

Ivan G. Costa, PhD
Pauwelsstrasse 19 • 52074 Aachen, Germany
Tel.: +49-241-80 80270
www: <http://www.costalab.org>

**IZKF Research Group Bioinformatics
Computational Biology Division**

Institute for Biomedical Engineering
RWTH Aachen University Medical School

Aachen Institute for Advanced Study in
Computational Engineering Science
RWTH Aachen University

Dear Dr. Rusk,

Thank you very much for your email on January 23, 2016 regarding our paper by Gusmao et al., "Analysis of computational footprinting methods for DNase sequencing experiments" (NMEMH-AS24264B-Z). We would like to thank you and the Reviewers for the thorough review of our paper and the constructive comments.

Reviewer 3

“Overall, this manuscript is acceptable once terminology in the manuscript is changed to clarify some key points.

Through-out the manuscript the authors use the term "experimental bias" to mean difference between Dnase-seq experiments due to local chromatin conformation, sequence specificity, and fragmentation. The fact that the authors call chromatin conformation a bias is truly troubling to this reviewer, since the assay's goal to measure chromatin accessibility. The authors need to change this term, because it is misleading and incorrect. I suggest they use "footprint normalization" or "experimental normalization" to describe their approach. Normalization is exactly what they are trying to perform by correcting for the different experimental conditions (one-hit vs two-hit, DNaseI concentrations) and for DNaseI sequence specificity.”

We agree with the reviewer that term “experimental bias” was misleading. We have also removed from the text any reference of “local chromatin conformation” as being an “experimental artifact”. We think that the term “experimental normalization” is too abstract. We have decided to use the term “sequence bias” as this better describes what is being measured. We have renamed our strategies as following: (1) sequence bias estimated on DHS regions as proposed by He et al., 2014, which we call “DHS sequence bias”; and (2) sequence bias estimated on naked DNA experiments, which we call “naked DNA sequence bias”.

We have also incorporated in the manuscript all changes requested by the Editor.

- 1) We now refer to the tag count statistics to TC and to the baseline method, which is based on ranking of tag count statistics as “TC-Rank”. Similarly, we call now other baseline methods as FS-Rank and PWM-Rank.

- 2) We have changed all text passages to indicate that we evaluate 10 computational footprinting methods. The discrepancies in numbers arise from the fact that our final evaluation also included baseline approaches and HINT method variants.
- 3) We have improved the first paragraph of **Section Results – Impact of experimental artifacts** to reflect yours and referee 3 requests.

Modified text passages have been marked in the text. We should be happy to answer any further question, which you might have.

With best regards,

A handwritten signature in blue ink, appearing to read 'I. G. Costa'.

Ivan G. Costa
Computational Biology Division
Institute for Biomedical Engineering
RWTH Aachen University