bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Evolution of Natural Lifespan Variation and Molecular Strategies of Extended Lifespan

Alaattin Kaya[1,2*], Cheryl Zi Jin Phua[3], Mitchell Lee[4], Lu Wang[5], Alexander Tyshkovskiy[2,6], Siming Ma[3], Weiqiang Liu[7], Benjamin Harrison[4], Xiaqing Zhao[4], Xuming Zhou[2,7], Brian M. Wasko[8], Theo K. Bammler[5], Daniel E. L. Promislow[4,9], Matt Kaeberlein[4], Vadim N. Gladyshev[2*]

[1] Virginia Commonwealth University, Department of Biology, Richmond, VA, 23284, USA

[2] Division of Genetics, Brigham and Women's Hospital, Boston, MA, 02115, USA

[3] Genome Institute of Singapore, Agency for Science, Technology and Research (A*STAR), Singapore

[4] Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, 98195, USA

[5] Department of Environmental and Occupational Health Sciences, University of Washington, Seattle WA 98105, USA

[6] Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow, Russia

[7] Key Laboratory of Animal Ecology and Conservation Biology, Chinese Academy of Sciences, Institute of Zoology, Beijing, China

[8] Department of Biology, University of Houston-Clear Lake, Houston, TX 77058, USA

[9] Department of Biology, University of Washington, Seattle, WA 98195, USA

[*]Corresponding Authors: kayaa@vcu.edu, vgladyshev@rics.bwh.harvard.edu

KEYWORDS: aging, natural lifespan variation, adaptation, gene-environment interaction, caloric restriction

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

The question of why and how some species or individuals within a population live longer than others is among the most important questions in the biology of aging. A particularly useful model to understand the genetic basis and selective forces acting on the plasticity of lifespan are closely related species or ecologically diverse individuals of the same species widely different in lifespan. Here, we analyzed 76 diverse wild isolates of two closely related budding yeast species *Saccharomyces cerevisiae and Saccharomyces paradoxus* and discovered a diversity of natural intra-species lifespan variation. We sequenced the genomes of these organisms and analyzed how their replicative lifespan is shaped by nutrients and transcriptional and metabolite patterns. We identified sets of genes and metabolites to regulate aging pathways, many of which have not been previously associated with lifespan regulation. We also identified and characterized long-lived strains with elevated intermediary metabolites and differentially regulated genes for NAD metabolism and the control of epigenetic landscape through chromatin silencing. Our data further offer insights into the evolution and mechanisms by which caloric restriction regulates lifespan by modulating the availability of nutrients without decreasing fitness. Overall, our study shows how the environment and natural selection interact to shape diversity of lifespan.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

INTRODUCTION

Organisms can respond to environmental cues by modifying the genotype to arrive at more than one phenotype, a process known as phenotypic plasticity [1, 2]. Consequently, many morphological, behavioral and physiological phenotypes vary within and between species in natural populations [3-7]. For example, genetic variation in natural populations of many organisms can differentially affect their neural and endocrine functions, leading to variation in quantitative life-history traits such as fitness and age at maturation [8, 9]. Variation in another fitness trait, lifespan, has also attracted much attention [10-12]. Across eukaryotes, species longevity can differ over many orders of magnitude, from the 2-day replicative lifespan (RLS) of budding yeast (*Saccharomyces cerevisiae*) to organisms that live for centuries, such as the Greenland shark (*Somniosus microcephalus*), which can live for more than 500 years [13-15]. Longevity also consistently varies among genetically diverse individuals of the same species, indicating that variability of lifespan is not constrained at the level of species, and that the molecular determinants of lifespan appear to be phenotypically plastic within the same genetic pool [16-21].

What, then, are the factors that favor long-lived or short-lived individuals? The molecular pathways that underlie the genetic and environmental determinants of lifespan are among the most intensely studied areas in the aging field [18, 19, 22-26]. Animal models of aging have shown that longevity of laboratory animals can be extended by environmental [27-30], dietary [31-33], pharmacological [34-37] and genetic interventions [38-40]. However, many of these laboratory-adapted populations are constrained by genetic and environmental background [41-43]. For example, artificially created mutant strains may show longer lifespan under laboratory setting, but demonstrate reduced fitness in their natural environment [44-47]. Therefore, a more integrated approach is needed to understand how the natural environment and natural selection interact to shape lifespan and associated life-history traits.

In order to understand the experiments that nature has carried out, modifying genotypes to arrive at different lifespans, we need to identify how environmental factors can trigger changes in the genome which in turn impact lifespan plasticity. Towards this goal, we employed 76 wild isolates of the budding yeast to capture evolved diversity of natural lifespan. This collection includes 40 diploid isolates of *Saccharomyces cerevisiae* and 36 diploid isolates of *Saccharomyces paradoxus* [48, 49]. These two species are closely related (sharing a common

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ancestor between 0.4 and 3 million years ago) with 90% genome identity and can mate and produce viable progeny [50, 51]. Earlier genome sequencing has shown that while some of these strains fall into distinct lineages with unique genetic variants, almost half of the strains have mosaic recombinant genomes arising from outcrosses between these genetically distinct lineages [49]. Therefore, these strains offer a great genetic pool to understand gene-environment interaction in shaping lifespan variation.

Accordingly, we assayed replicative lifespan (RLS) of ~4,000 individual cells from these isolates under three different conditions: yeast peptone dextrose (YPD, with 2% glucose), yeast peptone glycerol (YPG, 3% glycerol as a respiratory carbon source), and caloric restriction (CR-0.05% glucose). We identified up to 10-fold variation in median RLS. Although little is known about the natural and life histories of these wild isolates, they occupy diverse ecological niches ranging from trees to fruits to soil and face different evolutionary pressures for adaptation; thus, their natural lifespan variation must be encoded in their respective genomes [49, 51, 52]. We explored the relationship between gene expression, metabolite abundance, and longevity phenotypes to characterize the molecular patterns associated with condition-specific natural lifespan variation. We characterized genes and metabolites with significant positive or negative correlation to longevity and identified pathways associated with median RLS across these isolates under each condition. Our data showed that the environment can lead to different genotypes with wide differences in lifespan, which are regulated by distinct sets of genes, metabolites and connected pathways that have not been previously characterized in the laboratory setting. Altogether, this is the first comprehensive analysis of how the environment and natural selection interact to shape aging and the associated life-history traits.

RESULTS AND DISCUSSION

Wild isolates of Saccharomyces

Natural variation offers a powerful tool to assess the forces shaping sequence variation that underlies phenotypic diversity of natural populations. This approach could help explain how the phenotype is shaped by genotype and environment [2, 53]. Because of their wide ecological, geographical and genetic diversity, natural isolates of the budding yeast *Saccharomyces* have become an important model system for population/evolutionary genomics [52]. We took advantage of natural variation across 76 diploid wild yeast isolates of two closely related

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Saccharomyces species, including 40 *S. cerevisiae* and 36 *S. paradoxus* isolates, to examine variation in longevity. Their niches include human-associated environments, such as breweries and bakeries, and different types of wild ecological niches, such as trees, fruits, vineyards, and soils across different continents. There was also a group of pathogenic *S. cerevisiae* strains isolated from immunocompromised patients **(Supplementary Table 1)**. To understand the forces that shape diversity of lifespan and the associated mechanisms, we first sequenced and assembled the genomes of these 76 isolates. Consistent with a previously published report [49], we found substantial genetic variation among these strains (~655,000SNPs). We also analyzed ploidy states of these strains and characterized those with increased copy number of specific chromosomes (**Supplementary Fig. 1A, B**).

Growth characteristics of wild isolates in liquid culture

We monitored growth characteristics of natural yeast isolates on standard glucose media (i.e. during fermentation) and on media with a non-fermentable substrate, glycerol, as a carbon source (i.e. during respiration), using an automated growth analyzer, and calculated the doubling time. Most wild isolates grew faster than the diploid laboratory wild type (WT) BY4743 strain under both conditions, with an average doubling time of 65 minutes in YPD and 125 minutes in YPG **(Fig. 1A, Supplementary Table 1)**. Importantly, most of these strains grew at a similar rate on YPD, and variation in growth rate on YPG was also relatively small, indicating that these laboratory-optimized culture conditions are suitable for the nutritional needs of these strains and for RLS analysis **(Fig. 1A)**.

Lifespan under glucose, glycerol, and calorie restricted growth conditions

We assayed RLS of these isolates at 30 °C using 3 different growth conditions: YPD medium, YPG medium, and medium with restricted glucose (calorie restriction, CR, 0.05% glucose). Under YPD conditions, we observed a remarkable ~10-fold variation in median RLS and 9-fold variation in maximum RLS across these isolates (**Fig. 1B**, Pearson correlation coefficient = 0.65 between medium and maximum lifespans, $P < 0.00001$). The average median RLS of *S. paradoxus* strains (29.7) was significantly higher (adjusted P value $=2.7\times10^{-9}$) than the average median RLS of *S. cerevisiae* strains (22.7), showing that the *S. paradoxus* mother cells can produce more daughter cells in this medium (**Fig. 1B, C**). *S. paradoxus* strain Q32.3 had the

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

longest median RLS (median RLS=38.5), whereas *S. cerevisiae* strains Y10 and YS2 exhibited the shortest RLS (median RLS=4 for Y10 and 6 for YS2) **(Supplementary Table 1).**

Glycerol as a growth substrate can extend both RLS and chronological lifespan (CLS) of yeast [54, 57]. In the case of CLS, the increased longevity is clearly caused by a switch from fermentative metabolism to respiration [54]; however, for RLS the mechanism is unclear, as respiratory metabolism is not required for lifespan extension from CR in laboratory WT strains [58], and the requirement of respiratory metabolism for lifespan extension by glycerol has not been established. While we observed a median RLS increase in 39 strains on YPG; most other 27 strains showed no change and few strains decreased median RLS (**Fig. 1D, Supplementary Fig. 2A, Supplementary Table 1**). For example, *S. cerevisiae* strain BC187 showed an increased maximum and medium RLS when comparing growth on glucose and glycerol (maximum RLS = 62 in YPD and 92 in YPG; median R.LS = 32 in YPD and 37.5 in YPG). *S. paradoxus* strain KPN3829 also showed a similar increase (maximum RLS = 39 in YPD and 60 in YPG; median RLS = 21 in YPD and 35 in YPG) **(Supplementary Table 1, Supplementary Fig. 2A)**. On the other hand, *S. cerevisiae* strain YJM975 showed a decrease in both measures (maximum RLS= 70 in YPD and 60 in YPG; median RLS = 31 in YPD and 23 in YPG) **(Supplementary Table 1, Supplementary Fig. 2A).**

We further dissected the role of carbon source metabolism in RLS by comparing median RLS changes between YPD and YPG conditions. Interestingly, strains with the shortest RLS on YPD tended to achieve the largest lifespan gains from switching to YPG, while the long-lived strains on YPD generally did not show a further RLS increase (**Supplementary Fig. 2B)**. For example, 42 out of 51 short-lived strains (compared to laboratory WT strain on YPD) increased their median RLS when switched to glycerol **(Supplementary Fig. 2A, Supplementary Table 1),** whereas 17 out of 23 long-lived strains (compared to laboratory WT strain on YPD) did not show further median RLS extension **(Supplementary Fig. 2A, Supplementary Table 1**). This correlation matches prior studies of genetically diverse single gene deletions in the BY4742 background where the shortest-lived strains tended to receive the largest lifespan extension in response to CR [57]. Further analyses of a subset of wild isolates (36 strains) on the CR medium (0.05% glucose) revealed a similar pattern, with the shortest-lived strains on YPD tending to receive the largest lifespan increase from CR **(Supplementary Fig. 2C, Supplementary Table**

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1). Consistent with this, the effects of CR and YPG on lifespan relative to YPD for each strain were significantly correlated (r= 0.62, *Padj*=2.29x10^{-4}) **(Supplementary Fig. 2C).** CR is associated with increased respiration across different organisms, including yeast [59-61]. While we observed a 13% median RLS increase in the diploid laboratory WT strain, wild yeast isolates showed a wide range of median RLS under CR conditions **(Fig. 1D).** Among the 36 strains assayed under CR conditions, 16 did not respond to CR and decreased median RLS, while the other 20 strains showed various degrees of median RLS extension **(Supplementary Table 1, Supplementary Fig. 3).** For example, YJM978 strain showed a decreased median RLS (50% decrease), whereas CR caused a 75% median RLS increase of Y9 strain, compared to the 2% glucose medium **(Supplementary Table 1, Supplementary Fig. 3)**. Overall, our data suggest that long-lived strains have distinct niche-specific nutritional and environmental adaptations, similar to CR condition that resulted in altered lifespan. Therefore, there was no consistent RLS extension in long-lived strains when switched to glycerol or subjected to CR.

Endophenotype variation across wild isolates

While many previous large-scale omics studies on aging focused on genome-wide association [62-65], recent comparative studies on transcriptomics [66, 67], proteomics [68-70], metabolomics [71-73] and ionomics [72] have begun to shed light on molecular patterns and mechanisms that link to endophenotypes. For example, it has been suggested that natural variation is associated with extensive changes in gene expression, translation and metabolic regulation, which in turn may affect selection under different stress conditions [74, 75]. In fact, gene expression variation has repeatedly been postulated to play a major role in adaptive evolution and phenotypic plasticity [75, 76]. Evolution of gene expression is known to exhibit selective constraints [77, 78], thereby supporting specific phenotypic outcomes such as changes in morphology [79] and lifespan [80]. Similarly, comparative studies on molecule abundance by metabolite profiling have been utilized to describe the genotype to phenotype relations in model organisms [81, 82]. Accordingly, we aimed to explain lifespan differences among these natural populations of closely related yeast isolates by analyzing their endophenotype differences that include gene expression variation (transcriptomics) and differences in their metabolite levels (metabolomics).

For the transcriptome analyses, we obtained ≥ 5 million 150-bp paired-end RNA-seq reads for each strain. For metabolomics analysis, we applied targeted metabolite profiling using

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

liquid chromatography-mass spectrometry (LC-MS). After filtering and quality control, the data set contained RNA-seq reads for 5,376 genes and 166 metabolites identified commonly across all isolates (**Supplementary Table 2**). The expression profiles were similar to one another, with Spearman correlation coefficients of strain pairs ranging between 0.59 and 0.93 (except the pairing to Q59.1, CBS5829, YPS606, and UFRJ50791, with the range between 0.21 and 0.79) **(Supplementary Fig. 4, Supplementary Table 2).** To determine whether the previously published sequence-based evolutionary relationship [49] was reflected in their gene expression variation, we constructed gene expression phylograms using a distance matrix of 1 minus Spearman correlation coefficients and the neighbor-joining method [83]. The resulting topology was largely consistent with their phylogeny with a clear separation between *S. cerevisiae* and *S. paradoxus* strains **(Fig. 2A).**

To visualize endophenotypic variation between these two species and across the strains of the same species, we performed Principal Component analysis (PCA) on each type of data. PCA of the transcriptome revealed a pattern resembling the phylogenetic relationship, with the first three PCs explaining ~49 % of total variance in gene expression **(Fig. 2B, Supplementary Fig. 5A)**. Although some *S. paradoxus* strains clustered with *S. cerevisiae*, we observed clear species segregation based on PC2 (except for some outlier strains that were separated by PC1). PCA of metabolomics data revealed a similar structure with the first three PCs explaining ~41% of total variance in metabolite levels, and PC2 somewhat separating the species **(Fig. 2C, Supplementary Fig. 5B).** However, the strains were segregated along PC1 largely according to their median RLS, and the effect was stronger across *S. paradoxus* strains.

