Project Summary

RUI: Testing the network hypothesis of cellular aging in *Saccharomyces cerevisiae*

What is aging? This question remains elusive despite over 150 years of research. In this proposal, it is hypothesized that cellular aging is an emerged property of gene/protein networks at the cellular level. The plasticity of cellular aging, 70~80% of natural variation in life span, is mainly attributed to the stochasticity of detrimental changes that can propagate through the networks. Polymorphic variations can tinker the stochastic processes of aging by influencing the strength and dynamics of interactions in the networks and contribute to 20~30% of natural variation in life span. This hypothesis predicts that the rate of cellular aging ought to be proportional to the robustness of the network, which will be tested here both computationally and experimentally. The first aim is to examine the statistical associations of the effects of genes on aging and their effects in robustness, to look for shared characteristics of genes associated with aging and phenotypic capacitors, and to predict new genes associated with aging. The second aim is to experimentally test the link between network robustness and the rate of cellular aging. Several proxies of network robustness, such as tolerances to Hsp90 and TOR inhibitors and oxidative stress, will be measured in a collection of wild isolates of yeast. The associations between these proxies of robustness and the rate of cellular aging, defined by the Gompertz coefficient, will then be evaluated by regression analysis. Predicted new genes associated with aging will be experimentally verified.

The ***intellectual merit*** of this project is attributed to its novel hypothesis. By recognizing cellular aging as an emergent property, this network model of aging can provide a mechanistic foundation for other theories of aging, especially the antagonistic pleiotropy theory of aging. Because cellular aging is a system-level property, it cannot be pinpointed to specific genes, which explains a paradox in the field of aging: No gene with direct functional link to aging has been found, even though over 80 genes are known to influence the aging process in yeast. The general principle of emergent property of gene/protein networks will also offer new sights on the transition between genotype and phenotype spaces.

The ***broader impacts*** of this project can be found in the benefits to the research program and student learning at a historically black women's college. This project will largely be carried out by undergraduate students from a minority background. By engaging these students in hypothesis driven research and intense interdisciplinary training, we strive to cultivate their love of research and encourage their pursuits of science related careers. This project will also invigorate the research program and enhance the research capacity at this college.

RUI: Testing the network hypothesis of cellular aging in *Saccharomyces cerevisiae*

1. Results from Prior NSF Support

From February 2007 to July 2009, Dr. Hong Qin was supported by a HBCU-UP grant awarded to Tuskegee University (NSF Award #0411464, $2,498,843, P.I. Williams), titled "Enhanced Communication and Collaboration among STEM Disciplines through Undergraduate Curriculum Development and Research Opportunities". This support enabled Dr. Qin to study age-dependent genomic instability in wild isolates of the budding yeast.

Since January 2009, Dr. Qin was supported by a CCLI grant (NSF Award #0837075, $110,024, P.I. Qin), titled "Computing in Life Sciences through Hands-on Experience and Case Studies at Tuskegee University". Under this grant, Dr. Qin developed computing modules and materials based on his research in natural variation in life span and comparative genomics. These modules and materials have been offered at Tuskegee University and Spelman College (a historically black women's college).

Overall, three publications and two manuscripts resulted from these grants:

- Qin, H. and D.S. Goldfarb, Cellular aging is an emergent property of gene/protein regulatory networks. Manuscript in preparation.
- Qin, H. and A. Driks, Phylogenomic insights on the Bacillus spore coat. Mol. Biol. Evol., in revision.
- Qin, H., Teaching computational thinking through bioinformatics to biology students. Proceedings of 40th ACM Technical Symposium on Computer Science Education, 2009: p. 188-191.
- Qin, H., M. Lu, and D.S. Goldfarb, Genomic instability is associated with natural life span variation in *Saccharomyces cerevisiae*. PLoS ONE, 2008. 3(7): p. e2670.
- Qin, H. and L. Yang, Detection of changes in transitive associations by shortest-path analysis of protein interaction networks integrated with gene expression profiles. Proceedings of the IEEE International Conference on Biomedical Engineering and Informatics, 2008. 1: p. 418-423.

2. Goals and Objectives

What is aging? This important question in biology remains elusive despite 150 years of effort. We address this question from a unique approach based on our novel hypothesis that cellular aging is an emergent property of gene/protein networks (Qin & Goldfarb, in preparation). Our hypothesis has its root in the reliability model of aging developed by Gavrilov and Gavrilova [1]. Our hypothesis predicts that the rate of cellular aging ought to be proportional to the robustness of the gene/protein networks (see Fig 2 below). We define the rate of aging as the Gompertz coefficient (see section 4.1). We define robustness as the ability of cells to maintain homeostasis despite stochastic fluctuations, environmental changes, or polymorphic and mutation changes [2, 3]. We will test the prediction of our hypothesis both computationally and experimentally:

Aim 1 (Computational Tests): *We will examine whether the effects of genes on life span are statistically associated with their effects on network robustness, and whether genes associated with aging and phenotypic capacitors share similar characteristics. We will predict a list of novel candidate genes associated with aging for experimental tests based on the discovered patterns.* Here, we will use regression to analyze the ~500 genes with known effects on life span to explore what factors are informative on their roles in aging. These factors include network topological features, expression patterns, and competitive growth fitness. The phenotypic capacitors are known to influence network robustness, and hence, our hypothesis implies that they likely influence the rate of aging.

Aim 2 (Experimental Tests): *We will examine whether reduced network robustness will lead to slower rate of cellular aging. Specifically, we will test whether tolerances to Hsp90 and TOR inhibitors and oxidative stress in a collection of natural isolates of yeast are positively linked to their rates of aging, and whether deletion of the predicted candidate genes with possible effects on aging can influence the rate of aging.* Tolerances to Hsp90 and TOR inhibitors and oxidative stress indicate the capability of cells to maintain homeostasis in the face of external challenges, and are hence proxies of robustness.

These aims are based on the P.I.'s experience and expertise, our preliminary findings, our collection of yeast natural isolates, and the significance of expected results. This project will also be partitioned into modules and small projects and will be integrated into student learning and research.

3. Background and Significance

3.1 Overview of yeast aging

Aging has been studied for over 150 years, and tremendous strides have been made toward the mechanistic understanding of aging over the past two decades. Yet, even the very concept of aging is still under debate (For example, see [4]).

The aging of cells that undergo asymmetric divisions likely arose early in the evolution of both prokaryotes and eukaryotes [5-8]. As a unicellular organism, *Saccharomyces cerevisiae* (the budding yeast) has proven to be a good model system for studying aspects of cellular aging. Many key features of cellular aging were first discovered in yeast before they were established in metazoan cells [7, 9-11]. The life span of yeast can be measured in two ways: replicative and chronological life spans. Replicative life span (RLS) is the number of cell cycles that individual mother cells produce before they senesce and cease dividing [8, 12, 13]. The actual number of daughters produced by a cohort of mother cells is determined by microdissection. Chronological life span (CLS) is how long cells can survive without dividing in stationary phase [14, 15]. The number of surviving cells in a population is assessed over time by quantifying colony-forming units. Both replicative aging and chronological aging exhibit the familiar demographic characteristic of aging, that is, the nearly-exponential increase of mortality rate over time during the dying off phase of the population [16].

Studies of yeast aging have yielded many genes that are important in life span regulation, such as SIR2, TOR1, and SCH9 [17-19]. The roles of these genes on life span are often verified in other species. Paradoxically, none of these genes with known effects on life span suggests molecular mechanisms that are directly linked to aging. The effect of SIR2 on life span is attributed to the "toxic" effect due to accumulation of extrachromosomal rDNA circles [20], a concept that is itself mechanistically obscure and has recently been thrown into doubt [21]. The effect of TOR pathway on replicative life span is either attributed to the decreasing ribosome function and translation [9] or to the hyper-activation of cellular functions [4, 22]. The mechanism of TOR on chronological life span remains unclear [23]. It is speculated that bona fide aging genes do not exist because there are no conserved causes of aging [24-26]. It seems that the biology of aging becomes evasive once we delve into the molecular mechanisms. Why? Furthermore, why are there existing universal characteristics of aging despite the diverse genotypic and environmental factors that can influence the aging process? We argue that aging is an emergent property of biological complexity, and specifically, cellular aging is an emergent property of gene/protein regulatory networks. This novel hypothesis can provide a unifying framework for both the plasticity and the universality of aging, and can point to new research directions in aging and network biology.

3.2 Cellular aging as an emergent property of gene/protein networks – A novel hypothesis.

Complex systems often behave as a collective due to the interaction of individual components. For example, the response of individual fishes to their neighboring fishes can lead to the formation of a shoal of fishes. The activity of a protein can be attributed to its amino acid composition and their spatial interactions. In both examples, the collective activities exceed the sums of individual components, which are examples of emergent properties of biological complexities [27]. Although formal definition of biological complexity is challenging, it is generally accepted that gene/protein networks are examples of biological complexity. Conceptually, gene/protein networks are proxies of biological complexity at the cellular level, especially for single-cell organisms such as *S. cerevisiae*.

