



SYMPOSIUM

A Network Perspective on Metabolism and Aging

Quinlyn A. Soltow,* Dean P. Jones* and Daniel E. L. Promislow^{1,†}

*Department of Medicine, Emory University, Atlanta, GA 30322, USA; [†]Department of Genetics, University of Georgia, Athens, Georgia, 30602, USA

From the symposium “Metabolism, Life History and Aging” presented at the annual meeting of the Society for Integrative and Comparative Biology, January 3–7, 2010, at Seattle, Washington.

¹E-mail: promislow@uga.edu

Synopsis Aging affects a myriad of genetic, biochemical, and metabolic processes, and efforts to understand the underlying molecular basis of aging are often thwarted by the complexity of the aging process. By taking a systems biology approach, network analysis is well-suited to study the decline in function with age. Network analysis has already been utilized in describing other complex processes such as development, evolution, and robustness. Networks of gene expression and protein–protein interaction have provided valuable insight into the loss of connectivity and network structure throughout lifespan. Here, we advocate the use of metabolic networks to expand the work from genomics and proteomics. As metabolism is the final fingerprint of functionality and has been implicated in multiple theories of aging, metabolomic methods combined with metabolite network analyses should pave the way to investigate how relationships of metabolites change with age and how these interactions affect phenotype and function of the aging individual. The metabolomic network approaches highlighted in this review are fundamental for an understanding of systematic declines and of failure to function with age.

Introduction

Since the 1950s, much work has been dedicated to understanding the aging process and what affects it (Medawar 1946, 1952; Orgel 1963, 1970, 1973; Medvedev 1990). Although oversimplified, aging can be envisioned as a shift from fully functional to failing-to-function. In youth, organisms operate at their peak performance. Over time, life as a whole begins to lose coordination between the parts, and the systems begin to fail. Reductionist approaches have tested these ideas mechanistically through the failure of genetic replication [e.g., telomere shortening (Oeseburg et al. 2010)], mitochondrial dysfunction (Van Remmen and Jones 2009), accumulation of oxidative stress (Jones 2006), uncontrolled inflammation (Chung et al. 2008), and hormone dysregulation (Panowski and Dillin 2009)) as potential theories of aging. However, the system-wide deterioration that we observe as organisms age likely involves a combination of some or all of these mechanisms working together to create the complex phenomenon that we call aging.

To better understand the complexity of aging, researchers have begun to advocate a systems biology approach, in which we would study both causes and consequences of aging at genome-wide, transcriptome-wide, proteome-wide, and metabolome-wide levels (Barabasi and Oltvai 2004; Hood et al. 2004). Based upon the characteristics of life, there are four general ways ‘systems’ could decline in function with age: (1) decline in the capture and utilization of energy and in the efficiency with which it is extracted; (2) decline in structural or metabolic organization; (3) decline in barrier functions (i.e., delineation from the environment); and (4) decline in the fidelity of storage and transfer of information (i.e., fidelity of DNA during reproduction). Proteomics can be used to investigate the first three characteristics while genomics and gene-expression methods can be used to investigate the fourth characteristic (Fig. 1). In this view, genomics cannot give us a complete picture of aging because it accounts for only reproduction and for DNA encoded in the nucleus and perhaps also for

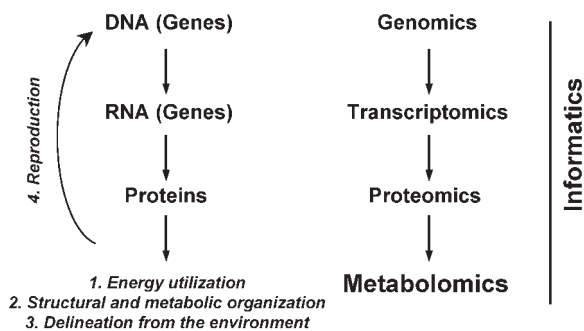


Fig. 1 The central dogma of life studied with “-omics” technologies provide a global perspective on how systems begin to fail with age.

mitochondria. However, the metabolome is the furthest down the line from gene to function and is most characteristic of the entire organismal state and phenotype. Therefore, changes in the metabolome provide the best representation of the characteristic declines in function with age.

Systems biology teaches us not only that phenotypes are influenced by many important genetic and environmental factors, but also that we need to understand the extremely complex interactions between these components. Interactions between the parts, as well as influences from the environment, give rise to new features, such as network behavior (Alm and Arkin 2003), which are absent in the isolated components. As we become more aware of the roles of the network and its components, we can begin to decode the aging process by connecting the molecular components to discrete or multiple signaling pathways. Here, we argue that network analysis of metabolomic data can help to determine how the failure of networks to maintain stability and homeostasis within living systems can lead to aging.

Network approaches

Systems biologists have made great progress in characterizing the structure and function of molecular networks. For example, gene network structure and function can be used to predict formation of patterns during early development (von Dassow et al. 2000; Salazar-Ciudad and Jernvall 2002; Albert and Othmer 2003), and to explain why networks are resistant to damage (Albert et al. 2000; Wagner 2000; Bergman and Siegal 2003).

To use and understand network analysis in research on aging, we first need a brief explanation of the components and structure of networks.

A network is a collection of units potentially interacting as a system, and each network is comprised of a finite set of nodes and edges; nodes are the individual elements (e.g., genes, proteins, and metabolites) within the network, and the edges are the connections (e.g., correlations and enzymatic reactions) that relate the nodes. Edges can be either directed, whereby one node acts upon another (as in gene regulatory networks), or undirected (as in protein–protein interaction [PPI] networks). A node’s connectivity, k , is the number of edges that connect the focal node to other nodes, and nodes with high connectivity are considered hubs. In most biological networks, k is distributed among nodes according to a power-law function $P(k)=\zeta(-\gamma)k^{-\gamma}$, where $P(k)$ is the probability that a selected node has exactly k connections (degrees) with other nodes (e.g., proteins), γ is the degree exponent, a characteristic value for a given network which determines many properties of the system, and $\zeta(-\gamma)$, Riemann’s zeta function, is a normalizing constant. The smaller the γ value, the more important is the role of hubs in the network (Barabasi and Oltvai 2004). Mark Newman (2003) provides a thorough review of network structure and function. There are numerous software packages that allow one to draw networks and evaluate network structure, including specific network programs (e.g., Pajek, UCINET, etc.), as well as components of general statistical packages (e.g., BioConductor in R).

