



HAL
open science

SOS1 MUTATIONS IN NOONAN SYNDROME: AN UPDATE ON MOLECULAR SPECTRUM, STRUCTURAL INSIGHTS ON PATHOGENIC EFFECTS, AND GENOTYPE-PHENOTYPE CORRELATIONS

Francesca Lepri, Alessandro de Luca, Lorenzo Stella, Cesare Rossi, Giuseppina Baldassarre, Francesca Pantaleoni, Viviana Cordeddu, Bradley J Williams, Maria L Dentici, Viviana Caputo, et al.

► **To cite this version:**

Francesca Lepri, Alessandro de Luca, Lorenzo Stella, Cesare Rossi, Giuseppina Baldassarre, et al.. SOS1 MUTATIONS IN NOONAN SYNDROME: AN UPDATE ON MOLECULAR SPECTRUM, STRUCTURAL INSIGHTS ON PATHOGENIC EFFECTS, AND GENOTYPE-PHENOTYPE CORRELATIONS. *Human Mutation*, 2011, 32 (7), pp.760. 10.1002/humu.21492 . hal-00636633

HAL Id: hal-00636633

<https://hal.science/hal-00636633>

Submitted on 28 Oct 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



SOS1 MUTATIONS IN NOONAN SYNDROME: AN UPDATE ON MOLECULAR SPECTRUM, STRUCTURAL INSIGHTS ON PATHOGENIC EFFECTS, AND GENOTYPE-PHENOTYPE CORRELATIONS

Journal:	<i>Human Mutation</i>
Manuscript ID:	humu-2010-0595.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	14-Feb-2011
Complete List of Authors:	<p>Lepri, Francesca; IRCCS Casa Sollievo della Sofferenza, Laboratorio Mendel De Luca, Alessandro; IRCCS Casa Sollievo della Sofferenza, Laboratorio Mendel Stella, Lorenzo; Universita' "Tor Vergata", Dipartimento di Scienze e Tecnologie Chimiche Rossi, Cesare; Policlinico S.Orsola-Malpighi, UO Genetica Medica Baldassarre, Giuseppina; Universita' di Torino, Pediatria Pantaleoni, Francesca; Istituto Superiore di Sanita', Ematologia, Oncologia e Medicina Molecolare Cordeddu, Viviana; Istituto Superiore di Sanita', Ematologia, Oncologia e Medicina Molecolare Williams, Bradley; GeneDx Dentici, Maria; IRCCS Casa Sollievo della Sofferenza, Laboratorio Mendel Caputo, Viviana; Istituto Superiore di Sanita', Ematologia, Oncologia e Medicina Molecolare Venanzi, Serenella; Istituto Superiore di Sanita', Ematologia, Oncologia, Medicina Molecolare Bonaguro, Michela; Policlinico S.Orsola-Malpighi, UO Genetica Medica Kavamura, Maria; Federal University of Sao Paulo, Medical Genetics Faienza, Maria; University of Bari, Department of Biomedicine of Developmental Age Pilotta, Alba; Ospedale Pediatrico, Auxoendocrinologia Stanzial, Franco; Ospedale di Bolzano, Servizio aziendale di Consulenza Genetica Faravelli, Francesca; Galliera Hospital, Department of Human Genetics Gabrielli, Orazio; Università Politecnica delle Marche, Istituto di Scienze Materno-Infantili Marino, Bruno; Policlinico "Umberto I", Universita' "La Sapienza",</p>

	<p>Pediatria Neri, Giovanni; Universita Cattolica del Sacro Cuore Silengo, Margherita; Universita' di Torino, Pediatria Ferrero, Giovanni; Universita' di Torino, Pediatria Torrente, Isabella; IRCCS Casa Sollievo della Sofferenza, Laboratorio Mendel Selicorni, Angelo; Paediatric Clinical Genetics, MBBM AO S. Gerardo Monza Foundation, University of Milano Bicocca Mazzanti, Laura; Università di Bologna, Policlinico S. Orsola- Malpighi, Pediatria Digilio, Maria; Ospedale Pediatrico Bambino Gesù Zampino, Giuseppe; Universita' Cattolica del Sacro Cuore, Clinica Pediatria Dallapiccola, Bruno; Ospedale Pediatrico "Bambino Gesù", IRCCS Gelb, Bruce; Mount Sinai School of Medicine, Pediatrics and Human Genetics Tartaglia, Marco; Istituto Superiore di Sanita', Ematologia, Oncologia e Medicina Molecolare</p>
Key Words:	Noonan syndrome, SOS1, mutation analysis, genotype-phenotype correlations, structural analysis

SCHOLARONE™
Manuscripts

Review

Humu-2010-0595

Research Article

Supporting Information for this preprint is available from the
Human Mutation editorial office upon request (humu@wiley.com)

***SOS1* mutations in Noonan syndrome: molecular spectrum, structural insights on pathogenic effects, and genotype-phenotype correlations**

Francesca Lepri,^{1,17} Alessandro De Luca,¹ Lorenzo Stella,² Cesare Rossi,³ Giuseppina Baldassarre,⁴ Francesca Pantaleoni,⁵ Viviana Cordeddu,⁵ Bradley J Williams,⁶ Maria L. Dentici,^{1,17} Viviana Caputo,⁵ Serenella Venanzi,⁵ Michela Bonaguro,³ Ines Kavamura,⁷ Maria F. Faienza,⁸ Alba Pilotta,⁹ Franco Stanzial,¹⁰ Francesca Faravelli,¹¹ Orazio Gabrielli,¹² Bruno Marino,¹³ Giovanni Neri,¹⁴ Margherita Cirillo Silengo,⁴ Giovanni B. Ferrero,⁴ Isabella Torrente,¹ Angelo Selicorni,¹⁵ Laura Mazzanti,¹⁶ Maria C. Digilio,¹⁷ Giuseppe Zampino,¹⁸ Bruno Dallapiccola,¹⁷ Bruce D. Gelb,¹⁹ and Marco Tartaglia^{5*}

¹IRCCS Casa Sollievo della Sofferenza, Laboratorio Mendel, San Giovanni Rotondo, Italy;

²Dipartimento di Scienze e Tecnologie Chimiche, Università "Tor Vergata", Rome, Italy; ³UO

Genetica Medica, Policlinico S.Orsola-Malpighi, Bologna, Italy; ⁴Dipartimento di Pediatria,

Università di Torino, Turin, Italy; ⁵Dipartimento di Ematologia, Oncologia e Medicina

Molecolare, Istituto Superiore di Sanità, Rome, Italy; ⁶GeneDx, Gaithersburg, MD; ⁷Medical

Genetics, Federal University of Sao Paulo, Sao Paulo, Brazil; ⁸Department of Biomedicine of

1
2
3 Developmental Age, University of Bari, Bari, Italy; ⁹Auxoendocrinologia, Ospedale Pediatrico,
4 Brescia, Italy; ¹⁰Servizio aziendale di Consulenza Genetica, Ospedale di Bolzano, Italy; ¹¹S.C.
5
6 Genetica Umana, Ospedali Galliera, Genova, Italy ¹²Istituto di Scienze Materno-Infantili,
7
8 Università Politecnica delle Marche, Ancona, Italy; ¹³Division of Pediatric Cardiology,
9
10 Department of Pediatrics, “Sapienza” University, Rome, Italy; ¹⁴Istituto di Genetica Medica,
11
12 Università Cattolica del Sacro Cuore, Rome, Italy; ¹⁵Clinica Pediatrica I, IRCCS, Fondazione
13
14 Policlinico Milano, Milan, Italy; ¹⁶Dipartimento di Pediatria, Università degli Studi di Bologna,
15
16 Bologna, Italy; ¹⁷Ospedale Pediatrico “Bambino Gesù”, IRCCS, Rome, Italy; ¹⁸Istituto di Clinica
17
18 Pediatrica, Università Cattolica del Sacro Cuore, Rome, Italy; ¹⁹Child Health and Development
19
20 Institute, Mount Sinai School of Medicine, New York, NY.
21
22
23
24
25
26
27
28
29
30
31

32 ***Correspondence:**

33
34 Marco Tartaglia, Ph.D.

35
36 Department of Hematology, Oncology and Molecular Medicine

37
38 Istituto Superiore di Sanità

39
40 Viale Regina Elena, 299

41
42
43 00161 Rome, Italy

44
45
46 phone: +39 06 49902569, fax: + 39 06 49902850, email: mtartaglia@iss.it
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Noonan syndrome (NS) is among the most common non-chromosomal disorders affecting development and growth. NS is caused by aberrant RAS-MAPK signaling and is genetically heterogeneous, which explains, in part, the marked clinical variability documented for this Mendelian trait. Recently, we and others identified *SOS1* as a major gene underlying NS. Here, we explored further the spectrum of *SOS1* mutations and their associated phenotypic features. Mutation scanning of the entire *SOS1* coding sequence allowed the identification of 30 different variants deemed to be of pathological significance, including 13 novel missense changes and in-frame indels. Various mutation clusters destabilizing or altering orientation of regions of the protein predicted to contribute structurally to the maintenance of autoinhibition were identified. Two previously unappreciated clusters predicted to enhance *SOS1*'s recruitment to the plasma membrane, thus promoting a spatial reorientation of domains contributing to inhibition, were also recognized. Genotype-phenotype analysis confirmed our previous observations, establishing a high frequency of ectodermal anomalies and a low prevalence of cognitive impairment and reduced growth. Finally, mutation analysis performed on cohorts of individuals with non-syndromic pulmonic stenosis, atrial septal defects and ventricular septal defects excluded a major contribution of germline *SOS1* lesions to the isolated occurrence of these cardiac anomalies.

Key Words: Noonan syndrome; NS; *SOS1*; mutation analysis; structural analysis; genotype-phenotype correlations.

INTRODUCTION

Noonan syndrome (NS; MIM# 163950) is a relatively common and clinically variable disorder characterized by postnatal reduced growth, facial dysmorphism, and congenital heart defects (CHDs) [Allanson, 1987; Noonan, 1994; van der Burgt, 2007; Tartaglia et al., 2010]. The distinctive and most recurrent facial features consist of a broad forehead, hypertelorism, down-slanting palpebral fissures, ptosis, high arched palate, and low-set, posteriorly rotated ears. Cardiac involvement is present in up to 80-90% of affected individuals, with pulmonic stenosis (PS), septal defects and hypertrophic cardiomyopathy (HCM) occurring most commonly [Burch et al., 1993; Marino et al., 1999]. Other associated features include multiple skeletal defects (chest and spine deformities), webbed/short neck, variable cognitive deficits, cryptorchidism, lymphatic dysplasia, bleeding diathesis, and, rarely, predisposition to certain hematologic malignancies during childhood.

NS is genetically heterogeneous, and, based upon the recent discoveries of the underlying disease genes, is now regarded as a disorder caused by enhanced signal flow through the RAS-MAPK pathway [Tartaglia et al., 2010]. This signaling cascade is known to mediate diverse biological functions, including cell proliferation, survival, fate determination and differentiation. It is activated in response to cytokine, hormone and growth factor stimulation, and is a major mediator of early and late developmental processes including morphology determination, organogenesis, synaptic plasticity processes and growth. In approximately 50% of affected individuals, NS is caused by heterozygous missense mutations in the *PTPN11* gene [Tartaglia et al. 2001, 2002], which encodes a cytoplasmic protein tyrosine phosphatase positively modulating RAS function. Activating mutations in five additional genes coding for transducers or modulatory proteins participating in this signaling pathway (*i.e.*, *KRAS*, *NRAS*, *SOS1*, *RAF1*,

1
2
3 and *BRAF*) account for an additional one-fourth of NS cases [Tartaglia et al., 2010]. Mutations in
4 these and other functionally related genes (*i.e.*, *CBL*, *NF1*, *KRAS*, *HRAS*, *BRAF*, *SPRED1*,
5
6 *SHOC2*, *MAP2K1*, and *MAP2K2*) have been reported to underlie clinically related disorders
7
8 [Aoki et al., 2008; Tidyman and Rauen, 2009; Tartaglia and Gelb, 2010]. Genotype-phenotype
9
10 correlation surveys have documented that the substantial phenotypic variation characterizing NS
11
12 can be ascribed, in part, to the gene mutated and even to the specific molecular lesion involved.
13
14
15

16
17 We and others recently reported that missense mutations in *SOS1* (MIM# 182530) account
18
19 for a significant proportion of NS [Roberts et al., 2007; Tartaglia et al., 2007; Zenker et al.,
20
21 2007a]. *SOS1* encodes a guanine nucleotide exchange factor (GEF) responsible for stimulating
22
23 the conversion of RAS from the inactive, GDP-bound to the active, GTP-bound form [Nimnual
24
25 and Bar-Sagi, 2002]. *SOS1* is a large multidomain protein characterized by an *N*-terminal
26
27 regulatory portion including tandem histone-like folds (HF), which are followed by a Dbl-
28
29 homology (DH) domain and a pleckstrin-homology (PH) domain, and a *C*-terminal catalytic
30
31 region including the RAS exchanger motif (REM) and CDC25 domains, followed by a tail
32
33 providing docking sites for adaptor proteins required for receptor anchoring (Figure 1). The
34
35 majority of NS-causing *SOS1* mutations were observed to affect residues predicted to be
36
37 implicated in the maintenance of *SOS1* in its autoinhibited conformation, and the first
38
39 biochemical characterizations of mutants consistently documented enhanced protein function and
40
41 increased signal flow through RAS [Roberts et al., 2007; Tartaglia et al., 2007]. These surveys
42
43 also indicated that subjects heterozygous for a mutated *SOS1* allele tend to exhibit a distinctive
44
45 phenotype that is characterized by ectodermal abnormalities generally associated with an
46
47 absence of cognitive deficits [Tartaglia et al., 2007; Zenker et al., 2007a]. We also observed that
48
49 height was less frequently below the 3rd centile compared with the overall NS population
50
51
52
53
54
55
56
57
58
59
60

1
2
3 [Tartaglia et al., 2007]. While available information supports the view that *SOS1* is not mutated
4
5 in cardiofaciocutaneous syndrome (CFCS; MIM# 115150) [Zenker et al., 2007a], a condition
6
7 clinically related to NS, a few individuals with ectodermal manifestations and distinctive facial
8
9 dysmorphism that might be suggestive of CFCS have recently been reported as having *SOS1*
10
11 mutations [Narumi et al., 2008; Nyström et al., 2008]. In these subjects, cognitive deficits were
12
13 generally absent or minor, but at least one case with mental retardation was reported [Narumi et
14
15 al., 2008].
16
17
18
19

20 Here, we characterized further the molecular spectrum and distribution of *SOS1* mutations,
21
22 as well as the phenotypic features associated with those molecular defects. We also explored
23
24 possible involvement of germline *SOS1* mutations in an opportunely selected group of isolated
25
26 CHDs that occur as an associated feature in subjects with NS and heterozygous for a *SOS1*
27
28 mutation. Our results provide a more accurate description of the spectrum of *SOS1* gene defects,
29
30 their consequence on *SOS1* structure and function, and their associated clinical features.
31
32
33
34
35

36 MATERIAL AND METHODS

37 Patients

38
39
40 Four cohorts were included in the study. The first cohort (group 1) included 143 clinically
41
42 well-characterized patients with NS enrolled in research protocols. Nearly all subjects of this
43
44 cohort were of European ancestry, with the majority being Italian. Within this group, subjects
45
46 were assessed by clinical geneticists experienced with NS and clinically related disorders (G.Z.,
47
48 M.C.D., L.M., B.D., G.B.F., M.C.S., A.S., I.K., G.N., M.F.F., A.P., F.S. and O.G.). Clinical
49
50 assessment included physical, anthropometric, neurologic and cardiac evaluations, as well as
51
52 accurate examination for craniofacial features, ophthalmologic and otorhinolaryngologic
53
54
55
56
57
58
59
60

1
2
3 defects, and ectodermal and musculoskeletal anomalies. Clinical features for the majority of
4 these individuals satisfied diagnostic criteria reported for NS [van der Burgt et al. 1994], but a
5 few individuals with a highly suggestive phenotype who lacked sufficient features to receive a
6 definitive diagnosis were also included. Based on scanning of the coding exons by denaturing
7 high-performance liquid chromatography (DHPLC) analysis and/or direct sequencing, no subject
8 within this cohort harbored a mutation in *PTPN11*, *KRAS* or *RAF1*. For approximately half of the
9 cases, mutations in *BRAF*, *MAP2K1*, *SHOC2*, *CBL* and *NRAS* had also been excluded. Besides
10 this large cohort, nine subjects with features fitting CFCS and no mutation in *KRAS*, *BRAF*,
11 *MAP2K1* or *MAP2K2* (group 2) [Sarkozy et al., 2009a] were also included in the study. The third
12 cohort (group 3) (N= 358) comprised anonymous samples from individuals with phenotypes
13 suggestive of NS for whom commercial DNA diagnostic testing was performed. Clinical data
14 were not available for these patients, and *PTPN11* mutations had not systematically been
15 excluded in all the subjects included in this group. While the output obtained from genotyping
16 this cohort of subjects could not be used in genotype-phenotype correlation studies or to estimate
17 *SOS1* mutation prevalence in the NS population, the mutation data were utilized to provide a
18 more detailed picture about the molecular spectrum of disease-causing mutations affecting the
19 *SOS1* gene. Finally, a cohort of 59 subjects with non-syndromic CHDs (PS, N= 21; atrial septal
20 defects (ASDs), N= 23; ventricular septal defects (VSDs), N= 15) was included in the study
21 (group 4). Clinical assessment of patients with isolated CHDs included complete physical
22 evaluation of dysmorphism and malformations, anthropometric measurements, renal
23 ultrasonography, and radiological studies. Cardiac evaluation included preoperative physical
24 evaluation, chest X-ray film, 12-lead electrocardiogram, and two-dimensional transthoracic
25 echocardiography with color flow Doppler. Karyotype analysis was performed in all patients of
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 this cohort. Inclusion criteria were based on the absence of any association with other clinical
4
5 features, and chromosomal anomalies, including the 22q11 deletion.
6
7