A large body of experimental data in yeast suggests complex mechanisms of aging. In fact, the very reasons that cause aging are still under debate, including the popular ROS theory [4, 28]. A common criticism on the known genes associated with aging is that they are not "real" ones because they have no direct role in aging. All of the known genes with effects on life span, such as SIR2, TOR1, and SCH9,

have no direct known functional link to aging. Similarly, in *Caenorhabditis elegans*, the life span extension mechanisms of Indy, daf2 are explained through the elusive mechanism of calorie restriction. Furthermore, many pathways are known to influence life span and are probably intertwined together. Although SIR2 and TOR are argued to be parallel pathways, there are evidences that they are inter-connected through MSN2/4 [29]. Both SIR2 and TOR pathways are thought to promote genomic stability during aging [29]. In a large scale screen, deletion of 90 genes were found to extend chronological life span in BY laboratory strains, and only 16 of them are TOR related [30]. The rest of them belong to iron homeostasis, cell wall organization and biogenesis, transport, and many have unknown functions [30]. Deletion of 300 genes can shorten chronological life span [30]. In another screen of replicative life span, 20% of the gene deletions were found to shorten replicative life span, whereas 10 out 564 genes significantly extend replicative life span [18]. Six of them are implicated in the TOR pathway, four others are a ubiquitin protease, an isocitrate dehyodrogenase, and two proteins with unknown functions. These seemingly complicated and often inconsistent observations argue for new models of aging.

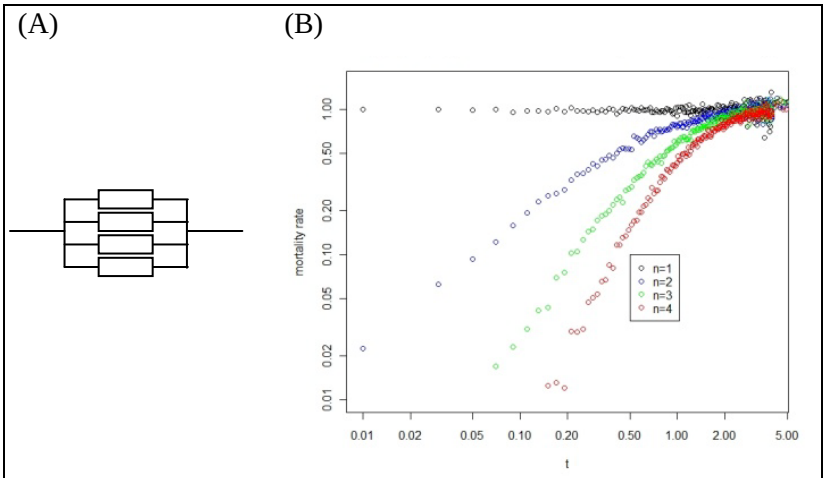

Fig 1. Increase of mortality rate over time, a characteristic of aging, can arise in systems with parallel non-aging components. (A) A system with 4 parallel components (n=4). (B) Simulation of parallel systems with 1,2,3, and 4 non-aging components (constant aging rates).

We propose that cellular aging is an emergent property of gene/protein regulatory networks, which will provide a conceptual framework to explain the seemingly inconsistent experimental data and the universal characteristics of aging. As an emergent property of networks, cellular aging is a system-level property. The hierarchical nature of the emergent property explains the difficulty of pinpointing "real" genes involved directly in aging. The universal characteristics of aging can be attributed to the common patterns of gene/protein networks shared among most species.

Viewing aging as a network emergent property is a logical extension of the reliability model of aging. Gavrilov and Gavrilova (2001) demonstrated that characteristics of aging can arise from a system with parallelly connected non-aging components (Fig 1) [1]. In other words, aging is an emergent property of a system with redundancy [1]. Furthermore, Gavrilov and Gavrilova found that a serial-parallel system model with initial heterogeneity can reproduce the Gompertz feature of aging. This model system consists of K serially linked modules with parallel components. Initially, there are a random number of functional components in each module based on a Poisson distribution (mean=λ) or binomial distribution (mean=np). The mortality rate of the this system is

$$m = \mu\lambda KCe^{-\lambda}\sum_{i=1}^{n}\frac{(\lambda\mu t)^{i-1}}{(i-1)!} \approx I\,e^{Gt}\text{ , when t} \ll 1/\mu.$$

where $I = CK\lambda\mu e^{-\lambda}$, and $G = \lambda\mu$, μ is the constant mortality rate of each component, C is a normalization constant, and n is the number of components in each module [1]. This model indicates that the rate of aging, G, is proportional to the average number of initial functional units, λ, in the system. Hence, the rate of aging is proportional to the redundancy of the system. A biological example of this serial-parallel model with the initial heterogeneity is the low number of transcription factors and redundant binding motifs in promoter regions. Notice that this serial-parallel model with initial heterogeneity cannot produce the known $\ln I \sim G$ negative correlation in realistic ranges of I and G, an indication of its shortcoming due to over simplicity. At the core of the reliability model is the functional compensation among network components. In the much more complex gene/protein networks , both network buffering and overlapping

roles of genes can lead to functional compensation [31, 32], which is generally termed "robustness" [2, 3]. Formally, robustness can be defined as insensitivity to noises due to stochastic fluctuations, environmental changes, or polymorphic and mutation changes [2, 3]. Given that cellular aging is hypothesized as an emergent property of gene/protein networks, it is therefore expected that rate of cellular aging ought to be proportional to the robustness of the network.

How do the characteristics of aging arise from a network-based model? Because only ~20% of the genes are essential in the yeast genomes, failure of nodes alone cannot explain the full complexity of yeast aging. Notice that failures of redundant nodes in the reliability model implicitly lead to heterogeneous system states with different failure risks. These implicit heterogeneous system states are theoretically related to the frailty model [33], which explicitly assumes heterogeneous system states. We therefore think that heterogeneity in system states is the key for the emergence of aging from systems. Conceptually, cells can be viewed as dynamic complex systems, which possess many attractor states [34]. Homeostasis of different cell states, such as the G0 phase and other mitotic phases, can be viewed as attractor states. Death can be viewed as an absorbing state in the sense that cell can transition from other states into, but not from, the state of death. Environmental factors and genes can expedite or delay the transition from the homeostasis states to the death state. In other words, "aging" is a "structural property" of cellular systems as dynamic complex systems. Over time, cells will move, in the state space of cells, into the "attractor of death".

Alternatively, aging at the network level may lead to poor attracting basins of homeostasis states. Increasing noises due to detrimental changes of protein molecules (network components) and decreasing network buffering will lead to failure to pass through cell cycle (equivalent to death in replicative aging) or failure to re-initiate cell cycle (equivalent to death in chronological aging). Notice that this second definition of aging is essentially equivalent to the decrease of network robustness. Aging in this sense is essentially the "default and flattened" system state space.

The role of networks in aging has also been discussed by others, explicitly or implicitly. Franschi argued that aging is associated with cellular networks [35]. Kowald and Kirkwood proposed that the effect of defective mitochondria, aberrant proteins, and free radicals should be integrated as a "network theory of aging" [36, 37]. Soti and Csermely proposed a "weak link theory of aging"- age-related damage affects the low affinity, transient interactions (weak links) and lead to increased noise, destabilization and diversity [38]. Eventually, these noises cannot be distinguished from normal signals. Hence, aging is "deterioration of emergent network properties", which is similar to the second scenario of yeast aging that we are discussing [38]. Rattan suggested that many genes influencing life span were not "selected" as "genes for aging" and their effect on aging is due to a number of "functionally coupled genes" [39]. Rattan argued that most hypothesized aging models imply "directly or indirectly that progressive failure of homeostatic mechanism is crucial for the process of aging" [39].

Our network model recognizes that cellular aging is itself an emergent property of networks. Hence, aging is a system-level property that can be influenced by propagation of changes through networks. These changes may be instigated by ROS damages or other factors, such as stochastic noises. Due to the low copy numbers of many genes, stochastic noise plays a significant role in the plasticity of aging, which is consistent with the large portion of non-genic variation in life span [16]. Genotypic variation only contribute 20-35% of the life span variation, as shown in fruit flies, budding yeast, and humans [16, 40-42].

Viewing cellular aging as an emergent property of gene/protein networks may provide mechanistic foundation for some evolutionary theories on aging, such as the antagonistic pleiotropy theory [43] and the disposable soma theory [44]. Not only can we better understand the concept of pleiotropy from the network perspective, we will probably also look to networks to understand the connection between advantages at early life and detrimental effects at late life [43]. This kind of trade-off is also the central argument of the disposable soma theory [44].

As an emergent property of complex systems, aging is hence a ubiquitous property in all organisms. Viewing aging as a system-level emergent property has some immediate implications theoretically and practically. First, it naturally leads to plasticity of life span. The plasticity of life span

likely provides the raw materials of variation that may be acted upon by natural selection. Another important implication is to reconcile the seemingly contradictory experimental results on cellular aging in various laboratory strains of yeast. Due to polymorphic variation, the roles of two pathways on aging can vary to an extent that the effect of one pathway may be easily measurable in one strain but not in another strain. For example, the role of SIR2 on calorie restriction is discernable in W303 background but not in the BY background [17, 45, 46]. Most yeast strains show extended replicative life span in 0.5% glucose, but not the BY laboratory strains [45], suggesting that variance in nutrient sensing pathway and presumably the relative extents that TOR1 and SIR2 play in calorie restriction response in different yeast strains.

3.3 Comparison with other studies on aging from network perspectives

The importance of network in aging is also recognized by many others. Xue and colleagues discovered in *C. elegans* that "aging genes" are more enriched at the interfaces between network modules and argues that "aging is linked to the dynamic network stability" [47]. Bell and colleagues studied the human protein interaction network and found genes involved in life span regulation tend to be inter-connected hubs. They experimentally verified 18 genes in worm and found a third of them can extend worm's life span [48]. Budovsky and colleagues explored the human orthologs of genes associated with aging in model organisms. They found that hub genes are often involved in age-related diseases [49]. They further found that tumor suppressors tend to extend life span, whereas oncogenes tend to shorten life span, which are argued as evidence for the evolutionary and molecular links between aging and cancer. Smith and colleagues compared genes associated with aging in yeast and worm and found that yeast orthologs of worm genes associated with aging also tend to alter life span in yeast, indicating a conserved mechanism of life span regulation between yeast and worm [50, 51]. Lorenz and colleagues focused on a 10-gene network from the Snf1 signaling pathway in yeast and engineered perturbation to infer causal network interactions and found new interaction associated with aging [52]. Barea and Bonatto argued that protein interaction network can synthesize different theories of aging and used yeast protein network to explore the connection between replicative and chronological aging [53]. Promislow compared yeast protein associated with aging to five traits and found that protein associated with aging are more connected than expected by chance, which is argued as evidence of antagonistic pleiotropy theory for the evolution of senescence [54]. Promislow also argued that the network characteristics of proteins associated with aging may be useful to predict new genes associated with aging [54].

