Stimulation of adaptive gene amplification by origin firing under replication fork constraint

Alex J. Whale [1], **Michelle King** [1], **Ryan M. Hull** [1], **Felix Krueger** [2] **and Jonathan Houseley** [1,*]

[1]Epigenetics Programme, Babraham Institute, Cambridge, UK and [2]Babraham Bioinformatics, Babraham Institute, Cambridge, UK

ABSTRACT

Adaptive mutations can cause drug resistance in cancers and pathogens, and increase the tolerance of agricultural pests and diseases to chemical treatment. When and how adaptive mutations form is often hard to discern, but we have shown that adaptive copy number amplification of the copper resistance gene *CUP1* occurs in response to environmental copper due to *CUP1* transcriptional activation. Here we dissect the mechanism by which *CUP1* transcription in budding yeast stimulates copy number variation (CNV). We show that transcriptionally stimulated CNV requires TREX-2 and Mediator, such that cells lacking TREX-2 or Mediator respond normally to copper but cannot acquire increased resistance. Mediator and TREX-2 can cause replication stress by tethering transcribed loci to nuclear pores, a process known as gene gating, and transcription at the *CUP1* locus causes a TREX-2-dependent accumulation of replication forks indicative of replication fork stalling. TREX-2-dependent *CUP1* gene amplification occurs by a Rad52 and Rad51-mediated homologous recombination mechanism that is enhanced by histone H3K56 acetylation and repressed by Pol32 and Pif1. *CUP1* amplification is also critically dependent on late-firing replication origins present in the *CUP1* repeats, and mutations that remove or inactivate these origins strongly suppress the acquisition of copper resistance. We propose that replicative stress imposed by nuclear pore association causes replication bubbles from these origins to collapse soon after activation, leaving a tract of H3K56-acetylated chromatin that promotes secondary recombination events during elongation after replication fork re-start events. The capacity for inefficient replication origins to promote copy number variation renders certain genomic regions more fragile than others, and therefore more likely to undergo adaptive evolution through *de novo* gene amplification.

INTRODUCTION

Adaptive mutations can enable organisms to tolerate or even thrive in hostile environments. Although all kinds of mutation can be adaptive, CNV - the loss or duplication of segments of genetic material - often underlies adaptation in eukaryotic cells from fungi to mammals ([1]). Adaptive mutation is frequently reported in chemotherapy resistant cancers and infections ([2–7]), or treatment resistant animal and plant pests ([8,9]), so the mechanisms by which adaptive mutations form is of considerable medical, economic and societal interest.

Three major classes of mechanism are implicated in *de novo* CNV (reviewed in ([10–12])): firstly, non-allelic homologous recombination either in mitosis or meiosis can occur when a double strand break (DSB) forms within a region homologous to multiple sites in the genome. Strand invasion of the resected DSB into an unmatched homologue may result in duplication, deletion or translocation depending on resolution (reviewed in ([13])). Secondly, non-homologous end joining can ligate unmatched DSB ends to create deletions and translocations (reviewed in ([14])). Thirdly, replication fork switching between either homologous or microhomologous templates creates discontinuities in the sequence of a daughter chromatid, resulting in CNV or translocations (reviewed in ([11]) and ([15])). All three classes can initiate further complex genome rearrangements by forming unstable species such as dicentric chromosomes or extrachromosomal DNA (reviewed in ([12]) and ([14])), and given that adaptive mutations are normally observed only after extended selection it is often difficult to confirm formation mechanisms.

*To whom correspondence should be addressed. Tel: +44 1223 496113; Email: jon.houseley@babraham.ac.uk
Present addresses:
Michelle King, Department of Biomedical Science, University of Sheffield, Sheffield, UK.
Ryan M. Hull, SciLifeLab, Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Solna, Sweden.

DNA replication has particular potential to invoke genetic change, and the copious CNV events induced by replication inhibitors such as hydroxyurea or aphidicolin show that non-conservative repair is not an uncommon outcome of replication fork stalling (16–18). Stalled replication forks that cannot be restarted by other means need to be repaired through recombination (19). The simplest model for recombinational repair invokes cleavage of the recombination fork by a structure specific endonuclease (SSE), often Mus81, to create a single ended DSB that can invade the sister chromatid in a process known as Break Induced Replication (BIR) (20,21). Studies using defined nuclease-induced breaks have determined the replication mechanism of BIR: an initial strand invasion mediated by Rad51 creates a D-loop that is extended by DNA polymerase δ (22–25). Replication proceeds in a migrating D-loop rather than a classical replication fork (26,27), with delayed second strand synthesis through copying of the newly synthesized leading strand (27,28). Elongation of the leading strand in the migrating D-loop relies on polymerase δ subunit Pol32, which is dispensable during normal DNA replication, and also the helicase Pif1 (25–27).

Replication forks stalled at an inducible barrier follow a similar but not identical mechanism; replication after restart is performed by DNA polymerase δ and becomes highly error prone (29–31), but the uncoupled semiconservative lagging strand synthesis observed is different to DSB-induced BIR systems and not easily reconciled with the migrating D-loop model (32,33). Indeed, stalled forks can initiate strand invasion via fork reversal without cleavage to form a DSB, which likely explains the difference in elongation mechanism (34,35), although replication fork cleavage by MUS81 and repair by a POLD3 (human Pol32)-dependent BIR mechanism is observed in cancer cells exposed to aphidicolin (36,37). Therefore, repair of stalled replication forks in different organisms and conditions seems to occur through related but non-identical BIR-type mechanisms involving Pol32-dependent synthesis by DNA polymerase δ with uncoupled second strand synthesis and high error rates.

Irrespective of this mechanistic variation, the use of BIR-type mechanisms for replication fork repair seems to be rare and stalled forks are normally resolved by a converging replication fork (38–40). Restricting the use of BIR and related mechanisms makes sense given that BIR-type elongation is prone to mutations (reviewed in (41)), and impaired by encounters with transcribed loci and chromatin marked with Histone 3 Lysine 56 acetylation (H3K56ac) (42,43). Potentially mutagenic BIR-type mechanisms are therefore used only as a last resort, when a converging replication fork does not reach the stalled fork by late G2, and even in this case only minimal chromosomal regions are replicated before encountering a converging fork.

Adaptive mutations (including CNV) emerge through natural selection acting on random mutations. However, all types of mutation have a mechanistic cause that delimits frequency and genomic location, even if the phenotypic outcome of a given mutation is random. Mutation rate in any given genomic window may therefore be constant across time if the environment is constant or if all potentially mutagenic mechanisms acting at that locus are unaffected by environmental change. However, environmental change may disrupt normal DNA processing genome-wide or at specific genomic locations, making use of potentially mutagenic BIR-type mechanisms more frequent. For example, induction of a gene in response to environmental change can impede oncoming replication forks, leading to site specific, environmentally-stimulated mutation ((44,45) and reviewed in (46,47)).

Indeed, we and others have demonstrated that CNV events at the budding yeast *CUP1* locus are stimulated by transcriptional induction of the *CUP1* gene and are tightly localized to the *CUP1* region (48–50). *CUP1* encodes a metallothionein that protects yeast from environmental copper, and the copy number of the *CUP1* gene defines copper resistance such that adaptation to toxic levels of environmental copper occurs primarily through *CUP1* gene amplification (51–53). It is known that RNA polymerase II transcription can impair replication (reviewed in (54)) through head-on collisions between RNA and DNA polymerases (55), generation of torsional stresses (56,57), formation of RNA:DNA hybrids (R-loops) (44,45,58) or formation of secondary DNA structures (59,60), all of which block fork progression and/or cause replication slippage. However, we found that *CUP1* CNV absolutely requires the histone modification Histone 3 Lysine 56 acetylation (H3K56ac) (49), which is not obviously related to any of these outcomes.

Here, we investigate the mechanism by which *CUP1* transcriptional induction causes CNV via H3K56ac, showing critical roles for TREX-2 and Mediator as well as late-firing replication forks adjacent to the *CUP1* genes. We propose a model involving replication origin firing, stalling and collapse that links *CUP1* transcription, errors in BIR and local histone modification to *de novo* CNV.

MATERIALS AND METHODS

Yeast strains and media

Yeast strains used are listed in Supplementary Table S1. Deletion strains were produced by standard deletion protocols using oligonucleotides listed in Supplementary Table S2 and validated by PCR. Construction of 3x*CUP1* new ARS and control strains: fragments of pRS316 containing *URA3* region ± ARS were amplified using oligonucleotides listed in Supplementary Table S2 and integrated in YRH23. Construction of *3xCUP1* no ARS strain: pJH285 containing one repeat *GFP-CUP1* was formed by ligating the XmaI BglII-(blunt) fragment of pFA6a-GFP-KanMX6 into pJH254 (49) digested with XmaI EcoRV. The three repeat plasmid pJH287 was formed by ligating three fragments—pJH285 ClaI SalI, pJH285 XhoI BglII and pJH285 BamHI EcoRI—simultaneously into EcoRI ClaI digested pJH264 (49). This construct was integrated into genome of YRH15 as described in (49).

All cells were cultured in shaking incubators at 30°C, 200 rpm. Overnight cultures and cells used in standard experiments were grown in yeast nitrogen base (YNB) media (which contains 250 nM CuSO$_4$) that was supplemented with CSM amino acids and 2% glucose (or 2% raffinose with 0.02% galactose when stated). YNB media and supplements were purchased from Formedium. Pre-cultures for all copper experiments were grown to saturation (∼2 days) in

4 ml cultures of YNB media and then diluted 1:2000 for subsequent treatments. For nicotinamide treatment, cells were cultured for one week in 4 ml YNB media with a final concentration of 5 mM Nicotinamide (Sigma I17451). For copper treatment, cells were grown in 4 ml YNB media $\pm$ 0.3 mM $CuSO_4$ for 1 week. For northern blot analysis, cells were grown in 4 ml YNB media with 2% glucose for 6 h, diluted and grown overnight in 25 ml same media to 0.6–0.8 $\times$ 10^7 cells/ml. Un-induced cells were harvested, cells were diluted to 0.15 $\times$ 10^7 cells/ml in 25 ml with 0.3 mM $CuSO_4$ and grown for 6 h before harvesting 2 $\times$ 10^7 cells by centrifugation and freezing on N_2. For TrAEL-seq experiments, cells were pre-cultured by inoculation in 4 ml yeast peptone broth containing 2% raffinose (YP Raf) for ~6 h at 30°C with shaking at 200 rpm. These cells were then diluted in 100 ml YP Raf (wild-type and *rad52Δ* cells ~1:500, *sac3Δ* cells ~1:50) and growth continued at 30°C 200 rpm for ~16 h until OD_{600} reached ~0.2. These 100 ml cultures were then split, with 50 ml transferred into 50 ml YP Raf media and the other 50 ml being transferred into 50 ml YP Raf media containing 0.02% galactose for 6 h at 30°C 200 rpm. Cells were centrifuged 1 min at 4600 rpm, resuspended in 70% ethanol at 1 $\times$ 10^7 cells/ml and stored at −70°C. YP media, raffinose and galactose were purchased from Formedium.

Adaptation assay

From saturated cultures grown ±0.3 mM $CuSO_4$, a 1:80 dilution in 200 µl of YNB media was placed in every well in a flat-bottomed 96-well cell culture plate with $CuSO_4$ at the required concentration. Plates were sealed using a gas-permeable membrane and incubated at 30°C with shaking for 3 days. Cells were resuspended and OD_{660} was measured by a BD FLUOstar Omega plate reader. Area-Under-Curve for plots of OD_{660} against [$CuSO_4$] were calculated for each sample and compared by one way ANOVA with Sidak's multiple comparison correction in GraphPad Prism (v8.2.1).

DNA extraction and Southern blotting

From a saturated culture, 2 ml of cells were washed in 50 mM EDTA and then spheroplasted using 250 µl of 0.34 U/ml lyticase (Sigma L4025) in 1.2 M sorbitol, 50 mM EDTA and 10 mM DTT at 37°C for 45 min. These cells were centrifuged at 1000 rcf, gently resuspended in 400 µl of 100 µg/ml RNase A (Sigma R4875), 50 mM EDTA and 0.3% SDS, and then incubated for 30 min at 37°C. After this, 4 ul of 20 mg/ml proteinase K (Roche 3115801) was added, mixed by inverting the samples and heated at 65°C for 30 min. The samples were then left to cool to room temperature before adding 160 µl of 5 M KOAc, then mixed by inversion and chilled on ice for 1 h. These samples were centrifuged at 20 000 rcf for 10 min before the supernatant was poured into a new tube containing 500 µl of phenol:chloroform (pH8) and placed on a rotating wheel for 30 min. After centrifugation at 10 000 rcf for 10 min, the upper phase was extracted using wide bore pipette tips and precipitated in 400 µl isopropanol. Pellets were then washed in 70% ethanol, left to air-dry and then digested overnight at 37°C in 50 µl

TE with 20 U EcoRI-HF (NEB). Samples were extracted with 50 µl phenol:chloroform, then ethanol precipitated in a 1.5 ml tube containing 112.5 µl 100% ethanol and 4.5 µl 3 M NaOAc before centrifugation at 20 000 rcf for 15 min. After washing in 70% ethanol, the pellets were dissolved for 1 h in 20 µl TE. Loading dye was added and samples were separated on 25 cm 0.8% or 1% TBE gels at 120 V for 16.5 h. Gels were denatured in 0.25 N HCl for 15 min, neutralized in 0.5 N NaOH for 45 min and washed twice in 1.5 M NaCl, 0.5 M Tris (pH 7.5) for 20 min each. Samples were transferred to HyBond N + membrane in 6× SSC through capillary action overnight and fixed by UV crosslinking using a Stratagene UV Stratalinker. Membranes were probed using random primed probes (listed in Supplementary Table S2) in 10 ml UltraHyb (AM8669 ThermoFisher Scientific) at 42°C then washed with 0.1× SSC 0.1% SDS at 42°C. Quantification of Southern blot bands was performed using ImageQuant (Version 7.0, GE), and CNV calculated as (intensity of all CNV bands/intensity of CNV and parental bands) × 100. Statistical analysis of CNV levels was performed by one-way ANOVA with Sidak's multiple comparison correction in GraphPad Prism (v8.2.1).

RNA extraction and northern blotting

Frozen cell pellets were lysed by 5 min vortexing at 4°C with ~50 µl glass beads and 40 µl GTC-phenol (2.1 M guanidine thiocyanate, 26.5 mM Na citrate pH 7, 5.3 mM EDTA, 76 mM β-mercaptoethanol, 1.06% *N*-lauryl sarcosine, 50% phenol pH 7). 600 µl GTC-phenol was added, mixed, and samples were heated at 65°C for 10 min then placed on ice for 10 min. 160 µl 100 mM NaOAc pH 5.2 and 300 µl chloroform:isoamyl alcohol (24:1) were added, samples were vortexed and centrifuged at top speed for 5 min. The upper phase was re-extracted first with 500 µl phenol:chloroform pH 7 (1:1) and then with 500 µl chloroform:isoamyl alcohol (24:1) before precipitation with 1 ml ethanol. Pellet was washed with 70% ethanol, re-suspended in 6 µl water and quantified using Quant-IT RiboGreen (ThermoFisher, R11490). 1 µg RNA was resolved per lane on 1.2% glyoxal agarose gels, blotted and probed with a random primed probe against *CUP1* ORF (Supplementary Table S2) as described (61).

Candidate genetic screen

A total of 206 strains from the Yeast Deletion Collection (Invitrogen 95401.H2) and other sources were streaked out on YPD agar plates and then re-streaked for single colonies on YPD plates containing 300 µg/ml G418. Pre-cultures were grown to saturation in 4 ml YNB media at 30°C, diluted 1:2000 in 4 ml YNB media with and without 5 mM Nicotinamide and then incubated for 1 week at 30°C with shaking. DNA extraction and Southern blotting was performed as described above. CNV rates of mutants were obtained by comparing the percentage of CNV alleles in nicotinamide-treated mutants to nicotinamide-treated wildtype cells, and calculating the fold change in CNV.

For network analysis, factors from the CNV screen and their first neighbours were imported into Cytoscape (v3.7.2) using stringApp (v1.6.0) (62) to retrieve *Saccharomyces*

cerevisiae protein interaction data and to construct the network. The Edge-weighted Spring Embedded layout was applied using 'stringdb score' to determine edge length, and node size and colour were mapped to fold-change in CNV from the CNV screen with labels applied to the strongest CNV enhancers and suppressors (>2 or <0.5 fold-change in CNV, respectively). Cluster Analysis was performed on this network using the ClusterONE app (v1.0) (63) which was used to identify clusters of proteins, and clusters with a similar impact on fold-change in CNV have their functional categories displayed.

