The American Journal of Human Genetics, Volume *106*

Supplemental Data

De Novo SOX6 Variants Cause a Neurodevelopmental Syndrome Associated with ADHD, Craniosynostosis, and Osteochondromas

Dara Tolchin, Jessica P. Yeager, Priya Prasad, Naghmeh Dorrani, Alvaro Serrano Russi, Julian A. Martinez-Agosto, Abdul Haseeb, Marco Angelozzi, G.W.E. Santen, Claudia Ruivenkamp, Saadet Mercimek-Andrews, Christel Depienne, Alma Kuechler, Barbara Mikat, Hermann-Josef Ludecke, Frederic Bilan, Gwenael Le Guyader, Brigitte Gilbert-Dussardier, Boris Keren, Solveig Heide, Damien Haye, Hilde Van Esch, Liesbeth Keldermans, Damara Ortiz, Emily Lancaster, Ian D. Krantz, Bryan L. Krock, Kieran B. Pechter, Alexandre Arkader, Livija Medne, Elizabeth T. DeChene, Eduardo Calpena, Giada Melistaccio, Andrew O.M. Wilkie, Mohnish Suri, Nicola Foulds, Genomics England Research Consortium, Amber Begtrup, Lindsay B. Henderson, Cara Forster, Patrick Reed, Marie T. McDonald, Allyn McConkie-Rosell, Julien Thevenon, Pauline Le Tanno, Charles Coutton, Anne C.H. Tsai, Sarah Stewart, Ales Maver, Rudolf Gorazd, Olivier Pichon, Mathilde Nizon, Benjamin Cogné, Bertrand Isidor, Dominique Martin-Coignard, Radka Stoeva, Véronique Lefebvre, and Cédric Le Caignec

SUPPLEMENTAL MATERIALS

Supplemental Tables

Subject Numbers	Internal IDs	Referring Centers	Methods of diagnostic
#1	PIT-1	Pittsburgh	Agilent 180K Human Genome CGH + SNP ISCA
#2	PIT-2	Pittsburgh	Agilent 180K Human Genome CGH + SNP ISCA
#3	PIT-3	Pittsburgh	Agilent 180K Human Genome CGH + SNP ISCA
#4	CHUN-1	Nantes	Agilent 60K ISCA v.2
#5	CHLM-1	Le Mans	Agilent 60K ISCA v.2
#6	LJU-1	Ljubljana	Agilent 180K
#7	IHG-1	Essen	Affy CytoScan HD
#8	CHUP-1	Poitiers	Agilent 105K
#9	UK-1	Nottingham	WGS (Genomics England 100,000 Genomes Project)
#10	UK-2	Nottingham	WGS (Genomics England 100,000 Genomes Project)
#11	CHUN-2	Nantes	WES; Agilent SureSelect Human All Exome 50 Mb Kit
#12	GDX-3	Sickkids	WES; Agilent SureSelect Clinical Research Exome Kit
#13	GDX-1	UCLA	WES; Agilent SureSelect Clinical Research Exome Kit
#14	GDX-2	Duke	WES; Agilent SureSelect Clinical Research Exome Kit
#15	PS-1	Pitié	WES; Medexome Nimblegen 47 Mb Kit
#16	CHOP-1	Philadelphia	WES; Agilent SureSelectXT Clinical Exome version 1
#17	LEUV-1	Leuven	WES; Medexome Nimblegen, custom EZ choice XL V4
#18	LEID-1	Leiden	WES; Agilent SureSelect V5
#19	CHUG-1	Grenoble	WGS

Table S1. Referring Centers and Methods of Diagnostic for the 19 Affected Individuals.

Protein	Reference sequence
SRY	CAA37790.1
SOX1	NP_005977.2
SOX2	NP_003097.1
SOX3	NP_005625.2
SOX4	AAH72668.1
SOX5	AAH60773.1
SOX6	AAK26115.1
SOX7	CAC84226.1
SOX8	NP_055402.2
SOX9	CAA86598.1
SOX10	CAG38808.1
SOX11	BAA88122.1
SOX12	AAH67361.1
SOX13	AAD50120.1
SOX14	AAC95380.1
SOX15	AAH72003.1
SOX17	BAB83867.1
SOX18	BAA94874.1
SOX21	AAC95381.1
SOX30	BAA37146.1

Table S2. Accession Numbers for the Human SOX Sequences Aligned in Figure 4.

