EMBO reports

Dissecting the co-transcriptome landscape of plants and their microbiota

Tatsuya Nobori, Yu Cao, Frederickson Entila, Eik Dahms, Yayoi Tsuda, Ruben Garrido-Oter, and Kenichi Tsuda
DOI: 10.15252/embr.202255380

Corresponding author: Kenichi Tsuda (tsuda@mail.hzau.edu.cn)

Review Timeline:

Submission Date:	10th May 22
Editorial Decision:	13th Jun 22
Revision Received:	13th Aug 22
Editorial Decision:	5th Sep 22
Revision Received:	9th Sep 22
Accepted:	13th Sep 22

Editor: Achim Breiling

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

Dear Prof. Tsuda,

Thank you for the submission of your research manuscript to EMBO reports. I have now received the reports from the three referees that were asked to evaluate your study, which can be found at the end of this email.

As you will see, the referees think that these findings are of interest. However, all referees have several comments, concerns, and suggestions, indicating that revision of the manuscript is necessary to allow publication of the study in EMBO reports. As the reports are below, and all their points need to be addressed, I will not detail them here.

Given the constructive referee comments, we would like to invite you to revise your manuscript with the understanding that all referee concerns must be addressed in the revised manuscript and in a detailed point-by-point response. Acceptance of your manuscript will depend on a positive outcome of a second round of review. It is EMBO reports policy to allow a single round of revision only and acceptance of the manuscript will therefore depend on the completeness of your responses included in the next, final version of the manuscript.

Revised manuscripts should be submitted within three months of a request for revision. Please contact me to discuss the revision (also by video chat) if you have questions or comments regarding the revision, or should you need additional time.

When submitting your revised manuscript, please also carefully review the instructions that follow below.

PLEASE NOTE THAT upon resubmission revised manuscripts are subjected to an initial quality control prior to exposition to re-review. Upon failure in the initial quality control, the manuscripts are sent back to the authors, which may lead to delays. Frequent reasons for such a failure are the lack of the data availability section (please see below) and the presence of statistics based on n=2 (the authors are then asked to present scatter plots or provide more data points).

When submitting your revised manuscript, we will require:

1) a .docx formatted version of the final manuscript text (including legends for main figures, EV figures and tables), but without the figures included. Figure legends should be compiled at the end of the manuscript text.

2) individual production quality figure files as .eps, .tif, .jpg (one file per figure), of main figures and EV figures. Please upload these as separate, individual files upon re-submission.

The Expanded View format, which will be displayed in the main HTML of the paper in a collapsible format, has replaced the Supplementary information. You can submit up to 5 images as Expanded View. Please follow the nomenclature Figure EV1, Figure EV2 etc. The figure legend for these should be included in the main manuscript document file in a section called Expanded View Figure Legends after the main Figure Legends section. Additional Supplementary material should be supplied as a single pdf file labeled Appendix. The Appendix should have page numbers and needs to include a table of content on the first page (with page numbers) and legends for all content. Please follow the nomenclature Appendix Figure Sx, Appendix Table Sx etc. throughout the text, and also label the figures and tables according to this nomenclature.

For more details, please refer to our guide to authors:
http://www.embopress.org/page/journal/14693178/authorguide#manuscriptpreparation

Please consult our guide for figure preparation:
http://wol-prod-cdn.literatumonline.com/pb-assets/embo-site/EMBOPress_Figure_Guidelines_061115-1561436025777.pdf

See also the guidelines for figure legend preparation:
https://www.embopress.org/page/journal/14693178/authorguide#figureformat

3) a complete author checklist, which you can download from our author guidelines (https://www.embopress.org/page/journal/14693178/authorguide). Please insert page numbers in the checklist to indicate where the requested information can be found in the manuscript. The completed author checklist will also be part of the RPF.

Please also follow our guidelines for the use of living organisms, and the respective reporting guidelines:
http://www.embopress.org/page/journal/14693178/authorguide#livingorganisms

4) that primary datasets produced in this study (e.g. RNA-seq, ChIP-seq, structural and array data) are deposited in an appropriate public database. If no primary datasets have been deposited, please also state this in a dedicated section (e.g. 'No primary datasets have been generated and deposited'), see below.

See also: http://embor.embopress.org/authorguide#datadeposition

Please remember to provide a reviewer password if the datasets are not yet public.

The accession numbers and database should be listed in a formal "Data Availability" section (placed after Materials & Methods) that follows the model below. This is now mandatory (like the COI statement). Please note that the Data Availability Section is restricted to new primary data that are part of this study. This section is mandatory. As indicated above, if no primary datasets have been deposited, please state this in this section

Data availability

The datasets produced in this study are available in the following databases:

- RNA-Seq data: Gene Expression Omnibus GSE46843 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE46843)
- [data type]: [name of the resource] [accession number/identifier/doi] ([URL or identifiers.org/DATABASE:ACCESSION])

*** Note - All links should resolve to a page where the data can be accessed. ***

Moreover, I have these editorial requests:

6) We strongly encourage the publication of original source data with the aim of making primary data more accessible and transparent to the reader. The source data will be published in a separate source data file online along with the accepted manuscript and will be linked to the relevant figure. If you would like to use this opportunity, please submit the source data (for example scans of entire gels or blots, data points of graphs in an excel sheet, additional images, etc.) of your key experiments together with the revised manuscript. If you want to provide source data, please include size markers for scans of entire gels, label the scans with figure and panel number, and send one PDF file per figure.

7) Our journal encourages inclusion of *data citations in the reference list* to directly cite datasets that were re-used and obtained from public databases. Data citations in the article text are distinct from normal bibliographical citations and should directly link to the database records from which the data can be accessed. In the main text, data citations are formatted as follows: "Data ref: Smith et al, 2001" or "Data ref: NCBI Sequence Read Archive PRJNA342805, 2017". In the Reference list, data citations must be labeled with "[DATASET]". A data reference must provide the database name, accession number/identifiers and a resolvable link to the landing page from which the data can be accessed at the end of the reference. Further instructions are available at: http://www.embopress.org/page/journal/14693178/authorguide#referencesformat

8) Regarding data quantification and statistics, please make sure that the number "n" for how many independent experiments were performed, their nature (biological versus technical replicates), the bars and error bars (e.g. SEM, SD) and the test used to calculate p-values is indicated in the respective figure legends (also for potential EV figures and all those in the final Appendix). Please also check that all the p-values are explained in the legend, and that these fit to those shown in the figure. Please provide statistical testing where applicable. Please avoid the phrase 'independent experiment', but clearly state if these were biological or technical replicates. See also:
http://www.embopress.org/page/journal/14693178/authorguide#statisticalanalysis

9) Please also note our reference format:
http://www.embopress.org/page/journal/14693178/authorguide#referencesformat

10) We updated our journal's competing interests policy in January 2022 and request authors to consider both actual and perceived competing interests. Please review the policy https://www.embopress.org/competing-interests and update your competing interests if necessary. Please name this section 'Disclosure and Competing Interests Statement' and put it after the Acknowledgements section.

I look forward to seeing a revised version of your manuscript when it is ready. Please let me know if you have questions or comments regarding the revision.

Please use this link to submit your revision: https://embor.msubmit.net/cgi-bin/main.plex

Best,

Achim Breiling
Senior Editor
EMBO Reports

Referee #1:

This paper carries out a very ambitious plan to understand the early transcriptional responses in both bacteria and plants when put in contact on leaves. The authors used a variety of commensal bacteria (and one pathogen) to gain phylogenetic and phenotypic diversity. The bacteria were inoculated one by one onto the leaves. The bacterial data are complex, yet certain patterns emerge, in particular when comparing commensal and pathogenic strains of bacteria. The downregulation of ribosomal genes in commensal bacteria is very compelling as it is highly reproducible. Other interesting results are the correlation of plant transcriptome with both flg22 and commensals, and the upregulation of several plant-associated genes whereas downregulation of the other genes. The lack of consistent pattern also is a "signal" that shows diversity of bacterial strategies. We see from the results the importance of transcriptomic studies rather than mere gene presence/absence. There is little overall correlation between plant and bacterial transcription, although some notable exceptions exist.

One importance of this paper to the field is that it generated many testable hypotheses based on the rich dataset and complex analysis that was performed. Notably, there is a whole class of genes with unknown annotation that is a rich treasure chest of future results. Overall, the manuscript summarizes the data and is well fit for publication from a scientific perspective. There are some points to clarify and a few missing items, but I see no major issues. I do feel that the figures feel very busy, and even with such a complex dataset, could be somewhat simplified.

Overall, I enjoyed reading this paper!

Comments on scientific content:
1. What is the difference between figure 1C (left) and S2C? They are both fold changes, and seem to have different axes somehow. I can see how figure 1C (right) and figure S2D complement each other as one is fold change and the other shows each transcriptome separately. There is something strange about this.
Fig 1C should be complemented with the same analysis of all replicates to see that they are clustered together, as expected.

2. Line 116: I would not use the term monocolonized if the plant has not undergone full sterilization.

3. Lines 121-122: There are 9 commensal strains in the text, and there is a reference to fig1B, where there are 19 strains in figure - This is confusing and unclear, I suggest marking the 9 commensal strains on the tree in a certain way.

