SUPPLEMENTARY INFORMATION

doi:10.1038/nature22034

Supplementary Table 1. pLoF variants observed in the homozygous null state

After calling and quality control, we annotated variants using Variant Effect Predictor and the Loftee plugin.[1,2] We subsequently extracted, when observed, the number of individuals per mutation who were carriers of two copies of each high confidence pLoF mutation. The numbers of unique homozygous pLoF carriers per gene are also tabulated; when this value is less than the sum of unique carriers per mutation, this indicates the presence of individuals with more than one homozygous pLoF variant for the same gene. The table is sorted by the number of unique homozygotes per gene in descending order. cDNA changes are prepended by Ensembl transcript accession ID. Similarly, amino acid changes, where applicable, are prepended by Ensembl protein accession IDs. We intersected this list with the NHGRI-EBI Catalog of Published Genome-Wide Association Studies and note any genome-wide-significant ($P \leq 5x10^{-8}$) association fell into the region of -/+200kbp transcription starting/ending sites of the gene. We also performed manual curation of all discovered homozygous pLoF variants (using criteria listed in Supplementary Table 8) and provide our impressions about variant call fidelity and anticipated biochemical impact.

Please refer to *Supplementary Table 1.xlsx*.

Supplementary Table 2. Phenotypes analyzed in recessive model pLoF association analysis.

Adiponectin, log	Citrulline, log	Lipoprotein-associated phospholipase A2 mass
Alanine, log	ClDC / C4DC, log	Lipoprotein(a), log
Albumin	Complement factor H binding to malondialdehyde modified LDL	Macrophage inflammatory protein 1 alpha, log
Alkalinine phophatase	Creatinine, log	Macrophage inflammatory protein 1 beta, log
ALT, log	Current smoker	Malondialdehyde modified LDL
Angiopoeitin-like 3, log	Cystatin C, log	Malondialdehyde modified LDL antibody, log
Angiopoeitin-like 4, log	Diabetes mellitus, type 2	Matrix metallopeptidase 3
Apolipoprotein AI	Diastolic blood pressure	Matrix metalloproteinaise 9, categorical
Apolipoprotein AII, log	E-selectin, log	Matrix metalloproteinase 2, categorical
Apolipoprotein AV, log	Eotaxin 1, log	Matrix metalloproteinase 9
Apolipoprotein B	Factor VII	Methionine, log
Apolipoprotein CIII, log	Fasting glucose, log	Monocyte chemoattractant protein 1
Apolipoprotein E, log	Ferritin, log	Myeloperoxidase, log
Arginine, log	Fibroblast growth factor 21, log	Myocardial Infarction
Asparagine, log	Glutamine, log	NanoMPC, log
AST, log	Glycine, log	Nerve growth factor beta, categorical
BMI	Granulocyte colony-stimulating factor, log	Neuronal cell adhesion molecular, log
Brain derived neurotrophic factor	Granulocyte macrophage colony-stimulating factor, categorical	Non-esterified fatty acid, log
C-reactive protein, log	Growth differentiation factor 15, log	NT-pro B-natriuretic peptide, log
C10 1, log	HDL cholesterol	Ornithine, log
C10 2	HDL2, log	Oxidized phospholipids on apolipoprotein B-100
C10 3, log	HDL3, log	P-selectin, log
C10, log	Heart rate, EKG	Peripheral arterial disease
C10OH / C8DC, log	Heart-type fatty acid binding protein, categorical	Phenylalanine, log
C12 1, log	Hemoglobin A1c	Phosphorus
C12, log	Histidine, log	Placental growth factor

