HGNC Newsletter Winter 2013-2014

< Previous Issue

There are currently 38082 approved symbols

Next Issue>

In this newsletter

Improvements to our website

As regular visitors to genenames.org will have noticed, we have made a number of changes to improve the functionality of our site and search tools and redesigned the layout of our homepage and Symbol Reports. We look forward to your feedback on these improvements. The main features of this new release are:

Improved homepage

Our homepage now features text and links that are designed to help those unfamiliar with our site to view our resources and quickly find what they are looking for. We have replaced the karyotype image shown on our previous homepage with a word cloud that displays the most common root symbols. Clicking on a root symbol within the image runs a search of our database e.g. selecting "**SLC**" will find all genes in our database with symbols that begin with SLC. The karyotype image has moved to the "Statistics & Downloads" page and provides the same functionality as before.

New solr search

One of the biggest changes is the implementation of a solr search that replaces both our previous Quick Gene Search and Advanced Gene Search. The search is available on the banner of every page on genenames.org and provides a choice between searching our Symbol Reports or searching within the text of our site. Solr allows users to select more relevant results by supporting searches with the following:

a. wildcards - users can add an asterisk (*****) to searches to stand in place for one or more characters, or a question mark (**?**) to stand in place for a single character substitution. e.g. B?GALT* will find all genes with the root symbol B3GALT and also all genes with the root symbol B4GALT.

b. the use of logic operators "AND", "OR" and "NOT". The default search uses "OR".

c. phrases - users can search for one specific phrase by using quotation marks e.g. "A2ML1 antisense RNA 1" finds only the one gene with this gene name, while A2ML1 antisense RNA 1 finds all genes that contain any of the individual terms within that name.

d. a specific field - users can do this by entering the field name immediately followed by a colon (:) and then the query, e.g. refseq_accession:NM_033360 or alias_name:"A-kinase anchor protein, 350kDa".

Results are ordered by relevance and each result shows the **approved gene symbol, approved gene name, HGNC ID**, **chromosome location** (if specified), **locus group** and the name of the field that the keyword/ID **matches** e.g. Approved Symbol or Synonyms. Users can also filter down the results by **locus type** or **locus group** and can change the number of hits displayed per page.

For a full description of our search including a list of our field names, please visit our Search help page.

New REST web service

Our new REST web service is a quick and easy way of searching and fetching data from our database within a script/program. Users may request results as either XML or JSON making our data easier to parse. For more information view our REST web-service help page.

Improved "List Search" now called "Symbol checker"

The tool "**List Search**" has been renamed "**Symbol checker**" and the interface to the application has been revamped for clarity and ease of use. The tool contains the same functionality as the old "List Search" but we are happy to report that we have increased the speed of the search for large symbol lists and have added a sortable results table. As well as a change in name, the URL has changed from **http://www.genenames.org/cgi-bin/hgnc_bulkcheck.pl** to http://www.genenames.org/cgi-bin/symbol_checker. For more information about our new "Symbol checker" visit our Symbol checker help page.

A number of other URLs have been changed within our site, please see our New features and changes page for a full list of these.

Coming soon: a new HCOP

We are currently developing a new version of the HCOP (HGNC Comparison of Orthology Predictions) tool. In addition to the species currently featured in HCOP we will be incoporating data from pig and anole lizard and adding orthology calls from the OrthoDB and PANTHER databases. We will be releasing a beta version of this tool in the next few days and would welcome your feedback, please visit our beta page for more information.

Functional pseudogenes

Although pseudogenes have traditionally been thought of as inactive, there have been a number of studies showing that transcribed pseudogenes can have a regulatory role. For example, **PTENP1** regulates levels of **PTEN** by binding to PTEN-targeting miRNA (see PMID: 20577206). Where we are aware of these cases we now add "(functional)" to the end of the gene name so that these genes can be searched e.g. the full name of **PTENP1** is "**phosphatase and tensin homolog pseudogene 1 (functional)**".

Genes on alternate loci

We are making a slight change to the way genes annotated on alternate reference loci are displayed within Symbol Reports. The Chromosomal Location field will now display the location plus the term "alternate reference locus" instead of the assembly unit name e.g. see the Symbol Report for **C4B_2**. For a full description of how we select genes to be named on alternate loci, please see our recent publication "**Vive la difference: naming structural variants in the human reference genome**." PMCID: PMC3648363.

Genes Symbols in the News

A recent study has found that Neanderthal ancestry is not always an advantage - it turns out that a variant copy of the **SLC16A11** gene associated with an increased risk of type 2 diabetes originally occurred in Neanderthals and was passed down to humans following breeding between the two species. Another study reported that there may be some truth to the idea that obese people have a slower metabolism. Some individuals carrying a **KSR2** variant associated with increased body mass do indeed have slower metabolism but they also have increased appetites. There has also been a breakthrough in research on a severe form of the muscle disease nemaline myopathy that causes death days after birth - the causative mutation has been found in the **KLHL40** gene. This means that it will be possible to offer screening for affected families in the future.

Publications

Heit C, Jackson BC, McAndrews M, Wright MW, Thompson DC, Silverman GA, Nebert DW, Vasiliou V. **Update of the human and mouse SERPIN gene superfamily.** Hum Genomics. 2013 Oct 30;7:22. PMID:24172014 PMCID:PMC3880077

Macarthur JA, Morales J, Tully RE, Astashyn A, Gil L, Bruford EA, Larsson P, Flicek P, Dalgleish R, Maglott DR, Cunningham F. **Locus Reference Genomic: reference sequences for the reporting of clinically relevant sequence variants.** Nucleic Acids Res. 2014 Jan 1;42(1):D873-8 PMID:24285302

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