TrAEL-seq library preparation and sequencing

$1–3 \times 10^7$ cells fixed in ethanol were pelleted by centrifugation for 30 s at 20 000 g, rinsed in 1 ml PFGE wash buffer (10 mM Tris HCl pH 7.5, 50 mM EDTA), and resuspended in 60 µl PFGE wash buffer containing 1 µl lyticase (17 U/µl 10 mM KPO$_4$ pH 7, 50% glycerol Merck L2524 > 2000 U/mg) then incubated for 10 min at 50°C. 40 µl of molten CleanCut agarose (Bio-Rad 1703594) cooled to 50°C was added, samples were vortexed vigorously for 5 s and pipetted into a plug mould (Bio-Rad 1703713), then left to solidify for 30 min at 40°C. Plugs were transferred into a 2 ml Eppendorf that contained 500 µl PFGE wash buffer containing 10 µl 17 U/ml lyticase and left to incubate at 37°C for 1 h. This solution was removed and replaced with 500 µl PK buffer (100 mM EDTA pH 8, 1 mg/ml Proteinase K, 1% sodium *N*-lauroyl sarcosine, 0.2% sodium deoxycholate) at 50°C overnight. The plugs were then rinsed in 1 ml TE and washed with 1 ml TE for 1 h with rocking. Plugs were then washed twice with 1 ml TE containing 10 mM PMSF (Merck 93482) for 1 h with rocking. Finally, plugs were digested in 200 µl TE containing 1 µl 1000 U/ml RNase T1 (Thermo EN0541) at 37°C for 1 h before being stored at 4°C in 1 ml TE.

A $\frac{1}{2}$ plug was used for each sample (referred to here on in as plugs). Plugs were equilibrated once in 100 µl 1× TdT buffer (NEB) for 30 min at room temperature, then incubated for 2 h at 37°C in 100 µl 1× TdT buffer containing 4 µl 10 mM ATP and 1 µl Terminal Transferase (NEB M0315L). Plugs were rinsed with 1 ml tris buffer (10 mM Tris–HCl pH 8.0), equilibrated in 100 µl 1× T4 RNA ligase buffer (NEB) containing 40 µl 50% PEG 8000 for 1 h at room temperature then incubated overnight at 25°C in 100 µl 1× T4 RNA ligase buffer (NEB) containing 40 µl 50% PEG 8000, 1 µl 10 pM/µl TrAEL-seq adaptor 1 (64) and 1 µl T4 RNA ligase 2 truncated KQ (NEB M0373L). Plugs were then rinsed with 1 ml tris buffer, transferred to 15 ml tubes and washed three times in 10 ml tris buffer with rocking at room temperature for 1–2 h each, then washed again overnight under the same conditions. Plugs were equilibrated for 15 min with 1 ml agarase buffer (10 mM Bis–Tris–HCl, 1 mM EDTA pH 6.5), then the supernatant removed and 50 µl agarase buffer added. Plugs were melted for 20 min at 65°C, transferred for 5 min to a heating block preheated to 42°C, 1 µl β-agarase (NEB M0392S) was added and mixed by flicking without allowing sample to cool, and incubation continued at 42°C for 1 h. DNA was ethanol precipitated with 25 µl 10 M NH$_4$OAc, 1 µl GlycoBlue, 330 µl of ethanol and resuspended in 10 µl 0.1× TE. 40 µl re-

action mix containing 5 µl Isothermal amplification buffer (NEB), 3 µl 100 mM MgSO$_4$, 2 µl 10 mM dNTPs and 1 µl Bst 2 WarmStart DNA polymerase (NEB M0538S) was added and sample incubated 30 min at 65°C before precipitation with 12.5 µl 10 M NH$_4$OAc, 1 µl GlycoBlue, 160 µl ethanol and re-dissolving pellet in 130 µl 1× TE. The DNA was transferred to an AFA microTUBE (Covaris 520045) and fragmented in a Covaris E220 using duty factor 10, PIP 175, Cycles 200, Temp 11°C, then transferred to a 1.5 ml tube containing 8 µl pre-washed Dynabeads MyOne streptavidin C1 beads (Thermo, 65001) re-suspended in 300 µl 2× TN (10 mM Tris pH 8, 2 M NaCl) along with 170 µl water (total volume 600 µl) and incubated 30 min at room temperature on a rotating wheel. Beads were washed once with 500 µl 5 mM Tris pH 8, 0.5 mM EDTA, 1 M NaCl, 5 min on wheel and once with 500 µl 0.1× TE, 5 min on wheel before re-suspension in 25 µl 0.1× TE. Second end processing and library amplification were performed with components of the NEBNext Ultra II DNA kit (NEB E7645S) and a NEBNext Multiplex Oligos set (e.g. NEB E7335S). 3.5 µl NEBNext Ultra II End Prep buffer, 1 µl 1 ng/µl sonicated salmon sperm DNA (this is used as a carrier) and 1.5 µl NEBNext Ultra II End Prep enzyme were added and reaction incubated 30 min at room temperature and 30 min at 65°C. After cooling, 1.25 µl 10 pM/µl TrAEL-seq adaptor 2 (64), 0.5 µl NEBNext ligation enhancer and 15 µl NEBNext Ultra II ligation mix were added and incubated 30 min at room temperature. The reaction mix was removed and discarded and beads were rinsed with 500 µl wash buffer (5 mM Tris pH 8, 0.5 mM EDTA, 1 M NaCl) then washed twice with 1 ml wash buffer for 10 min on wheel at room temperature and once for 10 min with 1 ml 0.1× TE. Libraries were eluted from beads with 11 µl 1× TE and 1.5 µl USER enzyme (NEB) for 15 min at 37°C, then again with 10.5 µl 1× TE and 1.5 µl USER enzyme (NEB) for 15 min at 37°C, and the two eluates combined. An initial test amplification was used to determine the optimal cycle number for each library. For this, 1.25 µl library was amplified in 10 µl total volume with 0.4 µl each of the NEBNext Universal and any NEBNext Index primers with 5 µl NEBNext Ultra II Q5 PCR master mix. Cycling program: 98°C 30 s then 18 cycles of (98°C 10 s, 65°C 75 s), 65°C 5 min. Test PCR was cleaned with 8 µl AMPure XP beads (Beckman A63881) and eluted with 2.5 µl 0.1× TE, of which 1 µl was examined on a Bioanalyser high sensitivity DNA chip (Agilent 5067-4626). Ideal cycle number should bring final library to final concentration of 1–3 nM, noting that the final library will be 2–3 cycles more concentrated than the test anyway. 21 µl of library was then amplified with 2 µl each of NEBNext Universal and chosen Index primer and 25 µl NEBNext Ultra II Q5 PCR master mix using same conditions as above for calculated cycle number. Amplified library was cleaned with 40 µl AMPure XP beads (Beckman A63881) and eluted with 26 µl 0.1× TE, then 25 µl of this was again purified with 20 µl AMPure XP beads and eluted with 11 µl 0.1× TE. Final libraries were quality controlled and quantified by Bioanalyser (Agilent 5067-4626) and KAPA qPCR (Roche KK4835). Libraries were sequenced on an Illumina NextSeq 500 as High Output 75 bp Single End by the Babraham Institute Next Generation Sequencing facility.

TrAEL-seq data processing

Unique Molecular Identifier (UMI) deduplication and mapping: Scripts used for UMI-handling as well as more detailed information on the processing are available here: https://github.com/FelixKrueger/TrAEL-seq). Briefly, TrAEL-seq reads carry an 8 bp in-line barcode (UMI) at the 5′-end, followed by a variable number of 1–3 thymines (T). Read structure is therefore NNNNNNNN(T)nSEQUENCESPECIFIC, where NNNNNNNN is the UMI, and (T)n is the poly(T). The script TrAELseq_preprocessing.py removes the first 8 bp (UMI) of a read and adds the UMI sequence to the end of the readID. After this, up to 3 T (inclusive) at the start of the sequence are removed. Following this UMI and Poly-T pre-processing, reads underwent adapter- and quality trimming using Trim Galore (v0.6.5; default parameters; https://github.com/FelixKrueger/TrimGalore). UMI-pre-processed and adapter-/quality trimmed files were then aligned to the respective genome using Bowtie2 (v2.4.1; option: –local; http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) using local alignments. Finally, alignment results files were deduplicated using UmiBam (v0.2.0; https://github.com/FelixKrueger/Umi-Grinder). This procedure deduplicates alignments based on the mapping position, read orientation as well as the UMI sequence.

To assist with the interpretation of aligned multi-copy sequences, the P_{GAL1}-HA cup1 samples were treated in a more specialized way before entering the TrAEL-seq processing procedure outlined above: Prior to TrAEL-seq pre-processing, sequences were deduplicated based on the first 23 bp on their 5′-end (using the script TrAELseq_sequence_based_deduplication.py). This region contains both the UMI sequence as well as the first 15 bp of genomic sequence, and should thus help identify (and remove) PCR amplified multi-copy sequences that would under normal conditions survive the UMI-aware deduplication procedure by aligning to several different genomic regions at random. Following deduplication-by-sequence and TrAEL-seq pre-processing, these sequences were aligned to a modified version of the yeast genome containing the additional P_{gal}-HA control sequences. To avoid multi-mapping artefacts arising from the integration of these sequences, the following two stretches of genomic sequence were masked by Ns: a) *CUP1* (chromosome VIII:212266–216251), and b) P_{GAL1} (chromosome II:278352–279023). Reads were then trimmed and mapped as above.

De-duplicated mapped reads were imported into Seq-Monk v1.47 (https://www.bioinformatics.babraham.ac.uk/projects/seqmonk/) and immediately truncated to 1 nucleotide at the 5′ end, representing the last nucleotide 5′ of the strand break. Reads were then summed in running windows as described in figure legends. Windows overlapping with non-single copy regions of the genome were filtered (rDNA, 2μ, mtDNA, sub-telomeric regions, Ty elements and LTRs), and total read counts across all included windows were normalized to reads per million mapped. A further enrichment normalization (20–90%) was applied to match the read count distributions of the P_{GAL1}-HA cup1 libraries. Read counts were exported for the consensus P_{GAL1}-HA cup1 region and plotted in GraphPad Prism

8. Comparison of datasets was performed using edgeR implemented in SeqMonk (65). For read polarity plots, forward and reverse read counts were quantitated in running windows as specified in the relevant figure legends before export for plotting using R v4.0.0 in RStudio. Read polarity values were calculated and plotted as either dots (individual samples) or as a continuous line (multiple sample display) for each quantification window using the formula read polarity = $(R - F)/(R + F)$, where F and R relate to the total forward and reverse read counts respectively. The R code to generate these plots can also be found here: https://github.com/FelixKrueger/TrAEL-seq.

Image processing

Gel images were processed in ImageQuant TL (v7.0) which involved cropping, rotating and altering contrast of whole images to improve visualization of bands.
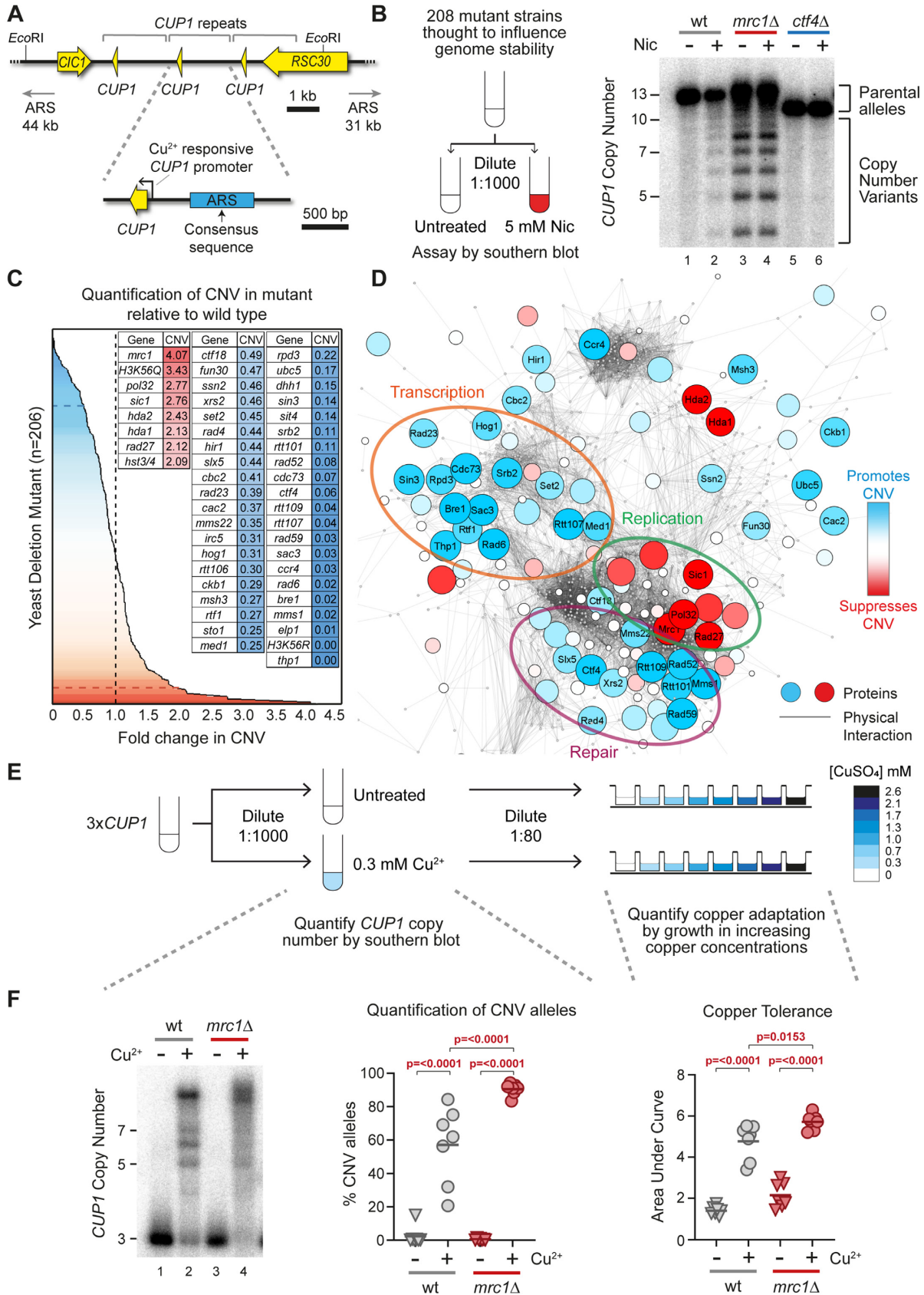
RESULTS

A genetic screen identifies enhancers and suppressors of *CUP1* CNV

The *CUP1* locus on chromosome VIII is composed of 1 or more tandem copies of a 2 kb sequence containing the *CUP1* gene with a copper responsive *CUP1* promoter and a poorly defined replication origin (an autonomously replicating sequence or ARS) (Figure 1A). Transcriptional activation of the locus promotes *de novo* CNV events including gene amplifications that increase copper resistance, such that exposure of budding yeast to copper-rich environments leads to the formation of new *CUP1* alleles that improve resistance (49).

To identify factors important for *CUP1* CNV, we performed a candidate genetic screen using 208 mutants drawn largely from the Yeast Deletion Collection (66), but with selected double deletion mutants, hypomorphs and histone point mutants (see Supplementary Table S3 for data on all strains tested). Mutants were selected to survey a wide range of mechanisms involved in RNA processing, DNA replication, DNA repair and genome stability. The S288C genetic background of these mutants includes a 13 copy *CUP1* array that rarely amplifies but does undergo transcriptionally stimulated contraction through an H3K56ac-dependent mechanism (49). Nicotinamide, an inhibitor of H3K56 deacetylases, accentuates CNV stimulation by transcription, such that even basal *CUP1* expression in the absence of copper becomes sufficient to cause measurable contractions (49). Fold-change in CNV relative to wild type can then be quantified by Southern blot after 10 generations ± nicotinamide to reveal enhancers or suppressors of CNV such as *MRC1* or *CTF4* (Figure 1B).

