6

Ageing research in the post-genome era: New technologies for an old problem

João Pedro de Magalhães

1 Introduction

'New horizons are never discovered by following old roads.' – John O'Reilly

The sequencing of a growing number of genomes, and of the human genome in particular, has paved the way for remarkable technological innovations that are revolutionizing research across the life sciences. Large-scale sequencing and emerging techniques and approaches, like microarrays and systems biology, are changing the way researchers study life and its intrinsic complexity. Cells and the molecules that compose them, in particular proteins programmed from the genome to give rise to complex interactions and functions, can now be studied in a high-throughput fashion. Genomics and its derivative technologies promise to transform biomedical research and be the preeminent area of scientific discovery during the 21st century.

New technologies are particularly useful in scientific fields that have already been the subject of considerable research, and thus for which substantial observations exist, but that remain largely unsolved (de Magalhães and Toussaint, 2004a). Biogerontology is one of such fields and the post-genome era promises to be a time of unprecedented breakthroughs in our understanding of ageing. The goal of this chapter is to review some of the key technological and methodological advances made possible by the recent sequencing of genomes and summarize their relevance to the study of complex traits like longevity, ageing and age-related diseases.

2 The genomic revolution

Clearly, genomic technologies are changing the way research on the science of ageing is conducted (Kaeberlein, 2004; Lim and Ng, 2007; Raghothama *et al.*, 2005). One fitting example of a breakthrough made possible by having the sequence of the human genome was the association of mutations in the *LMNA* gene with Hutchinson-Gilford progeria syndrome. Succinctly, using a combination of genetic techniques, Eriksson *et al.* (2003) identified a candidate interval of 4.82 Mb on chromosome 1. Though there

are roughly 80 known genes in this region, the researchers immediately focused on *LMNA* because mutations in this gene had been previously associated with various disorders. Sequencing of *LMNA* revealed a heterozygous base substitution in 18 of 23 patients (Eriksson *et al.*, 2003).

The work on *LMNA* is an excellent example of how useful it is for genetic research to have the sequence of the human genome, yet it is only the tip of the iceberg. As will be further detailed ahead, having the genome sequence for a given organism allows genome-wide studies and genetic screens. One example is the powerful technology of RNA interference (RNAi), which can now be used to manipulate gene expression on a whole-genome level (Hannon, 2002). In fact, with the availability of genome sequences for several model organisms, RNAi technology can be used to screen thousands of genes for their effects on a given phenotype. Not surprisingly, large-scale RNAi longevity studies have already been conducted in *C. elegans* and have allowed the identification of new genes and pathways that influence lifespan (Hamilton *et al.*, 2005; Lee *et al.*, 2003).

In addition to facilitating genetic research, the sequencing of genomes has allowed the development of several large-scale techniques for the quantification of the basic molecular components of living organisms. Having the genome sequence of a given organism opens the door to the development of microarrays used in gene expression studies (Brown and Botstein, 1999) and modern proteomics like mass spectrometry (Aebersold and Mann, 2003; Dierick *et al.*, 2002). These technologies allow researchers to generate massive amounts of data and help understand how the molecular components of cells, and the genes that give rise to them, produce the complexity of life. Some of their uses in the context of research on ageing, a process that can be seen as the corruption of life, are discussed below.

2.1 *Functional genomics*

The goal of functional genomics is to understand the function of genes, in particular the function of the roughly 25 000 genes that make up the human genome. Functional genomics focuses mostly on analyses of the expression of genes and/or proteins and their interactions (i.e. protein-protein interactions and protein-DNA interactions).

Clearly, one of the most powerful genomic technologies is the microarray, which is normally used to study gene expression. Microarrays allow researchers to quantify the mRNA levels of thousands of genes. Typically, gene expression studies of ageing aim to identify genes whose expression levels change with age. Briefly, many ageing tissues have been studied in animals and, although ageing changes in gene expression are often tissue-specific, a few common patterns have emerged. For instance, some pathways appear to change with age in different model organisms, such as genes linked to cellular detoxification (McElwee *et al.*, 2007). A complementary approach is to identify genes differentially expressed in long-lived strains (e.g. in long-lived knock-outs) or conditions (e.g. under caloric restriction). In an interesting work, Murphy *et al.* (2003) used microarray analysis to identify genes differentially expressed in long-lived *C. elegans* mutant strains, many of which they later showed to influence ageing.

One problem with gene expression studies of ageing is that genes differentially expressed may not necessarily be related to the ageing process but rather to specific age-related diseases. They may also reflect programmed responses to the loss of func-

tion and tissue homeostasis that normally represents ageing, rather than an underlying mechanism of ageing (de Magalhães and Toussaint, 2004a). As such, interpreting ageing microarray data is not as straightforward as when studying other processes, like the cell cycle.

Microarrays have numerous pharmacological applications, including for age-related diseases. Briefly, one example comes from the Connectivity Map, a collection of gene-expression profiles obtained from cultured human cells treated with 164 small molecules. The Connectivity Map can be used to find connections among small molecules, diseases and physiological processes and has been used to identify candidate small molecules that might be used to help treat Alzheimer's disease (Lamb *et al.*, 2006). Another area in biogerontology in which microarrays may be valuable is in finding biomarkers of ageing. Although the search has so far been largely fruitless and even somewhat controversial, being able to quantify how aged a given organism or patient is would be extremely useful to basic and applied research (Miller, 2001). Gene expression studies – alone or integrated with other parameters – may help identify biomarkers of ageing that allow us to quantify ageing (Kriete, 2006). For example, genes differentially expressed with age whose changes are attenuated in long-lived strains are potentially good candidates. Interestingly, it has been proposed that a cheaper and quicker approach to screen candidate life-extending therapies in mice is to employ gene expression biomarkers at earlier ages rather than conduct full longevity studies (Spindler and Mote, 2007).

