Ageing Research Reviews
3 (2004) 125–141

ageing
research
reviews

www.elsevier.com/locate/arr

Review

How bioinformatics can help reverse engineer human aging

João Pedro de Magalhães*, Olivier Toussaint

Unit of Cellular Biochemistry and Biology (URBC), Department of Biology, University of Namur (FUNDP), Rue de Bruxelles 61, B-5000 Namur, Belgium

Received 5 March 2003; received in revised form 28 August 2003; accepted 29 August 2003

Abstract

To study human aging is an enormous challenge. The complexity of the aging phenotype and the near impossibility of studying aging directly in humans oblige researchers to resort to models and extrapolations. Computational approaches offer a powerful set of tools to study human aging. In one direction we have data-mining methods, from comparative genomics to DNA microarrays, to retrieve information in large amounts of data. Afterwards, tools from systems biology to reverse engineering algorithms allow researchers to integrate different types of information to increase our knowledge about human aging. Computer methodologies will play a crucial role to reconstruct the genetic network of human aging and the associated regulatory mechanisms.

Keywords: Aging; Bioinformatics; Computational biology; Gerontology; Systems biology

1. Introduction

Aging is an intrinsic age-related process of loss of viability and increase in vulnerability (Comfort, 1964). Studying human aging has two major difficulties: the complexity of the aging phenotype with its widespread changes and pathologies associated with chronological age and the near impossibility of performing in vivo studies. Consequently, most researchers resort to models that may or may not be accurate representations of the human aging process. So, to understand human aging is an enormous challenge, not only due to the complexity inherent to aging but also since our hypotheses are based on extrapolations and our theories will have to be tested indirectly.

* Corresponding author. Tel.: +32-81-724133; fax: +32-81-724135.
E-mail address: jpnitya@senescence.info (J.P. de Magalhães).

1568-1637/$ – see front matter
doi:10.1016/j.arr.2003.08.006

It is not surprising then that human aging remains a mysterious process. We still cannot answer the most important questions: why do species age at different paces? What determines the rate of aging? How can we distinguish causes from effects of aging? What changes occur in an adult human being to make the chances of dying duplicate approximately every 8 years? In the end, why do we age? So far, we already know several genes that modulate rate of aging in animals, such as $p66^{shc}$ (Migliaccio et al., 1999), but optimistically, only a few in humans such as the progeroid genes responsible for Werner's (Yu et al., 1996) and Hutchinson–Gilford's syndromes (Eriksson et al., 2003). Yet these genes alone show the extraordinary influence of genetic mechanisms on aging. Just discovering more genetic players in the human aging process would be a major breakthrough.

Scientific discovery has always been limited by the available technologies. Given the modern methods to gather huge amounts of data, such as the recent genome sequencing efforts (Lander et al., 2001; Venter et al., 2001; Waterston et al., 2002), computational methods offer a powerful set of tools to help understand the aging process and its genetic network (reviewed in Luscombe et al., 2001; Yaspo, 2001). In this article, our objective is to focus on computational tools that may be useful for the study of human aging by: (1) data-mining methods, such as comparative genomics, phylogenetic footprinting, and DNA microarrays, that may be employed to gather information about aging; (2) in silico methodologies aimed at interpreting information, such as algorithms to understand complex networks, that may be useful to model the genetic network of human aging.

2. Data-mining methods

Computational data-mining approaches are particularly appropriate in areas with much data but few explanations, such as gerontology. If researchers can find patterns in data to perceive information, then information may enhance our knowledge over aging. The goal of applying computational data-mining approaches is to extract useful information from large amounts of data by employing mathematical methods that should be as automated as possible.

2.1. Comparative genomics of aging

Having the human genome offers the digital genetic code that is the source of aging while having several fully sequenced genomes offers data-mining opportunities to decipher that code. Since genomes had a common ancestor, every base pair in each organism can be explained as a combination of the ancestral genome with evolution, and so comparing different genomes is a powerful way to analyze and interpret genome sequence. Comparative genomics allows researchers to, for instance, discover new genes, assign function to unknown genes, and gather information on protein interactions (reviewed in Wei et al., 2002; Ureta-Vidal et al., 2003).

Early studies using the several available microbial genomes already showed how comparative genomics can be used to gather information about gene function (reviewed in Yaspo, 2001; Wei et al., 2002). One method involves determining the presence or absence of proteins across several genomes to build a phylogenetic profile for each one. Functionally related

proteins are expected to have similar profiles and thus can be grouped together, giving hints on the function of several unknown proteins (Pellegrini et al., 1999). Another approach to find functionally related proteins involves a "Rosetta stone" analogy. In brief, if the homologues of a pair of proteins in an organism are found fused into a single protein chain in another organism's genome, it can be used to infer a functional interaction (Marcotte et al., 1999a). Lastly, more detailed comparisons may be performed. In one example, the DNA repair proteins of *Escherichia coli* and *Saccaromyces cerevisiae* were compared to the entire sets of protein sequences of several fully sequenced genomes. Multiple alignments of the protein families found were constructed using algorithms such as ClustalW (Thompson et al., 1994). This method allowed an identification of novel enzymatic and DNA-binding domains involved in repair, as well as an improved view on the evolution of repair systems and how these depend on environmental conditions and the cell's physiology. By identifying domains with disrupted functional motifs, researchers can also predict proteins without enzymatic activity (Aravind et al., 1999).

It may be possible to discover novel proteins involved in aging through phylogenetic profiles using as reference, for example, proteins involved in progeroid syndromes. In addition, many different organisms exhibit extraordinary similar aging phenotypes at radically different paces. For example, some primates such as baboons (Bronikowski et al., 2002) or rhesus monkeys exhibit a similar aging process to humans but twice as fast and mice age 25–30 times faster than humans do (Finch, 1990); on the other hand, whales appear to age slower than humans (George et al., 1999). These observations are independent of environmental conditions, suggesting that genetic factors are largely responsible for rate of aging in mammals (Miller, 1999; de Magalhaes, 2003). Although different rates of aging may also be due to metabolic rates, these alone do not explain the variety in mammalian rates of aging (Austad, 1997). Eventually, it will be possible to compare whole sets of protein families across mammalian species with different rates of aging in hope of finding some sort of correlation, for example, in the form of expanded or contracted protein families.

Since the digital core of information from which aging arises is ultimately knowable (Hood, 2003), comparative genomics offers an accessible set of tools to study aging in humans. The major limitation in applying the described methodologies to mammals is the need for several fully sequenced genomes, which should occur within a reasonable future. For example, expansion and contraction of protein families in mice—mostly reproduction, immunity, and olfaction—do not suggest any involvement with aging (Waterston et al., 2002). Yet with the prospect of having several fully sequenced mammalian genomes within the next few years—namely, human, mouse, rat, chimpanzee, rhesus macaque, cow, pig, and dog genomes (http://www.genome.gov/page.cfm?pageID=10002154)—comparative genomics offers a powerful approach to study human aging.

