TCR repertoire as biomarker of immune diseases : applications to autoimmune diseases and COVID-19

Paul Stys, Pierre Barennes, Michèle Barbié, Roberta Lorenzon, Marie Surroque, Alexandra Roux, Michelle Rozenzwagj, Adrien Six, Encarnita Mariotti-Ferrandiz, David Klatzmann

TCR repertoire as biomarker of immune diseases : applications to autoimmune diseases and COVID-19

Paul Stys, Pierre Barennes, Encarnita Mariotti-Ferrandiz, David Klatzmann

TCR repertoire as biomarker of immune diseases : applications to autoimmune diseases and COVID-19

Paul Stys[1★], Pierre Barennes[1,2★], Michèle Barbie[2], Roberta Lorenzon[2], Marie Surroque[2], Alexandra Roux[2], Transimmunom Consortium, Michelle Rozenzwagj[1,2], Adrien Six[1], Encarnita Mariotti-Ferrandiz[1♦], David Klatzmann[1,2♦]

1. Sorbonne Université, INSERM, Immunology-Immunopathology-Immunotherapy (i3), Paris, France ; 2. AP-HP, Hôpital Pitié-Salpêtrière, Biotherapy (CIC-BTi) and Inflammation-Immunopathology-Biotherapy Department (i2B), Paris, France

★ These authors contributed equally to this work ; ♦ These authors contributed equally to this work