Vertebrate species	Scientific name	Reference sequence
Human	*Homo sapiens*	NM_033326.3
Mouse	*Mus musculus*	AAC52263.1
Dog	*Canis lupus familiaris*	XP_022263597.1
Cow	*Bos taurus*	NP_001178347.1
Armadillo	*Dasypus novemcinctus*	XP_023438566.1
Platypus	*Ornithorhynchus anatinus*	XP_028916408.1
Chicken	*Gallus gallus*	NP_001305380.1
Western clawed frog	*Xenopus tropicalis*	XP_017945061.1
Zebrafish	*Danio rerio*	NP_001116481.1

Table S3. Accession numbers of SOX6 ortholog sequences aligned in Figure S1.

Variant	Forward primer	Reverse primer
WT	CAAAAGAT**TGG**AAGGAAAAAATGGA	TTTCCTT**CCA**ATCTTTTGAAAGTAG
p.Trp161Cys	CAAAAGAT**TGC**AAGGAAAAAATGGA	TTTCCTT**GCA**ATCTTTTGAAAGTAG
WT	CGACCA**ATG**AATGCGTTCATGGTTTG	CGCATT**CAT**TGGTCGCTTGATGTGG
p.Met605Thr	CGACCA**ACG**AATGCGTTCATGGTTTG	CGCATT**CGT**TGGTCGCTTGATGTGG
WT	AGGATCTCGC**TGG**AAATCCATGTCCAACC	ATGGATTT**CCA**GCGAGATCCTAAGATTTTGC
p.Trp639Arg	AGGATCTCGC**CGG**AAATCCATGTCCAACC	ATGGATTT**CCG**GCGAGATCCTAAGATTTTGC
WT	GCACC**TCT**GCCAGCCCCGAGCCCA	GGCTGGC**AGA**GGTGCTAGAGCAGT
p.Ser746Leu	GCACC**TTG**GCCAGCCCCGAGCCCA	GGCTGGC**CAA**GGTGCTAGAGCAGT

Table S4. Primers Used to Generate Expression Plasmids for SOX6 Variants. Wild-type (WT) sequences are shown as references. Codons affected by variants are in bold font and underlined.

Affected individual	Body part	Region of the body part exhibiting endochondroma(s)
PIT-1	right arm	5th proximal phalanx distal 4th middle phalanx distal metaphysis of the 3rd metacarpal bone trapezoid bone
	left arm	proximal humerus distal 4th metacarpal bone metaphysis proximal end of 1st metacarpal bone
	left leg	proximal end of 1st metatarsal bone
	spine	right 12th vertebral body (transverse process)
PIT-2	right leg	distal femur metaphysis
	left leg	proximal fibula proximal phalanx of 4th toe
CHOP-1	right arm	proximal humerus*
	left arm	proximal humerus
	right leg	distal femur**
	left leg	distal tibia
	thorax	right rib*

Table S5. List of Endochondromas Found in Affected Individuals. *Osteochondromas no longer palpable, but still visible on X-rays two years later. **Osteochondromas no longer palpable and no longer visible on X-rays two years later.

Subject number	Internal identifier	Genomic variant (NM_033326.3)	Telomeric boundary	Centromeric boundary	Size of the deletion
#1	PIT-1	del ex. 1 to 2	16,358,927	16,497,834	138,907 bp
#2	PIT-2	del ex. 1 to 2	16,358,927	16,497,834	138,907 bp
#3	PIT-3	del ex. 1 to 2	16,358,927	16,497,834	138,907 bp
#4	CHUN-1	del ex. 1 to 4	16,238,166	16,554,120	315,954 bp
#5	CHLM-1	del ex. 2 to 3	16,309,520	16,419,313	109,793 bp
#6	LJU-1	del ex. 2 to 13	16,023,626	16,446,051	422,425 bp
#7	IHG-1	del ex. 2 to 12	16,049,440	16,399,572	350,132 bp
#8	CHUP-1	del ex. 5 to 7	16,128,647	16,239,865	111,218 bp

Table S6. Location, Boundaries and Size of the SNVs Affecting *SOX6* in Eight Study Subjects (hg19).

