**Author’s response to reviews**

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

**Authors:**

Marta Marzotto (marta.marzotto@univr.it)
Debora Olioso (debora.olioso@univr.it)
Maurizio Brizzi (maurizio.brizzi@unibo.it)
Paola Tononi (paola.tononi@univr.it)
Mirko Cristofoletti (mirkocristofoletti@yahoo.it)
Paolo Bellavite (paolo.bellavite@univr.it)

**Version:** 3  **Date:** 6 February 2014

**Author’s response to reviews:**

We included in cover letter but for your convenience we also paste here belo:

Ref. MS: 8021020051005359 "Extreme sensitivity of gene expression in human SH-SY5Y neurocytes to ultra-low doses of Gelsemium sempervirens”

Marta Marzotto, Debora Olioso, Maurizio Brizzi, Paola Tononi, Mirco Cristofoletti, Paolo Bellavite

Reply to Reviewer’s comments

Referee #1: Iris Bell

General comments: This is an important and carefully done study on the gene expression effects of various ultra low doses of the plant Gelsemium sempervirens. The target cell genes were in neuroblastoma cells. The question posed by the authors is well defined. The methods are appropriate and well described. The data appear sound and carefully presented. The manuscript adheres to relevant standards for reporting. The discussion and conclusions are well balanced and supported by the data and the contextual literature cited, with the potential enhancement as listed below. The authors acknowledge key work on which they built their study. The title and abstract convey what has been found in an accurate way. The writing is acceptable with the exception of the need for paragraphs that are much shorter and focused on relevant points within a heading or subheading.

R. The authors thanks the Referee for this general comment. We reduced the length of paragraphs and used suitable sub-headings

“The primary compulsory recommendation is that the authors need to put in paragraphs throughout the text. This is already a dense paper with many essential details, but the lack of visible paragraphs makes it harder to read. Perhaps they were once there, but in the version sent this reviewer, they were
not visible.
The authors ensure the presence of visible paragraphs in the manuscript layout”.
R. The text was divided in clear paragraphs and shortened without losing essential details.

“Discretionary revisions: The points below will hopefully enhance the paper in context of the current literature.

1. ...it would be valuable for the authors to write a clear summary statement in the Methods section for the reader less familiar with the issues. … It would help for them to add a brief comment to explain…”

R. Thanks to this suggestion, we added at the beginning of methods the following paragraph: “The homeopathic dilutions/dynamizations were prepared in a manner comparable to methods used by commercial manufacturers, i.e. using 30% ethanol for all dilution/succussion steps. Since ethanol at higher concentration may be toxic for cells, last dilution/succussion was made in pure water. The detailed procedure was carefully repeated in all experiments and precisely reported below, since it is relevant as basic science research on homeopathic medicine progresses.”

“2. In the places of their Discussion where they mention nanoparticles, …they can cite other studies related to this type of point, including one on nano-hypericum (a different herbal source but one in which doses were reduced simply by creating a nano vs bulk form and testing anti-anxiety effects [1]).

R: This has been done in the following text in the discussion: “As an example, Prakash and colleagues [78] compared in model animals the anti-anxiety effects of hypericum prepared as gold nanoparticles versus a bulk form and observed more significant effects with the nano-hypericum, even at a 10-fold lower dose.”

3. Again, to support the plausibility that homeopathic Gelsemium in the potencies tested could contain crudely formed nanoparticles, they could cite some additional papers. Etc.

R. The authors thanks the Reviewer for the interesting comments and inserted the cited suggested references in the Discussion section as follow: “Recent evidence supports the plausibility that homeopathic Gelsemium s. in the potencies tested could contain crudely formed nanoparticles. Bel-Haaj et al. [75] demonstrated that just extended ultrasonication of plant starch can create starch nanoparticles in water. Moreover, electron microscopic evidence of nanoparticles has been obtained in several different plants prepared homeopathically [76]. Gelsemium mother tincture itself, like many other plant extracts, can biosynthesize nanoparticles of silver metal from precursor substrate [77]. Nanoparticles have unique biological and physicochemical properties, including increased catalytic reactivity, protein and DNA adsorption, bioavailability, dose-sparing, electromagnetic, and quantum effects that are different from those of bulk-form materials [23].”

