Figure S1: The cumulative distribution of 5’ splice site score differences caused by 5’ splice site SNPs in the seven HapMap lymphoblastoid cell lines.

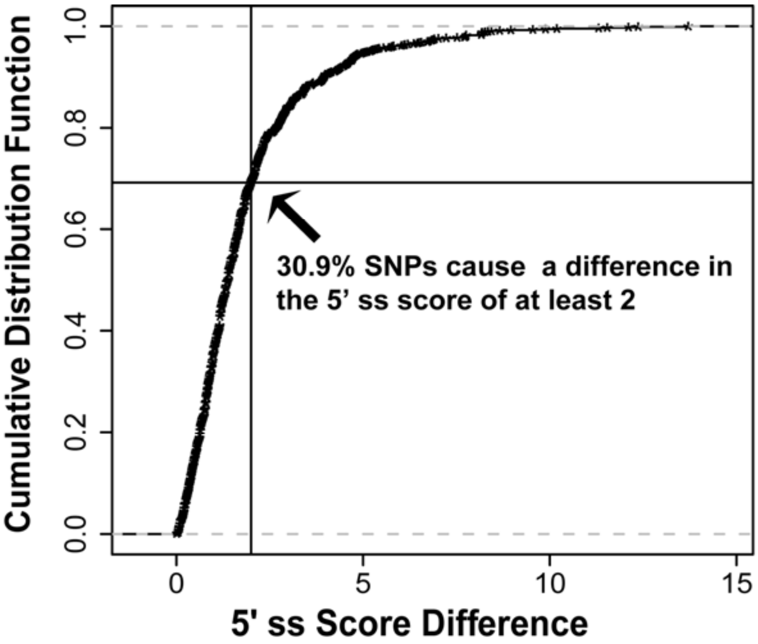


Figure S2: Minigene experiments confirm the causal impact of 5'splice site SNPs on splicing of five exons.

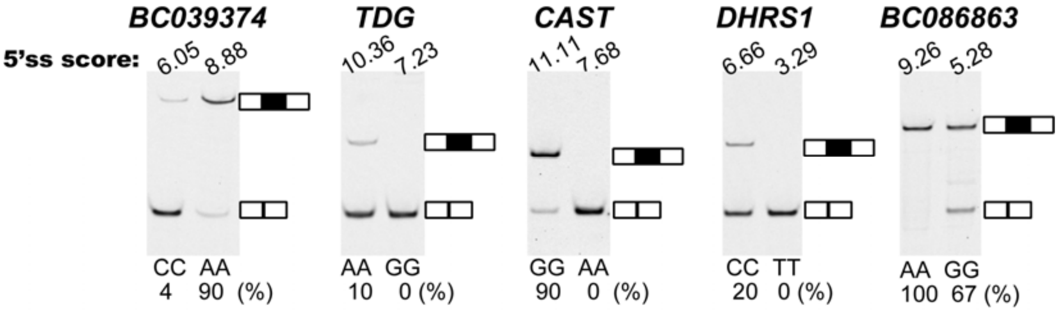
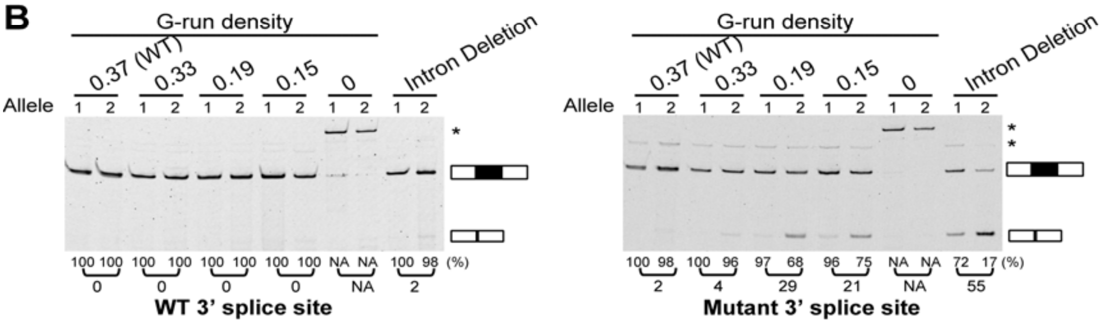


Figure S3: Results of *TRIM62* minigene experiments by site-directed G-to-A disruptions of poly-G runs. (A) Schematic diagrams of mutant minigene constructs with serial disruptions of poly-G runs. (B) Disruptions of poly-G runs in the minigene constructs containing the wildtype 3’ splice site have no impact on exon splicing (left panel). Disruptions of poly-G runs in the minigene constructs containing the mutant (weakened) 3’ splice site significantly reduce the exon inclusion level of the *TRIM62* exon.

A

G-run density_0.37(WT)	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
G-run density_0.33	--	G2	G3	G4	G5	G6	G7	G8	G9	G10
G-run density_0.19	--	G2	--	--	--	--	G7	G8	G9	G10
G-run density_0.15	--	--	--	--	--	--	G7	G8	G9	G10
G-run density_0	--	--	--	--	--	--	--	--	--	--



WT 3’ splice site

Mutant 3’ splice site

Figure S4: The effects of siRNA knockdown of *hnRNP H (HNRNPH1)* on the splicing of the *TRIM62* minigene constructs. (A) The mRNA concentrations of *hnRNP H* after treatments by siRNA for *hnRNP H* or by control siRNA. Error bars represent the standard error of the mean of the mRNA concentration estimated by four replicate treatments. (B) Western blot analysis of HNRNPH1 protein level after treatments by siRNA for *hnRNP H* or by control siRNA. (C) The siRNA knockdown of *hnRNP H* affects the splicing of the *TRIM62* minigene construct with the mutant (weakened) 3' splice site, but not the splicing of the minigene construct with the wildtype 3' splice site.

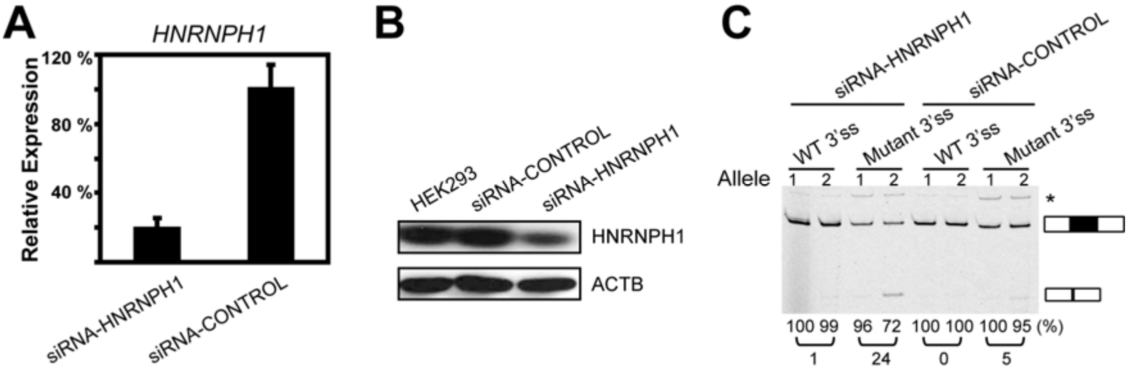


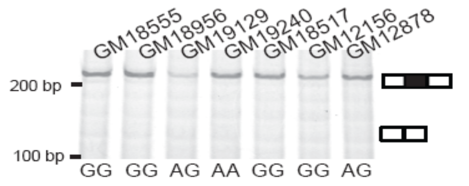
Table S1: Information of seven HapMap lymphoblastoid cell lines.

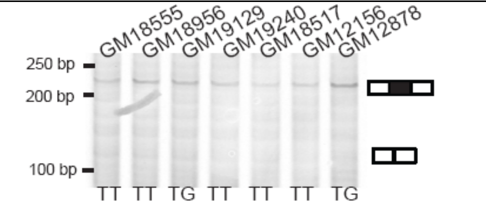
Individual ID	**Catalog ID**[a]	**Ancestry**	**Source of SNP data**
NA18555	GM18555	CHB (Han Chinese in Beijing, China)	HapMap, Exome sequencing
NA18956	GM18956	JPT (Japanese in Tokyo, Japan)	HapMap, Exome sequencing
NA19129	GM19129[b]	YRI (Yoruba in Ibadan, Nigeria)	HapMap, Exome sequencing
NA19240	GM19240[b]	YRI	HapMap, Exome sequencing
NA18517	GM18517[b]	YRI	HapMap, Exome sequencing
NA12156	GM12156[b]	CEU (U.S. Utah residents with ancestry from northern and western Europe)	HapMap, Exome sequencing
NA12878	GM12878[b]	CEU	HapMap, Exome sequencing

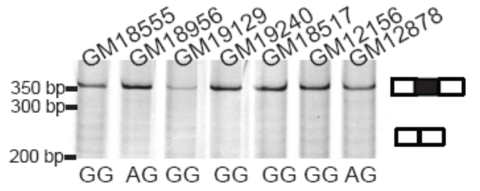
a. Samples’ catalog IDs at the Coriell Institute.