The shape of the power law distribution tells us much about networks, but it cannot reveal the underlying hierarchical structure that characterizes many networks. For example, many network analyses have used hierarchical clustering to identify co-regulated groups or modules within the larger network. This concept of modularity assumes that the function of the network can be better understood if we partition the network into a collection of modules. Each module is a discrete entity of several components, which act together to perform a unique function (Hartwell et al. 1999; Lauffenburger 2000; Hasty et al. 2001; Holter et al. 2001; Rao and Arkin 2001; Shen-Orr et al. 2002). This notion of modularity is especially important in biological networks. As Bruce Alberts (1998) noted, we can think of a cell as a collection of machines. Since each machine is made up of interacting molecules, we can imagine that each machine forms a discrete module within the larger network acting within a cell.

In a highly cited article on the topic, Girvan and Newman (2002) refer to this aspect of networks as their “community structure”. The community structure of the network (that is, the modules that

underlie the network) can be identified through hierarchical clustering algorithms (Girvan and Newman 2002). More recently, a variety of other approaches have been used, from independent component analysis (e.g., Gong et al. 2007) to iterative machine learning approaches (Segal et al. 2003) to approaches derived from graph theory (Newman's (2007) eigenvector analysis and Stone & Ayroles' (2009) Modulated Modularity Clustering). Implementations for many of these methods are available either as packages in R (R Foundation for Statistical Computing, Vienna, Austria), or as standalone software or websites (e.g., Stone and Ayroles 2009). Importantly, to our knowledge, no systematic review has yet been done to determine relative power, speed, biological relevance, and degree of overlap (or lack thereof) in the modules identified by these different statistical methods.

Researchers have used structural analysis of biological networks to better understand a broad range of traits. Studies have found, for example, that the most highly connected nodes are more likely to be essential for organismal function (Jeong et al. 2001), tend to evolve more slowly (Fraser et al. 2002; Hahn et al. 2004; Wuchty 2004), are less likely to be lost over evolutionary time (Krylov et al. 2003), and have a higher probability of being associated with senescence (Promislow 2004; Ferrarini et al. 2005; Budovsky et al. 2007). Results from both theoretical (Albert and Othmer 2003) and empirical (Promislow 2005) studies have shown that network structure is an important predictor of patterns of gene expression. Thus, network studies can provide a unique perspective on the cross-talk between biological modules, identify the most susceptible nodes for loss in communication, and model how changes in these interactions are associated with phenotypes of aging.

Networks and aging

From the above, one can anticipate that highly connected nodes in biological networks would be important in aging. Indeed, proteins associated with aging in yeast have significantly higher connectivity than expected by chance (Promislow 2004), suggesting that a loss in these links would have much greater effects on the entire network than loss of less connected proteins (Albert et al. 2000).

This review focuses on the way in which the structure and function of an organism's underlying molecular networks can affect its phenotype. In the case of aging, though, we can also reverse the causal arrow, asking instead how aging affects network

structure and function. If aging is the gradual failure not simply of individual biological components but rather of complex networks, studies of the structure and function of networks can provide a useful framework to understand both the causes and consequences of aging. Thus, as we turn to a systems biology approach to the study of aging, we shift our focus from studies of network function to studies of network failure.

Ferrarini et al. (2005) developed a theoretical model of network failure, asking in particular what role highly connected nodes, or hubs, play in maintaining network function. Assuming a power-law distribution of network connectivity, their theoretical model showed that restoration of hubs rescued 35% of the lost function, while restoration of weakly connected nodes led to very little improvement in network function. Interestingly, in an analysis of genes that were both highly connected and associated with aging, they found a disproportionate number associated with metabolism—a topic to which we will turn our focus later in this article. In marked contrast with the results of Ferrarini et al. (2005), Csermely and Soti (2006) suggested that the failure of weak links, leading to increased noise and loss of overall network integrity, may account for senescence.

In the previous section, we discussed the community structure of networks. Network approaches have also been used to identify subcomponents of networks associated with aging. As an example of the use of modularity in aging studies, Xue et al. (2007) looked for genes whose expression levels are significantly correlated across age, either positively or negatively. Using expression data from studies of human brains and of whole fruit flies, they were able to identify large modules of co-regulated, age-associated genes. They were further able to show that these modules tended to be associated with particular biological processes, and that expression levels in some genes (such as those associated with cell proliferation) tended to increase with age, while others (such as those associated with differentiation) decreased with age.

Furthermore, Xue et al. (2007) looked at the effect of dietary restriction on network structure. Dietary restriction is the most widely accepted method of extending lifespan in a range of laboratory organisms, from yeast, worms, and flies to mice and rats. Xue et al. (2007) found that dietary restriction altered the structure of the gene co-expression network in *Drosophila*.

In silico modeling, based on gene arrays or PPIs that have already been established, can also be used

to identify components of a network that are associated with lifespan. For example, to construct a “longevity network” via analysis of PPI, Budovsky et al. (2007) used the human orthologs of “longevity assurance genes” that have been established in model organisms (e.g., *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, and *Drosophila melanogaster*) in combination with a list of longevity-associated genes to determine whether the encoded proteins could be organized into a network. A core longevity network was constructed that contained 153 of the 211 longevity-associated proteins (LAPs) identified. The core network was characterized by the presence of highly interconnecting nodes (i.e., hubs), in which 15 of the 17 hubs were LAPs. Interestingly, 33 proteins “not” associated with longevity (“non-LAPs”) had connections with several LAPs, such that an extended network including these proteins revealed an additional 40 LAPs that did not appear in the core network (Budovsky et al. 2007).

