Classification Biological Sciences, Genetics.

Trajectory and Uniqueness of Mutational Signatures in Yeast Mutators

Sophie Loeillet[a,b], Mareike Herzog[c], Fabio Puddu[c], Patricia Legoix[d], Sylvain Baulande[d], Stephen P. Jackson[c] & Alain Nicolas[a,b,*]

[a] Institut Curie, PSL Research University, CNRS, UMR3244, 26 rue d'Ulm, 75248 Paris Cedex 05, France.

[b] Sorbonne Universités, UPMC University Paris 06, CNRS, UMR3244, 26 rue d'Ulm, 75248 Paris Cedex 05, France.

[c] Wellcome/Cancer Research UK Gurdon Institute, Henry Wellcome Bldg. of Cancer and Developmental Biology, Tennis Court Road, Cambridge CB2 1QN, UK.

[d] ICGEX NGS Platform, Institut Curie, 26 rue d'Ulm, 75248 Paris Cedex 05, France.

[*] Correspondence: alain.nicolas@curie.fr

Abstract

The acquisition of mutations plays critical roles in adaptation, evolution, senescence and tumorigenesis. Massive genome sequencing has allowed extraction of specific features of many mutational landscapes but it remains difficult to retrospectively determine the mechanistic origin(s), selective forces and trajectories of transient or persistent mutations and genome rearrangements. Here, we conducted the prospective reciprocal approach to inactivate 13 single or multiple evolutionary-conserved genes involved in distinct genome maintenance processes and characterize *de novo* mutations in 274 diploid *S. cerevisiae* mutation accumulation lines. This approach revealed the diversity, complexity and ultimate uniqueness of mutational landscapes, differently composed of base substitutions, small InDels, structural variants and/or ploidy variations. Several landscapes parallel the repertoire of mutational signatures in human cancers while others are either novel or composites of sub-signatures resulting from distinct DNA damage lesions. Notably, the increase of base substitutions in the homologous recombination deficient Rad51 mutant, specifically dependent on the Polζ translesion polymerase, yields COSMIC Signature 3 observed in BRCA1/BRCA2-mutant breast cancer tumors. Furthermore, “mutome” analyses in highly polymorphic diploids and single-cell bottleneck lineages revealed a diverse spectrum of loss-of-heterozygosity (LOH) signatures characterized by interstitial and terminal chromosomal events resulting from inter-homolog mitotic crossovers. Following the appearance of heterozygous mutations, the strong stimulation of LOHs in the *rad27/FEN1* and *tsa1/PRDX1* backgrounds leads to fixation of homozygous mutations or their loss along the lineage. Overall, these “Mutomes” and their trajectories provide a mechanistic framework to understand the origin and dynamics of genome variations that accumulate during clonal evolution.

Keywords | Mutator genes | Mutational profiles | Polζ | Loss of Heterozygosity | Dynamic of mutations accumulation

Significance Statement

Deficiencies in genome maintenance genes result in increased mutagenesis and genome rearrangements that impacts cell viability, species adaptation and evolvabilty. The accumulation of somatic mutations is also a landmark of most tumor cells but it remains difficult to retrospectively determine their mechanistic origin(s). Here, we conducted the prospective reciprocal approach to inactivate evolutionary-conserved genes involved in various genome maintenance processes and characterize *de novo* mutations in diploid *S. cerevisiae* mutation accumulation lines. Our results revealed the diversity, trajectory, complexity and ultimate uniqueness of the clonal mutational landscapes. Some mutational signatures ressemble those found in human tumors.

Introduction

Acquired and transitory mutations, broadly genome instability, can be evolutionary advantageous in contributing to the adaptation of species in changing environments, or detrimental in reducing short- and long-term fitness (1–3). Mechanistically, spontaneous mutations in normal cells, exposure to environmental genotoxic compounds and deficiencies in genome maintenance genes are prominent sources of subtle or drastic genome changes/rearrangements and eventually functional and phenotypic variations (4, 5). A paradigm for this phenomenon is the accumulation of a variable burden of passenger and driver somatic mutations in tumor cell lineages (6–11). Thus, genome sequencing and mutational landscape analyses of germline and somatic mutations have permitted the retrospective identification of the most likely environmental sources of mutagen exposures,

such as UV exposure in melanoma and smoking in lung cancers or genetic features such as deficiency in DNA mismatch repair in colon cancers and homologous recombination defects in breast and ovarian cancers (6–14). However, it remains puzzling that in numerous instances, an environmental factor and/or defective mutator gene(s) is not found, although numerous relevant and evolutionary conserved genome maintenance genes and pathways are known (5, 15, 16). Here, we conducted the reciprocal functional approach to inactivate one or several genes involved in distinct genome maintenance processes (replication, repair, recombination, oxidative stress response or cell cycle progression) in *S. cerevisiae* diploids, establish the genome-wide mutational landscapes of mutation accumulation (MA) lines, explore the underlying mechanisms and characterize the dynamics of mutation accumulation (and disappearance) along single-cell bottleneck passages.

Results and Discussion

Variety of mutational landscapes

Overall, we established the mutational landscapes of 274 MA lines generated in the isogenic BY and/or hybrid SK1/BY wild-type (WT) backgrounds. Strains assessed included wild-type (WT), 11 single deletion mutants (hereafter abbreviated by gene name) and 3 double-mutants covering various genome maintenance processes. Compared with WT, we analysed the following mutant strains: *pif1Δ*, *pol32Δ* and *rad27Δ* (replication), *msh2Δ* (mismatch repair), *mre11Δ, rad51Δ, tho2Δ* (recombination and repair), *lig4Δ* (Non-Homologous End Joining, *tsa1Δ* (oxidative stress response), *cac1Δ cac3Δ* (nucleosome deposition), *clb5Δ* and *sic1Δ* (cell cycle progression) (http://www.yeastgenome.org). The strain genotypes are indicated in Dataset S1. All the genes assessed are evolutionary conserved and most are implicated in human diseases and/or tumor development (Fig. 1*A*) (http://www.yeastgenome.org, http://www.genecards.org/). To ensure the recovery of independent events, 4-16 individual

colonies per strain were derived in parallel MA lines (Fig. 1*B*), and sequenced after a minimum of 180 single-cell bottleneck passages (*SI Appendix*, Materials and Methods, Dataset S2 for individual clones). One passage corresponds to ~25 generations. Our bioinformatics analyses of the next-generation sequencing (NGS) reads allowed identification of base substitutions (SNPs), multi-nucleotide polymorphisms (MNPs), small (1-44 bp) insertions/deletions (InDels), combinations of SNP and small InDels (Complex), structural variants (SVs), as well as chromosomal ploidy variations and LOH regions (*SI Appendix*, Materials and Methods, *SI Appendix,* Fig. S1). The coordinates and annotations of the 8,876 *de novo* mutations identified in this study are reported in Datasets 3 to 8. The number of mutations detected in the parallel MA lines of the same genotype was similar (*SI Appendix*, Fig. S2*A* and Dataset S9), thus excluding clonal effects. Except for few common homopolymer InDels in *msh2* and *rad27* backgrounds, all mutations were different from one another.

Functionally, 4,103/5,416 (75.8%) base substitutions were located in a gene-coding region (Datasets 3 and 4), similar to random expectation (76.5%). At the protein level, 2,879 (53.2%) modified the amino acid with a presumptive moderate functional impact according to SnpEff annotation (17), and 199 (3.7%) created a premature stop codon. Among these protein truncating mutations, 58 were located in an essential gene (https://www.yeastgenome.org/) and all were heterozygous, likely phenotypically recessive.

The mutation frequencies *per* strain genotype (normalized per clone and passage) are reported in Fig. 1*C*. Both BY and SK1/BY wild-type strains accumulated few SNPs corresponding to a frequency of 0.11 mutations/clone/passage or 1.8×10^{-10} mutations/nucleotide/generation, similar to previous measurements (18). Not surprisingly, the vast majority were heterozygous (allelic ratio of ~0.5) but a few appeared as homozygous (see below). The mutation frequencies and genome rearrangements in the mutant MA lines varied

119 up to 83-fold compared to WT (Fig. 1*C*; Dataset S10), and delineated 6 classes of mutational
120 profiles. The first class, comprising *cac1 cac3* (chromatin assembly factors), *lig4* (non-
121 homologous end-joining) and *pol32* (Polδ replication) accumulated few base substitutions,
122 similar to WT (Fig. 1*C*). The second class represented by *rad51* (homologous recombination)
123 and *tsa1* (oxidative stress) specifically increased base substitutions (20.5- and 13.6-fold,
124 respectively) but seemingly via different mechanisms (see below). Distinctively, *rad51* more
125 than *tsa1* (2.8×10^{-2} and 0.2×10^{-2} SV/clone/passage, respectively) enhanced structural
126 variations (SVs). All were heterozygous intra-chromosomal deletions (Dataset S8). Their
127 length varied between 488-59,200 nt and most particularly (11/17 cases) occurred between
128 transposable (Ty/LTR) elements and the others between repeated homeologous genes
129 (Dataset S8). This is typical of single-strand annealing (SSA) events, known to be Rad51-
130 independent (19). The third class is defined by *msh2* (mismatch repair) that exhibited a strong
131 increase of base substitutions (26.4-fold) and small InDels (495-fold) with a slight excess
132 (58.7%) of small InDels over base substitutions, as previously observed (20, 21). Notably, as
133 reported for haploid strains (22, 23), there was an excess (81%) of deletions *vs.* additions
134 within homopolymer tracks. Among all small variants, the complex base substitutions were
135 rare (9/2824) (Fig. 1*C*; Dataset S10). The fourth class of mutant represented by *clb5* and *sic1*
136 (cell cycle progression), *mre11* (double-strand break repair), and *tho2* (transcription coupled
137 recombination) exhibited a slight increase (1.5- to 3.6-fold) of base substitutions but also
138 aneuploidies. The fifth class, defined by *rad27* (lagging strand replication and base excision
139 repair) yielded the broadest spectrum of mutational events. It exhibited an increase of base
140 substitutions (8.1-fold increase) including few complex substitutions events (76/1208), small
141 InDels (63-fold increase) mostly located in homopolymers and microsatellites (518/564) with
142 an excess of insertion vs. deletion (73%) but also SVs represented by 26 large deletions (62-
143 23,614 nt) and 2 small duplications (520 and 542 bp) (Dataset S8) as well as aneuploidies

