Please cite this article in press as: Kim et al., De Novo Mutations in *SON* Disrupt RNA Splicing of Genes Essential for Brain Development and Metabolism, Causing an Intellectua..., The American Journal of Human Genetics (2016), http://dx.doi.org/10.1016/j.ajhg.2016.06.029

REPORT

De Novo Mutations in *SON* Disrupt RNA Splicing of Genes Essential for Brain Development and Metabolism, Causing an Intellectual-Disability Syndrome

Jung-Hyun Kim,[1,45] Deepali N. Shinde,[2,45] Margot R.F. Reijnders,[3,45] Natalie S. Hauser,[4] Rebecca L. Belmonte,[5] Gregory R. Wilson,[5] Daniëlle G.M. Bosch,[3] Paula A. Bubulya,[6] Vandana Shashi,[7] Slavé Petrovski,[8,9] Joshua K. Stone,[1] Eun Young Park,[1] Joris A. Veltman,[3,10] Margje Sinnema,[10] Connie T.R.M. Stumpel,[10] Jos M. Draaisma,[11] Joost Nicolai,[12] University of Washington Center for Mendelian Genomics, Helger G. Yntema,[3] Kristin Lindstrom,[13] Bert B.A. de Vries,[3] Tamison Jewett,[14] Stephanie L. Santoro,[15,16] Julie Vogt,[17] Deciphering Developmental Disorders Study,[18] Kristine K. Bachman,[19] Andrea H. Seeley,[19] Alyson Krokosky,[20] Clesson Turner,[20]

(Author list continued on next page)

The overall understanding of the molecular etiologies of intellectual disability (ID) and developmental delay (DD) is increasing as next-generation sequencing technologies identify genetic variants in individuals with such disorders. However, detailed analyses conclusively confirming these variants, as well as the underlying molecular mechanisms explaining the diseases, are often lacking. Here, we report on an ID syndrome caused by de novo heterozygous loss-of-function (LoF) mutations in *SON*. The syndrome is characterized by ID and/or DD, malformations of the cerebral cortex, epilepsy, vision problems, musculoskeletal abnormalities, and congenital malformations. Knockdown of *son* in zebrafish resulted in severe malformation of the spine, brain, and eyes. Importantly, analyses of RNA from affected individuals revealed that genes critical for neuronal migration and cortex organization (*TUBG1*, *FLNA*, *PNKP*, *WDR62*, *PSMD3*, and *HDAC6*) and metabolism (*PCK2*, *PFKL*, *IDH2*, *ACY1*, and *ADA*) are significantly downregulated because of the accumulation of mis-spliced transcripts resulting from erroneous SON-mediated RNA splicing. Our data highlight SON as a master regulator governing neurodevelopment and demonstrate the importance of SON-mediated RNA splicing in human development.

Recent advances in whole-exome and whole-genome sequencing have accelerated the identification of the genetic etiologies of intellectual disability (ID) and developmental delay (DD), facilitating appropriate care and therapy for affected individuals and their families. So far, mutations in more than 1,500 genes have been implicated in ID and DD disorders,[1–9] and de novo single-nucleotide variants and copy-number variations (CNVs) have been identified as a major cause of severe ID and/or DD.[5,7] Recently, two independent studies reported on a single individual with ID and/or DD and a de novo mutation in *SON* (SON DNA binding protein [MIM: 182465]), which encodes a protein required for proper RNA splicing. However, the level of evidence required for securely implicating mutations in this gene as disease causing was lacking.[5,10,11] Including these two individuals, we recruited a total of 20 unrelated individuals with mild to severe ID and/or DD (Figure 1A and Table S1) and report on the delineation of an ID syndrome caused by de novo LoF mutations in *SON*. This study was approved by the local institutes under the realm of diagnostic testing.

We compared in detail the phenotypic characteristics of all 20 individuals with *SON* LoF mutations. Clinical examination showed that all affected individuals had mild to moderate facial dysmorphisms, including facial asymmetry, midface retraction, low-set ears, downslanting palpebral fissures, deep-set eyes, horizontal eyebrows, a broad and/or depressed nasal bridge, and a short philtrum (Figures 1B and Figure S1). Interestingly, brain MRI, available for 19 affected individuals, revealed that 17 of them had significant abnormalities, including abnormal gyration patterns (polymicrogyria, simplified gyria, and periventricular nodular heterotopia), ventriculomegaly, Arnold-Chiari malformations, arachnoid cysts, hypoplasia of the

[1]Mitchell Cancer Institute, University of South Alabama, Mobile, AL 36604, USA; [2]Ambry Genetics, Aliso Viejo, CA 92656, USA; [3]Department of Human Genetics, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Medical Center, 6500 HB Nijmegen, the Netherlands; [4]Medical Genetics and Metabolism, Valley Children's Hospital, Madera, CA 93636, USA; [5]Department of Biological Sciences, California State University, Chico, Chico, CA 95929, USA; [6]Department of Biological Sciences, Wright State University, Dayton, OH 45435, USA; [7]Division of Medical Genetics, Department of Pediatrics, Duke University School of Medicine, Durham, NC 27710, USA; [8]Department of Medicine, University of Melbourne, Austin Hospital and Royal Melbourne Hospital, Parkville, VIC 3010, Australia; [9]Institute for Genomic Medicine, Columbia University, New York, NY 10027, USA; [10]Department of Clinical Genetics and School for Oncology & Developmental Biology (GROW), Maastricht University Medical Center, 6202 AZ Maastricht, the Netherlands; [11]Department of Pediatrics, Radboudumc Amalia Children's Hospital, 6500 HB Nijmegen, the Netherlands; [12]Department of Neurology, Maastricht University Medical Center, 6299 HX Maastricht, the Netherlands; [13]Division of Genetics and Metabolism, Phoenix Children's Hospital, Phoenix, AZ 85016, USA; [14]Section on Medical Genetics, Department of Pediatrics, Wake Forest School of Medicine, Winston-Salem, NC 27157, USA; [15]Nationwide Children's Hospital, Columbus, OH 43205, USA; [16]Ohio State University College of Medicine, Columbus, OH 43210, USA; [17]West Midlands

(Affiliations continued on next page)

Please cite this article in press as: Kim et al., De Novo Mutations in *SON* Disrupt RNA Splicing of Genes Essential for Brain Development and Metabolism, Causing an Intellectua..., The American Journal of Human Genetics (2016), http://dx.doi.org/10.1016/j.ajhg.2016.06.029

Luis Rohena,[21,22] Maja Hempel,[23] Fanny Kortüm,[23] Davor Lessel,[23] Axel Neu,[24] Tim M. Strom,[25,26] Dagmar Wieczorek,[27,28] Nuria Bramswig,[28] Franco A. Laccone,[29] Jana Behunova,[29] Helga Rehder,[29] Christopher T. Gordon,[30,31] Marlène Rio,[32] Serge Romana,[33,34] Sha Tang,[2] Dima El-Khechen,[2] Megan T. Cho,[35] Kirsty McWalter,[35] Ganka Douglas,[35] Berivan Baskin,[35] Amber Begtrup,[35] Tara Funari,[35] Kelly Schoch,[7] Alexander P.A. Stegmann,[10] Servi J.C. Stevens,[10] Dong-Er Zhang,[36,37,38] David Traver,[38] Xu Yao,[36] Daniel G. MacArthur,[39,40,41] Han G. Brunner,[3,10] Grazia M. Mancini,[42] Richard M. Myers,[43] Laurie B. Owen,[1] Ssang-Taek Lim,[44] David L. Stachura,[5] Lisenka E.L.M. Vissers,[3,45,*] and Eun-Young Erin Ahn[1,44,45,**]

corpus callosum, hypoplasia of the cerebellar hemispheres, and loss of periventricular white matter (Figures 1C–1E). 11 of 20 individuals developed seizures and/or epilepsy with an age of onset ranging from 1 to 6 years. 17 of 20 individuals presented with musculoskeletal abnormalities, comprising hemivertebrae, scoliosis or kyphosis, contractures, hypotonia, and hypermobility of the joints. Vision problems, including cerebral visual impairment, hypermetropia, optic atrophy, and strabismus, were present in 15 of 20 individuals. In addition, the vast majority of individuals showed congenital malformations consisting of urogenital malformations (6/20), heart defects (5/20), gut malformations (3/20), and a high and/or cleft palate (2/20). Short stature was present in ten individuals, and craniosynostosis involving both the metopic (n = 1) and the sagittal sutures (n = 2) was noted in 3 of 20 individuals. Metabolic screening was performed in 9 of 20 individuals, confirming mitochondrial dysfunction in individuals 2 and 11 and an O-glycosylation defect in individual 20 (a clinical summary is provided in Table 1, and details are listed in Table S2). Apart from individuals 13 (II-1 in family 13; Figure 1A), 15 (II-3 in family 15), and 20 (II-1 in family 20), none of the individuals had additional coding-sequence mutations that explained (part of) the phenotype (Table S2). Individual 13 was clinically diagnosed with dyskeratosis congenita, for which a maternally inherited pathogenic *TERT* (MIM: 187270) mutation was identified (Table S2). Individual 13 was, however, more severely affected than could be explained by a *TERT* mutation alone. Similarly, none of the other coding variants identified in individual 15 or the additional genes deleted by the 384 kb deletion CNV in individual 20 were likely to explain the phenotype of these individuals (Table S2).

SON (GenBank: NM_138927.2) is composed of 12 exons (Figure 2A) and encodes a protein (GenBank: NP_620305.2) containing an arginine/serine (RS)-rich domain and two RNA-binding motifs (a G-patch and a double-stranded RNA binding motif) (Figure 2A).[12–14] 17 of 20 mutations are frameshift mutations, including a recurrent 4-bp deletion (c.5753_5756del) in four independent individuals (Table S1 and Figures 2A and 2B). The remaining mutations include a nonsense mutation, an in-frame deletion of eight amino acids, and a complete gene deletion (Table S1). Importantly, parental DNA was available for testing in 19 of 20 individuals and indicated that all mutations had occurred de novo (Figure S2 and Table S1). Interestingly, de novo truncating mutations in *SON* have not been observed in over 2,000 control individuals,[4,15–18] and *SON*, with a Residual Variation Intolerance Score of −1.88, belongs to the 2% most intolerant human protein-coding genes.[19] Furthermore, interrogation of large databases (such as the Exome Aggregation Consortium [ExAC] Browser) has shown that, after sequence context and mutability are considered, *SON* is significantly depleted of LoF variants according to multiple LoF metrics (pLI = 1.00, and the false-discovery rate of the LoF depletion score is $p = 1.68 \times 10^{-6}$).[20,21] Although these population genetic signatures of intolerance cannot

