**The American Journal of Human Genetics, Volume *99***

## Supplemental Data

# De Novo Truncating Variants in *SON* Cause Intellectual Disability, Congenital Malformations, and Failure to Thrive

**Mari J. Tokita, Alicia A. Braxton, Yunru Shao, Andrea M. Lewis, Marie Vincent, Sébastien Küry, Thomas Besnard, Bertrand Isidor, Xénia Latypova, Stéphane Bézieau, Pengfei Liu, Connie S. Motter, Catherine Ward Melver, Nathaniel H. Robin, Elena M. Infante, Marianne McGuire, Areeg El-Gharbawy, Rebecca O. Littlejohn, Scott D. McLean, Weimin Bi, Carlos A. Bacino, Seema R. Lalani, Daryl A. Scott, Christine M. Eng, Yaping Yang, Christian P. Schaaf, and Magdalena A. Walkiewicz**

Supplemental Data

Supplemental Case Reports:

Subject 1 (Subject 1, II-1 in figure 1) is a 6 year old girl who was born to non-consanguineous parents at 32 weeks gestational age by caesarean section for fetal distress. Pregnancy was complicated by placenta previa and intrauterine growth restriction. In the neonatal period, she had respiratory failure necessitating mechanical ventilation, was hypotonic, and had feeding difficulties. Her subsequent development was delayed. She started crawling at 22 months, walking at 2 years, and acquired her first words at 18 months. She developed seizures at 1 year of age during febrile illness. Brain MRI performed at 2 years showed global cerebral volume loss with thinning of the corpus callosum, subtle periventricular gliosis, and external enlargement of subarachnoid space. By age 3 she was having up to 20 seizures per day and EEG performed at that time showed intermittent slow activity in the central regions and excessive slow activity in the occipital regions consistent with diffuse disturbance of cerebral function. An electrographic seizure discharge originating in the right central region was recorded. At her most recent follow-up visit at 6 years of age, she had persistent severe developmental delay, was nonverbal, and had been diagnosed with autism spectrum disorder. Seizure frequency was greatly reduced and daily anti-epileptics had been discontinued. She had persistent feeding difficulties with poor weight gain, as well as a history of marked joint laxity. Additional pertinent medical concerns included a history of delayed gastric emptying, intermittent exotropia, nystagmus, recurrent otitis media requiring myringotomy tube placement, a resolved atrioseptal defect, and a history of deep vein thrombosis. On physical exam, she was at the $2^{nd}$ - $3^{rd}$ percentile for height, weight, and head circumference. Distinctive features included frontal bossing, bitemporal narrowing, bilateral epicanthal folds, a small nose, horizontal eyebrows, and a smooth philtrum and thin upper lip [Figure 1]. She was globally hypotonic with diminished patellar reflexes bilaterally and had an unsteady gait.

Subject 2 (Subject 2, II-1 in figure 1) is a 23 year old male who was born at 41 weeks gestational age by caesarean section for fetal distress following a pregnancy complicated by low-grade maternal hypertension and increased fetal activity. A submucous cleft palate, bifid uvula, and jaundice were noted at birth. He had feeding difficulties and contracted RSV pneumonia at 5 weeks of age. Subsequent motor and speech development were delayed and he had onset of seizures at age 2 years. An MRI performed at 18 years of age showed a left paracerebellar arachnoid cyst and progressive ventricular and subarachnoid space dilatation. At his most recent follow-up at 23 years of age, he was living in a semi-independent environment with full-time caretakers. He was reported to have auditory hallucinations and behavioral issues, including aggressive and self-injurious behavior that had responded well to medical therapy. Other health concerns included ongoing dysphagia, a history of pancreatic lipase insufficiency, pancreatitis, and hypertension. He also had a history of poor enamel and dentition, and of exotropia and progressive vision loss. On physical exam, he was at the 40$^{th}$ percentile for height, the 1$^{st}$ percentile for weight, and 50$^{th}$ percentile for head circumference. Notable features included arachnodactyly and dolichostenomelia.

