Supp. Table S1. Tools used for *in silico* predictions

Tool	Link	Features	Output	Ref
CADD	http://cadd.gs.washington.edu/info	Ensemble score	PHRED-like scaled C-score	(Kircher et al., 2014)
CONDEL	http://bg.upf.edu/fannsdb/	Ensemble score	Condel score-based label (D/N)	(Gonzalez-Perez et al., 2011)
I-Mutant	http://folding.biofold.org/i-mutant/i-mutant2.0.html	3D protein structure & Gibbs free energy-based	Decreased/Increased Stability	(Capriotti et al., 2005)
MAPP	http://mendel.stanford.edu/sidowlab/downloads/MAPP/	Physico-chemical properties-based (hydropathy, polarity, volume...etc)	deleterious/neutral	(Stone, 2005)
MutationAssessor	http://mutationassessor.org/	Phenomenological features from protein familiy	Functional Impact (FI): high, medium,low, neutral + an FI score	(Reva et al., 2011)
MutPred	http://mutpred.mutdb.org/	14 structural and functional protein properties	deleterious/neutral	(Li et al., 2009)
nsSNP Analyzer	http://snpanalyzer.uthsc.edu/	Random forest machine learning	disease/neutral	(Bao et al., 2005)
PANTHER	http://www.pantherdb.org/tools/csnpScoreForm.jsp?	Phylogenetic tree-based Hidden Markov models machine learning	probability of deleteriousness (Pdeleterious)	(Tang, 2016)

PhD-SNP	http://snps.biofold.org/phd-snp/phd-snp.html	SVM machine learning	disease/neutral	(Capriotti et al., 2006)
PolyPhen-2	http://genetics.bwh.harvard.edu/pph2/	8-protein sequence features, 3 structure features, Bayesian classification	damaging/benign, sensitivity & specificity	(Adzhubei et al., 2010)
PON-P2	http://structure.bmc.lu.se/PON-P2/	Ensemble score	pathogenic/neutral/unknown	(Niroula et al., 2015)
PoP MuSiC	http://babylone.ulb.ac.be/popmusic/	3D protein structure & Gibbs free energy-based	decreased/increased Stability	(Dehouck et al., 2011)
Predict SNP	http://loschmidt.chemi.muni.cz/predictsnp/	Ensemble score	deleterious/neutral	(Bendl et al., 2014)
PROVEAN	http://provean.jcvi.org/index.php	Clustering, applicable to insertion and deletion mutation	deleterious/neutral	(Choi, 2015)
SIFT	http://sift.jcvi.org/	Protein sequence conservation	damaging/tolerated	(Kumar et al., 2009)
SNAP2	https://rostlab.org/services/snap2web/	Neural network machine learning	effect/neutral, a score, expected accuracy	(Hecht et al., 2015)

Supp Table S2. Non-Synoymous & Consensus Splice-Site Variants in All Demographic/Racial Groups

Position (GRCh37/hg19)	cDNA[a] / Amino Acid[b] Change	CpG dinucleotide	Allele Frequency				
			Caucasians[c]	Latinos	Africans	East Asians	South Asians
A. Non-Synonymous Variant:							
118955746	c.3G>A / p.M1I	–	0.0000136	0	0	0	0
118955750	c.7G>A / p.G3S	–	0.0000273	0	0	0	0
118955751	c.8G>T / p.G3V	–	0.0000136	0	0	0	0
118955769	c.26C>A / p.A9E	–	0.0000545	0	0	0	0
118955772	c.29C>A / p.T10K	–	0.0000409	0	0	0.000233	0
118955772	c.29C>T / p.T10M	+	0.0000136	0	0	0	0
118955774	c.31G>A / p.A11T	–	0	0	0.001713	0	0
118958966	c.35A>C / p.E12A	–	0.0000273	0	0	0	0
118958995	c.64C>T / p.R22C	+	0.0000273	0	0	0	0
118958996	c.65G>A / p.R22H	+	0	0	0	0.000231	0
118958998	c.67G>A / p.V23M	+	0	0.000086	0	0	0
118959351	c.94C>T / p.R32C	+	0.0000273	0	0	0	0
118959352	c.95G>A / p.R32H	+	0	0	0	0	0.000303

