

Table S4. NMD efficacy predictions for SATB1 truncating variants.

(Hg19/GRCh37)g.DNA-position	g.DNA-position of introduced (downstream) stopcodon	c.DNA-position (NM_001131010.4)	Protein effect**	NMDetective A¥ (1)	NMDetective A¥ (2)	NMDetective B¥ (1)	NMDetective B¥ (2)	Conclusion based on predictions with NMDetectiveA/B	Prediction based on canonical# and non-canonical§ NMD rules
Chr3:g.18456634_18456635delCT	Chr3:g.18436407	c.607_608delAG	p.S203Ffs*49	0.63	0.45	0.65	0.41	Conflicting; NMDetectiveA/B (1): triggers NMD, NMDetectiveA/B (2): intermediate NMD efficacy.	Triggers NMD, none of (non)-canonical NMD rules applicable
Chr3:g.18436155_18436156delTC	Chr3:g.18436098	c.1004_1005delGA	p.R335Tfs*20	0.51	0.52	0.41	0.41	Intermediate NMD efficacy	Might escape from NMD. None of canonical NMD rules applicable, non-canonical long-exon rule applicable (exon 7; 454 nucleotides).
Chr3:g.18428082G>A	Chr3:g.18428082	c.1228C>T	p.R410*	0.6		0.65		Triggers NMD	Triggers NMD, none of (non)-canonical NMD rules applicable
Chr3:g.18419777delG	Chr3:g.18419762	c.1460delC	p.P487Qfs*6	0.62	0.62	0.65	0.65	Triggers NMD	Triggers NMD, none of (non)-canonical NMD rules applicable
Chr3:g.18393687C>T	Chr3:g.18393611	c.1576G>A	p.(?)	0.57	0.6	0.65	0.65	Triggers NMD	Triggers NMD, none of (non)-canonical NMD rules applicable
Chr3:g.18391077delG	Chr3:g.18390837	c.1877delC	p.P626Hfs*81	0.08	0.26	0	0	Conflicting; NMDetectiveA/B (1) and NMDetectiveB (2): escapes NMD; NMDetectiveA (2): intermediate NMD efficacy,	Escapes NMD based on canonical last exon rule
Chr3:g.18390921_18390922delCA	Chr3:g.18390797	c.2032_2033delCT	p.L678Vfs*42	0.18	0.17	0	0	Escapes NMD	Escapes NMD based on canonical last exon rule
Chr3:g.18390874G>A	Chr3:g.18390874	c.2080C>T	p.Q694*	0.2		0		Escapes NMD	Escapes NMD based on canonical last exon rule
Chr3:g.18390747delT	Chr3:g.18390726	c.2207delA	p.N736Ifs*8	0.16	0.16	0	0	Escapes NMD	Escapes NMD based on canonical last exon rule

**For frameshift mutations, scores for NMDetectiveA and NMDetectiveB were assigned both based on the genomic location of the indel (1) and based on the genomic location of the first downstream stopcodon in the new reading frame (2; first nucleotide of introduced stopcodon) (PMID: 31659324). For splice site mutations, NMDetectiveA and NMDetectiveB were assigned based on the effect predicted by spliceAI (PMID: 30661751).

¥NMDetectiveA and NMDetectiveB cut-off scores (v2):

<0.25 predicted to escape NMD

≥0.25 - ≤0.52 predicted intermediate NDM efficacy

>0.52 predicted to trigger NMD (PMID: 31659324)

#Canonical rules of NMD (PMID: 27618451):

NMD is typically not triggered when the location of the protein truncating variant is

1. less than 50 nucleotides upstream of last exon-exon junction; or
2. in the last exon.

§Non-canonical rules of NMD (PMID: 27618451):

NMD is not triggered when the location of the protein truncating variant is

1. in a very long exon (> ±400 nucleotides); or
2. within 150 nucleotides from the start codon.

\$ - predicted amino acid sequences of NMD-escaping truncating variants in SATB1

Amino acid sequence of SATB1 (NM_002971.4/NM_001131010.4) in the normal situation

MDHLNEATQGKEHSEMSNNVSDPKGPPAKIARLEQNGSPLGRGRLGSTGAKMQGVPLKHSGLMKTNLRKGTMLPVFCVVEH
YENAIEYDCKEEHAEFVLVRKDMFLNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNVPVLSYVTDAPDATVADMLQDVYHVVTLK
IQLHSCPKLEDLPPEQWSHTTVRNALKDLLKDMNQSSLAKECPLSQSMISSIVNSTYYANVSAKACQEFGRWYKHFKKTKDMMV
EMDSLSELSQQGANHVNFGQQPVPNTAEQPPSPAQLSHGSQPSVRTPLPNLHPGLVSTPISPQLVNQQLVMAQLLNQQYAVN
RLLAQQSLNQQLYNHPPPVSRSMNKPLeQQVSTNTEVSSEIYQWVRDELKRAGISQAVFARVAFNRTQGLLSEILRKEEDPKTAS
QSLVLNLRAMQNFLLQPEAERDRIYQDERERSLNAASAMGPALISTPPSRPPQVKVTATATERNKGKPNNTMNINASIYDEIQQE
MKRAKVSQALFAKVAATKSQGWLCCELLRWKEDPSPENRTLWENLSMIRRFSLPQPERDAIYEQESNAVHHHGDPRPHIIVPA
EQIQQQQQQQQQQQQQQQAPPQPQQQPQTGPRLPPRQPTVASPAESDEENRQKTRPRTKISVEALGILQSFQIDVGLYPDE
EAIQTLSAQLDLPKYTIKFFQNRYYLKHGKGLKDNSGLEVDVAEYKEEELLKDLEESVQDKNTNTLFSVKLEEEELSVEGNTDINT
DLKD

