ProntoPCR

Handbook

**Contents**

[**Overview** 2](#_Toc177377068)

[**Instructions for Use** 2](#_Toc177377069)

[**Local Access (Recommended)** 2](#_Toc177377070)

[For New R Users: 2](#_Toc177377071)

[For Existing R Users: 3](#_Toc177377072)

[**Online Access** 3](#_Toc177377073)

[**How to Use** 3](#_Toc177377074)

[**Data Input** 3](#_Toc177377075)

[**Calculations** 6](#_Toc177377076)

[**Statistics** 11](#_Toc177377077)

[**Graphs** 19](#_Toc177377078)

[**Error Messages** 32](#_Toc177377079)

[**Input Data Errors** 32](#_Toc177377080)

[**Version History** 36](#_Toc177377081)

[**Feature Requests and Issues** 36](#_Toc177377082)

[**References** 36](#_Toc177377083)



# **Overview**

ProntoPCR is a software application that aims to perform routine qPCR calculations efficiently and automatically, such as averaging the housekeeper genes, calculating ΔCq, ΔΔCq, relative fold-change (2^-ΔCq), and fold-change (2^-ΔΔCq) of the target gene. It also provides the user with options to perform statistics and graphing of the data. ProntoPCR is an open-source, R shiny (reference Shiny) software application available both online and locally as an R package (reference R), making it accessible to users with minimal programming experience. It is also designed to be compatible with the output generated by PCR machines, such as the Quantstudio 5 (Thermo Fisher Scientific) enabling the user to get instant results. Users only require a comma separated values (.csv) file of the raw data, including the quantification cycle (Cq) values, the target (gene) and sample name to get started.

# **Instructions for Use**

ProntoPCR has been designed to operate both online and locally. Whilst we aim to maintain online availability, the hosted platform may change or become unavailable. Note the online application times out after 5 minutes of inactivity. Therefore, it is recommended to rely on the local version, which functions with the same features as the online option. The local version also does not require internet access once installed.

To run the application locally, the user needs to download R: <https://cran.r-project.org/> and RStudio: <https://posit.co/downloads/>. A step-by-step guide on how to install the ProntoPCR application locally is available below, as well as this video. The guide is aimed at users with minimal programming expertise.

We request that users of ProntoPCR cite the associated journal article if the application has been used in analysis. The citation is as follows:

## **Local Access (Recommended)**

### **For New R Users:**

1. Download R and RStudio.

* <https://rstudio-education.github.io/hopr/starting.html>

1. Open RStudio
2. Within the Console, type:

Library(devtools)

Devtools::install\_github(“MarnieMaddock/ProntoPCR”)

1. Once installed, this will appear in your console:
2. To run the app, inside the console section, type:

ProntoPCR()

1. The ProntoPCR Interface will now appear in a separate window.

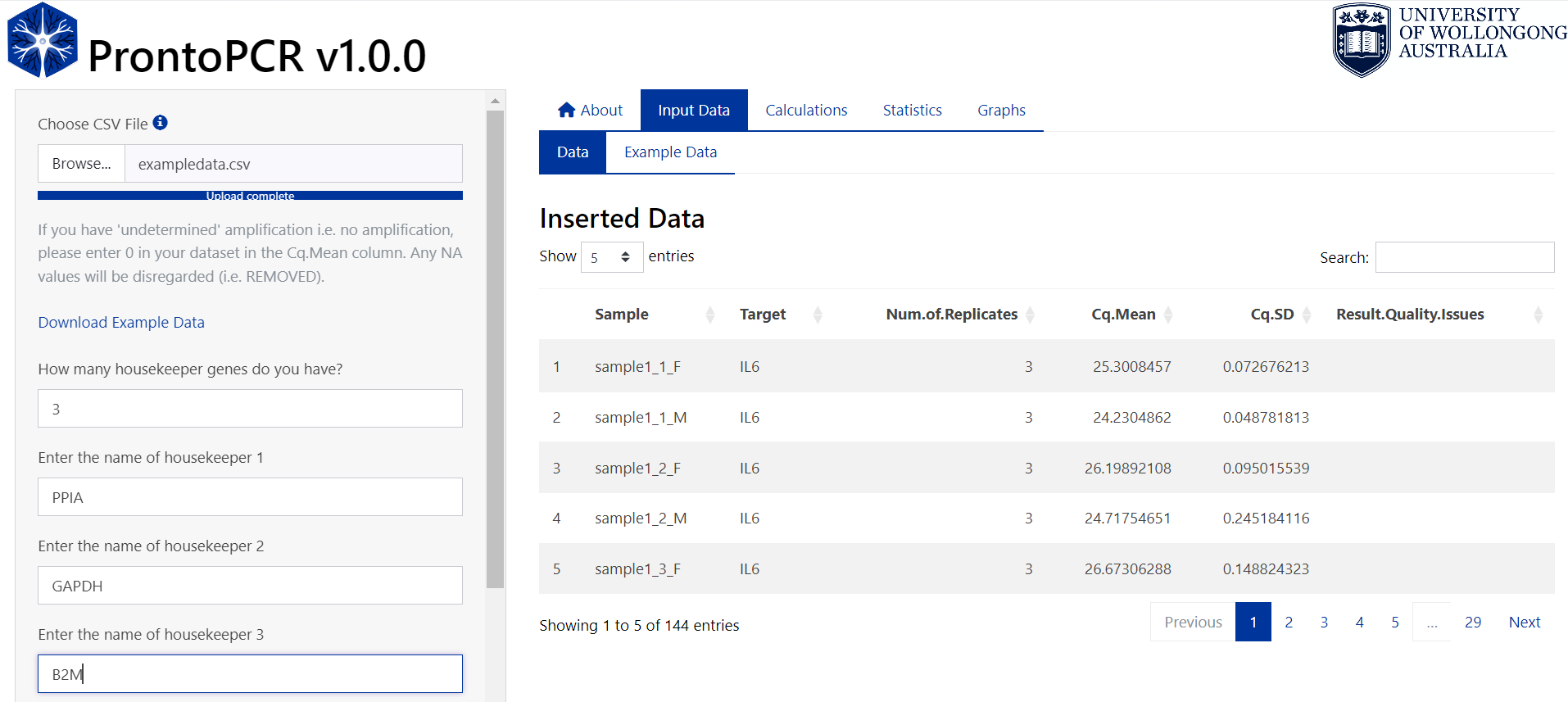
### **For Existing R Users:**

## **Online Access**

Access to ProntoPCR is available here: <https://marniem.shinyapps.io/ProntoPCR/>

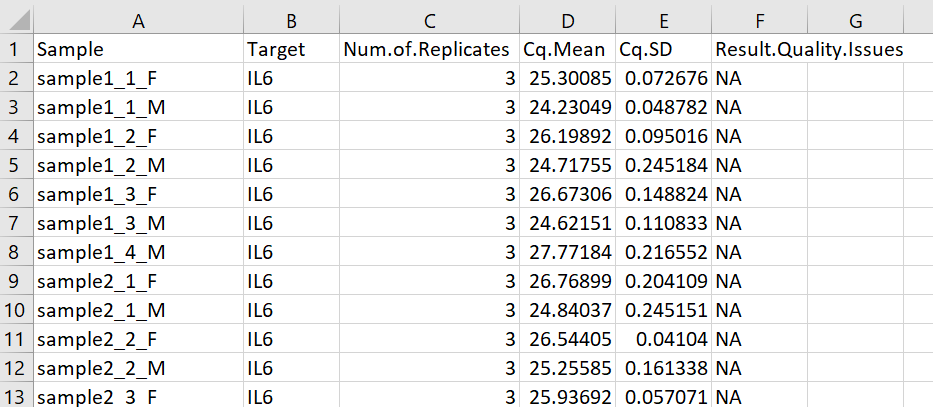
## **How to Use**

An example data file can be directly downloaded from the ProntoPCR app under the Input Data Tab. All example outputs in the images below have come from the example data .csv file.



### **Data Input**

1. Collate all PCR Cq values into one .csv file. It must be in the following format:

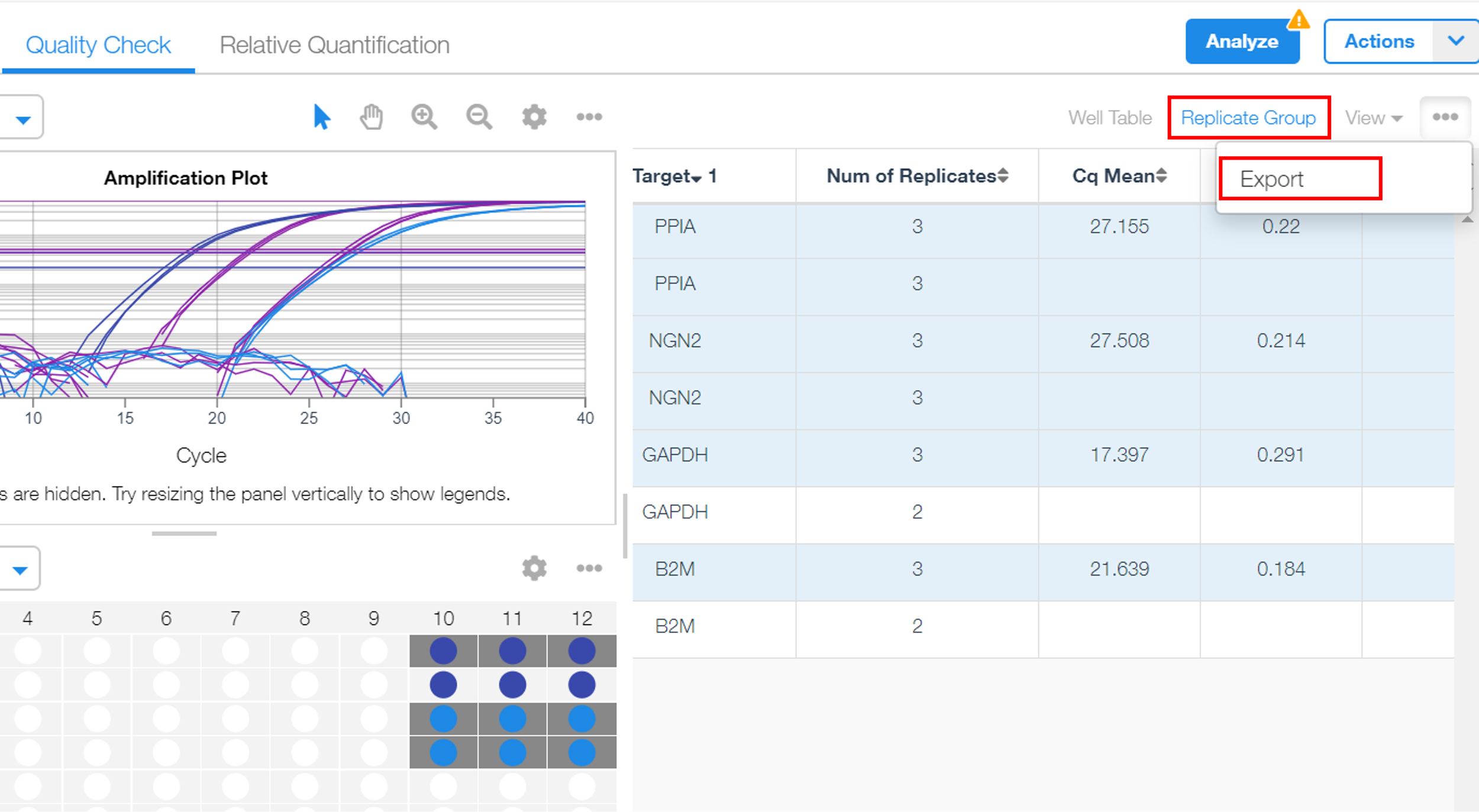


Checklist:

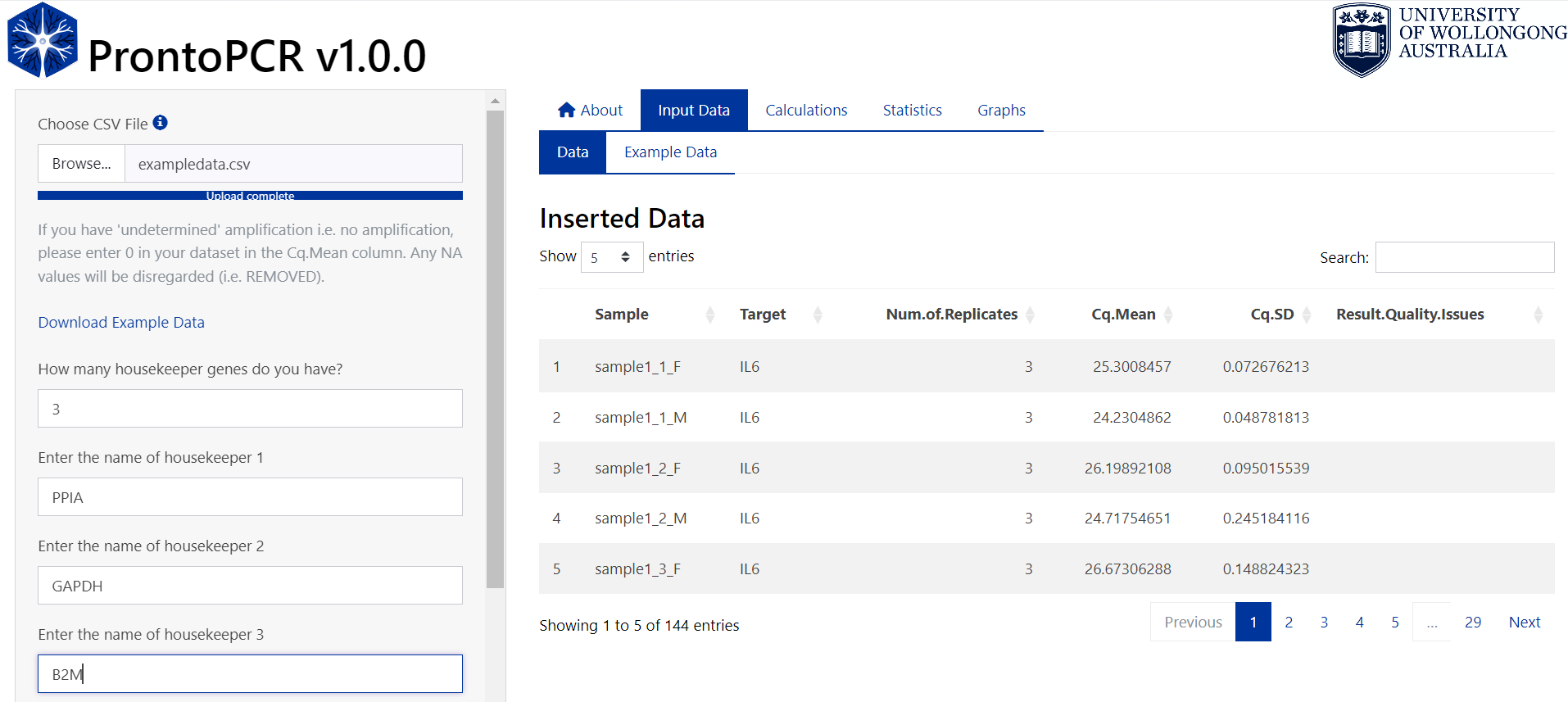
* CSV file containing PCR data with the formatting given above?
* Column names have capital letters when specified?
* Column names have full-stops . when specified?
* There are no spaces present in the column names?
* There are no gaps between rows in your dataset. Each row should have at least one measurement.
* If you have 'undetermined' amplification i.e. no amplification, 0 has been entered in your dataset for those instances. Any NA values will be disregarded.
* All Sample names have no spaces and use underscores (\_) in the naming system. Note the format is given by group1\_biologicalReplicateNumber\_group2. If there is no second group, write NIL. E.g. sample2\_1\_NIL.
* All non-template controls (NTC) and no reverse transcriptase controls (-RT) have been removed.

Note: This file format is compatible with the output given by the QuantStudio5 PCR machine (ThermoFisher Scientific). To do this, set up the plate format similar to below:

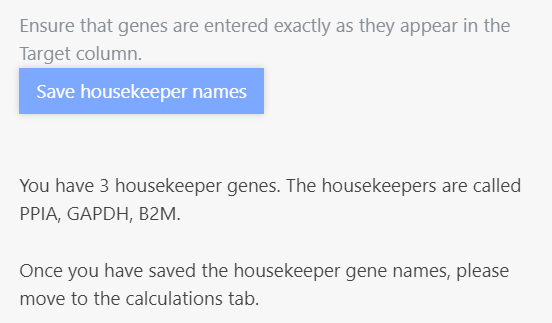
To export the results go to the Quality Check tab 🡪 Replicate Group 🡪 Export:



1. Open ProntoPCR and select the Input Data Tab.
2. Insert .csv file of PCR data by clicking on Browse…
3. Enter how many housekeeping genes to normalise the data to.
4. Enter the name of the housekeeping gene – exactly as it appears in the .csv file.