Because of their ability to generate large amounts of data in a relatively short period of time, gene expression studies have thus far been more used than protein studies, even though protein modifications can alter protein function and go undetected by gene expression studies. With proteomics, however, it is possible not only to analyse protein levels but inquire about post-translational modifications (Gafni, 2004). Although proteomic studies are less prevalent in ageing research, and arguably in biomedical research in general, there have been some successes. The most widely used rationale for applying proteomic analyses to ageing is the same for doing gene expression studies, that is, to identify protein differences with age or between short- and long-lived strains (Gafni, 2004). Recently, Dong *et al.* (2007) employed mass spectrometry to identify 86 proteins more or less abundant in a long-lived *daf-2* mutant *C. elegans* strain. They also showed that some of these proteins act in processes regulated by DAF-2, including development and ageing. In another study, Sowell *et al.* (2007) recently reported a mass-spectrometry-based survey of the *Drosophila* proteome at nine time points across adulthood. Overall, while not as high-throughput as gene expression studies, and in spite of advances in protein microarrays to assess protein abundance and function (LaBaer and Ramachandran, 2005), proteomics offers a complementary set of tools capable of providing unique insights.

Ultimately, one goal of functional genomics is to obtain gene expression maps for whole organisms: in other words, to know the expression patterns of nearly all genes in different cells or tissues and at different stages of the lifespan. Deriving such maps for lower organisms appears within reach, in particular for *C. elegans* since it only has 959 adult somatic cells and its cell lineage has been described. To achieve this objective, Dupuy *et al.* (2004) generated a 'Promoterome', a genome-wide resource of about 6000 *C. elegans* promoters which can be used to drive the expression of protein-encoding genes, such as localization markers like the green fluorescent protein (GFP). In mice,

one fascinating resource is the Allen Brain Atlas (http://www.brain-map.org/), a digital atlas of gene expression patterns of approximately 20 000 genes in the adult mouse brain. Although it does not presently contain different time points, the Allen Brain Atlas can serve as a reference point for comparisons with ageing and/or disease states (Lein *et al.*, 2007).

Another area of intense research is the study of transcriptional regulation and regulatory networks. This is a major challenge of the post-genome era because, even knowing the building blocks of life, genes can regulate their own expression and/or the expression of other genes in complex networks and regulatory loops. Phenotypes often arise from these multiple components and their complex interactions. One powerful technology to study transcriptional regulation is the ChIP-microarray or ChIP-chip where ChIP stands for chromatin immunoprecipitation. Succinctly, to identify targets of a given transcription factor, a whole-genome microarray can be used to find genomic regions to which the transcription factor is bound (Buck and Lieb, 2004). For example, Meier *et al.* (2007) used ChIP-chip technology to study mammalian DNA-damage response factors in damaged telomeres. Other studies of ageing have also taken into account gene regulatory regions. For example, one study in the ageing human brain reported that genes with reduced expression with age had increased DNA damage in their promoters, and in cell culture these same promoters were more susceptible to damage by oxidative stress (Lu *et al.*, 2004).

Understanding transcriptional regulation is further complicated, however, because gene expression can be regulated at many levels and our knowledge of these mechanisms is still incomplete. For example, the recent discovery of microRNAs and their capacity to regulate gene expression has added a new dimension to the problem and it is likely that other unknown mechanisms will still surprise researchers. Interestingly, Boehm and Slack (2005) showed that a microRNA regulates lifespan in *C. elegans*, so studies of these mechanisms have a huge potential to provide insights about ageing.

Although the above technologies are powerful by themselves, they can of course be used in combination with each other or with other techniques. For example, Hubner *et al.* (2005) focused on two rat tissues relevant to the pathogenesis of the metabolic syndrome and integrated gene expression profiles with linkage analysis to identify 73 candidate genes for hypertension. Studies combining genetic screens with microarray analysis have also been proposed to identify longevity genes in mice (de Haan and Williams, 2005). Another interesting example combining basic biology in a model organism with genome-wide approaches comes from the *SLC24A5* gene. A mutation in its zebrafish homolog was found to be responsible for pigmentation. Because it was known that the human gene showed a strong difference between populations, the researchers then demonstrated a strong association between an allele in *SLC24A5* and variation in pigmentation between European and African populations (Lamason *et al.*, 2005). Lastly, in *Drosophila*, Lai *et al.* (2007) studied gene expression in two strains for which quantitative trait loci (QTLs) affecting lifespan had been previously identified. Combining information on genes differentially expressed with age and between the two strains, they obtained a set of 49 candidate genes.

To cope with the large amounts of data being generated, including the growing number of genes being associated with ageing, several computational resources have been developed to assist researchers in navigating the seas of information (see *Table 1*). Indeed, we developed the Human Ageing Genomic Resources (de Magalhães *et al.*,

Table 1. *Major databases of relevance to biogerontologists.*

Name	URL	Type of data
AGEMAP	http://server1-kimlab.stanford.edu/cgi-bin/index.cgi?Home	Gene expression database for ageing in mice
Aging Gene Nexus	http://gan.usc.edu/	Microarray gene expression data
AGINGDB	http://aging.pharm.pusan.ac.kr/	Oxidative stress and calorie restriction in the study of ageing
AnAge	http://genomics.senescence.info/species/	Life history traits, longevity records, and ageing-related observations in organisms
Baltimore Longitudinal Study of Aging	http://blswww.grc.nia.nih.gov/	Repository for most of the longitudinal data collected during this study
GenAge	http://genomics.senescence.info/genes/	Ageing- and longevity-associated genes in model organisms and in humans
Human Mortality Database	http://www.mortality.org/	Mortality and population data
SAGE KE Database	http://uwaging.org/genesdb/	Genes and interventions that have been studied with respect to their effects on life-span or age-related neurological diseases

2005). One of its key components is GenAge (http://genomics.senescence.info/genes/), the first curated database of genes possibly related to human ageing. Recently, we also added a database of genes associated with ageing and/or longevity in model organisms as well as a list of genes analysed for their possible association with human longevity. Our aim in developing GenAge is to collect, manage and analyse the data. Not only do we want to summarize basic information about genes related to ageing, but we also want GenAge to be useful for researchers in many types of studies, including genetic association studies, or even for exploiting clinical interventions in human ageing. Another database of ageing- and longevity-associated genes is the SAGE KE database, which is now being hosted at the University of Washington (http://uwaging.org/genesdb/index.php). Interestingly, Lunetta *et al.* (2007) recently reported a genome-wide association study of longevity genes and performed an analysis focusing on candidate genes obtained from GenAge and from the SAGE KE database. Finally, another potentially useful database in biogerontology is the Gene Aging Nexus, a database of ageing gene-expression data (Pan *et al.*, 2007).

