Application Note

Improving the Diagnostic Path of Complex Neurological Disorders

How using trusted data from HGMD Professional increases diagnostic yield and reduces diagnostic error

Introduction to genetic testing in neurological diseases

The first study on global gene expression profiling from the healthy human brain showed that more than 80% of all human genes are actively expressed in brain structures (1). Neurodegenerative diseases show considerable genetic heterogeneity with different sequences in the same gene that may cause different phenotypes and in the same phenotype that can be created by many different genes. Even though the contribution of genetic factors to neurological disorders has long been recognized, most patients did not receive a molecular diagnosis because of old sequencing technologies and the phenotypic/genotypic heterogeneity underlying these disorders. Patients received little or no prognostic and therapeutic information, and no risk calculations for disease recurrence or implications for family members.

These challenges were addressed with the next generation sequencing (NGS) approach, which significantly increased the rate of molecular diagnosis in neurological disorders (2). NGS has not only enabled wider and rapid molecular diagnostics but has also helped identify many previously undiagnosed and unrecognized neurogenetic disorders, revealing an increasing demand for comprehensive genetic testing in patients with neurological diseases (3). Thanks to these advancements, inappropriate therapies for neurodegenerative diseases have been gradually eliminated and new targeted therapies are being offered to match patients` unique genetic profiles.

Whole-genome sequencing/whole-exome sequencing (WES/WGS) in neurological disorders- opportunities and challenges

Targeted genetic testing for particular types of mutations in neurological disorders might yield false-negative results because rare genetic variants might be associated with atypical forms of diseases. By covering over 98% of coding sequences, WGS is the most powerful method for identifying genetic

variants. A successful example of WGS in neurological diseases was the identification of a mutation in a Charcot-Marie-Tooth disease case (4). For premature epilepsy and sensory and motor neuropathy with microcephaly, WGS yielded important results revealing new genes and improving molecular diagnostics (5,6). WGS could be used as a reliable ally to detect expanded repeats which are common causes for some disorders such as Huntington's disease or Amyotrophic lateral sclerosis (7). In addition, WGS may offer detection in non-coding genomic regions which might be important for neurological syndromes which show a high prevalence of non-coding variations (8). The downside of WGS, however, is that it yields 4.5–5.0 million single-nucleotide and insertion-deletion variants per sample most of which are of uncertain clinical significance. Even when likely benign and benign polymorphisms have been filtered out, more than 400,000 variants are left for clinical interpretation—making this an extremely challenging task for clinicians.

WES examines coding regions of more than 20,000 known human genes and is the current method of choice for diagnosing rare diseases. This technique is particularly efficient in phenotypically variable conditions such as neurological disorders and has been used for the molecular diagnosis of many neurological phenotype categories, such as intellectual disabilities and other neurodevelopmental disorders, cerebellar ataxias, and epilepsies (9,10). In the cases when disease phenotypes are properly identified, molecular diagnosis established by WES can be as high as 94% (11).

Regardless of whether the chosen genetic test encompasses only a couple of genes, exome or whole genome, the processing of NGS data requires computationally sophisticated bioinformatics analysis for raw data processing as well as strong software and man-power for variant classification and clinical interpretation.

During the bioinformatic and clinical processing of NGS data that will result in the final medical assessment, all detected variants must be checked for if and how they are annotated in genomic databases. However, the existence of significant discrepancies in variant reporting and their classification requires additional scrutiny and caution. Some of the existing discrepancies between databases can only be resolved by manually scanning conflicting journals, assessing and reviewing supplementary materials in the articles, or directly contacting the authors.

HGMD combines electronic and human search procedures for data curation in order to provide high-quality information. HGMD is regularly updated by a team of expert curators. They screen peer-reviewed biomedical literature on an ongoing basis via manual inspection of over 250 journals to classify variants as disease-causing or possibly disease-causing (for Mendelian conditions), or as disease-associated (for multifactorial diseases). HGMD professional version users can assess the most up to date database with additional information and extra features such as additional literature reports, chromosomal coordinates, population frequency data, or functional prediction.

Use HGMD to design a gene panel for Alzheimer's disease and to optimize mutational screening strategies

Genetic factors may explain many of the elements influencing the risk of familial and early-onset Alzheimer's disease (EOAD). Known genes included in the pathogenesis of EOAD are amyloid precursor protein (APP), presenilin 1 (PSEN1), and presenilin 2 (PSEN2). Interestingly, late-onset Alzheimer's disease (LOAD) is genetically more complex than EOAD showing other genetic sites associated with various types of dementia. WES helped reveal more than 50 non-synonymous variants in LOAD risk factor genes (12) and rare variants in SORL1 and ABCA7 genes in both EOAD and LOAD (13).

Having such complexity underlying a disorder makes the decision on the genetic test difficult. The choice of a genetic test can be of crucial importance for proper molecular diagnostics of any genetic disorder. Since most of the available databases are gene-oriented and contain not just disease-associated variants but VUS and benign polymorphisms as well, it might be difficult for clinicians to get fast and easy views of the mutational profile of a particular disease. Phenotype search in HGMD can help in choosing how deep and how wide genetic testing should be. For example, HGMD shows 843 mutations in more than 100 genes associated with Alzheimer's disease indicating the need for comprehensive genetic analysis (Figure 1). The information on the number of genes and the number of variants that are associated with a disorder is very important for performing population studies and choosing the gene panels for analysis. It also helps in the optimization of mutational screening strategies.