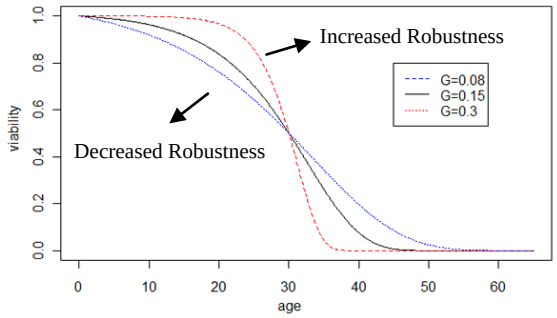

Fig 2. Prediction of our hypothesis, the effect of network robustness on the rate of aging *G*, is shown by changes of the Gompertz survival curves. Increased robustness leads to larger *G* and faster aging, shown by the sharper transition of the dying off phase, whereas decreased robustness leads to smaller *G* and slower aging, shown by the slower dying off phase. For illustration purposes, the average life spans are unchanged.

These previous studies provide many helpful leads to us. By recognizing that cellular aging is an emergent property of gene/protein networks, we are able to move beyond the current thinking of aging. In this project, we will test the most important prediction of our hypothesis: The rate of cellular aging ought to be proportional to the robustness of the network (Fig 2), both computationally and experimentally.

4. Previous and Preliminary results

4.1 Quantifying the aging process based on the Gompertz model

We quantified yeast survival curves using the two-parameter Gompertz aging model:

$$m = -\frac{1}{s}\frac{ds}{dt} = Ie^{Gt} \qquad \text{(Eq 1)}$$

$$s = e^{\left(I/G\right)\left(1-e^{Gt}\right)} \qquad \text{(Eq 2)}$$

where, *m* is mortality rate; *s* is the survival fraction of a population; *t* is time. Here, mortality rate is defined as the normalized declining rate of *s*. The initial mortality rate, *I*, describes the innate susceptibility to dying. The Gompertz coefficient, *G*, determines acceleration rate of the mortality rate over time and is viewed as the rate of aging. Examples of Gompertz survival curves can be seen in Fig 2.

Given a Gompertz model, the average life span is defined as:

$$Average\ Life\ Span = \frac{1}{G}\ln\left(1+\ln 2 \bullet \frac{G}{I}\right) \qquad \text{(Eq 3)}$$

.

This equation shows that for a given mean life span, there are many pairs of *I* and *G* values, as shown by the contour lines in Figure 3.

For a given set of measured life spans of N cells ($\vec{t}$), the likelihood function based on the Gompertz model is:

$$L\left(I,G\,|\,\vec{t}\right) = \prod_{i=1}^{N} s_i m_i = \prod_{i=1}^{N} Ie^{-Gt}\, e^{\left(I/G\right)\left(1-e^{Gt}\right)}$$

.

Hence, the log-likelihood is:

$$\ln L = \sum_{i=1}^{N}(\ln s_i + \ln m_i) = \sum_{i=1}^{N}\left[\ln I + Gt + \frac{I}{G}(1-e^{Gt})\right] .$$

Maximization of the above log-likelihood function will yield the maximal likelihood estimations of the Gompertz parameters for a data set of life span. We have implemented the maximization procedure in the R-statistics language and environment. This likelihood approach enables us to design nested models to statistically compare *I* and *G* values of different strains (see section 4.3).

The Gompertz model can be extended by the introduction of a morality component that is independent of time, represented by a constant as an extra term inserted to Equation 1. This extension will then use three parameters and is often called Gompertz-Makeham model [55, 56]. Environmental factors, such as nutrients, radiations, and toxic influences, are known to influence the parameters in both the Gompertz model [57] and the Gompertz-Makeham model [58]. Other alternative models include the Weibull model based on a power function [56] and the logistical model [59].

Departure from the Gompertz model is often observed at late age when the acceleration of mortality rate slows down [59, 60]. In other words, a small percentage of long-lived individuals are often important to determine whether the Gompertz model or other aging models is the best-fit model. As a result, large sample sizes are often required to examine the differences of various aging models. The Gompertz model is often indistinguishable from other models when sample sizes are less than 100

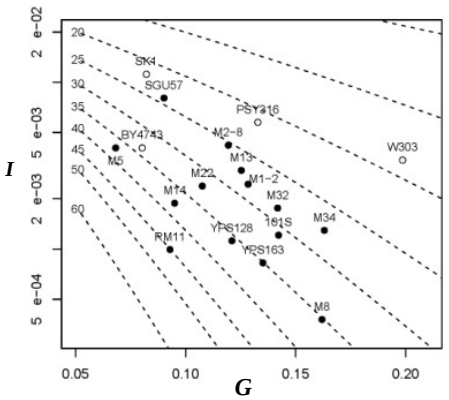

Fig 3. Replicative aging described in the Gompertz parameter space. The ARLS values are represented by the dashed contour lines with the values labeled at the ends. Solid black circles are natural isolates. White circles are laboratory strains.

[59, 60]. The sample sizes of most yeast RLS assays range from 30 to 60. Therefore, the Gompertz model is a reasonable choice to study the yeast replicative aging.

4.2 Quantifying the aging processes in a collection of yeast strains

We have systematically investigated the life span in a collection of yeast lab strains and wild isolates (Table 1) [16]. We observed a considerable amount of RLS variation among the 14 natural isolates. The average RLS (ARLS) values of the isolates follow a normal distribution. The mean ARLS for the 14 natural isolates is 31.7 cell divisions, and the standard deviation is 5.8 cell divisions. RM11 has the largest ARLS value at 44.5 cell divisions, and SGU57 has the smallest value at 23.6 cell divisions. The two extreme values differ by 1.9-fold.

The pattern of natural RLS variation described by the Gompertz model is presented by a scatter plot of $I \sim G$ (Fig. 3). The G values, which are related to the rate of aging, follow a normal distribution. The mean and standard deviation of the G value distribution are 0.12 and 0.025, respectively. The I values, the initial mortality rates, follow a lognormal distribution. The mean and standard deviation of the natural-logarithm transformations of I are -6.3 and 0.72, respectively. Using multiple regression analysis, negative correlations are observed between ARLS ~ I , ARLS ~ G, and I ~ G, with partial correlation coefficients at -0.90 (p-value = 2.7×10^{-5}), -0.86 (p-value = 1.8×10^{-4}), and -0.87 (p-value = 1.1×10^{-3}), respectively. Intuitively, large I and G values ought to shorten the average life span. The negative partial correlation between I (or ln I) and G is often interpreted as tradeoff from the evolutionary perspective. This correlation can be further verified by the straightforward regression between ln I and G. Linear regression shows that correlation between ln $I \sim G$ has a p-value of 0.065 and an R^2 value of 0.26. This negative correlation is often called Strehler-Mildvan correlation and is a characteristic of aging in higher organisms [61]. This observation demonstrates that yeast replicative aging shares similar characteristics of aging in higher organisms at the population level.

We used linear regression models to evaluate genotypic influence on RLS. In these linear models, we investigated whether the RLS values of individual cells could be predicted by variables corresponding to their genotypes, experimental groups, and plates. We found that genotypic variation is a significant factor on RLS and contributes to ~ 22% of the total RLS variation observed at the cellular level. This genotypic influence on life span is comparable to the estimation made in *Drosophila melanogaster* and in human population based on identical twins [40-42].

Table 1. Replicative and chronological life spans in yeast strains

Strain	ARLS	*I*	*G*	CLS
101S	31.3±0.8	0.0012±0.0007	0.14±0.02	3.4±0.2
M1-2	27.9±1.1	0.0024±0.0011	0.13±0.02	10.4±3.2
M13	26.7±1.2	0.0030±0.0011	0.13±0.01	16.3±3.7
M14	36.6±1.5	0.0019±0.0007	0.09±0.01	7.2 ±0.2
M22	31.8±1.3	0.0024±0.0009	0.11±0.01	5.2±2.1
M2-8	24.8±0.8	0.0042±0.0010	0.12±0.01	4.1±0.7
M32	28.1±0.8	0.0018±0.0005	0.14±0.01	6.4±0.8
M34	26.8±1.0	0.0013±0.0007	0.16±0.02	5.2±0.4
M5	36.7±1.0	0.0040±0.0008	0.07±0.01	4.9±0.5
M8	35.2±0.9	0.0004±0.0002	0.16±0.02	10.5±0.1
SGU57	23.6±1.5	0.0080±0.0022	0.09±0.01	9.3±0.7
RM11	44.6±1.5	0.0010±0.0004	0.09±0.01	12.7±2.9
YPS128	35.0±1.2	0.0011±0.0005	0.12±0.01	3.3 ±0.1
YPS163	34.4±0.8	0.0008±0.0003	0.14±0.01	4.2±1.1
BY4743	33.2±0.9	0.0040±0.0013	0.08±0.01	9.7±1.7
SK1	22.0±1.3	0.011 ±0.003	0.08±0.01	5.0±0.9
W303	18.7±0.6	0.0034±0.0011	0.20±0.02	17.2±3.9

ARLS (Average replicative life span), CLS (Chronological life span). Standard deviations are estimated by bootstrapping. The numbers of bootstraps equal the sample sizes. These strains are a unique and important resource of this project.