8 Informed consent for the genetic analyses was obtained from all patients or their legal
9
10 guardians.
11
12

13 14 15 **Mutation analysis**

16
17 Genomic DNA was isolated from peripheral blood leukocytes according to standard
18
19 procedures. The entire *SOS1* coding sequence, as well as the exon/intron boundaries and flanking
20
21 intronic portions were scanned for mutations by direct sequencing (group 3) or DHPLC analysis
22
23 (groups 1, 2, and 4) with the use of the 3100 and/or 3500HT Wave DNA Fragment Analysis
24
25 System (Transgenomic, Omaha, NE), at column temperatures recommended by the Navigator
26
27 version 1.5.4.23 software (Transgenomics), as previously described [Tartaglia et al., 2007].
28
29 Amplimers having abnormal elution profiles were re-amplified, purified (Qiagen, Hilden,
30
31 Germany) and sequenced using the ABI BigDye Terminator Sequencing Kit v.1.1 (Applied
32
33 Biosystems, Foster City, CA) and an ABI 3700 Capillary Array Sequencer or ABI 3100 Genetic
34
35 Analyzer (Applied Biosystems). Primer pair sequences as well as PCR and DHPLC analysis
36
37 settings are available upon request. Length of deletions and dinucleotide mutations were
38
39 determined by cloning purified PCR products in a pCR 2.1 TOPO vector (Invitrogen, Carlsbad,
40
41 CA) and sequencing purified clones (Plasmid Mini Kit, Qiagen). Nucleotide numbering of the
42
43 mutations and exonic disease-unrelated variants reflects cDNA numbering with 1 corresponding
44
45 to the A of the ATG translation initiation codon in the reference sequence (NM_005633.3), while
46
47 position of the intronic variants were numbered according to the reference genomic sequence
48
49 (NG_007530.1).
50
51
52
53
54
55
56
57
58
59
60

1
2
3 The level of conservation of affected residues among orthologous *SOS1* genes was
4 evaluated by using the NCBI HomoloGene tool (<http://www.ncbi.nlm.nih.gov/homologene>),
5 while the SIFT (Sorting Intolerant From Tolerant, <http://blocks.fhrc.org/sift/SIFT.html>) [Ng and
6 Henikoff, 2001] and PolyPhen (Polymorphism Phenotyping,
7 <http://genetics.bwh.harvard.edu/pph/>) [Sunyaev et al., 2000] software programs were used to
8 predict the biological relevance of the identified missense variants on protein function.
9

10
11
12
13
14
15
16
17
18 When available, parental DNAs were sequenced to establish whether the identified changes
19 in sporadic cases were *de novo* or to confirm co-segregation between the variant and phenotype
20 in families transmitting the trait. Paternity was confirmed using the AmpDESTER ProfilerPlus
21 kit (Applied Biosystems). All missense changes proven to be *de novo* by parental DNA
22 genotyping were deemed mutations causally linked to the disorder. To exclude the possibility
23 that variants co-segregating with NS were neutral polymorphic changes occurring in the
24 population, at least 300 population-matched DNAs obtained from unaffected subjects were
25 screened (DHPLC analysis and sequencing of variant elution profiles). When parental DNAs
26 were not available, we considered a novel non-synonymous variant as a causative mutation when
27 it affected a residue already reported to be mutated. Novel variants involving residues localized
28 in close proximity to residues mutated in NS, predicted to affect protein structure/function by
29 Polyphen, SIFT and structural evidence, but without sufficient genetic evidence (*i.e.*, cases with
30 unreported ethnicity or missing clinical evaluation and/or genetic testing of other members of
31 their families) were considered as probably pathogenic. These missense changes were found not
32 to occur in controls, and resulted in a non-conservative substitution at an invariant or highly
33 conserved residue (*Homo sapiens* [NP_005624.2], *Pan troglodytes* [XP_515425.2], *Canis*
34 *familiaris* [XP_540157.2], *Bos taurus* [XP_617859.4], *Mus musculus* [NP_033257.2], *Rattus*
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 *norvegicus* [XP_233820.4], and *Drosophila melanogaster* [NP_476597.2]). Finally, a third
4
5 category grouping functionally unclassified variants was considered for those novel changes for
6
7 which structural analysis did not allow to infer any functional effect.
8
9

10 The position of mutated amino acid residues was modeled by using the recently generated
11
12 crystal structure of the SOS1 protein truncated at the C-terminus (residues 1-1049) deposited in
13
14 the RCSB Protein Data Bank (<http://www.rcsb.org/pdb/home/home.do>) (PDB ID: 3KSY)
15
16 [Guerasko et al., 2010]. Figures were prepared using the UCSF Chimera 1.5 software package
17
18 (<http://www.cgl.ucsf.edu/chimera/>) [Pettersen et al., 2004]. Electrostatic potential calculations
19
20 were performed with the APBS software [Baker et al., 2001].
21
22
23
24
25
26

27 **Statistical analyses**

28
29 Confidence intervals for proportions were calculated by means of VassarStats software
30
31 (<http://faculty.vassar.edu/lowry/VassarStats.html>) using the Newcombe-Wilson method
32
33 including continuity correction [Wilson, 1927; Newcombe, 1998]. Genotype-phenotype
34
35 correlations were performed using 2x2 contingency-table analysis. The significance threshold
36
37 was set at $P = 0.05$.
38
39
40
41
42

43 **RESULTS**

44 ***SOS1* mutation scanning in NS, CFCS and isolated CHD**

45
46 DHPLC analysis and bidirectional direct sequencing of the entire *SOS1* coding sequence
47
48 on peripheral blood leukocyte genomic DNA specimens of groups 1 and 3 identified 41 different
49
50 heterozygous missense nucleotide substitutions and 3 small in-frame indels in 108 unrelated
51
52 subjects. Among them, five were private or common variants unrelated with the trait. Three of
53
54
55
56
57
58
59
60

1
2
3 them, c.1964C>T (p.Pro655Leu), c.3032A>G (p.Asn1011Ser) and c.3959A>G (p.His1320Arg,
4
5 had previously been reported in unaffected individuals [Roberts et al., 2007; Tartaglia et al.,
6
7 2007]. Similarly, the c.1705C>G (p.Leu569Val) was documented to occur as a homozygous
8
9 change in an unaffected parent, and the c.2122G>A (p.Ala708Thr) was found to occur as a
10
11 polymorphic change in ethnic-matched unaffected individuals of Hispanic ancestry, and were
12
13 therefore regarded as benign polymorphisms. A full list of the disease-unrelated changes,
14
15 including silent nucleotide substitutions and intronic variants, is reported in Supp. Table S1.
16
17
18
19

20 In these two cohorts, twenty-five *SOS1* sequence variants were deemed to be of
21
22 pathological significance (Table 1). Among these, 17 had previously been established as
23
24 causative mutations (Supp. Table S2), whereas 8 were novel. Among the latter, 5 mutations were
25
26 single nucleotide substitutions, but three small in-frame indels affecting two mutational hot spots
27
28 within the PH domain (p.Lys427_Asp430delinsAsn and p.Trp432_Glu433del) and the helical
29
30 linker connecting the PH and REM domains (PH-REM linker) (p.Leu554_Met558delinsLys)
31
32 were also identified. The c.1300_1301GG>AA (p.Gly434Lys) and c.1310T>C (p.Ile437Thr)
33
34 changes were demonstrated to occur as *de novo* events by genotyping of parental DNAs. The
35
36 c.2681C>G transition (p.Pro894Arg) was found to co-segregate with the disease, while the
37
38 remaining two variants, c.1430A>G (p.Gln477Arg) and c.1655G>T (p.Arg552Met), affected
39
40 amino acid residues that had previously been reported to be mutated in NS (Supp. Table S2).
41
42 Eight novel missense changes (p.Pro112Arg, p.Ile252Thr, p.Met422Val, p.Glu424Lys,
43
44 p.Gly482Arg, p.Leu490Arg, p.Arg497Gln and p.Thr549Lys) were considered as probably
45
46 pathogenic mutations. The majority of these variants were not observed in more than 300
47
48 population-matched unaffected individuals, and were non-conservative, the majority affecting
49
50 invariant residues among *SOS1* orthologs and predicted to be “damaging” or have functional
51
52
53
54
55
56
57
58
59
60

1
2
3 relevance according to the PolyPhen or/and SIFT programs. Moreover, affected residues were
4 located in mutational hot spots or in close spatial proximity to residues mutated in NS and were
5 predicted to have structural/functional consequences on protein structure (see below). For these
6 variants, however, the evidence provided by the theoretical models and structural analysis was
7 not unambiguously supported by genetic evidence. Finally, six missense nucleotide substitutions
8 (p.Thr37Ala, p.Pro478Leu, p.Ile784Thr, p.Arg1131Lys, p.Leu1140Ile and p.Thr1257Ala) were
9 classified as variants of unknown significance. For those, parental DNAs were not submitted for
10 testing or no clinical information was available for the carrier parent, and the predicted amino
11 acid change was generally not considered as having functional relevance according to the
12 PolyPhen or/and SIFT programs. Most of these changes affected the C-terminal portion of the
13 protein for which no structural information is currently available, precluding any consideration
14 regarding possible consequences on protein function.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

31
32 *SOS1* mutations were identified in 26 of the 143 subjects (18.2%) of group 1, which
33 included patients with clinical features satisfying the diagnostic criteria of NS or highly
34 suggestive for the disorder, and negative for mutations in *PTPN11*, *RAF1*, and *KRAS*. Based on
35 the accurate clinical assessment of this cohort and the relative prevalence of mutations affecting
36 those disease genes in NS, this finding supports the view that mutations in the *SOS1* gene
37 account for approximately 10% of NS cases, which is in line with our previous estimate
38 [Tartaglia et al., 2007], but slightly below those obtained from two other clinically well
39 characterized NS cohorts [Roberts et al., 2007; Zenker et al., 2007a].
40
41
42
43
44
45
46
47
48
49

50
51 DHPLC mutation scanning of coding exons, splice junctions, and flanking intronic portions
52 of *SOS1* did not reveal any putative pathogenic mutation in the nine subjects with a diagnosis of
53 CFCS and negative for mutations in *BRAF*, *KRAS*, *MAP2K1*, and *MAP2K2*. Similarly, no
54
55
56
57
58
59
60

1
2
3 mutation was identified among the 59 patients with non-syndromic CHDs included in the study,
4
5 excluding a major involvement of germline *SOS1* mutations in PS (N= 21; 95% CI= 0.000-
6
7 0.192), ASD (N= 23, 95% CI= 0.000-0.178), and VSD (N= 15, 95% CI= 0.000-0.253).
8
9

10 11 12 ***SOS1* mutation diversity, cluster distribution, and molecular modeling analysis**

13
14
15 The present data and available published records (updated to July 2010) (Supp. Table S2)
16
17 were utilized to analyze the diversity of NS-causing mutations, their cluster distribution, and to
18
19 explore their effects on *SOS1* structure and function. The 183 reported missense mutations
20
21 (including those with probably damaging effects) were observed to affect 32 residues, which
22
23 were not randomly scattered along the protein but tended to cluster in specific regions (Figure 1).
24
25
26 Approximately forty percent of *SOS1* defects affected four residues located in the PH-REM
27
28 linker (Ser548, Thr549, Leu550 and Arg552), with substitutions of residue Arg552 accounting
29
30 for one-third of all mutations. Another mutation cluster, involving short stretches within the PH
31
32 domain (*i.e.*, Glu424, Trp432, Glu433 and Gly434, and Gln477, Pro478 and Gly482) explained
33
34 roughly 20% of mutations. A third functional cluster (see below) resided at the interacting
35
36 regions of the DH (Thr266 and Met269) and REM (Lys728, Trp729 and Ile733) domains (16%
37
38 of total mutations). Among the most recurrent mutations, the c.2536G>A transition, predicting
39
40 the substitution of Glu846 by lysine within the CDC25 domain, accounted for 10% of defects.
41
42
43 Such a skewed distribution of affected residues implies specific perturbing consequences on
44
45
46
47
48
49 protein function.

50
51 *SOS1*'s GEF activity is controlled principally by two binding sites for RAS: the catalytic
52
53 site, which is located entirely within the CDC25 domain and promotes GDP release from RAS,
54
55 and a distal site, which is bracketed by the CDC25 domain and REM domains and positively
56
57
58
59
60

1
2
3 modulates GEF activity through promotion of a conformational change at the active site that
4 allows substrate GDP-RAS to bind [Margarit et al., 2003]. Basally, the catalytic output of SOS1
5 is constrained by the *N*-terminal regulatory HF domain and DH-PH unit [Sondermann et al.,
6 2004; Gureasko et al., 2010], and structural data indicate that this autoinhibitory effect is exerted
7 through direct HF and DH domain-mediated blockade of the allosteric site. An extensive
8 interdomain binding network also involving the PH-REM helical linker and the PH domain has
9 been recognized to stabilize this inhibitory conformation. Following the translocation of SOS1 to
10 the membrane, the inhibitory effect of the HF and DH domains is relieved allowing RAS binding
11 to the allosteric site, which in turn promotes a conformational rearrangement of the CDC25
12 domain allowing RAS binding to the catalytic site. Based on this allosteric mechanism of
13 activation and recent evidence indicating that the HF domain and the DH-PH unit are
14 conformationally coupled to control SOS1's recruitment to the plasma membrane and release of
15 autoinhibition [Gureasko et al., 2010], the distribution of identified mutations in NS supports a
16 picture in which the vast majority of disease-causing lesions affect the stability of the
17 intramolecular interactions that maintain SOS1 in an autoinhibited state by at least two major
18 distinct mechanisms.

19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41 A first class of mutations includes lesions predicted to promote conformational
42 rearrangements of domains that reduce the enzyme self-inhibition by impairing proper masking
43 of the distal RAS binding site or by acting on the allosteric control of catalytic activity. Within
44 this class, a first group of mutations (Class 1A: p.Thr266Lys, p.Met269Arg/Thr, p.Lys728Ile,
45 p.Trp729Leu and p.Ile733Phe) involve residues that participate in the autoinhibitory interaction
46 of the DH and REM domains blocking RAS access at the allosteric site, or are closely located to
47 them (Figure 2A). These mutations are predicted to affect the stability of the inactive
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 conformation of the protein directly by disrupting the inhibitory interdomain bonding network at
4 the distal site. Among these, the most recurrent mutations affect Met269 (10% of total cases),
5 which interacts directly with residues of the REM domain implicated in RAS binding
6 [Sondermann et al., 2004]. Of note, Lys728 and Trp729 are among these, and Ile733 is very
7 close to these residues; therefore, their substitution could also affect the REM's affinity for GTP-
8 bound RAS in addition to the stability of the DH-REM interface. Indeed, the drastic
9 p.Trp729Glu substitution has been shown to inhibit RAS binding to the REM domain
10 [Sondermann et al., 2004]. This might be related to the fact that within this group, mutations
11 affecting residues located at the DH surface appear to be more common (15% of total changes)
12 compared to those affecting the REM surface (2%). Among the probably pathogenetic variants,
13 the p.Ile252Thr substitution could be included in this group. Ile252 is an invariant residue placed
14 within a hydrophobic core of the DH domain, which also involves Thr215, Leu219, Ile249,
15 Tyr295 and Tyr298. Structural perturbation of the DH fold is expected to destabilize the masking
16 of the distal RAS binding site.

17
18 A second mutation group includes lesions affecting the interaction between the HF, DH
19 and PH domains, thus perturbing the overall autoinhibited conformation in which the HF and DH
20 domains block the distal RAS binding site [Guerasko et al., 2010] (Class 1B: p.Lys170Glu,
21 p.Tyr337Cys, p.Ile437Thr, p.Cys441Tyr, p.Ser548Arg, p.Leu550Pro, p.Arg
22 552Gly/Thr/Met/Lys/Ser, p.Leu554_Met558delinsLys, and probably pathogenetic variants
23 p.Met422Val, p.Arg497Gln and p.Thr549Lys,) (Figure 2B). Within this group, mutations of
24 residues 548-558 are the most recurrent. This stretch corresponds to the helical linker connecting
25 the PH and REM domains, which contributes structurally to the maintenance of the autoinhibited
26 conformation by interacting with the HF and DH domains [Guerasko et al., 2010]. For

1
2
3 instance, Arg552 interacts directly with the side chains of Asp140 and Asp169 of the HF
4 domain, and mutation of either Arg552 or Asp140 was demonstrated to completely abolish the
5 interaction between these domains [Sondermann et al., 2005]. A similar effect might be predicted
6 for the recurrent Lys-to-Glu change affecting codon 170 within the HF domain, and possibly for
7 the p.Arg497Gln substitution in the PH domain, since the electrostatic interactions between these
8 two residues located at the HF-PH domain interface (but also close to the helical linker)
9 contribute to stabilizing the HF domain's orientation. Similarly, substitution of residues Ile437
10 and Cys441 might affect the PH domain's structure in the region of the domain facing towards
11 the HF domain and possibly cause a reorientation of the latter. Finally, residue Tyr337 is located
12 at the PH-DH interface, and thus probably contributes to stabilizing the DH domain orientation.
13
14 The p.Phe78Cys change (and possibly p.Met422Val) can also be included in this group.
15
16 Substitution of the highly conserved Phe78, which is an unexposed HF residue contributing to a
17 hydrophobic core (involving Leu55, Leu59, Val74, Val133 and Ile137), might structurally
18 perturb the region of the HF domain at its interface with the helical linker. The p.Met422Val
19 change, on the other hand, might perturb the structure of the PH domain, thus perturbing the PH-
20 HF interaction.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40
41 The third group of class 1 mutations is formed by residues of the REM domain, interacting
42 with the helical hairpin of the CDC25 domain, which contributes to the allosteric structural
43 switch [Freedman et al., 2006], and whose conformation is essential for both RAS binding to the
44 active site and nucleotide exchange [Hall et al., 2001; Margarit et al., 2003] (Class 1C:
45 p.Phe623Ile and p.Tyr702His) (Figure 2C). A single mutation has been identified to affect
46 Phe623, a highly conserved residue of the REM domain located at the interface with the CDC25
47 domain. This residue is part of an extended hydrophobic groove of the REM domain that
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 accommodates the side chains of two hydrophobic residues (Ile956 and Phe958) of the helical
4
5 hairpin of the CDC25 domain. Such a hydrophobic interaction has been demonstrated to be
6
7 important for the correct orientation of the helical hairpin, and for SOS1's catalytic activity [Hall
8
9 et al., 2001]. Based on these observations, it can be hypothesized that the p.Phe623Ile
10
11 substitution might perturb the catalytic activity of the CDC25 domain by acting either on RAS
12
13 binding at the catalytic site, on the domain's catalytic efficiency or on allosteric control. A
14
15 similar effect might be associated with substitution of the invariant Tyr702, which is located
16
17 within the REM domain at the interface with the CDC25 domain, in close proximity to the
18
19 region of the latter that undergoes the conformational switch promoted by RAS binding at the
20
21 distal site.
22
23
24
25
26

27 While the primary anchorage of SOS1 to the plasma membrane is mediated by docking of
28
29 its C-terminal region to SH3 domain-containing adaptor proteins (*e.g.*, GRB2) that bind to the
30
31 activated receptors, additional anchorage sites at the membrane are provided by the
32
33 phosphatidylinositol-4,5-phosphate (PIP₂)- and phosphatidic acid (PA)-binding sites within the
34
35 PH domain [Chen et al. 1997; Zheng et al., 1997; Zhao et al., 2007], and by an extended
36
37 positively charged surface of the HF domain that interacts with anionic membranes, and possibly
38
39 also binds PA and PIP₂ [Guerasko et al., 2010; Yadav and Bar-Sagi, 2010]. These sites appear to
40
41 have complementary roles: while those within the PH domain mediate SOS1's targeting to the
42
43 membrane, the latter are thought to activate the GEF at the membrane. In the autoinhibited
44
45 structure of SOS1, the positively charged site of the HF domain is not oriented in a way that
46
47 would allow membrane binding [Guerasko et al., 2010]. This finding suggests the possibility that
48
49 the reorientation of the PA-binding site of the HF domain might be coupled to the destabilization
50
51 of the auto-inhibited conformation of the protein, permitting SOS1's activation through RAS
52
53
54
55
56
57
58
59
60