1 Background

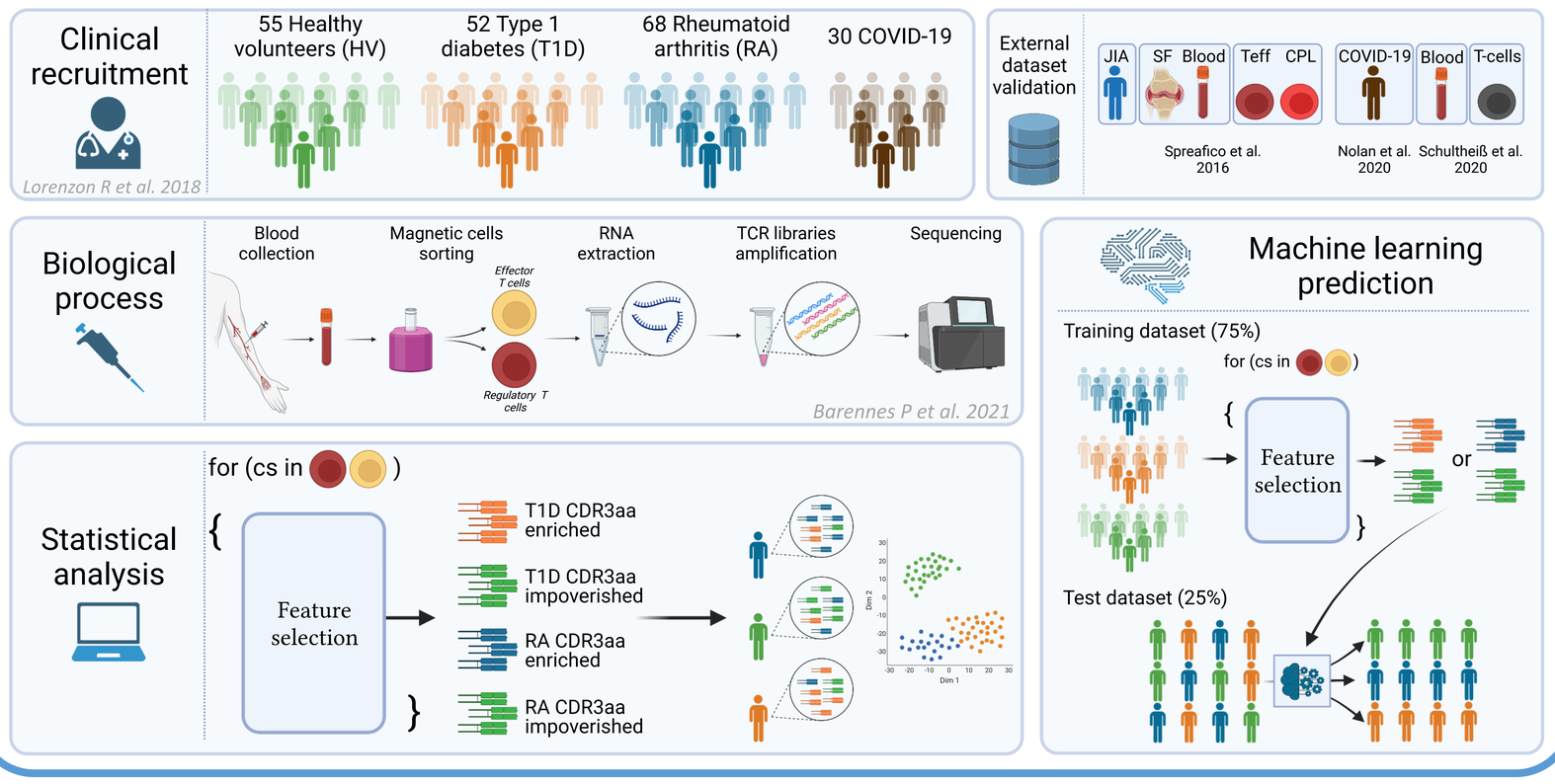
Adaptative T-cell repertoire diversity

The T-cell repertoire (TCR) is a dynamic biological object whose modifications will dependent on cell populations and the amplitude of the response according to the types of antigens encountered (microbial, self-antigens). We launched two observational trial, TRANSIMMUNOM (NCT02466217) to revisit the nosology of autoimmune diseases (AIDs) and SirocCo (ANR-21-CO12-0005) to investigate the mechanisms of cell subsets in COVID-19 patients. For each trial, we sorted the CD4 T-cell subpopulations from peripheral blood: effector T-cells (Teff) and regulatory T-cells (Treg), and sequenced their TCR$\alpha\beta$ repertoire.