Amino acid sequence of SATB1 (NM_002971.4/NM_001131010.4) in patient 5

*Chr3:g.18428082G>A; c.1228C>T; p.R410**

MDHLNEATQGKEHSEMSNNVSDPKGPPAKIARLEQNGSPLGRGRLGSTGAKMQGVPLKHSGLMKTNLRKGTMLPVFCVVEH
YENAIEYDCKEEHAEFVLVRKDMFLNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNVPVLSYVTDAPDATVADMLQDVYHVVTLK
IQLHSCPKLEDLPPEQWSHTTVRNALKDLLKDMNQSSLAKECPLSQSMISSIVNSTYYANVSAKACQEFGRWYKHFKKTKDMMV
EMDSLSELSQQGANHVNFGQQPVPNTAEQPPSPAQLSHGSQPSVRTPLPNLHPGLVSTPISPQLVNQQLVMAQLLNQQYAVN
RLLAQQSLNQQLYNHPPPVSRSMNKPLeQQVSTNTEVSSEIYQWVRDELKRAGISQAVFARVAFNRTQGLLSEILRKEEDPKTAS
QSLVLNLRAMQNFLLQPEAERDRIYQDERERSLNAASAMGPALISTPPSRPPQVKVTATATERNKGKPNNTMNINASIYDEIQQE
MKRAKVSQALFAKVAATKSQGWLCCELLRWKEDPSPENRTLWENLSMIRRFSLPQPERDAIYEQESNAVHHHGDPRPHIIVPA
EQIQQQQQQQQQQQQQQQAPPQPQQQPQTGPRLPPRQPTVASPAESDEENRQKTRPRTKISVEALGILQSFQIDVGLYPDE
EAIQTLSAQLDLPKYTIKFFQNRYYLKHGKGLKDNSGLEVDVAEYKEEELLKDLEESVQDKNTNTLFSVKLEEEELSVEGNTDINT
DLKD

Amino acid sequence of SATB1 (NM_002971.4/NM_001131010.4) in patient 8

*Chr3:g.18391077del; c.1877del; p.P626Hfs*81*

MDHLNEATQGKEHSEMSNNVSDPKGPPAKIARLEQNGSPLGRGRLGSTGAKMQGVPLKHSGLMKTNLRKGTMLPVFCVVEH
YENAIEYDCKEEHAEFVLVRKDMFLNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNVPVLSYVTDAPDATVADMLQDVYHVVTLK
IQLHSCPKLEDLPPEQWSHTTVRNALKDLLKDMNQSSLAKECPLSQSMISSIVNSTYYANVSAKACQEFGRWYKHFKKTKDMMV
EMDSLSELSQQGANHVNFGQQPVPNTAEQPPSPAQLSHGSQPSVRTPLPNLHPGLVSTPISPQLVNQQLVMAQLLNQQYAVN
RLLAQQSLNQQLYNHPPPVSRSMNKPLeQQVSTNTEVSSEIYQWVRDELKRAGISQAVFARVAFNRTQGLLSEILRKEEDPKTAS
QSLVLNLRAMQNFLLQPEAERDRIYQDERERSLNAASAMGPALISTPPSRPPQVKVTATATERNKGKPNNTMNINASIYDEIQQE
MKRAKVSQALFAKVAATKSQGWLCCELLRWKEDPSPENRTLWENLSMIRRFSLPQPERDAIYEQESNAVHHHGDPRPHIIVPA
EQIQQQQQQQQQQQQQQQAPPQPQQQPQTGPRLPPRQPTVASPAESDEENRQKTRPRTKISVEALGILQSFQIDVGLYPDE
EAIQTLSAQLDLPKYTIKFFQNRYYLKHGKGLKDNSGLEVDVAEYKEEELLKDLEESVQDKNTNTLFSVKLEEEELSVEGNTDINT
DLKD

Amino acid sequence of SATB1 (NM_002971.4/NM_001131010.4) in patient 9 and 10 *Chr3:g.18390921_18390922del; c.2032_2033del; p.L678Vfs*42*