Regional Genetics Service, Birmingham Women's NHS Foundation Trust, Birmingham B15 2TG, UK; [18]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK; [19]Geisinger Medical Center, Danville, PA 17822, USA; [20]Division of Genetics, Department of Pediatrics, Walter Reed National Military Medical Center, Bethesda, MD 20889, USA; [21]Division of Genetics, Department of Pediatrics, San Antonio Military Medical Center, Fort Sam Houston, TX 78234, USA; [22]Department of Pediatrics, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA; [23]Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; [24]Department of Pediatrics, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; [25]Institute of Human Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany; [26]Institute of Human Genetics, Technical University of Munich, 81675 Munich, Germany; [27]Institute of Human Genetics, University Clinic Düsseldorf, Heinrich-Heine-University, 40225 Düsseldorf, Germany; [28]Institute of Human Genetics, University Clinic Essen, University Duisburg-Essen, 45147 Essen, Germany; [29]Institute of Medical Genetics, Medical University of Vienna, Waehringer Strasse 10, 1090 Vienna, Austria; [30]Laboratory of Embryology and Genetics of Congenital Malformations, INSERM UMR 1163, Institut Imagine, 75015 Paris, France; [31]Université Paris Descartes, Sorbonne Paris Cité, Institut Imagine, 75015 Paris, France; [32]Département de Génétique, Hôpital Necker-Enfants Malades, 75015 Paris, France; [33]Service de Cytogénétique, Hôpital Necker-Enfants Malades, 75015 Paris, France; [34]Paris Descartes-Sorbonne Paris Cité University, Institut Imagine, 75015 Paris, France; [35]GeneDx Inc., 205 Perry Parkway, Gaithersburg, MD 20877, USA; [36]Moores Cancer Center, University of California, San Diego, La Jolla, CA 92093, USA; [37]Department of Pathology, University of California, San Diego, La Jolla, CA 92093, USA; [38]Division of Biological Sciences, University of California, San Diego, La Jolla, CA 92093, USA; [39]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA; [40]Analytical and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA 02114, USA; [41]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA; [42]Department of Clinical Genetics, Erasmus University Medical Center, 3015 CN Rotterdam, the Netherlands; [43]HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA; [44]Department of Biochemistry and Molecular Biology, College of Medicine, University of South Alabama, Mobile, AL 36688, USA
[45]These authors contributed equally to this work
*Correspondence: lisenka.vissers@radboudumc.nl (L.E.L.M.V.), eahn@health.southalabama.edu (E.-Y.E.A.)
http://dx.doi.org/10.1016/j.ajhg.2016.06.029.

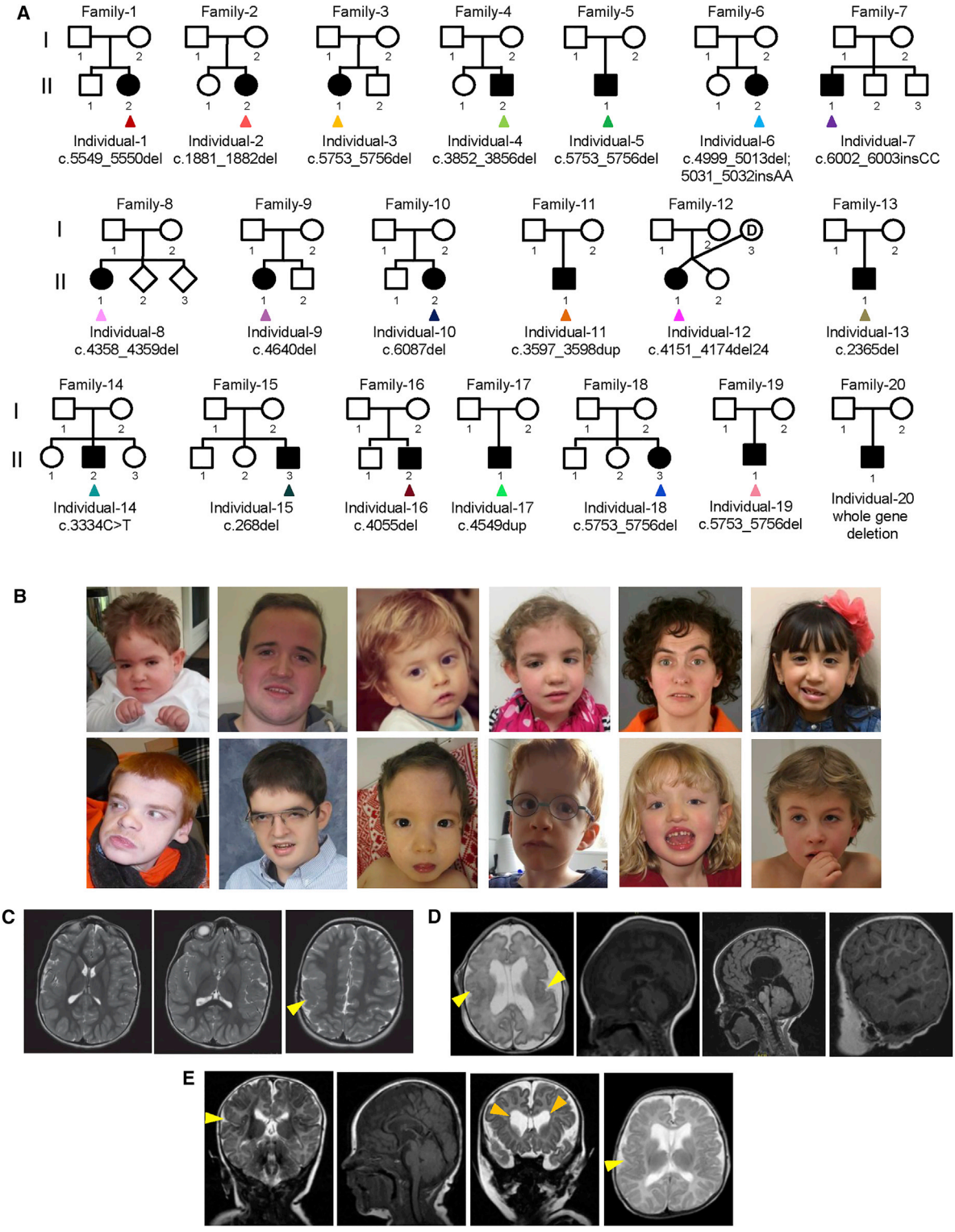

Figure 1. Pedigree Structures, Photos, and Brain MRI of Individuals with *SON* Mutations

(A) Family pedigrees of individuals carrying mutations in *SON*.

(B) Top row from left to right: photos of individuals 2 (at age 5 years), 4 (age 19 years), 5 (age 2 years), 6 (age 6 years), 8 (age 34 years), and 10 (age 6 years). Bottom row from left to right: photos of individuals 11 (age 21 years), 13 (age 14 years), 15 (age 15 months), 16 (age 5 years), 18 (age 6 years), and 19 (age 10 years). Shared facial dysmorphisms include facial asymmetry, midface retraction, low-set ears, downslanting palpebral fissures, deep-set eyes, horizontal eyebrows, a broad and/or depressed nasal bridge, and a short philtrum.

(C) Axial T2-weighted fast spin-echo MRI of the brain of individual 1 at age 3 years. Three panels show ascending images (left to right) revealing that the individual's insular cortex on the right is thickened and featureless. Less impressive areas of similar change were noted in the posterior aspect of the left insular cortex, which revealed bilateral perisylvian and parietal polymicrogyria (yellow arrowhead).

(legend continued on next page)

Table 1. Clinical Features of Individuals with *SON* Haploinsufficiency

	Percentage	Number of Affected Individuals
Intellectual disability	100%	20/20
Brain malformation	89%	17/19
Ventricular enlargement	74%	14/19
Corpus callosum abnormality	53%	10/19
Cortex malformation	37%	7/19
White-matter abnormalities	21%	4/19
Cerebellar abnormalities	21%	4/19
Other	11%	2/19
Neurological features	85%	17/20
Seizures	55%	11/20
Hypotonia	75%	15/20
Musculoskeletal abnormalities	85%	17/20
Hypermobility	40%	8/20
Scoliosis or kyphosis	20%	4/20
Hemivertebrae	10%	2/20
Contractures	10%	2/20
Other	85%	17/20
Eye and/or vision abnormality	75%	15/20
Strabismus	55%	11/20
Suspicion of CVI	20%	4/20
Hypermetropia	30%	6/20
Heart defect	25%	5/20
Gastrointestinal malformation	15%	3/20
Urogenital malformation	30%	6/20
Horseshoe kidney	10%	2/20
Other	20%	4/20
Facial dysmorphism	100%	20/20
Short stature	50%	10/20
Craniosynostosis	15%	3/20

The following abbreviation is used: CVI, cortical visual impairment.

be considered sufficient evidence of causality on their own, they support the hypothesis that *SON* LoF mutations are under strong purifying selection in the human population and that their occurrence most likely contributes to severe clinical phenotypes.

Transcripts bearing a premature stop codon are likely to be degraded by nonsense-mediated mRNA decay. To confirm that LoF mutations result in reduced dosage of *SON*, we used three different PCR primer sets (Table S3) to perform qRT-PCR to determine the amounts of the *SON* transcript in peripheral-blood mononuclear cells (PBMCs) isolated from trio 1 (I-1, I-2, and II-2 in family 1), trio 3 (I-1, I-2, and II-1 in family 3), individual 5 (II-1 in family 5; Figure 1A), and an unrelated healthy donor (Figure 2C). All three primer sets showed that compared to mRNA from the parental samples and the unrelated healthy donor, *SON* mRNA in the affected individuals was significantly downregulated (Figure 2C). Subsequent western blotting using PBMC lysates from trio 1 and two different SON antibodies consistently showed the reduction of SON in individual 1 (Figures 2D and 2E), indicating that de novo *SON* LoF mutations result in *SON* haploinsufficiency.

To examine the effect of *SON* haploinsufficiency on embryonic development, we utilized *Danio rerio* (zebrafish), which has a well-conserved homolog of human *SON* (NCBI Gene: LOC565999; Figures S3 and S4). We assessed the developmental effects of *SON* haploinsufficiency in vivo with morpholino (MO)-mediated knockdown of *son* in zebrafish embryos. Interestingly, embryos injected with a *son* MO had a host of developmental defects that ranged from bent tails (63.6%) to eye malformations and microcephaly (17.1%) and shortened or gnarled tails, deformed body axes, and massive body curvatures (2.1%) 24 hr post-injection (hpi) (Figure 3A and Figure S5). Embryos that survived 72 hpi progressed to more severe phenotypes including extreme spinal malformations (22.2%), head and eye malformations with brain edema (37.2%), and profound developmental abnormalities (10.1%; Figure 3B), mimicking features observed in the affected individuals.

SON is a nuclear speckle protein able to bind to both DNA and RNA, and its cellular functions include regulation of RNA splicing and gene transcription, as well as proper cell-cycle and embryonic stem cell maintenance.[13,22–25] To identify molecular mechanisms underlying the clinical features of individuals with *SON* haploinsufficiency, we examined global expression patterns upon *SON* knockdown in cellular systems. Hereto, we re-analyzed microarray-based RNA-expression profiling and RNA-sequencing datasets generated upon *SON* knockdown in HeLa cells[13,22] and human embryonic

(D) Sagittal T1-weighted and axial T2-weighted MRI of the brain of individual 2. The two images on the left (age 1 day; gestational age 34 + 6 weeks) reveal enlarged lateral ventricles, cavum septum pellucidum, a hypoplastic cerebellar hemisphere, a broad cistern magna, a small fourth ventricle, and a thin corpus callosum. The cortex shows a simplified gyration pattern, and the perisylvian and frontotemporal areas are suspect for polymicrogyria (yellow arrowheads). The two panels on the right (age 2 years) show the fissure Sylvie with an abnormal cortical border, an Arnold Chiari malformation, and hydrocephalus.
(E) Frontal T2-weighted, sagittal T1-weighted, axial T2-weighted, and sagittal T1-weighted MRI of the brain of individual 7 (II-1 in family 7) at the age of 2 months. The cortex shows deep sulci and perisylvian areas suspect for polymicrogyria (yellow arrowheads), as well as discrete heterotopic nodules (orange arrowheads). A thin corpus callosum, a small fourth ventricle, enlarged frontal horns of the lateral ventricles, and cavum septum pellucidum are present.

