**Statistical Report for Protein Expression in Normal, Failed, and Rescued Learning in Murine Models**

Prepared by:

Andre Guerra, Ryan Kollitz, and Phillip Young

Arizona State University

**Executive Summary**

In this report we will provide statistical analysis for the results of the experiments performed by Ahmed MM, et al. Reported in ‘Protein Dynamics Associated with Failed and Rescued Learning in the Ts65Dn Mouse Model of Down Syndrome’. Additionally we will describe methods performed for statistical analysis of this data, performed using JMP Pro 13[[1]](#endnote-2).

The experimental setup consisted of 8 unbalanced groups of 72 total mice with groups representing different combinations of 2 level categorical variables; genotype, behavior, and treatment. We found 45 proteins had statistically significant different levels of expression between experimental groups (alpha\_UT=0.025).

The Statistical Methods section will cover the analysis methods, including assumptions and hypotheses. The section following will analyze each protein in detail, including visual analysis and description. A table of summarized statistics for each protein will be included in Appendix A. Finally, a conclusion will provide discussion of results and future topics of study.

**Introduction**

The purpose of this report is to provide statistical analysis for results of the experiments performed by Ahmed MM, et al. Reported in ‘Protein Dynamics Associated with Failed and Rescued Learning in the Ts65Dn Mouse Model of Down Syndrome’[[2]](#endnote-3) with additional analysis reported by Higuera C, et al.[[3]](#endnote-4) Data from this experiment was retrieved from the University of California Irvine Machine Learning Repository[[4]](#endnote-5).

The basis for this study is to identify proteins associated with a phenotype of rescued learning previously demonstrated in trisomic (Ts65Dn) mice that have been acutely treated with the drug memantine[[5]](#endnote-6). Learning is observed in mice using context fear conditioning (CFC) where a mouse is introduced to a novel environment and is either shocked immediately before exploring the environment, or is allowed to explore for a given amount of time then shocked. Upon reintroduction to that environment, normal mice that have been subject to context-shock will freeze in place for an extended period of time before exploring; these mice have been stimulated to learn. However, normal mice that have been subject to shock-context are less inhibited to explore when reintroduced to the environment; these mice have not been stimulated to learn. When the same procedures are applied to Ts65Dn mice, the context-shock mice fail to learn and are uninhibited from exploring. As mentioned above, learning can be rescued in Ts65Dn mice by treatment with injection of memantine prior to CFC, demonstrating the ‘freeze in place’ behavior that is seen in normal mice after context-shock.

The study was conducted using 8 experimental groups consisting of genotypic control (c) mice and trisomic (t) Down syndrome mice, further divided by behavior and drug treatment. For behavior, mice were either stimulated to learn using (CS) or not stimulated to learn using shock-context (SC). The groups were then partitioned according to drug treatment, either having been treated with memantine (m) or an equal volume of a control solution, saline (s).Experimental groups as follows; genotype-behavior-treatment (group size):

c-CS-s (9) c-SC-s (9) c-CS-m (10) c-SC-m (10)

t-CS-s (7) t-SC-s (9) t-CS-m (9) t-SC-m (9)

Reverse Phase Protein Array (RPPA) was utilized to measure expression levels of 77 proteins in the cortex, taking 15 measurements of each protein per mouse. The analysis provided in this report will focus on the protein expression levels between groups to identify factors that contribute to differential protein expression, i.e. those due to genotype, behavior, and drug treatment. As such our statistical research questions are:

1. Is there a difference in the levels of protein expression between different classes of mice?
2. Is there a difference in the levels of protein expression in control and trisomic mice?
3. Is there a difference in the levels of protein expression in mice that have been stimulated to learn vs. mice that have not been stimulated to learn?
4. Is there a difference in the levels of protein expression in mice that have been treated with memantine vs. mice that have been treated with saline?
5. Do mixed factors (genotype, behavior, behavior) have effects on the levels of protein expression in mice?

**A Brief Analysis of Variables**

The data set contains a total of 79 variables, 76 being continuous variables, each a unique protein with an associated level of protein expression for each measurement, and 3 categorical variables.

The categorical variables are:

1. Genotype – Let G represent the genotype of a given mouse with values of either normal (c) or trisomic (t)
2. Behavior – Let B represent the behavioral treatment a given mouse was subjected to, with values of either Shock/Context (S/C) or Context/Shock (C/S)
3. Treatment – Let T represent the type of solution injected into a given mouse, with values of either memantine (m) or saline (s)

**Statistical Methods**

To answer the first statistical research question (SRQ1), we used Kruskal-Wallis test to assess effect of experimental group on protein expression and used a threshold of unusualness of αUT = 0.025 to evaluate p values.

Let tau represent the treatment affect, class, with 8 levels indexed by i. This allows us to look at differences between groups, however, we will not be able to make statistical inference based on single treatment effects. We also ran Steel-Dwass All Pairs comparisons, however, the sheer number of pairings deemed significant in this way was overwhelmingly high, to enumerate and describe each pairing in this report may not provide any meaningful understanding. When we were able to reject the null hypothesis we determined practicality of the difference using the effect sizes η² and ε².

The assumptions made for this model include that our response variable is continuous and that we have independence between measurements for that variable. The continuity assumption is valid, however we had to make changes to the original data file to ensure we were not violating assumptions of independence. As the original data file included 15 replicates of expression level per protein per mouse, we took the value of the Sample Arithmetic Mean of the replicates to represent each mouse in our testing.

We were unable to perform statistical analysis for SRQs 2 through 5 due to methods being outside the scope of our knowledge. Potential methods to investigate each factor and the potential interactions of those factors are Threeway ANOVA with Interaction or Self-Organizing Maps. The Threeway ANOVA would likely have to be performed using ranked-sum, if possible, as there are protein expressions within this dataset that do not follow a normal distribution, and issues with homoscedasticity may arise as well.

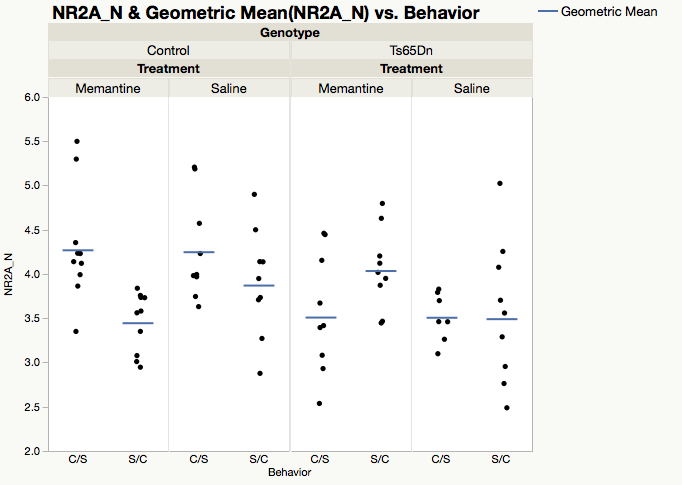
**Analysis**

In this paper we used Kruskal-Wallis tests for all of our analysis by comparing the expression levels of each of the proteins with the class (genotype-behavior-drug) of the mice. The use of Kruskal-Wallis over One-Way ANOVA was required by the nature of the distributions of the protein expression levels in the mice, examples of which are shown below.

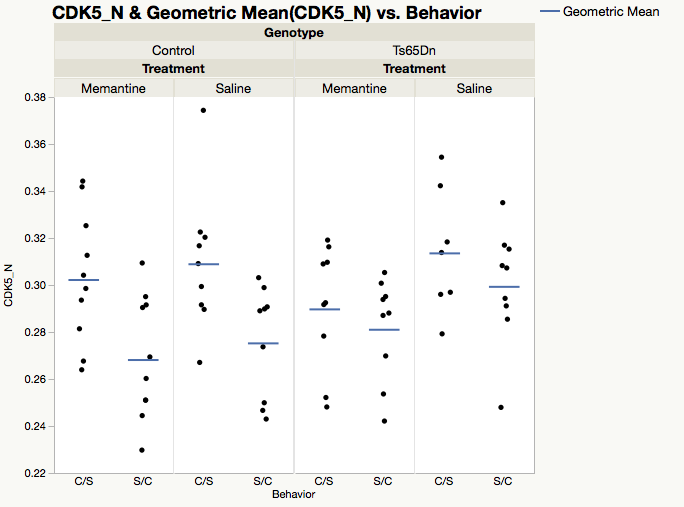


These two data visualizations show the distributions, outlier box plots, and shadowgrams of proteins pPKCAB\_N (left) and DYRK1A\_N (right). These are just two of the proteins that don’t follow a normal distribution so we already eliminate the use of parametric tests.

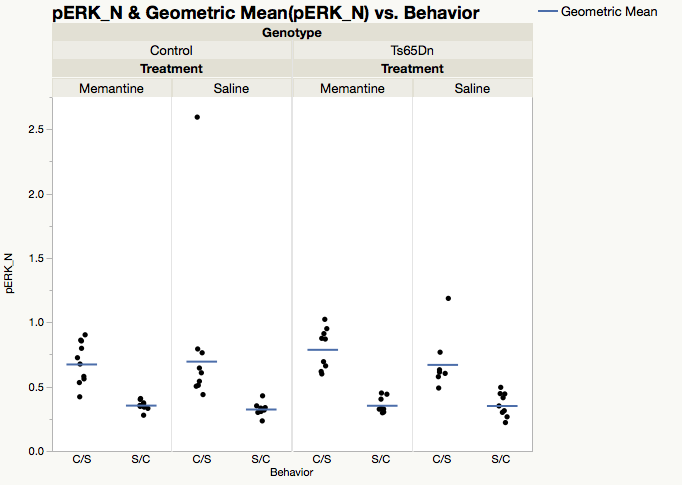
For the following visual analysis, we created scatter plots grouped by experimental group, and used a floating line representing the value of the group’s Sample Geometric Mean (SGM) for protein expression of the given protein. The reason we used SGM is due to the unbalanced design of the groups, there were varying amounts of mice in each experimental group as well as varying levels of protein expression.



