



HGNC Newsletter Winter 2014-2015

[<Previous Issue](#)

There are currently 39,400 approved symbols

[Next Issue>](#)



In this newsletter

[Welcome to Susan](#)

[New improved Gene Families \(beta release\)](#)

[Coming soon: an improved Search facility](#)

[Gene Symbols in the News](#)

[Meeting News](#)

[Publications](#)

Welcome to Susan

We would like to introduce a new member of our team, Susan Tweedie, who joined the project as a nomenclature advisor in December. Susan studied Genetics at the University of Glasgow and completed a molecular biology PhD at University College London. Following postdoctoral research on a variety of topics including parasitology and eukaryotic DNA methylation, she left the lab and completed an MRes in Bioinformatics at the University of York. Most recently she has worked as a curator at the Drosophila model organism database FlyBase where she was responsible for annotating gene function using the Gene Ontology and curating fly models of human disease.

New improved Gene Families (beta release)

Here at the HGNC we have been working hard over the past few months to develop new improved [gene family pages](#), which are now available as a beta release. We would very much like all of our newsletter readers to visit these new pages over the next six weeks and provide us with feedback. In the future the gene family pages will also be incorporated into our new Search, as described below.

Here is a summary of the improvements we have made:

Gene family index

Our [gene family index](#) is now ordered alphabetically according to family name, with root symbols shown in a separate column. Results are paginated and the user can choose to alter the number of results per page. Typing in either search box narrows down the results based on the input, as shown in the example below.

Fig. 1 The new gene family index

Gene Families

We strongly encourage naming families and groups of genes related by sequence and/or function using a "root" symbol. This is an efficient and informative way to name related genes, and already works well for a number of established gene families. The gene family index below is ordered alphabetically according to family name, with root symbol shown in a separate column. Where applicable the family pages include a curated display of hierarchical relationships between families and allow users to browse easily through each hierarchy. Many family pages contain a description of the family, often sourced from Wikipedia or UniProt, and where they share a particular protein domain we show a graphical display of the domain structure sourced from Pfam via UniProt ID. Each gene family page has a download link at the bottom of the page that generates a text file with all gene symbols and extra data fields such as "Approved Name" and "HGNC ID".

Proposals for new gene nomenclature schemes should be sent to: hgnc@genenames.org

Results: 1 to 11 of 636		Page 1 of 26	25
Name	Root symbol		
<input type="text" value="G protein-coup"/>	<input type="text"/>		
5-hydroxytryptamine (serotonin) receptors, G protein-coupled			
Adhesion G protein-coupled receptors	ADGR		
G protein-coupled receptors			
G protein-coupled receptors, Class A orphans			
G protein-coupled receptors, Class A rhodopsin-like			
G protein-coupled receptors, Class B orphans			
G protein-coupled receptors, Class B secretin-like			
G protein-coupled receptors, Class C metabotropic glutamate/pheromone			
G protein-coupled receptors, Class C orphans			
G protein-coupled receptors, Class F frizzled			
Receptor (G protein-coupled) activity modifying proteins	RAMP		
Results: 1 to 11 of 636		Page 1 of 26	25

New gene family pages

Our gene family pages have also been redesigned to provide more information, conform to a more standardised format, and make it easier to browse between families. An example gene family page for [Cholinergic receptors, muscarinic](#) is shown below.

Fig. 2 The new Gene Family page for "Cholinergic receptors, muscarinic"

Gene Family: Cholinergic receptors, muscarinic (CHRM)

Also known as : "Muscarinic acetylcholine receptors", "Muscarinic receptors", "mAChRs"

A subset of ⓘ : [Amine receptors](#), [Cholinergic receptors](#)

Gene family hierarchy map

```
graph LR; A[G protein-coupled receptors] --> B[G protein-coupled receptors, Class A rhodopsin-like]; B --> C[Amine receptors]; B --> D[Cholinergic receptors]; C --> E[Cholinergic receptors, muscarinic]; D --> E;
```

Muscarinic acetylcholine receptor: Muscarinic receptors, or mAChRs, are acetylcholine receptors that form G protein-receptor complexes in the cell membranes of certain neurons and other cells. They play several roles, including acting as the main end-receptor stimulated by acetylcholine released from postganglionic fibers in the parasympathetic nervous system. Muscarinic receptors were named as such because they are more sensitive to muscarine than to nicotine. Their counterparts are nicotinic acetylcholine receptors (nAChRs), receptor ion channels that are also important in the autonomic nervous system. Many drugs and other substances (for example pilocarpine and scopolamine) manipulate these two distinct receptors by acting as selective agonists or antagonists. [Source: [Wikipedia](#)]

