CHAPTER

# 6

# Integrative genomics of aging

*João Pedro de Magalhães*[1], *Cyril Lagger*[1] *and Robi Tacutu*[2]

[1]Integrative Genomics of Ageing Group, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, United Kingdom [2]Systems Biology of Aging Group, Department of Bioinformatics and Structural Biochemistry, Institute of Biochemistry of the Romanian Academy, Bucharest, Romania

OUTLINE



## Introduction

The sequencing of genomes has revolutionized biological and biomedical research. Thanks to various technologies and approaches that take advantage of genome sequence knowledge, researchers can now focus on whole biological systems rather than being limited to studying isolated parts. Because most biological processes are complex in the sense that they involve the interplay of multiple genes and proteins with each other and with the environment, surveying systems as a whole is imperative to fully comprehending them, and more accurately pinpointing how to intervene in them. Recent breakthroughs in developing cheaper and quicker sequencing technologies have given further power to our capacity to survey biological systems in a holistic way with multiple applications in aging research (reviewed in de Magalhaes, Finch, & Janssens, 2010). In addition to genomics, other omics approaches like transcriptomics, proteomics, and epigenomics have allowed for a systematic profiling of biological processes and disease states.

Aging is widely acknowledged as a complex process involving changes at various biological levels, interactions between them, and feedback regulatory circuits. The underlying mechanistic causes of aging remain a subject of debate, and it is likely that multiple degenerative processes are involved, including organ-specific processes but also interacting cell- and organ-level communications (Cevenini et al., 2010; de Magalhaes, 2011; Lopez-Otin, Blasco, Partridge, Serrano, & Kroemer, 2013). While there are simple triggers to complex biological processes, such as telomere shortening triggering replicative senescence in human fibroblasts (de Magalhaes, 2004), most researchers would agree that organismal aging involves multiple processes and possibly the interplay between various causal mechanisms. Likewise, hundreds of genes have been associated with aging in model organisms (Tacutu et al., 2018) and yet the pathways involved are complex and often interact in nonlinear ways (de Magalhaes, Wuttke, Wood, Plank, & Vora, 2012). One hypothesis is that aging and longevity cannot be fully understood by studying individual components

*Handbook of the Biology of Aging.*
DOI: https://doi.org/10.1016/B978-0-12-815962-0.00006-8

and processes (Cevenini et al., 2010). To understand aging we must then account for the intrinsic complexity of biological systems.

Our goal in this chapter is to review potential large-scale technologies in the context of aging and longevity research and how data can be analyzed and integrated to advance our understanding of these complex processes. We first review the major technologies available for researchers to survey biological systems in a systematic fashion and their applications to advance the biology and genetics of aging, debate issues in data analysis and statistics, and discuss data integration between different sources, as this is one of the major challenges of the post-genome era, and also one of the most promising. Various sources of data and approaches are discussed in this context.

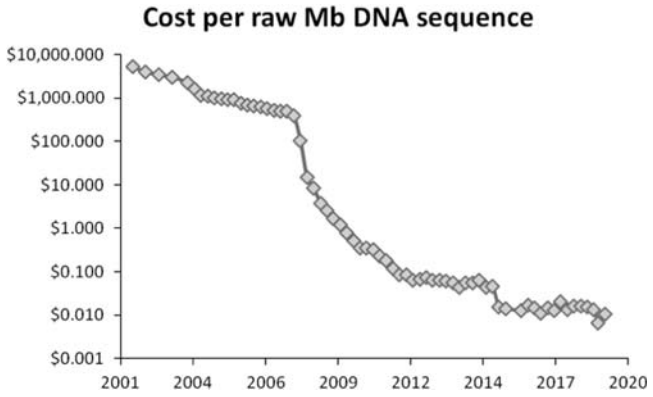


FIGURE 6.1 Exponential growth in sequencing capacity as reflected in the dropping costs of sequencing from 2001 to 2019. *Source: NHGRI (https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost).*

## Post-genome technologies and biogerontology

There are many open questions in biogerontology, but arguably most researchers focus on two key questions (de Magalhaes & Toussaint, 2004b): (1) what are the genetic determinants of aging, both in terms of longevity differences between individuals and species differences in aging? and (2) which changes occur across the lifetime to increase vulnerability, for example in a person from age 30 to age 70 to increase the chance of dying by roughly 30-fold? Post-genome technologies may help us answer both these questions.

### Genome-wide approaches and the genetics of aging and longevity

Understanding human phenotypic variation in aging and longevity has been a long-term research goal. Studies in twins have shown that longevity in humans has a genetic component, and the heritability of longevity has been estimated at ~25% (Christensen, Johnson, & Vaupel, 2006). If we could identify genetic variants associated with exceptional human longevity, these would likely be suitable for drug discovery (de Magalhaes et al., 2012). In 1994, *APOE* was associated with longevity in a French population (Schachter et al., 1994). The sequencing of the human genome in 2001 allowed for much more powerful whole-genome genotyping platforms capable of surveying hundreds of thousands of genetic variants in a cost-effective way (de Magalhaes, 2009). In spite of these recent technological advances, the genetics of human longevity remains largely misunderstood. Several genome-wide association studies (GWAS) have been performed with hundreds of individuals, with largely disappointing results. For example, one landmark study involving several European populations with a total of over 2000 nonagenarian sibling pairs identified only *APOE* as associated with longevity (Beekman et al., 2013); and although *APOE* has been consistently associated with longevity, it only modestly explains the heritability of longevity. GWAS focused on complex diseases and processes have been on many occasions equally disappointing to date, suggesting that common genetic variants have a modest contribution to longevity and complex diseases (de Magalhaes & Wang, 2019; Manolio et al., 2009).

The dropping costs of DNA sequencing (Fig. 6.1) mean that sequencing a human genome is rapidly becoming affordable. Therefore, researchers are moving from genotyping platforms based on known genetic variants to genome sequencing of thousands of individuals. It is possible that this will reveal rare variants with strong effects on longevity, as has been predicted to be the case for complex diseases (Manolio et al., 2009). Nonetheless, considering that only *APOE* has been associated with confidence with longevity, our understanding of the genetics of longevity lags behind our understanding of the genetics of complex age-related diseases, in itself made difficult by numerous factors like multiple genes with small effects. Intrinsic difficulties in longevity studies (e.g., lack of appropriate controls) or because longevity is a more complex trait may explain why our understanding of the heritability of longevity is still poor (de Magalhaes, 2014b).

An even greater source of variation in aging and longevity than that observed between humans is observed across species. We know that mice, for example, age 25–30 times faster than humans, even under the best environmental conditions (Finch, 1990). Even when compared to chimpanzees, our closest living relative whose genome is about 95% similar to our own, aging is significantly retarded in humans (de Magalhaes, 2006).

Therefore there must be a genomic basis for species differences in aging, and again the dropping costs of sequencing have permitted much more affordable de novo sequencing of genomes (de Magalhaes et al., 2010). For example, the sequencing of long-lived species, such as the naked mole-rat and bats (Keane et al., 2014; Kim et al., 2011; Zhang et al., 2013a) can provide candidate genes for selection in long-lived species, and it is interesting to observe that genes involved in DNA damage responses and repair have emerged from such studies (de Magalhaes & Keane, 2013). For example, Keane et al. reported the sequencing and comparative analysis of the genome of the bowhead whale, the longest-lived known mammal. The analysis revealed a number of genes under positive selection and bowhead-specific mutations in genes linked to cancer and aging. The analysis also identified gene changes associated with DNA repair, cell-cycle regulation, cancer, and aging, and potentially relevant changes in genes associated with thermoregulation, sensory perception, dietary adaptations, and immune response (Keane et al., 2015).

In addition to the analysis of genomes from long-lived species, comparative analyses of genomes from species with different lifespans are also beginning to provide further candidate genes for a role in aging. We developed a method to identify candidate genes involved in species differences in aging based on detecting proteins with accelerated evolution in multiple lineages where longevity is increased (Li & de Magalhaes, 2013). Our results revealed ~100 genes and functional groups that are candidate targets of selection when longevity evolves (Li & de Magalhaes, 2013). These include DNA damage response genes and the ubiquitin pathway and thus provide evidence that at least some repair systems were selected for, and arguably optimized, in long-lived species. Other methods aimed at discovering genes associated with longevity have focused on genes showing a stronger conservation in long-lived species (Jobson, Nabholz, & Galtier, 2010), searching for protein residues that are conserved in long-lived species but not in short-lived ones (Semeiks & Grishin, 2012). One recent work employed two methods, one based on residue change that co-occurs with the evolution of longevity and the other based on changes in rates of protein evolution, to identify candidate genes and biological processes modulating longevity across primates (Muntane et al., 2018). Because all these methods are conceptually different from each other, little overlap has been observed in the results. Nonetheless, it seems that genetic alterations contributing to the evolution of longevity in mammals have common patterns (or signatures) that are detectable using cross-species genome comparisons, though much work remains in order to improve the signal-to-noise ratio of these methods. One caveat of these studies is the lack of experimental validation, and thus all of these genes must be seen as candidates. Given the declining costs of sequencing we can expect many more such studies in the near future.

Aging is a particularly difficult process to unravel because it is much harder to study in humans than most other processes and diseases. Observational studies have been conducted but are extremely time-consuming, and clinical trials for longevity itself are challenging, even though trials for aging markers are possible (de Magalhaes, Stevens, & Thornton, 2017; de Magalhaes et al., 2012). Therefore most biogerontologists rely on model systems: human cells; unicellular organisms such as the yeast *Saccharomyces cerevisiae*; the roundworm *Caenorhabditis elegans*; the fruit fly *Drosophila melanogaster*; and rodents, in particular mice (*Mus musculus*) and rats (*Rattus norvegicus*). The small size and short life cycles of these organisms—even mice do not commonly live more than 4 years—make them inexpensive subjects for aging studies, and the ability to genetically manipulate them gives researchers ample opportunities to test their theories and unravel molecular and genetic mechanisms of aging.

The aforementioned traditional biomedical model organisms are widely used in other fields and not surprisingly a variety of tools are available to study them, and recently many of these powerful tools take advantage of omics approaches. While the genetics of aging was initially unraveled using traditional genetic approaches (reviewed in Johnson, 2002), large-scale forward genetic screening approaches now allow for hundreds of genes to be tested simultaneously for phenotypes of interest, including longevity and age-related traits. Genome-wide screens for longevity have been performed (McCormick & Kennedy, 2012), in particular in worms (Hamilton et al., 2005; Hansen, Hsu, Dillin, & Kenyon, 2005; Samuelson, Carr, & Ruvkun, 2007). Hundreds of genes have been associated with life extension in this way, although the overlap between these studies has been smaller than expected. Moreover, this type of aging studies will probably become even more accessible, with the advent of specialized, automated screening devices that monitor lifespan. For example, the *Caenorhabditis elegans* Lifespan Machine was the first automated system dedicated to lifespan assays based on video tracking of worms (Stroustrup et al., 2013). Since then, a more advanced system has been created, the WorMotel, a microfabricated device for long-term cultivation and automated longitudinal imaging of large numbers of worms confined to individual wells (Churgin et al., 2017). One important observation from the worm phenotypic screening of genetic interventions is that it seems that the most important pathways that modulate lifespan when disrupted in worms (and possibly in model organisms)

have been identified by now, even though there is still ample room to identify individual components.

Although usually more laborious, screens for genes affecting lifespan have also been performed in yeast, including for replicative lifespan (Smith et al., 2008), chronological lifespan (Fabrizio et al., 2010), and using pooled screen approaches (Matecic et al., 2010). Technical limitations in flies impede screens at the genome-wide level, but lifespan screens have been performed using the P-element modular-misexpression system (Paik et al., 2012) and Gene Search misexpression vector system (Funakoshi et al., 2011). Costs impede large-scale screens in mice, although large-scale knockout mouse repositories like the Knockout Mouse Project (www.komp.org) and the International Mouse Phenotype Consortium (www.mousephenotype.org) can facilitate such studies; one large-scale profiling of mouse mutants for aging-related phenotypes, called the Harwell Aging Screen, has been performed (Potter et al., 2016).

A variety of genome-wide screens have also been performed in vitro, in particular using RNAi-based technologies (Echeverri & Perrimon, 2006; Moffat & Sabatini, 2006; Mohr, Bakal, & Perrimon, 2010). These include screens focused on traits of interest for aging and longevity. For example, screens for cell lifespan have been performed in human fibroblasts revealing that senescent cells activate a self-amplifying secretory network involving CXCR2-binding chemokines (Acosta et al., 2008). A variety of readouts can be employed to assay for specific traits. For instance, screens have been performed for genes modulating resistance to oxidative stress in mammalian cells (Nagaoka-Yasuda, Matsuo, Perkins, Limbaeck-Stokin, & Mayford, 2007; Plank et al., 2013) and antioxidant responses (Liu et al., 2007). The possibilities are immense and provide another large-scale tool for deciphering biological processes.

One of the goals of biogerontology is to develop interventions that postpone degeneration, preserve health, and extend life (de Magalhaes, 2014a). Large-scale drug screening is now widespread in the pharmaceutical industry (Macarron et al., 2011). While life extension is harder and more expensive to assay than targets in high-throughput screening, systematic screens for life-extending compounds are now a distinct possibility. Petrascheck et al. assayed 88,000 chemicals for the ability to extend worm lifespan (Petrascheck, Ye, & Buck, 2007); while the success of this approach was modest (only 115 compounds significantly extended lifespan and only 13 by more than 30%), it provides proof-of-concept for large-scale screens in the context of life-extending drugs. Further investigations into drug-mediated worm longevity, using a similar protocol, even if with a smaller compounds library of known or suspected mammalian targets (many already approved for use in humans), revealed 60 promising drugs, which might provide beneficial effects on aging in mammals (Ye, Linton, Schork, Buck, & Petrascheck, 2014).

## Surveying the aging phenotype on a grand scale

In addition to understanding the genetic basis for phenotypic variation in aging and longevity, it is also crucial to elucidate the changes that contribute to age-related degeneration. Several age-related changes have been described and historically this focused on broad physiological and morphological aspects and the molecular and biochemical changes for which assays existed. Thanks to genome-wide approaches we can now survey the aging phenotype with unprecedented detail (de Magalhaes, 2009; Valdes, Glass, & Spector, 2013). In particular, advances in transcriptomics have allowed researchers to survey the expression levels of all genes in the genome in a single, relatively inexpensive, experiment.

