Help Page – Confidence Application

Getting Started

Steps to using Confidence are very easy! If at any point you are stuck, hover over the button and it will give you more information. The workflow is also described here.

Step 0. Prepare your count matrix and metadata files

This is the most important step! Make sure that your matrix and counts files are formatted according to the format shown in the example. A formatted reference dataset from the publicly available dataset GSE263226 provided by Lemay et.al, 2025 is available for download as a reference. Please ensure that your files are formatted in the same manner.

Step 1. Upload your metadata file

Take your formatted metadata file and upload it in Step 1. Wait for the application to read the file. Your metadata will then be shown as a table.

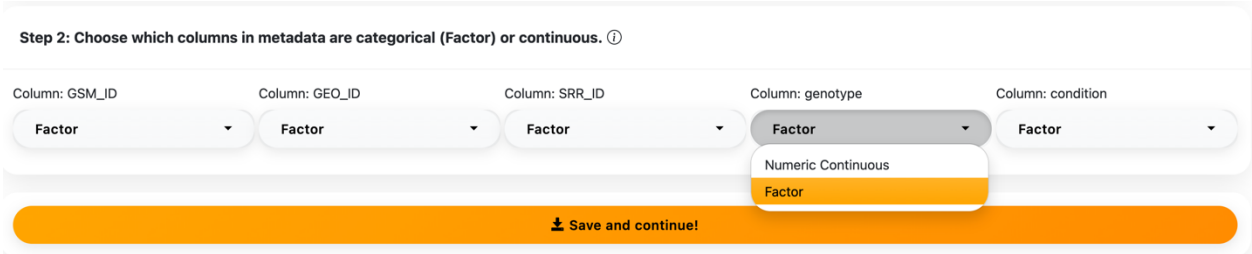
Show 10 entries Search:

	GSM_ID	GEO_ID	SRR_ID	genotype	condition
1	GSM8188131	GSE263226	SRR28557449	CTRL_PASMC	Control
2	GSM8188132	GSE263226	SRR28557448	CTRL_PASMC	Control
3	GSM8188133	GSE263226	SRR28557447	CTRL_PASMC	Control
4	GSM8188134	GSE263226	SRR28557446	CTRL_PASMC	Control
5	GSM8188135	GSE263226	SRR28557445	PAH PASMC	PAH
6	GSM8188136	GSE263226	SRR28557444	PAH PASMC	PAH
7	GSM8188137	GSE263226	SRR28557443	PAH PASMC	PAH
8	GSM8188138	GSE263226	SRR28557442	PAH PASMC	PAH
9	GSM8188139	GSE263226	SRR28557441	PAH PASMC	PAH

Showing 1 to 9 of 9 entries Previous 1 Next

Step 2. Choose which columns are categorical or continuous

This step is important for downstream analysis. Make sure you choose the correct option for each metadata column as it will affect the statistical tests used downstream.

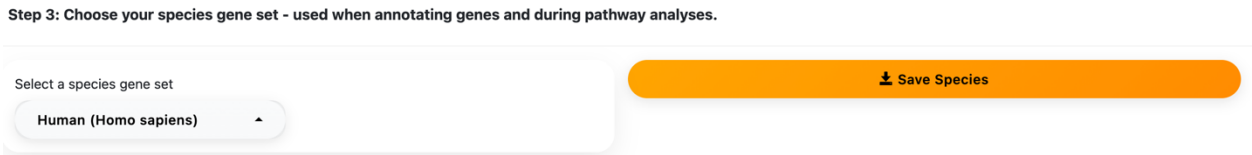


Step 3. Choose your gene species set

Based on your data, select the correct species reference genome dataset. Confidence currently supports 4 species: Rat, Mouse, Rabbit, and Human, all containing updated reference datasets.

