



HGNC Newsletter Winter 2009-2010

[< Previous Issue](#)

There are currently **28676** approved symbols

[Next Issue >](#)

HGNC would like to wish you all the best for 2010

In this newsletter: we present a new piRNA cluster nomenclature, a new bHLH gene family page, changes to HCOP and a focus on the difference between mapped and curated data in our database.

piRNA clusters

Piwi-interacting RNAs (piRNAs) are the largest class of small RNAs expressed in vertebrates. They are found in clusters throughout the genome, with each cluster containing from around ten up to thousands of piRNAs. [piRNABank](#), a web resource that classifies piRNAs and groups them into clusters, have so far identified 114 piRNA clusters in the human genome. We have now named each of these clusters with the symbol PIRC# for 'piwi-interacting RNA cluster #'. The PIRC# symbols can be viewed in full at our [RNA family page](#).

New links to Pseudogene.org

We have recently started providing links to annotations at the [Pseudogene.org](#) database from the Specialist Database Links section of relevant pseudogene symbol reports. An example of a symbol report that contains such a link is [ATF1P1](#). For more information on the pseudogene.org project, see the following publication: "Pseudogene.org: a comprehensive database and comparison platform for pseudogene annotation." (PMID: [17099229](#)). **Breaking news:** we are currently in the process of refining our pseudogene symbols to produce a more consistent nomenclature and will be providing details on this in the near future.....

New gene family page

We now have a gene family page for the [basic helix-loop-helix \(BHLH\) family](#). The basic helix-loop-helix structural motif is commonly found in transcription factors, many of which are involved in developmental processes. All of the genes encoding members of this family, as classified in the 2008 article from Stevens, Roalson & Skinner, "Phylogenetic and expression analysis of the basic helix-loop-helix transcription factor gene family: genomic approach to cellular differentiation" (PMID [18557763](#)), have a bHLH# symbol listed in the corresponding entry in the HGNC database. As many of these genes already had an alternate established nomenclature the bHLH# symbol is mostly included as an alias to enable retrieval of all the family members, but where it has been adopted as the approved symbol the root has been capitalised to BHLH# in line with our nomenclature guidelines.

Mapped vs. curated data

It has come to HGNC's attention that some users may be unaware of the difference between the mapped and curated data contained in our Gene Symbol Reports and downloads. Please note that mapped data are derived from external sources and as such are not subject to our strict manual checking and curation procedures. Therefore, the HGNC are unable to guarantee the same high quality for mapped data as for our curated data. Some identifiers, such as UCSC IDs and UniProt IDs, are always mapped. Others such as EntrezGene IDs and Ensembl IDs may be either curated by HGNC or mapped by external sources, but curated data will always be displayed in preference to mapped IDs. Mapped data are identified in Gene Symbol Reports by the disclaimer "mapped data supplied by [source]" in the header of the relevant field. On our [custom downloads](#) page, curated and mapped data are separate download options, and consequently appear in separate fields. Fields beginning with "gd" contain curated data, whereas fields beginning with "md" contain mapped data. For example, the field `gd_pub_ensembl_id` contains HGNC curated Ensembl IDs, while `md_ensembl_id` contains mapped data supplied by Ensembl.

New Species in HCOP

The [HCOP](#) (HGNC Comparison of Orthology Predictions) tool has been updated with orthology predictions between human and three new species: macaque, opossum and horse. This now raises the total number of species to 15 in HCOP, with orthology predictions also available between human and mouse, rat, cow, chicken, dog, platypus, fruitfly, chimp, zebrafish, *C. elegans* and *S. cerevisiae*. In addition to being able to search HCOP on a gene-by-gene basis, the complete orthology assertion data for each species set are available as text files from the Bulk Downloads section of the HCOP home page.

Meeting News

Matt recently attended the [Plant and Animal Genome XVIII Conference](#) held in San Diego, USA, from 9th-13th January. He presented a poster about gene nomenclature across species entitled "If we could talk to the animals", but also enjoyed talking to the humans at the meeting.

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

[Return To HGNC Newsletter Archive](#)