1 to 100 of 102 results page 1 of 2 Show 100 results

Phenotype	Gene symbol	Number of mutations
Affective/Apathetic syndromes, in Alzheimer disease, association with	APOE	1 mutation
Aggression/Agitation, in Alzheimer disease, association with	APOE	1 mutation
Agitation/Aggression-delusion, in Alzheimer disease, association with	APOE	1 mutation
Alzheimer disease	PSEN1 ABCA7 GRN ANGPTL3 CR1 APP PSEN2 SQSTM1 LRRK2 NOTCH3 HTRA1 DNMBP PRNP VCP DNAH14 RPS16 SORL1 C9orf72 bcl2l11tv13 GLIS3 RIN3 KANSL1 TSPOAP1 PLCG2 TM2D3 INPP5D ACE EPHA1 COBL PDGFRL MS4A6A OR5I1I2 CELSR1 HELZ2 ohcentv2 USP32 LAMC3 DOCK4 MTHFD1 OBSCN UBAP2 SPHK2 KMT5B GTSE1 DLEC1 TREM2 CHMP2B CRMP1 EPHA5 MAPT CDH2 EPHA6 ABCD4 SPATA7 TAS1R3 SCFD1 CCDC18 KIF19 THNSL1 CLDN7 FAM171A2 PPP1R14A TTN tmnovex3 L3MBTL2 CHRNB2 CHRNA4 APH1A PSENEN MEOX2 SRCAP HFE CSF1R SERPINI1 UBQLN2 DCTN1 GRIN2B dpp6i1 OGG1 UNC13C SORCS2 CLU PCDH11X FUS CPE PIN1 TBK1 CHCHD10 TARDBP ARC BIN1 psen1mv ADAM17 trem2v NLGN1 CUBN CD33 SETX SIGMAR1 EWSR1 SPG11 GAK CALHM1 APOE mapttv6 LRP10 TREML4 TREML2 MS4A1 MS4A7 EIF2AK3 NAMPT MRTFB THBS2 DAAM2 plekhg5tv6 AKAP9 GALR3 MIEF1 CLEC1 CTNNA1 CD163L1 UNC5C inpp5dtvx1 NME8 OPTN VLDLR ERMP1 CEP290 SEZ6 ANG CLN5 CD2AP ALPL OR56B1 C14orf28 PLCD1 IL6 RTN3 GIMAP2 CHCHD2	843 mutations
Alzheimer disease / frontotemporal dementia, increased risk	TREM2	1 mutation
Alzheimer disease and frontotemporal dementia	TREM2	2 mutations
Alzheimer disease in African Americans, association with	DIO2	1 mutation
Alzheimer disease in APOE4 non-carriers, association	BDNF	2 mutations
Alzheimer disease in e4 carriers, association	PSENEN	1 mutation
Alzheimer disease in Han Chinese, association with	C7	1 mutation

Figure 1. Phenotype search in HGMD shows 843 mutations (disease-associated mutations and disease-associated functional polymorphisms) detected in more than 100 genes associated with Alzheimer's disease.

Use HGMD to find up-to-date information on variants that are not available in other clinical databases

HGMD offers in depth information on variants and contains those that are not available in other clinical databases. SORL1 gene has been identified to associate with Alzheimer's disease (AD) through replicated genetic studies. Studies indicate its possible role in the progression of this disease making SORL1 a potential target for AD therapy (14). Thus, missing the clinically important SORL1 variations might negatively impact patient care. Gene-specific search in HGMD shows 164 reported variants in SORL1 that have clinical implications (Figure 2).

Figure 2. HGMD contains 164 reported variants in SORL1 that have potential clinical implications. All variants are sorted according to the variant type and variant class.

Disease-associated SORL1 variants are easy to find in HGMD either through their clinical effect or the type of mutation. The list of disease-associated mutations in SORL1 gene is presented in the Figure 3.

If a mutation such as c.372C>A, p.S124R is detected in SORL1, its classification and association with phenotype might be difficult to establish. dbSNP contains an entry for this particular variant (rs1306611994) but has no information on its clinical significance. There is no information on scientific publications related to this variant in dbSNP, either. This variant is not reported in the ClinVar. However, SORL1 c.372C>A, p.S124R is reported as the disease-associated in HGMD for EOAD. This classification has been backed up with the latest publications that investigate causative mutations and genetic risk factors in sporadic EOAD before 51 years (15) (Figure 4).

