Author's response to reviews

Title: Enrichment of ultraconserved elements among genomic imbalances causing mental delay and congenital anomalies

Authors:

Francisco Martinez (francisco@gva.es)
Sandra Monfort (monmem@alumni.uv.es)
Monica Rosello (rosello_mpi@gva.es)
Silvestre Oltra (jsoltra@uv.es)
David Blesa (dblesa@cipf.es)
Ramiro Quiroga (rquiroga21@yahoo.es)
Sonia Mayo (somadean@alumni.uv.es)
Carmen Orellana (orellana_car@gva.es)

Version: 3 Date: 21 October 2010

Author's response to reviews:

Dear Sir/Madam,

Thank you for your comments about our paper. Please find a revised version of the manuscript, and a point-by-point response to the referees

Reviewer 1

Major compulsory revisions:

Q) I am still confused about the number of UCEs included in the CNVs. In table 1, the total number of UCEs in the CNVs is 38, not 29! Since the same set of 9 UCEs are within the CNVs of cases 2 and 7 both should be included, or not? If not, it should be mentioned as ‘29 different UCEs’ but then I still don’t see why these 9 ‘same’ UCEs should not be included in the enrichment calculations.

A: The same set of UCEs only can be counted once, as we are testing how many DIFFERENT UCEs out of the total are present among pathogenic imbalances. We are not counting cases. The inclusion of the repeated elements would have no sense in the enrichment ratio, because the expected value comes from the proportion of the genome contained in the set of pathogenic imbalances. In this sense, a recurrent rearrangement affects the same region, and the same regulatory elements. We do not try to assess for frequencies including recurrent rearrangements, but for association (enrichment/depletion) to two categories: benign vs. pathogenic.

In the page 4, 3rd paragraph: “We found a total of 29 different elements…”

Q) I did not find any reference to the database of UCEs except for the original paper. If anyone would like to check whether any UCE is associated with their CNV, he/she should be able to find these UCEs? Did I overlook this reference? In case of a database please provide the URL; in case of a feature within a
A: In addition to the original paper, these elements are linked to the UCSC Genome Browser as a custom track. The way to directly access to this track is just clicking in the supplementary Table of the paper by Bejerano et al., Science 304:1321-1325 (2004), in order to display the ‘Ultra’ Track (100% ultra conserved elements (hg16-mm3-rn3) of length 200bp or more). Furthermore, among the publicly available custom tracks in hg18, we can found the Transcribed ultraconserved regions (T-UCRs) reblatted to hg18 (http://genome.ucsc.edu/goldenPath/customTracks/custTracks.html).

Minor essential revisions:

Q) It should be mentioned that 114 UCEs are not correctly mapped and why.
A: It has been mentioned, as suggested, in the introduction.

Q) If the paternity of the CNV is important, then why is it not discussed as mentioned in the answer to my question?
A: Just for not lengthening unnecessarily the text. The results were analyzed on this sense, but no conclusion could be drawn.

Q) In the discussion, I don’t see where the enrichment ratios of 3.7 and 1.7 come from. Please explain or refer to a table, if present there.
A: “In the 2.67% of the genome are present 10% of the isolated elements (13/131) and 4.6% of the clustered elements (16/350)...” So we can directly compute the observed/expected ratio (or enrichment ratio) as (10%)/(2.67%)= 3.7 and (4.6%)/(2.67%)=1.7. In fact these calculations are identical to those explained in tables 2 and 3.

Reviewer 2

Minor Essential Revisions:

Q1) The change regarding to my question (Q1 on the original version of the manuscript, the problem of resolving different versions of human reference genomes for CNVs and UCEs) is not reflected in the text although it said the conversion was done in the cover letter.
A: We are sorry for the confusion. In the Statistical Analysis section (see page 9), we said “…were performed through the UCSC Genome Browser [12], based on build hg18, except for the ultraconserved elements, which were examined in the hg16 version, previous conversion of the delimiting positions of CNVs.” For shake of clarity, this has been changed to: “…were performed through the UCSC Genome Browser [12], based on build hg18. The delimiting positions of all the ultraconserved elements, tracked in the hg16 version, were converted to built hg18 through the ‘Convert’ Feature (http://genome.ucsc.edu/cgi-bin/hgConvert) and compiled in an in-home excel sheet to facilitate visual inspection.”

Q2. More detailed information is needed in the Laboratory Analyses of Methods section. It's still not clear how the clone-based array results are integrated with oligo-based ones from the text (Methods, page 9). What do the authors mean (in
the coverletter) by "our own laboratory data"? Have they used oligo-based arrays only for the data presented in the manuscript?

A: The following paragraph has been added, as requested: “The CNVs detected with both kinds of arrays were collected together, defined by the distal ends of the first and last probe altered. It is worth to note that most benign and pathogenic CNVs were detected by the commercial oligonucleotides-based array because of its higher resolution and because it was applied to 95% patients. On the other hand, they can be considered complementary, as many of the small polymorphic CNVs previously detected in the clone-based array (Monfort et al., 2007) could not be refined in the oligo-array because of lack of probes in such regions, designed in order to avoid frequent polymorphic CNVs.”

“Our own laboratory data” means our patient series data, and the present study is based in the data from the whole series of patients.