“4. The authors are quite correct in emphasizing possible mechanisms such as
stochastic resonance and time-dependent sensitization in low dose amplification within the complex adaptive system (cell or organism). They may also want to acknowledge the possible involvement of hormesis as an adaptive response process in generating effects that might modulate gene expression.[8-10]"

R. Ok, this has been done in the Discussion section as follow: “This evidence is in agreement with recent hypotheses explaining the homeopathic effects (in the range of very low doses) in the framework of hormesis, where substance which are toxic at high doses turn into therapeutic when diluted to low and ultra-low doses [61]. According to the hermetic theory, ultra-diluted drugs and nanoparticles will act as low-dose stress conditions that could possibly evoke an adaptive response process producing effects that might modulate gene expression [62-64].”

Referee# 2: Leoni V Bonamin

This Reviewer asked only minor essential revisions. We thanks the Referee to have rated very positively the level of interest of our study.

“Abstract: Review "pharmacological activity" to UHD substances? What about the effects observed in dilutions above 2c? Could they be called as "pharmacological effects"? It must be discriminated in the abstract. ”

R. We agree with this observation and to avoid misunderstandings we have removed the adjective “Pharmacological” from the Abstract.

“Background: page 4 – it is important to mention the consensus about the use of succussed vehicle as control. However, this is still a controversy theme, since there are some articles suggesting unspecific effects of this kind of preparation. This particularity must be mentioned.”

R. Thanks to this suggestion, we mentioned this particularity in the “Background” section as follows: “the recent consensus recommendation among researchers in this field is for protocols that use the diluted and succussed vehicle solution as a control, however this is still a debated theme and has been done only in few cases [24,25]. “

“Please, explain in page 11, if the usefulness of UV spectra analysis for each dilution was only focused to the measurement of a “molecular marker” for lower potencies. How to interpret these data in face to the biological effects?”

R. OK, to follow this indication, in the beginning of results we clarified this point: “The spectrum of the lowest dilution (1c) was considered as marker for the actual presence of plant extract. Spectra of the subsequent dilutions checked the effectiveness of the 100x dilution steps, i.e. verified that a) the lower dilutions (from pharmaceutical factory and prepared in the laboratory) were comparable and b) the provided higher potencies were effectively diluted. “

“Results: Exploring the Table 3 more accurately, the authors could found some genes with an interesting pattern of expression in function of Gelsemium dilution. For instance, the EN 2 gene that is under expressed in treated cells exhibits a bell-shaped curve and the ALPK3 gene that is over expressed exhibits an
inverse bell-shaped curve (see illustrative figures in the attached text). This pattern can be seen in many other genes shown in this table, and is evident in Figure 8 and in the cluster analysis. In discussion, authors hypothesize that these genes are very sensible and able to be modulated by few molecules. However, it seems that after 9c, another wave of expressions or no-expressions is recovered, similarly to that observed in 2c dilution. Maybe testing even higher dilutions, such as 100 or 200c, the bell-shaped curve could be more evident and, thus, the hypothesis of “ultra-sensible genes” could be checked.”

R. Yes, this can be true and could be further studied. This suggestion was added in the Discussion (page 22): “Exploring results accurately, some genes show an interesting pattern of expression in function of Gelsemium dilution. For instance, the EN 2 gene that is under expressed in treated cells exhibits a bell-shaped curve and the ALPK3 gene that is over expressed exhibits an inverse bell-shaped curve. This pattern can be seen in many other genes in the cluster analysis. Moreover, it seems that after 9c, another wave of expressions or no-expressions is recovered. Maybe future testing even higher dilutions, such as 100c or 200c, the bell-shaped curve could be more evident and, thus, the hypothesis of “ultra-sensible genes” could be checked.”

“In page 21, discussion, an interesting approach is done regarding the putative mechanisms involved with high dilutions effects. The physical theories, nanostructures knowledge and the microarray results can lead to important hypothesis toward the mechanism of action of these preparations. Nevertheless, one must consider that they still are hypothesis to be demonstrated. In this sense, the title “Biological mechanisms” must be changed to “Hypothesis of biological mechanisms”.”