One major breakthrough in transcriptomics was the development of the microarray, which allows for the quantification of all annotated genes simultaneously. Briefly, this led in the past 20 years to a large number of gene expression profiling studies of aging (de Magalhaes, 2009; Glass et al., 2013; Lee, Klopp, Weindruch, & Prolla, 1999; Zahn et al., 2007). In a sense, however, these have been disappointing in that relatively few genes are differentially expressed with age in most tissues and few insights have emerged. As an exception, Zahn et al. observed a degree of coordination in age-related changes in gene expression. In mice, different tissues age in a coordinated fashion so that a given mouse may exhibit rapid aging while another ages slowly across multiple tissues (Zahn et al., 2007). In addition, our 2009 meta-analysis of aging gene expression studies revealed a conserved molecular signature of mammalian aging across organs and species consisting of a clear activation of inflammatory pathways accompanied by a disruption of collagen and mitochondrial genes (de Magalhaes, Curado, & Church, 2009). This molecular signature of aging maps well into established hallmarks of aging (Lopez-Otin et al., 2013). It should be noted, however, that transcriptional changes during aging may represent responses to aging rather than underlying causative mechanisms and thus their interpretation is not straightforward.

The dropping costs of sequencing have also allowed for gene expression profiling approaches that are digital in nature, as opposed to microarrays that are analog. Sequencing the transcriptome, usually referred to as RNA-seq, allows for unprecedented accuracy and power. A number of reviews

have focused on the advantages of RNA-seq as compared to microarrays (de Magalhaes et al., 2010; Mortazavi, Williams, McCue, Schaeffer, & Wold, 2008; Wang, Gerstein, & Snyder, 2009), and it is very clear that RNA-seq has a superior dynamic range and provides more data than microarrays. Our lab performed one of the first RNA-seq profiling experiments in the context of aging, which revealed gene expression changes in the rat brain in various previously unknown genes, including noncoding genes and genes not yet annotated in genome databases (Wood, Craig, Li, Merry, & de Magalhaes, 2013). Although it is exciting that many changes were observed in the so-called "dark matter" transcripts, because most of these are not annotated or have little information, follow-up is complicated; this emphasizes the need to study the new genomic elements that may be phenotypically important. In this context, large-scale efforts, such as ENCODE (www.encodeproject.org), which aims to identify all functional elements in the human genome (Dunham et al., 2012), are crucial to annotate and elucidate the function of all genomic elements. Similarly, the Genotype-Tissue Expression (GTEx) Consortium (gtexportal.org) provides data from multiple human tissues and ages (GTEx Consortium, 2015), and can be used to study transcriptional changes with age in human tissues (Chatsirisupachai, Palmer, Ferreira, & de Magalhaes, 2019; Yang et al., 2015).

A number of studies have also focused on profiling gene expression changes in life-extending interventions or in long-lived strains (de Magalhaes, 2009; Lee et al., 1999; Tyshkovskiy et al., 2019; Wood et al., 2015), as well as in short-lived and/or progeroid animals. For example, a large number of studies have focused on caloric restriction (CR) to identify specific genes and processes associated both with CR and whose age-related change is ameliorated in CR. In contrast to studies of aging, CR studies have revealed substantial gene expression changes, some of which can be associated with specific pathways and processes (Lee et al., 1999; Tsuchiya et al., 2004; Wood et al., 2015). A meta-analysis of gene expression studies of CR revealed a number of conserved processes associated with CR effects like growth hormone signaling, lipid metabolism, immune response, and detoxification pathways (Plank, Wuttke, van Dam, Clarke, & de Magalhaes, 2012). In another study, midlife gene expression profiling of mice of different lifespans due to different dietary conditions revealed a possible contribution of peroxisome to aging, which was then tested experimentally in invertebrates (Zhou et al., 2012). Arguably, gene expression profiling of manipulations of aging has been more successful in providing insights than profiling of aging per se.

Another area that has recently flourished due to technological developments is comparative transcriptomics focused on species with different lifespans. In a meta-analysis, Fushan et al., for example, analyzed the RNA-seq gene expression in liver, kidney, and brain, for 33 mammalian species from different families and with varying lifespans. This uncovered parallel evolution of gene expression and lifespan, and revealed genes evolving in agreement with the gradient of life-history variation as well as several processes and pathways (Fushan et al., 2015). In a recent study, the de novo assembled transcriptome of the gray whale (considered among the top 1% longest-lived mammals) has also been compared with that of other long- and short-lived mammals (including bowhead and minke whales). The authors show that long-lived mammals share common gene expression patterns, including high expression of DNA maintenance and repair, autophagy, ubiquitination, apoptosis, and immune responses (Toren et al., 2020).

Technological and methodological advances promise to allow even more powerful surveys of the molecular state of cells. Ribosome profiling is one recent approach, also based on next-generation sequencing platforms, consisting of sequencing ribosome-protected mRNA fragments. Compared to RNA-seq using total mRNA, ribosome profiling has the advantage that it is surveying active ribosomes, known as the translatome, and thus can be used to quantify the rate of protein synthesis, which is thought to be a better predictor of protein abundance (Ingolia, Ghaemmaghami, Newman, & Weissman, 2009). Advances in sequencing technology have also allowed for quantitative surveys of changes at the DNA level, including quantifying mutation accumulation with age in the genome and at the level of the mitochondrial genome (reviewed in de Magalhaes et al., 2010). One recent study found an age-related increase in human somatic mitochondrial mutations inconsistent with oxidative damage (Kennedy, Salk, Schmitt, & Loeb, 2013). Another study in aging mice found no increase in mitochondrial DNA point mutations or deletions, questioning whether these play a role in aging (Ameur et al., 2011).

Another level of changes during the life course comes from epigenetics. These are heritable changes that are not caused by changes in the DNA sequence. Large-scale profiling of epigenetic changes with age is now becoming more common, and with the dropping costs of sequencing will no doubt become even more widespread. It is clear that epigenetic changes, like methylation, are associated with age as well as with age-related diseases (Johnson et al., 2012). For example, two recent studies found epigenetic (methylation) marks highly predictive of chronological age in humans (Hannum et al., 2013; Horvath, 2013). Recently, an additional epigenetic biomarker of aging (DNAm

PhenoAge), incorporating several clinical measures of phenotypic age assayed in whole blood, has been developed to predict a variety of aging outcomes, including all-cause mortality, cancers, healthspan, physical functioning, and Alzheimer's disease (Levine et al., 2018). While there is still debate concerning the causality of epigenetic changes, and whether they are causes or effects of age-related degeneration, this is another field that has received increased attention lately (Field et al., 2018; Schnabl, Westermeier, Li, & Klingenspor, 2018). Modern approaches that allow the epigenome to be surveyed on a genome-wide scale however, such as methyl-DNA immunoprecipitation (MeDIP) for surveying DNA methylation and ChIP-Seq for studying histone modifications, provide the tools for researchers to study the epigenetics of aging (de Magalhaes et al., 2010). As such, the epigenome is yet another layer of genomic regulation that can be studied in a high-throughput fashion across the lifespan and in manipulations of longevity.

For all the success of transcriptomics, proteins are of course the actual machines of life and the correlation between transcripts and protein levels is not perfect. Transcriptomics provides a snapshot of transcriptional responses but in the context of aging we need proteomics to truly assay what changes occur with age. Proteomics approaches are still limited, however, in that they do not allow a comprehensive survey of the proteome in a single experiment (de Magalhaes, 2009). There have been some advances, though the number of proteins surveyed is often small compared to transcriptional profiling. For example, protein profiling of aging has been performed in the mouse heart, revealing 8 and 36 protein spots whose expression was, respectively, upregulated and downregulated due to aging (Chakravarti et al., 2008); comparable results in terms of number of proteins were also found in the mouse brain (Yang et al., 2008). Insights can be gained, however, and for instance proteome profiling of aging in mice kidney revealed functional categories associated with aging related to metabolism, transport, and stress response (Chakravarti et al., 2009). More recently, plasma proteomic profiling has revealed a proteomic signature of age (Tanaka et al., 2018).

Another emerging approach to profile age-related changes involves surveying the metabolome. One study compared metabolic parameters of young and old mice, which was then integrated with gene expression and biochemical data to derive a metabolic footprint of aging (Houtkooper et al., 2011). Another study determined the sera metabolite profile of mice of different ages and different genetic backgrounds to derive a metabolic signature that predicts the biological age in mice (Tomas-Loba, Bernardes de Jesus, Mato, & Blasco, 2013). In humans, a panel of 22 metabolites was found to significantly correlate with age and with age-related clinical conditions independent of age (Menni et al., 2013), while another group found 71 and 34 metabolites significantly associated with age in, respectively, women and men (Yu et al., 2012). In another study of more than 300 unique compounds, significant changes in the relative concentration of more than 100 metabolites were associated with age (Lawton et al., 2008). Complementarily, centenarians have also been shown to be characterized by a metabolic phenotype that is distinct from that of elderly subjects, in particular regarding amino acids and lipid species (Collino et al., 2013; Montoliu et al., 2014).

The relevance of large-scale approaches is likely to increase in the near future thanks to the rapid development of single-cell sequencing. As currently the most developed class among these techniques, single-cell RNA sequencing (scRNA-seq) can typically measure the transcriptome of hundreds of thousands of individual cells simultaneously (Regev et al., 2017). Methods targeting other molecular profiles, such as the epigenome, also tend to show promising development (Clark, Lee, Smallwood, Kelsey, & Reik, 2016), with in some cases the possibility of measuring several profiles in parallel in the same cell (Angermueller et al., 2016). An important opportunity given by such technology is the capacity to better define the notion of cell type and to build a cell atlas, such as the recently created murine atlas (Tabula Muris Consortium, 2018; Han et al., 2018) and the future human cell atlas (Regev et al., 2017).

One important limitation in age-related omics profiling, for example, transcriptomics, is that changes could be due to changes in cell types within a given tissue. As such, biogerontology will benefit from single-cell sequencing and it is encouraging to notice that a few scRNA-seq studies have already been specifically designed to investigate the aging process. While results have recently been published for the human pancreas (Enge et al., 2017), the mouse lung (Angelidis et al., 2019), and the mouse brain (Ximerakis et al., 2019), two other studies provide aging data for, respectively, 3 and 18 tissues obtained from the same mice (Kimmel et al., 2019; Pisco et al., 2020). A common conclusion of these early investigations is that aging gives rise to distinct transcriptional trajectories among cell types rather than to a universal pattern. For instance, Ximerakis et al. and Kimmel et al. both report more than 3000 genes differentially expressed with age, but most of them are specific to only one cell type and some of them show opposite regulations in different cell types. Along the same lines, Kimmel et al. observe that changes in transcriptional noise depend on cell identity (however this is in contrast with Enge et al. and Angelidis et al., both reporting an overall increase of such noise in the cell

types they considered). As an interesting example of multi-omics analysis, Angelidis et al. also show that the combination of scRNA-seq with proteomics reveals the cellular origin of extracellular matrix remodeling in the lungs of old mice. Although we have only considered here a few examples of the results contained in the aforementioned articles, we can already start appreciating the relevance and the potential of single-cell sequencing for aging research and large-scale analysis.

## Challenges in data analysis

Although large-scale omics approaches facilitate a broad array of studies and provide an incredible amount of data, the sheer volume of data generated creates challenges in turning the data into meaningful results and novel insights. From a statistical perspective, large-scale approaches also increase the chances of false hits that need to be accounted for when analyzing and interpreting the results. The uncertainties concerning potential false results in large-scale approaches emphasize the need for further experimental validation using a different, usually low-scale, approach. In gene expression studies, qPCR validation is usually used as the gold standard (Derveaux, Vandesompele, & Hellemans, 2010). Some types of studies, like genetic association studies of longevity, are not simple to validate, and often depend on further studies in other populations, which may or may not be feasible.

In a sense, the bottleneck in research using post-genome technologies is moving away from generating data toward interpreting data. As an example, a single 3-day run from an Illumina HiSeq X platform generates up to 900 Gb of data, which must be stored, processed, quality-controlled, and analyzed. This means that the standard experiment using next-generation sequencing platforms must account for a substantial amount of time for the bioinformatics and statistical processing of the data. Although several software tools exist now for this in silico work, labs not experienced with bioinformatics might struggle to develop a suitable pipeline and have to rely on core facilities, collaborators, or commercial services. Another problem is that for many next-generation sequencing approaches, especially in the case of single-cell sequencing, there is still no gold standard for the bioinformatics and statistical analysis and multiple alternative analysis pipelines still exist. Modest alterations in statistical parameters, for which there is no established standard, can also result in significant changes in results. For example, it is important to mention that microarray platforms for gene expression profiling are at present much quicker in terms of data analysis than approaches based on next-generation sequencing; because microarrays have been used for longer, standard methods are available for them and this is not yet the case for RNA-seq. This is even more relevant for single-cell sequencing for which comparative studies and standardization have only started to be proposed very recently (Luecken & Theis, 2019; Wang, Li, Nelson, & Nabavi, 2019). Researchers planning experiments need to carefully balance the advantages of the latest next-generation sequencing platforms with the price and simpler bioinformatics and statistics of array-based platforms.

One major and long-recognized problem of large-scale approaches is multiple hypothesis testing. Even a low-density microarray platform with a few hundred genes is testing for effects a few hundred times, which by chance will generate false positives. Modern genomic approaches, for example, in GWAS studies that survey millions of genetic variants, must adequately cope with this problem to generate biologically relevant results. A standard way of dealing with multiple hypothesis testing is the Bonferroni correction, in which the $P$ value cutoff (typically.05) is divided by the number of hypotheses being tested (e.g., for an array with 20,000 genes use .05/20,000 as cutoff). Bonferroni correction can be deemed as too stringent, and alternative methods for correcting for false positives have been developed (Storey & Tibshirani, 2003). Benjamini correction is also widely used, and is less stringent and straightforward to calculate (Benjamini & Hochberg, 1995). False discovery rate estimates based on simulations and scrambling of data have also been widely used, including by our lab (de Magalhaes et al., 2009; Plank et al., 2012, 2013), and although it requires some customization to the specific experiments, it provides an estimate of false positives based on real data captured from the experiment.