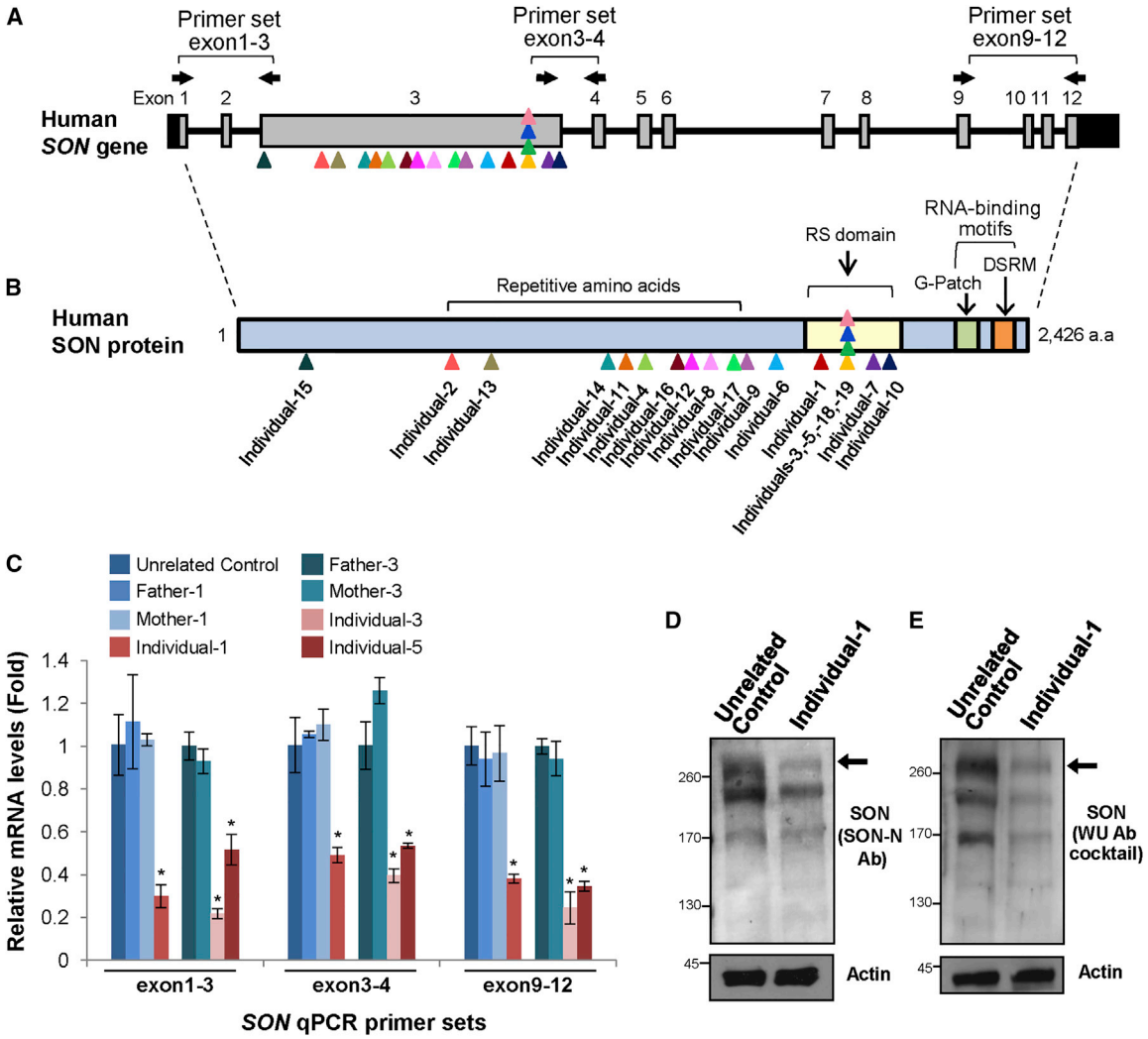

Figure 2. *SON* Mutations and Their Functional Effect at the RNA and Protein Levels
(A and B) Schematic representation of *SON* (A) and SON (B) shows the position of the mutations identified in the 20 affected individuals with color-coded arrowheads. The locations of the PCR primer sets are indicated by black arrows.
(C) Real-time qPCR with three different primer pairs showed that *SON* mRNA from the affected individuals was overall downregulated in comparison to mRNA from the parents and unrelated normal individual. Error bars represent mean ± SD. *$p < 0.001$.
(D and E) Western blotting demonstrated reduced expression of SON. SON-N antibody (1:1,000) was generated against amino acids 74–88 of the human SON (amino acid sequence DTELRYKPDLKEGSR). The cocktail of WU SON antibodies was a mixture of three different SON antibodies (WU09 [1:100], WU14 [1:2,000], and WU21 [1:200]). The epitopes of WU SON antibodies were as follows: MDSQMLATSS for WU09, CEESESKTKSH for WU14, and SMMSAYERS for WU21. SON-N antibody (D) and the cocktail of SON WU antibodies (E) showed similar results. The bands indicated by the black arrow represent full-length SON. Other bands, which could represent potential isoforms, were also detected. Besides the bands present in samples from both normal and affected individuals, no other specific bands were detected in the affected individuals.

stem cells.[25] Surprisingly, from these previous datasets, we noticed that a group of genes playing pivotal roles in neuronal cell migration, embryonic survival, metabolism, and mitochondrial function, including *TUBG1* (MIM: 191135), *FLNA* (MIM: 300017), *PNKP* (MIM: 605610), *WDR62* (MIM: 613583), *PSMD3*, *HDAC6* (MIM: 300272), *PCK2* (MIM: 614095), *PFKL* (MIM: 171860), *IDH2* (MIM: 147650), *ACY1* (MIM: 104620), and *ADA* (MIM: 608958), showed significantly decreased expression upon *SON* knockdown (Tables S4 and S5).[13,22,25] To investigate whether genes involved in regulating brain development and in metabolism are also downregulated in individuals with *SON* LoF mutations, we measured the levels of RNA expression of these genes in PBMCs from trio 1, trio 3, and individual 5, as well as from an unrelated healthy donor (primers are listed in Table S3). Using qPCR analysis, we confirmed that all 11 genes were indeed significantly downregulated in individuals with *SON* haploinsufficiency (Figures 4A and 4B).

SON functions as a splicing co-factor that promotes correct and efficient RNA splicing of weak splice sites and alternative splice sites by facilitating spliceosome recruitment to the elongating RNA polymerase II complex.[13] Prominent features observed upon *SON* knockdown in HeLa cells

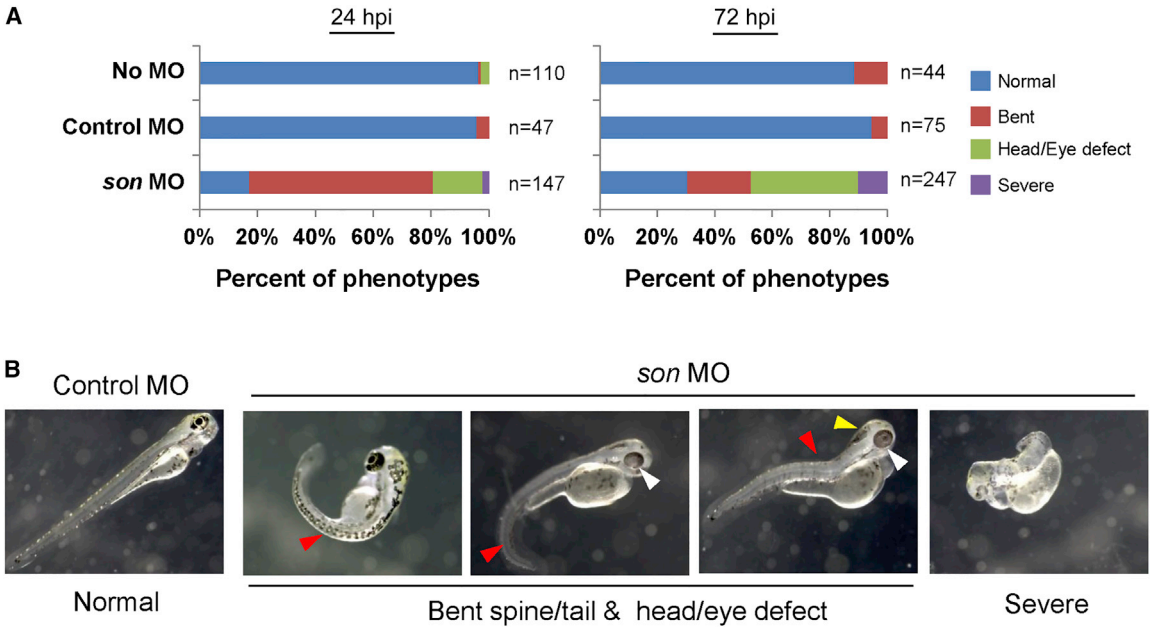

Figure 3. Targeted *son* Knockdown in Developing Zebrafish Causes Impaired Head Development and Spinal Malformations
(A) Zebrafish injected with a splice-blocking *son* morpholino (MO; 5′-TGGTCCTGGATATAACAGACAGATT-3′, 6.25 ng) that targeted the junction between intron 9 and exon 10, a control MO (5′-CCTCTTACCTCAGTTACAATTTATA-3′, 6.25 ng), or no MO showed a normal phenotype, a bent spine or tail, a head or eye defect, or a severe phenotype at 24 and 72 hpi. The percentages of embryos with each phenotype are shown in the bar graphs, and the number of embryos examined is listed next to each bar.
(B) Representative images of the phenotype observed 72 hr after MO injection (red arrow, bent spine or tail; white arrow, eye defects; and yellow arrow, brain edema).

and human embryonic stem cells have included intron retention and exon skipping, which have been shown at the gene level for *TUBG1*, *HDAC6*, and *ADA*.[13,22,25] We next sought to determine whether RNA splicing of these 11 genes is also impaired in our individuals with *SON* haploinsufficiency. To this end, we analyzed the pre-mRNA sequences of the remaining eight genes to predict weak splice sites that could be potential targets of SON-mediated RNA splicing (Table S6). We performed RT-PCR by using DNase-treated RNA samples isolated from trio 1, trio 3, individual 5, and an unrelated healthy donor and using primers designed to detect intron retention and exon skipping (Tables S6 and S7). We not only confirmed that these sites were indeed mis-spliced in HeLa cells upon *SON* knockdown (Figure S6) but also found that all three affected individuals showed significant intron retention (*TUBG1*, *FLNA*, *PNKP*, *WDR62*, *PSMD3*, *PCK2*, *PFKL*, *IDH2*, and *ACY1*) and exon skipping (*HDAC6* and *ADA*) at the predicted sites of the target pre-mRNAs and that this resulted in the accumulation of mis-spliced products (Figures 4C and 4D). In contrast, mis-spliced RNA products were absent in the parents and unrelated donor (Figures 4C and 4D). Together, these results indicate that SON-mediated RNA splicing is severely compromised in individuals with *SON* haploinsufficiency.

