Mutations in *SMG9*, Encoding an Essential Component of Nonsense-Mediated Decay Machinery, Cause a Multiple Congenital Anomaly Syndrome in Humans and Mice

Ranad Shaheen,[1,7] Shams Anazi,[1,7] Tawfeg Ben-Omran,[2,7] Mohammed Zain Seidahmed,[3] L. Brianna Caddle,[5] Kristina Palmer,[5] Rehab Ali,[2] Tarfa Alshidi,[1] Samya Hagos,[1] Leslie Goodwin,[5] Mais Hashem,[1] Salma M. Wakil,[1] Mohamed Abouelhoda,[1] Dilek Colak,[4] Stephen A. Murray,[5,*] and Fowzan S. Alkuraya[1,6,*]

Nonsense-mediated decay (NMD) is an important process that is best known for degrading transcripts that contain premature stop codons (PTCs) to mitigate their potentially harmful consequences, although its regulatory role encompasses other classes of transcripts as well. Despite the critical role of NMD at the cellular level, our knowledge about the consequences of deficiency of its components at the organismal level is largely limited to model organisms. In this study, we report two consanguineous families in which a similar pattern of congenital anomalies was found to be most likely caused by homozygous loss-of-function mutations in *SMG9*, encoding an essential component of the SURF complex that generates phospho-UPF1, the single most important step in NMD. By knocking out *Smg9* in mice via CRISPR/Cas9, we were able to recapitulate the major features of the *SMG9*-related multiple congenital anomaly syndrome we observed in humans. Surprisingly, human cells devoid of SMG9 do not appear to have reduction of PTC-containing transcripts but do display global transcriptional dysregulation. We conclude that SMG9 is required for normal human and murine development, most likely through a transcriptional regulatory role, the precise nature of which remains to be determined.

Introduction

The field of transcriptional regulation has grown rapidly in the recent past, and completely novel cellular mechanisms, e.g., control of transcript levels by other RNA species, have emerged [1]. Transcripts destined for translation have the potential to encode harmful peptides if the reading frame is altered, a mechanism that underpins the pathogenesis of many human genetic diseases[2,3]. Fortunately, the transcriptional regulatory machinery has the capacity to address at least some of these errors. One noteworthy process in this regard is nonsense-mediated decay (NMD). Through NMD, transcripts that contain a premature stop codon (PTC) are degraded such that the cell is protected from a potential dominant-negative effect or toxic gain of function exerted by the truncated protein[4–6].

NMD is a complex process that involves RNA-protein and protein-protein interactions, and it has been observed in all eukaryotes studied to date[6]. The current model suggests that exon-exon junctions are flagged by a UPF2-UPF3B complex (exon junction complex [EJC])[7,8]. If this complex is encountered downstream of a stop codon, this will trigger a series of events. The single most critical step in this process involves the phosphorylation of UPF1 by SMG1, a process that requires the dual action of two ancillary proteins (SMG8 and SMG9) in the SURF (SMG1, UPF1, and the eukaryotic release factors 1 and 3 [eRF1 and eRF3]) complex, which assembles on ribosomes that encounter a PTC[9]. Upon its phosphorylation, UPF1 releases eRF1 and eRF3, recruits SMG5, SMG6, and SMG7 and halts translation. SMG5–SMG7-mediated exonucleolytic decay and SMG6-mediated endonucleolytic decay lead to degradation of the PTC-containing transcript[10–15]. It is worth highlighting that NMD also operates in a partially redundant EJC-independent manner[15,16].

Importantly, NMD also acts on "physiological" transcripts, and at least 10% of human transcripts are subject to NMD regulation[17–19]. These transcripts are not limited to those harboring PTCs introduced normally through alternative splicing, but also include transcripts with a multitude of other features, including those in which the normal stop codon is separated from the poly(A)-binding protein by a long 3′ UTR, those with a 3′ UTR that spans an intron, and those with an upstream open reading frame[20–23].

The consequences of deficiency of NMD components in humans in the developmental context are largely unknown, with the exception of *UPF3B* (MIM: 300298), loss-of-function mutations of which result in intellectual disability (MIM: 300676) [23]. In this study, we show that deficiency of another NMD component, SMG9, although compatible with embryonic viability, is associated with a number of major malformations.

[1]Department of Genetics, King Faisal Specialist Hospital and Research Center, Riyadh 11211, Saudi Arabia; [2]Department of Genetics, Hamad Medical Corporation, Doha, Qatar; [3]Department of Pediatrics, Security Forces Hospital, Riyadh 12625, Saudi Arabia; [4]Department of Biostatistics, Epidemiology, and Scientific Computing, King Faisal Specialist Hospital and Research Center, Riyadh 11211, Saudi Arabia; [5]The Jackson Laboratory, Bar Harbor, ME 04609, USA; [6]Department of Anatomy and Cell Biology, College of Medicine, Alfaisal University, Riyadh 11533, Saudi Arabia
[7]These authors contributed equally to this work
*Correspondence: steve.murray@jax.org (S.A.M.), falkuraya@kfshrc.edu.sa (F.S.A.)
http://dx.doi.org/10.1016/j.ajhg.2016.02.010.

Materials and Methods

Human Subjects

Affected individuals were phenotyped via standard clinical evaluation, including appropriate imaging and laboratory tests. All available family members were recruited after they gave informed consent and were enrolled in an institutional review board (IRB)-approved research protocol (KFSHRC RAC no. 2080006). Blood was collected in EDTA tubes for DNA extraction and in Na-heparin tubes for the establishment of lymphoblastoid cell lines.

Autozygome Mapping and Linkage Analysis

We determined the entire set of autozygosity intervals per genome (autozygome) as described before[24,25]. In brief, runs of homozygosity (ROH) $\geq$2 Mb in size, used as surrogates of autozygosity, were determined from genome-wide SNP genotypes obtained on the Axiom SNP Chip platform (Affymetrix) via AutoSNPa[26]. We performed linkage analysis with the GeneHunter multipoint linkage analysis algorithm by employing a fully penetrant autosomal-recessive model[27].

Exome Sequencing

Exome capture was performed with a TruSeq Exome Enrichment kit (Illumina) according to the manufacturer's protocol. Samples were prepared as an Illumina sequencing library, and in the second step, the sequencing libraries were enriched for the desired target with the Illumina Exome Enrichment protocol. The captured libraries were sequenced with an Illumina HiSeq 2000 Sequencer. The reads were mapped by the Burrows-Wheeler Aligner against the UCSC Genome Browser hg19 reference genome. The SNPs and indels were detected by SAMtools. Variants from whole-exome sequencing were filtered such that only novel or very low frequency (0.1%) coding or splicing homozygous variants within the shared autozygome of the three affected individuals were considered as likely causal variants. The frequency of variants was determined via publicly available variant databases (1000 Genomes, the Exome Variant Server, and the Exome Aggregation Consortium [ExAC] Browser), as well as a database of 734 in-house ethnically-matched exomes.

Real-Time RT-PCR and Immunoblotting

For RT-PCR and relative qRT-PCR, total RNA from affected-individual and control-individual lymphoblastoid cell lines was extracted with the QIAamp RNA Mini Kit (QIAGEN), and DNase was treated by the RNase-Free DNase Set (QIAGEN), according to the manufacturer's recommendations. Preparation of the cDNA was carried out with the iScriptTM cDNA synthesis kit and Poly T oligonucleotide primers (Applied Biosystems). Two sets of primers for *SMG9* (MIM: 613176) cDNA were designed to specifically amplify the cDNA (*SMG9* exons 5–7 and 5–8). Relative qRT-PCR for the expression of *SMG9*, *VIM* (MIM: 193060), *TNS3* (MIM: 606825), *EGR1* (MIM: 128990), *UCHL1* (MIM: 191342), *SPINT2* (MIM: 605124), *RORA* (MIM: 600825), and *VCAN* (MIM: 118661) was performed with SYBR Green and an Applied Biosystems 7500 Fast Real-Time PCR System (Table S3).

For immunoblotting, total protein was extracted from one affected-individual and three control-individual lymphoblastoid cells. Anti-SMG9 antibody was purchased from Sigma Aldrich (SAB2107730). The membrane was blocked with 5% milk powder in PBS with Tween for 1 hr at room temperature (RT) and incubated with the primary antibody (dilution 1/500) overnight at 4°C, followed by stringency washes and treatment with secondary antibody for signal detection.

Global Transcriptional Profiling and Statistical Analysis

Lymphoblastoid cells from one affected individual and three healthy control individuals (two females and one male, age range 2–6 years) were used for global expression profiling, performed in triplicate for each sample, via Affymetrix's GeneChip Human Genome U133 Plus 2.0 Arrays. Sample handling, cDNA synthesis, cRNA labeling and synthesis, hybridization, washing and scanning of chips, and all related quality controls were performed according to the manufacturer's instructions. In brief, a total of 200 ng of RNA was reverse transcribed to synthesize the first strand of cDNA with the help of oligo-dT primers containing T7 flanking sequence. After second-strand synthesis, in vitro transcription was carried out with labeled aRNA. After purification and fragmentation of aRNA, the samples were hybridized overnight onto U133 Plus 2.0 Arrays. Finally, after being washed to remove the unbound transcripts, the hybridized microarrays were scanned and the intensity (CEL) files with the acquisition and initial quantification of array images were generated with Expression Console 1.3. Significantly modulated genes were defined as those with absolute fold change (FC) > 2 and adjusted p value < 0.05. We used the Benjamini-Hochberg step-up procedure to control the false discovery rate (FDR) at 5% [28]. Functional pathway, gene ontology (GO), and network analyses were performed with Ingenuity Pathways Analysis (IPA) 6.3 (Ingenuity Systems) and DAVID Bioinformatics Resources[29]. A right-tailed Fisher's exact test was used to calculate a p value determining the probability that the biological function (or pathway) assigned to that dataset is explained by chance alone. For GO analysis, we used the background list of genes expressed in lymphoblastoid cell lines that was obtained from The Gene Expression Barcode project. Statistical analyses were performed with the MATLAB software packages (Mathworks) and Partek Genomics Suite. P values were calculated via two-sided tests. P values of less than 0.05 were considered statistically significant.

Generation of *Smg9*$^{-/-}$ Mice

A null mutation was generated in mouse *Smg9* via CRISPR/Cas9 mutagenesis. A guide RNA was designed against exon 2 (first coding exon) with the online tool CRISPR Design, and a single guide with minimal off-target potential and sequence TCTACGGGATA GAGCGGCGG was selected. Production of sgRNA and microinjection was performed as previously described[30]. In brief, an oligonucleotide that includes a T7 promoter and the guide sequence above is annealed to a reverse primer that includes the entire stem-loop sequence and is amplified by PCR. The resulting PCR products were in vitro transcribed and purified, and the quality of the RNA products was confirmed with an Agilent Bioanalyzer 2100.

All animal procedures were conducted according to relevant national and international guidelines (AALAC and IACUC) and have been approved by The Jackson Laboratory animal care and use committee (protocol no. 99066). C57BL/6NJ and CB6F1/J mouse strains were used as embryo donors and pseudopregnant recipient dams, respectively. Cas9 mRNA (100 ng/μl) (TriLink Biotechnologies) and sgRNA (50 ng/μl) were injected into the pronuclei of zygotes. Oviduct transfers into pseudopregnant dams were performed on the same day. Founder mice were screened by PCR and Sanger sequencing, identifying several putative mutants. A single allele was chosen for expansion and analysis, which was

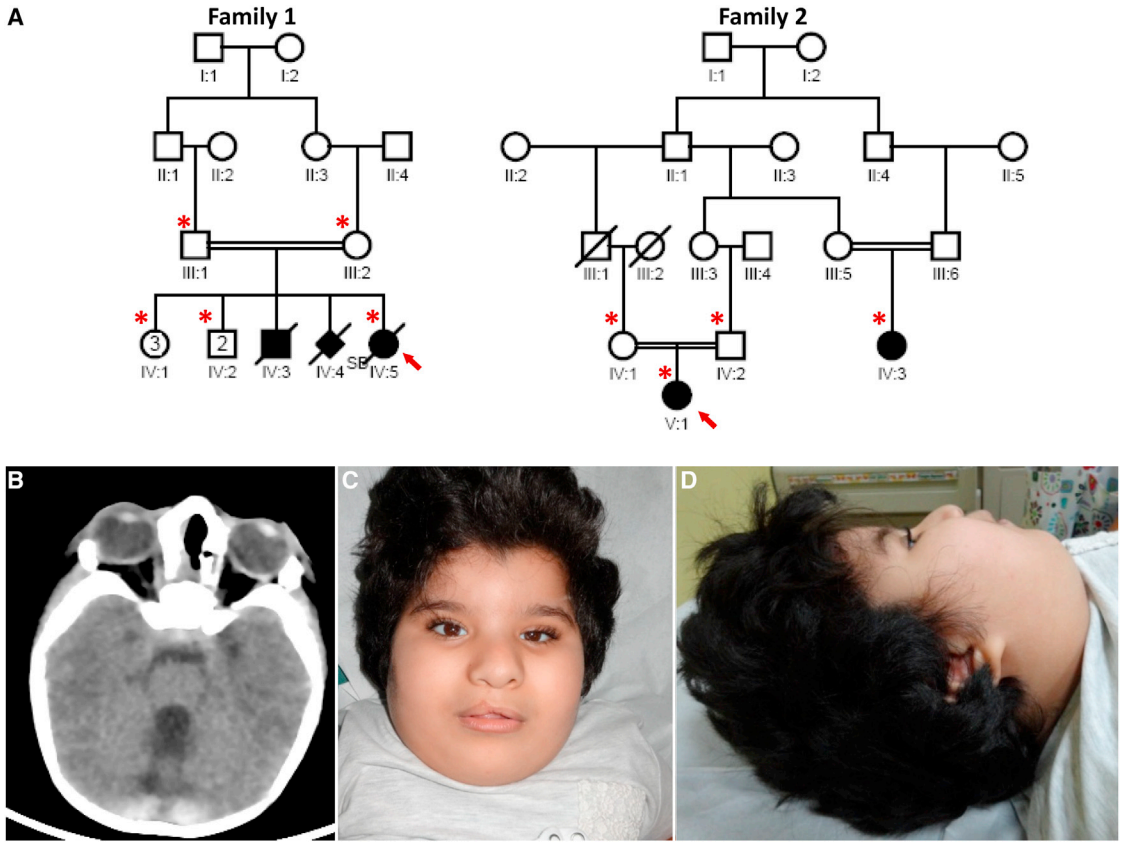


Figure 1. Identification of an Autosomal-Recessive Multiple Congenital Anomalies Syndrome
(A) Pedigrees of the two study families showing the consanguineous nature of the parents. The index individual is indicated in each pedigree by an arrow, and asterisks denote individuals whose DNA was available for analysis.
(B) CT scan for the index individual in family 1 (IV:5) showing cerebellar vermis hypoplasia.
(C and D) Facial images of the index individual in family 2 (V:1) showing a right-sided cleft lip (status post repair), a full and everted lower lip, a widow's peak, hypertelorism, small eyes, posteriorly rotated ears with attached lobules, and a broad nasal bridge.

confirmed by TOPO TA cloning and sequencing (Life Technologies). Genotyping protocols and mice are available at The Jackson Laboratory mouse strain datasheet.

