LEM domain proteins control the efficiency of adaptation through copy number variation

Paolo Colombi, Diane E. King, Jessica F. Williams, C. Patrick Lusk and Megan C. King

Department of Cell Biology, Yale School of Medicine, New Haven, CT USA 06520

Correspondence to: megan.king@yale.edu

Abstract

While loss of genome integrity is at the basis of numerous pathologies, including cancer, genome plasticity is necessary to adapt to a changing environment and thus is essential for long-term organismal success. Here we present data supporting a targeted mechanism that promotes adaptation to environmental stress by driving site-specific genome instability tied to transcriptional induction and the formation of RNA-DNA hybrids. Using an *in vitro* evolution assay we observe that the inner nuclear membrane LEM domain proteins Heh1 and Heh2 play antagonistic roles in inhibiting or promoting adaptation through copy number expansion, respectively, which is also reflected in their genetic interaction networks with genes responsible for transcription-dependent genome instability. Taken together, our data suggest the existence of a LEM domain protein-mediated mechanism by which an immediate transcriptional response to a changing environment drives targeted genome instability to promote increased variation on which selection can act to support long-term adaptation.

Introduction

The success of an organism relies in part on its ability to maintain fitness in the face of a changing environment, often sensed by increased cellular stress. It has been proposed that multiple mechanisms drive adaptation, differing in the time scale over which they take place, how long they can be sustained, their heritability, and their reversibility(Yona et al., 2015). These mechanisms fall into two major classes: non-genetic changes rooted in physiological responses that drive transcriptional and/or epigenetic modifications, and genetic changes occurring through point mutations and/or gene copy number variation (CNV). A model in which rapid, non-genetic physiological responses could drive adaptive genetic changes has been suggested to couple these two mechanisms [1]. Consistent with this model, recent work suggests that genes poised to respond transcriptionally to environmental challenges may also be subject to "stimulated" adaptive copy number variation of these genes [2]. However, the mechanisms coupling transcriptional up-regulation to local genome instability that can drive copy number variation have yet to be fully delineated.

In neo-Darwinian theory, genetic change occurs gradually and constantly in a stochastic fashion independent of the environment, becoming fixed in the genome as a consequence of providing a fitness advantage. However, in principle an inducible system where bursts of mutations are produced only in conditions of stress [3, 4] provides advantages over a system based on a constitutive, high mutator phenotype that could prove deleterious under conditions where an organism is well-adapted [5]. Indeed, Barbara McClintock provided the first evidence that elevated insertion-deletion

(indel) events, CNV, and other genome rearrangements are driven by elevated transposon activity in response to stress, which acts independently of stochastic environmental or metabolic damage to the DNA [6]. A similar phenomenon has been observed in prokaryotes, in which a mutagenic “SOS response” and DNA break repair mechanisms are up-regulated in response to stress (reviewed in [7]. Interestingly, the role of “stress induced mutagenesis” (SIM) might occur in diverse eukaryotes from yeasts [8-10], to algae [11], nematodes [12], and human cancer cells [13, 14], suggesting its broad conservation. However, as the SOS response appears to be prokaryote-specific, we lack a coherent framework that explains the mechanistic details of how stress may drive mutagenic processes in eukaryotic organisms.

Physiological changes including transcriptional induction and the landscape of epigenetic modifications that facilitate the rapid response to stress have also been suggested to drive genome instability [15, 16], thereby coupling non-genetic and genetic sources of adaptive potential, although whether high levels of transcription drive an increase in mutation rate in the absence of selection remains debated [17-19]. The ribosomal DNA repeats in *S. cerevisiae* provide an example of a genomic locus undergoing controlled genetic changes in a transcription-dependent manner that is linked to epigenetic modifications [20-24]. In this case, the repetitive nature of the rDNA supports frequent CNV (also called “repeat instability”); likely fitness advantages provide a selective pressure for maintaining an ideal rDNA copy number, for example cultures of strains engineered to have decreased rDNA copy number rapidly “acquire” additional rDNA copies [25]. Hints that an active mechanism may explain such observations have

emerged. For example, it has been suggested that cells possess mechanisms to control rDNA copy number by modulating either the frequency of local DNA stability within the locus and/or the mechanism of homology-directed repair (HDR) used to repair local DNA lesions [26, 27]. One influence on this process may be the sub-nuclear compartmentalization of the rDNA in the nucleolus, which is associated with the inner nuclear membrane (INM) in yeasts. Indeed, disrupting an rDNA tether to the nuclear envelope, provided by the integral inner nuclear membrane LEM (LAP2, emerin, MAN1) domain protein Heh1/Src1, increases rDNA CNV [28, 29], although the mechanism remains obscure. Moreover, it remains largely unexplored if similar mechanisms act elsewhere in the genome to modulate the frequency of CNV or point mutations, and, if so, how such mechanisms might contribute to adaptation.

Here we identify a stress-response pathway that modulates cellular adaptation through CNV. Interestingly, the extent of CNV is controlled at the INM, where genetic deletion of *HEH1* drives increased CNV through a mechanism that requires its ohnologue, *HEH2* [30]. Our findings support a model in which Heh2 acts antagonistically to Heh1 through a pathway that leads to the transcription-dependent formation of RNA-DNA hybrids. We suggest a model in which non-genetic changes at specific genetic loci that associate with the nuclear periphery drive local CNV to promote adaptation.

Results

To study the innate mechanisms by which cells adapt to changes in their environment through CNV, we adopted the *S. cerevisiae* multicopy *ENA* gene locus as an

experimental model, which comparative genomics approaches had suggested to undergo high levels of CNV [31]. Genome assemblies of budding yeast strains suggest that up to five *ENA* gene copies (with ≈98/99% identity) reside in a tandem array[32-35]; our WT W303 strain has four tandem *ENA* coding sequences, as confirmed by PCR, which we refer to here as *ENA1-4* (Figure 1A; Supplemental Figure 1A-B). Consistent with the idea that the *ENA* locus encodes pumps responsible for salt efflux, it is rapidly induced ≈3-fold in the presence of 100 mM LiCl (Figure 1B) and is essential under high salt conditions [36, 37] (Figure 1C).

To identify pathways that impact *ENA* CNV, we devised an *in vitro* evolution assay wherein cells are grown in liquid medium containing 100 mM LiCl by serial culturing for ≈200 generations (see Materials and Methods)(Figure 1D). Prolonged culturing in LiCl resulted in WT adapted ("WT-A") strains that tolerate high concentrations (up to 300 mM) of LiCl compared to the WT parental ("WT-P") strain (Figure 1E). To test the underlying basis for improved fitness of WT-A, we measured *ENA* gene copy number by qPCR. The adapted strain expanded its *ENA* gene dosage by 500%, possessing an average of 20 copies per genome (Figure 1F). To differentiate whether the increase in gene copy number occurred through intrachromosomal or extra-chromosomal expansion (historically referred to as a "homogenous staining region" (HSR) or double minutes (DMs), respectively), we tested whether salt tolerance was inherited through a Mendelian or random genetic segregation in meiosis (Figure 1G). The WT-A strains was crossed with a WT strain with a genetic marker integrated adjacent to the *ENA* locus to monitor segregation; WT::*URA3*-P) and meiosis was induced. We observed a

Mendelian inheritance pattern where two of the four progeny (F1 and F4) maintained ≈20 copies of the *ENA* locus (Figure 1F) with concomitant salt tolerance (Figure 1H). These data support the conclusion that cells adapt to high salt through an intrachromosomal expansion of the *ENA* genomic locus (≈20 *ENA* genes; Figure 1F-H), and not through the production of extrachromosomal DMs.

Copy number expansion is typically driven by non-conservative homology-directed repair processes downstream of replication fork collapse or the formation of DNA double-strand breaks[38]. As it has been suggested that high levels of transcription can drive local genome instability [15], and *ENA* expression is transcriptionally induced upon growth in high salt (Figure 1B), we next tested if transcription drives *ENA* locus instability. We devised a genome stability assay similar to that used to reveal the non-uniform rate of mutation across the budding yeast genome[39, 40]. Here, counter-selection for Ura3 activity, which converts 5-Fluoroorotic acid (5-FOA) to the toxic metabolite, 5-fluorouracil, allows the frequency of mutations in (or loss of) *URA3* to be defined. We inserted the *URA3* gene in three different positions at the *ENA* locus (Figure 2A) and measured the spontaneous instability at each position over a single cell cycle (2 hours growth before plating on 5-FOA; loss of *URA3* must occur prior to plating). In all positions, Ura3 activity is lost at rates at least an order of magnitude higher than at the endogenous *URA3* locus ($\approx 10^{-8}$; Figure 2B), suggesting a high level of local instability. As expected due to flanking tandem repeats [41-44], the highest level of instability ($\approx 10^{-5}$) is observed when *URA3* is inserted within the *ENA* copies (Figure 2B ("In"); in this case, the *URA3* gene is always lost, presumably by non-allelic

homologous recombination; Supplemental Figure 2A). Interestingly, the instability downstream (3') of the *ENA* gene locus (≈10^{-6}) is markedly higher than upstream (5'; ≈10^{-7}), suggesting a potential role for *ENA* transcription in modulating *URA3* stability. Consistent with this idea, inducing *ENA* gene expression by addition of LiCl during the single generation (2 hours) of release (from –uracil to 5-FOA) drives a higher frequency of 5-FOA resistant colonies specifically at *URA3* inserted downstream of the *ENA* locus (≈30%). Consistent with the idea that this effect is a direct result of transcription, it was abolished in cells harboring a loss of function allele of RNA polymerase II, *rpb1-1* [45, 46] (Figure 2C). In contrast, the region upstream of the *ENA* locus showed a lower frequency of 5-FOA resistant colonies in the presence of LiCl (≈40%) while the stability of the endogenous *URA3* locus is unaffected (Figure 2C).

5-FOA resistant colonies derived from strains with *URA3* inserted downstream of the *ENA* locus almost all map to the *URA3* gene, with an increase in the number of mutations per *URA3* gene upon exposure to LiCl (from ≈5% to ≈25% of 5-FOA-resistant clones have 2-4 mutations in *URA3*; Figure 2D). By contrast, the mutation profile of the upstream position is very similar to that of the endogenous locus, where ~30% of the 5-FOA-resistant colonies occur outside the *URA3* gene (Figure 2D) and does not change in response to LiCl. A more detailed mutational signature analysis shows that mutations driving 5-FOA-resistance at *URA3* integrated downstream of *ENA* are enriched in transversions (≈90%), which shifts to an increase of transitions (≈30%) and indels (≈15%) after LiCl exposure (Figure 2E); this contrasts with *URA3* integrated upstream of the *ENA* locus, which is unaffected by addition of LiCl and matches the

transition/transversion rate typical of the budding yeast genome as a whole (≈0.6)[47]. At the endogenous *URA3* locus in the absence of LiCl we observe a mutational profile in line with previous studies [39](Figure 2E). Here we also observe a shift from transversions to transitions after LiCl treatment, although these events remain relatively rare given the low rate of recovering mutations at *URA3* (Figure 2B,D).

These observations suggest that local factors might drive genome instability and/or the fidelity of repair [40, 48]. To gain further insight we turned to a genome-wide analysis of single nucleotide polymorphism (SNP) distribution across twenty-one *S. cerevisiae* strains [49, 50], in which we defined peaks of significantly locally high SNP density ("islands"; Figure 2F; see Methods). Ranking SNP islands according to their integrated peak height and width to give rise to a SNP island score (see Methods), we found that the highest ranked islands (18 of the top 20; 41 of the top 50, Supplemental Figure 2B, Table S1) harbor genes that fall into at least one of two categories: subtelomeric genes and genes with paralogues (many of which are ohnologues arising from the whole genome duplication in the *S. cerevisiae* lineage), both of which are established to undergo rapid changes in sequence content [51-53]; indeed, none of the genes located in the top 20 SNP islands are essential. Surprisingly, the essential *KRS1* gene, which lies immediately downstream of the *ENA* locus, scored 21st of all ranked islands (Table S1); the unusually high mutational frequency within *KRS1* was in fact noted in the initial comparative genomic sequence analysis of budding yeast strains [54]. More generally, the entire genomic region downstream of the *ENA* locus has a high SNP load compared to the region upstream of the *ENA* genes (Figure 2F), consistent with the heightened

instability of *URA3* at this position.

To gain insight into the factors that might underlie the high SNP load and relative instability of the *KRS1-ENA* region, we investigated its sub-nuclear localization by integrating a lac Operator array just upstream of *ENA1* and monitored its position in cells expressing GFP-lacI and Hmg1-mCherry, a nuclear envelope/ER marker. Interestingly, the *ENA* locus is more strongly enriched at the nuclear envelope compared to the *URA3* locus (Supplemental Figure 3), raising the possibility that the nuclear periphery might play a role in modulating the stability of this genomic region. As Heh1 was previously implicated in modulating the stability of the rDNA [29], we investigated if Heh1 (and/or its ohnologue, Heh2; Figure 3A) influences the chromosomal expansion of the *ENA* locus in response to salt stress by carrying out the *in vitro* LiCl evolution experiment for WT, *heh1*Δ*, heh2*Δ*,* and *heh1*Δ*heh2*Δ strains. All four genetic backgrounds acquired improved fitness on media containing LiCl (100 mM and 300 mM) after 200 generations of culturing in 100 mM LiCl (Figure 3B). However, we uncovered marked differences in *ENA* copy number in the adapted strains by qPCR depending on genotype. In WT strains, *ENA* copy number is stable at 4 copies over 60 generations but increases to ≈6 at around 90 generations, and ultimately reaches ≈20 by 200 generations (Figure 3C). Interestingly, the *heh1*Δ strain undergoes a more rapid expansion of *ENA* copies, doubling to ≈8 at 60 generations and reaching a maximum copy number of ≈32 by 90 generations, which is maintained until ≈200 generations. Surprisingly, the ablation of the paralogous *HEH2* has the opposite effect, delaying the *ENA* copy number expansion at all time points. Cells lacking both *HEH1* and *HEH2*

behave nearly like WT, suggesting the possibility of antagonism between these two ohnologues on *ENA* locus copy number.

The observed *ENA* copy numbers directly correlate with growth fitness in the presence of LiCl as only *heh1*Δ is resistant to growth in LiCl at ≈60 generations, while at ≈90 generations the *heh2*Δ is the least fit as it only has 4 copies (on average) of the *ENA* genes (Figure 3C-D). The copy number increase is driven by intra-chromosomal expansion as for WT strains in all cases (Supplemental Figure 4) and is reproducible (Supplemental Figure 5). Importantly, the influence of *HEH1* and *HEH2* on copy number expansion cannot be explained by changes in inherent salt tolerance or *ENA* expression at baseline as assessed by RT-qPCR (Figure 3B and E). Moreover, levels of the *ENA* transcript increase proportionally with gene copy number in the adapted strains, although the relative influence of LiCl on *ENA* transcript levels is diminished with increasing copy number (in the WT and *heh1*Δ backgrounds, compare parental to adapted strains, Figure 3E). Lastly, to gain insight into the stability of the expanded *ENA* copy number in the absence of salt stress, we serially re-streaked the adapted strains on rich media. Interestingly, only the *heh1*Δ strain shows a loss of *ENA* copy number (almost 50%), while the other strain backgrounds remain stable (Figure 3F, Supplemental Figure 5). This suggests that loss of Heh1 increases CNV in an unbiased fashion, while selection acts to determine if the *ENA* copy number expands or contracts depending on fitness for the environment.

