Understanding repertoire sequencing data through a multiscale computational model of the germinal center

Rodrigo García-Valiente[1,2,3#*], Elena Merino Tejero[1,2,3#], Maria Stratigopoulou[4,7], Daria Balashova[1,2,3], Aldo Jongejan[1,2,3], Danial Lashgari[1,2,3], Aurélien Pélissier[5,6], Tom G. Caniels[7,8], Mathieu A. F. Claireaux[7,8], Anne Musters[9,10], Marit J. van Gils[7,8], María Rodríguez Martínez[5], Niek de Vries[9,10], Michael Meyer-Hermann[11,12], Jeroen E.J. Guikema[13,4], Huub Hoefsloot[14], Antoine H.C. van Kampen[1,2,3,14]

[1]Amsterdam UMC location University of Amsterdam, Epidemiology and Data Science, Meibergdreef 9, Amsterdam, Netherlands; [2]Amsterdam Public Health, Methodology, Amsterdam, The Netherlands; [3]Amsterdam Infection and Immunity, Inflammatory Diseases, Amsterdam, The Netherlands; [4]Cancer Center Amsterdam, Amsterdam, The Netherlands; [5]IBM Research Zurich, 8803 Rüschlikon, Switzerland; [6]Department of Biosystems Science and Engineering, ETH Zurich, 4058 Basel, Switzerland; [7]Amsterdam UMC location University of Amsterdam, Medical Microbiology and Infection Prevention, Meibergdreef 9, Amsterdam, Netherlands; [8]Amsterdam Infection and Immunity, Infectious Diseases, Amsterdam, The Netherlands; [9]Amsterdam UMC location University of Amsterdam, Experimental Immunology, Meibergdreef 9, Amsterdam, Netherlands; [10]Amsterdam Rheumatology & Immunology Center, Amsterdam, The Netherlands; [11]Department for Systems Immunology and Braunschweig Integrated Centre of Systems Biology, Helmholtz Centre for Infection Research, Braunschweig, Germany; [12]Institute for Biochemistry, Biotechnology and Bioinformatics, Technische Universität Braunschweig, Braunschweig, Germany; [13]Amsterdam UMC location University of Amsterdam, Pathology, Lymphoma and Myeloma Center Amsterdam, Meibergdreef 9, Amsterdam, Netherlands; [14]Biosystems Data Analysis, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, the Netherlands

[#]These authors contributed equally

*r.garciavaliente@amsterdamumc.nl

#1 INTRODUCTION

- **Germinal centers**[1] (GCs) are microanatomical structures found in secondary lymphoid organs and are formed when an **adaptive response** is initiated against an antigen (Ag).
- The GC reaction begins with the activation of a limited number of Ag-specific B cells (founder clones) that start to proliferate (**clonal expansion**) to form the so-called GC dark zone (DZ). During the proliferation of these B cells, now called centroblasts (CBs), their BcR is changed due to **somatic hypermutations** (SHMs), which increase or decrease the binding affinity of the BcR for the Ag. A **clone represents a (large) lineage** of B cells comprising **subclones with unique BcRs** created by SHMs, all stemming from the same unmutated common ancestor.
- The CBs differentiate to centrocytes (CCs) and migrate to the GC light zone (LZ) where they collect Ag presented by follicular dendritic cells (FDCs) and, subsequently, interact with T follicular helper (Tfh) cells to become **positively selected** to return to the DZ to undergo further rounds of proliferation and SHM (**affinity maturation**). This cycle continues during the span of several weeks. **Memory B cells** (MBCs) and **plasma cells** (PCs) are output cells from the GC. **PCs express up to 100 times more BcR** than the other cells.
- **Sequencing of B cell immune receptor repertoires** helps us to **understand the adaptive immune response**, **although only provides information about the clonotypes and their frequencies**. Generally, **additional time-consuming or expensive experiments are required to further characterize the identified (dominant) clones** by measuring, for example, their affinity or function. Here, **we present a multiscale model of the germinal center to gain general insight in the interpretation of B cell repertoires by establishing:**
  - **1/ the relationship between clonal abundance and affinity**
  - **2/ the variability of affinity within a clone**
  - **3/the extent to which PCs with high BcR mRNA content may disturb the identification of dominant clones in RNA-Seq repertoires**

#2 METHODS

We extended and modified our previous **GC multiscale model**[2] to include:

- **BcR representation**, so every cell has its own BcR Fab sequence with its own (theoretical) affinity.
- **SHM fate tree**, so every BcR can mutate according to experimental results[3]. These mutations have a different effect (affinity change, lethal, neutral) depending on their type (replacement, R, or silent, S) and the region of the BcR where they take place (complementarity-determining regions, CDRs, that bind the antigen, or framework regions, FWRs, that are structural).

