Single cell profiling reveals modulatory effects of immunotherapeutic RNA approaches in cancer and autoimmune disease models

DISSERTATION

Zur Erlangung des Grades

Doktor der Naturwissenschaften

(Dr. rer. nat.)

Am Fachbereich Biologie

Der Johannes-Gutenberg-Universität in Mainz

Michael Streuber

geb. am 30.04.1991 in Bogotá, Kolumbien

Mainz, 2024

Dekan: Prof. Dr. Eckhard Thines

Tag der mündlichen Prüfung: 19.11.2024

EIDESSTAATLICHE ERKLÄRUNG

Hiermit versichere ich, dass die vorliegende Dissertation von mir selbstständig und ohne entgeltliche Hilfe Dritter verfasst wurde.

Alle benutzen Hilfsmittel sind vollständig angegeben worden und keine anderen, als die aufgeführten Quellen wurden verwendet. Alle der Literatur entnommenen Stellen wurden als solche gekennzeichnet. Personen welche mir Unterstützungsleistungen gaben wurde in der Dissertation benannt.

Diese Dissertation wurde bisher nicht für eine Prüfung oder Promotion oder für einen ähnlichen Zweck zur Beurteilung eingereicht. Ich erkläre, dass ich nicht schon anderweitig einmal die Promotionsabsicht angemeldet oder ein Promotionseröffnungsverfahren beantragt habe.

Marburg, den 29.03.2024 ……..……..

Michael Streuber

“Never forget that you are learning only for yourself. Use it to make the world a little better. And if you get stuck I know that you will find a way, use your mind and your instinct.”

Helmut Rossel (1939-2017)
Grandfather, firefighter and locksmith

SUMMARY

This work focused on establishing a single-cell RNA sequencing (scRNA-seq) pipeline for comprehensive immune cell analysis to reveal cell types and states as well as to dissect the gene expression (GE) and regulatory relationships between genes and cells in health and disease. Our scRNA-seq pipeline is based on the Chromium Controller, a device capturing and amplifying a minute amount of RNA and analysis tools for clustering (Seurat) and enriched pathways or genes (Cluster Profiler and gene set variation analysis). In this thesis, researchers used the pipeline to investigate the effects of immunotherapeutic RNA approaches in cancer and autoimmune disease models.

Chapter one demonstrates the pipeline's ability to identify cell types and the robustness of the protocol through the consistent gene expression profiles of PBMC samples from the same donors. Comparisons between 3'gene expression protocol (3'GE) and 5'GE protocols showed no significant differences in cell composition and cell-type-specific genes. Therefore, single cell analysis pipeline based on 3’GE was applied to study transcriptional dynamics of immune cells in disease models for cancer and multiple sclerosis (MS) upon RNA vaccination in three different experimental settings:

Chapter two describes the studies related to the adjuvant effect of RNA-Lipoplex (LPX) vaccination encoding the reporter gene enhanced green fluorescent protein (EGFP) on murine splenocytes, indicating an increase in the frequency of B cells and monocyte/granulocyte fraction. This led to the triggering of chemokines and inflammatory cytokines including upregulation of genes involved in interferon alpha (IFNα) and beta (INFβ) response, suggesting a heightened immune response and confirming the adjuvant effect of RNA by modulating innate immune system.

Chapter three outlines the impact of autoantigen-encoding non-immunogenic RNA-LPX vaccination in experimental autoimmune encephalomyelitis (EAE), a mouse model for MS. A successful tolerance induction in EAE sick mice was observed in form of high effector regulatory T (Treg) cell frequencies using a myelin oligodendrocyte glycoprotein peptide incorporated with 1-methylpseudouridine (MOG35_55_m1Ψ). The main investigation for this thesis chapter was further to dissect the effect of MOG35-55_U in comparison to MOG35-55_m1Ψ and irrelevant m1Ψ and to study functional differences with respect to inflammatory and immunosuppressive properties of the present CD4+ T cell subpopulations. MOG35-55_U in comparison to MOG35-55_ m1Ψ led to downregulation of cell cycle genes and resulted into

less cytotoxic signature but therefore higher level of stress was observed. MOG35_55_U as a more immunogenic mRNA construct did not lead to an enrichment of pathways demonstrating a (stronger) tolerance induction in EAE sick mice.

Lastly, chapter four illustrates the therapeutic potential of liposomally-formulated RNA encoding the modified IL2var alone or in combination with a monoclonal antibody against programmed cell death 1 (aPD1) in a MC38 tumor mouse model to assess the immune response in CD45+ cells. The treatment demonstrated increased tumor infiltrating lymphocytes (TILs) frequency, a strong effector phenotype, reduction of exhaustion markers and an increased expression of transcription factors, suggesting a promising therapeutic cancer vaccine potential.

In conclusion, the scRNA-seq pipeline developed in this thesis allowed in-depth analysis of individual cell subsets at single cell level and demonstrated the potential of various RNA-based treatments in cancer and autoimmune disease models. Further research is suggested to fully exploit the pipeline's capabilities and enhance therapeutic strategies.

ZUSAMMENFASSUNG

Diese Arbeit konzentrierte sich auf die Etablierung einer Einzelzell-RNA-Sequenzierung (scRNA-seq) Pipeline für eine umfassende Analyse von Immunzellen, um Zelltypen und -Zustände aufzudecken sowie die Genexpression (GE) und regulatorische Beziehungen zwischen Genen und Zellen in Gesundheit und Krankheit zu untersuchen. Unsere scRNA-seq-pipeline basiert auf dem Chromium Controller, einem Gerät zur Erfassung und Amplifikation einer geringen Menge RNA, sowie auf Analysetools für das Clustering (Seurat) und angereicherte Signalwege oder Gene (Cluster Profiler und Gen-Set-Variationsanalyse). In dieser Arbeit nutzten die Forscher die Pipeline, um die Auswirkungen immuntherapeutischer RNA-Ansätze in Krebs- und Autoimmunkrankheitsmodellen zu untersuchen.

Kapitel eins zeigt die Fähigkeit der Pipeline, Zelltypen zu identifizieren und die Robustheit des Protokolls durch die konsistenten Genexpressionsprofile von PBMC-Proben derselben Spender. Vergleiche zwischen 3'-Genexpressionsprotokoll (3'GE) und 5'-GE-Protokollen zeigten keine signifikanten Unterschiede in der Zellzusammensetzung und den zelltypspezifischen Genen. Daher wurde die Einzelzellanalyse-Pipeline auf der Grundlage von 3'-GE angewendet, um die Transkriptionsdynamik von Immunzellen in Krankheitsmodellen für Krebs und Multiple Sklerose (MS) nach RNA-Impfung in drei verschiedenen experimentellen Einstellungen zu untersuchen.

Kapitel zwei beschreibt die Studien zur adjuvanten Wirkung der RNA-Lipoplex (LPX)-Impfung, die für das Reportergen Enhanced Green Fluorescent Protein (EGFP) kodiert, auf Mäusesplenozyten, die eine Erhöhung der Häufigkeit von B-Zellen und der Monozyten-/Granulozytenfraktion erkennen lassen. Dabei wurde eine Zunahme der Häufigkeit von B-Zellen und der Monozyten/Granulozyten-Fraktion festgestellt. Dies führte zur Auslösung von Chemokinen und entzündlichen Zytokinen, einschließlich der Hochregulation von Genen, die an der Interferon-alpha (IFNα)- und -beta (INFβ)-Reaktion beteiligt sind. Dies legt eine verstärkte Immunantwort nahe und bestätigt den adjuvanten Effekt von RNA durch Modulation des angeborenen Immunsystems.

Kapitel drei behandelt die Auswirkungen einer Autoantigen-kodierenden nicht-immunogenen RNA-LPX-Impfung bei experimenteller autoimmuner Enzephalomyelitis (EAE), einem Mausmodell für MS. Es wurde eine erfolgreiche Toleranzinduktion bei erkrankten EAE-Mäusen in Form hoher Effektor-Regulatorischer T-Zellen (Treg-Zellen) beobachtet, bei dem ein Myelin-Oligodendrozyten-Glykoprotein-Peptid mit 1-Methylpseudouridin

(MOG35_55_m1Ψ) inkorporiert wurde. Die Hauptuntersuchung dieses Thesis-Kapitels zielte darauf ab, den Effekt von MOG35-55_U im Vergleich zu MOG35-55_m1Ψ und irrelevantem m1Ψ zu analysieren und funktionelle Unterschiede hinsichtlich entzündlicher und immunsuppressiver Eigenschaften der vorliegenden CD4+ T-Zell-Subpopulationen zu untersuchen. Im Vergleich zu MOG35-55_m1Ψ führte MOG35-55_U zu einer Herunterregulierung von Zellzyklus-Genen und resultierte in einer geringeren zytotoxischen Signatur, jedoch wurde ein höheres Stressniveau beobachtet. MOG35_55_U als ein immunogeneres mRNA-Konstrukt führte nicht zu einer Anreicherung von Signalwegen, die eine (stärkere) Toleranzinduktion bei erkrankten EAE-Mäusen zeigen.

Schließlich wird in Kapitel 4 das therapeutische Potenzial von liposomal formulierter RNA, die für das modifizierte IL2var kodiert, allein oder in Kombination mit einem monoklonalen Antikörper gegen den programmierten Zelltod 1 (aPD1) in einem MC38-Tumormausmodell veranschaulicht, um die Immunantwort in CD45+-Zellen zu bewerten. Die Behandlung zeigte eine erhöhte Häufigkeit tumorinfiltrierender Lymphozyten (TILs), einen starken Effektorphänotyp, eine Reduktion von Erschöpfungsmarkern und eine gesteigerte Expression von Transkriptionsfaktoren, was auf ein vielversprechendes therapeutisches Potenzial eines Krebsimpfstoffs hinweist.

Zusammenfassend ermöglichte die in dieser Arbeit entwickelte scRNA-seq-pipeline eine eingehende Analyse einzelner Zellsubsets auf Einzelzellebene und zeigte das Potenzial verschiedener RNA-basierter Behandlungen in Krebs- und Autoimmunkrankheitsmodellen auf. Weitere Forschung ist nötig, um die Fähigkeiten der Pipeline vollständig auszuschöpfen und therapeutische Strategien zu verbessern.

ABBREVIATIONS

ACT	Adoptive T cell therapy
ADT	Antibody derived tags
AdjRNA	Adjuvant Ribonucleoside acid
aggr	aggregated
AgRNA	Antigen-coding Ribonucleoside acid
ATAC-seq	Assay for transposase-accessible chromatin with sequencing
BBB	Blood-Brain-Barrier
BC	Buffy coat
BCR	B cell receptor
BNT	BioNTech
bp	Base pairs
CAR	Chimeric antigen receptors
cDCs	Conventional dendritic cells
cDNA	Complementary Deoxyribonucleic acid
CITE-seq	Cellular indexing of transcriptomes and epitopes sequencing
Cl. Mono	Classical monocytes
CNS	Central nervous system
CPI	Checkpoint inhibitor
CRISPR	Clustered regularly interspaced short palindromic repeats
CTLs	Cytotoxic T lymphocytes
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
DCs	Dendritic cells

ABBREVIATIONS

DE	Differentially expressed
DGEA	Differential gene expression analysis
DNA	Deoxyribonucleic acid
DPBS	Dulbecco's phosphate buffered saline
dsRNA	Double-stranded Ribonucleoside acid
EAE	Experimental autoimmune encephalomyelitis
EBV	Epstein Barr virus
EGFP	Enhanced green fluorescent protein
FACS	Fluorescence activated cell sorting
FDA	Food and Drug Administration
FF	Fresh frozen
FSC	Forward scatter
FU	Functional Units
GC	Guanin-Cytosine
GE	Gene expression
GEMs	Gel bead-in-emulsions
GO	Gene Ontology
hAlb	Human serum albumin
HBV/HCV	Hepatitis B/C virus
hPBMC	Human peripheral blood mononuclear cell
HPV	Human papilloma virus
HTO	Hashtag-oligos
ICB	Immune checkpoint blockade

ICI	Immune checkpoint inhibitor
IFNα	Interferon alpha
INFβ	Interferon beta
i.v.	Intravenous
IDs	Identities
KEGG	Kyoto Encyclopedia of Genes and Genomes
KIR	Killer Ig-like Receptors
KNN	k-Nearest Neighbor
LD	Live/Dead
LNP	Lipid Nanoparticles
LPX	Lipoplex
m1Y	1-methyl-pseudouridine
mAbs	Monoclonal antibodies
MACS	Magnetic Activated Cell Sorting
MDSCs	Myeloid-Derived Suppressor Cells
MHC	Major Histocompatibility Complex
MOG	Myelin Oligodendrocyte Glycoprotein
mRNA	Messenger ribonucleoside acid
MS	Multiple sclerosis
NGS	Next-Generation-Sequencing
NK cells	Natural-killer cells
ORA	Over-representation-analysis
OV	Oncolytic virus

OVA	Ovalbumin
PBMC	Peripheral blood mononuclear cell
PCA	Principal component analysis
PCs	Principal components
PD-1	Programmed Cell Death Protein 1
pDCs	Plasmacytoid dendritic cells
PHS	Pooled human serum
PMT	Photomultiplier tubes
Polydt	Poly deoxythymine
PPMS	Primary progressive multiple sclerosis
qPCR	quantitative-polymerase-chain-reaction
Rep	Replicate
RMS	Relapsing forms of multiple sclerosis
RNA	Ribonucleoside acid
RNA-LPX	RNA-lipoplexes
RPMI	Roswell park memorial institute
RT	Room temperature
RT reaction	Reverse transcriptase reaction
SNN	Sheared-nearest neighbor
SSC	Side scatter
s.c.	Subcutaneous
seq	Sequencing
scRNA-seq	Single cell RNA sequencing

ABBREVIATIONS

TAA	Tumor-associated self antigens
TCR	T cell receptor
Tfh cells	T follicular helper cells
Th cells	T helper cells
TILs	Tumor infiltrating lymphocytes
TME	Tumor microenvironment
Treg cells	Regulatory T cells
TSA	Tumor-specific antigens
tSNE	t-distributed stochastic neighbor embedding
TSO	Template switching oligo
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique molecular identifier
UVB	Ultraviolet B light
VLA-4	Very late activation antigen 4
γδ T cells	Gamma delta T cells

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

1 Chapter One: Establishing high-throughput single cell transcriptome sequencing in biomedical and preclinical research to explore immunotherapeutic notions

1.1 Introduction

1.1.1 Single cell RNA sequencing

For the validation of a cancer vaccine treatment and its effect on the different immune cell subtypes, cell subset-specific activation markers are routinely investigated by fluorescence-activated-cell-sorting (FACS). Even though it is a widely used method for protein analysis, there is an incomplete understanding of the immune responses based on protein analysis which requires prior knowledge. Some further limitations of FACS include the provision or creation for specific antibodies and the possible interference with downstream analyses [1]. The quantitative-polymerase-chain-reaction (qPCR) method is another common method for the investigation of the transcriptional state of immune cells, but it can cause difficulties with regions of extreme GC% (Guanin-Cytosine), resulting in inefficiently amplified base-pairs [2].

The gap mentioned above was to be closed by the development of cheaper and faster Next-Generation-Sequencing (NGS) technologies which led to the creation and completion of the Human genome project [3]. NGS methods do not require electrophoretic separation of sequencing products and bacterial cloning of Deoxyribonucleic acid (DNA) fragments [2,4]. Traditional Ribonucleoside acid sequencing (RNA-seq) techniques measure the expression level averaged across hundreds to millions of cells as input, also called bulk RNA-seq techniques. However, they have some drawbacks due to averaged gene expression profile from bulk RNA-seq such as signals driving tumorigenesis or therapeutic resistance from a rare cell types can be obscured [4] and they require a quite high minimal amount of RNA [5].

To overcome these issues new method regarding creation of the sequencing library or sample treatment evolved since 2009, in which Fuchou Tang and his group succeeded in establishing the first method for genome wide single cell RNA sequencing (scRNA-seq) [6]. This method enabled the observation of the real state of individual cells, their gene expression dynamics, differentiation trajectories and regulatory processes as well transcriptional kinetics. It also allowed to identify novel cell types. The value of single cell sequencing platforms gained more

attention and led to the development of different non-commercial and commercial techniques by institutes and companies. Using a platform like scRNA-seq has become especially relevant since cells are temporally organized and interact with each other to orchestrate self-assembly and response to stimuli as a whole [1]. Studying single cell gene expression and comparing gene expression profiles between individual cells allows the discovery of previously undetected populations and unveiling new regulatory paths. It is especially useful to identify subgroups of the investigated single cell populations, such as individual T lymphocytes expressing highly diverse T-cell receptors (TCRs) [7]. It further allows the identification of co-regulated gene modules and inference of gene-regulatory networks that underlie functional heterogeneity and cell-type specification [8].

The dissociation of cells from a tissue and the isolation of single cells with specific devices are the first steps of a typical scRNA-seq experimental workflow. Various cell-isolation techniques like serial dilution, robotic-controlled micropipettes, laser-capture microdissection, rotating magnets with antibodies, flow sorting and microfluidic platforms are able to isolate cells with different outcome and advantages/disadvantages. Concerning cell isolation, in this thesis a microfluidic platform was chosen since the probability of isolating multiple cells is lower compared to serial dilution. In addition, it has a higher throughput compared to robotic micromanipulation, there is no risk that cells are able to adhere to nanofilters and there is no risk of cell splicing due to UV damage to DNA or RNA in case of laser-based applications as well as isolation is not biased due to the use of markers (using magnets with nanoparticles conjugated to antibodies e.g. CellSearch) [9]. Although the required consumables are more expensive and cell size must be below a certain size threshold, microfluidic platforms show high throughput of thousands to millions of cells, the reagent costs can be reduced since some reactions are done on-chip and reactions can be multiplexed and the [10].

In order to make a sound decision to choose an appropriate technique for the (i) isolation, (ii) cell barcode labeling and (iii) generation of sequencing libraries, it is necessary to compare the different platforms with special emphasis on their characteristics. These fundamental characteristics are: maximum cell throughput number, capture efficiency of single cells, double rate, needed reaction volume and the maximum sequencing depth per cell (Table 1).

Concerning the **throughput number,** most plate-based technologies can only handle lower maximum cell input e.g. 1k cells (except SPLIT-seq with more than 100k cells), which is very inefficient for big cohorts of samples. In contrary, nanowell arrays (CellenONE-ICELL8, SEQ-well, BD Rhapsody) and microfluidic platforms (Drop-seq, InDrop and 10x Genomics) are able to process between 25k – 40k single cells.

Technologies differ also in terms of **sequencing depth** (generated reads per single cell). It requires from up to 10k reads per cell (SPLIT-seq) to 1 million reads (Smart-seq2 and MATQ-seq) while all other platforms have an intermediate range of 100 – 300k reads. These differences are based on the varying numbers of processable single cells which will be barcoded. Smart-seq2 and MATQ-seq are the best performers in this category.

Methods vary greatly in their **reaction volume.** This is an important characteristic as it determines the minimum quantity/volume of a sample needed to process it with the appropriate protocol. All plate-based platforms have a rather high reaction volume of probe (microliter), which can be challenging if rare cell types shall be assessed or isolated or if a sample consists of a relatively small reaction volume. Among the investigated technologies, nanowell arrays and microfluidics show the best results as they only require a few nanoliter reaction volume [11–21].

Table 1: Overview of various scRNA-seq platforms (sources: [11–21])

Protocol	Transcript Data	Platform	Throughput number	Read depth (per cell)	Reaction volume	Reference
CellenONE-ICELL8	Full length	Nanowell array	100 - 5k	300k	Nanoliter	[11]
Smart-seq2	Full length	Plate-based	100 - 1k	1 mil	Microliter	[12]
MATQ-seq	3´sequencing	Plate-based	100 - 1k	1 mil	Microliter	[13]
MARS-seq	3´sequencing	Plate-based	100 - 1k	10k - 100k	Microliter	[14]
CEL-seq	3´sequencing	Plate-based	100 - 1k	10k - 100k	Nanoliter	[15]
SPLIT-seq	3´sequencing	Plate-based	1k - 100k	10k	Microliter	[16]
SEQ-well	3´sequencing	Nanowell array	1k - 10k	10k - 100k	Nanoliter	[17]
BD Rhapsody	3´sequencing	Nanowell array	2k - 40k	50k	Nanoliter	[18]
Drop-seq	3´sequencing	Microfluidics	1k - 10k	10k - 100k	Nanoliter	[19]
InDrop	3´sequencing	Microfluidics	1k - 10k	10k - 100k	Nanoliter	[20]
Chromium	3´ and 5´sequencing	Microfluidics	500 - 25k	10k - 100k	Nanoliter	[21]

Regarding **capture efficiency**, microfluidic devices like InDrop or Chromium (10x Genomics) have only moderate capture efficiency between 7-50% [10] whereas plate based platforms like Smart-seq2 and MARS-seq and the nanowell array CellenONE-ICELL8 have much higher capture efficiency of more than 80%, which is useful when you have a sample with limited cell number and most of it shall be analyzed [10,11].

The **double rate** represents the percentage of double or multiple cells (multiplets) in one well/oil droplet and increases with higher number of loaded single cells. This is important as some protocols can process more than 10k cells and a higher amount of multiplets need to be taken into account later for the analysis to not make false statements. While Rhapsody has the lowest double rate with 0.6% and DropSeq has a double rate of 11.3%, the platform C1 from Fluidigm can even reach 30% double rate. All other scRNA-seq platforms show quite intermediate and low double rates which makes the data analysis easier dealing with less multiplets. Double rates: 1.3-4% (ICell8), 2.27% (MARS-seq), 4% (InDrop) and 1.6-3% (Chromium). The double rate for Smart-seq2 was not reported [10].

In general, there is also a difference for final library constructs or transcript data as mentioned above in Table 1: the major difference between two protocols of 3´ GE (Gene expression) or 5'GE is the fact that final library constructs consist of different primer sequences (polydT, poly deoxythymine tail or template switching oligo, TSO), generated during the complementary DNA (cDNA) construction. This primer sequence variations allows different library constructions and downstream investigations (Figure 1 and Figure 2).

Droplet-based microfluidics can encapsulate thousands of single cells in individual partitions, each containing all the necessary reagents for cell lysis, reverse transcription and molecular tagging for final library preparation to sequence these molecules [21]. A device that fulfills this mode of operation is the Chromium Controller provided by 10X Genomics. More 10X Genomics protocols were established and can be used to not only survey the RNA expression in single cells, but also to profile the cell surface protein expression (using barcoded labelled antibodies), CRISPR (Clustered regularly interspaced short palindromic repeats) perturbations, full length TCRs and BCRs (B cell receptor) of single cells, antigen specificity, chromatin accessibility using ATAC-seq (assay for transposase-accessible chromatin with sequencing) and targeted gene expression analysis [22,23].

Taking into consideration all the above-mentioned characteristics, microfluidics technologies show the best results for cell throughput number, double rate, reaction volume and sequencing depth concerning the study of high number of single cells. Although the plate-based protocols show higher capture efficiency, all other characteristics were more in line with our expectations and requirements to study different types and numbers of cells. We decided for the microfluidic platform Chromium Controller (10x genomics) which will be explained in more detail.

For the application of the Chromium Controller and gene expression analysis, cells first need to be isolated, filtered and/or sorted to have viable single cells. Lab workflow continues with the sample loading in a microfluidic chip (max. 8 samples in parallel). The chip encapsulates single gel beads, which with barcoded oligonucleotides consisting of: sequencing adapters, 16 base-pairs (bp) "10X barcode" drawn from ~750,000 designed sequences to mark cells, a 10 bp randomer to mark molecules (unique molecular identifier, UMI) and 30 bp oligo-dT anchor sequence for binding poly-adenylated RNA transcripts (Single Cell 3´Reagents Kits v2 at that time) with single (each 8bp long) which was used for this thesis. In each microfluidic channel thousands of cells are encapsulated with reagents (e.g. partioning oil and lysis reagents) and gel beads to form the Gel beads in emulsion (GEMs). Cells are lysed and cDNA is generated inside the oil droplets via reverse transcriptase (RT) reaction. GEMs are broken, cDNA is pooled and amplified in one tube. Another sequencing adaptor and 10X Index sequence is added, compatible with NGS protocols (Figure 1)[21].

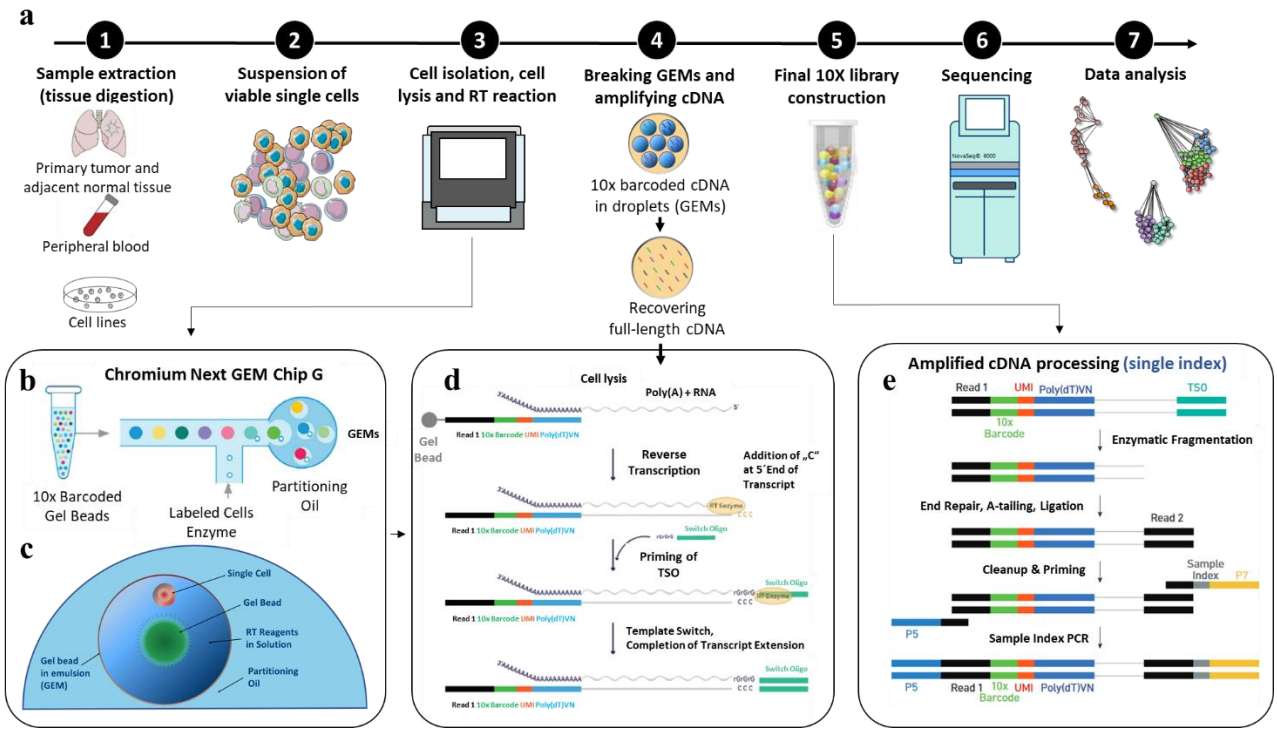


Figure 1: 10x Genomics concept of scRNA-seq workflow for 3´ capture assay. Short description of the workflow (modified from [21]). **a)** The workflow includes 7 main steps: (1) Cell/sample recovery, (2) generation of viable single cell suspension, (3) loading cells on microfluidic chip including cell lysis and RT reaction reagents, (4) pooling and cDNA amplification, (5) library construction, (6) sequencing and (7) data analysis. **b)** Chromium Next GEM Chip G's consists of two main channels, providing gel beads, single cells and enzyme reagents to form the GEMs. **c)** GEMs contain one barcoded-gel bead, optimally one single cell and RT reagents. **d)** Polyadenylated RNA transcript is primed by oligo-dT sequence, RT reaction generates cDNA with following sequences (single index protocol): TruSeq Read 1 (sequencing adaptor), 10X barcode (for cell identification), UMI sequence (for transcript quantification), polydT tail and a template-switch-oligo (TSO) at 3´end (for 5'GEne expression, TSO is at 5´end). First-strand cDNA is purified and full-length cDNA is amplified via PCR. **e)** Enzymatic fragmentation and size selection are done for optimal fragment size. Remaining TruSeq Read 2, sample index primer (i7) is incorporated via End Repair, A-tailing, Adaptor Ligation, and PCR. P5 and P7 primers used for sequencing bridge amplification are added as well. Dual index libraries contain also another index primer (i5) (adapted from [21]).

The 3´ Gene expression (GE) and 5'GE protocols differ in the captured ends of the polyadenylated transcript in the final library. In the 3'GE protocol, the polydT sequence is located on the gel bead oligo. For the 5'GE protocol the polydT is supplied as an RT primer. Final library structures are illustrated below (Figure 2). Single cell 3'GE libraries are suitable for sequencing on the Illumina sequencing platforms (MiSeq, HiSeq4000, NovaSeq6000) due to standard Illumina paired-end constructs beginning with P5 and end with P7 sequence. Read 1 encodes the 10x Barcode and the UMI sequence, whereas Read 2 encodes the cDNA fragment [24].

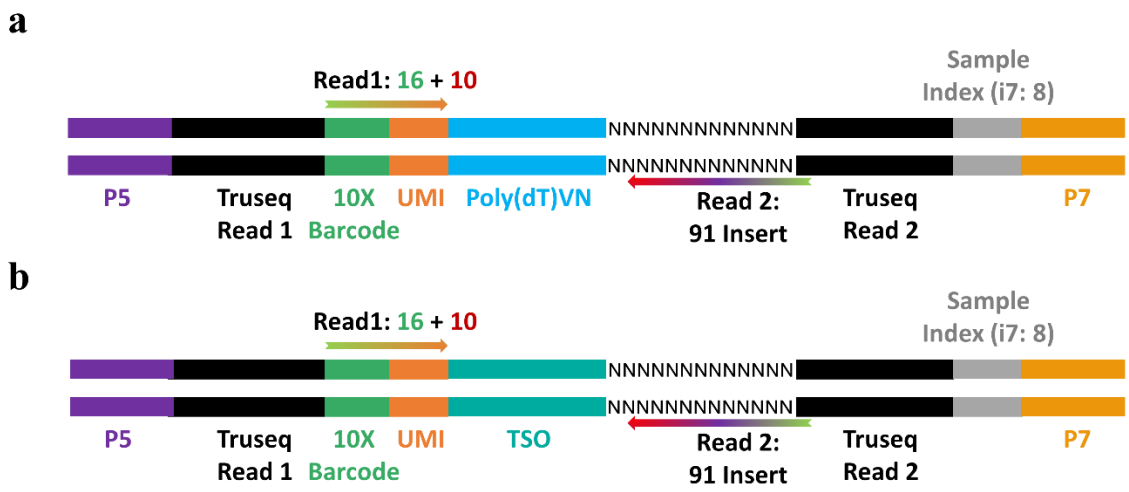


Figure 2: Final 10x Genomics library structures. a) Single Cell 3' v2 Gene Expression Library. **b**) Single Cell 5' v2 Gene Expression Library. The polydT is supplied by an RT primer and final library structure contains a TSO sequence for 5'GE protocol.

Following the sequencing, raw data analysis (demultiplexing) was done using Cell Ranger and output was further analyzed using Seurat analysis tool (Figure 6). Another tool which was used during this doctoral thesis was "Gene-set-enrichment-analysis" (GSEA, which is used to check if a set of genes shows statistically significant differences between two biological states) and "Gene-set-variation-analysis (GSVA) [26].

1.1.2 Aim of the establishment experiments

The main goal of the establishment experiments was the successful detection of immune and tumor cells with results meeting the published standards. For the experiments, tumor cells of the Colo205 cell line [27] were provided by A. Kölsch and K. Claude. Cryogenic storage vials containing human peripheral blood mononuclear cells (PBMCs) were supplied by Ö. Akili-Öztürk and Buffy coat preparation was conducted by R. Rae. All other procedures, such as the cultivation of cell lines, single cell preparation, and the application of our single cell analysis pipeline, was shared between me, L. Kolb and E. Diken. An exception was the loading of single cell libraries onto the HiSeq4000, performed by J. Grimm and L. Giese, for the sequencing of the samples. In detail PBMCs and a tumor cell line Colo205 were mixed at predefined ratios mimicking tumor-infiltrating immune cells from the tumor environment. The cell type ratios in the mixed PBMC and Colo205 cells were confirmed by FACS through S. Attig. Additionally, standard sequencing result parameters (median genes per cells, estimated number of cells etc.)

for PBMC samples were tested to be comparable with results provided by 10x Genomics with established sample preparation procedures.

The present work deals also with the comparison of the performance of the two different existing scRNA-seq capture assays, namely 10X Genomics 3'GE protocol and 5'GE protocol. The comparison will provide a better understanding if there are differences in recovered cell numbers and sequencing performance (e.g. higher level of median genes per cell for one protocol). Based on the results, it should be decided if one method would be more suitable than the other for the investigation of transcriptional differences between various processed samples or if there will be discrepancies if different datasets based on 3'GE and 5'GE respectively will be analyzed. 5'GE protocol was executed by A. Obermann and samples were analyzed by S. Newrzela in BioNTech.

In addition to the main goal of the experiment, several minor experimental parameters were taken into account and compared to get in general a better understanding of the factors that have to be carefully considered in each experimental design. On one hand, human PBMC samples were processed simultaneously with murine cells in the same microfluidic chip to rule out (species) cross contaminations to show that the robustness of the protocol for laboratory and bioinformatics data analysis is given. Furthermore, a comparison between the preparation methods of PBMC samples (fresh prepared of thawing fresh frozen samples) was done to assess the influence of the freezing procedure on cell viability and quality. Between the fresh frozen human PBMC samples a total of three donors were analyzed to see if data from different donors are comparable. In addition, different number of cells (500, 5000 or 10,000) were loaded to evaluate the minimal number of cells required for a robust analysis. Besides testing different cell preparations protocols (thawing fresh frozen or using fresh prepared cells) capture assays and different sample conditions, another focus is also on establishing the data analysis pipeline. For this purpose, different bioinformatics tools were tested and used for cell clustering and gene expression analysis.

1.2 Results

1.2.1 Mimicking single cell profiling of tumor and its microenvironment

The following chapter summarizes the dissection of human immune cells and a tumor cell line Colo205. Here, sequence results of the PBMCs were compared with publicly available datasets provided by the developer 10X Genomics to ensure that the overall quality of the experiment

meets standard requirements. For this experiment three different cell types were prepared as followed (Figure 3): i) tumor cell line Colo205 cells were used for single cell sequencing establishment experiments after two passages; ii) cryoconserved hPBMCs were thawed and washed several times and iii) murine splenocytes (mice treated with EGPP mRNA as "immunized" or with PBS buffer as "control", these samples are part of Chapter Two: Intrinsic adjuvanticity of RNA vaccines modulates innate immune system) were sorted. Here, murine cells were loaded with human samples alternatingly to prove the robustness of the protocol for laboratory and bioinformatics data analysis and to evaluate the risk of cross-contamination. Two replicates contained mixed PBMC and Colo205 cells in different ratios (80% PBMCs and 20% Colo205 cells or 50% PBMCs and 50% Colo205 cells). All cells were filtered for size selection and counted before loading on the Single cell chip.

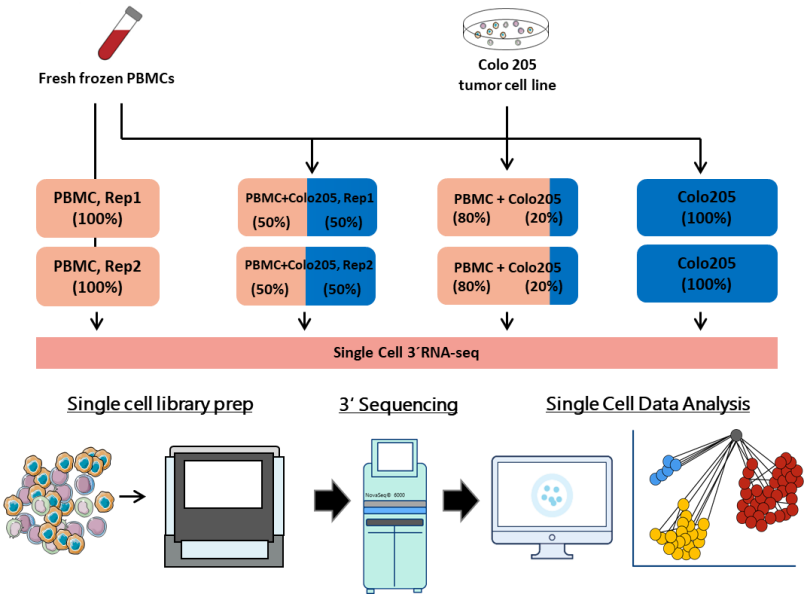


Figure 3: Workflow overview of establishment scRNA-seq experiment – Processing pure human PBMC and mixed PBMC samples with tumor cell line Colo205. Human samples (pure hPBMCs or mixture at indicated ratios with tumor cell line Colo205 cells) were loaded on the single cell chip. Cell input was 10,000 cells for alle samples. Samples were processed by the single cell sequencing pipeline.

After barcode labeling, reverse transcriptase and cDNA amplification successful cDNA synthesis was verified via assessment of the cDNA fragment profiles in the BioAnalyzer and concentration via Qubit (Figure 4). Same steps were done for final library synthesis after fragmentation, end repair, A-tailing, adaptor ligation, sample index PCR and several bead cleanups.

Almost all samples have a **peak fragment profile** at approximately 1200 bp wherein the signals of the mixed samples PBMC and Colo205 cells show particularly large deflections and the replicate 2 of the PBMC 100% have a peak between 200 bp and ~220 bp. In Figure 4a of the bioanalyzer results, significantly higher concentrated cDNA fragments were detected for both mixed samples of hPBMCs and Colo205 cells. **cDNA concentrations** was determined for the mix samples of PBMCs and Colo205 cells which are also higher concentrated due to size difference (immune cells much smaller than cancer or epithelial cells) than for all other samples [28]. Results can be seen in Table 2. The concentration of these samples is two-fold higher than the pure 100 % human PBMC samples.

This could be most likely to the fact that the proportion of cDNA for Colo205 tumor cell line is much higher than that of immune cells with lower amount of RNA [29]. The peak at ~220bp in the same plot may indicate carryover adaptor/primer dimers which were not purified during bead purification [30]. However, these primer dimers could be purified during library preparation. All samples could be further processed for library preparation. Figure 4b shows the fragment profile of all final libraries after adapter ligation, index PCR etc. again with the human samples mixed with the Colo205 cells as the highest signal peaks. All final libraries have a uniform average fragment size between 452-483 bp. Based on the electropherograms, no residual primers were detected that could negatively affect the sequencing.

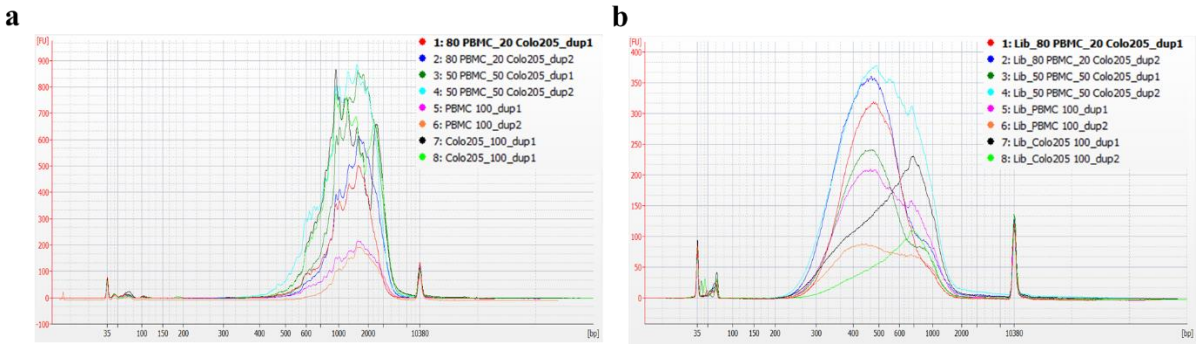


Figure 4: Electropherogram of cDNA and final libraries of human PBMC samples. The fragment length is shown on the x-axis [bp] whereas the amount of the fluorescence signals is illustrated on the y-axis [FU]. Both standards can be seen at 35 bp and 10380 bp. The signals of the human and mouse amplified (**a**) cDNA samples and (**b**) final libraries are illustrated in different colors (80%PBMC_20%Colo205_dup1: red; 80%PBMC_20%Colo205_dup2: dark blue; 50%PBMC_50%Colo205_dup1: dark green; 50%PBMC_50%Colo205_dup2: light blue; 100%PBMC_dup1: pink; 100%PBMC_dup2: orange; Colo205_100%_dup1: black; Colo205_100%_dup2: light green).

Table 2: DNA concentration of human samples of cDNA and final libraries. Concentrations in ng/µl were measured by Bioanalyzer and Qubit.

		Qubit	Bioanalyzer	
Sample type	**Sample name**	**Concentration (ng/µl)**	**Fragment size (bp)**	**Concentration (ng/µl)**
cDNA	PBMC_80_Colo205cells_20_dup1	26.65	1454	25.67
	PBMC_80_Colo205cells_20_dup2	36.35	1669	34.46
	PBMC_50_Colo205cells_50_dup1	53.00	1685	54.14
	PBMC_50_Colo205cells_50_dup2	89.00	1436	67.26
	PBMC_100_dup1	12.80	1600	12.83
	PBMC_100_dup2	12.00	1744	11.30
	Colo205_100_dup1	97.00	1538	47.23
	Colo205_100_dup2	82.00	1574	52.03
Final library	PBMC_80_Colo205cells_20_dup1	38.20	511	37.53
	PBMC_80_Colo205cells_20_dup2	45.90	513	45.63
	PBMC_50_Colo205cells_50_dup1	28.70	525	29.63
	PBMC_50_Colo205cells_50_dup2	57.80	570	59.54
	PBMC_100_dup1	26.70	562	30.52
	PBMC_100_dup2	15.00	563	14.46
	Colo205_100_dup1	26.60	635	24.08
	Colo205_100_dup2	8.25	665	9.41

Samples were sequenced using the Illumina Hiseq4000 device and a HiSeq3000/4000 PE flowcell. Demultiplexing was done by Cell Ranger software to determine the raw sequencing results. After the sequencing run, the generated clusters were checked with the Illumina software "Sequencing Analysis Viewer" and some selected parameters are summarized in the table below (Table 3). Except the parameters %>=Q30 and the Error rate, the results for all the other parameters were the same over all three reads. More than 68% of the clusters and more than 331 million reads passed filtering. For read 3 more than 77 % of all bases had a quality score of Q30 (1 error in 1000 bases) or more. Optionally, the error rate could be determined via the use of a PhiX control.

Table 3: Raw sequencing results of human PBMC samples. QC results and statistics of HiSeq4000 run are given as means and were summarized by Sequencing Analysis Viewer. In detail: Cluster PF, Reads PF, Quality Score of 30%.

Read	Clusters PF (%)	Reads PF (M)	% >= Q30
1	68.81	331.15	96.29
2	68.81	331.15	82.36
3	68.81	331.15	77.38

Following sequencing, demultiplexing was done: Generation of FASTQs, UMI tags extraction, alignment of the cDNA inserts against a proper reference genome, barcodes- and UMIs filtering. PCR duplicates were identified and number of reads that provide meaningful information is calculated [21]. The estimated cell numbers were approximately 6,000 cells for the human samples and between 4,300 - 5,288 cells for the murine samples. The level of median genes per cell was below 1,000 genes for all samples (except for the Colo205 cells). All samples showed adequate sequencing saturation at ~90%, whereas only 50-60% of the reads mapped against the transcriptome. The cell recovery rate was above 40% for all samples.

Table 4: Summary of sequencing results of hPBMC samples. Following results after demultiplexing for each sample are shown: Estimated number of sequenced cells, mean reads per cell, median genes per cell, total number of detected genes, median UMI counts per cell, sequencing saturation [%], reads which mapped confidently to transcriptome [%] and the determined recovery rate of cells [%].

Sample	Input number of cells	Estimated number of cells	Mean reads per cell	Median gene per cell	Median UMI counts per cell	Seq saturation (in %)	Reads mapped confidently to transcriptome (in %)	Recovery rate (in %)
PBMC100_dup1	10,000	6,273	45,258	478	1,277	93.9	55.7	62.7
PBMC100_dup2	10,000	6,074	43,411	580	1,729	91.5	55.8	60.7
50_50_PBMCColo205_dup1	10,000	6,000	43,112	3,425	14,814	47.3	70.6	60.74
50_50_PBMCColo205_dup2	10,000	6,000	44,122	3,392	15,275	45.5	70.0	60.0
10x_dataset_3k (fresh) PBMC	n.a.	2,690	69,137	833	2,309	94.3	73.1	n.a.
10x_dataset_4k (fresh) PBMC	n.a.	4,340	87,433	1,235	3,866	90.5	62.3	n.a.
10x_dataset_8k (fresh) PBMC	n.a.	8,381	93,552	1,297	4,084	90.5	61.4	n.a.

Results were set in correlation with public available datasets of 10x Genomics with different numbers of recovered cells [31–33]. Both 10x public datasets with 4k and 8k fresh PBMCs show higher amount of median gene per cell level compared to measured samples.

For the detection of Colo205 cells, Cell Ranger needed to be run with “Force Cell Count” function. In the plot below (Figure 5) a correlation between Barcodes and UMI Counts is illustrated. Instead of one shoulder (if only PBMCs would be present), two shoulders can be observed. Second shoulder was not included in the beginning, that means no Colo205 cells were analyzed in the data. This was corrected by including second shoulder as well using “force cell count” function.

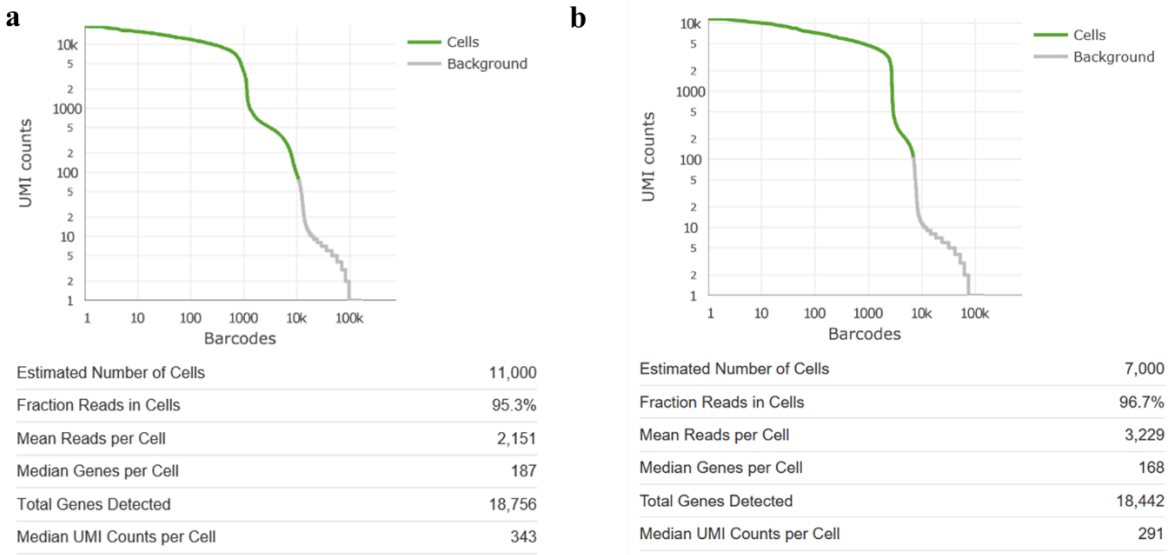


Figure 5: Correlation between barcodes and UMI counts for human PBMC samples. (**a**) Results for 80% PBMCs and 20% Colo205 cells sample duplicate 1 and (**b**) results for 50% PBMCs and 50% Colo205 cells sample duplicate 1 illustrated as kneeplot. The Cell Ranger summary illustrates results of datasets like estimated number of recovered cells (here number was manually set to 11,000 and 7,000 cells to include also PBMCs since they were masked with low RNA counts by Colo205 cell RNA content); fraction reads in cells; mean reads per cell, median genes per cells; total number of detected genes and median UMI counts per cell. The green line in cure visualizes the number of barcodes and UMI counts which have been used for downstream analysis.

Output was further analyzed using Seurat analysis tool (Figure 6): Excluding outlier cells with very high gene count (due to the possibility of multiple cells in one GEM) or very high percentage of mitochondrial genes (dying or dead cells) [34]. A global-scaling normalization method "LogNormalize" was used and followed by the identification and usage of the top 1,000 variable genes. Reduction of the data dimensionality from high- to two dimensionality. Cell clustering was done through graph-based approach (KNN, k-Nearest Neighbor) graph and pair-wise correlation between each cluster was calculated for hierarchical clustering. Cell are visualized in t-distributed stochastic neighbor embedding (tSNE) or in a new dimensional reduction technique called Uniform Manifold Approximation and Projection (UMAP) [35,36]. Using Seurat, gene expression in all cell clusters can be verified.

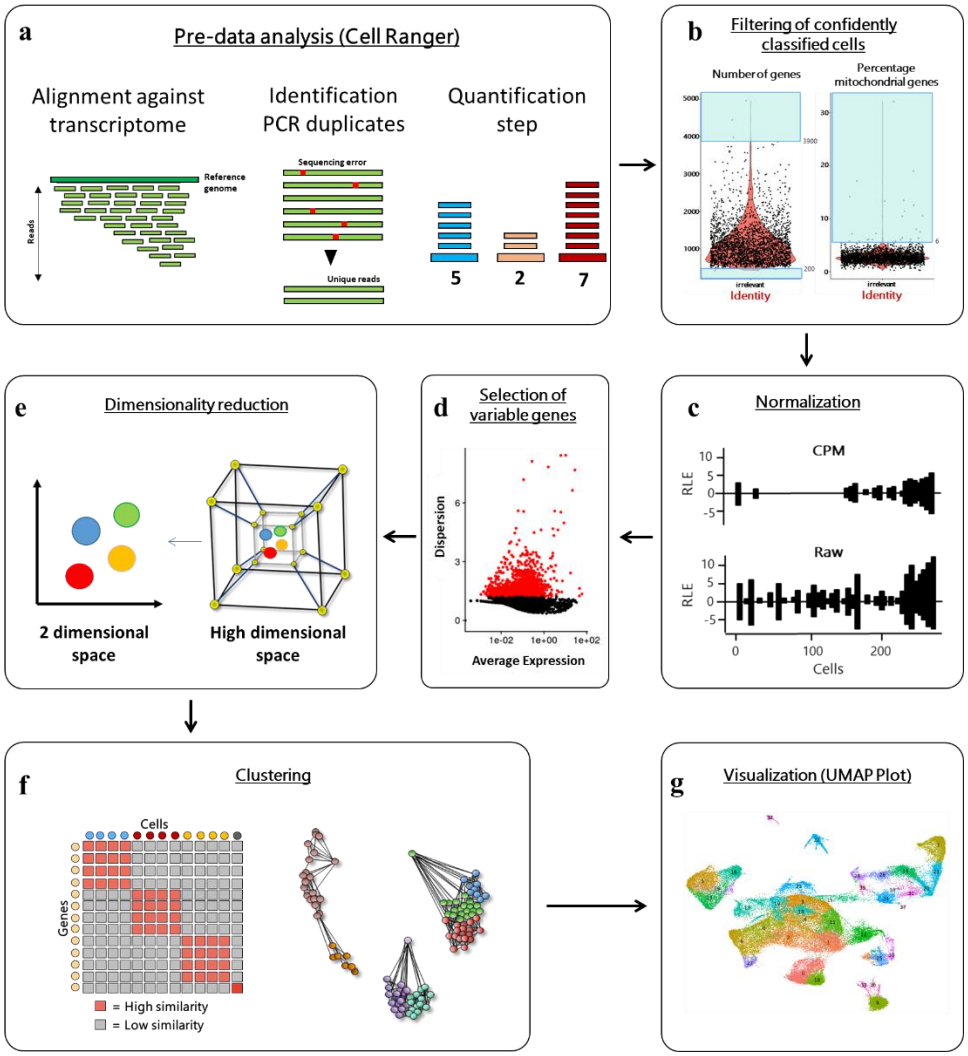


Figure 6: Computational analysis strategy for scRNA-seq data. a) Sequenced cDNA inserts are aligned to an appropriate reference genome using STAR, followed by PCR duplicate identification and read quantification, **b**) Filtering out cells containing high amount of genes (potential multiplets) or high percentage of mitochondrial genes (potential ruptured low quality cells), **c**) Identification of top 1,000 variable genes, **d**) Normalization of the feature expression measurements, **e**) Dimensionality reduction step, **f**) Cells clustering based on the similar or different gene expression profile, **g**) Cell clustering is visualized by UMAP.

For processing a single cell sequencing dataset, essential QC steps had to be performed as described in chapter 6.5.1. After raw data processing, a single cell cluster analysis was done using Seurat (method part 6.5.2) in which the expression levels of cell type specific gene markers were plotted to identify the present cell types in the dataset. During this analysis, cells were filtered out based on very high mitochondrial gene percentage or very high gene count (method part 6.5.2). Specific parameters were used for clustering as final dimension of 14 and a resolution of 0.6. These results are illustrated below in Figure 7 for a mix of hPBMCs and

Colo205 cells replicate 1 (here 80% PBMCs and 20% Colo205 cells) as an example (using tSNE dimensional reduction).

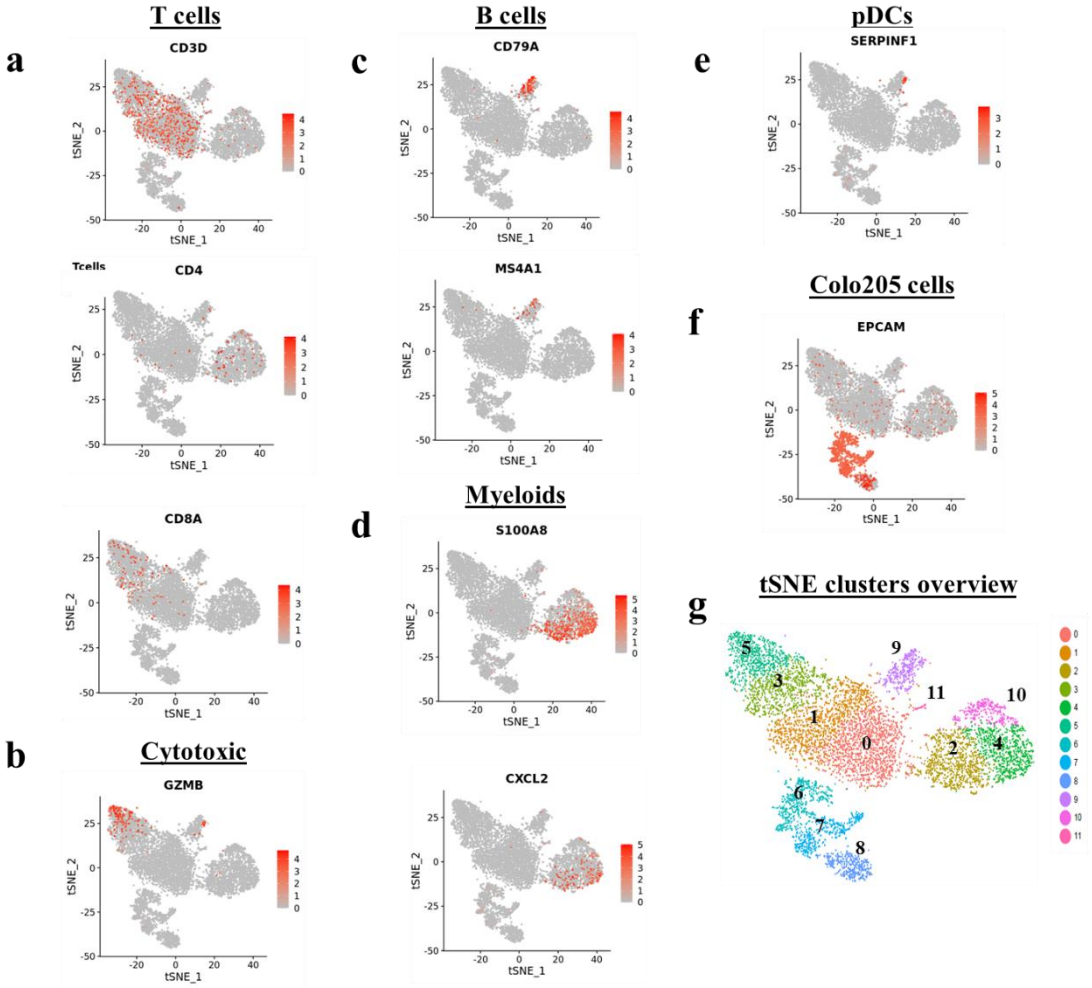


Figure 7: Identification of cell subsets in the human PBMC and Colo250 datasets. Dataset consists of 80% PBMCs and 20% Colo205 cells. Selected marker expression as features visualized on tSNE dimension reduction plot (function called FeaturePlot). At that time tSNE dimension reduction was done for data analysis since UMAP was not developed. Each dot here represents a single cell. The color gradient shows the expression level from low to high with grey to red color. Various representative markers were plotted for the detection of (**a**) T cells, (**b**) Cytotoxic lymphocytes, (**c**) B cells, (**d**) Myeloid cells, (**e**) pDCs, (**f**) Colo205 cells and (**g**) an overview of all identified cell clusters by Seurat [dimension: 14; resolution: 0.6].

