

# **Gene ID Conversion**

## **Using R Bioconductor and NCBI gene\_info**

**CFDE: Gene Working Group**  
**Biomedical Data Commons Workbench (BDCW)**  
**Metabolomics Workbench DCC**

**Mano Maurya and Shankar Subramaniam**  
**2022-08-22**

**Funding: NIH Common Funds (Metabolomics Workbench and Common Fund  
Data Ecosystem (CFDE))**

# Motivation

- Genomic and gene expression data is integral to biomolecular data analysis.
- The types of identifiers used for genes may differ across different resources (such as GEO, ENCODE, GTEx) providing such data sets.
- The ability to use a single type of gene identifier is crucial for integrating data from two or more resources. Most downstream analysis tools work with (or internally convert IDs to) ENTREZ ID.
- This gene ID conversion tool will facilitate the use of a common gene identifier.
- NGS/RNASeq mappers generally provide read-counts (or equivalent quantities) at the **gene-level** or transcript-level.
- The most common gene IDs are gene **SYMBOL** (e.g., KLF4) AND **Ensembl GENE ID** (e.g., ENSG00000136826.*n*).
- The most common transcript IDs are NCBI Refseq ID (e.g., NM\_001314052.*n*) or Ensembl Transcript ID (ENST00000420475.*n*).
- SYMBOLs are more human readable and change for yet not well studied genes. If a new SYMBOL is assigned to a gene, its current SYMBOL generally gets included in its ALIAS.

# User Interface

Web-based interface: A user can specify the name of the organism, enter the list of genes or upload a file containing the gene list and select the ID type for the gene list provided (e.g., SYMBOL or ENTREZ ID). Then, the user can select the types of the converted IDs to be included in the output.

**Species (currently human, mouse and rat)**

**List of genes (Symbol, Ensembl gene ID, etc.; only one ID type in a list)**

as a typed list (if just a few genes; white-space or comma separated)

or as a file with one column or row (white-space, comma separated)

or as a file with 2 columns (if additional information is available)

2 columns: GENEID/SYMBOL CHR

The type of gene identifier provided (SYMBOL, ENTREZ GENE ID, ALIAS, ENSEMBL GENE ID, REFSEQ TRANSCRIPT ID).

# Challenges

- Most current focus on SYMBOL to ENTREZ ID conversion.
- Approach
  - ❑ R packages AnnotationDbi and org.Xy.eg.db are used to convert the IDs.
  - ❑ If Gene ID type is “SYMBOL or ALIAS” (most suitable choice), then it is searched in SYMBOL first and if not found then it is searched in ALIAS and the results are combined. *The SYMBOLs not found in ALIAS too are searched (case-insensitive) in Homologene database to see if it exists in another organism.*
- **When searching in ALIAS, a SYMBOL may appear in ALIAS of several genes.** For example, AIM1 appears in ALIAS of CRYBG1, AURKB and SLC45A2. Which one is correct? In this case, additional information such as chromosome number will be useful to resolve it to one (hopefully) or two genes.
- **For a very small set of genes, a SYMBOL may give two different ENTREZ IDs.** For example, searching for MEMO1 on NCBI gene website (<https://www.ncbi.nlm.nih.gov/gene/?term=human+MEMO1>) lists:

Name/Gene ID	Description	Location	Aliases	MIM
<input type="checkbox"/> <a href="#">MEMO1</a> ID: 51072	mediator of cell motility 1 [ <i>Homo sapiens</i> (human)]	Chromosome 2, NC_000002.12 (31867188..32011008, complement)	C2orf4, CGI-27, MEMO, NS5ATP7	611786
<input type="checkbox"/> <a href="#">MEMO1</a> ID: 7795	Methylation modifier for class I HLA [ <i>Homo</i> <i>sapiens</i> (human)]			601201

Use of chromosome number will resolve this.

