



## HGNC Newsletter Winter 2011-2012

[< Previous Issue](#)

There are currently 32695 approved symbols

[Next Issue >](#)



### In this newsletter

[Renaming genes with C\\$orf# symbols](#)

[Spotlight on pseudogenes](#)

[New Gene Family Resources](#)

[Gene Symbols in the News](#)

[Meeting News](#)

[Publications](#)

### Renaming genes with C\$orf# symbols

As most of our users and contributors are aware, the HGNC aims to name every human gene based on the character or function of the gene. However, in some instances the only available information on a gene is at the sequence and annotation level. If the function of a gene is unknown, we then check for characterised orthologs in other species that we may name the gene after, or for structural information on the encoded protein that may provide the basis for a gene name. If there is no other source for a name and the gene is predicted to encode a protein, we assign a C\$orf# symbol, where \$ represents the chromosome on which the gene is located and # is the next number in a numerical series. Although C\$orf# symbols are user-friendly and many researchers choose to publish using these symbols, we regard them as temporary because they do not provide information on the character or function of the gene and are not easily propagated across species. Many C\$orf# symbols were assigned years ago, and we are now in the process of renaming C\$orf#s where new information is available. Louise used her experience from her previous role as a curator at [InterPro](#) to create a list of C\$orf# genes that encode proteins with a predicted functional domain or structural repeat, or that are predicted to be members of a protein family. We have also been searching the scientific literature and following gene annotation updates for C\$orf# genes. As a result, we renamed 155 C\$orf#s with more informative gene symbols and names during 2011 and we hope to rename many more in 2012. If you know of any genes with a C\$orf# symbol that you think could be renamed, please contact us via [hgnc@genenames.org](mailto:hgnc@genenames.org), or use our [gene symbol request form](#).

### Spotlight on pseudogenes

Pseudogenes now make up around 23% of all genes named by the HGNC and are the second largest class of genes with approved nomenclature after protein-coding genes. Pseudogenes are traditionally thought of as the dysfunctional members of gene families, often overlooked because they have no protein-coding ability. A recent Nature paper ([Poliseno et al. 2010](#)) suggests that some expressed pseudogenes do have a biological function, acting as decoys for microRNAs that would otherwise downregulate the mRNA of the protein-coding gene from which they are derived (e.g. [PTENP1](#) regulates the expression of [PTEN](#)). Even if most pseudogenes prove to be non-functional their proper annotation is extremely important since their analysis can provide an invaluable insight into the evolution of a gene family; equally these loci are also discussed in the literature and hence need a unique name. It is also useful to know which loci represent pseudogenes since they can be confused as protein-coding genes by both gene prediction tools and in molecular genetic experiments.

For each named pseudogene, we provide links to relevant annotation projects from the HGNC gene entry, including [Pseudogene.org](#), [Entrez Gene](#), [Vega](#) and [Ensembl](#). Where possible we name pseudogenes after protein-coding parents using the symbol of the parent gene, the letter 'P' and a unique number e.g. [CCNJ1P1](#) is a pseudogene of [CCNJ](#). In some cases the parent gene cannot be identified, for example where the pseudogene is present in a cluster with several protein-coding genes of the same family. In these cases, we name the pseudogene as part of that family, but denote its pseudogene status with the addition of a 'P' at the end of the symbol e.g. [ZNF890P](#). We also name unitary pseudogenes that have functional protein-coding orthologs in other species. In these cases, the pseudogene is assigned the same symbol as its functional ortholog, except a 'P' is usually added to the symbol, e.g. [GULOP](#). Certain loci encode functional proteins in some individuals, while other individuals carry non-functional, or pseudogenized, alleles of these genes. In these cases, we name the gene according to the protein-coding allele but add the term "gene/pseudogene" to the gene name to indicate that the gene is a segregating pseudogene, see the gene entry for [CASPI2](#). There are exceptions to the pseudogene symbol formats described here in cases where the pseudogene symbol has been used extensively in the literature, or where the pseudogene belongs to a nomenclature system devised by a specialised committee that we coordinate with, such as the T cell receptor and immunoglobulin pseudogenes. If you know of any specific pseudogenes we have not yet named that need a name please contact us ([hgnc@genenames.org](mailto:hgnc@genenames.org)).