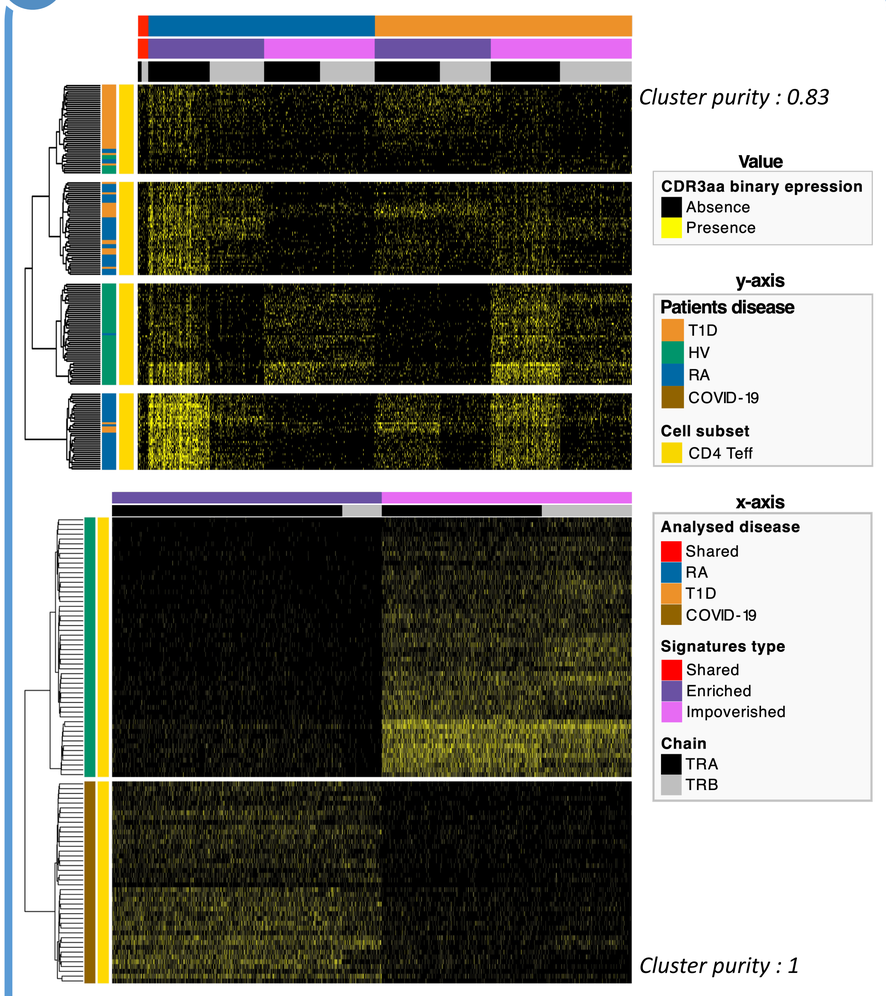
AIMS

- **Identify disease-specific TCRs** from the repertoire of sorted T cells
- Determine the **specificity and biological roles** of these TCRs to better understand the diseases
- Use these sequences to **classify and predict** patient outcome

2 Methods



3 Signatures

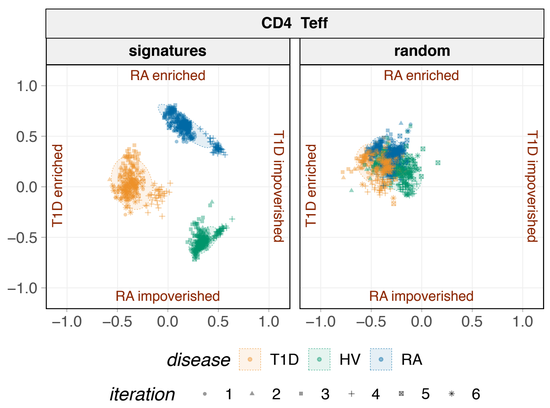


Ward clustering methods applied on AID (top) and COVID-19 (bottom) signatures discriminate samples with the same unique CDR3aa sequences from each identified signature (here, for CD4 Teff alpha and beta chains, there is one enriched and one depleted signature for each disease). The selection of CDR3aa works well on a viral disease like **COVID-19** since the **clustering shows a perfect separation** of patients from HV. The **strategy also separates AIDs patients from HVs** but to a lesser extent.

4 Signatures validation

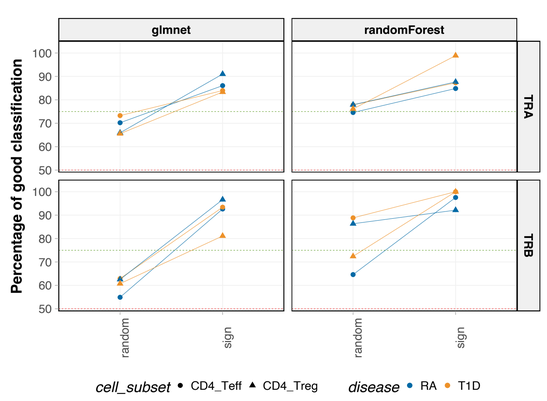
Laboratory dataset

Random



The radviz represents the CDR3aa expression of the 4 AID signatures for each patient. The signatures were calculated on 75% of the dataset of each disease. 6 random draws of the dataset were made (iterations). Finally, for each iteration and each signature, the equivalent number of unique CDR3aa was randomly selected from the corresponding dataset. The **random signatures do not separate the patients**, regardless of the iterations, which shows that the **strategy is reproducible** and that the results are **not due to chance**.