Mix of 80% hPBMCs and 20% Colo205 tumor cells is visualized in following Figure 8a-c with all identified cell types. It pursues the goal of this experiment which of the existing cell types can be identified and quantified in the next step. In order to detect the present cell types, the above shown cell type markers (genes in Figure 7) were used for identification and their expression was plotted. Colo205 tumor cells could be successfully identified beside further

human PBMC subsets like T cells, B cells, myeloids and even plasmacytoid dendritic cells (pDCs). The tSNE plot shows that most of the cells identified are Colo205 cells (here Epcam+ cells, Colo205 cells shown in light green). Proving that immune and tumor cells can be identified and distinguished from each other, which was one of the above listed aims under 1.1.2.

The exact cell type frequencies of the human PBMC and Colo205 cells are illustrated in Figure 8d-f. The boxplot illustrates that over 60% of the total cells are attributable to Colo205 and not 50% like originally aimed. Here, the mixing of the cells may not have been performed correctly. In terms of immune cells, T cells made up the majority of the total population followed by myeloids.

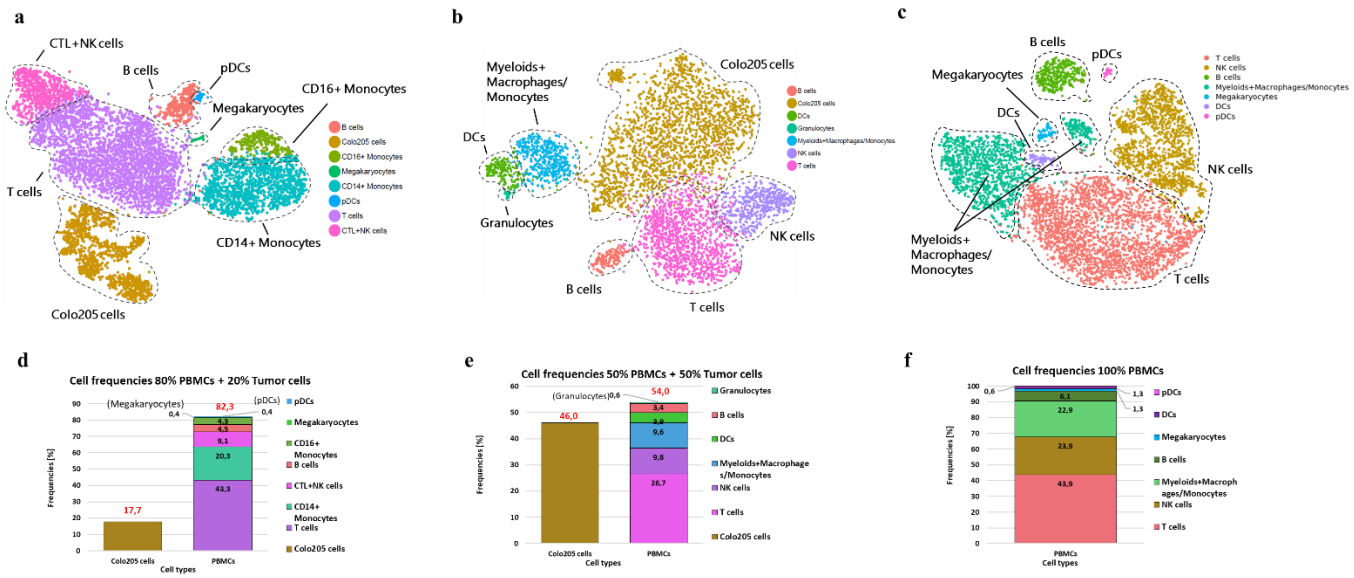


Figure 8: Final tSNE plots and cell frequencies for identified cell types of human PBMC and Colo205 cell dataset. These tSNE plots illustrate all identified cells types. Final tSNE are illustrated for (a) 80%PBMCs and 20%Colo205 cell ratio, (b) 50%PBMCs and 50%Colo205 cell ratio, (c) 100% PBMCs and (d) 100% Colo205 cells side by side. The boxplots illustrate all present cell subsets on the x-axis in different colors, whereas the y-axis features cell type frequencies in %. Cell frequencies of cell subsets are illustrated for (**d**) 80%PBMCs and 20%Colo205 cell ratio, (**e**) 50%PBMCs and 50%Colo205 cell ratio and (**f**) 100% PBMCs respectively.

The results showed that laboratory protocol (Single Cell 3´RNA sequencing by 10x Genomics) and the application of bioinformatics tools allow the differentiation of immune and cancer cells on single cell level. In this experiment, in addition to the percentage quantification of immune and cancer cells, the present immune cell subsets could be detected using canonical marker

genes from literature [21]. Some results like the low number of median genes per cell and the unequal cell type ratio of originally expected 50%/50% immune and cancer cells show that further optimization and testing is needed. These will be discussed in the following chapter 1.2.2.

1.2.2 Examination and comparison of two scRNA-seq capture assays

In order to compare and assess potential differences between the two laboratory protocols 3'GE and 5'GE protocol (capture assays) generating 3´and 5`GE libraries as described in 1.1.2 and 1.2.1 both protocols were used to analyze PBMC samples side by side. In order to verify the quality of the 3'GE and 5'GE protocols the aim was to investigate if any difference in cell recovery or the level of median genes per cell could be detected. In the pre-experiment (1.2.1) the sequencing results were comparable with public datasets from 10x Genomics [31–33] (see above) but could be optimized since the median genes per cell values of the different PBMC samples were half of the expected values than the values of public available datasets as well as the percentage of reads mapped to transcriptome. Therefore, the temperature sensitivity and donor-specific effects of PBMCs were tested in addition, in order to provide more insights if the cell preparation method (fresh or fresh frozen PBMCs) lead to more recovered cells and higher amount of detected transcripts.

The aim was to test the presumably gentler and faster execution of the processing of the PBMCs by adjusting the single cell prep protocol under 6.4.7. In order to compare the difference between preparation methods, temperature and sample size, the following investigations have been done:

- To investigate whether **freezing** impacts cell viability or quality, some PBMC samples were deep frozen in liquid nitrogen and thawed for the experiment, while the other samples were prepared freshly.

- To exclude **donor-specific effects** I compared PBMCs from 3 different donors.

- To test whether the **minimum size of a sample** is affecting the results, I used a smaller amount than the suggested 10,000 cells (for example 5,000 or 500 cells) on the Chromium Controller simulating a potential situation where we have a limited cell input.

All cells were filtered, counted and loaded together onto a single cell chip. Three samples (fresh frozen and fresh PBMC's) were loaded on another Chromium Controller to create 5'GE libraries (loading scheme in Figure 9):

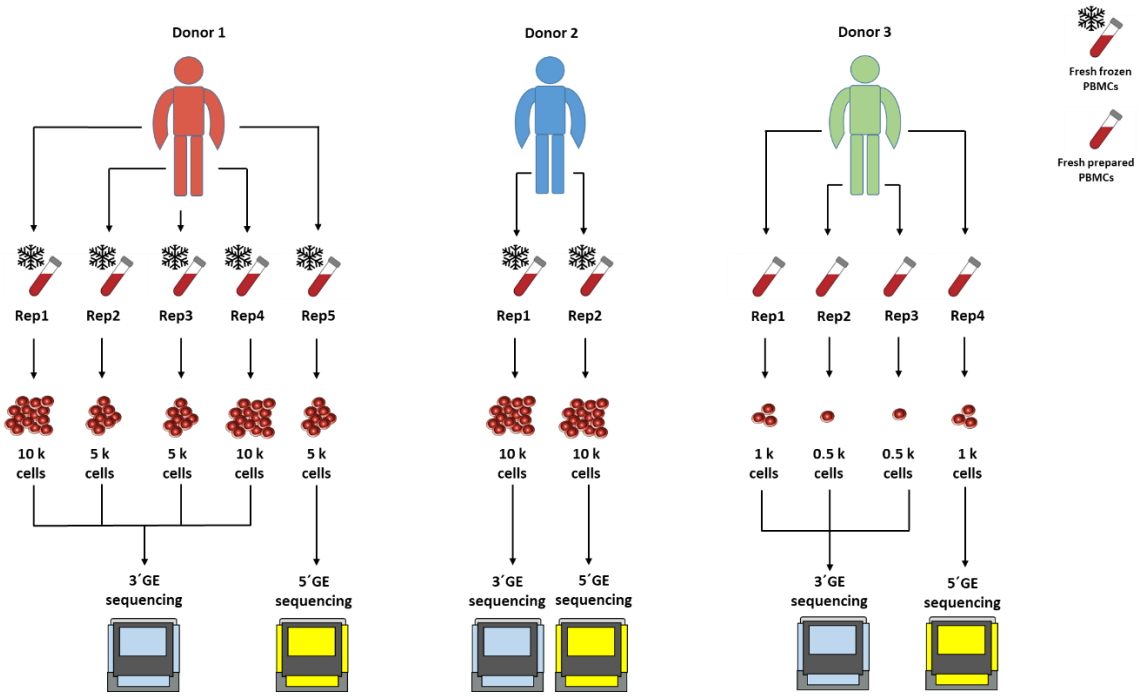


Figure 9: Workflow overview of establishment experiment – Processing three different PBMC donors with different conditions (3'GE and 5'GE protocol). From three different donors, PBMCs were thawed or prepared freshly. Various amounts of cells (0.5 k – 10k) cells were loaded and processed with 3'GE or 5'GE protocol.

During library prep, QC steps were again performed to determine the **concentration** and **fragment size** of **cDNA** and **final libraries** (summary in Table 5) but here in general some inconsistencies were detected: In Figure 10 and Figure 11 fragment profiles of all samples (determined by Bioanalyzer and TapeStation) of the cDNA and final library samples are shown. After the cDNA was amplified and purified, the electropherogram (Figure 10a) of the Bioanalyzer showed that four of the 3' cDNA samples were very low in concentration. Three of the four samples were the freshly collected PBMCs replicates and one was the replicate of the previously analyzed PBMC sample from PBMC donor 1 (FF_PUD_PBMC_D1.4). For the fresh PBMCs, it might be related to the fact that the cells were stored at room temperature for a long time after collection, thus driving down their transcription level. One of the fresh frozen PBMC samples was also low in concentration which could have been a result of improper purification or low cDNA concentration in general. Almost all cDNA samples in Figure 10a

show low concentrations of primer residues at 40-50bp. The mean value of all samples is approximately 1500 bp. In the same plot there was a shoulder behind the marker peak starting at 10380 bp which was seen in almost all samples. This could be PCR bubbles or an error due to the bioanalyzer measurement [37]. However, these fragments could be removed in the course of further purifications. In Figure 10b illustrates all electropherograms of the 3'GE final libraries before sequencing. Almost all samples illustrate a uniform peak profile with no minor primer residues. Two things were observed: First, sample Fresh_PBMC_D3.2 (black) is much less concentrated than the other samples (already as a cDNA sample). This could mean a lower cell recovery and therefore lower amount of detected transcripts. However, the other replicates from the same PBMC donor were similarly concentrated, indicating that the sample was not properly amplified or purified due to a potential dilution error. Second, samples FF_PBMC_PUD_D1.1 (red) and FF_PBMC_PUD_D1.2 (dark blue) show a profile with a smaller mean fragment size, from this it can be deduced that an inaccurate fragmentation of the sample occurred during the library prep. cDNA and final 5'GE libraries were analyzed with the TapeStation. All libraries were subsequently used for sequencing.

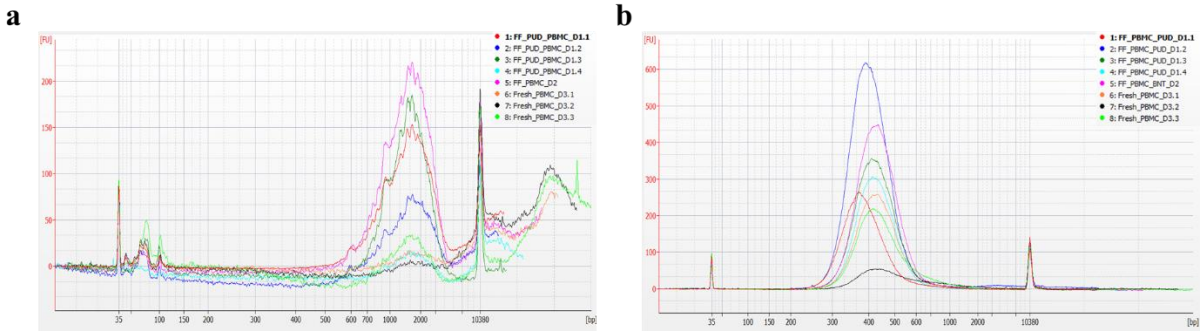


Figure 10: Electropherogram of cDNA and final libraries of human PBMC samples from different donors created with 3'GE protocol. The fragment length is shown on the x-axis [bp] whereas the amount of the fluorescence signals is illustrated on the y-axis [FU]. Both standards can be seen at 35 bp and 10380 bp. The signals of (a) the amplified cDNA samples (b) and the final libraries are illustrated in different colors (FF_PUD_PBC_D1.1: red; FF_PUD_PBMC_D1.2: dark blue; FF_PUD_PBMC_D1.3: dark green; FF_PUD_PBMC_D1.4: light blue; FF_PBMC_D2: pink; Fresh_PBMC_D3.1: orange; Fresh_PBMC_D3.2: black; Fresh_PBMC_D3.3: light green).

All cDNA samples probed by 5'GE protocol are shown in the figure below (Figure 11). The third sample "Fresh_PBMC_5p_D3" was discarded from the analysis because of a technical measurement issue and not analyzed further. The TapeStation run for the Fresh PBMC sample did not work. The HS D5000 should have been used for all three samples, which is why

additional fragments larger than 1.5kb would have been expected. Only two samples are shown here, both showing expected peak profiles at 100 and approximately 500-1500 bp. All the samples were further processed and led to sequence able final libraries.

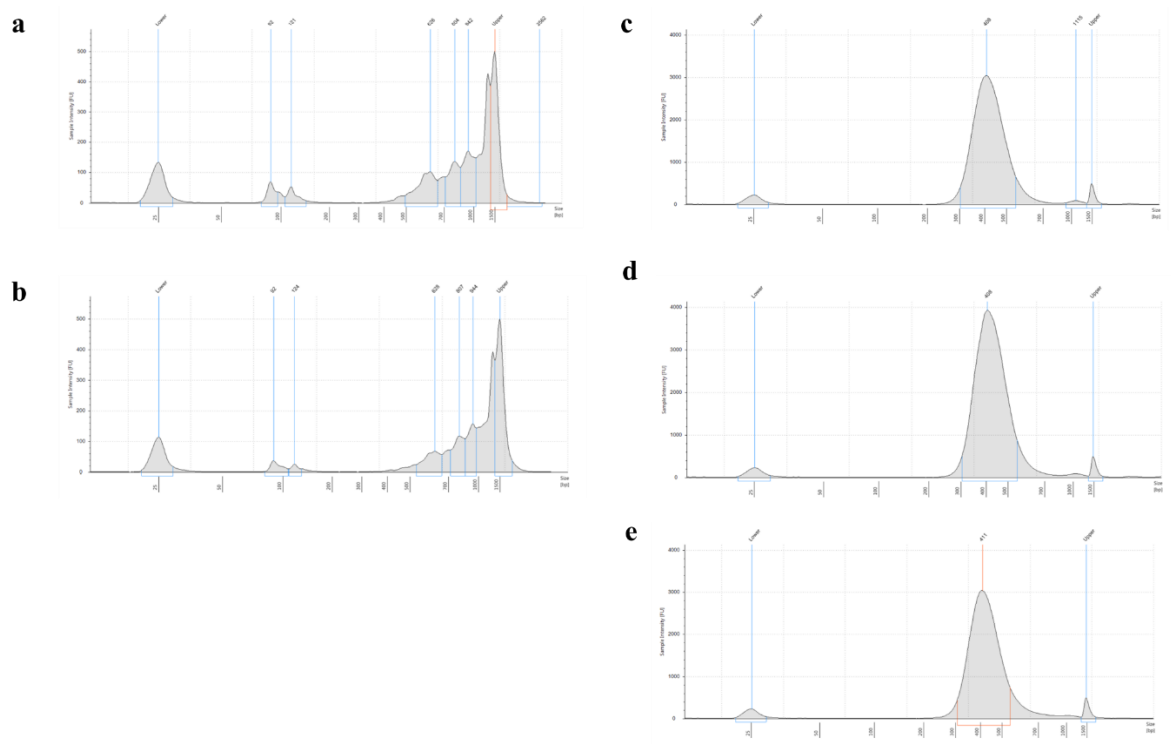


Figure 11: Electropherogram of cDNA and final libraries of human PBMC samples from different donors created with 5'GE protocol. The fragment length is shown on the x-axis [bp] whereas the amount of the fluorescence signals is illustrated on the y-axis [FU]. Both standards can be seen at 25 bp and 1500 bp. The signals of the human amplified cDNA samples are illustrated in (**a**) for sample FF_PUD_PBMC_Donor1 and (**b**) for sample FF_PBMC_Donor2. Profiles of human PBMC final libraries are shown in (**c**) for FF_PUD_PBMC_Donor1, (**d**) for FF_PUD_PBMC_Donor1 and (**e**) for Fresh_PBMC_Donor1.

All samples processed with the 5'GE protocol are much lower in concentration (see below in Table 5). No average value could be determined for these samples using the TapeStation. Almost all 3'GE and 5'GE final scRNA-seq libraries (except three) are comparable in terms of concentration and mean fragment size. Only the sample "Fresh_PBMC_D3.2" has a very low concentration of 1.23 ng/µl. The mean fragment size of the same sample was 559 bp, which was significantly larger than all other samples.

Table 5: DNA concentrations of human PBMC cDNA samples and final libraries. Concentrations in ng/µl were measured by Bioanalyzer, TapeStation and Qubit.

		Qubit	Bioanalyzer	
Sample type	**Sample name**	**Concentration (ng/µl)**	**Fragment size (bp)**	**Concentration (ng/µl)**
cDNA	FF_PUD_PBMC_D1.1	5.72	1626	0.55
	FF_PUD_PBMC_D1.2	6.06	1629	0.18
	FF_PUD_PBMC_D1.3	7.58	1860	1.77
	FF_PUD_PBMC_D1.4	4.3	1807	0
	FF_PUD_PBMC_5p_D1	2.96	N.A.	0.43
	FF_PBMC_BNT_D2	8.36	1426	1.74
	FF_PBMC_BNT_5p_D2	2.72	N.A.	0.26
	Fresh_PBMC_D3.1	6.22	1911	0.23
	Fresh_PBMC_D3.2	5.18	1379	0.68
	Fresh_PBMC_D3.3	7.62	1877	0.99
	Fresh_PBMC_5p_D3	N.A.	N.A.	N.A.
Final library	FF_PUD_PBMC_D1.1	3.72	406	3.64
	FF_PUD_PBMC_D1.2	8.58	417	8.73
	FF_PUD_PBMC_D1.3	6.02	441	5.79
	FF_PUD_PBMC_D1.4	4.5	441	4.62
	FF_PUD_PBMC_5p_D1	8.86	408	9.02
	FF_PBMC_BNT_D2	5.86	446	6.09
	FF_PBMC_BNT_5p_D2	10.05	408	11.1
	Fresh_PBMC_D3.1	3.84	449	4.29
	Fresh_PBMC_D3.2	1.23	559	0.87
	Fresh_PBMC_D3.3	3.04	477	2.79
	Fresh_PBMC_5p_D3	8.60	411	8.0

Subsequently, the samples were sequenced using the Illumina Hiseq4000 device and a HiSeq3000/4000 PE flowcell (Single Index, Paired-end run, Read 1: 26 cycles, Index Read i7: 8 cycles and Read 2: 98 cycles).

After sequencing, generated clusters were checked with the Illumina software "Sequencing Analysis Viewer". The results shown in Table 6 indicate a successful sequencing run with a good Q30 value of over 70%, which could be higher but is sufficient at around 71%. For read 3 more than 75 % of all bases had a quality score of Q30 (1 error in 1000 bases) or more. After filtering, more than 340 million reads could be generated for all three reads. With the error rate of 0.45% it can be seen that with this sequencing run a little less than 1% PhiX Control was pipetted to the sequencing pool.

Table 6: Raw sequencing results of human PBMC samples (created with Illumina HiSeq4000). QC results and statistics of HiSeq4000 run are given as means and were summarized by Sequencing Analysis Viewer. In detail: Cluster PF, Reads PF, Quality Score of 30% or higher and Error Rate.

Read	Clusters PF (%)	Reads PF (M)	% >= Q30	Error Rate (%)
1	71.11	343.45	96,75	0.13
2	71.11	343.45	77.78	0.00
3	71.11	343.45	75.32	0.45

After demultiplexing by Cell Ranger software (results see in Table 7) a level >1,000 median genes per cell could be achieved, except for the "FF_PUD_PBMC_D1.1" sample (not achieved in the last experiment in 1.2.1 The amount of "UMI counts per cell" was continuously >3,000 and comparable among the samples. Some samples had only a low percent match to the reference genome at ~50%. One reason could be a low diversity of different library frequency fragments, and not all fragments map (to a greater extent) against the reference genome. Another potential reason could be that the libraries were not accurately diluted during preparation for sequencing. Sequencing saturation was sufficient with more than 85% for most of the samples. Results show that the optimization of the single cell prep protocol lead to improved results compared to previous experiment (experiment in 1.2.1). A comparison between fresh and fresh frozen PBMCs is not completely possible since fewer cells were loaded between fresh and fresh frozen PBMC samples for fresh PBMC. Though, for both PBMC types (independent of cell input) a reasonable level of median genes per cell was achieved (>1,000 genes) indicating that the influence of a freezing and thawing cycle did not have a very prominent effect in terms of gene detection in this experiment.

Table 7: Summary of sequencing results of human PBMC samples (determined by Cell Ranger software). Following results after demultiplexing for each sample are depicted: Estimated number of sequenced cells, mean reads per cell, median genes per cell, total number of detected genes, median UMI counts per cell, sequencing saturation [%], reads which mapped confidently to transcriptome [%] and the determined recovery rate of cells [%].

Sample	Input number of cells	Estimated number of cells	Mean reads per cell	Median gene per cell	Median UMI counts per cell	Seq saturation (in %)	Reads mapped confidently to transcriptome (in %)	Recovery rate (in %)
FF_PUD_PBMC_D1.1	10,000	4,402	54,225	969	3,398	85,0	51.3	44,0
FF_PUD_PBMC_D1.2	5,000	2,471	93,79	1,048	3,801	90.1	53.2	49.4
FF_PUD_PBMC_D1.3	5,000	2,306	104,089	1,039	3,666	89.3	53.6	46.1
FF_PUD_PBMC_D1.4	10,000	4,027	58,509	1,036	3,561	85.4	49,0	40.2
FF_PUD_PBMC_5p_D1	10,000	4,316	40,298	1,135	3,876	81.9	62.6	43.1
FF_PBMC_BNT_D2	10,000	5,288	46,713	1,041	3,502	81.4	51.1	52.9
FF_PBMC_BNT_5p_D2	10,000	4,703	35,425	1,055	3,335	82.4	66.2	47,0
Fresh_PBMC_D3.1	1,000	698	289,979	1,394	4,517	95.3	52.9	69.8
Fresh_PBMC_D3.2	500	330	322,132	1,418	4,73	95.3	50.3	66,0
Fresh_PBMC_D3.3	500	371	465,445	1,429	4,897	94.3	52.4	74.2
Fresh_PBMC_5p_D3	1,000	654	265,707	1,700	5,302	96.5	70.2	65.4
10x_dataset_3k (fresh) PBMC	n.a.	2,900	24,722	722	2,117	94.3	71.9	n.a.
10x_dataset_4k (fresh) PBMC	n.a.	4,340	87,433	1,235	3,866	90.5	62.3	n.a.
10x_dataset_8k (fresh) PBMC	n.a.	8,381	93,552	1,297	4,084	90.5	61.4	n.a.

As described in the methods section, the Seurat analysis tool was used to perform multiple QC cuts and a reduction in the dimensionality of the data to visually display the data points as dots (single cells) in a dedicated tSNE plot. After bioinformatic data analysis, we also wanted to show that by an alignment of the PBMC donor datasets processed with 3'GE protocol (here subspaces across a given grouping variable are aligned using dynamic time warping, DTW), the identified cell populations for both donors had comparable cell type frequency in percentage (Figure 12, Figure 13a, Figure 14a). In addition, alignment was done between the PBMC donor 1 dataset processed with both 3'GE- and 5'GE protocol (Figure 13d-e, Figure 14b). The investigation of cell type frequencies and expression levels of known canonical markers aimed to show whether one of the two protocols or a PBMC donor has an advantage over its counterpart.

Single cell cluster analysis was done (chapter 6.5.2) using specific parameters as final dimension of 18 and a resolution of 1.8 was used for the alignment of PBMC donor 1 and 2 (3'GE protocol). The plots for the identification of the individual major cell types are shown

below (Figure 12). Besides B cells, myeloids/monocytes, DCs and NK cells, CD4+ T cells as well as CD8+ T cells could be identified.

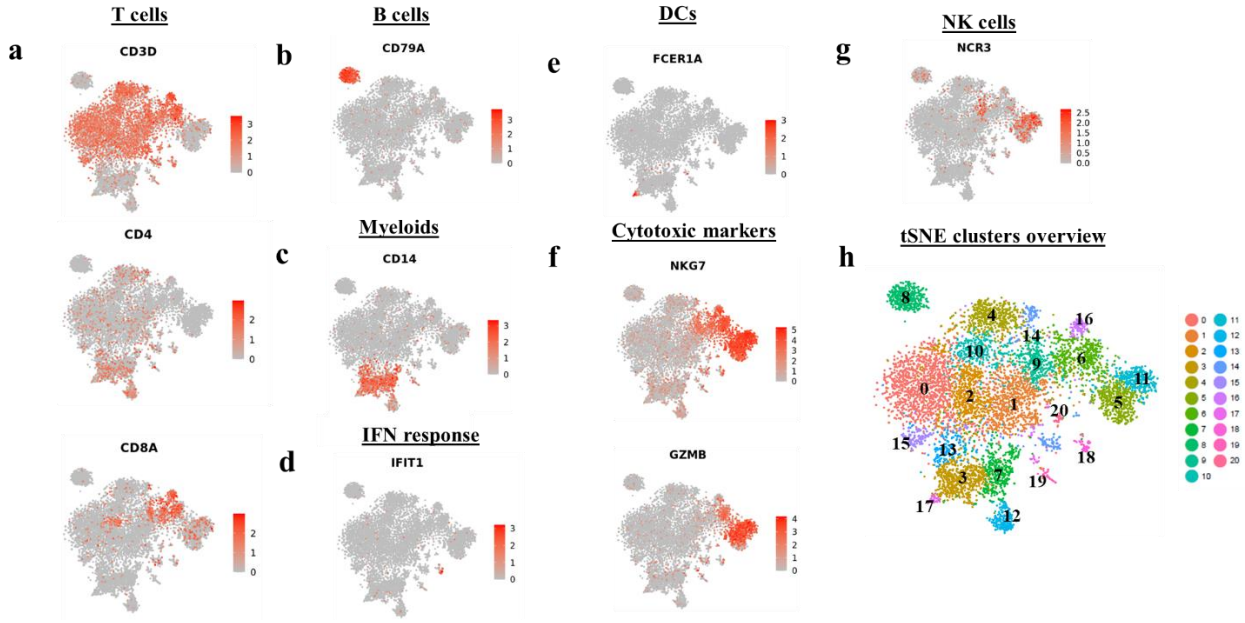


Figure 12: Identification of cell subsets based on the integrated human PBMC datasets donor 1 and donor 2. Marker expression as features visualized on tSNE dimension reduction plot (function called FeaturePlot). At that time tSNE dimension reduction was done for data analysis since UMAP was not developed. A gene which is higher expressed in a cell is colored in darker red. Various representative markers were plotted for the detection of (**a**) T cells, (**b**) B cells, (**c**) Myeloid/Monocytes cells, (**d**) IFN response cells, (**e**) DCs, (**f**) cytotoxic markers, (**g**) NK cells and (**h**) an overview of all identified cell clusters by Seurat[dimension: 18; resolution: 1.8].

After aligning two PBMC donor datasets (FF_PUD_PBMC_donor1 and FF_PUD_PBMC_donor2) and checking various canonical cell markers, 11 populations were identified and visualized in the following tSNE plot as well as their corresponding comparison group like donor or 3'GE/5'GE protocol (Figure 13d and Figure 13e). In general, different T cell subsets clustered not centrally but very dispersed over the tSNE plot (not shown here). For this reason, they were generally annotated as "T cells".

In the following section, the results between the comparison of the 3'GE- and the 5'GE protocols are illustrated. Data from PBMC donor 1 for each protocol was compared, as its PBMCs were already studied in the past experiment (1.2.1). For both data sets, an alignment was performed to compare them adequately (as mentioned above) using specific parameters as final dimension of 20 and a resolution of 0.8. Cell types present were identified by examining cell-specific canonical markers. Same cell types were identified as in the above comparison

between PBMC donors 1 and 2 (Figure 13a and Figure 13b). The selected cell type specific markers are identical to those in Figure 7 (comparison of donors 1 and 2). The clear expression of cell markers for CD8+ T cells, monocytes/myeloids, B cells and cytotoxic cells justifies the robustness of the used protocols and procedures for cell preparation, sequencing and data processing.

To determine PBMC donor specific characteristics, the frequencies of each cell type per donor were quantified and compared in percentage (Figure 13c). Although the frequencies for the CD8+ cytotoxic T cells differed slightly between the two PBMC donors 1 and 2, the data for the other cell types showed almost no deviations from each other. Therefore, it can be concluded that the PBMC samples of donors 1 and 2 do not differ greatly in their composition of immune cells.

A comparison of the detected cell type frequencies between the two data sets (3'GE and 5'GE protocol) should further indicate whether there is a large difference between the protocols (Figure 13f). All cell type ratios were quite similar between the two protocols. Yet there is a small reduction of T cell frequency for 5'GE protocol and slight increase of CD14+ cell frequency upon 3'GE protocol. For both protocols, T cells made up the majority of the cell populations, followed by CD14+ monocytes/granulocytes, which illustrates the comparability of both protocols and no clear striking difference/advantage of one of the GE protocols.

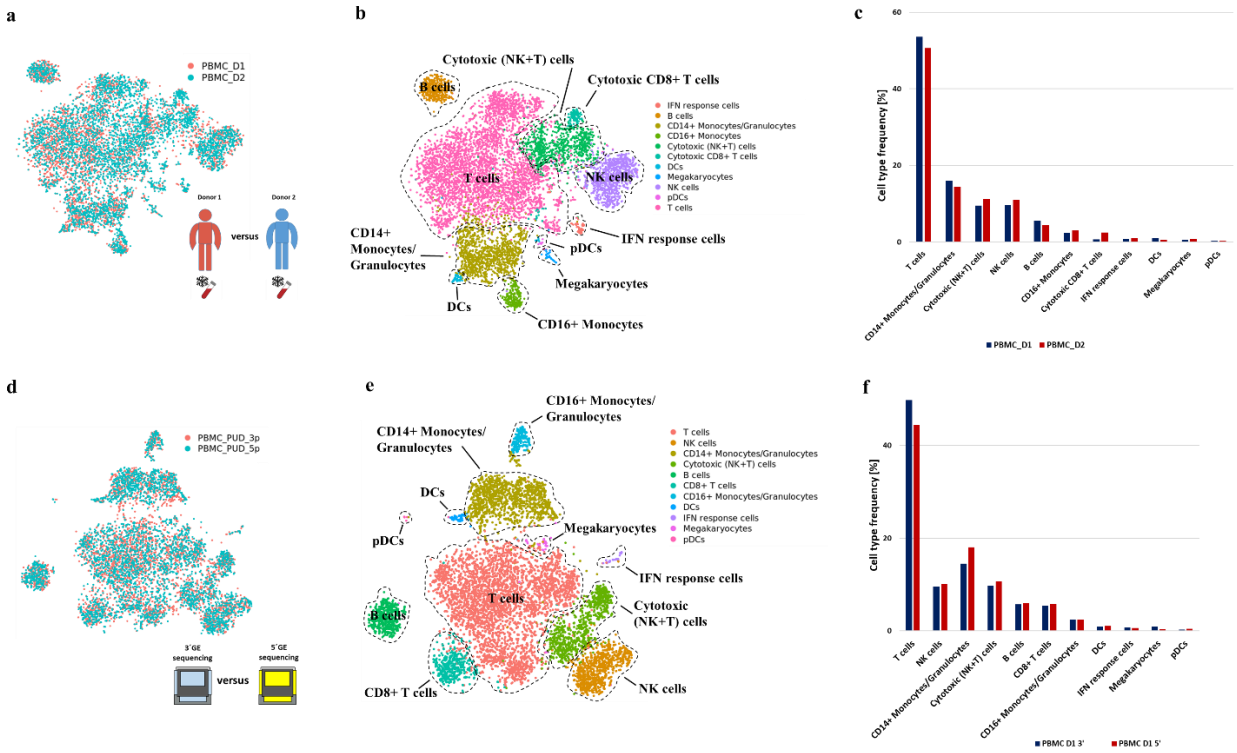


Figure 13: Integrated final tSNE plots and cell type frequencies with identified cell types of human PBMC donor 1 and donor 2 dataset (3`GE protocol) and of human PBMC donor 1 dataset (3'GE and 5'GE protocols).** This tSNE plot shows (**a**) the cells derived from PBMC donor 1 (red) and 2 (blue, 3`GE protocol), (b**) all present cell types of human PBMC donor 1 and donor 2 dataset, (**c**) all determined cell type frequencies for PBMC donor 1 and 2 in different colors (PBMC donor 1: dark blue; PBMC donor 2: red) of all cell subsets on the x-axis in a barplot, whereas the y-axis features cell type frequencies in % of total cells, (**d**) 3'GE and 5'GE protocol information, (**e**) all present cell types from human PBMC donor 1 of 3'GE and 5`GE protocol and (**f**) all determined cell type frequencies for PBMCs donor 1 on the x-axis for 3'GE and 5'GE protocol in different colors (PBMC D1 3'GE protocol: dark blue; PBMC D1 5'GE protocol: red, 3'GE protocol). The y-axis features cell type frequencies in % of total cells.

For a more detailed investigation of the expression profiles between both PBMC donors 1 and 2, a comparative gene expression analysis was performed using the Seurat analysis tool. Cell type specific canonical markers for each cell type were examined to see if and how far the expression of these markers differed between the PBMC donors. In the dotplot below (Figure 14a), there are hardly any major expression differences between the two PBMC donors 1 and 2 in the different cell types. The markers *S100A9* as well as *CD68* were expressed in more cells in the DCs of donor 2 than in donor 1. In terms of phenotype, there is another small difference: the cytotoxic marker *NKG7*, as well as *CD8A*, *CD8B* and *GNLY* are expressed in lower percentages in the cytotoxic CD8+ T cells in the PBMC donor 1 than in donor 2.

This small variation gives indication that there might be a small difference in the cytotoxic cell level between both donors.

The gene expression results (Figure 14b) between the 3`GE and 5`GE protocols in the dot plot show hardly any striking differences. Almost every expression of the canonical markers for each cell type was the same/similar in percentage and intensity between the two protocols. This underlines the above findings concerning similar cell type percentage between the protocols.

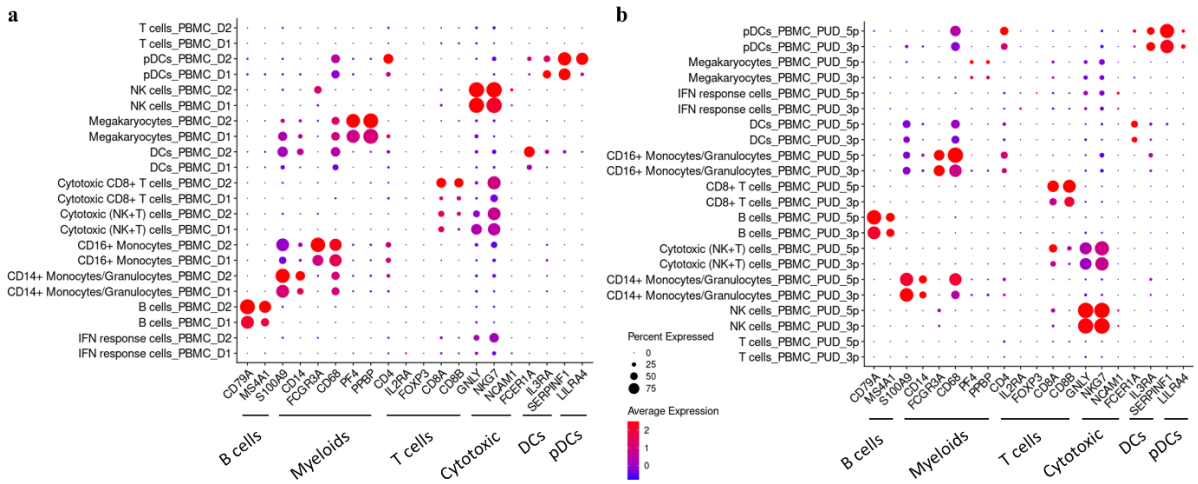


Figure 14: Gene expression levels of canonical cell type markers of human PBMC donor 1 and donor 2 dataset (3`GE protocol) and of human PBMC donor 1 dataset (3'GE and 5'GE protocols). The dotplots display the expression levels of selected canonical markers for all assigned cell types between (**a**) PBMCs donors 1 and 2 , as well as between (**b**) the 3'GE and 5'GE protocols for PBMCs donors 1 on the x-axis, whereas the y-axis the respective treatment group of cells in which a gene is expressed. The size of the dot features in how many cells a gene is expressed [in %] and the color code illustrates the expression level (from blue to red).

In this section, single cell preparation protocols were optimized to improve the robustness of the laboratory procedure and gene expression detection. The optimization led to an improvement of the raw results after sequencing (a higher level of median genes per cell). The presented and similar results comparing two PBMC donors for the 3'GE protocol as well as the comparison of PBMC donor 1 between the 3'GE and 5'GE protocol show that the 5`GE protocol or the 3'GE protocol has no significant advantage over each other. The results of gene expression analysis between the two donors were comparable, supporting the robustness of the 3'GE protocol workflow, including the analysis pipeline.

1.3 Discussion

For better understanding of the immune responses in cancer treatment (or treatment of other diseases) and to close the gap between limited FACS and NGS technologies for transcriptome analysis, it was essential to establish a reliable scRNA-seq method. The establishment and application of this technology (in the laboratory and the analysis of the data) will be discussed.

The first part of chapter one (1.2.1) discusses the results of the establishment experiments with hPBMCs and Colo205 cells. The research aim was to identify all present major cell types and cell subsets, to verify if the adjusted ratios of the mixed cell types (50% PBMCs and 50% Colo205 cells) can be confirmed and if the level of median genes per cell of ~1,000 with 5,000 recovered cells can be achieved with this single cell suspension preparation protocol and data analysis pipeline.

The other part of chapter one (1.2.2) deals with the optimization of the protocol for single cell preparation, the comparison of two scRNA-seq protocols and the investigation/comparison of two different prepared human PBMCs (fresh isolated or frozen). I wanted to study if the optimization (faster and gentler cell preparation) has an improvement in the level of median genes per cell.

1.3.1 Mimicking single cell profiling of tumor and its microenvironment

In summary, after the study of hPBMCs and Colo205 tumor cells we showed that the 10x Genomics scRNA-seq technology Chromium Controller and all necessary bioinformatic processing steps like demultiplexing (Cell Ranger) and the analysis pipeline Seurat are established and ready to use at TRON. A review of the reads mapping against the human or mouse genome showed that there was no cross-contamination with the other species in either the hPBMC or the murine splenocyte samples (results not shown here). This is especially important when, as in this case, the complete single cell chip is loaded with samples. To clarify the scientific question of which cell types are involved in the samples of hPBMCs and murine splenocytes, the expression of specific cell markers was checked using the "Seurat" tool as mentioned in the method section. Here, it can be seen in the feature plots in Figure 7, contrary to general expectation, there was no full expression of the T cell marker CD3D in the T cell cluster. This observation can be explained by the fact that cells may modify specific processes in case of slight variations in their environment as trade-off and so do not show their full gene

expression potential every moment [38]. Sometimes markers are not high expressed at all. In addition, the experiment was done with low sequencing depth.
The method and evaluation of the PBMC and Colo205 cell frequencies show that Seurat is able to perform a quantitative determination of the cell types. With this experiment we could show that in a mixed sample with the different cell types we are able to identify and also quantify them. This would be essential later in the investigation of clinical samples when it comes to the evaluation of an immunotherapy in different experimental groups and the effect on different cell populations.

Following sequencing and demultiplexing by Cell Ranger, a lower level of median genes per cell (<600) was observed for all samples (except with the Colo205 cell mix). In comparison with the results of 10x Genomics (median gene level per cell ~1.000) an improvement needed to be done [31–33]. One reason for lower median gene per cell level may have been the preparation of the cell suspension before loading the cells onto the chip of the Chromium Controller. Cells might have been stored on ice for too long. It might be that the counting of the cells took too long. An optimization should aim at a fast but at the same time gentle cell preparation.

The scRNA-seq workflow proved to be very powerful and highly suitable for the dissection of heterogeneous samples containing different cell populations and species through the evaluation of different cell type frequencies as well as the analysis of the gene expression level of the present cell types and subsets.

1.3.2 Examination and comparison of two scRNA-seq capture assays

With the comparison of different human samples (fresh and frozen PBMCs) and different scRNA-seq methods (3’GE protocol vs. 5’GE protocol), it was verified that there are no remarkable differences in the sequencing results of the PBMC samples (gene expressions). In this work, three different PBMC donors were compared to see if the gene expression levels were similar to each other or if there were large differences between individual donors. At the same time, data were compared between PBMCs that were frozen and thawed for the experiment or freshly harvested and prepared for the experiment. In summary, there seems to be no real benefit of processing fresh PBMC samples directly (mainly based on median genes per cell). Since these samples have been stored at room temperature for longer time, this could

have potentially influenced the expression behavior of the cells. Freezing, storing and later thawing of PBMC may have a time management benefit for many samples depending on the situation. This can be useful if one wants to process more sample for a single run all together. With the intention of detecting any differences between the donors, PBMCs from donor 1 and 2 were examined using the 3`GE protocol. Notably, Kreher et al. 2003 showed in their study that cell content between four healthy individuals was quite similar [39]. During the analysis among the PBMC donors 1 and 2 (final tSNE in Figure 13b), it was found that a more precise assignment of CD4+ T cells was not readily possible. In Figure 13c, the cell frequencies of CD8+ cytotoxic T cells are slightly different between donors 1 and 2. Maybe there is a biological difference e.g. simply fewer CD8+ CTL in one donor. All other frequencies were well comparable. After a more detailed gene analysis between PBMC donors 1 and 2 (dotplot in Figure 14a), only a few expression profiles differed between the two donors. There was also no case where a gene was not expressed at all by one of the donors and was expressed by the other donor. Only the gene *NKG7* was a bit lower expressed by one of the donors. However, this could be related to the fact that not all genes were fully expressed at the time of cell isolation.

In parallel, it was tested whether the expected results of cell recovery of up to 56% could be achieved by using different cell numbers as input. Replicates of each PBMC sample were generated for better comparison. The results in Table 7 showed that a satisfying cell recovery of >50% could be achieved for all samples. This demonstrates the reliability of the protocol and the achievement of a high cell recovery >56% which is superior to the InDrop protocol in capture efficiency. Plate-based methods such as Smart-seq2 can analyze more isolated cells with over 80% capture efficiency, but this requires many more runs as their maximum throughput number is 100 - 1,000 cells per well, whereas the Chromium Controller can load >10,000 cells per sample.

Furthermore, two different protocols for scRNA-seq were compared to check whether one of the 3'or 5'scRNA-seq protocols has an advantage over the other in terms of the resulting amount of detected genes per single cell. Overall, only slightly higher median gene level could be detected for 5’GE protocol but no considerable differences in cell composition and cell-type specific genes were detected. This shows how comparable both protocols are to each other. Thus, both protocols can be used for the identification of existing cell types due to a subsequent

gene expression analysis as long as no further specific downstream analyses are required (more on this point in one of the next paragraphs).

At this stage, the analysis did not go beyond cell type gene expression analysis and its visualization due to the scientific question and the goals (see above, 1.1.2). However, it shall be mentioned that there are other methods of bioinformatic data analysis that can be applied to answer scientific questions. One method is the so-called "differential gene expression analysis" (DGEA). This can be used to investigate the influence of two or more treatments directly in a disease case. This technique can also be used to detect and reveal differences in the subtypes of a cell population of one or more samples. DGEA could have been done between fresh and frozen PBMCs to see if differences in gene expression (e.g. higher percentage stressed/dying cells). The results (also in chapter 1.2.1) have shown that the identified immune cell contents (by scRNA-seq) are comparable to FACS analysis. scRNA-seq can be used unsupervised and is able to (re)examine all detected gene expressions at any time even years after the experiment, e.g. for any gene expressions not initially looked at.

As a result of an internal optimization of the single cell preparation protocol, which includes a faster preparation of the cells, the sequencing parameter of the "median gene per cell" could be achieved to about 1000 average genes per cell, which corresponds to the exemplary results of 10x Genomics [40].

The work on the data analysis was extended in order to answer immunological and scientific questions. In these chapters, murine immune cells (whole CD45+ or antigen specific CD4+ T cells) are analyzed following mRNA treatment (see the following chapters 2, 3 and 0).

Since the 5'GE protocol does not have a significant advantage in terms of detection or gene expression level of the single cells (only a minor higher amount of detected genes and recovered single cells for 5`GE), I decided to continue with the 3'GE protocol for future experiments. If a combination of gene expression analysis (scRNA-seq) and/or only the investigation of T cell receptors at the single cell level (scTCR-seq) is considered, it is indispensable to use the 5'GE protocol. If these protocols can be further developed to analyze transcriptome, proteome, genome and epigenome, metabolome and microbiome in various combination or even all together (see more under FINAL CONCLUSION AND OUTLOOK) will be very useful to answer scientific questions and to accelerate research in cancer and vaccines development. One

approach would be CITE-seq (cellular indexing of transcriptomes and epitopes) using oligonucleotide-labeled antibodies allowing simultaneous measurement into single cell readout, which will be also explained in thesis chapter four (4.1.2 Hashtag oligonucleotide (HTO) application for multiplexing and doublet detection). The technology will be of tremendous value and has been already used for further projects and is an important tool for answering scientific questions in terms of immunotherapy not only in human also in other species like in mice (treated with different kind of mRNAs, see next paragraph and thesis sections).

2 Chapter Two: Intrinsic adjuvanticity of RNA vaccines modulates innate immune system

2.1 Introduction

2.1.1 Cancer and RNA vaccine

One of the most severe but at the same time often largely not completely understood and hard to cure diseases of our times is cancer. According to the World Health Organization (WHO) in 2019 cancer is the first or second leading cause of death before the age of 70 years in 112 of 183 countries [41]. Cancer is a result of mutational and genetic changes and genomic instability which leads to tumor-promoting inflammation and further promote the acquisition of cell death resistance, tumorigenesis, evasion of growth suppression, angiogenesis, proliferative signaling invasion and metastasis and evasion of immune destruction. Research shows that these mutational changes can be triggered by a number of physical, biological and chemical carcinogens or diseases. The physical carcinogens include ultraviolet light and ionization radiation [42], while the chemical carcinogens are related to smoking [41] and asbestos [43] and can promote cancer formation. Another cause are biological carcinogens such as oncogenic viruses or bacteria e.g. human papilloma virus (HPV), hepatitis B/C virus (HBV/HCV)[41] or Epstein Barr virus (EBV) [44].

Due to a lack of knowledge, many tumors are still surgically removed as a cancer therapy which is not a complete solution and involves high risks for the patient. In the worst case, not all cancer cells can be removed from the surgical site and distant tumor sites remain undetected, resulting in cancer recurrence at a later time point. Radiation therapy is one technique to eradicate cancer cells using high-energy x-rays, electron or proton beams from outside the body which are placed directly onto tumors [45]. Chemotherapy is another method and attempted to kill cancer cells in a systemic application since the drugs travels throughout the body, and can prevent proliferation and cell cycle of cancer cells [46]. Yet, it has severe side effects and also often do not lead to complete elimination of the cancer. Further cancer therapy strategies include i) the use of recombinant chimeric antigen receptors (CARs) to redirect immune cells e.g. T cells to recognize and target tumor cell surface molecules as CAR T cell therapy [47], ii) adoptive T cell therapy (ACT) to select, activate and expand *in vitro* highly reactive tumor-specific T cell sub populations and transfer them back into patients to reinvigorate anti-tumor immunity [48] and iii) immune checkpoint inhibitors (CPI) like programmed cell death protein-1 (PD-1) and

cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), are common inhibitory pathways that can be tackled using monoclonal antibodies directed against e.g. PD-1/PD-L1 or CTLA-1 to block the engagement of PD-1 with its ligands, which leads inhibition of immune suppression and an enhanced tumor recognition by CTLs [49]. In contrary, to this immunotherapy has emerged as promising therapy alternative so that research and development of cancer vaccines become an increasingly important field.

The efficiency of cancer vaccines is based on cancer antigens which are either tumor-associated self-antigens (TAA) (cancer/germline antigens) or truly tumor-specific antigens (TSAs) that are not subject to immune tolerance. Cancer vaccines induce tumor-specific T cell or B cell responses, which might lead to tumor rejection by strong immunomodulation. In fact, antigen T cell and NK cell receptors have shown to be of special use as recognition for immunologically target and active proteins in a variety of immune and non-immune cells [50]. DNA, mRNA, viral vectors or dendritic cells (DCs) can be used for antigen delivery. mRNA provides immune activation-mediated co-stimulation and antigen delivery at the same time [50]. mRNA is often used as therapeutic vaccine since it delivers genetic information and has innate immunostimulatory activity [51]. To attenuate the immunostimulatory activity of mRNA, Double-stranded RNA (dsRNA) can be either removed or [52,53] modified nucleosides like pseudouridine, N1-methylpseudouridine, 2-thiouridine, 5-methylcytidine or N6-methyladenosine can be added [54–57]. This will be further explained and assessed in chapter 3 and chapter 4, where mRNA was used to activate the immune system in autoimmune disease and cancer disease settings respectively.

Of the above mentioned the most promising are mRNA cancer vaccines as they are directly injected or administered by loading it ex vivo on autologous DCs and then reinfused via subcutaneous, intranodal or intravenous route [58]. To overcome challenges in delivery of antigens it was shown by Kranz et al. 2016 that DCs can be targeted effectively in vivo using intravenously administered RNA-lipoplexes (RNA-LPX). As RNA is protected by LPX from extracellular ribonucleases, it mediates efficient uptake and triggers interferon-α (IFNα) for the induction of functional anti-tumor T cells and rejection of advanced tumors in mouse models. RNA-LPX further [59] activate plasmacytoid DCs and macrophages [51] and are optimized for MHC class I and II presentation. RNA-LPXs can activate DC maturation in situ and inflammatory immune mechanisms, has an adjuvant effect by strong IFNα response and induce

antigen-specific T-cell responses which can be applied for cancer immunotherapy [51]. According to a study by Sahin et al. 2020 in which a cancer vaccine called melanoma FixVac (BNT111) was used, checkpoint-inhibitor (CPI)-experienced patients with unresectable melanoma show an immune reaction accompanied by the induction of strong CD4+ and CD8+ T cell immunity against the vaccine antigens [60]. Since antigen presentation in the absence of immune stimulation may result in tolerance, cancer vaccines are co-administered with adjuvants. These adjuvants like aluminum salts, monophosphoryl lipid A or toll like receptor agonists mimic evolutionary conserved pathogen- or danger-associated molecular patterns. Adjuvants further include liposomes, lipopolysaccharides, and double- and single-stranded RNA. The antigen-specific immune response can be increased by providing innate immune modulation through the presence of an adjuvant in conjunction with the vaccine. It was further shown that antigen-encoding single-stranded RNA combines the delivery of the antigen and adjuvanticity [53,54,61,62].

To study mechanistic effects of systemic, liposome-formulated RNA-LPX vaccines, several investigations were done by Kranz et al. In this chapter, the single cell genomics team contributed with the mechanistic studies of the induction of i) an adaptive immune response represented by vaccine antigen-specific T cells and ii) induction of the innate immune modulation by using our scRNA-seq pipeline is described. This chapter describes further data which accompanying the study (manuscript currently under preparation).

Conservative (bulk) next generation sequencing methods are not accurate enough since they reflect only average gene expression, loose cellular heterogeneity information and cannot detect cell specific markers [63,64]. The assessment of RNA-LPX effects required a more fine-grained analysis method like scRNA-seq. To understand the treatment effect of RNA-LPX administration on the individual immune cell populations at the site of accumulation and uptake by DCs in a lymphoid tissue such as the spleen, scRNA-seq, as state-of-the-art approach, offers the in-depth characterization of immune cell type composition and mechanistic effects. The aim was to investigate treatment effect on cell types (or immune stimulation) and induction of key molecules shaping the lymphoid environment in response to RNA-LPX treatment. This preliminary scRNA-seq analysis to characterize in depth the immune stimulation was followed by further studies (mentioned in discussion 2.3) including the decoupling of unspecific effects leading to innate immune stimulation and antigen-specific effects leading to adaptive immune

stimulation. This chapter highlights outcomes of this RNA-LPX investigation with its great potential for a solution in understanding and treating severe diseases [65].

2.1.2 Aim of the experiment

In order to investigate the effect of RNA-LPX vaccination on murine splenocytes at single cell level, the established single cell analysis pipeline was used to investigate the effect of RNA-LPX in comparison to a control (sodium chloride (NaCl) control, termed untreated), on murine CD45+ cells. The RNA-LPX used in this thesis encodes "EGFP" (enhanced green fluorescent protein) [66], to measure gene expression, as well for its ability to trigger an immune response in the human or animal body [67].

Murine cell subsets were identified in the vaccinated and untreated replicates in order to assess the impact of EGFP mRNA-LPX. A treatment effect of the RNA-LPX indicated by cell type identification and frequency changes between the two treatments groups was examined, followed by an examination of the gene expression levels of the immune cell subsets in the spleen. For instance, subset-specific co-stimulatory activation was investigated by assessment of activation markers (e.g. *CD40*, *CD86*, *CD69*) on immune cell subsets as well as other inflammatory immune responses compared to the control. Three hours after injection was chosen as an endpoint since the highest RNA content for activation markers was determined in a previous experiment at this time point (data not shown here). Primarily, a former PhD student in TRON constructed the mRNA-LPX, vaccinated the mice, isolated the splenocytes, and conducted the readout of samples through flow cytometry. The subsequent preparation of the single cells and analysis of splenocytes using our single cell pipeline was shared between members of the single cell genomics team at TRON gGmbH.