To understand how *HEH1* and *HEH2* achieve this effect, we investigated their synthetic

genetic interaction networks [55-60]. Interestingly, *HEH1* shows synthetic sickness with genes involved at different stages of transcription: *TOP1*, which prevents negative DNA supercoiling, the THO-TREX/TREX2 complex, which influence transcriptional elongation and termination (coordinated with mRNA export), and *XRN1* and *RRP6*, which participate in RNA surveillance and degradation. All these genes play a major role in preventing the formation or persistence of RNA-DNA hybrids (R-loops) that constitute a major threat to genome stability [61]. Intriguingly, *HEH2* consistently displays the opposite effect of *HEH1*, acting as a genetic suppressor of the same genetic network. For example, deletion of *HEH2* substantially rescues the growth of synthetically sick *heh1*Δ*sac3*Δ, *heh1*Δ*xrn1*Δ and *heh1*Δ*rrp6*Δ strains (Figure 4B, Supplemental Figure 6). Interestingly, the deletion of the Heh2 winged-helix domain is enough to partially rescue the fitness loss of *heh1*Δ*sac3*Δ (Figure 4B). These genetic interactions suggest the possibility that Heh1 and Heh2 might modulate R-loop formation or resolution, which we first tested genetically by examining interactions between *HEH1* and *HEH2* and the ribonuclease H1 and H2 enzymes involved in the removal R-loops in *S. cerevisiae*, encoded by *RNH1* and *RNH201* [62]. Loss of *HEH2* is suppressive of the growth defect of the *rnh1*Δ*rnh201*Δ genotype in the presence of hydroxyurea, driving a marked increase in fitness with the colony size increasing by ≈75% (Figure 4C).

To directly test how *HEH1* and *HEH2* influence the accumulation of R-loops at the *ENA* locus, we carried out targeted RNA-DNA immunoprecipitation (DRIP) using the monoclonal antibody S9.6 [63] followed by qPCR. We detected robust RNA-DNA hybrids across the *ENA* region (Figure 4D). However, there is a marked asymmetry,

with the gene *KRS1* accumulating almost twice the extent of hybrids detected at the *ENA* or *RSM10* genes, thus correlating with the stability of the exogenous *URA3* gene inserted at these locations (Figure 2). Interestingly, the deletion of *HEH1* drives an overall increase of RNA-DNA hybrids, most dramatically at *KRS1* (Figure 4D); *heh2*Δ*,* or the combined *heh1*Δ *heh2*Δ genetic backgrounds do not produce any measurable change compared to WT cells, suggesting that loss of *HEH1* only drives accumulation of RNA-DNA hybrids when *HEH2* is present. Taken together, these orthogonal approaches suggest that the presence of *HEH2* favors the formation RNA-DNA hybrids, while *HEH1* opposes this activity.

Discussion

This work provides evidence for pathway in which the act of transcription and the LEM domain protein *HEH2* promote CNV at the *ENA* locus, while *HEH1* opposes this activity. In cells lacking *HEH1*, RNA-DNA hybrids accumulate downstream of the *ENA* locus, which is associated with a more rapid rate of both CNV expansion (in the presence of selective pressure) and contraction (in the absence of selective pressure). This model is further supported by a network of opposing genetic interactions between *HEH1* and *HEH2* and genes encoding factors influencing transcription termination, RNA turnover, or metabolizing RNA-DNA hybrids. We speculate that the expansion of the *ENA* locus and salt tolerance provides just one example of the biological contexts in which specific sites in the genome undergo this type of stress-induced CNV pathway, although further studies will be required test this notion.

While it has previously been appreciated that tandem genomic repeats frequently undergo reciprocal exchanges that can drive changes in copy number [38] (which can then be acted upon by selection) and that accumulation of RNA-DNA hybrids is tied to local, increased genome instability [61], these processes are largely thought to take place due to faulty repair mechanisms or incidental (and often deleterious) effects. At the same time, substantial evidence supports the notion that CNV is exploited to promote adaptation in yeasts [64, 65], but also in a broad array of other organisms and context, for example resistance to antibiotics in prokaryotes or chemotherapy resistance in the setting of cancer (reviewed in [7, 66]). Here we provide evidence that the antagonistic functions of conserved LEM domain proteins [67], likely through modulating RNA-DNA hybrids, influence the local loss of "genome stability" and fuel CNV-dependent adaptation. Our findings extend previous recent observations that CNVs can be driven in response to environmental stress in a manner that requires both transcription [2] and its associated histone modifications [2, 68]. Importantly, here we find that such processes result not only from stochastic losses of genome integrity resulting from collisions between the replication fork and RNA-DNA hybrids and subsequent repair processes, but also from a balance of *HEH1*- and *HEH2*-dependent influences that regulate a local loss of genome stability in response to physiological cues. Importantly, in this context we conclude that spontaneous gain/loss of copy number and selection for copy gain (growth in LiCl) alone cannot explain our observations.

How might the balance of LEM domain protein activities influence RNA-DNA hybrid

formation? It is interesting to note that paralogues (in this case the *HEH1* and *HEH2* ohnologues arising from the whole genome duplication in the history of *S. cerevisiae*) often evolve to carry out antagonistic functions [69, 70]. The genetic interaction network of *HEH1* and *HEH2* (Fig. 4) supports the notion that these factors antagonize each other to regulate either transcriptional termination, leading to stalled RNA-DNA hybrids, or instead the removal of RNA-DNA hybrids by RNase H. Interestingly, the domain of Heh2 necessary to impart its activity, its C-terminal winged helix domain, is structurally shared among the RNA-interacting components of the THO-TREX complex [71], although further studies will be necessary to define the biochemical activities or interactions by which these LEM domain proteins influence the accumulation of RNA-DNA hybrids.

In addition to changes in copy number, here we also demonstrate that the ENA locus is also a hot spot for point mutations, in a manner that is at least partly influenced by its transcriptional response to high concentrations of LiCl. For example, the rates of mutation within *URA3* 5' and 3' to the *ENA* gene cluster is substantially higher than the median rate observed across tens of *URA3* insertion sites on chromosome VI [40], raising the question of why this locus is so unusually unstable. As the *ENA* locus is replicated early [72], and the occupancy of Pol II is similar both up and downstream of the ENA locus [73], we favor the idea that, as was described for the acquisition of suppressor of ochre mutations in budding yeast due to non-random mutations to one of the six tRNA-Tyr genes [74], the orientation of the *ENA* transcriptional unit (and the associated RNA-DNA hybrids) drives collisions with the replication fork that initiates

from a nearby ARS, only 21 kbps 5' to the *ENA* gene cluster. The change of the mutational signature in the presence of LiCl downstream of the *ENA* locus is compatible with the mutation profile observed for DNA polymerases involved in gene conversion (Polδ and Polε) interrogated at another highly unstable region of the budding yeast genome, the mating type locus [75], suggesting that they may be products of DNA repair acting after replication fork collapse, although translesion synthesis, which can drive single base pair substitutions, may also play a role [76-78]. This model echoes the observations that the recombinase RecA and DNA synthesis are required for "directed mutagenesis" in bacteria [79, 80] but supports the existence of mechanisms that drive a "regional" (rather than random) mutagenesis followed by selection rather than a preference for adaptive mutations.

Oncogene amplification represents a major driver of carcinogenesis and a challenge to cancer therapy [81-83]. The results presented here suggest that adaptive mechanisms that drive CNV can be disabled to combat oncogenesis and therapy resistance. The antagonistic effect of *HEH1* and *HEH2* on this process highlights the need for further investigation into how members of the LEM domain family in mammals, which includes LAP2, emerin, LEM2 and MAN1, impact on genome integrity; to date this potential connection has gone largely uninvestigated [84].

Materials and Methods

Yeast strain generation and culturing. All yeast strains used in this study and their derivation are listed in Table S2. Unless otherwise stated, all experiments were conducted at 30°C. The *rpb1-1* strain (derived from KWY1302, a gift K. Weis, ETH Zurich, Switzerland) was cultured at room temperature (RT) as described in the Figure Legends. All strains were grown in YP (1% yeast extract and 2% peptone) with 2% dextrose (YPD) with the addition of different concentration of LiCl, as described. Standard yeast manipulations including transformations, tetrad dissection, and PCR-based integration were performed as described[85]. The LacO integrations were generated as described[86].

***In vitro* evolution experiment.** The strains were initially grown overnight at 30°C in 2 ml YPD. In the morning, the density of the culture was quantified using an automated cell counter (Moxi^{z} - Orflo) and 50,000 cells were diluted into 50 ml of YPD with100mM LiCl and grown for 24h. This procedure of quantification and dilution of the culture was repeated for 18 days (~200 generations). Every day samples were collected and stored for fitness and gene copy number analysis.

Reverse transcription PCR and qPCR. RNA was prepared using the MasterPure Yeast RNA purification Kit (Epicentre) according to manufacturer's instructions from cells growing in exponential phase at 30°C in YPD, and after 10 min exposure to YPD with 100mM LiCl. DNA contamination was removed by treating samples with DNaseI for 45 min at 37°C. cDNA was synthesized from 500ng of total RNA using Superscript® III

First-Strand Synthesis (Invitrogen) with oligo-dT primers. cDNA was added to the iTaq™ Universal SYBR® Green supermix (BIORAD) with primers to amplify *ENA1-4* (the primers anneal in regions that are identical between all 4 gene copies) and *ACT1* as internal load control. Reactions mixes were cycled in a CFX96 Touch Real-Time PCR Detection system (BIORAD). To calculate relative gene expression we used the $2^{-\Delta\Delta Ct}$ method of analysis[87]. Log2 $2^{-\Delta\Delta Ct}$ values were from three independent experiments were normalized to control condition (0mM LiCl) and plotted as shown in Figs. 1B, 3E. Primers listed in Table S3.

Genomic DNA extraction and copy number quantification by qPCR. Genomic DNA was prepared using a modified Winston method. Cells were grown overnight to saturation in YPD, washed with 1 ml of water and resuspended in 10 mM Tris-HCl, pH 8.0, 2% TRITON X-100, 1% SDS, 100 mM NaCl, 1 mM EDTA. 200 µl of phenol:chloroform:isoamyl alchol (25:24:1) (Fisher Scientific) and 100 µl of glassbeads were added to the cell suspension and vortexed for 5 min. After adding 200 µl of 10 mM Tris-HCl pH 8.0, 1 mM EDTA (TE), the suspension was centrifuged for 5 min at 17000 x g, the supernatant transferred to a new tube and genomic DNA was precipitated with 100% ethanol. The pellets were washed with 70% ethanol, air dried and resuspended in TE containing 50 µg/µl RNase. The method adopted for measuring copy number was adapted from Weaver et al.[88] with minor modifications. Briefly, we used the following approach: $2^{-\Delta\Delta Cq}$ based on a target assay T (*ENA*) for the DNA segment being interrogated for copy number variation and a reference assay R (*ALG9*) for an internal control segment which is a single copy gene. The $\Delta C_q = (C_{q,ENA}-C_{q,ALG9})$ is a measure of

the copy number of the target segment (*ENA*) relative to the reference segment (ALG9). The next step in determining the relative copy number is to calibrate the ΔC_q value to a sample with a single copy number for the target (*ENA*) and for reference gene (*ALG9*) $\Delta C_{q,C} = (C_{q,ENA}\text{-}C_{q,ALG9})$. Assuming that the efficiencies of the target and reference assay are similar and close to 1[87], the relative copy number is calculated from the formula $2^{-\Delta\Delta Cq}$ where $\Delta\Delta C_q = \Delta C_q - \Delta C_{q,c}$. For every target and calibrator sample, three different concentrations of genomic DNA were tested, each in triplicate, thus allowing us to evaluate the efficiency for every sample analyzed.

Genome stability assay. Cells were grown in selective media (CSM-URA) for two days starting from a single colony obtained from a freshly re-streaked strain. After two days, the cells were resuspended in complete YPD media for a single cell cycle (2 hrs. for WT strains) and then plated on YPD plates (to determine the total number of cells plated) and on 5-fluoorotic acid (5-FOA)-containing plates to identify the number of cells that lost URA3 activity. In the case of LiCl treatment, cells were released for a single cell cycle (2 hrs. for WT) in YPD containing 100 mM LiCl. The data are reported as a ratio between the number of cells growing in 5-FOA plates versus the total number of cells measured on YPD plates.

Microscopy. For the imaging experiments, cells were grown to mid-log phase and immobilized on a 1.4% agarose pad containing complete synthetic medium (CSM) with 2% glucose and sealed with VALAP (1:1:1 Vaseline/lanolin/paraffin). The microscopy experiments were carried out on a wide-field deconvolution microscope (DeltaVision;

Applied Precision/GE Healthcare) equipped with a 100×, 1.40 NA objective lens and solid state illumination. The images were acquired using an Evolve EMCCD camera (Photometrics). Temperature control was achieved through the enclosure of the microscope within an environmental chamber. In all cases, a z-series of images with 200 nm spacing were acquired and further processed as described under "Image processing and analysis."

Image processing and analysis. The 3D reconstruction of the nuclear envelope, fitting of the LacI-GFP/LacO focus and the position of the LacI-GFP/LacO with respect to the nuclear envelope were determined as described in our published work[89].

Single nucleotide polymorphism analysis. Variant data was obtained for 21 *S. cerevisiae* strains, kindly provided by Dr. J. Michael Cherry[49], which includes BC187, BY4741, BY4742, CEN.PK2-1Ca, D273-10B, DBVPG6044, FL100, FY1679, JK9-3d, K11, L1528, RedStar, RM11-1A, SEY6210, Sigma1278b-10560-6B, SK1, UWOPS05_217_3, W303, X2180-1A, YPS163, and YS9. The genome build used as the reference and for plots was UCSC sacCer3. Individual SNPs were excluded if any of the following criteria were met: allele balance for heterozygous genotype > 0.75, read depth > 360, strand bias > -0.1, or those flagged for low quality. These 21 filtered files were combined and a variant frequency file was generated, which lists the number of strains for which a variant was identified at each position for each chromosome. The variant frequency file was divided by the total number of variants across the genome (729,305) to make a distribution file. Peaks were called (and island scores were

determined) using SICER[90] with the following settings: window size = 100 bp, gap size = 100 bp, E-value = 100, FDR of 2.1%.

Epistasis analysis. Genetic interactions were assessed by spotting equivalent numbers of cells in six 10-fold serial dilutions incubated at 30°C for 2-4 days. For epistasis analysis the size of 60 colonies from three independent growth assays were measured using Fiji[91]. For each replicate the strains tested were isolated directly from spores obtained from tetrad dissection.

DNA:RNA immunoprecipitation (DRIP). Yeast strains grown in YPD were harvested in exponential phase, washed with 1 ml of water and frozen in liquid nitrogen. The cells were lysed with RA1 buffer (Macherey-Nagel) supplemented with 100 mM NaCl and 1% β-mercapthoethanol. The cell suspension was mixed with phenol:chloroform:isoamyl alcohol (25:24:1) (Fisher Scientific) and glass beads. Cells were broken by mechanical shaking using a pulsing vortex mixer. After spinning, the upper phase was transferred to a new tube and nucleic acids were precipitated with 1 ml of isopropanol followed by centrifugation. Pellets were washed with 70% ethanol, air dried, resuspended in 50 mM Tris-HCl, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 5 mM EDTA and sonicated with a Bioruptor (Diagenode) to obtain 100-500 bp fragments. Fifteen micrograms of sonicated nucleic acids were diluted in 1 ml of IP buffer (0.1% SDS, 1% Triton X-100, 10 mM HEPES pH 7.7, 0.1% sodium deoxycholate, 275 mM NaCl) and incubated overnight on a rotating wheel at 4°C with 1 μg of S9.6 antibody (Kerafast), followed by precipitation with 25 μl of Dynabeads protein G (Life Technologies) pre-blocked with bovine serum

albumin and *E. coli* DNA. Beads were washed 5 times with IP buffer, and the DNA fragments were collected by PCR cleanup kit (QIAGEN). The collected DNA fragments were quantified by qPCR using the primers listed in Table S3, and values for DRIP were calculated using the formula ΔC_q"- antibody"= $2^{-(Cq \text{ "beads only"} - Cq \text{ "input chromatin"})}$ and ΔC_q"+ antibody"= $2^{-(Cq \text{ "beads only"} - Cq \text{ "input chromatin"})}$. The values were normalized to *CEN16*, which was set to a value of 1 in order to compensate for differences in immunoprecipitation efficiency[92].

