

Aus dem Institut für Klinische Neuroimmunologie der  
Ludwig-Maximilians-Universität München  
Vorstand: Prof. Dr. Martin Kerschensteiner



# **Immune repertoires of monozygotic twins discordant for multiple sclerosis**

Dissertation

zum Erwerb des Doktorgrades der Naturwissenschaften  
an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität zu München

Vorgelegt von

**JULIA HANSEN**

aus

Salzgitter-Bad

2021

**Mit Genehmigung der Medizinischen Fakultät  
der Universität München**

**Betreuer: PD Dr. rer. nat. Klaus Dornmair**

**Zweitgutachterin: Prof. Dr. rer. nat. Elfriede Nöbner**

**Dekan: Prof. Dr. med. Thomas Gudermann**

**Tag der mündlichen Prüfung: 30.11.2021**

## AFFIDAVIT



**Eidesstattliche Versicherung**

**Hansen, Julia**

---

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

**„Immune repertoires of monozygotic twins discordant for multiple sclerosis“**

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 30.11.2021

Julia Hansen

---

Ort, Datum

---

Unterschrift Doktorandin

## ACKNOWLEDGMENTS

I would like to thank all people who contributed to the success of this thesis.

First, I want to thank PD. Dr. Klaus Dornmair for giving me the opportunity to work in his group and providing this interesting research project. I am grateful for the stimulating discussions and the possibility of working in such an independent way. It was a pleasure to be a part of his group and work in the field of clinically relevant research.

In addition, I would kindly thank Prof. Dr. Reinhard Hohlfeld and Prof. Dr. Martin Kerschensteiner for giving me the opportunity to write my thesis in this institute.

Further, I would like to thank all the employees of the neuroimmunological ambulance for the kind collaborations and help with all the patients' samples. In particular, I would like to thank Dr. Lisa Ann Gerdes for being part of her deeply interesting twin study. My special thanks goes to all patients and control persons who provided cerebrospinal fluid and blood for our studies.

Furthermore, I want to thank Dr. Eduardo Beltran, who introduced me into technologies in a friendship based way and always encouraged me to incorporate and develop my own ideas. It was a pleasure to work with him on this project.

I also do not want to forget all members of the Dornmair group for the warm welcome in the group and their willingness to introduce me into new scientific techniques. Also thanks to all members of our great institute for their tireless readiness for scientific discussions, helpful criticisms and fun in the lab.

Moreover, I am more than deeply grateful for all my colleagues and friends, who I met during the journey of this PhD. Thank you Michelle, Miri, Bella, Cate, Stephan, Ramoni, Kathrin, Vanessa, Marta, Dr. Binder and last but not least Dr. Hütchen for sharing the enthusiasm and passion for science as well as the burdens of the last years and being there for me in between all the ups and downs (and not letting me give up). Thank you for all the lunching, coffee breaks, good talks, laughs and company. It would not have been such a blast without you.

Finally, I would like to express my very profound gratitude to my family, especially my mum and my sister, who always supported and encouraged me throughout the seemingly endless eleven years of studying and working in Munich. Nobody could have thought, I would stay here for so long :D. Without you, this chapter of my life would have not been possible! Thank you so much for your continuous love and patience and giving me this opportunity.

## ZUSAMMENFASSUNG

Multiple Sklerose (MS) ist eine Autoimmunerkrankung des zentralen Nervensystems, die durch perivaskuläre und parenchymale Immunzellinfiltration von CD8<sup>+</sup> und CD4<sup>+</sup> T-Zellen, B-Zellen und Makrophagen in das Gehirn gekennzeichnet ist. Diese eindringenden Zellen spielen eine entscheidende Rolle in Entzündungsprozessen aktiver Läsionen und führen schließlich zur Zerstörung neuronaler Strukturen und klinischen Symptomen. Die Funktion dieser Zellen, insbesondere der dominierenden CD8<sup>+</sup>-Zellpopulation, in der Krankheitspathogenese ist noch nicht geklärt. Zudem war die Identifizierung von antigen-spezifischen  $\alpha$ : $\beta$  T-Zellrezeptoren (TZR) oder H:L B-Zellrezeptoren (BZR) von Lymphozyten zwar möglich, aber technisch anspruchsvoll und konnte bisher nicht mit deren Zellliniensignaturen oder dem Aktivierungsstatus in Verbindung gebracht werden.

Die in dieser Arbeit etablierte neue Methode kombiniert eine auf Durchflusszytometrie basierende Hochdurchsatz-Indexsortierung von Einzelzellen mit einem modifizierten Protokoll der Smartseq2-RNA-Sequenzierungsmethode, um Transkriptomprofile von Liquor-ansässigen Lymphozytenzellen zusammen mit ihren antigenspezifischen Rezeptoren zu untersuchen. Zunächst wurde die Methode mit gefrorenem Liquor des Indexpatienten #4526 etabliert und validiert. Die erfolgreiche Identifizierung von  $\alpha$ : $\beta$  TZR- oder H:L Ig-Ketten mit bis zu 90 % Ausbeute bei gleichzeitiger Charakterisierung klonal expandierter Lymphozytenzellen mit einer noch nie dagewesenen Tiefe durch deren Genexpressionsprofile wurde demonstriert.

Anschließend wurde die Methode an einer einzigartigen Kohorte monozygoter Zwillinge, die für MS diskordant sind, angewandt, um die Immunzellzusammensetzung und -veränderungen des menschlichen Liquors in sehr frühen - sogar prodromalen - Stadien von MS zu untersuchen. Dabei zeigte sich das Vorhandensein von expandierten Plasmazellen bereits bei klinisch gesunden Zwillingen und die auffällige Präsenz einer dominierenden expandierten CD8<sup>+</sup> T-Zell-Population mit einem aktivierten Phänotyp gewebspezifischer Gedächtniszellen in MS Patienten.

Ferner wurden in den expandierten Plasmazellen von fünf Patienten antigeninduzierte Prozesse wie somatische Hypermutationen nachgewiesen. Zudem wurde der besonders interessante T-Zell-Klon TCR TRAV17\*01/TRBV4-1\*01 fünfzehnmal in dem MS erkrankten Zwilling #161, sowie auch viermal in dem klinisch gesunden Zwilling #162 identifiziert. Dessen HLA-Restriktion wurde mit zwei verschiedenen Methoden untersucht, der Rezeptor rekombinant exprimiert und mit einer Methode, die auf Plasmid-kodierten kombinatorischen Peptidbibliotheken basiert, nach Zielantigenen gesucht.

Schließlich wurden Transfektionsbedingungen des Antigen-suchprotokolls durch die Implementierung eines blau fluoreszierenden Reporterproteins optimiert und könnten in zukünftigen Experimenten zur Identifizierung von Kandidatenantigenen beitragen.

## SUMMARY

Multiple sclerosis (MS) is an autoimmune disease of the central nervous system that is characterized by perivascular and parenchymal immune cell infiltration of CD8<sup>+</sup> and CD4<sup>+</sup> T cells, B cells, and macrophages into the brain. Those invading cells play a crucial role in inflammatory processes of active lesions and eventually lead to tissue destruction of neuronal structures and clinical symptoms. The function of those cells, particularly of the dominating CD8<sup>+</sup> cell population in disease pathogenesis is not entirely understood yet. Moreover, the identification of matching antigen-specific  $\alpha:\beta$  T-cell receptors (TCR) or H:L B-cell receptor (BCR) of lymphocytes was possible but technically challenging. Furthermore, previous approaches could not relate the identified antigen receptors to cell lineage signatures or activation status.

Therefore, the new method combines high-throughput flow cytometry based single cell index sorting with a modified protocol of the widely used Smartseq2 RNA-sequencing approach to study whole transcriptomes of cerebrospinal fluid (CSF)-resident single lymphocyte cells along with their antigen-specific receptors. First, the method was established and validated with frozen CSF of index patient #4526 and the successful identification of matching  $\alpha:\beta$  TCR or H:L Ig chains with up to 90% yield, whilst characterization interesting lymphocyte cells with an unprecedented depth by their gene expression profiles, demonstrated.

Then, the method was applied to a unique cohort of monozygotic twins discordant for MS to investigate immune cell compositions and alterations of human CSF in very early - even prodromal - stages of MS on a single cell level. This revealed the presence of clonally expanded plasma cells already in clinically healthy co-twins and the striking relevance of a dominating expanded CD8<sup>+</sup> T cell population with a distinct activated tissue-resident memory cell phenotype in patients with clinical definite MS.

Further, in-depth sequence analysis of the dominantly expanded plasma cells of five patients revealed antigen-driven processes as somatic hypermutations. Moreover, the particularly interesting expanded T cell clone TCR TRAV17\*01/TRBV4-1\*01 was not only detected fifteen times in the MS diseased twin #161, but also four times in the clinically healthy co-twin #162. Additionally, HLA restrictions of TCR TRAV17\*01/TRBV4-1\*01 were investigated by two different approaches, the T cell clone recombinant expressed in a 58 T hybridoma cell line and searched for target antigens with an unbiased method that uses plasmid-encoded combinatorial peptide libraries.

Lastly, transfection conditions of the antigen search protocol were optimized by the implementation of a blue fluorescent reporter protein and might contribute to the identification of candidate antigens in future experiments.

<b>I.</b>	<b>Table of contents</b>	<b>VII</b>
I.	Table of contents.....	VII
II.	Abbreviations.....	XI
<b>1</b>	<b>Introduction.....</b>	<b>2</b>
1.1	Mammalian immune system.....	2
1.1.1	Players of the adaptive immunity.....	3
1.1.2	B cells.....	5
1.1.3	T cells.....	7
1.1.4	Antigen recognition of T cells.....	10
1.1.5	Long-term immunity of T cells.....	13
1.2	Autoimmune diseases.....	14
1.2.1	Multiple sclerosis.....	14
1.2.2	Immune cells in MS.....	16
1.3	Analysing immune repertoires.....	18
1.4	RNA-sequencing.....	20
1.5	Detection of T cell antigens.....	21
<b>2</b>	<b>Aims.....</b>	<b>24</b>
<b>3</b>	<b>Materials and Methods.....</b>	<b>25</b>
3.1	Materials.....	25
3.1.1	Devices.....	25
3.1.2	Chemicals and consumables.....	25
3.1.3	Buffers and Solutions.....	26
3.1.4	Antibodies.....	26
3.1.5	Antibiotics.....	28
3.1.6	Oligonucleotides.....	28
3.1.7	Enzymes & specific Kits.....	31
3.1.8	Plasmids.....	31

3.1.9	Cell lines.....	32
3.1.10	Biological samples.....	33
3.2	Identification and characterization of expanded lymphocytes in MS patients .....	36
3.2.1	Sample preparation.....	39
3.2.2	Fluorescence-activated cell sorting (FACS) .....	40
3.2.3	PCR enrichment .....	41
3.2.4	NGS preparation.....	44
3.2.5	NGS and Data analysis.....	45
3.2.6	Finding TCR TRAV17*01/TRBV4-1*01 in PBMCs.....	46
3.3	Molecular biological methods .....	47
3.3.1	DNA amplification.....	47
3.3.2	Analysis of DNA fragments.....	47
3.3.3	DNA isolation .....	48
3.3.4	Restriction cloning .....	49
3.3.5	Recombinant production of 58 $\alpha\beta$ T hybridoma cells.....	49
3.3.6	Fusion of HLA molecules with mtagBFP2 .....	50
3.4	Cell handling.....	51
3.4.1	Prokaryotic cells.....	51
3.4.2	Eukaryotic cells .....	52
3.4.3	Flow cytometry .....	53
3.5	Identification of target antigens .....	54
3.5.1	Investigating HLA restriction.....	54
3.5.2	Unbiased Identification of T cell antigens .....	54
3.5.3	T cell activation assay .....	55
3.5.4	B-cell antigens.....	56
<b>4</b>	<b>Results.....</b>	<b>57</b>
4.1	Single-cell profiling of patients with neuroimmunological diseases .....	57
4.1.1	Optimizing NGS preparation steps for high-throughput scRNA seq.....	57

4.1.2	Proof of concept: single-cell profiling of human CSF of index patient #4526..	60
4.2	Transcriptome profiling of patients with early and established MS .....	63
4.2.1	Cellular landscapes of CSF lymphocytes .....	63
4.2.2	Clonally expanded lymphocytes .....	66
4.2.3	In-depth characterization of expanded T cells .....	69
4.2.4	Expanded CD8 <sup>+</sup> T cells of individual twin pairs.....	73
4.2.5	Transcriptome profiling of patients with CIS diagnosis .....	75
4.3	Characterization of expanded lymphocyte clones .....	76
4.3.1	Expanded plasma cells .....	76
4.3.2	Expanded CD8 <sup>+</sup> TCR TRAV17*01/TRBV4-1*01 .....	80
4.3.3	Investigating HLA preferences of TCR TRAV17*01/TRBV4-1*01 .....	83
4.4	T cell antigen search .....	88
4.4.1	Antigen search experiments of TCR TRAV17*01/TRBV4-1*01 .....	88
4.4.2	Optimizing the T cell antigen search protocol .....	91
<b>5</b>	<b>Discussion .....</b>	<b>97</b>
5.1	Single-cell profiling of patients with neuroimmunological diseases .....	98
5.1.1	Optimizing NGS preparation steps for high-throughput scRNA seq.....	99
5.1.2	Proof of concept: single-cell profiling of human CSF of index patient #4526..	99
5.2	Transcriptome profiling of patients with early and established MS .....	101
5.2.1	Cellular landscapes of CSF lymphocytes .....	101
5.2.2	Clonally expanded lymphocytes .....	103
5.2.3	In-depth characterization of expanded T cells .....	103
5.2.4	Expanded CD8 <sup>+</sup> T cells of individual twin pairs.....	105
5.2.5	Transcriptome profiling of patients with CIS diagnosis .....	106
5.3	Characterization of expanded lymphocyte clones .....	107
5.3.1	Expanded plasma cells .....	107
5.3.2	Expanded CD8 <sup>+</sup> TCR TRAV17*01/TRBV4-1*01 .....	108
5.3.3	Investigating HLA preferences of TCR TRAV17*01/TRBV4-1*01 .....	109

5.4	T cell antigen search .....	112
5.4.1	Antigen search experiments of TCR TRAV17*01/TRBV4-1*01 .....	112
5.4.2	Optimizing the T cell antigen search protocol .....	113
5.5	Outlook .....	115
<b>6</b>	<b>References.....</b>	<b>116</b>
<b>7</b>	<b>Appendix.....</b>	<b>129</b>
7.1	Methodical part: Nucleotide sequences for cloning reactions .....	129
7.1.1	TRAV17*01 TRAJ23*01 .....	129
7.1.2	TRBV4-1*01 .....	129
7.1.3	TRDV*01 .....	130
7.1.4	pHSE3' section (full sequence: 8072 bp) .....	130
7.1.5	HLA A*24:02:01:01 (1098 bp).....	131
7.1.6	HLA B*07:02:01:01 (1089 bp).....	131
7.1.7	HLA C*07:02:01:01 (1101 bp).....	132
7.1.8	mtagBFP2 (702 bp) .....	132
7.1.9	HLA C*07:02:01:01-BFP construct.....	133
7.1.10	pcDNA6-V5 segment of N24 PECPL library (146 bp).....	133
7.2	Results part: T cell antigen search .....	134
7.2.1	Co-culture experiments: Single cell isolation and reactivation assays.....	134
<b>8</b>	<b>List of Publications .....</b>	<b>139</b>

## II. Abbreviations

Basic units were used in accordance with the international SI system. Amino acids were termed by using the single letter code.

+G	LNA-modified guanosine	ER	Endoplasmic reticulum
A	Adenosine	f	Female
aa	Amino acid	Fab	Fragment antigen binding
APC	Antigen presenting cell	FACS	Fluorescence-activated cell sorting
APC	Allophycocyanin fluorochrome	FCS	Fetal calf serum
$\beta$ 2-m	$\beta$ 2-microglobulin	Fc	Fragment crystallizable
BCR	B-cell receptor	FCS	Fetal calf serum
Blast	Basic local alignment search tool	FITC	Fluorescein isothiocyanate
bp	Base pairs	FOX-P3	Forkhead-box-protein 3
C	Constant	FR	Framework region
C	Cytosine	FSC	Forward scatter
CCR7	Chemokine receptor 7	Fwd.	Forward
CD	Cluster of differentiation	g	Gravitational acceleration constant ( $g \cong 9.81 \text{ m/s}^2$ )
cDNA	Complementary DNA	G	Guanine
CDR	Complementarity determining region	G418	Geneticin
CIS	Clinical isolated syndrome	GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
CNS	Central nervous system	HKG	House-keeping gene
CSF	Cerebrospinal fluid	HLA	Human leucocyte antigen
CSR	Class switch recombination	Hobit	Homolog of Blimp-1 in T cells
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4	HPLC	High performance liquid chromatography
Cy	Cychrome	HSC	Hematopoietic stem cells
D	Diversity	HTS	High throughput sequencing
DC	Dendritic cell	Ig	Immunoglobulin
<i>dd</i> H <sub>2</sub> O	Double deionized water	IgG	Immunoglobulin G
DMSO	Dimethyl sulfoxide	IgM	Immunoglobulin M
DNA	Deoxyribonucleic acid	IIH	Intracranial idiopathic hypertension
dNTP	Deoxynucleotide	IMGT	ImMunoGeneTics
dsDNA	Double stranded DNA	ITAM	Immunoreceptor tyrosine-based activation motif
DTT	Dithiothreitol	IVIG	Intravenous immunoglobulines
EAE	Experimental autoimmune encephalomyelitis	IUPAC	International union of pure and applied chemistry
EDTA	Ethylenediaminetetraacetic acid		
EDSS	Expanded disability status scale		

J	Joining	RACE	Rapid amplification of cDNA-ends
kb	Kilobase	Rev.	Reverse
KLF2	Krüppel-like Factor 2	rG	Riboguanosine
KLRF1	Killer cell lectin like receptor F1	RNA	Ribonucleic acid
L	Ladder (size marker)	RNase	Ribonuclease
LGI-1-E	Leucine-rich-glioma-inactivated-1 encephalitis	RNA-seq	RNA-sequencing
LNA	Locked nucleic acid	rpm	Rounds per minute
m	Male	RRMS	Relapsing remitting MS
m	Mouse	RT	Room temperature
MHC	Major histocompatibility complex	RT reaction	Reverse transcription reaction
M-MLV	Moloney Murine Leukemia Virus	RT-PCR	Reverse transcription polymerase chain reaction
mRNA	Messenger RNA	RX	Rituximab
MRI	Magnet resonance imaging	S1P	Sphingosine-1-phosphate
MS	Multiple sclerosis	S1PR1	Sphingosine-1-phosphate receptor 1
mTec	Medullary thymic epithelial	SC	Single cell
N	Stands for random insertion of any nucleotide	SCNI	Subclinical Neuroinflammation
NFAT	Nuclear factor of activated T cells	SCP	Single cell project
NGS	Next-generation sequencing	SDHA	Succinate dehydrogenase complex, subunit A
NIC	Non-inflammatory control	SHM	Somatic hypermutation
NK cell	Natural killer cell	Smart	Switching mechanism at the 5' end of RNA template
NMDA-R-E	N-Methyl-D-Aspartate-receptor encephalitis	SPMS	Secondary progressive MS
OCB	Oligoclonal bands	SPRI	Solid phase reversible immobilization
pA	Poly A	SSC	Side scatter
PacBlue	Pacific blue	T	Thymine
PBMC	Peripheral blood mononuclear cell	t-SNE	t-Distributed stochastic neighbour embedding
PBS	Phosphate-buffered saline	T <sub>H</sub>	Helper T cells
PCR	Polymerase chain reaction	T <sub>C</sub>	Cytotoxic T cells
PE	Phycoerythrin	T <sub>CM</sub>	Central memory T cells
PECPL	Plasmid-encoded combinatorial peptide libraries	T <sub>EM</sub>	Effector memory T cells
PEG	Polyethylene glycol	T <sub>RM</sub>	Tissue-resident memory cells
PenStrep	Penicillin-Streptomycin	T <sub>REG</sub>	Regulatory T cells
PerCP	Peridinin chlorophyll protein	TBE-Buffer	Tris-borate-EDTA Buffer
PI	Propidium iodide	TCR	T cell receptor
PLEX	Plasma exchange	Tris	Tris(hydroxymethyl)aminomethane
PPIA	Peptidylprolyl isomerase A	TS	Template switch
PPMS	Primary-progressive MS	TSO	Template-switching oligo

U	Uracil
UDG	Uracil-DNA glycosylase
UV	Ultraviolet
V	Variable
VN	'V' stands for either A,C or G and 'N' is any base

WML	White matter lesions
w/v	Weight per volume
w/w	Weight percentage
WT	Whole transcriptome

# 1 Introduction

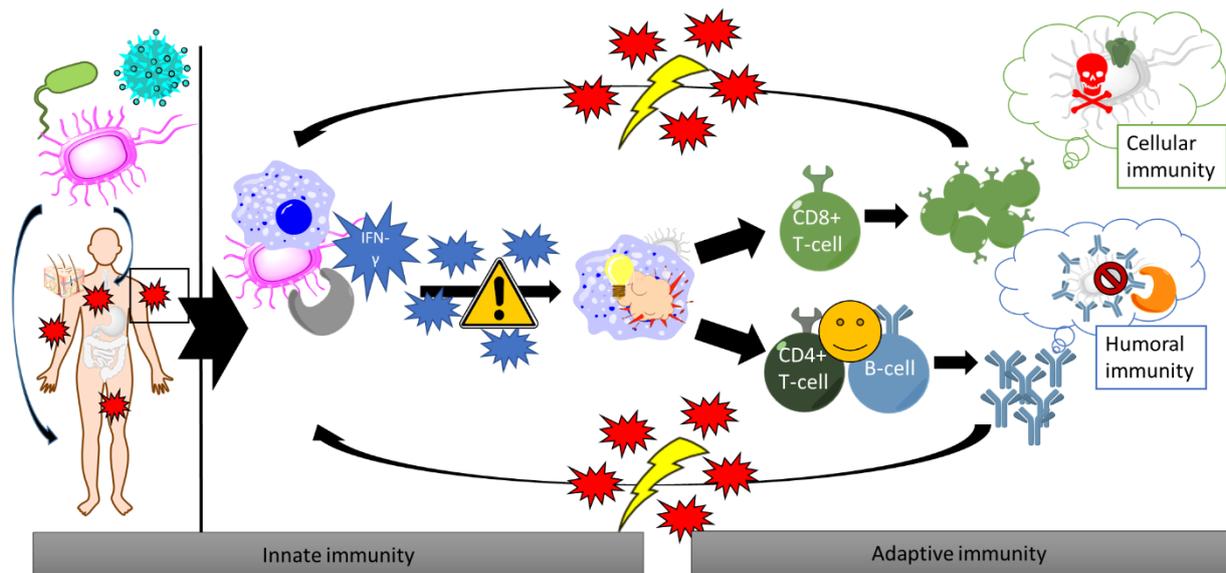
## 1.1 Mammalian immune system

Hazards caused by the infiltration of foreign microorganisms such as virus, bacteria, fungi or parasites have to be defeated by mammals, which have developed an effective immune system with different lines of defence mechanisms (Figure 1).<sup>[1-3]</sup>

First, mammals are naturally protected from invading pathogens by physical barriers (skin or mucosa) and chemical mechanisms (lipids, enzymes or the pH of acidic metabolic systems). Once these physical barriers are damaged or leaky, pathogens are confronted by the second line of mammalian defence, the innate immunity.<sup>[4-5]</sup> It is an immediate cellular defence that comprises a variety of phagocytic immune cells such as macrophages, neutrophils (phagocytes), natural killer cells (NK cells), and the complement system.<sup>[3, 6-7]</sup> The recognition of common pathogens is rather unspecific and limited by a fixed repertoire of germline encoded cell surface receptors and unspecific patterns.<sup>[8-10]</sup> The encounter of conspicuous cellular abnormalities (altered/stressed cells, microbial products, and cytokines) activates scavenger cells to produce an inflammatory environment (e.g. IFN- $\gamma$ ) and directly elicit the removal of the detected pathogens by proteolytic cascades such as intake or digestion.<sup>[11]</sup> Failed or incomplete eradication releases more cytokines by dying immune cells and alarms the third line of mammalian immune defence.<sup>[12-14]</sup>

This so-called adaptive immunity developed as a target-specific defence and is able to eliminate a various number of different pathogens. It is highly specific and diverse but takes more time to build up in order to effectively kill the identified target. Key characteristics of the adaptive immune system are lymphocytes, antibodies, antigen specific receptors and the ability of developing an immunological memory. This memory ability enables a more rapid and specific second response upon reinfection with the same pathogen and thus a shorter disease duration.<sup>[15-16]</sup>

Only the fruitful interaction of all these mechanisms makes the mammalian immune system so effective and ensures their successful survival against pathogenic organisms.



**Figure 1: Defence lines of the mammalian immune system.** Innate immune cells are able to recognize specific and fixed patterns of pathogenic surface structures. This initiates the phagocytosis of infiltrating microorganisms and the generation of an inflammatory environment. Antigen presenting cells recognize those signals and are able to transport this information and present the antigenic peptide to cells of the adaptive immunity in lymphoid organs ( $CD4^+$  T cells, dark green;  $CD8^+$  T cells, light green and B cells, light blue). Either this leads to the activation of the  $CD8^+$  T cell mediated cellular immunity or the B cell mediated humoral immunity to effectively clear the infiltrating pathogen. In both cases, activation leads to differentiation and clonal expansion of target-specific effector clones.  $CD8^+$  cytotoxic T cells (CTLs) can directly kill infected cells by releasing granzymes and perforin. In contrast,  $CD4^+$  helper T cells activate B cells to produce highly specific antibodies. Those are able to clear the pathogen by activating the different effector mechanism neutralization, opsonisation or complement activation.

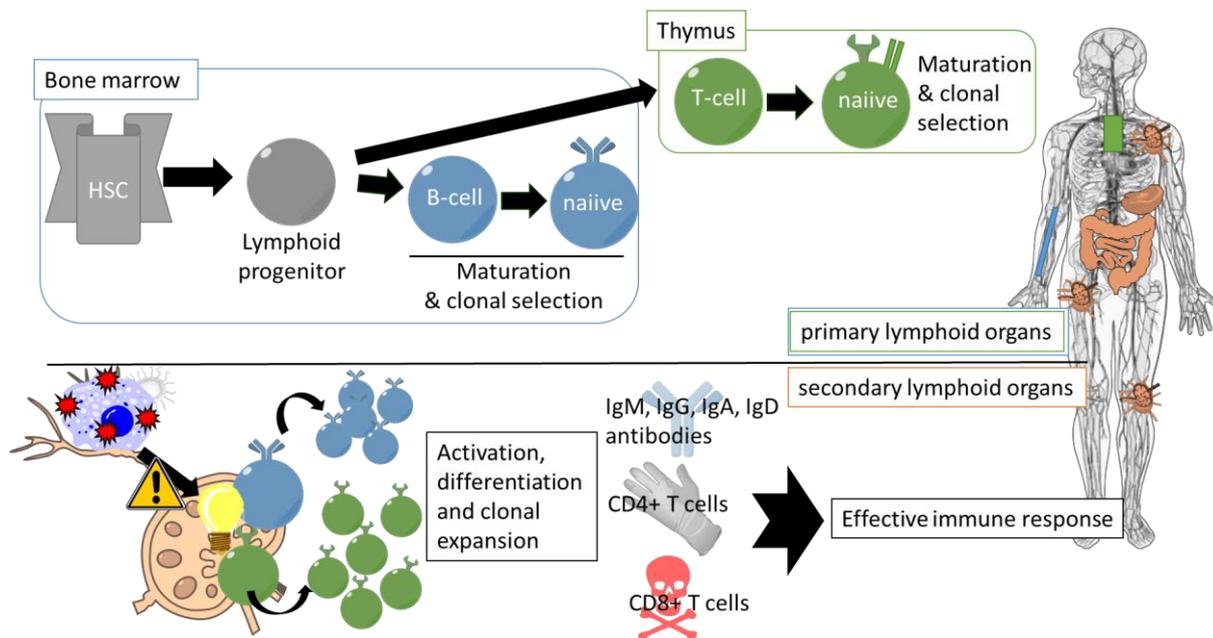
### 1.1.1 Players of the adaptive immunity

Lymphocytes such as T and B cells are main players of the adaptive immunity. Dependent on the type of invading pathogen an effective immune response is initiated. Key features of this acquired immunity are target specific recognition, antigen presentation, complex signal cascades and the formation of a long lasting immunological memory.<sup>[17]</sup>

All lymphocytes arise from pluripotent hematopoietic stem cells in the bone marrow and mature to their respective immune cells in bone marrow, lymph nodes and blood (Figure 2). Developmental stages and rearrangement steps are induced and closely controlled by cytokines and the presence of multiple checkpoints in the primary lymphoid organs (bone marrow, thymus).<sup>[18-19]</sup> This includes the development of a membrane-bound cell surface receptor that is capable of recognizing foreign antigens by distinguishing ‘self’ from ‘non-self’. Receptors are probed for functionality and self-antigen recognition in a process called central tolerance. Only around 10% of all produced lymphocytes survive this clonal selection process and are able to leave to secondary lymphoid organs as mature, naïve lymphocytes.<sup>[20-21]</sup> Those cells patrol throughout the entire organism and are crucial to maintain immune surveillance in the body.

Pathogenic antigens are presented to lymphocytes by antigen presenting cells (APCs). Those phagocytic cells are able to ingest, process and present pathogenic fragments as peptide: MHC complexes on their cell surfaces. Antigens that are presented in such a complex can be recognized by naïve lymphocytes and induce their activation, differentiation and proliferation into a pool of highly specific effector clones with the same antigen-specificity.<sup>[22-24]</sup>

The adaptive immune response can either involve the activation of cytotoxic  $CD8^+$  T cells that directly recognize and kill conspicuous or altered cells. In contrast to this T-cell mediated immunity, the B cell mediated humoral immune response involves the activation of  $CD4^+$  T helper cells. Activated T helper cells are able to induce the differentiation of naïve B cells into plasma cells. Those are able to produce high amounts of antigen specific antibodies that bind invading pathogens and induce their clearance by using different effector mechanisms (neutralization, opsonisation or complement activation).<sup>[25-26]</sup>

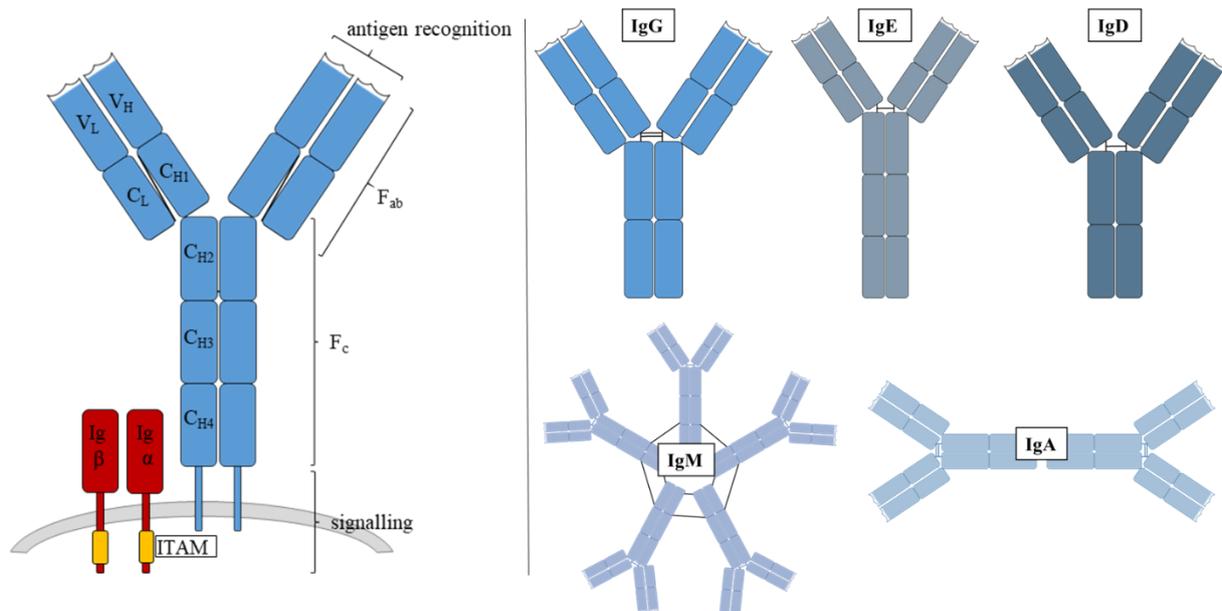


**Figure 2: Developmental stages of lymphocytes. (primary lymphoid organs)** Lymphoid progenitor cells arise from hematopoietic stem cells (HSC) in the bone marrow and can differentiate into different cells of the adaptive immune system. The maturation of B cells (blue) occurs in the bone marrow, while T cells (green) are developing in the thymus. Both cell types undergo maturation processes in which they develop and produce specific cell surface receptors. Those are then probed for functionality and self-antigen recognition in a process called central tolerance. Only 10% of all produced T and B cells are clonally selected and able to leave into secondary lymphoid organs as naïve, mature lymphocytes. **(secondary lymphoid organs)** Immune surveillance is maintained by mature but naïve immune cells that circulate throughout secondary lymphoid organs (lymph nodes, spleen and gut) via blood vessels and the lymphatic system. Those cells are activated by phagocytic cells of the innate immune system (e.g. dendritic cells) that encountered a pathogen in peripheral organs and transport this information to the resting cells in secondary lymphoid organs. Activation of T and B cells induces their differentiation and clonal expansion to highly specific effector clones with the same antigen specificity. Dependent on the type of invading pathogen, an effective immune response is induced. This can either involve the production of target-specific antibodies (IgM, IgG, IgA, IgD) by differentiated plasma cells or antigen-specific  $CD4^+$  T-helper and  $CD8^+$  cytotoxic T cells.

### 1.1.2 B cells

B cells arise and develop from lymphoid progenitor cells in the bone marrow. During maturation, they produce an antigen specific B cell surface receptor (BCR) that is able to detect pathogenic antigens. (Figure 3, left panel) BCRs are then probed for the recognition of self-antigens and 80-90% are directly removed from the repertoire in a process known as negative selection. Dependent on the type of self-antigen recognition, B cells are able to either edit their receptors, induce apoptosis, anergy (a state of functional un-responsiveness) or remain clonally ignorant (tolerance induction).<sup>[27]</sup> Finally, positive selection allows those mature, naïve  $IgM^+$   $IgD^+$  B cells to enter lymphoid follicles and spleen, which increases their half-life's to several weeks.<sup>[28]</sup>

The encounter of an antigen occurs in secondary lymphoid organs (lymph nodes, spleen, peyers patches) and induces B cell activation and differentiation into antibody-producing plasma cells or memory cells. Only resting plasma cells migrate back into the bone marrow and express high levels of IgG, IgA or IgE.<sup>[29-30]</sup>



**Figure 3: Structure of B-cell antibodies. Left panel:** The surface bound B cell receptor (BCR, blue) is composed of two heavy and light chains that are covalently linked by disulphide bridges. It can be subdivided into a variable part ( $F_{ab}$ ) that is crucial for antigen recognition and a constant part that ( $F_c$ ) is bound to the cell surface and transduces activation signals. The signalling is further mediated by the cytosolic immunoreceptor tyrosine-based activation motif (ITAM, yellow) regions of the BCR associated immunoglobulin heterodimer  $Ig_\alpha$  and  $Ig_\beta$  (red).  $F_{ab}$ - fragment of antigen binding –  $F_c$ : fragment crystallisable –  $V_L$ : Variable part of the light chain –  $V_H$ : Variable part of the heavy chain –  $C_L$ : constant part of the light chain –  $C_H$ : constant part of the heavy chain **Right panel:** Structure of the different types of soluble antibodies. The different isotypes are determined by their constant regions and define their effector function in an immune response. Secreted antibodies are produced from activated plasma cells and able to bind target antigens of pathogens and mediate their clearance by triggering the particular effector mechanism neutralization, opsonisation or complement activation.  $IgD$ ,  $IgE$  and  $IgG$  are monomeric, while  $IgA$  occur as a dimer and  $IgM$  in a pentameric form.

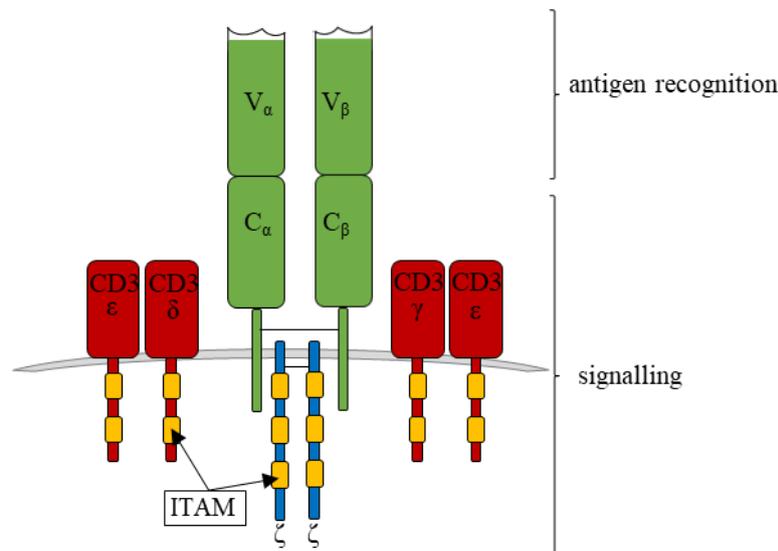
Antibodies can be either soluble or membrane bound to the cell surface of B-cells as BCRs (Figure 3). Either way, they are composed of two heavy and two light chains that are connected by disulphide bonds. According to their function, antibodies are split into a variable (F<sub>ab</sub>-fragment of antigen binding) and a constant part (F<sub>c</sub>- fragment crystallisable).<sup>[31]</sup> The variable part contains the highly variable antigen binding site and the F<sub>c</sub>-part determines its isotype. The F<sub>c</sub> part is crucial for signal transduction in the surface bound B cells or in case of soluble antibodies, it is triggering the respective effector mechanisms (neutralization, opsonisation, complement activation).<sup>[32-34]</sup>

The diverse antibody repertoire of individuals enables them to recognize a broad variety of different antigenic pathogens and structures.<sup>[35]</sup> Upon antigen encounter, B cells are even able to specifically shape their receptors and by this further enhance their antigen specificity. This process known as affinity maturation involves mechanisms of class switch recombination (CSR) and the introduction of somatic hypermutations (SHM).<sup>[36-37]</sup> CSR is regulated by cytokines and loops out switch domains in the constant regions of BCR heavy chains to switch their isotypes and thus their effector function. Additionally SHM induces random mutations into the variable regions of the BCR heavy and light chains, which shapes the antigen binding affinity. Only mutations leading to an increased BCR affinity survive and will produce clonally expanded antibody secreting plasma cells to defeat the hazardous invader.<sup>[38]</sup> In parallel, memory cells bearing the affinity matured BCR are generated and recirculate in the body in a resting state. Reencounter with the same antigen activates them to rapidly differentiate into antibody producing plasma cells and this induces a much faster second immune response.<sup>[39]</sup>

### 1.1.3 T cells

T cells are a crucial part of the adaptive immune system and derive from HSC in the bone marrow. They mature and differentiate in the thymus. During these processes, they produce an antigen-specific T cell receptor (TCR) that is able to recognize antigens and trigger an effective immune response. It consists of an  $\alpha$ - and  $\beta$ -chain, each with a variable part that is able to recognize foreign antigens and a constant region that is anchored in the T-cell membrane (Figure 4). Signal transduction is mediated by the associated signalling proteins  $\zeta$  homodimer and the CD3 ( $\epsilon$ ,  $\delta$ ,  $\gamma$ ) complex.

The TCR is rearranged in different steps during T cell maturation and its affinity subsequently probed for functionality. Only positive selected TCRs are then probed in the thymus by medullary thymic epithelial (mTec) cells for recognition of self-antigens (negative selection). This step is crucial to remove potentially dangerous autoreactive cells from the TCR repertoire and is known as central tolerance. Only around 10% of the produced TCRs will survive this process and remain immunological tolerant whilst all others are clonally deleted.<sup>[40-42]</sup>



**Figure 4: Structure of an  $\alpha$ : $\beta$  TCR complex.** The heterodimeric TCR protein (green) is composed of two glycosylated polypeptide chains that are linked with disulphide bridges and anchored into the T cell membrane (grey) by a cytoplasmic trans-membrane region. Each chain consist of two immunoglobulin (Ig) domains: a highly diverse variable (V) domain that is crucial for antigen recognition and a constant (C) domain that transmits T cell signals. Signal transduction is mediated by the immunoreceptor tyrosine-based activation motifs (ITAMs, yellow) of the associated  $\zeta$  homodimer (blue) and the four CD3 ( $\epsilon$ ,  $\delta$ ,  $\gamma$ ) subunits (red). (Adapted from MURPHY *et al.*)<sup>[43]</sup>

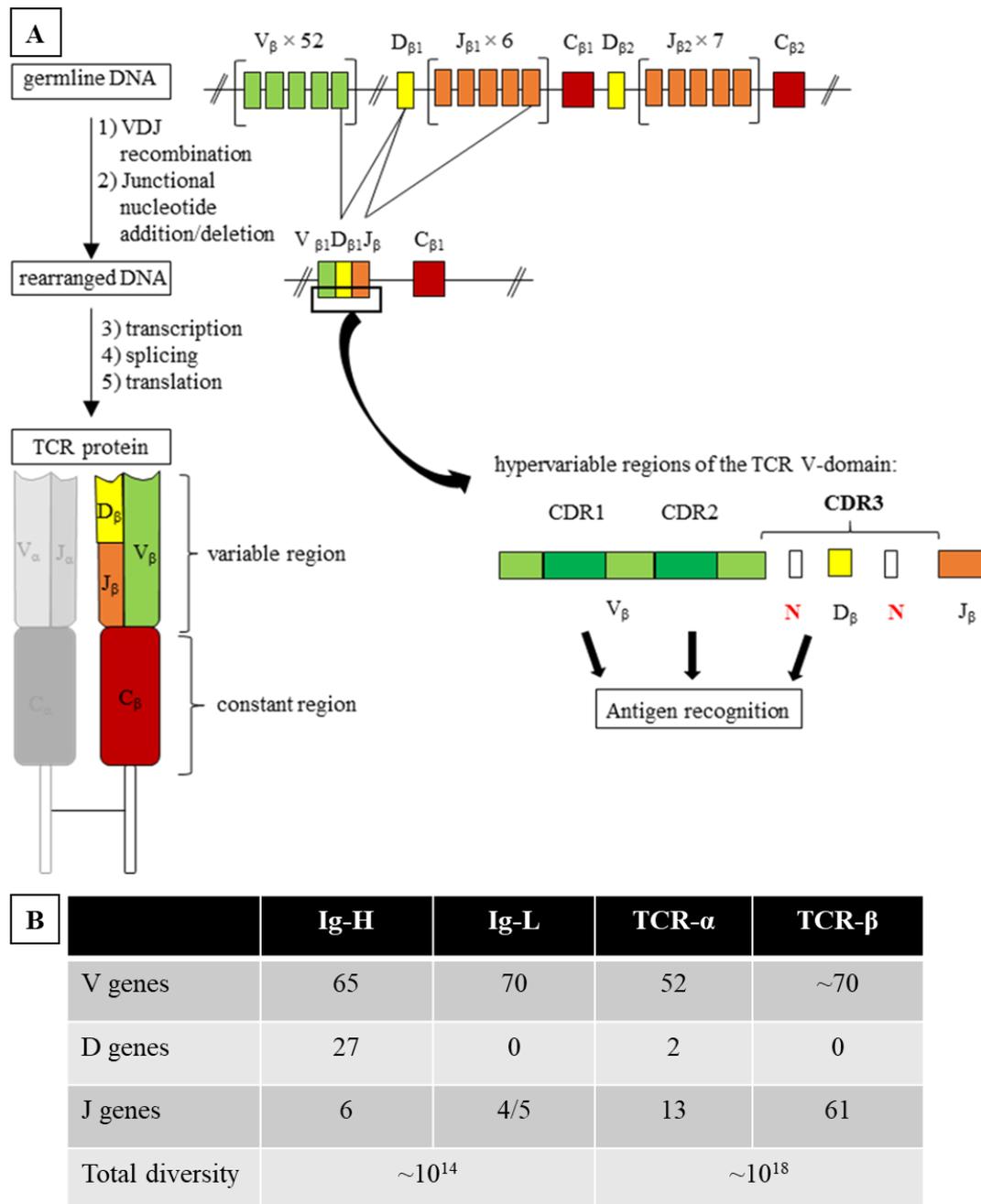
T cell precursors in the thymus also express the co-receptors  $CD4^+$  and  $CD8^+$  that determine their effector functions in an immune response. During maturation, those double-positive T-cells are stimulated to differentiate into  $CD4^+$  T-helper cells ( $T_H$ ) or  $CD8^+$  cytotoxic T cells ( $T_C$ ). Activation through certain stimulation signals induces the T cells to commit to one lineage, which then determines their effector function such as direct killing or calling for help.

Those mature, naïve T cells are able to migrate into the periphery and scan cells for abnormalities. The three major T cell subclasses are  $T_H$ ,  $T_C$  and regulatory T cells ( $T_{REG}$ ), and are characterized by expressing a different panel of cell surface molecules and cytokines.<sup>[44-47]</sup> In contrast to B cells, T cells can only detect antigens that are processed and presented to them in a complex with major histocompatibility complex (MHC) proteins on the surface of APCs.  $CD4^+$  T cells recognize peptides that are presented on MHC class II molecules, whereas  $CD8^+$  T cells require the presentation on MHC class I.<sup>[48]</sup> Binding of the TCR to the MHC:peptide complex activates the T-cells by inducing the phosphorylation of the ITAM regions of the associated signalling proteins. This signal cascade activates the transcription factor NFAT (nuclear factor of activated T cells) and other downstream signalling pathways that induces proliferation and differentiation in the cell.<sup>[49]</sup> Ultimately, this produces a clonally expanded pool of activated effector cells with a unique antigen specificity.

Activated  $CD8^+$   $T_C$  effector cells induce apoptosis in pathogenic cells by secretion of cytotoxic enzymes (perforin, granzymes) or expression of apoptosis-inducing receptors.<sup>[50-51]</sup> In contrast, activated  $CD4^+$   $T_H$  cells - including  $T_{H1}$ ,  $T_{H2}$ ,  $T_{H17}$ ,  $T_{FH}$  – have more diverse effector functions. They either help to activate APCs or macrophages to trigger phagocytosis of pathogenic cells or activate the B cell mediated humoral immune response that involves the production of soluble antibodies.<sup>[52]</sup>

Therefore, T cells and in particular the antigen recognition by their receptors are indispensable parts of the adaptive immunity. Thus, a huge TCR diversity is crucial to recognize a broad spectrum of antigenic peptides. It is created during T cell maturation in the thymus and involves several somatic rearrangement processes (Figure 5). In detail, germline DNA of the TCR  $\alpha$  and  $\beta$  loci comprises a distinct number of variable (V), diversity (D), joining (J) and constant (C) gene segments that are randomly rearranged in a process of V(D)J recombination.<sup>[53-54]</sup> Further, receptor diversity is enhanced by nucleotide deletion and the introduction of random nucleotides (N) into the junctions of the rearranged gene segments.<sup>[55]</sup> This is especially important for the hypervariable regions of the TCR variable domain that are directly involved in antigen-recognition. Whilst the complementarity-determining regions (CDR) 1 and 2 are binding to the MHC of another cell, the hypervariable CDR3 region is in direct contact with the peptide:MHC complex.<sup>[56]</sup> Altogether the combinatorial and junctional diversities in humans theoretically generates a TCR repertoire with up to  $10^{15}$  different  $\alpha:\beta$  TCRs.<sup>[23]</sup> So far only  $2.5 \times 10^7$  different TCR clonotypes could be experimentally detected in humans, which is linked to clonal selection processes during tolerance induction in the thymus.<sup>[57]</sup> The majority of TCRs

from this huge repertoire are distinct for each individual and shaped by personal genetically and environmental influences.