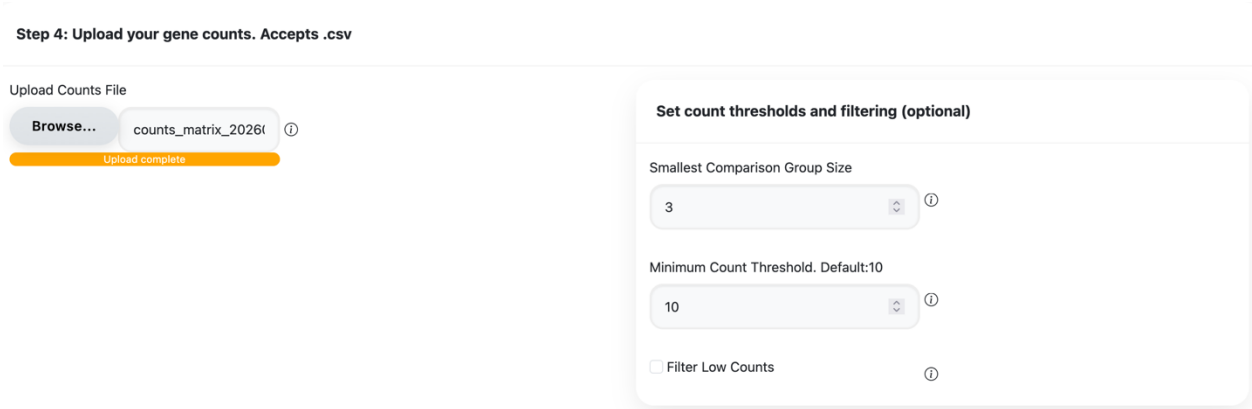
Warning: If you aligned your data using an older reference genome, you will get an error. Please verify your reference genome version before selecting.

Version Update: February 26th, 2026

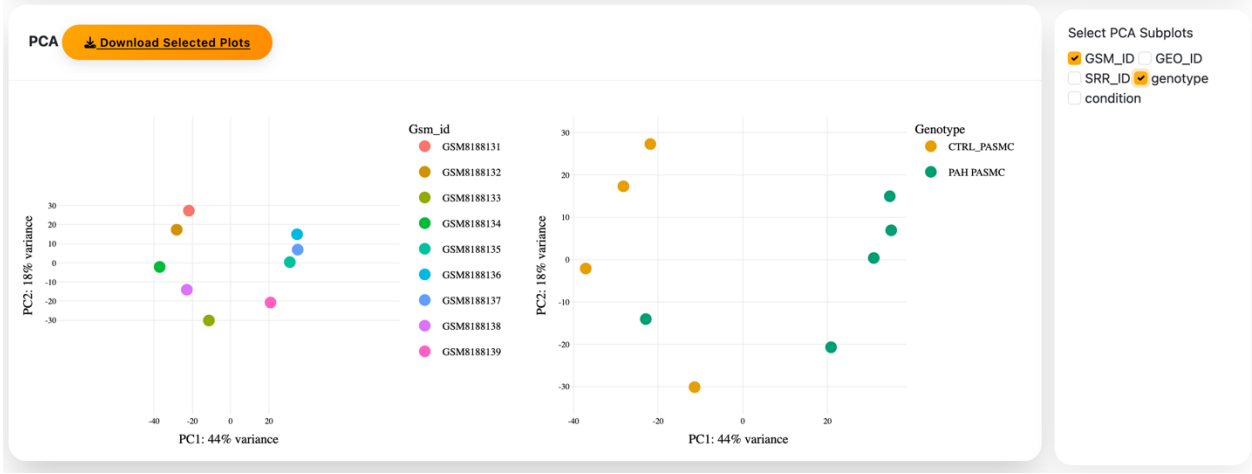


Step 4. Upload your counts matrix file

Upload your formatted counts file. Remember to make sure both your counts and metadata files are properly formatted before uploading.



Heads up: Confidence will flicker after uploading and this is completely normal! Do not hit refresh! Depending on file size it may take a few minutes. Once complete, PCA plots will appear so you can check for batch effects, filter low counts, and set minimum count thresholds. If you are filtering counts, the applicaiton will load again, this is normal as it is adjusts the counts matrix automatically.



