

# Word for Word

## Where to find approved gene names online and how to use them



Many genes have multiple symbols and spellings; for example, the gene commonly known as *HER2* is sometimes called *HER-2*, *NEU*, *NGL*, *ERBB2*, or *CD340*. The researchers who discover a gene usually name it, and there are few standards for these names—they can reflect the gene’s molecular product, its function, or a disease associated with it.<sup>1</sup>

This ambiguity in gene symbols can make searching the scientific literature difficult and pose reproducibility challenges. So how do researchers and authors know which gene symbol to use? And how should gene names be written to ensure clarity and specificity?

The HUGO Gene Nomenclature Committee (HGNC) is the authority charged with approving and standardizing human gene nomenclature.<sup>1,2</sup> Its guidelines, originally issued in 1979, are

continually updated to reflect the emerging scientific understanding of the human genome, with the latest update issued in 2020.<sup>3,4</sup> The principle behind the HGNC guidelines is “one human genome—one gene language.”<sup>1,2</sup> A corollary of this principle is that each gene should have a single standard symbol.

When writing about human genes, authors should use the HGNC-approved symbols found at [genenames.org](https://www.genenames.org) (previously known as Entrez Gene). These symbols are usually 3 to 7 characters long; are written using italicized capital letters; do not include superscripts, subscripts, or punctuation; and do not contain Roman numerals or Greek letters.<sup>2</sup> Many journals, such as the *Nature*, *Lancet*, and *JAMA* families of journals, require use of HGNC-approved gene symbols.

**Examples:**

Correct: *ADRB2*  
Incorrect: *ADRβ2*

Correct: *KRAS*  
Incorrect: *K-Ras*

If you are writing about a gene that is widely known by a symbol other than its HGNC-approved one or a gene whose symbol has changed (found under “alias symbols” on [genenames.org](https://www.genenames.org)), you should use the currently approved symbol. However, for clarity, it is acceptable to provide a previous symbol or widely used alias symbol in parentheses at first use.<sup>1</sup>

**Examples:**

*ERBB2* (previously *HER2/neu*)  
*CD274* (also known as *PD-L1*)

HGNC-approved gene symbols may be used without expansion; that is, it's not necessary to spell out the full name of the gene.

**Examples:**

Acceptable use of gene symbol alone: *MYC* is constitutively expressed in several cancers.

Unnecessary use of gene name: MYC proto-oncogene, bHLH transcription factor (*MYC*) is constitutively expressed in several cancers.

Some journals also require that a gene identification number such as an **Online Mendelian Inheritance in Man (OMIM)**, **GenBank**, or **HGNC ID** should be provided in parentheses. Doing so can further help reduce ambiguity regarding gene symbols.

**Examples:**

*BRD1* (HGNC: 1102)

*CTLA4* (OMIM: 123890)

The symbols for nonhuman vertebrate genes are assigned by the **Vertebrate Gene Nomenclature Committee (VGNC)**. The symbols for vertebrate genes with human homologs generally are the same as the human symbols but are not written in all capitals. Specific guidelines and databases of approved gene symbols are available for **mouse**, **rat**, **Xenopus**, and **zebrafish** genes.<sup>2</sup>

**Examples:**

Human gene: *CYP1A2*

Mouse gene: *Cyp1a2*

## References

1. Human Gene Nomenclature. *AMA Manual of Style*, pp. 729-754.
2. Bruford EA, Braschi B, Denny P, Jones TEM, Seal RL, Tweedie S. Guidelines for human gene nomenclature. *Nat Genet.* 2020;52:754-758.
3. Naming human genes [editorial]. *Nat Genet.* 2020;52:751. doi: 10.1038/s41588-020-0679-1
4. HUGO Gene Nomenclature Committee. Current guidelines for naming human genes. Accessed August 18, 2021. <https://www.genenames.org/about/guidelines/#!/#tocAnchor-1-1>

—Amy Ninetto

Image Credit: © **Can Stock Photo / NatStocker**