Supplementary files

# Identification of a *CNTN5* CNV deletion in a multiplex family with ASD

## Supplementary Note

The multiplex family was identified and recruited in Pakistan by a local psychiatrist (Dr. Brohi Qasim). The family consists of five brothers, a sister, and two parents. Individuals were interviewed by the psychiatrist using a face-to-face standardized evaluation based on DSM-IV criteria. Three of the brothers had a diagnosis of autism, and a fourth brother had learning disorder. Research Ethics Board approval was obtained from the involved institutions and written informed consent was obtained from the participants or their parents/guardians, in compliance with the Helsinki Declaration. Blood was drawn in Pakistan and delivered within 4 days to the laboratory of Dr. Rouleau (McGill University). DNA was extracted following a standard salting-out protocol. Genotyping was performed in the Genome Quebec Innovation Centre (Montréal, Quebec, Canada) using Illumina HumanOmniExpress BeadChip. The final reports were extracted from GenomeStudio after classical quality control and three CNV calling algorithms were used: QuantiSNP[1], PennCNV[2], and CNVPartition (GenomeStudio software, Illumina, San Diego, California, USA). QuantiSNP v2.2 was used with MATLAB Compiler Runtime v7.9 and default parameters. For PennCNV, we first generated a population B allele frequency (PFB) file using the whole genotyping dataset. Then the detect\_cnv.pl script was run using default parameters and the default lib/hh550.hmm model. CNVPartition was run directly from GenomeStudio with default parameters. An in-house Python script, SV-Segregation, was used to identify the segregating CNVs in the multiplex Pakistani family. The script was used to identify CNVs that were shared amongst all three affected brothers with ASD. The software is freely available at <https://bitbucket.org/guyrouleaulab/sv_segregation>. We defined exonic CNVs as those encompassing at least one exon of a gene, and intronic CNVs as those encompassing only an intronic part of a gene. All genome coordinates refer to hg19.

## Supplementary Figure

Chart

Description automatically generated

B.

A.

Graphical user interface, application, table

Description automatically generated

**Pedigree of multiplex Pakistani family with ASD**. A) In the multiplex family, three brothers had a diagnosis of ASD and a fourth brother had a diagnosis of learning disorder. Neither parent had a diagnosis of ASD. An 11q22.1 deletion, inherited from an unaffected father, was identified in all three affected brothers with ASD and their brother with learning disorder. B) Given the genomic coordinates of the CNV, the 11q22.1 deletion encompasses an intronic region of the CNTN5 gene between exon 2 and exon 3 (http://genome.ucsc.edu – human assembly: GRCh37/hg19).

## References

1. Colella S, Yau C, Taylor JM, Mirza G, Butler H, Clouston P, et al. QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. Nucleic Acids Res [Internet]. 2007 [cited 2019 Feb 9];35:2013–25. Available from: https://academic.oup.com/nar/article/35/6/2013/1034786

2. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SFA, et al. PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res [Internet]. 2007 [cited 2019 Feb 9];17:000–000. Available from: http://genome.cshlp.org/content/early/2007/10/05/gr.6861907