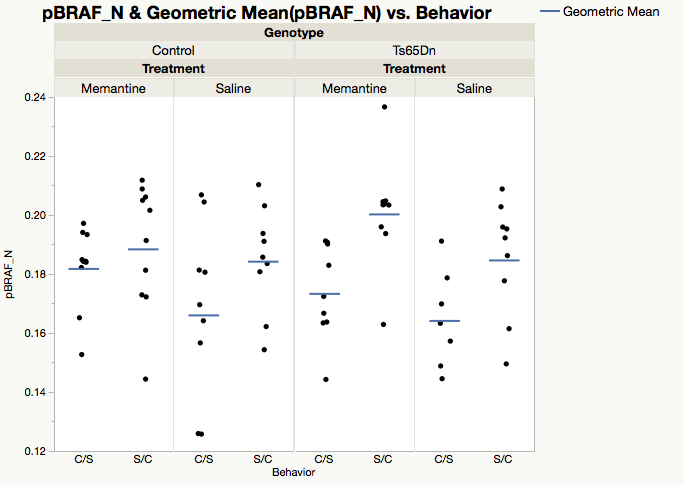
There is statistically significant evidence to support the belief that NR2A\_N is differentially expressed among experimental groups of mice (H(7)=20.8598, p=0.004). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.21656, ε²=0.2938). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding genotype steady, we can see that there are apparent differences in the behavioral groups of the control mice. The CS group has higher levels of expression than their SC counterparts.



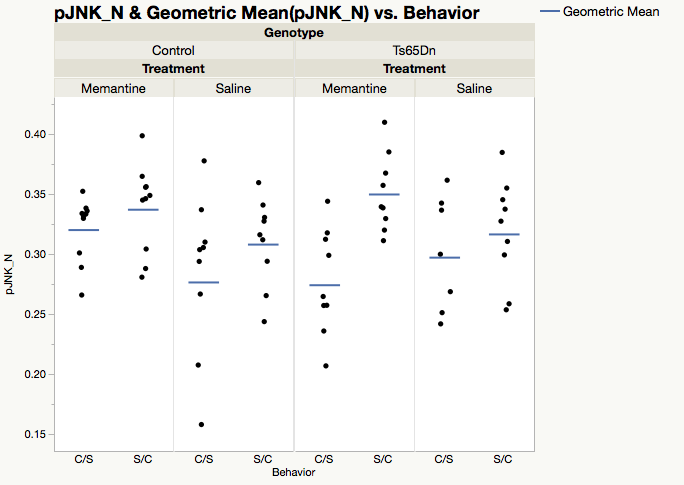
There is statistically significant evidence to support the belief that CDK5\_N is differentially expressed among experimental groups of mice (H(7)=19.3028, p=0.0073). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.19223, ε²=0.27187). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze a graph of the data to better understand which effects play a large role in protein expression. Visually, there are apparent differences in between behavioral groups. For each of these pairs, the CS group has higher levels of expression than their SC counterparts.



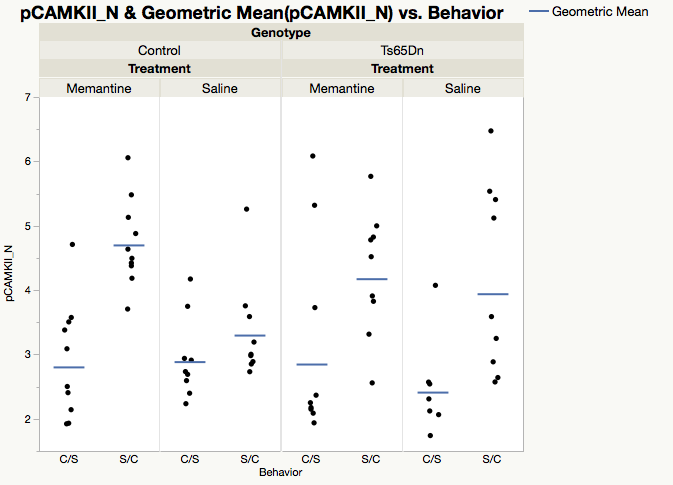
There is statistically significant evidence to support the belief that pERK\_N is differentially expressed among experimental groups of mice (H(7)=53.1294, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.72077, ε²=0.7483). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has notably higher levels of expression than their S/C counterparts.



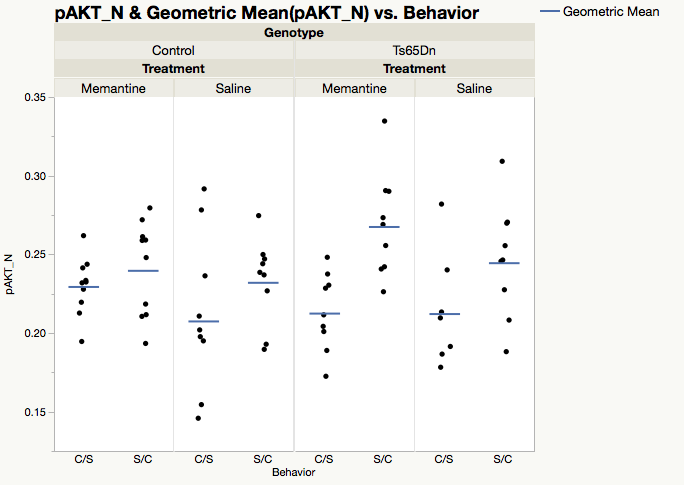
There is statistically significant evidence to support the belief that pBRAF\_N is differentially expressed among experimental groups of mice (H(7)=18.3947, p=0.0103). The effect size for this difference is medium, meaning this difference is also practically significant (η²= 0.17804, ε²=0.25908). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has notably lower levels of expression than their S/C counterparts.



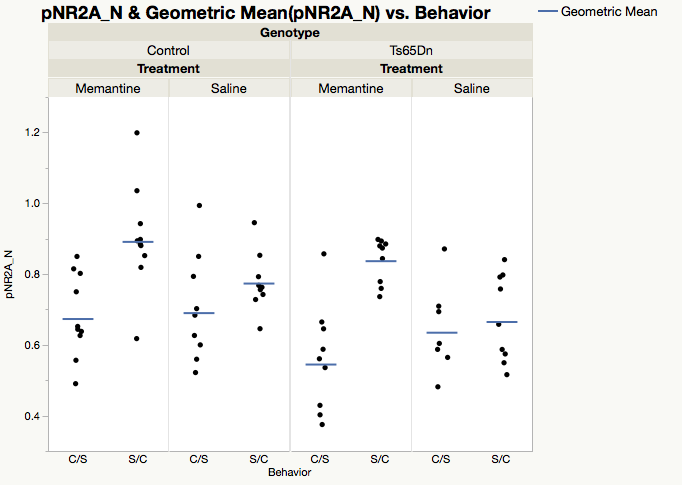
There is statistically significant evidence to support the belief that pJNK\_N is differentially expressed among experimental groups of mice (H(7)=17.8696, p=0.0126). The effect size for this difference is medium, meaning this difference is also practically significant (η²= 0.16984, ε²=0.25168). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has notably lower levels of expression than their S/C counterparts.



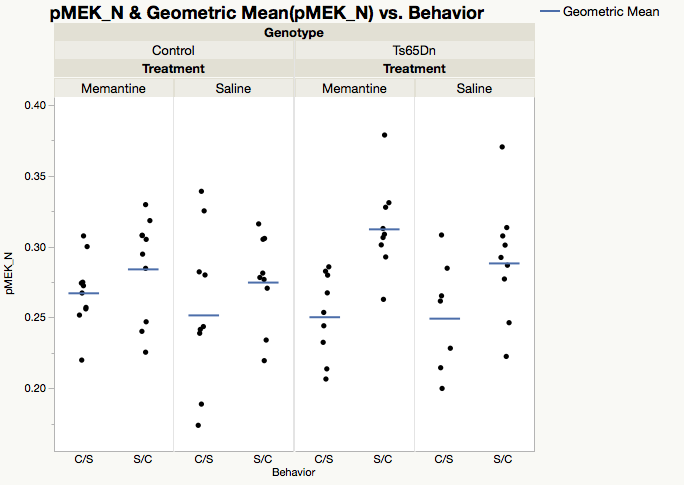
There is statistically significant evidence to support the belief that pCAMKII\_N is differentially expressed among experimental groups of mice (H(7)=28.5872, p=0.0002). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.33574, ε²=0.40123). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has notably lower levels of expression than their S/C counterparts.



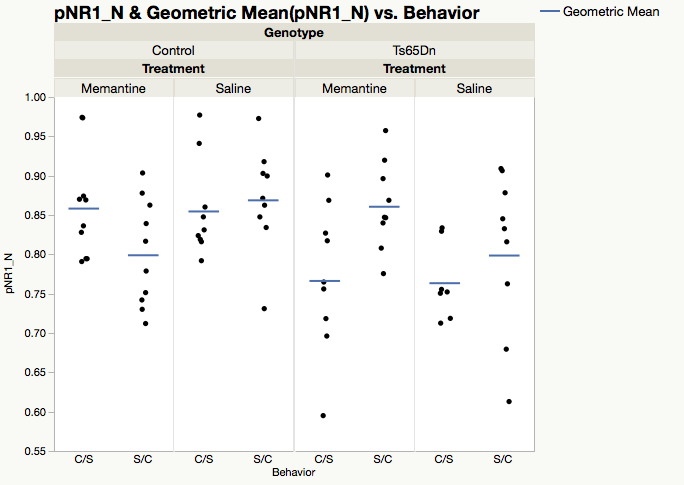
There is statistically significant evidence to support the belief that pAKT\_N is differentially expressed among experimental groups of mice (H(7)=16.876, p=0.0182). The effect size for this difference is medium, meaning this difference is also practically significant (η²= 0.15431, ε²=0.23769). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has notably lower levels of expression than their S/C counterparts.



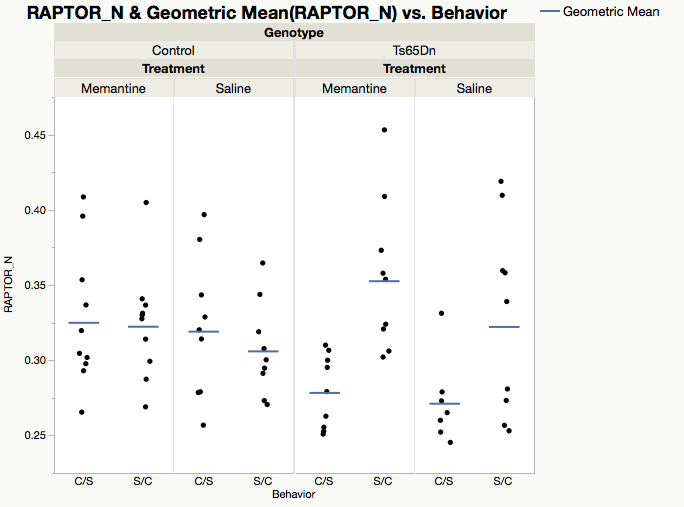
There is statistically significant evidence to support the belief that pNR2A\_N is differentially expressed among experimental groups of mice (H(7)=31.2753, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.3793, ε²=0.4405). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. If we hold genotype constant we can see that the S/C - m mice have an unusually high level of expression for this protein.