To produce homozygous embryos, timed matings were performed, where day "0" was assumed to be the midpoint of the prior light cycle after the appearance of a copulation plug. Dissections were performed at the time indicated. Embryos were photographed, fixed in 4% paraformaldehyde (24 hr for embryonic day 15.5 [E15.5] and 48 hr for E18.5), stabilized in an acrylamide-based hydrogel[31], and stained in Lugol's solution (Sigma catalog no. L6146-1L) for 24 hr (E15.5) or 48 hr (E18.5). Samples were then rinsed with 1× PBS and embedded in 1% agarose for imaging.

Imaging was conducted on a Bruker Skyscan 1172 X-ray scanner (Bruker BioSpin). Images were reconstructed with NRecon software (Bruker BioSpin) and surface rendered with Imaris software (Bitplane).

For the qRT-PCR of *Smg9*, RNA was purified from E15 embryonic tissue. Isolated RNA was subject to DNA degradation. cDNA was generated with a High Capacity RNA-to-cDNA Kit (Applied Biosystems), followed by RNA degradation. A TaqMan (Applied Biosystems) assay was designed to amplify cDNA for *Smg9* transcripts. Primers were designed to span the exon 2 and exon 3 junction with the probe lying directly over the exon 2 and exon 3 junction site to eliminate genomic DNA amplification. *Smg9* qPCR was performed with Applied Biosystems TaqMan reagents, and *Gapdh* was used as the internal control, multiplexed to optimize accuracy. TaqMan assay was amplified via conditions specified by Applied Biosystems. We used a reaction volume of 10 μl containing 40 ng cDNA. Thermocycler conditions included an initial hold at 50° for 2 min, polymerase activation hold at 95° for 20 s, 40 cycles of 95°, a 1 s denaturation, and a 20 s 60° anneal and extension time. The analysis was performed with Viia7 analysis software. Samples were run in triplicate. Wild-type expression was calculated with n = 1. Each biological replicate is represented on the bar graphs.

Results

Clinical Report

The proband in family 1(IV:5) is a Saudi Arabian female delivered at term by elective cesarean section to a 33-year-old gravida 12, para 8 mother with history of three abortions. The parents are first-degree cousins (Figure 1A). One sibling died at the age of one year and had multiple congenital anomalies similar to the proband's, as well as a cleft palate (no records available). There is history of a stillborn girl who also had a similar pattern of multiple congenital anomalies (no records available). She was found to have craniofacial dysmorphism, microphthalmia, and major brain and heart malformations and died at 7 weeks

Table 1. Clinical Details of the Three Affected Individuals Recruited in this Study

	Category	Subcategory	Features
Family 1, IV:5	family history	–	parents are 1st degree cousins, one sibling who had died at the age of one year with a similar disease
	head and neck	head	prominent forehead and occiput
		ears	low set malformed ears
		eyes	microphthalmia
		nose	depressed nasal bridge and anteverted nares
		mouth	high arched palate
	skeletal	hands	clenched hands with camptodactyly
	neurologic	CNS	Dandy-Walker malformation, cerebellar vermis hypoplasia, hypoplastic corpus callosum
	cardiovascular	heart	interrupted aortic arch, hypoplastic tricuspid and aortic valves, large muscular VSD
Family 2, V:1	family history	–	parents are 1st degree cousins, one similarly affected cousin
	head and neck	head	narrow forehead
		ears	posteriorly rotated ears with attached lobules
		eyes	hypertelorism and poor vision
		nose	broad nasal bridge
		mouth	full and everted lower lip, right-sided cleft lip
	skeletal	hands	syndactyly between 2nd and 3rd toes
	neurologic	–	Dandy-Walker malformation, decreased myelination, brain atrophy
	cardiovascular	heart	large VSD
Family 2, IV:3	head and neck	head	NA
	skeletal	hands	NA
	neurologic	–	truncal hypotonia with peripheral hypertonia, brisk deep tendon reflexes, adductor spasm in both lower limbs
	cardiovascular	heart	VSD

Abbreviations are as follows: NA, not available; VSD, ventricular septal defect.

of age. Family 2 consists of first-cousin healthy parents of Qatari origin with one child (index, V:1) and a first cousin (IV:3) who was also found to have a similar phenotype in the form of craniofacial dysmorphism, congenital heart disease, and brain malformation (Figure 1A). For clinical details of the three affected individuals, see the Supplemental Note and Table 1.

A Multiple Congenital Anomaly Syndrome Is Linked to Loss-of-Function Mutations in *SMG9*

The phenotypic overlap (especially the Dandy-Walker malformation and congenital heart disease) between the affected members of both families, their consanguineous nature, and different geographic origin in Arabia prompted us to consider the possibility that the disease is caused by homozygosity for two distinct founder haplotypes (one in family 1 and another in family 2) that harbor pathogenic variants in the same gene. Indeed, the autozygome of the three affected individuals overlapped on a single locus in which the two families displayed two different haplotypes, and this was further confirmed by linkage analysis (Figure 2A and Figure S1). To identify the likely causal variant within this critical locus, we performed exome sequencing on each of the two probands and achieved 80× mean depth of the target region (target was 52 Mb), and 97% of target regions were covered with at least 10× depth. We only considered homozygous coding and splicing novel variants within the critical locus (Figure S2). Only one variant per exome remained (c.520_521delCC and c.701+4A>G) according to these filters, both affecting the same gene, *SMG9* (GenBank: NM_019108.2). Sanger sequencing confirmed homozygosity for these mutations in the probands and heterozygosity in the parents (Figure 2B). The indel in family 1 predicts a frameshift and premature truncation, p.Pro174Argfs*12. Splice Site Prediction by Neural Network indicates that the intronic variant in family 2 reduces the score of the consensus donor site of exon 6 from 0.98 to 0.65. However, RT-PCR on lymphoblastoid cells from the index individual in family 2 revealed complete skipping of exon 6. The resulting aberrant transcript that predicts frameshift and premature truncation (p.Tyr197Aspfs*10) was the only

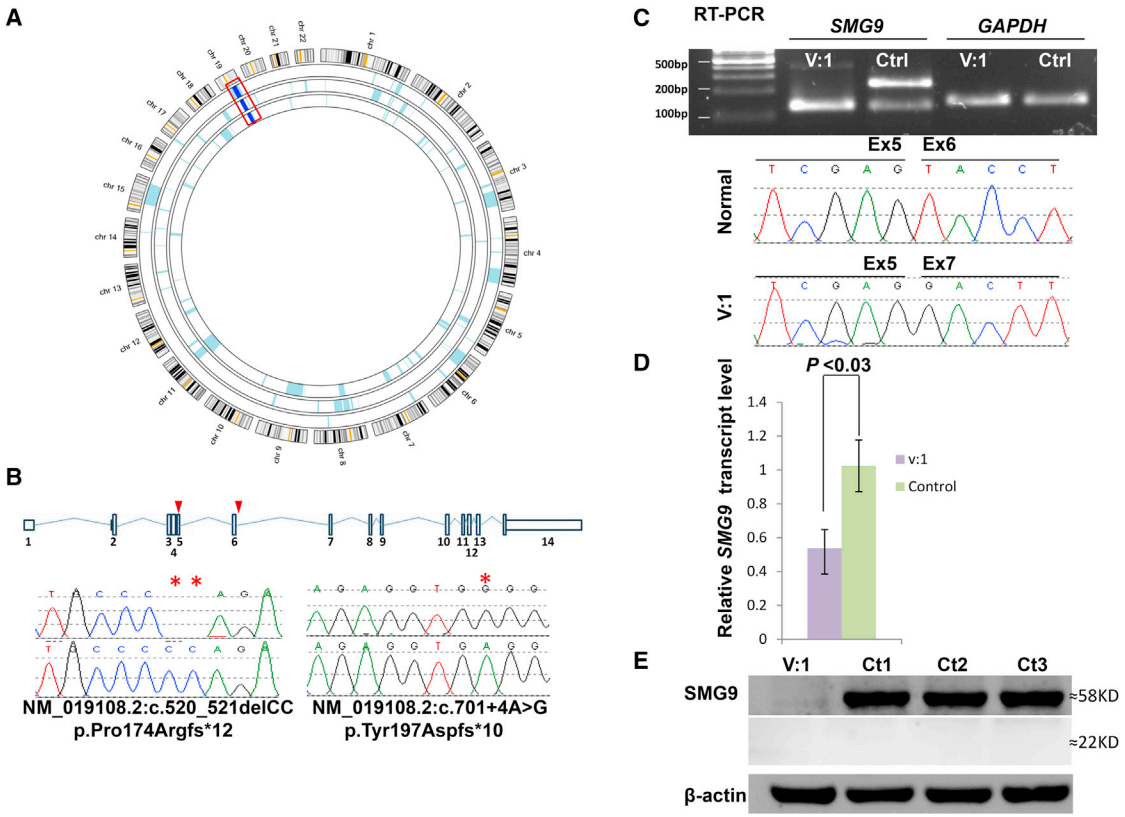


Figure 2. A Multiple Congenital Anomalies Syndrome Is Linked to *SMG9* Mutations
(A) AgileMultiIdeogram showing a single homozygous region, on chromosome 19, shared between the three affected individuals from the two families as indicated by the dark blue bars boxed in red.
(B) Upper panel: schematic representation of *SMG9* transcript (red triangles indicate the sites of the two mutations). Lower panel: sequence chromatograms of two variants in *SMG9* identified in the family (control tracing is shown for comparison and the location of the mutation is denoted by red stars).
(C) RT-PCR gel image and sequence chromatogram show the enhancement of aberrant *SMG9* transcript in lymphoblasts derived from affected individual (Aff) as compared with lymphoblasts from a healthy individual of similar age (control), and the absence of the normal *SMG9* transcript.
(D) Relative quantitative real-time PCR showing 50% reduction in *SMG9* expression in the affected individual in comparison to expression in the control individual. Result is the average for three independent experiments and three control individuals ($p < 0.03$ on t test).
(E) Immunoblotting with antibody against the N terminal of SMG9 (SAB2107730) showing no detectable band from cells derived from the index individual (Aff) (target mass KDa, 57.7, 54.8) as compared with those derived from the three normal control individuals (Ct1, Ct2, Ct3), as well as the absence of the truncated SMG9 isoform at 22 KDa.

recovered transcript in the affected individual, unlike control individuals, who possessed this band at a very low abundance in comparison to the normal exon-6-containing transcript (Figure 2C). Western blot analysis revealed complete lack of SMG9, and the predicted truncated isoform could not be recovered, indicating NMD and/or instability (Figure 2E and Figure S3). Indeed, qRT-PCR showed 50% reduction of the *SMG9* transcript, as compared to control individuals ($p < 0.03$) (Figure 2D). Taken together, our results are consistent with SMG9 deficiency as the likely cause of the multiple congenital anomaly syndrome we observed in the two probands.

No Evidence of Widespread Perturbation of NMD in SMG9 Deficiency

SMG9 has been found to be necessary for SMG1-mediated phosphorylation of UPF1, the single most important step in NMD, so we were surprised by the finding that the index individual in family 2 (V:1), who appears to have complete deficiency of SMG9, shows evidence of NMD of the PTC-containing *SMG9* transcript (Figure 2D). Therefore, we set out to test the entire transcriptional profile of the affected individual and compare it to that of control individuals for whom exome sequencing data were available. We first looked for evidence in control individuals that transcripts with PTC are significantly lower than other transcripts in general and found that, despite extreme variability, such was the case ($p < 0.01$). We then repeated the same test on the affected individual and found that the same pattern was observed (Figure S4 and Table S1). Comparing the expression of 27 common PTC-containing genes in control individuals and in the affected individual revealed no significant difference (p value > 0.05) (Figure S5 and Table S2). However, as shown in Figure 3, the global transcriptional profile of the affected individual and control individuals

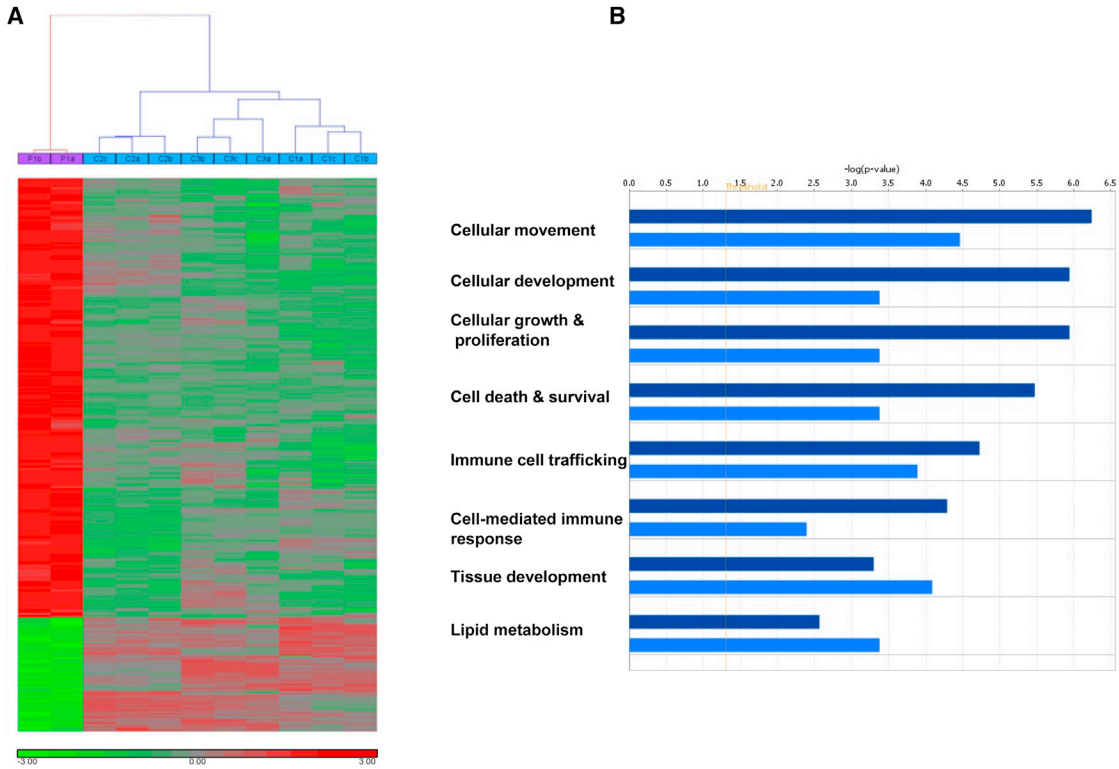


Figure 3. Global Transcriptional Changes Associated with SMG9 Deficiency
(A) Heat-map of genes that were significantly dysregulated in a SMG9-deficient individual in comparison to three control individuals. The hierarchical clustering clearly distinguished individuals as either affected or controls. The expression level of each gene across the samples is scaled to [-3, 3] interval. These mapped expression levels are depicted with a color scale as shown at the bottom of the figure, such that highly expressed genes are indicated in red, intermediate in black, and weakly expressed in green.
(B) GO and functional analysis of DE genes (up- or downregulated). x axis indicates the significance (−log p value) of the functional association that is dependent on the number of genes in a class as well as biological relevance. Dark bars represent upregulated genes, and light bars represent downregulated genes. The threshold line represents a p value of 0.05.

revealed a pattern of gene dysregulation that clearly distinguished individuals as either affected or controls (Figure 3A). We observed a prevalent upregulation of gene expression as a result of SMG9 deficiency. Indeed, with a stringent statistical threshold (FDR of 5% and FC $>$ 2), 300 probes were differentially expressed (DE), of which 238 probes (171 genes) were upregulated and 62 probes (54 genes) were downregulated, when comparing expression in the affected individual versus that in the control individuals (Figure 3A and Table S1). Validation was carried out on selected genes taken from the most dysregulated ones by quantitative real-time PCR, and the results were in agreement with the microarray (Figure S6). The GO and functional analyses of upregulated DE genes showed an over-representation of genes involved in cellular movement, development, growth and proliferation, immune response, and cell death. On the other hand, lipid metabolism and the tissue development process were significantly enriched categories among the downregulated genes (Figure 3B). Although we cannot exclude the possibility of an effect by SMG9 deficiency on NMD for specific transcripts, our data suggest that there is no widespread impairment of NMD although the transcriptional profile was dysregulated.