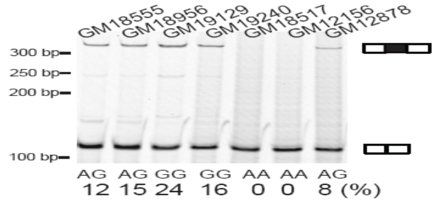
b. Affymetrix Exon 1.0 array data are available for these samples (see Materials and Methods).

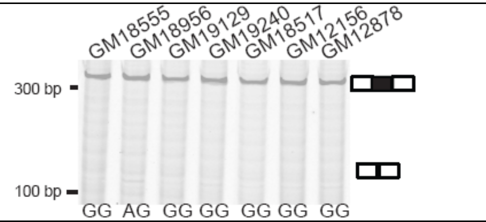
Table S2: Fluorescently labeled RT-PCR results of 84 exons in seven HapMap lymphoblastoid cell lines. The expected positions of the exon inclusion and skipping PCR products are indicated to the right of the gel. Exons whose 5' splice site SNPs cause splicing differences among seven individuals are highlighted in red.

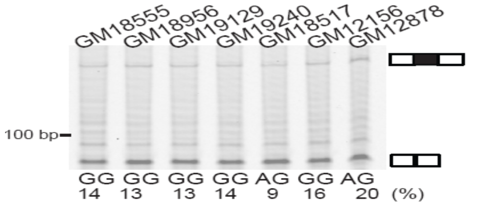
Gene Symbol	*PKN2*	
Exon Coordinates (hg18)	+chr1:89045800-89045875	
SNP ID	rs6671949	
Predicted Size (Skipping/Inclusion)	140/216	

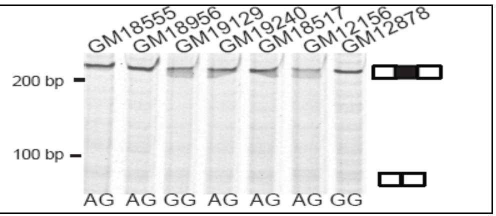
Gene Symbol	*TBC1D4*	
Exon Coordinates (hg18)	-chr13:74813262-74813372	
SNP ID	rs9565152	
Predicted Size (Skipping/Inclusion)	125/236	

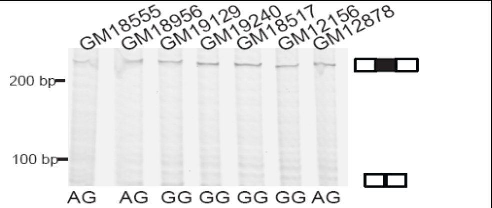
Gene Symbol	*AX747187*	
Exon Coordinates (hg18)	+chr12:95498506-95498598	
SNP ID	rs17025569	
Predicted Size (Skipping/Inclusion)	276/369	

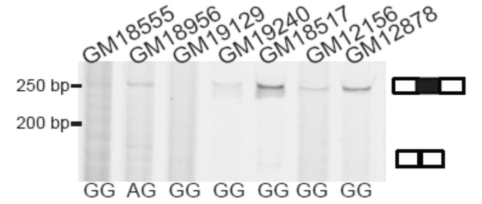
Gene Symbol	*C3orf31*	
Exon Coordinates (hg18)	-chr3:11861365-11861570	
SNP ID	rs392621	
Predicted Size (Skipping/Inclusion)	112/318	

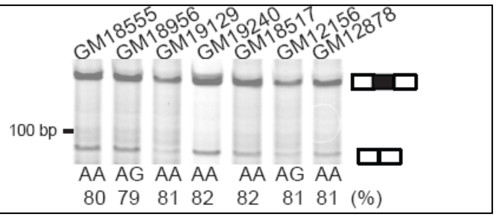
Gene Symbol	*IMMT*	
Exon Coordinates (hg18)	-chr2:86260067-86260256	
SNP ID	rs17027011	
Predicted Size (Skipping/Inclusion)	139/329	

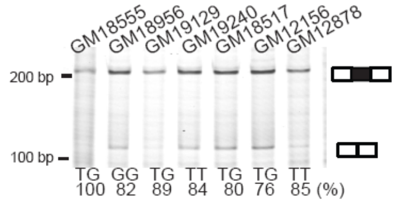
Gene Symbol	*MMAB*	
Exon Coordinates (hg18)	-chr12:108483431-108483517	
SNP ID	rs12309115	
Predicted Size (Skipping/Inclusion)	71/158	

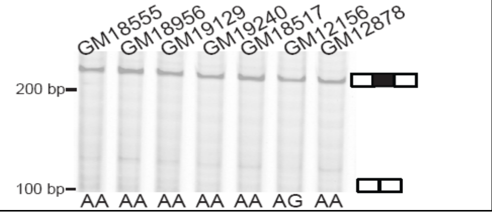
Gene Symbol	*UNC13D*	
Exon Coordinates (hg18)	-chr17:71342596-71342739	
SNP ID	rs17581728	
Predicted Size (Skipping/Inclusion)	88/232	

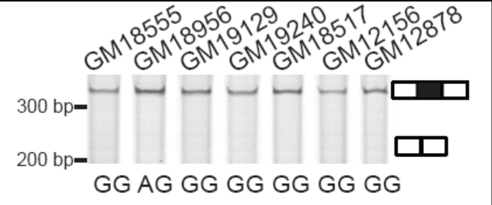
Gene Symbol	*FTSJ3*	
Exon Coordinates (hg18)	-chr17:59255176-59255331	
SNP ID	rs2584627	
Predicted Size (Skipping/Inclusion)	82/238	

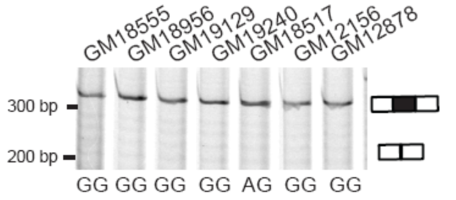
Gene Symbol	*SLC22A4*	
Exon Coordinates (hg18)	+chr5:131695348-131695442	
SNP ID	rs2304081	
Predicted Size (Skipping/Inclusion)	160/255	

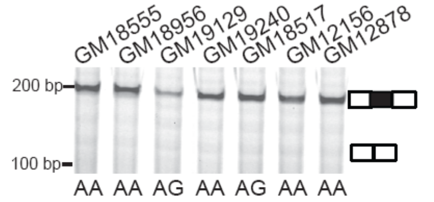
Gene Symbol	*TBL1XR1*	
Exon Coordinates (hg18)	-chr3:178247968-178248031	
SNP ID	rs3749234	
Predicted Size (Skipping/Inclusion)	83/147	

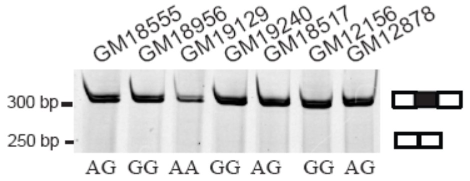
Gene Symbol	*PTPN14*	
Exon Coordinates (hg18)	-chr1:212637879-212637967	
SNP ID	rs11580603	
Predicted Size (Skipping/Inclusion)	111/200	

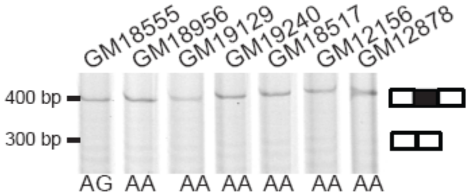
Gene Symbol	*KIF18B*	
Exon Coordinates (hg18)	-chr17:40364496-40364623	
SNP ID	rs17546822	
Predicted Size (Skipping/Inclusion)	101/229	

