Print this Page

Presentation Abstract

Title:	P16.52D - Need for speed in high-throughput sequencing data analysis
Keywords:	NGS data analysis; alignment; variant calling
Authors:	M. Plüss[1,2], S. M. Caspar[1], J. Meienberg[1], A. M. Kopps[1], I. Keller[3], R. Bruggmann[4], M. Vogel[2], **G. Matyas**[1]; [1]Center for Cardiovascular Genetics and Gene Diagnostics, Foundation for People with Rare Diseases, Schlieren-Zurich, Switzerland, [2]Institute of 4D Technologies, University of Applied Sciences and Arts Northwestern Switzerland, Windisch, Switzerland, [3]Department of Clinical Research, University of Berne, Berne, Switzerland, [4]Interfaculty Bioinformatics Unit and Swiss Institute of Bioinformatics, University of Berne, Berne, Switzerland.
Abstract:	Introduction: Rapidly evolving high-throughput sequencing (HTS) technologies produce huge amounts of data, requiring high-throughput data analyses with appropriate sensitivity/recall. However, widely-used alignment and variant calling tools are either too slow (e.g., BWA+GATK) or have insufficient sensitivity/recall (e.g., Isaac). In this study, we compared speed, disk footprint, and sensitivity/recall of the current gold standards BWA+GATK and Isaac with GENALICE MAP (genalice.com), a recently introduced ultra-fast HTS data analysis software solution. Materials and Methods: We performed alignment and variant calling on short-read (2x150bp) 60x PCR-free WGS data of NA12878 using BWA+GATK, Isaac, and GENALICE MAP. While measuring analysis time and disk footprint, we assessed the sensitivity/recall of the three pipelines according to GIAB v3.3 using RTG Tools. Results: Our data demonstrate more than 95 and 15 times higher speed as well as more than 45 and 20 times decreased disk footprint of GENALICE MAP compared to BWA+GATK and Isaac, respectively. The sensitivity/recall of GENALICE MAP was considerably higher than that of Isaac and close to BWA+GATK. Conclusion: The GENALICE MAP software offers hitherto unprecedented high speed and low disk footprint in alignment and variant calling with reasonable sensitivity/recall. Thus, it is a promising new tool in HTS data analysis either as a stand-alone/primary or a secondary solution, enabling realignments and reanalyses in ever-growing patient cohorts (e.g., due to updates of the reference genome such as hg19 to hg38).
Presentation Time:	Monday, May 29, 2017, 4:45 PM - 5:45 PM

The European Society of Human Genetics 2017