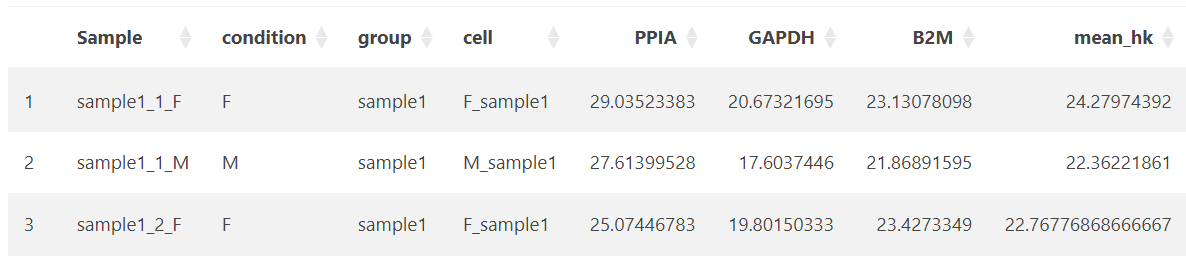


1. Save the names of the housekeeper genes by selecting the ‘Save housekeeper names’ button. An error message will appear if the housekeeper's name is incorrect or not found. If the names of the housekeepers are correct, they will be displayed below the button:

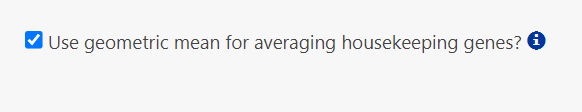


### **Calculations**

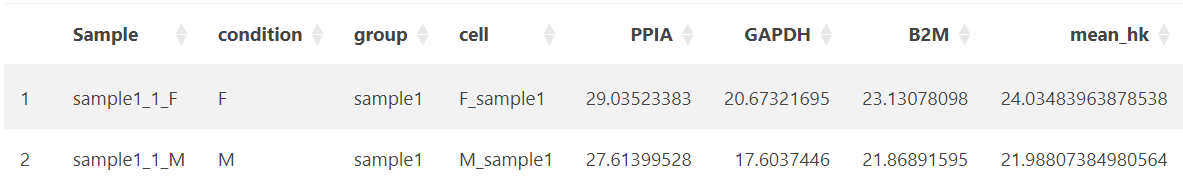
1. Select the Calculations Tab. The average of the housekeeper genes will automatically be calculated using arithmetic mean. This will be displayed as the mean\_hk column on the right. E.g. the mean\_hk for sample1\_1\_F is calculated as follows:



1. To use geometric mean, select this option on the side panel:



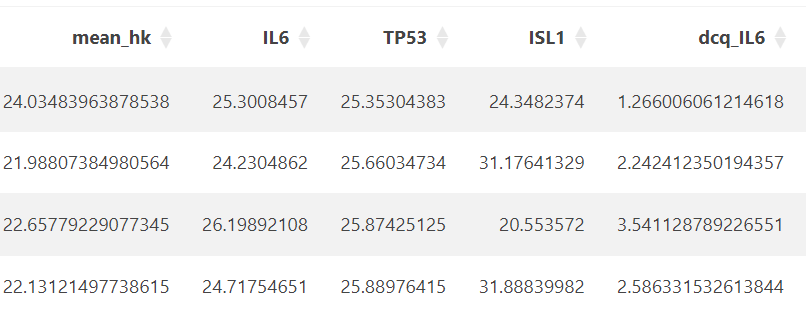
The geometric mean is calculated as follows for sample1\_1\_F:



1. The ΔCq is calculated as follows:

The ΔCq is displayed as dcq\_geneName within the table on the right. E.g. for the gene IL6, the ΔCq (dcq\_IL6) is calculated by:

25.3008457 – 24.0348396 = 1.2660060…

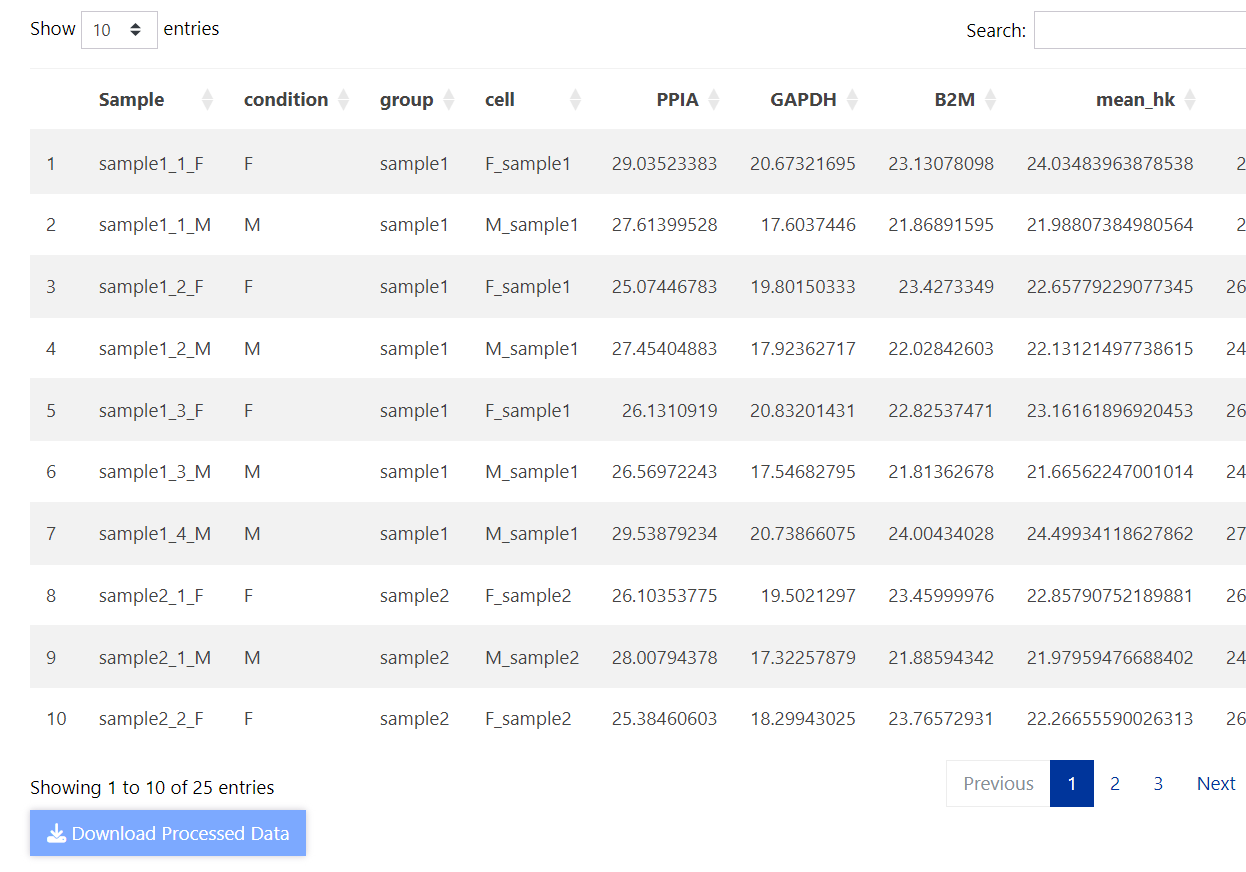


1. Relative fold change (2^(- ΔCq)) is automatically calculated and given as fc\_dcq\_gene within the table:

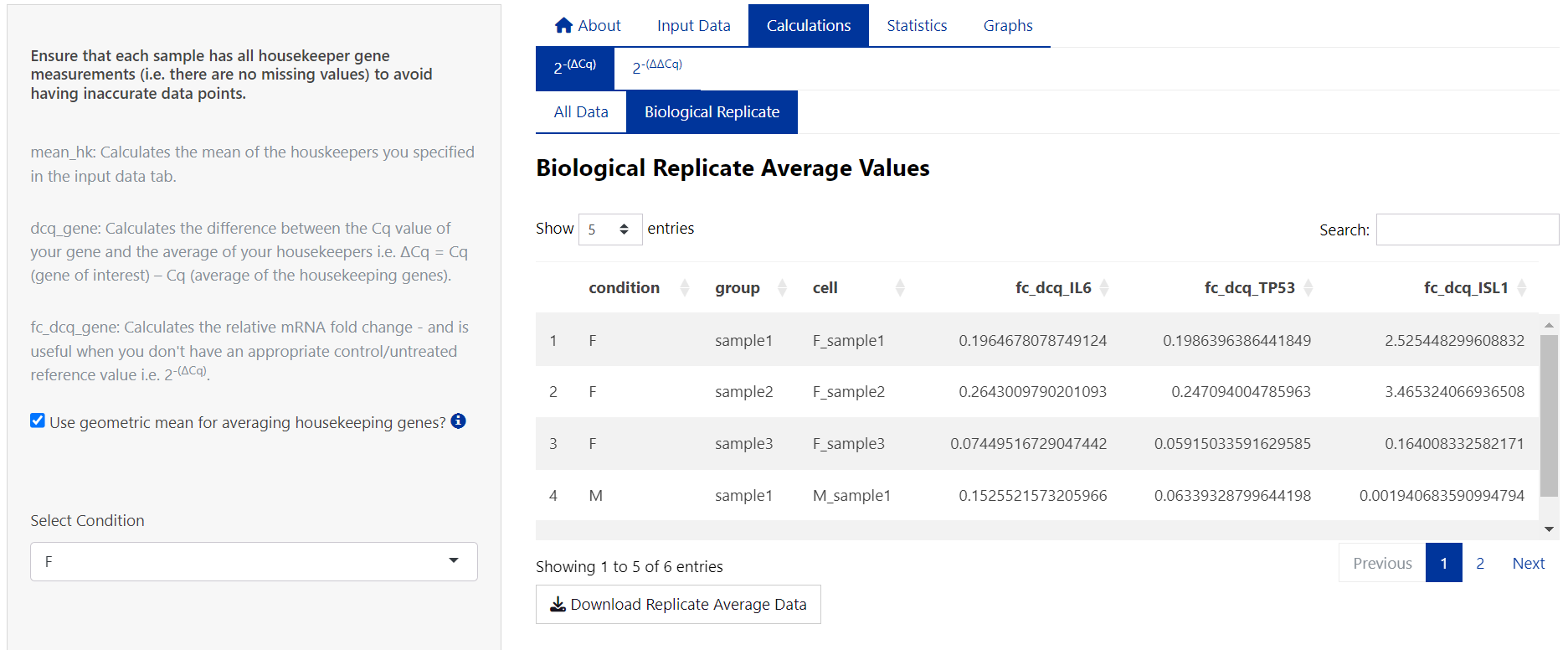
E.g. the relative fold change of IL6 (fc\_dcq\_IL6):



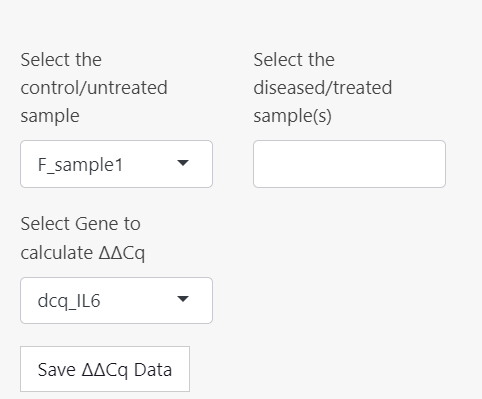
1. To save the relative fold change data, press the download processed data button:



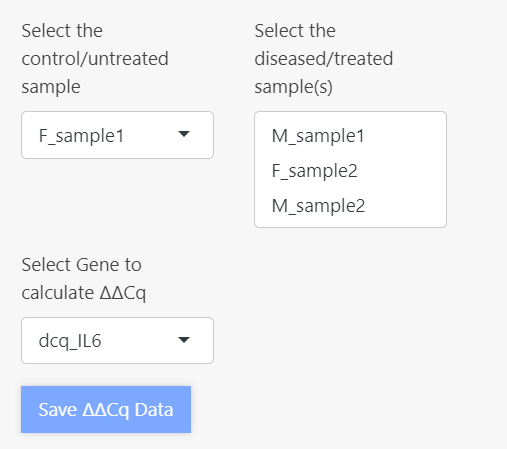
1. To save biological replicate average values from the 2^(- ΔCq) method, select the Biological Replicate tab and press the ‘Download Replicate Average Data’ Button.



1. For calculations using the 2^(- ΔΔCq) method, proceed to the 2- ΔΔCq tab.
2. Select the sample/group to be used as the control/reference sample. E.g. F\_sample1:

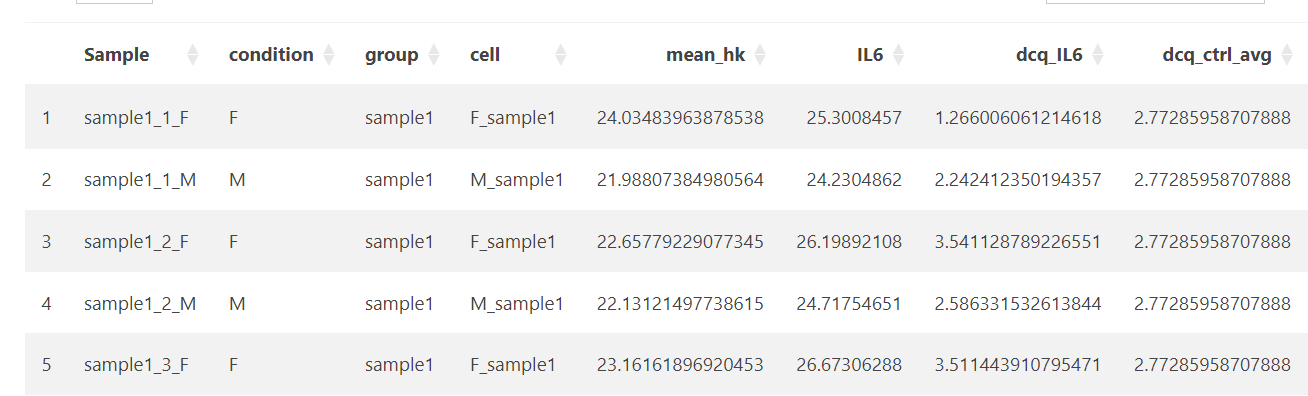


1. Select the remaining groups, and the gene/target to perform the calculation on:



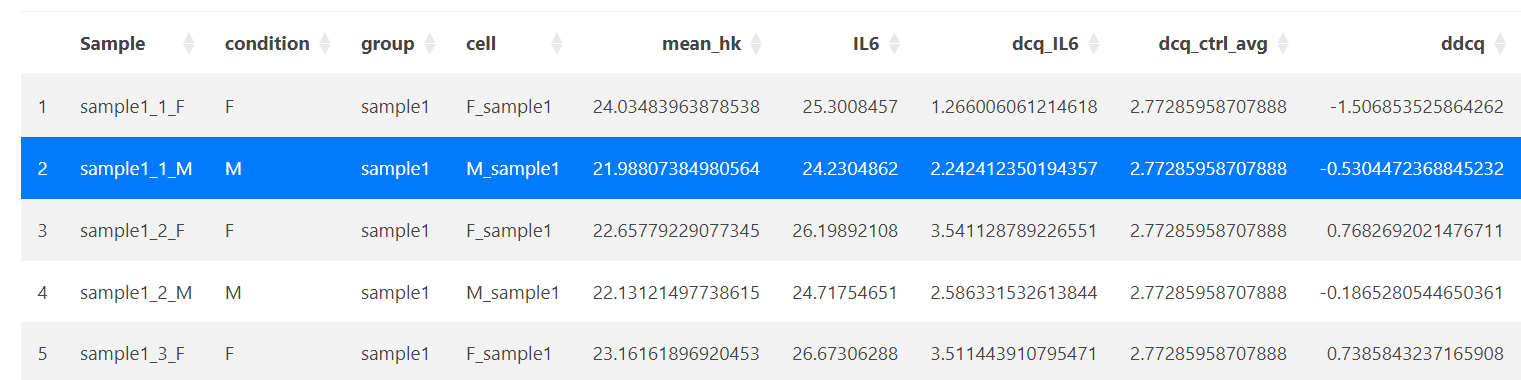
1. Save the 2^(- ΔΔCq) data set within the ProntoPCR environment by pressing the ‘Save ΔΔCq Data’ button. It will appear light blue if it has been saved. This is important to do if performing statistics and/or graphing.
2. 2^(- ΔΔCq) is calculated as follows:

Average the ΔCq values of the control group. This will appear as dcq\_ctrl\_avg, and the same number will be displayed down the column of the table. E.g. the dcq\_ctrl\_avg for the control sample (F\_sample1) is calculated by:

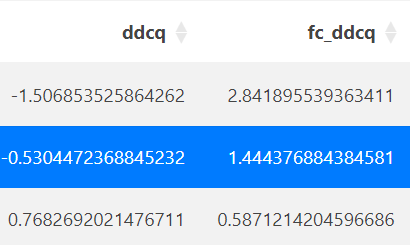


Then perform ΔΔCq (ddcq) by calculating the difference between the ΔCq of your gene of interest and the average of the control group (dcq\_ctrl\_avg):

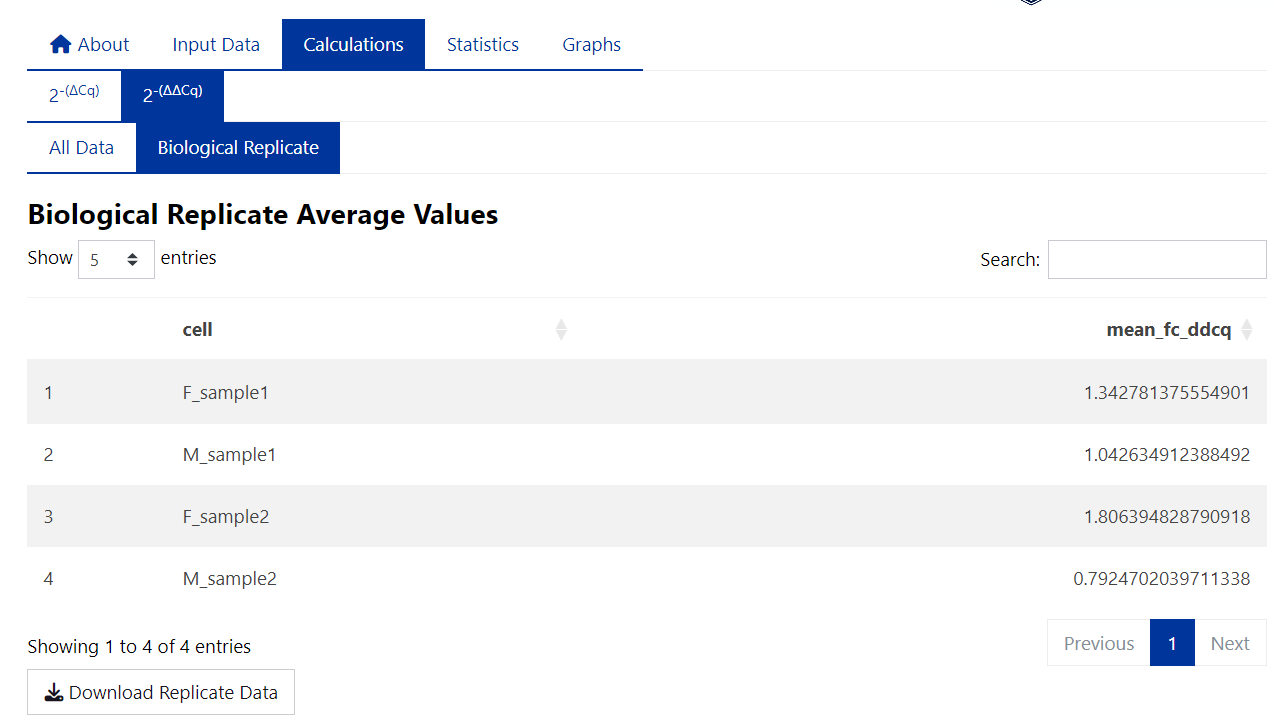
E.g. the ΔΔCq for IL6 and sample1\_1\_M is:



Therefore the fold change (fc\_ddcq) 2^(- ΔΔCq):



1. Press the ‘Download Processed Data’ Button to save the fold change data.
2. The biological replicate average data is displayed in the ‘Biological Replicate’ tab and can be saved using the ‘Download Replicate Data’ button.



### **Statistics**

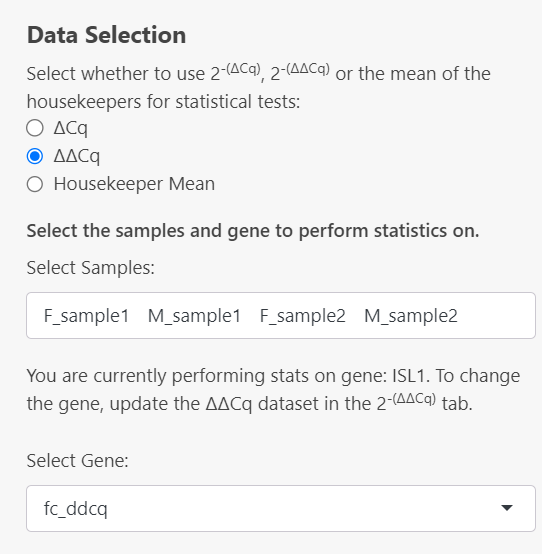
Note: There is no one-size-fits-all approach to statistical analysis. The statistical tests given within ProntoPCR are not exhaustive. It may be more appropriate to use another statistical program in conjunction with ProntoPCR.The p-value significance is denoted as follows, with the default α = 0.05: \*\*\* for *p*-values ≤ 0.001, \*\* for *p*-values ≤ 0.01, \* for *p*-values ≤ 0.05, and ‘ns’ (not significant) for *p*-values > 0.05.

