Framework for confidence-based somatic mutation detection

Bernhard Y. Renard*[,1], Martin Löwer*, Jos de Graaf, Mustafa Diken, Özlem Türeci, Cedrik Britten, Sebastian Kreiter, Michael Koslowski, John C. Castle, Ugur Sahin

TRON - Translational Oncology at the Johannes Gutenberg University of Mainz, Langenbeckstr. 1, Building 708, 55131 Mainz, Rhineland-Palatinate, Germany USA
[1] Current address: Research Group Bioinformatics (NG 4), Robert Koch-Institute, Nordufer 20, 13355 Berlin
*authors contributed equally

Next generation sequencing (NGS) has enabled the high throughput discovery of genetic variations and somatic mutations, through resequencing of entire genomes to targeted regions (e.g. protein coding exons). However, the NGS technology platform process is still prone to errors resulting in inaccurate base calls, including both false positive and negatives. While multiple algorithms, platforms, and bench protocols exist, there is not effective framework for method comparison. Furthermore, many of the computational methods incorporate ad hoc filtering steps that do not lend themselves straightforwardly to statistical confidence measures. Here we present the results of a study with the aim of identifying somatic mutations in the B16 melanoma cell line using NGS, including sequencing of both the B16 and the associated C57BL/6 (Black6) mouse genomes. In the course of the study, we developed a framework to effectively and efficiently compare platforms, methods, and algorithms. Using these samples and comparison framework, we evaluated new software tools, including GATK, SAMTOOLS, and SomaticSNiPer, and lab protocols for the detection of somatic mutations. We present the framework for comparison, the results of the evaluation, and the somatic mutations found in the B16 melanoma cells.