# Gene ID Conversion

Available as a web-based tool: <https://bdcw.org/geneid/geneidconv.php>

Also as an SmartAPI: <https://smart-api.info/ui/e712b9eb07e637a00ae468f757ce2a1f>

## For integration with other tools and scripts:

Uses php \_GET so that gene symbol can be embedded in the URL itself:

Example URL: targeted for programmatic/API access:

URL to use for json output with CLI (e.g., using [curl -L 'URL']; use &View=txt for text output):

[https://bdcw.org/geneid/geneid\\_proc\\_selcol\\_GET.php?species=hsa&GeneListStr=ITPR3\\_IL6\\_KLF4&GeneIDType=SYMBOL\\_OR\\_ALIAS&View=json](https://bdcw.org/geneid/geneid_proc_selcol_GET.php?species=hsa&GeneListStr=ITPR3_IL6_KLF4&GeneIDType=SYMBOL_OR_ALIAS&View=json)

REST format URL:

[https://bdcw.org/geneid/rest/species/hsa/GeneIDType/SYMBOL\\_OR\\_ALIAS/GeneListStr/ITPR3\\_IL6\\_KLF4/View/json](https://bdcw.org/geneid/rest/species/hsa/GeneIDType/SYMBOL_OR_ALIAS/GeneListStr/ITPR3_IL6_KLF4/View/json)



General purpose page: <https://bdcw.org/geneid/geneidconv.php>

- Single and multiple genes
- As a list or via file upload

→ ↻ 🏠 <https://bdcw.org/geneid/geneidconv.php> 90% ☆

Import bookmarks... ⚙ Most Visited 🌐 Getting Started

## Gene ID Conversion Tool

Genomic and gene expression data is integral to biomolecular data analysis. The types of identifiers used for genes differ across different resources providing such data sets. The ability to use a single type of gene identifier is imperative for integrating data from two or more resources. This gene ID conversion tool facilitates the use of a common gene identifier.

**Select Database:**  Selecting "Recent NCBI gene\_info" option (which uses precomputed tables integrating recent NCBI gene\_info and R Bioconductor package org\_Xy\_eg\_db) results in faster results. Coverage may not be complete with either database.

**Select species:**

**List of gene IDs separated by white space or comma:**

**or Upload a file of gene list:**  No file selected.  File contains header

If the list contains more than 10,000 genes, connection may time-out depending upon the server load. In such a case, consider submitting more than one list in separate sessions, and combine the results outside this tool.

**Gene ID type:**

GeneIDType: [SYMBOL\_OR\_ALIAS]

Please select the columns for the output: [Know more about the columns.](#)

SYMBOL,  ALIAS,  ENTREZID,  GENENAME,  ENSEMBL,  CHROMOSOME,  MIM,  HGNC  
 REFSEQ,  UNIPROT

URL to use for json output with CLI (e.g., using [curl -L 'URL']; use &View=txt for text output):  
[https://bdcw.org/geneid/geneid\\_proc\\_selcol\\_GET.php?species=hsa&GeneListStr=PNPLA3\\_AIM1&GeneIDType=SYMBOL\\_OR\\_ALIAS&View=json](https://bdcw.org/geneid/geneid_proc_selcol_GET.php?species=hsa&GeneListStr=PNPLA3_AIM1&GeneIDType=SYMBOL_OR_ALIAS&View=json)

**General purpose page:** <https://bdcw.org/geneid/geneidconv.php>

The first five rows of gene ID conversion results are given below. To further select, click on HTML.  
[Download txt](#), [HTML](#), [JSON](#).

	SYMBOL_OR_ALIAS	MATCH_SOURCE	ALIAS	SYMBOL	ENTREZID	ENSEMBL	GENENAME	REFSEQ	UNIPROT	KEGG	MARRVEL	LandingPages
113	PNPLA3	SYMBOL	ADPN C22orf20 iPLA(2)epsilon PNPLA3	<a href="#">PNPLA3</a>	<a href="#">80339</a>	<a href="#">ENSG00000100344</a>	patatin like phospholipase domain containing 3	<a href="#">NM_025225</a>	<a href="#">Q9NST1</a>	<a href="#">80339</a>	<a href="#">80339</a>	<a href="#">PNPLA3</a>
114	PNPLA3	SYMBOL	ADPN C22orf20 iPLA(2)epsilon PNPLA3	<a href="#">PNPLA3</a>	<a href="#">80339</a>	<a href="#">ENSG00000100344</a>	patatin like phospholipase domain containing 3	<a href="#">NP_079501</a>	<a href="#">Q9NST1</a>	<a href="#">80339</a>	<a href="#">80339</a>	<a href="#">PNPLA3</a>
1	AIM1	ALIAS	AIM1	<a href="#">CRYBG1</a>	<a href="#">202</a>	<a href="#">ENSG00000112297</a>	crystallin beta-gamma domain containing 1	<a href="#">NM_001371242</a>	<a href="#">Q9Y4K1</a>	<a href="#">202</a>	<a href="#">202</a>	<a href="#">CRYBG1</a>