118959356	c.99A>G / p.I33M	–	0	0	0	0	0.000121
118959383	c.126G>T / p.L42F	–	0	0	0	0.000116	0
118959391	c.134C>T / p.S45L	+	0.0004362	0	0.000096	0	0
118959406	c.149A>G / p.Q50R	–	0	0	0	0	0.000121
118959417	c.160A>C / p.I54L	–	0.0001227	0	0	0	0
118959797	c.166A>G / p.M56V	–	0	0	0	0	0.000061
118959807	c.176C>T / p.T59I	–	0.0000409	0	0	0	0.000606
118959824	c.193G>C / p.D65H	–	0.0000136	0	0	0	0
118959928	c.212T>C / p.I71T	–	0.0000273	0	0	0	0
118959930	c.214G>A / p.G72R	–	0.0000409	0.000086	0	0	0
118959940	c.224G>A / p.S75N	–	0	0	0.000096	0	0
118959972	c.256G>A / p.E86K	–	0.0000136	0	0	0	0
118959973	c.257A>T / p.E86V	–	0.0010633	0	0.000192	0	0
118960431	c.305C>T / p.T102I	–	0	0	0	0.000116	0
118960458	c.332G>C / p.G111A	–	0.0000136	0	0	0	0
118960469	c.343A>C / p.K115Q	–	0.0000818	0	0	0	0
118960709	c.354C>G / p.N118K	–	0.0000136	0	0	0	0
118960713	c.358C>A / p.H120N	–	0.0000136	0	0	0	0

118960719	c.364G>C / p.A122P	–	0.0000136	0	0	0	0
118960734	c.379C>G / p.P127A	–	0.0000136	0	0	0	0
118960735	c.380C>A / p.P127Q	–	0.0001909	0.000518	0	0	0.000484
118960750	c.395A>G / p.K132R	–	0	0	0	0.000116	0
118960751	c.396G>C / p.K132N	–	0.0000136	0	0	0	0
118960775	c.420G>T / p.K140N	–	0	0	0	0.000116	0
118960777	c.422G>A / p.S141N	–	0	0	0.000096	0	0
118960900	c.423T>A / p.S141R	–	0	0	0.000096	0	0
118960911	c.434C>A / p.T145N	–	0	0.000086	0	0	0
118960953	c.476C>T / p.P159L	+	0.0000136	0	0	0	0
118960959	c.482T>C / p.L161P	–	0.0000136	0	0	0	0
118960967	c.490A>G / p.R164G	–	0.0002045	0	0	0	0
118960969	c.492G>T / p.R164S	–	0.0000136	0	0	0	0
118962123	c.499C>T / p.R167W	+	0.0001091	0	0	0	0
118962124	c.500G>A / p.R167Q	+	0.0002045	0	0.000363	0	0
118962156	c.532G>A / p.D178N	–	0	0	0.00036	0	0.003665
118962207	c.583C>T / p.R195C	+	0.0001091	0	0	0	0
118962837	c.615C>G / p.I205M	–	0.0000273	0	0	0	0

118962854	c.632G>T / p.C211F	–	0	0	0.000096	0	0
118962856	c.634A>G / p.M212V	–	0	0.000087	0	0	0
118963117	c.655G>A / p.A219T	–	0	0.000086	0	0	0
118963118	c.656C>T / p.A219V	–	0	0	0	0	0.000061
118963123	c.661G>A / p.G221S	–	0.0000136	0.000086	0	0	0
118963136	c.674G>A / p.R225Q	+	0.0002181	0.000173	0.000096	0	0
118963150	c.688G>T / p.D230Y	–	0.0000136	0	0	0	0
118963152	c.690C>G / p.D230E	–	0	0	0	0	0.000182
118963171	c.709G>A / p.V237M	–	0.0000136	0	0	0	0
118963177	c.715C>T / p.H239Y	–	0	0.000086	0	0	0
118963180	c.718G>A / p.D240N	+	0	0.000173	0	0	0
118963186	c.724G>A / p.E242K	+	0.0000136	0	0	0	0.000061
118963198	c.736C>T / p.R246C	+	0.0000136	0	0.001154	0.000462	0.000121
118963199	c.737G>A / p.R246H	+	0.0000545	0	0.000096	0	0
118963207	c.745G>A / p.A249T	+	0	0.000086	0	0	0
118963215	c.753G>C / p.R251S	–	0.0000136	0	0	0	0
118963217	c.755C>T / p.A252V	–	0.0000136	0	0	0	0
118963230	c.768C>A / p.H256Q	–	0.0000136	0	0	0.000116	0

