Performance evaluation of indel calling tools using real short-read data

Mohammad Shabbir Hasan, Xiaowei Wu and Liqing Zhang

Human Genomics 2015 9:20
https://doi.org/10.1186/s40246-015-0095-3
© Hasan et al. 2015
Received: 9 April 2015  Accepted: 20 July 2015  Published: 19 August 2015

Abstract

Background

Insertion and deletion (indel), a common form of genetic variation, has been shown to cause or contribute to human genetic diseases and cancer. With the advance of next-generation sequencing technology, many indel calling tools have been developed, however, evaluation and comparison of these tools using large-scale real data are still scant. Here we evaluated seven popular and publicly available indel calling tools, GATK Unified Genotyper, VarScan, Pindel, SAMtools, Dindel, GTAK HaplotypeCaller, and Platypus, using 78 human genome low-coverage data from the 1000 Genomes project.

Results


UPS-indel: a Universal Positioning System for Indels

Mohammad Shabbir Hosen 1, Xiaowei Wu, Layne T. Watson & Liqing Zhang 2

Scientific Reports 7, Article number: 14106 (2017)

Abstract

Storing biologically equivalent indels as distinct entries in databases causes data redundancy, and misleads downstream analysis. It is thus desirable to have a unified system for identifying and representing equivalent indels. Moreover, a unified system is also desirable to compare the indel calling results produced by different tools. This paper describes UPS-indel, a utility tool that creates a universal positioning system for indels so that equivalent indels can be uniquely determined by their coordinates in the new system, which also can be used to compare different indel calling results. UPS-indel identifies 15%