To understand the basis of this segregation pattern, we performed pathway enrichment analysis by combining the 500 top genes (250 with positive weights and 250 with negative weights) and 40 top metabolites (20 with positive weights and 20 with negative weights) contributing to each PC, respectively. This integrative analysis of genes and metabolites contributing to PC1 revealed a distinct set of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, including RNA degradation, MAPK signaling pathway, cell cycle, pantothenate and CoA biosynthesis, ribosome biogenesis and pentose phosphate pathway **(Fig. 2D, Supplementary Table 3)**. The analysis of genes and metabolites for PC2 revealed the KEGG pathways related to ribosome, autophagy, endocytosis, cell cycle, mRNA surveillance, and nucleotide excision repair **(Fig. 2E, Supplementary Table 3)**. These results suggest that

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

these processes diverged most significantly across the wild isolates of two species of *Saccharomyces* genus and may account for their phenotypic diversity, including lifespan. It should be noted, however, that we do not know exactly what each PC represents, unless it perfectly aligns or correlates with some known variables. Also, either biological (e.g. phylogenetic structure) or technical (e.g. data normalization or batch effect) or mixed effects of both may render PCA biased [84].

Gene expression and metabolite abundance correlate with longevity

To identify the endophenotypes (genes and metabolites) correlating with natural lifespan variation across wild isolates, we applied the phylogenetic generalized least-squares (PGLS) method to account for the phylogenetic relationships among the strains and test for different models of traits evolution [85, 86]. Regression was performed between endophenotypic values and median RLS under different models of trait evolution and the best-fit model was then selected based on maximal likelihood (Materials and Methods). To assess robustness of the relationship, calculation of regression slopes was repeated by taking out one yeast strain at a time. This ensured the overall relationship did not depend on a particular isolate.

With the PGLS approach, we identified 73 transcripts with significant correlation to median RLS ($Padj \leq 0.01$; 39 with positive correlation and 34 with negative correlation) **(Supplementary Table 2).** Among the top hits with positive correlation were a putative zinc finger protein coding gene *CMR3* (*Padj*=3.3×10^{-9}), histone acetyltransferase (HAT) gene *HPA2* (*Padj*=0.0002), O-methyltransferase *COQ3* (*Padj*=0.006) and zing regulated protein gene *ZRG8* **(Fig. 3A, Supplementary Table 2).** The top hits with negative correlation included the genes coding for cyclin-dependent kinase Pho85p interacting proteins *PCL1* (*Padj*=0.0008) and *PCL2* (*Padj*=0.001), regulator of Ty1 transposon protein coding gene *RTT107* (*Padj*=0.007), and inositol monophosphatase gene *INM1* (*Padj*=0.006) **(Fig. 3B, Supplementary Table 2).**

To assess if any of our transcript hits were previously implicated in yeast lifespan, we compared our list of significant genes (357 genes at *Padj*=0.05) to the genes associated with RLS regulation in laboratory WT strain listed in the GenAge database [87]. The GenAge list contains 611 genes from the published literature that were reported for decreased or increased RLS of laboratory yeast strains (595 deleted and 16 overexpressed genes) **(Supplementary Table 4)**. There were 39 genes present in both our list and the GenAge list, 23 of which showed the same

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

direction of correlation with RLS. For example, *INM1*, *RTT107*, *PPH3* and *BSC1* genes increase RLS when deleted (GenAge database) and increase RLS when their transcript levels decrease across wild isolates (this study). **(Supplementary Table 4).** However, the overall pattern did not reach statistical significance (Fisher exact test, $p \geq 0.05$).

We further compared our results with the gene expression patterns across wild isolates and 1,376 laboratory knock-out strains (KO) strains [88], whose RLS was published [89] previously **(Supplementary Fig. 6A, B Supplementary Table 1)**. We calculated an association of gene expression with different measures of RLS (mean RLS, median RLS, and maximum RLS) across KO strains. Our analysis revealed around 400 significant genes (*Padj*=0.05) commonly up or down regulated in all types of RLS measures and more than 1000 genes associated with median RLS **(Fig. 4A).** We then calculated a correlation matrix of RLS-associated gene expression changes across KO strains and wild isolates **(Fig. 4B).** We found no positive correlation between RLS associated gene expression changes across KO strains and RLS associated gene expression changes across natural isolates **(Fig. 4B)**.

To further understand whether the same sets of biological pathways are involved in lifespan variation achieved by laboratory manipulations and by natural selection, we performed functional enrichment (GSEA) of genes associated with RLS across deletion and wild isolates **(Fig. 4C).** For the genes correlating positively with longevity across the KO strains, the enriched terms included cellular responses to stress, ribosome, translation, cellular senescence, and DNA repair **(Fig. 4C).** On the other hand, positive terms enriched in wild isolates included TCA cycle, oxidative phosphorylation, and lipid metabolic process, regulation of apoptosis, and autophagy **(Fig. 4C)**. Most of these pathways have previously been implicated in aging and lifespan regulation in laboratory conditions [90].

With regard to metabolite measurements, we identified 11 metabolites that showed correlation with median RLS (*Padj* ≤ 0.01) **(Supplementary Table 3).** Among them, seven metabolites (tryptophan, lactate, 2-hydroxyglutarate, 3-hydroxypropionic acid, 2-hydroxyisobutyrate, 2-hydroxybutyrate, and phenyllactic acid) correlated positively **(Fig. 5A, Supplementary Fig. 7B)** and four metabolites (lysine, quinolinic acid, propionate, Se-methylselenocysteine) showed negative correlation **(Fig. 5B, Supplementary Fig. 7B).** This points to the potential roles of tryptophan metabolism and NAD^+ biosynthesis in natural lifespan variation. Tryptophan degradation, also known as the kynurenine pathway, is required for *de*

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

novo NAD^+ biosynthesis from L-Tryptophan (Trp) [91, 92]. Recent studies have shown that inhibition of Trp degradation increases the lifespan in yeast [93], *C. elegans* [94, 95] and fruit flies [96]. Increased kynurenine pathway activity and alteration in kynurenine pathway metabolites with age have also been observed in rats [97] and humans [98, 99], indicating a conserved role of this pathway in lifespan regulation [100]. More recently, a study across 26 mammalian species also found that species characterized by kynurenine pathway activation were shorter-lived [72]. Interestingly, while our data showed that abundance of Trp correlates positively with RLS, we also found that quinolinic acid, an intermediate in the kynurenine pathway [91, 92], correlates negatively with median RLS, suggesting a possible inhibition of Trp degradation and inactivation of the kynurenine pathway. Similarly, increased quinolinic acid with age was associated with aging and neurodegeneration [98, 99]. Furthermore, our transcriptome data showed that the *BNA2* (indoleamine 2,3-dioxygenase) gene, required for the first rate-limiting step of Trp catabolism in the kynurenine pathway [92], was significantly downregulated (*Padj* = 0.02) in long-lived wild isolates **(Fig. 3B)**.

Other than being an electron carrier in metabolic pathways, NAD^+ is a co-substrate of several NAD^+-dependent enzymes responsible for lysine acetylation (KAC), a conserved protein post-translational modification (PTM) [101, 102]. Redox balance is critical for the availability of NAD^+, and it is directly regulated by mitochondrial respiration [103]. Decreased levels of NAD^+ are associated with aging [104] and age-related diseases, whereas NAD^+ supplementation was found to extend lifespan in mice [105], yeast and worms [104]. Increased activity of NAD^+-dependent histone deacetylase (KDAC), *SIR2* (*SIRT1* in mammals) has been associated with NAD^+-mediated lifespan extension across different organisms [106-108]. Even though our transcriptome data showed no significant association between *SIR2* transcript level and median RLS, *SIR2* activity might be enhanced by a different mechanism. For example, CR induced *PNC1* expression reduces cellular nicotinamide (NAM, an inhibitor of *SIR2* [109]. Decreased concentration of nicotinamide to nicotinic acid (NA) might enhance *SIR2* activity. Interestingly, our transcriptome data revealed a significant increase in mRNA abundance of *PNC1* in long-lived isolates **(Fig. 3A).** It is also possible that the increased *SIR2* activity might not be necessary for NAD+-regulated lifespan extension in wild isolates, as has been suggested previously [110-112]. Although our metabolite data showed no significant changes in NAD^+ level across wild isolates **(Fig. 5B)**, we found the increase in NADH, correlated (*Padj*=0.04) with increased

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

median RLS (**Supplementary Fig. 7B**). However, it has been long established that the assessment of cellular NAD^+ content and the NAD^+/NADH ratio is extremely challenging. They are predominantly bound to intracellular proteins so that the total concentrations determined in protein-free tissue extracts do not reflect true concentrations of these metabolites. In addition, the NAD^+ and NADH contents are likely altered in our samples, because they are highly sensitive to temperature, light, pH, or buffer contents during sample preparation, so that we do not know the exact effect of kynurenine pathway inhibition and increased *PNC1* expression on NAD^+ level [102].

Our metabolite list also included several related short chain fatty acids (SCFAs: 3-hydroxypropionic acid, 2-hydroxyisobutyrate, 2-hydroxybutyrate) and 2-hydroxyglutarate, which are known to be involved in histone modifications. Recently, two new types of post-translational modification of lysine (Lys) residues in histones (H2A, H2B, H3, and H), namely histone propionylation and histone 2-hydroxyisobutyrylation (Khib) (both also known as histone acylation), were identified as an evolutionarily conserved modification from yeast to human [113-115]. Earlier studies have shown that histone modification through Khib is directly related to metabolic regulation and can change the enzymatic activity of glycolytic enzymes to regulate glycolysis in response to the availability of carbon source [116, 117]. For example, 2-hydroxyisobutyrylation on histone H4K8 (H4K8hib) is found to be regulated by glucose availability in the budding yeast, and inhibition of H4K8hib was shown to reduce chronological RLS in yeast [117]. Regulation of H4K8hib in yeast was mediated by histone lysine deacetylases (KDAC) Rpd3p, Hos3p and Esa1p [117, 118]. Rpd3p and Hos3p have both deacetylase and de-2–hydroxyisobutyrylase activities; however, their regulation for selective removal of each group has not been documented. Yeast Esa1p (Tip60 human ortholog) also found to add Khib (2–hydroxyisobutyrylation) to substrate proteins [118]. Along with our metabolomics data, our transcriptomic data supports the distinct regulation of Khib in long-lived isolates, since transcript abundance of both *HOS3* and *ESA1* was found to be positively correlate with median RLS **(Supplementary Fig. 7A)**. Together, these results suggest that NAD^+ homeostasis and a regulatory network involving histone lysine acetylation (KAC) and Khib are involved in lifespan regulation in wild isolates**.** In addition, 2-hydroxyglutarate was found to enhance gene silencing through inhibition of specific H3K36 histone demethylases in yeast [1119]. Overall, the data point to the importance of transcriptional regulation through different types of histone

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

modifications in long-lived isolates. It should be also noted here that proper epigenetic regulation of ribosomal DNA (rDNA) contributes to transcriptional regulation of rDNA (constitute ~80% of the total mass of cellular RNA) and genome stability, which are required for normal life span of yeast [120]. The unstable rDNA region on chromosome XII is organized into a tandem array of 9.1-kb units that are repeated 100 to 200 times. Deacetylation of rDNA region by NAD-dependent histone deacetylase, Sir2, contributes to the rDNA gene silencing and suppress rDNA recombination to prevent formation of extra chromosomal rDNA circles [121, 122]. Increased rDNA repeats and formation of extra chromosomal rDNA circles are known to be a main factor for yeast replicative aging [120-122]. Consistently, our analysis of genomic reads revealed that increased rDNA copy number negatively correlates with median RLS (*Pearson correlation coefficient = -0.3, P = 0.01*) across these wild isolates (**Supplementary Fig. 8**)

A link between our transcriptome and metabolome data could also be seen for the Lys biosynthesis pathway. We found a negative correlation for Lys abundance, (i.e., long-lived strains tend to have less Lys). Our transcriptome data also showed that the two genes controlling the first rate-limiting step of Lys biosynthesis, *LYS20* and *LYS21*, are negatively correlated **(Supplementary Fig. 7A)**, possibly explaining the decreased Lys levels in long-lived strains **(Fig. 5B)**. It should also be noted these two enzymes catalyze condensation of Acetyl-CoA and 2-oxoglutarate (alpha-ketoglutarate) to produce homocitrate and CoA directly regulating Acetyl-CoA pool which is required Lys acetylation [101].

In addition to TCA cycle, oxidative phosphorylation and NAD^+ metabolism, our data point to lipid metabolism, autophagy and apoptotic processes as possible regulators of natural lifespan variation across wild isolates **(Fig. 4C)**. Interestingly, all these processes have some resemblance to pathways, regulated by CR/dietary restriction (DR) [59, 60], further supporting the findings that CR associated mechanisms may be responsible for lifespan extension in long-lived wild isolates. CR has been found to increase the NAD^+/NADH ratio [123-125], and the inhibition of the kynurenine pathway by CR has also been reported [130, 131]. These findings suggest an interesting mechanism by which CR-like physiological stimuli manifest their lifespan effects through increased respiration and kynurenine pathway inhibition. Although the involvement of Khib in CR-mediated lifespan regulation has not yet been thoroughly studied, our findings suggest histone Khib modification might be competing with NAD^+ regulated KAC to influence acetylation balance. Both NAD^+ and Khib have been shown to be regulated by glucose

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

availably [117, 123-125], indicating niche-specific nutritional and environmental cues for the long life of natural isolates. These long-lived strains might be under selection due to repeating seasonal and/or constant starvation cycles, mimicking CR conditions. As a proof of concept, we compared gene expression changes of CR responding strains (i.e. those that showed an increased RLS) and non-responding strains (decreased RLS). We found 207 genes differentially expressed, including 66 that significantly reduced expression and 139 that increased expression in the CR-responding group **(Supplementary Table 4)**. Next, we compared our CR gene list to GenAge data that are based on 112 genes with significant expression changes under CR conditions and/or associated with CR mediated RLS extension in a laboratory WT strain [126] **(Supplementary Table 4).** Interestingly, there were only common 3 genes, including *SCH9* (negatively correlated), which is a functional ortholog of mammalian S6 kinase known to be phosphorylated by Tor1p, *NPT1* (positively correlated), which is a nicotinate phosphoribosyl transferase that acts in the salvage pathway of NAD^+ biosynthesis and *MDH1* (positively correlated), a mitochondrial malate dehydrogenase responsible for interconversion of malate to oxaloacetate using NAD^+/NADH as a cofactor. All three genes were previously associated with the CR mediated RLS extension [111, 127, 128]. Furthermore, correlation of aggregated gene expression patterns (5,376 genes) of wild isolates to gene expression patterns of a laboratory WT strain, grown under CR condition, revealed no significant correlation between any types of pairwise comparison **(Supplementary Fig. 4).** However, our protein-protein interaction network of 207 genes, differentially expressed in CR-responding strains, revealed a highly connected dense network of pathways, such as ribosome (both mitochondrial and cytoplasmic), cell cycle, TCA cycle, oxidative phosphorylation, steroid biosynthesis and Trp metabolism **(Fig. 6).** All these pathways are known to be a hallmark of CR associated longevity mechanism [58-60]. The term “longevity regulating pathway” also appeared with connected 31 nodes of known aging regulating genes (out of 36) **(Fig. 6)**, indicating known longevity regulators are interacting with the genes associated with longevity regulation in wild isolates. Overall, we conclude that evolution of CR dependent longevity follows different genetic trajectories to regulate known genes and related pathways of CR to extend lifespan **(Fig. 6).** Most probably, CR associated genes in wild isolates are favored over through the course of long adaptive evolution under CR like condition induced by the genetic variability that adaptive mechanisms never been evolved under laboratory conditions **(Fig. 7).** It would be also interesting to experiment long term adaptive evolution of

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

laboratory WT strain under CR condition to see if similar genes would be targeted for adapted changes (mutations). Overall, our data indicates that life history trajectories of these strains adapted a mechanism, resembling CR at the level of pathways, although genes associated with these pathways differed from the genes induced by CR in laboratory WT strains.