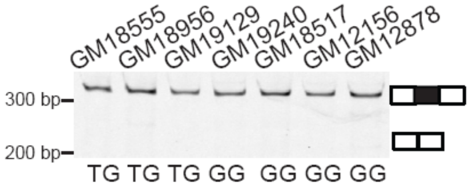
Gene Symbol	*PFAS*	
Exon Coordinates (hg18)	+chr17:8099848-8099953	
SNP ID	rs7221716	
Predicted Size (Skipping/Inclusion)	230/336	

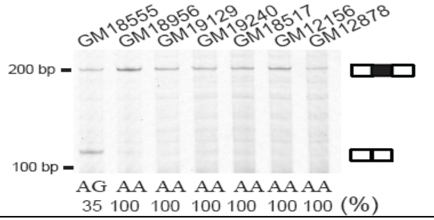
Gene Symbol	*ZNRD1*	
Exon Coordinates (hg18)	+chr6:30137980-30138089	
SNP ID	rs9261269	
Predicted Size (Skipping/Inclusion)	212/322	

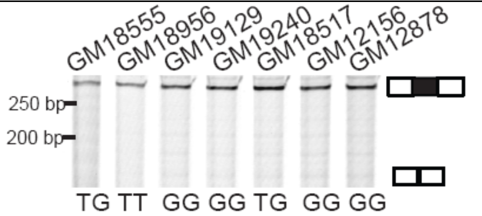
Gene Symbol	*MPHOSPH9*	
Exon Coordinates (hg18)	-chr12:122211281-122211349	
SNP ID	rs1260317	
Predicted Size (Skipping/Inclusion)	133/202	

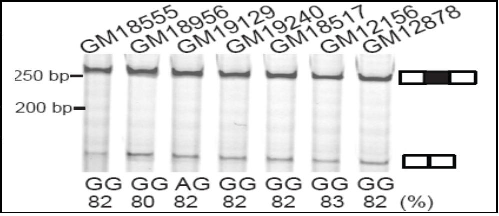
Gene Symbol	*DST*	
Exon Coordinates (hg18)	-chr6:56534895-56534957	
SNP ID	rs9367689	
Predicted Size (Skipping/Inclusion)	252/315	

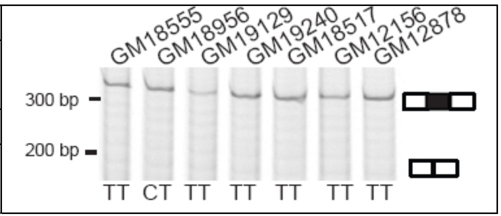
Gene Symbol	*MTF1*	
Exon Coordinates (hg18)	-chr1:38061969-38062071	
SNP ID	rs12751325	
Predicted Size (Skipping/Inclusion)	295/398	

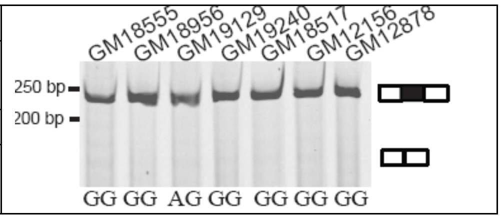
Gene Symbol	*TRIM62*	
Exon Coordinates (hg18)	-chr1:33403659-33403754	
SNP ID	rs2306257	
Predicted Size (Skipping/Inclusion)	217/313	

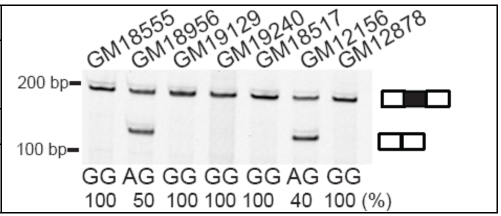
Gene Symbol	*BC086863*	
Exon Coordinates (hg18)	-chr1:217389904-217389996	
SNP ID	rs12079503	
Predicted Size (Skipping/Inclusion)	109/202	

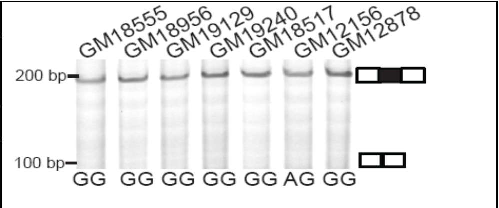
Gene Symbol	*DNASE1L3*	
Exon Coordinates (hg18)	-chr3:58161764-58161876	
SNP ID	rs3772985	
Predicted Size (Skipping/Inclusion)	167/280	

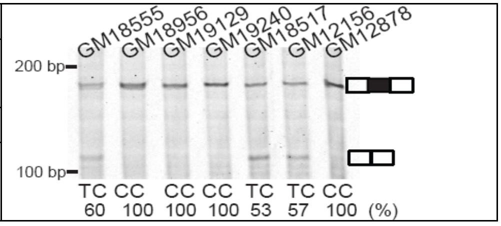
<table>
<tr><td>Gene Symbol</td><td><i>EIF4G1</i></td><td rowspan="4"> </td></tr>
<tr><td>Exon Coordinates (hg18)</td><td>+chr3:185531954-185532088</td></tr>
<tr><td>SNP ID</td><td>rs16858676</td></tr>
<tr><td>Predicted Size (Skipping/Inclusion)</td><td>121/256</td></tr>
</table>

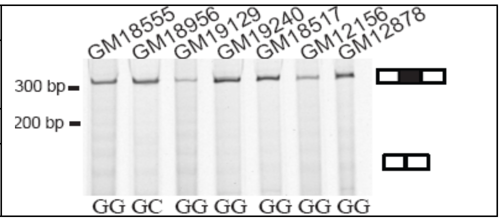
<table>
<tr><td>Gene Symbol</td><td><i>UTP15</i></td><td rowspan="4"> </td></tr>
<tr><td>Exon Coordinates (hg18)</td><td>+chr5:72904070-72904205</td></tr>
<tr><td>SNP ID</td><td>rs16870610</td></tr>
<tr><td>Predicted Size (Skipping/Inclusion)</td><td>181/317</td></tr>
</table>

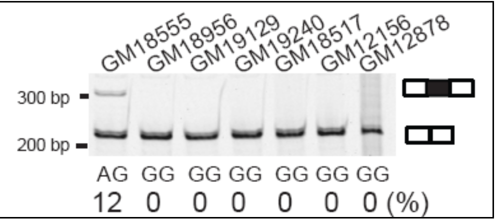
<table>
<tr><td>Gene Symbol</td><td><i>ATL3</i></td><td rowspan="4"> </td></tr>
<tr><td>Exon Coordinates (hg18)</td><td>-chr11:63157074-63157145</td></tr>
<tr><td>SNP ID</td><td>rs17656941</td></tr>
<tr><td>Predicted Size (Skipping/Inclusion)</td><td>172/244</td></tr>
</table>

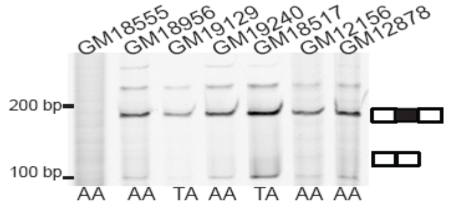
<table>
<tr><td>Gene Symbol</td><td><i>CAST</i></td><td rowspan="4"> </td></tr>
<tr><td>Exon Coordinates (hg18)</td><td>+chr5:96102205-96102243</td></tr>
<tr><td>SNP ID</td><td>rs7724759</td></tr>
<tr><td>Predicted Size (Skipping/Inclusion)</td><td>155/194</td></tr>
</table>

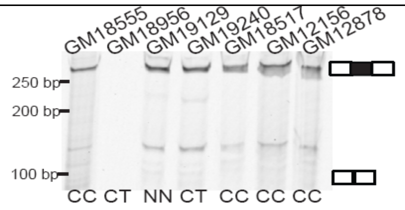
<table>
<tr><td>Gene Symbol</td><td><i>DHX38</i></td><td rowspan="4"> </td></tr>
<tr><td>Exon Coordinates (hg18)</td><td>+chr16:70700815-70700910</td></tr>
<tr><td>SNP ID</td><td>rs17666927</td></tr>
<tr><td>Predicted Size (Skipping/Inclusion)</td><td>100/196</td></tr>
</table>

