

CHLORELLA should not be a long noncoding RNA: it's already an alga!

David Roy Smith , Breno Dupin , Matheus Sanita Lima *

Department of Biology, Western University, London, ON N6A 5B7, Canada

*Corresponding author. Department of Biology University of Western Ontario, 1151 Richmond Street, London, ON N6A 5B7, Canada. E-mail: msanital@uwo.ca

Keywords *Chlorella*, lncRNA, *Arabidopsis*, gene name, species name, chloroplast

The only thing worse than bad health is a bad name. [1] — Gabriel Garcia Marquez

We read with enthusiasm the recent article by Kang *et al.* [2] describing an *Arabidopsis* nuclear-encoded, chloroplast-targeted long non-coding RNA (lncRNA) involved in chloroplast regulation. This lncRNA appears to act as a developmental switch within chloroplasts, mediating the transition from early biogenesis to late degeneration during leaf ageing [2]. These data are exciting because they show that anterograde signalling between the nucleus and chloroplast is more complicated and nuanced than previously thought [3].

However, one aspect of this novel *Arabidopsis* noncoding RNA struck us as odd: its name. The authors are calling it *CHLORELLA*, short for *CHLOROPLAST-RELATED LONG NONCODING RNA*. At first glance, this name sounds fine. Except for one important detail—it is already taken! As many microbial ecologists are aware, there is a genus of green alga called *Chlorella*. To be fair, *Chlorella* is not as famous as the research superstar *Arabidopsis*. But *Chlorella* species have been model research organisms for over a century [4], and once were the ‘workhorses’ of chloroplast studies along with other micro algae and crop species [5]. Melvin Calvin’s work on chloroplast carbon dioxide assimilation [6], for example, resulted in a Nobel prize in 1961. Nobel laureate Otto Heinrich Warburg (1931) also used *Chlorella* as his primary model for investigating the mechanisms of photosynthesis [7]. More recently, *Chlorella* species have taken centre stage in microbial ecology studies (published in ISME) investigating long-term symbioses [8] and interspecies competition [9].

We believe naming an important land plant ncRNA after a model green alga is not only confusing but counterproductive to research on microbial eukaryotes. *CHLORELLA* has not yet been identified in *Chlorella*, or any other green alga. But *Chlorella* has a chloroplast, so there is a chance it might have *CHLORELLA*; it would be even more perplexing if *Chlorella* lacked *CHLORELLA*. You get the point—gene nomenclature is a serious issue in the omics era [10].

To the best of our knowledge, *CHLORELLA* is the first published case of a name collision between an ncRNA and a valid genus name. As other cautionary tales from the media [11, 12] and primary literature can attest [13], the lack of standardization in gene nomenclature and naming practices are an obstacle to comparative omics studies. Indeed, a poorly chosen name can confound the evolutionary history of a gene (or any molecular locus) [13], as is epitomized by the name *CHLORELLA*. Moreover, unconventional gene names hinder text-mining efforts for ecological and biomedical research [14]. It is interesting to ponder how large language models will treat *CHLORELLA* the ncRNA when its semantic network is tied up with *Chlorella* the alga (and vice versa). The English linguist J. R. Firth said it best: ‘You shall know a word by the company it keeps’ [15].

Gene naming practices are of particular importance to biologists, including microbial ecologists, tackling research questions through a comparative approach. Just recently, for our own comparative genomics endeavours, we had to develop a bionformatics ‘gene dictionary tool’ (GDT for short) to address inconsistencies in gene names [16]. This is precisely why international consortia and committees, such as the Human Gene Nomenclature Committee (HGNC), have issued clear gene nomenclature guidelines. The HGNC stresses that ‘gene symbols should not spell out commonly used words’ [17], and it considers ‘[gene] symbols that affect data handling and retrieval’ one of the few cases for which gene symbols might be suitable for a change [18]. For instance, the gene names SEPT1 and MARCH1 had to be altered because they could be mistaken for dates in conventional spreadsheets. The genes WARS and CARS also had to be renamed because of their similarity to common English words. And the list goes on. Perhaps the most notorious examples of ill-advised gene names are Pokémon, Velcro, and lunatic fringe [11, 12], all of which had to be changed, in some cases under the threat of litigation. Name collision is also a great concern for the naming of species. In fact, how to handle and avoid nomenclatural confusion (greatly derived from homonyms) is one of the core goals of the International Code of Nomenclature for algae, fungi, and plants [19].

Received: 9 March 2026. **Revised:** 21 April 2026. **Accepted:** 23 April 2026

© The Author(s) 2026. Published by Oxford University Press on behalf of the International Society for Microbial Ecology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Regarding lncRNAs specifically, Mattick *et al.* [20] wrote that ‘unless a lncRNA is antisense to a protein-coding gene, we recommend naming lncRNAs for their own sake with allusion to a discerned characteristic or function (as has been traditional for proteins) (...) If no biological context is available, we recommend naming the lncRNA according to the GENCODE system’. Pointedly, TAIR (The Arabidopsis Information Resource) states that: ‘Gene names should convey some meaning as to the function of the gene product’ [21]. As it stands, the naming of *CHLORELLA* does not follow these recommendations.