## Data integration

As mentioned previously, the recent shift in biological research toward large-scale approaches has resulted in the capacity to generate huge amounts of data, much of which is publicly available. These data, however, are in most cases heterogeneous and obtained at different time scales and biological levels. Moreover, differences also often exist due to platform and methodology diversity. Still, if our aim is to obtain a global picture of complex processes, such as aging and most age-related diseases, we have to develop the computational methods and tools that allow us to integrate and analyze these diverse data. In this section, we give an overview of the online resources currently

available for aging research and discuss some of the studies that aim to integrate and analyze various types of data. This can be used on its own using public databases or in combination with data from one's own experiment(s).

## Data and databases

Before diving into aging-specific resources, it should be mentioned that one important prerequisite step for data integration, the existence of databases, has already seen a tremendous expansion in the past years and continues to develop at increasing speeds. Currently, there are a number of databases, for humans and model organisms, which host a plethora of information available in a standardized, computational-retrievable and-usable form (in many cases these data are even manually curated to improve quality). These databases provide access both to a wide range of -omes (including genomes, transcriptomes, proteomes, epigenomes, interactomes, reactomes, etc.) and to a multitude of functional data (including biological processes, molecular functions, appurtenance to molecular pathways, etc.). Obviously, integrating this type of information with aging-specific data leads to a more holistic perspective of the aging process and can help in a number of analyses. Although the field of biogerontology has seen a slower increase in integrative systems biology, a series of resources specific to aging have also been created in recent years (Table 6.1), in particular in the context of our Human Ageing Genomic Resources, which are arguably the benchmark in the field (Tacutu et al., 2018).

It should also be noted that although each of these databases acts as a stand-alone resource, focusing on certain facets of aging, in many cases they also show common patterns. For example, in the Human Ageing Genomic Resources, there are many genes that can be found in two or more databases (Fig. 6.2), hence also increasing the confidence of their association to aging.

Similarly, a number of databases for age-related diseases have been developed, though the quality and type of data vary greatly. For example, there are many very good databases for cancer, while the number of databases for heart diseases is still limited. A nonexhaustive list of databases for age-related diseases is provided in Table 6.2.

While some of the resources presented above integrate data related to more than one facet of aging and/or age-related diseases, the concept of multidimensional data integration, at least at the level of aging- and disease-specific databases, is still in its infancy and the task is usually left to the researchers performing integrative analyses. Some large resources, however, like NCBI and Ensembl integrate different types of data and are of course major resources for data integration.

One other aspect that should be kept in mind is that sometimes even the amount of high-throughput

TABLE 6.1 List of major online databases and resources related to aging.

| Name (citation) | Web address | Short description |
|---|---|---|
| AnAge (Tacutu et al., 2018) | http://genomics.senescence.info/species/ | Aging, longevity, and life history information in animals |
| CellAge (Avelar et al., 2020) | https://genomics.senescence.info/cells/ | Database of genes associated with cell senescence |
| Comparative Cellular and Molecular Biology of Longevity Database (Stuart et al., 2013) | http://genomics.brocku.ca/ccmbl/ | Database with cellular and molecular traits from vertebrate species collected to identify traits correlated with longevity |
| DrugAge (Barardo et al., 2017) | https://genomics.senescence.info/drugs/ | Database of drugs, compounds, and supplements that extend longevity in model organisms |
| GenAge (Tacutu et al., 2018) | http://genomics.senescence.info/genes/ | Genes associated with longevity and/or aging in model organisms and candidate aging-related human genes |
| GenDR (Wuttke et al., 2012) | http://genomics.senescence.info/diet/ | Genes associated with dietary restriction from mutations and gene expression profiling |
| AgeFactDB (Huhne, Thalheim, & Suhnel, 2014) | http://agefactdb.jenage.de/ | Observations on the effect of aging factors on lifespan and/or aging phenotype |
| Digital Ageing Atlas (Craig et al., 2014) | http://ageing-map.org/ | Database of molecular, physiological, and pathological age-related changes |
| LongevityMap (Budovsky et al., 2013) | http://genomics.senescence.info/longevity/ | Database of human genetic variants associated with longevity |

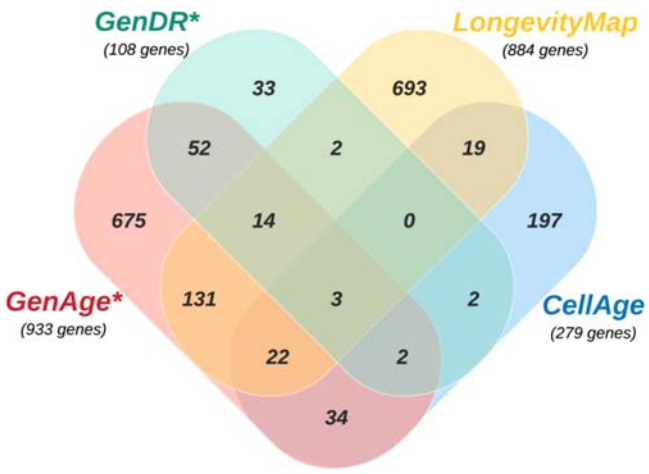


FIGURE 6.2 Venn diagram for the overlap of genomic databases in the Human Ageing Genomic Resources (Tacutu et al., 2018). Data sources used included: *GenAge, build 19*: 307 human genes and 2152 genes in model organisms; *GenDR, build 4*: 214 genes in model organisms; *LongevityMap, build 3*: 884 human genes; *CellAge, build 1*: 279 human genes. *Overlap between the four datasets was computed considering the human genes from LongevityMap and CellAge, and the nonredundant list of human genes and orthologs of genes in model organisms for GenAge and GenDR.

information for only one type of data may pose computational challenges, both in terms of handling and analyzing. Consequently, integrating and analyzing data from multiple sources will result in an even bigger challenge, the complexity increasing in most cases in a nonlinear fashion, and as such data integration comes at a cost, the haystack in which the needles have to be found increases exponentially.

## Finding needles in haystacks: Network approaches and multidimensional data integration

With the expansion of large-scale approaches, and the inevitable increase in age-related data available, new hypotheses of aging trying to integrate multidimensional information have been developed. More than 20 years ago, the idea that aging was caused not simply by the failure of individual components, but rather by a network of parallel and gradual dysregulations, was proposed (Kirkwood & Kowald, 1997). While the effects of each individual event could be relatively small, the authors argued that the integrative

TABLE 6.2 Selected online databases and resources related to age-related diseases.

| Name (citation) | Web address | Short description |
|---|---|---|
| Ageing-related Disease Genes (Fernandes et al., 2016) | https://genomics.senescence.info/diseases/ | Dataset obtained from the comparison of aging-related genes with age-related disease genes |
| OMIM (OMIM, 2014) | http://www.omim.org/ | Online Mendelian Inheritance in Man (part of NCBI) |
| Catalog of Published Genome-Wide Association Studies (Buniello et al., 2019) | https://www.ebi.ac.uk/gwas/ | NHGRI-EBI catalog of GWAS studies |
| Genetic Association Database (Zhang et al., 2010) | http://geneticassociationdb.nih.gov/ | Archive of human genetic association studies of complex diseases and disorders |
| AlzGene (Bertram, McQueen, Mullin, Blacker, & Tanzi, 2007) | http://www.alzgene.org/ | Database with genetic resources for Alzheimer's disease |
| The COSMIC Cancer Gene Census (Sondka et al., 2018) | https://cancer.sanger.ac.uk/census/ | Catalog of genes for which mutations have been causally implicated in cancer |
| COSMIC (Tate et al., 2019) | https://cancer.sanger.ac.uk/cosmic | Catalogue Of Somatic Mutations In Cancer |
| The Cancer Genome Atlas (TCGA) | http://cancergenome.nih.gov/ | Portal providing access to cancer-related large-scale data from the NCI and NHGRI |
| The Human Metabolome Database | http://www.hmdb.ca/ | Database of small-molecule metabolites found in the human body |
| Roadmap Epigenomics Project | http://www.roadmapepigenomics.org/ | Resource of human epigenomic data |
| TSGene (Zhao, Kim, Mitra, Zhao, & Zhao, 2016) | https://bioinfo.uth.edu/TSGene/ | Tumor suppressor gene database |
| Progenetix (Cai et al., 2014) | https://progenetix.org/ | Copy number abnormalities in human cancer from comparative genomic hybridization experiments |
| MethyCancer (He et al., 2008) | http://methycancer.psych.ac.cn/ | Human DNA methylation and cancer |
| PubMeth (Ongenaert et al., 2008) | http://www.pubmeth.org/ | Cancer methylation database |
| LncRNADisease (Bao et al., 2019) | http://www.rnanut.net/lncrnadisease/ | Long noncoding RNA and disease associations |
| HMDD 2.0 (Huang et al., 2019) | http://www.cuilab.cn/hmdd | Experimentally supported human microRNA and disease associations |

contribution of defective mitochondria, aberrant proteins, and free radicals, taken together, could explain many of the major changes that occur during aging. Although, since it was first proposed, many other types of aging factors have been taken into consideration (and the strife to integrate more will probably continue), the "network theory of aging" might have been the onset of studying aging in a holistic way.

Network biology provides a conceptual framework to study the complex interactions between the multiple components of biological systems (Barabasi, Gulbahce, & Loscalzo, 2011). In network biology, a network is defined as a set of nodes (as a mathematical model for genes, proteins, metabolites, etc.) with some node pairs being connected through directed/asymmetric or undirected/symmetric edges (as a model for physical interactions, coexpression relationships, metabolic reactions, etc.). Depending on the type of components and the nature of the interactions that are analyzed, there is currently a large variety of network types that can be constructed; perhaps the most used being protein interaction networks, gene regulatory networks, coexpression networks, and metabolic networks.

With regards to aging research, the idea of analyzing many longevity/aging determinants at the same time has been pushed forward, mostly due to the accumulating knowledge about the genetic determinants of aging (de Magalhaes & Toussaint, 2004a; Tacutu et al., 2018), but also by the development of bioinformatics tools pertaining to network biology like Cytoscape (Saito et al., 2012). Not surprisingly, network-based approaches have been increasingly used to study aging and age-related diseases (Barabasi et al., 2011; de Magalhaes et al., 2012; Fernandes et al., 2016; Ideker & Sharan, 2008; Smita, Lange, Wolkenhauer, & Kohling, 2016; Soltow, Jones, & Promislow, 2010). Common topics addressed by these approaches include network construction for aging/longevity or various related conditions, analysis of topological features, finding functional submodules, etc.

### *Construction of longevity networks*

The initial attempts to construct longevity networks date back more than 15 years. As a first step toward the construction of a human aging network, we used genes previously associated with aging and their interacting partners, using protein–protein interaction data, to construct networks related to DNA metabolism and the GH/IGF-1 pathway. We further suggested, based on a "guilt-by-association" methodology, that among the interacting partners of genes associated with aging there could also be other genes that are involved in aging. Additionally, functional analysis of the network revealed that many of the genes which are important during development might also regulate the rate of aging (de Magalhaes & Toussaint, 2004a).

One central question in aging research is whether genes and pathways associated with aging and longevity are evolutionary conserved. For example, in Fig. 6.3 is a schematic representation of longevity protein interaction networks across model organisms. However, the question of relevance arises: are aging-related data in one species also relevant in another species? This is an important issue since at times the data available in different species could be used complementarily. Results so far suggest that genes whose manipulation results in a lifespan effect tend to be highly evolutionary conserved across divergent eukaryotic species (Budovsky, Abramovich, Cohen, Chalifa-Caspi, & Fraifeld, 2007; de Magalhaes & Church, 2007; Fernandes et al., 2016; Smith et al., 2008; Yanai, Budovsky, Barzilay, Tacutu, & Fraifeld, 2017). Moreover, while not universal, some empirical data suggest that the effect on longevity of many of these genes is also conserved (Smith et al., 2008). As such, it is not completely senseless to integrate longevity-associated genes from multiple species. Using this premise, it was then shown that the human orthologs of longevity-associated genes from model organisms, together with their interacting partners, could act in a cooperative manner and form a continuous protein–protein interaction network, called the Human Longevity Network (Budovsky et al., 2007).

### *Topological features*

One important aspect in network biology is the analysis of a network's topological characteristics (i.e., studying the way in which the nodes and edges of a network are arranged). Particular focus has been on scale-free networks, a very common type of network among social and biological networks. The scale-free topology means that the nodes in the network have a connectivity distribution p(k) given by a power-law function $k^{-\gamma}$, where p(k) is the probability that a certain node has exactly k edges, and $\gamma$ is the degree exponent, a parameter value which for most of the studied networks is usually between 2 and 3 (Barabasi & Albert, 1999). The aforementioned Human Longevity Network has a scale-free topology, with a high contribution of hubs (highly connected genes) to the overall connectivity of the network. Interestingly, almost all of the hubs in the longevity network had been reported previously to be involved in at least one age-related pathology (Budovsky et al., 2007), suggesting a link between diseases and the mechanisms regulating longevity.