144 (7.8 fold increase) (Dataset S10). Similar to *rad51*, the SVs in *rad27* reached a spontaneous
145 frequency of 4.3 10^{-2}/clone/passage. The deletions involved homeologous repeated regions
146 located in *cis* but fewer (3/26 in *rad27* instead of 11/17 in *rad51*) involved Ty/LTR elements
147 (Dataset S8). The sixth class of mutational profile is represented by *pif1*, affecting various
148 DNA metabolism functions (SGD, http://www.yeastgenome.org), whose major feature is the
149 rapid and complete loss of mitochondrial DNA (*SI Appendix*, Fig. S2*B*). Further, *pif1*Δ MA
150 lines exhibit a slight increase of base substitutions (2.5-fold) (Figure 1 and Dataset S10),
151 consistent with the 2 to 3-fold increase of spontaneous mutagenesis previously observed in
152 WT cells lacking mitochondrial DNA (*rho0*) (24). Compared to our previous analyses of
153 haploid mutants (23), the mutational spectrum and the overall frequencies of SNPs and small
154 Indels per genome in the haploid and diploid cells are similar (*SI Appendix*, Fig. S3)
155 indicating no drastic effect of the ploidy variation.
156 To more broadly characterize all the mutational landscapes, we also examined variation of
157 mitochondrial DNA (mtDNA) and ribosomal DNA (rDNA) copy number. It was substantially
158 variable in the WT and mutant parental strains with 0 to 103 mtDNA copies and 42 to 122
159 rDNA copies. In the MA lines, slight changes of mitochondrial DNA and rDNA copy number
160 (~20 copies) occurred from clone to clone (*SI Appendix*, Fig. S2*B* and *C*), compared with
161 parent. *lig4* clones increased and *mre11* clones decreased median copy number of mtDNA (34
162 and 40 copies, respectively). The study of additional MA lines issued from independent
163 parental strains, preferentialy with variable amount of starting mtDNA, would be required to
164 conclude if this is a mutant-specific effect, as observed in other yeast mutants (25) and
165 determine its impact on mutational profiles. In summary, this set of mutator profiles illustrates
166 a variety of mutator behaviours, leading to a considerable variety of mutation loads and
167 mutational landscapes.
168

Mutational signatures

The landscape of somatic mutations in tumor genomes has been correlated with distinct mutational processes, via mathematical and statistical methods able to distinguish different mutation signatures (6, 12, 14, 26–28). It has allowed identification of >30 cancer-derived patterns called COSMIC signatures, (http://cancer.ac.uk/cosmic/signatures) based on the relative incidences of base substitution changes within a trinucleotide context (12, 26). Similarly, we established the base substitution profile of our yeast mutants that yielded ≥500 SNP mutations (Fig. 2*A*) and the relative contribution of the "COSMIC" signatures (Fig. 2*B*). *tsa1*, one of the strongest single-gene mutant mutator in yeast (29), doesn't exhibit a predominant signature but a near equal contribution of Signatures 1, 3, 9, 18 and 30 (Fig. 2*B*). Thus, loss of Tsa1 – the major thioredoxin peroxidase that scavenges hydrogen peroxide in *S. cerevisiae* (30) – which yielded C>A and C>T mutations was not associated with a specific COSMIC signature. Mutations in the human ortholog gene *PRDX1* has not been associated with disease or tumors, perhaps due to extensive functional redundancy of thioredoxin peroxidases in mammals (31). Robustly, the *msh2* signature (C>T, C>A and T>C) supplemented with homopolymers/microsatellite instability was most similar to Signatures 14 and 20 (Fig. 2*B*), consistent with MMR-deficient cancer-derived signatures associated with elevated rates of colorectal and uterine cancers. Our analysis of the *msh2Δ* base substitutions identified by Lujan et *al*. (21) yielded a similar mutational signature (Fig. 2B). By contrast, *rad27* exhibits Signature 8 associated with breast cancer and medulloblastoma (https://cancer.sanger.ac.uk/signatures/). Since, *rad27* yields all kinds of mutational events (Fig. 1*C-D*) including a large spectrum of base substitutions (Fig. 2*A*), Signature 8 might be the sum of several lesion-specific sub-signatures. On the other hand, the *rad51* profile predominantly involving C>A, C>G and C>T changes exhibited Signature 3 (Fig. 2*B*)

consistent with its prominent role in homologous recombination (32). Also differently, our analyses of the base substitutions in the mutator DNA polymerase mutants *pol1-L868M*, *pol2-M644G* and *pol3-L612M* (21) yielded the predominant signatures 8, 22 and 12, respectively (*SI Appendix*, Fig. S4). Altogether, these results outline the uniqueness of the base substitutions signatures to specific genes, and retrospectively inform on the molecular defects underlying the accumulation of mutations in specific tumors (6–15).

Base substitution in the absence of Rad51 specifically requires Polζ

Several decades ago, the elevated mutagenesis of a *rad51* mutant was found to decrease when cells were also mutated in *REV3* (33), a gene now known to encode a component of the error-prone translesion synthesis (TLS) Rev1/Rev3/Rev7 polymerase ζ (zeta) complex (34). To further explore *rad51* mutagenesis, we associated the *rad51* deletion with each TLS polymerase deletion mutant, and measured mutation frequencies with the sensitive *CAN1*^{R} mutational assay (35). This revealed that *rad51* enhanced mutagenesis was reduced essentially to WT levels in combination with *rev1*, *rev3 or rev7* but remained unchanged with *pol4* (Polλ) or *rad30* (Polη) (Fig. 2*C*). Consistently, we did not find significant additive or synergic effects of combining *rad51* with the *rev1 rev3 pol4 rad30* quadruple mutant.

Since, Rev3 carries the catalytic activity of Polζ while the Rev1 and Rev7 proteins might also serve as "recruitment platforms" involved in other related but distinct biological functions - the mammalian REV7 is involved in controlling DNA end resection and DNA damage responses via the Shieldin complex (36–38) - we also combined *rad51* with the catalytically dead *rev3-D1142A,D144A* polymerase mutant (39). *rad51*-induced mutagenesis was reduced to the WT level (Fig. 2*C*), demonstrating a role for Rev3 TLS activity. Thus, Polζ appeared specifically involved in the default repair of DNA lesions in the absence of Rad51-dependent homologous recombination, most likely during replication. As Polζ is

evolutionary conserved (40), these results raise the possibility that Polζ is responsible for enhanced mutational loads observed in HR-deficient *BRCA1/2* mammalian cells, as well as in patients with *RAD51* mutations and Fanconi anemia-like phenotypes (41).

For comparison, we also combined *rev3* with the other base-substitution mutators. We found no reduction of $CAN1^R$ cells in the *tsa1* background, indicating that Rad52 foci accumulating in this mutant (42) result from a different lesion(s) than in the *rad51* setting. In contrast, the inactivation of *REV3* yielded a partial decrease (58%) of $CAN1^R$ cells when combined with *rad27* (Fig. 2*C*), suggesting that its deficiency in Okazaki fragment processing during lagging strand replication generates DSBs and/or single-strand gaps similar to *rad51*. The remaining Rev3-independent base substitution mutations may result from default base excision repair of apurinic/apyrimidinic (AP) sites (43), thus partially contributing to the composite Signature 8. Finally, similar to *tsa1*, the lack of Pol ζ had no discernible effect on *msh2* mutagenesis (Fig. 2*C*). In conclusion, Polζ genetic dependency appears specifically connected to formation and/or resolution of lesions arising in a HR-deficient context.

Occurrence of homozygous *de novo* mutations

Beyond heterozygous mutations, we found some base substitutions and InDels with an allelic ratio of 1.0, implying loss of the wild-type allele. This mostly occurred in the *msh2, tsa1* and *rad27* diploids, representing 2.3%, 6.6%, and 13.4% of the total frequency of base substitutions and small indels mutations, respectively (Fig. 3*A*, Dataset S10). In such situations, various types of genomic events, distinguishable by the state of the homologous chromosomes, could be invoked (Fig. 3*B*). In *msh2*, 70/71 cases occurred in full diploid cells and resulted from two identical (18 cases) or two distinct (48 cases) InDels, located within the same homopolymer tract on the homologs (Fig. 3*A*, Datasets 3 and 6). It can be explained from the >1 nucleotide length of these motifs and high rate of polymerase slippage within

244 homopolymers during replication (44). In *tsa1*, the homozygous SNPs were also mostly found
245 on chromosomes with 2 copies (54/60 cases) but all were located in non-repeated nucleotide
246 sequences (Fig. 3*A*, Datasets 3 and 4). This was rather similar in *rad27*, except that 22/163
247 cases were associated with a change of the local copy number (1 or >2). We hypothesize that
248 along the lineages, the heterozygous *de novo* mutations were rendered homozygous upon a
249 subsequent LOH event.
250

251 ## Detection of LOH signatures in hybrid yeasts

252 LOH can result from mitotic inter-homolog recombination, short tract mitotic gene
253 conversions and/or break-induced replication (BIR) events that are difficult to detect in
254 isogenic strains. To comprehensively detect LOHs, we generated additional WT and mutant
255 MA lines from the polymorphic SK1/BY diploid that carries >53,000 constitutive SNPs
256 markers, distributed on each chromosome with one marker every 218 bp in average (Dataset
257 S11). Compared to the isogenic and hybrid WT, the mutant MA lines exhibited similar
258 mutation frequencies and specific mutational landscapes (compare Fig. 1*C* and 1*D*, *SI*
259 *Appendix*, Fig. S2*A*) but revealed the presence of numerous LOH regions, robustly defined to
260 involve ≥3 adjacent markers (Fig. 4*A*, *SI Appendix*, Fig. S5 to 9). In WT and *pif1*, LOHs were
261 rare (0.09 and 0.12 LOH/clone/passage, respectively) (*SI Appendix*, Fig. S7 and 8). It was
262 modestly increased in *rad51* (0.18 LOH/clone/passage), while many arose in *tsa1* and *rad27*
263 (1.2 and 2.5 LOH/clone/passage corresponding to a 12.7- and 27.3-fold increase,
264 respectively). In numerous instances, these LOH events involved several chromosomes in the
265 same clone (Fig. 4*B*; *SI Appendix*, Fig. S5*A* and 6*A*). Considering all the clones, the LOHs
266 covered a large fraction of the genome in *tsa1* and almost all the genome in *rad27* (Fig. 4*C*).
267 Regarding the occurrence of homozygous mutations, again this was most frequent in *tsa1* and
268 *rad27* cells (Fig. 4*D*; Dataset S10). Notably, 16/22 in *tsa1* and 95/114 in *rad27* were located

in LOH regions with 2 copies of the chromosome (Fig. 4*C*, *E*; Datasets 3 and 4), consistent with the hypothesis that along the cell lineage, mutations arose as heterozygous and passively became homozygous as part of a subsequent overlapping LOH event (Fig. 4*E*). Among the remaining events, 2 cases in *tsa1* and 10 cases in *rad27,* resulted from the occurrence of a *de novo* mutation on one homolog and an overlapping *de novo* deletion on the homolog (Fig. 3*B*, Fig. 4*D*; Dataset S4), as frequently found in tumor cells that carried a germline susceptibility mutation and then acquired a secondary somatic deletion on the homologous chromosome (45). Thus, the highly mutagenic *tsa1* and *rad27* strains stimulated SNPs and LOH events, a dual signature that accelerate and enlarge the spectrum of genome modifications.