2.2 *Comparative and evolutionary genomics*

The human genome can be seen as the digital code from which proteins, the molecular machines of life, emerge (Hood and Galas, 2003). Comparative genomics, which is

the analysis and comparison of different genomes (either from different species or from different individuals of the same species), provides opportunities to decipher that code by allowing researchers to associate DNA variation with phenotypic variation.

Ageing and longevity are quantitative traits. Both are complex in the sense that they are modulated by two or more genes and their interactions with each other and with the environment. Associating complex diseases with DNA variants is challenging but researchers are now benefiting from the availability of large numbers of genetic markers and new genotyping technologies (Cardon and Bell, 2001). One common type of genetic variation is the single-nucleotide polymorphism (SNP). SNPs are now widely used genetic markers in genetic association studies and one major resource is HapMap, a public database with several million SNPs (International HapMap Consortium, 2005). Thanks to this and other projects, massive genome-wide association studies are ongoing to search for disease genes. One example is the Wellcome Trust Case Control Consortium (2007), which examined about 500 000 SNPs in 14 000 individuals affected by major diseases and in 3000 controls, yielding evidence for the involvement of many loci in the diseases.

Such studies promise to help bridge the gap between genotype and phenotype for hundreds of complex diseases. By focusing on longevity, genome-wide studies have a huge potential to help unravel the genetic basis of longevity, age-related diseases, and possibly of the ageing process itself. Of course, genes associated with longevity may or may not be related to the basic ageing process. For example, one of the first genes associated with longevity was *APOE*, with one of its alleles known to promote premature atherosclerosis reported to be less frequent in centenarians than in controls (Schachter *et al.*, 1994). Whether *APOE* is involved in the basic process of ageing remains to be determined. Nonetheless, associating genes with longevity and age-related diseases has a remarkable potential to improve human health and in some cases may provide mechanistic clues about ageing (Barzilai and Shuldiner, 2001). Moreover, with the dropping costs of resequencing the human genome, knowing the genetic determinants of human longevity is crucial for developing personalized medicine (Hood, 2003).

Certainly, an exciting area of research is identifying the genetic basis for individual variance in longevity and susceptibility to age-related diseases. An even greater phenotypic variation than that observed between individuals, however, is of course observed between species, and this is also reflected in the biology of ageing. Even similar species, like primates, can age at remarkably different rates (*Figure 1*). As in population studies, the rationale behind evolutionary genomics is that by studying different phenotypes of the process under study we can gather clues about its underlying genetics and mechanisms. It is clear that the genomic information determines the features of each species' ageing process to a large extent (Austad, 2005; de Magalhães, 2003). Multiple sequenced genomes provide data-mining opportunities to at least derive testable hypotheses. Identifying some of the genes involved in species differences in longevity would help understand the cellular and molecular basis for differences in the ageing process and in diseases like cancer, type II diabetes and neurodegenerative diseases. Eventually, it would inform about human susceptibility to such diseases.

Due to its much smaller size, sequencing the mitochondrial DNA (mtDNA) is considerably cheaper and quicker than sequencing the nuclear genome. As such, the mtDNA of hundreds of vertebrates has now been sequenced. Because mitochondria

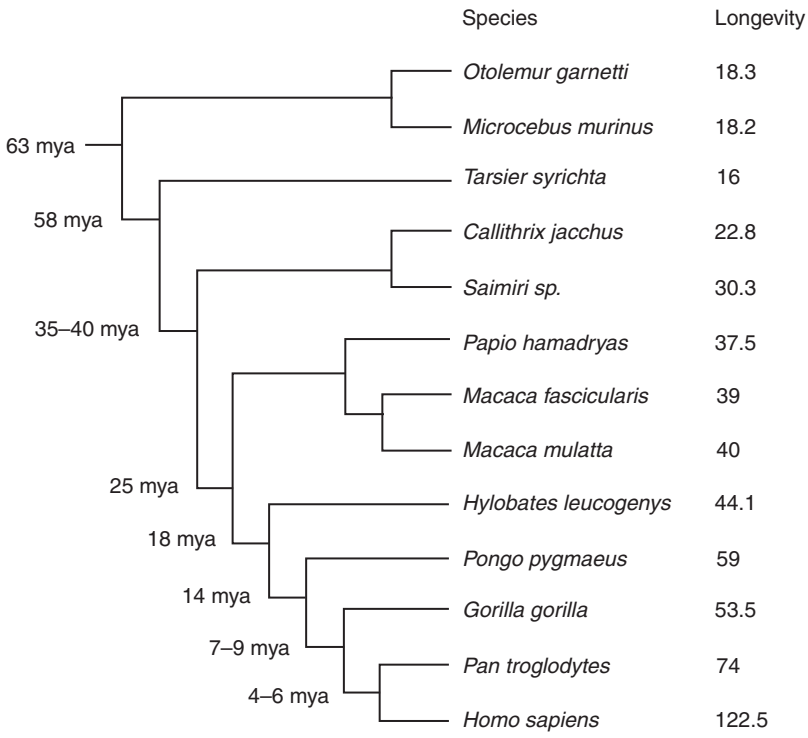

Figure 1. *Phylogeny and longevity of primate species approved for sequencing at the time of writing (http://www.genome.gov/10002154). The strategy behind evolutionary genomics of ageing, which can be seen as a form of comparative genomics, is to take advantage of the variation in longevity between similar species and of the growing number of sequenced genomes to understand not only the evolutionary pressures shaping longevity but obtain biologically significant clues about the genetics of ageing and age-related diseases. Longevity refers to maximum longevity and was obtained from build 9 of the AnAge database (de Magalhães* et al.*, 2007). Because it is not yet possible to quantify rate of ageing, maximum longevity under protected conditions is often used as an approximation to compare differences in ageing between species (de Magalhães, 2006). Human longevity, of course, may be overestimated due to the much larger sample size. Phylogeny from Goodman, 1999. Branch sizes are not to scale. mya, million years ago.*

have been associated with ageing, and hypothesized to play a causal role in the human ageing process, some studies have been conducted trying to associate some feature of the mtDNA with the longevity of different species (de Magalhães, 2005a; Khaidakov *et al.*, 2006; Lehmann *et al.*, 2006; Rottenberg, 2007). For example, in one intriguing study Samuels (2005) demonstrated a relation between the free energy of mitochondrial DNA, which can be seen as a factor in its susceptibility to mutation, and longevity. Although there are alternative explanations for all these studies, they could suggest a general pattern of selection on the mitochondrion that is associated with the evolution of longevity.