2.2 Results

This chapter describes the influence of EGFP mRNA treatment in murine splenocytes (CD45+ cells). EGFP mRNA treatment was compared with an untreated control to consider its application for cancer immunotherapy and to ensure vaccine efficacy, optimal-antigen-specific immune response (immunogenicity).

To understand the treatment effect of RNA-LPX murine immune cells were isolated three hours after injection with EGFP mRNA or control and analyzed by 3´scRNA-seq.

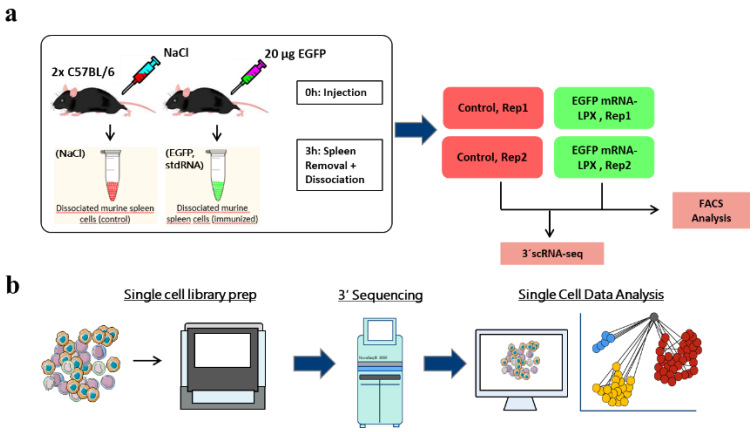


Figure 15: The overview of experimental workflow – Processing treated/not treated murine samples (**a**) Murine samples, treated/not treated with EGFP mRNA. (**b**) Murine samples (splenocytes) were loaded one by one on the single cell chip. Cell input was 10,000 cells for all samples. Samples were processed by our previously established single cell sequencing pipeline.

As a results of the analysis of the QC parameters of the cDNA and final libraries of murine splenocyte samples, almost all samples showed a peak fragment profile at approximately 1200 bp, indicating that so far applied experimental steps led to an expected cDNA fragment profile. cDNA concentrations determined for all samples were quite equal among the samples can be seen in Table 8. As shown in Figure 16b all final libraries have a uniform average fragment size between 452-483 bp with no residual primers accounting for final constructs which can be sequenced. Concentrations can be seen in the table below and were quite equal among all four samples.

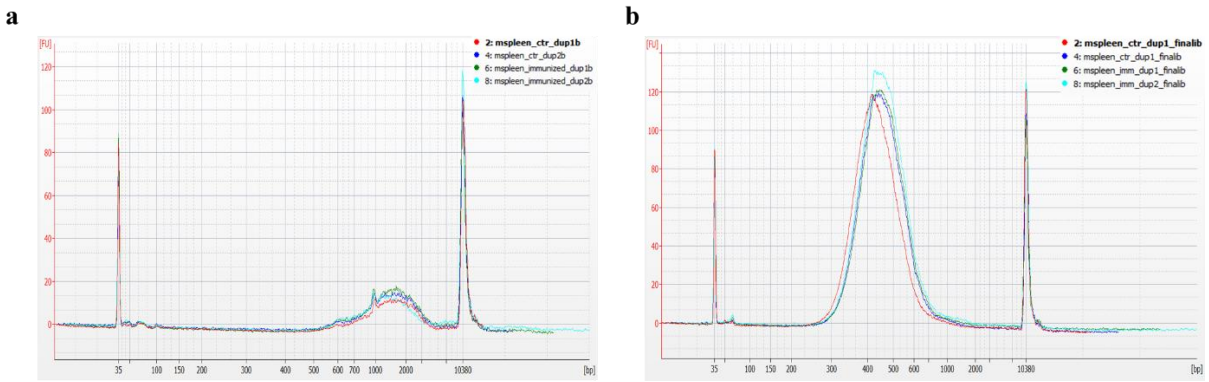


Figure 16: Electropherogram of (a) cDNA and (b) final libraries of mouse samples. The fragment length is shown on the x-axis [bp] whereas the amount of the fluorescence signals is illustrated on the y-axis [FU]. Both standards can be seen at 35 bp and 10380 bp. The signals of the mouse (**a**) amplified cDNA- and (**b**) final library samples are illustrated in different colors (mspleen_ctrl_dup1b: red; mspleen_ctrl_dup2b: dark blue; mspleen_immunized_dup1b: dark green; mspleen_immunized_dup2b: light blue).

Table 8: DNA concentration of mouse samples of cDNA samples and final libraries. Concentrations in ng/µl were measured by Bioanalyzer and Qubit.

		Qubit	Bioanalyzer	
Sample type	**Sample name**	**Concentration (ng/µl)**	**Fragment size (bp)**	**Concentration (ng/µl)**
cDNA	mspleen_ctr_dup1	2.38	1629	0.55
	mspleen_ctr_dup2	2.36	1807	0.93
	mspleen_imm_dup1	2.68	1911	11.96
	mspleen_imm_dup2	1.89	1877	0.70
Final library	mspleen_ctr_dup1	15.7	452	15.55
	mspleen_ctr_dup2	18.7	475	16,37
	mspleen_imm_dup1	16.7	477	19.04
	mspleen_imm_dup2	14.4	483	16.39

Subsequently, samples were sequenced using the Illumina Hiseq4000 device. Generated clusters were checked with the Illumina software "Sequencing Analysis Viewer" (Table 9). More than 68% of the clusters and more than 331 million reads passed filtering. For read 3 more than 77 % of all bases had a quality score of Q30 (1 error in 1000 bases) or more. The error rate was 0 because no PhiX control was used at that time.

Table 9: Raw sequencing results of human and mouse samples. QC results and statistics of HiSeq4000 run are given as means and were summarized by Sequencing Analysis Viewer. In detail: Cluster PF, Reads PF, Quality Score of 30% or higher and Error Rate.

Read	Clusters PF (%)	Reads PF (M)	% >= Q30	Error Rate (%)
1	68.81	331.15	96.29	0,00
2	68.81	331.15	82.36	0,00
3	68.81	331.15	77.38	0,00

Cell Ranger software was performed for demultiplexing to determine the raw sequencing results. As Table 10 shows, the estimated number of cells was between 4,300 – 5,288 cells for the murine samples as expected range for cell recovery. Yet, the level of median genes per cell was below 1,000 genes for all samples (except for the Colo205 cells) indicating that murine splenocytes might have reduced their expression potential during the cell preparation procedure. As all samples showed adequate sequencing saturation at ~90%, whereas only 50-60% of the reads mapped well against the transcriptome. Of special importance is the cell recovery rate of above 40% for all samples as it proves the above-mentioned expected range of cells isolated and recovered by the Chromium Controller.

Table 10: Summary of sequencing results of human and mouse samples. Following results after demultiplexing for each sample are shown: Estimated number of sequenced cells, mean reads per cell, median genes per cell, total number of detected genes, median UMI counts per cell, sequencing saturation [%], reads which mapped confidently to transcriptome [%] and the determined recovery rate of cells [%].

Sample	Input number of cells	Estimated number of cells	Mean reads per cell	Median gene per cell	Median UMI counts per cell	Seq saturation (in %)	Reads mapped confidently to transcriptome (in %)	Recovery rate (in %)
Spleen_control_dup1	10,000	4,321	56,844	729	1,738	91.7	58.0	60.0
Spleen_control_dup2	10,000	4,306	74,004	801	1,956	93.1	59.0	43.21
Spleen_immunized_dup1	10,000	5,288	47,639	675	1,589	93.1	59.0	43.06
Spleen_immunized_dup2	10,000	4,373	62,723	684	1,763	92.1	60.3	52.88
10x_dataset_3k (fresh) PBMC	n.a.	2,900	24,722	722	2,117	94.3	71.9	n.a.
10x_dataset_4k (fresh) PBMC	n.a.	4,340	87,433	1,235	3,866	90.5	62.3	n.a.
10x_dataset_8k (fresh) PBMC	n.a.	8,381	93,552	1,297	4,084	90.5	61.4	n.a.

The results were set in correlation with public available datasets of 10x Genomics with different numbers of recovered cells. Both 10x Genomics datasets with 4k and 8k fresh PBMCs show higher amount of median gene per cell level with more than 60% reads mapped confidently to transcriptome [31–33].

For processing a single cell sequencing dataset, some QC steps had to be performed as described in chapter 6.5.1.. After raw data processing, a single cell cluster analysis (Seurat analysis tool, chapter 6.5.2.) for the integrated murine cell dataset using UMAP dimensional reduction was conducted (Figure 17). The UMAP was previously implemented in the analysis pipeline used as it maintains better the cell-to-cell neighbor information than tSNE and retains better the global structure in scRNA-seq data analysis. It also visualizes better the local structure due to cross-entropy object function and Laplacian Eigenmaps initialization [68]. Using UMAP, T and B cells, macrophages, erythrocytes, cDCs and pDCs were identified (the latter was removed from the downstream analysis and not considered further). Several canonical markers were used for the identification and confirmation of present cell types (dotplot in Figure 18). The UMAP plot below (Figure 17) represents the comparison of both treatment groups (control group and EGFP mRNA) and the identified cell types.

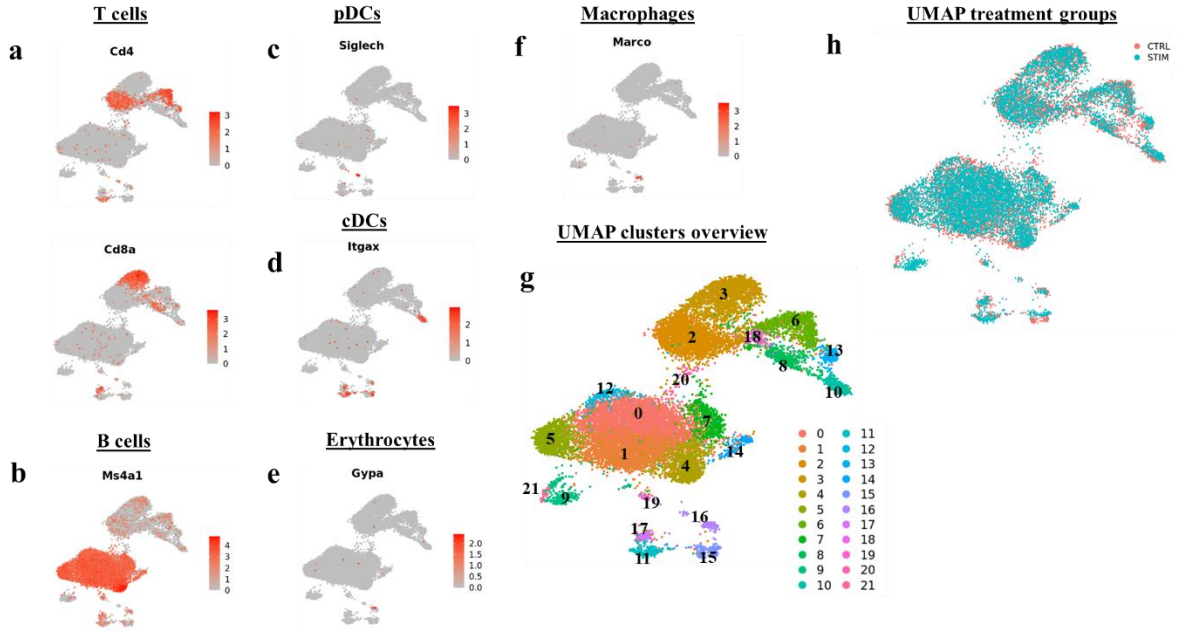


Figure 17: Spatial gene expression levels, cluster and treatment group overview on a UMAP dimension reduction plot. Selected marker expression visualized in UMAP, a function called Featureplot. A gene which is higher expressed in a cell is colored in darker red. Various markers were plotted for the detection of (**a**) T cells, (**b**) B cells, (**c**) pDCs, (**d**) cDCs, I Erythrocytes, (**f**) Macrophages, (**g**) an overview of all identified cell clusters by Seurat [dimension: 22; resolution: 0.6] and (**h**) an overview about the treatment groups for each cell.

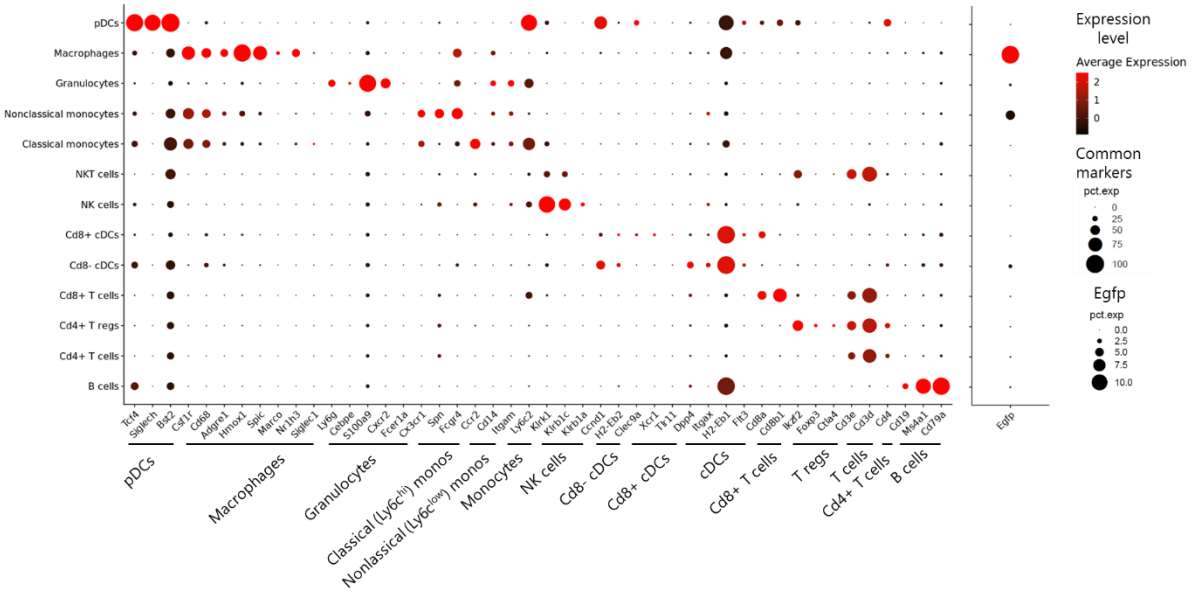


Figure 18: Canonical marker expression of identified murine immune cell types. The dotplot shows the expression of presentative genes on the x-axis, whereas the y-axis the respective treatment group of cells in which a gene is expressed. The size of the dot features indicates in how many cells a gene is expressed [in %] and the color code illustrates the expression level. The EGFP marker is separately illustrated with separate legend for the expression in percentage of the cells.

All major cell types and more refined cell subset such as NKT cells, CD4+ T reg cells or even granulocytes and pDCs can be recognized. In addition to macrophages, classical, non-classical monocytes, CD8+ cDCs (associated with cDC1) and CD8- cDCs (associated with cDC2) were identified [69]. The majority of the identified immune cells consisted of B cells. Even though different B cell subsets (including germinal center, marginal zone, memory and transitional B cells as well as plasma cells, data not shown) were detected they were combined to an overall B cell population due to irrelevance to the analysis. The barplot in Figure 19b illustrates, with respect to the total cell population, an increased frequency of B and CD4+ T cells followed by CD8+ T cells (determined by two methods: FACS and scRNA-seq). The frequency of B cells was slightly increased after stimulation by RNA-LPX treatment. In another study, GC B cell and CD4+ T follicular helper cells (Tfh) frequencies were increased after RNA-LNPs (lipid nanoparticles) and were associated with long-lived and high-affinity neutralizing antibodies and durable protection. Strong mRNA-LPN CD8+ T cell response was detected for two doses in form of INFγ, TNFα and CD107a expression [70]. CD4+ T cells were slightly decreased after RNA-LPX stimulation in this experiment. In general, RNA-LPX vaccines induced memory T-cell and strong effector responses [51]. The frequencies determined between the two protocols

FACS and scRNA-seq were comparable and did not differ greatly. These show a good correlation confirming the overall accuracy of the cell type annotation via scRNA-seq.

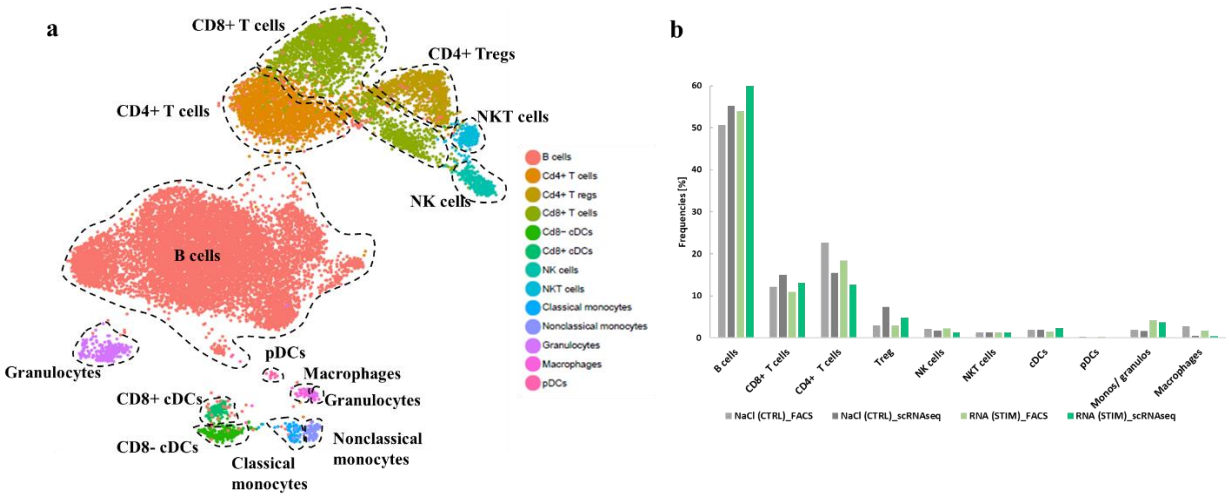


Figure 19: Identified murine immune cell types and determined cell type frequencies. (a) The UMAP plot shows all identified cell types illustrated as cluster of specific color. (**b**) The barplot illustrates all present cell subsets on the x-axis and each treatment group as well both protocols in different colors (Control sample processed by FACS: light grey; control sample processed by scRNA-seq: dark grey; stimulated sample processed by FACS: light green; stimulated sample processed by scRNA-seq: dark green). The y-axis features cell type frequencies in percentage of total cells.

For the differential gene expression analysis clusters/cell types were compared individually between the two treatment groups as described in 6.5.3 25 significantly expressed genes were selected and are presented in a heatmap in Figure 20a because of the role in response to RNA-LPX as chemokines or interleukins. CD4+ and CD8+ T cells are not shown because no significant difference (adjusted p-value > 0.05) was detected. Although CD4+ and CD8+ T cell frequencies are reported to be induced after RNA-LPX treatment [71], it could be that this would happen at a later time point than three hours after injection. An upregulation of chemokines like *Ccl3*, *Ccl4*, *Ccl5*, *Cxcl9* and Cxcl10 could be observed in CD8- DCs, classical monocytes and macrophages. Ccl6 was downregulated in both monocyte subsets after stimulation with RNA-LPX. Classical monocytes showed upregulation of interleukins (*Il15*, *Il18* and *Il27*) and Il-1-like molecules (*Il1b* and *Il1m*) as well as the aforementioned chemokines. Besides this, stimulation with RNA-LPX led to upregulation of *Csf2* and *Ifng* in NK cells.

As enriched pathways are an important indicator for (RNA-LPX) treatment effect, Figure 20b shows the result in form of a heatmap. The enrichment of interferon signaling (type I and type II) was observed in cell types after RNA-LPX stimulation. Granulocytes showed enriched pathways like cytokine signaling in immune system, *Tnfα* signaling via *Nkfb* and Il6_Jak_Stat3 signaling, proving an treatment effect of EGFP RNA-LPX three hours after injection.

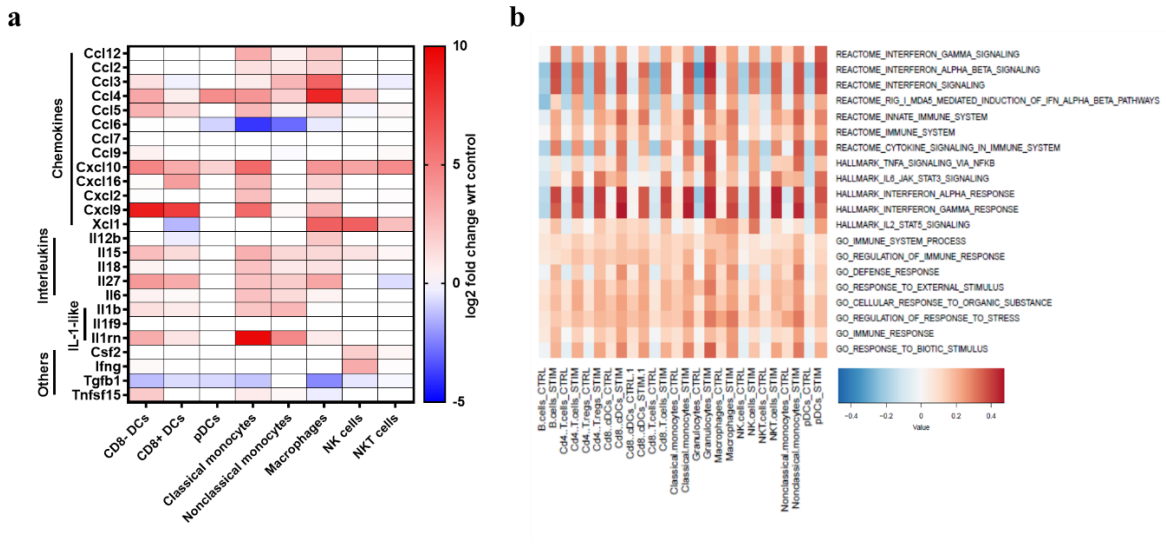


Figure 20: Differential gene expression (a) and pathway analysis results (b) for murine immune cells. The heatmap for top 25 DE genes (**a**) display the gene expression for cells of stimulated group (RNA) and features: Cell types are plotted on the x-axis (except CD4+ and CD8+ T cells), while the DE genes are plotted on the y-axis (based on their log2fold change with control). Genes, which are higher expressed, are colored red, genes which are downregulated are colored blue. (**b**) The heatmap for pathway analysis results further illustrates on the x-axis cell types and treatment information. On the y-axis all enriched pathways from different reference datasets are shown. Enriched pathways are colored in red. Value is also "enrichment score" which is the maximum deviation from zero.

Further relevant genes in specific cell types like CD8- cDCs, CD8+ cDCs and CD8+ T cells with a role in e.g. IFN-response, apoptosis or proliferation are highlighted in the dotplot below in Figure 21 in order to confirm an immune induction by EGFP mRNA-LPX. As CD8- cDCs in Figure 21a showed upregulation of genes of the IFN*α* and IFN*β* response as well as some chemokines involved in the recruitment of immune cells were observed. Further, Il1-like genes are also upregulated as a consequence of RNA-LPX stimulation and classified as defense response to viruses. After RNA-LPX treatment, CD8+ cDCs exhibited apoptosis and proliferation in addition to the pathways mentioned above (Figure 21b). In addition, an upregulation of genes related to cell differentiation was observed in the CD8+ T cells (Figure 21c).

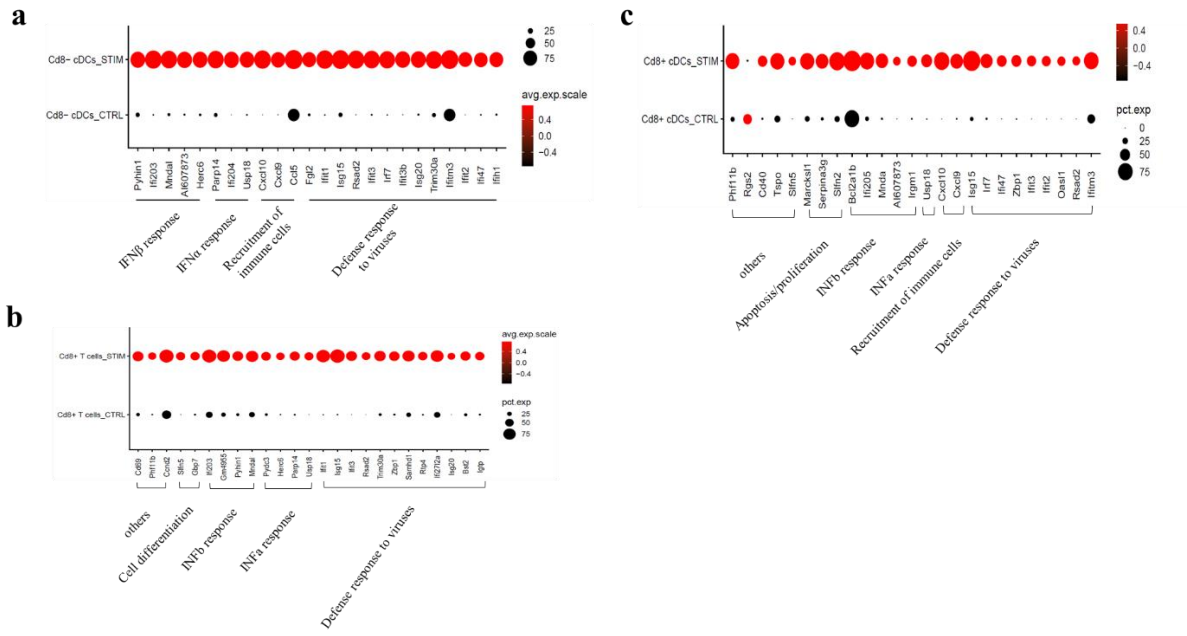


Figure 21: Expression of upregulated genes in three selected cell subsets in response to EGFP RNA-LPX treatment. The dotplot shows the top 25 differentially expressed genes on the x-axis for the stimulated CD8- and Cd8+ cDCs as well as CD8+ T cell subsets, split by treatment group of cells on the y-axis (illustrated by Seurat analysis tool). The genes are ordered according to indicated function. The size of the dot features in how many cells a gene is expressed [in %] and the color code illustrates the expression level.

2.3 Discussion

scRNA-Seq technology was applied in this RNA-LPX study to provide deeper insights into the effect of EGFP RNA-LPX onto the spleen as exemplary lymphoid environment and used to confirm the adjuvant effect of RNA modulating innate immune system. These results are part of a manuscript which is currently under preparation. In this manuscript, two effects of systemic, liposome-formulated RNA (RNA-lipoplex, RNA-LPX) vaccines were studied: i) adaptive induction of vaccine antigen-specific T cells and ii) innate immune modulation. In brief, our single cell RNA sequencing analysis study preceded to demonstrate adjuvant effect of RNA modulating innate immune system at single cell level. Then followed by further studies to assess whether the innate immune modulatory effect observed with RNA-LPX is dose-dependent and to deconvolute the contribution of both components of an antigen-encoding RNA-LPX vaccine by engineering adjuvant RNA (AdjRNA) (follow-up studies mentioned below). Although it was shown that antigen-coding RNA (AgRNA) dose correlates with the strength of antigen-specific T cell immunity and activation of the lymphoid compartment, more

deeper investigations were required: There was a lack of information which exact mechanisms in which cell types might be affected and responsible for an adjuvant effect of RNA modulating innate immune system.

As mentioned in the aims section in 2.1.2, the comparison of cell subset frequencies between the treatment control and EGFP mRNA-LPX treated cells was one major task. It showed that mainly the frequency of B cells and monocyte/granulocyte fractions was slightly increased by EGFP RNA-LPX treatment (Figure 19b). Beside that a slight increase of CD8+ T cells could be observed. Potential further changes in the cell frequencies e.g. for CD4+ T cells might be detected at a later time point since previous studies by Kranz et al. 2016 proved changes in the cell frequencies (higher level of CD4+ and CD8+ T cells) 6-24h after RNA-LPX treatment [51]. The cell frequencies determined by FACS and scRNA-seq, two orthogonal methods, were in a very good agreement and confirmed the accurate identification of cell subsets by scRNA-seq.

Also on the transcript expression level, a potential EGFP mRNA-LPX treatment effect is shown by an upregulation of certain genes involved in IFNα- and INFβ response (*Ifi204*, *Ifi203*) [72,73], recruitment of immune cells or chemokines (*Cxcl9* and *Cxcl10*)[74] or defense responses to viruses (*Irf7*) [75] in murine splenocytes after RNA-LPX treatment (Figure 21). Therefore, the developed RNA-LPX vaccine with preserved adjuvanticity is able to trigger the expression of chemokines and inflammatory cytokines in splenocytes. RNA-LPX vaccine targets selectively DCs resident within lymphoid compartments and is internalized by DCs. Especially, DCs and lymphocytes are expected to be activated through the application of this RNA-LPX treatment. CD8- cDCs, CD8+ cDCs and CD8+ T cells showed an expected upregulation of differentially expressed genes associated with an IFNα and β response. A strong type I IFN innate immune response is reported to be triggered by the RNA and can induce functional anti-tumor T cells and at the same time positively influences the rejection of advanced tumors in mouse models [51,59,76–78]. The same cells expressed some DE genes as a result of RNA-LPX treatment which are summarized in the literature under the “defense response to viruses” [51]. Further DE genes in CD8+ cDCs are mentioned in the literature to play a role in proliferation [50], apoptosis, cell differentiation [79] or in the recruitment of immune cells [50,80]. This reflects and confirms the treatment effect/induction of cell activation by RNA-LPX.

The heatmap in Figure 20a shows upregulation of some chemokines such as *Cxcl9* and *Cxcl10*, which are known as IFNy-inducible gene and involved in the recruitment of CD8+ T cells [81] and NK cells [80]. Macrophages and NK cells also express *Xcl1*, produced by NK cells and activated Cd8+ T cells in order to recruit additional cDC1 cells [82] and induce the CD8+ T cell response [83]. The heatmap shows among the top 25 DE genes also *Il1b*, a pro-inflammatory cytokine promoting the ability to induce T cell activation [84]. Interestingly, IFNα and γ signaling pathways have been enriched, maybe due to some genes which overlapping in GSEA pathways

Furthermore, *Ifng* was also detected as one of the 25 DE genes (up-regulated) in NK cells, as well as *Tgfb1* which was detected down-regulated in CD8- cDCs, classical monocytes and macrophages. This downregulation has an advantage for the treatment, as *Tgfb1* has suppressive functions as a regulatory cytokine in cancer. It is known to promote tumor growth in various skin cancers, so it's down-regulation is a beneficial effect of RNA-LPX treatment [85].

When comparing the results of the GSEA pathway analysis (Figure 20b), some overlaps to the DGEA and the results in the heatmap and dotplot become apparent: Initially, a treatment effect of RNA-LPXs was observed as all enriched pathways were mostly observed to a higher extent in the immune cells under RNA-LPX treatment. As already shown in literature, the IFNα pathway was particularly prominent [51]. RNA-LPX treatment also induces innate immune response and cytokine signaling, which is also reflected in the results and literature [51]. Other expected pathways that were not enriched include upregulation of MHC expression, promotion of maturation, and inhibition of regulatory T cell functions [51]. A potential reason for this could be the above-mentioned phenomenon that the time between RNA-LPX injection and loading the cells was too short. The study of Kranz et al. 2016 in which injection and cell isolation was between 6-24h suggests that the genes for these pathways might not have potentially been expressed yet [51].

The above-mentioned follow-up studies included the analysis whether the innate immune modulatory effect observed with RNA-LPX is dose-dependent. For studying dose-dependency of RNA-LPX mediated adaptive immunity, activation and expansion of CD8+ T cells, mice were immunized with four weekly injections of three different doses and the utilization of Ovalbumin (OVA) agRNA. RNA-LPX immunization resulted in a dose-dependent expansion of circulating antigen-specific CD8+ T cells. The results indicated that antigen-encoding RNA-

LPX concurrent with type I IFN-driven innate activation augments expansion and differentiation of antigen-specific CD8+ T cells in a dose-dependent manner. Another aim was to deconvolute the contribution of both components of an antigen-encoding RNA-LPX vaccine by engineering RNA-LPX consisting of the full RNA backbone but lacking the antigen-encoding open-reading frame (adjuvant RNA [AdjRNA]). AdjRNA immunization, in contrast to the OVA RNA-LPX (antigen RNA [AgRNA]), contributes merely the adjuvant component, but not the antigen. However, since single-stranded RNA is a natural immune sensor agonist, RNA-LPX concurrently triggers a type I interferon (IFN)-dominated, strong innate response. Previous research studies confirmed that type I IFN is of major importance in the induction of functional anti-tumor T cells and rejection of advanced tumors in mouse models treated with RNA-LPX [51,59,76–78]. In summary, results showed that expansion and differentiation of antigen-specific T cells is positively correlated with AgRNA-LPX dose. In addition, once sufficient AgRNA is provided to initiate induction of a low number of antigen-specific T cells, the AdjRNA fill-up can compensate for the AgRNA dose dependency.

The RNA-LPX approach is already applied for the treatment of melanoma by Sahin et al. 2020 but could be even more efficient in combination with other therapies such as checkpoint inhibitors (CPI), monoclonal antibodies (mAbs) or radiotherapy [86]. For mice and human beings, the frequencies of Th1 cells were very high and a strong type I IFN response was observed after this treatment, as also shown in other studies [51]. The necessary regulation or inhibition of regulatory T cells can be achieved by a strong type I IFN response and contributes to an immune-antitumor therapy. This makes it an important tool for the development of new drugs and therapies against diseases such as cancer (chapter 4). In conclusion, the study by Kranz et al. including our scRNA-seq analysis could confirm inflammatory properties and could show that RNA-LPX is a promising tool for treating disease like cancer. On the contrary, in absence of the inflammatory properties and lack of the adjuvant effect, autoimmune diseases like multiple sclerosis can be addressed via RNA-LPX treatment as will be shown in the following chapter 3. These findings illustrated the potential of the scRNA-seq technology for providing deeper insights into the lymphoid environment and their responses after RNA-LPX treatment. They further indicated that the RNA-LPX treatment is a promising immunotherapy in patients which needs to be further assessed and investigated to understand its full potential and precise impact on immune environment as well as well as factors influencing the treatment effects i.e. attenuating or potentiating them.

This chapter illustrated the immune stimulation and induction of key molecules shaping the lymphoid environment in response to RNA-LPX treatment by scRNA-seq. This characterization indicated which impact and responses a treatment with RNA-LPX might have for future treatment of various diseases. Immunogenicity mechanism of activation and the innate immune response have been shown to have an adjuvant effect on the murine splenocytes on their gene expression level compared to the control.

The adjuvant effect of RNA-Lipoplex (LPX) vaccination encoding the reporter gene enhanced green fluorescent protein (EGFP) on murine splenocytes was explored, indicating an increase in the frequency of B cells and monocyte/granulocyte fraction. This led to the triggering of chemokines and inflammatory cytokines including upregulation of genes involved in interferon alpha (IFNα) and beta (INFβ) response, suggesting a heightened immune response and confirming the adjuvant effect of RNA by modulating innate immune system.

3 Chapter Three: Investigation of antigen-specific CD4+ T cell heterogeneity in a mouse model for multiple sclerosis demonstrates induced tolerance response

3.1 Introduction

3.1.1 Multiple Sclerosis and Experimental Autoimmune Encephalomyelitis (EAE)

The human body has developed protective responses against pathogens by the innate and adaptive immune systems. These tolerance mechanisms like presentation of self-antigens of immature DCs maintain immune homeostasis and control self-reactive cells [87], but tolerance regulatory pathways can fail, leading to a breakdown of self-tolerance (autoimmunity). The first case of autoimmunity was recognized in 1904 through the observation of autoreactive antibodies reacting to self-blood cells of patients [88].

Multiple sclerosis (MS) is a chronic, immune-mediated, and multifocal demyelinating autoimmune disease of the central nervous system (CNS) and it is one of the most common causes of neurological disability in young adults. Environmental factors, in particular vitamin D or ultraviolet B light (UVB) exposure, Epstein–Barr virus (EBV) infection, obesity and smoking, but also many genes modestly increase MS disease susceptibility [89]. MS leads to large focal lesions in the white matter of the brain and spinal cord, characterized by primary demyelination with a variable extent of axonal loss [90]. Autoreactive CD4+ T cells can trigger MS disease by escaping negative selection and clonal deletion pass through a disrupted blood–brain barrier and enter the CNS [91]. Those activated CD4+ T cells lead to the recruitment of other inflammatory cells, such as microglia, macrophages, and B cells, which in turn drive the production of the antibodies and pro-inflammatory cytokines that destroy the myelin sheath (Figure 22) [92,93]. As a result, there is an imbalance between effector T cells which are driving the disease and regulatory T cells (Treg cells), suppressing effector T cells [94]. This results in an inability of nerve cells to send impulses to the synapses due to damaged myelin sheaths. Another important immune cell subsets are DCs acting as sentinel by presenting antigen linked to their major histocompatibility complex (MHC). They become mature DCs and upregulate costimulatory molecules [95]. DCs can reactivate autoreactive CD4+ T cells and memory T cells, and prompt their polarization into CD4+ Th1 and Th17 effector T cells but they also

mediate and activate Treg cells as immune supervisors which inhibit the inflammatory immune response [96].

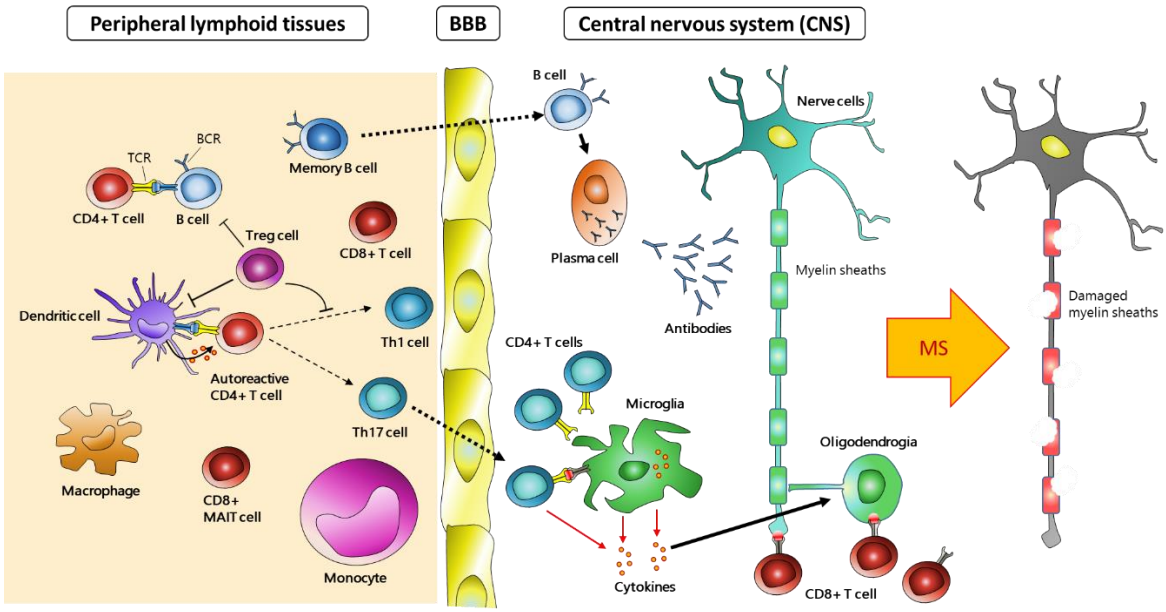


Figure 22: Immunopathology of multiple sclerosis. In the periphery, CD4+ effector T cells, Th1 and Th17 cells are activated by antigen-presenting cells (APCs) via molecular mimicry or bystander activation. Effector cells can migrate across the BBB into the CNS and get reactivated by resident APCs like DCs and microglia. CD4+ effector T cells can directly damage the myelin sheaths by the expression of cytokines or secrete proinflammatory mediators to recruit and activate further immune cell subsets like CD8+ T cells, B cells or plasma cells, which can damage the CNS as well (adapted from [97])

MS can affect any part of the CNS, and thus, its clinical manifestations are often diverse, with signs and symptoms that vary widely depending on the extent and location of the damaged areas [98]. The characteristic clinical symptoms include neurological defects such as sensory disturbances, temporary loss of vision, motor impairments, fatigue, pain and cognitive deficits [99].

3.1.2 Immunotherapeutic strategies for MS

In fact, most of the available therapies for MS target CD4+ responses, but the use of new techniques has made it possible to test the CD4+ model more stringent, showing that some of its predictions are not borne out, and providing evidence for the involvement of other immune cells, such as CD8+ cells and B lymphocytes [100]. Since B cells play a central role in the pathogenesis of MS, depletion of B cells can be aimed as one alternative approach for MS treatment. Anti-CD20 chimeric antibody Rituximab is known to be used for patients with relapsing-remitting MS by complete B cell depletion. Mack et al. 2021 showed that another

antibody is much more effective namely anti-CD79b which led to partial B cell depletion and abrogates MOG-specific antibody production. It also may preserve regulatory B cells and lead to an increased MOG-specific release of IL-10 and significantly reduced the development of EAE progression in mice, but it does not directly affect CD4+ effector T cells driving the autoimmune disease [101]. The humanized monoclonal antibody Natalizumab is directed against very late activation antigen 4 (VLA-4). Blockade of VLA-4 with natalizumab is thought to not only interfere with autoimmunological mechanisms but also with central nervous system immune surveillance [102]. Ocrelizumab is an intravenously administered, humanized anti-CD20 monoclonal antibody approved from the WHO for the treatment of adults with relapsing forms of multiple sclerosis (RMS) or primary progressive multiple sclerosis (PPMS) [103]. The study of Dobson and Giovannoni (2018) highlighted an overview of disease-modifying therapies, but this thesis focus on the aspect of an mRNA application explained below [104].

For animal model-based investigations of MS, a mouse model was established called experimental autoimmune encephalomyelitis (EAE). EAE can be triggered by immunization with a myelin oligodendrocyte glycoprotein peptide (MOG 35–55) and strong adjuvant. MOG is an autoantigen associated with multiple sclerosis. In EAE, MOG-specific T cells drive disease, causing progressive nerve damage [105]. An alternative approach has been investigated by Kariko and colleagues, who developed a modification of mRNA-based vaccines that incorporated 1-methylpseudouridine (m1Ψ), a naturally occurring component of eukaryotic 18S ribosomal RNA. In place of uridine, they noticed that 1-methylpseudouridine (m1Ψ) mRNA had a higher translational capacity and lower innate immune activation [106,107]. Furthermore, due to the replacement of uridine (U) by incorporation of m1Y during in vitro transcription and subsequent removal of double-stranded mRNA contaminants, TLR7-triggering activity is abrogated and inflammatory properties of single stranded mRNA is reduced [53,54,107]. To study the effects of m1Ψ mRNA in an autoimmune disease MS, EAE mouse model was used in which tolerance induction was already demonstrated by selectively expressing MOG35-55, the epitope of myelin oligodendrocyte glycoprotein, in dendritic cells [108]. Krienke et al. showed the effect of antigen-encoding m1Y mRNA treatment on Naïve Thy1.2+ C57BL/6 mice after immunization with MOG35-55 m1Ψ or U mRNA. Both endogenous T cells and adoptively transferred MOG35-55–T cell receptor transgenic Thy1.1+ CD4+ T cells from 2D2mice T cell were expanded. However, CD4+ T cells of MOG35-55U

mRNA or control-treated mice showed little to no suppressive activity. Whereas, MOG35-55U mRNA–expanded T cells secreted IFNy and thereby exhibited a functional TH1 effector profile.

MOG35-55–specific splenic CD4+ T cells from MOG35-55U mRNA treated animals showed down-regulation of the activation marker CD44 and strong expression of co-inhibitory molecules. With application of non-immunogenic m1Ψ mRNA, co-stimulatory signals got lost and autoreactive Teff cells were not induced anymore. Instead, Treg cells expanded and regulate effector T cells negatively. Importantly, the m1Ψ mRNA formulation also stimulated proliferation of suppressive, MOG-specific FoxP3+ Treg cells (Figure 23) [105]. As part of the study, Krienke et al. collaborated with us, the single cell genomics team at TRON gGmbH, using our scRNA-seq analysis pipeline, to analyze distinct antigen-specific CD4+ T cell subsets which might be expanded in EAE treated with antigen-encoding m1Ψ mRNA. In more detail, previous research has been done to identify the most affected immune cells and their location. The group of Stephen Miller suggested that treatment-induced Treg cells are capable of regulating the trafficking of effector T cells, which could explain the observed accumulation of lymphocytes in the spleen [109]. Thereby, Treg cells dissociate Teff cells in the spleen, preventing them from trafficking to sites of inflammation [110,111]. The main research goal of Krienke et al. was to test the non-immunogenic autoantigen-encoding m1Ψ RNA-LPX hypothesis and if systemic tolerogenic antigen presentation in lymphoid tissue in a noninflammatory context is possible. This should be achieved using nucleoside-modified and purified mRNA (as mentioned above) coding for disease-related autoantigens, resulting in antigen presentation on splenic CD11c+ antigen presenting cells in the absence of costimulatory signals. Moreover, Krienke et al. wanted to assess the impact of autoantigen-encoding m1Ψ RNA-LPX (MOG35-55_m1Ψ) vaccination on EAE disease development in comparison with immunogenic autoantigen-encoding RNA (irrelevant m1Ψ RNA-LPX, irrelevant_ m1Ψ). For deeper characterization of these individual cells upon the respective treatments they collaborated with us and we used our scRNA-seq pipeline since it enables detection of complex and rare cell populations as well as the investigation of regulatory relationships and enrichments of specific genes in distinct antigen-specific CD4+ T cell subsets. Hence, we investigated the effect and the differences between MOG35-55_ m1Ψ RNA-LPX and irrelevant m1Ψ RNA-LPX. Beside this published research, another treatment group, autoantigen-encoding U RNA-

LPX (MOG35-55_U), was included in our collaborative scRNA-seq experiment. However, it was not discussed in the published manuscript. This thesis chapter deals with unraveling the role of MOG35-55_U in comparison to the in depth studied MOG35-55_ m1Ψ RNA-LPX and irrelevant m1Ψ RNA-LPX groups on EAE disease development.

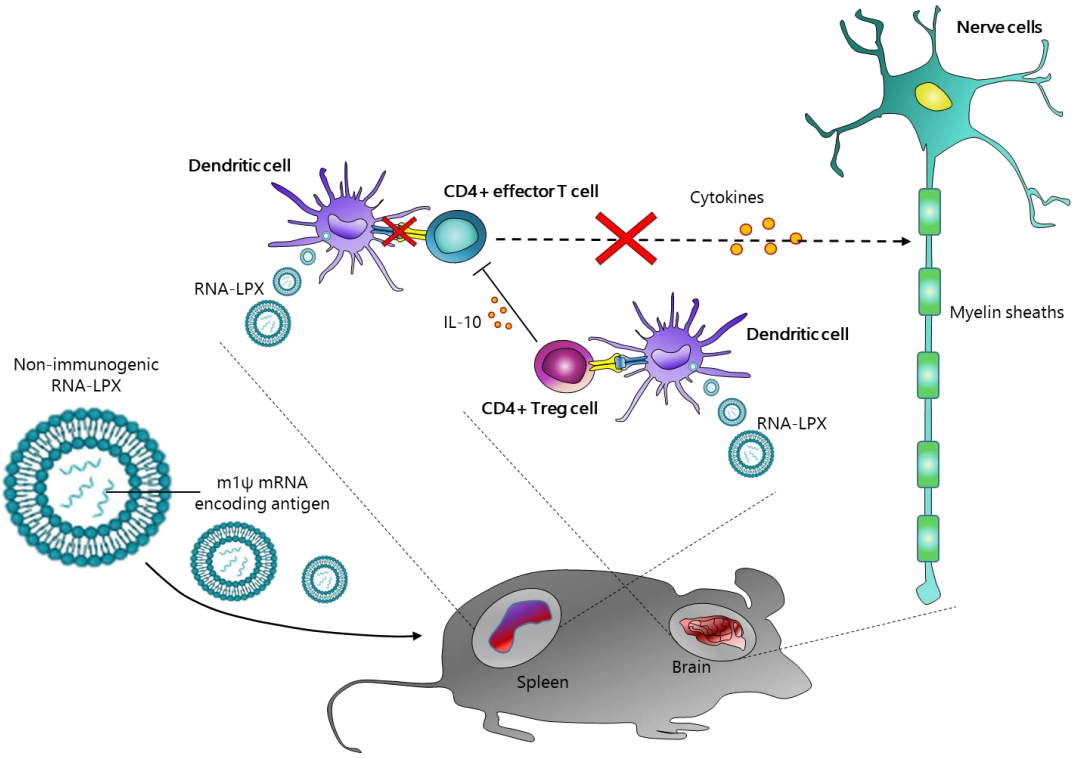


Figure 23: Tolerance induction of non-immunogenic MOG35_55_m1Y mRNA construct in EAE sick mice After intravenous injection in mice, myelin oligodendrocyte glycoprotein (MOG) m1Ψ mRNA in anti-inflammatory lipid nanoparticles are taken up by DCs. MOG m1Ψ mRNA as non-immunogenic RNA-LPX triggers the proliferation of CD4+ Treg cells, which mediate antigen-specific suppression and inhibit co-stimulation of CD4+ effector T cells (Th1 and Th17 cells). CD4+ Treg cells express IL-10 that suppresses also CD4+ effector T cells, preventing further myelin damage and paralysis (adapted from [112])

3.1.3 Aim of the experiment

Krienke et al. aimed for the study of MOG35-55_m1Ψ RNA-LPX and if tolerance induction can be confirmed in the experimental autoimmune encephalomyelitis (EAE) murine model of multiple sclerosis (MS). As part of Krienke et al. study we supported their investigations by using our scRNA-seq pipeline to look at the effect of immunogenic autoantigen-encoding RNA as well and compared their impacts on splenic CD4+ T cells at single cell resolution and gene expression levels. In this thesis, I further aimed to dissect effect of MOG35-55_U to assess the

role of MOG35-55_U and its impact on EAE sick mice in comparison to the in depth studied MOG35-55_ m1Ψ RNA-LPX and irrelevant m1Ψ RNA-LPX groups.

After successful tolerance induction in EAE mice among MOG35-55_ m1Ψ treatment, the aim was to investigate functional differences with respect to inflammatory and immunosuppressive properties of the different cell subpopulations present among antigen-specific CD4+ T cells between the different treatment groups by performing scRNA-seq. One hypothesis was that in vivo delivery of non-immunogenic RNA encoding disease related antigens into DCs, would enable systemic tolerogenic antigen-presentation in lymphoid tissues. After the treatments, only antigen-specific CD4+ T cells were investigated as they are known to be the key players and strong effectors in EAE or multiple sclerosis, especially the effector T helper cell subsets Th1 and Th17 [97]. Notably, two of these conditions (MOG35-55_ m1Ψ and irrelevant_ m1Ψ) were compared in a previous study by Krienke et al. 2021 [105]. Here, in contrast to Krienke et al. study, the comparison was extended to three conditions including the immunogenic autoantigen-encoding MOG35-55_U in order to shed light on the effect on autoantigen presentation under immunogenic conditions. Some preparations have been done beforehand: the cloning and final construction of the mRNA, utilization of synthetic peptides, the final generation of non-immunogenic mRNA, inducing EAE using the MOG35-55 peptide, monitoring and determining protective immunity in the mice, anesthesia and sacrifice of mice, mouse tissue preparation for single cell suspension, magnetic-activated cell sorting (MACS), tetramer staining and cell sorting via flow cytometry. Following preparation of the single cells, implementing and using single cell analysis to investigate enriched pathways and gene, using our single cell pipeline fell under my purview.

3.2 Results

This study focused on the investigation of the antigen-specific T cell heterogeneity on experimental autoimmune encephalomyelitis (EAE) mouse model for multiple sclerosis after RNA vaccination. In detail, we examined differential gene expression in different detected T cell subpopulations after treatment with three different (modified) MOG35_55 encoding mRNA-LPX: the MOG35-55_ m1Ψ, the MOG35-55_U and the irrelevant_ m1Ψ. After treatment, murine spleens were removed and antigen-specific CD4+ T cells were isolated, filtered to be singlets (see methods 6.3.3, 6.3.4 and 6.4.7) and processed by our single cell pipeline. From each treatment group replicates were prepared and loaded on a single cell chip

(Figure 24). The exact number of loaded cells is unknown since they were sorted, filtered and cell concentration was too low to be counted.

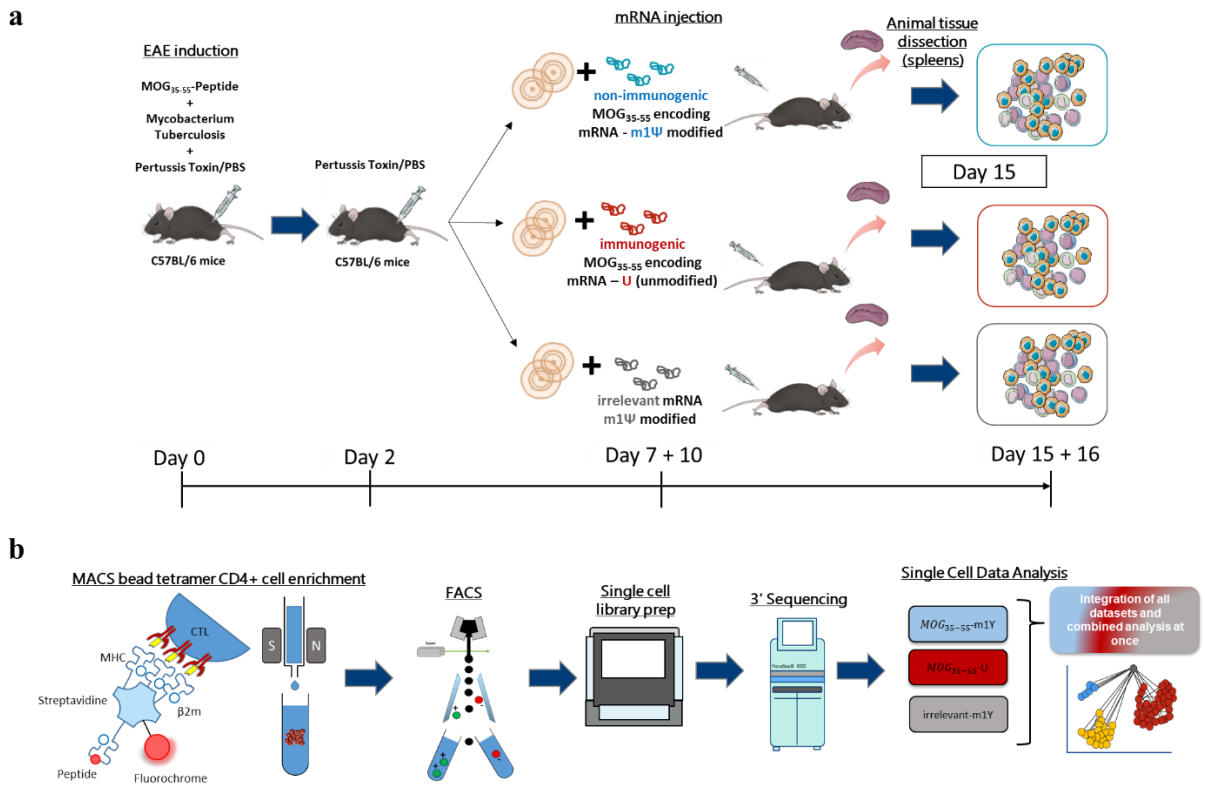


Figure 24: Experimental workflow overview – Processing antigen-specific CD4+ T cells. (**a**) Mice samples (splenocytes) with different treatments were loaded on the single cell chip. All three treatment groups were loaded as duplicates. Sample 1 and 2 (MOG35_m1Ψ Rep1 and Rep2) were processed one day earlier. (**b**) All cells have been tetramer stained, MACS enriched, sorted for MOG35_55 peptide, filtered and processed in the single cell sequencing pipeline. For data analysis all three datasets were integrated.