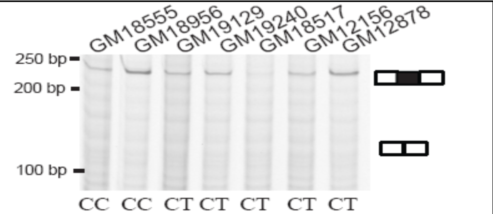
<table>
<tr><td>Gene Symbol</td><td><i>DHRS1</i></td><td rowspan="4"> </td></tr>
<tr><td>Exon Coordinates (hg18)</td><td>-chr14:23830602-23830671</td></tr>
<tr><td>SNP ID</td><td>rs10134537</td></tr>
<tr><td>Predicted Size (Skipping/Inclusion)</td><td>109/179</td></tr>
</table>

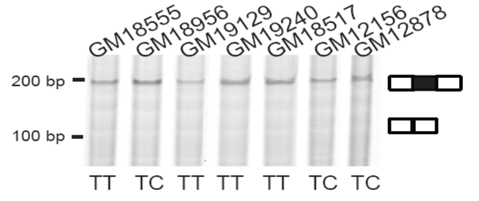
<table>
<tr><td>Gene Symbol</td><td><i>GALC</i></td><td rowspan="4"> </td></tr>
<tr><td>Exon Coordinates (hg18)</td><td>-chr14:87477492-87477655</td></tr>
<tr><td>SNP ID</td><td>rs448805</td></tr>
<tr><td>Predicted Size (Skipping/Inclusion)</td><td>142/306</td></tr>
</table>

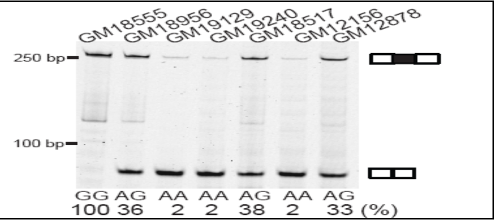
Gene Symbol	*TDG*	
Exon Coordinates (hg18)	+chr12:102890868-102890941	
SNP ID	rs17035056	
Predicted Size (Skipping/Inclusion)	229/303	

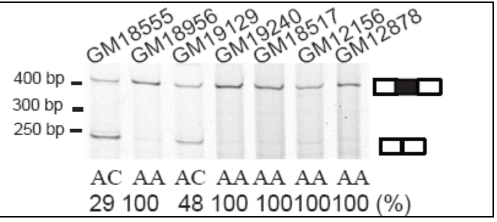
Gene Symbol	*MOXD1*	
Exon Coordinates (hg18)	-chr6:132683461-132683520	
SNP ID	rs9493285	
Predicted Size (Skipping/Inclusion)	120/180	

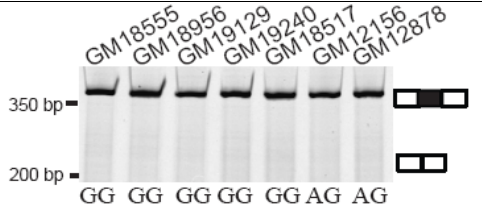
Gene Symbol	*FAM129C*	
Exon Coordinates (hg18)	+chr19:17509185-17509352	
SNP ID	rs8107859	
Predicted Size (Skipping/Inclusion)	102/270	

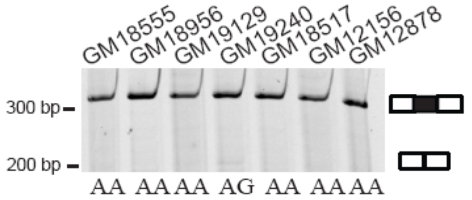
Gene Symbol	*RAP1GAP2*	
Exon Coordinates (hg18)	+chr17:2815584-2815687	
SNP ID	rs12950923	
Predicted Size (Skipping/Inclusion)	128/232	

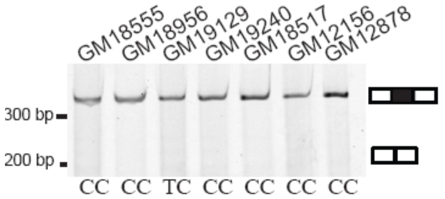
Gene Symbol	*NKAP*	
Exon Coordinates (hg18)	-chrX:118949927-118950002	
SNP ID	rs194302	
Predicted Size (Skipping/Inclusion)	121/197	

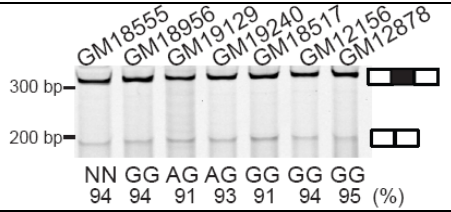
Gene Symbol	*HMSD*	
Exon Coordinates (hg18)	+chr18:59771568-59771741	
SNP ID	rs9945924	
Predicted Size (Skipping/Inclusion)	77/251	

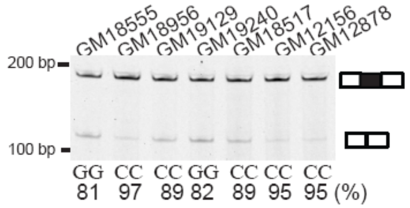
Gene Symbol	*BC039374*	
Exon Coordinates (hg18)	+chr2:97684049-97684199	
SNP ID	rs11894651	
Predicted Size (Skipping/Inclusion)	247/398	

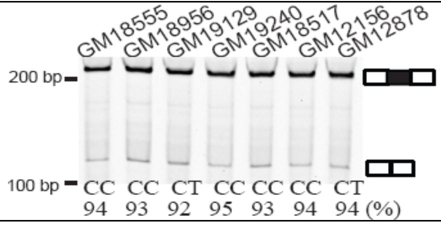
Gene Symbol	*PSAP*	
Exon Coordinates (hg18)	-chr10:73249228-73249385	
SNP ID	rs11000016	
Predicted Size (Skipping/Inclusion)	223/381	

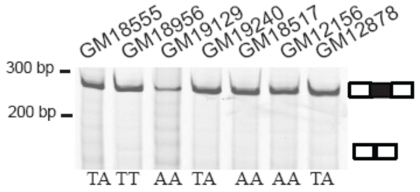
Gene Symbol	*IGF2R*	
Exon Coordinates (hg18)	+chr6:160399045-160399146	
SNP ID	rs8191828	
Predicted Size (Skipping/Inclusion)	203/305	

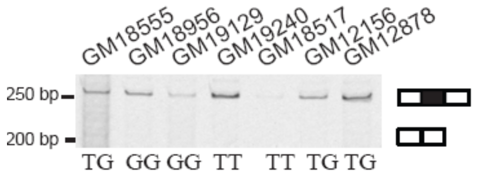
Gene Symbol	*POLE*	
Exon Coordinates (hg18)	-chr12:131724673-131724798	
SNP ID	rs5744973	
Predicted Size (Skipping/Inclusion)	206/332	

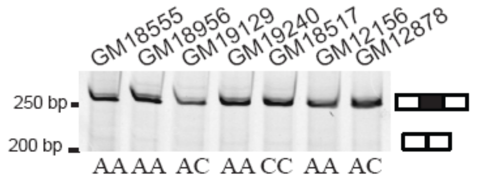
Gene Symbol	*AP2S1*	
Exon Coordinates (hg18)	-chr19:52034562-52034675	
SNP ID	rs312186	
Predicted Size (Skipping/Inclusion)	192/306	

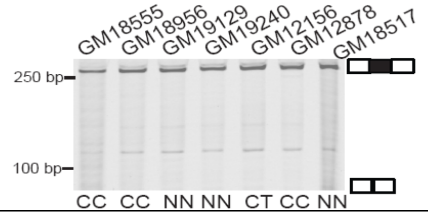
Gene Symbol	*DHRSX*	
Exon Coordinates (hg18)	-chrX:2336786-2336854	
SNP ID	rs5939175;rs73632310	
Predicted Size (Skipping/Inclusion)	119/188	

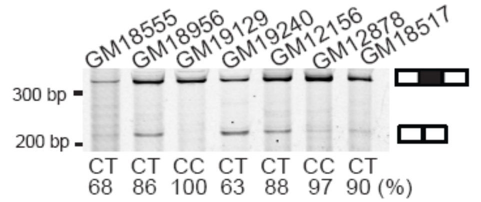
Gene Symbol	*NCKAP1L*	
Exon Coordinates (hg18)	+chr12:53197584-53197693	
SNP ID	rs2270581	
Predicted Size (Skipping/Inclusion)	109/219	