4.3 Diploid cells tend to age faster than haploid cells based on likelihood ratio tests of nested models.

A particular challenge in studying variation in life span is its stochasticity, which demands rigorous quantitative modeling and analysis. Our likelihood approach enables us to statistically test *I* and *G* values of different yeast strains, which is often sidestepped by other investigators in yeast aging. We developed likelihood ratio tests based on nested models, implemented in the R-language and environment.

Diploid cells are much larger in size than haploid cells, indicating that there more protein molecules (network components) in gene/protein networks in diploid cells. Hence, networks are more robust in diploid cells than those in haploid ones. Our hypothesis will then predict that diploid cells should age fasters than haploid cells. We found that haploid wildtype strains (BY4741 and 4742) share the same rate of aging (Gompertz coefficient) (Table 2A). We then found that diploid wildtype strain (BY4743) has significantly higher rate of aging than do the two haploid wildtype strains (Table 2B). Therefore, these findings are in agreement with our hypothesis (see Fig 2). Interestingly, the diploid strain lives longer despite higher rate of aging, because of its lower initial mortality rate. This kind of tradeoff is known as the Strehler-Mlidvan correlation [61].

Table 2A
Rates of aging are the same in haploid BY4741 and BY4742 based on nest model tests.

Models	lnL	
H0 $I_{BY4741}= I_{BY4742}$, $G_{BY4741}= G_{BY4742}$	-587.3	
$H1_i$ $I_{BY4741}\neq I_{BY4742}$, $G_{BY4741}= G_{BY4742}$	-579.4	Best parsimonious model p-value=$7.2x10^{-5}$(chi-square, df=1) $G_{BY4741}= G_{BY4742}= 0.0653$
$H1_g$ $I_{BY4741}= I_{BY4742}$, $G_{BY4741}\neq G_{BY4742}$	-581.4	
H2 $I_{BY4741}\neq I_{BY4742}$, $G_{BY4741}\neq G_{BY4742}$	-579.4	

Table 2B
Aging rate is faster in diploid BY4743 than in haploid strains based on nest model tests.

Models	lnL	
H0 $I_{BY4741} = I_{BY4742} = I_{BY4743}$, $G_{BY4741}= G_{BY4742}= G_{BY4743}$	-952.7	
$H1_{g2}$ $I_{BY4741} = I_{BY4742} = I_{BY4743}$, $G_{BY4741}= G_{BY4742}\neq G_{BY4743}$	-946.1	
$H1_{i2}$ $I_{BY4741} = I_{BY4742} \neq I_{BY4743}$, $G_{BY4741}= G_{BY4742} = G_{BY4743}$	-939.4	
$H2_{ig2}$ $I_{BY4741} = I_{BY4742} \neq I_{BY4743}$, $G_{BY4741}= G_{BY4742}\neq G_{BY4743}$	-937.2	
$H3_{i3ig2}$ $I_{BY4741}\neq I_{BY4742}\neq I_{BY4743}$, $G_{BY4741}= G_{BY4742}\neq G_{BY4743}$	-929.3	Best parsimonious model p-value=$3.6x10^{-10}$ (chi-square, df=3) $G_{BY4741}= G_{BY4742} = 0.0653$ $G_{BY4743} = 0.0786$
H6 $I_{BY4741}\neq I_{BY4742}\neq I_{BY4743}$, $G_{BY4741}\neq G_{BY4742}\neq G_{BY4743}$	-929.3	

5. Research Plans, Expected Results, and Alternatives Approaches

Aim 1 (Computational Tests): We will examine whether the effects of genes on life span are statistically associated with their effects on network robustness, and whether genes associated with aging and phenotypic capacitors share similar characteristics, such as network topological characteristics, gene expression patterns, and competitive growth fitness. We will predict a list of novel candidate genes associated with aging for experimental tests based on the discovered patterns.

5.1 Testing the association between robustness and effects on aging.

The first computational goal is to test the association between gene's role in network robustness and gene's effect on cellular aging. Several proxies of network robustness will be used: variance of morphological changes, variance of expression levels, and growth rates of deletion mutants. Genes associated with aging will be based on two large-scale yeast aging assays using the yeast deletion collection, one is the replicative life span measure and the other is the chronological life span measure [30, 51, 62]. Statistical associations will be explored among these data.

The concept of robustness can be interpreted in many ways. In general, robustness is defined as insensitivity to noises due to either stochastic fluctuations, environmental changes, or polymorphic and mutation changes [2, 3]. To achieve robustness, cells can either use network buffering mechanism or use simple gene duplication [32, 63]. One often used proxy of a gene's role in robustness is the phenotypic variances in the mutants of this gene, such as morphological phenotypes and expression levels [2, 3, 54, 64]. Paradoxically, this measure can be interpreted in opposite ways for singleton and duplicated genes, which has been neglected in previous works. For a singleton gene, phenotypic variance for its deletion mutant indicates how much robustness can be attributed to the wildtype allele. Hence, large phenotypic variances in deletion mutants suggest large effects on robustness for singleton genes. For duplicated genes, deletion of one copy of the duplicated genes often leads to phenotypic variation to a much less extent than singleton genes because duplicated genes tend to functionally compensate for each other [32]. Hence, small phenotypic variances in deletion mutants suggest large role in robustness for duplicate genes. A recent study by Cooper and colleagues confirmed that single deletions of duplicated genes tend to have similar fitness. However, they concluded that duplication effect is only ~1.5% and does not substantially contribute to gene's robustness (pleiotropy in the authors' words) [64]. These previous reports suggest that we need to control for singleton and duplicate genes in our analysis.

The *Saccharomyces cerevisiae* Morphological Database (SCMD, http://scmd.gi.k.u-tokyo.ac.jp/) provides a list of 501 morphological parameters for every yeast gene in four groups - cell shapes, bud sizes, nucleus locations, and actin localizations [65]. Recall that the replicative life span is essentially the maximal potential of mitotic division and the chronological life span is the maximal potential length of pause of the cell division, measures about cell divisions are likely informative for aging. Many of the morphological parameters are conceived with respective to the budding process of *S. cerevisiae*, and hence, are also likely to offer insights on the replicative aging process. Principal components will be calculated to remove inter-correlation of these parameters. The variance of the principal components for a given deletion mutant can be used as a proxy for how much robustness can be attributed to this gene because the wildtype allele of this gene has "buffered" these morphological phenotypes [64].

Another proxy of robustness is the variance of gene expression levels in a deletion mutant of a gene because this variance is due to the "buffering effect" of the deleted gene [66, 67]. Preliminary search of the NCBI GEO database showed that many datasets of gene expression are potentially suitable for our project. We are particularly interested in series of cell cycles (GDS2347), heat shock stress (GDS2343, 1711, 281, 36, and 35), oxidative stress (GDS108), rapamycin effect (GDS2338), various growth inhibitions (GDS2196 and 1636), polymorphic variation (GDS1115 and 1116), and nutrient limitations (GDS777, 115, 112, and 111). We have experiences with large scale gene expression analysis and network inference [68, 69]. We have established a batch query procedure to GEO in the R environment through a client-server exchange protocol implemented in the GEOquery package [68, 70].

The third proxy of robustness is the competitive growth fitness [71]. If deletion of a gene substantially reduces robustness, competitive growth of the cells should be significantly lowered. Hence, competitive growth fitness of a deletion mutant should be inversely correlated with gene's role in network robustness. The relative growth fitness in the presence of rapamycin is of great interest because rapamycin targets the TOR pathway that is known to regulate life span [72].

A large scale RLS data for 564 yeast gene deletion mutants are available from the Kaeberlein group [62]. Among them, 42 gene deletions can extend RLS with p-value < 0.05 (26 if p-value<0.01), and 57 gene deletions shortened RLS with p-value < 0.05. This means that 464 genes, when deleted, show no detectable effect on RLS (p-value > 0.05). We did a regression on effect on life span ~ competitive growth fitness in these 564 genes. For lifespan extension deletions, partial correlation is observed between life-span extension effect and growth fitness in rich media (YPD) (p-value =0.038). For the entire data set, deletion effect on RLS is significantly correlated with fitness measured in media with ethanol as the main carbon source (YPE) (Fig 4, p-value = 0.00159, R-square = 0.04). This is related to the superoxide dismutase (SOD) activity and the reactive oxygen species (ROS), because YPE gives the lowest SOD activity [73]. Interestingly, over-expression of SOD can extend life span of laboratory strains of fruit flies with shortened life spans, but fails to do so in fruit flies that are recently isolated from the wild [74, 75]. Based on these preliminary results, it is very promising that larger scale analysis can yield interesting and new insights on yeast aging.

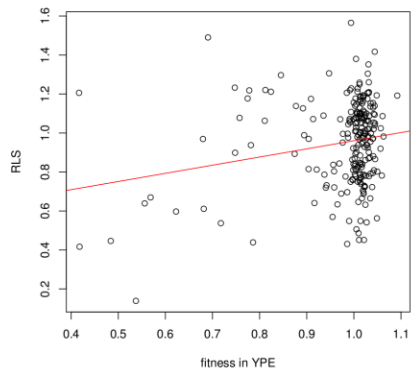

Fig. 4. RLS is correlated with growth fitness in YPE, suggesting the role of SOD. However, this correlation is largely due to deletion mutant with severe defects.