Subject 3 (Subject 3, II-1 in figure 1) is a 9 year old female who was born to non-consanguineous parents at 40 weeks gestation by induced vaginal delivery for failure to progress, following a pregnancy complicated by decreased fetal movements. At birth, she was small for gestational age, jaundiced, and required short-term hospitalization for apathy to nursing, hypoglycemia, and temperature instability. During her first two years of life she had failure to thrive, chronic diarrhea, and recurrent otitis media requiring myringotomy tubes, with laboratory work showing both IgA and IgG deficiency. At 2 years of age, she was noted to have speech delay, poor interaction with her siblings and self-stimulatory behavior, and ultimately received a diagnosis of autism spectrum disorder at age 3 years. At age 5 years,

she developed staring episodes. An interictal EEG at that time showed intermittent bitemporal rhythmic slowing. Repeat EEG and brain MRI performed at age 8 years were normal. At her most recent follow-up at age 9 years she was in mainstream classes with a one-on-one teaching assistant due to lack access to special education and/or life skills classes. She had persistent feeding difficulties and self-stimulatory behaviors. Her parents reported her dentist had noted graying of the teeth despite good dental hygiene. On exam, she was at the 25$^{th}$ percentile for height, less than the 5$^{th}$ percentile for weight (z-score: -2.3), and less than the 3$^{rd}$ percentile (z-score: -4) for head circumference. Notable facial features included slightly down-slanting palpebral fissures, a down-turned mouth with a thin upper lip, thin extremities with low body fat, lower back dimples, small joint hypermobility (Beighton score 6/9), and a palpable right cervical rib.

Subject 4 (not shown in figure 1) is a 3 year old female who was born to non-consanguineous parents at 33 weeks by caesarean section for fetal distress following a pregnancy complicated by IUGR, maternal borderline diabetes, and Factor V Leiden deficiency. Postnatally, she was hypotonic and required intubation. She was noted to have poor suck, swallow, and breathe likely related to presence of a submucous cleft palate, a type 1 laryngeal cleft, laryngomalacia, and oropharyngeal dysphagia. Developmental delay was first noted at age 15 months; she did not consistently sit independently until 17 months. An eye exam performed at that time showed difficulty tracking, pseudostrabismus, and cortical visual impairment. Brain MRI performed at 17 months showed periventricular leukomalacia with mild dilatation of the lateral ventricle secondary to volume loss. Around the same time, she was seen by cardiology and found to have abnormal placement of the carotid arteries in the neck. She was seen for follow-up at three years of age. Pertinent medical concerns addressed at that time included a history of IgA deficiency that had resolved over time with normalization of her response to vaccination by around 2 years of age. IgG was normal for her age. She had a history of persistent feeding difficulties leading to

placement of a g-tube for feeding. This was ultimately found to be a gastric motility disorder and she started eating by mouth following treatment with reglan. She also had findings concerning for congenital myasthenic syndrome based on abnormal EMG studies. She additionally had a history of photosensitivity, mild vesicoureteral reflux with urinary tract infections, PE tube placement, and past upper respiratory tract infections. On exam at age 3 years, she was crawling and pulling to stand but not walking. She had routine usage of approximately 5 words. She was noted to have dramatically improved growth parameters including reaching the 75$^{th}$ percentile for height, 85$^{th}$ percentile for weight, and 60$^{th}$ percentile for head circumference. Notable dysmorphic features included a broad forehead, flattened nasal bridge, bilateral epicanthal folds, cupid's bow upper lip with a thin lower lip, and prominent finger pads with a positive thumb sign and thumb to forearm sign.