Distributions and mechanisms of interstitial and terminal LOHs

In *tsa1* and *rad27*, the majority of LOHs were interstitial (81% and 76%, respectively; Fig. 4*F*) with a length varying from 33 bp to 419 kb and 17 bp to 846 kb, respectively (Dataset S12). The remaining LOHs were terminal, with lengths varying from 659 bp to 1,052 kb in *tsa1* and 55 bp to 1,079 kb in *rad27* (Dataset S12). Globally, the interstitial LOHs are shorter than the terminal LOH (*SI Appendix*, Fig. S10*A*), consistent with their origin resulting from gene conversion like events and/or double crossovers rather than a single crossover. The LOHs size ranges were similar to those observed in a previous study (46). In both mutants, the LOHs were from one or the other parental haplotype, with a slight BY vs. SK1 excess genotype (58 and 55%, respectively). Due to the extended polymorphism of the BY and SK1 genomes, this slight bias may result from intrinsic and emerging lethal allele incompatibilities when part of the genome become homozygous, a somatic manifestation of the spore inviability observed in the SK1/S288C haploid segregants (47, 48). The annotation of the LOH breakpoint regions did not localize to specific functional elements except in *pif1* where they often were in proximity to a LTR/Ty region and/or the rDNA locus (*SI Appendix*, Fig.

S10*B*). Thus, after only 25 single bottleneck passages, the stimulation of LOH created mosaic diploid genomes (Fig. 4*C* and *SI Appendix*, Fig. S5*A* and 6*A*) that reached 4.7-28.9% homozygosity per clone in *tsa1*, and 26.6-60.7% in *rad27*.

The formation of terminal LOHs is a hallmark of break-induced-replication (BIR) (49, 50), whereas both terminal and interstitial LOHs can result from mitotic crossover recombination and/or gene conversion. Since BIR specifically depends on the activity of *POL32* and *PIF1* (51–54), we examined the effect of deleting these genes in the *tsa1* mutant (Dataset S1). Similar to *tsa1*, the *tsa1 pol32* and *tsa1 pif1* SK1/BY MA lines displayed increased base substitutions (13.9 and 32.3-fold *vs.* WT, respectively) and LOHs (11.1- and 17.5- fold *vs* WT, respectively). The absolute frequency of terminal LOHs, however, was not significantly reduced (0.22, 0.25 and 0.29 /clone/passage in *tsa1*, *tsa1 pol32* and *tsa1 pif1*, respectively) and the large excess of interstitial *vs.* terminal LOHs was retained (81%, 72%, 81% in *tsa1*, *tsa1 pol32* and *tsa1 pif1*, respectively) (Fig. 4*F*). Thus, such LOHs result from stimulation of mitotic recombination, rather than BIR, explaining the synthetic lethality of the *tsa1 rad51* double mutant (42). We examined the length of the terminal LOH in the *tsa1*, *tsa1 pif1* and *tsa1 pol32* (*SI Appendix*, Fig. S10*A)* and observed no significant difference between *tsa1* and *tsa1 pif1* but a significant increase of terminal LOH length in *tsa1 pol32* suggesting a role of Pol32 in the distribution of the initiating events although the annotation of the terminal LOH breakpoints in the three *tsa1* strain is similar (*SI Appendix*, Fig. S10*C*). The contribution of BIR in the stimulation of the *rad27* LOHs could not be examined due to the synthetic lethality of the *rad27 pol32* double mutant (55). Nevertheless, the synthetic lethality of *rad27* (alike *tsa1*) with *rad51* (55), suggests that *rad27* LOHs also largely result from inter-homolog mitotic recombination, albeit not necessarily stimulated by identical initiating lesion(s).

Trajectory of base substitution and LOH along lineages

319 To determine trajectories of mutation accumulation, we sequenced the genomes of the *tsa1*
320 clone N and *rad27* clone C cells collected at each of the 25 bottleneck passages (*SI Appendix*,
321 Fig. S5*B* and 6*B*, and Movies S1 and S2, respectively). In both mutants, the accumulation of
322 heterozygous mutations (SNPs and small Indels) appeared essentially regular; in *tsa1*, 15/25
323 passages yielded one or two *de novo* mutations and 7/25 passages three or five mutations; in
324 *rad27*, 7/25 passages yielded one or two mutations but a majority of passages (15/25) yielded
325 three to nine mutations. These multiple mutation events did not necessarily arise within one
326 cell division, since in our experimental protocol, each bottleneck passage correspond to ~25
327 generations (Fig. 5, Datasets 13-14). Also, to note, 3/16 *tsa1 pol32* clones (C12, D12, O12)
328 that exhibited a LOH in both *MSH2* and *PMS1* regions showed a higher number of *de novo*
329 mutations (34, 30 and 46 mutations, respectively) compared to 19 mutations in average in the
330 other clones. This can be explained by the presence of *MLH1-D161* homozygous allele from
331 BY and *PMS1-K818* homozygous allele from SK1, previously reported to confer a mismatch
332 repair deficient phenotype in haploid strains (56). This case illustrates the occurrence of a
333 secondary mutator phenotype occurring during the clonal drift.
334 Along the lineages, several *de novo* heterozygous mutations (2/45 and 9/90 in *tsa1* and
335 *rad27*, respectively) chronologically became homozygous in a single bottleneck passage as a
336 consequence of an overlapping LOH, while others (1/45 and 6/90 in *tsa1* and *rad27*,
337 respectively) were eliminated in favour of the WT allele (Fig. 6*A-B*; Dataset S15-16, Movies
338 S1-2, http://xfer.curie.fr/get/i2FK7JDkQgs/SupAnim.zip). This opposite outcome is explained
339 by the occurrence of an overlapping LOH mediated by an inter-homolog recombination event,
340 followed by the segregation of the non-sister chromatids carrying both WT or mutant allele in
341 the daughter cells (Fig. 4*E*). Multiple fixation and elimination of mutations, as well as
342 extension of LOH tracts, also occurred in a single passage (Fig. 6*C*). The biological impact of
343 such a mutator phenotype is functionally important because during cell proliferation,

stimulation of LOHs will allow the phenotypic expression of recessive *de novo* mutations when fixed but also erase heterozygous mutations that transiently occurred during clonal evolution. In cancer settings, such a mutator phenotype could be initially advantageous to enhance the genetic diversity to stimulate proliferation of pre-tumoral cells, while afterwards the restoration of the WT allele could be beneficial to restore cell physiology. A similar scenario for a dominant mutator gene mutation will permit a wave of cell genetic diversification and its subsequent elimination, avoiding the accumulation of additional disadvantageous mutations (2, 57, 58). Retrospectively, in contrast to reversible epigenetic events that may not leave long term molecular scars, a transient mutation can remain detectable as a LOH event. This "archaeological signature" raises the prospect that one or more of LOH embedded genes may have been transiently mutated during the evolutionary history of a cell lineage.

Conclusion

The mutation of genes controlling genome stability and/or the epigenetic deregulation of their expression contributes to create the genetic diversity on which the Darwinian selection can act. Our study has illustrated the large variety of mutational profiles generated by genetic deficiencies in genome-stability genes, and described the dynamics of *de novo* mutations and genome rearrangements (fixation and disappearance) during vegetative growth. This knowledge suggests ways to mechanistically interpret tumor cells genome evolution and genetic sensitivity (6–11, 59, 60), as well as genome evolution in species (1–3). On the evolutionary scale, impaired function of genes such as *RAD27/FEN1* and *TSA1/PRDX1* may allow the generation of genetic diversity, including occasional beneficial mutations (or suppressors of less fit mutant states), while additional recombination-dependent changes may be beneficial to resolve burdens of allelic incompatibilities in polymorphic and hybrid

species. In the future, extending analyses of "mutomes" in yeast should allow refinement of the mutator scope of additional genome maintenance genes and graph the complexity of the genes/pathways and their interactions (4). It will also likely suggest how related phenomena operate in other organisms such as *Caenorhabditis elegans* (61) and engineered human cell lines (62) amenable to "mutome" analyses.

Materials and Methods

Strains and mutation accumulation lines. Mutations accumulations lines were obtained from BY or BY/SK1 diploid mutants carrying homozygous deletions of the genes listed in Fig. 1*A*. The details of the strain constructions are described in *SI Appendix* and the complete strain genotypes are listed in Dataset S1. All strain constructions were checked by PCR, Southern blot or Sanger sequencing. Proper gene deletions were confirmed by the lack of read coverage upon whole genome sequencing of the parent and MA lines.

Generation of mutation accumulation lines. The mutation accumulation lines (MA) were obtained as described in (23). Briefly, 4 to 16 colonies of each diploid parental strain was subjected to 12 to 100 single-cell bottlenecks (Datasets S2 and 10). One single-cell bottleneck is performed by picking one colony of average size and by streaking it to individual colonies on YPD plate (1% yeast extract, 2% peptone, 2% dextrose), incubated for 3 days of growth at 30°.