Seven deletion mutants exhibited a >2-fold increase in CNV, including four DNA replication mutants (*mrc1Δ*, *pol32Δ*, *sic1Δ* and *rad27Δ*) and three histone deacetylase mutants (*hda1Δ*, *hda2Δ* and *hst3Δ hst4Δ*). Forty-one mutants suppressed CNV >2-fold, representing a wider range of biological processes including regulation of transcription (*thp1Δ*, *med1Δ*, *cdc73Δ* etc.), DNA repair (*rtt107Δ*, *rad59Δ*, *rad52Δ* etc.), nucleosome assembly (*rtt106Δ*, *cac2Δ*, *hir1Δ* etc.) and various histone modi-

Figure 1. A screen for genes regulating transcriptionally-stimulated CNV. (**A**) Schematic of the *CUP1* array and the surrounding region of Chromosome VIII. ORFs are shown in yellow, and grey brackets indicate the repeated region. Each repeat contains a replication origin (ARS); the ARS consensus

fiers (*bre1Δ, rtt109Δ, rpd3Δ* etc.) (Figure 1C). ClusterONE, which identifies clusters of interacting proteins in a physical interaction network, identified clusters of proteins that have similar effects on CNV. Clusters of proteins that promote CNV are involved in chromatin/transcription (orange circle) and DNA repair (purple circle), while a cluster of proteins that facilitate DNA replication & cell cycle progression all tend to suppress CNV (green circle) (Figure 1D). These clusters are coherent with the genetic dependencies we have previously observed (49).

To validate the importance of these genes in adaptive *CUP1* amplification, selected genes were deleted in a tester strain that carries only three copies of *CUP1* (3x*CUP1*) (Figure 1E and Supplementary Figure S1). During growth in sub-lethal copper, 3x*CUP1* cells that undergo *CUP1* amplification gain a selective advantage and become dominant in the population. *CUP1* amplification allows growth in higher concentrations of copper, so cultures grown in sub-lethal copper acquire increased resistance as cells with amplified alleles proliferate. These phenotypes can be quantified by Southern blot and copper resistance assays respectively (Figure 1F). For example, deletion of *MRC1* in the 3x*CUP1* background increases the proportion of cells that acquire adaptive *CUP1* amplifications over 10 generations growth in sub-lethal copper, which increases adaptation of the population to higher concentrations of copper.

Therefore, our candidate genetic screen for enhancers and suppressors of *CUP1* CNV provides a genetic profile consistent with a replication-linked recombination mechanism driven by transcription.

TREX-2 and Mediator are required for transcriptionally stimulated CNV

Many recent studies have shown that RNA:DNA hybrids called R-loops impair replication fork progression, providing a well-validated mechanism for transcription-associated recombination ((67), reviewed in (68)). In consequence, we expected mutants affecting R-loop formation or processing to show strong phenotypes in the genetic screen, but *hpr1Δ*, *tho2Δ* and *thp2Δ* cells (THO Complex mutants) which accumulate R-loops had little effect (69–71) (Supplementary Table S3). Furthermore, *rnh1Δ rnh201Δ* cells known to accumulate high levels of R-loops underwent copper adaptation and *CUP1* CNV at normal rates in the 3x*CUP1* background (Figure 2A) (72–74). We therefore find no evidence that R-loops are involved in transcriptionally stimulated *CUP1* CNV.

In contrast, two of the strongest suppressor mutants found in the screen were deletions of *SAC3* and *THP1,* which encode components of Transcription and RNA Export complex 2 (TREX-2) (75,76). Remarkably, 3x*CUP1 thp1Δ* and *3xCUP1 sac3Δ* mutants underwent no detectable *CUP1* CNV and were completely unable to adapt to copper despite showing normal resistance to sub-lethal concentrations of copper (Figure 2B, C and Supplementary Figure S2A). Although TREX-2 mutations alter expression of some genes, we did not detect any difference in the induction of *CUP1* in response to copper (Supplementary Figure S2B) (77,78). This implicates TREX-2 as a link between transcription and recombination events at *CUP1*.

TREX-2 is physically associated with the Mediator complex (77) and mediator mutant *med1Δ* also suppressed *CUP1* CNV in the genetic screen. Med1 is a component of the middle module of Mediator but has not been placed within the structure (reviewed in (79)), whereas the TREX-2/Mediator interface has been mapped to Med31, which projects out from the middle module (77,80,81). Just as we observed in TREX-2 mutants, 3x*CUP1 med31Δ* cells did not undergo detectable CNV or adaptation during growth in sub-lethal copper, despite normal *CUP1* mRNA induction (Figure 2D and Supplementary Figure S2B). We further analysed cells lacking Srb2, a component of the head module located at the interface between Mediator and RNA polymerase II (82,83), and found that 3x*CUP1 srb2Δ* cells were similarly impaired both in *CUP1* amplification and acquisition of increased copper resistance (Supplementary Figure S2C). These results reveal a critical role for both TREX-2 and Mediator in the mechanism of *CUP1* amplification.

←——

sequence was defined in (124) and the blue rectangle denotes a sequence with ARS activity experimentally validated in (104) though the actual limits of the ARS element have not been determined. Schematic shows cells with three copies of the *CUP1* gene although the BY4741 background used as in the initial screen contains 13 *CUP1* copies. The nearest adjacent replication origins (ARS elements) in both directions are indicated. (**B**) Schematic of the candidate genetic screen for regulators of CNV and a representative Southern blot used for quantification of CNV. Wild type (wt) and indicated mutant cells from yeast haploid deletion collection (Invitrogen 95401.H2) and other sources (Supplementary Table S3) were grown to saturation then diluted 1:1000 and re-grown to saturation (10 generations) ±5 mM nicotinamide (Nic) and subject to southern analysis of *CUP1* copy number. Quantification of % CNV alleles was calculated as the percentage of alleles deviating from the parental copy number and fold-change was calculated relative to % CNV alleles in nicotinamide-treated wild-type cells. (**C**) Summary of genetic screen plotting fold-change in nicotinamide-induced CNV of 206 deletion strains relative to wild-type. Grey dashed line indicates wild-type fold-change in CNV. The blue dashed line represents cut-off for CNV-suppressing mutations with a <0.5 fold change in CNV and red dashed lines indicates the cut-off for CNV enhancing mutations with a >2-fold change in CNV. Genes called as enhancers (blue) or suppressors (red) are shown in inset table, full results are given in Supplementary Table S3. (**D**) Protein–Protein Physical Interaction Network of factors from the CNV screen and their first neighbours (visualized in Cytoscape Version 3.7.2). Nodes represent proteins and edges represent high-confidence physical interactions between proteins imported from stringApp (v1.6.0). The size and colour of nodes indicates deviation from wild-type fold-change in CNV, with red nodes representing increasing rates of CNV in mutant and blue nodes representing decreasing rates of CNV in mutant. First neighbours of screen factors are shown as small grey nodes. Clusters of interacting proteins (identified in the physical interaction network using ClusterONE v1) which showed similar fold-change in CNV are circled. (**E**) Methodology for assessing CNV and copper adaptation in cells containing three copies of the *CUP1* gene (*3xCUP1*). Saturated cultures of 3x*CUP1* wild-type and mutant cells were diluted 1:1000 in media ±0.3 mM CuSO$_4$ and incubated for 1 week at 30°C. Half the culture was used for Southern blot analysis, while the other half was diluted 1:80 and grown in a 96-well plate containing a range of concentrations of CuSO$_4$ to assay for copper tolerance. (**F**) Example southern blot analysis of *CUP1* locus and copper tolerance assay in wild type (wt) and *mrc1Δ* mutant strain with three copies of the *CUP1* gene, as outlined in (E); Southern blot quantification shows the percentage of alleles deviating from the parental copy number, for adaptation OD$_{600}$ was plotted against [CuSO$_4$] and copper tolerance quantified as area-under-curve for each culture; n = 7, *P*-values calculated by one-way ANOVA.
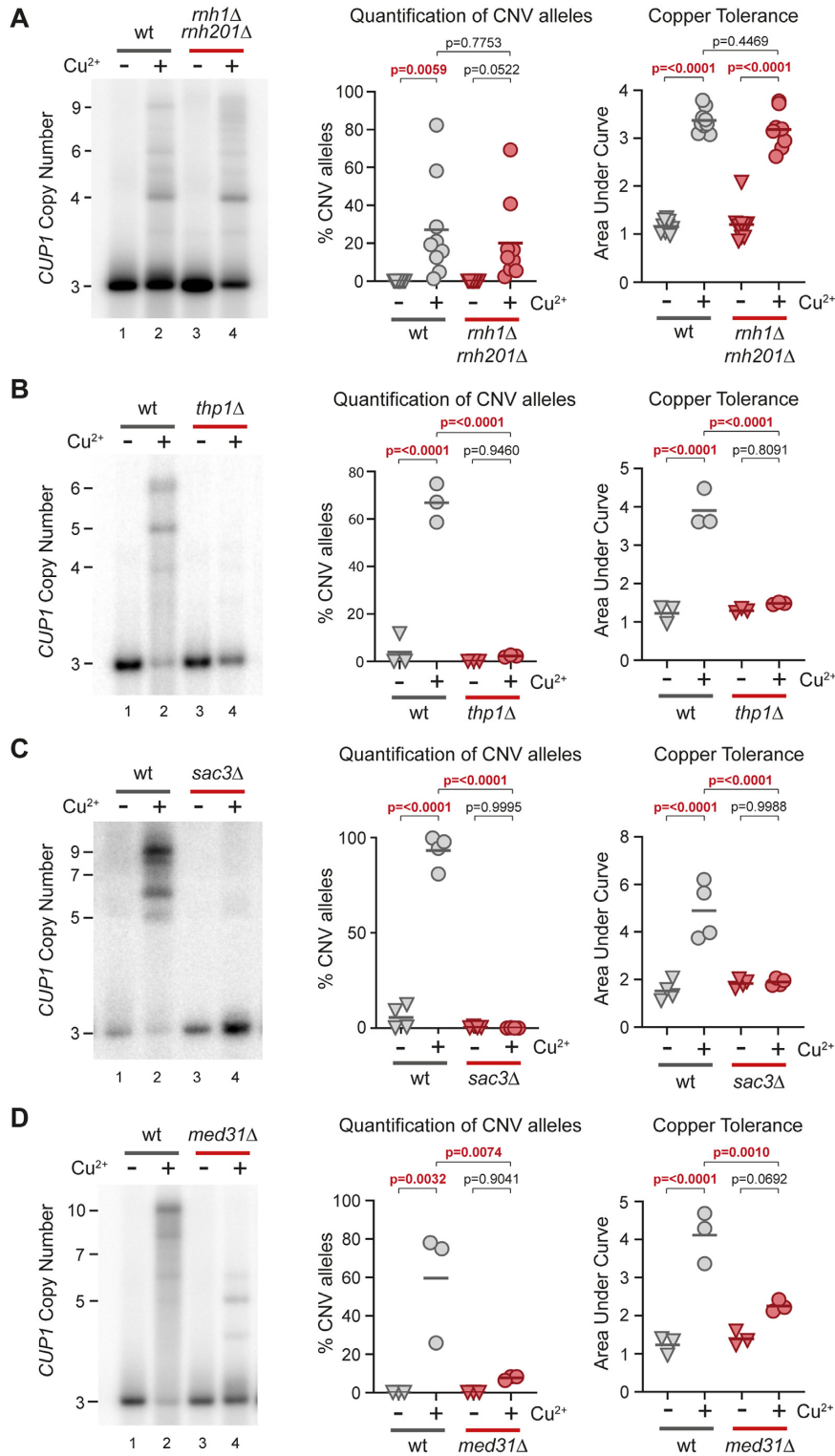
Figure 2. TREX-2 and Mediator are required for transcription stimulated *CUP1* CNV. (**A**) Southern blot analysis of *CUP1* copy number in wild-type (wt) and *rnh1Δ rnh201Δ* cells with three copies of the *CUP1* gene, grown to saturation in YNB media ±0.3 mM CuSO$_4$; $n = 9$. Quantification shows the percentage of alleles deviating from the parental copy number. Copper adaptation was assessed by diluting these same cells in 96-well plates with varying concentrations of CuSO$_4$ and incubating for 3 days (see Figure 1E). Final OD$_{600}$ was plotted against [CuSO$_4$] and copper tolerance quantified as area-under-curve for each culture; all *P*-values were calculated by one-way ANOVA. (**B**) Southern analysis of *CUP1* copy number in *3xCUP1* wild-type (wt) and *thp1Δ* cells after 10 generations ±0.3 mM CuSO$_4$ (Quantification of CNV alleles and Copper adaptation as in A); $n = 3$. (**C**) Southern analysis of *CUP1* copy number in *3xCUP1* wild-type (wt) and *sac3Δ* cells after 10 generations ±0.3 mM CuSO$_4$ (quantification of CNV alleles and copper adaptation as in A); $n = 4$. (**D**) Southern analysis of *CUP1* copy number in *3xCUP1* wild-type (wt) and *med31Δ* cells after 10 generations ±0.3 mM CuSO$_4$ (Quantification of CNV alleles and copper adaptation as in A); $n = 3$.

Mediator connects transcription factors to the core RNA polymerase II machinery at active promoters (reviewed in (84)), while TREX-2 associates with nuclear pores to promote mRNA processing and export (85). However, the gene expression and RNA export functions of Mediator and TREX-2 are separable (77), so the requirement for both complexes in *CUP1* amplification implicates their shared physical interaction. This interaction connects actively transcribed genes to the nuclear pore in a process termed gene gating (77,86–88). Gene gating can impair DNA replication by constraining DNA topology, causing replication forks rendered already fragile by HU treatment to collapse in a Rad53 mutant lacking checkpoint activity (56). This process, albeit only previously characterized under considerable replicative stress, provides a mechanism by which *CUP1* transcriptional induction acting through TREX-2 and Mediator could impede replication forks without a requirement for R-loop formation.

Evidence for replication fork stalling caused by *cup1* transcription

We recently developed TrAEL-seq to detect replication fork stalling and replication intermediates (64). Unfortunately, TrAEL-seq is ineffective for copper-treated cells as copper-induced apoptotic fragments generate a high background (64,89) (example data is deposited at GEO GSE154811). However, we have also created a P_{GAL1}-*HA cup1* strain in which *GAL1* promoters replace all *CUP1* promoters, and *CUP1* ORFs are replaced by *3HA* ORFs, allowing transcriptional induction at the *cup1* locus using the non-toxic sugar galactose (Figure 3A) (49).

TrAEL-seq reads accumulate at sites of replication fork stalling, including at the endogenous *GAL1* promoter when transcriptionally active, providing a measure of disruption to replication fork progression (64). The activity of the *GAL1* promoter is exceptionally strong under normal 2% galactose induction, leading to saturating levels of Sac3- and Thp1-independent CNV in 10 generations that probably result from direct collisions between RNA polymerase and the replication fork (Supplementary Figure S3A). However, CNV is Sac3- and Thp1-dependent under moderate 0.02% galactose induction that should better reflect *CUP1* gene expression (Figure 3B), so we constructed TrAEL-seq libraries from P_{GAL1}-*HA cup1* wild-type and *sac3*Δ cells induced with 0.02% galactose.

We expected to observe a pronounced accumulation of TrAEL-seq reads at the P_{GAL1}-*HA* promoter as a result of replication fork stalling, but the promoter peaks formed in 0.02% galactose were modest and Sac3-independent (Figure 3C), although further enhanced in 2% galactose (Figure 3D). In contrast, the region upstream of the P_{GAL1} promoter containing the ARS element accumulated significantly more TrAEL-seq reads in wild type than *sac3*Δ cells on induction with 0.02% galactose (right hand side and far left of Figure 3C). Normalization errors could also give rise to such a baseline signal increase, however global TrAEL-seq read distributions were unaffected by addition of galactose (Supplementary Figure S3B shows all equivalent windows in single copy regions of chromosome VIII), and if anything, the normalization process globally increases

*sac3*Δ signals compared to wild-type. Therefore, we observe a small Sac3-dependent increase in TrAEL-seq read density outside the transcribed region and particularly around the ARS element in P_{GAL1}-*HA cup1* suggesting that replication forks stall but not at well-defined sites.

The Sac3-dependent increase in TrAEL-seq reads across P_{GAL1}-*HA cup1* is consistent with replication forks stalling due to DNA topological constraint, which would not happen at a defined location but rather happen at random across a wide area. Stalled replication forks are most often rescued by a converging replication fork, but in areas of widespread replicative stress the converging fork may be delayed and alternative replication fork restart pathways such as BIR invoked.

CUP1 amplification occurs in response to defective DNA polymerase δ elongation

We previously proposed a model for *CUP1* CNV in which BIR-type events are initiated at the *CUP1* locus to resolve stalled replication forks, but are impeded on encountering H3K56ac chromatin, leading to secondary non-allelic recombination events that result in CNV (illustrated in Figure 4A) (49). This mechanism explains two key genetic dependencies of transcriptionally stimulated CNV at *CUP1*: the absolute dependence of CNV on the H3K56 acetyltransferase Rtt109, and the dramatic enhancement of CNV observed in cells lacking Pol32. The genetic screen confirmed the importance of H3K56ac through the impact of H3K56Q and R mutations (Figure 1C) and identified recombination factors important for *CUP1* CNV (Rad52, Rad59), but certain factors had much less effect than we expected (Rad51, Mus81) (Figure 1C, Supplementary Table S3) leading us to probe the recombination mechanism in more detail.

