UPDATED REPORT ON INCONSISTENCIES AND DISAGREEMENT IN EXPERIMENTAL RESULTS AND DATA ANALYSIS IN THE DEPARTMENT OF BIOMEDICAL AND NEUROMOTOR SCIENCES, UNIVERSITY OF BOLOGNA



The following report was created in order to express the doubts concerning the to-be-published paper by Dr Claudia Fuchs et al. as well as other concerns regarding the results obtained in the laboratory in the last 2 years. The authors of the report, as the co-authors of the manuscripts mentioned below, are concerned that some results presented there may be based on unconfirmed hypotheses and on data which were not correctly acquired and/ or analysed and/ or presented.

It should be underlined that in the past Dr	and	communicated
orally their doubts to Prof. Ciani, however	no action was taken so far.	

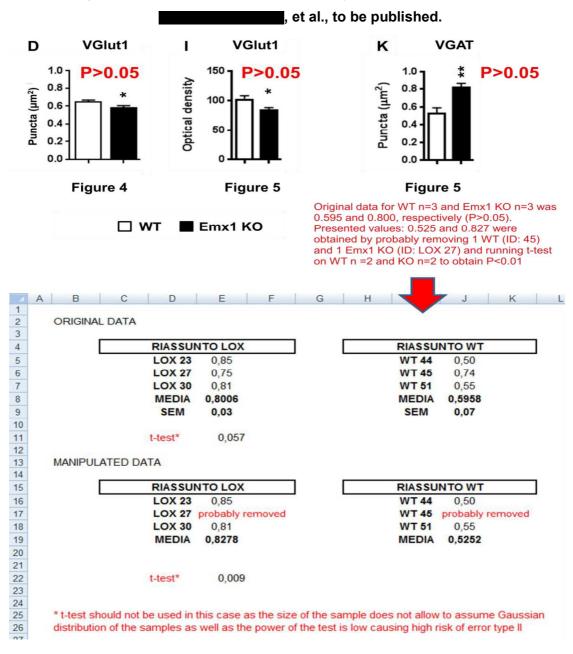
Since the research of the lab is funded and enabled by the CDKL5 children parents' associations, Telethon foundation, external grants, and the University of Bologna, the authors are convinced that it should be as much transparent as possible, leaving no space for unconfirmed results. Since the lab members have no free access to the experimental data from other researchers of the group, it creates space for potential undetected errors, data manipulation etc. That is why the authors strongly believe that the highlighted doubts should not be left undisputed as well as all the raw data from mentioned and future experiments should be presented to other lab members, without any restrictions.

Additionally, because of a very specific profile of patients the lab's research is targeting, all the group members and donors must be absolutely sure that nothing is going to be published that potentially could lead to an ineffective or even harmful therapy. As there are several concerns regarding datasets presented in the to-be-published paper of Fuchs et al., the authors of the report postulate to repeat all the presented experiments prior to the publication.

Below there is a list of doubts and inconsistencies regarding the research work of the lab. As it also covers results that were already published, the authors insist that all these discrepancies are thoroughly discussed in order to detect potential errors and, if it is a case, to correct them as soon as possible.

1. <u>Doubts about the paper: Loss of cdkl5 in forebrain excitatory neurons impairs hippocampal function (et al. to be published)</u>

The final version of the paper sent to Dr. prior to publication contained 3 graphs where non-significant differences between groups were marked as significant (Fig. 4D, 5I, 5K). In the Materials and Methods the level of significance was set for p<0.05. In the mentioned cases the p value for difference between means was higher than 0.05 but still was marked as significant (with asterisks). In one case (Fig. 5K) re-analysis of raw data done by Dr. revealed that the graph represents values as if 2 animals out of 6 were removed manually in order to obtain significant effect. Such manipulation was not mentioned in Materials and methods and caused decrease of the sample size to N=2 per group. Two of mentioned graphs were presented on the poster on the international conference. The graphs were crucial for the main article hypothesis.



