Peer Community In Genomics

The same but different: How small scale hidden variations can have large effects

Francois Sabot *based on peer reviews by* ***Sandra Corjito*** *and 1 anonymous reviewer*

Diego A. Hartasánchez, Annamaria Kiss, Virginie Battu, Charline Soraru, Abigail Delgado-Vaquera, Florian Massinon, Marina Brasó-Vives, Corentin Mollier, Marie-Laure Martin-Magniette, Arezki Boudaoud, Françoise Monéger (2023) Expression of cell-wall related genes is highly variable and correlates with sepal morphology. bioRxiv, ver. 4, peer-reviewed and recommended by Peer Community in Genomics.
https://doi.org/10.1101/2022.04.26.489498

Submitted: 15 March 2023, Recommended: 14 September 2023

Cite this recommendation as:
Sabot, F. (2023) The same but different: How small scale hidden variations can have large effects. *Peer Community in Genomics*, 100243. 10.24072/pci.genomics.100243

Published: 14 September 2023

For ages, we considered only single genes, or just a few, in order to understand the relationship between phenotype and genotype in response to environmental challenges. Recently, the use of meaningful groups of genes, e.g. gene regulatory networks, or modules of co-expression, allowed scientists to have a larger view of gene regulation. However, all these findings were based on contrasted genotypes, e.g. between wild-types and mutants, as the implicit assumption often made is that there is little transcriptomic variability within the same genotype context.

Hartasànchez and collaborators (2023) decided to challenge both views: they used a single genotype instead of two, the famous *A. thaliana* Col0, and numerous plants, and considered whole gene networks related to sepal morphology and its variations. They used a clever approach, combining high-level phenotyping and gene expression to better understand phenomena and regulations underlying sepal morphologies. Using multiple controls, they showed that basic variations in the expression of genes related to the cell wall regulation, as well as the ones involved in chloroplast metabolism, influenced the global transcriptomic pattern observed in sepal while being in near-identical genetic background and controlling for all other experimental conditions.

The paper of Hartasànchez et al. is thus a tremendous call for humility in biology, as we saw in their work that we just understand the gross machinery. However, the Devil is in the details: understanding those very small variations that may have a large influence on phenotypes, and thus on local adaptation to environmental challenges, is of great importance in these times of climatic changes.

References:

Hartasánchez DA, Kiss A, Battu V, Soraru C, Delgado-Vaquera A, Massinon F, Brasó-Vives M, Mollier C, Martin-Magniette M-L, Boudaoud A, Monéger F. 2023. Expression of cell-wall related genes is highly variable and correlates with sepal morphology. bioRxiv, ver. 4, peer-reviewed and recommended by Peer Community in Genomics. https://doi.org/10.1101/2022.04.26.489498

Reviews

Evaluation round #1

DOI or URL of the preprint: https://doi.org/10.1101/2022.04.26.489498
Version of the preprint: 2

Authors' reply, 18 July 2023

Download author's reply
Download tracked changes file

Decision by Francois Sabot, posted 30 May 2023, validated 31 May 2023

Minor revisions

Dear Dr Hartasànchez and Dr Monéger

Your manuscript "Expression of cell-wall related genes is highly variable and correlates with sepal morphology" has been reviewed by two independent reviewers, and they found it of high interest and valuable for publication. However, they propose some minor corrections, in order to improve it a slight bit.

If you are able to perform these revisions as requested, I will accept it without anymore revisions.

Sincerely yours

Francois Sabot

Reviewed by anonymous reviewer 1, 12 May 2023

The manuscript entitled "Expression of cell-wall related genes is highly variable and correlates with sepal morphology", by Hartasanchez et al., addresses the question of whether the variability in gene expression in parallel with variability in organ morphology could be used to identify gene regulatory networks associated with organ size and shape determination.

The authors used a transcriptomic and morphological dataset obtained from 27 individual petals collected on 3 independent plants of the same background and grown on the same condition. Each sepal was imaged under a confocal microscope and allowing the evaluation of 7 morphological traits. RNA was extracted from the same samples to perform RNA-seq analysis. The analysis of the same individual sample for the morphological and expression analysis will allow to correlate gene regulatory networks with morphology.