#### **Selecting the Data**

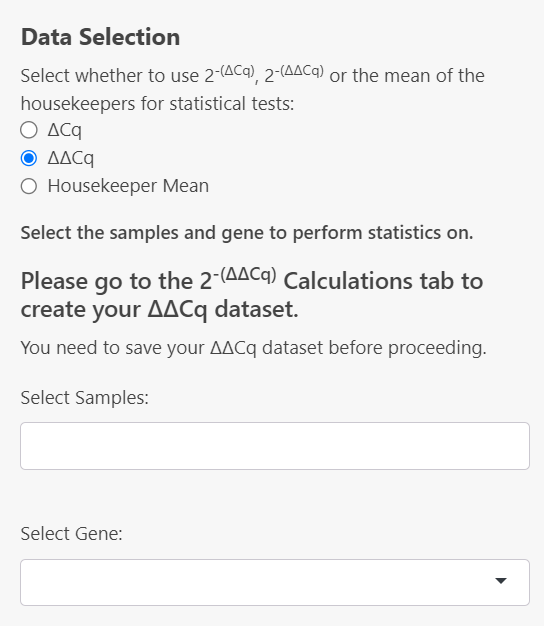
Statistics can be performed on the 2^(- ΔCq), 2^(- ΔΔCq) or housekeeper gene mean values.

1. Select the data to be used for statistics:
   1. If 2^(- ΔCq) is selected, choose the gene and samples to perform statistics on:

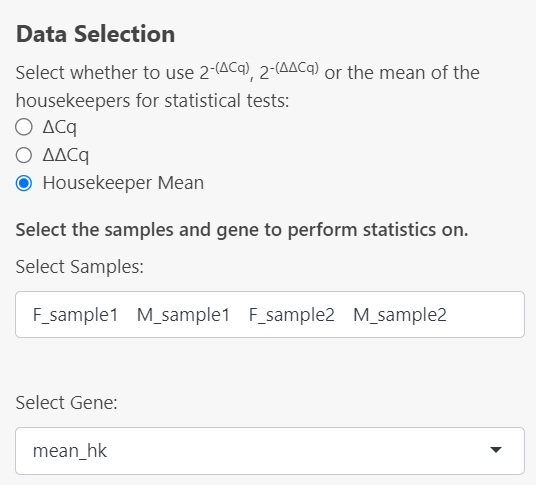


* 1. If 2^(- ΔΔCq) is selected, the gene will automatically be inputted based on the selected options in the calculations 🡪 2^(- ΔΔCq) tab. Select your samples in the dropdown menu. If you don’t see a sample in this section, it must be added to the dataset created in the Calculations 🡪 2^(- ΔΔCq) tab: 

If the 2^(- ΔΔCq) dataset has not been created yet, this error message will be displayed:



* 1. If Housekeeper mean is selected, the mean\_hk column will be used as the ‘Gene’ to perform statistics on:

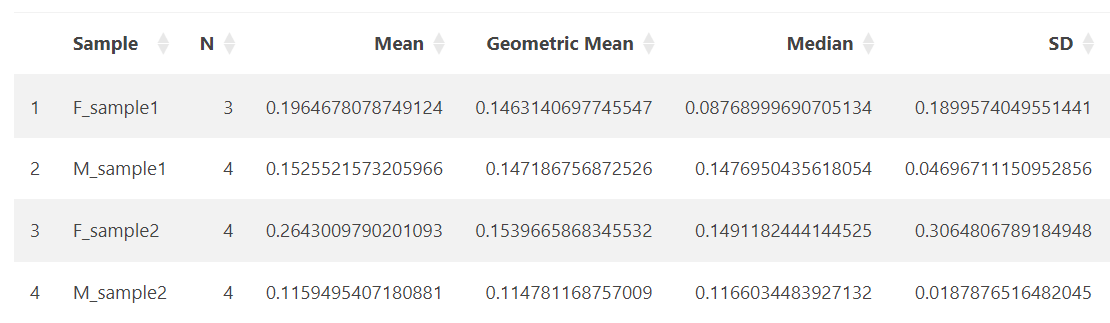


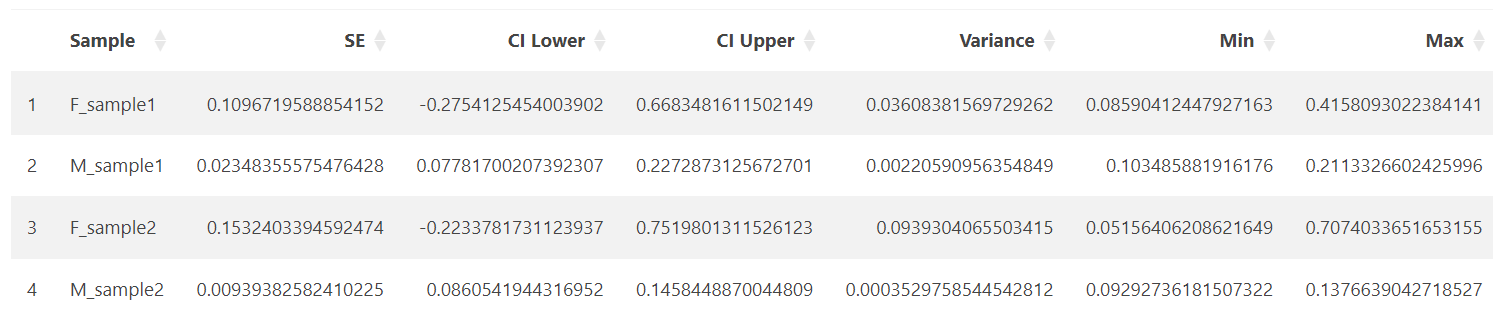
#### **Descriptive Statistics**

From this point onwards, all output given has been performed on the 2^(- ΔCq) dataset for the gene *IL6* from the example data using the geometric mean for averaging the housekeepers.

1. Descriptive Statistics can be computed by selecting the checkboxes in the side panel. Descriptive statistics available include:
   1. Sample size:

* Calculated by counting the number of valid (non-missing) data points within each specified group for that gene.
  1. Mean:
* The mean values are calculated for each specified group by averaging the valid (non-missing) data points using the arithmetic mean for that gene.
  1. Geometric Mean:
* The geometric means are calculated for each specified group by taking the exponent of the average of the logarithms of the valid (non-missing) data points.
  1. Median:
* The median values are calculated for each specified by finding the middle value of the sorted valid (non-missing) data points within each group.
  1. Standard Deviation:
* The standard deviations (SD) are calculated for each specified group by measuring the spread of the valid (non-missing) data points around the mean within each group.
  1. Standard Error:
* The standard errors (SE) are calculated for each specified group by dividing the standard deviation (SD) of the valid (non-missing) data points by the square root of the number of valid observations in each group (i.e. Sample Size (*n*)).
  1. 95% Confidence Interval of the Mean:
* Confidence intervals (CI) are calculated for each group by first determining the standard error based on the standard deviation and sample size, then applying either the Z-score for large samples (*n* ≥ 30) or the t-score for smaller samples to compute the lower and upper bounds of the interval around the mean. These intervals provide a range where the true population mean is likely to fall with a specified level of confidence (95%).
* For large samples (*n* ≥ 30):
  + - * is the sample mean.
      * is the standard deviation.
      * is the sample size.
      * is the Z-score corresponding to the confidence level (i.e. 1.96 for 95% confidence).
* For smaller samples (*n* < 30):
  + - * is the sample mean.
      * is the standard deviation.
      * is the sample size.
      * is the t-score corresponding to the desired confidence level and degrees of freedom ().
* In both cases, the confidence interval is expressed as:
  1. Variance:
* The variances are calculated for each specified group by measuring the average of the squared differences between each valid (non-missing) data point and the group mean. This provides a measure of how spread out the data points are within each group.
  + - * is the sample variance.
      * ∑ denotes the summation over all observations.
      * represents each individual observation.
      * is the mean of the observations.
      * is the number of observations in the sample.
      * represents the degrees of freedom.
  1. Minimum/Maximum:
* The minimum and maximum values are determined for each group by identifying the smallest and largest valid data point for that gene.



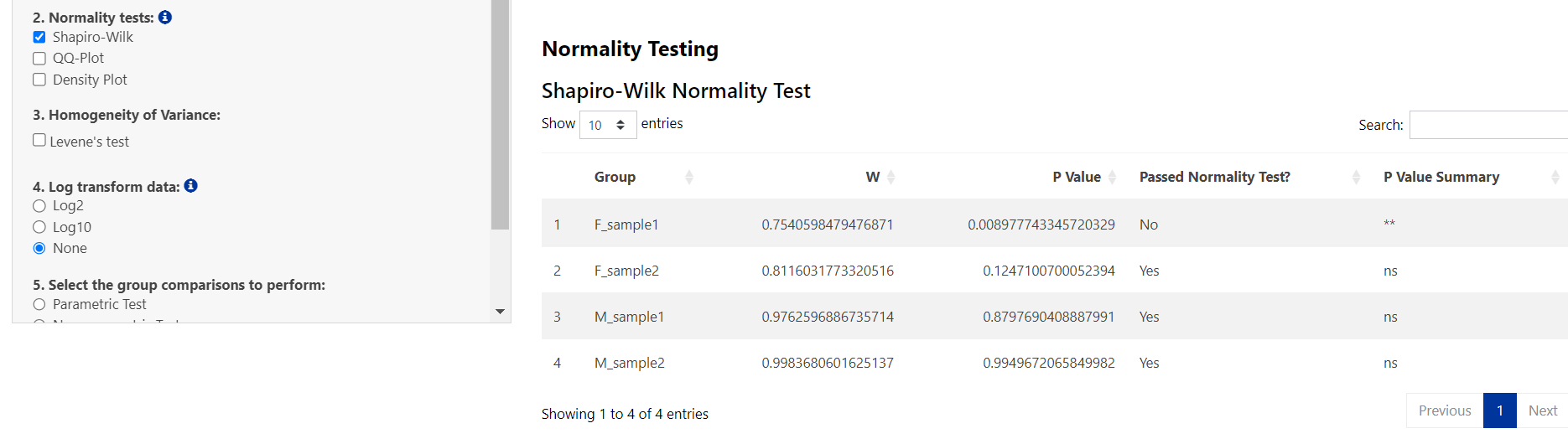


#### **Normality Testing**

The normality of data can be tested using a Shapiro-Wilk test, QQ plots and density plots in ProntoPCR. By default, the normality is tested per group on the raw data. If one-way ANOVA is selected, the normality tests will be performed on the residuals, not the raw data. A Residuals vs Fitted plot will also be displayed if One-way ANOVA and a normality test are both selected. The raw values are the default if no comparison of groups test, or any other group comparison tests are selected. Note that normality testing is usually considered to be unreliable for small sample sizes.

1. A Shapiro-Wilk test can be performed by selecting the checkbox on the side panel. It is computed using `stats::shapiro.test` in R.

* Raw data:



* Residuals:



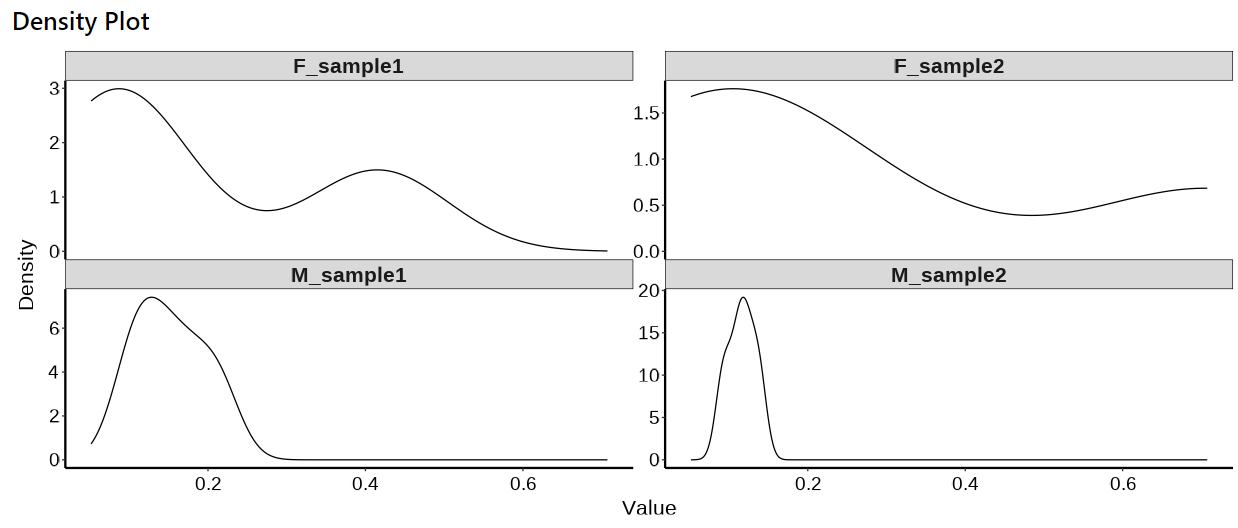
1. A QQ Plot will be generated by selecting the checkbox on the side panel. It is computed using:





1. A density plot will be generated by selecting the ‘Density Plot’ checkbox on the side panel. It is computed using:



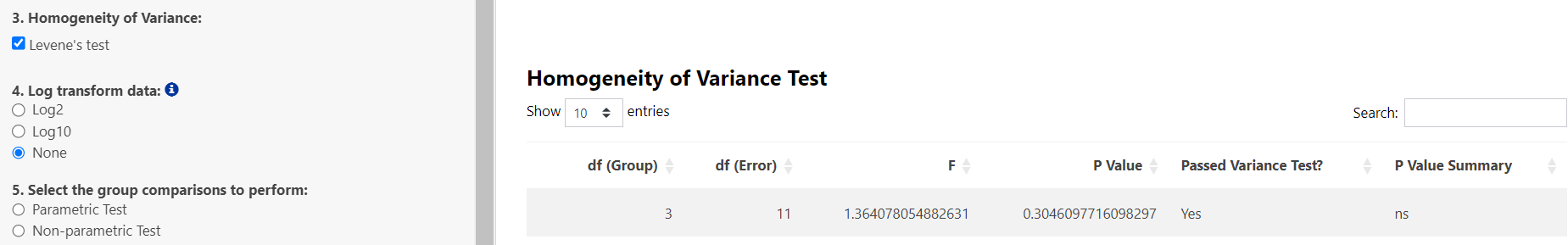


#### **Homogeneity of Variance Testing:**

The homogeneity of variance can be tested using a Levene’s Test.

1. To perform a levene’s test select this option on the side panel. It is computed using:

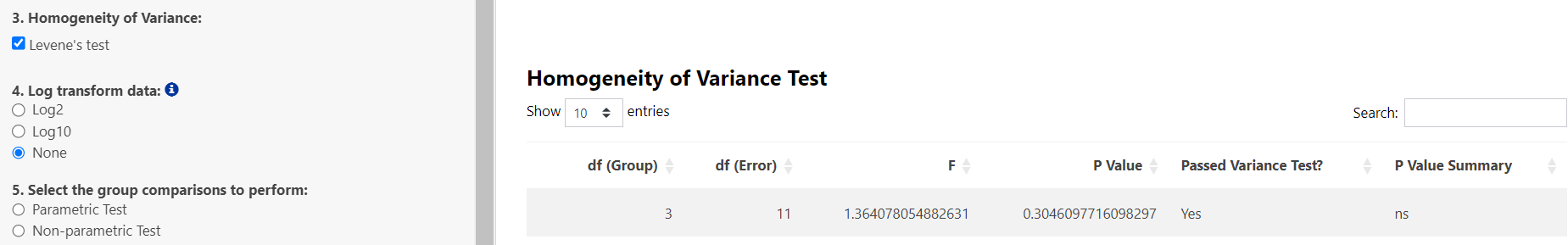




#### **Log Transformation:**

If required, data can be log-transformed using log2() or log10(). For PCR data, it is recommended to use log2() transformations, however, log10() is given as an additional option if there are large differences in expression. If log2 or log10 is selected, all statistics will be performed on the transformed values e.g. log2(2^-ΔΔCq). Log transformation may help with normalizing data and reducing skewness.

1. To perform a log transformation, select Log2 or Log10 in the side panel:



#### **Group Comparison Testing:**

Multiple tests are provided to compare groups. If two samples are selected:

* Independent t-test 🡪 Parametric Test
* Mann Whitney-U test 🡪 Non-parametric Test
* Welch t-test 🡪 Welch’s Test

If more than two samples are selected:

* One-way ANOVA 🡪 Parametric Test

Post-hoc tests:

* + Tukey’s HSD
  + Pairwise t-test with Bonferroni adjustment for multiple comparisons
  + Pairwise t-test with Holm adjustment for multiple comparisons
  + Pairwise t-test with Benjamini-Hochberg adjustment for multiple comparisons
  + Scheffé's post-hoc
* Kruskal Wallis Test 🡪 Non-parametric Test

Post-hoc test:

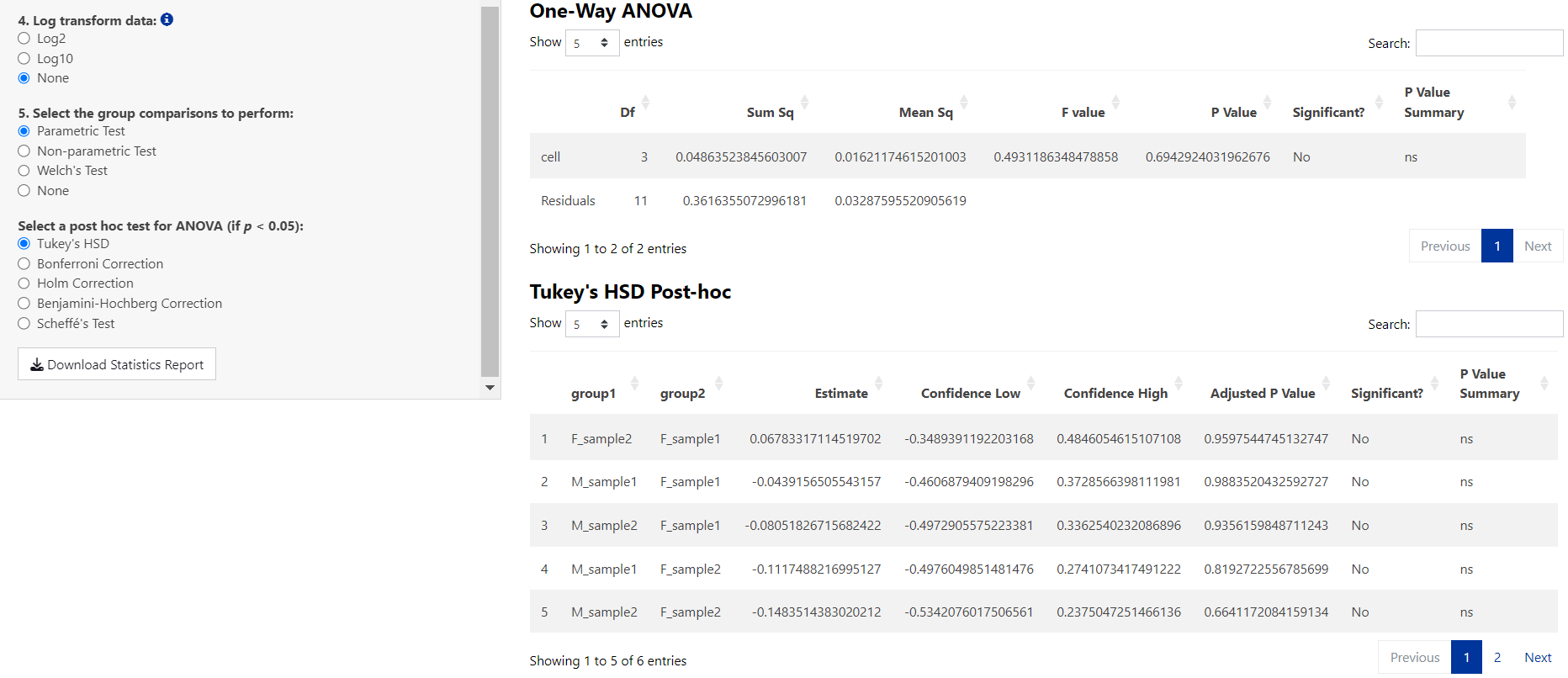
* + Dunn’s test with:
    - Bonferroni adjustment for multiple comparisons
    - Šidákadjustment for multiple comparisons
    - Holm adjustment for multiple comparisons
    - Holm-Šidák adjustment for multiple comparisons
    - Benjamini-Hochberg adjustment for multiple comparisons
    - Hochberg’s Step-Up adjustment for multiple comparisons
* Welch ANOVA 🡪 Welch’s Test