References

1. Yona AH, Frumkin I, Pilpel Y. A Relay Race on the Evolutionary Adaptation Spectrum. Cell. 2015;163(3):549-59. doi: 10.1016/j.cell.2015.10.005.
2. Hull RM, Cruz C, Jack CV, Houseley J. Environmental change drives accelerated adaptation through stimulated copy number variation. PLoS Biol. 2017;15(6):e2001333. doi: 10.1371/journal.pbio.2001333.
3. Ram Y, Hadany L. The evolution of stress-induced hypermutation in asexual populations. Evolution. 2012;66(7):2315-28. doi: 10.1111/j.1558-5646.2012.01576.x.
4. Ram Y, Hadany L. Stress-induced mutagenesis and complex adaptation. Proceedings of the Royal Society of London B: Biological Sciences. 2014;281(1792):20141025. doi: 10.1098/rspb.2014.1025.
5. Funchain P, Yeung A, Stewart JL, Lin R, Slupska MM, Miller JH. The consequences of growth of a mutator strain of Escherichia coli as measured by loss of function among multiple gene targets and loss of fitness. Genetics. 2000;154(3):959-70.
6. McClintock B. The significance of responses of the genome to challenge. Science (New York, NY). 1984;226(4676):792-801.
7. Fitzgerald DM, Hastings PJ, Rosenberg SM. Stress-Induced Mutagenesis: Implications in Cancer and Drug Resistance. Annual Review of Cancer Biology. 2017;1(1):119-40. doi: 10.1146/annurev-cancerbio-050216-121919.
8. Heidenreich E. Adaptive Mutation in Saccharomyces cerevisiae. Critical Reviews in Biochemistry and Molecular Biology. 2008. doi: 10.1080/10409230701507773.
9. Rodriguez GP, Romanova NV, Bao G, Rouf NC, Kow YW, Crouse GF. Mismatch repair-dependent mutagenesis in nondividing cells. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(16):6153-8. doi: 10.1073/pnas.1115361109.
10. Shor E, Fox CA, Broach JR. The yeast environmental stress response regulates mutagenesis induced by proteotoxic stress. PLoS Genetics. 2013;9(8):e1003680. doi: 10.1371/journal.pgen.1003680.
11. Goho S, Bell G. Mild environmental stress elicits mutations affecting fitness in Chlamydomonas. Proceedings of the Royal Society of London B: Biological Sciences. 2000;267(1439):123-9. doi: 10.1098/rspb.2000.0976.
12. Matsuba C, Ostrow DG, Salomon MP, Tolani A, Baer CF. Temperature, stress and spontaneous mutation in Caenorhabditis briggsae and Caenorhabditis elegans. Biology Letters. 2013;9(1):20120334. doi: 10.1098/rsbl.2012.0334.
13. Bindra RS, Glazer PM. Genetic instability and the tumor microenvironment: towards the concept of microenvironment-induced mutagenesis. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 2005;569(1-2):75-85. doi: 10.1016/j.mrfmmm.2004.03.013.
14. Bristow RG, Hill RP. Hypoxia and metabolism: Hypoxia, DNA repair and genetic instability. Nature Reviews Cancer. 2008;8(3):180. doi: 10.1038/nrc2344.
15. Azvolinsky A, Giresi PG, Lieb JD, Zakian VA. Highly Transcribed RNA Polymerase II Genes Are Impediments to Replication Fork Progression in Saccharomyces cerevisiae. Molecular Cell. 2009;34(6):722-34. doi: 10.1016/j.molcel.2009.05.022.

16. Zeller P, Padeken J, van Schendel R, Kalck V, Tijsterman M, Gasser SM. Histone H3K9 methylation is dispensable for *Caenorhabditis elegans* development but suppresses RNA:DNA hybrid-associated repeat instability. Nature Genetics. 2016;48(11):1385-95. doi: 10.1038/ng.3672.
17. Chen X, Zhang J. Yeast mutation accumulation experiment supports elevated mutation rates at highly transcribed sites. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(39):E4062-E. doi: 10.1073/pnas.1412284111.
18. Park C, Qian W, Zhang J. Genomic evidence for elevated mutation rates in highly expressed genes. EMBO Reports. 2012;13(12):1123-9. doi: 10.1038/embor.2012.165.
19. Zhu YO, Siegal ML, Hall DW, Petrov DA. Precise estimates of mutation rate and spectrum in yeast. Proceedings of the National Academy of Sciences. 2014;111(22):E2310-E8. doi: 10.1073/pnas.1323011111.
20. Ganley AR, Kobayashi T. Highly efficient concerted evolution in the ribosomal DNA repeats: total rDNA repeat variation revealed by whole-genome shotgun sequence data. Genome Res. 2007;17(2):184-91. doi: 10.1101/gr.5457707.
21. Ganley AR, Kobayashi T. Monitoring the rate and dynamics of concerted evolution in the ribosomal DNA repeats of Saccharomyces cerevisiae using experimental evolution. Mol Biol Evol. 2011;28(10):2883-91. doi: 10.1093/molbev/msr117.
22. Gottlieb S, Esposito RE. A new role for a yeast transcriptional silencer gene, SIR2, in regulation of recombination in ribosomal DNA. Cell. 1989;56(5):771-6.
23. Kobayashi T, Ganley AR. Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. Science. 2005;309(5740):1581-4. doi: 10.1126/science.1116102.
24. Kobayashi T, Horiuchi T, Tongaonkar P, Vu L, Nomura M. SIR2 Regulates Recombination between Different rDNA Repeats, but Not Recombination within Individual rRNA Genes in Yeast. Cell. 2004;117(4):441-53. doi: 10.1016/S0092-8674(04)00414-3.
25. Kobayashi T, Heck DJ, Nomura M, Horiuchi T. Expansion and contraction of ribosomal DNA repeats in Saccharomyces cerevisiae: requirement of replication fork blocking (Fob1) protein and the role of RNA polymerase I. Genes & Development. 1998;12(24):3821-30.
26. Houseley J, Tollervey D. Repeat expansion in the budding yeast ribosomal DNA can occur independently of the canonical homologous recombination machinery. Nucleic Acids Res. 2011;39(20):8778-91. doi: 10.1093/nar/gkr589.
27. Kobayashi T. Regulation of ribosomal RNA gene copy number and its role in modulating genome integrity and evolutionary adaptability in yeast. Cellular and Molecular Life Sciences. 2011;68(8):1395-403. doi: 10.1007/s00018-010-0613-2.
28. Mekhail K, Moazed D. The nuclear envelope in genome organization, expression and stability. Nature Reviews Molecular Cell Biology. 2010;11(5):317-28. doi: 10.1038/nrm2894.
29. Mekhail K, Seebacher J, Gygi SP, Moazed D. Role for perinuclear chromosome tethering in maintenance of genome stability. Nature. 2008;456(7222):667-70. doi: 10.1038/nature07460.

30. King MC, Lusk CP, Blobel G. Karyopherin-mediated import of integral inner nuclear membrane proteins. Nature. 2006;442(7106):1003-7. doi: 10.1038/nature05075.
31. Ibáñez C, Pérez-Torrado R, Chiva R, Guillamón JM, Barrio E, Querol A. Comparative genomic analysis of Saccharomyces cerevisiae yeasts isolated from fermentations of traditional beverages unveils different adaptive strategies. International Journal of Food Microbiology. 2014;171:129-35. doi: 10.1016/j.ijfoodmicro.2013.10.023.
32. Haro R, Garciadeblas B, Rodríguez-Navarro A. A novel P-type ATPase from yeast involved in sodium transport. FEBS Letters. 1991;291(2):189-91.
33. Rodríguez-Navarro A, Quintero FJ, Garciadeblás B. Na+-ATPases and Na+/H+ antiporters in fungi. Biochimica et Biophysica Acta (BBA) - Bioenergetics. 1994;1187(2):203-5. doi: 10.1016/0005-2728(94)90111-2.
34. Strope PK, Skelly DA, Kozmin SG, Mahadevan G, Stone EA, Magwene PM, et al. The 100-genomes strains, an S. cerevisiae resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. Genome Research. 2015;25(5):762-74. doi: 10.1101/gr.185538.114.
35. Wieland J, Nitsche AM, Strayle J, Steiner H, Rudolph HK. The PMR2 gene cluster encodes functionally distinct isoforms of a putative Na+ pump in the yeast plasma membrane. EMBO Journal. 1995;14(16):3870-82.
36. Garciadeblás B, Rubio F, Quintero FJ, Bañuelos MA, Haro R, Rodríguez-Navarro A. Differential expression of two genes encoding isoforms of the ATPase involved in sodium efflux in Saccharomyces cerevisiae. Molecular and General Genetics MGG. 1993;236(2-3):363-8. doi: 10.1007/BF00277134.
37. Ruiz A, Ariño J. Function and regulation of the Saccharomyces cerevisiae ENA sodium ATPase system. Eukaryotic Cell. 2007;6(12):2175-83. doi: 10.1128/EC.00337-07.
38. Hastings PJ, Lupski JR, Rosenberg SM, Ira G. Mechanisms of change in gene copy number. Nature Reviews Genetics. 2009;10(8):551-64. doi: 10.1038/nrg2593.
39. Lang GI, Murray AW. Estimating the Per-Base-Pair Mutation Rate in the Yeast Saccharomyces cerevisiae. Genetics. 2008;178(1):67-82. doi: 10.1534/genetics.107.071506.
40. Lang GI, Murray AW. Mutation rates across budding yeast chromosome VI are correlated with replication timing. Genome Biology and Evolution. 2011;3:799-811. doi: 10.1093/gbe/evr054.
41. Gangloff S, Zou H, Rothstein R. Gene conversion plays the major role in controlling the stability of large tandem repeats in yeast. The EMBO Journal. 1996;15(7):1715-25.
42. Lobachev KS, Shor BM, Tran HT, Taylor W, Keen JD, Resnick MA, et al. Factors affecting inverted repeat stimulation of recombination and deletion in Saccharomyces cerevisiae. Genetics. 1998;148(4):1507-24.
43. Ronne H, Rothstein R. Mitotic sectored colonies: evidence of heteroduplex DNA formation during direct repeat recombination. Proceedings of the National Academy of Sciences. 1988;85(8):2696-700.
44. Verstrepen KJ, Jansen A, Lewitter F, Fink GR. Intragenic tandem repeats generate functional variability. Nature Genetics. 2005;37(9):986-90. doi: 10.1038/ng1618.

45. Green EM, Jiang Y, Joyner R, Weis K. A negative feedback loop at the nuclear periphery regulates GAL gene expression. Molecular Biology of the Cell. 2012;23(7):1367-75. doi: 10.1091/mbc.E11-06-0547.
46. Morrissey JP, Deardorff JA, Hebron C, Sachs AB. Decapping of stabilized, polyadenylated mRNA in yeast pab1 mutants. Yeast (Chichester, England). 1999;15(8):687-702.
47. Lynch M, Sung W, Morris K, Coffey N, Landry CR, Dopman EB, et al. A genome-wide view of the spectrum of spontaneous mutations in yeast. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(27):9272-7. doi: 10.1073/pnas.0803466105.
48. Hawk JD, Stefanovic L, Boyer JC, Petes TD, Farber RA. Variation in efficiency of DNA mismatch repair at different sites in the yeast genome. Proceedings of the National Academy of Sciences. 2005;102(24):8639-43. doi: 10.1073/pnas.0503415102.
49. Sheppard TK, Hitz BC, Engel SR, Song G, Balakrishnan R, Binkley G, et al. The Saccharomyces Genome Database Variant Viewer. Nucleic Acids Research. 2016;44(D1):D698-D702. doi: 10.1093/nar/gkv1250.
50. Song G, Dickins BJA, Demeter J, Engel S, Dunn B, Cherry JM. AGAPE (Automated Genome Analysis PipelinE) for Pan-Genome Analysis of Saccharomyces cerevisiae. PLoS One. 2015;10(3):e0120671. doi: 10.1371/journal.pone.0120671.
51. Brown CA, Murray AW, Verstrepen KJ. Rapid Expansion and Functional Divergence of Subtelomeric Gene Families in Yeasts. Current Biology: 2010;20(10):895-903. doi: 10.1016/j.cub.2010.04.027.
52. di Fagagna FdaA, Reaper PM, Clay-Farrace L, Fiegler H, Carr P, von Zglinicki T, et al. A DNA damage checkpoint response in telomere-initiated senescence. Nature. 2003;426(6963):194-8. doi: 10.1038/nature02118.
53. Mefford HC, Trask BJ. The complex structure and dynamic evolution of human subtelomeres. Nature Reviews Genetics. 2002;3(2):91-102. doi: 10.1038/nrg727.
54. Schacherer J, Shapiro JA, Ruderfer DM, Kruglyak L. Comprehensive polymorphism survey elucidates population structure of Saccharomyces cerevisiae. Nature. 2009;458(7236):342-5. doi: 10.1038/nature07670.
55. Collins SR, Miller KM, Maas NL, Roguev A, Fillingham J, Chu CS, et al. Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature. 2007;446(7137):806-10. doi: 10.1038/nature05649.
56. Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, et al. The genetic landscape of a cell. Science (New York, NY). 2010;327(5964):425-31. doi: 10.1126/science.1180823.
57. Grund SE, Fischer T, Cabal GG, Antúnez O, Pérez-Ortín JE, Hurt E. The inner nuclear membrane protein Src1 associates with subtelomeric genes and alters their regulated gene expression. The Journal of Cell Biology. 2008;182(5):897-910. doi: 10.1083/jcb.200803098.
58. Reid RJD, González-Barrera S, Sunjevaric I, Alvaro D, Ciccone S, Wagner M, et al. Selective ploidy ablation, a high-throughput plasmid transfer protocol, identifies new genes affecting topoisomerase I-induced DNA damage. Genome Research. 2011;21(3):477-86. doi: 10.1101/gr.109033.110.
59. Wilmes GM, Bergkessel M, Bandyopadhyay S, Shales M, Braberg H, Cagney G, et al. A genetic interaction map of RNA-processing factors reveals links between

