Identification of longevity genes with systems biology approaches

Yuanyou Tan¹,³
John M Bush¹
Weijiu Liu²
Fusheng Tang¹

¹Department of Biology, University of Arkansas, Little Rock, AR, USA; ²Department of Mathematics, University of Central Arkansas, Conway, AR, USA; ³Department of Bioengineering, Wuhan University of Science and Engineering, Hubei, China

Abstract: Identification of genes involved in the aging process is critical for understanding the mechanisms of age-dependent diseases such as cancer and diabetes. Measuring the mutant gene lifespan, each missing one gene, is traditionally employed to identify longevity genes. While such screening is impractical for the whole genome due to the time-consuming nature of lifespan assays, it can be achieved by in silico genetic manipulations with systems biology approaches. In this review, we will introduce pilot explorations applying two approaches of systems biology in aging studies. One approach is to predict the role of a specific gene in the aging process by comparing its expression profile and protein–protein interaction pattern with those of known longevity genes (top-down systems biology). The other approach is to construct mathematical models from previous kinetics data and predict how a specific protein contributes to aging and antiaging processes (bottom-up systems biology). These approaches allow researchers to simulate the effect of each gene’s product in aging by in silico genetic manipulations such as deletion or over-expression. Since simulation-based approaches are not as widely used as the other approaches, we will focus our review on this effort in more detail. A combination of hypothesis from data-mining, in silico experimentation from simulations, and wet laboratory validation will make the systematic identification of all longevity genes possible.

Keywords: systems biology, yeast, aging, in silico genetic manipulation, modeling

Introduction

Aging is the age-dependent decline of cellular and organismal functions. While this decline is not selected by evolution and there are no genes dedicated to longevity, alterations of some genes with pleiotropic phenotypes shorten or extend the life span and led to the proposal of multiple aging pathways. SIR2 is one example of such longevity gene. Deletion of SIR2 shortens the life span and elevated expression of SIR2 extends the life span of yeast,¹,² nematodes,³ and flies.⁴ While the deacetylase function of SIRT1 (a mammalian homolog of yeast SIR2) protects heart cells from reactive oxygen species (ROS) damage,⁵ improves insulin sensitivity and decreases plasma glucose,⁶ suppresses colon tumor growth⁷ and exerts multiple other beneficial effects for cells, feeding midlife mice with resveratrol, which upregulates the expression of SIRT1, does not extend the life span.⁸ Therefore, identification of more genes involved in aging is an emergent task for aging researchers since the onset of most life-threatening disease are age-dependent.

At the post-genome era, completion of the genome-wide knock-out strains in yeast or knock-down strains in nematodes makes it possible to screen the whole genome for genes involved in the aging process. Multiple screens have been performed for
Caenorhabditis elegans, yielding non-SIR2 aging-related genes. Kaeberlein and Kennedy’s12 team screened a subset of the yeast knock-out collection and discovered other longevity genes including TOR1, whose product negatively controls autophagy, a vacuole-/lysosome-dependent degradation of damaged or obsolete materials. Our group conducted a partial screen of autophagy mutants and discovered that intra-vacuolar degradation of lipids is a key step in the anti-aging process.13 Intriguingly, these different screens identified different aging genes, indicating the need for a more exhaustive screen. However, the labor intensity of the life span assays and the interactions of multiple proteins during aging make it difficult to exhaustively identify longevity genes by traditional screening of single or double knock-out or knockdown mutants. A revolutionary strategy for identification of longevity genes is thus in demand.

The vast amount of ‘-omics’ data allowed us to establish mathematical models to simulate genuine intracellular networks and to estimate the effect of each gene on longevity with in silico approaches and then test their effect with real experiments. Genome-wide studies revealed networks of various types including regulatory networks, physical interaction networks, metabolic networks, etc. These networks can impact aging studies at multiple levels. First, these networks themselves provide ample information about the relationship between aging genes and other genes. Second, partitioning proteins in networks into clusters based on graph theories can predict functions of proteins since proteins with similar functions tend to form clusters. Third, integration of the networks with kinetics data into mathematical models can simulate the age-dependent variation of the function of aging proteins. Therefore, analyses of the ‘-omics’ data from multiple levels will guide the search for longevity genes.

