Gain-of-function *SOS1* mutations cause a distinctive form of Noonan syndrome

Marco Tartaglia[1], Len A. Pennacchio[2,3], Chen Zhao[4], Kamlesh K. Yadav[4], Valentina Fodale[1], Anna Sarkozy[5], Bhaswati Pandit[6], Kimihiko Oishi[6], Simone Martinelli[1], Wendy Schackwitz[2,3], Anna Ustaszewska[2], Joel Martin[2,3], James Bristow[2,3], Claudio Carta[1], Francesca Lepri[5], Cinzia Neri[5], Isabella Vasta[7], Kate Gibson[8], Cynthia J. Curry[9], Juan Pedro López Siguero[10], Maria Cristina Digilio[11], Giuseppe Zampino[7], Bruno Dallapiccola[5], Dafna Bar-Sagi[4], Bruce D. Gelb[6]

[1]Dipartimento di Biologia Cellulare e Neuroscienze, Istituto Superiore di Sanità, Viale Regina Elena, 299, 00161Rome, Italy. [2]Genomics Division, MS 84-171, Lawrence Berkeley National Laboratory, Berkeley, California 94720 USA. [3]U.S. Department of Energy Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, California 94598, USA. [4]Department of Biochemistry, New York University School of Medicine, 550 First Avenue, New York, New York 10016. [5]IRCCS-CSS, San Giovanni Rotondo and CSS-Mendel Institute, Rome, and Department of Experimental Medicine and Pathology, University "La Sapienza", Viale Regina Elena 261, 00198, Rome, Italy. [6]Center for Molecular Cardiology and Departments of Pediatrics and Human Genetics, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, New York 10029, USA. [7]Istituto di Clinica Pediatrica, Università Cattolica del Sacro Cuore, Largo A. Gemelli 8, 00168, Rome, Italy. [8]Royal Children's Hospital, Herston Road, Herston, Queensland 4029, Australia. [9]Genetic Medicine Central California, 351 East Barstow #106, Fresno, California 93710, USA. [10]Endocrinología Pediátrica, Hospital Materno-Infantil, Avida

Arroyo de los Ángeles, 29011 Málaga, Spain. . [11]Genetica Medica, Ospedale "Bambino Gesù", Piazza S. Onofrio 4, 00165, Rome, Italy.

Corresponding Authors:

Marco Tartaglia, Ph.D.

Dipartimento di Biologica Cellulare e Neuroscienze

Istituto di Superiore Sanità

Viale Regina Elena, 299

00161 Rome, Italy

Phone: +39-06-4990-2569

Fax: +39-06-4938-7143

Email: mtartaglia@iss.it

Bruce D. Gelb, M.D.

Center for Molecular Cardiology

Mount Sinai School of Medicine

One Gustave L. Levy Place, Box 1040

New York, NY 10029

Tel: 212-241-3302

Fax: 212-241-3310

Email: bruce.gelb@mssm.edu

Noonan syndrome (NS) is a developmental disorder characterized by short stature, facial dysmorphia, congenital heart defects and skeletal anomalies[1]. Increased RAS-mitogen-activated protein kinase (MAPK) signaling due to *PTPN11* and *KRAS* mutations cause 50% of NS[2-6]. Here, we report that 22 of 129 NS patients without *PTPN11* or *KRAS* mutation (17%) have missense mutations in *SOS1*, which encodes a RAS-specific guanine nucleotide exchange factor (GEF). *SOS1* mutations cluster at residues implicated in the maintenance of SOS1 in its autoinhibited form and ectopic expression of two NS-associated mutants induced enhanced RAS activation. The phenotype associated with *SOS1* defects is distinctive, although within NS spectrum, with a high prevalence of ectodermal abnormalities but generally normal development and linear growth. Our findings implicate for the first time gain-of-function mutations in a RAS GEF in inherited disease and define a new mechanism by which upregulation of the RAS pathway can profoundly change human development.

NS is an autosomal dominant, genetically heterogeneous trait. *PTPN11*, the first NS-associated gene identified[6], was discovered through positional cloning. It encodes the non-membranous protein tyrosine phosphatase, SHP-2, that primarily serves positive regulatory roles in signal transduction, particularly via the receptor tyrosine kinase (RTK)-mediated RAS-MAPK pathway. Most mutations perturb the switch between the basally inactive and phosphotyrosine-bound active conformations of SHP-2, shifting the equilibrium towards the latter[3,6-8]. Similar somatic *PTPN11* mutations underlie several hematopoietic disorders, particularly juvenile myelomonocytic leukemia (JMML)[9]. Gain-of-function *RAS* mutations and second allele loss of

NF1 also cause JMML, emphasizing the effects of gain-of-function *PTPN11* mutations in increasing RAS-MAPK signaling.

Recent disease gene discovery established that NS and some phenotypically related traits are classed etiologically as disorders of dysregulated RAS-MAPK signaling. LEOPARD syndrome arises from a functionally distinct class of *PTPN11* mutations[10,11]. Gain-of-function germline mutations in *HRAS* have been found in Costello syndrome and in *KRAS* in severe NS and cardio-facio-cutanteous (CFC) syndrome[2,4,12,13]. In addition, *BRAF, MEK1* and *MEK2* mutations have been observed in CFC[13,14].