A large scale CLS data for yeast deletion collection are available from the Kaeberlein and the Fields groups [30]. We found that CLS in this dataset is highly correlated with growth fitness in YPD and YPE (p-value<0.003), which is likely due to the re-growth technique used during the screen. However, we did not find this correlation in the top 100 ranked genes, which is similar to what we found in the RLS data set.

It is also possible that a gene's role in robustness is linked to the number of its functional pathways. Thus, the number of GO annotations of each gene will also be included in the regression analysis of robustness.

5.2 Testing the association between genes associated with aging and phenotypic capacitors

Our hypothesis predicts that genes associated with aging likely play roles in network robustness. Phenotypic capacitors, such as Hsp90, are also known to play important roles in network robustness [3, 76]. Hence, we speculate that genes associated with aging and phenotypic capacitors tend to share some common characteristics, such as network topological features (connectivity, betweenness, etc), expression profiles, enriched GO categories, or similar growth fitness. Interestingly, genes associated with aging and phenotypic capacitors were both found to have more interactions in two separate reports [54, 76]. Hence, these two groups of genes will at least share some network topological features.

Levy and Siegal ranked 1583 genes' potential in phenotypic capacitance [76]. Among this list, 186 and 548 genes have been assayed for RLS and CLS, respectively. These samples are reasonably large, and will be used in our multiple regression analysis.

5.3 Predicting new genes associated with aging.

Based on the discovered associations in 5.1 and 5.2, we will use machine learning techniques to generate a list of new candidate genes associated with aging for experimental tests. Tentatively, we propose to use support vector machine (SVM) based on libsvm (http://www.csie.ntu.edu.tw/~cjlin/libsvm) with an R interface implemented in the e1071 package. SVM is a supervised learning method for classification [77]. It first needs a training dataset, in this case, a list of genes with known effects on cellular aging marked as "shortening" or "extension" of life spans. The trained SVM model will then be run on a large dataset to make predictions.

As a proof of principle, we built a SVM model using a short list of genes with strong effect on aging [78] and used principal components of one set of gene expression data for prediction [79]. Remarkably, this simple SVM model could predict some new genes associated with aging with reasonable biological functions, including glucose sensing and glycolysis pathways. We therefore propose to build a more sophisticated SVM model based on multiple large-scale datasets. There are several ways to integrate different datasets into SVM. We can either extract principal components from all of them, or we can train and run them separately and then reconcile the different predictions.

We will give extra weights to some datasets which we think are informative to aging and robustness. For example, we are particularly interested in synthetic lethal gene pairs (SLGP). These synthetic lethal pairs are pair of genes that result in lethal phenotypes when both are deleted, but non-lethal if only one of them is deleted. Conceptually, when only one copy of SLGP is deleted, the robustness of the disrupted gene network is likely reduced. (It is apparent that gene's contribution to robustness is non-linear, because single deletion of SLGP often has competitive growth fitness > 0.8.) We have curated 441 synthetic lethal gene pairs (SLGP) in yeast. Among them, RLS of 22 deletions are currently measured. Five of them (AKR1, ARP1, EFT1, PAC10, SAC6) can significantly shorten RLS ($p<0.05$) and one (TIF1) can significantly extend RLS ($p<0.05$). Interestingly, both EFT1 and TIF1 are duplicated genes. EFT1 is nearly identical to EFT2, both encode the Elongation factor 2 which catalyzes ribosomal translocation during protein synthesis. TIF1 is nearly identical to TIF2, both encode the translation initiation factor elF4A. Despite of their shared roles in protein translation and the same evolutionary history of duplication, deletions of EFT1 and TIF1 have opposite effects on replicative life span. This preliminary analysis shows that more sophisticated analyses with much larger scale data sets are very promising.

5.4 Expected computational outcomes

Based on the hypothesis that cellular aging is an emergent property of the gene/protein networks, characteristics of cellular aging are expected to be associated with proxies of robustness. Genes associated with aging and phenotypic capacitors are expected to share some characteristics. Given the complexity of the aging process, correlations are likely to be found among multiple factors. Predicted genes with potential roles in cellular aging will be tested experimentally, which in turn, can verify and improve the computational analyses and predictions.

5.5 Alternative computational strategies

It is important to emphasize that our proposed explorative study include multiple approaches and multiple data sets. Here, we discuss a few strategic alternatives.

In addition to the proposed hypothesis driven research plan, we may consider an "unbiased and non-subjective" approach to look for association between characteristics of cellular aging with *any* available phenotypic, functional, and sequence measures. As a model organism of cell biology and molecular genetics, abundant genome-scale data are available for *S. cerevisiae*, including sequence-based measures using comparative genomics, sub-cellular localization, half-lives of mRNA and proteins, and many published expression data sets. Discovered associations from this approach will then force us to re-examine our hypothesis and the current thinking of aging.

In addition to gene/protein networks, we may consider whether features of metabolic pathways and co-expression networks are associated with characteristics of aging. We can also incorporate cis-regulations into the gene/protein networks. Recently, Huang and Fraenkel integrated protein network, genetic network and transcription network using a constraint optimization method and found many unexpected changes during the yeast pheromone response [80]. Similar study on cellular aging may also reveal novel findings.

In addition to SVM, another predictive method is the Bayesian approach, similar to the naïve Bayesian approach used to predict protein interactions [81]. Training of the Bayesian model can be achieved by the relatively large number of known genes associated with aging.

As are the cases in all computational analyses, critical evaluation by human experts is an important step during interpretation and conceptualization. For example, we will choose Hsp90 as our top priority in experimental tests, because it is a phenotypic capacitor that has been well-studied by Dr. Suzannah Rutherford (see her support letter). Interestingly, yeast Hsp90 genes (Hsp82 and Hsc82) were not identified as phenotypic capacitors during the screen conducted by Levy and Siegal, which is either due to false negatives or limitation of the morphological measures.

We are aware that some genes can significantly influence the rate of aging but not the average life span, and some genes can significantly influence the average life span but not the rate of aging. The first case will lead to "false negatives" during the large scale screens based on life span. The second case will lead to "false positives" in our analysis. These two caveats will be considered during the discussion and interpretation of our results. We like to emphasize that our experimental verifications can identify the potential false positives.

Aim 2 (Experimental Tests): We will examine whether reduced network robustness will lead to slower rate of cellular aging. Specifically, we will test whether tolerances to Hsp90 and TOR inhibitors and oxidative stress in a collection of natural isolates of yeast are positively linked to their rates of aging, and whether deletion of the predicted candidate genes with possible effects on aging can influence the rate of aging.

The predicted effect of network robustness on the rate of cellular aging can be seen in Fig 2. We will study several measurable proxies of robustness: tolerance to Hsp90 inhibitors, tolerance to TOR inhibitors, and tolerance to oxidative stress. One important resource is our collection of natural isolates of yeast with measured RLS and CLS and quantified Gompertz parameters (Table 1). Another resource is the yeast gene deletion collection. A Bioscreen C is requested to facilitate high throughput growth curve assays.

5.6 Is tolerance to Hsp90 inhibitors associated with the rate of aging?

Hsp90 is a phenotypic capacitor that can buffer mutations in its substrate proteins [82, 83]. As a result, Hsp90 plays an important role in network robustness. Hsp90 inhibitors include geldanamycin, 17AAG, and radicicol (available from SIGMA). We will examine whether tolerances to these Hsp90 inhibitors are associated with the rate of aging, defined by the Gompertz coefficient, in a collection of yeast natural isolates. Two copies of the Hsp90 gene exist in yeast genome: HSC82 and HSP82, which makes molecular manipulation a complicated process [84]. Using engineered haploid lab strains (W303), Harris et al found that reduced Hsp90 activities do not change the mean RLS [85]. However, neither Gompertz modeling nor sophisticated statistical tests such as LRT was used. In fact, based on visual examination of their data (Figure 1 in their report), the drop of viability is sharper in the wildtype but less so in the mutants with reduced Hsp90 activities [85]. This indicates that G is larger in the wildtype but is smaller in the mutants, which is consistent with our hypothesized role of Hsp90 in cellular aging. Notice that the rate of aging can change significantly without changing the mean life span (See Eq3).

We found many single-nucleotide polymorphisms in the promoters and coding regions of HSC82 and HSP82 (http://www.sanger.ac.uk/gbrowse/gbrowse/cere_dmc/#search). It is also known that certain

fungal Hsp90 is resistant to radicicol inhibition due to a leucine-to-isoleucine substitution [86]. Hence, reasonable yeast natural variation should exist in responses to Hsp90 inhibitors.

One way to assay the tolerance to these inhibitors is to grow yeast strains in different concentrations of these inhibitors using a high throughput growth curve assay instrument – such as the Bioscreen C (http://www.growthcurvesusa.com/) . From the lag, mid-point, and plateau of the growth curves, we can estimate each strain's dose-dependent responses, such as the lethal dose at 50% (LD50), (for similar applications, see [30, 87]).

We will use regression to test whether tolerances to Hsp90 inhibitors are associated with their rates of aging in these natural isolates of yeast. For one or two strains, we will perform replicative life span assay under several inhibitor doses, for example, LD10, LD25, and LD50. We will then use regression to test whether aging rates are associated with doses of Hsp90 inhibitors.

Alternatively, flow cytometry can be used to detect the fraction of live and dead cells. The LIVE/DEAD® Cell Viability kit from Invitrogen will be used to stain yeast cells. Finally, plating assays and colony-forming units will be used to verify the high-throughput measures.

Because of the complexities of aging, we will look for partial correlation between tolerance to Hsp90 inhibitors and aging rates by controlling for other factors, such as responses to oxidative stress (see below), tolerance to TOR inhibitors (see below), genomic instability and mitotic asymmetry (see our previous publication [88]).