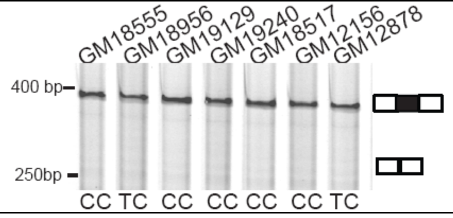
Gene Symbol	*HACE1*	
Exon Coordinates (hg18)	-chr6:105351223-105351324	
SNP ID	rs6918700	
Predicted Size (Skipping/Inclusion)	178/280	

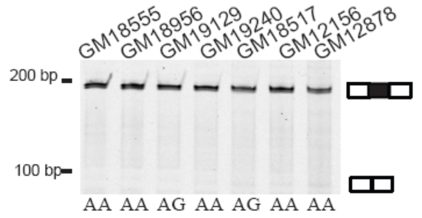
Gene Symbol	*CAPN14*	
Exon Coordinates (hg18)	-chr2:31264688-31264724	
SNP ID	rs1465972	
Predicted Size (Skipping/Inclusion)	217/254	

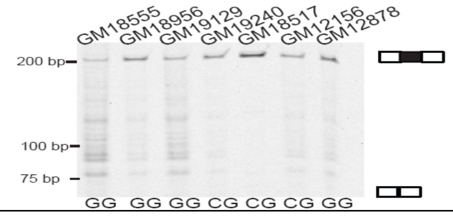
Gene Symbol	*OSBPL3*	
Exon Coordinates (hg18)	-chr7:24877886-24877939	
SNP ID	rs6952815	
Predicted Size (Skipping/Inclusion)	217/271	

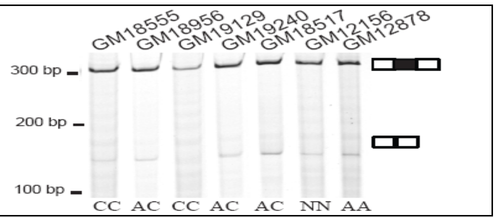
Gene Symbol	KIAA0556	
Exon Coordinates (hg18)	+chr16:27617150-27617322	
SNP ID	rs12920169	
Predicted Size (Skipping/Inclusion)	91/264	

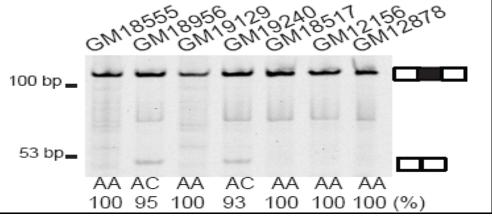
Gene Symbol	*EVC*	
Exon Coordinates (hg18)	+chr4:5797832-5797969	
SNP ID	rs2286343	
Predicted Size (Skipping/Inclusion)	214/352	

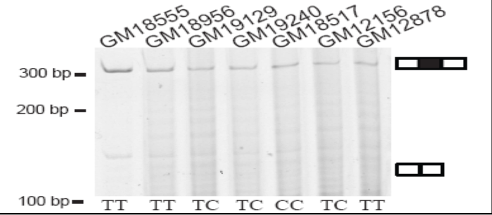
Gene Symbol	*TLN1*	
Exon Coordinates (hg18)	-chr9:35702002-35702121	
SNP ID	rs2295795	
Predicted Size (Skipping/Inclusion)	271/391	

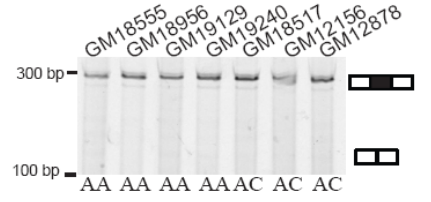
Gene Symbol	*KIAA1731*	
Exon Coordinates (hg18)	+chr11:93048385-93048480	
SNP ID	rs10831088	
Predicted Size (Skipping/Inclusion)	98/194	

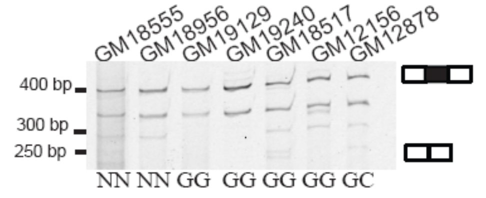
Gene Symbol	*PPFIA3*	
Exon Coordinates (hg18)	+chr19:54328306-54328440	
SNP ID	rs2287761	
Predicted Size (Skipping/Inclusion)	67/202	

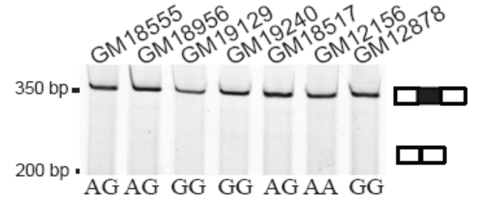
Gene Symbol	*PBK*	
Exon Coordinates (hg18)	-chr8:27736499-27736641	
SNP ID	rs727813	
Predicted Size (Skipping/Inclusion)	166/309	

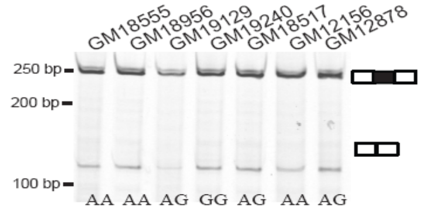
Gene Symbol	*AFTPH*	
Exon Coordinates (hg18)	+chr2:64650261-64650317	
SNP ID	rs2287531	
Predicted Size (Skipping/Inclusion)	48/105	

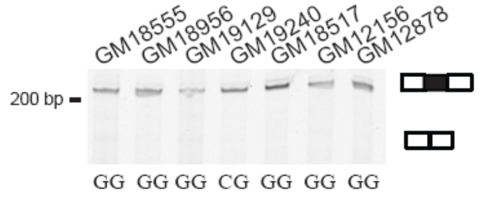
Gene Symbol	*NPC1*	
Exon Coordinates (hg18)	-chr18:19373773-19373963	
SNP ID	rs1140458	
Predicted Size (Skipping/Inclusion)	116/309	

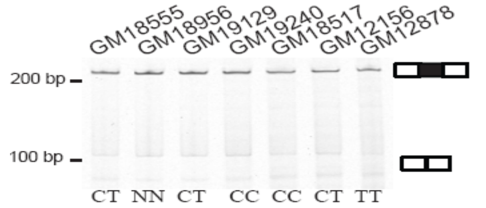
Gene Symbol	*CNTROB*	
Exon Coordinates (hg18)	+chr17:7788519-7788681	
SNP ID	rs11650083	
Predicted Size (Skipping/Inclusion)	135/295	

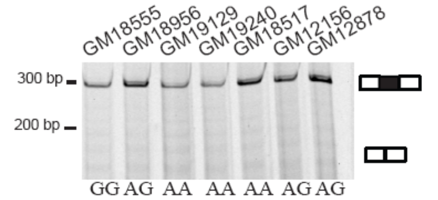
Gene Symbol	*AX746648*	
Exon Coordinates (hg18)	-chr7:155926685-155926835	
SNP ID	rs2272194	
Predicted Size (Skipping/Inclusion)	243/394	

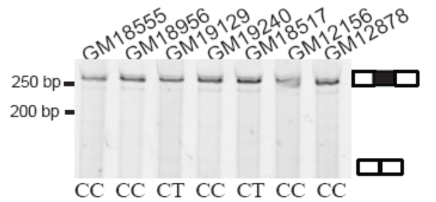
Gene Symbol	*CEPT1*	
Exon Coordinates (hg18)	+chr1:111527607-111527732	
SNP ID	rs694180	
Predicted Size (Skipping/Inclusion)	237/363	

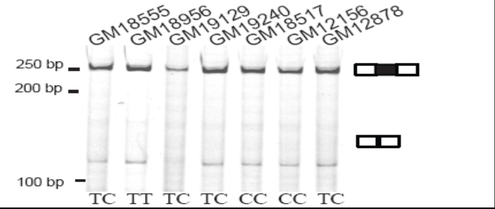
Gene Symbol	*SNX2*	
Exon Coordinates (hg18)	+chr5:122158860-122158977	
SNP ID	rs2161391	
Predicted Size (Skipping/Inclusion)	130/248	

