Utilisation des clusters de calcul en cancérologie

application à l'étude de l'hétérogénéité cellulaire des tumeurs

Frédéric PONT, Marie Tosolini, Jean Jacques Fournié.
Pôle Technologique du Centre de Recherches en Cancérologie de Toulouse

JCAD 2019

Toulouse, octobre 2019

Inserm
La science pour la santé
From science to health

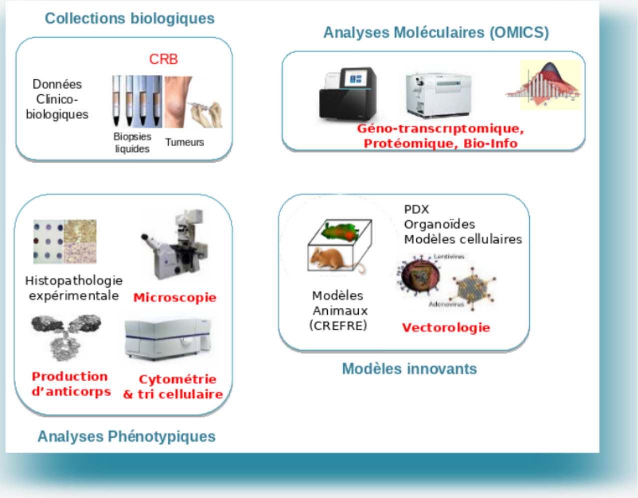
CRCT
Centre de Recherches en Cancérologie Toulouse
ONCOPOLE

Pôle
TECHNOLOGIQUE
CRCT

cnrs

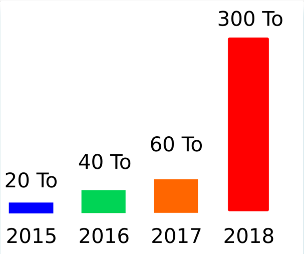
UNIVERSITÉ
TOULOUSE III
PAUL SABATIER
Université de Toulouse

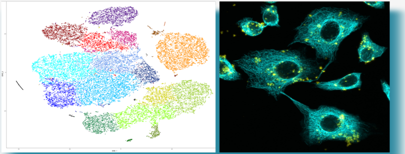
Le Pôle Technologique du CRCT (www.poletechno-crct.inserm.fr)


Collections biologiques
CRB
Données Clinico-biologiques
Biopsies liquides
Tumeurs
Analyses Moléculaires (OMICS)
Géno-transcriptomique, Protéomique, Bio-Info
PDX
Organoïdes
Modèles cellulaires
Lentivirus
Adenovirus
Modèles Animaux (CREFRE)
Vectorologie
Modèles innovants
Histopathologie expérimentale
Microscopie
Production d'anticorps
Cytométrie & tri cellulaire
Analyses Phénotypiques

- Technologies et compétences accessibles aux labos publics et privés

Evolution du volume de données au CRCT





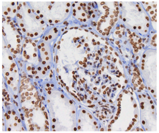
Nouvelles Technologies

OPEN DATA

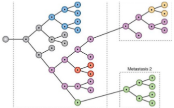
Le cancer : hétérogénéité à tous niveaux