118963230	c.768C>G / p.H256Q	–	0	0	0	0	0.0001156
118963469	c.773A>G / p.E258G	–	0	0	0	0	0.000303
118963481	c.785G>A / p.S262N	–	0.0000136	0	0	0	0
118963490	c.794T>C / p.V265A	–	0	0	0	0	0.000061
118963493	c.797C>G / p.A266G	–	0	0.000087	0	0	0
118963505	c.809C>T / p.A270V	–	0	0	0	0	0.000061
118963507	c.811A>G / p.M271V	–	0	0	0.000096	0	0
118963514	c.818A>G / p.D273G	–	0.0000273	0	0	0	0
118963517	c.821G>C / p.G274A	–	0.0000136	0	0	0	0
118963663	c.844G>C / p.V282L	–	0.0000273	0	0	0	0
118963672	c.853C>A / p.L285I	–	0	0	0	0.000116	0
118963673	c.854T>G / p.L285R	–	0	0.000173	0.000192	0	0
118963678	c.859G>A / p.G287S	+	0	0.000086	0	0	0
118963706	c.887A>G / p.Q296R	–	0.0000136	0	0	0	0
118963708	c.889G>T / p.A297S	–	0.0000136	0	0	0	0
118963826	c.919G>A / p.D307N	–	0	0	0	0	0.000061
118963832	c.925C>T / p.P309S	–	0.0000682	0	0	0	0.000061
118963835	c.928G>C / p.E310Q	–	0	0	0.000096	0	0

118963841	c.934G>A / p.D312N	–	0	0	0	0	0.000061
118963854	c.947T>C / p.V316A	–	0	0	0	0	0.000424
118963868	c.961C>T / p.R321C	+	0.0000136	0	0	0	0
118963869	c.962G>A / p.R321H	+	0.0019222	0.000432	0.000288	0.000116	0.000242
118963869	c.962G>C / p.R321P	–	0	0	0.000096	0	0.000061
118963881	c.974G>A / p.R325Q	+	0.0000273	0	0	0	0
118963891	c.984G>C / p.Q328H	–	0	0.000346	0	0	0
118963899	c.992C>T / p.A331V	–	0.0000273	0	0	0	0
118963950	c.1043A>G / p.K348R	–	0.0000409	0	0	0	0
118963967	c.1060G>T / p.A354S	–	0	0	0.000097	0	0
118963970	c.1063C>T / p.R355W	+	0.0000545	0	0	0	0
118963971	c.1064G>A / p.R355Q	+	0.0000136	0	0.000097	0	0
118963982	c.1075G>A / p.D359N	+	0.0000136	0	0.000785	0	0
B. Consensus Splice-Site Variant:							
118960778	c.422+1G>A	–	0.0000136	0	0	0	0
118962834	c.613-1G>T	–	0	0	0.000096	0	0
118963111	c.652-2delA	–	0.0000136	0	0	0	0
Total Allele Frequency:			0.0057529	0.0026770	0.0062970	0.0017380	0.0072366

[a] Nucleotide numbering uses +1 as the A of the ATG translation initiation codon in NM_000190.3; [b] Amino acid numbering uses +1 as the first amino acid of the NP_000181.2; [c] Caucasians sequences include European and Finnish data

Supp Table S3. *In silico* pathogenicity prediction

Variant	CADD	CONDEL	I-Mutant	MAPP	Mutation Assessor	MutPred	nsSNP Analyzer	PANTHER	PhD-SNP	PolyPhen-2	PON-P2	PoP MuSiC	Predict SNP	PROVEAN	SIFT	SNAP2	Program counts			Supporting Decision for Pathogenicity	Consensus Deleterious score	Pathogenicity*
																	D*	N*	U*			
p.M1I	N	D	D	N	U	D	U	N	N	D	N	U	N	N	D	N	5	8	3	5/13	0.38	A
p.G3V	N	D	D	N	D	N	U	N	D	D	U	U	D	N	D	N	7	6	3	7/13	0.54	A
p.G3S	N	D	D	N	N	N	U	N	D	N	N	U	N	N	N	N	3	11	2	3/14	0.21	B
p.A9E	N	D	D	N	N	N	N	N	N	N	N	U	N	N	D	N	3	12	1	3/15	0.2	B
p.T10K	N	D	D	N	N	N	N	N	N	N	N	U	N	N	D	N	3	12	1	3/15	0.2	B
p.T10M	N	D	N	N	N	N	N	N	D	D	N	U	D	N	D	N	5	10	1	5/15	0.33	A
p.E12A	N	D	D	N	N	N	N	N	N	D	N	U	N	N	N	N	3	12	1	3/15	0.2	B
p.R22C	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	16	0	0	16/16	1	D
p.R32C	D	D	D	D	N	N	U	D	D	D	D	D	D	N	D	N	11	4	1	11/15	0.73	A
p.S45L	N	N	N	N	N	N	D	D	N	N	N	D	N	N	N	N	3	13	0	3/16	0.19	B
p.I54L	N	D	D	N	N	N	U	D	N	N	U	D	N	N	N	N	4	10	2	4/14	0.29	A
p.T59I	N	D	N	N	N	D	U	D	D	N	N	U	N	N	N	N	4	10	2	4/14	0.29	A
p.D65H	D	D	N	N	D	N	U	D	D	D	U	U	D	D	D	D	10	3	3	10/13	0.77	D
p.I71T	D	D	D	D	D	D	U	D	D	D	U	U	D	D	D	D	13	0	3	13/13	1	D