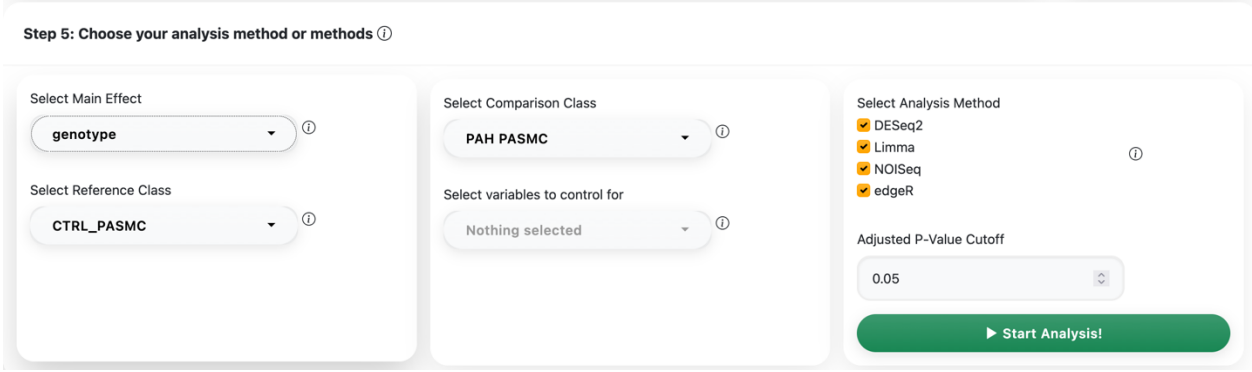
Step 5. Select your analysis methods

- **Main effect:** Select the variable you are studying. Options are drawn from your uploaded metadata.
- **Reference class:** Select the reference/control class
- **Comparison class:** Select the class you are comparing against (e.g disease/treatment)
- **Control variables (optional):** Add batch numbers or other covariates here.

Check the analysis methods you are interested in and hit the Start Analysis button.

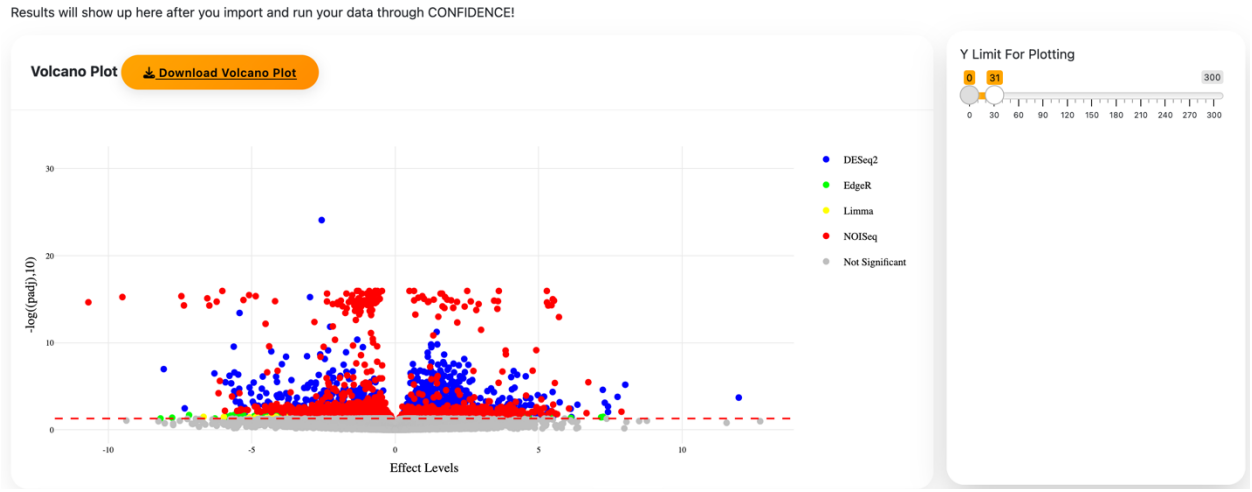
Version Update: February 26th, 2026

Important: Confidence will navigate to the Results section and show a loading spinner. Do not hit refresh as this will break the session. Let it run through.