CONCLUDING COMMENTS

The budding yeast contributed significantly to our understanding of genetics and cell biology and has become an important model of aging, since Mortimer introduced the yeast RLS phenotype [13]. With the power of genetics and experimental tools, yeast has provided clues for understanding the aging process in eukaryotes and yielded hypotheses that can be tested in other organisms, including mammals [129, 130]. However, the yeast aging community has adopted single reference isolates or strains derived from the S288c strain, introduced to aging research by Mortimer in 1959 [13]. He derived the S288c strain from the diploid strain isolated from a rotten fig in California in 1938 [131]. Since then, it has been over 60 years that the S288c strain and its derivatives have been kept in prolonged culture under laboratory conditions that induced relaxed selection for different traits [132]. Relaxed selection might also induce detrimental mutations which may in turn lead to lifespan reduction as has been shown for different organisms [133, 134]. For example, a widely used yeast knockout library has been created from the S288c derivate BY strain background (BY4716, BY4741, BY4742, BY4743), and genome sequencing revealed 39 mutations between the original S288c isolate and the diploid BY4716 strain, a progenitor of the yeast knockout library [135]. In addition, the BY strain series inherited a mutated copy of Hap1 and an allelic variant of MIP1, which are important for mitochondrial respiration and mitochondrial genome stability. Accordingly, this strain manifests several genotype-specific traits, including decreased sporulation and increased rates of mitochondrial genome loss [136]. It is important to note that mitochondria have an important role in a wide variety of metabolic and cellular processes, including energy production, amino acid synthesis, lipid metabolism, cell cycle regulation, apoptosis, autophagy and signaling processes and many of these processes are directly linked to lifespan regulation and aging [137, 138]. Finally, the presence of auxotrophic markers (i.e., mutations that prevent yeast cell unable to synthesize an essential compound such as amino acids) in its genome has been shown to have potential confounding effects for many traits [1136, 139]. These characteristics of S288c and the

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

laboratory adapted BY strain series illustrate that the species has the potential genotype-specific lifespan traits, regulated by particular genes in laboratory settings. But these longevity-associated conditions might be detrimental and not be employed in the natural environment. For example, most of the long-lived yeast mutant strains, from the knockout library have been shown to have reduced fitness relative to isogenic wild-type cells [47]. Similar trade-off was observed across many long-lived laboratory mutants of other model organisms of aging, including nematodes, flies and mice, which may show reduced lifespan, reduced fecundity, slower movement or reduced rates of larval development under naturalistic conditions. These findings indicate that these long-lived strains lack a competitive advantage in the wild and might be eliminated because of selection pressure in their natural environment. These issues of laboratory adapted model organisms have renewed our interest in understanding how the environment modulates lifespan diversity, leading to extended lifespan without significant reduction in fitness or fecundity [45, 46, 140, 141].

From this perspective, the natural isolates we analyze in the current study offer an excellent tool that can serve as a new model for yeast aging studies, allowing to leverage gene-environment impact on lifespan variation. In fact, our comparison of gene expression changes and longevity signatures across laboratory-adapted long-lived mutants and long-lived natural isolates showed that different genes and pathways may be associated with longevity in wild isolates and deletion strains. Interestingly, the pathways associated with lifespan regulation in wild isolates are known to be regulated by CR in laboratory conditions, although the regulation differs at the gene expression level across wild yeast isolates.

Our findings also provide insights into the evolution of adaptive metabolic shift from fermentation to respiration under CR-like condition to regulate NAD^+ metabolism to extend lifespan without decreasing the fitness of strains in the wild **(Fig. 7)**. Although it is unknown if the CR-induced lifespan extension is achieved through common mechanisms in different species, strong conservation of the NAD^+ pathway across different taxa suggests that the same mechanisms regulate NAD^+ homeostasis under CR across other organisms. In addition, our data suggest co-regulation of Lys modification by Khib in long-lived strains, although its association with CR and NAD^+ metabolism needs further investigation. Regulation of NAD^+ and Khib suggests a contribution of epigenetic factors to evolution of gene expression that possibly originated as an adaptation to food availability and other niche-specific environmental

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

conditions. Overall, our research has uncovered one of the experiments of Nature that employs environment to modify genotype and gene expression, arriving at different lifespans. Further understanding of how gene-environment interaction modulates genes and pathways associated with longevity may open new therapeutic applications to slow aging and age-related diseases through diet, lifestyle, or pharmacological interventions.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Materials and Methods

Yeast strains and growth conditions

Many of the diploid wild isolates of *S. cerevisiae* and *S. paradoxus* (68 isolates) were obtained from the Sanger Institute [49] and the remaining (8 isolates) was gifted by Justin Fay from Washington University [48]. Detailed information about strains used in this study is in Supplementary Table 1. The diploid laboratory WT strain BY4743 was purchased from American Type Culture Collection (ATCC). Growth rates were determined using a BioTEK Epoch 2 instrument by the analysis of optical density in the OD_{600} range, and doubling times were calculated with an R script by analyzing fitting spline function from growth curve slopes [142]. The maximum slope of the spline fit was used as an estimate for the growth rate and doubling time for each evolved line, in combination with the YODA software package [143].

Replicative lifespan assay

RLS was determined using a modification of our previously published protocol [144]. Yeast cell cultures for each strain were freshly started from frozen stocks on yeast extract peptone dextrose (YPD) plates and grown for 2 days at 30 ^{0}C prior to dissections. Several colonies were streaked onto new YPD with 2% glucose, YPD with 0.05% glucose or YPG plates with 3% glycerol using pipette tips. After overnight growth, ~100 dividing cells were lined up. After the first division, newborn daughter cells were chosen for RLS assays using a dissection microscope. For each natural isolate, at least two independent assays were performed. Each assay contained 20–80 mother cells of BY4743 strain, which was used in every experiment as a technical control.

RNA-sequencing and data analysis

Three independent cultures for each strain were collected at the OD_{600}=0.4 to isolate DNA and RNA from each culture using Master Pure yeast DNA purification kit (Lucigene, Cat. number: MPY80200) and Quick-RNA 96 Kit from Zymo Research (Cat. number: R1053). Illumina TruSeq DNA and RNA library preparation kit according to user manual and loaded on Illumina HiSeq 4000 platform to produce 150 bp paired-end sequences for both DNA and RNA libraries. The average depth of coverage was ~100X per strain for both DNA and RNA sequencing. After quality control and adapter removal, STAR software package [145] was used to map the reads

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

against a pseudo reference genome of each strains, in which we replaced identified nucleotide changes in S288c reference genome. Genome Analysis Toolkit (GATK) [146] was used to identify SNPs for each strain. Read alignment rate for transcriptome data against pseudo genome was varied between 92-97% across *S. cerevisiae* strains and 93-99% across *S. paradoxus* strains (**Supplementary Fig. 9**). Read counts per gene were calculated using Feature Counts [147]. To filter out genes with the low number of reads, we used filterByExpr function from edgeR package and only the genes with at least 1 count per million (cpm) in at least half of the strains were retained, resulted in an expression set of 5,376 genes across replicates of wild isolates. Filtered data was then subjected to trimmed mean of M-values (TMM) normalization [148]. Differential gene expression analysis was performed using the R package EdgeR [149]. We used Benjamini-Hochberg multiple-testing procedure and selected genes with a false discovery rate (FDR) of 5%.

Metabolite profiling and data analysis

We collected 25 ml cells from the same culture that was used for sample collection for RNA-seq analysis. 1 ml of MeOH:H_2O mixture (8:2, v/v) was added to the samples, swirled at 550 rpm on a mixer for 5 minutes and then transferred to an Eppendorf tube, they were sonicated in an ice bath for 10 min, centrifuged at 4°C at 14,000 rpm for 15 min, and 600 µl of supernatant was collected into a new tube and dried in a vacuum centrifuge at 30 °C for 2.5 hrs. Samples were reconstituted in 1 mL and injected into a chromatography system consisting of a dual injection valve setup allowing injections onto two different LC columns with each column dedicated to an ESI polarity. 5 µL were injected on the positive mode column and 10 µL on the negative side column. The columns were a matched pair from the same production lot number and were both a Waters BEH amide column (2.1 x 150 mm). Auto sampler was maintained at 4 °C and column oven was set to 40 °C. Solvent A (95% H_2O, 3% acetonitrile, 2% methanol, 0.2% acetic acid with 10 mM ammonium acetate and 5 µM medronic acid and Solvent B (5% H2O, 93% acetonitrile, 2% methanol, 0.2% acetic acid with 10 mM ammonium acetate 5 µM medronic acid) were used for sample loading. After completion of the 18-minute gradient, injection on the opposite column was initiated and the inactive column was allowed to equilibrate at starting gradient conditions. A set of QC injections for both instrument and sample QC were run at the beginning and end of the sample run. Data was integrated by Multiquant 3.0.2 software. Peaks

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

were selected based on peak shape, a signal-to-noise of 10 or better and retention times consistent with previously run standards and sample sets. Analysis of the dataset was performed using R (version 3.6.0). All the metabolites with ≥ 40% missingness were excluded, and a total of 166 metabolites were included in the imputation step. We imputed the remaining missing values using the K-nearest neighbors imputation method implemented in the R impute package [150]. The log2-transformed abundance was median normalized prior to imputation.

Principal component analysis (PCA)

Principal component analysis was performed on preprocessed data (e.g. normalized and imputed log2 abundance of the metabolomic data, and the log2-counts per million (CPM) values of the filtered and TMM normalized RNAseq data) using the R prcomp function. To identify the underlying pathways, the factors in each of the first three principal components (PCs) were ranked by their contributions, and pathway enrichment analysis was performed on the top 500 transcripts using Network Analyst [151] and on the top 60 metabolites using Metabol Analyst [152] platforms.

Phylogenetic regression by generalized least squares

R packages 'nmle' and 'phylolm'were used to perform phylogenetic regression by generalized least squares method [153]. We tested four models of trait evolution: (i) complete absence of phylogenetic relationship ('Null'); (ii) Brownian Motion model ('BM'); (iii) BM transformed by Pagel's lambda ('Lambda'); and (iv) Ornstein–Uhlenbeck model ('OU'). The parameters for Lambda and OU models were estimated simultaneously with the coefficients using maximum likelihood. The best-fit model was selected based on maximum likelihood. Strength of correlation was based on the p-value of regression slope. To confirm robustness of results, regression was performed by leaving out each strain, one at a time, and computing *P* values using the remaining strains.

Differential expression between CR responding/non-responding groups

The differential expression of genes between CR-responsive (extending lifespan on CR conditions) and non-responsive (decreasing lifespan on CR conditions) strains of *S. cerevisiae* species were determined using R package limma [154]. Briefly, we used a linear mixed model approach, fitting the CR responding group as a fixed effect and the strain as the random effect

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

where the intra-strain correlation was incorporated into the covariance. To use linear mixed models, we first converted the gene counts to log2 counts per million (log-CPM) using the TMM normalization data and then estimated both observation-level and sample-specific weights using the voom "WithQualityWeights" function from the limma package [154]. The observation-level weights allow us to use a linear mixed model (by accounting for the dependence between mean and variance), and the sample-specific weights enable us to weigh individual samples up or down. Benjamini-Hochberg (BH) adjustment was performed to account for multiple hypotheses [155]. Genes with adjusted p-value < 0.05 were considered significant.

Gene expression signature associated with RLS across deletion strains

Gene expression data on deletion mutants was obtained from GSE45115, GSE42527 and GSE42526 [88]. The corresponding RLS lifespan data for mutant strains was from [89]. Based on the raw data from the number of replicates, we calculated median, mean and maximum RLS, together with corresponding standard errors (SE) for each deletion strain. In total, this resulted in 1,376 deletion strains, for which both RLS and gene expression data were available. logFC of individual genes corresponding to each mutant strain compared to control samples were used for subsequent analysis.

To identify genes associated with RLS across KO strains linear models in limma were used [154]. We found genes associated with median, mean and maximum RLS both in linear and logarithmic scale. Benjamini-Hochberg (BH) adjustment was performed to account for multiple hypotheses [155]. Genes with adjusted p-value < 0.05 were considered significant. To determine statistical significance of the overlap between genes associated with different metrics of RLS, we performed Fisher exact test separately for up- and downregulated genes, considering 6,170 genes as background.

Comparison between signatures of RLS across deletion and natural strains

To compare gene expression signatures associated with different metrics of RLS across deletion and natural *Saccharomyces* strains, we calculated Spearman correlation coefficients between corresponding gene expression slope coefficients in a pairwise manner. Clustering within correlation matrix was performed with complete hierarchical approach and Spearman correlation distance.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To increase the signal of correlation matrix, the union of top 1000 statistically significant genes from each of the two signatures in a pair was used to calculate Spearman correlation coefficient. To get an optimal gene number for removal of noise, we looked how the total number of significantly correlated pairs of signatures depended on the number of genes used to calculate the correlation coefficient. As a threshold, we considered BH adjusted p-value < 0.05 and Spearman correlation coefficient > 0.1.

To determine statistical significance of the overlap between genes associated with different metrics of RLS across deletion and natural strains, we performed Fisher exact test, considering 4,712 genes as background. To identify genes, whose deletions are associated with longer or shorter lifespan in *S. cerevisiae* strains, we compared the distribution of RLS across samples corresponding to certain deletion strains with the distribution of median RLS across all measured deletion strains. For that we used Mann-Whitney test. Genes with BH adjusted p-value < 0.05 were considered significant. Overlap of these genes with lifespan-associated genes across natural strains was assessed with Venn diagram and Fisher exact test p-value (with BH adjusted p-value threshold of 0.05).

Functional enrichment analysis

For the identification of functions enriched by genes associated with RLS across deletion and natural strains, we performed gene set enrichment analysis (GSEA) [168] on a pre-ranked list of genes based on $\log_{10}$(p-value) corrected by the sign of regulation, calculated as:

$$-(pv) \times sgn(b),$$

where *pv* and *b* are p-value and slope of expression of certain gene, respectively, and *sgn* is signum function (is equal to 1, -1 and 0 if value is positive, negative and equal to 0, respectively). REACTOME, KEGG and GO biological process (BP) from Molecular Signature Database (MSigDB) have been used as gene sets for GSEA [156]. We utilized fgsea package in R for GSEA analysis. Adjusted p-value cutoff of 0.1 was used to select statistically significant functions.