1
2
3 binding at the allosteric site [Guerasko et al., 2010; Yadav and Bar-Sagi, 2010]. Class 2
4
5 mutations include changes that are predicted to enhance SOS1's catalytic function by membrane-
6
7 dependent mechanisms. Within this class, a first group is predicted to perturb the self-inhibiting
8
9 orientation of the HF domain by affecting solvent exposed residues located within the positively
10
11 charged surface of the HF domain, thus favoring its membrane binding (Class 2A: p.Pro102Arg
12
13 and p.Glu108Lys, and probably pathogenetic variant p.Pro112Arg) (Figure 2D). Significantly,
14
15 these mutations introduce a positively charged amino acid (arginine or lysine) enhancing the
16
17 positive electrostatic potential of the HF surface. In particular, Glu108 is adjacent to the PA-
18
19 binding motif and its substitution by lysine generates a contiguous patch of positively charged
20
21 residues that was recently demonstrated to potentiate membrane binding [Yadav and Bar-Sagi,
22
23 2010]. A similar perturbing effect would be predicted for the c.305C>G and c.335C>G missense
24
25 changes affecting the closely located Pro102 and Pro112 residues.
26
27
28
29
30
31

32 A second group includes mutations affecting the membrane-binding surface of the PH
33
34 domain (Class 2B: p.Lys427_Asp430delinsAsn, p.Trp432Arg, p.Trp432_Glu433del,
35
36 p.Glu433Lys, p.Gly434Arg/Lys, p.Gln477Arg/His, p.Pro478Arg,
37
38 p.Pro481_Gly482insArgLeuPro, and probably pathogenetic variants p.Glu424Lys, p.Gly482Arg
39
40 and p.Leu490Arg) (Figure 2E). As anticipated, two sites within the PH domain have been
41
42 identified [Chen et al. 1997; Zheng et al., 1997; Zhao et al., 2007]. Both sites contribute to
43
44 membrane recruitment of SOS1 and are required for efficient GEF activity. We noticed that the
45
46 vast majority of mutations affecting the PH domain involved solvent-exposed residues within
47
48 two adjacent regions (*i.e.*, residues 424-436 and 473-493) that do not overlap spatially with the
49
50 PIP₂-binding site, which is constituted by residues Lys456, Arg459, Lys472 and Arg489 [Zheng
51
52 et al., 1997]. A first cluster of mutations affected residues Gln477, Pro478, Pro481 and
53
54
55
56
57
58
59
60

1
2
3 Gly482, which are placed within the PA-binding site encompassing residues 472-483 [Zhao et
4 al., 2007], while the second cluster affect residues Glu424, Lys427-Asp430, Trp432-Gly434 and
5
6 Leu490. Of note, the nature of the substitution seemed to be critical since all missense mutations
7
8 affecting these regions resulted in the introduction of a positively charged residue. In two cases
9
10 (Gly434, Gln477), multiple lesions were predicted to introduce either an arginine or a lysine,
11
12 further indicating a specific role for the introduced charged residue. In additional two cases,
13
14 including the recurrent c.1297G>A (p.Glu433Lys), the positively charged residue replaced a
15
16 negatively charged amino acid. Similarly, the three identified in-frame indels were predicted to
17
18 increase the electrostatic potential of the postulated membrane interaction surface. Based on
19
20 established evidence supporting an absolute requirement of SOS1's recruitment to the plasma
21
22 membrane for RAS activation in response to an extracellular stimulus [Zhao et al., 2007], this
23
24 non-random distribution and type of lesions strongly suggest that their pathogenetic effect is
25
26 related to a strengthened binding of SOS1 to PA, PIP₂ and/or other membrane phospholipids, and
27
28 consequently to a more stable recruitment of the protein at the membrane, which in turn would
29
30 enhance the duration and/or amplitude of GEF function.
31
32
33
34
35
36
37
38

39 Only two disease-causing mutations have been found to affect the CDC25 domain so far.
40
41 The affected residues (*i.e.*, Glu846 and Pro894) are placed in close proximity to each other, and
42
43 distally from the two RAS binding sites and regions implicated in the conformational
44
45 rearrangement of the domain promoting RAS's binding to the active site (Figure 2F). The high
46
47 recurrence of the p.Glu846Lys change indirectly documents a relevant role of this region in
48
49 SOS1's function. This amino acid substitution has recently been documented to profoundly
50
51 perturb intracellular signaling [Chen et al., 2010]. Available crystal data, however, do not allow
52
53 us to infer any functional effect for these mutations, as well as for unclassified variants
54
55
56
57
58
59
60

1
2
3 affecting residues in this domain (Ile784, Gln977 and Ser1000) or at the C-terminus (Arg1131,
4
5 Leu1140 and Thr1257). We noted that Arg1131 is adjacent to Ser1132, which is one of the
6
7 identified growth factor-induced, MAPK-mediated phosphorylation sites of SOS1 [Corbalan-
8
9 Garcia et al., 1996]. These sites cluster within a small region that contains the proline-rich SH3-
10
11 binding sites implicated in GRB2 binding. Of note, phosphorylation of these sites has been
12
13 demonstrated to reduce SOS1's binding affinity for GRB2, and it could be speculated that
14
15 mutations altering the recognition motif at these sites might affect phosphorylation and promote
16
17 enhanced GRB2 binding. Similarly, Leu1140 and Thr1257 cluster within this region. Based on
18
19 evidence suggesting that this region possibly exerts an autoinhibitory effect on SOS1's activity
20
21 [Aronheim et al., 1994; Wang et al., 1995], it can be speculated that lesions affecting this region
22
23 might perturb such an inhibitory mechanism.
24
25
26
27
28

29 Finally, available structural data did not allow us to identify any functional clue for the
30
31 p.Asp309Tyr NS-causing substitution as well as for the unclassified sequence variants affecting
32
33 Thr37 and Thr378 residues.
34
35
36
37
38

39 **Phenotypic spectrum and genotype-phenotype correlations of *SOS1* mutations**

40 Extensive clinical information was obtained for 39 subjects with NS and *bona fide SOS1*
41
42 mutations, including individuals recruited in this study (group 1, cases NS01 to NS25) and re-
43
44 examined subjects of our original study (NS26 to NS39) [Tartaglia et al., 2007] (Supp. Table
45
46 S3). Overall, analysis of the clinical data confirmed previous observations from our group and
47
48 others indicating that mutations in *SOS1* are associated with a distinctive phenotype that
49
50 unambiguously falls within the NS clinical spectrum, but is characterized by a high prevalence of
51
52 ectodermal features (keratosis pilaris/hyperkeratotic skin, sparse eyebrows, and sparse,
53
54
55
56
57
58
59
60

1
2
3 generally thin and curly scalp hair) (84% of cases), and a relatively low occurrence of cognitive
4
5 deficits (11% of cases) compared to what is observed in the NS general population. Of note, in
6
7 two cases mental retardation was potentially attributable to critical illness during infancy. In two
8
9 additional subjects, the intelligence quotient was border-line to the lower limit of the normal
10
11 range with impairment of linguistic skills and oculo-manual coordination in one subject, and
12
13 delay of speech in the other (Supp. Table S3). We also confirmed our previous observation
14
15 indicating a relatively low prevalence of subjects exhibiting reduced growth (length/stature
16
17 below the 3rd centile) (29% of cases). In these subjects, reduced growth was invariably associated
18
19 with delayed bone age. Remarkably, more than one-third of subjects with mutated *SOS1* allele
20
21 exhibited fetal macrosomia, which, however, did not appear to correlate with the extent of their
22
23 postnatal growth, being length/stature in these subjects below the 3rd centile in a comparable
24
25 proportion of cases, compared with subjects without this feature (3/14 vs. 7/24, Fisher's exact
26
27 probability = 0.45). Of note, *SOS1* mutation-positive subjects displayed a more pronounced
28
29 growth failure (45% of cases) as newborns, which was not generally secondary to poor sucking
30
31 and/or swallowing.
32
33
34
35
36
37

38
39 *SOS1* mutation-positive subjects displayed typical facial features (Figure 3). Macrocephaly
40
41 was, however, overrepresented compared to the general NS population (61% vs. 12%) [Sharland
42
43 et al., 1992]. Among these subjects, cardiac defects were frequently observed (89% of cases),
44
45 with PS, ASD and VSD being the most recurrent anomalies, while prevalence of HCM was
46
47 comparable to that observed in *PTPN11* mutation-associated NS, occurring in less than 10% of
48
49 cases [Sarkozy et al., 2009b]. Of note, PS was found to be frequently associated with ASD (35%
50
51 of cases). Finally, two subjects were documented to present with mandibular multiple giant cell
52
53
54
55
56
57
58
59
60

1
2
3 lesions (MGCLs), which were associated with multiple tumors, including abdominal
4
5 rhabdomyosarcoma, cerebral glioma, and skin granular cell tumors in one subject.
6
7

8 The analysis of distribution of the major clinical features among *SOS1* mutation-positive
9
10 subjects documented a significantly higher prevalence of fetal macrosomia in subjects with class
11
12 1B mutations compared to individuals with class 1A mutations (Fisher's exact probability =
13
14 0.024). We failed in identifying any other significant genotype-phenotype correlation. Based on
15
16 the relatively small size of the cohort analyzed, however, the occurrence of other associations
17
18 between specific features and individual *SOS1* lesions or mutation clusters cannot be ruled out.
19
20
21
22
23

24 DISCUSSION

25
26
27 In this report, we have expanded the available information about the molecular diversity of
28
29 *SOS1* mutations underlying NS, and have provided a more comprehensive assessment of the
30
31 clinical features associated with those molecular lesions. We also explored systematically the
32
33 predicted structural consequences of NS-causing *SOS1* defects, developing a classification of
34
35 these gene lesions based on the predicted role of affected residues and functional consequences
36
37 derived from the nature of the amino acid change.
38
39
40

41 Combined with data from previous surveys, our findings indicate that *SOS1* mutations are
42
43 almost always missense changes, although we documented that small in-frame indels occur in a
44
45 small proportion of cases. Available mutation records (data updated to July 2010) revealed a
46
47 complete absence of nonsense, frameshift, and splicing defects, which is consistent with
48
49 functional characterization of a panel of NS-causing *SOS1* mutations [Roberts et al., 2007;
50
51 Tartaglia et al., 2007; Chen et al., 2010; Guerasko et al., 2010], and further support their
52
53 activating role on *SOS1* functional dysregulation in disease pathogenesis. Moreover, the
54
55
56
57
58
59
60

1
2
3 accumulated data strongly indicated that specificity in the amino acid substitution is relevant to
4
5 the functional dysregulation of the protein. Specifically, occurrence of invariant amino acid
6
7 changes was observed for several residues and accounted for approximately 50% of total events,
8
9 strongly suggesting a specific role for the introduced residue. Exemplifying this is the
10
11 observation of positively charged amino acids that were observed invariantly to replace solvent
12
13 exposed residues located within the HF and PH domains (*i.e.*, Glu108, Trp432, Glu433 and
14
15 Gly434), which provides strong evidence for the electrostatic nature of the perturbing effect of
16
17 those mutations on SOS1 functional dysregulation. On the other hand, the identity of substitution
18
19 did not seem to be critical in other cases (*e.g.*, most mutations affecting the helical linker
20
21 connecting the DH and PH domain), indicating a crucial role for the amino acid residue being
22
23 replaced. This is the case of Arg552, for which all but one amino acid substitutions resulting
24
25 from a single base change affecting this codon have been documented to cause NS.
26
27
28
29
30

31
32 Based on SOS1's crystal structure and mechanism of activation [Guerasko et al., 2008,
33
34 2010], the non-random distribution pattern of altered residues and nature of substitutions indicate
35
36 that NS-causing mutations dysregulate SOS1's GEF function by at least two major mechanisms.
37
38 As originally reported by Roberts *et al.* [2007] and Tartaglia *et al.* [2007], a large fraction of
39
40 mutations was found to cluster in regions of the protein that participate in the interdomain
41
42 binding network that maintain SOS1 in its catalytically inactive conformation (class 1
43
44 mutations). Among these, mutations can affect either residues directly implicated in the
45
46 autoinhibitory interaction between the DH and REM domains that impair GTP-RAS binding at
47
48 the distal site (or residues closely located to them), or residues located within the DH-PH helical
49
50 linker or regions of the HF, DH and PH domains that contribute to stabilizing the HF, DH and
51
52 REM interdomain binding network that maintains SOS1 in its catalytically inactive
53
54
55
56
57
58
59
60

1
2
3 conformation. Of note, this study recognized a third group of class 1 mutations that were found
4
5 to specifically affect residues of the REM domain located at the interface with regions of the
6
7 CDC25 domain identified to undergo the GTP-RAS binding-induced conformational
8
9 rearrangement required for both RAS binding to the active site and nucleotide exchange reaction.
10
11 While the effect of these amino acid substitutions on RAS binding at the catalytic site or on the
12
13 domain's catalytic efficiency cannot be ruled out, we hypothesize that these mutations might
14
15 upregulate SOS1's catalytic activity by weakening the inhibitory allosteric control on the active
16
17 site.
18
19
20
21

22 Strikingly, our analysis also allowed us to discern a previously unrecognized class of
23
24 mutations affecting solvent exposed residues located within the membrane oriented surface of
25
26 the PH domain and a positively charged surface of the HF domain, and predicted to enhance
27
28 SOS1's catalytic function by distinct membrane-dependent mechanisms. While the complex
29
30 process of SOS1's catalytic activation has not been characterized in detail, available structural
31
32 and functional data indicate that the two major events controlling SOS1's function, *i.e.*, protein
33
34 recruitment to the plasma membrane and release of autoinhibition, are linked. Membrane
35
36 translocation of SOS1 is promoted by binding of the protein to activated cell surface receptors
37
38 *via* SH3 domain recognition sites at the C-terminus of SOS1 that mediate its interaction with
39
40 adaptor proteins. Additional anchorage sites at the membrane, however, are also provided by the
41
42 PH and HF domains. Experimental and structural data indicate that two distinct sites within the
43
44 PH domains bind to PIP₂ and PA, and contribute significantly to SOS1's targeting to the
45
46 membrane [Chen et al., 1997; Zhao et al., 2007]. On the other hand, in the SOS1's autoinhibited
47
48 structure, the positively charged surface of the HF domain is not oriented in a way that would
49
50 allow membrane binding, suggesting the possibility that reorientation of the HF domain in the
51
52
53
54
55
56
57
58
59
60

1
2
3 membrane-bound conformation might destabilize HF's interaction with the REM domain,
4
5 unmasking the distal RAS binding site [Guerasko et al., 2010; Yadav and Bar-Sagi, 2010]. Based
6
7 on this model, the molecular spectrum of disease-causing mutations affecting the PH and HF
8
9 domains are predicted to favor the membrane-dependent electrostatic switch leading to a
10
11 productive reorientation of the protein in the plane of the membrane and consequent release of
12
13 auto-inhibition at the regulatory RAS binding site. In this scenario, class 2 mutations would
14
15 enhance the simultaneous engagement of the membrane by the PIP₂- and PA-binding pockets of
16
17 the PH domain and the positively charged surface of the HF domain, synergizing to increase the
18
19 stability of SOS1 translocation at the membrane and catalytic activation. Specifically, the
20
21 pathogenetic effect of this class of mutations affecting the PH domain would be related to a more
22
23 stable recruitment of the protein at the membrane, while the introduction of an arginine or lysine
24
25 residue within the positively charged surface of the HF domain would favor a spatial
26
27 reorientation of the domain weakening its inhibitory function. Consistent with this model,
28
29 Guerasko and co-workers [2010] demonstrated that the Glu108Lys substitution promotes
30
31 enhanced SOS1's GEF activity in a PIP₂-dependent manner, and elegantly provided evidence for
32
33 a role of non-specific electrostatic interactions of the HF domain with the membrane in release of
34
35 autoinhibition and activation of SOS1.
36
37
38
39
40
41
42

43 The present mutation scanning of a clinically well-characterized cohort of subjects with
44
45 features fitting or highly suggestive of NS, and negative for mutations in *PTPN11*, *RAF1* and
46
47 *KRAS*, confirmed our previous estimate of *SOS1* mutation prevalence, indicating that defects in
48
49 this gene account for approximately 10% of NS. Moreover, detailed clinical examination of 39
50
51 mutation-positive individuals, including data from clinical re-examination of 14 subjects
52
53 constituting our original cohort [Tartaglia et al., 2007], provided a more complete assessment
54
55
56
57
58
59
60

1
2
3 of the phenotypic variation associated with these molecular lesions, and strengthened the specific
4 association with ectodermal involvement (84%), and decreased prevalence of cognitive deficits
5 (11%) documented in previous reports [Roberts et al., 2007; Tartaglia et al., 2007; Zenker et al.,
6 2007a; Denayer et al., 2010]. A major finding also regarded the relatively lower prevalence of
7 short stature/length below the 3rd centile in these subjects (29%) compared to the overall NS
8 population, for which an estimate of 70% is generally reported. This percentage is lower than in
9 other studies, in which the prevalence ranged from 31% to 64% [Roberts et al., 2007; Zenker, et
10 al., 2007a; Denayer, et al., 2010]. Based on the different recruitment strategies utilized, this
11 discrepancy might be due, in part, to differences in patient selection (*e.g.*, recruitment performed
12 by pediatric endocrinology units vs. cardiology units). It should be noted, however, that this
13 disagreement might be also related to a sampling bias depending on the age distribution among
14 cohorts. In NS, while birth length is typically normal, growth parameters usually drop below the
15 3rd centile during the first years of life. Since there is often some attenuation and/or delay of the
16 pubertal growth spurt, the prevalence of short stature in NS is expected to be highest during the
17 age of normal puberty. Because of the delay in bone age, however, many patients have some
18 catch-up growth in their late teens. As we documented a delay in bone age in all subjects with
19 length/stature below the 3rd centile, a significant variation in the estimate of the growth status
20 between cohorts differing considerably in age distribution would be expected, since the estimate
21 of reduced growth prevalence in any NS cohort would be sensitive to the composition of the
22 cohort in terms of proportion of pediatric vs adult subjects.