When studying the evolution of the much-larger nuclear genome to gather clues about ageing, one major drawback is the fact that ageing is a largely mysterious process. The genetic basis of longevity is also mostly unknown, which means researchers are virtually searching needles in a haystack. An example may be useful. Because humans tend to live longer and exhibit signs of ageing at later ages than chimpanzees, we recently studied the evolutionary forces acting on human-chimpanzee orthologous gene pairs. With only two species, however, our signal-to-noise ratio was still low and hence our analyses were limited to protein-coding sequences and to pathways and mechanisms previously associated with ageing. Surprisingly, we found that ageing-associated genes tend to evolve under stronger-than-average functional constraints. Albeit interesting from an evolutionary perspective, however, we failed to obtain significant insights regarding the genetic basis for differences in ageing and longevity between these two species (de Magalhães and Church, 2007).

Many different theoretical frameworks seeking to explain species differences in ageing have been put forward. One of the most popular explanations is the free radical theory of ageing, which succinctly argues that species with higher generation of reactive oxygen species (ROS) accumulate damage faster and thus age faster. Because mitochondria, the cells' powerhouses, are one major source of ROS, the free radical theory offered an elegant explanation to the long-held view that short-lived mammals have higher metabolic rates (Harman, 1981). Recently, however, this explanation has been put into question by the demonstration that metabolic rates, when statistically corrected for the effects of body mass and phylogeny, do not correlate with longevity in placental mammals or in birds (de Magalhães *et al.*, 2007a).

In addition to ageing being largely mysterious, another problem in evolutionary genomics (of ageing and otherwise) is that we have few clues regarding which genomic features are involved in species differences and in adaptation. New genes can obviously give rise to new phenotypes. For example, transgenic mice expressing a CPD-photolyase, enzymes involved in repair of UV-induced damage that were evolutionarily lost in placental mammals, were protected from sunlight-induced tumorigenesis (Jans *et al.*, 2005). Probably, though, more subtle adaptive changes in gene products or in their regulatory regions underlie the augmented human longevity. Considering that humans are thought to have essentially the same set of genes as chimpanzees, it is unlikely that differences in ageing are due to the loss or evolution of new genes, at least among such closely related species. It has long been thought that differences between similar species may be due to transcriptional regulation (King and Wilson, 1975), though probably a combination of changes in regulatory and protein-coding regions are involved (Hoekstra and Coyne, 2007). The challenge is to identify such changes from the background of neutral mutations.

In spite of these difficulties, it appears plausible that, as the genomes of more organisms are sequenced, it will be possible to associate selection patterns in one or more genes with the evolution of longevity in mammalian and/or primate lineages. In that context, sequencing the genome of long-lived mammals would be a tremendous help (de Magalhães *et al.*, 2007b). New genomic methods also promise to aid in these efforts. Because differences in gene expression might play a role in species differences in ageing, gene expression studies can be used in combination with genome comparisons to study regulatory networks. Using a multispecies cDNA microarray, Gilad *et al.* (2006) identified genes with expression levels significantly elevated or reduced in the human

lineage when compared to other primates. Transcription factors were particularly prevalent among the identified genes. Overall, I am convinced that increases in the signal-to-noise ratio through better methods and/or more data will make it possible to identify candidate human longevity genes. Eventually, with the dropping costs of resequencing the human genome (Church, 2006), such genes can be further studied in genetic association studies in human populations and may provide important clues about the basic cellular and molecular mechanisms of ageing.

3 Systems biology

The aforementioned technologies have allowed researchers to survey biological systems and their molecular components on a global scale. Systems biology is not only the study of these components but how they interact to give rise to the system under study (Hood, 2003; Kriete *et al.*, 2006). In biogerontology, many genes have been related to ageing in model systems but it is now necessary to study how these genes interact and how they exert their influence as an aggregate to modulate the ageing process. As such, the ageing process is suitable for a system-level approach (Kirkwood *et al.*, 2003; Vijg and Suh, 2003). GenAge, in fact, was developed as a tool for systems biology by providing a core dataset that allows researchers to employ data-mining algorithms and computational approaches to study ageing from a genomic, proteomic and evolutionary perspective.

One simple way to study a complex system is to observe the effects of disrupting each of its components. Indeed, based on GenAge, I developed a system-level interpretation of genetic interventions that alter the ageing process in model systems to help comprehend the mechanisms of ageing. One of my conclusions was that DNA alterations appear to be more relevant to ageing than other forms of molecular damage (de Magalhães, 2005b).

As aforementioned, a key aspect of systems biology involves the study of how the system's components – typically genes and/or proteins, but they could be metabolites or cells – interact. Some potentially useful network studies of ageing have already been conducted (Bergman *et al.*, 2007; Promislow, 2004). Of particular interest, a number of researchers have studied networks of ageing-associated proteins in order to identify new candidate genes or proteins that may play a role in ageing (Budovsky *et al.*, 2007; de Magalhães and Toussaint, 2004b; Ferrarini *et al.*, 2005). In this 'guilt-by-association' method, the principle is that proteins that tend to interact with proteins previously associated with ageing have a higher probability of also being involved in ageing than what would be expected by chance, and are thus promising targets for further experiments (see *Figure 2*).