p.G72R	D	D	D	D	D	D	U	D	D	D	D	U	D	D	D	D	14	0	2	14/14	1	D
p.E86V	D	D	N	D	N	D	U	D	D	N	U	N	N	N	N	N	6	8	2	6/14	0.43	A
p.E86K	D	D	D	N	N	N	U	D	D	N	U	D	N	N	N	N	6	8	2	6/14	0.43	A
p.G111A	D	D	D	N	N	N	U	D	N	N	U	D	N	D	N	N	6	8	2	6/14	0.43	A
p.K115Q	N	D	D	N	N	N	U	D	N	N	U	D	N	N	N	N	4	10	2	4/14	0.29	A
p.N118K	N	D	D	D	N	N	U	D	D	N	U	D	N	D	D	N	8	6	2	8/14	0.57	A
p.H120N	N	D	D	N	N	N	U	D	D	N	N	D	N	N	N	N	5	10	1	5/15	0.33	A
p.A122P	D	D	N	D	D	D	D	D	D	D	U	D	D	D	D	D	14	1	1	14/15	0.93	D
p.P127A	N	D	D	N	N	N	N	N	N	N	U	D	N	N	N	N	3	12	1	3/15	0.2	B
p.P127Q	D	D	D	N	N	N	N	N	N	N	U	D	N	N	N	N	4	11	1	4/15	0.27	A
p.K132N	N	D	N	N	N	N	U	D	N	N	U	D	N	D	N	N	4	10	2	4/14	0.29	A
p.P159L	D	D	N	D	D	N	N	D	D	D	U	D	D	D	D	N	11	4	1	11/15	0.73	A
p.L161P	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	16	0	0	16/16	1	D
p.R164S	N	D	D	N	N	N	N	D	N	N	U	D	N	N	N	N	4	11	1	4/15	0.27	A
p.R164G	N	D	D	N	N	N	D	D	N	N	U	D	N	D	D	N	7	8	1	7/15	0.47	A
p.R167W	D	D	D	D	D	D	N	D	D	D	D	D	D	D	D	D	15	1	0	15/16	0.94	D
p.R167Q	D	D	D	D	D	D	N	D	D	D	D	D	D	D	D	D	15	1	0	15/16	0.94	D
p.R195C	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	16	0	0	16/16	1	D
p.I205M	N	D	D	N	D	N	U	D	N	N	U	D	N	N	D	N	6	8	2	6/14	0.43	A
p.G221S	D	D	D	N	D	N	D	D	D	N	U	D	D	N	D	N	10	5	1	10/15	0.67	A