Mutation calling, LOH detection and mutational signatures. Illumina whole-genome sequencing were performed on parents and mutations accumulation lines. The paired-end reads were aligned on R64-1-1 S288c SGD reference sequence (http://www.yeastgenome.org). Our analysis pipeline outlined in *S1 Appendix*, Fig. S1

allowed the detection of all kind of mutations and genome rearrangements and ploidy variations. The base substitution mutational signatures were extracted using the R/Bioconductor MutationalPatterns package (63). In the hybrid strains, the LOH regions were detected by genotyping the 53,523 polymorphisms that distinguish the BY and SK1 strain backgrounds (Dataset S11). The LOH were robustly defined as regions showing at least 3 consecutive homozygous markers of the same haplotype (See details in *SI Appendix*).

Canavanine mutator assay. To measure the rate of mutations in the *CAN1* gene, the occurrence of canavanine resistant colonies in BY haploids were measured as previously described (64). The fluctuation test assays were performed from 5 independent cultures. The mutation rate was calculated using bz-rates (65) tool (http://www.lcqb.upmc.fr/bzrates). Reported mutation rates are the average of at least 3 experiments.

Acknowledgements

We thank A. Londono-Vallejo and all laboratory members for stimulating discussions. We also thank the reviewers for helpful suggestions. Research in the Institut Curie was funded by the “Ligue contre le Cancer” to A.N. and to the NGS-ICGEX platform by ANR-10-EQPX-03 (Equipex) and ANR-10-INBS-09-08 (France Génomique Consortium) grants from the Agence Nationale de la Recherche ("Investissements d’Avenir" program). We thanks the Bioinformatics platform of the Institut Curie for the sequencing read mapping. Research in the S.P.J. lab is supported by Wellcome Strategic Award 101126/Z/13/Z (COMSIG); Wellcome Investigator Award 206388/Z/17/Z; Wellcome PhD Fellowship 098051 to MH; Cancer Research UK Programme Grant C6/A18796; Cancer Research UK C6946/A24843 and Wellcome WT203144 Institute Core Funding. We thank the Cancer Genome Project and DNA sequencing pipelines at the Wellcome Sanger Institute for help with sample submission,

tracking, library preparation and sequencing.

Author contributions

A.N. designed research. S.L., M.M. and F.P. performed research. P.L and S.B. performed the NGS done at the Institut Curie. All authors contributed to the analyses of the data. S.L and A.N. wrote the paper with input from F.P. and S.P.J.

Competing interests

The authors declare no conflict of interest.

Footnotes. nt: nucleotide; bp: base pair; YPD (Yeast Peptone Dextrose rich media); CAN: Canavanine.

References

1. P. D. Sniegowski, P. J. Gerrish, R. E. Lenski, Evolution of high mutation rates in experimental populations of E. coli. *Nature* **387**, 703–705 (1997).
2. D. A. Thompson, M. M. Desai, A. W. Murray, Ploidy Controls the Success of Mutators and Nature of Mutations during Budding Yeast Evolution. *Curr. Biol.* **16**, 1581–1590 (2006).
3. A. N. Nguyen Ba, *et al.*, High-resolution lineage tracking reveals travelling wave of adaptation in laboratory yeast. *Nature* **575**, 494–499 (2019).
4. K. Myung, C. Chen, R. D. Kolodner, Multiple pathways cooperate in the suppression of genome instability in Saccharomyces cerevisiae. *Nature* **411**, 1073–1076 (2001).
5. C. D. Putnam, R. D. Kolodner, Pathways and mechanisms that prevent genome instability in Saccharomyces cerevisiae. *Genetics* **206**, 1187–1225 (2017).

6. M. R. Stratton, P. J. Campbell, P. A. Futreal, The cancer genome. *Nature* **458**, 719–724 (2009).
7. I. Martincorena, P. J. Campbell, Somatic mutation in cancer and normal cells. *Science (80-.).* **349**, 1483–1489 (2015).
8. P. J. Campbell, *et al.*, Pan-cancer analysis of whole genomes. *Nature* **578**, 82–93 (2020).
9. E. D. Pleasance, *et al.*, A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* **463**, 191–196 (2010).
10. S. Nik-Zainal, *et al.*, Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).
11. M. S. Lawrence, *et al.*, Mutational heterogeneity in cancer and search for new cancer genes. *Nature* **499**, 214–218 (2013).
12. L. B. Alexandrov, *et al.*, The repertoire of mutational signatures in human cancer. *Nature* **578**, 94–101 (2020).
13. E. D. Pleasance, *et al.*, A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463**, 184–90 (2010).
14. L. B. Alexandrov, *et al.*, Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
15. E. M. Kass, M. E. Moynahan, M. Jasin, When Genome Maintenance Goes Badly Awry. *Mol. Cell* **62**, 777–787 (2016).
16. T. A. Knijnenburg, *et al.*, Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. **23**, 239–254 (2018).
17. D. M. R. P. Cingolani, A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain

w1118 ; iso-2; iso-3. *Fly (Austin).* **6**, 80–92 (2012).

18. N. P. Sharp, L. Sandell, C. G. James, S. P. Otto, The genome-wide rate and spectrum of spontaneous mutations differ between haploid and diploid yeast. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E5046–E5055 (2018).

19. E. L. Ivanov, N. Sugawara, J. Fishman-Lobell, J. E. Haber, Genetic requirements for the single-strand annealing pathway of double-strand break repair in Saccharomyces cerevisiae. *Genetics* **142**, 693–704 (1996).

20. G. I. Lang, L. Parsons, A. E. Gammie, Mutation rates, spectra, and genome-wide distribution of spontaneous mutations in mismatch repair deficient yeast. *G3 (Bethesda).* **3**, 1453–1465 (2013).

21. S. A. Lujan, *et al.*, Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition. *Genome Res.* **24**, 1751–1764 (2014).

22. K. T. Nishant, *et al.*, The Baker's yeast diploid genome is remarkably stable in vegetative growth and meiosis. *PLoS Genet.* **6** (2010).

23. A. Serero, C. Jubin, S. Loeillet, P. Legoix-Né, A. G. Nicolas, Mutational landscape of yeast mutator strains. *Proc. Natl. Acad. Sci.* **111**, 1897–1902 (2014).

24. A. K. Rasmussen, A. Chatterjee, L. J. Rasmussen, K. K. Singh, Mitochondria-mediated nuclear mutator phenotype in Saccharomyces cerevisiae. *Nucleic Acids Res.* **31**, 3909–3917 (2003).

25. F. Puddu, *et al.*, Genome architecture and stability in the Saccharomyces cerevisiae knockout collection. *Nature* (2019).

26. L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, P. J. Campbell, M. R. Stratton, Deciphering Signatures of Mutational Processes Operative in Human Cancer. *Cell Rep.* **3**, 246–259 (2013).

27. S. Nik-Zainal, S. Morganella, Mutational signatures in breast cancer: The problem at

494 the DNA level. *Clin. Cancer Res.* **23**, 2617–2629 (2017).
495 28. F. Maura, *et al.*, A practical guide for mutational signature analysis in hematological
496 malignancies. *Nat. Commun.* **10** (2019).
497 29. M.-E. Huang, A.-G. Rio, A. Nicolas, R. D. Kolodner, A genomewide screen in
498 Saccharomyces cerevisiae for genes that suppress the accumulation of mutations. *Proc.*
499 *Natl. Acad. Sci. U. S. A.* **100**, 11529–11534 (2003).
500 30. I. Iraqui, *et al.*, Peroxiredoxin Tsa1 is the key peroxidase suppressing genome
501 instability and protecting against cell death in Saccharomyces cerevisiae. *PLoS Genet.*
502 **5** (2009).
503 31. Z. A. Wood, E. Schröder, J. R. Harris, L. B. Poole, Structure, mechanism and
504 regulation of peroxiredoxins. *Trends Biochem. Sci.* **28**, 32–40 (2003).
505 32. K. S. . Jensen R.B, Carreira A., Purified human BRCA2 stimulates RAD51-mediated
506 recombination. *Nature* **467**, 678–683 (2010).
507 33. S.-K. Quah, R. C. Von Borstel, P. J. Hastings, DNA sequence analysis of spontaneous
508 mutagenesis in Saccharomyces cerevisiae. *Genetics* **96**, 1491–1505 (1980).
509 34. J. R. Nelson, C. W. Lawrence, D. C. Hinkle, Thymine-thymine dimer bypass by yeast
510 DNA polymerase ζ. *Science (80-.).* **272**, 1646–1649 (1996).
511 35. G. I. Lang, A. W. Murray, Estimating the per-base-pair mutation rate in the yeast
512 Saccharomyces cerevisiae. *Genetics* **178**, 67–82 (2008).
513 36. R. Gupta, *et al.*, DNA Repair Network Analysis Reveals Shieldin as a Key Regulator
514 of NHEJ and PARP Inhibitor Sensitivity. *Cell* **173**, 972–988 (2018).
515 37. D. G. T. & R. D. W. Junya Tomida, Kei-ichi Takata, Sarita Bhetawal, Maria D Person,
516 Hsueh-Ping Chao, FAM35A associates with REV7 and modulates DNA damage
517 responses of normal and BRCA1-defective cells. *EMBO J.* **37**, 1–14 (2018).
518 38. H. Dev, *et al.*, Shieldin complex promotes DNA end-joining and counters homologous

519 recombination in BRCA1-null cells. *Nat. Cell Biol.* **20**, 954–965 (2018).
520 39. H. M. Siebler, A. G. Lada, A. G. Baranovskiy, T. H. Tahirov, Y. I. Pavlov, A novel
521 variant of DNA polymerase zeta, Rev3ΔC, highlights differential regulation of Pol32
522 as a subunit of polymerase delta versus zeta in Saccharomyces cerevisiae. *DNA Repair*
523 *(Amst).* **24**, 138–149 (2014).
524 40. G. N. Gan, J. P. Wittschieben, B. Wittschieben, R. D. Wood, DNA polymerase zeta
525 (pol) in higher eukaryotes. *Cell Res.* **18**, 174–183 (2008).
526 41. A. T. Wang, *et al.*, A dominant mutation in human RAD51 reveals its function in DNA
527 interstrand crosslink repair independent of homologous recombination. **59**, 478–490
528 (2015).
529 42. S. Ragu, *et al.*, Oxygen metabolism and reactive oxygen species cause chromosomal
530 rearrangements and cell death. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 9747–9752 (2007).
531 43. X. Wu, Z. Wang, Relationships between yeast Rad27 and Apn1 in response to
532 apurinic/apyrimidinic (AP) sites in DNA. *Nucleic Acids Res.* **27**, 956–962 (1999).
533 44. M. Strand, T. A. Prollat, R. M. Liskayt, T. D. Petes, Destabilization of tracts of simple
534 repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* **365**,
535 274–276 (1993).
536 45. Y. Li, *et al.*, Patterns of somatic structural variation in human cancer. *Nature* **578**, 112–
537 121 (2020).
538 46. E. Yim, K. E. O'Connell, J. St. Charles, T. D. Petes, High-resolution mapping of two
539 types of spontaneous mitotic gene conversion events in saccharomyces cerevisiae.
540 *Genetics* **198**, 181–192 (2014).
541 47. R. Laureau, *et al.*, Extensive Recombination of a Yeast Diploid Hybrid through
542 Meiotic Reversion. *PLoS Genet.* **12** (2016).
543 48. E. Martini, R. L. Diaz, N. Hunter, S. Keeney, Crossover homeostasis in yeast meiosis.