23
24
25 We observed the occurrence of multiple tumors (*i.e.*, MGCLs, abdominal
26 rhabdomyosarcoma, cerebral glioma and granular cell tumors of the skin) in one of the *SOS1*
27 mutation-positive subjects included in the study. It has been established that NS patients are at

1
2
3 increased risk of developing childhood myeloproliferative disease (*i.e.*, juvenile myelomonocytic
4 leukemia) and acute leukemia [Tartaglia and Gelb, 2005], and that this specific association is
5 strictly linked to a specific class of activating germline mutations in the *PTPN11* gene [Tartaglia
6 et al., 2006]. Occurrence of neuroblastoma, glial tumors and rhabdomyosarcoma in subjects with
7 NS, however, has also been reported [Kratz, 2009]. Remarkably, while mutations in *SOS1* have
8 previously been considered to be more benign in terms of risk of malignancy than those in other
9 genes in the RAS/MAPK pathway (*e.g.*, *PTPN11*, *HRAS*, *KRAS* and *NF1*) [Swanson et al.,
10 2008], a significant high occurrence of solid tumors, particularly embryonal rhabdomyosarcoma,
11 has recently been reported in subjects with NS due to a mutated *SOS1* allele [Denayer et al.,
12 2010; Jongmans et al., 2010; Hastings et al., 2010]. The present finding is in line with these
13 recent reports, and provides further evidence that subjects heterozygous for *SOS1* mutations may
14 have a significant risk for certain solid tumors. Of note, MGCLs were observed in one additional
15 *SOS1* mutation-positive subject of the present cohort. MGCLs are benign tumor-like lesions
16 most frequently affecting the jaws but also occurring in other bones or soft tissues. They consist
17 of an osteoblast-like cell population representing the proliferating tumor cells producing
18 cytokines inducing the maturation of a subset of phagocytes into osteoclast-like giant cells [de
19 Lange et al., 2007]. While the incidence of MGCLs in patients with NS is not known, they were
20 originally linked to a narrow spectrum of germline *PTPN11* mutations [Tartaglia et al., 2002;
21 Jafarov et al., 2005; Lee et al., 2005]. In recent years, however, their co-occurrence in NS (and
22 other RAS-opathies) has been extended to other genes encoding transducers with roles in the
23 RAS-MAPK pathway, with *SOS1* being the most frequently mutated gene [Beneteau et al., 2009;
24 Hanna et al., 2009; Neumann et al., 2009], indicating that *SOS1* is a major predisposing gene for
25 MGCLs in NS.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 One patient in the present cohort, who was heterozygous for the *de novo* c.1297G>A
4 missense change (p.Gly434Lys), presented with gingival hypertrophy. Interestingly, a *SOS1*
5 frameshift mutation (c.3248-3249insC) predicting a truncated protein lacking the C-terminal
6 region encompassing the SH3 domain-binding and MAPK phosphorylation sites, was previously
7 documented to be responsible for a form of hereditary gingival fibromatosis (HGF1; MIM#
8 135300), a genetically heterogeneous benign gingival overgrowth condition, in a large family
9 [Hart et al., 2002]. Consistent with the negative modulatory role of the C-terminal region on
10 SOS1 function [Wang et al., 1995; Corbalan-Garcia et al., 1998], the SOS1 mutant was
11 demonstrated to localize at the plasma membrane constitutively and sustain the activation of
12 RAS-MAPK signaling, linking the gingival overgrowth to enhanced SOS1-mediated signal flow
13 through RAS [Jang et al., 2007]. While further studies are required to appreciate more precisely
14 the peculiar effect of the HGF1-associated mutation on SOS1 functional dysregulation and its
15 specific effect on gingival growth, the present finding provides further evidence in support of a
16 relation between RAS/MAPK signaling dysregulation and gingival hypertrophy.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

36 Excluding Down syndrome, NS is the most frequent developmental disorder associated
37 with congenital cardiac disease. Among *SOS1* mutation-positive individuals, cardiac defects
38 occurred in the majority of cases (> 85%), which is consistent with previous studies [Roberts, et
39 al., 2007; Tartaglia, et al., 2007; Zenker, et al., 2007a; Denayer, et al., 2010]. Similarly to what
40 observed in NS subjects with *PTPN11* mutations [Tartaglia et al., 2002; Zenker et al., 2004],
41 *SOS1* mutation-positive patients exhibit a high prevalence of PS (68% of cases), which was
42 typically associated with ASD (35% of cases), a relatively high occurrence of atrial and
43 ventricular septal defects (39% of cases), and a low prevalence of HCM (< 10% of cases). Based
44 on these findings and preliminary observations suggesting that a distinct class of mutations in
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 NS disease might be implicated in isolated CHDs [Pandit et al., 2007; Greenway, et al., 2009],
4
5 and given the relatively “mild” developmental and growth related phenotype documented in a
6
7 significant proportion of *SOS1* mutation-positive subjects, we evaluated the possible contribution
8
9 of germline *SOS1* gene lesions in apparently non-syndromic PS, ASD and VSD. Mutation
10
11 analysis revealed no *bona fide* mutation within these cohorts, which does not support the idea of
12
13 a major contribution of *SOS1* gene defects in the pathogenesis of these CHDs, in line with other
14
15 studies documenting only a minor contribution of germline *PTPN11* mutations in isolated CHDs
16
17 and HCM [Sarkozy, et al., 2003, 2005; Roberts, et al., 2005; Weismann, et al., 2005; Limongelli,
18
19 et al., 2006].
20
21
22
23

24
25 Subjects heterozygous for a mutated *SOS1* allele present with ectodermal manifestations
26
27 and distinctive facial dysmorphism that might be suggestive of CFCS in some individuals
28
29 (Figure 3) [Narumi et al. 2008; Nystrom et al. 2008]. In these subjects, however, cognitive
30
31 deficits are generally absent or minor [Tartaglia et al., 2007; Zenker et al., 2007a; Narumi et al.
32
33 2008; Denayer et al., 2010; present study]. Consistent with these observations, our mutational
34
35 screening on a clinically well-characterized CFCS cohort failed in identifying any *SOS1*
36
37 mutation, confirming a previous survey indicating that *SOS1* does not represent a major gene for
38
39 this disorder [Zenker et al. 2007a]. While the clinical features of *SOS1* mutation-positive subjects
40
41 appear to be less severe compared to what is generally observed in CFCS, the identification of
42
43 subjects with an “overlapping” phenotype suggests that a clinical continuum might be associated
44
45 with defects in *SOS1*, as previously documented for other disease genes implicated in RAS-
46
47 opathies [Zenker et al., 2007b; Sarkozy et al., 2009a]. On the other hand, based on the increasing
48
49 evidence documenting co-occurrence of mutations in functionally related genes that contribute to
50
51 the severity of the phenotype [Nystrom et al., 2009; Tang et al., 2009; Thiel et al., 2009;
52
53
54
55
56
57
58
59
60

1
2
3 Longoni et al., 2010], the possibility that mutations concomitant to those affecting *SOS1* might
4
5 modulate the phenotype in these subjects cannot be ruled out. Overall, these findings further
6
7 emphasize the difficulty in identifying efficient clinical criteria to define these disorders
8
9 nosologically, and make evident the clinical relevance of a molecular-based definition of these
10
11 clinically overlapping disorders.
12
13
14
15
16

17 **ACKNOWLEDGMENTS**

18
19 We are indebted to the patients and families who participated in the study, and the physicians
20
21 who referred the subjects. This research was funded by grants from Telethon-Italy (GGP07115
22
23 and GGP10020), ERA-Net for research programmes on rare diseases 2009 (“European network
24
25 on Noonan Syndrome and related disorders”) and *Associazione Italiana Sindromi di Costello e*
26
27 *Cardiofaciocardutanea* to M.T., NIH (HL71207) to B.D.G., Italian Ministry of Health (RC2009 and
28
29 RC2010) to F.L., and Italian Ministry of Education, University and Research (FIRB
30
31 RBIP06PMF2_005) to A.D.L.
32
33
34
35
36
37
38

39 **REFERENCES**

40
41 Allanson JE. 1987. Noonan syndrome. *J Med Genet* 24:9-13.
42
43
44 Aoki Y, Niihori T, Narumi Y, Kure S, Matsubara Y. 2008. The RAS/MAPK syndromes: novel
45
46 roles of the RAS pathway in human genetic disorders. *Hum Mutat* 29:992-1006.
47
48
49 Aronheim A, Engelberg D, Li N, al-Alawi N, Schlessinger J, Karin M. 1994. Membrane
50
51 targeting of the nucleotide exchange factor Sos is sufficient for activating the Ras signaling
52
53 pathway. *Cell* 78:949-961.
54
55
56
57
58
59
60

1
2
3 Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems:
4 application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98:10037-10041.
5
6

7
8 Beneteau C, Cave H, Moncla A, Dorison N, Munnich A, Verloes A, Leheup B. 2009. SOS1 and
9 PTPN11 mutations in five cases of Noonan syndrome with multiple giant cell lesions. *Eur J Hum*
10 *Genet* 17:1216-1221.
11
12
13

14
15 Burch M, Sharland M, Shinebourne E, Smith G, Patton M, McKenna W. 1993. Cardiologic
16 abnormalities in Noonan syndrome: phenotypic diagnosis and echocardiographic assessment of
17 118 patients. *J Am Coll Cardiol* 22:1189-1192.
18
19
20
21

22
23 Chen RH, Corbalan-Garcia S, Bar-Sagi D. 1997. The role of the PH domain in the signal-
24 dependent membrane targeting of Sos. *EMBO J* 16:1351-1359.
25
26
27

28
29 Chen PC, Wakimoto H, Conner D, Araki T, Yuan T, Roberts A, Seidman CE, Bronson R, Neel
30 BG, Seidman JG, Kucherlapati R. 2010. Activation of multiple signaling pathways causes
31 developmental defects in mice with a Noonan syndrome-associated *Sos1* mutation. *J Clin Invest*,
32 in press (PMID: 21041952).
33
34
35
36
37

38
39 Corbalan-Garcia S, Yang SS, Degenhardt KR, Bar-Sagi D. 1996. Identification of the mitogen-
40 activated protein kinase phosphorylation sites on human *Sos1* that regulate interaction with
41 *Grb2*. *Mol Cell Biol* 16:5674-5682.
42
43
44
45

46
47 Corbalan-Garcia S, Margarit SM, Galron D, Yang SS, Bar-Sagi D. 1998. Regulation of *Sos*
48 activity by intramolecular interactions. *Mol Cell Biol* 18:880-886.
49
50

51
52 de Lange J, van den Akker HP, van den Berg H. 2007. Central giant cell granuloma of the jaw: a
53 review of the literature with emphasis on therapy options. *Oral Surg Oral Med Oral Pathol Oral*
54
55
56

1
2
3 Radiol Endod 104:603-615.
4

5
6 Denayer E, Devriendt K, de Ravel T, Van Buggenhout G, Smeets E, Francois I, Sznajer Y,
7
8 Craen M, Leventopoulos G, Mutesa L, Vandecasseye W, Massa G, Kayserili H, Sciot R, Fryns
9
10 JP, Legius E. 2010. Tumor spectrum in children with Noonan syndrome and SOS1 or RAF1
11
12 mutations. *Genes Chromosomes Cancer* 49:242-252.
13
14

15
16 Ferrero GB, Baldassarre G, Delmonaco AG, Biamino E, Carta C, Rossi C, Silengo MC. 2008.
17
18 Clinical and molecular characterization of 40 patients with Noonan syndrome. *Eur J Med Genet*
19
20 51:566-72.
21
22

23
24 Freedman TS, Sondermann H, Friedland GD, Kortemme T, Bar-Sagi D, Marqusee S, Kuriyan J.
25
26 2006. A Ras-induced conformational switch in the Ras activator Son of sevenless. *Proc Natl*
27
28 *Acad Sci USA* 103:16692-16697.
29
30

31
32 Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, Mesquita SM, Ergul E, Conta JH,
33
34 Korn JM, McCarroll SA, Gorham JM, Gabriel S, Altshuler DM, Quintanilla-Dieck Mde L,
35
36 Artunduaga MA, Eavey RD, Plenge RM, Shadick NA, Weinblatt ME, De Jager PL, Hafler DA,
37
38 Breitbart RE, Seidman JG, Seidman CE. 2009. De novo copy number variants identify new
39
40 genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet* 41:931-935.
41
42

43
44 Guerasko J, Galush WJ, Boykevisch S, Sondermann H, Bar-Sagi D, Groves JT, Kuriyan J. 2008.
45
46 Membrane-dependent signal integration by the Ras activator Son of sevenless. *Nat Struct Mol*
47
48 *Biol* 15:452-461.
49

50
51 Guerasko J, Kuchment O, Makino DL, Sondermann H, Bar-Sagi D, Kuriyan J. 2010. Role of the
52
53 histone domain in the autoinhibition and activation of the Ras activator Son of Sevenless. *Proc*
54
55

1
2
3 Natl Acad Sci USA 107:3430-3435.
4

5
6 Hall BE, Yang SS, Boriack-Sjodin PA, Kuriyan J, Bar-Sagi D. 2001. Structure-based
7
8 mutagenesis reveals distinct functions for Ras switch 1 and switch 2 in Sos-catalyzed guanine
9
10 nucleotide exchange. *J Biol Chem* 276:27629-27637.
11

12
13
14 Hanna N, Parfait B, Talaat IM, Vidaud M, Elsedfy HH. 2009. SOS1: a new player in the
15
16 Noonan-like/multiple giant cell lesion syndrome. *Clin Genet* 75:568-571.
17

18
19 Hart TC, Zhang Y, Gorry MC, Hart PS, Cooper M, Marazita ML, Marks JM, Cortelli JR, Pallos
20
21 D. 2002. A mutation in the SOS1 gene causes hereditary gingival fibromatosis type 1. *Am J*
22
23 *Hum Genet* 70:943-954.
24

25
26
27 Hastings R, Newbury-Ecob R, Ng A, Taylor R. 2010. A further patient with Noonan syndrome
28
29 due to a SOS1 mutation and rhabdomyosarcoma. *Genes Chromosomes Cancer* 49:967-968.
30

31
32 Jafarov T, Ferimazova N, Reichenberger E. 2005. Noonan-like syndrome mutations in PTPN11
33
34 in patients diagnosed with cherubism. *Clin Genet* 68:190-191.
35

36
37 Jang SI, Lee EJ, Hart PS, Ramaswami M, Pallos D, Hart TC. 2007. Germ line gain of function
38
39 with SOS1 mutation in hereditary gingival fibromatosis. *J Biol Chem* 282:20245-20255.
40

41
42 Jongmans MC, Hoogerbrugge PM, Hilkens L, Flucke U, van der Burgt I, Noordam K,
43
44 Ruitenkamp-Versteeg M, Yntema HG, Nillesen WM, Ligtenberg MJ, van Kessel AG, Kuiper RP,
45
46 Hoogerbrugge N. 2010. Noonan syndrome, the SOS1 gene and embryonal rhabdomyosarcoma.
47
48 *Genes Chromosomes Cancer* 49:635-641.
49

50
51
52 Ko JM, Kim JM, Kim GH, Yoo HW. 2008. PTPN11, SOS1, KRAS, and RAF1 gene analysis,
53
54 and genotype-phenotype correlation in Korean patients with Noonan syndrome. *J Hum Genet*
55
56

1
2
3 53:999-1006.
4

5
6 Kratz C. 2009. Myeloproliferative disease and cancer in persons with Noonan syndrome and
7 related disorders. In: Zenker M, editor. Monogr Hum Genet. Noonan Syndrome and Related
8 Disorders. 17:119–127, Karger, Basel.
9

10
11 Lee JS, Tartaglia M, Gelb BD, Fridrich K, Sachs S, Stratakis CA, Muenke M, Robey PG, Collins
12 MT, Slavotinek A. 2005. Phenotypic and genotypic characterisation of Noonan-like/multiple
13 giant cell lesion syndrome. *J Med Genet* 42:e11.
14
15

16
17 Limongelli G, Hawkes L, Calabro R, McKenna WJ, Syrris P. 2006. Mutation screening of the
18 PTPN11 gene in hypertrophic cardiomyopathy. *Eur J Med Genet* 49:426-430.
19
20

21
22 Longoni M, Moncini S, Cisternino M, Morella IM, Ferraiuolo S, Russo S, Mannarino S,
23 Brazzelli V, Coi P, Zippel R, Venturin M, Riva P. 2010. Noonan syndrome associated with both
24 a new Jnk-activating familial SOS1 and a de novo RAF1 mutations. *Am J Med Genet A*
25 152A:2176-2184.
26
27

28
29 Margarit SM, Sondermann H, Hall BE, Nagar B, Hoelz A, Pirruccello M, Bar-Sagi D, Kuriyan J.
30 2003. Structural evidence for feedback activation by Ras.GTP of the Ras-specific nucleotide
31 exchange factor SOS. *Cell* 112:685-695.
32
33

34
35 Marino B, Digilio MC, Toscano A, Giannotti A, Dallapiccola B. 1999. Congenital heart diseases
36 in children with Noonan syndrome: An expanded cardiac spectrum with high prevalence of
37 atrioventricular canal. *J Pediatr* 135:703-706.
38
39

40
41 Narumi Y, Aoki Y, Niihori T, Sakurai M, Cave H, Verloes A, Nishio K, Ohashi H, Kurosawa K,
42 Okamoto N, Kawame H, Mizuno S, Kondoh T, Addor MC, Coeslier-Dieux A, Vincent-Delorme
43
44
45
46
47
48
49
50
51

1
2
3 C, Tabayashi K, Aoki M, Kobayashi T, Guliyeva A, Kure S, Matsubara Y. 2008. Clinical
4 manifestations in patients with SOS1 mutations range from Noonan syndrome to CFC syndrome.
5
6 J Hum Genet 53:834-841.
7
8

9
10 Neumann TE, Allanson J, Kavamura I, Kerr B, Neri G, Noonan J, Cordeddu V, Gibson K,
11
12 Tzschach A, Kruger G, Hoeltzenbein M, Goecke TO, Kehl HG, Albrecht B, Luczak K, Sasiadek
13
14 MM, Musante L, Laurie R, Peters H, Tartaglia M, Zenker M, Kalscheuer V. 2009. Multiple giant
15
16 cell lesions in patients with Noonan syndrome and cardio-facio-cutaneous syndrome. Eur J Hum
17
18 Genet 17:420-425.
19
20

21
22
23 Newcombe RG. 1998. Two-sided confidence intervals for the single proportion: comparison of
24
25 seven methods. Stat Med 17:857-872.
26
27

28
29 Ng PC, Henikoff S. 2001. Predicting deleterious amino acid substitutions. Genome Res 11:863-
30
31 874.
32

33
34 Nimmual A, Bar-Sagi D. 2002. The two hats of SOS. Sci STKE 145:pe36.
35

36
37 Noonan JA. 1994. Noonan syndrome. An update and review for the primary pediatrician. Clin
38
39 Pediatr (Phila) 33:548-555.
40

41
42 Nystrom AM, Ekvall S, Berglund E, Bjorkqvist M, Braathen G, Duchon K, Enell H, Holmberg
43
44 E, Holmlund U, Olsson-Engman M, Annerén G, Bondeson ML. 2008. Noonan and cardio-facio-
45
46 cutaneous syndromes: two clinically and genetically overlapping disorders. J Med Genet 45:500-
47
48 506.
49

50
51
52 Nystrom AM, Ekvall S, Strömberg B, Holmström G, Thuresson AC, Annerén G, Bondeson ML.
53
54 2009. A severe form of Noonan syndrome and autosomal dominant café-au-lait spots - evidence
55
56

1
2
3 for different genetic origins. *Acta Paediatr* 98:693-698.