At this point, we see an urgent need to understand how senescence affects the structural and functional interactions that underlie complex networks. A recent study of age-related changes in gene expression in mice (Southworth et al. 2009) points the way forward. Southworth et al.’s study relied on the ‘AGEMAP’ dataset (Zahn et al. 2007), which measured genome-wide levels of gene expression at four different ages in 16 different tissues in C57BL/6 mice. Using these data, Southworth et al. (2009) constructed gene co-regulatory networks in 16-month-old and 24-month-old mice. They found that the integrity of these networks, measured as the number of connections between genes, declined from 16 to 24 months. In particular, a differential co-expression network analysis from 13 different tissues revealed that 24-month-old mice had 26% fewer total edges than did 16-month-old mice, and large interconnected gene groups were lost with age in a non-uniform fashion. Those genes that tended to lose correlations were associated with specific transcriptional factors and tended to be clustered in specific regions of the chromosome. The clusters of genes that lost correlation with age included genes involved in steroid production, memory, and mitochondrial function while DNA-damaged genes became more correlated with other genes in older mice (Southworth et al. 2009).

Overall, the literature reviewed thus far on genome and proteome network modules demonstrates that aging is associated with (1) failure of hubs; (2) failure of weak links; and (3) redistribution of node structure. The challenge that we now face is to determine what kinds of network components, in

terms of both structure and function, are most likely to lose integrity with age.

Metabolic networks and aging

In this section, we will first describe the history of metabolomics and why this particular “-omic” technology has special advantages in the study of aging. Second, we will discuss ways of identifying age-related changes in the structure of any molecular networks (including genomic and proteomic as well as metabolomic networks). Last, we discuss, in particular, the study of senescence in metabolomic networks.

History of metabolomics

Beginning in 1998, the first published study of metabolomic technology emerged as a method to detect hundreds of small molecules, including antioxidants, oxidative stress markers, purines, and indoles, in rat-liver mitochondria (Kristal et al. 1998). Since then, the technology has flourished, increasing the number of publications on the topic from one in 1998 to hundreds by 2003 (Kell 2004), and almost 1000 in 2009 alone. In the early years of metabolomics research, no clear definition of the “metabolome” had yet been claimed. Oliver et al. (1998) made the first attempt by suggesting that the metabolome is a comprehensive measure of low-molecular-weight molecules present in cells in a particular physiological or developmental state. The definition has now evolved into a more generalized term including all small molecules characterizing a biological system (Boccard et al. 2010). In any case, the small molecules detected while measuring the metabolome consist of those generated *in vivo*, including those affected by diet, drugs, pesticides, and other environmental exposures.

Ideally, a metabolic profile would detect and quantify all small molecules present in a biological system at a given time (i.e., a snapshot of cellular metabolism or of components in circulating blood). However, this is not possible with current technology. As discussed below, the techniques for separation and detection currently in use are unable to identify all *in vivo* metabolites due to inadequate extraction, chemical diversity, and the dynamic range of the metabolome (Goodacre et al. 2004). These complications have limited our knowledge with respect to known metabolites, but in just 2 years, the number of metabolites identified in humans has more than tripled, from 2180 to almost 7000 by 2009 (Wishart et al. 2009). This number does not include the additional thousands of metabolites from drugs,

common toxins or pollutants, and food additives. Thus, it is obvious that we have much to learn about the metabolome.

The progress made thus far in our ability to characterize the metabolome has been made possible by advances in the technology platforms in analytical chemistry. The primary analytical platforms include nuclear magnetic resonance (NMR), mass spectrometry (MS), and a number of separation techniques, such as capillary electrophoresis (CE), gas chromatography (GC), and high performance liquid chromatography (HPLC). Chemical separation by GC or HPLC increases the resolution and sensitivity of MS while enabling selectivity of species based on their chemical properties (Goodacre et al. 2004). In addition, the availability of a variety of columns for HPLC has made the detection of comprehensive metabolic profiles more feasible. NMR spectroscopy has been developed as a valuable resource for metabolomics due its reproducibility and its ability to provide structural information. However, NMR has low sensitivity and low throughput [there are exceptions (Soininen et al. 2009)], which limits our ability to detect the maximum number of metabolites in a short period of time. Mass spectrometry is a more promising technique to generate high-throughput metabolomic data because it overcomes the weaknesses of NMR with higher sensitivity in a relatively shorter time required for analysis. For example, the Fourier transform-MS (FTMS) has a high resolving power and sub-parts-per-million mass accuracy that facilitates identification with less dependence upon chromatography. Further, the measurement of accurate mass by the FTMS has proved useful for characterization of unknown metabolites by the unambiguous assignment of elemental formulas (Aharoni et al. 2002). Thus, the performance characteristics of MS, particularly FTMS, are ideal for the complex mixtures encountered in high-throughput-metabolomics applications.

Metabolomics platforms include both “top-down” versus “bottom-up” approaches. In a bottom-up model, the method is targeted to detect specific molecules or pathways in a hypothesis-driven manner. Instrumentation and bioinformatics are not restrictive for bottom-up models since a finite number of metabolites are targeted and less-sensitive instrumentation is needed. On the other hand, a top-down approach is untargeted and more reliant on sensitivity and robustness of the instrument in order to generate extensive metabolic profiles. Complex bioinformatics are then needed to mine the resultant data and to generate hypotheses. Although both approaches have drawbacks, in systems biology,

bottom-up metabolomics is most limiting, as the focus on a few selected metabolites can lead the researcher to ignore critically important interactions with other molecules (Forster et al. 2003). Top-down metabolomic approaches are more useful for drawing general conclusions regarding aging, or any multi-system effect, because they take a global snapshot of the organism’s metabolic state.