2. General doubts about the main hypothesis studied in the laboratory:

There is not enough proof that P-GSK3β protein levels in the P45 Cdkl5 KO mice is lower than for P45 WT mice. On the contrary, there are data suggesting that it is equal or even increased in Cdkl5 KO mice compared to WTs (in contrast to what was published so far by the laboratory). The list of experiments supporting this statement is as follows:

- Phosphoarray on GSK3β (2 independent assays).
- Western blots on GSK3β protein.
- Preliminary results from other research group which show toxicity of the GSK3β inhibitor in the same strain of WTs (Results of Prof. Bartesaghi).
- Corn oil (vehicle) itself proved to have a small positive effect itself (10% on dendrites) however necessary vehicle controls were omitted in the presented study.
- Moreover, it is biologically unlikely that reported phosphorylation levels of GSK3β change drastically as shown, between age of P75 and P90 in the Cdkl5 KO mice.

Presented doubts refer to paragraphs 2.1., 2.2., 2.3 and 2.4. mentioned below.

2.1 <u>Doubts about the paper: HDAC4: a key factor underlying brain developmental alterations in CDKL5 disorder (Trazzi et al. 2016)</u>

General doubt:

It is uncertain that the statistical analyses were performed correctly i.e. the data was checked whether they meet the assumptions of the parametrical tests used (e.g., normal distribution, equality of variance between compared groups). This doubt was already raised by the reviewers in a revision of the manuscript that, before being published in Hum Mol gen, had been previously sent to Acta Neuropathologica (and rejected). Dr. _______, as a co-author of discussed manuscript, was refused to be shown raw data for the control experiments presented in the paper on the figure 10. All detected discrepancies and severity of doubts and potential errors heavily influence main hypotheses of the article.

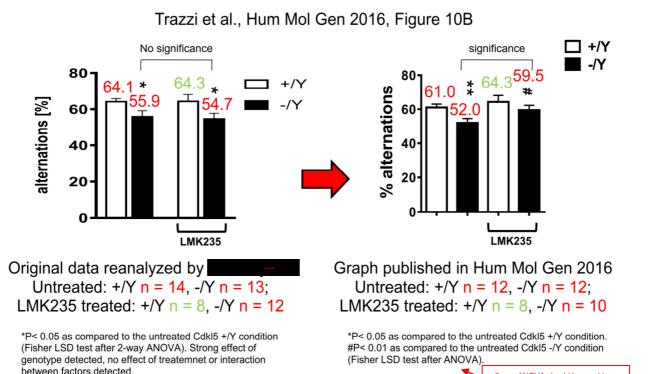
Specific doubts:

Y maze (Fig. 10 A-B):

Presented Y maze data do not seem to match the raw data in several aspects.

• LMK 235 treated -/Y and +/Y animals were tested in September and November 2015 in separate experiments and pooled for final analysis. In the Materials and Methods section it is stated that there were 10 LMK 235 treated -/Y (12 in raw data) and 8 LMK 235 treated +/Y (8 in raw data). It is not explained why and which animals were removed and/or added to the experiment. This experiment was about to prove that treatment with LMK235 restores memory in Y-maze test. However, re-analysis of raw data by Dr.

from the final analysis without an explanation what allowed for obtaining significant differences between untreated and treated -/Y groups:



Numbers of animals presented in the published graph do not correspond to raw data (except of +/Y group where n = 8 - marked in green). Some animals were removed from the analysis in order to obtain a significant result. It was not explained in Materials and Methods.