Sem1/Dss1-containing complexes and mRNA export and splicing. Molecular Cell. 2008;32(5):735-46. doi: 10.1016/j.molcel.2008.11.012.
60. Yewdell WT, Colombi P, Makhnevych T, Lusk CP. Lumenal interactions in nuclear pore complex assembly and stability. Mol Biol Cell. 2011;22(8):1375-88. doi: 10.1091/mbc.E10-06-0554.
61. Santos-Pereira JM, Aguilera A. R loops: new modulators of genome dynamics and function. Nature Reviews Genetics. 2015;16(10):583-97. doi: 10.1038/nrg3961.
62. Cerritelli SM, Crouch RJ. Ribonuclease H: the enzymes in eukaryotes. FEBS J. 2009;276(6):1494-505. doi: 10.1111/j.1742-4658.2009.06908.x.
63. Boguslawski SJ, Smith DE, Michalak MA, Mickelson KE, Yehle CO, Patterson WL, et al. Characterization of monoclonal antibody to DNA · RNA and its application to immunodetection of hybrids. Journal of Immunological Methods. 1986;89(1):123-30. doi: 10.1016/0022-1759(86)90040-2.
64. Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, et al. Characteristic genome rearrangements in experimental evolution of Saccharomyces cerevisiae. Proc Natl Acad Sci U S A. 2002;99(25):16144-9. doi: 10.1073/pnas.242624799.
65. Gresham D, Usaite R, Germann SM, Lisby M, Botstein D, Regenberg B. Adaptation to diverse nitrogen-limited environments by deletion or extrachromosomal element formation of the GAP1 locus. Proc Natl Acad Sci U S A. 2010;107(43):18551-6. doi: 10.1073/pnas.1014023107.
66. Hastings PJ. Adaptive amplification. Crit Rev Biochem Mol Biol. 2007;42(4):271-83. doi: 10.1080/10409230701507757.
67. Barrales RR, Forn M, Georgescu PR, Sarkadi Z, Braun S. Control of heterochromatin localization and silencing by the nuclear membrane protein Lem2. Genes Dev. 2016;30(2):133-48. doi: 10.1101/gad.271288.115.
68. Black JC, Manning AL, Van Rechem C, Kim J, Ladd B, Cho J, et al. KDM4A lysine demethylase induces site-specific copy gain and rereplication of regions amplified in tumors. Cell. 2013;154(3):541-55. doi: 10.1016/j.cell.2013.06.051.
69. Furumizu C, Alvarez JP, Sakakibara K, Bowman JL. Antagonistic roles for KNOX1 and KNOX2 genes in patterning the land plant body plan following an ancient gene duplication. PLoS Genetics. 2015;11(2):e1004980. doi: 10.1371/journal.pgen.1004980.
70. Shum EY, Jones SH, Shao A, Dumdie J, Krause MD, Chan W-K, et al. The Antagonistic Gene Paralogs Upf3a and Upf3b Govern Nonsense-Mediated RNA Decay. Cell. 2016;165(2):382-95. doi: 10.1016/j.cell.2016.02.046.
71. Ellisdon AM, Dimitrova L, Hurt E, Stewart M. Structural basis for the assembly and nucleic acid binding of the TREX-2 transcription-export complex. Nature Structural & Molecular Biology. 2012;19(3):328-36. doi: 10.1038/nsmb.2235.
72. Raghuraman MK, Winzeler EA, Collingwood D, Hunt S, Wodicka L, Conway A, et al. Replication dynamics of the yeast genome. Science. 2001;294(5540):115-21. doi: 10.1126/science.294.5540.115.
73. Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol Cell. 2006;24(5):735-46. doi: 10.1016/j.molcel.2006.10.023.

74. Ito-Harashima S, Hartzog PE, Sinha H, McCusker JH. The tRNA-Tyr gene family of Saccharomyces cerevisiae: agents of phenotypic variation and position effects on mutation frequency. Genetics. 2002;161(4):1395-410.
75. Hicks WM, Kim M, Haber JE. Increased mutagenesis and unique mutation signature associated with mitotic gene conversion. Science (New York, NY). 2010;329(5987):82-5. doi: 10.1126/science.1191125.
76. Holbeck SL, Strathern JN. A Role for REV3 in Mutagenesis During Double-Strand Break Repair in Saccharomyces cerevisiae. Genetics. 1997;147(3):1017-24.
77. Rattray A, Strathern J. Homologous Recombination Is Promoted by Translesion Polymerase Polη. Molecular Cell. 2005;20(5):658-9. doi: 10.1016/j.molcel.2005.11.018.
78. Yang W, Woodgate R. What a difference a decade makes: insights into translesion DNA synthesis. Proceedings of the National Academy of Sciences. 2007;104(40):15591-8. doi: 10.1073/pnas.0704219104.
79. Cairns J, Foster PL. Adaptive reversion of a frameshift mutation in Escherichia coli. Genetics. 1991;128(4):695-701. Epub 1991/08/01.
80. Foster PL, Cairns J. Mechanisms of directed mutation. Genetics. 1992;131(4):783-9.
81. Albertson DG. Gene amplification in cancer. Trends Genet. 2006;22(8):447-55. doi: 10.1016/j.tig.2006.06.007.
82. Ercan D, Zejnullahu K, Yonesaka K, Xiao Y, Capelletti M, Rogers A, et al. Amplification of EGFR T790M causes resistance to an irreversible EGFR inhibitor. Oncogene. 2010;29(16):2346-56. doi: 10.1038/onc.2009.526.
83. Turner KM, Deshpande V, Beyter D, Koga T, Rusert J, Lee C, et al. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. Nature. 2017;543(7643):122-5. doi: 10.1038/nature21356.
84. Montes de Oca R, Shoemaker CJ, Gucek M, Cole RN, Wilson KL. Barrier-to-autointegration factor proteome reveals chromatin-regulatory partners. PLoS One. 2009;4(9):e7050. doi: 10.1371/journal.pone.0007050.
85. Amberg DC, Burke D, Strathern JN. Methods in yeast genetics. Cold Spring Harbor, NY: CSHL Press; 2005.
86. Rohner S, Gasser SM, Meister P. Modules for cloning-free chromatin tagging in Saccharomyces cerevisae. Yeast. 2008;25(3):235-9. doi: 10.1002/yea.1580.
87. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402-8. doi: 10.1006/meth.2001.1262.
88. Weaver S, Dube S, Mir A, Qin J, Sun G, Ramakrishnan R, et al. Taking qPCR to a higher level: Analysis of CNV reveals the power of high throughput qPCR to enhance quantitative resolution. Methods. 2010;50(4):271-6. doi: 10.1016/j.ymeth.2010.01.003.
89. Zhao Y, Schreiner SM, Koo PK, Colombi P, King MC, Mochrie SGJ. Improved Determination of Subnuclear Position Enabled by Three-Dimensional Membrane Reconstruction. Biophysical Journal. 2016;111(1):19-24. doi: 10.1016/j.bpj.2016.05.036.
90. Xu S, Grullon S, Ge K, Peng W. Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. Methods Mol Biol. 2014;1150:97-111. doi: 10.1007/978-1-4939-0512-6_5.

91. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9(7):676-82. doi: 10.1038/nmeth.2019.
92. El Hage A, Webb S, Kerr A, Tollervey D. Genome-wide distribution of RNA-DNA hybrids identifies RNase H targets in tRNA genes, retrotransposons and mitochondria. PLoS Genet. 2014;10(10):e1004716. doi: 10.1371/journal.pgen.1004716.

Figure Legends

Figure 1. *ENA* copy number expansion favors adaptation to high salt. (A) Schematic of the *ENA* genomic locus, including the *ENA* genes (green, numbered from 1 to 4), *RSM10* (upstream of the *ENA* locus, violet), and *KRS1* (downstream of the *ENA* locus, grey). Arrowheads indicate the direction of transcription. (B) The level of the *ENA1-4* transcript increases after 10 min. in the presence of 100 mM LiCl. RT-qPCR analysis of *ENA1-4* transcript in 100 mM LiCl normalized to 0 mM LiCl. Mean ± SD from 3 independent experiments. (C) Deletion of *ENA1-4* leads to cell death in the presence of 100 mM LiCl. Serial dilutions, 1:10, were grown on YPD plates with or without 100 mM LiCl. (D) Schematic of the *in vitro* evolution experiment. Cells were cultured for ≈200 generations in YPD containing 100 mM LiCl; every 24 h cells were diluted to 1,000 cells/ml in fresh media. (E) The adapted strain (WT-A) has a fitness advantage over the parental strain (WT-P) when grown on YPD plates containing 100 mM and 300 mM LiCl. Serial dilutions as in (C). Note that plates containing LiCl were imaged after growth for an additional 24 h. (F) The copy number of the *ENA* locus is increased in WT-A and its meiotic progeny (described in G-H) as measured by qPCR. Mean ± SD of three independent experiments. (G) Schematic of the predicted genotypic and phenotypic (relative growth in LiCl denoted by + and -) differences of progeny (F1) derived through Mendelian inheritance (homogenous staining region/HSR; expansion on the chromosome) and non-Mendelian double minutes (DM; generation of episomes) during meiotic segregation. The chromosome (in red) can contain the parental *ENA* genomic locus (in blue), or the adapted *ENA* locus (in green, HSR as single band or DM as circles). (H) The spores inheriting the adapted ENA locus (F1-1 and F1-4) retain

improved growth on media containing 300 mM LiCl. Serial dilutions as in (C).

Figure 2. *ENA* locus instability is biased toward the 3' region and is promoted by transcription. (A) Schematic representing the location of *URA3* insertions at the *ENA* locus. (B) The *ENA* genomic locus and its flanking regions have higher spontaneous instability compared to *URA3* at its endogenous locus. Measure of the loss of *URA3* activity as the rate of obtaining clones resistant to 5-FOA compared to total cells plated. Mean ± SD of three independent experiments. (C) Addition of LiCl to induce *ENA* expression (Fig. 1B) drives increased instability specifically downstream of the *ENA* locus. This effect is abrogated in cells harboring a mutation that ablates Pol II function (*rpb1-1*). Data is expressed as the ratio for LiCl-treated to untreated cells. Mean ± SD of three independent experiments. (D) Increasing *ENA* gene expression affects the number of mutational events within *URA3* specifically downstream of the *ENA* locus. In blue are events driving 5-FOA resistance outside the *URA3* gene. (E) Addition of LiCl to induce *ENA* expression alters the mutational signature specifically downstream of the *ENA* locus. (F) The SNP load is biased downstream of the *ENA* locus. SNP depth derived from 21 *S. cerevisiae* strains (black) processed to identify SNP peaks (gray), which were further analyzed to identify significant SNP islands (shaded in blue tones by score; see Methods).

Figure 3. LEM domain proteins control adaptation efficiency through *ENA* copy number variation. (A) Schematic of the topology and domain architecture of Heh1 and Heh2. The conserved domains LAP2-emerin-MAN1 (LEM, blue), MAN1 C-terminal

homology domain/winged helix (WH, green), and transmembrane domains (grey) are indicated. INM is inner nuclear membrane. (B) The adapted (A) strains have better fitness than the parental strains (P) growing on YPD plates containing 100 mM and 300 mM LiCl. The adapted strains were obtained after serial culturing (~ 200 generations) in YPD containing 100 mM LiCl. Serial dilutions as in Fig. 1C. (C) Deletion of *HEH1* and/or *HEH2* influences the rate of *ENA* copy number expansion. qPCR measures of the *ENA* copy number through the course of the *in vitro* evolution experiment. (D) Fitness comparison of the four genetic backgrounds at ~60 and ~90 generations. Serial dilutions as in Fig. 1C. (E) *ENA1-4* transcript levels increase in the adapted strains. Transcript levels of parental and adapted strains in control conditions (YPD) and after incubation for 10 min with 100 mM LiCl. The data are normalized to WT-P grown in YPD. Mean ± SD of three independent experiments. (F) *HEH1* influences the stability of the adapted *ENA* locus in the absence of selective pressure. qPCR analysis of *ENA* copy number of adapted strains before and after 4 serial re-streaking on media lacking LiCl.

Figure 4. *HEH1* and *HEH2* are part of a genetic pathway modulating R-loop formation. (A) Schematic of the tested genetic interactions (red and green lines) and those reported in the literature (red dashed lines). Red lines indicate a synthetic growth defect and green lines indicate a synthetic growth rescue for the connected gene pairs. (B) Synthetic genetic interactions between *SAC3* and *HEH1/2.* Serial dilutions, 1:10, cells were grown on YPD plates. Quantification of the colony size (expressed in cm^2), for 60 colonies for each strain from three independent experiments are plotted. Mean ±

SEM. **** $p < 0.0001$ (ANOVA). (C) Synthetic genetic interactions between *RNH1/201* and *HEH1/2.* Serial dilutions, 1:10, cells were grown on YPD and on YPD containing 50 mM HU. Measurements of the colony size (expressed in cm^2) for 60 colonies for each strain from three independent experiments are plotted. Mean ± SEM. **** $p < 0.0001$ (ANOVA). n.s. is not significant. (D) R-loop analysis by DRIP-qPCR across the *ENA* genomic locus in *WT*, *heh1*Δ*, heh2*Δ*, heh1*Δ*heh2*Δ strains. *CEN16*, *KRS1, ENA1-4,* and *RSM10* were analyzed by qPCR. The values for no-antibody (-Ab) and antibody S9.6 (+Ab) were calculated as described in Methods, and normalized to *CEN16* to compensate for differences in IP efficiency. Mean ± SD of three biological replicates, each analyzed by qPCR three times.

Figure 1

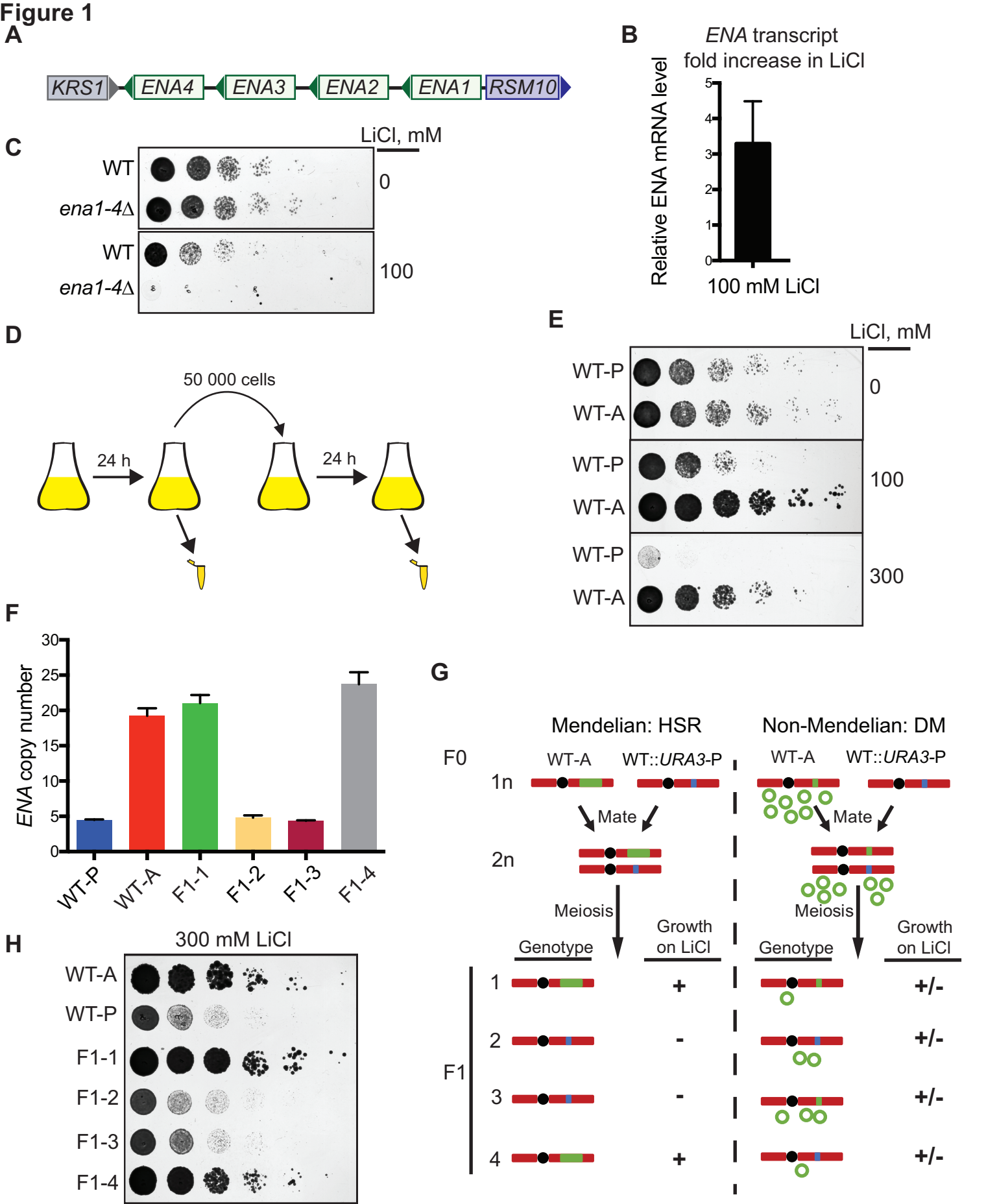

A
KRS1
ENA4
ENA3
ENA2
ENA1
RSM10
B
ENA transcript
fold increase in LiCl
Relative ENA mRNA level
0
1
2
3
4
5
100 mM LiCl
C
LiCl, mM
WT
ena1-4Δ
0
WT
ena1-4Δ
100
D
50 000 cells
24 h
24 h
E
LiCl, mM
WT-P
WT-A
0
WT-P
WT-A
100
WT-P
WT-A
300
F
ENA copy number
0
5
10
15
20
25
30
WT-P
WT-A
F1-1
F1-2
F1-3
F1-4
G
Mendelian: HSR
Non-Mendelian: DM
F0
WT-A
WT::URA3-P
WT-A
WT::URA3-P
1n
Mate
Mate
2n
Meiosis
Meiosis
Genotype
Growth on LiCl
Genotype
Growth on LiCl
F1
1
+
+/-
2
-
+/-
3
-
+/-
4
+
+/-
H
300 mM LiCl
WT-A
WT-P
F1-1
F1-2
F1-3
F1-4