R. Yes, this has been corrected, indicating that these are “hypotheses”.

Referee# 3: Tadahiro Numakawa

“Comment 1. Fold change in the mRNAs levels regulated by extract of Gelsemium sempervirens are very small (|log2 fold change| is less than 0.8). It is very important to determine whether proteins encoded by the affected genes are similarly changed by high dose (2c) of the extract.”

R. The goal of this study was to assess the capacity of the dilution to perturb cell activity using a whole genomic approach as a first screening. Since we tested very low doses of compounds (even in 2c), we hypothesized to have very small changes and choose to investigate gene expression changes, the analytical approach that was putatively more sensible on the basis of our experience and biotechnological literature. We agree with the observation of the Reviewer and acknowledged it in the text of Results: “In general, mean fold changes in the mRNAs levels of cells treated with Gelsemium s. were small and only 4 genes showed |log2 fold change| > 0.8” We addressed this point also in the discussion including the paragraph: “In general, mean fold changes in the mRNAs levels of cells treated with Gelsemium s. were small and only 4 genes showed |log2 fold change| > 0.8”.” We also agree that it would be important to check if proteins are similarly changed, although, it is known that the sensitivity of Western blot is
small. To address this point raised by the Reviewer, we added a paragraph to the Discussion: “These microarray findings can be regarded as a preliminary screening of the sensitivity of SH-SY5Y cellular system to Gelsemium s., while more robust conclusions about the possible role of the implicated genes will require to determine whether proteins encoded by the affected genes are similarly changed, through proteomic and phosphoproteomic approaches, and/or further studies using plant purified active compounds.”

“Comment 2. It is also important whether Gelsemium sempervirens-induced changes in the mRNA expressions have physiological significance or not. Although the extract did not affect viability of SH-SY5Y cells in the basal condition, it is possible that this extract elicit a protective effect against cell-death inducing conditions such as serum deprivation or oxidative stress.”

R. The possible physiological significance of our results with 2c dilution is extensively discussed in the manuscript. We agree with the Reviewer that the small changes induced by high dilutions are debatable, and we recognized it in Discussion: “The physiological or pharmacological implications of this observation remain to be clarified, but the rejection of the null hypothesis furnishes a new input for the open debate on this kind of therapeutic approach.” We agree that possibly this extract could elicit a protective effect against cell-death inducing conditions, but our experimental conditions were designed through preliminary experiments to keep the cells fully viable for the time of the experiment (24 h) and this was confirmed by the results reported in figure 2. In further studies it could be assessed the effect of Gelsemium in more stressful, cell-death inducing, conditions.

“Comment 3. In association with Comment 2, Gelsemium may change neuronal response to a transmitter analog carbachol used in this study because genes encoding transcriptional factors and G-protein coupled receptors were affected. Ca2+ response to carbachol in SH-SY5Y cells after the highest dose of extract treatment (2c) should be determined.”

R. We thank the Reviewer for this suggestions. In this preliminary investigation, carbachol was used to demonstrate that the cells are functionally responsive as neurons. As shown in figure 1A, the fluorescent-based plate reader method that we applied was sufficient to demonstrate a neuronal response, but traces were too disturbed by noise, when applied to a long kinetic measurement and were not suitable to estimate small changes induced by Gelsemium. Anyway, our further microarray results (manuscript in preparation) confirmed the action of carbachol on gene expression in these cells and an increased effect of Gelsemium 2c in stimulated cells than in resting ones.

“Comment 4. In the discussion section, authors mentioned that the modulation of gene expression by ultra-low dose (30c) of Gelsemium could be a nonlinear association between extract dose and gene expression. Although it is difficult to demonstrate chaotic effect in the regulation of gene expression by ultra-low dose of the extract, linearity of the effect of the extract in higher doses can be determined. To confirm this, expression levels of two or three mRNA expressions affected by Gelsemium (2c) should also be determined by RT-qPCR method in
10-3, 10-5 and 3c dilutions.”

