**Immune repertoire annotation**

*Gainova Kristina, Ivanova Eugenia (Strebulaeva)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sample | Donor | Subset | Phenotype | CMVstatus |
| s13 | D1 | CD4 | naive | CMV+ |
| s2 | D1 | CD4 | naive | CMV+ |
| s1 | D1 | CD4 | memory | CMV+ |
| s5 | D1 | CD4 | memory | CMV+ |
| s15 | D2 | CD4 | naive | CMV- |
| s16 | D2 | CD4 | naive | CMV- |
| s10 | D2 | CD4 | memory | CMV- |
| s8 | D2 | CD4 | memory | CMV- |
| s9 | D1 | CD8 | naive | CMV+ |
| s14 | D1 | CD8 | naive | CMV+ |
| s3 | D1 | CD8 | memory | CMV+ |
| s7 | D1 | CD8 | memory | CMV+ |
| s11 | D2 | CD8 | naive | CMV- |
| s12 | D2 | CD8 | naive | CMV- |
| s4 | D2 | CD8 | memory | CMV- |
| s6 | D2 | CD8 | memory | CMV- |

**Detection of the replicas**

Firstly, we sorted our samples by replicas. This could be done on the basis of the clustering by variable segment usage profile - same replicas are grouping together clearly and also demonstrate high correlation level on the heatmap of the V-segment profile overlap and on the diversity histograms.

**Identification of the CD4+/CD8+ cells and CMV+ donor CD8+ cells**

It is known that CD4+ and CD8+ T-cells have different target receptors - CD4+ cells recognize MHCII receptor whereas CD8+ cells recognize MHCI, so their receptor variable segment profiles should be quite distinct. And on the V usage profile heatmap we saw division on two clusters. It is also known that it is difficult to distinguish CD8+ and CD4+ profiles between two donors, so we can consider those clusters as CD4+ and CD8+ and not D1 and D2.

To define which cluster belongs to CD8+ cells we looked on the histogram for the HLA recognition. s3/s7 samples have the highest specificity to HLA B\*07 forming MHCI, so they are supposed to belong to CD8+ cells as all the samples in the cluster with s3/s7. Besides, we looked at the histogram of the antigen recognition and found that s3/s7 samples have high specificity to CMV, which means the donor of these samples was CMV+ and low diversity in these samples defines them as memory cells.

**Identification of the donors**CD8+ TCR repertoires are different between donors, as are the CD4+ TCR repertoires. This fact allows us to suggest that CD4+ and CD8+ repertoires form separate clusters for different donors on the heatmap of the V-segment profile overlap. For example, s9/s14 and s3/s7 samples, assigned by us as CD8+ cells, form one cluster, while s11/s12 and s4/s6, assigned also as CD8+ cells, form another cluster.

**How we differed mature and naive T cells**

Firstly, we analyzed “heatmap with dendrogram” (page 9)) and found 4 groups (in each 4 samples where 2 samples are replicas) which have common ancestors. Hence, all of these groups contain mature cells and naive cells which have the same cell type - CD4/8 (from one donor). It is known that the naive T-cell repertoire is highly diverse, providing the ability to recognize the universe of exogenous and potentially dangerous antigens. Thus, to differentiate naive and mature T-cells we analyzed “plot parent species of putative antigen recognized by TCR”. Those samples which contain less recognized species (were less diverse) were assigned to mature T-cells and the second pair of cells were assigned to naive cells.