Subject 5 (Subject 5, II-1 in figure 1) is a 15-year old girl who was born to non-consanguineous parents at 35 weeks by caesarean section for maternal hypertension. Her neonatal course was complicated by feeding problems requiring nasogastric tube feeds, gastroesophageal reflux, and need for transient respiratory support with oxygen and CPAP. Brain MRI performed at age 1 year showed prominent extra-axial CSF spaces, abnormal signal in the peri-atrial white matter most consistent with gliosis, dysgenesis of the rostrum of the corpus callosum, and mild delay in peripheral white matter fiber myelination. Her development was delayed and she had a diagnosis of moderate-severe intellectual disability. She had a history of behavioral concerns described as tantrums with self-injurious behavior that occurred without provocation. Additional pertinent medical concerns included a history of pneumonia requiring intubation at age 12 years, a history of urolithiasis, and congenital absence of the right kidney. She had also had periods of poor feeding followed by periods of excessive eating. At her most recent follow-up visit at 15 years of age, she was at the 3$^{rd}$ centile for height, 12$^{th}$ centile for weight, and 72$^{nd}$ centile for

head circumference. She was nonverbal but able to follow simple commands. Pertinent findings on physical exam included laterally flared eyebrows and exaggerated lumbar lordosis.

Subject 6 (Subject 6, II-1 in figure 1) is a 9- year- old female who was born at 36 weeks gestation to non-consanguineous parents via vaginal delivery. Her prenatal course was complicated by IUGR, abnormal prenatal ultrasound imaging of the kidneys, oligohydramnios and maternal preeclampsia. She was hospitalized for 11 days following delivery for feeding problems requiring NG tube feeds. She was diagnosed with a right multicystic dysplastic kidney and congenital lobar emphysema. At 6 months of age, she was admitted to the hospital for respiratory failure and had a very long course including the need for a tracheostomy and G-tube. The tracheostomy was closed at 5 years of age. She was also noted to have significant hypotonia and developmental delays. She suffered a middle cerebral artery stroke in 2010, and since then has had multiple transient ischemic attacks. Multiple CT scans of the brain were unremarkable with the exception of mild prominence of the lateral and third ventricles. MRI and MRA studies of the brain showed slight asymmetry of the temporoparietal lobes with a few foci of white matter T2 signal abnormalities from prior right MCA stroke which have been stable. Although she has not had documented seizures, she has had an abnormal EEG. The EEG was limited due to artifact, but background appeared mildly slow and disorganized with more prominent slowing on the right. No frank epileptiform discharges were noted. At her most recent follow-up at 9 years of age, she had down-slanting palpebral fissures and a long face with full cheeks. She had a history of multiple surgeries for strabismus. Hearing tests had been inconclusive and a sedated exam was planned. Developmentally, she was progressing slowly. She was speaking in simple sentences and able to walk with minimal dragging of her left lower extremity. Her height and weight consistently tracked below the $5^{th}$ centile, with height now at 3 standard deviations below the mean, and a head circumference at the $12^{th}$ centile.

For subjects 1-6, we obtained consent for research studies under a protocol approved by the BCM institutional review board. These subjects were identified by querying our database of 6,000+ consecutive WES cases referred to the Baylor Miraca Genetics Laboratory for clinical whole exome sequencing between October 2011 and February 2016. Whole exome sequencing and data processing were performed as previously described[1]. This test targets coding and untranslated region exons of approximately 20,000 genes to a mean coverage of greater than 130x with 95% of targeted regions achieving at least 20x coverage[1]. All detected variants in the *SON* gene were confirmed by Sanger sequencing of probands, and *de novo* status was confirmed by sequencing of maternal and paternal samples in all cases [Figure 1]. Written informed consent for publication was obtained for all described subjects.

Subject 7 (Subject 7, II-1 in figure 1) is a 3-year-old girl who was born to non-consanguineous parents following a pregnancy complicated by severe intrauterine growth restriction and ascertainment of congenital heart defects and left lung agenesis on prenatal ultrasound. Delivery was induced at 36 weeks for IUGR and a cesarean section was performed because of abnormal fetal heart rate. Brief ventilation was required after delivery. Postnatal examination confirmed left lung agenesis, multiple ventricular septal defects, and a patent ductus arteriosus. She was also noted to have left thumb agenesis and right thumb hypoplasia, bilateral 2-3 toe syndactyly, gallbladder agenesis, and a right cervical chondroma. Radiographic studies revealed a T4 hemivertebra, sagittal slot from T1 to T5, synostosis of the 1st - 2nd and 3rd - 4th ribs on the left side without deviation of the spine. Cardiac surgery was performed at age 1 month and reconstructive hand surgery at age 1 year. Because of feeding difficulties and poor oral intake, placement of a gastrostomy tube was recommended but declined. During subsequent follow-up, she was noted to have central hypotonia in addition to psychomotor and speech delays. She sat independently at age 17 months, walked at age 25 months,

