

GUEST EDITORIAL

Nomenclature: Genes, Weights and Measures, Animals, Elements, and Planets

Nomenclature is a controversial issue in any area, but particularly within science. However, the importance of standardization of terminology over the last two centuries has been recognized in many different fields. To this end, various authorities have introduced standardization in the forms of SI units [Conférence Générale des Poids et Mesures (CGPM), established in 1875], names and symbols for the chemical elements [IUBMB-IUPAC Joint Commission on Biochemical Nomenclature (JCBN), re-established in 1977], names for the planets [International Astronomical Union (IAU) Nomenclature Committee, established in 1922], scientific names for organisms [International Commission on Zoological Nomenclature (ICZN), established in 1895], and approved gene symbols. The authority charged with approving and implementing human gene names and symbols is the Human Gene Nomenclature Committee (HGNC), which is a subcommittee of HUGO (The Human Genome Organisation). This was formed in the 1970s [first reported by Shows *et al.* (1)] when it was realized that a consistent, standardized gene nomenclature would be vital to the collaborative effort of maintaining databases of genes. Nomenclature committees were soon established for both human and mouse genes and guidelines (2, 3) published in conjunction with the reports from the HGM (Human Genome Mapping) meetings. In a further development, databases, initially the Genome Database (GDB) and the Mouse Genome Database (MGD), also implemented “approved” gene nomenclature as part of their curation system.

Since the inception of the HGNC, the number of databases has continued to expand, with many more now available online. This means that standardized gene symbols are even more important as they ensure correct labeling of any individual gene across the breadth of databases. All the major databases containing information on human genes now use the HGNC standard for approved symbols. Thus, when searching with the symbol *ABCCI*, information about only that gene will be retrieved from LocusLink, GDB, OMIM, GeneCards, GenBank and GENATLAS. This standardization will become more crucial once the whole of the human genome is sequenced; there are now over 10,400 approved gene symbols in the HGNC online database, but there are at least another 70,000 genes still to name.

However, it is not just human gene databases which are making use of the information and invaluable curation

skills of the HGNC, but also many other organism gene databases. Where possible, these try to use the same gene symbols for confirmed orthologous genes to maintain consistency and searchability. This becomes clear when searching the approved symbols *COL2A1* and *COL4A1*, members of the collagen gene family. *COL2A1* retrieves orthologous genes from the human, chicken, cattle, zebrafish, mouse, rat and sheep databases, while *COL4A1* retrieves orthologous genes from the human, mouse, pig and cattle databases.

In addition to these resources, there are a number of journals which now insist on the use of approved gene symbols. These include *Genomics*, *Nature Genetics*, *Annals of Human Genetics*, *Mammalian Genome*, *Cytogenetics and Cell Genetics*, and of course *Radiation Research*. There are also a number of other journals which mention nomenclature in their instructions to authors. Thus the penultimate barrier to accurate information retrieval is being overcome, as approved symbols can now be found in the literature databases such as PubMed, contained in the titles and abstracts. Some journals have tried to include all known aliases of a particular gene in a publication to ensure understanding by all parties. This is unnecessary and confusing, so our recommendation is to use the approved symbol alongside the alias favored by the author in the title and/or abstract and to use their favorite symbol (preferably the approved one) in the rest of the paper.

However, the final barrier to finding information about a specific gene lies in the hands of the researcher, prior to publication. If everyone contacted the relevant Nomenclature Committee to obtain approved gene symbols and then used these in the title and abstract of their paper, this would greatly enhance information retrieval in all areas. The usual arguments against approved symbols are that they are memorable, unpronounceable, unknown, illogical and not the original symbol. While some of these arguments are occasionally true, approved symbols are always unique, and the HGNC endeavors to ensure that they are accurate, meaningful and memorable and reflect the wishes of the relevant scientific community. We are always eager for new volunteers to act as Specialist Advisors for different gene families to give us a greater range of expertise to call upon. Thus approved gene symbols aim to be key identifiers agreed by the community and not an imposition by a faceless bureaucratic committee.

The Nomenclature Committee Web page can be found at URL <http://www.gene.ucl.ac.uk/nomenclature/>, and we can be contacted via nome@galton.ucl.ac.uk.