The composition of samples selected for metabolomic analysis can also limit and complicate interpretations. Samples of tissues and cells often require rigorous protocols for handling and extraction in order to isolate metabolites; these methods can result in inconsistent and more targeted metabolic profiles. In such a case, one must be careful when drawing conclusions in systems biology studies, since the modeling is not comprehensive and will be biased toward the metabolites that were extracted (Goodacre et al. 2004). On the other hand, biofluids are perhaps the ideal type of sample to be analyzed because they are easily obtained and can be analyzed by MS and NMR with little preparation of the sample (Brindle et al. 2002; Lindon et al. 2003; Nicholson and Wilson 2003). Biofluids, specifically plasma and serum, can provide a clear representation of a global metabolic profile because these fluids have contact with organ systems throughout the body.

Although the identities of the individual components of the metabolome are less well understood than that of the transcriptome and proteome, metabolomics offers distinct advantages as a means of studying the molecular state of an organism. In particular, theoretical (Fell 1996; Kell and Mendes 2000) and experimental (Raamsdonk et al. 2001) studies of metabolic control reveal that changes in the quantities of individual enzymes have significant effects on the concentrations of numerous individual metabolites. In addition, there are countless chemical alterations and regulatory mechanisms that occur downstream of the genome. As a result, the metabolome is effectively an integration not only of cellular activities at a functional level, but also of extrinsic, environmental factors (Fig. 2). Because the metabolome is essentially the “end of the line” in terms of biological processes, changes in the metabolome may provide a better mirror of the consequences of genotypic and phenotypic changes than do changes of the actual transcriptome or proteome. To summarize, the metabolome is the final fingerprint of genetic regulation and enzymatic activity, and the resulting global metabolic profile can best be measured by metabolomics.

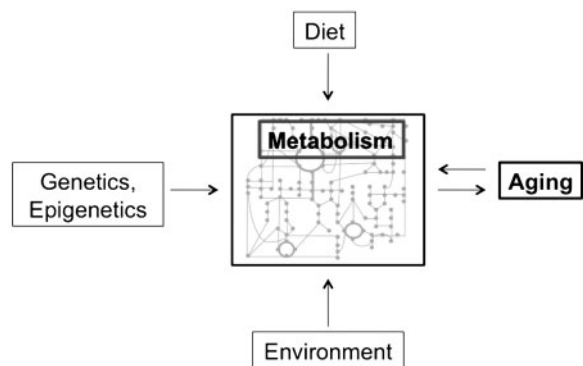


Fig. 2 Metabolic profiles integrate the effects of diet, environment, and genetics for mechanistic studies of aging.

Recent advances in metabolomics have provided insights into certain aspects of aging, and many trends of metabolites during aging are emerging. Using ^1H NMR, urine samples from aging rats, dogs, and humans were analyzed, which showed that resonances associated with creatinine, amino acids, and fatty acids increase with age while glucose/myoinositol (Williams et al. 2005), glycoproteins (Wang et al. 2007), succinate, and other Krebs-cycle intermediates decrease with age (Schnackenberg et al. 2007; Gu et al. 2009). These results demonstrate the utility of NMR to find variability in macromolecules with age, and HPLC–MS studies of aging expand on these results by detecting more than 11,000 ions in a biological sample (Williams et al. 2006). HPLC–MS analysis of urine from aging Sprague Dawley rats not only confirms ^1H NMR results of declining Krebs-cycle intermediates but also identifies increased levels of oxidized antioxidants in older rats (Schnackenberg et al. 2007). HPLC–MS mass chromatograms detect significantly more ions in urine from 4-week-old Wistar-derived rats than from 20-week-old rats (Williams et al. 2005). Ions identified as carnitine, ascorbic acid, and urate increase in ion intensity with age, but the large majority of age-associated ions remain unidentified. Another study of aging using HPLC–MS to perform metabolomic characterization in the serum of 4, 10, 18, and 24-month-old rats found that levels of carnosine, cholesterol, and various fatty acids among other unidentified metabolites were associated with age (Yan et al. 2009). These studies demonstrate that HPLC–MS-based metabolomics can provide a wealth of information (e.g., thousands of metabolites detected), although much is to be learned concerning the identity of the ions. One way to attack the problem of unknown molecules is to construct a network connecting metabolites that change with age. The relationships between the

known and unknown metabolites may assist in elucidation of pathways and in identification of previously unrecognized ions.

Metabolic networks

While most studies of biological networks and aging have focused on protein–protein or gene regulatory interactions, we suggest here that metabolic networks may prove particularly valuable in furthering our understanding of the biological and phenotypic changes that occur as organisms age. Metabolites are the final fingerprint of genetic regulation and enzymatic reactions, and provide a clear picture of phenotype. The potential value of studying metabolic networks is further supported by the fact that the most widely tested intervention to prolong lifespan, dietary restriction, has obvious effects on energy metabolism (Masoro 2001), as demonstrated by a life-long metabolomics study of caloric restriction in dogs (Wang et al. 2007).

There are a variety of ways to construct metabolic networks (e.g., Batagelj and Mrvar 1998; Stone and Ayroles 2009). Ravasz et al. (2002) reconstructed a network by using the information about the *E. coli* enzymes for all metabolic reactions available in the ERGO database. This network includes a few highly connected molecules, such as ATP, that tell us little about the biological functions of the specific nodes and pathways with which they interact. Thus, Ravasz et al. (2002) developed an algorithm that would remove these ‘redundant’ edges, and also simplified the connections between substrates.

Ravasz et al. (2002) suggested that in reconstructing these networks, we need to recognize not only the global patterns of the network (i.e., the number of nodes to which each node is connected), but also the underlying modular substructure. One way that we can measure overall modularity in a network is by estimating a network’s clustering coefficient, $C(k)$. The clustering coefficient measures the degree to which groups of nodes in a network tend to cluster together to form tightly connected neighborhoods. For example, in a network with a high clustering coefficient, if nodes A and B are both connected to node C, there is a high probability that nodes A and B are connected with one another. In a study of metabolic networks in 43 different organisms, Ravasz et al. (2002) found that the clustering coefficient was about an order of magnitude larger than one would expect in a network in which the number of connections is distributed according to a power law, but at random with respect to modularity. This

result suggests that metabolic networks are characterized by high intrinsic modularity.