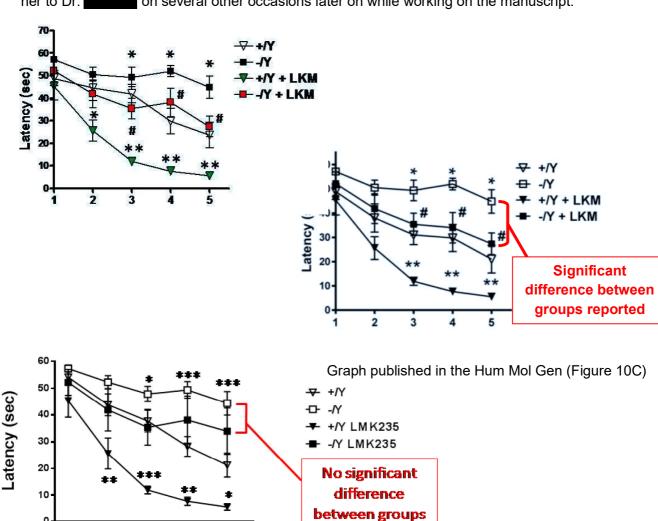
Morris water maze (Fig. 10 C-D):

Presented water maze data do not seem to match the raw data in several aspects.

- Animal numbers stated in the paper (LMK 235 treated +/Y n = 6 and LMK 235 treated -/Y n = 6) do not correspond to animal numbers used in the study (raw data indicate that there were 5 LMK 235 treated +/Y mice and 5 LMK 235 treated -/Y mice).
- The final version on the graph published in Hum Mol Gen (Fig. 10C) is not consistent with the final draft version sent to Dr. by Prof. Ciani the 17th of May 2016.

2-way ANOVA should be used here

Below the discrepancy between the two mentioned graphs id presented. Morris water maze results (learning curve) sent by Prof. Ciani to Dr. on the 5th of May 2016 and re-sent by her to Dr. on several other occasions later on while working on the manuscript:



The graph above was obtained by Dr. as a result of reanalysis of raw data. The main difference reported by the authors between untreated -/Y mice (white squares □), and -/Y LMK235 (black squares ■) could not be proved. This is due to severe errors made in the original analysis.

2

ż

It should be mentioned here that before the publication of the article in Hum Mol Gen, a version of it was rejected from other journal. In that version the first graph shown above (with coloured data points) was included with a following description: "In the MWM test, while Cdkl5 +/Y mice learned to find the platform by the 2nd day, no significant learning was detected in untreated Cdkl5 -/Y mice until the 5th day, clearly indicating a learning deficit in Cdkl5 KO mice". The reviewer pointed out (revision from the 5th of June 2016): "Also they [authors] state that 'Cdkl5 +/Y learned to find the platform by the 2nd day. However, in the plot of the latency [on] the second day they barely changed their execution,

and in fact it is not until the fourth day that their latency becomes clearly reduced. Instead the slope of the learning curve is a bit steeper in the knockout mice. This means that in fact the learning differences between genotypes appear at day four". In the article published in the Hum Mol Gen the description stayed the same, however the graph was modified so the +/Y curve and the description now correspond. To my best knowledge no more experiments were performed in order to obtain new data points that could lead to such a change of this curve.

