

Age-specific genomic and transcriptomic variation reveals limited evidence for cis-regulatory interactions modulating aging in *Saccharomyces cerevisiae*

Kaitlin M. McHugh ,* Felipe S. Barreto , Molly K. Burke *

Department of Integrative Biology, Oregon State University, 2701 SW Campus Way, Corvallis, OR 97331, United States

*Corresponding authors: Kaitlin M. McHugh, Department of Integrative Biology, Oregon State University, 2701 SW Campus Way, Corvallis, OR 97331, United States. Email: kaitlin.m.mchugh@gmail.com; Molly K. Burke, Department of Integrative Biology, Oregon State University, 2701 SW Campus Way, Corvallis, OR 97331, United States. Email: Molly.Burke@oregonstate.edu

The budding yeast *Saccharomyces cerevisiae* is a well-established model for studying the genetic basis of complex traits, and it is a powerful system for investigating mechanisms of aging. Here, we examine the genomic and transcriptomic factors contributing to increased replicative age in recombinant yeast populations harboring standing genetic variation. Using fluorescence-activated cell sorting, we isolated young and aged cohort pairs across 12 biological replicates and sequenced their progeny to assess patterns of differentiation at the nucleotide and transcription levels. Most differentiated alleles were located in coding regions, including significant variants within 132 unique genes. Transcriptomic analysis revealed 60 differentially expressed genes in aged populations, including 18 genes with increased expression in aged cohorts, and 42 genes with decreased expression. Although only 2 genes (*RFA3* and *WSC4*) were implicated in both genomic and transcriptomic analyses, functional overlap associated with protein homeostasis, DNA repair, and cell cycle regulation was evident across datasets. Notably, we found no strong evidence that differentially expressed genes were more likely to occur in close proximity to significant gene variants. This suggests that late-life survival is not predominantly governed by background-independent local cis-regulatory interactions. These findings are consistent with a polygenic architecture in which distinct genotypes may promote longevity through diverse regulatory mechanisms.

Keywords: GWAS; lifespan; gene expression; gene regulation; evolutionary genetics; yeast

Introduction

Aging is a nearly universal process, yet the mechanisms that regulate its onset and progression remain incompletely understood. Identifying the genes and pathways that determine the timing and severity of aging offers the possibility of predicting or even delaying its physiological consequences. While many evolutionarily conserved regulators of aging have been described (López-Otín et al. 2013, 2023; Lemoine 2021), less is known about how sequence-level variation alters gene expression to shape aging. Deeper insight into this process could shed light on how natural genetic variation contributes to the evolution of aging phenotypes.

Gene expression is governed by both cis- regulatory and trans-regulatory elements. Cis- regulatory elements (eg promoters, enhancers, silencers) typically control the activity of linked genes through direct physical interactions, whereas trans- regulatory elements (eg transcription factors, regulatory RNAs) regulate unlinked genes via diffusion-mediated interactions. Cis-regulatory changes are of particular evolutionary interest because they are thought to contribute disproportionately to adaptive divergence (Wittkopp and Kalay 2011). Accordingly, dissecting the relative contributions of cis- and trans-acting factors provides a powerful

framework for evaluating how genetic variation shapes aging-related phenotypes, by way of gene expression.

In theory, integrating genome sequence data with gene expression profiles of individuals from phenotypically diverse populations offers a means to distinguish cis- from trans-acting regulatory effects (Li and Ritchie 2021). Given the scale necessary to achieve this goal of integrating data across phenotypic, genomic, and transcriptomic levels, laboratory studies with model organisms are especially useful. Studies with this objective have successfully identified regulatory variants underlying complex traits such as thermotolerance (Lecheta et al. 2020) and metabolic responses (Cubillos et al. 2017; Franco-Duarte et al. 2017), but applications to aging remain comparatively rare.

The yeast *Saccharomyces cerevisiae* has long been recognized as a powerful model system for aging research. Yeast share many conserved pathways with higher eukaryotes, and hundreds of genes have interchangeable human orthologs (Kachroo et al. 2015; Skrzypek et al. 2018). Replicative lifespan, defined as the total number of asexual divisions a cell undergoes, serves as a widely used proxy for aging in dividing cells and parallels aspects of human cellular senescence (Janssens and Veenhoff 2016). The replicative age of yeast mother cells can be measured via bud-scar staining, and fluorescence-activated cell sorting (FACS) enables

separation of young and old cohorts for molecular profiling (Chen and Contreras 2004; Hendrickson et al. 2018). These features, combined with extensive genomic resources, make *S. cerevisiae* an especially tractable system for dissecting the genetic regulation of aging.

Here we leverage an outbred yeast population (“4S”; Phillips et al. 2021) generated from repeated outcrossing of four diverse ancestral strains. This population maintains extensive heterozygosity and standing genetic variation, providing a unique opportunity to evaluate how *cis*- and *trans*-regulatory elements contribute to expression differences associated with replicative aging. By stratifying cells by age and combining pooled-population sequencing with transcriptomic profiling, we identify candidate regulatory variants that underlie differential expression in aged cohorts. Our analyses reveal both known and novel genetic contributors to longevity-associated pathways, including genes involved in genome stability, protein homeostasis, epigenetic modifications, and nutrient signaling. However, we find limited evidence that background-independent *cis*-regulatory effects—in other words, effects that act consistently across genotypes—are a primary driver of late-life survival in this population. This suggests that longevity is shaped by complex, non-local regulatory interactions.

Materials and methods

Creation of experimental replicates

All experimental populations used in this study were derived from the “4S” yeast population described by Phillips et al. 2021. Briefly, this population was created by combining 4 haploid strains from the barcoded SGRP collection (YPS128, Y12, DBVPG6765, and DBVPG6044; Cubillos et al. 2009). Linder et al. 2020 modified these strains to enable easy crossing and diploid recovery (*MATa* strains have genotype *ycr043C::NatMX*, *ura3::KanMX-barcode*, *ho Δ*, and *MATα* strains have genotype *ycr043C::HygMX*, *ura3::KanMX-barcode*, *ho Δ*). After initial crossing, the 4S base population was subsequently subjected to 12 iterations of further outcrossing to maximize genetic diversity and allow for some domestication to laboratory conditions (Burke et al. 2020; Phillips et al. 2021).

The 4S population was revived from -80°C storage and 1 mL was plated as a lawn of cells on YPD media (1% yeast extract, 2% peptone, 2% dextrose) containing both 300 $\mu\text{g/mL}$ hygromycin B (HYG) and 100 $\mu\text{g/mL}$ nourseothricin sulfate (NTC), to ensure that only diploid cells (containing both mating types and both resistance cassettes) grew. The lawn of cells was broadly sampled with a wooden applicator and the resulting cellular material was used to inoculate 10 mL of YPD; this process was repeated to create 12 biological replicates (Fig. 1a). Each replicate was grown for 9 d at $30^{\circ}\text{C}/200\text{ rpm}$ to generate a distribution of cells with a range of replicative ages (Fig. 1b). This length of time was extended from the time period used for initial bud scar quantification (see Supplementary Methods) to allow more aged cells to accumulate in the population.

Cell sorting and generation of paired age-cohorts

Prior to the experiment, we performed fluorescence imaging on sorted cells to confirm that the sorting process effectively separates aged cells from young cells (see Supplementary Methods, Supplementary Figs. 1 and 2, Supplementary Table 1). After 9 d of aging, each replicate was standardized to a concentration of approximately 10^7 cells/mL using an optical density measurement at 600 nm (an OD_{600} of 1.0 ± 0.1). Cell bud scars within each replicate were fluorescently stained with wheat-germ agglutinin Alexa

fluor 488 (WGA CF488A; Biotium) per manufacturer instructions and sorted using a SONY SH800 cell sorter (Fig. 1c). A “main gate” was created to filter out particles of unusual size (eg large clumps of cells) that appeared as outliers based on a visual inspection of the scattered light signals during sorting setup; this gate typically retained approximately 80% to 95% of particles (Supplementary Fig. 1). This gate was applied to both young and aged replicates. In addition to the main gate, the “aged” cohort for each replicate consisted of cells sorted from within the top 85% to 100% of fluorescence values while the “young” cohort consisted of cells sorted from within the bottom 50% of fluorescence values (to amplify differences between young and old cohorts, the gates used here are more stringent than those used in the Supplementary Methods). Sorting of each replicate was stopped once it reached a count of 250,000 cells.

The 24 sorted fractions (12 pairs of young and aged replicate-cohorts) were transferred to 10 mL of YPD media and allowed to reach mid-log-phase growth (Fig. 1d) before being archived to reduce potential noise from transient patterns of gene expression. This step was implemented to capture heritable, persistent patterns of gene expression rather than growth-phase-specific patterns, or transcriptional responses triggered by the sorting protocol itself (see Discussion). All samples were collected for RNA sequencing when OD_{600} values ranged between 4 and 7, as growth-curve data suggest these values best predict mid-log phase. Approximately 1 million cells were collected from each replicate-cohort and immediately spun down. They were then preserved in 800 μL TRIzol (Invitrogen), which killed cells while preserving the RNA present in the sample to capture gene expression at the time of preservation. Each replicate preserved in TRIzol was then homogenized by beadbeating for 30 s prior to storage at -80°C . Samples for DNA sequencing were collected immediately following sample collection for RNA; 1 mL of culture (~ 15 – 30 million cells) was transferred directly to a cryotube and immediately frozen at -80°C (Fig. 1e).

