Correlation of PRS with cell type composition after ILR transformation provides new insights into the relationship between immune-related diseases and PBMCs cell type proportions

Giuditta Clerici

Giuditta Clerici[1], Costanza Cantalini[1], Federica Santonastaso[1], Davide Bolognini[1], Matiss Ozols[2], Alessandro Raveane[1], Cecilia Dominguez Conde[1], Edoardo Giacopuzzi[1], Nicola Pirastu[1], Nicole Soranzo[1,2,3]

1. Motivation

Many diseases, such as Obesity, Type 1 Diabetes Mellitus, Asthma, Inflammatory Bowel Disease, show immune-related alterations. These alterations can result in an imbalanced composition of Peripheral Blood Mononuclear Cell (PBMCs) that can be measured by scRNA-seq experiments.

For most of these disorders, the involvement of specific PBMC types remains unclear.

2. Aims of the project

- ❑ Provide a **novel and robust method** to study the relationship between PRS and cell type composition estimated from scRNA-seq experiments
- ❑ Study the influence of **genetic predisposition to inflammatory traits** on PBMCs compositions

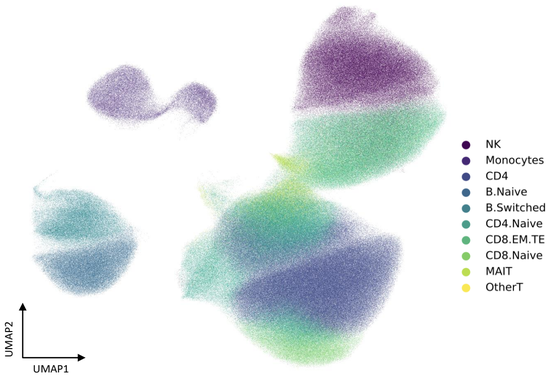
3. Data & Methods

OneK1K cohort[1]
Blood from 982 donors with Northern European ancestry

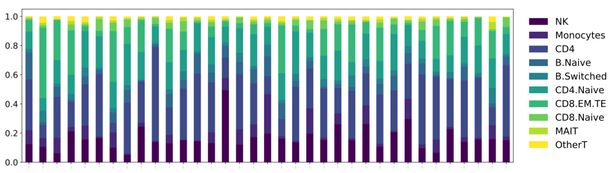
~ 1.27 million PBMCs

mRNA capture and sequencing

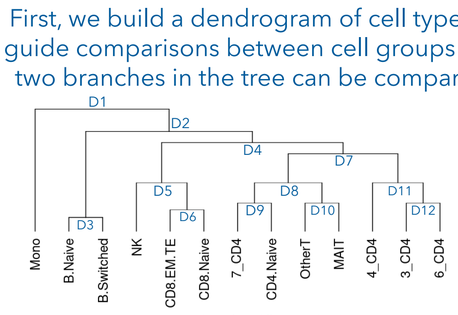
We applied the most recent best practices for single-cell quality control[2] and annotated cell types using **Celltypist**



Since compositional data is intrinsically biased, we applied **ILR** (isometric log-ratio)[3] transformation on cell type proportions



First, we build a dendrogram of cell types to guide comparisons between cell groups (any two branches in the tree can be compared)



* CD4 are in separate groups because they derive from different Leiden clusters
** OtherT include: gamma-deltaT, ILC, IL22

	Mono	B Naive	B Switched	NK	CD8 EM & TE	CD8 Naive	7_CD4*	CD4 Naive	OtherT**	MAIT	4_CD4*	3_CD4*	6_CD4*
D1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
D2	0	1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
D3	0	1	-1	0	0	0	0	0	0	0	0	0	0
D4	0	0	0	1	1	1	-1	-1	-1	-1	-1	-1	-1
D5	0	0	0	1	-1	-1	0	0	0	0	0	0	0
D6	0	0	0	0	1	-1	0	0	0	0	0	0	0
D7	0	0	0	0	0	0	1	1	1	1	-1	-1	-1
D8	0	0	0	0	0	0	1	1	-1	-1	0	0	0
D9	0	0	0	0	0	0	1	-1	0	0	0	0	0
D10	0	0	0	0	0	0	0	0	1	-1	0	0	0
D11	0	0	0	0	0	0	0	0	0	0	1	-1	-1
D12	0	0	0	0	0	0	0	0	0	0	0	1	-1

Each comparison is transformed into a ILR direction

Data for PRS weights
GWAS catalog[4]
PGS catalog[5]

Validation traits
- Monocytes counts
- Lymphocytes count

Traits tested
- Inflammatory Bowel Disease (IBD)
- Ulcerative Colitis (UC)
- Crohn's Disease
- Asthma
- Type 1 Diabetes Mellitus
- Body Mass Index (BMI)

Genomic DNA and SNP genotyping

w

X

$$PRS = \sum_{i=1}^{n} w_i X_i$$

PRS

Donors

Linear model fitting

$$ILR\ directions \sim age + sex + pool_{group} + PRS$$

with 926 donors

ILR directions

Donors

Finally, we compute the weighted sum of proportions for each donor *i*

$D1_i = pA_i * wA1 + pB_i * wB1 + pC_i * wC1 + pD_i * wD1 + pE_i * wE1$

$D2_i = pA_i * wA2 + pB_i * wB2 + pC_i * wC2 + pD_i * wD2 + pE_i * wE2$

...