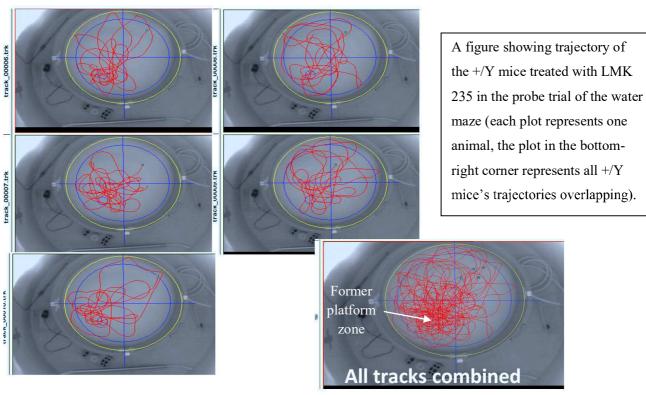
- LMK 235 treated -/Y and LMK 235 treated +/Y animals were examined in May 2015 while control experiments including untreated -/Y and untreated +/Y were run independently (what was not mentioned in Materials and Methods). Dr. could not identify these control experiments among any of the experiments run on age matching untreated animals between December 2013 and September 2014 as groups' mean values presented in the result section do not fit to any raw data set.
- Exactly the same controls (untreated -/Y and +/Y animals) were used in the already published paper "Inhibition of GSK3β rescues hippocampal development and learning in a mouse model of CDKL5 disorder" (Fuchs et al. 2015, doi: http://dx.doi.org/10.1016/j.nbd.2015.06.018, Fig. 9A). Curves for untreated -/Y and untreated +/Y overlap between the mentioned publication and the final draft version of the Hum Mol Gen paper sent to the co-authors by Prof. Ciani (first graph above, with coloured data points), but not with the graph eventually published in Hum Mol Gen where +/Y curve seems to be changed for days 3, 4 and 5.
- As for the probe test (Fig. 10D), the published graph represents the "latency to enter the target quadrant", not the "latency to enter the former platform zone" as incorrectly stated in Materials and Methods and figure caption. This error is crucial for understanding the result of this experiment, as presented "latency to enter the target quadrant" is a very uncommon and disputable indicator of memory retention in the water maze, much less common (if actually used at all) than indicators such as percentage of time spent in the former platform quadrant, latency to enter the former platform zone, number of crossings through the former platform zone or mean proximity to the former platform zone (please refer to a comprehensive article "What is the most sensitive measure of water maze probe test performance?" from Paul Frankland's laboratory, doi: http://dx.doi.org/10.3389/neuro.07.004.2009).

Analysis of raw data on LMK 235 treated +/Y and LMK 235 treated -/Y done by Dr. revealed that for the probe test: percentage of time spent in the former platform quadrant was 42.9 \pm 4.6 for treated +/Y and 28.2 \pm 5.7 (chance level) for treated -/Y; Latency [in seconds] to enter the former platform zone was 11.2 \pm 3.7 and 33.4 \pm 10.2, respectively (latency measured for -/Y treated mice is commonly observed in young untreated -/Y animals; since the control experiment was not run simultaneously, the experiment is not conclusive); Number of crossings through the

former platform zone was 5.2 ± 0.6 and 1.6 ± 0.9 (very low), respectively. These results raise a severe doubt about the efficiency of the LMK 235 treatment on -/Y mice in terms of restoration of learning and memory as measured in the Morris water maze test, which was one of the main hypotheses of the article. Lack of control animals in the same experiment makes it inconclusive.

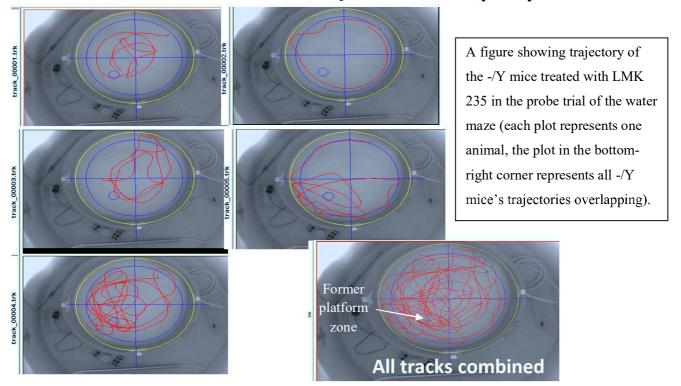
Below, the animals' trajectory during the probe test is presented for LMK 235 treated +/Y and LMK 235 treated -/Y mice, respectively (each plot represents performance of one animal on the course of 60 seconds of the probe trial in the Morris water maze). In each figure the plot in the bottom right corner represents the overlapping tracks obtained for all +/Y mice or all -/Y mice, respectively. The plots were generated by the Ethovision software version 3.0 from the raw files identified as the ones recorded for that experiment. The pictures clearly show memory retention in LMK 235 treated +/Y and no memory retention for LMK 235 treated -/Y mice. Additionally, as no visual (cued) probe was performed (or at least it was not reported in the article and no appropriate file exists in the experiment directory on the computer used for data acquisition), it is not possible to explain the source of low motility also observed in -/Y mice. All these doubts were not mentioned or discussed in the article.

