

# Yale University

*MB&B*  
260/266 Whitney Avenue  
PO Box 208114  
New Haven, CT 06520-8114

*Telephone:*  
203 432 6105  
360 838 7861 (fax)  
Mark.Gerstein@yale.edu  
<http://bioinfo.mbb.yale.edu>

10 February 2016

Rafal Marszalek, Ph.D.

Senior Editor

*Genome Biology*

236 Gray's Inn Road

London WC1X 8HB

United Kingdom

Dear Dr. Marszalek,

Please find enclosed our revised manuscript entitled “The real cost of sequencing: linking data generation, analysis, and interpretation”, a solicited opinion piece for Genome Biology. We have incorporated the recommended revisions from the reviewer and the Genome Biology editorial board.

The editors asked that we “discuss the cost of the assembly in more detail. The referee also raises the issue of long reads and how their use will shift the cost from mapping to the assembly part of the genomic pipeline. We feel this is an important point that should be acknowledged.”

With regard to the reviewer’s comment on genome assembly as well as long reads we thank the reviewer. We have added a more detailed discussion on the computational costs associated with assembly as well as the impact of long reads on assembly algorithms and the sequencing depth required for genome assembly. Additionally, we have incorporated a new panel in figure 3 that illustrates the running time and memory requirements for a set of assembly algorithms applied to genome of different sizes.

The editors asked that we “discuss the full economic cost at more length. The referee noted as well that the role of bioinformaticians and their inclusion in the cost considerations should be included. We think this discussion should fall in the ‘full economic cost’ section, and ask that you expand the manuscript to include it.”

We thank the reviewer for their comment on expanding the analysis of sequencing costs to include a full economic costing. We have added a discussion of the full economic costing of sequencing projects and the importance of better incorporating bioinformatics costs for developing an accurate cost metric of sequencing projects. We also call attention to the fact that often this is not the price commonly associated with sequencing a genome. Furthermore, we have developed a new figure (figure 2) in which we break down the cost

of sequencing into the relative contributions of various components such as labor, reagents, indirect costs, etc. for both exome and whole genome sequencing.

Yours sincerely,

Mark Gerstein  
Albert L. Williams Professor  
of Biomedical Informatics