acquired her first words at age 2 years and first spoke in sentences at age 3 years. At her most recent follow-up at age 3 years, her height was 85 cm (-2.2SD), her weight was 9.6 kg (-3SD), and her head circumference was 45.5 cm (-2.5SD). Notable dysmorphic features included plagiocephaly, a prefrontal angioma, small mouth with thin lips and hypoplasia of the triangular muscle of the upper lip, flat philtrum, full cheeks, bulbous nose, low-set and dysplastic ears, left ptosis and synophrys.

For subject 7, we obtained written informed consent for genetic analyses according to French ethical guidelines. Whole-exome trio sequencing capture was performed using the BGI Human 59M Exon kit based on Combinatorial Probe-Anchor Ligation (cPAL™) technology and the captured material was sequenced on Complete Genomics platform. Sequencing reads were aligned to the reference genome sequence GRCh37 using Teramap base-calling software (BGI). Variants were called using two different pipelines: 1. an in-house method from BGI scoring hypotheses by a Bayesian framework, 2. a combination of Picard software (version 1.119) for removal of PCR duplicates, GATK (version 3.2.2) for indel realignment and base recalibration, and GATK (UnifiedGenotyper or HaplotypeCaller) and Samtools mpileup (version 1.0.29-g68ca977) for variant calling. Results obtained by the two pipelines were compared to each other, and a higher confidence score was attributed to variants found by both pipelines. Variants with a frequency >0.01% were filtered out against 1000 Genomes data (version 2014 Oct, 2577 individuals), Genome of the Netherlands data [Genome of the Netherlands, 2014] (SNPs and Indels release 5, 769 individuals), the NHLBI GO Exome Sequencing Project (ESP) data (ESP6500SI-V2, 6503 individuals; NIEHS Environmental Genome Project, Seattle, WA (URL: http://evs.gs.washington.edu/niehsExome/) [accessed on April 2016]), and the Exome Aggregation Consortium (ExAC) (Cambridge, MA (URL: http://exac.broadinstitute.org) [accessed on April 2016]). Variants were then selected according to the hypotheses of a *de novo* or a recessive event. Remaining variants were removed if found in a local database of 300 healthy control individuals from CHU Nantes.

Potential pathogenicity of variants was determined using SIFT[2] (version 5.2.2, released November 7, 2014), PolyPhen2[3] (version 2.2.2), Align GVGD, Mutation Taster[4], and CADD v1.3[4] programs.

Subject 8

Consent for research studies was obtained under a protocol approved by the BCM institutional review board. This subject was studied by a 400k customized oligonucleotide-SNP microarray, designed by BMGL and manufactured by Agilent Technologies (Santa Clara, CA, USA). The array targets over 4,200 genes at the exon level and includes 60,000 SNP probes. The entire genome is covered with an average resolution of 30 kb, excluding repetitive sequences. The procedures for DNA digestion, labeling and hybridization, and data analysis, were performed as previously described[5].