HGMD accession	HGMD codon change	HGMD amino acid change	HGVS (nucleotide)	HGVS (protein)	Variant class	Reported phenotype	Reference	Extra information
CM1719336	AGC-AGA	Ser124Arg	c.372C>A	p.S124R	DM	Alzheimer disease, early onset	Bellenguez (2017) Neurobiol Aging 59, 220.e1 Lacour (2019) J Alzheimers Dis 71: 227 [Additional report]	hg38 hg19 COM dbSNP
CM166729	AGC-AGG	Ser124Arg	c.372C>G	p.S124R	DM	Alzheimer disease, early onset	Nicolas (2016) Mol Psychiatry 21, 831 Bellenguez (2017) Neurobiol Aging 59: 220.e1-e9 [Additional report]	hg38 hg19
CM166726	CGA-TGA	Arg268Term	c.802C>T	p.R268*	DM	Alzheimer disease, early onset	Nicolas (2016) Mol Psychiatry 21, 831 Bellenguez (2017) Neurobiol Aging 59: 220.e1-e9 [Additional report]	hg38 hg19 CpG dbSNP
CM164095	CGA-TGA	Arg416Term	c.1246C>T	p.R416*	DM	Alzheimer disease, early onset	Verheijen (2016) Acta Neuropathol 132, 213 Holstege (2017) Eur J Hum Genet 25: 973 [Additional report] Thonberg (2017) Acta Neuropathol Commun 5: 43 [Additional report]	hg38 hg19 CpG dbSNP
CM166731	TGT-TCT	Cys473Ser	c.1418G>C	p.C473S	DM	Alzheimer disease, early onset	Nicolas (2016) Mol Psychiatry 21, 831 Bellenguez (2017) Neurobiol Aging 59: 220.e1-e9 [Additional report]	hg38 hg19
CM200314	CGG-TGG	Arg490Trp	c.1468C>T	p.R490W	DM	Alzheimer disease, early-onset	Park (2020) Neurobiol Aging 85, 155.e5	hg38 hg19 CpG dbSNP
CM123106	GGA-CGA	Gly511Arg	c.1531G>C	p.G511R	DM	Alzheimer disease, early onset	Pottier (2012) Mol Psychiatry 17, 875 Caglayan (2014) Sci Transl Med 6: 223ra20 [Functional characterisation] Bellenguez (2017) Neurobiol Aging 59: 220.e1-e9 [Additional report] 1 more reference(s)...	hg38 hg19 COM
CM166732	GGA-GAA	Gly543Glu	c.1628G>A	p.G543E	DM	Alzheimer disease, early onset	Nicolas (2016) Mol Psychiatry 21, 831 Bellenguez (2017) Neurobiol Aging 59: 220.e1-e9 [Additional report]	hg38 hg19

Figure 3. The list of SORL1 disease-associated mutations in HGMD.

HGMD accession	Reported disease/phenotype	Variant class	Gene symbol	Codon change	Amino acid change	Codon number	Feedback
CM1719336	Alzheimer disease, early onset	DM	SORL1	AGC-AGA	Ser-Arg	124	Feedback

The S124R substitution exhibits a shift in polarity from polar to positively charged and displays a decrease in Kyte-Doolittle hydrophobicity from -0.8 to -4.5. Approximately 0.67% of missense mutations in HGMD are Ser-Arg. The mutation occurs 2091 amino acids from the end of the protein.

Literature citation	Citation type	Support	Comments/notes
1. Bellenguez (2017) *Neurobiol Aging* **59:** 220.e1 PubMed: 28789839 Contribution to Alzheimer's disease risk of rare variants in TREM2, SORL1, and ABCA7 in 1779 cases and 1273 controls.	Primary literature report		see Supplementary Table 7.
2. Lacour (2019) *J Alzheimers Dis* **71:** 227 PubMed: 31381512 Causative Mutations and Genetic Risk Factors in Sporadic Early Onset Alzheimer's Disease Before 51 Years.	Additional literature report		Described as risk factor.

Extra information	
Coding strand genomic sequence (GRCh38)	CAACGTGATCGTGGCCTTGGCCCGAGATAG(C/A)CTGGCATTGGCGAGGCCCAAGAGCAGTGAT
Genomic coordinate (GRCh38)	chr11:121470093
Genome viewers	UCSC; UCSC (codon); NCBI Genome Data Viewer; NCBI SeqViewer
HGVS nomenclature	NM_003105.6: c.372C>A; NP_003096.2: p.S124R
Variant Call Format (VCF)	CHROM POS ID REF ALT 11 121470093 CM1719336 C A
Protein structures	Q92673; INstruct;
dbSNP number	rs1306611994
HGMD variant class	Disease causing mutation
HGMD computed rankscore	0.79000
CpG	No

Figure 4. SORL1 c.372C>A, p.S124R in HGMD. Detailed information on the evidence used for classification of this variant is available for users

Once this situation happens in clinical practice, clinicians find themselves facing different interpretations for the detected variant. If the variant is misclassified as VUS instead of likely pathogenic, patients might be left without diagnosis and appropriate treatment. Not having up-to-date information might cause clinicians to miss the important implication of this variant to the EOAD, particularly if the patient exhibits an uncommon clinical phenotype.

