

	10	20	30	40	50
<b>Q9UPS6-1</b>	<b>MENSHPPHHH</b>	<b>HQPPPPQGP</b>	<b>SGERRNHWR</b>	<b>SYKLMIDPAL</b>	<b>KKGHHKLYRY</b>
A0A0A0MQV9-1	MENSHPPHHH	HQPPPPQGP	SGERRNHWR	SYKLMIDPAL	KKGHHKLYRY 50
	60	70	80	90	100
	<b>DGQHFSLAMS</b>	<b>SNRPVEIVED</b>	<b>PRVVGIWTKN</b>	<b>KELELSVPKF</b>	<b>KIDEFYVGPV</b>
	DGQHFSLAMS	SNRPVEIVED	PRVVGIWTKN	KELELSVPKF	KIDEFYVGPV 100
	110	120	130	140	150
	<b>PPKQVTFAKL</b>	<b>NDNIRENFLR</b>	<b>DMCKKYGEVE</b>	<b>EVEILYNPKT</b>	<b>KKHLGIAKVV</b>
	PPKQVTFAKL	NDNIRENFLR	DMCKKYGEVE	EVEILYNPKT	KKHLGIAKVV 150
	160	170	180	190	200
	<b>FATVRGAKDA</b>	<b>VQHLHSTSV</b>	<b>GNIIHVELDT</b>	<b>KGETRMRFYE</b>	<b>LLVTGRYTPQ</b>
	FATVRGAKDA	VQHLHSTSV	GNIIHVELDT	KGETRMRFYE	LLVTGRYTPQ 200
	210	220	230	240	250
	<b>TLPVGELEDAV</b>	<b>SPIVNETLQL</b>	<b>SDALKRLKDG</b>	<b>GLSAGCGSGS</b>	<b>SSVTPNSGGT</b>
	TLPVGELEDAV	SPIVNETLQL	SDALKRLKDG	GLSAGCGSGS	SSVTPNSGGT 250
	260	270	280	290	300
	<b>PFSQDTAYSS</b>	<b>CRLDTPNSYG</b>	<b>QGTPLTPRLG</b>	<b>TPFSQDSSYS</b>	<b>SRQPTPSYLF</b>
	PFSQDTAYSS	CRLDTPNSYG	QGTPLTPRLG	TPFSQDSSYS	SRQPTPSYLF 300
	310	320	330	340	350
	<b>SQDPAVTFKA</b>	<b>RRHESKFTDA</b>	<b>YNRRHEHHYV</b>	<b>HNSPAVTAVA</b>	<b>GATAAFRGSS</b>
	SQDPAVTFKA	RRHESKFTDA	YNRRHEHHYV	HNSPAVTAVA	GATAAFRGSS 350
	360	370	380	390	400
	<b>DLFFGAVGGT</b>	<b>GGSSGPPFKA</b>	<b>QPQDSATFAH</b>	<b>TPPPAQATPA</b>	<b>PGFKSAFSPY</b>
	DLFFGAVGGT	GGSSGPPFKA	QPQDSATFAH	TPPPAQATPA	PGFKSAFSPY 400
	410	420	430	440	450
	<b>QTPVAHFPPP</b>	<b>PEEPTATAAF</b>	<b>GARDSGEFRR</b>	<b>APAPPPLPPA</b>	<b>EPLAKEKPGT</b>
	QTPVAHFPPP	PEEPTATAAF	GARDSGEFRR	APAPPPLPPA	EPLAKEKPGT 450
	460	470	480	490	500
	<b>PPGPPPPDTN</b>	<b>SMELGGRPTF</b>	<b>GWSPEPCDSP</b>	<b>GTPTLESSPA</b>	<b>GPEKPHDSL</b>
	PPGPPPPDTN	SMELGGRPTF	GWSPEPCDSP	GTPTLESSPA	GPEKPHDSL 500
	510	520	530	540	550
	<b>SRIEMLLKEQ</b>	<b>RTKLLFLREP</b>	<b>DSDTELOMEG</b>	<b>SPISSSSSQL</b>	<b>SPLAPFGTNS</b>
	SRIEMLLKEQ	RTKLLFLREP	DSDTELOMEG	SPISSSSSQL	SPLAPFGTNS 550
	560	570	580	590	600