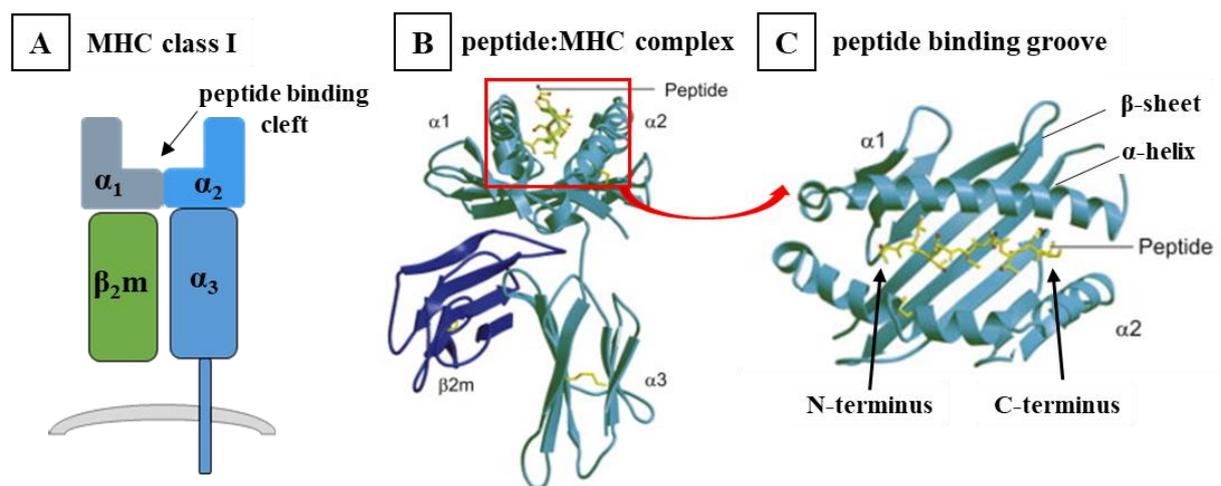
**Figure 5: Receptor rearrangement during lymphocyte development.** (A) Germline DNA of TCR $\alpha$  and  $\beta$ -genes are encoded in a different number of variable (V, green), diversity (D; only present in  $\beta$ -loci; orange), joining (J, yellow) and constant (C; red) gene segments. During maturation it is randomly rearranged by a step known as V(D)J recombination. This comprises the irreversible genetically recombination of gene segments and the junctional addition or deletion of random nucleotides (N) into the hypervariable complementarity-determining regions (CDR) of the TCR variable domain. Those mechanisms are crucial for generating the high variety of different antigen specificities. The rearranged DNA of TCR $\alpha$  and  $\beta$ -chains is further processed by transcription, splicing and translation to produce the a highly diverse  $\alpha\beta$  heterodimeric TCR protein with two unique antigen binding sites at the variable region and a distinct constant region. Each rearranged variable DNA sequence comprises the sequence information for three hypervariable complementarity-determining regions (CDR) that are crucial for antigen selectivity of the emerging TCR protein. Whilst CDR1 and CDR2 loops are encoded by sequences of the V gene segment, CDR3 loop is derived from the end of the V segment and the beginning of the J segment. (Adapted from MURPHY *et al.*)<sup>[43]</sup> (B) The total diversity of different lymphocyte receptors is created by combinatorial and junctional diversity. It is higher for TCR $\alpha\beta$  genes than for BCR IgH:L. However, BCR diversity is increased upon antigen encounter by shaping its specificity with affinity maturation mechanisms (CSR and SHM). (Adapted from JANEWAY *et al.*)<sup>[58]</sup>

### 1.1.4 Antigen recognition of T cells

T cells recognize antigens that are presented in a complex with MHC molecules on the cell surface of APCs such as macrophages, dendritic cells (DC) or B cells. Antigens are taken up by those APCs in the periphery, further processed in their cytosol and loaded on MHC glycoproteins that are transported to the cell surface. This process also induces maturation and activation of the APCs, which results in the expression of certain cytokines. In particular, it increases the expression of the chemokine receptor 7 (CCR7<sup>+</sup>) that is crucial to guide those cells home along a chemokine gradient. Ultimately, APCs reach their distinct lymph nodes via afferent lymphatic vessels to present their antigens to lymphocytes and initiate an effective adaptive immune response.<sup>[59-60]</sup>

CD8<sup>+</sup> T cells recognize fragmented antigens that are presented on MHC class I proteins. MHC class I may be expressed by all nucleated cells of the body beside of erythrocytes and is encoded in the three classical human leucocyte antigen (HLA) alleles A, B and C. Each of these alleles is present twice in each person and all are highly polymorphic throughout the population. This ensures an immense diversity to successfully defeat all kinds of potential pathogenic invaders and thus survival of the organism.<sup>[61-62]</sup>

MHC class I proteins are heterodimers that consist of a membrane-spanning  $\alpha$ -chain complex ( $\alpha_{1-3}$ ) and a non-covalently bound  $\beta_2$ -microglobulin ( $\beta_2m$ ) chain (Figure 6A-B).<sup>[63]</sup> The transmembrane domain of the  $\alpha_3$ -subunit is anchoring the complex in the cell membrane, whilst the  $\alpha_1$  and  $\alpha_2$  domains fold into a peptide binding cleft (Figure 6B-C). This binding groove is formed by two separated  $\alpha$ -helices and eight antiparallel  $\beta$ -sheets in which a peptide fragment can be embedded.<sup>[64]</sup>



**Figure 6: Antigen presentation by MHC class I molecules.** (A) Schematic structure of a MHC class I heterodimer. It consists of three  $\alpha$ -chains (different shades of blue) and a non-covalently attached  $\beta_2$ -microglobulin chain ( $\beta_2m$ , green). The molecule is anchored in the membrane of nucleated cells by the transmembrane-domain of the  $\alpha_3$ -chain. The peptide-binding cleft is formed by the subunits  $\alpha_1$  and  $\alpha_2$ . (B) Ribbon model of a MHC class I molecule with a bound peptide in the middle of the MHC binding groove. The three  $\alpha$ -subunits are displayed in light blue, the  $\beta_2m$  subunit in purple and the peptide in yellow. (C) Top view of a ribbon model of the peptide-binding groove that is formed by a MHC class I molecule (blue). A peptide is bound in

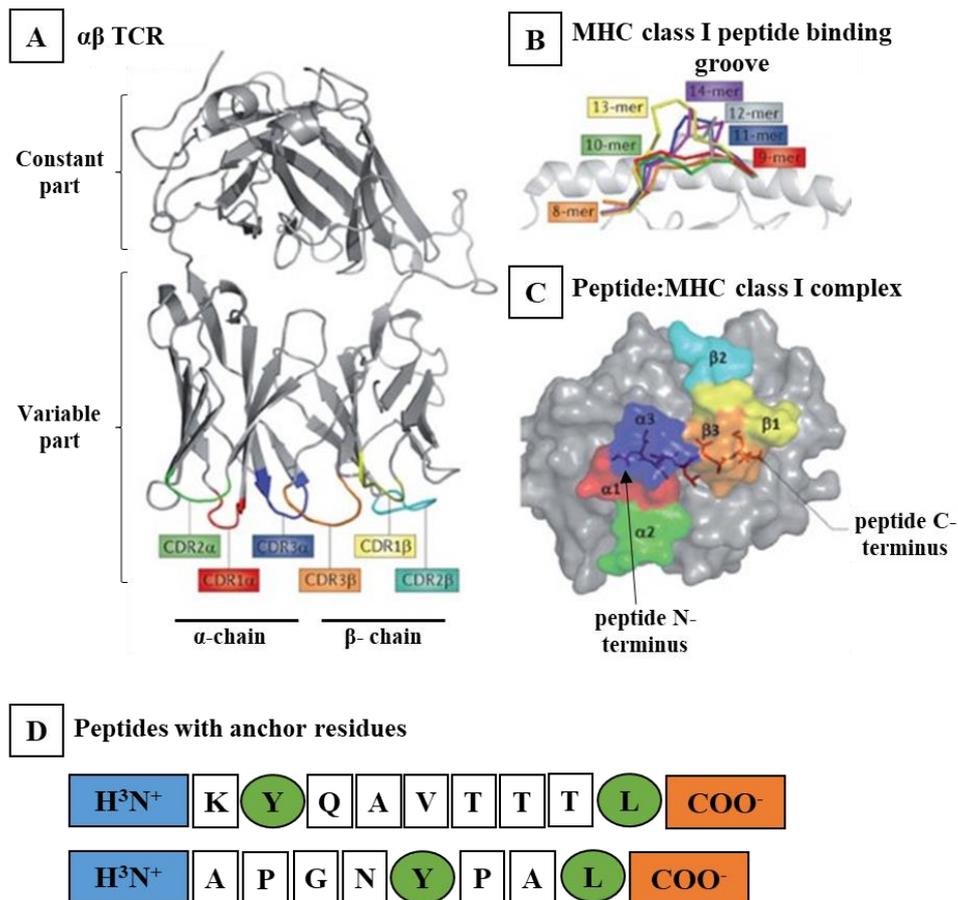
the groove (yellow). The binding groove is formed by the  $\alpha_1$  and  $\alpha_2$  domains that fold into two separated  $\alpha$ -helices on top of eight antiparallel  $\beta$ -sheets. The MHC class I molecule interacts with the bound peptide via ionic interactions and hydrogen bonds. (A adapted from JANEWAY *et al.*)<sup>[58]</sup> (B-C taken and modified from TAK WAH MAK *et al.*)<sup>[65]</sup>

Antigens are processed by APCs in the cytosol and loaded on top of this MHC class I complex. Processing involves cutting the antigenic peptide into small fragments of 8-14 amino acids (aa) by the cytosolic proteasome. Further, the fragments are loaded on a pre-MHC complex in the endoplasmic reticulum (ER) of the cell. The ER completes folding of the complex and enables the peptide-MHC complex to reach the cell surface. Notably, this complex is only stable and expressed on the cell in a peptide bound form.<sup>[66-68]</sup>

The peptide is thereby loaded onto the formed peptide-binding cleft of the  $\alpha_1$  and  $\alpha_2$  domains of the MHC molecule. It is spatially fixed by hydrogen bonds and ionic interactions between the amino and carboxyl ends of the peptide with the backbone of the MHC molecule (Figure 6B-C).<sup>[69]</sup> This binding pattern with a fixed peptide N- and C-terminus is known as closed-end conformation and characteristic for MHC class I molecules. Whilst shorter peptides are entirely embedded in the binding cleft, the centre of longer peptides is distorted and will bulge out of the cleft (Figure 7B).<sup>[70-71]</sup>

MHC molecules of different HLA alleles have distinct peptide binding preferences that vary with the germline encoded amino acid composition of their binding groove. Those allele-specific binding signatures differ from each other by certain peptide lengths and specific anchor positions. Anchor residues are amino acids within the peptide that don't necessarily have to be the same, but have to be similar in charge and size (Figure 7D).<sup>[72-73]</sup>

Peptides that are presented by MHC class I molecules in this way can be detected by CD8<sup>+</sup> T cells. Thereby antigens are recognized by the variable region of the  $\alpha$ : $\beta$  TCR that is formed by the CDR1-3 domains of the  $\alpha$  and  $\beta$  chain (Figure 7A). The CDR1 and 2 loops are in direct contact with the germline-encoded domains of the MHC:peptide complex, whereas the hypervariable CDR3 loop is binding to the non-germline encoded highly variable bound peptide (Figure 7C). The N-terminal end of the presented peptide is in contact with the CDR3 of the TCR  $\alpha$ -chain and the C-terminus with the TCR  $\beta$ -chain. Ultimately, TCR ligation induces structural changes in the T cell that activate the intracellular signalling cascade to trigger differentiation and activation of the CD8<sup>+</sup> T cell.<sup>[74-75]</sup>



**Figure 7: CD8<sup>+</sup> T cell antigen recognition of peptide:MHC class I complexes.** (A) Ribbon model of an  $\alpha\beta$  TCR depicting the orientation of the variable complementarity-determining regions 1-3 (CDR1-3) of each TCR chain. (B) MHC class I peptide-binding groove with bound peptides of different length (8- 14 aa). MHC class I molecules bind peptides at their ends in a 'closed-end' conformation. Due to these fixed points, larger peptides are bulging out of the binding groove and this effect increases with the number of amino acids in the peptide. (C) Top view of the peptide binding groove of a peptide:MHC class I complex bound to an  $\alpha\beta$  TCR. The variable CDR loops of the TCR are binding different regions of the presented peptide complex. CDR1 (red and yellow) and CDR2 (green and light blue) of both TCR chains are mainly interacting with the MHC molecule, whilst the hypervariable CDR3 (blue and orange) loop is in direct contact with the cognate peptide. In general, the N-terminus of the peptide is bound by the variable part of the  $\alpha$ -chain and the C-terminus is in contact with the TCR  $\beta$ -chain. (D) Shown are examples of peptides with MHC specific anchor residues. Each peptide sequence starts with the N-terminal amino group (blue), followed by a varying long sequence of amino acids – generally 8-14 aa - and terminates with the C-terminal carboxyl group (orange). Different MHC molecules prefer binding and presentation of peptides with a distinct length and specific anchor positions. Those anchor amino acids (green) do not need to be the same, but similar in size and charge. (A-C taken and modified from SEWELL *et al.*)<sup>[75]</sup> (D adapted from JANEWAY *et al.*)<sup>[58]</sup>

It is widely known that HLA types are associated with the susceptibility to certain diseases such as virus infections or autoimmune diseases.<sup>[76-78]</sup> In particular the autoimmune disease multiple sclerosis (MS) with a familiar accumulation of cases is suggesting a genetic involvement of HLA types in disease pathogenesis.<sup>[79]</sup> Genome-wide association studies of MS patients have demonstrated a strong genetic linkage of disease incidences with certain HLA loci.<sup>[80-81]</sup> Up to now, the most dominant correlation is located within genes of the MHC class II locus. It comprises the HLA-DR2 haplotype, which involves the alleles HLA-DRB1\*15:01 and HLA-DQB1\*06.<sup>[82-83]</sup> Further it was shown that MHC class I genes are also linked to disease susceptibility. HLA-A03:01 seems to increase the risk of developing MS, whilst HLA-A02:01 and HLA-C05 might have a protective role.<sup>[84-86]</sup>

### 1.1.5 Long-term immunity of T cells

Upon effective clearance of a pathogen, T cells differentiate into long-lived memory cells that are crucial for immune surveillance. Those cells get activated by a re-infection with the same antigen and elucidate a fast immune response that confers protective immunity. Memory T cells either migrate home to the lymph nodes as central memory T cells ( $T_{CM}$ ) or circulate in blood and the lymphatic system as effector memory cells ( $T_{EM}$ ). This migration is regulated by expression of the lymph node homing receptors CCR7 and sphingosine-1-phosphate receptor 1 (S1PR1). For a long time it was thought that all memory T cells leave their tissue of first antigen encounter to patrol throughout the rest of the body as  $T_{CM}$  or  $T_{EM}$  cells.<sup>[87-88]</sup>

Recently, a third subpopulation of memory cells was identified that permanently reside in the tissue of a previous pathogen infiltration. Those cells are known as tissue-resident memory cells ( $T_{RM}$ ) and can be seen as an immediate ‘frontline defence’ of the affected tissue that confers local immunity.  $T_{RM}$  cells were mainly found in tissues with direct contact to pathogens such as skin, lung or the intestines, but also in immune-privileged tissues as the brain.<sup>[89-90]</sup>

$T_{RM}$  cells are characterized as a unique subpopulation of memory cells that do not circulate in blood or the lymphatic system and exhibit a specific expression profile. In particular, expression markers that are crucial for lymphocyte homing and T cell egress as CCR7, S1PR1 and CD62L are down-regulated and tightly controlled.<sup>[91]</sup>

S1PR1 is an essential receptor for lymphocyte trafficking and particularly important to mediate T cell egress and recirculation in lymphoid organs.<sup>[92-93]</sup> This downregulation is controlled by different regulatory mechanisms and ensures tissue residency of the cells. First, CD69 that is also known as an early marker for lymphocyte activation, is upregulated in  $T_{RM}$  cells and this inhibits formation of S1PR1.<sup>[94]</sup> Second, the key transcriptional regulator homolog of Blimp-1 in T cells (Hobit) downregulates the expression of transcription factor krüppel-like Factor 2 (KLF2). KLF2 expression is essential for S1PR1 transcription, so no S1PR1 can be produced.<sup>[95-96]</sup> Whilst Hobit and CD69 are seen as key markers for the identification of  $T_{RM}$  cells, their unique phenotype vary in accordance with the particular microenvironment of the residing tissue.<sup>[97]</sup>

Besides of maintaining long-term immunity in the destined tissue,  $T_{RM}$  cells might also play a vital role in sustained inflammation of certain tissues and autoimmune processes.  $T_{RM}$  cells seem to be involved in propagating chronic inflammations in the autoimmune diseases psoriasis and rheumatoid arthritis.<sup>[98-99]</sup>

## 1.2 Autoimmune diseases

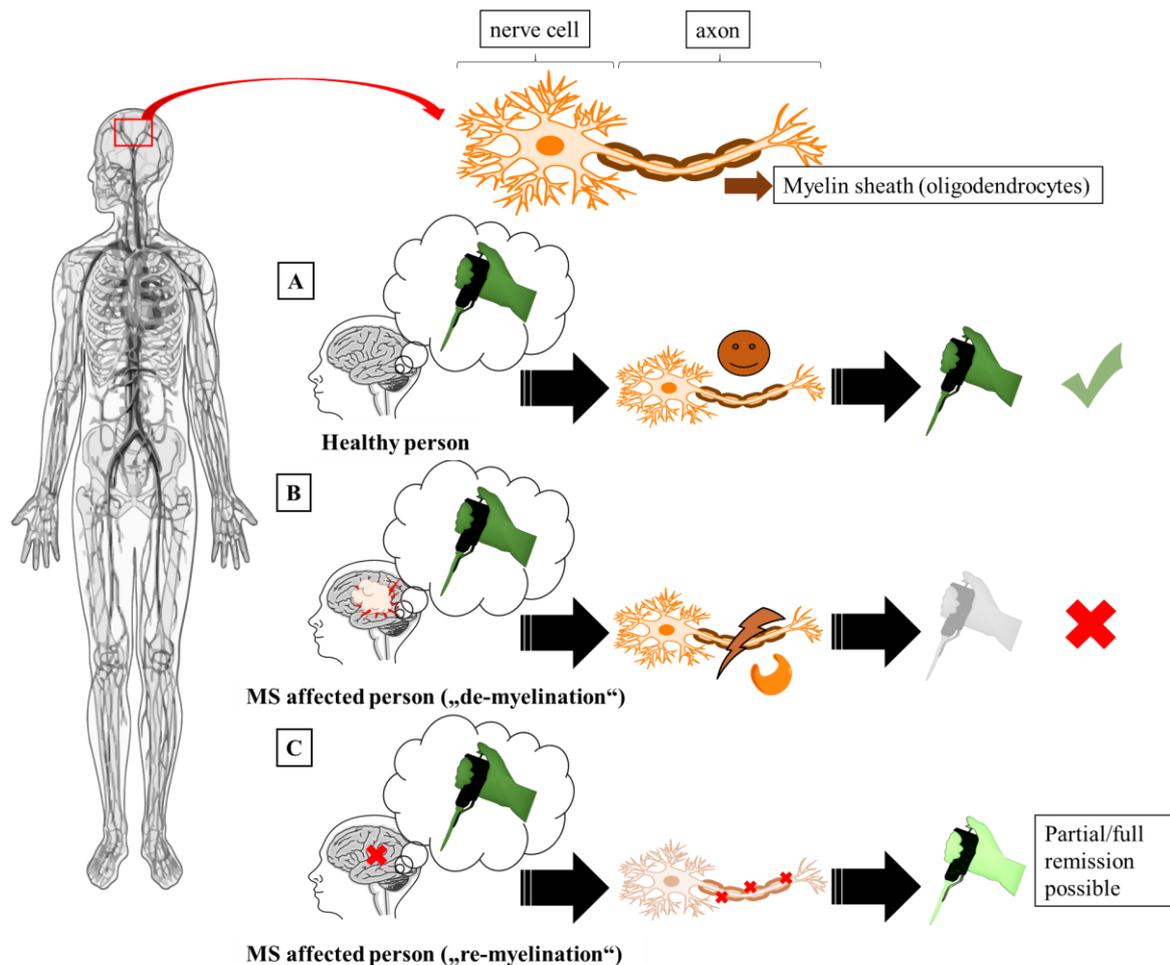
The mammalian immune system developed several regulatory mechanisms to ensure an effective differentiation of harmful pathogenic structures from their own body tissues. A sophisticated control system is essential for survival of the host organism and the occurrence of malfunctions were first pictured as “horror autotoxicus” by PAUL EHRLICH in the early 20<sup>th</sup> century.<sup>[100]</sup> In autoimmune diseases the discrimination between self and non-self is defective, so misguided immune cells trigger an aberrant immune response against their own body tissues.<sup>[101]</sup> Those immune cells are termed ‘auto-reactive’ and are usually deleted from the immune repertoire by passing several developmental checkpoints. Lymphocytes with a receptor that binds too strong or too weak to presented self-antigens are directly removed from the repertoire during their maturation in the thymus (1.1.3).<sup>[20, 102]</sup> Beside of this central tolerance mechanism, tolerance is also induced in mature lymphocytes in the periphery and causes either deletion, suppression by regulator cells or the induction of anergy (a state of functional unresponsiveness).<sup>[103]</sup>

Depending on the affected regions, autoimmune diseases can be clinically classified into systemic diseases that affect the entire organism (e.g. systemic lupus erythematosus) or into local diseases that affect specific organs or tissues of the body such as in diabetes type-1 or MS. In general, autoimmunity arises prevalently in young adults and several factors including genetic predisposition and environmental triggers as well as infections might play a role in their pathogenesis.<sup>[104-106]</sup> Current treatments primarily focus on medicating the devastating symptoms that result as a consequence of the occurring immune attacks. This emphasizes the relevance of a deeper understanding of their disease pathogenesis, which is crucial to develop treatments that are more effective and might be able to cure those diseases in the future.

### 1.2.1 Multiple sclerosis

MS is a chronic inflammatory disease of the central nervous system (CNS) and was first described by JEAN-MARTIN CHARCOT in the late 19<sup>th</sup> century.<sup>[107]</sup> It is affecting 0.1% of the global population and primarily young females between the age of 20-40 years.<sup>[108]</sup> MS is pathologically characterized by neuronal damage, axonal loss and demyelination in cerebrospinal fluid (CSF) and brain tissues.<sup>[109-110]</sup>

During the course of the disease, lesions are enriching that are formed by de-myelinated regions and scar tissues. Those lesions are known as ‘plaque’ and affecting neural networks as well as signal transduction processes in the body. Ultimately, this leads to disease progression and physiological disabilities (Figure 8).



**Figure 8: Schematic representation of signal transduction processes in the human body.** (A) The human brain is connected to all cells of the body (neurons, muscles, glands) by a diverse and complex signalling network of nerve cells. This comprises neurons with a cell body and long branches of axons that transmit these signals efficiently and fast throughout the body. In healthy individuals those axons are protected by a layer of oligodendrocyte cells – the myelin sheath - that insulate the axon and ensures a fast transmission of electrical impulses. (B) In MS patients, inflammatory processes damage this insulating myelin sheath. Demyelination decreases or even prevents the transduction of signals and can cause neurological symptoms. (C) In non-inflammatory phases of the disease the body is able to partially or fully restore signal transduction (“re-myelination”). The relapsing-remitting course of MS is characterized by alternating episodes of de- and remyelination. Yet each acute attack leaves scars in the remaining brain tissue. The accumulation of de-myelinated axons and scar tissue in the brain ultimately lead to progressive deterioration of signal transduction and increasing disabilities during the disease course.

Clinically, lesion formation can be detected by magnetic resonance imaging (MRI) and is associated with the occurrence of so-called “relapses”. Typical symptoms of MS are vision problems, sensory disturbances, paralysis or mobility impairments and they appear to correlate with affected regions of the brain. Usually, MS starts with a first episode of clinical symptoms, which is defined as clinical isolated syndrome (CIS). OKUDA *et al.* were even able to define an earlier onset of the disease that is known as radiologically isolated syndrome (RIS). It is characterized by MRI anomalies, but without clinical symptoms.<sup>[111]</sup> Although each disease course is highly variable, the most common forms of MS are primary-progressive MS (PPMS) and relapsing-remitting MS (RRMS). RRMS is characterized by alternating phases of clinical attacks (‘relapses’) and remissions. Some RRMS patients with a longer disease duration convert to a secondary progressive MS (SPMS) disease course. In SPMS, relapses seem to play a minor

role and it is characterized by progressive worsening of symptoms and remaining disabilities.<sup>[112-113]</sup>

Whilst the focus of current research is addressing the complex immunological processes during the course of the disease, the etiology and pathogenesis of MS remains enigmatic. Several risk parameters have been discussed that might contribute to disease susceptibility, including genetic predisposition (e.g. higher familial risk, correlation with HLA genes), environmental factors, stress and viral infections.<sup>[114-117]</sup> Since the risk for developing MS is enriched in genetically predisposed individuals, it appears advantageous to study the disease in homozygotic twins with discordance for MS. This not only removes confounding factors as genetically heterogeneity, age and gender, but also reduces the effect of environmental influences to a minimum as twins were raised in the same households. Further, “healthy” co-twins of genetically identical individuals with manifest MS have the highest familial risk for developing MS during their lifetimes. To this end, the group of LISA GERDES (Institute of Clinical Neuroimmunology, LMU Munich) was able to collect a unique cohort of - up to now - 85 genetically identical homozygotic twin pairs with discordance regarding their MS disease status. Moreover, this unique cohort was already used for several twin studies as comparing gut microbial compositions and epigenetics.<sup>[118-119]</sup>

However, as long as the factors that trigger the outbreak of MS are still not entirely understood, it remains an incurable disease. Thus, currently available treatments focus on reducing the tremendous inflammatory side effects of the disease but have only limited clinical efficacy.<sup>[120]</sup>

### **1.2.2 Immune cells in MS**

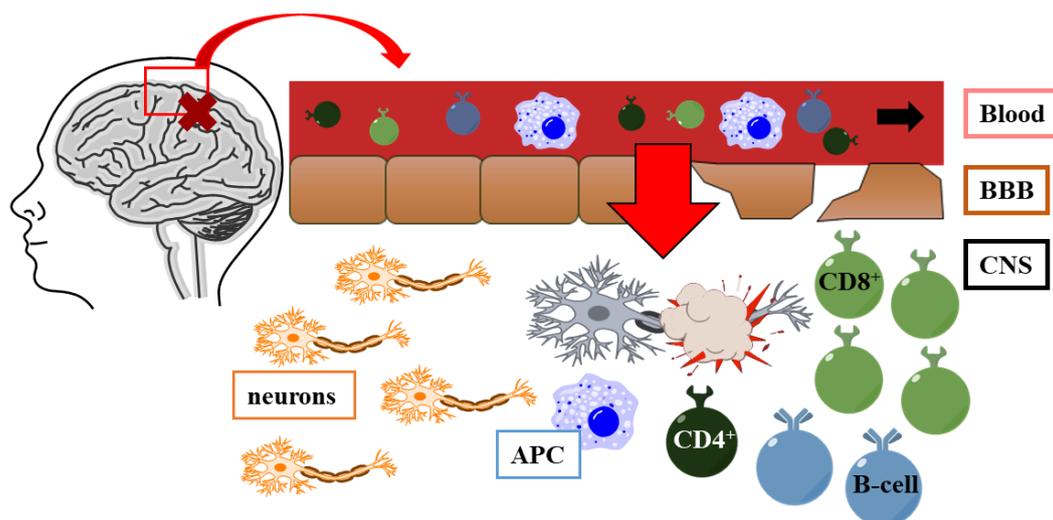
MS is histologically characterized by multifocal inflammation, lesion formations (‘plaques’) in the white matter of the brain, destruction of myelin sheaths and the loss of oligodendrocytes.

In general, the human brain is an immune-privileged area of the CNS and spatially protected from invading pathogens and the influx of harmful cells by the so-called blood-brain barrier (BBB; Figure 9).<sup>[121]</sup> Preserving the precious neuronal networks of the CNS is crucial for the survival of the organism and thus tightly controlled by the BBB. It comprises a semi-permeable layer of endothelial cells that allows selective transport of nutrients (e.g. water, ions, and glucose) and prevents the passage of pathogens or white blood cells. During inflammatory conditions, it gets more permeable to allow immune cells like phagocytes or lymphocytes to enter the CNS and clear the putatively harmful infection. This ‘disrupted’ BBB and the perivascular infiltration of immune cells including activated lymphocytes (B-cells, plasma

cells, CD4<sup>+</sup> and CD8<sup>+</sup> T cells) and APCs (macrophages) into the CNS can be seen in MS patients.<sup>[122-123]</sup>

Because of the strong genetic linkage of MS with the HLA-DR2 haplotype (1.1.4), MS was long associated as a MHC class II driven disease. In parallel, most of the current MS animal models like experimental autoimmune encephalomyelitis (EAE) are based on myelin-reactive CD4<sup>+</sup> T cells, as CD4<sup>+</sup> T cells – but not CD8<sup>+</sup> T cells – may easily be expanded *in vitro* and thus used to induce EAE in mice.<sup>[124]</sup> However, activated CD8<sup>+</sup> T cells are representing the dominating T-lymphocyte population in chronic, active MS lesions and outnumber CD4<sup>+</sup> T cells by a factor of two to three.<sup>[125-127]</sup> Further, clonally expanded CD8<sup>+</sup> T cells from brain tissues could be also tracked in CSF and blood of two MS patients over several years, thus implicating a putatively relevance in disease progression.<sup>[128]</sup>

Moreover, expanded plasma cells may be found in brain tissue and the CSF of MS patients and are a hallmark of MS as well as a marker for chronic inflammation. They produce oligoclonal band antibodies (OCBs), which consist mainly of clonally expanded IgG<sup>+</sup> plasma cells that underwent affinity maturation and are persisting in the CSF.<sup>[129-130]</sup> OCBs are a major diagnostic tool and used as a key clinical criteria in diagnosing MS.<sup>[131]</sup> Around 40% of MS patients have also OCBs with expanded IgM<sup>+</sup> plasma cells that underwent SHM.<sup>[132]</sup> The role of those antigen-experienced plasma cells during the disease course is still not entirely clear. Recent findings suggest they might recognize intracellular autoantigens and thus are involved in removing cell debris that is released during tissue destruction in the inflamed CNS.<sup>[133]</sup>



**Figure 9: Schematic representation of immune cell infiltrations into the CNS of MS patients.** The human brain is an immune-privileged area that is physically protecting its precious neuronal structures (orange) from the infiltration of pathogens by the blood-brain barrier (BBB, brown). The BBB is a semi-permeable layer of endothelial cells that regulates the exchange of essential nutrients (e.g. glucose, ions or water) from blood and the CNS, but prevents the influx of pathogens or immune cells into the brain. Inflammations (marked with a red X) within the CNS can open up this tightly protected barrier, so immune cells are able to enter and control those harmful inflammatory conditions (big red arrow). MS is a neuroimmunological disease of the CNS that is characterized by chronic inflammation and neuronal loss (grey). Due to a 'leaky' BBB, putatively harmful immune cells can enter the CNS and trigger the destruction of neuronal myelin sheaths. Though the pathogenesis of MS is not clear yet, the infiltrating immune cells that might be involved in tissue destruction were identified as APCs (macrophages, dark

blue) and activated lymphocytes (B-cells, light blue; CD4<sup>+</sup> and predominantly CD8<sup>+</sup> T cells, dark and light green). Gaining a deeper understanding of the complex immune mechanisms that lead to the infiltration of lymphocytes and especially unravelling the role of the activated CD8<sup>+</sup> cells, that represent the dominating population in MS lesions, is crucial to shed light on the pathogenesis.

### 1.3 Analysing immune repertoires

The identification and characterization of pathogenic immune cells in autoimmune diseases such as MS is crucial for a better understanding of the underlying immune processes and the first step to detect target antigens.<sup>[134]</sup> Thereby it is essential to distinguish disease-relevant lymphocytes from bystander cells that are only attracted by the inflammatory environment of the immune response. Key criteria are close proximity, the expression of activation markers such as CD69 and an increased frequency of uniform lymphocyte clones.<sup>[94]</sup> In general, T and B cells get activated by recognition and binding of cognate antigens, which ultimately lead to proliferation and expansion of lymphocyte clones with the same receptor specificity (1.1.3). In particular, the CDR3 regions of the TCR and BCR variable domains that are in direct contact with the cognate antigen can be tracked as a unique identifier for clonal expansion (1.1.4). As clonal expansions result from antigen-driven processes, it is a well-established indicator for putative pathogenicity and emphasizes the importance of immune repertoire studies.<sup>[135]</sup>

However, immune repertoire approaches have to cope with the huge diversity of each individual's antigen receptor repertoire that was estimated to comprise  $2.5 \times 10^{18}$  different TCR combinations.<sup>[23, 136]</sup> As the genetically encoded diversity of TCR  $\beta$  chains is higher than for  $\alpha$  chains, widely used bulk approaches were primarily focussing on the investigation of the TCR  $\beta$  chain repertoire.<sup>[58]</sup> Further, a mechanism known as allelic exclusion that occurs during somatic rearrangement of the TCR  $\beta$  chain prevents the production and expression of an additional second  $\beta$  chain.<sup>[137]</sup> In contrast, dual  $\alpha$  chains are commonly known and thus the evaluation of clonal expansion based on the  $\beta$  chain repertoire was more accurate and reliable.<sup>[137]</sup>

Initial attempts in determining the complexity of the TCR  $\beta$  repertoire in humans were performed on the protein level by using flow cytometry, but they were limited by the availability of antibodies against variable regions of TCR  $\beta$  chains.<sup>[138]</sup> In addition, flow cytometry based methods were lacking deeper sequence information of the CDR3 diversity of the TCRs, which is crucial to identify disease-relevant clones.<sup>[139]</sup> Further developments in the field of repertoire sequencing were establishing CDR3 spectratyping for assessing TCR expansions.<sup>[138, 140]</sup> It is a transcriptome based approach that uses PCR amplification of TCR variable genes and subsequent electrophoretic analysis to reveal length distributions of CDR3 regions of TCR  $\beta$

chains. In healthy person with a heterogeneous polyclonal immune repertoire, the CDR3 length distribution is reflected by a bell like pattern. Deviations from this Gaussian distribution are associated with oligoclonal expansions. With this, it was possible to detect individual expanded clones in blood, CSF and brain tissue of MS patients.<sup>[127-128]</sup>

However, CDR3 spectratyping gives only a first overview on the overrepresented TCR  $\beta$  families and has a low resolution. To reveal CDR3 sequences on a nucleic acid level, the combination with conventional DNA sequencing methods such as Sanger sequencing was required.<sup>[141]</sup> Further, the identification of both matching  $\alpha:\beta$  TCR and H:L Ig chains is crucial for subsequent antigen screening experiments and this emphasizing the necessity for single cell (SC) based methods. By combining CDR3 spectratyping with SC analysis, JUNKER *et al.* were able to identify CD8<sup>+</sup> T cell infiltrations and clonal TCR expansions in active MS brain lesions.<sup>[142]</sup> In contrast to bulk repertoire analysis, the used SC methods have still been laborious (e.g. manual single cell isolation) and for a long time limited by either sufficient sensitivity or throughput.<sup>[138, 143]</sup> In particular, the analysis of rare human samples with relatively low cell numbers such as CNS tissue or CSF was still a major technical challenge.

A huge breakthrough in repertoire studies occurred with the development of high throughput sequencing (HTS) tools such as ROCHE 454 pyrosequencing and ILLUMINA next-generation sequencing (NGS).<sup>[144-145]</sup> HTS enables massive parallel sequencing of samples to low costs, but requires pre-enrichment of the target DNA. First approaches used multiplex PCR to enrich the particular antigen-receptor variable families with a pool of primers from genomic DNA.<sup>[138]</sup> Since genomic DNA comprises long stretches of introns and this increases the probability for sequencing errors and PCR artefacts, it is not an ideal target for HTS analysis. Thus, later repertoire sequencing analysis were targeting and enriching full-length mRNA transcripts by using 5' rapid amplification of cDNA ends (RACE) and anchor PCR.<sup>[146-147]</sup> Finally HTS methods enabled the unbiased identification of TCR and BCR CDR3 nucleotide sequences in peripheral blood, CSF and active brain lesions of MS patients.<sup>[148-151]</sup>

Even with the new developed HTS approaches, it was not possible to identify pathogenic T-cells and their target antigens in MS. First, it might be that the used repertoire methods have so far been not sensitive enough to depict the full immune repertoire in patient's samples. Further, the selection criteria for disease relevancy might have been too biased or the focus too selective and thereby missing out relevant lymphocytes. In addition, to gain a deeper understanding of the complex immunological mechanisms that occur during MS, it is crucial to simultaneously analyse the TCR and BCR repertoires in the same specimen. Especially for the analysis of precious human material such as in CNS specimens, this was a huge technical limitation. For

example, flow cytometry as the method of choice for high throughput applications is useful to reveal a realistic picture of blood samples. In contrast, when samples with a low lymphocyte cell count such as CSF are used, particularly the rare subpopulations of cells are lost and thus the overall picture of relevant cells is skewed. Therefore, further technical improvements that can handle samples such as CSF whilst maintaining the throughput of current repertoire analysis methods have to be developed.

In order to identify disease-relevant cells, a greater sequencing depth that comprises not only the immune repertoire, but also link it to its gene expression profile e.g. by combining it with RNA-sequencing tools would be beneficial. This would not only reveal further information about the activation status of interesting clones but also would make it possible to phenotypically cluster subpopulations and reveal a bigger picture of immunological relevant mechanisms. Possibly this enables the identification of disease-relevant immune cells and their antigen-specific receptor with a greater certainty and might lead to the identification of a target MS antigen.

#### **1.4 RNA-sequencing**

RNA-sequencing (RNA-seq) is a NGS technology with unprecedented power that reveals the whole transcriptome of single cells and thus enables their in-depth characterization.<sup>[152]</sup> Compared to the huge complexity of the human genome that remains a major challenge of genome-based approaches or the limited informational value of protein-based techniques, analysing RNA transcriptomes offer the unique opportunity to study current cellular processes and phenotypical changes in cells as cell cycle status or activation profiles. In general RNA-seq comprises the following steps: cell isolation by a high-throughput method of choice e.g. flow cytometry, RNA isolation e.g. by targeting all pA<sup>+</sup> mRNA molecules, generation of more stable double-stranded cDNA by reverse transcription, followed by a fragmentation and labelling step, and finally sequencing as well as bioinformatics analysis by mapping the data to a reference genome. Although conventional Sanger sequencing reproduces reliable DNA sequencing results with up to several hundreds of nucleotides length, it is laborious and time-consuming and therefore not suitable for such HTS applications.<sup>[153]</sup> In contrast, NGS based Illumina sequencing is technically limited by short sequencing reads, but suitable to analyse huge amounts of different cDNA sequences simultaneously. This enables massively parallel sequencing on Illumina chips in a much shorter time and with reduced costs. Those short cDNA sequences are produced in a process known as tagmentation that produces fragmented, uniquely labelled cDNA libraries ready for sequencing. In the end, the enormous amount of sequencing

information that is generated by RNA-seq technologies has to be analysed by complex bioinformatics tools to extract the desired information and remains a major technical challenge.

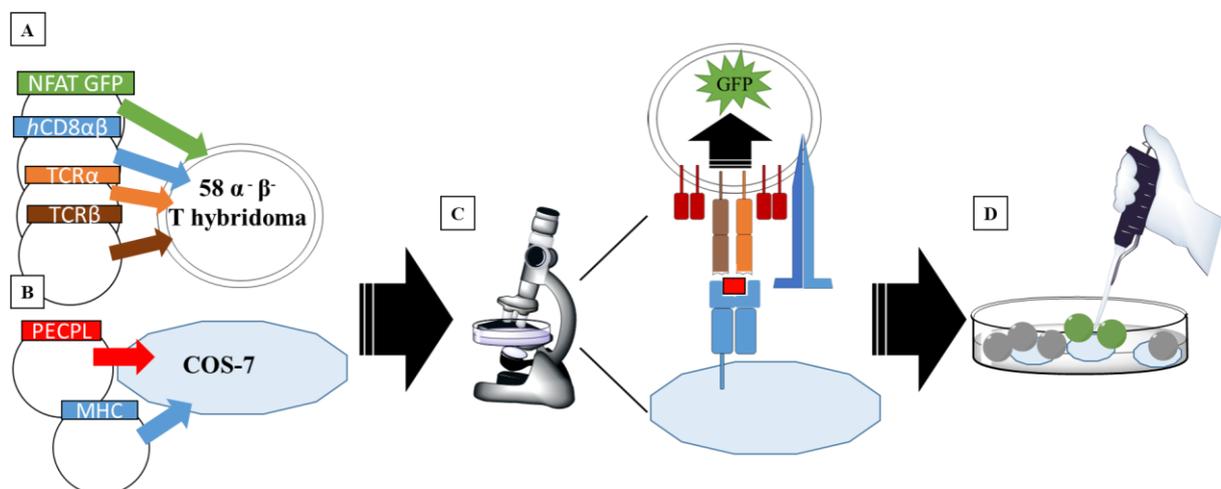
### 1.5 Detection of T cell antigens

T cells play an important role in infectious diseases, tumours and autoimmunity, but their relevant target antigens are largely unknown.<sup>[154]</sup> The identification of T cell antigens is especially important to improve therapeutic treatments and for the development of new vaccines. Further, in autoimmune diseases like MS, detecting the recognized antigens that trigger the presumable pathogenic CD8<sup>+</sup> T cells would reveal tremendous new insights to the disease pathogenesis.<sup>[155]</sup>

Antigen detection of T cells is characterized by complex mechanisms of antigen processing and presentation that are still technical challenging for current T cell profiling methods.<sup>[156]</sup> First, T cells recognize and bind peptide:MHC complexes that are presented to them by APCs with low affinity.<sup>[157]</sup> Those weak interactions make it difficult for affinity-based purification methods that have been used for instance to identify target B cell antigens.<sup>[133, 158]</sup> Second, the TCR diversity is huge and was generated to identify a broad spectrum of possible target antigens (1.1.3). Thus, T cells are polyspecific and can be activated by multiple different short peptide fragments.<sup>[159]</sup> Further, the intracellular processing of antigens, which includes post-translational modifications and their presentation, is MHC dependent.<sup>[160]</sup> It was also shown that the stereochemistry of antigen presentation affects T cell binding.<sup>[161]</sup> Lastly, the huge diversity of possible antigens is even enhanced by the polygenic nature of the HLA and its distinct peptide binding preferences (1.1.4). Altogether, this leads to a low predictability of native-like antigens and emphasizes the need of unbiased large-scale identification methods.

The first T cell profiling technologies were using cDNA libraries that were transfected into melanoma cells and subsequently tested for CD8<sup>+</sup> T cell activation. With this, they were lacking *in vivo* occurring intracellular processing steps, which was complicating the detection of native-like antigens.<sup>[162]</sup> Others were using short synthesized combinatorial peptide libraries with certain peptide lengths and partly fixed anchor positions. However, those positional scanning methods were still missing native-like intracellular antigen processing.<sup>[163-165]</sup> Additionally the ratio of activating to non-activating peptides of those methods is very low and made the identification of cognate antigens challenging. Hence, this led to the development of unbiased plasmid-encoded combinatorial peptide libraries (PECPL), in which T cell activation was investigated *in vitro* by PECPL transfected APCs in different model systems such as baculovirus or yeast.<sup>[161, 166]</sup>

Based on those PECPL approaches, KATHERINA SIEWERT established an unbiased method for the identification of presumably disease-relevant T cell antigens in 2012 (Figure 10).<sup>[167]</sup> The method combines PECPL with a novel single-cell detection system and was first validated in a well characterized TCR antigen model of the influenza virus. It was then successfully applied for the identification of target CD8<sup>+</sup> T cell candidate antigens of a patient with psoriasis vulgaris and MS.<sup>[168-169]</sup> Clonally expanded and thus potentially disease-relevant TCRs were first identified in patient material and then recombinant expressed in a hybridoma tumour cell line. Those 58  $\alpha\beta^-$  cells express neither a TCR  $\alpha$ - nor  $\beta$ -chain and were transfected with the cloned target TCR of interest along with plasmids that encode for the TCR co-receptors *mCD3 $\epsilon$*  and *hCD8 $\alpha\beta$* . As a readout for successful antigen binding that leads to activation of the T cell, the transcription factor NFAT was fused to a GFP reporter protein (1.1.3). Antigen-presentation of the PECPLs was mediated by an adherent COS cell line that was pre-transfected with the particular patient specific MHC molecules. Plasmid libraries with a defined length, but random amino acid constitutions were used for transfection of those COS cells. Ultimately, co-culturing of both cell lines combined with capillary micromanipulation under the fluorescence microscope makes this unbiased approach a powerful tool to discover unknown T cell antigens with a high sensitivity. Compared to other methods, the method overcomes low-affinity binding of TCR and peptide:MHC complexes by simple gravity. Further, the enrichment of a potential activating candidate antigen by micromanipulation out of a pool of millions of plasmids is enormous and speeding up the so far laborious PECPL based antigen search strategies.



**Figure 10: Schematic presentation of the PECPL based method for the unbiased identification of CD8<sup>+</sup> T cell antigens.** (A) 58  $\alpha\beta^-$  T hybridoma cells were stably transfected with the following plasmids: (1) sGFP under control of the NFAT promoter (green), (2) human CD8 $\alpha$ - and  $\beta$ -chains (blue), (3) the potentially disease-relevant TCR $\alpha$ - and  $\beta$ -chains (orange and brown). (B) Adherent COS-7 cells were used as APCs and transfected with plasmids encoding for the randomized PECPL with a defined amino acid length (red) and the patient-specific MHC molecule (blue). (C) Co-culturing of both cell lines and analysis under a fluorescence microscope reveals productive T-cell activation by fluorescent green lighting up hybridoma cells. (D) The APC underneath of the green fluorescing T hybridoma cells bears the potential antigen-coding plasmid and can be isolated by micromanipulation with a capillary. (modified from SIEWERT *et al.*)<sup>[167]</sup>

---

However, the success of the PECPL based assay methodically depends on two critical key parameters: the *in vitro* reconstitution of functional CD8<sup>+</sup> T cells and the availability of APCs that are presenting candidate antigens. On the T cell side, the presence of all stably transfected plasmids can be monitored straightforward by flow cytometry analysis of cell surface expressions and enables the selection of a suitable T cell clone with high activation potential for co-culture experiments. In contrast, peptide:MHC complexes in the MHC class I presentation pathway solely reach the cell surface of APCs (e.g. COS-7 cells) when antigenic peptides are bound and is crucial for the outcome of prospective co-culturing experiments. Since antibodies against human HLA-antigens are still limited, MHC surface expression and thus the successful presentation of the randomised PECPLs in peptide:MHC complexes, can hardly be monitored with the currently existing antigen search method.<sup>[170]</sup>

## 2 Aims

Though the importance of the adaptive immune systems in MS is well known and expanded lymphocytes were detected in the CNS of patients, their cognate antigens could not be identified yet. This might result from not sensitive enough current repertoire analysis methods that were not able to reliably identify putatively disease-relevant immune cells in specimens of MS patients. However, the identification of target MS antigens would provide tremendous new insights into disease pathogenesis and could lead to the development of new therapeutic targets in the future.

To address this and identify pathogenic immune cells in MS, the major aim of this work is to establish a suitable high throughput technique that enables the identification and molecular characterization of lymphocytes along with their antigen-specific receptors in SC from human CSF. In preliminary experiments that were performed in the group of K. DORNMAIR in 2016, it was already possible to establish and optimize the whole-transcriptome Smart-seq2 approach of PICELLI *et al.* in human PBMCs.<sup>[147, 171]</sup> Additionally, it was possible to successfully test and validate the first steps of the method in frozen, human CSF of a test patient with a CIS diagnosis (index patient #4526).