+/Y + LMK 235 (N=5)



LMK 235 treated +/Y mice showed memory retention in the probe trial of the Morris Water Maze

-/Y + LMK 235 (N=5)



LMK 235 treated -/Y mice did not show memory retention in the probe trial of the Morris Water Maze

• Lastly, a new experiment designed to prove the effect of LMK 235 compound on restoration of the hippocampus-dependent behaviour, including spatial memory, in -/Y mice and performed in November 2016, failed to prove it. In other words, the positive behavioural effect of the LMK 235 treatment on -/Y mice published in Hum Mol Gen could not be replicated. In the mentioned experiment, again only treated animals were tested, no vehicle or untreated control animals were used (these controls were about to be run separately in the future). It must be mentioned here though, that the experimental design was slightly different: animals were injected with LMK 235 starting from P5 with 14 doses of the drug (every second day). Behavioural experiments were performed around P45. This change might have influenced the overall effect of the drug, but in the light of presented doubts it may also prove the lack of effect of LMK 235 treatment on the behaviour of -/Y mice.

2.2 <u>Doubts about the paper: Inhibition of GSK3ß with Tideglusib rescues hippocampal</u> defects in juvenile, but not in adult Cdkl5 mutant mice (Fuchs et al. to be published)

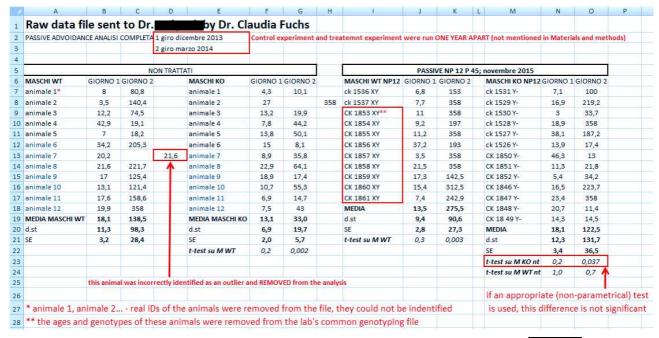
Morris water maze (Fig. 2A-B):

• Presented water maze data do not match the collected data (Fig. 2). Treated animals were examined in May 2016, control experiment including vehicle treated -/Y and vehicle treated +/Y mice was run independently (what was not mentioned in materials and methods and is against good laboratory practice). Moreover, although control animals of matching age were tested in similar conditions in December 2013, February 2014 and September 2014, none of those are presented on the graph in the manuscript, indicating that the presented control groups were taken from yet different experiment that could not be identified. 2 experiments with only treated animals were identified. Experiments from October 2015 and May 2016 were merged to obtain the final results. Control animals were evaluated in a separate experiment that could not be identified so far. For the learning part (days 1-5), 2 treated -/Y animals (1527 and 1530) were included in the final analysis although they should not be (they did not pass the visual probe test). On the other hand, for the probe test (day 6), 6 treated -/Y animals out of 13 were excluded (1527 and 1530 excluded correctly as they did not pass the visual probe plus 4 more without justification: 1526, 1528, 1529, 1531), what decreased the mean latency value for the treated -/Y group for nearly 30% (the excel files with raw data can be made available on demand).

