

Naming human genes

Gene nomenclature can be complicated, and the official naming of genes requires rational standards to avoid confusion and to maximize clarity. The HUGO Gene Nomenclature Committee has released updated guidelines for the naming of human genes, and we encourage the community to adopt these recommendations.

The advent of new genomics technologies has afforded us with unprecedented ability to accurately sequence human genomes, and cataloging the order of the A, T, G and C DNA nucleotides is mostly straightforward. However, assigning definitions and functions to genomic regions, and standardizing classifications and nomenclature can be fraught. Therefore, systematic guidelines are needed to help the community with the official naming of human genes to ensure that naming proceeds in a rational and sustainable manner.

The HUGO Gene Nomenclature Committee (HGNC) has prepared some updated guidelines for the naming of human genes (<https://www.genenames.org/>). A **Comment** in this issue summarizes the current practices of the HGNC and provides advice to researchers on the naming and reporting of novel genes. We endorse these guidelines and encourage the community to follow these suggestions when naming new human genes.

In addition to standardizing nomenclature for protein-coding genes, the HGNC also presents guidelines for

naming RNA genes and pseudogenes. Further recommendations for how to name non-coding RNAs, readthrough transcripts, gene segments, genomic regions and population-specific genes are also included. The **Comment** in this issue also provides helpful information about the main factors that HGNC considers when assigning gene nomenclature and discusses the specific conventions used in gene names (such as characters, letters and numbers).

For clarity and specificity, each gene is assigned only one symbol. Isoforms such as splice variants or alternate transcripts for the most part do not receive their own names.

We fully support these standards and recommendations. For new genes, authors are asked to contact the HGNC to approve any new symbols before publication. The HGNC notes that they do not consider publication precedence when approving nomenclature. They aim to minimize changes to gene symbols in the future.

Broadly, we endorse the following guidelines as well, and we ask authors to adhere to them in submissions. We ask that the approved gene symbols (<https://www.genenames.org/>) always be used. To avoid ambiguity, the HGNC ID (HGNC:number)

can be used; each gene has a unique identity associated with the sequence. We ask authors to use the approved gene symbol at least once in the manuscript's abstract. We also agree with the HGNC's guidance to use italics to denote genes, alleles and RNAs, to distinguish them from proteins.

Although naming human genes is complex, the naming of genes in different species also presents quite a challenge. There is also a Vertebrate Gene Nomenclature Committee (VGNC) in charge of determining frameworks to name genes from non-human organisms. One of their guiding principles is ensuring that human-specific traits are not referred to in the nomenclature, if possible.

We believe that defining the taxonomy of the genome should be approached with clear, practical guidelines that aim to avoid ambiguity and redundancy in gene nomenclature; thus, we encourage the community to abide by the HGNC recommendations for the naming and reporting of human genes. □

Published online: 3 August 2020
<https://doi.org/10.1038/s41588-020-0679-1>