

May 21, 2014

To the Editors:

We are submitting our manuscript “RVD2: An ultra-sensitive variant detection model for low-depth heterogeneous next-generation sequencing data” for consideration as an original article in *Bioinformatics*. It describes a novel statistical model and inference algorithm for identifying low-frequency single nucleotide mutations in heterogeneous samples.

Next-generation sequencing technology is increasingly being used for clinical diagnostic tests including non-invasive prenatal diagnostic tests and circulating tumor cell tests for chemotherapy monitoring. Unlike research cell lines, clinical samples are often genomically heterogeneous due to low sample purity or the presence of genetic subpopulations. Current, single nucleotide variant calling algorithms do not perform well to detect low frequency mutations in low read depth sequence data and a variant calling algorithm for calling low-frequency polymorphisms in heterogeneous samples is needed.

We have developed a novel statistical model and inference algorithm, RVD2, which uses a hierarchical Bayesian statistical model to estimate the frequency of variants in mixed populations. We show that RVD2 performs better on mixed samples than many other popular variant calling algorithms on a chemically synthesized DNA construct. We demonstrate the usefulness of RVD2 on a matched tumor-normal sample with known incomplete purity and identify several mutations in the *PAXIP1* gene that are not evident using other algorithms.

We are able to ensure our software implementation is available on our website for two years after publication. All supporting data for our manuscript is publically available.

Thank you for your consideration,
Patrick Flaherty
Assistant Professor
Department of Biomedical Engineering
Program in Bioinformatics and Computational Biology
Worcester Polytechnic Institute