5.7 Is tolerance to oxidative stress associated with the rate of aging?

Oxidative stress is a probable environmental factor that is often experienced by yeast. Yeast strain's tolerance to oxidative stress indicates its ability to maintain homeostasis in the face of environmental changes, and is therefore a reasonable indicator of its robustness. We did a preliminary survey of strains' tolerance to paraquat, a compound that can induce oxidative stress (Fig. 5). We found four strains with LD50 of paraquat < 0.2 mM and three strains with LD50 > 0.2mM. The CLSs between these two groups are different with p-value=0.057 (Wilcoxon test). This preliminary survey was done using spot assay, a quick but inaccurate method. We propose to do high throughput growth curves using Bioscreen C to titrate the responses of these strains to various doses of paraquat and H_2O_2.

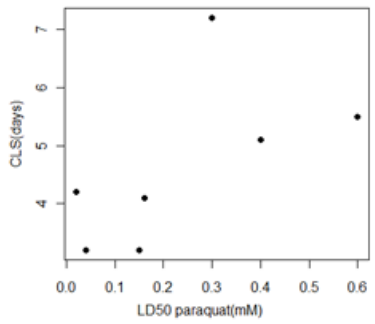

Fig. 5. Resistance to paraquat induced oxidative stress is correlated with CLS.

5.8 Is tolerance to TOR inhibitors associated with the rate of aging?

TOR is involved in nutrient sensing pathway and is a key player in the response to calorie restriction [4, 18, 89]. Yeast genome contains two copies of TOR proteins, TOR1 and TOR2. Rapamycin is a highly specific inhibitor of TOR proteins [90]. Decreased TOR activities extend both RLS and CLS in yeast [18, 30]. TOR and Sch9/PKA act in concert in nutrient response, and regulates cell cycle progression, translation, stress response, and autophagy [9]. Given the pleiotropic role of TOR pathway, it is likely that TOR pathway plays a role in cell's robustness in response to various environmental changes. Hence, tolerance to TOR inhibitors, such as rapamycin, may be indicative of TOR involvement in cell's robustness.

Assays of rapamycin tolerance in different strains will be conducted in similar ways to Hsp90 inhibitors. It is possible that there are inter-connections among TOR, Hsp90, and oxidative response pathways. We will test this possible link by regression analysis of tolerances to Hsp90 and TOR inhibitors and oxidative stress in our collection of natural isolates of yeast.

5.9 Can deletion of predicted genes influence the rate of aging?

From the computational analysis and empirical verification, we will generate a list of genes that can potentially influence the rate of yeast aging. To experimentally test some candidate genes, we will perform replicative life span assays for the deletion mutants of these genes. (Deletion mutants will be ordered from the yeast gene deletion collection, available at Open Biosystems). First, haploid mutants in BY4742 (MATalpha) background will be assayed. If a deletion mutant in MATalpha background can significantly influence the rate of aging, haploid MATa and diploid deletion strains will be evaluated.

The rates of aging, defined here as the Gompertz coefficients, will be compared among samples using the likelihood ratio test of nested models (See our preliminary results). It is possible that a gene deletion can significantly alter *I* and *G* but not the mean life span. This paradox can be understood from the G-I plot in Figure 3 and Eq 3, and to the best of our knowledge, has not been addressed by others. Hence, rigorous quantification and likelihood based comparison are unique aspects and strength of our proposed study.

5.10 Expected experimental outcomes and alternative plans.

Based on our hypothesis that cellular aging is an emergent property of gene/protein networks, it is expected that *G* is correlated with tolerances to Hsp90 or TOR inhibitors, or oxidative stress. We also expect inter-correlations of tolerances to Hsp90 and TOR inhibitors and oxidative stress. Given the complexity of aging, we expect partial correlations in most cases. When cells age in the presence of Hsp90 inhibitors, we expected smaller *G* values and slower dying-off phases (Fig 2). In the event that no correlation was found, the finding will be quite startling to not only us but also many other researchers. For example, it will force us to rethink many current models on aging, especially the ROS theory of aging.

Currently, the number of our phenotyped natural isolates of yeast is 14. Although we have uncovered some significant and interesting associations of traits using this collection, many significant but weak associations have likely been missed. We plan to increase the sample size of our phenotyped natural isolates to 30~40, by purchasing the yeast wild isolates sequenced by the Sanger center (http://www.sanger.ac.uk/Teams/Team118//sgrp/).

An alternative direction is the tolerance to Sir2 inhibitors, such as splitomicin and its analogoues [91-93]. The controversial role SIR2 in yeast aging and calorie restriction is strain-dependent, suggesting variations of SIR2 activities among strains. Another alternative direction is to survey the variation in responses of calorie restriction and examine their effect on *I* and *G* parameters.

In the event that computational prediction does not provide promising candidates, we will rely on the evaluations of human experts. For example, the role SSD1-V allele has recently been found to be pleiotropic and is important for longevity [94].

5.11 Summary

Our proposed computational and experimental components are closely related and complementary. Computational component will be refined by experimental data, and interpretation of experimental data demands rigorous quantification. We anticipate that new experiments will likely be designed to test new findings from computational analysis, and new computational analysis and modeling will be needed based on experimental findings. The integration of computing, mathematics and biology in this project is also an important education aspect for the undergraduate student participants.

6. Project Management, Milestones, and Assessment.

P.I. Qin will recruit five undergraduate volunteers during the Fall and Spring semesters, three undergraduate researchers during the Summer semesters, and a full-time research technician to join this project. We will focus on sophomores and juniors to ensure some continuity of the personnel. (Budget is requested for three summer students and a full time technician). Both the P.I. and the research technician will coordinate the efforts and schedule experiments with the students.

In Year 1, we expect to train the technician and the first batch of students, conduct computational analysis and part of the experiments. In Year 2, we will conduct the bulk of experimental tests. The

technician and students from year 1 are expected to play key roles in both training new students and designing and carrying out experiments.

Under P.I.'s guidance, students will learn to analyze, interpret, and synthesize their findings. Their contributions are encouraged to the extent that can lead to co-authorships in manuscripts. Students will attend research meetings to learn from their peers in Year 1. They are expected to present their own work at various meetings in Year 2. Possible meetings include the annual student research day at Spelman College, the HBCU-UP national meeting, and the National Conferences on Undergraduate Research (NCUR).

7. Intellectual Merit and Long-term Goals

The ***intellectual merit*** of this project lies in our novel hypothesis of cellular aging. Outcomes of this project have the potential to provide a fresh network perspective to extend the classic view of genotype space and phenotype space. The classic view of transition from genotype space to phenotype space was often credited to Richard Lewontin [95]. Given a complex trait such as aging, and especially given that aging is an emergent property of gene/protein networks, networks should be considered in the transition between genotype and phenotype spaces. It is unclear how connection patterns and dynamics of genes/proteins networks shape the transition from genotypic space to the phenotypic space. Plasticity of yeast aging is the manifestation of genotypic variation, stochastic fluctuation, and environmental cues. Some of these variations can be canalized by gene/protein networks, whereas others can be amplified through gene/protein networks. The propose project will add insights on how gene/protein networks can shape complex phenotypes.

This project is part of the P.I.'s ***long-term plan*** to study the gene/protein networks and complex traits and to understand the transition between genotypes and phenotypes from the perspective of networks. Dr. Qin is using simulation to study the emergence of aging and the theoretic implication of network on aging, such as the roles of different network configurations on characteristics of aging. He is also interested in the genetic mechanism of natural variation in lifespan and age-dependent traits, and their evolutionary and ecological implications. Overall, this project will open the gate to many new research directions.

8. Broader Impacts

The ***broader impacts*** can be found on student experiences, curricula improvement, and the research program at Spelman College. As a Biology faculty member at a historically black women's liberal arts college, this project will be carried out by our students. The interdisciplinary nature and genome-scale studies of this project will add a sense of excitement to students, as well as providing the much-needed training for these future scientists. Dr. Qin teaches Genomic and Proteomics, Microbiology, the HHMI phage genomics course, and the student seminar series. These courses target students from freshmen to seniors, and offer a diverse pool for recruitment. Since his start at Spelman in August, 2009, two junior students have been working with him on yeast aging. One biology student, Meighan Parker, focuses on the experimental side, while another student, Lolade Bolaji, a mathematics major, focuses on the computational side. Part of this project will also be integrated into Dr. Qin's courses, which will enrich the learning experiences of many more students.

We like to emphasize that this project is both practically assessable by undergraduates and intellectually challenging. Students will learn to deal with the frustrations and rewards that will inevitably arise from thinking through the challenging questions and problems posted in this project. Dr. Qin will also strive to instill the fundamental values and ethics of research into the students. By taking ownership of this project, students may discover their love of research and may choose science-related careers. Dr. Qin especially encourages his students to learn more about genomics and computational biology. Finally, this project will invigorate the research program and improve the research capacity at a historically black college.