Our data have revealed that the complex neurodevelopmental disorder observed in these affected individuals is due to compromised SON function, which causes insufficient production of downstream targets as a result of erroneous SON-mediated RNA splicing. Moreover, the roles of several downregulated genes are well-known causes of ID and/or DD in humans (Tables S4 and S5).[4,6,26–35] For instance, *FLNA* haploinsufficiency is the most common cause of periventricular nodular heterotopia (MIM: 300049),[35] a rare brain malformation that we also found among our cohort with *SON* LoF mutations. Similarly, de novo LoF mutations in *TUBG1* are known to result in cortical malformations (MIM: 615412),[32] also frequently observed in our cohort of affected individuals. Because we have shown that a substantial number of essential developmental genes are significantly downregulated upon *SON* haploinsufficiency, SON thus represents a master regulator of genes essential for human neurodevelopmental processes.

In summary, we have identified de novo LoF mutations in *SON* as a cause of a complex neurodevelopmental disorder associated with ID and/or DD and severe brain malformations. In addition, we have revealed the underlying molecular mechanism by showing that *SON* haploinsufficiency leads to defective RNA splicing of multiple genes critical for brain development, neuronal migration, and metabolism. Our findings thus greatly contribute to our understanding of how defective RNA splicing leads to human neurodevelopmental disorders.

Supplemental Data

Supplemental Data include six figures and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2016.06.029.

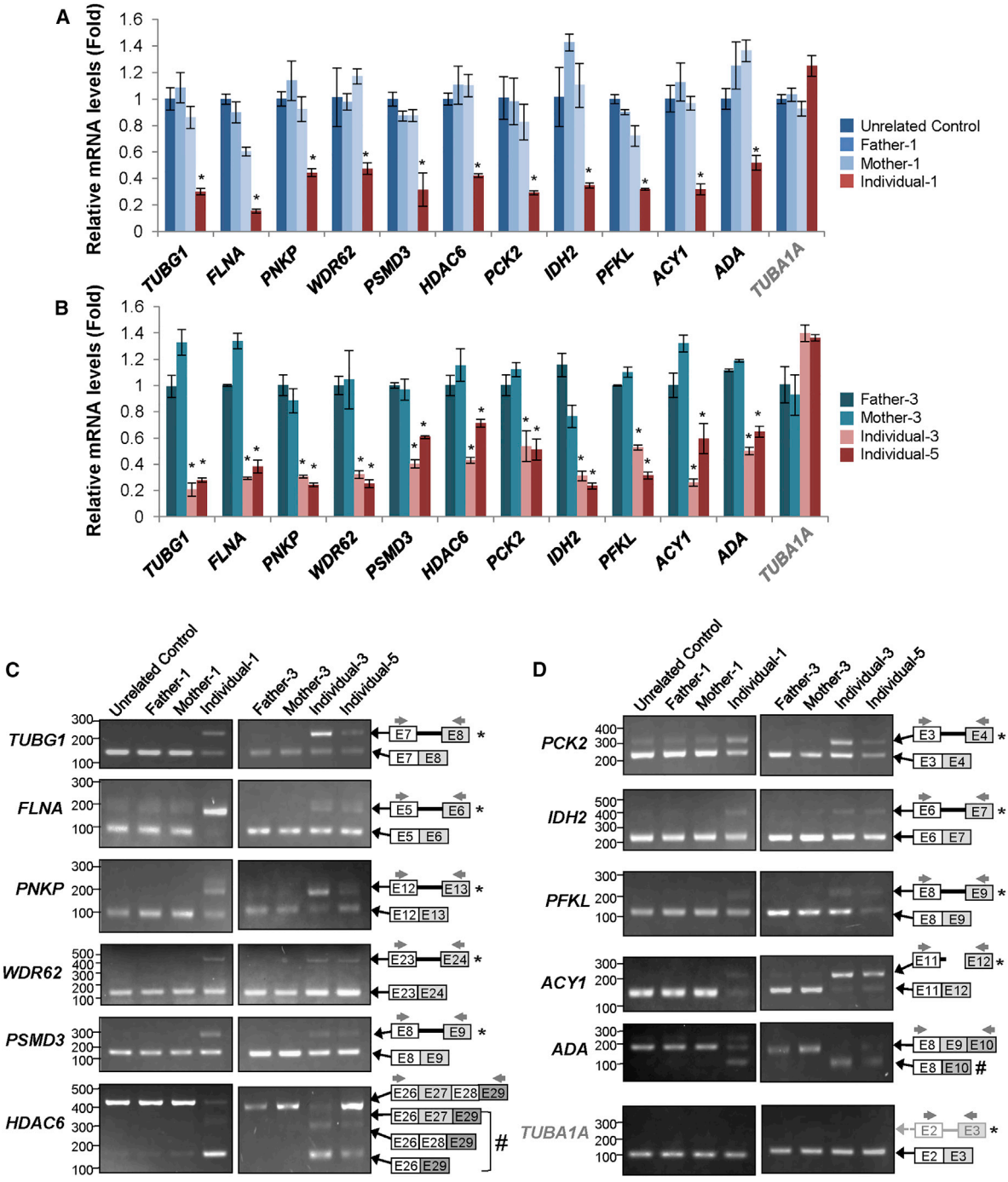

Figure 4. Individuals Carrying Heterozygous *SON* LoF Mutations Have Defective RNA Splicing of Genes Associated with the Pathophysiology of ID and/or DD and Metabolic Disorders, Resulting in Their Reduced Expression

(A and B) Multiple genes associated with the pathophysiology of ID and/or DD (A) and metabolic disorders (B) in the affected individuals were downregulated in comparison to genes from parents and unrelated healthy individuals. *TUBA1A* mRNA served as a negative control (unaffected transcript). Error bars represent mean ± SD. *$p < 0.001$.

(C and D) Intron retention and exon skipping of genes involved in ID and/or DD when mutated (C) and genes involved in metabolic disorders when mutated (D) in the individuals with *SON* mutations. The locations of the primers used for PCR are marked by gray arrows above the exons. Analysis of *TUBA1A* pre-mRNA, which served as a negative control, demonstrated that splicing of this transcript is not impaired in the affected individuals. *, intron-retained products; #, exon-skipped products.

Conflicts of Interest

D.N.S., S.T., and D.E. are employees of Ambry Genetics, Inc. M.T.C., K.M., G.D., B.B., A.B., and T.F. are employees of GeneDx, Inc.

Acknowledgments

We are deeply grateful to all individuals and their families for participating in this study. We thank Dr. Michael Markey (Wright State University) for transcript splicing analysis, Dr. David Goldstein,

Jennifer Sullivan, and Nicole Walley (Duke University) for their help in collecting clinical information, Petra de Vries (Radboudumc) and Hermann-Josef Lüdecke (University Duisburg-Essen) for technical support, and Dr. Christian Gilissen (Radboudumc) for bioinformatic support. We thank the University of Washington Center for Mendelian Genomics (funded by the National Human Genome Research Institute and National Heart, Lung, and Blood Institute grant 1U54HG006493 to Drs. Debbie Nickerson, Jay Shendure, and Michael Bamshad) and the Radboudumc Genomics Technology Center for family-based sequencing. This work was supported by NIH grants (CA190688 and CA185818 to E.-Y.E.A. and GM084407 to P.A.B.), a start-up fund from the University of South Alabama Mitchell Cancer Institute (to E.-Y.E.A.), the Ohio Board of Regents (to P.A.B.), a California State University Chico Internal Research Grant (to D.L.S.), grants from Stichting ODAS and Vereniging Bartiméus-Sonneheerdt (5781251 to B.B.A.d.V. and D.G.M.B.), and a grant from the Dutch Organization for Health Research and Development (912-12-109 to B.B.A.d.V. and J.A.V.). This study used data generated by the DECIPHER Consortium. A full list of centers who contributed to the generation of the data is available at http://decipher.sanger.ac.uk and via email at decipher@sanger.ac.uk. Funding for the DECIPHER project was provided by the Wellcome Trust.

Received: May 4, 2016
Accepted: June 24, 2016
Published: August 18, 2016

Web Resources

Clinical Genome Database, http://research.nhgri.nih.gov/CGD/
DECIPHER, https://decipher.sanger.ac.uk/
ESEfinder, http://rulai.cshl.edu/tools/ESE/
ExAC Browser, http://exac.broadinstitute.org/
Genic Intolerance, http://genic-intolerance.org/
NCBI Gene, http://www.ncbi.nlm.nih.gov/gene
OMIM, http://www.omim.org/
RefSeq, http://www.ncbi.nlm.nih.gov/refseq/

