Title: A Short Review of Variants Calling for Single-cell-sequencing Data with Applications

Authors: Zhuohui Wei, Chang Shu, Changsheng Zhang, Jingying Huang, Hongmin Cai

PII: S1357-2725(17)30251-0
DOI: https://doi.org/10.1016/j.biocel.2017.09.018
Reference: BC 5225

To appear in: *The International Journal of Biochemistry & Cell Biology*

Received date: 1-1-2017
Revised date: 19-9-2017
Accepted date: 23-9-2017

Please cite this article as: Wei, Zhuohui., Shu, Chang., Zhang, Changsheng., Huang, Jingying., & Cai, Hongmin., A Short Review of Variants Calling for Single-cell-sequencing Data with Applications.*International Journal of Biochemistry and Cell Biology* https://doi.org/10.1016/j.biocel.2017.09.018

A Short Review of Variants Calling for Single-cell-sequencing Data with Applications

Zhuohui Wei, Chang Shu, Changsheng Zhang, Jingying Huang, Hongmin Cai

School of Computer Science & Engineering, South China University of Technology, Guangzhou, China
E-mail: hmcai@scut.edu.cn

Abstract The field of single-cell sequencing is fleetly expanding, and many techniques have been developed in the past decade. With this technology, biologists can study not only the heterogeneity between two adjacent cells in the same tissue or organ, but also the evolutionary relationships and degenerative processes in a single cell. Calling variants is the main purpose in analyzing single cell sequencing (SCS) data. Currently, some popular methods used for bulk-cell-sequencing data analysis are tailored directly to be applied in dealing with SCS data. However, SCS requires an extra step of genome amplification to accumulate enough quantity for satisfying sequencing needs. The amplification yields large biases and thus raises challenge for using the bulk-cell-sequencing methods. In order to provide guidance for the development of specialized analyzed methods as well as using currently developed tools for SNS, this paper aims to bridge the gap. In this paper, we firstly introduced two popular genome amplification methods and compared their capabilities. Then we introduced a few popular models for calling single-nucleotide polymorphisms and copy-number variations. Finally, break-through applications of SNS were summarized to demonstrate its potential in researching cell evolution.

Abbreviations: CNVs; SNPs; MDA; MALBAC; GATK; ADO; WGA; SNVs; NGS; aCGH; HMM; CNAs; cn.MOPS;

Keywords: single-cell sequencing; multiple-displacement amplification; multiple-displacement amplification and multiple-annealing, looping-based amplification cycling; single-nucleotide polymorphisms; copy-number variations

1 Introduction

In the past few years, heterogeneity at the molecular level among single cells in organs and tissues has aroused the interest of biologists. For example, tumor genetic heterogeneity is quite common and is important for reconstruction of the evolutionary history of the disease. The gene sequences and expression levels show large variations between adjacent tumor cells, and important differences between embryonic cells in several developmental stages have not yet been comprehensively understood [1]. However, such subtle differences due to cell heterogeneity cannot be accurately measured by bulk-cell sequencing [2], which is designed to measure an averaged profile of multiple cells.

The invention of single-cell sequencing technology carves out a new way to delineate intra tumor heterogeneity [3-5] and traces the evolution of single cells at the molecular level. This