Biomolecular Detection and Quantification 5 (2015) 23–29

Contents lists available at ScienceDirect

Biomolecular Detection and Quantification

journal homepage: www.elsevier.com/locate/bdq

Research paper

Feasibility of a workflow for the molecular characterization of single cells by next generation sequencing

Francesca Salvianti[a], Giada Rotunno[b], Francesca Galardi[c], Francesca De Luca[c], Marta Pestrin[c], Alessandro Maria Vannucchi[b], Angelo Di Leo[c], Mario Pazzagli[a], Pamela Pinzani[a,*]

[a] Department of Clinical, Experimental and Biomedical Sciences, University of Florence, Florence, Italy
[b] Department of Clinical and Experimental Medicine, University of Florence, Florence, Italy
[c] Sandro Pitigliani Medical Oncology Department, Hospital of Prato, Istituto Toscano Tumori, Prato, Italy

ARTICLE INFO

Article history:
Received 29 April 2015
Received in revised form 24 July 2015
Accepted 27 July 2015
Available online 12 September 2015

Keywords:
Single cell
Circulating tumor cells
NGS
Somatic mutations
Microgenomics

ABSTRACT

The purpose of the study was to explore the feasibility of a protocol for the isolation and molecular characterization of single circulating tumor cells (CTCs) from cancer patients using a single-cell next generation sequencing (NGS) approach.

To reach this goal we used as a model an artificial sample obtained by spiking a breast cancer cell line (MDA-MB-231) into the blood of a healthy donor.

Tumor cells were enriched and enumerated by CellSearch® and subsequently isolated by DEPArray™ to obtain single or pooled pure samples to be submitted to the analysis of the mutational status of multiple genes involved in cancer.

Upon whole genome amplification, samples were analysed by NGS on the Ion Torrent PGM™ system (Life Technologies) using the Ion AmpliSeq™ Cancer Hotspot Panel v2 (Life Technologies), designed to investigate genomic "hot spot" regions of 50 oncogenes and tumor suppressor genes.

We successfully sequenced five single cells, a pool of 5 cells and DNA from a cellular pellet of the same cell line with a mean depth of the sequencing reaction ranging from 1581 to 3479 reads.

We found 27 sequence variants in 18 genes, 15 of which already reported in the COSMIC or dbSNP databases. We confirmed the presence of two somatic mutations, in the *BRAF* and *TP53* gene, which had been already reported for this cells line, but also found new mutations and single nucleotide polymorphisms. Three variants were common to all the analysed samples, while 18 were present only in a single cell suggesting a high heterogeneity within the same cell line.

This paper presents an optimized workflow for the molecular characterization of multiple genes in single cells by NGS. The described pipeline can be easily transferred to the study of single CTCs from oncologic patients.

1. Introduction

The molecular analysis of solid tumors is of great impact in the field of oncology research notwithstanding some unsolved challenging aspects in both sample selection and technical approaches. Solid tumors are in fact a heterogeneous mixture of tumor and normal cells differentially contributing to the final results of molecular biology tests on nucleic acids. As a result molecular signatures represent an average value masking the individual signals deriving from each single cancer cell [1]. The analysis of somatic mutations at the single-cell level can evidence whether one or more variants detected in the bulk tissue are simultaneously present in the same cell or if, on the contrary, derive from different cell clones.

Next-generation sequencing (NGS) has become a major tool for nucleic acid analysis and promises to disclose the basis of diseases, opening perspectives for increased personalized therapies and screening [2]. NGS has been so far applied to nucleic acids deriving from a high number of cells providing global information on the average state of the entire population of cells [1]. By combining whole genome amplification (WGA) with NGS, it is now possible to achieve individual cancer cell sequencing [3].

* Corresponding author at: Dipartimento di Scienze Biomediche, Sperimentali e Cliniche Università Degli Studi di Firenze Viale Pieraccini, 6 50139 Firenze, Italy.
E-mail address: p.pinzani@dfc.unifi.it (P. Pinzani).

http://dx.doi.org/10.1016/j.bdq.2015.07.002

Single-cell sequencing has the potential to address some important questions in oncology such as understanding tumor heterogeneity, reconstructing cell lineages and defining the genome of rare tumor cell populations [3–5].

In addition, single-cell technologies will contribute to elucidate rare cell events such as scarce cancer cells in clinical samples and the dissemination of circulating tumor cells (CTCs) in the blood of cancer patients [1,6]. Since CTCs reflect the heterogeneity of primary or metastatic tumors, their analysis at the single-cell level represent a real time liquid biopsy of the disease giving an insight in cancer progression, helping in the choice of target therapies or in the prediction of the efficacy of chemotherapeutic agents [1] and thus representing a key tool in the development of precision medicine.

Another aspect that can be considered as an application of single-cell sequencing is the analysis of cell lines derived from tumors that are commonly used as models in cancer research. This approach can be successfully adopted to identify those cell lines that more closely resemble the characteristics of the tumor under study [7]. Even if at a much lower extent than tissues, the cell line population usually considered as composed of identical cellular elements can hide a heterogeneity that could be disclosed only by single-cell analysis.