Hence, the objective of this work is to first identify disease-relevant lymphocyte clones in CSF of MS patients and then investigate their target antigens:

- 1) Continuing the validation of the newly established method in the CSF of index patient #4526: At first, this involves optimizing the NGS preparation steps, particularly the costly tagmentation reaction, for high-throughput applications. Further, samples of a whole 96-well plate shall be prepared for a NGS Hiseq test run and subsequent bioinformatics analysis of the whole transcriptome RNA seq data validated.
- 2) Applying the method to frozen CSF samples of a uniquely chosen cohort of monozygotic twins with discordancy regarding their MS diagnosis.
- 3) Subsequently, clonally expanded immune cells with a particular interesting activation status shall be chosen, further characterized and their functional roles investigated.
- 4) The putatively disease-relevant CD8<sup>+</sup> T cell clone TCR TRAV17\*01/TRBV4-1\*01 shall be *in vitro* reconstituted and used to identify potential target antigens with already established unbiased antigen search protocols.<sup>[133, 167]</sup> Prior, the peptide:MHC presentation performance from COS-7 cells used in the T cell antigen search method of KATHERINA SIEWERT shall be optimized. This includes the establishment of a new BFP-HLA construct that enabled the production of stable HLA-expressing COS cells and optimization of the used FuGene transfection conditions.

### 3 Materials and Methods

#### 3.1 Materials

##### 3.1.1 Devices

**Table 1: Used devices**

Device	Model	Company
Agarose gel documentation	Molecular Imager Gel Doc XR	BIO-RAD
Agarose gel electrophoresis chamber	Pharmacia LKB-ECPS 3000/150	AMERSHAM PHARMACIA BIOTECH
Agarose gel visualization	LED Blue light Transilluminator (UVT-22 BE-LED)	HEROLAB
	System UV-1600PC	VWR
Automated electrophoresis system	Bioanalyzer 2100	AGILENT TECHNOLOGIES
Autoclave	DX-150	SYSTEC
Centrifuges	Heracus™ Megafuge™ 1.0 R	THERMO FISHER SCIENTIFIC
Flow cytometer	Aria III cell sorter	BD BIOSCIENCE
	FACSVerser™	BD BIOSCIENCE
<b>Fluorescence microscope</b>	Axio Vert 200 M	ZEISS
Objective	Plan-Apochromat® (5x/0.16, 10x/0.45, ∞/0.17, 20x/0.8)	ZEISS
CCD Camera	CoolSNAP HQ	PHOTOMETRICS
Electroporation system	Gene pulser	BIO-RAD
Heating block	Thermomixer comfort	EPPENDORF
Incubator shakers	HT Multitron II	INFORS AG
pH-meter	pH 521	WTW
Thermal cycler	Mastercycler® Pro	EPPENDORF
UV/Vis Spectrophotometer	Nano-Drop 2000 Spectrophotometer	THERMO FISHER SCIENTIFIC
	Qubit 3.0 Fluorometer	THERMO FISHER SCIENTIFIC
UV-Transilluminator	BL-312T	BIOS CORPORATION

##### 3.1.2 Chemicals and consumables

All chemicals, except where otherwise specified were purchased from commercial companies like MERCK and SIGMA-ALDRICH and used without further purification. Consumables such as pipettes, tips and reaction tubes were procured from BIOZYM and EPPENDORF.

### 3.1.3 Buffers and Solutions

The double-deionized water (*ddH*<sub>2</sub>O) was obtained by purification with a Milli-Q® system (resistivity 18.2  $\frac{M\Omega}{cm}$ ; MERCK).

**Table 2: Composition of the used buffers and solutions**

TBE buffer (5×)	900 mM Trizma Base
	900 mM Boric acid
	2 mM EDTA Titriplex
	in <i>ddH</i> <sub>2</sub> O
Freezing buffer	10% (v/v) DMSO
	in FCS (heat inactivated, sterile filtered**)
Stop solution	10% (v/v) FCS
	in RPMI-1640 medium with glutamine
FACS buffer*	1% (v/v) FCS
	in 1× PBS (pH 7.4, THERMO FISHER SCIENTIFIC)
RPMI-1640 co <sup>10</sup>	10% (v/v) FCS (heat inactivated**; sterile filtered)
	1% (v/v) Sodium pyruvate
	1% (v/v) PenStrep
	1% (v/v) Non-essential amino acids
	in RPMI-1640 medium with glutamine
* was sterilized by filtration through Millipore Express® PLUS membrane filter (0.22 μm)	
**heat inactivation: 30 min, 56°C	

### 3.1.4 Antibodies

**Table 3: Antibodies for extracellular staining**

Specificity	Fluorochrome	Clone	Isotype	Company	Dilution
SYTOX® Blue Dead Cell Stain	$\lambda_{\text{emission}}$ : 480 nm $\lambda_{\text{excitation}}$ : 444 nm	-	-	INVITROGEN	1:1000
Propidium iodide (PI, viability dye)	$\lambda_{\text{emission}}$ : 632 nm $\lambda_{\text{excitation}}$ : 493 nm	-	-	SIGMA ALDRICH	1:500
<b>Non-CD4<sup>+</sup> panel (SC RNAseq)</b>					
anti-human CD197 (CCR7)	PE	G043H7	Mouse IgG2a, κ	EBIOSCIENCE	1:20
anti-human CD56 (NCAM)	FITC	HCD56	Mouse IgG1, κ	BIOLEGEND	1:40
anti-human CD19	APC*Fire	HIB19	Mouse IgG1, κ	BIOLEGEND	1:40
anti-human CD27	BV605	O323	Mouse IgG1, κ	BIOLEGEND	1:40

anti-human CD38	Blue (eF450)	HB7	Mouse IgG1, κ	EBIOSCIENCE	1:40
anti-human CD24	APC	eBioSN3	Mouse IgG1, κ	EBIOSCIENCE	1:50
anti-human CD3	AF700	OKT 3	Mouse / IgG2a, κ	EBIOSCIENCE	1:50
anti-human CD8	PerCP	SK1	Mouse IgG1, κ	BIOLEGEND	1:50
Fc-Receptor Blocking, human				MILTENYI BIOTEC	1:50
<b>CD4<sup>+</sup> panel (SC RNaseq)</b>					
anti-human CD45RO	FITC	UCHL1	Mouse / IgG2a, κ	EBIOSCIENCE	1:40
anti-human CD197 (CCR7)	APC	3D12	Rat / IgG2a, κ	EBIOSCIENCE	1:40
anti-human CD127	AF700	eBioRDR5	Mouse IgG1, κ	EBIOSCIENCE	1:40
anti-human CD25	Blue*BV421	BC96	Mouse IgG1, κ	BIOLEGEND	1:40
anti-human HLA-DR	APC*Fire	L243	Mouse IgG2a, κ	BIOLEGEND	1:40
anti-human CD196 (CCR6)	BV605	G034E3	Mouse IgG2b, κ	BIOLEGEND	1:40
anti-human CD183 (CXCR3)	PerCP	G02547	Mouse IgG1, κ	BIOLEGEND	1:20
anti-human CD185 (CXCR5)	PE	MU5UBEE	Mouse / IgG2b, κ	EBIOSCIENCE	1:40
Fc-Receptor Blocking, human				MILTENYI BIOTEC	1:50
<b>Panel for PBMC sorting</b>					
anti-human CD45RO	FITC	UCHL1	Mouse / IgG2a, κ	EBIOSCIENCE	1:40
anti-human CCR7	APC	3D12	Rat IgG2a, κ	EBIOSCIENCE	1:40
anti-human CD19	APC*Fire	HIB19	Mouse IgG1, κ	BIOLEGEND	1:40
anti-human CD38	PE*Cy7	MIB19	Mouse IgG1, κ	EBIOSCIENCE	1:40
anti-human CD27	BV605	O323	Mouse IgG1, κ	BIOLEGEND	1:40
anti-human CD8	PerCP	SK1	Mouse IgG1, κ	BIOLEGEND	1:50
anti-human CD3	AF700	OKT 3	Mouse / IgG2a, κ	EBIOSCIENCE	1:50
anti-human CD4	Pacific Blue	S3.5	mIgG2a *PacBlue	THERMO FISHER SCIENTIFIC	0.1 $\frac{mg}{mL}$
Fc-Receptor Blocking, human				MILTENYI BIOTEC	1:50
<b>Immunofluorescence antibodies used in antigen search experiments</b>					
anti-mouse CD3ε	APC	145-2C11	Hamster/IgG	EBIOSCIENCE	1:50
anti-mouse CD3ε	∅ (for CD3 activation assay)	145-2C11	Hamster/IgG	EBIOSCIENCE	1:50
anti-human CD8β	PE	2ST8.5H7	Mouse BALB/c IgG2a, κ	BD BIOSCIENCES	1:50

### 3.1.5 Antibiotics

**Table 4: Antibiotics used for plasmid selection in cell culture**

Antibiotic (Eukaryotes)	Stock conc.	Cell culture conc.	Selected plasmid	Company
Blasticidin S	10 000 $\frac{\mu\text{g}}{\text{mL}}$	3 $\frac{\mu\text{g}}{\text{mL}}$	NFAT-sGFP (in hybridoma cells)	INVITROGEN
Geneticin (G418) Sulfate	100 $\frac{\text{mg}}{\text{mL}}$	1.5 $\frac{\text{mg}}{\text{mL}}$	1) TCR- $\beta$ chain (in hybridoma cells) 2) HLA-BFP (in COS cells)	SANTA CRUZ BIOTECHNOLOGY
Hygromycin B	50 $\frac{\text{mg}}{\text{mL}}$	0.3 $\frac{\text{mg}}{\text{mL}}$	TCR- $\alpha$ chain (in hybridoma cells)	THERMOFISHERSCIENTIFIC
Puromycin	1000 $\frac{\mu\text{g}}{\text{mL}}$	1 $\frac{\mu\text{g}}{\text{mL}}$	<i>h</i> CD8 $\beta$ (in hybridoma cells)	BIOMOL
Antibiotic (Prokaryotes)	Stock conc.	Cell culture conc.		Company
Ampicillin	100 $\frac{\text{mg}}{\text{mL}}$	100 $\frac{\mu\text{g}}{\text{mL}}$		SIGMA

### 3.1.6 Oligonucleotides

All primers listed in Table 5 were purchased from the company METABION (stock solution: 100  $\mu\text{M}$ ; HPLC purified) and diluted with nuclease-free water (AMBION). Primers were stored at -20 °C, besides of TSO\_Smart-seq2 that contains RNA nucleotides and was stored at -80 °C. The terms OUT and IN of the primers refer to their position in a nested PCR. Ambiguous wobble nucleotides were introduced according to the standard IUPAC nomenclature<sup>[172]</sup>: R (A or G), Y (C or T), K (G or T), S (G or C), W (A or T), B (G or T or C), N (any base) and V (A, C or G). The modified nucleotides riboguanosine (rG) and locked nucleic acid (LNA) guanosine (+G) were included in the 3'-end of template switching oligo (TSO) sequences to increase the affinity towards cDNA and thereby facilitating the template switching mechanism of the enzyme in the reverse transcription reaction (RT reaction) (3.2.3.1).<sup>[173-174]</sup>

**Table 5: Oligonucleotides**

Oligonucleotide	Nucleotide-Sequence (5' → 3' direction)
<b>Smart-seq2-specific Primer</b>	
TSO_Smart-seq2	Biotin_AAG CAG TGG TAT CAA CGC AGA GTG AAT rGrG+G
Oligo-dT <sub>30</sub> VN	AAG CAG TGG TAT CAA CGC AGA GTA CT <sub>30</sub> V N
TSO-PCR_67°C (fwd.)	AAG CAG TGG TAT CAA CGC AGA GT
<b>Illumina Index Primer</b>	
S5xx	5' → 3'

S501	AATGATACGGCGACCACCGAGATCTACAC TAGATCGC TC
S502	AATGATACGGCGACCACCGAGATCTACAC CTCTCTAT TC
S503	AATGATACGGCGACCACCGAGATCTACAC TATCCTCT TC
S504	AATGATACGGCGACCACCGAGATCTACAC AGAGTAGA TC
S505	AATGATACGGCGACCACCGAGATCTACAC GTAAGGAG TC
S506	AATGATACGGCGACCACCGAGATCTACAC ACTGCATA TC
S507	AATGATACGGCGACCACCGAGATCTACAC AAGGAGTA TC
S508	AATGATACGGCGACCACCGAGATCTACAC CTAAGCCT TC
<b>N7xx</b>	<b>5' → 3'</b>
N701	CAAGCAGAAGACGGCATAACGAGAT TCGCCTTA GTCTCGT
N702	CAAGCAGAAGACGGCATAACGAGAT CTAGTACG GTCTCGT
N703	CAAGCAGAAGACGGCATAACGAGAT TTCTGCCT GTCTCGT
N704	CAAGCAGAAGACGGCATAACGAGAT GCTCAGGA GTCTCGT
N705	CAAGCAGAAGACGGCATAACGAGAT AGGAGTCC GTCTCGT
N706	CAAGCAGAAGACGGCATAACGAGAT CATGCCTA GTCTCGT
N707	CAAGCAGAAGACGGCATAACGAGAT GTAGAGAG GTCTCGT
N708	CAAGCAGAAGACGGCATAACGAGAT CCTCTCTG GTCTCGT
N709	CAAGCAGAAGACGGCATAACGAGAT AGCGTAGC GTCTCGT
N710	CAAGCAGAAGACGGCATAACGAGAT CAGCCTCG GTCTCGT
N711	CAAGCAGAAGACGGCATAACGAGAT TGCCTCTT GTCTCGT
N712	CAAGCAGAAGACGGCATAACGAGAT TCCTCTAC GTCTCGT
<b>Primers for target-specific TCR and BCR PCRs</b>	
C $\alpha$ IN (rev.)	AGT CTC TCA GCT GGT ACA CG
C $\alpha$ OUT (rev.)	GCA GAC AGA CTT GTC ACT GG
C $\beta$ IN (rev.)	TCT GAT GGC TCA AAC ACA GC
C $\beta$ OUT_mid4 (rev.)	TGG GTG TGG GAG ATC TCT G
C $\mu$ IN (rev.)	AAG GGT TGG GGC GGA TGC
C $\mu$ OUT (rev.)	GGG AAT TCT CAC AGG AGA CG
HG IN (rev.)	AGA GGT GCT CTT GGA GGA G
HG OUT (rev.)	AGT TCC ACG ACA CCG TCA C
Kappa IN (rev.)	GAA GAT GAA GAC AGA TGG TGC
Kappa OUT (rev.)	CAC ACA ACA GAG GCA GTT CC
Lambda IN (rev.)	GCT TGG AGC TCC TCA GAG G
Lambda OUT (rev.)	CAC CAG TGT GGC CTT GTT GGC
Nextera_TSO-PCR_60°C (fwd.)	TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG NNA ACG TGA TAA GCA GTG GTA TCA ACG CAG
TSO-PCR_60°C (fwd.)	AAG CAG TGG TAT CAA CGC AG
<b>Finding TCR TRAV17*01/TRBV4-1*01 in PBMCs</b>	
Rev-spec3	TGNGTYTCCCANGCNCNGG
Trappist-161-TRB for	CTTGGGCTGCAGGTCGACATGGGCTGCAGGCTGCTC

TRBV4_1*01 for	GACACTGAAGTTACCCAGAC
$\alpha$ -chain RT	GCCACAGCACTGTTGC
$\beta$ -chain RT	GWAGAAGCCTGTGGCC
V $\alpha$ 3 for OUT	GAAAATGCCACCATGAACTGC
TRAJ 23 rev	CCGAAGATAAGCTTTCCTCC
V $\alpha$ 3 for IN	TTCAGGTAGAGGCCTTGTC
TRA spec Rev	GCTTTCCTCCCTGGTTRTAR
<b>Primers for cloning of antigen search experiments</b>	
162_PB-01-H-for	GTTTCCGCGGTGGGTCCTGTCCCAGGTGCAGGTGGTGGAG
162_PB-01-K-for	AACGGGCGCGCGATGTGATATTGTGATGACCCAGAGT
301_PB-01-H-for	GTTTCCGCGGTGGGTCCTGTCCCAGGTGCAGCTGGTGCAG
4626_PB-01-K-for	AACGGGCGCGCGATGTGATGTTGTGATGACTCAGTC
4X26_PB-H-01-for	GTTTCCGCGGTGGGTCCTGTCCGAGGTGCAGCTGGTGGAG
CHeavy SalI rev *	GTGCCCCAGAGGTcgaCTTGG
C $\alpha$ rev IN	AGTCTCTCAGCTGGTACACG
C $\beta$ rev IN	TCTGATGGCTCAAACACAGC
IGKV1_PB-01-K-for	AACGGGCGCGCGATGTGACATCCAGATGACCCAGTC
Kappa-V,C KasI rev *	AGACAGATGGcGCcGCCACAG
TRAPPIST-161_aII-for (alpha II)	CTTGGGCTGCAGGTCGACATGCTGTTCTCCAGCCTGC
TRAPPIST-161_TRA-for	CTTGGGCTGCAGGTCGACATGGAACTCTCTGGGAGT
TRAPPIST-161_TRB-for	CTTGGGCTGCAGGTCGACATGGGCTGCAGGCTGCTC
<b>Cloning HLA~BFP constructs</b>	
A101_lead fwd	CTCGTCGAC / ATGGCCGTCATGGCGCCC
A23_NotI rev	TGCGGCCGCCACTTTACAAGCTGTGAGAGAC
B0801_lead fwd	CTCGTCGAC / ATGCTGGTCATGGCGCCC
B27:02_Not rev	AAGCGGCCGCAGCTGTGAGAGACACATCAG
HLA_C*07:02:01 NotI rev	CGCG / GATCCTCATGC / GGCCGCGATGAGAGACTCATCAGAGC
HLA_C_for_SalI_for **	ACA CGT CGA CAT GCG GGT CAT GGC GCC
mtagBFP2_Not_fwd	TTTGC / GGCCGCAAGCGAGCTGATTAAGGAGAAC
mtagBFP2_stop_BamHI rev	CGCG / GATCCTTAATTAAGCTTGTGCCCCAG
pHSE 3' rev	AAAACATCAAGGGTCCCATA
<b>Primers for PECPL recovery in T cell activation assays</b>	
pcDNA-2 <sup>nd</sup> rev10	CTAGACTCGAGCGGCCGC
pcDNA-2 <sup>nd</sup> -for-TOPO	CACCTCCGGCGCGCCACCATG
pcDNA-for 1	CACTGCTTACTGGCTTATCG
pcDNA-rev1	ACTAGAAGGCACAGTCGAGG
pcDNA-rev3	TGGTGATGGTGATGATGACC

\*BIRGIT OBERMAIR

\*\*KATHERINA SIEWERT

### 3.1.7 Enzymes & specific Kits

**Table 6: Enzymes and specific Kits**

Enzyme /Kit name	Company
2× KAPA HiFi HotStart ReadyMix	KAPA BIOSYSTEMS
2× Phusion High-Fidelity PCR Master Mix with HF Buffer	Thermo Fisher Scientific
Agencourt AMPure XP	BECKMANN COULTER
<i>AsCI</i> restriction enzyme	NEW ENGLAND BIOLABS
<i>BamHI</i> restriction enzyme	NEW ENGLAND BIOLABS
EasySep™ human CD4 positive selection kit II	STEMCELL™ TECHNOLOGIES
FuGene HD Transfection mix	PROMEGA
iProof High Fidelity Master Mix	BIO-RAD
MinElute Gel extraction Kit	QIAGEN
<i>NdeI</i> restriction enzyme	NEW ENGLAND BIOLABS
Nextera DNA library preparation kit	ILLUMINA
<i>NotI</i> restriction enzyme	NEW ENGLAND BIOLABS
pcDNA 3.1 directional TOPO expression kit	INVITROGEN
PureLink™ RNA Micro Kit)	INVITROGEN
<i>PvuII</i> restriction enzyme	NEW ENGLAND BIOLABS
QIA HiSpeed® Plasmid Maxi Kit	QIAGEN
QIA prep Spin Miniprep Kit	QIAGEN
QIAquick® Nucleotide Removal Kit	QIAGEN
Rapid DNA Dephos and Ligation Kit	ROCHE
RNaseOUT Recombinant Ribonuclease Inhibitor (40 $\frac{Units}{\mu L}$ )	THERMO FISHER SCIENTIFIC
<i>SalI</i> restriction enzyme	NEW ENGLAND BIOLABS
SMARTScribe Reverse Transcriptase	TAKARA
SuperScript® II Reverse Transcriptase (200 $\frac{Units}{\mu L}$ )	THERMO FISHER SCIENTIFIC
T4 DNA ligase (1 U/ $\mu$ L)	INVITROGEN
Taq DNA polymerase (5 U/ $\mu$ L)	ROCHE
<i>XmnI</i> restriction enzyme	NEW ENGLAND BIOLABS

### 3.1.8 Plasmids

**Table 7: Plasmids**

Plasmid	Size	Specificity
BFP_pHSE3'- HLA XY	9.9 kb	HLA construct as above with additional mtag-BFPs tag for stable transfectants amp <sup>R</sup> , neo <sup>R</sup>

pcDNA-NFAT-sGFP	7.4 kb	NFAT-GFP construct as a T cell activation reporter that was used for stable transfection of 58 $\alpha\beta$ T hybridoma cells; amp <sup>R</sup> , bls <sup>R</sup>
pcDNA-sGFP	5.1 kb	COS-7 transfection control
pHSE3'-HLA XY	9.2 kb	HLA construct for expression of a patient-specific HLA XY in COS-7 APCs amp <sup>R</sup> , neo <sup>R</sup>
pRSV5hygro		TCR $\alpha$ chain cloning
pRSV5-neo		TCR $\beta$ chain cloning
<b>Plasmid-encoded combinatorial peptide libraries</b>		
8L (3.9 x 10 <sup>8</sup> clones)	5.1 kb	CAGGGAAGGCGCGCCACC <b>ATG</b> NNK NNK NNK NNK NNK NNK NNK <b>CTG</b> TGAGCGGCCGCTAAACTAT
Cw06	5.1 kb	CAGGGAAGGCGCGCCACC <b>ATG</b> TWY AGR NNK NNK NNK NNK NNK NNK MTH <b>TGAGCGGCCGCTAA</b> ACT AT
N24 (2x10 <sup>8</sup> clones)	5.1 kb	CAGGGAAGGCGCGCCACC <b>ATG</b> NNK NNK NNK NNK NNK NNK NNK NNK <b>TGAGCGGCCGCTAAACTAT</b>
N27 (10 <sup>8</sup> clones)	5.1 kb	CAGGGAAGGCGCGCCACC <b>ATG</b> NNK NNK NNK NNK NNK NNK NNK NNK <b>TGAGCGGCCGCTAAACTAT</b>
N27L (1.8 x 10 <sup>8</sup> clones)	5.1 kb	CAGGGAAGGCGCGCCACC <b>ATG</b> NNK NNK NNK NNK NNK NNK NNK NNK <b>CTN</b> TGAGCGGCCGCTAA ACT AT
N30L	5.1 kb	CAGGGAAGGCGCGCCACC <b>ATG</b> NNK NNK NNK NNK NNK NNK NNK NNK <b>CTN</b> TGAGCGGCCGCTAA ACT AT

### 3.1.9 Cell lines

**Table 8: Eukaryotic and prokaryotic cell lines**

<b>Eukaryotic cell lines</b>	
COS-7	Immortalized fibroblast-like cell line from African green monkey kidney tissue; carrying SV40 origin; adherent cells
A24:02-BFP COS	COS-7 cells (as above) stably transfected with HLA-A24:02-BFP construct
B07:02-BFP COS	COS-7 cells (as above) stably transfected with HLA-B07:02-BFP construct
C07:02-BFP COS	COS-7 cells (as above) stably transfected with HLA-C07:02-BFP construct
58 $\alpha\beta$ T hybridoma cells	mouse T hybridoma cell line with TCR deficiency for transfection of disease-specific human $\alpha\beta$ TCR
CD8 $\alpha\beta$ -GPE cells	GP+E86 packaging cell line transfected with <i>h</i> CD8 $\alpha\beta$ plasmid (SEITZ et al.) <sup>[175]</sup> that enables retroviral CD8 transfer to 58 cells
58V <sub><math>\alpha</math>17</sub> V <sub><math>\beta</math>4.1</sub> cells	58 $\alpha\beta$ T cells stably transfected with plasmids for TCR $\alpha\beta$ chains V <sub><math>\alpha</math>17</sub> and V <sub><math>\beta</math>4</sub> , NFAT-sGFP and <i>h</i> CD8 $\beta$
EBV immortalized B cells #162	EBV transformed B cells from PBMCs of patient #162 (Klaus Dornmair)

JM22 58V <sub>α10.1</sub> V <sub>β17</sub> cells	58αβ <sup>+</sup> T cells stably transfected with plasmids for JM22 TCRαβ chains V <sub>α10.1</sub> and V <sub>β17</sub> , NFAT-sGFP and <i>h</i> CD8β (KATHERINA SIEWERT)			
COS-7~A02	Cos-7 cells (as above) stably transfected with HLA-A02:01 (KATHERINA SIEWERT)			
<b>Prokaryotic <i>E. coli</i> strains</b>				
ElectroMAX™ DH10B™	T1 Phage resistant cells that can be transformed by electroporation			INVITROGEN
MAX Efficiency™ DH5α Competent Cells	Chemically competent cells			INVITROGEN
One Shot™ TOP10	Chemically competent <i>E. coli</i>			INVITROGEN
<b>EBV B cell lines</b>	<b>HLA-A</b>	<b>HLA-B</b>	<b>HLA-Cw</b>	<b>EBV</b>
#36	24:03, 29:02	18:01, 44:03	12:03, 16:01	
FE	01:01	08:01	07:01	
IBi	11:01, 24:02	18:01, 51:01	04:01, 07:01	Z-
KK	01:01, 23:01	50:01, 57:02	06:02, 18	Z-
TG	23:01, 24:02	07:02, 07:05	07:02, 15:05	Z-
#162	24:02	07:02	07:02	

### 3.1.10 Biological samples

Human PBMCs and CSF samples were obtained from patients and either processed freshly or flash frozen in liquid nitrogen and provided by PD DR. TANIA KÜMPFEL from the Institute of Clinical Neuroimmunology (Klinikum Großhadern, Munich, Germany).

Informed consent was obtained from all patients and voluntary donors.

Here, the unique cohort (Table 9, upper part) comprised four monozygotic twins with discordance for their MS disease status, four co-twins with subclinical neuroinflammation (SCNI) and four controls. Twin pairs were genetically identical individuals of whom only one developed clinical manifested MS, whereas the ‘healthy’ co-twin was asymptomatic. Co-twins were clinically healthy, but CSF of the related twin sibling with MS was not available. In-depth clinical analysis of those healthy SCNI co-twins revealed small MRI lesions and partly OCBs, which might be an indicator of a very early pre-clinical stage in MS pathogenesis. Controls were separated into a group of non-inflammatory controls (NIC), comprising two patients with intracranial idiopathic hypertension, and an inflammatory, but non-MS-related control group with two encephalitis [anti-leucine-rich-glioma-inactivated-1 encephalitis (anti-LGI-1-E) and anti-N-Methyl-D-Aspartate-receptor encephalitis (NMDA-R-E)] individuals.

**Table 9: Clinical characteristics of the analysed CSF from human patients (first part modified from BELTRÁN *et al.*)<sup>[176]</sup>**

Code internal	Code (Beltrán <i>et al.</i> )	Disease type	Age	Sex	disease duration (yrs)	Disease modifying treatment	MRI findings	CSF results	
#155	T1	AR-MS	RRMS	40	f	4.0	interferon-beta	> 5 WML*#	OCB+, 2/μl
#156		AR-H	SCNI	40	f	4.0	none	> 5 WML*	OCB+, 5/μl
#161	T2	AU-MS	RRMS	21	f	1.5	natalizumab	> 5 WML*	OCB+, 3/μl
#162		AU-H	SCNI	21	f	1.5	none	2-5 WML	OCB+, 13/μl
#163	T3	AV-MS	SPMS	50	f	17.0	intrathecal steroids	> 5 WML*	OCB-, 1/μl
#164		AV-H	Healthy	50	f	17.0	none	no WML	OCB-, 3/μl
#301	T4	BJ-MS	RRMS	45	f	1.0	teriflunomide	> 5 WML*	OCB+, 3/μl
#302		BJ-H	Healthy	45	f	1.0	none	no WML	OCB-, 1/μl
#114	co-twin BF-H	SCNI	28	f	1.5	none	2-5 WML*	OCB+, 3/μl	
#116	co-twin V-H	SCNI	38	m	7.0	none	1 WML	OCB+, 9/μl	
#186	co-twin D-H	SCNI	32	f	13.0	none	2-5 WML	OCB-, 2/μl	
#78	co-twin W-H	SCNI	69	f	28.0	none	1 WML	OCB-, 4/μl	
Anti-LGI1 Enc.	Enc-1	LGI-1-E	62	m	0.3	steroids, PLEX	no WML	OCB-, 3/μl	
Anti-NMDA Enc.	Enc-2	NMDARE	20	f	0.4	steroids, PLEX, RX, IVIG	2-5 WML	OCB-, 1/μl	
SCP-14	NIC-14	IIH	22	f	1.0	acetazolamide	no WML	OCB-, 3/μl	
SCP-15	NIC-15	IIH	30	f	0.5	acetazolamide	no WML	OCB-, 3/μl	
Code internal	Disease type		Age	Sex	Disease modifying treatment		CSF results		
#4526	GAD antibody associated encephalitis		47	m	-		OCB+, 14/μl		
#4626	CIS/ RRMS (since 2014)		26	f	-		OCB+, 9/μl		
#6014	RRMS/ early MS		32	m	-		OCB+, 8/μl		
#6137	RRMS+ CAPS/ FMF		46	f	-		OCB+, 11/μl		

f=female, m=male, MRI=magnet resonance imaging, CIS=clinical isolated syndrome, RRMS=relapsing remitting MS, SPMS=secondary progressive MS, PLEX=plasma exchange, RX=rituximab, IVIG=intravenous

immunoglobulines, WML=white matter lesions, \*=dissemination in space with a least two lesions in two different regions according to the revised McDonald criteria (2017), #=dissemination in time with detection of new lesions in a follow-up MRI scan, OCB=oligoclonal bands, LIG-1-E=anti-leucine-rich-glioma-inactivated-1 encephalitis, NMDA-R-E=anti-N-Methyl-D-Aspartate-receptor encephalitis, NIC=non-inflammatory control, SCNI=subclinical neuroinflammation, IHH= intracranial idiopathic hypertension, SCP= single cell project, CIS= Clinical isolated syndrome, CAPS=cryopyrin-associated autoinflammatory syndromes, FMF=familial mediterranean fever, GAD=glutamic acid decarboxylase

**Table 10: HLA class I alleles of analysed PBMCS from HLA-matching RIS patients**

Internal Code	HLA-A alleles		HLA-B alleles		HLA-C alleles	
#74	A03	A68	B41	B44	C05	C17
<b>#78</b>	A02	A03	<b>B07</b>	B37	<b>C07*</b>	C06
<b>#102</b>	A02	A68	B44	B47	<b>C07</b>	C06
<b>#112</b>	A24	A26	B27	B49	<b>C07</b>	C01
<b>#114</b>	A03		<b>B07</b>	B51	C07	
#120	A11	A31	B35	B35	C04	
<b>#130</b>	A03	A24	B08	B39	<b>C07</b>	
<b>#140</b>	A01	A03	<b>B07</b>	B35	<b>C07</b>	C04
<b>#147</b>	A01	A02	B08	B51	<b>C07</b>	C15
<b>#148</b>	A01	A02	B08	B51	<b>C07</b>	C15
<b>#157</b>	A02	A30	B08	B42	<b>C07</b>	C17
<b>#161</b>	<b>A24</b>		<b>B07</b>		<b>C07</b>	
<b>#162</b>	<b>A24</b>		<b>B07</b>		<b>C07</b>	
<b>#301</b>	A33	A02	B44	<b>B07</b>	C02	<b>C07</b>
<b>#302</b>	A33	A02	B44	<b>B07</b>	C02	<b>C07</b>

\*HLA-matching patients

### 3.2 Identification and characterization of expanded lymphocytes in MS patients

A workflow of the strategy to identify heterodimeric TCR and Ig receptors together with their transcriptomes by scRNA-seq was outlined below (Figure 11).

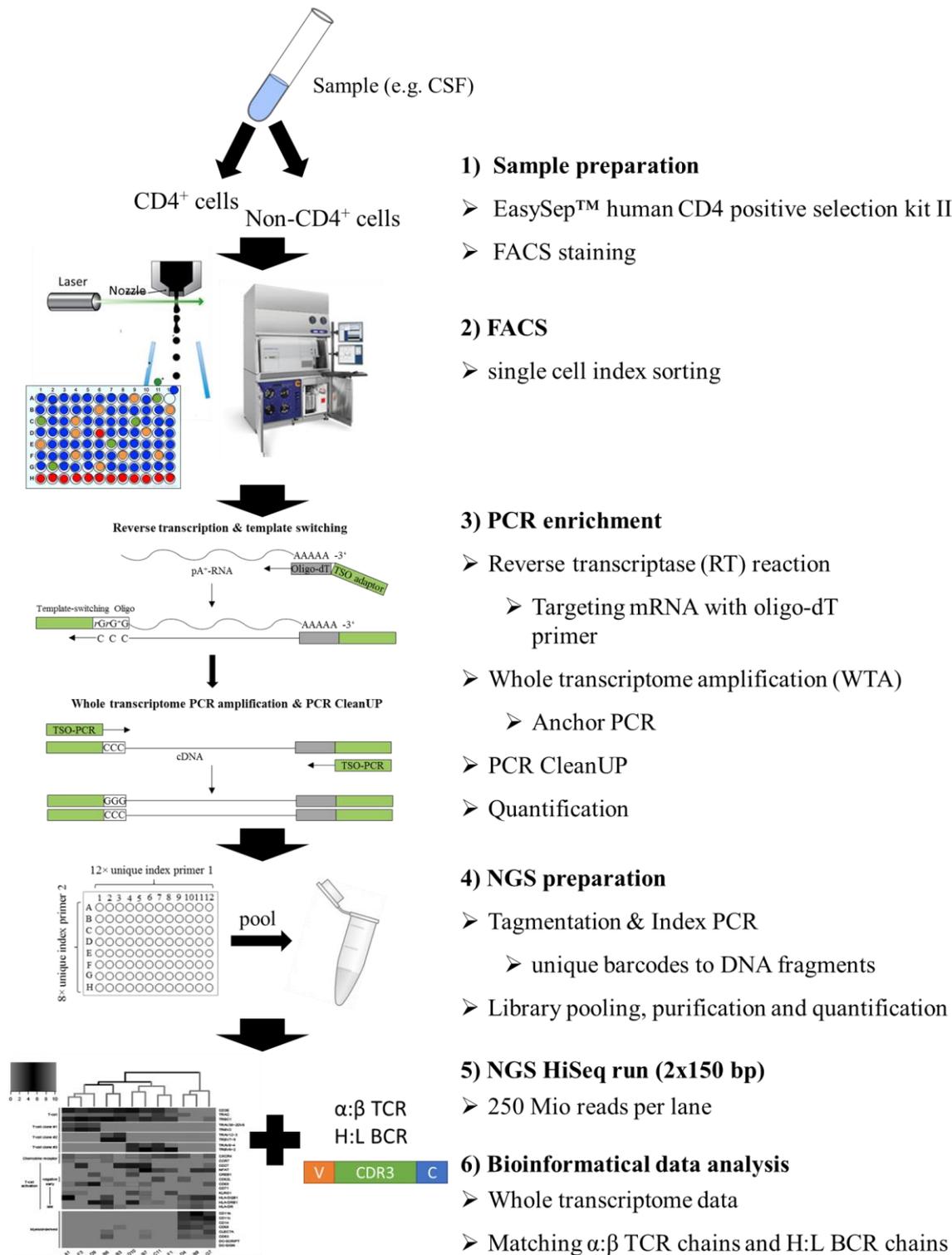
CD4<sup>+</sup> cells that constitute the overwhelming majority of cells in CSF were analysed separately from non-CD4<sup>+</sup> cells to increase the chance of detecting rare cell subpopulations such as plasma cells in the remaining population. Thus, human samples were prepared by using the EasySep™ human CD4 positive selection kit II that isolates CD4<sup>+</sup> cells with antibody labelled magnetic beads (3.2.1). The remaining non-CD4<sup>+</sup> population and the CD4<sup>+</sup> samples were then stained with an individual panel of cell-type specific antibodies against cell-surface markers (Table 3). This enabled the identification of the FACS sorted SC on the protein level and added data to the later transcriptome analysis.

In the next step, both stained samples were processed on a BD Aria III cell sorter (3.2.2). Usually sorting a particular cell population of interest resulted in the loss of all other precious cells that bypass the nozzle of the FACS at the same time. In contrast, sorting every SC of the sample so far meant losing crucial information about bound antibodies that identified and further characterized cell subtypes. These issues were bypassed with using the index sorting feature of the FACS and reduced the loss of cells from the precious human CSF samples to a minimum. Single lymphocytes were selected according to their cell shape and size by using forward and side scatter properties of the FACS and SC isolated in cell lysis buffer into 96-well plates. Every sorted SC that passed the nozzle of the FACS was hit by the laser beam and by this the information of bound labelling antibodies was retained. The index sorting data was mapped to each single well and gave an overview of the prevalent subpopulations in the samples. It was also used to assess the quality of the partly longer stored frozen samples and evaluate potential treatment effects of the MS patients.

For whole transcriptome (WT) data, RNA of lysed SC was transcribed to DNA and enriched in different PCR steps as described in the Smart-seq2 protocol of PICELLI *et al.*<sup>[147]</sup>(3.2.3) In short, pA<sup>+</sup> RNA was reversed transcribed into cDNA with oligo-dT primer that contain a 5'-overhang template-switch oligo (TSO) adaptor sequence and a TSO-Oligo that is added to the cDNA 3' end (3.2.3.1). By this TSO anchors, transcribed cDNA molecules were amplified in an unbiased manner with WT-PCR (3.2.3.2). Afterwards, PCR reaction ingredients, excess enzymes and primers were removed from the enriched WT fragments by size selection with XP Agencourt magnetic beads. Following the PCR CleanUP, DNA samples were quantified and the quality of PCR enrichment assessed with a Qubit fluorometer and an Agilent Bioanalyzer (3.2.3.3).

Successful amplified samples showed concentrations between  $0.5 \frac{ng}{\mu L}$  and  $1 \frac{ng}{\mu L}$  and a main peak between 1.5 kb and 2 kb.

Next, the transcriptomes of the SC were fragmented into smaller pieces and uniquely barcoded in the NGS preparation steps (3.2.4). For fragmentation, the protocol of BAYM *et al.* was used and optimized, so that the used reaction volume of the original Smart-seq2 method was decreased to 2  $\mu L$  (3.2.4.1).<sup>[177]</sup> This reduced the cost of the most expensive step – the Nextera DNA library preparation kit (Illumina) reagents- in this analysis to a minimum. cDNA libraries were barcoded in an index PCR with a combination of different unique index primers that identified transcripts of each single well in the final data analysis (3.2.4.2). Libraries were pooled and purified in several steps and the quality of the final pool controlled (3.2.4.3). cDNA libraries with a concentration between  $0.8 \frac{ng}{\mu L}$  and  $10 \frac{ng}{\mu L}$  and an equal size distribution between 200 kb and 1 kb were sent for paired-end 2x 150 bp NGS HiSeq run (3.2.5). Usually 2x 100 bp sequencing was commonly performed for sequencing of WT from SCs and used in the Smart-seq2 protocol. But only 2x 150 bp sequencing comprised long enough reads to reconstitute full CDR3 sequences of matching  $\alpha:\beta$  TCR and H:L BCR chains in addition to the transcriptome profiles in the bioinformatics analysis.



**Figure 11: Overview on the workflow of the modified “Smart-seq2” single cell transcriptome method.** (1) First, human samples such as CSF or PBMCs were drawn from patients and prepared for the analysis as follows. CD4<sup>+</sup> cells were isolated from the remaining non-CD4<sup>+</sup> population with the EasySep™ human CD4 positive selection kit II. Each sample was prepared for fluorescence-activated cells sorting (FACS) by staining with individual panels of cell specific antibodies against extracellular surface markers. (2) Single cells were isolated into 96-well plates with the index sorting feature of FACS BD Aria III. (3) The transcriptome of each cell was targeted with oligo-dT primers that comprised a 5’-overhang TSO adaptor sequence. pA<sup>+</sup>-RNA was reverse transcribed and the TSO adaptor sequence added to the 3’ end of the cDNA. Subsequently the whole transcriptome of the cell was amplified, purified with Agencourt XP beads and quantified in a PCR enrichment step. (4) NGS libraries were prepared by cutting the enriched transcriptomes into small fragments during tagmentation reaction and uniquely labelled with index primers. Barcodes enabled the identification of transcriptome data of each single well. (5) Finally libraries of each well from the 96-well plate were pooled, purified and quantified. The final sample was sent for a paired-end 2x 150 bp NGS HiSeq run. (6) Bioinformatic data analysis yielded the transcriptome of each SC along with their antigen-specific receptors.

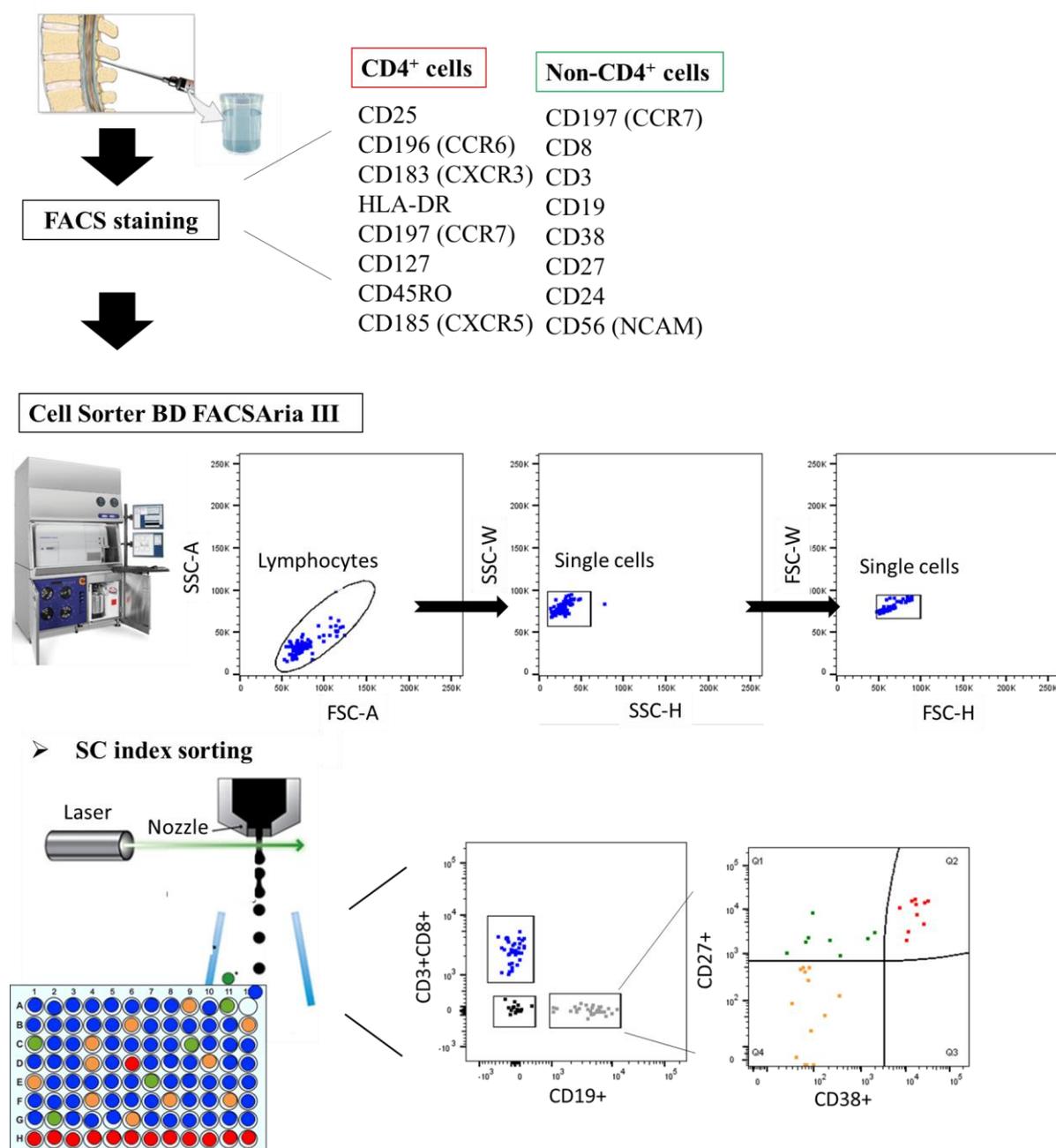
### 3.2.1 Sample preparation

Whole blood (~10 mL) was drawn from patients and controls into EDTA and PBMCs prepared by Ficoll density gradient centrifugation. The blood was diluted up to a final volume of 35 mL with PBS (pH: 7.4) and carefully layered on top of 15 mL Pancoll human solution (Ficoll 400; PAN BIOTECH). Samples were centrifuged (400 g, 30 min, RT, no brake), the PBMC layer removed and washed. Therefore, PBS was added to a final volume of 50 mL, samples were centrifuged (300 g, 10 min, RT, with brake) and the supernatant discarded. Platelets were removed by washing the pellet twice with 50 mL PBS, centrifugation (200 g, 10 min, RT, with brake) and discarding the supernatant. The remaining cell pellet with the isolated PBMCs was resuspended in freezing buffer and aliquoted in 2× 1 mL aliquots in cryo tubes with each containing ~ 10× 10<sup>6</sup> cells. Samples were frozen at -80 °C in isopropyl alcohol containing freezing containers and for longer storage transferred to a liquid nitrogen tank.

CSF was drawn from patients and CD4<sup>+</sup> cells depleted by using the EasySep™ human CD4 positive selection kit II according to the manufacturer's instructions. In short, 3 mL CSF was drawn freshly from patients, centrifuged (300 g, 10 min, 20 °C) and the supernatant discarded. The cell pellet was resuspended in 250 µL EasySep Buffer (2% FBS, 1 mM EDTA in 1x PBS) and transferred to a 2 mL Protein LoBind tube. 12.5 µL human CD4 selection cocktail was added to the cell suspension and incubated at RT for 3 min. RapidSpheres™ beads were vortexed for 30 sec, 12.5 µL added to the cells, gently mixed and incubated for 3 min at RT. The final volume was increased to 1 mL by adding 700 µL EasySep Buffer and the tube incubated within the magnetic stand for 3 min at RT. The supernatant containing non-CD4<sup>+</sup> cells were transferred to a new 2 mL Protein LoBind tube. To collect remaining non-CD4<sup>+</sup> cells, RapidSpheres™ beads were washed with 1 mL EasySep Buffer and 3 min incubation on the magnetic stand. The supernatant was transferred to the Protein LoBind tube containing the non-CD4<sup>+</sup> cells. 2 mL FACS Buffer was added to the remaining beads to yield the CD4<sup>+</sup> cell sample. Subsequently both samples were centrifuged (310 g, 7 min, 4 °C) and stained for flow cytometry sorting. Each sample was stained with an individual panel of fluorochrome-conjugated anti-human antibodies against lymphocyte surface markers (Table 3). Staining antibodies were diluted in 100 µL FACS buffer and incubated for 30 min on ice in the dark. Afterwards, 300 µL FACS buffer was added, the samples centrifuged (310 g, 7 min, 4 °C) and the supernatant discarded. The cell pellet was resuspended in 300 µL FACS buffer and kept on ice until further processing with the FACS sort.