**Expression profiling**

Comparisons of the genome-wide expression between young organisms and old individuals were achieved in multiple organisms.14 While this approach revealed which genes are up- or downregulated in old individuals, several constraints limit the application of this approach in identifying longevity genes. Most studies compared tissues from young animals and old animals.14 Since tissues from old animals are a combination of young and old cells, these comparisons likely miss some longevity-associated genes. The uni-cellular organism budding yeast thus leads the studies on expression profiling. Gordon’s group15 compared the expression pattern of replicative aged cells from a long life span mutant (snf4Δ) and a short life span mutant (sip2Δ) of the Snf1p glucose-sensing pathway and inferred that aged cells have elevated gluconeogenesis. With microanalytic biochemical methods, they confirmed that aged cells indeed shift away from glycolysis toward gluconeogenesis and energy storage. These results put Snf1p in a key position in the process of aging (Figure 1).

A variation of sampling is to compare the expression of normal and long-lived strains. In budding yeast Saccharomyces cerevisiae, Longo’s group identified that several stress-responsible transcription factors (Msn2p, Msn4p, Gis1p) are highly active in strains (sch9Δ, ras2Δ, sch9Δ)
that had longevity phenotype.\textsuperscript{16} However, the chronological longevity of \textit{sch9A} is not halted by deletion of \textit{MSN2} and/or \textit{MSN4} genes.\textsuperscript{17} Msn2p and Msn4p are transcription factors activated in stress conditions. The combination of microarray data, proteomics data, protein-protein interaction data, and other -omics data, in another word, studying the interrelationship of all the components that comprise a system is a trend in identifying genes involved in the aging process.\textsuperscript{18}

**Construction and analysis of aging networks**

Contrary to the reduction biology approaches, systems biology provides approaches from the top-down view. These approaches can be further divided into data-driven (from-omics to model) and hypothesis-driven (from kinetics data to mathematical model to predictions) approaches. A cell is composed of multiple structural and functional networks. The physical interactions of the cytoskeletal network, membrane trafficking network, and protein–protein interaction network provide the structural basis of a cell. The logic interactions of signal transduction network, gene regulatory network, and metabolic network provide the functional basis for a cell. Searching for an interesting gene by starting out using the network view is emerging as a profitable approach. A group of yeast cell biologists\textsuperscript{19} constructed a protein–protein interaction map for the development of yeast cell polarity and predicted the involvement of novel genes in this process. They predicted 20 new proteins in cell polarity development and confirmed that 13 proteins indeed involve in cell polarity by fluorescence microscopy. This pioneering work revealed that the systems approach can be a productive one in the search for longevity-related genes. Among different types of networks, the protein–protein interaction (PPI) network is the most straightforward one in terms of construction and application. By combing the known PPIs with the known longevity-related proteins, Budovsky and colleagues\textsuperscript{20} constructed a longevity PPI network. Their network predicts many aging-related proteins. For budding yeast, several sets of genome-wide protein-protein interaction data are available. Through the years, scientists also accumulated vast amount of PPI data for yeast cells (These data can be downloaded from the website of \textit{Saccharomyces} Genome Database, http://www.yeastgenome.org/). Analysis of the PPI network has multiple potential applications for aging studies for the following reasons. First, functions of genes are exerted by their protein products. Second, proteins in the same functional group tend to form clusters in the PPI network.\textsuperscript{21,22} Third, the PPI network provides a global view of the cell and can display the effect of cross-talks among different pathways. Fourth, the PPI network of yeast is the best-studied system compared with similar studies in other organisms.

Cell senescence is the accumulation of molecular damages.\textsuperscript{23} A cell will die when these damages attack to proteins that occupy hub positions in network. Thus, in a PPI network, most longevity genes are hubs. For example, Sir2p has about 70 interacting partner proteins. By simply extracting proteins interacting with the known longevity genes, we can construct an ‘aging network’ (Figure 1). Consistent with Gordon’s prediction,\textsuperscript{15} Snf1p functions as one hub in this aging network (Figure 1). Moreover, this network revealed some novel features of known aging proteins (genes). This figure shows that some longevity proteins are connected through common proteins such as Dbf2p and Pmr1p. While Dbf2p is involved in mitosis and stress responses,\textsuperscript{24} Pmr1p is involved in intracellular calcium and manganese homeostasis.\textsuperscript{25} Since these two proteins link several known longevity proteins, it is predictable that these two are also involved in aging. Intriguingly, one of our mathematical modeling approaches also predicted that Pmr1p is involved in longevity (see below).