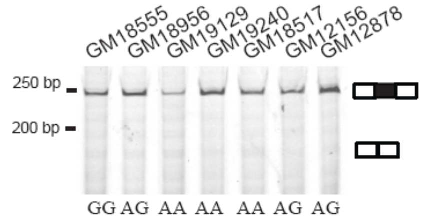
Gene Symbol	*TTK*	
Exon Coordinates (hg18)	+chr6:80778165-80778237	
SNP ID	rs239559	
Predicted Size (Skipping/Inclusion)	140/213	

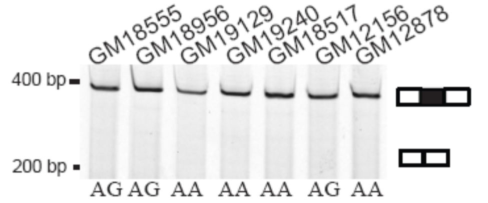
Gene Symbol	*OSBPL2*	
Exon Coordinates (hg18)	+chr20:60297658-60297781	
SNP ID	rs3746657	
Predicted Size (Skipping/Inclusion)	94/218	

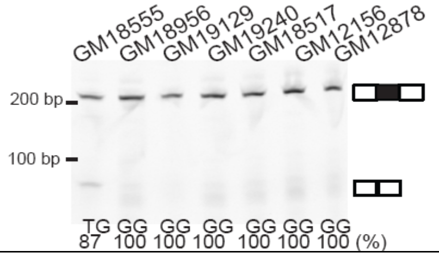
Gene Symbol	*ARHGAP23*	
Exon Coordinates (hg18)	+chr17:33891789-33891910	
SNP ID	rs11867891	
Predicted Size (Skipping/Inclusion)	175/297	

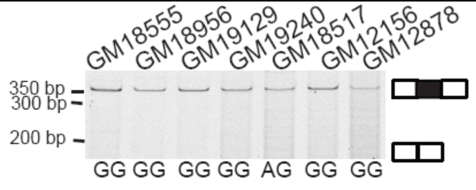
Gene Symbol	*C6orf89*	
Exon Coordinates (hg18)	+chr6:36995332-36995455	
SNP ID	rs11542163	
Predicted Size (Skipping/Inclusion)	137/261	

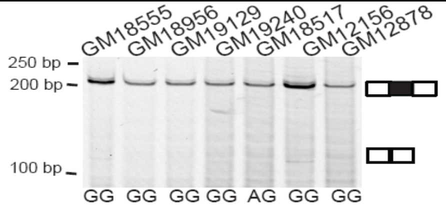
Gene Symbol	*INO80D*	
Exon Coordinates (hg18)	-chr2:206619473-206619581	
SNP ID	rs2909111	
Predicted Size (Skipping/Inclusion)	143/252	

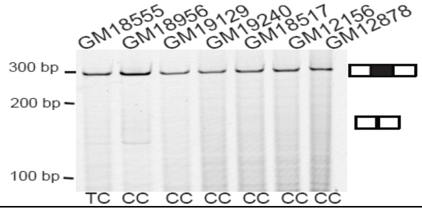
Gene Symbol	*FIG4*	
Exon Coordinates (hg18)	+chr6:110212866-110212924	
SNP ID	rs10499054	
Predicted Size (Skipping/Inclusion)	187/246	

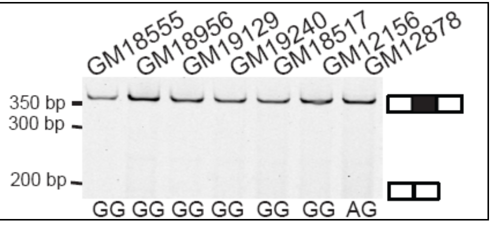
Gene Symbol	*NCAPG*	
Exon Coordinates (hg18)	+chr4:17450397-17450558	
SNP ID	rs2286537	
Predicted Size (Skipping/Inclusion)	234/396	

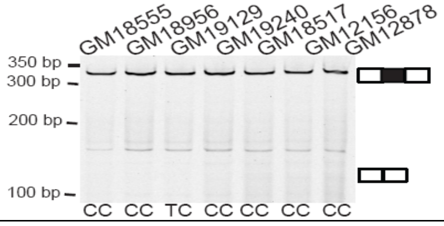
Gene Symbol	*ACAD11*	
Exon Coordinates (hg18)	-chr3:133820168-133820306	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	73/212	

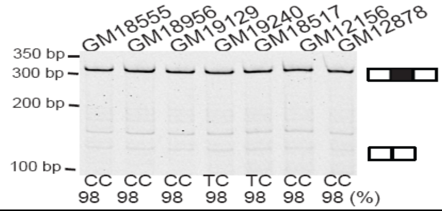
Gene Symbol	*TNFAIP6*	
Exon Coordinates (hg18)	+chr2:151930816-151930977	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	190/352	

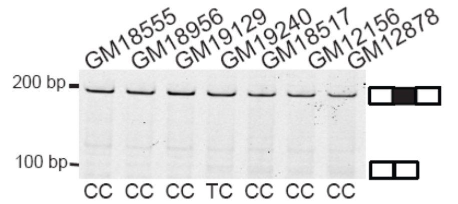
Gene Symbol	*PTPN14*	
Exon Coordinates (hg18)	-chr1:212634865-212634952	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	121/209	

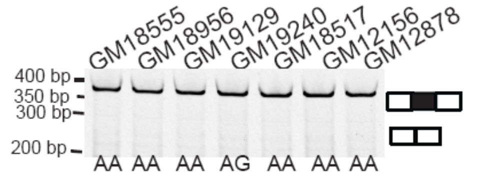
Gene Symbol	*PDGFRB*	
Exon Coordinates (hg18)	-chr5:149480960-149481078	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	172/291	

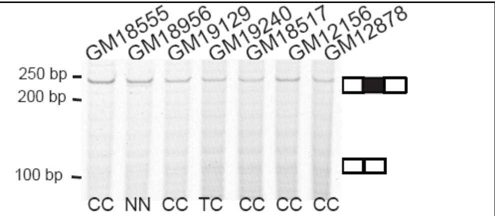
Gene Symbol	*NID1*	
Exon Coordinates (hg18)	-chr1:234220809-234220981	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	194/367	

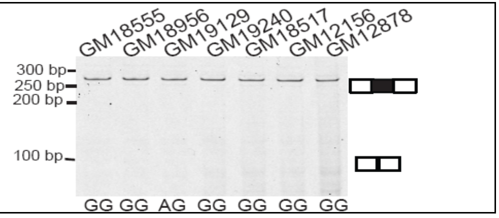
Gene Symbol	*MED16*	
Exon Coordinates (hg18)	-chr19:826244-826454	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	124/335	

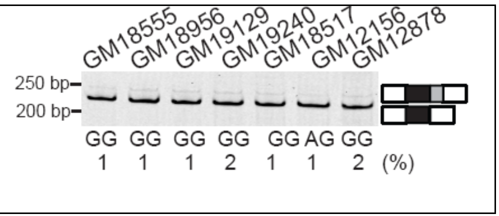
Gene Symbol	*GCN1L1*	
Exon Coordinates (hg18)	-chr12:119077428-119077617	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	121/311	

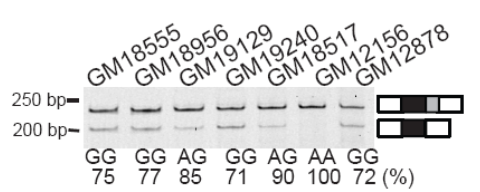
Gene Symbol	*ATP6V1E1*	
Exon Coordinates (hg18)	-chr22:16460960-16461054	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	101/198	

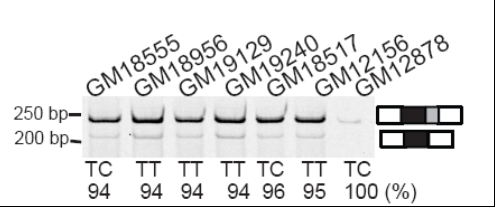
Gene Symbol	*C2CD3*	
Exon Coordinates (hg18)	-chr11:73436898-73437031	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	252/386	

Gene Symbol	*SLC4A11*	
Exon Coordinates (hg18)	-chr20:3159579-3159704	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	116/242	

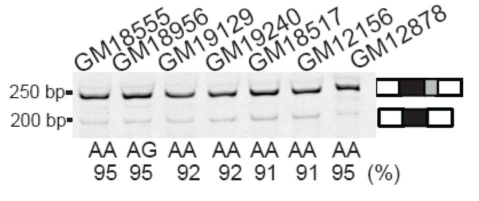
Gene Symbol	*AARS2*	
Exon Coordinates (hg18)	-chr6:44380360-44380532	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	97/270	