### 3.2.2 Fluorescence-activated cell sorting (FACS)

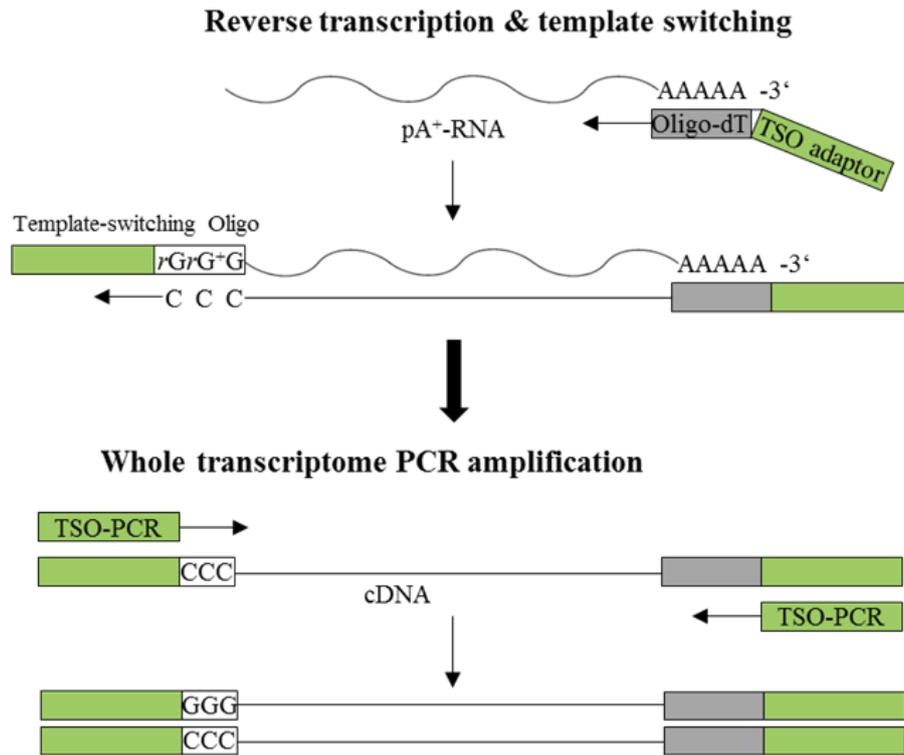
Cells were isolated from FACS stained CD4<sup>+</sup> and Non-CD4<sup>+</sup> samples by using the index sorting feature of the BD FACSAria™ flow cytometer. Single lymphocytes were selected according to their forward scatter and side scatter properties (Figure 12). Index sorting enables to sort all SC by retaining information of labeled antibodies. Each SC was placed into one well of a 96-well plate with 5 µL SC mix (0.4 Unit RNase OUT; 0.038% Triton-X-100; 1 µM oligo dT primer; 1 mM dNTP mix). After sorting, the plates were vortexed, centrifuged (700 g, 10 sec, RT) and immediately placed on ice to preserve a high-quality RNA. Linearization and binding of the oligo dT primer was facilitated by incubating the plates on a thermal cycler (3 min, 72 °C) and immediately keeping the plates on ice. Plates were centrifuged (700 g, 10 sec, RT) and immediately put back on ice until reverse transcription (RT).



**Figure 12: Flow-chart of the involved steps in fluorescence-activated cell sorting (FACS).** Human samples such as CSF or PBMCs were drawn from patients and prepared as described in section 3.2.1. The CD4<sup>+</sup> subpopulation and the remaining non-CD4<sup>+</sup> cells were stained with a specific antibody panel of fluorescent dyes against common cell surface markers as depicted above. Both subpopulations were separately processed on a BD FACSaria™ Fusion cell sorter. Single lymphocytes were selected according to their size and cell shape by using forward scatter (FSC) and side scatter (SSC) properties in the flow cytometer. Finally, all single lymphocytes were sorted into wells of a 96-well plates by using single cell (SC) index sorting. All SC that passed the nozzle were targeted by a laser beam and this made it possible to subsequently match sorted cells in the FACS plot with their location on the 96-well plate.

### 3.2.3 PCR enrichment

PCR enrichment steps were performed according to the protocol of PICELLI *et al.* and the workflow depicted in Figure 13.<sup>[147]</sup> All steps were performed in 96-well plates on a thermal cycler and the plates kept on ice in between the steps. Used primers were listed in Table 5.



**Figure 13: Schematic presentation of the transcriptome enrichment strategy.** pA<sup>+</sup>- RNA of the sorted and lysed single cells was targeted by an oligo-dT primer with 5'-adaptor template-switching oligo (TSO) sequence. The SuperScriptII reverse transcriptase is adding cytosine (C) nucleotides to the 3' end of reverse transcribed cDNA, switch template and add the TSO adaptor sequence to the 3' end. Full length cDNA was enriched by PCR amplification with the TSO-PCR primer.

### 3.2.3.1 RT reaction

The transcriptome of pA<sup>+</sup> RNA from sorted single lymphocytes was tagged and transcribed into cDNA with the anchored Oligo-dT<sub>30</sub>VN primer that contained a 5'-TSO adaptor sequence overhang. SuperScriptII reverse transcriptase had a terminal transferase activity and added a few non-templated cytosine nucleotides to the 5'-end of the RNA template. Additionally it is able to switch template and added the specific TSO\_Smart-seq2-oligo to the 3'-end of the synthesized full length cDNA. Biotin was included into the 5'-end of the TSO\_Smart-seq2 primer sequence to prevent multiple elongations of the sequence.

5  $\mu$ L RT mix (1 $\times$  SuperScriptII first strand buffer (CLONTECH), 5 mM DTT (CLONTECH), 1 M Betaine, 6 mM MgCl<sub>2</sub>, 10 Unit RNase OUT, 1  $\mu$ M TSO\_Smart-seq2, 100 Unit or 50 Unit SuperScript II reverse transcriptase) was added to sorted SC of section 3.2.2. Therefore, the final volume of the RT reaction was 10  $\mu$ L, the samples mixed and centrifuged (700 g, 10 sec, RT). The following program was used for the reaction and the synthesized cDNA either directly processed in the WT PCR or stored at -20  $^{\circ}$ C:

RT reaction & template switching	42 °C, 90 min	
1. Unfolding RNA sec. structure	50 °C, 2 min	} 10 ×
2. Completing RT & Template switch	42 °C, 2 min	
Enzyme inactivation	70 °C, 15 min	

### 3.2.3.2 WT amplification

Full-length cDNA that was tagged at both sides with the TSO anchor sequence, was amplified in a WT-PCR with the TSO-PCR\_67 °C primer (final concentration 0.2 μM). 15 μL PCR mix (0.2 μM TSO-PCR primer, 1x KAPA HiFi HotStart ReadyMix) was added to samples of section 3.2.3.1., the samples mixed and centrifuged (700 g, 10 sec, RT). The TSO-PCR primer was used as forward and reverse primer and the PCR performed in a final volume of 25 μL with the following program:

Initial denaturation	98 °C, 3 min	23 - 30 ×
1. Denaturation	98 °C, 20 sec	} (Depending on the RNA quality of the sample)
2. Primer annealing	67 °C, 15 sec	
3. Elongation	72 °C, 6 min	
Final extension	72 °C, 5 min	

### 3.2.3.3 Purification and quantification of PCR products

Primers and components of PCR kits that could affect the efficiency of following PCRs were removed after WT-PCR with the Agencourt AMPure XP Kit. PCR Cleanup was performed according to the manufacturer's instructions. It is based on a solid phase reversible immobilization (SPRI) technique that binds negatively charged DNA with carboxyl coated paramagnetic particles.<sup>[178]</sup> DNA was washed by applying a magnetic field that physically separates DNA bound magnetic beads and contaminants.

Primer-dimer carryover after WT-PCR was prevented with a bead to sample ratio of 0.6:1 and PCR products eluted in 10 μL TE buffer with low EDTA (AFFYMETRIX).

The DNA concentration of PCR products was determined fluorescently with a Qubit 3.0 Fluorometer according to the manufacturer's instructions.

### 3.2.4 NGS preparation

The WT of SC was fragmented with tagmentation as described by BAYM *et al.*<sup>[177]</sup> Each single well was uniquely barcoded during index PCR by using index primers with an Illumina sequencing adaptor as a 5'-overhang. Barcoded samples were pooled into one final sample, purified and the quality controlled before it was send for 2x 150 bp paired-end NGS sequencing (Figure 14).

#### 3.2.4.1 Tagmentation reaction

Tagmentation was performed with the Nextera DNA library preparation kit in a final volume of 2.5  $\mu\text{L}$ . An initial testing of different tagmentation input concentrations (concentrations: 1.1 ng/ $\mu\text{L}$ , 1 ng/ $\mu\text{L}$ , 0.8 ng/ $\mu\text{L}$ , 0.8 ng/ $\mu\text{L}$  in double the reaction volume, 0.5 ng/ $\mu\text{L}$ ) was performed and the fragment-size distributions assessed on a Bioanalyzer. Afterwards, tagmentation was set to 0.8 ng/ $\mu\text{L}$  and for samples with lower input concentration, the amount of used TDE1 enzyme downscaled accordingly.

PCR products after WT were diluted to a final concentration of 0.8 ng/ $\mu\text{L}$  and 1 $\mu\text{L}$  of this dilution added to 1.50  $\mu\text{L}$  tagmentation mix (1.25  $\mu\text{L}$  tagmentation DNA buffer, 0.25  $\mu\text{L}$  TDE1) that was prepared in a new 96-well plate on ice. Samples were mixed by pipetting up and down 10x and the tagmentation performed on a thermal cycler for 10 min at 55  $^{\circ}\text{C}$ . The plates were immediately put on ice and directly continued with the index PCR to avoid further fragmentation of WT by the still active TDE1, which would produce too short sequences.

#### 3.2.4.2 Index PCR

Illumina TrueSeq primers (Table 5, N701-N712 for column 1-12 and S501-S508 for row A-H) were used in the index PCR to uniquely barcode and add Illumina adaptors to each well of the 96-well plate. 11.20  $\mu\text{L}$  KAPA HiFi HotStart ReadyMix (2x) was added to each sample after tagmentation on ice. 4.40  $\mu\text{L}$  S5xx-primer (5  $\mu\text{M}$ ) was added to each row and 4.40  $\mu\text{L}$  N7xx-primer (5  $\mu\text{M}$ ) to each column of the plate. Samples were mixed and centrifuged (700 g, 10 sec, RT). Index PCR was performed on a thermal cycler with the following program:

Enzyme activation	72 $^{\circ}\text{C}$ , 3 min	
Stripping off/ Inactivation of TDE1	98 $^{\circ}\text{C}$ , 5 min	
1. Denaturation	98 $^{\circ}\text{C}$ , 10 sec	} 8 $\times$
2. Primer annealing	62 $^{\circ}\text{C}$ , 30 sec	

---

3. Elongation	72 °C, 30 sec
Final extension	72 °C, 5 min

### 3.2.4.3 Library pooling, purification and quantification

For high throughput analysis of 96-well plates it was not practical to do PCR purification and quality control for all samples. Samples with the same DNA concentration as tagmentation input gave a similar post-index-PCR DNA concentration. Thus, same amounts (4  $\mu$ L) of cDNA libraries were first pooled row-wise into 8-strip tubes. Next, libraries in 8-strip tubes were purified (bead to sample ratio of 0.75:1) with Agencourt XP beads and quantified with a Qubit Fluorometer as described in chapter 3.2.3.3. The quality of randomly selected cDNA libraries was checked on a high sensitivity DNA chip in an Agilent 2100 Bioanalyzer. Good cDNA library pools showed an equal fragment-length distribution between 200 bp and 1 kb and were then pooled in equimolar concentrations into a final pool. To minimize primer-dimer carryover, the final library pool was further purified (bead to sample ratio of 0.75:1) and quantified.

### 3.2.5 NGS and Data analysis

2x150 bp paired-end sequencing of the final library was performed on a HighSeq 1500 sequencer by the group of DR. HELMUT BLUM at the gene center Munich. An average sequence depth of 2-3 million reads (per lane) per single-cell was obtained. Bioinformatic analysis was performed in collaboration with EDUARDO BELTRÁN (comprising marker-gene expression matrix, heat-maps and t-distributed stochastic neighbour embedding (tSNE) projections).

Briefly, raw sequencing reads of scRNAseq data of each SC were analyzed with FastQC<sup>[179]</sup> for quality control and filtering. Next, reads were aligned to the UCSC hg19 reference genome using HISAT2<sup>[180]</sup> and used as featureCounts<sup>[181]</sup> input to count mapped reads. TCR and Ig receptor repertoire from scRNA-Seq data was obtained by using MIGEC-CdrBlast.<sup>[182]</sup> More information on the bioinformatical analysis, including the detailed R script and Seurat package to generate t-SNE plots for the single-cell data, can be found in the Supplemental Material of BELTRÁN *et al.*<sup>[176]</sup>

Marker-gene expression matrix were generated by the single-cell consensus clustering (SC3) tool. tSNE projections show dimensionally reduced data of scRNAseq and are preferentially used to visualize complex high-dimensional datasets in two dimensional plots.<sup>[183]</sup> Heat-maps show normalized expression levels of the NGS data from the selected cells and display mean

expression levels and thus average values of main discriminatory genes. In contrast, violin plots were used to further resolve the expression levels of certain cell markers on a single cell level.

### 3.2.6 Finding TCR TRAV17\*01/TRBV4-1\*01 in PBMCs

PBMCs from HLA-matching RIS patients were analysed with specific primers that were designed to specifically bind to the CDR3 region of TCR TRAV17\*01/TRBV4-1\*01 and cover all possible combinations of nucleotide positions by using wobble nucleotides.

CD8<sup>+</sup> (naïve and CCR7<sup>+</sup> antigen-experienced subpopulations), CD4<sup>+</sup> and CD19<sup>+</sup> B-cell populations were isolated from frozen or freshly prepared human PBMCs (3.2.1) by FACS sort. PBMCs were centrifuged (300x g, 6 min, 4 °C) and stained in 100 µL FACS staining solution (antibodies for the PBMC panel from Table 3 were diluted in FACS buffer) for 30 min on ice in the dark. Cells were washed with 300 µl FACS buffer and resuspended in 300 µL FACS buffer. CD8<sup>+</sup>, CD4<sup>+</sup> and B-cell subpopulations were sorted into tubes by FACS and bulk RNA of the samples preserved in 500 µL Trizol at -80°C.

RNA was extracted with the PureLink™ RNA Micro Kit according to the manufacturer's instructions for samples in TRIzol® Reagent. 25 ng carrier RNA was added before the chloroform and all steps after RNA transfer performed on ice. RNA was eluted in 12 µL RNase free water and either directly processed in the RT or stored at -80°C.

cDNA was synthesized from 3 µL crude RNA as follows. First 1 µL primer mastermix (α-chain RT and β-chain RT, final conc. In 10 µL of 1 µM) was added to the RNA on ice, incubated for 3 min at 72°C on a thermal cycler and cooled down to 42°C for 2 min. 6.25 µL RT mastermix (1.2 µM TSO\_Smart-seq2, 1 U RNaseOUT, 10 U SmartScribe enzyme, 1 mM dNTPs, 2 mM DTT, in 1x First strand buffer) was added and the samples incubated for 60 min at 42°, followed by an enzyme activation step for 10 min at 70°C. cDNA was purified by a 1:1 ratio with Agencourt AMPure XP Beads as described in section 3.2.3.3 and eluted in 10 µL TE.

The TRBV4-1\*01 specific CDR3 region was targeted by two rounds of nested PCR amplification with 1 µL purified cDNA, the primers (Trappist-161-TRB for/ Rev-spec3 and TRBV4\_1\*01 for/ Rev-spec3) and Phusion enzyme in a final volume of 25 µL as described in section 3.3.1. PCR products were analysed on a 2% agarose gel (3.3.2).

### 3.3 Molecular biological methods

#### 3.3.1 DNA amplification

DNA (~1 ng) was amplified by PCR in 1x Phusion High-Fidelity PCR Master Mix with HF Buffer and a final volume of 50  $\mu$ L. Primers were used in a final concentration of 0.15  $\mu$ M, samples mixed and centrifuged (700 g, 10 sec, RT). The following program was used for amplification:

Initial denaturation	98 °C, 30 sec	
1. Denaturation	98 °C, 20 sec	} 25 ×
2. Primer annealing	60 °C, 20 sec	
3. Elongation	72 °C, 30 sec	
Final extension	72 °C, 10 min	

DNA (~1 ng) was amplified by PCR in 1x iProof™ High-Fidelity PCR Master Mix and a final volume of 100  $\mu$ L. Primers were used in a final concentration of 0.5  $\mu$ M, samples mixed and centrifuged (700 g, 10 sec, RT). The following program was used for amplification:

Initial denaturation	98 °C, 3 min	
1. Denaturation	98 °C, 20 sec	} 40 ×
2. Primer annealing	53 °C, 20 sec	
3. Elongation	72 °C, 30 sec	
Final extension	72 °C, 10 min	

#### 3.3.2 Analysis of DNA fragments

Plasmid DNA, amplified PCR products and DNA fragments were analyzed for the correct fragment size and purified by agarose gel electrophoresis. Gels (0.9% - 2.5%) were prepared with LE agarose (20  $\frac{mg}{mL}$ , BIOZYME) in 1× TBE buffer and PeqGREEN (1:25000; VWR PEQLAB) added for nucleic acid visualization. Samples were prepared with 1× DNA loading buffer (10 mM Tris-HCl (pH 7.6), 0.03% bromophenol blue, 0.03% xylene cyanol, 60% glycerol and 60 mM EDTA; VWR PEQLAB) and run in 1× TBE buffer at 100V- 150 V until sufficient electrophoretic separation. Fragment sizes were estimated by comparison to loaded DNA Ladders (50 bp peqGOLD DNA Ladder, 0.5  $\frac{mg}{mL}$ , VWR PEQLAB). DNA bands were detected

with UV-light and documented on a gel imaging system with Vision-Capt Software (Quantum ST5; VILBER LOURMAT).

The quantity and quality of double-stranded DNA was determined by measuring the absorbance at 260 nm with a *NanoDrop 2000* UV-Vis spectrophotometer.

PCR products and plasmids were sequenced by the Genomics Service Unit of the Biology at the LMU Munich, with the protocol for "Cycle, Clean and Run BigDye v3.1". The data was analysed with the software CHROMAS (Version 2.5.1, TECHNELYSIUM PTY LTD.) and the online websites of IMGT/V-QUEST (ImMunoGeneTics) (INTERNATIONAL IMMUNOGENETICS INFORMATION SYTEM®)<sup>[184]</sup> and the Nucleotide Basic Local Alignment Search Tool (Blast) (NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION)<sup>[185]</sup>.

### 3.3.3 DNA isolation

Plasmid DNA from bacterial cells was isolated with QIAprep Spin Miniprep Kit or QIAHiSpeed® Plasmid Maxi Kit according to the manufacturer's instructions.

PCR products and digested plasmids were purified by agarose gel electrophoresis (3.3.2), and DNA bands visualized with a LED Blue light Transilluminator. Bands were excised with a scalpel and DNA extracted by using the MinElute® Gel Extraction Kit according to the manufacturer's instructions. DNA was eluted in 13 µL EB Buffer and either further processed or frozen at -20°C.

Highly concentrated DNA was obtained by ethanol precipitation. Glycogen (20 µg) and Pellet Paint™ Co-Precipitant (2 µL, NOVAGEN) was used as carriers to facilitate DNA precipitation. 10% 3 M sodium acetate (v/v) and 2.5% (v/v) ethanol was added in the given order and samples incubated (30 min, -80°C or O/N, -20°C) to precipitate the DNA. DNA was pelleted by centrifugation (4°C, 20.800 x g, 1 h) and washed twice with 200 µL 80% ethanol and centrifugation (4°C, 20.800 x g, 15 min).

For the generation of stable transfected eukaryotic cell lines (3.4.2), DNA was precipitated without the co-precipitant Pellet Paint. Further, remaining ethanol was removed under a sterile hood and the DNA pellet dried for 5 mins with open lid. The dried pellet was soaked in 2 µL

*ddH*<sub>2</sub>O and resuspended in RPMI medium without supplements to a final concentration of approximately  $50 \frac{\text{ng}}{\mu\text{L}}$ .

### 3.3.4 Restriction cloning

Digests were performed (16 U enzyme per 1 μg DNA, in 1x CutSmart or NEB 3.1 Buffer) in a final volume of 50 μL and samples incubated for two hours at 37°C.

Usually double-digests were performed, but when restriction sites were less than 100 bp apart, digests were done sequential with an enzyme heat inactivation (65°C, 20 min) step in between both digests. For restriction cloning, vectors and inserts were digested with the same restriction enzymes, samples purified by agarose gel electrophoresis and gel extraction. Before ligation, the restriction sites of vectors were dephosphorylated with the Rapid DNA Dephos and Ligation Kit according to the manufacturer's instructions.

Ligation was performed with 1 μL T4 DNA ligase in 1x T4 DNA Ligation Buffer and a final volume of 20 μL. 100 ng dephosphorylated vector was used for each ligation and the amount of insert determined as follows:  $\frac{\text{Insert length [bp]}}{\text{Vector length [bp]}} \times 10 \times 3 = \text{ng Insert per } 100 \text{ ng Vector}$ .

Control ligations without insert were performed to examine self-ligation of the vector. Samples were incubated (20 min, RT) and kept O/N at 16°C. Ligation products were transformed by heat shock into MAX Efficiency™ DH5α Competent Cells or One Shot™ TOP10 cells.

Single bacterial clones were isolated into 15 mL falcons with LB<sup>AMP</sup> medium and grown (1 hour, 37 °C, 225 rpm).

### 3.3.5 Recombinant production of 58αβ T hybridoma cells

Patient-specific α and β chains of TCR<sub>α1</sub> TRAV17\*01/TRBV4-1\*01 and TCR<sub>α2</sub> TRDV\*01/TRBV4-1\*01 were cloned individually into expression vectors as follows. PCR was performed with Phusion enzyme as described in section 3.3.1 and as a template 1 μL purified cDNA of SC-WT products from section 3.2.3 were used. Restriction cloning was performed as described in section 3.3.4.

TCR<sub>α1</sub> (TRAV17\*01 TRAJ23\*01) chain was amplified with the primers TRAPPIST-161\_TRA-for, Cα rev IN and cloned with the restriction enzymes Sall/ PvuII into the pRSV5hygro vector.

TCR<sub>β</sub> (TRBV4-1\*01) chain was amplified with primers TRAPPIST-161\_TRB-for, Cβ rev IN and cloned into pRSV5-neo vector by Sall/AvaI restriction digest.

TCR $\alpha_2$  (TRDV\*01) sequence was amplified with the primers TRAPPIST-161\_aII-for, C $\alpha$  rev IN and restriction cloned into pRSV5hygro with enzymes Sall/ PvuII.

Plasmid DNA was isolated (3.3.3) and analysed by sequencing (3.3.2). Then, T cell plasmids were linearized by XmnI digest (3.3.4), plasmids for  $\alpha$  and  $\beta$ -chains co-precipitated with ethanol (3.3.3) and electroporated under sterile conditions in 58 $\alpha\beta$  hybridoma cells (3.4.2).

Successful TCR surface expression was monitored by staining cells with anti-mouse CD3 $\epsilon$ -APC antibody and analysed by FACS (3.4.3).

Further, hCD8 $\alpha\beta$  and sGFP-NFAT plasmids were transfected as described in section 3.4.2 and stable 58V $\alpha_{17}$ V $\beta_{4.1}$  cells produced. CD8 expression was monitored by staining with anti-human CD8 $\beta$ -PE antibody and FACS analysis. sGFP-NFAT incorporation was checked by TCR activation with coated anti-mouse CD3 $\epsilon$  antibody for 16 hours. GFP-positive cells were analysed by fluorescence microscopy and FACS analysis (3.4.3).

### 3.3.6 Fusion of HLA molecules with mtagBFP2

To quantify and monitor HLA transfection efficiencies, mtagBFP2, which is a monomeric blue fluorescent protein with a highly stable blue fluorescence, was fused as a reporter-protein to HLA molecules.<sup>[186]</sup> Therefore, mtagBFP2 cDNA was cloned into the cytosolic domain of HLA-C07:02:01 in two steps by introducing a rarely cutting Not-I restriction site (Figure 28). The used primers, enzymes and plasmids were listed in Table 5-6.

First, HLA C\*07:02 was amplified from HLA\_C07:02\_pHSE-3' plasmid with the primers HLA\_C\_for\_Sal fwd and HLA C\*07\_Not rev.(3.3.1) The PCR product (1150 bp) was purified on a 2% agarose gel and sequenced (3.3.2). Double-digest of the PCR product and the HLA\_C07:02\_pHSE-3' plasmid was performed with Sall-HF and BamHI-HF in CutSmart Buffer. Digested insert (1128 bp) and vector (8072 bp) were purified on a 1.5% agarose gel and ligated. Ligation products were transformed via heat shock into DH5 $\alpha$  cells and grown colonies analysed by DNA sequencing.

Next, mtagBFP2 was amplified with the primers mtagBFP2\_Not\_fwd and mtagBFP2\_Stop\_BamHI\_rev (3.3.1). The PCR product (720 bp) was purified on a 2% agarose gel and sequenced (3.3.2). Subsequently the BFP-PCR product and the \*new\* HLA\_C07\*02 pHSE vector with the NotI restriction site from the prior step were sequentially digested with Not-I and BamHI in NEB 3.1 Buffer. Digested products were purified on an agarose gel and ligated overnight. Ligation products were transformed by heat-shock into DH5 $\alpha$  cells and grown colonies analysed by DNA sequencing.

Plasmids were transfected stably into COS-7 cells by electroporation (3.4.2) and the HLA-BFP expression monitored by FACS analysis (3.4.3).

All other BFP-HLA constructs were produced by adding the NotI restriction site into the cytosolic domain of the HLA sequences with an overhang reverse primer. PCR was performed with Phusion enzyme as described in section 3.3.1 and as a template 1  $\mu$ L purified cDNA of SC-WT products from section 3.2.3 were used. HLA A\*24:02 was amplified with the primers A101\_lead fwd and A23\_NotI rev. HLA B\*07:02 was amplified with the primers B0801\_lead fwd and B0801\_lead fwd. These new inserts were cloned into the BFP-HLA\_C07:02\_pHSE-3' plasmid in only one-step by using the restriction enzymes Sall/NotI as described above.

### 3.4 Cell handling

#### 3.4.1 Prokaryotic cells

Bacterial cells were cultured in autoclaved LB medium at 37°C and 225 rpm or plated on LB agar plates (0.15% Agar) at 37°C. Grown colonies on plates were counted manually and the concentration of cultures in suspension determined by measuring the OD (600 nm) on a UV/Visible spectrophotometer. Cells were diluted with LB medium, when the OD<sub>600nm</sub> was not within the linear range of the spectrophotometer. An OD<sub>600nm</sub> of 1.0 was estimated to  $8 \times 10^8 \frac{\text{cells}}{\text{mL}}$ .

The antibiotic Ampicillin ( $100 \frac{\text{mg}}{\text{mL}}$ ) was used for plasmid selection and diluted 1:1000 in LB medium. Transformed cells were kept at 4° for short-term storage and glycerol stocks (50% v/v) prepared for long-term storage at -20°.

*E.coli* cells were transformed with DNA plasmids either by heat shock transformation of MAX Efficiency™ DH5 $\alpha$  Competent Cells or One Shot™ TOP10 cells. For electroporation ElectroMAX™ DH10B™ cells were used and a Gene pulser.

For heat shock transformation, bacterial cells were thaw on ice and 5  $\mu$ L crude ligation product added. The transformation reaction was incubated (30 min, ice), the heat shock performed in a pre-heated water bath (30 sec, 42°C) and cooled down on ice for 2 min. 250  $\mu$ L SOC medium was added and the samples incubated (1 hour, 37°C, 225 rpm). 200  $\mu$ L was plated on LB<sup>AMP</sup> agar plates and incubated overnight at 37°C.

For electroporation, 2  $\mu$ L ethanol precipitated ligation product was added on ice to 20  $\mu$ L thawed bacterial cells and the sample transferred into pre-cooled 0.1 cm cuvettes (Gene

Pulser/MicroPulser Electroporation Cuvettes, 0.1 cm gap, BIO-RAD). The dry cuvettes were pulsed (2 kV, 200  $\Omega$ , 25  $\mu$ F) and 1 mL pre-warmed SOC medium immediately added. Transformed cells were transferred into 15 mL falcons and incubated for 1 hour at 37°C. Bacteria were either plated onto pre-warmed LB<sup>Amp</sup> agar plates and incubated at 37°C overnight or grown in 20 mL LB<sup>Amp</sup> medium.

### 3.4.2 Eukaryotic cells

Eukaryotic cells were cultured in RPMI-1640 co<sup>10</sup> medium under sterile cell culture hoods and grown at 37°C, 5% CO<sub>2</sub>.

Living cells were counted manually in counting chambers after trypan blue (0.1% trypan blue solution SIGMA, in 1x PBS) was added.

For long-term storage, cells (~1.5 Mio per vial) were frozen quickly in 1 mL freezing buffer in isopropyl alcohol containing freezing containers at -80°C and transferred to liquid nitrogen tanks after 2 days.

Frozen cells were taken in culture by thawing them quickly at RT, centrifugation (RT, 5 min, 300x g) to remove the DMSO containing freezing buffer. Cells were recovered for 2 days in RPMI co<sup>10</sup> medium before antibiotic selection was started.

Adherent cells such as COS-7 were loosened from vessel surfaces by the proteolytic enzyme trypsin. Cells were washed with 1x PBS, before trypsin-EDTA solution (GIBCO) was added. After incubation (7-10 min, 37°C, 5% CO<sub>2</sub>), the double volume stop solution was added and cells were centrifuged (RT, 5 min, 300x g) in 50 mL falcons.

Eukaryotic cells were transfected transiently for co-culture experiments (3.5.3) with FuGene® HD reagent as recommended by the manufacturer and by electroporation under sterile conditions when stable cell lines were prepared.

FuGene® HD transfection was used to transiently transfect COS cells with HLA plasmids and/or PECPLs. COS-7 cells (50.000 COS  $\frac{cells}{cm^2}$ ) were plated 3 hours before transfection. For transfection of a 6-well plate, 2  $\mu$ g DNA was prepared in 100  $\mu$ L RPMI medium without supplements and 7  $\mu$ L at RT pre-equilibrated FuGene reagent added. After an incubation of 12 min at RT, the transfection mix was added to the cells. For subpool analysis in a 24-well setting, the transfection conditions were scaled down proportionally. Transfected COS cells were used for co-culture experiments (3.5.3) 24-72 hours post-transfection.

For electroporation,  $2.5 \times 10^6$  cells were washed twice with RPMI medium without supplements and resuspended in 700  $\mu$ L RPMI medium without supplements. 15-30  $\mu$ g DNA was precipitated and washed as described in section 3.3.3 and resuspended in 100  $\mu$ L RPMI without supplements. Washed cells were transferred to pre-cooled 0.4 cm cuvettes (Gene Pulser/MicroPulser Electroporation Cuvettes, 0.1 cm gap, BIO-RAD) on ice and the pre-cooled DNA added carefully without producing air bubbles in the cuvette. Samples were incubated for 10 minutes on ice and dry cuvettes pulsed (280 V, 960  $\mu$ F) with a Gene pulser. After 10 minutes recovery on ice, electroporated cells were transferred into 10 mL pre-warmed RPMI  $co^{10}$  medium and incubated for 30 minutes at RT. Transfected cells were grown in RPMI  $co^{10}$  medium for 2 days to fully recover and then selected by different antibiotic concentrations.

Stable hybridoma cell lines for co-culture-experiments were produced in several steps. First, empty 58 $\alpha$ - $\beta$ - were transfected with co-precipitated and patient-specific TCR $\alpha$  and TCR $\beta$  plasmids. The TCR  $\beta$  chain of electroporated cells was selected by different G418 antibiotic concentrations ( $3 \frac{mg}{mL}$ ,  $2.5 \frac{mg}{mL}$ ,  $2 \frac{mg}{mL}$ ). 7-10 days later grown single clones were isolated individually and the TCR  $\alpha$  chain selected by different Hygromycin concentrations ( $6 \frac{\mu g}{mL}$ ,  $4 \frac{\mu g}{mL}$ ). Next, only 58 cells with successful TCR expression were co-cultured for 2 days with confluent CD8 $\alpha$  $\beta$ -GPE cells that transduces *h*CD8 $\alpha$  $\beta$ . Cells were selected with ( $0.5 \frac{\mu g}{mL}$ ). Finally, sGFP-NFAT plasmids were transfected into the cells by another round of electroporation and selected with Blasticidin S ( $6 \frac{\mu g}{mL}$ ). The successful incorporation of all plasmids was monitored by FACS analysis (3.4.3) and when necessary, the yield of CD8-expressing cells increased by FACS SC sorting.

BFP-HLA constructs (3.3.6) were transfected into COS-7 cells by electroporation as described above and selected with G418 ( $3 \frac{mg}{mL}$ ,  $2.5 \frac{mg}{mL}$ ,  $2 \frac{mg}{mL}$ ). To increase the transfection yield, high expressing BFP cells were selected under sterile conditions by FACS SC sorting.

### 3.4.3 Flow cytometry

Surface expression of certain markers or transfected plasmids of eukaryotic cells were analysed and monitored by analytical flow cytometry on a BD FACSVerse™. Cells from culture were washed twice with FACS Buffer and either directly proceeded with FACS analysis or stained with fluorescent-bound antibodies (Table 3). For staining, the antibodies were diluted 1:50 in

50  $\mu$ L FACS Buffer and the cells incubated for 30 min on ice in the dark. Samples were washed twice to remove excess fluorescent dyes and proceeded with flow cytometry analysis. Data was assessed with FlowJo (version 10.6.2, BDBIOSCIENCE).

### 3.5 Identification of target antigens

The expanded T-cell clone TCR TRAV17\*01/TRBV4-1\*01 that was identified in twin pair #161/#162 with the method from section 3.2 was *in vitro* reconstituted. TCR  $\alpha$  and  $\beta$  chains were cloned and stable transfected into 58  $\alpha\beta$  hybridoma cells along with *hCD8 $\alpha\beta$*  and NFAT-sGFP reporter plasmids (section 3.3.5).<sup>[167]</sup>

Further patient-specific HLAs were cloned to be expressed with randomized PECPL in COS cells as APCs during co-culture experiments (section 3.5.3). An overview of the method is depicted in Figure 10. To increase the transfection efficiency of PECPL libraries, a reporter-BFP tag was fused to already cloned HLAs and stable transfected BFP-HLA-COS cells prepared.

#### 3.5.1 Investigating HLA restriction

The EBV cell line was prepared from blood of patient #162 and immortalized by K. DORNMAIR. Those cells express the patient-specific HLAs (A24:02:01, B07:0201 and C07:02:01) and might present endogenous antigens on their surface that can be detected by 58V $_{\alpha 17}$ V $_{\beta 4.1}$  hybridoma cells.

100.000 EBV cells were added carefully into 100 $\mu$ L RPMI co<sup>10</sup> medium into wells of a 96-well round-bottom plate. The plate was centrifuged (100xg, 10 min, RT) and 100.000 58V $_{\alpha 17}$ V $_{\beta 4.1}$  hybridoma cells added. The samples were mixed with a pipette and co-cultured for 16 hours at 37°C. Cells were stained with anti-mouse CD3 $\epsilon$ -APC antibody to distinguish hybridoma cells from EBV cells in the subsequent FACS analysis. T cell activation was analysed on the fluorescence microscope and by analytic flow cytometry.

#### 3.5.2 Unbiased Identification of T cell antigens

For unbiased identification of CD8<sup>+</sup> T cell antigens from the *in vitro* reconstituted TCRs, the method of K. SIEWERT was used (schematic overview in Figure 10).<sup>[167]</sup>

In short, randomized PECPLs with a defined amino acid length were transiently transfected with FuGene (section 3.4.2) into HLA-BFP-COS-7 cells (50.000 COS  $\frac{cells}{cm^2}$ , 3.3.6). Single cell isolation experiments were performed in 3.5 cm dishes and scanning experiments either in 6-well or 24-well plates. After an incubation (37°C, 5% CO<sub>2</sub>) time of 24, 48 or 72 hours post-transfection, COS cells expressed those peptides on their cell surface and were used as functional APCs in T cell activation assays. Excess reagents were removed by washing the cells twice with 1x PBS and RPMI co<sup>10</sup> medium (for 6-well: 2mL, for 24-well: 0.5 mL) without phenol red added. T-cells (125.000  $\frac{cells}{cm^2}$ , 3.4.2, 3.3.5) were added in 1 mL RPMI co<sup>10</sup> medium without phenol red to the prepared COS-cell-APCs and incubated for 16 hours before analysis of the co-cultures on an inverse fluorescence microscope. sGFP fluorescence was detected (Filter<sub>excitation</sub>: 472/30 nm; Filter<sub>emission</sub>: 520/35 nm) and autofluorescent artefacts detected with the Cy3 channel (Filter<sub>excitation</sub>: 545/25 nm; Filter<sub>emission</sub>: 6050/70 nm). Additionally transmission light (TM) was used to differentiate fluorescent hybridoma cells from COS cells or artefacts. Images were taken with TM light, GFP and Cy 3 channel and either recorded by the software MetaMorph V6.3r6 and analysed with ImageJ 1.48i or the scan software of ZEN2pro (ZEISS) and analysed with the program ZEN 2 (blue edition, ZEISS).

### 3.5.3 T cell activation assay

Co-cultures were analysed after 16 hours by fluorescence microscopy. Detection was done either manually or with the scan software Zen2pro (ZEISS).

COS cells underneath green fluorescent hybridoma cells were isolated with a micromanipulator and glass capillaries into 7  $\mu$ L ddH<sub>2</sub>O and stored on ice for maximum three hours before proceeding with DNA amplification of pcDNA6-V5 plasmids. Antigen coding plasmids with putative activating mimotope sequences were amplified by a nested Outer and Inner PCR. Outer PCR was performed with primers pcDNA-for-1 and pcDNA-rev-1 (product size: 279 bp) and the iProof polymerase (section 3.3.1). Inner PCR was performed with pcDNA-2<sup>nd</sup>-for and pcDNA-2<sup>nd</sup>-rev10 (product size: 62bp) and the iProof polymerase, but in a final volume of 50  $\mu$ L (section 3.3.1). To control successful fragment enrichment, 10  $\mu$ L outer PCR product and 5  $\mu$ L inner PCR products were loaded on an analytical 2% agarose gel after each amplification (section 3.3.2).

Enriched sequences from different isolated samples were individually restriction cloned as described in section 3.3.4. Samples were first precipitated with ethanol (section 3.3.3) and restriction digest performed sequentially with AscI and NotI. Excess reagents and small digest

fragments (<17 nt) were removed with the QIAquick Nucleotide Removal Kit according to the manufacturer's instructions and ligated into *AscI/NotI* digested pcDNArc spacer vector. Ligated products were purified by ethanol precipitation and transformed into DH10B *E.coli* cells by electroporation (section 3.4.1).

Transformed cells were grown overnight in 20 mL LB<sup>Amp</sup> medium and 50 µL of a 1:1000 dilution plated on LB<sup>Amp</sup> agar plates, to estimate the plasmid number. The total number of clones was calculated from the bacterial colonies of the plates. Then plasmid DNA of the bacteria from the 50 mL falcons was isolated by Miniprep (section 3.3.3). The concentration of the plasmid DNA was determined by NanoDrop (section 3.3.2) and remaining transformed bacterial cells stored at 4°C. Isolated DNA from each sample was subsequently tested for re-activation or another round of single cell isolation in a T cell activation assay by transfection into COS-cells and co-culturing as described above.

In the first reactivation step, the pool with best activation capacity was chosen by the highest quantity of green fluorescent hybridoma cells. Antigen-encoding plasmids were enriched in the next step by splitting the original bacterial pool that was stored at 4°C randomly into 30 subpools, Miniprep isolation of each subpool and another round of COS-cell transfection and co-culturing. In each of this subpool rounds the potential mimotopes get enriched, which was evaluated by an increasing number of green cells in each step. Finally, the method was aiming for re-activation from single bacterial clones and sequencing of antigenic mimotopes.

### 3.5.4 B-cell antigens

Expanded plasma cell clones of patients #162 (IGHV3-33\*01; IGKV2-24\*01), #4526 (IGHV4-59\*01; IGKV1-39\*01), #4626 (IGHV3-15\*01; IGKV2-30\*01), #301 (IGHV1-69\*01; IGKV1-16\*02) were amplified from 1 µL purified cDNA of SC-WT products from section 3.2.3. PCR was performed with primers from Table 11 and Phusion enzyme as described in section 3.3.1. PCR products were analysed on a 2% agarose gel. Samples were handed over to MD student ISABELLE BARLIANTO for production of recombinant antibodies and testing them for antigen activation on ProtoArrays (Version 5.0; Human Protein Microarrays, INVITROGEN).

**Table 11: Combinations of primers for the amplification of expanded plasma cell DNA**

Patients	Heavy chain		Light chain	
	#162	162_PB-01-H-for	CHeavy SalI rev	162_PB-01-K-for
#4526-1	4X26_PB-H-01-for	CHeavy SalI rev	IGKV1_PB-01-K-for	Kappa_V,CKasIrev
#4626	4X26_PB-H-01-for	CHeavy SalI rev	4626_PB-01-K-for	Kappa_V,CKasIrev
#301-1	301_PB-01-H-for	CHeavy SalI rev	IGKV1_PB-01-K-for	Kappa_V,CKasIrev

## 4 Results

### 4.1 Single-cell profiling of patients with neuroimmunological diseases

Characterization of disease-relevant lymphocytes is fundamental for the understanding of basic immunological processes in autoimmune-, tumor- and infectious diseases. Typically, lymphocyte subpopulations are distinguished by immunohistochemistry or flow cytometry. Thus, their identification relies on the availability of antibodies to extracellular surface markers or - in some cases - intracellular proteins and is limited to a small number of molecular markers. Moreover, information on the hypervariable antigen-specific TCR or Ig regions, particularly on matching antigen-specific receptors cannot usually be obtained by flow cytometry.

Therefore, a single-cell profiling method that enables identification of antigen-specific receptors along with classical surface markers and whole transcriptomes was needed (schematic overview in Figure 11). The method combines single cell index sorting (scIS) with a modified protocol of the Smart-seq2 method for high-throughput sequencing of RNA-seq data (scRNA-seq).<sup>[147]</sup> scIS allowed to cope with the limited cell number of lymphocytes in CSF samples. Further, the optimized scRNA-seq protocol is suitable for the characterization of CSF resident immune cells and their receptor repertoires.

As part of my master thesis project, FACS index sorting and whole transcriptome amplification were already tested and established with PBMCs of healthy donors and CSF of an untreated patient with a CIS diagnosis (#4526).<sup>[171]</sup> Furthermore, it was confirmed by target-specific PCRs that full TCR and BCR sequences could be reconstituted from WT PCR products.

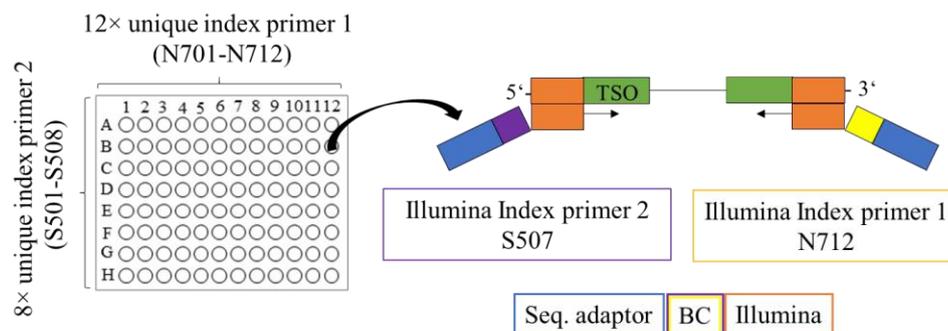
Here, it was continued with these promising results. First the protocol for NGS preparation was optimized to reduce the costs per sample (4.1.1). Next, scRNAseq was performed and validated with the CSF of index patient #4526 (4.1.2). Finally, the method was applied to analyse a unique cohort of monozygotic twins with discordance for their MS diagnosis and controls (4.2). The NGS HiSeq Runs (paired-end 2x150 bp) were performed by the group of DR. HELMUT BLUM and bioinformatics analysis in collaboration with EDUARDO BELTRÁN (3.2.5).

#### 4.1.1 Optimizing NGS preparation steps for high-throughput scRNA seq

Sample preparation for scRNA seq NGS run with the Illumina Nextera XT kit from the original Smart-seq2 protocol was too expensive for high-throughput analysis of human CSF samples. Thus, the protocol of BAYM *et al.* <sup>[177]</sup> for inexpensive multiplexed library preparation was established (Figure 14).

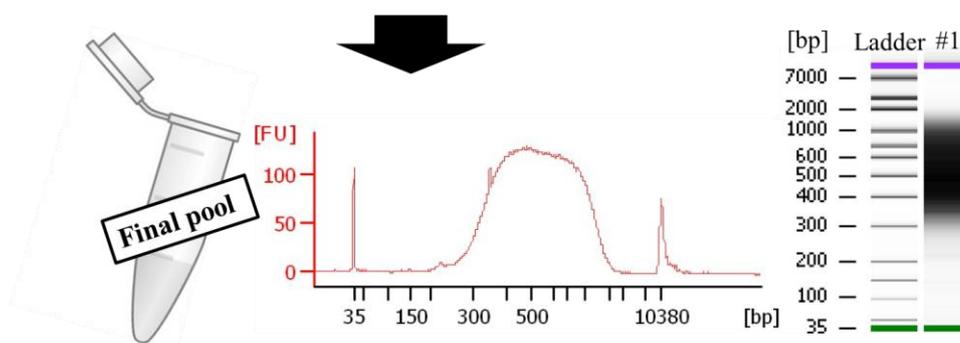
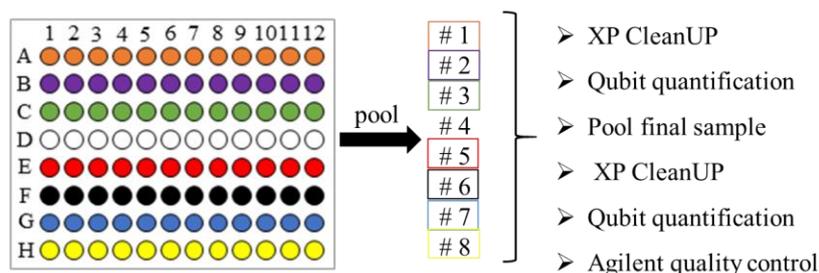
The volume of the tagmentation reaction was significantly reduced to 2.5  $\mu$ L and immediately proceeded with index PCR. Further, BAYM *et al.* demonstrated that using the same input DNA concentrations for tagmentation produced fragmented cDNA libraries with similar fragment-size distributions. Therefore, quality control of each single cell of the 96-well plate on an Agilent Bioanalyzer was omitted and only performed for randomly selected samples. This decreased expenses and work load as well as facilitating sample handling for NGS preparation. Same amounts of barcoded cDNA libraries from each row were directly pooled “crude” and cleaned up with XP beads (0.75x). The quality and quantity of each row-pool was assessed on a Qubit Fluorometer and Agilent Bioanalyzer. The final pool was assembled by pooling equimolar concentrations of row-pools. Hereafter, another purification step with XP beads (0.75x) was necessary to remove all remaining primer-dimers and fragments below 200 bp.

