This file will demonstrate how to analyze .xml files exported from a Bioplex 200 with the 'ANC' and 'CNA' programs.

The example files provided were generated by stimulating neurons with NMDA for 5 minutes (similar to Lautz et al, 2018). The experiment consisted of stimulating acute brain slices from a wildtype animal with NMDA or control media, refered to as ACSF. (The experiment also involved stimulation with DHPG or glutamate, but those data will be ignored here).

The sample layouts were as follows (the order is pseudo-randomized to avoid artifacts due to PiSCES degrading over time or from other plate effects)):

	Row 1 and 5:	Row 2 and 6:	Row 3 and 7:	Row 4 and 8:
Expt1:	Glutamate	NMDA	DHPG	ACSF
Expt 2:	Glutamate	DHPG	NMDA	ACSF
Expt3:	DHPG	ACSF	NMDA	Glutamate
Expt4:	Glutamate	ACSF	NMDA	DHPG

So, for Expt 1:
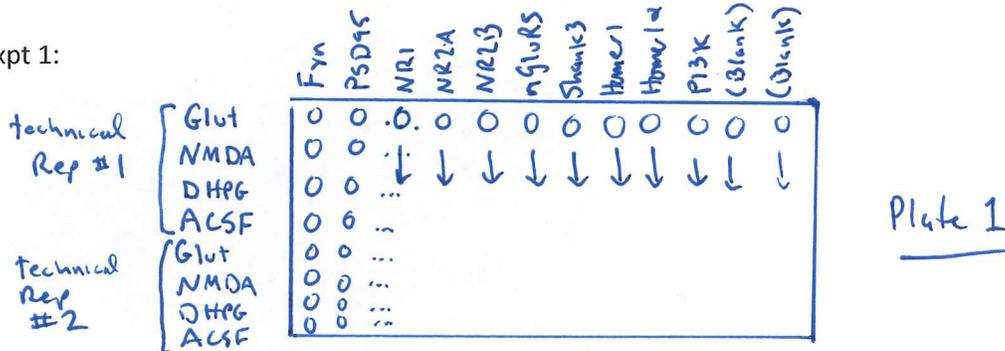
The probe antibodies in plate 1, by column: Fyn, PSD95, NMDAR1, NMDAR2A, NMDAR2B, mGlur5, Shank3, Homer1, Homer1a, PI3K.
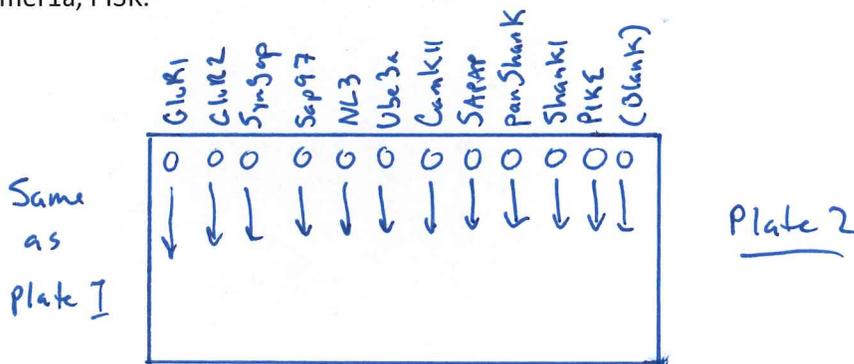
Probe antibodies in plate 2, by column: GluR1, GluR2, SynGAP, SAP97, NL3, Ube3a, CamKII, SAPAP, panShank, Shank1, PIKE.

The bead mixture in all plates consisted of the following IP antibodies conjugated to bead regions in (parenthesis): Fyn (12), PSD95 (08), NMDAR1 (13), NMDAR2A (26), NMDAR2B (28), mGlur5 (77), Shank3 (43), Homer1 (30), PI3K (15), GluR1 (64), GluR2 (68), SynGAP (85), SAP97 (45), NL3 (89), Ube3a (91), CamKII (47), PIKE (70), Shank1 (96).