Genomic sequencing and analysis

Frozen samples were thawed and prepped using the Gentra Puregene Yeast/Bact Kit (Qiagen) following the manufacturer’s protocol. Genomic DNA libraries were prepared using the Illumina Nextera Library Prep Kit (Illumina), with some modifications to the protocol to increase throughput (Baym et al. 2015). All samples were dual-indexed and pooled for sequencing, with pool sizes expected to exceed sequencing coverage by 2 or more orders of magnitude. A size-selection step for 350 to 550 bp was performed using a Qiagen MinElute Gel Extraction kit. The genomic DNA library was sequenced with the NextSeq 2000 (P2, PE150 reads) at the Center for Quantitative Life Sciences (CQLS) at Oregon State University.

We used a GATK-based variant calling pipeline (Phillips et al. 2021) to produce a merged VCF file containing all single-nucleotide polymorphisms (SNPs) and short insertion-deletion polymorphisms (indels) identified across all replicates and cohorts. Raw fastq files were aligned to the *S. cerevisiae* S288C reference genome (R64-2-1) using GATK v4.0 with standard best practices workflows and default filtering options (Van der Auwera and O’Connor 2020). The merged VCF file was converted into tables of variant frequencies by extracting the allele depth (AD) and unfiltered depth (DP) fields for all variants passing quality filters. AD was used as the alternate allele count at a presumed variant, and the DP was used as the total coverage at that site. The VCF file was also used as an input for SnpEff v4.3 (Cingolani et al. 2012) to extract potential functional effects of individual SNPs

(dataset deposited in Dryad; pipeline information available as described in Data Availability Statement). This variant frequency table was further filtered for minimum coverage values of 20 and maximum coverage of 1,000 at every site in each replicate-cohort, and a minor allele frequency of at least 5% across the dataset (scripts available on GitHub, as described in Data Availability Statement). These filtering parameters meet or exceed recommendations by Guo et al. (2013) to reduce error rates in pooled-sequencing data. Differences in allele frequencies between the young and aged pairs of each replicate were evaluated via Cochran–Mantel–Haenszel (CMH) tests using the pool-seq package in R (Taus et al. 2017). The resulting P-values were log-transformed and a Bonferroni multiple testing correction ($\alpha=0.05$) was used to generate a genome-wide significance threshold (scripts on GitHub). The SNPs and indels exceeding this threshold make up our “candidate genomic variant list.” This list was used to identify potentially enriched Genome Ontology (GO) terms through the Saccharomyces Genome Database Gene-Ontology Term Finder (Ashburner et al. 2000; The Gene Ontology Consortium et al. 2023) with an FDR correction.

Using genome sequences from the four founder strains (Wing et al. 2020), we were able to estimate the most likely founder genotype at each position along the genome given our observed SNP frequencies (script available on GitHub). SNPs within a 5-kb sliding window (using a 2-kb step size) were used to estimate haplotype frequencies and the average divergence (“D”) between founder haplotypes across the genome (as described in Burke et al. 2014a, 2014b). The D statistic is the mean absolute difference in allele frequency between a haplotype’s frequency in the young cohort and its frequency in the aged cohort, averaged across all replicates. These values were plotted against expected (normally distributed) values to identify regions of unusually high divergence (script available on GitHub). We assessed signatures of selection (rather than genetic drift) by comparing the genome-wide SNP/indel and haplotype data, evaluating whether regions with elevated differentiation between young and aged replicates overlapped with regions of ancestral haplotype divergence consistently across replicates.

Transcriptomic sequencing and differential gene expression analysis

Direct-zol RNA Miniprep Kits (Zymo Research) were used per manufacturer instructions to isolate total RNA. Each sample was treated with TURBO DNase enzyme (Invitrogen) and further purified and concentrated using the Clean and Concentrator kit (Zymo Research). Samples were quantified with a Qubit fluorometer (3.0) and integrity assessed with an Agilent Bioanalyzer 2100 prior to library preparation using QuantSeq 3’ mRNA-Seq FWD kit (Lexogen, Vienna, Austria). The 24 barcoded libraries were sequenced as single-read 100-bp in an Illumina NextSeq2000 P1 flow cell at the CQLS.

We filtered the RNA-seq data using the recommendations provided by Lexogen QuantSeq for 3’ mRNA-seq (described in detail in code/transcriptome/README.md on GitHub). The bbduk tool was used to trim adapter sequences, polyA sequences, and bases with Phred scores below 20; we also removed reads shorter than 25 nucleotides. We used FastQC to check the quality of the filtered reads. After filtering, we performed genome alignment (*S. cerevisiae* S288C reference genome; R64-2-1) using the STAR alignment tool (Dobin et al. 2013), and used featureCounts (Liao et al. 2014) for quantification. The combined count matrix was filtered to remove genes with fewer than 5 mapped reads across all replicates

and genes that have no expression (zero reads mapped) in 13 or more replicates. This retained 5,620 genes for analysis (dataset archived in Dryad; normalized count data is also available on Dryad). We used Combat-seq (Zhang et al. 2020) to correct for a possible batch effect affecting pairs 1 to 6 and pairs 7 to 12, as these 2 groups were cultured and sorted on different days (Supplementary Figs. 3 and 4, script available on GitHub). We then performed differential expression analysis using DESeq2 (Love et al. 2014) in R. We compare expression (E) at each gene between aged and young cohorts for the i^{th} pair (P) within the j^{th} treatment (T), where the value of T_j is 0 for young and 1 for aged (Equation 4.1). P-values were corrected for false-discovery rate (FDR) and significance level assessed at $\text{FDR} \leq 0.1$ (script available on GitHub).

$$E_{ij} = \mu + \beta_1 P_i + \beta_2 T_j + \epsilon_{ij} \quad (4.1)$$

Integrating genomic and transcriptomic analyses

Candidate gene lists generated from genome and transcriptome analyses were compared to determine if any genes were identified by both sequencing methods, as such overlap might implicate a cis-regulatory mechanism. We use 3 methods to identify potential cis-regulatory interactions using these 2 datasets: (i) we identified variants from our candidate genomic variant list located within 10 kb of differentially expressed transcripts, (ii) we analyzed predicted functional annotations of our genomic candidate variant list using SNPeff, and (iii) we performed a GO-term analysis on the genes that were implicated in both the genomic and transcriptomic candidate variant lists.

We first investigated gene variants that were near differentially expressed genes in our transcriptomic candidate variant list, and thus may be candidates for local cis-regulatory interactions. Local regulatory variation is commonly identified within promoter regions and the 3’ untranslated region (Ronald et al. 2005), and 10 kb is often used as a threshold for detecting cis-regulatory interactions (eg Leach et al. 2007). The *S. cerevisiae* genome has documented high recombination rates, and linkage disequilibrium has been shown to decline sharply in the strains most closely related to our ancestral haplotypes over this distance (Schacherer et al. 2009). Furthermore, SNPeff uses 5 kb as a cutoff value for upstream and downstream gene variants (which should include most promoters and 3’ UTRs). However, many studies of cis-regulatory interactions set even stricter limits (eg 1 kb; Renganaath et al. 2020; Shih and Fay 2021). Thus, we used 10, 5, and 1 kb as progressively conservative thresholds to detect evidence of potential local cis-regulatory interactions (all scripts for this analysis are available on GitHub). This was done simply by mapping our differentially expressed genes to the genome using the Saccharomyces Genome Database (Wong et al. 2023) and the UCSC Genome Browser (Perez et al. 2025), and finding each gene variant located within 10 kb upstream or downstream of that gene in our candidate genomic variant list. This list also included variants located within differentially expressed genes.

Second, we looked at the predicted functional effects of significantly differentiated SNPs, comparing these to the predicted functional effects of all SNPs across the genome. We extracted the first annotation from SNPeff (sorted by deleteriousness) for each SNP that exceeded the significance threshold. We further grouped the SNP annotations so that any annotation category with a count of 1 or fewer in our candidate genetic variant list was grouped as “other” for both our candidate SNP list and the whole-genome

annotation data, leaving us with 5 functional annotation categories (downstream gene variants, upstream gene variants, nonsynonymous variants, synonymous variants, and “other.” A Fisher’s exact test was used to assess whether the proportion of SNPs assigned to each annotation category in our significant data differed from all annotated SNPs in the genome (all scripts required for this analysis are available in GitHub).

Finally, genes that were implicated in both the genomic and transcriptomic candidate gene lists were used for an additional GO-term enrichment analyses through the Saccharomyces Genome Database Gene-Ontology Term Finder ($P < 0.1$ with an FDR correction).

Results

We used a paired-cohort design to identify genomic and transcriptomic correlates of replicative aging in a genetically diverse yeast population. Our approach assumes that bud scar fluorescence is a reliable proxy for replicative age (supported by [Supplementary Figs. 1 and 2](#)), and that genotypes conferring greater longevity become enriched among high-fluorescence cells over the 9-d incubation period. Under these assumptions, consistent allele frequency and expression differences between aged and young cohorts across 12 independent replicates reflect heritable contributions to late-life survival.