### 1) Tagmentation & Index PCR



- Illumina index primer 1&2 with barcodes (BC) for each well of a 96-well plate
- identifies DNA fragments from each SC

### 2) Library pooling, purification, quantification

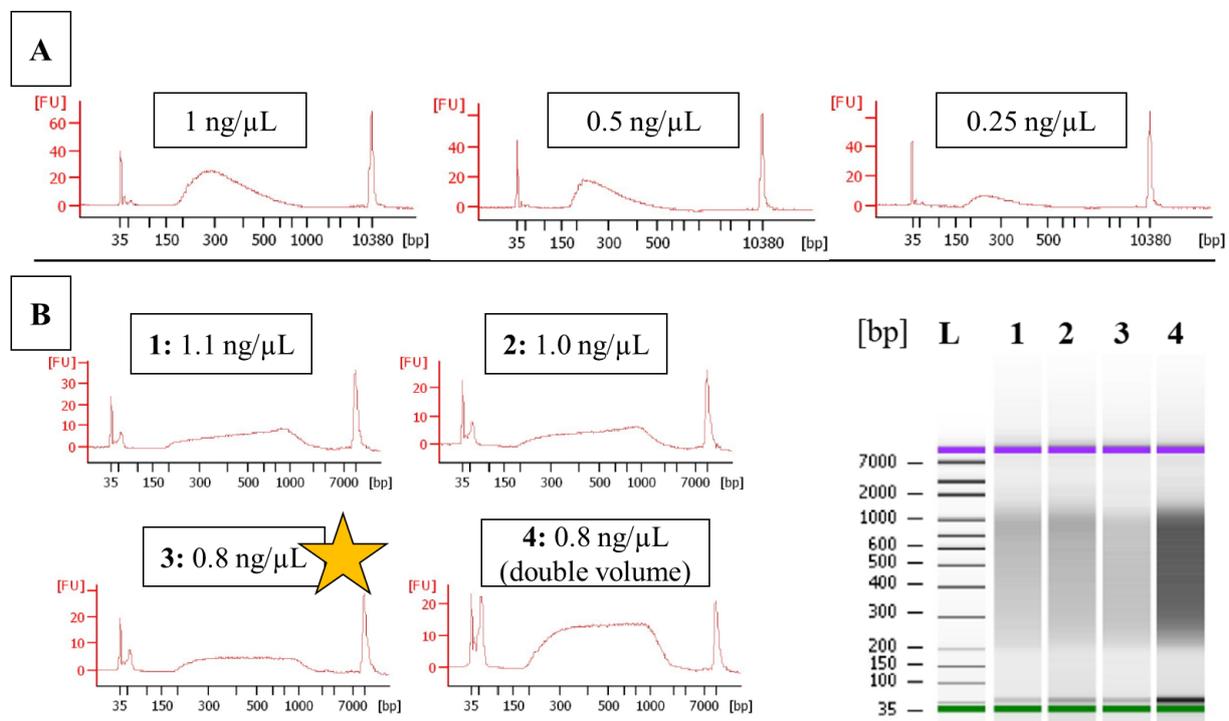


**Figure 14: Workflow of the library preparation steps for whole transcriptome analysis of a NGS HiSeq run.** (1) cDNA libraries were prepared for paired-end 2x 150 bp run as follows. Enriched and purified whole transcriptomes were enzymatically

fragmented into smaller pieces during tagmentation. Subsequent Index PCR with a unique combination of index primers for each single well of the 96-well plate tags all DNA fragments. Transcripts of each well were identified by this barcode in the bioinformatics data analysis. Then, barcoded libraries were purified with XP Agencourt beads and quantified with a Qubit fluorometer. (2) cDNA libraries of each well from the 96-well plate were pooled row-wise to 8 samples (#1-#8), purified and quantified. Equal amounts of each pool were combined to one final sample and purified again. The concentration of this final pool was determined and its quality assessed on an Agilent Bioanalyzer. Fragmented libraries showed an equal length distribution between 200 bp and 1 kb.

BAYM *et al.* found that an input concentration of 0.5 ng/ $\mu$ L works best for genomic DNA of gram-negative bacteria. Fragment sizes of final cDNA libraries should be equally distributed between 200 bp and 1 kb to ensure efficient scRNA sequencing. Smaller fragment sizes are too short to produce sufficiently long sequence reads for data analysis and bigger fragments are unable to bind to the Hi seq chip. Thus, the fragment size distribution which mainly depends on the tagmentation reaction is critical for the outcome of SC transcriptome analysis. Different input DNA concentrations from one single cell with different dilutions (0.25 ng/ $\mu$ L- 1 ng/ $\mu$ L) were screened, samples cleaned up with XP Beads (0.75x ratio) and analyzed on an Agilent Bioanalyzer Hiseq chip (Figure 15). Here, tagmentation produced the best results with 0.8 ng/ $\mu$ L. Hence, higher concentrated samples were diluted to 0.8 ng/ $\mu$ L prior to tagmentation and lower concentrated samples were used with proportional less tagmentation enzyme.

To ensure high enough DNA concentrations of all samples for this tagmentation step, the XP cleanUP after WT-PCR from the Smart-seq2 protocol was changed to a 0.6x ratio and the elution volume decreased to 10  $\mu$ L.



**Figure 15: Establishing and optimizing the critical tagmentation step of scRNAseq transcriptome preparation.** (A) To reduce NGS preparation costs, tagmentation was performed according to the protocol of BAYM *et al.* in a final volume of only 2.5  $\mu$ L. 1  $\mu$ L of whole transcriptome DNA from one single cell was differently diluted (concentrations: 1 ng/ $\mu$ L, 0.5 ng/ $\mu$ L and 0.25 ng/ $\mu$ L) and used as input DNA for tagmentation reaction and subsequent index PCR. Resulting fragment size distributions were analyzed on a high sensitivity DNA chip in an Agilent 2100 Bioanalyzer. (B) Input DNA concentrations (1.1 ng/ $\mu$ L, 1.0

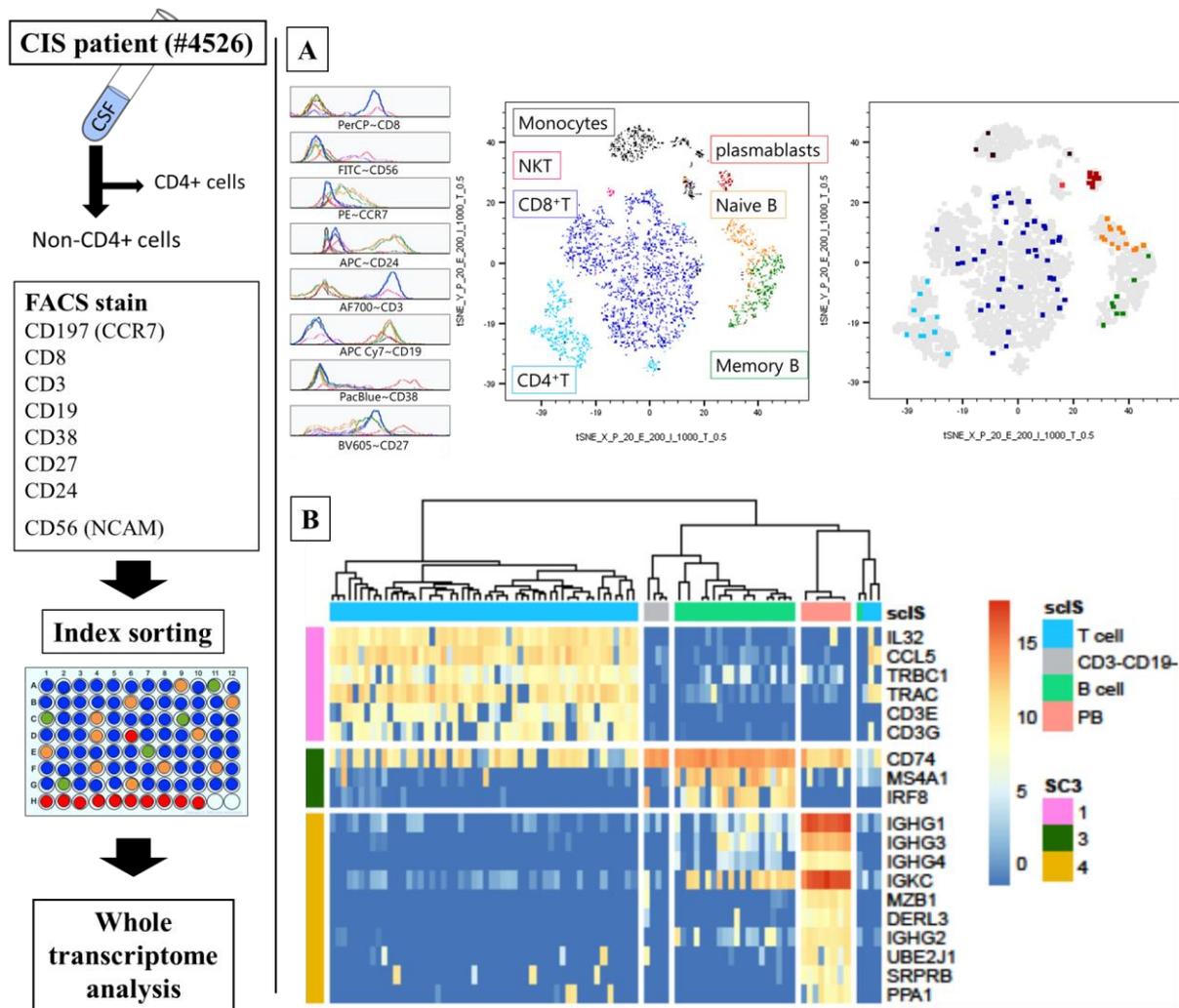
ng/ $\mu$ L, 0.8 ng/ $\mu$ L and 0.8 ng/ $\mu$ L in double the reaction volume) were fine-tuned with different dilutions of another single cell as described for (A).

Lastly, sequencing according to the original protocol from PICELLI *et al.* was performed with paired-end 100 bp NGS HiSeq. Although this gives sufficient sequence information for the transcriptome profile of SC, sequence reads are too short for getting full CDR3 identity of cells. With this, clonal expansions could only be identified with uncertainty by reconstitution with computational prediction methods. So, 2x 150 bp paired end sequencing was performed. As a proof of concept, the modified transcriptome approach was validated on the CSF sample of index patient #4526 (4.1.2).

#### **4.1.2 Proof of concept: single-cell profiling of human CSF of index patient #4526**

Exemplarily the method was applied to the CSF of CIS patient #4526. In this first test, CD4<sup>+</sup> T cells were depleted prior to flow cytometry because they constitute the overwhelming majority of cells in CSF (Figure 16). The remaining non-CD4<sup>+</sup> cells, i.e. CD8<sup>+</sup> T cells, diverse B cell subpopulations, NK(T) cells, macrophages, etc., were then stained for the lymphocyte surface markers CD3, CD8, CD19, CD27, CD38, and CD56 and subjected to flow cytometric scIS. In total, 94 single-cells were isolated and particularly “rare” cell types of interest enriched in the last row of the 96-well plate. For each well, the staining pattern indicates the phenotypic characterization of established cell types and lineages (CD8<sup>+</sup> cells in blue, CD19<sup>+</sup>CD27<sup>-</sup>CD38<sup>-</sup> naïve B-cells in orange, CD19<sup>+</sup>CD27<sup>+</sup>CD38<sup>+</sup> plasmablasts in red, CD19<sup>+</sup>CD27<sup>+</sup>CD38<sup>-</sup> memory B-cells in green, Figure 16A). Then, cells were processed for whole transcriptome analysis and scRNA-seq was performed by paired end 2x150 bp NGS HiSeq.

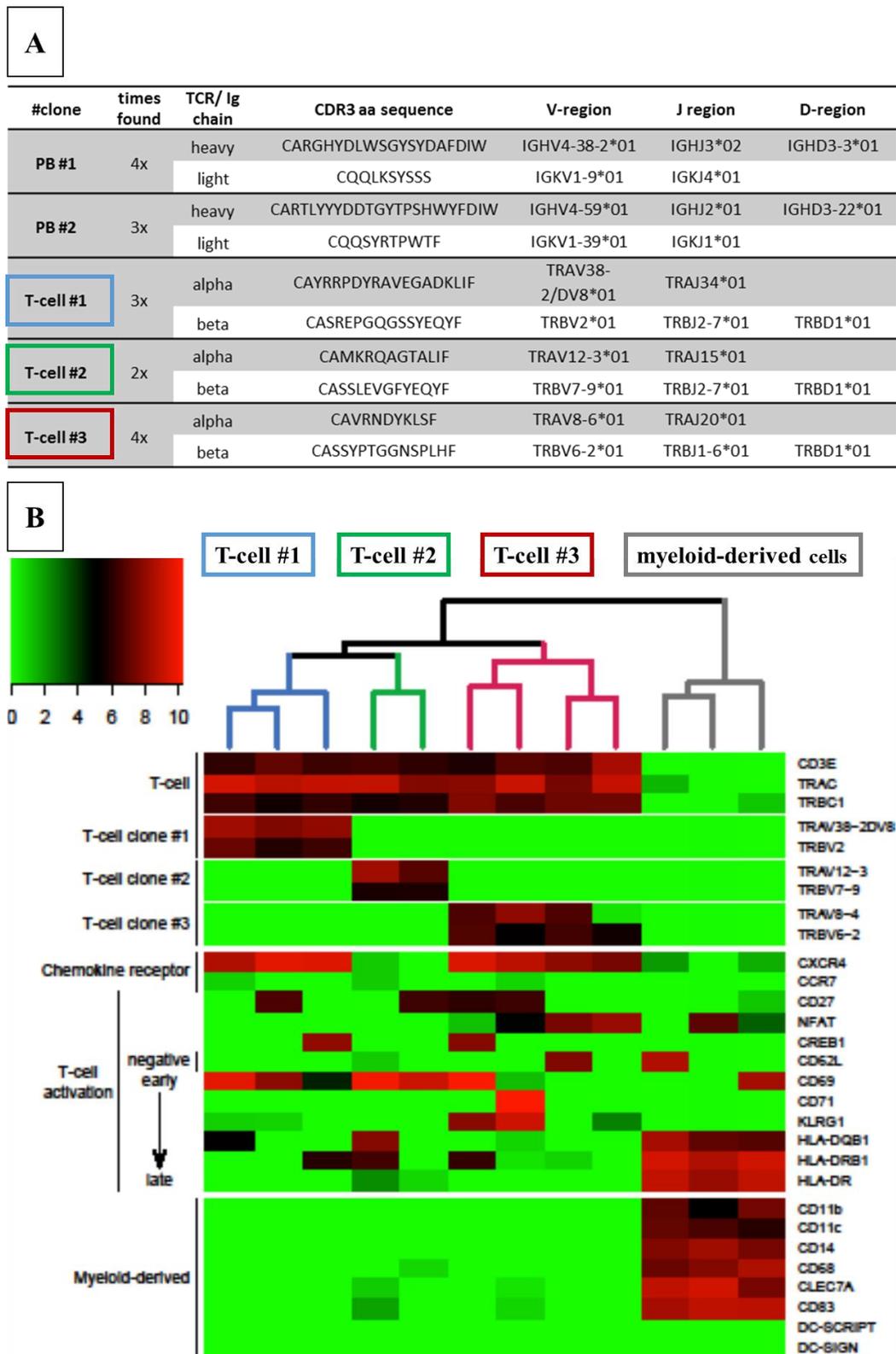
Analysis of the scRNA-seq data set using the single-cell consensus clustering (SC3) tool allowed the generation of a marker-gene expression matrix that indicates clonal relations among individual cells (Figure 16B).<sup>[187]</sup> T and B cells are clearly separated from each other and from other cells. Note that staining patterns from scIS (upper lane) correspond very well to the results obtained by scRNA-seq analysis. To show that it is possible to enrich particularly rare cell types of interest, ten CD19<sup>+</sup>CD27<sup>+</sup>CD38<sup>+</sup> cells, which amounted to less than 1% of CSF cells, were collected in wells H1 to H10 of the 96-well plate.



**Figure 16: Validation of the modified Smart-seq2 method with CSF from CIS patient #4526.** (Left panel) Brief overview of the performed steps for analyzing the CSF of index patient #4526. (Right panel) (A) tSNE analysis of flow cytometry scIS data of all single lymphocytes ( $n=3813$ ) and the index sorted cells ( $n=94$ ) are shown. (B) Marker-gene expression matrix were generated by the single-cell consensus clustering (SC3) tool. Five similarity-based clusters were identified by using SC3 that are separated by white vertical lines. A color scale (right panel) indicates the cell type classification with both scIS and SC3 (3 populations). Lane 1 shows the scIS cell classification: T cells ( $CD3^+$ , blue), B cells ( $CD19^+$ , green), PB (Plasmablast,  $CD19^+CD27^+CD38^+$ , salmon) and other cells ( $CD3^-CD19^-$ , grey). SC3 found 19 marker transcripts (lines 2 to 20) that distinguish cells in T cells (pink), B cells (dark green) and PB (light brown).

Altogether it was possible to identify 37  $\alpha:\beta$  TCR pairs from 41 collected  $CD8^+$  T cells (90% yield), and 25 H:L Ig receptor pairs from 31  $CD19^+$  cells (80% yield). Examples are given in Figure 17A. Even variations between distinct single T cells could be analyzed (Figure 17B). A heatmap displays T cell specific transcripts, early and late activation markers, and non-related transcripts. It not only distinguishes T cells and non-T cells, but also differentiates activation patterns of individual T cells.

In conclusion, the method enables the identification of matching  $\alpha:\beta$  TCR, or H:L Ig chains, and concomitantly the mRNA profiles of each individual cell from scRNAseq data. This might allow for the identification of presumably disease-relevant clones and reveal additional characteristics as cell lineages, subfamilies, activation status or cytokine profiles to further characterize interesting clones.



**Figure 17: Matching  $\alpha:\beta$  TCR and H:L Ig chains along with whole transcriptome analysis at single cell level from CSF of CIS patient #4526 (n= 94 single-cells). (A) Matching  $\alpha:\beta$  TCR and H:L Ig chain nucleotide sequences with full CDR3 sequences of single lymphocytes. Clonal expansion was defined as identical paired  $\alpha:\beta$  TCR and H:L Ig chain sequences in more than one single cell and examples were given for three expanded CD8<sup>+</sup> T-cells and two plasmablasts (PB). (B) Heatmap of three clonally-related T cell clones (TCR sequences given in panel A) and three CD3<sup>+</sup>C19<sup>-</sup> myeloid-derived cells reflecting gene expression values of TCR-specific (top 9 lines), activation marker (middle) and control (bottom) transcripts.**

## 4.2 Transcriptome profiling of patients with early and established MS

To study molecular peculiarities of immune cells at different stages of MS pathogenesis, the new approach was used to compare gene expression of single lymphocytes in CSF samples from early and established MS patients and controls (detailed clinical information in Table 9, section 3.1.10).

Briefly, the study comprised four MS patients with clinical manifested MS, six clinically healthy subjects with conspicuous signs of prodromal MS – defined here as subclinical neuroinflammation (SCNI, 5.2) – and four controls. As controls, CSF from two non-inflammatory subjects (NIC) with idiopathic intracranial hypertension (IIH) and two individuals with encephalitis (Enc.), an inflammatory, but non-MS-related disease was used.

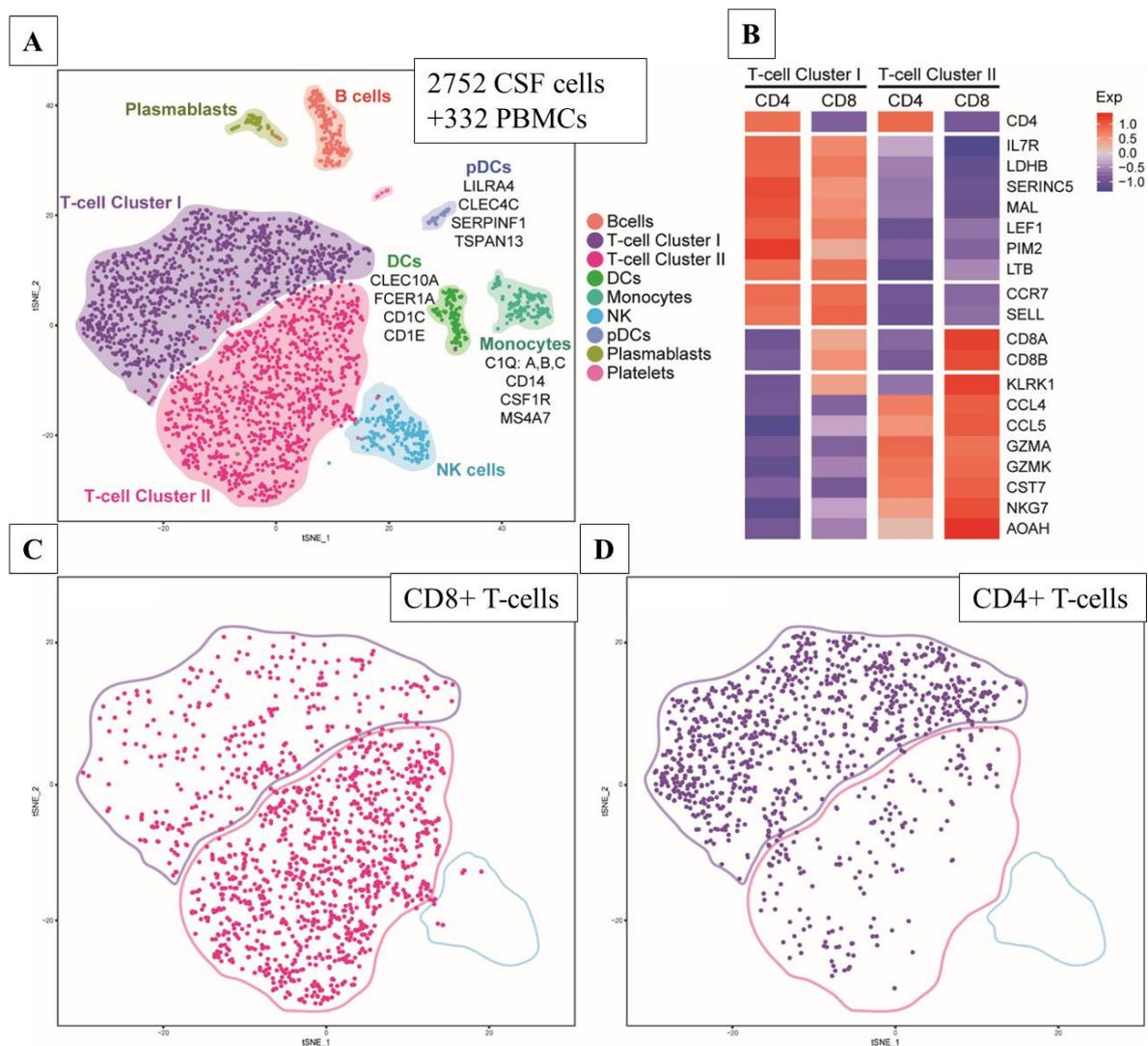
The NGS HiSeq Runs (paired-end 2x150 bp) were performed by the group of DR. HELMUT BLUM and all bioinformatics analysis in collaboration with EDUARDO BELTRÁN (3.2.5).

### 4.2.1 Cellular landscapes of CSF lymphocytes

The newly established approach that combines high-throughput scRNA seq with flow cytometry based index sorting was used to analyse a total of 2,752 CSF cells and 332 PBMC cells. Lymphocyte distributions in human blood are well studied, therefore PBMCs were used as spiked-in controls to facilitate and validate unbiased clustering in the performed tSNE analysis (Figure 18A). Of note, transcriptome data from PBMCs were excluded from any subsequent analysis. The comprehensive analysis of gene expression patterns of all single CD4<sup>+</sup> and non-CD4<sup>+</sup> cells combined identified nine clusters of lymphocyte subpopulations in human CSF that were coloured accordingly (Figure 21A). For the majority of clusters including plasmablasts, B-cells, NK cells and the two T-cell clusters, flow cytometric index sorting data were used to validate the particular population identity. The three remaining clusters were identified as DC, pDC and monocytes by common expression markers and the top discriminating genes listed underneath each cluster. For those cell types, no surface markers for extracellular staining were included.

T-cells were differentiated into two clusters that were classified by their gene expression profiles with a heat map (Figure 18B) and by using the individual index sorting data of CD8<sup>+</sup> (Figure 18C) and CD4<sup>+</sup> cells (Figure 18D). Cells of T cell cluster I show high expression levels of IL7R, CD4<sup>+</sup>, and homing molecules CCR7 and SELL which indicate a central memory phenotype. In contrast, T cell cluster II is characterized by high expression of transcripts for CD8 $\alpha$  and CD8 $\beta$ . It also comprises a minor population of CD4<sup>+</sup> expressing cells that might

result from an incomplete removal of CD4<sup>+</sup> cells in the positive selection steps that separates them from the remaining non-CD4<sup>+</sup> population. Further, T cells of T cell cluster II show a cytotoxic expression profile (granzymes, cytotoxic molecules) with a presumable effector functionality. Index sorting data confirmed the identity of both clusters and, moreover, revealed a spillover of cells in both clusters. Overall, T cell cluster I was assigned as predominately CD4<sup>+</sup> population, and T cell cluster II as predominately CD8<sup>+</sup>. Spillovers might result from incomplete separation of CD4<sup>+</sup> cells during positive selection or high similarity of gene expression markers. Colouring of cells was performed according to index sorting data.

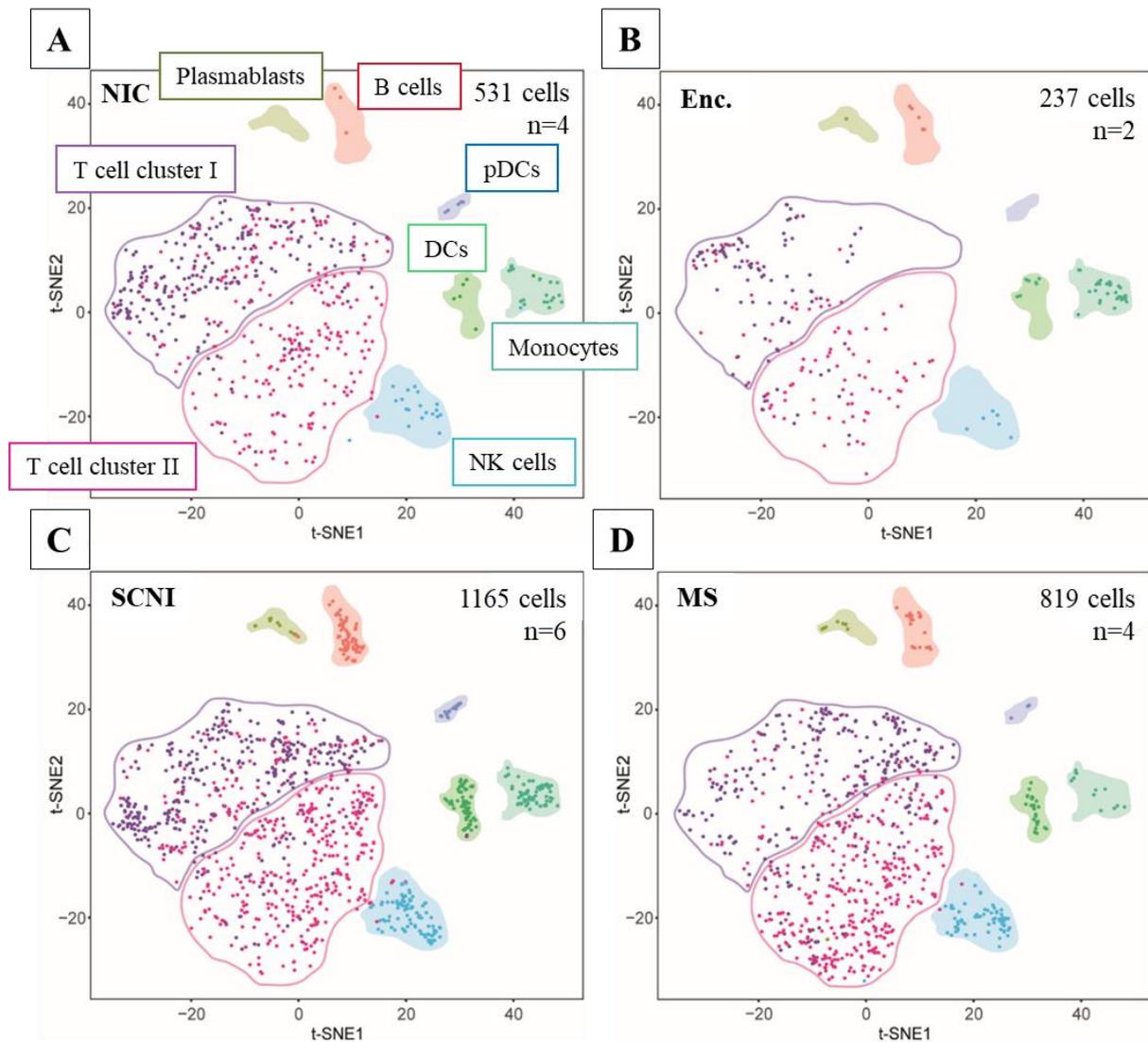


**Figure 18: Cellular landscape of lymphocytes out of all analysed cells by scRNA-seq.** (A) In total 2752 CSF cells and 332 PBMCs from 16 human subjects were processed with the scRNA-seq method and analysed bioinformatically (bioinformatic NGS analysis performed by EDUARDO BELTRÁN). The cellular composition of PBMCs is well studied, thus, their transcriptome data was used here as spike-in controls to perform unbiased clustering by tSNE analysis. PBMC data was then excluded from all following analysis and panels. 9 clusters (B-cells, T-cell cluster I, T-cell cluster II, DC, monocytes, NK cells, pDC, plasmablasts and platelets) were identified according to extracellular FACS staining and common cell-type specific expression markers. DC, pDC and monocytes were not stained by our FACS panel, therefore, the main discriminating genes were listed underneath. Of note, roughly the same numbers of CD4<sup>+</sup> cells and Non-CD4<sup>+</sup> cells were analysed, this does not reflect the actual proportions of CD4<sup>+</sup> cells to Non-CD4<sup>+</sup> cells in human CSF. Each dot represents one single cell and was highlighted with the colour of the respective cluster. (B) Heat-map of normalized mean expression levels of main discriminatory genes found in T-cell cluster I and II from panel (A). (C) t-SNE projection showing only CD8<sup>+</sup> cells in T-cell clusters I and II according to the FACS index sorting data. (D) t-SNE projection showing only CD4<sup>+</sup> cells in T-cell cluster I and II. CD4<sup>+</sup> cells were selected according to extracellular FACS index sorting data. (A-D) taken and modified from BELTRÁN *et al.*<sup>[176]</sup>

It was possible to resolve cellular landscapes of lymphocytes from human CSF on a single cell level for different MS disease stages (Figure 19A-D). Analysed individuals were separated into groups as described above, though two healthy co-twins that did not show signs of subclinical neuroinflammation were included in the NIC subgroup. Detailed t-SNE projections showing only cells for the individual subgroups NIC (n=4; 531 cells; Figure 19A), Enc (n=2; 237 cells; Figure 19B), SCNI (n=6; 1165 cells; Figure 19C) and established MS (n=4; 819 cells; Figure 19D) reveal a similar composition of cell types throughout all study groups. Immune cells of the adaptive immune system such as T-cells, B-cells and NK cells as well as cells of the innate immunity such as monocytes and DC were present in all groups.

As expected, the overwhelming majority of cells in CSF were T cells, and the ratio of cells from T cell cluster I and II was preserved in all disease stages. Of note, the absolute numbers of CD4<sup>+</sup> cells here do not reflect their proportions in CSF. Interestingly, plasmablasts (PB) are already present in SCNI subjects, which was so far an exclusive discriminatory characteristic of patients with manifest MS. Detailed analysis of the individual SCNI subjects revealed a strict correlation of the existence of PB with a positive OCB<sup>+</sup> status. Further, elevated numbers of DC and pDC were found in SCNI and MS subjects.

Taken together, the results of the tSNE analysis reveal no striking differences between the different disease stages. However, the presence of PB already in SCNI subjects was distinctive and so far unknown.



**Figure 19: Cellular landscapes of lymphocytes from human CSF in different disease stages analysed by scRNA-seq. (A-D) t-SNE projections showing the 9 different lymphocyte cell type clusters, (as assigned in Figure 18) of each analysed subgroup (A: NIC, B: Enc., C: SCNI, D: MS) in individual panels. The respective number of cells and studied individuals in each subgroup are stated in each panel. (A-D taken and modified from BELTRÁN *et al.*)<sup>[176]</sup>**

#### 4.2.2 Clonally expanded lymphocytes

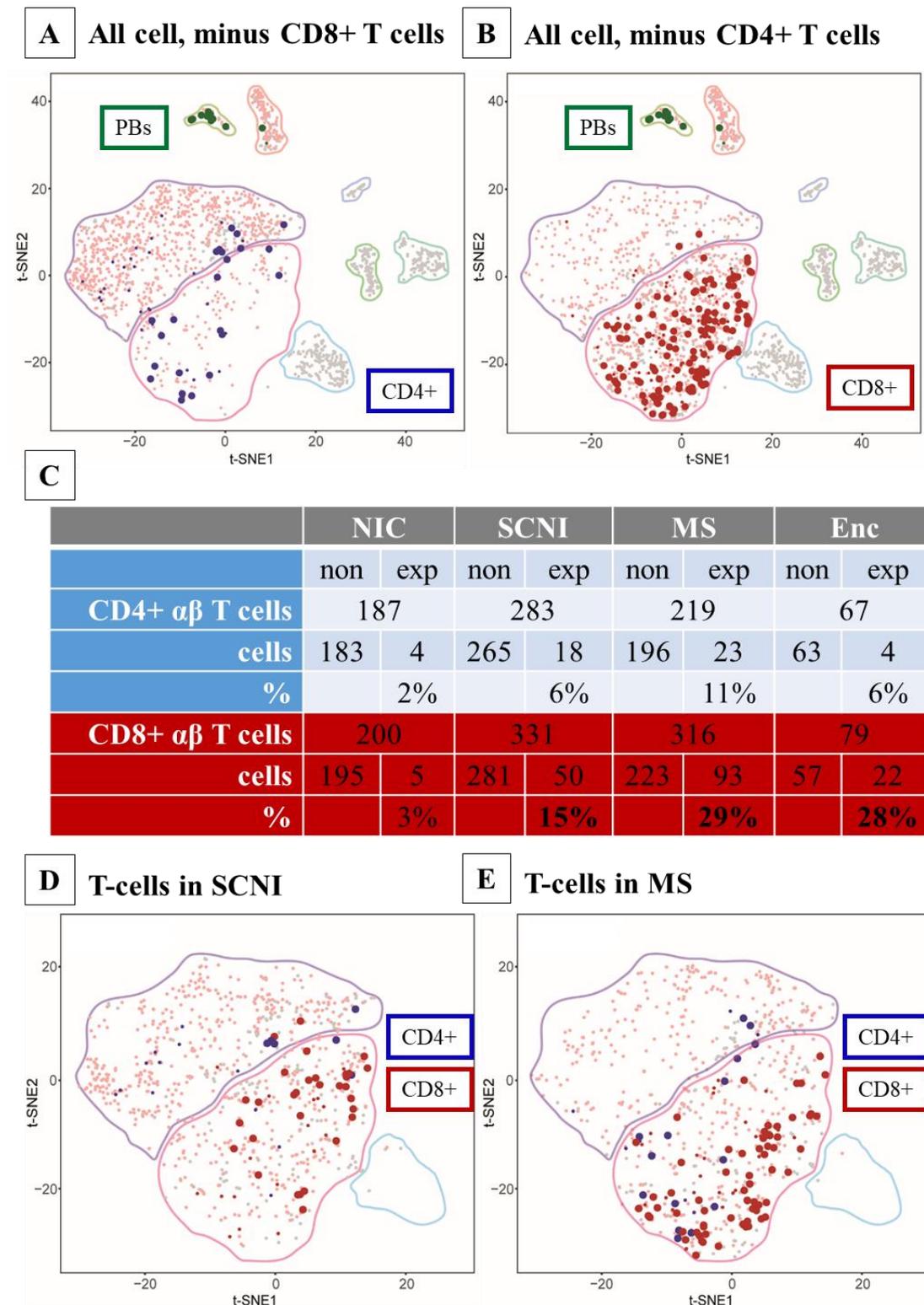
This state-of-the-art method not only allows the characterization of cell subtypes by their gene expression profiles, but 2x150 bp NGS sequencing also provides sufficient sequence information to link these transcriptome data to their antigen-specific receptors (Figure 20).

This enables tracking of the clonal history of single lymphocytes in the particular tSNE projection and was used to identify clonal expansion of receptor chains in T and B cells of all subject groups (Figure 20A+B). Clonal expansion of lymphocytes result from antigen-driven processes and are associated with pathogenic relevance. Hence,  $\alpha$ : $\beta$  TCR and H:L BCR chain sequences of T and B cells were analysed and clonal expansion identified when CDR3 sequence identity was detected in more than one individual lymphocyte cell. Clonal expansions were

detected in all clusters containing T and B lymphocytes with a strikingly dominant expansion of CD8<sup>+</sup> T cells. Moreover, CD4<sup>+</sup> and CD8<sup>+</sup> expansions were detected mainly in T cell cluster II. Interestingly, 90% of all PB from MS patients and 20% of PB in SCNI subjects were clonally expanded.

A detailed distribution analysis of expanded CD4<sup>+</sup> and CD8<sup>+</sup> cells in the individual study groups (Figure 20C) reveals CD8<sup>+</sup> and CD4<sup>+</sup> expansions in all inflammatory conditions (SCNI, MS, Enc) with an overall lower frequency of CD4<sup>+</sup> expansions. Since percentages of expanded cells were normalized to the individual number of all detected cells for each cell type, the absolute numbers of analysed CD4<sup>+</sup> cells and the thus distorted CD4<sup>+</sup>/CD8<sup>+</sup> ratios from this study are not relevant here. While 29% of all CD8<sup>+</sup> T cells in MS and 15% in SCNI subjects are clonally expanded, only 11% of CD4<sup>+</sup> cells in MS and 6% in SCNI are expanded. In general, expanded CD8<sup>+</sup> cells from SCNI subjects (Figure 20D) group in the upper right area of T cell cluster II, while expanded CD8<sup>+</sup> cells from MS patients (Figure 20E) predominately accumulate in the lower and central left area of T cell cluster II. Similar tendencies can be seen for the distribution of expanded CD4<sup>+</sup> cells. This rather uneven distribution in the tSNE projection reveals different gene expression profiles, which cannot be further resolved by tSNE analysis alone.

Overall, clonal expansions were found in all three lymphocyte populations (B-cells, CD4<sup>+</sup> and CD8<sup>+</sup> T cells) of the adaptive immune system and are a substantial characteristic of CSF from subjects with SCNI.



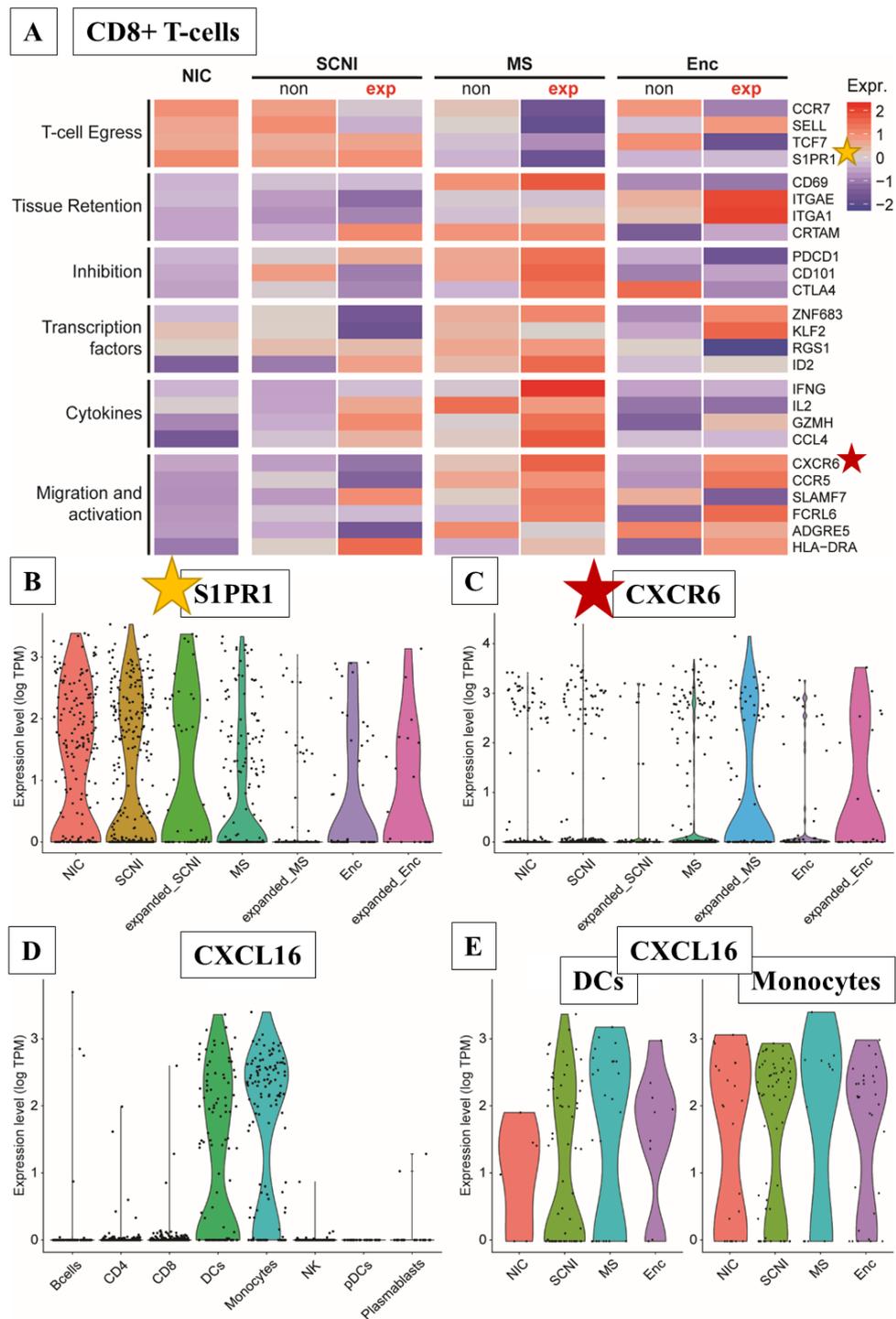
**Figure 20: Identification of clonally expanded lymphocytes of the adaptive immune system in human CSF.** (A+B) t-SNE projections of all clonally expanded lymphocytes. Expanded PB are depicted with green dots, expanded CD4<sup>+</sup> T cells in blue and expanded CD8<sup>+</sup> T cells in red. Light red dots indicate T-cells with at least one identified transcript of TCR or BCR chains and bigger dots indicate clonal expansion. Dot sizes correlate with the number of found clones with an identical lymphocyte receptor. Cells of the innate immune system are shown in grey. CD4<sup>+</sup> and CD8<sup>+</sup> T cells were identified according to index sorting staining data and the presence of at least one receptor chain. (C) Differentiated analysis of absolute numbers and relative amounts of expanded CD4<sup>+</sup> and CD8<sup>+</sup> T cells in all study groups (NIC, SCNI, MS, Enc). Percentages of expanded clones for each cell population were normalized to the absolute number of cells found in the individual subgroups. Bold letters indicate statistical significance ( $P < 0.05$ ) according to analysis by  $\chi^2$  test (NIC: not significant; SCNI:  $P = 0.00009$ ; MS:  $P = 0.00001$ ; Enc:  $P = 0.00013$ ). (D+E) Detailed t-SNE projections showing only clonally expanded CD4<sup>+</sup> and CD8<sup>+</sup> T-cells from SCNI (D) and MS (E) subjects. (A-E taken and modified from BELTRÁN *et al.*)<sup>[176]</sup>

### 4.2.3 In-depth characterization of expanded T cells

To gain deeper insights into the different expression profiles of expanded CD8<sup>+</sup> T cells and characterize those presumably pathogenic cells in more detail, mean expression levels of particular selected T cell marker genes were compared in a heat-map (Figure 21A). 25 T cell markers were selected to discriminate T cell egress, tissue retention, inhibition, transcription factors, cytokines, the migratory and activation status of expanded CD8<sup>+</sup> clones and non-expanded CD8<sup>+</sup> cells in all study groups (NIC, SCNI, MS, Enc). For NIC, cell numbers of expanded clones were too low to be differentiated into two individual groups. Discriminative markers involved in T cell egress such as homing receptors S1PR1, CCR7, TCF7 (TCF-1) and SELL (CD62L) are gradually downregulated from NIC to non-expanded and expanded cells of SCNI subjects to non-expanded and expanded cells of patients with MS (top panel). In contrast, expression of CD69, known as a very early indicator for T cell activation as well as a key marker for tissue resident memory cells (T<sub>RM</sub>), is strongly upregulated in expanded clones of MS patients, but not in SCNI (second panel). Further, expression of tissue retention markers ITGAE (CD103), ITGA1 (CD49a) and CRTAM which are also involved in the development of T<sub>RM</sub> cells are slightly upregulated only in expanded and non-expanded CD8<sup>+</sup> cells from MS subjects (second panel). The upregulation of T cell inhibition markers is a distinct characteristic of a T<sub>RM</sub> phenotype and comprises the gene expression markers PDCD1 (PD-1), CD101 and CTLA4, which are notably upregulated in expanded CD8<sup>+</sup> cells of MS (third panel). The recently identified T<sub>RM</sub> cell key marker CXCR6 is also upregulated solely in expanded cells of MS patients and cannot be found on cells of SCNI subjects (bottom panel). This tendency towards gene upregulation throughout the subject groups seems to also be consistent on a transcriptional level. The key transcriptional regulator of T<sub>RM</sub> cells ZNF683 (Hobit) is strikingly upregulated in cells of MS subjects, but not in SCNI (fourth panel). Particularly, it is directly involved in downregulation of the downstream transcription factor KLF2 and ultimately controls transcripts that are involved in T cell egress (fourth panel). This pattern can strikingly only be found in expanded CD8<sup>+</sup> cells of the MS subgroup. The same applies to the upregulation of pro-inflammatory cytokines such as INF $\gamma$  and IL-2 (fifth panel) and T cell activation markers CXCR6 and CCR5 (bottom panel). Other T cell markers for activation and migration including SLAMF7 and FCRL6 are also particularly upregulated in the expanded CD8<sup>+</sup> T cell population of MS subjects (bottom panel). This upregulation can also be seen to a smaller extent in cells of SCNI but not as concisely and pronouncedly as in the MS group.

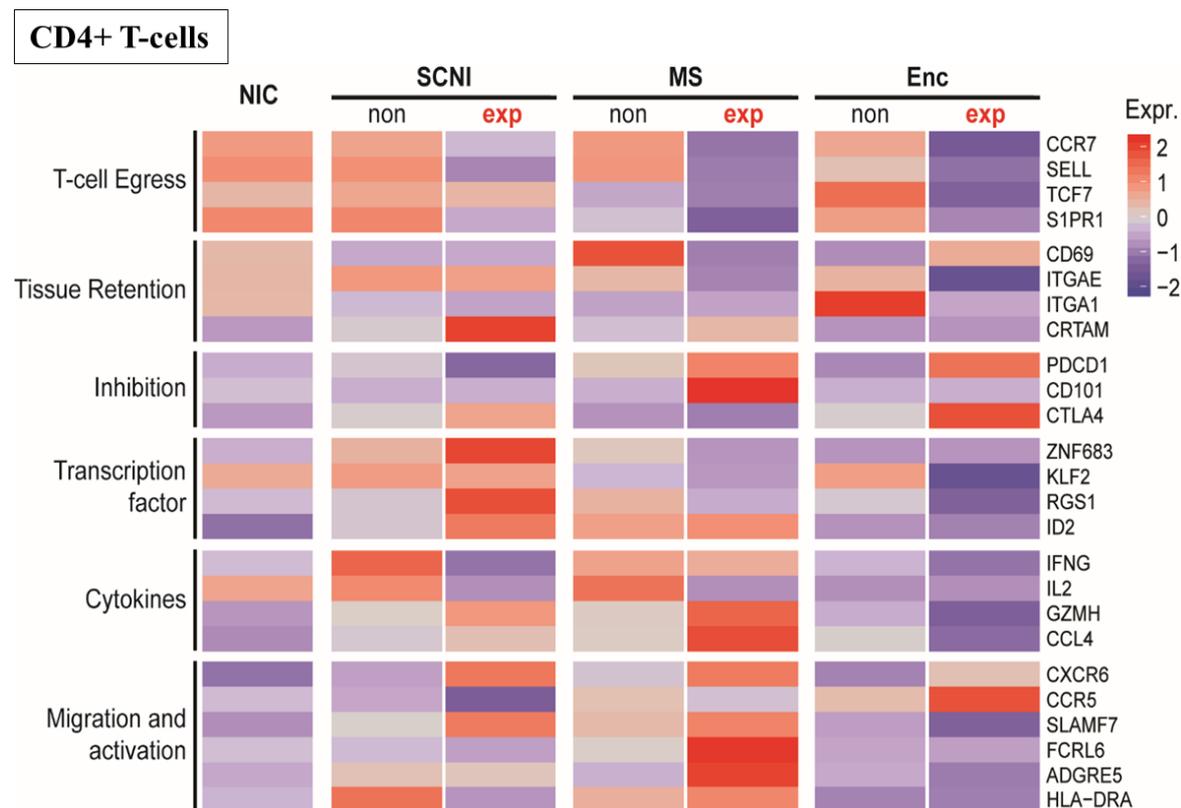
Taken together, the heat map analysis of expanded CD8<sup>+</sup> T cells from all subgroups reveals distinct features of an activated T<sub>RM</sub> phenotype in expanded CD8<sup>+</sup> T cells of patients with definite MS, which can be already detected to a lesser extent in SCNI subjects.