The initial and rate limiting step in the activation of the RAS-MAPK pathway by extracellular signals is the ligand-dependent conversion of RAS-GDP to RAS-GTP. In the context of RTK signaling, this reaction is catalyzed by the RAS-specific guanine nucleotide exchange factor (GEF) Son of Sevenless (SOS)[15]. Therefore, mutational activation of SOS provides the potential for the upregulation of RAS signaling, an apparent requisite for NS disease pathogenesis. The human genome contains two *SOS* genes, *SOS1* and *SOS2*, encoding highly similar multi-domain proteins (Fig. 1). Structure function studies of SOS1 have indicated that in the basal state the protein is maintained in an autoinhibited conformation due to complex regulatory intra- and inter-molecular interactions[16-18]. Following RTK stimulation, SOS1 is recruited to the plasma membrane where it acquires a catalytically active conformation by mechanisms that are not fully understood.

To explore whether *SOS1* mutations have a role in NS, we assembled genomic DNAs from 96 individuals who were negative for mutations in the two established NS genes (Cohort A) and performed high throughput resequencing of the *SOS1* coding region (exons 2-24) and flanking intronic boundaries. We identified 33 sequencing variants, including 12 non-synonymous

changes observed in 15 samples (Table 1 and Supplemental Table 1). Strikingly, three variants, affecting six subjects, altered Arg^{552} and a fourth resided nearby, altering Leu^{550}. Both residues are evolutionarily conserved.

To provide further evidence that nonsynonymous variants were mutations, we leveraged the sporadic cases. Among the seven variants for which we possessed both parental samples, we failed to observe the relevant sequence change in either parent in five instances; paternity was confirmed in each, providing final proof that these were *de novo* mutations (Table 1). In single cases, P655L and H1320R were inherited from unaffected parents. W432R, L550P, and Y702H were observed in families with two to three affected individuals and co-segregated with disease. For the exons with variants that were not *de novo* (exons 11, 13, 14, 17 and 24), we analyzed 155 control individuals, identifying only the P655L variant in the population. Thus, we concluded that ten missense changes were disease-causing mutations. We suspect that the H1320R change is a rare polymorphism but cannot exclude incomplete penetrance in the unaffected parent. The prevalence of *SOS1* mutations in the cohort was 13/96 or 12.5% (95% C.I.: 7.4-22%), which can be considered a lower limit due to the incomplete coverage inherent with our high throughput approach.

To elucidate further the range of molecular defects associated with disease, *SOS1* exons 2-24 were scanned with DHPLC in an additional 33 NS samples without *PTPN11* or *KRAS* mutation (Cohort B). Sequencing of amplimers with abnormal denaturing profiles revealed an additional seven missense mutations among nine subjects as well as another probable rare nonsynonymous polymorphism, Q977R, that a proband inherited from an unaffected mother but was not found among the controls (Table 1).

In Cohort B, two additional mutations altering Arg^{552} and two independent S548R alleles were observed, providing further evidence for the importance of that region. A second mutation cluster in SOS1's pleckstrin homology (PH) domain became apparent with the identification of an additional instance of E433K as well as a C441Y mutant. A third functional cluster residing in the interacting regions of the Dbl homology (DH) and RAS exchanger motif (Rem) domain was apparent with the identification of M269R, which joined W729L and I733F identified in Cohort A.

The GEF activity of SOS is principally controlled by two regulatory determinants: A catalytic site where the dissociation of nucleotide from RAS occurs and an allosteric site which stimulates exchange activity through the binding of nucleotide-bound RAS[19]. Whereas the catalytic site is located entirely within the Cdc25 domain, the allosteric site is bracketed by the Cdc25 domain and Rem domains. In resting conditions, SOS is autoinhibited due to a blockade of the allosteric site by the DH-PH unit[18]. The three NS-associated *SOS1* mutation clusters reside in regions within the molecule that are predicted to contribute structurally to the maintenance of the protein in its autoinhibitied conformation. For example, Arg^{552} lies in the helical linker between the PH and Rem domains (Fig. 1) and is predicted to interact directly with the side chains of Asp^{140} and Asp^{169} in the histone domain of SOS1[17]. The disruption of this interaction could affect the relative orientation of the DH-PH unit and the Rem domain. The mutation cluster represented by E433K and C441Y may disrupt the autoinhibited conformation by causing a structural destabilization of the PH domain. The third cluster (M269R, W729L and I733F) consists of residues that mediate the interaction of the DH and Rem. Trp^{729} interacts directly with Met^{269}, thereby positioning the DH domain in its autoinhibitory conformation. Trp^{729} is also critical for the binding of Ras at the allosteric site. Sondermann and co-workers engineered a W729E SOS1

mutant, which proved unable to bind RAS-GTP at the allosteric site and had severely reduced nucleotide release from RAS-GDP at the catalytic site[18]. Since the NS-associated W729L substitution is more conservative, we suspected that it would preferentially affect autoinhibition.

To examine directly the effects of these mutations on SOS1 function, we chose two representative *SOS1* mutants, R552G and W729L, which were expressed transiently in Cos-1 cells. RAS activation, as a read-out of GEF activity, was assessed using a RAF-RBD binding assay (Fig. 2). When wild type SOS1 was expressed, RAS activation was low in starved cells, increased 26-fold rapidly after EGF stimulation and returned to basal levels by 30 min. In contrast, expression of R552G or W729L SOS1 resulted in the accumulation of activated RAS in starved cells and the response to EGF was prolonged. These results confirmed our predictions based on structural information that the NS-associated *SOS1* mutations would principally abrogate autoinhibition, resulting in increased RAS activation.

