

Whole brain data compilation and quality control

Purpose

This is a protocol for R shiny apps that accept output from the QUINT workflow (or similar mouse brain registration and segmentation data) and allows you to create grid and anatomical heatmaps for mouse brain data. The two applications are compiled onto a singular webpage at:

<https://github.com/MXHend/Mouse-Brain-Heatmap-Website>

Preparation

First, create a folder which will contain all your data for this project and name it whatever you would like. Within this folder, create two folders labeled 'left' and 'right' for hemispheres. If you only have one hemisphere of data, just create a singular folder labeled 'data'.

After completing the QUINT workflow, you should have a folder (or more likely, multiple folders) called "Reports" which contains another folder with an ending label of RefAtlasRegions. In this folder will be at least two files with one of them having the same name as the folder. Ignore that file and copy the rest into the appropriate folder within the project folder you created in the beginning. Do this for every report folder you have until all the files are located in those two (or one) folders in your project folder. You may now open the website.

Using N2U for the first time:

1. Begin by hitting the browse button labeled for the '**left side**' and select all files in the 'left' folder.
 - a. If you have a 'data' folder, select all files in that folder instead.
2. Hit the 'open' button to start loading the data.
3. While that is loading you can hit the browse button labeled for the '**right side**' and select the files in the 'right' folder and hit the 'open' button to begin loading that data.
4. Wait until there is a message on the right side of the app that says the last side you loaded is done.
5. Double check that the above step has appeared and then click the button labeled '**click this when all files are done loading**'.
6. Wait until the message on the right side changes and mentions the 'blank annotation' and the 'checkpoint' are done.
7. Press the button labeled '**Download Checkpoint 1**' and save it within your project folder. This will be used the next time you load the data.
8. Press the button labeled '**Download blank annotation file**' and save it within your project folder.
9. Open your file explorer and navigate to your project folder.
10. Open the **annotation** file you just downloaded and **fill out** all columns labeled with "NA" (if you don't have a column, then you may leave it NA)
 - a. Mouse – Mouse name
 - b. Sex – Mouse sex
 - c. Treatment – Mouse treatment type
 - d. MPI – Months post injection

- e. Genotype – Mouse genotype
 - f. Marker – What you are comparing in the mice, such as a cell type or stain.
 - g. Include – Use this column if you wish to subset the full dataset.
 - i. Y = include
 - ii. N = do not include
11. **Save** this file as a CSV file with a different name so that you know you have completed it.
 12. **Reload** the webpage or stop and restart the RStudio app.

Using N2U when you have saved checkpoint 1 and completed filling out the annotation file:

1. Begin by clicking the Browse button labeled by the phrase **'Input checkpoint 1 here'** and select "checkpoint1.rds" you saved earlier and click 'open'.
2. Check the right side of the screen for confirmation that the checkpoint is done loading and then click the browse button to **input in the annotation file** you finished creating.
3. Once that has completed (check the right side for text saying has finished), **download Checkpoint 2** somewhere you have easy access to
 - a. This will be for if you would like to come back to the app again later or if the app manages to fail at any later step.
4. Once that is complete you may notice that the box labeled for selecting your **'x axis variables'** now has options to choose from. Please note the following when you select these variables:
 - a. These variables will be on the heatmap going left and right .
 - b. They will be combined and if they don't make sense, such as combining 'daughter' (region) with 'mouse', they will display inaccurate heatmaps or error out the program.
 - c. These variables will also be displayed individually in the form of a color key on the **outside left** of the heatmap.
 - d. **Common Examples:**
 - i. 'Mouse'
 - ii. 'Marker' and 'Mouse'
 - iii. 'Marker', 'Mouse', 'Genotype', 'Sex', 'Treatment', 'Batch', and 'MPI'
5. Next, select the **'y axis variables'** and note the following when selecting them:
 - a. These variables will be on the heatmap going up and down.
 - b. They will be combined, so double check they make sense avoiding a similar issue such as described for the x axis.
 - c. These will not have to match the display variables for the y axis.
 - d. **Common Examples:**
 - i. 'Daughter'
 - ii. 'Side' and 'Daughter'
 - iii. 'Side' and 'Parent'
 - iv. 'Side' and 'Specials'
 1. 'Specials' is a special variable containing the following ABA regions: CTXsp, fiber tracts, HPF, HY, Isocortex, OLF, PAL, STR, TH, MB, P, MY, and CB
6. Select the **'display' y axis variables** next noting the following:
 - a. These variables will be displayed in the form of a color key on the **outside top** of the heatmap.
 - b. These must make sense for this axis such not using regions and mice together.

