Whole transcriptomics analyses of mimicking Circulating Tumor Cells (CTCs) by single-cell RNA sequencing (scRNAseq)

Jessica Garcia[1,2,3], François Monjaret[5], Florence Geiguer[1,2,3], Jean-Philippe Aurel[7], Alain Puisieux[2], Anne Pierre Morel[2], Claire Bardel[6], Gaelle Vilchez[6], Brigitte Manship[2], Sébastien Couraud[1,4], Arnaud Vigneron[2], Guillhem Tourniaire[5] & Léa Payen[1,2,3]

1. Laboratoire de Biochimie et Biologie Moléculaire, Groupe Hospitalier Sud, Hospices Civils de Lyon, Lyon, 69003, France / 2. Université Claude Bernard Lyon 1, INSERM 1052, CNRS 5286, Centre Léon Bérard, Cancer Research Center of Lyon, Equipe Labellisée Ligue contre le Cancer, Lyon, France / 3. Cancer Institute of Hospices Civils de Lyon, CIRculating CANcer (CIRCAN) program, 69495 Pierre Bénite, France / 4. Service de Pneumologie Aigue Spécialisée et Cancérologie Thoracique, Groupement Hospitalier Sud, Cancer institute of Hospices Civils de Lyon, Pierre Bénite, 69495, France / 5. Cellenion SASU, LYON, France / 6. Laboratoire de Bio-informatique, Groupe Hospitalier, Hospices Civils de Lyon, Lyon, 69003, France / 7. Clearbridge Biomedics Inc., Singapore

Purposes

Ultra-precision medicine is emerging as a cornerstone of cancer biology, as it aims a providing a better comprehension of cancer cells and their clinical response to immunotherapy. Consistently, single-cell sequencing of CTCs is a powerful tool to further decipher tumor plasticity and to identify new pharmacological targets influencing clinical outcome and response to treatment. The aim of this study is to validate an original workflow allowing the isolation at the single-cell level of mimicking CTCs without interfering with single-cell RNAseq analysis.

Methodology

The technological advances in microfluidic systems and isolation technologies have resulted in the enriched extraction of mimicking CTCs from healthy whole blood samples. In the present study, the ClearCell Fx (Biolidics Limited) was used as a label-free microfluidic system for enrichment of wholly intact CTCs, while the cellenONE F1.4 system (Cellenion) was used to isolate single CTCs. The latter allowed high-throughput automated isolation and dispensing of single CTCs in 96-well plates containing cell lysis buffer. From these 96-well plates, scRNA libraries were prepared with the NebNext Single Cell/Low Input kit (New England Biolabs) and sequenced with paired-end sequencing (2x75 bp) on the NextSeq 500/550 from Illumina. FASTQ files, obtained by the demultiplexing of Base-Call Files (BCL) were used for reads alignment and gene annotation based on the Hg19 (GRCh37.p5) genome reference. The mapped and annotated reads were stored in BAM files, processed using the Gfold (Generalized FoldChange) algorithm to produced normalized read counts. Those were then uploaded in a web-based platform named ASAP (Automated Single-cell Analysis Pipeline) to generate PCA analyses in order to distinguish cell populations and to produce HEATmaps to observe the differential expression of genes between cells.

1 The h-TERT-shp53/RAS HMEC (Morel et al., 2012) breast cancer cell line was used as mimicking-CTC cell line to evaluate transcriptomic changes. This cell line contains three different sub-populations : Mesenchymal cells, epithelial cells and cells undergoing EMT. This enabled us to evaluate whether the complete workflow provides information on tumoral heterogeneity.

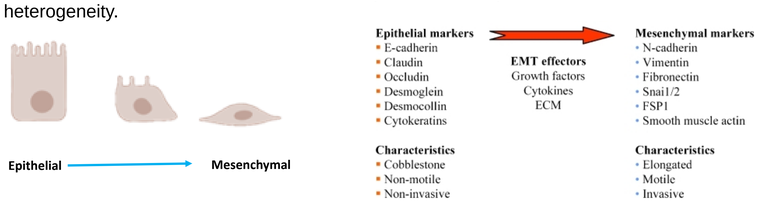

Figure 1 : Description of tumor cell line h-TERT-shp53/RAS HMEC. Epithelial-to-Mesenchymal Transition (EMT) is a morphogenetic process in which epithelial cells loose their characteristics and gain mesenchymal properties during embryogenesis and cancer progression.

