BMC Bioinformatics

MEETING ABSTRACT

Open Access

A short tutorial in analyzing NGS data of cancer genomes for somatic mutation calling

Huy Vuong[1], Zhongming Zhao[1,2,3,4*]

From 12th Annual UT-ORNL-KBRIN Bioinformatics Summit 2013
Buchanan, TN, USA. 22-24 March 2013

Background

Somatic mutation is the key element of tumorigenesis as these changes in nucleotide sequence of the cancer genome in somatic cells acquired throughout life can lead to protein alteration, cellular damage and thus cause cancer. The advent of next generation sequencing has significantly improved our ability to identify somatic mutations in cancer genomes paving the way for the comprehensive online catalogue for somatic mutation in human cancer (COSMIC)[1] which contains more than 820,000 mutations so far. Nevertheless, there are still many challenges in detecting somatic mutation in cancer especially for low frequency mutation due to either tumor heterogeneity or contamination with normal cells. Here, in this short tutorial, I will present a recent somatic mutation caller tool developed by the Broad Institute called Mutect[2] as part of the GATK (Genome Analysis Toolkit). I will use my own NGS dataset to demonstrate the tools and address some issues of troubleshooting input data and interpreting output.

Materials and methods

1. MuTect can be downloaded here: http://www.broadinstitute.org/cancer/cga/mutect. You must register an account with Broad Institute to download.

2. Instruction for running MuTect: http://www.broadinstitute.org/cancer/cga/mutect_run

3. User support forum: http://gatkforums.broadinstitute.org/categories/mutect

4. Because this workshop uses real world NGS data set, it's not required for participants to follow the walkthrough. However, participants are encouraged to run MuTect with their own data later (whole genome sequencing, whole exome sequencing, RNA sequencing data, etc.).

Authors' details

[1]Department of Biomedical Informatics, Vanderbilt University, Nashville, TN 37232, USA. [2]Department of Psychiatry, Vanderbilt University, Nashville, TN 37232, USA. [3]Department of Cancer Biology, Vanderbilt University, Nashville, TN 37232, USA. [4]Center for Quantitative Sciences, Vanderbilt University School of Medicine, Nashville, TN 37232, USA.

References

1. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, Jia M, Shepherd R, Leung K, Menzies A, Teague JW, Campbell PJ, Stratton MR, Futreal PA: **COSMIC: Mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer.** *Nucleic Acids Res* 2009, **39**(Database issue):D945.
2. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G: **Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples.** *Nat Biotechnology* 2013, **31(3)**:213-219.
3. DePristo M, Banks E, Poplin R, Garimella K, Maguire J, Hartl C, Philippakis A, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell T, Kernytsky A, Sivachenko A, Cibulskis K, Gabriel S, Altshuler D, Daly M: **A framework for variation discovery and genotyping using next-generation DNA sequencing data.** *Nature Genetics* 2011, **43(5)**:491-498.

* Correspondence: zhongming.zhao@vanderbilt.edu
[1]Department of Biomedical Informatics, Vanderbilt University, Nashville, TN 37232, USA
Full list of author information is available at the end of the article

BioMed Central