***MBB***

Yale University

***266 Whitney Avenue***

***PO Box 208114***

***New Haven, CT 06520-8114***

***Telephone:***

***203 432 6105  
360 838 7861 (fax)***

***mark@gersteinlab.org***

***GersteinLab.org***

Today’s date

Dear Editor of Nature Biotechnology,

Please find enclosed our manuscript entitled “Sensitive information leakage from functional genomics data: Theoretical quantifications & practical file formats for privacy preservation”, which we hope will be considered for publication in Nature Biotechnology.

Our study relates to genomic privacy, which is becoming an increasingly important topic with the rise of personalized medicine and direct-to-consumer genetic testing. The advancement of technologies for high throughput biomedical data acquisition at en ever increasing pace is bringing a surge of datasets. Consequently, one of the biggest limitations in biotechnology is how we will deal with this large-scale human data that contain private information.

By instantiating simple linking attacks, we show that small number of DNA / RNA sequencing reads can result in large privacy breaches that makes it possible to identify individuals. This technology will soon have large scale practical use as We are also planning on providing standards and software support in a user-friendly environment through privaseq3.gersteinlab.org.

We strongly believe this study will be of great interest to Nature Biotechnology readership, because it provides a new biotechnological advancement that will have impacts on a broad audience from the bioinformatics community to experimental biomedical data producers. We were a bit unsure of the appropriate format of the manuscript for Nature Biotechnology. We have submitted this as a full-length article. However, if re-structuring of the manuscript is necessary to consider it for review, we would be happy to revise it.

We list a number of suitable reviewers for the paper.

Yours sincerely,

Mark Gerstein

Albert L. Williams Professor

of Biomedical Informatics