4
5
6 Pandit B, Sarkozy A, Pennacchio LA, Carta C, Oishi K, Martinelli S, Pogna EA, Schackwitz W,
7
8 Ustaszewska A, Landstrom A, Bos JM, Ommen SR, Esposito G, Lepri F, Faul C, Mundel P,
9
10 López Sigüero JP, Tenconi R, Selicorni A, Rossi C, Mazzanti L, Torrente I, Marino B, Digilio
11
12 MC, Zampino G, Ackerman MJ, Dallapiccola B, Tartaglia M, Gelb BD. 2007. Gain-of-function
13
14 RAF1 mutations cause Noonan and LEOPARD syndromes with hypertrophic cardiomyopathy.
15
16 *Nat Genet* 39:1007-1012.

17
18
19
20
21 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004.
22
23 UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*
24
25 25:1605-1612.

26
27
28 Roberts AE, Hult B, Rehm HL, Rehm HL, McDonough B, Barr S, Seidman CE, Seidman JG,
29
30 Kucherlapati RS. 2005. The PTPN11 gene is not implicated in nonsyndromic hypertrophic
31
32 cardiomyopathy. *Am J Med Genet A* 132A:333-334.

33
34
35
36 Roberts AE, Araki T, Swanson KD, Montgomery KT, Schiripo TA, Joshi VA, Li L, Yassin Y,
37
38 Tamburino AM, Neel BG, Kucherlapati RS. 2007. Germline gain-of-function mutations in SOS1
39
40 cause Noonan syndrome. *Nat Genet* 39:70-74.

41
42
43 Sarkozy A, Conti E, Esposito G, Pizzuti A, Dallapiccola B, Mingarelli R, Marino B, Digilio MC,
44
45 Paoletti V. 2003. Nonsyndromic pulmonary valve stenosis and the PTPN11 gene. *Am J Med*
46
47 *Genet A* 116A:389-390.

48
49
50
51 Sarkozy A, Conti E, Lepri FR, Pizzuti A, Dallapiccola B, Autore C, Tartaglia M. 2005.
52
53 Hypertrophic cardiomyopathy and the PTPN11 gene. *Am J Med Genet A* 136:93-94.

1
2
3 Sarkozy A, Carta C, Moretti S, Zampino G, Digilio MC, Pantaleoni F, Scioletti AP, Esposito G,
4
5 Cordeddu V, Lepri F, Petrangeli V, Dentici ML, Mancini GM, Selicorni A, Rossi C, Mazzanti L,
6
7 Marino B, Ferrero GB, Silengo MC, Memo L, Stanzial F, Faravelli F, Stuppia L, Puxeddu E,
8
9 Gelb BD, Dallapiccola B, Tartaglia M. 2009a. Germline BRAF mutations in Noonan,
10
11 LEOPARD, and cardiofaciocutaneous syndromes: molecular diversity and associated phenotypic
12
13 spectrum. *Hum Mutat* 30:695-702.
14
15

16
17
18 Sarkozy A, Digilio MC, Marino B, Dallapiccola B. 2009b. Genotype-Phenotype Correlations in
19
20 Noonan Syndrome. In: Zenker M, editor. *Monogr Hum Genet. Noonan Syndrome and Related*
21
22 *Disorders*. 17:40–54, Karger, Basel.
23
24

25
26 Sharland M, Burch M, McKenna WM, Paton MA. 1992. A clinical study of Noonan syndrome.
27
28 *Arch Dis Child* 67:178-83.
29

30
31 Sondermann H, Soisson SM, Boykevisch S, Yang SS, Bar-Sagi D, Kuriyan J. 2004. Structural
32
33 analysis of autoinhibition in the Ras activator Son of sevenless. *Cell* 119:393-405.
34
35

36
37 Sondermann H, Nagar B, Bar-Sagi D, Kuriyan J. 2005. Computational docking and solution x-
38
39 ray scattering predict a membrane-interacting role for the histone domain of the Ras activator son
40
41 of sevenless. *Proc Natl Acad Sci USA* 102:16632-16637.
42
43

44
45 Sunyaev S, Ramensky V, Bork P. 2000. Towards a structural basis of human non-synonymous
46
47 single nucleotide polymorphisms. *Trends Genet* 16:198-200.
48

49
50 Swanson KD, Winter JM, Reis M, Bentires-Alj M, Greulich H, Grewal R, Hruban RH, Yeo CJ,
51
52 Yassin Y, Iartchouk O, Montgomery K, Whitman SP, Caligiuri MA, Loh ML, Gilliland DG,
53
54 Look AT, Kucherlapati R, Kern SE, Meyerson M, Neel BG. 2008. SOS1 mutations are rare in
55
56
57
58
59
60

1
2
3 human malignancies: implications for Noonan Syndrome patients. *Genes Chromosomes Cancer*
4
5 47:253-259.
6
7

8
9 Tang S, Hoshida H, Kamisago M, Yagi H, Momma K, Matsuoka R. 2009. Phenotype-genotype
10
11 correlation in a patient with co-occurrence of Marfan and LEOPARD syndromes. *Am J Med*
12
13 *Genet A* 149A:2216-2219.
14
15

16
17 Tartaglia M, Gelb BD. 2005. Noonan syndrome and related disorders: genetics and pathogenesis.
18
19 *Annu Rev Genomics Hum Genet* 6:45-68.
20

21
22 Tartaglia M, Gelb BD. 2010. Disorders of dysregulated signal traffic through the RAS-MAPK
23
24 pathway: phenotypic spectrum and molecular mechanisms. *Ann NY Acad Sci*, in press (PMID:
25
26 20958325).
27

28
29 Tartaglia M, Zampino G, Gelb BD. 2010. Noonan syndrome: clinical aspects and molecular
30
31 pathogenesis. *Mol Syndromol* 1:2-26.
32
33

34
35 Tartaglia M, Mehler EL, Goldberg R, Zampino G, Brunner HG, Kremer H, van der Burgt I,
36
37 Crosby AH, Ion A, Jeffery S, Kalidas K, Patton MA, Kucherlapati RS, Gelb BD. 2001.
38
39 Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan
40
41 syndrome. *Nat Genet* 29:465-468.
42
43

44
45 Tartaglia M, Kalidas K, Shaw A, Song X, Musat DL, van der Burgt I, Brunner HG, Bertola DR,
46
47 Crosby A, Ion A, Kucherlapati RS, Jeffery S, Patton MA, Gelb BD. 2002. PTPN11 mutations in
48
49 Noonan syndrome: molecular spectrum, genotype-phenotype correlation, and phenotypic
50
51 heterogeneity. *Am J Hum Genet* 70:1555-1563.
52
53

54
55 Tartaglia M, Martinelli S, Stella L, Bocchinfuso G, Flex E, Cordeddu V, Zampino G, Burgt I,
56
57
58
59
60

1
2
3 Palleschi A, Petrucci TC, Sorcini M, Schoch C, Foa R, Emanuel PD, Gelb BD. 2006. Diversity
4 and functional consequences of germline and somatic PTPN11 mutations in human disease. *Am*
5
6
7
8 *J Hum Genet* 78:279-290.

9
10 Tartaglia M, Pennacchio LA, Zhao C, Yadav KK, Fodale V, Sarkozy A, Pandit B, Oishi K,
11
12 Martinelli S, Schackwitz W, Martin J, Bristow J, Carta C, Lepri F, Neri C, Vasta I, Gibson K,
13
14 Curry CJ, Siguero JP, Digilio MC, Zampino G, Dallapiccola B, Bar-Sagi D, Gelb BD. 2007.
15
16 Gain-of-function SOS1 mutations cause a distinctive form of Noonan syndrome. *Nat Genet*
17
18
19
20 39:75-79.

21
22
23 Thiel C, Wilken M, Zenker M, Sticht H, Fahsold R, Gusek-Schneider GC, Rauch A. 2009.
24
25 Independent NF1 and PTPN11 mutations in a family with neurofibromatosis-Noonan syndrome.
26
27
28
29 *Am J Med Genet A* 149A:1263-1267.

30
31 Tidyman WE, Rauen KA. 2009. The RASopathies: developmental syndromes of Ras/MAPK
32
33
34
35 pathway dysregulation. *Curr Opin Genet Dev* 19:230-236.

36
37 van der Burgt I. 2007. Noonan syndrome. *Orphanet J Rare Dis* 2:4.

38
39 van der Burgt I, Berends E, Lommen E, van Beersum S, Hamel B, Mariman E. 1994. Clinical
40
41 and molecular studies in a large Dutch family with Noonan syndrome. *Am J Med Genet* 53:187-
42
43
44 191.

45
46 Wang W, Fisher EM, Jia Q, Dunn JM, Porfiri E, Downward J, Egan SE. 1995. The Grb2 binding
47
48 domain of mSos1 is not required for downstream signal transduction. *Nat Genet* 10:294-300.

49
50
51 Weismann CG, Hager A, Kaemmerer H, Maslen CL, Morris CD, Schranz D, Kreuder J, Gelb
52
53
54
55
56
57
58
59
60 BD. 2005. PTPN11 mutations play a minor role in isolated congenital heart disease. *Am J Med*

1
2
3 Genet A 136:146-151.
4

5
6 Wilson EB. 1927. Probable inference, the law of succession, and statistical inference. J Am Stat
7
8 Assoc 22:209-212.
9

10
11 Yadav KK, Bar-Sagi D. 2010. Allosteric gating of Son of sevenless activity by the histone
12
13 domain. Proc Natl Acad Sci USA 107:3436-3440.
14

15
16 Zenker M, Buheitel G, Rauch R, Koenig R, Bosse K, Kress W, Tietze HU, Doerr HG, Hofbeck
17
18 M, Singer H, Reis A, Rauch A. 2004. Genotype-phenotype correlations in Noonan syndrome. J
19
20 Pediatr 144:368-374.
21
22

23
24 Zenker M, Horn D, Wiczorek D, Allanson J, Pauli S, van der Burgt I, Doerr HG, Gaspar H,
25
26 Hofbeck M, Gillessen-Kaesbach G, Koch A, Meinecke P, Mundlos S, Nowka A, Rauch A, Reif
27
28 S, von Schnakenburg C, Seidel H, Wehner LE, Zweier C, Bauhuber S, Matejas V, Kratz CP,
29
30 Thomas C, Kutsche K. 2007a. SOS1 is the second most common Noonan gene but plays no
31
32 major role in cardio-facio-cutaneous syndrome. J Med Genet 44:651-656.
33
34

35
36 Zenker M, Lehmann K, Schulz AL, Barth H, Hansmann D, Koenig R, Korinthenberg R, Kreiss-
37
38 Nachtsheim M, Meinecke P, Morlot S, Mundlos S, Quante AS, Raskin S, Schnabel D, Wehner
39
40 LE, Kratz CP, Horn D, Kutsche K. 2007b. Expansion of the genotypic and phenotypic spectrum
41
42 in patients with KRAS germline mutations. J Med Genet 44:131-135.
43
44

45
46 Zhao C, Du G, Skowronek K, Frohman MA, Bar-Sagi D. 2007. Phospholipase D2-generated
47
48 phosphatidic acid couples EGFR stimulation to Ras activation by Sos. Nat Cell Biol 9:706-712.
49
50

51
52 Zheng J, Chen RH, Corblan-Garcia S, Cahill SM, Bar-Sagi D, Cowburn D. 1997. The solution
53
54 structure of the pleckstrin homology domain of human SOS1. A possible structural role for the
55
56

1
2
3 sequential association of diffuse B cell lymphoma and pleckstrin homology domains. J Biol
4
5 Chem 272:30340-30344.
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

FIGURE LEGENDS

FIGURE 1. SOS1 domain structure and location of residues altered in Noonan syndrome. (A) Schematic structure of SOS1 and variants identified in the present study. SOS1 protein domains are indicated (DH, DBL homology domain; PH, pleckstrin homology domain; REM, RAS exchanger motif; CDC25, CDC25 domain). Disease-causing mutations and probably pathogenetic/unclassified variants are shown above and below the cartoon, respectively. Residues affected by class 1 mutations/variants are shown in red, while those affected by class 2 and class 3 changes are shown in yellow and green, respectively. Residues affected by substitutions with unpredictable effect on SOS1 function are shown in black. Novel amino acid substitutions are underlined. (B) Location of affected residues in SOS1 represented in its inactive conformation, according to the crystal structure of the protein truncated at the C-terminus (residues 1-1049) (PDB ID: 3KSY) [Guerasko et al., 2010]. C α ribbon trace of the HF (sky blue), DH (sandy brown), PH (plum), REM (dark green), and CDC25 (blue) domains, and the helical linker connecting the PH and REM domains (gray). Mutated residues are indicated with their side chains as thick lines and colored as reported above. Residue Asp309 (uncharacterized mutation p.Asp309Tyr) is shown in purple. Affected residues are listed in Supp. Table S2.

FIGURE 2. Detailed analysis of structural perturbations resulting from Noonan syndrome-causing amino acid substitutions. (A) Class 1 mutations affecting residues at the distal RAS binding site. The cartoon includes the DH (sandy brown) and REM (dark green) domains only.

1
2
3 Affected residues are shown in red. The autoinhibitory binding network includes the
4 hydrophobic interaction between residues Met269 (DH) and Trp729 (REM), both mutated in NS,
5 and between the former and Leu687 (green). Leu690, Val697, Ile718 and Ile736 are hydrophobic
6 residues (cyan) that interact with Ile733 (REM). Mutated residues Thr266 (DH) and Lys728
7 (REM) face each other. Substitution of Thr266 by lysine would create an electrostatic repulsion
8 with Lys728. The unclassified variant Leu252 contributes to a hydrophobic core with residues
9 Tyr215, Leu219, Ile249, Tyr295 and Tyr298 (purple), whose disruption is expected to perturb
10 considerably the DH domain surface interacting with the REM domain. (B) Class 1B mutations.
11 The cartoon includes the HF (sky blue), DH (sandy brown) and PH (magenta) domains and the
12 PH-REM helical linker (grey). Relevant affected residues are shown in red. Met422 and Ile437
13 (PH) participate in a hydrophobic bonding network with residues Ile425, Phe464 and Leu467
14 (cyan, see also the inset). Hydrophobic interaction between Tyr337 and Met538 (green)
15 contributes to the binding network stabilizing the interaction between the PH and DH domains.
16 Other interdomain interactions involve Leu550 and residues of the DH and PH domains, Val225,
17 Leu221, Phe226 and Tyr546 (grey), Arg552 and residues of the HF domain, Asp140 and Asp169
18 (orange), Ser548 and Asp169, and Lys170 and residues of the PH domain, Arg497 and Lys498
19 (blue). Phe78 participates to the hydrophobic interaction involving residues of the HF domain
20 core located close to the PH domain and PH-REM linker (Leu55, Leu59, Val74, Val133 and
21 Ile137; purple). (C) Class 1 mutations affecting the REM domain region (dark green) interacting
22 with the CDC25 domain (blue). The helical hairpin (residues 929-978) implicated in the
23 conformational switch (green) and residues interacting with RAS at the active site (light blue) are
24 shown. Phe623 hydrophobically interact with Ile 601, Leu613, Phe627, Ile956 and Phe958
25 (cyan). The hydrogen bond between Tyr702 and Ser802 (orange) is also shown. (D) Class 2
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 mutations affecting the HF domain. The left panel shows the HF (light blue), DH (sandy brown)
4 and PH (plum) domains. The HF surface colored according to the electrostatic potential (from
5 and PH (plum) domains. The HF surface colored according to the electrostatic potential (from
6 red at $-3kT/e$ to blue at $+3kT/e$) is also shown (right panel). Mutations affect solvent exposed
7 residues (yellow side chains, left panel; yellow circles, right panel) located in a region that has a
8 positive electrostatic potential (right panel), and has been implicated in membrane binding. (E)
9 Class 2 mutations affecting the PH domain. The left panel includes the PH domain (plum) only.
10 The PH surface, colored according to the electrostatic potential (from red at $-5kT/e$ to blue at
11 $+5kT/e$) is also shown (left panel). Affected residues are shown in yellow (side chains, left panel;
12 circles, right panel). Residues that are predicted to bind to PIP_2 [Zheng et al., 1997], Lys456,
13 Arg459, Lys472 and Arg489, are shown in green (circled in the right panel), while the PA-
14 interacting region (residues 472-483) [Zhao et al., 2007] is shown in orange (circled in the right
15 panel). (F) Class 3 mutations. The cartoon includes the REM (dark green) and CDC25 (blue)
16 domains only (residues 567-1049). Affected residues are shown in green. Residues implicated in
17 RAS binding at the catalytic site are shown (light blue). Glu846 and Pro894 are placed distally
18 from the active site and regions implicated in the conformational rearrangement of the CDC25
19 domain. Glu846 electrostatically interacts with Arg1026 and Lys 1029 (orange).