Gene Symbol	*WDR46*	
Exon Coordinates (hg18)	-chr6:33363347-33363473	
SNP ID	rs465151	
Predicted Size (Skipping/Inclusion)	226/248	

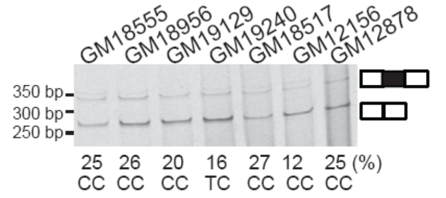
Gene Symbol	*PLD2*	
Exon Coordinates (hg18)	+chr17:4669690-4669843	
SNP ID	rs3764897	
Predicted Size (Skipping/Inclusion)	204/237	

Gene Symbol	*C9orf78*	
Exon Coordinates (hg18)	-chr9:131631336-131631403	
SNP ID	rs3818553	
Predicted Size (Skipping/Inclusion)	216/247	

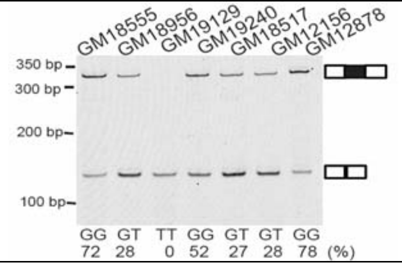
Gene Symbol	*PKN3*
Exon Coordinates (hg18)	+chr9:130522303-130522339
SNP ID	rs7874430
Predicted Size (Skipping/Inclusion)	199/243



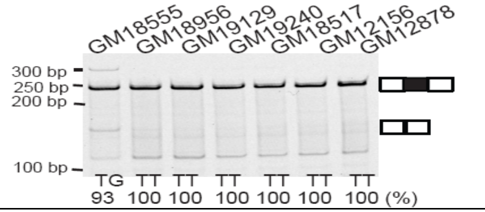
Gene Symbol	*PAQR6*
Exon Coordinates (hg18)	-chr1:154481176-154481326
SNP ID	exome sequencing
Predicted Size (Skipping/Inclusion)	277/348



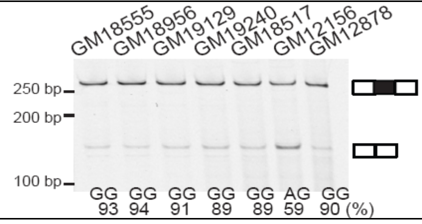
Gene Symbol	*WDR67*
Exon Coordinates (hg18)	+chr8:124223683-124223877
SNP ID	rs10101626
Predicted Size (Skipping/Inclusion)	146/341



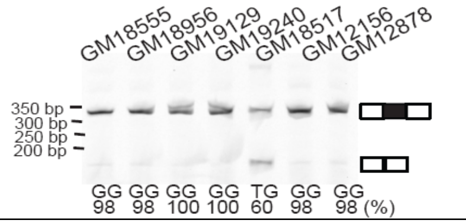
Gene Symbol	*RCC2*
Exon Coordinates (hg18)	-chr1:17621286-17621374
SNP ID	exome sequencing
Predicted Size (Skipping/Inclusion)	153/242



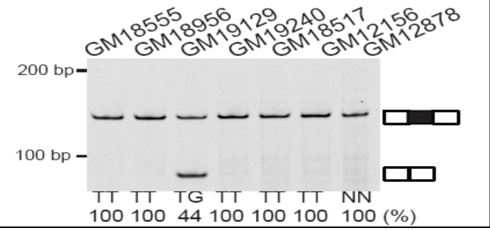
Gene Symbol	*FGGY*
Exon Coordinates (hg18)	+chr1:59695221-59695336
SNP ID	exome sequencing
Predicted Size (Skipping/Inclusion)	144/260



Gene Symbol	*CRYZ*
Exon Coordinates (hg18)	-chr1:74957481-74957644
SNP ID	exome sequencing
Predicted Size (Skipping/Inclusion)	162/326



Gene Symbol	*CASC5*
Exon Coordinates (hg18)	+chr15:38685883-38685942
SNP ID	exome sequencing
Predicted Size (Skipping/Inclusion)	82/142



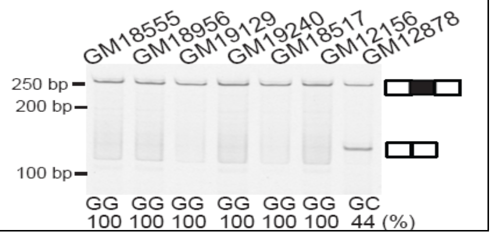
Gene Symbol	*IFIH1*	
Exon Coordinates (hg18)	-chr2:162844752-162844868	
SNP ID	exome sequencing	
Predicted Size (Skipping/Inclusion)	137/254	

Table S3: Primer list.

Primers for RT-PCR of 84 exons				
Gene Symbol	**Forward primer (5’ to 3’)** [a]	**Reverse primer (5’ to 3’)**	**Predicted skipping form size (bp)**	**Predicted inclusion form size (bp)**
PKN2	caagcctgatactcctcagtca	cgagccacaatatctcctttc	140	216
TBC1D4	tctacccaccaagagccaag	tgttgcattttctgggattg	125	236
AX747187	aaagaaaagcagagtccccc	gtcaggaacatccgagcagt	276	369
C3orf31	gctgagtctggctttcgtct	tgcgacagggtcatctactg	112	318
IMMT	agtttgtcctccgtccattg	tttttggagttgtgaggcag	139	329
UNC13D	gcagaagacgtacaacgagg	gcagtacaccagggccag	88	232
TBL1XR1	ggcaagatgttccaagcaa	ggcctttatgctgccctaag	83	147
PTPN14	ggatggatttggacaggaaa	gggcagtctcttctttgttga	111	200
ZNRD1	ttgaggggaaggttgtgaag	accagacaggatggacaagg	212	322
MTF1	tgacttgagccttctgtcca	caggcacttctggaggttgt	295	398
UTP15	gatgcacgaacgagtgagag	gcaagggcaagactcaaaat	181	317
TDG	tctctggggttgtcttaccg	cttctggcatttgctgttca	229	303
NKAP	ttctggaaaatccctggaag	gccaatttcacctcttcgtg	121	197
HMSD	gtctcacttcgctcctggg	tctccatcttcacctccgat	77	251
EVC	gcctccttgaccttcttcct	gaatcgcacaatagcccttc	214	352
MPHOSPH9	tacagcaaaaattcggaccc	tagcgtcatgcgaactgaac	133	202
DNASE1L3	atcaggatggagacgcagat	caaacaaacctggggtcagt	167	280
DHRS1	ggggattgtgcagacagaac	gcaccttaccactcaggctc	109	179
CAST	ccacctacatagaggaattggg	ttcagttttctttccagcagc	155	194
DHX38	accagctcccttttggaat	cttcctctttggcccgac	100	196
FTSJ3	aagcaaaggcactggacatc	atcttcagctccacacgctc	82	238
TLN1	ttgatacggccagtgatgtg	aggaaggtgctgaagtcctg	271	391
KIAA0556	gctgtattcccagaccagga	ctcttccttctcctccagca	91	264
MMAB	tcgacaagtacaccagccag	cgagctgatcttgcctcc	71	158
KIF18B	cagaggatgaggatgaaggc	gtcaggaggttggctgctt	101	229
EIF4G1	ggtcctaaaaaggtggggac	tctgctgctcactcaggttg	121	256
PFAS	catccagagtttctcccctg	aatttgaggacgttgttggg	230	336