We visualized several manually chosen statistically significant functions with a heatmap colored based on normalized enrichment score (NES). Clustering of functions has been performed with hierarchical complete approach and Euclidean distance.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgements

This work was supported by NIH AG067782 and AG064223 to V.N.G. and 1K01AG060040 to A.K. Studies performed by D.P. and M.K. were funded by a University of Washington Nathan Shock Center of Excellence in the Basic Biology pilot award to A. K. (P30AG013280). M.B.L was supported by the HHMI Gilliam Fellowship for Advanced Study and the UW NIH Alzheimer's Disease Training Program (T32 AG052354).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REFERENCES

1. Via, S. and R. Lande, *Genotype-Environment Interaction and the Evolution of Phenotypic Plasticity.* Evolution, 1985. **39**(3): p. 505-522.
2. Li, X., et al., *Genomic and environmental determinants and their interplay underlying phenotypic plasticity.* Proc Natl Acad Sci U S A, 2018. **115**(26): p. 6679-6684.
3. Wagner, G.P. and J. Zhang, *The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms.* Nat Rev Genet, 2011. **12**(3): p. 204-13.
4. Sambandan, D., et al., *Phenotypic plasticity and genotype by environment interaction for olfactory behavior in Drosophila melanogaster.* Genetics, 2008. **179**(2): p. 1079-88.
5. Malagon, J.N., et al., *Evolution of Drosophila sex comb length illustrates the inextricable interplay between selection and variation.* Proc Natl Acad Sci U S A, 2014. **111**(39): p. E4103-9.
6. Caswell, H., *Extrinsic mortality and the evolution of senescence.* Trends Ecol Evol, 2007. **22**(4): p. 173-4.
7. Fusco, G. and A. Minelli, *Phenotypic plasticity in development and evolution: facts and concepts. Introduction.* Philos Trans R Soc Lond B Biol Sci, 2010. **365**(1540): p. 547-56.
8. Bonier, F., et al., *Do baseline glucocorticoids predict fitness?* Trends Ecol Evol, 2009. **24**(11): p. 634-42.
9. Finch, C.E. and M.R. Rose, *Hormones and the physiological architecture of life history evolution.* Q Rev Biol, 1995. **70**(1): p. 1-52.
10. Arking, R., et al., *Factors contributing to the plasticity of the extended longevity phenotypes of Drosophila.* Exp Gerontol, 1996. **31**(6): p. 623-43.
11. Libert, S. and S.D. Pletcher, *Modulation of longevity by environmental sensing.* Cell, 2007. **131**(7): p. 1231-4.
12. Ratikainen, II and H. Kokko, *The coevolution of lifespan and reversible plasticity.* Nat Commun, 2019. **10**(1): p. 538.
13. Mortimer, R.K. and J.R. Johnston, *Life span of individual yeast cells.* Nature, 1959. **183**(4677): p. 1751-2.
14. Nielsen, J., et al., *Eye lens radiocarbon reveals centuries of longevity in the Greenland shark (Somniosus microcephalus).* Science, 2016. **353**(6300): p. 702-4.
15. Jones, O.R., et al., *Diversity of ageing across the tree of life.* Nature, 2014. **505**(7482): p. 169-73.
16. Finch, C.E. and M.C. Pike, *Maximum life span predictions from the Gompertz mortality model.* J Gerontol A Biol Sci Med Sci, 1996. **51**(3): p. B183-94.
17. Klass, M.R., *A method for the isolation of longevity mutants in the nematode Caenorhabditis elegans and initial results.* Mech Ageing Dev, 1983. **22**(3-4): p. 279-86.
18. Kaya, A., et al., *Defining Molecular Basis for Longevity Traits in Natural Yeast Isolates.* NPJ Aging Mech Dis, 2015. **1**.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

19. Ma, S., et al., *Comparative transcriptomics across 14 Drosophila species reveals signatures of longevity.* Aging Cell, 2018. **17**(4): p. e12740.

20. Dato, S., et al., *The genetic component of human longevity: New insights from the analysis of pathway-based SNP-SNP interactions.* Aging Cell, 2018. **17**(3): p. e12755.

21. Wright, K.M., et al., *A Prospective Analysis of Genetic Variants Associated with Human Lifespan.* G3 (Bethesda), 2019. **9**(9): p. 2863-2878.

22. de Magalhaes, J.P., et al., *Genome-environment interactions that modulate aging: powerful targets for drug discovery.* Pharmacol Rev, 2012. **64**(1): p. 88-101.

23. Li, Y. and J.P. de Magalhaes, *Accelerated protein evolution analysis reveals genes and pathways associated with the evolution of mammalian longevity.* Age (Dordr), 2013. **35**(2): p. 301-14.

24. Ma, S. and V.N. Gladyshev, *Molecular signatures of longevity: Insights from cross-species comparative studies.* Semin Cell Dev Biol, 2017. **70**: p. 190-203.

25. van Dongen, J., et al., *Genetic and environmental influences interact with age and sex in shaping the human methylome.* Nat Commun, 2016. **7**: p. 11115.

26. Dato, S., et al., *The impact of nutrients on the aging rate: A complex interaction of demographic, environmental and genetic factors.* Mech Ageing Dev, 2016. **154**: p. 49-61.

27. Dillon, M.E., L.R. Cahn, and R.B. Huey, *Life history consequences of temperature transients in Drosophila melanogaster.* J Exp Biol, 2007. **210**(Pt 16): p. 2897-904.

28. Johnston, R.K. and T.W. Snell, *Moderately lower temperatures greatly extend the lifespan of Brachionus manjavacas (Rotifera): Thermodynamics or gene regulation?* Exp Gerontol, 2016. **78**: p. 12-22.

29. Packer, L. and K. Fuehr, *Low oxygen concentration extends the lifespan of cultured human diploid cells.* Nature, 1977. **267**(5610): p. 423-5.

30. Bisschops, M.M., et al., *Oxygen availability strongly affects chronological lifespan and thermotolerance in batch cultures of Saccharomyces cerevisiae.* Microb Cell, 2015. **2**(11): p. 429-444.

31. Alic, N. and L. Partridge, *Death and dessert: nutrient signalling pathways and ageing.* Curr Opin Cell Biol, 2011. **23**(6): p. 738-43.

32. Stefana, M.I., et al., *Developmental diet regulates Drosophila lifespan via lipid autotoxins.* Nat Commun, 2017. **8**(1): p. 1384.

33. Kitada, M., et al., *The impact of dietary protein intake on longevity and metabolic health.* EBioMedicine, 2019. **43**: p. 632-640.

34. Bitto, A., et al., *Transient rapamycin treatment can increase lifespan and healthspan in middle-aged mice.* Elife, 2016. **5**.

35. Novelle, M.G., et al., *Metformin: A Hopeful Promise in Aging Research.* Cold Spring Harb Perspect Med, 2016. **6**(3): p. a025932.

36. Rajman, L., K. Chwalek, and D.A. Sinclair, *Therapeutic Potential of NAD-Boosting Molecules: The In Vivo Evidence.* Cell Metab, 2018. **27**(3): p. 529-547.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

37. Nadon, N.L., et al., *NIA Interventions Testing Program: Investigating Putative Aging Intervention Agents in a Genetically Heterogeneous Mouse Model.* EBioMedicine, 2017. **21**: p. 3-4.

38. Zwaan, B.J., *The evolutionary genetics of ageing and longevity.* Heredity (Edinb), 1999. **82 (Pt 6)**: p. 589-97.

39. Vijg, J. and Y. Suh, *Genetics of longevity and aging.* Annu Rev Med, 2005. **56**: p. 193-212.

40. Singh, P.P., et al., *The Genetics of Aging: A Vertebrate Perspective.* Cell, 2019. **177**(1): p. 200-220.

41. Harper, J.M., *Wild-derived mouse stocks: an underappreciated tool for aging research.* Age (Dordr), 2008. **30**(2-3): p. 135-45.

42. de Magalhaes, J.P., *Why genes extending lifespan in model organisms have not been consistently associated with human longevity and what it means to translation research.* Cell Cycle, 2014. **13**(17): p. 2671-3.

43. Reichard, M., *Evolutionary ecology of aging: time to reconcile field and laboratory research.* Ecol Evol, 2016. **6**(9): p. 2988-3000.

44. Bartke, A., et al., *Genes that prolong life: relationships of growth hormone and growth to aging and life span.* J Gerontol A Biol Sci Med Sci, 2001. **56**(8): p. B340-9.

45. Jenkins, N.L., G. McColl, and G.J. Lithgow, *Fitness cost of extended lifespan in Caenorhabditis elegans.* Proc Biol Sci, 2004. **271**(1556): p. 2523-6.

46. Chen, J., et al., *A demographic analysis of the fitness cost of extended longevity in Caenorhabditis elegans.* J Gerontol A Biol Sci Med Sci, 2007. **62**(2): p. 126-35.

47. Delaney, J.R., et al., *Quantitative evidence for early life fitness defects from 32 longevity-associated alleles in yeast.* Cell Cycle, 2011. **10**(1): p. 156-65.

48. Hyma, K.E. and J.C. Fay, *Mixing of vineyard and oak-tree ecotypes of Saccharomyces cerevisiae in North American vineyards.* Mol Ecol, 2013. **22**(11): p. 2917-30.

49. Liti, G., et al., *Population genomics of domestic and wild yeasts.* Nature, 2009. **458**(7236): p. 337-41.

50. Liti, G., D.B. Barton, and E.J. Louis, *Sequence diversity, reproductive isolation and species concepts in Saccharomyces.* Genetics, 2006. **174**(2): p. 839-50.

51. Tirosh, I., et al., *A yeast hybrid provides insight into the evolution of gene expression regulation.* Science, 2009. **324**(5927): p. 659-62.

52. Liti, G., *The fascinating and secret wild life of the budding yeast S. cerevisiae.* Elife, 2015. **4**.

53. Saltz, J.B., et al., *Why does the magnitude of genotype-by-environment interaction vary?* Ecol Evol, 2018. **8**(12): p. 6342-6353.

54. Burtner, C.R., et al., *A molecular mechanism of chronological aging in yeast.* Cell Cycle, 2009. **8**(8): p. 1256-70.

55. Wei, M., et al., *Tor1/Sch9-regulated carbon source substitution is as effective as calorie restriction in life span extension.* PLoS Genet, 2009. **5**(5): p. e1000467.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

56. Delaney, J.R., et al., *Sir2 deletion prevents lifespan extension in 32 long-lived mutants.* Aging Cell, 2011. **10**(6): p.1089-91.

57. Schleit, J., et al., *Molecular mechanisms underlying genotype-dependent responses to dietary restriction.* Aging Cell, 2013. **12**(6): p. 1050-61.

58. Kaeberlein, M., et al., Increased life span due to calorie restriction in respiratory-deficient yeast. PLoS Genet, 2005. **1**(5): p. e69.

59. Koubova, J. and L. Guarente, *How does calorie restriction work?* Genes Dev, 2003. **17**(3): p. 313-21.

60. Ungvari, Z., et al., *Mechanisms underlying caloric restriction and lifespan regulation: implications for vascular aging.* Circ Res, 2008. **102**(5): p. 519-28.

61. Lin, S.J., et al., *Calorie restriction extends Saccharomyces cerevisiae lifespan by increasing respiration.* Nature, 2002. **418**(6895): p. 344-8.

62. Stumpferl, S.W., et al., *Natural genetic variation in yeast longevity.* Genome Res, 2012. **22**(10): p. 1963-73.

63. Burke, M.K., et al., *Genome-wide association study of extreme longevity in Drosophila melanogaster.* Genome Biol Evol, 2014. **6**(1): p. 1-11.

64. Wilson, K.A., et al., *GWAS for Lifespan and Decline in Climbing Ability in Flies upon Dietary Restriction Reveal decima as a Mediator of Insulin-like Peptide Production.* Curr Biol, 2020.

65. Deelen, J., et al., *A meta-analysis of genome-wide association studies identifies multiple longevity genes.* Nat Commun, 2019. **10**(1): p. 3669.

66. Fushan, A.A., et al., *Gene expression defines natural changes in mammalian lifespan.* Aging Cell, 2015. **14**(3): p. 352-65.

67. Seim, I., S. Ma, and V.N. Gladyshev, *Gene expression signatures of human cell and tissue longevity.* NPJ Aging Mech Dis, 2016. **2**: p. 16014.

68. Yang, S., et al., *Comparative proteomic analysis of brains of naturally aging mice.* Neuroscience, 2008. **154**(3): p. 1107-20.

69. Tanaka, T., et al., *Plasma proteomic signature of age in healthy humans.* Aging Cell, 2018. **17**(5): p. e12799.

70. Heinze, I., et al., *Species comparison of liver proteomes reveals links to naked mole-rat longevity and human aging.* BMC Biol, 2018. **16**(1): p. 82.

71. Laye, M.J., et al., *The effects of age and dietary restriction on the tissue-specific metabolome of Drosophila.* Aging Cell, 2015. **14**(5): p. 797-808.

72. Ma, S., et al., *Organization of the Mammalian Metabolome according to Organ Function, Lineage Specialization, and Longevity.* Cell Metab, 2015. **22**(2): p. 332-43.

73. Cheng, S., et al., *Distinct metabolomic signatures are associated with longevity in humans.* Nat Commun, 2015. **6**: p. 6791.

74. Whitehead, A. and D.L. Crawford, *Neutral and adaptive variation in gene expression.* Proc Natl Acad Sci U S A, 2006. **103**(14): p. 5425-30.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

75. Vu, V., et al., *Natural Variation in Gene Expression Modulates the Severity of Mutant Phenotypes.* Cell, 2015. **162**(2): p. 391-402.

76. Gilad, Y., A. Oshlack, and S.A. Rifkin, *Natural selection on gene expression.* Trends Genet, 2006. **22**(8): p. 456-61.

77. Holloway, A.K., et al., *Adaptive gene expression divergence inferred from population genomics.* PLoS Genet, 2007. **3**(10): p. 2007-13.

78. Thompson, D.A., et al., *Evolutionary principles of modular gene regulation in yeasts.* Elife, 2013. **2**: p. e00603.

79. Beldade, P., P.M. Brakefield, and A.D. Long, *Contribution of Distal-less to quantitative variation in butterfly eyespots.* Nature, 2002. **415**(6869): p. 315-8.

80. Whitaker, R., et al., *Dietary switch reveals fast coordinated gene expression changes in Drosophila melanogaster.* Aging (Albany NY), 2014. **6**(5): p. 355-68.

81. Guijas, C., et al., *Metabolomics activity screening for identifying metabolites that modulate phenotype.* Nat Biotechnol, 2018. **36**(4): p. 316-320.

82. Harrison, B.R., et al., *The metabolome as a link in the genotype-phenotype map for peroxide resistance in the fruit fly, Drosophila melanogaster.* BMC Genomics, 2020. **21**(1): p. 341.

83. Brawand, D., et al., *The evolution of gene expression levels in mammalian organs.* Nature, 2011. **478**(7369): p. 343-8.

84. Uyeda, J.C., D.S. Caetano, and M.W. Pennell, *Comparative Analysis of Principal Components Can be Misleading.* Syst Biol, 2015. **64**(4): p. 677-89.

85. Felsenstein, J., *Phylogenies and the Comparative Method.* The American Naturalist, 1985. **125**(1): p. 1-15.

86. R. P. Freckleton, P. H. Harvey, and M. Pagel, *Phylogenetic Analysis and Comparative Data: A Test and Review of Evidence.* The American Naturalist, 2002. **160**(6): p. 712-726.

87. de Magalhaes, J.P., J. Curado, and G.M. Church, *Meta-analysis of age-related gene expression profiles identifies common signatures of aging.* Bioinformatics, 2009. **25**(7): p. 875-81.

88. Kemmeren, P., et al., *Large-scale genetic perturbations reveal regulatory networks and an abundance of gene-specific repressors.* Cell, 2014. **157**(3): p. 740-52.

89. McCormick, M.A., et al., *A Comprehensive Analysis of Replicative Lifespan in 4,698 Single-Gene Deletion Strains Uncovers Conserved Mechanisms of Aging.* Cell Metab, 2015. **22**(5): p. 895-906.

90. Denoth, L.A., T. Julou, and Y. Barral, Budding yeast as a model organism to study the effects of age. FEMS Microbiol Rev, 2014. **38**(2): p. 300-25

91. Platten, M., et al., *Tryptophan metabolism as a common therapeutic target in cancer, neurodegeneration and beyond.* Nat Rev Drug Discov, 2019. **18**(5): p. 379-401.

92. Pinson, B., et al., *Dual control of NAD(+) synthesis by purine metabolites in yeast.* Elife, 2019. **8**.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

93. He, C., et al., *Enhanced longevity by ibuprofen, conserved in multiple species, occurs in yeast through inhibition of tryptophan import.* PLoS Genet, 2014. **10**(12): p. e1004860.

94. Edwards, C., et al., *Mechanisms of amino acid-mediated lifespan extension in Caenorhabditis elegans.* BMC Genet, 2015. **16**: p. 8.

95. van der Goot, A.T., et al., *Delaying aging and the aging-associated decline in protein homeostasis by inhibition of tryptophan degradation.* Proc Natl Acad Sci U S A, 2012. **109**(37): p. 14912-7.