The scale-free design can be found in a wide range of molecular and cellular systems, largely governing their internal organization (Barabasi & Oltvai, 2004), and it appears to have been also favored by evolution (Oikonomou & Cluzel, 2006). Although a more detailed

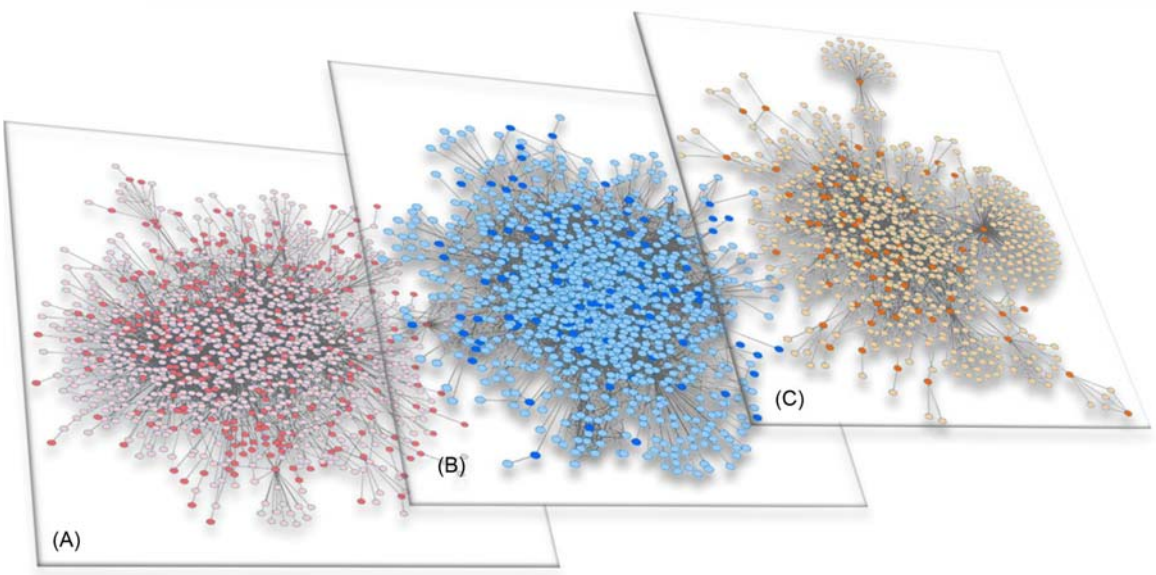


FIGURE 6.3 Schematic view on longevity networks across species. (A) Worm longevity network. (B) Fly longevity network. (C) Mouse longevity network. (A–C) Networks include longevity-associated genes (LAGs) from the GenAge database, and their protein–protein interaction from the BioGRID database (Stark et al., 2006). Dark/light colors depict LAGs and LAG-interacting partners. The number of nodes in each network (based on GenAge, build 19 and BioGRID, release 3.5.178) is summarized in the table below.

| Species | LAGs in GenAge | LAGs with interactions | Longevity network | LAGs in the network |
|---|---|---|---|---|
| Worm | 875 | 545 | 2534 | 525 |
| Fly | 191 | 162 | 2180 | 161 |
| Mouse | 136 | 99 | 834 | 94 |

discussion about the evolvability of complex networks is beyond the scope of this chapter, it should be mentioned that the properties of scale-free networks confer some net advantages in solving cellular tasks. For example, this type of architecture permits an efficient local dissipation of external perturbations, while at the same time reliably transmitting signals (and discriminating against noise) between distant elements of the network (Csermely & Soti, 2006). Additionally, the scale-free property offers an unexpected degree of robustness, maintaining the ability of nodes to communicate even under extremely high fault rates, by minimizing the effect of random failures on the entire network (Albert, Jeong, & Barabasi, 2000; Wagner, 2000).

On one hand, analyzing the aging/longevity networks can provide a framework for the conceptualization of the aging process and may reveal fundamental traits and constraints of biological systems. On the other hand, networks can help in assessing the importance of genes in a certain process. For example, differentiating between the hubs of a longevity network and all other nodes is often a very attractive way of reducing a candidate list. This approach comes as no surprise, as some components of a cellular network are more important than others with regard to aging. It was previously shown that longevity-associated genes in model organisms have a higher average connectivity, with many being network hubs (Budovsky et al., 2007; Ferrarini, Bertelli, Feala, McCulloch, & Paternostro, 2005; Promislow, 2004). Moreover, it has been established that there is a positive correlation between a protein's connectivity and its degree of pleiotropy, an elevated degree being common among proteins associated with senescence (Promislow, 2004). As such, it makes sense in choosing highly connected longevity candidates. Still, it should also be kept in mind that other topological measures besides degree also exist (e.g., closeness, eigenvector centrality, betweenness, bridging centrality, etc.) and their usage could result in a different sorting order. Ultimately, no matter what the selection criteria are, experimental validation is warranted.

### *Network modularity*

Focusing on entire categories of genes or on network modules, and on the cross-talk between these modules, could provide valuable and unique hints regarding the system's susceptibility to failure. In relation to this, Xue et al. examined the modular structure

of protein–protein interaction networks during brain aging in flies and humans. Interestingly, they found two large modules of coregulated genes, both associated with the proliferation–differentiation switch, displaying opposite age-related expression changes. A few other modules found to be associated with the oxidative–reductive metabolic switch were found, but only during fly aging. Overall, the authors found that aging is associated with a limited number of modules which are interlinked through genes more likely to affect aging/longevity (Xue et al., 2007).

### *Multidimensional data integration*

Age-related changes can be found at many levels (expression changes, posttranslational modifications, cross-linking, or alterations in protein interactions), yet integration of multidimensional data is still in its early stages. Attempts to integrate protein–protein interaction networks with transcriptional data have already been made with relative success. As partly mentioned previously, a new analytic method permitting the integration of both transcriptome and interactome information has been employed to study network modularity in aging (Xue et al., 2007). In another study, a human protein interaction network for longevity was used in conjunction with transcriptional data from muscle aging in humans for the prediction of new longevity candidates (Bell et al., 2009).

Our meta-analysis of CR microarray studies in mammals integrated co-expression data, information on genetic mutants, and analysis of transcription factor binding sites to reveal promising candidate regulators, providing a comprehensive picture of the changes that occur during CR. In addition to the several processes previously associated with CR mentioned above, we also found novel associations, such as strong indications of the effect that caloric restriction has on circadian rhythms (Plank et al., 2012). A cross comparison of insulin signaling and CR models, using transcriptomics and metabolomics, identified some key pathways and metabolite fingerprints shared in long-lived strains, highlighting the potential role of increased amino acid metabolism and purine upregulation in longevity (Gao et al., 2018). Addressing another crucial aim in gerontology, the need to have reliable biomarkers of aging can also be done by using network-based approaches, and the integration of networks with gene expression data to create modular biomarkers of aging has been carried out (Fortney, Kotlyar, & Jurisica, 2010).

Functional classification analysis, using for example web tools like DAVID (Huang da, Sherman, & Lempicki, 2009) which analyze Gene Ontology and pathway annotations, can also generate useful information regarding the nature of longevity-associated genes. For example, several studies have already shown that inhibition of translation can be an effective modulator of lifespan extension (Curran & Ruvkun, 2007; Hansen et al., 2007; Pan et al., 2007). The integration of large-scale lists of genes with gene annotation data is therefore common in analyzing omics experiments and can provide insights concerning mechanisms, processes, and pathways (reviewed in de Magalhaes et al., 2010).

The study of aging is strongly linked to that of age-related diseases. This becomes obvious when looking at the overlap between the genes associated with major age-related diseases (including atherosclerosis, cancer, type II diabetes, and Alzheimer's disease) and the genes involved in lifespan regulation (Budovsky et al., 2007, 2009; Fernandes et al., 2016; Tacutu, Budovsky, Yanai, & Fraifeld, 2011; Wolfson, Budovsky, Tacutu, & Fraifeld, 2009), as well as when analyzing the many direct and indirect molecular interactions which exist between them (Simko, Gyurko, Veres, Nanasi, & Csermely, 2009; Tacutu et al., 2011). Networks have been extensively used for the study of diseases (Goh & Choi, 2012; Ideker & Sharan, 2008). Recently, examples of analyses of multidimensional data for age-related diseases have also started to amass. For example, based on a network of genes and diseases created by Goh et al. (2007), structural facets of proteins, such as the intrinsic disorder content, and epigenetic aspects as alternative splicing, have been studied (Midic, Oldfield, Dunker, Obradovic, & Uversky, 2009). Models of diseases–genes–drugs have also been constructed, and new insights have been found about the usage of drugs (Yildirim, Goh, Cusick, Barabasi, & Vidal, 2007). However, outside the scope of this chapter, gene–drug interaction data are thus another type of data that can be used, and indeed there are successful examples in aging research (Barardo et al., 2017; Calvert et al., 2016; Donertas, Fuentealba Valenzuela, Partridge, & Thornton, 2018; Ziehm et al., 2017). In fact, a network-based view of drug discovery and biomarkers is starting to emerge to also account for the complexity of human biology (de Magalhaes et al., 2012; Erler & Linding, 2010).

In the context of GWAS, combining GWAS with phylogenetic conservation and a complexity assessment of co-occurring transcription factor binding sites can identify *cis*-regulatory variants and elucidate their mechanistic role in disease. This has been recently carried out for type II diabetes successfully linking genetic association signals to disease-related molecular mechanisms (Claussnitzer et al., 2014). For Parkinson disease, integrative analyses of gene expression and GWAS data have also provided key insights into the genetic etiology of the disease (Edwards et al., 2011). Lastly, constructing molecular networks based on whole-genome gene expression profiling and genotyping data, together with the use of Bayesian inference, has helped to identify key

causal regulators in late-onset Alzheimer's disease (Zhang et al., 2013b). Studies of proteomic and metabolomic networks, although in their infancy, may be another key step in constructing an integrative framework to study aging and age-related disease. Such integration has been suggested as needed, complementarily to the more classical genome–transcriptome–phenotype model (Hoffman, Lyu, Pletcher, & Promislow, 2017).

The above examples are only a selected few since many different types of network analyses and data integration can be performed. It is not surprising that integrative approaches are starting to be used to combine disease and aging-related data. While some studies have focused on a particular disease and its links to aging/longevity (Budovsky et al., 2009; Miller, Oldham, & Geschwind, 2008), others have attempted in a broader way to look at the common signatures of aging/longevity and all major age-related diseases (Wang, Zhang, Wang, Chen, & Zhang, 2009; Wolfson et al., 2009).

In order to better understand the gene expression and protein level changes that occur with age, other genomic and epigenetic layers should be considered. For example, age-related changes in miRNA expression profiles can have a significant impact on protein levels. In terms of data integration, combining miRNA data with protein–protein interactions has been used to analyze the molecular links between aging, longevity, and age-related diseases, and to suggest the potential role for miRNAs in targeting certain genes with features of antagonistic pleiotropy, implying thus a preferability to initiate longevity-promoting interventions in adult life (Tacutu, Budovsky, Wolfson, & Fraifeld, 2010). In another study, interpreting the methylation patterns in cancer and aging has been done using an integrative system. By developing a novel epigenome–interactome approach with differential methylation data, tissue-independent age-associated methylation hot spots targeting stem-cell differentiation pathways have been recently discovered (West, Beck, Wang, & Teschendorff, 2013).

One important aspect of data integration is that integrating multiple data sources will significantly expand our view of the aging process, and it is possible that some of the current well-accepted hypotheses will even be challenged. For example, although at the network level of protein–protein interactions it seems that hub genes are of utmost importance for the robustness of the entire network, when looking at an epigenetic level it has been suggested that the age-associated drift in DNA methylation occurs preferentially in genes that occupy peripheral network positions of exceptionally low connectivity (West, Widschwendter, & Teschendorff, 2013). Only by having a complete, multilayered picture of the aging process can we hope to fully understand its intricacies.

The emerging concept of *multilayer networks* looks therefore relevant and promising to shed new light on integrative approaches. The main idea is to go beyond the usual one-dimensional representation of networks containing only one type of nodes and one type of edges. Indeed, the above discussion makes it clear that biological systems are complex, heterogeneous, and involve various types of agents and relations between them. Interestingly, the last few years have shown increasing efforts to develop new mathematical tools to adapt the features of usual networks (such as topology and modularity) to more complex structures. As such, Kivelä et al. offer both a comprehensive review of the subject and a unifying classification of the various types of multilayer networks (Kivelä et al., 2014). Of interest, Halu et al. recently applied such techniques to create and analyze a multiplex network of human diseases (Halu, De Domenico, Arenas, & Sharma, 2019). Their model is made of two layers of similar nodes, each node representing a disease. In the first layer, diseases are connected by an edge if they share common genes, whereas in the second layer they are connected if they share common symptoms. Considering the two layers as part of a single system offers an integrative genotype–phenotype approach. The authors report, for example, that diseases sharing common genes tend to share common symptoms or also that multiplex community detection allows to confirm and find new disease associations. Inspired by such results, it is expected that similar studies applied to aging/longevity networks will appear in the future.

## Predictive methods and models

Given the intrinsic costs of performing animal aging studies, particularly in mammals, developing predictive computational tools is of utmost importance. Indeed, to identify suitable drug targets with antiaging properties, methods for prioritizing them are necessary (de Magalhaes et al., 2012). Fortunately, many computational tools are already available for prioritizing candidates (Moreau & Tranchevent, 2012), and could be of great use to biogerontologists.

One of the main assumptions for many predictive methods is based on the "guilt-by-association" principle, in which new genes or drugs are considered candidates due to their relation with genes that are already known to be associated with aging or longevity. Though this premise is common to many strategies, they usually differ in the type of associations that are considered. For example, based on the finding that hubs and centrally located nodes have a higher likelihood to be associated with aging/longevity, Witten and Bonchev used a *C. elegans* network to predict new

longevity-associated genes (Witten & Bonchev, 2007). Likewise, other topological measures have been employed for similar goals. Using a proximity measure in a yeast network (the shortest path to an already known gene reported to be associated with an increased lifespan), Managbanag et al. identified a set of single-gene deletions predicted to affect lifespan. Testing this experimentally, their validation showed that the predicted set was enriched for mutations conferring either increased or decreased replicative lifespan (Managbanag et al., 2008). In another example, using machine learning and classification techniques, Freitas et al. devised a predictive model to discriminate between aging-related and nonaging-related DNA repair genes. In this analysis, they found that gene connectivity together with specific gene ontology terms, having an interaction with the XRCC5 protein, and a high expression in T lymphocytes are good predictors of aging association for human DNA repair genes (Freitas, Vasieva, & de Magalhaes, 2011). Because machine learning is outside the scope of this chapter we refer readers to a recent review of machine learning in aging research (Fabris, Magalhaes, & Freitas, 2017).

In *C. elegans*, various properties of longevity genes have been analyzed and then used to verify the prediction of new longevity regulators (Li, Dong, & Guo, 2010). In one study, the authors found that longer genomic sequences, co-expression with other genes during the transition from dauer to nondauer state, enrichment in certain functions and RNAi phenotypes, higher sequence conservation, and a higher connectivity in a functional interaction network, are all predictors of an association with longevity. While the validation of the prediction was computational only, based on the precision calculated with a 10-fold cross method for a set of known positive and negative longevity-associated genes, the authors found that a few of the predicted genes had been in the meantime experimentally validated in the scientific literature (Li et al., 2010).