AP2S1	tccgctttatcctcatccag	ccaggtccagttcacagaca	192	306
TRIM62	caaggtcaagctcttctgcc	ctccagctcctctagcatgg	217	313
PPFIA3	aggagtcgttgtatcggcag	caaaattcccatgacgttcc	45	180
SLC22A4	atctccccgatggctgata	ggagcatccagagacagagc	160	255
MOXD1	tcgaaaagggaaggaaatga	ctcctggtgcttagtcctcc	120	180
FAM129C	ctctctgccctgaatccttg	caaagtggccttggtgct	102	270
PSAP	aagtccctgtcggaagagtg	caagcacacgaaggaaggat	223	381
NCKAP1L	ctggagactgtgttggctga	tttgatgtgtcttcggacca	109	219
CEPT1	gacatttggttttgtgtctgct	ccacccccattctccttaat	237	363
NCAPG	cccctgcatcttctccttta	ttgtagtcaaggtggcatcc	234	396
DST	ataagtgctgtgaccacccc	tgtaggtgcaggtgtctgct	252	315
IGF2R	ccaccccatcttttctctca	ctggcttccaattcttgagc	203	305
SNX2	cgactttgaggatctggagg	caggagaaggttccgaggat	130	248
POLE	ggtgatgcacttctaccgct	ttcctttgattccgccatag	206	332
OSBPL3	ggaagtggtggaaggactga	atggtagatgtgctcctcgg	217	271
NPC1	catcgcagtcctgaacaaag	ggcttcacccagtcgaaata	118	309
CNTROB	aaatggcagaacgagaggaa	ggaggcatagtccagaccct	132	295
C6orf89	ctgttttcactcacctgcca	ggacccctctggctataacg	137	261
TTK	cggttcacttgggcatttac	ttcacatcacaatctgggct	140	213
PBK	tggctttggtactggggtaa	gctgctggaaaaggatcttg	166	309
HACE1	tggtcagagagatacagcacaga	tgtagcaacttcagcaaggc	178	280
GALC	cgagcatcacttcacgctac	gctgtaacttcaacacgtcctaa	142	306
BC086863	cgtggaagtatgctacgcaa	tctccccaaaatgtcctcac	109	202
ATL3	atctgcctcaccccaagtc	agctcctgctggtaacgaaa	172	244
RAP1GAP2	ggaggaggaagaggacaacc	gatccgctcatgtaccgtct	128	232
BC039374	cgacagggatttaagagagca	tcgcattgtggtttcaagtc	247	398
DHRSX	atagtgacgggagggacaga	atggaagtcatggaagccaa	119	188
CAPN14	agactgccccctgagttctt	tcctgcctttcattttggtc	217	254
AX746648	ctctcccaccacatacagca	tgagcacaggagttcgtgac	243	394
ARHGAP23	agttcctcaagcagagtgcg	cctgtaggctggacaccact	175	297
INO80D	tgaaagccacagccttctct	agctcctactctccgcatca	143	252

FIG4	atcccactgaagggaaacct	atggccgaaagaactcattg	187	246
AFTPH	cttgggaatagacacccgaa	tccttggtgggctctaacat	48	105
OSBPL2	tgctctggaggatcaacacc	ccgtgcttctctctgcttct	94	218
KIAA1731	ccacctcctcctccaactct	tcttgtgctgcctgtttttg	98	194
WDR46	ggcagtttggaccctacaga	agctcaatgccctgattgtc	226	248
PLD2	ctacatccacagcaaggtgc	aagtctggccgggtatttg	204	237
C9orf78	cctaccaacatggctgtgaa	tcccactctgcacaactcag	216	247
PKN3	caggggacacagaggaagag	cacacagggtaggcacgaag	199	243
WDR67	cccagagccagaaaactcag	ttcagcatgttctacagccg	146	341
TNFAIP6	accacagagaagcacggtct	tactcatttgggaagcctgg	190	352
PTPN14	ttggcattttctttatgggg	aaacttgtgtcgtgtggcaa	121	209
PDGFRB	ggacatgagcaaggacgagt	ttgccttcacagatgagcac	172	291
NID1	ctactgctggtgcgtggat	aatgatggtggttggctctc	194	367
MED16	tgcagcccagtatggtacag	cagcacaaattcctccgtct	124	335
GCN1L1	tgagaagaatggcctgaacc	taactcgtcccaaagcatcc	121	311
ATP6V1E1	tgattgttcgttgcaggaaa	tccagggtgttggaaacct	101	196
C2CD3	ggggcagataaaagttgctg	tcagaatggaggtttgggag	252	386
SLC4A11	cccttgtaccccttggactt	gtggtcagcagaatcaccaa	116	242
AARS2	acgacagccccaagtacaac	ggatgaaacctccacagacc	97	270
PAQR6	ccattcctgttcgacaacct	ctctgcgtgatcccatatca	277	348
RCC2	tcatcgagggtcttagccac	gaggtttcctttgcagtcca	153	242
CRYZ	ccccgtggagacatacattc	ccccatgaaccagaacactc	162	326
IFIH1	aagcaagccaaagctgaaga	gcccattgttcatagggttg	137	254
CASC5	gagacctgttagaagacggca	atctgcaaagctgactcgac	82	142
ACAD11	aattcagtggacaagtgggg	tgcagaacctccatattccc	73	212
FGGY	taaggcgggacatttctttg	ggagaaggccaaggtctctt	144	260

Primers for minigene construction and RT-PCR

Gene Symbol	**Forward primer (5’ to 3’)**	**Reverse primer (5’ to 3’)**	**Size of amplified genomic fragment (bp)**
BC039374	gggcatgacttctaggggcccttcaccgtcttagccaggat	gtcgacggcctctagagatctatcagaatcacccaggctgt	579
TDG	gggcatgacttctaggggccccactttttattcagccaactgtg	gtcgacggcctctagagatcttcccagctagttcaggatgc	438
CAST	gggcatgacttctaggggcccacttgattctataaggaacctatctca	gtcgacggcctctagagatcttcaggagcactgagacaacc	394
DHRS1	gggcatgacttctaggggcccggcactaccagatccttcca	gtcgacggcctctagagatctagcctgctcctaaggctctc	495
BC086863	gggcatgacttctaggggccctcgaattctcaaagggtagca	gtcgacggcctctagagatcttggtgaagcatgggatcaat	497
TRIM62	gggcatgacttctagccatgtgcagttcacactcc	gtcgacggcctctagcaggcaatcttcccatctgt	1220
HMSD	gggcatgacttctagcaccatccctgtgttccttc	gtcgacggcctctagtgccttttgcaaaccctaac	1360
Minigene_RT-PCR	cgtcgccgtccagctcgaccaggaattcgagctcactctcttc	aagcttaggctcttggcgtt	122 (skipping form)

Primers for realtime qPCR assay

Gene Symbol	**Forward primer (5’ to 3’)**	**Reverse primer (5’ to 3’)**	**Fragment Size (bp)**
HNRNPH1	cttgccctggtcttgctc	cttggtctgccttctctggt	111
ACTB	gagctacgagctgcctgacg	gtagtttcgtggatgccacag	120

a. A 22 nt universal tag sequence (5’-cgtcgccgtccagctcgaccag-3’), derived from the GFP N-terminal region, was added to the 5’ end of the forward gene-specific primer. A fluorescently labeled universal primer 5’FAM-cgtcgccgtccagctcgaccag-3’ was added as the third primer in RT-PCR reaction. Predicted fragment size does not include the tag sequence in the primers.

Table S4: Splice site sequences used in minigene splicing reporters.

Splice site sequence	Splice site score[a]	Note
`aaa|gtgagg`	4.44	5’ splice site sequences used in *TRIM62* and *HMSD* minigene experiments (Fig. 3C).
`aag|gttagc`	5.61	
`aag|gtcagc`	7.09	
`cag|gtacat`	7.79	
`aaa|gtgagt`	8.40	
`aag|gtgagc`	9.60	
`cag|gtacgt`	10.65	
`aag|gtaagt`	11.00	
`caaattaatatttgtttaag|aaa`	1.59	3’ splice site sequences used in *CAST* minigene experiments (Fig. 4C).
`caaattaatatttgatttag|aaa`	3.39	
`caaattaatatttgtttcag|aaa`	5.40	
`caaattaatatttgtttcag|gta`	6.97	
`caaattaatttttttttcag|gta`	8.84	
`catttttttttttgtttcag|aaa`	10.90	
`catttttttttttgtttcag|gta`	12.47	
`cggccctgtgtgtgtgtaag|agg`	2.07	3’ splice site sequences used in *TRIM62* minigene experiments (Fig. 4D, Fig. 5, Fig. S3, Fig. S4).
`cggccctgtgtgtgggtcag|agg`	4.03	
`cggccctgtgtgtgtttaag|agg`	5.93	
`cggccctgtgtatatttcag|agg`	8.00	
`cggccctgtgtgtgtttcag|agg`	9.70	

[a]: Splice site score is calculated by MAXENT (see Materials and Methods).