- c. Please note that the more elements in a variable. the less distinguishable the color key is going to be (i.e. don't use 'daughter') This is a good place to use 'specials'.
 - d. **Common Examples:**
 - i. 'parent'
 - ii. 'side' and 'parent'
 - iii. 'side', 'parent', and 'specials'
7. When you have selected variables for all the above options, click the button labeled **'Get Options'**.
- a. This button will use the data you just gave it to give you more specific options for your heatmap.
8. Next, look at all the options that are selected in the box labeled with **'variables to see'** and confirm you want to see all of the above variables on the heatmap.
- a. You will only ever touch this in an instance where you want to calculations to be the same, but the display to only show part of it.
 - b. To not include something in the calculations, remove it using the annotation file.
9. **This step will determine the values of the heatmap, so pay close attention:** From the dropdown list labeled **'select the nutil variable'**, select the variable you will be using as your value (i.e. Load)
- a. As a note: load is object area/region area so if you are looking for that calculation, this is it.
10. **If you would like to change the variable in any way:**
- a. Select **Multiply by 100** if you would like to do as stated.
 - b. Select **Use variables** if you would like to select variables **to create a percentage**. Select the variables based on the following information:
 - i. Segments by each variable labeled here and then creates percentages of any other variable you selected above that is not in this box.
 - ii. Here the program will group your data based on your selections and create a percentage separating from the largest to the smallest designation. For example:
 - 1. If your rows were 'mouse' and 'marker' (marker has two options per mouse), your selected nutil variable was 'object count', and you selected, 'Mouse', 'side', and 'daughter' here; it would first separate the data by mouse, then by side, then by daughter. Then it would look at the designations of the row and see that it's separated by mouse and marker so it calculates the percent each marker takes up as it is already segmented by mouse.
 - iii. If you are hoping to compare the percentage of elements in a variable (such as 'pSyn+' to 'pSyn-' in the variable 'marker'), then do not include that variable in this segmentation as it would then turn out to be 100% for each of them
 - iv. **Common Examples:**
 - 1. 'side', 'mouse', and 'daughter'
 - 2. 'mouse', 'daughter'
 - c. **Note:** If you try to select both "Multiply" and "Use variables", it will do neither.
11. **If you would like the data to be in log 10 scale:** check the corresponding box.