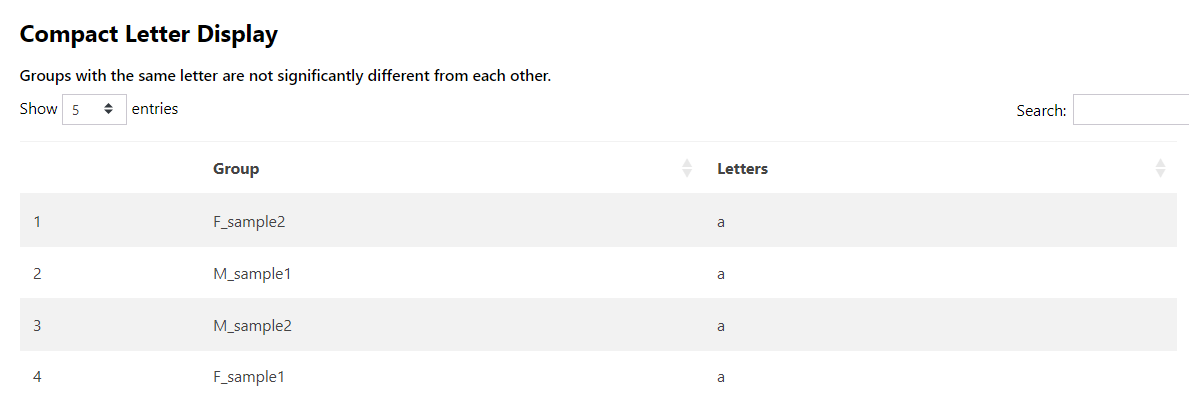
Post-hoc test:

* + Games-Howell post-hoc

Note: Post-hoc test results will be performed and displayed regardless of the group comparison p-value. This decision was made to ensure that all comparisons are available for review, as some users may wish to explore trends or patterns even in the absence of statistically significant differences. However, it is only appropriate to interpret the post-hoc results if the group comparison p-value is less than 0.05, as this threshold indicates a statistically significant difference between the groups.

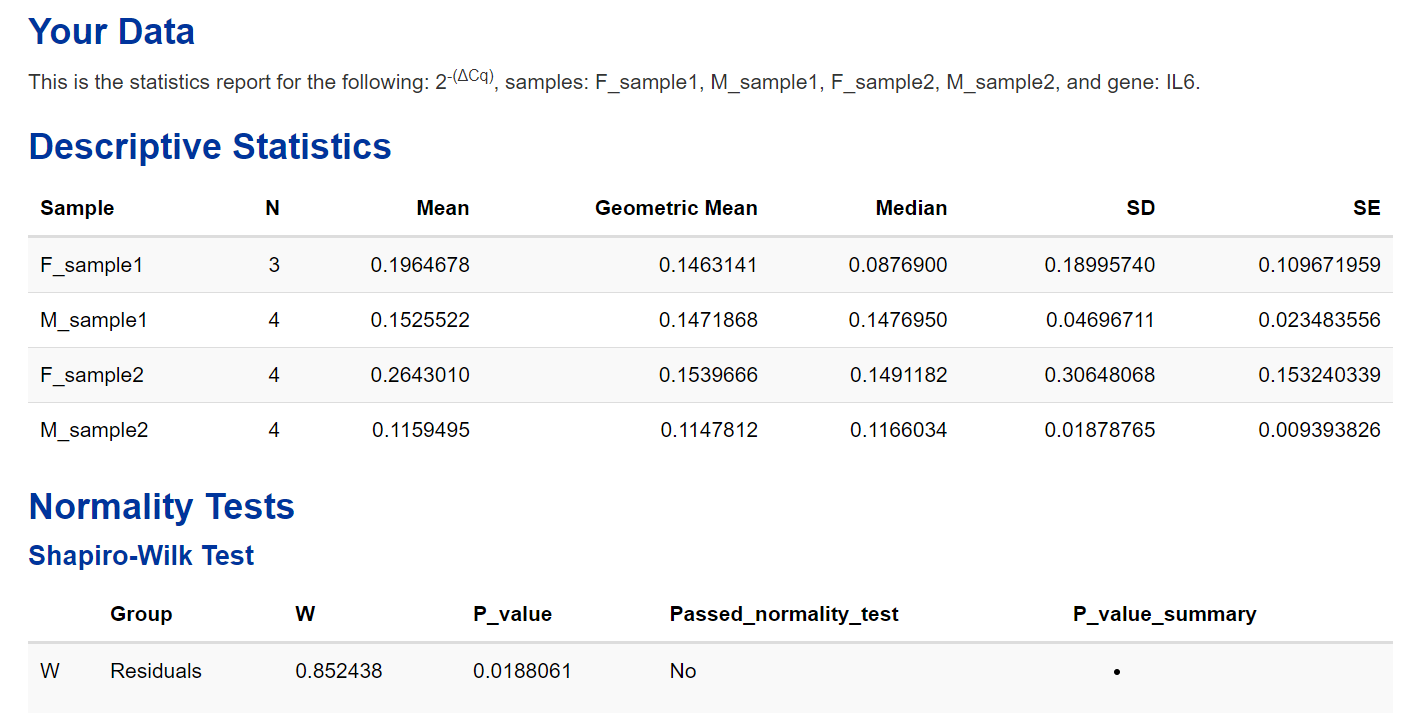
1. To perform group comparisons, select an option from the side panel:



1. If a group comparison test is selected, a compact letter display results table will be provided. Samples sharing the same letter are not significantly different from one another, based on the selected post-hoc test results.

#### **Saving Results:**

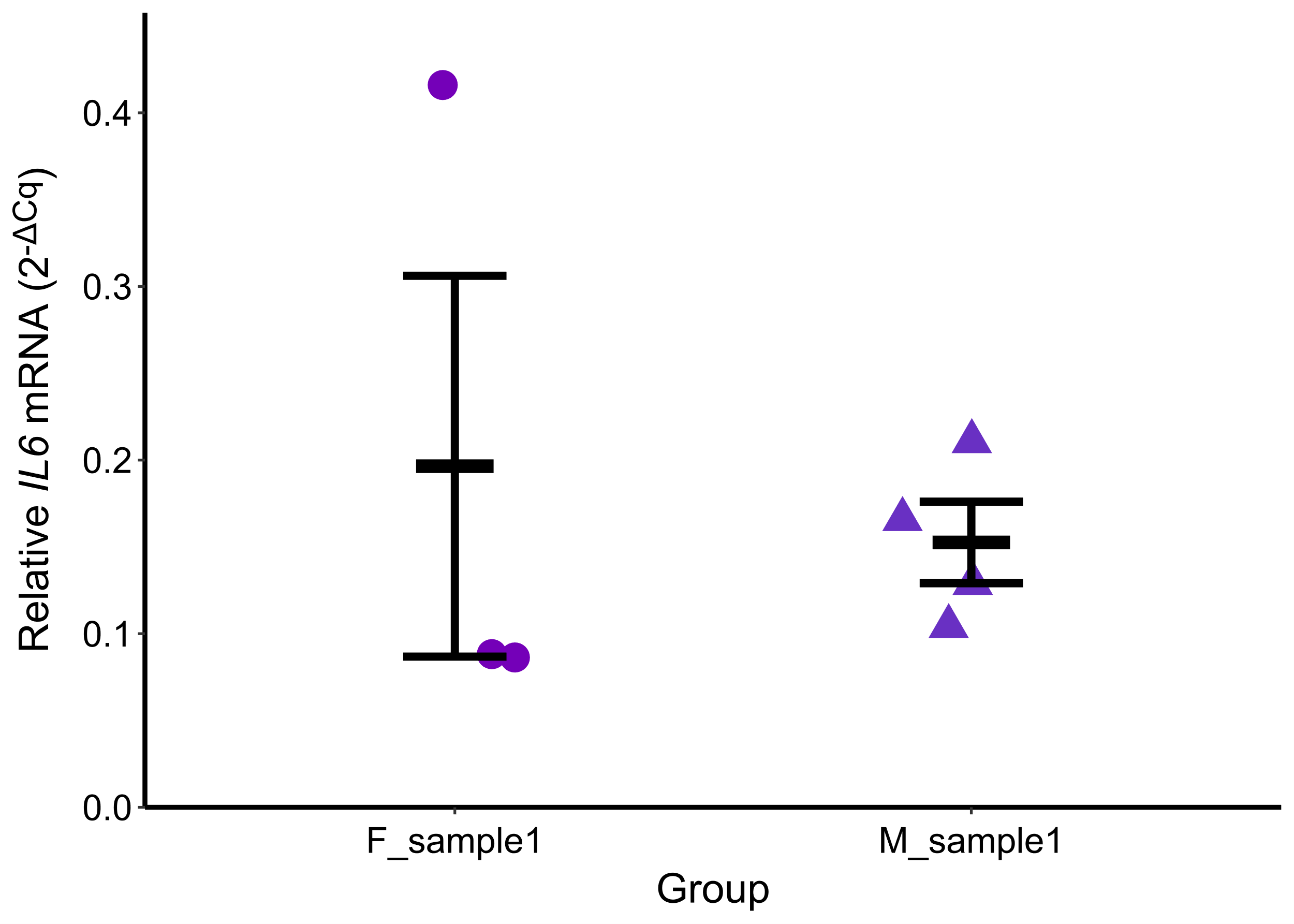
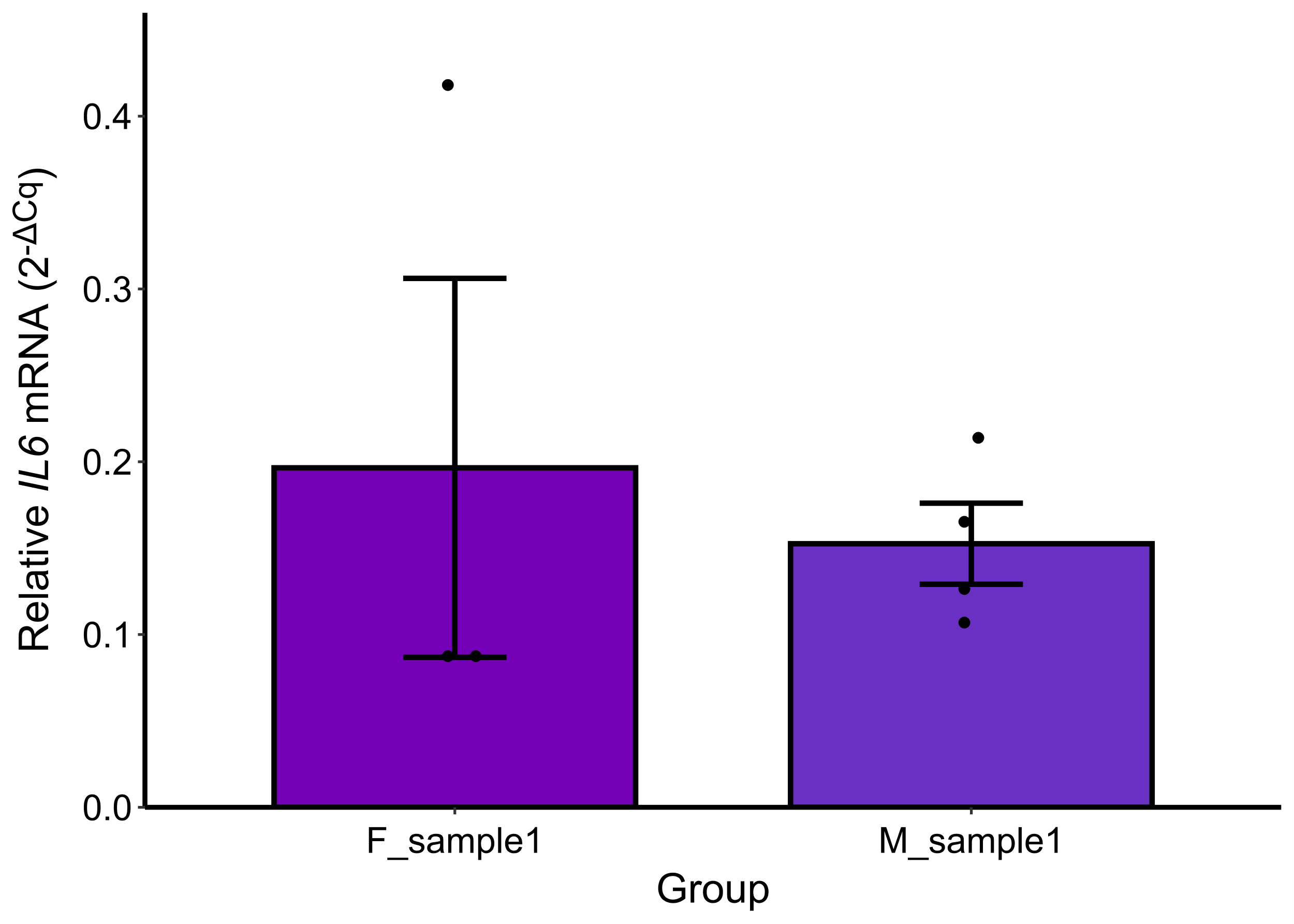
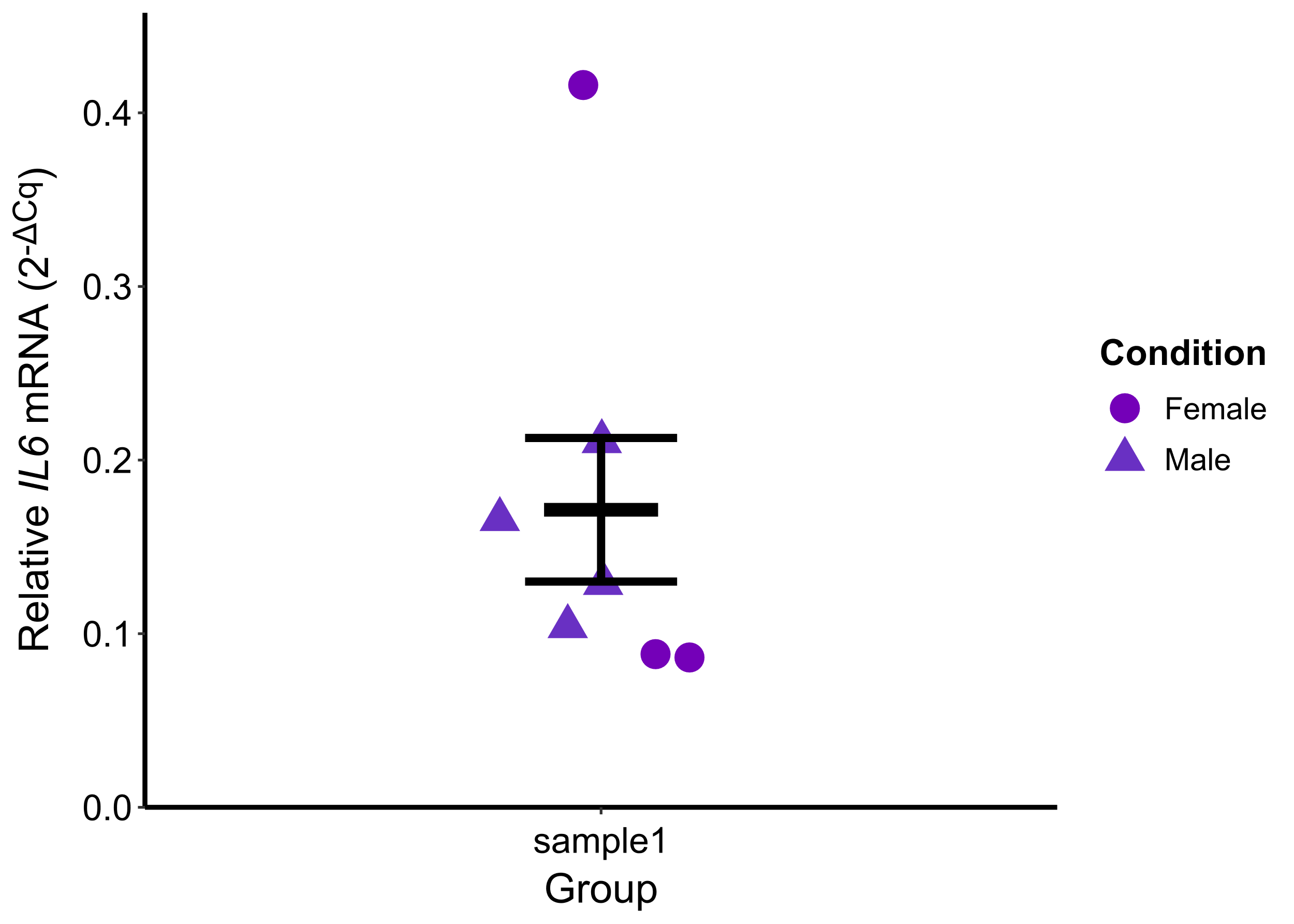
All results can be saved as an HTML file by selecting the Download Statistics Report Button on the side panel. This can be opened in any web browser for viewing.



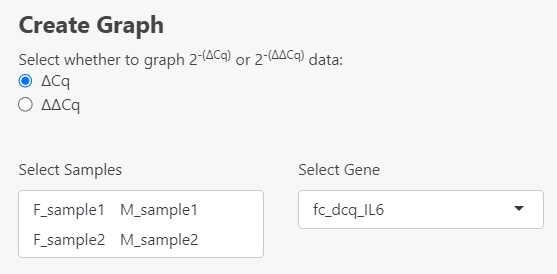
### **Graphs**

Graphs can be generated within the Graphs Tab. They are automatically generated after selecting the Samples and gene to graph. Each graph is fully customisable.