Allele and haplotype frequency differentiation

We identified 67,049 biallelic SNPs that met our filtering criteria, of which >98% remained segregating within each individual replicate ([Supplementary Fig. 5](#)). Complete SNP and indel results have been deposited in the Dryad repository. Genome-wide coverage across all replicates averaged 117x (range: 40x–277x; [Supplementary Table 2](#), [Supplementary Fig. 6](#)). Additionally, we identified 5,129 insertions/deletions (indels) that met filtering criteria with an average genome-wide coverage of 116x (range: 36x–278x; [Supplementary Table 2](#); deposited in Dryad). After applying genome-wide thresholds, we identified a total of 669 SNPs and 132 indels with significant difference in frequency between conditions ([Fig. 2a](#), [Supplementary Table 3](#), [Supplementary Fig. 7](#)). The list of 801 significant SNPs and indels were located within 233 unique genomic features, including 132 protein-coding genes. Together, we refer to these as our “candidate genomic variant list” ([Supplementary Table 3](#)). We utilized SNPeff to find the gene feature most closely associated with each SNP. Two of the indels were not annotated within the SNPeff tool, but were noted as being within an autonomously replicating sequence (ARS319) in the UCSC genome browser. The resulting list of gene features was used for GO-term enrichment analysis, which detected functions associated with glycoprotein metabolism and the cell periphery ([Fig. 3](#), [Supplementary Table 4](#)).

Haplotype-level analyses revealed distinct ancestral backgrounds associated with regions of the genome harboring significant SNPs ([Fig. 2b](#)). Visual comparisons of [Fig. 2a and 2b](#) led to the designation of 3 main peaks of interest, on chromosomes 5, 10, and 11. In addition to surpassing the threshold in [Fig. 2a](#), each of these peaks shows a single ancestral haplotype rising to prominence ([Fig. 2b](#)) and the average percent divergence between all haplotypes is greater than 6% ([Supplementary Fig. 8](#)). Notably, each of these peaks originated from a different ancestral haplotype; the DBVPG6765 European wine strain was elevated in the chromosome 5 peak, the YPS128 North American wild strain was isolated in the chromosome 10 peak, and the Y12 Japanese sake strain was elevated in the chromosome 11 peak

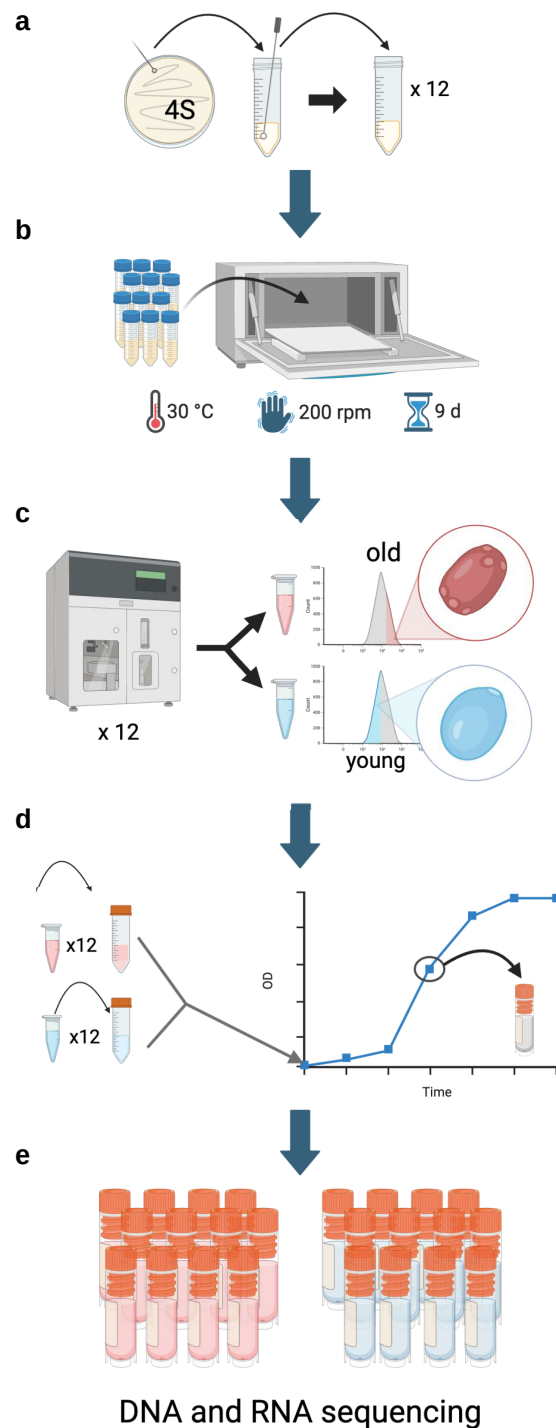


Fig. 1. Experimental design: a) Replicates were created by broadly sampling the 4S lawn to inoculate 12 liquid cultures. b) Each replicate was aged for 9 days at 30°C and 200 rpm. c) Each replicate was fluorescently stained and sorted into 2 cohorts of aged and young cells. d) Each cohort was grown to mid-log phase and e) preserved for sequencing. Created in BioRender. McHugh, K. (2026) <https://BioRender.com/0nkbpu7>.

([Supplementary Table 5](#)). Additionally, several potentially interesting features were identified under each significant peak. Chromosome 5 included variants within *MNN1*, *YND1*, *FMP52*, *NOP16*. In chromosome 10, the most significant genic variants were found within *IMA5*, *MNN5*, *SWI3*, *BBC1*, and *RFA3*. The most significant genic variants within chromosome 11 included

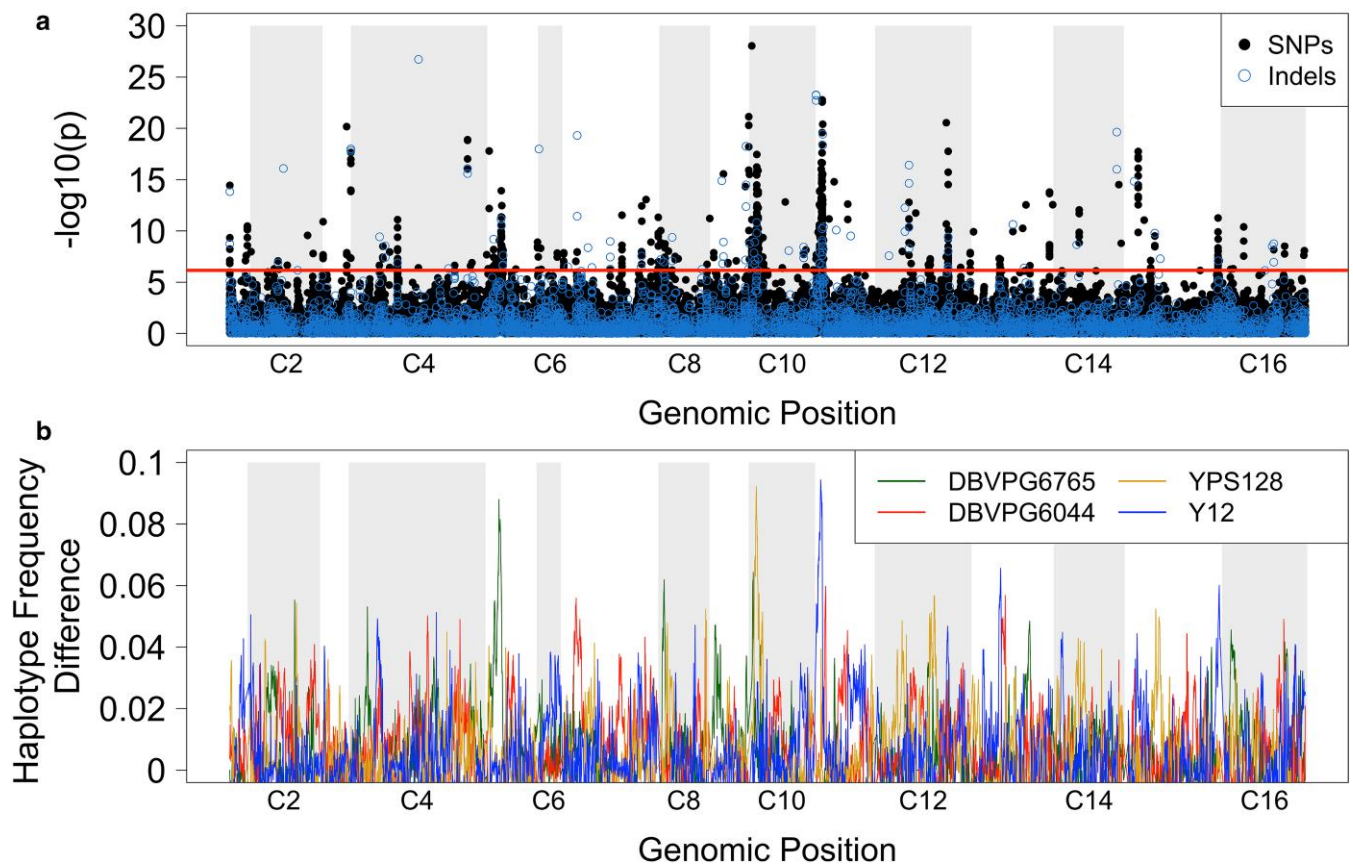


Fig. 2. Gene variant frequency differentiation across the genome. a) Manhattan plot from a paired CMH-test of SNP and indel frequencies across the genome reveals regions of significant differentiation. The red line indicates a Bonferroni corrected threshold of $\alpha = 0.05$ differentiating the most significant variants. b) A sliding-window analysis of variation in ancestral haplotype frequencies between young and aged cohorts shows 3 primary regions of the genome in which a single ancestral haplotype is favored.