2.2. Transcriptional regulation of aging

One of the major discoveries of the sequencing of the mouse and human genomes is the extraordinary genetic similarity between mice and humans. According to the latest projections, mice share roughly 99% of the total number of human genes (Waterston et al., 2002), and some scientists argue that humans and their close relatives, e.g. chimpanzees, have the same set of genes (for example, Wade, 2001). Recent evidence indicates that

alterations in transcription discriminate humans from other primates, not different genes (Enard et al., 2002). It appears that the genes may well be very similar in humans and other animals, but they are used or transcribed differently and the way they are transcribed determines the differences between humans and other animals (Hood and Galas, 2003; Levine and Tjian, 2003). Therefore, transcriptional regulation may be a key in mammalian aging, as already suggested (Roy et al., 2002; Hood, 2003).

Transcriptional regulation is a complex process that only recently began to be understood (reviewed in Arnone and Davidson, 1997; Fickett and Wasserman, 2000). Importantly, transcriptional regulation is also digital in nature. It is located in the non-coding genetic sequence, largely in the form of *cis*-regulatory sequences that are critical in development and cellular differentiation. These regulatory sequences are specific targets of transcription factors (TFs) that, together with proteins that bind them, control gene activity. Since transcriptional regulation is located in the genetic information, bioinformatics is also attempting to tackle regulatory networks.

One promising computational tool for the identification and characterization of regulatory regions is phylogenetic footprinting (reviewed in Hardison et al., 1997; Pennacchio and Rubin, 2001; Ureta-Vidal et al., 2003). As with comparative genomics, the principle behind phylogenetic footprinting is that regulatory regions tend to be conserved across different species. One example of how phylogenetic footprinting can be used involves the stem cell leukemia (*SCL*) gene. Briefly, clones containing the human, mouse, and chicken *SCL* gene were sequenced. Comparing human and mouse sequences revealed several homology peaks, a subset of which corresponded to previously known *SCL* enhancers. Chicken sequences corresponding to human–mouse peaks identified a conserved non-coding region with no known functions. Using a transgenic reporter assay, it was possible to characterize that region as a new neural transcriptional enhancer (Gottgens et al., 2000). In other experiments, phylogenetic footprinting has been used in the discovery of mammalian regulatory elements of genes such as Bruton's tyrosine kinase, β-globin (Hardison et al., 1997), and interleukins 4, 5, and 13 (Loots et al., 2000). With the availability of more genomes, phylogenetic footprinting should gradually become more powerful (for example, Hardison et al., 1997).

Since TFs bind specific sequences, it is possible to find putative transcription factor binding sites (TFBS) in genome sequence through computational algorithms (reviewed in Stormo, 2000; Ohler and Niemann, 2001; Pennacchio and Rubin, 2001). One major problem with putative TFBS identified by computational methods, and to a lesser degree, in phylogenetic footprinting, is the presence of false positives. Experimental verification is often necessary. In addition, TFs in metazoans often interact with each others, making it difficult to find the complete set of regulatory interactions. To minimize this problem, more efficient algorithms, or combination of algorithms, can search clusters of TFBS or TFBS conserved between mouse and human sequence (Wagner, 1999; Levy and Hannenhalli, 2002). Although these methods are becoming increasingly efficient, the true power of putative TFBS detection emerges when taken together with other types of data, as will become apparent ahead.

The principle in studying transcriptional regulation may also involve a comparative biology approach in an attempt to understand why different mammals age at different rates. Only instead of searching genes, we would be searching regulatory sequences. The importance of transcriptional regulation should not be underestimated. For example, as much as 40% of human TFBS are not functional in rodents (Dermitzakis and Clark, 2002). Using

phylogenetic footprinting it may be possible to find the TFs that control key genes involved in aging while detecting putative TFBS may help us gather hints about mechanisms of aging. If indeed the differences in rate of aging are due to subtle transcriptional differences amongst mammals, then the study of transcriptional proteins and regulatory sequences will be relevant for aging research. In other words, if aging and cancer can be seen as the corruption of the genetic program, then the study of transcriptional regulation might allow researchers to understand why it becomes corrupted with age.

2.3. DNA microarrays

DNA microarrays based on the quantification of mRNA levels are a growing technology (reviewed in Lockhart and Winzeler, 2000). Although the correlation between mRNA and protein levels is not always linear (Gygi et al., 1999), DNA microarrays have the technological edge over proteomics and protein microarrays (reviewed in Pandey and Mann, 2000; MacBeath, 2002) due to their capacity to produce large amounts of gene expression data from different conditions in relatively little time. The principle is that changes in mRNA levels between different conditions also reflect changes in the system under study and often at a protein level.

Studies of complex processes in yeast have showed the power of DNA microarrays by measuring mRNA levels for practically every yeast gene. One experiment attempted to relate the gene expression program to the sequence of events in sporulation by measuring gene expression at $t = 0, 0.5, 2, 5, 6, 7, 9$, and 11.5 h during sporulation. By grouping genes according to their expression profiles during sporulation, it was possible to find functional links and provide clues about the function of previously uncharacterized genes. Clustering of gene expression patterns allowed hints about the function of hundreds of genes and, for example, a 10-fold increase in the number of identified genes that participate at the middle stages of sporulation. DNA microarrays have also become crucial to understand transcriptional regulation. In the same experiment, consensus sequences for the *USR1* transcriptional regulator were found by computational approaches upstream of the start codon of many genes clustered according to expression patterns. These results allowed an association between a temporal pattern of gene expression, a stage in sporulation, and a transcriptional regulator with its respective regulatory sequence. The role of the Ndt80 TF was also investigated through either ectopically expressing Ndt80 or eliminating it. In theory, this approach can be used to find nearly every gene regulated by a TF under the experimental conditions used, though false negatives may exist since genes are often controlled by multiple TFs. Together with analysis of putative Ndt80 binding sites, this work allowed the identification of several genes presumably controlled by Ndt80 (Chu et al., 1998).

Clustering genes with similar expression patters with age may allow researchers to find functional links. Yet the objective in applying DNA microarrays to study aging is to relate the gene expression program with the sequence of events of the aging process in hope that will allow us to determine the regulation of aging. One major problem is that, unlike yeast sporulation, the aging process may not be programmed. Others have suggested that age-related changes in gene expression are deleterious for they represent a shift from what is assumed to be a young pattern (for example, Lee et al., 1999; Jiang et al., 2001). Yet instead of indicating causes of aging, gene expression profiles as animals age may represent the

tissue's response to aging. For instance, if aging derives from damage accumulation, then gene expression patterns will change with age as a response to damage. Therefore, due to the unique and unclear basis of aging, data obtained from gene expression patterns must be carefully interpreted. For example, one study on cellular senescence found disparate changes in gene expression between two different cell types despite the same telomere-dependent mechanism. Yet up-regulated senescence-specific genes showed chromosomal clustering between both cell types, leading to the suggestion that chromatin changes are involved in cellular senescence (Zhang et al., 2003). It is clear that changes occur in human cells as we age, but gene expression profiles may only reflect a response to those changes rather than the changes themselves.

