SUPPORTING INFORMATION

Supp. Table S1. Predicted effects on protein function of GNE missense variants associated with GNE myopathy[1]

#	hGNE1[2]	hGNE2[2]	PolyPhen2[3]	SIFT[4]	Align[5]	PMut[6]	Overall[7]
1	E2G	E33G	Probably Damaging	Damaging	65	Neutral	Severe
2	R11W	R42W	Probably Damaging	Damaging	65	Pathological	Severe
3	C13S	C44S	Possibly Damaging	Damaging	65	Neutral	Medium
4	A26P	A57P	Benign	Tolerated	25	Neutral	Mild
5	P27L	P58L	Possibly Damaging	Tolerated	65	Neutral	Medium
6	P27S	P58S	Possibly Damaging	Tolerated	65	Neutral	Medium
7	I28M	I59M	Possibly Damaging	Damaging	0	Neutral	Medium
8	M29T	M60T	Possibly Damaging	Damaging	65	Neutral	Medium
9	M29R	M60R	Probably Damaging	Damaging	65	Pathological	Severe
10	E35K	E66K	Benign	Tolerated	55	Pathological	Medium
11	P36L	P67L	Probably Damaging	Damaging	65	Neutral	Severe
12	E40K	E71K	Possibly Damaging	Tolerated	55	Pathological	Medium
13	I51M	I82M	Probably Damaging	Tolerated	0	Neutral	Medium
14	M60V	M91V	Benign	Tolerated	15	Neutral	Mild
15	R71W	R102W	Probably Damaging	Damaging	65	Pathological	Severe
16	G89R	G120R	Probably Damaging	Damaging	65	Pathological	Severe
17	G89S	G120S	Probably Damaging	Tolerated	55	Neutral	Medium
18	R101C	R132C	Probably Damaging	Damaging	65	Neutral	Severe
19	R101H	R132H	Probably Damaging	Damaging	25	Neutral	Medium
20	I106T	I137T	Possibly Damaging	Damaging	65	Neutral	Medium
21	R129Q	R160Q	Probably Damaging	Tolerated	35	Pathological	Medium
22	H132Q	H163Q	Probably Damaging	Damaging	15	Neutral	Medium
23	G135V	G166V	Probably Damaging	Damaging	65	Neutral	Severe
24	G136R	G167R	Probably Damaging	Damaging	65	Pathological	Severe
25	I142T	I173T	Probably Damaging	Damaging	65	Neutral	Severe
26	I150V	I181V	Probably Damaging	Tolerated	25	Neutral	Medium
27	Y156H	Y187H	Probably Damaging	Tolerated	65	Neutral	Medium
28	R162C	R193C	Probably Damaging	Damaging	65	Pathological	Severe
29	M171V	M202V	Probably Damaging	Damaging	15	Neutral	Medium
30	D176V	D207V	Possibly Damaging	Tolerated	65	Neutral	Medium
31	R177C	R208C	Probably Damaging	Damaging	65	Neutral	Severe
32	I178N	I209N	Probably Damaging	Damaging	65	Pathological	Severe
33	I178M	I209M	Probably Damaging	Damaging	0	Neutral	Medium
34	L179F	L210F	Probably Damaging	Tolerated	15	Neutral	Medium
35	Y186C	Y217C	Probably Damaging	Damaging	65	Pathological	Severe
36	D187G	D218G	Probably Damaging	Damaging	65	Pathological	Severe
37	I200F	I231F	Benign	Damaging	15	Neutral	Medium
38	R202L	R233L	Benign	Damaging	65	Pathological	Medium
39	G206S	G237S	Possibly Damaging	Tolerated	55	Neutral	Medium
40	D208N	D239N	Benign	Tolerated	15	Neutral	Mild
41	D213V	D244V	Possibly Damaging	Tolerated	65	Neutral	Medium
42	V216A	V247A	Probably Damaging	Damaging	65	Neutral	Severe
43	Q219K	Q250K	Possibly Damaging	Tolerated	45	Neutral	Medium