	<b>QPGFRGPTPP</b>	<b>SSRPSSTGLE</b>	<b>DISPTPLPDS</b>	<b>DEDEELDLGL</b>	<b>GPRPPPEPGP</b>
	QPGFRGPTPP	SSRPSSTGLE	DISPTPLPDS	DEDEELDLGL	GPRPPPEPGP 600
	610	620	630	640	650
	<b>PDPAGLLSQT</b>	<b>AEVALDLVGD</b>	<b>RTPTSEKME</b>	<b>GQQSSGEDME</b>	<b>ISDDEMP</b>
	PDPAGLLSQT	AEVALDLVGD	RTPTSEKME	GQQSSGEDME	ISDDEMP 650
	660	670	680	690	700
	<b>ITSADCPKPM</b>	<b>VVTPGAAAVA</b>	<b>APSVLAPTL</b>	<b>LPPPPGFPP</b>	<b>PPPPPPPPQ</b>
	ITSADCPKPM	VVTPGAAAVA	APSVLAPTL	LPPPPGFPP	PPPPPPPPQ 700
	710	720	730	740	750
	<b>PGFPMPPPLP</b>	<b>PPPPPPPPAH</b>	<b>PAVTVPPPPL</b>	<b>PAPPGVPPP</b>	<b>ILPPLPPFP</b>
	PGFPMPPPLP	PPPPPPPPAH	PAVTVPPPPL	PAPPGVPPP	ILPPLPPFP 750
	760	770	780	790	800
	<b>GLFPVMQVDM</b>	<b>SHVLGGQWGG</b>	<b>MPMSFQMOTQ</b>	<b>VLSRLMTGQG</b>	<b>ACPYPFFMAA</b>
	GLFPVMQVDM	SHVLGGQWGG	MPMSFQMOTQ	VLSRLMTGQG	ACPYPFFMAA 800
	810	820	830	840	850
	<b>AAAAASAGLQ</b>	<b>FVNLPPYRGP</b>	<b>FSLNSGPGR</b>	<b>GQHWPPLPKF</b>	<b>DPSVPPPGYM</b>
	AAAAASAGLQ	FVNLPPYRGP	FSLNSGPGR	GQHWPPLPKF	DPSVPPPGYM 850
	860	870	880	890	900
	<b>PRQEDPHKAT</b>	<b>VDGVLLVVLK</b>	<b>ELKAIMKRDL</b>	<b>NRKMVEVAF</b>	<b>RAFDEWWDKK</b>
	PRQEDPHKAT	VDGVLLVVLK	ELKAIMKRDL	NRKMVEVAF	RAFDEWWDKK 900
	910	920	930	940	950
	<b>ERMAKASLTP</b>	<b>VKSGEHKDED</b>	<b>RPKPKDRIAS</b>	<b>CLLESWGKGE</b>	<b>GLGYEGLGLG</b>
	ERMAKASLTP	VKSGEHKDED	RPKPKDRIAS	CLLESWGKGE	GLGYEGLGLG 950
	960	970	980	990	1000
	<b>IGLRGAIIRLP</b>	<b>SFKVVRKEPP</b>	<b>DTTSSGDKR</b>	<b>LRPSTSVDEE</b>	<b>DEESERERDR</b>
	IGLRGAIIRLP	SFKVVRKEPP	DTTSSGDKR	LRPSTSVDEE	DEESERERDR 1000
	1010	1020	1030	1040	1050

ITSADCPK

<b>DMADTPCELA</b>	<b>KRDPKGVGVR</b>	<b>RRPARPLELD</b>	<b>SGGEEDEKES</b>	<b>LSASSSSSAS</b>	
DMADTPCELA	KRDPKGVGVR	RRPARPLELD	SGGEEDEKES	L-----	1041
<b>1060</b>	<b>1070</b>	<b>1080</b>	<b>1090</b>	<b>1100</b>	
<b>SSSGSSTTSP</b>	<b>SSSASDKEEE</b>	<b>QESTEEEEEA</b>	<b>EEEEEEVPR</b>	<b>SQLSSSSTSS</b>	
-----	-----EEE	QESTEEEEEA	EEEEEE---	-----	1061
<b>1110</b>	<b>1120</b>	<b>1130</b>	<b>1140</b>	<b>1150</b>	
<b>TSDKDDDDDD</b>	<b>SDDRDESEND</b>	<b>DEDTALSEAS</b>	<b>EKDEGDSDEE</b>	<b>ETVSIVTSKA</b>	