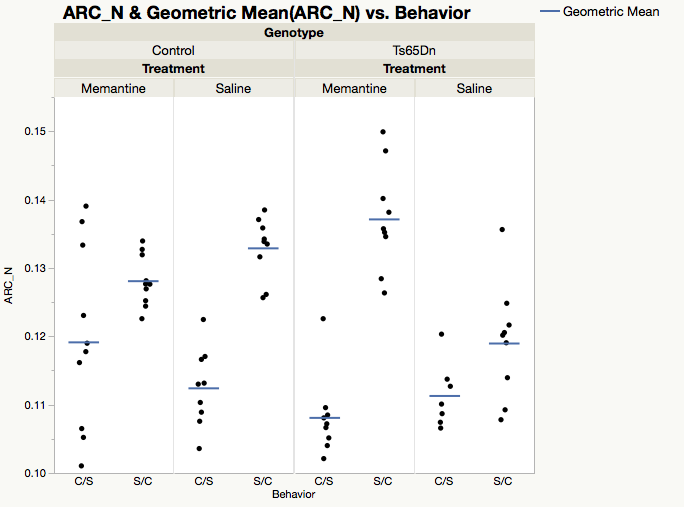


There is statistically significant evidence to support the belief that pMEK\_N is differentially expressed among experimental groups of mice (H(7)=18.0359, p=0.0118). The effect size for this difference is medium, meaning this difference is also practically significant (η²= 0.17244, ε²=0.25403). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has notably lower levels of expression than their S/C counterparts.

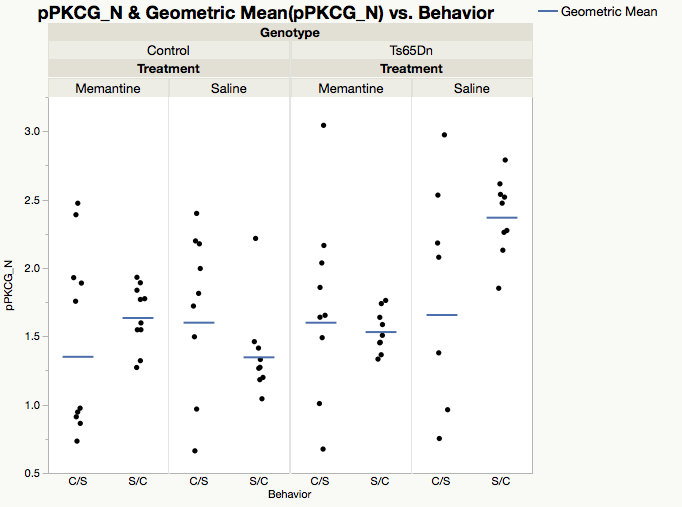


There is statistically significant evidence to support the belief that pNR1\_N is differentially expressed among experimental groups of mice (H(7)=17.428, p=0.0148). The effect size for this difference is medium, meaning this difference is also practically significant (η²= 0.16294, ε²=0.24546). Upon examination of this graph there is no discernable cause for the variation in the expression levels of this protein in these mice.

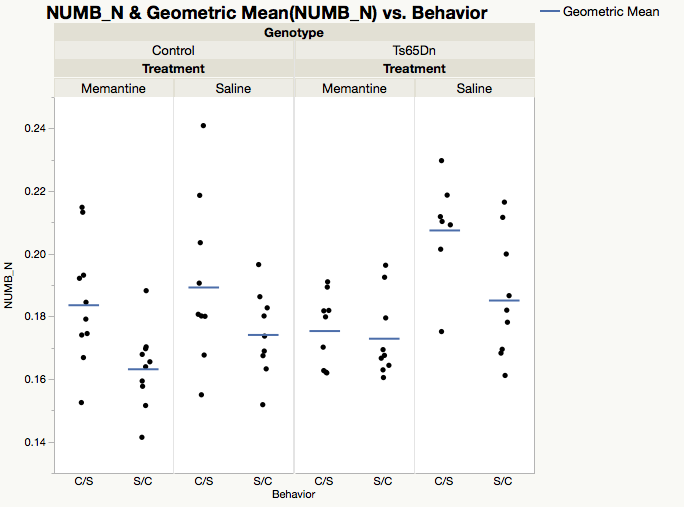
There is statistically significant evidence to support the belief that RAPTOR\_N is differentially expressed among experimental groups of mice (H(7)=20.921, p=0.0039). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.21752, ε²=0.29466). Upon examination of this graph we can see a large difference in the levels of expression of this protein in trisomy mice with the C/S group have a lower level of expression than their S/C counterparts.



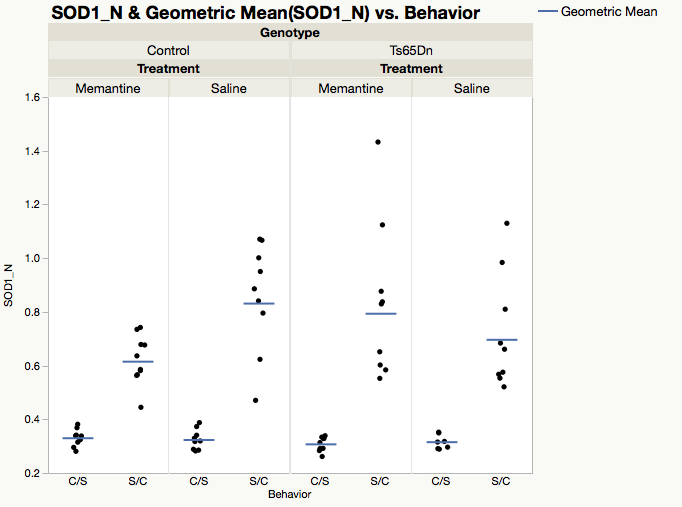
There is statistically significant evidence to support the belief that ARC\_N is differentially expressed among experimental groups of mice (H(7)=46.5048, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.61726, ε²=0.655). Evaluating the visual, there appears to be a difference in protein expression due to behavioral treatment with C/S having lower levels when compared to S/C groups.



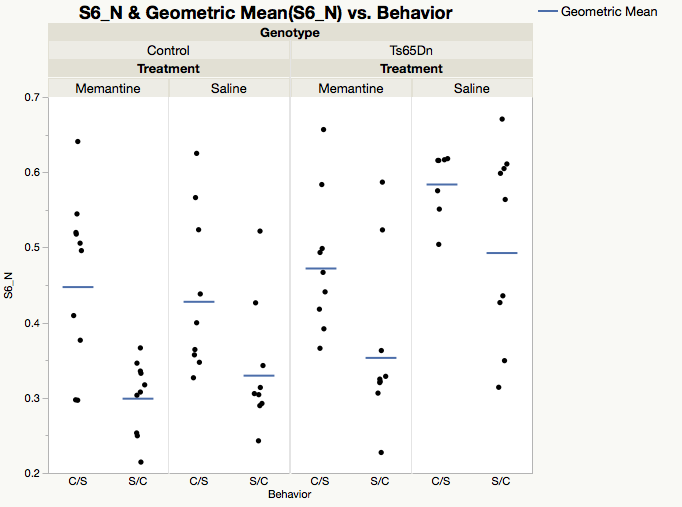
There is statistically significant evidence to support the belief that pPKCG\_N is differentially expressed among experimental groups of mice (H(7)=20.3404, p=0.0049). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.20844, ε²=0.28648). Visually a trend is difficult to discern in this data, however, there may be some difference between t-S/C-s and other experimental groups with a value of the SGM around 2.5 while all other groups have values of around 1.5.



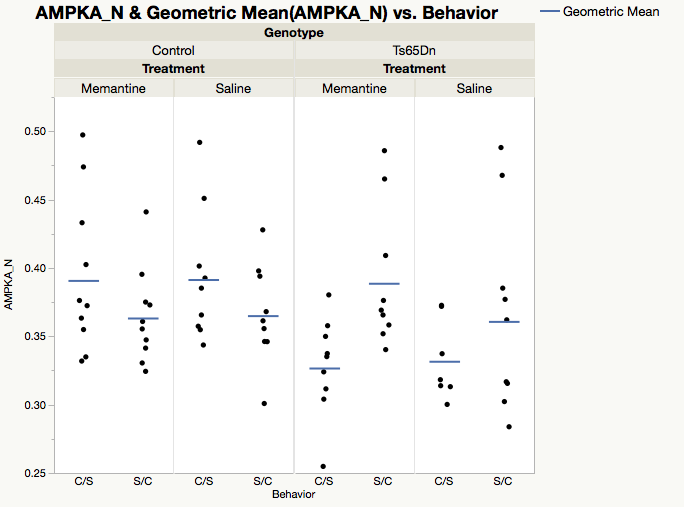
There is statistically significant evidence to support the belief that NUMB\_N is differentially expressed among experimental groups of mice (H(7)=22.2586, p=0.0023). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.23842, ε²=0.3135). Upon examination of this graph we were unable to find any discernable cause for the variation in levels of expression of this protein in these mice.

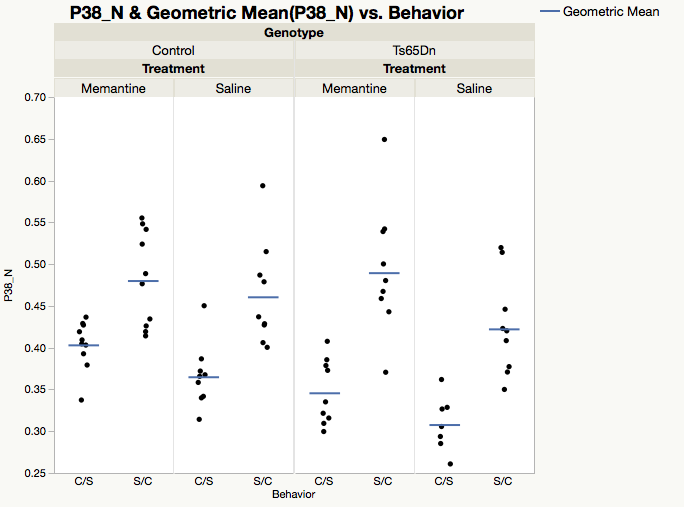


There is statistically significant evidence to support the belief that SOD1\_N is differentially expressed among experimental groups of mice (H(7)=55.4942, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.75772, ε²=0.78161). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding the genotype and drug constant we notice a large difference in the levels of expression of this protein in different behavioral groups.