Supplemental Figures

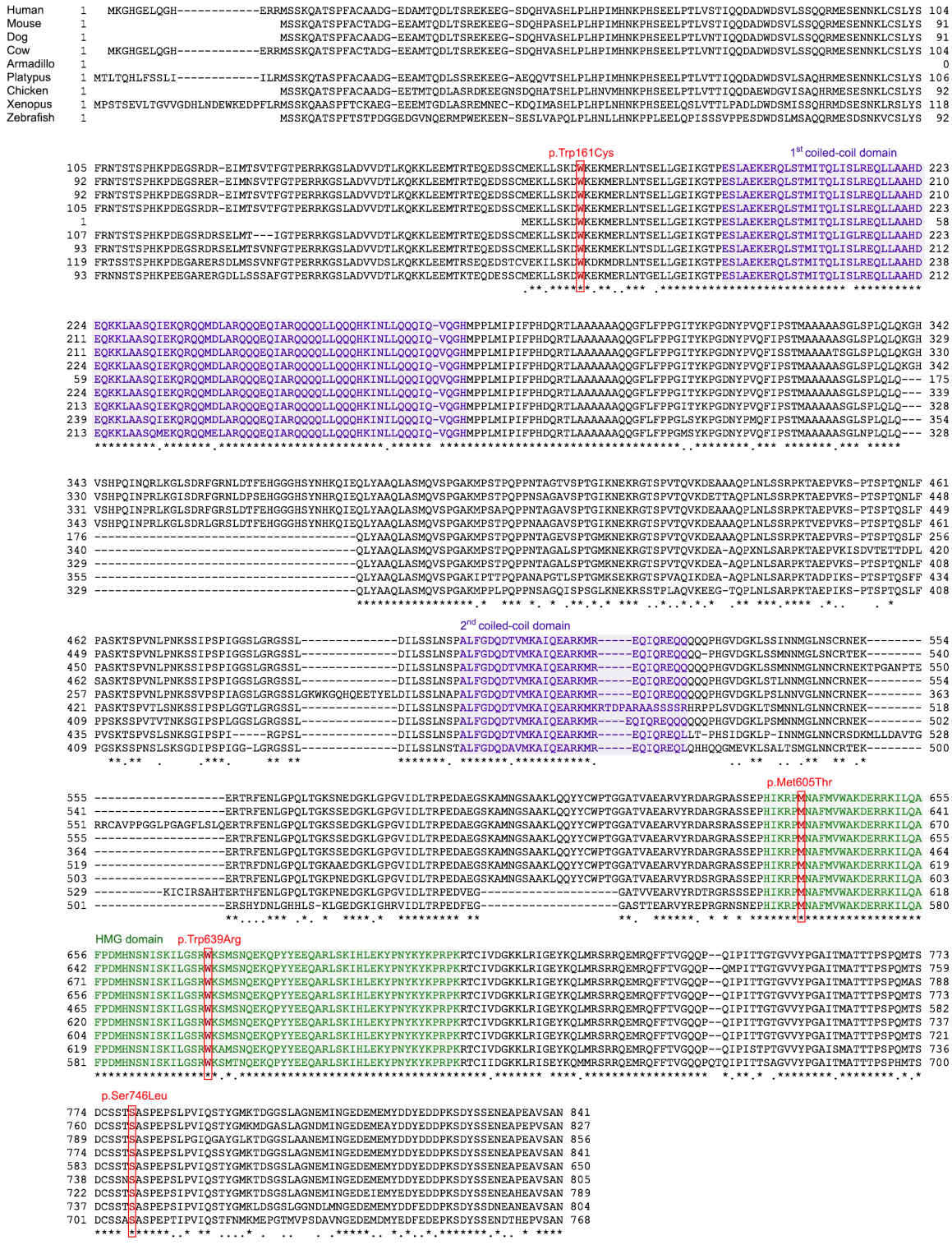


Figure S1. ClustalW Alignment of SOX6 Vertebrate Orthologs. The known functional domains of SOX6 are colored and labeled. Fully conserved residues are marked with asterisks and semi-conserved residues with dots underneath the sequences. Residues altered in study subjects are indicated in red.