4. Results

Validation traits

$ILR\ directions \sim age + sex + pool_{group} + PRS_{Lymphocyte_{count}}$

The PRS for **lymphocyte** count is **negatively** related to **D1** (p.val = 0.040, t.val = -2.05)

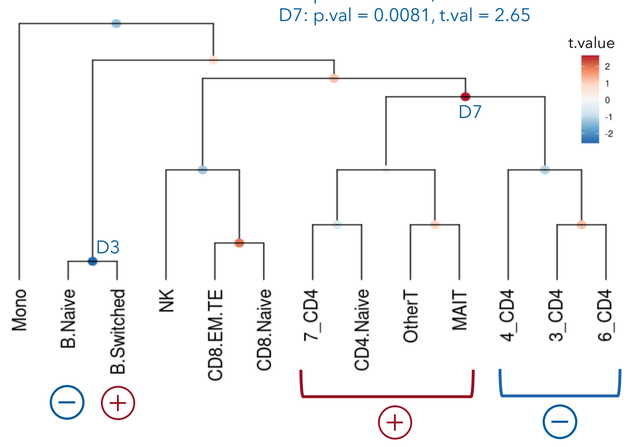
$ILR\ directions \sim age + sex + pool_{group} + PRS_{Monocyte\_count}$

The PRS monocyte count is **positively** related to the ratio monocytes/lymphocites (p.val = 0.05, t.val = 1.94)

Traits tested

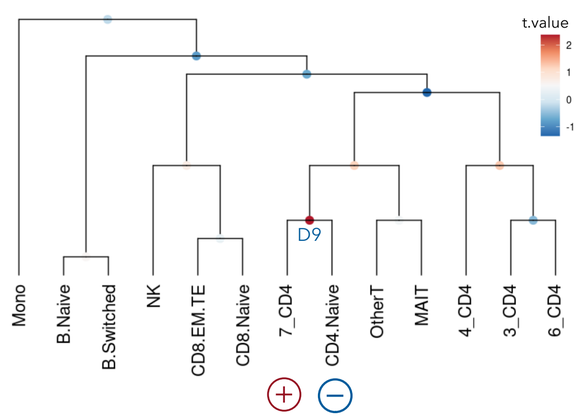
$ILR\ directions \sim age + sex + pool\_group + PRS_{BMI}$

D3: p.val = 0.0098, t.val = -2.58
D7: p.val = 0.0081, t.val = 2.65



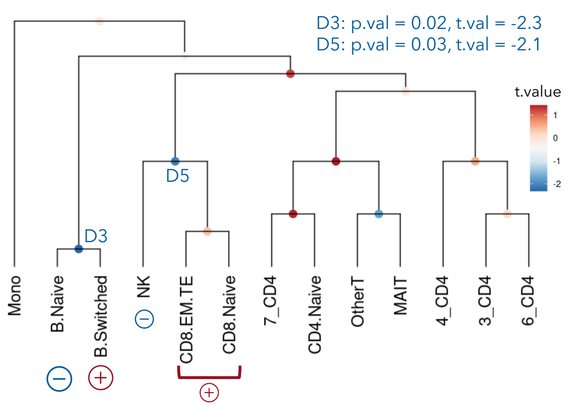
$ILR\ directions \sim age + sex + pool\_group + PRS_{IBD}$

D9: p.val = 0.02, t.val = 2.4



$ILR\ directions \sim age + sex + pool\_group + PRS_{T1DM}$

D3: p.val = 0.02, t.val = -2.3
D5: p.val = 0.03, t.val = -2.1



How to read the plots

Each dot (**node**) on the dendrogram represent a **contrast or direction** between two groups of cells. The colors represent the **t.value** of the association between that ILR Direction and the PRS for the trait of interest, with red for positive and blue for negative. If there is a **positive relationship** (red dot), then you can assume that the cell type on the left of that node increase (respect to the one on the right) together with the increase in PRS for the trait tested.

5. Conclusions and future perspectives

In summary, our approach provides a novel and robust method to correlate a phenotype of interest and cell-type composition. We discovered some interesting associations between disease PRS and alteration in specific cell type proportions. This can lead to new insights into the immune-related component of these diseases and the mechanisms of immune dysregulation.

Future perspectives

- ❑ Test our method on a bigger dataset
- ❑ Extend the analysis to other phenotypes
- ❑ Wrap the method in a package

giuditta.clerici@external.fht.org
@GiudittaClerici

Affiliations

1. Human Technopole, Milan, Italy
2. Wellcome Sanger Institute, Hinxton, UK
3. University of Cambridge, Cambridge, UK

References

1. Yazar S. et al, Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease (2022)
2. Lukas Heumos et al, Best practices for single-cell analysis across modalities (2023)
3. Vera Pawlowsky-Glahn, Juan José Egozcue, Raimon Tolosana-Delgado, Modeling and Analysis of Compositional Data (2015)
4. GWAS catalog GCST90002393 and GCST90002388
5. PGS catalog PGS003981, PGS002066, PGS001331, PGS002727, PGS002025,