Reviewer’s report

Title: Extreme sensitivity of gene expression in human SH-SY5Y neurocytes to ultra-low doses of Gelsemium sempervirens

Version: 1 Date: 8 January 2014

Reviewer: Guy Brock

Reviewer’s report:

Overall comment: The authors investigate the dose-response of human SH-SY5Y neurocytes to Gelsemium sempervirens treatment. For the most part, the statistics are done adequately. However, several places need significant improvement, in particular the description of the Friedman/Wilcoxon analysis for the identified gene-sets. Also, the authors could improve the analysis using limma by incorporating the arrays from all of the doses in a linear model, rather than analyzing single doses at a time relative to control samples. This would substantially increase the sample size and potentially improve power, and further allow for testing a dose-response effect. Specific comments and further details are given below.

Major Revisions

1. Analysis using limma: As stated above, authors could significantly improve power by including all of the dilutions in their analysis of differential expression. A dose-response effect could then be tested by a linear model, and additional terms (quadratic, cubic) could be included to allow flexibility in the shape of the response. In lieu of this, the authors could test for an overall difference among treatment groups by treating dilution as categorical and using an ANOVA analysis (unadjusted p-values from ANOVA F-statistics will be reported in ‘F.p.value’ from the resulting limma fit). Likely the resulting list of genes will be very similar to what the authors currently found using just the 2c dilution, but expanded due to the increased sample size and power. Authors could follow-up any significant ANOVA findings with contrasts comparing each dilution with control. The blocked nature of the dilutions (i.e., each treatment dilution was compared to a corresponding control dilution) can be accounted for by treating dilution level as a blocking variable (see the ‘block’ argument in ‘lmFit’ and the ‘duplicateCorrelation’ function) or main effect (e.g., two-way ANOVA). This approach may also supplant the need for the Wilcoxon/Friedman analysis (see comments below), as the ANOVA or dose-response model would identify genes that are responding across treatment groups (dilutions).

2. Comments for ‘Statistics’ section in Methods: The authors need to explicitly state the study design here, and what hypotheses are being tested. This is more clearly stated in the Results under ‘Gene expression changes induced in SH-SY5Y cells’ but should be included here as well.
3. Comments for Wilcoxon/Friedman analysis: I have multiple comments for this analysis which are listed below:

a. The whole purpose of the Friedman / Wilcoxon test is not described and motivated properly in the ‘Statistics’ section of the Methods. After reading through all of the Results, it becomes clear that the point of this analysis is to determine whether the ‘direction’ of effect for the differentially expressed (DE) genes detected in the 2c concentration is maintained across all other dilutions. This needs to be clearly stated up-front.

b. After looking at how the data for frequency of downregulated vs upregulated genes are presented (page 14), would it not be simpler and clearer to analysis this using Fisher’s exact test?

c. The response variable for this analysis is not the same as for the limma analysis, since the authors average over the replicates. While the authors eventually state this on page 14 under the ‘Wilcoxon test for paired data’ section, it should be clearly stated in the Methods.

d. The design and hypotheses for the Friedman test need to be more clearly stated. e.g., the Friedman test is being used as a substitute for a one-way ANOVA to see if the distribution of differences is shifted from zero across all the dilutions.

e. The details concerning how the Wilcoxon test statistic and p-value (page 10) are calculated are not needed. The important aspect is to clearly state the purpose for the Wilcoxon test and what hypotheses are being tested.

f. Figure 6 can be relegated to Supplementary Material.

g. Authors should not report a p-value for the 2c concentration, as this is ‘significant’ by definition.

h. Final sentence on page 15 under ‘Wilcoxon test for paired data’ (“Due to the small sample size … ”) – it seems this would be addressed by the more comprehensive linear model detailed in Comment #1.

Minor Comments:

1. Statement concerning gene-set results in Abstract is confusing.

2. No statement concerning treatment of missing values on the NimbleGen arrays and pre-screening of transcripts is given. Did the authors really test all 45,033 transcripts? If so, this would contribute to their lack of power to detect differentially expressed genes, and the authors are advised to pre-screen invariant transcripts using e.g. the ‘genefilter’ package in Bioconductor. Even if no data imputation or pre-screening is done, authors need to include a statement concerning this.

3. Analysis of cell viability assays: Authors include a figure (Figure 3) concerning the results of the assay and state that no statistically significant differences were
found, but do not state what statistical test was used. This should be included in the Methods.

4. Pg. 13, under ‘Analysis from Gelsemium s. dilutions and controls’ – give actual numbers instead of ‘most’ and ‘some’ (“were detected in most genes of cells treated with Gelsemium s. 3c and in some genes of cells treated with … ”)

5. Cluster analysis: The cluster analysis is nicely presented, however how did the authors decide upon 5 clusters to group the data? Was some cluster validation (e.g. Brock et al. (2008) Journal of Statistical Software 25:4 and Handl et al. (2005) Bioinformatics 21(15), 3201-12) used or did the authors investigate a range of clusters and pick the best one by visual inspection? Some statement concerning this needs to be included.

6. There are several misspellings that should be corrected: “Fold Discovery Rate” (pg. 9) should be “False Discovery Rate”, “Benjamini-Hoeckberg” (caption to Table 1, pg 34) should be “Benjamini-Hochberg”

7. The commas for the decimal place on the y-axis for Figures 3 and 4 should be replaced with periods.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.