544 *Cell* **126**, 285–95 (2006).
545 49. B. Llorente, C. E. Smith, L. S. Symington, Break-induced replication: What is it and
546 what is it for? *Cell Cycle* **7**, 859–864 (2008).
547 50. A. Malkova, G. Ira, Break-induced replication: Functions and molecular mechanism.
548 *Curr. Opin. Genet. Dev.* **23**, 271–279 (2013).
549 51. J. R. Lydeard, S. Jain, M. Yamaguchi, J. E. Haber, Break-induced replication and
550 telomerase-independent telomere maintenance require Pol32. *Nature* **448**, 820–823
551 (2007).
552 52. N. Saini, *et al.*, Migrating bubble during break-induced replication drives conservative
553 DNA synthesis. *Nature* **502**, 389–392 (2013).
554 53. M. A. Wilson, *et al.*, Pif1 helicase and Polδ promote recombination-coupled DNA
555 synthesis via bubble migration. *Nature* **502**, 393–396 (2013).
556 54. R. A. Donnianni, *et al.*, DNA Polymerase Delta Synthesizes Both Strands during
557 Break- Induced Replication. *Mol Cell.* **76**, 371–381 (2019).
558 55. S. Loeillet, *et al.*, Genetic network interactions among replication, repair and nuclear
559 pore deficiencies in yeast. *DNA Repair (Amst).* **4**, 459–468 (2005).
560 56. A. Demogines, A. Wong, C. Aquadro, E. Alani, Incompatibilities involving yeast
561 mismatch repair genes: A role for genetic modifiers and implications for disease
562 penetrance and variation in genomic mutation rates. *PLoS Genet.* **4**, 1–11 (2008).
563 57. A. Couce, *et al.*, Mutator genomes decay, despite sustained fitness gains, in a long-term
564 experiment with bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E9026–E9035 (2017).
565 58. A. Giraud, *et al.*, Costs and benefits of high mutation rates: Adaptive evolution of
566 bacteria in the mouse gut. *Science (80-.).* **291**, 2606–2608 (2001).
567 59. E. Guo, *et al.*, FEN1 endonuclease as a therapeutic target for human cancers with
568 defects in homologous recombination. *Proc. Natl. Acad. Sci.*, 202009237 (2020).

60. H. Farmer, *et al.*, Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005).

61. B. Meier, *et al.*, Mutational signatures of DNA mismatch repair deficiency in C. elegans and human cancers. *Genome Res.* **28**, 666–675 (2018).

62. X. Zou, *et al.*, Validating the concept of mutational signatures with isogenic cell models. *Nat. Commun.* **9**, 1–16 (2018).

63. F. Blokzijl, R. Janssen, R. van Boxtel, E. Cuppen, MutationalPatterns: Comprehensive genome-wide analysis of mutational processes. *Genome Med.* **10**, 1–11 (2018).

64. R. A. G. Reenan, R. D. Kolodner, Characterization of Insertion Mutations in the Saccharomyces cerevisiae MSHl and MSH2 Genes: Evidence for Separate Mitochondrial and Nuclear Functions. *Genetics* **132**, 975–985 (1992).

65. A. Gillet-Markowska, G. Louvel, G. Fischer, bz-rates : A Web Tool to Estimate Mutation Rates from Fluctuation Analysis . *G3: Genes|Genomes|Genetics* **5**, 2323–2327 (2015).

66. V. Boeva, *et al.*, Control-FREEC: A tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* **28**, 423–425 (2012).

Figures legends

Fig. 1. Mutational landscapes. (*A*) List of genes studied and their functions. (*B*) Experimental strategy to generate mutation accumulation lines. The *WT* diploid strains (BY/BY or SK1/BY background) were deleted for both copies (Δ/Δ) of the potentially mutator gene(s). Then, 4-16 independent clones of the WT and Δ/Δ diploids were grown mitotically and derived for up to 100 single-cell bottlenecks passages on YPD rich medium at 30°C (23). The genome of the resulting accumulation lines was individually sequenced by NGS and the reads analysed for detection of *de novo* mutations and genome rearrangements

(see *SI Appendix,* Materials and Methods and Fig. S1). (*C-D*) Mutational profiles in BY/BY and SK1/BY strains, respectively. N. of mutations: total number of *de novo* mutations detected in each strain, including single nucleotide variants (SNP), small indels, multi-nucleotide variant (MNP), "complex" events referring to combination of SNPs and small Indels, chromosome aneuploidies and structural variants (SV) (large deletions/insertions). The SNP and small indels comprise both heterozygous (allelic ratio ~0.5) and apparently homozygous events (allelic ratio ~1.0). For each mutant, the class of mutator profile, the number of clones, passages and mutations are indicated. The mean number of mutations *per* clone normalized to the number of passages and standard error are shown. The mutational fold variation compared to the corresponding WT is shown into parenthesis. Mann-Whitney-Wilcoxon test was performed to compare each mutant with WT (ns: not significant, **p-value<0.01).

Fig. 2. Mutational signatures and Polζ-dependent mutagenesis. (*A*) Mutational profiles of WT, *tsa1, rad51, msh2* and *rad27* mutants obtained with MutationalPatterns (63). N: number of base substitutions examined (sum of heterozygous and homozygous SNP found in BY and SK1/BY backgrounds). Count for *tsa1* is the sum of SNPs observed in the *tsa1*Δ/Δ, *tsa1*Δ/Δ *pol32*Δ/Δ and *tsa1*Δ/Δ *pif1*Δ/Δ strains (BY and SK1/BY backgrounds). *: WT data includes our data (110 SNPs) and 719 *de novo* SNP detected by Sharp *et al.* (18) in another WT diploid *S. cerevisiae* MA lines. (*B*) Relative contribution of COSMIC signatures in WT, *tsa1, rad51, msh2* and *rad27* mutational profiles (dataset as in *A*), calculated with MutationalPatterns (63). **: data from Lujan *et al*. (21). (*C*) Canavanine resistance (CAN^R) assay of WT and haploid mutants (BY background). The mutation rate is the average of at least 3 fluctuation tests, each made with 5 independent cultures. It is calculated according to

618 Reenan and Kolodner (64), using bz-rates web-tool (http://www.lcqb.upmc.fr/bzrates (65).
619 Error bars are standard deviation.
620
621 **Fig. 3. Occurrence and potential origin of «homozygous» *de novo* mutations.** (*A*) Allelic
622 ratio of mutations from MA lines (BY and SK1/BY backgrounds). Sum of heterozygous
623 (allelic ratio ~0.5) or homozygous (allelic ratio ~1.0) SNPs, small indels, MNP and complex
624 mutations. Copy number is calculated with Control-FREEC (66). (*B*) Molecular events
625 leading to *de novo* mutations with an allelic ratio of ~1.0, associated or not with local or
626 chromosomal copy number variation.
627
628 **Fig. 4. Detection of LOHs in the SK1/BY MA lines.** (*A*) Total number of LOH regions per
629 clone normalized *per* passage. (*B*) Examples of allelic profiles in a WT, *tsa1* and *rad27* MA
630 lines at final passage 25. The genotype of the 53,523 SK1 *vs.* BY polymorphisms are plotted
631 on the 16 chromosomes. A minimum of ≥3 adjacent markers of the same parental genotype
632 was retained to define the local haplotype (see *SI Appendix*) being either heterozygous
633 SK1/BY (grey), homozygous SK1 (blue) or homozygous BY (red). The triangles indicate the
634 location of the heterozygous (black) or "homozygous" (purple) *de novo* mutations (SNP,
635 MNP, complex and small indels, see *SI Appendix,* Materials and Methods). (*C*) Heatmap of
636 the genome wide occurrence of homozygosity among SK1/BY *tsa1* and *rad27* clones. (*D*)
637 Number of homozygous mutations originated from mitotic recombination or BIR (labelled
638 REC), chromosome loss or deletion identified in the SK1/BY mutants, being located or not in
639 a LOH region. (*E*) Two steps occurrence of homozygous *de novo* mutations upon inter-
640 homolog mitotic recombination or BIR. (*F*) Percentage of interstitial and terminal LOH–REC
641 tracks, N: Total number of LOH events.
642

Fig. 5. **Dynamics of LOH formation in the *tsa1* clone N and *rad27* clone C lineages.** Trajectory of the *de novo* LOH events in the SK1/BY *tsa1* clone N (left) and *rad27* clone C (right) from passage 1 to 25. First column is the parental clone. Grey: SK1/BY heterozygous markers, red: homozygous BY markers; blue: SK1 homozygous markers. In *rad27*, the passages 2 and 3 were mosaic and are shown in *SI Appendix*, Fig. S6; passage 8 is omitted because the cells could not be recovered after storage.