Use HGMD to reach the decision on variant classification with the in-depth information on causative variants

HGMD offers additional information such as literature reports, chromosomal coordinates, population frequency data, and functional prediction for every variant. HGMD curators adopt a policy of continual content curation, commenting and annotating new information to the users. When new evidence suggests a benign nature of a variant, it may be removed from the database at the discretion of experienced curators. Variant reclassification continually takes place in the HGMD giving high-quality data to the end-users, making sure they don't waste valuable time on benign variants or polymorphisms. HGMD is the only database that pursues a policy of continuous curation and reclassification wherever necessary not relying solely on the original submitter updating their submission.

Use HGMD to enhance the understanding of different variants associated with neurological disorders

Even in the cases of variants in known disease-associated genes, variant classification in neurological disorders may be complicated. Mutations in the PSEN1 gene are found to be a common cause of familial Alzheimer's disease. Variants in this gene could be difficult to interpret due to the conflicting or incomplete information in available databases. c.356C>T (p.Thr119Ile) missense variant in PSEN1 gene is classified as likely pathogenic in ClinVar with the latest update in April 2019 and incomplete evidence. HGMD offers newly discovered data that go in favor of its pathogenicity in EOAD (Figure 5). These novel data may help to enhance our understanding of different variants associated with dementia in the study population.

Figure 5. PSEN1 c.356C>T (p.Thr119Ile) in HGMD. This mutation is classified as disease-associated in HGMD. The classification is backed up with the latest scientific publications no older than a year

Use HGMD to find ethnically relevant variants in Parkinson's disease

Parkinson's disease (PD) affects millions of people and is the second most common neurodegenerative disorder. Several genes including PRKN, PINK, and ATP13A2 have been involved in PD pathogenesis and traditionally investigated in this disease. Findings have shown that only about 5–10% of PD patients have uniform forms of the disease. NGS can be used to determine the effect of genes on PD and to discover genes associated with different forms of the disease. HGMD phenotype search shows more than 80 genes and 800 genetic variations associated with PD. Unlike others, which are gene/locus/variant-specific databases, HGMD can provide a quick overview of the number of disease-related genes and variants (Figure 6).

Parkinson disease	LRRK2 PSAP GALC GBA ACMSD PRKN ABCB1 COQ2 EIF4G1 PLA2G6 SYNJ1 PINK1 VPS35 mapttv6 NOTCH3 GCH1 NOTCH2NLC SNCA NR4A2 SMPD1 NUBPL NOD2 GIGYF2 FBXO7 ATP13A2 SLC41A1 TBP ATG7 POLG PARK7 HTRA2 ATG5 PITX3 VPS13C NPC1 UCHL1 DCTN1 RHOT1 SH3GL2 STAB1 ATG12 arsbtv2 APOE LRP10 TARDBP CSMD1 TMEM230 HSPA9 DNAJC13 PTPRH UHRF1BP1L GPATCH2L CHCHD2 ms4a6atv1 MAPT RAB39B ANK2 ovos2 NDUFV2 SCARB2 TNR tnk2tv1 ANG ATP7B PRNP NUS1 C9orf72 SIRT1 SNCAIP NPC2 CR1 SORL1 ALS2 SETX NME8 SPG11 APP TREM2 ABCA7 PIK2B FUS NDUFAF5 FLNA dstlv2 NEFM ETH1 LAMP2 PARL PLXNA4 EEF1D LRRK1 TENM4 GLA ARSA	980 mutations
Parkinson disease / frontotemporal lobar degeneration	NPC2	1 mutation
Parkinson disease / Gaucher disease 3	GBA	1 mutation
Parkinson disease & dementia	SNCA	3 mutations
Parkinson disease & optic atrophy	SLC25A46	4 mutations
Parkinson disease 15	FBXO7	1 mutation
Parkinson disease 17	VPS35	1 mutation
Parkinson disease and ADHD	SLC6A3	2 mutations
Parkinson disease and paraquat use, association with	GSTT1	1 mutation
Parkinson disease association with pesticide exposure	ABCB1	1 mutation
Parkinson disease dementia	LRRK2 LRP10 GBA	4 mutations
Parkinson disease dementia, association with	SORL1 PITX3	2 mutations
Parkinson disease in IBMPFD	VCP	1 mutation
Parkinson disease related pain, association with	SCN9A	1 mutation
Parkinson disease with dementia	GBA PRKN PSEN2	7 mutations
Parkinson disease with dementia, association with	CDKN1A	2 mutations

Figure 6. Phenotype search in HGMD shows more than 80 genes and more than 800 genetic variations (disease-associated mutations and disease-associated functional polymorphisms) associated with PD

Most of the known PD mutations are found through research conducted in European, North American, North African Arab or Asian populations. Limited studies exist on the genetics of PD in the Black African populations even though African populations have more private alleles than any other population. NGS is a perfect approach to identifying novel genetic variants and disease-associated mutations in such populations. However, novel variants are difficult to classify especially since most of the databases contain neither entries nor evidence to help elucidate their clinical importance.

Heterozygous missense variant in one of the known PD genes, ATP13A2 (S1004R) was one of the rare variants detected in the study of Parkinson's disease in Black South African and Nigerian patients (16). It was found in a 39 years old patient from South Africa. None of the currently available databases contain information on this variant, except HGMD. This variation is classified as likely pathogenic in HGMD (Figure 7,8).