44	D225N	D256N	Probably Damaging	Tolerated	15	Neutral	Medium
45	F233S	F264S	Possibly Damaging	Damaging	65	Neutral	Medium
46	I241S	I272S	Benign	Tolerated	65	Pathological	Medium
47	R246W	R277W	Probably Damaging	Damaging	65	Pathological	Severe
48	R246Q	R277Q	Possibly Damaging	Tolerated	35	Neutral	Mild
49	M261V	M292V	Possibly Damaging	Tolerated	15	Neutral	Mild
50	M261I	M292I	Benign	Tolerated	0	Neutral	Mild
51	M265T	M296T	Benign	Damaging	65	Neutral	Medium
52	I270N	I301N	Possibly Damaging	Tolerated	65	Pathological	Medium
53	I270T	I301T	Benign	Tolerated	65	Pathological	Medium
54	R277C	R308C	Probably Damaging	Tolerated	65	Pathological	Medium
55	R277G	R308G	Possibly Damaging	Tolerated	65	Pathological	Medium
56	P283S	P314S	Probably Damaging	Tolerated	65	Neutral	Medium
57	H293R	H324R	Possibly Damaging	Damaging	25	Pathological	Medium
58	G295D	G326D	Probably Damaging	Tolerated	65	Pathological	Medium
59	G295R	G326R	Probably Damaging	Tolerated	65	Pathological	Medium
60	M297T	M328T	Benign	Damaging	65	Pathological	Medium
61	I298T	I329T	Probably Damaging	Damaging	65	Pathological	Severe
62	N300K	N331K	Probably Damaging	Damaging	65	Neutral	Severe
63	C303V	C334V	Possibly Damaging	Tolerated	65	Neutral	Medium
64	G304R	G335R	Probably Damaging	Damaging	65	Pathological	Severe
65	R306Q	R337Q	Probably Damaging	Tolerated	35	Pathological	Medium
66	A310P	A341P	Probably Damaging	Damaging	25	Pathological	Severe
67	V315M	V346M	Probably Damaging	Damaging	15	Neutral	Medium
68	N317D	N348D	Probably Damaging	Tolerated	15	Neutral	Medium
69	R321C	R352C	Probably Damaging	Damaging	65	Pathological	Severe
70	V331A	V362A	Probably Damaging	Damaging	65	Neutral	Severe
71	H333R	H364R	Probably Damaging	Tolerated	25	Pathological	Medium
72	R335W	R366W	Probably Damaging	Damaging	65	Pathological	Severe
73	L347P	L378P	Probably Damaging	Damaging	65	Pathological	Severe
74	H348N[8]	H379N[8]	Benign	Tolerated	65	Neutral	Medium
75	V367I	V398I	Probably Damaging	Tolerated	25	Neutral	Medium
76	D378Y	D409Y	Probably Damaging	Damaging	65	Pathological	Severe
77	L379H	L410H	Probably Damaging	Damaging	65	Neutral	Severe
78	P390S	P421S	Probably Damaging	Tolerated	65	Pathological	Medium
79	V421A	V452A	Possibly Damaging	Tolerated	65	Neutral	Medium
80	Y434C	Y465C	Possibly Damaging	Tolerated	65	Pathological	Medium
81	C453F	C484F	Possibly Damaging	Damaging	65	Pathological	Severe
82	A460V	A491V	Possibly Damaging	Tolerated	65	Pathological	Medium
83	L463P	L494P	Probably Damaging	Damaging	65	Pathological	Severe
84	G469R	G00R	Probably Damaging	Damaging	65	Pathological	Severe
85	I472T	I503T	Possibly Damaging	Damaging	65	Pathological	Severe
86	L508S	L539S	Probably Damaging	Damaging	65	Neutral	Severe
87	H509Y	H540Y	Possibly Damaging	Tolerated	65	Neutral	Medium
88	P511H	P542H	Probably Damaging	Damaging	65	Pathological	Severe
89	P511L	P542L	Probably Damaging	Damaging	65	Pathological	Severe
90	N519S	N550S	Possibly Damaging	Tolerated	45	Pathological	Medium
91	A524V	A555V	Probably Damaging	Damaging	65	Pathological	Severe
92	F528C	F559C	Probably Damaging	Damaging	65	Pathological	Severe
93	L556S	L587S	Probably Damaging	Damaging	65	Pathological	Severe

