REPORT

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

Mari J. Tokita,[1,2] Alicia A. Braxton,[1,2] Yunru Shao,[1,3] Andrea M. Lewis,[1,3] Marie Vincent,[4] Sébastien Küry,[4] Thomas Besnard,[4] Bertrand Isidor,[4,5] Xénia Latypova,[4] Stéphane Bézieau,[4] Pengfei Liu,[1,2] Connie S. Motter,[6] Catherine Ward Melver,[6] Nathaniel H. Robin,[7] Elena M. Infante,[8] Marianne McGuire,[2,8] Areeg El-Gharbawy,[8] Rebecca O. Littlejohn,[9] Scott D. McLean,[9] Weimin Bi,[1,2] Carlos A. Bacino,[1,3] Seema R. Lalani,[1,3] Daryl A. Scott,[1,3,10] Christine M. Eng,[1,2,3] Yaping Yang,[1,2,12] Christian P. Schaaf,[1,3,11,12,*] and Magdalena A. Walkiewicz[1,2,12,*]

SON is a key component of the spliceosomal complex and a critical mediator of constitutive and alternative splicing. Additionally, SON has been shown to influence cell-cycle progression, genomic integrity, and maintenance of pluripotency in stem cell populations. The clear functional relevance of SON in coordinating essential cellular processes and its presence in diverse human tissues suggests that intact SON might be crucial for normal growth and development. However, the phenotypic effects of deleterious germline variants in *SON* have not been clearly defined. Herein, we describe seven unrelated individuals with de novo variants in *SON* and propose that deleterious variants in *SON* are associated with a severe multisystem disorder characterized by developmental delay, persistent feeding difficulties, and congenital malformations, including brain anomalies.

Whole-exome sequencing (WES) is an essential tool in the diagnostic evaluation of individuals with suspected genetic disorders for which a genetic etiology has not been established by conventional approaches. Studies of large cohorts of individuals have demonstrated a diagnostic yield of 25%–30% when WES is applied to otherwise perplexing cases.[1–3] The additional benefit of the unbiased sequencing approach of WES is the ability to ascertain genes in which variants have not been previously reported to cause disease. In our clinical WES cohort of over 6,000 unrelated individuals—the majority of whom have neurologic manifestations and are of pediatric age—we identified six individuals (subjects 1–6) with truncating variants in *SON* (SON DNA binding protein [MIM: 182465]) and overlapping clinical features. We analyzed parental samples by Sanger sequencing or WES and confirmed de novo status of all six variants. Subsequently, we ascertained one additional individual (subject 7) with two de novo missense variants in *SON* and similar features. Herein, we comprehensively phenotype all seven individuals and propose that deleterious variants in *SON* are associated with severe developmental outcomes.

This study was performed in accordance with a protocol that was prospectively reviewed and approved by the Baylor College of Medicine Institutional Review Board. Written informed consent was obtained from all study participants. The key clinical features of our cohort are summarized in Table 1. Detailed clinical summaries for all subjects are provided in the Supplemental Data, and photographs are included in Figure 1. All subjects had dysmorphic features including, for example, mild midface retrusion with apparently deep-set eyes (n = 6), frontal bossing and bitemporal narrowing (n = 2), downslanting palpebral fissures (n = 5), and epicanthal folds (n = 3). All subjects had either a smooth or short philtrum (n = 7), and a subset had thin lips (n = 5) and/or a short mouth (n = 3). All subjects exhibited developmental delay, which appeared to progress with age into moderate to severe intellectual disability. All but one individual had additional neurological features including regression (n = 3), epilepsy or other electroencephalography (EEG) abnormalities (n = 4), autism spectrum disorder (n = 3), and hyper- or hypotonia (n = 5). Additionally, five of six subjects evaluated had abnormalities detected on brain imaging; features suggestive of volume loss specifically were seen in all five. Five subjects had congenital malformations. An atrial septal defect, ventricular septal defect, patent ductus arteriosus, left lung agenesis, single kidney, dysplastic kidney, and agenesis of the gallbladder were each seen in a single individual; several subjects had more than one malformation. All subjects had a history of feeding difficulties, which were evident as early as the

[1]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA; [2]Baylor Miraca Genetics Laboratories, Houston, TX 77021, USA; [3]Texas Children's Hospital, Houston, TX 77030, USA; [4]Service de Génétique Médicale, Centre Hospitalier Universitaire de Nantes, 44093 Nantes Cedex 1, France; [5]INSERM UMR-S 957, 1 Rue Gaston Veil, 44035 Nantes, France; [6]Genetic Center, Akron Children's Hospital, Akron, OH 44302, USA; [7]Department of Genetics, University of Alabama at Birmingham, Birmingham, AL 35294, USA; [8]University of Pittsburgh School of Medicine and Children's Hospital of Pittsburgh, Pittsburgh, PA 15224, USA; [9]Children's Hospital of San Antonio, San Antonio, TX 78207, USA; [10]Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, TX 77030, USA; [11]Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX 77030, USA
[12]These authors contributed equally to this work
*Correspondence: schaaf@bcm.edu (C.P.S.), mwalkiew@bcm.edu (M.A.W.)
http://dx.doi.org/10.1016/j.ajhg.2016.06.035.