References Cited

1. Gavrilov, L.A. and N.S. Gavrilova, *The reliability theory of aging and longevity.* J Theor Biol, 2001. **213**(4): p. 527-45.
2. Balouri, H., *Computational modeling of gene regulatory networks - a primer*. 2008, London: Imperial College Press.
3. Masel, J. and M.L. Siegal, *Robustness: mechanisms and consequences.* Trends Genet, 2009. **25**(9): p. 395-403.
4. Blagosklonny, M.V., *Aging: ROS or TOR.* Cell Cycle, 2008. **7**(21): p. 3344-54.
5. Stewart, E. and F. Taddei, *Aging in Escherichia coli: signals in the noise.* Bioessays, 2005. **27**(9): p. 983.
6. Stewart, E.J., R. Madden, G. Paul, and F. Taddei, *Aging and death in an organism that reproduces by morphologically symmetric division.* PLoS Biol, 2005. **3**(2): p. e45.
7. Henderson, K.A. and D.E. Gottschling, *A mother's sacrifice: what is she keeping for herself?* Curr Opin Cell Biol, 2008. **20**(6): p. 723-8.
8. Mortimer, R.K. and J.R. Johnston, *Life span of individual yeast cells.* Nature, 1959. **183**(4677): p. 1751-2.
9. Kaeberlein, M., C.R. Burtner, and B.K. Kennedy, *Recent developments in yeast aging.* PLoS Genet, 2007. **3**(5): p. e84.
10. Kennedy, B.K., K.K. Steffen, and M. Kaeberlein, *Ruminations on dietary restriction and aging.* Cell Mol Life Sci, 2007. **64**(11): p. 1323-8.
11. Finkel, T., M. Serrano, and M.A. Blasco, *The common biology of cancer and ageing.* Nature, 2007. **448**(7155): p. 767-74.
12. Park, P.U., M. McVey, and L. Guarente, *Separation of mother and daughter cells.* Methods Enzymol, 2002. **351**: p. 468-77.
13. Steffen, K.K., B.K. Kennedy, and M. Kaeberlein, *Measuring replicative life span in the budding yeast.* J Vis Exp, 2009(28).
14. Fabrizio, P. and V.D. Longo, *The chronological life span of Saccharomyces cerevisiae.* Aging Cell, 2003. **2**(2): p. 73-81.
15. Fabrizio, P. and V.D. Longo, *The chronological life span of Saccharomyces cerevisiae.* Methods Mol Biol, 2007. **371**: p. 89-95.
16. Qin, H. and M. Lu, *Natural variation in replicative and chronological life spans of Saccharomyces cerevisiae.* Exp Gerontol, 2006. **41**(4): p. 448-56.
17. Kaeberlein, M., M. McVey, 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.
18. Kaeberlein, M., R.W. Powers, 3rd, K.K. Steffen, E.A. Westman, D. Hu, N. Dang, E.O. Kerr, K.T. Kirkland, S. Fields, and B.K. Kennedy, *Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients.* Science, 2005. **310**(5751): p. 1193-6.
19. Fabrizio, P., F. Pozza, S.D. Pletcher, C.M. Gendron, and V.D. Longo, *Regulation of longevity and stress resistance by Sch9 in yeast.* Science, 2001. **292**(5515): p. 288-90.
20. Sinclair, D.A. and L. Guarente, *Extrachromosomal rDNA circles--a cause of aging in yeast.* Cell, 1997. **91**(7): p. 1033-42.
21. Ganley, A.R., S. Ide, K. Saka, and T. Kobayashi, *The effect of replication initiation on gene amplification in the rDNA and its relationship to aging.* Mol Cell, 2009. **35**(5): p. 683-93.

22. Blagosklonny, M.V., *Aging-suppressants: cellular senescence (hyperactivation) and its pharmacologic deceleration.* Cell Cycle, 2009. **8**(12): p. 1883-7.
23. Wei, M., P. Fabrizio, F. Madia, J. Hu, H. Ge, L.M. Li, and V.D. Longo, *Tor1/Sch9-regulated carbon source substitution is as effective as calorie restriction in life span extension.* PLoS Genet, 2009. **5**(5): p. e1000467.
24. Sinclair, D.A., S.J. Lin, and L. Guarente, *Life-span extension in yeast.* Science, 2006. **312**(5771): p. 195-7; author reply 195-7.
25. Steinkraus, K.A., M. Kaeberlein, and B.K. Kennedy, *Replicative aging in yeast: the means to the end.* Annu Rev Cell Dev Biol, 2008. **24**: p. 29-54.
26. Kirkwood, T.B. and C.E. Finch, *Ageing: the old worm turns more slowly.* Nature, 2002. **419**(6909): p. 794-5.
27. Mazzocchi, F., *Complexity in biology. Exceeding the limits of reductionism and determinism using complexity theory.* EMBO Rep, 2008. **9**(1): p. 10-4.
28. de Magalhaes, J.P. and G.M. Church, *Cells discover fire: employing reactive oxygen species in development and consequences for aging.* Exp Gerontol, 2006. **41**(1): p. 1-10.
29. Medvedik, O., D.W. Lamming, K.D. Kim, and D.A. Sinclair, *MSN2 and MSN4 link calorie restriction and TOR to sirtuin-mediated lifespan extension in Saccharomyces cerevisiae.* PLoS Biol, 2007. **5**(10): p. e261.
30. Powers, R.W., 3rd, M. Kaeberlein, S.D. Caldwell, B.K. Kennedy, and S. Fields, *Extension of chronological life span in yeast by decreased TOR pathway signaling.* Genes Dev, 2006. **20**(2): p. 174-84.
31. Wagner, A., *Robustness against mutations in genetic networks of yeast.* Nat Genet, 2000. **24**(4): p. 355-61.
32. Gu, Z., L.M. Steinmetz, X. Gu, C. Scharfe, R.W. Davis, and W.H. Li, *Role of duplicate genes in genetic robustness against null mutations.* Nature, 2003. **421**(6918): p. 63-6.
33. Vaupel, J.W., K.G. Manton, and E. Stallard, *The impact of heterogeneity in individual frailty on the dynamics of mortality.* Demography, 1979. **16**(3): p. 439-54.
34. Kauffman, S., *The Origins of Order: Self-organization and Selection in Evolution*. 1993, New York, U.S.A.: Oxford University Press.
35. Franceschi, C., *Cell proliferation, cell death and aging.* Aging (Milano), 1989. **1**(1): p. 3-15.
36. Kirkwood, T.B. and A. Kowald, *Network theory of aging.* Exp Gerontol, 1997. **32**(4-5): p. 395-9.
37. Kowald, A. and T.B. Kirkwood, *A network theory of ageing: the interactions of defective mitochondria, aberrant proteins, free radicals and scavengers in the ageing process.* Mutat Res, 1996. **316**(5-6): p. 209-36.
38. Soti, C. and P. Csermely, *Aging cellular networks: chaperones as major participants.* Exp Gerontol, 2007. **42**(1-2): p. 113-9.
39. Rattan, S.I., *Gerontogenes: real or virtual?* FASEB J, 1995. **9**(2): p. 284-6.
40. Fukui, H.H., L. Xiu, and J.W. Curtsinger, *Slowing of age-specific mortality rates in Drosophila melanogaster.* Exp Gerontol, 1993. **28**(6): p. 585-99.
41. McGue, M., J.W. Vaupel, N. Holm, and B. Harvald, *Longevity is moderately heritable in a sample of Danish twins born 1870-1880.* J Gerontol, 1993. **48**(6): p. B237-44.
42. Herskind, A.M., M. McGue, N.V. Holm, T.I. Sorensen, B. Harvald, and J.W. Vaupel, *The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870-1900.* Hum Genet, 1996. **97**(3): p. 319-23.

43. Williams, G.C., *Pleiotropy, natural selection and the evolution of senescence.* Evolution, 1957: p. 398-411.
44. Kirkwood, T.B., *Evolution of ageing.* Nature, 1977. **270**(5635): p. 301-4.
45. Kaeberlein, M., K.T. Kirkland, S. Fields, and B.K. Kennedy, *Sir2-independent life span extension by calorie restriction in yeast.* PLoS Biol, 2004. **2**(9): p. E296.
46. Kaeberlein, M. and R.W. Powers, 3rd, *Sir2 and calorie restriction in yeast: a skeptical perspective.* Ageing Res Rev, 2007. **6**(2): p. 128-40.
47. Xue, H., B. Xian, D. Dong, K. Xia, S. Zhu, Z. Zhang, L. Hou, Q. Zhang, Y. Zhang, and J.D. Han, *A modular network model of aging.* Mol Syst Biol, 2007. **3**: p. 147.
48. Bell, R., A. Hubbard, R. Chettier, D. Chen, J.P. Miller, P. Kapahi, M. Tarnopolsky, S. Sahasrabuhde, S. Melov, and R.E. Hughes, *A human protein interaction network shows conservation of aging processes between human and invertebrate species.* PLoS Genet, 2009. **5**(3): p. e1000414.
49. Budovsky, A., A. Abramovich, R. Cohen, V. Chalifa-Caspi, and V. Fraifeld, *Longevity network: construction and implications.* Mech Ageing Dev, 2007. **128**(1): p. 117-24.
50. Smith, E.D., M. Tsuchiya, L.A. Fox, N. Dang, D. Hu, E.O. Kerr, E.D. Johnston, B.N. Tchao, D.N. Pak, K.L. Welton, D.E. Promislow, J.H. Thomas, M. Kaeberlein, and B.K. Kennedy, *Quantitative evidence for conserved longevity pathways between divergent eukaryotic species.* Genome Res, 2008. **18**(4): p. 564-70.
51. Smith, E.D., B.K. Kennedy, and M. Kaeberlein, *Genome-wide identification of conserved longevity genes in yeast and worms.* Mech Ageing Dev, 2007. **128**(1): p. 106-11.
52. Lorenz, D.R., C.R. Cantor, and J.J. Collins, *A network biology approach to aging in yeast.* Proc Natl Acad Sci U S A, 2009. **106**(4): p. 1145-50.
53. Barea, F. and D. Bonatto, *Aging defined by a chronologic-replicative protein network in Saccharomyces cerevisiae: an interactome analysis.* Mech Ageing Dev, 2009. **130**(7): p. 444-60.
54. Promislow, D.E., *Protein networks, pleiotropy and the evolution of senescence.* Proc Biol Sci, 2004. **271**(1545): p. 1225-34.
55. Comfort, A., *The biology of senescence*. 3rd ed. 1979, New York: Elsevier.
56. Finch, C.E., *Longevity, Senescence, and the Genome*. 1990, Chicago: The University of Chicago Press.
57. Sacher, G.A., *Life table modification and life prolongation.*, in *Handbook of the Biology of Aging*, C.E. Finch and L. Hayflick, Editors. 1977, Van Nostrand Reinhold Company: New York. p. 582-638.
58. Boxenbaum, H., *Gompertz mortality analysis: aging, longevity hormesis and toxicity.* Arch Gerontol Geriatr, 1991. **13**(2): p. 125-37.
59. Wilson, D.L., *A comparison of methods for estimating mortality parameters from survival data.* Mech Ageing Dev, 1993. **66**(3): p. 269-81.
60. Wilson, D.L., *The analysis of survival (mortality) data: fitting Gompertz, Weibull, and logistic functions.* Mech Ageing Dev, 1994. **74**(1-2): p. 15-33.
61. Strehler, B.L. and A.S. Mildvan, *General theory of mortality and aging.* Science, 1960. **132**: p. 14-21.
62. Managbanag, J.R., T.M. Witten, D. Bonchev, L.A. Fox, M. Tsuchiya, B.K. Kennedy, and M. Kaeberlein, *Shortest-path network analysis is a useful approach toward identifying genetic determinants of longevity.* PLoS One, 2008. **3**(11): p. e3802.

