

## Supplementary Materials and Methods

### **T cell receptor repertoire characteristics both before and following immunotherapy correlate with clinical response in mesothelioma**

Vroman H.<sup>1,2\*</sup>, Balzaretto G.<sup>3,4\*</sup>, Belderbos R.<sup>1,2</sup>, Klarenbeek P.L.<sup>3,4</sup>, van Nimwegen M.<sup>1</sup>, Bezemer K.<sup>1</sup>, Cornelissen R.<sup>1</sup>, Niewold I.T.G.<sup>5</sup>, van Schaik B.D.C.<sup>6</sup>, van Kampen A.H.C.<sup>6</sup>, Aerts J.G.J.V.<sup>1,2</sup>, de Vries N.<sup>3</sup>, Hendriks R.W.<sup>1</sup>

<sup>1</sup>Department of Pulmonary Medicine, Erasmus Medical Center Rotterdam, 's-Gravendijkwal 230, 3015 CE, Rotterdam, The Netherlands.

<sup>2</sup>Erasmus MC Cancer Institute, Erasmus Medical Center Rotterdam, 's-Gravendijkwal 230, 3015 CE, Rotterdam, The Netherlands.

<sup>3</sup>Department of Clinical Immunology & Rheumatology, Amsterdam Rheumatology and Immunology Center (ARC), Amsterdam UMC, location AMC, University of Amsterdam, Meibergdreef 9, 1105AZ, Amsterdam, the Netherlands.

<sup>4</sup> Department of Experimental Immunology | Amsterdam Infection & Immunity Institute, Amsterdam UMC, location AMC, University of Amsterdam, Meibergdreef 9, 1105AZ, Amsterdam, The Netherlands.

<sup>5</sup> Laboratory of Genome Analysis | Amsterdam UMC, location AMC, University of Amsterdam, Meibergdreef 9, 1105AZ, Amsterdam, The Netherlands.

<sup>6</sup> Bioinformatics Laboratory, Department of Clinical Epidemiology, Biostatistics, and Bioinformatics | Amsterdam UMC, location AMC, University of Amsterdam, Meibergdreef 9, 1105AZ, Amsterdam, The Netherlands

\*Shared 1<sup>st</sup> authorship

### ***Study Design***

The 9 MPM patients described in this study were enrolled in a first-in-human open-label nonrandomized dose escalation phase I study<sup>1</sup>. Briefly, MPM patients were included at least 6 weeks after their last chemotherapy treatment, or received no previous treatment if they refused chemotherapy treatment. Upon inclusion in the study, patients underwent leukapheresis. The leukapheresis product was used as source for the generation of autologous monocyte-derived DCs (moDCs) that were subsequently loaded with a allogeneic tumor cell lysate consisting of a mixture of five *in vitro* cultured MPM cell lines<sup>1</sup>. Patients received three vaccinations biweekly, and two booster vaccinations after 3 and 6 months. For the dose escalation study, patients 1-3 received  $10 \times 10^6$  moDCs per vaccination, patients 4-6 received  $25 \times 10^6$  moDCs per vaccination, and patients 7-9 received  $50 \times 10^6$  moDCs per vaccination. Patients 7 and 9 did not receive a third booster vaccination due to shortage of the generated moDCs. All other patients completed the full treatment.

Ethylene diamine tetra acetic acid (EDTA)-anticoagulated peripheral blood was drawn from all patients at baseline prior to the first vaccination (pre) and 5 weeks after the first vaccination (post). Peripheral blood mononuclear cells (PBMCs) were isolated from the peripheral blood by standard Ficoll density gradient centrifugation and stored at  $-196^\circ\text{C}$  for further analysis.

### ***Flowcytometric sorting of T cell subsets***

Stored PBMCs were stained for extracellular markers, CD3-APC-eF780 (eBioscience, San Diego, CA, USA), CD4-FITC (BD Biosciences, San Jose, CA, USA), CD8-PerCP-Cy5.5 (eBioscience), CD25-PE-Cy7 (BD Biosciences), CD127-PE (BD Biosciences), PD1-APC (Biolegend, San Diego, CA, USA). DAPI (Invitrogen, Carlsbad, CA, USA) was used as live/dead marker. T cell subsets were sorted using a FACS Aria equipped with BD FACS

Diva software (BD Biosciences). Cells were selected on negativity for DAPI. Doublets were depleted using side scatter and forward scatter width and height. First, 500.000 CD3<sup>+</sup> T cells were sorted, followed by 50.000 cells per T cell subset fraction, gated as depicted in **Fig. S1**. Sorted T cell fractions were collected in RNA lysis buffer (Qiagen, Venlo, the Netherlands) containing 2-mercaptoethanol and stored at -80°C for RNA extraction.