We have recently used a combined approach, first reasoning that the interaction partners of longevity-associated genes are more likely to modulate longevity, and second narrowing down the candidate list based on features of antagonistic pleiotropy. Although by the time of this study several genome-wide longevity assays had been performed in *C. elegans*, our prediction method, followed by experimental validation, resulted in the discovery of new longevity regulators at a frequency much higher than previously achieved (Tacutu et al., 2012).

Combining a network-based approach with transcriptional data from human aging has also been used as a method of prediction. Using a human longevity network constructed based on homologs from invertebrate species, and comparing the result with age-related transcriptional data from human muscle aging, Bell et al. determined a set of human interaction partners potentially involved in aging. Testing the homologs of these genes in *C. elegans* revealed that 33% of the candidates extended lifespan when knocked-down (Bell et al., 2009). In another, more recent study, fly longevity-associated genes from GenAge were analyzed using networks in order to predict key pathways and genes involved in lifespan regulation, which have been shown to have significant transcriptional changes in aging (Li et al., 2019).

Focusing on CR, network and systems biology approaches have also been used to predict genes necessary for the life-extending effects of CR. By looking at genes that are more connected to already-known CR-related genes, Wuttke et al. successfully predicted a set of novel genes mediating the life-extending effects of CR. Nine novel genes related to CR were validated experimentally in yeast. This revealed three novel CR mimetic genes (Wuttke et al., 2012).

In terms of metabolomics, a recent study has used network reconstruction and network analysis of metabolite relationships to associate measured plasma metabolites with sex and age, as well as to analyze the variations in the regulation of metabolic activity of amino acids, lipids, and ketone bodies (Vignoli, Tenori, Luchinat, & Saccenti, 2018). This type of methods could be used as diagnostic and predictive tools in the investigation of the human aging phenotype at a metabolic level.

While in the last few decades many studies in model organisms have successfully identified genetic factors that affect lifespan, the effect of combined interventions (epistasis), whether synergistic or antagonistic, has been evaluated to a much more limited degree. Using network features and biochemical/physicochemical features, a two-layer deletion network model has been recently developed and used for predicting the epistatic effects of double deletions on yeast longevity. Results showed that the functional features (such as mitochondrial function and chromatin silencing), the network features (such as the edge density and edge weight density of the deletion network), and the local centrality of deletion gene are important predictors for the deletion effects on longevity (Huang et al., 2012).

Candidate gene prioritization methods, such as the ones described in this section, have been instrumental in guiding various experiments that provided important insights into aging mechanisms (Lorenz, Cantor, & Collins, 2009; Wuttke et al., 2012; Xue et al., 2007). The accuracy and specificity of these in silico predictive methods is still limited, however. Similarly, while computational methods have been developed for predicting candidate drugs from gene expression data

(Iorio et al., 2010; Lamb et al., 2006; Sirota et al., 2011), these have only been partly implemented in the context of aging (Calvert et al., 2016; Donertas et al., 2018), in spite of their widespread interest.

## Concluding remarks

Biological and medical research has often failed to capture the whole picture of the disease or process under study. Researchers have traditionally focused on a limited number of players that either had the greatest impact or, by chance, happened to be associated with the phenotype of interest. For some diseases (e.g., antibiotics were developed as therapies with only a modest understanding of the mechanisms involved) and processes (e.g., overcoming replicative senescence with ectopic telomerase expression), a limited mechanistic understanding may suffice to develop interventions, but for many others our understanding and models are imperfect at best and possibly even flawed. Researchers do not see the forest for the trees and for complex processes like aging, and age-related diseases like cancer and neurodegenerative diseases, this impedes the development of successful interventions. Not surprisingly, the overall rate of success of clinical trials is only about 20% (DiMasi, Feldman, Seckler, & Wilson, 2010), and to date there is no established approach to retard, even if slightly, human aging. While serendipitous discoveries like antibiotics, are always possible, it is widely acknowledged that the study of complex processes like aging stands a better chance of developing clinical interventions based on broad biological understanding (de Magalhaes, 2014a; de Magalhaes et al., 2017). In addition, the discoveries in the genetics of aging and technological advances in large-scale methodologies, like high-throughput profiling and screening, mean that it is vital now to cope with the growing amount of data in the context of drug discovery (de Magalhaes et al., 2012). As new layers of genomic regulation are uncovered (e.g., noncoding RNAs) this raises new challenges and further emphasizes the need to study biological systems in a comprehensive fashion to capture and decipher their intrinsic complexity.

Overall, our belief is that the combination of large-scale approaches to unravel both age-related changes as well as identify the causes for variability across individuals and species will drive the field forward. These require, however, adequate data and statistical analysis to avoid biases and false results. The integration of different types of data provides opportunities for synergy and discovery that we believe will result in a much deeper understanding of aging and the development of interventions to extend lifespan and preserve health.

## Acknowledgments

We are thankful to Gabriela Bunu and Dmitri Toren for helping in the creation of Figs. 6.2 and 6.3. JPM is grateful to the UK Biotechnology and Biological Sciences Resource Council (BB/R014949/1), the Wellcome Trust (208375/Z/17/Z), the Royal Society (RG120543), the Leverhulme Trust (RPG-2016-015), and LongeCity for supporting his lab's development of integrative genomic resources and methods described in this chapter. RT is grateful for the funding support received through the Competitiveness Operational Programme 2014–2020, POC-A.1-A.1.1.4-E-2015. CL is grateful for the funding provided by the Human Frontier Science Program (fellowship LT000741/2019-C). We apologize to those whose relevant papers are not cited owing to lack of space.

## References

Acosta, J. C., O'Loghlen, A., Banito, A., Guijarro, M. V., Augert, A., Raguz, S., ... Gil, J. (2008). Chemokine signaling via the CXCR2 receptor reinforces senescence. *Cell*, *133*(6), 1006–1018. Available from https://doi.org/10.1016/j.cell.2008.03.038.

Albert, R., Jeong, H., & Barabasi, A. L. (2000). Error and attack tolerance of complex networks. *Nature*, *406*(6794), 378–382. Available from https://doi.org/10.1038/35019019.

Ameur, A., Stewart, J. B., Freyer, C., Hagstrom, E., Ingman, M., Larsson, N. G., & Gyllensten, U. (2011). Ultra-deep sequencing of mouse mitochondrial DNA: Mutational patterns and their origins. *PLoS Genetics*, *7*(3), e1002028. Available from https://doi.org/10.1371/journal.pgen.1002028.

Angelidis, I., Simon, L. M., Fernandez, I. E., Strunz, M., Mayr, C. H., Greiffo, F. R., ... Schiller, H. B. (2019). An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nature Communications*, *10*(1), 963. Available from https://doi.org/10.1038/s41467-019-08831-9.

Angermueller, C., Clark, S. J., Lee, H. J., Macaulay, I. C., Teng, M. J., Hu, T. X., ... Reik, W. (2016). Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nature Methods*, *13*(3), 229–232. Available from https://doi.org/10.1038/nmeth.3728.

Avelar, R. A., Ortega, J. G., Tacutu, R., Tyler, E. J., Bennett, D., Binetti, P., ... de Magalhaes, J. P. (2020). A multidimensional systems biology analysis of cellular senescence in aging and disease. *Genome Biol*, *21*(1), 91. https://doi.org/10.1186/s13059-020-01990-9.

Bao, Z., Yang, Z., Huang, Z., Zhou, Y., Cui, Q., & Dong, D. (2019). LncRNADisease 2.0: An updated database of long non-coding RNA-associated diseases. *Nucleic Acids Research*, *47*(D1), D1034–D1037. Available from https://doi.org/10.1093/nar/gky905.

Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512.

Barabasi, A. L., Gulbahce, N., & Loscalzo, J. (2011). Network medicine: A network-based approach to human disease. *Nature Reviews. Genetics*, *12*(1), 56–68. Available from https://doi.org/10.1038/nrg2918.

Barabasi, A. L., & Oltvai, Z. N. (2004). Network biology: Understanding the cell's functional organization. *Nature Reviews in Genetics*, *5*(2), 101–113. Available from https://doi.org/10.1038/nrg1272.

Barardo, D., Thornton, D., Thoppil, H., Walsh, M., Sharifi, S., Ferreira, S., ... de Magalhaes, J. P. (2017). The DrugAge database of aging-related drugs. *Aging Cell*, *16*(3), 594–597. Available from https://doi.org/10.1111/acel.12585.

Barardo, D. G., Newby, D., Thornton, D., Ghafourian, T., de Magalhaes, J. P., & Freitas, A. A. (2017). Machine learning for predicting lifespan-extending chemical compounds. *Aging*, *9*(7), 1721–1737. Available from https://doi.org/10.18632/aging.101264.

Beekman, M., Blanche, H., Perola, M., Hervonen, A., Bezrukov, V., Sikora, E., ... Franceschi, C. (2013). Genome-wide linkage analysis for human longevity: Genetics of Healthy Aging Study. *Aging Cell*, *12*(2), 184–193. Available from https://doi.org/10.1111/acel.12039.

Bell, R., Hubbard, A., Chettier, R., Chen, D., Miller, J. P., Kapahi, P., ... Hughes, R. E. (2009). A human protein interaction network shows conservation of aging processes between human and invertebrate species. *PLoS Genetics*, *5*(3), e1000414. Available from https://doi.org/10.1371/journal.pgen.1000414.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*(1), 289–300.

Bertram, L., McQueen, M. B., Mullin, K., Blacker, D., & Tanzi, R. E. (2007). Systematic meta-analyses of Alzheimer's disease genetic association studies: The AlzGene database. *Nature Genetics*, *39*(1), 17–23. Available from https://doi.org/10.1038/ng1934.

Budovsky, A., Abramovich, A., Cohen, R., Chalifa-Caspi, V., & Fraifeld, V. (2007). Longevity network: Construction and implications. *Mechanisms of Ageing and Development*, *128*(1), 117–124. Available from https://doi.org/10.1016/j.mad.2006.11.018.

Budovsky, A., Craig, T., Wang, J., Tacutu, R., Csordas, A., Lourenco, J., ... de Magalhaes, J. P. (2013). LongevityMap: A database of human genetic variants associated with longevity. *Trends in Genetics: TIG*, *29*(10), 559–560. Available from https://doi.org/10.1016/j.tig.2013.08.003.

Budovsky, A., Tacutu, R., Yanai, H., Abramovich, A., Wolfson, M., & Fraifeld, V. (2009). Common gene signature of cancer and longevity. *Mechanisms of Ageing and Development*, *130*(1-2), 33–39. Available from https://doi.org/10.1016/j.mad.2008.04.002.

Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., ... Parkinson, H. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research*, *47*(D1), D1005–D1012. Available from https://doi.org/10.1093/nar/gky1120.

Cai, H., Kumar, N., Ai, N., Gupta, S., Rath, P., & Baudis, M. (2014). Progenetix: 12 years of oncogenomic data curation. *Nucleic Acids Research*, *42*(1), D1055–D1062. Available from https://doi.org/10.1093/nar/gkt1108.

Calvert, S., Tacutu, R., Sharifi, S., Teixeira, R., Ghosh, P., & de Magalhaes, J. P. (2016). A network pharmacology approach reveals new candidate caloric restriction mimetics in C. elegans. *Aging Cell*, *15*(2), 256–266. Available from https://doi.org/10.1111/acel.12432.

Cevenini, E., Bellavista, E., Tieri, P., Castellani, G., Lescai, F., Francesconi, M., ... Franceschi, C. (2010). Systems biology and longevity: An emerging approach to identify innovative anti-aging targets and strategies. *Current Pharmaceutical Design*, *16*(7), 802–813.

Chakravarti, B., Oseguera, M., Dalal, N., Fathy, P., Mallik, B., Raval, A., & Chakravarti, D. N. (2008). Proteomic profiling of aging in the mouse heart: Altered expression of mitochondrial proteins. *Archives of Biochemistry and Biophysics*, *474*(1), 22–31. Available from https://doi.org/10.1016/j.abb.2008.02.001.

Chakravarti, B., Seshi, B., Ratanaprayul, W., Dalal, N., Lin, L., Raval, A., & Chakravarti, D. N. (2009). Proteome profiling of aging in mouse models: Differential expression of proteins involved in metabolism, transport, and stress response in kidney. *Proteomics*, *9*(3), 580–597. Available from https://doi.org/10.1002/pmic.200700208.

Chatsirisupachai, K., Palmer, D., Ferreira, S., & de Magalhaes, J. P. (2019). A human tissue-specific transcriptomic analysis reveals a complex relationship between aging, cancer, and cellular senescence. *Aging Cell*, *18*(6), e13041. Available from https://doi.org/10.1111/acel.13041.

Christensen, K., Johnson, T. E., & Vaupel, J. W. (2006). The quest for genetic determinants of human longevity: Challenges and insights. *Nature Reviews. Genetics*, *7*(6), 436–448. Available from https://doi.org/10.1038/nrg1871.

Churgin, M. A., Jung, S. K., Yu, C. C., Chen, X., Raizen, D. M., & Fang-Yen, C. (2017). Longitudinal imaging of *Caenorhabditis elegans* in a microfabricated device reveals variation in behavioral decline during aging. *Elife*, 6, e26652. Available from https://doi.org/10.7554/eLife.26652.

Clark, S. J., Lee, H. J., Smallwood, S. A., Kelsey, G., & Reik, W. (2016). Single-cell epigenomics: Powerful new methods for understanding gene regulation and cell identity. *Genome Biology*, *17*, 72. Available from https://doi.org/10.1186/s13059-016-0944-x.

Claussnitzer, M., Dankel., Simon, N., Klocke, B., Grallert, H., Glunk, V., Berulava, T., ... Laumen, H. (2014). Leveraging cross-species transcription factor binding site patterns: From diabetes risk loci to disease mechanisms. *Cell*, *156*(1–2), 343–358. Available from https://doi.org/10.1016/j.cell.2013.10.058.