Supplemental Figures

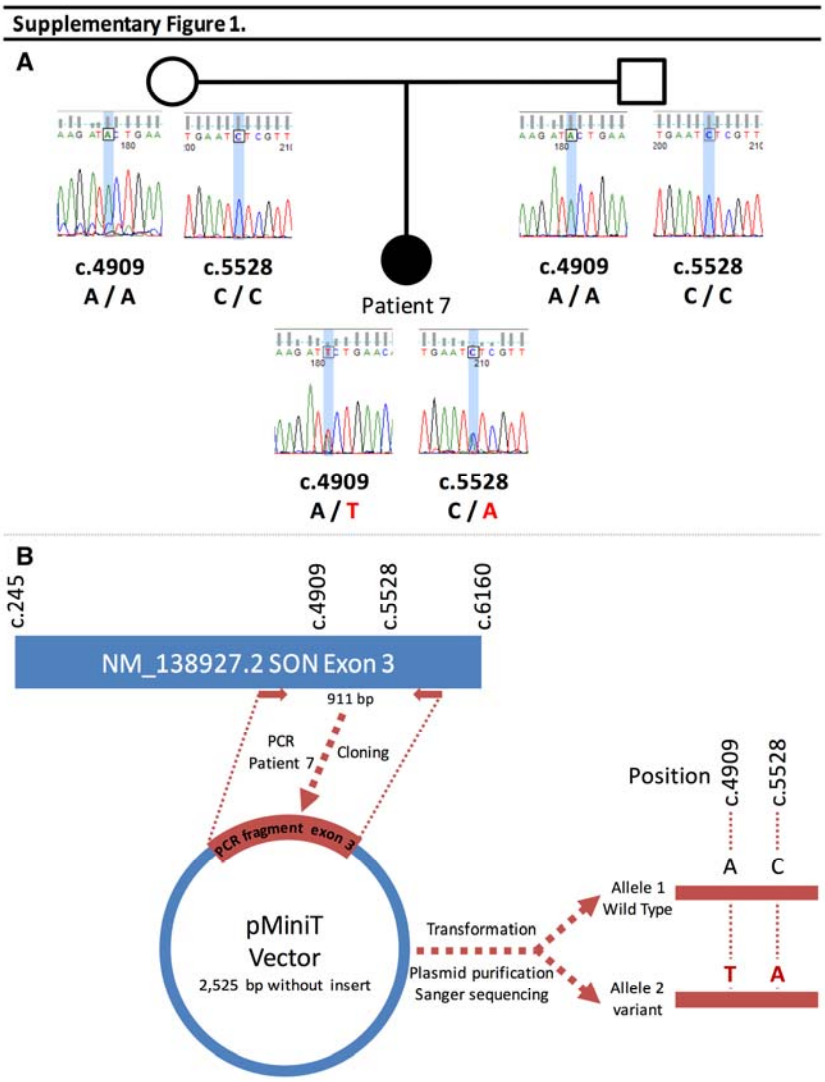


Figure S1. Haplotype phasing of the *de novo* missense variants detected in Subject 7.

A) Chromatograms showing that mother and father are negative for the missense variants detected in Subject 7. B) The region encompassing the two variants was PCR-amplified, sub-cloned, and Sanger sequenced (as depicted) to assess phase of the two *de novo SON* variants, c.4909A>T (p.Thr1637Ser) and c.5528C>A (p.Ser1843Tyr), detected in Subject 7. Analysis of sequence data revealed that the two variants are in *cis* configuration.

