

## Genew

---

Human Gene Nomenclature Database

Maja Jakóbič

## What is Genew

---

- Genew is the only resource that provides data for all human genes which have approved symbols.
- It's managed by HUGO Gene Nomenclature Committee (HGNC)
- First published in 1979
- An estimated total of 26,000-40,000 genes

## Implementation

---

- Microsoft Access 97 relational DBMS
- 13 tables, over 170 fields
- search engine based on a Perl front-end querying a PostgreSQL database, data in text files exported from off-line database  
<http://www.gene.ucl.ac.uk/cgi-bin/nomenclature/searchgenes.pl>

## Data flow

---

- **Source of data**
  - Data submissions (authors, journals, literature review, chromosome sequencing projects, gene families, homologs from other species)
  - Other database files
    - LocusLink
    - SWISS-PROT
    - The Genome Database - GDB
    - Mouse Genome Database - MGD (not exported)

## Data flow

---

### ○ Output

- Online Genew Search Engine
  - Search: gene symbol, gene name, aliases
  - Retrieved: gene symbol, gene name, aliases, cytogenetic location, OMIM (Online Mendelian Inheritance in Man) number, PMID (PubMed ID)
- Online text files

## Other information

---

- LBLAST
- Confidentiality
- LocusLink updates
- Sequence analysis
- <http://www.gene.ucl.ac.uk/nomenclature/guidelines.html>

## Rules

---

- Each gene record
  - Has a unique HGNC ID
  - Has a status...

## Rules

---

- Status
  - Pending
  - Approved
  - Reserved
  - Symbol Withdrawn
  - Entry Withdrawn
  - Approved Non-Human

## Rules

- Each gene symbol
  - Must be unique
  - Is a short-form representation (or abbreviation) of the descriptive gene name
  - should only contain Latin letters and Arabic numerals
  - should not contain punctuation
  - should not end in "G" for gene
  - not contain any reference to species, for example "H/h" for human

## Let's see...

The screenshot shows the Geneweb Human Gene Nomenclature Database Search Engine. It features a 'Simple Text Search' section with a search box and a 'Search' button. Below it, there are 'Display Options' for showing 50 records in HTML format, sorted by Symbol. An 'Advanced Search' section includes dropdown menus for 'Approved Symbols', 'Approved Gene Names', 'All Records', and 'Chromosomes', each with 'Start', 'End', and 'begin with' options. There are also 'AND' operators and a 'Submit search' button.

The screenshot shows the 'Search Hints' page. It provides instructions on how to use multiple search terms in the advanced search, such as separating values with commas. It also mentions that gene names are written using American Spelling and that data is available as a tab-delimited flat file. A 'Useful Links' section includes links to the 'Gene symbol submission form' and 'Guidelines for Human Gene Nomenclature'. The page is maintained by nome@qmulon.ucl.ac.uk and was last updated in Sept 2002.

The screenshot shows the search results for the query 'MYC'. It indicates that 31 records were returned. A table lists the results, including the Approved Gene Symbol, Approved Gene Name, Location, Sequence Accession IDs, Previous Symbols, and Aliases.

Approved Gene Symbol	Approved Gene Name	Location	Sequence Accession IDs	Previous Symbols	Aliases
<a href="#">ATF50</a>	ATP synthase, H <sup>+</sup> -transporting, mitochondrial F1 complex, O subunit (oligonucleotide sensitivity conferring protein)	21q22.1-q22.2	<a href="#">NM_001697</a>		<a href="#">OSCF</a> , ATPO
<a href="#">BLMH</a>	bilecanin hydrolase	17q11.2	<a href="#">X92106</a> <a href="#">NM_000386</a>		BH
<a href="#">C2orf15</a>	chromosome 3 open reading frame 15	3q12-q13.3	<a href="#">AB063296</a> <a href="#">NM_033364</a>		AAT1, AAT1alpha, MYCBP-binding protein
<a href="#">DDX18</a>	DEAD (Asp-Glu-Ala-Asp) box polypeptide 18	2q21.2	<a href="#">X98743</a> <a href="#">NM_006773</a>		MDb
<a href="#">FRA5E</a>	fragile site, distal on chromosome 5, rare, fra(5)(q24.1)	8q24.1			
<a href="#">FRA11F</a>	fragile site, distal on chromosome 11, rare, fra(11)(p15.1)	11p15.1			
<a href="#">FRA16B</a>	fragile site, distal on chromosome 16, rare, fra(16)(q22.1)	16q22.1			
<a href="#">FRA16E</a>	fragile site, distal on chromosome 16, rare, fra(16)(p12.1)	16p12.1			
<a href="#">FRA17A</a>	fragile site, distal on chromosome 17, rare, fra(17)(p12)	17p12			