C12OH / C10DC	Hydroxy-butyrate, log	Potassium
C14 1, log	Hypertension	PR interval, EKG
C14 2, log	IDL1	Pregnancy-associated plasma protein A, categorical
C14, log	IDL2	Pregnancy-associated plasma protein A, log
C141OH / C121DC	IgM immune complexes, log	Proline, log
C14OH / C12DC	IgM PCBS, log	QRS complex, log, EKG
C16 1, log	Insulin C-peptide, log	Rheumatoid factor, categorical
C16 2, log	Insulin-like growth factor 1, log	Rheumatoid factor, log
C16, log	Insulin, log	S100 calcium binding protein B, categorical
C161OH / C141DC	Intercellular adhesion molecule 1, log	Serine, log
C16OH / C14DC	Interferon gamm, categorical	Serum amyloid A, log
C18 1, log	Interleukin 1 alpha, categorical	Sex-hormone binding globulin, log
C18 2, log	Interleukin 1 beta, categorical	Sodium
C18, log	Interleukin 1 receptor antagonist, categorical	Soluble receptor for advanced glycation end-products, log
C181OH / C161DC	Interleukin 10, log	Soluble vascular endothelial growth factor, log
C182OH	Interleukin 12 p40, categorical	Sortilin, log
C18OH / C16DC	Interleukin 12 p70, categorical	Stem cell factor, log
C2, log	Interleukin 15, categorical	Stromal cell-derived factor 1
C20	Interleukin 17, categorical	Superoxide dismutase 1, log
C20 4, log	Interleukin 18, log	Systolic blood pressure
C201OH / C181DC	Interleukin 2, categorical	Thrombomodulin, log
C20OH / C18DC	Interleukin 23, categorical	Thyroid stimulating hormone, log
C22	Interleukin 3, categorical	Total cholesterol
C3, log	Interleukin 4, categorical	Triglycerides, log
C4Cl4, log	Interleukin 5, categorical	Tumor necrosis factor alpha, categorical
C4OH, log	Interleukin 7, categorical	Tumor necrosis factor beta, categorical
C5 1, log	Interleukin 8, log	Tyrosine, log
C5DC, log	Iron	Uric acid
C5OH / C3DC, log	Ketones, log	Valine, log
C5S, log	Lactate, log	Vascular cell adhesion molecule 1, log
C6, log	LDL cholesterol	Vascular endothelial growth

		factor, log
C61DC / C81OH, log	LDL peak diameter	VLDL, intermediate, log
C6DC, log	LDL1, log	VLDL, large, log
C7DC, log	LDL2A, log	VLDL, small, log
C8 1, log	LDL2B, log	Waist:hip ratio
C8, log	LDL3A, log	
C81DC, log	LDL3B, log	
Calcium	LDL4A, log	
CD40 ligand, log	LDL4B, log	
CD40, log	LDL4C, log	
Ceruloplasmin	Lecithin-type oxidized LDL receptor 1	
Chloride	Leucine isoleucine, log	
Cholesterol efflux	Lipoprotein mid-zone, log	
Chorionic gonadotropin alpha, log	Lipoprotein-associated phospholipase A2 activity	

We measured 201 categorical and quantitative clinic traits and biomarkers across participants listed above alphabetically that were available to test for association in an autosomal pLoF recessive model.
Traits in red indicate those obtained by history, physical or EKG.
HDL = high density lipoprotein; LDL = low density lipoprotein; VLDL = very low density lipoprotein; IDL = intermediate density lipoprotein; EKG = electrocardiogram; BMI = body mass index

Supplementary Table 3. Significant gene-trait recessive model pLoF association results.