Extensive phenotype data were available for 16 individuals with *SOS1* missense mutations. These individuals had cardiac disease (primarily pulmonary valve stenosis), pectus deformities, shorted and webbed neck, and dysmorphic facial features ranging from typical for NS to an appearance resembling cardio-facio-cutaneous syndrome (Table 2 and Supplemental Figure 1). Ectodermal features including facial keratosis pilaris, hypoplastic eyebrows and curly hair were significantly more prevalent among individuals with a *SOS1* mutation compared to the general NS population, particularly during childhood. Short stature with a height below the 3rd centile was observed in only 2 of 15 individuals with a *SOS1* mutation, whereas prevalence is 70-76% among NS in general and *PTPN11* mutation-negative NS. In contrast, macrocephaly was overrepresented among those with SOS1 mutations. Only one individual with a *SOS1* mutation had mental retardation, potentially attributable to critical illness as a newborn. In comparison, 30

and 35% of all children with NS and those without a *PTPN11* mutation, respectively, require special education.

The analysis of 129 probands with NS from two cohorts identified 14 different molecular lesions among 22 independent cases (17% of *PTPN11-/KRAS*-mutation-negative cases), making *SOS1* the 2nd most common cause of this disorder thus far. Like *PTPN11*, *SOS1* mutations were found in sporadic and familial NS and engendered a high prevalence of pulmonary valve disease. The *SOS1*-associated phenotype, while clearly within the NS spectrum, resembled CFC somewhat in its dysmorphia, macrocephaly and ectodermal manifestations but differed notably with preserved development and linear growth. *SOS1* mutations were invariably missense defects and clustered at specific regions implicated in the complex regulation of SOS catalytic activity. Among mutations causing developmental disorders with dysregulated RAS-MAPK signaling, *SOS1* defects are notable for affecting a protein functioning entirely upstream of RAS.

The results reported here represent the first examples of inherited gain-of-function mutations in *SOS1*. A frameshift mutation in exon 21 of *SOS1* has been reported in one family inheriting hereditary gingival fibromatosis[20]. No additional case inheriting a *SOS1* mutation with this autosomal dominant trait has been reported, deferring final proof of causality. Complete loss of Sos1 in mice is embryonic lethal due to placental and cardiac defects, but haploinsufficient mice develop normally[21,22]. Transgenic mice expressing a dominant active form of Sos1 in keratinocytes develop skin papillomata; this epidermal proliferation can be suppressed with reduced Egfr signaling[23]. In contrast with several other genes now associated with inherited disorders with dysregulated RAS signaling, somatic *SOS1* mutations have not been reported in cancer.

The biochemical analysis of two NS-related SOS1 proteins revealed gain-of-function effects resulting in increased RAS activation. Since many of the *SOS1* mutations altered residues related to the autoinhibition of SOS, either through interaction of the histone folds with the PH-Rem linker or interaction of the DH domain at the allosteric RAS binding site, the predominant pathogenetic mechanism appears to be increased availability of the allosteric RAS binding site causing increased GEF activity. This increased GEF activity, in turn, increases RAS activation, which will lead to downstream signaling. SOS proteins also possess GEF activity through their DH and PH domains towards Rho GTPases such as RAC[15]. Aside from genetic evidence implicating RAS signaling in NS and related disorders, the positions of the *SOS1* mutations implicate RAS signaling specifically.

The two highly conserved *SOS* genes in vertebrates are widely expressed[24]. Sos1 and Sos2 bind a docking protein, Grb2, with different affinities[25] and Sos2 could not compensate for the loss of Sos1 in the *Sos1* knockout mice, suggesting that these proteins play unique roles. We examined the hypothesis that *SOS2* mutations similar to those in *SOS1* can also cause NS but failed to identify sequence changes at homologous positions.

Discovery of disease-causing mutations is challenging for genetic disorders presenting primarily sporadically or in small kindreds. Marked genetic heterogeneity further complicates such efforts due to statistical power issues. Our results provide proof of principal for high throughput resequencing with large cohorts, particularly when the relevant biological pathway can be identified. With three disease genes identified for NS but more than 40% of cases unexplained, future efforts will be directed towards exploiting this strategy in order to further advance diagnostics and prognostication for this disorder.

METHODS

High-Throughput Resequencing of *SOS1*. We assembled a cohort of 96 human subjects with NS from whom genomic DNAs were obtained from peripheral blood leukocytes. Nearly all subjects were Caucasian and of European ancestry, with the majority being Italian. The subjects did not harbor *PTPN11* or *KRAS* mutation based on scanning of the coding exons with DHPLC (Wave 2100 System, Transgenomic) and/or bidirectional DNA sequencing as previously described[2,5]. For sporadic cases, which represented the vast majority of the subjects, we obtained both parental DNAs whenever possible. All non-anonymous samples were collected under Institutional Review Board-approved protocols and with informed consent.

We chose a cohort of this size with the assumption that *SOS1* would account for at least 1% of NS (or 2% of *PTPN11*-/*KRAS*-negative NS). Based on Collins and Schwartz [26], this powered the study to detect a mutation in an NS gene at approximately 80% with α=0.05. If the gene accounted for 5% of *PTPN11*-/*KRAS*-negative NS, then the power to detect it with a cohort of this size would exceed 95%.