63. Dean, E.J., J.C. Davis, R.W. Davis, and D.A. Petrov, *Pervasive and persistent redundancy among duplicated genes in yeast.* PLoS Genet, 2008. **4**(7): p. e1000113.
64. Cooper, T.F., A.P. Morby, A. Gunn, and D. Schneider, *Effect of random and hub gene disruptions on environmental and mutational robustness in Escherichia coli.* BMC Genomics, 2006. **7**: p. 237.
65. Ohya, Y., J. Sese, M. Yukawa, F. Sano, Y. Nakatani, T.L. Saito, A. Saka, T. Fukuda, S. Ishihara, S. Oka, G. Suzuki, M. Watanabe, A. Hirata, M. Ohtani, H. Sawai, N. Fraysse, J.P. Latge, J.M. Francois, M. Aebi, S. Tanaka, S. Muramatsu, H. Araki, K. Sonoike, S. Nogami, and S. Morishita, *High-dimensional and large-scale phenotyping of yeast mutants.* Proc Natl Acad Sci U S A, 2005. **102**(52): p. 19015-20.
66. Promislow, D., *A regulatory network analysis of phenotypic plasticity in yeast.* Am Nat, 2005. **165**(5): p. 515-23.
67. Dudley, A.M., D.M. Janse, A. Tanay, R. Shamir, and G.M. Church, *A global view of pleiotropy and phenotypically derived gene function in yeast.* Mol Syst Biol, 2005. **1**: p. 2005 0001.
68. Qin, H. and L. Yang, *Detection of changes in transitive associations by shortest-path analysis of protein interaction networks integrated with gene expression profiles.* Proceedings of the IEEE International Conference on Biomedical Engineering and Informatics, 2008. **1**: p. 418-423.
69. Gilchrist, M.A., H. Qin, and R. Zaretzki, *Modeling SAGE tag formation and its effects on data interpretation within a Bayesian framework.* BMC Bioinformatics, 2007. **8**: p. 403.
70. Sean, D. and P.S. Meltzer, *GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor.* Bioinformatics, 2007. **23**(14): p. 1846-7.
71. Steinmetz, L.M., C. Scharfe, A.M. Deutschbauer, D. Mokranjac, Z.S. Herman, T. Jones, A.M. Chu, G. Giaever, H. Prokisch, P.J. Oefner, and R.W. Davis, *Systematic screen for human disease genes in yeast.* Nat Genet, 2002. **31**(4): p. 400-4.
72. Xie, M.W., F. Jin, H. Hwang, S. Hwang, V. Anand, M.C. Duncan, and J. Huang, *Insights into TOR function and rapamycin response: chemical genomic profiling by using a high-density cell array method.* Proc Natl Acad Sci U S A, 2005. **102**(20): p. 7215-20.
73. Macierzynska, E., A. Grzelak, and G. Bartosz, *The effect of growth medium on the antioxidant defense of Saccharomyces cerevisiae.* Cell Mol Biol Lett, 2007.
74. Orr, W.C., R.J. Mockett, J.J. Benes, and R.S. Sohal, *Effects of overexpression of copper-zinc and manganese superoxide dismutases, catalase, and thioredoxin reductase genes on longevity in Drosophila melanogaster.* J Biol Chem, 2003. **278**(29): p. 26418-22.
75. Spencer, C.C., C.E. Howell, A.R. Wright, and D.E. Promislow, *Testing an 'aging gene' in long-lived drosophila strains: increased longevity depends on sex and genetic background.* Aging Cell, 2003. **2**(2): p. 123-30.
76. Levy, S.F. and M.L. Siegal, *Network hubs buffer environmental variation in Saccharomyces cerevisiae.* PLoS Biol, 2008. **6**(11): p. e264.
77. Guyon, I., J. Weston, S. Barnhill, and V. Vapnik, *Gene Selection for Cancer Classification using Support Vector Machines.* Mach. Learn., 2002. **46**(1-3): p. 389-422.
78. Kaeberlein, M. and B.K. Kennedy, *Large-scale identification in yeast of conserved ageing genes.* Mech Ageing Dev, 2005. **126**(1): p. 17-21.
79. Gasch, A.P., P.T. Spellman, C.M. Kao, O. Carmel-Harel, M.B. Eisen, G. Storz, D. Botstein, and P.O. Brown, *Genomic expression programs in the response of yeast cells to environmental changes.* Mol Biol Cell, 2000. **11**(12): p. 4241-57.

80. Huang, S.S. and E. Fraenkel, *Integrating proteomic, transcriptional, and interactome data reveals hidden components of signaling and regulatory networks.* Sci Signal, 2009. **2**(81): p. ra40.
81. Jansen, R., H. Yu, D. Greenbaum, Y. Kluger, N.J. Krogan, S. Chung, A. Emili, M. Snyder, J.F. Greenblatt, and M. Gerstein, *A Bayesian networks approach for predicting protein-protein interactions from genomic data.* Science, 2003. **302**(5644): p. 449-53.
82. Rutherford, S.L., *Between genotype and phenotype: protein chaperones and evolvability.* Nat Rev Genet, 2003. **4**(4): p. 263-74.
83. Rutherford, S.L. and S. Lindquist, *Hsp90 as a capacitor for morphological evolution.* Nature, 1998. **396**(6709): p. 336-42.
84. Nathan, D.F. and S. Lindquist, *Mutational analysis of Hsp90 function: interactions with a steroid receptor and a protein kinase.* Mol Cell Biol, 1995. **15**(7): p. 3917-25.
85. Harris, N., M. MacLean, K. Hatzianthis, B. Panaretou, and P.W. Piper, *Increasing Saccharomyces cerevisiae stress resistance, through the overactivation of the heat shock response resulting from defects in the Hsp90 chaperone, does not extend replicative life span but can be associated with slower chronological ageing of nondividing cells.* Mol Genet Genomics, 2001. **265**(2): p. 258-63.
86. Prodromou, C., J.M. Nuttall, S.H. Millson, S.M. Roe, T.S. Sim, D. Tan, P. Workman, L.H. Pearl, and P.W. Piper, *Structural basis of the radicicol resistance displayed by a fungal hsp90.* ACS Chem Biol, 2009. **4**(4): p. 289-97.
87. Fraud, S., J.Y. Maillard, M.A. Kaminski, and G.W. Hanlon, *Activity of amine oxide against biofilms of Streptococcus mutans: a potential biocide for oral care formulations.* J Antimicrob Chemother, 2005. **56**(4): p. 672-7.
88. Qin, H., M. Lu, and D.S. Goldfarb, *Genomic instability is associated with natural life span variation in Saccharomyces cerevisiae.* PLoS One, 2008. **3**(7): p. e2670.
89. Blagosklonny, M.V., *TOR-driven aging: speeding car without brakes.* Cell Cycle, 2009. **8**(24): p. 4055-9.
90. Bjornsti, M.A. and P.J. Houghton, *The TOR pathway: a target for cancer therapy.* Nat Rev Cancer, 2004. **4**(5): p. 335-48.
91. Posakony, J., M. Hirao, S. Stevens, J.A. Simon, and A. Bedalov, *Inhibitors of Sir2: evaluation of splitomicin analogues.* J Med Chem, 2004. **47**(10): p. 2635-44.
92. Posakony, J., M. Hirao, and A. Bedalov, *Identification and characterization of Sir2 inhibitors through phenotypic assays in yeast.* Comb Chem High Throughput Screen, 2004. **7**(7): p. 661-8.
93. Hirao, M., J. Posakony, M. Nelson, H. Hruby, M. Jung, J.A. Simon, and A. Bedalov, *Identification of selective inhibitors of NAD+-dependent deacetylases using phenotypic screens in yeast.* J Biol Chem, 2003. **278**(52): p. 52773-82.
94. Li, L., Y. Lu, L.X. Qin, Z. Bar-Joseph, M. Werner-Washburne, and L.L. Breeden, *Budding yeast SSD1-V regulates transcript levels of many longevity genes and extends chronological life span in purified quiescent cells.* Mol Biol Cell, 2009. **20**(17): p. 3851-64.
95. Lewontin, R., *The genetic basis of evolutionary change*. 1974: Columbia University Press.