Gene	Variants	Trait	Phenotyped Carriers	Beta	SE	*P*
USP45	c.2101dupA(p.Arg701LysfsTer20), c.1636A>T(p.Lys546Ter)	QRS complex, log, EKG	2	0.48	0.05	1.40E-23
APOC3	c.55C>T(p.Arg19Ter)	Apolipoprotein CIII, log	4	-2.20	0.22	4.82E-23
A3GALT2	c.315_318delGACT(p.Thr106Serfs)	Insulin C-peptide, log	2	-3.68	0.53	6.17E-12
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Intermediate density lipoprotein 1. log	4	-1.85	0.28	9.67E-11
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Very low density lipoprotein, intermediate, log	4	-1.54	0.26	3.74E-09
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Very low density lipoprotein, small, log	4	-1.55	0.27	5.98E-09
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Low density lipoprotein 2A, log	4	-2.05	0.35	7.62E-09
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Low density lipoprotein 2B, log	4	-2.11	0.37	1.78E-08
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Midzone lipoprotein, log	4	-1.57	0.28	2.84E-08
APOC3	c.55C>T(p.Arg19Ter)	High density lipoprotein cholesterol	4	26.95	4.85	2.88E-08
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Intermediate density lipoprotein 2. log	4	-1.69	0.31	9.09E-08
OR6C75	c.305delT(p.Phe102SerfsTer6)	Apolipoprotein AII, log	2	-0.83	0.16	1.58E-07
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Low density lipoprotein 1, log	4	-1.91	0.36	1.63E-07

UBA7	c.1189G>T(p.Glu397Ter)	QRS complex, log, EKG	5	0.25	0.05	1.98E-07
PLA2G7	c.663+1G>A	Lipoprotein-associated phospholipase A2 activity	2	-245.3	47.12	1.99E-07
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Low density lipoprotein 4A, log	4	-1.62	0.31	2.05E-07
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Low density lipoprotein 4B, log	4	-1.58	0.31	4.70E-07
CIB3	c.198+1G>T	Apolipoprotein CIII, log	2	-1.64	0.31	1.75E-07
FAM151A	c.1003_1031delCCTGGGGATGACGGTCTGAATGTG GAGTG(p.Pro335AlafsTer4), c.940+1G>A	Interleukin 10, log	2	1.59	0.31	5.05E-07
PIF1	c.145G>T(p.Glu49Ter)	QRS complex, log, EKG	2	0.24	0.05	5.19E-07
PHLDB3	c.853delG(p.Val285SerfsTer20)	High density lipoprotein 3, log	2	-2.58	0.53	9.41E-07
APOC3	c.55C>T(p.Arg19Ter)	Apolipoprotein AI	4	58.90	12.06	1.06E-06
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Low density lipoprotein 3A, log	4	-1.99	0.41	1.25E-06
TREH	c.90-9_106delTCTCTGCAGTGAGATTTACTGCCACG	Very low density lipoprotein, large, log	4	-1.26	0.26	1.35E-06
CYP2F1	c.1295-2A>G	Interleukin 8, log	2	3.72	0.78	1.75E-06
KRTAP5-4	c.98_99delCC(p.Ser33TrpfsTer110)	C22	2	0.005	0.001	2.78E-06

After ascertaining the unique homozygous pLoF carriers for each gene, we tested each gene separately for autosomal recessive association with each of 201 traits (Supplementary Table 4). We restricted analyses to gene-trait pairs where there were at least two phenotyped carriers yielding 18,959 gene-trait pairs. To reduce the likelihood of observing false associations, we employ a Bonferroni *P* threshold of 0.05 / 18,959 = 3 x 10^{-6}. Here, we report the associations that have *P* values less than the prespecified threshold.
EKG = electrocardiogram

Supplementary Table 4. Association of pLOF mutations in LDLR and PCSK9 with risk for myocardial infarction.

Gene	MI Cases (Carriers / Total)	Controls (Carriers / Total)	MI OR	*P**
LDLR	17 / 4,793	1 / 5,710	20.3	2×10^{-4}
PCSK9	2 / 4,793	11 / 5,710	0.22	0.05

pLoF alleles in Mendelian genes previously associated with LDL cholesterol and risk of myocardial infarction are demonstrated as positive controls.
**P*-values were estimated by logistic regression accounting sex and principal components of ancestry.

Supplementary Table 5. Significant gene-trait recessive model associations for proteomic scan.