Table 1. Clinical Features of Subjects with De Novo *SON* Variants

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7
Current age	6 years	23 years	9 years	3 years	15 years	9 years	3 years
Sex	female	male	female	female	female	female	female
Pregnancy	IUGR, placenta previa	maternal hypertension	IUGR	IUGR, maternal borderline diabetes, factor V deficiency	maternal hypertension	IUGR, oligohydramnios, pre-eclampsia, fetal anomalies	IUGR, fetal anomalies
Age at birth	32 weeks	full term	full term	33 weeks	35 weeks	36 weeks	36 weeks
Delivery	C-section for fetal distress	C-section for fetal distress	wrapped cord, variable heart rate, failure to progress	C-section for fetal distress	C-section for maternal hypertension	vaginal delivery	C-section for fetal distress
Postnatal course	respiratory failure, feeding difficulties	feeding difficulties	feeding difficulties, hypoglycemia	respiratory failure, feeding difficulties	feeding difficulties, respiratory issues	feeding difficulties, respiratory issues	respiratory distress, feeding difficulties
Height	2nd percentile	40th percentile	25th percentile	75th percentile	3rd percentile	−3 (*Z* score)	1st percentile
Weight	3rd percentile	1st percentile	−2.29 (*Z* score)	85th percentile	12th percentile	2nd percentile	−3 (*Z* score)
OFC	2nd percentile	50th percentile	−4 (*Z* score)	60th percentile	72nd percentile	12th percentile	−2.5 (*Z* score)
Distinctive features	frontal bossing, bitemporal narrowing, epicanthal folds, thin lip, smooth philtrum	downslanting palpebral fissures, bifid uvula, submucous cleft palate, short philtrum	downslanting palpebral fissures, downturned mouth, short philtrum, thin lip, thin limbs	submucous and laryngeal cleft, frontal bossing, bitemporal narrowing, epicanthal folds, thin lip, smooth philtrum	downslanting palpebral fissures, laterally flared eyebrows, short philtrum	downslanting palpebral fissures, long face, full cheeks, short philtrum, thin lips	downslanting palpebral fissures, epicanthal folds, smooth philtrum, thin lips
Developmental delay	yes	yes	yes	yes	yes	yes	yes
Regression	yes	yes	no	yes	no	no	no
ASD	yes	yes	yes	NA	NA	no	no
Seizures	yes	yes	no (abnormal EEG)	staring spells	NA	no (abnormal EEG)	no
Tone	hypotonia	NA	hypotonia and spasticity	hypotonia	normal	hypotonia	hypotonia
Brain imaging	global volume loss, thin corpus callosum, mild periventricular gliosis	progressive ventricular and subarachnoid space dilatation, arachnoid cyst	unremarkable	periventricular leukomalacia with mild dilation of the lateral ventricle	prominent extra-axial spaces, dysgenesis of corpus callosum	evidence of prior MCA stroke, prominent ventricles	not done
Congenital malformation	atrial septal defect (resolved)	NA	NA	abnormal placement of carotid arteries in the neck	single kidney	dysplastic kidney, congenital lobar emphysema	VSD, PDA, agenesis of the left lung, gallbladder agenesis
Vision	exotropia, nystagmus	progressive vision loss, myopia, exotropia	NA	esotropia, CVI, blue sclera, segmental optic nerve hypoplasia	history of bilateral eye surgery	strabismus	no concerns

(*Continued on next page*)

Table 1. *Continued*

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7
Hearing	PE tubes	auditory hallucination	PE tubes	PE tubes	NA	inconclusive hearing assessment	no concerns
Gastro-intestinal features	delayed gastric emptying, feeding difficulties	pancreatic lipase insufficiency, dysphagia	failure to thrive, chronic diarrhea, feeding difficulties	failure to thrive, G-tube feeding, diarrhea, reflux, gastric dysmotility	feeding difficulties	dysphagia, G-tube feeding	failure to thrive, G-tube recommended
Musculo-skeletal features	joint laxity	scoliosis, arachnodactyly, dolichostenomelia	joint laxity, cervical rib	joint laxity, cervical ribs, mild syndactyly	exaggerated lumbar lordosis	none	hemivertebrae, rib fusion, thumb agenesis, syndactyly
Hematologic features	DVT	NA	IgG and IgA deficiency, recurrent infection	IgA deficiency, recurrent infection	borderline IgG levels	prior MCA infarct, multiple TIAs	NA

Abbreviations are as follows: ASD, autism spectrum disorder; C-section, cesarean section; CVI, cortical visual impairment; DVT, deep-vein thrombosis; EEG, electroencephalography; G-tube, gastrostomy tube; IgA, immunoglobulin A; IgG, immunoglobulin G; IUGR, intrauterine growth restriction; MCA, middle cerebral artery; NA, not ascertained; OFC, occipitofrontal circumference; PDA, patent ductus arteriosus; PE tubes, pressure-equalizing tubes; TIA, transient ischemic attack; and VSD, ventricular septal defect.

neonatal period and associated with growth failure in most cases. Several subjects required a gastrostomy feeding tube. Most subjects also had ophthalmologic concerns including strabismus (n = 4) and vision loss (n = 2). Six subjects had skeletal abnormalities including joint laxity (n = 3), cervical ribs (n = 2), scoliosis (n = 1), and thumb agenesis (n = 1). Pregnancy and delivery complications were common in the cohort. Five of the seven subjects had intrauterine growth restriction, at least four had significant fetal distress requiring delivery via cesarean section, and five were born prematurely. Three subjects had a history of borderline low or frankly deficient immunoglobulin levels, and two subjects had episodes of suspected abnormal clotting, including unprovoked deep-vein thrombosis in subject 1 and a history of a right middle cerebral artery infarct and multiple transient ischemic attacks in subject 6.