A high throughput approach to the resequencing of *SOS1* was performed at the Joint Genome Institute. The resequencing protocol was as follows: Oligonucleotide primers (sequences available upon request) for amplifying the *SOS1* coding exons (n=23) were designed to give a product size in the range of 200-700 bp with a minimum of 40 bp flanking the splice sites using the Exon Primer program (http://ihg.gsf.de/ihg/ExonPrimer.html), which is bundled with the UCSC Genome Browser (hg17 genome build: http://genome.ucsc.edu/). M13F and M13R tags were added to the forward and reverse primers, respectively. Five nanograms of genomic DNA from each NS sample was amplified in a 8 µl PCR reaction using AmpliTaq Gold (Applied

Biosystems) using PE 9700 machines and subsequently cleaned using a diluted version of the Exo-SAP based PCR product pre-sequencing kit (USB Corporation) dispensed by a nanoliter dispenser (Deerac Fluidics Equator). All PCR set-up procedures were performed in a 384-well format using a Biomek NX workstation following their optimization. Sequencing reactions were then performed using the M13 primers along with BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and cleaned with BET before separation on an ABI 3730xl DNA Analyzer. Base calling, quality assessment and assembly were carried out using the Phred, Phrap, Polyphred, Consed software suite (http://www.phrap.org/). All sequence variants identified were verified by manual inspection of the chromatograms and putative causative mutations were verified using another independent sequencing reaction.

When available, parental DNAs were then sequenced to establish whether the change was *de novo*. Paternity was confirmed by simple tandem repeat (STR) genotyping using the AmpF/STR Identifier PCR Amplification Kit (Applied Biosystems). Anonymous Caucasian control genomic DNAs were screened for *SOS1* coding exons in which putative mutations had been identified using DHPLC analysis of PCR-generated amplimers at column temperatures recommended by the Navigator version 1.5.4.23 software (oligonucleotide primer sequences and DHPLC conditions available upon request). Amplimers having abnormal denaturing profiles were purified (Microcon PCR, Millipore) and sequenced bi-directionally using the ABI BigDye terminator Sequencing Kit v.1.1 (Applied Biosystems) and an ABI Prism 310 Genetic Analyzer (Applied Biosystems). Eighty-five additional Caucasian control DNAs were digested with *Mne*I (New England Biolabs) or *Bsrs*I (Promega) to further exclude occurrence of the 1297G→A and 1649T→C missense changes, respectively. Informatics analysis of sequences to predict splice acceptor and donor sites as well as exonic splice enhancers was performed using programs

available at the following websites: http://www.cbs.dtu.dk/services/NetGene2, http://www.fruitfly.org/seqtools/splice.html, http://rulai.cshl.edu/ tools/ESE.

Analysis of the Second NS Cohort. We assembled a 2nd panel of 33 *PTPN11*-negative/*KRAS*-negative NS genomic DNAs. This panel was used as confirmatory of the results of the 1st panel and to extend the range of mutations associated with NS. These DNAs were scanned for *SOS1* mutations using DHPLC and abnormal amplimers were sequenced bi-directionally as described above.

RAS Activation Assay. GST-RAF-RBD fusion proteins were expressed in *Escherichia coli* by induction with 0.5 mM of isopropyl-1-thio-β-D-galactopyranoside (IPTG) for 5 hours. The expressed fusion proteins were isolated from bacteria lysates by affinity chromatography with glutathione agarose beads for 1 h at 4 °C. Cos-1 cells were co-transfected with HA-tagged RAS wild type and wild type (WT) or mutant SOS1. Twenty-four hours after transfection, cells were switched to serum-starvation medium (0% DMEM) for 16 h. Following stimulation with EGF (10 ng/ml) for the indicated intervals at 37 ºC, the cells were collected in RBD lysis buffer containing 25mM Tris-HCl (pH7.4), 120 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, 10 mg/ml pepstatin, 50 mM NaF, 1% aprotinin, 10 µg/ml leupeptin, 1 mM Na_3VO_4, 10 mM benzamidine, 10 µg/ml soybean trypsin inhibitor, 1% NP40, 0.25% sodium deoxycholic acid. For each condition, 400 µg of whole cell lysate was pre-cleared with 10 µl 50% GST for 5 min at 4 ^oC. The samples were then centrifuged and supernatants were transferred to Eppendorf tubes containing 20 µg GST-RAF-RBD immobilized beads. Samples were incubated for 1.5 h at 4 ^oC. The complexes were collected by centrifugation and washed six times with buffer containing 25

mM Tris-HCl (pH 7.4), 120 mM NaCl, 10 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, 50 mM NaF, 1% NP40. Protein complexes were eluted with SDS sample buffer, separated by SDS-12.5% PAGE, and transferred to nitrocellulose membrane. The proteins were detected by Western blot with anti-HA antibody (12CA5; 1:10,000) and goat anti-mouse HRP conjugated secondary antibody (Cappel; 1:10,000).

Acknowledgments

We are indebted to the patients and families who participated in the study, the physicians who referred the subjects, and the Joint Genome Institute's production sequencing group. This work was supported by Telethon-Italy grant GGP04172 and "Programma di Collaborazione Italia-USA/malattie rare" (M.T.), US National Institutes of Health Grants HL71207, HD01294, and HL074728 (B.D.G.), CA55360 and CA28146 (D.B.-S.), and Italian Ministry of Health Grant RC 2006 (B.D.). Research conducted at the E.O. Lawrence Berkeley National Laboratory and the Joint Genome Institute was performed under Department of Energy Contract DE-AC0378SF00098, University of California (LAP).