2

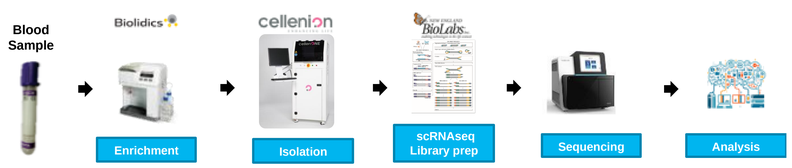

Figure 2: Overview of the CTC isolation workflow at the single cell level

3

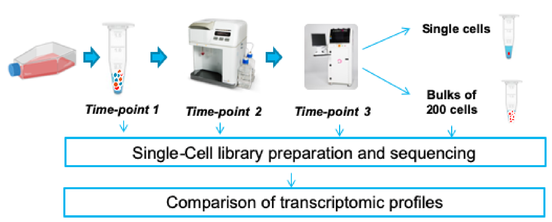

Figure 3: Overview of the CTC isolation workflow workflow at the single cell level and comparison with a bulk of 200 mCTCs

Acknowledgement & funding
Authors would like to thank all patients, families and prescribing physicians.
This work was supported by Astra Zeneca, Sysmex Inostics, ClearBridge Biomedics, Merck, Sophia Genetics, and *Ligue Nationale contre le Cancer of Saône et Loire, the CLARA, The Region Rhône Alpes Auvergne, Innovotion*
Contact: jessica.garcia02@chu-lyon.fr / lea.payen-gay@chu-lyon.fr / sebastien.couraud@chu-lyon.fr

Results

4

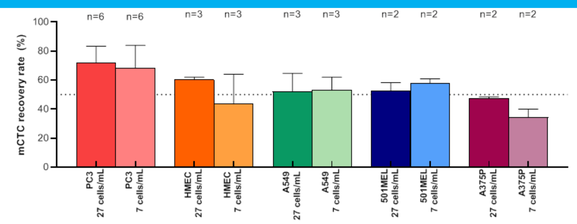

Figure 4 : Enrichment of Circulating Tumor Cells (CTS's) by a Label-Free Inertial microfluidic method: the red blood cells were eliminated with a lysis buffer. The residual white blood cells and mCTC's were loaded in the ClearCell CTChip FR1. mCTC enrichment is based on the Dean Flow Fractionation principle. The smaller hematological cells [8 – 15 µm] are affected by the Dean Drag and migrate to waste outlet whereas the mCTC's [>15 µm] migrate to the upper outlet.

ClearCell recovery
- Recovery of mCTCs superior to 60% for most of the cell lines
- Recovery of 1 mCTC/ml (the clinical frequency in lung cancer is 7 CTCs/ml whole blood; internal data, Garcia et al. Jove 2019)

5

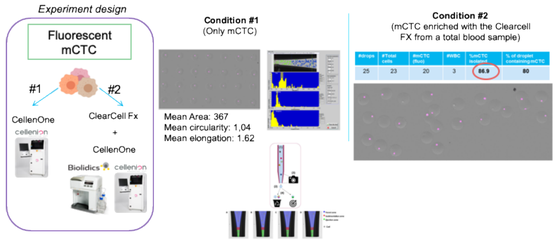

Figure 5: Experimental design to evaluate the CellenOne sorting impact on mCTC morphology and recovery rate

Physical integrity of the cell shape after sorting
- Conservation of cell circularity and elongation after the ClearCell and CellenOne steps
- Achievement of a recovery rate of 86.9% of isolated mCTCs mixed with WBC

6

	mCTC 1% (Assay 1)	mCTC 1% (Assay 2)	mCTC 0.1%
CTC negative selection based on CD45 intensity	yes	yes	yes
Retained size (µm)	from 12 to 50	from 15 to 50	from 16 to 50
Recovered mCTC/ Number of total isolations	19/52	26/30	7/7
Recovery rate (%)	37	87	100
Identity checking on glasslide			

Figure 6: Table summarizing mCTC recovery rates of the complete workflow (A) before experimental condition optimization (B) after optimization at 1% mCTCs and 0.1% mCTCs in the enrichment from Clearcell FX1 instrument

Optimization of cell piking based on :
- Exclusion fluorescence labelling of WBC cells
- The size inclusion rate of mCTCs

7

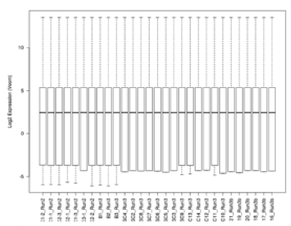

Figure 7: Average of Log2 expression by Voom. Single cells and bulk mCTCs samples

Average Log2 expression normalized by Voom
- Average of around 10, 000 transcripts expressed per sample at each time point both in single cell and bulk
- High homogeneity of the number of transcripts: single cells and bulks