p.R225Q	D	D	D	N	D	D	D	D	D	D	U	D	D	D	D	D	14	1	1	14/15	0.93	D
p.D230Y	N	D	D	N	N	N	U	D	N	N	U	D	N	N	N	N	4	10	2	4/14	0.29	A
p.V237M	N	D	D	N	D	N	U	D	N	N	U	D	N	N	N	N	5	9	2	5/14	0.36	A
p.E242K	N	D	D	N	N	N	U	D	D	N	U	D	N	N	N	N	5	9	2	5/14	0.36	A
p.R246H	N	D	D	N	D	N	U	D	N	N	U	D	N	D	D	N	7	7	2	7/14	0.5	A
p.R246C	D	D	D	N	N	N	U	D	N	N	U	D	N	D	N	N	6	8	2	6/14	0.43	A
p.R251S	D	D	D	D	D	D	U	D	D	D	D	D	D	D	D	D	15	0	1	15/15	1	D
p.A252V	D	D	N	N	D	D	U	D	D	N	U	D	N	D	D	N	9	5	2	9/14	0.64	A
p.H256Q	N	D	D	N	N	N	U	D	N	N	N	D	N	N	N	N	4	11	1	4/15	0.27	A
p.S262N	D	D	N	N	D	N	U	D	N	D	D	D	D	N	D	N	9	6	1	9/15	0.6	A
p.D273G	N	D	D	N	N	N	U	D	N	N	U	N	N	N	N	N	3	11	2	3/14	0.21	B
p.G274A	N	D	D	N	N	N	U	D	N	N	N	D	N	N	N	N	4	11	1	4/15	0.27	A
p.V282L	D	D	D	D	D	N	U	D	D	D	U	D	D	D	D	N	12	2	2	12/14	0.86	D
p.Q296R	N	D	D	N	N	N	U	D	N	N	U	N	N	N	N	N	3	11	2	3/14	0.21	B
p.A297S	N	N	D	N	N	N	U	D	N	N	N	D	N	N	N	N	3	12	1	3/15	0.2	B
p.P309S	N	D	D	N	N	N	N	N	N	N	N	U	N	N	N	N	2	13	1	2/15	0.13	B
p.R321H	N	D	D	D	N	D	N	N	D	N	N	D	N	N	N	N	6	10	0	6/16	0.38	A
p.R321C	N	D	D	D	N	N	N	N	N	N	U	D	N	N	N	N	4	11	1	4/15	0.27	A
p.R325Q	N	D	D	D	N	N	N	N	N	N	N	D	N	N	N	N	4	12	0	4/16	0.25	B
p.A331V	D	D	D	N	D	N	U	D	D	D	U	D	D	D	D	N	11	3	2	11/14	0.79	D

p.K348R	N	D	N	N	N	N	U	D	N	N	N	N	N	N	N	N	2	13	1	2/15	0.13	B
p.R355W	D	D	D	N	D	N	U	D	D	D	D	D	D	D	D	D	13	2	1	13/15	0.87	D
p.R355Q	D	D	D	N	D	N	U	D	N	D	D	D	N	N	D	N	9	6	1	9/15	0.6	A
p.D359N	D	D	D	N	D	N	N	N	D	D	U	U	D	N	N	N	7	7	2	7/14	0.5	A

*: D = deleteriousness; N = neutral; U = unknown; A = ambiguous; B: benign

Supp Table S4. HGMD Reported Mutations Determined Benign by *in Silico* and/or Expression Assays

***HMBS* Variant Encoding:**	**U-ALA***	**U-PBG***	**RBC-HMBS Activity* (% of Mean Normal)**	**Reference**
p.T59I	NR*	Normal	60%	(Schneider-Yin et al., 2008)
p.E86V	Normal	Normal	78%	(Floderus et al., 2002)
p.E86V	Normal	Normal	89%	(Floderus et al., 2002)
p.E86V	Normal	Normal	96%	(Floderus et al., 2002)
p.R225Q	fluctuated	fluctuated	NR*	(Floderus et al., 2002)
p.R225Q	Normal	Normal	>50%	(von Brasch et al., 2004)
p.D230Y	NR*	NR*	NR*	(Whatley et al., 2009)
p.R321H /c.423-1G>C (p.R321H / IVS7$^{-1G>C}$)	↑↑	↑↑	NR*	(von Brasch et al., 2004)
p.R321H	slightly ↑	slightly ↑	Normal	(Schuurmans et al., 2001)
p.R321H	↑↑	↑↑	Normal	(Anyaegbu et al., 2012)
p.R321H	NR*	NR*	84%	(Cerbino et al., 2015)
p.R321H	Normal	Normal	53%	(Cerbino et al., 2015)
p.R321H/c.826-1G>A (p.R321H/ IVS12$^{-1G>A}$)	↑↑	↑↑	69%	(Cerbino et al., 2015)
p.R321H/c.826-1G>A (p.R321H/ IVS12$^{-1G>A}$)	↑↑	↑↑	37%	(Cerbino et al., 2015)
D359N	NR	NR	NR	(Di Pierro et al., 2006)

* U-ALA:Urinary ALA concentration; U-PBG: Urinary PBG concentration; RBC-HMBS Activity: Erthryocyte HMBS enzymatic activity;
NR: Not reported