In this mechanism, impairment of DNA polymerase δ elongation at H3K56ac chromatin is the critical step leading to CNV. If H3K56ac is absent then elongation should proceed until eventual resolution by an oncoming replication fork, but mutations that impair DNA polymerase δ elongation should induce CNV irrespective of H3K56ac. To test this, we combined *rtt109*Δ, which abrogates H3K56ac, with *pol32*Δ, which impairs BIR elongation. *rtt109*Δ completely suppressed CNV and copper adaptation as we have previously shown (49) (Figure 4B, lanes 1–4), whereas *pol32*Δ significantly enhanced CNV (Figure 4B, lanes 5–6). Importantly, the double *rtt109*Δ *pol32*Δ mutant underwent high levels of CNV equivalent to the *pol32*Δ single mutant (Figure 4B, lanes 7–8), consistent with the idea that H3K56ac causes CNV by impairing elongation.

A specific requirement for an inefficient BIR-type mechanism is unusual and Pol32-dependence alone is open to other interpretations. We therefore examined the helicase Pif1, which is also critical for BIR elongation, and observed that deletion of *PIF1* both alone and in combination with *pol32*Δ increased amplification of the 3x*CUP1* allele compared to wild type (Figure 4C). Unfortunately, slow growth and extensive flocculation prevented determination of copper resistance in *pif1*Δ cells, but the gene amplification phenotype is consistent with the hypothesis that defects in DNA polymerase δ elongation promote *CUP1* CNV.
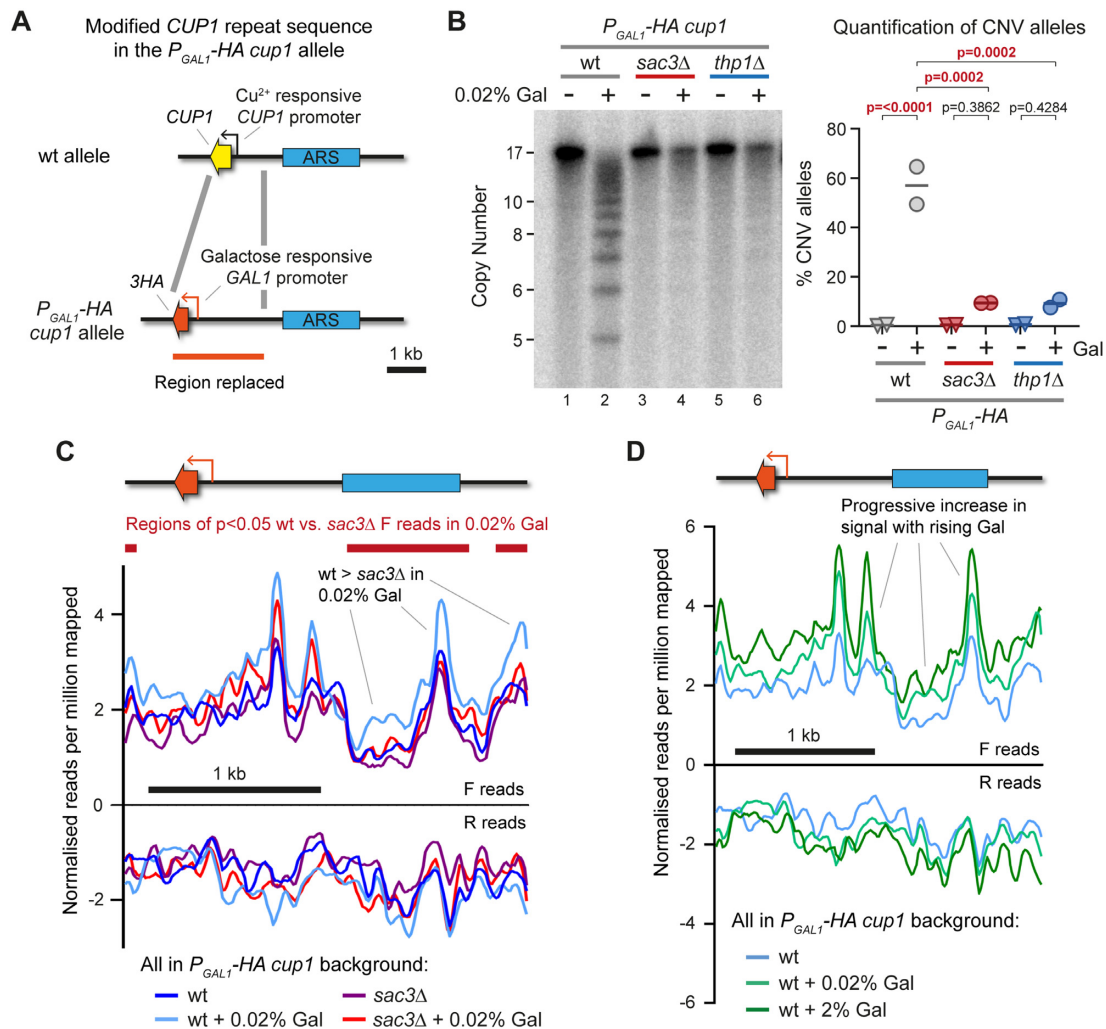
Figure 3. Replication fork stalling and cleavage at *CUP1* locus. (**A**) Schematic of modified *CUP1* repeat in P_{GAL1}-*HA cup1* cells, where every *CUP1* ORF is replaced by a 3HA coding sequence (which is phenotypically neutral) and every *CUP1* promoter is replaced by the galactose-inducible *GAL1* promoter. (**B**) Southern blot analysis of P_{GAL1}-*HA cup1* cells comparing *sac3Δ* and *thp1Δ* cells to wild-type cells grown for 10 generations in 2% raffinose ±0.02% galactose. Quantification shows the percentage of alleles deviating from the parental copy number of 17 copies; p-values were calculated by one-way ANOVA, n = 2. (**C**) Plots of TrAEL-seq read density on forward and reverse strands across a P_{GAL1}-*HA cup1* repeat. Forward TrAEL-seq reads in replicating cells arise primarily from replication forks moving right-to-left, reverse reads from forks moving left-to-right. The dominance of forward over reverse reads shows that replication direction is primarily right-to-left, accumulations of forward reads without a decrease in reverse reads most likely represents increased average fork residency indicative of slower fork progression or more frequent stalling, see Kara *et al.* for more details (64). TrAEL-seq profiles are an average of two biological replicates of P_{GAL1}-*HA cup1* wild-type and *sac3Δ* cells grown to mid-log in 2% raffinose, with or without a 6 h 0.02% galactose induction. Reads were quantified per million reads mapped in 50 bp windows spaced every 10 bp, and an enrichment normalization applied to make overall read count distributions as uniform as possible, see Supplementary Figure S3B for the distribution of reads across single copy regions of chromosome VIII. Regions of significant difference between wild type and *sac3Δ* in 0.02% galactose were called using edgeR (65), with forward reads quantified in 200 bp windows spaced every 50 bp, no normalization was applied prior to the Edge algorithm. (**D**) Plots of TrAEL-seq read density on forward and reverse strands across a P_{GAL1}-*HA cup1* repeat for wild-type cells induced for 6 h with 0%, 0.02% or 2% galactose. TrAEL-seq profiles are an average of two biological replicates, 0% and 0.02% conditions are the same data as (C), data is processed as in (C).

Whereas Pif1 and Pol32 are required for elongation in BIR-type mechanisms, the helicase Mph1 suppresses BIR through unwinding duplexes formed by strand invasion (90,91). The mechanism shown in Figure 4A requires efficient strand invasion both to initiate synthesis by DNA polymerase δ and to facilitate second strand invasion leading to CNV. This requirement is validated by the phenotype of *mph1Δ* mutants, which underwent increased *CUP1* CNV and adaptation in the 3x*CUP1* background (Figure 4D).

The phenotypic impacts of *pol32Δ*, *pif1Δ* and *mph1Δ* on CNV provide strong support for the hypothesis that *CUP1*

CNV arises through a BIR-type mechanism. However, it is not efficient processive elongation after fork restart that leads to CNV, but rather situations in which elongation is impeded by chromatin structure or genetic defects.

Requirement for structure specific endonucleases and recombination proteins in *CUP1* CNV

Once DNA polymerase δ elongation is impeded, the replication fork can be resolved either by an oncoming replication fork, which would maintain parental copy number, or
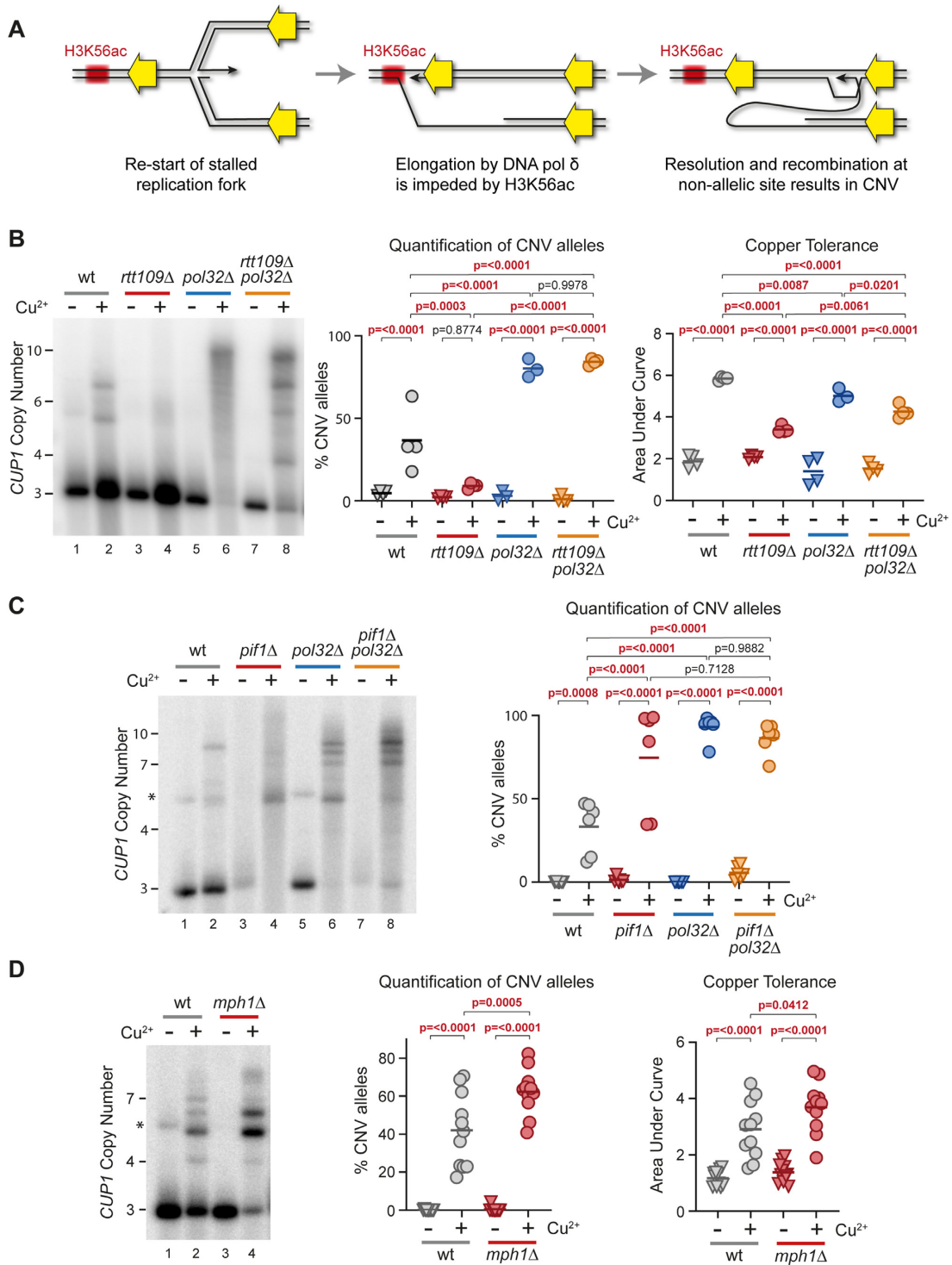
Figure 4. Importance of impaired DNA polymerase δ elongation in *CUP1* CNV. (**A**) Proposed mechanism leading to *CUP1* CNV. Left—a replication fork stalls and is re-started after reversal using DNA polymerase δ in a BIR-type mechanism. Middle—elongation is impaired on encounter with H3K56ac chromatin. Right—cleavage of replication fork yields a free end capable of strand invasion at a non-allelic locus, resulting in CNV. (**B**) Southern analysis of *CUP1* copy number in *3xCUP1* wild-type (wt), *rtt109Δ*, *pol32Δ* and *rtt109Δ pol32Δ* cells after 10 generations ±0.3 mM CuSO₄. Quantification shows the percentage of alleles deviating from the parental copy number, *P*-values calculated by one-way ANOVA. Copper adaptation was assessed by treated cells with varying concentrations of CuSO₄ and incubating for 3 days. Final OD_{600} was plotted against [CuSO₄] and copper tolerance quantified as area-under-curve for each culture; *P*-values were calculated by one-way ANOVA; $n = 4$. (**C**) Southern blot analysis of *CUP1* copy number in *3xCUP1* wild-type (wt), *pif1Δ*, *pol32Δ* and *pif1Δ pol32Δ* cells after 10 generations ±0.3 mM CuSO₄, analysed as in (B), $n = 6$. Asterisk indicates non-specific band in some non-copper treated samples migrating just above the 5x*CUP1* mark; this band is occasionally visible and likely represents a rare rearranged form of the locus. It does not bestow altered copper resistance and is never observed in the copper treated samples, and is therefore excluded from the quantification. (**D**) Southern blot analysis of *CUP1* copy number and adaptation test in *3xCUP1* wild-type (wt) versus *mph1Δ* cells after 10 generations ±0.3 mM CuSO₄, analysed as in (B), $n = 11$. Asterisk indicates non-specific band, see (C) for details.

by structure specific endonucleases (SSEs). In a nuclease-induced DSB model, resolution of BIR intermediates by an SSE results in a half-crossover product containing a free DNA end that could initiate strand invasion events at non-allelic sites (23,24).

Three SSE complexes with overlapping substrate specificities have been characterized in budding yeast—Mus81–Mms4, Yen1 and Slx1–Slx4 (reviewed in (92)). *mus81Δ* and *yen1Δ* mutations individually had no effect in the genetic screen and little effect on *CUP1* amplification in the 3x*CUP1* background (Figure 5A, lanes 3–6), however, the double mutant *mus81Δ yen1Δ* displayed a robust CNV suppression phenotype and reduced copper adaptation (Figure 5A, lanes 7–8). Additional deletion of *SLX4* did not further reduce *CUP1* amplification or adaptation in a *mus81Δ yen1Δ* background (Supplementary Figure S4A), so therefore Mus81 and Yen1 act redundantly to provide SSE activity required for efficient *CUP1* amplification.

Constitutive impairment of elongation in *pol32Δ* overcomes the CNV suppression caused by *rtt109Δ*, but SSE activity should be important for resolution irrespective of whether elongation is impeded by H3K56ac chromatin or lack of Pol32. As predicted, deletion of *POL32* did not restore *CUP1* CNV or adaptation in 3x*CUP1* *mus81Δ yen1Δ*, showing that SSE activity is an intrinsic and critical step in the mechanism (Figure 5B).

SSE-mediated resolution creates a free DNA end that is predicted to be highly recombinogenic. We expected that secondary strand invasion events instigated by the free DNA end would be mediated by Rad52 and Rad51, but *rad51Δ* had no effect in the genetic screen whereas *rad59Δ* abrogated CNV. We therefore analysed these mutants in the 3x*CUP1* system. In contrast to the genetic screen, which measures only contraction of the *CUP1* array, *CUP1* amplification and copper adaptation in 3x*CUP1* were strongly reduced in *rad51Δ* and almost absent in *rad52Δ* (Figure 5C, lanes 1–6), whereas CNV significantly increased in *rad59Δ* (Figure 5C lanes 7,8). This shows that *CUP1* copy number amplifications and contractions are mediated by separate mechanisms that act in competition: a Rad59-mediated pathway causes the majority of contraction events, as previously observed (93), but amplifications depend on Rad51, and some DNA ends that would be processed to contractions by Rad59 in the wild type are diverted to Rad51-mediated amplifications in *rad59Δ*.

