Peer Review File

Article information: http://dx.doi.org/10.21037/aob-20-79

Reviewer comments

Major comments:

Comment 1: p5r21 — "As seen in Fig 1C, ... MCHC and Platelet Count showed negative correlation to aging"
Such claims should be accompanied by a mention of Pearson's r and the corresponding P-value. On multiple occasions, the significance of the reported correlations is not visually evident. Figure 1C shows a flat line for MCHC. Perhaps, transforming the axes could make the figures more presentable.
Table 1.3, in which MCHC is mentioned once again, never explicitly reports Pearson's r with chronological age across the total sample. Moreover, the median values for all age groups are within the 1.5% deviation from the total sample median. Quartile values are extremely similar as well, so the reported P-value <0.001 is surprising. Table 3.3 shows that the correlation is significant only in the white people and is extremely small (-0.02). Such results hardly qualify as significant correlations.
Reply 1: Thanks for your suggestions. Pearson's r = -0.044, p=0.138, Spearman's r = -0.156, p<0.001. It is the result of different statistical methods. Because the data is not normal, we all use Spearman rank correlation to evaluate the relationship between the data.
Changes in the text: p5r21 and figure 1C.

Comment 2: p9r207-8 — "Samples were divided into two groups based on age, with >60 being Old group, and <60 being Young group"
Three age groups are used in other parts of the manuscript, so the rationale for using only two age groups in the DEG section.
Reply 2: This study explores the relationship between age and changes in gene expression. 60 years old is used as a routine to define the age limit of the elderly, and it is suitable as a grouping standard. At the same time, due to the small number of individuals in the sample, if divided into three groups, the middle-aged group is too small.
Changes in the text: p9r207-8.

Comment 3: p42-47 — Given the high variability between races and sexes, it is recommended to use a method of statistical analysis that can account for it. I suggest using mixed-effects linear models. Parametrizing sex and race as random effects should provide more reliable results on which blood parameters significantly change with age.
Reply 3: Thanks for your comments. It's really a good advice. The mixed linear model was used to eliminate the interference of gender and race. After analyzing the

relationship between age and biological indicators, we found some similar results and also found some different results. See the attached table (*The correlation of biological parameters to aging differed in a linear mixed*) for details.

Minor comments:

Comment 1: p2r54 — "...whole genome data of samples were used 54for enrichment analysis of differential gene sets"
This sentence is misleading. The authors carry out differential gene expression analysis, but this sentence reads as if the authors used genomic information, as in enrichment SNP analysis.
Reply 1: Your comments are really useful. We've modified our manuscript. You can see changes in p2r54.

Comment 2: p3r78 — While methylation aging clocks are indeed very accurate, the cited work by Hannum et al. is not the most accurate instance. Also, note that some solutions for transcriptomic aging clock yield similar performance (see 10.3389/fgene.2018.00242).
Reply 2: We're totally agree with your point. The sentences have been revised and you can see that in p3r78.

Comment 3: p11r233-52 — Some studies also report high individual variability in these discussed blood parameters, which may obscure the effect of aging. To be more specific, Lin et al. wrote: "It is unclear whether inter-individuals differences are due to individual's characteristics that remain stable with aging or result from the different rates of changes in different types of lymphocytes in across individuals" (see 10.1186/s12979-016-0079-7).
I wonder if the data the authors have allows them to test their findings in a longitudinal setting.
Reply 3: At present, our data comes from ImmPort, and does not mark the time of different individuals and research. If the subsequent database is updated, perhaps we can conduct further longitudinal research.

Comment 4: p14r307— "Further work will be needed to investigate their potential as biomarkers for aging"
In fact, this work has already been done. Mamoshina et al. published a hematologic aging clock that uses the specified blood parameters to predict chronological age (see 10.1093/gerona/gly005). It would be interesting to see if the results presented within this manuscript are in line with that article's findings.
Reply 4: That's really a good advice. We've checked the paper and found that the biomarkers they mentioned changed quite similarly to that in our manuscript. It provides us a proper direction for the next research.

Comment 5: p20 — Figure 1A is extremely compressed. Please make sure that the actual submission has a version with proper resolution.
Reply 5: Sorry about that. We don't know the reason but we'll send you the figure again.

Comment 6: p30-47 — In certain cases, it is not clear to what statistics the reported P-value relates to. For example, Table 1.5 is filled with median values; do P-values correspond to an equal medians test or to the omitted Pearson's r values?
Also, make sure asterisks are properly described. Do they mark the significant results?
Reply 6: Thanks for the comments. Because most of the data is not normally distributed, we used Spearsman's rank sum test to analyze the relationship between aging and different biological parameters.