Gene	Analyte	Carrier	Non-Carrier	Relative effect	95% CI	*P*
KIAA1919	Oncostatin-M	31249.1 [16113.5, 46384.8]	1042.9 [921.5, 1149.8]	7.34	4.11, 13.11	3.31E-09
SLC9A3R1	Calcium/calmodulin-dependent protein kinase type II subunit beta	3991.4 [3241.2, 4551.4]	543.5 [422.2, 721.8]	5.38	3.11, 9.31	6.86E-08
SLC9A3R1	Calcium/calmodulin-dependent protein kinase type II subunit delta	18246.3 [14697.6, 22838.5]	2060.1 [1632.3, 3016.7]	6.27	3.33, 11.8	2.54E-07
SLC9A3R1	Cytoplasmic tyrosine-protein kinase BMX	371.9 [316.3, 1958.8]	207.4 [183.7, 229.2]	3.01	2.05, 4.41	3.22E-07
SLC9A3R1	cAMP-regulated phosphoprotein 19	2374.4 [1840.2, 2952.8]	393.1 [338.6, 520.8]	4.33	2.59, 7.21	3.44E-07
SLC9A3R1	ICOS ligand	3892 [3358, 4768.6]	653.4 [597.2, 891.6]	4.88	2.79, 8.52	4.08E-07
SLC9A3R1	Calcium/calmodulin-dependent protein kinase type II subunit alpha	1051 [644.9, 1491.7]	318.6 [285.1, 378.2]	2.94	1.99, 4.32	6.48E-07
CST9L	Betacellulin	9519.8 [4962.9, 14076.8]	389.9 [355.3, 473.2]	6.44	3.29, 12.63	7.35E-07
KRTAP5-4	Hemoglobin	75015.1 [37840, 112190.2]	567.2 [381.6, 858.7]	14.82	5.52, 39.8	9.96E-07
SLC9A3R1	Signal transducer and activator of transcription 3	2300.5 [1621.7, 3090.8]	273.2 [220.2, 457.6]	5.86	3.05, 11.28	1.20E-06
SLC9A3R1	Allograft inflammatory factor 1	5649.2 [3105, 8240.5]	1497.7 [1292.3, 1612.8]	2.64	1.84, 3.78	1.26E-06
SLC9A3R1	Mitogen-activated protein kinase 14	3063.6 [2354.3, 3948.9]	596 [497.9, 809.1]	4.16	2.44, 7.08	1.50E-06
SLC9A3R1	40S ribosomal protein S7	3655.8 [2562.9, 4782]	617 [540.6, 814.7]	3.81	2.3, 6.31	1.81E-06
A3GALT2	Protein disulfide-isomerase	1522.8 [843.1, 2202.6]	197.7 [174.4, 227.4]	3.68	2.24, 6.03	2.16E-06
SLC9A3R1	Proto-oncogene vav	13683 [9381, 19803.1]	1260.8 [809, 2014.8]	8.15	3.67, 18.11	2.20E-06
SLC9A3R1	Signal transducer and activator of transcription 6	1117.6 [809.5, 1459.6]	281.9 [223.4, 370.2]	3.2	2.05, 4.99	2.30E-06
SLC9A3R1	Signal transducer and activator of transcription 1-alpha/beta	12812.4 [7620.3, 18920.9]	1195 [828, 2060.8]	8.6	3.77, 19.63	2.51E-06
KIAA1919	TATA-box-binding protein	5042.8 [2694.4, 7391.3]	266.1 [234, 329.4]	5.42	2.82, 10.42	3.04E-06
SLC9A3R1	Cofilin-1	11804.8 [9540.7, 13583.9]	2603.9 [1973.1, 3446.7]	3.31	2.08, 5.27	3.41E-06
SLC9A3R1	Mitogen-activated protein kinase 8	2720.1 [2376.8, 2978.1]	427.6 [330, 596.6]	4.7	2.56, 8.63	4.02E-06