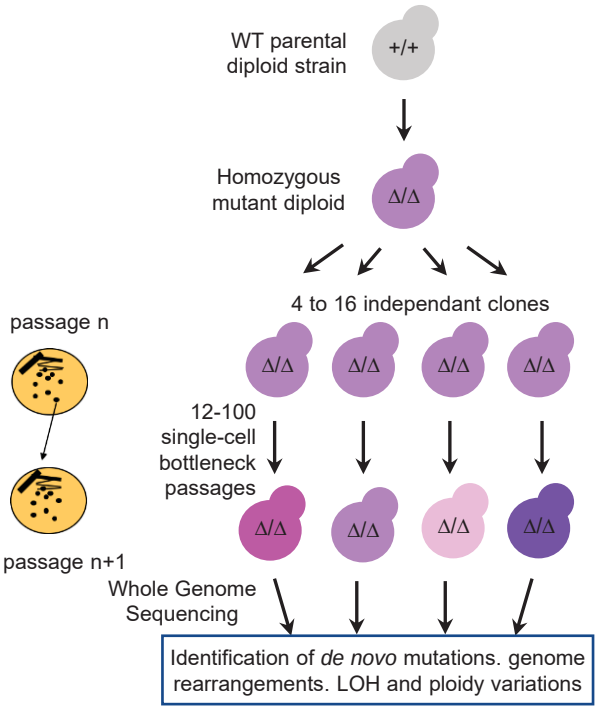
Fig. 6. **Dynamics of *de novo* mutations and LOH formation in the *tsa1* clone N and *rad27* clone C lineages.** (*A*-*B*) Trajectory and heterozygous (grey) *vs.* homozygous (purple) status of the 45 *de novo* mutations detected in the *tsa1* clone N and of the 90 *de novo* mutations detected in the *rad27* clone C, respectively, from passage 1 to 25. Green: mutation found in a 3 copy-number region and exhibiting a 1/3 allelic ratio. The coordinates of the mutations (chromosome number, position, nucleotides in the parental BY reference, nucleotides in the mutant clone) are shown. Green star: heterozygous mutation that became homozygous; Orange star: mutations eliminated in a single passage. Numbers in parenthesis refer to chromosomes shown in (*C*). (*C*), Examples of fixation and elimination of mutations upon LOH and of mutations associated with the occurrence or extension of a nearby LOH event.

Figure 1

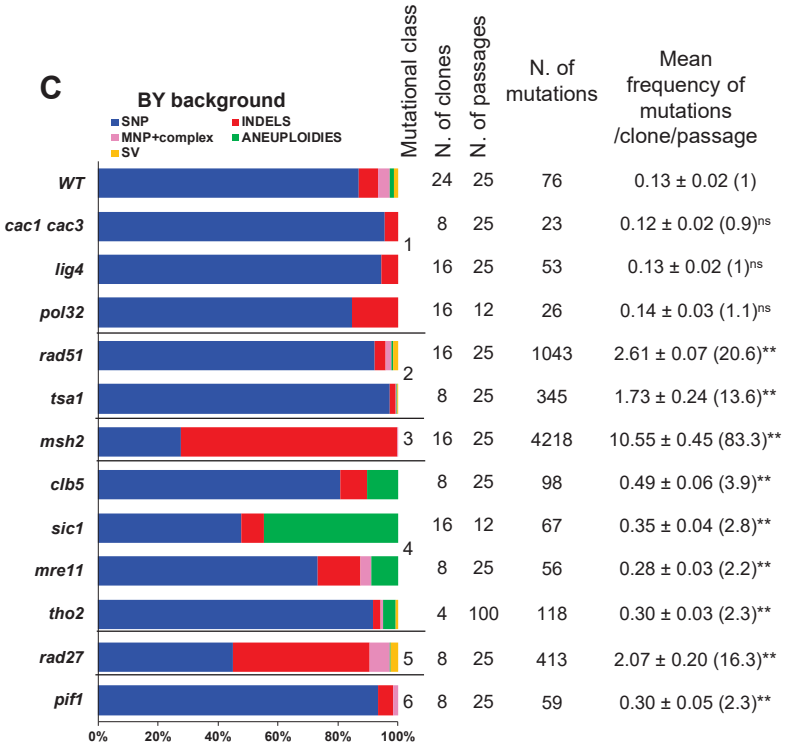
A

Deleted yeast gene	Human ortholog	Biological function	
RAD27	*FEN1*	Replication/repair	*Flap endonuclease involved in DNA replication and long patch base excision repair*
POL32	*POLD3*	Replication	*Subunit of DNA polymerase δ*
MRE11	*hMRE11*	Recombination/repair	*Protein involved in DSB repair. NHEJ and telomere metabolism*
RAD51	*hRAD51*	Recombination	*Strand exchange protein involved in recombinational repair of DNA double-strand breaks*
LIG4	*hLIG4*	Non-homologous end-joining	*DNA ligase required for NHEJ*
MSH2	*hMSH2*	Mismatch repair	*Protein involved in mismatch repair process*
TSA1	*PRDX1*	Oxydative stress response	*Thioredoxin peroxidase*
CAC1 and CAC3	*CAF-1*	Chromatin-assembly factor	*Nucleosome deposition*
PIF1	*hPIF1*	Telomere/mitochondria/ G quadruplex	*5'-3' helicase; telomerase inhibitor; Involved in maintenance of mitochondrial genome. processing of G-quadruplex secondary structures and DNA synthesis*
THO2	*hTHO2*	Transcription	*Subunit of the THO complex involved in transcriptional elongation-associated recombination*
CLB5	*CCNB1*	Cell cycle / Chromosome segregation	*B-type cyclin involved in DNA replication during S-phase*
SIC1	*KIP1*	Cell cycle / Chromosome segregation	*Cyclin-dependent kinase inhibitor*

B



C



D

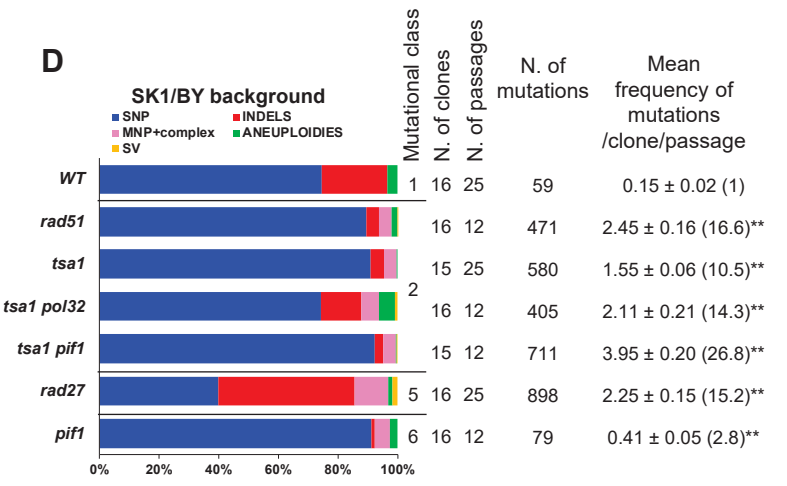


Figure 2

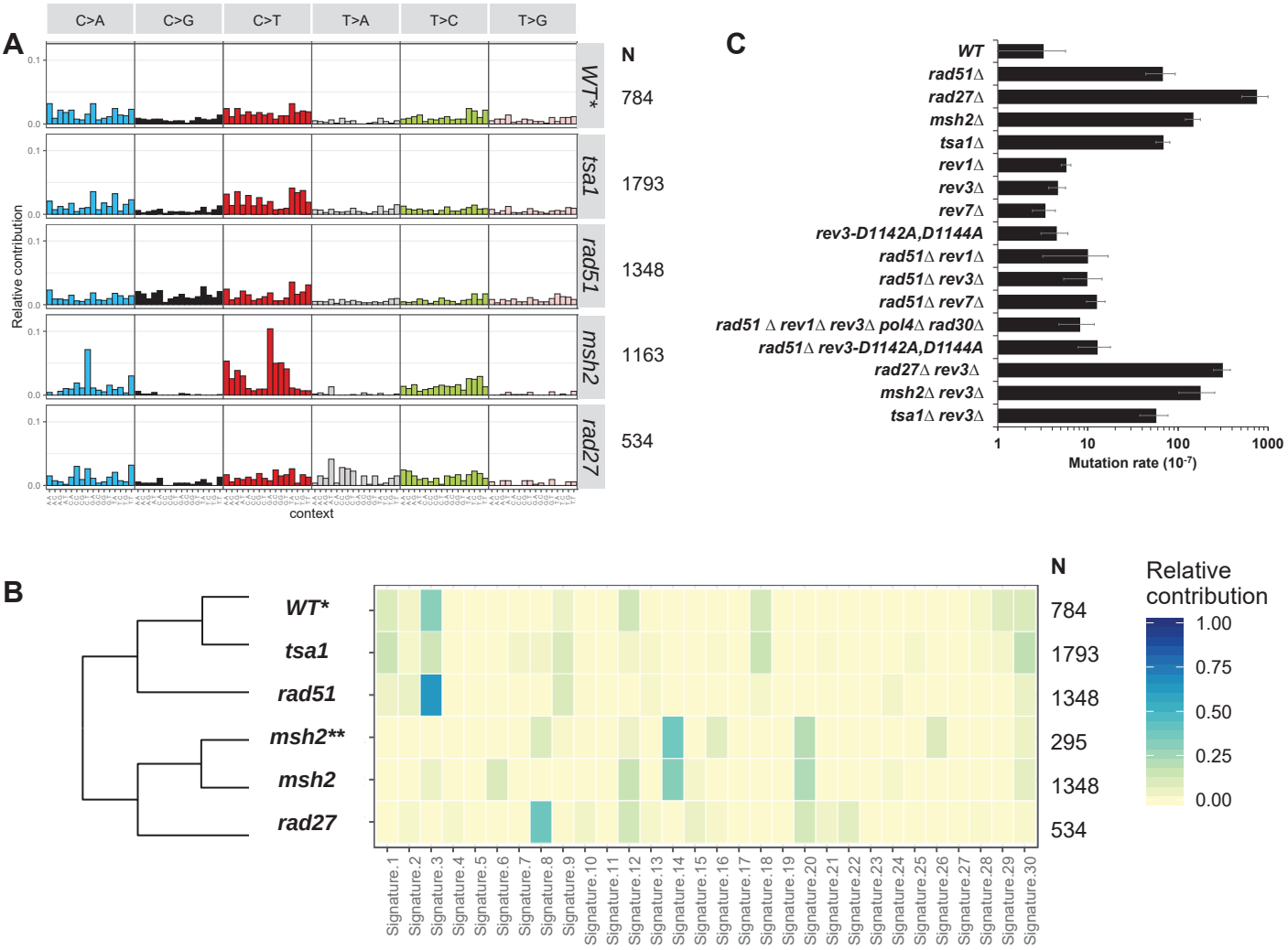

A
C>A
C>G
C>T
T>A
T>C
T>G
N
WT*
784
tsa1
1793
rad51
1348
msh2
1163
rad27
534
Relative contribution
context
C
WT
rad51Δ
rad27Δ
msh2Δ
tsa1Δ
rev1Δ
rev3Δ
rev7Δ
rev3-D1142A,D1144A
rad51Δ rev1Δ
rad51Δ rev3Δ
rad51Δ rev7Δ
rad51 Δ rev1Δ rev3Δ pol4Δ rad30Δ
rad51Δ rev3-D1142A,D1144A
rad27Δ rev3Δ
msh2Δ rev3Δ
tsa1Δ rev3Δ
1
10
100
1000
Mutation rate (10-7)
B
WT*
tsa1
rad51
msh2**
msh2
rad27
N
784
1793
1348
295
1348
534
Relative contribution
1.00
0.75
0.50
0.25
0.00
Signature.1
Signature.2
Signature.3
Signature.4
Signature.5
Signature.6
Signature.7
Signature.8
Signature.9
Signature.10
Signature.11
Signature.12
Signature.13
Signature.14
Signature.15
Signature.16
Signature.17
Signature.18
Signature.19
Signature.20
Signature.21
Signature.22
Signature.23
Signature.24
Signature.25
Signature.26
Signature.27
Signature.28
Signature.29
Signature.30