### New Gene Family Resources

We have added quite a few new gene family pages to [genenames.org](http://genenames.org) in the last couple of months. Here is a list:

[A-L](#)

[M-R](#)

[S-Z](#)

<a href="#">Carboxylesterases</a>	<a href="#">Mex-3 homologs</a>	<a href="#">S100 calcium binding proteins</a>
<a href="#">CDGSH iron sulfur domain containing</a>	<a href="#">Mitochondrial ribosomal proteins</a>	<a href="#">Signal-regulatory proteins</a>
<a href="#">Cyclin-dependent kinases</a>	<a href="#">Myosins</a>	<a href="#">Small G protein signaling modulators</a>
<a href="#">Engulfment and cell motility proteins</a>	<a href="#">N(alpha)-acetyltransferase subunits</a>	<a href="#">Sorting nexins</a>
<a href="#">Glycerol kinases</a>	<a href="#">ORAI calcium release-activated calcium modulators</a>	<a href="#">Specificity protein/Krüppel-like transcription factors</a>
<a href="#">HAUS augmin-like complex subunits</a>	<a href="#">Pellino homologs</a>	<a href="#">Trafficking protein particle complex</a>
<a href="#">Late cornified envelopes</a>	<a href="#">Proline-rich transmembrane proteins</a>	<a href="#">Transmembrane and coiled-coil domain containing</a>
<a href="#">Latent transforming growth factor, beta binding proteins</a>	<a href="#">Ribonucleases, RNase A</a>	<a href="#">Voltage-dependent anion channels</a>
<a href="#">Lipocalins</a>		<a href="#">WAS protein homologs</a>
		<a href="#">Zinc finger, AN1-type domain containing</a>

---

## Gene Symbols in the News

There have been a number of international news reports featuring approved gene symbols over the last few months, a selection of which is featured here. There have been several reports of associations between genes and disorders. Individuals carrying a particular variant of the [CYP27B1](#) gene have a [greater risk of developing multiple sclerosis](#), while a mutated form of the [PRRT2](#) gene was found in [19 out of 23 families studied with cases of benign familial infantile epilepsy](#), and a study found that [15% of patients analysed with chronic lymphocytic leukaemia](#) carried a mutated copy of the [SF3B1](#) gene. Recent work shed more light onto [the relationship between the omentum and ovarian cancer](#); upregulation of the [FABP4](#) gene was found in all human ovarian omental metastases studied, while *Fabp4* knockout mice models for ovarian cancer showed reduced tumour burden. A child born with [no pancreas](#) was found to carry a defective copy of the [GATA6](#) gene; researchers now hope that this knowledge could help in future work on producing pancreatic beta cells to treat type 1 diabetes. Finally, there was more insight into the relationship between genes and characteristic human behaviour: a study found that individuals carrying a variant of the [ABCC9](#) gene [need greater levels of sleep](#) than those without the variant.

---

## Meeting News

The HGNC will be attending [HGM 2012](#) in Sydney, Australia from 11th-14th March. Matt will be presenting a poster entitled "ncRNA gene nomenclature: The long and small of it". All human ncRNA genes named to date can be found at the HGNC ncRNA webpage: [www.genenames.org/rna](http://www.genenames.org/rna).

---

## Publications

Jackson BC, Thompson DC, Wright MW, McAndrews M, Bernard A, Nebert DW, Vasiliou V. **Update of the human secretoglobin (SCGB) gene superfamily and an example of 'evolutionary bloom' of androgen-binding protein genes within the mouse Scgb gene superfamily.** Hum Genomics. 2011 Oct 1;5(6):691-702. PMID: [22155607](#)

---

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](mailto:hgnc@genenames.org)

[View the HGNC Newsletter Archive](#)