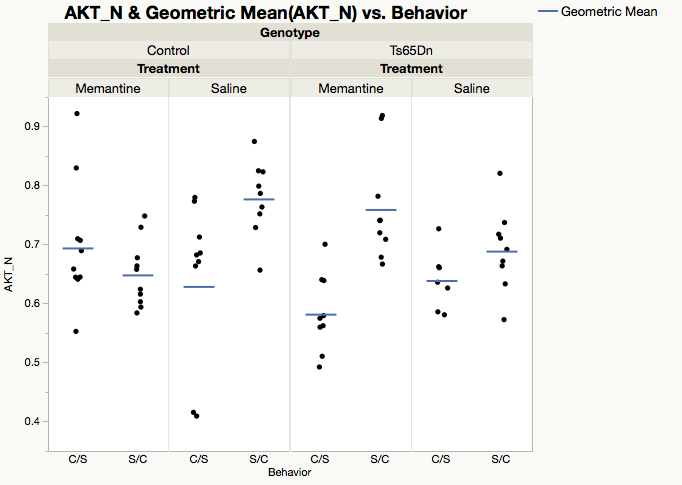


There is statistically significant evidence to support the belief that S6\_N is differentially expressed among experimental groups of mice (H(7)=34.1883M, p<0.0001). The effect size for this difference is large, meaning these results are also practically significant (η²= 0.4282, ε²=0.48152). Behavioral treatment appears to be a large factor in difference between groups, where C/S treated groups have notably higher levels of protein expression than S/C groups. However, one S/C group appears to be higher than every other C/S group with the exception of the group of mice that shared drug treatment and genotype.

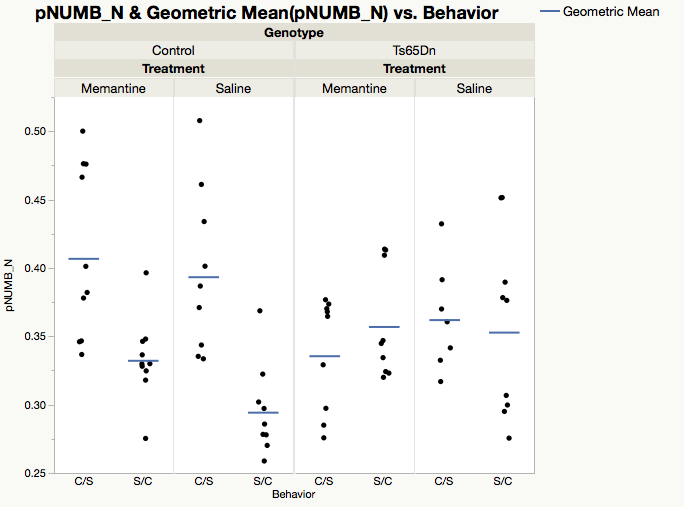


There is statistically significant evidence to support the belief that AMPKA\_N is differentially expressed among experimental groups of mice (H(7)=16.6124, p=0.0201). The effect size for this difference is medium, meaning this difference is also practically significant (η²= 0.15019, ε²=0.23398). Upon examination of this graph we are unable to find any discernable cause for the difference in variation in expression levels for this protein.

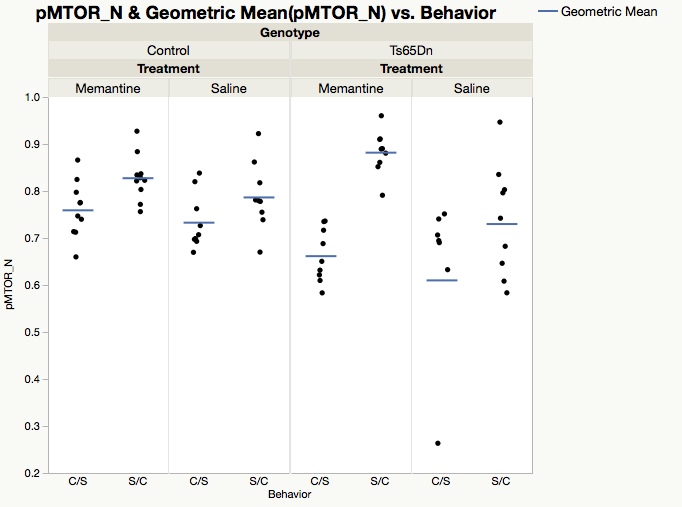
There is statistically significant evidence to support the belief that P38\_N is differentially expressed among experimental groups of mice (H(7)=46.8048, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.62195, ε²=0.65922). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. For each of these pairs, the C/S group has notably lower levels of expression than their SC counterparts.



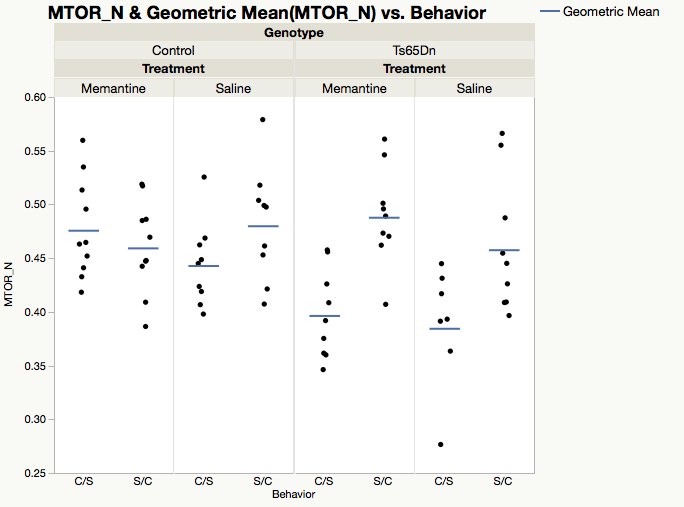
There is statistically significant evidence to support the belief that AKT\_N is differentially expressed among experimental groups of mice (H(7)=30.2377, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.36309, ε²=0.42588). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. In most cases the C/S group had lower levels of expression than the S/C group with the exception of the memantine control group.



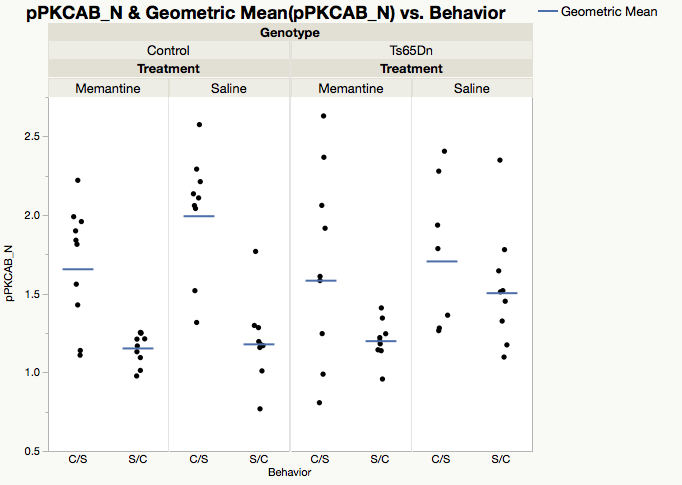
There is statistically significant evidence to support the belief that pNUMB\_N is differentially expressed among experimental groups of mice (H(7)=28.8591, p=0.0005). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.29467, ε²=0.36421). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. In the control mice the C/S groups have noticably higher levels of expression of this protein than their S/C counterparts.



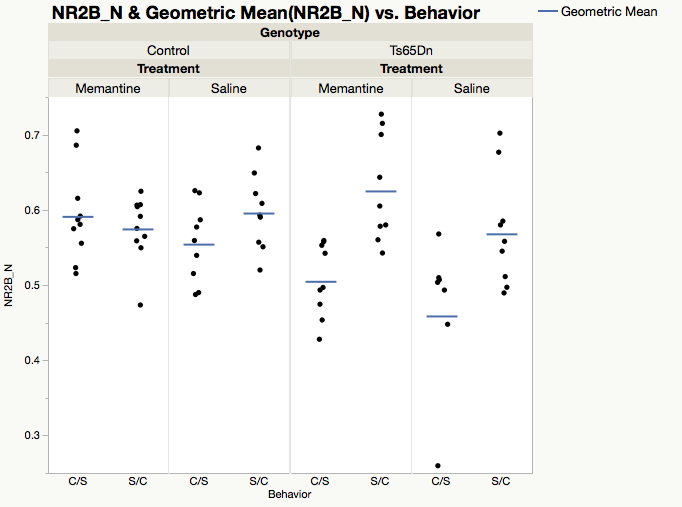
There is statistically significant evidence to support the belief that pMTOR\_N is differentially expressed among experimental groups of mice (H(7)=38.0964, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.48588, ε²=0.53567). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding genotype and drug constant we observe a difference in the levels of expression in the two behavioral groups.



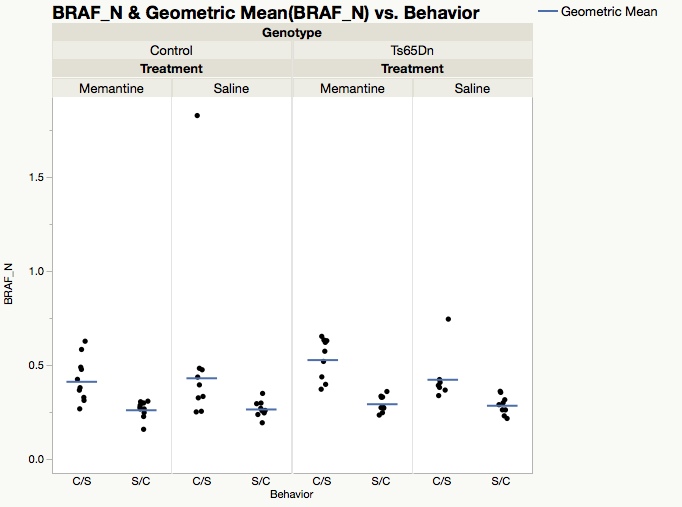
There is statistically significant evidence to support the belief that MTOR\_N is differentially expressed among experimental groups of mice (H(7)=25.698, p=0.0006). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.29216, ε²=0.36194). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. From the graph we are unable to discern a definite cause for the variation in levels of expression of this protein.