The constructed aging network could be used to predict longevity genes from different angles. One application is to integrate one’s interested metabolic pathway into the aging network (Figure 1). The synthesis of ergosterol, the yeast counterpart of mammalian cholesterol, requires more than 20 enzymes. These enzymes interact with each other and form complexes (Figure 1). As shown in Figure 1, some Erg proteins (Erg6p) has more connections than other Erg proteins to predicted aging proteins. These connections suggest that Erg6p may be more important than other Erg enzymes in life span extension. Indeed, we found that deletion of \textit{ERG6} produces a strain that has a shorter life span than strains missing either \textit{ERG3} or \textit{ERG5} genes grown on calorie-restricted medium, a known longevity manipulation.\textsuperscript{15}

A more productive application of the PPI network is the prediction of functional protein clusters. One can partition proteins in the PPI network into multiple clusters based on the connectivity. If the connectivity density of a group of proteins is higher than certain threshold or than the connectivity between this group and other proteins, one can assign this group into a cluster of functional related proteins.\textsuperscript{26} Multiple algorithms can achieve such clustering.\textsuperscript{27} We invented a new algorithm and divided the yeast PPI network into about 200 clusters. Most of these clusters contain proteins sharing at least one common function according to gene ontology.\textsuperscript{28} These algorithms allowed us to do \textit{in silico}
genetic manipulations such as gene deletion. Deletion of a node from the network corresponds to the knock-out of a gene from the genome of an organism. After the *in silico* manipulation, one can use the clustering algorithms to analyze the mutated network. If a gene negatively controls longevity, deletion of that gene may result in clusters similar to clusters caused by *SIR2*-overexpression, *fob1*Δ or other longevity manipulations. While this perspective is very attractive, the static nature of most current PPI networks limits the application. The current PPI networks only represent the occurrence of protein-protein interactions; it does not include constraints such as where the proteins are located, when the proteins are synthesized and activated, etc. Thus, several groups are constructing dynamic networks, which include locations, time, and intensities of protein-protein interactions.29 Other groups are trying to incorporate the genome-wide expression data with the PPI network.30 In addition to temporal and spatial control of protein activities, the molecular heterogeneity during aging is another constraint. For example, aged (56-day-old) female rabbits have three-fold higher enzyme activities and different isoforms of NADP-dependent 17β-hydroxysteroid dehydrogenase than young (28-day-old) rabbits.31 The molecular heterogeneity should also be considered in constructing the dynamic PPI networks. *In silico* manipulation of the dynamic PPI network will allow a researcher to systematically knock-out each gene, or combinations of genes from the genome and to predict the outcome. Efforts in this direction will be very insightful for aging studies.

**Aging pathway**

Cells respond to environmental and intracellular stresses by re-programming the expression of related genes. In addition to data-driven approaches (genomics, proteomics, etc), these genes can also be identified by mathematical simulation and prediction, another aspect of systems biology. Control theories can successfully predict the effect of upregulation or downregulation of proteins involved in one function.32 At the cellular level, when a cell receives a stimulus, it will re-program its metabolic pathways to maintain the steady state. Inability to antagonize the stimuli will result in cell death. Aging is the process of losing such antagonizing capacities. The cytosol of a cell maintains cellular homeostasis including redox potential, ionic levels, etc. Disruptions of such homeostasis (biomarkers of aging) will initiate damage to cells.

Multiple biomarkers of cell senescence have been connected with aging to date. These biomarkers include increases in cytosolic calcium and ROS levels, which may in turn cause mutations in mitochondrial DNA or loss of heterozygosity in chromosomal DNA and lead to cell death. Although the cause–effect relationships for some of these markers are still issues of debate, an outline of progression of aging is emerging (Figure 2). ROS as a marker for cell senescence are generated by mitochondria and several other intracellular sources. The electron

---

**Figure 2** Cytosolic redox homeostasis and aging. Dotted arrow: potential causing agents. Arrow: stimulation. ⊥: inhibition. Cell senescence is caused by elevated calcium and/or ROS, which are in turn caused by a malfunction in the intracellular redox homeostasis.

**Abbreviations:** ER, endoplasmic reticulum; Gpx, glutathione peroxidase; SOD, superoxide dismutase.
transport chain of mitochondria is the main source. Other sources include a wide range of extramitochondrial enzymes, such as NADPH-oxidase and myeloperoxidase, and the endoplasmic reticulum, where the superoxide is generated by a leakage of electrons from NADPH-cytochrome-P450 reductase. In addition, elevated ROS was caused by elevated redox potentials including elevated GSSG levels and NADP+ levels.