One important application of DNA microarrays for the study of aging would be to find markers of aging. For example, DNA microarray analysis of malignant lymphomas has proven useful in detecting genes that can be employed to classify and predict the prognosis of tumors (reviewed in Schwaenen et al., 2003). Finding gene expression patterns capable of serving as indicators of how aged an animal or a human is would prove useful for research. For instance, if the recent advances in sequencing power are anything to go by, then in 10 years we may be able to sequence a human genome in 1 day (for example, Hood, 2003). Having a method to measure biological aging would allow us to calculate the pace of aging in individuals and possibly find clues about the genetic players that modulate the aging process. Even now, such method would allow us to compare, for example, single-nucleotide polymorphisms amongst individuals with different paces of aging.

Other technical variants and applications of DNA microarrays exist (reviewed in Lockhart and Winzeler, 2000). One powerful technique for the understanding of transcriptional regulation is genome-wide location analysis (reviewed in Wyrick and Young, 2002). In yeast, the technique involves a DNA microarray with the complete set of yeast intergenic regions. DNA enriched with an antibody against the TF of interest is labeled and hybridized against the microarray revealing which promoters, and often which genes, are regulated by the TF. This technique has also been used to study the E2F TF family in human cells using a selection of about 1200 human promoter sequences. Several genes were identified as potentially activated by E2F, including many genes with no previous connection to E2F (Ren et al., 2002).

Applying genome-wide location analysis to study aging is a potentially powerful approach not only in understanding cellular senescence but also the aging process. It could shed light on what genes are regulated by TFs suspect of participating in aging—for example, the redox-regulated TFs such as AP-1, Sp1, and NF-κB (Lavrovsky et al., 2000). Furthermore, it could help elucidate the roles of these and other TFs at different ages. For instance, understanding which genes are activated at different ages by *p53* could help clarify *p53*'s role on aging and cancer mechanisms. Of course, we may face the same problem of finding responses to aging rather than causes. If genes do not exist to cause aging, then neither do regulatory sequences. Yet these respond to stimulus and if these responses change during aging then finding them may help understand what causes aging.

When applying DNA microarrays to the study of aging we face the problem of how to focus on human aging. It appears difficult to obtain data on gene expression at different ages for humans using longitudinal studies. One idea could be to use cross-sectional studies. Although such studies are obviously biased by the variability amongst individuals (Hofer and Sliwinski, 2001), if done on a sufficient number, cross-sectional studies may prove useful

(Weindruch et al., 2002). Of course, cross-sectional studies appear the only solution to study long-lived animals, such as whales or turtles, and they have already been employed to study, for example, rhesus monkeys (Kayo et al., 2001). Therefore, at least using longitudinal studies, we are likely to depend on model organisms.

Studying aging in animals has the additional problem of having to deal with several different cells types—over 200 in vertebrates (Alberts et al., 1994). As far as we know, aging can be caused simultaneously in all tissues, it can be a result of changes in a particular organ, or it can even involve different genetic programs at different tissues. Indeed, previous studies already indicated that gene expression profiles are specific for the aging process of each organ (reviewed in Weindruch et al., 2002). Interpreting gene expression profiles in multicellular organisms is a difficult problem, which will rely on having a substantial amount of data from different tissues to allow researchers to isolate which age-related gene expression changes are passive effects, causes, or responses to aging. Techniques for measuring tissue-specific gene expression have been applied to *Caenorhabditis elegans* (reviewed in Reinke, 2002) and could be employed in mammals to discriminate tissue-specific age-related gene expression changes. One example is the use of laser capture microdissection. Lastly, different forms of data-analysis, such as chromosomal clustering (Zhang et al., 2003), may reveal information that is otherwise unobvious.

In conclusion, DNA microarrays may provide much data and even information for the study of aging. The great advantage of DNA microarrays is that it is not necessary to know which genes are important in the process under study. Large amounts of age-related gene expression data from different tissues of, for instance, mice, would prove an invaluable resource for the study of aging. At present, microarray databases, such as Standford's (Sherlock et al., 2001), already contain a few datasets regarding aging (reviewed in Jennings and Young, 1999). With time, researchers can hope to have the age-related transcriptome of different tissues from several model organisms.

3. Modeling human aging

Computational methods offer several data-mining tools, as previously mentioned, and an exponentially increasing amount of data (Table 1). To understand biological systems, the amount of data needed is colossal because biological systems are intrinsically complex with diverse, often multifunctional, elements that interact in non-linear ways (Toussaint et al., 1991). In turn, data and information can be analyzed and interpreted by way of computer and mathematical methods to create models that are easier to study than biological systems. Ultimately, these models allow us to increase our knowledge about the process under study. If we could simulate, for instance, cellular senescence in silico, it would be a major development. Yet the grail of gerontology is the reconstruction of the genetic network of human aging: the identification of the causal structure of the aging process's gene network.

3.1. System structure and identification

Uncovering a complex process such as human aging will depend on the employment of both computational tools and experimental approaches. The integration of these two forms

Table 1
List of major databases and bioinformatics websites and selection of databases that may be useful to gerontologists

Name	Website
Major databases and bioinformatics websites	
EMBL	http://www.embl-heidelberg.de/
EMBL's genome browser	http://www.ensembl.org/
European Bioinformatics Institute	http://www.ebi.ac.uk/
NCBI	http://www.ncbi.nlm.nih.gov/
TIGR	http://www.tigr.org
USCS Genome Bioinformatics	http://genome.ucsc.edu/
Databases useful to gerontologists	
Baltimore Longitudinal Study of Aging	http://blswww.grc.nia.nih.gov/
BodyMap	http://bodymap.ims.u-tokyo.ac.jp/
GeneCards	http://bioinformatics.weizmann.ac.il/cards/
HPRD	http://www.hprd.org
Protein Data Bank	http://www.rcsb.org/pdb/
SAGE KE database of genes/interventions	http://sageke.sciencemag.org/cgi/genesdb
Swiss-Prot	http://www.us.expasyorg/sprot/
Telomere Database	http://www.genlink.wustl.edu/teldb/index.html

of information requires a systems biology approach (reviewed in Ideker et al., 2001a; Kitano, 2002a,b). Systems biology is based on information, e.g. the quantification of a gene product, obtained from a given biological system under different genetic and/or environmental conditions. Information is then mathematically treated to construct a model that explains the system. For example, insights into the regulation of galactose use (*GAL*) in yeast have been obtained through systems biology. To study the *GAL* pathway, mRNA and protein data were obtained from yeast strains under different environmental and genetic conditions. Exemplifying, strains were examined each with a different *GAL* gene deleted. Using previously known protein–protein and protein–DNA interactions, previous models of the *GAL* pathway, and the new data, it was possible to build an integrated physical-interaction network for over 300 genes. Several putative interactions were also identified through gene expression analysis combined with TFBS scans or simply by searching genes with correlated expression profiles. New regulatory phenomena were also proposed, some of which later experimentally verified (Ideker et al., 2001b).

