Genotypes to be tested via metabolomics:

C= Cha-Gal4/+; silent/silent

E= Cha-Gal4/+; G85Rcr/G85Rcr

A= Cha-Gal4/+; UAS-Tkt-silent/silent

F= Cha-Gal4/+; UAS-Tkt-G85Rcr/G85Rcr

B= Cha-Gal4/+; Df-8143(tkt) silent/silent

D= Cha-Gal4/+; Df-8143(tkt) G85Rcr/G85Rcr

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| |  | | --- | | **Ram Prosad Chakrabarty** | | AttachmentsFeb 24, 2022, 7:20 PM |  |  |
| |  | | --- | | to Yuliya, me, Navdeep | | | |

Hi Julia,

I'm sorry for taking a longer than usual turnaround time; we were optimizing our target list. Please find the results attached and note that I relabel the samples as follows:

A1-A4 (1, 7, 13, 19)

B1-B4 (2, 8, 14, 20)

C1-C4 (3, 9, 15, 21)

D1-D4 (4, 10, 16, 22)

E1-E4 (5, 11, 17, 23)

F1-F4 (6, 12, 18, 24)

Please let me know if you need any help in analyzing the data; I'd be happy to help. Thanks much!

Best,

Ram

**Ram Prosad Chakrabarty**

PhD candidate

[Chandel Lab](https://www.chandellab.com/)

Driskill Graduate Program (DGP) in Life Sciences

Northwestern University Feinberg School of Medicine

**Cell:**773 837 8436

303 E. Superior Street, Simpson Querry 5-300-F

Chicago, IL 60611

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| |  | | --- | | **Ram Prosad Chakrabarty** | | Fri, Feb 25, 10:30 AM |  |  |
| |  | | --- | | to Yuliya, me, Navdeep | | | |

Hi Julia,

Sure, please feel free to email me for any help. Here's a step-by-step basic metabolomics data analysis guide you may follow initially:

1. You've received an excel file with the calculated peak areas corresponding to approximately 200 soluble metabolites. The samples are organized into columns and the metabolites are organized into rows  
2. Select all and copy.  
3. Open a new excel file and use “Paste Values” in this new sheet  
4. In the new file, delete the column with the KEGG IDs  
5. Change the name of the TIC or normalization factor to “TIC” and delete the other row, this will be used for normalization as a proxy for the total number of ions detected by the instrument for each sample. If these numbers are very different between the samples, then there was probably a problem with the run or with your extraction or your estimation of starting material. You can use either the normalization factor or the TIC for this, they are just scaled proportions of one another.  
6. Delete the other rows (with sample information etc.) at the top so you just have TIC on row one, then the first metabolite on row 2.  
7. Insert 2 rows above the row of the TIC  
a. On row one, column one, write “Sample” then label your samples/columns on the first row (each sample must have a unique name.)  
b. On row two write “Group” then label each sample/column with a group title. You must have at least 3 samples per group in order to load the file into metaboanalyst, if you do not have 3 samples per group, you will need to do the analysis in excel/graphpad looking at individual metabolites of interest or running t-tests or something like that in excel after manual normalization to TIC.  
8. Remove “0” values  
a. Hit ctrl+F for Find and Replace function  
b. Go to the “Replace” tab  
c. Select “Options”  
d. Select “Match entire cell contents”  
e. In the “Find what” box type 0  
f. Make sure there is nothing written in the “Replace with” box  
9. Save as a tab delimited text file (.txt)  
10. Open [metaboanalyst.ca](http://metaboanalyst.ca/) website  
11. Hit Click to start  
12. Hit Statistical Analysis (one factor)  
13. For ‘Data Type:’ select “Peak Intensity Table”  
14. For ‘Format:’ select “Samples in Columns (unpaired)”

Select data file  
15. Hit Submit  
16. On the Data Integrity Check page, click “Skip” this will automatically impute small values for the zeroes that were removed  
17. On the Data Filtering page, click “None” then “Submit” then “Proceed”  
18. On the Normalization page, select “Normalize by reference feature”, then click “specify” and enter “TIC”, then click “Normalize” then click “Proceed”  
19. Now click on “One-way Analysis of Variance (ANOVA)”  
20. You can adjust the FDR as appropriate, or click on the little spreadsheet icon next to the paint pallet icon to see the list of metabolites that meet the specificed significance threshold. You can also then click on the individual metabolites to see graphs of the data for each individual metabolite, you can also click on the little green excel/table icon that will show you the ranked list of significant metabolites and the significant post hoc comparisons that met your threshold. If you want to see the p values for all metabolites, simply adjust the significance threshold to 1 and you will be able to see a ranked list of metabolites by ANOVA or Ttest p-value by again clickon on the excel/sheet icon.  
21. Next on the left side click on “PCA” look through the 2D Score Plot tab and the Biplot tab to see how well the groups cluster. You can click the paint pallet icon to generate downloadable plots.  
22. Next on the left side click on “Heatmap” select “Use top 25” or adjust to other numbers of metabolites for inclusion and hit submit to see a smaller list of metabolites based on the significance value of one-way ANOVAs. You can also stop the heatmap from automatically clustering the samples/columns by selecting “Do not reorganize: Samples” if needed. This can also be downloaded with the paint pallet icon.  
23. You can also generate a report of everything you have done with detailed descriptions of the analyses in the “Download” page on the left. When you hit “Generate Report” a link will appear to the right of the button that says “Analysis Report” you can save this linked PDF  
  
Pathway Analysis:  
1. Define your list of differentially abundant metabolites using an FDR of 0.05 or 0.1 in the ANOVA in step 17/18 above  
2. Go back to the original Excel sheet from Peng  
3. Find the KEGG ID’s corresponding to your list of metabolites  
4. Go back to the metaboanalyst start page, and instead of clicking on Statistical Analysis, choose Pathway Analysis  
5. You will need to remove the KEGG IDs that have multiple entries that Peng can’t differentiate between  
6. Paste your list of KEGG IDs  
7. Under Input Type, select KEGG ID  
8. Hit submit  
9. You will probably get some KEGG ID’s that aren’t recognized, you can look those up on KEGG and see if there is a more specific metabolite that you think might make more sense to replace.

Have a nice rest of the day!

Best,

Ram

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| |  | | --- | | **Ram Prosad Chakrabarty** | | Attachments5:01 AM (5 hours ago) |  |  |
| |  | | --- | | to Yuliya, me, Navdeep | | | |

Hi Julia,

I was just playing with the data to see its quality; it looks really good. I just wanted to share a few findings in case it's useful for you.

Different genotypes clustered nicely (please see the attached heatmap). It looks like genotypes A, B, and C cluster together, and genotypes D, E, and F cluster together (did you expect such clustering based on genotypes?). Also, I performed one-way ANOVA with an FDR cutoff of 0.05, and 66 metabolites were significantly different between the groups (please see the excel file). Many of these metabolites are really very interesting (please see the attached excel file and the heatmap)!

Please let us know if you have any questions or need any help. Thanks much and have a nice day!

Best,

Ram