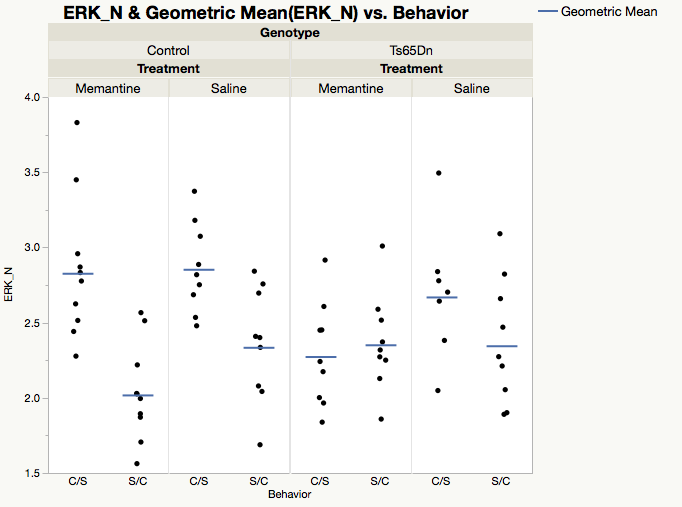
There is statistically significant evidence to support the belief that pPKCAB\_N is differentially expressed among experimental groups of mice (H(7)=31.4138, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.38147, ε²=0.44245). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. For each of these pairs, the C/S group has notably higher levels of expression than their S/C counterparts.



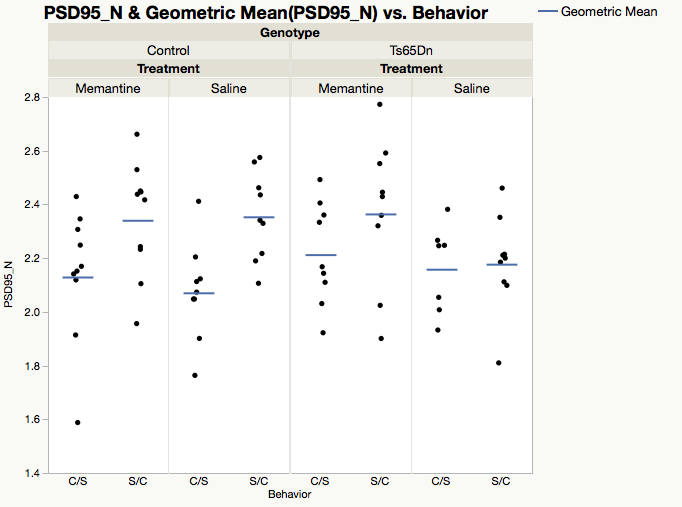
There is statistically significant evidence to support the belief that NR2B\_N is differentially expressed among experimental groups of mice (H(7)=26.1435, p=0.0005. The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.29912, ε²=0.36822). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. In the trisomy mice the C/S group has a noticeably lower level of expression than their S/C counterparts.



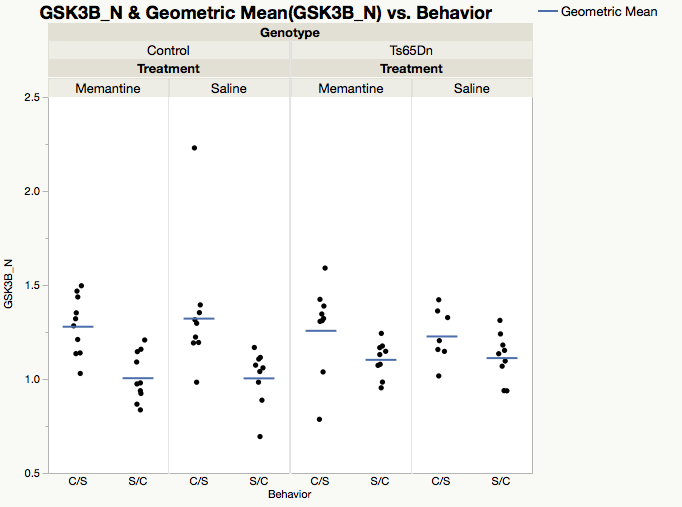
There is statistically significant evidence to support the belief that BRAF\_N is differentially expressed among experimental groups of mice (H(7)=41.8159, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.544, ε²=0.58896). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. For each of the pairs, the C/S group had higher levels of expression than their S/C counterpart.



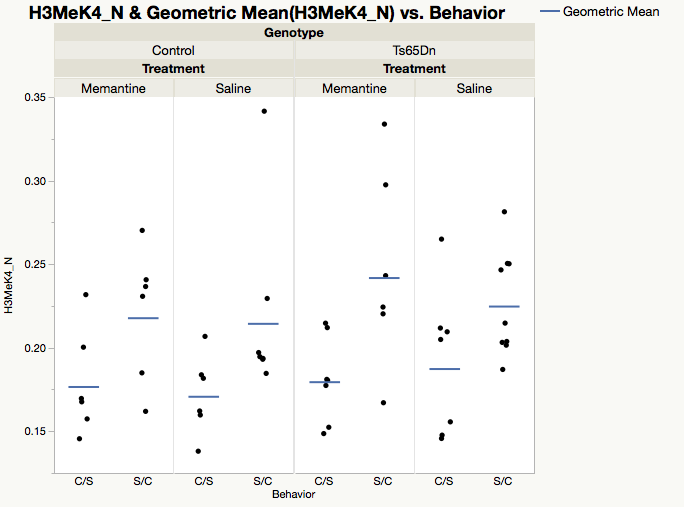
There is statistically significant evidence to support the belief that ERK\_N is differentially expressed among experimental groups of mice (H(7)=27.9094, p=0.0002). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.32671, ε²=0.39309). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. In most cases the C/S group had much higher levels of expression than their S/C counterpart with the exception of the memantine trisomy mice.



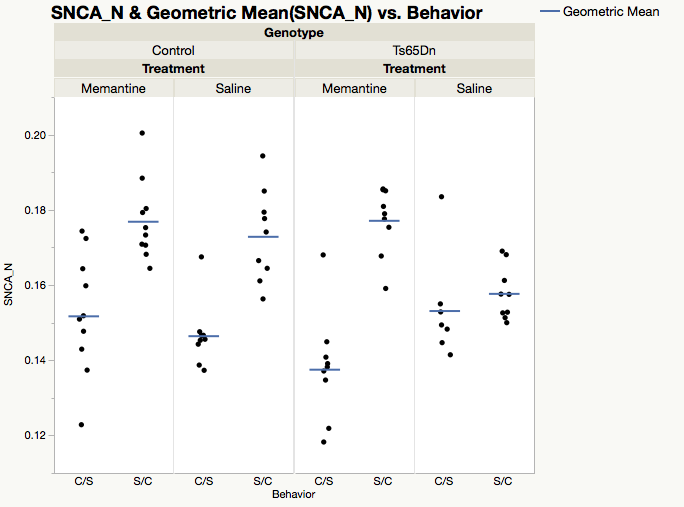
There is statistically significant evidence to support the belief that PSD95\_N is differentially expressed among experimental groups of mice (H(7)=16.3661, p=0.022). The effect size for this difference is medium, meaning this difference has some practical significance (η²= 0.14634, ε²=0.23051). Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has lower levels of expression than their S/C counterparts. An exception may be in the groups of trisomic mice treated with saline where PSD95\_N expression appears to be even with only a slight increase from C/S behavior.



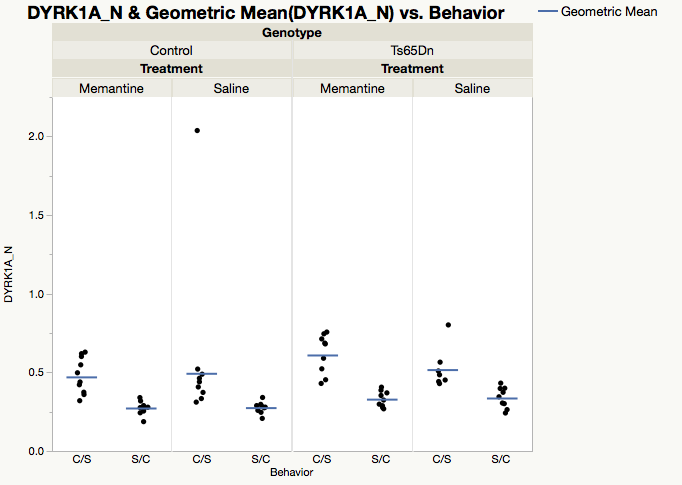
There is statistically significant evidence to support the belief that GSK3B\_N is differentially expressed among experimental groups of mice (H(7)=28.2133, p=0.0002). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.33146, ε²=0.39737). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. For each of the pairs, the C/S group has notably higher levels of expression than their S/C counterparts



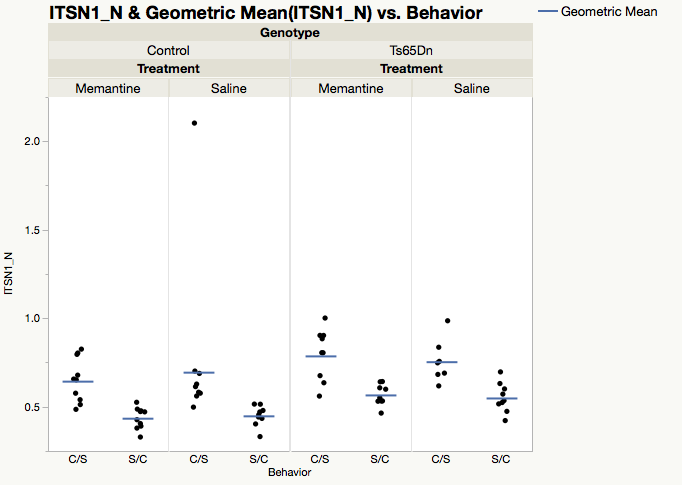
There is statistically significant evidence to support the belief that H3MeK4\_N is differentially expressed among experimental groups of mice (H(7)=17.847, p=0.0127). The effect size for this difference is medium, meaning this difference also has some practical significance (η²= 0.2358, ε²=0.33674). Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has notably lower levels of expression than their S/C counterparts.