Figure 7. ATP13A2 S1004R missense mutation classified as likely pathogenic in HGMD.

Amino acid comparison		
Trait	Ser (S)	Arg (R)
Amino acid name	serine	arginine
Polarity/charge	polar	positively charged
pH	neutral	basic
Residue weight	87	156
Hydrophobicity score	-0.8	-4.5
Hydrophilicity score	0.3	3.0
Secondary structure propensity	α indifferent β breaker	α indifferent β indifferent
Grantham difference	110	
MutPred likelihood of being deleterious	NONE CALCULATED	

dbNSFP3.5 predictions	
PolyPhen2 prediction B (benign), P (possibly damaging), D (probably damaging)	Variable: D;P;D
SIFT prediction	Damaging
LRT prediction	Deleterious
MutationTaster prediction	Disease causing
MutationAssessor prediction	Medium impact
FATHMM	Damaging
fathmm-MKL	Damaging
M-CAP	Damaging
CADD The larger the score the more likely the SNP is damaging (PHRED-like)	28.9
MetaSVM	Damaging
MetaLR	Damaging
PhyloP 20way The larger the score, the more conserved the site (max 1.199000).	1.049000
PhyloP 100way The larger the score, the more conserved the site (max 10.003000).	4.248000
GERP RS The larger the score, the more conserved the site (max 6.17).	5.12
1000 Genomes	No data
gnomAD	No data
Interpro domain	P-type ATPase, transmembrane domain

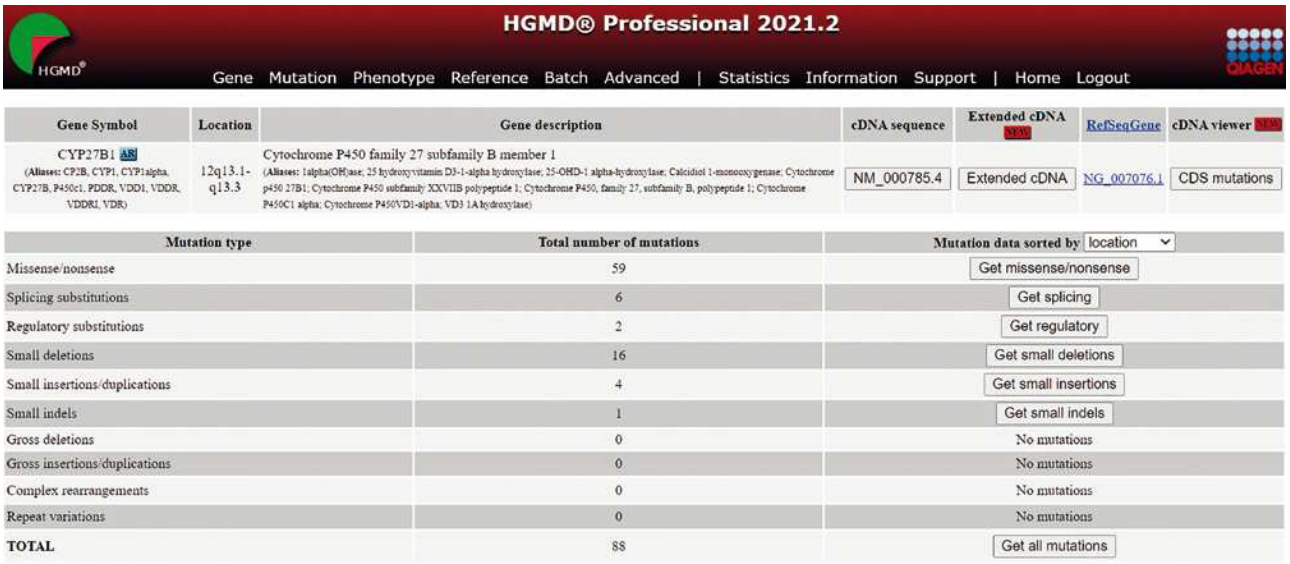
Figure 8. Amino acid comparison and in-silico analysis are presented in the HGMD for ATP13A2 S1004R missense mutation as supporting evidence for its classification

HGMD and VUS reclassification

Sequencing heterogeneous neurological diseases will yield many novel variants whose classification should be available for clinicians for fast and easy review. In addition to rare deleterious variants, many VUS will be identified. VUS reclassification may take a lot of valuable clinician time. HGMD contains only variants that have clinical importance which can help in filtering out those whose effect on the disease is not yet known. Unlike other sources that include practically all submitted VUS without further inspecting them, HGMD reclassifies VUS when enough manually curated evidence is present. Compared to other sources, the decision of whether to include a variant into the HGMD database is not an arbitrary one. The decision involves an exercise of expert curation and these variants are specifically marked in the HGMD to indicate that some degree of uncertainty exists.

Use HGMD to identify susceptibility loci for a multiple sclerosis (MS) panel

NGS has helped in revealing the association of MS with the human leukocyte antigen HLA genes and also revealed novel variants related to this disease (17,18). These newly discovered genes might help not only in molecular diagnostics of MS but can give information on new therapeutic strategies. Eighty-four disease-associated genes have been reported in HGMD for MS. Using HGMD, it is easy to find a gene of interest and investigate the reported variants associated with the disease (Figure 9,10).



