HGNC Newsletter Spring 2011

< Previous Issue

There are currently 31021 approved symbols

Next Issue >

In this newsletter

Introducing our new look website

Those of you who have visited genenames.org recently may have had a pleasant surprise. We have completely redesigned our website to make navigation more intuitive and to give every page a consistent look and feel. We hope you all agree that we have achieved this aim and we look forward to your feedback. Here is a quick guide to the new site:

The new HGNC banner

Each page on genenames.org has the same banner at the top with tabs and drop-down menus, meaning that you can now navigate quickly to all of our pages and tools without having to return to the homepage. The banner includes a simple version of the Quick Gene Search that searches for results that contain any part of your search term. The HGNC logo on the left side of the banner works as a "home" button.

The new homepage

Our homepage features the full version of the Quick Gene Search, which allows you to choose between searching for terms that equal, begin or contain your search term. You can also select how many results you want to view per page. At the bottom left of the page, there is a website search box that allows searching of gene family pages, HGNC documentation and information pages. Please note that you should use Quick Gene Search, Advanced Gene Search or the List Search to search for gene Symbol Reports.

There is a brand new feature, "Browse approved symbols by chromosome", an interactive graphic of the human chromosomes and the mitochondrion. Clicking on a chromosome graphic will take you to the HGNC Statistics and Downloads page for that chromosome. There you can download the total number of approved symbols for the chromosome or you can browse symbols by locus type e.g. you can view the number of protein-coding genes vs. pseudogenes with approved nomenclature on chromosome 7. You can also link through to our custom downloads page which will allow you to specify the exact data fields to download for the chromosome that you selected.

We also now have a 'News' section on the bottom right of the homepage. Here we will report on important updates to the HGNC project and include a few links to gene symbols that have featured in the international media. The box on the left features a list of updated FAQs.

The new Symbol Report

We have completely redesigned our Symbol Reports so that the HGNC nomenclature is more prominent and related data are grouped together. Our core HGNC data are now displayed at the top of the Symbol Report in a shaded box; this features the approved nomenclature, the unique HGNC ID for the symbol, previous nomenclature, synonyms, the locus type and the chromosomal location. You will notice that each core data field title and the titles of the following sections are accompanied by orange plus signs; these are links to the relevant part of our updated Symbol Report documentation. One of the many improvements to our Symbol Reports is that we now provide direct links to our curated gene family pages. If the gene is a member of one of these families, the link is shown underneath the core data section.

In addition to the HGNC data mentioned above, our Symbol Reports include links to a wide variety of relevant data. Data links that have been checked and manually curated by members of the HGNC are indicated by the letter "**C**" after the link. Data that have been downloaded from external sources are indicated by the letter "**D**". Please note that the HGNC cannot guarantee the reliability of downloaded data. Data links are organised into the following sections:

GENE FAMILY – only present if the gene is a member of a curated HGNC Gene Family page.

SPECIALIST DATABASE –this section provides links to databases that are relevant to only certain classes of gene. The section is not displayed on Symbol Reports that do not have any of these links.

HOMOLOGS – we have grouped all of our homology-related links into one section, including a link to our own HCOP orthology data mining tool. In addition to providing links to the gene pages at MGI and RGD for the mouse and rat orthologs, we now also display the approved mouse and rat gene symbols.

NUCLEOTIDE SEQUENCES – as before, we provide links to curated accessions, RefSeq transcripts and the CCDS (consensus coding sequence) records, where

applicable. We now also provide links to the Vega gene sequence, which were not available on our previous site.

GENE RESOURCES – in addition to linking to gene annotation pages at Entrez Gene, Ensembl, UCSC and Vega, we now also provide direct links to the Genome Browsers supported by these four projects.

PROTEIN RESOURCES – we continue to link to the UniProt record for the encoded protein product. A new feature for our Symbol Reports is a link to the InterPro Protein Match page; this page shows all predicted protein signature for the encoded protein by the InterPro member databases.

