HGNC Newsletter Winter 2012-2013

< Previous Issue

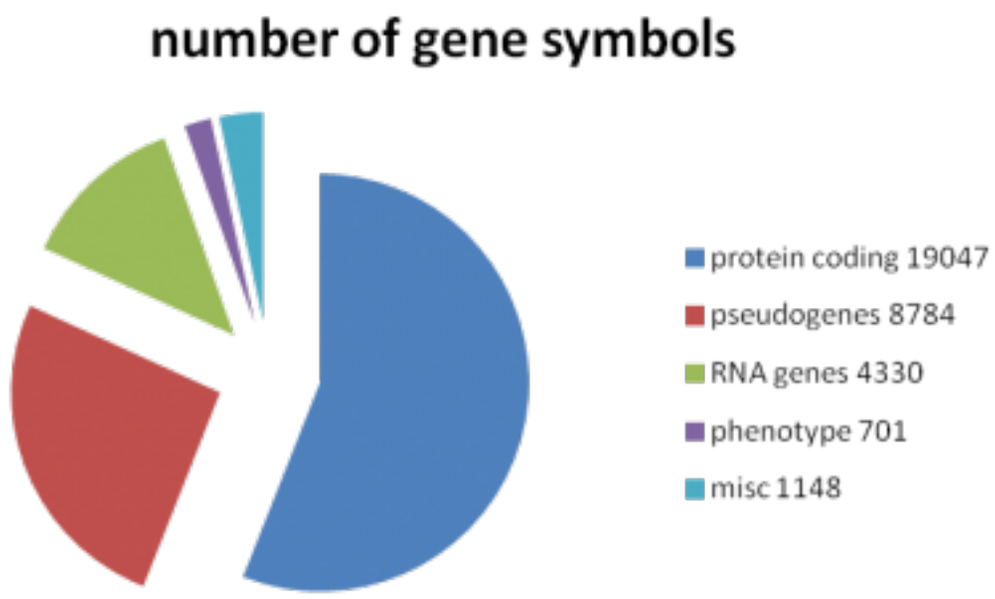
There are currently 34010 approved symbols

Next Issue>

In this newsletter

Gene Symbol Landmark

We are pleased to announce that we have now approved over 34,000 gene symbols. The pie chart below shows the proportion of approved symbol per main locus type. For a full breakdown of our approved symbols by locus type, please visit our Statistics and Downloads page.



Displaying ambiguous locus types

There are some genes where the locus type is not certain, often where the gene has been annotated as protein-coding by one annotation group and as a pseudogene or a non-coding RNA by another annotation group. We want to make our users aware of these instances and have done this by adding a double dagger symbol ‡ next to the gene symbol at the top of the relevant Symbol Reports. Scrolling over the double dagger triggers a pop-up box with the words "ambiguous locus type", while clicking on the double dagger shows the full definition "Ambiguous locus type. This symbol is shown when the protein coding status is uncertain." See the **ASAH2C** Symbol Report for an example. We would welcome any further information/data on these loci that would help us decide to which locus type they should be assigned.

Improvements to our Custom Download Resources

We have updated our Custom Downloads page to become more organised and hopefully easier to use. For example, we have separated out the data that are downloaded from external resources into a separate section away from our HGNC-curated data. We have also provided symbols next to each data column; clicking on each symbol opens a box containing information on that particular data column. We have a new output format for the Custom Downloads tool; in addition to text and Perl code, users can now choose to create a tiny URL to bookmark and view their results at any convenient time.

Additionally, the output for the Locus Specific Databases data column has been improved. Each individual Locus Specific Database name and link is now enclosed within double quotation marks, which means that the data for each database will not be separated if the link or name contains a comma. For example, the Locus Specific Database custom downloads output for the **RS1** gene now contains data in the following format:

"X-Linked Juvenile Retinoschisis|http://www.LOVD.nl/RS1","Mutations of the X-linked Retinoschisis Gene|http://www.retina-international.org/files/sci-news/xlrsmut.htm","Mental Retardation database|http://grenada.lumc.nl/LOVD2/MR/home.php?select_db=RS1","LOVD - Leiden Open Variation Database|http://grenada.lumc.nl/LOVD2/eye/home.php?select_db=RS1"

instead of:

X-Linked Juvenile Retinoschisis|http://www.LOVD.nl/RS1,Mutations of the X-linked Retinoschisis Gene|http://www.retina-international.org/files/sci-news/xlrsmut.htm,Mental Retardation database|http://grenada.lumc.nl/LOVD2/MR/home.php?select_db=RS1,LOVD - Leiden Open Variation Database|http://grenada.lumc.nl/LOVD2/eye/home.php?select_db=RS1