Sequencing results are summarized in Table 2. Variant nomenclature is consistent with *SON* transcript GenBank: NM_138927.2 (UCSC Genome Browser hg19). All truncating and missense variants were confirmed by Sanger sequencing and found to be de novo by parental testing. Subjects 1–6 had truncating variants including one premature stop variant in exon 3 (c.286C>T [p.Gln96*]), three frameshift variants in exon 3 (c.3073dupA [p.Met1025Asnfs*6], c.3852_3856delGGTAT [p.Met1284Ilefs*2], and c.5753_5756delTTAG [p.Val1918Glufs*87]), and one frameshift variant in exon 4 (c.6233delC [p.Pro2078Hisfs*4]) (Figure 2). Of note, the c.5753_5756delTTAG (p.Val1918Glufs*87) variant was observed in two unrelated subjects. Subject 7 had two de novo missense changes in *cis* configuration in exon 3 (Figure S1). SIFT and PolyPhen-2 predicted the c.4909A>T (p.Thr1637Ser) missense variant to be deleterious and benign, respectively, and the c.5528C>A (p.Ser1843Tyr) variant to be deleterious and damaging, respectively. It is unclear whether this is a complex allele or whether an individual variant contributes to the disease phenotype.

SON is located in human chromosomal region 21q22.11 and consists of 12 exons.[5] A striking feature of the gene's structure is the size of exon 3, which accounts for 82% of the entire coding region (GenBank: NM_138927.2). According to the Exome Aggregation Consortium (ExAC) Browser, *SON* is predicted to be intolerant to loss-of-function mutations given that 49.1 loss-of-function variants are expected but only one loss-of-function variant is observed (pLI = 1.00).[6] *SON* does not appear to be intolerant to missense variation,[6] however, suggesting that cautious interpretation of the missense changes detected in subject 7 is warranted.

The canonical SON isoform (GenBank: NP_620305.2, isoform F) encoded by GenBank: NM_138927.2 is a 2,426 amino acid protein that is ubiquitously present in human tissues and highly conserved[7,8] and has an estimated 84% sequence homology between human *SON* and mouse *Son*.[9] SON contains several recognizable domains implicating it as a modulator of RNA processing; these include an

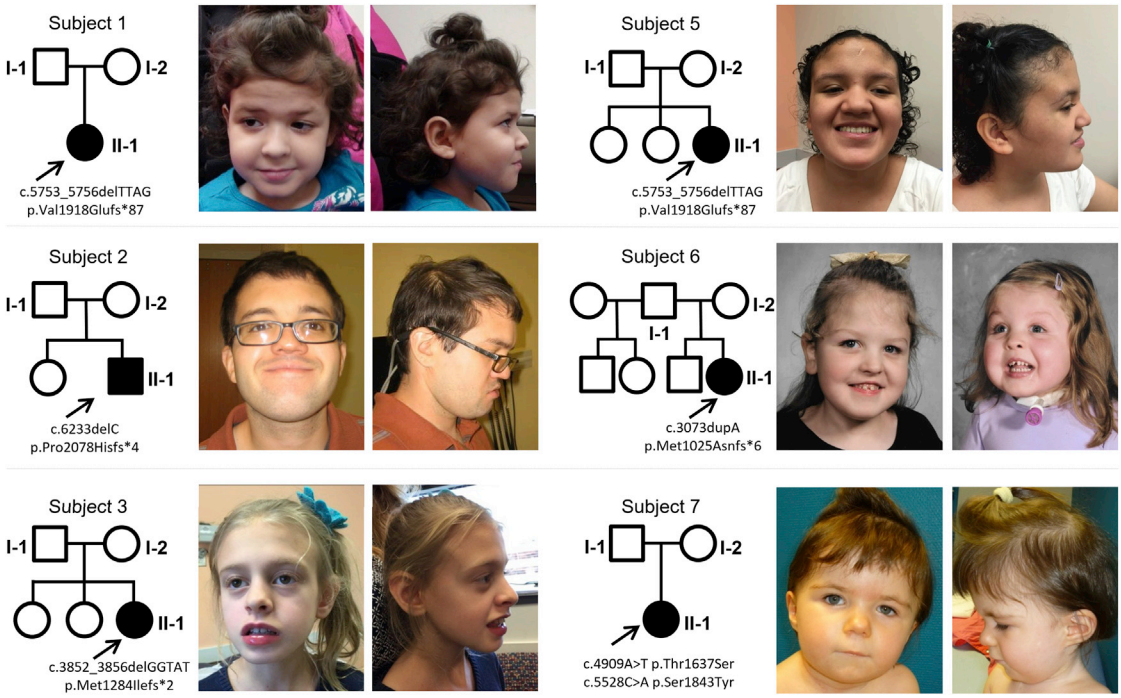

Figure 1. Photographs and Pedigrees of Subjects with *SON* Variants
Photographs show subjects reported in this article, and pedigrees illustrate the de novo status of all detected *SON* variants. Shaded symbols represent affected individuals.

arginine/serine (RS)-rich domain, a G-patch domain, and a double-stranded RNA-binding motif (Figure 2).[4,9,10] The RS domain is involved in protein-protein interactions and RNA processing.[11–14] Interestingly, the c.5528C>A substitution affects the serine at amino acid 1,843 within the RS region (Figure 2), thus altering the composition of a crucial functional domain of SON.