PTK1, MNN4, TOR2, and YKL202W. Peaks in each of these 3 chromosomes also contained a number of non-genic significant variants found within 5 kb of an annotated gene.

Differential gene expression between aged and young cohorts

RNA-seq libraries yielded an average of 21.8 million reads after quality trimming, and 16.5 million counted per sample after mapping (Supplementary Table 6). Mapping rates ranged from 57.3 to 80.1%.

The gene matrix generated from these reads initially included 6,607 genes, but further filtering to remove rare transcripts resulted in 5,620 genes for analysis (complete results deposited in Dryad). A total of 60 differentially expressed genes between the young and aged cohorts were identified at an FDR < 0.1. We used a more liberal FDR threshold (as used in other exploratory genome-wide transcriptomic studies, e.g. Zhang et al. 2021, Li et al. 2024) to more thoroughly investigate potential cis-regulatory candidate genes that may have small effect sizes. Recognizing that this also increases the risk of identifying false positives, we refer to these as our “transcriptomic candidate list” and report individual P-values for each in Supplementary Table 7. We found 42 of these genes were downregulated in the aged populations, while 18 were upregulated (Fig. 4, Supplementary Fig. 4b, Supplementary Table 7).

Within our transcriptomic candidate list, YDR543C had the greatest $\log_2$ fold change (approximately 125% reduction in expression in aged cohorts); however, YDR543C also had the

lowest normalized expression count than any other feature at 20.9; by comparison, the mean expression value was 9,154.16 and the median was 1,180.323. The second highest change in expression was FIT2, with approximately 101% average decrease in expression in the aged cohorts compared to the young. SPG5 and YNR066C also showed large reductions in expression, 81 and 79%, respectively. WSC4 has the largest increase in expression in aged cohorts, at 41%. YRO2 had the next highest increase in expression in the aged cohorts, also nearing 41%, followed by NDT80 and SPS4 at 40 and 38% (Fig. 4, Supplementary Fig. 4b).

No GO terms were found to be significantly overrepresented within our transcriptomic candidate gene list, so we examined the GO terms directly to characterize functional patterns. We found a few functions shared among sets of differentially expressed genes. Within the 18 upregulated genes, GO terms were commonly associated with DNA replication and/or repair (POL12, CLB6, CDC45), transcription (GCR1, NDT80, OPI1), translation (DIA4, MRP1, SUI2), and cell cycle regulation (CLB6, NDT80, SPS4, TOS2). DNA repair (ACT1, MAG1, RFA3, UBC13) and cell cycle regulation (ACT1, CWP1, IRC18, RFA3) were also associated with several downregulated genes. Other terms linked to downregulated genes included protein modification, targeting, folding, and breakdown (EMA19, SSA1, UBC13, VID24, CPR6, RFA3, RPN6), lipid metabolism (CYB5, ERG25, IPT1, TMA7), ion transport and homeostasis (ATX1, FIT2, ATX2, VMA10), and cell wall organization and biogenesis (ACT1, CCW12, CWP1, IRC18).

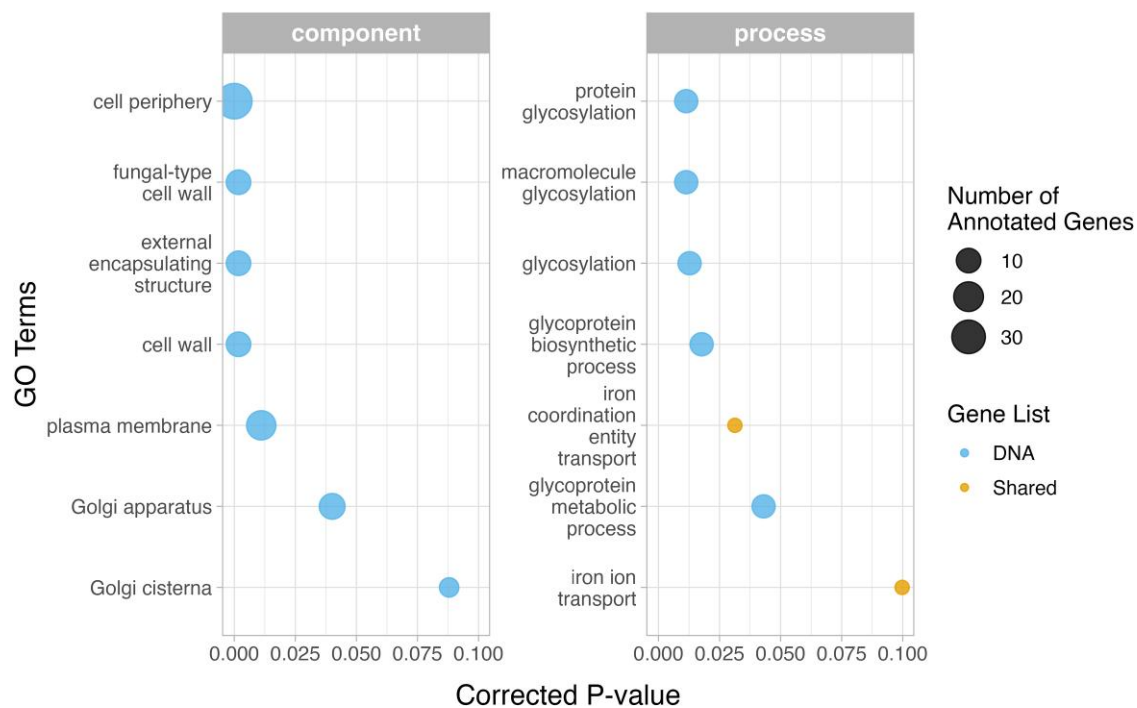


Fig. 3. Results of GO-term analysis for the genomic and transcriptomic gene lists. GO-term analyses were conducted separately for the genomic variant list, and the shared list of genes found in both genomic and transcriptomic variant lists. No significant terms were found for the transcriptomic variant list alone.

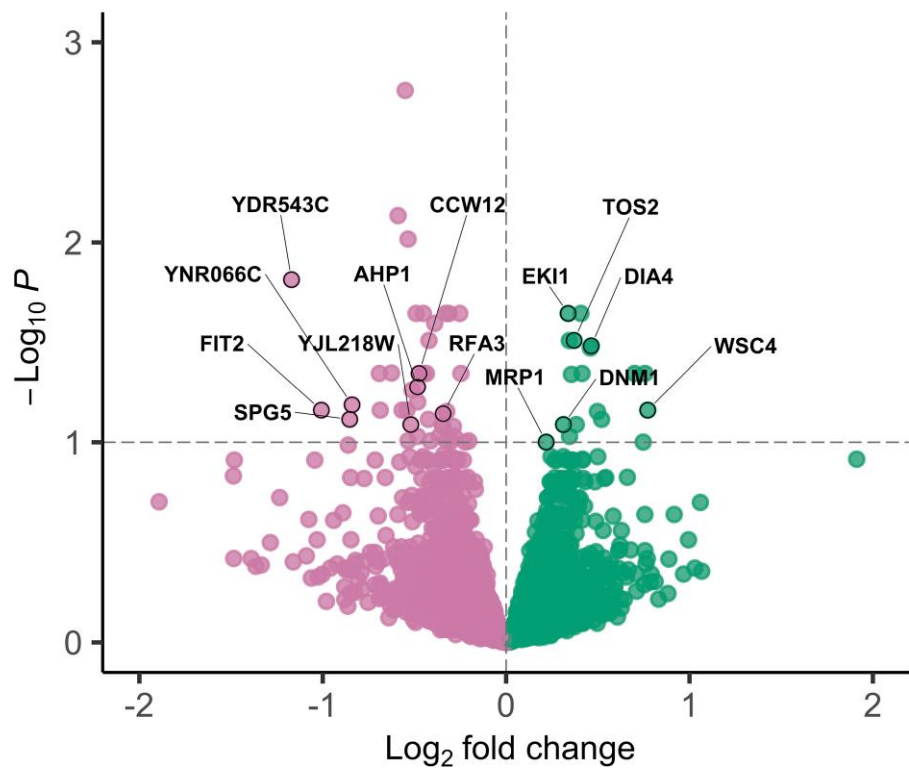


Fig. 4. Volcano plot of differentially expressed genes. The horizontal dashed line represents significance threshold (adjusted P -value < 0.1). Green points represent genes that are upregulated in the aged cohorts ($\log_2$ fold change > 0), and pink points represent genes that are downregulated in the aged cohorts ($\log_2$ fold change < 0). Genes labelled in the plot ($n = 14$) were identified as candidates in both the genomic and transcriptomic analyses (Table 1).

