Journal of Statistics on Health Decision

EXTENDED ABSTRACT

P5

An exploratory analysis of aging gene expression levels using gene functions' combinations: As expected the heart has age!

Adelaide Freitas[1], Jorge Cabral[2], Adela Iutis[2], Joana Antão[2], Sérgio Marques[2], Tiago Costa[2], Gabriela Moura[3]

[1] Department of Mathematics, University of Aveiro. Center for Research & Development in Mathematics and Applications (CIDMA), University of Aveiro. Portugal.

[2] Department of Mathematics, University of Aveiro, Portugal.

[3] Department of Medical Sciences, University of Aveiro. Institute for Research in Biomedicine (iBiMED), University of Aveiro. Portugal.

Introduction

Aging is an important biological process and a known risk factor for many diseases in humans. Nevertheless, individuals of the same chronological age can vary in health status from fit to frail [1]. Aging gene expression (GE) signatures are very tissue specific [2]. At least 19 studies have profiled changes in GE with age in nine separate mouse tissues and found that similar classes of genes are regulated with age [3]. Genes involved in the inflammatory response and heat shock factors have been found to increase expression levels with age in mouse brain, muscle, and heart [4,5,6]. Studies of differentially expressed genes (DEG) are traditionally performed by the identification of DEG using multiple testing and then a functional analysis is executed. However, when the data set consists of a few samples, multiple testing does not tend to identify DEG.

Objective

In this work, based on cardiac tissue extracted from a small number of mice and where statistical inference did not detect DGE, an exploratory approach (using tools from R software) is proposed to detect potential DEG, starting with a partition of genes by their functionalities, and then evaluating GE levels by combination of their biological functionalities.

Methods

Expression levels, as the number of reads, of 20738 heart tissue genes from four 12-month-old mice (Mus musculus) and four 24-month-old mice were analyzed. Firstly, average GE levels by age group were plotted. Due to their different ranges among genes (e.g., high and low GE), all the subsequent analyses were conducted using expression levels standardized by gene. Groups of genes belonging to the same biological process were identified based on GroupGo functionality provided by ClusterProfiler R-function on org.Mm.eg.db.: a genome wide annotation for mouse. The level of specificity selected was 7 so as not to compromise the number of genes. The 12 functions more frequent were chosen. All the combinations (212+1) of these functions (present or absent) were generated. Functionality combinations without genes were removed remaining 64 of them (C1,…,C64). For each non-empty combination, a boxplot was constructed to visually explore the difference values (D) of the mean expression level between 12 and 24-month-old mice by gene. Mann-Whitney U test was used to detect significant differences by gene.

Results

Comparing the mean values of (non-standardized) expression levels of the 20738 genes by age group, a tendency of higher values for 12-month-old mice was displayed (Figure 1).

All the genes were clustered by functionality. From the 20 300 functions obtained, the following the 12 most frequent were selected: RNA metabolic process (17.70%), cellular protein modification process (16.37%), regulation of cellular macromolecule biosynthetic process (15.18%), regulation of RNA metabolic process (14.39%), RNA biosynthetic process (13.69%), nucleic acid-templated transcription

Keywords:
Gene expression level, GroupGo functionality.

Corresponding author:
Adelaide Freitas
adelaide@ua.pt

First published: 23 OCT 2020

(13.62%), transcription – DNA-templated (13.58%), regulation of RNA biosynthetic process (13.24%), regulation of nucleic acid-templated transcription (13.20%), regulation of transcription – DNA-templated (13.16%), regulation of cellular protein metabolic process (11.83%) and transcription by RNA polymerase II (9.65%).

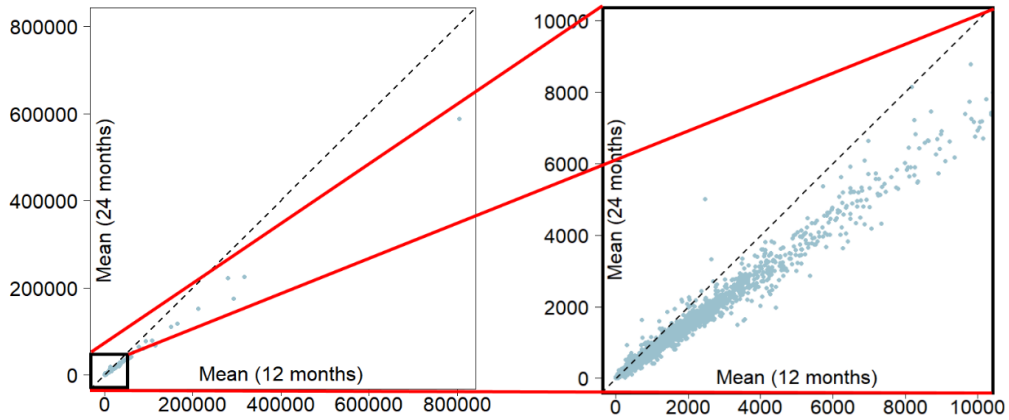

Figure 1 - Mean values of the expression level of 20738 genes by age group.