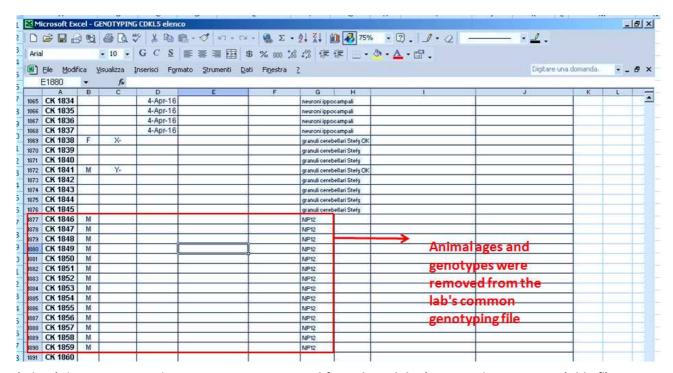
Passive avoidance (Fig. 10 A-B):

Passive avoidance data do not match the data collected by (Fig. 2).

2 experiments with only treated animals were identified. Experiments from November 2015 and May 2016 were merged to obtain the final results (control animals were evaluated in a separate experiment run one year earlier that could not be identified as the specific ID numbers of control -/Y and +/Y mice are omitted in the result file). From the final analysis 2 animals were removed while only 1 was correctly identified as an outlier. Below there is a screenshot pasted from the raw data file sent by Dr. Claudia Fuchs to Dr. (where all the discrepancies are depicted. Because of basic but severe errors in statistical analysis, a non-existing effect was "detected" (difference between latency in untreated and treated -/Y animals), supposedly proving a beneficial effect of the tested drug on memory in passive avoidance test. Examples below very clearly depict and demonstrate major problems in the laboratory. Making raw data commonly inaccessible causes such errors to be proceed further and possibly eventually published.



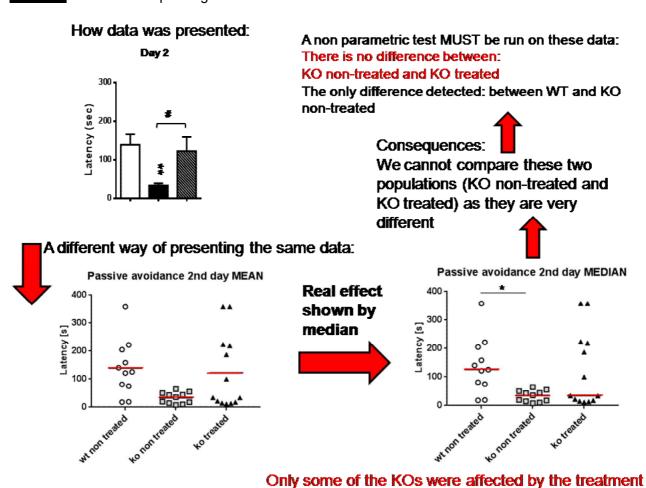
Several errors of data analysis were detected in a raw file sent by Dr. Fuchs to Dr.



Animal data as age and genotype were removed from the original genotyping common lab's file.

The final conclusion from the analysis of the passive avoidance experiment is invalid and not justified by the data. Dr Fuchs stated: "In particular, early GSK3β inhibition with NP-12 restores neuronal survival, dendritic development and hippocampal connectivity in Cdkl5 knockout mice aged P45 (Fig. 3,4). These effects on neuroanatomical defects were accompanied by a complete recovery in

<u>hippocampus-dependent memory performance</u>". The raw data clearly shows that only 5 -/Y animals out of 13 treated animals could be considered as showing "restored recovery in hippocampus-dependent memory performance". When analyzed properly (non-parametric statistical test instead of parametrical), the data do not reveal a significant difference between treated and untreated -/Y animals, in contrary to what was presented in the manuscript. Below there is a diagram sent by Dr.



- Length of dendrites (Fig. 3) there are 2 different versions of data points for WT group, they
- are in the same file sent by Dr Claudia Fuchs to without explanation why the numbers were changed.
- In the manuscript it is never stated how many animals were used for particular procedures what makes it impossible to verify the composition of the experimental groups.
- Improper data acquisition and archivization was detected:
 - Confocal images: samples acquired with z-stack were compared to samples without zstack (samples from treated animals).
 - o CDs with confocal images, at least part of them, are kept at home, not in the lab.
- In some cases appropriate paired statistical tests should be applied. In the materials and methods section no such tests are mentioned.