The MDA-MB-231 is an adherent human breast cancer cell line of epithelial origin, which is highly aggressive, invasive and poorly differentiated. MDA-MB-231 cells display a mesenchymal-like phenotype thus representing a good model for metastatic breast cancer [8].

In the present study we investigated the mutational status of multiple genes on an artificial sample obtained by spiking the MDA-MB-231 cell line into the blood of a healthy donor. The tumor cells were immunomagnetically isolated by CellSearch® (Veridex, USA), sorted by a dielectrophoretic system (DEPArray™, Silicon Biosystems, Italy) and then submitted to WGA and targeted deep sequencing. We used this model with the aim to demonstrate the feasibility of a workflow for the isolation and molecular characterization of single circulating tumor cells from cancer patients using a single-cell NGS approach.

2. Methods

2.1. Cell culture and spiking experiments

Experiments were conducted on the MDA-MB-231 human breast cancer cell line. MDA-MB-231 cells were obtained from Di.V.A.L. Toscana (Laboratory for Drug Validation and Antibody production, University of Florence) and maintained in DMEM (Lonza, Basel, Switzerland) + 10% FBS (PAA Laboratories, Austria) at 37 °C in 5% CO_2. Cells were detached from plastic surface by Accutase (PAA Laboratories) to better preserve membrane antigens. MDA-MB-231 cells ($n = 500$) were spiked in blood samples from a healthy volunteer collected in CellSave® tubes (Veridex LLC, Raritan, NJ) and processed as described below. The remaining cells from the same culture were centrifuged and the resulting cell pellet was submitted to DNA extraction by QIAamp DNA Mini Kit (Qiagen, Germany).

2.2. Cell enrichment and enumeration

Cell enrichment and enumeration was performed by the CellSearch® System. 7.5 mL of whole blood were processed using the CellSearch™ Epithelial Cell kit (Veridex LLC), which selects EpCAM positive cells using ferrofluids particles coated with EpCAM antibody. Cells were stained with the nuclear dye 4′,6′-diamino-2-phenylindole (DAPI), anti-cytokeratin 8, 18 and 19-phycoerythrin (PE) labelled antibodies, and CD45 antibody labelled with allophycocyanin (APC). After enrichment, isolated and stained cells were resuspended in the MagNest Device (Veridex LLC), labelled cells were analyzed in the CellTracks® Analyzer II (Veridex LLC) and cells identified and enumerated according to the criteria specified by the manufacturer's instructions.

2.3. Single cell sorting

Following CellSearch® enrichment the sample was recovered from the Veridex cartridge and loaded into the DEPArray™ A300K chip (Silicon Biosystems) according to the manufacturer's instructions. The chip is a single-use, microfluidic cartridge containing an array of individually controllable electrodes, each one with embedded sensors. Chip scanning was performed by an automated fluorescence microscope to generate an image gallery, with cells selected according to their morphology (round shape, round nucleus within the cytoplasm) and staining pattern deriving from that of the CellSearch® system: DAPI positive, PE positive (CK8, CK18, CK19 positive cells), APC negative (CD45 negative cells).

After tumor cell identification, single cells or groups of few cells were recovered into 200 μl tubes.

2.4. Whole genome amplification

Single cells or pools of cells were submitted to WGA using the Ampli1™ WGA kit (Silicon Biosystems) according to the manufacturer's instructions, in order to obtain a sample suitable for sequencing analysis. The procedure is based on a ligation-mediated PCR following a site specific DNA digestion by *MseI* enzyme. The kit has no needs for precipitation steps avoiding DNA loss and provides a library of fragments of about 0.2-2 kb representing the entire genome. The kit uses a mixture of Taq polymerase with a proofreading enzyme, Pwo polymerase, that has been reported to have error rates more than 10 times lower than the error rate observed with *Taq* polymerase [9].

2.5. WGA quality control

The quality of the output product of the WGA reaction was assessed by the Ampli1™ QC kit (Silicon Biosystems) according to the manufacturer's instructions. The kit is a PCR-based assay which implies the amplification of two distinct regions of the human genome to produce two amplicons (A and B) of 373 and 167 bp respectively. PCR products A and B were analyzed by capillary electrophoresis on the Agilent 2100 Bionalyzer. The presence of both amplification products indicates a successful WGA and consequently the suitability of the sample for downstream analysis.

2.6. Next generation sequencing

Sequencing analysis was performed on the Ion Torrent PGM™ system (Life Technologies). We amplified the samples using the Ion AmpliSeq™ Cancer Hotspot Panel v2 (Life Technologies) designed to target 207 amplicons covering mutations from 50 oncogenes and tumor suppressor genes. DNA quantification was assessed using the Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). Ten nanograms of DNA were used to prepare barcoded libraries using the Ion AmpliSeq™ Library kit 2.0 and Ion Xpress™ barcode adapters (Life Technologies). The libraries were purified with Agentcourt AMPure XP (Beckman Coulter) and quantified with the Ion Library Quantitation Kit (Life Technologies) on the StepOne Plus system (Applied Biosystem).

Template preparation was performed with the Ion OneTouch™ 2 System and Ion One Touch ES. Finally sequencing was performed on PGM using Ion PGM™ Sequencing 200 kit v2 (Life Technologies)