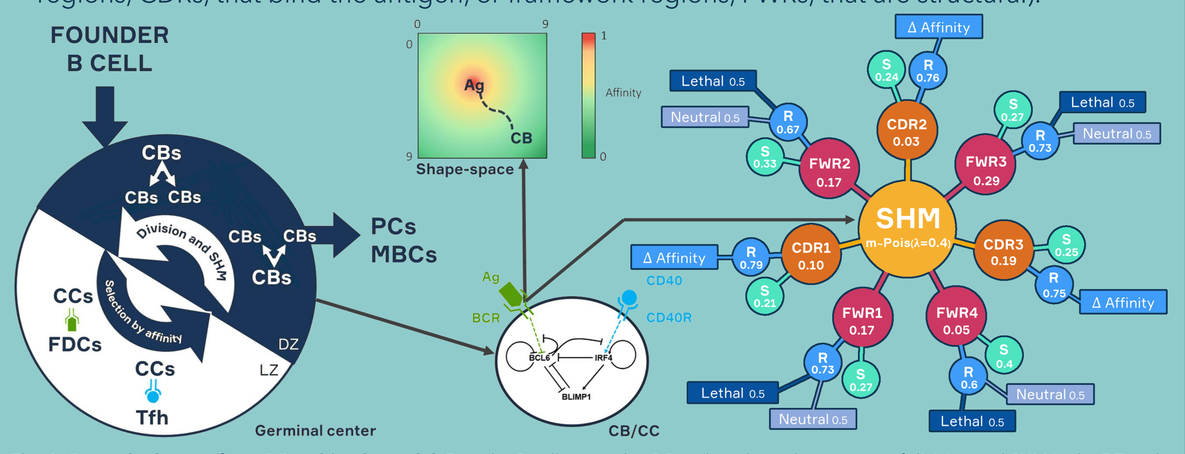

Fig. 1. General scheme of our GC multiscale model. Founder B cells enter the GC and go through a process of division and SHM in the DZ and selection in the LZ, based on the affinity of their BcRs per a theoretical antigen. Our multiscale model includes a GRN that drives PC differentiation. The affinity of the BcRs is based on the distance between the BcR sequence and the optimal BcR in a continuous shape-space. After each cell division m mutations happen on a daughter cell. The SHM fate tree shows the probabilities of a mutation happening on each region, the probability of it being of a specific type and their consequent effect on the cell.

#3 RESULTS

Affinity and abundance

Can we safely assume that the higher the affinity of a (sub)clone, the higher its abundance and viceversa?

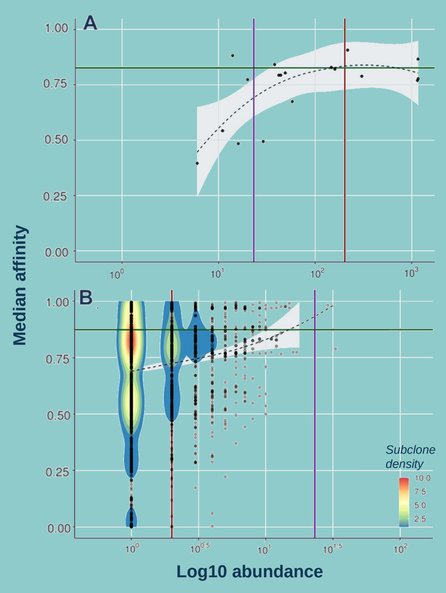

Fig. 2. Relation between clone (A) and subclone (B) abundance and median affinity at day 21 of the GC reaction for a representative simulation. Each dot represents a (sub)clone. Horizontal green line denotes the 75th percentile threshold. Vertical red line denotes the 75th percentile threshold. Vertical purple line denotes the 0.5% of the total cell counts threshold. Black dotted line denotes a lowess fit. The density map represents the concentration of subclones.

DNA-Seq vs RNA-Seq

Does the higher BcR mRNA content in PCs affect the determination of dominant clones?

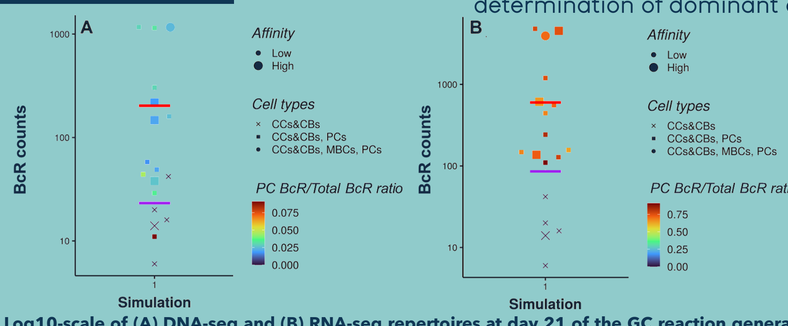

Fig. 3. Log10-scale of (A) DNA-seq and (B) RNA-seq repertoires at day 21 of the GC reaction generated by a representative simulation. We assume that PCs express 100-fold higher levels of BcR mRNA. Each dot represents a clone, some of which are a mixture of B cells, MBCs and/or PCs. In both cases we find 5 and 13 dominant clones using the >=75th percentile (red horizontal line) or >=0.5% of the counts (purple horizontal line) as a threshold, respectively. Dot colors indicate the fraction of PC BcR sequences within each clone, whose range is about a factor of 100 times larger for the RNA-based repertoire. The size of the symbol represents the median affinity of that clone (small symbol: affinity below the 75th percentile).