Table 1: Missense Variants

Exon	DNA Sequence Variant	Amino Acid Substitution	SOS1 Domain	Cohort (Observations)	Mut/Poly; Basis[1]
11	1294T→C	W432R	PH	A (1)	Mut; Ctrl
11	1297G→A	E433K	PH	A (1)	Mut; Ctrl
11	1649T→C	L550P	PH-Rem Linker	A (1)	Mut; Ctrl
11	1654A→G	R552G	PH-Rem Linker	A (4)	Mut; DN
11	1655G→A	R552K	PH-Rem Linker	A (1)	Mut; DN
11	1656G→C	R552S	PH-Rem Linker	A (1)	Mut; DN
13	1964C→T	P655L	Rem	A (1)	Poly
14	2104T→C	Y702H	Rem	A (1)	Mut; Ctrl
15	2186G→T	W729L	Rem	A (1)	Mut; DN
15	2197A→T	I733F	Rem	A (1)	Mut; DN
17	2536G→A	E846K	Cdc25	A (1)	Mut; Ctrl
24	3959A→G	H1320R	C-Term	A (1)	Poly
4	322G→A	E108K	HF	B (2)	Mut; Ctrl
7	806T→G	M269R	DH	B (1)	Mut; DN
11	1297G→A	E433K	PH	B (1)	Mut; Ctrl
11	1322G→A	C441Y	PH	B (1)	Mut; DN
11	1642A→C	S548R	PH-Rem Linker	B (2)	Mut; DN
11	1654A→G	R552G	PH-Rem Linker	B (1)	Mut; DN
11	1656G→C	R552S	PH-Rem Linker	B (1)	Mut; DN
19	2930A→T	Q977R	Cdc25	B (1)	Poly

[1]Ctrl- Controls; DN- *de novo*

Table 2: Genotype-Phenotype Correlation

Clinical Feature	No./Total (%) of Subjects		
	***SOS1* Mutation**	**All[a]**	**Without *PTPN11* Mutation[b]**
Polyhydramnios	8/15 (53)	43/130 (33)	NA
Fetal Macrosomia	9/15 (60)	NA	NA
Short Stature (<3rd centile)	2/15 (13)	84/115 (73)***	45/64 (70)***
Macrocephaly	9/16 (56)	19/151 (12)***	NA
Downslanting Palpebral Fissures	15/16 (94)	NA	NA
Ptosis	16/16 (100)	NA	NA
Low-Set Ears with Thickened Helix	16/16 (100)	NA	NA
Thick Lips/Macrostomia	14/16 (88)	NA	NA
Short/Webbed Neck	15/16 (94)	NA	NA
Abnormal Pectus	16/16 (100)	144/151 (95)	46/61 (75)*
Cardiac Involvement	13/16 (81)	132/151 (87)	42/66 (64)
Pulmonary Valve Stenosis	10/16 (62)	93/151 (62)	30/65 (46)
Septal Defect	4/16 (25)	29/151 (19)	11/63 (18)
HCM	2/16 (12)	30/151 (20)	17/65 (26)
Facial Keratosis Pilaris	8/16 (50)	21/151 (14)***	NA
Curly Hair	14/16 (88)	44/151 (29)***	NA
Cryptorchidism	6/9 (67)	64/83 (77)	25/35 (71)
Mental Retardation	1/16 (6)	32/105 (30)*	21/59 (36)*
Bleeding Diathesis	5/16 (31)	37/151 (25)	NA

[a]From Ref. [27]. [b]From Ref. 5. Significance: *, < .05; **, < .01; ***, < .001. Definitions: HCM, hypertrophic cardiomyopathy; NA, not available.

References

1. Noonan, J.A. Hypertelorism with Turner phenotype. A new syndrome with associated congenital heart disease. *Am J Dis Child* **116**, 373-80 (1968).
2. Carta, C. et al. Germline missense mutations affecting KRAS isoform B are associated with a severe Noonan syndrome phenotype. *Am J Hum Genet* **79**, 129-135 (2006).
3. Fragale, A., Tartaglia, M., Wu, J. & Gelb, B.D. Noonan syndrome-associated SHP2/PTPN11 mutants cause EGF-dependent prolonged GAB1 binding and sustained ERK2/MAPK1 activation. *Hum Mutat* **23**, 267-77 (2004).
4. Schubbert, S. et al. Germline KRAS mutations cause Noonan syndrome. *Nat Genet* **38**, 331-6 (2006).
5. Tartaglia, M. et al. PTPN11 mutations in Noonan syndrome: molecular spectrum, genotype-phenotype correlation, and phenotypic heterogeneity. *Am J Hum Genet* **70**, 1555-63 (2002).
6. Tartaglia, M. et al. Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat Genet* **29**, 465-8 (2001).
7. Keilhack, H., David, F.S., McGregor, M., Cantley, L.C. & Neel, B.G. Diverse biochemical properties of Shp2 mutants. Implications for disease phenotypes. *J Biol Chem* **280**, 30984-93 (2005).
8. Tartaglia, M. et al. Diversity and functional consequences of germline and somatic PTPN11 mutations in human disease. *Am J Hum Genet* **78**, 279-90 (2006).
9. Tartaglia, M. et al. Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. *Nat Genet* **34**, 148-50 (2003).