Figure 2

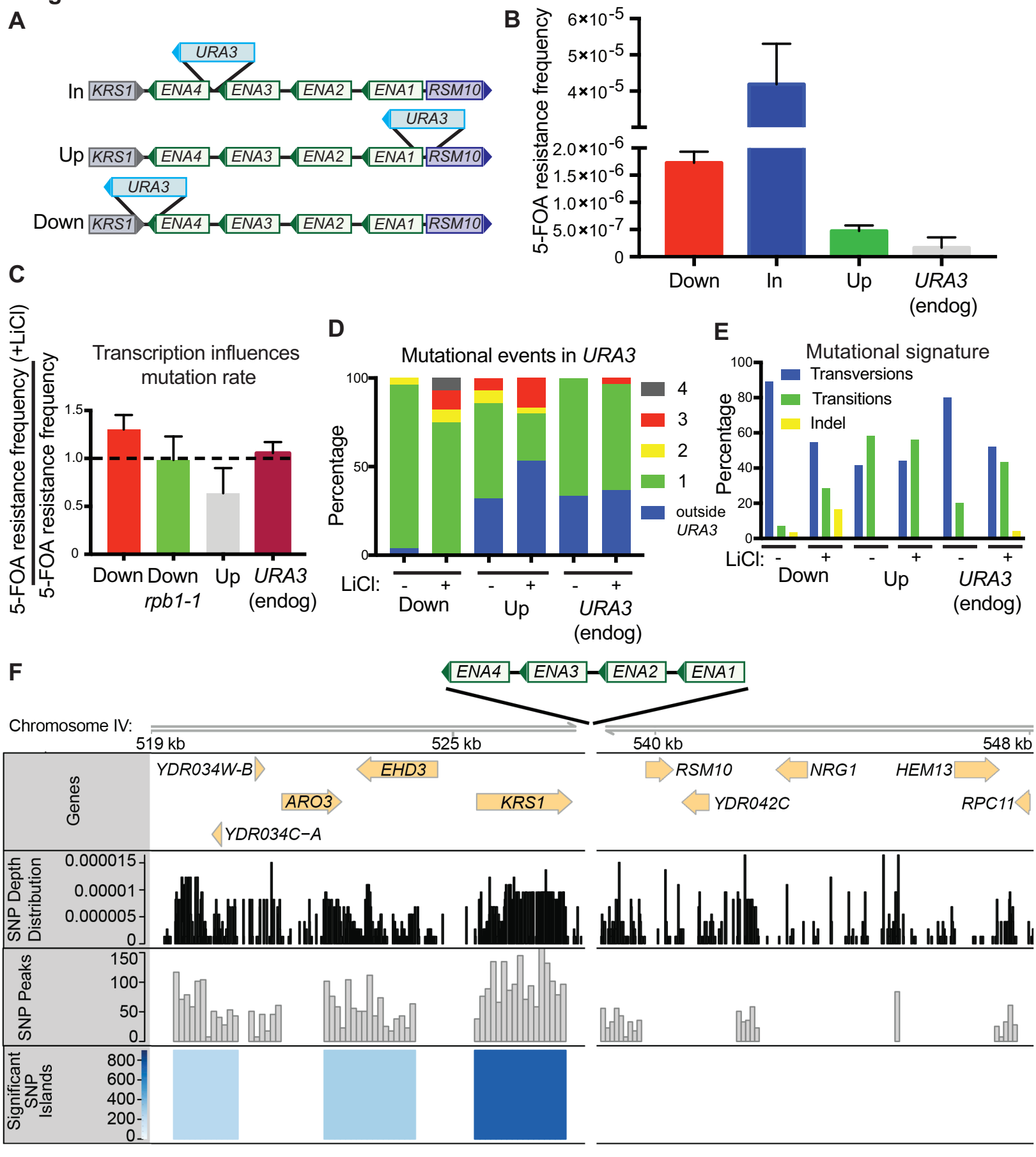

A
URA3
In KRS1 ENA4 ENA3 ENA2 ENA1 RSM10
URA3
Up KRS1 ENA4 ENA3 ENA2 ENA1 RSM10
URA3
Down KRS1 ENA4 ENA3 ENA2 ENA1 RSM10
B
5-FOA resistance frequency
6×10-5
5×10-5
4×10-5
2.0×10-6
1.5×10-6
1.0×10-6
5.0×10-7
0
Down
In
Up
URA3 (endog)
C
5-FOA resistance frequency (+LiCl)
5-FOA resistance frequency
Transcription influences mutation rate
1.5
1.0
0.5
0
Down
Down rpb1-1
Up
URA3 (endog)
D
Mutational events in URA3
Percentage
100
50
0
4
3
2
1
outside URA3
LiCl: - + - + - +
Down
Up
URA3 (endog)
E
Mutational signature
Percentage
100
80
60
40
20
0
Transversions
Transitions
Indel
LiCl: - + - + - +
Down
Up
URA3 (endog)
F
ENA4 ENA3 ENA2 ENA1
Chromosome IV:
519 kb
525 kb
540 kb
548 kb
Genes
YDR034W-B
EHD3
ARO3
KRS1
YDR034C-A
RSM10
NRG1
HEM13
YDR042C
RPC11
SNP Depth Distribution
0.000015
0.00001
0.000005
0
SNP Peaks
150
100
50
0
Significant SNP Islands
800
600
400
200
0

Figure 3

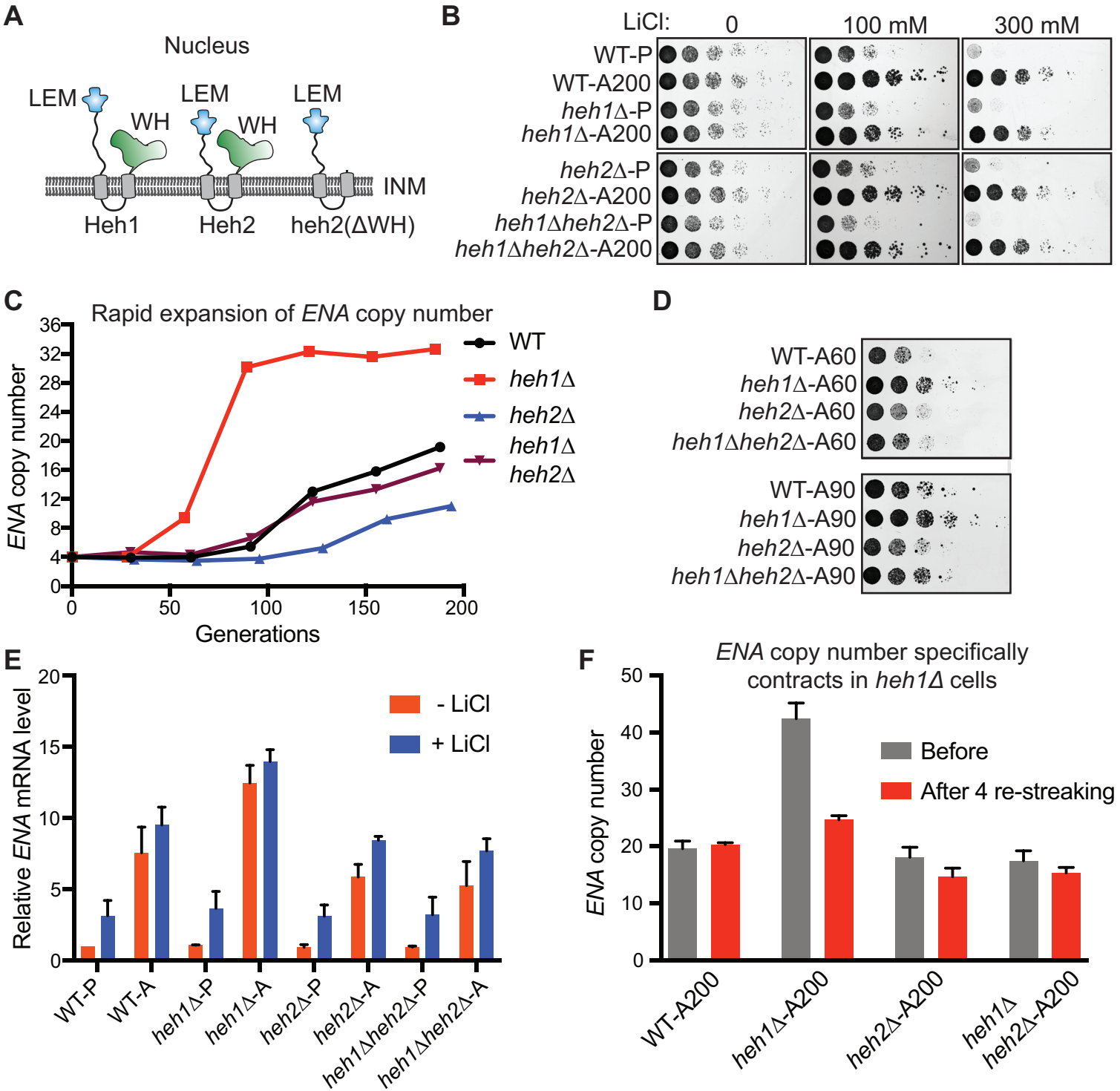
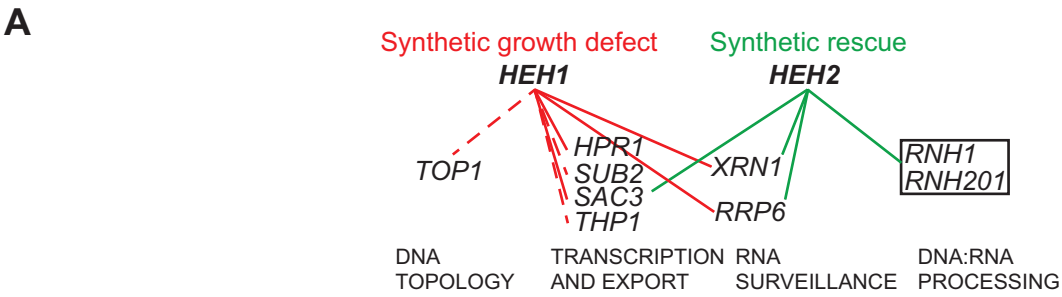
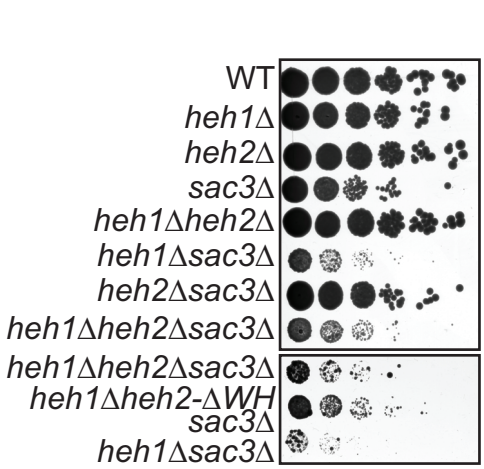
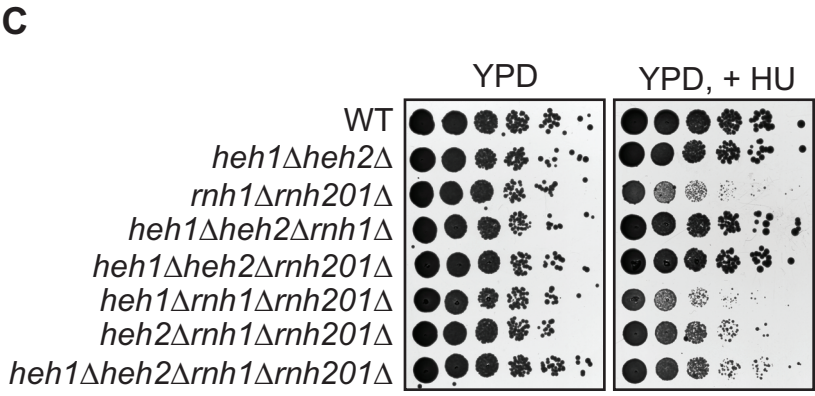
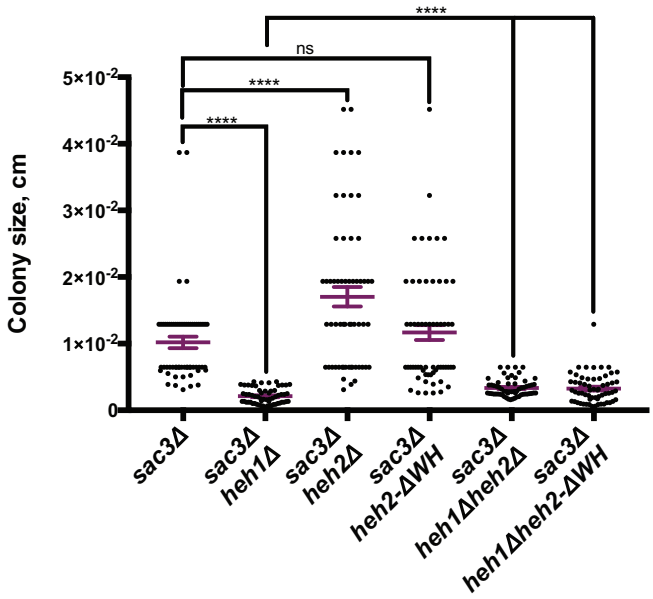

A
Nucleus
LEM
WH
LEM
WH
LEM
INM
Heh1
Heh2
heh2(ΔWH)
B
LiCl:
0
100 mM
300 mM
WT-P
WT-A200
heh1Δ-P
heh1Δ-A200
heh2Δ-P
heh2Δ-A200
heh1Δheh2Δ-P
heh1Δheh2Δ-A200
C
Rapid expansion of ENA copy number
ENA copy number
Generations
WT
heh1Δ
heh2Δ
heh1Δ heh2Δ
D
WT-A60
heh1Δ-A60
heh2Δ-A60
heh1Δheh2Δ-A60
WT-A90
heh1Δ-A90
heh2Δ-A90
heh1Δheh2Δ-A90
E
Relative ENA mRNA level
- LiCl
+ LiCl
WT-P
WT-A
heh1Δ-P
heh1Δ-A
heh2Δ-P
heh2Δ-A
heh1Δheh2Δ-P
heh1Δheh2Δ-A
F
ENA copy number specifically contracts in heh1Δ cells
ENA copy number
Before
After 4 re-streaking
WT-A200
heh1Δ-A200
heh2Δ-A200
heh1Δ heh2Δ-A200