Gene Symbol	Location	Gene description	cDNA sequence	Extended cDNA NEW	RefSeqGene	cDNA viewer NEW
CYP27B1 AR (Aliases: CP2B, CYP1, CYP1alpha, CYP27B, P450c1, PDDR, VDD1, VDDR, VDDR1, VDR)	12q13.1-q13.3	Cytochrome P450 family 27 subfamily B member 1 (Aliases: 1alpha(OH)ase, 25 hydroxyvitamin D3-1-alpha hydroxylase, 25-OHD-1 alpha-hydroxylase, Calcidiol 1-monooxygenase, Cytochrome p450 27B1, Cytochrome P450 subfamily XXVIIB polypeptide 1, Cytochrome P450, family 27, subfamily B, polypeptide 1, Cytochrome P450C1 alpha, Cytochrome P450VD1-alpha, VD3 1A hydroxylase)	NM_000785.4	Extended cDNA	NG_007076.1	CDS mutations

Mutation type	Total number of mutations	Mutation data sorted by location
Missense/nonsense	59	Get missense/nonsense
Splicing substitutions	6	Get splicing
Regulatory substitutions	2	Get regulatory
Small deletions	16	Get small deletions
Small insertions/duplications	4	Get small insertions
Small indels	1	Get small indels
Gross deletions	0	No mutations
Gross insertions/duplications	0	No mutations
Complex rearrangements	0	No mutations
Repeat variations	0	No mutations
TOTAL	88	Get all mutations

Figure 9. CYP27B1 gene has 86 variants with potential clinical implications reported in HGMD. All variants are sorted according to the mutation type and variant class.

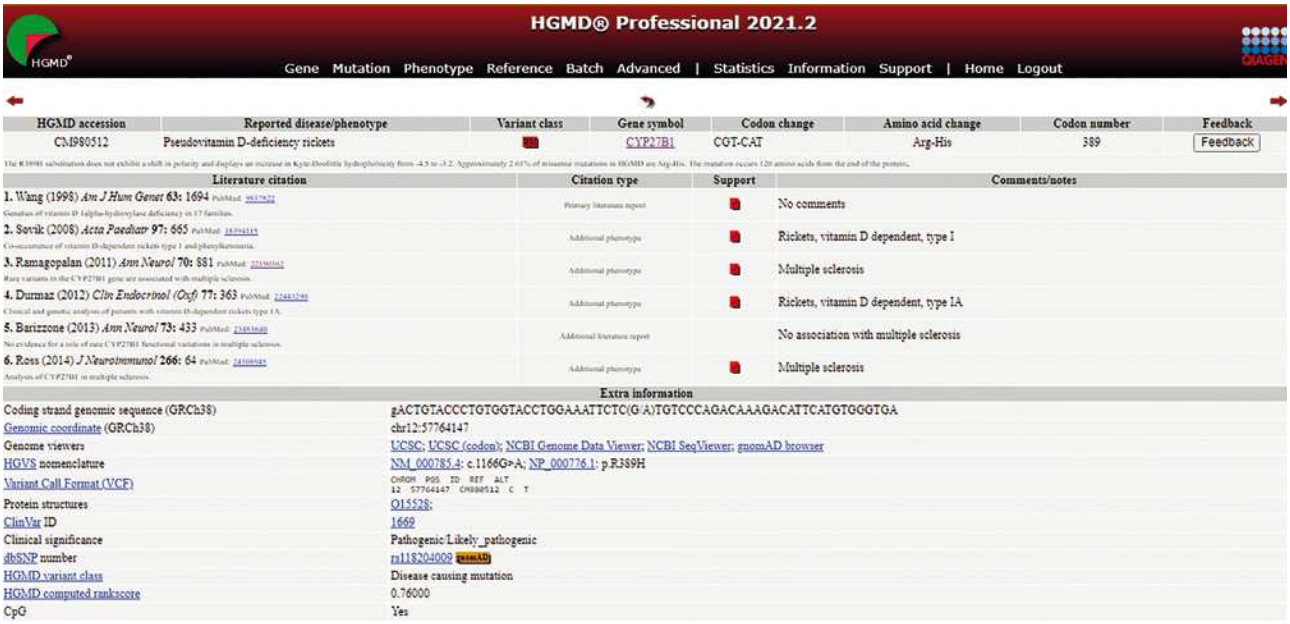


Figure 10. Direct links to evidence used for variant classification are available for user review

However, it seems that known alleles are not sufficient to induce MS and that rare variants with greater effect sizes may still not be identified. GWAS studies have revealed over 230 MS risk alleles across the human genome, highlighting its complex genetic architecture. The great challenges remain regarding the translation of these findings into an etiological framework and actionable clinical understanding. New susceptibility loci have been discovered spanning hundreds of kilobases (kb) and many tens of genes on different chromosomes (19). It has been shown that 2% of MS heritability resides in the newly investigated genomic regions. HGMD offers an advanced feature for chromosome search and easy access to potential susceptibility loci.