Previous analyses of murine and human cells have shown nuclear staining of SON in a stippled pattern consistent with localization to the nuclear speckle.[9,10,15,16] The nuclear speckle is a subcellular intranuclear compartment that is enriched with pre-mRNA splicing factors, including small nuclear ribonucleoprotein particles[17] and SR protein family members known to be involved in RNA splicing.[4] SON's functional domains and its localization in the nuclear speckle suggest that it plays a role in pre-mRNA processing. Functional studies have confirmed that SON is an important mediator of both constitutive and alternative splicing[4,18] and that it is specifically involved in splicing short introns with suboptimal or weak splice sites.[8,19] Known targets of SON-mediated splicing include cell-cycle and microtubule genes, as well as genes involved in DNA repair.[4,19] Indeed, depletion of SON by RNAi leads to an array of adverse cellular consequences, including mitotic arrest, disordered spindle architecture with abnormal chromosomal alignment, aneuploidy in cells that continue to divide,[16,19] and loss of genomic integrity, as evidenced by increased double-stranded DNA breaks and micronuclei formation in cells lacking functional SON.[19] In addition, the regulatory effect of SON on splicing has been shown to be essential for the maintenance of pluripotency and self-renewal in human embryonic stem cells.[8]

In spite of extensive work showing a critical functional role for SON in coordinating splicing and evidence that aberrant splicing contributes to human disease,[20,21] mutations in *SON* have not yet been definitively linked to a phenotype in humans. The first de novo truncating variant in *SON* was identified in a large cohort of individuals with severe intellectual disability.[22] Zhu et al. later described another individual with a de novo truncating variant in *SON*.[23] This individual had developmental delay, epilepsy, minor dysmorphic features, macrocephaly, brain white-matter abnormalities, intestinal atresia, and a ventricular septal defect. However, this individual also had a de novo missense change in a second candidate gene, *C5AR1* (MIM: 113995), confounding the clinical relevance of the *SON* change. This published individual and the seven subjects in our cohort exhibit many of the same features, suggesting that deleterious variants in *SON* cause a consistent phenotype. Moreover, two of our subjects (1 and 5) share the same frameshift variant as the individual described by Zhu et al., indicating that this is a recurrent pathogenic change.

Variants in genes encoding other components of the spliceosomal machinery have been implicated in several developmental disorders, including Guion-Almeida type mandibulofacial dysostosis (MFDGA [MIM: 610536]) and Nager syndrome (MIM: 154400), among others (recently

Table 2. Putative Pathogenic Variants in *SON*

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7[a]	Subject 7[a]
DNA variant	c.5753_5756delTTAG	c.6233delC	c.3852_3856delGGTAT	c.286C>T	c.5753_5756delTTAG	c.3073dupA	c.4909A>T	c.5528C>A
Protein change	p.Val1918Glufs*87	p.Pro2078Hisfs*4	p.Met1284Ilefs*2	p.Gln96*	p.Val1918Glufs*87	p.Met1025Asnfs*6	p.Thr1637Ser	p.Ser1843Tyr
Inheritance	de novo	de novo	de novo	de novo	de novo	de novo	de novo	de novo
ExAC Browser	novel	novel	novel	novel	novel	novel	novel	novel
SIFT	–	–	–	–	–	–	deleterious	deleterious
PolyPhen-2	–	–	–	–	–	–	benign	damaging
CADD	–	–	–	–	–	–	15.19	15.67

Variant nomenclature is based on GenBank: NM_138927.2.
[a]These variants are in *cis* configuration, and both are confirmed de novo changes.

reviewed by Lehalle et al.[24]). MFDGA is caused by mutations in *EFTUD2* (MIM: 603892), which encodes a highly conserved spliceosomal GTPase.[24] The phenotype associated with MFDGA mirrors both in breadth and severity the features common to our cohort, including psychomotor delay, growth retardation, musculoskeletal anomalies, and cardiac, brain, and visceral malformations.[24] Nager syndrome, which is caused by mutations in *SF3B4* (MIM: 605593), is characterized by midface retrusion, downslanting palpebral fissures, and thumb anomalies.[25] All of these features were present in one or more of our subjects with *SON* variants. This phenotypic overlap with established spliceosomal disorders confers plausibility to the hypothesis that defects in *SON* cause the features seen in our cohort.

Orthogonal evidence of the potential clinical relevance of *SON* haploinsufficiency derives from reports of individuals with copy-number variants (CNVs) involving this gene. Non-recurrent microdeletions encompassing 21q22.11, the locus harboring *SON*, have been extensively described in the literature.[26–36] Roberson et al. performed genotype-phenotype correlations for 46 individuals with partial 21q monosomy and, consistent with other reports,[27,36] found that individuals with deletions encompassing the 21q22.11 locus manifest severe phenotypes.[35] Lindstrand et al. compared 26 individuals who had partial 21q monosomy and for whom reliable molecular data were available.[34] Alignment of the deleted regions and comparison of phenotypes showed a narrow 159 kb region of overlap in 21q22.11 among individuals with