URLs

Annals of Human Genetics: <http://www.cup.cam.ac.uk/journals/hge/>

Cattle database: <http://www.ri.bbsrc.ac.uk/cgi-bin/arkdb/browsers/browser.sh?species=cattle>

CGPM, Conférence Générale des Poids et Mesures: International System of Units: <http://physics.nist.gov/cuu/Units/international.html>

Chicken database: <http://www.ri.bbsrc.ac.uk/cgi-bin/arkdb/browsers/browser.sh?species=chicken>

Cytogenetics and Cell Genetics: <http://www.karger.ch/journals/ccg/>

GDB Genome Database: <http://gdbwww.gdb.org/gdb/>

GENATLAS: <http://bisance.citi2.fr/GENATLAS/>

GenBank: <http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=Nucleotide>

GeneCards: <http://bioinformatics.weizmann.ac.il/cards/>

Genomics: <http://apnet.com/www/journal/ge.htm>

HGNC, Human Gene Nomenclature Committee: <http://www.gene.ucl.ac.uk/nomenclature/>

HGNC guidelines: <http://www.gene.ucl.ac.uk/nomenclature/guidelines.html>

HGNC online database: <http://www.gene.ucl.ac.uk/cgi-bin/nomenclature/searchgenes.pl>

HGNC Specialist Advisors: <http://www.gene.ucl.ac.uk/nomenclature/advisors.html>

HUGO, The Human Genome Organisation: <http://www.gene.ucl.ac.uk/hugo/>

IAU, International Astronomical Union: Working group for planetary system nomenclature: <http://www.flag.wr.usgs.gov/USGSFlag/Space/nomen/nomen.html>

ICZN, International Commission on Zoological Nomenclature: <http://www.iczn.org/>

JCBN, IUPAC-IUBMB Joint Commission on Biochemical Nomenclature: <http://www.chem.qmw.ac.uk/iupac/jcfn/>

LocusLink: <http://www.ncbi.nlm.nih.gov/LocusLink/index.html>

Mammalian Genome: <http://link.springer.de/link/service/journals/00335/>

MGD nomenclature guidelines: <http://www.informatics.jax.org/mgihome/nomen/table.shtml>

MGD, Mouse Genome Database: <http://www.informatics.jax.org/>

Nature Genetics: <http://www.nature.com/ng/>

OMIM: <http://www3.ncbi.nlm.nih.gov/Omim/>

Pig database: <http://www.ri.bbsrc.ac.uk/cgi-bin/arkdb/browsers/browser.sh?species=pig>

PubMed: <http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=PubMed>

Radiation Research: <http://www.radres.org/>

Rat database: <http://ratmap.gen.gu.se/>

Sheep database: <http://www.ri.bbsrc.ac.uk/cgi-bin/arkdb/browsers/browser.sh?species=sheep>

Zebrafish database: <http://zfin.org/>

H. Wain, E. Bruford, A. Duncanson,
R. Lovering and S. Povey

The Galton Laboratory

Department of Biology

University College London

Wolfson House, 4, Stephenson Way

London NW1 2HE

Email nome@galton.ucl.ac.uk

REFERENCES

1. T. B. Shows, C. A. Alper, D. Bootsma, M. Dorf, T. Douglas, T. Huisman, S. Kit, H. P. Klinger, C. Kozak, P. A. Lalley, D. Lindsley, P. J. McAlpine, J. K. McDougall, P. Meera Khan, M. Meisler, N. E. Morton, J. M. Opitz, C. W. Partridge, R. Payne, T. H. Roderick, P. Rubinstein, F. H. Ruddle, M. Shaw, J. W. Spranger and K. Weiss, International system for human gene nomenclature (1979) ISGN (1979). *Cytogenet Cell Genet.* **25**, 96–116 (1979).
2. J. A. White, P. J. McAlpine, S. Antonarakis, H. Cann, J. T. Eppig, J. Frazer, J. Frezal, D. Lancet, J. Nahmias, P. Pearson, J. Peters, A. Scott, H. Scott, N. Spurr, C. Talbot, Jr. and S. Povey, Guidelines for human gene nomenclature (1997). HUGO Nomenclature Committee. *Genomics* **45**, 468–471 (1997).
3. L. J. Maltais, J. A. Blake, J. T. Eppig and M. T. Davisson, Rules and guidelines for mouse gene nomenclature: A condensed version. *Genomics* **45**, 471–476 (1997).