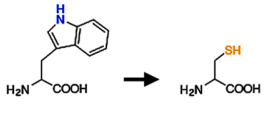
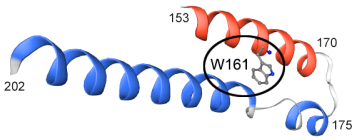
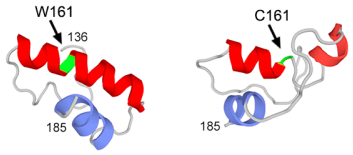
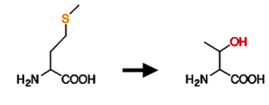
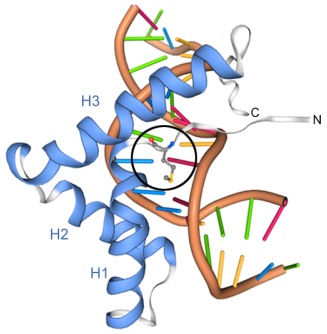
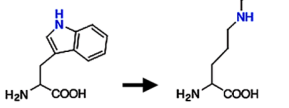
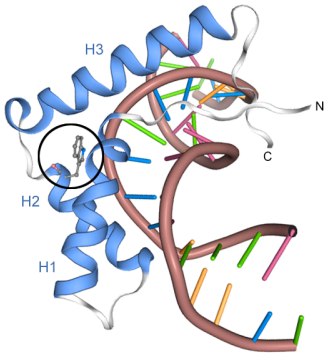
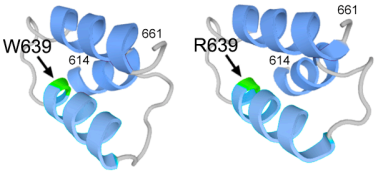
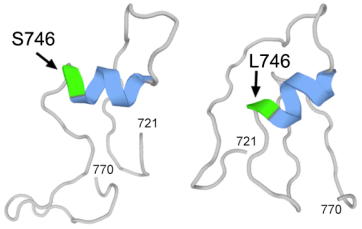
Variant	Amino acid structure	Swiss-Model	PEP-FOLD3
p.Trp161Cys			
p.Met605Thr			
p.Trp639Arg			
p.Ser746Leu			

Figure S2. Prediction of Structural Changes Imposed by Missense Variants. The first column identifies the missense variants detected in four study subjects. The second column shows schematics of wild-type and missense variant residues. The third column shows structural models proposed by Swiss-Model for the SOX6 regions containing three missense variants. For p.Met605Thr and p.Trp639Arg, the models correspond to the SOX5 HMG domain (with the α-helices H1, H2, and H3 in blue, and flanking sequences in grey) wrapped around the DNA helix (backbone in brown and nucleotide pairs colored). Altered amino acids are circled in black and their side chains represented. The last column shows structural models predicted by PEP-FOLD3 for the wild-type and mutant SOX6 regions that encompass missense variants. The wild-type and variant residues are indicated, as are the boundaries of the sequences used for modeling.

Genomics England Research Consortium

Ambrose J. C. [1], Bleda M. [1], Boardman-Pretty F. [1,2], Boissiere J. M. [1], Boustred C. R. [1], Caulfield M. J.[1,2], Chan G. C. [1], Craig C. E. H. [1], Daugherty L. C. [1], de Burca A. [1], Devereau, A. [1], Elgar G. [1,2], Foulger R. E. [1], Fowler T. [1], Furió-Tarí P. [1], Hackett J. M. [1], Halai D. [1], Holman J. E. [1], Hubbard T. J. P. [1], Kasperaviciute D. [1,2], Kayikci M. [1], Lahnstein L. [1], Lawson K. [1], Leigh S. E. A. [1], Leong I. U. S. [1], Lopez F. J. [1], Maleady-Crowe F. [1], Mason J. [1], McDonagh E. M. [1,2], Moutsianas L. [1,2], Mueller M. [1,2], Need A. C. [1,2], Odhams C. A. [1], Patch C. [1,2], Perez-Gil D. [1], Polychronopoulos D. [1], Pullinger J. [1], Rahim T. [1], Rendon A. [1], Rogers T. [1], Ryten M. [1], Savage K. [1], Scott R. H. [1], Siddiq A. [1], Sieghart A. [1], Smedley D. [1,2], Smith K. R. [1,2], Sosinsky A. [1,2], Spooner W. [1], Stevens H. E. [1], Stuckey A. [1], Thomas E. R. A. [1,2], Thompson S. R. [1], Tregidgo C. [1], Tucci A. [1,2], Walsh E. [1], Watters, S. A. [1], Welland M. J. [1], Williams E. [1], Witkowska K. [1,2], Wood S. M. [1,2], Zarowiecki M.[1].

1. Genomics England, London, UK

2. William Harvey Research Institute, Queen Mary University of London, London, EC1M 6BQ, UK.