4. Line 161: I wouldn't say D36E "fell between" Pto and the commensals. Visually, I understand D36E may be considered a distinct group but it is not "between" the two. If anything the D36E is closer to the commensals, But still seems relatively far from them so it is hard to assume that "commensal and disarmed pathogenic bacteria trigger common plant immune responses. "

5. This is related to the previous point: Figure 1C does not include the D36E and Pto strains. Why?

6. Is figure S2E and S2F on the same Y-axis? It is unclear. If it is supposed to be, the columns on S2E and S2F are in a different order. It makes it hard to compare.

7. Figure 2A: the numbers of genes (also in later figures) could go to the supplementary materials. The number of genes is important for understanding why or why not an enrichment may be statistically significant for technical reasons but it is really not necessary for the main figure.

8. Figure 2: I would like to see a discussion about certain conserved transcription factors. These can support the results on downregulation of ribosomal proteins (namely, transcription is also downregulated in the commensals and maybe upregulated among Pto). Similarly, I would love to see some analysis on mobile genetic elements. Levy et al. (Nature Genetics 2018) reported that mobile genetic elements were found to be depleted in plant-associated bacterial genomes and it would be interesting if these are also downregulated in planta.

9. Figure 3A: the reader will understand the result reported in text if the bottom genes will be marked, maybe in a square (PA genes upregualted, non-PA genes downregulated).

10. Lines 205-207: I don't see a p-value for Catalase genes in the figure or the supplementary figure.

11. There are no p-values for flagellar assembly. There are only values for motility, and chemotaxis.

12. Figure S4A and S4C: why was k=8 chosen for the k-means clustering? There is no explanation in the text about this.

13. I don't see a p-value for nitrate transport (line 333).

14. Line 347-348: A result of Pto is mentioned but is not in the referenced figure (4E)!

15. Figure 5: it would be important to correlate the plant response with the actual flg22 and elf18 bacterial expression across strains (not only with the protein similarity).

16. Figure 5a: It is unclear if we see all diff. expressed genes or just the ones that correlate well between commensal response and flg22 response. It should be stated clearly.

17. Discussion: it would be interesting to discuss negative results. Namely, genes that were expected to be up- or down-regulated based on the literature on plant-microbe interactions, yet the rich data the authors produced clearly refuted these expectations.

Minor grammar and visual issues:

18. Overall, the figures are very busy, with tons of colors. If I had to make one big change, I would label the NAMES of the strains (e.g. Leaf 1) in red, rather than putting a color strip on top of many figures in the paper, which adds 6 more colors to an already busy figure. For example, figure 2A does not need the top color strip, just color the names on the bottom of the figure.

19. Line 83: To make the English clearer, I would say "however, there are only a few available...". Without this, the sentence is a bit awkward.

20. Figures phylum/class level colors: Using very similar colors - it is almost impossible to distinguish between the green colors that mark Alphaproteobacteria and Gammaproteobacteria.

21. Figure 1C: Confusing to see the title "plant" and see in the graph names of bacterial lines, so I would add information that the bacterial names represent plant lines plus the same bacterium.

22. Lines 153-155: The sentence is unclear, and it is unclear when plant transcriptome is meant, and when bacterial transcriptome is meant.

23. The log pvalue scale bar in Figure 2A goes in both directions to 10. It is not strictly necessary to change this but I had to look at it twice to understand.

24. Figure 2E suddenly is not left to right. Figure 2B, C, D, are all left to right, and then suddenly 2E is a square going right, down, left, up. Not strictly necessary to change but it was confusing.

25. Figure 4A has a spelling error- metabolism is spelled wrong.

26. Figure 4C also is spelled wrong - intracellular and Extracellular (with two L's)

27. Lines 304-305 and 316-317 are a bit awkwardly phrased and I only understood the meaning when I got to the end of the paragraph. The way it is written, I didn't realize immediately that it's genes that are Up AND PA and Down AND nonPA. I would phrase it more clearly like you do in figure 4A (Up and PA, and Down and nonPA). Maybe write "as well as" instead of "and" or maybe put in parentheses (Figure 4A-C, "Up and PA") so it's clearer. Some of the confusion stems from the fact that "Sulfur metabolism" genes are not classified only as PA (there are some yes and some no). So just reading line 303, I am already confused.. Then only after I realized you mean they are considered PA only in these proteobacteria.

Referee #2:

The study described in the manuscript by authors, Nobori, et al. aimed to explore the landscape of plant and bacterial transcriptome changes in response to commensal and pathogenic bacterial infection in Arabidopsis. The authors colonized Arabidopsis leaf tissue with a handful of commensal bacterial strains that span many plant-associated phyla, and utilized the co-transcriptomics workflow to capture both plant and microbial RNA for sequencing. To me, the major findings are that 1) despite having a similar phenotypic impact on plants, commensal microbes elicit distinct molecular responses, 2) bacteria also display highly heterogeneous responses to existing within the plant apoplast, and 3) that genes which are strongly induced upon plant infection also tend to be more broadly conserved among plant-associated strains. While descriptive, I thought the manuscript to be a tour-de-force of a co-transcriptome survey that offers a useful dataset for further exploration of plant/microbial interactions, and warrants publication.

I only have a few minor points that the authors should consider addressing:

1. Figure 1c shows a PCA plot with each strain represented as a single data point for the bacterial and plant transcriptomes. The main conclusion drawn from this is that there exists substantial variation among strains, and that transcriptome similarities among some bacteria do not correlate with changes in their effect on the plant transcriptome. However (unless I'm misinterpreting the PCA plot), I'm not sure that PCA alone can really suggest this, as large distances in PCA space can exist for small changes, if the total amount of variation is small. One way this might be made more clear is if the authors were to add in each replicate sample into this plot (I gather that only 1 representative is shown, currently?) If the inter-replicate distances are

much smaller than the inter-species differences, then this might be a more convincing argument. I don't think this matters much to the overall message of the manuscript, but thought it might be good to show.

2. Similarly, I don't think that the statement on line 167 "revealed commonalities between commensals and a disarmed pathogen" can really be supported by PCA alone - whether or not the response to the disarmed pathogen is "between" the pathogen and commensal strains is likely very complex, and I think would need to be verified with looking at shared sets of genes that are affected among the three groups.

3. On line 202-204, the authors state, "PTI is likely responsible for suppressing bacterial metabolism in planta, and only pathogens can overcome PTI to be metabolically active" based on their observation that the D36E mutant (ineffective in suppressing PTI) shares the commensal phenotype of suppressing ribosome biosynthesis genes, while the virulent Pto strain does not. I think this is an interesting idea, however at some point won't commensal strains need to overcome PTI to survive in the host? It might be more fair to say that pathogens are quicker to suppress PTI (as is implied by the previous sentence).

Referee #3:

The manuscript of Nabori et al. presents a comprehensive analysis of transcriptional responses detected in Arabidopsis leaves and in the infiltrated commensal bacteria. The aim is to characterize the molecular response of the host and of commensal bacteria during leaf apoplast colonization. For this, the authors selected commensals previously identified as being part of Arabidopsis leaf microbiota. Individual commensals from different taxa have been infiltrated into Arabidopsis leaf and transcriptional responses were monitored. The main finding of this study is a differential response in the different host-commensal combinations. The authors performed an impressive study that monitored both plant and bacterial responses in nine independent interactions. These have been carefully analyzed and compared with those of a well-known pathogen (Pto) as well as an effector-disarmed form of Pto (D36). Interestingly, the overall pattern emerging from this study is that Arabidopsis responds to commensals and to the effector-disarmed pathogen similarly.

The authors performed a detailed analysis of bacterial functions differentially regulated in planta compared to in vitro and identified both common responses across infecting bacteria as well as individual ones. Comparisons across and between phyla are performed to finely dissect these responses.
Overall, this is a comprehensive study, carefully executed, analyzed, and interpreted which clearly opens the possibility for numerous in-depth and detailed analyses at the isolate or function level.