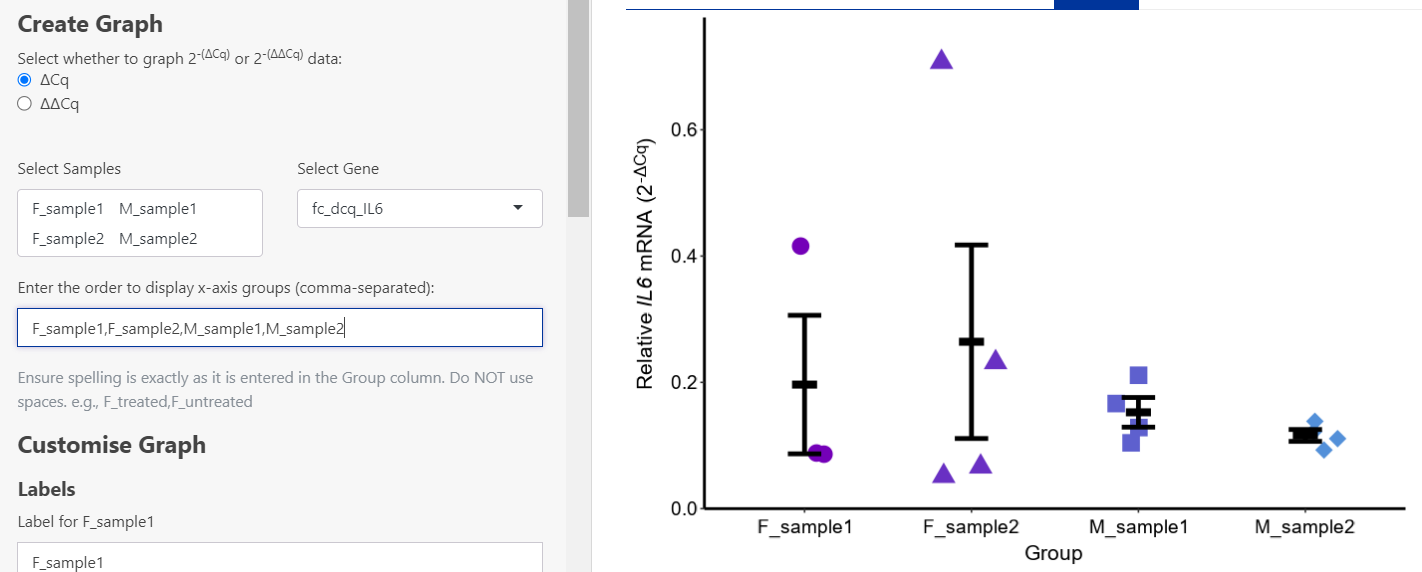
Currently, there are three types of graphs available, scatter plot, grouped scatter plot and column graph:



1. To get started, select whether to graph (2^-ΔCq) or (2^-ΔΔCq) data in the side panel, and select the samples and gene:



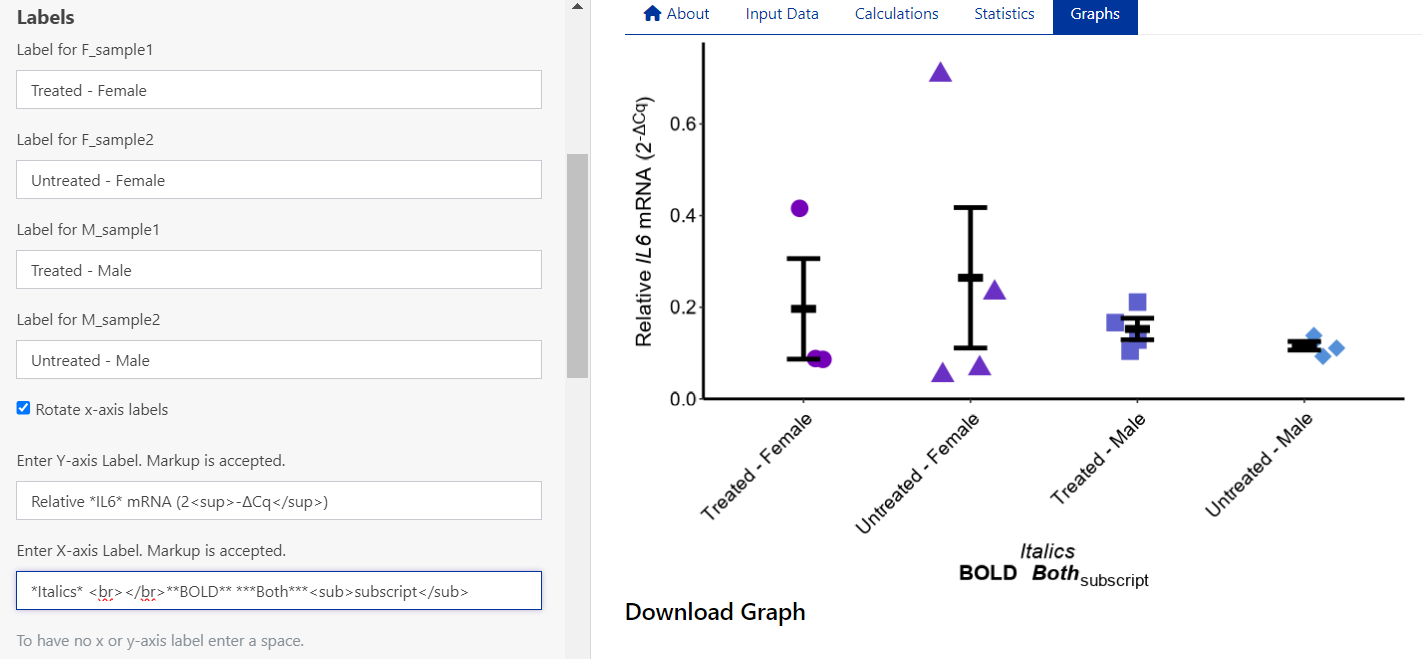
1. To generate the graph, type the sample names in the order to be displayed on the x-axis. The samples must be separated by a comma with no spaces. It is case-sensitive.



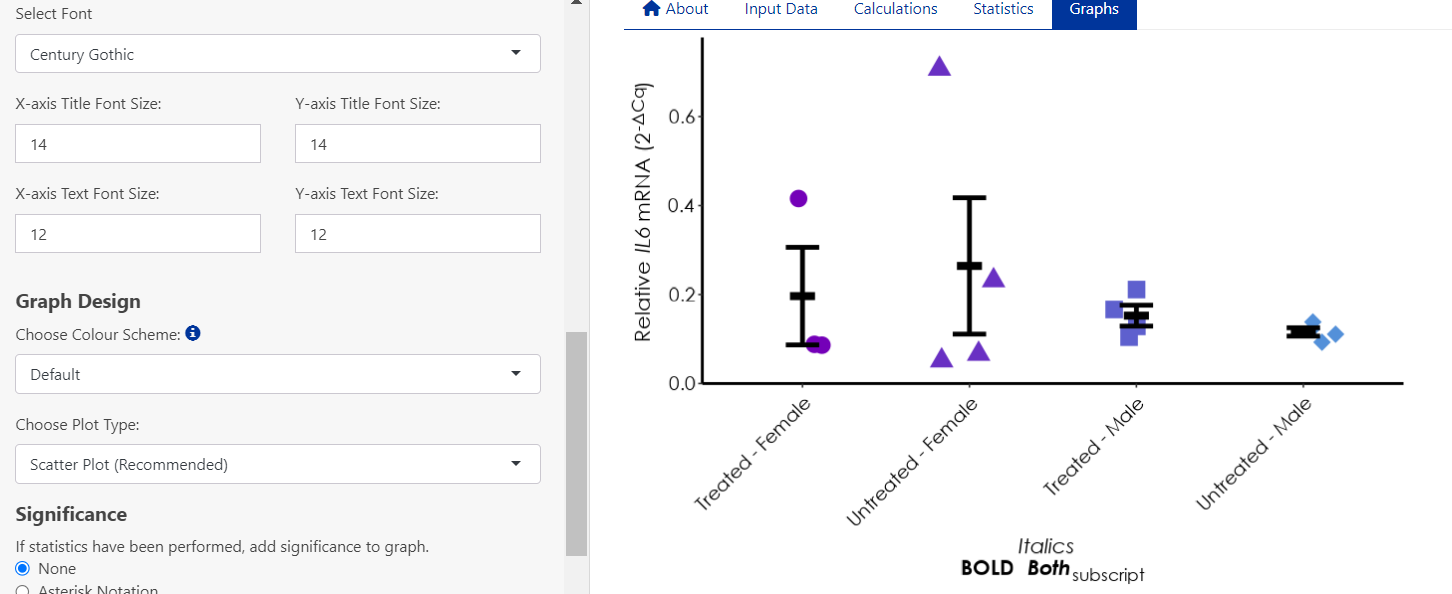
#### **Customising Labels:**

1. Labels for the x and y axes can be changed in the side panel. To remove a label, enter a space in the text input box. The x and y-axis titles accept markup.

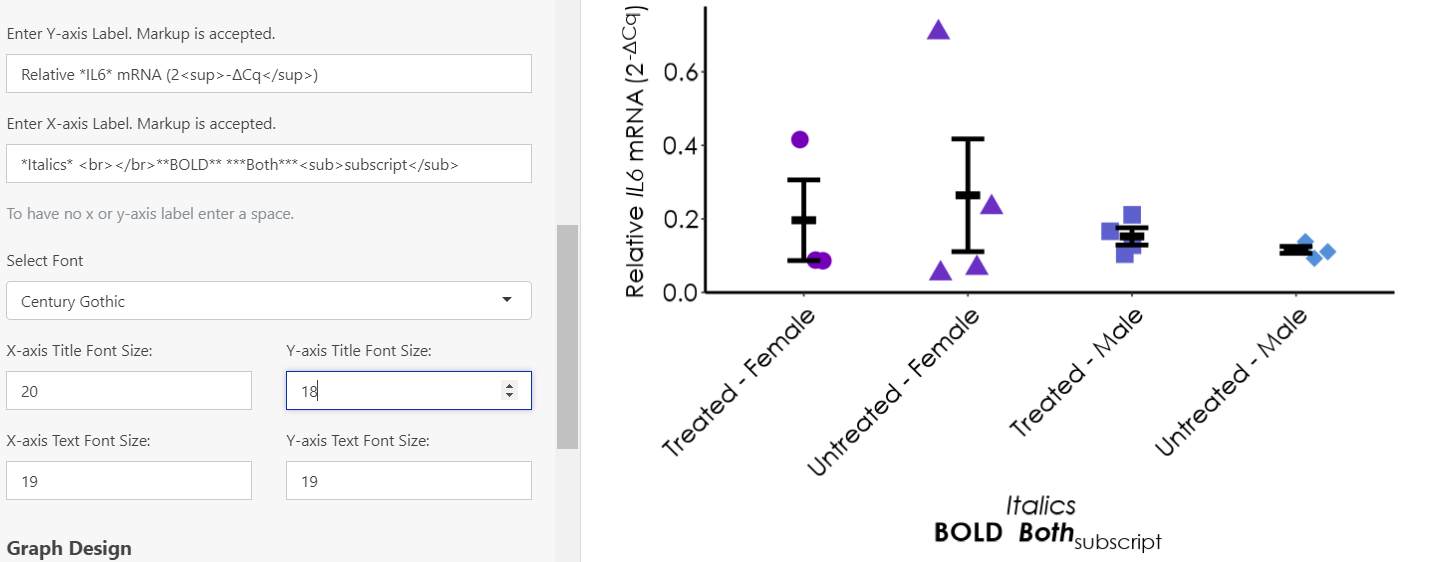
* \*italics\*
* \*\*bold\*\*
* \*\*\*bold and italics\*\*\*
* <br> </br> Line Break
* <sup> superscript </sup>
* <sub> subscript </sub>



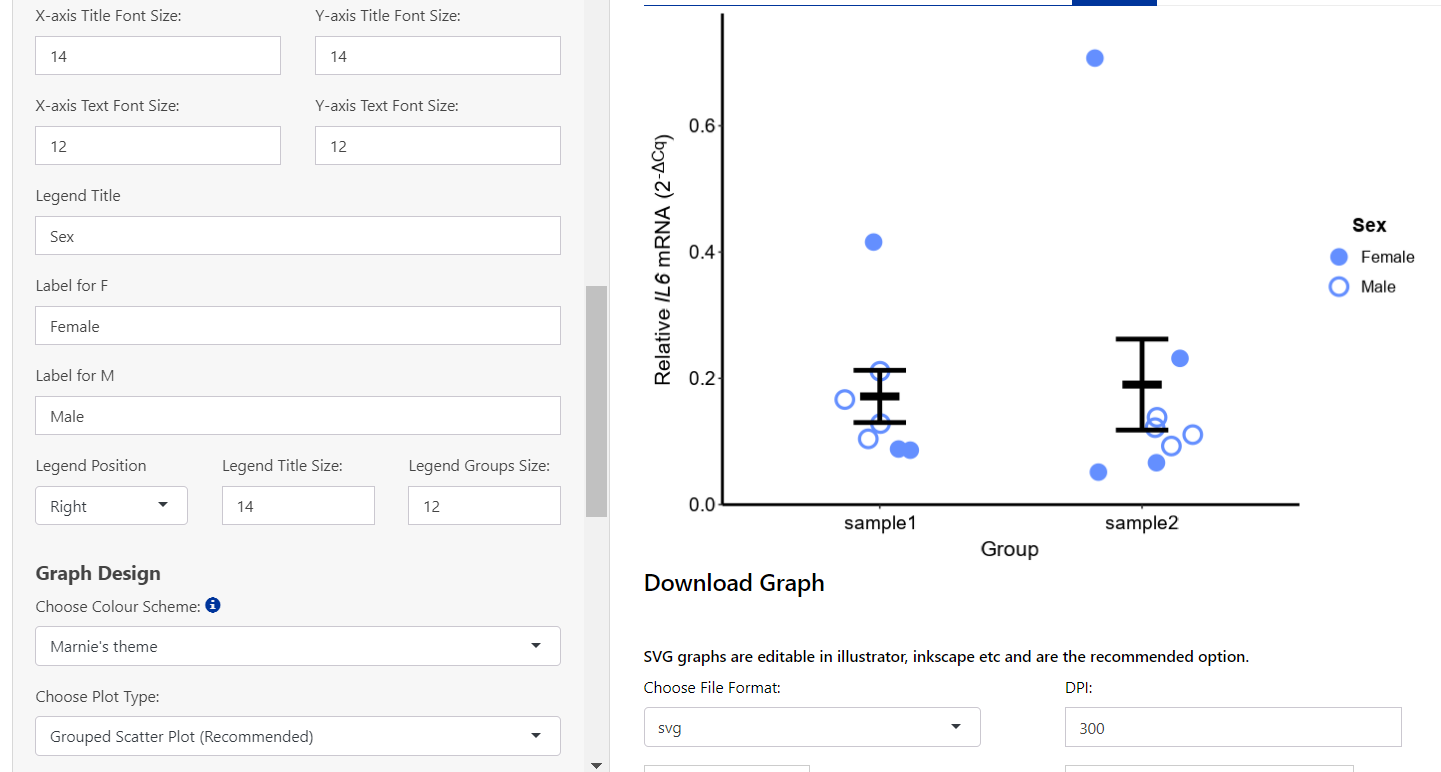
1. The x-axis labels can be rotated by selecting the checkbox.
2. The font can be changed in the dropdown menu. If your desired font is not listed, download the graph as a scalable vector graphics file (svg) and edit it in Inkscape or Illustrator. Alternatively, suggest new fonts on [GitHub Issues](https://github.com/MarnieMaddock/ProntoPCR/issues).



1. Text size can be changed by pressing the up or down arrow, or by typing in the required size in the respective box on the side panel:



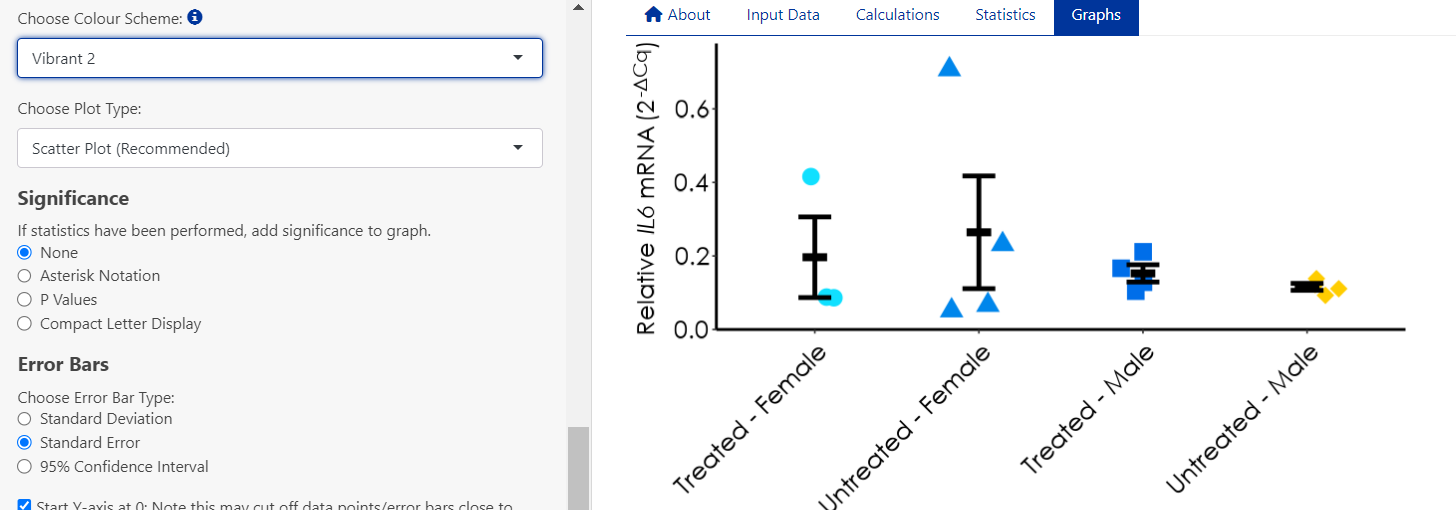
1. If Grouped Scatter Plot is selected, the legend can be customised or removed:



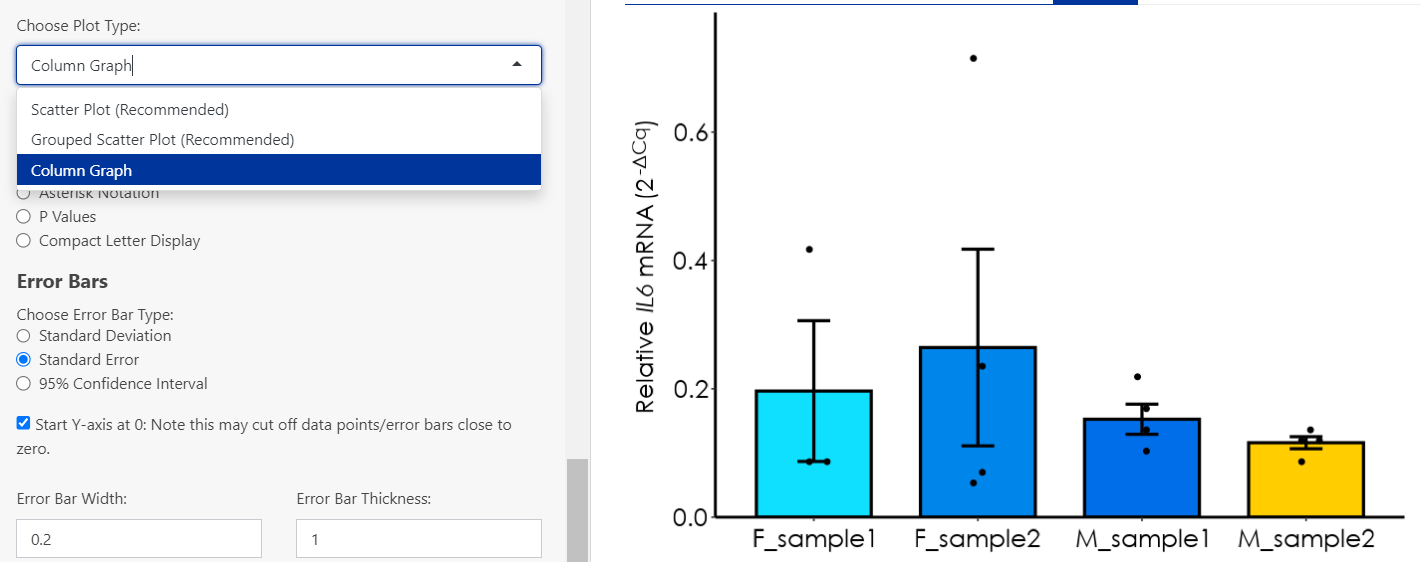
#### **Graph Design:**

1. To change the colours of the graph, select from the drop-down menu. An error message will be displayed if there are too many groups for the selected palette. Colour palettes available:



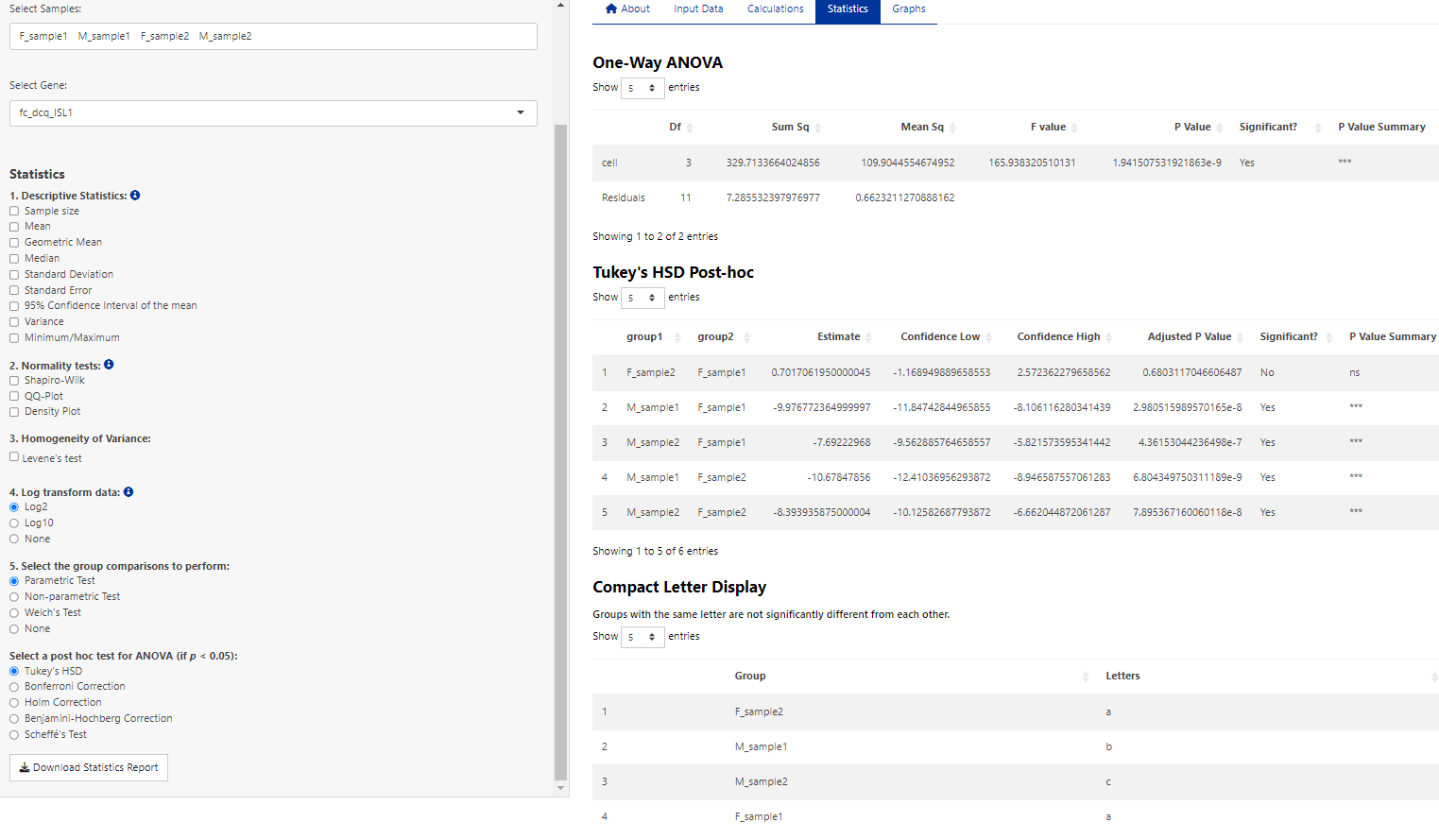


1. To change the graph type, select the preferred option in the Choose Plot Type menu:

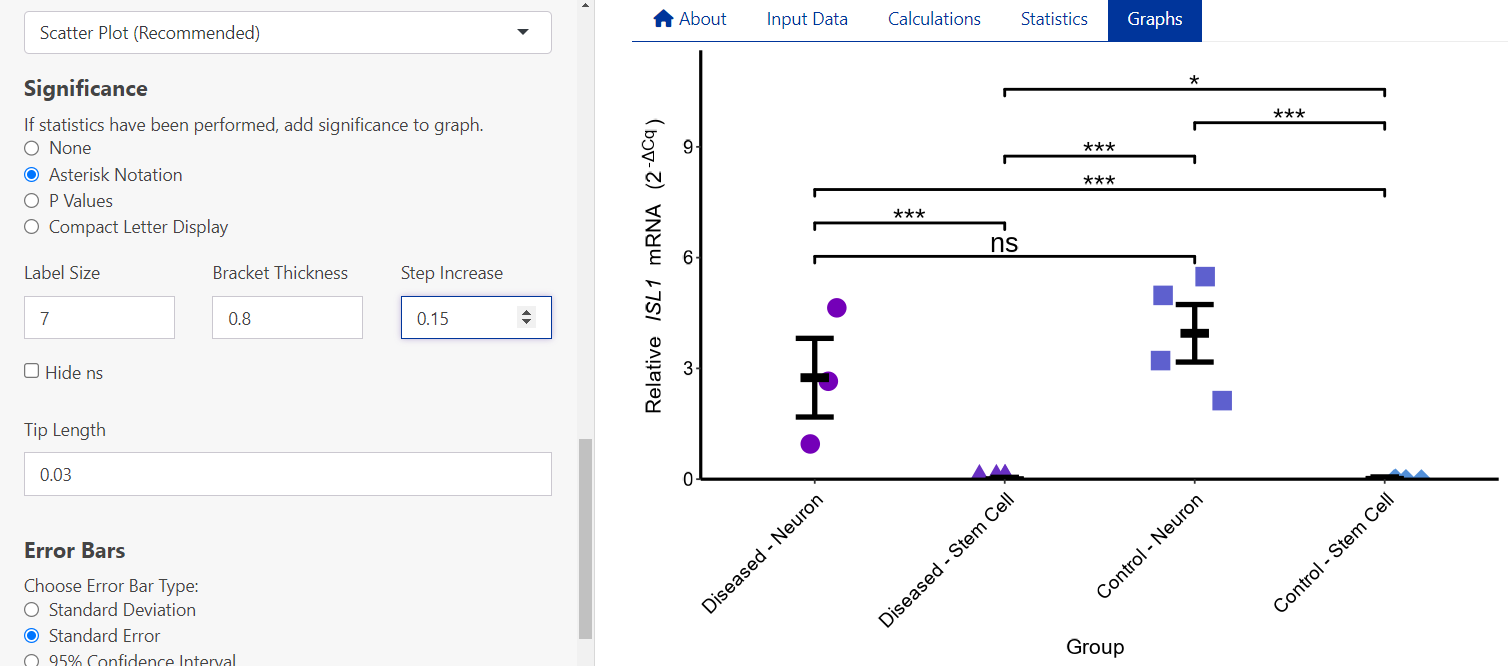


#### **Significance:**

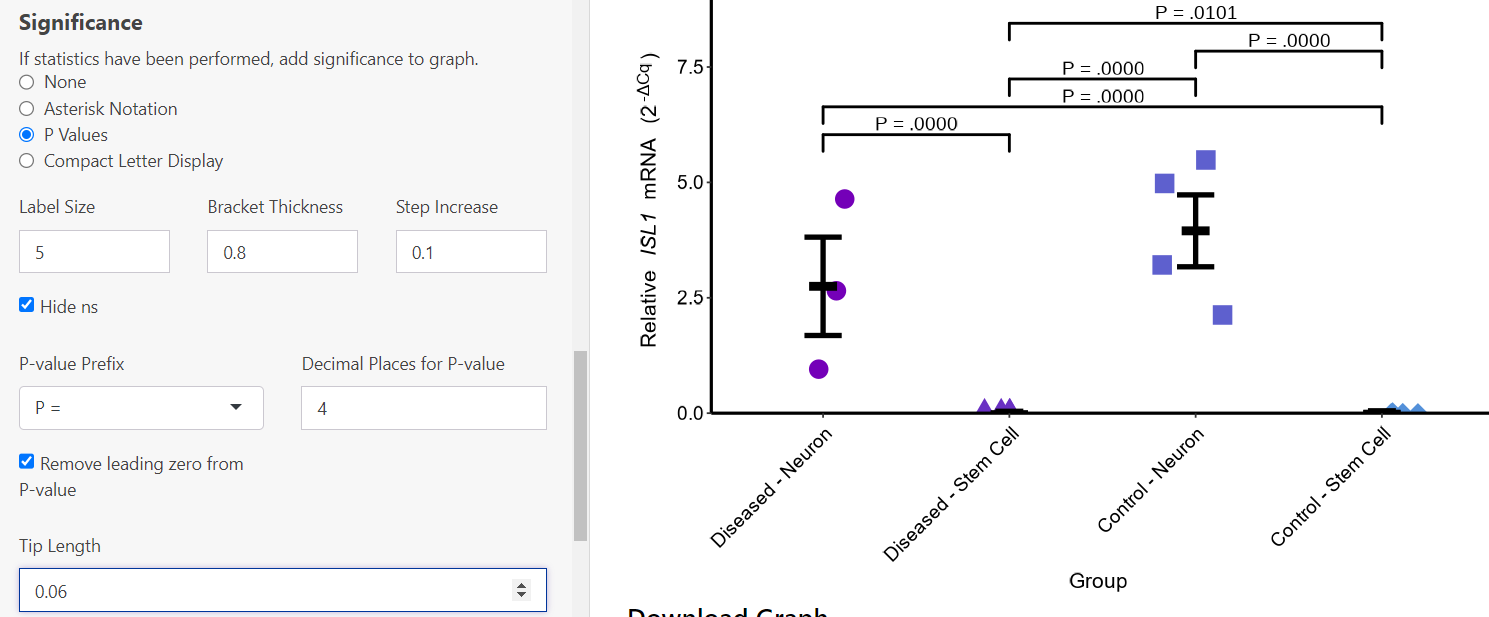
A group comparison test must be performed in the Statistics Tab to add significance options to a graph. The results from the post-hoc test will be used on the graph. The chosen calculation (i.e. 2-(ΔCq) or 2-(ΔΔCq)), samples and gene must be identical in the Statistics and Graphs tab for the correct p-values to be plotted. Error messages (Discrepancy detected) will appear if the options do not match, and will prompt the user to alter the selection. If the graph is empty, it may mean that a discrepancy was detected and needs to be changed for the graph to generate. For example, to add significance to a graph for the samples: F\_sample1, M\_sample1, F\_sample2, M\_sample2 and gene *ISL1* with a log2 transformation:



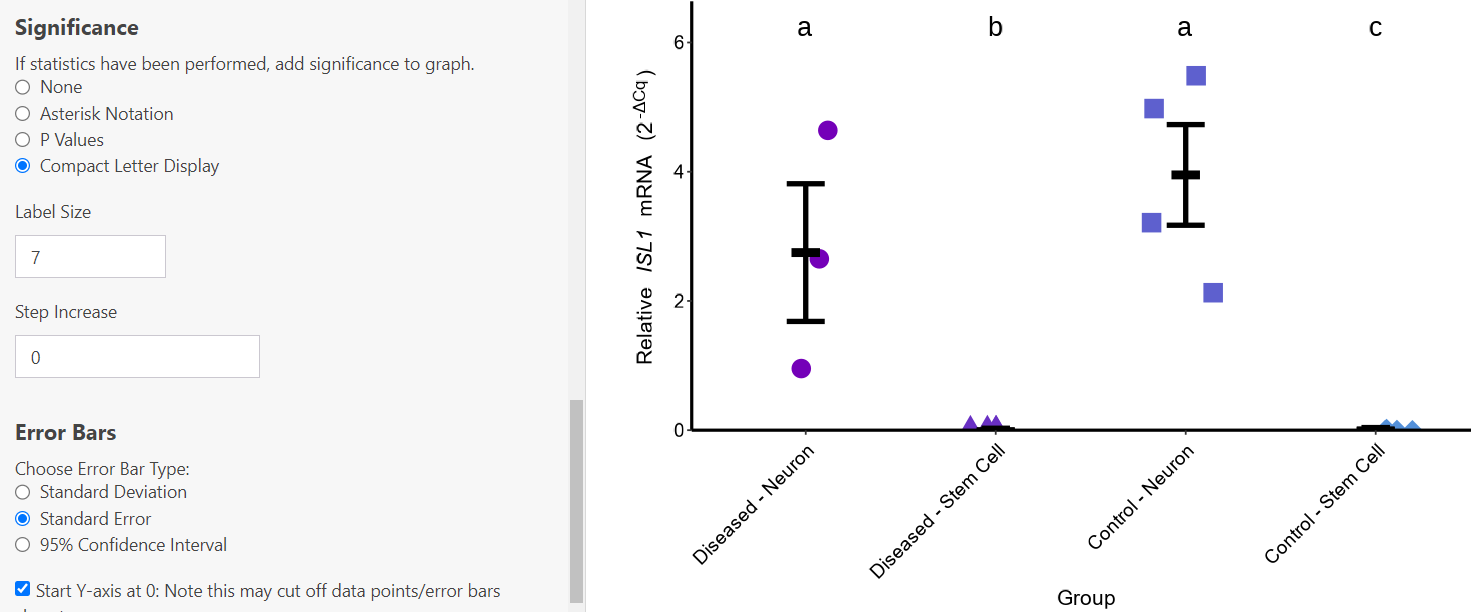
1. Three ways to denote significance on a graph are available, including asterisk notation (\*, \*\*, \*\*\*, ns), p-values (0.001), or compact letter display (a, b, c, ac).
   1. Asterisk Option: The label size (size of \*), bracket thickness (line thickness), step increase (the vertical distance between each bracket) and tip length (vertical bracket length) can all be altered by changing the up and down arrows, or by entering in a number value. Non-significant comparisons can be hidden by pressing the Hide ns checkbox.



* 1. P Values Option: The label size (size of 0.002), bracket thickness (line thickness), step increase (the vertical distance between each bracket), tip length (vertical bracket length) and p-value decimal places can all be altered by changing the up and down arrows, or by entering in a number value. Non-significant comparisons can be hidden by pressing the Hide ns checkbox. The leading 0 from the p-value can also be removed by pressing the corresponding checkbox. P = can also be added before the p-value by selecting this option in the P-value prefix selection.

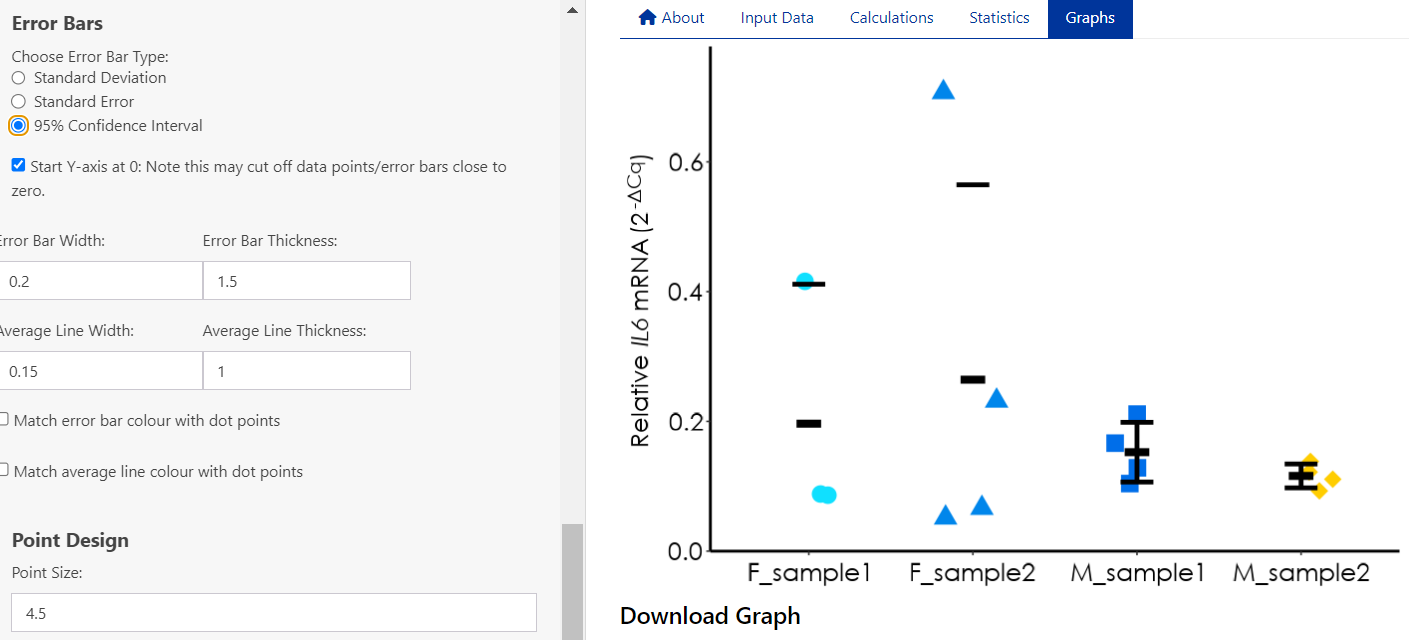


* 1. Compact Letter Display: Letters denoting differences between groups come from the output given in the Statistics tab. The Label Size and Step Increase can be changed by using the up and down errors within the box, or entering a number.