MDHLNEATQGKEHSEMSNNVSDPKGPPAKIARLEQNGSPLGRGRLGSTGAKMQGVPLKHSGLMKTNLRKGTMLPVFCVVEH
YENAIEYDCKEEHAEFVLVRKDMFLNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNVPVLSYVTDAPDATVADMLQDVYHVVTLK
IQLHSCPKLEDLPPEQWSHTTVRNALKDLLKDMNQSSLAKECPLSQSMISSIVNSTYYANVSAKACQEFGRWYKHFKKTKDMMV
EMDSLSELSQQGANHVNFGQQPVPNTAEQPPSPAQLSHGSQPSVRTPLPNLHPGLVSTPISPQLVNQQLVMAQLLNQQYAVN
RLLAQQSLNQQLYNHPPPVSRSMNKPLeQQVSTNTEVSSEIYQWVRDELKRAGISQAVFARVAFNRTQGLLSEILRKEEDPKTAS
QSLVLNLRAMQNFLLQPEAERDRIYQDERERSLNAASAMGPALISTPPSRPPQVKVTATATERNKGKPNNTMNINASIYDEIQQE
MKRAKVSQALFAKVAATKSQGWLCCELLRWKEDPSPENRTLWENLSMIRRFSLPQPERDAIYEQESNAVHHHGDPRPHIIVPA
EQIQQQQQQQQQQQQQQQAPPQPQQQPQTGPRLPPRQPTVASPAESDEENRQKTRPRTKISVEALGILQSFQIDVGLYPDE
EAIQTLSAQLDLPKYTIKFFQNRYYLKHGKGLKDNSGLEVDVAEYKEEELLKDLEESVQDKNTNTLFSVKLEEEELSVEGNTDINT
DLKD

Amino acid sequence of SATB1 (NM_002971.4/NM_001131010.4) in patient 11

*Chr3:g.18390874G>A; c.2080C>T; p.Q694**

MDHLNEATQGKEHSEMSNNVSDPKGPPAKIARLEQNGSPLGRGRLGSTGAKMQGVPLKHSGLMKTNLRKGTMLPVFCVVEH
YENAIEYDCKEEHAEFVLVRKDMFLNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNVPVLSYVTDAPDATVADMLQDVYHVVTLK
IQLHSCPKLEDLPPEQWSHTTVRNALKDLLKDMNQSSLAKECPLSQSMISSIVNSTYYANVSAKACQEFGRWYKHFKKTKDMMV
EMDSLSELSQQGANHVNFGQQPVPNTAEQPPSPAQLSHGSQPSVRTPLPNLHPGLVSTPISPQLVNQQLVMAQLLNQQYAVN
RLLAQQSLNQQLYNHPPPVSRSMNKPLeQQVSTNTEVSSEIYQWVRDELKRAGISQAVFARVAFNRTQGLLSEILRKEEDPKTAS
QSLVLNLRAMQNFLLQPEAERDRIYQDERERSLNAASAMGPALISTPPSRPPQVKVTATATERNKGKPNNTMNINASIYDEIQQE
MKRAKVSQALFAKVAATKSQGWLCCELLRWKEDPSPENRTLWENLSMIRRFSLPQPERDAIYEQESNAVHHHGDPRPHIIVPA
EQIQQQQQQQQQQQQQQQAPPQPQQQPQTGPRLPPRQPTVASPAESDEENRQKTRPRTKISVEALGILQSFQIDVGLYPDE
EAIQTLSAQLDLPKYTIKFFQNRYYLKHGKGLKDNSGLEVDVAEYKEEELLKDLEESVQDKNTNTLFSVKLEEEELSVEGNTDINT
DLKD

Amino acid sequence of SATB1 (NM_002971.4/NM_001131010.4) in patient 12

*Chr3:g.18390747del; c.2207del; p.N736Ifs*8*

MDHLNEATQGKEHSEMSNNVSDPKGPPAKIARLEQNGSPLGRGRLGSTGAKMQGVPLKHSGLMKTNLRKGTMLPVFCVVEH
YENAIEYDCKEEHAEFVLVRKDMFLNQLIEMALLSLGYSHSSAAQAKGLIQVGKWNVPVLSYVTDAPDATVADMLQDVYHVVTLK
IQLHSCPKLEDLPPEQWSHTTVRNALKDLLKDMNQSSLAKECPLSQSMISSIVNSTYYANVSAKACQEFGRWYKHFKKTKDMMV
EMDSLSELSQQGANHVNFGQQPVPNTAEQPPSPAQLSHGSQPSVRTPLPNLHPGLVSTPISPQLVNQQLVMAQLLNQQYAVN
RLLAQQSLNQQLYNHPPPVSRSMNKPLeQQVSTNTEVSSEIYQWVRDELKRAGISQAVFARVAFNRTQGLLSEILRKEEDPKTAS
QSLVLNLRAMQNFLLQPEAERDRIYQDERERSLNAASAMGPALISTPPSRPPQVKVTATATERNKGKPNNTMNINASIYDEIQQE
MKRAKVSQALFAKVAATKSQGWLCCELLRWKEDPSPENRTLWENLSMIRRFSLPQPERDAIYEQESNAVHHHGDPRPHIIVPA
EQIQQQQQQQQQQQQQQQAPPQPQQQPQTGPRLPPRQPTVASPAESDEENRQKTRPRTKISVEALGILQSFQIDVGLYPDE
EAIQTLSAQLDLPKYTIKFFQNRYYLKHGKGLKDNSGLEVDVAEYKEEELLKDLEESVQDKNTNTLFSVKLEEEELSVEGNTDINT
DLKD