Supplementary Table 2. Summary of MaPSy validation in patient samples

HGMD or SNP id	*gene*	hg19 position	nucleotide change	AA change	MaPSy result	validation result	Patient tissue/ RNA source	class	Reference / source
CM990215	*ATM*	chr11:108186638:+	c.6095G>A	p.R2032K	loss of splicing *in vivo* (0 fold) and *in vitro* (0 fold)	exon 44 skipping	lymphoblastoid cell line	TP	[46]
CM980147	*ATM*	chr11:108183151:+	c.5932G>T	p.E1978*	loss of splicing *in vivo* (0 fold) and *in vitro* (0 fold), but neither were significant due to low counts	exon 42 skipping	lymphoblastoid cell line	FN	[46]
CM940140	*ATP7A*	chrX:77266736:+	c.1933C>T	p.R645*	loss of splicing *in vivo* (0.4 fold) and *in vitro* (0.4 fold)	exon 8 skipping	fibroblasts	TP	Coriell
CM994455	*BRCA1*	chr17:41258474:-	c.211A>G	p.R71G	loss of splicing *in vivo* (0 fold) and *in vitro* (0.5 fold)	exon 5 skipping	T-lymphocytes	TP	[47]
CM950153	*BRCA1*	chr17:41215920:-	c.5123C>A	p.A1708E	loss of splicing *in vivo* (0.3 fold) and *in vitro* (0.3 fold)	exon 18 skipping	T-lymphocytes	TP	[47]
CM980231	*BRCA1*	chr17:41215963:-	c.5080G>T	p.E1694*	loss of splicing *in vivo* (0.03 fold) and *in vitro* (0.1 fold)	exon 18 skipping	T-lymphocytes	TP	[48]
CM101977	*BRCA1*	chr17:41267746:-	c.131G>A	p.C44Y	no change	no change	lymphoblastoid cell line	TN	Amanda Toland
CM053138	*BRCA2*	chr13:32921033:+	c.673A>G	p.T225A	no change	no change	T-lymphocytes	TN	[49]
CM042309	*BRCA2*	chr13:32903621:+	c.7007G>A	p.R2336H	loss of splicing *in vivo* (0.03 fold) and *in vitro* (0.03 fold)	exon 13 skipping	T-lymphocytes	TP	[47]
CM051414	*COL1A1*	chr9:130579482:-	c.608G>T	p.G203V	no change	no change	fibroblast	TN	Joan Marini
CM051415	*COL1A1*	chr9:130579482:-	c.634G>A	p.G212R	Loss of splicing *in vivo* (0.1 fold) and *in vitro* (0.2 fold)	no change	fibroblast	FP	Joan Marini
CM111552	*ENG*	chr9:130579482:-	c.1687G>T	p.E563*	loss of splicing *in vivo* (0.38 fold) and *in vitro* (0.2 fold)	loss of exon 13	T-lymphocytes	TP	Authors
CM950449	*FBN1*	chr15:48729559:-	c.6339T>G	p.Y2113*	loss of splicing *in vivo* (0.04 fold) and *in vitro* (0.2 fold)	exon 51 skipping	fibroblasts	TP	[50]
CM930246	*FBN1*	chr15:48758037:-	c.4766G>T	p.C1589F	loss of splicing *in vivo* (0.3 fold) and *in vitro* (0.6 fold)	no change	fibroblasts	FP	Coriell
CM910169	*GALT*	chr9:34648167:+	c.563A>G	p.Q188R	loss of splicing *in vivo* (0.01 fold) and *in vitro* (0.4 fold)	missplicing	lymphoblastoid cell line	TP	Coriell
CM002632	*HPRT1*	chrX:133632695:+	c.590A>T	p.E197V	loss of splicing *in vivo* (0 fold) and *in vitro* (0.3 fold)	exon 8 skipping	lymphoblastoid cell line	TP	[51]
CM002633	*HPRT1*	chrX:133632707:+	c.602A>T	p.D201V	loss of splicing *in vivo* (0.2 fold) and *in vitro* (0.3 fold)	exon 8 skipping	T-lymphocytes	TP	[51]
CM004361	*HPRT1*	chrX:133632674:+	c.569G>A	p.G190E	no change	no change	lymphoblastoid cell line	TN	[51]
CM920361	*HPRT1*	chrX:133632706:+	c.601G>A	p.D201N	no change	no change	lymphoblastoid cell line	TN	[51]
CM920362	*HPRT1*	chrX:133632706:+	c.601G>T	p.D201Y	loss of splicing *in vivo* (0.18 fold) but not *in vitro*	no change	lymphoblastoid cell line and T lymphocytes	TN	[51]
CM022413	*ITPA*	chr20:3193842:+	c.94C>A	p.P32T	loss of splicing *in vivo* (0.1 fold) and *in vitro* (0.3 fold)	exon 3 skipping	lymphoblastoid cell line and T-lymphocytes	TP	Coriell [52]
CM045463	*MLH1*	chr3:37053590:+	c.677G>A	p.R226Q	loss of splicing *in vivo* (0 fold) and *in vitro* (0.01 fold)	loss of exon 8	whole blood	TP	[53]
CM960965	*MLH1*	chr3:37045935:+	c.350C>T	p.T117M	no change	no change	lymphoblastoid cell line	TN	[54]
CM068362	*MLH1*	chr3:37045923:+	c.338T>A	p.V113D	loss of splicing *in vivo* (0.6 fold) but no change *in vitro*	no change	lymphoblastoid cell line	TN	[54]
CM000176	*MLH1*	chr3:37053550:+	c.637G>A	p.V213M	loss of splicing *in vivo* (0.6 fold) and *in vitro* (0.65 fold)	no change	lymphoblastoid cell line	FP	[54]
CM041068	*PSEN1*	chr14:73653628:+	c.548G>T	p.G1883V	loss of splicing *in vivo* (0.3 fold) and *in vitro* (0.1 fold)	brain-specific loss of exons 6 and 7	frontal cortex (post mortem)	TP	[55]
CM064173	*PSEN1*	chr14:73673096:+	c.871A>C	p.T291P	no change	exon 9 skipping	Whole blood	FN	[56]
CM054825	*PSEN1*	chr14:73653589:+	c.509C>T	pS170F	no change *in vivo* but loss of splicing *in vitro* (0.5 fold)	no change	frontal cortex and cerebellum (post mortem)	TN	Alison Goate
rs2295682	*RBM23*	chr14:23374862:-	c.408G>A	p.R136	loss of splicing *in vivo* (0.3 fold) and *in vitro* (0.5 fold)	exon 6 skipping	lymphoblastoid cell line	TP	[57]
CM075012	*SLC12A3*	chr16:56902212:+	c.433C>T	p.R145C	loss of splicing *in vivo* (0.4 fold) and *in vitro* (0.6 fold) but neither were significant	exon 7 skipping	whole blood	FN	[58]
CM115034	*TAZ*	chrX:153647918:+	c.497T>A	p.L166*	loss of splicing *in vivo* (0.3 fold) and *in vitro* (0.5 fold)	exon 6 and 7 skipping	lymphoblastoid cell line	TP	Coriell
CM067040	*TAZ*	chrX:153641901:+	c.367C>T	p.R123*	loss of splicing *in vivo* (0.6 fold) and *in vitro* (0.4 fold)	intron retention	lymphoblastoid cell line	TP	Coriell