Violin plots were used to further characterize the expression of the key T<sub>RM</sub> cell markers CXCR6 and S1PR1 in single expanded CD8<sup>+</sup> cells of all study groups. Compared to heatmaps that only display average values of all cells from each group, violin plots resolve gene expression on a single cell level (Figure 21B-E). As seen in the heatmap, the homing receptor expression of S1PR1 is only downregulated significantly in expanded CD8<sup>+</sup> T cells from MS subjects (Figure 21B). Moreover, the key T<sub>RM</sub> cell marker CXCR6 is only significantly upregulated in expanded CD8<sup>+</sup> T cells of MS and Enc patients (Figure 21C). In proinflammatory conditions, e.g. high levels of IFN- $\gamma$  as seen for expanded CD8<sup>+</sup> T cells in MS (Figure 21A), CXCR6 with its sole ligand CXCL16 is crucial for recruitment of CD8<sup>+</sup> T cells. Therefore, expression levels of CXCL16 in all cell types (B-cells, CD4<sup>+</sup> and CD8<sup>+</sup> T cells, DC, monocytes, NK cells, pDC, PB) were analysed. Strikingly, CXCL16 is only significantly upregulated in DC and monocytes (Figure 21D) throughout all patient groups (Figure 21E). As IFN- $\gamma$  is only highly expressed in expanded cells from MS patients, these results might indicate an involvement of CXCL16-expressing DC and monocytes in the recruitment of activated, expanded CD8<sup>+</sup> T cells with a distinct T<sub>RM</sub> phenotype into the CSF of patients.



**Figure 21: Phenotypical characterization of clonally expanded CD8<sup>+</sup> T-cells in patients with early and established MS and controls.** (A) Heatmap showing gene expression levels of selected T cell markers for expanded CD8<sup>+</sup> T cell clones in all study subjects. CD8<sup>+</sup> T cells were identified by extracellular staining information obtained by flow cytometry and the presence of at least one TCR chain from the transcriptome data. 25 selected T cell marker genes are represented in this heat-map for non-expanded and expanded CD8<sup>+</sup> T cell clones from each subpopulation (NIC, SCNI, MS and Enc). Only for NIC, the number of expanded cells was too low, thus, no distinction could be made. Clonal expansion was defined when CDR3 identity was detected in more than one single CD8<sup>+</sup> T cell. The heat map was created with normalized expression levels of the NGS data from the selected cells and mean expression levels coloured according to their z-scores ranging from -2.5 (blue) to +2.5 (red). For better visualization, a yellow and red star highlight selected genes that were analysed in more detail in panels (B-C). (B-E) Violin plots showing gene expression levels of selected genes on a single cell level. Each dot represents one single cell and violin shapes were only formed when differences were statistically significant. Expression levels of tissue retention marker S1PR1 (B) and CXCR6 (C) of expanded CD8<sup>+</sup> T-cells (as selected in panel A) throughout all study groups (NIC, SCNI, MS, Enc.) are shown. The solute agonist of CXCR6 receptor is CXCL16. Therefore, its gene expression was analysed in the different clusters of cell types (B-cells, CD4, CD8, DC, monocytes, NK cells, pDC, PB) in panel (D) and more detailed in DC and monocytes from the different study groups (NIC, SCNI, MS, Enc.) in panel (E). (A-E taken and modified from BELTRÁN *et al.*)<sup>[176]</sup>

Although the majority of all expanded cells in CSF were CD8<sup>+</sup>, expanded CD4<sup>+</sup> T cells were also detected in all patient groups and mean gene expression levels of selected T cell marker genes compared by heat map analysis as shown for CD8<sup>+</sup> T cells (Figure 21A). The same 25 T cell markers used for in-depth analysis of CD8<sup>+</sup> cells were selected to differentiate expanded CD4<sup>+</sup> clones and non-expanded CD4<sup>+</sup> cells from all study groups (NIC, SCNI, MS, Enc). Overall, similar distinctive T<sub>RM</sub> cell signatures for up- and downregulation of the selected T cell markers can be seen. This phenotype also gradually increases from NIC to non-expanded and expanded CD4<sup>+</sup> T cells of SCNI to those from MS. Compared to the T<sub>RM</sub> phenotype of CD8<sup>+</sup> T cells, the characteristic phenotype in the CD4<sup>+</sup> T cell population is not as pronounced. Taken together, these results indicate a clear distinction of expression markers of clonally expanded and non-expanded T cells. The expression profile of clonally expanded CD8<sup>+</sup> T cells in MS patients shows a characteristic phenotype of activated T<sub>RM</sub> cells that was also detectable in expanded CD4<sup>+</sup> T cells of MS, though not as pronounced. Further, the detected clonally expanded T cells in SCNI subjects show a less strong but detectable phenotypic shift to a T<sub>RM</sub>-like expression profile.



**Figure 22: Heatmap showing mean gene expression levels of selected T cell markers for expanded CD4<sup>+</sup> T cells.** CD4<sup>+</sup> T cells were identified by extracellular staining information obtained by flow cytometry and the presence of at least one TCR chain from the transcriptome data. 25 selected T cell marker genes are represented in this heat-map for non-expanded and expanded CD4<sup>+</sup> T cell clones from each subpopulation (NIC, SCNI, MS and Enc). Clonal expansion was defined when CDR3 identity was detected in more than one single CD4<sup>+</sup> T cell. The heat map was created with normalized expression levels of the NGS data from the selected cells and mean expression levels coloured according to their z-scores ranging from -2.5 (blue) to +2.5 (red). (taken and modified from BELTRÁN *et al.*)<sup>[176]</sup>

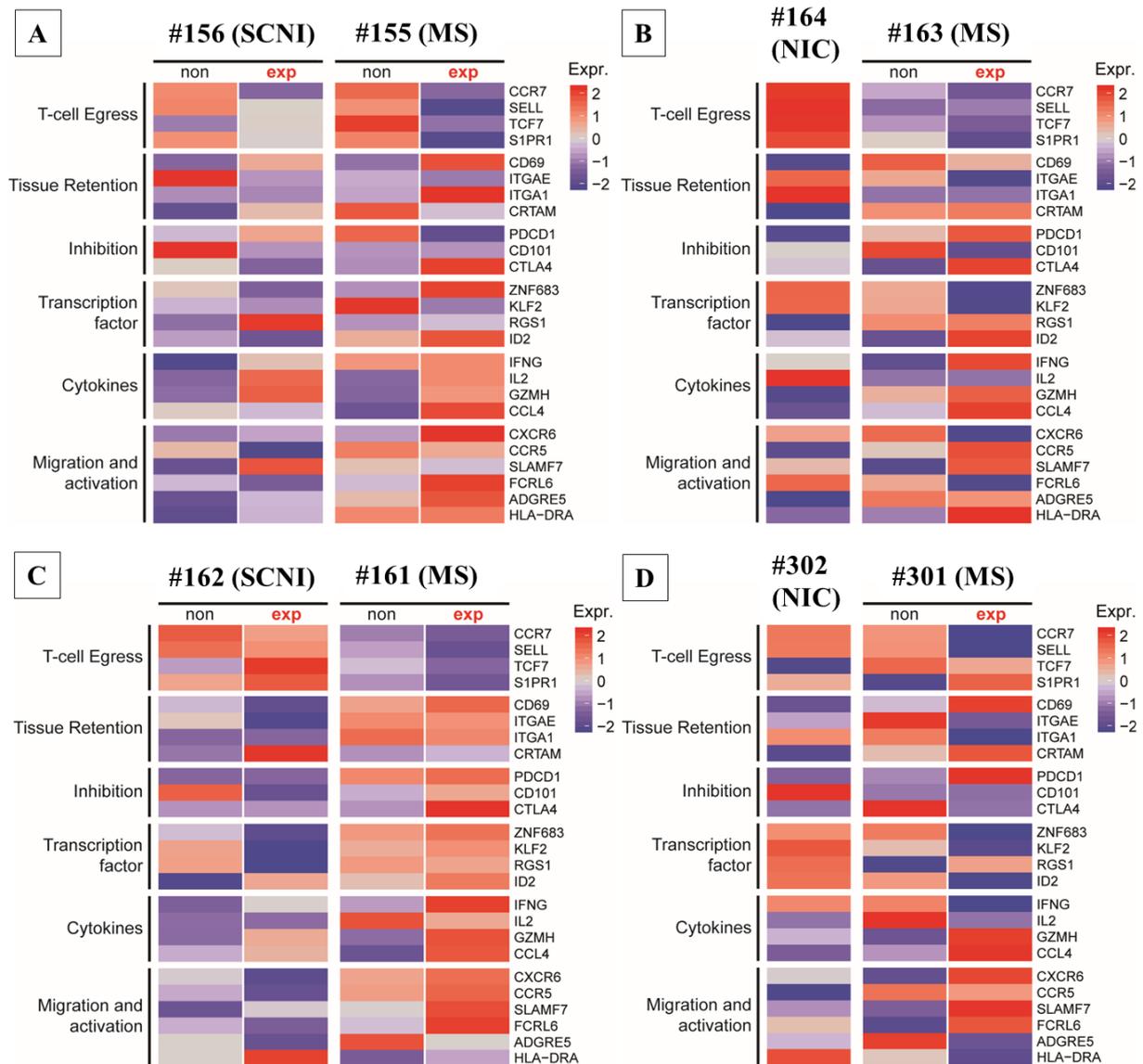
#### 4.2.4 Expanded CD8<sup>+</sup> T cells of individual twin pairs

Finally, the unique cohort of monozygotic twins discordant for MS disease was used to compare mean expression profiles of expanded CD8<sup>+</sup> T cells by heat map analysis in genetically identical individuals (Figure 23A-D). Each twin pair comprises a diseased twin (Figure 23 panel A: #155, panel B: #163, panel C: #161, panel D: #301) with established MS and a clinically healthy co-twin. All MS twins received distinct MS therapies (Table 9), whereas all co-twins were untreated. The two co-twins #156 (Figure 23A) and #162 (Figure 23C) fulfilled criteria for SCNI while the two co-twins #164 (Figure 23B) and #302 (Figure 23D) were found to be clinically and subclinically healthy NIC. Average data from all subject groups were differentially analyzed solely for expression patterns of CD8<sup>+</sup> T cells of distinct twin pairs. Of note, #156 developed MRI evidence for radiologically isolated syndrome (RIS) years after CSF sampling.

The individual comparison of expression patterns of non-expanded CD8<sup>+</sup> T cells and expanded CD8<sup>+</sup> T cells from the “healthy” co-twins and the diseased MS twin show overall similar average patterns and tendencies of the previously defined T<sub>RM</sub> signatures. The gradual downregulation of T<sub>RM</sub> key markers S1PR1 and upregulation of CD69 from non-expanded cells of the “healthy” co-twin to expanded CD8<sup>+</sup> cells of MS subjects is concise throughout all individuals and most pronounced in expanded CD8<sup>+</sup> T cells from MS patients. High expression levels of proinflammatory cytokines and activation markers can also be found and emphasize this cytotoxic and activated T<sub>RM</sub> phenotype. Moreover, it is remarkable that these cytotoxic, activated profiles are consistent within all twin pairs even though they received different MS medications. The only patient that differs from the patterns seen in all other individuals is MS twin #301 who was treated with teriflunomide, a disease-modifying drug that reduces the proliferation of activated T and B cells in patients.<sup>[188]</sup>

In short, the data reveal a distinct activated T<sub>RM</sub>-like phenotype particularly striking in expanded CD8<sup>+</sup> T cells of MS patients but already noticeable in SCNI patients and even a concise feature of disease-discordant twin pairs with identical genetically backgrounds. The upregulation of CXCL16 genes in DC and monocytes of all study groups was also of note, especially as slightly increased numbers of those cell types were found in the tSNE projections of cellular CSF landscapes that were found in MS subjects. Further, expanded CD4<sup>+</sup> cells show an activated and proinflammatory expression profile compared to non-expanded cells which is more pronounced in MS subjects than in SCNI or healthy controls. Finally, the presence of clonally expanded PB that were already found in SCNI individuals was unexpected.

Overall, these findings emphasize the involvement of all cells of the adaptive immune system and even of cells of the innate immune system in MS pathogenesis. In particular, the importance of clonally expanded T cells with notable prevalence of the CD8<sup>+</sup> T cell population even in very early pre-clinical phases of MS could be demonstrated and seemed to be important in pathogenic processes contributing to clinically definite MS.



**Figure 23: Heatmaps showing gene expression levels of selected T cell markers for expanded CD8<sup>+</sup> T cell clones of all twin-pairs. (A-D)** CD8<sup>+</sup> T cells were identified by extracellular staining information obtained by flow cytometry and the presence of at least one TCR chain from the transcriptome data. 25 selected T cell marker genes are represented in this heatmap for non-expanded and expanded CD8<sup>+</sup> T cell clones in the homozygotic twin pairs (**A**: #156/#155, **B**: #164/#163, **C**: #162/#161, **D**: #302/#301) Clonal expansion was defined when CDR3 identity was detected in more than one single CD8<sup>+</sup> T cell. The heatmap was created with normalized expression levels of the NGS data from the selected cells and mean expression levels coloured according to their z-scores ranging from -2.5 (blue) to +2.5 (red). (A-D) taken and modified from BELTRÁN *et al.*<sup>[176]</sup>

---

#### **4.2.5 Transcriptome profiling of patients with CIS diagnosis**

To fill the gap and understand immune processes that occur in patients between prodromal and established MS, CSF samples from untreated CIS patients with first clinical symptoms were analysed. In prodromal phases of the disease SCNI or RIS subjects are solely identified by chance - or in this case by purposefully looking for genetically predisposed subjects with maximal familial risks. In contrast, CIS patients are diagnosed at the very early onset MS with first clinical symptoms and thus in highly active, inflammatory phases of the disease. Therefore, transcriptome profiling of three patients with CIS diagnosis (#4626: 381 SC, #6014: 475 SC, #6137: 95 SC) and an additional control patient with lymphoma diagnosis (SCP-10: 95 SC) was performed (detailed clinical information in Table 9, section 3.1.10). Bioinformatics analysis by EDUARDO BELTRÁN is still ongoing, but preliminary data (not shown here) suggest a particular CIS specific phenotype.

### 4.3 Characterization of expanded lymphocyte clones

The local expansion of lymphocyte clones with the same receptor specificity is a commonly accepted indicator for an antigen-driven process and thus presumable involvement in disease pathology. To shed more light on and evaluate the potential of the CSF resident expanded T and B cell clones found in the prior chapter (Figure 20), the receptors of individual lymphocyte clones were further investigated and characterized. Therefore, full length sequences of  $\alpha:\beta$  TCR and H:L BCR chains from clonally expanded CD8<sup>+</sup> T cells and plasma cells were first identified from WT enriched cDNA of SC by target specific PCR amplification and conventional Sanger sequencing. Additionally, sequences were confirmed with the NGS data by EDUARDO BELTRÁN.

#### 4.3.1 Expanded plasma cells

Earlier studies on target antigens of expanded plasma cells that were isolated from OCBs of MS patients have already suggested their functional role as bystander cells that might play a role in debris removal of the inflammatory CNS processes during the disease.<sup>[133, 149]</sup>

However, with the newly established transcriptome approach expanded plasma cells were already detected in the CSF of very early patients with MS and subjects with SCNI (Figure 19). The function of those cells in disease pathogenesis was not determined yet and addressed in the following.

Expanded plasma cells were detected in four out of five patients (early MS: #301; CIS: #4526, #4626, #6014 and SCNI: #162; Table 9) with present OCBs. Full length CDR3 sequences of H:L BCR chains were identified by PCR amplification and Sanger sequencing as described in section 3.5.4 and sequences listed in Table 12. In all patients except from #6014, a dominant expansion of at least one patient-specific IgG plasma cell clone was detected.

Table 12: Matching H:L BCR chains from CSF resident expanded plasma cell clones.

#301(early RRMS)		CDR3 amino acid sequence		V	J	D
#1: 3/14 (21%)	3x	IgG	CARASRGCARTTCSVTLYYYGLDVW	HV1-69*12	HJ6*02	HD2-2*02
		κ	CQQYNIYPLTF	KV1-16*02	KJ4*01	
#2: 4/14 (29%)	4x	IgG	CARVPFDPIVWFGDPLGYCGMDVW	HV4-4*02	HJ6*02	HD3-10*01
	3x	λ	CQAWDSSTAVF	LV3-1*01	LJ2*01 or LJ3*01	
#162 (SCNI)		CDR3 amino acid sequence		V	J	D
5/14 (36%)	3x	IgG	CARHKWAAADSW	HV3-33	HJ4	HD6-13
	5x	κ	CMQATQFPITF	KV2-24	KJ3	
#4526 (CIS)		CDR3 amino acid sequence		V	J	D
#1: 2/10 (20%)	2x	IgG	CARTLYYYDDTGYTPSHWYFDIW	HV4-59*01	HJ2*01	HD3-22*01
		κ	CQQSYRTPWTF	KV1-39*01 or KV1D-39*01	KJ1*01	
#2: 3/10 (30%)	3x	IgG	CARGHYDLWSGYSDAFDIW	HV4-38-2*01	HJ3*02	HD3-3*01
	2x	κ	CQQLKSYSSS	KV1-9*01	KJ4*01 or IGKJ4*02	
#4626 (CIS)		CDR3 amino acid sequence		V	J	D
2/5 (25%)	2x	IgG	CTTDRMVVVAGLNDYW	HV3-15*01	HJ1*01	HD2-21*01, HD2-21*01
		κ	CMQGTHWPSITF	KV2-30*01, KV2D-30*01	KJ5*01	
#6014 (CIS)*						

\*#6014: 4 plasma cells identified, but no expansions detectable

Further, a closer look to the whole variable BCR regions of both chains reveal somatic mutations in germline encoded frame regions of heavy and light chains from expanded clones (Table 13). Both, the presence of clonal expansion and SHM suggest antigen driven processes and thus putative disease relevancy.

Table 13: Amino acid sequences of H:L BCR variable chains from expanded plasmablast clones that were identified in the CSF of patients.

Patients	FR1	CDR1	FR2	CDR2	FR3	CDR3
<b>#162</b>						
IGHV3-33*01	QVQ <b>V</b> VESGG.GVV QPG <b>K</b> SLRLSCT <b>A</b> S	GFTF.... <b>S</b> NYG	MHWVRQAPG KGLEWV <b>A</b> T	IWYD..G <b>S</b> K <b>T</b>	YYADSVK.GRFTISRDN <b>S</b> <b>R</b> K <b>T</b> <b>V</b> YLQMNSLR <b>V</b> GDTAVY <b>H</b> C	ARHKWAAADSWGQGT LTVSS
IGKV2-24*01	DIVMTQ <b>S</b> PLSSPVT LGQPASISCRSS	QSL <b>A</b> H <b>S</b> .DGNTY	LSW <b>I</b> QQRPGQ PPR <b>P</b> LIY	KI.....S	NRFSGVP.DRFSGSG..AGTDFT LKISRVEAEDVGVYYC	MQATQFPITSA
<b>#301-1</b>						
IGHV1-69*12	QVQLVQSGA.EVKK PGSSVKV <b>S</b> CK <b>T</b> F	GGTF.... <b>S</b> R <b>N</b> A	<b>I</b> NWVRQAPG <b>H</b> GLEWMGG	<b>L</b> IP <b>S</b> .. <b>V</b> G <b>L</b> A	NYAQKFQ.GRVTITAD <b>A</b> STST <b>A</b> NMELSSLR <b>S</b> DDTAVYYC	ARCSVTLYYYGLDVWG QGTTVTVSS
IGKV1-16*02	DIQMTQSPSSLSAS VGDRVITICRAS	<b>Q</b> D <b>I</b> ..... <b>G</b> F <b>H</b>	LAWFQQKPG <b>E</b> APKSLIY	AA..... <b>T</b>	<b>T</b> LQSGVP. <b>F</b> KFSGSG..SGTDFT LTISSLQPEDFATYYC	QQYNS <b>I</b> PLTFGGXPXXD
<b>#301-2</b>						
IGHV4-4*02	QVQLQESGP. <b>R</b> L <b>V</b> K PSGTL <b>S</b> LTC <b>A</b> VS	<b>G</b> D <b>L</b> IS... <b>S</b> R <b>N</b> W	WSWVRQPPG KGLEWIGE	<b>I</b> F <b>H</b> A...G <b>T</b> I	<b>K</b> YNPSL <b>R</b> .SRVTISVDKSKNQF <b>S</b> L <b>R</b> L <b>N</b> SVTAADTAVYYC	ARVPFDPIVWFGDPLGY CGMDVWGP <b>G</b> TTVTVSS
IGLV3-1*01	SYELTQ <b>P</b> L <b>S</b> .VSVSP GQTASITCS <b>G</b> H	KLG..... <b>D</b> K <b>H</b>	<b>A</b> FWYQQKPG QSPVLVIY	QD.....S	KRPSGIP.ERFSGSN..SGNTATL TISGTQAMDEADY <b>F</b> C	QAWDSSTAVFGGGTKLT VL

#4526-1						
IGHV4-59*01	QVQLQESGP.GL <b>GR</b> PSETLSLTCTVS	GGSI.... <b>NNYY</b>	WSW <b>LR</b> QPPG K <b>A</b> LEWIGY	<b>SYYR...</b> <b>GNT</b>	NYNPS <b>L</b> E.SRVTISVDTS <b>R</b> NQF SLKL <b>T</b> SLTAADTA <b>L</b> YYC	ARTLYYYDDTGYTPSHW YFDIWGRGTLVTVSS
IGKV1-39*01	DIQMTQSPSSLSAS VGDE <b>ESHHL</b> PGS	QSI..... <b>TTH</b>	<b>L</b> SWYQ <b>H</b> KPG KAPKLLIY	AA.....S	<b>NLQ</b> TGVP.SRFSGSG..SGTDFT LT <b>T</b> SLQPEDFATY <b>H</b> C	QQSY <b>R</b> TPWTFGRGTRVE IK
#4526-2						
IGHV4-38-2*01	QVQLQESGP.GLVK PSETLSLT <b>C</b> DVS	<b>D</b> NSIS...SGYY	WGWIRQPPG KGLEWIGS	<b>I</b> HHS...G ST	YYNPSLK.SRVT <b>L</b> SVDTSK <b>N</b> L FSLKL <b>T</b> SVTAADTAVY <b>F</b> C	ARGHYDLWSGYSYDAF DIWGQGMVTVSQ
IGKV1-9*01	DIQLTQSPSFLSASV GDRV <b>X</b> ITCRAS	<b>Q</b> DI..... <b>S</b> NF	<b>L</b> VWYQQKPG KAP <b>N</b> LLIY	AA.....S	<b>S</b> LQSGVP.SRFSGSG..SGTEFTL TISSLQPEDFATYYC	QQL <b>K</b> SYPSSSPSGTKVEI K

\*SHM into the germline encoded frame regions of the antigen-specific receptors were highlighted with bold, red letters

### 4.3.2 Expanded CD8<sup>+</sup> TCR TRAV17\*01/TRBV4-1\*01

Dominant clonal expansions of CD8<sup>+</sup> T cells were identified with scRNAseq in human CSF samples from MS patients and even in subjects with SCNI (Figure 20). Those cells even showed an activated T<sub>RM</sub> phenotype and might be a result from sustained activation with putative disease-relevant autoantigens (Figure 21). Thus, the identification of target antigens from those expanded CD8<sup>+</sup> T cell clones is crucial to clarify their potential relevance in disease pathogenesis and was addressed in this work.

The expanded CD8<sup>+</sup> T cell clone TCR TRAV17\*01/TRBV4-1\*01 was chosen for further characterization. Interestingly, the CDR3 $\alpha\beta$  sequences of TCR TRAV17\*01/TRBV4-1\*01 were found fifteen times in the CSF of MS patient #161 and thus represented the most dominantly expanded CD8<sup>+</sup> T cell clone of this patient. Strikingly, the same CDR3 sequence was also found four times in the CSF of the corresponding clinically healthy co-twin #162. Sequences were first identified by T cell specific PCR with C $\alpha$ -rev IN/ C $\beta$ -rev IN and TSO fwd primers and conventional Sanger sequencing from WT enriched cDNA of SC. Further, those findings were validated with the NGS data by EDUARDO BELTRÁN.

To exclude cross-contamination of samples during NGS preparation, the detected CDR3 sequences of TCR TRAV17\*01/TRBV4-1\*01 from both patients were analysed in more detail and revealed different nucleotide usage for the same amino acid sequence (Table 14). As the nucleotide usage is specific for each individual and was consistent for the particular sequences found in different wells of each patient, cross-contamination could be excluded. Further phenotypic characterization of this clone by NGS data analysis (EDUARDO BELTRÁN) even revealed characteristics of an activated T<sub>RM</sub> phenotype (data not shown).

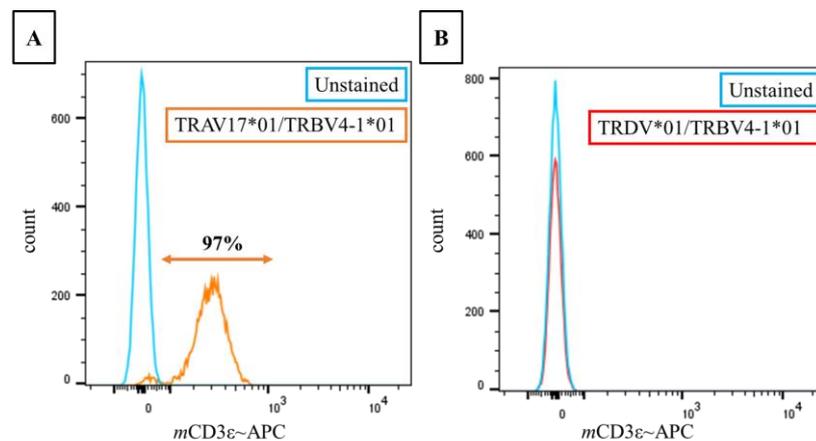
In addition to the expanded TRAV17\*01  $\alpha$ -chain, in 9 out of 15 cells of patient #161 the additional alpha chain TRDV\*01 was detected (Table 14). Expression of dual alpha chains is known, since no allelic exclusion exists for TCR  $\alpha$  chain gene rearrangement.<sup>[189]</sup> Whether both  $\alpha$  chains are also expressed in functional TCR complexes on the surface of cells was assessed by *in vitro* reconstitution of both  $\alpha$  chains together with the  $\beta$  chain in 58 hybridoma cells.

Table 14: CDR3 sequences of  $\alpha$ : $\beta$  TCR variable chains from the most dominating expanded CD8<sup>+</sup> T cell clone sequences that were identified in the CSF of twin pair #161/ #162.

patient		CDR 3 nucleotide sequence	CDR3 amino acid sequence	V	J	D
#161	TRAV17*01 (15x)	TGT GCT ACG <b>GAC TAC</b> AAC CAG GGA GGA AAG CTT ATC TTC	CAT <b>D Y</b> NQGGKLIF	AV17	J23	
	TRDV*01 (9x)	TGT GCT CTT GGG GAA CTG TTT TTC TAC AGC AGT GCT TCC AAG ATA ATC TTT	CALGELFFYSSASKIIF	DV1	J3	
	TRBV4-1*01 (15x)	TGC GCC AGC AGC CAA GAT CCG <b>GGG</b> GCC TGG GAG ACC CAG TAC TTC	CASSQDP G AWETQYF	BV4-1	J2-5	D1
#162	TRAV17*01 (4x)	TGT GCT ACG <b>GAT TAT</b> AAC CAG GGA GGA AAG CTT ATC TTC	CAT <b>D Y</b> NQGGKLIF	AV17	J23	
	TRBV4-1*01 (4x)	TGC GCC AGC AGC CAA GAT CCG <b>GGA</b> GCC TGG GAG ACC CAG TAC TTC	CASSQDP G AWETQYF	BV4-1	J2-5	D2

\*Red, bold letters indicate relevant sequence differences of T cell clones TCR TRAV17\*01/TRBV4-1\*01 found in #161 and #162

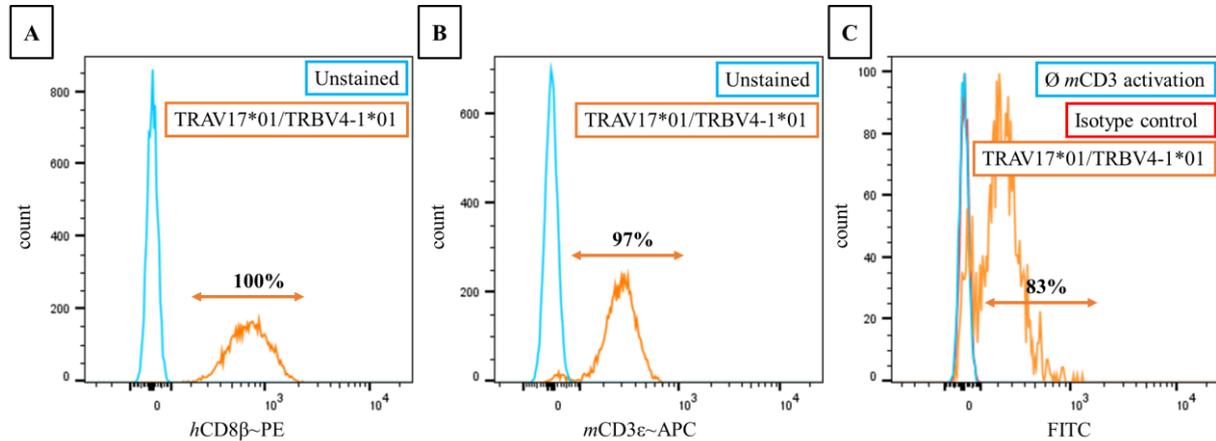
Therefore, TCR chain sequences for TRAV17\*01 (Trappist  $\alpha$  fwd/ C $\alpha$ -IN rev), TRDV\*01 (Trappist -161 all for/ C $\alpha$ -IN rev) and TRBV4-1\*01 (Trappist  $\beta$  fwd/ C $\beta$ -IN rev) were amplified from WT cDNA and cloned into pRSV5 vectors for antigen search as described in section 3.3.5. TCR TRAV17\*01/TRBV4-1\*01 and TRDV\*01/TRBV4-1\*01 plasmids were then stably transfected into 58 hybridoma cells that already expressed *hCD8* and the reporter gene NFAT-sGFP. Functional TCRs only reach cell surfaces in a complex with the co-receptor molecule CD3, thus productive TCR expression was confirmed by extracellular staining with *mCD3 $\epsilon$* -APC conjugated fluorescence antibodies. Flow cytometry analysis of sixteen hybridoma clones from two differently transfected hybridoma cell lines of each TCR combination (TRAV17\*01/TRBV4-1\*01 and TRDV\*01/TRBV4-1\*01) was performed and *mCD3 $\epsilon$* -APC fluorescence compared to an unstained control (Figure 24, exemplarily shown for one clone of each combination). Compared to TCR TRAV17\*01/TRBV4-1\*01 (97% APC<sup>+</sup>, Figure 24A), no surface expression of *mCD3 $\epsilon$*  could be detected for any of the analysed clones from TRDV\*01/TRBV4-1\*01 (Figure 24B). This revealed that the additionally detected second alpha chain TRDV\*01 does not reach the cell surface as a functional TCR complex and was thus not relevant for the following experiments.



**Figure 24: Flow cytometry analysis of reconstituted TCR TRAV17\*01/TRBV4-1\*01 and TCR TRDV\*01/TRBV4-1\*01 hybridoma cell lines.** Surface expression of both expanded T cell complexes was monitored by staining with *mCD3 $\epsilon$* -APC conjugated antibodies. In total, 32 grown clones from two different transfections of 58 $\alpha$ - $\beta$ - hybridoma cells with different cell line histories (Frankfurt and Hamburg origin) were analysed individually for TCR TRAV17\*01/TRBV4-1\*01 (A, orange) and TCR TRDV\*01/TRBV4-1\*01 (B, red) expression. A representative plot is indicating the fraction of cells of the respective alpha chains. Percentages show the amount of APC<sup>+</sup> cells of each condition compared to the negative control (unstained cells, blue).

Further, CD8 surface expression from all reconstituted cells of TCR TRAV17\*01/TRBV4-1\*01 was assessed by surface staining with *hCD8 $\beta$* -PE conjugated antibodies and only cells with high CD8 expression (100%; Figure 25A) selected for further experiments. Lastly, T cell activation through CD3 crosslinking was induced by incubation with coated *mCD3 $\epsilon$*  antibodies for 16 hours and monitoring of GFP expression as indicator for successful NFAT signalling (97%; Figure 25B). Induced expression of the reporter gene sGFP as a readout for T cell

activation was compared to a non-stimulated control by measuring the GFP fluorescence in the FITC channel of a flow cytometer (83%; Figure 25C). Flow cytometry analysis of the characterized T cell hybridoma clone that was selected for antigen experiments are shown in Figure 25.



**Figure 25: Flow cytometry analysis of the selected TCR TRAV17\*01/TRBV4-1\*01 clone that was used for following antigen search experiments.** Surface expression of transfected plasmids were monitored after staining with fluorescent dye-conjugated antibodies (A: *hCD8β*~PE, B: *mCD3ε*~APC). Further, T cell activation was induced by incubation with *mCD3* antibodies (orange) or an isotype control (red) for 16 hours and the resulting NFAT-sGFP expression evaluated in the FITC channel of the flow cytometer (C). Unstained cells (A+B) and cells without *mCD3* activation (C) were used as negative controls (blue) and percentages are given to indicate the amount of positive cells in each channel compared to the negative control.

### 4.3.3 Investigating HLA preferences of TCR TRAV17\*01/TRBV4-1\*01

T cells only recognize target antigens that are presented to them as peptide:MHC complexes on the cell surface of APCs. Hence, knowing the HLA restriction of the expanded T cell clone, would significantly simplify antigen search experiments since TCR activation have to be screened for various PECPLs in combination with all different HLA molecules. Therefore, prior to antigen search experiments, preliminary experiments were conducted to investigate HLA preferences of TCR TRAV17\*01/TRBV4-1\*01. Remarkably, twin pair #161/#162, in which the TCR TRAV17\*01/TRBV4-1\*01 was detected, is homozygous for all three loci of MHC class I and this reducing the variability to a minimum. The HLA homozygosity (A24:02, B07:02 and C07:02) was determined by classical genotyping of PROF SPANNAGL (Laboratory for Immunogenetics, LMU Munich) as well as validated with the NGS WT data by EDUARDO BELTRÁN.

First, it was tested whether the TCR  $\beta$  chain sequence of TRBV4-1\*01 could be identified in PBMCs of fifteen different partially and fully HLA matching patients with MS or RIS diagnosis. Compared to CSF resident T cells, the TCR repertoire of human PBMCs is much higher and might reveal TCRs with CDR3 sequence similarities. Thus, PBMCs from fifteen different patients (listed in Table 10) were prepared from whole blood and CD8<sup>+</sup> T cells (naïve

and CCR7<sup>+</sup> antigen-experienced) isolated by bulk cell sorting on a FACS BD AriaIII. RNA was extracted by conventional Trizol isolation, transcribed into cDNA with TCR specific RT primers and the TRBV4-1\*01 CDR3 sequence (CASSQDPGAWETQYF) amplified by PCR with target-specific primers (Trappist-161-TRB for/ Rev-spec3). Since the receptor diversity for the TCR beta loci is much higher than for alpha, chances for the identification of sequence similarities are higher for the TCR  $\beta$  chain. Therefore first, PCR amplification with beta specific reverse primers was performed that were designed to specifically bind in the TRBV4-1\*01 CDR3 region. To maximize variability, the Rev-spec3 (5'-TGNGTYTCCCANGCNCNGG-3') primer was designed with the usage of wobble nucleotides that covers all possibilities for nucleotide combinations of the TRBV4-1\*01 CDR3 sequence. Additionally, for eight of the most relevant patients (including #161, #162, #120 and #130) the TCR alpha chain TRAV17\*01 CDR3 sequence (CATDYNQGGKLIF) was amplified by PCR with target-specific primers (V $\alpha$ 3 for IN/ TRA spec Rev), but revealed no results.

In conclusion, sequences with homologies to the one of the TCR TRAV17\*01/TRBV4-1\*01 CDR3 sequences could only be detected for the TCR  $\beta$  chain in PBMCs from patient #130 (C\*07) (Table 15). Although the CDR3 sequence of the  $\beta$ -chain of patient #130 matched with the expanded sequences found in twin-pair #161/#162, the  $\alpha$ -chain could not be detected. Since  $\alpha$ : $\beta$  TCR pairing is fundamental for antigen recognition, it seems like the T cell clone TRAV17\*01/TRBV4-1\*01 is private to the twin pair #161/ #162.

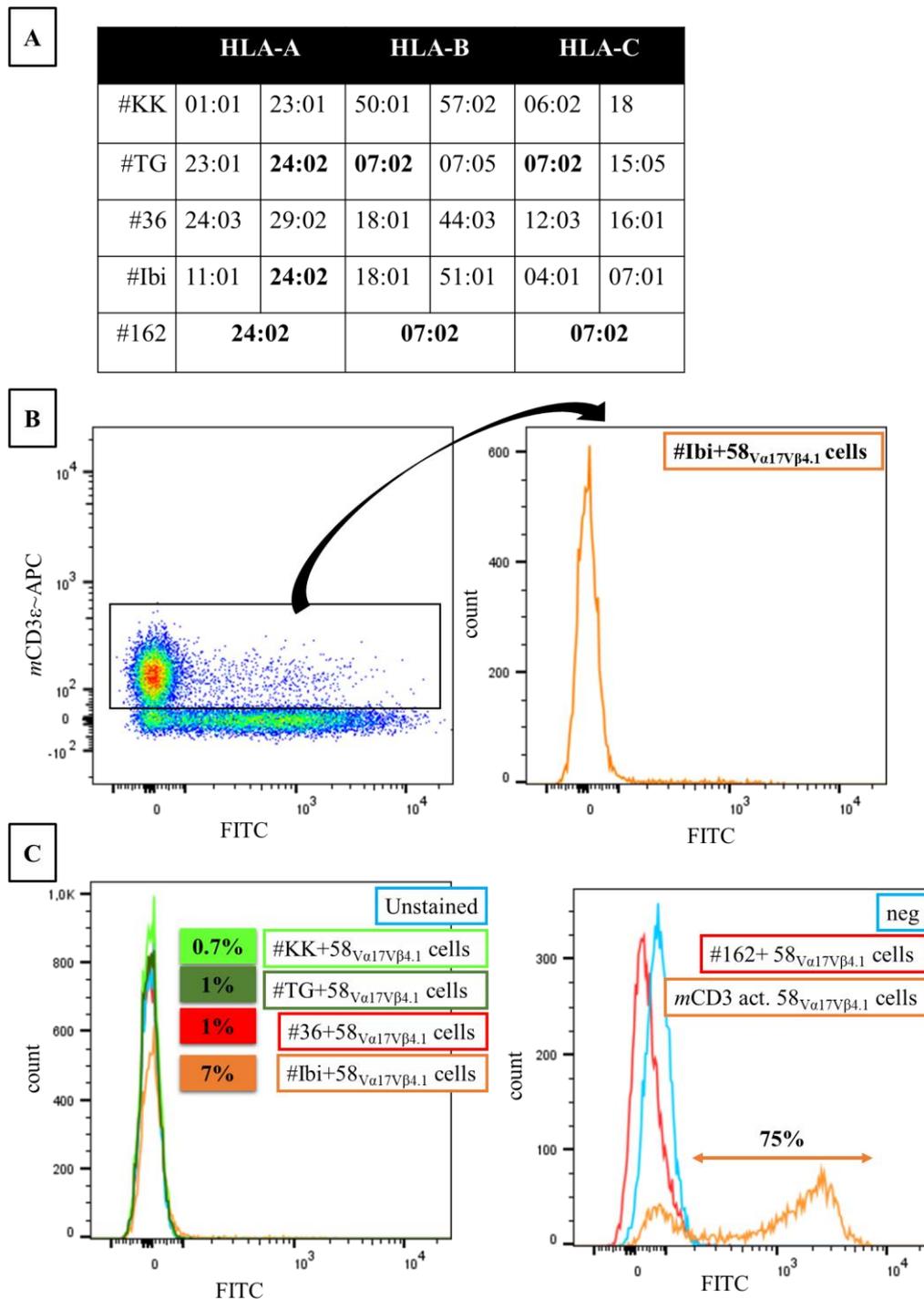
Table 15: CDR3 sequence comparison of TRBV4-1\*01  $\beta$  chain clones amplified from PBMCs of HLA-matching patients.

Patient		CDR3 nucleotide sequence	CDR3 amino acid sequence	V	J	D
#161	$\beta$	TGC GCC AGC AGC CAA GAT CCG <b>GGG</b> GCC TGG GAG ACC CAG TAC TTC	CASSQDP <b>G</b> AWETQYF	<b>BV4-1</b>	<b>J2-5</b>	<b>D1</b>
#162	$\beta$	TGC GCC AGC AGC CAA GAT CCG <u>GGA</u> <u>GCC</u> TGG GAG ACC CAG TAC TTC	<u>CASSQDP</u> <u>G</u> AWETQYF	<b>BV4-1</b>	<b>J2-5</b>	<b>D2</b>
#130 (C07)	$\beta$	TGC GCC AGC AGC CAA GAT CCG <b>GGC</b> GCN <u>TGG</u> GAA <u>ACC</u> CAA	<u>CASSQDP</u> <u>G</u> AWETQ	<b>BV4-1</b>	J2-4	-

\*Bold letters indicate relevant sequence differences and underlined sequences highlight similarities, when compared to the TRBV4-1\*01 CDR3 sequence of #161

Anyhow, to further investigate the HLA restriction of TCR TRAV17\*01/TRBV4-1\*01, co-culture experiments with EBV cells were performed. Given that EBV infection has a strong linkage with MS, it was tested whether TCR TRAV17\*01/TRBV4-1\*01 might recognize endogenous EBV antigens on the cell surface of five different EBV cell lines that express a heterogeneous set of HLA molecules (Figure 26A). The EBV cell lines were provided by the group of PD J. MAUTNER (Helmholtz Centre Munich) and produced by immortalization of patient's PBMCs. Since EBV lines were produced with GFP vector selection, *mCD3ε* staining was performed after the co-culture experiment to distinguish activated green T cells from green fluorescing EBV cells (Figure 26B). For the TCR activation assay, 50.000 EBV cells were co-cultured with 100.000 58V $\alpha$ 17V $\beta$ 4.1 cells in 96-well plates. After 16 hours of incubation, cells were stained and analysed by flow cytometry. T cell activation ranged from 0.7 % for #KK EBV cells with no matching HLA molecules to 7 % for #Ibi EBV cells that share the A24:02 allele (Figure 26C, left panel). Even in co-culture experiments with EBV cells that were produced by K. DORNMAIR from PBMCs of patient #162 and therefore express all patient-specific HLAs, no T cell activation could be detected (Figure 26C, right panel). Since no statistically significant activation could be detected, TCR TRAV17\*01/TRBV4-1\*01 seem not to recognize any peptides, neither of endogenous sources, nor from EBV.

Further, no clear HLA preference for TCR TRAV17\*01/TRBV4-1\*01 could be identified with this preliminary experiments. So antigen search experiments were performed by screening T cell activation of all three HLA (A24:02, B07:02 and C07:02) alleles in combination with different PECPLs.



**Figure 26: HLA restriction experiments of TCR TRAV17\*01/TRBV4-1\*01 by co-culture experiments with different EBV cell lines.** (A) HLA A, B and C alleles of the five EBV cell lines #KK, #TG, #36, #Ibi and #162 that were used for co-culture experiments with TCR TRAV17\*01/TRBV4-1\*01. T cell activation by potential expression of endogenous EBV antigens was validated 16 hours after co-culturing of EBV and 58V $\alpha$ 17V $\beta$ 4.1 cells by flow cytometry. Bold letters highlight HLA compliance with patient-specific HLAs of #161. (B) Exemplary flow cytometry gating strategy of #Ibi EBV co-culture with 58V $\alpha$ 17V $\beta$ 4.1 cells that was used to distinguish GFP expressing green fluorescing Z- EBV cells and FITC positive, activated T hybridoma cells by surface staining with mCD3 $\epsilon$ -APC antibody. (C) Flow cytometry analysis of the T cell activation assays. 58V $\alpha$ 17V $\beta$ 4.1 cells were co-cultured with EBV cells for 16 hours, stained with mCD3 $\epsilon$ -APC conjugated antibodies and T cell activation analysed by flow cytometry. (Left panel) T cell activation after co-culture with #KK (light green), #TG (dark green), #36 (red), #Ibi (orange) was compared to an unstained control (blue). Percentages are given to indicate the amount of FITC-positive hybridoma cells compared to the negative control. (Right panel) T cell activation after co-culture with EBV cells of patient #162 (red) was compared to mCD3 activated 58V $\alpha$ 17V $\beta$ 4.1 cells (orange) and cells without mCD3 activation (negative control, blue). Percentages are given to indicate the amount of FITC-positive hybridoma cells compared to the negative control.

