Project 2

Estimation of T-cell receptor diversity after personalized vaccination of glioblastoma patients

Data overview

Glioblastoma (GBM) is a very aggressive brain tumor with no standard therapy after recurrence and median patients' survival shorter than 12 months. In a recent study, GBM patients have been treated with personalized cancer vaccines to stimulate the recognition of tumor cells by T cells [1]. T cells can eliminate cancer cells after interaction of the T-cell receptor (TCR) with tumor antigens bound to the Human Leukocyte Antigen (HLA) molecules of tumor cells. TCR are composed of two different protein chains: alpha and a beta. In this study, blood samples collected from a responder patient before vaccination and after relapse were sequenced to reconstruct the TCR alpha and beta chains.

In the Supplementary Material of this article, is available a multi-sheet Excel file containing the estimated alpha or beta sequences ("Alpha/Beta CDR3 amino acid seq" column) and their counts ("UMI_counts"), before treatment and after relapse: https://static-content.springer.com/esm/art%3A10.1038%2Fs41586-018-0792-9/MediaObjects/41586_2018_792_MOESM8_ESM.xlsx. The more abundant sequences (i.e. with higher counts) reported on top, might represent clones of T cells with the same TCR that have expanded after recognition of a tumor antigens (e.g. after vaccination).

Analysis to be performed

Diversity indexes can be used to measure the compositional complexity of a set of TCR clones. *Diversity* increases when the number of different clones and with the evenness of their relative abundances. For a given number of different clones (also called *richness*), diversity is maximal when all clones are equally abundant.

Define two functions to compute the following indexes:

Richness ***R***, defined as the number of different TCR clones.

Diversity ***D***, computed with the Inverse- Simpson index:

$$D = \frac{1}{\sum_{i=1}^{R} {p_i}^2}$$

Where $i, \ldots, R$ are the different TCR clones and p_i are their relative counts, computed in R as:

```
pi <- UMI_counts/sum(UMI_counts)
```

Apply the functions to the four sets of TCR clones of the Excel table (alpha or beta chains, before treatment and after relapse).

Separately for alpha and beta chains, show the differences in richness and diversity of the samples before treatment and after relapse using barplots.

Separately for alpha and beta chains, calculate how many unique clones (with exactly the same sequence) are in common between the samples before treatment and after relapse.

Discuss the results considering the fact that anticancer vaccines might induce the expansion of antigen-reactive clones (i.e. augmentation of p_i for some clones) and/or appearance of new clones.

Report

The analysis above should be described in a short report that will be evaluated and considered for the final grade. The report should:

- Contain the **code** implemented to run the analysis above, together with the **results** (as tables, plots, or just numbers reported within the text) and their **description/discussion**.
- Be at **maximum 6 pages** long.
- Be saved in a Word doc named **"Surname_Name_Project3.doc"**.
- Be sent by e-mail to Dr. Finotello not later than **June 03, 2019**, using **"RProject3_2019"** as e-mail object.

References

[1] Keskin DB, Anandappa AJ, Sun J, Tirosh I, Mathewson ND, Li S, Oliveira G, Giobbie-Hurder A, Felt K, Gjini E, Shukla SA, Hu Z, Li L, Le PM, Allesøe RL, Richman AR, Kowalczyk MS, Abdelrahman S, Geduldig JE, Charbonneau S, Pelton K, Iorgulescu JB, Elagina L, Zhang W, Olive O, McCluskey C, Olsen LR, Stevens J, Lane WJ, Salazar AM, Daley H, Wen PY, Chiocca EA, Harden M, Lennon NJ, Gabriel S, Getz G, Lander ES, Regev A, Ritz J, Neuberg D, Rodig SJ, Ligon KL, Suvà ML, Wucherpfennig KW, Hacohen N, Fritsch EF, Livak KJ, Ott PA, Wu CJ, Reardon DA. ***Neoantigen vaccine generates intratumoral T cell responses in phase Ib glioblastoma trial***. Nature. 2019 Jan;565(7738):234-239.