20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46 **FIGURE 3.** Facial dysmorphism and other features of subjects with Noonan syndrome
47 heterozygous for mutations in the *SOS1* gene. *SOS1* mutation-positive subjects generally exhibit
48 typical facial features, including macrocephaly, hypertelorism, ptosis, downslanting palpebral
49 fissures, sparse eyebrows with keratosis pilaris, a short and broad nose with upturned tip, low-
50 set and posteriorly angulated ears, and high forehead commonly associated with bitemporal
51
52
53
54
55
56
57
58
59
60

1
2
3 narrowing and prominent supraorbital ridges. Curly hair is present in most of the patients. Other
4
5 common features include pectus anomalies (NS10, NS19, NS37), short and/or webbed neck
6
7 (NS6, NS10, NS19, NS22, NS38), and cubitus valgus (NS37). Keloid scars (NS16), recurrent
8
9 hemorrhages (NS18) and deep plantar creases (NS38) also occur in these subjects. In some
10
11 infants, the face is suggestive of cardiofaciocutaneous syndrome due to the coarseness of features
12
13 (NS39).
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

Table 1. *SOS1* exonic indels and missense changes identified in the study

Exon	Nucleotide change	Amino acid change	Domain	Notes	Number of cases
<i>Mutations</i>					
4	c.322G>A	p.Glu108Lys	HF		2, fam.unknown
5	c.508A>G	p.Lys170Glu	HF		1, sporadic; 2, fam.unknown
7	c.797C>A	p.Thr266Lys	DH		1, sporadic; 3, fam.unknown
7	c.806T>C	p.Met269Thr	DH		3, sporadic; 2, fam.unknown
7	c.806T>G	p.Met269Arg	DH		1, fam.unknown
11	c.1281_1289delGAATATTGA	p.Lys427_Asp430delinsAsn	PH		1, fam.unknown
11	c.1294T>C	p.Trp432Arg	PH		1, sporadic; 1, fam.unknown
11	c.1294_1299delTGGGAG	p.Trp432_Glu433del	PH	CTRL, NPS	1, sporadic
11	c.1297G>A	p.Glu433Lys	PH		3, sporadic; 3, fam.unknown
11	c.1300G>A	p.Gly434Arg	PH		2, fam.unknown
11	c.1300_1301delGGinsAA	p.Gly434Lys	PH	<i>de novo</i> , PPhen+, SIFT+, CON	1, sporadic
11	c.1310T>C	p.Ile437Thr	PH	<i>de novo</i> , PPhen++, SIFT+, CON	1, sporadic; 1, familial
11	c.1322G>A	p.Cys441Tyr	PH		1, fam.unknown
11	c.1430A>G	p.Gln477Arg	PH	PPhen-, SIFT-, CON	2, fam.unknown
11	c.1433C>G	p.Pro478Arg	PH		1, sporadic
11	c.1642A>C	p.Ser548Arg	PH-REM linker		4, fam.unknown
11	c.1654A>G	p.Arg552Gly	PH-REM linker		3, sporadic; 8, fam.unknown
11	c.1655G>A	p.Arg552Lys	PH-REM linker		2, sporadic; 2, fam.unknown
11	c.1655G>T	p.Arg552Met	PH-REM linker	NPS, PPhen++, SIFT+, CON	2, sporadic
11	c.1655G>C	p.Arg552Thr	PH-REM linker		1, fam.unknown
11	c.1656G>C	p.Arg552Ser	PH-REM linker		2, sporadic; 7, fam.unknown
11	c.1660_1673delCTTGATGTAACAATinsAA	p.Leu554_Met558delinsLys	PH-REM linker		1, fam.unknown
15	c.2197A>T	p.Ile733Phe	REM		1, fam.unknown
17	c.2536G>A	p.Glu846Lys	CDC25		1, sporadic, 6 fam.unknown
18	c.2681C>G	p.Pro894Arg	CDC25	CTRL, PPhen+, SIFT-, CON	1, familial [§]
<i>Possibly pathogenic variants</i>					
4	c.335C>G	p.Pro112Arg	HF	NPS, PPhen+, SIFT-	1, sporadic
7	c.755T>C	p.Ile252Thr	DH	CTRL, PPhen++, SIFT+, CON	1, familial
11	c.1264A>G	p.Met422Val	PH	PPhen++, SIFT-, CON	1, fam.unknown

Table 1. *SOS1* exonic indels and missense changes identified in the study (continued)

Exon	Nucleotide change	Amino acid change	Domain	Notes	Number of cases
11	c.1270G>A	p.Glu424Lys	PH	PPhen+, SIFT+, CON	1, fam.unknown
11	c.1444G>C	p.Gly482Arg	PH	PPhen+, SIFT+, CON	1, fam.unknown
11	c.1469T>G	p.Leu490Arg	PH	CTRL, NPS, PPhen++, SIFT+, CON	1, sporadic
11	c.1490G>A	p.Arg497Gln	PH	CTRL, PPhen+, SIFT+, CON	1, sporadic*
11	c.1646C>A	p.Thr549Lys	PH-REM linker	PPhen+, SIFT-, CON	1, fam.unknown
17 <i>Unclassified</i>					
18 <i>variants</i>					
3	c.109A>G	p.Thr37Ala	HF	PPhen-, SIFT-, CON	1, fam.unknown
11	c.1433C>T	p.Pro478Leu	PH	PPhen+, SIFT-, CON	1, fam.unknown
15	c.2351T>C	p.Ile784Thr	REM	PPhen++, SIFT+, CON	1, fam.unknown ^{&}
23	c.3392G>A	p.Arg1131Lys	C-terminus	PPhen-, SIFT-, CON	1, fam.unknown
23	c.3418T>A	p.Leu1140Ile	C-terminus	PPhen-, SIFT-, CON	1, fam.unknown
24	c.3769A>G	p.Thr1257Ala	C-terminus	PPhen-, SIFT-, CON	1, fam.unknown

Nucleotide numbering reflects cDNA numbering with 1 corresponding to the A of the ATG translation initiation codon in the reference sequence (NM_005633.3). Exon 2 corresponds to the first protein coding exon. Novel mutations are in bold. HF, histone folds; DH, DBL homology domain; PH, plekstrin homology domain; REM, RAS exchanger motif, CDC25, CDC25 domain. Fam.unknown, familial status unknown; NPS, unavailable parental DNA samples. CTRL, variant not occurring in ≥ 300 population-matched unaffected subjects; *de novo*, variant demonstrated to occur *de novo* by DNA genotyping of unaffected parents; PPhen++, amino acid change predicted to be “probably damaging” by PolyPhen; PPhen+, amino acid change predicted to be “possibly damaging” by PolyPhen; PPhen-, amino acid change predicted to be “benign” by Polyphen; SIFT+, amino acid change predicted to “affect protein function” by SIFT; SIFT-, amino acid change predicted to be “tolerated” by SIFT; CON, variant at a conserved residue. §, variant inherited from an affected parent; *, variant inherited from an apparently unaffected parent; &, variant concomitant with the disease-causing p.Met269Arg change.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

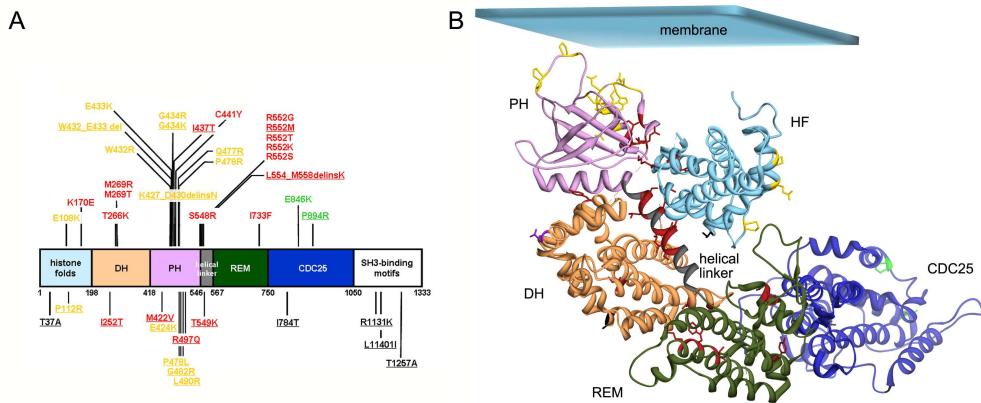


FIGURE 1. *SOS1* domain structure and location of residues altered in Noonan syndrome.

Peer Review

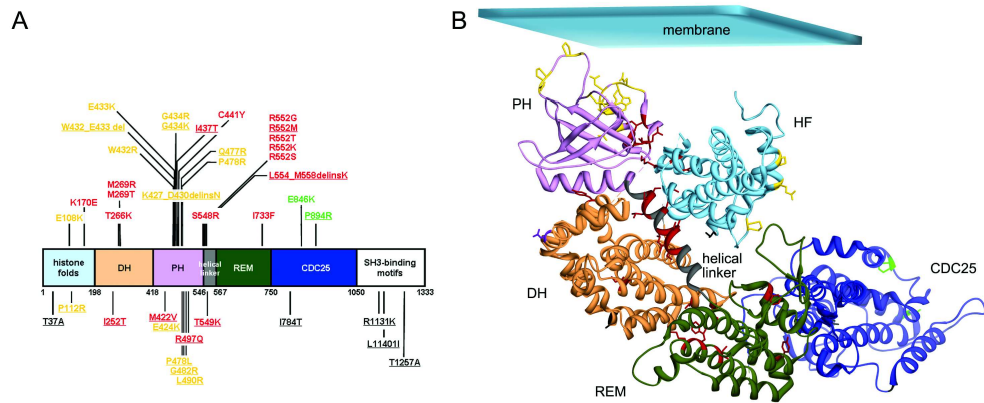


FIGURE 1. *SOS1* domain structure and location of residues altered in Noonan syndrome.

Peer Review

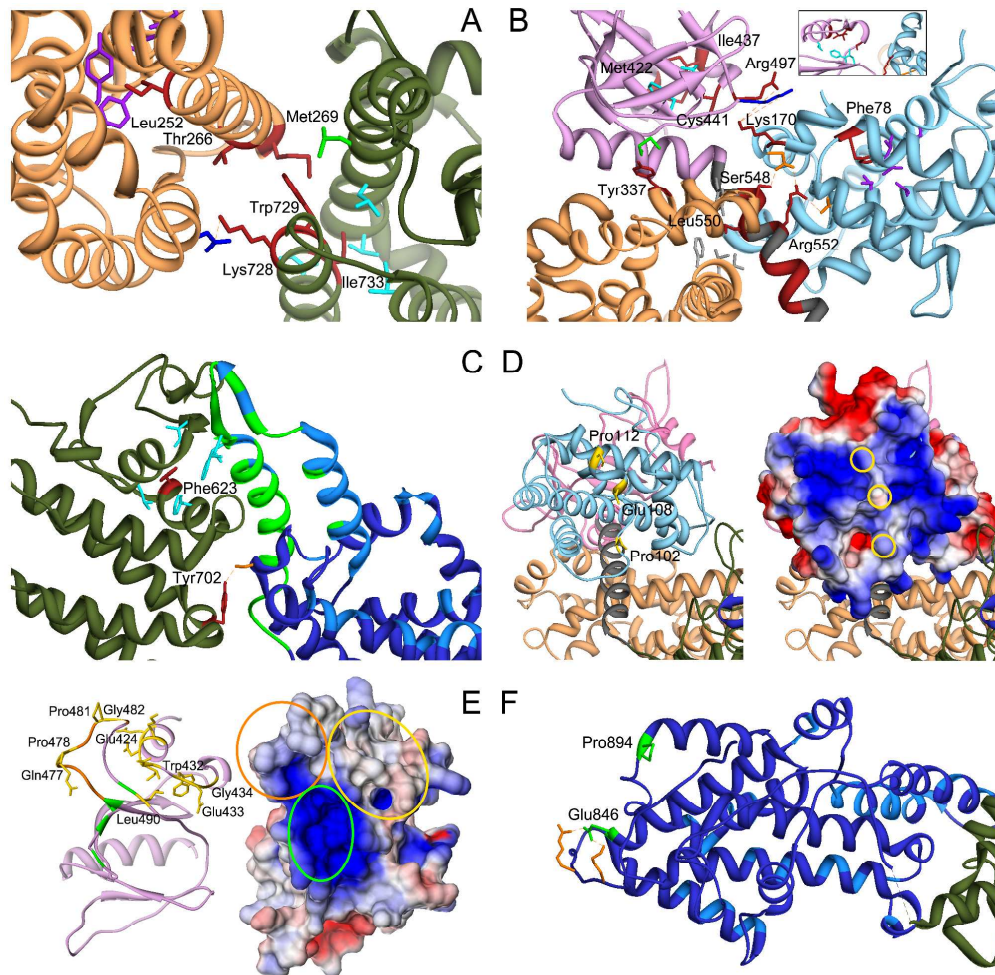


FIGURE 2. Detailed analysis of structural perturbations resulting from Noonan syndrome-causing amino acid substitutions.

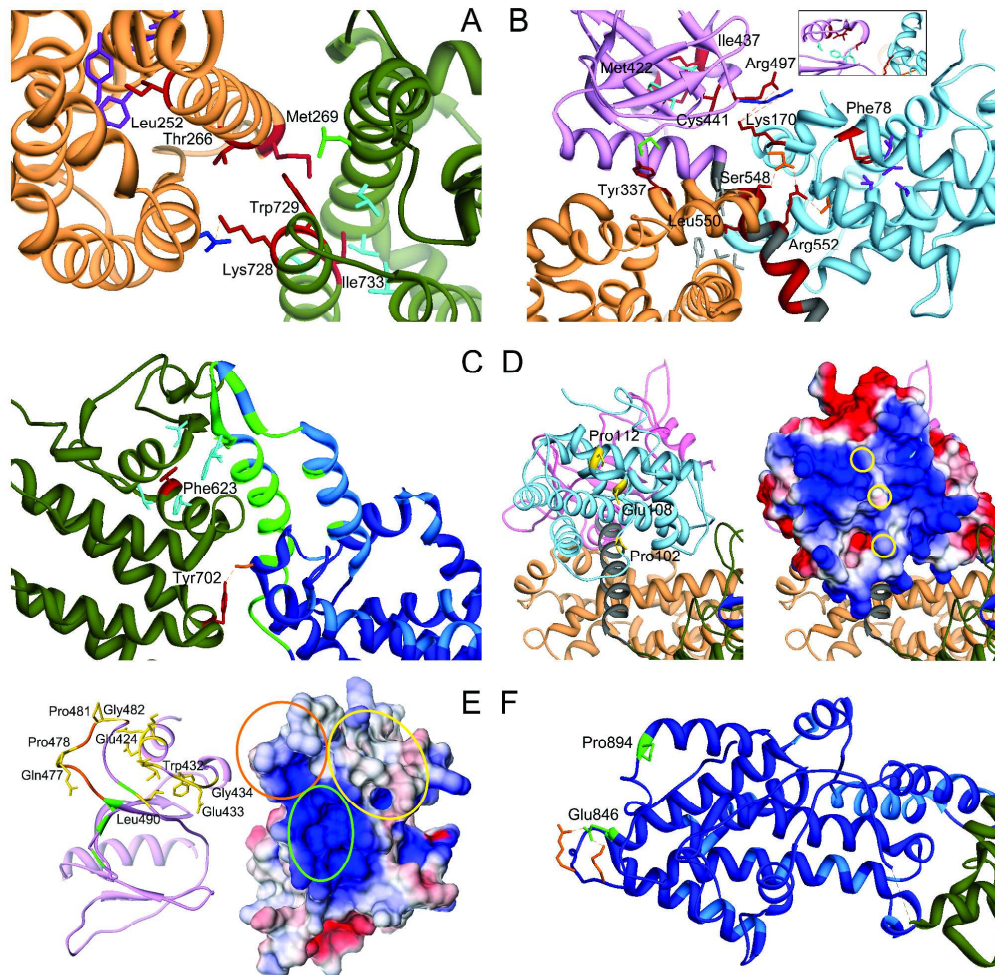


FIGURE 2. Detailed analysis of structural perturbations resulting from Noonan syndrome-causing amino acid substitutions.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



FIGURE 3. Facial dysmorphism and other features of subjects with Noonan syndrome heterozygous for mutations in the SOS1 gene.

Supp. Table S1. List of disease-unrelated *SOS1* sequence variants identified in the study

Location	Nucleotide change	Amino acid change
Intron 2	c.87+86delG	-
Intron 4	c.345+52C>G	-
Intron 5	c.510+24T>C	-
	c.510+59C>G	-
Intron 6	c.720+25C>G	-
	c.721-69T>C	-
	c.721-18G>C	-
Intron 7	c.864+23C>T	-
	c.864+24G>A	-
Intron 9	c.1074+5G>C	-
	c.1075-25C>G	-
Intron 10	c.1202+25A>T	-
	c.1203-20T>C	-
Exon 11	c.1230G>A	Silent change (Gln410)
	c.1705C>G	p.Leu569Val
	c.1800T>A	silent change (Ile600)
	c.1854C>G	silent change (Tyr618)
Intron 11	c.1859-40T>C	-
	c.1859-37T>C	-
Exon 13	c.1964C>T	p.Pro655Leu
Exon 14	c.2122G>A	p.Ala708Thr
Intron 16	c.2511-98C>G	-
Intron 17	c.2673+14T>C	-
	c.2674-22C>G	-
Exon 18	c.2760G>A	silent change (Arg920)
Intron 18	c.2791+53C>T	-
Intron 19	c.2964+43C>A	-
Exon 20	c.3032A>G	p.Asn1011Ser
Exon 21	c.3093T>C	silent change (Tyr1031)
Intron 21	c.3346+72+73delCT	-
	c.3347-143T>C	-
	c.3347-127T>C	-
Exon 22	c.3357C>T	silent change (Thr1119)
Intron 22	c.3391+7A>G	-
	c.3392-89T>G	-
Exon 24	c.3959A>G	p.His1320Arg

Nucleotide numbering of the exonic variants reflects cDNA numbering with 1 corresponding to the A of the ATG translation initiation codon in the reference sequence (NM_005633.3). Exon 2 corresponds to the first protein coding exon. Position of intronic variants is numbered according to the reference genomic sequence (NG_007530.1).