----DDDDDD	SDDRDESEND	DEDTALSEAS	EKDEGDSDEE	ETVSIVTSKA	1107
<b>1160</b>	<b>1170</b>	<b>1180</b>	<b>1190</b>	<b>1200</b>	
<b>EATSSSESSE</b>	<b>SSEFESSSES</b>	<b>SPSSSEDEEE</b>	<b>VVAREEEEE</b>	<b>EEEEEMVAEES</b>	
EATSSSESSE	SSEFESSSES	SPSSSEDEEE	VVAREEEEE	EEEEEMVAEES	1157
<b>1210</b>	<b>1220</b>	<b>1230</b>	<b>1240</b>	<b>1250</b>	
<b>MASAGPEDFE</b>	<b>QDGEEAALAP</b>	<b>GAPAVDSLGM</b>	<b>EEEVDIETEA</b>	<b>VAPEERPSML</b>	
MASAGPEDFE	QDGEEAALAP	GAPAVDSLGM	EEEVDIETEA	VAPEERPSML	1207
<b>1260</b>	<b>1270</b>	<b>1280</b>	<b>1290</b>	<b>1300</b>	
<b>DEPPLPVGVE</b>	<b>EPADSREPE</b>	<b>EPGLSQEGAM</b>	<b>LLSPEPPAKE</b>	<b>VEARPPLSPE</b>	
DEPPLPVGVE	EPADSREPE	EPGLSQEGAM	LLSPEPPAKE	VEARPPLSPE	1257
<b>1310</b>	<b>1320</b>	<b>1330</b>	<b>1340</b>	<b>1350</b>	
<b>RAPEHDLEVE</b>	<b>PEPPMMLPLP</b>	<b>LOPPLPPRP</b>	<b>PRPPSPPPEP</b>	<b>ETTDASHP</b>	<b>PSV</b>
RAPEHDLEVE	PEPPMMLPLP	LOPPLPPRP	PRPPSPPPEP	ETTDASHP	PSV 1307
<b>1360</b>	<b>1370</b>	<b>1380</b>	<b>1390</b>	<b>1400</b>	
<b>PPEPLAEDHP</b>	<b>PHTPGLCGSL</b>	<b>AKSQSTETVP</b>	<b>ATPGGEPPLS</b>	<b>GGSSGLSLSS</b>	
PPEPLAEDHP	PHTPGLCGSL	AKSQSTETVP	ATPGGEPPLS	GGSSGLSLSS	1357
<b>1410</b>	<b>1420</b>	<b>1430</b>	<b>1440</b>	<b>1450</b>	
<b>PQVPGSPFSY</b>	<b>PAPSPSLSSG</b>	<b>GLPRTPGRDF</b>	<b>SFTPTFSEPS</b>	<b>GPLLLPVCPL</b>	
PQVPGSPFSY	PAPSPSLSSG	GLPRTPGRDF	SFTPTFSEPS	GPLLLPVCPL	1407
<b>1460</b>	<b>1470</b>	<b>1480</b>	<b>1490</b>	<b>1500</b>	
<b>PTGRRDERSG</b>	<b>PLASPVLEET</b>	<b>GLPLPLPLPL</b>	<b>PLPLALPAVL</b>	<b>RAQARAPTPL</b>	
PTGRRDERSG	PLASPVLEET	GLPLPLPLPL	PLPLALPAVL	RAQARAPTPL	1457
<b>1510</b>	<b>1520</b>	<b>1530</b>	<b>1540</b>	<b>1550</b>	
<b>PPLLPAPLAS</b>	<b>CPPPMKRKPG</b>	<b>RPRRSPPSML</b>	<b>SLDGPLVRPP</b>	<b>AGAALGRELL</b>	
PPLLPAPLAS	CPPPMKRKPG	RPRRSPPSML	SLDGPLVRPP	AGAALGRELL	1507
<b>1560</b>	<b>1570</b>	<b>1580</b>	<b>1590</b>	<b>1600</b>	
<b>LLPGQPQTPV</b>	<b>FPSTHDPRTV</b>	<b>TLDFRNAGIP</b>	<b>APPPPLPPQP</b>	<b>PPPPPPPPVE</b>	
LLPGQPQTPV	FPSTHDPRTV	TLDFRNAGIP	APPPPLPPQP	PPPPPPPPVE	1557