Analysis of cDNA and final libraries QC parameters of murine splenocyte samples after barcode labeling, reverse transcriptase and cDNA amplification and library construction was done using Qubit and Bioanalyzer (Figure 25 and Table 11). Most of the samples showed a peak fragment profile at ~1800 bp (Figure 25a). For some cDNA samples primer signals could be observed at ~60bp but all samples were further processed and primer signals should be cleaned as a result of further bead purification steps during the rest of the protocol. cDNA concentrations between duplicates were quite equal but differed slightly between the treatment groups itself. Except for MOG35_55_m1Ψ_rep1, concentrations between replicates were quite similar (Table 11). Figure 25b illustrates the fragment profile of all final libraries. Sample 4

(MOG35_55_U_Rep2) showed a very thin peak profile, which could be due to a different bead cleanup performance but it will not negatively affect a sequencing run. Sample 1 and 2 (MOG35_55_m1Ψ_Rep1 and Rep2) are lower concentrated, which might be due to the sample origin itself. All final libraries have an average fragment size between 429-449 bp with no influencing primer signals and could be sequenced.

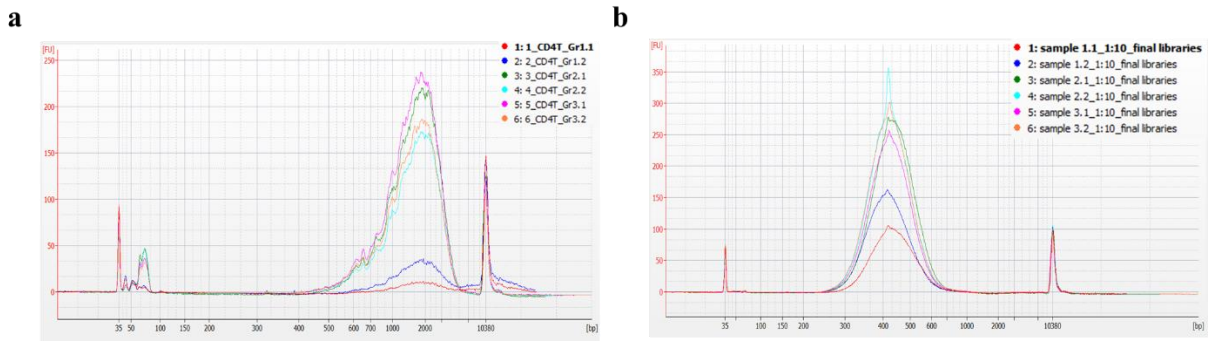


Figure 25: Electropherogram of cDNA and final libraries of murine antigen-specific CD4+ T cell samples (measured by Bioanalyzer). The fragment length is shown on the x-axis [bp] whereas the amount of the fluorescence signals is illustrated on the y-axis [FU]. Both standards can be seen at 35 bp and 10380 bp. The signals of the mouse (**a**) amplified cDNA- and (**b**) final library samples are illustrated in different colors (MOG35_m1Ψ_Rep1: red; MOG35_m1Ψ_Rep2: dark blue; MOG35_U_Rep1: dark green; MOG35_U_Rep2: light blue; irrelevant_m1Ψ_Rep1: pink; irrelevant_m1Ψ_Rep2: orange).

In general, final libraries of MOG35_55_m1Ψ treatment group were a bit lower concentrated compared to the other 4 samples. One reason could be that replicates of MOG35_55_m1Ψ treatment group received two less amplification cycles. All final libraries could be sequenced.

Table 11: DNA concentration of antigen-specific CD4+ murine samples of cDNA samples and final libraries. Concentrations in ng/µl were measured by Bioanalyzer and Qubit.

		Qubit	Bioanalyzer	
Sample type	**Sample name**	**Concentration (ng/µl)**	**Fragment size (bp)**	**Concentration (ng/µl)**
cDNA	MOG35_55_m1Ψ_Rep1	0.30	2510	0.06
	MOG35_55_m1Ψ_Rep2	0.62	2225	0.18
	MOG35_55_U_Rep1	1,59	1824	1.62
	MOG35_55_U_Rep2	1,58	1791	1.52
	irrelevant_m1Ψ_Rep1	1,72	1758	2.00
	irrelevant_m1Ψ_Rep2	1,59	1749	1.83
Final library	MOG35_55_m1Ψ_Rep1	17,6	449	18.16
	MOG35_55_m1Ψ_Rep2	35,8	429	26.16
	MOG35_55_U_Rep1	51	446	41.70
	MOG35_55_U_Rep2	48	433	41.09
	irrelevant_m1Ψ_Rep1	51,4	434	39.72
	irrelevant_m1Ψ_Rep2	43,6	433	47.21

All six final libraries were sequenced using the Illumina Hiseq4000 device and a HiSeq3000/4000 PE flowcell. Generated clusters were checked with the Illumina software "Sequencing Analysis Viewer". More than 71% of the clusters and more than 344 million reads passed filtering. For read 3 more than 77 % of all bases had a quality score of Q30 (1 error in 1000 bases) or more. The error rate in read 3 was less than the expected 1% added PhiX control (with zero to 0.51%).

Table 12: Raw sequencing results of antigen-specific CD4+ murine samples. QC results and statistics of HiSeq4000 run are given as means and were summarized by Sequencing Analysis Viewer. In detail: Cluster PF, Reads PF, Quality Score of 30% or higher and Error Rate.

Read	Clusters PF (%)	Reads PF (M)	% >= Q30	Error Rate (%)
1	71.32	344.48	96.77	0.12
2	71.32	344.48	77.64	0,00
3	71.32	344.48	79.68	0.51

Following the sequencing, Cell Ranger software was performed for demultiplexing. Approximately 1,185 – 1,690 cells could be recovered (estimated number of cells). The level of median genes per cell was between 405 – 1,109 (Table 13). Two samples which were lower

in median genes per cell level were MOG35_55_m1Ψ_Rep1 and MOG35_55_U_Rep2. Sample MOG35_55_m1Ψ_Rep1 was almost two times less concentrated than its duplicate MOG35_55_m1Ψ_Rep2, so a sequencing issue can be excluded and reason for lower median gene per cell level might be probably due to different sample concentration and maybe a lower library concentration in general since PhiX percentage was also lower than 1% for all three reads (Table 12).

Sequencing saturation was good with >96% for all samples, and more than 63% of the reads mapped against the transcriptome. Output data of technical replicated for each group were aggregated via "cell ranger aggr" function so that three datasets with one sample could be used for downstream analysis with Seurat.

Table 13: Summary of sequencing results of antigen-specific CD4+ murine samples. Following results after demultiplexing for each sample are shown. In detail: Estimated number of sequenced cells, mean reads per cell, median genes per cell, total number of detected genes, median UMI counts per cell, sequencing saturation [%], reads which mapped confidently to transcriptome [%] and number of amplification samples.

Sample	Input number of cells	Estimated number of cells	Mean reads per cell	Median gene per cell	Median UMI counts per cell	Seq saturation (in %)	Reads mapped confidently to transcriptome (in %)	Recovery rate (in %)
MOG35_55_m1Y_dup1	3,018	1,690	266,187	405	585	99.0	66.0	56.0
MOG35_55_m1Y_dup2	2,668	1,494	299,158	1,058	1,965	98.5	68.8	56.0
MOG35_55_U_dup1	2,116	1,185	149,535	1,109	2,131	96.7	63.6	56.0
MOG35_55_U_dup2	2,305	1,291	166,364	955	1,737	97.2	63.2	56.0
irrelevant_m1Y_dup1	2,320	1,299	163,552	1,024	1,907	96.7	65.6	56.0
irrelevant_m1Y_dup1	2,338	1,309	136,148	1,039	1,978	96.3	66.7	56.0

Seurat analysis was done starting with raw data processing and QC steps (described in in chapter 6.5.1). Technical replicates were aggregated from the beginning and integration of all three antigen-specific CD4+ datasets were done for better comparison. Dimensional reduction was applied and cells were visualized using UMAP (Figure 26). CD8+ T cells (expression of *Cd8a* gene) and DC's (marker expression not shown here) were identified as a contamination. CD8+ T cells could be isolated together and DCs with tetramer stained CD4+ T cells during the FACS sorting process, since DCs, CD4+ T cells and CD8+ T cells interact with each other [113]. Cluster 10 consisting of CD8+ T cells and DCs and was not be taken into account for downstream analysis and therefor excluded as a contamination since only antigen-specific CD4+ T cells are the focus of this study. Furthermore, a low quality cell cluster (cells with low

amount of RNA and high amount of mitochondrial genes)[114] was found and excluded from downstream analysis as well so that the analysis is not affected by these cells. Different markers were used to examine the presence for general T cell phenotypes like exhaustion (*Cd160* and *Pdcd1*) or proliferation (*Top2a* and *Mki67*), as well as for specific CD4+ T cell subsets like Effector Treg phenotype (*Foxp3*, *Ikzf2* and *Il10*), Th 1 phenotype (*Tbx21*, *Ifng* and *Csf2*) and Th17 phenotype (*Il17a*, *Il7f* and *Rorc*) were observed. The determination of especially these CD4+ T cell subsets confirms the strength of scRNA-seq to identify subgroups of major cell types and can give deeper insights into the treatment effect of MOG335_55_m1Ψ, MOG35_55_U and irrelevant_m1Ψ mRNA.

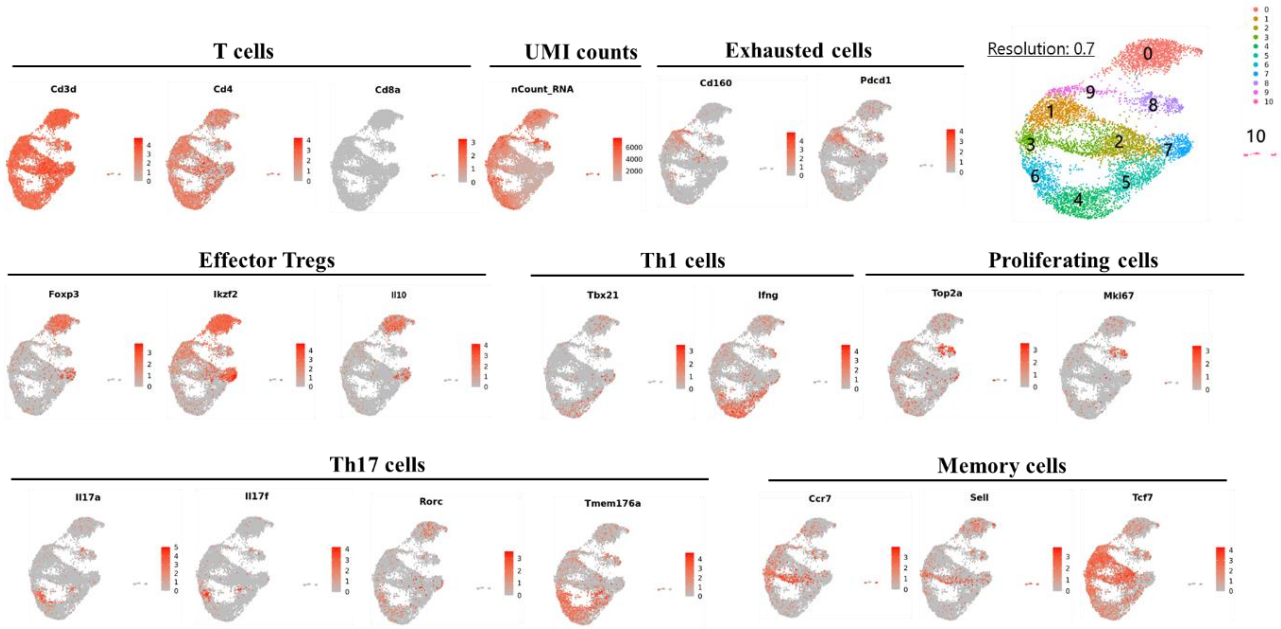


Figure 26: Identification of cell subsets in murine antigen-specific CD4+ T cells. Selected marker expression visualized in the UMAP as Featureplot. A gene which is higher expressed in a cell is colored in darker red. Various markers were plotted for the detection of T cells, Exhausted cells, Effector Treg cells, Th1 cells, Th17 cells and Proliferating cells as well as an overview of all identified cell clusters by Seurat [dimension: 16; resolution: 0.7].

In Figure 27 an intermediate UMAP plot is shown which illustrates all 11 identified clusters with a preliminary alignment as well as their treatment information. Beside effector Treg cells, exhausted T cell, proliferating cells, memory T cells and Th17 cells, a mix of Th1/Th17 cells could be identified in cluster 4 and 5. An exact alignment of Th17 and Th1 or Th1/Th17 mixed cell cluster was not directly possible since Th1 and Th17 marker expression was widely spread in cluster 4-6. The reason for this might be that the algorithm was influenced by other CD4+ T cell subsets and similar gene expression profiles for both effector subsets. Therefore, a more

separated Th1 and Th17 specific clustering was not possible. For a more stringent clustering and Th1 and Th17 cell identification a "subclustering" was done (method section 6.5.4). Briefly, only cluster 4-6 were taken and a re-analysis with only these three clusters was done so that the algorithm can cluster Th1 and Th17 cell subsets better without other influencing CD4+ T cell subset signals (Figure 28).

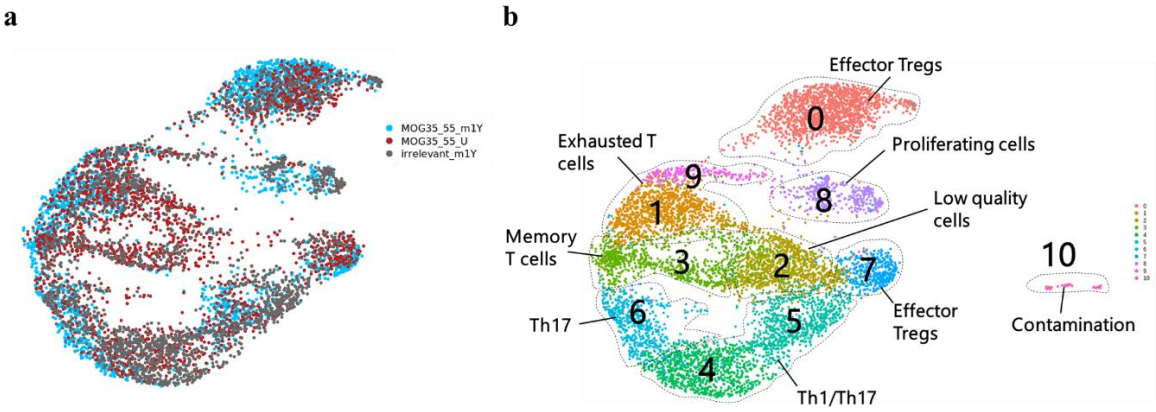


Figure 27: Intermediate UMAP with treatment information and identified murine cell types antigen-specific CD4+ T cells. (**a**) UMAP illustrating treatment information and (**b**) intermediate UMAP plot shows all present cells (not subclustered).

After subclustering cells were separated in 12 different clusters and signals of Th1 and Th17 phenotype (Figure 28d and Figure 28e) could be identified in a more separated form. Cluster 3 and 7 showed expression of more Th17 markers, whereas cluster 4 was assigned solely as Th1 cell subset. The remaining clusters 0, 1, 2 and 5-11 were assigned as mixed Th1/Th17 cell population. This result confirms also a separate Th1 cell subset which could not be identified before with all CD4+ T cell types and confirming subclustering as powerful strategy to obtain higher granularity for subsets of interest.

The UMAP and barplot below (Figure 28a and b) illustrates the final assignment of all antigen-specific CD4+ T cells subsets and the determined cell type frequencies among all cell types and three treatment groups. The expected increase of effector Treg cells and decrease of Th1 cells and Th1/Th17 mix upon MOG35_55_m1Ψ treatment proves the tolerance induction and reduction of effector T cells (Figure 28b). The disease effect could be observed with high effector T cell frequencies (Th1 and Th1/Th17) after irrelevant_m1Ψ treatment which led to sickness in the mice group since they are the driving force of this disease [115]. The frequencies of Th17 cells were quite equal for all three treatments, but since it was the smallest population

of all present cell types, striking differences might not be expected to be identified. Therefore, an enrichment of these cells would be necessary for deeper investigations. Exhausted T cell frequencies were similar between MOG35_55_m1Ψ and MOG35_55_U treatment with almost 20% (in comparison to ~14% after irrelevant_m1Ψ treatment) and illustrates a potential result of these (un)modified mRNA constructs, leading to an increased exhaustion. Higher memory T cell frequency after MOG35_55_U might be explained by the fact that MOG35_55_U treatment is more immunogenic than MOG35_55_m1Ψ. Therefore, MOG35_55_m1Ψ could prevent less the formation of autoreactive memory T cells. That is why the frequency of memory T cells upon MOG35_55_m1Ψ treatment is lower after treatment compared to control (irrelevant_m1Ψ mRNA) [116]. Proliferating cell frequencies were lowest after MOG35_55_U and higher upon MOG35_55_m1Ψ and irrelevant_m1Ψ treatment, which could be a result of present 1-methylpseudouridine. For MOG35_55_m1Ψ a 1-methylpseudouridine-5'-triphosphate, m1ΨTP was used instead of uridine-5'-triphosphate, UTP [52], which might be the reason that MOG35_55_U mRNA treatment might not result in such a high effector Treg cell frequency and low effector T cell frequency as for MOG35_55_m1Ψ mRNA construct. The results could confirm that MOG35_55_U mRNA is more immunogenic leading to higher co-stimulation, more disease-driving response.

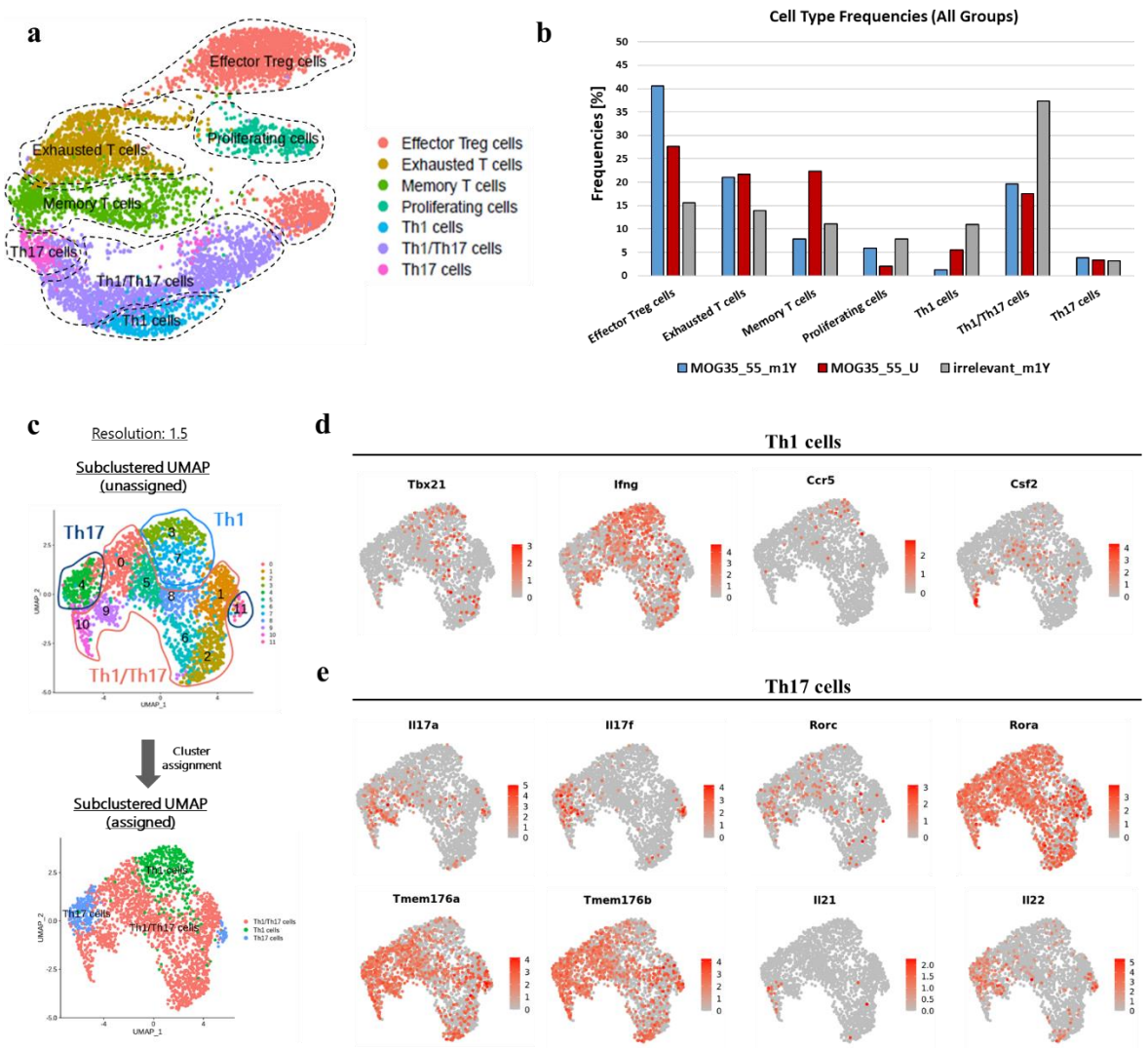


Figure 28: Final cell type assignment of murine antigen-specific CD4+ T cells and the respective cell type frequencies among treatment groups. (**a**) Final UMAP plot shows all identified cells. This UMAP contains the cell identities (their barcodes) of previous subclustered Th17/Th17 cells which have been transferred to the entire antigen-specific CD4+ dataset. (**b**) This boxplot illustrates the determined frequencies with all present cell subsets on the x-axis and each treatment group in different colors (MOG35_55_m1Ψ: blue; MOG35_55_U: red and irrelevant_m1Ψ: grey). The y-axis features cell type frequencies in %. (**c**) Subclustering procedure with raw clusters and assigned Th1 and Th17 cell populations. Featureplots with (**d**) Th1 gene expression and (**e**) Th17 gene expression.

In addition to cell type frequency assessment, differential gene expression analysis was done to demonstrate and confirm a treatment effect of MOG35_55_m1Ψ mRNA application in EAE sick mice also on transcriptional level. Here cell types were grouped together to evaluate a general difference between the three treatment groups. More potential differential expressed genes between the other two mRNA constructs were investigated using Seurat. Like in method part explained (6.5.3), the top 20 significantly expressed genes (ranked by adjusted p-value <

0.05) were selected and shown in Figure 29. Comparison between all three groups was done groupwise as followed: MOG35_55_m1Ψ against irrelevant_m1Ψ (Figure 29a), MOG35_55_m1Ψ against MOG35_55_U (Figure 29b) and MOG35_55_U against irrelevant_m1Ψ (Figure 29c).

Comparison between MOG35_55_m1Ψ and irrelevant_m1Ψ (Figure 29a) revealed the expected Treg cell activity (*Ccl5* and *Tnfrsf1b*), anti-inflammatory cytokine effect (*Il10*), expression of effector molecule (*Gzmb*) after MOG35_55_m1Ψ treatment as well as down regulation of proliferation processes (*Malat1*, *Wdr26*, *Abtb1* and *Prkca*). This confirms the expected tolerance induced impact of MOG35_55_m1Ψ in EAE sick mice [105]. Using irrelevant_m1Ψ, cells seemed to maintain their survival by the expression of *Ms4a4b*, inducing cytotoxicity (*Pfn1, Klrc1*) as well as apoptosis (*Lgals1*). The gene Bhlhe40 was reported to induce autoimmune disease, which would reflect also the disease state of EAE sick mice. *Cd69* expression demonstrate a regulation of Treg cell differentiation to prevent a potential tolerance induction against the disease.

A consequence of MOG35_55_m1Ψ treatment compared to MOG35_55_U (Figure 29b) is a downregulation of proliferation (*Wdr26*, *Abtb1*, *Madd*) and an increasement of Treg cell function plus Treg cell development and cytokine effect (*Ccl5*, *Tnfrsf1b*, *Mid1*, *Gzmb*, *Akna* and *Ikzf2*). Antigen-specific CD4+ T cells treated with MOG35_55_U expressed genes involved in the (down)regulation of cell cycle (*Btg1, Ddx5*) as well as Treg differentiation (*Cd69*) or apoptosis (*Ms4a4b*). They show cytotoxic signature but to a lesser extent than murine cells after MOG35_55_m1Ψ. A lot of ribosomal and mitochondrial genes could be found under the top 20 differential expressed genes, which might be a sign of stressed or less viable cells [117].

The comparison between MOG35_55_U and irrelevant_m1Ψ (Figure 29c), illustrated again suppression of proliferation, growth and apoptosis (*Malat1*, *Ypel3* and *Prkca*) upon MOG35_55_U treatment. Cells further showed memory formation (*Cd27*), which is expected after MOG35_55_U application [116]. EAE sick mice treatment with irrlevant_m1Ψ instead of MOG35_55_U led to induction of apoptosis (*Lgals1*, *Ifitm2*) and autoimmune disease (*Bhlhe40*). Cells express effector molecules like *Klrc1* and *Nkg7* and also show upregulation of

activation compared to MOG35_55_U. A deeper investigation on affected pathways between the treatment groups was done as well and results are shown in Figure 30.

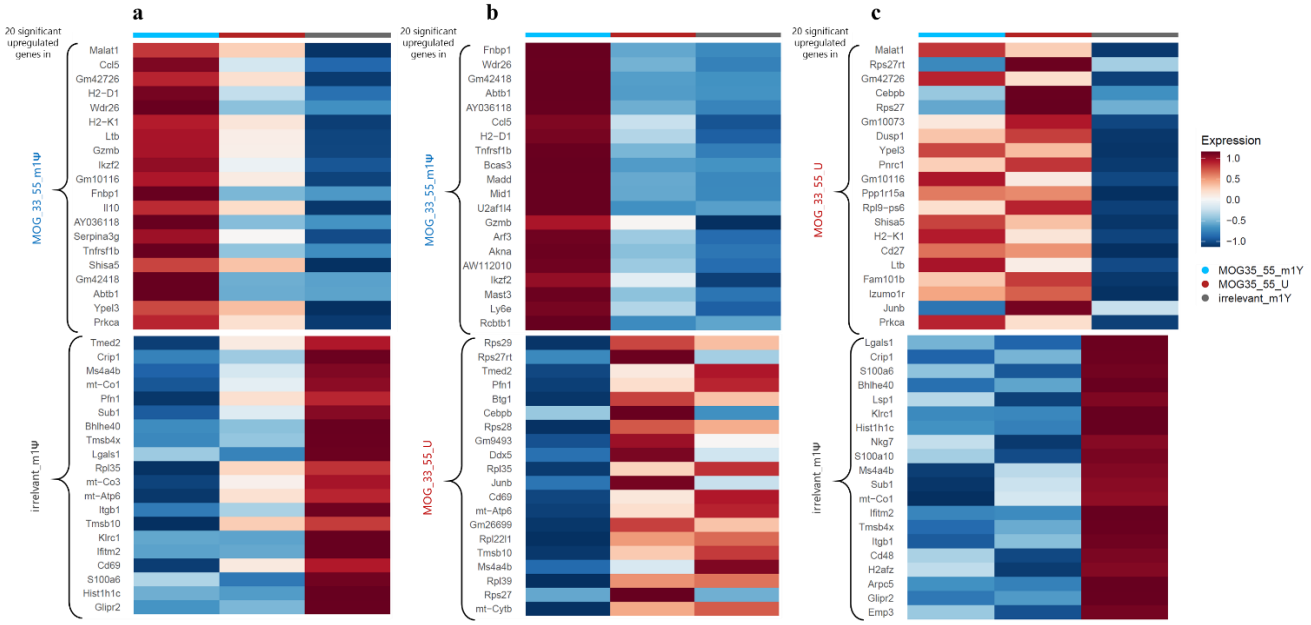


Figure 29: Differential gene expression analysis results between all three mRNA treatments for murine antigen-specific CD4+ T cells. The heatmap displays the expression of 20 top significant expressed genes for combined cell types revealed after the comparison between the treatment groups: (**a**) MOG35_55_m1Ψ against irrelevant_m1Ψ, (**b**) MOG35_55_m1Ψ against MOG35_55_U and (**c**) MOG35_55_U against irrelevant_m1Ψ. For combined cells, all three treatment groups are plotted on the x-axis (MOG35_55_m1Ψ: light blue; MOG35_55_U: red; irrelevant_m1Ψ: grey), while the differential expressed genes are plotted on the y-axis (based on their adjusted p-value). Genes, which are higher expressed, are colored more (dark) red, genes which are downregulated are colored more (dark) blue.

To answer the overall scientific question if there are more potential differences especially between MOG35_55_m1Ψ and MOG35_55_U mRNA treatment, a pathway analysis was done using the algorithm clusterProfiler 4.0, a universal interface for functional enrichment analysis in thousands of organisms [118]. The tool uses the latest annotation databases and offers different analysis methods like “ORA” (over-representation-analysis) to determine whether genes from pre-defined sets (belonging to a specific GO term or KEGG pathway) are present more than would be expected (over-represented) in a subset of the data. This can be an experimentally-derived gene list e.g. a list of differentially expressed genes (DEGs) of this dataset (not shown here).

For MOG35_55_m1Ψ treatment in Effector Treg cells, clusterProfiler4.0 identified enrichment of pathways involved in the regulation of cell proliferation and immune effector process (Figure

30). Previous differential gene expression analysis confirmed that some differential expressed genes like *Malat1*, *Abtb1*, *Madd* and *Wdr26* could be identified as genes which have negative regulation functions on the cell proliferation after applying MOG35_55_m1Ψ. Beside regulation of proliferation, defense response was positive regulated (affected genes: *Ccl5*, *H2-T23*, *Tnfrsf1b*) as well as cytokine production involved in immune response (affected genes: *Ccl5*, *H2-T23* and *Stat1*) which confirms the expected effect on MOG35_55_m1Ψ. Differential gene expression analysis between MOG35_55_U and MOG35_55_m1Ψ treatments could not really confirm higher immunogenicity or higher pathogenic effector response for MOG35_55_U (supplementary Figure 60). As expected, further evaluations of pathway analysis showed no further differences between both MOG35_55_m1Ψ and MOG35_55_U treatments and between MOG35_55_U and irrelevant_m1Ψ in the different cell types (supplementary Figure 49 to Figure 55). In effector Treg cells only cell killing (enriched by *Gzmb* gene) was found as one enriched pathway for MOG35_55_U against irrelevant_m1Ψ (supplementary Figure 49). This may be due to the similarity of the mRNA constructs MOG35_55_m1Ψ and MOG35_55_U as the former was only slightly modified with a 1-methylpseudouridine-5'-triphosphate.

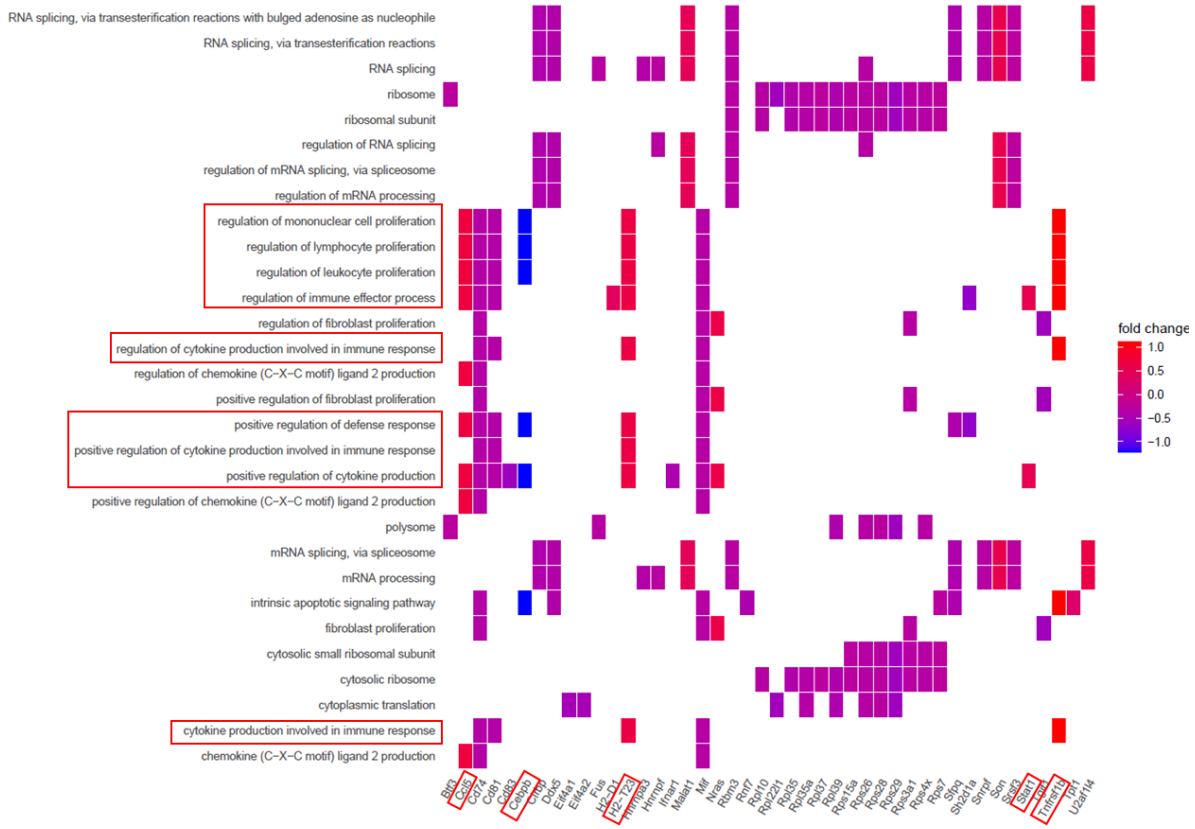


Figure 30: Pathway analysis results for effector Treg cells (comparison: MOG35_55_m1Ψ against MOG35_55_U) of murine antigen-specific CD4+ T cells. The heatmap displays expression patterns for effector Treg cells and features of the respected comparison: All genes which are involved in these significant terms are depicted on the x-axis, while the enriched pathways are plotted on the y-axis (by clusterProfiler4.0 algorithm). Genes, which were found to be involved in an enriched pathway, are colored as follow: More pink means a higher fold change (upregulated), whereas more blue means a lower fold change (downregulated gene). Selected pathways linked to proliferation, defense/immune response or cytokine production are highlighted in red frames.

3.3 Discussion

The established scRNA-seq pipeline was used to investigate three different mRNA-LPX treatment groups in EAE sick mice and their effect on CD4+ T cell populations. Whereas in the previously published manuscript by Krienke et al. the effect of MOG35-55_m1Ψ RNA-LPX compared to irrelevant group was studied and tolerance induction in the EAE murine model could be confirmed, we went beyond and supported Krienke et al. studies by using our scRNA-seq pipeline to look at the effect of an additional treatment group, immunogenic autoantigen-encoding RNA. Here, we assessed differences and similarities of their effects on splenic CD4+ T cells on gene expression levels at single cell resolution. My own investigation for this thesis chapter was further to dissect the effect of MOG35-55_U in comparison to MOG35-55_m1Ψ

RNA-LPX and irrelevant m1Ψ RNA-LPX and to study functional differences with respect to inflammatory and immunosuppressive properties of the present CD4+ T cell subpopulations. The established scRNA-seq pipeline (from chapter one) was used for deeper characterization of involved immune cells and to study enriched pathways and their relationships among the different treatments. As already shown by Krienke et al. a successful tolerance induction in EAE sick mice upon MOG35_55_m1Ψ treatment compared to irrelevant_m1Ψ treatment could be observed. This was accounted by high effector Treg cell frequencies and low effector T cell frequencies upon MOG35_55_m1Ψ treatment. Results further revealed a milder immunosuppressive and effector response like not such a high enrichment of pathogenic effector cells in antigen-specific CD4+ T cells after using MOG35_55_U mRNA compared to MOG35_55_m1Ψ group. This was observed by an intermediate frequency level of almost all cells (except for memory T cells) compared to MOG35_55_m1Ψ group and irrelevant group. A treatment effect of MOG35_55_m1Ψ, in addition to CD4+ T cell frequency changes, was also demonstrated at the transcriptional level in the form of an upregulation of *Ccl5* and *Tnfrsf1b* (Treg cell activity), *Il10* (anti-inflammatory cytokine action), *Gzmb* (effector molecule expression) as well as down regulation of *Malat1* or *Prkca* (proliferation processes). *Wdr26*, known to be involved in cell migration [119], was down regulated and could potentially prevent cell trafficking to site of inflammation leading to a reduction of EAE disease level.

My investigations and comparisons between MOG35_55_U and MOG35_55_m1Ψ treatment group as well as irrelevant m1Ψ RNA-LPX treatment group demonstrated that on the one side MOG35_55_U compared to MOG35_55_m1Ψ led to downregulation of cell cycle genes (*Btg1*, *Ddx5*), but on the other side it resulted in an upregulation of apoptosis marker *Ms4a4b*, being counterproductive in the treatment of the disease. However, this observation has to be taken with care since it was the expression of one marker. It could be also related to negative regulation of mouse T cell proliferation [120]. MOG35_55_U further resulted into less cytotoxic signature (less expression of effector molecules) compared to MOG35_55_m1Ψ but therefore higher level of stress was observed (more expression of ribosomal and mitochondrial genes) [117]. MOG35_55_U suppresses cell proliferation, growth and apoptosis by the expression of *Malat1*, *Ypel3* and *Prkca* similar to MOG35_55_m1Ψ but only latter showed enrichment of pathways connected to a tolerance induction like regulation of cell proliferation, immune effector process, positive regulation of defense response or cytokine production

involved in immune response. MOG35_55_U as a more immunogenic mRNA construct did not lead to an enrichment of pathways demonstrating a (stronger) tolerance induction in EAE sick mice. In line with the findings for two of the three conditions in Krienke et al. (2021) [105], the impact of immunogenic mRNA (MOG35_55_U) does not show great enhancements of the treatment before (prophylactic) or after EAE disease induction compared to non-immunogenic mRNA (MOG35_55_m1Ψ). This could be explained by the low effector Treg expansion, an unstable effector Treg cell phenotype during inflammatory conditions as well as a strong Teff cell expansion [94], which I could also observed in the frequency levels (Figure 28b). On the other hand, Th1/Th17 phenotype seems to be unstable and rather tend to convert to effector Treg cells in m1Ψ as there is regulation of regulatory T cell differentiation up in Th1/Th17 m1Ψ against U (suppl. Figure 50). With regards to the question what are potential effects of MOG35_55_U on the pathogenic cells, the application and comparison of MOG35_55_U with irrelevant_m1Ψ showed some immune activity like slightly higher frequency of effector Treg cells and memory T cells, suppression of proliferation and apoptosis but therefore did not benefit over the application of MOG35_55_m1Ψ. Taking this into account my analysis confirmed the importance of noninflammatory autoantigen-encoding constructs which are required to elicit the tolerogenic effect since MOG35_55_m1Ψ as more non-immunogenic mRNA construct seems to be more promising as a technique to treat MS patients in the future [105]. The non-immunogenic mRNA construct MOG35_55_m1Ψ delivered autoimmune antigens to APCs without an inflammatory response resulting in a tolerance to the delivered antigen. Further unbiased investigation and comparison of co-stimulatory interactions between DCs and T cell compartment in MOG35_55_m1Ψ and MOG35_55_U condition using tools like CellChat would be highly interesting. This could be used to confirm expected co-stimulatory interactions but also to discover new interactions upon this novel RNA treatment strategies. Furthermore, if tolerance was successfully induced, autoreactive T cells were suppressed (bystander tolerance) resulting in a prevention of symptomatic disease (paralysis), reduction of early-stage disease progression as well as restoring of motoric functions [105]. Alternative treatments and future perspectives for MS treatment, include peptide or protein delivery in anti-inflammatory formats administered in oral, nasal or subcutaneous way, or apply the delivery on carriers such as cells or major histocompatibility complex molecules. Though not all of them has led to a striking successful treatment so far [121]. According to the results, it seems that the approach of the RNA-LPX antigen delivery system approach has many

advantages over alternative treatment techniques (in terms of production, cost and adaptation to other protein antigens) [105]. As previous studies had shown the role of CD4+ T cells in the pathogenesis of MS for driving the disease, this exploration mainly focused on the study of antigen-specific CD4+ T cells and their role in MS disease upon RNA-LPX treatment. Along with CD4+ effector Th1 and Th17 cells, CD8+ T cells can also lyse myelin-specific CD4+ T cells and might have an impact in the pathogenic of MS lesions as effector cells [122]. Yet Li et al. (2022) showed evidence that a subset of human CD8+ T cells (expressing the inhibitory killer cell immunoglobulin-like receptors KIR3DL1 and KIR2DL3) represents only 0-3% of blood CD8+ T cells in healthy individuals but show an increased amount in patients with autoimmune diseases. CD8+ T cells expressing KIRs can regulate NK cells killing activity, an important factor for targeting CD4+ effector T cells and direct killing of autoreactive or pathogenic CD4+ T cells [123]. These CD8+ KIR+ T cells would be worth to be investigated further (in addition to CD4+ T cells investigated here) to explore if and how they could potentially support the tolerance induction in EAE sick mice or MS patients. More studies on CD8+ effector T cells and regulatory CD8+ T cells are necessary in order to develop alternative strategies to suppress self-reactivity in MS.

In order not to leave potentially important cell populations unexamined further investigations on other immune cell subsets like B cells are relevant as since B cells could prospectively contribute to CNS pathology [124] by secreting factors damaging CNS or promoting inflammatory polarization of microglia [97]. As B cells can pass the Blood-Brain-Barrier (BBB) by support of CD4+ Th17 cells [125]. Contrary to this, B cells play a critical role in maintaining peripheral tolerance and suppressing the development of autoimmune diseases [95]. Yet regulatory B cells can suppress pro-inflammatory responses through the expression of *IL-10*, *TGF-β*, and *IL-35* [126], support Treg cell expansion, inhibit effector T cell expansion and regulate CNS inflammation [127]. Therefore, regulatory B cells with a restored production of IL-10 and an anti-inflammatory phenotype [97] play an important role for the suppression of MS disease and are worth to be further investigated for the development of more efficient MS treatment strategies. As NK cells are valuable protectors of the innate immune system, their role in the suppression of autoimmune disease like MS is to be highlighted. They express a number of important receptors which are essential for different mechanisms like natural cytotoxicity (*NKp30*, *NKp44*, and *NKp46*). They are also responsible for cell activation

(*CD94/NKG2*) and inhibition (*KIR* receptors, killer Ig-like receptors). Hence, NK cells show a great ability to suppress proliferation of autologous T cells and should be investigated deeper for potential detection of other immunoregulatory capabilities [95,128]. As to medication, the FDA (Food and Drug Administration) approved in March 2020 the oral capsules Zeposia (Ozanimod) for the treatment of adults (but not for children and adolescents younger than 18) with relapsing forms of MS. Zeposia bind to sphingosine 1-phosphate (S1P) protein receptors on immune cells to catch and prevent inflammatory effector cells to enter brain and spinal cord. One drawback of Zeposia is that it may increase the risk of infections, such as serious cases of shingles (herpes zoster) since it prevents immune cells from entering the blood [129].

More research needs to be done for finding solutions to treating MS patients with different self-antigens. Using scRNA-seq technology provided deeper insights into the cell populations and their responses after MOG35_55_m1Ψ treatment. Nucleoside-modified and purified mRNA are well defined clinical stage compounds in pharmaceutical context and are tested in clinical human trials for potential new solutions in the treatment of autoimmune diseases like MS [130].

4 Chapter Four: Combined IL-2 variant and anti-PD 1 treatment induces antitumoral changes in the immune landscape in a mouse tumor model

4.1 Introduction

4.1.1 Ribocytokines

Boosting the immune system to destroy cancer cells and to ensure permanent antitumor immune response is the major aim of immunotherapy. Although there have been promising breakthroughs for different immunotherapies, unpredictable efficacy, tumor heterogeneity and acquired treatment resistance still challenge researchers in the development of efficient immunotherapies and cancer vaccines [131]. Cancer vaccines induce tumor-specific T cell or B cell responses, which might lead to tumor rejection by strong immunomodulation. In fact, antigen T cell and NK cell receptors have shown to be of special use as recognition for immunologically target and active proteins in a variety of immune and non-immune cells [50].

Inhibition of Programmed Cell Death Protein 1 (PD-1) is one approved treatment across various cancer indications. It blocks the interaction between programmed cell death-Ligand 1 (PD-L1) (Figure 31a and b) and further prevents PD-1 pathway–mediated immunosuppressive responses and anti-tumor immune responses [132]. However, only a small fraction of patients develop durable anti-tumor responses and many cancer patients are resistant (or partially susceptible) to PD-1/PD-L1 blockade due of low T cell infiltration [133,134].

Besides ICI (immune checkpoint inhibitor) like PD-1 or Cytotoxic T-Lymphocyte-Associated Protein 4 (CTLA4) blockade, there are further different kinds of Food and Drug Administration (FDA)-approved immunotherapies like adoptive cellular therapy (ACT) [135], oncolytic virus (OV) therapy [136] or cytokine therapy [137].

Cancer vaccines with promising results designed to treat infections have moved into late-stage clinical trials [138]. Despite some examples of vaccines that induced systemic regression of large tumors [139,140] and prolonged survival [141], small clinical trial sizes, marginal survival benefits and resource-intense approaches have held the field back from greater success. Cancer vaccines in general shall help recognizing and killing tumor cells, but sometimes they fail when the immune system is not fully activated or when tumor cells become resistant by the upregulation of a negative immune regulator like PD-1 which further decrease immunogenicity.

mRNA-encoded antibodies or cytokines called (Ribocytokines) have also been developed for LNPs. Ribocytokines can be used as a platform for the delivery of mRNA encoding bi-specific antibodies or cytokines [142]. In comparison to cytokines, ribocytokines can target CD3 and TAAs and therefore bind T cells to cancer cells, leading to improved T cell-mediated antitumor activity [143]. Ribocytokines aim to activate the immune system, recruiting various immune cells or targeting tumor cells. One of these cytokines is Interleukin-12 (IL-12) [144]. IL12 is considered to have good anti-tumor activity and activates natural killer cells (similar to IL10 as an anti-inflammatory cytokine, mentioned in discussion chapter three (3.3)), T lymphocytes, and macrophages, promotes DCs maturation, and inhibits vascular formation. IL12 also induces the production of interferon-γ (IFN-γ) and other immune factors, thus exerting strong anti-tumor effects [145,146]. Although IL12 has successfully been used for cancer treatment in mice, it was associated with toxicity when administered systemically in cancer patients [147]. Severe toxic side effects are largely due to excessive systemic IFN-γ expression. Cytokine localized delivery is also strongly focused in order to achieve effective anti-tumor therapy. Therefore, IL12 was delivered as NHS-IL-12 fusion protein (targeting to the necrotic part of metastatic solid tumors) [148], as a recombinant cytokine injected locally in combination with cetuximab (in patients with head and neck tumors) [149], or as adjuvant with a tumor vaccine [150]. Due to the potentially lethal and dose-limiting toxicities, so far IL12 is not FDA approved and tumor-targeted delivery and/or in situ vaccination need to be further explored. Further promising but so far not FDA-approved cytokines like IFNγ, granulocyte-macrophage colony-stimulating factor (GM-CSF), IL15, and IL21, are currently evaluated for their anticancer potential in preclinical and clinical models [131].

Interleukin-2 (IL2) is one promising and FDA-approved immunotherapy treatment, but its efficacy and utility in a clinical setting is limited by its short in vivo half-life, severe toxicity and propensity to amplify immune suppressive T regulatory (Treg) cells in addition to effector natural killer (NK) and T cells [151]. Thus, novel approaches to overcome the limitations of IL2 therapy are required (Figure 31e). In this work, a modified variant of IL2 developed (encoding a mutant variant of human IL-2 fused to human serum albumin) from the colleagues in BioNTech (here just named as IL2var) is used to expand specifically CD8+ T cells and not Treg cells to overcome the Treg cell expansion issue since they can suppress the activation of many immune cell types including autoreactive conventional T cells and can further be intrinsically defective in many common autoimmune diseases [152]. IL2var can further enhance

the ex vivo expansion of T cells [153] and preliminary data showed an increased cell type frequency of macrophages and tumor infiltrating lymphocytes (TILs). Previous experiments (data not shown here) showed a transient and short growth retardation in mice with only IL2var, whereas a combination with IL2var and the antibody anti-PD-L1 led to a tumor rejection. Researchers from BioNTech (BNT) and from single cell genomics team at TRON gGmbH collaborated to investigate the IL2var treatment effect via unbiased immune cell profiling using scRNA-seq. The following two hypotheses were investigated in this chapter: i) Is there an increased lymphocytes frequency (especially CD8+ T cells and CD4+ Treg cells) compared to IL2var alone or/and treatment with IL2var in combination with anti-PD-L1 observable? ii) Can we confirm a stronger effector phenotype among combination treatment (i.e. higher expression of cytotoxic molecules) in CD8+ T cells compared to treatment with IL2var alone and which therefore might result in a complete tumor rejection? We used our scRNA-seq pipeline to detect relevant changes in gene expression and examine cell subset phenotypes among IL2var treatment and combination treatment with anti-PD-L1 at single cell level.

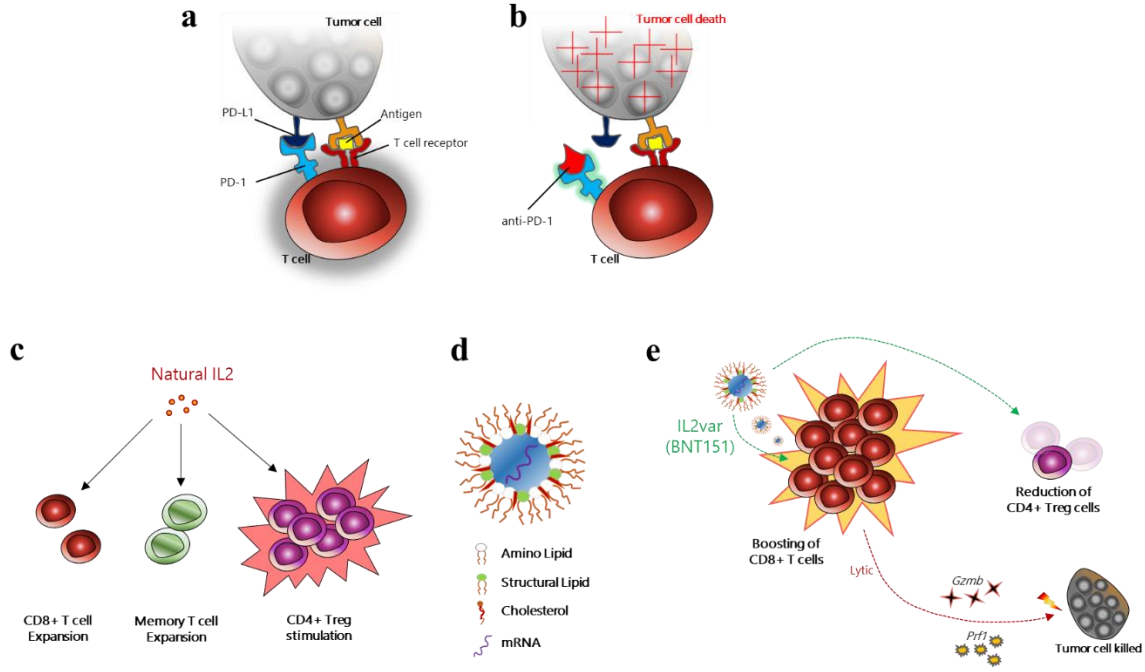


Figure 31: Mechanism of PD-1 inhibitor, IL2 mode of action for modified IL2var. (**a**) Binding of PD-1 and PD-L1 can inhibit the immune response and (**b**) anti-PD-1 antibody can reverse the inhibition [adapted from [154]]. (**c**) IL2 is consumed by activated CD8+ T cells, which can lead to the expression of cytotoxic molecules. It not only further enhances CD8+ effector and memory T cells but also stimulates Treg cells. (**d**) Novel lipid nanoparticle-formulated, nucleoside-modified RNA encoding a mutant variant of human IL-2 (IL2var) fused to human serum albumin (hAlb). (**e**) Modified version of IL2 which is fused with human albumin has the following advantages: expansion of effector T cells (secreting cytotoxic molecules like *Prf1* or *Gzmb* to kill tumor cells), reduced expansion of Treg cells, extended serum half-life and increased bioavailability in tumors (adapted from [155]).

4.1.2 Hashtag oligonucleotide (HTO) application for multiplexing and doublet detection

With the benefit of single cell genomics approaches to reconstruct unsupervised taxonomies of cell types, other challenges remained for quite some time in the past: missing application of massively parallel technologies to detect rare cell types [21] as well as the robust identification of artifactual signals arising from cell multiplets or technology-dependent batch effects [156]. After the development of the CITE-Seq technique (Cellular Indexing of Transcriptomes and Epitopes by Sequencing), the detection of cell surface proteins was integrated into the scRNA-seq readout using oligonucleotide-tagged antibodies. Antibodies can label different experimental samples and thereby discriminate replicates in one experimental setting [157]. A 12bp barcode makes HTO unique (no UMI sequence) and can be sequenced alongside the cellular transcriptome. HTO's and scRNA-seq libraries can be independently amplified and pooled at desired quantities.

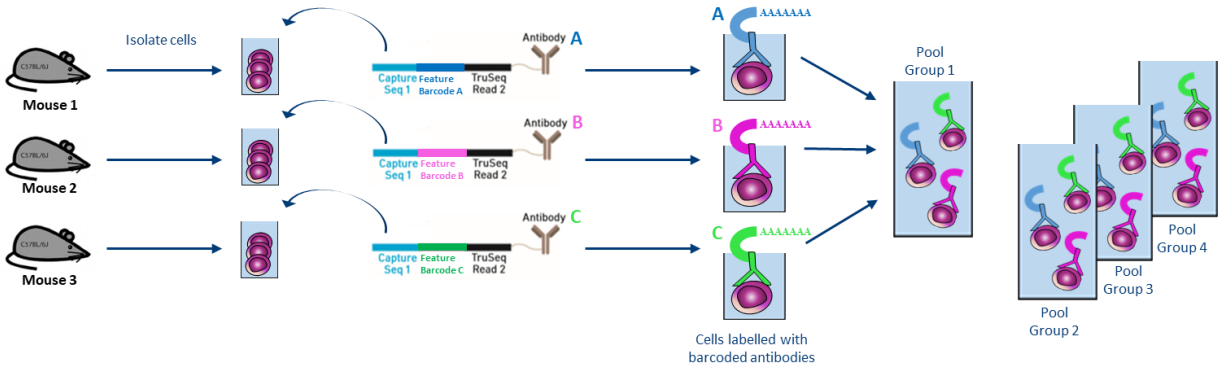


Figure 32: Schematic overview of sample multiplexing by Cell Hashing using barcoded antibodies. Cells from different mice are incubated with DNA-barcoded antibodies (Biolegend) recognizing ubiquitous cell surface proteins. Pooling of multiple samples into one sample pool is possible via distinct barcodes (referred to as hashtag-oligos, HTO) on the antibodies. Following sequencing, cells can be assigned to their sample of origin based on HTO levels (adapted from [158]).

4.1.3 Aim of the experiment

To demonstrate therapeutic potency of IL2var alone or in combination with antibody aPD-1, CD45+ cells in a MC38 tumor mouse model were isolated 3 days after treatment and characterized by scRNA-seq. Following the identification of present murine immune cells, a treatment effect of the mRNA with or without aPD-1 was determined by the difference in cell type percentages as well as gene expression levels for each sample and compared to a control.