Comments:
1. The authors analyze the data for each commensal based on differential abundances of transcripts in planta versus in vitro, but the interpretation is then made between commensals or between plants inoculated with different commensals. Bacterial functions are differently expressed by commensals already in the in vitro state, as shown in Figure S9, thus analyzing the data only from the perspective of in planta versus in vitro provides a unilateral view.
2. The degree of plant gene expression varies between inoculations with different commensals. Could this be explained by the observed differences in the bacterial population (Fig.S1), or by the differences in their native levels of expression of specific functions (point1)
3. Line135. Transcriptional responses are analyzed at 6hpi. This is fine and well justified by the authors, however, it is important to mention in the abstract and consider throughout the manuscript that these are the early responses of the interacting partners.
4. Line 139. The authors chose to compare expression levels of bacterial functions in planta versus in vitro, alias growth in rich media. While this can be explained by the rationale of the experimental setup it is important to explain to the reader why a rich media was chosen as a reference. Having the baseline as growth in a nutrient-rich environment explains the reduction in "ribosome" activity as mentioned by the authors, thus maybe creating a distorted view on which functions are activated or deactivated in the apoplast. One might imagine bacteria rarely encounter rich media in natural environments, thus the comparison baseline will be completely different.
5. Line 186. Results presented here come as an argument for comment 4. Pto varies in responses when grown in rich versus minimal media, thus it will be relevant to discuss which functions in bacteria and in the plant remain constant independent of the baseline. Is it possible to identify those for Pto and draw comparisons with responses in commensals?
6. Line 204. The authors likely mean that overcoming PTI is a prerequisite for an active metabolism in infectious bacteria. It would be interesting to know if plants impaired in PTI would enable commensals to activate a larger metabolic activity, and which ones would those be.
7. Line 217. It is not entirely clear which physiological responses the authors refer to here. This sentence might require revision.
8. Line 261. It is relevant to emphasize that upregulated genes were primarily strain-specific, while those downregulated are general. In this regard, it will be relevant to add to figures S4B and D the number of genes which are unique to the analyzed isolates. Which functions do these genes have? Do they contribute to other differences depicted in the other figures?
9. Line 290. Are these functions activated in planta or just have a higher level of expression? It will be important to make a distinction between activated versus higher/lower levels of expression.
10. Line296. "Ribosome" suppression is mentioned here as a being a consequence of plant-imposed control of bacterial growth, however, there is no data supporting this statement. This could be a native response of bacteria when transiting from rich to low

nutrient conditions and not a plant-driven effect. A relevant reference could be the response in the presence of Arabidopsis apoplastic fluid.
11. Line 324. Could it be envisioned that both degradation and nutrient sources for growth are at play?
12. Line 354. The statement regarding sugar transporters of Pto being not differentially expressed in planta is difficult to understand in the light of the given reference.
13. Line 394. Co-transcriptome analyses identified that most genes of Arabidopsis and bacteria are not correlated. This result is surprising, and the question is whether this is a result of the input data used for the analysis (DEGs for in planta vs in vitro). Would the analysis give similar results if genes expressed in planta will be considered instead?
14. Line 579- Methods. It is not clear if the expression of bacterial genes in vitro is performed on washed and water resuspended cells or directly from the cultures grown on rich media.

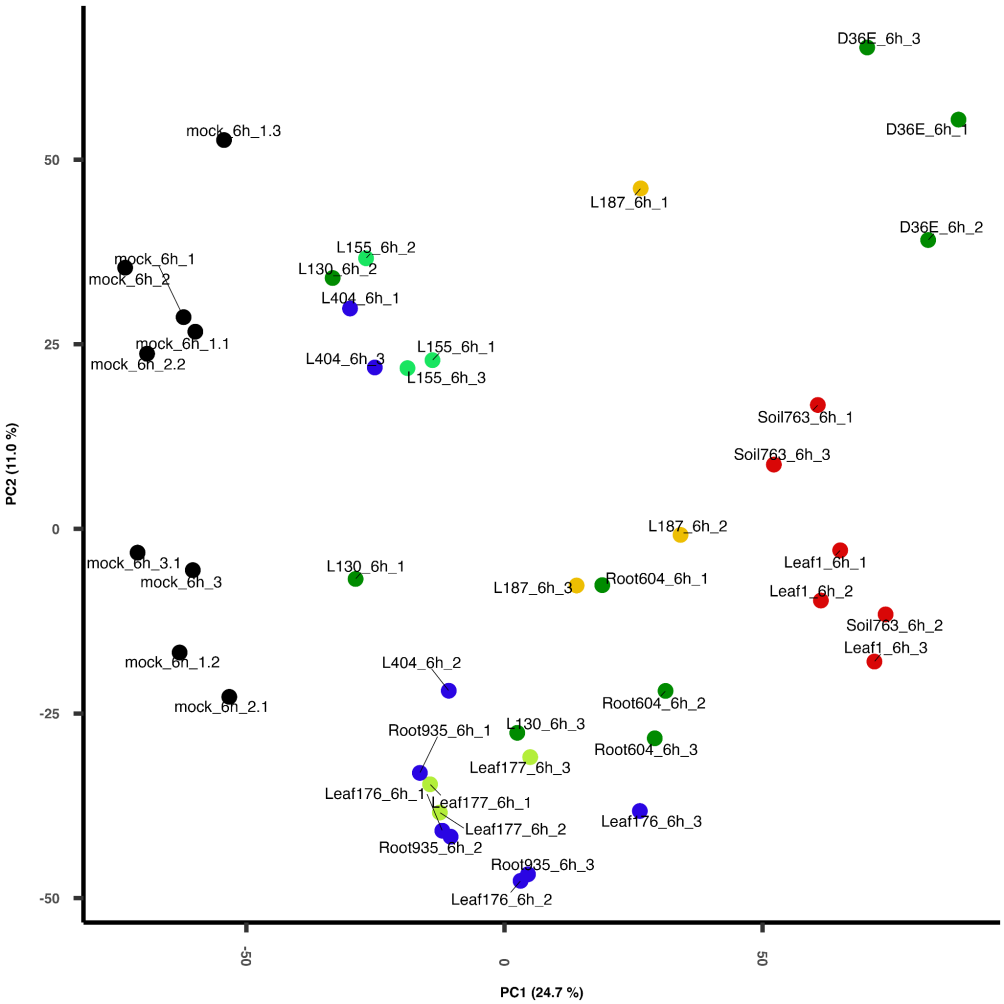
Referee #1:

This paper carries out a very ambitious plan to understand the early transcriptional responses in both bacteria and plants when put in contact on leaves. The authors used a variety of commensal bacteria (and one pathogen) to gain phylogenetic and phenotypic diversity. The bacteria were inoculated one by one onto the leaves. The bacterial data are complex, yet certain patterns emerge, in particular when comparing commensal and pathogenic strains of bacteria. The downregulation of ribosomal genes in commensal bacteria is very compelling as it is highly reproducible. Other interesting results are the correlation of plant transcriptome with both flg22 and commensals, and the upregulation of several plant-associated genes whereas downregulation of the other genes. The lack of consistent pattern also is a "signal" that shows diversity of bacterial strategies. We see from the results the importance of transcriptomic studies rather than mere gene presence/absence. There is little overall correlation between plant and bacterial transcription, although some notable exceptions exist.

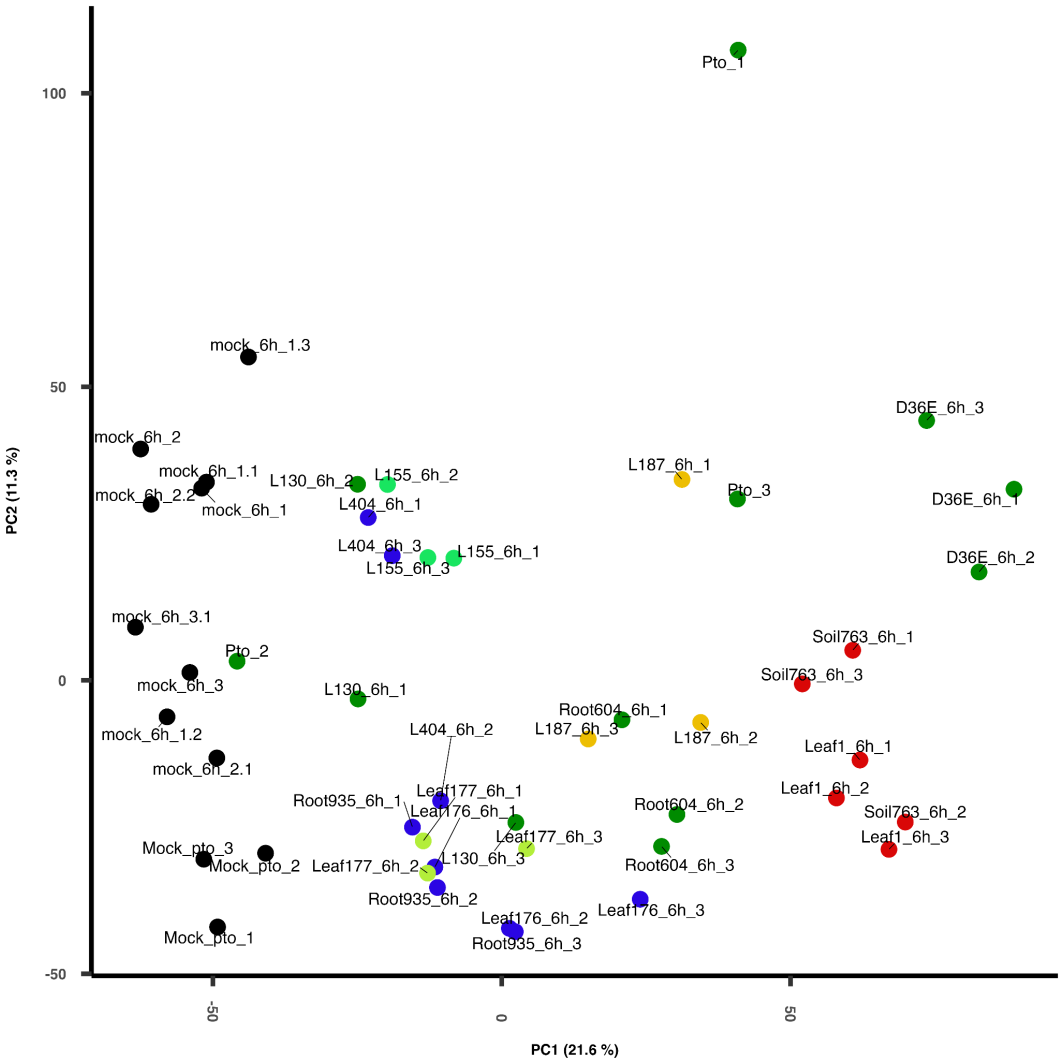
One importance of this paper to the field is that it generated many testable hypotheses based on the rich dataset and complex analysis that was performed. Notably, there is a whole class of genes with unknown annotation that is a rich treasure chest of future results. Overall, the manuscript summarizes the data and is well fit for publication from a scientific perspective. There are some points to clarify and a few missing items, but I see no major issues. I do feel that the figures feel very busy, and even with such a complex dataset, could be somewhat simplified.