Applying systems biology to study human aging is not straightforward. The accuracy and detail of a model is dependent on how much data we can gather in how many different circumstances (Selinger et al., 2003). From a mathematical perspective, we can imagine modeling aging as mapping the rules that make a young organism old. If a large number of genes are involved in the process, as it appears likely, then a large number of measurements are necessary to understand the rules governing those genes. The duration of aging even in animal models makes aging a difficult subject of study when compared to, for instance, yeast sporulation. In addition, a major limitation to study aging in humans or model organisms, such as mice, is the relatively small amount of ways to change aging (Table 2). Therefore, and since our understanding of the aging process is still in its infancy, modeling human aging will have to be accomplished in phases. The first phase must be the identification of the elements involved in the human aging process, possibly with knowledge about interactions

Table 2
Major perturbations of the aging phenotype in mice and potential perturbations of human aging

Name	Perturbation	Reference
Mice		
Ames dwarf	Homozygous mice show over 50% increases in life-span.	Brown-Borg et al., 1996
Caloric restriction	Delay of the aging process.	Weindruch and Walford, 1988
GHR (growth hormone receptor)	Increase in life-span of 40–50% in homozygous knock-outs.	Coschigano et al., 2000
Ghrhr (growth hormone releasing hormone receptor)	Life-span increase of about 20% in homozygous knock-out mice.	Flurkey et al., 2001
IGF-1R (insulin-like growth factor receptor)	Heterozygous mice live 26% longer than wild-type.	Holzenberger et al., 2003
klotho	Possible accelerated aging phenotype of homozygous knock-outs.	Kuro-o et al., 1997
p53	Heterozygous mutant mice display signs of accelerated aging.	Tyner et al., 2002
$p66^{shc}$	Roughly 30% increase in life-span in −/− mice.	Migliaccio et al., 1999
Snell dwarf mice	Life-span increase of 42% in homozygous mice.	Flurkey et al., 2001
Telomere dysfunction and ATM deficiency	Possible accelerated aging in double mutant mice.	Wong et al., 2003
urokinase-type plasminogen activator	Roughly 20% increase in life-span of transgenic mice.	Miskin and Masos, 1997
XPD (xeroderma pigmentosum, group D)	Possible accelerated aging phenotype due to homozygous mutation.	de Boer et al., 2002
Humans		
CKN1 (Cockayne Syndrome Type I)	Possible accelerated aging due to mutation.	Henning et al., 1995
WRN (Werner Syndrome gene)	Premature aging due to recessive mutation.	Yu et al., 1996
LMNA (lamin A)	Possible premature aging due to dominant mutation.	Eriksson et al., 2003

between them. In addition, we must attempt to learn more about the structure of the aging process. For instance, does aging derive from damage accumulation, programmed gene expression changes, changes in DNA structure, or some other process?

Experimental approaches have already proven useful in identifying a few genes that may be involved in human aging. The new computational tools described earlier will be determinant to find novel genes involved in aging. DNA microarrays, in particular, are a powerful approach. For example, studying cellular processes such as stress response and DNA repair may help us gather clues about functional interactions between proteins suspect of being involved in aging (Table 2) and previously uncharacterized proteins. If indeed certain progeroid syndromes in both mice and men are cases of accelerated aging, then finding new functional links involving these proteins is a promising approach. Also, large-scale gene expression profiles of aging animals may help clarify the structure of the aging process, something that has so far eluded researchers.

3.2. System-level perturbations in model organisms

As we identify the elements of the aging process we can attempt to predict its progress by perturbing each component of the system. Due to the duration of human aging, model organisms will play a critical role. Namely, perturbations of aging in model organisms by, for instance, genetic interventions will be crucial to obtain the amount of data necessary to understand aging. Another type of perturbation results from evolution and the way several species have different rates of aging.

The major problem of using animal models is that the genetics of aging in model organisms may or may not be similar to the genetics of aging in humans. For example, mutations in the mouse homologue to the Werner's syndrome gene have no visible effect on their aging process (Lombard et al., 2000; Wang et al., 2000). One critical paradox in studying aging is that as we move across evolution in search of models closer to humans, these have increasingly longer life spans, making them increasingly more difficult and expensive to study. For example, studying genetic perturbations is much easier in *C. elegans* than in mice, but genes found in mice have more chances of being successfully extrapolated to humans than genes found in *C. elegans*. Although studies in, for instance, invertebrates will continue to yield putative genes involved in human aging (for example, Tower, 2000; Hekimi and Guarente, 2003), we must have a coherent view of aging mechanisms in several species, including mammals, before we can extrapolate conclusions to human aging.

Model organisms have already given useful insights on biological complexity and on the organization and dynamics of the aging process. Some successful work using DNA microarrays has been done in *Drosophila melanogaster* (for example, Zou et al., 2000), *C. elegans* (for example, Lund et al., 2002), and mice (for example, Lee et al., 1999, Cao et al., 2001; Jiang et al., 2001; Miller et al., 2002). Interestingly, some of the studies employing DNA microarrays on mice also studied caloric restriction and/or life-extending mutations (for example, Lee et al., 1999; Cao et al., 2001; Miller et al., 2002). If the mechanisms of aging are similar amongst mammals, as it may be the case (de Magalhaes and Toussaint, 2002), then mammalian aging is a combination of the basic mechanisms of aging, metabolic rates, and species-specific traits. Gene expression studies comparing different primates as well as different rodents have already been used to identify species-specific patterns of expression (Enard et al., 2002). Performing such studies using a temporal resolution of aging, at least in rodents, would allow us to determine the progression of gene expression as animals age, thus helping us find common responses to aging or even causative factors in mammalian aging.

If aging and development are two independent processes (for example, Miller, 1999), it is also important to identify the changes in animals before vulnerability starts to increase. In other words, to identify what changes occur prior to the sexual peak that make aging commence. If we aim to develop ways to stop aging, then this is the question we must ask, not what changes drive senescence in adulthood. Therefore, gene expression changes prior to the increase of age-related vulnerability may reveal important clues about the aging process, though care must be taken to discriminate such changes from developmental changes.

Although we focus on mRNA quantification and genomics, other data can be gathered such as protein studies (for example, MacBeath, 2002), metabolic fluxes (for example, Nielsen, 1998; Strohman, 2002), and even the study of age-related changes or clinical