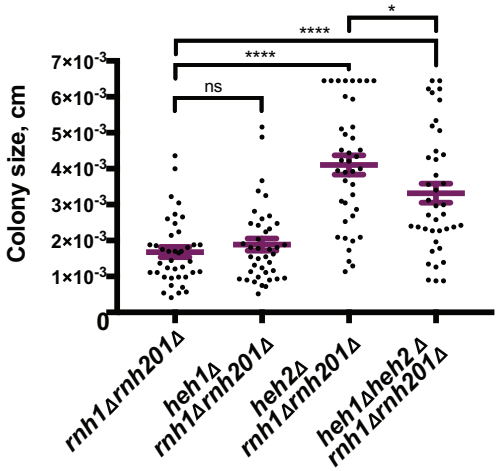
Figure 4

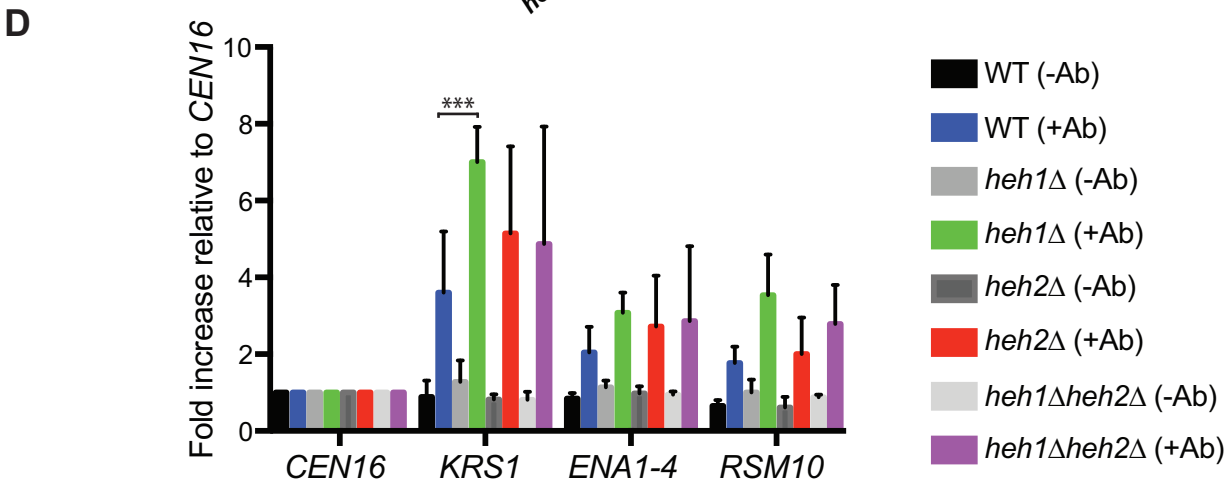

A
Synthetic growth defect
Synthetic rescue
HEH1
HEH2
TOP1
HPR1
SUB2
SAC3
THP1
XRN1
RRP6
RNH1
RNH201
DNA TOPOLOGY
TRANSCRIPTION AND EXPORT
RNA SURVEILLANCE
DNA:RNA PROCESSING


B
WT
heh1Δ
heh2Δ
sac3Δ
heh1Δheh2Δ
heh1Δsac3Δ
heh2Δsac3Δ
heh1Δheh2Δsac3Δ
heh1Δheh2Δsac3Δ
heh1Δheh2-ΔWH
sac3Δ
heh1Δsac3Δ


C
YPD
YPD, + HU
WT
heh1Δheh2Δ
rnh1Δrnh201Δ
heh1Δheh2Δrnh1Δ
heh1Δheh2Δrnh201Δ
heh1Δrnh1Δrnh201Δ
heh2Δrnh1Δrnh201Δ
heh1Δheh2Δrnh1Δrnh201Δ


Colony size, cm
5×10-2
4×10-2
3×10-2
2×10-2
1×10-2
0
****
ns
****
****
sac3Δ
sac3Δ heh1Δ
sac3Δ heh2Δ
sac3Δ heh2-ΔWH
sac3Δ heh1Δheh2Δ
sac3Δ heh1Δheh2-ΔWH


Colony size, cm
7×10-3
6×10-3
5×10-3
4×10-3
3×10-3
2×10-3
1×10-3
0
****
*
****
ns
rnh1Δrnh201Δ
heh1Δ rnh1Δrnh201Δ
heh2Δ rnh1Δrnh201Δ
heh1Δheh2Δ rnh1Δrnh201Δ


D
Fold increase relative to CEN16
10
8
6
4
2
0
***
CEN16
KRS1
ENA1-4
RSM10
WT (-Ab)
WT (+Ab)
heh1Δ (-Ab)
heh1Δ (+Ab)
heh2Δ (-Ab)
heh2Δ (+Ab)
heh1Δheh2Δ (-Ab)
heh1Δheh2Δ (+Ab)

Colombi et al., Supplemental Figures 1-6, Supplemental Tables 1-3

Supplemental Figure 1. Characterization of *ENA* genomic locus in W303. (A) PCR amplification of the *ENA* genomic locus from genomic DNA extracted from a strain containing a single *ENA* gene. The locus is schematized above the gel and the red arrows indicate the position of the PCR primers. (B) PCR amplification of the *ENA* genomic locus of our WT W303 strain finds four *ENA* gene copies (validated by qPCR, Fig. 1F).

Supplemental Figure 2. (A) The *URA3* gene inserted between the *ENA* repeats ("In") is lost in the clones that become resistant (FOA-R) to growth on 5-FOA. 1 to 5 represents five different clones analyzed. The PCR in the lower panel shows the presence of the *ENA* genes even in clones that lost *URA3*. On the right are the schematics of the PCR primer binding sites (red) at the starting locus prior to 5-FOA selection. (B) The genomic region 3' to the *ENA* locus has a very high SNP load as assessed by comparative genomics (red, see also Table 1).

Supplemental Figure 3. The *ENA* locus is associated with the nuclear periphery. (A) Fluorescent micrographs of one *z* section of cells expressing Hmg1-mCherry and LacI-GFP. The Lac operator (LacO) array was inserted in proximity to the *ENA* locus. Green and red channels are shown, in addition to the merge. Dotted lines denote cell boundaries. (B) Distribution of the normalized distance between the LacO array and the nuclear envelope (NE) derived from the 3D reconstructions. Here the *ENA* locus is

enriched at the NE and it is depleted the nuclear interior, while the endogenous *URA3* locus is relatively depleted from the NE and enriched at the nuclear interior. (C) Cumulative distribution of the normalized distance of the LacO array from the NE. The dashed lines represent the raw data while the continuous lines represent the binned data.

Supplemental Figure 4. The *ENA* locus undergoes intra-chromosomal expansion in *heh1*Δ, *heh2*Δ *and heh1*Δ*heh2*Δ cells. (A, C, and E) The adapted *ENA* locus segregates in a Mendelian fashion in the adapted (A) *heh1*Δ, *heh2*Δ and *heh1*Δ*heh2*Δ strains, respectively. The *ENA* gene copy number of the four spores (F1-1-F1-4) was assessed by qPCR. (B, D, F) The spores from (A, C and E) that inherit the adapted *ENA* locus also show improved fitness on media containing 300 mM LiCl over those inheriting the WT *ENA* locus. Serial dilutions as in Fig. 1C.

Supplemental Figure 5. Independent biological replicates of *ENA* expansion and contraction. (A) qPCR to determine *ENA* copy number through the *in vitro* evolution experiment of the indicated strains. (B) Second independent experiment showing that *HEH1* influences the stability of the adapted *ENA* locus in the absence of selective pressure. qPCR analysis of *ENA* copy number of adapted strains before and after 4 serial re-streakings on complete media.

Supplemental Figure 6. Synthetic genetic interactions between *XRN1* (top), *RRP6* (bottom) and *HEH1/2.* Serial dilutions as in Fig. 1C. Measurement of the colony size

(expressed in cm^2); 60 colonies from each strain from three independent experiments are plotted. Mean ± SEM. *** $p < 0.001$, ** $p < 0.01$ (ANOVA).

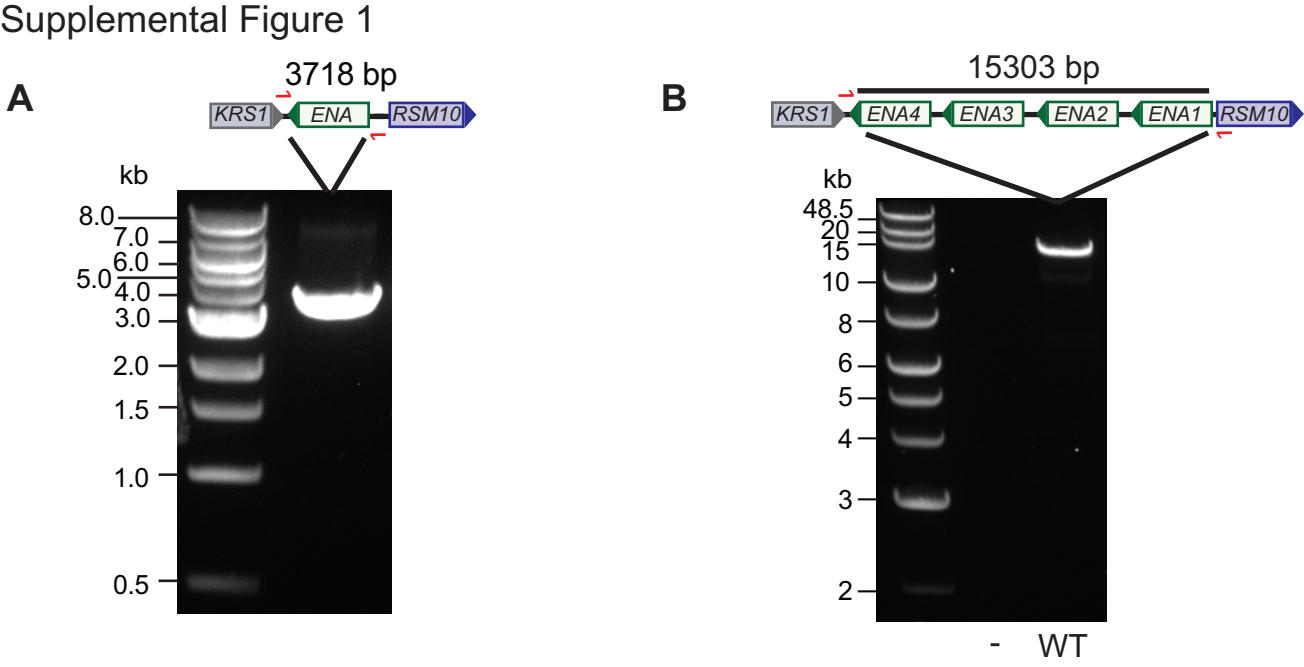
Supplemental Tables

Table 1. Top SNP islands (see Methods). Coordinates for each island, as well as gene features, whether the region is in the subtelomere (defined as within 25 kb of the chromosome end) and whether the region contains a gene with a paralogue are indicated. The *KRS1/ENA* region, ranked 21st, is highlighted.

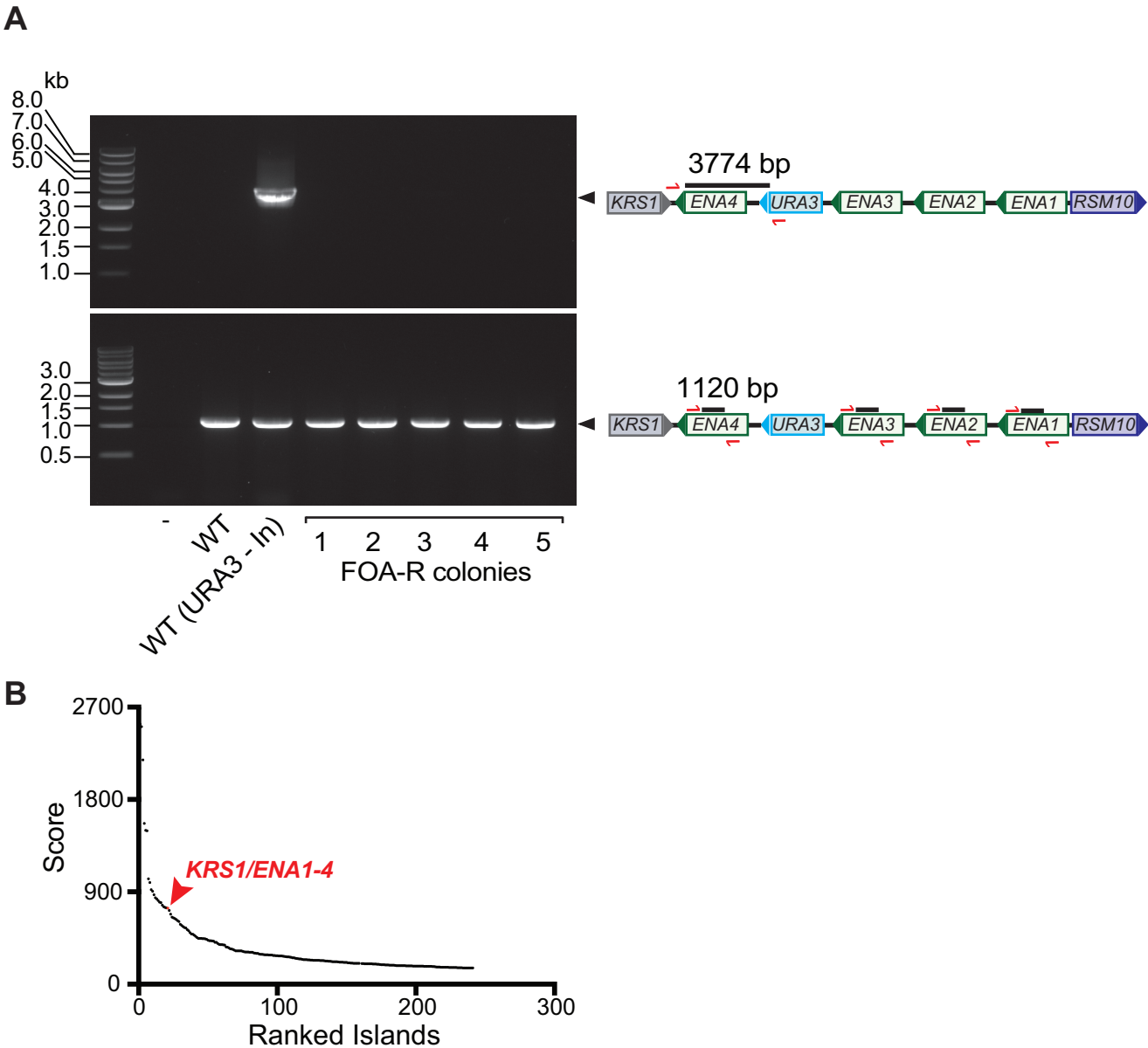
Table 2. *S. cerevisiae* strains used in this study.

Table 3. Oligonucleotide primers used in this study.