Overall, I enjoyed reading this paper!
We are thankful for the appreciative comments and are pleased that the Reviewer enjoyed our work.

Comments on scientific content:
1. What is the difference between figure 1C (left) and S2C? They are both fold changes, and seem to have different axes somehow. I can see how figure 1C (right) and figure S2D complement each other as one is fold change and the other shows each transcriptome separately. There is something strange about this.
Fig 1C should be complemented with the same analysis of all replicates to see that they are clustered together, as expected.
Our response 1
Figure S2C (now Appendix Figure S2C) contains *Pto* and D36E in addition to all commensal trains while Figure 1C (left) does not have Pto and D36E. This affected PCA. According to the suggestion, we provided PCA plots with each replicate for Figure 1C and Figure S2C (now Appendix Figure S2C). While replicates of the same samples were generally clustered together, there are variations among independent replicates. This is not unusual and can be explained by various factors (e.g., growth chamber conditions and soil batches). To correct such random effects, we fit the data to a linear mixed model (see method), and all the data we presented were fitted data.



Individual replicates of plant transcriptome used in Fig1C are shown.



Individual replicates of plant transcriptomes used in FigS2C (now Appendix Figure S2C) are shown.

2. Line 116: I would not use the term monocolonized if the plant has not undergone full sterilization.

Our response 2

We have replaced “monocolonized” with “colonized”. In addition, we have removed the term “monoassociation” througout the text.

3. Lines 121-122: There are 9 commensal strains in the text, and there is a reference to fig1B, where there are 19 strains in figure - This is confusing and unclear, I suggest marking the 9 commensal strains on the tree in a certain way.

Our response 3

Those 9 strains used for co-transcriptome analysis are marked with stars in Figure 1B. In Figure legend, it reads: “Stars indicate the strains used for co-transcriptome analysis.” We hope that this is sufficient for understanding.

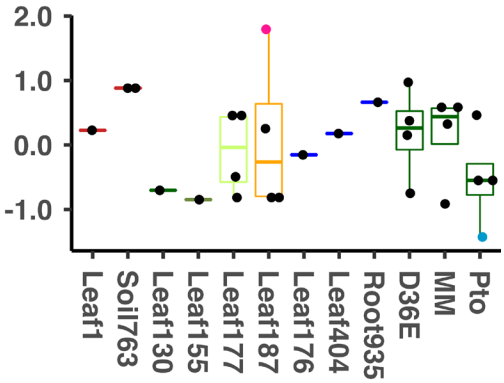
4. Line 161: I wouldn't say D36E "fell between" Pto and the commensals. Visually, I understand D36E may be considered a distinct group but it is not "between" the two. If anything the D36E is closer to the commensals, But still seems relatively far from them so it is hard to assume that "commensal and disarmed pathogenic bacteria trigger common plant immune responses. "
Our response 4
Thank you for pointing this out. We have revised the text to more accurately describe our result. Now it reads: "*The virulent pathogen Pto triggered a highly different plant transcriptome response compared with commensals and the disarmed pathogen D36E (* Appendix Figure S2C)*).*"

5. This is related to the previous point: Figure 1C does not include the D36E and Pto strains. Why?
Our response 5
We did not include *Pto* and D36E data in Figure 1C as the difference between Pto and commensal strains would mask the difference among commensal strains. Figure S2C (now Appendix Figure S2C) includes *Pto* and D36E in addition to the commensal strains. Indeed, the resolution of the differences among the commensal strains decreases in Figure S2C (now Appendix Figure S2C).

6. Is figure S2E and S2F on the same Y-axis? It is unclear. If it is supposed to be, the columns on S2E and S2F are in a different order. It makes it hard to compare.
Our response 6
We have removed Figure S2E as Figure S2F (now Appendix Figure S2E) provides similar information.

7. Figure 2A: the numbers of genes (also in later figures) could go to the supplementary materials. The number of genes is important for understanding why or why not an enrichment may be statistically significant for technical reasons but it is really not necessary for the main figure.
Our response 7
While including the number of genes makes this figure more complex, we think that having this information together with the heat map is important for understanding. Thus, we would like to keep it.

8. Figure 2: I would like to see a discussion about certain conserved transcription factors. These can support the results on downregulation of ribosomal proteins (namely, transcription is also downregulated in the commensals and maybe upregulated among Pto). Similarly, I would love to see some analysis on mobile genetic elements. Levy et al. (Nature Genetics 2018) reported that mobile genetic elements were found to be depleted in plant-associated bacterial genomes and it would be interesting if these are also downregulated in planta.
Our response 8
We analyzed conserved sigma factors but did not see a clear pattern (please see the plot below).



Expression fold changes (*in planta* vs. *in vitro*) of transcriptional sigma factors that are conserved among the strains. Log_2 normalized values are shown.

We were not able to identify mobile genetic elements (prophages and transposons) shared across our strains in our gene annotation. In both cases (TFs and mobile elements), we believe an improved and detailed genome annotation would enable more extensive analysis in the future.

9. Figure 3A: the reader will understand the result reported in text if the bottom genes will be marked, maybe in a square (PA genes upregualted, non-PA genes downregulated).
Our response 9
Thank you very much for the suggestion. We have added rectangles indicating that PA up and non-PA down genes in Figure 3A.

10. Lines 205-207: I don't see a p-value for Catalase genes in the figure or the supplementary figure.
Our response 10
The red dots in Figure 2C (now Appendix Figure S2C) (Catalase) indicate that they are statistically significantly induced. We hope that this is sufficient for understanding.

11. There are no p-values for flagellar assembly. There are only values for motility, and chemotaxis.
Our response 11
The red dots in Figure 2D (Flagellar assembly proteins and Chemotaxis) indicate that they are statistically significantly induced. We hope that this is sufficient for understanding.

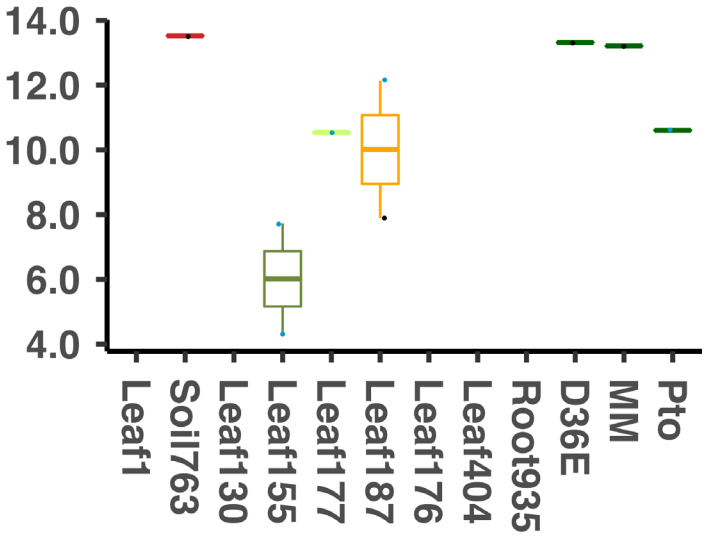
12. Figure S4A and S4C: why was k=8 chosen for the k-means clustering? There is no explanation in the text about this.
Our response 12
Thank you for pointing this out. The k-means clustering was chosen based on the elbow method. We have now described this in the legend. Now it reads: "*Gene clusters defined by k-means clustering are shown (k = 8). The number of clusters was chosen based on the elbow method.*"

13. I don't see a p-value for nitrate transport (line 333).
Our response 13

The red dots in Figure S7 (now Appendix Figure S7) (Nitrate transport) indicate that they are statistically significantly induced. We hope that this is sufficient for understanding.

14. Line 347-348: A result of Pto is mentioned but is not in the referenced figure (4E)!
Our response 14
Thank you very much for pointing this out. We have now included *Pto* in Figures 4E and 4G.

15. Figure 5: it would be important to correlate the plant response with the actual flg22 and elf18 bacterial expression across strains (not only with the protein similarity).
Our response 15
The plot below shows expression of FliC, which encodes flg22 epitopes. D36E showed the induction of FliC, which is consistent with this strain's ability to trigger strong immune responses in plants (Fig. 5C). In contrast, Leaf155 showed lower expression of FliC, which is consistent with this strain's relatively weak ability to trigger plant immunity (Fig. 5C). Overall, the in planta FliC expression level of each strain was well correlated with the strain's ability to induce plant immunity. For instance, D36E and Soil763 shows high relative expression of FliC in planta and induced plant immunity strongly (Fig. 5C); Leaf155 showed a lower level of FliC expression in planta and did not trigger strong immune responses in plants (Fig. 5C). Therefore, the actual expression level of MAMPs could also explain differences in plant responses against different strains. However, many strains did not express a detectable level of FliC, thus comprehensive analysis was not possible, and we did not include this result in our revised manuscript.