Supp. Table S2. List of *SOS1* exonic indels and missense changes (updated to July 2010)

Nucleotide change	Amino acid change	Status and functional class	N	References
c.109A>G	p.Thr37Ala	Unclassified variant	1	<i>p.s.</i>
c.233T>G	p.Phe78Cys	Probably pathogenic, 1B	1	3
c.305C>G	p.Pro102Arg	Mutation, 2A	1	11
c.322G>A	p.Glu108Lys	Mutation, 2A	4	2, <i>p.s.</i>
c.335C>G	p.Pro112Arg	Probably pathogenic, 2A	1	<i>p.s.</i>
c.508A>G	p.Lys170Glu	Mutation, 1B	5	6, 11, <i>p.s.</i>
c.755T>C	p.Ile252Thr	Probably pathogenic, 1A	1	<i>p.s.</i>
c.797C>A	p.Thr266Lys	Mutation, 1A	8	1, 6, 7, 11, <i>p.s.</i>
c.806T>C	p.Met269Thr	Mutation, 1A	9	3, 6, 11, <i>p.s.</i>
c.806T>G	p.Met269Arg	Mutation, 1A	8	1, 2, 3, 6, 8, <i>p.s.</i>
c.925G>T	p.Asp309Tyr	Mutation, uncharacterized	3	1, 3, 5
c.1010A>G	p.Tyr337Cys	Mutation, 1B	1	1
c.1132A>G	p.Thr378Ala	Unclassified variant	1	11
c.1264A>G	p.Met422Val	Probably pathogenic, 1B	1	<i>p.s.</i>
c.1270G>A	p.Glu424Lys	Probably pathogenic, 2B	1	<i>p.s.</i>
c.1281_1289delGAATAT TGA	p.Lys427_Asp430de linsAsn	Mutation, 2B	1	<i>p.s.</i>
c.1294T>C	p.Trp432Arg	Mutation, 2B	5	2, 3, 10, <i>p.s.</i>
c.1294_1299delTGGGAG	p.Trp432_Glu433del	Mutation, 2B	1	<i>p.s.</i>
c.1297G>A	p.Glu433Lys	Mutation, 2B	14	2, 3, 4, 6, 8, 11, <i>p.s.</i>
c.1300G>C	p.Gly434Arg	Mutation, 2B	1	1
c.1300G>A	p.Gly434Arg	Mutation, 2B	3	3, <i>p.s.</i>
c.1300_1301delGGinsAA	p.Gly434Lys	Mutation, 2B	1	<i>p.s.</i>
c.1310T>C	p.Ile437Thr	Mutation, 1B	2	<i>p.s.</i>
c.1322G>A	p.Cys441Tyr	Mutation, 1B	4	6, <i>p.s.</i>
c.1430A>G	p.Gln477Arg	Mutation, 2B	2	<i>p.s.</i>
c.1431G>T; 1433C>T	p.Gln477His; Pro478Leu	Mutation, 2B	1	3
c.1433C>G	p.Pro478Arg	Mutation, 2B	2	3, <i>p.s.</i>
c.1433C>T	p.Pro478Leu	Unclassified variant (2B)	1	<i>p.s.</i>
c.1442_1443insAAGACT TCC	p.Pro481_Gly482 insArgLeuPro	Mutation, 2B	1	5

Nucleotide change	Amino acid change	Status and functional class	N	References
c.1444G>C	p.Gly482Arg	Probably pathogenic, 2B	1	<i>p.s.</i>
c.1469T>G	p.Leu490Arg	Probably pathogenic, 2B	1	<i>p.s.</i>
c.1490G>A	p.Arg497Gln	Probably pathogenic, 1B	1	<i>p.s.</i>
c.1642A>C	p.Ser548Arg	Mutation, 1B	8	1, 2, 6, <i>p.s.</i>
c.1646C>A	p.Thr549Lys	Probably pathogenic, 1B	1	<i>p.s.</i>
c.1649T>C	p.Leu550Pro	Mutation, 1B	2	2, 8
c.1654A>G	p.Arg552Gly	Mutation, 1B	29	1, 2, 3, 8, 11, <i>p.s.</i>
c.1655G>C	p.Arg 552Thr	Mutation, 1B	2	9, <i>p.s.</i>
c.1655G>T	p.Arg 552Met	Mutation, 1B	2	<i>p.s.</i>
c.1655G>A	p.Arg 552Lys	Mutation, 1B	7	2, 3, <i>p.s.</i>
c.1656G>C	p.Arg 552Ser	Mutation, 1B	15	2, 6, 9, <i>p.s.</i>
c.1656G>T	p.Arg 552Ser	Mutation, 1B	4	3, 5, 8
c.1660_1673delCTTGAT GTAACAATinsAA	p.Leu554_Met558 delinsLys	Mutation, 1B	1	<i>p.s.</i>
c.1705C>G	p.Leu569Val	Polymorphism		11, <i>p.s.</i>
c.1867 T>A	p.Phe623Ile	Mutation, 1C	1	3
c.1964C>T	p.Pro655Leu	Polymorphism		1, 2, 3, 11, <i>p.s.</i>
c.2104T>C	p.Tyr702His	Mutation, 1C	2	2, 3
c.2122G>A	p.Ala708Thr	Polymorphism		<i>p.s.</i>
c.2183A>T	p.Lys728Ile	Mutation, 1A	1	12
c.2186G>T	p.Trp729Leu	Mutation, 1A	1	2
c.2197A>T	p.Ile733Phe	Mutation, 1A	2	2, <i>p.s.</i>
c.2351T>C	p.Ile784Thr	Unclassified variant	1	<i>p.s.</i>
c.2536G>A	p.Glu846Lys	Mutation, 3	19	1, 2, 3, 5, 8, <i>p.s.</i>
c.2681C>G	p.Pro894Arg	Mutation, 3	1	<i>p.s.</i>
c.2930A>G	p.Gln977Arg	Unclassified variant	1	2
c.2999G>A	p.Ser1000Asn	Unclassified variant	1	3
c.3032A>G	p.Asn1011Ser	Polymorphism		<i>p.s.</i>
c.3392G>A	p.Arg1131Lys	Unclassified variant	1	<i>p.s.</i>
c.3418T>A	p.Leu1140Ile	Unclassified variant	1	<i>p.s.</i>
c.3769A>G	p.Thr1257Ala	Unclassified variant	1	<i>p.s.</i>
c.3959A>G	p.His1320Arg	Polymorphism	1	2, <i>p.s.</i>

Nucleotide numbering of the exonic variants reflects cDNA numbering with 1 corresponding to the A of the ATG translation initiation codon in the reference sequence (NM_005633.3). *p.s.*, present study.

References

- Beneteau C, Cave H, Moncla A, Munnich A, Verloes A, Leheup B. 2009. SOS1 and PTPN11 mutations in five cases of Noonan syndrome with multiple giant cell lesions. *Eur J Hum Genet* 17:1216-21.
- Denayer E, Devriendt K, de Ravel T, Van Buggenhout G, Smeets E, Francois I, Sznajder Y, Craen M, Leventopoulos G, Mutesa L et al. 2010. Tumor spectrum in children with Noonan syndrome and SOS1 or RAF1 mutations. *Genes Chromosomes Cancer* 49:242-52.
- Ferrero GB, Baldassarre G, Delmonaco AG, Biamino E, Carta C, Rossi C, Silengo MC. 2008. Clinical and molecular characterization of 40 patients with Noonan syndrome. *Eur J Med Genet* 51:566-72.
- Hanna N, Parfait B, Talaat IM, Vidaud M, Elsedfy HH. 2009. SOS1: a new player in the Noonan-like/multiple giant cell lesion syndrome. *Clin Genet* 75: 568-71.
- Jongmans MC, Hoogerbrugge PM, Hilkens L, Flucke U, van der Burgt I, Noordam K, Ruitkamp-Versteeg M, Yntema HG, Nillesen WM, Ligtenberg MJ et al. 2010. Noonan syndrome, the SOS1 gene and embryonal rhabdomyosarcoma. *Genes Chromosomes Cancer* 49:635-41.
- Ko JM, Kim JM, Kim GH, Yoo HW. 2008. PTPN11, SOS1, KRAS, and RAF1 gene analysis, and genotype-phenotype correlation in Korean patients with Noonan syndrome. *J Hum Genet* 53:999-1006.
- Narumi Y, Aoki Y, Niihori T, Sakurai M, Cave H, Verloes A, Nishio K, Ohashi H, Kurosawa K, Okamoto N et al. 2008. Clinical manifestations in patients with SOS1 mutations range from Noonan syndrome to CFC syndrome. *J Hum Genet* 53:834-41.
- Neumann TE, Allanson J, Kavamura I, Kerr B, Neri G, Noonan J, Cordeddu V, Gibson K, Tzschach A, Kruger G et al. 2009. Multiple giant cell lesions in patients with Noonan syndrome and cardio-facio-cutaneous syndrome. *Eur J Hum Genet* 17:420-5.
- Nystrom AM, Ekvall S, Berglund E, Bjorkqvist M, Braathen G, Duchon K, Enell H, Holmberg E, Holmlund U, Olsson-Engman M et al. 2008. Noonan and cardio-facio-cutaneous syndromes: two clinically and genetically overlapping disorders. *J Med Genet* 45:500-6.
- Roberts AE, Araki T, Swanson KD, Montgomery KT, Schiripo TA, Joshi VA, Li L, Yassin Y, Tamburino AM, Neel BG et al. 2007. Germline gain-of-function mutations in SOS1 cause Noonan syndrome. *Nat Genet* 39:70-4.
- Tartaglia M, Pennacchio LA, Zhao C, Yadav KK, Fodale V, Sarkozy A, Pandit B, Oishi K, Martinelli S, Schackwitz W et al. 2007. Gain-of-function SOS1 mutations cause a distinctive form of Noonan syndrome. *Nat Genet* 39:75-9.
- Zenker M, Horn D, Wiczorek D, Allanson J, Pauli S, van der Burgt I, Doerr HG, Gaspar H, Hofbeck M, Gillissen-Kaesbach G et al. 2007. SOS1 is the second most common Noonan gene but plays no major role in cardio-facio-cutaneous syndrome. *J Med Genet* 44:651-6.

Supp.Table S3. Clinical features of *SOS1* mutation-positive subjects

Case	NS1	NS2	NS3	NS4	NS5	NS6	NS7	NS8	NS9	NS10
Code	S1021	S1036	S1052	BO-2564	BO-0138	BO-4043	BO-1767	BO-2905	BO-0235	BO-1834
Sex	M	M	F	M	M	M	M	M	M	M
Age at evaluation	11 y	12 m	1 y	12 m	8 y, 10 m	5 m, 13 d	2 y, 1 m	14 y, 5 m	4 y, 2 m	8 y, 2 m
Amino acid change	R552K	W432_E433del	R552S	E433K	R552G	E433K	T266K	P478R	M269T	M269T
Polyhydramnios	+	-	+	+	-	-	+	-	-	+
Fetal macrosomia	-	-	+	+	-	+	-	-	-	-
Neonatal/infantile growth failure	-	-	-	-	+	-	-	-	+	+
Poor sucking	-	-	+	-	-	-	-	-	+	-
Poor swallowing	-	-	+	-	-	-	-	-	-	-
Apneas	-	-	-	+	-	-	-	-	-	-
Short stature (<3rd centile)	-	-	-	-	-	-	-	+	-	+
Height/length (cm)				71	127	65	98	147	105	113
Age				12 m	8 y, 10 m	5 m, 13 d	2 y, 1 m	14 y, 5 m	4 y, 2 m	8 y, 2 m
Centile	5 th	50 th	75 th	5 th – 10 th	25 th – 50 th	25 th – 50 th	>98 th	<2 nd	50 th	<2 nd
Delayed bone age								+		+
GH deficiency					+			+		
Craniofacial anomalies	+	+	+	+	+	+	+	+	+	+
- Macrocephaly	-	-	-	+	+	-	+	-	-	-
- Scaphocephaly	-	-	-	-	-	-	-	-	-	-
- High forehead	+	+	-	-	+	+	+	+	+	+
- Bitemporal narrowing	+	+	-	-	+	+	+	-	+	+
- Prominent supraorbital ridge	-	-	-	-	+	+	+	+	-	+
- Downslanting palpebral fissures	+	-	+	+	+	+	+	+	+	+
- Hypertelorism	-	-	+	+	+	+	+	-	+	+
- Epicanthal folds	+	-	+	+	+	+	+	+	+	+
- Palpebral ptosis	+	+	+	+	+	-	+	-	-	+
- Flat nasal bridge	+	-	+	+	+	+	+	-	+	+
- Broad nasal root	-	-	+	-	+	+	+	-	+	+
- Prominent philtrum	+	-	+	-	+	+	+	+	-	+
- Thick lips/macrostomia	-	-	-	-	+	-	+	-	-	+
- Low-set ears	+	+	+	+	+	+	+	+	+	+
- Thickened helix	-	-	+	+	+	+	+	-	+	+
- Large, thick ear lobe	-	-	+	+	+	+	+	-	+	+

Lepri et al., *Human Mutation*

6

Case	NS1	NS2	NS3	NS4	NS5	NS6	NS7	NS8	NS9	NS10
- Micrognathia	-	-	+		-	+	-	+	-	+

Supp. Table S3. Clinical features of *SOS1* mutation-positive subjects (continued)

Case	NS11	NS12	NS13	NS14	NS15	NS16	NS17	NS18	NS19	NS20
Code	BO-1651	BO-0596	BO-2635	BO-4338	HD1256	HD290	05-0175	07-0934	HD1284	HD696
Sex	M	M	F	M	F	M	M	F	F	F
Age at evaluation	4 y, 11 m	16 y, 7 m	26 y	1 y, 3 m	9 y, 7 m	22 y	13 y	10 y	6 y, 6 m	9 y
Amino acid change	W432R	E433K	R552G	E846K	R552S	K170E	L490R	R552G	I437T	R552G
Polyhydramnios	-	+	-	+	+	+	-	-	-	-
Fetal macrosomia	-	-	-	-	+	-	-	-	-	-
Neonatal/infantile growth failure	-	+	+	+	+	-	+	+	+	-
Poor sucking	-	-	+	+	-	-	-	-	N.R.	-
Poor swallowing	-	-	-	+	-	-	+	-	N.R.	-
Apneas	-	-	-	-	-	-	-	-	-	-
Short stature (<3rd centile)	-	+	-	+	+	-	-	-	+	-
Height/length (cm)	114	161	154	72	123	172	138.5	107	93	132.5
Age	4 y, 11 m	16 y, 7 m	26 y	1 y, 3 m	9 y, 7 m	22 y	11 y	6 y, 8 m	4 y, 6 m	8 y, 6 m
Centile	90th-97th	2nd	10th	2nd	< 3rd	25th	25th	3rd	<2nd	50th – 75th
Delayed bone age	-	+	+		+		-		+	-
GH deficiency	-	+	-		-	-	-			-
Craniofacial anomalies	+	+	+	+	+	+	+	+	+	+
- Macrocephaly	+	+	+	+	-	+	+	+	+	+
- Scaphocephaly	-	-	-	-	-	-	-	-	-	-
- High forehead	+	+	+	+	+	+	+	+	+	+
- Bitemporal narrowing	-	-	-	-	+	-	+	+	+	-
- Prominent supraorbital ridge	-	+	+	+	+	+	-	-	+	-
- Downslanting palpebral fissures	+	+	+	+	-	+	+	+	+	+
- Hypertelorism	+	+	+	+	-	+	+	+	+	+
- Epicanthal folds	-	+	+	-	+	-	+	-	+	+
- Palpebral ptosis	+	+	+	+	+	+	+	+	-	-
- Flat nasal bridge	-	-	-	-	+	-	+	+	+	-
- Broad nasal root	-	+	-	-	+	-	-	+	+	+
- Prominent philtrum	-	+	-	-	+	-	+	+	-	+

Lepri et al., *Human Mutation*

7

Case	NS11	NS12	NS13	NS14	NS15	NS16	NS17	NS18	NS19	NS20
- Thick lips/macrostomia	-	-	+	+	-	-	+		+	+
- Low-set ears	+	+	+	+	-	+	+	+	+	+
- Thickened helix	-	-	-	-	+	+	-	+	+	+
- Large, thick ear lobe	-	+	+	+	-	-	+	+	+	+
- Micrognathia	+	+	-	-	+	-	-	+	+	-

Supp. Table S3. Clinical features of *SOS1* mutation-positive subjects (continued)

Case	NS21	NS22	NS23	NS24	NS25	NS26	NS27	NS28	NS29	NS30
Code	N1183	N118_3_father	08-0045	N1157	N1212	N1185	HD227	HD215_father	HD215	HD715
Sex	M	M	M	M	M	F	F	M	F	M
Age at evaluation	12 y, 8 m	40 y	3 y	28 y	11 y, 7 m	15 y, 3 m	7 y, 4 m	37 y	3 y	1 y, 5 m
Amino acid change	P894R	P894R	G434K	M269T	R552M	C441Y	S548R	L550P	L550P	R552G
Polyhydramnios	+	N.R.	-	N.R.	-	+	-	-	+	+
Fetal macrosomia	+	-	-	-	-	+	+	-	-	+
Neonatal/infantile growth failure	-	-	-	N.R.	+	+	+	-	+	+
Poor sucking	-	-	-	-	-	+	-	-	+	+
Poor swallowing	-	-	-	-	-	+	-	-	+	+
Apneas	-	-	-	-	-	-	-	-	-	+
Short stature (<3rd centile)	+	+	-	-	+	+	-	-	-	+
Height/length (cm)	137cm	160cm	96cm	170cm	131.8cm	147 cm	118	171	93	79
Age	12 y, 8 m	40 y	3 y	28 y	11 y, 7 m	15 y, 3 m	7 y, 4 m	37 y	3 y	1 y, 5 m
Centile	0.4th – 2nd	0.4th – 2nd	75th	9th – 25th	0.4th – 2nd	< 3rd	25th	25th	50th	25th
Delayed bone age	+	-	-	-	+	+	+	-	+	+
GH deficiency	-	-	-	-	-	-	-	-	-	-
Craniofacial anomalies	+	+	+	+	+	+	+	+	+	+
- Macrocephaly	+	+	-	+	-	-	-	+	+	+
- Scaphocephaly	-	-	-	-	-	-	-	-	-	-
- High forehead	+	-	-	+	+	-	+	+	+	-
- Bitemporal narrowing	-	-	-	+	+	-	+	+	+	-
- Prominent supraorbital ridge	-	-	+	+	-	+	+	+	+	-
- Downslanting palpebral fissures	-	+	-	+	+	+	+	+	+	+
- Hypertelorism	+	+	+	+	-	+	+	+	+	+
- Epicanthal folds	-	-	-	-	+	+	+	-	-	-

Lepri et al., *Human Mutation*

8

Case	NS21	NS22	NS23	NS24	NS25	NS26	NS27	NS28	NS29	NS30
- Palpebral ptosis	+	+	+	+	+	+	+	+	+	+
- Flat nasal bridge	-	-	-	+	+	+	+	-	+	+
- Broad nasal root	-	-	-	-	-	-	-	-	-	-
- Prominent philtrum	-	+	+	-	+	+	-	-	+	+
- Thick lips/macrostomia	+	+	+	+	-	+	+	-	+	+
- Low-set ears	+	+	+	+	+	+	+	+	+	+
- Thickened helix	+	+	+	+	+	+	+	+	+	+
- Large, thick ear lobe	+	-	-	-	+	+	-	-	+	+
- Micrognathia	-	-	-	-	+	+	-	-	-	-

Supp. Table S3. Clinical features of *SOS1* mutation-positive subjects (continued)

Case	NS31	NS32	NS33	NS34	NS35	NS36	NS37	NS38	NS39	Total
Code	HD303	04-0400	HD099	S1073	S1089	HD1090	S1108	S1121	S1135	
Sex	M	M	F	M	M	F	M	F	F	25 M, 14F
Age at evaluation	3 y, 5 m	8 y, 2 m	5 y, 0 m	18 y	34 y	13 y, 7 m	17 y	6 y	3 y	8 y, 10 m ⁺
Amino acid change	R552G	R552G	W729L	R552G	R552G	R552S	R552S	R552K	I733F	
Polyhydramnios	-	+	+	-	+	+	-	-	+	18/37 (49%)
Fetal macrosomia	+	+	-	+	+	-	+	+	-	14/38 (37%)
Neonatal/infantile growth failure	+	-	+	-	-	-	-	-	-	17/35 (49%)
Poor sucking	+	-	+	-	-	-	-	-	-	9/38 (24%)
Poor swallowing	+	-	+	-	-	-	-	-	-	8/37 (22%)
Apneas	-	-	+	-	+	-	-	-	-	5/37 (12%)
Short stature (<3rd centile)	-	-	-	-	-	-	-	-	-	11/39 (28%)
Height/length (cm)	100	115	102	73.5	165	154		118		
Age	3 y, 5 m	8 y, 2 m	5 y, 0 m	20 y	34 y, 7 m	13 y, 7 m		8 y, 1 m		
Centile	75th	<3rd	10th	25th-50th	5th	25th	5th	5th	10th	
Delayed bone age	+	+	-	-	-	+	-	-	-	15/26 (58%)
GH deficiency	-	-	-	-	-	-	-	-	-	3/26 (12%)
Craniofacial anomalies	+	+	+	+	+	+	+	+	+	39/39 (100%)
- Macrocephaly	+	+	+	-	+	-	+	-	+	24/39 (61%)
- Scaphocephaly	-	-	-	-	-	-	-	-	-	0/39 (0%)
- High forehead	-	-	+	-	-	-	-	-	+	25/39 (64%)
- Bitemporal narrowing	-	-	+	-	-	-	-	-	-	17/39 (44%)