Chlorella is a kind-hearted phototroph and not litigious by nature, so we don’t think Kang *et al.* [2] have much to worry about. But this freshwater alga has had a hard time of late apropos its name. It’s a touch overused. Walk into any health food store and you’ll be bombarded with wellness products boasting the word *Chlorella* in their titles, from vitamins to creams to shampoos. There are also lines of *chlorella* clothing and sneakers. And some ‘health’ gurus are even promoting *Chlorella* injections. Given all this, Kang *et al.* [2] should consider renaming this fascinating functional noncoding RNA to something more redolent of *Arabidopsis* or land plants in general, following the guidelines quoted above. In no way, however, do we want this small criticism to detract from the work of Kang *et al.* [2]. In a recent article, we speculated how lncRNAs could participate in anterograde/retrograde signalling and even organelle nucleoid topology [22]. In other words, it would be an understatement to say we have been eagerly awaiting the identification of a bona fide functional noncoding RNA in a chloroplast system. Thankfully one has arrived. But to be clear: it is not an alga.

Acknowledgements

The authors wish to acknowledge the painstaking work and unabated commitment of researchers, such as Kang *et al.* 2025, who continue to unravel the mysteries and wonders of plant noncoding RNAs.

Conflicts of interest

None declared.

Funding

D.R.S. is supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Data availability

Not applicable.

References

- Garcia MG. *Love in the Time of Cholera*. New York: Vintage International, 2007.
- MH Kang, J Lee, J Kim, HB Mohammad, J Park, HJ Jung *et al.* The chloroplast-targeted long noncoding RNA *CHLORELLA* mediates chloroplast functional transition across leaf ageing via anterograde signalling. *Nat Plants* 2025;11:2217–29 <https://doi.org/10.1038/s41477-025-02129-z>
- Woodson J, Chory J. Coordination of gene expression between organellar and nuclear genomes. *Nat Rev Genet* 2008;9:383–95. <https://doi.org/10.1038/nrg2348>
- Kutschera U, Pieruschka R, Farmer S *et al.* The Warburg-effects: basic metabolic processes with reference to cancer development and global photosynthesis. *Plant Signal Behav* 2020;15:1776477. <https://doi.org/10.1080/15592324.2020.1776477>
- Benson AA, Calvin M. The path of carbon in photosynthesis: VII respiration and photosynthesis. *J Exp Bot* 1950;1:63–8. <https://doi.org/10.1093/jxb/1.1.63>
- Calvin M. The path of carbon in photosynthesis. *Science* 1962;135:879–89. <https://doi.org/10.1126/science.135.3507.879>
- Warburg O. Über die Geschwindigkeit der photochemischen Kohlensäurezersetzung in lebenden Zellen. In: Warburg O. (ed.), *Über die Katalytischen Wirkungen der Lebendigen Substanz*. Berlin, Heidelberg: Springer, 1928, 308–40.
- He M, Wang J, Fan X *et al.* Genetic basis for the establishment of endosymbiosis in *Paramecium*. *ISME J* 2019;13:1360–9. <https://doi.org/10.1038/s41396-018-0341-4>
- Song H, Lavoie M, Fan X *et al.* Allelopathic interactions of linoleic acid and nitric oxide increase the competitive ability of *Microcystis aeruginosa*. *ISME J* 2017;11:1865–76. <https://doi.org/10.1038/ismej.2017.45>
- Seringhaus MR, Cayting PD, Gerstein MB. Uncovering trends in gene naming. *Genome Biol* 2008;9:401. <https://doi.org/10.1186/gb-2008-9-1-401>
- Simonite T. Pokémon blocks gene name. *Nature* 2005;438:897. <https://doi.org/10.1038/438897a>
- Hopkin M. Troublesome gene names get the boot. *Nature* 2006. <https://www.nature.com/news/2006/061106/full/news061106-2.html>
- Theofanopoulou C, Gedman G, Cahill JA *et al.* Universal nomenclature for oxytocin–vasotocin ligand and receptor families. *Nature* 2021;592:747–55. <https://doi.org/10.1038/s41586-020-03040-7>
- Chen L, Liu H, Friedman C. Gene name ambiguity of eukaryotic nomenclatures. *Bioinformatics* 2005;21:248–56. <https://doi.org/10.1093/bioinformatics/bth496>
- Firth JR. A synopsis of linguistic theory 1930–1955. In: Firth J.R., Palmer F.R. (eds.), *Studies in Linguistic Analysis*. London: Longman, 1952, 1–32.
- Dupin B, Sanita Lima M, Paschoal AR *et al.* Protocol for creating a gene dictionary for organelle genomes using the Gene Dictionary Tool. *STAR Protoc* 2025;6:104187. <https://doi.org/10.1016/j.xpro.2025.104187>
- Wright MW. A short guide to long non-coding RNA gene nomenclature. *Hum Genomics* 2014;8:7. <https://doi.org/10.1186/1479-7364-8-7>
- Bruford EA, Braschi B, Denny P *et al.* Guidelines for human gene nomenclature. *Nat Genet* 2020;52:754–8. <https://doi.org/10.1038/s41588-020-0669-3>
- The International Code of Nomenclature for algae, fungi, and plants (Madrid Code). https://www.iaptglobal.org/_functions/code/madrid (17 April 2026, date last accessed).
- Mattick JS, Amaral PP, Carninci P *et al.* Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat Rev Mol Cell Biol* 2023;24:430–47. <https://doi.org/10.1038/s41580-022-00566-8>
- The Arabidopsis Information Resource. <https://phoenixbioinformatics.atlassian.net/wiki/spaces/COM/pages/42215953/Nomenclature> (20 February 2026, date last accessed).
- Sanita Lima M, Paschoal AR, Silva Domingues D *et al.* Pervasive transcription of plant organelle genomes: functional noncoding transcriptomes? *Trends Plant Sci* 2024;29:626–9. <https://doi.org/10.1016/j.tplants.2024.01.004>