Step-by-step analysis instructions:

1) Open Matlab, and add the 'ANC Program' folder and all sub-folders to the active directory path.
2) Open the 'ANC_input.m' file (which has been pre-populated with the above information about the experimental setup) and run the program. The program will run for a couple of minutes.
3) Open the file "WT_aCSF_v_NMDA_Hits.csv". These are your ANC-significant PiSCES at two levels of stringency- 4 of 4 experiments (22 hits), and at least 3/4 experiments (33 hits).
4) Open the "WT_aCSF_v_NMDA_MFI.csv" file. These are the MFI for each IP_probe across all experiments for input into R.
5) Paste-transpose the column titles of the data (ACSF_1 to NMDA_4) into the first row of a new excel sheet. Add the columns "experiment" for experiment number, and "treatment", for weather or not NMDA was added. Save this file as "traits.csv". Compare with the example file provided.
6) Open R Studio and set the active directory to the folder containing the "traits.csv" and "WT_aCSF_v_NMDA_MFI.csv" files.
7) Run through the R program to generate CNA modules. The attached R program has been configured to run with the attached "input" and "traits" files. See comments within the program for instructions, or see the tutorial for WGCNA here:
   https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/Tutorials/

   Below is a brief description of each section of code
   a. Load inputs
   b. Eliminate MFI values below an arbitrary cut-off point (MFI = 100) to reduce noise contributed by undetected PiSCES.
   c. Run the ComBAT function to eliminate batch effects. Compare the data clustering before (right) and after (left) ComBAT yields:
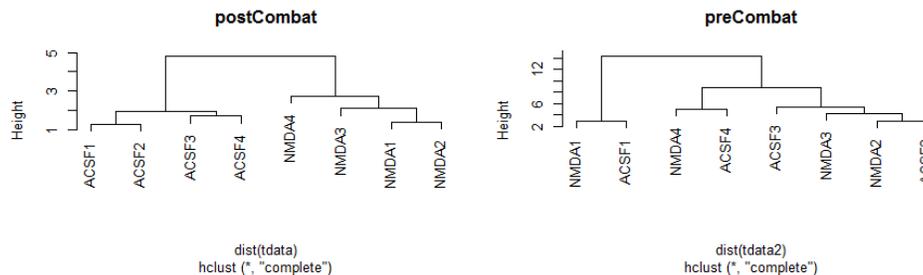

   d. Re-format the data from ComBAT output to WGCNA input
   e. Run PCA on data to see large-scale effects of treatment or batch. Here, the data separate across PC1 by treatment (ACSF vs. NMDA).
   f. Run CNA

g. Check the TOM plot. You should see that the color-coded modules form boxes that corelate well within-module and not between-module. This example is acceptable. Note, for example, the large red square that corresponds to the black module.

h. Generate module-trait relationships. This is the most important graphical output of CNA, as it identifies which modules correlate with "Treatment", i.e. NMDA vs aCSF. "Black" is highly correlated with treatment, and greenyellow is also significant ($p<0.05$) as well. <u>We will consider all PiSCES with module membership greater than 0.7 and $p<0.05$ in these modules as "CNA HITS"</u>.

Module-trait relationships

	experiment	treatment
MEbrown	0.03 (0.9)	0.15 (0.7)
MEpurple	0.0076 (1)	-0.39 (0.3)
MEmagenta	-0.0079 (1)	0.3 (0.5)
MEgreen	-0.00011 (1)	-0.17 (0.7)
MEturquoise	0.042 (0.9)	-0.45 (0.3)
MEblue	-0.022 (1)	-0.68 (0.07)
MEblack	0.0047 (1)	-0.94 (8e-04)
MEgreenyellow	-0.041 (0.9)	-0.76 (0.03)
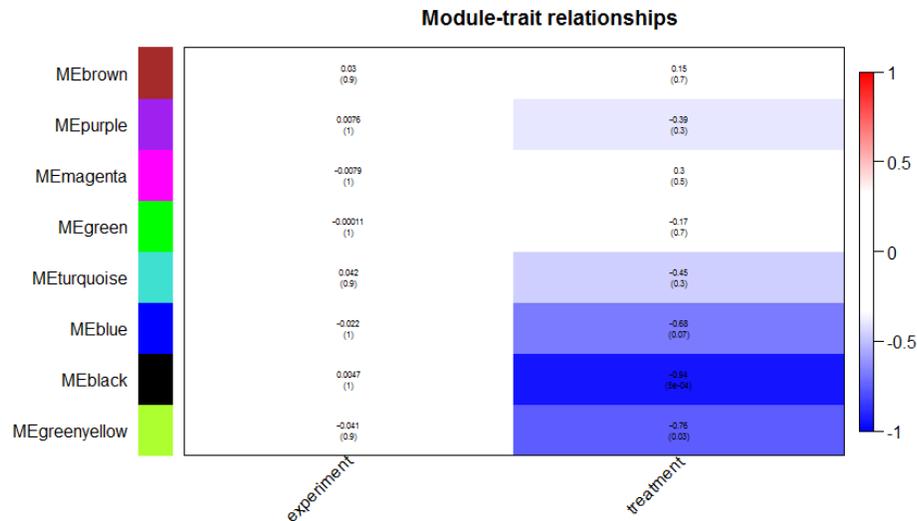
i. Define the variables for the output file and run the code to produce the file.