Smg9$^{-/-}$ Mice Have Reduced Viability and Display Major Brain, Eye, and Cardiovascular Malformations

To determine its role in a mammalian model, we attempted to generate a null allele of *Smg9* in the mouse by using CRISPR/Cas9. We targeted exon 2 (first coding exon) of *Smg9* by using a single guide sequence designed to minimize off-target potential. From eight founders, a single mutation containing a combination of a 2 bp deletion and 10 bp insertion (Figures 4A and 4B) was chosen for expansion and analysis. This mutation is predicted to result in a frameshift and a premature stop at amino acid position 104 (Figure 4C). qRT-PCR on E15.5 embryos showed that *Smg9* is reduced by 60%–80%, indicating that the mouse line is at least a severe hypomorph (Figure S7).

Intercrosses of heterozygous animals yielded no homozygous pups, suggesting embryonic lethality. To identify and characterize embryonic phenotypes of *Smg9*$^{-/-}$, we performed timed matings and harvested embryos at E14.5, E15.5, and E18.5 (Table 2). As shown in Figure 5 and Table 2, phenotypes were identified at all stages, and the exact presentation varied significantly between specimens. For example, homozygous embryos examined at E15.5 showed a range of phenotypes including edema, hemorrhage, and exencephaly (Figures 5C and 5D) and in one case,

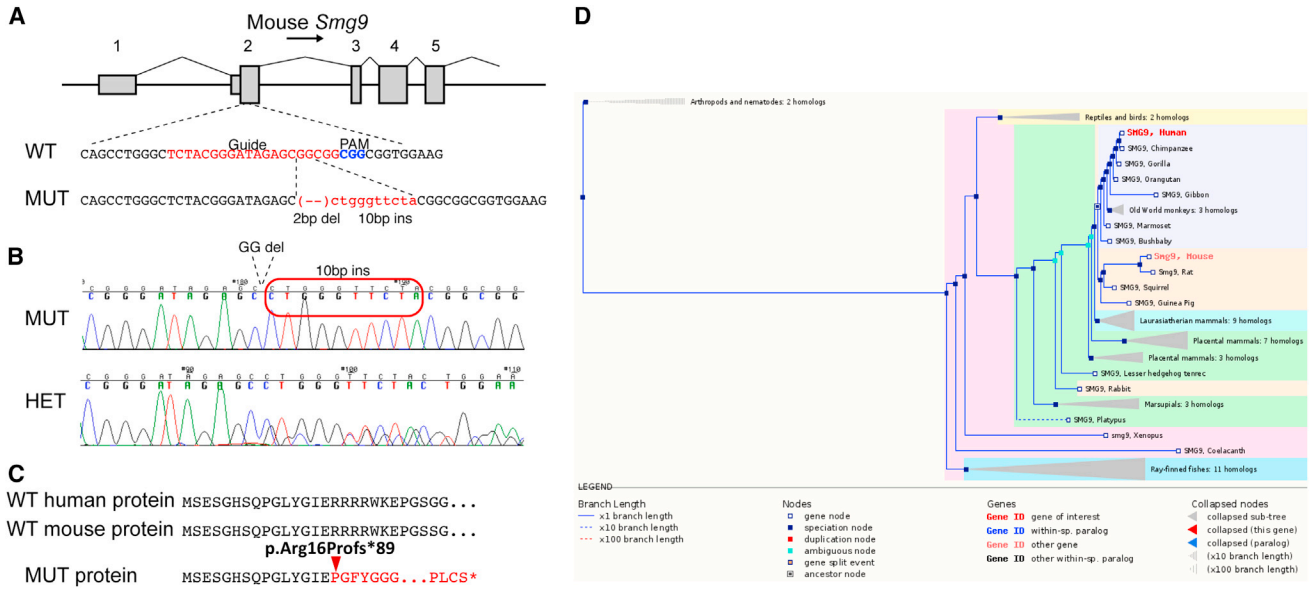


Figure 4. Generation of an *Smg9* Null Allele
(A) Schematic of mouse *Smg9*, showing the position of the CRISPR guide position in exon 2 (first coding exon; exon numbers indicated above the gene model). The mutant allele chosen for analysis contains a 2 bp deletion and simultaneous 10 bp insertion, resulting in a frameshift mutation beginning at 7:24403443 (GRCm38/mm10).
(B) Sequence chromatogram of a TOPO-TA clone of the mutant allele (top) and a heterozygous mutant (bottom), confirming the allele structure. Red box indicates inserted sequence.
(C) Structure of the predicted protein product showing the frameshift at residue 16 and the predicted premature termination at position 104. The aligned human protein is shown above the mouse sequence.
(D) SMG9 phylogenetic tree and conservation across species.

hind-limb preaxial polydactyly (Figure 5D, arrow). These phenotypes are variable and incompletely penetrant, but the majority of homozygotes display some clear phenotypic abnormality, unlike wild-type and heterozygous embryos (Figures 5A and 5B). To examine these phenotypes in more detail, we performed iodine-contrast microCT (computed tomography), which provides a high-resolution 3D dataset with sufficient contrast to observe internal structures. Mutant embryos were clearly smaller and developmentally delayed. Several other features were noted, including edema, observed during gross inspection (Figures 5H–5J, arrowheads). We also noted reduced size in the mid- and hindbrain relative to their stage of development (Figure 5H), microphthalmia (Figure 5I), thin myocardium, and a very prominent atrioventricular septal defect (AVSD) (Figure 5J, arrow). At E18.5, the two homozygotes identified displayed variable levels of edema (Figures 5L and 5M), with evidence of hemorrhage. One specimen was dead (Figure 5M) and showed thin myocardium in sections of microCT volumes (Figure 5S, arrow) and anophthamia (Figures 5M and 5R, arrow) of one eye, consistent with the phenotypes observed at E15.5. Thus, our findings support a role for *Smg9* in brain, heart, and eye development.

Discussion

The widespread role of NMD in transcriptional regulation suggests that defects in NMD can have pathological consequences at the organismal level. Interrogating a developmental role of NMD is challenging because many of its components have additional physiological roles. Moreover, NMD, although conserved across eukaryotes as a mechanism, differs in its details between different species such that the results observed in one species are not necessarily applicable to others[6,32].

Table 2. Summary of Genotypes and Phenotypes Observed in Smg9 Mouse Embryos

	Genotype			Mutant Phenotypes					
Stage (n)	**+/+**	**−/+**	**−/−**	**Developmental Delay**	**Brain (uCT)**	**Edema**	**Heart (uCT)**	**Eye (uCT)**	**Limb**
E14.5 (4)	1	1	2	2/2	2/2	2/2	NA	NA	1/4
E15.5 (38)	13	17	8	8/8	7/7	7/8	7/7	7/7	1/8
E18.5 (13)	6	5	2	2/2	2/2	1/2	1/2	1/2	0/2

NA, not applicable.

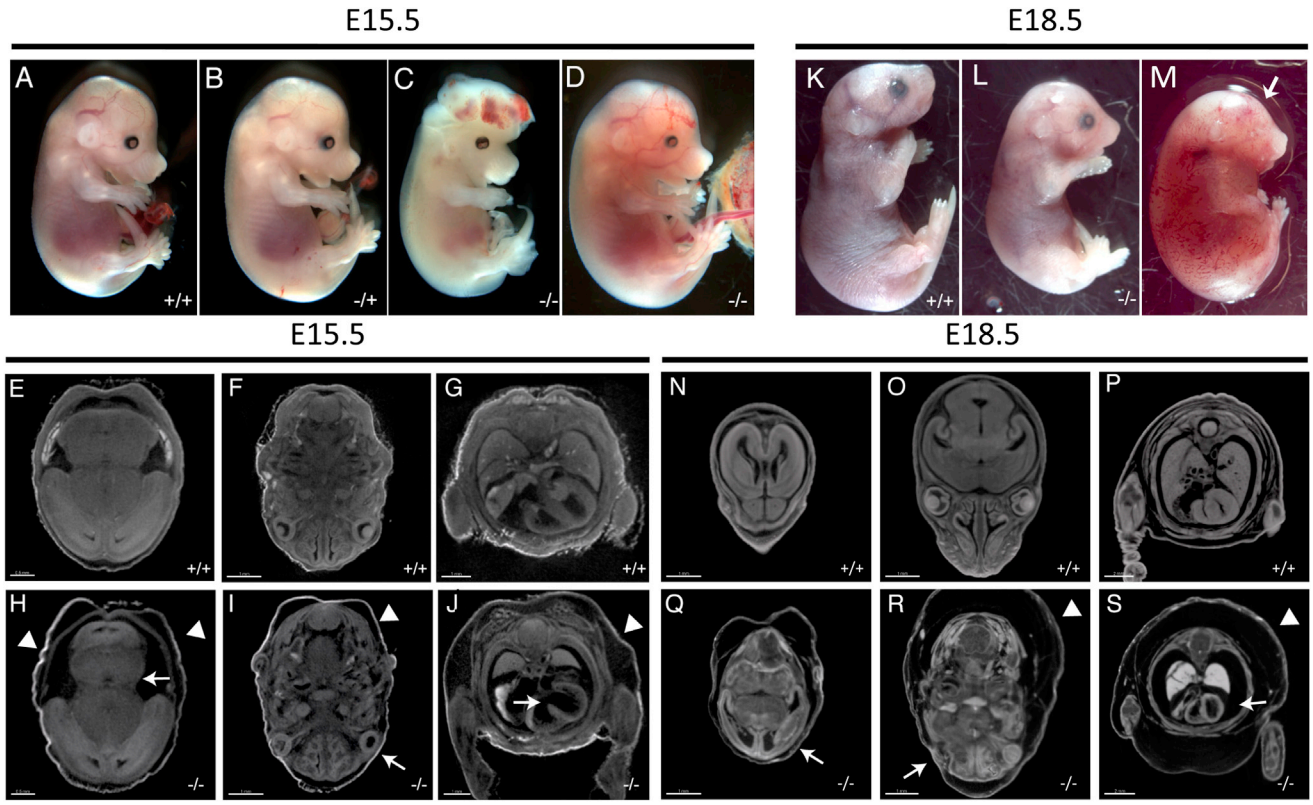


Figure 5. Multiple Developmental Abnormalities in *Smg9*$^{-/-}$ Mouse Embryos
(A–D) Mutant embryos at E15.5 display multiple developmental defects, including exencephaly (C) and sporadic hemorrhage and edema (D), versus wild-type (A) or heterozygous (B) littermates.
(E–J) Volume slices of iodine-contrast microCT images of the mutant in (D) show multiple defects, including abnormal brain development (arrow, H), microphthalmia (arrow, I), and an atrioventricular septal defect and thin myocardium (arrow, J). Clear edema in the mutant images is indicated by the arrowheads (H–J). Control E15.5 sections from similar planes are shown in (E)–(G).
(K–S) Some embryos survived or were recovered dead at E18.5 (L and M versus K). As with earlier stages, both mutants were developmentally delayed and showed a variable degree of edema. One mutant had anophthalmia of one eye (arrow in M and R). MicroCT images of the mutant in (M) show edema (arrowheads in R and S), a severely malformed brain (Q), and abnormally thin myocardium in (S). Control sections through similar planes are shown in (N)–(P).
Scale bars for (E)–(J) are 0.5 mm (E and H) and 1.0 mm (F, G, I, and J) and for (N)–(S) are 1.0 mm (N, O, Q, and R) and 2.0 mm (P and S).

Nonetheless, reduced embryonic viability as a converging phenotype among animal models deficient for at least some components of NMD appears to suggest that NMD does contribute to normal development. For example, knockdown of SMGL-1 and SMGL-2 is embryonically lethal in *C. elegans*[33]. In fruitflies, depletion of UPF1 and UPF2 is also embryonically lethal[34]. Zebrafish morphants for *upf1*, *upf2*, *smg5*, or *smg6* have high rates of embryonic lethality[35]. Finally, mice knocked out for *Smg1*, *Upf1*, or *Upf2* display early embryonic lethality as well [36–38].

SMG9 was first identified as a component of NMD in 2009 when it was found, along with SMG8, to co-purify with SMG1, forming a stable SMG1-SMG8-SMG9 complex[11]. Further characterization of this protein revealed that it is part of the SURF complex and is necessary for the phosphorylation of UFP1 by SMG1. Deletion mapping identified that the presence of both N and C termini is needed for the proper binding to SMG1[39]. Although the consequences of SMG9 deficiency on the NMD machinery are clearly demonstrated, virtually nothing is known about its developmental relevance. As mentioned above, deficiency of various NMD components has been found to result in abnormal development in model organisms, and UPF3B deficiency in humans is linked to intellectual disability[40]. However, SMG9 deficiency states have not been established in humans or other model organisms. Our study, therefore, fills an important gap in our knowledge of the developmental role of this component of NMD.

The identification of a single disease locus and two independent, homozygous, apparently loss-of-function mutations therein involving the same gene is highly suggestive of a causal relationship between *SMG9* mutations and the multiple congenital anomaly syndrome we describe. This is further corroborated by the recapitulation of key developmental defects in the knockout mice we generated via CRISPR/Cas9 mutagenesis in this study. Interestingly, the predominant involvement of the brain is reminiscent of the zebrafish morphants that are knocked down for various NMD components in which the hindbrain is particularly affected [35]. However, the mechanism through

which SMG9 deficiency results in abnormal embryogenesis remains unclear.