We performed proteomic analyses in 84 individuals. In these participants, 9 genes (*A3GALT2*, *APOC3*, *CST9L*, *CYP2F1*, *KIAA1919*, *KRTAP5-4*, *SIGLEC9*, *SIGLECL1*, and *SLC9A3R1)* were observed at least twice in the homozygous pLoF state. We associated the homozygous pLoF genotype across 1,310 protein analytes. Median [interquartile range] concentrations by homozygous pLoF genotype status are presented. Our threshold for significance was 0.05 / (1,310 analytes x 9 genes) = 4.3 x 10^{-6}. Effect estimates are for log transformed analytes. Bonferroni-adjusted significant associations are shown here.

Supplementary Table 6. Summary of manual curation approach for homozygous pLoF variants.

Category	Flag	Criteria
Technical	Mapping	Large proportion of polymorphisms within mapped reads in a region of low mapping quality (by human chained self alignment[6]).
	Genotyping	Low number (<7) of reads at variant site or low proportion (<80%) of reads with variant.
Rescue	Rescuing in-phase polymorphism	Presence of an additional polymorphism within the same codon rescuing a putative nonsense mutation, or an additional nearby frameshift mutation restoring the reading frame of another frameshift mutation.
	Splice site rescue	Evidence of redundancy of splice sites ('NAGNAG motifs')[7] indicating capability of alternative splicing.
Impact	Affects minority of transcripts	Less than half of the annotated transcripts are affected despite the canonical transcript being affected as annotated by Ensembl.
	Weak exon/splice site conservation	Evidence of weak conservation throughout the affected exon or splice site indicating potential lack of biological function; base-wise conservation was complementarily assessed in mammals with Genomic Evolutionary Rate Profiling (GERP)[8] and vertebrates with Phylogenetic *P*-values (phyloP)[9].

All discovered 1,580 homozygous pLoF genotypes from the 1,317 pLoF genes were manually curated to estimate genotyping errors and evidence of mitigation of putative pLoF mutations.

Supplementary Table 7. Summary of manual curation flags for homozygous pLoF variants.

Category	Flag	Flagged Variants N = 1,580
Technical	Mapping	70 (4.4%)
	Genotyping	56 (3.5%)
Rescue	Rescuing in-phase polymorphism	10 (0.6%)
	Splice site rescue	29 (1.8%)
Impact	Affects minority of transcripts	15 (0.9%)
	Weak exon/splice site conservation	8 (0.5%)
Total Unique		177 (11.2%)

A minority of variants had multiple flags. Most overlap occurred within the Technical category and there were 111 unique flags within this category.

Supplementary Table 8. Observed homozygous pLoF genes that are predicted to be not tolerated.

Gene	Mouse knockout complete lethality	Predicted human essential by mouse/human conservation	Predicted human LoF intolerant
ADRA2B	N	Y	N
AFF3	N	N	Y
APPL2	Y	N	N
ARID3A	N	Y	N
ARID4A	N	N	Y
ATRN	N	N	Y
ATXN2	N	Y	Y
BDKRB2	N	Y	N
BDNF	Y	Y	Y
BLM	Y	Y	N
CDK12	N	N	Y
CHGA	N	Y	N
CLIP1	N	N	Y
CLK3	N	N	Y
CNTF	N	Y	N
CNTN1	Y	Y	Y
COL1A2	Y	Y	Y
COL5A3	Y	N	N
COL7A1	Y	Y	N
CPVL	Y	Y	N
CUBN	N	Y	N
CUL9	N	N	Y
DMBT1	N	Y	N
DOCK1	Y	Y	N
DSC3	N	Y	N
EP400	Y	Y	Y
FLG	Y	Y	NA
FMN2	N	N	Y
FMNL2	N	N	Y
GDNF	Y	Y	N
GOLGA8M	NA	NA	Y
GPHN	Y	Y	Y
HBA2	Y	N	N
HSPG2	N	Y	N
HTT	Y	Y	Y
IQSEC3	N	N	Y
IRF2BPL	N	N	Y
JAK1	Y	Y	Y
KDM2B	N	Y	Y