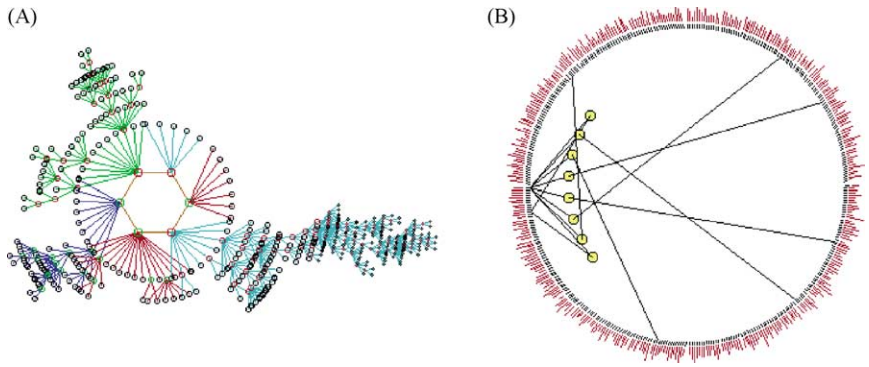

Fig. 1. (A) Basin of attraction of a random Boolean network with six attractors. The attractors (center) are the outcome of the system—for instance, an attractor can be a step in cellular differentiation, a phenotype, even a disease—while the wiring represents biological connections with their own rules describing the level of interaction. Each state can represent, for example, gene activity. Particularly useful for modeling cellular or even yeast cell cycles, Boolean networks also take into account the robustness of biological systems where small perturbations to one state may not affect the general outcome, which is in accordance to what we know about aging—for instance, many diseases accelerate mortality without affecting rate of aging. (B) Cellular automata, or cells, are computer simulations that try to model nature and life by employing a single set of rules. This image shows the network wiring between two time-steps for a large number of cells ($n = 400$, on the perimeter), highlighting just one cell. Although these methods are somewhat rough ways of portraying life, they can be much useful in understanding the dynamics and organization of complex genetic networks such as aging. Images generated using DDLAB by Andy Wuensche: http://www.santafe.edu/~wuensch/.

features of aging (for example, Mitnitski et al., 2002). The aim is to obtain as much data as possible under as many different conditions to help explain the observed differences. For example, metabolic control analysis may be applied to study metabolic fluxes in animals under caloric restriction and ad libitum-fed controls. Studying what metabolites differ between controls and caloric restricted mice may contribute to our understanding of aging.

3.3. Reconstructing the genetic network of human aging

Eventually, we would like to create models that allow us to control the aging process. Due to the limitations on gathering data about aging, great accuracy in describing human aging will be impossible in the near future. Therefore, we propose Boolean networks to define the genetic network of human aging. Simple mathematical descriptions such as Boolean networks classify gene interactions as 1 or 0 (Kauffman, 1993). Given the large amounts of genes presumably involved in a complex process such as aging, Boolean networks offer a simple and accessible way to model the genetic network of aging. Although they fail to take into account genes at intermediary levels and are sometimes seen by biologists as inaccurate, Boolean networks make it easy to model complex networks from large amounts of data (Fig. 1 for an example). Given the complexity of the aging phenotype, Boolean networks appear as a realistic way to simulate the genetic network of aging within a reasonable future. A Boolean model of human aging would already be incredibly useful for medicine in predicting potential anti-aging interventions.

Several algorithms have been proposed to reconstruct the genetic network of complex processes based on DNA microarray data (D'Haeseleer et al., 2000; Wahde and Hertz, 2000; Wagner, 2001; de la Fuente et al., 2002; Yeung et al., 2002). Unfortunately, these statistical approaches require amounts of data unrealistic for the present status of aging research. In addition, DNA microarray data will probably be obtained from model organisms, not humans. Therefore, the greatest challenge in reconstructing the genetic network of aging will be to integrate data coming from different sources, e.g. DNA microarrays and mutagenesis experiments in model organisms, cellular senescence, and genomics, into one coherent framework, as suggested by others (Jazwinski, 2002; Kirkwood et al., 2003; Hood, 2003).

As mentioned, gene expression data must be combined with other sorts of information, e.g. TFBS prediction, in order for us to understand a system. For example, in one experiment, regulatory motif pairs were used together with gene expression data to study the synergism between them. This approach allowed the identification of novel motif combinations regulating transcription and the modeling of regulatory networks in yeast (Pilpel et al., 2001). If indeed transcriptional regulation plays a key role in aging, then both data-mining strategies to find the relevant transcriptional information and the integration of DNA microarray data from model organisms, such as mice, are crucial. In another example of data integration, researchers combined protein information from experimental data, phylogenetic profiles, correlated mRNA expression levels, and patterns of domain fusions—involving the "Rosetta stone" analogy method—for practically all *S. cerevisiae* proteins. The result was the discovery of 93,000 functional links between proteins which allowed them to assign function to a previously uncharacterized protein family, a yeast homologue of human colon-cancer genes, and the yeast prion Sup35 (Marcotte et al., 1999b). In the end, modeling human aging will require multiple approaches with several feedback loops. To reconstruct the genetic network of human aging we need a strategy integrating the discovery of the genes involved with systematic perturbations of aging (Fig. 2).

4. Conclusion: is it possible to reverse engineer human aging?

Reverse engineering is "the process of analyzing a subject system to identify the system's components and their interrelationships and create representations of the system in another form or at a higher level of abstraction." (Chikofsky and Cross, 1990). To reverse engineer human aging would be to reconstruct the genetic network of aging; to find the mechanisms by which a human becomes old and find how to delay and perhaps even reverse the aging process. In practice, reverse engineer of aging would allow us to predict which genes actively regulate rate of aging and eventually what genes could we target to delay human aging and postpone age-related pathology and degeneration.

In theory, it is possible to reverse engineer a complex process (D'Haeseleer et al., 2000). Yet given the, previously mentioned, limitations in studying human aging, we find it unlikely that the genetic network of aging will be understood in detail within a near future. The number of experiments required to fully understand a complex phenotype such as aging is at present beyond our technology (Wagner, 2001; Krupa, 2002). Yet just finding a few more genes involved in aging would be a major breakthrough; as it would be to understand the structure of the aging process. Indeed, others have claimed long ago how the goal of

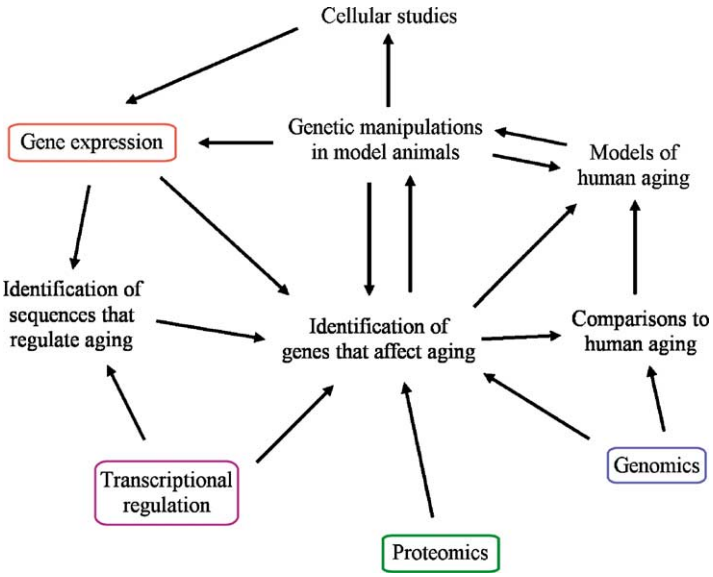

Fig. 2. Understanding the genetic network of human aging is based on two general phases: (1) identifying the genetic players involved; (2) systematically perturbing aging and elements/pathways suspect of being involved in aging of model organisms by, for instance, genetic interventions (for example, Case, 2003); the integration of results allows us to continually improve our models. A variety of experimental and computational approaches can be used for both aims. Through genetic manipulations, gene expression studies can help discover novel genes involved in aging as well as clarify the role of genes under study. Databases too can be used by, for instance, comparative genomics, to find potential genetic players affecting aging. Comparative genomics can also be used to assess if genes found in model organisms may play a role in human aging. Cellular studies can be useful, for instance, in conjunction with animal models to investigate specific cellular processes. Lastly, studies on transcriptional regulation may be employed to investigate the regulatory signals affecting the genes in question as well as find new elements involved in aging. Color indicates large-scale databases. Arrows indicate the flow of information.