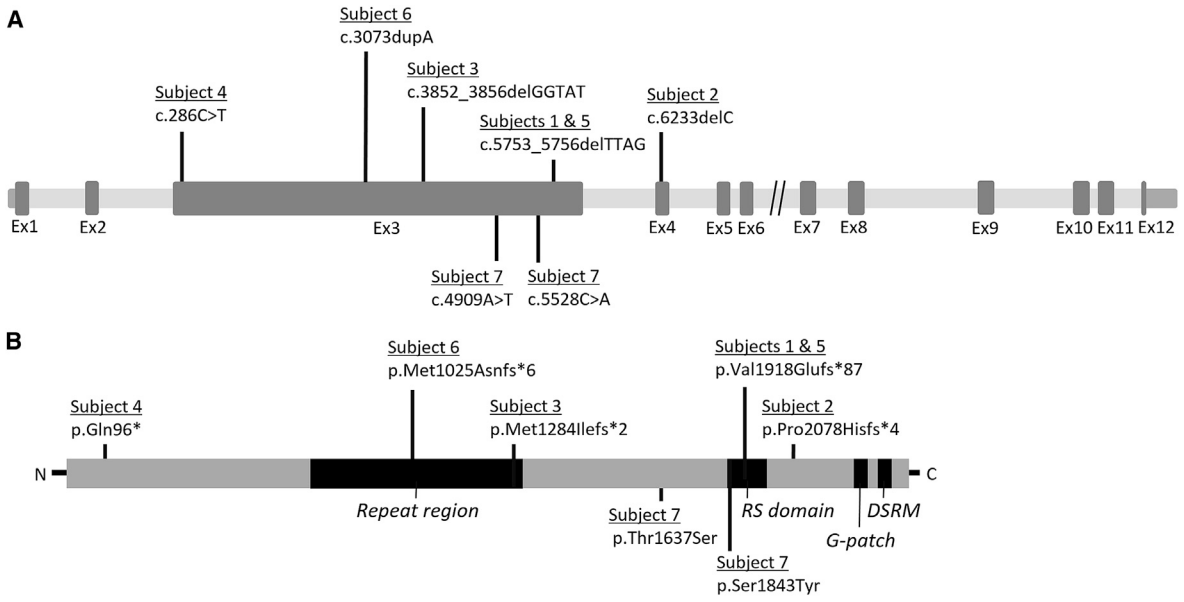

Figure 2. Intragenic Location of *SON* Variants and Key Protein Functional Domains
(A) All but one of the *SON* variants in the described individuals localize to exon 3 of *SON* (GenBank: NM_138927.2).
(B) Approximate location of amino acid changes in relation to SON's key functional domains, which include a unique central highly repetitive region, an RS-rich domain, a G-patch domain, and a double-stranded RNA-binding motif (DSRM). Data were extracted from GenBank: NP_620305.2. This panel was adapted from Hickey et al.[4]

Table 3. *SON*-Associated Clinical Features in Reported Subjects with Deletions Encompassing *SON*

	Subjects 1–7	Izumi et al.[26]	Fukai et al.[27]	Beri-Dexheimer et al.[28] (Patient 2)	Hoyer et al.[29]	Shinawi et al.[30] (Patient 2)	Thevenon et al.[31]	Carrascosa-Romero et al.[32]	Katzaki et al.[33] (Patient 2)	dbVar: nssv577822	Subject 8
Deletion (Mb)	NA	1.9	1.4	3.3	3	1.81	2.97	2.84	2.9	0.341	0.825
Developmental delay	+ (7/7)	NR	+	+	+	+	+	+	+	+	+
IUGR and/or low birth weight	+ (5/6)	+	–	+	NR	+	+	+	+	NR	NR
Short stature	+ (5/7)	+	+	+	+	+	+	+	+	+	+
Respiratory problems	+ (5/7)	+	NR	NR	NR	NR	NR	+	NR	NR	NR
Feeding problems	+ (7/7)	+ (G-tube)	NR	NR	NR	+ (G-tube)	+	+ (G-tube)	NR	NR	–
Seizures	+ (2/6)	NR	+	–	–	–	+	–	+	+	+
Abnormal tone	+ (5/6)	+	+	+	NR	NR	+	+	NR	NR	–
Brain anomalies	+ (5/6)	+	+	+	NR	+	+	+	+	NR	–
Congenital malformations	+ (5/5)	heart	heart	heart	–	NR	heart	renal	–	NR	heart

Abbreviations are as follows: G-tube, gastrostomy tube; NA, not applicable; and NR, not reported.

intellectual disability.[34] This region contains only five genes and encompasses the entirety of all isoforms of *SON*, suggesting that loss of *SON* might contribute specifically to the intellectual disability in these individuals.[34]

To further explore the potential implications of *SON* copy-number loss, we queried our internal clinical database of chromosomal microarrays (n = ~70,000 affected individuals) and identified an individual (subject 8; Table 3) with a ~825 kb deletion encompassing *SON* and ten additional RefSeq genes. This individual was reported to have global developmental delay, seizures, and a congenital heart defect—features also seen in the described subjects with *SON* sequence variants. We then selectively reviewed published reports of phenotypically characterized individuals with <5 Mb 21q22.11 deletions that partially or completely involve *SON* (Figure S2)[26–33] and found substantial phenotypic overlap between individuals with deletions encompassing *SON* and the seven subjects with *SON* variants reported herein (Table 3). The individuals with microdeletions included both male and female probands, which is notable given the clear predominance of female subjects in our cohort. Seven of eight individuals with deletions of *SON* had developmental delay; all eight individuals had growth failure with short stature, seven of eight had brain anomalies, and six of eight had a history of intrauterine growth restriction and/or low birth weight. Four individuals were reported to have feeding difficulties, which required G-tube placement in three. Table 3 also includes a single individual with a small 341 kb de novo deletion reported in ClinVar (ClinVar: SCV000080160.5; dbVar: nssv577822). This individual is reported to have global developmental delay, seizures, and short stature—all features seen in our subjects with *SON* variants. Thus, although we cannot exclude the possibility that other genes in this region (e.g., *GART* [MIM: 138440], *DONSON* [MIM: 611428], *CRYZL1* [MIM: 603920], and *ITSN1* [MIM: 602442]) contribute to the phenotype in individuals with large deletions, the existing CNV data on this well-studied region strengthens the supposition that *SON* haploinsufficiency is in fact pathogenic.

In summary, we have characterized a clinical phenotype associated with pathogenic variants involving *SON*. The similarity in phenotype between subjects with truncating variants and those with CNVs suggests that haploinsufficiency of *SON* could be the underlying disease mechanism. Although additional studies will be necessary to confirm the functional relevance of heterozygous loss of *SON* and to capture the full phenotypic spectrum, the available human data compellingly support the assertion that deleterious variants in *SON* are associated with a severe human phenotype.

Accession Numbers

The accession numbers for the truncating variants reported in this paper are ClinVar: SCV000297718, SCV000297719, SCV000297720, SCV000297721, and SCV000297722.

Supplemental Data

Supplemental Data include a detailed clinical history of each reported subject and two figures and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2016.06.035.