Comparing genomic and transcriptomic analyses

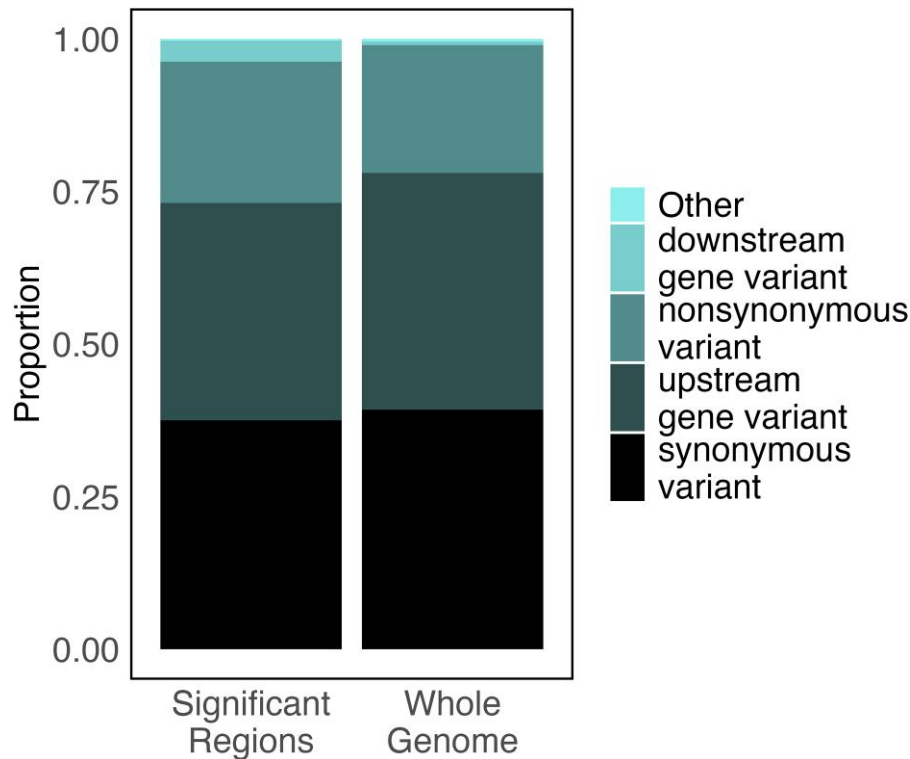
First, we identified 124 significant genetic variants that were located within 10 kb of a differentially expressed gene; these local

variants were distributed among 14 of the 60 differentially expressed genes from our transcriptomic candidate gene list. The other 46 genes were not associated with any significant local

Table 1. Differentially expressed genes with significant gene variants located within 10 kb of the gene.

Gene variants	Variant types (variants within 10 kb)					
	Variants within gene	Variants within 10 kb	Nonsynonymous SNPs	Synonymous SNPs	Intergenic SNPs	Indels
Gene						
AHP1	0	10	6	0	0	4
CCW12	0	10	6	0	0	4
DIA4	0	3	0	0	3	0
DNM1	0	1	0	0	1	0
EKI1	0	2	0	1	0	1
FIT2	0	14	4	3	6	1
MRP1	0	1	0	1	0	0
RFA3	4	39	10	17	12	0
SPG5	0	1	0	0	1	0
TOS2	0	1	0	0	1	0
WSC4	2	14	4	10	0	0
YDR543C	0	1	0	0	1	0
YJL218W	0	26	7	18	1	0
YNR066C	0	1	0	1	0	0
Totals	6	124	37	51	26	10

Many upstream and downstream gene variants associated with these genes are also located within the coding regions of other genes, resulting in them being annotated as nonsynonymous and synonymous variants, rather than intergenic SNPs. However, 6 of these variants are located within the transcript of the differentially expressed gene: RFA3 contains 2 nonsynonymous and 2 synonymous variants, while WSC4 contains 2 synonymous variants.

**Fig. 5.** Downstream gene variants and nonsynonymous SNPs are overrepresented among our significant SNPs compared to upstream gene variants and synonymous variants. (Fisher's test, $P < 0.001$).

variants. Of the 14 genes with local genomic variants, 8 have variants within 5 kb, and 4 within 1 kb. Only 2 of these genes included variants within the gene transcript itself. Using SNPeff, we investigated the predicted functional implications of each of these variants (Table 1). While most of these variants were synonymous or intergenic, there were a substantial proportion of nonsynonymous variants (including 2 of the variants within the RFA3 transcript).

The second method we used to search for potential cis-regulatory interactions was to investigate the representation of

variant identities within the significant SNP data (eg nonsynonymous variants, synonymous variants, upstream gene variants, downstream gene variants) within our candidate genomic variant list and the annotated genome as a whole. These results should be interpreted with caution, since many of the annotations returned by SNPeff are loosely defined. Still, using a simulated Fisher's exact test, we find that the distribution of annotation categories in our candidate list is statistically significant compared to annotations across the whole genome ($P < 0.001$). Directly comparing the proportional representation of categories between our

candidate gene list and the whole genome suggests that upstream gene variants and synonymous gene variants were underrepresented in our candidate gene list, while downstream gene variants and nonsynonymous variants were overrepresented compared to the whole genome (Fig. 5).

Finally, we performed a GO-term analysis on the initial list of 14 genes that were candidates for local cis-regulatory interactions ($P < 0.1$, FDR correction). We returned 2 terms, iron ion transport and iron coordination entity transport (both corresponding to the genes *FIT2* and *DNM1*).

Discussion

We used bud scar staining and FACS to separate twelve paired replicates of aged and young populations for genomic and transcriptomic comparisons. To detect heritable age-associated variation, we sequenced sorted replicates after a short period of growth rather than immediately after sorting to standardize for growth stage. While our sorting approach achieved modest differences in replicative age phenotypes, genomic and transcriptomic data show clear differentiation between treatments. Haplotype analyses support that this differentiation is driven by selection rather than drift, as highly differentiated variants consistently map to specific founder haplotypes across independent replicates.

Genomic differentiation implicates established and emergent aging pathways

Studies across diverse yeast strains have revealed consistent molecular hallmarks of cellular aging (reviewed by Janssens and Veenhoff 2016). Using existing genomic resources (SNPeff, SGD, UCSC Genome Browser), we found that many of our candidate genes fall within these well-established pathways, including protein homeostasis (eg *MNN* genes, *VID24*), nutrient sensing (eg *TOR2*), and DNA repair (eg *RFA3*).

Genes involved in protein synthesis, folding, targeting, and degradation were overrepresented in our dataset. A particularly striking pattern is the broad enrichment of glycosylation-related genes (eg *MNN* family genes, *PSA1*) alongside reduced expression of genes encoding abundant cell wall glycoproteins (*CCW12* and *CWP1*) in aged cohorts. Although cell wall staining could potentially bias detection toward cell wall-related processes, concordance between genomic and expression changes and prior literature suggest a biological basis. For example, *CCW12* and *CWP1* are both mannoproteins involved in new cell wall growth at bud sites (Smits et al. 2006; Ragni et al. 2011), and have shown binding affinity for WGA (Kung et al. 2009). If cells were upregulating WGA-binding proteins to escape the sorting protocol, we might expect both *CCW12* and *CWP1* show increased expression in our aged cohorts; instead, we find reduced expression in both genes. *CCW12* in particular makes up over 40% of cell wall glycoproteins and is essential for maintaining cell structure and morphogenesis (Ghanegolmohammadi et al. 2021), so its reduced expression implicates age-related shifts in cell wall composition that may impair structural integrity. While direct analogs are lacking in metazoans without cell walls, altered glycan composition and glycosylation patterns have also been associated with aging in mammals (Krištić et al. 2014; Dall'Olio 2018; Cindrić et al. 2021; Wu et al. 2025), supporting the evolutionary conservation of glycan metabolism as an aging modulator. The mechanistic links between the glycome and aging remain poorly understood (Miura and Endo 2016; Palmer 2022), but our results reinforce its relevance.

Consistent with prior work linking vacuolar dysfunction to declines in proteostasis and lifespan (Hughes and Gottschling 2012; Ghavidel et al. 2018; Chen et al. 2020), we also observed differentiation in genes involved in protein degradation (*VID24*, *VMA10*). *WSC4* (also called *YHC8*), one of only 2 genes implicated in both genomic and transcriptomic analyses, encodes a stress-response factor believed to influence protein targeting to the endoplasmic reticulum (Mamoun et al. 1999). These results support the hypothesis that maintenance of proteome integrity, via mechanisms mitigating the accumulation of misfolded or aggregated proteins, is a target of selection for increased replicative age (Aguilaniu et al. 2003; Saarikangas and Barral 2015).

Genes involved in DNA replication and repair also showed consistent differentiation. Upregulated genes in the aged cohorts included those essential for DNA replication initiation (*POL12*, *CLB6*, *CDC45*) and genotoxic stress response (*OPI1*, *CDC45*, *SPS4*), whereas downregulated genes included those involved in DNA repair (*MAG1*, *UBC13*, *RFA3*). *RFA3*, which encodes a subunit of Replication Protein A involved in both repair and telomere maintenance, is the second of the 2 genes implicated in both genomic and transcriptomic analyses. While previous work found *RFA3* upregulated in old cells (Lesur and Campbell 2004), we found it to be downregulated. This discrepancy likely reflects distinct dynamics in a recombinant genetic background where polygenic variation and interactions can emerge.

Metal ion homeostasis, particularly of iron, also emerged as a potential contributor to lifespan variation. Genes involved in iron and manganese transport and storage (*FIT2*, *ATX1*, *ATX2*) showed consistent expression differences. *FIT2*, located near highly differentiated genomic variants, has been linked to increased survival following DNA damage (Chen et al. 2020). These results suggest that metal ion availability influences DNA repair capacity in aging cells, potentially through interactions with polyamines such as spermine, which can stabilize DNA structures and modulate chromatin accessibility (Pedreño et al. 2005).