gerontology should be to discriminate causes from effects of aging and find the one or few physiological processes that control aging (Medawar, 1955). Although aging is a complex process that involves many genes and pathways, different genes influence aging, directly or indirectly, in different ways. In fact, the recent history of the analysis of complex processes shows that even in the most complex of processes we are likely to find key controlling nodes (Risch, 2000; Kitano, 2002b). Both in practice (Migliaccio et al., 1999) and in theory (Wagner, 2001), it appears possible to locate such nodes in human aging. Even if we cannot reconstruct the entire genetic network of aging, locating a few key regulatory processes would be invaluable. For example, if we could develop a draft Boolean model of what genes determine rate of aging in primates, it would allow us a better understanding of human aging and possibly the identification of potential therapeutic targets; optimistically, such findings would open the path for the development of ways to delay age-related pathology and senescence.

New tools provide new goals. Until recently, it was unthinkable to attempt to identify all genes involved in aging. Yet the modern high-throughput technologies allow us to consider the possibility of defining most if not all players involved in aging as well as how they

interact to form the aging process. Eventually, we will be able to predict the effect of each gene on the aging process.

Acknowledgements

J.P. de Magalhães is funded by the Fundação para a Ciência e a Tecnologia, Portugal, and O. Toussaint is a Research Associate from the FNRS, Belgium.

References

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J.D., 1994. Molecular Biology of the Cell, third ed. Garland Publishing, New York/London.

Aravind, L., Walker, D.R., Koonin, E.V., 1999. Conserved domains in DNA repair proteins and evolution of repair systems. Nucl. Acids Res. 27, 1223–1242.

Arnone, M.I., Davidson, E.H., 1997. The hardwiring of development: organization and function of genomic regulatory systems. Development 124, 1851–1864.

Austad, S.N., 1997. Comparative aging and life histories in mammals. Exp. Gerontol. 32, 23–38.

Bronikowski, A.M., Alberts, S.C., Altmann, J., Packer, C., Carey, K.D., Tatar, M., 2002. The aging baboon: comparative demography in a non-human primate. Proc. Natl. Acad. Sci. U.S.A. 99, 9591–9595.

Brown-Borg, H.M., Borg, K.E., Meliska, C.J., Bartke, A., 1996. Dwarf mice and the ageing process. Nature 384, 33.

Cao, S.X., Dhahbi, J.M., Mote, P.L., Spindler, S.R., 2001. Genomic profiling of short- and long-term caloric restriction effects in the liver of aging mice. Proc. Natl. Acad. Sci. U.S.A. 98, 10630–10635.

Case, C.C., 2003. Transcriptional tools for aging research. Mech. Ageing Dev. 124, 103–108.

Chikofsky, E.J., Cross II, J.H., 1990. Reverse engineering and design recovery: a taxonomy. IEEE Software 7, 13–17.

Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P.O., Herskowitz, I., 1998. The transcriptional program of sporulation in budding yeast. Science 282, 699–705.

Comfort, A., 1964. Ageing: The Biology of Senescence. Routledge & Kegan Paul, London.

Coschigano, K.T., Clemmons, D., Bellush, L.L., Kopchick, J.J., 2000. Assessment of growth parameters and life span of GHR/BP gene-disrupted mice. Endocrinology 141, 2608–2613.

D'Haeseleer, P., Liang, S., Somogyi, R., 2000. Genetic network inference: from co-expression clustering to reverse engineering. Bioinformatics 16, 707–726.

de Boer, J., Andressoo, J.O., de Wit, J., Huijmans, J., Beems, R.B., van Steeg, H., Weeda, G., van der Horst, G.T., van Leeuwen, W., Themmen, A.P., Meradji, M., Hoeijmakers, J.H., 2002. Premature aging in mice deficient in DNA repair and transcription. Science 296, 1276–1279.

de la Fuente, A., Brazhnik, P., Mendes, P., 2002. Linking the genes: inferring quantitative gene networks from microarray data. Trends Genet. 18, 395–398.

de Magalhaes, J.P., Toussaint, O., 2002. The evolution of mammalian aging. Exp. Gerontol. 37, 769–775.

de Magalhaes, J.P., 2003. Is mammalian aging genetically controlled? Biogerontology 4, 119–120.

Dermitzakis, E.T., Clark, A.G., 2002. Evolution of transcription factor binding sites in mammalian gene regulatory regions: conservation and turnover. Mol. Biol. Evol. 19, 1114–1121.

Enard, W., Khaitovich, P., Klose, J., Zollner, S., Heissig, F., Giavalisco, P., Nieselt-Struwe, K., Muchmore, E., Varki, A., Ravid, R., Doxiadis, G.M., Bontrop, R.E., Paabo, S., 2002. Intra- and inter-specific variation in primate gene expression patterns. Science 296, 340–343.

Eriksson, M., Brown, W.T., Gordon, L.B., Glynn, M.W., Singer, J., Scott, L., Erdos, M.R., Robbins, C.M., Moses, T.Y., Berglund, P., Dutra, A., Pak, E., Durkin, S., Csoka, A.B., Boehnke, M., Glover, T.W., Collins, F.S., 2003. Recurrent de novo point mutations in lamin A cause Hutchinson–Gilford progeria syndrome. Nature 423, 293–298.