Differences among the median values of D across the 64 non-empty combinations are observed (Figure 2). The highest ones are in C17 with four genes (median=1.08) and in four combinations with single gene, C27, C59, C57 and C29 (median= -1.26, -1.21, 1.05, 1.01).

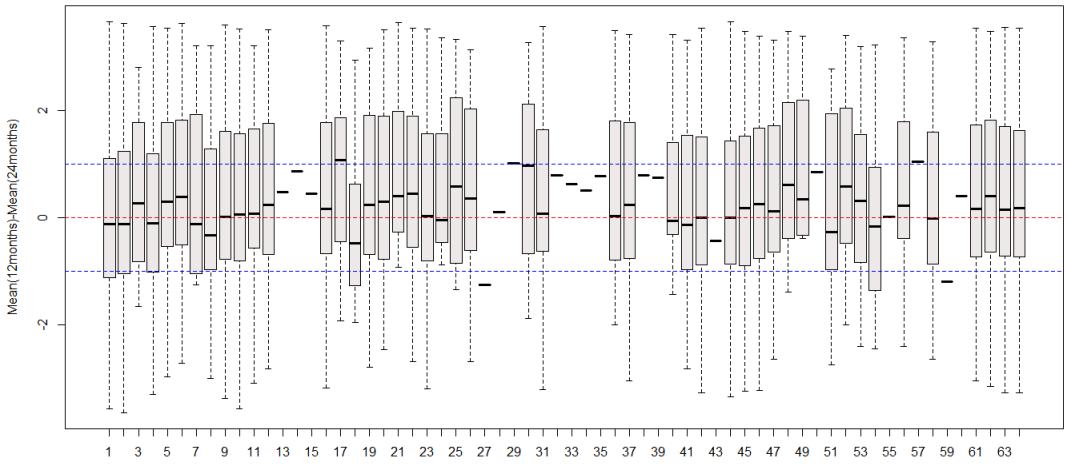

Figure 2 - Comparative boxplots of the D-values for all the 64 non-empty functionality combinations.

For the five combinations with the most extreme D-values, no statistically significant differences were detected between 12 and 24-month old mice. However, analyzing in detail the 8 expression levels for each four singular combinations, mice with expression levels consistent with the opposite age group were identified (mouse #5 in C27 and mouse #7 in C59). Moreover, differences became statistically significant (α=0.05) when the "abnormal" mouse was removed (Table 1).

Table 1 - Singular combinations and p-values from Mann Whitney U test.

		p-value	
Combination	**ID Gene**	**8 mice**	**7 mice (mouse removed)**
C27	78784	0.1038	0.0436 (#5)
C59	19733	0.1416	0.0477 (#7)
C29	68776	0.4857	0.4000 (#7 or #8)
C57	11727	0.3094	0.2118 (#7)

Due to limitation of space, exploratory analyses for the other non-singular functionality combinations are not herein presented.

Conclusion

The reduced sample size limits the ability to capture statistical regularity of gene expression, and consequently, the detection of statistically significant differences between age groups. Nevertheless, a new approach focused on functionality combinations (with single gene) provided the identification of (potential) DEG.

Acknowledgements

The Center for Research and Development in Mathematics and Applications (CIDMA) supported this work through the Portuguese Foundation for Science and Technology (FCT), references UIDB/04106/2020 and UIDP/04106/2020.

References

1. Jansen H, Moghtadaei M, Mackasey M, Rafferty S, Bogachev O, Sapp J et al. Atrial structure, function and arrhythmogenesis in aged and frail mice. Sci. Rep. 2017;7:44336. https://doi.org/10.1038/srep44336
2. Yang J, Huang T, Petralia F, Long Q, Zhang B, Argmann C et al. Synchronized age-related gene expression changes across multiple tissues in human and the link to complex diseases. Sci. Rep. 2015;5:15145. https://doi.org/10.1038/srep15145
3. Zahn J, Poosala S, Owen A, Ingram D, Lustig A, Carter A et al. AGEMAP: A gene expression database for aging in mice. PLoS Genet. 2007;3(11):e201. https://doi.org/10.1371/journal.pgen.0030201
4. Lee C, Weindruch R, Prolla T. Gene-expression profile of the ageing brain in mice. Nat. Genet. 2000;25:294–297. https://doi.org/10.1038/77046
5. Edwards M, Sarkar D, Klopp R, Morrow J, Weindruch R, Prolla T. Impairment of the transcriptional responses to oxidative stress in the heart of aged C57BL/6 mice. Ann. N. Y. Acad. Sci. 2004;1019:85–95. https://doi.org/10.1196/annals.1297.017
6. Dhahbi J, Tsuchiya T, Kim H-J, Mote P, Spindler S. Gene expression and physiologic responses of the heart to the initiation and withdrawal of caloric restriction. The Journals of Gerontology A. 2006;61:218–231. https://doi.org/10.1093/gerona/61.3.218