Figure 3

A

Parental strain:	Number of heterozygous mutations: *1 alternate allele + reference allele*	Number of heterozygous mutations: *2 alternate alleles + no reference allele*	Number of homozygous mutations: *1 alternate allele + no reference*		Copy number: 2	Copy number: <2	Copy number: >2
WT	125	0	3	*(2.3%)*	3	0	0
rad27	1045	9	163	*(13.4%)*	141	16	6
tsa1	848	1	60	*(6.6%)*	54	2	4
pif1	116	0	2	*(1.7%)*	1	1	0
clb5	82	0	1	*(1.2%)*	1	0	0
mre11	44	0	5	*(10.2%)*	4	1	0
cac1 cac3	21	0	2	*(8.7%)*	2	0	0
pol32	25	0	1	*(3.8%)*	1	0	0
sic1	34	0	1	*(2.9%)*	1	0	0
msh2	2801	48	23	*(0.8%)*	22	1	0
lig4	53	0	0	*(0.0%)*	0	0	0
rad51	1421	2	18	*(1.2%)*	7	11	0
tho2	107	0	5	*(4.5%)*	4	1	0
tsa1 pol32	323	1	40	*(11.0%)*	17	20	3
tsa1 pif1	654	1	27	*(4.0%)*	25	2	0

B

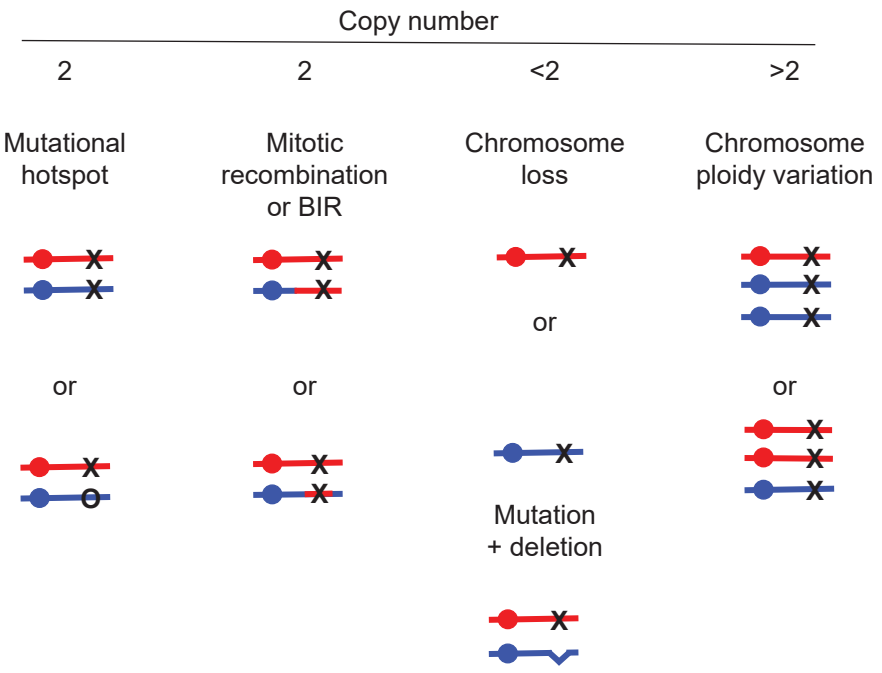
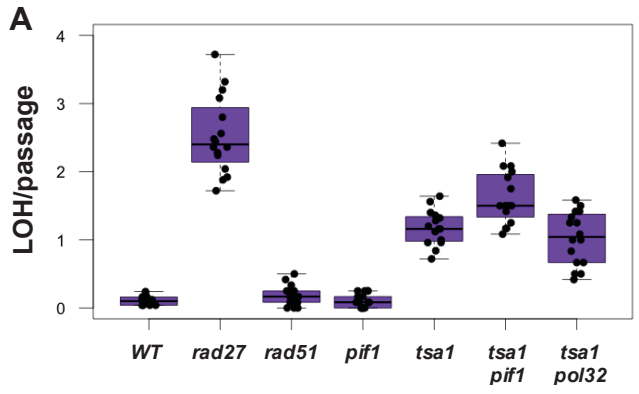
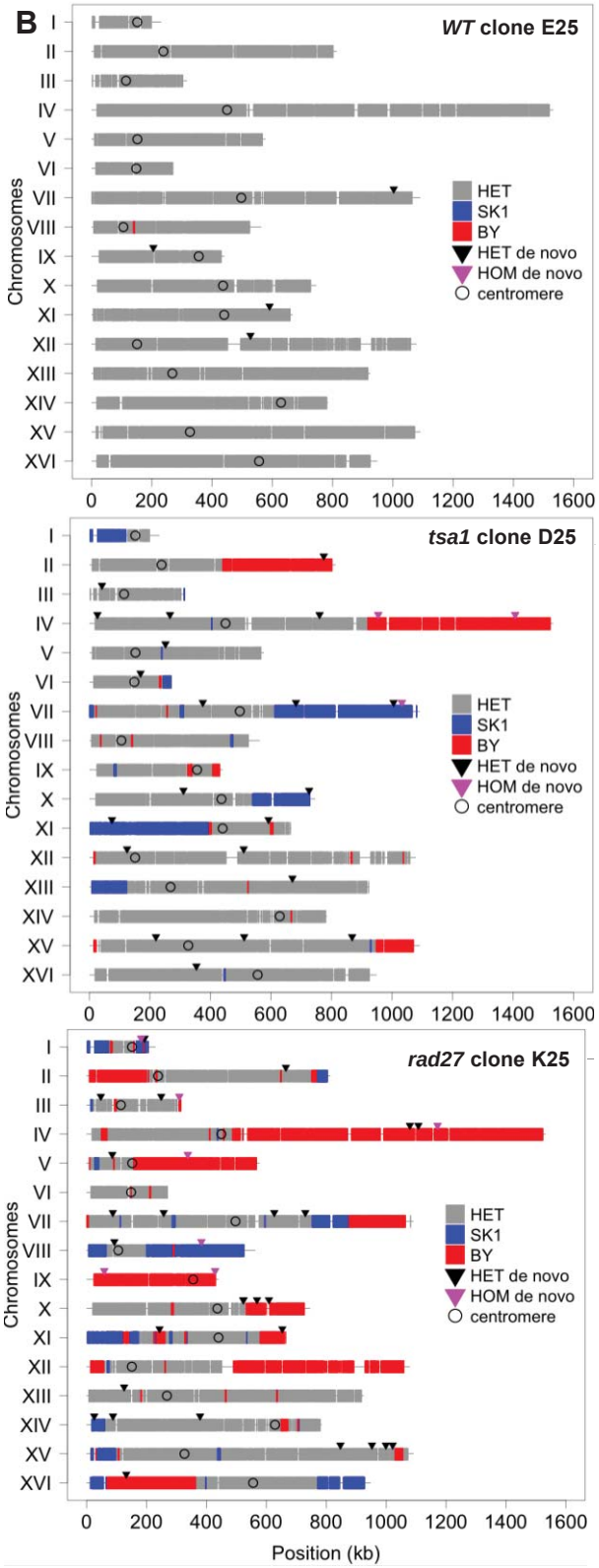
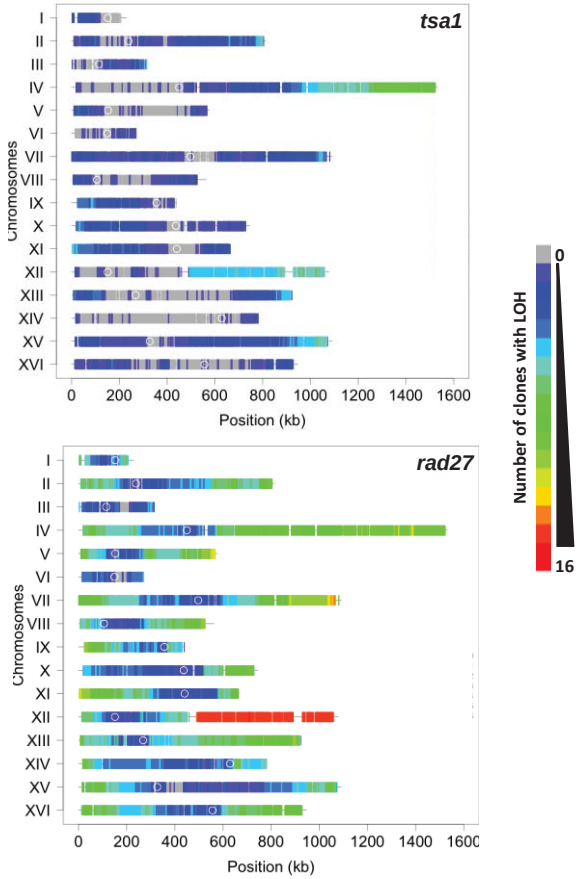