The future of systems biology and of biological research in general is in integrating different sources of information to understand complex processes. Indeed systems biology is not limited to a few technologies or approaches. For example, as discussed above, evolutionary biology provides a unique perspective on gene action, and evolutionary studies can be integrated into systems biology. In fact, evolutionary systems biology is a major emerging field whose broad goal is to study how gene networks evolve (Medina, 2005). Understanding regulatory networks also benefits from the integration of different types of data, such as combining microarray data with literature mining, which can be seen as a form of systems biology (Werner, 2007). Likewise, some

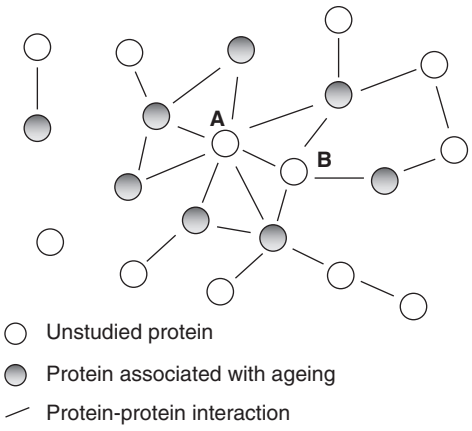

Figure 2. *'Guilt-by-association' method in a simple hypothetical network of protein-protein interactions. Each circle represents a protein with grey circles representing proteins previously associated with ageing and white circles representing proteins not previously studied in the context of ageing. Just from visualizing the network, it is clear that proteins A and B are the best targets for future experiments because they interact with a large number of proteins already linked to ageing, which makes A and B more likely to play a role in ageing. Typically, though, protein networks tend to be more complex with more proteins and interactions. It is also possible to define the type of the interactions and take into consideration regulatory interactions, such as activation or repression, which require more advanced algorithms to infer the best targets for future experiments.*

of the integrative genomic approaches described above can be classified as systems biology approaches.

Although promising, applying systems biology to ageing research still faces some significant hurdles. For instance, ageing research suffers from a lack of appropriate experimental models (de Magalhães and Toussaint, 2004a). In cancer biology, applying the principles of systems biology is much easier because cancer researchers have at their disposal multiple cancer cell lines in which to conduct a battery of large-scale analyses using '-omic' technologies, which can then be integrated to identify cancer markers that make it possible to develop personalized cancer therapies (Weinstein, 2004). In fact, one exciting project is the Cancer Genome Atlas, a comprehensive effort to explore the spectrum of genomic changes associated with cancer (http://cancergenome.nih.gov). Systems biology is also appropriate to study infectious diseases, which perturb the network structure of a system (e.g. the immune system), and are themselves caused by agents that are made of interacting components and whose genome can be (or has been) sequenced. Although cellular models of ageing are suitable for systems biology, applying systems biology to complex organisms like mammals still suffers from a lack of experimental paradigms. Other applications and problems of systems biology in the study of ageing are discussed below.

3.1 *Modelling the ageing process*

One goal of systems biology is to develop computer models of the whole system from its parts and their relationships. *In silico* models are easier to interpret and may allow novel insights into the process under study. The ultimate aim of modelling is to predict phenotypes that may help guide experiments and may even have clinical applications. For example, as mentioned above, quantifying different aspects of ageing may reveal a fingerprint of ageing that allows researchers to quantify the whole ageing process (Kriete, 2006).

Models can be constructed at different levels, focusing on different components of the system, and with varying degrees of mathematical complexity (de Jong, 2002). The type of model depends on what is sought from the model and on the data available, with better data of course allowing more complete models. Modelling can follow a top-down or a bottom-up approach (or a combination of the two). A top-down model begins with an overview of the system. Its parts are then refined until a level of detail is reached that provides suitable predictions. In bottom-up modelling, the components of the system are first identified and characterized and their relationships are then put together until the entire system is modelled.

Probably the main problem in modelling ageing is that we still do not understand the essence of the ageing process (de Magalhães, 2005b). Some researchers have even argued that there is no basic ageing process and that what we call ageing is in fact the product of different age-related diseases and degenerative processes occurring approximately in parallel (Holliday, 1995). Recent breakthroughs in the genetics of ageing showing it is possible to modulate the entire ageing process by manipulating single genes, however, have cast doubts on this view (de Magalhães, 2003).

Others view ageing as the result of one or more forms of molecular damage accumulating with age, in particular damage caused by chemical reactions such as oxidative damage or DNA mutations. But even if ageing is caused by the accumulation of molecular damage and inefficient repair systems, there are different mechanisms by which this may occur and these need to be modelled in different ways. Bell (1998) argued that there could be conceptually different types of repair. DNA, for example, can be repaired based on an original blueprint. Another way of eliminating errors is to have a large number of DNA copies, for example in different cells, and then subject those cells to a given test that eliminates cells with a number of DNA mutations above a certain threshold (in this case, the outcome might be apoptosis). To put it another way, evolution may increase longevity by optimizing repair mechanisms – and even a slight increase in DNA repair will have major consequences over the large number of cell divisions necessary to produce an adult mammal – or evolution might operate by optimizing cell selection (e.g. acting on processes like apoptosis and mitosis).

The point has also been made that degeneration is not caused by damage to the components of a system but rather by the inability to replace those components (Williams, 1957). Ageing can be seen as a result of systems being redundant in irreplaceable elements (Gavrilov and Gavrilova, 2001), and some authors have suggested a 'network theory of ageing' integrating different forms of molecular damage (Franceschi *et al.*, 2000; Kowald and Kirkwood, 1996). Alternatively, a completely different hypothesis is that some developmental mechanisms shaped by evolution to optimize reproduction have an impact on ageing and age-related diseases (de Magalhães

and Church, 2005). The free radical theory of ageing can, in fact, be interpreted in light of the antagonistic roles of ROS in development/growth and in ageing (de Magalhães and Church, 2006). It appears that at least some forms of damage may be caused by genetic programmes (e.g. as a form of antagonistic pleiotropy) rather than inefficient repair systems.

Overall, there is no consensus on what the driving force of ageing is or what type of components play a role in it, which makes it nearly impossible to model ageing in higher organisms with precision. For now, modelling ageing has been restricted to lower organisms or simple pathways. A model of yeast ageing has been proposed (Gillespie *et al.*, 2004), as has a model of the role of chaperones in ageing (Proctor *et al.*, 2005). Ultimately, however, the goal of systems biology is to reconstruct the causal structure of the human ageing process at different levels and understand its genetic network.