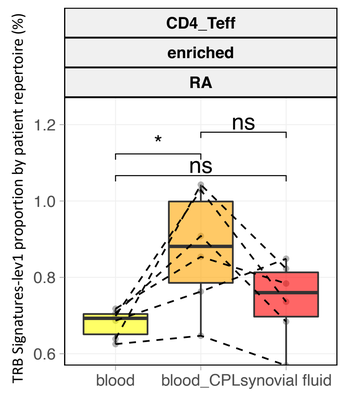
Classification



Machine learning models were trained on the signatures obtained on 75% of the dataset. The models were then tested on the remaining 25% of the data. The results are compared to a random selection of unique CDR3aa whose size is equivalent to the signature of each condition (disease and cell subsets). The boxplots represent the percentage of patients, from the naive data, correctly classified.
In all cases, the models derived from the signatures predict better than random signature. Moreover, the **signatures correctly predicted the disease of a patient in 81% to 100%** depending on the conditions.

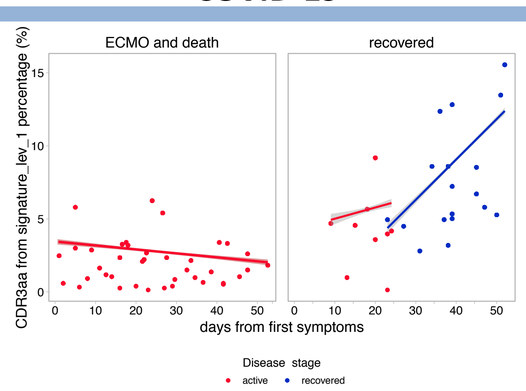
External dataset

RA



The RA Teff β signature was searched in the Spreafico et al. dataset with a Leveinstein distance of 1. Samples were taken from blood and synovial fluid. Circulating pathogenic lymphocytes (CPL) identified by Spreafico in juvenile idiopathic arthritis (JIA) patients are correlated with disease activity. The RA signature from Transimmunom is statistically more present in CPL compared to whole blood but also more found in synovial fluid than in whole blood without being significant. The **signature is thus found in another dataset**, treated differently, and also **assimilated to the disease**.

COVID-19



A signature was identified on patients recovering from COVID-19 on the Nolan et al. dataset. It was compared to the Schultheiß et al. dataset with a Leveinstein distance of 1. This dataset has the particularity to contain patients who died from COVID-19 or had severe forms (ECMO) and patients in recovery. The identified signature is much more found in the latter, and more and more over time. This allows us to **predict patient outcomes**.

5 Conclusions & perspectives

- ✓ We developed a general pipeline that **identifies discriminant TCR signatures from circulating blood**
- ✓ This strategy works well for a viral disease (COVID-19) as well as for several AIDs **applied to Teff as well as Treg** TCR repertoires. **Analyzing cell subtypes will allow a better understanding** of the underlying mechanisms.
- ✓ We show how a signature can be **used as a biomarker of immune diseases and their progression**
- ❑ We still need to characterize the specificity of these signatures

References

- Lorenzon R, Mariotti-Ferrandiz E, Aheng C, et al (2018) Clinical and multi-omics cross-phenotyping of patients with autoimmune and autoinflammatory diseases: the observational TRANSIMMUNOM protocol BMJ Open 2018;8:e021037
- Barennes, P., Quiniou, V., Shugay, M. et al. (2021) Benchmarking of T cell receptor repertoire profiling methods reveals large systematic biases. Nat Biotechnol 39, 236–245
- Spreafico R, Rossetti M, van Loosdregt J, et al. A circulating reservoir of pathogenic-like CD4+ T cells shares a genetic and phenotypic signature with the inflamed synovial micro-environment. Ann Rheum Dis. 2016;75(2):459-465. doi:10.1136/annrheumdis-2014-206226
- Nolan, Sean et al. “A large-scale database of T-cell receptor beta (TCRβ) sequences and binding associations from natural and synthetic exposure to SARS-CoV-2.” Research squarers.3.rs-51964. 4 Aug. 2020, doi:10.21203/rs.3.rs-51964/v1. Preprint.
- Schultheiß, Christoph et al. “Next-Generation Sequencing of T and B Cell Receptor Repertoires from COVID-19 Patients Showed Signatures Associated with Severity of Disease” Immunity, Volume 53, Issue 2, 2020, Pages 442-455.e4, ISSN 1074-7613, doi:10.1016/j.immuni.2020.06.024.