We also identified *TOR2*, a central component of the mTOR pathway and part of both *TORC1* and *TORC2* complexes (Kapahi et al. 2010), among our top genomic candidates. The TOR/mTOR pathway is one of several nutrient sensing pathways that has been linked to aging (Fontana et al. 2010), and has specifically been tied to replicative lifespan in yeast (Kaeberlein et al. 2005). Inhibition of *TORC1* has been linked to lifespan extension (e.g. Seo et al. 2024) and prevention of vacuolar fragmentation (Michaillat and Mayer 2013; Stauffer and Powers 2015, 2017), reinforcing its conserved role in aging. TOR (and DNA repair pathways) also play an important role in cell cycle regulation, which is inherently linked to replicative aging in yeast, as checkpoint mechanisms assess genomic integrity before division. Another candidate, *SWE1* (*Wee1*) regulates the G2/M transition and cell polarity, but has not been previously linked to aging. Our data suggest that *SWE1* may influence lifespan by modulating asymmetric inheritance of damaged cellular components, a process critical to rejuvenation in budding yeast (Lengefeld et al. 2017).

Our results highlight diverse molecular pathways, including cell cycle regulation, proteostasis, glycan metabolism, and metal ion homeostasis as components of the polygenic architecture of replicative aging. Unlike expression-profiling studies of isogenic strains (eg Hendrickson et al. 2018), or temporal analyses of single lineages (Janssens and Veenhoff 2016), our use of recombinant populations allowed us to interrogate natural genetic variation underlying extended replicative lifespan. The overlap between our findings and known aging hallmarks validates our approach,

while the discovery of novel contributors emphasizes the power of studies of standing genetic variation to deepen our understanding of complex traits.

Regulatory mechanisms underlying polygenic variation for lifespan

Having characterized coding-region differentiation, we next asked whether noncoding or regulatory variation contributed to aging-associated differences. While most candidate variants (~58%) occurred in genic regions, a substantial fraction were intergenic. Because RNA regulatory networks can influence aging (Montano and Long 2011; Čáp and Palková 2024), we searched for enrichment in regions involved in RNA surveillance but found none linked to canonical regulatory molecules (miRNAs, lncRNAs, snoRNAs). We did, however, detect age-associated upregulation of *AIR2*, a component of the TRAMP nuclear surveillance complex, consistent with regulation of RNA degradation mechanisms rather than noncoding RNA sequence variation. Notably, RNA exosome activity, including *AIR2*, has been implicated in maintaining cell wall integrity through regulation of protein glycosylation (Novačić et al. 2021).

Among the 191 unique genomic features implicated in our genomic and transcriptomic analyses, 9 are annotated as “dubious ORFs,” which are unlikely to encode functional proteins. While these could represent artifacts or misannotated protein-coding genes, prior studies suggest some act as regulatory RNAs (Havilio et al. 2005). One of the ORFs we identified, *YDR543C*, lies within a telomeric region and may reflect altered transcription of telomeric RNAs (Aguado et al. 2020; Zeinoun et al. 2023). However, expression of this gene was very low and should be interpreted with caution. No other significant differences in expression were seen in telomeric regions. Overall, we find limited evidence that noncoding sequence variation directly drives aging-related regulatory differences between cohorts.

If local cis-regulatory interactions were a dominant mechanism promoting late life survival, we would expect substantial overlap between genes implicated in our genomic and transcriptomic analyses, or an enrichment of genetic variants in noncoding regions in close proximity to differentially expressed transcripts. To test this, we examined the distribution of significant genomic variants within 10, 5, and 1 kb of differentially expressed genes. Only 14 differentially expressed genes (24%) had nearby variants, with fewer under stricter thresholds, suggesting limited cis-regulatory influence.

Among significant candidate SNPs, upstream gene variants were underrepresented and downstream gene variants were overrepresented relative to genome-wide expectations. This pattern could indicate that enhancer-like or trans-regulatory sequences in downstream regions contribute more strongly to expression variation than promoter or more proximal variants (Savinkova et al. 2009; Wittkopp and Kalay 2011). Given evidence that trans-regulatory interactions are a more common source of new and pleiotropic variation in yeast (Emerson et al. 2010; Siddiq and Wittkopp 2022; Vande Zande et al. 2022), our data support the hypothesis that trans-acting and post-transcriptional mechanisms, rather than local sequence variation, underlie most regulatory divergence associated with replicative aging.

Design considerations and trade-offs

Several design choices involve trade-offs that bear on interpretation. In each case, the intention was to favor heritable, reproducible signal over signal that could not be reliably disentangled

from technical noise, but these decisions involve trade-offs that bear on the interpretation of our results.

Allowing sorted cohorts to recover to mid-log phase prior to nucleic acid extraction was a deliberate design choice intended to capture heritable patterns of gene expression rather than the transient transcriptional state of cells immediately after sorting. Aged cells exiting FACS are simultaneously responding to multiple stressors, so profiling them at this stage would conflate these stress responses with true genotype-driven differences. This approach limits our ability to interrogate extremely aged, non-dividing cells, but ensures that any differentiation we report reflects variants whose effects persist through subsequent cell divisions.

The FACS gating thresholds were similarly chosen to balance enrichment against the asymmetry of the underlying biological distribution. The aged gate (top 15%) and the young gate (bottom 50%) are disjoint ranges on the fluorescence axis, separated by a buffer of ~35% of the distribution in which no cells were collected. The asymmetry reflects the underlying biology: most cells in a 9-d culture have low bud-scar counts, so a wider gate can be used to quickly retrieve a representative young cohort of comparable size, while the upper tail is sparse. The top 15% is the most aggressive feasible threshold that still yields 250,000 cells. Our bud-scar image data confirm that WGA fluorescence correlates with bud scar count across genotypes (Supplementary Figs. 1 and 2, Supplementary Table 1), but also show within-cohort variance in true replicative age. This heterogeneity dilutes the age-associated signal. Because we still detect significant genomic and transcriptomic differentiation, the underlying heritable contribution is at least as large as what we report.

The use of WGA staining and FACS itself introduces the potential for bias, since cell wall composition and morphology may influence both in a manner independent of replicative age. However, several features of our results argue against this. The WGA-binding mannoproteins CCW12 and CWP1 were both down-regulated in the aged fraction, which is the opposite of what we would expect if the protocol were enriching for genotypes with elevated baseline WGA affinity. The paired structure of our design further controls for genotype-dependent sort survival, since both fractions of each replicate experience identical sorting conditions, depleting any sorting-intolerant genotype symmetrically from both fractions. And the 3 major peaks of differentiation map to distinct ancestral haplotypes on different chromosomes encoding functionally diverse genes, a pattern difficult to reconcile with a sorting artifact, which would more plausibly concentrate in loci associated with structural elements of the cell wall. A genotype's WGA binding affinity could nonetheless bias its representation in the aged fraction, though this effect is expected to be modest given the single round of sorting used here.

The modest number of differentially expressed genes and relatively small effect sizes we observe are consistent with the highly polygenic architecture of replicative lifespan suggested by our genomic results. In a genetically heterogeneous population, different genotypes likely promote longevity through distinct molecular mechanisms, each involving different sets of genes and regulatory interactions. When expression is profiled from pooled cohorts, as in our design, these genotype-specific expression signatures are averaged across individuals, meaning that only expression differences that are consistent across many genetic backgrounds will emerge as significant. The fact that we nonetheless recover clear genomic differentiation, with 3 distinct haplotypes rising in frequency across independent replicates, demonstrates that our design can detect reproducible signals

when they exist. The limited overlap between genomic and transcriptomic candidates therefore reflects the biology of the trait; if strong, background-independent cis-regulatory effects were a primary driver of late-life survival in this population, we would expect to detect them. We cannot exclude the possibility that cis-regulatory variants contribute to longevity in a genotype-specific manner; dissecting such effects would require haplotype-resolved expression data or allele-specific approaches. Instead, our results suggest that the polygenic architecture of replicative aging is characterized by many small-effect loci acting through diverse, largely non-local regulatory mechanisms.

Conclusions

While identifying genetic variants associated with aging is valuable, their biological relevance must be interpreted within regulatory contexts. Integrating genomic and transcriptomic data from the same populations enables inference of such relationships, revealing both shared candidates and broader regulatory trends that would be invisible using a single data type. Here, this approach clarified spatial relationships between differentially expressed genes and nearby variants, and provided a framework for linking genomic variation to transcriptional outcomes associated with aging.

Future research could extend this framework through iterative cycles of sorting and experimental evolution to select for progressively longer-lived populations, allowing tests of whether aging-related signals remain stable or shift over time. Incorporating additional -omic layers, such as epigenetic, proteomic, and metabolomic analyses, could further refine our understanding of how natural genetic variation shapes aging. Epigenetic modifications, for instance, may contribute to aging on short evolutionary timescales, but over how many generations or cycles of selection might those be sustained? Likewise, investigating proteomic and metabolic profiles could provide deeper insights into the functional consequences of age-associated genetic changes, particularly in relation to glycan metabolism and protein composition.

While our study was limited in its ability to isolate extremely old cells, we identified unambiguous signals from a comparison of populations following a single instance of bud-scar sorting into young and aged cohorts. Additionally, haplotype data revealed that the allelic shifts observed in our aged replicates originated from 3 distinct ancestral haplotypes; this suggests that our approach of using genetically diverse populations of yeast allowed for the discovery of more complex adaptive responses that could not be identified from more traditional isogenic designs. Extending this approach across multiple age cohorts and selection cycles may ultimately disentangle how naturally-occurring variants and gene expression patterns jointly influence replicative lifespan.