In order to examine the gene expression levels of CD45+ cells for differences between treatments regarding T cell phenotype/functionality, various cell state markers were examined, primarily cytotoxic and exhaustion markers. One focus was to assess a potentional increased cell type frequency of macrophages and TILs after IL2var or the combination treatment of aPD-1 and IL2var compared to control. In combination with aPD-1, we expected an expansion of TILs and a stronger effector phenotype in CD8+ T cells (higher expression of cytotoxic markers). In addition, we determined whether the modified IL2var leads to a reduction of CD4+ Treg cells. The following preparations have been done in BioNTech: MC38 tumor inoculation, modified IL2 treatment on MC38 tumor bearing mice, mice observation, anesthesia and sacrifice of mice, mouse tissue preparation, MACS, dilution of barcoded-antibodies, staining of and flow cytometry. All further activities, including preparation of single cell solution samples, loading samples on 10X Genomics device Chromium Controller, library preparation of 3´GE samples and barcoded-antibody samples, preparation for sequencing on NovaSeq and following single cell analysis was performed by me with support of my supervisors.

4.2 Results

To explore efficacy and mode of action of IL2var alone or in combination with aPD-1 in MC38 mouse tumor model, we performed scRNA-seq (Figure 33). Briefly, four groups of C57BL/6 mice were injected with MC38 tumor cell line and received single treatment (1: Control which was rat Ig2a and human Albumin LNP-formulated, 2: aPD-1, 3: IL2var and 4: combination treatment aPD-1 + IL2var). Tumors were dissociated 14 days after the treatment using gentleMACS, stained with feature barcoding antibody to be able to pool individual samples groupwise, isolated via flow cytometry for viable CD45+ cells and processed by our scRNA-seq pipeline using the Chromium Controller and subsequent sequencing and computational analysis. One single sample per treatment group was prepared and loaded on a single cell chip.

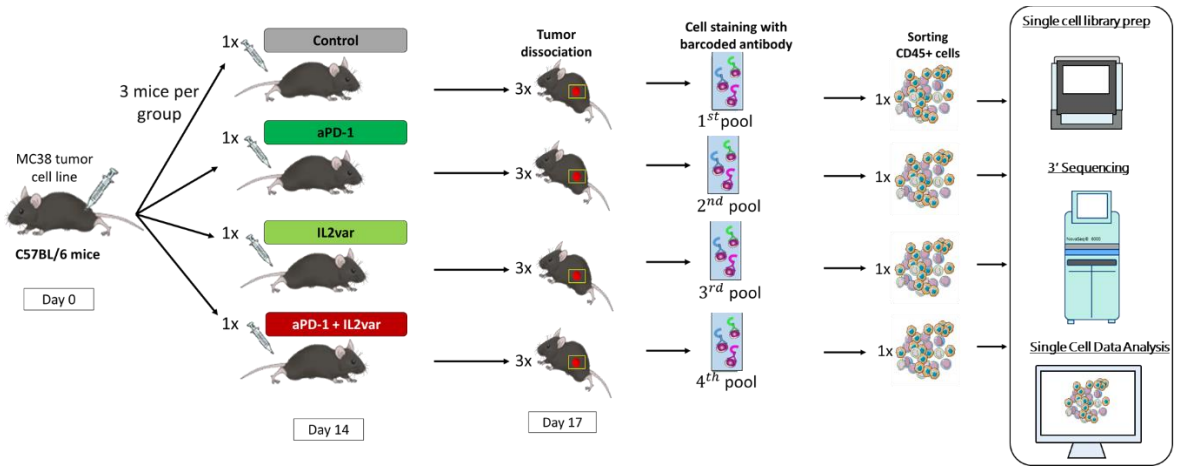


Figure 33: Experimental overview of aPD1 + IL2var treatment on MC38 tumor inoculated mice and workflow of 3´-single cell RNA sequencing. Time schedule mRNA treatment of C57BL/6 mice (injected with MC38 tumor cell line, received one treatment) and experimental approach including preparation of single cells: Tumor dissociation (GentleMACS), CD45 TIL selection (MACS), CD45 staining and barcode labeling (HTO), FACS sort of CD45+, barcoding and sequencing of single cell 3´-libraries and established data analysis pipeline (adapted from [158,159]).

Following barcode labeling, reverse transcriptase and cDNA amplification and final library construction (Gene expression and HTO), the analysis of cDNA and final libraries QC parameters of CD45+ cell samples was done using Qubit and Bioanalyzer (Figure 34a, b and c and Table 14). For the cDNA samples in Figure 34a, all profiles showed a peak fragment profile at ~1800 bp and library construction could be executed. cDNA concentrations for cells treated with IL2var or aPD1 and IL2var were slightly lower concentrated (

Table 14), which could indicate potentially a lower loaded cell number. No primer dimers could be detected in Figure 34b. During the library preparation, samples were washed using SPRIselect beads and the supernatant was taken to prepare the smaller libraries consisting of HTO sequences of the barcoded antibody labeling. Electropherograms of final HTO libraries showed quite equal profiles between each other and no indication of disturbing fragments (Figure 34c). Final libraries for IL2var and the combination treatment aPD1 and IL2var were lower concentrated for GE and HTO, probably due to lower amount of loaded cells (

Table 14) but all samples could be sequenced.

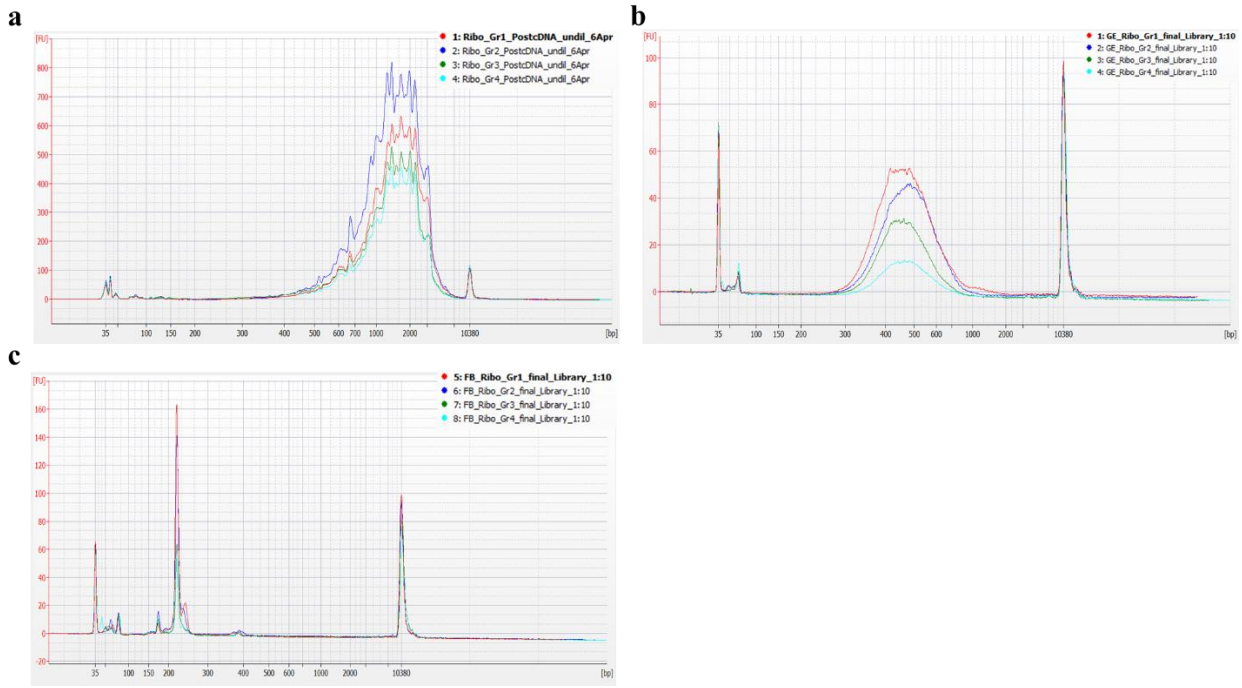


Figure 34: Electropherogram of cDNA and final libraries (GE and HTO) of CD45+ murine TILs. The fragment length is shown on the x-axis [bp] whereas the amount of the fluorescence signals is illustrated on the y-axis [FU]. Both standards can be seen at 35 bp and 10380 bp. The signals of the mouse (**a**) amplified cDNA- and (**b**) final GE library samples and (**c**) final HTO library samples are illustrated in different colors (Control: red; aPD-1: dark blue; IL2var: dark green; aPD-1 and IL2var: light blue).

Table 14: DNA concentration of CD45+ murine TILs of cDNA and final libraries (GE and HTO). Concentrations in ng/µl were measured by Bioanalyzer and Qubit. This table contains data from both used protocols (GE and HTO).

		Qubit	Bioanalyzer	
Protocol	**Sample**	**Concentration (ng/µl)**	**Fragment size (bp)**	**Concentration (ng/µl)**
Gene expression (cDNA)	Control	9,36	1710	7.97
	aPD-1	11,2	1639	8.58
	IL2var	7,44	1593	6.03
	aPD-1 + IL2var	6,2	1637	5.25
Gene expression (final library)	Control	8,82	493	10.35
	aPD-1	9,42	492	7.59
	IL2var	6,5	474	4.99
	aPD-1 and IL2var	2,54	485	2.18
HTO (final library)	Control	5,74	226	3.95
	aPD-1	5,08	228	4.34
	IL2var	3,16	221	1.83
	aPD-1 and IL2var	2,68	221	1.84

All final GE and HTO libraries were sequenced using the Illumina NovaSeq6000 device. Generated clusters were checked with the Illumina software "Sequencing Analysis Viewer". More than 79% of the clusters passed filtering. For read 3 more than 95.3 % of all bases had a quality score of Q30 (1 error in 1000 bases) or more. The error rate in read 3 was only slightly higher than the expected 1% added PhiX control (with 1.23%).

Table 15: Raw sequencing results of CD45+ murine TILs. QC results and statistics of HiSeq4000 run are given as means and were summarized by Sequencing Analysis Viewer. In detail: Cluster PF, Quality Score of 30% or higher and Error Rate.

Read	Clusters PF (%)	% >= Q30	Error Rate (%)
1	79.37	96.74	1.23
2	79.37	93.7	0
3	79.37	95.3	1.23

Following the sequencing, Cell Ranger software was performed for demultiplexing. Between 6,511 – 7,556 cells could be recovered (estimated number of cells listed in Table 16), whereas IL2var and combination treatment sample (aPD-1 and IL2var) were not as low concentrated as assumed before. However, the level of median genes per cell was between 756 – 1,730 and sample IL2var as well as aPD-1 and IL2var treatment group showed very low amount of median genes per cell. Sequencing saturation was good with >78% for all samples, and more than 58.8% of the reads mapped against the transcriptome.

Table 16: Summary of sequencing results of CD45+ murine TILs. Following results after demultiplexing for each sample are depicted: Estimated number of sequenced cells, mean reads per cell, median genes per cell, total number of detected genes, median UMI counts per cell, sequencing saturation [%], reads which mapped confidently to transcriptome [%] and number of amplification samples.

Sample	Input number of cells	Estimated number of cells	Mean reads per cell	Antibody Reads Usable per Cell	Median gene per cell	Median UMI counts per cell	Antiobody Median UMIs per cell	Seq saturation (in %)	Reads mapped confidently to transcriptome (in %)	Recovery rate (in %)
Control	15,000	6,511	179,308	2796	1,229	2,831	66.9	85.7	66.9	43.4
aPD-1	15,000	6,889	110,679	3473	1,730	4,259	63.7	78,0	63.7	45.9
IL2var	15,000	7,057	101,58	2187	970	1,923	58.8	81.6	58.8	47,0
aPD-1 and IL2var	15,000	7,556	107,141	2040	756	1,410	60.9	84.5	60.9	50.3

HTO demultiplexing for the feature barcoding libraries was done to determine the number and percentage of single cells labeled with a barcoded antibody (**singlets**), the amount of cells not labeled (**negatives**) and labelling of more than one feature barcoding antibody which a high amount of cells indicating more than one single cell (**doublets**). All four groups showed a higher level of negatives than expected (supplementary Figure 56b). Especially the combination treatment group where more than 58% of all cells were annotated as negatives, cells were also low in media gene per cell level. All other three groups with cells consisted of 26.8 – 37.6% negatives. In general, too high amounts of total cells were not correctly stained with the barcoded antibodies since the amounts of negatives was quite high (ambient RNA might have been labeled and amplified with Total-seq antibodies). For optimization cells with antibodies could be washed more thoroughly next time for instance by introduction of an additional cell washing step. In addition, the determination of the antibody-titer might potentially also lead to a more efficient antibody staining. After these results, some further investigations were done: Singlets, negatives and doublets were clustered in a tSNE plot (Figure 56a) and the level of RNA UMI counts was determined for each group. A high amount of ambient RNA was amplified during the library preparation (not shown here) and later counted as “negatives”. Combination treatment group showed a high amount of this ambient RNA. In addition, the HTO clustering revealed that cells which were counted as singlets were clustering together with “doublets”. From their gene count profile (not shown here) there was no indication that a doublet would contain more than one cell and did not show higher level of gene counts. This might be the reason for joint clustering with the singlets. Due to this findings, singlets and doublets could not be convincingly distinguished by the HTO method and both were used for downstream analysis as “singlets”. However, doublets could also be distinguished by marker-

based analysis as described later. The cells assigned as "negatives" indicating ambient RNA stemming from empty droplets were excluded from downstream analysis.

Raw data processing and QC steps were done using Seurat analysis tool (described in chapter 6.5.1). All four treatment groups were integrated for better comparison. Dimensional reduction was done and cells were visualized using UMAP. Various canonical markers were used (only a few are shown in Figure 35) to examine the presence T cells (*Cd3d*, *Cd4* and *Cd8a*), macrophages (*Cd14*), B cells (*Cd79a*), NK cells (*Ncr1*), γδ T cells (*Trdc*), proliferating cells (Mki67), dendritic cells (*Cd74*), classical monocytes (*Ly6c2*), cells with an exhausted phenotype (*Pdcd1*), neutrophils (*G0s2*), mast cells (*Cpa3*), cells with an effector phenotype (*Prf1*), low quality cells (low number of genes) and cells with a high level of mitochondrial genes. The analysis revealed two doublet cluster (cluster 18 and 21) since these cells expressed lymphocyte and myeloid marker at the same time. These clusters were excluded as well as cluster 15 which consist of low quality cells with low amount of genes [114].

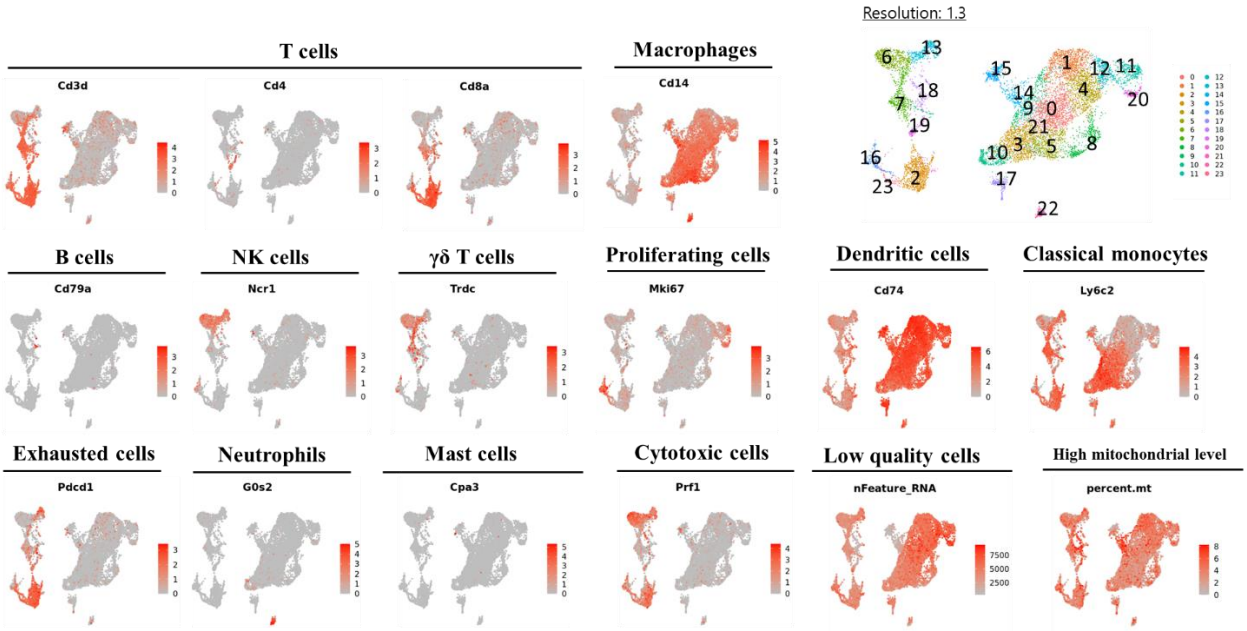


Figure 35: Identification of cell subsets in cells of CD45+ murine TILs. Selected marker expression visualized in UMAP, as Featureplot. A gene which is higher expressed in a cell is colored in darker red. Various markers were plotted for the detection of T cells, macrophages, B cells, NK cells, γδ T cells, proliferating cells, dendritic cells, classical monocytes, exhausted cell, neutrophils, mast cells, cytotoxic cells, low quality cells and cells with a high level of mitochondrial genes as well as an overview of all identified cell clusters by Seurat [dimension: 18; resolution: 1.3].

Figure 36a illustrates 12 preliminary identified cell populations. Although M1 and M2 macrophage marker were investigated (not shown here), it was not possible for more fine determination since the marker expression was not distinct. Cells with *Cd14* gene expression were annotated as "macrophages". Beside macrophages, neutrophils, mast cells, a classical monocyte population and a small cluster with mature cDCs (expression of *Fscn1*, *Csf3* and *Ccr7* genes, expression not shown here) and proliferating cells (myeloids) were identified. For CD4+ T cells a deeper CD4+ T cell subset discrimination through Seurat was not possible due to small CD4+ population and other interfering lymphocytes and myeloids. Three different CD8+ effector T cell populations were found: A cytotoxic T lymphocyte (CTL) CD8+ population which clustered together with γδ T cells, one CD8+ CTL with a more proliferating phenotype and a CD8+ CTL population expressing also exhaustion markers (termed CD8+ Late Effector). More γδ T cells clustered in a separate part of the UMAP together with NK cells. Due to the various different immune population's finer distinction between similar cell subsets like cytotoxic lymphocytes was not possible (i.e. NK cells or CD8+ CTLs clustered together with γδ T cells). Therefore, subclustering strategy was applied: All lymphocyte clustered were separated and re-analyzed (Figure 36b).

After the re-analysis of the subclustering it was possible to separate the NK cells and Effector CD8+ from the γδ T cells (Figure 36b). Two clusters consisted of cells with a CD8+ Late Effector phenotype which were combined due to high similarity. A double negative T cell population (DN T cells) was identified as well. These cells are TCRαβ+CD3+ but CD4- and CD8-. This poor investigated subset is described to possess an immunoregulatory/suppressive function and can prevent graft-versus-host disease and autoimmunity [160,161]. DN T cells have significant anti-cancer effects in patient-derived xenograft (PDX) models of lung cancer and leukemia [162–164] and inhibit proliferation and invasion of human pancreatic cancer cells [165]. CD4+ T cells could not further discriminate into T helper or Treg cell subsets due to too low cell number.

After the new assignment of the subclustered lymphocytes (Figure 35), these results were taken and fused with the original UMAP plot. Therefore, the previously assigned lymphocytes clusters were discarded and the newly assigned lymphocyte clusters were "merged" with the myeloids (information about merging strategy see 6.5.5). The new shape of the UMAP with all identified cell types and cell type frequencies are shown in the plot below (Figure 36c and d).

Macrophages represent the majority of the total population between lymphocytes and other myeloid subsets. All other populations show a frequency of 1-6 %. With regard to the scientific question of this project, whether a treatment effect in the form of a higher CTLs frequency was achieved by IL2var alone or by the combination of aPD-1 + IL2var, a second calculation of the numbers of cells per gram of tumor was done (Figure 37a and b).

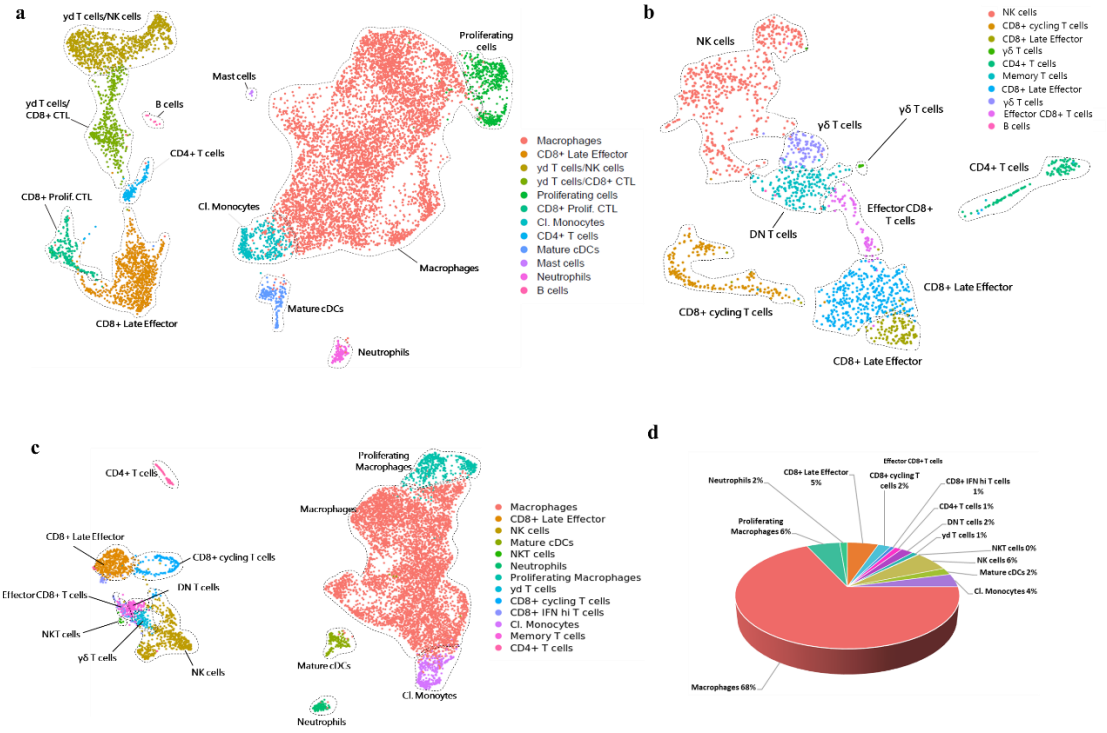


Figure 36: Intermediate UMAP plots with merged cell subsets of CD45+ murine TILs and respective final cell type frequencies. (**a**) This intermediate UMAP plot shows all preliminary identified lymphocytes and myeloids which are illustrated as single dots. (**b**)This UMAP plot shows all subclustered lymphocytes. (**c**) The final UMAP plot shows all identified cells. This UMAP contains the previous subclustered lymphocytes which have been merged with myeloid cells. (**d**) The pie chart plot illustrates all identified cell subsets in different colors with frequencies in %.

Lymphocyte frequencies (Figure 37a) were assessed, whereas higher frequencies for CD8+ Late Effector, CD8+ cycling T cells, DN T cells and γδ T cells could be proven after IL2var or combination treatment with aPD-1 + IL2var. For control and aPD-1 alone a higher frequency was observed in the NK cells. CD4+ T cell frequency was also a bit enhanced upon aPD-1 treatment and more importantly reduced after IL2var and combination treatment. If these CD4+ T cells might have a Treg cell phenotype, a successful reduction of Treg cells by modified IL2var would be proven as well. Certainly, this population consisted of too less cells for downstream CD4+ T cell subset analysis (control group of around 61 cells) and could not be further discriminated. Figure 37b illustrates the numbers of cells (myeloids and lymphocytes)

per gram of tumor (normalized to tumor content) among all four treatments for all identified cell types. The treatment effect of IL2var alone or in combination with aPD-1 (i.e. higher cell type frequency) could be observed and confirmed on cell level for macrophages, CD8+ Late Effector, NK cells, γδ T cells, CD8+ cycling T cells, Effector CD8+ T cells, DN T cells and CD4+ T cells with higher number of cells per gram tumor compared to control and/or aPD-1 treatment. For a clearer confirmation of the treatment effect in TILs, the calculation was performed separately again only for the lymphocytes (Figure 37c). The result answered the first scientific question of the study and could show that the use of IL2var leads to an increase in effector cells such as CD8+ CTLs or NK cells, as well as γδ T cells and DN T cells.

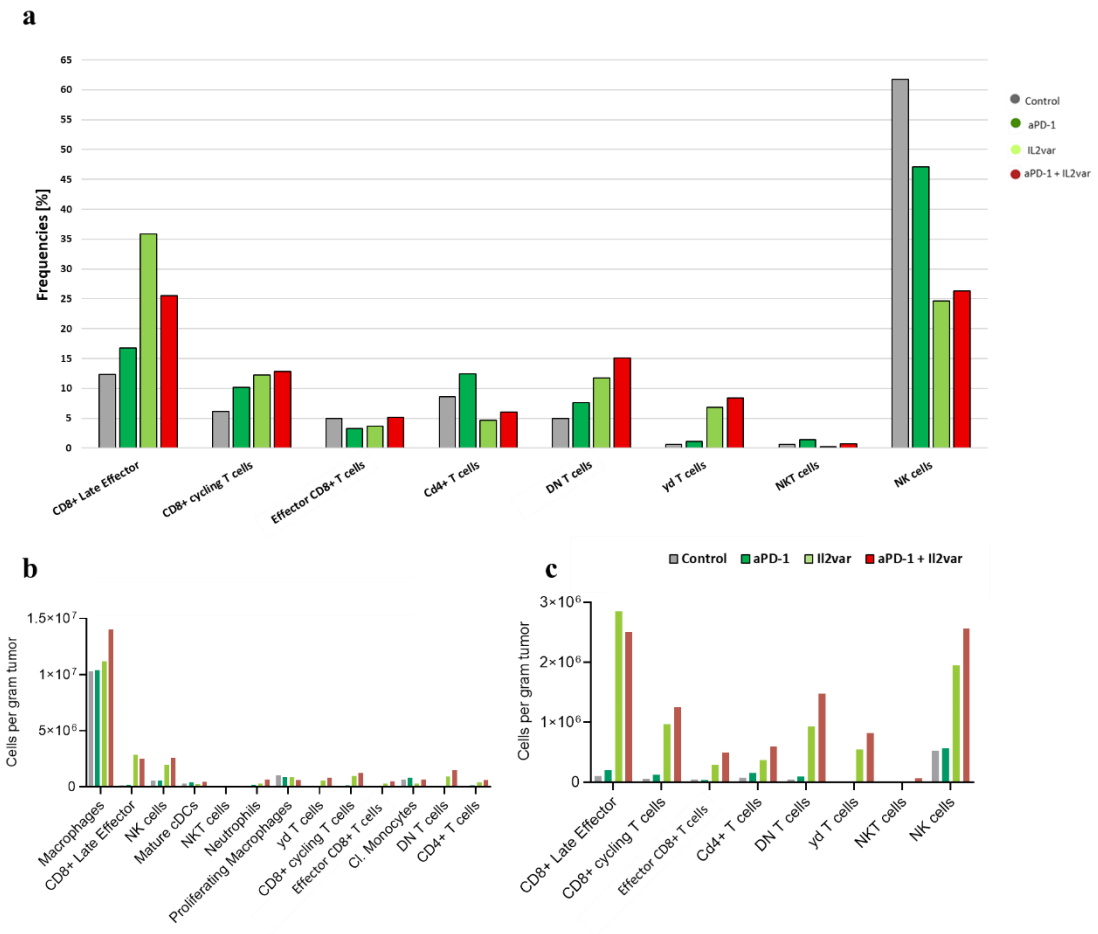


Figure 37: Cells per gram tumor of CD45+ murine TILs and lymphocytes and lymphocyte frequencies. These boxplots illustrate (**a**) lymphocyte frequencies in % are illustrated as well as the amount of cells per gram tumor for (**b**) all identified cell subsets and **(c**) all lymphocytes. On the x-axis and each treatment group in different colors (Control: red; aPD-1: dark blue; IL2var: dark green; aPD-1 and IL2var: light blue).

Furthermore, to investigate and confirm whether the TILs in the tumor have a stronger effector phenotype (upregulation of cytotoxic markers), a gene expression analysis was performed. With regard to the CD8+ Late Effector cells, the influence of the IL2var and/or the combination with aPD-1 on the exhaustion of the cells was also investigated (Figure 38). The focus of the analysis in Figure 38 was on DN T cells, effector CD8+ T cells, CD8+ cycling T cells and CD8+ late effectors since they are considered the main executors of the immune effect. Here, the expression of the transcription factor *Tcf7* (*TCF1*) was found to be specifically upregulated in DN T cells and Effector CD8+ T cells upon IL2var and the combination treatment. In CD8+ cycling T cells the effect could be only observed after the combination treatment. The transcription factor Tcf7 is of such importance because it has been associated with an improved cancer immunotherapy outcomes [166] and to be involved in the effector differentiation [167], which would reflect an successful inhibition and a decrease of the tumor growth through the induction of CD45+ TILs by IL2var and the combination treatment.

In addition to *Tcf7,* a lower expression of the transcription factor *Tox* was observed as well in all four cell populations (Figure 38). The single cell transcriptome study by Kim et al. 2020 showed that *Tox* can be seen as a predictor of anti-PD-1 response in melanoma cancer patients [168] and regulate the development of CD4+ T cells, NK cells and lymphoid tissue inducer cells [169]. Decrease or inhibition of *Tox* results in down-regulation of IC molecules such as PD-1, TIM-3, TIGIT, and CTLA-4 that can promote exhaustion and hinder anti-tumor response by CD8+ cell [168]. The downregulation of *Tox* in all four immune cell subsets illustrates the potential of IL2var to induce a decrease in exhaustion as an anti-tumor response.

In this single cell transcriptome study, the expression of additional exhaustion markers and various effector molecules was analyzed in combined CD8+ T cells (DN T cells, Effector CD8+ T cells, CD8+ cycling T cells and CD8+ Late Effector) to support the hypothesis of a stronger effector phenotype in CD45+ TILs after IL2var or the combination treatment (Figure 38). The expression of markers related to TIL exhaustion (e.g., *Tox*, *Lag3*, *Tigit*) was strongly reduced with all treatments compared to control. Some markers like *Entpd1*, *Nt5e*, *Ctla4* and *Havcr2* (*TIM3*) were not downregulated after IL2var treatment but rather upon the combination treatment with aPD-1, illustrating that IL2var alone could potentially not be sufficient enough to reduce the exhaustion effect in TILs on a broader level. Overall, combined CD8+ T cells showed a heightened expression of genes related to cytotoxicity (e.g., *Gzma*, *Gzmbc*, *Gzmd*,

Gzme, *Gzmf*, *Gzmf*, *Gzmg*, *Gzmk* and *Gzmm*), which supports the theory of IL2 being able to enhance the effector potential of cytotoxic molecules. The results show that some *Gzm* molecules are differentially expressed for IL2var or the combination treatment like e.g. *Gzmc*, which is quite upregulated with IL2var alone, whereas *Gzmg* seems to be higher expressed through IL2var and the addition of aPD-1.

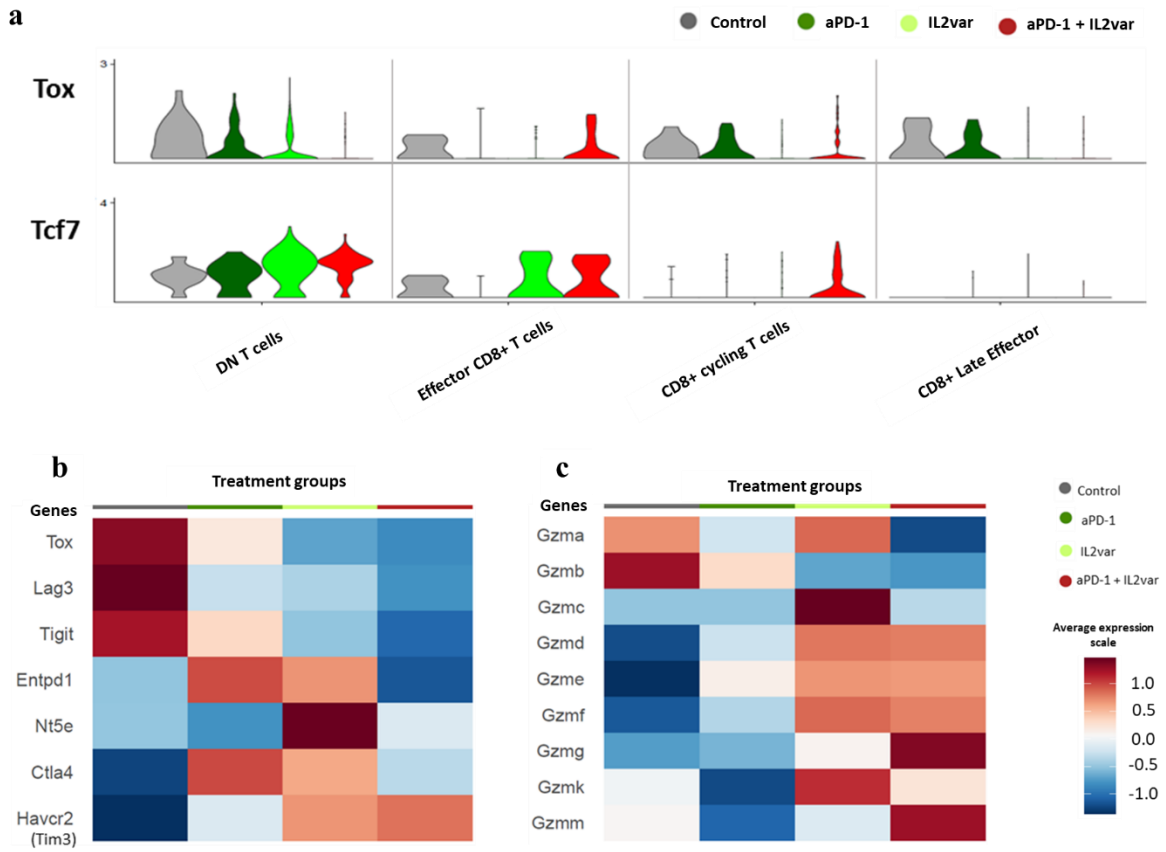


Figure 38. Naïve/memory, effector and exhaustion marker expression of murine combined CD8+ T cells. (**a**) The violinplot shows expression probability distributions of *Tox* (exhaustion marker) and *Tcf7* (TCF1, naïve/memory marker) across clusters. The violin shape comes from the data's density plot. The thicker part means the values in that section of the violin have higher frequency, and the thinner part implies lower frequency. On the x-axis cell subsets (DN T cells, CD8+ IFN hi T cells, CD8+ cycling T cells and CD8+ Late Effector) are separated into treatment groups. The heatmap display the expression of (**b**) exhaustion and (**c**) cytotoxic markers for combined CD8+ T cells. Combined CD8+ T cells: DN T cells, Effector CD8+ T cells, CD8+ cycling T cells and CD8+ Late Effector. Each treatment group is plotted on the x-axis, while the respective genes are plotted on the y-axis (based on their adjusted p-value). Scale: log fold-change of the average expression between the groups. Genes, which are higher expressed, are colored in darker red, genes which are downregulated are colored in darker blue.

4.3 Discussion

The aim of the present research was to study IL2var in the context of a potential cancer treatment in the MC38 tumor mouse model and to evaluate the following hypotheses: i) Is there an increased lymphocytes frequency visible (especially CD8+ T cells and CD4+ regulatory T cells) among IL2var or/and the combination treatment with IL2var and anti-PD-L1 and ii) does the treatment of MC38 tumor bearing mice with IL2var alone or together with anti-PD-L1 lead to stronger effector phenotype (i.e. higher expression of cytotoxic molecules) in CD8+ T cells. Therapeutic cancer vaccine potential could be demonstrated in MC38 tumor mouse model with an increased TILs frequency and strong effector phenotype (upregulation of cytotoxic markers) after treatment with IL2var as well as combining IL2var with aPD-1. Single cell data analysis of >9,000 immune cells revealed 13 distinct CD45+ subpopulations. Overall, tumors of mice treated with IL2var alone or in combination with aPD-1 had an increased number of CD8+ T cells, CD4+ T cells, γδ T cells, DN T cells, NK cells and macrophages compared to Control. The expression of markers related to TIL exhaustion (e.g. *Tox, Lag3, Pdcd1*) was strongly reduced with combination treatment. Importantly, IL2var increased the expression of the transcription factor Tcf7, particularly on DN T cells and a subset of CD8+ T cells.

Regarding the CD4+ T cell frequencies (presumably including CD4+ Treg cells), the recovered population was too small to perform downstream analysis (around 61 cells in control group for T cells). Verification of the hypothesis on reduction of Treg cells could not be studied, but is important with regard to the side effect of IL2 (IL2 is known to expand Treg cells which can suppress the activation of many immune cell types). Due to a widely dispersed *Foxp3* and *Ikzf2* expression (supplementary Figure 61) the entire CD4+ T cell cluster might be a Treg cell population, the results of cell type frequencies (Figure 37c) would show that cells of the treatment group aPD-1 and combination of aPD-1 and IL2var have a lower Treg cell frequency than control or aPD-1 alone. This would confirm a successful reduction of the CD4+ Treg population and show the expected effect of IL2var. However, a slight Th1 marker expression (*T-bet*) and *Ifng* makes it difficult to assign the population exclusively as a Treg cell population and to make a definitive statement about the treatment effects in the current setting (supplementary Figure 61). A follow-up experiment with T cell or CD4+ T cell enrichment might help to study a Treg population and their response to IL2var treatment.

This integrated analysis helped already to reveal and confirm key players and mechanism in a broader perspective taking all immune cell populations into account. Initially a manuscript was

planned as well. However, due to the low cell number in the control group, statistical evaluation for differential gene expression is very limited. Therefore, a follow-up study with a focus on the T cell compartment is planned. This could be possible with the enrichment of T cells to increase the population of CD4+ Treg cells and have fewer interfering myeloids in the downstream analysis. By that a reduction effect in CD4+ Treg cells by IL2var could be potentially confirmed and an alternative approach for the depletion of accumulating intratumoral Treg cells by CTLA4 ICB (immune checkpoint blockade) would be not necessary [170].

The size of tumors for all mice used per group was also examined (supplementary Figure 57). No treatment effect could be observed at that point, but it could be shown later on cell and gene level. In general, an investigation of the CD45+ cells would have been interesting at a later time point when tumor growth had reached a plateau and had not progressed. However, dissection of the tumor at a later time point was not possible because the mice already showed too much discomfort and the tumor had to be removed at day 17. Yet, the collection of CD45+ cells from shrunken tumors also carries the risk that after isolation no accurate analysis can be performed with a too low number of processed single cells (see previous point regarding the too low number of CD4+ T cells for further downstream analysis). Through the fact that some tumors reached a plateau and tumor growth was stopped, there is the possibility that IL2var successfully inhibited cancer cell proliferation. However, Valle-Mendiola et al. 2014 showed that cancer cell proliferation depends on the IL2 dose. The authors in Valle-Mendiola et al. proved that high IL2 dose reduce cancer cell proliferation while low dose can enhance it in cervical cancer cells [171].

After an optimized washing step for barcoded antibodies (results 4.2), barcoded antibodies can be used to distinguish potential outliers of cells in different mice or groups [158]. This would be necessary for a follow-up experiment since there was an issue of the identification of singlets, doublets and neagtives. Another benefit and additional application for Total-seq antibodies would be the simultaneously detection of transcripts and surface proteins. This enables targeted proteomics together with flow cytometry, mass cytometry and transcriptomics by scRNA-seq in general. Since the barcode sequence space for the labelling of different CITE-seq antibodies is virtually unlimited, this approach has a clear advantage in the number of different proteins that can be detected simultaneously [158]. However, commercial Total-seq antibodies are

conjugated to 1–2 oligos per antibody and so antibodies need to be labeled with a higher number of oligos per antibody. By that, sensitivity would be increased and a more robust detection in cells might be potentially achieved. Kleino et al. 2022 produced sets of antibodies with a mean labelling varied from 1 to 6 oligos per antibody [172]. Their modified oligo-labelling improved identification of weak protein targets in scRNA-seq multimodal experiments. The optimized oligo-labelling reduced sequencing costs due to a more balanced amplification of different antibody signals and was stable in fixed cells [172]. This further demonstrates the feasibility of a stain, fix, store, and analyze later for multimodal scRNA-seq which could be of important use for further investigations of IL2var to assess e.g. the amount of gene produced *Ifng* or *Il2* by CD45+ cells and its effect in the TME (tumor microenvironment) after IL2var treatment at different time points. On a transcriptional level, *Ifng* and *Il2* expression was compared between all treatments, whereas cells treated with IL2var or in combination with a-PD-1 treatment showed higher expression of *Ifng* compared to aPD-1 alone or control treatment (supplementary Figure 58b). No striking difference in Il2 expression was observed between the treatment groups (supplementary Figure 58a). Using barcoded antibodies against those translated gene products could allow to uncover differences on protein level. With all these optimization and later potential application of barcoded antibodies, results could be included in a follow-up study.

Another strategy for doublet identification is the tool “DoubletFinder” by McGinnis and colleagues (McGinnis et al, 2019) which is able to predict doublets according to each single cell's vicinity in gene expression space to artificial doublets created by averaging the transcriptional profile of randomly chosen cell pairs. As soon as doublets are removed, the identification of differentially expressed genes is enhanced. The authors also provided a method for estimating DoubletFinder input parameters. By that, the algorithm can be applied across scRNA-seq datasets with diverse distributions of cell types [173].

Downstream discrimination of CD8+ T cells, CD4+ T cells, γδ T cells, DN T cells, NK cells was only possible due to subclustering of lymphocytes. The subclustered and refined lymphocytes were then merged back into the initial dataset with the remaining myeloid populations. Due to reclustering, the UMAP shape changed and the final UMAP looked different to the initial UMAP after merging. Therefore, a transfer of cell IDs would have been another alternative to keep the UMAP shape the same. Another important aspect would be the

precise analysis of further myeloid cell subsets with e.g. a M1 and M2 macrophages phenotype. Another important study was also supported by our single cell genomics team at TRON with scRNA-seq data analysis performed by me, in which J. Beck et al. (2024) demonstrated that a combined treatment of a monoclonal antibody targeting Trp1 (mAb) and IL2 led to an inflammatory reaction by *Ifng*-producing CD8+ T cells in the TME and M1 polarization [174]. IL2 can further reinstate the effector function of (dysfunctional) CD8+ T cells together with PD-1 ICB (immune checkpoint blockade) [175]. Taken together, those findings show that also in this ribocytokine study a subclustering of myeloid cell types would be useful to learn more about the effects of myeloid subsets and their polarization after IL2var and or combination treatment.

A combination therapy with ribocytokine IL2var with antigen-specific RNA-encoded cancer vaccines could potentiate the anti-tumoral effect and would be worth further investigation [176–178]. The combination of different (modified) cytokines like e.g. IL2 and IL7 or IL12 might potentially lead to an enhanced immune response of TILs in TME. IL7 plays a role for the development, maintenance, and proliferation of T lymphocytes [179]. IL7 was demonstrated to assist T cells to restore homeostasis [180] and exerts superior activity to induce the expansion of specific T cells against breast carcinoma than IL2 [181]. IL12 stimulates T and NK cell activity and induces interferon gamma production [182]. A combination of two or even multiple cytokines for treatment would be of interest for potentially more efficient tumor treatment strategies in future to enhance the effect of IL2var.

Another aspect is the optimization of treatment schedule for antigen-cytokine vaccination itself. IL2var (with or without aPD-1) was applied once as single treatment and on cell frequency and gene expression level seemed to have an impact of CD45+ cells in tumors. However, the question stills remains if a single treatment would be sufficient on long term aspect and efficient to achieve complete tumor reduction. Therefore, further studies on the optimization of treatment schedule and the selection of vaccination time point need to be investigated in the future as well.

5 FINAL CONCLUSION AND OUTLOOK

The thesis describes the establishment of the scRNA-seq pipeline as well as its application to investigate the modulatory effects of immunotherapeutic RNA approaches in cancer and autoimmune disease models. The scRNA-seq pipeline consists of the device "Chromium Controller" for capturing single cells and allowing the amplification of a minute amount of RNA for sequencing, which can be a benefit to study the effect of mRNA-LPX vaccinations on a single cell level. The pipeline further consists of several data analysis tools like "Seurat, Cluster Profiler and Gene Set Variation Analysis" for investigation of immune cells in depth and their transcriptional activities among various mRNA-LPX treatments.

In summary, the evaluation of the PBMC and Colo205 cell frequencies showed that the presented single cell analysis pipeline in this thesis is able to perform a quantitative determination of the cell types. It was further verified that there are no remarkable differences in gene expression profiles of PBMC samples from same donors which showed the robustness of the protocol. Furthermore, the 3'GE and 5'GE protocols were compared and overall, no considerable differences in cell composition and cell type specific genes were detected.

In addition to the investigation of human samples as an establishment experiment, murine immune cells in naïve and tumor settings as well as autoimmune disease model were analyzed in three different experimental settings and with the following different scientific backgrounds:

i) In order to investigate the effect of RNA-LPX vaccination on murine splenocytes compared to a control, the established single cell analysis pipeline was used to investigate transcriptional profiles of murine CD45+ cells after the treatment (Chapter Two: Intrinsic adjuvanticity of RNA vaccines modulates innate immune system). These results are part of a manuscript which is currently under preparation. It was shown that mainly B cells and monocyte/granulocyte fraction were slightly increased in frequency by EGFP RNA-LPX treatment. The developed RNA-LPX vaccine with preserved adjuvanticity is able to trigger the expression of chemokines and inflammatory cytokines in splenocytes demonstrated by an enrichment of IFN α pathway, innate immune response and cytokine signaling. Importantly, Kranz et al. 2016 showed a high level of immune response in cells between 6 to 24 h after injection in form of a higher frequencies of CD8+ T cells [51], which suggests follow-up experiments with a longer time frame between injection and capturing of the immune cells. Using this approach a higher level of immune response through EGFP RNA-LPX could be assessed and more potentially enriched

pathways connected to an induction of inflammatory response could be detected. The results presented here provide important insights into the effect of applying RNA-LPX. Those effect could be even more pronounced in combination with checkpoint inhibitors [183].

ii) Furthermore, the impact of autoantigen-encoding non-immunogenic RNA-LPX (MOG35-55_ m1Ψ) vaccination in EAE investigated in comparison to autoantigen-encoding immunogenic RNA-LPX (MOG35-55_U) and irrelevant non-immunogenic RNA-LPX (irrelevant_ m1Ψ) to assess a tolerance induction (Chapter Three: Investigation of antigen-specific CD4+ T cell heterogeneity in a mouse model for multiple sclerosis demonstrates induced tolerance response). The findings proved a successful tolerance induction in EAE sick mice upon MOG35_55_m1Ψ treatment compared to irrelevant_m1Ψ and MOG35_55_U treatment accounted by a low effector T cell frequencies upon MOG35_55_m1Ψ treatment. This was further demonstrated at the transcriptional level in the form of an upregulation of genes connected to Treg activity, anti-inflammatory cytokine action, effector molecule expression as well as down regulation of genes associated with proliferation processes. Since this study focused only on antigen-specific CD4+ T cells further investigations on other immune cell subsets like B cells and NK cells might be of important relevance and should be further studied in follow-up experiments as well. In addition to MOG35-55_ m1Ψ vaccination, other research groups also investigated the possibility to apply CAR T cells to eliminate self-reactive immune cells driving MS. Yi and colleagues worked with an MOG35-55 construct containing a peptide-MHCII chimeric antigen receptor (MOG35_55 pMHCII-CAR) targeting only higher-affinity TCRs to prevent the induction of EAE. Although the results are promising the group indicated that more research needs to be done in this field [184].

iii) Lastly, transcriptomic profiling at single cell level was used to demonstrate the therapeutic potency of a modified IL2var alone or in combination with the antibody aPD-1 of CD45+ cells in a MC38 tumor mouse model (Chapter Four: Combined IL-2 variant and anti-PD 1 treatment induces antitumoral changes in the immune landscape in a mouse tumor model). In summary, the therapeutic cancer vaccine potential could be demonstrated in the MC38 tumor mouse model with an increased TILs frequency and strong effector phenotype (upregulation of cytotoxic markers) after treatment with IL2var as well as combining IL2var with aPD-1. However, the expected reduction of Treg cell frequency could not be completely proven since CD4+ Treg cells were not identified. Therefore, follow-up studies with an enrichment of T cells are necessary to increase the population of CD4+ Treg cells and to potentially confirm the

reduction effect in CD4+ Treg cells by IL2var. In addition, further studies could pave the way to enhance the immune response by combining other antigens like OVA mRNA, E7 mRNA or Trp1 mRNA due to their anti-tumor response potential [145–147] or using other (modified) cytokines like IL7 or IL12.

At the moment, many more newly developed single cell genomics methods are available facing new challenges in order to answer more scientific questions in various scientific fields in the future. This opens many more possibilities to deeply investigate the characteristics of complex diseases with recent advances in single cell transcriptional approaches:

In the study of Andrews et al. 2021 the application of single-nucleus RNA sequencing (snRNA-seq) enabled the detection of rare subtypes of liver mesenchymal cells e.g. of cholangiocyte progenitors that had only been observed during in vitro differentiation experiments [185]. This approach bypasses the cell dissociation step required for scRNA-seq by using detergents to release nuclei from intact cells [185] and could potentially give the opportunity to investigate so far unknown but important cancer subsets or immune cells which might be responsible for a treatment response. One drawback of droplet-based microfluidics is the potential reduction in certain cell types as neutrophils. snRNA-seq could be considered to circumvent this for specific projects where sensitive subsets are in focus [186,187].

Another method that assesses concomitant high-quality accessible chromatin profile of thousands of single cells is called single cell assay for transposase-accessible chromatin with sequencing (scATAC-seq). This method allows the inference of mtDNA heteroplasmy, clonal relationships, cell state and accessible chromatin variation in individual cells to link epigenomic variability to subclonal evolution and deduce cellular dynamics of differentiating hematopoietic cells in vitro and in vivo [188]. The latter could be very useful in future investigations to understand the molecular mechanisms and epigenetic regulation of immune cell differentiation. However, in standard scATACseq, mtDNA muation calling is not included. Scientists adapted this method in order to have this additional information (mtscATAC-seq) [188]. Some research groups even use the advantage of integrating scRNA-seq and scATAC-seq data (Zhang et al. 2022) to learn more about the relationship between chromatin region and genes and identify unknown cell types in their datasets [189].

Notably, new technologies simultaneously retrieving the surface proteome, the transcriptome and the chromatin accessibility of individual cells are part of the so called “single cell

multimodal omics" approach. By leveraging multiplexed fluorescence, DNA, RNA and isotope labeling, the detection of tens to thousands of cancer subclones or molecular biomarkers [190] is enabled. Single cell multimodal omics method could be used to study the causalities and e.g. in patients with neuroinflammatory diseases by capturing different modality of cellular functions at the same time (DNA, RNA and protein surface detection etc.).

In the context of cancer, the investigation of archived samples based on Paraformaldehyde-fixed (PFA) and formalin fixed, paraffin embedded tissues (FFPE) have a tremendous value to get more information on the tumor microenvironment and surrounding tissues and cells. "Chromium Single Cell Gene Expression Flex" (10x Genomics) is a whole transcriptome approach working on the instrument "Chromium X", recently added to TRON's platform technologies, and enabling transcriptome profiling in these critical and complex sample types [191]. Furthermore, in order to measure all the gene activity in a tissue sample and map its location within the tissue, the spatial capture technology "Visium Spatial Gene Expression" (10x Genomics) was developed, using spatially barcoded mRNA-binding oligonucleotides [192] and also recently added to TRON's platform technologies. This can be applied to gain a better understanding of gene and protein expression in the tissue microenvironment of e.g. skin cancer patients and assess potential impacts of future treatment strategies on the cell microenvironment and all involved cell compartments.

The generation of all these data provides new opportunities to accelerate e.g. the identification of cell types due the use or own building of large single cell atlases. These atlases enable the generation of an integrated map of all cells across a tissue type or even whole organisms [193]. As mentioned above they can be also used to map single cells to a reference and even to align datasets of different single cell sequencing methods [194]. All these technologies provide a tremendous outlook on how individual cells could be characterized in the future, far beyond of only transcriptomes and epitope analysis.

6 MATERIALS AND METHODS

This thesis contains methods which were used for the establishment of the single cell sequencing technology of 10x Genomics as well as the application of this technique for the assessment of biological questions concerning the autoimmune disease multiple sclerosis and immunotherapy treatment of cancer. For the establishment of scRNA-seq, different conditions like PBMC donors, sample material (fresh frozen and fresh prepared PBMCs), different species like human and murine samples were compared in various experiments as well as different sequencing techniques (3´scRNA-seq vs. 5´scRNA-seq).

All involved methods under section **6.1**, **6.2** and **6.3** were kindly undertook by various PhD students and scientists from TRON and BioNTech.

Cell sorting was performed by scientists from TRON and technicians at BioNTech.

6.1 mRNA constructs and *in vitro* transcription

In vitro transcribed mRNA is a single-stranded mRNA (ssRNA) that consists of a 5'cap, 5'-UTR and 3'-UTR that both flank the coding sequence (CDS) of a specific protein or peptide of interest and a 3'poly(A) tail of specific length. Altogether, structurally it resembles naturally occurring eukaryotic mRNA. All structural elements, nucleotide modifications as well as the CDS itself can influence the stability, translation efficiency and immunogenicity of mRNA.

6.1.1 Preparation of RNA constructs

RNA construct for the investigation of EAE

The sequences corresponding to mRNA for EAE investigation was cloned into the multiple-cloning site (MCS) of pST1-hAg-MCS-FI-A30LA70 plasmid. The plasmid-backbone contained sequences corresponding to: human alpha globin 5'UTR (hAg), a 3' UTR of FI element and a poly(A) tail of 100 nucleotides, interrupted by a short linker after 30 nucleotides (A30LA70, L = GCAUAUGACU). The coding sequences of the antigen-encoding templates MOG35-55 was fused to an upstream mmsec and downstream mmMITD sequence (pST1-hAg-mmsec(opt)- MCS-mmMITD (opt)-A30LA70), respectively, as the addition of a leader peptide and an MHC trafficking signal (MITD) strongly improves antigen presentation by APCs [195].

RNA construct for impact assessment of EGFP mRNA treatment

RNA constructs were generated by in vitro transcription [196]. Plasmid templates were based on derivatives of the pCMV-Script-Vector (Stratagene) [197]. All vectors are pharmacologically optimized for RNA stability and protein translation 7 and encode a T7 promoter, a 5' human α-globin-UTR, a 3' UTR of two sequential sequences of human β-globin and a poly(A) tail of either 100 nucleotides, with a linker after 70 nucleotides (OVA MHC class I and II), or of 120 adenosines nucleotides (OVA MHC class I). The OVA constructs encode the H-2Kb-restricted, immunodominant epitope OVA257-264 (SIINFEKL) (OVA MHC class I), or a fusion of the H-2Kb-restricted epitope with the I-Ad-restricted epitope OVA323-339 (ISQAVHAAHAEINEAGR) (OVA MHC class I and II). Both are equipped with the secretion signal for routing to the endoplasmic reticulum and the transmembrane domain derived from human MHC class I (MITD) for improved presentation of MHC class I and II epitopes as described previously (OVA MHC class I) [195], or the analogous murine sequences (OVA MHC class I and II). Luciferase (LUC) RNA and enhanced green fluorescent protein (EGFP) RNA were described previously [51]. AdjRNA encodes the optimized backbone identical to the OVA MHC class I and II construct but the antigen sequence is replaced by a GS linker (GGSGGGGSGGGGSGGGGSGG). RNA was capped with a β-S-ARCA cap 0, eluted in H2O and stored at ¬-80 °C until further use. RNA was formulated with liposomes composed of DOTMA and DOPE to yield RNA-lipoplexes (RNA-LPX) with a negative net charge for targeting of lymphoid compartments by intravenous administration [51].

RNA construct for the investigation of Ribocytokine modified Il2 treatment on MC38 mice

This RNA construct is called here just IL2var and currently in development as a promising treatment candidate [198,199].

6.1.2 Synthetic peptides EAE

For EAE disease induction, MOG35-55 peptide (amino acid sequence: MEV GWY RSP FSR VVH LYR NGK) was needed and obtained from Jerini Peptide Technologies, Berlin, Germany. Before usage, peptides were dissolved in H2O to a final concentration of 5 mg/ml or 20 mg/ml. Peptides were stored at -20 °C until further use [105].