**Disambiguation:** AIM1 is alias for CRYBG1, AURKB and SLC45A2

	Select	SYMBOL_OR_ALIAS	MATCH_SOURCE	ALIAS	SYMBOL	ENTREZID	ENSEMBL	GENENAME	REFSEQ	UNIPROT	KEGG	MARRVEL	LandingPages
113	<input type="checkbox"/>	PNPLA3	SYMBOL	ADPN C22orf20 iPLA(2)epsilon PNPLA3	<a href="#">PNPLA3</a>	<a href="#">80339</a>	<a href="#">ENSG00000100344</a>	patatin like phospholipase domain containing 3	<a href="#">NM_025225</a>	<a href="#">Q9NST1</a>	<a href="#">80339</a>	<a href="#">80339</a>	<a href="#">PNPLA3</a>
114	<input type="checkbox"/>	PNPLA3	SYMBOL	ADPN C22orf20 iPLA(2)epsilon PNPLA3	<a href="#">PNPLA3</a>	<a href="#">80339</a>	<a href="#">ENSG00000100344</a>	patatin like phospholipase domain containing 3	<a href="#">NP_079501</a>	<a href="#">Q9NST1</a>	<a href="#">80339</a>	<a href="#">80339</a>	<a href="#">PNPLA3</a>
1	<input type="checkbox"/>	AIM1	ALIAS	AIM1	<a href="#">CRYBG1</a>	<a href="#">202</a>	<a href="#">ENSG00000112297</a>	crystallin beta-gamma domain containing 1	<a href="#">NM_001371242</a>	<a href="#">Q9Y4K1</a>	<a href="#">202</a>	<a href="#">202</a>	<a href="#">CRYBG1</a>
459	<input type="checkbox"/>	AIM1	ALIAS	AIM1	<a href="#">AURKB</a>	<a href="#">9212</a>	<a href="#">ENSG00000178999</a>	aurora kinase B	<a href="#">XP_016880799</a>	<a href="#">Q96GD4</a>	<a href="#">9212</a>	<a href="#">9212</a>	<a href="#">AURKB</a>
473	<input type="checkbox"/>	AIM1	ALIAS	AIM1	<a href="#">AURKB</a>	<a href="#">9212</a>	<a href="#">ENSG00000178999</a>	aurora kinase B	<a href="#">XP_016880800</a>	<a href="#">Q96GD4</a>	<a href="#">9212</a>	<a href="#">9212</a>	<a href="#">AURKB</a>
486	<input type="checkbox"/>	AIM1	ALIAS	AIM1	<a href="#">SLC45A2</a>	<a href="#">51151</a>	<a href="#">ENSG00000164175</a>	solute carrier family 45 member 2	<a href="#">NM_001012509</a>	<a href="#">Q9UMX9</a>	<a href="#">51151</a>	<a href="#">51151</a>	<a href="#">SLC45A2</a>
492	<input type="checkbox"/>	AIM1	ALIAS	AIM1	<a href="#">SLC45A2</a>	<a href="#">51151</a>	<a href="#">ENSG00000164175</a>	solute carrier family 45 member 2	<a href="#">NM_001297417</a>	<a href="#">Q9UMX9</a>	<a href="#">51151</a>	<a href="#">51151</a>	<a href="#">SLC45A2</a>

# Gene ID Conversion Tool

Genomic and gene expression data is integral to biomolecular data analysis. The types of identifiers used for genes differ across different resources providing such data sets. The ability to use a single type of gene identifier is imperative for integrating data from two or more resources. This gene ID conversion tool facilitates the use of a common gene identifier.

**Select Database:** Recent NCBI gene\_info Selecting "Recent NCBI gene\_info" option (which uses precomputed tables integrating recent NCBI gene\_info and R Bioconductor package org\_Xy\_eg\_db) results in faster results. Coverage may not be complete with either database.

**Select species:** Human (Homo sapiens)

List of gene IDs separated by white space or comma:

or Upload a file of gene list: Browse... No file selected.  File contains header

If the list contains more than 10,000 genes, connection may time-out depending upon the server load. In such a case, consider submitting more than one list in separate sessions, and combine the results outside this tool.

