#=================================================================================================================#

# Exploratory analysis of fetal human skin

# Author: Mahima Arunkumar

# Date: 28.01.2022

# Adapted from https://satijalab.org/seurat/pbmc3k\_tutorial.html

# Data for fetal human skin from: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE156972

# Rscript purpose:

# - Identify the alpha and beta T-cells in the skin

# - Compare this to healthy donor (EX0004)

# - DEG in order to identify a signature for "naive" T cells from the author's data

# (of the paper above)

# - Take this signature to identify "naive" T cells in our healthy donor (EX0004) and allo-HSCT donor

# - Check for developmental pathway using velocity and diffusion map tool

# - Clonality analysis to assess the expansion of the cells

# - Test for possible ligands

# - UMAP and violin plots to check for enrichment of residency signatures and the markers CD69, CD103 etc.

# in the fetal T cells and compare to adult

# Goals and Background:

# - We would like to see whether naive T cells can persist/be resident in the adult skin

# - We want to know if they might even be remnants from fetal live (persisting precursor cells)

# - We want to see if there is a developmental pathway as there are also some memory T cells in the fetal skin

# - Only cells with the same TCR belong to the same family that might have differentiated. Expectation:

# Naive T cells should not be clonal, which is why this analysis is relevant to see what could have stimulated

# them in utero

# - We want to see if certain residency signatures are enriched

#=================================================================================================================#

List of TODOs:

1. Download and structure all relevant Gustavo scripts in local laptop:

* 383: tcf7 expression for each subset in blood and skin and merged [202004]
* 458: currently investigting
* 539: include only Tm and Tn sigs and exclude non-Tcells [202011]
* (463): look kurz [202007]

1. Take gdT cells out
2. Find out number of differentially expressed genes when comparing fetal T cells and adult T cells
3. Take top X up and down signatures from fetus and apply to adult to see if we have remnants
4. TCF7 expression
5. Take fetal skin and apply signatures to look at heterogeneity in fetal skin to show that there are no memory cells. Doing so we can see what is truly naïve.
6. These truly naïve signatures we again take and check in adult skin and check if they are naïve and if they are also resident.
7. Check residency in 2 ways: 1. Apply residency signatures 2. Look at host vs. donor (EX0008)

* If we find naïve cells in adult skin, and if they even have same identity as cells from fetus and they are host, that would be nice! 😉
* We are looking at scRNAseq data, so make sure to take individual cells when comparing!

Später TODOs:

* Mail to 10X Kontaktperson regarding data completeness and clonality
* Restructure Rscript
* Annotate Präsi with most important findings
* Upload präsi, What has been done liste and data

ToDo 18.03.2022:

* Seurat workflow with Gustavo marker genes mentioned in R-script GA\_AN0152 + TCF7 as naïve and as separate (**check**)
* Generate UMAP/tSNE for before and after removing all non T-cells (**check**)
* Print table showing proportion of each subset, once for both and once for each condition (fetal/ adult) (**check**)
* Update presentation slides (**check**)
* Find differentially expressed genes -> how many (**check**)
* Write Chang Feng about my results (**check**)
* Basically, do the open Tab (make it a bookmark!) tutorial (**check**)

Questions:

* Clonality: Should I write mail to 10X kontaktperson?
* Differentially expressed genes for naïve signature only from Tn cluster or looking at all clusters
* Perhaps assign cluster labels again or try for all clusters to really find DEGs
* When we take top X up and top X downregulated genes to build the naïve signature, even though all of them are significant, ranked by padjusted value…
* What does it mean: Apply these naïve signatures to adult data to check for fetal remnants -> can this not happen by chance, even if we do find this to be the case? What exactly should I do in this step?
* Eventuell, falls hie rim annot Tn cluster nichts gefunden, dann noch bei anderen subsets suchen
* When can we say we really have remnants/ we found remnants when applying fetal sig to adult?
* Warum Unterteilung in Tn und Tn and Tcm, das sind doch unterschiedliche Kombis oder? Soll ich die top 20 für Tn and Tcm subset auch machen und mit Tn DEGs liste vergleichen und ne neue Signature liste erstellen? Oder soll ich von vorne herein schon Tn and Tcm als Tn bezeichnen und dann werden sie automatisch in DEG liste mit reingenommen?

Tell Christina:

1. I started with raw files, from GEO for fetus and for raw EX0004 adult healthy sample
2. Blockteil wie es lief and screen timer; vers. Blockteilgruppen
3. Aisha in Propra, Lily harte Zeiten, Propra damlas zero Erfahrung, es war sehr hart; Blockteil leichter (Mehr Erfahrung, Übung macht den Meister, Hiwi hilft sehr)
4. 2 Klausuren verpasst wegen Blockteil -> schreib sie zum Nachtermin wegen zeitlicher Überschneidung, aber keine Sorge, dafür genug Zeit zum Vorbereiten
5. GlobalProtect probleme gefixt