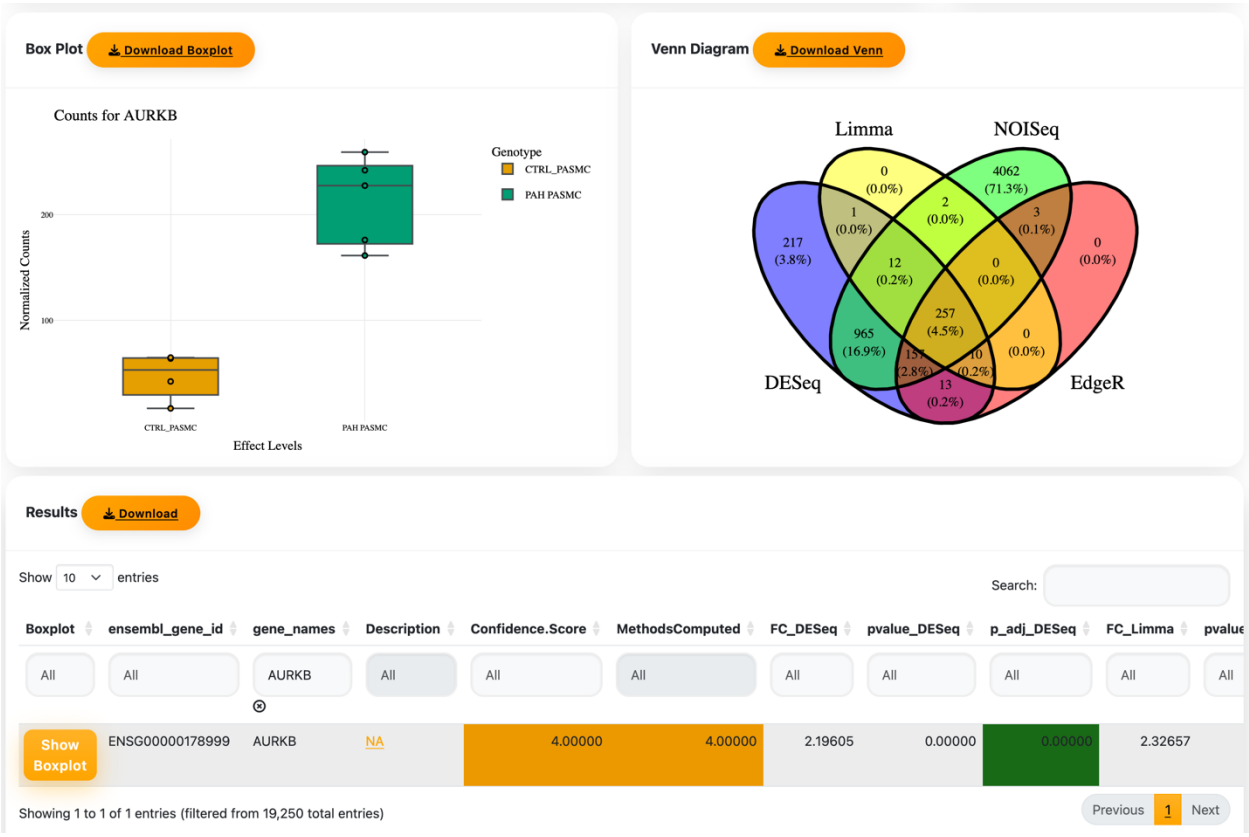


Step 6. Download and view results

- The volcano plot is interactive. NOISeq may produce many outliers causing results to cluster near the x-axis.
- Adjust the Y-axis limit to improve the plot. The app will reload briefly, do not refresh.

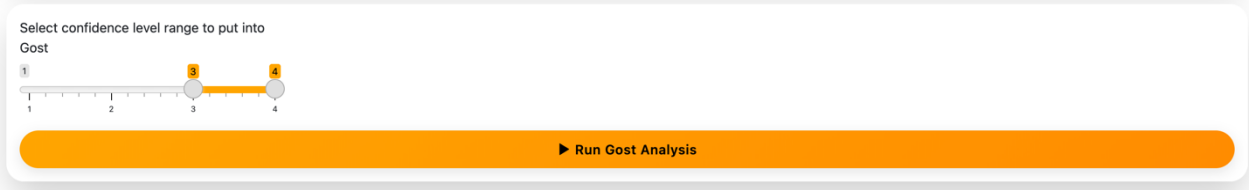


- Double-click a legend entry to isolate results from a specific method.
- The Venn diagram shows DEGs found across your selected methods.
- Scroll down to the results table. Enter a gene name to view its normalised counts box plot.