R. Actually the linearity of dose-dependence of the Gelsemium effects changed according to the sub-groups of genes considered (see cluster analysis). Considering the whole down-regulated gene-set, there was an apparent relation between dilution and decrease of activity [figure 7], but since each dilution step corresponds to a 102 fold dilution, the highly diluted samples showed an activity clearly independent on the dose. We agree with the reviewer that possibly chaotic dynamics are involved in this phenomenon and that this may preclude linearity of effects, as we also suggested in papers cited in the Discussion. We performed RT-PCR on some genes in Gels 3c treated samples and this confirmed qualitatively the data of Gels 2c. E.g. three genes DDI1, PRSS54 and TAC4 showed negative fold changes compared to the corresponding control (DDI1 log fc -0.33, PRSS54 log fc -0.01, -0.14 and TAC4 log fc -0.21, -0.54). We did not report the data in the manuscript since, again, those small changes did not allow statistic confirmation. These evidences prompted us to adopt Friedman test and Fisher exact test (see response to Reviewer n. 4) to analyze data of the 49 selected genes considered as statistical unit to assess the trend of increasing dilutions (negative fold changes).

Referee# 4: Guy Brock

“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).”

R. We thanks the Referee for this expert commentary that we addressed carefully with the goal to improve our manuscript. Our experiments had a dose response setup, but we expected that the effect decrease rapidly with increasing dilution. So we have focused the aim of this report, that was to select and analyze the most soundly changing genes in the highest dose (Gels 2c) by applying restrictive parameters to the DEGs, such as FDR-adjusted p values and a cutoff on the fold change values. For this reason we used a direct pairwise comparison between the drug potencies and the control (notably the mean of
controls). For this reason, we added to the “Methods” an initial paragraph to explain the overall conceptual and statistical approach (see point n. 2).

Anyway, as suggested by the Referee we reconsidered the results from the analysis of the overall difference among treatment groups (unadjusted and BH-adjusted ‘F.p.values’), obtained as additional information in our limma R package. The results based on the F-test on our data is that the number of significantly DE genes, with adj.p.values < 0.05, dropped to 2 (instead of 56 considering limma adj.p. val. < 0.05 in the pairwise comparison of the G2c sample vs the average of controls), so this approach seemed to drop the sensitivity of the analysis to a unacceptable level. Moreover, no significant genes (with adjusted statistic p<0.05) were found in higher dilutions, neither observing the F.p.value nor observing the limma p.values of pairwise comparison. This would contrast with the evident changes induced by Gelsemium and reported in table 1 and figures 4 and 5. This because the expression changes were very small (even if significant) and the FDR correction erased all these variations as null. So in this first report we choose to discuss only the undisputable changes found in G2c treatment and then to follow-up their behavior in highest 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.”

R. As suggested by the Referee, we revised the first paragraph “Statistics” of the Method section as follow (page 10-11): “The experimental model had dose-response setup, including 6 dilutions of Gelsemium s. and 6 corresponding controls. The main working variable was the Log2-transformed fluorescence value of microarray analysis of gene expression. Data from 4 independent experiments were considered. Expecting effects to diminish with increasing dilution, we focused to a pair-wise comparison between Gelsemium s. dilutions and the vehicle controls instead of an overall comparison analysis. Two consecutive statistical approaches were followed. The first approach analyzed the complete transcriptome dataset and was aimed to select the DEGs that were most significantly affected by treatment at the highest dose; linear model (Limma) was applied to compare the expression values from Gelsemium 2c treated and the mean of controls (see details below). The second approach analyzed only the expression values of the selected DEGs when treated with highly diluted drugs or the corresponding controls. The main focus was to verify the null hypothesis that the higher Gelsemium s. dilutions did not affect the expression of the genes compared to control. For this analysis we used Friedman test as nonparametric ANOVA and Fisher’s exact test (see details below). The tests analyzed the distributions of the fold changes in the down- or up-regulated DEGs and determined whether the direction of effect for the DEGs detected in the 2c concentration was maintained across all other dilutions (3c-30c).”

We believe that after this clarification the whole approach is more clear and Statistical methods are better justified.

“3. Comments for Wilcoxon/Friedman analysis: I have multiple comments for this
analysis which are listed below:

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.”