Supplementary Table 4. ENCODE Accession numbers used for the purpose of validating MaPSy results

ENCODE Accession Number
SRR307897
SRR307898
SRR307901
SRR307902
SRR307903
SRR307904
SRR307905
SRR307906
SRR307907
SRR307908
SRR307911
SRR307912
SRR307920
SRR307926
SRR307927
SRR307932
SRR307933
SRR315301
SRR315302
SRR315313
SRR315314
SRR315315
SRR315316
SRR315327
SRR315328
SRR315329
SRR315330
SRR315331
SRR315336
SRR315337
SRR534301
SRR534302
SRR534309
SRR534310
SRR534319
SRR534320
SRR534321
SRR534322
SRR534323
SRR534324
SRR545695
SRR545696
SRR545697
SRR545698
SRR545699
SRR545700

References

46. Teraoka, S.N. *et al.* Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences. *American Journal of Human Genetics* **64**, 1617-31 (1999).
47. Sanz, D.J. *et al.* A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients. *Clin Cancer Res* **16**, 1957-67 (2010).
48. Mazoyer, S. *et al.* A BRCA1 nonsense mutation causes exon skipping. *Am J Hum Genet* **62**, 713-5 (1998).
49. Sharp, A., Pichert, G., Lucassen, A. & Eccles, D. RNA analysis reveals splicing mutations and loss of expression defects in MLH1 and BRCA1. *Hum Mutat* **24**, 272 (2004).
50. Dietz, H.C. *et al.* The skipping of constitutive exons in vivo induced by nonsense mutations. *Science* **259**, 680-3 (1993).
51. Tu, M., Tong, W., Perkins, R. & Valentine, C.R. Predicted changes in pre-mRNA secondary structure vary in their association with exon skipping for mutations in exons 2, 4, and 8 of the Hprt gene and exon 51 of the fibrillin gene. *Mutat Res* **432**, 15-32 (2000).
52. Arenas, M., Duley, J., Sumi, S., Sanderson, J. & Marinaki, A. The ITPA c.94C>A and g.IVS2+21A>C sequence variants contribute to missplicing of the ITPA gene. *Biochim Biophys Acta* **1772**, 96-102 (2007).
53. Pagenstecher, C. *et al.* Aberrant splicing in MLH1 and MSH2 due to exonic and intronic variants. *Human Genetics* **119**, 9-22 (2006).
54. Auclair, J. *et al.* Systematic mRNA analysis for the effect of MLH1 and MSH2 missense and silent mutations on aberrant splicing. *Hum Mutat* **27**, 145-54 (2006).
55. Dermaut, B. *et al.* A novel presenilin 1 mutation associated with Pick's disease but not beta-amyloid plaques. *Ann Neurol* **55**, 617-26 (2004).
56. Dumanchin, C. *et al.* Biological effects of four PSEN1 gene mutations causing Alzheimer disease with spastic paraparesis and cotton wool plaques. *Hum Mutat* **27**, 1063 (2006).
57. Hull, J. *et al.* Identification of common genetic variation that modulates alternative splicing. *PLoS Genet* **3**, e99 (2007).
58. Riveira-Munoz, E. *et al.* Transcriptional and functional analyses of SLC12A3 mutations: new clues for the pathogenesis of Gitelman syndrome. *J Am Soc Nephrol* **18**, 1271-83 (2007).