KIF14	Y	N	N
KRT10	Y	Y	N
KRT4	N	Y	N
LGALS9C	N	N	Y
LPIN1	N	Y	N
LTN1	N	Y	N
MAML2	N	N	Y
MAPK8IP3	Y	Y	Y
MBIP	Y	N	N
MCM9	Y	Y	N
MCOLN3	N	Y	N
MEOX2	N	Y	N
MEST	N	Y	Y
MIER3	N	N	Y
MINK1	N	N	Y
MTERF	Y	N	N
MYH15	Y	N	N
NCOA3	N	Y	Y
NCOR1	N	Y	Y
NR3C1	Y	Y	N
NUMB	N	Y	N
ORAI1	Y	Y	N
PDZD2	N	N	Y
PHACTR4	Y	Y	N
PHRF1	N	N	Y
PKD1L1	N	Y	N
PKD2	N	Y	Y
POP4	Y	N	N
PSMC6	N	N	Y
RAD54L	Y	N	N
RBM26	N	N	Y
RFC1	Y	N	N
RGPD3	Y	N	NA
RILP	Y	N	N
RYR1	Y	Y	N
SATB1	Y	Y	N
SCNN1A	Y	Y	N
SLC9A3R1	N	Y	N
SLX4	N	Y	N
SPINK5	Y	Y	N
SUCO	Y	NA	N
TBP	N	Y	N
TEP1	Y	N	N

TERF1	N	Y	N
TGFBR3	N	Y	N
TGIF1	N	Y	N
TMPRSS6	N	Y	N
TTN	Y	Y	N
UNC5B	N	Y	N
UPK2	N	Y	N
XIRP2	Y	Y	N
YME1L1	N	N	Y
ZC3H12C	N	N	Y
ZFHX3	Y	N	Y
ZNF384	N	N	Y

We intersected the list of 1,317 homozygous pLoF genes observed in the PROMIS cohorts with prior lists of genes where a homozygous pLoF genotype is predicted to not be tolerated in humans. Genes that are present in at least one of the three presented lists here are displayed. We obtained the list of phenotypes from mouse knockout experiments from the Mouse Genome Informatics (MGI) database.[3] We extracted genes, when knocked out, have been associated with "complete lethality" in the "embryonic," "prenatal," "perinatal," "postnatal," "neonatal" phase in the mouse any experiment. *GOLGA8M* is coded as "NA" since there is no mouse ortholog. We next intersected a dataset of predicted human essential genes based on mouse/human conservation data.[4] Neither *SUCO* nor *GOLGA8M* are in this dataset and are coded as "NA." Next, we intersected the list with LoF constraint scores derived from human whole exome sequencing studies.[5] RGPD3 and FLG were not in the dataset and are coded as "NA."

Supplementary Table 9. Genes with relative depletion of homozygous pLoF genotypes.

