Seminars in Cell & Developmental Biology 70 (2017) 177–189

Contents lists available at ScienceDirect

Seminars in Cell & Developmental Biology

journal homepage: www.elsevier.com/locate/semcdb

Review

What have we learned on aging from omics studies?

Alessandro Cellerino[a,b,*], Alessandro Ori[b,*]

[a] Laboratory of Biology (BIO@SNS), Scuola Normale Superiore, Pisa, Italy
[b] Leibniz Institute on Aging – Fritz Lipmann Institute (FLI), Jena, Germany

ARTICLE INFO

Article history:
Received 7 April 2017
Accepted 13 June 2017
Available online 16 June 2017

Keywords:
Aging
Genomics
Proteomics
Transcriptomics
Killifish
Nothobranchius furzeri
Pathways
Protein complex
Stoichiometry
Data integration
Proteostasis
Ribosome
Proteasome
Lysosome
RNAseq
Mitochondria

ABSTRACT

Aging is a complex process. Transcriptomic studies of the last decade have identified genes and pathways that are regulated during aging in multiple species and organs. Yet, since a manifold of pathways are regulated and the amplitude of regulation is often small, reproducibility across studies is moderate and disentangling cause-consequence relationships has proven challenging. Here, we review a number of consistent findings in the light of more recent, longitudinal studies and of studies combining transcriptomics and proteomics that identified deregulation of protein biosynthetic pathways as an early event and likely driver of aging.

Contents

* Corresponding authors.
E-mail addresses: alessandro.cellerino@sns.it (A. Cellerino), alessandro.ori@leibniz-fli.de (A. Ori).

http://dx.doi.org/10.1016/j.semcdb.2017.06.012
1084-9521/

1. Introduction

Despite multiple attempts, very few genetic variants have been associated with human longevity [1]. Indeed, twin studies have shown that longevity has low heritability (20–30%) [2]. This suggests that external factors influence the aging of an individual after its conception, starting from very early in life [3]. These factors include diet, environment, physical activity and the interaction with commensal and pathogenic microbiomes, all of which have been shown to influence healthspan in multiple model organisms and humans [4–7]. The life history of an individual can thus influence the rate of the aging process, and has detectable consequences at both physiological and molecular level.

Lifestyle factors can influence the physiological status of multiple organs, predominantly via epigenetic mechanisms. Indeed, epigenetic DNA methylation patterns seem to be more correlated with age than gene expression. This led to the development of the so-called "epigenetic clock" [8–10], a set of CpG islands whose methylation can predict the chronological age across human tissues and survival in longitudinal studies [11]. Epigenetic mechanisms of aging linked to modification of chromatin structure and their influence on gene expression and metabolism have been the subject of recent reviews [12–14] and will not be treated here. In the first part of this review, we will summarize the extensive information that was gathered from studies of genome-scale transcript regulation in aging tissues, while in the second part we will focus on post-transcriptional mechanisms of aging that are emerging from the combined application of genomic and proteomic approaches.

2. Aging as a complex process

Many reviews on the biology of aging contain at some point –mostly in the introductory phrases- a sentence stating that aging is a complex phenomenon. Under this definition, it is often understood that many different biological processes are involved in aging and it is not possible to pinpoint a single cause and molecular mechanism.

The development of so-called "omics" technologies brought about the expectation that an exhaustive molecular description of aging-regulated processes could have been possible, thereby shedding light on its mechanisms.

Since the first genome-scale microarray study on muscle aging in 1999 [15], a number of different papers described aging of various tissues in different species, employing microarrays, next-generation sequencing and mass spectrometry-based proteomics.

Although some level of consistency could be detected among these studies (see below), a reader not educated in quantitative biology would rather be surprised by the differences that at first sight seem to emerge by a comparison. Differences in the technological platforms being used in different studies as well as statistical and biological limitations contribute to these discrepancies.

Statistical limitations. All methods that were developed to analyze genome-scale data face the problem of data arrays, where the number of genes whose expression is quantified is in the order of 10^4 and the number of individual samples is in the order of a few units to tenths. This leads to: i) a severe problem of type I errors (false calling of differential-expression) due to the extent of *multiple statistical testing*, and ii) a significant problem of type II errors (failure to detect a true differential expression), due to the low statistical power of a typical genome-wide experiment.

The statistical tools employed for analysis of genome-scale expression data are tailored to contain type I errors at the expense of type II errors. This is a logical choice given the large number of comparisons possible and the risk that looser thresholds would result is a substantial proportion of false positives. However, a high rate of type II errors poses a hurdle in comparisons between datasets.

Biological limitations. A first biological issue is related to the amplitude of the gene expression changes observed during aging. Although aging is characterized by clear morphological and physiological changes, the associated changes in gene expression tend to be small. As an example, we show here RNA-seq data from brain of three different species: the short-lived fish *N. furzeri* [16], mouse [17] and human [18] (Fig. 1A). In the case of *N. furzeri*, 5 young samples were contrasted with 11 old samples [16]. In this comparison only ~20% of the 7863 differentially-expressed genes show an absolute change larger than 100% and half of the significant fold changes are smaller than 50% (Fig. 1A). A similar analysis of mouse and human data shows a similar pattern (Fig. 1A).

A second biological issue is related to inter-individual variability in expression, particularly in old individuals. It is well known that the variance of many physiological variables increases with age [19]. At least two causes can underlie the variability of the aging phenotypes: i) individual aging trajectories differ, and distinct aging-related pathways can be modulated to different extents across individuals. This can lead to discrepancy between the biological age of an individual and its chronological age, which is most often used to define sample groups in omics experiments. ii) The cellular composition of organs may change (e.g., due to inflammation and fibrosis, or changes of proportion between cell types, as known for the blood [20]), and these changes may affect different individuals to a varying extent. This clearly creates a problem for genome-scale studies that normally suffer from a limited amount of samples. Firstly, it decreases the power to detect changes. Secondly, the "old" samples from two different studies are randomly extracted sets from highly heterogeneous populations and may be considerably different. We illustrate here an increase in the coefficient of variation for gene expression in three datasets: whole brain of killifish and mouse, and human prefrontal cortex (Fig. 1B). It is also interesting that increased variability might be reduced in very old individuals [21]. Systematic studies with large number of samples are needed to better describe the general change of gene expression variance across tissues and the sources of increased variation.

All these difficulties notwithstanding, a number of conserved pathways and genes were identified by meta-analysis and will be presented below.

3. When is old enough? A link between development and aging

A further problem linked to genome-scale studies of aging is that, in many cases, and especially in animal studies, a simple pairwise comparison between "young" and "old" samples is performed. The choice of these two ages is to some extent arbitrary, and it is not without consequences: the young samples,