Data availability

Raw sequencing data have been deposited in the NCBI Sequence Read Archive under accession PRJNA1368902. Code and processed data used for all analyses is available in a GitHub repository, along with README files explaining script sequence and usage (https://github.com/Katie-McHugh/regulation_of_aging/tree/Dryad_v3). The Burke Lab SNP calling pipeline is separately available at <https://github.com/mollyburke/Burke-Lab-SNP-calling-pipeline>. Raw and filtered or normalized datasets will be archived with Dryad (DOI: 10.5061/dryad.34tmpg4x4).

Supplemental material available at [G3](#) online.

Acknowledgments

We would like to thank Matt Powers for valuable statistical advice, Virginia Weis, Chris Curtin, and Aaron Liston and members of the Burke Lab (Megan Sandoval-Powers, Giovanni Crestani, Hannah Dugo, Brysyn Goodson, and Felix Vik), for their comments on earlier drafts of the manuscript. We would also like to acknowledge Jeff Bishop and Adam Fries at the GC3F facility at the University of Oregon for their lending their assistance and expertise with techniques for cell sorting and microscopy, respectively. AI platforms (ChatGPT-4o, Claude) were used occasionally to troubleshoot R code (primarily related to figure creation) or re-phrase text written by the authors for clarity and conciseness. All text and scripts were independently verified for accuracy by the authors.

Funding

This work was supported by National Institutes of Health award [R35GM147402 to MKB] as well as a seed grant awarded to MKB by the Collins Medical Trust and startup funds to MKB provided by the College of Science at Oregon State University.

Conflicts of interest

None declared.

Literature cited

- Aguado J, d'Adda di Fagnagna F, Wolvetang E. 2020. Telomere transcription in aging. *Ageing Res Rev.* 62:101115. <https://doi.org/10.1016/j.arr.2020.101115>.
- Aguilani H, Gustafsson L, Rigoulet M, Nyström T. 2003. Asymmetric inheritance of oxidatively damaged proteins during cytokinesis. *Science.* 299:1751–1753. <https://doi.org/10.1126/science.1080418>.
- Ashburner M et al. 2000. Gene ontology: tool for the unification of biology. *Nat Genet.* 25:25–29. <https://doi.org/10.1038/75556>.
- Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM. 2015. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One.* 10:e0131262. <https://doi.org/10.1371/journal.pone.0131262>.
- Burke MK, King EG, Shahrestani P, Rose MR, Long AD. 2014a. Genome-wide association study of extreme longevity in *Drosophila melanogaster*. *Genome Biol Evol.* 6:1–11. <https://doi.org/10.1093/gbe/evt180>.
- Burke MK, Liti G, Long AD. 2014b. Standing genetic variation drives repeatable experimental evolution in outcrossing populations of *Saccharomyces cerevisiae*. *Mol Biol Evol.* 31:3228–3239. <https://doi.org/10.1093/molbev/msu256>.
- Burke MK, McHugh KM, Kutch IC. 2020. Heat shock improves random spore analysis in diverse strains of *Saccharomyces cerevisiae*. *Front Genet.* 11:597482. <https://doi.org/10.3389/fgene.2020.597482>.
- Čáp M, Palková Z. 2024. Non-coding RNAs: regulators of stress, ageing, and developmental decisions in yeast? *Cells.* 13:599. <https://doi.org/10.3390/cells13070599>.
- Chen C, Contreras R. 2004. The bud scar-based screening system for hunting human genes extending life span. *Ann N Y Acad Sci.* 1019:355–359. <https://doi.org/10.1196/annals.1297.061>.
- Chen KL et al. 2020. Loss of vacuolar acidity results in iron-sulfur cluster defects and divergent homeostatic responses during aging in *Saccharomyces cerevisiae*. *Geroscience.* 42:749–764. <https://doi.org/10.1007/s11357-020-00159-3>.

- Cindrić A, Krištić J, Martinić Kavur M, Pezer M. 2021. Glycosylation and aging. *Adv Exp Med Biol.* 1325:341–373. https://doi.org/10.1007/978-3-030-70115-4_17.
- Cingolani P et al. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin).* 6:80–92. <https://doi.org/10.4161/fly.19695>.
- Cubillos FA et al. 2017. Identification of nitrogen consumption genetic variants in yeast through QTL mapping and bulk segregant RNA-Seq analysis. *G3 (Bethesda).* 7:1693–1705. <https://doi.org/10.1534/g3.117.042127>.
- Cubillos FA, Louis EJ, Liti G. 2009. Generation of a large set of genetically tractable haploid and diploid *Saccharomyces* strains. *FEMS Yeast Res.* 9:1217–1225. <https://doi.org/10.1111/j.1567-1364.2009.00583.x>.
- Dall'Olio F. 2018. Glycobiology of aging. *Subcell Biochem.* 90:505–526. https://doi.org/10.1007/978-981-13-2835-0_17.
- Dobin Alexander, Davis Carrie A, Schlesinger Felix, Drenkow Jorg, Zaleski Chris, Jha Sonali, Batut Philippe, Chaisson Mark, Gingeras Thomas R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 29: 15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
- Emerson JJ et al. 2010. Natural selection on cis and trans regulation in yeasts. *Genome Res.* 20:826–836. <https://doi.org/10.1101/gr.101576.109>.
- Fontana L, Partridge L, Longo VD. 2010. Extending healthy life span—from yeast to humans. *Science.* 328:321–326. <https://doi.org/10.1126/science.1172539>.
- Franco-Duarte R et al. 2017. Genomic and transcriptomic analysis of *Saccharomyces cerevisiae* isolates with focus in succinic acid production. *FEMS Yeast Res.* 17:057. <https://doi.org/10.1093/femsyr/fox057>.
- Ghanegolmohammadi F et al. 2021. Defining functions of mannoproteins in *Saccharomyces cerevisiae* by high-dimensional morphological phenotyping. *J Fungi.* 7:769. <https://doi.org/10.3390/jof7090769>.
- Ghavidel A et al. 2018. Rapid nuclear exclusion of Hcm1 in aging *Saccharomyces cerevisiae* leads to vacuolar alkalization and replicative senescence. *G3 (Bethesda).* 8:1579–1592. <https://doi.org/10.1534/g3.118.200161>.
- Guo Y et al. 2013. Evaluation of allele frequency estimation using pooled sequencing data simulation. *ScientificWorldJournal.* 2013:895496. <https://doi.org/10.1155/2013/895496>.
- Havilio M, Levanon EY, Lerman G, Kupiec M, Eisenberg E. 2005. Evidence for abundant transcription of non-coding regions in the *Saccharomyces cerevisiae* genome. *BMC Genomics.* 6:93. <https://doi.org/10.1186/1471-2164-6-93>.
- Hendrickson DG et al. 2018. A new experimental platform facilitates assessment of the transcriptional and chromatin landscapes of aging yeast. *eLife.* 7:e39911. <https://doi.org/10.7554/eLife.39911>.
- Hughes AL, Gottschling DE. 2012. An early age increase in vacuolar pH limits mitochondrial function and lifespan in yeast. *Nature.* 492:261–265. <https://doi.org/10.1038/nature11654>.
- Janssens GE, Veenhoff LM. 2016. Evidence for the hallmarks of human aging in replicatively aging yeast. *Microb Cell.* 3:263–274. <https://doi.org/10.15698/mic2016.07.510>.
- Kachroo AH et al. 2015. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. *Science.* 348: 921–925. <https://doi.org/10.1126/science.aaa0769>.
- Kaeberlein M et al. 2005. Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients. *Science.* 310:1193–1196. <https://doi.org/10.1126/science.1115535>.
- Kapahi P et al. 2010. With TOR, less is more: a key role for the conserved nutrient-sensing TOR pathway in aging. *Cell Metab.* 11: 453–465. <https://doi.org/10.1016/j.cmet.2010.05.001>.
- Krištić J et al. 2014. Glycans are a novel biomarker of chronological and biological ages. *J Gerontol A Biol Sci Med Sci.* 69:779–789. <https://doi.org/10.1093/gerona/glt190>.
- Kung LA et al. 2009. Global analysis of the glycoproteome in *Saccharomyces cerevisiae* reveals new roles for protein glycosylation in eukaryotes. *Mol Syst Biol.* 5:308. <https://doi.org/10.1038/msb.2009.64>.
- Leach L. J, Zhang Z., Lu C, Kearsley M. J, Luo Z. 2007. The Role of Cis-Regulatory Motifs and Genetical Control of Expression in the Divergence of Yeast Duplicate Genes. *Molecular Biology and Evolution.* 24:2556–2565. <https://doi.org/10.1093/molbev/msm188>.
- Lecheta MC et al. 2020. Integrating GWAS and transcriptomics to identify the molecular underpinnings of thermal stress responses in *Drosophila melanogaster*. *Front Genet.* 11:658. <https://doi.org/10.3389/fgene.2020.00658>.
- Lemoine M. 2021. The evolution of the hallmarks of aging. *Front Genet.* 12:693071. <https://doi.org/10.3389/fgene.2021.693071>.
- Lengefeld J, Hotz M, Rollins M, Baetz K, Barral Y. 2017. Budding yeast Wee1 distinguishes spindle pole bodies to guide their pattern of age-dependent segregation. *Nat Cell Biol.* 19:941–951. <https://doi.org/10.1038/ncb3576>.
- Lesur I, Campbell JL. 2004. The transcriptome of prematurely aging yeast cells is similar to that of telomerase-deficient cells. *Mol Biol Cell.* 15:1297–1312. <https://doi.org/10.1091/mbc.e03-10-0742>.
- Li B, Ritchie M. 2021. From GWAS to gene: transcriptome-wide association studies and other methods to functionally understand GWAS discoveries. *Front Genet.* 12:713230. <https://doi.org/10.3389/fgene.2021.713230>.
- Li L et al. 2024. A cross-tissue transcriptome wide association studies identifies new susceptibility genes for insomnia. *J Neurophysiol.* 133:572–581. <https://doi.org/10.1152/jn.00490.2024>.
- Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 30:923–930. <https://doi.org/10.1093/bioinformatics/btt656>.
- Linder RA, Majumder A, Chakraborty M, Long A. 2020. Two synthetic 18-way outcrossed populations of diploid budding yeast with utility for complex trait dissection. *Genetics.* 215:323–342. <https://doi.org/10.1534/genetics.120.303202>.
- López-Otín C, Blasco M, Partridge L, Serrano M, Kroemer G. 2013. The hallmarks of aging. *Cell.* 153:1194–1217. <https://doi.org/10.1016/j.cell.2013.05.039>.
- López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2023. Hallmarks of aging: an expanding universe. *Cell.* 186:243–278. <https://doi.org/10.1016/j.cell.2022.11.001>.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
- Mamoun CB, Beckerich JM, Gaillardin C, Kepes F. 1999. Disruption of YHC8, a member of the TSR1 gene family, reveals its direct involvement in yeast protein translocation. *J Biol Chem.* 274: 11296–11302. <https://doi.org/10.1074/jbc.274.16.11296>.
- Michaillat L, Mayer A. 2013. Identification of genes affecting vacuole membrane fragmentation in *Saccharomyces cerevisiae*. *PLoS One.* 8: e54160. <https://doi.org/10.1371/journal.pone.0054160>.
- Miura Y, Endo T. 2016. Glycomics and glycoproteomics focused on aging and age-related diseases—glycans as a potential biomarker for physiological alterations. *Biochim Biophys Acta.* 1860: 1608–1614. <https://doi.org/10.1016/j.bbagen.2016.01.013>.
- Montano M, Long K. 2011. RNA surveillance—an emerging role for RNA regulatory networks in aging. *Ageing Res Rev.* 10:216–224. <https://doi.org/10.1016/j.arr.2010.02.002>.