References

1. Ropers, H.H. (2010). Genetics of early onset cognitive impairment. Annu. Rev. Genomics Hum. Genet. *11*, 161–187.
2. Mefford, H.C., Batshaw, M.L., and Hoffman, E.P. (2012). Genomics, intellectual disability, and autism. N. Engl. J. Med. *366*, 733–743.
3. Vissers, L.E., de Ligt, J., Gilissen, C., Janssen, I., Steehouwer, M., de Vries, P., van Lier, B., Arts, P., Wieskamp, N., del Rosario, M., et al. (2010). A de novo paradigm for mental retardation. Nat. Genet. *42*, 1109–1112.
4. Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T., Albrecht, B., Bartholdi, D., Beygo, J., Di Donato, N., et al. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. Lancet *380*, 1674–1682.
5. Gilissen, C., Hehir-Kwa, J.Y., Thung, D.T., van de Vorst, M., van Bon, B.W., Willemsen, M.H., Kwint, M., Janssen, I.M., Hoischen, A., Schenck, A., et al. (2014). Genome sequencing identifies major causes of severe intellectual disability. Nature *511*, 344–347.
6. Deciphering Developmental Disorders Study (2015). Large-scale discovery of novel genetic causes of developmental disorders. Nature *519*, 223–228.
7. de Ligt, J., Willemsen, M.H., van Bon, B.W., Kleefstra, T., Yntema, H.G., Kroes, T., Vulto-van Silfhout, A.T., Koolen, D.A., de Vries, P., Gilissen, C., et al. (2012). Diagnostic exome sequencing in persons with severe intellectual disability. N. Engl. J. Med. *367*, 1921–1929.
8. Vissers, L.E., Gilissen, C., and Veltman, J.A. (2016). Genetic studies in intellectual disability and related disorders. Nat. Rev. Genet. *17*, 9–18.
9. Lelieveld, S.H., Reijnders, M.R.F., Pfundt, R., Yntema, H.G., Kamsteeg, E.-J., de Vries, P., de Vries, B.B.A., Willemsen, M.H., Kleefstra, T., Löhner, K., et al. (2016). Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. Nat. Neurosci. Published online August 1, 2016. http://dx.doi.org/10.1038/nn.4352.
10. Zhu, X., Petrovski, S., Xie, P., Ruzzo, E.K., Lu, Y.F., McSweeney, K.M., Ben-Zeev, B., Nissenkorn, A., Anikster, Y., Oz-Levi, D., et al. (2015). Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. Genet. Med. *17*, 774–781.
11. MacArthur, D.G., Manolio, T.A., Dimmock, D.P., Rehm, H.L., Shendure, J., Abecasis, G.R., Adams, D.R., Altman, R.B., Antonarakis, S.E., Ashley, E.A., et al. (2014). Guidelines for investigating causality of sequence variants in human disease. Nature *508*, 469–476.
12. Sharma, A., Takata, H., Shibahara, K., Bubulya, A., and Bubulya, P.A. (2010). Son is essential for nuclear speckle organization and cell cycle progression. Mol. Biol. Cell *21*, 650–663.
13. Ahn, E.Y., DeKelver, R.C., Lo, M.C., Nguyen, T.A., Matsuura, S., Boyapati, A., Pandit, S., Fu, X.D., and Zhang, D.E. (2011). SON controls cell-cycle progression by coordinated regulation of RNA splicing. Mol. Cell *42*, 185–198.
14. Hickey, C.J., Kim, J.H., and Ahn, E.Y. (2014). New discoveries of old SON: a link between RNA splicing and cancer. J. Cell. Biochem. *115*, 224–231.
15. Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature *515*, 216–221.
16. Genome of the Netherlands Consortium (2014). Whole-genome sequence variation, population structure and demographic history of the Dutch population. Nat. Genet. *46*, 818–825.
17. Gulsuner, S., Walsh, T., Watts, A.C., Lee, M.K., Thornton, A.M., Casadei, S., Rippey, C., Shahin, H., Nimgaonkar, V.L., Go, R.C., et al.; Consortium on the Genetics of Schizophrenia (COGS); PAARTNERS Study Group (2013). Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. Cell *154*, 518–529.
18. Xu, B., Ionita-Laza, I., Roos, J.L., Boone, B., Woodrick, S., Sun, Y., Levy, S., Gogos, J.A., and Karayiorgou, M. (2012). De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. Nat. Genet. *44*, 1365–1369.
19. Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet. *9*, e1003709.
20. Petrovski, S., Gussow, A.B., Wang, Q., Halvorsen, M., Han, Y., Weir, W.H., Allen, A.S., and Goldstein, D.B. (2015). The Intolerance of Regulatory Sequence to Genetic Variation Predicts Gene Dosage Sensitivity. PLoS Genet. *11*, e1005492.
21. Exome Aggregation Consorsium (2016). Analysis of protein-coding genetic variation in 60,706 humans. bioRxiv http://dx.doi.org/10.1101/030338.

22. Sharma, A., Markey, M., Torres-Muñoz, K., Varia, S., Kadakia, M., Bubulya, A., and Bubulya, P.A. (2011). Son maintains accurate splicing for a subset of human pre-mRNAs. J. Cell Sci. *124*, 4286–4298.
23. Kim, J.H., Baddoo, M.C., Park, E.Y., Stone, J.K., Park, H., Butler, T.W., Huang, G., Yan, X., Pauli-Behn, F., Myers, R.M., et al. (2016). SON and Its Alternatively Spliced Isoforms Control MLL Complex-Mediated H3K4me3 and Transcription of Leukemia-Associated Genes. Mol. Cell *61*, 859–873.
24. Ahn, E.E., Higashi, T., Yan, M., Matsuura, S., Hickey, C.J., Lo, M.C., Shia, W.J., DeKelver, R.C., and Zhang, D.E. (2013). SON protein regulates GATA-2 through transcriptional control of the microRNA 23a~27a~24-2 cluster. J. Biol. Chem. *288*, 5381–5388.
25. Lu, X., Göke, J., Sachs, F., Jacques, P.E., Liang, H., Feng, B., Bourque, G., Bubulya, P.A., and Ng, H.H. (2013). SON connects the splicing-regulatory network with pluripotency in human embryonic stem cells. Nat. Cell Biol. *15*, 1141–1152.
26. Bilgüvar, K., Oztürk, A.K., Louvi, A., Kwan, K.Y., Choi, M., Tatli, B., Yalnizoğlu, D., Tüysüz, B., Cağlayan, A.O., Gökben, S., et al. (2010). Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. Nature *467*, 207–210.
27. Chen, J.F., Zhang, Y., Wilde, J., Hansen, K.C., Lai, F., and Niswander, L. (2014). Microcephaly disease gene Wdr62 regulates mitotic progression of embryonic neural stem cells and brain size. Nat. Commun. *5*, 3885.
28. Jamuar, S.S., Lam, A.T., Kircher, M., D'Gama, A.M., Wang, J., Barry, B.J., Zhang, X., Hill, R.S., Partlow, J.N., Rozzo, A., et al. (2014). Somatic mutations in cerebral cortical malformations. N. Engl. J. Med. *371*, 733–743.
29. Jensen, L.R., Lenzner, S., Moser, B., Freude, K., Tzschach, A., Wei, C., Fryns, J.P., Chelly, J., Turner, G., Moraine, C., et al. (2007). X-linked mental retardation: a comprehensive molecular screen of 47 candidate genes from a 7.4 Mb interval in Xp11. Eur. J. Hum. Genet. *15*, 68–75.
30. Nicholas, A.K., Khurshid, M., Désir, J., Carvalho, O.P., Cox, J.J., Thornton, G., Kausar, R., Ansar, M., Ahmad, W., Verloes, A., et al. (2010). WDR62 is associated with the spindle pole and is mutated in human microcephaly. Nat. Genet. *42*, 1010–1014.
31. Nofech-Mozes, Y., Blaser, S.I., Kobayashi, J., Grunebaum, E., and Roifman, C.M. (2007). Neurologic abnormalities in patients with adenosine deaminase deficiency. Pediatr. Neurol. *37*, 218–221.
32. Poirier, K., Lebrun, N., Broix, L., Tian, G., Saillour, Y., Boscheron, C., Parrini, E., Valence, S., Pierre, B.S., Oger, M., et al. (2013). Mutations in TUBG1, DYNC1H1, KIF5C and KIF2A cause malformations of cortical development and microcephaly. Nat. Genet. *45*, 639–647.
33. Shen, J., Gilmore, E.C., Marshall, C.A., Haddadin, M., Reynolds, J.J., Eyaid, W., Bodell, A., Barry, B., Gleason, D., Allen, K., et al. (2010). Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. Nat. Genet. *42*, 245–249.
34. Yu, T.W., Mochida, G.H., Tischfield, D.J., Sgaier, S.K., Flores-Sarnat, L., Sergi, C.M., Topçu, M., McDonald, M.T., Barry, B.J., Felie, J.M., et al. (2010). Mutations in WDR62, encoding a centrosome-associated protein, cause microcephaly with simplified gyri and abnormal cortical architecture. Nat. Genet. *42*, 1015–1020.
35. Fox, J.W., Lamperti, E.D., Ekşioğlu, Y.Z., Hong, S.E., Feng, Y., Graham, D.A., Scheffer, I.E., Dobyns, W.B., Hirsch, B.A., Radtke, R.A., et al. (1998). Mutations in filamin 1 prevent migration of cerebral cortical neurons in human periventricular heterotopia. Neuron *21*, 1315–1325.

The American Journal of Human Genetics, Volume *99*

Supplemental Data

De Novo Mutations in *SON* Disrupt RNA Splicing of Genes Essential for Brain Development and Metabolism, Causing an Intellectual-Disability Syndrome

Jung-Hyun Kim, Deepali N. Shinde, Margot R.F. Reijnders, Natalie S. Hauser, Rebecca L. Belmonte, Gregory R. Wilson, Daniëlle G.M. Bosch, Paula A. Bubulya, Vandana Shashi, Slavé Petrovski, Joshua K. Stone, Eun Young Park, Joris A. Veltman, Margje Sinnema, Connie T.R.M. Stumpel, Jos M. Draaisma, Joost Nicolai, University of Washington Center for Mendelian Genomics, Helger G. Yntema, Kristin Lindstrom, Bert B.A. de Vries, Tamison Jewett, Stephanie L. Santoro, Julie Vogt, Deciphering Developmental Disorders Study, Kristine K. Bachman, Andrea H. Seeley, Alyson Krokosky, Clesson Turner, Luis Rohena, Maja Hempel, Fanny Kortüm, Davor Lessel, Axel Neu, Tim M. Strom, Dagmar Wieczorek, Nuria Bramswig, Franco A. Laccone, Jana Behunova, Helga Rehder, Christopher T. Gordon, Marlène Rio, Serge Romana, Sha Tang, Dima El-Khechen, Megan T. Cho, Kirsty McWalter, Ganka Douglas, Berivan Baskin, Amber Begtrup, Tara Funari, Kelly Schoch, Alexander P.A. Stegmann, Servi J.C. Stevens, Dong-Er Zhang, David Traver, Xu Yao, Daniel G. MacArthur, Han G. Brunner, Grazia M. Mancini, Richard M. Myers, Laurie B. Owen, Ssang-Taek Lim, David L. Stachura, Lisenka E.L.M. Vissers, and Eun-Young Erin Ahn

Supplemental Figures

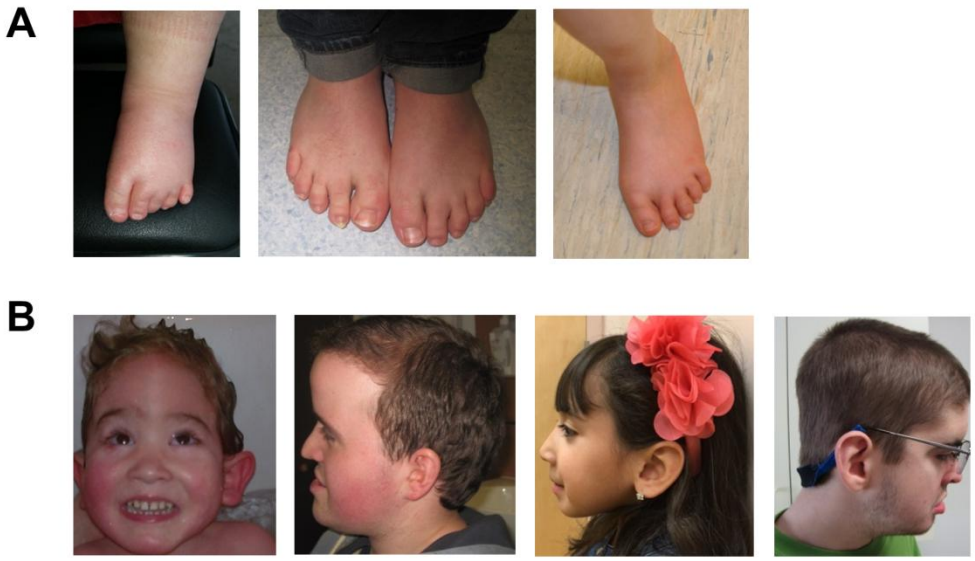

Figure S1. Photographs of individuals with *SON* mutations.

(**A**) Photos showing small feet and short toes of Individual-2 at age 9 years, Individual-4 at age 19 years, and Individual-19 at age 10 years (from left to right). (**B**) Photos showing low set ears in Individual-2, Individual-4, Individual-10 and Individual-13 (from left to right). In addition, midface retraction and short philtrum are noticeable in Individual-4 and Individual-13.

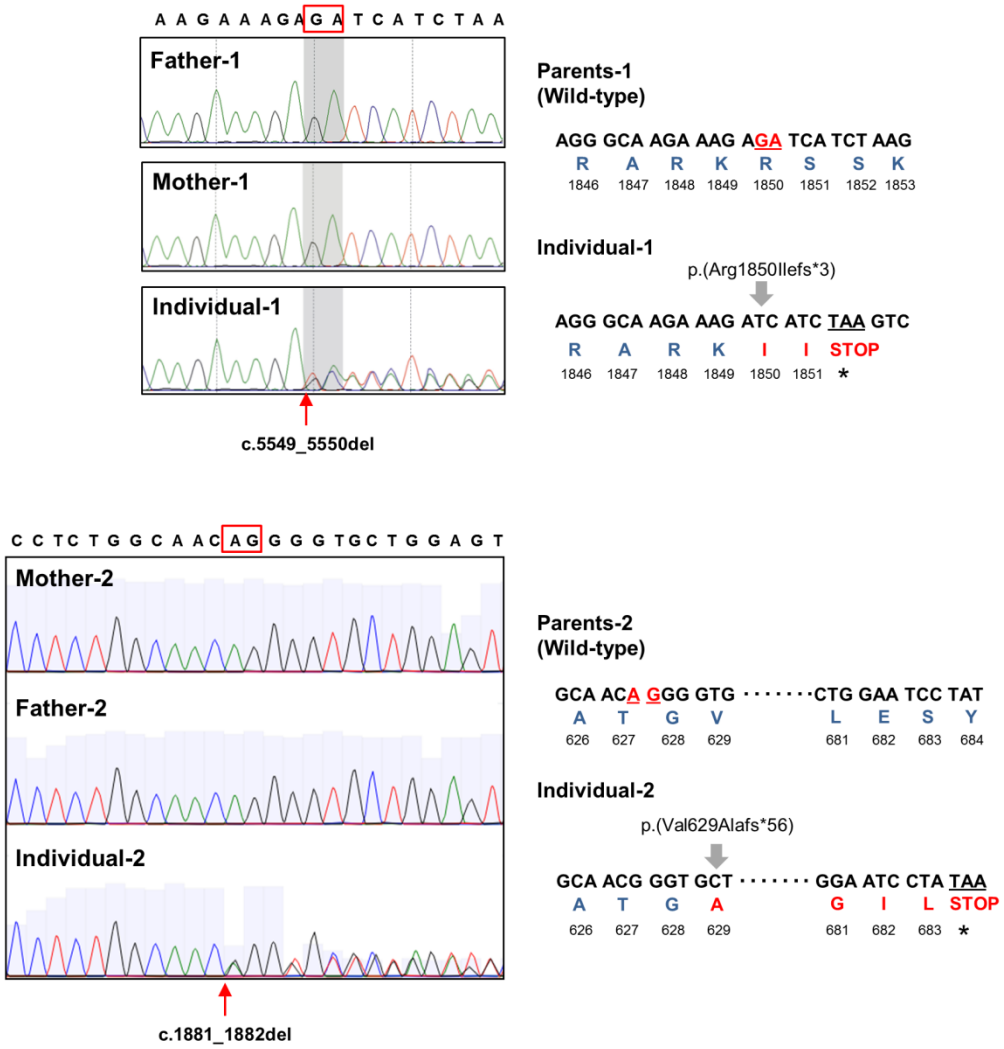

Figure S2. Sequence chromatogram of two families with *SON* heterozygous mutations.

The sequence chromatograms (left panel) of Trio-1 and Trio-2. Heterozygous nucleotide deletion was detected at c.5549_5550 (Individual-1; top panel) and c.1881_1882 (Individual-2; bottom panel) position of *SON* transcript (NM_138927.2). Mutated (or deleted) sites are indicated in the chromatograms by red arrows. The predicted amino acids (right panel) corresponding to each codon are represented. For each sequence, the first mutated amino acid sequence is marked by gray arrow, and mutated amino acids are depicted in red.

Figure S3. Alignment of human and zebrafish SON proteins

Homo sapiens SON isoform F (GenBank accession number: NP_620305) and Danio rerio son isoform X1 (GenBank accession number: XP_694356) were aligned with MUSCLE[36] using default parameters. Alignment was visualized in Jalview[37]. Sequence homology is indicated using BLOSUM62 shading. Sequence identity for total protein is 47%, whereas the RS domain, G-patch and DSRM have a percentage of sequence identity of 48%, 79% and 67% respectively. The total homology for the C-terminus – from the RS domain to the end – is 59%.

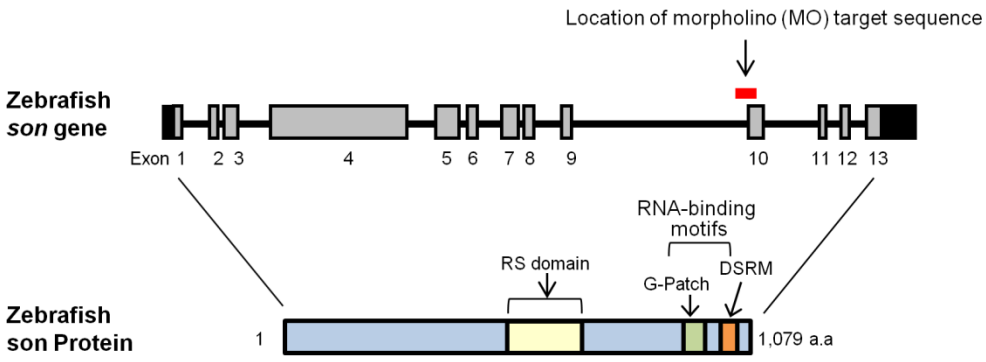

Figure S4. Schematic representation of the zebrafish *son* gene and son protein.

Top: schematic representation of the zebrafish *son* gene (Gene ID: LOC565999, GenBank accession number: XM_689264). The zebrafish *son* contains 13 exons (gray boxes) and several introns (bold lines). Red bar indicates target sequence (intron9-exon 10 junction) of morpholino.

Bottom: Structural feature of the zebrafish son protein (GenBank accession number: XP_694356).

Similar to human SON, zebrafish son has the RS domain (Ser/Arg-rich domain) and two RNA-binding motifs (G-Patch and DSRM).

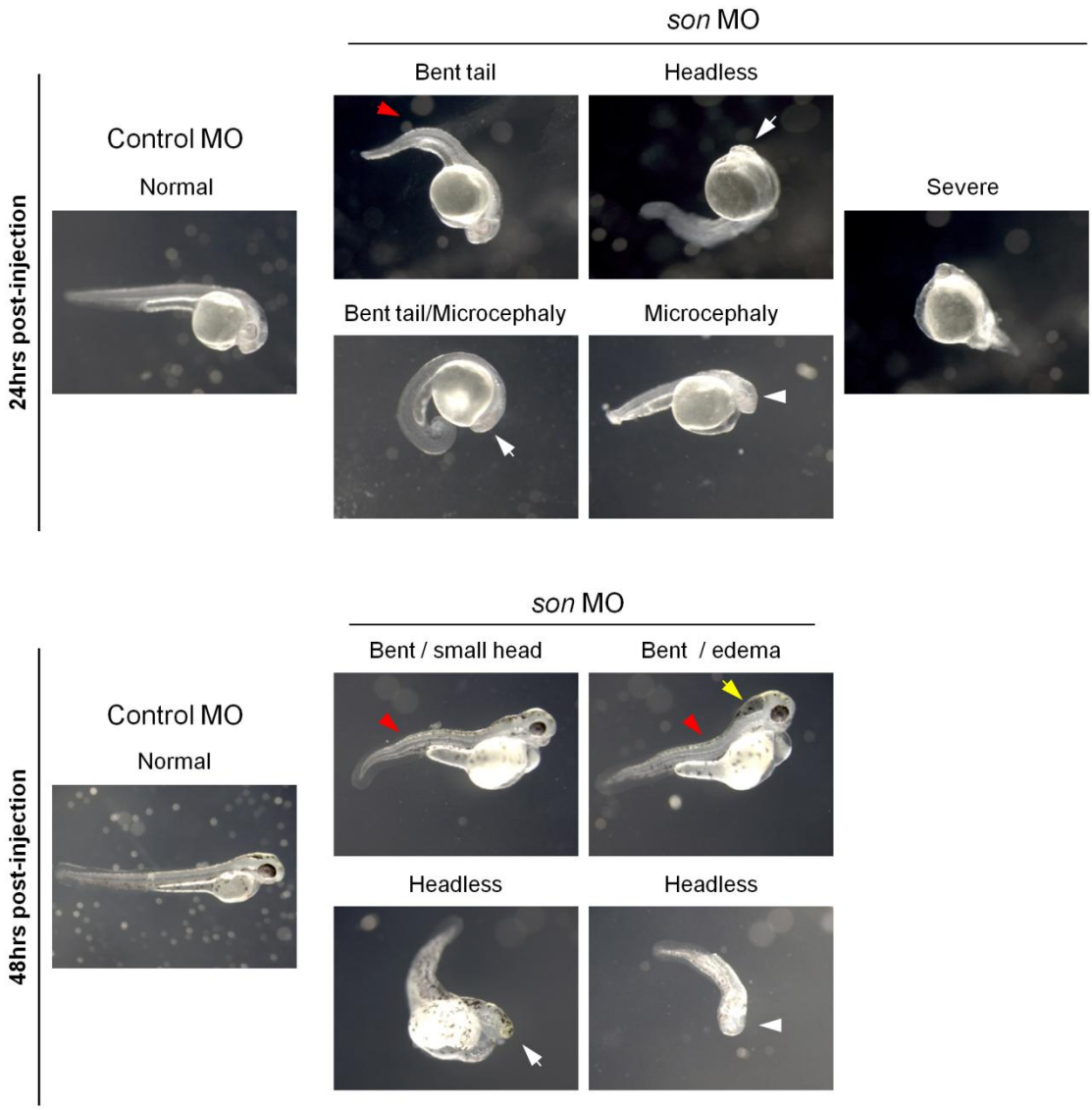

Figure S5. Targeted *son* knockdown in developing zebrafish causes impaired head development and spinal malformations.

Representative images of phenotype observed 24 hours and 48 hours after MO injection. Zebrafish injected with *son* morpholino (MO, 6.25ng) showed bent spine/tail (indicated by red arrow), head/eye defect, no head or small head (white arrow) and brain edema (yellow arrow). Severely malformed fish and headless fish were also observed.

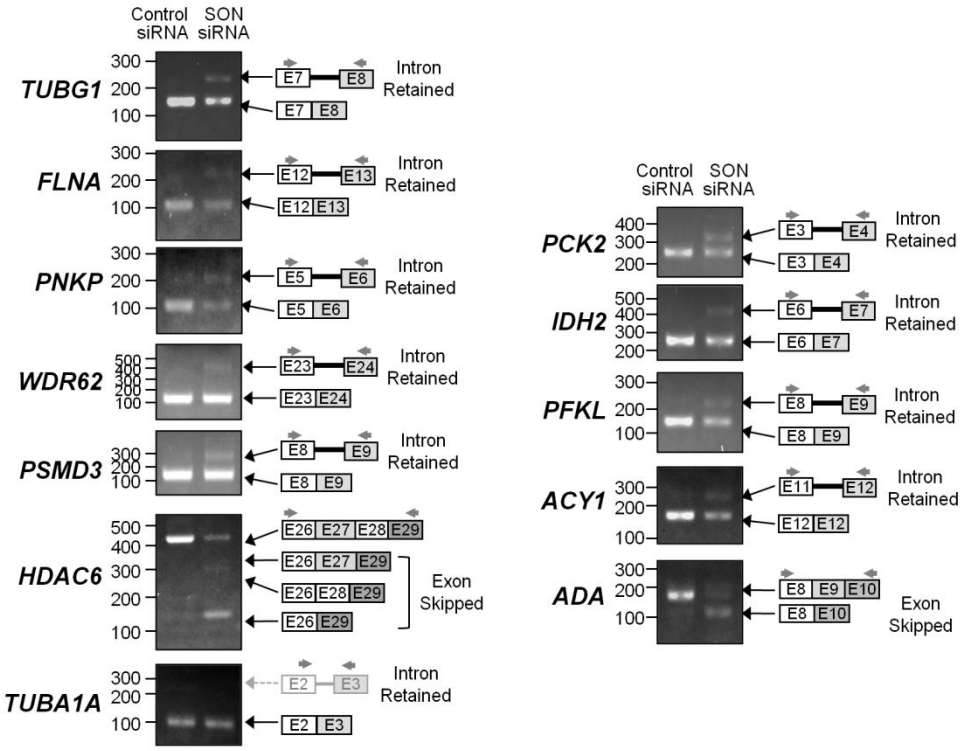

Figure S6. *SON* knockdown effect on intron retention and exon skipping of SON target genes in HeLa cells.

RT-PCR was performed to detect splicing defects of selected SON target genes in control or SON siRNA-transfected HeLa cells. Increased intron retention of nine SON target genes (*TUBG1, FLNA, PNKP, WDR62, PSMD3, PCK2, IDH2, PFKL* and *ACY1*) and exon skipping of two SON target genes (*HDAC6* and *ADA*) were detected in SON siRNA-transfected HeLa cells using specific primer sets (indicated by gray arrows above the exons). *TUBA1A* splicing was analyzed as a negative control that does not require SON for correct splicing.

Supplemental Tables

Table S1. Mutations identified in *SON*

Individual	Gender	Age	cDNA (NM_138927.2)	Protein	Origin
Individual-1	F	5 yrs	c.5549_5550del	p.(Arg1850Ilefs*3)	*De novo*
Individual-2*	F	9 yrs	c.1881_1882del	p.(Val629Alafs*56)	*De novo*
Individual-3*	F	4 yrs 11 mos	c.5753_5756del	p.(Val1918Glufs*87)	*De novo*
Individual-4	M	19 yrs	c.3852_3856del	p.(Met1284Ilefs*2)	*De novo*
Individual-5	M	6 yrs	c.5753_5756del	p.(Val1918Glufs*87)	*De novo*
Individual-6	F	8 yrs	c.4999_5013del; 5031_5032insAA	p.([Asp1667_Asn1671del; Asp1678Lysfs*9])	*De novo*
Individual-7	M	6 yrs 5 mos	c.6002_6003insCC	p.(Arg2002Glnfs*5)	*De novo*
Individual-8	F	32 yrs	c.4358_4359del	p.(Thr1453Serfs*11)	*De novo*
Individual-9	F	7 yrs	c.4640del	p.(His1547Leufs*76)	*De novo*
Individual-10	F	6 yrs 10 mos	c.6087del	p.(Ser2029Argfs*22)	*De novo*
Individual-11	M	21 yrs	c.3597_3598dup	p.(Pro1200Argfs*17)	*De novo*
Individual-12	F	2 yrs 9 mos	c.4151_4174del24	p.(Leu1384_Val1391del)	*Unknown***
Individual-13	M	14 yrs	c.2365del	p.(Ser789Alafs*8)	*De novo*
Individual-14	M	7 yrs	c.3334C>T	p.(Arg1112*)	*De novo*
Individual-15	M	1 yr 9 mos	c.268del	p.(Ser90Valfs*59)	*De novo*
Individual-16	M	4 yrs 9 mos	c.4055del	p.(Pro1352Glnfs*14)	*De novo*
Individual-17	M	4 yrs 11 mos	c.4549dup	p.(Glu1517Glyfs*6)	*De novo*
Individual-18	F	10 yrs 5mos	c.5753_5756del	p.(Val1918Glufs*87)	*De novo*
Individual-19	M	10 yrs	c.5753_5756del	p.(Val1918Glufs*87)	*De novo*
Individual-20	M	10 yrs	Whole gene deletion***	-	*De novo*

M: Male; F: Female; yrs: years; mos: months; **Individual-2 from Gilissen et al. 2014; Individual-3 from Zhu et al. 2015; **Individual-12 is conceived using an ovum donor, but mutation is absent in father and twin sister.***This individual carries a small deletion copy number variant including SON and five other genes: [hg19] chr21:g.(34877993_34894566-35278567_3559909)del (ISCN arr 21q22.11q22.11(34894566-3527867)x1 dn.*

Table S2. Detailed clinical features of the individuals carrying SON mutations

Excel File

Table S3. Primers used for qPCR analysis

Gene	Forward primer (5'→3')	Reverse primer (5'→3')
SON (Exon1-3)	CAGATTTTTAGGTCTTTCGTGGT	TTTTTCTGGAGCCCTCTTTC
SON (Exon3-4)	TCTCCCATCCGTCGTAAAAG	GGTTTGGTGGTAAAGGGACA
SON (Exon9-12)	GCAATAAGGAACCCATCCTAGT	AGGGCTCCATTTCTCAATACC
TUBG1	AGCTGGTGTCTACCATCATGT	CGTAGTGAGAGGGGTGTAGC
FLNA	GCCAAATGCAGCTTGAGAAC	AGGATCAGGGTCCAGATGAG
PNKP	ATCTTCCTGCCCTCGGA	GGGTTAACTCCCAGCTGTTT
WDR62	GTCCATGGGCTACCAACAT	AGGGCAATGACTCTACAAGATAC
PSMD3	CGCCTCAACCACTATGTTCT	CGGAACTGTAAATCAGCCTCT
HDAC6	AAGAAGACCTAATCGTGGGACT	GCTGTGAACCAACATCAGCTC
TUBA1A	GCAAAGAAGATGCTGCCAATAA	CAACTGAGAGACGTTCCATGAG
PCK2	GGCTGAGAATACTGCCACACT	ACCGTCTTGCTCTCTACTCGT
IDH2	CGGCACTTTCAAAATGGTCT	GCATACTGGAAGCAGCTGTG
PFKL	TCGACTGCAGGACCAATGTC	AGCTTCTCCGACAACCACAG
ACY1	CGGGTCTGCAAGGATATGAA	TCATGGGTGAGAAGCCTAGA
ADA	GCCTTCGACAAGCCCAAAGTA	CTCTGCTGTGTTAGCTGGGAG
GAPDH	GGCGCTGAGTACGTCGTGGAGTCCA	AAAGTTGTCATGGATGACCTTGG

Table S4. Significantly downregulated genes in all three studies on *SON* knockdown

Excel file.

A: List of significantly downregulated genes in all three studies on *SON* knockdown (Ahn *et al*.; Sharma *et al.*; Lu *et al*.) that have previously been implicated of human disease based on the clinical genomic database of the National Human Genome Research Institute.

B: List of significantly downregulated genes in all three studies upon *SON* knockdown (Ahn *et al*.; Sharma *et al.*; Lu *et al*.) without prior implication in human disease.

Table S5. Selected SON target genes that are directly implicated in ID / DD and metabolism.

Gene name (symbol)	Gene functions	Mutations previously found in affected individuals	Individual phenotype caused by mutations	Downregulated by *SON* knockdown?			References
				Ahn *et al*	Sharma *et al*	Lu *et al.*	
TUBG1	-Microtubule nucleation -Proper mitotic spindle formation -Knockdown causes impaired migration of neuronal precursors in developing mouse brain	Missense	Post. Agyria and frontal pachygyria Microcephaly	Yes	Yes	Yes	Poirier *et al*. (2013)[32]
FLNA	- Actin-binding and crosslinking actin filaments - Cytoskeleton remodeling	Missense	Periventricular nodular heterotopia	Yes	Yes	Yes	Jamuar *et al*. (2014)[28] Fox *et al*. (1998)[35]
PNKP	- Nucleotide-excision repair, DNA damage removal -Knockdown causes apoptosis of neuronal precursors and differentiated neurons	Missense Frameshift	Microcephaly, seizure Developmental delay	Yes	Yes	Yes	Shen *et al*. (2010)[33]
WDR62	- A mitotic spindle pole protein - Spindle organization and mitotic progression - Knockdown leads mitotic arrest and cell death of neural progenitor cells in mouse brain	Missense Frameshift	Microcephaly Pachygyria, Hypoplasia, Lissencephaly,Schizencephaly, Polymicrogyria	Yes	Yes	Yes	Bilguvar *et al*. (2010)[26] Nicholas *et al*. (2010)[30] Yu *et al*. (2010)[34]
PSMD3	- Proteosomal degradation - A subunit of the 26S proteosome complex - Knockdown shows smaller ear primordia and mis-patterned central nervous system neurons in zebrafish	Missense	Microcephaly, hypotonia, seizure, growth abnormality	Yes	Yes	Yes	DDD Study (2015)[6]
HDAC6	- Deacetylation of histones and other proteins - α-tubulin deacetylation to promote cell motility and proper microtubule network	Missense	Intellectual disability	Yes	Yes	Yes	Jensen *et al*. (2007)[29]
ACY1	- Hydrolysis of acylated L-amino acids	Missense Frameshift	Autistic feature, dystonia, hypotonia	Yes	Yes	Yes	Rauch et al. (2012)[4]
ADA	- Catalysis of the hydrolysis of adenosine to inosine	Deficiency (missense?)	Hypotonia, nystagmus, seizure, developmental delay	-	Yes	-	Nofech-Mozes *et al.* (2007)[31]
PCK2	- Catalysis of the conversion of oxaloacetate to phosphoenolpyruvate in mitochondria - A rate-limiting enzyme in gluconeogenesis	N/A	N/A	Yes	Yes	-	-
IDH2	- Catalysis of the oxidative decarboxylation of isocitrate to produce α–ketoglutarate (2-oxoglutarate) during the TCA cycle in mitochondria	Specific missense	Rare neurometabolic disorder characterized by supraphysiological levels of D-2-HG	Yes	Yes	Yes	Kranendijk et al.(2010)[38]
PFKL	- Catalysis of the conversion of D-fructose 6-phosphate to D-fructose 1,6-bisphosphate - A rate-limiting enzyme in glycolysis	N/A	N/A	Yes	Yes	Yes	-

TUBG1 = tubulin, gamma-1; *FLNA* = filamin A; *PNKP* = polynucleotide kinase 3'-phosphatase; *WDR62* = WD repeat domain 62; *PSMD3* = proteasome 26S subunit, non-ATPase 3; *HDAC6* = histone deaceylase 6; *ACY1* = aminoacylase 1; *ADA* = adenosine deaminase; *PCK2* = phosphoenolpyruvate carboxykinase 2 (mitochondrial); *IDH2* = isocitrate dehydrogenase 2 (NADP+), mitochondrial; *PFKL* = phosphofructokinase, liver.