The mapped domains of [P11229](#), encoded by the [CHRM1](#) gene, an example gene within the family. [Source: [Pfam](#) & [UniProt](#)]

Genes contained within the family ⓘ

Approved Symbol	Approved Name	Previous Symbols	Synonyms	Chromosome
CHRM1	cholinergic receptor, muscarinic 1			11q12-q13
CHRM2	cholinergic receptor, muscarinic 2			7q35-q36
CHRM3	cholinergic receptor, muscarinic 3			1q43
CHRM4	cholinergic receptor, muscarinic 4			11p12-p11.2
CHRM5	cholinergic receptor, muscarinic 5			15q26

Downloads ⓘ [Current set only](#)

Gene family names, IDs and aliases

Each gene family now has a unique numerical ID that forms the last part of the gene family page URL to aid linking and downloading. For example the numerical ID for "Cholinergic receptors, muscarinic" is 180. Each family also has a unique gene family name and, where a family has a root symbol, this is shown in

parentheses next to the name. Please note that not all gene family pages equate to a particular set of genes with the same root symbol; in these cases only a gene family name is displayed. Other commonly-used gene family names and abbreviations are listed following the text "Also known as:" e.g. "[Cholinergic receptors, muscarinic](#)" are also known as "Muscarinic acetylcholine receptors", "Muscarinic receptors" and "mAChRs".

Gene family hierarchy map

The new gene family pages provide an improved display of curated hierarchical relationships between families and allow users to browse easily through each hierarchy. As shown in Fig. 2 all gene families that fall into hierarchies include a "Gene family hierarchy map". Clicking on any gene family within the map opens a pop-up that contains a link to that particular gene family page and highlights the current gene family and its direct relatives, and connects them with a red outline. Users can move each gene family within the page by clicking and dragging the box. In addition to the map, text links to any related families are provided within the page. For example, in Fig 2 there are direct links to the [Amine receptors](#) and [Cholinergic receptors](#).

Gene family descriptions

Many of our gene family pages now contain a description of the family. These are often from Wikipedia (as shown in Fig 2) or UniProt (e.g. [Integrins](#)), in which case the source is clearly marked with a link through to the original page; in some cases the descriptions have been written by HGNC curators and these can be identified by [Source: HGNC] (e.g. [IGH orphans](#)); if they come from another source this will be clearly displayed within square brackets.

Example gene mapped domains graphic

Where gene family members share a particular protein domain we often show a graphical display of the protein domain structure for an example gene family member, which is sourced from [Pfam](#) via [UniProt](#) ID. In Fig. 2 the domain structure is shown for the product of the [CHRM1](#) gene. Clicking on a domain within the graphic takes the user through to the Pfam description page for that domain.

Genes within the family

HGNC Symbol Reports for each gene within a family can be accessed by clicking on the Approved Symbol. By default the table of family members is sorted by Approved Symbol, but where the family shares a root symbol the members can be sorted by that symbol even where it is a synonym or previous symbol, e.g. in the DEAD box polypeptides (DDX) family INTS6 is sorted by its previous symbol of DDX26. Note that the symbol used for sorting is highlighted in green to make this clear.

Above the table of family members there is a small pie chart symbol next to the text "Genes contained within the family". Clicking on the pie chart opens a pop-up box containing statistics on the locus types of the genes included in the family. For example, the [Tubulin](#) gene family contains 22 protein-coding genes and four pseudogenes.

Gene family downloads

We are now able to provide a much improved way of downloading gene families as data sets, allowing users to choose between downloading a single family or the entire family hierarchy. For example, users can choose between downloading only the [Cholinergic receptors, muscarinic](#) genes shown in Fig 2 or they can browse through to the [G protein-coupled receptors](#) and download all genes belonging to that hierarchy. Each gene family page has a download link at the bottom of the page that generates a text file with all gene symbols and extra data fields such as "Approved Name" and "HGNC ID". Please note that some gene family pages do not contain a list of genes because these are included to complete the hierarchical structure; these pages enable users to download all the genes from further down the hierarchy, e.g. [Amine receptors](#) and [Cholinergic receptors](#).

We look forward to [receiving your comments](#) on our new Gene Family pages.