8A

8B

Comparaison between B and C1 samples

Database "Gene Ontology" (GO)	N° GO	Description	Regulation of B samples	Odd-Ratio	p Value
Biological Processes	GO:0030595	leukocyte chemotaxis	↗	7,9	8,46E-08
Biological Processes	GO:0050900	leukocyte migration	↗	4,8	9,62E-07
Biological Processes	GO:0030593	neutrophil chemotaxis	↗	12,2	1,12E-06
Biological Processes	GO:0097529	myeloid leukocyte migration	↗	7,8	1,12E-06
Biological Processes	GO:0060326	cell chemotaxis	↗	5,9	1,17E-06
Biological Processes	GO:1990266	neutrophil migration	↗	10,5	3,43E-06
Cellular component	x	WholeBlood	↗	4,6	4,43E-04
Cellular component	x	CD14+_Monocytes	↗	3,9	3,97E-02

Figure 8: (A) ACP of different experimental conditions : time-point 1 named C1; time-point 2 named C2 time point 3 named B. **(B)** Principle dysregulated transcriptomics signatures in the Bulk condition (B) compared to trypsinized cells (C1)

PCA of time-point 1 (C1); time-point 2 (C2) time point 3 (B)
- Very high reproducibility intra-conditions
- The bulk conditions are between the freshly trypsinized cells and immune C2 cells (signature of white blood cells highlighted)
- In conclusion, the spiking of HMECs (breast cells) into a healthy whole blood slightly modifies the transcriptomics of the cell line.
- Regulation of signatures linked to immune response stimulation. The HMECs are recognized as a pathogen.

9

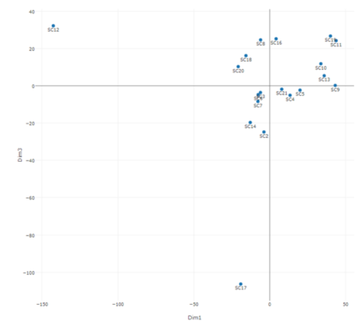

Figure 9: PCA of single cell transcriptome: comparison of transcript levels of single HMEC isolated cells (SC)

Difference between bulk (B) and single cells (SC)
- The bulk transcriptomics are slightly different from SC transcriptomics, due to technical differences in the extraction of total RNA (data not shown)
- A clear heterogeneity in transcriptomics is observed in the SC transcriptome

10

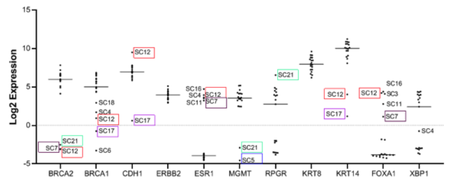

Figure 10: Log2 fold-changes of markers, expressed in breast cancer cells

Tumoral heterogeneity studied at single mCTC level
- SC7, SC12 – highly expressed ESR1; under-expressed BRCA2
- High heterogeneity of the HMEC population in agreement with its characteristic high plasticity

11

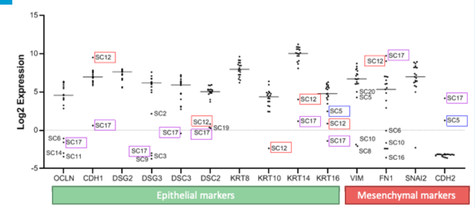

Figure 11: Log2 fold-changes of markers, involved into the epithelia-mesenchymal transition (EMT)

Tumoral heterogeneity studied at single mCTC level
- SC17- Mesenchymal profile
- SC12 – Epithelial profile
- The majority of the cells are in transition (intermediate profile). This shift might be due to the stimulation of immune response of white blood cell of the donor against the human breast HMEC cells

12

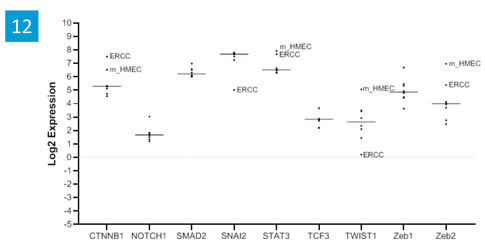

Figure 12: Log2 fold-changes of transcription factors in bulk of 200 HMECs with and without the similar frozen step between Clearcell FX1 enrichment and CellenOne isolation. ERCC positive control is a pool of 50 cell lines and m_HMEC is the transcriptomics of the mesenchymal sub-population of HMECs.

None impact of the frozen step on the transcriptomics of the bulk of HMEC cells
- ZEB1, ZEB2 and twist 1 are over expressed in mesenchymal HMEC cells

Conclusion & Perspectives

In conclusion, we propose this workflow as a standard protocol to analyze CTCs owing to its specificity and reliability at single-cell level. Such knowledge of single-cell biology may lead to the development of specific therapies to limit tumor progression and seeding of tumoral cells into secondary healthy organs by blocking newly identified targets