Figure 4

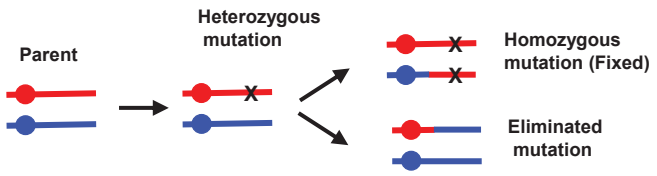






D

SK1/BY mutant	Number of homozygous *de novo* mutations	in LOH region				Not in LOH
		in REC	in chr loss	in deletion	Total	
WT	2	0	0	0	0 (0%)	2 (100%)
rad27	114	95	3	10	108 (95%)	6 (5%)
rad51	11	3	5	0	8 (73%)	3 (27%)
pif1	1	0	1	0	1 (100%)	0 (0%)
tsa1	22	16	0	2	18 (82%)	4 (18%)
tsa1 pif1	27	22	0	3	25 (93%)	2 (7%)
tsa1 pol32	40	20	17	2	39 (98%)	1 (3%)



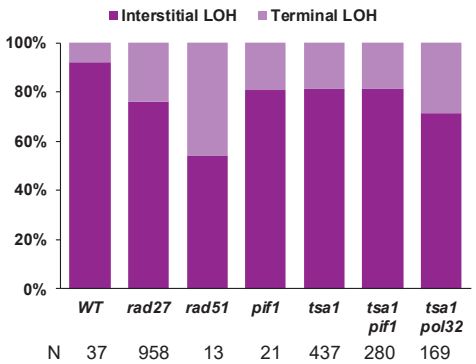


Figure 5

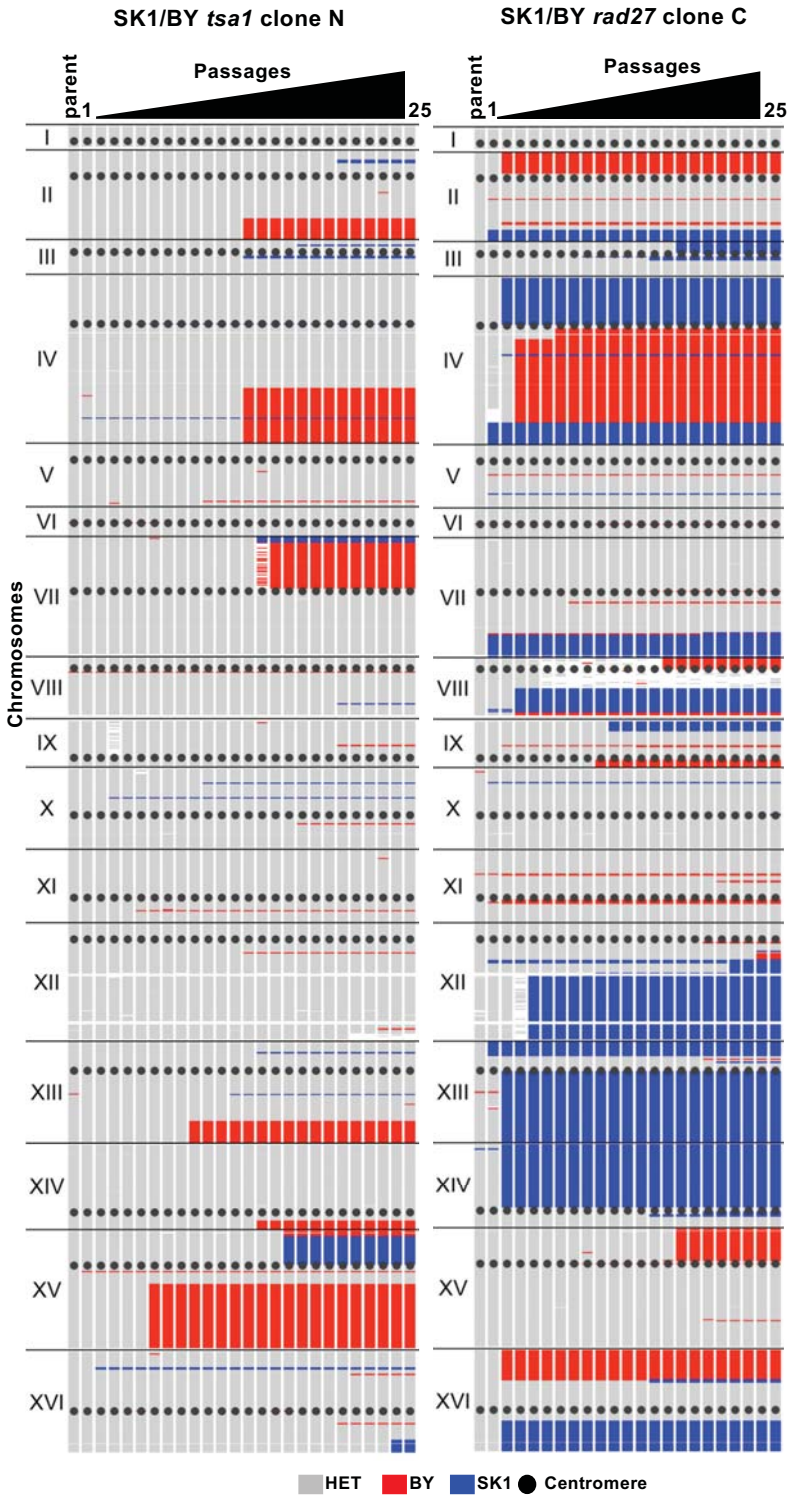
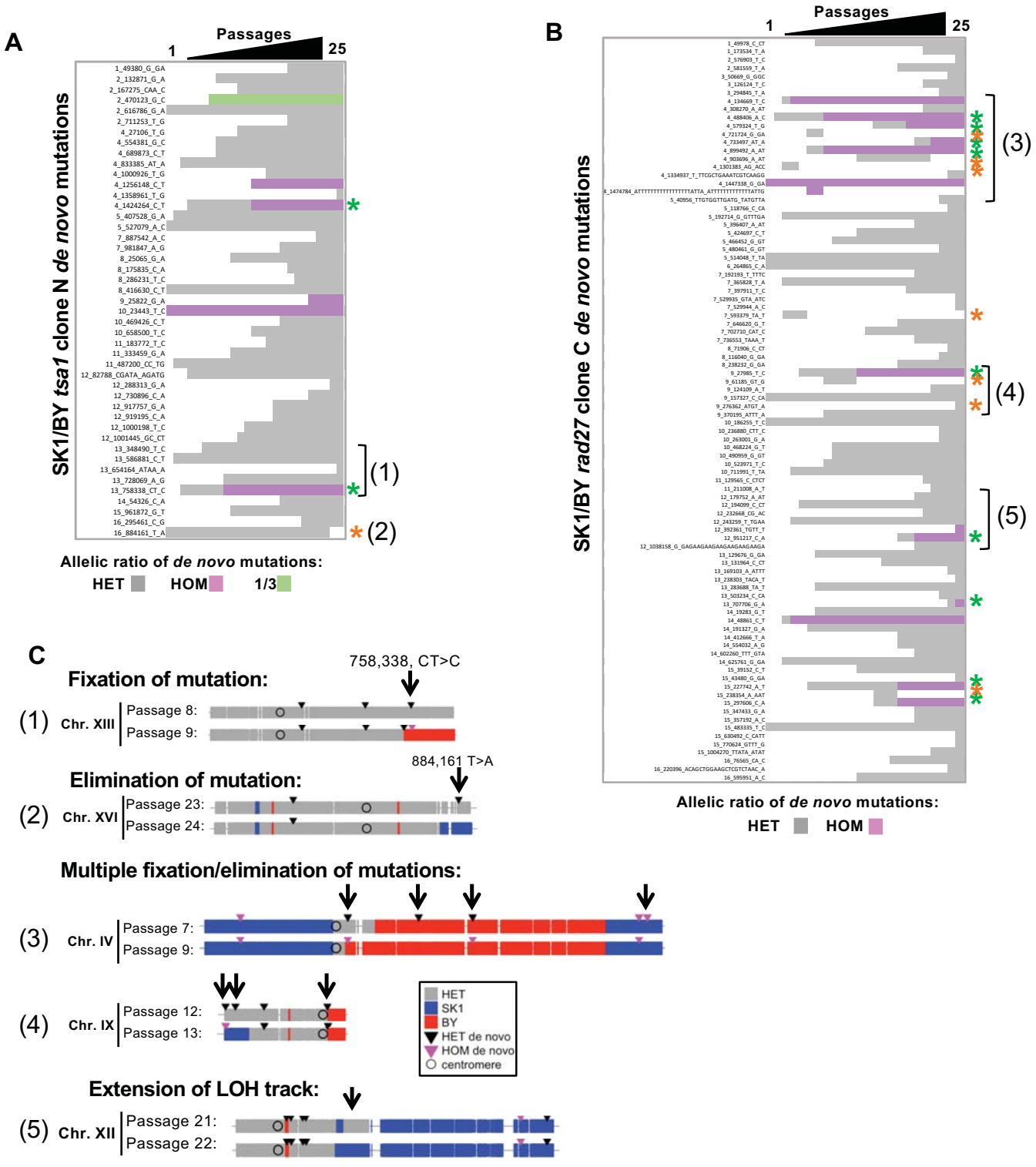

SK1/BY tsa1 clone N
SK1/BY rad27 clone C
parent
Passages
1
25
Chromosomes
I
II
III
IV
V
VI
VII
VIII
IX
X
XI
XII
XIII
XIV
XV
XVI
HET
BY
SK1
Centromere

Figure 6


A
Passages
1
25
SK1/BY tsa1 clone N de novo mutations
(1)
(2)
Allelic ratio of de novo mutations:
HET
HOM
1/3
B
Passages
1
25
SK1/BY rad27 clone C de novo mutations
(3)
(4)
(5)
Allelic ratio of de novo mutations:
HET
HOM
C
Fixation of mutation:
758,338, CT>C
(1) Chr. XIII
Passage 8:
Passage 9:
Elimination of mutation:
884,161 T>A
(2) Chr. XVI
Passage 23:
Passage 24:
Multiple fixation/elimination of mutations:
(3) Chr. IV
Passage 7:
Passage 9:
(4) Chr. IX
Passage 12:
Passage 13:
HET
SK1
BY
HET de novo
HOM de novo
centromere
Extension of LOH track:
(5) Chr. XII
Passage 21:
Passage 22: