Zprávy od Ondry ohledně dat od Thomase Boehma

Ja myslim, ze je dulezite se podivat hlavne na opakujici se TCR

A podivat se, jak moc se ty opakujici TCR nachazeji pouze v CD8 nebo v CD4 (idealne napric mysmi) a jak moc jsou CD4 i CD8.

Muze byt relativne komplikovane prijit na nejlepsi zpusob, jak to kvantifikovat. V prvnim kroku bych se jen na ne podival. Predpokladam, ze maji UMI, ale je to bulk. Takze u stejne sekvence s ruznymi UMI nebudeme stejne schopni rict, jestli pochazi z jedne nebo vice bunek (a i kdyby z vice, tak to muze byt klonalni expanze v ramci te mysi).

Postupoval bych tak, ze bych tedy vytahnul vsechny neunikatni sekvence (tzn. zachyceny pod vice UMI) a pak se podival, jak jsou sdileny mezi vzorky a mezi CD8 a CD4.

Muzes pak udelat count table techto sekvenci, v radku bude sekvence ve slupci konkretni vzorek (CD4 1, CD4 2, CD4 3, CD8 1, CD8 2, CD8 3 atp.).

Pak by se z toho udelala pod tabulka, kde se vezmou jen ty sekvence, kde alespon ve 2 sloupcich nejsou nuly.

Na to bychom se podivali, a pak rozhodli, jestli je treba to kvantifikovat a pokud ano, tak jak a jestli a jak to normalizovat na celkovy pocet UMI na kazdy vzorek.