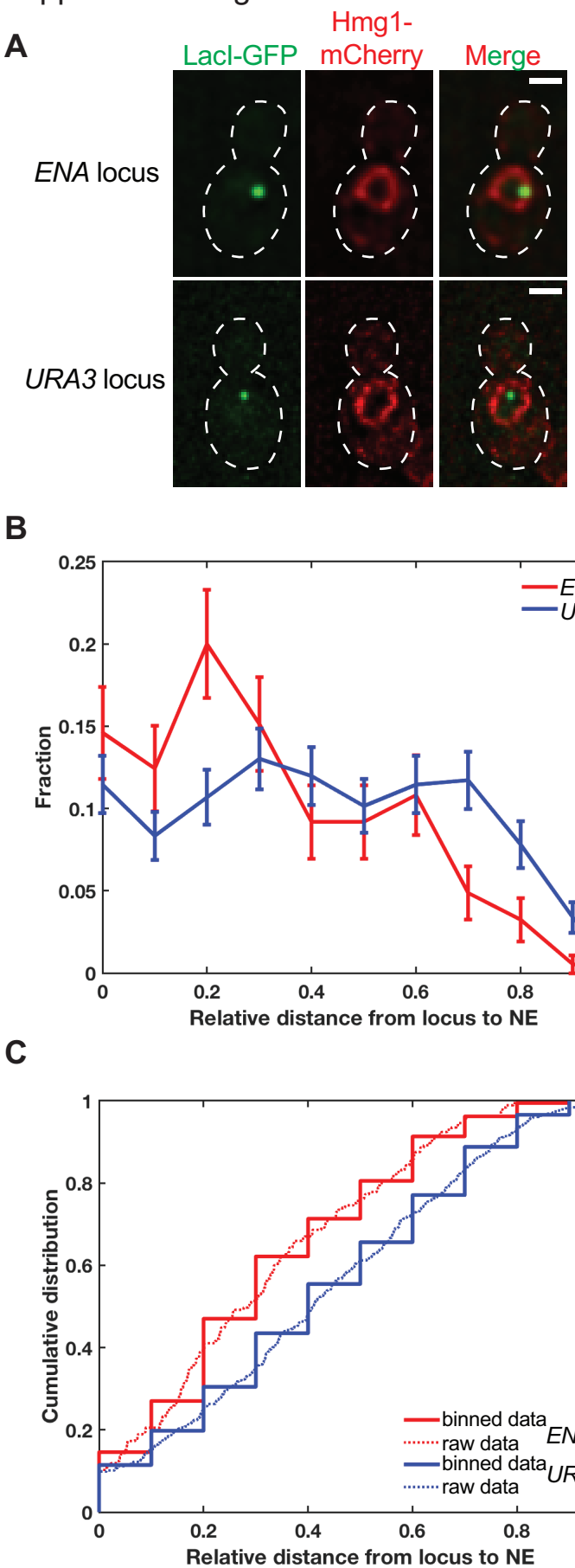
Supplemental Figure 1



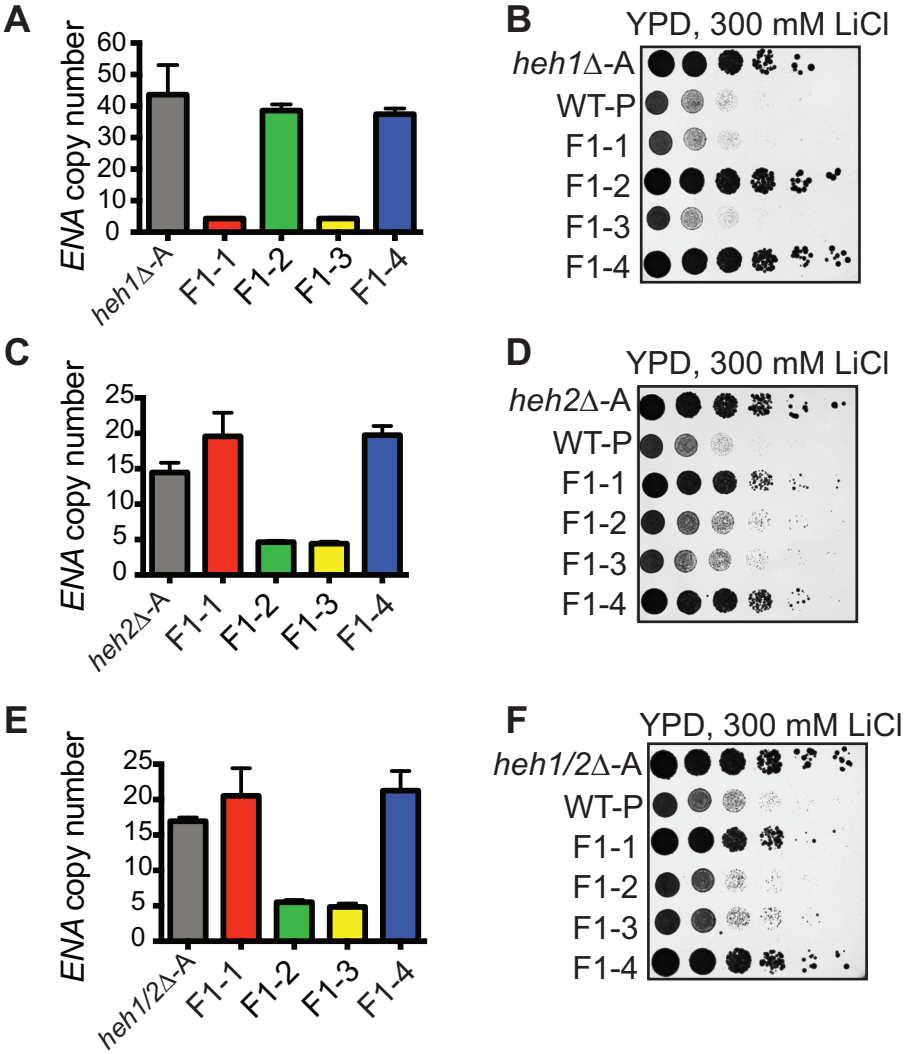
Supplemental Figure 2



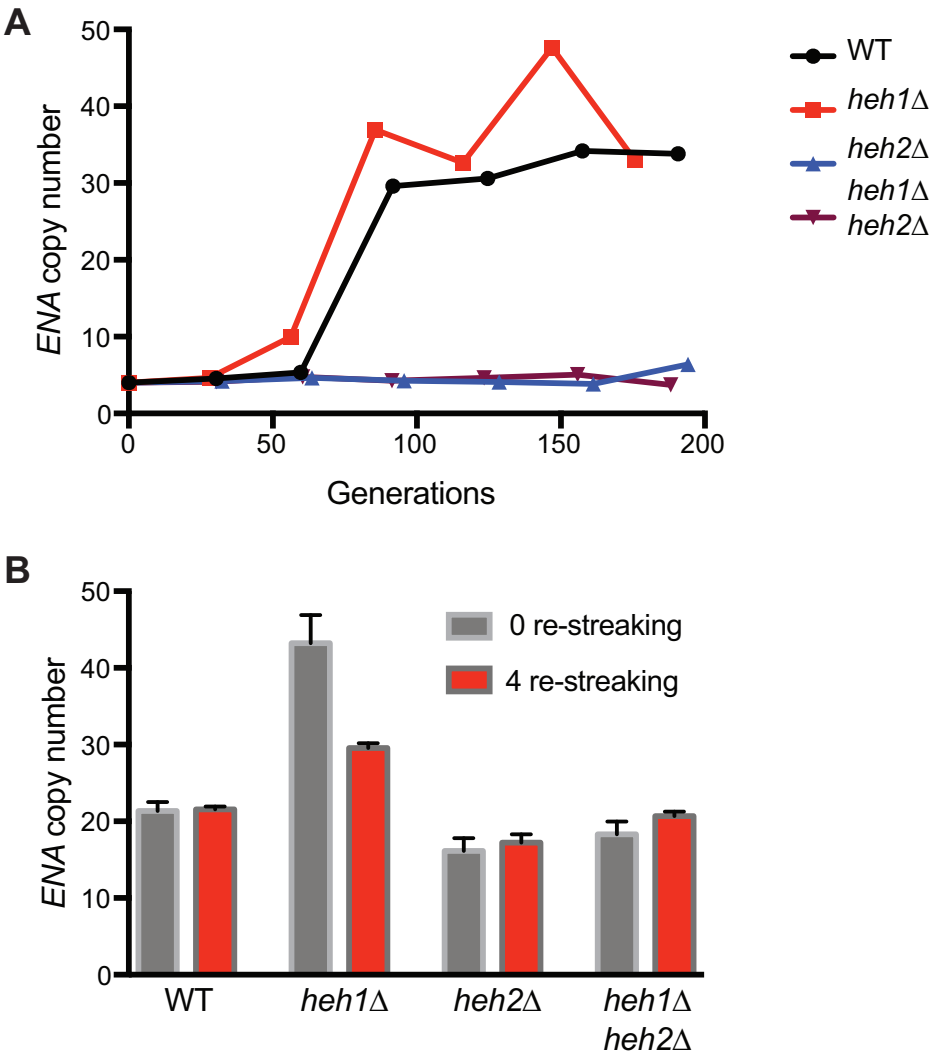
Supplemental Figure 3



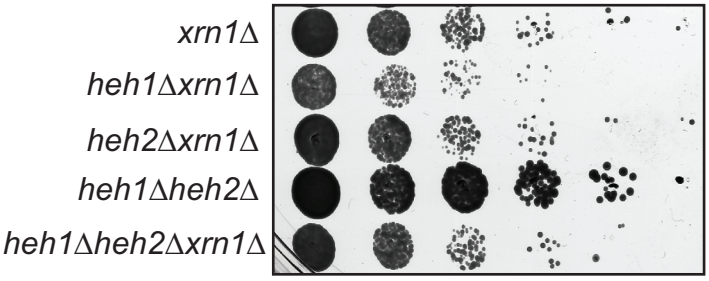
Supplemental Figure 4

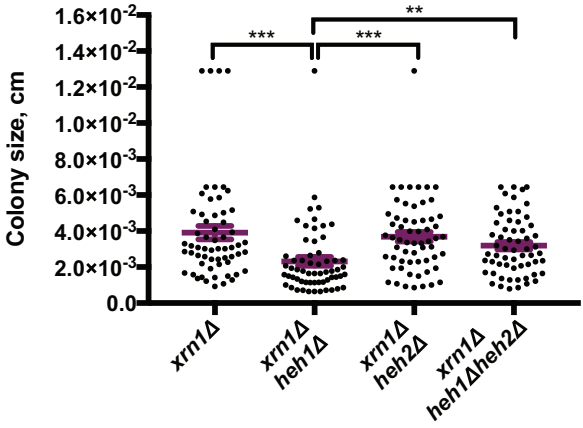


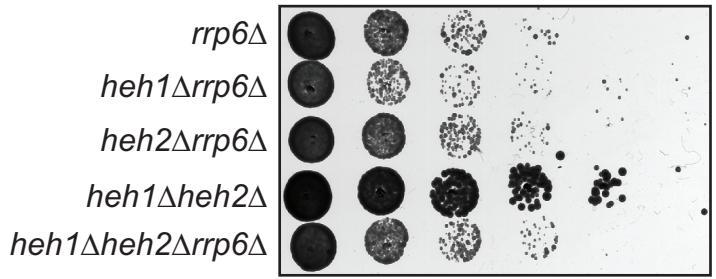
Supplemental Figure 5



Supplemental Figure 6







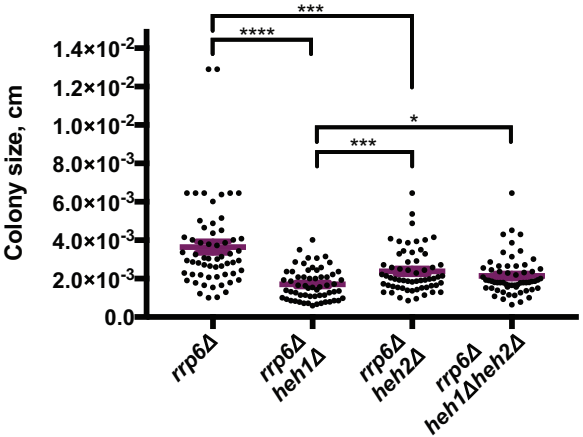


Table 1							
Island 1 (I:26800-32299)	Score=2535.74205583						
	FLO9	I:24000-27968	-	INTERSECTS on 5' side	Subtelomere	Paralog=FLO1(203403,208016,within_species_paralog)	
	GDH3	I:31567-32940	+	INTERSECTS on 5' side	Paralog=GDH1(1041678,1043042,within_species_paralog)		
Island 2 (VIII:92100-94499)	Score=2510.93794873						
	YHL008C	VIII:92627-94510	-	INTERSECTS on 3' side			
Island 3 (I:195900-199099)	Score=2185.10497897						
	SWH1	I:192619-196185	+	INTERSECTS on 3' side	Paralog=OSH2(417663,421514,within_species_paralog)		
Island 4 (VII:8800-13499)	Score=1568.13180526						
	YPS5	VII:8470-8967	+	INTERSECTS on 3' side	Subtelomere		
	YGL258W-A	VII:9162-9395	+	CONTAINED	Subtelomere		
	VEL1	VII:11110-11730	+	CONTAINED	Subtelomere	Paralog=YOR387C(1069621,1070241,within_species_paralog)	
	MNT2	VII:12481-14157	-	INTERSECTS on 3' side	Subtelomere	Paralog=MNT4(736803,738545,within_species_paralog)	
Island 5 (X:23900-26699)	Score=1500.59105166						
	REE1	X:23133-23729	+	BORDERS on its 3' side	Subtelomere		
	IMA5	X:24341-26086	-	CONTAINED	Subtelomere	Paralog=IMA2(22525,24294,within_species_paralog)	
	YJL215C	X:26412-26771	-	INTERSECTS on 3' side			
	HXT8	X:26887-28596	+	BORDERS on 5' side	Paralog=HXT7(1154216,1155928,within_species_paralog)		
Island 6 (VII:15300-20299)	Score=1494.78069198						
	ADH4	VII:15159-16307	+	INTERSECTS on 3' side	Subtelomere		
Island 7 (VI:191100-194599)	Score=1027.4792738						
	FAB1	VI:184502-191338	+	INTERSECTS on 3' side			
	YFR020W	VI:192737-193435	+	CONTAINED			
	ATG18	VI:194812-196314	+	BORDERS on 5' side			
Island 8 (XIV:735700-739299)	Score=990.60266807						
	BIO3	XIV:734291-735733	+	INTERSECTS on 3' side			
	MNT4	XIV:736803-738545	+	CONTAINED	Paralog=MNT2(12481,14157,within_species_paralog)		
Island 9 (I:202300-204299)	Score=921.303304748						
	FLO1	I:203403-208016	+	INTERSECTS on 5' side	Subtelomere	Paralog=FLO9(24000,27968,within_species_paralog)	
Island 10 (IX:424200-429699)	Score=905.210691146						
	GTT1	IX:423809-424513	-	INTERSECTS on 5' side	Subtelomere		
Island 11 (XVI:19400-24199)	Score=871.83879674						
	SAM3	XVI:22938-24701	+	INTERSECTS on 5' side	Subtelomere	Paralog=MMP1(17956,19707,within_species_paralog)	
Island 12 (IX:392600-397699)	Score=842.282242453						
	FLO11	IX:389572-393675	-	INTERSECTS on 5' side	Paralog=BSC1(384601,385587,within_species_paralog)		
	YIR020C	IX:394255-394557	-	CONTAINED			
	YIR020W-A	IX:394917-395159	+	CONTAINED			
	YIR020C-B	IX:397215-397949	-	INTERSECTS on 3' side			
	MRS1	IX:397294-398385	+	INTERSECTS on 5' side	Paralog=CCE1(420155,421216,within_species_paralog)		
Island 13 (XV:42800-44599)	Score=832.127565367						
	FRE7	XV:40748-42610	+	BORDERS on its 3' side	Paralog=FRE5(1061564,1063648,within_species_paralog)		
	GRE2	XV:43694-44722	+	INTERSECTS on 5' side	Paralog=YDR541C(1519664,1520698,within_species_paralog)		
	YOL150C	XV:44473-44784	-	INTERSECTS on 3' side			
	DCP1	XV:44938-45633	+	BORDERS on 5' side			
Island 14 (VI:17600-21599)	Score=815.921605548						
	AGP3	VI:17004-18680	+	INTERSECTS on 3' side	Subtelomere	Paralog=ALP1(135940,137661,within_species_paralog)	
	YFL054C	VI:20847-22787	-	INTERSECTS on 3' side	Subtelomere		
Island 15 (V:560600-566799)	Score=799.622588886						
	PUG1	V:559454-560365	+	BORDERS on its 3' side	Subtelomere	Paralog=RTA1(918512,919465,within_species_paralog)	
	YER186C	V:561705-562625	-	CONTAINED	Subtelomere		
	YER187W	V:566230-566655	+	CONTAINED	Subtelomere		
Island 16 (XIII:907400-908699)	Score=796.390381503						
	YMR316C-B	XIII:907321-907629	-	INTERSECTS on 5' side	Subtelomere		
	YMR317W	XIII:907364-910786	+	SURROUNDS	Subtelomere		
Island 17 (XIV:742100-744499)	Score=770.538949961						
	FRE4	XIV:739951-742110	+	INTERSECTS on 3' side	Paralog=FRE2(9091,11226,within_species_paralog)		
	YNR061C	XIV:742881-743540	-	CONTAINED			
	YNR062C	XIV:744360-745343	-	INTERSECTS on 3' side			
Island 18 (XV:29600-32399)	Score=754.549508367						
	HPF1	XV:28703-31606	-	INTERSECTS on 5' side	Paralog=YIL169C(23119,26106,within_species_paralog)		
Island 19 (IX:26200-28499)	Score=747.593492399						
	YIL169C	IX:23119-26106	-	BORDERS on its 5' side	Subtelomere	Paralog=HPF1(28703,31606,within_species_paralog)	
Island 20 (IX:385400-390299)	Score=744.156960663						
	YAP5	IX:384609-385346	+	BORDERS on its 3' side	Paralog=YAP7(270633,271370,within_species_paralog)		
	YIR018C-A	IX:385564-385701	-	CONTAINED			
	FLO11	IX:389572-393675	-	INTERSECTS on 3' side	Paralog=BSC1(384601,385587,within_species_paralog)		
Island 21 (IV:525400-527099)	Score=741.861880103						
	KRS1	IV:525440-527215	+	INTERSECTS on 5' side			
	ENA5	IV:527422-530697	-	BORDERS on 3' side			
Island 22 (VII:20500-22399)	Score=718.621667044						
	ZRT1	VII:20978-22108	+	CONTAINED	Subtelomere	Paralog=ZRT2(402794,404062,within_species_paralog)	
	FZF1	VII:22304-23203	+	INTERSECTS on 5' side	Subtelomere		
Island 23 (XII:788300-789399)	Score=683.448164996						
	CHS5	XII:787664-789679	+	SURROUNDS			
Island 24 (X:26900-31199)	Score=652.656962559						
	YJL215C	X:26412-26771	-	BORDERS on its 5' side			
	HXT8	X:26887-28596	+	INTERSECTS on 3' side	Paralog=HXT7(1154216,1155928,within_species_paralog)		
Island 25 (XV:384200-386599)	Score=647.244160489						
	CIN5	XV:383533-384420	-	INTERSECTS on 5' side	Paralog=YAP6(974631,975782,within_species_paralog)		
	YOR029W	XV:384600-384935	+	CONTAINED			
	DFG16	XV:386825-388684	+	BORDERS on 5' side			
Island 26 (XIII:21000-21599)	Score=639.536122198						
	ERG13	XIII:19060-20535	-	BORDERS on its 5' side	Subtelomere		
	PGA3	XIII:20761-21699	-	SURROUNDS	Subtelomere	Paralog=CBR1(274072,274926,within_species_paralog)	