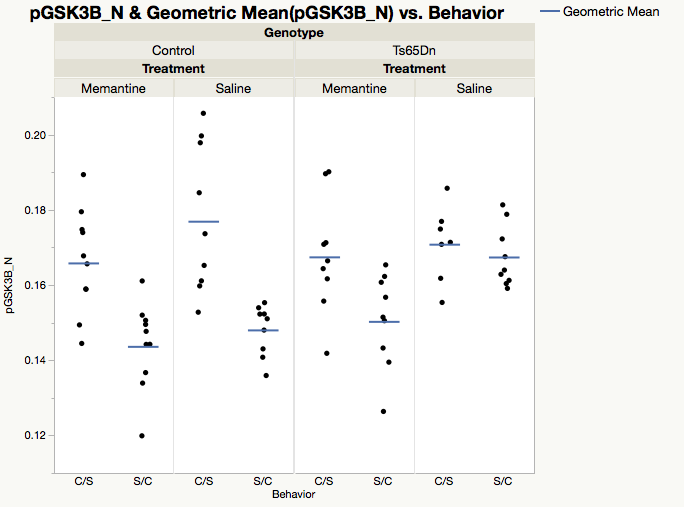
There is statistically significant evidence to support the belief that SNCA\_N is differentially expressed among experimental groups of mice (H(7)=45.7141, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.60491, ε²=0.64386). Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has notably lower levels of expression than their S/C counterparts.



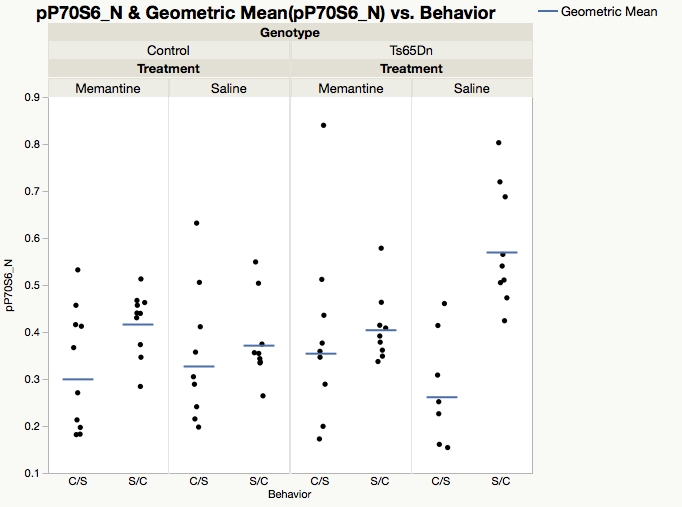
There is statistically significant evidence to support the belief that DYRK1A\_N is differentially expressed among experimental groups of mice (H(7)=50.8442, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.68507, ε²=0.71612). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. The C/S group has higher levels of expression than their SC counterparts.



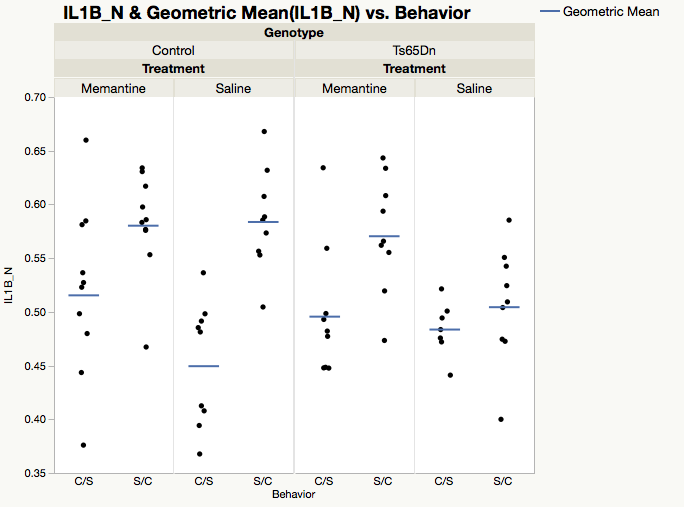
There is statistically significant evidence to support the belief that ITSN1\_N is differentially expressed among experimental groups of mice (H(7)=48.1445, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.64288, ε²=0.67809). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. The C/S group has higher levels of expression than their SC counterparts.



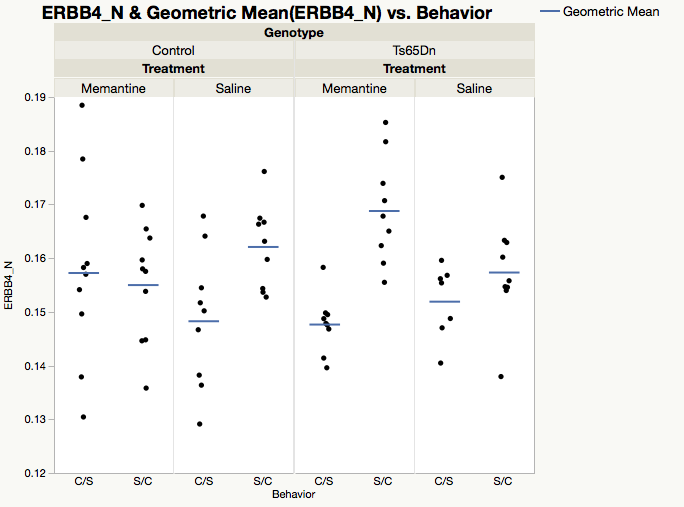
There is statistically significant evidence to support the belief that pGSK3B\_N is differentially expressed among experimental groups of mice (H(7)=37.5002, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.47657, ε²=0.52817). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding the behavior and drug constant there is a noticeable difference in the level of expression of this protein based on the genotype of the mice.



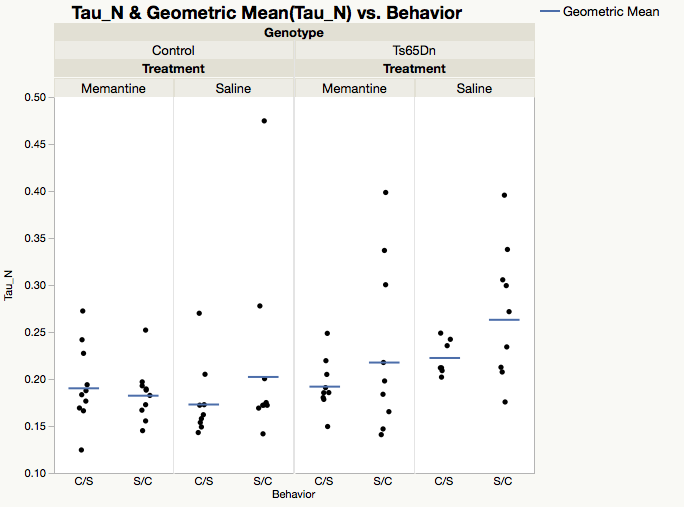
There is statistically significant evidence to support the belief that pP70S6\_N is differentially expressed among experimental groups of mice (H(7)=22.9011, p=0.0018). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.24845, ε²=0.32255). Upon examination of this graph we are unable to find any discernable cause for the variation in levels of expression of this protein in these mice.

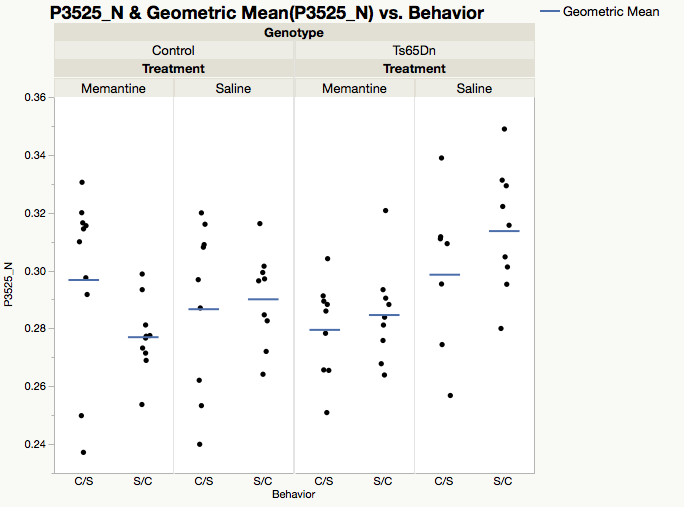


There is statistically significant evidence to support the belief that IL1B\_N is differentially expressed among experimental groups of mice (H(7)=31.0524, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.37582, ε²=0.43736). Visual analysis of the data shows there appears to be a differences in the level of expression between the two levels of behavioral treatment. C/S treatment appears to be lower in nearly every case when compared to S/C, though differences vary, which may be an indication that other interactions may be at play here as well.

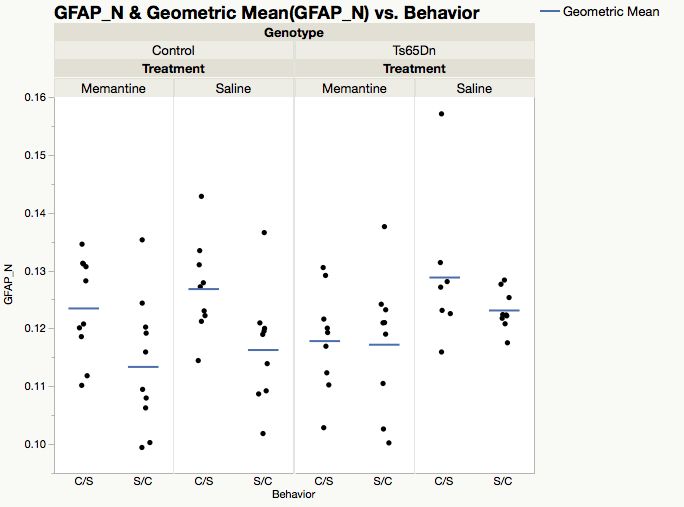


There is statistically significant evidence to support the belief that ERBB4\_N is differentially expressed among experimental groups of mice (H(7)=22.5383, p=0.0021). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.24279, ε²=0.31744). Due to relatively high variation as spread in these groups, discerning any trends between treatment effects is difficult. Using the value of the SGM indicated by the blue line for each group, there appears to be a slight difference in protein expression due to behavior.