Collino, S., Montoliu, I., Martin, F. P., Scherer, M., Mari, D., Salvioli, S., ... Rezzi, S. (2013). Metabolic signatures of extreme longevity in northern Italian centenarians reveal a complex remodeling of lipids, amino acids, and gut microbiota metabolism. *PLoS One*, *8*(3), e56564. Available from https://doi.org/10.1371/journal.pone.0056564.

GTEx Consortium. (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*, *348*(6235), 648–660. Available from https://doi.org/10.1126/science.1262110.

Tabula Muris Consortium. (2018). Single-cell transcriptomics of 20 mouse organs creates a *Tabula muris*. *Nature*, *562*(7727), 367–372. Available from https://doi.org/10.1038/s41586-018-0590-4.

Craig, T., Smelick, C., Tacutu, R., Wuttke, D., Wood, S. H., Stanley, H., ... de Magalhaes, J. P. (2015). The Digital Ageing Atlas: Integrating the diversity of age-related changes into a unified resource. *Nucleic Acids Research*, *43*, D873–8. Available from https://doi.org/10.1093/nar/gku843.

Csermely, P., & Soti, C. (2006). Cellular networks and the aging process. *Archives of Physiology and Biochemistry*, *112*(2), 60–64. Available from https://doi.org/10.1080/13813450600711243.

Curran, S. P., & Ruvkun, G. (2007). Lifespan regulation by evolutionarily conserved genes essential for viability. *PLoS Genetics*, *3*(4), e56. Available from https://doi.org/10.1371/journal.pgen.0030056.

de Magalhaes, J. P. (2004). From cells to ageing: A review of models and mechanisms of cellular senescence and their impact on human ageing. *Experimental Cell Research*, *300*(1), 1–10.

de Magalhaes, J. P. (2006). Species selection in comparative studies of aging and antiaging research. In P. M. Conn (Ed.), *Handbook of models for human aging* (pp. 9–20). Burlington, MA: Elsevier Academic Press.

de Magalhaes, J. P. (2009). Aging research in the post-genome era: New technologies for an old problem. In C. H. Foyer, R. Faragher, & P. J. Thornalley (Eds.), *Redox metabolism and longevity relationships in animals and plants* (pp. 99–115). New York and Abingdon: Taylor and Francis.

de Magalhaes, J. P. (2011). The biology of ageing: A primer. In I. Stuart-Hamilton (Ed.), *An Introduction to Gerontology* (pp. 21–47). Cambridge: Cambridge University Press.

de Magalhaes, J. P. (2014a). The scientific quest for lasting youth: Prospects for curing aging. *Rejuvenation Research*, *17*(5), 458–467. Available from https://doi.org/10.1089/rej.2014.1580.

de Magalhaes, J. P. (2014b). Why genes extending lifespan in model organisms have not been consistently associated with human longevity and what it means to translation research. *Cell Cycle*, *13*(17), 2671–2673. Available from https://doi.org/10.4161/15384101.2014.950151.

de Magalhaes, J. P., & Church, G. M. (2007). Analyses of human-chimpanzee orthologous gene pairs to explore evolutionary hypotheses of aging. *Mechanisms of Ageing and Development*, *128*(5–6), 355–364. Available from https://doi.org/10.1016/j.mad.2007.03.004.

de Magalhaes, J. P., Curado, J., & Church, G. M. (2009). Meta-analysis of age-related gene expression profiles identifies common signatures of aging. *Bioinformatics*, *25*(7), 875–881. Available from https://doi.org/10.1093/bioinformatics/btp073.

de Magalhaes, J. P., Finch, C. E., & Janssens, G. (2010). Next-generation sequencing in aging research: Emerging applications, problems, pitfalls and possible solutions. *Ageing Research Reviews*, *9*(3), 315–323. Available from https://doi.org/10.1016/j.arr.2009.10.006.

de Magalhaes, J. P., & Keane, M. (2013). Endless paces of degeneration--applying comparative genomics to study evolution's moulding of longevity. *EMBO Reports*, *14*(8), 661–662. Available from https://doi.org/10.1038/embor.2013.96.

de Magalhaes, J. P., Stevens, M., & Thornton, D. (2017). The business of anti-aging science. *Trends in Biotechnology*, *35*(11), 1062–1073. Available from https://doi.org/10.1016/j.tibtech.2017.07.004.

de Magalhaes, J. P., & Toussaint, O. (2004a). GenAge: A genomic and proteomic network map of human ageing. *FEBS Letters*, *571*(1-3), 243–247.

de Magalhaes, J. P., & Toussaint, O. (2004b). How bioinformatics can help reverse engineer human aging. *Ageing Research Reviews*, *3*(2), 125–141.

de Magalhaes, J. P., & Wang, J. (2019). The fog of genetics: What is known, unknown and unknowable in the genetics of complex traits and diseases. *EMBO Reports*, *20*(11), e48054. Available from https://doi.org/10.15252/embr.201948054.

de Magalhaes, J. P., Wuttke, D., Wood, S. H., Plank, M., & Vora, C. (2012). Genome-environment interactions that modulate aging: Powerful targets for drug discovery. *Pharmacological Reviews*, *64*(1), 88–101. Available from https://doi.org/10.1124/pr.110.004499.

Derveaux, S., Vandesompele, J., & Hellemans, J. (2010). How to do successful gene expression analysis using real-time PCR. *Methods*, *50*(4), 227–230. Available from https://doi.org/10.1016/j.ymeth.2009.11.001.

DiMasi, J. A., Feldman, L., Seckler, A., & Wilson, A. (2010). Trends in risks associated with new drug development: Success rates for investigational drugs. *Clinical Pharmacology and Therapeutics*, *87*(3), 272–277. Available from https://doi.org/10.1038/clpt.2009.295.

Donertas, H. M., Fuentealba Valenzuela, M., Partridge, L., & Thornton, J. M. (2018). Gene expression-based drug repurposing to target aging. *Aging Cell*, *17*(5), e12819. Available from https://doi.org/10.1111/acel.12819.

Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., … Lochovsky, L. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74. Available from https://doi.org/10.1038/nature11247.

Echeverri, C. J., & Perrimon, N. (2006). High-throughput RNAi screening in cultured cells: A user's guide. *Nature Reviews. Genetics*, *7*(5), 373–384.

Edwards, Y. J., Beecham, G. W., Scott, W. K., Khuri, S., Bademci, G., Tekin, D., … Vance, J. M. (2011). Identifying consensus disease pathways in Parkinson's disease using an integrative systems biology approach. *PLoS One*, *6*(2), e16917. Available from https://doi.org/10.1371/journal.pone.0016917.

Enge, M., Arda, H. E., Mignardi, M., Beausang, J., Bottino, R., Kim, S. K., & Quake, S. R. (2017). Single-cell analysis of human pancreas reveals transcriptional signatures of aging and somatic mutation patterns. *Cell*, *171*(2), 321–330, e314. Available from https://doi.org/10.1016/j.cell.2017.09.004.

Erler, J. T., & Linding, R. (2010). Network-based drugs and biomarkers. *The Journal of Pathology*, *220*(2), 290–296. Available from https://doi.org/10.1002/path.2646.

Fabris, F., Magalhaes, J. P., & Freitas, A. A. (2017). A review of supervised machine learning applied to ageing research. *Biogerontology*, *18*(2), 171–188. Available from https://doi.org/10.1007/s10522-017-9683-y.

Fabrizio, P., Hoon, S., Shamalnasab, M., Galbani, A., Wei, M., Giaever, G., … Longo, V. D. (2010). Genome-wide screen in *Saccharomyces cerevisiae* identifies vacuolar protein sorting, autophagy, biosynthetic, and tRNA methylation genes involved in life span regulation. *PLoS Genetics*, *6*(7), e1001024. Available from https://doi.org/10.1371/journal.pgen.1001024.

Fernandes, M., Wan, C., Tacutu, R., Barardo, D., Rajput, A., Wang, J., … de Magalhaes, J. P. (2016). Systematic analysis of the gerontome reveals links between aging and age-related diseases. *Human Molecular Genetics*, *25*(21), 4804–4818. Available from https://doi.org/10.1093/hmg/ddw307.

Ferrarini, L., Bertelli, L., Feala, J., McCulloch, A. D., & Paternostro, G. (2005). A more efficient search strategy for aging genes based on connectivity. *Bioinformatics*, *21*(3), 338–348. Available from https://doi.org/10.1093/bioinformatics/bti004.

Field, A. E., Robertson, N. A., Wang, T., Havas, A., Ideker, T., & Adams, P. D. (2018). DNA methylation clocks in aging: Categories, causes, and consequences. *Molecular Cell*, *71*(6), 882–895. Available from https://doi.org/10.1016/j.molcel.2018.08.008.

Finch, C. E. (1990). *Longevity, senescence, and the genome*. Chicago and London: The University of Chicago Press.

Fortney, K., Kotlyar, M., & Jurisica, I. (2010). Inferring the functions of longevity genes with modular subnetwork biomarkers of *Caenorhabditis elegans* aging. *Genome Biology*, *11*(2), R13. Available from https://doi.org/10.1186/gb-2010-11-2-r13.

Freitas, A. A., Vasieva, O., & de Magalhaes, J. P. (2011). A data mining approach for classifying DNA repair genes into ageing-related or non-ageing-related. *BMC Genomics*, *12*, 27. Available from https://doi.org/10.1186/1471-2164-12-27.

Funakoshi, M., Tsuda, M., Muramatsu, K., Hatsuda, H., Morishita, S., & Aigaki, T. (2011). A gain-of-function screen identifies wdb and lkb1 as lifespan-extending genes in Drosophila. *Biochemical and Biophysical Research Communications*, *405*(4), 667–672. Available from https://doi.org/10.1016/j.bbrc.2011.01.090.

Fushan, A. A., Turanov, A. A., Lee, S. G., Kim, E. B., Lobanov, A. V., Yim, S. H., … Gladyshev, V. N. (2015). Gene expression defines natural changes in mammalian lifespan. *Aging Cell*, *14*(3), 352–365. Available from https://doi.org/10.1111/acel.12283.

Gao, A. W., Smith, R. L., van Weeghel, M., Kamble, R., Janssens, G. E., & Houtkooper, R. H. (2018). Identification of key pathways and metabolic fingerprints of longevity in *C. elegans*. *Experimental Gerontology*, *113*, 128–140. Available from https://doi.org/10.1016/j.exger.2018.10.003.

Glass, D., Vinuela, A., Davies, M. N., Ramasamy, A., Parts, L., Knowles, D., … Spector, T. D. (2013). Gene expression changes with age in skin, adipose tissue, blood and brain. *Genome Biology*, *14*(7), R75. Available from https://doi.org/10.1186/gb-2013-14-7-r75.

Goh, K. I., & Choi, I. G. (2012). Exploring the human diseasome: The human disease network. *Brief Function in Genomics*, *11*(6), 533–542. Available from https://doi.org/10.1093/bfgp/els032.

Goh, K. I., Cusick, M. E., Valle, D., Childs, B., Vidal, M., & Barabasi, A. L. (2007). The human disease network. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(21), 8685–8690. Available from https://doi.org/10.1073/pnas.0701361104.

Halu, A., De Domenico, M., Arenas, A., & Sharma, A. (2019). The multiplex network of human diseases. *NPJ Systems Biology and Applications*, *5*, 15. Available from https://doi.org/10.1038/s41540-019-0092-5.

Hamilton, B., Dong, Y., Shindo, M., Liu, W., Odell, I., Ruvkun, G., & Lee, S. S. (2005). A systematic RNAi screen for longevity genes in C. elegans. *Genes & Development*, *19*(13), 1544–1555. Available from https://doi.org/10.1101/gad.1308205.

Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., ... Guo, G. (2018). Mapping the mouse cell atlas by microwell-seq. *Cell*, *172*(5), 1091–1107, e1017. Available from https://doi.org/10.1016/j.cell.2018.02.001.

Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., ... Zhang, K. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular Cell*, *49*(2), 359–367. Available from https://doi.org/10.1016/j.molcel.2012.10.016.

Hansen, M., Hsu, A. L., Dillin, A., & Kenyon, C. (2005). New genes tied to endocrine, metabolic, and dietary regulation of lifespan from a *Caenorhabditis elegans* genomic RNAi screen. *PLoS Genetics*, *1*(1), 119–128. Available from https://doi.org/10.1371/journal.pgen.0010017.

Hansen, M., Taubert, S., Crawford, D., Libina, N., Lee, S. J., & Kenyon, C. (2007). Lifespan extension by conditions that inhibit translation in *Caenorhabditis elegans*. *Aging Cell*, *6*(1), 95–110. Available from https://doi.org/10.1111/j.1474-9726.2006.00267.x.

He, X., Chang, S., Zhang, J., Zhao, Q., Xiang, H., Kusonmano, K., ... Wang, J. (2008). MethyCancer: The database of human DNA methylation and cancer. *Nucleic Acids Res*, *36*(Database issue), D836–D841. Available from https://doi.org/10.1093/nar/gkm730.

Hoffman, J. M., Lyu, Y., Pletcher, S. D., & Promislow, D. E. L. (2017). Proteomics and metabolomics in ageing research: From biomarkers to systems biology. *Essays in Biochemistry*, *61*(3), 379–388. Available from https://doi.org/10.1042/EBC20160083.

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, *14*(10), R115. Available from https://doi.org/10.1186/gb-2013-14-10-r115.

Horvath, S., & Raj, K. (2018). DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews. Genetics*, *19*(6), 371–384. Available from https://doi.org/10.1038/s41576-018-0004-3.

Houtkooper, R. H., Argmann, C., Houten, S. M., Canto, C., Jeninga, E. H., Andreux, P. A., ... Auwerx, J. (2011). The metabolic footprint of aging in mice. *Scientific Reports*, *1*, 134. Available from https://doi.org/10.1038/srep00134.

Huang da, W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. Available from https://doi.org/10.1038/nprot.2008.211.

Huang, T., Zhang, J., Xu, Z. P., Hu, L. L., Chen, L., Shao, J. L., ... Chou, K. C. (2012). Deciphering the effects of gene deletion on yeast longevity using network and machine learning approaches. *Biochimie*, *94*(4), 1017–1025. Available from https://doi.org/10.1016/j.biochi.2011.12.024.