These data are consistent with a model in which Mus81 or Yen1 cleave the replication fork after DNA polymerase δ is impeded, and the resulting free DNA end is resected then annealed by Rad59 or Rad51 to yield *CUP1* copy number contractions or amplifications respectively.

Replication fork progression is a critical mediator of *CUP1* CNV

Gene gating is common (78) and if each gene gating event caused sufficient replication stress to instigate error-prone BIR-type mechanisms then dramatic genome instability would ensue. This suggests that the *CUP1* locus is particularly prone to CNV, and to gain further insights into the properties of this locus we explored the potent CNV-stimulatory effect of loss of Mrc1, the strongest suppressor in the genetic screen (Figure 1F).

We first asked whether Mrc1 suppresses *CUP1* CNV by mediating the DNA replication checkpoint (94,95), however the absence of key replication checkpoint proteins Mec1 and Rad53 did not increase CNV or adaptation in 3x*CUP1* cells (Supplementary Figure S5A). This result was unexpected due to the involvement of Mec1 in both maintaining replication fork stability and releasing gated genes from the nuclear pore. Loss of Mec1 would therefore be predicted to increase CNV, but this may be compensated for by reduced activation in *mec1Δ* cells of the DNA repair pathways that generate CNV. Mrc1 also regulates replication fork speed in concert with Tof1 and Csm3 (96,97), and deletion of *TOF1* in 3x*CUP1* phenocopied *mrc1Δ* in significantly accelerated *CUP1* CNV, suggesting that replication fork speed is a major determinant of CNV rate (Figure 6A). Reducing replication fork speed would increase the chance that late replicating regions are not replicated prior to G2/M, requiring emergency repair by BIR-type mechanisms as at mammalian fragile sites (37) and promoting the use of inefficient late firing replication origins (98,99). The *CUP1* locus is located in a particularly late replicating region of Chromosome VIII and contains poorly characterized ARS elements (Supplementary Figure S5B and Figure 1A), so both delayed replication of the locus and activation of local replication origins are plausible contributors to CNV.

If either late replication or local origin activation is important, then altering the replication profile should impact CNV. We therefore deleted the gene encoding Clb5, a cyclin that regulates late firing replication origins; replication fork progression is normal in *clb5Δ* mutants (100) but loss of Clb5 prevents firing of late origins and S phase is extended to allow completion of replication (101,102). We observed that *CUP1* CNV decreased substantially in *clb5Δ* cells with a concurrent reduction in copper adaptation (Figure 6B), showing that Clb5 activity is important for *CUP1* CNV. To confirm this result we tested two other mutants, *sic1Δ* and *dia2Δ*, which cause similar shifts in replication and origin use to *clb5Δ* (98) albeit through a variety of mechanisms, and obtained the same result (Supplementary Figure S5C). Therefore, Clb5 activity is required for efficient *CUP1* CNV.

If *clb5Δ* suppresses CNV by extending S-phase such that the *CUP1* locus has more time to replicate, then slowing replication forks by removing Mrc1 or Tof1 should restore CNV. Unfortunately, the *mrc1Δ clb5Δ* double mutant was not viable in this background, however, deletion of *CLB5* in 3x*CUP1 tof1Δ* completely suppressed *CUP1* CNV and copper adaptation (Figure 6C). This is hard to reconcile with a model in which *CLB5* deletion simply extends replication timing to allow replication forks more time to finish DNA synthesis as we would not expect this to completely offset the replication fork slowing caused by *tof1Δ*, so other effects of *CLB5* deletion likely underlie the suppression of CNV.

Late firing replication origin activity promotes *CUP1* CNV

Since extension of S-phase in *clb5Δ* and *sic1Δ* mutants is attributed to suppression of late-firing replication origins
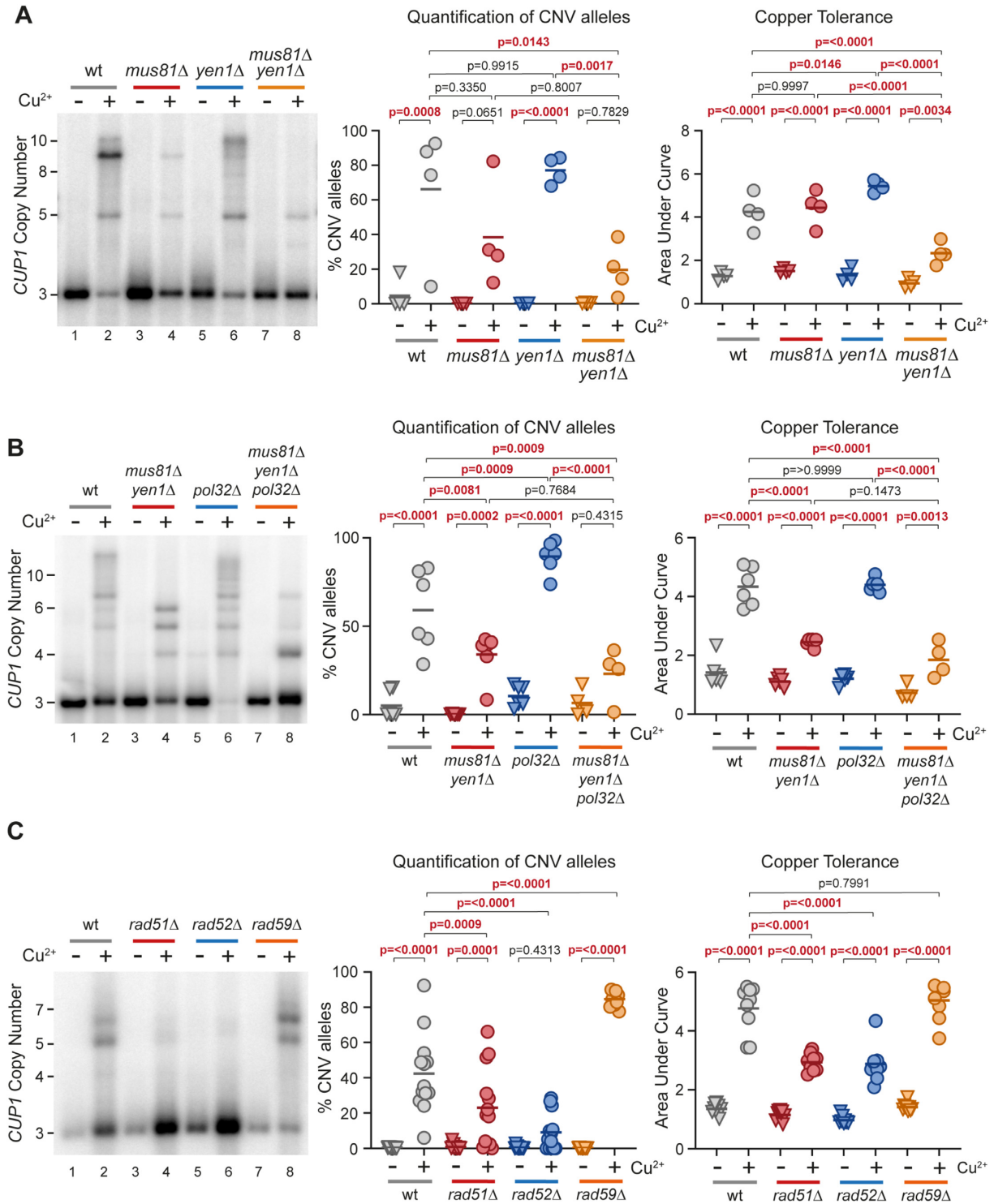
Figure 5. Importance of SSEs and Rad proteins in *CUP1* CNV. (**A**) Southern analysis of *CUP1* copy number in *3xCUP1* wild-type (wt), *mus81Δ*, *yen1Δ* and *mus81Δ yen1Δ* cells after 10 generations ±0.3 mM CuSO$_4$. Quantification shows the percentage of alleles deviating from the parental copy number, *P*-values calculated by one-way ANOVA. Copper adaptation was assessed by treated cells with varying concentrations of CuSO$_4$ and incubating for 3 days. Final OD$_{600}$ was plotted against [CuSO$_4$] and copper tolerance quantified as area-under-curve for each culture; *P*-values were calculated by one-way ANOVA; $n = 4$. (**B**) Southern blot analysis of *CUP1* copy number and adaptation test in *3xCUP1* wild-type (wt), *mus81Δ yen1Δ*, *pol32Δ* and *mus81Δ yen1Δ pol32Δ* cells after 10 generations ±0.3 mM CuSO$_4$, analysed as in (A), $n = 6$. (**C**) Southern blot analysis of *CUP1* copy number and adaptation test in *3xCUP1* wild-type (wt), *rad51Δ*, *rad52Δ* and *rad59Δ* cells after 10 generations ±0.3 mM CuSO$_4$, analysed as in (A), $n = 12$.
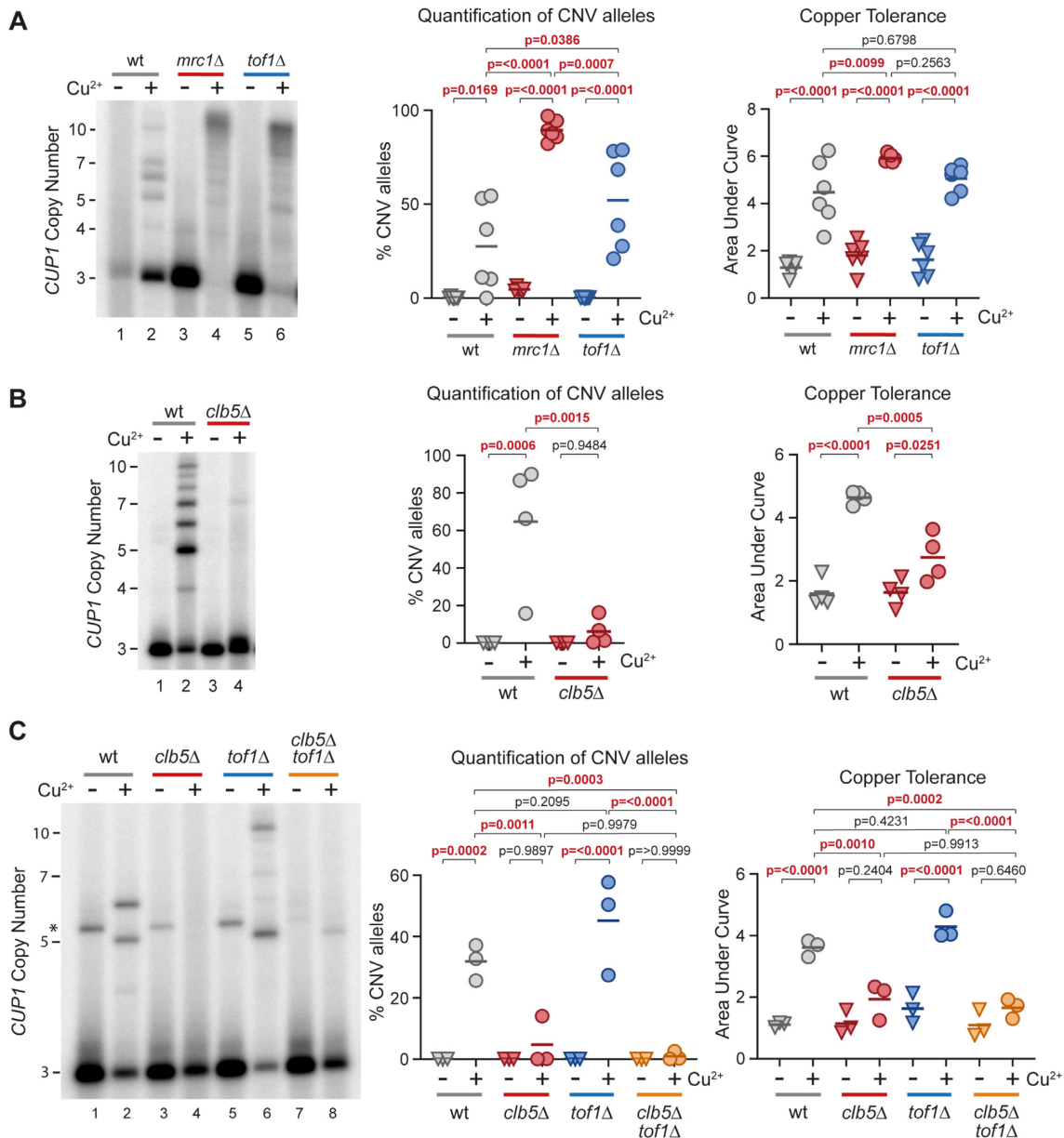
Figure 6. Effect of replication timing and replication fork progression on *CUP1* CNV. (**A**) Southern analysis of *CUP1* copy number in *3xCUP1* wild-type (wt), *mrc1Δ* and *tof1Δ* cells after 10 generations ±0.3 mM CuSO₄. Quantification shows the percentage of alleles deviating from the parental copy number, *P*-values calculated by one-way ANOVA. Copper adaptation was assessed by treating cells with varying concentrations of CuSO₄ for 3 days, final OD₆₀₀ was plotted against [CuSO₄] and copper tolerance quantified as area-under-curve for each culture; p-values were also calculated by one-way ANOVA; *n* = 6. (**B**) Southern blot analysis of *CUP1* copy number and adaptation test in *3xCUP1* wild type (wt) and *clb5Δ* mutants. Cells were cultured and CNV alleles and copper adaptation quantified as in (A); *n* = 4. (**C**) Southern blot analysis of *CUP1* copy number and adaptation test in *3xCUP1* wild type (wt), *clb5Δ*, *tof1Δ* and *clb5Δ tof1Δ* cells. Cells were cultured, CNV alleles and copper adaptation quantified as in (A); *n* = 3. Asterisk indicates non-specific band, see Figure 4C for details.

(98,102,103), and conversely deletion of *MRC1* increases usage of late firing origins (95,98), it is possible that local origin activity is the principle determinant of *CUP1* CNV, rather than replication timing or replication fork speed *per se*. Replication origins (ARS elements) are present in each of the *CUP1* repeats although whether these are active during S phase under normal conditions is unclear (Figure 1A).