Perturbation of NMD was a prime candidate mechanism in view of the established role of NMD in protecting the cells from the toxic effect of PTC-containing proteins. However, our finding that the *SMG9* mutant transcript itself is subject to NMD in the affected individual's cells suggested to us that this might not be the case. Indeed, subsequent global transcriptional profiling also failed to show evidence of impaired NMD as a result of SMG9 deficiency. Although different transcripts are known to have different sensitivity to NMD, so we cannot exclude the possibility that NMD is defective for certain PTC-containing transcripts, our data suggest that NMD was generally functional. Given that NMD affects at least 10% of human transcripts, many of which lack PTC, it is possible that the generalized transcriptional dysregulation we observed is still reflective of an abnormality in NMD that we could not measure. Interestingly, it has been suggested that SMG9 might play other physiological roles independent of its role in NMD because it was identified in homodimers that are not part of the SURF complex, and this could also be relevant to the transcriptional dysregulation of affected individuals' cells[39].

In conclusion, this study shows that *SMG9* mutations most likely cause a distinctive multiple congenital anomaly syndrome in humans and that its deficiency in mice also leads to abnormal embryogenesis and an overlapping developmental profile. Despite the established role of SMG9 in NMD, we show that PTC-containing transcripts do undergo efficient degradation in the context of SMG9 deficiency, although the latter state is still associated with widespread transcriptional dysregulation. Future studies will be needed to further delineate *SMG9*-related phenotypes in humans and the underlying mechanism, which might be related to a non-NMD role played by this protein.

Supplemental Data

Supplemental Data include a Supplemental Note, seven figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2016.02.010.

Acknowledgments

We thank the families in this study for their enthusiastic participation. We thank the Sequencing Core Facility at the King Faisal Specialist Hospital and Research Center for their technical help. This study was supported by King Abdulaziz City for Science and Technology (KACST) grant 13-BIO1113-20 (F.S.A.), NIH grant OD011185 (S.A.M.), KACST grant 11-BIO2072-20 (D.C.), and the King Salman Center for Disability Research (F.S.A.).

Received: October 27, 2015
Accepted: February 11, 2016
Published: March 24, 2016

Web Resources

The URLs for data presented herein are as follows:

Berkeley Drosophila Genome Project NNSplice 0.9, http://www.fruitfly.org/seq_tools/splice.html
Burrows-Wheeler Aligner, http://bio-bwa.sourceforge.net/
CRISPR Design, http://crispr.mit.edu
ExAC Browser, http://exac.broadinstitute.org/
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
OMIM, http://www.omim.org/
SAMtools, http://samtools.sourceforge.net/
The Gene Expression Barcode project, http://barcode.luhs.org
The Jackson Laboratory mouse strain datasheet, https://www.jax.org/strain/027253
UCSC Genome Browser hg19, http://genome.ucsc.edu

References

1. Pennisi, E. (2012). Genomics. ENCODE project writes eulogy for junk DNA. Science *7*, 1159–1161.
2. Holoch, D., and Moazed, D. (2015). RNA-mediated epigenetic regulation of gene expression. Nat. Rev. Genet. *16*, 71–84.
3. Bhuvanagiri, M., Schlitter, A.M., Hentze, M.W., and Kulozik, A.E. (2010). NMD: RNA biology meets human genetic medicine. Biochem. J. *430*, 365–377.
4. Stalder, L., and Mühlemann, O. (2008). The meaning of nonsense. Trends Cell Biol. *18*, 315–321.
5. Chang, Y.-F., Imam, J.S., and Wilkinson, M.F. (2007). The nonsense-mediated decay RNA surveillance pathway. Annu. Rev. Biochem. *76*, 51–74.
6. Isken, O., and Maquat, L.E. (2007). Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev. *21*, 1833–1856.
7. Hwang, J., and Maquat, L.E. (2011). Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. Curr. Opin. Genet. Dev. *21*, 422–430.
8. Behm-Ansmant, I., and Izaurralde, E. (2006). Quality control of gene expression: a stepwise assembly pathway for the surveillance complex that triggers nonsense-mediated mRNA decay. Genes Dev. *20*, 391–398.
9. Lykke-Andersen, J., Shu, M.-D., and Steitz, J.A. (2000). Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. Cell *103*, 1121–1131.
10. Ohnishi, T., Yamashita, A., Kashima, I., Schell, T., Anders, K.R., Grimson, A., Hachiya, T., Hentze, M.W., Anderson, P., and Ohno, S. (2003). Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. Mol. Cell *12*, 1187–1200.
11. Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., and Ohno, S. (2006). Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes Dev. *20*, 355–367.
12. Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T., et al. (2009). SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. Genes Dev. *23*, 1091–1105.

13. Clerici, M., Mourão, A., Gutsche, I., Gehring, N.H., Hentze, M.W., Kulozik, A., Kadlec, J., Sattler, M., and Cusack, S. (2009). Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. EMBO J. *28*, 2293–2306.
14. Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., and Ohno, S. (2012). N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. Nucleic Acids Res. *40*, 1251–1266.
15. Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature translational termination. Nat. Rev. Mol. Cell Biol. *13*, 700–712.
16. Metze, S., Herzog, V.A., Ruepp, M.-D., and Mühlemann, O. (2013). Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. RNA *19*, 1432–1448.
17. Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat. Genet. *36*, 1073–1078.
18. Wittmann, J., Hol, E.M., and Jäck, H.-M. (2006). hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. Mol. Cell. Biol. *26*, 1272–1287.
19. Chan, W.K., Huang, L., Gudikote, J.P., Chang, Y.F., Imam, J.S., MacLean, J.A., 2nd, and Wilkinson, M.F. (2007). An alternative branch of the nonsense-mediated decay pathway. EMBO J. *26*, 1820–1830.
20. Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., and Jacobson, A. (2004). A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature *432*, 112–118.
21. Eberle, A.B., Stalder, L., Mathys, H., Orozco, R.Z., and Mühlemann, O. (2008). Posttranscriptional gene regulation by spatial rearrangement of the 3′ untranslated region. PLoS Biol. *6*, e92.
22. Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W., and Kulozik, A.E. (2008). Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. EMBO J. *27*, 736–747.
23. Rebbapragada, I., and Lykke-Andersen, J. (2009). Execution of nonsense-mediated mRNA decay: what defines a substrate? Curr. Opin. Cell Biol. *21*, 394–402.
24. Alkuraya, F.S. (2010). Autozygome decoded. Genet. Med. *12*, 765–771.
25. Alkuraya, F.S. (2012). Discovery of rare homozygous mutations from studies of consanguineous pedigrees. In Current Protocols in Human Genetics (Wiley) http://dx.doi.org/10.1002/0471142905.hg0612s75, Unit 6.12.
26. Carr, I.M., Flintoff, K.J., Taylor, G.R., Markham, A.F., and Bonthron, D.T. (2006). Interactive visual analysis of SNP data for rapid autozygosity mapping in consanguineous families. Hum. Mutat. *27*, 1041–1046.
27. Lindner, T.H., and Hoffmann, K. (2005). easyLINKAGE: a PERL script for easy and automated two-/multi-point linkage analyses. Bioinformatics *21*, 405–407.
28. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Series B Methodol. *57*, 289–300.
29. Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. *4*, 3.
30. Guimier, A., Gabriel, G.C., Bajolle, F., Tsang, M., Liu, H., Noll, A., Schwartz, M., El Malti, R., Smith, L.D., Klena, N.T., et al. (2015). MMP21 is mutated in human heterotaxy and is required for normal left-right asymmetry in vertebrates. Nat. Genet. *47*, 1260–1263.
31. Wong, M.D., Spring, S., and Henkelman, R.M. (2013). Structural stabilization of tissue for embryo phenotyping using micro-CT with iodine staining. PLoS ONE *8*, e84321.
32. Conti, E., and Izaurralde, E. (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. Curr. Opin. Cell Biol. *17*, 316–325.
33. Longman, D., Plasterk, R.H., Johnstone, I.L., and Cáceres, J.F. (2007). Mechanistic insights and identification of two novel factors in the C. elegans NMD pathway. Genes Dev. *21*, 1075–1085.
34. Avery, P., Vicente-Crespo, M., Francis, D., Nashchekina, O., Alonso, C.R., and Palacios, I.M. (2011). Drosophila Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. RNA *17*, 624–638.
35. Wittkopp, N., Huntzinger, E., Weiler, C., Saulière, J., Schmidt, S., Sonawane, M., and Izaurralde, E. (2009). Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. Mol. Cell. Biol. *29*, 3517–3528.
36. Medghalchi, S.M., Frischmeyer, P.A., Mendell, J.T., Kelly, A.G., Lawler, A.M., and Dietz, H.C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. Hum. Mol. Genet. *10*, 99–105.
37. McIlwain, D.R., Pan, Q., Reilly, P.T., Elia, A.J., McCracken, S., Wakeham, A.C., Itie-Youten, A., Blencowe, B.J., and Mak, T.W. (2010). Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. Proc. Natl. Acad. Sci. USA *107*, 12186–12191.
38. Thoren, L.A., Nørgaard, G.A., Weischenfeldt, J., Waage, J., Jakobsen, J.S., Damgaard, I., Bergström, F.C., Blom, A.M., Borup, R., Bisgaard, H.C., and Porse, B.T. (2010). UPF2 is a critical regulator of liver development, function and regeneration. PLoS ONE *5*, e11650.
39. Fernández, I.S., Yamashita, A., Arias-Palomo, E., Bamba, Y., Bartolomé, R.A., Canales, M.A., Teixidó, J., Ohno, S., and Llorca, O. (2011). Characterization of SMG-9, an essential component of the nonsense-mediated mRNA decay SMG1C complex. Nucleic Acids Res. *39*, 347–358.
40. Tarpey, P.S., Raymond, F.L., Nguyen, L.S., Rodriguez, J., Hackett, A., Vandeleur, L., Smith, R., Shoubridge, C., Edkins, S., Stevens, C., et al. (2007). Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. Nat. Genet. *39*, 1127–1133.

The American Journal of Human Genetics, Volume *98*

Supplemental Information

Mutations in *SMG9*, Encoding an Essential Component of Nonsense-Mediated Decay Machinery, Cause a Multiple Congenital Anomaly Syndrome in Humans and Mice

Ranad Shaheen, Shams Anazi, Tawfeg Ben-Omran, Mohammed Zain Seidahmed, L. Brianna Caddle, Kristina Palmer, Rehab Ali, Tarfa Alshidi, Samya Hagos, Leslie Goodwin, Mais Hashem, Salma M. Wakil, Mohamed Abouelhoda, Dilek Colak, Stephen A. Murray, and Fowzan S. Alkuraya

'Supplemental Note: Case Reports'

Family 1: IV:5 (the index) had a birth weight of 2.130 kg (5^{th} centile), length of 44 cm (-2.4SD), and head circumference of 31.5 cm (-2.2SD). Antenatal ultrasound scan had revealed polyhydramnios, abnormal umbilical artery Doppler (absent diastolic flow), splaying of the cerebellum and VSD. Apgar scores were 6 and 8 at one and five minutes respectively. Physical examination showed craniofacial abnormalities consisting of prominent forehead and occiput, low set malformed ears, wide anterior fontanelle, depressed nasal bridge and anteverted nares, microphthalmia, high arched palate, clenched hands with camptodactyly. Central nervous system involvement included Dandy-Walker malformation, cerebellar vermis hypoplasia and hypoplastic corpus callosum on CT brain (Figure 1B). Cardiac evaluation revealed interrupted aortic arch, hypoplastic tricuspid and aortic valves, and large muscular VSD. She developed seizures and sepsis, and was mechanically ventilated. Despite aggressive management, she died at the age of 7 weeks. Laboratory investigations included normal hematologic indices, liver functions and renal functions. Her karyotype was 46,XX. Plasma amino acids, carnitine and acylcarnitines were normal. TORCH serology for congenital infections was negative. Renal ultrasound was normal.

Family 2: V:1 (the index) had a birth weight of 2.26 kg, length (-2.2SD) was 47 cm (15^{th} centile) and head circumference was 32 cm (-1.8SD). She was found to have right-sided cleft lip, and a large ventricular septal defect (VSD). Additional dysmorphic features included narrow forehead, prominent metopic suture, widow's peak, hypertelorism, posteriorly rotated ears with attached lobules, broad nasal bridge, full and everted lower lip, small eyes, and syndactyly between 2^{nd} and 3^{rd} toes (Figure 1C and D). She underwent VSD closure at 1 year of age and repair of cleft

lip at 18 months of age. She showed features of global developmental delay. At 3 years of age, she was only able to roll over, but not sit independently. Examination at 14 months of age revealed weight of 6 kg (-4.3SD), length of 67 cm (-3SD) and head circumference of 40 cm (-4.4SD). Neurological examination revealed truncal hypotonia and exaggerated deep tendon reflexes with clonus. Ophthalmological examination was significant for poor vision necessitating corrective glasses. Laboratory investigations showed normal electrolytes, renal and hepatic functions and the hematologic indices. Creatine kinase (CK) was normal. Tandem mass spectrometry (TMS) for metabolic disorders screen was unremarkable. Serum ammonia, lactate, plasma sterol profile and transferrin isoelectric focusing–CDG screening were normal. Array CGH and clinical exome sequencing were negative. Abdominal ultrasound revealed no abnormalities. Her most recent echocardiography showed status post VSD closure, otherwise unremarkable study. Brain MRI performed at 15 months revealed brain atrophy, decreased myelination and Dandy-Walker malformation. IV:3: The first cousin of the index was also found to have a similar phenotype. She was born at term after an uneventful pregnancy and delivery. She was found to have ventricular septal defect (VSD) and dysmorphic features. She underwent VSD closure at 13 months of age. She was also found to have major gastroesophageal reflux and recurrent aspiration necessitating NG feeding, bronchial stenosis and laryngeal cleft type 1. She showed features of global developmental delay. At 2 years of age, there was no eye contact, and she was only able to roll over, but not sit independently. Examination at 2 years of age revealed weight of 6.3 kg (-5.2SD), length of 69 cm (-4.8SD) and head circumference of 39 cm (-6SD). Neurological examination showed truncal hypotonia with peripheral hypertonia, brisk deep tendon reflexes and adductor spasm in both lower limbs. Laboratory investigations showed normal array CGH, plasma sterol profile and lysosomal

studies. Abdominal ultrasound revealed no abnormalities. Her most recent echocardiography showed status post VSD closure, and mild tricuspid regurgitation. Brain MRI performed at 7 months revealed generalized brain atrophy, prominent ventricular system and thin corpus