Huang, Z., Shi, J., Gao, Y., Cui, C., Zhang, S., Li, J., ... Cui, Q. (2019). HMDD v3.0: A database for experimentally supported human microRNA-disease associations. *Nucleic Acids Research*, *47*(D1), D1013–D1017. Available from https://doi.org/10.1093/nar/gky1010.

Huhne, R., Thalheim, T., & Suhnel, J. (2014). AgeFactDB--the JenAge Ageing Factor Database--towards data integration in ageing research. *Nucleic Acids Research*, *42*(1), D892–D896. Available from https://doi.org/10.1093/nar/gkt1073.

Ideker, T., & Sharan, R. (2008). Protein networks in disease. *Genome Research*, *18*(4), 644–652. Available from https://doi.org/10.1101/gr.071852.107.

Ingolia, N. T., Ghaemmaghami, S., Newman, J. R., & Weissman, J. S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, *324*(5924), 218–223. Available from https://doi.org/10.1126/science.1168978.

Iorio, F., Bosotti, R., Scacheri, E., Belcastro, V., Mithbaokar, P., Ferriero, R., ... di Bernardo, D. (2010). Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(33), 14621–14626. Available from https://doi.org/10.1073/pnas.1000138107.

Jobson, R. W., Nabholz, B., & Galtier, N. (2010). An evolutionary genome scan for longevity-related natural selection in mammals. *Molecular Biology and Evolution*, *27*(4), 840–847. Available from https://doi.org/10.1093/molbev/msp293.

Johnson, A. A., Akman, K., Calimport, S. R., Wuttke, D., Stolzing, A., & de Magalhaes, J. P. (2012). The role of DNA methylation in aging, rejuvenation, and age-related disease. *Rejuvenation Research*, *15*(5), 483–494. Available from https://doi.org/10.1089/rej.2012.1324.

Johnson, T. E. (2002). A personal retrospective on the genetics of aging. *Biogerontology*, *3*(1–2), 7–12.

Keane, M., Craig, T., Alfoldi, J., Berlin, A. M., Johnson, J., Seluanov, A., ... de Magalhaes, J. P. (2014). The Naked Mole Rat Genome Resource: Facilitating analyses of cancer and longevity-related adaptations. *Bioinformatics*, *30*, 3558–3560. Available from https://doi.org/10.1093/bioinformatics/btu579.

Keane, M., Semeiks, J., Webb, A. E., Li, Y. I., Quesada, V., Craig, T., ... de Magalhaes, J. P. (2015). Insights into the evolution of longevity from the bowhead whale genome. *Cell Reports*, *10*(1), 112–122. Available from https://doi.org/10.1016/j.celrep.2014.12.008.

Kennedy, S. R., Salk, J. J., Schmitt, M. W., & Loeb, L. A. (2013). Ultra-sensitive sequencing reveals an age-related increase in somatic mitochondrial mutations that are inconsistent with oxidative damage. *PLoS Genetics*, *9*(9), e1003794. Available from https://doi.org/10.1371/journal.pgen.1003794.

Kim, E. B., Fang, X., Fushan, A. A., Huang, Z., Lobanov, A. V., Han, L., ... Gladyshev, V. N. (2011). Genome sequencing reveals insights into physiology and longevity of the naked mole rat. *Nature*, *479*(7372), 223–227. Available from https://doi.org/10.1038/nature10533.

Kimmel, J. C., Penland, L., Rubinstein, N. D., Hendrickson, D. G., Kelley, D. R., & Rosenthal, A. Z. (2019). Murine single-cell RNA-seq reveals cell-identity- and tissue-specific trajectories of aging. *Genome Res*, *29*(12), 2088–2103. https://doi.org/10.1101/gr.253880.119.

Kirkwood, T. B., & Kowald, A. (1997). Network theory of aging. *Experimental Gerontology*, *32*(4-5), 395–399.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, *2*(3), 203–271. Available from https://doi.org/10.1093/comnet/cnu016.

Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., ... Golub, T. R. (2006). The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science*, *313*(5795), 1929–1935. Available from https://doi.org/10.1126/science.1132939.

Lawton, K. A., Berger, A., Mitchell, M., Milgram, K. E., Evans, A. M., Guo, L., ... Milburn, M. V. (2008). Analysis of the adult human plasma metabolome. *Pharmacogenomics*, *9*(4), 383–397. Available from https://doi.org/10.2217/14622416.9.4.383.

Lee, C. K., Klopp, R. G., Weindruch, R., & Prolla, T. A. (1999). Gene expression profile of aging and its retardation by caloric restriction. *Science*, *285*(5432), 1390–1393.

Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., ... Horvath, S. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging*, *10*(4), 573–591. Available from https://doi.org/10.18632/aging.101414.

Li, J. Q., Duan, D. D., Zhang, J. Q., Zhou, Y. Z., Qin, X. M., Du, G. H., & Gao, L. (2019). Bioinformatic prediction of critical genes and pathways involved in longevity in *Drosophila melanogaster*. *Molecular Genetics and Genomics: MGG*, *294*(6), 1463–1475. Available from https://doi.org/10.1007/s00438-019-01589-1.

Li, Y., & de Magalhaes, J. P. (2013). Accelerated protein evolution analysis reveals genes and pathways associated with the

evolution of mammalian longevity. *Age*, *35*(2), 301–314. Available from https://doi.org/10.1007/s11357-011-9361-y.

Li, Y. H., Dong, M. Q., & Guo, Z. (2010). Systematic analysis and prediction of longevity genes in *Caenorhabditis elegans*. *Mechanisms of Ageing and Development*, *131*(11–12), 700–709. Available from https://doi.org/10.1016/j.mad.2010.10.001.

Liu, Y., Kern, J. T., Walker, J. R., Johnson, J. A., Schultz, P. G., & Luesch, H. (2007). A genomic screen for activators of the antioxidant response element. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(12), 5205–5210. Available from https://doi.org/10.1073/pnas.0700898104.

Lopez-Otin, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013). The hallmarks of aging. *Cell*, *153*(6), 1194–1217. Available from https://doi.org/10.1016/j.cell.2013.05.039.

Lorenz, D. R., Cantor, C. R., & Collins, J. J. (2009). A network biology approach to aging in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(4), 1145–1150. Available from https://doi.org/10.1073/pnas.0812551106.

Luecken, M. D., & Theis, F. J. (2019). Current best practices in single-cell RNA-seq analysis: A tutorial. *Molecular Systems Biology*, *15*(6), e8746. Available from https://doi.org/10.15252/msb.20188746.

Macarron, R., Banks, M. N., Bojanic, D., Burns, D. J., Cirovic, D. A., Garyantes, T., … Sittampalam, G. S. (2011). Impact of high-throughput screening in biomedical research. *Nature Reviews. Drug Discovery*, *10*(3), 188–195. Available from https://doi.org/10.1038/nrd3368.

Managbanag, J. R., Witten, T. M., Bonchev, D., Fox, L. A., Tsuchiya, M., Kennedy, B. K., & Kaeberlein, M. (2008). Shortest-path network analysis is a useful approach toward identifying genetic determinants of longevity. *PLoS One*, *3*(11), e3802. Available from https://doi.org/10.1371/journal.pone.0003802.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., … Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature*, *461*(7265), 747–753. Available from https://doi.org/10.1038/nature08494.

Matecic, M., Smith, D. L., Pan, X., Maqani, N., Bekiranov, S., Boeke, J. D., & Smith, J. S. (2010). A microarray-based genetic screen for yeast chronological aging factors. *PLoS Genetics*, *6*(4), e1000921. Available from https://doi.org/10.1371/journal.pgen.1000921.

McCormick, M. A., & Kennedy, B. K. (2012). Genome-scale studies of aging: Challenges and opportunities. *Current Genomics*, *13*(7), 500–507. Available from https://doi.org/10.2174/138920212803251454.

Menni, C., Kastenmuller, G., Petersen, A. K., Bell, J. T., Psatha, M., Tsai, P. C., … Valdes, A. M. (2013). Metabolomic markers reveal novel pathways of ageing and early development in human populations. *International Journal of Epidemiology*, *42*(4), 1111–1119. Available from https://doi.org/10.1093/ije/dyt094.

Midic, U., Oldfield, C. J., Dunker, A. K., Obradovic, Z., & Uversky, V. N. (2009). Protein disorder in the human diseasome: Unfoldomics of human genetic diseases. *BMC Genomics*, *10*(Suppl. 1), S12. Available from https://doi.org/10.1186/1471-2164-10-S1-S12.

Miller, J. A., Oldham, M. C., & Geschwind, D. H. (2008). A systems level analysis of transcriptional changes in Alzheimer's disease and normal aging. *The Journal of Neuroscience*, *28*(6), 1410–1420. Available from https://doi.org/10.1523/JNEUROSCI.4098-07.2008.

Moffat, J., & Sabatini, D. M. (2006). Building mammalian signalling pathways with RNAi screens. *Nature Reviews. Molecular Cell Biology*, *7*(3), 177–187. Available from https://doi.org/10.1038/nrm1860.

Mohr, S., Bakal, C., & Perrimon, N. (2010). Genomic screening with RNAi: Results and challenges. *Annual Review of Biochemistry*, *79*, 37–64. Available from https://doi.org/10.1146/annurev-biochem-060408-092949.

Montoliu, I., Scherer, M., Beguelin, F., DaSilva, L., Mari, D., Salvioli, S., … Collino, S. (2014). Serum profiling of healthy aging identifies phospho- and sphingolipid species as markers of human longevity. *Aging*, *6*(1), 9–25. Available from https://doi.org/10.18632/aging.100630.

Moreau, Y., & Tranchevent, L. C. (2012). Computational tools for prioritizing candidate genes: Boosting disease gene discovery. *Nature Reviews of Genetics*, *13*(8), 523–536. Available from https://doi.org/10.1038/nrg3253.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, *5*(7), 621–628. Available from https://doi.org/10.1038/nmeth.1226.

Muntane, G., Farre, X., Rodriguez, J. A., Pegueroles, C., Hughes, D. A., de Magalhaes, J. P., … Navarro, A. (2018). Biological processes modulating longevity across primates: A phylogenetic genome-phenome analysis. *Molecular Biology and Evolution*, *35*(8), 1990–2004. Available from https://doi.org/10.1093/molbev/msy105.

Nagaoka-Yasuda, R., Matsuo, N., Perkins, B., Limbaeck-Stokin, K., & Mayford, M. (2007). An RNAi-based genetic screen for oxidative stress resistance reveals retinol saturase as a mediator of stress resistance. *Free Radical Biology & Medicine*, *43*(5), 781–788. Available from https://doi.org/10.1016/j.freeradbiomed.2007.05.008.

Oikonomou, P., & Cluzel, P. (2006). Effects of topology on network evolution. *Nature Physics*, *2*, 532–536.

*OMIM. (2014). Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine*, Johns Hopkins University (Baltimore, MD), www.omim.org.

Ongenaert, M., Van Neste, L., De Meyer, T., Menschaert, G., Bekaert, S., & Van Criekinge, W. (2008). PubMeth: A cancer methylation database combining text-mining and expert annotation. *Nucleic Acids Research*, *36*(Database issue), D842–D846. Available from https://doi.org/10.1093/nar/gkm788.

Paik, D., Jang, Y. G., Lee, Y. E., Lee, Y. N., Yamamoto, R., Gee, H. Y., … Park, J. J. (2012). Misexpression screen delineates novel genes controlling *Drosophila* lifespan. *Mechanisms of Ageing and Development*, *133*(5), 234–245. Available from https://doi.org/10.1016/j.mad.2012.02.001.

Pan, F., Chiu, C. H., Pulapura, S., Mehan, M. R., Nunez-Iglesias, J., Zhang, K., … Zhou, X. J. (2007). Gene Aging Nexus: A web database and data mining platform for microarray data on aging. *Nucleic Acids Research*, *35*(Database issue), D756–D759. Available from https://doi.org/10.1093/nar/gkl798.

Petrascheck, M., Ye, X., & Buck, L. B. (2007). An antidepressant that extends lifespan in adult *Caenorhabditis elegans*. *Nature*, *450*(7169), 553–556. Available from https://doi.org/10.1038/nature05991.

Pisco, A. O., Schaum, N., McGeever, A., Karkanias, J., Neff, N. F., Darmanis, S., … Quake, S. R. (2020). A single cell transcriptomic atlas characterizes aging tissues in the mouse. *Nature*, *583*, 590–595.

Plank, M., Hu, G., Silva, A. S., Wood, S. H., Hesketh, E. E., Janssens, G., … Church, G. M. (2013). An analysis and validation pipeline for large-scale RNAi-based screens. *Scientific Reports*, *3*, 1076. Available from https://doi.org/10.1038/srep01076.

Plank, M., Wuttke, D., van Dam, S., Clarke, S. A., & de Magalhaes, J. P. (2012). A meta-analysis of caloric restriction gene expression profiles to infer common signatures and regulatory mechanisms. *Molecular Biosystems*, *8*(4), 1339–1349. Available from https://doi.org/10.1039/c2mb05255e.

Potter, P. K., Bowl, M. R., Jeyarajan, P., Wisby, L., Blease, A., Goldsworthy, M. E., … Brown, S. D. (2016). Novel gene function revealed by mouse mutagenesis screens for models of age-related disease. *Nature Communications*, *7*, 12444. Available from https://doi.org/10.1038/ncomms12444.

Promislow, D. E. (2004). Protein networks, pleiotropy and the evolution of senescence. *Proceedings of the Royal Society. Biological Sciences*, *271*(1545), 1225–1234. Available from https://doi.org/10.1098/rspb.2004.2732.

Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., & Human Cell Atlas Meeting, P. (2017). The human cell atlas. *Elife*, 6, e27041. Available from https://doi.org/10.7554/eLife.27041.

Saito, R., Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., Lotia, S., … Ideker, T. (2012). A travel guide to *Cytoscape plugins*. *Nature Methods*, *9*(11), 1069–1076. Available from https://doi.org/10.1038/nmeth.2212.

Samuelson, A. V., Carr, C. E., & Ruvkun, G. (2007). Gene activities that mediate increased life span of C. elegans insulin-like signaling mutants. *Genes & Development*, *21*(22), 2976–2994. Available from https://doi.org/10.1101/gad.1588907.