4 Future prospects

'By the year 2030, we will have (1) developed a complete model of all human cell types, obviating the need for many laboratory experiments [by doing computer simulations instead]; (2) lowered the cost of doing a complete genomic sequence for an human individual to less than $1,000 each; and (3) catalogued all the genes involved in ageing. Therefore, human clinical trials to extend lifespan could already be underway by this date.' – Francis Collins

It is my belief that understanding ageing will require what is known in reverse engineering as 'a grey box approach': a combination of white and black box. A white box is the source code, the binary code being decoded: in this case, the genome. The black box approach involves observing results of different inputs in the system being analysed to understand its functioning: in this case, genetic manipulation experiments followed by observations from gross phenotypes to changes in genes and proteins. In this context, microarrays and proteomics provide new high-throughput ways to generate massive amounts of detailed observations of the system under study, while a systems-biology thinking allows researchers to fit these immense amounts of data into coherent models.

The lack of a theoretical framework that explains ageing (or the large numbers of frameworks that may or may not be correct) is a difficulty in applying genomic technologies to ageing, from interpreting gene expression results to modelling ageing, passing by evolutionary genomics. On the other hand, genomic technologies have so far been mostly used for hypothesis generation rather than hypothesis testing and may even provide new conceptual frameworks that have so far eluded biogerontologists.

Biomedical discovery will continue to benefit from the extraordinary progress of technology and provide researchers with new horizons. For example, as mentioned above, comparative genomics is an emerging field and it will soon be possible to resequence the genome of thousands of patients. As resequencing costs continue to drop, however, even more ambitious projects become feasible. One endeavour that may allow researchers to address the role of mutations in ageing would be to produce a phylogenetic tree of cells. The idea would be to catalogue somatic mutations in the genomes of individual cells during development and ageing of mammals, to sequence the genome of many cells in a mouse. This would allow researchers to infer the history

of cell divisions and obtain a mammalian cell-fate map describing cellular lineages during the entire lifespan (Salipante and Horwitz, 2007).

Ultimately, biomedical research will gradually transform biologists and physicians from observers to architects. At present, there are inherited limitations to radically engineering biological systems and not many success stories. In non-biological fields, engineers have an in-depth understanding that permits computer models and an abstraction of the task at hand. The situation is quite different in biology, at least for now (Heinemann and Panke, 2006). Progress in the emergent field of synthetic biology, however, coupled with a more complete understanding of biological processes made possible by genomic technologies, promises gradually to augment our capacity to engineer biology. De Grey (2003), for example, has argued that it will be possible within a foreseeable future to engineer human ageing to the point of reversing it.

In conclusion, this is an exciting time for research in the biomedical sciences with unprecedented opportunities for discovery. When compared to other biomedical fields, ageing research is still in its infancy. Biogerontologists have now high-throughput technologies with which to prod the mechanisms of ageing, and the computational capacity to build models of the ageing process with an increasing ability to predict phenotypes. In the next decade or two, I anticipate we will learn more about the basic ageing process and its genetic, cellular and molecular underpinnings than ever before.

Acknowledgements

I would like to thank everyone at the Annual Scientific Meeting of the Society for Experimental Biology in Glasgow, March–April 2007, for excellent discussions on these topics. Further thanks to the participants of the Workshop on Complexities of Aging in Biological Systems in Santa Fe, March 2007, for useful suggestions and ideas particularly on applying systems biology to ageing. My work is supported by a NIH-NHGRI CEGS grant to George Church.

References

Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422:** 198–207.

Austad, S.N. (2005) Diverse aging rates in metazoans: targets for functional genomics. *Mech. Ageing Dev.* **126:** 43–49.

Barzilai, N. and Shuldiner, A.R. (2001) Searching for human longevity genes: the future history of gerontology in the post-genomic era. *J. Gerontol. A Biol. Sci. Med. Sci.* **56:** M83–87.

Bell, G. (1988) *Sex and Death in Protozoa: the History of Obsession*. Cambridge University Press, Cambridge.

Bergman, A., Atzmon, G., Ye, K., MacCarthy, T. and Barzilai, N. (2007) Buffering mechanisms in aging: a systems approach toward uncovering the genetic component of aging. *PLoS Comput. Biol.* **3:** e170.

Boehm, M. and Slack, F. (2005) A developmental timing microRNA and its target regulate life span in *C. elegans*. *Science* **310:** 1954–1957.

Brown, P.O. and Botstein, D. (1999) Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* **21:** 33–37.

Buck, M.J. and Lieb, J.D. (2004) ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics* **83:** 349–360.

Budovsky, A., Abramovich, A., Cohen, R., Chalifa-Caspi, V. and Fraifeld, V. (2007) Longevity network: construction and implications. *Mech. Ageing Dev.* **128:** 117–124.

Cardon, L.R. and Bell, J.I. (2001) Association study designs for complex diseases. *Nat. Rev. Genet.* **2:** 91–99.

Church, G.M. (2006) Genomes for all. *Sci. Am.* **294:** 46–54.

de Grey, A.D. (2003) An engineer's approach to the development of real anti-aging medicine. *Sci. Aging Knowledge Environ.* **2003:** VP1.

de Haan, G. and Williams, R.W. (2005) A genetic and genomic approach to identify longevity genes in mice. *Mech. Ageing Dev.* **126:** 133–138.

de Jong, H. (2002) Modeling and simulation of genetic regulatory systems: a literature review. *J. Comput. Biol.* **9:** 67–103.

de Magalhães, J.P. (2003) Is mammalian aging genetically controlled? *Biogerontology* **4:** 119–120.

de Magalhães, J.P. (2005a) Human disease-associated mitochondrial mutations fixed in nonhuman primates. *J. Mol. Evol.* **61:** 491–497.

de Magalhães, J.P. (2005b) Open-minded scepticism: inferring the causal mechanisms of human ageing from genetic perturbations. *Ageing Res. Rev.* **4:** 1–22.

de Magalhães, J.P. (2006) Species selection in comparative studies of aging and anti-aging research. In: *Handbook of Models for Human Aging* (ed. P.M. Conn). Elsevier Academic Press, Burlington, MA, pp. 9–20.