***In planta* expression (not fold changes) of bacterial *FliC* gene.**

16. Figure 5a: It is unclear if we see all diff. expressed genes or just the ones that correlate well between commensal response and flg22 response. It should be stated clearly.
Our response 16
We used all of the differentially expressed genes in our dataset. We have revised the figure legend. Now it reads: "*(Green/purple heatmap) Gene expression fold changes (FCs) of differentially expressed plant genes at least in one sample between bacteria-inoculated plants and water-inoculated plants.*"

17. Discussion: it would be interesting to discuss negative results. Namely, genes that were expected to be up- or down-regulated based on the literature on plant-microbe interactions, yet the rich data the authors produced clearly refuted these expectations.
Our response 17
We lack sufficient knowledge of how bacterial genes behave (induced or suppressed) in planta. Therefore, at this stage, it is not possible to provide such comparisons. We think that our dataset will be the basis of such future comparisons.

Minor grammar and visual issues:

18. Overall, the figures are very busy, with tons of colors. If I had to make one big change, I would label the NAMES of the strains (e.g. Leaf 1) in red, rather than putting a color strip on top of many figures in the paper, which adds 6 more colors to an already busy figure. For example, figure 2A does not need the top color strip, just color the names on the bottom of the figure.
Our response 18
We think that the color bar is easier to recognize the different taxa for each strain. Thus, we would like to keep as they are.

19. Line 83: To make the English clearer, I would say "however, there are only a few available...". Without this, the sentence is a bit awkward.
Our response 19
Thank you for the suggestion. We have revised the text accordingly.

20. Figures phylum/class level colors: Using very similar colors - it is almost impossible to distinguish between the green colors that mark Alphaproteobacteria and Gammaproteobacteria.
Our response 20
This is the color scheme used in the previous paper (Bai et al Nature 2015). We would like to use the consistent color scheme. Table 1 provides a detail taxonomy for each strain. We hope that this is sufficient for understanding.

21. Figure 1C: Confusing to see the title "plant" and see in the graph names of bacterial lines, so I would add information that the bacterial names represent plant lines plus the same bacterium.
Our response 21
Thank you for the comment. To avoid confusion, we have changed the title “Plant” and “Bacteria” to “Plant transcriptome” and “Bacterial transcriptome”

22. Lines 153-155: The sentence is unclear, and it is unclear when plant transcriptome is meant, and when bacterial transcriptome is meant.
Our response 22
To avoid confusion, we have revised the text. Now it reads: “*Also, Bacteroidetes strains (Leaf176, Leaf404, and Root935) showed similar bacterial transcriptional changes in plants, but plant transcriptome changes triggered by these strains were distinct (Fig. 1C).*”

23. The log pvalue scale bar in Figure 2A goes in both directions to 10. It is not strictly necessary to change this but I had to look at it twice to understand.
Our response 23
We would like to include information for p-values and up or down regulation. We hope that this is OK.

24. Figure 2E suddenly is not left to right. Figure 2B, C, D, are all left to right, and then suddenly 2E is a square going right, down, left, up. Not strictly necessary to change but it was confusing.
Our response 23
We agree that the panel ordering might be confusing, but we wanted to fit all the important information in the Figure. We hope that this is OK.

25. Figure 4A has a spelling error- metabolism is spelled wrong.
26. Figure 4C also is spelled wrong - intracellular and Extracellular (with two L's)
Our responses 25 and 26.
Thank you very much for pointing these out. We have corrected the errors.

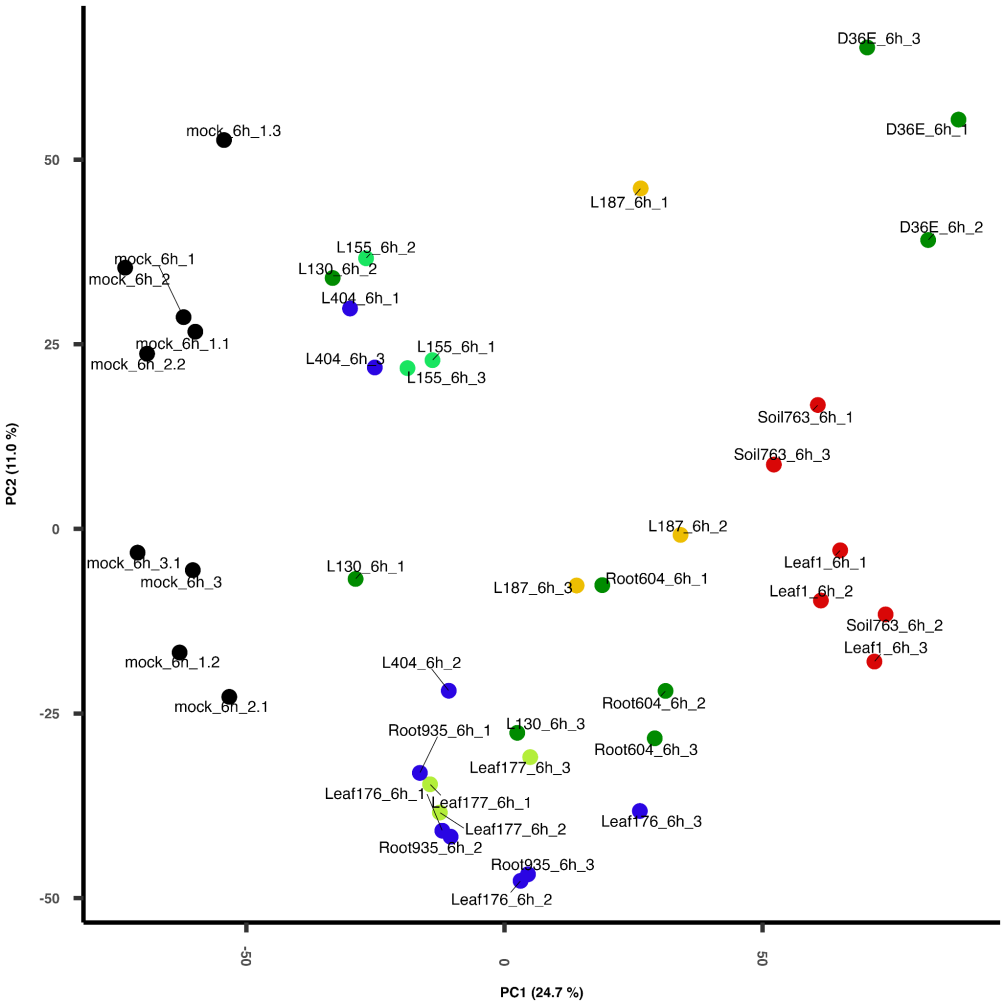
27. Lines 304-305 and 316-317 are a bit awkwardly phrased and I only understood the meaning when I got to the end of the paragraph. The way it is written, I didn't realize immediately that it's genes that are Up AND PA and Down AND nonPA. I would phrase it more clearly like you do in figure 4A (Up and PA, and Down and nonPA). Maybe write "as well as" instead of "and" or maybe put in parentheses (Figure 4A-C, "Up and PA") so it's clearer. Some of the confusion stems from the fact that "Sulfur metabolism" genes are not classified only as PA (there are some yes and some no). So just reading line 303, I am already confused.. Then only after I realized you mean they are considered PA only in these proteobacteria.
Our response 27
Thank you very much for the comments. We have revised the text accordingly.
"*We found a significant number of "sulfur metabolism"-related genes that are PA genes and induced in planta in the three Proteobacteria strains (Fig. 4A-4C).* "
" *On the other hand, a significant number of sulfur metabolism-related genes were non-PA genes and suppressed in planta in the Bacteroidetes strains Leaf176 and Leaf404 (Fig. 4A and 4B).*"

Referee #2:

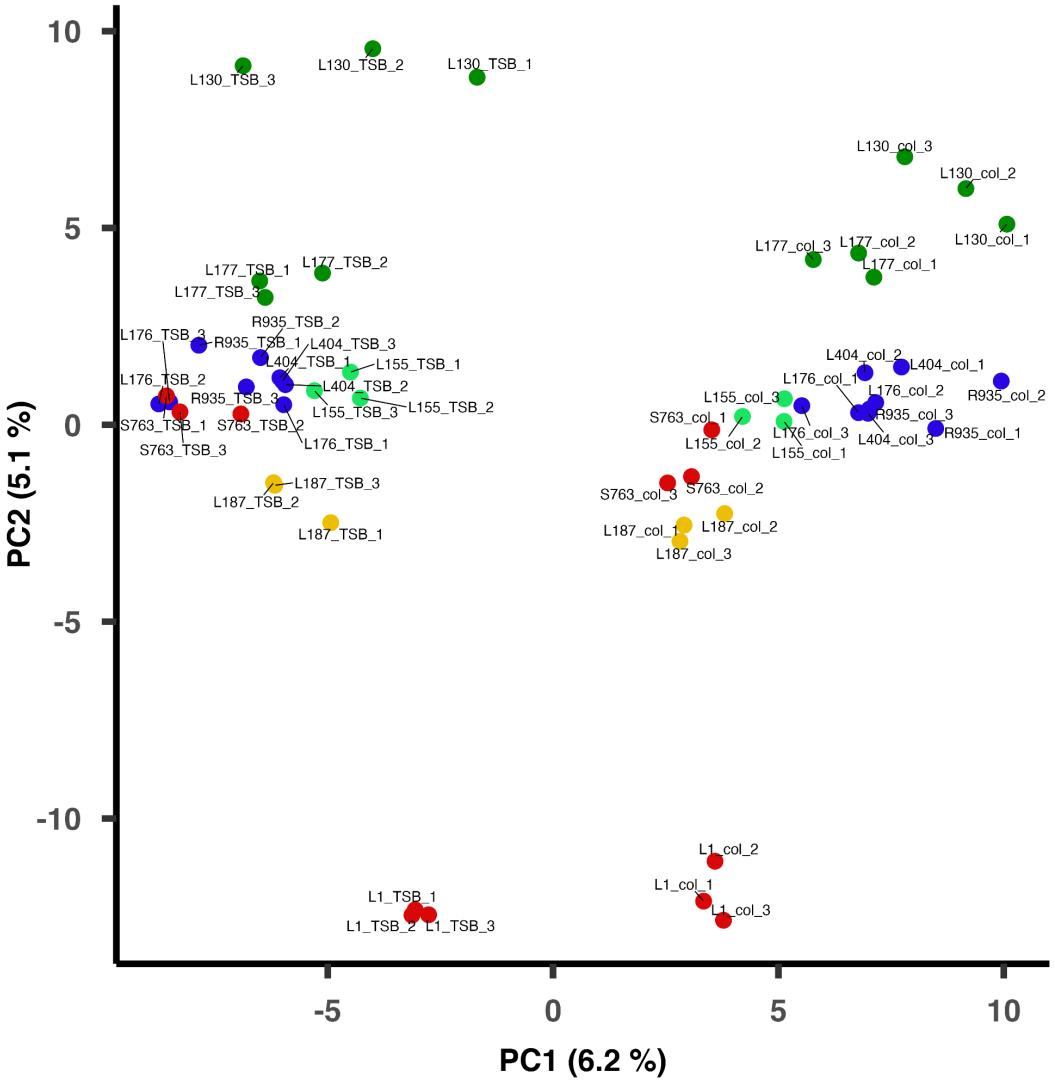
The study described in the manuscript by authors, Nobori, et al. aimed to explore the landscape of plant and bacterial transcriptome changes in response to commensal and pathogenic bacterial infection in Arabidopsis. The authors colonized Arabidopsis leaf tissue with a handful of commensal bacterial strains that span many plant-associated phyla, and utilized the co-transcriptomics workflow to capture both plant and microbial RNA for sequencing. To me, the major findings are that 1) despite having a similar phenotypic impact on plants, commensal microbes elicit distinct molecular responses, 2) bacteria also display highly heterogeneous responses to existing within the plant apoplast, and 3) that genes which are strongly induced upon plant infection also tend to be more broadly conserved among plant-associated strains. While descriptive, I thought the manuscript to be a tour-de-force of a co-transcriptome survey that offers a useful dataset for further exploration of plant/microbial interactions, and warrants publication.

I only have a few minor points that the authors should consider addressing:
We thank the Reviewer for the comments and for appreciating the value of our dataset and analysis.

1. Figure 1c shows a PCA plot with each strain represented as a single data point for the bacterial and plant transcriptomes. The main conclusion drawn from this is that there exists substantial variation among strains, and that transcriptome similarities among some bacteria do not correlate with changes in their effect on the plant transcriptome. However (unless I'm misinterpreting the PCA plot), I'm not sure that PCA alone can really suggest this, as large distances in PCA space can exist for small changes, if the total amount of variation is small. One way this might be made more clear is if the authors were to add in each replicate sample into this plot (I gather that only 1 representative is shown, currently?) If the inter-replicate distances are much smaller than the inter-species differences, then this might be a more convincing argument. I don't think this matters much to the overall message of the manuscript, but thought it might be good to show.
Our response 28
In the original figures, we used mean expression values from the linear model analysis. According to the suggestion, we provided PCA plots with each replicate for Figure 1C. Gene expression fold changes (FC) are not used because FC cannot be calculated before taking the mean values of replicates in each condition. The reproducibility of bacterial transcriptome data (bottom plot) is high. Plant transcriptome data (top) showed some variability among replicates, which could be explained by various factors (e.g., growth chamber conditions and soil batches). To correct such random effects, we fit the data to a linear mixed model (see method), and all the data we presented (both plant and bacteria) were fitted data.



Individual replicates of plant transcriptome used in Fig1C are shown.



Individual replicates of bacterial transcriptome used in Fig1C are shown.
"TSB" and "col" in the labels indicate *in vitro* and *in planta*, respectively.

2. Similarly, I don't think that the statement on line 167 "revealed commonalities between commensals and a disarmed pathogen" can really be supported by PCA alone - whether or not the response to the disarmed pathogen is "between" the pathogen and commensal strains is likely very complex, and I think would need to be verified with looking at shared sets of genes that are affected among the three groups.

Our response 29

Thank you for pointing this out. We have revised the text to more accurately describe our result. Now it reads: "*The virulent pathogen Pto triggered a highly different plant transcriptome response compared with commensals and the disarmed pathogen D36E (Appendix Fig. S2C).*"

3. On line 202-204, the authors state, "PTI is likely responsible for suppressing bacterial metabolism in planta, and only pathogens can overcome PTI to be metabolically active" based on their observation that the D36E mutant (ineffective in suppressing PTI) shares the commensal phenotype of suppressing ribosome biosynthesis genes, while the virulent Pto strain does not. I think this is an interesting idea, however at some point won't commensal

strains need to overcome PTI to survive in the host? It might be more fair to say that pathogens are quicker to suppress PTI (as is implied by the previous sentence).

Our response 30

We never observed that the used commensal strains actually multiplicate in the leaf apoplast even 7 days after hand-infiltration. At this stage, we really don't know how metabolic acitivity of commensal bacteria change in the host. Nevertheless, to be accurate, we have revised the text. Now it reads: "*Since D36E is a mutant of Pto lacking PTI-suppressing effector molecules, PTI is likely responsible for suppressing bacterial metabolism in planta, and only pathogens can overcome PTI to be metabolically active at six hours after infection.*"

Referee #3:

The manuscript of Nabori et al. presents a comprehensive analysis of transcriptional responses detected in Arabidopsis leaves and in the infiltrated commensal bacteria. The aim is to characterize the molecular response of the host and of commensal bacteria during leaf apoplast colonization. For this, the authors selected commensals previously identified as being part of Arabidopsis leaf microbiota. Individual commensals from different taxa have been infiltrated into Arabidopsis leaf and transcriptional responses were monitored. The main finding of this study is a differential response in the different host-commensal combinations. The authors performed an impressive study that monitored both plant and bacterial responses in nine independent interactions. These have been carefully analyzed and compared with those of a well-known pathogen (Pto) as well as an effector-disarmed form of Pto (D36). Interestingly, the overall pattern emerging from this study is that Arabidopsis responds to commensals and to the effector-disarmed pathogen similarly.

The authors performed a detailed analysis of bacterial functions differentially regulated in planta compared to in vitro and identified both common responses across infecting bacteria as well as individual ones. Comparisons across and between phyla are performed to finely dissect these responses.
Overall, this is a comprehensive study, carefully executed, analyzed, and interpreted which clearly opens the possibility for numerous in-depth and detailed analyses at the isolate or function level.
We thank the Reviewer for the appreciation of the quality and robustness of our work and the value of our dataset and analysis.

Comments:
1. The authors analyze the data for each commensal based on differential abundances of transcripts in planta versus in vitro, but the interpretation is then made between commensals or between plants inoculated with different commensals. Bacterial functions are differently expressed by commensals already in the in vitro state, as shown in Figure S9, thus analyzing the data only from the perspective of in planta versus in vitro provides a unilateral view.
Our response 31
We do not know whether gene expression levels and/or expression changes are important for bacterial functions. To compare different commensals, we chose to show expression changes. Nevertheless, to address this, we also analyzed data with gene expression levels and compared with expression changes (Figure S9) (now Appendix Figure S9). Generally, we could capture similar trends, indicating that biologically meaningful information could be captured by expression changes.

2. The degree of plant gene expression varies between inoculations with different commensals. Could this be explained by the observed differences in the bacterial population (Fig.S1), or by the differences in their native levels of expression of specific functions (point1)
Our response 32
We apologize for the confusion. In Figure S1 (now Appendix Figure S1), we measured colonization abilities of the commensal strains at 3 dpi, but we took transcriptome data at 6 hpi to avoid the effects of differential bacterial populations on plant and bacterial

transcriptome. We have revised the text. Now it reads: “These strains could colonize in the leaf endosphere to various degrees three days after inoculation when inoculated on the leaf surface (Appendix Fig. S1).”

