Yale University

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

Tel: 203 432 6105
Email: Mark.Gerstein@yale.edu

May 29, 2014

Dear Editor,

please find enclosed our manuscript, entitled "**Analysis of structural variation breakpoints from 1,092 humans reveals details of mutation mechanisms**", for consideration for publication in Nature Genetics. In this manuscript we derived confident set of breakpoints for 8,943 structural variants detected in 1,092 human individuals sequenced by the 1000 Genomes Project. We further performed a comprehensive analysis of the breakpoints including comparison in the context of genome functional and epigenetic contents. For this purpose we used data generated by the ENCODE and by the NIH Roadmap Epigenomics Mapping Consortium. Our study provides insights into the mutational mechanism leading to formation of structural variations. In particular, we hypothesize that structural variations mediated by non-allelic homologous recombination could originate in early embryonic and germ cells without replication and then are passed on through the germline. Furthermore, for structural variants generated by template switching during replication we suggest existence of particular spatial and temporal configurations for DNA to generate a variant.

Our analysis is of importance not only for fundamental understanding of mutations in the genome but also for clinical use – in particular, for understanding genetic alteration associated with cancer, senescence, and aging. We, therefore, believe that our manuscript will be of great interest to a broad audience and, especially, to the readers of Nature Genetics.

We have recently noticed the "Call for data analysis papers". We feel that out manuscript fits the announcement. We realize that we have missed the deadline for a presubmission inquiry; however, we completed and submit the manuscript before the manuscript submission deadline. We are kindly asking you to consider the manuscript for this call.

We list a number of suitable reviewers for the manuscript:

Aleksandar Milosavljevic, Baylor College of Medicine, amilosav@bcm.edu, 713-798-8719
Lars Fuek, Uppsala University, lars.feuk@igp.uu.se, +4618-471 4827
Stephen Scherer, University of Toronto, stephen.scherer@sickkids.ca, (416) 813-7613
Gabor Marth, Boston College, gabor.marth@bc.edu, (617)-552-3571
Kai Ye, The Genome Institute at Washington University, kye@genome.wustl.edu

We would prefer if the manuscript would not be sent to

James Lupski, Baylor College of Medicine
Evan Eichler, University of Washington

Yours sincerely,

Mark Gerstein
Albert L. Williams Professor of
Biomedical Informatics