de Magalhães, J.P. and Church, G.M. (2005) Genomes optimize reproduction: aging as a consequence of the developmental program. *Physiology (Bethesda)* **20:** 252–259.

de Magalhães, J.P. and Church, G.M. (2006) Cells discover fire: employing reactive oxygen species in development and consequences for aging. *Exp. Gerontol.* **41:** 1–10.

de Magalhães, J.P. and Church, G.M. (2007) Analyses of human-chimpanzee orthologous gene pairs to explore evolutionary hypotheses of aging. *Mech. Ageing Dev.* **128:** 355–364.

de Magalhães, J.P. and Toussaint, O. (2004a) How bioinformatics can help reverse engineer human aging. *Ageing Res. Rev.* **3:** 125–141.

de Magalhães, J.P. and Toussaint, O. (2004b) GenAge: a genomic and proteomic network map of human ageing. *FEBS Lett.* **571:** 243–247.

de Magalhães, J.P., Costa, J. and Toussaint, O. (2005) HAGR: the Human Ageing Genomic Resources. *Nucleic Acids Res.* **33 Database Issue:** D537–543.

de Magalhães, J.P., Costa, J. and Church, G.M. (2007a) An analysis of the relationship between metabolism, developmental schedules, and longevity using phylogenetic independent contrasts. *J. Gerontol. A Biol. Sci. Med. Sci.* **62:** 149–160.

de Magalhães, J.P., Sedivy, J.M., Finch, C.E., Austad, S.N. and Church, G.M. (2007b) A proposal to sequence genomes of unique interest for research on aging. *J. Gerontol. A Biol. Sci. Med. Sci.* **62:** 583–584.

Dierick, J.F., Dieu, M., Remacle, J., Raes, M., Roepstorff, P. and Toussaint, O. (2002) Proteomics in experimental gerontology. *Exp. Gerontol.* **37:** 721–734.

Dong, M.Q., Venable, J.D., Au, N., Xu, T., Park, S.K., Cociorva, D., Johnson, J.R., Dillin, A. and Yates, J.R. 3rd (2007) Quantitative mass spectrometry identifies insulin signaling targets in *C. elegans*. *Science* **317:** 660–663.

Dupuy, D., Li, Q.R., Deplancke, B., *et al.* (2004) A first version of the *Caenorhabditis elegans* Promoterome. *Genome Res.* **14:** 2169–2175.

Eriksson, M., Brown, W.T., Gordon, L.B., *et al.* (2003) Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature* **423:** 293–298.

Ferrarini, L., Bertelli, L., Feala, J., McCulloch, A.D. and Paternostro, G. (2005) A more efficient search strategy for aging genes based on connectivity. *Bioinformatics* **21:** 338–348.

Franceschi, C., Valensin, S., Bonafe, M., Paolisso, G., Yashin, A.I., Monti, D. and De Benedictis, G. (2000) The network and the remodeling theories of aging: historical background and new perspectives. *Exp. Gerontol.* **35:** 879–896.

Gafni, A. (2004) Proteomics in aging-related research. *Sci. Aging Knowledge Environ.* **2004:** pe41.

Gavrilov, L.A. and Gavrilova, N.S. (2001) The reliability theory of aging and longevity. *J. Theor. Biol.* **213:** 527–545.

Gilad, Y., Oshlack, A., Smyth, G.K., Speed, T.P. and White, K.P. (2006) Expression profiling in primates reveals a rapid evolution of human transcription factors. *Nature* **440:** 242–245.

Gillespie, C.S., Proctor, C.J., Boys, R.J., Shanley, D.P., Wilkinson, D.J. and Kirkwood, T.B. (2004) A mathematical model of ageing in yeast. *J. Theor. Biol.* **229:** 189–196.

Goodman, M. (1999) The genomic record of Humankind's evolutionary roots. *Am. J. Hum. Genet.* **64:** 31–39.

Hamilton, B., Dong, Y., Shindo, M., Liu, W., Odell, I., Ruvkun, G. and Lee, S.S. (2005) A systematic RNAi screen for longevity genes in *C. elegans*. *Genes Dev* **19:** 1544–1555.

Hannon, G.J. (2002) RNA interference. *Nature* **418:** 244–251.

Harman, D. (1981) The aging process. *Proc. Natl Acad. Sci. USA* **78:** 7124–7128.

Heinemann, M. and Panke, S. (2006) Synthetic biology – putting engineering into biology. *Bioinformatics* **22:** 2790–2799.

Hoekstra, H.E. and Coyne, J.A. (2007) The locus of evolution: evo devo and the genetics of adaptation. *Evol. Int. J. Org. Evol.* **61:** 995–1016.

Holliday, R. (1995) *Understanding Ageing*. Cambridge University Press, Cambridge.

Hood, L. (2003) Systems biology: integrating technology, biology, and computation. *Mech. Ageing Dev.* **124:** 9–16.

Hood, L. and Galas, D. (2003) The digital code of DNA. *Nature* **421:** 444–448.

Hubner, N., Wallace, C.A., Zimdahl, H., *et al.* (2005) Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nat. Genet.* **37:** 243–253.

International HapMap Consortium (2005) A haplotype map of the human genome. *Nature* **437:** 1299–1320.

Jans, J., Schul, W., Sert, Y.G., *et al.* (2005) Powerful skin cancer protection by a CPD-photolyase transgene. *Curr. Biol.* **15:** 105–115.

Kaeberlein, M. (2004) Aging-related research in the '-omics' age. *Sci. Aging Knowledge Environ.* **2004:** pe39.

Khaidakov, M., Siegel, E.R. and Shmookler Reis, R.J. (2006) Direct repeats in mitochondrial DNA and mammalian lifespan. *Mech. Ageing Dev.* **127:** 808–812.

King, M.C. and Wilson, A.C. (1975) Evolution at two levels in humans and chimpanzees. *Science* **188:** 107–116.

Kirkwood, T.B., Boys, R.J., Gillespie, C.S., Proctor, C.J., Shanley, D.P. and Wilkinson, D.J. (2003) Towards an e-biology of ageing: integrating theory and data. *Nat. Rev. Mol. Cell Biol.* **4:** 243–249.