96. Oxenkrug, G.F., *The extended life span of Drosophila melanogaster eye-color (white and vermilion) mutants with impaired formation of kynurenine.* J Neural Transm (Vienna), 2010. **117**(1): p. 23-6.

97. Braidy, N., et al., *Changes in kynurenine pathway metabolism in the brain, liver and kidney of aged female Wistar rats.* FEBS J, 2011. **278**(22): p. 4425-34.

98. de Bie, J., et al., *Central kynurenine pathway shift with age in women.* J Neurochem, 2016. **136**(5): p. 995-1003.

99. Chatterjee, P., et al., *Alterations in serum kynurenine pathway metabolites in individuals with high neocortical amyloid-beta load: A pilot study.* Sci Rep, 2018. **8**(1): p. 8008.

100. Castro-Portuguez, R. and G.L. Sutphin, *Kynurenine pathway, NAD(+) synthesis, and mitochondrial function: Targeting tryptophan metabolism to promote longevity and healthspan.* Exp Gerontol, 2020. **132**: p. 110841.

101. Choudhary, C., et al., *The growing landscape of lysine acetylation links metabolism and cell signalling.* Nat Rev Mol Cell Biol, 2014. **15**(8): p. 536-50.

102. Houtkooper, R.H., et al., *The secret life of NAD+: an old metabolite controlling new metabolic signaling pathways.* Endocr Rev, 2010. **31**(2): p. 194-223.

103. Zhu, X.H., et al., *In vivo NAD assay reveals the intracellular NAD contents and redox state in healthy human brain and their age dependences.* Proc Natl Acad Sci U S A, 2015. **112**(9): p. 2876-81.

104. Verdin, E., *NAD(+) in aging, metabolism, and neurodegeneration.* Science, 2015. **350**(6265): p. 1208-13.

105. Zhang, H., et al., *NAD(+) repletion improves mitochondrial and stem cell function and enhances life span in mice.* Science, 2016. **352**(6292): p. 1436-43.

106. Kaeberlein, M., McVey, M, and L. Guarente, The SIR2/3/4 complex and SIR2 alone promote longevity in Saccharomyces cerevisiae by two different mechanisms. Genes Dev. 1999. **13**(19): p. 2570-80.

107. Lin, S.J., P.A. Defossez, and L. Guarente, *Requirement of NAD and SIR2 for life-span extension by calorie restriction in Saccharomyces cerevisiae.* Science, 2000. **289**(5487): p. 2126-8.

108. Imai, S., et al., *Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase.* Nature, 2000. **403**(6771): p. 795-800.

109. Anderson, R.M., et al., *Nicotinamide and PNC1 govern lifespan extension by calorie restriction in Saccharomyces cerevisiae.* Nature, 2003. **423**(6936): p. 181-5.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

110. Kaeberlein, M., et al., *Sir2-independent life span extension by calorie restriction in yeast.* PLoS Biol, 2004. **2**(9): p. E296.

111. Kaeberlein, M., et al., *Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients.* Science, 2005. **310**(5751): p. 1193-6.

112. Tsuchiya, M., et al., *Sirtuin-independent effects of nicotinamide on lifespan extension from calorie restriction in yeast.* Aging Cell, 2006. **5**(6): p. 505-14.

113. Dai, L., et al., *Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark.* Nat Chem Biol, 2014. **10**(5): p. 365-70.

114. Zhang, K., et al., *Identification and verification of lysine propionylation and butyrylation in yeast core histones using PTMap software.* J Proteome Res, 2009. **8**(2): p. 900-6.

115. Chen, Y., et al., *Lysine propionylation and butyrylation are novel post-translational modifications in histones.* Mol Cell Proteomics, 2007. **6**(5): p. 812-9.

116. Dong, H., et al., *Protein lysine de-2-hydroxyisobutyrylation by CobB in prokaryotes.* Sci Adv, 2019. **5**(7): p. eaaw6703.

117. Huang, J., et al., *2-Hydroxyisobutyrylation on histone H4K8 is regulated by glucose homeostasis in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 2017. **114**(33): p. 8782-8787.

118. Huang, H., et al., *Landscape of the regulatory elements for lysine 2-hydroxyisobutyrylation pathway.* Cell Res, 2018. **28**(1): p. 111-125.

119. Janke, R., A.T. Iavarone, and J. Rine, *Oncometabolite D-2-Hydroxyglutarate enhances gene silencing through inhibition of specific H3K36 histone demethylases.* Elife, 2017. **6**.

120. Sinclair, D.A. and L. Guarente, *Extrachromosomal rDNA circles--a cause of aging in yeast.* Cell, 1997. **91**(7): p. 1033-42.

121. Guarente, L., *Sir2 links chromatin silencing, metabolism, and aging.* Genes Dev, 2000. **14**(9): p. 1021-6.

122. Kwan, E.X., *A natural polymorphism in rDNA replication origins links origin activation with calorie restriction and lifespan.* PLoS Genet, 2013. **9**(3): e1003329.

123. Canto, C., et al., *AMPK regulates energy expenditure by modulating NAD+ metabolism and SIRT1 activity.* Nature, 2009. **458**(7241): p. 1056-60.

124. Canto, C., et al., *Interdependence of AMPK and SIRT1 for metabolic adaptation to fasting and exercise in skeletal muscle.* Cell Metab, 2010. **11**(3): p. 213-9.

125. Lu, S.P. and S.J. Lin, *Regulation of yeast sirtuins by NAD(+) metabolism and calorie restriction.* Biochim Biophys Acta, 2010. **1804**(8): p. 1567-75.

126. Wuttke, D., et al., *Dissecting the gene network of dietary restriction to identify evolutionarily conserved pathways and new functional genes.* PLoS Genet, 2012. **8**(8): p. e1002834.

127. Imai, S.I. and L. Guarente, *It takes two to tango: NAD(+) and sirtuins in aging/longevity control.* NPJ Aging Mech Dis, 2016. **2**: p. 16017.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

128. Easlon, E., et al., *The malate-aspartate NADH shuttle components are novel metabolic longevity regulators required for calorie restriction-mediated life span extension in yeast.* Genes Dev, 2008. **22**(7): p. 931-44.

129. Kennedy, B.K., *The genetics of ageing: insight from genome-wide approaches in invertebrate model organisms.* J Intern Med, 2008. **263**(2): p. 142-52.

130. Fontana, L., L. Partridge, and V.D. Longo, *Extending healthy life span--from yeast to humans.* Science, 2010. **328**(5976): p. 321-6.

131. Mortimer, R.K. and J.R. Johnston, *Genealogy of principal strains of the yeast genetic stock center.* Genetics, 1986. **113**(1): p. 35-43.

132. Pizarro, F., F.A. Vargas, and E. Agosin, *A systems biology perspective of wine fermentations.* Yeast, 2007. **24**(11): p. 977-91.

133. Cui, R., et al., *Relaxed Selection Limits Lifespan by Increasing Mutation Load.* Cell, 2019. **178**(2): p. 385-399 e20.

134. Promislow D.E. and M. Tatar, *Mutation and senescence: where genetics and demography meet.* Genetica. 1998. **103**(1-6): p. 299-314.

135. Schacherer, J., et al., *Genome-wide analysis of nucleotide-level variation in commonly used Saccharomyces cerevisiae strains.* PLoS One, 2007. **2**(3): p. e322.

136. Young, M.J. and D.A. Court, *Effects of the S288c genetic background and common auxotrophic markers on mitochondrial DNA function in Saccharomyces cerevisiae.* Yeast, 2008. **25**(12): p. 903-12.

137. Wang, Y. and S. Hekimi, *Mitochondrial dysfunction and longevity in animals: Untangling the knot.* Science, 2015. **350**(6265): p. 1204-7.

138. Jang, J.Y., et al., *The role of mitochondria in aging.* J Clin Invest, 2018. **128**(9): p. 3662-3670.

139. Mulleder, M., et al., *A prototrophic deletion mutant collection for yeast metabolomics and systems biology.* Nat Biotechnol, 2012. **30**(12): p. 1176-8.

140. Maklakov, A.A. and T. Chapman, *Evolution of ageing as a tangle of trade-offs: energy versus function.* Proc Biol Sci, 2019. **286**(1911): p. 20191604.

141. Ramani, A.K., et al., *The majority of animal genes are required for wild-type fitness.* Cell, 2012. **148**(4): p. 792-802.

142. Kahm, M., et al., *grofit: Fitting Biological Growth Curves with R.* 2010, 2010. **33**(7): p. 21.

143. Olsen, B., C.J. Murakami, and M. Kaeberlein, *YODA: software to facilitate high-throughput analysis of chronological life span, growth rate, and survival in budding yeast.* BMC Bioinformatics, 2010. **11**: p. 141.

144. Steffen, K.K., B.K. Kennedy, and M. Kaeberlein, *Measuring replicative life span in the budding yeast.* J Vis Exp, 2009. **25**(28): p. 1209.

145. Dobin, A. and T.R. Gingeras, *Mapping RNA-seq Reads with STAR.* Curr Protoc Bioinformatics, 2015. **51**: p. 11 14 1-11 14 19.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

146. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.* Genome Res, 2010. **20**(9): p. 1297-303.

147. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.* Bioinformatics, 2014. **30**(7): p. 923-30.

148. Maza, E., et al., *Comparison of normalization methods for differential gene expression analysis in RNA-Seq experiments: A matter of relative size of studied transcriptomes.* Commun Integr Biol, 2013. **6**(6): p. e25849.

149. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-40.

150. Troyanskaya, O., et al., *Missing value estimation methods for DNA microarrays.* Bioinformatics, 2001. **17**(6): p. 520-5.

151. Zhou, G., et al., *NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis.* Nucleic Acids Res, 2019. **47**(W1): p. W234-W241.

152. Chong, J., D.S. Wishart, and J. Xia, *Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis.* Curr Protoc Bioinformatics, 2019. **68**(1): p. e86.

153. Ho, L. and C. Ane, *A linear-time algorithm for Gaussian and non-Gaussian trait evolution models.* Syst Biol, 2014. **63**(3): p. 397-408.

154. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies.* Nucleic Acids Res, 2015. **43**(7): p. e47.

155. Benjamini, Y. and Y. Hochberg, *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.* Journal of the Royal Statistical Society. Series B (Methodological), 1995. **57**(1): p. 289-300.

156. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

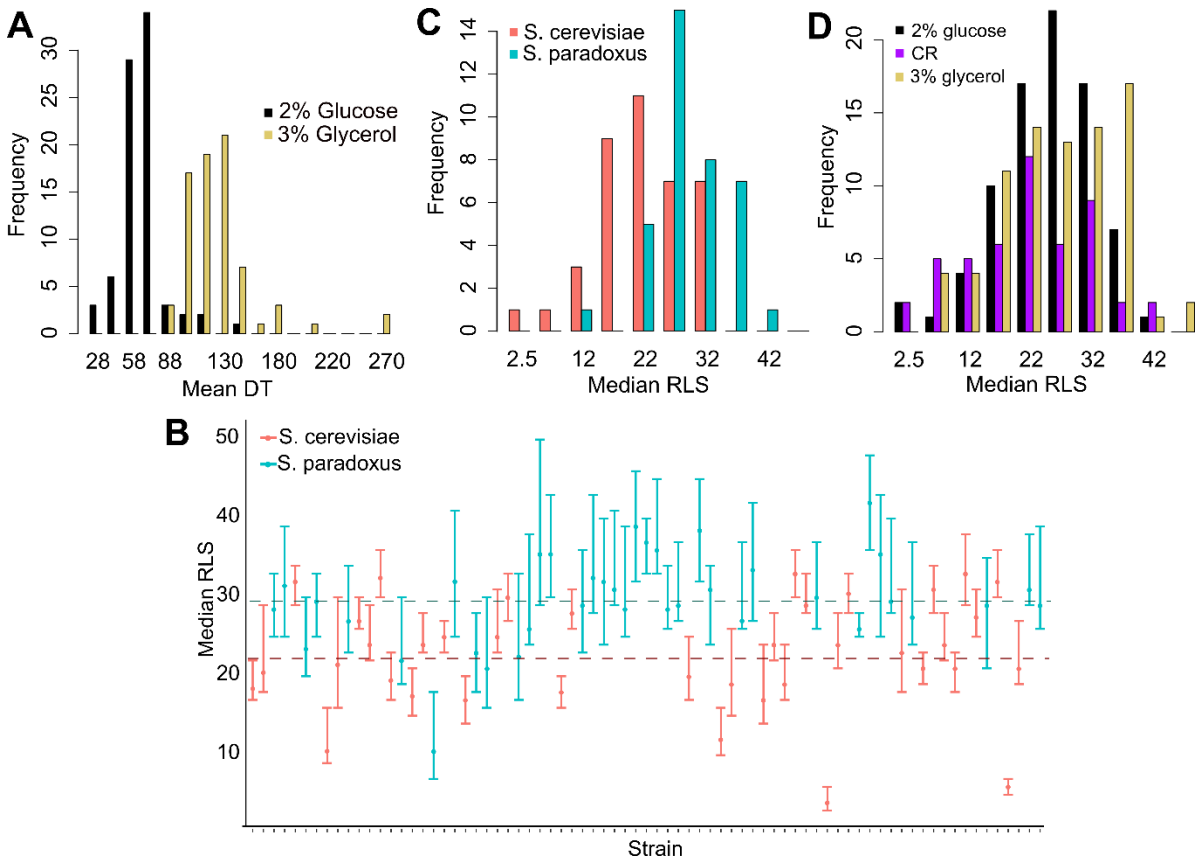


Figure 1: Doubling time and replicative lifespan of yeast wild isolates. A) Distribution of mean doubling time on YPD (2% glucose) and YPG (3% glycerol). **B)** Median RLS (Error bars" are the 95%CI of the median lifespan) and **C)** its distribution across *S. cerevisiae* and *S. paradoxus* isolates grown in YPD. Dashed lines represent average median RLS of *S. cerevisiae* (red) and *S. paradoxus* strains (turquoise). **D)** Distribution of median RLS across different conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

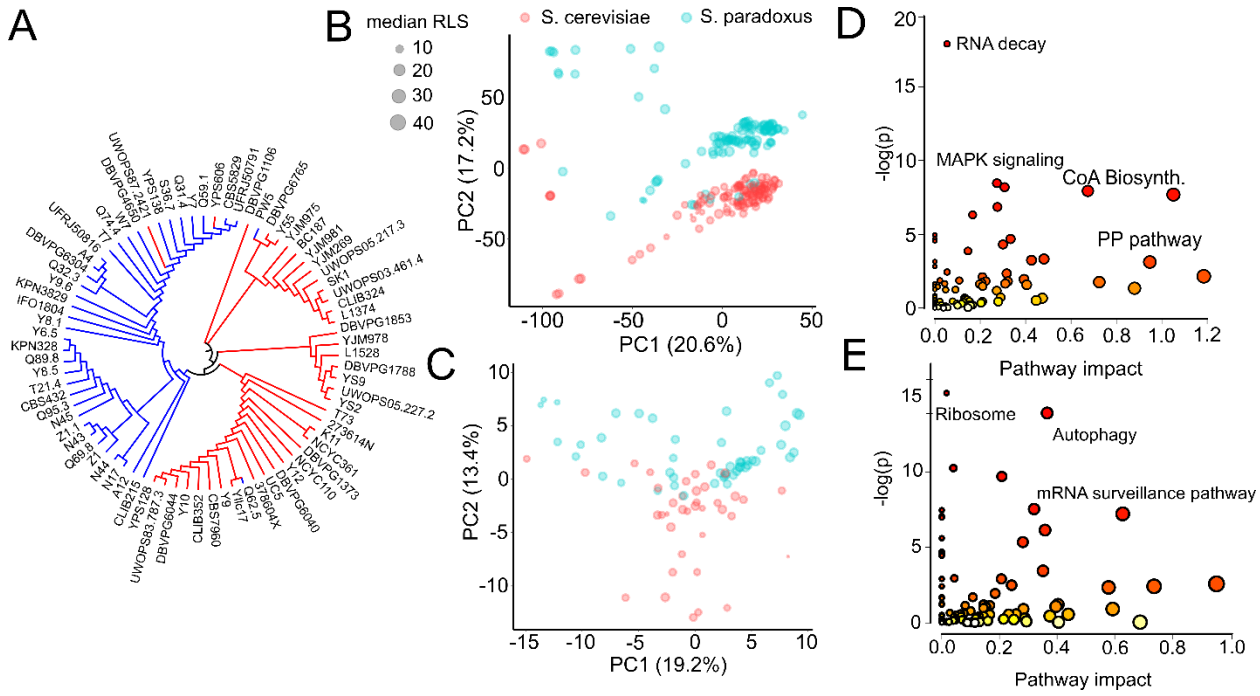


Figure 2: Endophenotypic variation across strains. (A) Phylogenetic relationship based on the transcriptome data of 76 strains. Principal component analysis (PCA) of **(B)** transcriptomics and **(C)** metabolomics. Percent variance explained by each principal component (PC) is shown in parentheses. Pathway enrichment analysis for combined top genes and metabolites contributing to **(D)** PC1 and **(E)** PC2. Some of the enriched KEGG pathways are shown in each panel. The whole list of enriched pathways can be found in Supplementary Table 2.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