2.3 <u>Doubts about the paper: Inhibition of GSK3ß rescues hippocampal development and learning in a mouse model of CDKL5 disorder (Fuchs et al. 2015)</u>

- Although in the paper there are vehicle controls mentioned, in fact these were non-treated mice (as stated by a person who personally did the experiment).
- Presented water maze data do not match the collected data (Fig. 9). Treated and control
 animals were examined in September 2014 (when no effect of treatment was shown),
 however these are not the data shown in the paper. Data shown in the paper could not be
 identified.

2.4 <u>Doubts about the paper: Characterization Loss of CDKL5 impairs survival and dendritic</u> growth of newborn neurons by altering AKT/GSK-3β signaling (Fuchs et al. 2014)

- Other groups (Prof. M. Giustetto) were not able to repeat the results / had different results concerning p-AKT, VGLUT1, PSD95.
- No-one else was able to obtain comparable data involving Caspase-3: 10-fold difference in the counting was also pointed out during the Telethon convention in 2014 by another researcher who asked Dr
 what antibody was used to visualize so many cells.

3. Doubts concerning other projects from the lab:

• GENE THERAPY

The particle produced by Dr Fuchs during the internship in the USA (2015) did not prove to work. After several months and many additional experiments performed in the laboratory in Bologna by Dr it occurred that the particle that Dr Fuchs provided was never shown to be functional. The whole experimental work had to be repeated by Dr Fuchs was not able to provide explanations for the exact experimental protocols and steps she has undertaken, but claimed that she run all the necessary controls to prove that the particle was functional. Moreover, Dr Fuchs was not able to explain the cloning strategy she had been supposed to apply during her internship. Dr Fuchs also claimed that her constructs were sequenced during her internship, however one year later she declined it. Verification of her experiments was additionally difficult as Dr Fuchs claimed that the original data from experiments were left in the lab in the USA. Since one of the projects of the lab was based on the mentioned particle, it has been delayed for one year already according to the previous research plan.

PROTEIN THERAPY

- Spines counting performed in a blinded manner did not prove the results obtained when the genotype and treatments were known to the researcher.
- P-AKT optical density in brain slices (images shown to Kininska or Amicus experts and part of the Patent data) was not confirmed in various, independent, western blot

- analysis: neither the putative positive effect of the treatment, nor the difference between untreated KO and WT mice at that age (6 months).
- TAT-CDKL5 diffusion in the brain after ICV injection: images were presented with concentrical diffusion from the ventricle to the other brain areas, which is biologically doubtful considering the very low levels of protein and that the optical signal was amplified.
- Lab members, except of the PI, have no access to legal copies of statistical tools (e.g. SPSS, Graphpad) on common computers. Instead, we are encouraged to use the illegal copies on our private computers. This forms a part of the problem of transparency of the final data and raise doubts about discrepancies between raw and final data that were detected above.

CONCLUSIONS

Taking into account the number and severity of doubts listed above as well as having in mind the best interest of the CDKL5 patients, the authors of this report claim that immediate actions are taken in order to verify the results and correct potential errors.

First of all we insist on MAKING AVAILABLE ALL THE RAW DATA CONCERNING THE EXPERIMENTS DISCUSSED IN THIS REPORT, with all necessary explanations about the changes in the data, removal/ addition of data points, statistical tests used etc. It should be presented in a form that can be understood without further explanations from the side of the person who makes the data available.

Secondly, we insist on REPEATING THE ENTIRE EXPERIMENT that led to the conclusions stated in the to-be-published manuscript by Fuchs et al. sent to the lab members in July 2016.

^{*} Data used to prepare the following report as well as detailed explanation of the process that led to the stated conclusions can be presented by the authors when necessary *