Formation of a replication bubble at a replication origin creates two replication forks that move in opposite directions, and replication origins are therefore detectable as

sites of sharp changes in replication fork direction. Replication fork direction is revealed by the polarity of TrAEL-seq reads (64), and TrAEL-seq data from BY4741 wild-type cells with 13 copies of *CUP1* shows a change in polarity from negative to positive at the *CUP1* origin indicative of ARS activity (Figure 7A upper panel, highlighted in orange). This signal is weak compared to other ARS elements in the region and becomes almost undetectable in 3x*CUP1* cells (Supplementary Figure S6A, upper panel), consistent with the *CUP1* ARS elements being functional
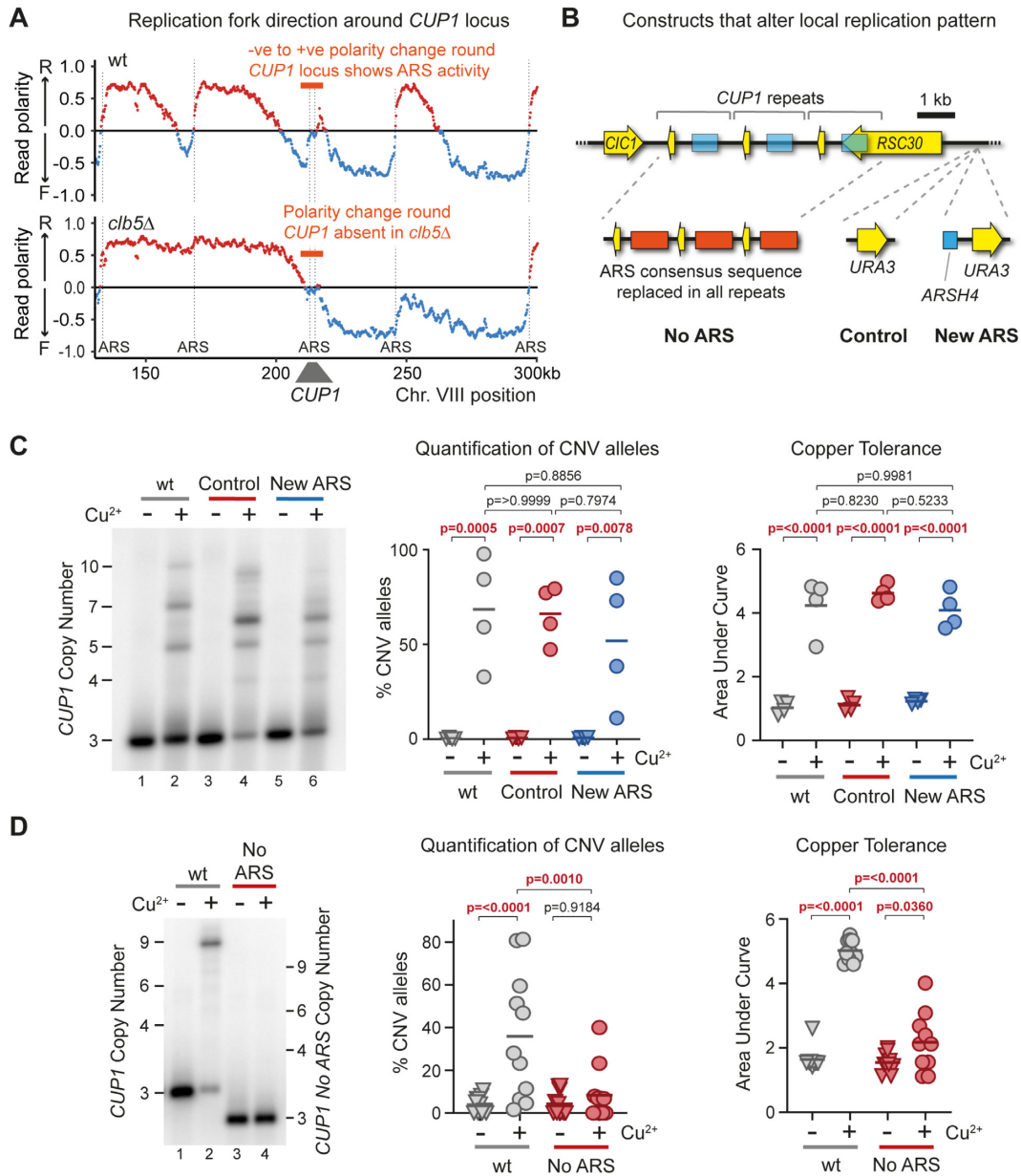
Figure 7. Local Replication origin firing regulates CNV at *CUP1* locus. (**A**) TrAEL-seq read polarity plots for wild type and *clb5Δ* cells grown to mid log in 2% Glucose. TrAEL-seq detects replication direction, an excess of reverse (R) reads at any site (indicated by positive polarity, red), results from replication forks moving left-to-right, the opposite for replication forks moving right-to-left (blue). Sharp transitions from −ve to +ve occur at active replication origins (ARS elements), gradual transitions from +ve to −ve are regions where forks converge. Plots show regions surrounding *CUP1* on chromosome VIII, quantified by (R-F)/(R + F) with R and F referring to Reverse and Forward reads respectively; $n = 2$, data from GSE154811. (**B**) Schematic of the three copy *CUP1* array and the modifications introduced to influence local replication pattern. In 'No ARS' cells, the region containing the endogenous ARS upstream of every *CUP1* repeat have been replaced by a non-expressed sequence derived from a *GFP* tagging plasmid. 'New ARS' and 'Control' cells have a *URA3* marker integrated upstream of the *RSC30* promoter, either with or without an efficient replication origin (*ARSH4*) respectively. TrAEL-seq data confirming activity of the new ARS is shown in Supplementary Figure S5A. (**C**) Southern blot analysis of *CUP1* copy number in *3xCUP1* wild-type (wt), Control and New ARS cells (described in B). Cells were cultured for 10 generations ±0.3 mM CuSO₄. Quantification shows the percentage of alleles deviating from the parental copy number, *P*-values calculated by one-way ANOVA. Copper adaptation was assessed by treating cells with varying concentrations of CuSO₄ and incubating for 3 days, final OD₆₀₀ was plotted against [CuSO₄] and copper tolerance quantified as area-under-curve for each culture; p-values were also calculated by one-way ANOVA; $n = 4$. (**D**) Southern blot analysis of *CUP1* copy number and growth curve analysis in *3xCUP1* wild-type (wt) and *3xCUP1* cells with the region containing the ARS sites replaced by unrelated sequence (described in B). Cells were cultured, CNV alleles and copper adaptation quantified as in (B), $n = 10$ for Southern blot analysis, $n = 11$ for growth curve analysis.

but rarely active. Very little change in replication fork polarity is detected at *CUP1* in *clb5Δ* cells with 13 copies of *CUP1*, placing the *CUP1* ARS elements amongst those origins dependent on Clb5 (Figure 7A lower panel), whereas TrAEL-seq of 3x*CUP1 mrc1Δ* reveals pronounced replication origin activity at *CUP1* that is not detected in the 3x*CUP1* wild type, consistent with increased usage of late-acting replication origins (Supplementary Figure S6A lower panel) (95,98).

The presence of Clb5-dependent ARS elements at *CUP1* is therefore consistent with the possibility that *CUP1* ARS activity promotes *CUP1* CNV. If so then loss of Clb5 would suppress CNV by preventing local ARS activity rather than by extending S phase. To distinguish these possibilities, we created origin deletion and insertion strains that either prevent *CUP1* ARS activity or force early replication (Figure 7B).

Firstly, to promote early replication of *CUP1* we integrated the efficient and early *ARSH4* origin adjacent to the *CUP1* locus to create a 'New ARS' strain, along with a control strain containing the selectable marker but no ARS element (Figure 7B). TrAEL-seq analysis confirmed that this origin is efficient and dominates the local replication profile (Supplementary Figure S6B). However, this had no detectable impact on copper adaptation or *CUP1* repeat amplification, and repeat amplification remained sensitive to *clb5Δ* (Figure 7C and Supplementary Figure S6C). Therefore, *CUP1* amplification does not depend on replication forks arriving at the locus late in S-phase from distant origins, and the extension of S-phase in *clb5Δ* cannot explain the strong suppression of CNV.

Secondly, to prevent replication origin firing at *CUP1* we mutated the ARS element in each *CUP1* repeat. As there are considerable discrepancies in the locations assigned to the *CUP1* ARS element (104,105) we deleted a substantial region of each *CUP1* repeat upstream of the *CUP1* promoter, replacing this with unrelated sequence derived from a *GFP* construct (Figure 7B, 'no ARS'). This change did not alter basal copper resistance compared to 3x*CUP1* cells containing wild type *CUP1* repeats (Supplementary Figure S6D, top panel), confirming that *CUP1* function was unaffected. However, when pre-exposed to copper, loss of the ARS suppressed *CUP1* CNV and copper adaptation to a similar extent as *CLB5* deletion (Figure 7D). This shows that activity of local ARS elements in the *CUP1* repeats, which is suppressed in *clb5Δ*, drives *CUP1* CNV. Interestingly, efficient *CUP1* amplification was restored in the No ARS strain when H3K56ac was rendered constitutive by removal of the H3K56 deacetylases Hst3 and Hst4, indicating that *CUP1* ARS activity contributes to CNV primarily through deposition of H3K56ac, although the growth defect caused by *hst3Δ hst4Δ* meant that we could not determine whether this also restored adaptation (Supplementary Figure S6E).

These experiments resolve the contribution of Clb5 and Mrc1 to the *CUP1* CNV phenotype, showing that activation of inefficient replication origins at *CUP1* rather than the late replication timing of the locus underlies the replication-dependence of transcriptionally stimulated *CUP1* amplification. Taken together with our previous re-

sults, this suggests that replication origin activation, in regions of high topological strain caused by promoter-nuclear pore interactions, leads to increased CNV through an inefficient BIR-type mechanism.

DISCUSSION

Here, we have dissected the mechanism by which transcriptional induction of the copper resistance gene *CUP1* stimulates CNV events that cause *CUP1* amplification and thereby increase copper resistance. We demonstrate critical roles for the TREX-2 and Mediator complexes that link transcribed loci to the nuclear pore, and for local replication origin activity, in addition to the known importance of H3K56ac.

A two-step mechanism for *CUP1* CNV

Given recent discoveries of replication fork instability resulting from collisions between replication forks and RNA polymerase II or R-loops, the results of our genetic screen for *CUP1* CNV modulators were not as we expected (44,45,54,55,58). The importance of TREX-2 and Mediator, deletions of which are even stronger suppressors of copper adaptation than *rad52Δ*, suggests that topological constraint of replication fork progression causes CNV, rather than direct interactions with the transcription unit, and this is coherent with the small but widespread increase in TrAEL-seq read density when the locus is transcriptionally active. Nonetheless, topological impairment of replication fork progression is known (56,57) and it is not too surprising that this increases the frequency of homologous recombination events that can result in CNV.

However, the requirement for H3K56ac needs to be explained (49). H3K56ac impairs DNA polymerase δ elongation during BIR (43), but given that H3K56ac should only be present on new histones that are deposited behind the replication fork, it is unclear why DNA polymerase δ encounters H3K56ac during replication fork restart as the histones ahead of the replication fork should not carry H3K56ac (106,107). One possibility is that abortive replication origin firing deposits H3K56ac: if both forks in the nascent replication bubble stall due to high topological strain then the replication bubble could dissolve by fork reversal and/or nascent strand degradation (illustrated in Supplementary Figure S7A). During the formation and initial progression of the bubble, histone exchange would leave an 'epigenetic scar' of H3K56ac chromatin in this non-replicated genomic region (also illustrated in Supplementary Figure S7A). We acknowledge that this would require a very pronounced fork reversal or degradation, but the restoration of adaptive *CUP1* CNV in the no ARS strain by deletion of *HST3* and *HST4*, which causes H3K56ac to be present throughout the genome (108), does suggest that the *CUP1* origins facilitate CNV through H3K56ac deposition.

The overall mechanism we propose for transcriptionally-stimulated *CUP1* CNV is illustrated in Figure 8. Firstly, replication is initiated from a replication origin in the *CUP1* region but rapidly stalls and the replication bubble collapses
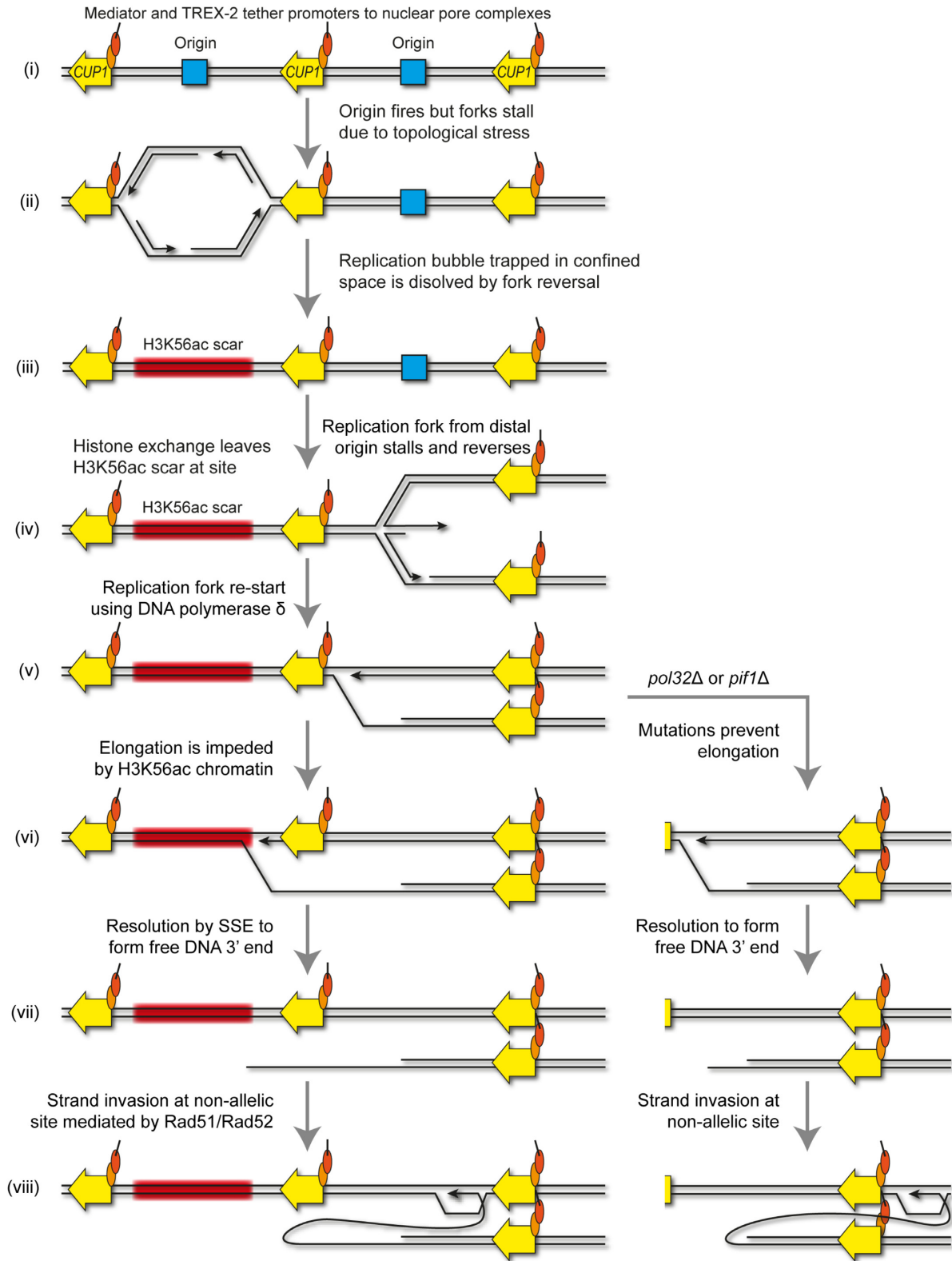
Figure 8. A model for origin-dependant stimulated-CNV of *CUP1* locus. DNA strands are shown in black, the *CUP1* gene shown in yellow, replication origins are shown in blue, chromatin containing H3K56ac shown in red, and mediator and TREX-2 complexes are shown as orange and red circles respectively. See Discussion for more details. Replication fork re-start is represented following the model described in (33), resulting in elongation by DNA polymerase δ without concurrent Okazaki fragment synthesis. The elongating re-started fork or the new fork formed by strand invasion will eventually be resolved by a replication fork converging from the left (not shown), allowing semi-conservative synthesis replication of the lagging strand by DNA polymerase δ (32). The slightly more complex outcome, Rad59-mediated resolution with a reversed converging fork, is illustrated in Supplementary Figure S7B.

due to replicative stress imposed by TREX-2 and Mediator, leaving a tract of H3K56ac chromatin in an un-replicated genomic region (Figure 8, steps i–iii). A second replication fork arriving from a distal site stalls due to local replicative stress and initiates fork re-start using DNA polymerase δ through fork reversal and strand invasion (Figure 8, steps iv–v), but elongation is impaired on encounter with the H3K56ac chromatin (Figure 8, step vi). Resolution by a SSE releases a free DNA end (Figure 8, step vii) that can undergo Rad51-mediated strand invasion with the sister chromatid to yield an amplification or a contraction if invasion occurs at a non-allelic site (Figure 8, step viii). If H3K56ac has not been deposited ahead of the replication fork that arrives in step iv (as in *rtt109Δ*), the stalling and re-start process would still occur, but the SSE resolution and non-allelic strand invasion steps (vi–viii) that result in CNV are avoided. Irrespective of any impediments to elongation and secondary strand invasion events, resolution by a converging replication fork (coming from the left in Figure 8) is finally required to complete chromosomal replication following step viii.

In *pol32Δ* and *pif1Δ* mutants with defective DNA polymerase δ elongation, SSE resolution must occur close to the site of re-start, but the outcome is otherwise the same as a free DNA end that can undergo non-allelic strand invasion is again formed (Figure 8 alternate pathway steps vi–viii). This event is however no longer dependent on H3K56ac, and is probably more efficient than in wild-type as it seems unlikely that H3K56ac inhibits DNA polymerase δ elongation as effectively as the absence of Pif1 or Pol32. It should be noted that this requirement for resolution by an oncoming fork, which is common in repair of stalled forks by BIR-type mechanisms (38–40), means that the BIR fork does not need to elongate any significant distance, and in *pol32Δ* or *pif1Δ* mutants the SSE resolution and strand invasion (steps vii and viii) would occur immediately following the initial strand invasion, and even if the intermediate in step viii is not elongated, resolution by an oncoming fork would still result in amplification.