Figure S1

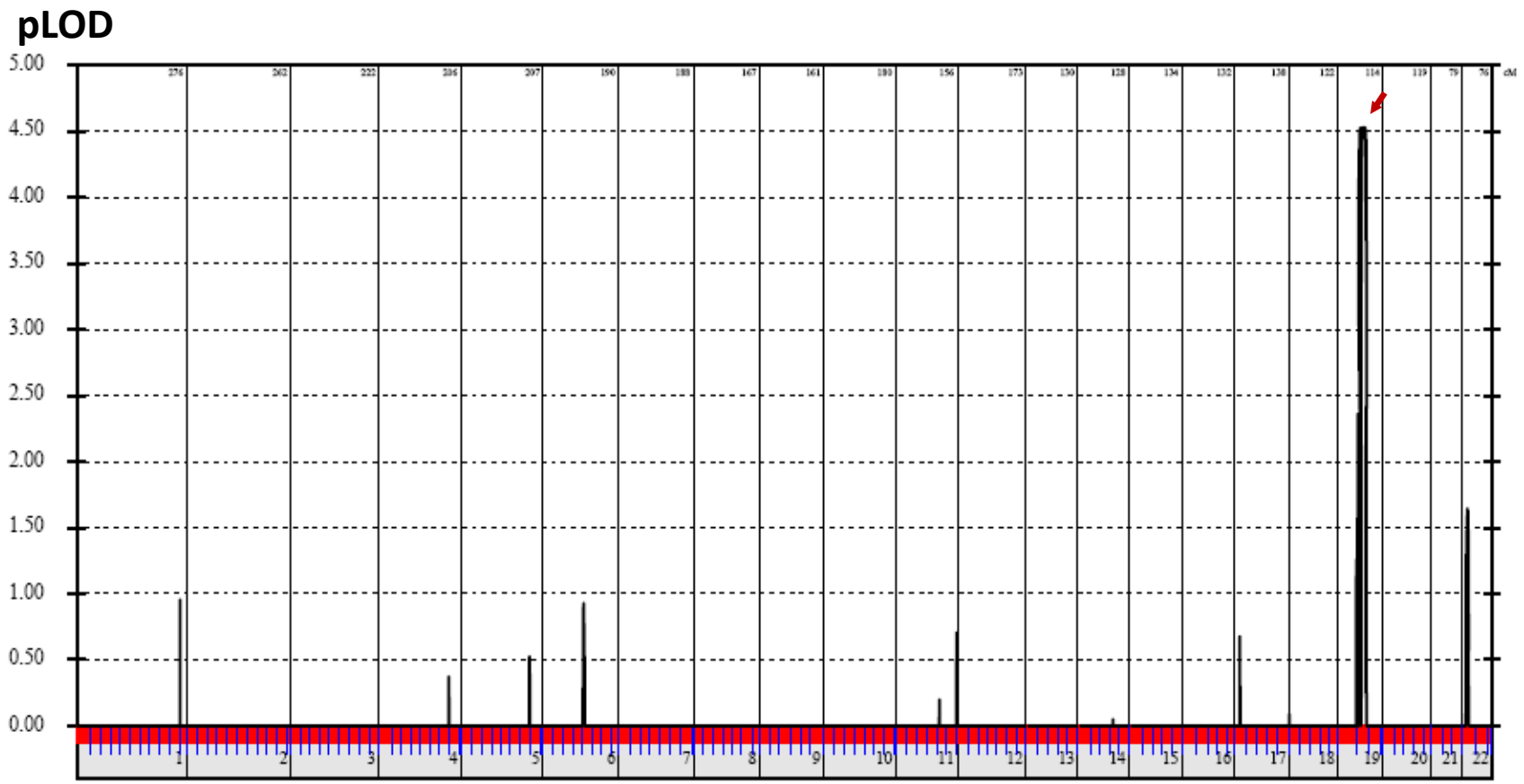

pLOD
5.00
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

Figure S1. Genome-wide linkage analysis revealed a single maximal peak with a LOD score of ~4.5 on chromosome 19.

Figure S2

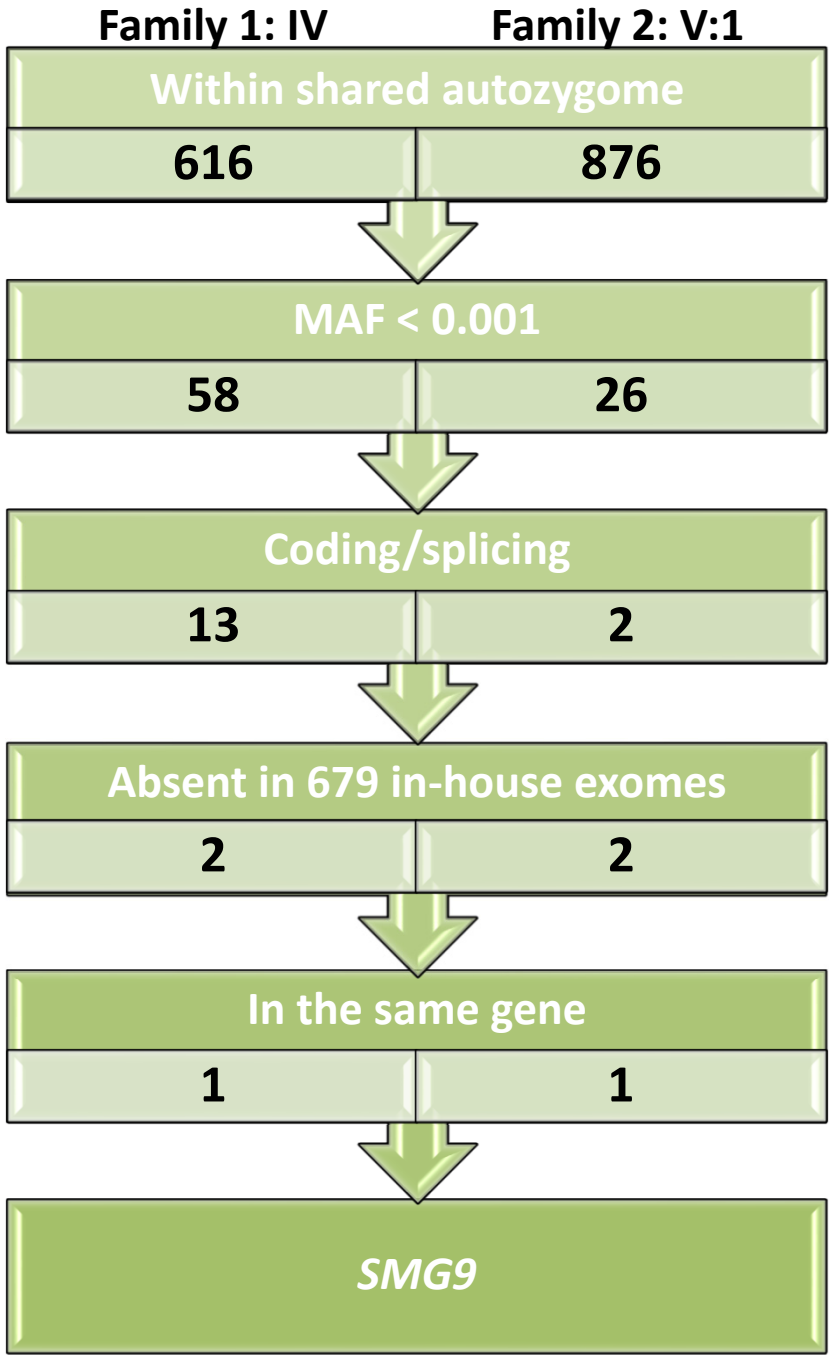

Family 1: IV
Family 2: V:1
Within shared autozygome
616
876
MAF < 0.001
58
26
Coding/splicing
13
2
Absent in 679 in-house exomes
2
2
In the same gene
1
1
SMG9

Figure S2. Illustration of the exome filtering scheme and the number of survived variants in each step in both families.

Figure S3

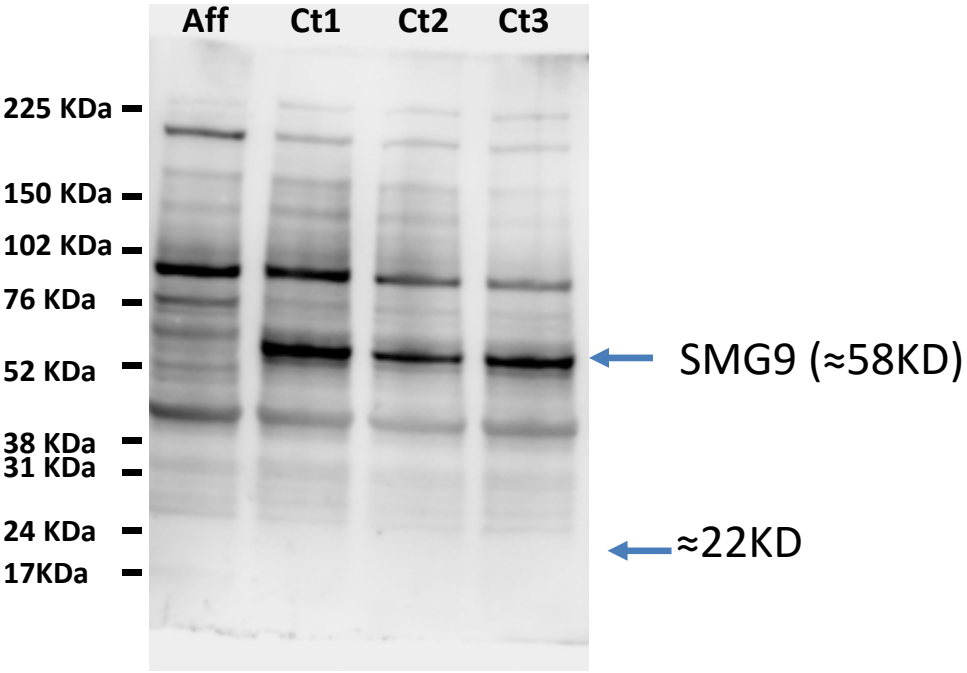
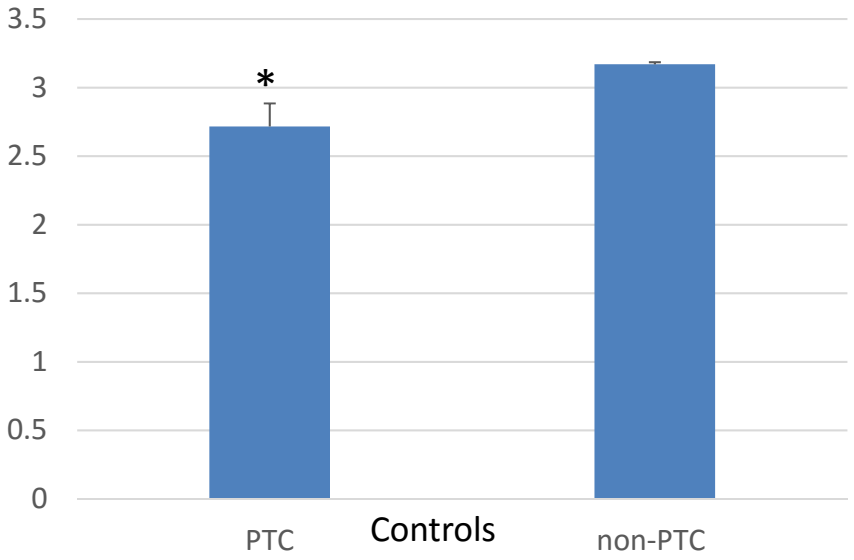

Aff
Ct1
Ct2
Ct3
225 KDa
150 KDa
102 KDa
76 KDa
52 KDa
38 KDa
31 KDa
24 KDa
17KDa
SMG9 (≈58KD)
≈22KD

Figure S3. Immunobloting using antibody against the N terminal of SMG9 antibody (SAB2107730). The image shows no detectable band from cells derived from affected individual (V:1) as compared with the three normal controls (Ct1, Ct2, Ct3) at the target mass (57.7-54.8) kDa as well as from the truncated transcript (predicted mass approximately 22KDa).

Figure S4

A


3.5
3
2.5
2
1.5
1
0.5
0
*
PTC
Controls
non-PTC

B

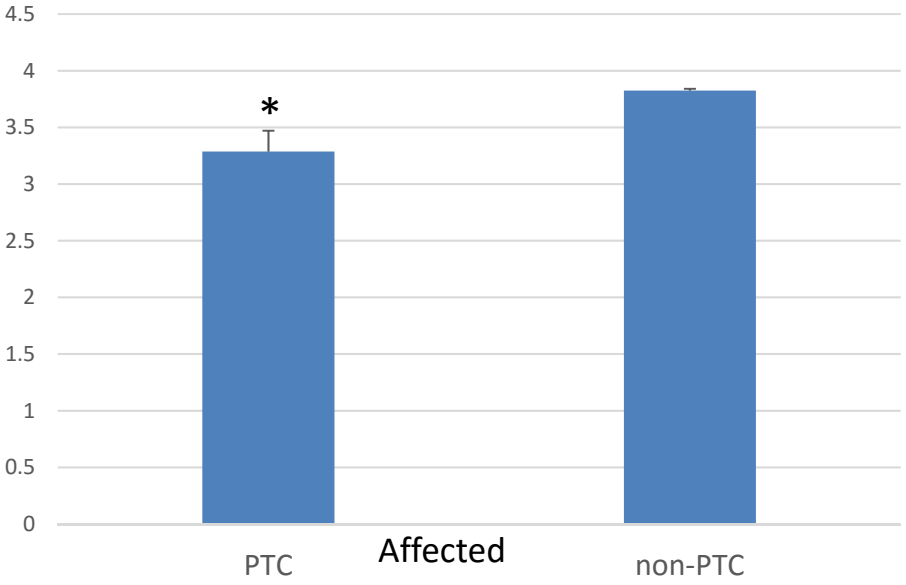
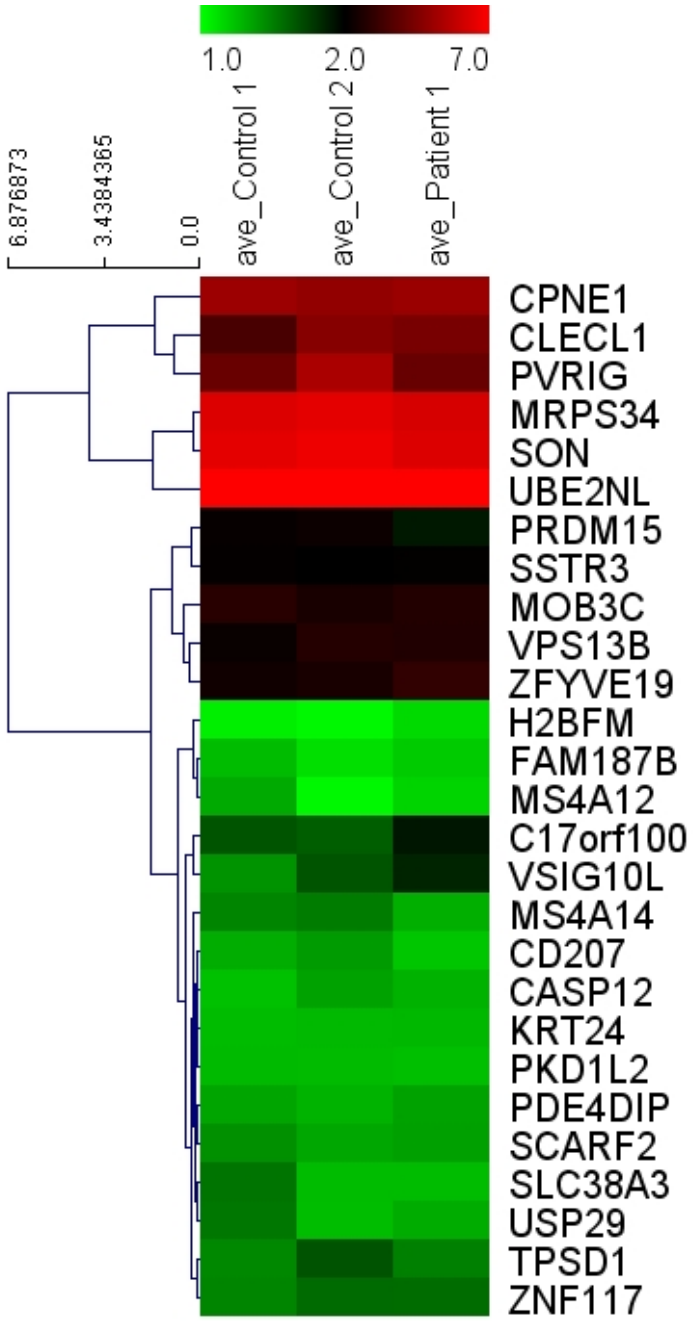

4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
*
PTC
Affected
non-PTC

Figure S4. Comparison of PTC-containing transcripts to other transcripts in controls and affected individual obtained from the expression microarray. Transcripts with PTC are significantly lower than other transcripts in general in controls (averaged) (A) and affected individual (B) ($p < 0.01$). Data shown represent mean ± SEM.