3. Line135. Transcriptional responses are analyzed at 6hpi. This is fine and well justified by the authors, however, it is important to mention in the abstract and consider throughout the manuscript that these are the early responses of the interacting partners.
Our response 33
Thank you for the suggestion.
We have added six hours after infection in Abstract or indicated “early responses” in several parts.

4. Line 139. The authors chose to compare expression levels of bacterial functions in planta versus in vitro, alias growth in rich media. While this can be explained by the rationale of the experimental setup it is important to explain to the reader why a rich media was chosen as a reference. Having the baseline as growth in a nutrient-rich environment explains the reduction in "ribosome" activity as mentioned by the authors, thus maybe creating a distorted view on which functions are activated or deactivated in the apoplast. One might imagine bacteria rarely encounter rich media in natural environments, thus the comparison baseline will be completely different.
Our response 34
For the pathogen *Pto*, expression levels of ribosome genes decreased in the minimum medium while increased upon transition from the rich medium to plant. On the other hand, in the disarmed strain of *Pto* (D36E), expression levels of ribosome genes decreased upon transition from the rich medium to plant. This suggests that the reduction in ribosome activity in the commensal strains cannot be simply explained by the fact that we used the rich medium for *in vitro* samples. Please also see Our response 40.

5. Line 186. Results presented here come as an argument for comment 4. Pto varies in responses when grown in rich versus minimal media, thus it will be relevant to discuss which functions in bacteria and in the plant remain constant independent of the baseline. Is it possible to identify those for Pto and draw comparisons with responses in commensals?
Our response 35
For instance, in Figure S3A (now Appendix Figure S3A), we showed bacterial processes which are not significantly affected upon the transition from the rich medium but affected upon the transition from the rich medium to the host plant. For instance, sulfur metabolism genes are specifically affected in planta (please also see Figure 4A and 4B). We think that readers can extract information of their interest.

6. Line 204. The authors likely mean that overcoming PTI is a prerequisite for an active metabolism in infectious bacteria. It would be interesting to know if plants impaired in PTI would enable commensals to activate a larger metabolic activity, and which ones would those be.
Our response 36
This will indeed be one of our future research topics.

7. Line 217. It is not entirely clear which physiological responses the authors refer to here. This sentence might require revision.

Our response 37
We have added (Fig. 2D) at the end of the sentence.

8. Line 261. It is relevant to emphasize that upregulated genes were primarily strain-specific, while those downregulated are general. In this regard, it will be relevant to add to figures S4B and D the number of genes which are unique to the analyzed isolates. Which functions do these genes have? Do they contribute to other differences depicted in the other figures?
Our response 38
We have revised the text according to the suggestion. Now it reads: "*In both of these phyla, a large number of genes were commonly suppressed among the three strains in plants but, induced genes were primarily strain-specific (Appendix Fig. S4B, S4D, and S5).*"
Regarding the strain specific genes, in Figure S4, we used shared genes among the three strains of Bacteriodetes or Proteobacteria in order to compare expression changes of the same genes in different strains.

9. Line 290. Are these functions activated in planta or just have a higher level of expression? It will be important to make a distinction between activated versus higher/lower levels of expression.
Our response 39
We have replaced "activated" with "induced".

10. Line296. "Ribosome" suppression is mentioned here as a being a consequence of plant-imposed control of bacterial growth, however, there is no data supporting this statement. This could be a native response of bacteria when transiting from rich to low nutrient conditions and not a plant-driven effect. A relevant reference could be the response in the presence of Arabidopsis apoplastic fluid.
Our response 40
For the pathogen *Pto*, expression levels of ribosome genes decreased upon transition from the rich medium to minimum medium while increased upon transition from the rich medium to plant. On the other hand, in the disarmed strain of *Pto* (D36E), expression levels of ribosome genes decreased upon transition from the rich medium to plant. This suggests that reduction in ribosome activity in the commensal strains cannot be simply explained by the transition to the low nutrient condition. To make this point clearer, we have added a new sentence: "*The contrasting pattern between commensals and Pto supports that suppression of these genes in commensals is not merely an artifact caused by the transition from the rich media to the plant apoplast.*"

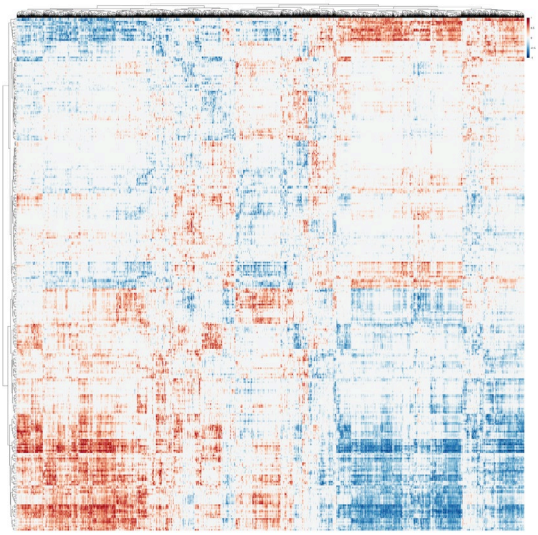
11. Line 324. Could it be envisioned that both degradation and nutrient sources for growth are at play?
Our response 41
Yes, we think both are possible. We hope that our message is clear in the original text.

12. Line 354. The statement regarding sugar transporters of Pto being not differentially expressed in planta is difficult to understand in the light of the given reference.
Our response 42
We have removed the statement as it was confusing.

13. Line 394. Co-transcriptome analyses identified that most genes of Arabidopsis and bacteria are not correlated. This result is surprising, and the question is whether this is a result of the input data used for the analysis (DEGs for in planta vs in vitro). Would the analysis give similar results if genes expressed in planta will be considered instead?

Our response 43

We used all shared bacterial genes for co-transcriptome analysis in Figure 6. We also tested using only *in planta* bacterial transcriptome data and obtained a similar result.



A map of correlation coefficients between plant genes and bacterial OGs. Rows and columns are bacterial OGs and plant genes, respectively.

14. Line 579- Methods. It is not clear if the expression of bacterial genes in vitro is performed on washed and water resuspended cells or directly from the cultures grown on rich media

Our response 44

We analyzed bacteria from the cultures grown in rich media. We revised the text. Now it reads: “0.1 volume of the stop buffer (95% EtOH, 5% Phenol) was directly added to liquid bacterial cultures before centrifuging to collect bacterial cells.“

Dear Prof. Tsuda,

Thank you for the submission of your revised manuscript to our editorial offices. I have now received the reports from the three referees that I asked to re-evaluate your study, you will find below. As you will see, the referees now fully support the publication of your study.

Before proceeding with formal acceptance, I have these editorial requests I ask you to address in a final revised manuscript:

- I would suggest this more concise title:
Dissecting the co-transcriptional landscape of plants and related microbiota

- Please add up to 5 keywords to the title page.

- Please provide the abstract written in present tense throughout.

- We updated our journal's competing interests policy in January 2022 and request authors to consider both actual and perceived competing interests. Please review the policy https://www.embopress.org/competing-interests and update your competing interests if necessary. Please name this section 'Disclosure and Competing Interests Statement' and put it after the Acknowledgements section.

- We now use CRediT to specify the contributions of each author in the journal submission system. CRediT replaces the author contribution section. Please use the free text box to provide more detailed descriptions. Thus, please remove the author contributions section from the manuscript text file. See also guide to authors:
https://www.embopress.org/page/journal/14693178/authorguide#authorshipguidelines

- Please order the manuscript sections like this (using these names):
Title page - Abstract - Keywords - Introduction - Results - Discussion - Materials & Methods - DAS (data availability section) - Acknowledgements - Disclosure and Competing Interests Statement - References - Figure legends

- Please make sure that the number "n" for how many independent experiments were performed, their nature (biological versus technical replicates), the bars and error bars (e.g. SEM, SD) and the test used to calculate p-values is indicated in the respective figure legends (main and Appendix figures), and that statistical testing has been done where applicable. Please avoid phrases like 'independent experiment', but clearly state if these were biological or technical replicates. Please add complete statistical testing to all diagrams (main, EV and Appendix figures). Please also indicate (e.g. with n.s.) if testing was performed, but the differences are not significant. In case n=2, please show the data as separate datapoints without error bars and statistics.

- Please add the legend for Table EV1 to the file itself and add a title. Then please upload the revised file. Finally, please remove the legend for the EV table from the manuscript text file.

- There are six source data files uploaded. Please provide these with a title and legend on the first TAB of the excel file. Moreover, please indicate there and in the file name to which figure these source data belong. Finally, please remove the mention/titles of these source data files from the Appendix file.

- Please format the references and the callouts according to our reference style:
http://www.embopress.org/page/journal/14693178/authorguide#referencesformat

- Finally, please find attached a word file of the manuscript text (provided by our publisher) with changes we ask you to include in your final manuscript text, and some queries, we ask you to address. Please provide your final manuscript file (using the attached file as basis) with track changes, in order that we can see any modifications done.