6.1.3 Generation of non-immunogenic mRNA.

For in vitro transcription of mRNA from the DNA template, 1-methylpseudouridine-5'-triphosphate (m1ΨTP) (TriLink) was used instead of uridine-5'-triphosphate (UTP), and mRNA was purified by HPLC or cellulose [52]. For HPLC purification the protocol of

Weissman et al., 2013 [200] was adapted and elution of the mRNA was performed with a gradient of 38 % - 70 % of Buffer B. mRNA quality control was performed by spectrophotometry on a 2100 Bioanalyzer (Agilent technologies). Absence of double-stranded RNA (dsRNA) was confirmed using the dsRNA-specific mAb J2 (Scicons) as described elsewhere [53]. For the generation of standard IVT mRNA (immunogenic mRNA) uridine-5'-triphosphate (UTP) was used during the reaction.

6.1.4 EAE induction and clinical assessment

Acute experimental autoimmune encephalomyelitis (EAE) was induced in female C57BL/6 mice (8-10 weeks) with 50 µg of MOG35-55 peptide emulsified in CFA (Difco Laboratories) supplementaryed with 10 mg/ml of heat-inactivated Mycobacterium tuberculosis H37RA (Difco Laboratories) subcutaneously at the base of the tail. The mice received 200 ng Pertussis toxin (PTX) (List Biological Laboratories, INC., Campbell, CA) intraperitoneally on the day of immunization and 2 days later. Starting on day 10 after immunization mice were daily weighed and scored for clinical symptoms according to the following criteria: 0, no disease; 1, decreased tail tone; 2, impaired righting reflex; 3, partial hind limb paralysis; 4, complete hind limb paralysis; 5, hind limb paralysis with partial fore limb paralysis; and 6, moribund or dead [105].

6.1.5 Application of mRNA on mice

<u>EAE investigation</u>

To determine protective immunity in C57BL/6 EAE mice, mice were treated with 20 µg MOG35-55 m1Ψ, 20 µg irrelevant m1Ψ RNA-LPX or saline unless stated otherwise on days 7 and 10 or at an EAE score of 1-2 after disease induction [105].

<u>Assessment of EGFP mRNA treatment</u>

Mice were vaccinated with RNA-LPX i.v. once, or with three to four weekly doses as indicated. In experiments where AgRNA-LPX and AdjRNA-LPX were combined, RNA-LPX consisted of mixtures of AgRNA-LPX (encoding OVA MHC class I and II) and titrated amounts of AdjRNA-LPX as described. IFNα was analyzed after the first vaccination at the time points indicated. For the assessment of costimulatory marker expression, splenocytes were isolated, stained and analyzed by flow cytometry 24 hours after vaccination. For the characterization of T-cell responses, blood and splenocytes were analyzed five days after vaccination by flow cytometry. For single cell RNA sequencing, splenocytes were isolated 3 hours after vaccination and further processed as described below.

Ribocytokine modified Il2 treatment on MC38 mice

Nanoparticulate formulated RNA vaccines in isotonic buffer solution were performed i.v. into the retrobulbar venous plexus or the mice. It was administered under 2.5% isoflurane/oxygen anesthesia into the retrobulbar venous plexus (maximum volume 200 µl) using a syringe with a 29 G needle attached. The animal was returned to its cage and reawakening from anesthesia was ensured.

6.1.6 Antibody treatment

C57BL/6 EAE mice were treated with PD-1 (500 µg first two treatments, 250 µg following treatments, clone RMP1-14, BioXcell), or CTLA-4 (500 µg first two treatments, 250 µg following treatments, clone 9H10, BioXcell) blocking antibody or isotypematched control antibodies (rat IgG2a and syrian hamster IgG, BioXcell) on days 7, 10, 14 and 17 after EAE induction. Antibodies were diluted in PBS and applied intraperitoneally [105].

6.2 Animals

6.2.1 Mouse strains

EAE sick mice

C57BL/6 mice were purchased from ENVIGO RMS GmbH, Netherlands and Janvier Laboratories, France, respectively. C57BL/6 mice for the investigation of EAE were maintained underspecific pathogen-free (SPF) conditions at the animal facility of BioNTech AG Mainz, Germany. Animal experiments were conducted in accordance to German federal and state animal experimentation regulations and approved by the responsible national authority (national investigation office Rhineland-Palatinate, permission number G15-8-024) [105].

Impact assessment of EGFP mRNA treatment

Female 8–12 weeks old C57BL/6 mice (Envigo) were kept and treated as approved by the Ethics Committee for animal research of Rhineland-Palatinate, Germany.

MC38 tumor bearing mice for Ribocytokine investigation

Female C57BL/6JOlaHsd (Envigo) were treated as approved by the Ethics Committee for animal research of Rhineland-Palatinate, Germany

6.2.2 Anesthesia and sacrification of mice

EAE mice

Inhalation anesthesia of mice with 2.5 vol % isoflurane (Baxter, Deerfield, IL, USA) and O2 was induced for blood retrieval via the retrobulbar venous plexus, in vivo bioluminescence imaging experiments and intravenous (i.v.) injections for adoptive cell transfer as well as RNA-LPX application via the retrobulbar venous plexus. Constant anesthesia was ensured by an anesthesia machine (UniVet Mobil, Gropper Medizintechnik, Deggendorf, Germany) and mice were kept in the anesthesia chamber until the respiratory rate stayed constantly slow. For the perfusion of mice, mice were anaesthetized by intraperitoneal (i.p.) injection of a mixture of 120 mg/kg ketamine (Life Technologies, Carlsbad, CA, USA) and 16 mg/kg xylazin (Life Technologies, Carlsbad, CA, USA). As control for deep anesthesia the footpad reflex was applied, and mice were regarded as under anesthesia when the reflex was no longer observed [105].

MC38 tumor bearing mice for Ribocytokine investigation

A 2.5% isoflurane/oxygen mixture was introduced into the anesthesia chamber for short-term inhalation anesthesia. The animal was transferred to the anesthesia chamber and remained there until it was fully anesthetized and regular, shallow breathing was established. After the animal was removed from the anesthesia chamber, inhalation anesthesia with isoflurane provided a surgical tolerance of 30-45 seconds. During this period, the appropriate applications could be performed.

6.2.3 MC38 tumor inoculation

MC38 tumor bearing mice for Ribocytokine investigation

MC38 tumor cell line was cultured for 9 days until tumor inoculation with $7.5x10^5$ MC38 s.c. in 100 µL PBS was done in C57BL/6JOlaHsd mice.

6.2.4 Mouse tissue preparation for single cell suspension

EAE mice

All cell-based analyses were performed on single cell suspensions of blood, spleen, brain and spinal cord. In brief, peripheral blood was collected from the vena facialis. Spleens were mashed through a 70-µm cell strainer (BD-Falcon) using the plunger of a 5-ml syringe (BD Biosciences) while rinsing with PBS. Erythrocytes were removed by hypotonic lysis. The single cell suspension was resuspended in 70 % Percoll (GE Healthcare) and layered under a 30:37 percoll gradient. The final percoll gradient was centrifuged at 300 g for 40 min at room temperature. The lymphocyte cell layer was collected at the interphase between 70 % and 37 % percoll and washed with 2 % FCS/PBS before further analyses [105].

EGFP mice

Spleen single cell suspension were prepared in PBS by mashing the tissue against the surface of a 70 µm cell strainer (BD Falcon) using the plunger of a 3 ml syringe (BD Biosciences). Erythrocytes were removed by hypotonic lysis. Serum was generated at the time points as indicated and stored at -20 °C.

MC38 tumor bearing mice for Ribocytokine investigation

MC38 tumor sizes were measured and harvested until tumors reached a size of 10-70 mm^3. Tumors were weighted and transferred to C tube (Miltenyi Biotec) with 2.5 mL tumor dissociation buffer (Miltenyi Biotec). Dissociation step was done two times: Cutted pieces were dissociated with gentleMACS dissociator (Miltenyi Biotec) and incubated (20 min at 37°C). Digested tumors were sheared with gentleMACS two times and shortly spinned to collect suspension at the tube bottom. Suspension was filtered through 70 µm cell strainer (Miltenyi Biotec) and C tubes as well as strainers were rinsed RPMI1640 (American Type Culture Collection). After centrifugation (460 x g, 6 min) 1 mL erylysis buffer (8,25 g NH4CL; 1g KHCO3; 0,2 ml EDTA, 1 l distilled water, sterile filtered) was added, resuspend, filled up to 5 ml and incubated (3 min at RT). Further 15 ml PBS+2mM were added and tube was centrifugated (460 x g, 6 min). Pellet was resuspended in 1 ml PBS+2mM EDTA and filled up to 10 ml. Cell suspension was transferred through 70 µm cell strainer and centrifuged (460 x g, 6 min) before CD45 MACS TIL isolation (section 6.3.1).

6.3 *ex vivo* techniques

6.3.1 Magnetic activated cell sorting (MACS)

All antibody-coupled microbeads, LS-columns and magnetic stands were obtained from Miltenyi Biotec, Bergisch Gladbach, Germany. Reagents, as well as MACS buffer (5 mM EDTA and 5 % FBS in DPBS) and cells were kept on ice throughout the cell separation procedure.

EAE investigation

Splenic CD4+ T cells of C57BL/6 mice treated on days 0, 3, 7 and 10 with MOG35-55 m1Ψ or U mRNA-LPX, irrelevant m1Ψ mRNALPX or saline were isolated 3 days after the last immunization using L3T4 microbeads and MACS LS columns (Miltenyi Biotec). 3x 10^4 CD4+ T cells were stimulated for 48 h with MOG35-55 peptide-loaded (0, 5, 10, 20 and 100 µg/ml) BMDCs (3x 10^4 cells) and the supernatants were analyzed for cytokine content with a Luminex multiplex cytokine assay (17- plex mouse ELISA assay kit, Thermo Fisher Scientific) [105].

Ribocytokine investigation

Cells were resuspended in 100 µl MACS buffer/Microbeads stock (1:10 dilution of Microbeads from Miltenyi Biotec in MACS buffer) per 107 cells (1 g tumor = 108 total cells) and incubated (15 min at 4°C). LS column (Miltenyi Biotec) was rinsed with 3 ml MACS buffer and sample was filed up with MACS buffer to final volume of 500 µl. Cell suspension was transferred to LS column and tube was rinsed two times with 1 ml MACS buffer. LS Column was placed on collection tube and eluted with 3 ml by pushing down plunger proper fast. Effluent was centrifuged (460 x g, 6 min) and resuspension was done in 550 µl DC medium. Cell suspension was counted and 106 cells per sample were transferred to 96 well plate for FACS staining (section 6.3.2).

6.3.2 Staining for flow cytometry

EAE investigation

Flow cytometry surface antibodies were used in accordance with the manufacturer's protocol. Briefly, for extracellular staining approximately 2x10^6 splenocytes or LN cells (for single cell preparation see section 6.4.7) were used and stained in 100 µl antibody mix in FACS buffer (5 mM EDTA and 5 % FBS in DPBS) for 30 min at 4 °C in dark. Cells were washed with 2 ml DPBS, centrifuged (450 x g, 6 min, RT; Allegra X-15R centrifuge, Beckman-Coulter, Krefeld,

Germany) and cell pellet was resuspended in 200 µl FACS buffer for flow cytometric measurement. Staining of murine blood-derived cells was performed directly on 50 µl total blood cells by adding antibody master mix into whole blood. Furthermore, staining involved an additional erythrocyte lysis step by incubation of cells in 200 µl BD blood lysis buffer (BD Biosciences, Heidelberg, Germany). Stained cells were then washed twice with 2 ml PBS and resuspended in 100-200 µl FACS buffer for analysis [105].

EGFP mRNA treated mice

Staining of immune cells for flow cytometry was performed as previously described [51]. Briefly, viability of splenocytes was determined by staining with LIVE/DEAD™ fixable yellow dead cell stain kit (Invitrogen). CCR7 was stained for 30 min at 37 °C and fixed with stabilizing fixative (BD Bioscience) for 30 min at 4 °C prior to staining of further antigens. Peripheral blood was lysed using BD FACS™ lysing solution (BD Biosciences). The following antibodies were used: CD49b (DX5), CD45 (30-F11), CD40 (3/23), CD69 (H1.2F3), CD4 (RM4-5), CD11b (M1/70), PD-1 (J43) (all BD Bioscience), CD19 (1D3), CD86 (GL-1), F4/80 (BM8) (all Biolegend), CD8 (5H10) (Invitrogen), PD-L1 (MIH5), CCR7 (4B12), KLRG1 (2F1) (all eBioscience) and CD11c (N418) (Miltenyi).

Ribocytokine investigation

MACS isolated TILs were centrifuged (460 x g, 4 min) and uptaken in 100 µl LD stain + Fc block followed by incubation (15 min at 4°C). 100 µl PBS were added and suspension was centrifuged (460 x g, 4 min). Furthermore, 50 µl Barcoded antibody (Biolegend, stock: 0.5 mg/ml; 1:100 diluted in FACS buffer) was added and stained (20 min at 2-8°C). Cells were washed once with 150 µl PBS (centrifugation: 460 x g, 5min), 50 µl antibody Mastermix was added and shortly vortexed. Solution was stained (30 min at 2-8°C) and washed again once with 150 µl PBS (centrifugation: 460 x g, 5min). Cells were resuspended in 100 µl FACS buffer (500 ml PBS, 5 ml 0.5 M EDTA, 25 ml 5% FCS) and samples were pooled per group to obtain 4 samples in total.

6.3.3 Tetramer staining for EAE study

Transcriptome profiles and phenotype of endogenous MOG35-55-specific CD4+ T cells upon vaccination of EAE mice as well as naïve mice with different RNA-LPX was analyzed by tetramer staining. MOG35-55-specific APC-conjugated pMHC class II tetramer was obtained from the National Institutes of Health Tetramer Core Facility, Emory University Vaccine

Center, Atlanta, GA, USA. 10x106 cells from single cell suspension of mouse splenocytes were incubated with MOG35-55 APCconjugated tetramer (1:50 dilution) for 60 min at RT in 500 µl complete DC medium. Cells were washed once with DPBS and anti-APC MACS was performed (section 6.3.1) according to the manufacturer's recommendation. Cells were incubated with the remaining extracellular or intracellular antibodies as described in sections [105].

6.3.4 Flow cytometry

Flow cytometry or fluorescence activated cell sorting (FACS) is a useful technique to define multiple characteristics of individual cells in both qualitative and quantitative ways from a heterogeneous cell suspension. Different cell components are analyzed by fluorescently labeled monoclonal antibodies binding to their respective targets. While single cells pass in a liquid stream through several lasers, cells scatter light from the laser and the fluorochromes are excited by the laser beams to emit light at varying wavelengths. Thereby, particle's relative granularity, size and fluorescence intensity are measured. Forward scatter (FSC) light signals and side scatter (SSC) light signals measure size as well as granularity and fluorescence detectors measure the fluorescence signal intensity emitted from positively stained cells. All of these different light signals and the emitted fluorescent light are then detected by sensitive photomultiplier tubes (PMT) and digitized for computer analysis.

EAE investigation

Flow cytometric data for EAE study were acquired on a FACSCanto II or LSRFortessa (BD Biosciences, Heidelberg, Germany) depending on the number of fluorophores used per staining panel and analyzed with FlowJo V7.6.5 and V10.4 software (Tree Star, Ashland, OR, USA), respectively. Murine monoclonal antibodies were used for staining of cell surface markers (section 6.3.2). Staining was performed as described in the sections below and the following intracellular antibodies were used: BTLA, CCR6, CCR7, CD11b, CD11c, CD19, CD25, CD3, CD4, CD40, CD44, CD49b, CD5, CD62L, CD69, CD8, CD80, CD86, CD90.1, CTLA-4, Foxp3, ICOS, IFNγ, IL-17A, LAG-3, KLRG1, MHC-II, PD-1, PD-L1, RORγt, T-bet, TIGIT and TIM-3. These antibodies were purchased from eBioscience, Bioledgend or BD Pharmingen and used in accordance with the manufacturer's protocol [105].

Single cell suspensions were stained for 30 min at 4 °C for extracellular markers. Chemokine receptors were stained at 37 °C for 30 min and subsequently fixed with stabilizing fixative (BD

Biosciences) for 30 min at 4 °C. After performing live-dead staining (fixable viability dye APCeFluor® 780, eBioscience) and staining of cell-surface markers, cells were fixed and permeabilized using Cytofix/Cytoperm and Perm/Wash buffer from BD Biosciences according to the manufacturer's protocol. Cells were incubated for 30 min at 4 °C with intracellular antibodies and washed twice with Perm/Wash before FACS analysis. Samples were acquired on a BD FACSCanto II, BD FACSCelesta or BD LSR Fortessa and analyzed using FlowJo 7.6.5 or FlowJo 10.4 software (Tree Star). For intracellular staining of Foxp3, CTLA-4, RORγt and T-bet the Foxp3 transcription factor staining buffer set (Thermo Fisher Scientific) was used for fixation and permeabilization. Extracellular staining was performed according to the manufacturer's recommendation. MOG35-55-specific CD4+ T cells were detected by tetramer staining using an APC-conjugated pMHC class II tetramer (National Institutes of Health Tetramer Core Facility). Single cell suspensions of mouse splenocytes were incubated with MOG35-55 APC-conjugated tetramer for 1 h at room temperature in culture medium. Cells were washed and anti-APC staining and subsequent anti-APC MACS were performed. Cells were incubated with extracellular or intracellular antibodies as described above. For additional cell sorting of MOG35-55-specific CD4+ T cells, cells were stained with live-dead dye (APCeFluor® 780, eBioscience) and MOG35-55-specific CD4+ T cells were then sorted on a FACSAria cell sorter (BD Biosciences) according to their surface marker expression. Absolute cell counts (total cell infiltrates, CD3+, CD8+, CD4+, Thy1.1+ 2D2 CD4+ T cells) in single cell suspensions of spleen, brain and spinal cord were determined with Trucount™ Tubes (BD Biosciences) by flow cytometry [105].

Spleens were isolated, tetramer staining was performed (section 6.3.3 and MOG35-55-specific CD4+ T cells were pre-enriched by anti-APC MACS (section 6.3.1). Cells were resuspended in FACS buffer, sorted on a FACSAria cell sorter (BD Biosciences, Heidelberg, Germany) according to their surface marker expression (live cells/ CD4+/ tetramer+ cells) and collected in 500 µl FACS buffer [105].

EGFP mRNA treated mice

Gating was performed as follows: cDCs: CD11chi F4/80-, NK: CD49b+ CD3-, CD4+ T cells: CD3+ CD4+, CD8+ T cells: CD3+ CD8+, B cells: CD3- CD19+. OVA-specific CD8+ T cells were quantified by staining with H2-Kb-restricted OVA257-264 MHC class I tetramers (MBL International Corporation). Flow cytometric data were acquired on a BD FACSCanto II or a BD LSRFortessa and analyzed with FlowJo 10.5 software (Tree Star).

MC38 tumor bearing mice for Ribocytokine investigation

For sorting the following antibodies were used: barcoded antibodies (Biolegend, clone 30-F11, dilution: 1:200), LD (eBioscience, dilution: 1:800), Fc block CD16/CD32 (Becton Dickinson, clone 2.4G2, dilution: 1:100). In total 2x 15,000 cells per pooled sample were sorted by BD LSR Fortessa into 2x 500 µL PBS + 0.04% BSA.

6.4 Cell biological methods

6.4.1 Cell culture

All methods in cell culture were performed in a laminar airflow cabinet (HeraSafe, Heraues Instruments, Hanau, Germany) under sterile conditions. Cells and cell lines were cultivated in a sterile incubator (HERAcell 240i incubator, Thermo Scientific, Waltham, MA, USA) at 37°C and 5% CO_2. Centrifugation steps were performed with Allegra X-15R centrifuge (Beckman-Coulter, Krefeld, Germany) and Heraeus Megafuge X3R Centrifuge (Thermo Scientific, Dreieich, Germany) for specific preparation of Buffy coats (section 6.4.3).

6.4.2 Thawing of cryoconserved cells

Frozen PBMCs in cryogenic storage vials from in total two donors (D1 and D2) were provided by TRON scientist and removed from liquid nitrogen storage tanks (-120°C) and carefully defrosted in a 37°C water bath (Aqualine AL5, Lauda, Lauda-Königshofen, Germany) until some small ice crystals were visible in the medium. Cryogenic storage vials were disinfected with 70% ethanol and cell suspension was transferred under laminar airflow cabinet to 15 ml reaction tube (Greiner Bio-One, Frickenhausen, Germany) containing 9 ml culture medium. The cryogenic storage vial was rinsed with 1 ml pre-warmed thawing medium containing RPMI 1640 GlutaMAX (Gibco) + 10% PHS (One Lambda). Rinses medium was added to the 15 ml reaction tube while gently shaking the tube. Tube was centrifuged (300 x g, 4min for PBMCs and 5min for Colo205 cells), cell pellets were resuspended in 1 ml culture medium. An additional amount of 9 ml culture medium was added for final volume of ~10 ml. This washing step was done once for Colo205 cells and three times for PBMCs. Colo205 cells were counted with C-Chips and PBMCs with CASY counter (section 6.4.6).

6.4.3 Buffy coat preparation

The bag with the fresh blood sample (BC) was disinfected comprehensively from the outside with Bode Bacillol AF tissues (Bode) by TRON technician. Under a sterile bench, the lower end of the tubing was cut off with scissors and disinfected with Terralin liquid (Schülke). Blood

from one BC was divided equally among three 50 ml reaction tubes by tipping. Two of the 50 ml reaction tubes were filled up to 50 ml (up to 50 mL mark on the tube) and one to 40 ml using a 25 ml rod pipette (Greiner Bio-One, Frickenhausen, Germany) with PBS and inverted several times to obtain a homogeneous solution. Four new 50 ml reaction tubes were filled with 15 ml of Ficoll using 25 ml rod pipette, followed by a slowly pipetting of 35 ml from the diluted onto the Ficoll. Here, the diluted blood must overlay the Ficoll to create 2 phases. All 50 ml tubes were centrifuged using (800 x g, 25 min), whereas the brake of the centrifuge was switched off. For washing, all interphase and supernatant (Platelet, plasma and PBMCs) of each tube was removed and each transferred to new 50 mL tubes using 10 mL rod pipette and filled up to 50 ml with PBS. Tubes were centrifuged (300 x g, 8 min) and afterwards supernatants were immediately transferred to new 50 ml tubes by decanting. Using a 1 ml pipette, each cell pellet in the remaining 50 ml tubes were resuspended with 1 ml PBS. Thus, eight 50 mL tubes were obtained for each BC (4 tubes with resuspended cell pellets + 4 tubes with the supernatants). This washing step was done two times in total. After centrifugation, supernatants were discarded immediately by decantation. After resuspension of cell pellets with 1 ml, all cells were combined in a 50 ml reaction tube. The remaining tubes were rinsed with 10 ml PBS using a 10 ml rod pipette. The resulting cell suspensions were also added to the existing cell suspension from the cell pellets. The obtained cell suspension was filled up to 50 ml with PBS and cell suspension was homogenized by turning it upside down three times. Finally, cell counting is performed (section 6.4.6) and the cells were stored at room temperature for further preparation for scRNA-seq (section 6.4.7).

6.4.4 Cell culture media

If not otherwise stated culture medium (RPMI + Glutamax + 10% FCS, not heat inactivated!) was used for culturing single cell suspension for Colo205 tumor cell line. Medium was pre-warmed for 15-20 min at 37°C in incubator.

6.4.5 Cultivation of cell lines

The Colo205 cell line was purchased from American Type Culture Collection (Virginia, United States of America). Cell lines were thawed (section 6.4.2) and cultivated in the appropriate medium (from section 6.4.4). The cell line was seeded in T75 flasks (Greiner Bio-One, Frickenhausen, Germany) for sub-cultivation with 9 ml culture medium and carefully mixed with pipetting up and down. After centrifugation (300 x g, 4 min) Colo205 cells were resuspended with 1 ml culture medium. Transfer of 1 ml cell suspension was done into a 15 ml

reaction tube with culture media. Cells were counted (section 6.4.6), a new T75 flask was filled with 20 ml culture medium and an appropriate volume for 1.5 x10^6 cells was added to the T75 flask for sub-cultivation (at 37°C and 5% CO2 for total 72h). Cells from two passages (7 days incubation) were used for single cell sequencing establishment experiments. For cell line expansion, T175 flasks (Greiner Bio-One, Frickenhausen, Germany) were used. Culture medium with cells was aspirated and transferred into a 50 ml reaction tube. The T75 flask was further washed with 10 ml sterile PBS and combined with supernatant. Cells were centrifuged (300 x g, 4min), resuspended in 1 ml culture medium and kept at RT. To detach adherent cells, 10 ml Accutase (Sigma Aldrich, Schnelldorf, Germany) was added and cells were incubated (37°C, 5-10 min). Detached cells were transferred into a new 50 ml reaction tube and the flask was washed with 10ml culture medium and combined with detached cells. Reaction tube was centrifuged (300 x g, 4 min) and cells were resuspended in 4 ml culture medium. In the end, 1 ml suspension cell solution and 4 ml detached cell solution were combined to obtain 5 ml cell solution. Colo205 cell suspension was washed two times with 1 ml PBS + 0.04% BSA and filtered using a 30 µm strainer (Miltenyi). Cell numbers were determined via cell counting (section 6.4.6).

6.4.6 Determination of cell number and cell viability

Cell counting using C-Chips

Cell number and viability were determined by manual counting using C-Chips (Peqlab, Erlangen, Germany) [201] in between sub-cultivation cycles, after thawing frozen cell pellets and after single cell suspension preparation. Cells were detached (section 6.4.5) and resuspended in the respective growth medium at approximately 1-2x10^6 cells/ml. Sample was diluted under sterile conditions with cell viability dye Erythrosin B (1 mg/ml in DPBS) [202] and incubated for approximately 30 secs. The diluted cell suspension was transferred to the C-Chip, all four quadrants were counted and the mean determined and multiplied with the dilution factor two, which equaled the concentration of 10^4 cells per ml. The percentages of living cells were determined by dividing the number of live cells by the total cell number. For single cell sequencing suspensions, a cutoff of > 85 % live cells were set.

Cell counting using CASY counter

PBMCs (section Chapter Two:) were counted on CASY TTC (OMNI Life Science GmbH & Co KG, Bremen, Germany). First, CASY was set up by putting a new CASYcup (OLS OMNI

Life Science) with 10 ml fresh CASYton (OLS OMNI Life Science) under the measurement capillary. A background measurement (<100 counts/ml) must be performed for accuracy in beforehand. For PBMCs a cell specific setup was adjusted under "Set up management" before measurement (Size scale: 0-20 µm; Evaluation cursors: 5.61 to 20.00 µm; Normalization cursors: 4.59 to 20.00 µm). Sample volume of 10 µl was added to CASYcup with 10 ml CASYton (dilution 1:1000) and CASYcup was gently inverted three times. CASYcup was placed under capillary and measurement was started. Cursor settings (Normalization cursor in blue for cell debris and evaluation cursor in red for dead cells) had to be specified manually according to the picture below to identify debris, dead cells and living cells and to determine total number of cells, number of living and dead cells as well as cell viability in %.

6.4.7 Preparation of single cell suspensions for scRNA-seq

Establishment scRNA-seq of human and mouse

First, PBMCs and Colo205 cells were washed beforehand. In Brief, PBMCs and Colo205 cells in 15 ml reaction tubes were centrifuged (300 x g, PBMCs: 5min and Colo205 cells: 4min) and cell pellets were resuspended carefully with 1ml PBS + 0.04% BSA. Cells were transferred into a 2 ml reaction tube (Eppendorf). The 15 ml reaction tube were rinsed with 500 µl PBS + 0.04% BSA. Cells were washed a second time, followed by a filtering step using a 30 µm strainer (Miltenyi). Cells were counted again (see section 6.4.6) and a cell concentration of $1x10^6$cells/ml was obtained.

For the establishment of scRNA-seq we prepared a mix of Colo205 cells (cultivation procedure in section 6.4.5) and hPBMCs (from Donor1, thawing procedure in section 6.4.2) in a 50:50 ratio (each sample $0.5x10^6$cells/ml were mixed together for total $1x10^6$cells/ml). In addition, a pure PBMC sample (c = $1x10^6$cells/ml) was prepared as well. Splenocytes for the establishment of scRNA-seq were prepared as previously described [51]. Per sample,$1x10^4$ cells were loaded to the Chromium Controller Chip. Pure hPBMCs, Colo205 + hPBMC mix and murine splenocytes were filled alternately onto the chip wells.

Single cells were processed according to 10x Genomics Chromium Single Cell 3´ Reagent Guidelines (v2 chemistry: CG00052 Rev D). Therefore, cells were partitioned into nanoliter-scale gel bead-in-emulsions (GEMs) using 10X Chromium Controller (10x Genomics, Pleasanton, CA, USA). As a result of reverse transcription (RT) reaction using a C1000 touch thermal cycler (Bio-Rad, Munich, Germany), polyadenylated mRNA's from partitioned and

lysed cells were incubated with primers containing (i) an Illumina R1 sequence, (ii) a 16 bp 10X barcode, (iii) a 10 bp unique molecular identifier (UMI) and (iv) a poly-dT primer sequence resulting in barcoded, full-length cDNA. cDNA was purified using silane magnetic beads (Thermo Scientific, Waltham, MA, USA) and amplified by PCR. During library preparation via enzymatic fragmentation, end repair, A-tailing, adaptor ligation, post-ligation cleanup with SPRIselect (Beckman-Coulter Genomics, Krefeld, Germany), index PCR and cleanup, the following sequences were added: P5 and P7 (Illumina, San Diego, CA, USA), a sample index (i7), and R2 (read 2 primer sequence). Quantification of cDNA and final libraries were performed using Qubit dsDNA HS Assay Kit (Life Technologies, Carlsbad, CA, USA) and high-sensitivity DNA chips (Agilent Technologies, Santa Clara, CA, USA). Final libraries were diluted to 250 pM, pooled and sequenced on an Illumina HiSeq4000 platform. Therefore, a HiSeq3000/4000 SBS Kit (150 cycles) and HiSeq3000/4000 PE Cluster Kit (containing also Hiseq3000/4000 PE Flowcell, HiSeq cBot 2 Manifold and Hiseq3000/4000 Accessory Kit from Illumina) were used. For sequencing (Single Index, Paired-end run) the following read lengths were used: 26 cycles (Read 1), 8 cycles (i7 Index) and 98 cycles (Read 2) [21,203].

Establishment scRNA-seq: Comparing 3´scRNA-seq vs. 5´scRNA-seq

For the comparison of different PBMC donors, settings and scRNA-seq methods (3'GE vs. 5`GE), PBMCs were washed as described above and $1x10^4$ cells per sample were loaded to the Chromium Controller Chip in TRON according to the scheme in Figure 9 to create 3'GE libraries later:

Three of the samples (fresh frozen and fresh PBMC's) were loaded on another Chromium Controller in BioNTech to create 5'GE libraries.

At this point we followed the 10x Genomics Chromium Single Cell 3´ Reagent Guidelines (v2 chemistry: CG00052 Rev D) to create 3'GE libraries and the Chromium Single Cell V(D)J Reagent Kits User Guide (chemistry: CG000086, Rev E) to generated 5'GE libraries. The Single Cell V(D)J Reagent Kit differs in the fact that one generates 5' GE libraries from amplified cDNA via a 5' UTR to constant region. For sequencing on the HiSeq4000 all final eleven libraries (eight for 3'GE and three for 5'GE) were diluted to 250 pM, pooled and loaded on an Illumina HiSeq4000 platform as described above with the following read lengths: 26 cycles (Read 1), 8 cycles (i7 Index) and 98 cycles (Read 2).

EAE investigation

FACS sorted antigen-specific CD4+ T cells were filtered with a 40 µm Flowmi™ cell strainer (Bel-Art, Wayne, NJ, USA), centrifuged (300 x g, 5 min) and resuspended with 0.04 % BSA in PBS. Final single cell suspensions were processed according to 10x Genomics chromium single cell 3´ reagent guidelines as described above (v2 chemistry; 10x Genomics, Pleasanton, CA, USA). Final libraries were diluted to 250 pM, pooled and loaded on an Illumina HiSeq4000 platform (Illumina, San Diego, CA, USA) with 2x 150 paired-end kits to achieve a median of ~150,000 reads per cell with the following read lengths: 26 cycles (Read 1), 8 cycles (i7 Index) and 98 cycles (Read 2).

MC38 tumor bearing mice for Ribocytokine investigation

CD45+ sorted TILs with 500 µl PBS + 0.04%BSA in 1.5 ml reaction tube (Eppendorf) were washed once with PBS + 0.04% BSA (centrifugation: 400 x g, 5min at 4°C). Supernatant was removed until slightly less than ~35µl liquid level. Using a 200 µl Rainin pipette (Rainin), the cell pellet was carefully resuspended 15 times (with wide-bore tips). Entire sample volume was aspirated with Rainin pipette (set to 35 µl, normal tips) and filled remaining volume up to 35 µl by aspirating from PBS/0.04%BSA stock tube and put back to original tube. An appropriate sample volume with ~$1.5x10^4$ cells mixed with Nuclease-free water (Thermo Fisher), final volume: 43.2 µl) was loaded on each well of the Single Cell Chromium Controller chip. The following protocol was done according to the 10x Genomics Chromium Single Cell 3´Reagent Guidelines (v3 chemistry with Feature Barcoding technology for Cell Surface Protein: CG000185 Rev C). Here the only difference compared to the protocols before was the creation of barcoded, full-length cDNA and barcoded DNA from the cell surface protein Feature Barcode molecules at the same time due to individual TotalSeq antibodies. Size selection was done using SPRIselect beads (Beckman Coulter, #B23318) to separate amplified cDNA products for 3'GEne Expression and for Cell Surface Protein library preparation. Final libraries for 3'GEne Expression and Cell Surface Protein were diluted to 200 pM, pooled and loaded on two lanes of an Illumina NovaSeq 6000 S2 flowcell. A 1x 100 cycle paired-end kit (Illumina) was used for sequencing. The goal was to achieve around 80,000 reads per cell for Gene Expression libraries and 5,000 reads per cell for Cell Surface Protein libraries with the following read lengths (adapted to the v3 chemistry): 28 cycles (Read 1), 8 cycles (i7 Index) and 91 cycles (Read 2).

6.5 Single cell RNA sequencing data analysis

6.5.1 Raw data processing

Establishment scRNA-seq of human and mouse (EGFP investigation)

For demultiplexing, barcode processing and quantifying transcript counts, the 10x Genomics Cell Ranger single cell pipeline (v2.1.0) was used using the parameters, as described by Zheng et al. 2017 [21]. In brief, demultiplexing of all samples was performed via the Cell Ranger mkfastq pipeline based on the 8 bp read (index read) to generate FASTQs for the Read 1 and Read 2 ends, as well as the 16 bp 10X barcode. Cell Ranger count pipeline was used to align the generated FASTQs to a human reference genome GRCh38 (v3.0.0). For mouse data the reference genome mm10 (v1.2.0) was used. Aligned reads were filtered for valid UMIs (maximal one mismatch allowed) with sequencing quality score > 10 (90 % base accuracy), barcodes with UMI counts that fell within the 99th percentile of the range, to generate gene-barcode matrices, whereas PCR duplicates were removed. Principal component analysis (PCA) was done to consider genes with at least one UMI count in at least on cell, whereas the top 1000 most variable genes were identified based on mean and dispersion. A median between 50,128 and 74,004 reads per cell, respectively, were obtained, for hPBMCs which has been demonstrated to be sufficient for unbiased cell-type classification and states. For mouse data, technical replicates for each group were aggregated via the "Cell Ranger aggr" function and producing a single gene-barcode matrix. Here, a median between 33,122 and 74,004 reads per cell could be determined.

Establishment scRNA-seq: Comparing 3´scRNA-seq vs. 5´scRNA-seq

Demultiplexing was done using 10x Genomics Cell Ranger single cell pipeline (v2.1.0) as described above and FASTQs were mapped to the human reference genome GRCh38 (v3.0.0) [21]. After demultiplexing a 3'GE median between 46,713 and 104,089 reads per cell was achieved for PBMC donor 1 and 2. For Donor 3 a 3'GE median between 289,979 and 465,445 was obtained (lower number of cells per sample: 0.5k – 1k cells loaded). For 5'GE libraries a comparable median between 35,425 and 265,707 was determined for Donor 1, 2 and 3.

EAE investigation

The 10x Genomics Cell Ranger single cell pipeline (v2.1.1) was used in this experiment to perform sample demultiplexing, barcode processing and single cell 3´ gene counting as

described before [21]. Afterwards generated FASTQs were aligned to the mouse mm10 reference genome (v1.2.0) using Cell Ranger count pipeline. A median between 136,148 and 299,158 reads per cell was obtained. Output data of technical replicated for each group were aggregated via "cell ranger aggr" function and were produced a single gene-barcode matrix containing all the data.

Ribocytokines investigation

Demultiplexing, barcode and UMI extraction and counting of single cell 3´ genes was performed by running 10x Genomic Cell Ranger data analysis pipeline (here v3.0.2) as described previously [21]. FASTQs were aligned to a mouse mm10 reference genome (v1.2.0). Sequencing and demultiplexing resulted in a median between 101,158 and 179,308 reads per cell for Gene Expression libraries and between 2,040 and 3,473 reads per cell for Cell Surface Protein libraries.

6.5.2 Single cell sequencing cluster analysis

Establishment scRNA-seq of human and mouse

In this first section, the cluster analysis procedure for hPBMCs and Colo205 cells is summarized. The cluster analysis for mouse splenocytes is described in the section below (EGFP investigation). Demultiplexing was followed by single cell clustering analysis using the Seurat (v2.2.0) R toolkit (with R version 3.4.3.) for quality check of the data, normalization, reduction of high dimensional space, finding cell neighbors by comparing gene expression per cell and visualization of the present single cell populations in clusters using t-SNE (t-stochastic neighbor embedding) or UMAP (Uniform-Manifold-Approximation and Projection) [194,204]. If not further specified, functions were run with default parameters like from the tutorials (https://satijalab.org/seurat/). Briefly, a Seurat object was set up by keeping all genes which were expressed in ≥ 3 cells and keeping all cells with at least 200 detected genes.

For filtering low quality cells (dead cells with low gene content or multiplets with very high number of genes) a threshold was manually chosen. Here, parameters for number of genes, number of UMIs and percentage of present mitochondrial genes per cell set like explained in the tutorial to filter out outliers as shown in the figure and the table with the values in the supplementary (Figure 39).

All cells not passing these criteria's were filtered out, followed by normalization method called "LogNormalize". Gene expression values are normalized for each cell by total expression and multiplied by the scale factor (default: 10,000) to be log-transformed [194]. The identification of variable expressed genes (here the top 2,000) was done using "FindVariableGenes" function to have meaningful data [35].

To remove unwanted sources of variation like batch effects or technical noises and to improve downstream analysis like dimensionality reduction and clustering, "scaling" had to be done. Here the variables nUMIs and percentages of mitochondrial genes were regressed out [194].

A linear dimensional reduction on highly variable genes was done for single cell data to further improve the performance by performing PCA (Principal Component Analysis) on scaled data [35].

Since cell clusters in Seurat are based on PCA scores, hence one has to choose a number of significant PCs with low p-value, which is determined by plotting the distribution of p-values for each PC with a null distribution (dashed line) in a "JackStrawPlot" (supplementary Figure 40).

A second more heuristic plot was examined in parallel called "ElbowPlot". This plot shows standard deviations of the principal components. A cutoff (PC value) can be chosen where there is a clear elbow in the graph (supplementary Figure 41):

After examining these results, the following PC values were chosen for hPBMC/mix samples: 14 (PBMC 100%, replicate 1), 12 (PBMC 100%, replicate 2) and 14 (50/50% PBMCs and Colo205 cells replicate 1 and 2).

To break down cells into clusters the Louvain algorithm was used to construct a sheared-nearest neighbor (SNN) graph based on a cell-cell distance matrix as an input. This was done using the "FindClusters" function in Seurat in which a resolution parameter had to be indicated too. This resolution parameter sets the 'granularity' of the downstream clustering, with lower values leading to less number of clusters. At the end, the following resolution values were selected: 0.8 (PBMC 100%) and 0.6 (50/50% PBMCs and Colo205 cells) [19,194].

Additionally, clustering results are shown in two dimensions using the t-distributed stochastic neighbor embedding (t-SNE) algorithm, which aims to co-localize cells with similar local neighborhoods. For tSNE the same PCs were used as input as above determined [194].

For the identification of DE (differentially expressed) genes, Seurat provides the "FindMarkers" function which finds positive and as well as negative markers of a single cluster compared to all other cells [25]. This function requires the specification of two parameters: "min.pct" to indicate a threshold of a gene to be detected at a minimum percentage in either of the two groups of cells. The second parameter is "thres.use" which describes by which amount a gene should be differentially expressed between two groups. For 100% PBMC and 50/50 PBMC and Colo205 cell mix a threshold of 0.25 was chosen for both parameters [35].

To visualize the expression of genes in cell cluster and to determine present cell types, Seurat provides various tools like "VlnPlot" (shows expression probability distributions across clusters with median and interquartile range), "FeaturePlot" (local gene expression visualization on a tSNE or UMAP plot) or "DoHeatmap" (plotting the expression of genes and cells per cluster). For the heatmap the top 10 expressed genes per cluster were plotted. Following the examination of canonical markers and the top expressed genes per cluster, the present clusters were assigned with a cell type name. The cell numbers per assigned cell cluster could be extracted and the percentages of cell frequencies could be determined as follows: Division of cells in one cluster by the total number of cells multiplied by 100 percent. The before mentioned procedures remained the same for the other experiments if not further specified.

Establishment scRNA-seq: Comparing 3´scRNA-seq vs. 5´scRNA-seq

The analysis was done as described above. All changes are mentioned below in the text. These data were processed using Seurat (v2.3.0) R toolkit (with R version 3.5.0.). The following thresholds were manually chosen to filter low quality cells (values are shown in the table in supplementary Table 18).

After performing PCA, examining the PC values and in addition running the FindClusters function in Seurat, specific parameters (listed in supplementary Table 19) were chosen to processed with RunTSNE etc. [194] For the identification of DE genes by running FindMarkers function the two parameters "min.pct and "thres.use" were run for all datasets with a value of 0.25 [35]. The rest of the clustering analysis was done like mentioned above. For the alignment of PBMC donor 1 and 2 dataset (3'GE protocol) the RunCCA command was used (with 30 dimensions) as well as later the AlignSubspace function using "cca" as reduction type and 18 dimensions. To align the PBMC donor 1 dataset of 3'GE protocol with PBMC donor 1 dataset of 5'GE protocol AlignSubspace function was done with 20 dimensions as well. RunTSNE and

FindClusters function (resolution = 1.8 and dimensions = 18 for alignment of PBMC donor 1 and 2; resolution = 0.8 and dimensions = 20 for alignment of donor 1 3'GE dataset and 5'GE dataset) were run as well afterwards.

EAE investigation

All three datasets were imported into Seurat (v4.0.1) R toolkit (R version 4.1.0.), classified as "RNA" assay for graph-based clustering and visualizations. All functions were run with default parameters, unless specified otherwise as above or described in the tutorials (https://satijalab.org/seurat/).

Low quality cells with less than 200 genes, more than three median absolute deviations (MADs) above the median mitochondrial genes and more than three MADs above the median gene counts were excluded from further analysis.

For the identification of variable expressed genes (here the top 2,000), the updated function FindVariableFeatures was run, here "vst" was chosen as selection.method. The parameter "vst" indicates that top variable genes should fit a line to the relationship of log(variance) and log(mean) using local polynomial regression (loess). Gene values are standardized using the observed mean and expected variance (given by the fitted line). Gene variance is calculated on the standardized values after clipping to a maximum [22,36].

"Integration" strategy was used to identify shared cell states (anchors) with the "FindIntegrationAnchors" function in Seurat (using the first 30 PC's as default) between the three single cell datasets to harmonize them as previously described [36]. Therefore, a list of three Seurat objects were taken as input and identified anchors were used to integrate the three datasets together using "IntegrateData" function (with the first 30 PC's as default). This creates a new assay called "integrated" in the data matrices [36].

Scaling was done for the "RNA" assay as well as for "integrated" assay with default parameters.

In addition to performing PCA (using the first 30 PC's) on scaled data, the PC values were determined as mentioned above. However, beside the above mentioned strategies to identify significant and meaningful PC values, a new calculation was done to determine a maximum and minimum PC value. To determine the maximum PC value, first the percentage of variation associated with each PC was determined as well as the cumulative percentage for each PC. Next, the final PC value was calculated which exhibits a cumulative percentage higher than

80% and which is associated with a percentage of variation less than 5. For the identification of the lower PC value, the difference between variation of PC and subsequent PC was determined in which the change in percentage of variation is more than 0.1 %. The minimum value of both calculation was 16 PC's and taken for further analysis. This calculation was developed by Andrea Cortini from BioNTech, Mainz, Germany.

A Seurat-supported Python package and Non-linear dimension reduction technique called Uniform-Manifold-Approximation and Projection (UMAP) was applied (function: "RunUMAP") to learn manifold of the data and place similar cells together in low-dimensional space prior to graph-based clustering approach for cell cluster identification [68]. This approach is based on shared nearest neighbor (SNN) modularity optimization using the FindNeighbors function and took the top 16 PC's. A two-dimensional non-linear embedding of the cells was generated on the first 16 PC dimensions using UMAP for the visualization. The cells of the different treatment groups can be visualized in a UMAP plot using the "DimPlot" function. Various resolution parameters from 0.0 to 2.0 were plotted using the FindClusters function as described above. After the investigation of canonical marker expression using FeaturePlots, Heatmaps and VlnPlots, a final community resolution parameter of 0.7 could be determined [205].

Via differential expression method, positive and negative markers of a single cluster compared to other clusters or to all other remaining cells were identified applying the "Wilcoxon rank sum test" (default) [22,36].

Assignment of cell type identity to clusters was done by using canonical markers and differentially top expressed genes.

Beside extracting the numbers of cells per cluster but also the cell type numbers per treatment group, the treatment information was extracted by using the "paste" function, which converts its information to character strings and separates and links them by the string e.g. "_" [206]. These new Idents were saved under a new column in the data matrix and can be fetched up using "Idents" function, which can manipulate an object's identity class. Using the "table" function, which builds a contingency table of the counts at each combination of factor levels, an excel csv file was extracted with the numbers of cells per treatment group [206].

Although quality steps as filtering low quality cells were done, a low quality cell cluster mainly defined by a low cytoplasmic transcript fraction (to a large extent ribosome associated genes)

and a small contamination of a DC population and CD8+ T cells were found and excluded for further analysis.

EGFP investigation

For clustering analysis, Seurat (v2.3.4) R toolkit (with R version 3.5.0.) was used and the two datasets (RNA-LPX vaccinated and NaCl control) were aligned [194]. Cells with more than 200 genes, mitochondrial gene content with three median absolute deviation (MADs) below the median and gene counts three MADs below the median were included for further analysis. In total 7,703 genes for the control group and 8,813 genes for the RNA vaccinated group passed the quality filter. Digital gene expression measurements were log-normalized and log-transformed afterwards. The number of detected molecules per cell and the percentage of mitochondrial gene content were regressed out as source of cellular variation for improved downstream dimensional reduction. Both samples were integrated via a canonical correlation analysis (CCA). For cell cluster identification, an algorithm was used based on shared nearest neighbor (SNN) modularity optimization, taking the top 22 correlation components (CCs) and a community resolution parameter of 1.4 into account. The same CCs were used for a two-dimensional non-linear embedding of the cells in a UMAP projection. “Wilcoxon rank sum test” (default) was used for differential expression analysis. Cell type identity to clusters was assigned by using canonical markers and differentially top expressed genes. Several different B cell populations (memory B cells, marginal zone B cells, follicular B cells, transitional B cells, germinal center B cells, cycling B cells and plasma cells) were detected and combined to a single B cell cluster. pDCs were selected based on marker expression (Siglech, Tcf4) by the Seurat lasso tool. For a better resolution, pDCs, monocytes, granulocytes and macrophages were subsetted and reclustered with 14 CCs at resolution 0.6. The main and selected subsetted data sets were merged and processed together as beforehand using 22 CCs and resolution 1.4. All identified cell type identities were retrieved. Despite the use of stringent cut-offs for total gene content, two doublet cell clusters co-expressing myeloid markers with B and T cell markers, respectively (e.g. S100a9 and Cd79a or Cd3d), and one doublet cluster expressing B and T cell markers (e.g. Cd79a and Cd3d) were identified (1.3%) in both groups. Erythrocyte contamination (expressing Hba-a1, Alas2 and Gypa) was less than 0.1%. These doublet and erythrocyte subsets were excluded from further analysis.

Ribocytokine investigation

Seurat (v4.0.1) R toolkit (with R version 4.0.4.) was used for clustering analysis as described above (section EAE investigation) [194]. After filtering low quality cells, digital gene expression measurements and Antibody-Derived-Tags (ADT, feature barcoding data) were used to identify cells with a single ADT barcode (classified as “Singlets”), cells with multiple ADT barcodes (classified as “Doublets”) and cells with no ADT barcode (classified as “Negative”). Cells which were classified as negative were not further taken into account for the analysis since those signals might consist of ambient RNA which was amplified and should not be used a single cell data. Therefore, only cells with more than zero ADT counts were used for downstream analysis using the “subset” function to exclude cells with zero counts. RNA and ADT assay were separately log-normalized and log-transformed afterwards. For the identification of the most variable genes, the FindVariableFeature function was run using the selection method “mean.var.plot”. This method calculates the average expression and dispersion for each gene, divides genes into num.bin (deafult 20) bins and calculates z-scores for dispersion within each bin. For the normalization of ADT data, the method “CLR” was chosen, which applies a centered log ratio transformation. The normalization of the ADT data further requires the “HTODemux” function to assign sample-of-origin for each cell and annotate doublets with a default value of 0.99 for the positive.quantile. The parameter “positive.quantile” describes the inferred 'negative' distribution for each hashtag over which the cell is considered 'positive'.

Only cells classified as “Singlets” and “Doublets” were kept with the subset function because negative labelled cells only show a low amount of total UMI, which would mean that those cells were just empty droplets. Doublets were kept as well because after a quick tSNE analysis they didn’t show typical doublet characteristics since they cluster together with Singlets based on their expression level. Furthermore, from previous investigations they show the same average level of number expressed genes per cell.

Integration strategy and Non-linear dimension reduction technique UMAP were used as described above [36,68].

The two-dimensional non-linear embedding of the cells was generated on the first 18 PC dimensions using UMAP for the visualization. All commands were run on the “integrated” assay so far. A resolution parameter of 1.3 could be determined resulting in 24 cell clusters. To

investigate the DE genes for each of the identity classes in a dataset, the "FindAllMarkers" function was run on the RNA assay. As a result of the investigation of canonical markers and differentially top expressed genes, a low quality cell cluster (low number of genes/UMIs) was identified as well as a cluster with Doublets showing various phenotypes like Fibroblasts (*Sparc* and *Col1a1*), conventional Dentric cells 1 and 2 (cDC1/cDC2; markers: *Cd74*, *Clec9a*), Non-Classical monocytes (Non.-Cl. Mono; marker; *Spn*), Cytotoxic T cells (CTLs; markers: *Gzmb* and *Perf1*), CD8+ and CD4+ T cells. Furthermore, a mast cell population (*Cpa3* and *Gata2*) was detected. All these populations were excluded from the analysis since the focus was on T cells and Myeloids resulting 10 final cell types: CD4+ T cells, CD8+ Late Effector, CD8+ Proliferating CTL, yd T cell (Gamma delta) which clustered together with CD8+ CTL, yd T cells mix with NK cells (Natural killer cells), B cells, Cl. Monocytes, Mature cDCs, Neutrophils, Macrophages and Proliferating Macrophages.

To improve the resolution of the identified cell cluster especially the mix of yd T cells/NK cells or yd T cells/CD8+ T cells etc. all lymphocyte clusters were analyzed separately from the myleoids to increase the chance that the algorithm is able to identify single populations (not mixed) based on a more similar gene expression which is e.g. for the lymphocytes not interfered from the DE genes of the myeloids. Therefore, a "Subclustering" strategy was applied (section 6.5.4) followed by "Merging" of more resolved lymphocyte and myeloid data (section 6.5.5). This was done to combine both cell data sets with all treatment groups for better downstream analysis (determination of cell type frequencies, investigation of local gene expression in clusters using FeaturePlots, VlnPlots and Heatmaps as well as assessing DE genes between cell clusters and treatment groups).

6.5.3 Differential gene expression analysis (DGEA)

First, a dataset has to be loaded either with only the cell type information or with the addition of the treatment information.

To extract the differential expressed genes between cell clusters or even treatment groups the "AverageExpression" function needed to be run, which determines the expression for an average single cell in each identity class and was saved under a new Seurat object name. Running the FindMarkers function on the RNA assay a differential gene expression analysis could be done between cell type or treatment groups or both (indication by ident.1 and ident.2). Identified DE genes can be extracted as excel csv file and be plotted in various ways like FeaturePlots, VlnPlots or Heatmaps.

6.5.4 Subclustering of single cell RNA data

Ribocytokine investigation

To isolate the lymphocyte from the main dataset and to analyze them independently from myeloids, the "subset" function was used. The following lymphocyte clusters were saved under a new output name and used for further cell clustering: "CD8+ Late Effector","yd T cells/NK cells", "yd T cells/CD8+ CTL", "CD8+ Prolif. CTL", "CD4+ T cells", "B cells". To ensure a clean lymphocyte dataset, all present lymphocytes which showed a gene expression of CD14 and CD68 > 0 counts (doublets) were excluded for this subclustering. There was no need to do a subclustering of myeloids since these cell populations were not clustered as a mix of different cell subsets. As a consequence of subclustering for lymphocytes, the following Seurat functions needed to be re-run again for a new clustering: FindVariableFeature, Scaling, RunPCA (Identification of a new PC value), FindNeighbors, RunUMAP and FindClusters to identify a meaningful resolution parameter. To be able to merge this subclustered lymphocyte dataset later with the original dataset (section 6.5.5) an integration step was done again to combine all treatment groups of the present lymphocytes. For this preparation all Cell Ids of each treatment groups were saved in the cell matrices under "Clusters". Afterwards, the FindVariableFeature function was run on each of the four treatment groups of lymphcoytes with selection.method = "vst", followed by the Integration step (functions: FindIntegratedAnchors and IntegrateData). Scaling was done again for RNA and integrated assay and RunPCA function was run as well as RunUMAP and FindNeighbors on the top 21 PCs. The rest of the procedure to identify makers and cell types was repeated like mentioned in the previous section 6.5.2. This resulted in 16 clusters (determined resolution: 2.0) from which cell types were combined showing a similar gene expression profile ending up with a better resolved clustering to separate especially NK cells from yd T cells. The following CD45+ TILs could be identified: NK cells, yd T cells, Memory T cells, CD4+ T cells, CD8+ cycling T cells, CD8+ IFN hi T cells (IFN-response genes were highly expressed here) and CD8+ Late Effector (expression of exhaustion markers like *Lag3*, *Pdcd1*, *Cd200*, *Tox* and effector markers at the same time: *Gzmb*, *Prf1*, *Nkg7* etc.).

EAE investigation

For a better resolved clustering of the Th1/Th17 cells, a Subclustering was separately done for cluster 4-6. Each of them was isolate from the main dataset using subset function and cell IDs were saved with Idents function. Re-clustering was done as mentioned above. The RunUMAP

and FindNeighbors function was ran on the top 16 PCs for Th1/Th17 cells. Final resolution for Th1/Th17 cells was 1.5.