10. Digilio, M.C. et al. Grouping of Multiple-Lentigines/LEOPARD and Noonan Syndromes on the PTPN11 Gene. *Am J Hum Genet* **71**, 389-94 (2002).

11. Legius, E. et al. PTPN11 mutations in LEOPARD syndrome. *J Med Genet* **39**, 571-4 (2002).

12. Aoki, Y. et al. Germline mutations in HRAS proto-oncogene cause Costello syndrome. *Nat Genet* **37**, 1038-40 (2005).

13. Niihori, T. et al. Germline KRAS and BRAF mutations in cardio-facio-cutaneous syndrome. *Nat Genet* **38**, 294-6 (2006).

14. Rodriguez-Viciana, P. et al. Germline mutations in genes within the MAPK pathway cause cardio-facio-cutaneous syndrome. *Science* **311**, 1287-90 (2006).

15. Nimnual, A. & Bar-Sagi, D. The two hats of SOS. *Sci STKE* **2002**, PE36 (2002).

16. Corbalan-Garcia, S., Margarit, S.M., Galron, D., Yang, S.S. & Bar-Sagi, D. Regulation of Sos activity by intramolecular interactions. *Mol Cell Biol* **18**, 880-6 (1998).

17. Sondermann, H., Nagar, B., Bar-Sagi, D. & Kuriyan, J. Computational docking and solution x-ray scattering predict a membrane-interacting role for the histone domain of the Ras activator son of sevenless. *Proc Natl Acad Sci U S A* **102**, 16632-7 (2005).

18. Sondermann, H. et al. Structural analysis of autoinhibition in the Ras activator Son of sevenless. *Cell* **119**, 393-405 (2004).

19. Margarit, S.M. et al. Structural evidence for feedback activation by Ras.GTP of the Ras-specific nucleotide exchange factor SOS. *Cell* **112**, 685-95 (2003).

20. Hart, T.C. et al. A mutation in the SOS1 gene causes hereditary gingival fibromatosis type 1. *Am J Hum Genet* **70**, 943-54 (2002).

21. Qian, X. et al. The Sos1 and Sos2 Ras-specific exchange factors: differences in placental expression and signaling properties. *Embo J* **19**, 642-54 (2000).

22. Wang, D.Z. et al. Mutation in Sos1 dominantly enhances a weak allele of the EGFR, demonstrating a requirement for Sos1 in EGFR signaling and development. *Genes Dev* **11**, 309-20 (1997).

23. Sibilia, M. et al. The EGF receptor provides an essential survival signal for SOS-dependent skin tumor development. *Cell* **102**, 211-20 (2000).

24. Bowtell, D., Fu, P., Simon, M. & Senior, P. Identification of murine homologues of the Drosophila son of sevenless gene: potential activators of ras. *Proc Natl Acad Sci U S A* **89**, 6511-5 (1992).

25. Yang, S.S., Van Aelst, L. & Bar-Sagi, D. Differential interactions of human Sos1 and Sos2 with Grb2. *J Biol Chem* **270**, 18212-5 (1995).

26. Collins, J.S. & Schwartz, C.E. Detecting polymorphisms and mutations in candidate genes. *Am J Hum Genet* **71**, 1251-2 (2002).

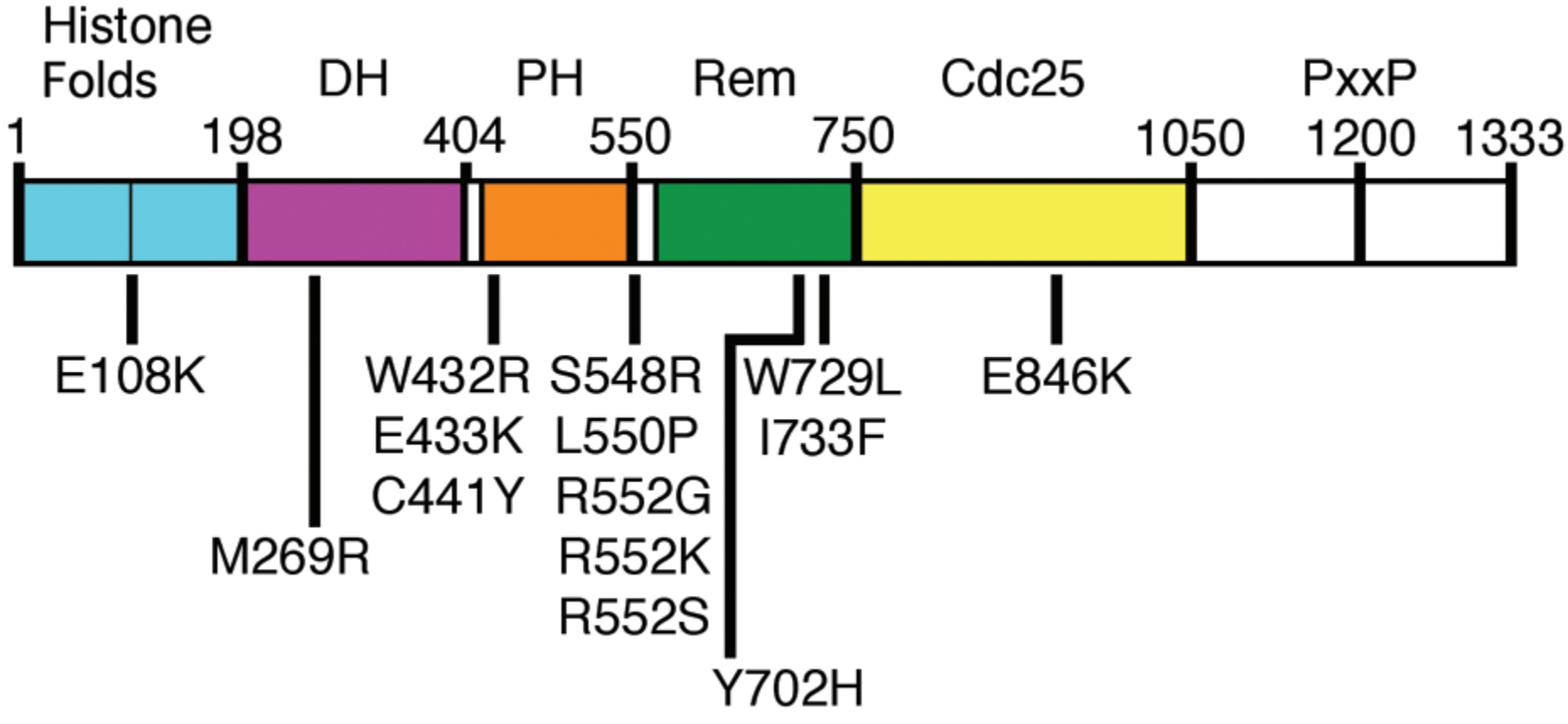
27. Sharland, M., Burch, M., McKenna, W.M. & Paton, M.A. A clinical study of Noonan syndrome. *Arch Dis Child* **67**, 178-83 (1992).