Conflicts of Interest

The Department of Molecular and Human Genetics at the Baylor College of Medicine derives revenue from molecular genetic testing offered at the Baylor Miraca Genetics Laboratories.

Acknowledgments

We greatly appreciate the study participants and their families, without whom this work would not have been possible. We also express our gratitude to Mahshid Azamian for excellent assistance with sample collection.

Received: April 27, 2016
Accepted: June 24, 2016
Published: August 18, 2016

Web Resources

ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/
ExAC Browser, http://exac.broadinstitute.org/
OMIM, http://www.omim.org/
RefSeq, http://www.ncbi.nlm.nih.gov/RefSeq
UCSC Genome Browser, https://genome.ucsc.edu/

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 patients referred for clinical whole-exome sequencing. JAMA *312*, 1870–1879.
2. Lee, H., Deignan, J.L., Dorrani, N., Strom, S.P., Kantarci, S., Quintero-Rivera, F., Das, K., Toy, T., Harry, B., Yourshaw, M., et al. (2014). Clinical exome sequencing for genetic identification of rare Mendelian disorders. JAMA *312*, 1880–1887.
3. Sawyer, S.L., Hartley, T., Dyment, D.A., Beaulieu, C.L., Schwartzentruber, J., Smith, A., Bedford, H.M., Bernard, G., Bernier, F.P., Brais, B., et al.; FORGE Canada Consortium; Care4Rare Canada Consortium (2016). Utility of whole-exome sequencing for those near the end of the diagnostic odyssey: time to address gaps in care. Clin. Genet. *89*, 275–284.
4. 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.
5. Khan, I.M., Fisher, R.A., Johnson, K.J., Bailey, M.E., Siciliano, M.J., Kessling, A.M., Farrer, M., Carritt, B., Kamalati, T., and Buluwela, L. (1994). The SON gene encodes a conserved DNA binding protein mapping to human chromosome 21. Ann. Hum. Genet. *58*, 25–34.
6. Lek, M., Karczewski, K., Minikel, E., Samocha, K., Banks, E., Fennell, T., O'Donnell-Luria, A., Ware, J., Hill, A., Cummings, B., et al. (2015). Analysis of protein-coding genetic variation in 60,706 humans. bioRxiv http://dx.doi.org/10.1101/030338.
7. Mattioni, T., Hume, C.R., Konigorski, S., Hayes, P., Osterweil, Z., and Lee, J.S. (1992). A cDNA clone for a novel nuclear protein with DNA binding activity. Chromosoma *101*, 618–624.
8. 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.
9. Wynn, S.L., Fisher, R.A., Pagel, C., Price, M., Liu, Q.Y., Khan, I.M., Zammit, P., Dadrah, K., Mazrani, W., Kessling, A., et al. (2000). Organization and conservation of the GART/SON/DONSON locus in mouse and human genomes. Genomics *68*, 57–62.
10. Saitoh, N., Spahr, C.S., Patterson, S.D., Bubulya, P., Neuwald, A.F., and Spector, D.L. (2004). Proteomic analysis of interchromatin granule clusters. Mol. Biol. Cell *15*, 3876–3890.
11. Birney, E., Kumar, S., and Krainer, A.R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. Nucleic Acids Res. *21*, 5803–5816.
12. Graveley, B.R. (2000). Sorting out the complexity of SR protein functions. RNA *6*, 1197–1211.
13. Sacco-Bubulya, P., and Spector, D.L. (2002). Disassembly of interchromatin granule clusters alters the coordination of transcription and pre-mRNA splicing. J. Cell Biol. *156*, 425–436.
14. Zhong, X.Y., Wang, P., Han, J., Rosenfeld, M.G., and Fu, X.D. (2009). SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. Mol. Cell *35*, 1–10.
15. 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.
16. Huen, M.S., Sy, S.M., Leung, K.M., Ching, Y.P., Tipoe, G.L., Man, C., Dong, S., and Chen, J. (2010). SON is a spliceosome-associated factor required for mitotic progression. Cell Cycle *9*, 2679–2685.
17. Spector, D.L., and Lamond, A.I. (2011). Nuclear speckles. Cold Spring Harb. Perspect. Biol. *3* http://dx.doi.org/10.1101/cshperspect.a000646.
18. Lu, X., Ng, H.H., and Bubulya, P.A. (2014). The role of SON in splicing, development, and disease. Wiley Interdiscip. Rev. RNA *5*, 637–646.
19. 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.
20. Wang, G.S., and Cooper, T.A. (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. Nat. Rev. Genet. *8*, 749–761.
21. Padgett, R.A. (2012). New connections between splicing and human disease. Trends Genet. *28*, 147–154.
22. 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.
23. 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.
24. Lehalle, D., Wieczorek, D., Zechi-Ceide, R.M., Passos-Bueno, M.R., Lyonnet, S., Amiel, J., and Gordon, C.T. (2015). A review of craniofacial disorders caused by spliceosomal defects. Clin. Genet. *88*, 405–415.