After construction of a network, we can glean important information from analysis of its structure and function. Like gene and protein networks, the most connected modules in metabolic networks play a critical role in overall network size (Jeong et al. 2000; Wagner and Fell 2001) and structure (Wagner and Fell 2001). The diameter of metabolic networks, defined as the average shortest distance between any two nodes in the network, does not change across organisms even with the increasing number of substrates that occur with increasing organismal complexity (Jeong et al. 2000).

In the following section, we describe our current work on the analysis of steady-state metabolic networks. While studies of complex flux-sum (Chung and Lee 2009) or kinetic processes (Coquin et al. 2008) are clearly of interest, a recent study of an *in silico* metabolic network model by Çakir et al. (2009) found that metabolic variation observed at steady-state, with no perturbation, can reveal connectivity of the underlying metabolic network. Our own work, discussed below, shows how one might use such steady-state comparisons across individuals to assess the effect of aging on metabolomic network structure.

Metabolic network and aging in the common marmoset

We are interested in top-down metabolic networks and how they change with age. Until now, most research on metabolism and aging has been focused on the relationship of dietary restriction to longevity and on mechanisms of age-related metabolic disease. Studies on metabolic mechanisms of aging have focused on a variety of physiological traits, including mitochondrial function, oxidative stress/redox balance and its downstream effects on lipids, proteins, and DNA; signaling in the insulin and insulin-like growth factor (IIS) and TOR pathways; glucocorticoid signaling; and inflammation. However, the overall process of aging likely involves all of these components, so a systems approach using metabolic networking can provide a more comprehensive insight into the interactions between these components.

As we have outlined above, a network approach should help us understand key metabolomic differences in longevity among species, and among individuals within species. In particular, our goal is to identify specific components of the underlying metabolomic network that are most likely to fail with age.

Previous progress in metabolic network reconstruction depended critically on the functional annotation of the genes that encode proteins with unknown functions and focused mostly on specific pathways such as fatty acid synthesis or aspartate metabolism (Chen et al. 2009; Curien et al. 2009). However, with the recent advances in metabolomics technologies, we can begin to construct metabolomic networks without relying on data extrapolated from gene-expression assays or on *in silico* modeling. Steady-state metabolomics can provide the necessary information to elucidate age-related changes in network connectivity and structure.

We have used a top-down metabolic profiling technique to compare the steady-state metabolic network of a set ($n=60$) of young (<6 years of age) and old (>6 years of age) common marmosets (*Callithrix jacchus*). The common marmoset has recently emerged as the premiere primate model for studies of aging. Because of their short lifespan (6–15 years) and small size (~400 g), marmosets are easy to handle, require little space, and make long-term studies of aging more feasible. Using our HPLC-FTMS-based metabolomics platform (Johnson et al. 2008), partnered with a metabolite peak extraction method (Yu et al. 2009), we can identify over 2500 metabolites in a single drop of blood. These metabolites cover the entire profile of the organism, including “dead-end” or inactive ones, as well as metabolites that have yet to be identified. In our preliminary analyses of metabolomic networks in marmosets varying between 2 and 13 years of age, we find that connectivity of metabolites decreases with age (Fig. 3), and that the abundance of numerous metabolites significantly changes with age (Fig. 4). By using these types of analyses (i.e., global metabolic profiling combined with network analysis), well-designed models of aging may reveal biomarkers that correlate with age-related phenotype and disease.

To carry out metabolic network analyses of aging, we can use both cross-sectional studies and longitudinal approaches. Cross-sectional approaches can be used to determine how metabolomic networks vary in response to differences in environment, genotype, sex, age, behavior, and so forth. Second, with respect to aging, longitudinal studies, where we track single individuals over time, are particularly valuable, as we can fully control for differences among individuals in genotype or rearing environment. This approach is also a powerful way to monitor effects of experimental manipulations, such as genetic knockdowns or dietary changes, on the metabolome.

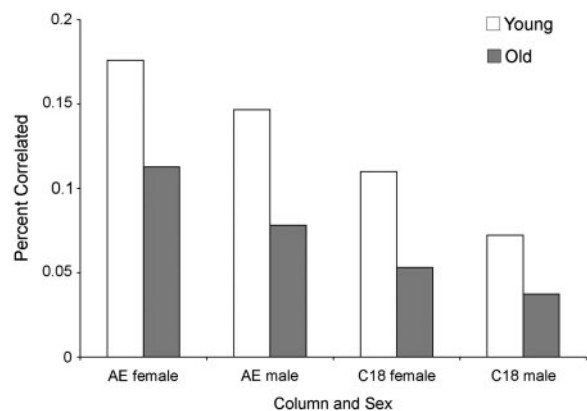


Fig. 3 Percent of correlated metabolites as a function of extraction column and sex. We calculated correlation coefficients between all possible pairs of metabolites among individuals for both the AE column ($n=765$ metabolites) and the C18 column ($n=714$ metabolites), for groups of young or old marmosets separately. In each case, there was a total of $n(n-1)$ possible pairwise correlations. We determined the percent of all correlations that was highly significant ($P < 10^{-7}$). For both columns, and for both sexes, we see more significant correlations among young individuals than among old individuals. In all four comparisons, the difference between old and young groups is highly significant ($P < 10^{-13}$).

Although this review highlights the progress that has been made thus far in using metabolomics data to generate biological networks, there are still key challenges in the field. The field of metabolomics has yet to adopt a set of standardized methods, which involve sampling, sample separation, MS-platform selection, and data extraction. At this point, perhaps the greatest challenge is that curated metabolomic databases are incomplete. Using LC-FTMS, we can obtain estimates of the quantities of thousands of metabolites, each with a unique mass:charge (m/z) and retention time, from a sample of just a few microliters. But without a substantial effort, we would not have detailed understanding of the chemical identities of all of these molecules. Furthermore, methods for quantification of metabolites have yet to be verified.