94	I557T	I588T	Possibly Damaging	Damaging	65	Neutral	Medium
95	G559R	G590R	Probably Damaging	Damaging	65	Pathological	Severe
96	G559A	G590A	Probably Damaging	Damaging	55	Neutral	Severe
97	G568S	G599S	Probably Damaging	Damaging	55	Neutral	Severe
98	G568V	G599V	Probably Damaging	Damaging	65	Pathological	Severe
99	V572L	V603L	Probably Damaging	Damaging	25	Neutral	Medium
100	G576E	G607E	Probably Damaging	Damaging	65	Pathological	Severe
101	C579Y	C610Y	Probably Damaging	Tolerated	65	Pathological	Severe
102	C581R	C612R	Probably Damaging	Damaging	65	Pathological	Severe
103	I587T	I618T	Benign	Damaging	65	Pathological	Medium
104	I587N	I618N	Possibly Damaging	Damaging	65	Pathological	Medium
105	A591T	A622T	Probably Damaging	Damaging	55	Neutral	Severe
106	A600E	A631E	Probably Damaging	Damaging	65	Pathological	Severe
107	A600T	A631T	Probably Damaging	Tolerated	55	Neutral	Medium
108	L603F	L634F	Probably Damaging	Damaging	15	Neutral	Medium
109	A630T	A661T	Probably Damaging	Tolerated	55	Neutral	Medium
110	A631T	A662T	Probably Damaging	Damaging	55	Neutral	Severe
111	A631V	A662V	Probably Damaging	Damaging	65	Pathological	Severe
112	N635K	N666K	Probably Damaging	Damaging	65	Neutral	Severe
113	A648V	A679V	Probably Damaging	Tolerated	65	Pathological	Medium
114	I656N	I687N	Probably Damaging	Damaging	65	Pathological	Severe
115	G669R	G700R	Probably Damaging	Damaging	65	Pathological	Severe
116	Y675H	Y706H	Probably Damaging	Tolerated	65	Neutral	Medium
117	V679G	V710G	Probably Damaging	Damaging	65	Pathological	Severe
118	V696M	V727M	Probably Damaging	Damaging	15	Neutral	Medium
119	S699L	S730L	Probably Damaging	Damaging	65	Pathological	Severe
120	G708S	G739S	Probably Damaging	Damaging	55	Neutral	Severe
121	M712T	M743T	Benign	Damaging	65	Pathological	Severe
Frequent missense SNPs[9]							
1	D208E	D239E	Benign	Tolerated	35	Neutral	Mild
2	I423V	I454V	Possibly Damaging	Tolerated	25	Neutral	Mild
3	V696L	V727L	Benign	Tolerated	25	Neutral	Mild

[1] See main text (**Table 1**) for sequence variants and literature references; Predicted effects on the protein calculated by each program are color-coded: red = severe; yellow = medium; green = mild/not deleterious.

[2] Variants are indicated in hGNE1 (NP_005467.1) and hGNE2 (NP_001121699.1) nomenclature. See main text and Table 1footnotes for nomenclature details.

[3] **POLYPHEN 2**: POLYmorphism PHENotyping progrm 2: (http://genetics.bwh.harvard.edu/pph2/). This program was used in January 2014, utilizing version 2.2.2 of the software, protein sequences from UniProtKB/UniRef100 Release 2011_12 (14-Dec-2011), structures from PDB/DSSP Snapshot 03-Jan-2012 (78,304 entries) and UCSC MultiZ multiple alignments of 45 vertebrate genomes with hg19/GRCh37 human genome (08-Oct-2009) [Adzhubei et al., 2010].

[4] **SIFT**: Sort Intolerant From Tolerant human Protein: (http://sift.jcvi.org/) [Ng and Henikoff, 2003].