Figure 11. The search for susceptibility loci and genes can go through Advanced option in HGMD- chromosome search

The classification conundrum – how to correctly interpret a variant? Use HGMD as an aid to variant classification

Looking at the variant frequency in the control population, one can easily misclassify a variant as benign/likely benign. c.1529C>T, p.Ala510Val is a common missense mutation in an SPG7 gene which has been associated with hereditary spastic paraplegia, and pure cerebellar ataxia (20). Literature data on the significance of this particular variant are conflicting and variant classification range from VUS to pathogenic. The frequency of this variant goes in favor of its benign nature. It has been reported in GnomAD in 820 of 282,858 alleles including homozygotes. Since clinical laboratories have been unable to reach a consensus on the interpretation, clinicians face the difficult task to decide how to counsel and treat patients with hereditary spastic paraplegia and this variant.

HGMD classifies this variant as disease-associated, backing up the classification with more than 30 up-to-date scientific publications. Amino acid comparisons and in-silico analysis also go in favor of its pathogenicity (Figure 12, 13).

Figure 12. c.1529C>T, p.Ala510Val missense mutation in a SPG7 gene. The classification is based upon more than 30 scientific publications

Amino acid comparison		
Trait	Ala (A)	Val (V)
Amino acid name	alanine	valine
Polarity/charge	non-polar	non-polar
pH	neutral	neutral
Residue weight	71	99
Hydrophobicity score	1.8	4.2
Hydrophilicity score	-0.5	-1.5
Secondary structure propensity	strong α former β indifferent	α former strong β former
Grantham difference	64	
MutPred likelihood of being deleterious	VERY HIGH RISK	

dbNSFP3.5 predictions	
PolyPhen2 prediction	Probably damaging
SIFT prediction	Damaging
LRT prediction	Deleterious
MutationTaster prediction	Disease causing
MutationAssessor prediction	Medium impact
FATHMM	Damaging
fathmm-MKL	Damaging
M-CAP	Damaging
CADD The larger the score the more likely the SNP is damaging (PHRED-like)	32
MetaSVM	Damaging
MetaLR	Damaging
PhyloP 20way The larger the score, the more conserved the site (max 1.199000).	0.935000
PhyloP 100way The larger the score, the more conserved the site (max 10.003000).	4.972000
GERP RS The larger the score, the more conserved the site (max 6.17).	5.42
1000 Genomes	0.0022/11
gnomAD	0.00288/709
Interpro domain	P-loop containing nucleoside triphosphate hydrolase

Figure 13. Amino acid comparison and in-silico analysis are presented in the HGMD for c.1529C>T, p.Ala510Val missense mutation in a SPG7 gene as evidence for its classification

Conclusion

HGMD provides clinicians with the most up-to-date information about disease-associated variants. In the realm of complex neurological disorders, it can be vital to adequate diagnostics.

Large quantities of data currently available might be erroneous or incomplete, and therefore of questionable value to clinical decision-making. Fortunately, HGMD offers high quality manually curated data and variants that are reliably classified and associated with the disease of interest. This ensures that little time is wasted going through polymorphisms, unrelated literature, or unverified information.

Clinical genomics has high demands in terms of quality because final results are as good or bad as the quality of data used. HGMD can also help in the optimization of mutational screening strategies as it provides valuable data for clinical interpretive use in exome screening studies.

HGMD Public vs. HGMD Pro

Feature	HGMD Public	HGMD Pro
Up-to-date content		
Displays mutations 3 years or older	X	
Updates mutations every 3 months		X
Search features		
Search by gene symbol	X	X
Search by gene description	X	X
Search by OMIM number	X	X
Search by disease/phenotype	X	X
Search missense/nonsense variants	X	X
Search splice mutations	X	X
Search regulatory mutations	X	X
Search small deletions	X	X
Search small indels	X	X
Search gross deletions	X	X
Search gross insertions	X	X
Search complex rearrangements	X	X
Search repeat variations	X	X
Search by chromosomal location		X
Search by HGNC/OMIM/GDB/Entrez ID		X
Search by RefSeq transcript		X
Search by gene ontology		X
Search using operators (+,-,*,"")		X
Search phenotype using UMLS semantic		X
Search phenotype using HGMD phenotype		X
Search references by first author		X
Search references by PubMed journal		X
Search references by PubMed ID		X
Search references by publication year		X
Search references by HGMD gene		X
Search references by Medline journal abbreviation		X
Batch search		X
Advanced search (by substitution, motif, function,etc.)		X

HGMD Public vs. HGMD Pro

Feature	HGMD Public	HGMD Pro
Display features		
HGMD accession ID	X	X
Codon change	X	X
Amino acid change	X	X
Codon number	X	X
Associated phenotype	X	X
References	X	X
Misense/nonsense mutations	X	X
Splicing mutations	X	X
Regulatory mutations	X	X
Small deletions	X	X
Small insertions	X	X
Gross deletions	X	X
Gross insertions/duplications	X	X
Complex rearrangements	X	X
Repeat variations	X	X
cDNA sequence	X	X
Extended cDNA		X
Mutation's first published report		X
Related genes		X
Gene ontology		X
Variant class (DM, DM?, FP, DP, DFP)		X
Gene aliases		X
Mutation sorted by location		X
Mutation sorted by phenotype		X
Mutation sorted by author		X
Mutation sorted by year		X
Mutation sorted by entrydate		X
Extra information (HGVS, VCF, rankscore, etc.)		X
Comparison between hg19 and hg38		X
Amino acid comparison		X
dbNSFP predictions (CADD, MutationTaster, SIFT, Polyphen, etc.		X
Orthologous amino acid conservation comparison		X