25. Bernier, F.P., Caluseriu, O., Ng, S., Schwartzentruber, J., Buckingham, K.J., Innes, A.M., Jabs, E.W., Innis, J.W., Schuette, J.L., Gorski, J.L., et al.; FORGE Canada Consortium (2012). Haploinsufficiency of SF3B4, a component of the pre-mRNA spliceosomal complex, causes Nager syndrome. Am. J. Hum. Genet. *90*, 925–933.
26. Izumi, K., Brooks, S.S., Feret, H.A., and Zackai, E.H. (2012). 1.9 Mb microdeletion of 21q22.11 within Braddock-Carey contiguous gene deletion syndrome region: dissecting the phenotype. Am. J. Med. Genet. A. *158A*, 1535–1541.
27. Fukai, R., Hiraki, Y., Nishimura, G., Nakashima, M., Tsurusaki, Y., Saitsu, H., Matsumoto, N., and Miyake, N. (2014). A de novo 1.4-Mb deletion at 21q22.11 in a boy with developmental delay. Am. J. Med. Genet. A. *164A*, 1021–1028.
28. Béri-Dexheimer, M., Latger-Cannard, V., Philippe, C., Bonnet, C., Chambon, P., Roth, V., Grégoire, M.J., Bordigoni, P., Lecompte, T., Leheup, B., and Jonveaux, P. (2008). Clinical phenotype of germline RUNX1 haploinsufficiency: from point mutations to large genomic deletions. Eur. J. Hum. Genet. *16*, 1014–1018.
29. Hoyer, J., Dreweke, A., Becker, C., Göhring, I., Thiel, C.T., Peippo, M.M., Rauch, R., Hofbeck, M., Trautmann, U., Zweier, C., et al. (2007). Molecular karyotyping in patients with mental retardation using 100K single-nucleotide polymorphism arrays. J. Med. Genet. *44*, 629–636.
30. Shinawi, M., Erez, A., Shardy, D.L., Lee, B., Naeem, R., Weissenberger, G., Chinault, A.C., Cheung, S.W., and Plon, S.E. (2008). Syndromic thrombocytopenia and predisposition to acute myelogenous leukemia caused by constitutional microdeletions on chromosome 21q. Blood *112*, 1042–1047.
31. Thevenon, J., Callier, P., Thauvin-Robinet, C., Mejean, N., Falcon-Eicher, S., Maynadie, M., de Maistre, E., Bidot, S., Huet, F., Beri-Dexheimer, M., et al. (2011). De Novo 21q22.1q22.2 deletion including RUNX1 mimicking a congenital infection. Am. J. Med. Genet. A. *155A*, 126–129.
32. Carrascosa-Romero, M.C., Suela, J., Pardal-Fernández, J.M., Bermejo-Sánchez, E., Vidal-Company, A., MacDonald, A., Tébar-Gil, R., Martínez-Fernández, M.L., and Martínez-Frías, M.L. (2013). A 2.84 Mb deletion at 21q22.11 in a patient clinically diagnosed with Marden-Walker syndrome. Am. J. Med. Genet. A. *161A*, 2281–2290.
33. Katzaki, E., Morin, G., Pollazzon, M., Papa, F.T., Buoni, S., Hayek, J., Andrieux, J., Lecerf, L., Popovici, C., Receveur, A., et al. (2010). Syndromic mental retardation with thrombocytopenia due to 21q22.11q22.12 deletion: Report of three patients. Am. J. Med. Genet. A. *152A*, 1711–1717.
34. Lindstrand, A., Malmgren, H., Sahlén, S., Schoumans, J., Nordgren, A., Ergander, U., Holm, E., Anderlid, B.M., and Blennow, E. (2010). Detailed molecular and clinical characterization of three patients with 21q deletions. Clin. Genet. *77*, 145–154.
35. Roberson, E.D., Wohler, E.S., Hoover-Fong, J.E., Lisi, E., Stevens, E.L., Thomas, G.H., Leonard, J., Hamosh, A., and Pevsner, J. (2011). Genomic analysis of partial 21q monosomies with variable phenotypes. Eur. J. Hum. Genet. *19*, 235–238.
36. Lyle, R., Béna, F., Gagos, S., Gehrig, C., Lopez, G., Schinzel, A., Lespinasse, J., Bottani, A., Dahoun, S., Taine, L., et al. (2009). Genotype-phenotype correlations in Down syndrome identified by array CGH in 30 cases of partial trisomy and partial monosomy chromosome 21. Eur. J. Hum. Genet. *17*, 454–466.

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 40th percentile for height, the 1st percentile for weight, and 50th 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 25th percentile for height, less than the 5th percentile for weight (z-score: -2.3), and less than the 3rd 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 75th percentile for height, 85th percentile for weight, and 60th 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 3rd centile for height, 12th centile for weight, and 72nd 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