Figure S5



p_value (C1,C2)	0.918
p_value (C2,P1)	0.827
p_value (P1,C1)	0.758
p_value (P1,Cntrls)	0.986

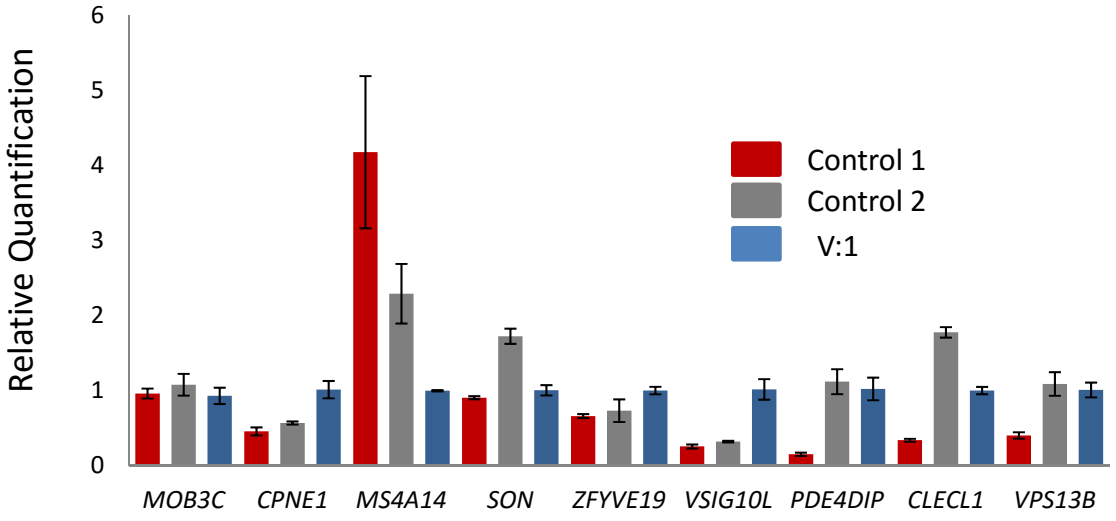


Figure S5. Heat-map and Venn diagram of common PTC-containing genes in controls and patient. There was no significant difference between in mRNA expression in affected individual and controls for the 27 shared PTC containing genes (p value >0.05). Purple circle in the Venn diagram represents affected individual (V:1), light pink represents control 1 and light green represents control 2.

Figure S6

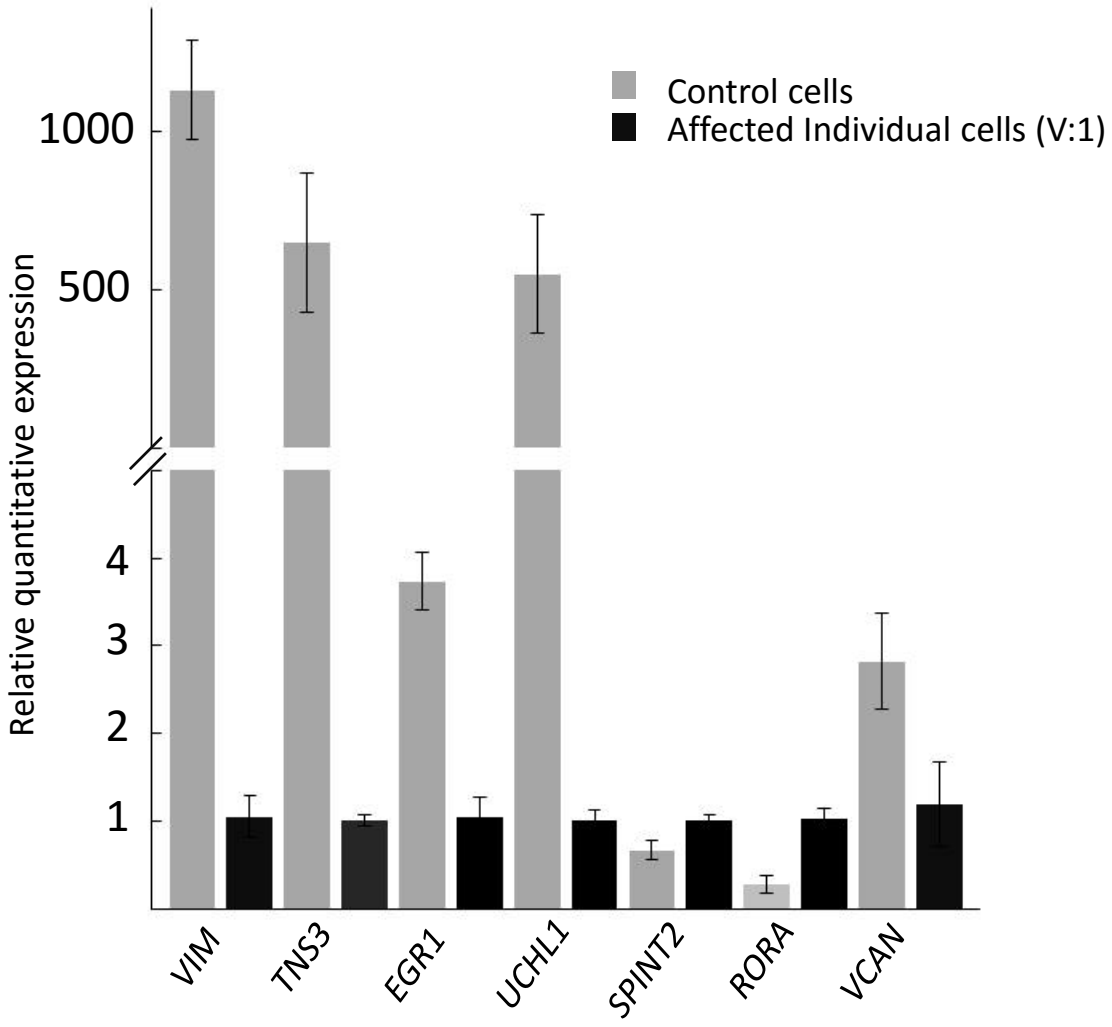

Control cells
Affected Individual cells (V:1)
Relative quantitative expression
1000
500
4
3
2
1
VIM
TNS3
EGR1
UCHL1
SPINT2
RORA
VCAN

Figure S6. Relative quantification Real time PCR result for selected genes found to be dysregulated in the microarray in affected individual (V:1) compared to control. Result is the average of triplicate experiments.

Figure S7

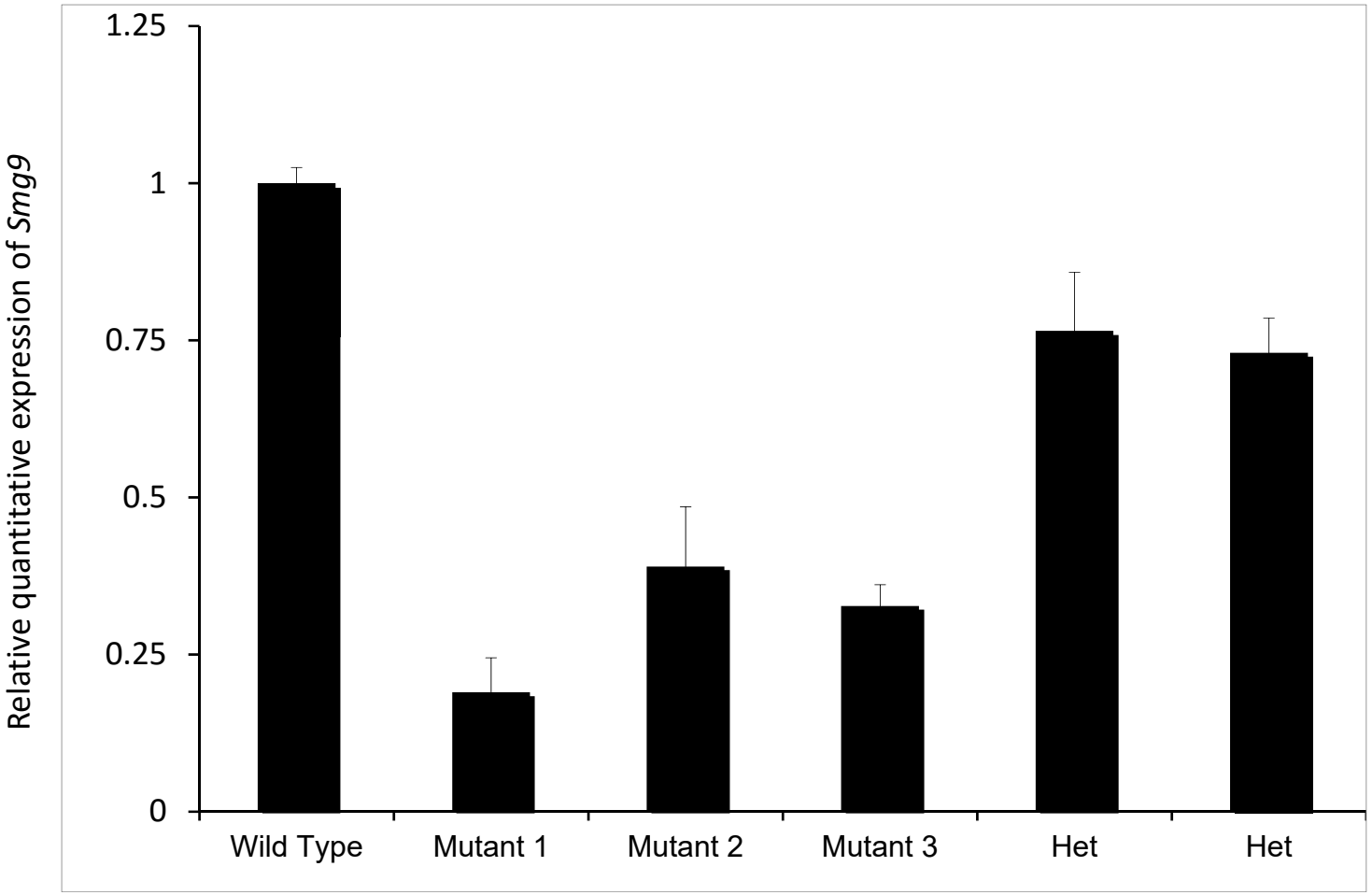

Relative quantitative expression of *Smg9*
1.25
1
0.75
0.5
0.25
0
Wild Type
Mutant 1
Mutant 2
Mutant 3
Het
Het

Figure S7. Relative quantification Real time PCR of *Smg9* shows 60-80% reduction in *Smg9* transcripts in mutant mouse compare to wild type. Each biological replicate is represented on the bar graphs.

Supplementary Table 1A. Significantly up-regulated genes in SMG9 deficient individual compared to controls.