population

patient

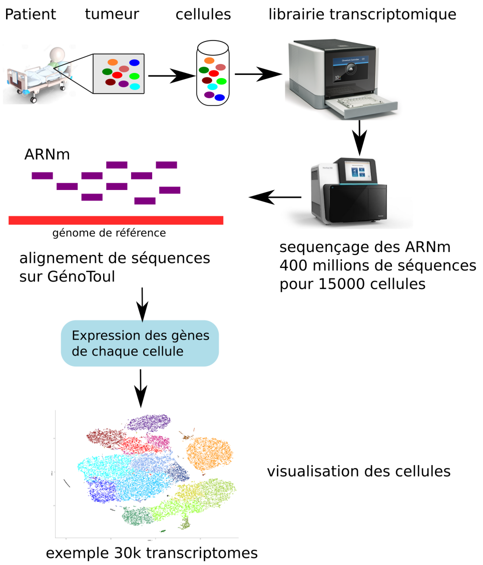


tumeur



clône
tumoral

Single Cell RNAseq et hétérogénéité cellulaire



Patient tumeur cellules librairie transcriptomique

ARNm

génome de référence

alignement de séquences
sur GénoToul

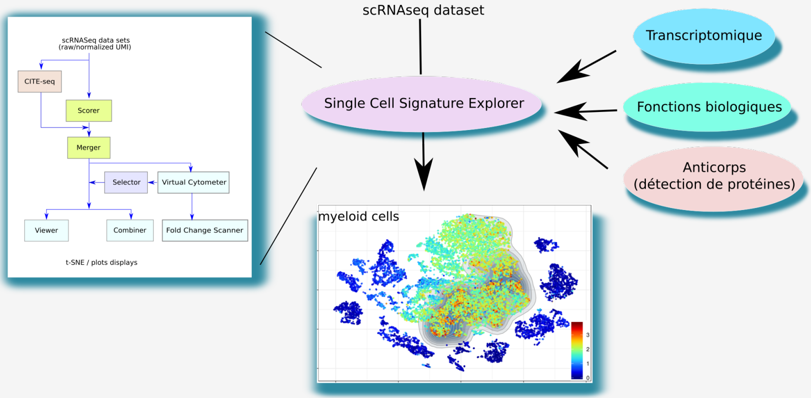
sequençage des ARNm
400 millions de séquences
pour 15000 cellules

Expression des gènes
de chaque cellule

visualisation des cellules

exemple 30k transcriptomes

Signification biologique : Single-Cell Signature Explorer



scRNAseq dataset

Single Cell Signature Explorer

Transcriptomique

Fonctions biologiques

Anticorps
(détection de protéines)

scRNASeq data sets
(raw/normalized UMI)
CITE-seq
Scorer
Merger
Selector
Virtual Cytometer
Viewer
Combiner
Fold Change Scanner
t-SNE / plots displays
myeloid cells
3
2
1
0

Tumeur du poumon

- Logiciels libres pour le Single-Cell RNAseq (Nucleic Acids Res. 2019 Jul 11, PNAS 2019 Jun 11, `https://sites.google.com/site/fredsoftwares/products`)
- Les calculs sont effectués sur CALMIP.

Calcul des scores des fonctions biologiques sur cluster

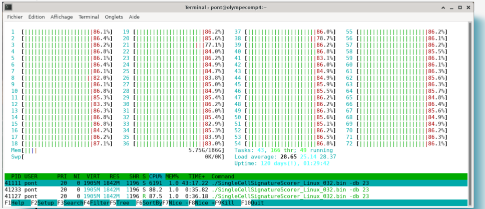
scRNAseq dataset : matrice cellules - gènes de [10-100 k cellules] x 20k gènes

Single Cell Signature Scorer

20 000 listes de gènes
= Fonctions biologiques

$$\text{score of } C_j \text{ cell for } GS_x \text{ gene set} = \frac{\sum_{i=1}^{k} UMI_{GS_{x_i}}}{\sum_{i=1}^{n} UMI_{C_{j_i}}} \times 100,$$