CLINICAL RESOURCES – this section provides links to associated phenotypes, diseases and mutations associated with the gene.

REFERENCES – this section displays the PubMed IDs for references curated by HGNC members for the gene. There is also a link to the same references at CiteXplore.

OTHER DATABASE LINKS – this section includes links to external resources that cannot be grouped into one of the categories above. There are two new sets of links in this section, which were not present in our previous Symbol Reports:

- A link to the Reactome project, a manually curated, peer-reviewed signaling pathway database.
- A link to a list of all Gene Ontology (GO) terms annotated for the gene product at the QuickGO project.

For an example of Symbol Report that contains data in all of sections, see ALDH1A1.

Gene Family Pages

We have been working on improving both the prominence and breadth of our gene family pages. It is now easier to find these pages within our website; there is a Gene Families tab on the new banner and gene Symbol Reports now contain links straight through to relevant gene family pages. We are continuing to add new families to this valuable resource. Here is a list of families that we have added to coincide with the release of the new website:

ADP-ribosylation factor-like genes; Amiloride-sensitive cation channels; BEN domain containing; Cas scaffolding protein family; Coiled-coil-helix-coiled-coil-helix domain containing; C-type lectin domain containing; Components of oligomeric golgi complex; Cytochrome b family; High mobility group family; 5-hydroxytryptamine receptors; Ribosomal protein L genes; Transfer RNAs

We have also improved how ncRNA (non-protein-coding RNAs) gene families are represented, by assigning individual pages to gene family groups.

Search Genes tools

Our new website allows you to access a page that includes all three of our "Search Genes" tools (Quick Gene Search, Advanced Gene Search and List Search) via the Search Genes tab on the banner. Alternatively, you can use the drop-down menu beneath this tab to navigate to separate versions of the tools. The List Search has been updated and provides more search options than the previous version. You can now choose whether to perform a case sensitive or case insensitive search, to view the results as an HTML table or as a tab delimited text file, to retain or exclude search terms that do not match any HGNC symbols and to sort your results in either the same order as your original list or by alphabetical order. Please read our updated documentation for more information on these three tools.

Gene Symbols in the news

Many of our approved gene symbols have featured in the international media in recent months. There have been several studies associating genes with disease: loss of DICER1 activity has been linked with macular degeneration; a mutated form of the *NKX2-1* gene has been linked to a poor prognosis for patients with lung tumours; while variants of the *APOE* gene are linked with the slow recovery regrowth of nerves outside the brain and spinal cord following trauma.

There have been several reports on genes that affect birth and child development. Variants of the *FSHR* gene were shown to be strongly associated with pre-term births, while a low carbohydrate diet in early pregnancy has been shown to affect epigenetic markers at the *RXRA* gene, which have been associated with childhood obesity. The *ROBO1* gene has been linked to the development of speech in infants.

In behavioral news, both the *CYP1A2* gene and its regulator, *AHR*, have been linked to high caffeine consumption.

Meeting News

Elspeth attended the Human Genome Meeting (HGM 2011) meeting in Dubai, United Arab Emirates from 14-17 March, while Matt attended Abcam 2011, Non-coding RNA, Epigenetic Memory, and the Environment in London, UK from 14-15 April 2011.

At the end of this month, Louise and Ruth will be attending the European Human Genetics Conference (ESHG 2011) in Amsterdam, Netherlands. They will present a poster entitles "Using HGNC to improve and expand Gene Family resources". Please come and visit the poster to chat to them if you are attending the meeting.

Publications

Wright MW, Bruford EA. **Naming 'junk': Human non-protein coding RNA (ncRNA) gene nomenclature**. Hum Genomics. 2011 Jan 1;5(2):90-8. PMID:21296742.

This paper describes our progress in actively engaging with the RNA research community in a quest to provide unique names for all the sequences encoding non-protein-coding RNAs (ncRNAs). Most of the classical small ncRNA genes have now been provided with a unique nomenclature, and work on naming the long (>200 nucleotides) non-coding RNAs (lncRNAs) is ongoing.

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