j. Name and save the output file.

k. To create a heatmap of a module of interest, enter the color-name of the module in line 249.

l. To create a heatmap of all ANC∩CNA hits (see steps 8-11), make a heatmap of the main significant module and write the output file in line 288. Open the aa.csv file in Excel, and delete all non-ANC significant PiSCES. You may need to repeat this step for other modules to include all ANC∩CNA hits. Then re-import your new list in line 289 and generate a heatmap.

8) Open the CNA_output file in Excel. Sort by p.MM.black (since black was the module highly correlated with treatment). Highlight all PiSCES with a $p<0.05$ and MM>0.7. Repeat with "greenyellow"

9) Open the WT_aCSF_v_NMDA.csv file that was generated by MatLab. We focus on the 3of4Union4of4 hits, which allows for a PiSCES to be non-significant in one experiment, but requires slightly higher p values to reach the statistical cut-off compared to 4 of 4. Copy-paste these 33 interactions to a new file labeled "ANCandCNA_hits".

10) Sort the CNA_output file alphabetically by Column 1 to make the next matching step easier.

11) Find each ANC hit in the list of CNA hits. If the hit is highlighted (i.e. in the black or greenyellow module), highlight it also in the "ANCandCNA_hits" sheet. If it is not highlighted, delete the hit from the "ANCandCNA_hits" sheet.

12) Consolidate the list by eliminating blank rows. Use "text to columns" to split column A into IP and probe by splitting using the "_" symbol. Label the columns IP, Probe, pVal1-4, Fold Change 1-4. Compare to the file provided "ANCandCNA.xlsx" to ensure accuracy.

13) Some interactions are only significant in 3 of 4, and we want to eliminate these non-significant outliers from our average fold change calculation. Use the if-then function to eliminate FC values whose p value is greater than 0.05. Simultaneously convert to log2. See ANCandCNA.xlsx, columns L-O for example formula.

14) Some interactions will be significant, but in the opposite direction in one of four experiments. We'll eliminate these outliers too. Cell L20 is the only example here.

15) Take the average of the 3 or 4 Log2FC numbers that remain.

16) Copy-paste values the IP, probe and averageLog2FC from step 18 into a new excel sheet. Name the sheet "cytoscape.xls".

17) Open Cytoscape, import->network->file and select cytoscape.xls

18) Label the IP as source node, Probe as target node, and keep AvgLog2FC as an edge attribute

19) Arrange it as you see fit.

20) Under "Style" tab and "Edge" subtab, use "continuous mapping" to map "AvgLog2FC" onto stroke color and edge width. Manipulate the colors and handle positions to get useful diagrams.

21) To make heatmaps, output the "aa" heatmap file in R using the write.csv command for the black module. See lines 250-292 of the R code. Repeat for the greenyellow module using a different name for the file.

22) Merge the two files in Excel, and manually delete all interactions that are NOT present in the ANCandCNA.xls list (sort by column A to make this easier).

23) Re-import the edited list into R and create a heatmap using the supplied function, R code lines 293-315.

ANCnCNA Hits