12. **If you would like your heatmap to be trimmed:** check the corresponding box. (This does not affect the calculations.)
 - a. A trimmed heatmap removes any column with more than 50% of the data missing.
13. If you would like to **remove ABA regions** from your heatmap, select any amount of them from the drop down labeled **'Select any regions to remove'**.
14. **If you would like row and column names:** select from the corresponding drop down boxes and select what you would like to see.
 - a. x and y are the combined column names that are created when you first select the x axis variables and y axis variables and will most likely contain the most information.
15. The **color scale** is **not** something you will need to adjust your **first time** loading the heatmap.
16. Now that you are done selecting the base variables for your heatmap, click the button labeled **'Create heatmap'** and wait for a preview to appear in the right side.
 - a. Once this is done, you may find that your data is not distributed along the full scale. This is due to the color scale I mentioned before to skip over.
17. Now that the heatmap is done, you will note some text at the top of the right side which says 'min' and 'max'. Please enter these values into the textboxes labeled **'Min Value'** and **'Max Value'** for the **color scale**.
18. At this point you can adjust the options under the label **'Custom Colors'**. **If you would like to**, then **select either the Viridis color scale checkbox or the 2 color scale checkbox or the 3 color scale box** and then choose or adjust the colors to what you would like.
19. When you are done, click the **'create heatmap'** button again and wait for the preview to appear.
20. **To save the heatmap:** you have two options you can choose from or download both.
 - a. Click the button labeled **'Download full Heatmap'** and this will download a heatmap with all information you selected above. Make sure to download it into a directory you can easily access to check if you like your chosen options.
 - b. Click the button labeled **'Download Just the Plot'** and this will download a heatmap without any of the legends for easy editing. Make sure to download it into a directory you can easily access to check if you like your chosen options.
21. Look at your heatmap and determine **if you need to change anything**. If you do, start from step 3 and check over each step at a time.
 - a. **Don't forget** to press the button in the middle while changing options before attempting to recreate the heatmap again.

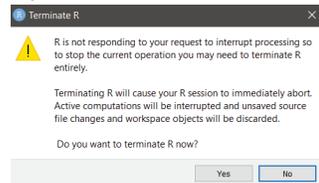
Using N2U when you have saved checkpoint 2:

1. Begin by clicking the Browse button labeled by the phrase **'Input checkpoint 2 here'** and select "checkpoint2.rds" you saved earlier and click 'open'.
2. Check the right side of the screen for confirmation that the checkpoint is done loading and then **continue from step 4 above**.

For integration with the Mouse Brain Heatmap (MBH) app for anatomical heatmaps OR additional numeric and/or annotation information for the heatmap:

1. If you would like to display data in anatomical space, **take it into the MBH app**, click the **'Download MBH Table'** button and save it wherever you can access it.

2. If you would like **the color scheme to also be the same in the MBH app**, click the **'Download Colors'** button.
3. If you would like the values of the heatmap or any of the axis annotation information of the heatmap, click the buttons labeled **'Download Data'**, **'Download AnnoX'**, and **'Download AnnoY'**.
4. Finally, **if you would like access to the many variables created during this program running or the data frame that contains the raw data at any point:**
 - b. Open the app in R Studio.
 - c. Run through all of the steps.
 - d. Once it has created a heatmap, wait a minute or two for everything to fully load.
 - e. Click back into R Studio without closing the app.
 - f. Find the console in the bottom left corner and click the button that looks like a stop sign **ONCE** and only once.
 - g. Wait very patiently for the app to close on its own.
 - h. If it gives you an error like below click 'No' and wait some more

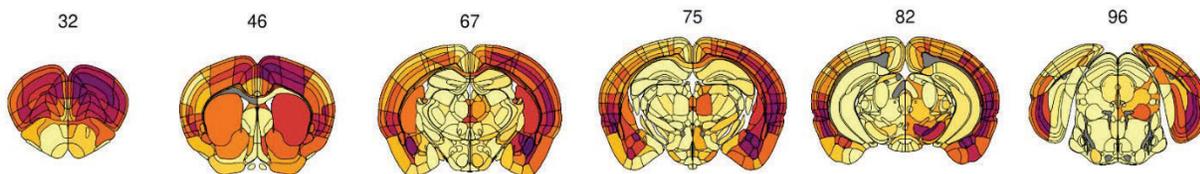


- i.
- i. Eventually it will stop the app and variables will appear in the top right box labeled 'environment'.
- j. To see these variables, double click their names to open a viewer pane.
- k. To save a variable as a csv file write the following code in the console:
 - i. `write.csv(variableName, "filename.csv")`
 - ii. **If you would like to designate the location of the file**, include the file path in front of the filename with **all back slashes replaced with forward slashes**
 1. `"~\projects\file.csv"` is **incorrect**
 2. `"~/projects/file.csv"` is **correct**

Anatomical heatmap

Purpose

This program plots regional quantification matrices composed of Nutli outputs in anatomical space (ABA CCFv3). This input matrix can be compiled using N2U or manually. Plots consist of coronal sections based off the Allen Brain Atlas and present options of single or multi-figure plots displaying your parameter of interest in anatomical regions and subregions.