Lepri et al., *Human Mutation*

9

Case	NS31	NS32	NS33	NS34	NS35	NS36	NS37	NS38	NS39	Total
- Prominent supraorbital ridge	-	-	+	-	+	-	+	+	+	22/39 (56%)
- Downslanting palpebral fissures	+	+	+	+	+	+	+	+	+	35/39 (90%)
- Hypertelorism	+	+	+	+	+	+	+	+	+	34/39 (87%)
- Epicanthal folds	+	-	+	+	+	+	+	+	+	26/39 (67%)
- Palpebral ptosis	+	+	+	+	+	+	+	+	+	34/39 (87%)
- Flat nasal bridge	+	+	+	+	+	+	+	+	+	27/39 (69%)
- Broad nasal root	-	-	-	-	+	-	-	+	+	14/38 (37%)
- Prominent philtrum	-	+	-	-	+	-	+	-	+	22/39 (56%)
- Thick lips/macrostomia	+	+	+	+	+	+	+	+	+	25/38 (66%)
- Low-set ears	+	+	+	+	+	+	+	+	+	38/39 (97%)
- Thickened helix	+	+	+	+	+	+	-	+	+	30/39 (77%)
- Large, thick ear lobe	-	-	+	+	+	+	+	+	+	25/38 (66%)
- Micrognathia	-	-	-	-	-	-	+	-	-	12/38 (32%)

Supp. Table S3. Clinical features of *SOS1* mutation-positive subjects (continued)

Case	NS1	NS2	NS3	NS4	NS5	NS6	NS7	NS8	NS9	NS10
Congenital heart defects	+	+	+	+	+	+	+	-	+	+
- Pulmonic stenosis (PS)	+	+	+	+	-	+	+	-	+	+
- Pulmonary valve dysplasia	-	-	-	-	-	-	-	-	-	-
- Mitralic valve dysplasia	-	-	-	-	-	-	-	-	-	-
- Supravalvular PS	-	-	-	-	-	-	-	-	-	-
- Atrial septal defect	-	-	+	-	-	-	-	-	-	-
- Hypertrophic cardiomyopathy	-	-	-	-	-	-	-	-	-	-
- Atrioventricular canal	-	-	-	-	-	-	-	-	-	-
- Tetralogy of Fallot	-	-	-	-	-	-	-	-	-	-
- Ventricular septal defect	-	-	-	-	+	-	-	-	-	+
- others	-	-	-	-	-	-	-	+	-	-
Skin features	-	-	+	-	-	-	+	-	+	+
- Dark skin	-	-	-	-	-	-	-	-	+	+
- palmo/plantar crease	-	-	+	-	-	-	-	-	-	-
- Pruritus	-	-	-	-	-	-	+	-	-	-
- others	-	-	-	-	-	-	-	-	-	-
Ectodermal features	+	-	+	-	+	+	+	+	+	+

Case	NS1	NS2	NS3	NS4	NS5	NS6	NS7	NS8	NS9	NS10
- Sparse/absent scalp hair	-	-	-	-	-	-	+	-	-	-
- Thin hair	-	-	-	-	-	+	+	-	-	-
- Curly hair	+	-	-	-	+	-	-	+	-	+
- Sparse eyebrows	+	-	+	-	+	+	+	+	+	+
- Keratosis pilaris faciei	+	-	-	-	-	-	-	-	-	-
- Thin dystrophic nails	-	-	-	-	-	-	-	-	-	-
Musculoskeletal abnormalities	+	-	+	+	+	+	+	+	+	+
- Short webbed neck	+	-	+	+	+	+	+	+	+	+
- Cubitus valgus	-	-	-	-	+	+	+	+	+	+
- Hyperextensible joints	-	-	-	-	-	-	-	-	-	+
- Pectus anomalies	-	-	+	+	+	+	+	+	+	+
Undescended testes	+	-	+	+	+	+	+	+	+	+
CNS involvement	-	+	-	+	-	-	-	-	-	+
- Mental retardation	-	+	-	-	-	-	-	-	-	+
- Hyperactivity	-	+	-	-	-	-	-	-	-	+
- CNS abnormalities	-	-	-	-	-	-	-	-	-	+
- Seizures/EEG abnormalities	-	-	-	-	-	-	-	-	-	+

24
25 **Supp. Table S3. Clinical features of *SOS1* mutation-positive subjects (continued)**

Case	NS11	NS12	NS13	NS14	NS15	NS16	NS17	NS18	NS19	NS20
Congenital heart defects	+	+	+	+	+	+	+	+	+	+
- Pulmonic stenosis (PS)	-	-	+	+	+	+	+	+	+	+
- Pulmonary valve dysplasia	-	-	-	-	+	-	-	-	+	-
- Mitralic valve dysplasia	-	-	-	-	+	-	-	-	-	-
- Supravalvular PS	-	-	-	-	-	-	-	-	-	-
- Atrial septal defect	-	-	+	-	-	-	-	+	+	+
- Hypertrophic cardiomyopathy	-	-	-	-	-	-	-	-	+	-
- Atrioventricular canal	-	-	-	-	-	-	-	-	-	-
- Tetralogy of Fallot	-	-	-	-	-	-	-	-	-	-
- Ventricular septal defect	-	+	-	-	-	-	-	-	+	-
- others	+	+	-	-	-	-	-	-	-	-
Skin features	-	-	+	+	+	+	-	+	-	-
- Dark skin	-	-	+	-	-	-	-	-	-	-

Case	NS11	NS12	NS13	NS14	NS15	NS16	NS17	NS18	NS19	NS20
- palmo/plantar crease	-	-	+	+	-	-	-	+	-	-
- Pruritus	-	-	-	+	-	-	-	-	-	-
- others	-	-	+	(SK)	+	(DrSk)	+	(K)	-	-
Ectodermal features	+	+	+	+	+	+	+	+	+	+
- Sparse/absent scalp hair	+	-	+	+	+	-	-	-	-	-
- Thin hair	+	+	+	+	+	-	-	+	+	+
- Curly hair	-	+	+	+	+	+	-	-	-	+
- Sparse eyebrows	+	+	+	+	+	+	+	+	+	+
- Keratosis pilaris faciei	-	-	+	-	-	+	-	-	-	-
- Thin dystrophic nails	-	-	-	-	+	-	-	-	-	-
Musculoskeletal abnormalities	+	-	+	+	+	+	+	+	+	+
- Short webbed neck	-	-	+	+	+	+	-	+	+	+
- Cubitus valgus	-	-	-	-	-	-	-	-	-	-
- Hyperextensible joints	-	-	+	+	+	-	-	+	-	-
- Pectus anomalies	+	-	+	+	-	+	+	(PE)	+	(PC)
Undescended testes	-	-	-	-	-	+	-	-	-	-
CNS involvement	-	-	-	-	-	-	-	+	-	-
- Mental retardation	-	-	-	-	-	-	-	+	-	(*)
- Hyperactivity	-	-	-	-	-	-	-	+	-	(***)
- CNS abnormalities	-	-	-	-	-	-	-	-	-	-
Seizures/EEG abnormalities	-	-	-	-	-	-	-	-	-	-

Supp. Table S3. Clinical features of *SOS1* mutation-positive subjects (continued)

Case	NS21	NS22	NS23	NS24	NS25	NS26	NS27	NS28	NS29	NS30
Congenital heart defects	-	-	+	+	+	+	+	-	+	+
- Pulmonic stenosis (PS)	-	-	+	+	-	+	-	-	+	+
- Pulmonary valve dysplasia	-	-	+	-	-	+	-	-	+	-
- Mitralic valve dysplasia	+	-	-	-	+	-	-	-	-	+
- Supravalvular PS	-	-	-	-	-	-	-	-	+	-
- Atrial septal defect	-	-	+	-	-	-	+	-	-	-
- Hypertrophic cardiomyopathy	-	-	-	-	-	-	-	-	-	-
- Atrioventricular canal	-	-	-	-	-	-	-	-	-	-
- Tetralogy of Fallot	-	-	-	-	-	-	-	-	-	-

Lepri et al., *Human Mutation*

12

Case	NS21	NS22	NS23	NS24	NS25	NS26	NS27	NS28	NS29	NS30
- Ventricular septal defect	-	-	-	-	-	-	-	-	-	-
- others	-	-	-	-	AVD	MVP,AVDRCoTA	-	-	-	-
Skin features	+	+	+	-	-	-	+	+	+	+
- Dark skin	+	+	+	-	-	-	+	+	+	+
- Palmo/plantar crease	-	-	-	-	-	-	-	-	-	+
- Pruritus	+	-	-	-	-	-	-	-	-	+
- others	-	-	-	-	-	-	-	-	-	+(CA)
Ectodermal features	+	+	+	+	-	+	+	+	+	+
- Sparse/absent scalp hair	-	-	-	-	-	-	-	-	-	-
- Thin hair	-	-	-	-	-	-	-	-	-	-
- Curly hair	+	+	+	+	-	+	+	+	+	+
- Sparse eyebrows	+	-	+	-	-	+	+	-	+	+
- Keratosis pilaris faciei	+	-	+	-	-	+	-	-	-	-
- Thin dystrophic nails	-	-	-	-	-	-	-	+	-	-
Musculoskeletal abnormalities	+	+	+	+	+	+	+	+	+	+
- Short/webbed neck	-	+	-	-	+	+	+	-	+	+
- Cubitus valgus	-	-	-	-	-	+	-	+	+	-
- Hyperextensible joints	+	+	+	-	-	+	+	+	+	+
- Pectus anomalies	+	-	-	+	+	+	+	+	+	+
Undescended testes	+	-	+	+	+	-	-	+	-	+
CNS involvement	-	-	-	-	-	+	-	-	-	-
- Mental retardation	-	-	-	-	-	+(*****)	-	-	-	-
- Hyperactivity	-	-	-	-	-	+	-	-	-	-
- CNS abnormalities	-	-	-	-	-	-	-	-	-	-
- Seizures/EEG abnormalities	-	-	-	-	-	-	-	-	-	-

Supp. Table S3. Clinical features of *SOS1* mutation-positive subjects (continued)

Case	NS31	NS32	NS33	NS34	NS35	NS36	NS37	NS38	NS39	Total
Congenital heart defects	+	+	+	+	+	+	+	+	+	35/39 (90%)
- Pulmonic stenosis (PS)	-	+	+	+	-	+	+	+	-	27/39 (69%)
- Pulmonary valve dysplasia	+	+	+	-	-	+	+	+	-	11/39 (28%)
- Mitralic valve dysplasia	+	-	-	-	-	-	-	-	+	6/39 (15%)
- Supravalvular PS	-	+	+	-	-	+	-	-	-	4/39 (10%)

Lepri et al., *Human Mutation*

13

Case	NS31	NS32	NS33	NS34	NS35	NS36	NS37	NS38	NS39	Total
- Atrial septal defect	-	+	+	+	-	-	-	-	-	10/39 (26%)
- Hypertrophic cardiomyopathy	+	-	-	-	+	-	-	-	-	3/39 (8%)
- Atrioventricular canal	-	-	-	-	-	-	-	-	-	0/39 (0%)
- Tetralogy of Fallot	-	-	-	-	-	-	-	-	-	0/39 (0%)
- Ventricular septal defect	-	-	-	-	-	+	-	-	-	5/39 (13%)
- others	-	-	-	-	-	-	-	-	-	
Skin features	+	+	+	-	+	+	+	-	-	22/39 (57%)
- Dark skin	+	+	+	-	+	+	+	+	+	18/39 (46%)
- palmo/plantar crease	-	-	-	-	+	-	-	+	-	7/38 (18%)
- Pruritus	+	+	+	-	-	-	-	+	-	8/36 (22%)
- others									+	(CA)
Ectodermal features	+	+	+	+	-	+	+	+	+	35/39 (90%)
- Sparse/absent scalp hair	-	-	-	-	-	-	-	+	+	7/39 (18%)
- Thin hair	-	-	-	-	-	-	-	+	+	12/39 (31%)
- Curly hair	+	+	+	+	-	+	+	+	+	27/39 (69%)
- Sparse eyebrows	-	+	+	-	-	+	+	+	+	30/39 (77%)
- Keratosis pilaris faciei	+	+	+	-	+	+	+	+	-	13/39 (33%)
- Thin dystrophic nails	-	-	-	-	+	-	-	+	-	4/38 (10%)
Musculoskeletal abnormalities	+	+	+	+	+	+	+	+	+	37/39 (95%)
- Short webbed neck	+	+	+	+	+	+	+	+	+	31/39 (79%)
- Cubitus valgus	-	-	+	+	+	+	+	+	+	16/37 (43%)
- Hyperextensible joints	+	+	+	-	-	-	-	+	-	17/37 (46%)
- Pectus anomalies	+	+	+	+	+	+	+	+	+	32/39 (82%)
Undescended testes	-	+		+	+					18/26 (69%)
CNS involvement	-	-	-	-	-	-	-	-	-	5/38 (13%)
- Mental retardation	-	-	-	-	-	-	-	-	-	4/39 (13%)
- Hyperactivity	-	-	-	-	-	-	-	-	-	1/37 (3%)
- CNS abnormalities	-	-	-	-	AC	-	-	-	-	2/37 (5%)
- Seizures/EEG abnormalities	-	-	-	-	-	-	-	-	-	1/39 (3%)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Lepri et al., *Human Mutation*

14

Supp. Table S3. Clinical features of *SOS1* mutation-positive subjects (continued)

Case	NS1	NS2	NS3	NS4	NS5	NS6	NS7	NS8	NS9	NS10
Ocular abnormalities	-	-	+(Sb)	+(Ast)	-	+(Sb)	-	+(My)	-	+
Gastrointestinal abnormalities	-	-	+(GER)	-	-	-	-	-	-	-
Renal abnormalities	-	-	-	-	-	-	-	-	-	-
Coagulation defects	-	-	-	-	+(VWD)	-	-	-	-	-
Mandibular giant cell lesions	-	-	-	-	-	-	-	-	-	-
Other tumors	-	-	-	-	-	-	-	-	-	-
Others	-	-	-	-	-	-	-	IH	-	-

Supp. Table S3. Clinical features of *SOS1* mutation-positive subjects (continued)

Case	NS11	NS12	NS13	NS14	NS15	NS16	NS17	NS18	NS19	NS20
Ocular abnormalities	-	-	+(IC, My)	-	-	-	-	+(My, Ast,Sb)	-	-
Gastrointestinal abnormalities	-	-	-	-	-	-	-	-	-	+(GER solv.)
Renal abnormalities	-	+(BH)	+(IN)	-	-	-	+(ERP)	-	-	-
Coagulation defects	-	-	-	-	-	+	-	+(CPD)	+(##)	-
Mandibular giant cell lesions	-	-	-	-	-	-	-	-	-	+
Other tumors	-	-	-	-	-	-	-	-	-	-
Others	-	-	V, CoHipD	luDy, AbU	SFF	LTT	-	-	-	-

Supp. Table S3. Clinical features of *SOS1* mutation-positive subjects (continued)

Case	NS21	NS22	NS23	NS24	NS25	NS26	NS27	NS28	NS29	NS30
Ocular abnormalities	-	-	-	-	-	+	-	-	-	-
Gastrointestinal abnormalities	-	-	-	-	-	-	-	-	-	-
Renal abnormalities	-	-	-	-	-	-	+	-	-	-
Coagulation defects	-	-	-	-	-	-	+	-	-	-
Mandibular giant cell lesions	+	-	-	-	-	-	-	-	-	-
Other tumors	abRMS, cGI, GCC	-	-	-	-	-	-	-	-	-
Other	N, L, HypThy	-	HTGum	-	-	-	-	-	-	-

Supp. Table S3. Clinical features of *SOS1* mutation-positive subjects (continued)

Case	NS31	NS32	NS33	NS34	NS35	NS36	NS37	NS38	NS39	Total
Ocular abnormalities	-	-	+ (Sb)	+	+	-	+	+	+	13/39 (33%)
Gastrointestinal abnormalities	-	-	-	-	+	-	-	-	-	3/38 (8%)
Renal abnormalities	-	-	-	-	-	-	-	-	-	4/38 (10%)
Coagulation defects	-	+	-	-	+	+	-	+	-	9/34 (26%)
Mandibular giant cell lesion	-	-	-	-	-	+	-	-	-	3/39 (8%)
Other tumors	-	-	-	-	-	-	-	-	-	1/39 (3%)
Others	-	-	-	-	-	ArS	-	-	-	

abRMS= abdominal rhabdomyosarcoma; AbU= absence of uvula; AC= Arnold-Chiari malformation; Ar= arrhythmia; ArCy= arachnoid cyst; ArS= articular synovitis; Ast= astigmatism; AVD= aortic valve dysplasia; BH= bilateral hydronephrosis; CoHipD= congenital hip dysplasia; CA= Cavernous angiomata; cGl= cerebral glioma; CPD= coagulation-platelet dysfunction; DrSk= dry skin; DUT= Double urinary tract; ERP= enlarged renal pelvis; F= female; GCC= granular cell tumors; GER= gastroesophageal reflux; HypoGly= Hypoglycemia; HypThy= hypothyroidism; HTGum= Hypertrophic gums; IH= inguinal hernia; K= keloid; L= lentigines; lC= left cataract; lN= left nephrolithiasis; luDy= 1st lumbar and 4th cervical vertebral dysmorphism; LV= left ventricle; LTT= large thumbs and toes; M= male; MVI= mitral valve insufficiency; MVP= mitral valve prolapse; My= Myopia; N= nevi; PC= pectus carenatum; PE= pectus excavatum; PP= pes planovalgus; RCoTA= right cor triatriatum; s= severe; Sb= strabismus; solv.= solved; SK= skin keloids; SFF= severe flat feet; TI= tricuspid insufficiency; V= valgism; VWD= von willebrand disease.

⁺Median value (range: 5 months to 40 years), [#]More prominent on the left eye; ^{##}Coagulation factors VII, IX, and XI deficiency, [^]First steps at 16-17 months of age; *The subject had hypoxia as a complication of the heart surgery; ** Developmental delay was documented, with unsupported walking reached at 18-20 months and first spoken words at 30 months. Cognitive development was impaired, with requirement of a full time special education setting. Poor fine motor skill was also reported; ***Speech delay; ****Wechsler Preschool and Primary Scale of Intelligence Total Intellectual Quotient = 81; at 4 years, according to the Griffith's scale, the General Quotient was border-line to the inferior normal limits, with impairment of linguistic skills and oculo-manual coordination; *****Cognitive deficits were secondary to surgery complication during the neonatal period (IQ score= 44, WISC-R scale).