De novo recurrent variants in U2AF2 RNA-binding domains in an intellectual disability syndrome.

Dong Li[1], Daniëlle G.M. Bosch[2], Elizabeth J. Bhoj[1], Mar Costa-Roger[3], Maji Debanjana[4], Qin Wang[1], Leslie Granger[5], Andrea Petersen[5], Carrie Hammond[6], Michael Parker[6], Michael E. March[1], Mark R. Battig[1], Hanna Faghfoury[7], L. Kristin Parsley[8], Emanuele Agolini[9], Maria Lisa Dentici[9], Meredith Wright[10], Rachel Palmquist[11], Khanh Lai[11], Marcello Scala[12,13], Alice Brooks[14], Michele Iacomino[13], Federico Zara[12,13], Meghan Towne[15], Kirsty McWalter[16], Raymond Louie[17], Michael Lyons[17], Miroslava Hancarova[18], Zdenek Sedlacek[18], Gemma Aznar-Laín[19], Martin Schwarz[18], Timothy J. Maarup[20], Tugce B. Balci[21], Carlo Marcelis[22], Bert B.A. de Vries[22], Jozef Gecz[23], Elaine H. Zackai[1], Yuanquan Song[1], Gijs van Haaften[2], Clara Kielkopf[4], Luis A. Pérez-Jurado[3,19,23], Hakon Hakonarson[1]

[1]Children's Hospital of Philadelphia; [2]University Medical Center Utrecht; [3]Universitat Pompeu Fabra; [4]University of Rochester School of Medicine; [5]Randall Children's Hospital; [6]Sheffield Children's NHS Foundation Trust; [7]University of Toronto; [8]University of Illinois College of Medicine; [9]Bambino Gesù Children's Hospital; [10]Rady Children's Institute for Genomic Medicine; [11]Universtiy of Utah; [12]University of Genoa; [13]IRCCS Istituto Giannina Gaslini; [14]Erasmus Medical Center; [15]Ambry Genetics; [16]GeneDx; [17]Greenwood Genetic Center; [18] Charles University Prague; [19]Hospital del Mar Research Institute; [20]Kaiser Permanente; [21]Western University, London Health Sciences Centre; [22]Radboud University Medical Centre; [23]University of Adelaide

Several evolutionarily conserved *cis*-acting sequences on pre-mRNA facilitate the recognition of exon-intron boundaries and identify the 3' splice site (SS) during the spliceosome assembly. In this complex, U2AF1 recognizes the conserved AG dinucleotides preceding the 3'SS, and concurrently C-terminal RNA binding domains (RBDs) of U2 auxiliary factor subunit 2 (U2AF2) bind to the sequential polypyrimidine tract, adjacent to the AG dinucleotides at the 3'SS. These early steps ensure the high fidelity of splicing, which is vitally important because mistakes in splicing could result in unintended effects leading to dysregulation of abundance of different gene isoforms expression and/or frameshift mutations. By exome sequencing and international matchmaking we identified 20 unrelated individuals with nine different *de novo* novel missense variants in the RBDs of U2AF2. Four out of the nine variants were recurrent in RNA binding motif 1 (RRM1), where most variants reside, and were found in 15 individuals [c.445C>T (p.Arg149Trp), c.448C>T (p.Arg150Cys), c.449G>A (p.Arg150His), c.556G>A (p.Val186Met)], representing mutation hotspots of *U2AF2*. All of the identified variants reside in the highly intolerant regions. Detailed clinical assessment of the affected individuals showed that they present with developmental delay (DD) or intellectual disability (ID) (20/20), febrile seizures (8/13), hypoplastic or agenesis of corpus callosum (5/12), and similar craniofacial dysmorphology. Mapping of the mutations on the three-dimensional structures of RBDs of U2AF2 predicted that six of the identified variants are likely to alter splice site RNA binding affinity and/or specificity, including the three recurrent ones - c.445C>T (p.Arg149Trp), c.448C>T (p.Arg150Cys), and c.449G>A (p.Arg150His), whereas three mutations, including the recurrent variant c.556G>A (p.Val186Met), could affect the stability of the protein, which we currently are validating experimentally. Preliminary transcriptome analyses from patients' cells suggest differentially spliced exons in some target genes. Together, these findings suggest that

these variants may exert a loss of function effect. To test our hypothesis, we are currently conducting additional functional assessments. Furthermore, we queried denovo-db and identified four additional individuals from DDD4K study harboring de novo variants in U2AF2, suggesting about 0.1% of affected individuals with DD/ID may result from a *de novo* variant in *U2AF2*. In summary, through multi-center collaborations, we demonstrate that *de novo U2AF2* mutations are associated with a novel neurodevelopmental disorder.