

Searched OMIM.org for "huntington's disease". Recorded the five identifiers (six-digit numbers) of those five records. The corresponding biomaRt filter name for these identifiers is "mim_morbid_accession". Used biomaRt to retrieve two tables with the following attributes, limiting to the five MIM values I found:

Top 5 MIM values found: 603218, 604802, 143100, 606438, 607136.

First table (2 points)

Entrez Gene ID

HGNC symbol

Ensembl Gene ID

First Table:

	entrezgene_id	hgnc_symbol	ensembl_gene_id
1	3064	HTT	ENSG00000197386
2	5621	PRNP	ENSG00000171867
3	57338	JPH3	ENSG00000154118
4	6908	TBP	ENSG00000112592

Second table (2 points)

HGNC symbol

Ensembl Gene ID

Ensembl Transcript ID

Second Table:

	hgnc_symbol	ensembl_gene_id	ensembl_transcript_id
1	HTT	ENSG00000197386	ENST00000680239
2	HTT	ENSG00000197386	ENST00000680956
3	HTT	ENSG00000197386	ENST00000680360
4	HTT	ENSG00000197386	ENST00000681528
5	HTT	ENSG00000197386	ENST00000647962
6	HTT	ENSG00000197386	ENST00000649900
7	HTT	ENSG00000197386	ENST00000680291
8	HTT	ENSG00000197386	ENST00000355072
9	HTT	ENSG00000197386	ENST00000648150
10	HTT	ENSG00000197386	ENST00000506137
11	HTT	ENSG00000197386	ENST00000512909
12	HTT	ENSG00000197386	ENST00000510626
13	HTT	ENSG00000197386	ENST00000649131
14	HTT	ENSG00000197386	ENST00000509618
15	HTT	ENSG00000197386	ENST00000650588
16	HTT	ENSG00000197386	ENST00000650595
17	HTT	ENSG00000197386	ENST00000513639
18	HTT	ENSG00000197386	ENST00000513326
19	HTT	ENSG00000197386	ENST00000509043
20	HTT	ENSG00000197386	ENST00000509751
21	HTT	ENSG00000197386	ENST00000512068
22	HTT	ENSG00000197386	ENST00000513806
23	HTT	ENSG00000197386	ENST00000508321
24	PRNP	ENSG00000171867	ENST00000430350
25	PRNP	ENSG00000171867	ENST00000379440
26	PRNP	ENSG00000171867	ENST00000424424
27	PRNP	ENSG00000171867	ENST00000457586
28	JPH3	ENSG00000154118	ENST00000537256
29	JPH3	ENSG00000154118	ENST00000301008
30	JPH3	ENSG00000154118	ENST00000284262
31	JPH3	ENSG00000154118	ENST00000563609
32	TBP	ENSG00000112592	ENST00000421512
33	TBP	ENSG00000112592	ENST00000230354
34	TBP	ENSG00000112592	ENST00000423353
35	TBP	ENSG00000112592	ENST00000636632
36	TBP	ENSG00000112592	ENST00000446829
37	TBP	ENSG00000112592	ENST00000392092
38	TBP	ENSG00000112592	ENST00000540980
39	TBP	ENSG00000112592	ENST00000616883

R Code (*also submitted separately as a file):

```
> if (!requireNamespace("BiocManager", quietly = TRUE))
+   install.packages("BiocManager")
> BiocManager::install("biomaRt")

> library(biomaRt)
> listMarts()

> ensembl = useMart("ensembl", dataset="hsapiens_gene_ensembl")
```

#first table

```
>getBM(attributes=c("entrezgene_id","hgnc_symbol","ensembl_gene_id"),  
filters="mim_morbid_accession", values=c("603218", "604802", "143100", "606438", "607136"),  
mart=ensembl)
```

#second table

```
>getBM(attributes=c("hgnc_symbol","ensembl_gene_id", "ensembl_transcript_id"),  
filters="mim_morbid_accession", values=c("603218", "604802", "143100", "606438", "607136"),  
mart=ensembl)
```

Last two points, different number of rows:

The second table contains more rows because the third column, “ensembl_transcript_id”, contains multiple entries for the same gene. The multiple entries for that column reflect the multiple transcripts available for the specific genes, as there is more than just one transcript recorded for each of those genes.