[5] **Align-GVGD:** (http://agvgd.iarc.fr/agvgd_input.php) [Tavtigian et al., 2005; Mathe et al., 2006]. This program was used in January 2014 in the Align-GVGD version of 30/10/2013. Scores: <16 = mild; 16-34= medium; >34= severe.

[6] **PMut: (**http://mmb2.pcb.ub.es:8080/PMut/). This program was used in the PMut version available in January 2014 [Ferrer-Costa et al., 2004].

[7] Overall severity was scored by combining the scores of the 4 prediction programs:
Red (Severe) when: 4 severe scores; 3 severe and 1 medium; 3 severe and 1 mild.
Green (Mild) when: 4 mild scores; 3 mild and 1 medium; 2 mild and 2 medium.
Yellow (Medium) when: any other combination of scores.

[8] Variant H379N (hGNE2) results from an indel sequence variant (L378del;H379N).

[9] As retrieved from exome databases. See text for details.

Supp. Table S2. Predicted splicing effects of selected *GNE* variants

Variant (NM_001128227.2)	Location	Exp. splice effect?[1]	Predicted splicing effect[2]		References
			wt	var	
Intronic variants					
c.710-4A>G	in 4	ND	67%	60%	[Cho et al., 2013]
c.862+4A>G	in 5	Yes	92%	81%	[Nishino et al., 2002]
c.1076-1delG	in 6	ND	60%	82%	[Cho et al. 2013]
c.1163+2dupT	in 7	Yes	98%	99%	[Broccolini et al., 2004]
c.1504+5G>A	in 9	ND	90%	14%	[Cho et al., 2013]
c.1505-4G>A	in 9	ND	5%	9%	[Cho et al., 2013]
c.1909+5G>A	in 11	Yes	16%	0%	[Boyden et al., 2011]
Exonic variants in proximity of splice junctions[3]					
c.709G>A/p.G237S	ex 4	ND	100%	94%	[Broccolini et al., 2004]
c.710delG/p.G237Vfs*3	ex 5	ND	100%	12%	[Broccolini et al., 2004])
c.715G>A/p.D239N	ex 5	ND	31%	30%	[Sim et al., 2013]
c.1508T>C/p.I503T	ex 10	ND	5%	5%	[Nishino et al. 2002; Yabe et al., 2003]
c.1727delG/p.G576Efs*11	ex 11	ND	88%	95%	[Park et al., 2012]

Severity predictions are visualized in the same colors as in **Supp. Table S1**.
Red (Severe), **Yellow** (Medium), **Green** (Mild).
The DNA numbering is based on cDNA sequence of the longest mRNA splice variant NM_001128227.2 (encoding hGNE2 protein). Nucleotide numbering uses +1 as the A of the ATG translation initiation codon, with the initiation codon as codon 1.

[1]Experimental (exp) splice site effect: as reported in literature; ND = Not Done. Yes= Splice site effect was demonstrated experimentally.

[2]Using Human Splice Site Prediction by Neural Network (http://www.fruitfly.org/seq_tools/splice.html), comparing wild type (wt) sequence to variant (var) sequence. The NNSPLICE 0.9 version (January 1997) of this splice site predictor was used (in January 2014) [Reese et al., 1997].

[3]Exonic cDNA variants resulting in protein variants located within 5-bp from a splice junction are listed.