Intraclonal affinity variance

Can we expect all subclones within a clone to have a similar affinity?

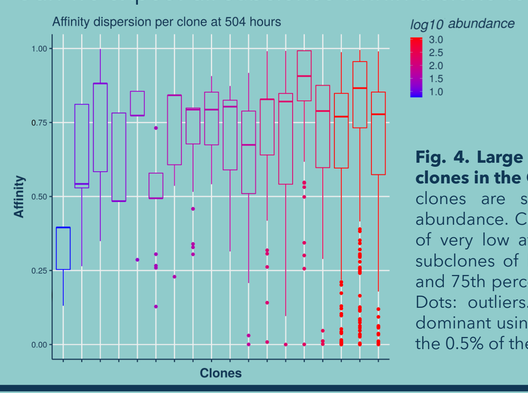

Fig. 4. Large variation in subclonal affinity for 18 surviving clones in the GC at day 21 of a representative simulation. The clones are sorted in ascending order according to their abundance. Clones with a higher abundance include subclones of very low affinity, while clones of low abundancy may have subclones of high affinity. Horizontal line: median. Boxes: 25th and 75th percentiles. Whiskers: 1.5 times the interquartile range. Dots: outliers. The top 5 and the top 13 of the clones are dominant using as threshold the 75th percentile of the clones or the 0.5% of the cell counts, respectively.

Comparison with experimental results

How accurately can we reproduce experimental data?

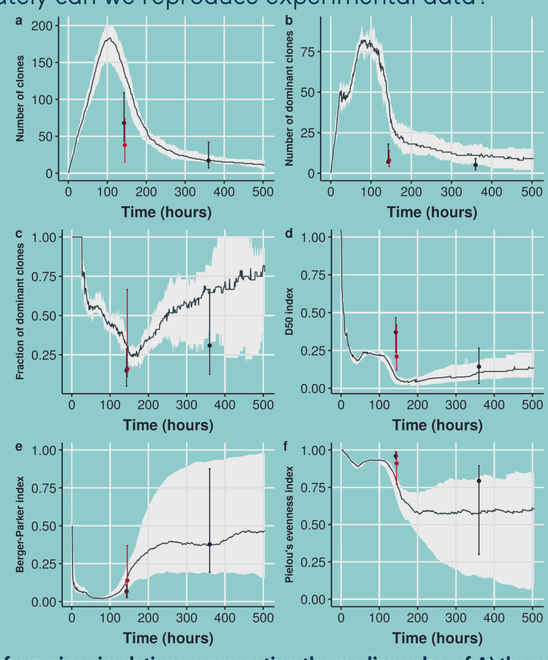

Fig. 5. Results from nine simulations representing the median value of A) the number of clones, B) the number of dominant clones; C) fraction of dominant clones; D) D50 index; E) Berger-Parker index, and F) Pielou's evenness index over time, compared to experimental data. The minimum and maximum values of all the simulations are delimited by the shadowed areas. Dominant clones were defined the clones accounting for at least as 0.5% of the repertoire. The dots and associated bars represent the median and maximum and minimum values for single-cell RNA-Seq samples from mice steady-state specific pathogen-free (SPF) gut-associated GCs[4] at their measured timepoints (black: GC data from mice immunized with chicken gamma globulin; red: data from mice immunized with ovalbumin, hemagglutinin or ovalbumin conjugated with 4-Hydroxy-3-nitrophenylacetyl hapten). The experimental values are close to the values in the corresponding timepoint of our simulation results.

References

[1]De Silva, N. S. & Klein, U. Dynamics of B cells in germinal centres. Nat Rev Immunol 15, 137-148, doi:10.1038/nri3804 (2015).
[2]Merino Tejero, E. et al. Multiscale Modeling of Germinal Center Recapitulates the Temporal Transition From Memory B Cells to Plasma Cells Differentiation as Regulated by Antigen Affinity-Based Tfh Cell Help. Front Immunol 11, 620716, doi:10.3389/fimmu.2020.620716 (2020).
[3]Cui, A. et al. A Model of Somatic Hypermutation Targeting in Mice Based on High-Throughput Ig Sequencing Data. J Immunol 197, 3566-3574, doi:10.4049/jimmunol.1502263 (2016).
[4]Nowosad, C. R. et al. Tunable dynamics of B cell selection in gut germinal centres. Nature 588, 321-326, doi:10.1038/s41586-020-2865-9 (2020).

- There is a **limited correlation between (sub)clonal abundance and affinity**
- There is **large affinity variability among same-ancestor subclones**
- **PCs do not significantly affect the number and identification of dominant clones** in single GCs by sequencing BcR mRNAs

Acknowledgments

This work is supported by COSMIC (https://www.cosmic-h2020.eu) which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 765158, by ARCAID (https://www.arcaid-h2020.eu), which has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 847551, and by the Human Frontier Science Program 570 (RGP0033/2015). We thank Davide Angeletti (University of Gothenburg) for sharing information about his recent experiments.