calcul parallèle :
1 cellule & 20 000 fonctions biologiques / CPU


Terminal - pont@olympecomp4:~
Fichier Édition Affichage Terminal Onglets Aide

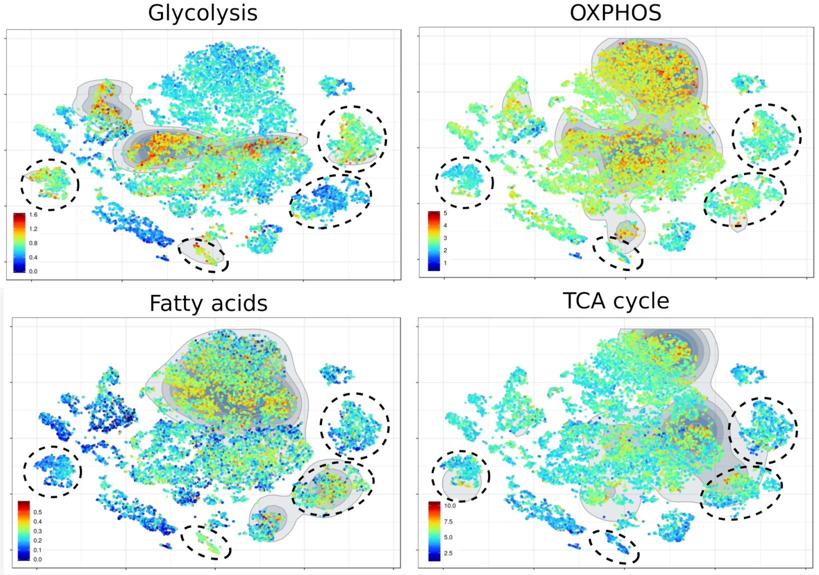
```
  1  [||||||||||||||||||||||86.1%]   19 [||||||||||||||||||||||86.2%]   37 [||||||||||||||||||||||86.0%]   55 [||||||||||||||||||||||86.2%]
  2  [||||||||||||||||||||||86.4%]   20 [||||||||||||||||||||||85.6%]   38 [||||||||||||||||||||||78.7%]   56 [||||||||||||||||||||||86.1%]
  3  [||||||||||||||||||||||86.2%]   21 [||||||||||||||||||||||77.1%]   39 [||||||||||||||||||||||86.2%]   57 [||||||||||||||||||||||86.2%]
  4  [||||||||||||||||||||||86.8%]   22 [||||||||||||||||||||||84.0%]   40 [||||||||||||||||||||||86.8%]   58 [||||||||||||||||||||||85.5%]
  5  [||||||||||||||||||||||86.1%]   23 [||||||||||||||||||||||86.2%]   41 [||||||||||||||||||||||83.1%]   59 [||||||||||||||||||||||86.1%]
  6  [||||||||||||||||||||||86.4%]   24 [||||||||||||||||||||||84.9%]   42 [||||||||||||||||||||||86.0%]   60 [||||||||||||||||||||||86.1%]
  7  [||||||||||||||||||||||86.1%]   25 [||||||||||||||||||||||84.7%]   43 [||||||||||||||||||||||84.9%]   61 [||||||||||||||||||||||86.3%]
  8  [||||||||||||||||||||||82.0%]   26 [||||||||||||||||||||||83.8%]   44 [||||||||||||||||||||||84.9%]   62 [||||||||||||||||||||||85.6%]
  9  [||||||||||||||||||||||86.1%]   27 [||||||||||||||||||||||85.0%]   45 [||||||||||||||||||||||85.9%]   63 [||||||||||||||||||||||86.3%]
  10 [||||||||||||||||||||||86.8%]   28 [||||||||||||||||||||||84.9%]   46 [||||||||||||||||||||||85.5%]   64 [||||||||||||||||||||||85.7%]
  11 [||||||||||||||||||||||85.3%]   29 [||||||||||||||||||||||85.4%]   47 [||||||||||||||||||||||86.2%]   65 [||||||||||||||||||||||84.8%]
  12 [||||||||||||||||||||||83.3%]   30 [||||||||||||||||||||||86.2%]   48 [||||||||||||||||||||||86.2%]   66 [||||||||||||||||||||||86.4%]
  13 [||||||||||||||||||||||86.3%]   31 [||||||||||||||||||||||86.0%]   49 [||||||||||||||||||||||86.3%]   67 [||||||||||||||||||||||85.6%]
  14 [||||||||||||||||||||||86.8%]   32 [||||||||||||||||||||||85.4%]   50 [||||||||||||||||||||||85.6%]   68 [||||||||||||||||||||||84.9%]
  15 [||||||||||||||||||||||86.8%]   33 [||||||||||||||||||||||82.9%]   51 [||||||||||||||||||||||85.1%]   69 [||||||||||||||||||||||84.9%]
  16 [||||||||||||||||||||||84.2%]   34 [||||||||||||||||||||||85.3%]   52 [||||||||||||||||||||||86.2%]   70 [||||||||||||||||||||||86.2%]
  17 [||||||||||||||||||||||86.2%]   35 [||||||||||||||||||||||83.9%]   53 [||||||||||||||||||||||86.5%]   71 [||||||||||||||||||||||86.3%]
  18 [||||||||||||||||||||||87.1%]   36 [||||||||||||||||||||||83.0%]   54 [||||||||||||||||||||||86.8%]   72 [||||||||||||||||||||||86.1%]
  Mem[||||                                    5.75G/186G]   Tasks: 43, 166 thr; 49 running
  Swp[                                           0K/0K]   Load average: 28.65 25.14 28.37
                                                            Uptime: 120 days(!), 01:29:42

  PID USER      PRI  NI  VIRT   RES   SHR S CPU% MEM%   TIME+  Command
41111 pont       20   0 1905M 1842M  1196 S 6191  1.0 43:17.22 ./SingleCellSignatureScorer_Linux_032.bin -db 23
41233 pont       20   0 1905M 1842M  1196 S 88.2  1.0  0:35.82 ./SingleCellSignatureScorer_Linux_032.bin -db 23
41127 pont       20   0 1905M 1842M  1196 R 87.5  1.0  0:36.18 ./SingleCellSignatureScorer_Linux_032.bin -db 23
F1Help  F2Setup F3SearchF4FilterF5Tree  F6SortByF7Nice -F8Nice +F9Kill  F10Quit
```