Kowald, A. and Kirkwood, T.B. (1996) A network theory of ageing: the interactions of defective mitochondria, aberrant proteins, free radicals and scavengers in the ageing process. *Mutat. Res.* **316:** 209–236.

Kriete, A. (2006) Biomarkers of aging: combinatorial or systems model? *Sci. Aging Knowledge Environ.* **2006:** pe1.

Kriete, A., Sokhansanj, B.A., Coppock, D.L. and West, G.B. (2006) Systems approaches to the networks of aging. *Ageing Res. Rev.* **5:** 434-448.

LaBaer, J. and Ramachandran, N. (2005) Protein microarrays as tools for functional proteomics. *Curr. Opin. Chem. Biol.* **9:** 14–19.

Lai, C.Q., Parnell, L.D., Lyman, R.F., Ordovas, J.M. and Mackay, T.F. (2007) Candidate genes affecting *Drosophila* life span identified by integrating microarray gene expression analysis and QTL mapping. *Mech. Ageing Dev.* **128:** 237–249.

Lamason, R.L., Mohideen, M.A., Mest, J.R., *et al.* (2005) SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science* **310:** 1782–1786.

Lamb, J., Crawford, E.D., Peck, D., *et al.* (2006) The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313:** 1929–1935.

Lee, S.S., Lee, R.Y., Fraser, A.G., Kamath, R.S., Ahringer, J. and Ruvkun, G. (2003) A systematic RNAi screen identifies a critical role for mitochondria in *C. elegans* longevity. *Nat. Genet.* **33:** 40–48.

Lehmann, G., Budovsky, A., Muradian, K.K. and Fraifeld, V.E. (2006) Mitochondrial genome anatomy and species-specific lifespan. *Rejuvenat. Res.* **9:** 223–226.

Lein, E.S., Hawrylycz, M.J., Ao, N., *et al.* (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445:** 168–176.

Lim, C.A. and Ng, H.H. (2007) Application of advanced technologies in ageing research. *Mech. Ageing Dev.* **128:** 149–160.

Lu, T., Pan, Y., Kao, S.Y., Li, C., Kohane, I., Chan, J. and Yankner, B.A. (2004) Gene regulation and DNA damage in the ageing human brain. *Nature* **429:** 883–891.

Lunetta, K.L., D'Agostino, R.B. Sr., Karasik, D., *et al.* (2007) Genetic correlates of longevity and selected age-related phenotypes: a genome-wide association study in the Framingham Study. *BMC Med. Genet.* **8 Suppl 1:** S13.

McElwee, J.J., Schuster, E., Blanc, E., *et al.* (2007) Evolutionary conservation of regulated longevity assurance mechanisms. *Genome Biol.* **8:** R132.

Medina, M. (2005) Genomes, phylogeny, and evolutionary systems biology. *Proc. Natl Acad. Sci. USA* **102 Suppl 1:** 6630–6635.

Meier, A., Fiegler, H., Munoz, P., Ellis, P., Rigler, D., Langford, C., Blasco, M.A., Carter, N. and Jackson, S.P. (2007) Spreading of mammalian DNA-damage response factors studied by ChIP-chip at damaged telomeres. *EMBO J.* **26:** 2707–2718.

Miller, R.A. (2001) Biomarkers of aging. *Sci. Aging Knowledge Environ.* **2001:** pe2.

Murphy, C.T., McCarroll, S.A., Bargmann, C.I., Fraser, A., Kamath, R.S., Ahringer, J., Li, H. and Kenyon, C. (2003) Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature* **424:** 277–283.

Pan, F., Chiu, C.H., Pulapura, S., *et al.* (2007) Gene Aging Nexus: a web database and data mining platform for microarray data on aging. *Nucleic Acids Res.* **35:** D756–759.

Proctor, C.J., Soti, C., Boys, R.J., Gillespie, C.S., Shanley, D.P., Wilkinson, D.J. and Kirkwood, T.B. (2005) Modelling the actions of chaperones and their role in ageing. *Mech. Ageing Dev.* **126:** 119–131.

Promislow, D.E. (2004) Protein networks, pleiotropy and the evolution of senescence. *Proc. Biol. Sci.* **271:** 1225–1234.

Raghothama, C., Harsha, H.C., Prasad, C.K. and Pandey, A. (2005) Bioinformatics and proteomics approaches for aging research. *Biogerontology* **6:** 227–232.

Rottenberg, H. (2007) Coevolution of exceptional longevity, exceptionally high metabolic rates, and mitochondrial DNA-coded proteins in mammals. *Exp. Gerontol.* **42:** 364–373.

Salipante, S.J. and Horwitz, M.S. (2007) A phylogenetic approach to mapping cell fate. *Curr. Top. Dev. Biol.* **79:** 157–184.

Samuels, D.C. (2005) Life span is related to the free energy of mitochondrial DNA. *Mech. Ageing Dev.* **126:** 1123–1129.

Schachter, F., Faure-Delanef, L., Guenot, F., Rouger, H., Froguel, P., Lesueur-Ginot, L. and Cohen, D. (1994) Genetic associations with human longevity at the APOE and ACE loci. *Nat. Genet.* **6:** 29–32.

Sowell, R.A., Hersberger, K.E., Kaufman, T.C. and Clemmer, D.E. (2007) Examining the proteome of *Drosophila* across organism lifespan. *J. Proteome Res.* **6:** 3637–3647.

Spindler, S.R. and Mote, P.L. (2007) Screening candidate longevity therapeutics using gene-expression arrays. *Gerontology* **53:** 306–321.

Vijg, J. and Suh, Y. (2003) Functional genomics of ageing. *Mech. Ageing Dev.* **124:** 3–8.

Weinstein, J.N. (2004) Integromic analysis of the NCI-60 cancer cell lines. *Breast Dis.* **19:** 11–22.

Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447:** 661–678.

Werner, T. (2007) Regulatory networks: linking microarray data to systems biology. *Mech. Ageing Dev.* **128:** 168–172.

Williams, G.C. (1957) Pleiotropy, natural selection, and the evolution of senescence. *Evolution* **11:** 398–411.