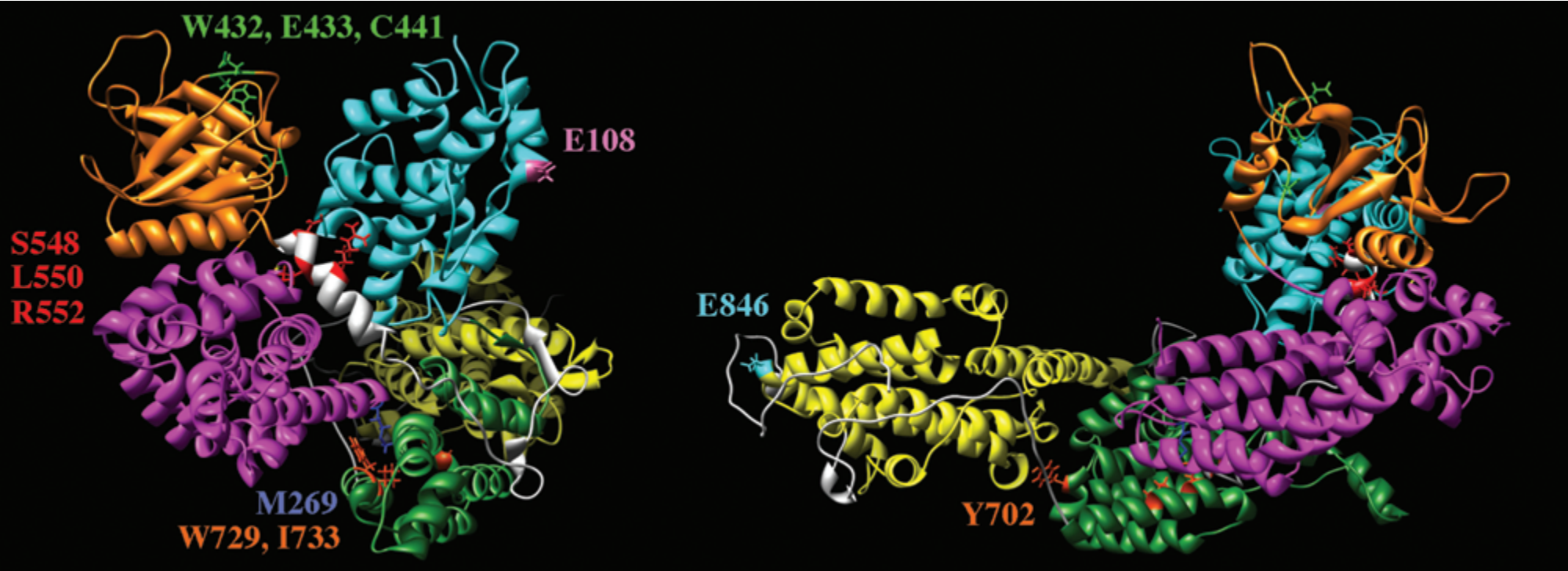
Figure Legends

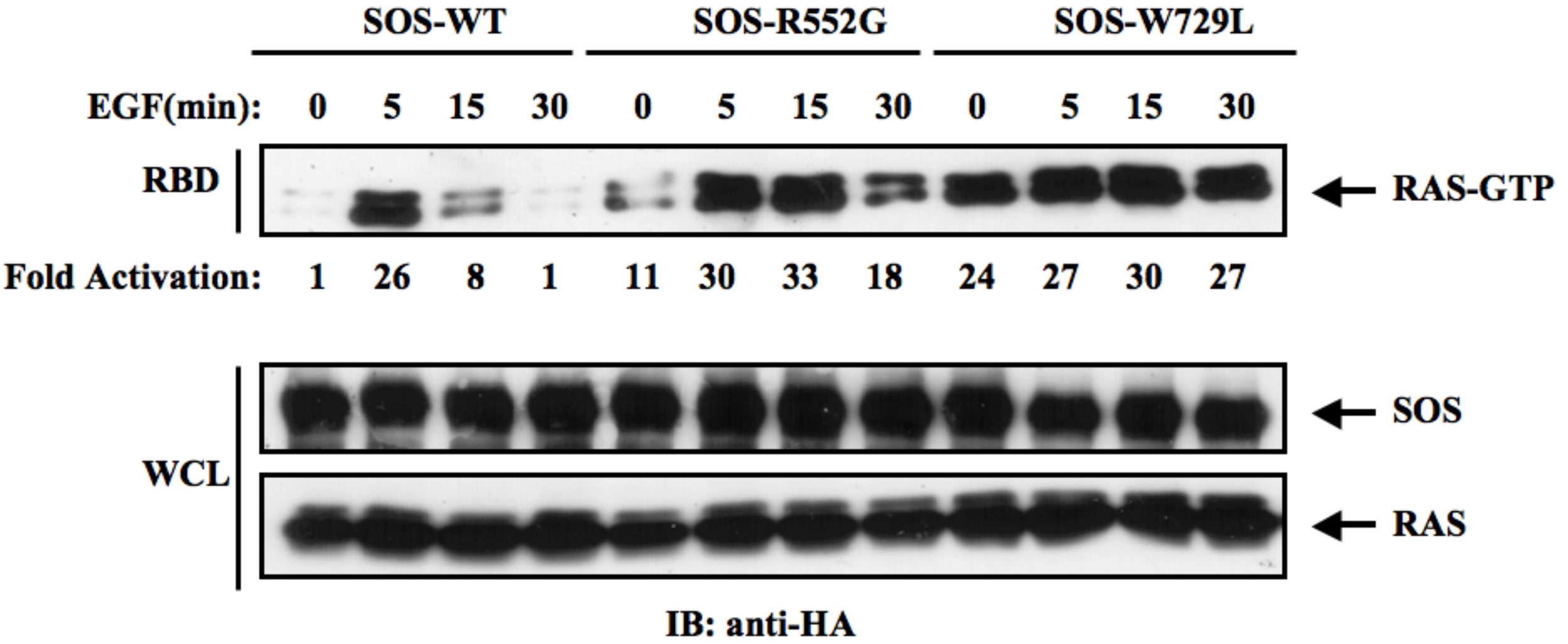
Figure 1 (a) SOS1 domain structure and location of residues altered in Noonan syndrome. The predicted amino acid substitutions from the 14 *SOS1* missense mutations are positioned below the cartoon of the SOS1 protein with its functional domains indicated above. Abbreviations: DH, Dbl homology domain; PH, plekstrin homology domain; Rem, RAS exchanger motif. **(b)** Location of the mutated residues on the three-dimensional structure of SOS1. The functional domains are color coded as follows: Histone folds, cyan; DH, magenta; PH, orange; Rem, green; Cdc25, yellow. Residues affected by mutations are indicated with their lateral chains and numbered. Based on Ref. 17, which utilized structural data and computational modeling.

Figure 2 RAS activation assay. Full-length, HA-tagged wild type (WT), R552G or W729L SOS1 were expressed in Cos-1 cells with HA-RAS. Binding of RAS to RAF-RBD was assayed