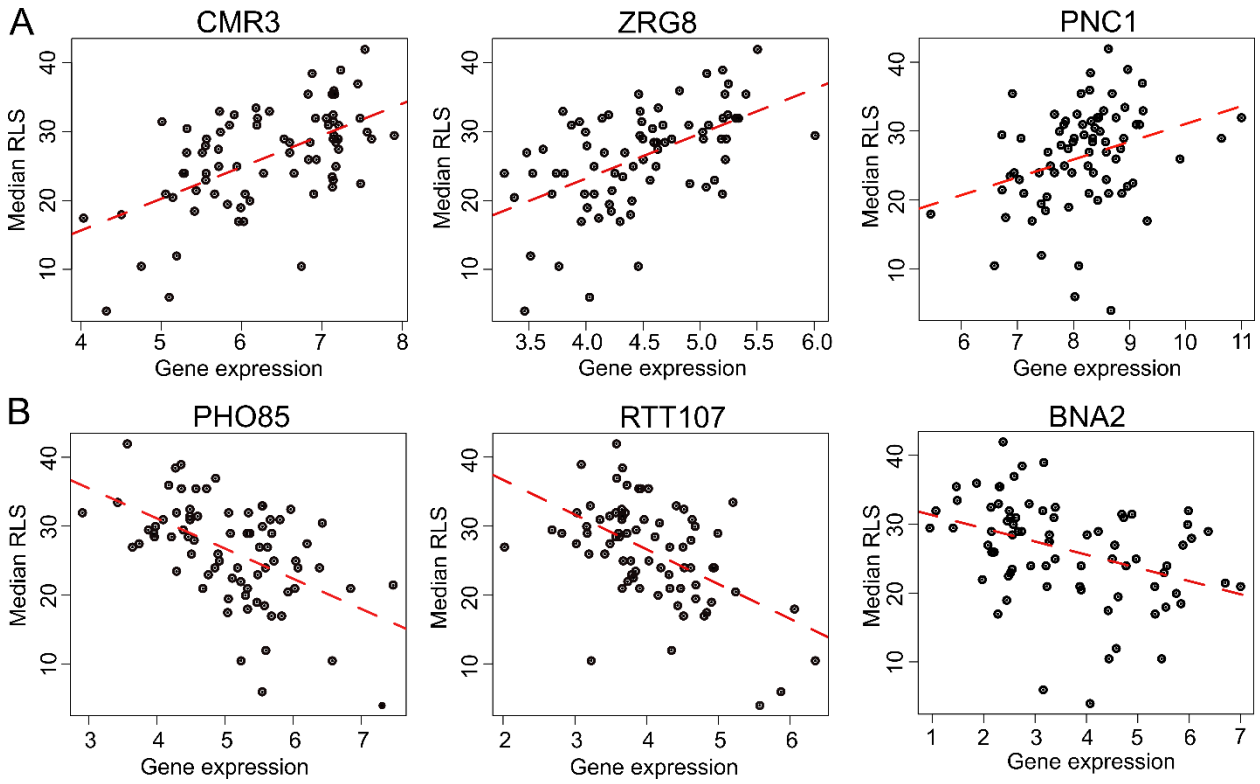


Figure 3: Selected genes whose expression correlates with median RLS. (A) Transcript abundance of *CMR3*, *ZRG8*, and *PNC1* that positively correlate with median RLS, **(B)** transcript abundance of *PHO85*, *RTT107*, and *BNA2* that negatively correlate with median RLS. Regression slope *P* values can be found in Supplementary Table 3.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

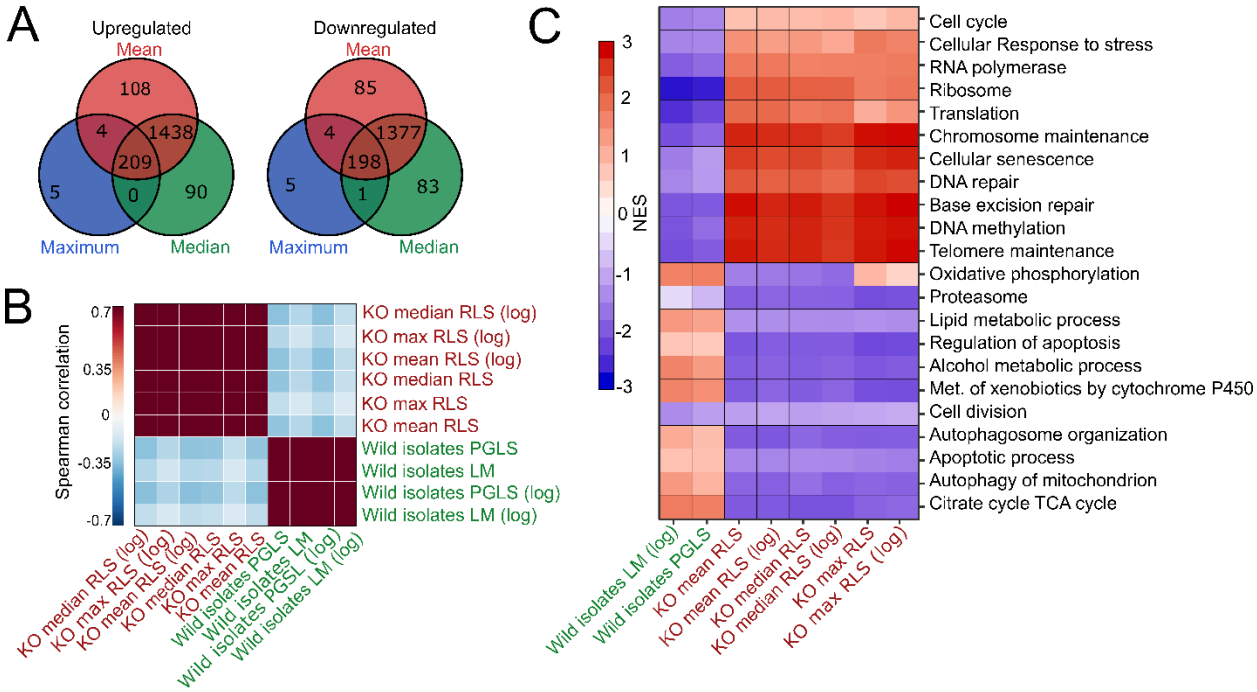


Figure 4: Comparative analysis of lab yeast KO and wild isolates. (A) Significant genes (BH adj. p-value < 0.05) associated with maximum, median and mean RLS across deletion strains based on transcriptomics data obtained from 1,376 KO strains. Genes positively and negatively associated with RLS are significantly shared across different metrics of RLS (BH adj. Fisher exact test p-value < 0.05). **(B)** Denoised correlation matrix of gene expression changes associated with RLS across deletion and natural isolates. Correlation coefficient is calculated using union of top 1,000 statistically significant genes for each pair of signatures with Spearman method. LM: Linear model; PGLS: phylogenetic regression least squares. **(C)** Functional enrichment (GSEA) of genes associated with RLS across deletion and natural strains. Cells are colored based on normalized enrichment score (NES). The whole list of enriched functions can be found in Supplementary Table 1.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

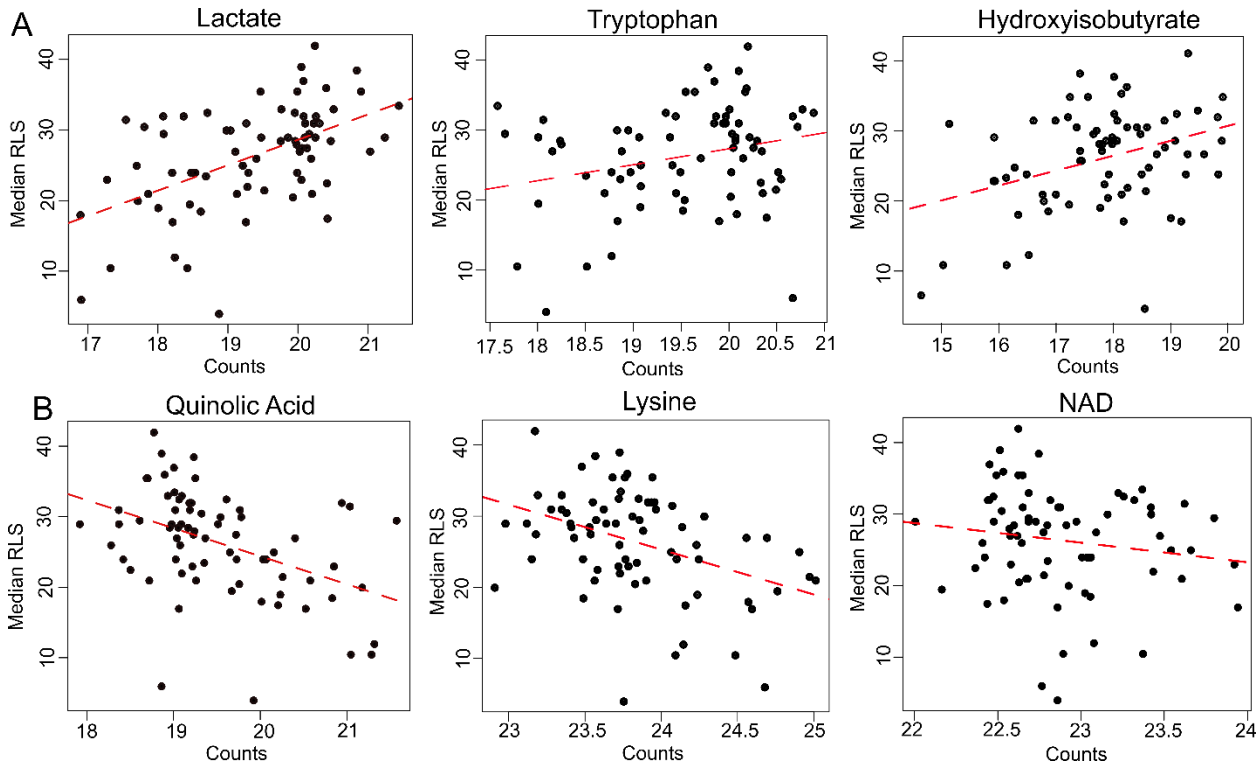


Figure 5: Selected metabolites correlating with median RLS. (A) Abundance of lactate, Trp, and hydroxyisobutyrate that positively correlate, and **(B)** abundance of quinolic acid, Lys, and NAD that negatively correlate with median RLS. Regression slope *P* values can be found in Supplementary Table 3.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

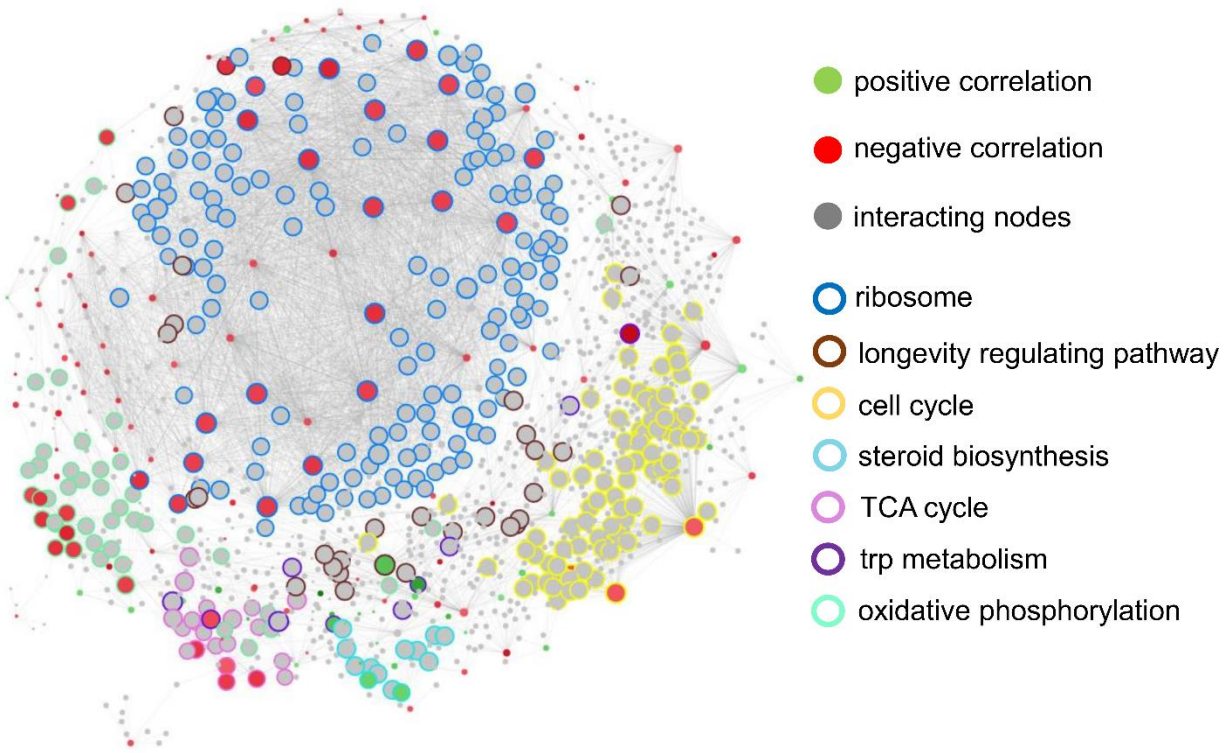


Figure 6: Protein-protein interaction network of genes, differentially expressed in strains responding to caloric restriction. Red nodes indicate decreased expression and green nodes increased expression for the network hub. Gray nodes are interacting noises of differentially expressed genes. Pathway enrichment analyses based on the network nodes are shown with different colors.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