ROS generation has been shown to be modulated by calcium. A rise of calcium can increase ROS. On the other hand, an elevated ROS may result in an increase in calcium. In mitochondria, it appears that calcium diminishes ROS from complexes I and III under normal conditions, but enhances ROS when the complex formations are inhibited. Deletion of yeast cytosolic thioredoxin peroxidase I greatly decreases the reduced glutathione (GSH)/the oxidized glutathione (GSSG) ratio in mitochondria upon calcium treatment. A low ratio of GSH/GSSG indicates a high oxidative potential within a cell. These previous observations suggest the contribution of cytosolic homeostasis to organelle functions and overall cellular redox homeostasis.

Another cell senescence marker is mutations in chromosomal DNA or mitochondrial DNA, which has been induced by ROS directly or by elevated cytosolic calcium. While the cytosolic calcium homeostasis is maintained by multiple mechanisms, attacks from ROS will decrease the calcium buffering capacities of cells. The steady state levels of each of the markers are balances between forces or metabolites that generate and that remove these markers. Therefore, fitting the proteins that produce these forces into the aging pathway (Figure 2) will shed lights on the aging mechanisms and likely provides new routes for aging studies.

Although the ROS theory of aging was proposed by Denham Harman in 1950s without any knowledge about the contributing molecules, recent studies at the molecular level revealed a tight association between ROS and aging genes. Deletion of the insulin-like growth factor receptordaf-2 extends the life span of C. elegans by two-fold. This enhanced longevity depends on a forkhead transcription factor (FoxO) daf-16. Adult worms with an increased daf-16 activity live longer than control worms and appear to possess augmented stress resistance, including resistance to ROS. Indeed, deletion of FoxO transcription factors (FoxO1, FoxO3, FoxO4) decreases the expression of ROS-removing enzymes such as superoxide dismutase 2 and accumulates ROS. Sir2p, a well-established longevity protein, is an NAD+-dependent histone deacetylase. A mammalian homolog of Sir2p, SIRT1, deacetylates the acetylated p53. Active p53 arrests cell cycle and induces apoptosis after detecting DNA damage. Acetylation alters p53’s activity depending on the acetylation site. Jung and colleagues found that ROS is increased during p53-induced senescence. Conversely, treating cells with hydrogen peroxide (H2O2) decreases the level of NAD+, decreases the activity of SIRT1, and lead to the accumulation of acetylated p53.

While longevity genetic manipulations decrease the ROS level, short-lived mutants such as activated Ras2 increased the intracellular ROS level. Therefore, mathematically modeling the intracellular homeostasis of NAD/NADH, NADP/NADPH, GSH/GSSG, ROS, and calcium will highlight the contribution of involved proteins.

Feedback control models

The mitochondrial aging theory suggests that an accumulation of defective mitochondria is a major contributor to the aging process. The mathematical model for the dynamics of the mitochondrial population was developed by Kowald and colleagues. The model divides the mitochondrial population into two major classes, intact mitochondria with no damage to their DNA and defective ones with some form of mtDNA damage. Radicals attack mitochondrial membranes and DNA, converting intact mitochondria into defective ones. The defective mitochondria further generate more superoxides. This positive feedback loop was modeled by a system of ordinary differential equations. With the model, mitochondrial turnover rate, cell division rate, and frequent fusion scenario were simulated by varying these parameters. While their model simulated the contribution of mitochondria in life span extension, they did not specify which proteins are responsible for the maintenance of the mitochondrial integrity.

Another strong candidate for the senescence factor in S. cerevisiae is the accumulation of extrachromosomal ribosomal DNA circles (ERCs) in old cells. Yeast rDNA is located on chromosome XII and consists of 100–200 copies of a 9.1 kb repeat sequence. ERCs are formed by the excision of one of the repeat sequences of the rDNA, followed by homologous recombination. Gillespie and colleagues established a probability model to simulate the ERCs accumulation. The model assumed that a cell acquires an ERC either through excision from the chromosome or by inheritance from its mother; ERCs replicate during successive generations of individual mother cells; ERCs are distributed asymmetrically between mother and daughter cells at division. Assuming that once a cell reaches 1000 ERCs the cell ceases to divide, the model is able to simulate survival curves of yeast cells, which fitted in experimental data if the probability of ERC excision increases quadratically with increasing cell generation.
number. This quadratical increase strongly indicates that ERC formation is not governed by a constant probability but by one that increases with the age of the mother. This suggests that some intrinsic aging process, other than ERC accumulation alone, must contribute to the aging of yeast cells.