### *Analysis of T-cell receptor repertoires*

**Next-Generation Sequencing.** To increase the yield of the RNA extraction procedure of the collected sorted T cell fractions, 10,000 non-TCR expressing HEK293T cells were added to each sample prior to RNA isolation (RNeasy Micro Kit, Qiagen). RNA-sequencing of the TCR $\beta$  chains was performed with a modified version of a custom protocol described previously<sup>2</sup>, and adapted in order to obtain unique molecular identifier (UMI)-tagged products. In short, specific complementary DNA (specific-cDNA) of TCR $\beta$  molecules was synthesized using a TCR  $\beta$ -chain Constant region reverse primer tagged with a 9 random nucleotide UMI and a consensus sequence. After specific-cDNA synthesis, Exonuclease I (Thermo Fisher Scientific, Breda, The Netherlands) treatment was performed to remove left over primers, followed by a multiplexed PCR with 23 forward primers covering all TCR  $\beta$ -chain Variable genes and a reverse primer binding to the consensus sequence previously introduced in the specific-cDNA and tagged with an 8 bp patient identifier (MID, Molecular Identifier). Obtained amplicons were purified using two rounds of AMPure XP beads clean-up (Beckman Coulter, Woerden, The Netherlands), quantified using Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific), dual-indexed with i5 and i7 adapters (Nextera XT Index Kit v2) and sequenced using the Illumina Miseq Kit v3 2 x 300 bp technology according to the manufacturer's manual (Illumina, San Diego, California, USA).

**Bioinformatic generation of TCR $\beta$  repertoires.** The obtained sequencing reads were analyzed with “RESEDA” (REpertoire SEquencing Data Analysis, <https://bitbucket.org/barbera/reseda>), using the following steps: 1) pairwise assembly of the paired-end reads using PEAR<sup>3</sup>, 2) identification of the 8 bp MID, 3) identification of the CDR3, 4) alignment to the IMGT database<sup>4</sup> to obtain the Variable and Joining gene assignment, 5) removal of reads with low quality bases (Q score < 30) in the CDR3, 6) clustering of reads in clones based on 100% amino acidic CDR3 identity, 7) UMI-based correction of clonal frequencies and 8) contamination check between samples from different individuals.

**Analysis of TCR $\beta$  clones.** The list of final clones generated by RESEDA was analyzed using in-house developed scripts in R studio (R version 3.3.2). The absolute number of mRNA molecules was indicated by the number of unique UMIs, hereafter referred to as reads. The frequency of each clone was calculated as percentage of the total number of UMIs. Clones with a frequency  $\geq 0.5\%$  were defined highly expanded clones (HECs), based on previous observations<sup>2</sup>. We defined the impact of a clone as its frequency in the repertoire, and the impact of a group of clones as their cumulative frequency. The repertoire diversity was calculated with the Chao-modified Sørensen index (CMSI) as described earlier<sup>5-7</sup>, with a value of 1 indicating perfect overlap between two samples and 0 indicating no overlap. For clones that were shared between the pre- and post-DC-therapy total CD3 T cell samples, a  $\log_2(\text{Fold change (FC)}) \geq 3$  defined the expanding clones (EXP) and a  $\log_2(\text{FC}) \leq -3$ , defined the decreasing clones (DEC). The remaining clones were designated stable. For the sorted five T cell subpopulations (PD1<sup>-</sup> and PD1<sup>+</sup> CD4 and CD8 T cells, and Tregs), thresholds of  $\log_2(\text{FC}) \geq 2$  and  $\log_2(\text{FC}) \leq -2$  were used to define the three groups of clones. TCR $\beta$  clones that were exclusively detected in the samples before or exclusively after treatment were

considered as disappearing clones (DIS) and newly circulating clones (NEW), respectively. All clones  $< 0.01\%$  for the total CD3 and  $< 0.1\%$  for the sorted T cell subpopulations were considered background.

#### *Statistical analysis*

Data were all normally distributed according to D'Agostino-Pearson test for normality and were therefore reported as mean and standard deviation (SD). Differences between T cell subsets were evaluated using T-test, and differences between pre- and post-DC-therapy samples were evaluated with a paired T-test. p-values  $< 0.05$  were considered statistically significant. Prism 7 software (Graph Pad, San Diego, CA, USA) was used to perform the statistical tests. **P-values were corrected using the Benjamini & Hochberg False Discovery Rate when multiple hypothesis were tested on the same T cell subsets.**

## References

- 1 Aerts JGJV, De Goeje PL, Cornelissen R, Kaijen-Lambers MEH, Bezemer K, Van Der Leest CH *et al.* Autologous dendritic cells pulsed with allogeneic tumor cell lysate in mesothelioma: From mouse to human. *Clin Cancer Res* 2018; **24**: 766–776.
- 2 Klarenbeek PL, Tak PP, van Schaik BDC, Zwinderman AH, Jakobs ME, Zhang Z *et al.* Human T-cell memory consists mainly of unexpanded clones. *Immunol Lett* 2010; **133**: 42–48.
- 3 Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 2014; **30**: 614–620.
- 4 Giudicelli V, Chaume D, Lefranc MP. IMGT/GENE-DB: A comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res* 2005; **33**. doi:10.1093/nar/gki010.
- 5 Chao A, Chazdon RL, Shen TJ. A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecol Lett* 2005; **8**: 148–159.
- 6 Chao A, Chazdon RL, Colwell RK, Shen TJ. Abundance-based similarity indices and their estimation when there are unseen species in samples. *Biometrics* 2006; **62**: 361–371.
- 7 Musters A, Klarenbeek PL, Doorenspleet ME, Balzaretto G, Esveldt REE, van Schaik BDC *et al.* In Rheumatoid Arthritis, Synovitis at Different Inflammatory Sites Is Dominated by Shared but Patient-Specific T Cell Clones. *J Immunol* 2018; **201**: 417–422.