Address: [http://www.gene.ud.ac.uk/cgi-bin/homendata/get\\_data.pl?hgnc\\_id=24010](http://www.gene.ud.ac.uk/cgi-bin/homendata/get_data.pl?hgnc_id=24010)

## Symbol Report: **C3orf15**

[Back To Search](#)

---

Approved Gene Symbol: C3orf15  
 Approved Gene Name: chromosome 3 open reading frame 15  
 HGNC ID: 24010  
 Location: 3q12-q13.3  
 Sequence Accession ID: [AB063296](#)  
 PubMed: [12923483](#), [14551891](#)  
 Status: Approved

---

**Synonyms**

Aliases: AAT1, AAT1alpha, MYCBP-binding protein  
 Previously Approved Symbols:  
 Previous Gene Names:

---

**Other Database Links**

- Ensembl: [C3orf15](#)
- Enzyme ID:
- GENATLAS: [C3orf15](#)
- GeneCards: [C3orf15](#)
- GeneClinics Gene Tests: [C3orf15](#)
- DMGT:
- LocusLink ID: [89876](#)
- MGI:
- OMIM:
- RefSeq ID: [NM\\_033364](#)
- Swiss-Prot ID:
- UCCS: [NM\\_033364](#)

---

Maintained by [nome@zepton.ac.uk](mailto:nome@zepton.ac.uk) Scripts last updated: July 2002 [HGNC Homepage](#)

Done Internet

NCBI Sequence Viewer - Microsoft Internet Explorer

Address: <http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?val=AB063296>

NCBI Nucleotide

Search: Nucleotide For:  Go Clear

Display: default Show: 20 Send to: File Get Subsequence Clipboard Details

1: [AB063296](#) Homo sapiens aat-1 [gi:14488276]

---

LOCUS AB063296 1365 bp mRNA linear PRI 20-NOV-2002  
 DEFINITION Homo sapiens aat-1 mRNA for AAT-1 alpha, complete cds.  
 ACCESSION AB063296  
 VERSION AB063296.1 GI:14488276  
 KEYWORDS .  
 SOURCE Homo sapiens (human)  
 ORGANISM [Homo sapiens](#)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 1  
 AUTHORS Yukitake,H., Furusawa,N., Taira,T., Iguchi-Arigo,S.M. and Arigo,H.  
 TITLE AAT-1, a novel testis-specific ANY-1-binding protein, forms a  
 quaternary complex with ANY-1, A-kinase anchor protein 94, and a  
 regulatory subunit of cAMP-dependent protein kinase and is  
 phosphorylated by its kinase  
 JOURNAL J. Biol. Chem. 277 (97), 45480-45492 (2002)  
 MEDLINE [12323318](#)  
 PUBMED [12323483](#)

REFERENCE  
 2 (bases 1 to 1365)  
 AUTHORS Arigo,H. and Yukitake,H.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-NOV-2001) Hiroyoshi Arigo, Graduate School of  
 Pharmaceutical Sciences, Hokkaido University, Molecular Biology/  
 Kita 17, Nishi 6, Kita-ku, Sapporo 060-0812, Japan

Done Internet