Table S6. Splice site score analysis of intron-retained genes by *SON* knockdown (5' splice site score / 3' splice site score) and the primer sequences (bold, underlined) used for RT-PCR to detect splicing defects

Target Gene (Gene symbol)	Primer location	Target sequence (Exon; Upper case / Intron; lower case)	Splice site score	PCR product size (bp)	
				Un spliced	Spliced
WDR62	Exon 23 Exon 24	CACCCCTCCTTCCTGCCCCAGCAGAAGGAATCATCTGAGGCCA**GTGAGCTCATCCTCTACTCTCT**GGAGGCAGAAGTGACAGTCACAGGGACAGACAGgtgggtgtcctttccaccaagggagccttagttggaggaaccccagctgatagctgcatcctggaagaagtgctctctg….	12.59	484	146
		….tccgggtggggctagctgttgagtctccagctgaagtccttgttccctctctgcccccactggcactgcagCCAGTATTGCAGGAAGGAGGTGGAGGCCGGGCCTGGAGACCAGCAGGGCGACTCCTACCTCAGGGTGTCCTCC**GACAGCCCAAAGGACCAG**AGCCCGCCTGAGG	**4.86***		
FLNA	Exon 12 Exon 13	CTGGAGGGCGGCG**TCGTTGGCAAGTCAGCAG**ACTTTGTGGTGGAGGCTATCGGGGACGACGTGGGCACGCTGGgtaagttggaggctgcagcatgggcacctggggacagacgatggcaaggacggcccaccctgaggctccagggcactg….	9.7	238	137
		….ggggactggtggctgttgtcagGCTTCTCGGTGGAAGGGCCATCGCAGGCTAAGATCGAATGTGACGACAAGGGCGACGG**CTCCTGTGATGTGCGCTAC**TGGCCGCAGGAGGCTGGCGAGTATGCCGTTCACGTGC	**4.2***		
PNKP	Exon 5 Exon 6	**GTGGCTGGCTTTGATCTGG**ACGGGACGCTCATCACCACACGCTCTGGGAAGGTCTTTCCCACTGGCCCCAGTGACTGGAGgtgataagaggcaaaacaagggagtgagtgagg….	**6.7***	210	123
		……gccccgggtcaccccTgccgcttcatacctgccgtagGATCTTGTACCCAGAGATTCCCC**GTAAGCTCCGAGAGCTGGAA**GCCGAGGGCTACAAG	**5.4***		
PCK2	Exon 3 Exon 4	CTGGCTGGCCCGCACAGACC**CCAAGGATGTGGCACGAGTAGAG**AGCAAGACGGTGATTGTAACTCCTTCTCAGCGGGACACGGTACAACTCCCGCCTGGTGGGGCCCGTGGGCAGCTGGGCAACTGGATGTCCCCAGCTGATTTCCAGCGAGCTGTGGATGAGAGGTTTCCAGGCTGCATGCAGGgtaaccagggcaggggcacagtggc…	**6.7***	358	271
		…..gcacggaagatgtgaacaggtttggaaccCttcatccaggggatgccttcctccacagGCCGCACCATGTATGTGCTTCCATTCAGCATGGGTCCTGTGGGCTCCCCGCTGTCCCGCATCGGGGTGCAGCTCACTGACTCAGCCTA**TGTGGTGGCAAGCATGCG**TATTATGACC	11.8		
IDH2	Exon 6 Exon 7	TCCATCTC**AGGTTTTGCGCACAGCTGC**TTCCAGTATGCCATCCAGAAGAAATGGCCGCTGTACATGAGCACCAAGAACACCATACTGAAAGCCTACGATGGGCGTTTCAAGGACATCTTCCAGGAGATCTTTGACAAgtaaagcctcatccatgtactctgtggcctttcttcccttcccccatgctgttcccatcctaccctgg……	**5.0***	445	270
		……gagtgcatttggctcagctccgaggctcagggagggatccccaacctgtcagccttctgccctctcccataacagacctttttactcccagGCACTATAAGACCGACTTCGACAAGAATAAGATCTGGTATGAGCACCGGCTCATTGATGACATGGTGGCTCAGGTCCTCAAGTCTTCGGGTGGCTTTGTGTGGGCCTGCAAGAACTAT**GACGGAGATGTGCAGTCAGACAT**CCT	12.9		
PFKL	Exon 8 Exon 9	ACTCGGAGCCGT**GGGTCCCGACTGAACATCATCA**TCATCGCTGAGGGTGCCATTGACCGCAACGGGAAGCCCATCTCGTCCAGCTACGTGAAGGACgtgcgtgtgggcctgggggtggccactgggcacctgctcctctaggccgtgt……	**5.3***	227	135
		……tggggctcagggctggtccttcccactgtcctgcagCTGGTGGTTCAGAGGCTGGGCTTCGACACCC**GTGTAACTGTGCTGGGCCAC**GTGCAGCGGGGAGGGACGCCCTCTGCCTTCGACCGGA	9.0		

ACY1	Exon 11 Exon 12	GCTGCAGTCAAACCCCCACCTGAAAGAGGGGTCCGTGAC**CT CCGTGAACCTGACTAA**GCTAGAGGGTGGCGTGGCCTATAAC GTGATACCTGCCACCATGAGCGCCAGCTTTGACTTCCGTGTG GCACCGGATGTGGACTTCAAGgtgccacctccacctgggt…….	**4.5***	316	175
		……agaggcctctggaaagcctgaaggatcagctcgtctcccttctcttagGCTTTTG AGGAGCAGCTGCAGAGCTGGTGCCAGGCAGCTGGCGAGGG GG**TCACCCTAGAGTTTGCTC**AG	12.4		
TUBG1	Exon 7 Exon 8	ACAACACAGCCCTGAACCGGATTGCCACAGACCGCCTGCAC ATCCAGAAC**CCATCCTTCTCCCAGATCAA**CCAGCTGgtgggccc ccactcctggactcctttggacttgaagccctccttgttggag……..	8.3	268	153
		……gggaggtcccacccaggctgaggcccataacatggcacgcctgtccccagGTG TCTACCATCATGTCAGCCAGCACCACCACCCTGCGCTACCCT GGCTACATGAACAATGACCTCATCGGCCTCATCGCCTCGCTC ATTCCCACCCCACGGCTCC**ACTTCCTCATGACCGGCTAC**ACC CCTCTC	**8.4 (Dual)**		
TUBA1A	Exon 2 Exon 3	ATTGGGGGAGGAGATGATTCCTTCAACACCTTCTTCAGTGAG ACGGGGGCTGGCAAGCATGTGCCCCG**GGCAGTGTTTGTAGA CTTGGA**ACCCACAGTCATTGgtgagttgacctcagtaacccaagtgagatcc cagggtgctgggacaggaggtctgtcctgggggctc……	9.2	257	105
		…….cgctcctcgtccctcctcctcccccctgctcctcccccatcatgttctccagATG AAGTTCGCACTGGCACCTACCGCCAGCTCTTCCACCCTGAGC AA**CTTATCACAGGCAAAGAAGATGC**TGCCAATAACTATGCCC GAGGGCACTACACCATTGGCAAGGAGATCATTG	7.9		
PSMD3	Exon8 Exon9	CTGTCAGGACAGGAAACCTAGCCAAGTTCAACCAGGTCCTGG ATCAGTTTGGGGAGAAGTTTCAAGCAGATGGGACCTACACCC TAATTATCCGGCTGCGGCACAACGTGATTAAGACAGgtgtggatc aggatctgaggggctgtt……	8.6	304	176
		…….gtccactctgcccaccccatcgctccttcctctcccagGTGTACGCATGATC AGCCTCTCCTATTCCCGAATCTCCTTGGCTGACATCGCCCAG AAGCTGCAGTTGGATAGCCCCGAAGATGCAGA	**13.9 (Dual)**		

The 5' splice site threshold score is 6.7 and the 3' splice site threshold score is 6.632. *, scores lower than threshold. Dual, splice sites that can be recognized as both 5' and 3' splice sites

Table S7. Sequence information of exon-skipped genes (5' splice site score / 3' splice site score) and the primer sequences (bold, underlined) used for RT-PCR to detect splicing defects

Target Gene (Gene symbol)	Primer location	Target sequence	PCR product size (bp)	
			All exon included	Exon Skipped
HDAC6	Exon 26 Exon 29	(exon 26) GGGGCCTCAGAATCTCAGGCCCCAGGAGAGGAGAACCTACTAGGAGAGGCAGCTGG**AGGTCAGGACATGGCTGATT**CGATGCTGATGCAGGGATCTAGGGGCCTCACTGATCAG (exon 27) GCCATATTTTATGCTGTGACACCACTGCCCTGGTGTCCCCATTTGGTGGCAGTATGCCCCATACCTGCAGCAGGCCTAGACGTGACCCAACCTTGTGGGGACTGTGGAACAATCCAAGAGAATTGGGTGTGTCTCTCTTGCTATCAG (exon 28) GTCTACTGTGGTCGTTACATCAATGGCCACATGCTCCAACACCATGGAAATTCTGGACACCCGCTGGTCCTCAGCTACATCGACCTGTCAGCCTGGTGTTACTACTGTCAGGCCTATGTCCACCACCAG (exon 29) GCTCTCCTAGATGTGAAGAACATCGCCCACCAGAACAAGTTTGGGGAGGATATGCCCCACCCA**CACTAAGCCCCAGAATACGG**TCCCTCTTCACCTTCTGAGGCCCACGATAGACCAGCTGTAGCTCATTCCAGCCTGTACCTTGGATGAGGGGTAGCCTCCCACTGCATCCCATCCTGAATA	417	288 270 141
ADA	Exon 8 Exon 10	(exon 8) GCTGTGGACATACTCAAGACAGAGCGGCTGGGACACG**GCTACCACACCCTGGAAGAC**CAGGCCCTTTATAACAGGCTGCGGCAGGAAAACATGCACTTCGAG (exon 9) ATCTGCCCCTGGTCCAGCTACCTCACTGGTGCCTGGAAGCCGGACACGGAGCATGCAGTCATTCG (exon 10) GCTCAAAAATGACCAGGCTAACTACTCGCTCAACACAGATGACCC**GCTCATCTTCAAGTCCACCC**TGGACACTGATTACCAGATGACCAAACGGGACATGGGCTTTACTGAAGAGGAGTTTAAAAGGCTG	195	140

Supplemental References

36. Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32, 1792-1797.

37. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189-1191.

38. Kranendijk, M., Struys, E.A., van Schaftingen, E., Gibson, K.M., Kanhai, W.A., van der Knaap, M.S., Amiel, J., Buist, N.R., Das, A.M., de Klerk, J.B., et al. (2010). IDH2 mutations in patients with D-2-hydroxyglutaric aciduria. Science 330, 336.