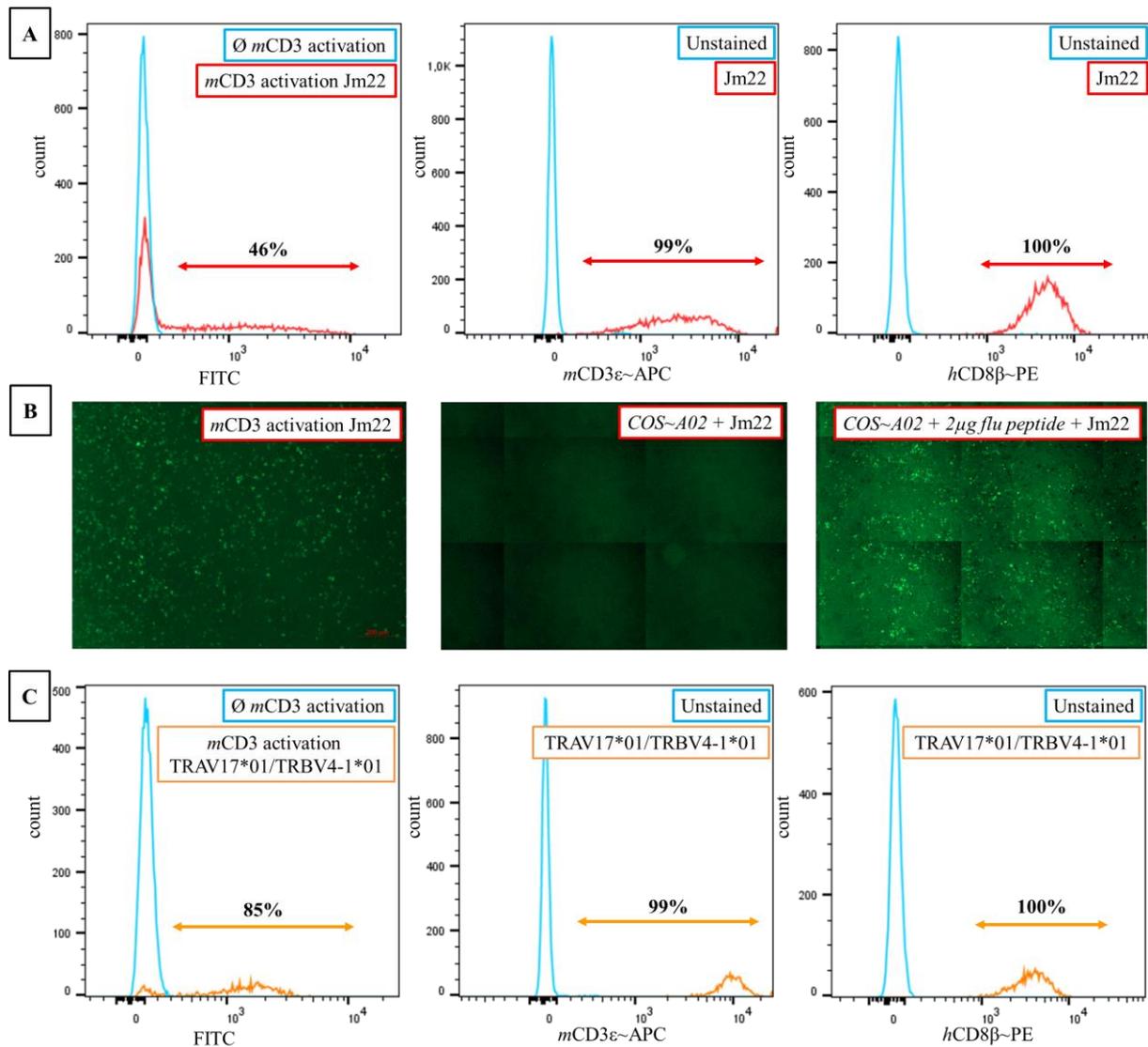
#### 4.4 T cell antigen search

The identification of target antigens is crucial to unravel the functional role of those expanded, activated lymphocytes in autoimmune diseases and might reveal new insights into disease pathology or might be used for the development of potential new therapeutic targets. To address these remaining questions, an unbiased, highly sensitive assay that enables the reliable identification of target antigens of CD8<sup>+</sup> T cells was successfully established by KATHERINA SIEWERT (Figure 10).<sup>[133, 167]</sup>

First, the results of KATHERINA SIEWERT with JM22 hybridoma cells and flu antigen were reproduced in a preliminary experiment to ensure reliable methodical handling. Then *in vitro* reconstituted 58V $\alpha$ 17V $\beta$ 4.1 cells were used to search for candidate antigens of TCR TRAV17\*01/TRBV4-1\*01 (4.4.1). Lastly, the T cell antigen search protocol was optimized by the implementation of a HLA~BFP construct that enabled the construction of stably expressing HLA~COS cell lines.

##### 4.4.1 Antigen search experiments of TCR TRAV17\*01/TRBV4-1\*01

For practicing experimental cell handling of the antigen search protocol, the T cell activation assay was tested with JM22 hybridoma cells and flu antigen.<sup>[167]</sup> JM22 T cell activation with flu antigen was confirmed and thus KATHERINA SIEWERT data reliably reproduced (Figure 27B). Further, cell characteristics such as GFP activation with *m*CD3~antibodies as well as CD8 and CD3 surface expression of the JM22 (Figure 27A) hybridoma clone was compared to the reconstituted TCR TRAV17\*01/TRBV4-1\*01 (Figure 27C) and revealed no significant differences.



**Figure 27: Methodical validation of T cell antigen search with JM22.** (A+B) To exclude methodical mistakes, the results of K SIEWERT with JM22 TCR and flu antigen were reproduced here. (A) Flow cytometry analysis of JM22 clone that was used by K SIEWERT for establishing unbiased T cell antigen search (red). T cell activation was induced by 16 h incubation with *mCD3* antibodies and the resulting NFAT-sGFP expression evaluated in the FITC channel of the flow cytometer (left panel). Surface expressions of transfected plasmids were monitored after staining with fluorescence dye-conjugated antibodies (central panel: *mCD3 $\epsilon$ -APC*, right panel: *hCD8 $\beta$ -PE*). Unstained JM22 cells and cells without *mCD3* activation were used as negative controls (blue) and percentages are given to indicate the amount of positive cells in each channel compared to the negative control. (B) Example sections of the co-culture experiments from JM22-58 cells were taken in the GFP channel of the fluorescence microscope. JM22 activation with pre-coated *mCD3* antibodies (left panel, 200  $\mu$ M scale) and JM22 co-culture experiment with untransfected COS cells (middle panel, 500  $\mu$ M scale) as controls are shown. JM22 co-culture with stable transfected COS~A02 cells that are presenting flu antigen (transfected with 2  $\mu$ g flu peptide 72 hours before analysis; right panel, 500  $\mu$ M scale). JM22 T cells were added 16 hours before analysis by fluorescence microscopy. (C) Flow cytometry analysis of used TCR TRAV17\*01/TRBV4-1\*01 clone for T cell antigen search for direct comparison with K SIEWERT JM22 clone. T cell activation was induced by incubation with coated *mCD3* antibodies (orange) for 16 hours and the resulting NFAT-sGFP expression evaluated in the FITC channel of the flow cytometer (left panel). Surface expressions of transfected plasmids were monitored after staining with fluorescence dye conjugated antibodies (central panel: *mCD3 $\epsilon$ -APC*, right panel: *hCD8 $\beta$ -PE*). Further, Unstained cells and cells without *mCD3* activation were used as negative controls (blue) and percentages are given to indicate the amount of positive cells in each channel compared to the negative control.

Thus, antigen search experiments with the reconstituted TCR TRAV17\*01/TRBV4-1\*01 were performed by screening 58<sub>V $\alpha$ 17V $\beta$ 4.1</sub> cell co-cultures with COS-7 cells that were transiently transfected with one of the three individual HLA alleles (A24:02, B07:02 and C07:02) in combination with seven different PECPL libraries (Table 16). Scan experiments of co-cultures

were performed in 3.5 cm cell culture dishes 24, 48 and 72 hours after FuGene transfection of COS-7 cells with 1  $\mu$ g HLA (A24:02, B07:02 or C07:02) and 1  $\mu$ g plasmid DNA of seven different PECPL libraries. Green, activated 58 $V_{\alpha}17V_{\beta}4.1$  cells were counted to determine the HLA restriction of TCR TRAV17\*01/TRBV4-1\*01.

Single activated 58 $V_{\alpha}17V_{\beta}4.1$  cells were identified with all possible HLA-PECPL combinations. In contrast, clusters of green fluorescing T cells were preferentially detected in HLA A24:02 and C07:02 transfected cells in combination with the 8L and 9L library. Since no clear HLA restriction was sticking out after HLA screening experiments, regular single cell isolation experiments to enrich putative candidate antigens were continued.

**Table 16: Determining the HLA restriction of TCR TRAV17\*01/TRBV4-1\*01 - Scan experiments**

Libraries	Surface screened	Detected clusters for each HLA					
		A24:02		B07:02		C07:02	
		Single	Double	Single	Double	Single	Double
8L	19.2 cm <sup>2</sup>	19	4	11	-	15	1
9L	19.2 cm <sup>2</sup>	14	1	8	-	11	-
10L	19.2 cm <sup>2</sup>	25	3	19	1	16	3
N24	19.2 cm <sup>2</sup>	15	1	14	-	8	-
Cw06	19.2 cm <sup>2</sup>	7	-	18	-	14	1
8L/9L (1:1)	19.2 cm <sup>2</sup>	28	-	24	-	19	-
N24/N27 (1:1)	19.2 cm <sup>2</sup>	28	1	23	-	21	-

Single cell isolation experiments were performed in 3.5 cm cell culture dishes 24, 48 and 72 hours after transfection of COS-7 cells with HLAs (A24:02, B07:02 or C07:02) and 8L/9L libraries. COS cells underneath green fluorescing clusters (singlets, doublets or triplets) of 58 $V_{\alpha}17V_{\beta}4.1$  cells were isolated with a micromanipulator. Most COS-7 cells were isolated in the HLA C07:02 transfected condition with the 8L/9L libraries, revealing a putative HLA C preference of TCR TRAV17\*01/TRBV4-1\*01 (Table 18, section 7.2.1). Altogether, plasmid DNA of 37 isolated PECPLs was enriched by PCR amplification with pcDNA-for1/ pcDNA-rev1 primers and pooled in equimolar concentrations to six different pools. The pools were re-cloned into pcDNA6 vectors and used for another round of single cell “re-isolation” to further enrich candidate antigens. Re-activations were mainly found in pools#3 and #5 that comprised solely original samples from HLA C of the first round of single cell isolation. The eleven original samples of pool#3 and pool#5 plus eleven re-isolated samples were then restriction cloned individually and tested for T cell reactivation in subpool experiments (Table 19, section 7.2.1). Co-culture assays of subpools were performed in 24-well cell culture plates 24, 48 and 72 hours after FuGene transfection of COS-7 cells with 0.5  $\mu$ g HLA~C07:02 and 0.5  $\mu$ g plasmid DNA of the re-cloned samples. The well with the highest number of activated, green fluorescing

T cells of each round was selected, divided into 30 different subpools and those again tested for T cell re-activation. In total, three rounds of subpool experiments were performed until T cell reactivation was not detectable anymore. Thus, no candidate antigens for TCR TRAV17\*01/TRBV4-1\*01 could so far be identified.

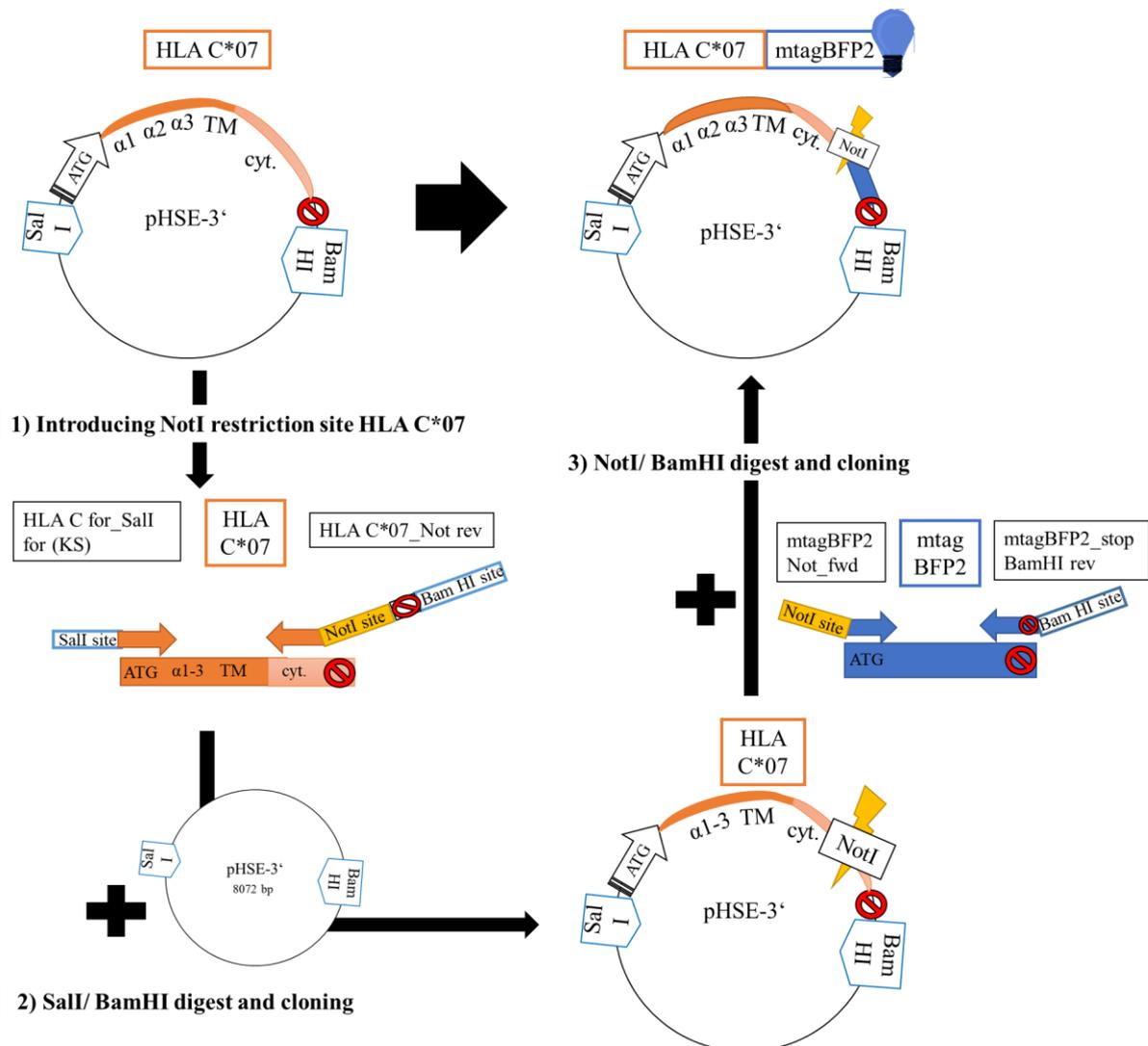
#### 4.4.2 Optimizing the T cell antigen search protocol

The density of expressed PECPL on the cell surfaces of APCs is critical for the chances of T cell activation and thus detection of candidate target antigens. A major limitation of the currently used antigen search method remained the monitoring of antigen presentation efficiencies of peptide:MHC surface complexes in the COS-7 cells. In contrast to the *in vitro* reconstituted T cells that are produced as stable transfectants and can be monitored straightforward by analytical flow cytometry, antibodies against human HLA-antigens are still limited.<sup>[170]</sup> Therefore, HLA and PECPL plasmids were so far only transiently transfected, which is inefficient as transfection efficiencies of each plasmid construct multiply. To produce stable transfected HLA COS cell lines – independent of the availability of allele-specific HLA antibodies – the blue fluorescent mtagBFP protein was introduced here as a fluorescence reporter.

In this project, mtagBFP2 was cloned into the intracellular domain of the existing HLA-C07:02~pHSE vector as outlined in Figure 28A. Since, the BFP fluorescence is not interfering with already existing fluorescent signals (FITC, Cy3 channels) of the antigen search protocol, it was used to monitor HLA expression as well as to establish stably transfected HLA~COS-7 cell lines.

For mtagBFP2 cloning, the rarely cutting restriction enzyme NotI was chosen. In a first step, the NotI site was introduced into the cytosolic domain of the existing HLA-C07:02~pHSE construct by restriction cloning. Then, NotI and BamHI restriction sites were added to the mtagBFP2 sequence by PCR amplification with overhang primers. Finally, NotI/BamHI restriction digestion and cloning were performed to produce the intended BFP-HLA-C07:02~pHSE vector. To easily apply the mtagBFP2 sequence to prospective HLA~pHSE constructs, cloning was designed using a strategy to exchange HLA sequences in only one restriction cloning step. Therefore, the respective HLA sequences were just amplified with reverse primers that contained an overhanging NotI restriction site. SalI/NotI restriction cloning of the BFP-HLA-C07:02~pHSE vector was used to replace HLA C07 with HLA sequences of A24:02:01:01 and B07:02:01:01. After all, stably transfected mtagBFP2-HLA-COS-7 cell lines

were produced for these three HLAs and might contribute to the successful identification of target antigens of TCR TRAV17\*01/TRBV4-1\*01 in prospective experiments.



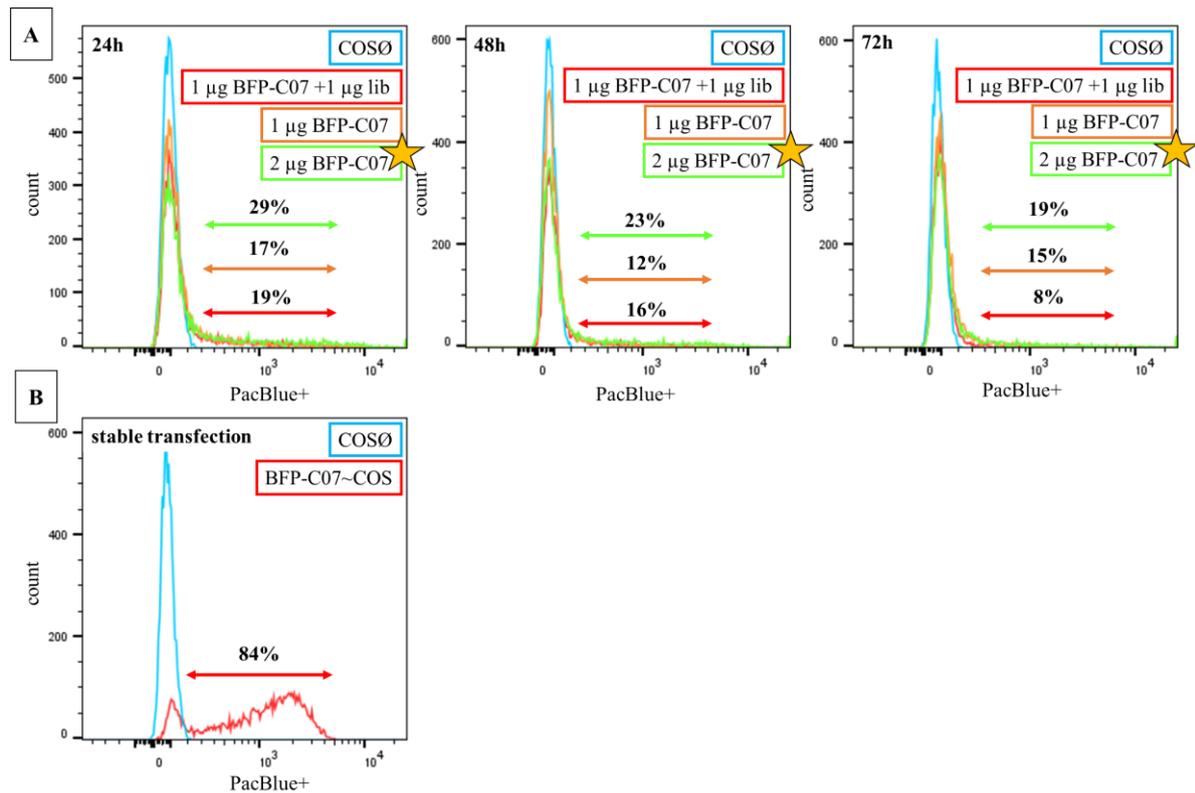
**Figure 28: Establishing stably transfected HLA-expressing APCs to optimize T cell antigen search.** Cloning strategy of mtagBFP2 into the HLA-C\*07:02\_pHSE-3' vector. 1) HLA-C\*07:02 was amplified with the primers HLA\_C\_for\_SalI\_for (KS) forward and HLA\_C\*07\_Not reverse that were designed to introduce a NotI restriction site into the cytosolic domain of the HLA protein. 2) The HLA C\*07 PCR product from step 1 was digested with the restriction enzymes SalI and BamHI and cloned into a SalI/ BamHI digested pHSE-3' vector. 3) mtagBFP2 was amplified with the designed primers mtagBFP2\_Not forward and mtagBFP2\_stop\_BamHI reverse to introduce a NotI and BamHI restriction site into the PCR product. The mtagBFP2 PCR product was digested with the restriction enzymes NotI and BamHI and cloned into the NotI/ BamHI digested \*new\* HLA\_C07\*02 pHSE vector with the NotI restriction site.

Besides of the production of stable transfected COS-7 cell lines, the newly constructed BFP-HLA-C07:02~pHSE vector, whose BFP fluorescence can be easily monitored by flow cytometry, could also be used to critically evaluate the efficiencies of the transient FuGene transfection conditions that still apply for the transfection of PECPLs and potentially could be further optimized.

In preliminary experiments GERALDINE RÜHL was already examining the effect of different transfection conditions on the number of transfected PECPLs in the APCs by real-time PCR.<sup>[190]</sup> Although, these experiments revealed first insights into the correlations of input DNA and

PECPL plasmids found in transfected COS-7 cells, so far it was not possible to directly screen and evaluate different transfection efficiencies by surface expression of APCs. As transient transfections with FuGene HD reagent are highly sensitive to input plasmid DNA concentrations and cell densities of transfected cells, the BFP-HLA-C07:02~pHSE vector could be used to systematically evaluate the COS-7 cell transfection conditions (Figure 29, Figure 30).

Different input concentrations of BFP-C07 plasmid (1  $\mu$ g, 2  $\mu$ g and 1  $\mu$ g+ 1 $\mu$ g PECPL) were used for transient FuGene transfection and the BFP expression of the COS-7 cells monitored 24, 48 and 72 hours post transfection by FACS (Figure 29A). Untransfected COS-7 cells were used as a negative control and transient transfection conditions compared to BFP expression of a stable transfected BFP-C07 COS cell line (Figure 29B). In general, the BFP expression efficiencies are higher with increased amounts of input DNA and in all cases decrease slightly during culturing time from 24 to 72 hours. The best transfection efficiencies with 29 % after 24 hours post transfection and 19 % 72 hours post transfection were detected with 2  $\mu$ g input DNA. Compared with the significantly higher BFP expression of 84 % in the stable expressing HLA-BFP COS cell line, this suggests the tremendous influence that the newly established stable HLA-expressing COS-7 cells might have for prospective antigen search experiments. Altogether, the usage of stably expressing HLA-BFP COS-7 cells increases the expression of HLA molecules on the cell surface of APCs, which indirectly also increases the chances for peptide presentation. Further, by dropping the HLA plasmid from the transient FuGene transfection protocol, this increases the amount of plasmid DNA of PECPLs, which was shown to also correlates with a higher transfection efficiency.

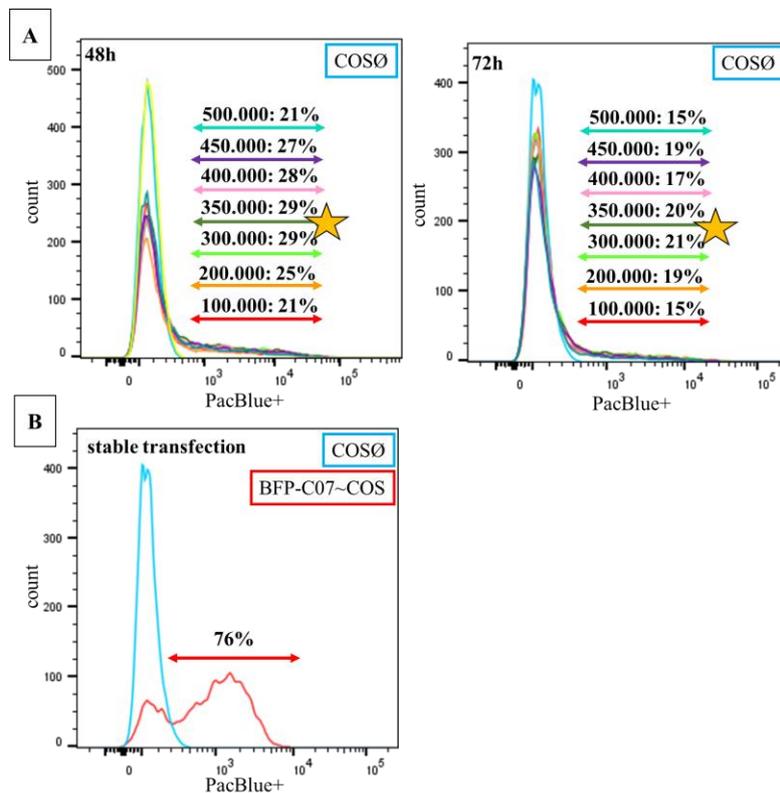


**Figure 29: Evaluating the potential effect of HLA-BFP constructs for T cell antigen search. (A)** The new BFP-C07~pHSE construct was used to monitor transfection conditions of COS-7 cells with FuGene HD reagent that were used so far. Thus, COS-7 cells were transfected with 1 µg BFP-C07~pHSE plasmid and 1 µg PEPCL library as described by KATHERINA SIEWERT (red). Additionally, transfection with different amounts of BFP-C07~pHSE plasmid (1 µg, orange and 2 µg, green) was tested. BFP expression levels were evaluated 24, 48 and 72 hours (left, middle and right panel) post-transfection by analysing PacificBlue positive (PacBlue+) cells with flow cytometry. **(B)** Untransfected COS-7 cells (COSØ, blue) were used as negative control and stable transfected BFP-C07~COS-7 cells (red) as positive control. Percentages show the amount of BFP+ cells of each condition compared to the negative control. Yellow stars highlight the condition with the highest amount of BFP+ cells.

Next, different cell densities of COS-7 cells (100.000, 200.000, 300.000, 350.000, 400.000, 450.000, 500.000) in a 6-well plate were analysed 48 and 72 hours post transfection by FACS (Figure 30A). Cells were plated 4 hours before FuGene transfection with 2 µg BFP-HLA plasmid. Untransfected COS-7 cells were used as a negative control and stably transfected BFP-C07 COS-7 cells as a positive control (Figure 30B). The relative transfection efficiency of both time points is similar and shows that an increase of cell numbers until 300.000 to 350.000 transfected COS cells also correlates with higher BFP expression levels. The expression levels decrease with higher cell numbers, which can be explained with a too high cell confluency. Thus, it could be demonstrated that the transfection conditions of the original antigen search protocol of KATHERINA SIEWERT with 40.000 cells/cm<sup>2</sup> were already chosen quite efficiently, but could be slightly optimized by reducing the COS-7 cell density to 31.000 - 36.000 cells/cm<sup>2</sup>.

Altogether, the amount of used plasmid DNA and densities of COS cells for optimal transfection efficiencies were tested and it was found that transfection of 31.000-36.000 cells/cm<sup>2</sup> with 2 µg plasmid DNA yielded the best transfection results. Further, the

comparison of the best transient transfection outcome with a maximal efficiency of around 29 % can be drastically increased by the usage of stable expressing HLA-COS cells. Indirectly this also increases the transfection efficiencies of PECPL plasmid DNA into COS-cells, by omitting the second plasmid in the transient transfection protocol. These optimized transfection conditions might increase chances for surface expression of MHC:peptide complexes and by this chances to identify target antigens.



**Figure 30: Optimizing transfection conditions of T cell antigen search with the new BFP-COS construct.** (A) The BFP-C07~pHSE construct was used to optimize cell densities of COS-7 cell transfection with FuGene HD reagent as described by the protocol of KATHERINA SIEWERT. Different amount of COS-7 cells (100.000 in red, 200.000 in orange, 300.000 in light green, 350.000 in dark green, 400.000 in pink, 450.000 in purple, 500.000 in turquoise and untransfected COS in blue) were plated 4 hours before and transfected with 2  $\mu$ g BFP-C07~pHSE plasmid. BFP expression levels were evaluated 48 hours (left panel) and 72 hours (right panel) post-transfection by analysing PacificBlue positive (PacBlue+) cells with flow cytometry. (B) Untransfected COS-7 cells (COSØ, blue) were used as negative control and stably transfected BFP-C07~COS-7 cells (red) as positive control. Percentages show the amount of BFP+ cells of each condition compared to the negative control. Yellow stars highlight the condition with the highest amount of BFP+ cells.

Lastly, antigen search experiments with 58V $\alpha$ 17V $\beta$ 4.1 cells and all newly constructed HLA-BFP~COS cells were performed with the optimized transfection conditions (Table 17). The 8L/9L library was chosen for this test, as it showed already the most promising results in the first round of T cell activation (section 4.4.1). Compared to an untransfected control, the highest amount of activated T cells was detected in HLA C07:02~BFP and it was also the only HLA with detectable clusters of more than two activated, green fluorescing T cells. This might be an indication for a presumable HLA C preference of TCR TRAV17\*01/TRBV4-1\*01, but this data have to be validated in future experiments.

**Table 17: Determining HLA restriction of TCR TRAV17\*01/TRBV4-1\*01 for T cell antigen search experiments with the new HLA~BFP COS cell lines**

HLA	Detected clusters			Ø library**
	+8L/9L library			
	Singlets	Doublets	Triplets	Singlets
A24:02~BFP	59	-	-	4
B07:02~BFP	38	-	-	3
C07:02~BFP	<b>77</b>	<b>3</b>	<b>1</b>	1

\*28.8 cm<sup>2</sup> screened per HLA

\*\*9.6 cm<sup>2</sup> screened per HLA

## 5 Discussion

The identification of pathogenic cells in autoimmune diseases such as MS is crucial for understanding cellular immune processes that would reveal tremendous new insights into disease pathogenesis. Even though the influence of several factors that might contribute to disease onset such as environmental triggers (viral infections e.g. EBV) and genetically effects were intensively investigated, no clear cause could be found yet and thus MS remains an incurable disease. Therefore, the main purpose of this work was the identification of pathogenic immune cells and particularly their target antigens.

First, a new high throughput method for the transcriptome profiling of single lymphocytes in human CSF was successfully established and validated by analysing the CSF of CIS patient #4526. Further, CSF samples of a uniquely chosen cohort of monozygotic twins with discordance for MS were used to resolve the plethora of CSF resident lymphocytic cells in different stages – even in very early pre-clinical phases – of MS on a single-cell level. Thereby, interesting clonally expanded plasma cells and particularly a dominance of expanded CD8<sup>+</sup> T cells that showed signs of an activated T<sub>RM</sub> cell phenotype were identified.

Consequently, five of the identified CSF resident expanded plasma cell clones were in-depth characterized and showed evidence for antigen-driven processes. Moreover the most frequently expanded CD8<sup>+</sup> T cell clone TCR TRAV17\*01/TRBV4-1\*01 of twin pair #161/#162 was *in vitro* reconstituted and its HLA preferences as well as its functional role assessed with different strategies and first attempts to identify potential target antigens performed. Lastly, transfection conditions of the promising CD8<sup>+</sup> T cell antigen search PECPL technology were further optimized by the introduction of a mtagBFP2 reporter construct that enabled the generation of stably HLA-expressing COS-7 cells as APCs. This foundation will be useful for more efficient prospective antigen search experiments and might reveal candidate antigens of TCR TRAV17\*01/TRBV4-1\*01 and other putatively pathogenic CD8<sup>+</sup> T cells.

## 5.1 Single-cell profiling of patients with neuroimmunological diseases

Single cell profiling methods not only open up further possibilities for detecting new, yet unknown cellular subtypes, but allow the in-depth characterization of specific cell type signatures and phenotypic clustering.<sup>[191-192]</sup> However, so far a major technical challenge of currently available methods were high samples costs and - in regards to applying it to human samples with low lymphocyte cell numbers such as CSF- a suitable high throughput method for single cell isolation. Moreover the identification of full length H:L BCR and  $\alpha:\beta$  TCR sequences directly from the transcriptome data of single cells remained elusive. Hence, it was unimaginable to link those pivotal information to cell lineage signatures, activation status, or cytokine patterns, which might reveal substantial new insights into the putative functional roles of interesting expanded lymphocyte clones.

As part of this work, the commonly used Smartseq2 scRNAseq protocol was successfully modified and this very potent analysis tool for the first time applied to resolve the heterogeneity of cellular landscapes in human CSF on a single cell level.<sup>[147]</sup>

The project arised from promising preliminary experiments that were attempting to establish a new high-throughput method for the concomitant analysis of B- and T-cell immune repertoires in CSF of MS patients. The method combined the advantages of the state-of-the-art Smartseq2 transcriptome approach of PICELLI *et al.* with flow cytometry scIS and it was possible to identify paired  $\alpha:\beta$  TCR and H:L BCR sequences from individual cells of frozen CSF samples with 80% yield.<sup>[147, 171]</sup>

Following up on this, here NGS sample preparation steps were first optimized to reduce costs and facilitate sample handling. Further, the method was validated by using the CSF of index patient #4526 and lastly applied to a uniquely chosen cohort of patients to study CSF alterations in different stages of MS disease.

### 5.1.1 Optimizing NGS preparation steps for high-throughput scRNA seq

The widely used, but expensive NGS preparation Kit from ILLUMINA for scRNA-seq applications is also part of the Smartseq2 method. However, here NGS samples preparation costs were reduced by the successful implementation of the method from BAYM *et al.*, which was designed for inexpensive multiplexed library preparation, into the original Smartseq2 protocol. With this, it was possible to significantly reduce the reaction volume for tagmentation that remained the most expensive step of NGS library preparation to a minimal volume of 2.5  $\mu\text{L}$ . The tagmentation reaction is not only the cost limiting process of NGS preparation, but an equal post-tagmentation fragment-size distribution of NGS libraries also pivotal for the sequencing results. Whereas too short fragment-sizes are removed from the final library pool during DNA purification steps, too long fragments do not bind to the NGS HiSeq chip and thus, would skew bioinformatics data interpretation. BAYM *et al.* showed that a tagmentation input concentration of  $0.5 \frac{\text{ng}}{\mu\text{L}}$  works best for genomic DNA of gram-negative bacteria. In contrast, here it could be demonstrated that for whole transcriptome amplified cDNA an input concentration of  $0.8 \frac{\text{ng}}{\mu\text{L}}$  yielded the best post-tagmentation fragment-size distributions of NGS libraries.

In this work it was not only possible to combine the advantages of the latest scRNAseq method Smartseq2 with more cost efficient NGS preparation methods, but also to omit several dispensable steps from the original Smartseq2 protocol such as quality control of every single sample on a Bioanalyzer. Altogether, the adjustments of currently available single-cell profiling methods reduces costs, facilitate sample handling and will enable an increased throughput as well as higher multiplexing of samples.

### 5.1.2 Proof of concept: single-cell profiling of human CSF of index patient #4526

Here, the strengths of the newly implemented transcriptome single-cell profiling method that combines state-of-the-art scRNA seq approaches and high-throughput flow cytometry based single cell isolation techniques could be demonstrated. The method could be successfully validated by analyzing 94 isolated, single lymphocytes from the CSF of index patient #4526.

So far, a major technical challenge of high throughput methods for single cell separation was the handling of samples with a limited number of cells such as CSF with typically less than 5 leukocytes/ $\mu\text{L}$ . Whilst in conventional flow cytometry sorting solely one population of interest is usually selected, other and in particular rare cell subtypes are lost during the sorting

process. In contrast, scIS reduces the cell loss to a minimum and therefore its implementation extremely valuable for the analysis of CSF samples. Notably, by separating the CSF preponderant CD4<sup>+</sup> T cell population from the remaining non-CD4 population, it was even possible to enrich and characterize particularly rare cell subtypes such as PBs.

Strikingly, matching of extracellular flow cytometry staining patterns and intracellular marker gene expression revealed a reliable and consistent identification of cellular subtypes and thus confirmed the reliability of the method. In contrast to currently available single-cell profiling methods that are restricted to classify cell subtypes according to their mRNA profiles, the method offers the opportunities for pre-selecting particular extracellular markers of interest that reveal additional surface expression information to transcriptome-clustered cellular subtypes. Furthermore, the adjustment of paired-end sequencing from conventional 2x 100 bp to 2x 150 bp in NGS HiSeq runs, offer the exceptional possibilities of getting full-length sequences of matching antigen-specific receptors from single lymphocytes without sophisticated bioinformatics tools that rely on the *in vitro* reconstruction of CDR3 sequences.<sup>[193]</sup> Here, it was possible to identify matching  $\alpha:\beta$  TCR and H:L BCR chain sequences from up to 90% of all analyzed lymphocytes. Concisely, even within this small sample size, it was already possible to identify clonal expansion of T and B cells. This is particularly valuable, as the group of K. DORNMAIR successfully implemented *in vitro* assays for the identification of target antigens from expanded CD8<sup>+</sup> T cell and plasma cell clones and might allow the assessment of their functional roles in prospective experiments.<sup>[133, 167]</sup>

In conclusion, the new method allowed to identify matching antigen receptor chains of single lymphocytes in combination with their whole mRNA profiles. The scRNA-seq data agree well with the scIS staining patterns, but the amount of information is much greater than that obtained with flow cytometry. Thus, it is possible to detect clonally expanded TCR and Ig receptors and characterize individual cell clones in great detail. Compared to current immunophenotyping methods, which are restricted by the limited availability of cell surface markers, the method allows a much deeper characterization of lymphocyte clones at the single-cell level.<sup>[194]</sup> Because almost no cells are discarded by scIS, the method is particularly suited for rare biological samples from any species with low lymphocyte content, e.g. human CSF. Compared to traditional approaches, that focus on the analysis of predefined lymphocyte subsets and are thus not completely unbiased, the method sheds light on the largely unknown and complex cellular processes of human autoimmune reactions.

## 5.2 Transcriptome profiling of patients with early and established MS

The cellular landscapes of CSF cells in early and established stages of MS were investigated by applying the newly implemented transcriptome approach to individuals with the highest familial risk for developing MS. The unique cohort comprised CSF samples from eight monozygotic twin pairs with discordance for MS disease. In contrast to the diseased MS siblings with clinical definite MS, the genetically identical “co-twins” were clinically unaffected and thus untreated. Of note, CSF from only four of the eight individuals with clinical manifested MS from the eight twin pairs was available. Moreover, in six of the analysed “healthy” co-twins, in-depth clinical examination of the healthy co-twins revealed evidence for subclinical neuroinflammation with conspicuous MRI peculiarities, but no dissemination in space, which is a key diagnostic criteria for MS. Whilst only two of the co-twins did not show any distinctive features of this SCNI phenotype, four individuals with MRI peculiarities even had OCBs, a hallmark diagnostic criteria of MS (Table 9). Overall, these are signs for SCNI and might represent a pre-clinical form of MS.

Here, the SCNI classification is not identical with the conventional RIS phenotype, since the clinical criteria for RIS diagnosis require spatial dissemination of MRI lesion.<sup>[195-196]</sup> In general, RIS patients are only detected by chance and therefore currently available MRI criteria strict in order to clinically differentiate them from other neuroinflammatory disease such as migraine or ischemia. In contrast, here particular study subjects with the highest familial risk for developing MS were purposefully chosen to specifically target the investigation of these very early stages of MS pathogenesis. However, it is noteworthy that the variety throughout the SCNI co-twin group was heterogeneous and comprised two subjects that fulfilled current diagnostic criteria for RIS as well as four subjects with no classical RIS phenotype. Since those four non-RIS subjects showed small MRI lesions and two even had existing OCBs, these patients might reflect a prodromal phase of MS and thus may be considered as “pre-RIS” (Table 9). Of note, one of the studied SCNI subjects of the cohort meanwhile developed a full RIS phenotype by fulfilling the MRI criteria over time. Similarly, to the widely accepted RIS stage of MS, SCNI might be recognized as an even earlier prodromal phase in MS pathogenesis, nonetheless it will remain impossible to exactly define the initial time point of lesion formation.

### 5.2.1 Cellular landscapes of CSF lymphocytes

The most striking peculiarities that were detected by comparing the cellular CSF landscapes from all study groups in the t-SNE projection was the early and consistent existence of clonally

expanded, activated CD8<sup>+</sup> T cells already in SCNI subjects. In MS the relevance of expanded CD8<sup>+</sup> T cells was long suspected as they represent the most dominating lymphocyte population in active brain lesions. <sup>[125, 127-128, 142, 151, 155]</sup> Moreover, those cells even fulfilled criteria for persistence, pervasiveness and antigen-driven clonal expansion, which even supports the assumptions of a putative role in pathogenic processes. Since up to now MS animal studies are dominated by CD4<sup>+</sup> T cell centered models as EAE, the role of CD8<sup>+</sup> T cells in animal models was not extensively studied and thus rather disregarded yet. <sup>[197]</sup> Despite of this, even their recruitment as bystander cells of a secondary immune response was discussed. In contrast, the data of this work strongly suggest the involvement of clonally expanded and activated CD8<sup>+</sup> T cells in crucial steps of MS and even emphasize their pathogenic relevance. Their presence already in SCNI subjects also support the notion of a putative contribution in very early immune processes of the disease.

The presence of CSF resident OCBs in four out of six analysed SCNI co-twins is consistent with several earlier studies that attributed OCBs a role as early immunological markers for neuroinflammation and in RIS patients, they are even used as prognostic marker. <sup>[198]</sup> Strikingly, the detection of expanded PBs in the t-SNE projections strictly correlated with the OCB status of the patients, which agrees with previous findings that identified intrathecal B cells as a source of OCBs. <sup>[199]</sup> In general, plasmablasts were found to be a typical immune cell component seen in the CSF of MS patients. <sup>[200-201]</sup> In earlier work it was even demonstrated that certain CSF resident OCBs are involved in debris removal processes that clear dying cells from previously occurring inflammatory events. <sup>[133]</sup> Altogether, the results of this work suggest that deleterious inflammatory processes are already a characteristic feature of very early - even pre-clinical-phases of MS and might also be a substantial part of even earlier stages of the disease than the SCNI state investigated here.

Earlier studies on intrathecally expanded B cell clones in MS revealed ongoing antigen-driven B cell maturation processes as IgG class switching and SHM. <sup>[149, 202-203]</sup> Whilst CD4<sup>+</sup> T cells play a major role in those processes, particular CD4<sup>+</sup> T cell lineage signatures of known effector cell types (T<sub>FH</sub>, T<sub>REG</sub>, TH1, TH2 or TH17 cells) in CSF were not detected in this work. This might be biased by the low resolution of analysed CD4<sup>+</sup> T cells in the study that is not reflecting their overall composition in human CSF and might get more differentiated with higher numbers of analysed cells. However, since scRNAseq is just about to evolve as a powerful tool, little by little traditional T cell lineage concepts started to blur and unravel an unexpected complexity and dynamic nature of cellular immune processes. <sup>[204-205]</sup> Therefore, certain effector CD4<sup>+</sup> T

cell subtypes, though not noticeably here, might actually be present and crucially involved in inflammatory CSF immune processes of MS patients. In fact, the t-SNE analysis suggest the presence of different types of CD4<sup>+</sup> T cells in SCNI and MS subjects as seen by their different positions in the t-SNE projections, and might reflect transitory stages in between different lineages of CD4<sup>+</sup> cell types.

### **5.2.2 Clonally expanded lymphocytes**

Interestingly, in-depth characterization of the expanded and non-expanded CD8<sup>+</sup> T cell populations in t-SNE projections revealed striking differences. Regarding the non-expanded CD8<sup>+</sup> T cells the observed variations between NIC, SCNI, MS and Enc subjects were only marginally and solely the analysis of specific gene expression markers for homing, migration and activation revealed insignificant differences. In stark contrast, clonally expanded CD8<sup>+</sup> T cells showed major alterations in t-SNE projections and marker gene expression between all subject groups. Though expanded cells were detected in all subjects with neuroinflammation, normalized percentages of clonal expansions were approximately twice as frequent in MS and Enc as in SCNI, and higher in SCNI than in NIC. Consistently throughout all conditions, clonal expansions were detected more frequently in the CD8<sup>+</sup> T cell population as compared to CD4<sup>+</sup> T cells. However, for both expanded T cell populations from SCNI and MS subjects subtle, but noticeable differences were detected in the t-SNE projections and thus indicating altered cellular phenotypes between both subject groups. Since t-SNE projections only represent dimensional reduced data of thousands of highly diverse genes, a more detailed analysis of specifically selected marker genes was required to further address those phenotypical differences.

### **5.2.3 In-depth characterization of expanded T cells**

Strikingly, a detailed heatmap analysis of selected gene expression markers related to migration, homing and activation of CD8<sup>+</sup> T cells, revealed strong downregulation of T cell egress markers S1PR1, CCR7, SELL and TCF7 - key features of a T<sub>RM</sub> like phenotype - in clonally expanded CD8<sup>+</sup> T cells of MS when compared to NIC. In contrast, those egress markers were also slightly downregulated in non-expanded CD8<sup>+</sup> cells of MS, indicating a putative transitional stage of those cells in the evolvement of a T<sub>RM</sub>-like phenotype that might be triggered by the present inflammatory milieu. For SCNI subjects, T cell egress marker of non-expanded CD8<sup>+</sup> T cells showed no apparent differences when compared to cells from NIC. Notably, though not as

concise as in expanded CD8<sup>+</sup> T cells of MS, these findings demonstrated the downregulation of egress markers CCR7 and SELL as well as the upregulation of proinflammatory cytokines in expanded CD8<sup>+</sup> T cells of SCNI. Evidentially the results of this work suggest that those SCNI cells are in a transitory state and progressing towards an effector phenotype, which agrees very well with their different localizations in the t-SNE projections (Figure 20D, E). This gradual transition of T cell egress marker expression changing from a T<sub>CM</sub> phenotype to a more definite effector T<sub>RM</sub> phenotype was concisely found during the course of MS disease. Moreover, this phenotypic changes move evenly from CD8<sup>+</sup> T cells of NIC to non-expanded cells of SCNI via expanded cells of SCNI and non-expanded cells of MS and ultimately concluding in a fully established T<sub>RM</sub> phenotype in expanded CD8<sup>+</sup> T cells of patients with definite MS. Remarkably, this agrees very well with the further detected phenotypic changes of expanded CD8<sup>+</sup> T cells as the gradual upregulation of inhibition and retention markers (such as PDCD1 or CD69), cytokines and activation markers (such as IFN- $\gamma$  or CXCR6). Those markers also gradually shift in the same way as described for T cell egress markers. Conclusively this further supported the notion that expanded CD8<sup>+</sup> T cells phenotypically change from a T<sub>CM</sub> phenotype to a more definite T<sub>RM</sub> phenotype.

The relevance of autoreactive lymphocytes in MS pathology is commonly accepted, but the underlying processes that lead to their recruitment and further on their downstream cellular targets remain fundamental yet unanswered key questions. Even in noninflammatory conditions those autoreactive T cells are known to persist in the brain parenchyma and recent studies suggested them a primarily T<sub>RM</sub> phenotype.<sup>[206-207]</sup> However, the results of this work revealed a striking upregulation of the chemokine CXCR6 in expanded CD8<sup>+</sup> T cells of MS and Enc subjects. CXCR6 is known to play a role in various human diseases and in particular under pathological conditions was shown to be involved in the recruitment of activated CD8<sup>+</sup> T cells, but not CD4<sup>+</sup> T cells.<sup>[208-210]</sup> It is exclusively binding to the CXCL16 receptor and consistently an upregulation of CXCL16 gene expression levels was found in monocytes and DCs. Further, even in expanded CD4<sup>+</sup> T cells of SCNI and MS elevated levels of CXCR6 were detected, which might also be linked to the involvement of mononuclear phagocytes and/or DCs that serve as APCs.<sup>[211]</sup> In accordance with earlier studies, pDCs and more prevalent myeloid DCs were already identified in CSF of NIC subjects, but numbers were strikingly increased in CSF of neuroinflammatory conditions as in MS subjects.<sup>[212]</sup> Further, a communication between CD8<sup>+</sup> T cells and mononuclear phagocytes was just recently suggested by a study that was analysing interaction partners of perforin-stained CD8<sup>+</sup> T cells in MS brain lesions.<sup>[213]</sup>

Although the data strongly suggest particular functional interactions of CD8<sup>+</sup> T cells and mononuclear phagocytes and/or DCs, further studies are needed to shed light on this particular correlation.