References

1. Negi, S. K., & Guda, C. (2017). Global gene expression profiling of healthy human brain and its application in studying neurological disorders. Scientific reports, 7(1), 897.
2. Xue, Y., Ankala, A., Wilcox, W. R., & Hegde, M. R. (2015). Solving the molecular diagnostic testing conundrum for Mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing. Genetics in medicine: official journal of the American College of Medical Genetics, 17(6), 444–451.
3. Boycott, K. M., et al. (2017). International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. American journal of human genetics, 100(5), 695–705.
4. Pang, S. Y., et al. (2017). The role of gene variants in the pathogenesis of neurodegenerative disorders as revealed by next generation sequencing studies: a review. Translational neurodegeneration, 6, 27.
5. Martin, H. C., et al. (2014). Clinical whole-genome sequencing in severe early-onset epilepsy reveals new genes and improves molecular diagnosis. Human molecular genetics, 23(12), 3200–3211.
6. Gonzaga-Jauregui, C., et al. (2013). Mutations in VRK1 associated with complex motor and sensory axonal neuropathy plus microcephaly. JAMA neurology, 70(12), 1491–1498.
7. Bahlo, M., et al. (2018). Recent advances in the detection of repeat expansions with short-read next-generation sequencing. F1000Research, 7, F1000 Faculty Rev-736.
8. McDonell, L. M.,et al. (2013). Mutations in STAMBP, encoding a deubiquitinating enzyme, cause microcephaly-capillary malformation syndrome. Nature genetics, 45(5), 556–562.
9. Fogel, B. L., et al. (2014). Exome sequencing in the clinical diagnosis of sporadic or familial cerebellar ataxia. JAMA neurology, 71(10), 1237–1246.
10. Pyle, A., et al. (2015). Exome sequencing in undiagnosed inherited and sporadic ataxias. Brain: a journal of neurology, 138(Pt 2), 276–283.
11. Vilboux, T., et al. (2017). Molecular genetic findings and clinical correlations in 100 patients with Joubert syndrome and related disorders prospectively evaluated at a single center. Genetics in medicine: official journal of the American College of Medical Genetics, 19(8), 875–882.
12. Vardarajan, B. N., et al. (2015). Rare coding mutations identified by sequencing of Alzheimer disease genome-wide association studies loci. Ann Neurol, 78(3): 487–498.
13. Cuccaro, M. L., et al. (2016). SORL1 mutations in early- and late-onset Alzheimer disease. Neurology. Genetics, 2(6), e116.
14. Yin, R. H., Yu, J. T., & Tan, L. (2015). The Role of SORL1 in Alzheimer's Disease. Molecular neurobiology, 51(3), 909–918.
15. Lacour, M., et al. (2019). Causative Mutations and Genetic Risk Factors in Sporadic Early Onset Alzheimer's Disease Before 51 Years. Journal of Alzheimer's disease : JAD, 71(1), 227–243.
16. Oluwole, O. G., et al. (2020). Targeted next-generation sequencing identifies novel variants in candidate genes for Parkinson's disease in Black South African and Nigerian patients. BMC medical genetics, 21(1), 23.
17. Ogawa, K., et al. (2019). Next-generation sequencing identifies contribution of both class I and II HLA genes on susceptibility of multiple sclerosis in Japanese. Journal of neuroinflammation, 16(1), 162.
18. Ramagopalan, S. V., et al. (2011). Rare variants in the CYP27B1 gene are associated with multiple sclerosis. Annals of neurology, 70(6), 881–886.
19. Cotsapas, C., & Mitrovic, M. (2018). Genome-wide association studies of multiple sclerosis. Clinical & translational immunology, 7(6), e1018.
20. Fogel, B. L., et al. (2014). Exome sequencing in the clinical diagnosis of sporadic or familial cerebellar ataxia. JAMA neurology, 71(10), 1237–1246.

With over 314,707 expert-curated disease-causing mutations and more than 11,000 detailed summary reports of disease-associated/functional polymorphisms\, HGMD is the most up-to-date and comprehensive collection of known and published pathogenic gene lesions responsible for human inherited disease. Cited in over 5000 publications in leading scientific journals, it is integral to any clinical assessment of germline variants. HGMD provides valuable data for clinical interpretive and reporting use in exome screening studies, and optimizes mutational screening strategies. HGMD is a widely used and trusted resource for medical and clinical geneticists, bioinformaticians, physicians, and genetic counselors.

Ordering www.qiagen.com/shop | Technical Support digitalinsights.qiagen.com/support | Website digitalinsights.qiagen.com