Fickett, J.W., Wasserman, W.W., 2000. Discovery and modeling of transcriptional regulatory regions. Curr. Opin. Biotechnol. 11, 19–24.
Finch, C.E., 1990. Longevity, Senescence, and the Genome. The University of Chicago Press, Chicago/London.
Flurkey, K., Papaconstantinou, J., Miller, R.A., Harrison, D.E., 2001. Lifespan extension and delayed immune and collagen aging in mutant mice with defects in growth hormone production. Proc. Natl. Acad. Sci. U.S.A. 98, 6736–6741.
George, J.C., Bada, J., WZeh, J., Brown, S.E., O'Hara, T., Suydam, R., 1999. Age and growth estimates of bowhead whales (*Balaena mysticetus*) via aspartic acid racemization. Can. J. Zool. 77, 571–580.
Gottgens, B., Barton, L.M., Gilbert, J.G., Bench, A.J., Sanchez, M.J., Bahn, S., Mistry, S., Grafham, D., McMurray, A., Vaudin, M., Amaya, E., Bentley, D.R., Green, A.R., Sinclair, A.M., 2000. Analysis of vertebrate *SCL* loci identifies conserved enhancers. Nat. Biotechnol. 18, 181–186.
Gygi, S.P., Rochon, Y., Franza, B.R., Aebersold, R., 1999. Correlation between protein and mRNA abundance in yeast. Mol. Cell Biol. 19, 1720–1730.
Hardison, R.C., Oeltjen, J., Miller, W., 1997. Long human–mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. Genome Res. 7, 959–966.
Hekimi, S., Guarente, L., 2003. Genetics and the specificity of the aging process. Science 299, 1351–1354.
Henning, K.A., Li, L., Iyer, N., McDaniel, L.D., Reagan, M.S., Legerski, R., Schultz, R.A., Stefanini, M., Lehmann, A.R., Mayne, L.V., et al., 1995. The Cockayne syndrome group A gene encodes a WD repeat protein that interacts with CSB protein and a subunit of RNA polymerase II TFIIH. Cell 82, 555–564.
Hofer, S.M., Sliwinski, M.J., 2001. Understanding ageing: an evaluation of research designs for assessing the interdependence of ageing-related changes. Gerontology 47, 341–352.
Holzenberger, M., Dupont, J., Ducos, B., Leneuve, P., Geloen, A., Even, P.C., Cervera, P., Le Bouc, Y., 2003. IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. Nature 421, 182–187.
Hood, L., Galas, D., 2003. The digital code of DNA. Nature 421, 444–448.
Hood, L., 2003. Systems biology: integrating technology, biology, and computation. Mech. Ageing Dev. 124, 9–16.
Ideker, T., Galitski, T., Hood, L., 2001a. A new approach to decoding life: systems biology. Annu. Rev. Genom. Hum. Genet. 2, 343–372.
Ideker, T., Thorsson, V., Ranish, J.A., Christmas, R., Buhler, J., Eng, J.K., Bumgarner, R., Goodlett, D.R., Aebersold, R., Hood, L., 2001b. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science 292, 929–934.
Jazwinski, S.M., 2002. Biological aging research today: potential, peeves, and problems. Exp. Gerontol. 37, 1141–1146.
Jennings, E.G., Young, R.A., 1999. Genome expression on the World Wide Web. Trends Genet. 15, 202–204.
Jiang, C.H., Tsien, J.Z., Schultz, P.G., Hu, Y., 2001. The effects of aging on gene expression in the hypothalamus and cortex of mice. Proc. Natl. Acad. Sci. U.S.A. 98, 1930–1934.
Kauffman, S.A., 1993. The Origins of Order, Self-Organization and Selection in Evolution. Oxford University Press, New York.
Kayo, T., Allison, D.B., Weindruch, R., Prolla, T.A., 2001. Influences of aging and caloric restriction on the transcriptional profile of skeletal muscle from rhesus monkeys. Proc. Natl. Acad. Sci. U.S.A. 98, 5093–5098.
Kirkwood, T.B., Boys, R.J., Gillespie, C.S., Proctor, C.J., Shanley, D.P., Wilkinson, D.J., 2003. Towards an e-biology of ageing: integrating theory and data. Nat. Rev. Mol. Cell Biol. 4, 243–249.
Kitano, H., 2002a. Looking beyond the details: a rise in system-oriented approaches in genetics and molecular biology. Curr. Genet. 41, 1–10.
Kitano, H., 2002b. Computational systems biology. Nature 420, 206–210.
Krupa, B., 2002. On the number of experiments required to find the causal structure of complex systems. J. Theor. Biol. 219, 257–267.
Kuro-o, M., Matsumura, Y., Aizawa, H., Kawaguchi, H., Suga, T., Utsugi, T., Ohyama, Y., Kurabayashi, M., Kaname, T., Kume, E., Iwasaki, H., Iida, A., Shiraki-Iida, T., Nishikawa, S., Nagai, R., Nabeshima, Y.I., 1997. Mutation of the mouse klotho gene leads to a syndrome resembling ageing. Nature 390, 45–51.
Lander, E.S., et al., 2001. Initial sequencing and analysis of the human genome: International Human Genome Sequencing Consortium. Nature 409, 860–921.
Lavrovsky, Y., Chatterjee, B., Clark, R.A., Roy, A.K., 2000. Role of redox-regulated transcription factors in inflammation, aging and age-related diseases. Exp. Gerontol. 35, 521–532.

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, 1390–1393.

Levine, M., Tjian, R., 2003. Transcription regulation and animal diversity. Nature 424, 147–151.

Levy, S., Hannenhalli, S., 2002. Identification of transcription factor binding sites in the human genome sequence. Mamm. Genome 13, 510–514.

Lockhart, D.J., Winzeler, E.A., 2000. Genomics, gene expression and DNA arrays. Nature 405, 827–836.

Lombard, D.B., Beard, C., Johnson, B., Marciniak, R.A., Dausman, J., Bronson, R., Buhlmann, J.E., Lipman, R., Curry, R., Sharpe, A., Jaenisch, R., Guarente, L., 2000. Mutations in the WRN gene in mice accelerate mortality in a *p53*-null background. Mol. Cell Biol. 20, 3286–3291.

Loots, G.G., Locksley, R.M., Blankespoor, C.M., Wang, Z.E., Miller, W., Rubin, E.M., Frazer, K.A., 2000. Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. Science 288, 136–140.

Lund, J., Tedesco, P., Duke, K., Wang, J., Kim, S.K., Johnson, T.E., 2002. Transcriptional profile of aging in *C. elegans*. Curr. Biol. 12, 1566–1573.

Luscombe, N.M., Greenbaum, D., Gerstein, M., 2001. What is bioinformatics? A proposed definition and overview of the field. Methods Inform. Med. 40, 346–358.

MacBeath, G., 2002. Protein microarrays and proteomics. Nat. Genet. 32 (Suppl.), 526–532.

Marcotte, E.M., Pellegrini, M., Ng, H.L., Rice, D.W., Yeates, T.O., Eisenberg, D., 1999a. Detecting protein function and protein–protein interactions from genome sequences. Science 285, 751–753.

Marcotte, E.M., Pellegrini, M., Thompson, M.J., Yeates, T.O., Eisenberg, D., 1999b. A combined algorithm for genome-wide prediction of protein function. Nature 402, 83–86.

Medawar, P., 1955. The definition and measurement of senescence, vol. I. In: Wolstenholme, G.E.W. (Ed.), Ciba Foundation Colloquia on Ageing. J. & A. Churchill, London, pp. 4–15.

Migliaccio, E., Giorgio, M., Mele, S., Pelicci, G., Reboldi, P., Pandolfi, P.P., Lanfrancone, L., Pelicci, P.G., 1999. The $p66^{shc}$ adaptor protein controls oxidative stress response and life span in mammals. Nature 402, 309–313.

Miller, R.A., 1999. Kleemeier award lecture: are there genes for aging? J. Gerontol. A: Biol. Sci. Med. Sci. 54, B297–B307.

Miller, R.A., Chang, Y., Galecki, A.T., Al-Regaiey, K., Kopchick, J.J., Bartke, A., 2002. Gene expression patterns in calorically restricted mice: partial overlap with long-lived mutant mice. Mol. Endocrinol. 16, 2657–2666.