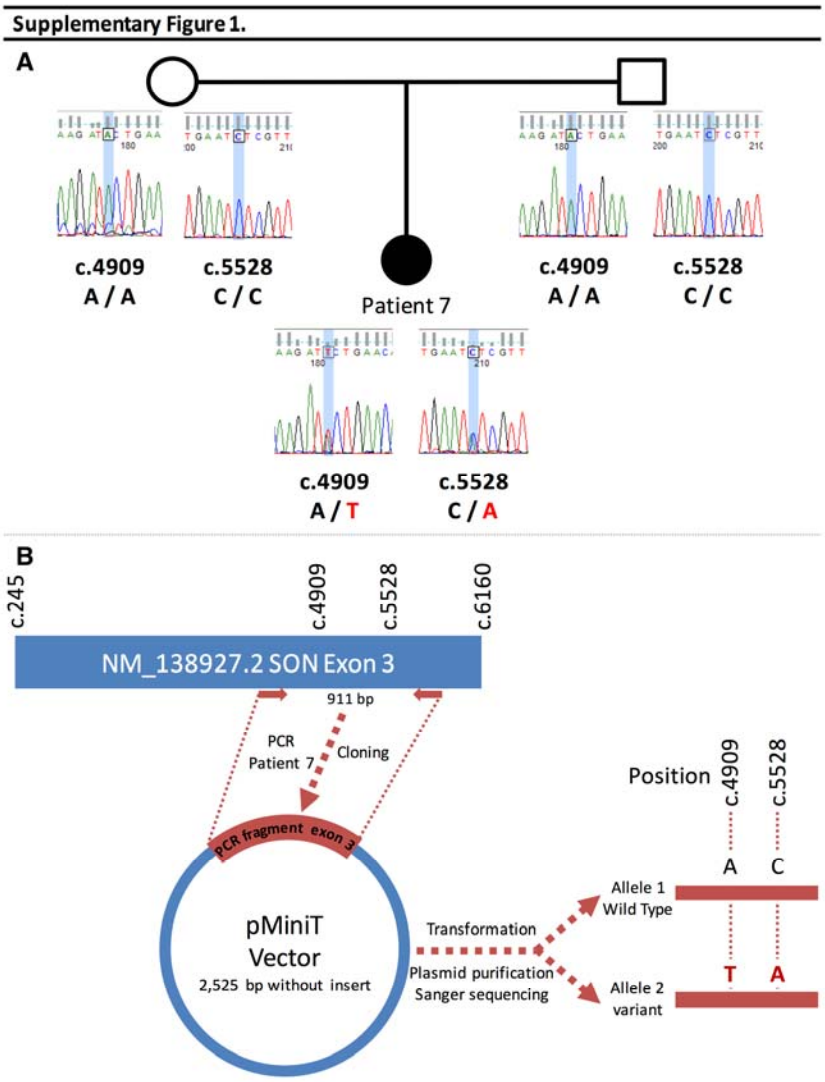

Figure S1. Haplotype phasing of the *de novo* missense variants detected in Subject 7.

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.

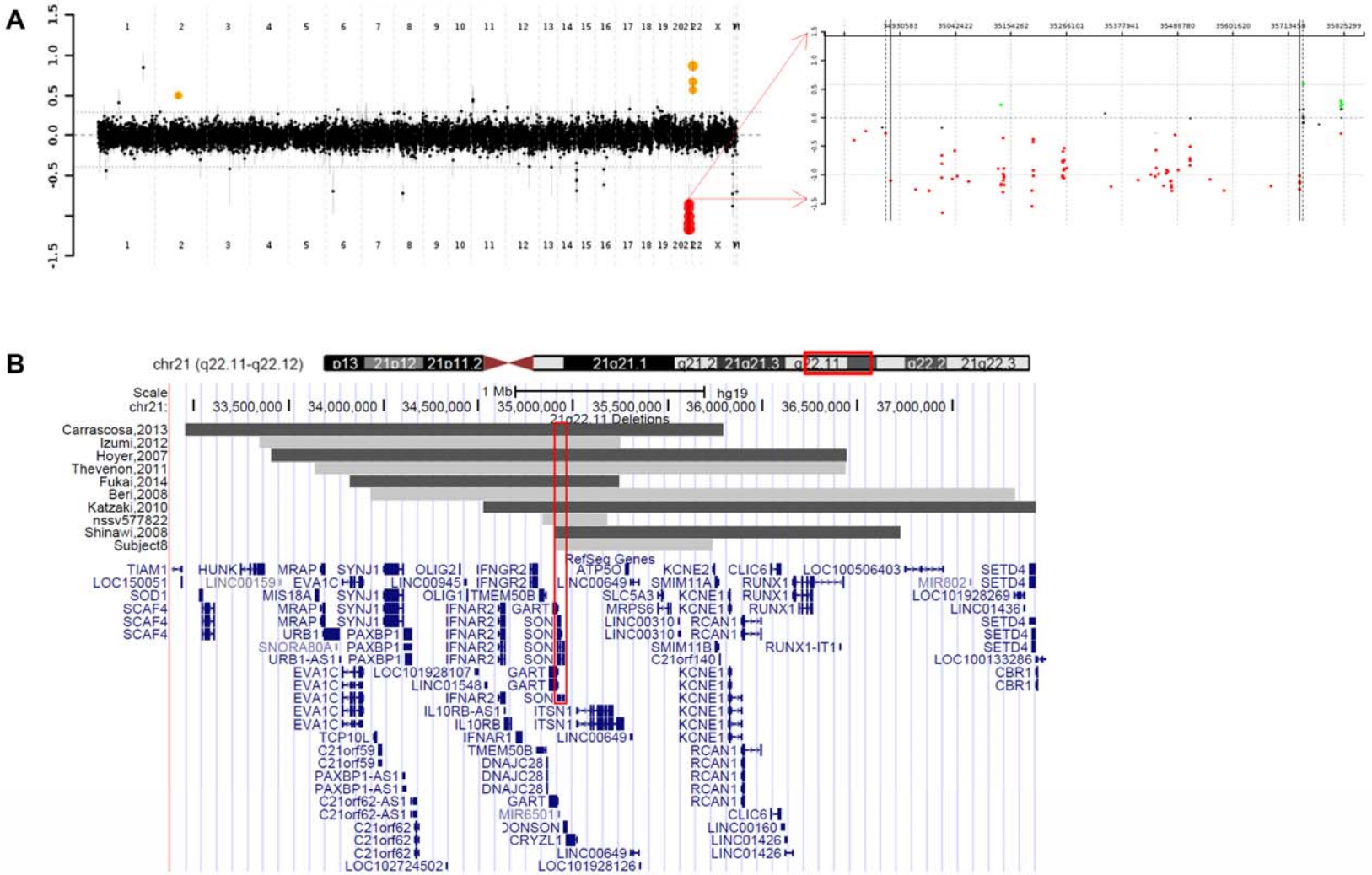
A
B
chr21 (q22.11-q22.12)
21q22.11 Deletions
RefSeq Genes
Carrascosa,2013
Izumi,2012
Hoyer,2007
Thevenon,2011
Fukai,2014
Beri,2008
Katzaki,2010
nssv577822
Shinawi,2008
Subject8

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.