- Novačić A et al. 2021. Yeast RNA exosome activity is necessary for maintaining cell wall stability through proper protein glycosylation. *Mol Biol Cell*. 32:363–375. <https://doi.org/10.1091/mbc.E20-08-0544-T>.
- Palmer RD. 2022. Aging clocks & mortality timers, methylation, glycomic, telomeric and more. A window to measure biological age. *Aging Med*. 5:120–125. <https://doi.org/10.1002/agm2.12197>.
- Pedreño E, López-Contreras A, Cremades A, Peñafiel R. 2005. Protecting or promoting effects of spermine on DNA strand breakage induced by iron or copper ions as a function of metal concentration. *J Inorg Biochem*. 99:2074–2080. <https://doi.org/10.1016/j.jinorgbio.2005.07.005>.
- Perez Gerardo, Barber Galt P, Benet-Pages Anna, Casper Jonathan, Clawson Hiram, Diekhans Mark, Fischer Clay, Gonzalez Jairo Navarro, Hinrichs Angie S, Lee Christopher M et al. 2025. The UCSC Genome Browser database: 2025 update. *Nucleic Acids Research*. 53:D1243–D1249. <https://doi.org/10.1093/nar/gkae974>.
- Phillips MA, Kutch IC, McHugh KM. 2021. Crossing design shapes patterns of genetic variation in synthetic recombinant populations of *Saccharomyces cerevisiae*. *Sci Rep*. 11:19551. <https://doi.org/10.1038/s41598-021-99026-0>.
- Ragni E et al. 2011. The genetic interaction network of CCW12, a *Saccharomyces cerevisiae* gene required for cell wall integrity during budding and formation of mating projections. *BMC Genomics*. 12:107. <https://doi.org/10.1186/1471-2164-12-107>.
- Renganaath K et al. 2020. Systematic identification of cis-regulatory variants that cause gene expression differences in a yeast cross. *Elife*. 9:62669. <https://doi.org/10.7554/eLife.62669>.
- Ronald James, Brem Rachel B, Whittle Jacqueline, Kruglyak Leonid, Frankel Wayne. 2005. Local Regulatory Variation in *Saccharomyces cerevisiae*. *PLoS Genetics*. 1:e25. <https://doi.org/10.1371/journal.pgen.0010025>.
- Saarikangas J, Barral Y. 2015. Protein aggregates are associated with replicative aging without compromising protein quality control. *Elife*. 4:e06197. <https://doi.org/10.7554/eLife.06197>.
- Savinkova LK et al. 2009. TATA box polymorphisms in human gene promoters and associated hereditary pathologies. *Biochemistry*. 74:117–129. <https://doi.org/10.1134/s0006297909020011>.
- Schacherer J, Shapiro JA, Ruderfer DM, Kruglyak L. 2009. Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature*. 458:342–345. <https://doi.org/10.1038/nature07670>.
- Seo Dongseong, Yalcin Gulperi, Jang Hyeonjun, Lee Han-Jun, Kim Deok Ho, Lee Cheol-Koo. 2024. TOR2 plays the central role in rapamycin-induced lifespan extension in budding yeast. *Biochemical and Biophysical Research Communications*. 734:150639. <https://doi.org/10.1016/j.bbrc.2024.150639>.
- Shih CS, Fay J. 2021. Cis-regulatory variants affect gene expression dynamics in yeast. *eLife*. 10:68469. <https://doi.org/10.7554/eLife.68469>.
- Siddiq MA, Wittkopp PJ. 2022. Mechanisms of regulatory evolution in yeast. *Curr Opin Genet Dev*. 77:101998. <https://doi.org/10.1016/j.gde.2022.101998>.
- Skrzypek MS et al. 2018. *Saccharomyces* genome database informs human biology. *Nucleic Acids Res*. 46:D736–D742. <https://doi.org/10.1093/nar/gkx1112>.
- Smits GJ, Schenkman LR, Brul S, Pringle JR, Klis FM. 2006. Role of cell cycle-regulated expression in the localized incorporation of cell wall proteins in yeast. *Mol Biol Cell*. 17:3267–3280. <https://doi.org/10.1091/mbc.e05-08-0738>.
- Stauffer B, Powers T. 2015. Target of rapamycin signaling mediates vacuolar fission caused by endoplasmic reticulum stress in *Saccharomyces cerevisiae*. *Mol Biol Cell*. 26:4618–4630. <https://doi.org/10.1091/mbc.E15-06-0344>.
- Stauffer B, Powers T. 2017. Target of rapamycin signaling mediates vacuolar fragmentation. *Curr Genet*. 63:35–42. <https://doi.org/10.1007/s00294-016-0616-0>.
- Taus T, Futschik A, Schlötterer C. 2017. Quantifying selection with pool-seq time series data. *Mol Biol Evol*. 34:3023–3034. <https://doi.org/10.1093/molbev/msx225>.
- The Gene Ontology Consortium et al. 2023. The gene ontology knowledgebase in 2023. *Genetics*. 224:iyad031. <https://doi.org/10.1093/genetics/iyad031>.
- Van der Auwera G, O'Connor BD. 2020. *Genomics in the cloud: using Docker, GATK, and WDL in terra*. 1st ed. O'Reilly Media.
- Vande Zande PV, Hill M, Wittkopp PJ. 2022. Pleiotropic effects of trans-regulatory mutations on fitness and gene expression. *Science*. 377:105–109. <https://doi.org/10.1126/science.abj7185>.
- Wing KM, Phillips MA, Baker AR, Burke MK. 2020. Consequences of cryopreservation in diverse natural isolates of *Saccharomyces cerevisiae*. *Genome Biol Evol*. 12:1302–1312. <https://doi.org/10.1093/gbe/evaa121>.
- Wittkopp PJ, Kalay G. 2011. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet*. 13:59–69. <https://doi.org/10.1038/nrg3095>.
- Wong Edith D, Miyasato Stuart R, Aleksander Suzi, Karra Kalpana, Nash Robert S, Skrzypek Marek S, Weng Shuai, Engel Stacia R, Cherry J Michael, Laceyfield S. 2023. *Saccharomyces* genome database update: server architecture, pan-genome nomenclature, and external resources. *GENETICS*. 224:iyac022. <https://doi.org/10.1093/genetics/iyac191>.
- Wu Y et al. 2025. A high-resolution N-glycoproteome landscape of aging mouse ovary. *Redox Biol*. 81:103584. <https://doi.org/10.1016/j.redox.2025.103584>.
- Zeinoun B, Teixeira MT, Barascu A. 2023. TERRA and telomere maintenance in the yeast *Saccharomyces cerevisiae*. *Genes (Basel)*. 14:618. <https://doi.org/10.3390/genes14030618>.
- Zhang X et al. 2021. Genome-wide transcriptome study using deep RNA sequencing for myocardial infarction and coronary artery calcification. *BMC Med Genomics*. 14:45. <https://doi.org/10.1186/s12920-020-00838-2>.
- Zhang Y, Parmigiani G, Johnson WE. 2020. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genom Bioinform*. 2:lqaa078. <https://doi.org/10.1093/nargab/lqaa078>.

Editor: S. Lee