Gene	Combined MAF	Heterozygotes Observed	Homozygotes Observed	Homozygotes Expected	Chi Square *P*
IRF2BPL	0.0377	277	3	19.49	1.83E-74
ATXN2	0.0231	160	1	8.45	2.00E-48
SLC25A36	0.0097	56	0	2.21	3.71E-24
NTSR2	0.0113	83	0	2.74	2.01E-22
TMEM78	0.018	187	3	5.65	5.16E-21
HRCT1	0.0097	70	2	2.19	3.36E-19
POP4	0.0096	72	1	2.15	2.48E-18
HLA-DPA1	0.0145	146	3	4	8.70E-18
DNAH14	0.03	432	3	13.13	8.11E-14
TMPRSS13	0.0105	110	2	2.48	2.57E-12
CRIPAK	0.0315	468	9	14.28	3.54E-11
MEF2A	0.0054	44	0	0.99	5.02E-10
MADCAM1	0.0046	33	0	0.8	9.93E-10
ATP11A	0.0063	61	0	1.21	3.73E-09
RAI1	0.0076	86	0	1.56	4.53E-08
PRKD3	0.0041	32	0	0.69	5.56E-08
POLR2A	0.0022	7	0	0.32	9.24E-08
GPATCH1	0.0038	29	0	0.62	1.51E-07
KRT1	0.0033	22	0	0.52	1.71E-07
TARS	0.0045	40	0	0.77	1.81E-07
PDIA2	0.0246	393	3	9.41	7.81E-07
MUC19	0.017	254	2	5.14	1.04E-06
ECD	0.0061	68	0	1.14	1.15E-06
GTF2IRD2	0.0034	28	0	0.55	1.88E-06
HDGFRP2	0.0019	7	0	0.27	2.07E-06
PTX4	0.0072	89	0	1.45	2.24E-06
SLC27A3	0.0118	170	0	2.95	3.18E-06
SRGAP1	0.0037	34	0	0.61	3.90E-06
DSPP	0.0091	123	1	2.01	9.63E-06
C2ORF15	0.0032	29	0	0.52	1.52E-05
FRYL	0.0068	88	0	1.32	3.67E-05
METTL20	0.0094	134	0	2.09	3.87E-05
CTTNBP2NL	0.0025	21	0	0.38	1.03E-04
HMX3	0.0039	43	0	0.64	1.32E-04
DZANK1	0.0109	164	1	2.62	1.38E-04
MAML3	0.003	30	0	0.47	1.66E-04
STT3A	0.0031	31	0	0.48	1.84E-04

YBX1	0.0066	90	0	1.28	2.06E-04
RBM23	0.0047	58	0	0.82	2.12E-04
CACNA1A	0.0026	25	0	0.4	2.56E-04
TGIF1	0.0065	88	1	1.24	4.36E-04
MESP2	0.0015	9	0	0.21	5.00E-04
HLA-A	0.0047	61	0	0.83	5.66E-04
TBK1	0.0019	15	0	0.27	6.54E-04
MAGEF1	0.0024	23	0	0.36	6.84E-04
C9orf173	0.0027	28	0	0.42	6.99E-04
CORO1C	0.0033	37	0	0.52	7.51E-04
LAMP1	0.0019	16	0	0.28	7.63E-04
SNX27	0.0014	9	0	0.2	9.27E-04
FAM155A	0.0032	36	0	0.5	9.91E-04
VWDE	0.0123	194	3	3.12	1.07E-03
CEP170	0.0014	9	0	0.19	1.29E-03
ARGFX	0.0069	100	1	1.36	1.70E-03
DUSP5	0.0024	25	0	0.36	1.88E-03
C15orf48	0.0026	28	0	0.39	1.92E-03
HIP1	0.006	86	0	1.12	2.01E-03
KRTAP9-6	0.0014	10	0	0.19	2.41E-03
KIAA1109	0.0032	39	0	0.51	3.01E-03
TIMELESS	0.004	52	0	0.66	3.05E-03
STOX2	0.0021	21	0	0.3	3.61E-03
PLCL1	0.0011	7	0	0.15	3.66E-03
ISPD	0.0016	15	0	0.24	4.07E-03
HCK	0.0013	10	0	0.18	4.49E-03
THAP2	0.0015	13	0	0.21	5.12E-03
MMD2	0.0011	8	0	0.16	5.44E-03
APOBR	0.0019	19	0	0.27	5.70E-03
SKOR1	0.0139	232	3	3.76	6.21E-03
KIAA1468	0.0012	9	0	0.16	6.76E-03
SEMA6B	0.0009	5	0	0.12	7.81E-03
SCAPER	0.0022	26	0	0.34	9.04E-03
HABP4	0.0014	13	0	0.19	1.36E-02
CDCP2	0.007	112	0	1.38	1.40E-02
DGKQ	0.002	24	0	0.3	1.62E-02
RBPMS	0.0017	18	0	0.24	1.76E-02
ADAM11	0.0013	12	0	0.18	1.85E-02
CLPP	0.0012	11	0	0.16	2.09E-02
NOMO3	0.0015	16	0	0.21	2.49E-02