6.5.5 Merging of single cell RNA data

Ribocytokine investigation

To prepare the merging of the resolved and subclustered lymphocytes (section 6.5.4) the main data set was loaded and previous lymphocyte clusters were all selected with the "CellSelector" function in Seurat, labeled with "Idents" function under new name (Old lymphocytes) and were excluded from the analysis via the subset function. Subsequently, the resolved subclustered lymphocyte populations were loaded and cell IDs were saved using Idents function under the column "Clusters" in the data matrix. The final step was to combine both Seurat objects (main dataset containing only the myeloids at this point and the resolved subclustered lymphcoytes) using the "merge" function. For a new clustering of the merged objects the following procedures had to be re-run: Normalization of RNA and integrated assay, FindVariableFeatures (selection.method = "mean.var.plot"), HTODemux, Scaling on RNA and integrated assay, RunPCA, RunUMAP and FindNeighbors on the top18 PC's. There was no need to define a new resolution parameter since saved Cell ID's could be retrieved back under "CellType" column using Idents function. The final merged UMAP was plotted with DimPlot function (indicating "CellType" under the parameter "group.by). B cells were excluded from the analysis since the focus was on T cells and myeloids. Following an intense cell cluster investigation of the preliminary assigned "Memory T cell" population, there was no expression of Cd8 and Cd4 in these cells, which was the reason why this cluster was assigned as Double negative T cells (DN T cells) using the "Rename" function in Seurat. The rest of the downstream analysis was the same as mentioned above in section 6.5.2 to determine the cell type frequencies and DE genes between the four treatment groups. Afterwards, cell ID transfer was done as described in the next section 6.5.6.

6.5.6 Transferring of cell identification barcodes

EAE investigation

First, the original main dataset was loaded (no subclustering) as well as the subclustered and final assigned datasets. The cluster names of the original main dataset and the subclustered final datasets had to be retrieved as a character vector. This was done using the "as.character" function which depletes characters including names and saved under a new Seurat object called

e.g. “idents.vector”. The column names of the datasets were saved in the created vector using the “colnames” function. These cluster names were replaced in the original main dataset with the cluster names from the subclustered datasets using the as.factor function, which converts a column into a factor column. By default idents are clusters but identities were put in a new column of meta.data called “Clusters”. The original UMAP shape with the new cluster identities of the subclustered data can be visualized by running the DimPlot function. All other features and investigations mentioned in section 6.5.2 like checking DE genes or determination of cell type frequencies was done.

6.5.7 Gene set enrichment

EAE investigation

Enrichment of selected marker genes for respective T cell subsets was assessed via a non-parametric and unsupervised gsva function from the R-package Gene Set Variation Analysis (GSVA) software using the ssgsea method [26].

EGFP investigation

Enrichment of selected gene sets from MSigDB v5.2 (REACTOME, HALLMARK and GO; human gene IDs were converted to mouse genes) for all differentially expressed genes in respective, averaged cell subsets was assessed using unsupervised ssGSEA method implemented in R package gsva (GSVA, v1.26.0) [26]. A minimal gene cut-off of at least 12 genes was chosen for REACTOME, and HALLMARK and at least 60 genes had to be enriched for GO gene sets. From 46 identified gene sets, the 20 most relevant gene sets were selected and plotted using the heatmap.2 function.

7 SUPPLEMENTARY

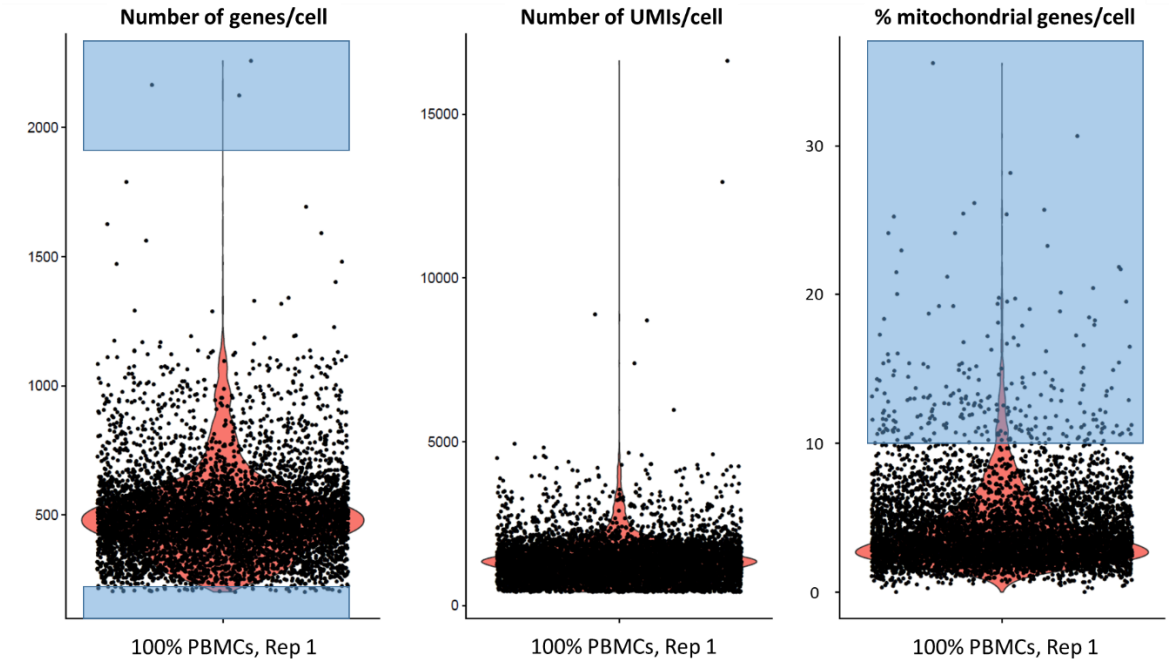


Figure 39: Visualization of number of genes, UMIs per cell and percentages of mitochondrial genes via violin plots for sample 100% PBMCs replicate 1. Cells are marked in blue were excluded from the analysis for being considered low quality cells.

Table 17: Filtering parameters for pure PBMCs and mix of PBMCs+Colo205 cells.

	Number of genes/cell		Percentages mitchondrial genes	
Sample	**low threshold**	**high threshold**	**low threshold**	**high threshold**
100% PBMCs	200	1900	-	10
50/50% PBMCs Colo205	200	6500	-	17

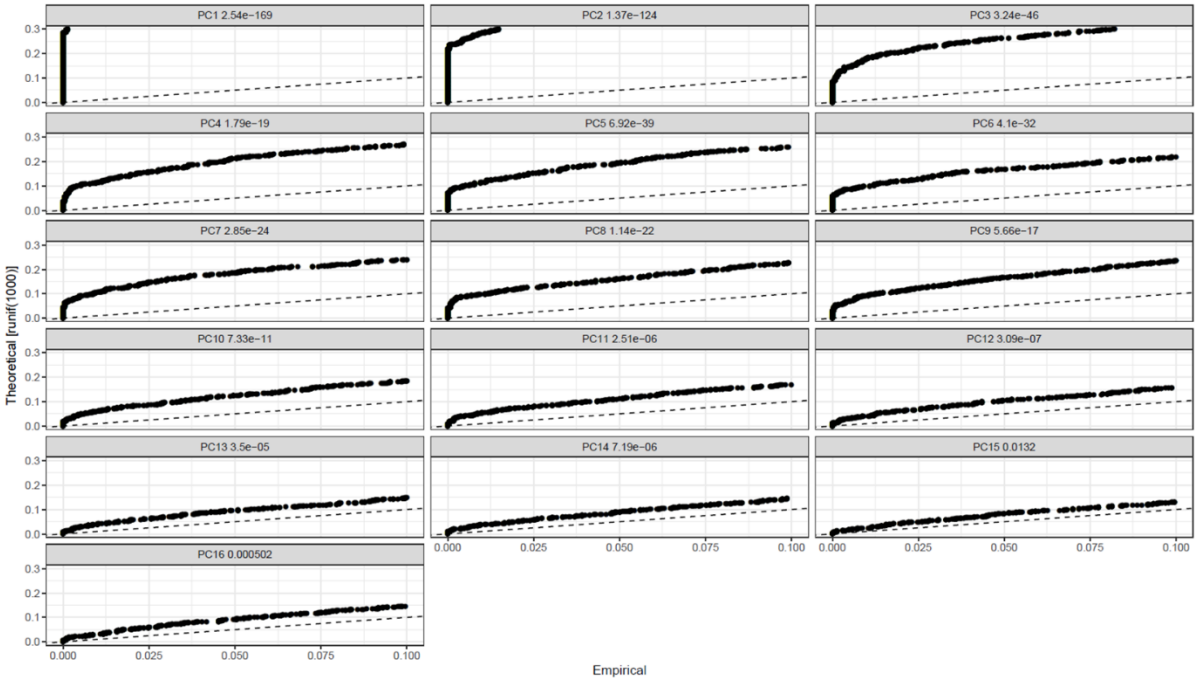


Figure 40: Visualization of significant PC's in JackStrawPlot for sample pure PBMCs replicate 1. Significant PCs show a p-value distribution (black curve) that is strongly skewed to the left compared to the null distribution (dashed line). The p-value for overall significance for each PC is based on a proportion test comparing the number of genes with a p-value below a particular threshold compared with the proportion of genes expected under a uniform distribution of p-values.

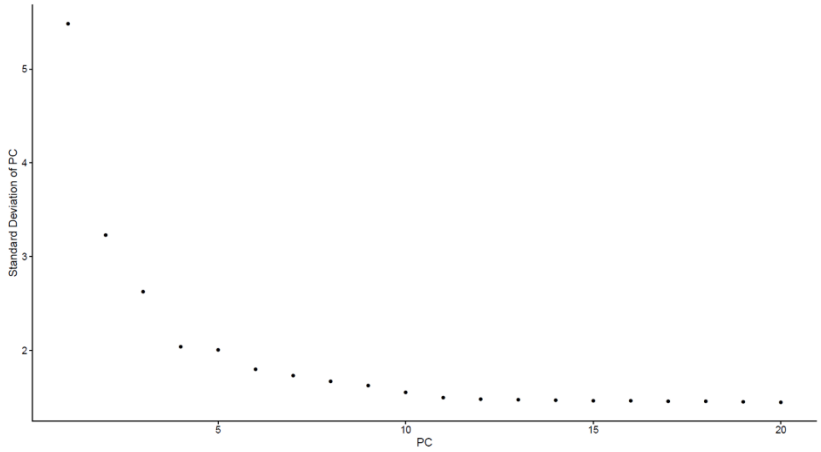


Figure 41: Visualization of standard deviations of the principle components in ElbowPlot for sample pure PBMCs replicate 1.

Table 18: Filtering parameters of low quality cells for fresh frozen and fresh PBMCs (different donors).

		Number of genes/cell		Percentages mitchondiral genes	
Sequencing	**Sample ID**	**low threshold**	**high threshold**	**low threshold**	**high threshold**
3´scRNAseq	1_FF PBMC, Donor 1, Rep1	200	2500	-	0.16
	2_FF PBMC, Donor 1, Rep2	200	6500	-	0.16
	3_FF PBMC, Donor 1, Rep3	200	2600	-	0.18
	4_FF PBMC, Donor 1, Rep4	200	2500	-	0.20
	5_Fresh PBMC, Donor 2	200	2500	-	0.22
	6_Fresh PBMC, Donor 3, Rep1	200	2600	-	0.07
	7_Fresh PBMC, Donor 3, Rep2	200	2600	-	0.05
	8_Fresh PBMC, Donor 3, Rep3	200	2600	-	0.07
5´scRNAseq	9_FF PBMC, Donor 1, Rep1	200	2500	-	0.19
	10_FF PBMC, Donor 2	200	3500	-	0.19
	11_Fresh PBMC, Donor 3, Rep1	200	3200	-	0.09

Table 19: Chosen PC and resolution values for fresh frozen and fresh PBMCs (different donors).

Sequencing	**Sample ID**	**PC's used**	**resolution**
3´scRNAseq	1_FF PBMC, Donor 1, Rep1	15	0.8
	2_FF PBMC, Donor 1, Rep2	12	0.8
	3_FF PBMC, Donor 1, Rep3	11	0.8
	4_FF PBMC, Donor 1, Rep4	12	0.8
	5_Fresh PBMC, Donor 2	15	1.2
	6_Fresh PBMC, Donor 3, Rep1	10	1.2
	7_Fresh PBMC, Donor 3, Rep2	10	1.2
	8_Fresh PBMC, Donor 3, Rep3	10	1.6
5´scRNAseq	9_FF PBMC, Donor 1, Rep1	16	0.8
	10_FF PBMC, Donor 2	14	0.8
	11_Fresh PBMC, Donor 3, Rep1	10	2

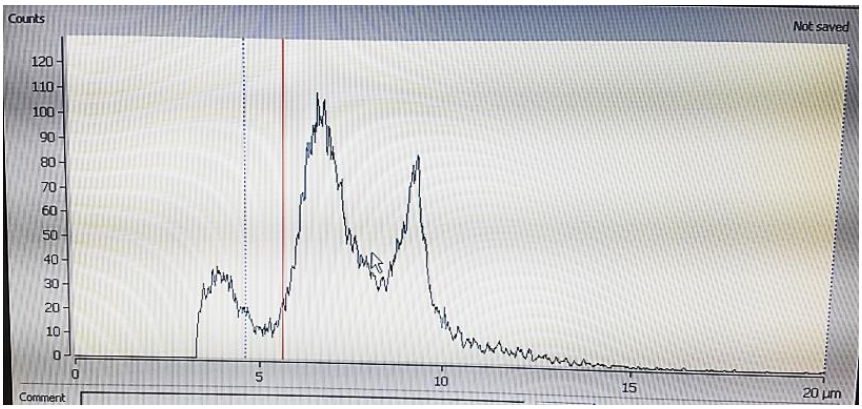
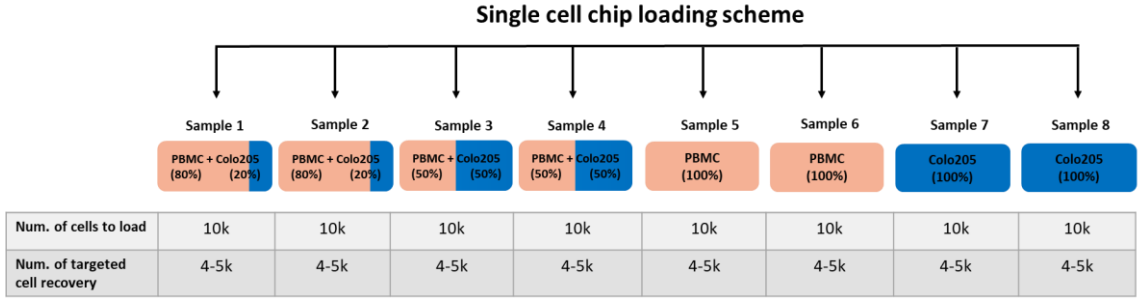


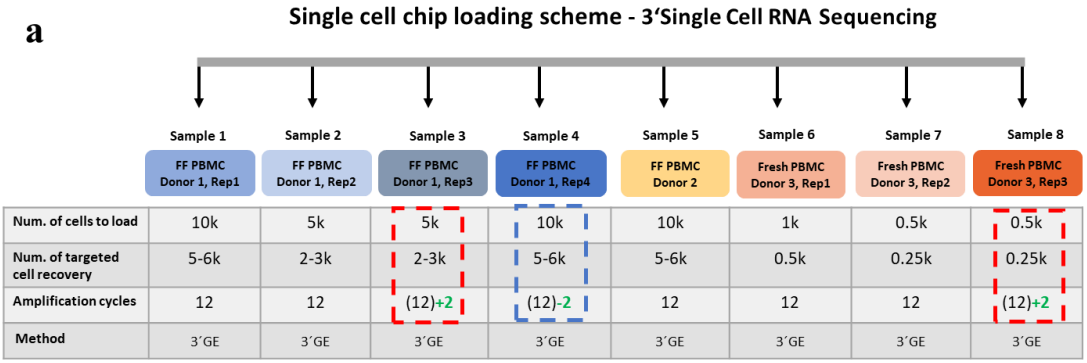
Figure 42: Determination of PBMC number and viability by monitoring cell size distribution with the CASY system. Debris: The peak on the left of the blue; dead cells: valley between blue and red cursor and living cells: Peaks in the right of the red cursor. X-axis: cell size in µm and y-axis: cell counts per ml.



Num. of cells to load	10k	10k	10k	10k	10k	10k	10k	10k
Num. of targeted cell recovery	4-5k	4-5k	4-5k	4-5k	4-5k	4-5k	4-5k	4-5k

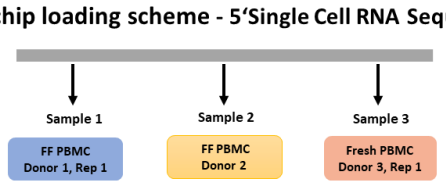
Figure 43: Chromium Controller Chip loading scheme of human and murine samples. Human samples (pure hPBMCs or mixture with tumor cell line Colo205 cells) and mice samples (splenocytes) were loaded one by one on the single cell chip. Cell input was 10,000 cells for all samples.

a



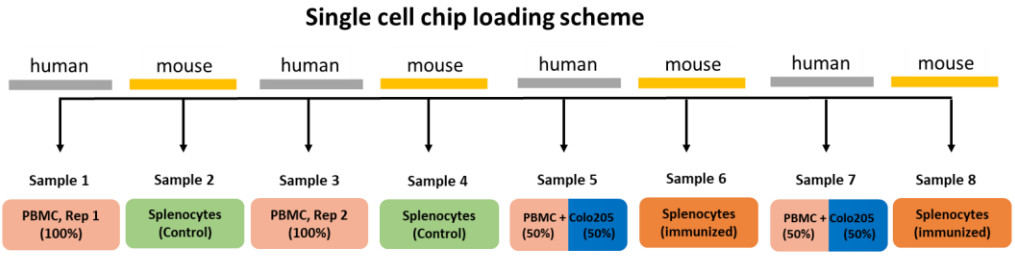
Num. of cells to load	10k	5k	5k	10k	10k	1k	0.5k	0.5k
Num. of targeted cell recovery	5-6k	2-3k	2-3k	5-6k	5-6k	0.5k	0.25k	0.25k
Amplification cycles	12	12	(12)+2	(12)-2	12	12	12	(12)+2
Method	3´GE	3´GE	3´GE	3´GE	3´GE	3´GE	3´GE	3´GE

b



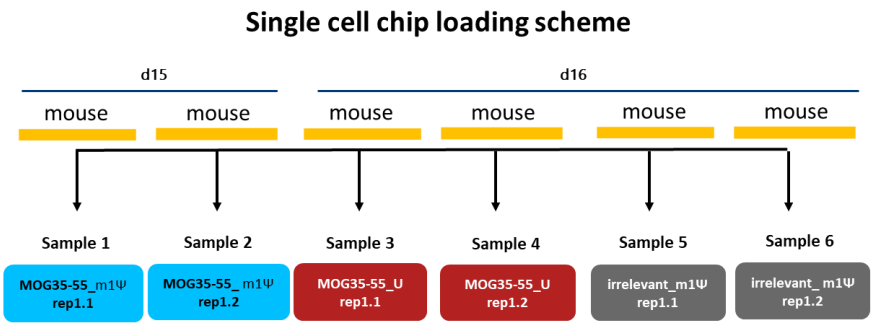
Num. of cells to load	10k	10k	1k
Num. of targeted cell recovery	5-6k	5-6k	0.5k
Amplification cycle	14	14	14
Method	5´GE	5´GE	5´GE

Figure 44: Chromium Controller Chip loading scheme for three different PBMC donors (3’GE and 5’GE protocol). One time fresh prepared PBMC sample and two times fresh frozen PBMC sample to create 3’GE and 5`GE libraries. Various settings like different amplification cycles and number of loaded cells should be compared. Chromium controller loading scheme for (**a**) 3’GE protocol and (**b**) for 5’GE protocol. Highlighted samples in frames, cycle number was increased or decreased for amplification.



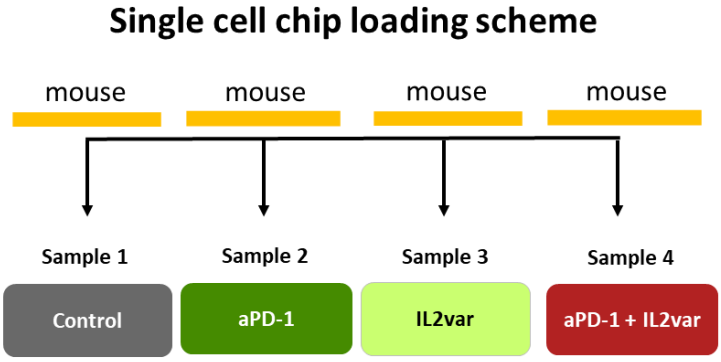
Num. of cells to load	10k	10k	10k	10k	10k	10k	10k	10k
Num. of targeted cell recovery	4-5k	4-5k	4-5k	4-5k	4-5k	4-5k	4-5k	4-5k

Figure 45: Chromium Controller Chip loading scheme of human and murine samples. Human samples (pure hPBMCs or mixture with tumor cell line Colo205 cells) and mice samples (splenocytes) were loaded one by one on the single cell chip. Cell input was 10,000 cells for alle samples.



Num. of cells loaded	?	?	?	?	?	?
Cell recovery	~56%	~56%	~56%	~56%	~56%	~56%
Amplification cycles	12	12	14	14	14	14
Recovered cell number	1.690	1.494	1.185	1.291	1.299	1.309

Figure 46: Chromium Controller Chip loading scheme of murine samples. Mice samples (splenocytes) with different treatments were loaded on the single cell chip. Cell input was not possible to determine after filtering step. All three treatment groups were loaded as duplicates. Sample 1 and 2 (MOG35_m1Ψ Rep1 and Rep2) were processed one day earlier.



Num. of cells loaded	15k	15k	15k	15k
Cell recovery	6-7k	6-7k	6-7k	6-7k

Figure 47: Chromium Controller Chip loading scheme of murine samples. Mice samples (CD45+ immune cells) with different treatments (1: Control which was rat Ig2a and human Albumin LNP-formulated, 2: aPD-1, 3: IL2var and 4: combination treatment aPD-1 + IL2var) were loaded on the single cell chip. Cell input was 15,000 cells for each sample.

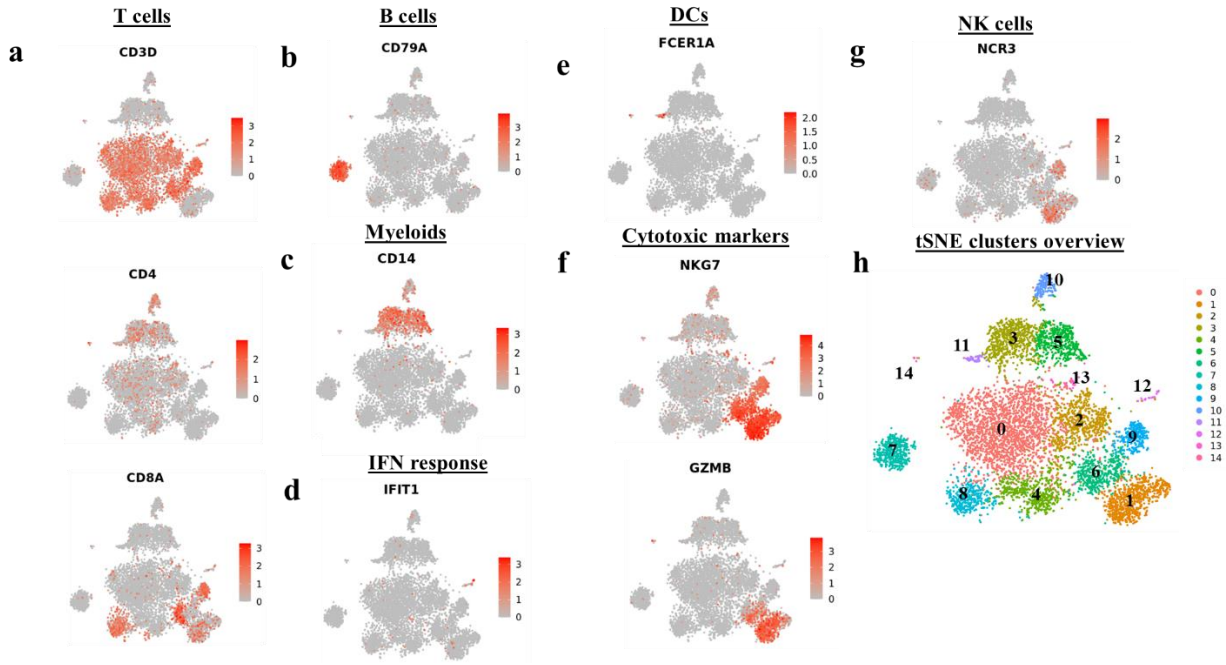


Figure 48: Identification of cell subsets using the human PBMC datasets 3´scRNA-seq versus 5´scRNA-seq. Selected marker expression as features visualized on UMAP dimension reduction plot (called FeaturePlot). A gene which is higher expressed in a cell is colored in darker red. Various representative markers were plotted for the detection of (**a**) T cells, (**b**) B cells, (**c**) myeloid/monocytes cells, (**d**) IFN response cells, (**e**) DCs, (**f**) cytotoxic markers, (**g**) NK cells and (**h**) an overview of all identified cell clusters by Seurat[dimension: 20; resolution: 0.8].

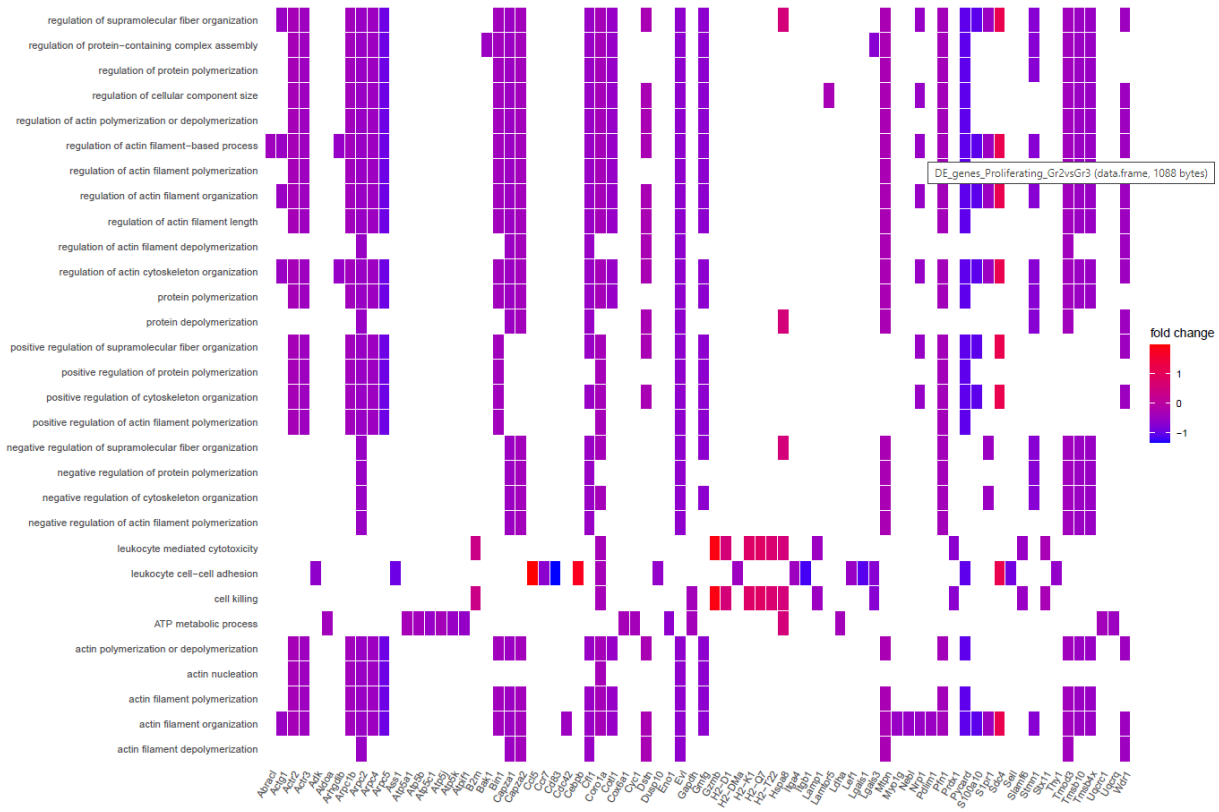


Figure 49: Heatmap of pathway analysis results for Effector Treg cells (comparison: MOG35_55_U against irrelevant_m1Ψ) of EAE mouse dataset (by clusterProfiler4.0 algorithm) The heatmap display expression patterns for effector Treg cells and features of the respected comparison: All genes which are involved in these significant terms are depicted on the x-axis, while the enriched pathways are plotted on the y-axis. Genes, which were found to be involved in an enriched pathway are colored as follow: More pink means a higher fold change (upregulated), whereas more blue means a lower fold change (downregulated gene).

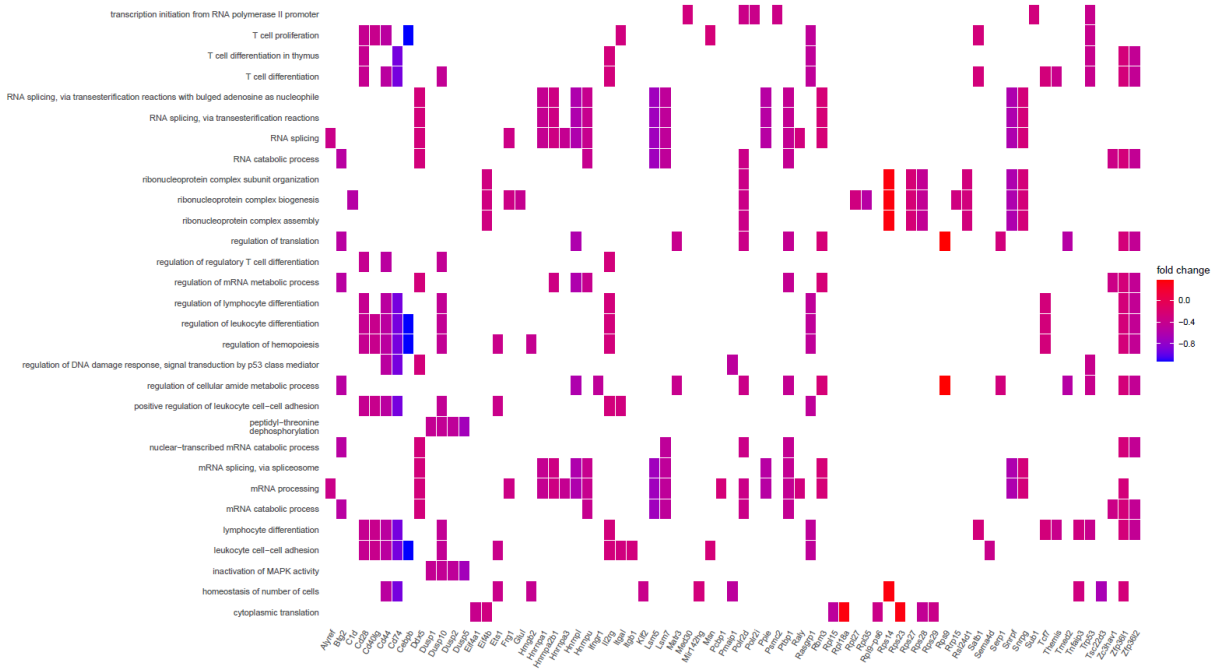


Figure 50: Heatmap of pathway analysis results for Th1/Th17 cells (comparison: MOG35_55_m1Ψ against MOG35_55_U) of EAE mouse dataset (by clusterProfiler4.0 algorithm) The heatmap display expression patterns for Th1/Th17 cells and features of the respected comparison: All genes which are involved in these significant terms are depicted on the x-axis, while the enriched pathways are plotted on the y-axis. Genes, which were found to be involved in an enriched pathway are colored as follow: More pink means a higher fold change (upregulated), whereas more blue means a lower fold change (downregulated gene).

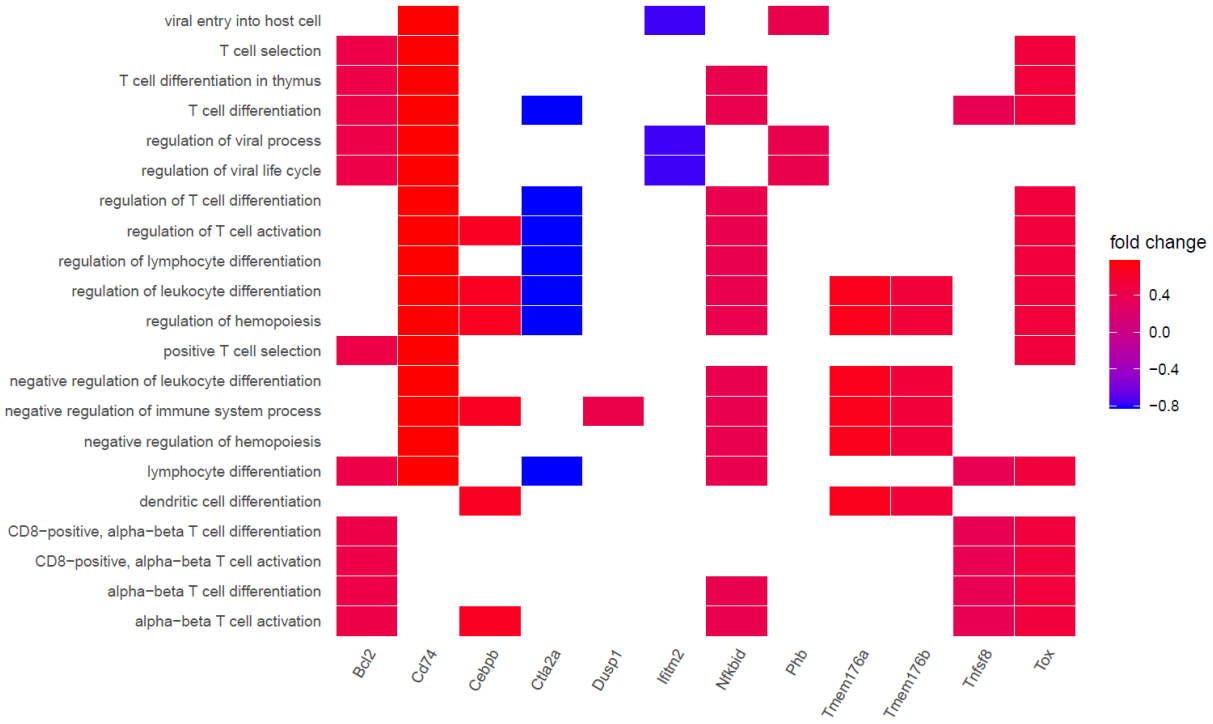


Figure 51: Heatmap of pathway analysis results for Th1/Th17 cells (comparison: MOG35_55_U against irrelevant_m1Ψ) of EAE mouse dataset (by clusterProfiler4.0 algorithm) The heatmap display expression patterns for Th1/Th17 cells and features of the respected comparison: All genes which are involved in these significant terms are depicted on the x-axis, while the enriched pathways are plotted on the y-axis. Genes, which were found to be involved in an enriched pathway are colored as follow: More pink means a higher fold change (upregulated), whereas more blue means a lower fold change (downregulated gene).

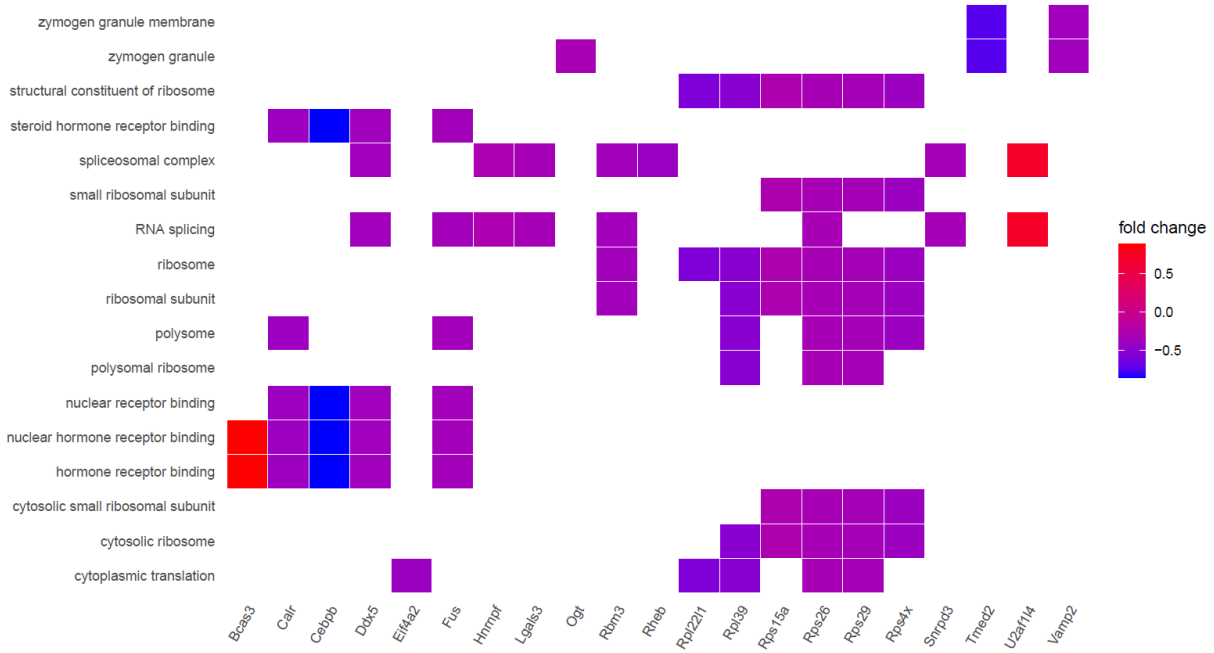


Figure 52: Heatmap of pathway analysis results for Exhausted cells (comparison: MOG35_55_m1Ψ against MOG35_55_U) of EAE mouse dataset (by clusterProfiler4.0 algorithm) The heatmap display expression patterns for exhausted cells and features of the respected comparison: All genes which are involved in these significant terms are depicted on the x-axis, while the enriched pathways are plotted on the y-axis. Genes, which were found to be involved in an enriched pathway are colored as follow: More pink means a higher fold change (upregulated), whereas more blue means a lower fold change (downregulated gene).

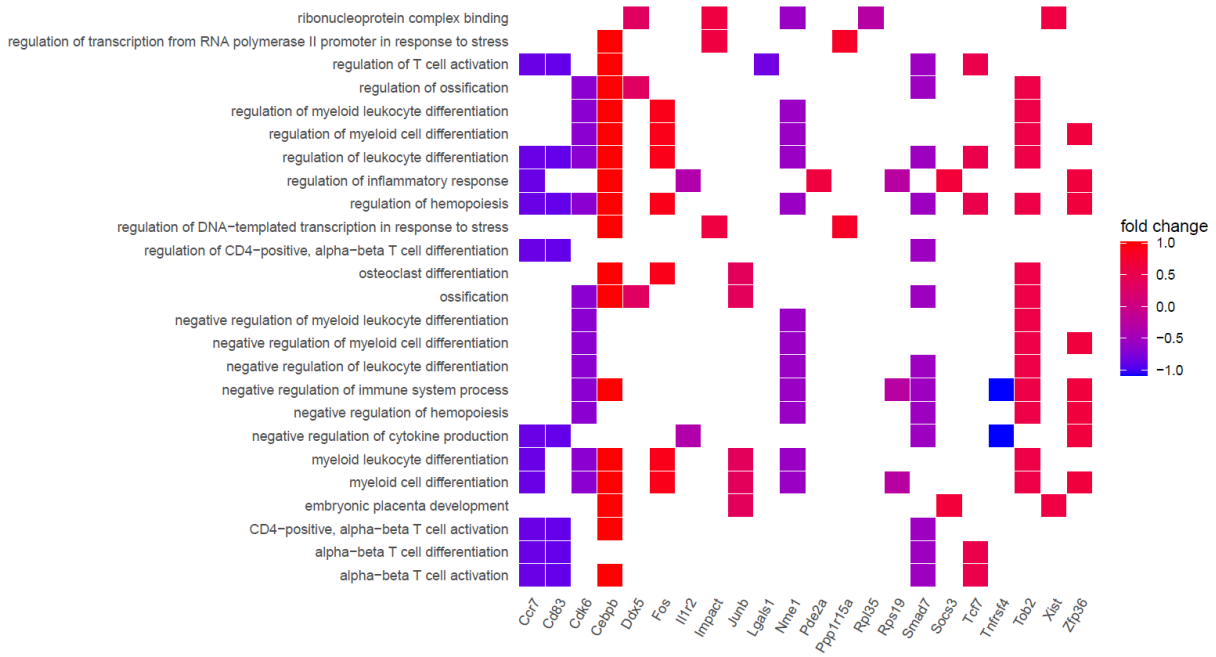


Figure 53: Heatmap of pathway analysis results for Exhausted cells (comparison: MOG35_55_U against irrelevant_m1Ψ) of EAE mouse dataset (by clusterProfiler4.0 algorithm) The heatmap display expression patterns for exhausted cells and features of the respected comparison: All genes which are involved in these significant terms are depicted on the x-axis, while the enriched pathways are plotted on the y-axis. Genes, which were found to be involved in an enriched pathway are colored as follow: More pink means a higher fold change (upregulated), whereas more blue means a lower fold change (downregulated gene).

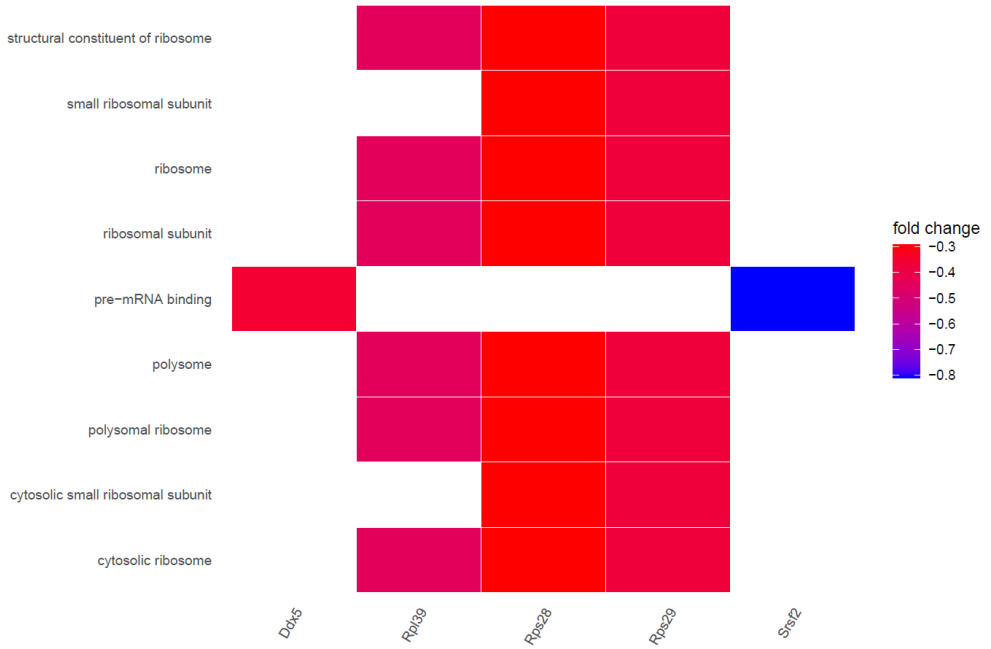


Figure 54: Heatmap of pathway analysis results for Memory cells (comparison: MOG35_55_m1Ψ against MOG35_55_U) of EAE mouse dataset (by clusterProfiler4.0 algorithm) The heatmap display expression patterns for memory cells and features of the respected comparison: All genes which are involved in these significant terms are depicted on the x-axis, while the enriched pathways are plotted on the y-axis. Genes, which were found to be involved in an enriched pathway are colored as follow: More pink means a higher fold change (upregulated), whereas more blue means a lower fold change (downregulated gene).

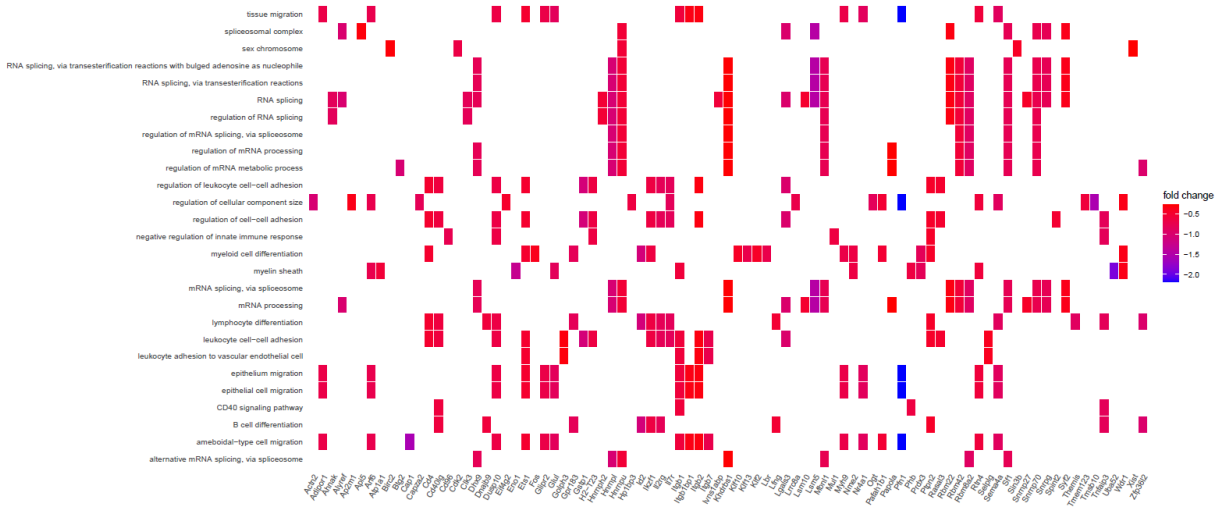


Figure 55: Heatmap of pathway analysis results for Proliferating cells (comparison: MOG35_55_m1Ψ against MOG35_55_U) of EAE mouse dataset (by clusterProfiler4.0 algorithm) The heatmap display expression patterns for memory cells and features of the respected comparison: All genes which are involved in these significant terms are depicted on the x-axis, while the enriched pathways are plotted on the y-axis. Genes, which were found to be involved in an enriched pathway are colored as follow: More pink means a higher fold change (upregulated), whereas more blue means a lower fold change (downregulated gene).

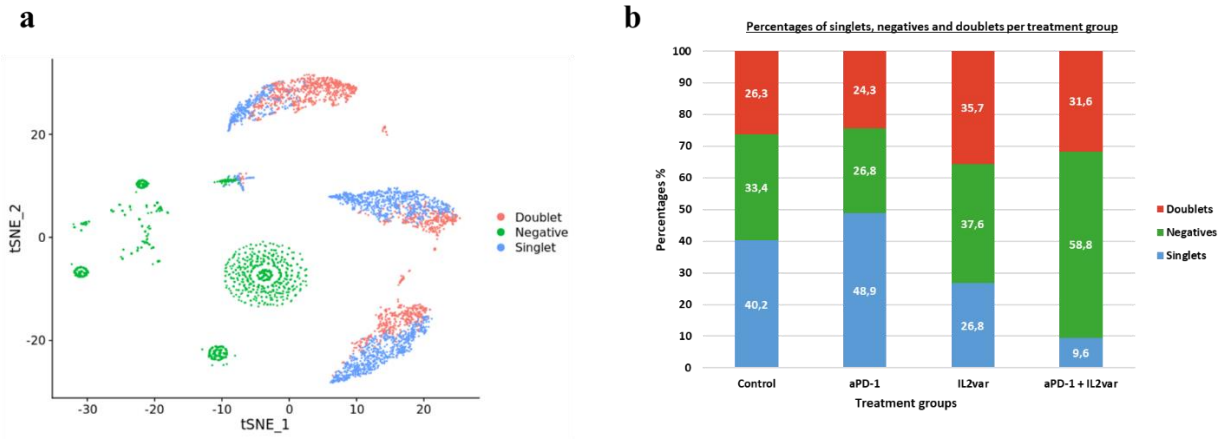


Figure 56: Overview percentage of identified singlets, doublets and negative labeled cells of CD45+ murine samples Determination of percentages was done for all four treatment groups (Control, aPD-1, IL2var and aPD-1 + IL2var). The figure illustrated the identified singlets, doublets and negatives in UMAP projection as well as the percentages for each of the four treatment groups.

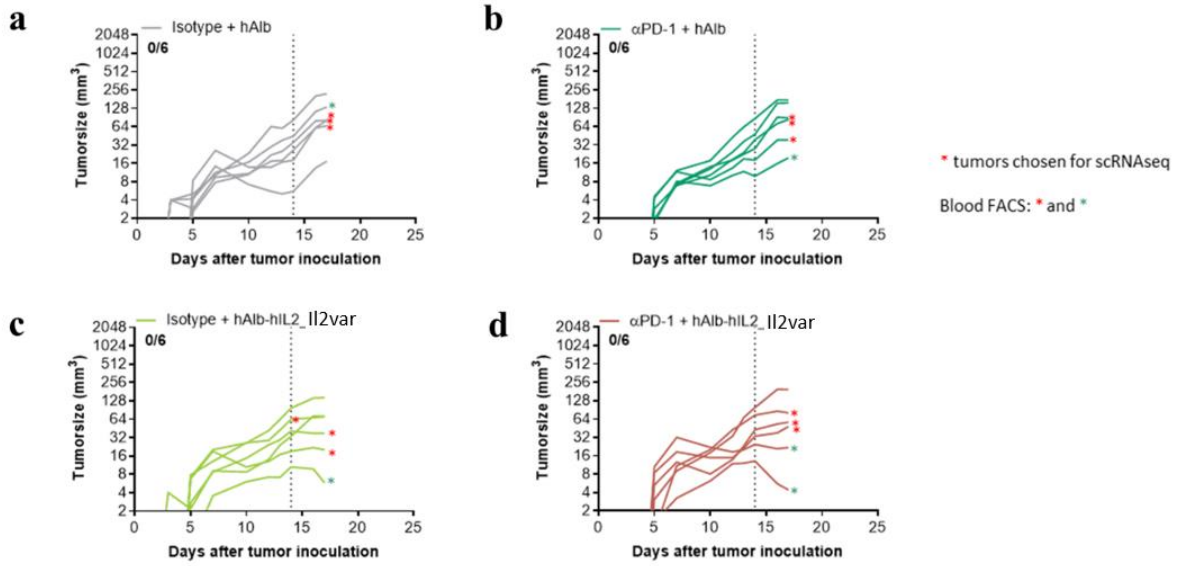


Figure 57: MC38 tumor growth curves for all four treatment groups. Tumor growth curve shows elevation of tumor size during 17 days in (**a**) control, (**b**) aPD-1, (**c**) IL2var and (**d**) aPD-1 + IL2var. Tumor volume (mm3) was measured every 5 days. The dashed line at day 14 shows the time of single treatment. Each experimental group consisted of 6 mice. The tumors with the curve trajectories that are marked with a red star were selected for scRNA-seq. The stars marked with green were additionally examined with FACS.

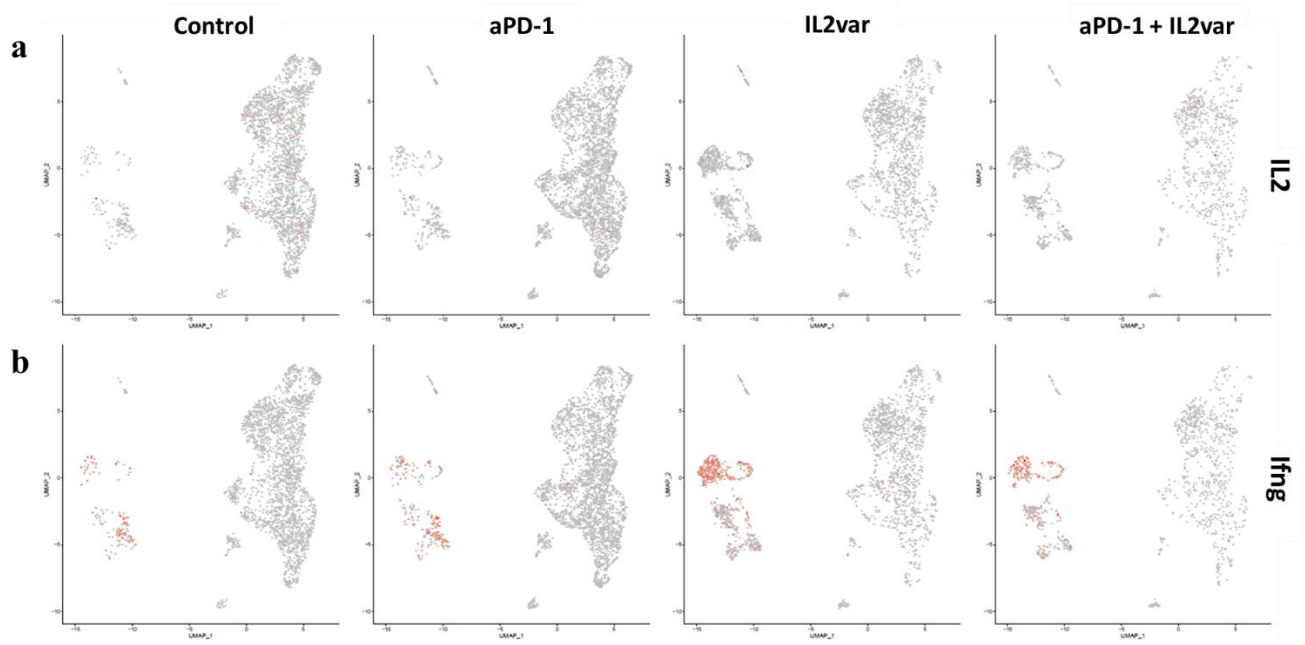


Figure 58: Featureplot local IL2 and Ifng gene expression in cells of CD45+ murine cells dataset (groupwise splitted). Selected marker expression visualized in a different kind of UMAP called Featureplot. Each dot here is a single cell. A gene which is higher expressed in a cell is colored in darker red. Plots are separated for Control, aPD-1, IL2var and aPD-1 + IL2var treatment.

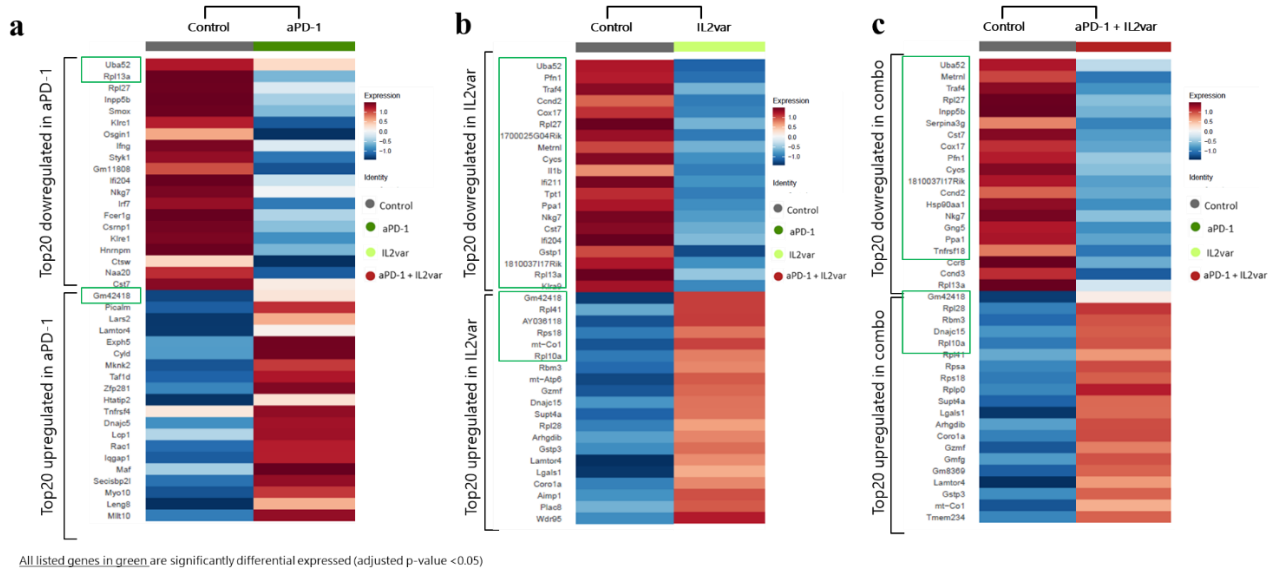


Figure 59: Illustration of differential expressed genes in heatmap for combined CD8+ T cells between treatment groups (by Seurat analysis tool). The heatmap display the expression of 20 top downregulated or upregulated genes for combined CD8+ T cells revealed after the comparison between the treatment groups: (**a**) Control against aPD-1, (**b**) Control against IL2var and (**c**) Control against aPD-1 + IL2var. Genes marked in a green frame are significant expressed (p value < 0.05). For combined CD8+ T cells, treatment groups are plotted on the x-axis (Control: grey; aPD-1: dark green; IL2var: light green and aPD-1 + IL2var: red), while the differential expressed genes are plotted on the y-axis (based on their adjusted p-value). Genes, which are higher expressed, are colored more (dark) red, genes which are downregulated are colored more (dark) blue.

