Erratum: SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler

Ruibang Luo[1,2†], Binghang Liu[1,2†], Yinlong Xie[1,2,3†], Zhenyu Li[1,2†], Weihua Huang[1], Jianying Yuan[1], Guangzhu He[1], Yanxiang Chen[1], Qi Pan[1], Yunjie Liu[1], Jingbo Tang[1], Gengxiong Wu[1], Hao Zhang[1], Yujian Shi[1], Yong Liu[1], Chang Yu[1], Bo Wang[1], Yao Lu[1], Changlei Han[1], David W. Cheung[2], Siu-Ming Yiu[2], Shaoliang Peng[4], Zhu Xiaoqian[4], Guangming Liu[4], Xiangke Liao[4], Yingrui Li[1,2], Huanming Yang[1], Jian Wang[1], Tak-Wah Lam[2*] and Jun Wang[1*]

Erratum

After the publication of 'SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* 2012 1:18' [1] this paper was used as case study by the ISA, Research Object and Nanopublication communities to use their data models to try to quantitatively assess the reproducibility of published research. While the supporting data and overall results were found to be accurate, a number of semantic problems were detected in a few points of interpretation, and in order to ensure the accuracy of the scientific we now outline these in this correction article.

1. While there are huge improvements to the quality of the resulting assemblies, other than the tables it was not stressed in the text that the speed of SOAPdenovo2 can be slightly slower than SOAPdenovo v1.
2. In the testing an assessment section (page 3), based on the correct results in table 1, where we say the scaffold N50 metric is an order of magnitude longer from SOAPdenovo2 versus SOAPdenovo1, this was actually 45 times longer.
3. Also in the testing an assessment section, based on the correct results in table 1, where we say SOAPdenovo2 produced a contig N50 1.53 times longer than ALLPATHS-LG, this should be 2.18 times longer.

4. Finally in this section, where we say the correct assembly length produced by SOAPdenovo2 was 3–80 fold longer than SOAPdenovo1, this should be 3–64 fold longer.

To see more about how these errors were detected see the case study article [2]. To access all the supporting materials, ISA metadata and pipelines implemented in Galaxy to test these results on your own, please see the supporting entry in GigaDB [3].

Author details

[1]BGI HK Research Institute, 16 Dai Fu Street, Tai Po, Industrial Estate, Tai Po, Hong Kong. [2]HKU-BGI Bioinformatics Algorithms and Core Technology Research Laboratory & Department of Computer Science, University of Hong Kong, Pok Fu Lam, Hong Kong. [3]School of Bioscience and Bioengineering, South China University of Technology, Guangzhou 510006, China. [4]School of Computer Science, National University of Defense Technology, No.47, Yanwachi street, Kaifu District, Changsha 410073, Hunan, China.

References

1. Luo R et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience. 2012;1:18.
2. González-Beltrán A, Li P, Zhao J, Avila-Garcia MS, Roos M, Thompson M, et al. (2015) From Peer-Reviewed to Peer-Reproduced in Scholarly Publishing: The Complementary Roles of Data Models and Workflows in Bioinformatics. PLoS ONE 10(7): e0127612. doi:10.1371/journal.pone.0127612.
3. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung D, Yiu SM, Liu G, Zhu X, Peng S, Li Y, Yang H, Wang J, Lam TW, Wang J: Software and supporting material for "SOAPdenovo2: an empirically improved memory-efficient short read de novo assembly". http://dx.doi.org/10.5524/100044.

* Correspondence: twlam@cs.hku.hk; wangj@genomics.org.cn
†Equal contributors
[2]HKU-BGI Bioinformatics Algorithms and Core Technology Research Laboratory & Department of Computer Science, University of Hong Kong, Pok Fu Lam, Hong Kong
[1]BGI HK Research Institute, 16 Dai Fu Street, Tai Po, Industrial Estate, Tai Po, Hong Kong
Full list of author information is available at the end of the article