To explore the intrinsic aging process, we developed a feedback control model for the dynamics of calcium in aging yeast cells. Malfunction of the calcium homeostasis may be a cause of intracellular senescence and aging. Indeed, the calcium hypothesis of brain aging is widely accepted. The transition from a robust control to malfunction of calcium homeostasis may signal or be the cause of aging. The cytosolic calcium concentration is the net results of pump proteins (Mid1p, Cch1p, Yvc1p, etc) that increase the concentration and pump proteins (Pmc1p, Pmr1p, Vcx1p, etc) that decrease the concentration. The functions of these two categories of proteins are coordinated by calmodulin and calcineurin. Yeast cells uptake calcium from the environment via Mid1p, Cch1p, and possibly other unidentified transporters. The rise of cytosolic calcium activates calmodulin which in turn activates the serine/threonine phosphatase calcineurin. The activated calcineurin de-phosphorylates Crz1p and suppresses the activity of Vcx1p. Activated Crz1p enters the nucleus and upregulates the expression of PMR1 and PMC1. Pmr1p pumps calcium ions into the organelle Golgi and possibly endoplasmic reticulum (ER). The calcium in ER and Golgi will be secreted along with the canonical secretory pathways. Pmc1p pumps calcium ions into vacuole, an organelle that stores excess ions and nutrients. While most calcium ions inside vacuoles form polyphosphate salts and are not re-usable, a small fraction of calcium ions can be channeled to the cytosol by Yvc1p. Yvc1p channels calcium to the cytosol and also contributes to the rise of cytosolic calcium concentration. In this intricate system, damage of one protein such as Pmc1p or Vcx1p may not affect cell’s ability to adjust to small variations of calcium burst. On the contrary, a decline of the whole system will ruin the robustness. Based on Cui and Kaandorp’s calcium model for log phase cells, we established a feedback control system to simulate cell cycle-dependent cytosolic calcium oscillations. Our model can reproduce calcium shocks and calcium accumulations observed in experiments. Since our model contains an aging factor, it can be used to quantitatively predict the calcium dynamics in aging cells. Under the assumption that a cell will die when the cytosolic calcium concentration exceeds a threshold, our model can be used to predict the life span (Figure 3). The theoretical and experimental analyses indicate that Pmr1p is a major contributor for calcium homeostasis during aging. Furthermore, our model predicts that the upregulation of Pmr1p by a constitutively highly

Figure 3 Upregulation of Pmr1p extends life span. The feedback control model of calcium dynamics developed by Tang and Liu predicts that the upregulation of Pmr1p by a calcium-independent promoter (Pvac8) can extend life span. The promoter of VAC8 (Pvac8) is likely constitutive, since aged mother cells and their daughter (young) cells did not show obvious different levels of vac8p-GFP signals.
expressed promoter, independent of calcium, can extend life span, as shown in Figure 3.

Conclusion
As stated above, searching for genes involved in aging can be achieved by both approaches of systems biology: data integration (from -omics to model) and mathematical modeling (from kinetics to model). Recently, Petranovic and Nielsen\textsuperscript{63} termed these two approaches as top-down and bottom-up systems biology. Unlike top-down systems biology, mathematical modeling can reveal genes involved in aging and predict how they involve in aging as well. The simulation of the accumulation of ERC molecules indicates that factors other than ERC are more important in the aging process. Studies revealed that disruption of cytosolic calcium homeostasis induces mitochondrial DNA mutations. When we simulate the intracellular calcium, we introduced a factor that damages calmodulin and calcineurin. This factor is most likely reactive oxygen species (ROS). One of our models predicts that Pmr1p is a key calcium pump in longevity. Pmr1p also pumps calcium into the lumen of endoplasmic reticulum, an organelle that generates ROS. These simulations formed testable hypotheses for wet lab experiments. Although most applications of systems biology are conducted in the model organism yeast, the conservation of intracellular functions make it suitable to reconstruct silicon mammalian cells and to conduct in silico manipulations. A combination of the top-down systems biology and the bottom-up systems biology will predict genes involved in aging, guide the experimental design, and greatly relieve labor in systematic identification of longevity genes.

Disclosure
The authors report no conflicts of interest in this work.

Acknowledgments
Research for FT is partially supported by a seed grant from the Office of Research and Sponsored Program of the University of Arkansas at Little Rock and the Kathleen Thomsen Hall Charitable Trust. WL is partially supported by a grant from the University Research Council Fund of the University of Central Arkansas. YT is partially supported by China Scholarship Council (2007–3075).

References