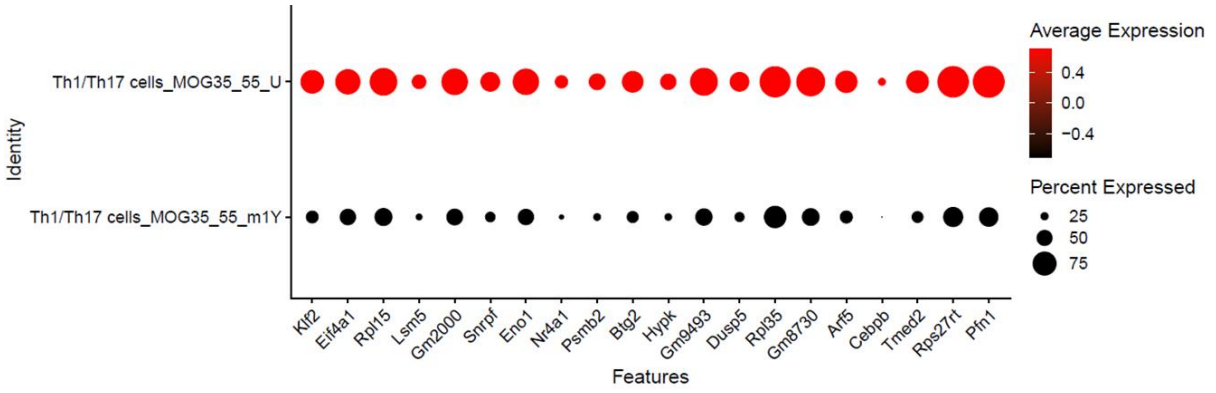


Figure 60: Expression of 20 upregulated genes in Th1/Th17 cells in response to MOG35-55_U treatment. The dotplot shows the top 20 differentially expressed genes on the x-axis Th1/Th17 cells, split by treatment group of cells on the y-axis (illustrated by Seurat analysis tool). The genes are ordered according to indicated function. The size of the dot features in how many cells a gene is expressed [in %] and the color code illustrates the expression level.

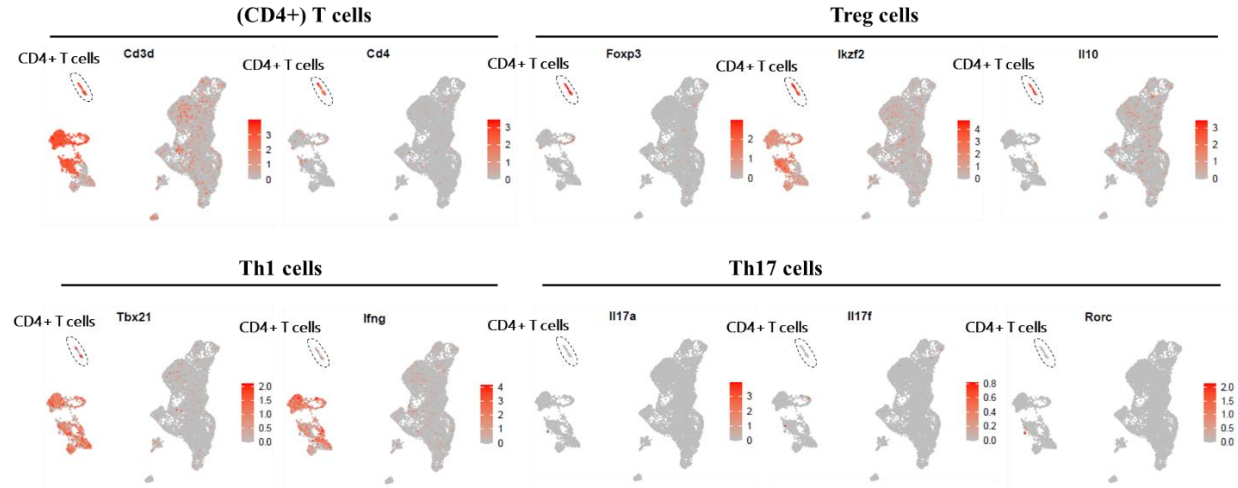


Figure 61: Featureplot local T cell gene expression in cells of CD45+ cells dataset. Selected marker expression for Th1 cell-, Th17 cell- and Treg cell subsets visualized in a different kind of UMAP called Featureplot. Each dot here is a single cell. A gene which is higher expressed in a cell is colored in darker red.

8 LITERATURE

1 *Alessandra Dal Molin, Barbara Di Camillo*. How to design a single-cell RNA-sequencing experiment: pitfalls, challenges and perspectives. Oxford University Press, Briefings in Bioinformatics 2018: 1–11

2 *van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C*. Ten years of next-generation sequencing technology. Trends in genetics TIG 2014; 30(9): 418–426

3 *J. Craig Venter et al.* The Sequence of the Human Genome. SCIENCE 2001(Vol 291): 1304–1351

4 *Li X, Wang C-Y*. From bulk, single-cell to spatial RNA sequencing. International journal of oral science 2021; 13(1): 36

5 *Kuksin M, Morel D, Aglave M et al.* Applications of single-cell and bulk RNA sequencing in onco-immunology. European journal of cancer (Oxford, England 1990) 2021; 149: 193–210

6 *Tang F, Barbacioru C, Wang Y et al.* mRNA-Seq whole-transcriptome analysis of a single cell. Nature methods 2009; 6(5): 377–382

7 *Stubbington MJT, Lönnberg T, Proserpio V et al.* T cell fate and clonality inference from single-cell transcriptomes. Nature methods 2016; 13(4): 329–332

8 *Wagner A, Regev A, Yosef N*. Revealing the vectors of cellular identity with single-cell genomics. Nature biotechnology 2016; 34(11): 1145–1160

9 *Wang Y, Navin NE*. Advances and applications of single-cell sequencing technologies. Molecular cell 2015; 58(4): 598–609

10 *Nguyen A, Khoo WH, Moran I et al.* Single Cell RNA Sequencing of Rare Immune Cell Populations. Frontiers in immunology 2018; 9: 1553

11 *Shomroni O, Sitte M, Schmidt J et al.* A novel single-cell RNA-sequencing approach and its applicability connecting genotype to phenotype in ageing disease. Scientific reports 2022; 12(1): 4091

12 *Picelli S, Björklund ÅK, Faridani OR et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nature methods 2013; 10(11): 1096–1098

13 *Sheng K, Cao W, Niu Y et al.* Effective detection of variation in single-cell transcriptomes using MATQ-seq. Nature methods 2017; 14(3): 267–270

14 *Jaitin DA, Kenigsberg E, Keren-Shaul H et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science (New York, N.Y.) 2014; 343(6172): 776–779

15 *Hashimshony T, Wagner F, Sher N, Yanai I*. CEL-Seq: Single-cell RNA-Seq by multiplexed linear amplification. Cell reports 2012; 2(3): 666–673

16 *Alexander B. Rosenberg, Charles M. Roco, Richard A. Muscat et al.* Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. SCIENCE 2018(Vol. 360, No. 6385): 176–182

17 *Gierahn TM, Wadsworth MH, Hughes TK et al.* Seq-Well: Portable, low-cost RNA sequencing of single cells at high throughput. Nature methods 2017; 14(4): 395–398

18 *Gao C, Zhang M, Chen L*. The Comparison of Two Single-cell Sequencing Platforms: BD Rhapsody and 10x Genomics Chromium. Current genomics 2020; 21(8): 602–609

19 *Macosko EZ, Basu A, Satija R et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 2015; 161(5): 1202–1214
20 *Klein AM, Mazutis L, Akartuna I et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 2015; 161(5): 1187–1201
21 *Zheng GXY, Terry JM, Belgrader P et al.* Massively parallel digital transcriptional profiling of single cells. Nature communications 2017; 8: 14049
22 *Hao Y, Hao S, Andersen-Nissen E et al.* Integrated analysis of multimodal single-cell data. Cell 2021; 184(13): 3573-3587.e29
23 *Yang L, Chan AKN, Miyashita K et al.* High-resolution characterization of gene function using single-cell CRISPR tiling screen. Nature communications 2021; 12(1): 4063
24 *10x Genomics company*. Chromium Single cell 3´Reagent Kits v2 User Guide. Rev F; 2019
25 *Trapnell C, Cacchiarelli D, Grimsby J et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nature biotechnology 2014; 32(4): 381–386
26 *Sonja Hänzelmann, Roberto Castelo and Justin Guinney*. GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinformatics 2013(14)
27 *Tsai C-H, Yang S-H, Chien C-M et al.* Mechanisms of cardiotoxin lll-induced apoptosis in human colorectal cancer colo205 cells. Clinical and experimental pharmacology & physiology 2006; 33(3): 177–182
28 *Laura Elizabeth Lansdowne*. Cancer cells vs normal cells. Technology Networks Cancer Research; 2020
29 *Long J, Wang D, Wang A et al.* A mutation-based gene set predicts survival benefit after immunotherapy across multiple cancers and reveals the immune response landscape. Genome medicine 2022; 14(1): 20
30 *10x Genomics company*. What are the additional peaks in my Single Cell Gene Expression library? 10x Genomics Biology at true resolution; Library Construction
31 *10x Genomics company*. Public dataset frozen PBMC Donor A. v1 chemistry
32 *10x Genomics company*. Public dataset 4k PBMC. v2 chemistry
33 *10x Genomics company*. Public dataset 8k PBMC. v2 chemistry
34 *Silvia Márquez-Jurado, Juan Díaz-Colunga, Ricardo Pires Neves et al.* Mitochondrial levels determine variability in cell death by modulating apoptotic gene expression. Nature communications; 2018(389)
35 *Satija R, Farrell JA, Gennert D et al.* Spatial reconstruction of single-cell gene expression data. Nature biotechnology 2015; 33(5): 495–502
36 *Stuart T, Butler A, Hoffman P et al.* Comprehensive Integration of Single-Cell Data. Cell 2019; 177(7): 1888-1902.e21
37 *Shuparski AG, Higgins BW, Miller KB et al.* Single cell qtSEQ: Cell-indexed quantitative and targeted RNA sequencing for sorted rare lymphocyte subpopulations. STAR protocols 2022; 3(1): 101064
38 *Wang X, He Y, Zhang Q et al.* Direct Comparative Analyses of 10X Genomics Chromium and Smart-seq2. Genomics, proteomics & bioinformatics 2021; 19(2): 253–266

39 *Kreher CR, Dittrich MT, Guerkov R et al.* CD4+ and CD8+ cells in cryopreserved human PBMC maintain full functionality in cytokine ELISPOT assays. Journal of immunological methods 2003; 278(1-2): 79–93
40 Freytag et al 2018_10x Data PBMCs median genes per cell
41 *Hyuna Sung, Jacques Ferlay, Rebecca L. Siegel et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA CANCER J CLI 2021(Volume 71, Number 3): 209–249
42 *Barnes JL, Zubair M, John K et al.* Carcinogens and DNA damage. Biochemical Society transactions 2018; 46(5): 1213–1224
43 *Nymark P, Wikman H, Hienonen-Kempas T, Anttila S.* Molecular and genetic changes in asbestos-related lung cancer. Cancer letters 2008; 265(1): 1–15
44 *Tsao S-W, Tsang CM, To K-F, Lo K-W.* The role of Epstein-Barr virus in epithelial malignancies. The Journal of pathology 2015; 235(2): 323–333
45 *Song G, Cheng L, Chao Y et al.* Emerging Nanotechnology and Advanced Materials for Cancer Radiation Therapy. Advanced materials (Deerfield Beach, Fla.) 2017; 29(32)
46 *Sun Y, Liu Y, Ma X, Hu H.* The Influence of Cell Cycle Regulation on Chemotherapy. International journal of molecular sciences 2021; 22(13)
47 *Lionel A. Kankeu Fonkoua, Olivia Sirpilla, Reona Sakemura et al.* CAR T cell therapy and the tumor microenvironment: Current challenges and opportunities. Molecular Therapy Oncolytics Review; 2022(Vol. 25): 69–77
48 *Arianna Pocaterra, Marco Catucci and Anna Mondino*. Adoptive T cell therapy of solid tumors: time to team up with immunogenic chemo/radiotherapy. Immunology; 2022(Volume 74): 53–59
49 *Enfu Hui.* Immune checkpoint inhibitors. The Journal of cell biology 2019; 218(3): 740–741
50 *Beck JD, Reidenbach D, Salomon N et al.* mRNA therapeutics in cancer immunotherapy. Molecular cancer 2021; 20(1): 69
51 *Kranz LM, Diken M, Haas H et al.* Systemic RNA delivery to dendritic cells exploits antiviral defence for cancer immunotherapy. Nature 2016; 534(7607): 396–401
52 *Baiersdörfer M, Boros G, Muramatsu H et al.* A Facile Method for the Removal of dsRNA Contaminant from In Vitro-Transcribed mRNA. Molecular therapy. Nucleic acids 2019; 15: 26–35
53 *Karikó K, Muramatsu H, Ludwig J, Weissman D.* Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translation of nucleoside-modified, protein-encoding mRNA. Nucleic acids research 2011; 39(21): e142
54 *Karikó K, Buckstein M, Ni H, Weissman D.* Suppression of RNA recognition by Toll-like receptors: The impact of nucleoside modification and the evolutionary origin of RNA. Immunity 2005; 23(2): 165–175
55 *Karikó K, Muramatsu H, Welsh FA et al.* Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. Molecular therapy the journal of the American Society of Gene Therapy 2008; 16(11): 1833–1840

56 *Karikó K, Muramatsu H, Keller JM, Weissman D.* Increased erythropoiesis in mice injected with submicrogram quantities of pseudouridine-containing mRNA encoding erythropoietin. Molecular therapy the journal of the American Society of Gene Therapy 2012; 20(5): 948–953
57 *Andries O, Mc Cafferty S, Smedt SC* de et al. N(1)-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice. Journal of controlled release official journal of the Controlled Release Society 2015; 217: 337–344
58 *Perez CR, Palma M* de. Engineering dendritic cell vaccines to improve cancer immunotherapy. Nature communications 2019; 10(1): 5408
59 *Kranz LM, Diken M, Fritz D, Haas H, Holzmann M, Meng M, Reuter K, Roth R, Selmi A, Kreiter S, et al.* Characterization of the IFNα response upon systemic administration of targeted mRNA vaccines. Cancer Immunotherapy (CIMT) Meeting; 2014
60 *Ugur Sahin, Petra Oehm, Evelyna Derhovanessian et al.* An RNA vaccine drives immunity in checkpoint-inhibitor-treated melanoma. Nature; 2020(585): 107–112
61 *Karikó K, Ni H, Capodici J, Lamphier M, Weissman D*. mRNA is an endogenous ligand for Toll-like receptor 3. Biol. Chem. 2004(279(13)): 12542–12550
62 *Coffman RL, Sher A, Seder RA*. Vaccine adjuvants: Putting innate immunity to work. Immunity; 2010(33(4)): 492–503
63 *Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y.* Single-cell RNA sequencing technologies and applications: A brief overview. Clin Transl Med 2022(12(3))
64 *Haque, A., Engel, J., Teichmann, S.A. et al.* A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. Genome medicine; 2017(75)
65 *L. Kranz, L. Hilscher, M. Diken, L. Kolb, E. Diken, M. Streuber, S. Kreiter and U. Sahin*. Intrinsic adjuvanticity drives T cell immunity induced by RNA vaccines. manuscript under preparation; 2024
66 *Arpino JAJ, Rizkallah PJ, Jones DD.* Crystal structure of enhanced green fluorescent protein to 1.35 Å resolution reveals alternative conformations for Glu222. PloS one 2012; 7(10): e47132
67 *Ansari AM, Ahmed AK, Matsangos AE et al.* Cellular GFP Toxicity and Immunogenicity: Potential Confounders in in Vivo Cell Tracking Experiments. Stem cell reviews and reports 2016; 12(5): 553–559
68 *Becht E, McInnes L, Healy J et al.* Dimensionality reduction for visualizing single-cell data using UMAP. Nature biotechnology 2018
69 *Ben Wylie, James Read, Anthony C. Buzzai, Teagan Wagner, Niamh Troy, Genevieve Syn, Shane R. Stone, Bree Foley, Anthony Bosco, Mark N. Cruickshank and Jason Waithman*. CD8+XCR1neg Dendritic Cells Express High Levels of Toll-Like Receptor 5 and a Unique Complement of Endocytic Receptors. Frontiers in immunology; 2019(9): 2990
70 *Pardi N, Hogan MJ, Naradikian MS et al.* Nucleoside-modified mRNA vaccines induce potent T follicular helper and germinal center B cell responses. The Journal of experimental medicine 2018; 215(6): 1571–1588
71 *Beck JD, Reidenbach D, Salomon N et al.* mRNA therapeutics in cancer immunotherapy. Molecular cancer 2021; 20(1): 69

72 *Chunfa L, Xin S, Qiang L et al.* The Central Role of IFI204 in IFN-β Release and Autophagy Activation during Mycobacterium bovis Infection. Frontiers in cellular and infection microbiology 2017; 7: 169
73 *Q Lu, N Shen, XM Li and SL Chen*. Genomic view of IFN-a response in pre-autoimmune NZB/W and MRL/lpr mice. Genes and Immunity 2007(8): 590–603
74 *Callahan V, Hawks S, Crawford MA et al.* The Pro-Inflammatory Chemokines CXCL9, CXCL10 and CXCL11 Are Upregulated Following SARS-CoV-2 Infection in an AKT-Dependent Manner. Viruses 2021; 13(6)
75 *Irving AT, Zhang Q, Kong P-S et al.* Interferon Regulatory Factors IRF1 and IRF7 Directly Regulate Gene Expression in Bats in Response to Viral Infection. Cell reports 2020; 33(5): 108345
76 *Grunwitz C, Salomon N, Vascotto F et al.* HPV16 RNA-LPX vaccine mediates complete regression of aggressively growing HPV-positive mouse tumors and establishes protective T cell memory. Oncoimmunology 2019; 8(9): e1629259
77 *Kranz LM, Diken M, Fritz D, Haas H, Holzmann M, Meng M, Reuter K, Roth R, Selmi A, Kreiter S, et al*. Novel RNA-lipoplexes with immunostimulatory and targeting properties induce potent anti-tumoral immunity. Abstract, 43rd Annu. Meet. Ger. Soc. Immunol. (DGfI), Mainz 2013
78 *Kreiter S, Vormehr M, van de Roemer N et al.* Mutant MHC class II epitopes drive therapeutic immune responses to cancer. Nature 2015; 520(7549): 692–696
79 *Sabado RL, Balan S, Bhardwaj N*. Dendritic cell-based immunotherapy. Cell research 2017; 27(1): 74–95
80 *Wendel M, Galani IE, Suri-Payer E, Cerwenka A*. Natural killer cell accumulation in tumors is dependent on IFN-gamma and CXCR3 ligands. Cancer research 2008; 68(20): 8437–8445
81 *Peng W, Liu C, Xu C et al.* PD-1 blockade enhances T-cell migration to tumors by elevating IFN-γ inducible chemokines. Cancer research 2012; 72(20): 5209–5218
82 *Böttcher JP, Bonavita E, Chakravarty P et al.* NK Cells Stimulate Recruitment of cDC1 into the Tumor Microenvironment Promoting Cancer Immune Control. Cell 2018; 172(5): 1022-1037.e14
83 *Matsuo K, Kitahata K, Kawabata F et al.* A Highly Active Form of XCL1/Lymphotactin Functions as an Effective Adjuvant to Recruit Cross-Presenting Dendritic Cells for Induction of Effector and Memory CD8+ T Cells. Frontiers in immunology 2018; 9: 2775
84 *van den Eeckhout B, Tavernier J, Gerlo S*. Interleukin-1 as Innate Mediator of T Cell Immunity. Frontiers in immunology 2020; 11: 621931
85 *Pérez-Lorenzo R, Markell LM, Hogan KA et al.* Transforming growth factor beta1 enhances tumor promotion in mouse skin carcinogenesis. Carcinogenesis 2010; 31(6): 1116–1123
86 *Ugur Sahin, Petra Oehm, Evelyna Derhovanessian et al.* An RNA vaccine drives immunity in checkpoint-inhibitor-treated melanoma
87 *Sun L, Zhang W, Zhao L et al.* Self-Tolerance of Vascular Tissues Is Broken Down by Vascular Dendritic Cells in Response to Systemic Inflammation to Initiate Regional Autoinflammation. Frontiers in immunology 2022; 13: 823853

88 *Mackay IR*. Travels and travails of autoimmunity: A historical journey from discovery to rediscovery. Autoimmunity reviews 2010; 9(5): A251-8
89 *Ascherio A*. Environmental factors in multiple sclerosis. Expert review of neurotherapeutics 2013; 13(12 Suppl): 3–9
90 *Lassmann H*. Multiple Sclerosis Pathology. Cold Spring Harbor perspectives in medicine 2018; 8(3)
91 *Fletcher JM, Lalor SJ, Sweeney CM et al.* T cells in multiple sclerosis and experimental autoimmune encephalomyelitis. Clinical and experimental immunology 2010; 162(1): 1–11
92 *Dendrou CA, Fugger L, Friese MA.* Immunopathology of multiple sclerosis. Nature reviews. Immunology 2015; 15(9): 545–558
93 *Wootla B, Denic A, Keegan BM et al.* Evidence for the role of B cells and immunoglobulins in the pathogenesis of multiple sclerosis. Neurology research international 2011; 2011: 780712
94 *Krienke C.* Nucleoside-Modified Autoantigen-Encoding mRNA-Vaccine as a Novel Approach for Antigen-Specific Tolerance Induction in Autoimmunity. Published Ph.D thesis 2020(Johannes Gutenber-Universität Mainz)
95 *Calahorra L, Camacho-Toledano C, Serrano-Regal M, Ortega M, Clemente D*. Regulatory Cells in Multiple Sclerosis: From Blood to Brain. Biomedicines; 2022(10(2)): 335
96 *Grigoriadis N, van Pesch V*. A basic overview of multiple sclerosis immunopathology. European journal of neurology 2015; 22 Suppl 2: 3–13
97 *Dolati S, Babaloo Z, Jadidi-Niaragh F et al.* Multiple sclerosis: Therapeutic applications of advancing drug delivery systems. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie 2017; 86: 343–353
98 *Feigin VL, Abajobir AA, Abate KH et al.* Global, regional, and national burden of neurological disorders during 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. The Lancet Neurology 2017; 16(11): 877–897
99 *Compston A, Coles A.* Multiple sclerosis. The Lancet 2002; 359(9313): 1221–1231
100 *Hemmer B, Archelos JJ, Hartung H-P*. New concepts in the immunopathogenesis of multiple sclerosis. Nature reviews. Neuroscience 2002; 3(4): 291–301
101 *Renner K, Neumayer S, Talke Y et al.* B-cell modulation with anti-CD79b antibodies ameliorates experimental autoimmune encephalitis in mice. European journal of immunology 2022; 52(4): 656–668
102 *Putzki N, Baranwal MK, Tettenborn B et al.* Effects of natalizumab on circulating B cells, T regulatory cells and natural killer cells. European neurology 2010; 63(5): 311–317
103 *Lamb YN*. Ocrelizumab: A Review in Multiple Sclerosis. Drugs 2022; 82(3): 323–334
104 *Dobson R, Giovannoni G*. Multiple sclerosis - a review. European journal of neurology 2019; 26(1): 27–40
105 *Krienke C., Kolb L., Diken E., Streuber M., Kirchhoff S., Bukur T., Akilli-Öztürk Ö., Kranz L., Berger H., Petschenka J., Diken M., Kreiter S., Yogev N., Waisman A., Karikó K., Türeci Ö., Sahin U.* A noninflammatory mRNA vaccine for treatment of experimental

autoimmune encephalomyelitis. SCIENCE; 8 Januar 2021(Vol371, Issue 6525): pp. 145-153

106 *Karikó K, Buckstein M, Ni H, Weissman D.* Suppression of RNA recognition by Toll-like receptors: The impact of nucleoside modification and the evolutionary origin of RNA. Immunity 2005; 23(2): 165–175

107 *Karikó K, Muramatsu H, Welsh FA et al.* Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. Molecular therapy the journal of the American Society of Gene Therapy 2008; 16(11): 1833–1840

108 *Yogev N, Frommer F, Lukas D, Kautz-Neu K, Karram K, Ielo D, von Stebut E, Probst HC, van den Broek M, Riethmacher D, Birnberg T, Blank T, Reizis B, Korn T, Wiendl H, Jung S, Prinz M, Kurschus FC, Waisman A.* Dendritic cells ameliorate autoimmunity in the CNS by controlling the homeostasis of PD-1 receptor(+) regulatory T cells. Immunity; 2012(37(2)): 264–275

109 *Neef T, Ifergan I, Beddow S, Penaloza-MacMaster P, Haskins K, Shea LD, Podojil JR, Miller SD.* Tolerance Induced by Antigen-Loaded PLG Nanoparticles Affects the Phenotype and Trafficking of Transgenic CD4+ and CD8+ T Cells. Cells; 2021(10(12)): 3445

110 *Getts DR, Turley DM, Smith CE et al.* Tolerance induced by apoptotic antigen-coupled leukocytes is induced by PD-L1+ and IL-10-producing splenic macrophages and maintained by T regulatory cells. Journal of immunology (Baltimore, Md. 1950) 2011; 187(5): 2405–2417

111 *Prasad S, Neef T, Xu D et al.* Tolerogenic Ag-PLG nanoparticles induce tregs to suppress activated diabetogenic CD4 and CD8 T cells. Journal of autoimmunity 2018; 89: 112–124

112 *Christine M. Wardell, Megan K. Levings*. mRNA vaccines take on immune tolerance

113 *Hoyer S, Prommersberger S, Pfeiffer IA et al.* Concurrent interaction of DCs with CD4(+) and CD8(+) T cells improves secondary CTL expansion: It takes three to tango. European journal of immunology 2014; 44(12): 3543–3559

114 *Ilicic T, Kim JK, Kolodziejczyk AA et al.* Classification of low quality cells from single-cell RNA-seq data. Genome biology 2016; 17: 29

115 *Kunkl M, Frascolla S, Amormino C et al.* T Helper Cells: The Modulators of Inflammation in Multiple Sclerosis. Cells 2020; 9(2)

116 *Croft M, Bradley L M, Swain S L*. Naive versus memory CD4 T cell response to antigen. Memory cells are less dependent on accessory cell costimulation and can respond to many antigen-pesenting cell types including resting B cells. The journal of immunology; 2022(152): 2675–2685

117 *Li N, Zhan X*. Mitochondrial Dysfunction Pathway Networks and Mitochondrial Dynamics in the Pathogenesis of Pituitary Adenomas. Frontiers in endocrinology 2019; 10: 690

118 *Wu T, Hu E, Xu S et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Cambridge (Mass.)) 2021; 2(3): 100141

119 *Runne C, Chen S.* WD40-repeat proteins control the flow of Gβγ signaling for directional cell migration. Cell adhesion & migration 2013; 7(2): 214–218
120 *Xu H, Yan Y, Williams MS et al.* MS4a4B, a CD20 homologue in T cells, inhibits T cell propagation by modulation of cell cycle. PloS one 2010; 5(11): e13780
121 *Bar-Or A, Vollmer T, Antel J, Arnold D, Caroline Bodner C, Campagnolo D, Gianettoni J, Jalili F, Kachuck N, Lapierre Y, Niino M, Oger J, Price M, Rhodes S, Robinson W, Shi F, Utz P, Valone F, Weiner L, Steinman L, Garren H*. Induction of Antigen-Specific Tolerance in Multiple Sclerosis After Immunization With DNA Encoding Myelin Basic Protein in a Randomized, Placebo-Controlled Phase 1/2 Trial. Arch Neurol; 2007(64(10)): 1407–1415
122 *Baughmann E, Mendoza J, Ortega S, Ayers C, Greenber B, Frohman E, Karandikar N.* Neuroantigen-Specific CD8+ Regulatory T-Cell Function is Deficient During Acute Exacerbation of Multiple Sclerosis. J Autoimmun; 2011(36(2)): 115–124
123 *Li J, Zaslavsky M, Su Y et al.* KIR+CD8+ T cells suppress pathogenic T cells and are active in autoimmune diseases and COVID-19. Science (New York, N.Y.) 2022; 376(6590): eabi9591
124 *Li R, Patterson KR, Bar-Or A.* Reassessing B cell contributions in multiple sclerosis. Nature immunology 2018; 19(7): 696–707
125 *Jain RW, Yong VW*. B cells in central nervous system disease: Diversity, locations and pathophysiology. Nature reviews. Immunology 2022; 22(8): 513–524
126 *Catalán D, Mansilla MA, Ferrier A et al.* Immunosuppressive Mechanisms of Regulatory B Cells. Frontiers in immunology 2021; 12: 611795
127 *Vasileiadis GK, Dardiotis E, Mavropoulos A et al.* Regulatory B and T lymphocytes in multiple sclerosis: Friends or foes? Auto- immunity highlights 2018; 9(1): 9
128 *Ahmadi A, Fallah Vastani Z, Abounoori M et al.* The role of NK and NKT cells in the pathogenesis and improvement of multiple sclerosis following disease-modifying therapies. Health science reports 2022; 5(1): e489
129 *Lamb YN*. Ozanimod: First Approval. Drugs 2020; 80(8): 841–848
130 *Dammes N PD*. Paving the Road for RNA Therapeutics. Trends Pharmacol Sci; 2020(41(10)): 755–775
131 *Yang M, Olaoba OT, Zhang C et al.* Cancer Immunotherapy and Delivery System: An Update. Pharmaceutics 2022; 14(8)
132 *Pardoll DM*. The blockade of immune checkpoints in cancer immunotherapy. Nature reviews. Cancer 2012; 12(4): 252–264
133 *Guan Wang, Xi Kang, Katherine S. Chen et al.* An engineered oncolytic virus expressing PD-L1 inhibitors activates tumor neoantigen-specific T cell responses. Nat Commun 2020(11): 1395
134 *Tumeh PC, Harview CL, Yearley JH et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature 2014; 515(7528): 568–571
135 *Franziska Füchsl and Angela M. Krackhardt*. Adoptive Cellular Therapy for Multiple Myeloma Using CAR- and TCR-Transgenic T Cells: Response and Resistance. Cells 2022(11(3)): 410

136 *Santos Apolonio J, Lima de Souza Gonçalves V, Cordeiro Santos ML et al.* Oncolytic virus therapy in cancer: A current review. World journal of virology 2021; 10(5): 229–255

137 *Conlon KC, Miljkovic MD, Waldmann TA.* Cytokines in the Treatment of Cancer. Journal of interferon & cytokine research the official journal of the International Society for Interferon and Cytokine Research 2019; 39(1): 6–21

138 *Kosinska, A. D. et al.* Synergy of therapeutic heterologous prime–boost hepatitis B vaccination with CpG-application to improve immune control of persistent HBV infection. SCIENCE 2019(Rep. 9, 10808)

139 *Brody, J. D. et al.* In situ vaccination with a TLR9 agonist induces systemic lymphoma regression: a phase I/II study. J. Clin. Oncol; 2010(28): 4324–4332

140 *Hammerich Lea*. Systemic clinical tumor regressions and potentiation of PD1 blockade with in situ vaccination. Nat. Med; 2019(25): 814–824

141 *Kantoff, P. W. et al.* Sipuleucel-T immunotherapy for castration-resistant prostate cancer. N. Engl. J. Med; 2010(363): 411–422

142 *Noubissi Nzeteu GA, Gibbs BF, Kotnik N, Troja A, Bockhorn M, Meyer NH*. Nanoparticle-based immunotherapy of pancreatic cancer. Front Mol Biosci; 2022(29;9): 948898

143 *Guevara, M. L., Persano, F., and Persano, S.* Advances in lipid nanoparticles for mRNA-based cancer immunotherapy. Front. Chem; 2020(8)

144 *Sahin U*. J.P. Morgan Healthcare Conference January 2020 Ugur Sahin, MD CEO and Co-Founder. J.P. Morgan Healthcare Conference 2020

145 *Leoni V, Vannini A, Gatta V et al.* A fully-virulent retargeted oncolytic HSV armed with IL-12 elicits local immunity and vaccine therapy towards distant tumors. PLoS pathogens 2018; 14(8): e1007209

146 *Kyoung-Ju Kim, Dahye Moon, So Jung Kong et al.* Antitumor effects of IL-12 and GM-CSF co-expressed in an engineered oncolytic HSV-1. Gene Therapy 2021(28): 186–198

147 *Emily M. Cheng, Noah W. Tsarovsky, Paul M. Sondel, Alexander L. Rakhmilevich*. Interleukin-12 as an in situ cancer vaccine component: a review. Cancer Immunology, Immunotherapy 2022(71): 2057–2065

148 *Strauss J, Heery CR, Kim JW et al.* First-in-Human Phase I Trial of a Tumor-Targeted Cytokine (NHS-IL12) in Subjects with Metastatic Solid Tumors. Clinical cancer research an official journal of the American Association for Cancer Research 2019; 25(1): 99–109

149 *McMichael EL, Benner B, Atwal LS et al.* A Phase I/II Trial of Cetuximab in Combination with Interleukin-12 Administered to Patients with Unresectable Primary or Recurrent Head and Neck Squamous Cell Carcinoma. Clinical cancer research an official journal of the American Association for Cancer Research 2019; 25(16): 4955–4965

150 *Vonderheide RH, Kraynyak KA, Shields AF et al.* Phase 1 study of safety, tolerability and immunogenicity of the human telomerase (hTERT)-encoded DNA plasmids INO-1400 and INO-1401 with or without IL-12 DNA plasmid INO-9012 in adult patients with solid tumors. Journal for immunotherapy of cancer 2021; 9(7)

151 *Jin D, Jiang Y, Wei J and Sun J*. New therapeutic strategies based on biasing IL-2 mutants for cancers and autoimmune diseases. International Immunopharmacology 2022(110): 108935

152 *Ffion Harris, Yoana Arroyo Berdugo, Timothy Tree*. IL-2-based approaches to Treg enhancement. Clinical and experimental immunology; 2023(Volume 211, Issue 2): 149–163

153 *Yong-Chen Lu, Linda L. Parker, Tangying Lu et al.* Treatment of Patients With Metastatic Cancer Using a Major Histocompatibility Complex Class II–Restricted T-Cell Receptor Targeting the Cancer Germline Antigen MAGE-A3. Journal of Clinical Oncology 2017(Volume 35, Issue 29)

154 *Lei Q, Wang D, Sun K et al.* Resistance Mechanisms of Anti-PD1/PDL1 Therapy in Solid Tumors. Frontiers in cell and developmental biology 2020; 8: 672

155 *Samaridou E, Heyes J and Lutwyche P*. Lipid nanoparticles for nucleic acid delivery: Current perspectives. Advanced Drug Delivery Reviews; 2020(154-155): 37–63

156 *Stegle O, Teichmann S and Marioni J.* Computational and analytical challenges in single-cell transcriptomics. nature reviews genetics 2015(16): 133–145

157 *Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, Satija R, Smibert P*. Simultaneous epitope and transcriptome measurement in single cells. Nature methods; 2017(14(9)): 865–868

158 *Stoeckius M, Zheng S, Houck-Loomis B et al.* Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. Genome biology 2018; 19(1): 224

159 *10x Genomics company*. Single Cell 3´Reagent Kit s v2 User Guide. User Guide Rev. B 2017

160 *Abraham S, Sach D and Sykes M*. Mechanism of protecion from graft-versus-host disease mortality by IL-2: III. Early Reductions in Donor T Cell Subsets and Expansion of a CD3+CD4-CD8- Cell Population. The journal of immunology 1992(Vol. 148): 3746–3752

161 *Strober S, Cheng L, Zeng D et al.* Double negative (CD4-CD8- alpha beta+) T cells which promote tolerance induction and regulate autoimmunity. Immunological reviews 1996; 149: 217–230

162 *Merims S, Li X, Joe B et al.* Anti-leukemia effect of ex vivo expanded DNT cells from AML patients: A potential novel autologous T-cell adoptive immunotherapy. Leukemia 2011; 25(9): 1415–1422

163 *Lee J, Minden MD, Chen WC et al.* Allogeneic Human Double Negative T Cells as a Novel Immunotherapy for Acute Myeloid Leukemia and Its Underlying Mechanisms. Clinical Cancer Research 2018; 24(2): 370–382

164 *Yao J, Ly D, Dervovic D et al.* Human double negative T cells target lung cancer via ligand-dependent mechanisms that can be enhanced by IL-15. Journal for immunotherapy of cancer 2019; 7(1): 17

165 *Chen Z, Cheng K, Walton Z et al.* A murine lung cancer co-clinical trial identifies genetic modifiers of therapeutic response. Nature 2012; 483(7391): 613–617

166 *Sade-Feldman M, Yizhak K, Bjorgaard SL et al.* Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. Cell 2018; 175(4): 998-1013.e20

167 *Villanueva L, Álvarez-Errico D, Esteller M.* The Contribution of Epigenetics to Cancer Immunotherapy. Trends in immunology 2020; 41(8): 676–691

168 *Kim K, Park S, Park SY et al.* Single-cell transcriptome analysis reveals TOX as a promoting factor for T cell exhaustion and a predictor for anti-PD-1 responses in human cancer. Genome medicine 2020; 12(1): 22

169 *Aliahmad P, Seksenyan A, Kaye J.* The many roles of TOX in the immune system. Current opinion in immunology 2012; 24(2): 173–177

170 *Simpson TR, Li F, Montalvo-Ortiz W et al.* Fc-dependent depletion of tumor-infiltrating regulatory T cells co-defines the efficacy of anti-CTLA-4 therapy against melanoma. The Journal of experimental medicine 2013; 210(9): 1695–1710

171 *Valle-Mendiola A, Weiss-Steider B, Rocha-Zavaleta L and Soto-Cruz I.* IL-2 Enhances Cervical Cancer Cells Proliferation and JAK3/STAT5 Phosphorylation at Low Doses, While at High Doses IL-2 Has Opposite Effects. Cancer Investigation 2014(Volume 32 Issue 4): 115–125

172 *Kleino I, Nowlan K, Kotimaa J, Kekäläinen E.* Optimising protein detection with fixable custom oligo-labelled antibodies for single-cell multi-omics approaches. Biotechnology journal 2022; 17(6): e2100213

173 *McGinnis CS, Murrow LM, Gartner ZJ.* DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. Cell systems 2019; 8(4): 329-337.e4

174 *J. Beck, M. Diken, M. Suchan, M. Streuber, E. Diken, L. Kolb, L. Allnoch, F. Vascotto, D. Peters, T. Beißert, Ö. Akilli-Öztürk, Ö. Türeci,. S. Kreiter, M. Vormehr and U. Sahin.* Long-lasting mRNA-encoded interleukin-2 restores CD8+ T cell neoantigen immunity in MHC class I-deficient cancers. Cancer Cell; 2024

175 *West EE, Jin H-T, Rasheed A-U et al.* PD-L1 blockade synergizes with IL-2 therapy in reinvigorating exhausted T cells. The Journal of clinical investigation 2013; 123(6): 2604–2615

176 *Persano S, Guevara ML, Li Z et al.* Lipopolyplex potentiates anti-tumor immunity of mRNA-based vaccination. Biomaterials 2017; 125: 81–89

177 *A. Le Moignic, V. Malard, T. Benvegnu, L. Lemiègre, M. Berchel, P.-A. Jaffrès, C. Baillou, M. Delost, R. Macedo, J. Rochefort, G. Lescaille, C. Pichon, F.M. Lemoine, P. Midoux, V. Mateo.* Preclinical evaluation of mRNA trimannosylated lipopolyplexes as therapeutic cancer vaccines targeting dendritic cells. Journal of Controlled Release 2018(Volume 278): 110–121

178 *Willem W. Overwijk David S. Lee Deborah R. Surman Kari R. Irvine Christopher E. Touloukian Chi-Chao Chan Miles W. Carroll Bernard Moss Steven A. Rosenberg and Nicholas P. Restifo.* Vaccination with a recombinant vaccinia virus encoding a “self” antigen induces autoimmune vitiligo and tumor cell destruction in mice: Requirement for CD4+ T lymphocytes. Proceedings of the National Academy of Sciences 1996(96): 2982–2987

179 *Zhao Y, Wei K, Chi H et al.* IL-7: A promising adjuvant ensuring effective T cell responses and memory in combination with cancer vaccines? Front. Immunol. 2022; 13: 56

180 *Li H-B, Tong J, Zhu S et al.* m6A mRNA methylation controls T cell homeostasis by targeting the IL-7/STAT5/SOCS pathways. Nature 2017; 548(7667): 338–342

181 *Cha E, Graham L, Manjili MH, Bear HD*. IL-7 + IL-15 are superior to IL-2 for the ex vivo expansion of 4T1 mammary carcinoma-specific T cells with greater efficacy against tumors in vivo. Breast cancer research and treatment 2010; 122(2): 359–369

182 *Emily M. Cheng, Noah W. Tsarovsky, Paul M. Sondel, Alexander L. Rakhmilevich*. Interleukin-12 as an in situ cancer vaccine component: a review. Cancer Immunol Immunother 2022(71): 2057–2065

183 *Ugur Sahin, Petra Oehm, Evelyna Derhovanessian et al.* An RNA vaccine drives immunity in checkpoint-inhibitor-treated melanoma. Nature; 2020(585): 107–112

184 *Jaeu Yi, Aidan T. Miller, Angela S. Archambault, Andrew J. Jones, Tara R. Bradstreet, Sravanthi Bandla, Yu-Sung Hsu, Brian T. Edelson, You W. Zhou, Daved H. Fremont, Takeshi Egawa, Nathan Singh, Gregory F. Wu, Chyi-Song Hsieh.* Antigen-specific depletion of CD4+ T cells by CAR T cells reveals distinct roles of higher- and lower-affinity TCRs during autoimmunity. Science Immunology; 2022(Volume 7, Issue 76)

185 *Andrews TS, Atif J, Liu JC et al.* Single-Cell, Single-Nucleus, and Spatial RNA Sequencing of the Human Liver Identifies Cholangiocyte and Mesenchymal Heterogeneity. Hepatology communications 2022; 6(4): 821–840

186 *Michal Slyper, Caroline B. M. Porter, Orr Ashenberg et al.* A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. Nature medicine; 2020(Vol 26): 792–802

187 *Gökcen Eraslan, Eugene Drokhlyansky, Shankara Anand, Evgenij Fiskin, Ayshwarya Subramanian, Michal Slyper, Jiali Wang, Nicholas Van Wittenberghe, John M. Rouhana, Julia Waldman, Orr Ashenberg, Monkol Lek, Danielle Dionne, Thet Su Win, Michael S. Cuoco, Olena Kuksenko, Alexander M. Tsankov, Philip A. Branton, Jamie L. Marshall, Anna Greka, Gad Getz, Ayellet V. Segrè, François Aguet, Orit Rozenblatt-Rosen, Kristin G. Ardlie, Aviv Regev*. Single-nucleus cross-tissue molecular reference maps toward understanding disease gene function. SCIENCE; 2022(376): 6594

188 *Caleb A. Lareau, Leif S. Ludwig, Christoph Muus et al.* Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. Nature biotechnology; 2021(39): 451–461

189 *Zhang Z, Yang C, Zhang X*. scDART: Integrating unmatched scRNA-seq and scATAC-seq data and learning cross-modality relationship simultaneously. Genome biology 2022; 23(1): 139

190 *Sabrina M. Lewis, Marie-Liesse Asselin-Labat, Quan Nguyen et al.* Spatial omics and multiplexed imaging to explore cancer biology. Nature methods; 2021(18): 997–1012

191 *10x Genomics company*. Chromium Fixed RNA Profiling regants Kits – User Guide. Rev C; 2022(Document number CG000527)

192 *Williams CG, Lee HJ, Asatsuma T et al.* An introduction to spatial transcriptomics for biomedical research. Genome medicine 2022; 14(1): 68

193 *Rasa Elmentaite, Cecilia Domínguez Conde, Lu Yang, Sarah A. Teichmann*. Single-cell atlases: shared and tissue-specific cell types across human organs. nature reviews genetics; 2022(23): 395–410

194 *Andrew Butler, Paul Hoffman, Peter Smibert et al.* Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature biotechnology; 2018(36(5)): 411–420

195 *S. Kreiter, A. Selmi, M. Diken, M. Sebastian, P. Osterloh, H. Schild, C. Huber, O. Türeci, U.* Increased antigen presentation efficiency by coupling antigens to MHC class I trafficking signals. Immunology; 2008(180): 309–318

196 *Kreiter S, Konrad T, Sester M et al.* Simultaneous ex vivo quantification of antigen-specific CD4+ and CD8+ T cell responses using in vitro transcribed RNA. Cancer immunology, immunotherapy CII 2007; 56(10): 1577–1587

197 *Holtkamp S, Kreiter S, Selmi A et al.* Modification of antigen-encoding RNA increases stability, translational efficacy, and T-cell stimulatory capacity of dendritic cells. Blood 2006; 108(13): 4009–4017

198 *Vormehr M*. An RNA encoded extended half-life Interleukin-2 variant unfolds tumor immunity by substantial increase of the effector to Treg ratio. CICON Poster Sess; 2019

199 *Vormehr M*. Substantial improvement of cancer immunotherapy by an RNA encoded extended half-life Interleukin-2 variant. SITC Poster Sess; 2019

200 *D. Weissman, N. Pardi, H. Muramatsu, K. Karikó*. HPLC purification of in vitro transcribed long RNA. Methods Mol. Biol; 2013(969): 43–54

201 *Strober W*. Monitoring cell growth. Current protocols in immunology; 2001(Appendix 3A)

202 *Strober W*. Trypan blue exclusion test of cell viability. Current protocols in immunology; 2001(Appendix 3B)

203 *Pollen AA, Nowakowski TJ, Shuga J et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. Nature biotechnology 2014; 32(10): 1053–1058

204 *Laurens van er Maaten and Geoffrey Hinton*. Visualizing Data using t-SNE. Journal of Machine Learning Research; 2008(9): 2579–2605

205 *Waltman L, van Eck NJ*. A smart local moving algorithm for large-scale modularity-based community detection. Eur. Phys. J. B 2013; 86(11): 75

206 *Richard A. Becker, John M. Chambers und Allan Reeve Wilks*. The new S language a programming environment for data analysis and graphics. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, Calif. 1988

9 PUBLICATION

Krienke C, Kolb L, Diken E, **Streuber M,** Kirchhoff S, Bukur T, Akilli-Öztürk Ö, Kranz LM, Berger H, Petschenka J, Diken M, Kreiter S, Yogev N, Waisman A, Karikó K, Türeci Ö, Sahin U. **A noninflammatory mRNA vaccine for treatment of experimental autoimmune encephalomyelitis**. Science. 2021 Jan 8;371(6525):145-153. doi: 10.1126/science.aay3638. PMID: 33414215.

J. Beck, M. Diken, M. Suchan, **M. Streuber**, E. Diken, L. Kolb, L. Allnoch, F. Vascotto, D. Peters, T. Beißert, Ö. Akilli-Öztürk, Ö. Türeci, S. Kreiter, M. Vormehr and U. Sahin, **Long-lasting mRNA-encoded interleukin-2 restores CD8+ T cell neoantigen immunity in MHC class I-deficient cancers**, Cancer Cell (2024), https://doi.org/10.1016/j.ccell.2024.02.013

10 ACKNOWLEDGMENT

The research and the writing of a doctoral thesis is like accompanying a small child from birth to adulthood. Behind this are not only a handful of people but a whole village of precious people to whom these words and my thanks are addressed from the bottom of my heart.

I thank my family, **my parents** and for your constant support, even if you probably don't know much about my work, you were always there when I needed you and that gave me a lot of strength. I thank **my twin brother**, for all the moments when something didn't work at all in the lab and I was able to puke to you via WhatsApp and you were able to understand and build me up through your lab experience. I thank **my partner**and **our**, as well as **grandma**. If it wasn't for you dear Doro, I would have lost faith in myself on January 20, 2020 and maybe even given up completely my PhD at some point. From you I have learned what it means at work, in everyday family life and in life to have a good time management and a well thought out structure. Above all, you showed me how important it is for body and mind to take a creative break and when my brain needs 100% for which tasks. I owe you so much.

A special thanks to my doctoral father **supervisor 1** and my second supervisor **2** and my second thesis reviewer **3** for the supervision of my thesis, the meetings I had with you and the feedbacks I got from you. I especially want to thank Ugur for moving me and our team in the direction of not only learning the techniques in the lab with regard to scRNA-seq, but also understanding how to work bioinformatically to get our own results. From Ugur I could also learn how to break down the most important aspects in a meeting from planned 20min to spontaneous 5min. I thank him for supervising my work as well as his wife Özlem Türeci from whom we learned a lot during her lectures in our PhD program.

A PhD without a team with nerves of steel and hearts of gold is almost unthinkable. I would especially like to thank my two supervisors and leaders of our Clinical Cancer Genomics functional unit, **A** and **B**. From them, I learned to prioritize my tasks so that at the end of the day, I can complete the most important tasks and walk out of the institute with a smile. They taught me to plan an experiment down to the smallest detail and beyond to think about it for those moments of "oh damn, what do we do if ..."! Their knowledge made me better and better when it came to preparing for an important meeting and defending my research results adequately in front of a scientific audience. Above all, they taught me that scientific feedback should never be taken personally, but that one should grow from it and see it rationally.

Furthermore, I would like to thank the rest of the Single Cell Genomics team: **C**. (without your help we still wouldn't get past Cell Ranger, thank you my dear 3rd supervisor, if it wasn't for you this Seurat analysis pipeline wouldn't exist in TRON), **D** (when I'm nervous before an experiment I hum like you to calm down, I'll never forget how we used to rejoice together when we finally got over an error message in Seurat), **E** (when you joined us in 2020 I was so happy not to be the only PhD in our group anymore, I will never forget our first Master Mainz experiment where you learned right away what can go wrong and how we mastered it together, I am proud how you evolved to a Senior PhD from Panda to Sensei Sifu, keep it up and dare to do something), **F** (every time I see you the sun comes up, you want to understand things down to the roots and inspire us a lot with that), **G** (since you joined us our single cell analysis pipeline is booming, I've never seen someone who gets into such a pipeline after only a few weeks with so much dedication and motivation, I always like to spend lunch breaks with you and the others), **H** (after only one day of lab inventory you already knew where everything was and started doing things right away, I'm glad you joined us after 6 months of waiting and you are now part of our team, here's to lots of fun and great experiments in the lab) as well as our wonderful PhD **I** and our excellent scientist **J**. Thank you so much!!!

At this point, I would like to highlight some well-known scientists without whom this work, these projects and the execution of many experiments would not have been possible: I thank **former PhD students** from TRON and **scientist** from BioNTech for their big support with the mRNA constructs and in vitro transcription. Many thanks as well to **K** and **L** from TRON and **M** from BioNTech for your help with Cell sorting. A big thank you to **N** for all the meetings we had before and after the amazing ribocytokine and EAE experiments. Your feedback was always productive and essential to understand whats happening on cell level after mRNA treatment.

An end always has a beginning. For this reason, I would like to thank the people who took me under their wing as a master's student in January 2017 and made TRON palatable as a future workplace through their warmth and knowledge of Next Generation Sequencing and Immunotherapy. I thank **O** and **P** who supervised my master thesis and from whom I learned a lot about NGS, which benefited me, Laura and Elif at the beginning of scRNA-seq establishment. I thank the whole Medical Genomics team (and all former members) as well as PCR Analytics. My time in the lab office room 338 started with you guys. You guys bailed us out more than once and were always a great support in the big lab 337 and with all the

sequencing. I would also like to thank the wonderfully talented women of the Cloning and Vectors Unit. I was able to sit in your office for a while and I still have fond memories of that time.

A PhD is always under the umbrella of a PhD program, at this point a big thank you for everything that made my PhD path easier. But a PhD is nothing without a group of insanely warm, crazy and funny PhDs. I would like to thank all PhDs of TRON and BioNTECH: and former PhDs. I would especially like to highlight a few PhDs like **1, 2, 3** and **4**. We talked a lot about good and bad in our PhDs, organized PhD retreats or even shared hobbies and danced together. The four of you were very special and important pillars in my PhD and I am sad if one day we might not work or spend time together. Four other PhDs who were like true sisters to me. When I finally came to your office and thus closer to my own unit, that was the beginning of a really nice time. I think back with a smile to our hike in the summer of 2019, to the cooking evenings at Marijana's flat, our cursing when something didn't work with RStudio or Phyton, as well as the nice hours of the Christmas parties or in our office where we laughed a lot and supported each other, in our projects as well as privately.

Without our cells we could not survive, they make us and require correct temperatures and nutrients. I thank the queens of cell culture for introducing your secrets under the hoods of cell culture. A lab can only run and function so well as long as someone takes care of the concerns and the machines with such passion as a sculptor takes care of his statue. I would like to thank the lab management team as well as **Q** from quality assurance and our good soul in the lab **R** for your help when we were not sure about certain things in the lab or the execution. Machines like computers can be unpredictable monsters and make your life hell. It's a good thing that we have a team of excellent IT experts. Especially I want to thank Andreas, I know you and your colleagues have a lot to do without our PC problems, but you were always with me with an angel's patience and always managed to get my PC up and running again. Thank you! I also thank all bioinformaticians and as well as all members of the Expression and Cell Biology Unit for your scientific and non-scientific exchange on beautiful sunny days in Mainz and in different places to take care of your physical well-being.

An experiment that is carried out is always part of a project that has to be created and managed. A very special thank you goes to the power women of the project management team (and former). Honestly, without you we would all be screwed and nothing would work. Experiments that are under the umbrella of a project need funding and an overview of the most important

things for our institute, as well as the products and reagents for our lab experiments. I thank **S** for making this PhD in TRON possible with your trust and support.

I would also like to especially thank **T and U** for your help in all matters of human resources and assistance in case of missions or illnesses. A big thank you goes to the masters of Immunology, Animal Models, Immunomonitoring and Biosampling and their whole teams,. Many times you helped me with immunology questions about projects and lab experiments and explaining different methods and techniques in the field of immunotherapy. I also thank our former PhD. I learned a lot from you about MS, working with you on the EAE experiment was exciting, educational and a long process. To the members of the Cancer Vaccine Unit of BioNTech I also want to thank you very much for the support and the many meetings on our Ribocytokines experiment. I would also like to thank another former TRON PhD student, thank you for your great support on the EGFP RNA-LPX experiment.

Last but not least, I would like to thank all the people behind the scientific scenes who stood by me especially when I didn't want to talk about my PhD problems with someone from work. I would like to thank my former partner who accompanied me for more than 10 years through my studies and a large part of my PhD time, as well as her whole family. What you all have done for me, that cannot be put into words and I am very happy about the time I had with you. I would like to thank my former fellow students. Much of what I know today we went through together, thank you for everything I was allowed to learn with and from you. I would also like to thank **V** (graduated 2010), your open ear together with Martin often let me calm down again when we talked over Teamspeak and enjoyed a game. Last but not least I would like to thank four people with whom I lived in a shared flat in Mainz for many years, we were just a super crazy shared flat and talked, laughed and cried about so much nonsense and beautiful things while cooking in the evenings. Because of you, my PhD was certainly a bit easier as well. Thank you, thank you, thank you all.