HGMD® : Human Gene Mutation Database

Example Queries for use with the MySQL Database

Common SQL Queries in the HGMD® Professional MySQL database

The tables allgenes and allmut contain data for easy querying of genes and mutations respectively. The content in these tables is extracted from other tables of the database.

Get **all mutations** from table “allmut”

```
SELECT * FROM allmut;
```

Get **all genes** from table “allgenes”

```
SELECT * FROM allgenes;
```

More Specific SQL Queries in the HGMD® Professional MySQL database

Get all mutations with more than one primary reference (extra references) with a pre-selected number of output fields such as mutation accession number, PMIDs from the primary and the extra references, disease name, gene symbol, and HGVS description.

Get **all mutations** from table “allmut”

```
SELECT allmut.acc_num, extrarefs.acc_num, allmut.pmid,
extrarefs.pmid, allmut.disease, allmut.gene, allmut.hgvs
FROM allmut, extrarefs WHERE allmut.acc_num=extrarefs.acc_num;
```

Example result line:

acc_num	acc_num	pmid	pmid	disease	gene	hgvs
BM1016096	BM1016096	20696889	23533228	Kallmann syndrome	NELF	1067C>G

More Specific SQL Queries in the HGMD® Professional MySQL database

Get **selected columns** for a particular gene, for example “PSEN1”, from the table “allmut”.

```
SELECT disease, gene, acc_num, tag, hgvs, pmid, omimid FROM allmut
WHERE gene='PSEN1';
```

Example result line:

disease	gene	acc_num	tag	hgvs	pmid	omimid
Alzheimer disease	PSEN1	CM004073	DM	1061C>T	11524469	104311

More Specific SQL Queries in the HGMD® Professional MySQL database

HGMD® Professional has adopted a policy of sub-categorizing mutations and polymorphisms. Seven different categories are captured.

Get the number of entries for a certain category

- DM disease-causing mutation
- DM? likely disease-causing mutation
- DP disease-associated polymorphism
- FP in vitro or in vivo functional polymorphism
- DFP disease-associated polymorphism with additional functional evidence
- R retired record

```
SELECT gene, acc_num, hgvs, tag FROM allmut WHERE tag='DM';
```

Example result line:

gene	acc_num	hgvs	tag
A2M	CM92001	2915G>A	DM

HGMD® Professional - Views

The schema HGMD_Views contains virtual table views provided to simply common queries. The views contain many joins between different tables.

- **data_for_ngs** : HGMD data presented in a way more suited to parsing NGS data
- **isoform_list** : Genes and mutations for which more than one isoform is recorded
- **mut_to_concept** : Mapping of curated phenotypes to common ontologies
- **with_additional_references** : Genes and mutations for which there is more than one cited reference in HGMD
- **with_dbsnp_rs** : Genes and mutations with a corresponding entry in dbSNP
- **with_edit_history** : Genes and mutations that have been updated in HGMD

HGMD® Professional - Views

data_for_ngs

HGMD data presented in a way more suited to parsing NGS data

Get **all records** from table “data_for_ngs”

SELECT * FROM hgmd_views.data_for_ngs;

Example result line:

hgmd accession	gene symbol	gene strand	chr	hg19 start	hg19 end	ref	alt	refseq	hgvs	variant class	dbSNP	primary pubmed	additional pubmed
CS991274	ABCA4	-	1	94485136	94485136	T	C	NM_000350.2	c.5196+2T>C	DM	rs61751405	9054934	9973280

HGMD® Professional - Views

isoform_list

Genes and mutations for which more than one isoform is recorded

Get **all records** from table “isoform_list”

SELECT * FROM hgmd_views.isoform_list;

Example result line:

hgmd gene id	gene symbol	gene description	entrezID	hgmd accession	variant class	mutation description	refseq	hgvs	chr	hg19 start	hg19 end	primary pubmed
914	gfapie	Glial fibrillary acidic protein, isoform epsilon	111111	CM134096	DM	Arg430His	NM_001131019.2	c.1289G>A	17	42987511	42987511	23634874

HGMD® Professional - Views

mut_to_concept

Mapping of curated phenotypes to common ontologies

Get **all records** from table “mut_to_concept”

SELECT * FROM hgmd_views.mut_to_concept;

Example result lines:

hgmd accession	gene symbol	phenotype	rela	cui	str	ispref	sab	code
CM134096	gfapie	Alexander disease, adult onset	root	C0270726	Alexander Disease	Y	MSH	D038261
CM134096	gfapie	Alexander disease, adult onset	is_a	C0007682	Disorder of central nervous system, unspecified	N	ICD10AM	G96.9

HGMD® Professional - Views

with_additional_references

Genes and mutations for which there is more than one cited reference in HGMD

Get **all records** from table “with_additional_references”

SELECT * FROM hgmd_views.with_additional_references;

Example result line:

hgmd accession	primary phenotype	additional phenotype	gene symbol	mutation description	refseq	hgvs	variant class	primary pubmed	additional pubmed	additional reference type
BM1016096	Kallmann syndrome	-	NELF	Ala356Gly	NM_015537.4	c.1067C>G	DM	20696889	23533228	SAR

HGMD® Professional - Views

with_dbsnp_rs

Genes and mutations with a corresponding entry in dbSNP

Get **all records** from table “with_dbsnp_rs”

SELECT * FROM hgmd_views.with_dbsnp_rs;

Example result line:

hgmd accession	gene symbol	mutation description	refseq	hgvs	variant class	dbSNP	1000G frequency
BM1267966	DNAH5	Gln2949Glu	NM_001369.2	c.8845C>G	DM	rs147688221	0.000/1

HGMD® Professional - Views

with_edit_history

Genes and mutations that have been updated in HGMD

Get **all records** from table “with_edit_history”

SELECT * FROM hgmd_views.with_edit_history;

Example result line:

gene symbol	hgmd accession	mutation description	hgvs	variant class	column updated	before update	after update	date updated
ABCG8	CM042913	Thr400Lys	c.1199C>A	DM	tag	FP	DFP	2014-04-03

Support: ts-bioinformatics@qiagen.com

HGMD homepage: https://digitalinsights.qiagen.com/products-overview/clinical-insights-portfolio/human-gene-mutation-database/