	TUB3	XIII:22048-23683	-	BORDERS on 3' side	Subtelomere		
Island 27 (XIV:702900-704599)	Score=632.245510189						
	MVD1	XIV:701895-703085	+	INTERSECTS on 3' side			
	AGA1	XIV:703699-705876	+	INTERSECTS on 5' side			
Island 28 (VIII:149300-152399)	Score=620.158130268						
	ECM12	VIII:149225-149680	+	INTERSECTS on 3' side			
	YHR022C	VIII:149575-150345	-	CONTAINED			
	YHR022C-A	VIII:151217-151306	-	CONTAINED			
	MYO1	VIII:151666-157452	+	INTERSECTS on 5' side			
Island 29 (XIV:750000-751399)	Score=611.306010611						
	YNR064C	XIV:749136-750008	-	INTERSECTS on 5' side			
	YNR065C	XIV:750350-753700	-	INTERSECTS on 3' side	Paralog=VTH2(11475,16124,within_species_paralog)		
Island 30 (XIV:746200-748299)	Score=584.351320922						
	YNR063W	XIV:746943-748766	+	INTERSECTS on 5' side			
Island 31 (I:100-1999)	Score=572.982241978						
	YAL069W	I:335-649	+	CONTAINED	Subtelomere		
	YAL068W-A	I:538-792	+	CONTAINED	Subtelomere	Paralog=YNL337W(7165,7419,within_species_paralog)	
	PAU8	I:1807-2169	-	INTERSECTS on 3' side	Subtelomere	Paralog=PAU11(6290,6652,within_species_paralog)	
	YAL067W-A	I:2480-2707	+	BORDERS on 5' side	Subtelomere	Paralog=YJL222W-A(9452,9679,within_species_paralog)	
Island 32 (IX:25000-25999)	Score=560.959682269						
	YIL169C	IX:23119-26106	-	SURROUNDS	Subtelomere	Paralog=HPF1(28703,31606,within_species_paralog)	
Island 33 (VIII:12300-15299)	Score=551.797057015						
	PAU13	VIII:11923-12285	-	BORDERS on its 5' side	Subtelomere	Paralog=PAU11(6290,6652,within_species_paralog)	
	YHL045W	VIII:12502-12849	+	CONTAINED	Subtelomere	Paralog=YMR324C(922202,922444,within_species_paralog)	
	YHL044W	VIII:13565-14272	+	CONTAINED	Subtelomere		
	ECM34	VIII:14901-15413	+	INTERSECTS on 5' side	Subtelomere	Paralog=COS12(2790,3932,within_species_paralog)	
	YHL042W	VIII:15667-16119	+	BORDERS on 5' side	Subtelomere	Paralog=COS8(6401,7546,within_species_paralog)	
Island 34 (VIII:525000-526999)	Score=546.968367354						
	FLO5	VIII:525392-528619	+	INTERSECTS on 5' side	Paralog=FLO9(24000,27968,within_species_paralog)		
Island 35 (IV:17800-18899)	Score=533.152197138						
	AAD4	IV:17577-18566	-	INTERSECTS on 5' side	Subtelomere	Paralog=AAD14(16118,17248,within_species_paralog)	
	YDL242W	IV:18959-19312	+	BORDERS on 5' side	Subtelomere		
Island 36 (VII:1081900-108429	Score=520.957923784						
	COS6	VII:1081584-1082729	-	INTERSECTS on 5' side	Subtelomere	Paralog=COS1(8330,9475,within_species_paralog)	
Island 37 (IX:37600-38699)	Score=503.720362376						
	YIL163C	IX:36899-37252	-	BORDERS on its 5' side			
	SUC2	IX:37385-38983	+	SURROUNDS			
Island 38 (XII:390600-393699)	Score=489.881611705						
	YPS3	XII:388744-390270	-	BORDERS on its 5' side	Paralog=YPS1(386511,388220,within_species_paralog)		
	YLR122C	XII:390954-391331	-	CONTAINED			
	YLR123C	XII:391078-391407	-	CONTAINED			
	YLR124W	XII:391600-391944	+	CONTAINED			
	YLR125W	XII:393484-393894	+	INTERSECTS on 5' side			
Island 39 (VI:29400-33399)	Score=487.16189157						
	YFL052W	VI:28232-29629	+	INTERSECTS on 3' side	Paralog=MAL33(800523,801929,within_species_paralog)		
	YFL051C	VI:30058-30540	-	CONTAINED	Paralog=YHR212W-A(538742,538945,within_species_paralog)		
	ALR2	VI:33272-35848	-	INTERSECTS on 3' side	Paralog=ALR1(74400,76979,within_species_paralog)		
Island 40 (XIII:370700-372599)	Score=477.837677818						
	ERB1	XIII:368094-370517	-	BORDERS on its 5' side			
	YMR050C	XIII:373057-378325	-	BORDERS on 3' side	Paralog=YOR142W-B(595112,600380,within_species_paralog)		
Island 41 (IX:32400-34099)	Score=467.19233059						
	YIL166C	IX:30938-32566	-	INTERSECTS on 5' side	Paralog=YOL163W(9597,10106,within_species_paralog)		
	YIL165C	IX:33718-34077	-	CONTAINED			
	NIT1	IX:34087-34686	-	INTERSECTS on 3' side			
Island 42 (IX:437800-439799)	Score=456.018774764						
Island 43 (II:802600-804799)	Score=447.741853573						
	MAL31	II:802631-804475	-	CONTAINED	Subtelomere	Paralog=MPH3(738008,739816,within_species_paralog)	
	YBR298C-A	II:805035-805256	-	BORDERS on 3' side	Subtelomere		
Island 44 (VII:986300-988599)	Score=447.558683868						
	SOL4	VII:985972-986739	+	INTERSECTS on 3' side	Paralog=SOL3(423724,424473,within_species_paralog)		
	MGA1	VII:988049-989419	+	INTERSECTS on 5' side	Paralog=SFL1(586981,589281,within_species_paralog)		
Island 45 (XII:1063400-106499	Score=446.638385731						
	PAU4	XII:1062919-1063281	+	BORDERS on its 3' side	Subtelomere	Paralog=PAU8(1807,2169,within_species_paralog)	
Island 46 (IX:246500-247899)	Score=445.05496791						
	YIL060W	IX:246392-246826	+	INTERSECTS on 3' side	Paralog=YOR192C-C(703988,704224,within_species_paralog)		
	YIL059C	IX:246550-246915	-	CONTAINED			
	YIL058W	IX:246914-247198	+	CONTAINED			
	RGI2	IX:247902-248396	-	BORDERS on 3' side	Paralog=RGI1(292066,292551,within_species_paralog)		
Island 47 (X:714600-715599)	Score=444.765619221						
	DAN4	X:712255-715740	-	SURROUNDS	Paralog=PAU9(7605,7733,within_species_paralog)		
Island 48 (VII:323400-325399)	Score=441.603181875						
	SRM1	VII:321782-323230	+	BORDERS on its 3' side			
	TOS8	VII:325331-326161	+	INTERSECTS on 5' side	Paralog=CUP9(213042,213962,within_species_paralog)		
Island 49 (II:190000-191799)	Score=437.778102701						
	PEP1	II:186844-191583	-	INTERSECTS on 5' side	Paralog=VTH2(11475,16124,within_species_paralog)		
Island 50 (XI:310300-312799)	Score=428.888525288						
	NUP100	XI:310199-313078	+	SURROUNDS	Paralog=NUP116(363364,366705,within_species_paralog)		

Table 2			
Name	**Genotype**	**Location**	**Source**
W303	*ade2-1 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1*		Euroscarf
PCCPL565	W303, *xrn1::HYG*		this work
PCCPL566	W303, *heh1::TRP1 xrn1::HYG*		this work
PCCPL567	W303, *heh2::natMX6 xrn1::HYG*		this work
PCCPL568	W303, *heh1::TRP1 heh2::natMX6 xrn1::HYG*		this work
PCCPL569	W303, *rrp6::HYG*		this work
PCCPL570	W303, *heh1::TRP1 rrp6::HYG*		this work
PCCPL571	W303, *heh2::natMX6 rrp6::HYG*		this work
PCCPL572	W303, *heh1::TRP1 heh2::natMX6 rrp6::HYG*		this work
PCCPL625	W303, *ena1/2/3/4Δ::kanMX6*		this work
PCCPL647	W303, *heh1Δ::natMX6*		this work
PCCPL648	W303, *heh2Δ::kanMX6*		this work
PCCPL649	W303, *heh1Δ::natMX6 heh2Δ::kanMX6*		this work
PCCPL654	W303, *LacI-GFP::HIS3 ENA1-LacO::TRP1 HMG1-mCherry::kanMX*	LacO inserted on ChrIV: 539221	this work
PCCPL666	W303, *sac3::HYG*		this work
PCCPL668	W303, *rnh1::HYG*		this work
PCCPL686	W303, *heh1Δ::natMX6* sac3::HYG		this work
PCCPL687	W303, *heh2Δ::kanMX6 sac3::HYG*		this work
PCCPL689	W303, *heh1Δ::natMX6 heh2Δ::kanMX6 sac3::HYG*		this work
PCCPL732	W303, *URA3*	*URA3* inserted ChrIV: 527281	this work
PCCPL733	W303, *URA3*	*URA3* inserted ChrIV: 531025	this work
PCCPL741	W303, *heh2::(1-570) HA::HIS sac3::HYG*		this work
PCCPL742	W303, *heh1Δ::natMX6 heh2::(1-570) HA::HIS sac3::HYG*		this work
PCCPL774	W303, *URA3*	*URA3* inserted ChrIV: 539591	this work
PCCPL775	W303, *URA3*	*URA3* at the endogenous locus	this work
PCCPL809	W303, *rnh201::HIS3*		this work
PCCPL814	W303, *URA3 rpb-1*	*URA3* inserted ChrIV: 527281	*rpb1-1* derived KWY1302 from Green et al.2012
PCCPL828	*W303, rnh1::HYG rnh201::HIS3*		this work
PCCPL829	*W303, heh1Δ::natMX6 rnh1::HYG rnh201::HIS3*		this work
PCCPL830	*W303, heh2Δ::kanMX6 rnh1::HYG rnh201::HIS3*		this work
PCCPL831	*W303, heh1Δ::natMX6 heh2Δ::kanMX6 rnh1::HYG*		this work
PCCPL832	W303, *heh1Δ::natMX6 heh2Δ::kanMX6 rnh201::HIS3*		this work
PCCPL833	*W303, heh1Δ::natMX6 heh2Δ::kanMX6 rnh1::HYG rnh201::HIS3*		this work
PCCPL841	W303, *LacI-GFP::HIS3 URA3-LacO::TRP1 HMG1-mCherry::kanMX*	LacO inserted on ChrV: 116227	this work

Table 3

Name	Sequence	Source	Target
PRPC355-f	GGGTCCTGTATGGCTTCATTTA	Lab.	ENA 1-4
PRPC355-r	GCCGCAGAACGTGATCTATAA	Lab.	ENA 1-5
PRPC500-f	GCACCGGTAGTGAATGTATGTA	Lab.	ALG9
PRPC500-r	CACCTGGAAGAAGACCATCAA	Lab.	ALG9
PRPC354-f	TACTGCGACCCAGACTCTTA	Lab.	KRS1
PRPC354-r	CTTCACCTTCACCACCTTTCT	Lab.	KRS1
PRPC358-f	CCGAAGTGCTACAGATGTTGATA	Lab.	RSM10
PRPC358-r	GAACCCAAGTCCAGCGATATT	Lab.	RSM10
PRPC359-f	GATTCCGGTGATGGTGTTACTC	Lab.	ACT1
PRPC359-r	TCAAATCTCTACCGGCCAAATC	Lab.	ACT1
PRPC722-f	TGAGCAAACAATTTGAACAG	Hage et al., 2014	CEN16
PRPC722-r	CCGATTTCGCTTTAGAAC	Hage et al., 2014	CEN16