An alternative pathway at step viii would be single strand annealing of the free DNA end with the oncoming fork, which results in a contraction (Supplementary Figure S7B). The high efficiency of single strand annealing compared to strand invasion probably explains the bias of *CUP1* CNV towards contractions that we have previously noted (49,93,109), but the Rad51-mediated pathway still yields sufficient amplification events to reproducibly emerge under the mild copper selection applied in the 3x*CUP1* adaptation system (Figure 1E). Therefore, transcriptionally stimulated CNV at the *CUP1* locus emerges through at least two competing pathways with partially overlapping outcomes. The mechanism by which CNV emerges, recombination resulting from elongation difficulties after replication fork re-start, is conceptually similar to the microhomology-mediated BIR (MMBIR) model, proposed to explain complex rearrangements in human cells (11). However, recombination at *CUP1* is strictly homology dependent rather than utilizing microhomology, and occurs in response to a defined chromatin signature that is at present unknown in MMBIR.

Replication stress as a driver of adaptation

Adaptive mutation at *CUP1* does not fit classical models of adaptation through selection of random pre-existing mutations, as *CUP1* CNV occurs in response to copper stress. However, neither does this constitute a stress-induced mutational pathway equivalent to the bacterial SOS response or recent reports of adaptive mutability during chemotherapy (110–113). Stress induced mutation pathways invoke a switch to the use of mutagenic repair pathways for repair of random DNA damage (reviewed in (114)), whereas the frequency and location of DNA damage is dramatically increased by the replication conflicts with a highly transcribed *CUP1* allele, driving CNV without necessitating a change in repair mechanism.

More generally, we suggest that DNA replication under environmental challenge is inherently mutagenic and prone to increase the genetic heterogeneity of a population from which adaptive mutations can be selected. Cells exposed to imperfect environments may be metabolically suboptimal and have reduced levels of dNTPs, ATP or histones needed for replication, may strongly induce stress responsive genes orientated head-on to replication forks (as in bacteria (115)) and may have increased levels of DNA lesions that impair replication fork progression. These would reduce fork progression and increase the frequency of local repair events, but also incur the widespread usage of inefficient origins that we find to be mutagenic at *CUP1* and increase the likelihood of fragile sites not completing replication at G2/M (37,116–121). Therefore, increased mutation rate does not need to be actively stimulated or initiated by cells under environmental challenge as long as some level of replication is maintained, since increased genetic heterogeneity should be an emergent property of replication under stress.

The question of when and how resistance mutations form is critical; as yeast pathogens become resistant to fungicides we will face major challenges in medicine and food production (reviewed in (122)). An example of the damage wrought by fungal pathogens is the progressive annihilation of amphibians by *Batrachochytrium dendrobatidis* (reviewed in (123)). Here, we provide proof of principle that the emergence of adaptive mutations is not inevitable as multiple genetic manipulations prevent budding yeast acquiring resistance to the agricultural fungicide copper sulfate. If adaptive mutations generally arise during fungicidal treatment in medicine or agriculture then a window of opportunity for avoiding resistance exists: by altering or supplementing treatment regimens to prevent replication or diminish replicative stress, the occurrence of *de novo* mutation should be reduced.

DATA AVAILABILITY

Sequencing data is available on GEO, accession number: GSE165163. Code for read processing is available on GitHub: https://github.com/FelixKrueger/TrAEL-seq and https://github.com/FelixKrueger/TrimGalore.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

REFERENCES

1. Katju,V. and Bergthorsson,U. (2013) Copy-number changes in evolution: rates, fitness effects and adaptive significance. *Front. Genet.*, **4**, 273.
2. Frei,E. 3rd, Rosowsky,A., Wright,J.E., Cucchi,C.A., Lippke,J.A., Ervin,T.J., Jolivet,J. and Haseltine,W.A. (1984) Development of methotrexate resistance in a human squamous cell carcinoma of the head and neck in culture. *Proc. Natl. Acad. Sci. U.S.A.*, **81**, 2873–2877.
3. Corcoran,R.B., Dias-Santagata,D., Bergethon,K., Iafrate,A.J., Settleman,J. and Engelman,J.A. (2010) BRAF gene amplification can promote acquired resistance to MEK inhibitors in cancer cells harboring the BRAF V600E mutation. *Sci. Signal*, **3**, ra84.
4. Huennekens,F.M. (1994) The methotrexate story: a paradigm for development of cancer chemotherapeutic agents. *Adv. Enzyme. Regul.*, **34**, 397–419.
5. Little,A.S., Balmanno,K., Sale,M.J., Newman,S., Dry,J.R., Hampson,M., Edwards,P.A., Smith,P.D. and Cook,S.J. (2011) Amplification of the driving oncogene, KRAS or BRAF, underpins acquired resistance to MEK1/2 inhibitors in colorectal cancer cells. *Sci. Signal*, **4**, ra17.
6. Sandegren,L. and Andersson,D.I. (2009) Bacterial gene amplification: implications for the evolution of antibiotic resistance. *Nat. Rev. Microbiol.*, **7**, 578–588.
7. Selmecki,A., Forche,A. and Berman,J. (2006) Aneuploidy and isochromosome formation in drug-resistant Candida albicans. *Science*, **313**, 367–370.
8. Lucas,E.R., Miles,A., Harding,N.J., Clarkson,C.S., Lawniczak,M.K.N., Kwiatkowski,D.P., Weetman,D., Donnelly,M.J. and Anopheles gambiae Genomes,C. (2019) Whole-genome sequencing reveals high complexity of copy number variation at insecticide resistance loci in malaria mosquitoes. *Genome Res.*, **29**, 1250–1261.
9. Jones,L., Riaz,S., Morales-Cruz,A., Amrine,K.C., McGuire,B., Gubler,W.D., Walker,M.A. and Cantu,D. (2014) Adaptive genomic structural variation in the grape powdery mildew pathogen, Erysiphe necator. *BMC Genomics*, **15**, 1081.
10. Zhang,F., Gu,W., Hurles,M.E. and Lupski,J.R. (2009) Copy number variation in human health, disease, and evolution. *Annu. Rev. Genomics Hum. Genet.*, **10**, 451–481.
11. Hastings,P.J., Ira,G. and Lupski,J.R. (2009) A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLos Genet.*, **5**, e1000327.
12. Gu,W., Zhang,F. and Lupski,J.R. (2008) Mechanisms for human genomic rearrangements. *Pathogenetics*, **1**, 4.
13. Liu,P., Carvalho,C.M., Hastings,P.J. and Lupski,J.R. (2012) Mechanisms for recurrent and complex human genomic rearrangements. *Curr. Opin. Genet. Dev.*, **22**, 211–220.
14. Hastings,P.J., Lupski,J.R., Rosenberg,S.M. and Ira,G. (2009) Mechanisms of change in gene copy number. *Nat. Rev. Genet.*, **10**, 551–564.
15. Zhang,F., Carvalho,C.M. and Lupski,J.R. (2009) Complex human chromosomal and genomic rearrangements. *Trends Genet.*, **25**, 298–307.
16. Mazzagatti,A., Shaikh,N., Bakker,B., Spierings,D.C.J., Wardenaar,R., Maniati,E., Wang,J., Boemo,M.A., Foijer,F. and McClelland,S.E. (2020) DNA replication stress generates distinctive landscapes of DNA copy number alterations and chromosome scale losses. bioRxiv doi: https://doi.org/10.1101/743658, 16 November 2020, preprint: not peer reviewed.
17. Arlt,M.F., Ozdemir,A.C., Birkeland,S.R., Wilson,T.E. and Glover,T.W. (2011) Hydroxyurea induces de novo copy number variants in human cells. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 17360–17365.
18. Arlt,M.F., Mulle,J.G., Schaibley,V.M., Ragland,R.L., Durkin,S.G., Warren,S.T. and Glover,T.W. (2009) Replication stress induces genome-wide copy number changes in human cells that resemble polymorphic and pathogenic variants. *Am. J. Hum. Genet.*, **84**, 339–350.
19. Petermann,E. and Helleday,T. (2010) Pathways of mammalian replication fork restart. *Nat. Rev. Mol. Cell Biol.*, **11**, 683–687.
20. Hanada,K., Budzowska,M., Davies,S.L., van Drunen,E., Onizawa,H., Beverloo,H.B., Maas,A., Essers,J., Hickson,I.D. and Kanaar,R. (2007) The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks. *Nat. Struct. Mol. Biol.*, **14**, 1096–1104.
21. Froget,B., Blaisonneau,J., Lambert,S. and Baldacci,G. (2008) Cleavage of stalled forks by fission yeast Mus81/Eme1 in absence of DNA replication checkpoint. *Mol. Biol. Cell*, **19**, 445–456.
22. Davis,A.P. and Symington,L.S. (2004) RAD51-dependent break-induced replication in yeast. *Mol. Cell. Biol.*, **24**, 2344–2351.
23. Deem,A., Barker,K., Vanhulle,K., Downing,B., Vayl,A. and Malkova,A. (2008) Defective break-induced replication leads to half-crossovers in Saccharomyces cerevisiae. *Genetics*, **179**, 1845–1860.
24. Smith,C.E., Lam,A.F. and Symington,L.S. (2009) Aberrant double-strand break repair resulting in half crossovers in mutants defective for Rad51 or the DNA polymerase delta complex. *Mol. Cell. Biol.*, **29**, 1432–1441.
25. Lydeard,J.R., Jain,S., Yamaguchi,M. and Haber,J.E. (2007) Break-induced replication and telomerase-independent telomere maintenance require Pol32. *Nature*, **448**, 820–823.
26. Wilson,M.A., Kwon,Y., Xu,Y., Chung,W.H., Chi,P., Niu,H., Mayle,R., Chen,X., Malkova,A., Sung,P. *et al.* (2013) Pif1 helicase and Poldelta promote recombination-coupled DNA synthesis via bubble migration. *Nature*, **502**, 393–396.
27. Saini,N., Ramakrishnan,S., Elango,R., Ayyar,S., Zhang,Y., Deem,A., Ira,G., Haber,J.E., Lobachev,K.S. and Malkova,A. (2013) Migrating bubble during break-induced replication drives conservative DNA synthesis. *Nature*, **502**, 389–392.
28. Donnianni,R.A. and Symington,L.S. (2013) Break-induced replication occurs by conservative DNA synthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 13475–13480.
29. Mizuno,K., Miyabe,I., Schalbetter,S.A., Carr,A.M. and Murray,J.M. (2013) Recombination-restarted replication makes inverted chromosome fusions at inverted repeats. *Nature*, **493**, 246–249.
30. Jalan,M., Oehler,J., Morrow,C.A., Osman,F. and Whitby,M.C. (2019) Factors affecting template switch recombination associated with restarted DNA replication. *Elife*, **8**, e41697.
31. Iraqui,I., Chekkal,Y., Jmari,N., Pietrobon,V., Freon,K., Costes,A. and Lambert,S.A. (2012) Recovery of arrested replication forks by homologous recombination is error-prone. *PLos Genet.*, **8**, e1002976.
32. Naiman,K., Campillo-Funollet,E., Watson,A.T., Budden,A., Miyabe,I. and Carr,A.M. (2021) Replication dynamics of recombination-dependent replication forks. *Nat. Commun.*, **12**, 923.
33. Miyabe,I., Mizuno,K., Keszthelyi,A., Daigaku,Y., Skouteri,M., Mohebi,S., Kunkel,T.A., Murray,J.M. and Carr,A.M. (2015)

Polymerase delta replicates both strands after homologous recombination-dependent fork restart. *Nat. Struct. Mol. Biol.*, **22**, 932–938.

34. Ait Saada,A., Lambert,S.A.E. and Carr,A.M. (2018) Preserving replication fork integrity and competence via the homologous recombination pathway. *DNA Repair (Amst.)*, **71**, 135–147.

35. Ait Saada,A., Teixeira-Silva,A., Iraqui,I., Costes,A., Hardy,J., Paoletti,G., Freon,K. and Lambert,S.A.E. (2017) Unprotected replication forks are converted into mitotic sister chromatid bridges. *Mol. Cell*, **66**, 398–410.

36. Bhowmick,R., Minocherhomji,S. and Hickson,I.D. (2016) RAD52 facilitates mitotic DNA synthesis following replication stress. *Mol. Cell*, **64**, 1117–1126.

37. Minocherhomji,S., Ying,S., Bjerregaard,V.A., Bursomanno,S., Aleliunaite,A., Wu,W., Mankouri,H.W., Shen,H., Liu,Y. and Hickson,I.D. (2015) Replication stress activates DNA repair synthesis in mitosis. *Nature*, **528**, 286–290.

38. Morrow,C.A., Nguyen,M.O., Fower,A., Wong,I.N., Osman,F., Bryer,C. and Whitby,M.C. (2017) Inter-fork strand annealing causes genomic deletions during the termination of DNA replication. *Elife*, **6**, e25490.

39. Pardo,B., Moriel-Carretero,M., Vicat,T., Aguilera,A. and Pasero,P. (2020) Homologous recombination and Mus81 promote replication completion in response to replication fork blockage. *EMBO Rep.*, **21**, e49367.

40. Mayle,R., Campbell,I.M., Beck,C.R., Yu,Y., Wilson,M., Shaw,C.A., Bjergbaek,L., Lupski,J.R. and Ira,G. (2015) DNA REPAIR. Mus81 and converging forks limit the mutagenicity of replication fork breakage. *Science*, **349**, 742–747.

41. Kramara,J., Osia,B. and Malkova,A. (2018) Break-induced replication: the where, the why, and the how. *Trends Genet.*, **34**, 518–531.

42. Liu,L., Yan,Z., Osia,B.A., Twarowski,J., Sun,L., Kramara,J., Lee,R.S., Kumar,S., Elango,R., Li,H. *et al.* (2021) Tracking break-induced replication shows that it stalls at roadblocks. *Nature*, **590**, 655–659.

43. Che,J., Smith,S., Kim,Y.J., Shim,E.Y., Myung,K. and Lee,S.E. (2015) Hyper-acetylation of histone H3K56 limits break-induced replication by inhibiting extensive repair synthesis. *PLos Genet.*, **11**, e1004990.

44. Lang,K.S., Hall,A.N., Merrikh,C.N., Ragheb,M., Tabakh,H., Pollock,A.J., Woodward,J.J., Dreifus,J.E. and Merrikh,H. (2017) Replication-transcription conflicts generate R-loops that orchestrate bacterial stress survival and pathogenesis. *Cell*, **170**, 787–799.

45. Hamperl,S., Bocek,M.J., Saldivar,J.C., Swigut,T. and Cimprich,K.A. (2017) Transcription-replication conflict orientation modulates R-loop levels and activates distinct DNA damage responses. *Cell*, **170**, 774–786.

46. Merrikh,H. (2017) Spatial and temporal control of evolution through replication-transcription conflicts. *Trends Microbiol.*, **25**, 515–521.

47. Zeman,M.K. and Cimprich,K.A. (2014) Causes and consequences of replication stress. *Nat. Cell Biol.*, **16**, 2–9.

48. Hull,R.M., King,M., Pizza,G., Krueger,F., Vergara,X. and Houseley,J. (2019) Transcription-induced formation of extrachromosomal DNA during yeast ageing. *PLoS Biol.*, **17**, e3000471.

49. Hull,R.M., Cruz,C., Jack,C.V. and Houseley,J. (2017) Environmental change drives accelerated adaptation through stimulated copy number variation. *PLoS Biol.*, **15**, e2001333.

50. Zhao,Y., Dominska,M., Petrova,A., Bagshaw,H., Kokoska,R.J. and Petes,T.D. (2017) Properties of mitotic and meiotic recombination in the tandemly-repeated CUP1 gene cluster in the yeast *Saccharomyces cerevisiae*. *Genetics*, **206**, 785–800.

51. Adamo,G.M., Lotti,M., Tamas,M.J. and Brocca,S. (2012) Amplification of the CUP1 gene is associated with evolution of copper tolerance in Saccharomyces cerevisiae. *Microbiology (Reading)*, **158**, 2325–2335.

52. Fogel,S., Welch,J.W., Cathala,G. and Karin,M. (1983) Gene amplification in yeast: CUP1 copy number regulates copper resistance. *Curr. Genet.*, **7**, 347–355.

53. Fogel,S. and Welch,J.W. (1982) Tandem gene amplification mediates copper resistance in yeast. *Proc. Natl. Acad. Sci. U.S.A.*, **79**, 5342–5346.

54. Garcia-Muse,T. and Aguilera,A. (2016) Transcription-replication conflicts: how they occur and how they are resolved. *Nat. Rev. Mol. Cell Biol.*, **17**, 553–563.

55. Helmrich,A., Ballarino,M., Nudler,E. and Tora,L. (2013) Transcription-replication encounters, consequences and genomic instability. *Nat. Struct. Mol. Biol.*, **20**, 412–418.

56. Bermejo,R., Capra,T., Jossen,R., Colosio,A., Frattini,C., Carotenuto,W., Cocito,A., Doksani,Y., Klein,H., Gomez-Gonzalez,B. *et al.* (2011) The replication checkpoint protects fork stability by releasing transcribed genes from nuclear pores. *Cell*, **146**, 233–246.