Coming soon: an improved Search facility

In addition to our improved gene family pages, we are in the process of developing an expanded Search facility that will include the ability to search our gene families. The new Search will have a similar look and feel to the current one, so this change should not disrupt the user experience but will make it easier for people to find what they want.

Our current Search has a dropdown menu that provides a choice between "Search symbol reports" and "Search Site". As shown below, the dropdown menu for the new Search will give the options "Search everything", "Symbol search", "Family search" and "Site Search".

When searching symbol reports the results filter on the left-hand side still provides the option to filter by locus group or locus type but the locus types are now displayed under the relevant locus group to make the relationships between the groups and types clearer e.g. "RNA, long non-coding" is listed beneath "Non-coding RNA".

Fig. 3 Example results shown for our new Gene Search

Results: 1 to 10 of 39194		Page 1	of 3920	Next >	Last >>
A1BG: alpha-1-B glycoprotein					
HGNC_ID:	HGNC:5	Location:	19q13.43	Locus type:	gene with protein product
A1BG-AS1: A1BG antisense RNA 1					
HGNC_ID:	HGNC:37133	Location:	19q13.43	Locus type:	RNA, long non-coding
A1CF: APOBEC1 complementation factor					
HGNC_ID:	HGNC:24086	Location:	10q21.1	Locus type:	gene with protein product
A2M: alpha-2-macroglobulin					
HGNC_ID:	HGNC:7	Location:	12p13.31	Locus type:	gene with protein product
A2M-AS1: A2M antisense RNA 1 (head to head)					
HGNC_ID:	HGNC:27057	Location:	12p13.31	Locus type:	RNA, long non-coding
A2ML1: alpha-2-macroglobulin-like 1					
HGNC_ID:	HGNC:23336	Location:	12p13	Locus type:	gene with protein product
A2ML1-AS1: A2ML1 antisense RNA 1					
HGNC_ID:	HGNC:41022	Location:	12p13.31	Locus type:	RNA, long non-coding
A2ML1-AS2: A2ML1 antisense RNA 2					
HGNC_ID:	HGNC:41523	Location:	12p13.31	Locus type:	RNA, long non-coding
A2MP1: alpha-2-macroglobulin pseudogene 1					
HGNC_ID:	HGNC:8	Location:	12p13.31	Locus type:	pseudogene
A3GALT2: alpha 1,3-galactosyltransferase 2					
HGNC_ID:	HGNC:30005	Location:	1p35.1	Locus type:	gene with protein product
Results: 1 to 10 of 39194		Page 1	of 3920	Next >	Last >>

This new Search will be released on genenames.org within the next two months, following testing and full release of the improved Gene Families.

Gene Symbols in the News

Several HGNC-approved symbols have appeared in the international media in recent months. A recent report gave [new hope for the treatment of triple-negative breast cancer](#) following the discovery that increased activity of the **BCL11A** gene is associated with this disease. Another study has provided hope that [a cholesterol-lowering drug may provide protection from heart disease](#) – the drug ezetimibe inhibits the function of the **NPC1L1** gene and the new study has found that individuals carrying an inactive copy of this gene are at a reduced risk of coronary heart disease. [A study on the genetics of Finnish prisoners who have committed violent crimes](#) found an association between violent behaviour and variants of the **MAOA** and **CDH13** genes, although it was noted that many individuals carry these variants without showing violent tendencies.

Meeting News

Beth and Elspeth attended the [Livestock Genomics Meeting 2014](#) in Cambridge, UK from 18th-20th September where Beth gave a talk on her recent work on HCO. Elspeth and Kris then attended [Genome Informatics 2014](#) also in Cambridge, UK from 21st-24th September, where they presented a poster entitled "Using HCOP To Standardise Gene Nomenclature Across Vertebrate Genomes".

Ruth attended [Computational RNA Biology](#) in Hinxton, UK from 11th-13th November where she presented a poster on "The challenges of naming long non-coding RNA genes".

Publications

Gray KA, Yates B, Seal RL, Wright MW, Bruford EA. **Genenames.org: the HGNC resources in 2015.** Nucleic Acids Res. 2015 Database Issue PMID:[25361968](#)

The RNAcentral Consortium (inc. Bruford, E and Wright, M). **RNAcentral: An international database of ncRNA sequences.** Nucleic Acids Res. 2015 Database Issue PMID:[25352543](#)

If you would like to be added to our HGNC Newsletter mailing list or if you have questions or comments on any human gene nomenclature issue, please email us at: hgnc@genenames.org

[View the HGNC Newsletter Archive](#)