<b>1610</b>	<b>1620</b>	<b>1630</b>	<b>1640</b>	<b>1650</b>	
<b>PTKLPFKELD</b>	<b>NQWPSEAIPP</b>	<b>GPRGRDEVTE</b>	<b>EYMEKAKSRG</b>	<b>PWRRPPKRRH</b>	
PTKLPFKELD	NQWPSEAIPP	GPRGRDEVTE	EYMEKAKSRG	PWRRPPKRRH	1607
<b>1660</b>	<b>1670</b>	<b>1680</b>	<b>1690</b>	<b>1700</b>	
<b>EDLVPPAGSP</b>	<b>ELSPPQPLFR</b>	<b>PRSEFEEMTI</b>	<b>LYDIWNGGID</b>	<b>EDDIRFLCVT</b>	
EDLVPPAGSP	ELSPPQPLFR	PRSEFEEMTI	LYDIWNGGID	EDDIRFLCVT	1657
<b>1710</b>	<b>1720</b>	<b>1730</b>	<b>1740</b>	<b>1750</b>	
<b>YERLLQQDNG</b>	<b>MDWLNDTLWV</b>	<b>YHPSTSLSSA</b>	<b>KKKKRDDGIR</b>	<b>EHVTGCARSE</b>	
YERLLQQDNG	MDWLNDTLWV	YHPSTSLSSA	KKKKRDDGIR	EHVTGCARSE	1707
<b>1760</b>	<b>1770</b>	<b>1780</b>	<b>1790</b>	<b>1800</b>	
<b>GFYTIDKKDK</b>	<b>LRYLNSSRAS</b>	<b>TDEPPADTQG</b>	<b>MSIPAQPHAS</b>	<b>TRAGSERRSE</b>	
GFYTIDKKDK	LRYLNSSRAS	TDEPPADTQG	MSIPAQPHAS	TRAGSERRSE	1757
<b>1810</b>	<b>1820</b>	<b>1830</b>	<b>1840</b>	<b>1850</b>	
<b>QRRLLSFTG</b>	<b>SCSDLLKFN</b>	<b>QLKFRKKKLL</b>	<b>FCKSHIHDWG</b>	<b>LFAMEPIAAD</b>	
QRRLLSFTG	SCSDLLKFN	QLKFRKKKLL	FCKSHIHDWG	LFAMEPIAAD	1807
<b>1860</b>	<b>1870</b>	<b>1880</b>	<b>1890</b>	<b>1900</b>	
<b>EMVIEYVGN</b>	<b>IRQVIADMRE</b>	<b>KRYEDEGIGS</b>	<b>SYMFVVDHDT</b>	<b>IIDATKCGNF</b>	
EMVIEYVGN	IRQVIADMRE	KRYEDEGIGS	SYMFVVDHDT	IIDATKCGNF	1857
<b>1910</b>	<b>1920</b>	<b>1930</b>	<b>1940</b>	<b>1950</b>	
<b>ARFINHSCNP</b>	<b>NCYAKVITVE</b>	<b>SQKKVIYSK</b>	<b>QHINVNEEIT</b>	<b>YDYKPIEDV</b>	
ARFINHSCNP	NCYAKVITVE	SQKKVIYSK	QHINVNEEIT	YDYKPIEDV	1907
<b>1960</b>	<b>1966</b>				
<b>KIPCLCGSEN</b>	<b>CRGTLN</b>				
KIPCLCGSEN	CRGTLN	1923			

**Supplementary Figure 1:** Location of rare coding variants listed in this paper. The canonical 1966 amino acid transcript (UniProtKB accession number Q9UPS6-1) is in boldface, the shorter 1923 amino acid transcript (UniProtKB accession number A0A0A0MQV9-1) is in plain text. Variant sites described in Figure 1 and in the text are outlined in black. SET domain is underlined.