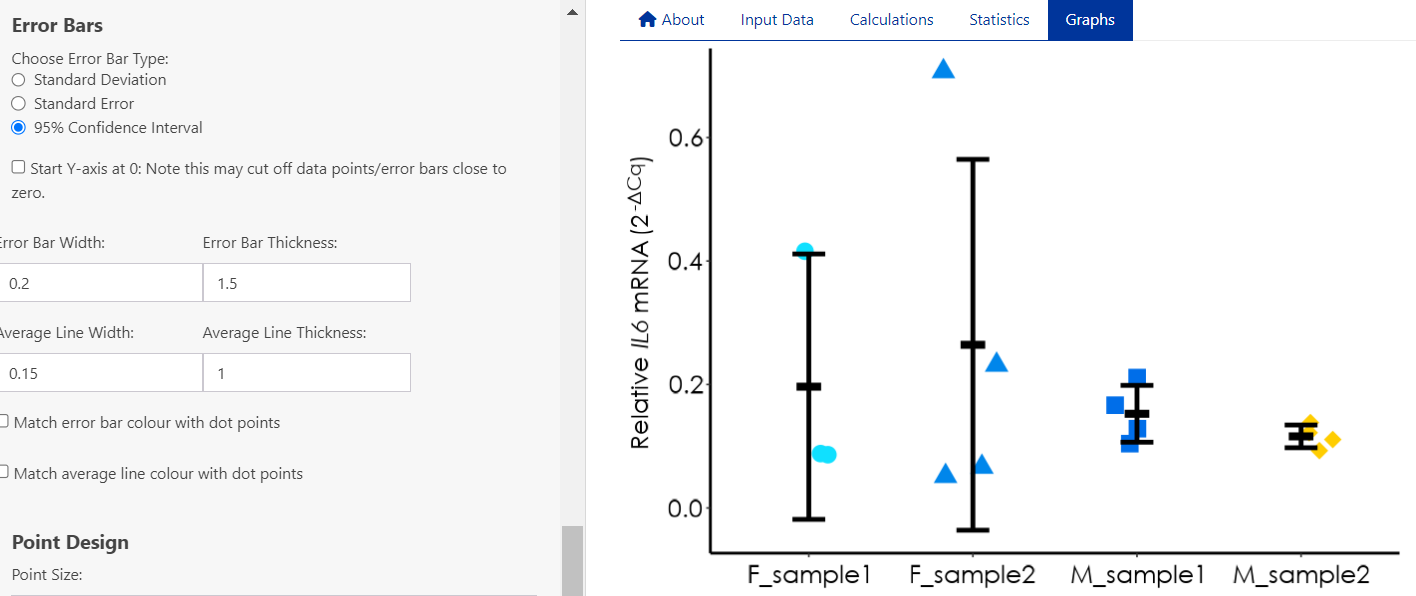


#### **Error Bars:**

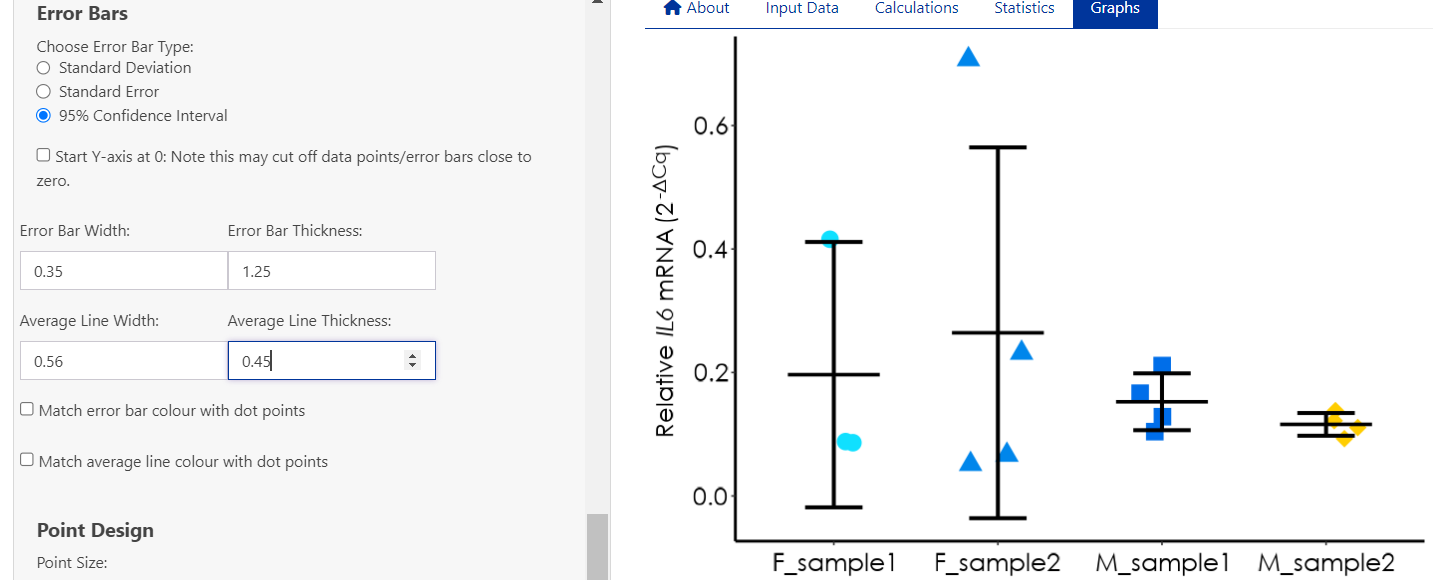
1. Three types of error bars can be selected in the Choose Error Bar section, including: Standard error, standard deviation and 95% Confidence Interval:



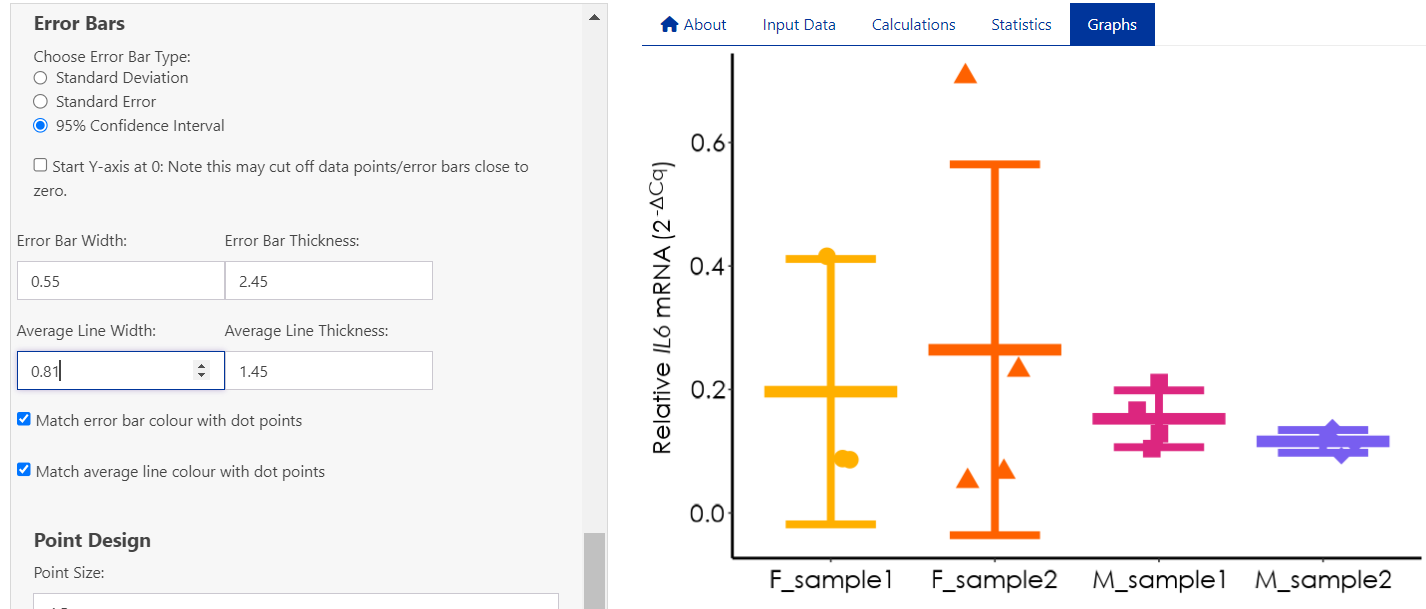
1. Often, error bars are missing if the Start Y-axis at 0 checkbox is selected. Uncheck this box to display error bars correctly:



1. Error bar and average bar width and thickness can be changed using the up and down arrows within the box, or by entering a number:



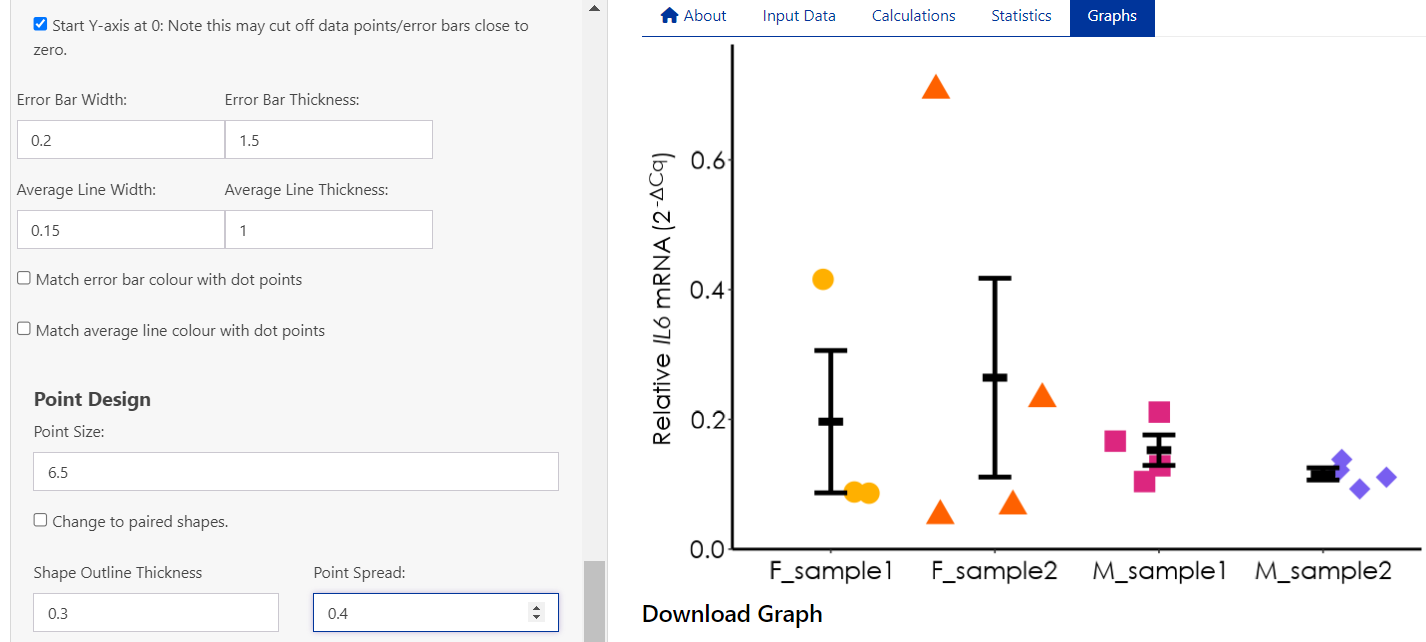
1. Error bar and average line colours can be matched to the dot points by selecting the corresponding checkboxes:



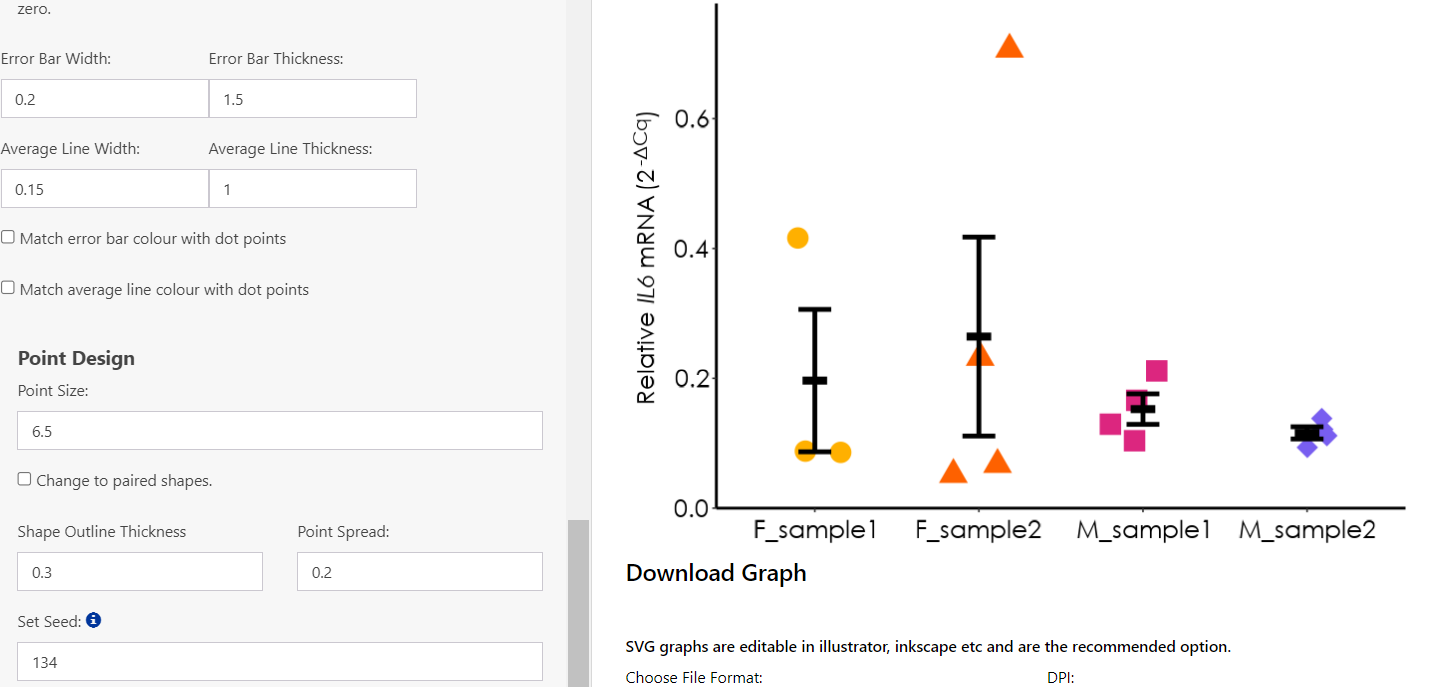
#### **Point Design:**

For scatter plots and grouped scatter plots:

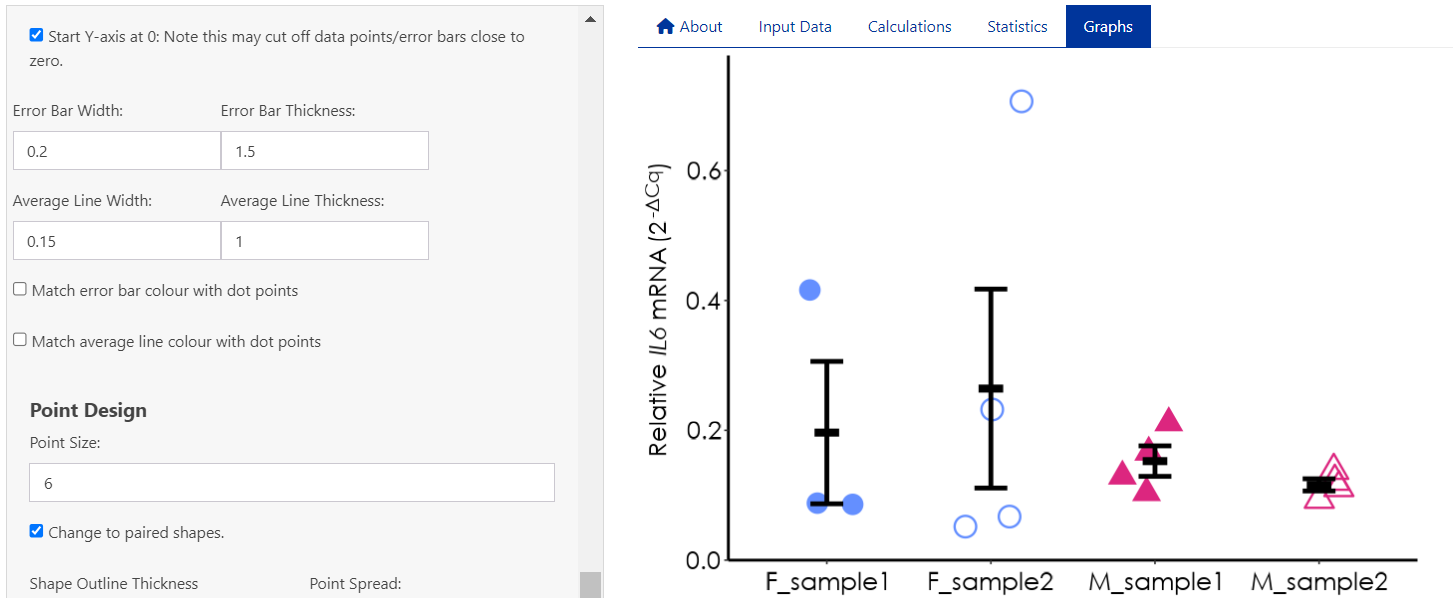
1. Data points on the graph can be customised by size, thickness and point spread by changing the number on the corresponding boxes:



1. If data points are overlapping the seed number can be changed. This is a random and arbitrary number. The same seed number will keep the points in the same position each time the graph generates:



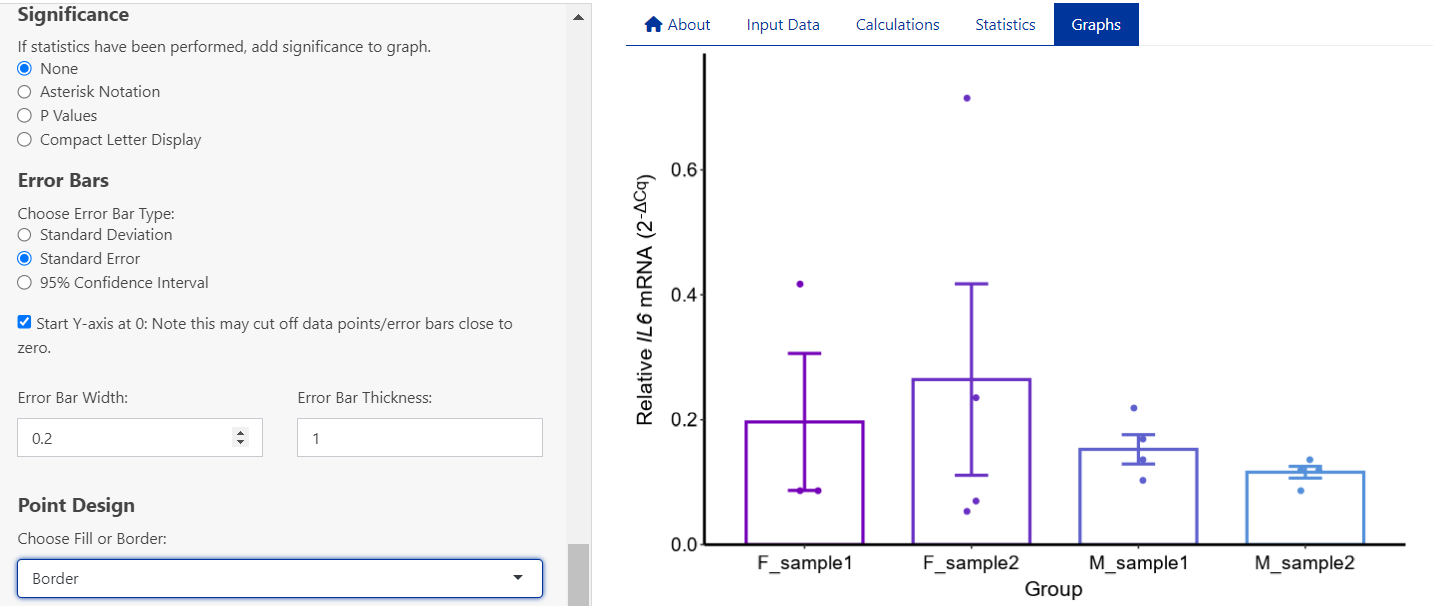
1. The shape of the points can be changed to paired shapes by selecting the checkbox:



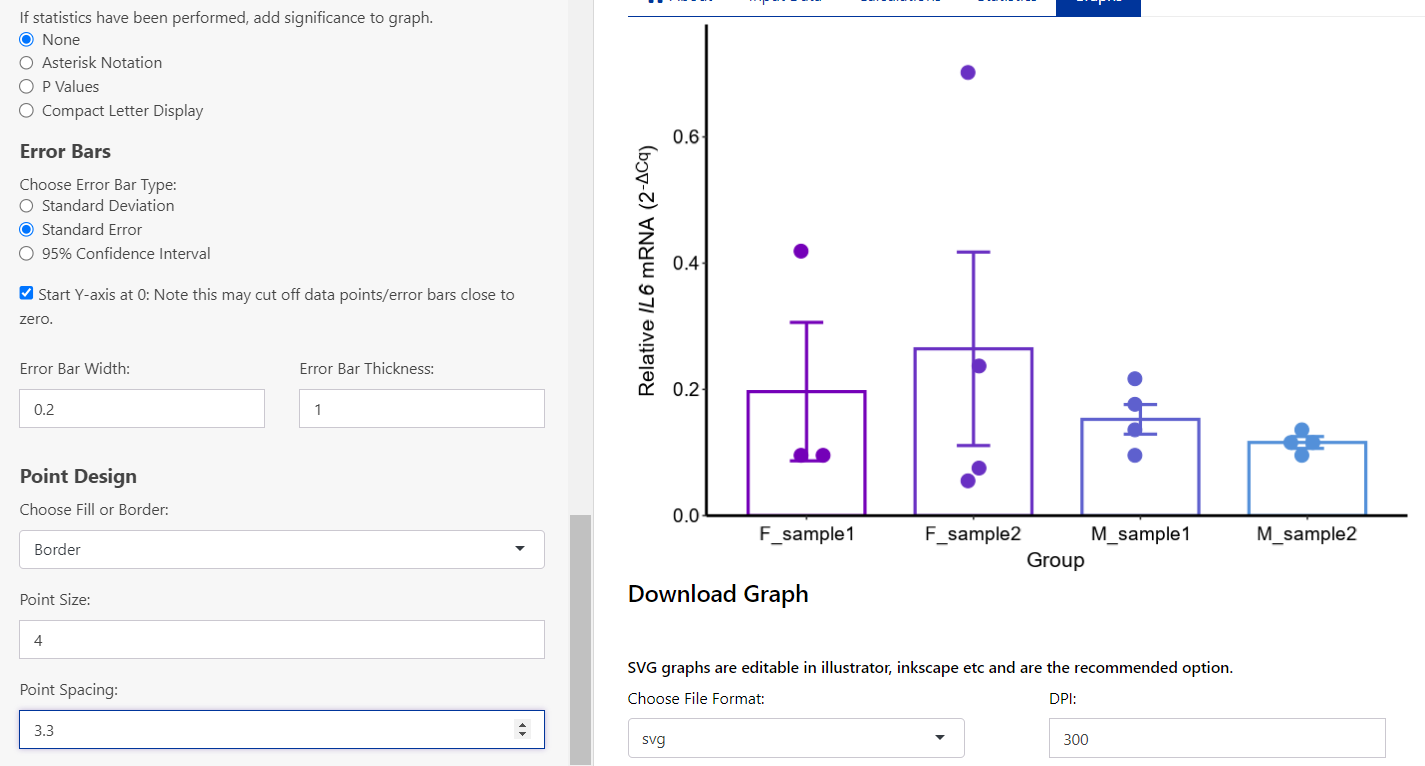
For column graphs:

1. The column graph can be customised by colour fill or border by toggling the Choose fill or Border dropdown menu:



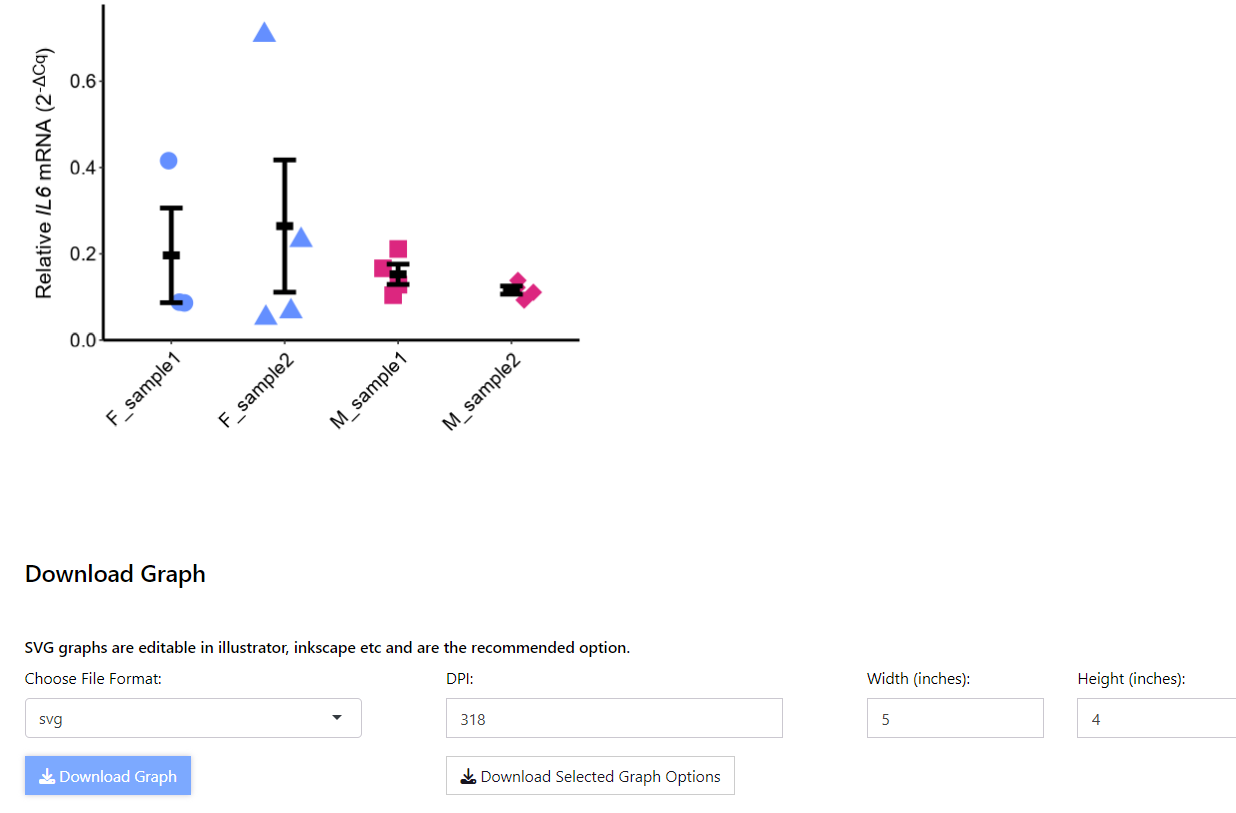


1. Point size and spread can be altered by changing the number of the corresponding box:

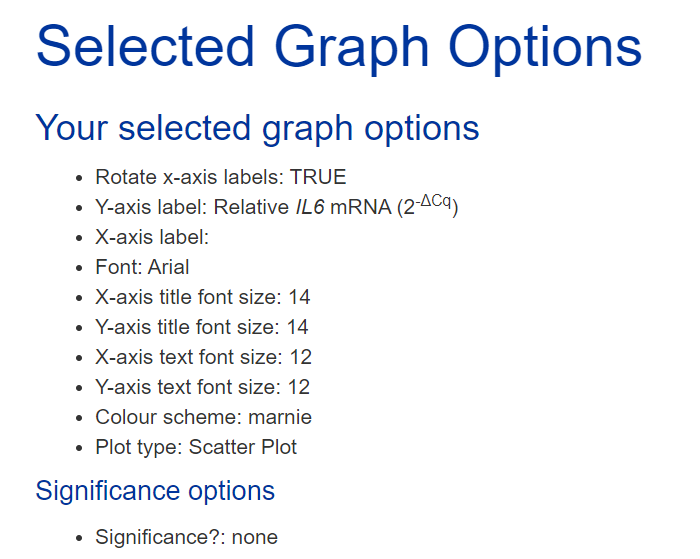


#### **Downloading Graphs:**

1. Graphs can be downloaded by clicking the Download Graph button. Graphs can be saved as an SVG, tiff, png, or jpeg by toggling the dropdown menu. It is recommended to download as an SVG, as these files can be opened in Inkscape or Illustrator to further customise graph aesthetics.
2. The width and height of the graph can be changed, as well as the dots per inch (DPI) by altering the numbers in the corresponding boxes. Note that by changing the DPI it alters the graph dimensions, and so may change the aesthetics of the graph. After downloading a graph check it meets your requirements.



1. To save the options selected to create the graph, press the download selected graph options button. This is useful as a reference to keep graph aesthetics similar across samples/genes etc. It is downloaded as a HTML file which can be opened in most web browsers.



# **Error Messages**

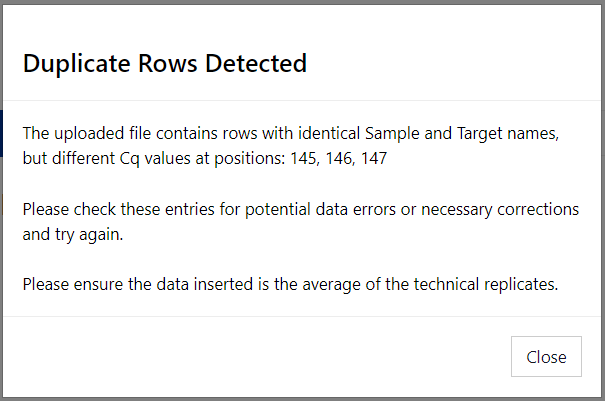
Numerous error messages may appear when using ProntoPCR. These are designed to guide the user into selecting the correct option or ensuring the correct data is inserted/used etc. If an error message appears within the app that is not described in the below section, please submit a screenshot of the issue and any information about how you got to this error in Github: [GitHub Issues](https://github.com/MarnieMaddock/ProntoPCR/issues). Please cover/anonymise sensitive information.

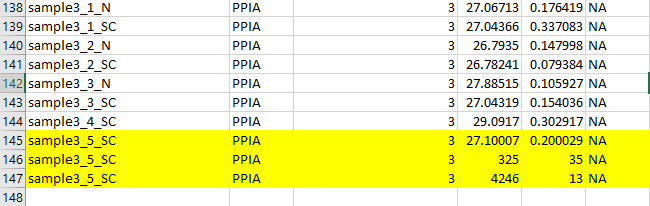
## **Input Data Errors**

When uploading your CSV file to ProntoPCR, it must meet certain requirements. If the file does not comply, error messages will prompt you to adjust the CSV file accordingly.

1. Duplicate Row Detected:

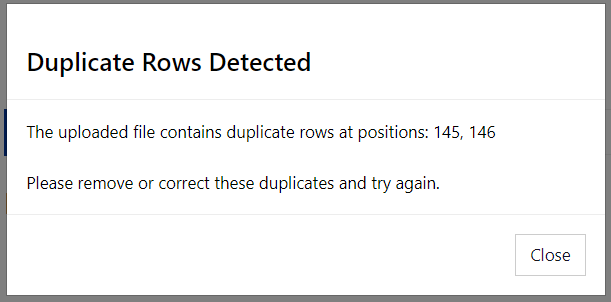
Often this error appears when inserting a file that includes technical replicates of PCR data. Please average these technical replicates and insert the averaged data. There should only be 1 sample for that gene in the file.

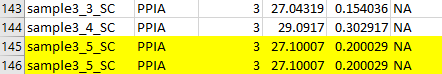




1. Duplicate Rows Detected:

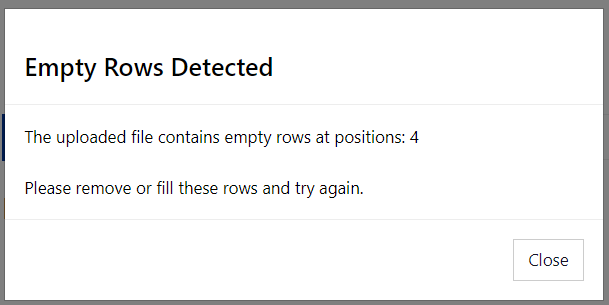
Two or more rows in the .csv file are identical. Remove the duplicates.

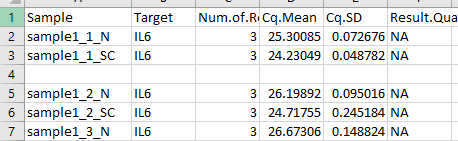


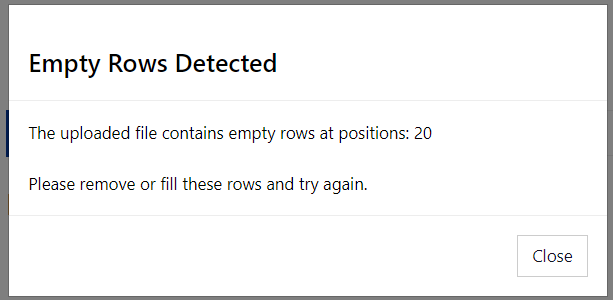


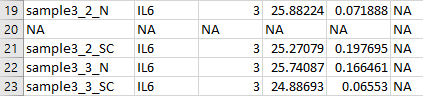
1. Empty Rows Detected:

Remove empty rows or NA rows to fix this error:



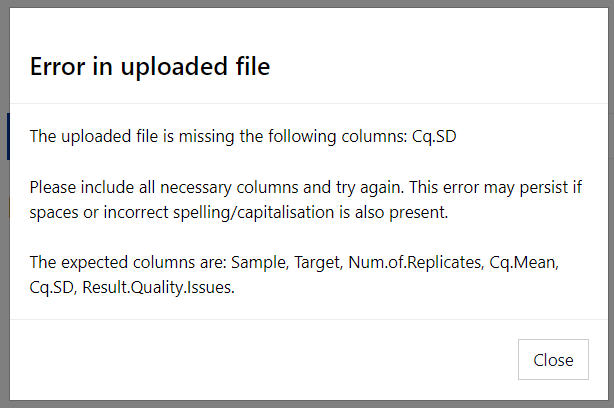


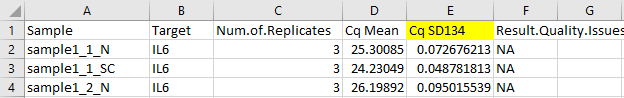




1. Error in uploaded file:

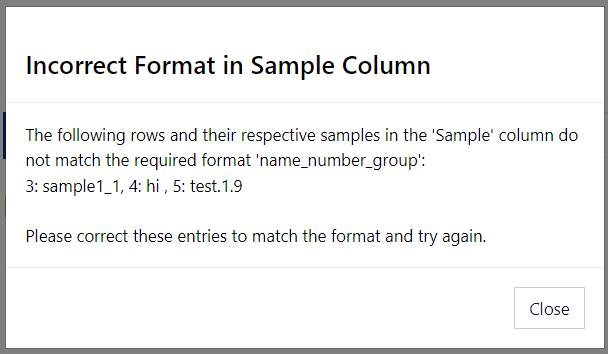
This appears when there is a discrepancy detected in the column names. It will prompt the user with the problematic column name. Correct the column name to resolve this issue.

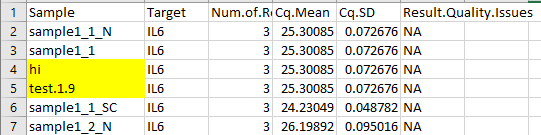




1. Incorrect Format in Sample Column:

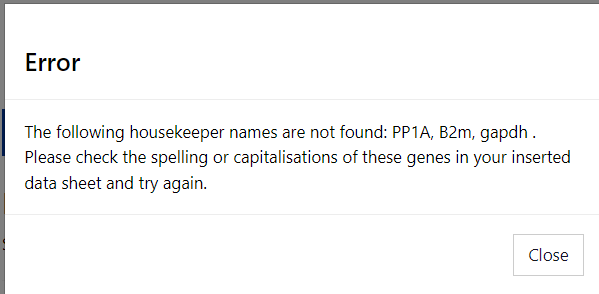
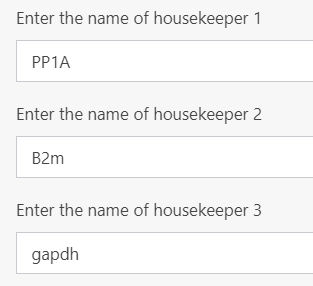
An error appears if the samples within the Sample column are in the incorrect format. It should be in the format name\_number\_group. See Input Data section for more information.



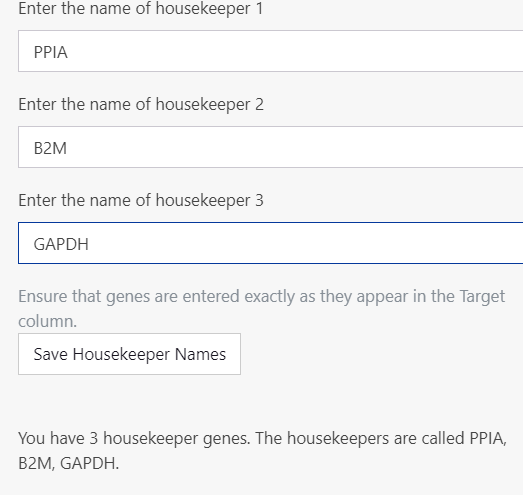


1. Error: The following housekeeper names are not found:

Appears if the spelling, capitalisation etc does not match the gene present in the Target Column. Rename the gene in the csv file, or alter the inputted name. If all genes are matched, they will appear in the side panel after pressing the Save Housekeeper Names button.

Correct:

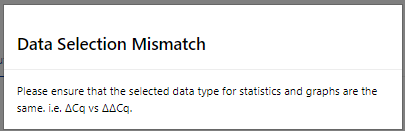


## **Statistics and Graph Errors**

Given the numerous options for performing statistics and creating graphs, error messages will appear if there are discrepancies between tabs, ensuring that the data used for statistical analysis matches the data being graphed.

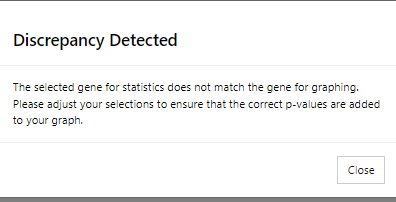
1. Data Selection Mismatch:

This error appears if the user has selected 2^-ΔCq for statistics and 2^-ΔΔCq for graphing (or vice versa). Change the stats/graphs tab to match either 2^-ΔCq or 2^-ΔΔCq.



1. Discrepancy Detected: The selected gene for statistics does not match the gene for graphing.

This error occurs if the gene in the statistics tab does not match the selected gene in the graphs tab. Change the stats/graphs tab to match the gene name.



# **Version History**

ProntoPCR was developed in R version 4.2.3, and R Shiny version 1.9.1. All code is available on the GitHub repository: <https://github.com/MarnieMaddock/ProntoPCR>

Version 1.0.0:

* Initial publicly available version.

# **Feature Requests and Issues**

If you'd like to request a new feature or report an issue, please submit it on GitHub: [GitHub Issues](https://github.com/MarnieMaddock/ProntoPCR/issues). Please include as much detail as possible to help us address your request effectively.

# **References**