to assess RAS activation in serum-starved cells (0 min) and after 5, 15 and 30 min of EGF stimulation. Total RAS and SOS1 proteins in the whole cell lysates (WCL), shown in the lower two panels, and activated RAS, upper panel, were detected immunologically with anti-HA. All fold activation ratios were compared to SOS-WT at 0 min.

Supplementary Figure 1 Facial dysmorphia in *SOS1*–associated Noonan syndrome. Photographs of 10 individuals with Noonan syndrome and the *SOS1* mutation indicated below each of them.


Histone
Folds
DH
PH
Rem
Cdc25
PxxP
1
198
404
550
750
1050
1200
1333
E108K
M269R
W432R
E433K
C441Y
S548R
L550P
R552G
R552K
R552S
Y702H
W729L
I733F
E846K


W432, E433, C441
E108
S548
L550
R552
M269
W729, I733
E846
Y702


SOS-WT
SOS-R552G
SOS-W729L
EGF(min): 0 5 15 30 0 5 15 30 0 5 15 30
RBD
RAS-GTP
Fold Activation: 1 26 8 1 11 30 33 18 24 27 30 27
WCL
SOS
RAS
IB: anti-HA

E108K

W432R

E433K

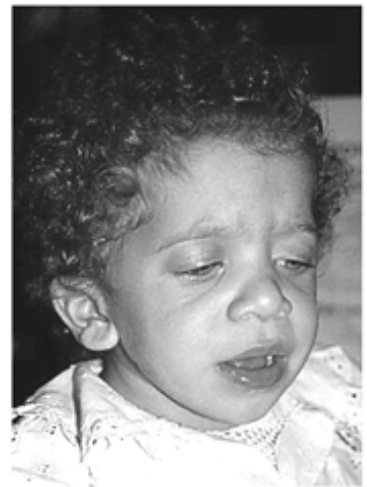
C441Y

S548R

R552G

R552G

R552K



R552S

I733F