**Gene ID type:** ENSEMBL

Upload gene list for gene ID conversion

The file genelist\_hg38\_GenCode\_V28\_ENSG\_subset.txt has been uploaded as uploads/genelist\_hg38\_GenCode\_V28\_ENSG\_subset.txt. GeneIDType: [ENSEMBL]

If a file is uploaded and chromosomal coordinates are provided (e.g., SYMBOL CHR START END), specify genome version. If the file has only two columns (SYMBOL CHR), then there is no need to specify genome version.

**Select closest genome version:** Select...

Please select the columns for the output: [Know more about the columns.](#)  
 SYMBOL,  ALIAS,  ENTREZID,  GENENAME,  ENSEMBL,  CHROMOSOME,  MIM,  HGNC  
 REFSEQ,  UNIPROT

Submit

## Ability to further select a gene subset.

The first five rows of gene ID conversion results are given below. To further select, click on HTML, Download txt, HTML, JSON.

	ENSEMBL.VER	ENSEMBL	SYMBOL	ENTREZID	LandingPages
1	ENSG00000131781.12	ENSG00000131781	FMO5	2330	FMO5
49	ENSG00000136244.11	ENSG00000136244	IL6	3569	IL6
99	ENSG00000131778.18	ENSG00000131778	CHD1L	9557	CHD1L
294	ENSG00000143167.11	ENSG00000143167	GPA33	10223	GPA33
298	ENSG00000164794.8	ENSG00000164794	KCNV1	27012	KCNV1

Total time for gene ID conversion: : 1.497 sec elapsed

	Select	ENSEMBL.VER	ENSEMBL	SYMBOL	ENTREZID
1	<input checked="" type="checkbox"/>	ENSG00000131781.12	ENSG00000131781	FMO5	2330
49	<input checked="" type="checkbox"/>	ENSG00000136244.11	ENSG00000136244	IL6	3569
99	<input type="checkbox"/>	ENSG00000131778.18	ENSG00000131778	CHD1L	9557
294	<input type="checkbox"/>	ENSG00000143167.11	ENSG00000143167	GPA33	10223
298	<input checked="" type="checkbox"/>	ENSG00000164794.8	ENSG00000164794	KCNV1	27012
300	<input type="checkbox"/>	ENSG00000186092.6	ENSG00000186092	OR4F5	79501
302	<input checked="" type="checkbox"/>	ENSG00000268020.3	ENSG00000268020	OR4G4P	79504
303	<input type="checkbox"/>	ENSG00000223972.5	ENSG00000223972	DDX11L2	84771
305	<input type="checkbox"/>	ENSG00000143194.12	ENSG00000143194	MAEL	84944
353	<input type="checkbox"/>	ENSG00000198842.9	ENSG00000198842	STYXL2	92235
357	<input type="checkbox"/>	ENSG00000173950.15	ENSG00000173950	XXYL1	152002
383	<input type="checkbox"/>	ENSG00000284742.1	ENSG00000284742	OR11P1P	282795
384	<input type="checkbox"/>	ENSG00000143195.12	ENSG00000143195	ILDR2	387597
402	<input type="checkbox"/>	ENSG00000196832.4	ENSG00000196832	OR11G2	390439
406	<input type="checkbox"/>	ENSG00000213068.3	ENSG00000213068	RPS17P6	391130
407	<input type="checkbox"/>	ENSG00000226653.3	ENSG00000226653	OR13Z1P	403228
408	<input type="checkbox"/>	ENSG00000179428.2	ENSG00000179428	IL6-AS1	541472
409	<input type="checkbox"/>	ENSG00000226015.2	ENSG00000226015	CCT8P1	644131
410	<input type="checkbox"/>	ENSG00000237613.2	ENSG00000237613	FAM138A	645520
411	<input type="checkbox"/>	ENSG00000184388.5	ENSG00000184388	PABPC1L2B	645974
413	<input type="checkbox"/>	ENSG00000227232.5	ENSG00000227232	WASH7P	653635
414	<input type="checkbox"/>	ENSG00000223972.5	ENSG00000223972	DDX11L16	727856
416	<input type="checkbox"/>	ENSG00000275520.1	ENSG00000275520	FAM236A	100129407



# For API-based access to integrate in user's existing tools:

**Important:** For some genes, multiple IDs may be listed. The users needs to select one or more for further use in a script or on the web.