YLPM1	0.0014	14	0	0.19	2.57E-02
MEF2B	0.0013	13	0	0.18	2.63E-02
GALNT9	0.0045	69	0	0.78	2.67E-02
GALNT6	0.0022	28	0	0.33	2.94E-02
ZNF195	0.0056	89	0	1.02	3.22E-02
PLEKHG5	0.0013	13	0	0.18	3.25E-02
LINC00955	0.0083	142	0	1.77	3.29E-02
BAI1	0.0009	7	0	0.12	3.42E-02
TRIM64B	0.0083	142	0	1.76	3.67E-02
AC096644.1	0.0058	95	0	1.09	3.67E-02
WDR73	0.0014	16	0	0.2	3.71E-02
MICAL1	0.0039	59	0	0.65	3.81E-02
RTN4	0.0015	18	0	0.22	4.23E-02

Using the combined minor allele frequency of null alleles and observed inbreeding, the predicted number of homozygotes was estimated and compared to the observed number of homozygotes per gene. Genes that demonstrated at least nominal significant ($P < 0.05$) for homozygote depletion are displayed.

Supplementary References

1 McLaren, W. *et al*. Deriving the Consequences of Genomic Variants with the Ensembl Api and Snp Effect Predictor. *Bioinformatics* **26**, 2069-2070, doi:10.1093/bioinformatics/btq330 (2010).

2 Karczewski, K. J. *Loftee (Loss-of-Function Transcript Effect Estimator)*, <https://github.com/konradjk/loftee> (2015).

3 Eppig, J. T. *et al*. The Mouse Genome Database (Mgd): Facilitating Mouse as a Model for Human Biology and Disease. *Nucleic Acids Res* **43**, D726-736, doi:10.1093/nar/gku967 (2015).

4 Georgi, B., Voight, B. F. & Bucan, M. From Mouse to Human: Evolutionary Genomics Analysis of Human Orthologs of Essential Genes. *PLoS Genet* **9**, e1003484, doi:10.1371/journal.pgen.1003484 (2013).

5 Samocha, K. E. *et al*. A Framework for the Interpretation of De Novo Mutation in Human Disease. *Nat Genet* **46**, 944-950, doi:10.1038/ng.3050 (2014).

6 Kent, W. J., Baertsch, R., Hinrichs, A., Miller, W. & Haussler, D. Evolution's Cauldron: Duplication, Deletion, and Rearrangement in the Mouse and Human Genomes. *Proc Natl Acad Sci U S A* **100**, 11484-11489, doi:10.1073/pnas.1932072100 (2003).

7 Hiller, M. *et al*. Widespread Occurrence of Alternative Splicing at Nagnag Acceptors Contributes to Proteome Plasticity. *Nat Genet* **36**, 1255-1257, doi:10.1038/ng1469 (2004).

8 Cooper, G. M. *et al*. Single-Nucleotide Evolutionary Constraint Scores Highlight Disease-Causing Mutations. *Nat Methods* **7**, 250-251, doi:10.1038/nmeth0410-250 (2010).

9 Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R. & Siepel, A. Detection of Nonneutral Substitution Rates on Mammalian Phylogenies. *Genome Res* **20**, 110-121, doi:10.1101/gr.097857.109 (2010).

10 Sulem, P. *et al*. Identification of a Large Set of Rare Complete Human Knockouts. *Nat Genet*, doi:10.1038/ng.3243 (2015).