**Figure S2.**

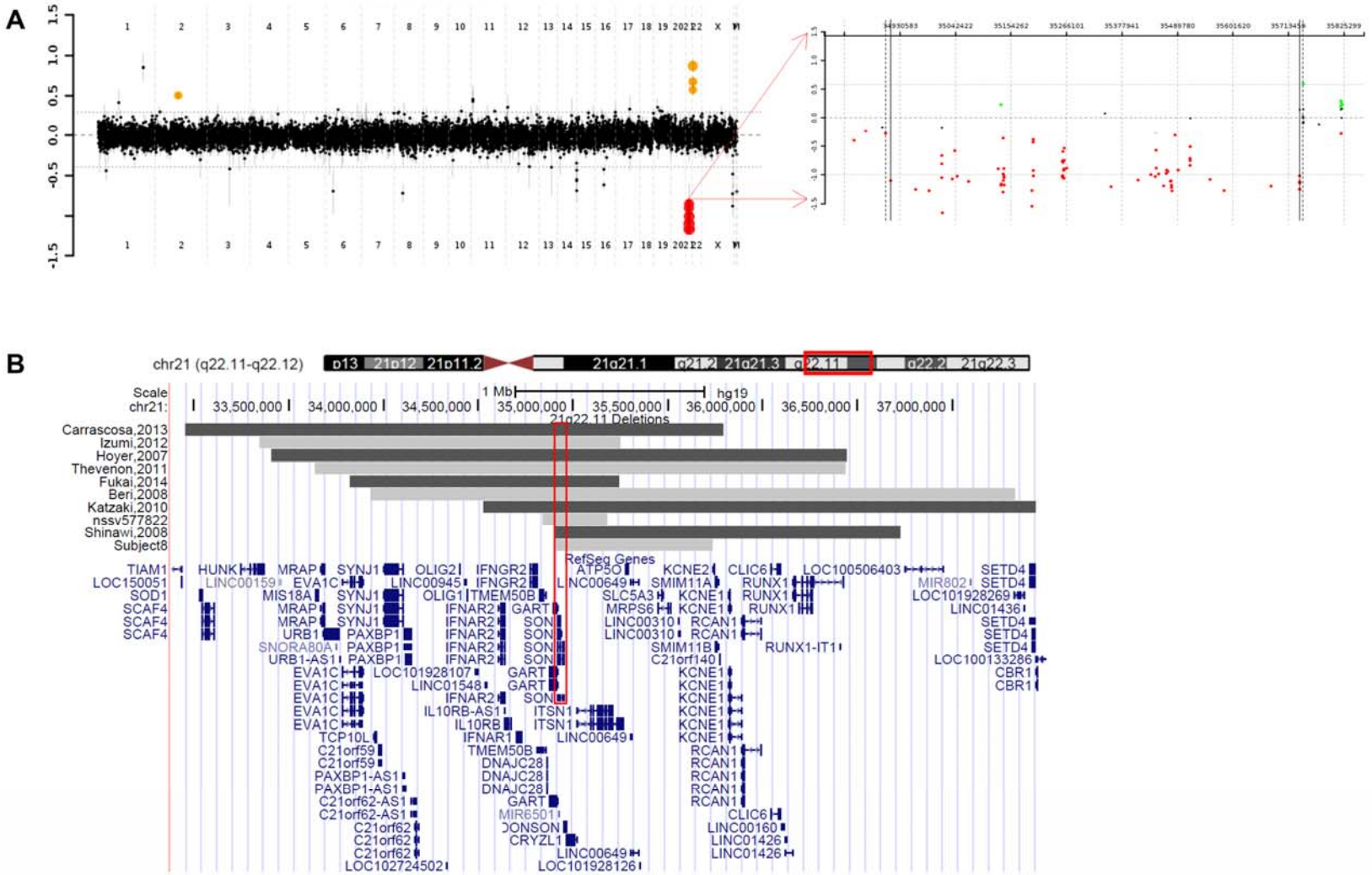
A
B
chr21 (q22.11-q22.12)
21q22.11 Deletions
RefSeq Genes
Carrascosa,2013
Izumi,2012
Hoyer,2007
Thevenon,2011
Fukai,2014
Beri,2008
Katzaki,2010
nssv577822
Shinawi,2008
Subject8

Figure S2. Copy number variants involving the *SON* gene

A) $Log_2$ ratio plots showing the 825kb deletion detected in Subject 8. B) Alignment of published subjects with deletions < 5Mb in size encompassing the *SON* gene. The deletion detected in Subject 8 and a small deletion reported in ClinVar (Accession # SCV000080160.5; dbVar nssv577822) are also depicted. RefSeq genes are displayed below the tracks (UCSC Genome Browser, GRCh37/hg19).

Supplemental References


1. Yang, Y., Muzny, D.M., Xia, F., Niu, Z., Person, R., Ding, Y., Ward, P., Braxton, A., Wang, M., Buhay, C., et al. (2014). Molecular findings among individuals referred for clinical whole-exome sequencing. JAMA 312, 1870-1879.
2. Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4, 1073-1081.
3. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat Methods 7, 248-249.
4. Schwarz, J.M., Cooper, D.N., Schuelke, M., and Seelow, D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods 11, 361-362.
5. Boone, P.M., Bacino, C.A., Shaw, C.A., Eng, P.A., Hixson, P.M., Pursley, A.N., Kang, S.H., Yang, Y., Wiszniewska, J., Nowakowska, B.A., et al. (2010). Detection of clinically relevant exonic copy-number changes by array CGH. Hum Mutat 31, 1326-1342.