

AUTHOR CORRECTION

Open Access



# Author Correction: scDesign2: a transparent simulator that generates high-fidelity single-cell gene expression count data with gene correlations captured

Tianyi Sun<sup>1</sup>, Dongyuan Song<sup>2</sup>, Wei Vivian Li<sup>3\*</sup> and Jingyi Jessica Li<sup>1,4,5,6\*</sup>

The original article can be found online at <https://doi.org/10.1186/s13059-021-02367-2>.

\*Correspondence: [vivian.li@rutgers.edu](mailto:vivian.li@rutgers.edu); [jli@stat.ucla.edu](mailto:jli@stat.ucla.edu); [ljiy03@g.ucla.edu](mailto:ljiy03@g.ucla.edu)

<sup>1</sup> Department of Statistics, University of California, 90095-1554 Los Angeles, CA, USA

<sup>2</sup> Interdepartmental Program of Bioinformatics, University of California, 90095-7246 Los Angeles, CA, USA

<sup>3</sup> Department of Biostatistics and Epidemiology, Rutgers School of Public Health, 08854 Piscataway, NJ, USA

<sup>4</sup> Department of Human Genetics, University of California, 90095-7088 Los Angeles, CA, USA

<sup>5</sup> Department of Computational Medicine, University of California, 90095-1766 Los Angeles, CA, USA

<sup>6</sup> Department of Biostatistics, University of California, 90095-1772 Los Angeles, CA, USA

Following the publication of the original paper [1], the authors identified two errors in the notation and formulas in the Methods section.

identifiable [111, 112]. To resolve this unidentifiability issue, scDesign2 uses the technique of distributional transform [113]: first draw  $V_{ij} \sim \text{Uniform}[0,1]$  independently for  $i=1, \dots, p$  and  $j=1, \dots, n$ ; second define  $U_{ij}$  as:

$$U_{ij} = V_{ij}F_i(X_{ij} - 1) + (1 - V_{ij})F_i(X_{ij}). \quad (1)$$

Equation (1) in the online version.

The first error is the  $U_{ij}$  expression. Specifically, Eq. (1) should be corrected as

$$U_{ij} = (1 - V_{ij})F_i(X_{ij} - 1) + V_{ij}F_i(X_{ij}). \quad (1)$$

Hence,  $\tilde{X}_{ij} \sim \tilde{F}_i$ ; that is, the continuous random variable  $\tilde{X}_{ij}$  constructed from  $X_{ij}$  and  $V_{ij}$  follows  $\tilde{F}_i$ . Defining  $U_{ij} = \tilde{F}_i(\tilde{X}_{ij})$ , we have  $U_{ij} \sim \text{Uniform}[0,1]$  and:

$$\begin{aligned} U_{ij} &= \tilde{F}_i(X_{ij} + V_{ij}) = F_i(X_{ij} - 1) + V_{ij} (F_i(X_{ij}) - F_i(X_{ij} - 1)) \\ &= V_{ij}F_i(X_{ij} - 1) + (1 - V_{ij})F_i(X_{ij}), \end{aligned}$$

The end of the derivation of Eq. (1) in the online version.



© The Author(s) 2023. **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.

Accordingly, at the end of the derivation of

$$\begin{aligned} U_{ij} &= \tilde{F}_i(X_{ij} + V_{ij}) = F_i(X_{ij} - 1) + V_{ij}(F_i(X_{ij}) - F_i(X_{ij} - 1)) \\ &= V_{ij}F_i(X_{ij} - 1) + (1 - V_{ij})F_i(X_{ij}), \end{aligned}$$

The last line should be changed to

$$\begin{aligned} U_{ij} &= \dots \\ &= (1 - V_{ij})F_i(X_{ij} - 1) + V_{ij}F_i(X_{ij}), \end{aligned}$$

The second error is one notation in the sentence just before the above-corrected expression. In “Hence,  $\tilde{X}_{ij} \sim \tilde{F}_i$ ; that is, the continuous random variable  $\tilde{X}_{ij}$  constructed from  $X_{ij}$  and  $V_{ij}$  follows  $\tilde{F}_i$ ”, at the end,  $\tilde{F}_i$  should be corrected as  $\tilde{F}_i$ .

The authors confirm that the notational errors corrected here had no effect on the conclusions or the performance of the method scDesign2.

Published online: 22 February 2023

#### Reference

1. Sun T, Song D, Li WV, et al. scDesign2: a transparent simulator that generates high-fidelity single-cell gene expression count data with gene correlations captured. *Genome Biol.* 2021;22:163. <https://doi.org/10.1186/s13059-021-02367-2>.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