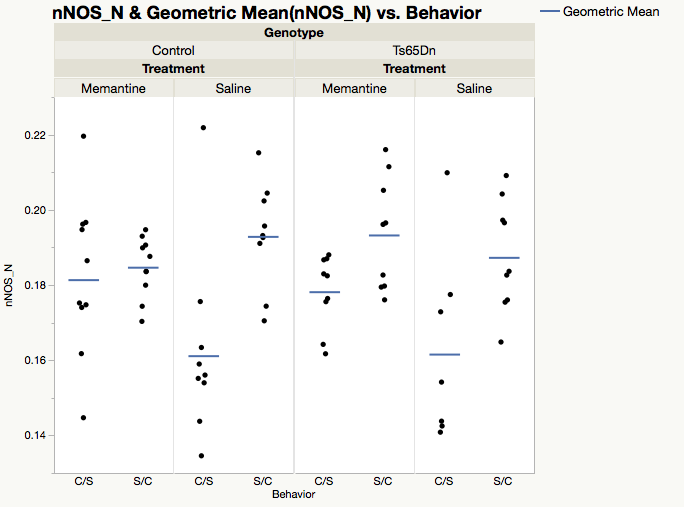


There is statistically significant evidence to support the belief that Tau\_N is differentially expressed among experimental groups of mice (H(7)=18.6315, p=0.0094). The effect size for this difference is large, meaning this difference is also of practical significance (η²= 0.18174, ε²=0.26242). Visually, the only difference between factors that can be gleaned seems to be the overall high expression in trisomic mice treated with saline, regardless of behavioral treatment. Further investigation is necessary.

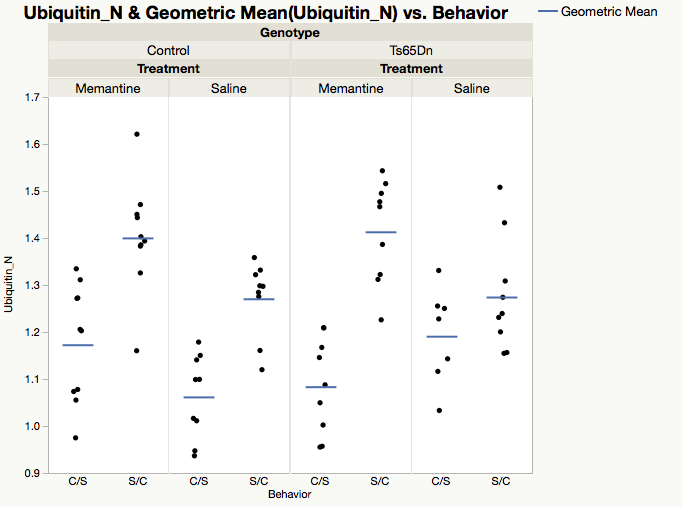
There is statistically significant evidence to support the belief that P3525\_N is differentially expressed among experimental groups of mice (H(7)=17.8817, p=0.0125). The effect size for this difference is medium, meaning this difference has some practical significance (η²= 0.17003, ε²=0.25185). Analyzing the visual, there do not appear to be any clear trends for any one factor. Interactions of effects may play a role, however, further statistical inference is necessary to shed light on this and make any statement of association with any certainty.

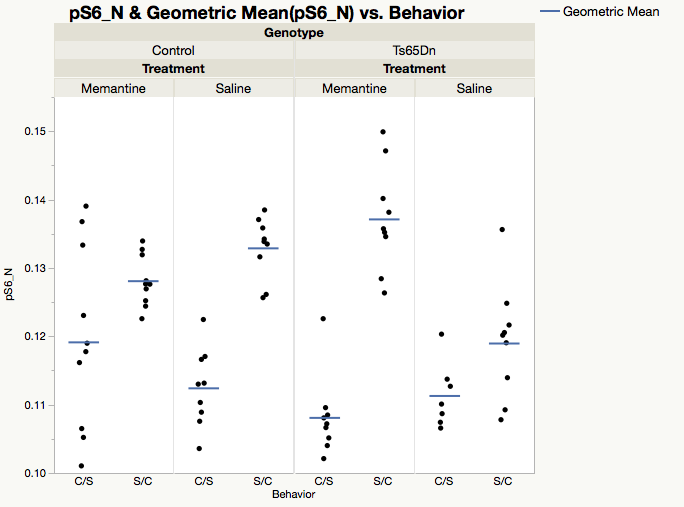


There is statistically significant evidence to support the belief that GFAP\_N is differentially expressed among experimental groups of mice (H(7)=17.3739, p=0.0151). The effect size for this difference is medium, meaning this difference may have some practical significance (η²= 0.16209, ε²=0.2447). Visual analysis of the data shows there appears to be a differences in the level of expression between the two levels of behavioral treatment. Levels of expression in C/S treatment appears to be higher in each case when compared to S/C, though differences vary, which may be an indication that other interactions may be at play here as well.

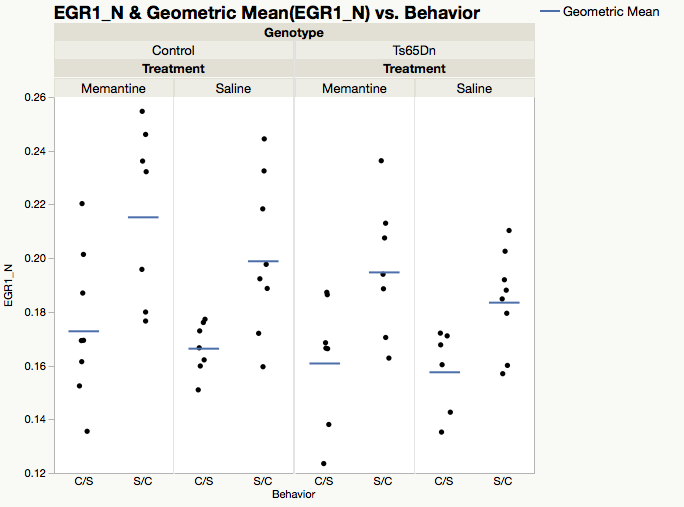


There is statistically significant evidence to support the belief that nNOS\_N is differentially expressed among experimental groups of mice (H(7)=20.9342, p=0.0039). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.21772, ε²=0.29485). Visual analysis of the data indicates C/S groups may have lower levels of expression than S/C groups, implicating behavior treatment as a large factor in difference between groups, however, more investigation is needed to state this with certainty.

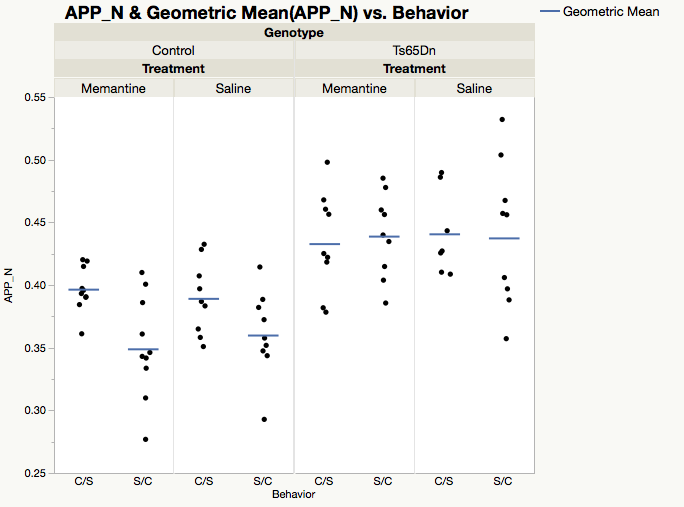


There is statistically significant evidence to support the belief that Ubiquitin\_N is differentially expressed among experimental groups of mice (H(7)=43.6074, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.57199, ε²=0.61419). Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has notably lower levels of expression than their S/C counterparts.

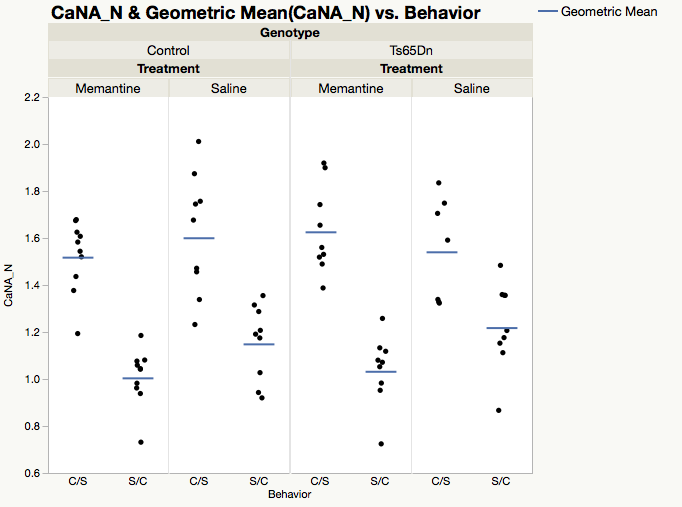
There is statistically significant evidence to support the belief that pS6\_N is differentially expressed among experimental groups of mice (H(7)=46.5048, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.0.61726, ε²=0.655). Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has notably lower levels of expression than their S/C counterparts.



There is statistically significant evidence to support the belief that EGR1\_N is differentially expressed among experimental groups of mice (H(7)=23.099, p=0.016). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.32198, ε²=0.40525). Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the C/S group has notably lower levels of expression than their S/C counterparts.



There is statistically significant evidence to support the belief that APP\_N is differentially expressed among experimental groups of mice (H(7)=36.397, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.45933, ε²=0.51263). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding behavior and drug constant we notice a large difference in the levels of expression of this protein in the different genotypes.



There is statistically significant evidence to support the belief that CaNA\_N is differentially expressed among experimental groups of mice (H(7)=50.7826, p<0.0001). The effect size for this difference is large, meaning this difference is also practically significant (η²= 0.6841, ε²=0.71525). While we were not able to use statistical inference to investigate treatment effects of each factor (Genotype, Behavior, Drug), we can analyze this graph to better understand which effects play a large role. Holding genotype and drug steady, we can see that there are apparent differences in the behavioral groups. For each of these pairs, the CS group has notably higher levels of expression than their SC counterparts.