URLs to use for json output with CLI (e.g., using [curl -L 'URL']; use &View=txt for text output):

[https://bdcw.org/geneid/geneid\\_proc\\_selcol\\_GET.php?species=hsa&GeneListStr=AIM1&GeneIDType=SYMBOL\\_OR\\_ALIAS&View=txt](https://bdcw.org/geneid/geneid_proc_selcol_GET.php?species=hsa&GeneListStr=AIM1&GeneIDType=SYMBOL_OR_ALIAS&View=txt)

[https://bdcw.org/geneid/geneid\\_proc\\_selcol\\_GET.php?species=hsa&GeneListStr=ITPR3\\_IL6\\_KLF4&GeneIDType=SYMBOL\\_OR\\_ALIAS&View=json](https://bdcw.org/geneid/geneid_proc_selcol_GET.php?species=hsa&GeneListStr=ITPR3_IL6_KLF4&GeneIDType=SYMBOL_OR_ALIAS&View=json)

REST format URL:

[https://bdcw.org/geneid/rest/species/hsa/GeneIDType/SYMBOL\\_OR\\_ALIAS/GeneListStr/AIM1/View/json](https://bdcw.org/geneid/rest/species/hsa/GeneIDType/SYMBOL_OR_ALIAS/GeneListStr/AIM1/View/json)

[https://bdcw.org/geneid/rest/species/hsa/GeneIDType/SYMBOL\\_OR\\_ALIAS/GeneListStr/ITPR3\\_IL6\\_KLF4/View/json](https://bdcw.org/geneid/rest/species/hsa/GeneIDType/SYMBOL_OR_ALIAS/GeneListStr/ITPR3_IL6_KLF4/View/json)

# Single gene case: [https://bdcw.org/geneid/geneid\\_proc\\_selcol\\_GET.php?species=hsa&GeneListStr=AIM1&GeneIDType=SYMBOL\\_OR\\_ALIAS&View=txt](https://bdcw.org/geneid/geneid_proc_selcol_GET.php?species=hsa&GeneListStr=AIM1&GeneIDType=SYMBOL_OR_ALIAS&View=txt)

[https://bdcw.org/geneid/geneid\\_proc\\_selcol\\_GET.php?species=hsa&GeneListStr=AIM1&GeneIDType=SYMBOL\\_OR\\_ALIAS&View=txt](https://bdcw.org/geneid/geneid_proc_selcol_GET.php?species=hsa&GeneListStr=AIM1&GeneIDType=SYMBOL_OR_ALIAS&View=txt)

Import bookmarks... 
 Most Visited 
 Getting Started

SYMBOL_OR_ALIAS	MATCH_SOURCE	ALIAS	SYMBOL	ENTREZID	ENSEMBL	GENENAME	REFSEQ	UNIPROT	KEGG	MARRVEL		
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			NM_001371242	Q9Y4K1	202	202
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			NM_001624	Q9Y4K1	202	202
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			NP_001358171	Q9Y4K1	202	202
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			NP_001615	Q9Y4K1	202	202
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			XM_017010333	Q9Y4K1	202	202
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			XP_016865822	Q9Y4K1	202	202
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			NM_001371242	B3KPT0	202	202
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			NM_001624	B3KPT0	202	202
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			NP_001358171	B3KPT0	202	202
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			NP_001615	B3KPT0	202	202
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			XM_017010333	B3KPT0	202	202
AIM1	ALIAS	AIM1	CRYBG1	202	ENSG00000112297	crystallin beta-gamma domain containing 1			XP_016865822	B3KPT0	202	202
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NM_001256834	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NM_001284526	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NM_001313950	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NM_001313951	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NM_001313952	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NM_001313953	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NM_001313954	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NM_001313955	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NM_004217	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NP_001243763	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NP_001271455	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NP_001300879	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NP_001300880	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NP_001300881	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NP_001300882	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NP_001300883	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NP_001300884	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NP_004208	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NR_132730	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	NR_132731	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XM_011524072	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XM_017025307	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XM_017025308	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XM_017025309	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XM_017025310	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XM_017025311	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XP_011522374	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XP_016880796	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XP_016880797	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XP_016880798	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XP_016880799	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	AURKB	9212	ENSG00000178999	aurora kinase B	XP_016880800	Q96GD4	9212	9212		
AIM1	ALIAS	AIM1	SLC45A2	51151	ENSG00000164175	solute carrier family 45 member 2			NM_001012509	Q9UMX9	51151	51151
AIM1	ALIAS	AIM1	SLC45A2	51151	ENSG00000164175	solute carrier family 45 member 2			NM_001297417	Q9UMX9	51151	51151
AIM1	ALIAS	AIM1	SLC45A2	51151	ENSG00000164175	solute carrier family 45 member 2			NM_016180	Q9UMX9	51151	51151
AIM1	ALIAS	AIM1	SLC45A2	51151	ENSG00000164175	solute carrier family 45 member 2			NP_001012527	Q9UMX9	51151	51151