57. Bermejo,R., Lai,M.S. and Foiani,M. (2012) Preventing replication stress to maintain genome stability: resolving conflicts between replication and transcription. *Mol. Cell*, **45**, 710–718.

58. Li,X. and Manley,J.L. (2005) Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell*, **122**, 365–378.

59. Wang,G. and Vasquez,K.M. (2017) Effects of replication and transcription on DNA structure-related genetic instability. *Genes (Basel)*, **8**, 17.

60. Kim,N. and Jinks-Robertson,S. (2011) Guanine repeat-containing sequences confer transcription-dependent instability in an orientation-specific manner in yeast. *DNA Repair (Amst.)*, **10**, 953–960.

61. Cruz,C. and Houseley,J. (2014) Endogenous RNA interference is driven by copy number. *Elife*, **3**, e01581.

62. Szklarczyk,D., Gable,A.L., Lyon,D., Junge,A., Wyder,S., Huerta-Cepas,J., Simonovic,M., Doncheva,N.T., Morris,J.H., Bork,P. *et al.* (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.*, **47**, D607–D613.

63. Nepusz,T., Yu,H. and Paccanaro,A. (2012) Detecting overlapping protein complexes in protein-protein interaction networks. *Nat. Methods*, **9**, 471–472.

64. Kara,N., Krueger,F., Rugg-Gunn,P. and Houseley,J. (2020) Genome-wide analysis of DNA replication and DNA double strand breaks by TrAEL-seq. *PLoS Biol*, **19**, e3000886.

65. Robinson,M.D., McCarthy,D.J. and Smyth,G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.

66. Giaever,G., Chu,A.M., Ni,L., Connelly,C., Riles,L., Veronneau,S., Dow,S., Lucau-Danila,A., Anderson,K., Andre,B. *et al.* (2002) Functional profiling of the Saccharomyces cerevisiae genome. *Nature*, **418**, 387–391.

67. Chappidi,N., Nascakova,Z., Boleslavska,B., Zellweger,R., Isik,E., Andrs,M., Menon,S., Dobrovolna,J., Balbo Pogliano,C., Matos,J. *et al.* (2020) Fork cleavage-religation cycle and active transcription mediate replication restart after fork stalling at co-transcriptional R-loops. *Mol. Cell*, **77**, 528–541.

68. Maffia,A., Ranise,C. and Sabbioneda,S. (2020) From R-loops to G-quadruplexes: emerging new threats for the replication fork. *Int. J. Mol. Sci.*, **21**, 1506.

69. Huertas,P. and Aguilera,A. (2003) Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol. Cell*, **12**, 711–721.

70. Pfeiffer,V., Crittin,J., Grolimund,L. and Lingner,J. (2013) The THO complex component Thp2 counteracts telomeric R-loops and telomere shortening. *EMBO J.*, **32**, 2861–2871.

71. Wellinger,R.E., Prado,F. and Aguilera,A. (2006) Replication fork progression is impaired by transcription in hyperrecombinant yeast cells lacking a functional THO complex. *Mol. Cell. Biol.*, **26**, 3327–3334.

72. Hartono,S.R., Malapert,A., Legros,P., Bernard,P., Chedin,F. and Vanoosthuyse,V. (2018) The affinity of the S9.6 antibody for double-stranded RNAs impacts the accurate mapping of R-loops in fission yeast. *J. Mol. Biol.*, **430**, 272–284.

73. Lockhart,A., Pires,V.B., Bento,F., Kellner,V., Luke-Glaser,S., Yakoub,G., Ulrich,H.D. and Luke,B. (2019) RNase H1 and H2 are differentially regulated to process RNA-DNA hybrids. *Cell Rep.*, **29**, 2890–2900.

74. Chang,E.Y., Tsai,S., Aristizabal,M.J., Wells,J.P., Coulombe,Y., Busatto,F.F., Chan,Y.A., Kumar,A., Dan Zhu,Y., Wang,A.Y. *et al.* (2019) MRE11-RAD50-NBS1 promotes Fanconi Anemia R-loop

suppression at transcription-replication conflicts. *Nat. Commun.*, **10**, 4265.

75. Ellisdon,A.M., Dimitrova,L., Hurt,E. and Stewart,M. (2012) Structural basis for the assembly and nucleic acid binding of the TREX-2 transcription-export complex. *Nat. Struct. Mol. Biol.*, **19**, 328–336.

76. Gonzalez-Aguilera,C., Tous,C., Gomez-Gonzalez,B., Huertas,P., Luna,R. and Aguilera,A. (2008) The THP1-SAC3-SUS1-CDC31 complex works in transcription elongation-mRNA export preventing RNA-mediated genome instability. *Mol. Biol. Cell*, **19**, 4310–4318.

77. Schneider,M., Hellerschmied,D., Schubert,T., Amlacher,S., Vinayachandran,V., Reja,R., Pugh,B.F., Clausen,T. and Kohler,A. (2015) The nuclear pore-associated TREX-2 complex employs mediator to regulate gene expression. *Cell*, **162**, 1016–1028.

78. Santos-Pereira,J.M., Garcia-Rubio,M.L., Gonzalez-Aguilera,C., Luna,R. and Aguilera,A. (2014) A genome-wide function of THSC/TREX-2 at active genes prevents transcription-replication collisions. *Nucleic. Acids. Res.*, **42**, 12000–12014.

79. Verger,A., Monte,D. and Villeret,V. (2019) Twenty years of Mediator complex structural studies. *Biochem. Soc. Trans.*, **47**, 399–410.

80. Lariviere,L., Plaschka,C., Seizl,M., Petrotchenko,E.V., Wenzeck,L., Borchers,C.H. and Cramer,P. (2013) Model of the Mediator middle module based on protein cross-linking. *Nucleic. Acids. Res.*, **41**, 9266–9273.

81. Guglielmi,B., van Berkum,N.L., Klapholz,B., Bijma,T., Boube,M., Boschiero,C., Bourbon,H.M., Holstege,F.C. and Werner,M. (2004) A high resolution protein interaction map of the yeast Mediator complex. *Nucleic. Acids. Res.*, **32**, 5379–5391.

82. Lariviere,L., Geiger,S., Hoeppner,S., Rother,S., Strasser,K. and Cramer,P. (2006) Structure and TBP binding of the Mediator head subcomplex Med8-Med18-Med20. *Nat. Struct. Mol. Biol.*, **13**, 895–901.

83. Schilbach,S., Hantsche,M., Tegunov,D., Dienemann,C., Wigge,C., Urlaub,H. and Cramer,P. (2017) Structures of transcription pre-initiation complex with TFIIH and Mediator. *Nature*, **551**, 204–209.

84. Casamassimi,A. and Napoli,C. (2007) Mediator complexes and eukaryotic transcription regulation: an overview. *Biochimie*, **89**, 1439–1446.

85. Lei,E.P., Stern,C.A., Fahrenkrog,B., Krebber,H., Moy,T.I., Aebi,U. and Silver,P.A. (2003) Sac3 is an mRNA export factor that localizes to cytoplasmic fibrils of nuclear pore complex. *Mol. Biol. Cell*, **14**, 836–847.

86. Dieppois,G. and Stutz,F. (2010) Connecting the transcription site to the nuclear pore: a multi-tether process that regulates gene expression. *J. Cell Sci.*, **123**, 1989–1999.

87. Ahmed,S. and Brickner,J.H. (2007) Regulation and epigenetic control of transcription at the nuclear periphery. *Trends Genet.*, **23**, 396–402.

88. Blobel,G. (1985) Gene gating: a hypothesis. *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 8527–8529.

89. Liang,Q. and Zhou,B. (2007) Copper and manganese induce yeast apoptosis via different pathways. *Mol. Biol. Cell*, **18**, 4741–4749.

90. Stafa,A., Donnianni,R.A., Timashev,L.A., Lam,A.F. and Symington,L.S. (2014) Template switching during break-induced replication is promoted by the Mph1 helicase in Saccharomyces cerevisiae. *Genetics*, **196**, 1017–1028.

91. Pham,N., Yan,Z., Yu,Y., Faria Afreen,M., Malkova,A., Haber,J.E. and Ira,G. (2021) Mechanisms restraining break-induced replication at two-ended DNA double-strand breaks. *EMBO J.*, **40**, e104847.

92. Rass,U. (2013) Resolving branched DNA intermediates with structure-specific nucleases during replication in eukaryotes. *Chromosoma*, **122**, 499–515.

93. Doi,G., Okada,S., Yasukawa,T., Sugiyama,Y., Bala,S., Miyazaki,S., Kang,D. and Ito,T. (2021) Catalytically inactive Cas9 impairs DNA replication fork progression to induce focal genomic instability. *Nucleic. Acids. Res.*, **49**, 954–968.

94. Osborn,A.J. and Elledge,S.J. (2003) Mrc1 is a replication fork component whose phosphorylation in response to DNA replication stress activates Rad53. *Genes Dev.*, **17**, 1755–1767.

95. Alcasabas,A.A., Osborn,A.J., Bachant,J., Hu,F., Werler,P.J., Bousset,K., Furuya,K., Diffley,J.F., Carr,A.M. and Elledge,S.J. (2001) Mrc1 transduces signals of DNA replication stress to activate Rad53. *Nat. Cell Biol.*, **3**, 958–965.

96. Somyajit,K., Gupta,R., Sedlackova,H., Neelsen,K.J., Ochs,F., Rask,M.B., Choudhary,C. and Lukas,J. (2017) Redox-sensitive alteration of replisome architecture safeguards genome integrity. *Science*, **358**, 797–802.

97. Yeeles,J.T.P., Janska,A., Early,A. and Diffley,J.F.X. (2017) How the eukaryotic replisome achieves rapid and efficient DNA replication. *Mol. Cell*, **65**, 105–116.

98. Koren,A., Soifer,I. and Barkai,N. (2010) MRC1-dependent scaling of the budding yeast DNA replication timing program. *Genome Res.*, **20**, 781–790.

99. Gispan,A., Carmi,M. and Barkai,N. (2014) Checkpoint-independent scaling of the *Saccharomyces cerevisiae* DNA replication program. *BMC Biol.*, **12**, 79.

100. McCune,H.J., Danielson,L.S., Alvino,G.M., Collingwood,D., Delrow,J.J., Fangman,W.L., Brewer,B.J. and Raghuraman,M.K. (2008) The temporal program of chromosome replication: genomewide replication in clb5{Delta} *Saccharomyces cerevisiae*. *Genetics*, **180**, 1833–1847.

101. Schwob,E. and Nasmyth,K. (1993) CLB5 and CLB6, a new pair of B cyclins involved in DNA replication in *Saccharomyces cerevisiae*. *Genes Dev.*, **7**, 1160–1175.

102. Donaldson,A.D., Raghuraman,M.K., Friedman,K.L., Cross,F.R., Brewer,B.J. and Fangman,W.L. (1998) CLB5-dependent activation of late replication origins in *S. cerevisiae*. *Mol. Cell*, **2**, 173–182.

103. Lengronne,A. and Schwob,E. (2002) The yeast CDK inhibitor Sic1 prevents genomic instability by promoting replication origin licensing in late G(1). *Mol. Cell*, **9**, 1067–1078.

104. Liachko,I., Bhaskar,A., Lee,C., Chung,S.C., Tye,B.K. and Keich,U. (2010) A comprehensive genome-wide map of autonomously replicating sequences in a naive genome. *PLos Genet.*, **6**, e1000946.

105. Wyrick,J.J., Aparicio,J.G., Chen,T., Barnett,J.D., Jennings,E.G., Young,R.A., Bell,S.P. and Aparicio,O.M. (2001) Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins. *Science*, **294**, 2357–2360.

106. Li,Q., Zhou,H., Wurtele,H., Davies,B., Horazdovsky,B., Verreault,A. and Zhang,Z. (2008) Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell*, **134**, 244–255.

107. Masumoto,H., Hawke,D., Kobayashi,R. and Verreault,A. (2005) A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature*, **436**, 294–298.

108. Celic,I., Masumoto,H., Griffith,W.P., Meluh,P., Cotter,R.J., Boeke,J.D. and Verreault,A. (2006) The sirtuins hst3 and Hst4p preserve genome integrity by controlling histone h3 lysine 56 deacetylation. *Curr. Biol.*, **16**, 1280–1289.

109. Fishman-Lobell,J., Rudin,N. and Haber,J.E. (1992) Two alternative pathways of double-strand break repair that are kinetically separable and independently modulated. *Mol. Cell. Biol.*, **12**, 1292–1303.

110. Cipponi,A., Goode,D.L., Bedo,J., McCabe,M.J., Pajic,M., Croucher,D.R., Rajal,A.G., Junankar,S.R., Saunders,D.N., Lobachevsky,P. *et al.* (2020) MTOR signaling orchestrates stress-induced mutagenesis, facilitating adaptive evolution in cancer. *Science*, **368**, 1127–1131.

111. Russo,M., Crisafulli,G., Sogari,A., Reilly,N.M., Arena,S., Lamba,S., Bartolini,A., Amodio,V., Magri,A., Novara,L. *et al.* (2019) Adaptive mutability of colorectal cancers in response to targeted therapies. *Science*, **366**, 1473–1480.

112. Foster,P.L. (2007) Stress-induced mutagenesis in bacteria. *Crit. Rev. Biochem. Mol. Biol.*, **42**, 373–397.

113. Rosenberg,S.M., Shee,C., Frisch,R.L. and Hastings,P.J. (2012) Stress-induced mutation via DNA breaks in *Escherichia coli*: a molecular mechanism with implications for evolution and medicine. *Bioessays*, **34**, 885–892.

114. Fitzgerald,D.M., Hastings,P.J. and Rosenberg,S.M. (2017) Stress-induced mutagenesis: implications in cancer and drug resistance. *Annu. Rev. Cancer Biol.*, **1**, 119–140.

115. Merrikh,C.N. and Merrikh,H. (2018) Gene inversion potentiates bacterial evolvability and virulence. *Nat. Commun.*, **9**, 4662.

116. Macheret,M., Bhowmick,R., Sobkowiak,K., Padayachy,L., Mailler,J., Hickson,I.D. and Halazonetis,T.D. (2020) High-resolution mapping of mitotic DNA synthesis regions and

common fragile sites in the human genome through direct sequencing. *Cell Res.*, **30**, 997–1008.

117. Sarni,D., Sasaki,T., Miron,K., Tur-Sinai,M.I., Rivera-Mulia,J.C., Magnuson,B., Lungman,M., Gilbert,D.M. and Kerem,B. (2019) Replication timing and transcription identifies a novel fragility signature under replication stress. bioRxiv doi: https://doi.org/10.1101/716951, 27 July 2019, preprint: not peer reviewed.

118. Brison,O., El-Hilali,S., Azar,D., Koundrioukoff,S., Schmidt,M., Nahse,V., Jaszczyszyn,Y., Lachages,A.M., Dutrillaux,B., Thermes,C. *et al.* (2019) Transcription-mediated organization of the replication initiation program across large genes sets common fragile sites genome-wide. *Nat. Commun.*, **10**, 5693.

119. Debatisse,M., Le Tallec,B., Letessier,A., Dutrillaux,B. and Brison,O. (2012) Common fragile sites: mechanisms of instability revisited. *Trends Genet.*, **28**, 22–32.

120. Durkin,S.G. and Glover,T.W. (2007) Chromosome fragile sites. *Annu. Rev. Genet.*, **41**, 169–192.

121. Le Beau,M.M., Rassool,F.V., Neilly,M.E., Espinosa,R. 3rd, Glover,T.W., Smith,D.I. and McKeithan,T.W. (1998) Replication of a common fragile site, FRA3B, occurs late in S phase and is delayed further upon induction: implications for the mechanism of fragile site induction. *Hum. Mol. Genet.*, **7**, 755–761.

122. Fisher,M.C., Hawkins,N.J., Sanglard,D. and Gurr,S.J. (2018) Worldwide emergence of resistance to antifungal drugs challenges human health and food security. *Science*, **360**, 739–742.

123. Fisher,M.C., Garner,T.W. and Walker,S.F. (2009) Global emergence of Batrachochytrium dendrobatidis and amphibian chytridiomycosis in space, time, and host. *Annu. Rev. Microbiol.*, **63**, 291–310.

124. Xu,W., Aparicio,J.G., Aparicio,O.M. and Tavare,S. (2006) Genome-wide mapping of ORC and Mcm2p binding sites on tiling arrays and identification of essential ARS consensus sequences in S. cerevisiae. *BMC Genomics*, **7**, 276.