**Conclusion/Discussion**

We were able to successfully identify 45 proteins which had statistically significant different levels of expression between the eight experimental groups. In this regard were able to reject the null hypothesis for the first statistical research question, ‘Is there a difference in the levels of protein expression between different classes of mice?’ We did this using the Kruskal-Wallis test H statistic with 7 degrees of freedom, evaluating the associated p value with an unusualness threshold of 0.025. The effect sizes η² and ε² were able to identify practical significance using the rule of thumb intervals to state, negligible, small, medium or large.

Our knowledge of statistical methods fell short when analyzing the remaining statistical research questions, as we were not able to provide results using appropriate statistical tests. However, we were able to identify factors as having potential roles in differential protein expression using visual analysis. The most prominent finding using visual analysis was a recurring disparity between groups receiving C/S treatment vs. S/C treatment, often finding one higher than the other for all pairings. This provides an interesting conclusion that should certainly be investigated further using appropriate statistical analysis such as SOM. The apparent conclusion is that many of these proteins have expression levels that are seemingly related to the behavioral treatment of the mice and could play large roles in learning. This is pertinent to the third SQR, ‘Is there a difference in the levels of protein expression in mice that have been stimulated to learn vs. mice that have not been stimulated to learn?’ The other SRQs, investigating genotype, drug, and the interactions of all three treatments, are not as easily identifiable as they do not demonstrate as grossly apparent trends as do the behavioral treatments. These will require a higher level of statistical analysis than we can provide.

**Appendix A**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Protein Name*** | ***H*** | ***p-value*** | ***η²*** | ***ε²*** |
| *DYRK1A\_N* | 50.8442 | <0.0001 | 0.68507 | 0.71612 |
| *ITSN1\_N* | 48.1445 | <0.0001 | 0.64288 | 0.67809 |
| *BDNF\_N* | 14.5585 | 0.0421 | --- | --- |
| *NR1\_N* | 13.1864 | 0.0677 | --- | --- |
| *NR2A\_N* | 20.8598 | 0.004 | 0.21656 | 0.2938 |
| *pAKT\_N* | 16.876 | 0.0182 | 0.15431 | 0.23769 |
| *pBRAF\_N* | 18.3947 | 0.0103 | 0.17804 | 0.25908 |
| *pCAMKII\_N* | 28.4872 | 0.0002 | 0.33574 | 0.40123 |
| *pCREB\_N* | 13.0141 | 0.0718 | --- | --- |
| *pELK\_N* | 14.2906 | 0.0462 | --- | --- |
| *pERK\_N* | 53.1294 | <0.0001 | 0.72077 | 0.7483 |
| *pJNK\_N* | 17.8696 | 0.0126 | 0.16984 | 0.25168 |
| *PKCA\_N* | 15.9422 | 0.0256 | --- | --- |
| *pMEK\_N* | 18.0359 | 0.0118 | 0.17244 | 0.25403 |
| *pNR1\_N* | 17.428 | 0.0148 | 0.16294 | 0.24546 |
| *pNR2A\_N* | 31.2753 | <0.0001 | 0.3793 | 0.4405 |
| *pNR2B\_N* | 14.5838 | 0.0417 | --- | --- |
| *pPKCAB\_N* | 31.4138 | <0.0001 | 0.38147 | 0.44245 |
| *pRSK\_N* | 7.3992 | 0.3885 | --- | --- |
| *AKT\_N* | 30.2377 | <0.0001 | 0.36309 | 0.42588 |
| *BRAF\_N* | 41.8159 | <0.0001 | 0.544 | 0.58896 |
| *CAMKII\_N* | 14.9845 | 0.0362 | --- | --- |
| *CREB\_N* | 11.4329 | 0.1208 | --- | --- |
| *ELK\_N* | 15.8244 | 0.268 | --- | --- |
| *ERK\_N* | 27.9094 | 0.0002 | 0.32671 | 0.39309 |
| *GSK3B\_N* | 28.2133 | 0.0002 | 0.33146 | 0.39737 |
| *JNK\_N* | 6.0746 | 0.5311 | --- | --- |
| *MEK\_N* | 7.2969 | 0.3986 | --- | --- |
| *TRKA\_N* | 7.2434 | 0.404 | --- | --- |
| *RSK\_N* | 5.8165 | 0.5613 | --- | --- |
| *APP\_N* | 36.397 | <0.001 | 0.45933 | 0.51263 |
| *Bcatenin\_N* | 15.098 | 0.0348 | --- | --- |
| *SOD1\_N* | 55.4942 | <0.001 | 0.75772 | 0.78161 |
| *MTOR\_N* | 25.698 | 0.0006 | 0.29216 | 0.36194 |
| *P38\_N* | 46.8048 | <0.0001 | 0.62195 | 0.65922 |
| *pMTOR\_N* | 38.0964 | <0.0001 | 0.48588 | 0.53567 |
| *DSCR1\_N* | 12.5838 | 0.0829 | --- | --- |
| *AMPKA\_N* | 16.6124 | 0.0201 | 0.15019 | 0.23398 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Protein Name*** | ***H*** | ***p-value*** | ***η²*** | ***ε²*** |
| *NR2B\_N* | 26.1435 | 0.0005 | 0.29912 | 0.36822 |
| *pNUMB\_N* | 28.8591 | 0.0005 | 0.29467 | 0.36421 |
| *RAPTOR\_N* | 20.921 | 0.0039 | 0.21752 | 0.29466 |
| *TIAM1\_N* | 14.4151 | 0.0443 | --- | --- |
| *pP70S6\_N* | 22.9011 | 0.0018 | 0.24845 | 0.32255 |
| *NUMB\_N* | 22.2586 | 0.0023 | 0.23842 | 0.3135 |
| *P70S6\_N* | 9.2984 | 0.2319 | 0.03591 | 0.13096 |
| *pGSK3B\_N* | 37.5002 | <0.0001 | 0.47657 | 0.52817 |
| *pPKCG\_N* | 20.3404 | 0.0049 | 0.20844 | 0.28648 |
| *CDK5\_N* | 19.3028 | 0.0073 | 0.19223 | 0.27187 |
| *S6\_N* | 34.1883 | <0.0001 | 0.42482 | 0.48152 |
| *ADARB1\_N* | 14.3697 | 0.045 | --- | --- |
| *AcetylH3K9\_N* | 12.76 | 0.0782 | --- | --- |
| *RRP1\_N* | 6.9304 | 0.4362 | --- | --- |
| *BAX\_N* | 10.3547 | 0.1693 | --- | --- |
| *ARC\_N* | 46.5048 | <0.0001 | 0.61726 | 0.655 |
| *ERBB4\_N* | 22.5383 | 0.0021 | 0.24279 | 0.31744 |
| *nNOS\_N* | 20.9342 | 0.0039 | 0.21772 | 0.29485 |
| *Tau\_N* | 18.6315 | 0.0094 | 0.18174 | 0.26242 |
| *GFAP\_N* | 17.3739 | 0.0151 | 0.16209 | 0.2447 |
| *GluR3\_N* | 11.0606 | 0.136 | --- | --- |
| *GluR4\_N* | 3.9214 | 0.7888 | --- | --- |
| *IL1B\_N* | 31.0524 | <0.0001 | 0.37582 | 0.43736 |
| *P3525\_N* | 17.8817 | 0.0125 | 0.17003 | 0.25185 |
| *pCASP9\_N* | 8.5804 | 0.2842 | --- | --- |
| *PSD95\_N* | 16.3661 | 0.022 | 0.14634 | 0.23051 |
| *SNCA\_N* | 45.7141 | <0.0001 | 0.60491 | 0.64386 |
| *Ubiquitin\_N* | 43.6074 | <0.0001 | 0.57199 | 0.61419 |
| *pGSK3B\_Tyr216\_N* | 8.3303 | 0.3044 | --- | --- |
| *SHH\_N* | 8.6839 | 0.2762 | --- | --- |
| *BAD\_N* | 14.4125 | 0.0443 | --- | --- |
| *BCL2\_N* | 6.4314 | 0.4914 | --- | --- |
| *pS6\_N* | 46.5048 | <0.0001 | 0.61726 | 0.655 |
| *pCFOS\_N* | 12.8787 | 0.0751 | --- | --- |
| *SYP\_N* | 6.9586 | 0.4332 | --- | --- |
| *H3AcK18\_N* | 13.8748 | 0.0535 | --- | --- |
| *EGR1\_N* | 23.099 | 0.0016 | 0.32198 | 0.40525 |
| *H3MeK4\_N* | 17.847 | 0.0127 | 0.2358 | 0.33674 |
| *CaNA\_N* | 50.7826 | <0.0001 | 0.6841 | 0.71525 |

1. JMP®, Version <x>. SAS Institute Inc., Cary, NC, 1989-2007. [↑](#endnote-ref-2)
2. Ahmed MM, Dhanasekaran AR, Block A, Tong S, Costa ACS, Stasko M, et al. (2015) Protein Dynamics Associated with Failed and Rescued Learning in the Ts65Dn Mouse Model of Down Syndrome. PLoS ONE 10(3): e0119491. https://doi.org/10.1371/journal.pone.0119491 [↑](#endnote-ref-3)
3. Higuera C, Gardiner KJ, Cios KJ (2015) Self-Organizing Feature Maps Identify Proteins Critical to Learning in a Mouse Model of Down Syndrome. PLoS ONE 10(6): e0129126. https://doi.org/10.1371/journal.pone.0129126 [↑](#endnote-ref-4)
4. Asuncion, A. & Newman, D.J. (2007). UCI Machine Learning Repository [http://www.ics.uci.edu/~mlearn/MLRepository.html]. Irvine, CA: University of California, School of Information and Computer Science. [↑](#endnote-ref-5)
5. Costa ACS, Scott-Mckean JJ, Stasko MR. Acute Injections of the NMDA Receptor Antagonist Memantine Rescue Performance Deficits of the Ts65Dn Mouse Model of Down Syndrome on a Fear Conditioning Test. Neuropsychopharmacology. 2007;33(7):1624-1632. doi:10.1038/sj.npp.1301535. [↑](#endnote-ref-6)