New Gene Family Resources

We have expanded our gene family resource by a considerable amount over the last few months. Here is a list of the new gene family pages:

Alkaline ceramidases
Ankyrin repeat domain containing
Apolipoproteins
Basic leucine zipper proteins
BTB domain containing
Collagens
EF-hand domain containing
Fatty acid desaturases
General transcription factor IIH complex subunits
G patch domain containing
Intermediate filaments
Maestro heat-like repeat containing
Nuclear hormone receptors
Neuroblastoma breakpoint family
OTU domain containing
Paraoxonases
Parvins
Phosphatase and actin regulators
Pleckstrin homology (PH) domain containing
PRAME family
RNA polymerase subunits
Septins
Serine/arginine-rich splicing factors
Sterile alpha motif (SAM) domain containing
Synaptotagmins
Tetratricopeptide (TTC) repeat domain containing
Tudor domain containing
U-box domain containing
WAP four-disulfide core domain containing
Zona pellucida glycoproteins
ZYG11 cell cycle regulator family

We have also added the Immunoglobulin superfamily domain containing family page, that is subdivided into the V-set domain containing, C1-set domain containing, C2-set domain containing, I-set domain containing and Immunoglobulin-like domain containing pages. The pages do not contain links to gene fragments, please see our immunoglobulins page and T cell receptors page for these genes. The immunoglobulin superfamily is a hugely diverse family and includes many different proteins, each containing at least one V-set, C1-set, C2-set, I-set, or immunoglobulin-like domain. V-set domains are found within the variable domains of antibodies but are also found in other proteins with a wide variety of functions such as cell adhesion. C1-set domains are found within the constant region of antibodies and are found within other classes of protein with an immune-related function. C2-set domains are similar to C1-set domains but are found mostly on cell surface proteins and I-set domains are found in many different proteins involved in cell adhesion.

We have consolidated our zinc finger family resource by creating one Zinc Finger master page from the many separate pages we had previously. In addition we have added several new subclasses, including: FLYWCH, DNL, C_4H_2, CCHC , GATA, MIZ and C_2CH types. C_2H_2 (Krüppel-type) is the best-characterised and largest class of zinc fingers, so we have created a dedicated C_2H_2page where you can view all the genes in this class or just those where there is an accompanying KRAB, BTB/POZ or SCAN domain. Wherever possible we have included links to the KZNF Gene Catalog.

Gene Symbols in the News

Approved gene symbols continue to appear in the international media, with several reports of links between genes and disease. A mutation in the **TREM2** gene, which has a role in immune responses and has been linked to chronic inflammation, is three to four times more common in patients with Alzheimer disease. Mutations in the **POLE** and **POLD1** genes, which both have a role in DNA repair, have been associated with a higher risk of developing bowel cancer, while a variant of the **RASGRF2** gene has been associated with binge drinking. In a demonstration of how genes and environment can interact to produce disease, a study has shown that individuals carrying a **SLC30A4** variant associated with a risk of developing type 2 diabetes seem to be at lower risk if they have high levels of beta carotene in their blood. There is promising gene therapy news following the successful transfection of the human **TBX18** gene into guinea pig hearts to create a new biological pacemaker that controls heartbeat. The **TBX18** gene is active in human embryonic hearts at the stage where pacemaker cells are formed. The study has led to hopes that the technique may work for human hearts in the future.

Meeting News

Elspeth attended the Plant and Animal Genome XXI Conference in San Diego where she presented a poster entitled "All animals are not equal: some genomes are more equal than others" about our new efforts to coordinate gene naming across vertebrate species. She was also delighted to be able to give a short presentation at the NSRP-8 (US National Animal Genome Research Program) Animal Genome business meeting to inform the program coordinators and collaborators of our work across vertebrates and enlist their help; this has already resulted in some valuable contacts. Please email us at hgnc@genenames.org if you are interested in gene naming in a vertebrate species you are working on.

Matt and Elspeth will be attending the 6th International Biocuration Conference in Cambridge, UK from 7th-10th April and will then be travelling soon afterwards to

the HGM 2013 and 21st International Congress of Genetics in Singapore, which runs from the 13th-18th April.

Publications

Gray KA, Daugherty LC, Gordon SM, Seal RL, Wright MW, Bruford EA. **Genenames.org: the HGNC resources in 2013.** Nucleic Acids Res. 2013 Jan 1;41(D1):D545-52. PMID: 23161694 PMCID: PMC3531211

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