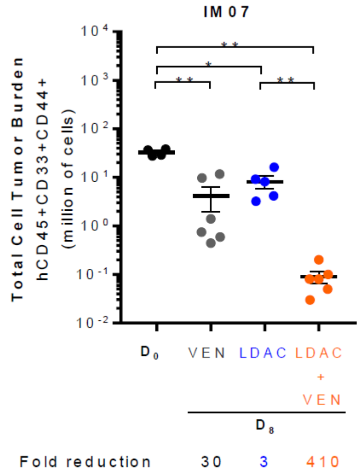
1h de calcul sur 3 noeuds CALMIP (216 CPUs)

Single Cell Signature Viewer

Profil métabolique de 6 tumeurs du poumon



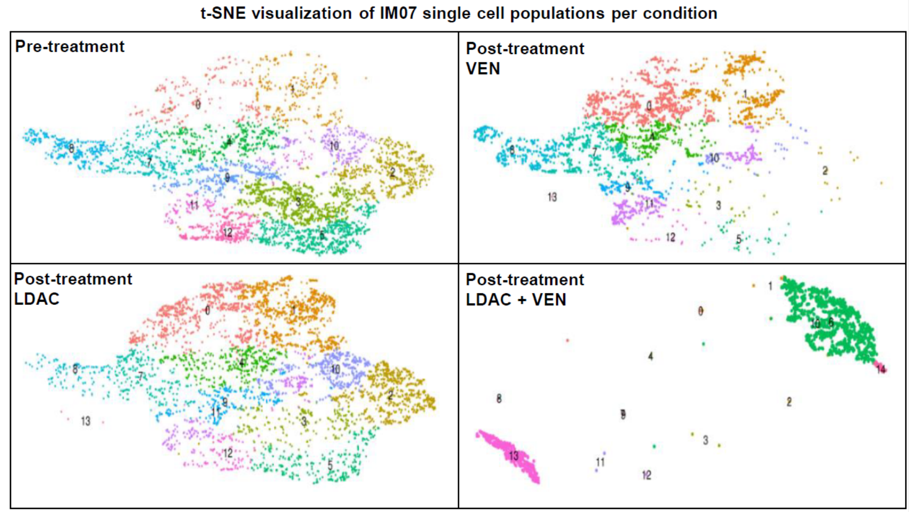
Réduction de la tumeur (LAM) après différents traitements



LDAC : chimiothérapie. VEN : inhibiteur BCL2

source : Claudie BOSC, Jean-Emmanuel Sarry, CRCT (2019)

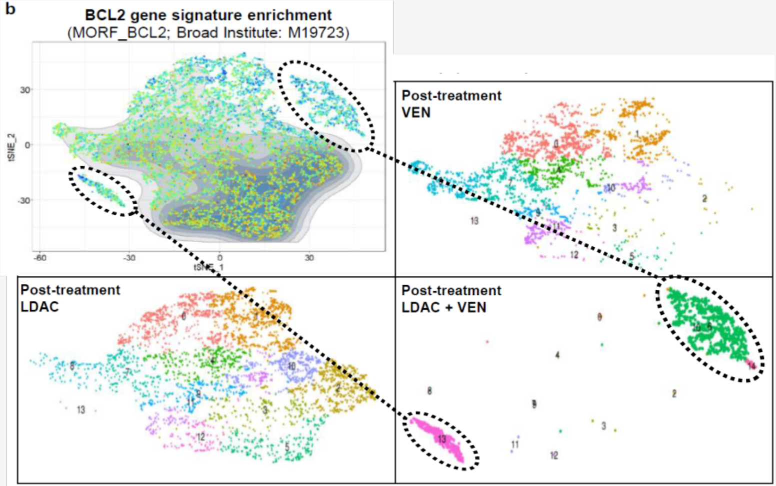
Impact du traitement sur les clones tumoraux