Gene Symbol	Gene Title	FC[a]
CCL20	chemokine (C-C motif) ligand 20	38.3
IGKC	immunoglobulin kappa constant	26.7
DSC2	desmocollin 2	22.4
CRYBB2 /// CRYBB2P1	crystallin, beta B2 /// crystallin, beta B2 pseudogene 1	20.0
HLA-DRB4	major histocompatibility complex, class II, DR beta 4	19.0
TMPRSS3	transmembrane protease, serine 3	16.2
SERPINB2	serpin peptidase inhibitor, clade B (ovalbumin), member 2	15.7
ADAMDEC1	ADAM-like, decysin 1	15.3
RORA	RAR-related orphan receptor A	14.3
CXCL9	chemokine (C-X-C motif) ligand 9	14.3
CRNDE	colorectal neoplasia differentially expressed (non-protein coding)	13.2
CASP1	caspase 1, apoptosis-related cysteine peptidase	12.3
DSC3	desmocollin 3	11.3
TACSTD2	tumor-associated calcium signal transducer 2	10.2
PRRX1	paired related homeobox 1	8.9
CMTM7	CKLF-like MARVEL transmembrane domain containing 7	8.5
TAGAP	T-cell activation RhoGTPase activating protein	6.8
LOC100506548 /// RPL37	uncharacterized LOC100506548 /// ribosomal protein L37	6.8
VCAN	versican	6.7
TCFL5	transcription factor-like 5 (basic helix-loop-helix)	6.7
PHF16	PHD finger protein 16	6.6
GLUL	glutamate-ammonia ligase	6.5
CMAHP	cytidine monophospho-N-acetylneuraminic acid hydroxylase, pseudogene	6.3

SPINT2	serine peptidase inhibitor, Kunitz type, 2	6.0
PIEZO2	piezo-type mechanosensitive ion channel component 2	6.0
CARD16 /// CASP1	caspase recruitment domain family, member 16 /// caspase 1, apoptosis-related cysteine	5.8
NOTCH2	notch 2	5.8
HSPA6	heat shock 70kDa protein 6 (HSP70B')	5.8
UXS1	UDP-glucuronate decarboxylase 1	5.6
C12orf79	chromosome 12 open reading frame 79	5.5
UGT8	UDP glycosyltransferase 8	5.4
LOC101060315 //NOTCH2 /// NOTCH2NL	neurogenic locus notch homolog protein 2-like /// neurogenic locus notch homolog protei	5.4
LOC389906	zinc finger protein 839 pseudogene	5.4
KRT7	keratin 7	5.2
SETD7	SET domain containing (lysine methyltransferase) 7	5.2
SOCS2	suppressor of cytokine signaling 2	5.1
CARD16	caspase recruitment domain family, member 16	5.0
APLP2	amyloid beta (A4) precursor-like protein 2	4.8
CR2	complement component (3d/Epstein Barr virus) receptor 2	4.7
OTTHUMG00000168533 /// RP11-277P12.20	NULL /// NULL	4.6
LINC00888	long intergenic non-protein coding RNA 888	4.4
NIPAL4	NIPA-like domain containing 4	4.3
LOC100506453	uncharacterized LOC100506453 /// zinc finger protein 839 pseudogene /// uncharacterized	4.3
NUDT12	nudix (nucleoside diphosphate linked moiety X)-type motif 12	4.3
ITM2A	integral membrane protein 2A	4.2
BAG3	BCL2-associated athanogene 3	4.2
JAK1	Janus kinase 1	4.2
CBS	cystathionine-beta-synthase	4.2

TBC1D1	TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1	4.1
ZAK	sterile alpha motif and leucine zipper containing kinase AZK	4.0
PGAP1	post-GPI attachment to proteins 1	4.0
LOC100996457	uncharacterized LOC100996457	4.0
RASA4 /// RASA4B /// RASA4CP	RAS p21 protein activator 4 /// RAS p21 protein activator 4B /// RAS p21 protein activa	3.9
JUN	jun proto-oncogene	3.9
FOXO3	forkhead box O3	3.8
KLF3	Kruppel-like factor 3 (basic)	3.8
ANKDD1A	ankyrin repeat and death domain containing 1A	3.8
DHRS7	dehydrogenase/reductase (SDR family) member 7	3.7
IFRD1	interferon-related developmental regulator 1	3.7
CD44	CD44 molecule (Indian blood group)	3.7
SPIN3	spindlin family, member 3	3.6
DNAJC18	DnaJ (Hsp40) homolog, subfamily C, member 18	3.6
IL1R2	interleukin 1 receptor, type II	3.5
CR1	complement component (3b/4b) receptor 1 (Knops blood group)	3.5
KCNN2	potassium intermediate/small conductance calcium-activated channel, subfamily N, member	3.4
PSPC1	paraspeckle component 1	3.4
ERAP2	endoplasmic reticulum aminopeptidase 2	3.4
BAG1	BCL2-associated athanogene	3.3
GOLM1	golgi membrane protein 1	3.3
---	---	3.2
TSC22D1	TSC22 domain family, member 1	3.2
GLB1L3	galactosidase, beta 1-like 3	3.2
CD55	CD55 molecule, decay accelerating factor for complement (Cromer blood group)	3.2

TMEM173	transmembrane protein 173	3.2
NFKBIZ	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	3.1
AZI2	5-azacytidine induced 2	3.1
LOC100507581	uncharacterized LOC100507581	3.1
FNDC3B	fibronectin type III domain containing 3B	3.1
GAS5 /// SNORD44 ///	growth arrest-specific 5 (non-protein coding) /// small nucleolar RNA, C/D box 44 /// s	3.1
SNHG12 /// SNORA16A /// SNORA44 /// SNORA61	small nucleolar RNA host gene 12 (non-protein coding) /// small nucleolar RNA, H/ACA bo	3.0
TET2	tet methylcytosine dioxygenase 2	3.0
GADD45B	growth arrest and DNA-damage-inducible, beta	3.0
MAN1A1	mannosidase, alpha, class 1A, member 1	3.0
CXCR4	chemokine (C-X-C motif) receptor 4	3.0
EPB41L4A-AS1	EPB41L4A antisense RNA 1	3.0
RASA4 /// RASA4B	RAS p21 protein activator 4 /// RAS p21 protein activator 4B	3.0
ZBTB10	zinc finger and BTB domain containing 10	3.0
MAP3K7CL	MAP3K7 C-terminal like	2.9
PLAU	plasminogen activator, urokinase	2.9
SNHG15 /// SNORA9	small nucleolar RNA host gene 15 (non-protein coding) /// small nucleolar RNA, H/ACA bo	2.9
C17orf58	chromosome 17 open reading frame 58	2.9
C6orf48	chromosome 6 open reading frame 48	2.9
GPR55	G protein-coupled receptor 55	2.9
IL2RA	interleukin 2 receptor, alpha	2.9
CBLB	Cbl proto-oncogene, E3 ubiquitin protein ligase B	2.8
BZW2	basic leucine zipper and W2 domains 2	2.8
LOC152225	uncharacterized LOC152225	2.8
TFDP2	transcription factor Dp-2 (E2F dimerization partner 2)	2.8

TP53BP2	tumor protein p53 binding protein, 2	2.8
THAP9-AS1	THAP9 antisense RNA 1	2.8
FAM189A1	family with sequence similarity 189, member A1	2.8
CCR7	chemokine (C-C motif) receptor 7	2.7
RAB27A	RAB27A, member RAS oncogene family	2.7
ZFAS1	ZNFX1 antisense RNA 1	2.7
EXOSC7	exosome component 7	2.7
UPF3B	UPF3 regulator of nonsense transcripts homolog B (yeast)	2.7
SNHG6 /// SNORD87	small nucleolar RNA host gene 6 (non-protein coding) /// small nucleolar RNA, C/D box 8	2.6
HNRNPA2B1	heterogeneous nuclear ribonucleoprotein A2/B1	2.6
ARSK	arylsulfatase family, member K	2.6
RGPD3 /// RGPD4 /// RGPD5 /// RGPD6 /// RGPD8	RANBP2-like and GRIP domain containing 3 /// RANBP2-like and GRIP domain containing 4 /	2.6
GADD45A	growth arrest and DNA-damage-inducible, alpha	2.6
BMP4	bone morphogenetic protein 4	2.6
POLR2J2 /// POLR2J3 /// POLR2J4 /// UPK3BL	polymerase (RNA) II (DNA directed) polypeptide J2 /// polymerase (RNA) II (DNA directed	2.6
GAB1	GRB2-associated binding protein 1	2.5
DOCK9	dedicator of cytokinesis 9	2.5
ERCC8	excision repair cross-complementing rodent repair deficiency, complementation group 8	2.5
DNAJB2	DnaJ (Hsp40) homolog, subfamily B, member 2	2.5
SARNP	SAP domain containing ribonucleoprotein	2.5
HNRNPA1	heterogeneous nuclear ribonucleoprotein A1	2.5
CD274	CD274 molecule	2.5
CD86	CD86 molecule	2.5
PIK3R5	phosphoinositide-3-kinase, regulatory subunit 5	2.5
DENND1A	DENN/MADD domain containing 1A	2.5

CCL5	chemokine (C-C motif) ligand 5	2.5
CLYBL	citrate lyase beta like	2.5
CFP	complement factor properdin	2.4
TIGIT	T cell immunoreceptor with Ig and ITIM domains	2.4
GALNT1	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 1 (Gal	2.4
UFM1	ubiquitin-fold modifier 1	2.3
LOC728613	programmed cell death 6 pseudogene	2.3
F11R	F11 receptor	2.3
MBNL3	muscleblind-like splicing regulator 3	2.3
RBMX /// SNORD61	RNA binding motif protein, X-linked /// small nucleolar RNA, C/D box 61	2.3
OARD1	O-acyl-ADP-ribose deacylase 1	2.3
RASSF1	Ras association (RalGDS/AF-6) domain family member 1	2.3
MDM2	MDM2 oncogene, E3 ubiquitin protein ligase	2.3
GPATCH2L	G patch domain containing 2-like	2.3
LOC100506453	uncharacterized LOC100506453 /// zinc finger protein 839 pseudogene /// uncharacterized	2.3
IL6	interleukin 6 (interferon, beta 2)	2.2
MICU3	mitochondrial calcium uptake family, member 3	2.2
PTBP2	polypyrimidine tract binding protein 2	2.2
RC3H1	ring finger and CCCH-type domains 1	2.2
ATP2B4	ATPase, Ca++ transporting, plasma membrane 4	2.2
COA1	cytochrome c oxidase assembly factor 1 homolog (S. cerevisiae)	2.2
C5orf56	chromosome 5 open reading frame 56	2.2
LETMD1	LETM1 domain containing 1	2.2
LRRC37B	leucine rich repeat containing 37B	2.2
ZNF79	zinc finger protein 79	2.2
TESPA1	thymocyte expressed, positive selection associated 1	2.2

TTC38	tetratricopeptide repeat domain 38	2.2
UHRF1BP1L	UHRF1 binding protein 1-like	2.2
LOC100996400 /// NOL4	uncharacterized LOC100996400 /// nucleolar protein 4	2.2
HSDL2	hydroxysteroid dehydrogenase like 2	2.2
RABEP1	rabaptin, RAB GTPase binding effector protein 1	2.2
VDR	vitamin D (1,25- dihydroxyvitamin D3) receptor	2.1
TTF2	transcription termination factor, RNA polymerase II	2.1
PRMT2	protein arginine methyltransferase 2	2.1
SNHG8 /// SNORA24	small nucleolar RNA host gene 8 (non-protein coding) /// small nucleolar RNA, H/ACA box	2.1
ISG20	interferon stimulated exonuclease gene 20kDa	2.1
LARP1B	La ribonucleoprotein domain family, member 1B	2.1
CRYBB2P1	crystallin, beta B2 pseudogene 1	2.1
RFX3	regulatory factor X, 3 (influences HLA class II expression)	2.1
FAM188A	family with sequence similarity 188, member A	2.1
LOC645513	uncharacterized LOC645513	2.1
NBN	nibrin	2.0
C7orf31	chromosome 7 open reading frame 31	2.0
MIR155 /// MIR155HG	microRNA 155 /// MIR155 host gene (non-protein coding)	2.0
ARNTL2	aryl hydrocarbon receptor nuclear translocator-like 2	2.0
SNHG1 /// SNORD22	small nucleolar RNA host gene 1 (non-protein coding) /// small nucleolar RNA, C/D box 2	2.0
IL10RB	interleukin 10 receptor, beta	2.0
BTG3	BTG family, member 3	2.0

[a]FC was calculated between the mean values of control and patient

Supplementary Table 1B. Significantly down-regulated genes in SMG9 deficient individual compared to controls.

Gene Symbol	Gene Title	FC[a]
HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	-39.2
IGHG1 /// IGHG2 /// IGHM /// IGHV4-31	immunoglobulin heavy constant gamma 1 (G1m marker) /// immunoglobulin heavy constant ga	-20.3
VIM	vimentin	-15.5
UCHL1	ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase)	-10.0
SLC12A8	solute carrier family 12 (potassium/chloride transporters), member 8	-7.8
DUSP4	dual specificity phosphatase 4	-6.1
MB21D2	Mab-21 domain containing 2	-5.6
CECR1	cat eye syndrome chromosome region, candidate 1	-5.1
COL5A1	collagen, type V, alpha 1	-4.8
LCK	lymphocyte-specific protein tyrosine kinase	-4.8
BIN1	bridging integrator 1	-4.7
ARL4C	ADP-ribosylation factor-like 4C	-4.6
CRIM1	cysteine rich transmembrane BMP regulator 1 (chordin-like)	-4.5
PLD6	phospholipase D family, member 6	-4.5
TNS3	tensin 3	-4.5
TSPO	translocator protein (18kDa)	-4.2
KDM4C	lysine (K)-specific demethylase 4C	-4.1
WARS2	tryptophanyl tRNA synthetase 2, mitochondrial	-3.5
CAPG	capping protein (actin filament), gelsolin-like	-3.3
SLC8A3	solute carrier family 8 (sodium/calcium exchanger), member 3	-3.1
NFIL3	nuclear factor, interleukin 3 regulated	-3.1

LILRB4	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member	-3.1
RNASET2	ribonuclease T2	-3.0
RHOB	ras homolog family member B	-3.0
ALOX5	arachidonate 5-lipoxygenase	-3.0
S100A11	S100 calcium binding protein A11	-3.0
EGR1	early growth response 1	-2.8
PMEPA1	prostate transmembrane protein, androgen induced 1	-2.7
TNFRSF1B	tumor necrosis factor receptor superfamily, member 1B	-2.6
TMC8	transmembrane channel-like 8	-2.6
TCF4	transcription factor 4	-2.5
KIAA1551	KIAA1551	-2.5
LOC285972	uncharacterized LOC285972	-2.4
PPAT	phosphoribosyl pyrophosphate amidotransferase	-2.4
MS4A1	membrane-spanning 4-domains, subfamily A, member 1	-2.3
LHFPL2	lipoma HMGIC fusion partner-like 2	-2.3
HLA-DPB1	major histocompatibility complex, class II, DP beta 1	-2.3
PSTPIP1	proline-serine-threonine phosphatase interacting protein 1	-2.3
GATM	glycine amidinotransferase (L-arginine:glycine amidinotransferase)	-2.2
RAPGEF5	Rap guanine nucleotide exchange factor (GEF) 5	-2.2
MBP	myelin basic protein	-2.2
ADD3	adducin 3 (gamma)	-2.2
LOC100996752 /// LOC101060261 /// NUDT4	uncharacterized LOC100996752 /// uncharacterized LOC101060261 /// nudix (nucleoside dip	-2.2
SMG9	smg-9 homolog, nonsense mediated mRNA decay factor (C. elegans)	-2.2

KCNN3	potassium intermediate/small conductance calcium-activated channel, subfamily N, member	-2.2
COTL1	coactosin-like 1 (Dictyostelium)	-2.1
NAPSB	napsin B aspartic peptidase, pseudogene	-2.1
MGC70870	C-terminal binding protein 2 pseudogene	-2.1
ANXA11	annexin A11	-2.0
KYNU	kynureninase	-2.0
KIAA0930	KIAA0930	-2.0
MAST3	microtubule associated serine/threonine kinase 3	-2.0
BTN3A3	butyrophilin, subfamily 3, member A3	-2.0
NEK6	NIMA-related kinase 6	-2.0