Supp. Table S3. Allele frequency of *GNE* variants in exome databases

Variants (NM_001128227.2)			Frequency per database			Total
hGNE2	**cDNA**	**Name**	**1000 Genomes**	**ESP**	**NIH-UDP**	
p.Q14Rfs*3	c.41_42delAA	TMP_ESP_9_36276997	-	1/10620*	-	1/10620
p.N23I	c.68A>T	rs199965140	1/2184	-	-	1/2184
p.R50H	c.149G>A	NA	-	-	1/1434	1/1434
p.L56F	c.166C>T	rs373398528	-	1/13006	-	1/13006
p.R108S	c.324G>T	rs141892824	-	1/13006	-	1/13006
p.L126M	c.376C>A	TMP_ESP_9_36246361	-	1/13006	-	1/13006
p.R132H	c.395G>A	rs144727134	-	1/13006	-	1/13006
p.R132C	c.394C>T	rs148523065	-	1/13006	-	1/13006
p.R160*	c.479C>T	rs372872777	-	1/13006	-	1/13006
p.D207V	c.620A>T	rs139425890	2/2184	-	-	2/2184
p.I231F	c.691A>T	TMP_ESP_9_36246046	-	1/13006	-	1/13006
p.D239E	**c.717T>G**	**rs35224402**	**11/2184**	**60/13006**	-	**71/15290**
p.A287T	c.859G>A	TMP_ESP_9_36236832	-	1/13006	-	1/13006
p.R294*	c.580C>T	rs200643106	1/2184	-	-	1/2184
p.R366W	c.1096C>T	rs150132839	-	1/13006	-	1/13006
p.D409Y	c.1276G>T	rs199877522	-	3/13006	-	3/13006
p.V443I	c.1327G>A	TMP_ESP_9_36227292	-	1/13006	-	1/13006
p.I454V	**c.1360A>G**	**rs35638832**	**1/2184**	**8/13006**	-	**9/15290**
p.F468S	c.1403T>C	TMP_ESP_9_36223471	-	1/13004	-	1/13004
p.I480V	c.1438A>G	TMP_ESP_9_36223436	-	1/13004	-	1/13004
p.Q482*	c.1444C>T	rs189454495	1/2184	-	-	1/2184
p.R512Q	c.1535G>A	rs138357804	1/2184	2/13006	-	3/15290
p.P534R	c.1601C>G	TMP_ESP_9_36222899	-	1/13006	-	1/13006
p.G564E	c.1691G>A	rs201808007	1/2184	-	-	1/2184
p.I572T	c.1715T>C	TMP_ESP_9_36222785	-	2/13006	-	2/13006
p.G578C	c.1732G>T	NA	-	-	2/1434	2/1430
p.E586A	c.1757A>C	rs200212703	1/2184	-	-	1/2184
p.V603L	c.1807G>C	rs121908632	1/2184	-	-	1/2184
p.L634F	c.1900C>T	TMP_ESP_9_36219844	-	1/13006	-	1/13006
p.A655V	c.1964C>T	rs200278654	-	2/13006	-	2/13006
p.R715S	c.2145C>A	rs139347806	-	1/13006	-	1/13006
p.R715C	c.2143C>T	NA	-	-	1/1434	1/1434
p.V727L	**c.2179G>T**	**rs121908627**	-	**17/13006**	-	**17/13006**
p.V727G	c.2180T>G	rs142031240	-	1/4552*	-	1/4552
p.V733A	c.2198T>C	TMP_ESP_9_36217426	-	1/13006	-	1/13006
p.A736T	c.2206G>A	rs201216576	1/2184	-	-	1/2184
	Total alleles*		**22/2184**	**112/13006**	**4/1434**	**138/16624**
	Total alleles without 3 SNPS		10/2184	27/13006	4/1434	41/16624
	q		0.00457875	0.00207596	0.0027894	0.00246631
	Allele frequency without 3 SNPs		**1/218**	**1/482**	**1/360**	**1/406**
	2pq		0.00915751	0.00415193	0.0055788	0.00493263
	Carrier rate (heterozygotes)		**1/109**	**1/241**	**1/180**	**1/203**
	q^2		2.096*10-5	4.31*10-6	7.78*10-6	6.08*10-6
	Prevalence of GNE myopathy		**1/47710**	**1/232019**	**1/128522**	**1/164474**

NA, Not Available. Red highlight, variant identified in GNE myopathy patient (listed in Table 1); Green highlight, frequent variants, considered SNPs. Databases were searched for missense, nonsense and indel variants located within the coding region of hGNE2 (intronic variants were not included). Allele frequency and prevalence was calculated according to the Hardy-Weinberg equation ($p^2 + 2pq + q^2 = 1$).
*Total alleles were (conservatively) counted based on the highest number of alleles in each database. Total alleles were not adjusted (i.e., in ESP database 1/4552 was counted as 1 variant in 13006 alleles)