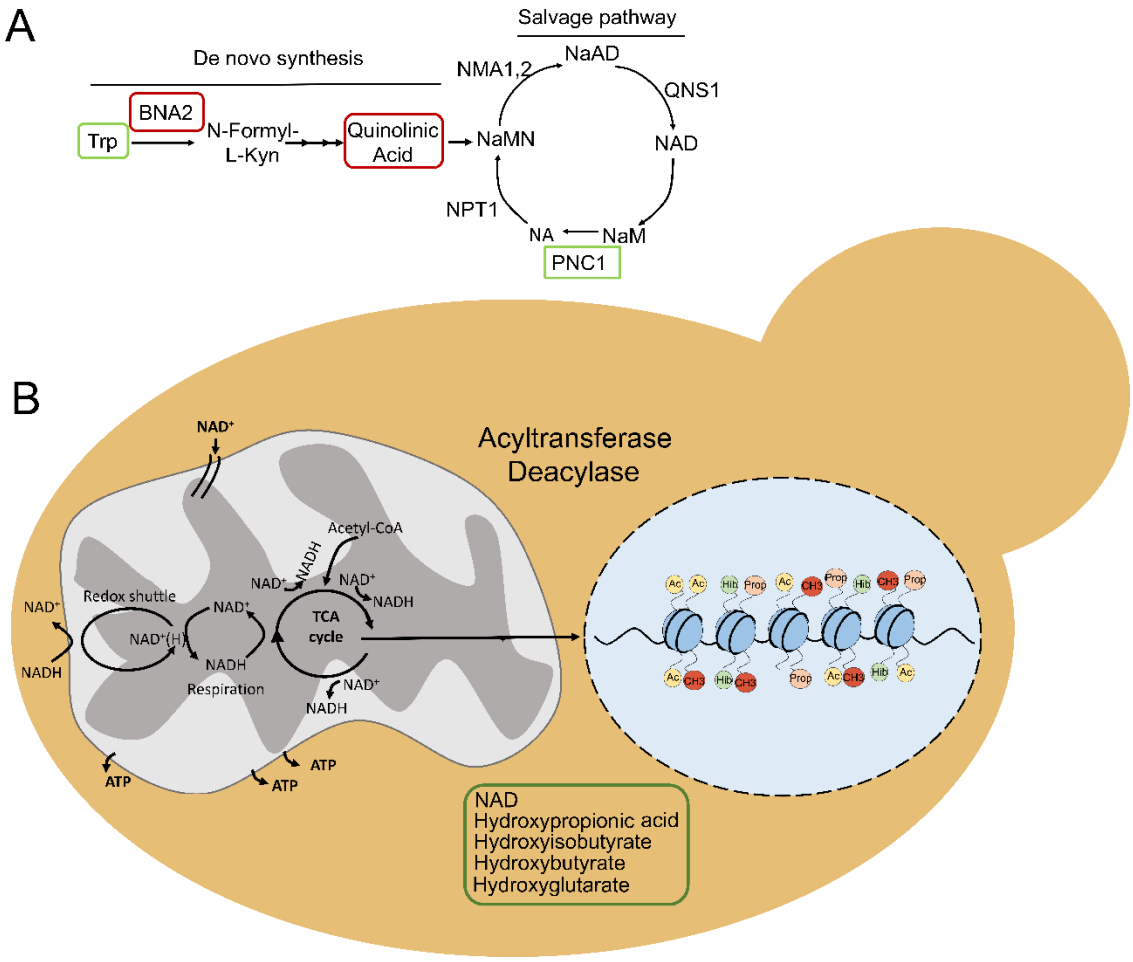


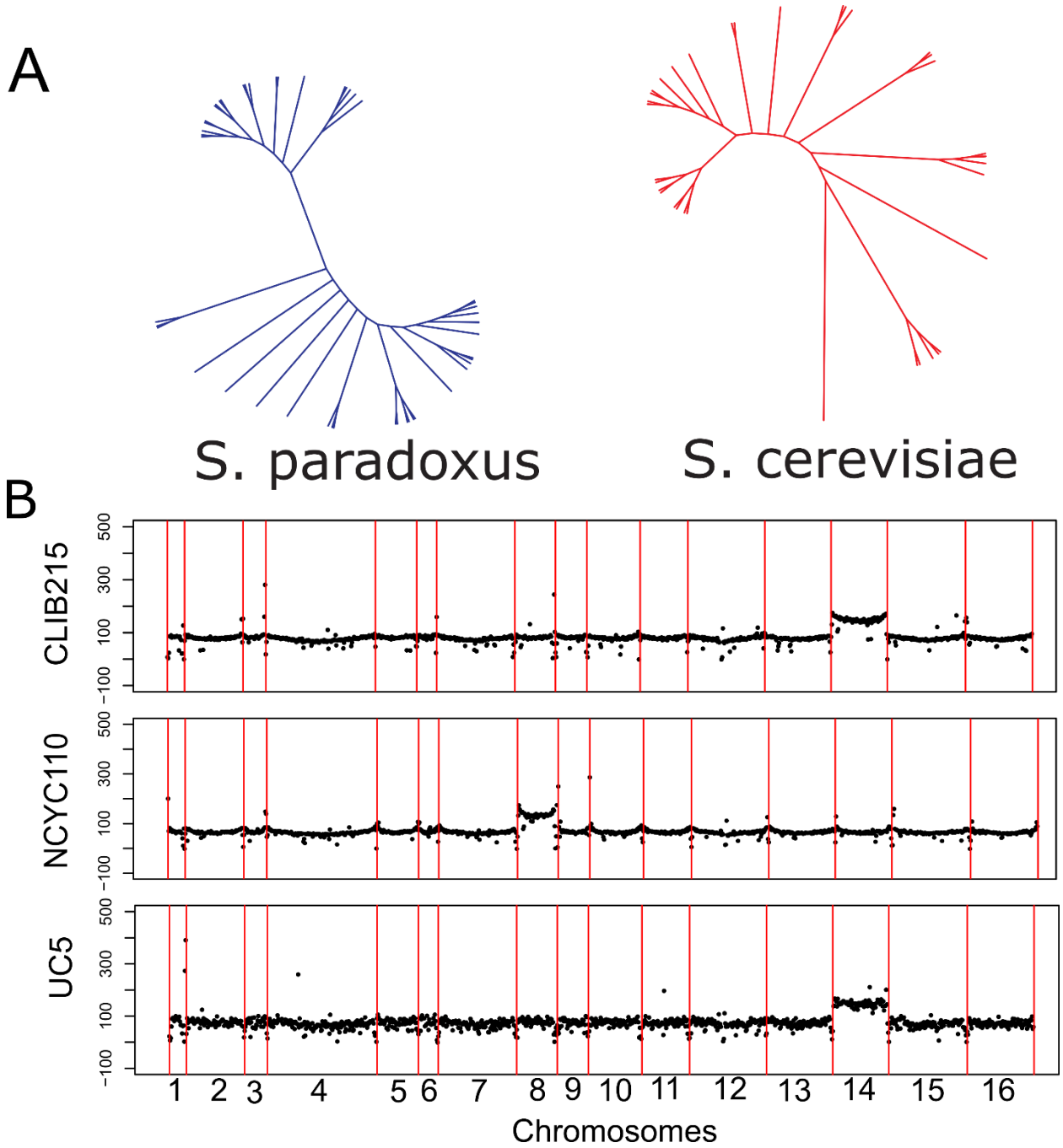
Figure 7: Adaptive genetic and metabolic changes in long-lived isolates. (A) Representation of Trp biosynthesis. Genes and metabolites associated with adaptation highlighted in red (decreased expression or decreased abundance) and green (increased expression or increased metabolite abundance). **(B)** Metabolic changes associated with altered respiration and TCA cycle. Adaptive changes in the abundance of TCA cycle metabolites are associated with histone lysine modifications.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary File:

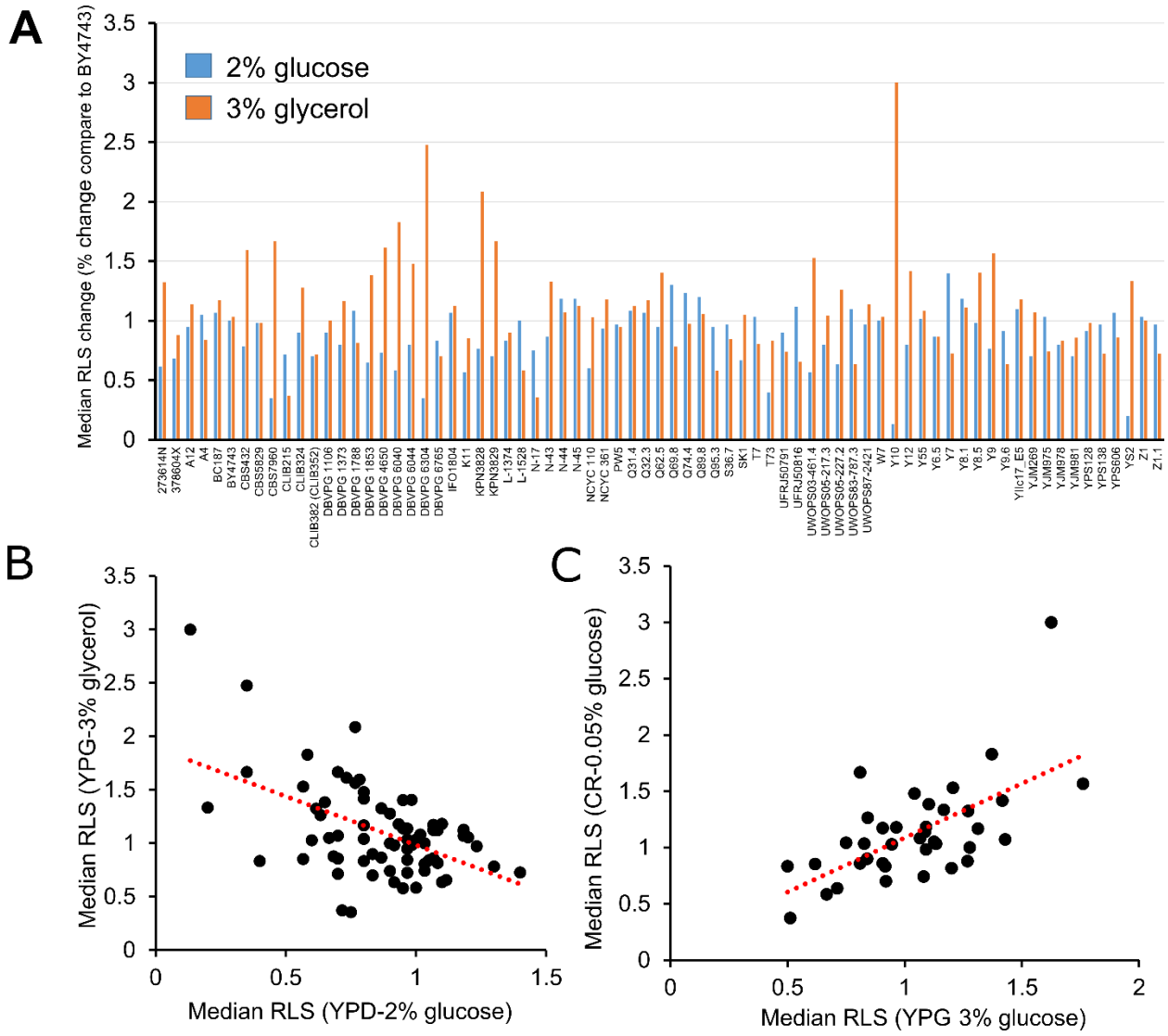
Evolution of Natural Lifespan Variation and Molecular Strategies of Extended Lifespan

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



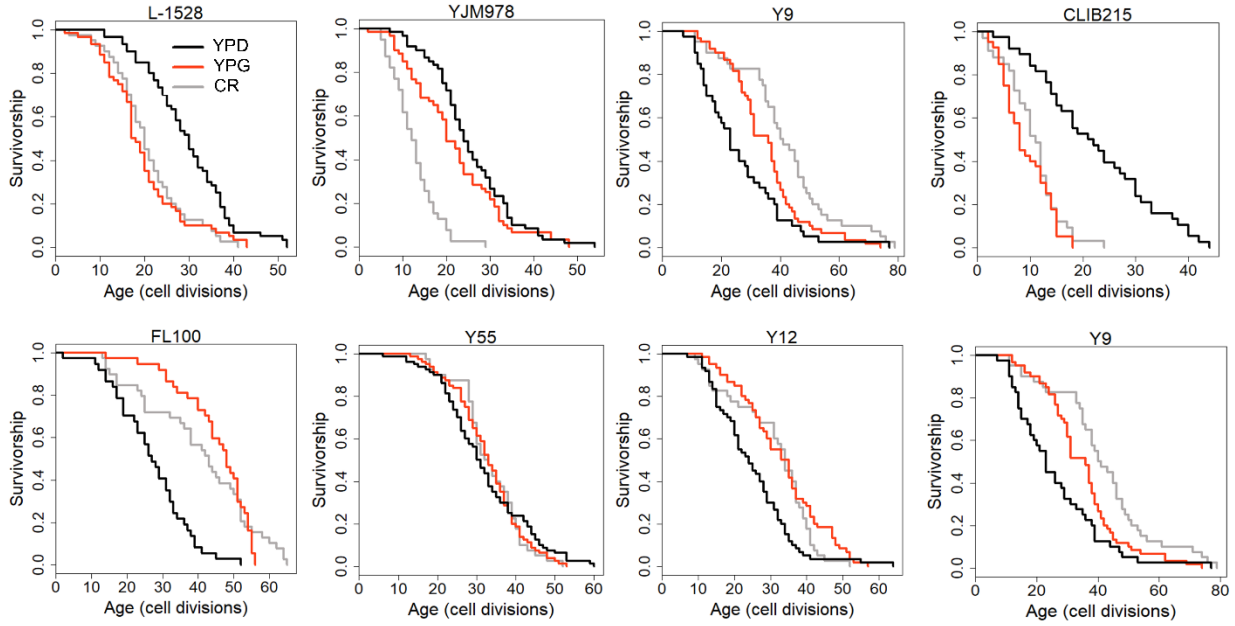
Supplementary Figure 1: Phylogeny and aneuploidy status of wild isolates. A) The neighbor-joining tree was constructed based on the alignment of complete genome sequences of the strains. **B)** Occurrence of aneuploidy in indicated wild isolates, identified by genome sequencing. Read depth was calculated in 100-bp windows. Genome Analysis Toolkit (GATK) to calculate the depth of coverage for each chromosome.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