[a]FC was calculated between the mean values of control and patient

Table S2: List of PTC-containing genes

Gene Symbol	Gene Title	Affymetrix	Type	zygosity	refGene functio	refGene exonic function	Patient1a	Patient1b	Average Expr
PLEKHA2	pleckstrin homo	225136_at	NMD	hom	exonic;splicing	frameshift deletion	8.469	8.474	8.471
SON	SON DNA binding	214988_s_	NMD	hom	exonic	frameshift deletion	8.398	8.472	8.435
SLAIN1	SLAIN motif fam	225619_at	NMD	hom	exonic	frameshift insertion	8.323	8.438	8.381
EFTUD2	elongation fact	222398_s_	NMD	het	exonic;splicing	frameshift insertion	8.124	8.153	8.139
NOP16	NOP16 nucleolar	214011_s_	NMD	hom	exonic	frameshift insertion	8.101	8.042	8.072
UBE2NL	ubiquitin-conju	217393_x_	NMD	hom	exonic	stopgain	8.064	8.001	8.033
CLECL1	C-type lectin-l	244413_at	NMD	het	exonic	frameshift insertion	8.097	7.933	8.015
RRM2B	ribonucleotide	223342_at	NMD	het	exonic	frameshift insertion	7.207	7.655	7.431
DCAF7	DDB1 and CUL4 a	224748_at	NMD	hom	exonic;splicing	frameshift deletion	7.215	7.280	7.248
P2RX5	purinergic rece	210448_s_	NMD	het	exonic	frameshift deletion	7.160	7.222	7.191
MRPS34	mitochondrial r	218112_at	NMD	het	exonic;splicing	frameshift insertion	6.814	6.773	6.794
KMT2C	lysine (K)-spec	222415_at	NMD	het	exonic	stopgain	6.629	6.858	6.743
SENP3	SUMO1/sentrin/S	203871_at	NMD	hom	exonic;splicing	frameshift deletion	6.213	6.271	6.242
CPNE1	copine I	206918_s_	NMD	het	exonic	frameshift insertion	6.029	6.075	6.052
HEY1	hes-related fam	44783_s_a	NMD	hom	exonic	frameshift deletion	5.795	5.861	5.828
CASP7	caspase 7, apop	207181_s_	NMD	het	exonic	frameshift deletion	5.804	5.677	5.740
IQSEC1	IQ motif and Se	203907_s_	NMD	hom	exonic;splicing	frameshift deletion	5.478	5.660	5.569
SLC37A4	solute carrier	202830_s_	NMD	hom	exonic;splicing	frameshift deletion	5.628	5.499	5.564
ZNF598	zinc finger pro	225104_at	NMD	hom	exonic;splicing	frameshift deletion	5.650	5.384	5.517
EML3	echinoderm micr	203442_x_	NMD	hom	exonic	stopgain	5.260	5.260	5.260
IL17RB	interleukin 17	224156_x_	NMD	hom	exonic	stopgain	5.445	4.932	5.188
VPS13B	vacuolar protei	213243_at	NMD	het	exonic	stopgain	4.846	5.297	5.072
KMT2B	lysine (K)-spec	203419_at	NMD	hom	exonic;splicing	frameshift insertion	5.040	4.792	4.916
NEIL2	nei endonucleas	226585_at	NMD	het	exonic;splicing	frameshift deletion	4.866	4.853	4.860
PVRIG	poliovirus rece	219812_at	NMD	hom	exonic	stopgain	4.948	4.743	4.846
SLC41A3	solute carrier	224931_at	NMD	het	exonic	frameshift deletion	4.655	4.918	4.787
ZFPM1	zinc finger pro	242282_at	NMD	hom	exonic	frameshift deletion	4.764	4.764	4.764
NOP9	NOP9 nucleolar	225514_at	NMD	het	exonic	stopgain	4.614	4.724	4.669
ZBTB7B	zinc finger and	235145_at	NMD	het	exonic	stopgain	4.670	4.598	4.634
BTN3A3	butyrophilin, s	204821_at	NMD	het	exonic	stopgain	4.559	4.663	4.611
ZNF211	zinc finger pro	205437_at	NMD	het	exonic	frameshift deletion	4.464	4.623	4.544
ZFYVE19	zinc finger, FY	225843_at	NMD	het	exonic	frameshift insertion	4.574	4.505	4.539
PRDM15	PR domain conta	230777_s_	NMD	hom	exonic;splicing	frameshift insertion	4.493	4.349	4.421
MOB3C	MOB kinase acti	227066_at	NMD	hom	exonic	stopgain	4.306	4.306	4.306
CRIPAK	cysteine-rich P	228318_s_	NMD	het	exonic	frameshift deletion	3.838	4.035	3.936
SARM1	sterile alpha a	213259_s_	NMD	hom	exonic	frameshift insertion	3.757	3.337	3.547
TMEM80	transmembrane p	65630_at	NMD	het	exonic	frameshift deletion	3.515	3.515	3.515

EBLN2	endogenous Born	219906_at	NMD	het	exonic	frameshift insertion	3.266	3.266	3.266
MAPK12	mitogen-activat	206106_at	NMD	het	exonic	frameshift insertion	3.482	2.928	3.205
SSTR3	somatostatin re	1553178_a	NMD	hom	exonic	frameshift insertion	3.211	3.097	3.154
LCN10	lipocalin 10	238071_at	NMD	het	exonic;splicing	stopgain	2.920	3.388	3.154
ZNF516	zinc finger pro	203604_at	NMD	hom	exonic;splicing	frameshift deletion	2.919	3.321	3.120
ZNF117	zinc finger pro	235408_x_	NMD	het	exonic	stopgain	3.070	3.070	3.070
NLRC3	NLR family, CAR	236295_s_	NMD	hom	exonic;splicing	frameshift deletion	2.684	3.316	3.000
ZNF480	zinc finger pro	222283_at	NMD	hom	exonic	frameshift deletion	3.091	2.876	2.983
PDE4DIP	phosphodiestera	205872_x_	NMD	het	exonic	stopgain	2.916	3.002	2.959
C17orf100	chromosome 17 o	229071_at	NMD	hom	exonic	frameshift insertion	2.648	2.981	2.814
NCAM1	neural cell adh	209968_s_	NMD	hom	exonic;splicing	frameshift deletion	2.795	2.795	2.795
TUT1	terminal uridyl	218965_s_	NMD	het	exonic	frameshift deletion	2.814	2.655	2.735
SLC46A1	solute carrier	1552279_a	NMD	hom	exonic;splicing	frameshift deletion	2.634	2.746	2.690
KHDC1	KH homology dom	230055_at	NMD	het	exonic	stopgain	2.543	2.814	2.679
FUT2	fucosyltransfer	208505_s_	NMD	het	exonic	stopgain	2.683	2.526	2.604
NEK3	NIMA-related ki	213116_at	NMD	hom	exonic;splicing	frameshift insertion	2.432	2.543	2.488
GAB4	GRB2-associated	1563816_a	NMD	het	exonic;splicing	stopgain	2.424	2.504	2.464
KCNMB3	potassium large	221125_s_	NMD	het	exonic	frameshift deletion	2.336	2.411	2.373
RHPN2	rhophilin, Rho	227196_at	NMD	het	exonic	stopgain	2.283	2.384	2.334
DNAH11	dynein, axonema	1553159_a	NMD	het	exonic	stopgain	2.317	2.317	2.317
EFCAB13	EF-hand calcium	231651_at	NMD	het	exonic	stopgain	2.206	2.363	2.284
VSIG10L	V-set and immun	238654_at	NMD	hom	exonic;splicing	frameshift insertion	2.055	2.486	2.270
SRRM3	serine/arginine	235880_at	NMD	hom	exonic;splicing	frameshift insertion	2.313	2.206	2.260
PKD1L2	polycystic kidn	1559261_a	NMD	hom	exonic	frameshift deletion	2.144	2.313	2.228
CC2D2A	coiled-coil and	234936_s_	NMD	het	exonic	stopgain	2.126	2.326	2.226
GGT6	gamma-glutamylt	236225_at	NMD	hom	exonic;splicing	stopgain	2.083	2.299	2.191
APOC3	apolipoprotein	205820_s_	NMD	het	exonic;splicing	stopgain	2.205	2.130	2.168
ZNF519	zinc finger pro	1564190_x	NMD	het	exonic	stopgain	2.219	2.111	2.165
LFNG	LFNG O-fucosylp	228762_at	NMD	het	exonic	stopgain	2.303	2.027	2.165
TPSD1	tryptase delta	214568_at	NMD	het	exonic	frameshift deletion	2.141	2.141	2.141
DDIT4L	DNA-damage-indu	228057_at	NMD	het	exonic	frameshift deletion	2.006	2.264	2.135
SSPO	SCO-spondin	234349_at	NMD	hom	exonic;splicing	frameshift insertion	2.079	2.150	2.115
ATRNL1	attractin-like	1569796_s	NMD	hom	exonic	frameshift insertion	1.921	2.248	2.085
ZNF283	zinc finger pro	243188_at	NMD	het	exonic	frameshift deletion	2.079	2.079	2.079
OPRM1	opioid receptor	211359_s_	NMD	het	exonic	stopgain	2.017	2.125	2.071
GRIA3	glutamate recep	1569290_s	NMD	hom	exonic	frameshift insertion	2.012	2.113	2.063
C14orf105	chromosome 14 o	1569434_a	NMD	het	exonic	stopgain	2.177	1.893	2.035
KRTAP7-1	keratin associa	1564960_a	NMD	hom	exonic;splicing	frameshift deletion	1.913	2.122	2.018
SCARF2	scavenger recep	227557_at	NMD	hom	exonic;splicing	frameshift insertion	2.201	1.826	2.014

P2RY4	pyrimidinergic	221466_at	NMD	het	exonic	stopgain	1.862	2.165	2.014
NR2E3	nuclear recepto	208385_at	NMD	hom	exonic;splicing	frameshift deletion	2.116	1.911	2.014
P4HA3	prolyl 4-hydrox	228703_at	NMD	het	exonic	stopgain	2.138	1.886	2.012
ACTN2	actinin, alpha	203864_s_	NMD	het	exonic	frameshift deletion	2.095	1.916	2.006
ADRA2C	adrenoceptor al	206128_at	NMD	het	exonic	frameshift insertion	2.117	1.849	1.983
DHDH	dihydrodiol deh	231416_at	NMD	het	exonic	frameshift insertion	2.123	1.838	1.981
RETNLB	resistin like b	223969_s_	NMD	het	exonic	frameshift insertion	1.889	1.951	1.920
ROBO3	roundabout, axo	219550_at	NMD	het	exonic	frameshift insertion	1.858	1.972	1.915
DCHS2	dachsous cadher	220373_at	NMD	het	exonic	frameshift deletion	1.884	1.884	1.884
FAM187B	family with seq	242259_at	NMD	hom	exonic	stopgain	1.908	1.825	1.866
KRT24	keratin 24	220267_at	NMD	het	exonic	frameshift deletion	1.892	1.820	1.856
TCEB3B	transcription e	220844_at	NMD	het	exonic	stopgain	1.788	1.876	1.832
SLC38A3	solute carrier	205972_at	NMD	hom	exonic;splicing	frameshift insertion	1.995	1.654	1.825
SPATA8	spermatogenesis	231006_at	NMD	het	exonic;splicing	stopgain	1.879	1.770	1.825
KRT37	keratin 37	207649_at	NMD	het	exonic	stopgain	1.935	1.702	1.819
ALLC	allantoicase	220365_at	NMD	het	exonic	stopgain	1.747	1.817	1.782
MS4A14	membrane-spanni	229510_at	NMD	hom	exonic	frameshift deletion	1.875	1.676	1.776
ZNF233	zinc finger pro	230919_at	NMD	het	exonic	frameshift deletion	1.844	1.707	1.776
TGM4	transglutaminas	206260_at	NMD	het	exonic	stopgain	1.896	1.592	1.744
CASP12	caspase 12 (gen	1564736_a	NMD	hom	exonic	stopgain	1.691	1.774	1.733
TRPM4	transient recep	219360_s_	NMD	het	exonic	stopgain	1.809	1.657	1.733
SLC22A24	solute carrier	1553923_a	NMD	het	exonic	stopgain	1.613	1.852	1.733
C11orf40	chromosome 11 o	1553086_a	NMD	het	exonic	frameshift insertion	1.689	1.764	1.727
PRM3	protamine 3	231758_at	NMD	hom	exonic	stopgain	1.633	1.731	1.682
USP29	ubiquitin speci	220895_at	NMD	hom	exonic	stopgain	1.587	1.691	1.639
PTTG2	pituitary tumor	214557_at	NMD	het	exonic	frameshift deletion	1.638	1.638	1.638
IDO2	indoleamine 2,3	1568638_a	NMD	hom	exonic	stopgain	1.753	1.495	1.624
ZNF852	zinc finger pro	1564662_a	NMD	hom	exonic;splicing	frameshift deletion	1.556	1.639	1.598
MS4A12	membrane-spanni	220834_at	NMD	hom	exonic	stopgain	1.598	1.598	1.598
CHST15	carbohydrate (N	203066_at	NMD	hom	exonic;splicing	frameshift deletion	1.553	1.636	1.595
COL6A5	collagen, type	1553835_a	NMD	het	exonic	stopgain	1.614	1.500	1.557
H2BFM	H2B histone fam	234899_at	NMD	het	exonic	stopgain	1.553	1.553	1.553
A2ML1	alpha-2-macrogl	1553505_a	NMD	het	exonic	frameshift deletion	1.555	1.483	1.519
CD207	CD207 molecule,	220428_at	NMD	hom	exonic;splicing	frameshift insertion	1.607	1.419	1.513
SLC7A13	solute carrier	238287_at	NMD	hom	exonic	frameshift deletion	1.537	1.445	1.491
GRP	gastrin-releasi	206326_at	NMD	het	exonic;splicing	frameshift deletion	1.527	1.455	1.491
CPN2	carboxypeptidas	216223_at	NMD	het	exonic	stopgain	1.311	1.388	1.350
VWDE	von Willebrand	239552_at	NMD	het	exonic	stopgain	1.382	1.250	1.316
MAL2	mal, T-cell dif	224650_at	NMD	hom	exonic;splicing	frameshift deletion	1.269	1.355	1.312

ZNF80	zinc finger pro	207272_at	NMD	hom	exonic	stopgain		1.350	1.066	1.208

Supplemental Table 3: List of RTPCR primers used in this study

Gene name	Forward	Reverse
SMG9	GAATTGGTGTGACAGTGCCA	GCAGTTTGCGGTCATTATTGA
VIM	TCA AGGGCCAAGGCA A	ATCTGAGCCTGCAGCTCC
TNS3	ACCAGGCCCTTG ACAGG	ATG ACA TCTCCC TTC AGA AGC
EGR1	ACCTTCAACCCTCAGGCG	CTA GGCCACTGA CCA AGCTG
UCHL1	CAG TGGCCA ATAATCAAGACA	CTT CAGCAGGGTGTC CTCT
SPINT2	CCATGC CTA GGTGGTGGT	GGAGTGGTCTTCAGA ATCCTG
RORA	GTAGAAACCGCTGCCAAC A	TGGTCTGGGGAAGGCTG
VCAN	GCC TTCCAA GTTATGTTG GTG	GCC AAATGATTA CAACACAGTCTT