Yale University

Bass Building, Rm 432A
260/266 Whitney Avenue
PO Box 208114
New Haven, CT 06520-8114
Telephone: 203 432 6105
Fax: 360 838 7861
Mark.Gerstein@yale.edu

Dec 17, 2014

To the editors of Nucleic Acids Research:

I am writing to submit our manuscript, titled "LARVA: an integrative framework for Large-scale Analysis of Recurrent Variants in noncoding Annotations", by Lucas Lochovsky, Jing Zhang, Yao Fu, Ekta Khurana, and Mark Gerstein, for the exclusive consideration of publication in Nucleic Acids Research.

In this paper, we introduce a new model for background somatic mutation accumulation in cancer genomes. Accurately modeling the background somatic mutation rate is important for separating those mutations that drive cancer progression from those that have arisen due to background mutation processes. Previous models assumed a constant background mutation rate, which implied a binomial distribution for somatic variants in cancer. By contrast, our model utilizes a variable background mutation rate, represented by a beta distribution. This gives rise to a beta-binomial distribution for the somatic variants counts in a specific region in cancer, and we demonstrate that this model vastly improves on the false positive rate of the previous models based on the binomial distribution. This represents a valuable contribution to the study of cancer genomes, and the process of finding the genetic basis of cancer.

Jing Zhang 1/19/2015 12:02 PM
Deleted: with

Jing Zhang 1/19/2015 12:03 PM
Deleted: spread of

We have prepared a software implementation of this method for identifying highly mutated genomic elements indicated by the intersection of whole genome cancer variant calls with noncoding annotations. Our goal is to address the lack of noncoding cancer mutation studies relative to coding mutation studies by facilitating noncoding mutation analysis. Such analyses offer insight into the regulatory disruptions brought about by cancer. We gathered whole genome variant calls for 760 cancer samples, and used our software to find highly mutated promoters, untranslated regions (UTRs), transcription start sites (TSSs), transcription factor peaks, and Dnase I hypersensitive sites, among others. Our software identified a number of previously established cancer drivers, such as TP53 and the TERT promoters, as highly mutated. Our results also indicated that some genes with few exon mutations are associated with regulatory elements that have a high number of mutations, suggesting that genes that may have been overlooked in coding mutation analyses may in fact be important when noncoding mutations are considered.

Jing Zhang 1/19/2015 12:03 PM
Deleted: recurrently

In addition, we directly criticize the paper by Weinhold *et. al* (Weinhold, N., Jacobsen, A., Schultz, N., Sander, C. and Lee, W. (2014) Genome-wide analysis of noncoding regulatory mutations in cancer. *Nature genetics*, **46**, 1160-1165). That paper used a

Jing Zhang 1/19/2015 12:04 PM
Deleted: In this paper

Jing Zhang 1/19/2015 12:10 PM
Formatted: Font:Italic

binomial distribution model by assuming a constant mutation rate, which we claimed might be problematic when there is huge overdispersion in the mutation count data. Consequently, we suggest that Dr. Sander and Dr. Lee should be avoided during the revision process of LARVA since they will obviously have a biased view of our manuscript.

We appreciate you taking the time to review and respond to our manuscript. Please address all correspondence concerning this manuscript to pi@gersteinlab.org. We also provide a list of suggested referees below.

Sincerely,

Mark B Gerstein
Albert L Williams Professor of Biomedical Informatics, Molecular Biophysics & Biochemistry, and Computer Science
Yale University

Jing Zhang 1/19/2015 12:11 PM
Deleted: , -> result

Jing Zhang 1/19/2015 12:12 PM
Deleted: .

Jing Zhang 1/19/2015 12:12 PM
Deleted: Consequenty

Jing Zhang 1/19/2015 12:14 PM
Deleted: suggest

Jing Zhang 1/19/2015 12:13 PM
Deleted: the

Jing Zhang 1/19/2015 12:13 PM
Deleted: (Make it strong point here)

Suggested referees:

Please note that we would like to exclude Dr. Chris Sander, Dr. William Lee, and Dr. Manolis Kellis as referees.

We would like to invite Dr. Gad Getz, Dr. Josh Stuart, and Dr. Haiyuan Yu as referees of our manuscript.

Lucas Lochovsky 1/19/2015 5:33 PM
Deleted: ... [1]

Lucas Lochovsky 1/19/2015 5:33 PM
Deleted:

Jing Zhang 1/19/2015 12:15 PM
Deleted: Peter Campbell and Michael Stratton

Jing Zhang 1/19/2015 12:16 PM
Formatted: Normal, Font Alignment: Auto

Jing Zhang 1/19/2015 12:16 PM
Formatted: Font:(Default) +Theme Body, 12 pt, Font color: Auto

Jing Zhang 1/19/2015 12:17 PM
Formatted: Font:(Default) +Theme Body, 12 pt, Font color: Auto, Pattern: Clear

Jing Zhang 1/19/2015 12:17 PM
Formatted: Font:

Jing Zhang 1/19/2015 12:17 PM
Formatted: Font:

Jing Zhang 1/19/2015 12:17 PM
Formatted: Font:(Default) +Theme Body, 12 pt, Bold, Font color: Auto