In addition, I would need from you:
- a short, two-sentence summary of the manuscript (not more than 35 words).
- two to four short bullet points highlighting the key findings of your study (two lines each).
- a schematic summary figure (in jpeg, png or tiff format with the exact width of 550 pixels and a height of not more than 400 pixels) that can be used as a visual synopsis on our website.

I look forward to seeing the final revised version of your manuscript when it is ready. Please let me know if you have questions regarding the revision.

Please use this link to submit your revision: https://embor.msubmit.net/cgi-bin/main.plex

Best,

Achim Breiling
Senior Editor
EMBO Reports

Referee #1:

The authors corrected the major issues that appeared in their original submission. My group and I are happy with the final manuscript which fits well EMBO reports scope and will be appreciated and nicely cited in our view.

Best regards,
Asaf Levy

Referee #2:

I thank the authors to their responses to my concerns, which have now been adequately addressed. I believe this manuscript warrants publication.

Referee #3:

The authors have revised the manuscript taking in comments and suggestions from the first review process. They also provided satisfactory responses to the reviewer's comments.

The authors performed the requested editorial changes.

Prof. Kenichi Tsuda
Huazhong Agricultural Univeristy
No.1 Shizishan Road, Hongshan
Wuhan, Hubei 430070
China

Dear Prof. Tsuda,

I am very pleased to accept your manuscript for publication in the next available issue of EMBO reports. Thank you for your contribution to our journal.

At the end of this email I include important information about how to proceed. Please ensure that you take the time to read the information and complete and return the necessary forms to allow us to publish your manuscript as quickly as possible.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File to accompany accepted manuscripts. As you are aware, this File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

If you do NOT want this File to be published, please inform the editorial office within 2 days, if you have not done so already, otherwise the File will be published by default [contact: emboreports@embo.org]. If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case." Please note that the author checklist will still be published even if you opt out of the transparent process.

Thank you again for your contribution to EMBO reports and congratulations on a successful publication. Please consider us again in the future for your most exciting work.

Yours sincerely,

Achim Breiling
Editor
EMBO Reports

****

THINGS TO DO NOW:

Once your article has been received by Wiley for production, the corresponding author will receive an email from Wiley's Author Services system which will ask them to log in and will present them with the appropriate license for completion.

You will receive proofs by e-mail approximately 2-3 weeks after all relevant files have been sent to our Production Office; you should return your corrections within 2 days of receiving the proofs.

Please inform us if there is likely to be any difficulty in reaching you at the above address at that time. Failure to meet our deadlines may result in a delay of publication, or publication without your corrections.

All further communications concerning your paper should quote reference number EMBOR-2022-55380V3 and be addressed to emboreports@wiley.com.

Should you be planning a Press Release on your article, please get in contact with emboreports@wiley.com as early as possible, in order to coordinate publication and release dates.

EMBO Press Author Checklist

Corresponding Author Name: Kenichi Tsuda	**USEFUL LINKS FOR COMPLETING THIS FORM**
Journal Submitted to: EMBO Reports	The EMBO Journal - Author Guidelines
Manuscript Number: EMBOR-2022-55380V1	EMBO Reports - Author Guidelines
	Molecular Systems Biology - Author Guidelines
	EMBO Molecular Medicine - Author Guidelines

Reporting Checklist for Life Science Articles (updated January 2022)

This checklist is adapted from Materials Design Analysis Reporting (MDAR) Checklist for Authors. MDAR establishes a minimum set of requirements in transparent reporting in the life sciences (see Statement of Task: 10.31222/osf.io/9sm4x). Please follow the journal's guidelines in preparing your manuscript.

Please note that a copy of this checklist will be published alongside your article.

Abridged guidelines for figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- ideally, figure panels should include only measurements that are directly comparable to each other and obtained with the same assay.
- plots include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n<5, the individual data points from each experiment should be plotted. Any statistical test employed should be justified.
- Source Data should be included to report the data underlying figures according to the guidelines set out in the authorship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements.
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.
- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:

  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Please complete ALL of the questions below.
Select "Not Applicable" only when the requested information is not relevant for your study.

Materials

Newly Created Materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
New materials and reagents need to be available; do any restrictions apply?	Not Applicable	

Antibodies	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
For **antibodies** provide the following information: - Commercial antibodies: RRID (if possible) or supplier name, catalogue number and or/clone number - Non-commercial: RRID or citation	Not Applicable	

DNA and RNA sequences	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Short novel DNA or RNA including primers, probes: provide the sequences.	Not Applicable	

Cell materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Cell lines: Provide species information, strain. Provide accession number in repository **OR** supplier name, catalog number, clone number, and/**OR** RRID.	Not Applicable	
Primary cultures: Provide species, strain, sex of origin, genetic modification status.	Not Applicable	
Report if the cell lines were recently **authenticated** (e.g., by STR profiling) and tested for mycoplasma contamination.	Not Applicable	

Experimental animals	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Laboratory animals or Model organisms: Provide species, strain, sex, age, genetic modification status. Provide accession number in repository **OR** supplier name, catalog number, clone number, **OR** RRID.	Not Applicable	
Animal observed in or captured from the field: Provide species, sex, and age where possible.	Not Applicable	
Please detail **housing and husbandry conditions**.	Not Applicable	

Plants and microbes	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Plants: provide species and strain, ecotype and cultivar where relevant, unique accession number if available, and source (including location for collected wild specimens).	Yes	Materials and Methods
Microbes: provide species and strain, unique accession number if available, and source.	Yes	Materials and Methods

Human research participants	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If collected and within the bounds of privacy constraints report on age, sex and gender or ethnicity for all study participants.	Not Applicable	

Core facilities	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If your work benefited from core facilities, was their service mentioned in the acknowledgments section?	Not Applicable	

Design

Study protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If study protocol has been **pre-registered, provide DOI in the manuscript**. For clinical trials, provide the trial registration number **OR** cite DOI.	Not Applicable	
Report the **clinical trial registration number** (at ClinicalTrials.gov or equivalent), where applicable.	Not Applicable	

Laboratory protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Provide DOI OR other citation details if **external detailed step-by-step protocols** are available.	Not Applicable	

Experimental study design and statistics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Include a statement about **sample size** estimate even if no statistical methods were used.	Yes	Materials and Methods
Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. **randomization procedure**)? If yes, have they been described?	Yes	Materials and Methods
Include a statement about **blinding** even if no blinding was done.	Not Applicable	
Describe **inclusion/exclusion criteria** if samples or animals were excluded from the analysis. Were the criteria pre-established? If sample or data points were omitted from analysis, report if this was due to attrition or intentional exclusion and provide justification.	Not Applicable	
For every figure, are **statistical tests** justified as appropriate? Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. Is there an estimate of variation within each group of data? Is the variance similar between the groups that are being statistically compared?	Yes	Materials and Methods

Sample definition and in-laboratory replication	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
In the figure legends: state number of times the experiment was **replicated** in laboratory.	Yes	Figure legends
In the figure legends: define whether data describe **technical or biological replicates**.	Yes	Figure legends

Ethics

Ethics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Studies involving **human participants**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval.	Not Applicable	
Studies involving **human participants**: Include a statement confirming that **informed consent** was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	Not Applicable	
Studies involving **human participants:** For publication of **patient photos**, include a statement confirming that consent to publish was obtained.	Not Applicable	
Studies involving experimental **animals**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval. Include a statement of compliance with ethical regulations.	Not Applicable	
Studies involving **specimen and field samples:** State if relevant **permits** obtained, provide details of authority approving study; if none were required, explain why.	Not Applicable	

Dual Use Research of Concern (DURC)	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Could your study fall under dual use research restrictions? Please check biosecurity documents and list of **select agents and toxins** (CDC): https://www.selectagents.gov/sat/list.htm	Not Applicable	
If you used a select agent, is the security level of the lab appropriate and reported in the manuscript?	Not Applicable	
If a study is subject to dual use research of concern regulations, is the name of the **authority granting approval and reference number** for the regulatory approval provided in the manuscript?	Not Applicable	

Reporting

The MDAR framework recommends adoption of discipline-specific guidelines, established and endorsed through community initiatives. Journals have their own policy about requiring specific guidelines and recommendations to complement MDAR.

Adherence to community standards	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
State if relevant guidelines or checklists (e.g., **ICMJE, MIBBI, ARRIVE, PRISMA**) have been followed or provided.	Not Applicable	
For **tumor marker prognostic studies,** we recommend that you follow the **REMARK** reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	Not Applicable	
For **phase II and III randomized controlled trials**, please refer to the **CONSORT** flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	Not Applicable	

Data Availability

Data availability	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Have **primary datasets** been deposited according to the journal's guidelines (see 'Data Deposition' section) and the respective accession numbers provided in the Data Availability Section?	Yes	Gene Expression Omnibus database (accession no. GSE150422)
Were **human clinical and genomic datasets** deposited in a public access-controlled repository in accordance to ethical obligations to the patients and to the applicable consent agreement?	Not Applicable	
Are **computational models** that are central and integral to a study available without restrictions in a machine-readable form? Were the relevant accession numbers or links provided?	Not Applicable	
If publicly available data were reused, provide the respective **data citations in the reference list**.	Yes	References