LDAC : chimiothérapie. VEN : inhibiteur BCL2

source : Claudie BOSC, Jean-Emmanuel Sarry, CRCT (2019)

Caractérisation des deux clones résistants au traitement



LDAC : chimiothérapie. VEN : inhibiteur BCL2

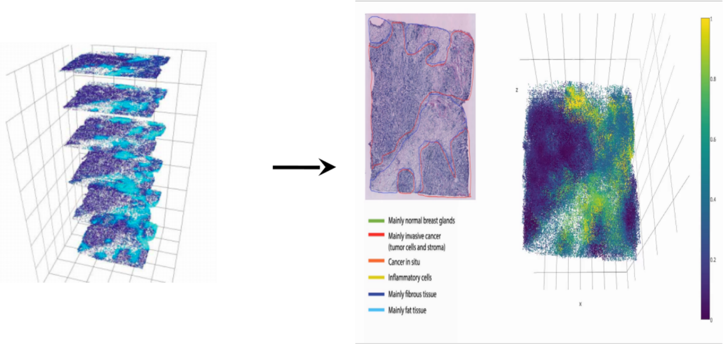
source : Claudie BOSC, Jean-Emmanuel Sarry, CRCT (2019)

Bilan et Perspectives

- Le Single Cell RNAseq est la technique la plus puissante pour explorer l'hétérogénéité des tumeurs.
- Une technique en plein développement avec de nombreuses variantes, et des besoins en calcul de plus en plus importants.
- Les clusters de calcul GénoToul et CALMIP apportent un gain très important en productivité et libèrent notre serveur local pour des tâches interactives.
- Les besoins en calcul de la recherche en cancérologie au CRCT sont en pleine expansion.

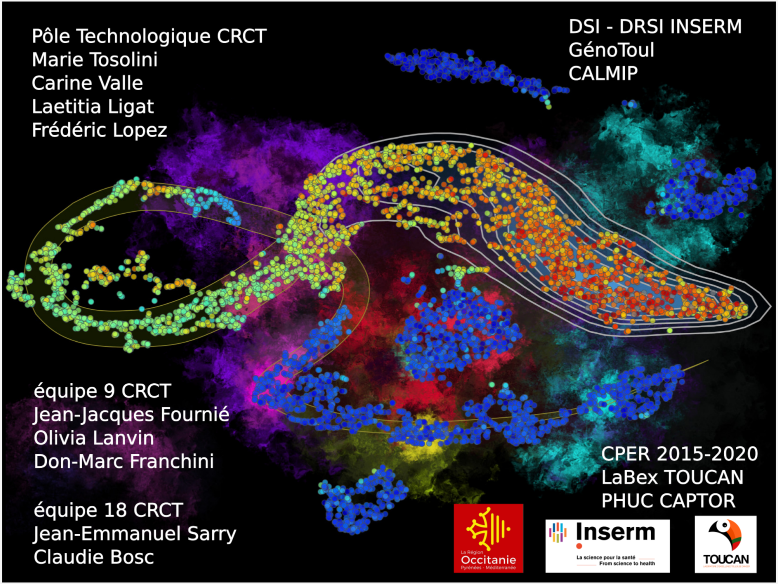
Bilan et Perspectives

spatial transcriptomics in 3D within human tumors


Mainly normal breast glands
Mainly invasive cancer (tumor cells and stroma)
Cancer in situ
Inflammatory cells
Mainly fibrous tissue
Mainly fat tissue
z
x
1
0.8
0.6
0.4
0.2
0

by courtesy of 10xGenomics (2019)

Remerciements



Pôle Technologique CRCT
Marie Tosolini
Carine Valle
Laetitia Ligat
Frédéric Lopez

DSI - DRSI INSERM
GénoToul
CALMIP

équipe 9 CRCT
Jean-Jacques Fournié
Olivia Lanvin
Don-Marc Franchini

équipe 18 CRCT
Jean-Emmanuel Sarry
Claudie Bosc

CPER 2015-2020
LaBex TOUCAN
PHUC CAPTOR

La Région Occitanie Pyrénées · Méditerranée

Inserm La science pour la santé From science to health

TOUCAN