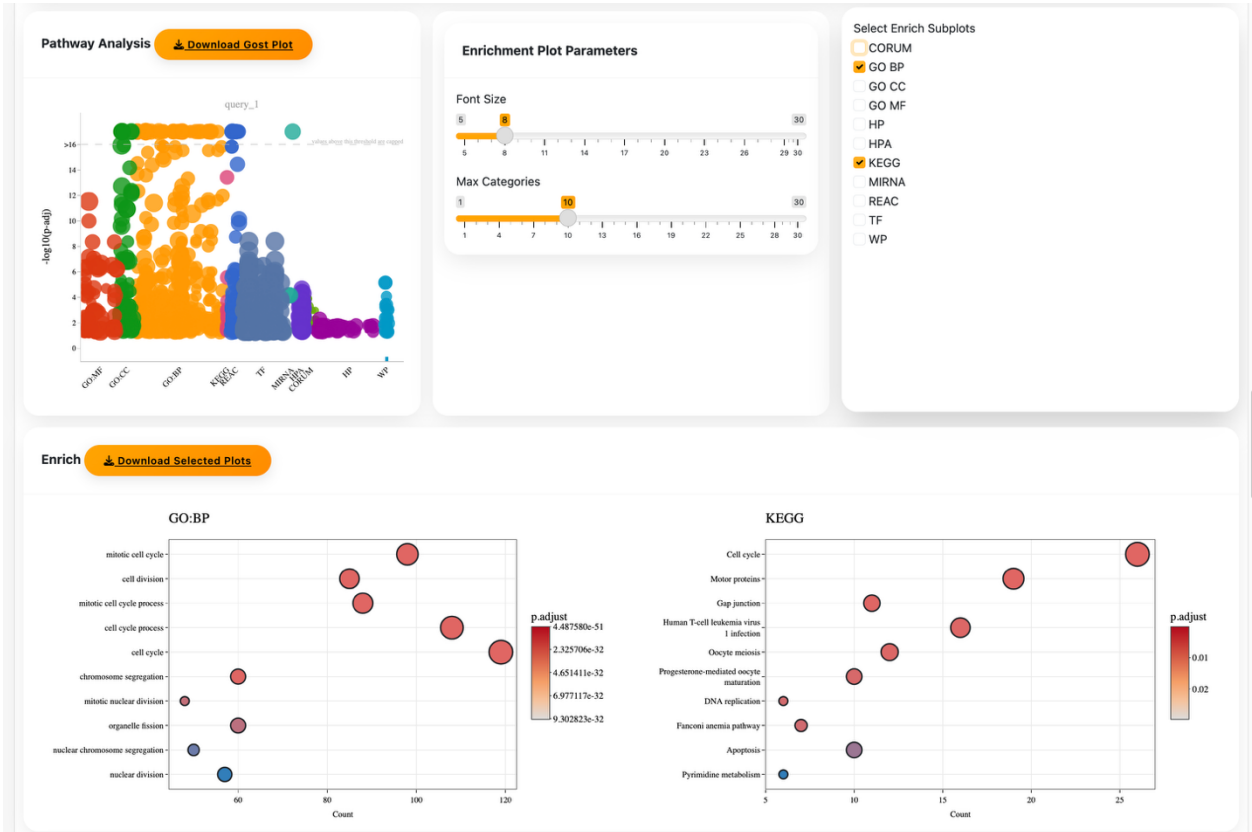


Step 7. Gene scoring and pathway analysis

Select genes using your desired Confidence score threshold and run the Gost analysis. Depending on the score and number of genes, this may take a while. If it appears stuck, adjust the score and try again.



Select your enrichment method (BP, KEGG, etc.) and download individual plots.



That's it! That's how you use Confidence.

Version Update: February 26th, 2026

FAQs

The most common reason Confidence fails to analyze data is related to formatting issues in the metadata and raw count files. Please review the common issues below. If Confidence does not perform as expected, please contact us at c.hindmarch@queensu.ca for assistance. In your help request, please include:

- Your count matrix file
- Your metadata file
- A description of the issue and/or a screenshot of the error message displayed

Q1.What format should my data be in?

Confidence requires the metadata and count matrix to be formatted in a specific structure. Please ensure that:

- The first column of the count matrix contains gene IDs (or EnsembleID).
- Column headers correspond exactly to sample IDs listed in the metadata file.
- The metadata file includes a column specifying the experimental condition.

Please refer to the processed dataset from GEO (GSE263226) as an example to model your file formatting.

Q2. Why is the app flickering or loading for a long time?

This is normal behaviour, especially after uploading files or starting analysis. The app is working in the background. Do not hit refresh as this will break your session.

Q3. Why does my volcano plot look strange?

NOISeq in particular can produce many outliers, causing results to cluster near the x-axis. Use the Y-axis limit adjustment to rescale the plot to a more readable range.

Q4. What do I do if I get an error after uploading counts file?

Two reasons are likely:

- Your counts file is not formatted correctly. Please check your formatting and refer to the processed dataset from GEO (GSE263226) as an example to model your file formatting.
- Your data was aligned using an older reference genome that does not match the current reference dataset in Confidence. Please check the genome version used during alignment and ensure it matches the selected species reference.

If you are experiencing issues, please contact us with your counts file and metadata file for further assistance.

Q5. The Page Is Broken or Not Reloading

This may be due to temporary server issues. Please email us so we can investigate and restore functionality as quickly as possible.

Q6. Some Gene Names Do Not Appear in the Volcano Plot

This is a known issue related to gene annotation retrieval from BioMart. If specific gene labels are missing, you can download the full results table to verify statistical results. Alternatively, you can use the gene-specific box plot function to visualize normalized counts for genes of interest.