"Loget" - a Uniform Differential Expression Unit to Replace "logFC" and "log2FC"

Alicja Pacholewska

Center for non-coding RNA in Technology and Health, University of Copenhagen

✉ **Correspondence**
a.pacholewska@gmail.com

Triple Blind Peer Review
The handling editor, the reviewers, and the authors are all blinded during the review process.

Full Open Access
Supported by the Velux Foundation, the University of Zurich, and the EPFL School of Life Sciences.

Abstract

Due to an inconsistency in the presentation of differential expression results in scientific articles, there is a need for a common and perspicuous way to report and discuss the differences in gene/protein expression levels, as well as metabolite levels. I suggest here a new term, a "loget" as a new differential expression unit that can be used in many of the "omics" fields.

Introduction

While reviewing yet another paper where changes in the expression levels were indicated incorrectly, I realized that it is not easy to retrieve the correct definition of fold change via commonly used web search engines. So what is a "change"? Change is a difference between one measure and another. A gene for example, may be up- or down-regulated in condition B versus A, and therefore, an increase or decrease in its expression level is observed (level in B minus level in A). A simple change is not very useful for genes, as their expression varies from lowly to highly expressed. A "fold change" helps to overcome this problem. It denotes how many times the expression increased or decreased, i.e. expression level in condition B divided by expression level in condition A, also known as "ratio". An increase has therefore, values above 1 and a decrease between 0 and 1. To better visualize the fold changes they are often shown on a logarithmic scale, so that a decrease is as visible as an increase. The most commonly used logarithm in differential expression analysis is the logarithm to the base 2 and its positive values denote an increase, and negative a decrease (0 stands for no change).

Objective

The objective of this observation was to unify the description of the differential expression among studies. A new term will not only uniform the results presented in scientific articles, but will also make it easier to communicate and discuss the results.
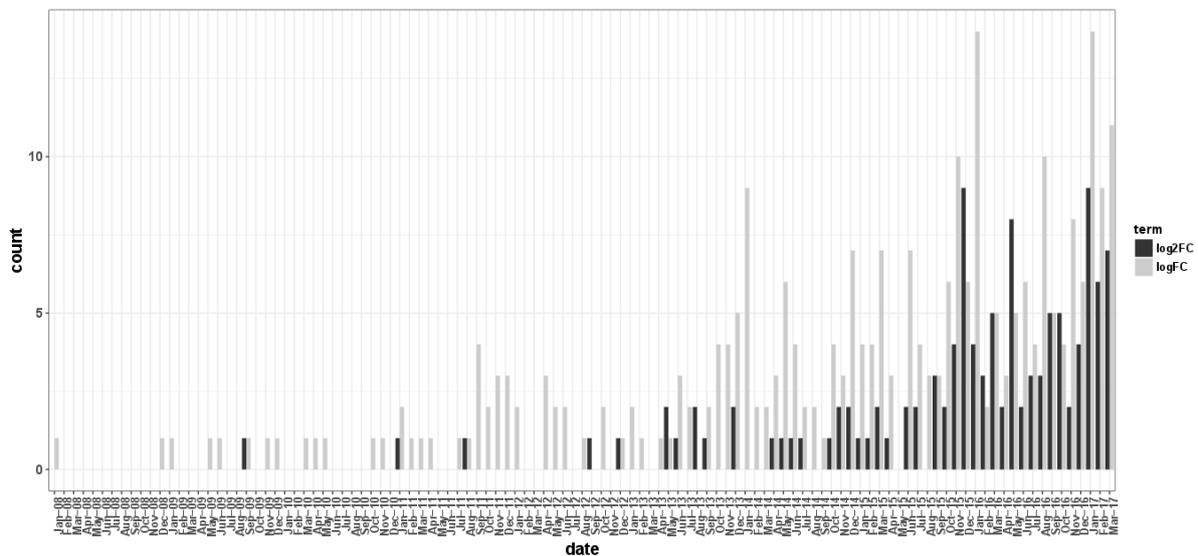
a

Figure Legend

Figure 1. The number of articles using the "logFC" or "log2FC" terms over time. Based on Springer Exemplar.

(http://www.springerexemplar.com).

Results & Discussion

A "log" term is usually referred to a natural logarithm, and the differential expression is usually calculated as a binary logarithm [1] [2] [3] [4].

Based on 11,472,063 documents collected by the Springer Exemplar, authors' tend to use more logFC abbreviation than log2FC, even today (Fig. 1). One reason for that is the use of "logFC" term by the authors in the manuals and output files generated by the most popular tools for the differential expression analysis [1] [4].

To overcome this problem, I suggest using a new term, "loget", that will always refer to the same unit.

A "loget" is calculated as follows:

loget = log2 ((level in test) / (level in reference))

Moreover, the "loget" term is easier to pronounce than the "logarithm two of fold change". Instead of always referring to the "log two fold change" or "binary logarithm of fold change", one could simply use the "loget" word. For example, instead of saying "There were 200 up-regulated genes with the log two-fold change range from 2 to 5" one could say "There were 200 up-regulated genes with the loget range from 2 to 5".

Conclusions

The "loget" term is a useful alternative to a rather long expression of "binary logarithm of fold change". As it is well defined, it may help to avoid often incorrectly used abbreviations in the description of gene or protein expression results. The use of this term will also help authors to verbally discuss their transcriptomics, proteomics, or other results. Instead of always referring to the "log 2 fold change" one could simply say "loget".

Additional Information

Methods and Supplementary Material

Please see https://sciencematters.io/articles/201706000011.

Ethics Statement

Not Applicable.

Citations

[1] Robinson et al. "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data". In: *Bioinformatics* 26.1 (Nov. 2010), pp. 139–140. DOI: 10.1093/bioinformatics/btp616. URL: https://doi.org/10.1093/bioinformatics/btp616.

[2] Simon et al. "Differential expression analysis for sequence count data". In: *Genome Biology* 11.10 (2010), R106. DOI: 10.1186/gb-2010-11-10-r106. URL: https://doi.org/10.1186/gb-2010-11-10-r106.

[3] Love Michael I, Huber Wolfgang, and Anders Simon. "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2". In: *Genome Biology* 15.12 (Dec. 2014), p. 550. DOI: 10.1186/s13059-014-0550-8. URL: https://doi.org/10.1186/s13059-014-0550-8.

[4] Ritchie M. E. et al. "limma powers differential expression analyses for RNA-sequencing and microarray studies". In: *Nucleic Acids Research* 43.7 (Jan. 2015), e47–e47. DOI: 10.1093/nar/gkv007. URL: https://doi.org/10.1093/nar/gkv007.

[5] Ross et al. "R: A Language for Data Analysis and Graphics". In: *Journal of Computational and Graphical Statistics* 5.3 (Sept. 1996), pp. 299–314. DOI: 10.2307/1390807. URL: https://doi.org/10.2307/1390807.

[6] Wickham Hadley. "ggplot2: Elegant Graphics for Data Analysis". In: *USE R! and Springer-Verlag New York* VIII (2009), p. 213. DOI: 10.1007/978-0-387-98141-3. URL: https://doi.org/10.1007/978-0-387-98141-3.