Schachter, F., Faure-Delanef, L., Guenot, F., Rouger, H., Froguel, P., Lesueur-Ginot, L., & Cohen, D. (1994). Genetic associations with human longevity at the APOE and ACE loci. *Nature Genetics*, *6*(1), 29–32. Available from https://doi.org/10.1038/ng0194-29.

Semeiks, J., & Grishin, N. V. (2012). A method to find longevity-selected positions in the mammalian proteome. *PLoS One*, *7*(6), e38595. Available from https://doi.org/10.1371/journal.pone.0038595.

Simko, G. I., Gyurko, D., Veres, D. V., Nanasi, T., & Csermely, P. (2009). Network strategies to understand the aging process and help age-related drug design. *Genome Medicine*, *1*(9), 90. Available from https://doi.org/10.1186/gm90.

Sirota, M., Dudley, J. T., Kim, J., Chiang, A. P., Morgan, A. A., Sweet-Cordero, A., … Butte, A. J. (2011). Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Science Translational Medicine*, *3*(96), 96ra77. Available from https://doi.org/10.1126/scitranslmed.3001318.

Smita, S., Lange, F., Wolkenhauer, O., & Kohling, R. (2016). Deciphering hallmark processes of aging from interaction networks. *Biochimica et Biophysica Acta*, *1860*(11 Pt B), 2706–2715. Available from https://doi.org/10.1016/j.bbagen.2016.07.017.

Smith, E. D., Tsuchiya, M., Fox, L. A., Dang, N., Hu, D., Kerr, E. O., … Kennedy, B. K. (2008). Quantitative evidence for conserved longevity pathways between divergent eukaryotic species. *Genome Research*, *18*(4), 564–570. Available from https://doi.org/10.1101/gr.074724.107.

Soltow, Q. A., Jones, D. P., & Promislow, D. E. (2010). A network perspective on metabolism and aging. *Integrative and Comparative Biology*, *50*(5), 844–854. Available from https://doi.org/10.1093/icb/icq094.

Sondka, Z., Bamford, S., Cole, C. G., Ward, S. A., Dunham, I., & Forbes, S. A. (2018). The COSMIC cancer gene census: Describing genetic dysfunction across all human cancers. *Nature Reviews in Cancer*, *18*(11), 696–705. Available from https://doi.org/10.1038/s41568-018-0060-1.

Stark, C., Breitkreutz, B. J., Reguly, T., Boucher, L., Breitkreutz, A., & Tyers, M. (2006). BioGRID: A general repository for interaction datasets. *Nucleic Acids Research*, *34*(Database issue), D535–D539. Available from https://doi.org/10.1093/nar/gkj109.

Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(16), 9440–9445. Available from https://doi.org/10.1073/pnas.1530509100.

Stroustrup, N., Ulmschneider, B. E., Nash, Z. M., Lopez-Moyado, I. F., Apfeld, J., & Fontana, W. (2013). The *Caenorhabditis elegans* lifespan machine. *Nature Methods*, *10*(7), 665–670. Available from https://doi.org/10.1038/nmeth.2475.

Stuart, J. A., Liang, P., Luo, X., Page, M. M., Gallagher, E. J., Christoff, C. A., & Robb, E. L. (2013). A comparative cellular and molecular biology of longevity database. *Age*, *35*(5), 1937–1947. Available from https://doi.org/10.1007/s11357-012-9458-y.

Tacutu, R., Budovsky, A., Wolfson, M., & Fraifeld, V. E. (2010). MicroRNA-regulated protein-protein interaction networks: How could they help in searching for pro-longevity targets? *Rejuvenation Research*, *13*(2–3), 373–377. Available from https://doi.org/10.1089/rej.2009.0980.

Tacutu, R., Budovsky, A., Yanai, H., & Fraifeld, V. E. (2011). Molecular links between cellular senescence, longevity and age-related diseases - a systems biology perspective. *Aging*, *3*(12), 1178–1191.

Tacutu, R., Shore, D. E., Budovsky, A., de Magalhaes, J. P., Ruvkun, G., Fraifeld, V. E., & Curran, S. P. (2012). Prediction of C. elegans longevity genes by human and worm longevity networks. *PLoS One*, *7*(10), e48282. Available from https://doi.org/10.1371/journal.pone.0048282.

Tacutu, R., Thornton, D., Johnson, E., Budovsky, A., Barardo, D., Craig, T., … de Magalhaes, J. P. (2018). Human ageing genomic resources: New and updated databases. *Nucleic Acids Research*, *46*(D1), D1083–D1090. Available from https://doi.org/10.1093/nar/gkx1042.

Tanaka, T., Biancotto, A., Moaddel, R., Moore, A. Z., Gonzalez-Freire, M., Aon, M. A., … Ferrucci, L. (2018). Plasma proteomic signature of age in healthy humans. *Aging Cell*, *17*(5), e12799. Available from https://doi.org/10.1111/acel.12799.

Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., … Forbes, S. A. (2019). COSMIC: The catalogue of somatic mutations in cancer. *Nucleic Acids Research*, *47*(D1), D941–D947. Available from https://doi.org/10.1093/nar/gky1015.

Tomas-Loba, A., Bernardes de Jesus, B., Mato, J. M., & Blasco, M. A. (2013). A metabolic signature predicts biological age in mice. *Aging Cell*, *12*(1), 93–101. Available from https://doi.org/10.1111/acel.12025.

Toren, D., Kulaga, A., Jethva, M., Rubin, E., Snezhkina, A. V., Kudryavtseva, A. V., … Fraifeld, V. E. (2020). Gray whale transcriptome reveals longevity adaptations associated with DNA repair and ubiquitination. Aging Cell, 19, e13158. https://doi.org/10.1111/acel.13158.

Tsuchiya, T., Dhahbi, J. M., Cui, X., Mote, P. L., Bartke, A., & Spindler, S. R. (2004). Additive regulation of hepatic gene expression by dwarfism and caloric restriction. *Physiological Genomics*, *17*(3), 307–315. Available from https://doi.org/10.1152/physiolgenomics.00039.2004.

Tyshkovskiy, A., Bozaykut, P., Borodinova, A. A., Gerashchenko, M. V., Ables, G. P., Garratt, M., … Gladyshev, V. N. (2019). Identification and application of gene expression signatures associated with lifespan extension. *Cell Metabolism*, *30*(3), 573–593, e578. Available from https://doi.org/10.1016/j.cmet.2019.06.018.

Valdes, A. M., Glass, D., & Spector, T. D. (2013). Omics technologies and the study of human ageing. *Nature Reviews. Genetics*, *14*(9), 601–607. Available from https://doi.org/10.1038/nrg3553.

Vignoli, A., Tenori, L., Luchinat, C., & Saccenti, E. (2018). Age and sex effects on plasma metabolite association networks in healthy subjects. *Journal of Proteome Research*, *17*(1), 97–107. Available from https://doi.org/10.1021/acs.jproteome.7b00404.

Wagner, A. (2000). Robustness against mutations in genetic networks of yeast. *Nature Genetics*, *24*(4), 355–361. Available from https://doi.org/10.1038/74174.

Wang, J., Zhang, S., Wang, Y., Chen, L., & Zhang, X. S. (2009). Disease-aging network reveals significant roles of aging genes in connecting genetic diseases. *PLoS Computational Biology*, *5*(9), e1000521. Available from https://doi.org/10.1371/journal.pcbi.1000521.

Wang, T., Li, B., Nelson, C. E., & Nabavi, S. (2019). Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data. *BMC Bioinformatics*, *20*(1), 40. Available from https://doi.org/10.1186/s12859-019-2599-6.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews in Genetics*, *10*(1), 57–63. Available from https://doi.org/10.1038/nrg2484.

West, J., Beck, S., Wang, X., & Teschendorff, A. E. (2013). An integrative network algorithm identifies age-associated differential methylation interactome hotspots targeting stem-cell differentiation pathways. *Scientific Reports*, *3*, 1630. Available from https://doi.org/10.1038/srep01630.

West, J., Widschwendter, M., & Teschendorff, A. E. (2013). Distinctive topology of age-associated epigenetic drift in the human interactome. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(35), 14138–14143. Available from https://doi.org/10.1073/pnas.1307242110.

Witten, T. M., & Bonchev, D. (2007). Predicting aging/longevity-related genes in the nematode *Caenorhabditis elegans*. *Chemistry & Biodiversity*, *4*(11), 2639–2655. Available from https://doi.org/10.1002/cbdv.200790216.

Wolfson, M., Budovsky, A., Tacutu, R., & Fraifeld, V. (2009). The signaling hubs at the crossroad of longevity and age-related disease networks. *The International Journal of Biochemistry & Cell Biology*, *41*(3), 516–520. Available from https://doi.org/10.1016/j.biocel.2008.08.026.

Wood, S. H., Craig, T., Li, Y., Merry, B., & de Magalhaes, J. P. (2013). Whole transcriptome sequencing of the aging rat brain reveals dynamic RNA changes in the dark matter of the genome. *Age*, *35*, 763–776. Available from https://doi.org/10.1007/s11357-012-9410-1.

Wood, S. H., van Dam, S., Craig, T., Tacutu, R., O'Toole, A., Merry, B. J., & de Magalhaes, J. P. (2015). Transcriptome analysis in calorie-restricted rats implicates epigenetic and post-translational mechanisms in neuroprotection and aging. *Genome Biology*, *16*, 285. Available from https://doi.org/10.1186/s13059-015-0847-2.

Wuttke, D., Connor, R., Vora, C., Craig, T., Li, Y., Wood, S., … de Magalhaes, J. P. (2012). Dissecting the gene network of dietary restriction to identify evolutionarily conserved pathways and new functional genes. *PLoS Genetics*, *8*(8), e1002834. Available from https://doi.org/10.1371/journal.pgen.1002834.

Ximerakis, M., Lipnick, S. L., Innes, B. T., Simmons, S. K., Adiconis, X., Dionne, D., … Rubin, L. L. (2019). Single-cell transcriptomic profiling of the aging mouse brain. *Nature Neuroscience*, *22*(10), 1696–1708. Available from https://doi.org/10.1038/s41593-019-0491-3.

Xue, H., Xian, B., Dong, D., Xia, K., Zhu, S., Zhang, Z., … Han, J. D. (2007). A modular network model of aging. *Molecular Systems Biology*, *3*, 147. Available from https://doi.org/10.1038/msb4100189.

Yanai, H., Budovsky, A., Barzilay, T., Tacutu, R., & Fraifeld, V. E. (2017). Wide-scale comparative analysis of longevity genes and interventions. *Aging Cell*, *16*(6), 1267–1275. Available from https://doi.org/10.1111/acel.12659.

Yang, J., Huang, T., Petralia, F., Long, Q., Zhang, B., Argmann, C., … Tu, Z. (2015). Synchronized age-related gene expression changes across multiple tissues in human and the link to complex diseases. *Scientific Reports*, *5*, 15145. Available from https://doi.org/10.1038/srep15145.

Yang, S., Liu, T., Li, S., Zhang, X., Ding, Q., Que, H., … Liu, S. (2008). Comparative proteomic analysis of brains of naturally aging mice. *Neuroscience*, *154*(3), 1107–1120. Available from https://doi.org/10.1016/j.neuroscience.2008.04.012.

Ye, X., Linton, J. M., Schork, N. J., Buck, L. B., & Petrascheck, M. (2014). A pharmacological network for lifespan extension in *Caenorhabditis elegans*. *Aging Cell*, *13*(2), 206–215. Available from https://doi.org/10.1111/acel.12163.

Yildirim, M. A., Goh, K. I., Cusick, M. E., Barabasi, A. L., & Vidal, M. (2007). Drug-target network. *Nature Biotechnology*, *25*(10), 1119–1126. Available from https://doi.org/10.1038/nbt1338.

Yu, Z., Zhai, G., Singmann, P., He, Y., Xu, T., Prehn, C., … Wang-Sattler, R. (2012). Human serum metabolic profiles are age dependent. *Aging Cell*, *11*(6), 960–967. Available from https://doi.org/10.1111/j.1474-9726.2012.00865.x.

Zahn, J. M., Poosala, S., Owen, A. B., Ingram, D. K., Lustig, A., Carter, A., … Becker, K. G. (2007). AGEMAP: A gene expression database for aging in mice. *PLoS Genetics*, *3*(11), e201. Available from https://doi.org/10.1371/journal.pgen.0030201.

Zhang, G., Cowled, C., Shi, Z., Huang, Z., Bishop-Lilly, K. A., Fang, X., … Wang, J. (2013a). Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science*, *339*(6118), 456–460. Available from https://doi.org/10.1126/science.1230835.

Zhang, B., Gaiteri, C., Bodea, L. G., Wang, Z., McElwee, J., Podtelezhnikov, A. A., … Emilsson, V. (2013b). Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell*, *153*(3), 707–720. Available from https://doi.org/10.1016/j.cell.2013.03.030.

Zhang, Y., De, S., Garner, J. R., Smith, K., Wang, S. A., & Becker, K. G. (2010). Systematic analysis, comparison, and integration of disease based human genetic association data and mouse genetic phenotypic information. *BMC Medical Genomics*, *3*, 1. Available from https://doi.org/10.1186/1755-8794-3-1.

Zhao, M., Kim, P., Mitra, R., Zhao, J., & Zhao, Z. (2016). TSGene 2.0: An updated literature-based knowledgebase for tumor suppressor genes. *Nucleic Acids Research*, *44*(D1), D1023–D1031. Available from https://doi.org/10.1093/nar/gkv1268.

Zhou, B., Yang, L., Li, S., Huang, J., Chen, H., Hou, L., … Han, J. D. (2012). Midlife gene expressions identify modulators of aging through dietary interventions. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(19), E1201–E1209. Available from https://doi.org/10.1073/pnas.1119304109.

Ziehm, M., Kaur, S., Ivanov, D. K., Ballester, P. J., Marcus, D., Partridge, L., & Thornton, J. M. (2017). Drug repurposing for aging research using model organisms. *Aging Cell*, *16*(5), 1006–1015. Available from https://doi.org/10.1111/acel.12626.