## Multiple genes case: [https://bcdw.org/geneid/rest/species/hsa/GeneIDType/SYMBOL\\_OR\\_ALIAS/GeneListStr/ITPR3\\_IL6\\_KLF4/View/json](https://bcdw.org/geneid/rest/species/hsa/GeneIDType/SYMBOL_OR_ALIAS/GeneListStr/ITPR3_IL6_KLF4/View/json)

← → ↻ 🏠 [https://bcdw.org/geneid/rest/species/hsa/GeneIDType/SYMBOL\\_OR\\_ALIAS/GeneListStr/ITPR3\\_IL6\\_KLF4/View/json](https://bcdw.org/geneid/rest/species/hsa/GeneIDType/SYMBOL_OR_ALIAS/GeneListStr/ITPR3_IL6_KLF4/View/json)

🔖 Import bookmarks... ⚙️ Most Visited 🌐 Getting Started

JSON Raw Data Headers

Save Copy Pretty Print

```
[
  {
    "SYMBOL_OR_ALIAS": "IL6",
    "MATCH_SOURCE": "SYMBOL",
    "ALIAS": "BSF-2|BSF2|CDF|HGF|HSF|IFN-beta-2|IFNB2|IL-6|IL6",
    "ENSEMBL": "ENSG00000136244",
    "ENTREZID": "3569",
    "GENENAME": "interleukin 6",
    "REFSEQ": "NM_000600",
    "SYMBOL": "IL6",
    "UNIPROT": "B4DVM1",
    "KEGG": "3569",
    "MARRVEL": "3569"
  },
  {
    "SYMBOL_OR_ALIAS": "IL6",
    "MATCH_SOURCE": "SYMBOL",
    "ALIAS": "BSF-2|BSF2|CDF|HGF|HSF|IFN-beta-2|IFNB2|IL-6|IL6",
    "ENSEMBL": "ENSG00000136244",
    "ENTREZID": "3569",
    "GENENAME": "interleukin 6",
    "REFSEQ": "NM_001318095",
    "SYMBOL": "IL6",
    "UNIPROT": "B4DVM1",
    "KEGG": "3569",
    "MARRVEL": "3569"
  },
  {
    "SYMBOL_OR_ALIAS": "IL6",
    "MATCH_SOURCE": "SYMBOL",
    "ALIAS": "BSF-2|BSF2|CDF|HGF|HSF|IFN-beta-2|IFNB2|IL-6|IL6",
    "ENSEMBL": "ENSG00000136244",
    "ENTREZID": "3569",
    "GENENAME": "interleukin 6",
    "REFSEQ": "NM_001371096",
    "SYMBOL": "IL6",
    "UNIPROT": "B4DVM1",
    "KEGG": "3569",
    "MARRVEL": "3569"
  },
  {
    "SYMBOL_OR_ALIAS": "IL6",
    "MATCH_SOURCE": "SYMBOL",
    "ALIAS": "BSF-2|BSF2|CDF|HGF|HSF|IFN-beta-2|IFNB2|IL-6|IL6",
    "ENSEMBL": "ENSG00000136244",
    "ENTREZID": "3569",
    "GENENAME": "interleukin 6",
    "REFSEQ": "NP_000591",
    "SYMBOL": "IL6",
    "UNIPROT": "B4DVM1",
    "KEGG": "3569",
    "MARRVEL": "3569"
  }
]
```

## How to use it in a python program:

A python script provides an example of how to use the gene ID conversion program. At the core, a URL-based query is constructed and executed using python packages “requests” and “bs4” (function “BeautifulSoup”). After some processing, the results are available as a pandas dataframe.

[https://bdcw.org/geneid/fetch\\_php\\_page.py](https://bdcw.org/geneid/fetch_php_page.py)