Miskin, R., Masos, T., 1997. Transgenic mice overexpressing urokinase-type plasminogen activator in the brain exhibit reduced food consumption, body weight and size, and increased longevity. J. Gerontol. A: Biol. Sci. Med. Sci. 52, B118–B124.

Mitnitski, A.B., Graham, J.E., Mogilner, A.J., Rockwood, K., 2002. Frailty, fitness and late-life mortality in relation to chronological and biological age. BMC Geriatr. 2, 1.

Nielsen, J., 1998. Metabolic engineering: techniques for analysis of targets for genetic manipulations. Biotechnol. Bioeng. 58, 125–132.

Ohler, U., Niemann, H., 2001. Identification and analysis of eukaryotic promoters: recent computational approaches. Trends Genet. 17, 56–60.

Pandey, A., Mann, M., 2000. Proteomics to study genes and genomes. Nature 405, 837–846.

Pellegrini, M., Marcotte, E.M., Thompson, M.J., Eisenberg, D., Yeates, T.O., 1999. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. Proc. Natl. Acad. Sci. U.S.A. 96, 4285–4288.

Pennacchio, L.A., Rubin, E.M., 2001. Genomic strategies to identify mammalian regulatory sequences. Nat. Rev. Genet. 2, 100–109.

Pilpel, Y., Sudarsanam, P., Church, G.M., 2001. Identifying regulatory networks by combinatorial analysis of promoter elements. Nat. Genet. 29, 153–159.

Reinke, V., 2002. Functional exploration of the *C. elegans* genome using DNA microarrays. Nat. Genet. 32 (Suppl.), 541–546.

Ren, B., Cam, H., Takahashi, Y., Volkert, T., Terragni, J., Young, R.A., Dynlacht, B.D., 2002. E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. Genes Dev. 16, 245–256.

Risch, N.J., 2000. Searching for genetic determinants in the new millennium. Nature 405, 847–856.

Roy, A.K., Oh, T., Rivera, O., Mubiru, J., Song, C.S., Chatterjee, B., 2002. Impacts of transcriptional regulation on aging and senescence. Ageing Res. Rev. 1, 367–380.

Schwaenen, C., Wessendorf, S., Kestler, H.A., Dohner, H., Lichter, P., Bentz, M., 2003. DNA microarray analysis in malignant lymphomas. Ann. Hematol. 82, 323–332.

Selinger, D.W., Wright, M.A., Church, G.M., 2003. On the complete determination of biological systems. Trends Biotechnol. 21, 251–254.

Sherlock, G., Hernandez-Boussard, T., Kasarskis, A., Binkley, G., Matese, J.C., Dwight, S.S., Kaloper, M., Weng, S., Jin, H., Ball, C.A., Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D., Cherry, J.M., 2001. The Stanford Microarray Database. Nucl. Acids Res. 29, 152–155.

Stormo, G.D., 2000. DNA binding sites: representation and discovery. Bioinformatics 16, 16–23.

Strohman, R., 2002. Maneuvering in the complex path from genotype to phenotype. Science 296, 701–703.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. 22, 4673–4680.

Toussaint, O., Raes, M., Remacle, J., 1991. Aging as a multi-step process characterized by a lowering of entropy production leading the cell to a sequence of defined stages. Mech. Ageing Dev. 61, 45–64.

Tower, J., 2000. Transgenic methods for increasing *Drosophila* life span. Mech. Ageing Dev. 118, 1–14.

Tyner, S.D., Venkatachalam, S., Choi, J., Jones, S., Ghebranious, N., Igelmann, H., Lu, X., Soron, G., Cooper, B., Brayton, C., Hee Park, S., Thompson, T., Karsenty, G., Bradley, A., Donehower, L.A., 2002. *p53* mutant mice that display early ageing-associated phenotypes. Nature 415, 45–53.

Ureta-Vidal, A., Ettwiller, L., Birney, E., 2003. Comparative genomics: genome-wide analysis in metazoan eukaryotes. Nat. Rev. Genet. 4, 251–262.

Venter, J.C., et al., 2001. The sequence of the human genome. Science 291, 1304–1351.

Wade, N., 2001. Life Script: How the Human Genome Discoveries Will Transform Medicine and Enhance Health. Simon & Schuster, London.

Wagner, A., 1999. Genes regulated cooperatively by one or more transcription factors and their identification in whole eukaryotic genomes. Bioinformatics 15, 776–784.

Wagner, A., 2001. How to reconstruct a large genetic network from n gene perturbations in fewer than $n(2)$ easy steps. Bioinformatics 17, 1183–1197.

Wahde, M., Hertz, J., 2000. Coarse-grained reverse engineering of genetic regulatory networks. Biosystems 55, 129–136.

Wang, L., Ogburn, C.E., Ware, C.B., Ladiges, W.C., Youssoufian, H., Martin, G.M., Oshima, J., 2000. Cellular Werner phenotypes in mice expressing a putative dominant-negative human WRN gene. Genetics 154, 357–362.

Waterston, R.H., et al., 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520–562.

Wei, L., Liu, Y., Dubchak, I., Shon, J., Park, J., 2002. Comparative genomics approaches to study organism similarities and differences. J. Biomed. Inform. 35, 142–150.

Weindruch, R., Walford, R.L., 1988. The Retardation of Aging and Disease by Dietary Restriction. C.C. Thomas, Springfield.

Weindruch, R., Kayo, T., Lee, C.K., Prolla, T.A., 2002. Gene expression profiling of aging using DNA microarrays. Mech. Ageing Dev. 123, 177–193.

Wong, K.K., Maser, R.S., Bachoo, R.M., Menon, J., Carrasco, D.R., Gu, Y., Alt, F.W., DePinho, R.A., 2003. Telomere dysfunction and Atm deficiency compromises organ homeostasis and accelerates ageing. Nature 421, 643–648.

Wyrick, J.J., Young, R.A., 2002. Deciphering gene expression regulatory networks. Curr. Opin. Genet. Dev. 12, 130–136.

Yaspo, M.L., 2001. Taking a functional genomics approach in molecular medicine. Trends Mol. Med. 7, 494–501.

Yeung, M.K., Tegner, J., Collins, J.J., 2002. Reverse engineering gene networks using singular value decomposition and robust regression. Proc. Natl. Acad. Sci. U.S.A. 99, 6163–6168.

Yu, C.E., Oshima, J., Fu, Y.H., Wijsman, E.M., Hisama, F., Alisch, R., Matthews, S., Nakura, J., Miki, T., Ouais, S., Martin, G.M., Mulligan, J., Schellenberg, G.D., 1996. Positional cloning of the Werner's syndrome gene. Science 272, 258–262.

Zhang, H., Pan, K.H., Cohen, S.N., 2003. Senescence-specific gene expression fingerprints reveal cell-type-dependent physical clustering of up-regulated chromosomal loci. Proc. Natl. Acad. Sci. U.S.A. 100, 3251–3256.

Zou, S., Meadows, S., Sharp, L., Jan, L.Y., Jan, Y.N., 2000. Genome-wide study of aging and oxidative stress response in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. 97, 13726–13731.