The variability of sepal morphology has been explored by PCA and pairwise comparison and showed that variation in sepal morphology is mainly explained by variation in area and aspect ratio. The transcriptomic dataset was explored using the squared coefficient variation for each morphological parameter and for each of the genes expressed. Highly and Low Variation Genes (HVG and LVG) were selected and analyzed for their enrichment in particular biological functions. This analyze pointed that a part of the HVGs genes are enriched

in carbohydrates metabolic process and cell-wall organization or biogenesis genes. A co-expression analysis performed with all expressed genes lead to the identification of 16 modules of co-expressed genes, 5 of which correlated with at least one morphological parameter. These last modules are enriched in cell-wall related genes. Finally, GRN of cell-wall related genes from one of the module was reconstruct and highlighted relation between key genes known as transcription factors and involved in secondary cell-wall deposition.

The work described here

The approach used in this work is original and allow to explore the natural variability in morphology and gene expression across wild-type sepals. Generally, the diversity relative to organ morphology is studied by using mutants and/or GWAS approach which exploit the variability between genotypes and potentially involving several regulatory pathways. The work presented here shows that even if many precautions are taken during the sampling, a variability of morphology and gene expression is observed for the same organ collected on the same condition. However, the authors propose a method to exploit this natural variability on a same accession and to highlight gene regulatory networks correlated with organ morphology. The procedure allows to propose some key genes that could be involved in determination of organ morphology and to help prioritize the choice of genes to further functional study.

The scientific procedure is efficient and clear. The statistic analysis is appropriate for each analysis. Presentation of the results/figures are consive and well presented.

Minor comments:

- Why the clustering co-expression analysis was not performed with uniquely the HVGs? Dis the authors make this analysis?
- Figure 2C : add a color information for the variable contribution. codeLegend Supp figure1. Add a description of the color code D, E and F
- Supp Fig1: add color significance in the legend (Red, Blue, green)
- Supp Fig3: indicate HVG and LVG on the figure

Reviewed by **Sandra Corjito**, 05 May 2023

In their article, Hartasánchez and colleagues link transcriptional and phenotypic variability between individual sepals in Arabidopsis thaliana. To do this, they measured several parameters describing sepal morphology for a total of 30 sepals and performed RNA-seq on these same individual sepals. In a first part of the article, they show that parameters related to the curvature of the sepal are more variable that the ones related to the size of sepals. They also identify genes with low or high sepal-to-sepal transcriptional variability and find that cell-wall related genes are highly variable. In the second part of the article, they directly link changes in expression and in morphology between sepals. For this, they clustered genes based on their expression in the sepals and then measured the correlation between the expression profile of each cluster and the morphological parameters in the different sepals. They could detect several significant correlations which they then explored further allowing them to identify several cell-wall related genes and transcription factors whom expression is correlated with sepal width.

This article is of great interest as it is the first to analyse transcriptional variability between organs in a plant, and to associate it with phenotypic variability. Moreover, their findings with cell-wall related genes provide targeted hypotheses that will be undoubtfully tested in the future. The work is also scientifically very robust and the article is well written. I would just suggest some minor changes, listed below, to improve some analyses and figures. In summary, I recommend this very good article.

Below are some minor changes I suggest:

- The authors indicate they do not see a correlation between CV2 and average expression level since they removed the lowly expressed genes. However, this is very hard to see in the Supplementary Figure 2 because of the different colors. Moreover, the fact that lowly variable genes (LVG) have a higher expression than

highly variable genes (HVG) as shown in Supplementary Figure 2 (top) contradicts this claim. To see if there is indeed a correlation or not, I recommend the authors to do one or several of the following suggestions: (i) add a regression line to the plot; (ii) aid visualization by showing the point density; (iii) provide the result of a correlation test. If the authors do find a correlation, I suggest to use the method described in Brennecke et al, 2013 (https://doi.org/10.1038/nmeth.2645) to identify HVG.

- The Supplementary Figure 3 is very difficult to understand. I would recommend the authors to explain a bit more what they did in the text and the legend. Also, an alternative to Venn diagrams, like the one proposed by the UpSetR package, might help understand the figure.

- In the Supplementary Figure 4, the authors do a Wilcoxon test to compare the CV2 of two groups of genes of very different size (718 genes vs 14085 genes). In order to actually be able to compare the CV2 of these two groups the number of genes should be similar. For this, I recommend to use a bootstrap strategy with 1000 random sets of 718 genes taken in the 14085 genes set and to compare the average and the 95% confidence interval obtained for this 1000 random sets with the CWRG list.

- In the Figures 5 and 6, the authors use a trait significance. However, they never explain in the text how it was measured and what it means.

- About the discussion on cell-wall related genes, I would like to add that this category of GO is also enriched among HVG detected between plants in Cortijo et al. 2019. The authors' observations might thus be more general and not specific to sepals, making it a phenomenon of wide interest.