Relative to work on gene-regulatory and PPI networks, relatively few researchers are working on the construction of metabolomic networks, and on tying together metabolomic network structure and function with physiological traits and environmental perturbations. Fortunately, this gap is likely to be filled in the coming years through research to define the human exposome (Wild 2005). In a recent National Academy of Sciences Workshop sponsored by the National Institute of Environment and Health Sciences, the possibility of creating a “Human

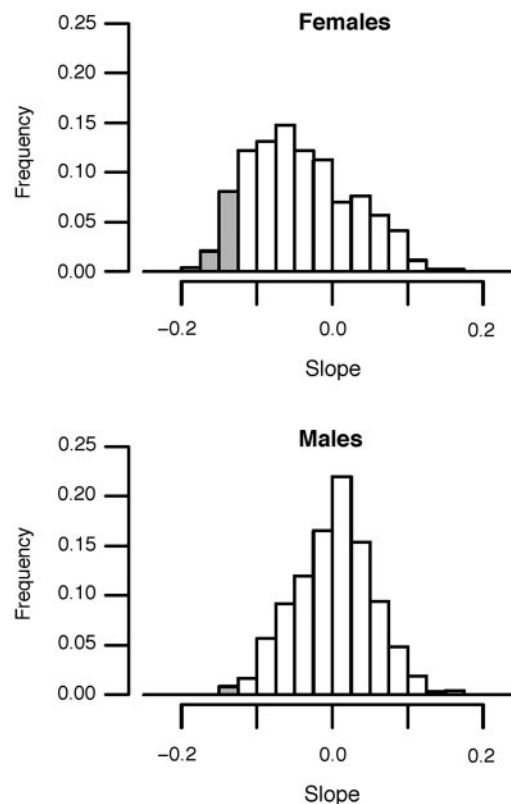


Fig. 4 Frequency distribution of correlation coefficients between quantity of metabolites and age of the marmoset, using extractions from an anion exchange column (results from a C18 column were comparable). Grey bars indicate correlation coefficients that were significantly different from zero ($P < 0.01$). Note that for females, a disproportionate number of metabolites decline with age (t-test of average slope across all metabolites, $P < 10^{-7}$), while on average, males are not significantly different from zero.

Exposome Project” was discussed. This projected could mimic the Human Genome Project to determine how environmental exposures—including diet, physical activity, stress, and drug use—contribute to human phenotype, aging, and disease. The field of metabolomics and network analysis are likely to play a key role in such a project.

We are still in the very early stages of applying metabolomic network analysis to the study of aging. While we know little, the potential for conceptual advances is substantial. We have described a study of aging in primates; we are likely to also reap significant benefits by applying these approaches to short-lived model organisms, such as fruit flies and nematode worms—the stalwarts of research on the biology of aging. Using model systems, we have identified a large number of genes that can extend longevity (reviewed in Kennedy 2008; Shimokawa et al. 2008; Ni and Lee 2010). In humans, however, our current best bet for extending both the quality and quantity of our lives is diet and exercise. As a recent

study demonstrates (Lewis et al. 2010), metabolomics may help us to understand just why such environmental interventions work.

The strength of connections between genes, proteins, or metabolites is an important system-wide property of a network that can be used to compare two or more states. Here, we compared young and old marmosets, but this approach can be used to compare many other states such as healthy versus diseased individuals, or wild-type versus mutant strains. Differential network analysis could potentially be used to compare network structure among species. In a study of wing polymorphism in ants, Abouheif and Wray (2002) were able to show that the loss of function in wing development was due to the failure of different network components in different species. As organisms age, do networks fail in similar ways from species to species, and from genotype to genotype, or is each individual unique? With apologies to Tolstoy, is it the case that all young, healthy individuals are alike, but every senescent individual is senescent in its own way? Network analysis should help us to answer this question.

Acknowledgments

Helpful comments on an earlier version of this manuscript were provided by members of the Promislow laboratory. The authors also acknowledge the generous support of the Division of Comparative Physiology and Biochemistry in the Society for Integrative and Comparative Biology for funding the symposium in which this work was first presented.

Funding

The National Institutes of Health (P01-ES016731 and T32-ES012870 to D.J., in part).