Example anatomical heatmap.

Plotting Options

Users will have the option to develop Raw data single-figure or multi-figure plots with either the mean or median displayed for each anatomical region. **Figure Numbers** refer to the image numbers on the [ABA](#) website.

Under the **Raw Data single-figure** tab in the MBH webpage, plots can be generated comparing regional data across conditions (MPI, Genotypes, Treatments, Markers, Sex) for a single figure. The **Raw data multi-figure** tab allows you to generate plots for multiple figures, consisting of coronal sections spanning across the brain, but will only display a single condition at a time.

Users also have the option to plot at the **daughter** level by default which includes all atlas regions and subregions down to the specific layer/part (e.g. MOp2/3, MOp5, BLAp), or the **parent** level which plots more generally the main regions at the top of atlas hierarchy inclusive of all layers/parts (e.g. MOp, BLA).

Protocol

1. **Input:** Determine whether you want to plot single or multi-figure plots. Upload your compiled matrix by hitting browse on the prompt '**Please upload the MBH table output**'.
 - a. This should be labeled as 'MBHTable' and is generated after selecting 'Download MBH table' in the N2U program.
 - b. In theory, any csv file following the same format as seen in the MBH table would suffice for this workflow.
2. **Heatmap Colors:** if you generated a custom color scheme in the N2U app, upload the .rds file in the brose button under the '**If desired, upload heatmapColors.rds here**' (not necessary step if using the default scheme).
3. **Figure Numbers:** select the Figure Number corresponding to the sections for your project.
 - a. This will only be one Figure Number if creating single-figure plots. Multi-figure Numbers should be separated by a comma *without spaces between them*.

- b. If you are unsure which figure you are working with, visit the [ABA](#) website and determine which section (out of 132) corresponds to the anatomical region(s) you are interested in plotting.
4. **Summary Level:** select the measure summary you would like to have plotted.
 - a. Selecting 'Mean' will display in each anatomical region the measure of interest averaged across all sections for the conditions set (genotype, treatment, etc). Selecting 'Median' will display median values.
5. **Region Level:** select the depth level for atlas regions.
 - a. By default, region level will be set to 'daughter' which will include anatomical regions split by layer/part. Choosing 'parent' will plot regions with all layers/parts combined as one region.
6. **Conditions:** the remaining dropdowns allow you to choose what you would like to plot.
 - a. single-fig allows you to compare across conditions
 - i. The MPI, genotype, treatments, and markers can be selected for when plotting. The data for mice within the conditions set here will be summarized and displayed in plots.
 - ii. By default, plots for each sex will be distinguished. If you would like to combine summary results across sexes, select 'Averaged' in the dropdown box under 'Sexes to Plot'.
 - b. multi-fig will only allow you to plot for a single condition.
 - i. Whereas single-fig can generate multiple plots for each condition in the same output, multi-fig only displays a single condition at a time.
7. **Plot Data**
 - a. Select 'Plot data' to create the heatmaps you have set. This may take time to load. Each time you adjust conditions or inputs you will need to reselect 'Plot data' to see the new plots.
 - b. 'Download Plot' generates .svg files for generated plots.
 - c. 'Download Data' generates a .csv with information on what anatomical regions are plotting and the values being displayed in each.

Notes:

-**Min/Max:** at the moment, the min and max can only be adjusted in N2U. If you would like to adjust the minimum and maximum values displayed, a new MBH Table will need to be created with adjusted min/max.