

EDITORIAL

Open Access



Standardized nomenclature and open science in *Human Genomics*

Vasilis Vasiliou¹, Kirill Veselkov^{1,2}, Elspeth Bruford^{3,4} and Juergen K. V. Reichardt^{5*}

Two recent papers have highlighted the vital importance of using standardized nomenclature in reporting data, especially when this is of clinical relevance. Higgins et al. [1] have drawn attention to the crucial matter of the appropriate nomenclature of DNA variants in scientific publications. It is critical that DNA variants can be identified unambiguously. And Fujiyoshi et al. [2] have called for gene products to be referenced using the approved gene symbol for the encoding gene, along with an appropriate database ID (HGNC ID, with UniProt ID where required, see Table 1 for resources for vertebrate genes). Confusion can impede data sharing and scientific progress, as well as potentially result in patient harm.

Human Genomics has always required the use of standardized gene symbols [3], and we now ask all authors, editors, reviewers, etc. to utilize the correct and verified nomenclature additionally for gene products and DNA variants in all submissions to this journal. Adherence to this policy will ensure full understanding by all readers and reproducibility of findings involving genes, gene products, and DNA variants. The usage of historic nomenclature in addition to this policy may be helpful in some fields to assist certain readers. We note that other

journals have already taken a keen interest in these matters [4].

We also note here in the broader context that *Human Genomics* strongly encourages the sharing of data to facilitate open science (https://en.wikipedia.org/wiki/Open_science), reproducibility, and full understanding of scientific advances. Therefore, depositing all relevant omic information in general, such as genomic, epigenomic, transcriptomic, metabolomic, and proteomic data, would be of great value to the scientific community. For example, the open-access MetaboLights repository of raw experimental metabolomic data and associated metadata has been recently re-designed to facilitate the growing demand for reproducibility and integration with other “omics” [5]. The recently engineered auto-deconvolution MSHub/GNPS platform has further enabled the community to store, process, share, annotate, compare, quantify reproducibility, and perform molecular networking of mass spectrometry metabolomic data in the context of multi-omics studies [6]. Accordingly, we strongly encourage depositing all relevant omic data relating to publications in our journal to aid advancement and reproducibility in science (Table 2).

* Correspondence: juergen.reichardt@jcu.edu.au

⁵Australian Institute of Tropical Health and Medicine, James Cook University, Smithfield, Queensland 4878, Australia

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Table 1 Resources for standardized identifiers

Type	Name	URL	Comments
Gene symbols	HGNC (HUGO Gene Nomenclature Committee)	www.genenames.org	Human gene symbols and IDs
Gene symbols	MGI (Mouse Genome Informatics)	www.informatics.jax.org	Mouse gene symbols and IDs
Gene symbols	RGD (Rat Genome Database)	https://rgd.mcw.edu	Rat gene symbols and IDs
Gene symbols	AgBase	https://agbase.arizona.edu	Chicken gene symbols and IDs
Gene symbols	VGNC (Vertebrate Gene Nomenclature Committee)	vertebrate.genenames.org	Gene symbols and IDs for selected vertebrate species
Gene symbols	Xenbase	www.xenbase.org	Xenopus gene symbols and IDs
Gene symbols	Zfin	www.zfin.org	Zebrafish gene symbols and IDs
Protein identifiers	UniProt	www.uniprot.org	Protein IDs for multiple species

Table 2 Examples of relevant omic databases

Type	Name	URL	Comments
DNA Variants	dbSNP	https://www.ncbi.nlm.nih.gov/snp/	NCBI SNP (single nucleotide polymorphism) database
	ClinVar	https://www.ncbi.nlm.nih.gov/clinvar/	NCBI database on genomic variation and clinical phenotypes
	LOVD (Leiden Open Variation Database)	https://www.lovd.nl	Open database for genetic variants
Proteomic	Protein	https://www.ncbi.nlm.nih.gov/protein/	NCBI protein databases
Proteomic	Human Variants Database	https://www.iit.ac.in/bioinfo/huvarbase/	Human protein variants
Metabolomic	Metabolights	https://www.ebi.ac.uk/metabolights/	Open database of metabolomic experiments, derived metabolites, their structures and biological roles
Metabolomic	GNPS	https://gnps.ucsd.edu	Mass spectrometry driven small molecule analysis

Authors' contributions

The authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Environmental Health Sciences, Yale School of Public Health, New Haven, CT 06510, USA. ²Department of Surgery and Cancer, Imperial College, London SW7 2AZ, UK. ³HUGO Gene Nomenclature Committee (HGNC), EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK. ⁴Department of Haematology, University of Cambridge, Cambridge CB2 0XY, UK. ⁵Australian Institute of Tropical Health and Medicine, James Cook University, Smithfield, Queensland 4878, Australia.

Published online: 22 February 2021

References

- Higgins J, Dagleish R, den Dunnen JT, et al. Verifying nomenclature of DNA variants in submitted manuscripts: guidance for journals. *Hum Mutat.* 2021; 42:3–7. <https://doi.org/10.1002/humu.24144>.
- Fujiyoshi K, Bruford EA, Mroz P, et al. Standardizing gene product nomenclature – a call to action. *PNAS.* 2021;118:e2025207118 <https://www.pnas.org/content/118/3/e2025207118>.
- Bruford EA, Braschi B, Denny P, et al. Guidelines for human gene nomenclature. *Nat Genet.* 2020;52:754–8 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7494048/>.
- Dengler VL. Committee presents guidelines for variant reporting. *Genetics in Medicine.* 2021;23:2.
- Haug K, Cochrane K, Nainala VC, et al. MetaboLights: a resource evolving in response to the needs of its scientific community. *Nucleic Acids Res.* 2020; 48:D440–4. <https://doi.org/10.1093/nar/gkz1019>.
- Aksenov AA, Laponogov I, Zhang Z, et al. Auto-deconvolution and molecular networking of gas chromatography–mass spectrometry data. *Nat Biotechnol.* 2020. <https://doi.org/10.1038/s41587-020-0700-3>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.