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, PMID 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:

<table>
<thead>
<tr>
<th>acc_num</th>
<th>acc_num</th>
<th>pmid</th>
<th>pmid</th>
<th>disease</th>
<th>gene</th>
<th>hgvs</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM1016096</td>
<td>BM1016096</td>
<td>20696889</td>
<td>23533228</td>
<td>Kallmann syndrome</td>
<td>NELF</td>
<td>1067C&gt;G</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>disease</th>
<th>gene</th>
<th>acc_num</th>
<th>tag</th>
<th>hgvs</th>
<th>pmid</th>
<th>omimid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer disease</td>
<td>PSEN1</td>
<td>CM004073</td>
<td>DM</td>
<td>1061C&gt;T</td>
<td>11524469</td>
<td>104311</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>gene</th>
<th>acc_num</th>
<th>hgvs</th>
<th>tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2M</td>
<td>CM92001</td>
<td>2915G&gt;A</td>
<td>DM</td>
</tr>
</tbody>
</table>
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”

$$\text{SELECT } * \text{ FROM hgmd_views.data_for_ngs;}$$

Example result line:

<table>
<thead>
<tr>
<th>hgmd_accession</th>
<th>gene_symbol</th>
<th>gene_strand</th>
<th>chr</th>
<th>hg19_start</th>
<th>hg19_end</th>
<th>ref</th>
<th>alt</th>
<th>refseq</th>
<th>hgvs</th>
<th>variant_class</th>
<th>dbSNP</th>
<th>primary_pubmed</th>
<th>additional_pubmed</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS991274</td>
<td>ABCA4</td>
<td>-</td>
<td>1</td>
<td>94485136</td>
<td>94485136</td>
<td>T</td>
<td>C</td>
<td>NM_000350.2</td>
<td>c.5196+2T&gt;C</td>
<td>DM</td>
<td>rs61751405</td>
<td>9054934</td>
<td>9973280</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>hgmd gene id</th>
<th>gene symbol</th>
<th>gene description</th>
<th>entrezID</th>
<th>hgmd accession</th>
<th>variant class</th>
<th>mutation description</th>
<th>refseq</th>
<th>hgvs</th>
<th>chr</th>
<th>hg19 start</th>
<th>hg19 end</th>
<th>primary pubmed</th>
</tr>
</thead>
<tbody>
<tr>
<td>914</td>
<td>gfapie</td>
<td>Glial fibrillary acidic protein, isoform epsilon</td>
<td>111111</td>
<td>CM134096</td>
<td>DM</td>
<td>Arg430His</td>
<td>NM_001131019.2</td>
<td>c.1289G&gt;A</td>
<td>17</td>
<td>42987511</td>
<td>42987511</td>
<td>23634874</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>hgmd accession</th>
<th>gene symbol</th>
<th>phenotype</th>
<th>rela</th>
<th>cui</th>
<th>str</th>
<th>ispref</th>
<th>sab</th>
<th>code</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM134096</td>
<td>gfnpe</td>
<td>Alexander disease, adult onset</td>
<td>root</td>
<td>C0270726</td>
<td>Alexander Disease</td>
<td>Y</td>
<td>MSH</td>
<td>D038261</td>
</tr>
<tr>
<td>CM134096</td>
<td>gfnpe</td>
<td>Alexander disease, adult onset</td>
<td>is_a</td>
<td>C0007682</td>
<td>Disorder of central nervous system, unspecified</td>
<td>N</td>
<td>ICD10AM</td>
<td>G96.9</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>hgmd accession</th>
<th>primary phenotype</th>
<th>additional phenotype</th>
<th>gene symbol</th>
<th>mutation description</th>
<th>refseq</th>
<th>hgvs</th>
<th>variant class</th>
<th>primary pubmed</th>
<th>additional pubmed</th>
<th>additional reference type</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM1016096</td>
<td>Kallmann syndrome</td>
<td>-</td>
<td>NELF</td>
<td>Ala356Gly</td>
<td>NM_015537.4</td>
<td>c.1067C&gt;G</td>
<td>DM</td>
<td>20696889</td>
<td>23533228</td>
<td>SAR</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>hgmd accession</th>
<th>gene symbol</th>
<th>mutation description</th>
<th>refseq</th>
<th>hgvs</th>
<th>variant class</th>
<th>dbSNP</th>
<th>1000G frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM1267966</td>
<td>DNAH5</td>
<td>Gln2949Glu</td>
<td>NM_001369.2</td>
<td>c.8845C&gt;G</td>
<td>DM</td>
<td>rs147688221</td>
<td>0.000/1</td>
</tr>
</tbody>
</table>
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:

<table>
<thead>
<tr>
<th>gene symbol</th>
<th>hgd accession</th>
<th>mutation description</th>
<th>hgvgs</th>
<th>variant class</th>
<th>column updated</th>
<th>before update</th>
<th>after update</th>
<th>date updated</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCG8</td>
<td>CM042913</td>
<td>Thr400Lys</td>
<td>c.1199C&gt;A</td>
<td>DM</td>
<td>tag</td>
<td>FP</td>
<td>DFP</td>
<td>2014-04-03</td>
</tr>
</tbody>
</table>
info@biobase-international.com
www.biobase-international.com