References

- Abouheif E, Wray GA. 2002. Evolution of the gene network underlying wing polyphenism in ants. *Science* 297:249–52.
- Aharoni A, Ric de Vos CH, Verhoeven HA, Maliepaard CA, Kruppa G, Bino R, Goodenowe DB. 2002. Nontargeted metabolome analysis by use of Fourier Transform Ion Cyclotron Mass Spectrometry. *OMICS* 6:217–34.
- Albert R, Jeong H, Barabasi AL. 2000. Error and attack tolerance of complex networks. *Nature* 406:378–82.
- Albert R, Othmer HG. 2003. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *J Theor Biol* 223:1–18.
- Alberts B. 1998. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* 92:291–4.
- Alm E, Arkin AP. 2003. Biological networks. *Curr Opin Struct Biol* 13:193–202.
- Barabasi AL, Oltvai ZN. 2004. Network biology: understanding the cell's functional organization. *Nat Rev Genet* 5:101–13.
- Batagelj V, Mrvar A. 1998. Pajek—Program for large network analysis. *Connections* 21:47–57.
- Bergman A, Siegal ML. 2003. Evolutionary capacitance as a general feature of complex gene networks. *Nature* 424:549–52.
- Boccard J, Veuthey JL, Rudaz S. 2010. Knowledge discovery in metabolomics: an overview of MS data handling. *J Sep Sci* 33:290–304.
- Brindle JT, et al. 2002. Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using ¹H-NMR-based metabolomics. *Nat Med* 8:1439–44.
- Budovsky A, Abramovich A, Cohen R, Chalifa-Caspi V, Fraifeld V. 2007. Longevity network: construction and implications. *Mech Ageing Dev* 128:117–24.
- Çakir T, Hendriks MM, Westerhuis JA, Smilde AK. 2009. Metabolic network discovery through reverse engineering of metabolome data. *Metabolomics* 5:318–29.
- Chen Y, et al. 2009. RRLC-MS/MS-based metabolomics combined with in-depth analysis of metabolic correlation network: finding potential biomarkers for breast cancer. *Analyst* 134:2003–11.
- Chung BK, Lee DY. 2009. Flux-sum analysis: a metabolite-centric approach for understanding the metabolic network. *BMC Syst Biol* 3:117.
- Chung JH, Seo AY, Chung SW, Kim MK, Leeuwenburgh C, Yu BP, Chung HY. 2008. Molecular mechanism of PPAR in the regulation of age-related inflammation. *Ageing Res Rev* 7:126–36.
- Coquin L, Feala JD, McCulloch AD, Paternostro G. 2008. Metabolomic and flux-balance analysis of age-related decline of hypoxia tolerance in *Drosophila* muscle tissue. *Mol Syst Biol* 4:233.
- Csermely P, Soti C. 2006. Cellular networks and the aging process. *Arch Physiol Biochem* 112:60–4.
- Curien G, Bastien O, Robert-Genthon M, Cornish-Bowden A, Cardenas ML, Dumas R. 2009. Understanding the regulation of aspartate metabolism using a model based on measured kinetic parameters. *Mol Syst Biol* 5:271.
- Fell DA. 1996. *Understanding the Control of Metabolism*. London: Portland Press.
- Ferrari L, Bertelli L, Feala J, McCulloch AD, Paternostro G. 2005. A more efficient search strategy for aging genes based on connectivity. *Bioinformatics* 21:338–48.
- Forster J, Famili I, Palsson BO, Nielsen J. 2003. Large-scale evaluation of in silico gene deletions in *Saccharomyces cerevisiae*. *OMICS* 7:193–202.
- Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C, Feldman MW. 2002. Evolutionary rate in the protein interaction network. *Science* 296:750–2.

- Girvan M, Newman ME. 2002. Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99:7821–6.
- Gong T, Xuan J, Wang C, Li H, Hoffman E, Clarke R, Wang Y. 2007. Gene module identification from microarray data using nonnegative independent component analysis. *Gene Regul Syst Bio* 1:349–63.
- Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB. 2004. Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol* 22:245–52.
- Gu H, Pan Z, Xi B, Hainline BE, Shanaiah N, Asiago V, Gowda GA, Raftery D. 2009. 1H NMR metabolomics study of age profiling in children. *NMR Biomed* 22:826–33.
- Hahn MW, Conant GC, Wagner A. 2004. Molecular evolution in large genetic networks: does connectivity equal constraint? *J Mol Evol* 58:203–11.
- Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999. From molecular to modular cell biology. *Nature* 402:C47–52.
- Hasty J, McMillen D, Isaacs F, Collins JJ. 2001. Computational studies of gene regulatory networks: in numero molecular biology. *Nat Rev Genet* 2:268–79.
- Holter NS, Maritan A, Cieplak M, Fedoroff NV, Banavar JR. 2001. Dynamic modeling of gene expression data. *Proc Natl Acad Sci USA* 98:1693–8.
- Hood L, Heath JR, Phelps ME, Lin B. 2004. Systems biology and new technologies enable predictive and preventative medicine. *Science* 306:640–3.
- Jeong H, Mason SP, Barabasi AL, Oltvai ZN. 2001. Lethality and centrality in protein networks. *Nature* 411:41–42.
- Jeong H, Tombor B, Albert R, Oltvai ZN, Barabasi AL. 2000. The large-scale organization of metabolic networks. *Nature* 407:651–4.
- Johnson JM, Strobel FH, Reed M, Pohl J, Jones DP. 2008. A rapid LC-FTMS method for the analysis of cysteine, cystine and cysteine/cystine steady-state redox potential in human plasma. *Clin Chim Acta* 396:43–8.
- Jones DP. 2006. Extracellular redox state: refining the definition of oxidative stress in aging. *Rejuvenation Res* 9:169–81.
- Kell DB. 2004. Metabolomics and systems biology: making sense of the soup. *Curr Opin Microbiol* 7:296–307.
- Kell DB, Mendes P. 2000. Snapshots of systems: metabolic control analysis and biotechnology in the post-genomic era. In: Cornish-Bowden A, Cardenas ML, editors. *Technological and medical implications of metabolic control analysis*. Dordrecht (The Netherlands): Kluwer Academic Publishers. p. 3–25.
- Kennedy BK. 2008. The genetics of ageing: insight from genome-wide approaches in invertebrate model organisms. *J Intern Med* 263:142–52.
- Kristal BS, Vigneau-Callahan KE, Matson WR. 1998. Simultaneous analysis of the majority of low-molecular-weight, redox-active compounds from mitochondria. *Anal Biochem* 263:18–25.
- Krylov DM, Wolf YI, Rogozin IB, Koonin EV. 2003. Gene loss, protein sequence divergence, gene dispensability, expression level, and interactivity are correlated in eukaryotic evolution. *Genome Res* 13:2229–35.
- Lauffenburger DA. 2000. Cell signaling pathways as control modules: complexity for simplicity? *Proc Natl Acad Sci USA* 97:5031–3.
- Lewis GD, et al. 2010. Metabolic signatures of exercise in human plasma. *Sci Transl Med* 2:33–37.
- Lindon JC, Holmes E, Nicholson JK. 2003. So what's the deal with metabolomics? *Anal Chem* 75:384A–391A.
- Masoro EJ. 2001. Dietary restriction: current status. *Aging (Milano)* 13:261–2.
- Medawar PB. 1946. Old age and natural death. *Modern Quart* 2:30–49.
- Medawar PB. 1952. *An unsolved problem in biology*. London: HK Lewis.
- Medvedev ZA. 1990. An attempt at a rational classification of theories of ageing. *Biol Rev Camb Philos Soc* 65:375–98.
- Newman MEJ. 2003. The structure and function of complex networks. *SIAM Rev* 45:167–256.
- Newman MEJ. 2007. Finding community structure in networks using the eigenvectors of matrices. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 73:036104.
- Ni Z, Lee SS. 2010. RNAi screens to identify components of gene networks that modulate aging in *Caenorhabditis elegans*. *Brief Funct Genomics* 9:53–64.
- Nicholson JK, Wilson ID. 2003. Opinion: understanding 'global' systems biology: metabolomics and the continuum of metabolism. *Nat Rev Drug Discov* 2:668–76.
- Oeseburg H, de Boer RA, van Gilst WH, van der Harst P. 2010. Telomere biology in healthy aging and disease. *Pflugers Arch* 459:259–68.
- Oliver SG, Winson MK, Kell DB, Baganz F. 1998. Systematic functional analysis of the yeast genome. *Trends Biotechnol* 16:373–8.
- Orgel LE. 1963. The maintenance of the accuracy of protein synthesis and its relevance to ageing. *Proc Natl Acad Sci USA* 49:517–21.
- Orgel LE. 1970. The maintenance of the accuracy of protein synthesis and its relevance to ageing: a correction. *Proc Natl Acad Sci USA* 67:1476.
- Orgel LE. 1973. Ageing of clones of mammalian cells. *Nature* 243:441–5.
- Panowski SH, Dillin A. 2009. Signals of youth: endocrine regulation of aging in *Caenorhabditis elegans*. *Trends Endocrinol Metab* 20:259–64.
- Promislow D. 2005. A regulatory network analysis of phenotypic plasticity in yeast. *Am Nat* 165:515–23.
- Promislow DE. 2004. Protein networks, pleiotropy and the evolution of senescence. *Proc Biol Sci* 271:1225–34.
- Raamsdonk LM, et al. 2001. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol* 19:45–50.
- Rao CV, Arkin AP. 2001. Control motifs for intracellular regulatory networks. *Annu Rev Biomed Eng* 3:391–419.

- Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabasi AL. 2002. Hierarchical organization of modularity in metabolic networks. *Science* 297:1551–5.
- Salazar-Ciudad I, Jernvall J. 2002. A gene network model accounting for development and evolution of mammalian teeth. *Proc Natl Acad Sci USA* 99:8116–20.
- Schnackenberg LK, Sun J, Espandiari P, Holland RD, Hanig J, Beger RD. 2007. Metabonomics evaluations of age-related changes in the urinary compositions of male Sprague Dawley rats and effects of data normalization methods on statistical and quantitative analysis. *BMC Bioinformatics* 8:S3.
- Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, Friedman N. 2003. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat Genet* 34:166–76.
- Shen-Orr SS, Milo R, Mangan S, Alon U. 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat Genet* 31:64–8.
- Shimokawa I, Chiba T, Yamaza H, Komatsu T. 2008. Longevity genes: insights from calorie restriction and genetic longevity models. *Mol Cells* 26:427–35.
- Soininen P, et al. 2009. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. *Analyst* 134:1781–5.
- Southworth LK, Owen AB, Kim SK. 2009. Aging mice show a decreasing correlation of gene expression within genetic modules. *PLoS Genet* 5:e1000776.
- Stone EA, Ayroles JF. 2009. Modulated modularity clustering as an exploratory tool for functional genomic inference. *PLoS Genet* 5:e1000479.
- Van Remmen H, Jones DP. 2009. Current thoughts on the role of mitochondria and free radicals in the biology of aging. *J Gerontol A Biol Sci Med Sci* 64:171–4.
- von Dassow G, Meir E, Munro EM, Odell GM. 2000. The segment polarity network is a robust developmental module. *Nature* 406:188–92.
- Wagner A. 2000. Robustness against mutations in genetic networks of yeast. *Nat Genet* 24:355–61.
- Wagner A, Fell DA. 2001. The small world inside large metabolic networks. *Proc Biol Sci* 268:1803–10.
- Wang Y, Lawler D, Larson B, Ramadan Z, Kochhar S, Holmes E, Nicholson JK. 2007. Metabonomic investigations of aging and caloric restriction in a life-long dog study. *J Proteome Res* 6:1846–54.
- Wild CP. 2005. Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev* 14:1847–50.
- Williams RE, Lenz EM, Lowden JS, Rantalainen M, Wilson ID. 2005. The metabonomics of aging and development in the rat: an investigation into the effect of age on the profile of endogenous metabolites in the urine of male rats using ¹H NMR and HPLC-TOF MS. *Mol Biosyst* 1:166–75.
- Williams R, Lenz EM, Wilson AJ, Granger J, Wilson ID, Major H, Stumpf C, Plumb R. 2006. A multi-analytical platform approach to the metabonomic analysis of plasma from normal and Zucker (fa/fa) obese rats. *Mol Biosyst* 2:174–83.
- Wishart DS, et al. 2009. HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res* 37(Database issue):D603–10.
- Wuchty S. 2004. Evolution and topology in the yeast protein interaction network. *Genome Res* 14:1310–14.
- Xue H, et al. 2007. A modular network model of aging. *Mol Syst Biol* 3:147.
- Yan S, Wu B, Lin Z, Jin H, Huang J, Yang Y, Zhang X, Shen Z, Zhang W. 2009. Metabonomic characterization of aging and investigation on the anti-aging effects of total flavones of *Epimedium*. *Mol Biosyst* 5:1204–13.
- Yu T, Park Y, Johnson JM, Jones DP. 2009. apLCMS—adaptive processing of high-resolution LC/MS data. *Bioinformatics* 25:1930–36.
- Zahn JM, et al. 2007. AGEMAP: a gene expression database for aging in mice. *PLoS Genet* 3:e201.