R. Thanks for this suggestion, we explained the purpose of the tests in the first paragraph of the Statistic paragraph (see point above).

“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?”

R. Yes, we agree that the Fisher exact test is a valid alternative to evaluate the significance of the difference between positive and negative fold changes in all the dilutions. Actually, as suggested by the Referee we applied the Fisher exact test to our data. It is reassuring that Fisher test analysis provided results similar to Wilcoxon.

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.”

R. We clarified this point in the revised version of the Statistic paragraph: “In the second part of the analysis, the significant DEGs in 2c treatment were divided in two groups (considered as gene-sets) according to their direction of change, including down- and up-regulated genes; data referred to the same dilution (from 2c to 30c of both treatments and respective controls) were joined, treating the single gene as a statistical unit and the mean of four replications as the corresponding datum.”

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.”

R. The design was explained in the first paragraphs of Statistics (see above). The hypotheses for the Friedman test was explained as follows: “The Friedman test is a nonparametric test for multiple related samples (in our case, the expression level of multiple matched samples from cells treated with 6 Gelsemium s. dilutions or 6 control solutions) that checks the null hypothesis that multiple ordinal responses come from the same population.”

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.”

R. As reported above, we used Fisher exact test instead of Wilcoxon. The former
was detailed as follows: “Following a significant result of Friedman test, frequency of down-regulated vs upregulated genes were calculated; |log2 fold change| lower than or equal to 0.05 were considered to be null. The significance of distributions for each dilution was analyzed by the Fisher’s exact test, which calculates the exact probability of getting, only by chance, the observed values or more extreme ones.”

“f. Figure 6 can be relegated to Supplementary Material.”
R. OK, this has been done.

g. Authors should not report a p-value for the 2c concentration, as this is ‘significant’ by definition.”
R. We accepted this request of the Reviewer and removed the p value for the 2c dilution from text and figures.

“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.”
R. Thanks to this suggestion, we have substituted Wilcoxon with Fisher, in any case we removed the unnecessary and potentially confounding indicated sentence from Results.

“Minor Comments: 1. Statement concerning gene-set results in Abstract is confusing.”
R. We thanks the Reviewer for this observation regarding the abstract. To clarify the point, in the revised version we wrote: “Fisher exact test, applied to the group of 49 genes down-regulated by Gelsemium s. 2c, showed that the direction of effects was significantly maintained across the treatment with high homeopathic dilutions ”

“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.”
R. As suggested we added the following sentence in the methods (page 10): “No pre-filtering to the dataset (variant-based or minimal expression-based) was applied to avoid a-priori loss of results when studying minimal doses of drug.”
Indeed, we also tried to pre-screen the dataset, by applying the p-detection criteria (Anais software); we obtained a reduced dataset (about 33000 transcript) but this seems not to improve the power to detect differentially expressed genes after FDR correction.

“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.”

R. As suggested, we included the statistical test at the end of the section: “Results of viability assay were analyzed by ANOVA and t-test comparing data from each Gelsemium s. dilution (G2c,G3c, G4c, G5c, G9c and G30c) with the corresponding controls (n=12 replicates for each group).”

“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 … ”)

R. We agree with this precise note of the Reviewer and removed the generic terms. The specific paragraph in the revised document is the following: “Inspection of data reported in Table 3, concerning the expression profiles of the 56 DEGs (49 down-regulated and 7 up-regulated by Gelsemium s. 2c), highlights small expression changes (i.e. |log2 fold change| from 0.05 to 0.6) in 52, 48, 39, 36 and 48 genes of cells treated with Gelsemium s. 3c, 4c, 5c, 9c and 30c respectively.”

“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.”

R. Our analysis is essentially descriptive and we did not make statistical inferences. On the other hand, we did’nt simply pick the best clusters by visual inspection while consulted the “figure of merit”. In order to clarify this aspect, we included a sentence as follows: “The application “Figure of merit” (FOM) was used to set the number of clusters that best fit the dataset variability [44]. The FOM measures the average intra-cluster variance of the observations, estimating the mean error using predictions based on the cluster averages [45].”

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”

R. The requested changes have been done and spelling was re-checked

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

R. OK, done