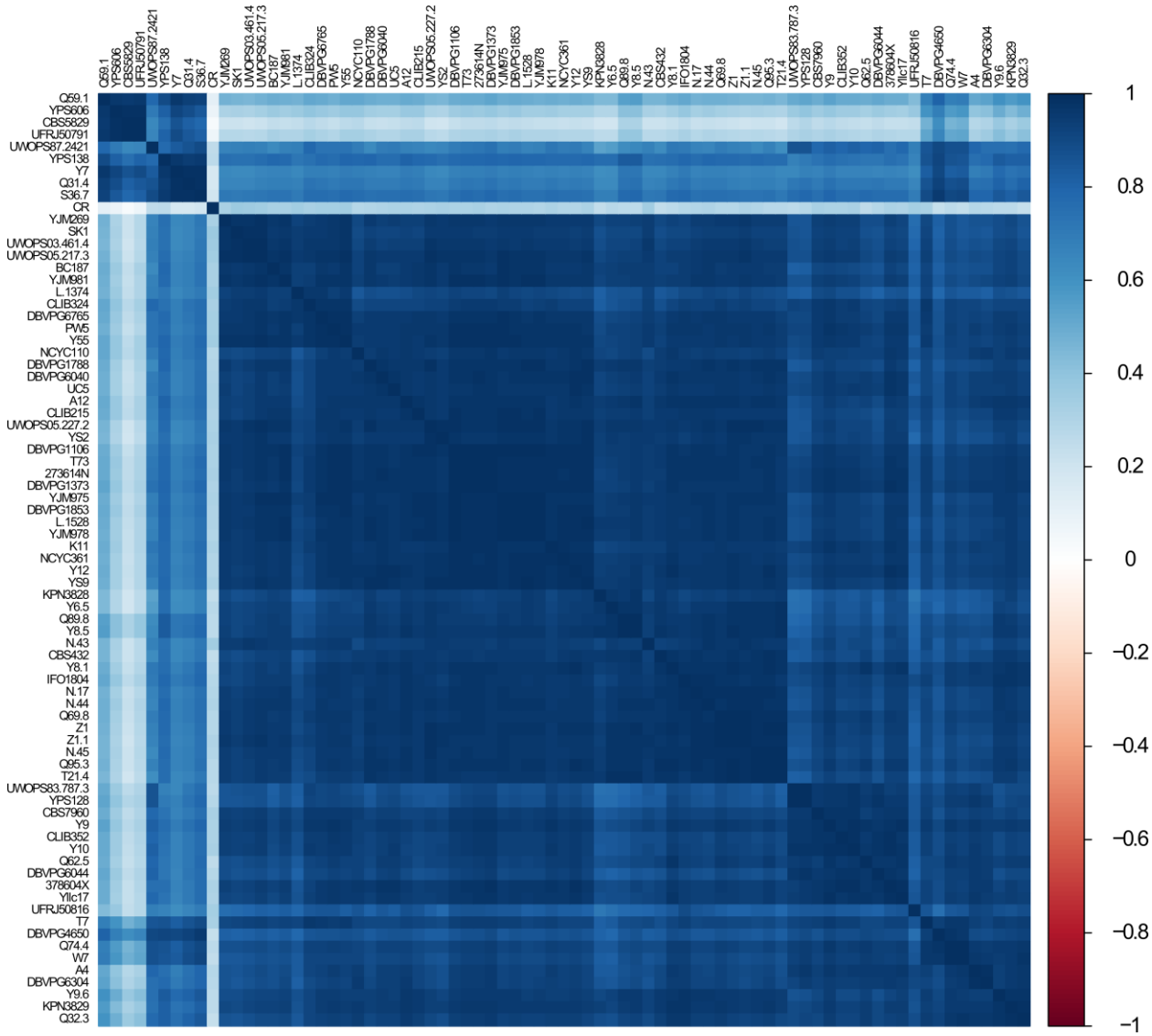
Supplementary Figure 2: RLS phenotype across wild isolates. A) Median RLS changes in YPD (glucose-blue) and YPG (glycerol-orange) across wild isolates normalized against laboratory diploid WT strain, BY4743 (median RLS X strain / median RLS BY4743). B) Significant negative correlation between median RLS YPD and median RLS YPG. Short-lived strains grown in YPD tend to have longer lifespan under YPG conditions, whereas lon-lived strains grown in YPD tend to have shorter median RLS when switched to YPG (Corr. coefficient = - 0.51, Padj value < 0.0001). C) Significant correlation was observed between median RLS in YPG and median RLS in CR, wherein glycerol-induced CR-induced respiration approximately equally extend RLS of those strains we tested (Corr. coefficient = 0.62, Padj value = 2.29×10^{-4}).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



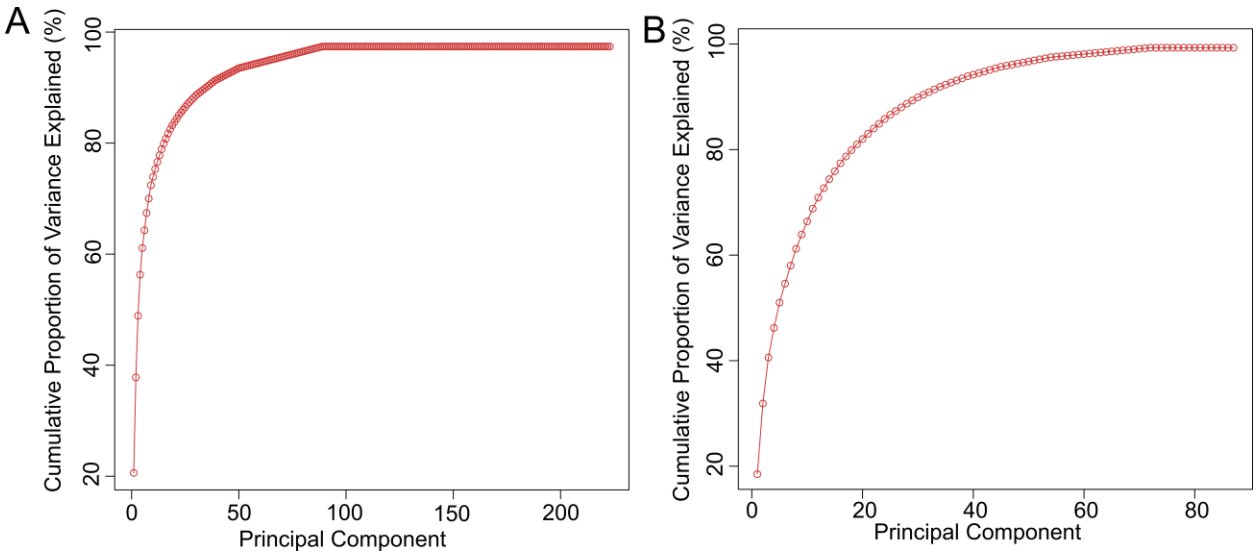
Supplementary Figure 3: RLS variation across wild isolates. Gompertz curves of RLS measures under YPD (black curves), YPG (red curves) and CR conditions are shown in each panel.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



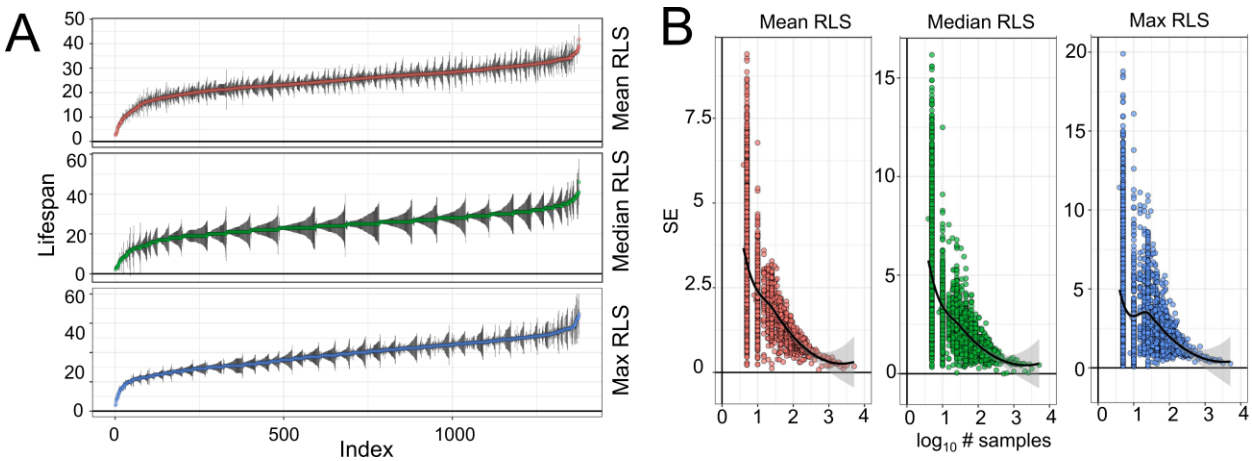
Supplementary Figure 4: Correlation of genome wide transcript levels across wild isolates. Heat map shows pairwise, 1-Pearson correlation matrix of gene expression data across wild isolates. Caloric restriction (CR) data was obtained from NCBI GEO, (SRA: SRX403444) and placed into the analysis.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



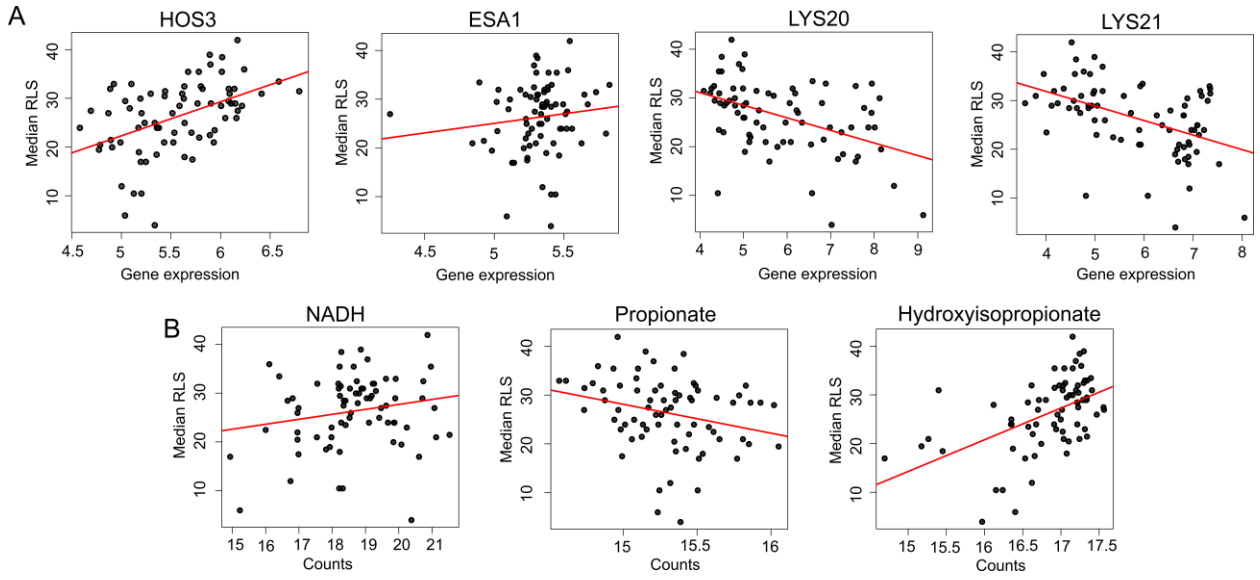
Supplementary Figure 5: Principal component analysis. Graphs show cumulative percentage of variance explained by Principal Components for (A) transcripts, (B) metabolites.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



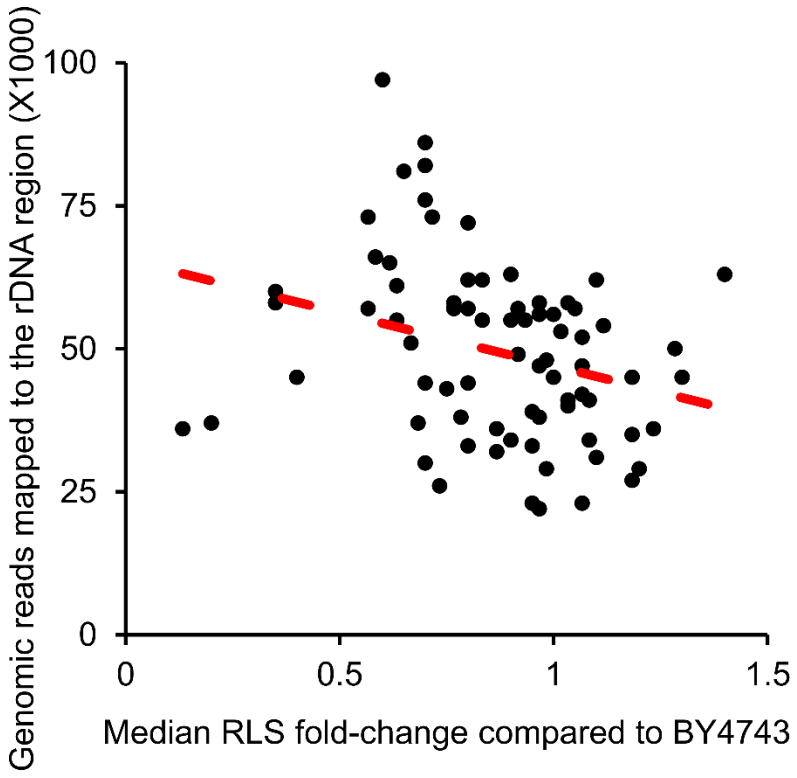
Supplementary Figure 6: RLS phenotype of yeast knock-out strains. A) Distribution of mean, median and maximum RLS across deletion strains with measured gene expression profile. RLS are sorted from the smallest to the greatest value. Black lines represent standard errors of the RLS estimates for corresponding strains. B) Dependence of standard error of RLS estimate on the number of strains used for evaluation of deletion mutant lifespan.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



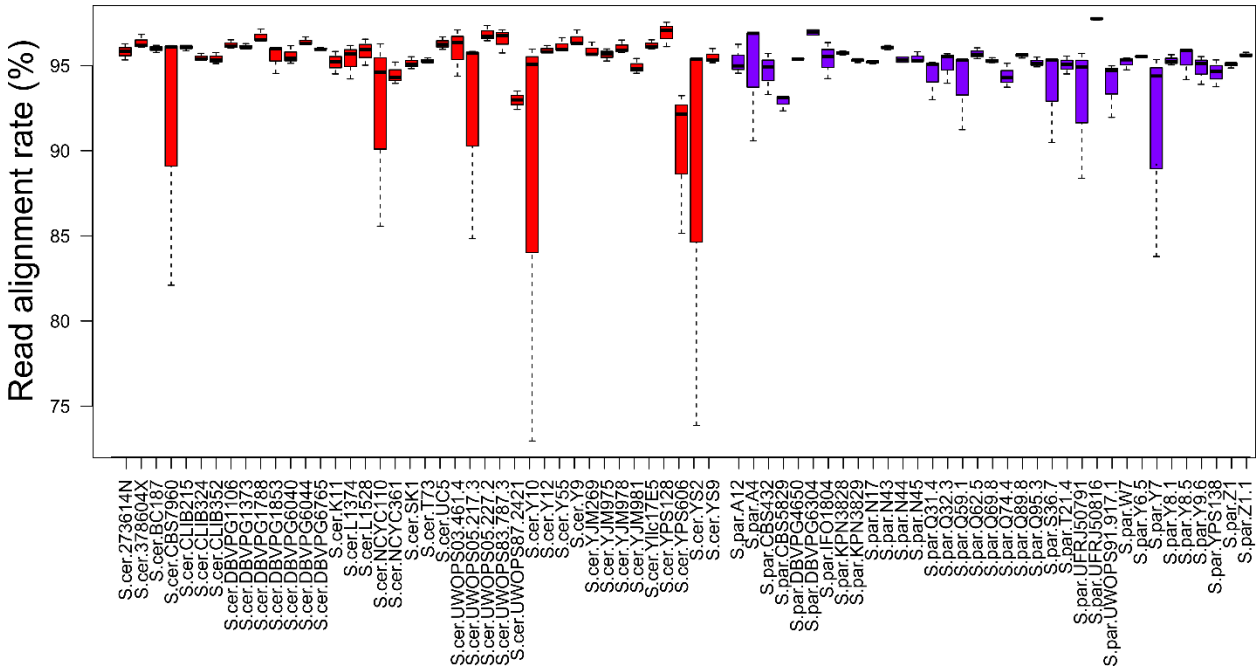
Supplementary Figure 6: Selected genes and metabolites correlating with median RLS. Transcript abundance of A) *HOS3 and ESA1* positively correlates with median RLS, and transcript abundance of *LYS20* and *LYS21* negatively correlates with median RLS. B) Correlation patterns of NADH, hydroxyisopropionate (positive correlation), and propionate (negative correlation) with median RLS. Regression slope *P* values can be found in Supplementary Table 3.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplementary Figure 8. Analysis of rDNA content in wild yeast isolates. Number of reads mapped to the rDNA region for each wild isolates were calculated using GATK and genome sequencing data. X axes shows median RLS fold change compared to the diploid laboratory WT BY4743 strain (Pearson correlation coefficient = -0.3, P = 0.01).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.09.374488; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplementary Figure 9. Read alignment rate for RNA-seq data. RNA-seq reads were mapped to the pseudo genome using STAR aligner software and mean value of alignment rate (3 replicates per strain) from STAR outputs are shown here. Error bars show standard error of mean. Red bars are indicates alignment rates for *S. cerevisiae* strains and purple bars indicate alignment rate for *S. paradoxus* strains.