#### **5.2.4 Expanded CD8<sup>+</sup> T cells of individual twin pairs**

A fundamental and obvious limitation of the study design is the comparatively small number of samples in the cohort group that just reveal a first glimpse into the complexity of inflammatory processes in CSF of very early stages of MS. Thus, these findings have to be validated in further studies with larger sample sizes. Furthermore, no patients with CIS diagnosis were included in the study as CSF samples of those patients are usually taken in close proximity to clinically active phases of the disease. To facilitate data comparability and thus reduce potential confounding factors that might skew the analysis, only study subjects outside of clinical activity were chosen. Whereas all SCNI subjects were untreated, the MS group comprised patients with each receiving a diverse immunomodulatory therapy. This treatment heterogeneity was surprisingly not significantly affecting the overall distributions of immune cell populations, clonal expansions or gene expression levels of CSF cells in MS individuals. The only remarkably exception was seen in MS twin #301, who received treatment with the drug teriflunomide, an inhibitor of lymphocyte proliferation that affected gene expression patterns and resulted in noticeably phenotypical changes (Figure 23). In contrast to the immune therapies of all other twins as interferone-beta (twin #155), natalizumab (twin #161) and intrathecal steroids (twin #163), teriflunomide is also the only drug that has an overall and direct effect on lymphocyte cells, thus phenotypic alterations are not surprising.<sup>[214-215]</sup> Whereas interferone-beta just indirectly inhibits the activation of T cells by affecting DCs, natalizumab solely blocks the entrance of autoreactive lymphocytes into the CNS by affecting BBB trafficking and steroids are universally acting anti-inflammatory, teriflunomide directly interferes with actively proliferating T and B cells by suppressing DNA replication enzymes. Obviously, these initial findings and a potential influence of treatment bias has to be further investigated and validated in studies with larger cohorts.

In conclusion, these results provide first unique insights into very early phases of neuroinflammation in MS that even occur before clinical onset of the disease. Fortunately, it was possible to study the complexity of CSF cells in clinically unaffected individuals with the highest familiar risk for developing definite MS as their associated monozygotic siblings were affected by MS. Substantial evidence for the collaborative participation of all three fundamental

immune cell populations (B cells, CD4<sup>+</sup> and CD8<sup>+</sup> T cells) of the adaptive immune system in very early stages of disease pathology were found. Particularly the clonally expanded CD8<sup>+</sup> T cell population of SCNI and even more pronounced in MS revealed conspicuous alterations in gene expression levels. Those phenotypical changes suggest a very early concomitant involvement of activated clonally expanded CD8<sup>+</sup> T cells with a T<sub>RM</sub>-like phenotype in disease pathogenesis. In tumours and other infectious diseases, T<sub>RM</sub>-like clonally expanded CD8<sup>+</sup> T cells could be directly linked to disease progression.<sup>[216-218]</sup> Consequently, also in MS the direct involvement of expanded CD8<sup>+</sup> T cells in disease progression and even further their potential role in CNS lesions, where they constitute the dominating and persisting lymphocyte population, could be speculated.<sup>[128, 142, 219-221]</sup> Even though further studies with larger cohorts need to be performed to validate this data and draw more concise conclusions, these results highly emphasize the need for studying interaction partners and target antigens of intrathecally expanded CD8<sup>+</sup> T cell clones in more depth. This might not only reveal implications for MS pathogenesis, but on the long run might contribute to more target-specific new therapeutic treatment options.

### **5.2.5 Transcriptome profiling of patients with CIS diagnosis**

CIS patients are usually diagnosed after the first clinical symptoms and thus in active and highly inflammatory phases of MS. Therefore, those subjects were not included in the previous analysis to facilitate data comparability and prevent skewed bioinformatics data interpretation. However, the in-depth phenotypic characterization of CSF resident and locally recruited immune cells in those highly active disease phases might reveal new insights and/or characteristic patterns of the involved lymphocytes especially in inflammatory processes of MS. Here, the scRNAseq approach was used to shed light on these open questions by analysing three patients with a CIS diagnosis, but data analysis is still ongoing.

### 5.3 Characterization of expanded lymphocyte clones

The identification of locally persistent, pervasive, clonally expanded PBs and CD8<sup>+</sup> T cells in the CSF of patients - as demonstrated in the previous chapter - clearly emphasized the involvement of particular target antigens in MS pathogenesis that drive lymphocyte outgrowth. Thus, in the second part of this work, six intrathecally expanded PB clones from five analysed subjects in different stages of MS disease (one SCNI, three CIS and one early MS), and one particularly interesting CSF-resident CD8<sup>+</sup> T cell clone of twin pair #161/#162 were successfully characterized in more detail. Moreover, the functional roles of the most prominent expanded CD8<sup>+</sup> T cell clone TCR TRAV17\*01/TRBV4-1\*01 in disease pathogenesis was assessed by investigating its HLA preferences with different strategies.

#### 5.3.1 Expanded plasma cells

Intrathecally expanded plasma cells are a hallmark of patients with MS and were found at least to some extent to be a source of CSF persistent OCBs that are commonly used as diagnostic criteria of clinically defined MS.<sup>[198-199]</sup> In the previous chapter of this work, it was already demonstrated that expanded PBs were - as expected- a consistent attribute in CSF of MS patients. In addition, the presence of expanded PBs already in patients with SCNI - a very early pre-clinical stage of MS pathogenesis - could be demonstrated and further its restriction to subjects with CSF resident OCBs. Whereas in MS antigen-driven B cell maturation processes were studied extensively and also their putative involvement in debris removal processes was controversially discussed, the functional roles of expanded plasma cells in pre-clinical stages of MS as in SCNI were not addressed so far.<sup>[133, 149, 202-203]</sup> Moreover, after flow cytometry scIS plasma cells were identified in three bioinformatically not yet analysed CIS patients. To compare functional roles of expanded PBs in different disease stages of MS, advantageously the transcriptome approach enabled the identification of H:L BCR sequences of those PBs from the CIS patients even with conventional Sanger sequencing.

Overall, the in depth characterization of all CSF-resident PBs from five patients with heterogeneous MS disease status (including one patient with early MS, three CIS patients and one subject with SCNI) revealed a striking IgG dominating clonal expansion of at least one PB clone with noticeable SHM in hypervariable CDR and frame regions of the variable part of H:L BCR chains. The only exception was CIS patient #6014, who only showed four PBs in total and thus numbers might not be significant. Clonal expansion, SHM and IgG class switching are characteristic criteria for antigen driven B cell maturation processes to produce target-specific

antigen receptors with high specificities.<sup>[149]</sup> Therefore, target antigens and the functional relevance from the four most dominant expanded PBs of each subject were further investigated *in vitro* by MD student ISABELLE BARLIANTO. Recombinant antibodies from expanded plasma cells #301-1, #162, #4526-1 and #4626 that were identified in section 4.3.1 were cloned, expressed and analysed for antigen recognition by ISABELLE BARLIANTO with an already established microarray based technique (as described by BRÄNDLE *et al.*).<sup>[133]</sup> In agreement with prior studies on target antigens of CSF resident PBs from MS patients, exclusively intracellular antigens such as fibroblast growth factor receptor 2 (FGFR2), PDZ binding kinase (PBK), neurotrophic tyrosine kinase receptor type 2 (NTRK2) and kinase insert domain receptor (KDR) were identified for all analysed PB clones- even for the subject with SCNI (data not published yet).<sup>[133]</sup>

Although the analysed clone numbers were not significantly high, this results strengthen the already existent ‘debris removal theory’. Conclusively, even in very early stages of disease pathology, these results emphasizes their role as bystander cells that are involved in a secondary immune response arising from the destructive inflammatory processes of the disease. Since expanded PBs seems to play a minor role in the pathogenic processes of MS and in fact CD8<sup>+</sup> T cells remain the most prevalently expanded immune cell population in CSF and brain tissues of MS patients, this emphasized the necessity of identifying CD8<sup>+</sup> T cell antigens.

### 5.3.2 Expanded CD8<sup>+</sup> TCR TRAV17\*01/TRBV4-1\*01

In order to shed light on the yet unknown functional roles of expanded CD8<sup>+</sup> T cell clones in MS pathogenesis, the most dominantly expanded CD8<sup>+</sup> T cell clone TCR TRAV17\*01/TRBV4-1\*01 that was detected in twin pair #161/#162 was *in vitro* reconstituted. Since the expanded TCR TRAV17\*01/TRBV4-1\*01 was not only detected in MS twin #161, but also in the clinically healthy SCNI subject #162, in depth sequence analysis was performed. Moreover, T cell repertoires are shaped during thymic selection processes and thus TCRs restricted to specific HLA alleles.<sup>[222-223]</sup> Beneficially, twin pair #161/#162 showed homozygosity for all three HLA class I loci. However, knowing the HLA restriction of TCR TRAV17\*01/TRBV4-1\*01 would substantially simplify T cell activation assays of subsequently performed antigen search experiments. Therefore, its HLA restriction was assessed by different strategies and lastly applied for unbiased identification of target antigens.

Antigen recognition of T cells is known to be MHC restricted and each individual's TCR repertoire shaped by genetic traits and lifetime environmental influences as viral exposures. The CD8<sup>+</sup> T cell clone TCR TRAV17\*01/TRBV4-1\*01 was detected 15 times in the CSF of MS twin #161 and unexpectedly also four times in the CSF of the SCNI co-twin #162. Moreover, in both subjects TCR TRAV17\*01/TRBV4-1\*01 was representing the most dominantly expanded CSF resident clone. Monozygotic twins and even non-related persons that share identical HLA alleles evidentially showed similarities in previous T cell repertoire studies and notably, increased frequencies of CDR3 overlaps were identified in monozygotic twins.<sup>[224-225]</sup> However, codon degeneracy that determine the nucleotide usage of amino acids is an inevitable feature even of monozygotic individuals.<sup>[226-227]</sup> Strikingly, in-depth analysis of TCR TRAV17\*01/TRBV4-1\*01 in patients #161 and #162 revealed identical CDR3 amino acid sequences, but different nucleotide usage in each individual.

Of note, besides of the TRAV17\*01 chain sequence that was found fifteen times in patient #161, a second, additional TRDV\*01 chain was detected in nine of the fifteen wells. Theoretically, the presence of dual, functional alpha chains in T cells is plausible, since allelic exclusion of the TCR locus is incomplete.<sup>[189]</sup> Further, a putative relevance of TCRs with dual alpha chains in certain autoimmune diseases was discussed as demonstrated in mouse models of diabetes.<sup>[228]</sup> However, the TRDV\*01 chain could not be functionally expressed in *in vitro* cell culture settings and conclusively not seem to play a role in antigen recognition processes and thus not relevant for following experiments.

### 5.3.3 Investigating HLA preferences of TCR TRAV17\*01/TRBV4-1\*01

Earlier findings of the group of K. DORNMAIR already demonstrated that T cell clonotypes were shared between MS brain lesions and peripheral blood.<sup>[128]</sup> Moreover, recent studies even found a strong link of the dominant, expanded clones found in active brain lesions of MS patients with the most frequent CD8<sup>+</sup> memory T cell clones in circulating blood.<sup>[128, 229-230]</sup> In contrast to these findings, TCR TRAV17\*01/TRBV4-1\*01 that represented the most dominantly expressed CD8<sup>+</sup> T cell clone in the CSF of twin pair #161/#162, could not be detected neither in naïve nor in CCR7<sup>+</sup> antigen-experienced CD8<sup>+</sup> T cells isolated from PBMCs of the patients. Antigen recognition of CD8<sup>+</sup> T cells is shaped during selection processes in the thymus and generates a highly diverse and individual TCR repertoire that is restricted to the personal MHC.<sup>[222-223]</sup> However, amongst the vast diversity of each individual's TCR repertoire, the existence and relevance of so called public TCRs that are typically shared between individuals with identical HLA alleles was already discussed in autoimmune diseases as Rasmussen

Encephalitis.<sup>[227, 231-232]</sup> Apart from genetically influences that usually shape immunological repertoires, thymic and peripheral processes seem to have the strongest impact on the higher abundancy of those public CD8<sup>+</sup> T cell responses in such diseases. As TCR TRAV17\*01/TRBV4-1\*01 was identified in two different, but genetically related individuals of different disease stages of MS, it was tempting to speculate that this T cell clone might also be involved in such a public adaptive immune response and thus potentially also present in other MS patients with matching HLA alleles. This hypothesis was even strengthened by recent findings of HOU *et al.* that identified conspicuously similar characteristics for highly expanded and public T cell clones such as typical short, rather germline-like encoded CDR3 regions.<sup>[233]</sup> Further, in agreement with the findings of TCR TRAV17\*01/TRBV4-1\*01, they even demonstrated that public clones with shared CDR3 amino acid sequences were commonly produced by the usage of multiple, different nucleotides. Conclusively, the existence of TCR TRAV17\*01/TRBV4-1\*01 in PBMCs of 15 patients with heterogeneity regarding their MS disease stage and partially as well as fully matching HLA class I alleles was investigated. This would not only clarify its role as a putative public TCR, but also might reveal HLA restriction of the T cell clone. Therefore primers for the PCR amplification of the TRBV4-1\*01 CDR3 $\beta$  region were designed by using wobble nucleotides to ensure a maximum of flexibility that covers all possible nucleotide sequences, whilst maintaining the specific CDR3 amino acid sequence. Nonetheless, TCR TRAV17\*01/TRBV4-1\*01 could not be identified in PBMCs of any of the analysed patients. Only in PBMCs of patient #130, that shared the HLA~C07:02 allele of twin pair #161/#162, a similar CDR3 $\beta$  amino acid sequence was found. Anyhow, the matching TRAV17\*01 CDR3 $\alpha$  sequence, which is essential for antigen recognition, could not be detected and thus no further conclusions on the HLA preference could be drawn from this preliminary experiments. Since around 87 % of the analysed patients shared the specific HLA alleles, but predominantly the HLA~C07:02 allele and with approximately 13% only a minor fraction sharing the HLA~A24:02 allele, this correlation might need to be further investigated and addressed in prospective experiments with a larger and more carefully selected patient group. Further, even though studies suggested a clonotype overlap of prominently expanded T cell clones in CNS and periphery, the results of this work suggest that TCR TRAV17\*01/TRBV4-1\*01 might only be locally present in CNS compartments and thus CSF samples of more patients might further clarify this particular correlation.

EBV infection is strongly linked to MS and neuropathological studies even suggest that persistent EBV infections might be directly involved in the generation of CNS injuries as local interaction of cytotoxic T cells and EBV infected cells were detected in post mortem brains of

MS patients.<sup>[234-235]</sup> Moreover, autologous EBV-transformed B cells presenting public EBV-specific antigen sequences were already used in earlier studies to evaluate potential roles of EBV antigens and T cell activation in MS.<sup>[236]</sup> Hereby, they even detected intrathecally enriched EBV specific CD8<sup>+</sup> T cells in MS patients. Therefore, EBV-transformed B cells from PBMCs of five patients with heterogeneous HLA alleles, including the TCR TRAV17\*01/TRBV4-1\*01 originating patient #162 and two patients with partially matching HLA alleles, were used for T cell activation assays. However, no significant T cell activation could be detected in the performed EBV related activation experiments. Hence, neither the HLA preference of TCR TRAV17\*01/TRBV4-1\*01 could be investigated here, nor a putative involvement of endogenously presented EBV antigens detected.

In conclusion, it was not possible to reliably identify the HLA preferences of TCR TRAV17\*01/TRBV4-1\*01. However, advantageously twin pair #161/#162 is homozygous for all three HLA class I loci and this already reduced the number of possibilities for T cell antigen search experiments significantly.

## 5.4 T cell antigen search

The most prominent expanded CD8<sup>+</sup> T cell clone TCR TRAV17\*01/TRBV4-1\*01 of twin pair #161/#162 was *in vitro* reconstituted and initial experiments to identify putative candidate antigens with *in vitro* based fluorescence assays performed.

In general, antigen driven immune processes play a key role in the pathogenesis of tumours, infectious diseases and autoimmunity.<sup>[237-238]</sup> Usually pathogenic antigens are taken up by APCs and presented as processed, antigenic peptide fragments in peptide:MHC complexes to cells of the adaptive immune system. Under inflammatory conditions lymphocytes are recruited to local tissues, get activated by binding their cognate antigens, which leads to differentiation and expansion into effector cell clones with the same receptor-specificity and conclusively triggers the effective clearance of the invading pathogen. This fundamental process was used previously to successfully establish an approach for the unbiased identification of target antigens of CD8<sup>+</sup> T cells in an *in vitro* model system.<sup>[167]</sup> The strengths and tremendous potential of this assay was already effectively demonstrated in different diseases as in the inflammatory skin disease psoriasis vulgaris and the muscle fiber autoimmune disease myositis.<sup>[168, 239-240]</sup> However, apart from the selection of an interesting and putatively disease-relevant CD8<sup>+</sup> T cell clone, a pivotal key criteria for the success of those antigen search experiments significantly depend on the proper presentation of MHC:peptide complexes by the used APC models. Thus, transfection conditions of COS-7 cells are critical and could be optimized.

Therefore currently used transfection conditions for transient transfection of COS-7 cells with FuGene reagent were optimized by implementation of a mtagBFP2 fluorescence reporter construct into the intracellular C-terminal domain of HLA~C07:02. This not only enabled the critical evaluation of currently used transfection conditions, but also the establishment of stably HLA expressing COS-7 cell lines and thus more efficient antigen search experiments. Altogether, this optimization will improve antigen-presentation performances in prospective experiments, increase chances for successful T cell activation and potentially contribute to the identification of more T cell antigens.

### 5.4.1 Antigen search experiments of TCR TRAV17\*01/TRBV4-1\*01

The PECPL based antigen search method for unbiased identification of target antigens from *in vitro* reconstituted CD8<sup>+</sup> TCRs was first implemented with T cell clone JM22 and flu antigen. In this experimental setup the HLA~A02 restriction of the T cell clone was already known and

this facilitated experiments drastically. In this work, first preliminary T cell activation experiments with the JM22 TCR and flu antigen were performed to reliably reproduce the data from K. SIEWERT and exclude methodical flaws for subsequently performed unbiased antigen search experiments.

Since the HLA preference of TCR TRAV17\*01/TRBV4-1\*01 remained elusive, scan experiments were performed with all three HLA alleles (A24:02, B07:02, C07:02) and seven different libraries. Preferentially C-terminal primary anchors with hydrophobic amino acid residues were recently identified as prominent peptide presentation motifs of all three TCR TRAV17\*01/TRBV4-1\*01 related HLA molecules, that was particularly valuable for the yet largely underestimated HLA~C07:02 allele.<sup>[241]</sup> These findings agree very well with the results from the T cell activation experiments that revealed a preference for HLA~C07:02 restricted peptide presentation of eight and nine amino acids length with fixed Leucine anchor positions. APCs synthesize antigenic peptides with a methionine start codon that is enzymatically cleaved by the methionin-aminopeptidase during cytosolic antigen processing steps. However, the efficiency of this enzyme crucially depends on the size of the initial amino acid and might be impaired when the amino acid following the start codon is sterically big such as Leucine.<sup>[242]</sup> This might lead to peptide variants with N-terminal methionine and could provoke misleded TCR activation or prevent proper presentation of candidate antigen by the APC. Overall, this might be potential reasons for difficulties in the T cell antigen search experiments and causative for the lost T cell activation signals after several rounds of single cell isolation and three rounds of subpool experiments, which made it impossible to identify any candidate mimotopes of TCR TRAV17\*01/TRBV4-1\*01 in this work.

In contrast to B cells, certain characteristics of T cells such as low TCR binding affinities and/or avidities to peptide:MHC complexes, polyspecificity and high receptor diversity, are pivotal for T cell activation and might directly impact T cell specificities, generally making T cell antigen searches methodically more challenging.<sup>[75, 158]</sup> Moreover, the “causing” antigen for TCR TRAV17\*01/TRBV4-1\*01 expansion might be an external infectious agent or triggered by endogenous retrovirus and thus not detectable with the PECPL based antigen search method. Even its role as a “bystander” cell that might just be reflecting the involvement in a secondary immune response, as proposed for OCB-resident plasma cells, might be possible.

#### **5.4.2 Optimizing the T cell antigen search protocol**

The functionality of the PECPL based technology for the identification of target antigens of CD8<sup>+</sup> T cells was already proven in multiple studies.<sup>[167-168, 240]</sup> Anyhow, the efficiencies of

HLA-dependent PECPL presentation of COS-7 cells is a significant influencing factor for the success of the method. Fluorescence antibodies against human HLA molecules besides of well-established exceptions such as HLA~A02 are only very rarely existent, highly cross-reactive and thus neither relevant nor existent for the majority of HLA molecules.<sup>[170]</sup> Moreover, with the currently used FuGene protocol, transfection efficiencies remained uncertain and could be just estimated by comparison to a GFP control. The implementation of a BFP reporter construct into the cytosolic domain of HLA~C07:02 for the first time enabled the reliably evaluation of the used transfection conditions of the PECPL based method. Moreover, the BFP construct was designed as a cassette vector that enables simple exchange of HLA sequences in a one-step cloning process and thus can be used to clone any HLA of interest for more efficient prospective antigen search experiments. Since mtagBFP is not interfering with already used fluorescence readout parameters of the antigen search method, it was not only used here to successfully improve transfections conditions, including the amount of plasmid DNA and COS-7 cell densities, but even enabled the establishment of stably HLA expressing COS-7 cell lines. Compared to transient transfections, these results demonstrated that the usage of stable COS-7 cell lines not only increase HLA expression by nearly threefold, but reduce the amount of plasmids for transient FuGene transfections from two to one and thereby increasing PECPL plasmid transfection efficiencies significantly.

Altogether, the used FuGene transfection conditions of the standard protocol were assessed and though yielding sufficient transfection efficiencies, could even be further improved. The introduction of mtagBFP seems to be a promising improvement of the method that does not interfere with current parameters of the PECPL method, and adds more transparency to the so far used FuGene transfection conditions. Moreover, the mtagBFP construct that was solely implemented in the cytosolic domain of the HLA proteins, enables the construction of stably transfected COS-7 cell lines regardless of the availability of HLA conjugated fluorescence antibodies. Overall, these new implementations lay the foundation for further methodical optimization of the standard antigen search protocol. Based on the findings of this work it can be assumed that chances of detecting candidate antigens can be increased by using the here implemented HLA BFP construct and newly optimized FuGene transfection conditions.

## 5.5 Outlook

The adaptive immune system, in particular the concomitant interplay of CD8<sup>+</sup> T cells and B cells play an essential role in active brain lesions of MS patients. The criteria for pathogenic relevant players are still vaguely, but remain the fundamental basis for the identification of their targeted structures in diseased tissues and might open up new possibilities for more target-specific therapeutic treatments on the long run.

First, the strengths of this newly established high throughput method for single-cell analysis of CSF resident lymphocytic cells, which will allow the study of MS patients in different disease stages in prospective experiments, was demonstrated. In particular, the characterization of CIS patients remains intriguingly and might not only reveal further insights into disease pathology, but also clarify the relevance of the here detected expanded T<sub>RM</sub>-like CD8<sup>+</sup> T cells in clinically active phase of the disease. Moreover, even beyond MS, the high-throughput method could be particularly valuable for studying rare biological samples with low lymphocyte numbers from any species e.g. CSF in other inflammatory or autoimmune diseases.

This might strengthen our knowledge of immune processes contributing to pathogenesis of such diseases and might lead to new, more precise criteria for the identification of pathogenic immune cells.

Here, the identified TCR TRAV17\*01/TRBV4-1\*01 with a dominant clonal expansion and a particular T<sub>RM</sub> phenotype was chosen for antigen search experiments. Though preliminary experiments on HLA restriction suggested a putative HLA C preference, data are not sufficiently enough to assure this and no target antigen could be identified in this work. Further experiments have to be performed for statistical significance and might contribute to the identification of putative candidate antigens by revealing the restricting HLA allele. Moreover, the implementation of stably transfected HLA-BFP COS-7 cell lines that also enabled the optimization of FuGene transfection conditions seems to be a promising methodical optimization for prospective antigen search experiments. Preliminary scan experiments with the new BFP constructs seemed promising and revealed a presumable HLA C restriction of TCR TRAV17\*01/TRBV4-1\*01, but this has to be further assessed by more experiments in the future. In case, the peptide binding preferences for HLA C could be proven, also more specific libraries with fixed anchor positions can be used to screen TCR TRAV17\*01/TRBV4-1\*01 for antigen recognition. The practical applicability of the methodical optimization have to be also still verified and could be tested with the already well-characterized JM22 model of the original antigen search protocol.

## 6 References

- [1] S. H. E. Kaufmann, A. Sher and R. Ahmed, *Immunology of infectious diseases*, American Society for Microbiology (ASM), Washington, **2002**, p. xi + 495 pp.
- [2] R. M. Zinkernagel, *Science* **1996**, *271*, 173.
- [3] B. P. Morgan and M. J. Walport, *Immunology Today* **1991**, *12*, 301-306.
- [4] C. A. Mims, A. Nash and J. Stephen, *Mims' Pathogenesis of Infectious Disease* **2001**, 67-83.
- [5] A. J. Husband in *Overview of the Mammalian Immune System*, Eds.: B. Woodward and H. H. Draper), Springer US, Boston, MA, **2001**, pp. 3-14.
- [6] S. Greenberg and S. Grinstein, *Current Opinion in Immunology* **2002**, *14*, 136-145.
- [7] D. Roos, H. Spits and C. E. Hack in *Innate immunity — phagocytes, natural killer cells and the complement system*, Eds.: F. P. Nijkamp and M. J. Parnham), Birkhäuser Basel, Basel, **2005**, pp. 63-80.
- [8] O. Takeuchi and S. Akira, *Cell* **2010**, *140*, 805-820.
- [9] S. Akira, S. Uematsu and O. Takeuchi, *Cell* **2006**, *124*, 783-801.
- [10] H. D. Brightbill and R. L. Modlin, *Immunology* **2000**, *101*, 1-10.
- [11] J. R. Dunkelberger and W. C. Song, *Cell Res* **2010**, *20*, 34-50.
- [12] R. Medzhitov and C. A. Janeway, *Current Opinion in Immunology* **1997**, *9*, 4-9.
- [13] R. M. Steinman and H. Hemmi in *Dendritic Cells: Translating Innate to Adaptive Immunity*, Eds.: B. Pulendran and R. Ahmed), Springer Berlin Heidelberg, Berlin, Heidelberg, **2006**, pp. 17-58.
- [14] A. Le Bon and D. F. Tough, *Current Opinion in Immunology* **2002**, *14*, 432-436.
- [15] J. Arsenio, B. Kakaradov, P. J. Metz, S. H. Kim, G. W. Yeo and J. T. Chang, *Nature Immunology* **2014**, *15*, 365-372.
- [16] K. K. McKinstry and T. M. Strutt in *Regulation and Maintenance of Adaptive Immunity*, Eds.: L. M. McManus and R. N. Mitchell), Academic Press, San Diego, **2014**, pp. 20-35.
- [17] F. A. Bonilla and H. C. Oettgen, *Journal of Allergy and Clinical Immunology* **2010**, *125*, S33-S40.
- [18] A. Y. Lai and M. Kondo, *Seminars in immunology* **2008**, *20*, 207-212.
- [19] V. F. J. Quesniaux, J. D. Down and H.-J. Schuurman in *Hematopoiesis, including lymphocyte development and maturation*, Eds.: F. P. Nijkamp and M. J. Parnham), Birkhäuser Basel, Basel, **2005**, pp. 3-17.
- [20] F. M. Burnet, *CA: A Cancer Journal for Clinicians* **1976**, *26*, 119-121.

- [21] H. von Boehmer, *Cell* **1994**, 76, 219-228.
- [22] G. Trinchieri, M. Kubin, G. Bellone and M. A. Cassatella, *Journal of Cellular Biochemistry* **1993**, 53, 301-308.
- [23] M. M. Davis and P. J. Bjorkman, *Nature* **1988**, 334, 395-402.
- [24] D. T. Fearon and R. M. Locksley, *Science* **1996**, 272, 50.
- [25] J. G. Hirsch, *Bacteriological reviews* **1959**, 23, 48-60.
- [26] T. P. Janeway CA Jr, Walport M, et al. , *Immunobiology: The Immune System in Health and Disease. 5th edition. New York: Garland Science* **2001**.
- [27] E. Edry and D. Melamed, *J Immunol* **2004**, 173, 4265-4271.
- [28] F. Melchers, A. Rolink, U. Grawunder, T. H. Winkler, H. Karasuyama, P. Ghia and J. Andersson, *Current Opinion in Immunology* **1995**, 7, 214-227.
- [29] F. Melchers, A. Rolink, U. Grawunder, T. H. Winkler, H. Karasuyama, P. Ghia and J. Andersson, *Curr Opin Immunol* **1995**, 7, 214-227.
- [30] F. Melchers, D. Haasner, U. Grawunder, C. Kalberer, H. Karasuyama, T. Winkler and A. G. Rolink, *Annu Rev Immunol* **1994**, 12, 209-225.
- [31] P. Parham, *J Immunol* **1983**, 131, 2895-2902.
- [32] A. L. DeFranco, J. D. Richards, J. H. Blum, T. L. Stevens, D. A. Law, V. W. Chan, S. K. Datta, S. P. Foy, S. L. Hourihane, M. R. Gold and et al., *Ann N Y Acad Sci* **1995**, 766, 195-201.
- [33] J. M. Woof and D. R. Burton, *Nature Reviews Immunology* **2004**, 4, 89-99.
- [34] D. R. Burton and J. M. Woof in *Human Antibody Effector Function, Vol. 51* (Ed. F. J. Dixon), Academic Press, **1992**, pp. 1-84.
- [35] C. J. Paige and G. E. Wu, *Faseb j* **1989**, 3, 1818-1824.
- [36] G. M. Griffiths, C. Berek, M. Kaartinen and C. Milstein, *Nature* **1984**, 312, 271-275.
- [37] H. N. Eisen and G. W. Siskind, *Biochemistry* **1964**, 3, 996-1008.
- [38] J. Stavnezer and C. T. Amemiya, *Seminars in Immunology* **2004**, 16, 257-275.
- [39] R. Ahmed and D. Gray, *Science* **1996**, 272, 54.
- [40] L. Klein, M. Hinterberger, G. Wirnsberger and B. Kyewski, *Nature Reviews Immunology* **2009**, 9, 833-844.
- [41] T. K. Starr, S. C. Jameson and K. A. Hogquist, *Annual Review of Immunology* **2003**, 21, 139-176.
- [42] G. Anderson, N. C. Moore, J. J. T. Owen and E. J. Jenkinson, *Annual Review of Immunology* **1996**, 14, 73-99.

- [43] K. Murphy and C. Weaver, *Janeway's Immunobiology, 9th edition*, **2016**, p.
- [44] H. Suzuki, J. A. Punt, L. G. Granger and A. Singer, *Immunity* **1995**, *2*, 413-425.
- [45] H. Cantor and E. A. Boyse, *Cold Spring Harb Symp Quant Biol* **1977**, *41 Pt 1*, 23-32.
- [46] T. R. Mosmann, L. Li, H. Hengartner, D. Kagi, W. Fu and S. Sad, *Ciba Found Symp* **1997**, *204*, 148-154; discussion 154-148.
- [47] A. K. Abbas, K. M. Murphy and A. Sher, *Nature* **1996**, *383*, 787-793.
- [48] T. J. Braciale, L. A. Morrison, M. T. Sweetser, J. Sambrook, M. J. Gething and V. L. Braciale, *Immunological reviews* **1987**, *98*, 95-114.
- [49] J. P. Shaw, P. J. Utz, D. B. Durand, J. J. Toole, E. A. Emmel and G. R. Crabtree, *Science* **1988**, *241*, 202.
- [50] M. J. Smyth and J. A. Trapani, *Immunology Today* **1995**, *16*, 202-206.
- [51] M. Barry and R. C. Bleackley, *Nature Reviews Immunology* **2002**, *2*, 401-409.
- [52] Y. Y. Wan and R. A. Flavell, *Journal of Molecular Cell Biology* **2009**, *1*, 20-36.
- [53] L. Rowen, B. F. Koop and L. Hood, *Science* **1996**, *272*, 1755-1762.
- [54] B. Arden, S. P. Clark, D. Kabelitz and T. W. Mak, *Immunogenetics* **1995**, *42*, 455-500.
- [55] F. W. Alt, E. M. Oltz, F. Young, J. Gorman, G. Taccioli and J. Chen, *Immunol Today* **1992**, *13*, 306-314.
- [56] M. M. Davis, J. J. Boniface, Z. Reich, D. Lyons, J. Hampl, B. Arden and Y. Chien, *Annu Rev Immunol* **1998**, *16*, 523-544.
- [57] T. P. Arstila, A. Casrouge, V. Baron, J. Even, J. Kanellopoulos and P. Kourilsky, *Science* **1999**, *286*, 958-961.
- [58] T. P. Janeway CA Jr, Walport M, et al., *Immunobiology: The Immune System in Health and Disease. 5th edition. New York: Garland Science; 2001.*
- [59] R. Förster, A. Schubel, D. Breitfeld, E. Kremmer, I. Renner-Müller, E. Wolf and M. Lipp, *Cell* **1999**, *99*, 23-33.
- [60] R. M. Steinman, *Annual Review of Immunology* **1991**, *9*, 271-296.
- [61] P. Parham, D. A. Lawlor, C. E. Lomen and P. D. Ennis, *The Journal of Immunology* **1989**, *142*, 3937.
- [62] P. Parham and T. Ohta, *Science* **1996**, *272*, 67.
- [63] P. J. Bjorkman, M. A. Saper, B. Samraoui, W. S. Bennett, J. L. Strominger and D. C. Wiley, *Nature* **1987**, *329*, 506-512.
- [64] E. Y. Jones, *Current Opinion in Immunology* **1997**, *9*, 75-79.

- [65] in *Chapter 6 - The Major Histocompatibility Complex*, Eds.: T. W. Mak, M. E. Saunders and B. D. Jett), Academic Cell, Boston, **2014**, pp. 143-159.
- [66] M. Bouvier, *Mol Immunol* **2003**, *39*, 697-706.
- [67] T. Wenzel, C. Eckerskorn, F. Lottspeich and W. Baumeister, *FEBS Letters* **1994**, *349*, 205-209.
- [68] L. Saveanu, D. Fruci and P. M. van Endert, *Molecular Immunology* **2002**, *39*, 203-215.
- [69] M. Bouvier and D. C. Wiley, *Science* **1994**, *265*, 398-402.
- [70] F. E. Tynan, S. R. Burrows, A. M. Buckle, C. S. Clements, N. A. Borg, J. J. Miles, T. Beddoe, J. C. Whisstock, M. C. Wilce, S. L. Silins, J. M. Burrows, L. Kjer-Nielsen, L. Kostenko, A. W. Purcell, J. McCluskey and J. Rossjohn, *Nature Immunology* **2005**, *6*, 1114-1122.
- [71] M. Probst-Kepper, H.-J. Hecht, H. Herrmann, V. Janke, F. Ocklenburg, J. Klempnauer, B. J. van den Eynde and S. Weiss, *The Journal of Immunology* **2004**, *173*, 5610.
- [72] V. H. Engelhard, *Current Opinion in Immunology* **1994**, *6*, 13-23.
- [73] K. C. Parker, M. A. Bednarek, L. K. Hull, U. Utz, B. Cunningham, H. J. Zweerink, W. E. Biddison and J. E. Coligan, *The Journal of Immunology* **1992**, *149*, 3580.
- [74] L. Kjer-Nielsen, C. S. Clements, A. W. Purcell, A. G. Brooks, J. C. Whisstock, S. R. Burrows, J. McCluskey and J. Rossjohn, *Immunity* **2003**, *18*, 53-64.
- [75] A. K. Sewell, *Nature Reviews Immunology* **2012**, *12*, 669-677.
- [76] B. A. Lie and E. Thorsby, *Current Opinion in Immunology* **2005**, *17*, 526-531.
- [77] W. M. Kast, A. M. Bronkhorst, L. P. de Waal and C. J. Melief, *Journal of Experimental Medicine* **1986**, *164*, 723-738.
- [78] J. Trowsdale, *Immunology Letters* **2011**, *137*, 1-8.
- [79] G. The Multiple Sclerosis Genetics, J. L. Haines, H. A. Terwedow, K. Burgess, M. A. Pericak-Vance, J. B. Rimmler, E. R. Martin, J. R. Oksenberg, R. Lincoln, D. Y. Zhang, D. R. Banatao, N. Gatto, D. E. Goodkin and S. L. Hauser, *Human Molecular Genetics* **1998**, *7*, 1229-1234.
- [80] D. A. Hafler, A. Compston, S. Sawcer, E. S. Lander, M. J. Daly, P. L. De Jager, P. I. de Bakker, S. B. Gabriel, D. B. Mirel, A. J. Ivinson, M. A. Pericak-Vance, S. G. Gregory, J. D. Rioux, J. L. McCauley, J. L. Haines, L. F. Barcellos, B. Cree, J. R. Oksenberg and S. L. Hauser, *N Engl J Med* **2007**, *357*, 851-862.
- [81] A. Kempainen, S. Sawcer and A. Compston, *Brief Funct Genomics* **2011**, *10*, 61-70.
- [82] C. Jersild, A. Svejgaard and T. Fog, *The Lancet* **1972**, *299*, 1240-1241.
- [83] L. F. Barcellos, J. R. Oksenberg, A. B. Begovich, E. R. Martin, S. Schmidt, E. Vittinghoff, D. S. Goodin, D. Pelletier, R. R. Lincoln, P. Bucher, A. Swerdlin, M. A. Pericak-Vance, J. L. Haines and S. L. Hauser, *The American Journal of Human Genetics* **2003**, *72*, 710-716.

- [84] T. W. Yeo, P. L. De Jager, S. G. Gregory, L. F. Barcellos, A. Walton, A. Goris, C. Fenoglio, M. Ban, C. J. Taylor, R. S. Goodman, E. Walsh, C. S. Wolfish, R. Horton, J. Traherne, S. Beck, J. Trowsdale, S. J. Caillier, A. J. Ivinson, T. Green, S. Pobywajlo, E. S. Lander, M. A. Pericak-Vance, J. L. Haines, M. J. Daly, J. R. Oksenberg, S. L. Hauser, A. Compston, D. A. Hafler, J. D. Rioux and S. Sawcer, *Annals of Neurology* **2007**, *61*, 228-236.
- [85] M. A. Friese, K. B. Jakobsen, L. Friis, R. Etzensperger, M. J. Craner, R. M. McMahon, L. T. Jensen, V. Huygelen, E. Y. Jones, J. I. Bell and L. Fugger, *Nature Medicine* **2008**, *14*, 1227-1235.
- [86] A. Fogdell-Hahn, A. Ligiers, M. Grønning, J. Hillert and O. Olerup, *Tissue Antigens* **2000**, *55*, 140-148.
- [87] E. D. Johnson and G. A. Cole, *Journal of Experimental Medicine* **1975**, *141*, 866-881.
- [88] E. J. Wherry, V. Teichgräber, T. C. Becker, D. Masopust, S. M. Kaech, R. Antia, U. H. von Andrian and R. Ahmed, *Nature Immunology* **2003**, *4*, 225-234.
- [89] T. Sathaliyawala, M. Kubota, N. Yudanin, D. Turner, P. Camp, Joseph J. C. Thome, Kara L. Bickham, H. Lerner, M. Goldstein, M. Sykes, T. Kato and Donna L. Farber, *Immunity* **2013**, *38*, 187-197.
- [90] L. M. Wakim, A. Woodward-Davis and M. J. Bevan, *Proceedings of the National Academy of Sciences* **2010**, *107*, 17872.
- [91] B. S. Sheridan and L. Lefrançois, *Nature Immunology* **2011**, *12*, 485-491.
- [92] M. Matloubian, C. G. Lo, G. Cinamon, M. J. Lesneski, Y. Xu, V. Brinkmann, M. L. Allende, R. L. Proia and J. G. Cyster, *Nature* **2004**, *427*, 355-360.
- [93] C. N. Skon, J.-Y. Lee, K. G. Anderson, D. Masopust, K. A. Hogquist and S. C. Jameson, *Nature Immunology* **2013**, *14*, 1285-1293.
- [94] D. Cibrián and F. Sánchez-Madrid, *European Journal of Immunology* **2017**, *47*, 946-953.
- [95] L. K. Mackay, M. Minnich, N. A. M. Kragten, Y. Liao, B. Nota, C. Seillet, A. Zaid, K. Man, S. Preston, D. Freestone, A. Braun, E. Wynne-Jones, F. M. Behr, R. Stark, D. G. Pellicci, D. I. Godfrey, G. T. Belz, M. Pellegrini, T. Gebhardt, M. Busslinger, W. Shi, F. R. Carbone, R. A. W. van Lier, A. Kallies and K. P. J. M. van Gisbergen, *Science* **2016**, *352*, 459.
- [96] L. K. Mackay, A. Braun, B. L. Macleod, N. Collins, C. Tebartz, S. Bedoui, F. R. Carbone and T. Gebhardt, *The Journal of Immunology* **2015**, *194*, 2059.
- [97] B. V. Kumar, W. Ma, M. Miron, T. Granot, R. S. Guyer, D. J. Carpenter, T. Senda, X. Sun, S.-H. Ho, H. Lerner, A. L. Friedman, Y. Shen and D. L. Farber, *Cell Reports* **2017**, *20*, 2921-2934.
- [98] C. O. Park and T. S. Kupper, *Nature Medicine* **2015**, *21*, 688-697.
- [99] R. A. Clark, *Science translational medicine* **2015**, *7*, 269rv261-269rv261.
- [100] G. L. Montgomery, *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences* **1957**, *42*, 322-323.

- [101] C. A. Janeway, *Immunology Today* **1992**, *13*, 11-16.
- [102] J. W. Kappler, N. Roehm and P. Marrack, *Cell* **1987**, *49*, 273-280.
- [103] C. C. Goodnow, J. Sprent, B. Fazekas de St Groth and C. G. Vinuesa, *Nature* **2005**, *435*, 590-597.
- [104] P. Marrack, J. Kappler and B. L. Kotzin, *Nat Med* **2001**, *7*, 899-905.
- [105] L. A. Zenewicz, C. Abraham, R. A. Flavell and J. H. Cho, *Cell* **2010**, *140*, 791-797.
- [106] R. Root-Bernstein and D. Fairweather, *Current Allergy and Asthma Reports* **2013**, *14*, 407.
- [107] J.-M. Charcot, *Histologie de la sclérose en plaques*, [s.n.], Paris, **1868**, p.
- [108] A. D. Sadovnick and G. Ebers, *Canadian Journal of Neurological Sciences* **1993**, *20*, 17-29.
- [109] B. Ferguson, M. K. Matyszak, M. M. Esiri and V. H. Perry, *Brain* **1997**, *120* ( Pt 3), 393-399.
- [110] A. Kutzelnigg, C. F. Lucchinetti, C. Stadelmann, W. Brück, H. Rauschka, M. Bergmann, M. Schmidbauer, J. E. Parisi and H. Lassmann, *Brain* **2005**, *128*, 2705-2712.
- [111] D. T. Okuda, E. M. Mowry, A. Beheshtian, E. Waubant, S. E. Baranzini, D. S. Goodin, S. L. Hauser and D. Pelletier, *Neurology* **2009**, *72*, 800.
- [112] A. Compston, H. Lassmann and I. McDonald, *McAlpine's Multiple Sclerosis*. London: Churchill Livingstone **2006**, 3-68.
- [113] R. Milo and A. Miller, *Autoimmunity Reviews* **2014**, *13*, 518-524.
- [114] R. A. Marrie, *Lancet Neurol* **2004**, *3*, 709-718.
- [115] A. K. Artemiadis, M. C. Anagnostouli and E. C. Alexopoulos, *Neuroepidemiology* **2011**, *36*, 109-120.
- [116] G. C. Ebers, A. D. Sadovnick and N. J. Risch, *Nature* **1995**, *377*, 150-151.
- [117] K. W. Wucherpfennig and J. L. Strominger, *Cell* **1995**, *80*, 695-705.
- [118] K. Berer, L. A. Gerdes, E. Cekanaviciute, X. Jia, L. Xiao, Z. Xia, C. Liu, L. Klotz, U. Stauffer, S. E. Baranzini, T. Kümpfel, R. Hohlfeld, G. Krishnamoorthy and H. Wekerle, *Proceedings of the National Academy of Sciences* **2017**, *114*, 10719.
- [119] N. Y. Souren, L. A. Gerdes, P. Lutsik, G. Gasparoni, E. Beltrán, A. Salhab, T. Kümpfel, D. Weichenhan, C. Plass, R. Hohlfeld and J. Walter, *Nat Commun* **2019**, *10*, 2094.
- [120] R. Ehling, T. Berger and M. Reindl, *Cent Nerv Syst Agents Med Chem* **2010**, *10*, 3-15.
- [121] M. J. Carson, J. M. Doose, B. Melchior, C. D. Schmid and C. C. Ploix, *Immunological reviews* **2006**, *213*, 48-65.

- [122] U. Traugott, E. L. Reinherz and C. S. Raine, *Science* **1983**, *219*, 308.
- [123] S. L. Hauser, A. K. Bhan, F. Gilles, M. Kemp, C. Kerr and H. L. Weiner, *Annals of Neurology* **1986**, *19*, 578-587.
- [124] C. S. Constantinescu, N. Farooqi, K. O'Brien and B. Gran, *British journal of pharmacology* **2011**, *164*, 1079-1106.
- [125] M. A. Friese and L. Fugger, *Ann Neurol* **2009**, *66*, 132-141.
- [126] J. Booss, M. M. Esiri, W. W. Tourtellotte and D. Y. Mason, *J Neurol Sci* **1983**, *62*, 219-232.
- [127] H. Babbe, A. Roers, A. Waisman, H. Lassmann, N. Goebels, R. Hohlfeld, M. Friese, R. Schröder, M. Deckert, S. Schmidt, R. Ravid and K. Rajewsky, *The Journal of experimental medicine* **2000**, *192*, 393-404.
- [128] C. Skulina, S. Schmidt, K. Dornmair, H. Babbe, A. Roers, K. Rajewsky, H. Wekerle, R. Hohlfeld and N. Goebels, *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 2428-2433.
- [129] J. W. Prineas and R. G. Wright, *Lab Invest* **1978**, *38*, 409-421.
- [130] A. S. Bellamy, V. L. Calder, M. Feldmann and A. N. Davison, *Clinical and experimental immunology* **1985**, *61*, 248-256.
- [131] A. Awad, B. Hemmer, H. P. Hartung, B. Kieseier, J. L. Bennett and O. Stuve, *J Neuroimmunol* **2010**, *219*, 1-7.
- [132] E. Beltrán, B. Obermeier, M. Moser, F. Coret, M. Simó-Castelló, I. Boscá, F. Pérez-Miralles, L. M. Villar, M. Senel, H. Tumani, R. Hohlfeld, B. Casanova and K. Dornmair, *Brain* **2014**, *137*, 2703-2714.
- [133] S. M. Brändle, B. Obermeier, M. Senel, J. Bruder, R. Mentele, M. Khademi, T. Olsson, H. Tumani, W. Kristoferitsch, F. Lottspeich, H. Wekerle, R. Hohlfeld and K. Dornmair, *Proceedings of the National Academy of Sciences* **2016**, 201522730.
- [134] A. Lossius, J. N. Johansen, F. Vartdal and T. Holmøy, *Annals of Clinical and Translational Neurology* **2016**, *3*, 295-306.
- [135] D. J. Woodsworth, M. Castellarin and R. A. Holt, *Genome Medicine* **2013**, *5*, 98-98.
- [136] T. P. Arstila, A. Casrouge, V. Baron, J. Even, J. Kanellopoulos and P. Kourilsky, *Science* **2000**, *288*, 1135.
- [137] M. Malissen, J. Trucy, E. Jouvin-Marche, P.-A. Cazenave, R. Scollay and B. Malissen, *Immunology Today* **1992**, *13*, 315-322.
- [138] A. Six, M. E. Mariotti-Ferrandiz, W. Chacara, S. Magadan, H.-P. Pham, M.-P. Lefranc, T. Mora, V. Thomas-Vaslin, A. M. Walczak and P. Boudinot, *Frontiers in Immunology* **2013**, *4*, 413.

- [139] J. M. Faint, D. Pilling, A. N. Akbar, G. D. Kitas, P. A. Bacon and M. Salmon, *Journal of Immunological Methods* **1999**, *225*, 53-60.
- [140] C. Fozza, F. Barraqueddu, G. Corda, S. Contini, P. Viridis, F. Dore, S. Bonfigli and M. Longinotti, *J Immunol Methods* **2017**, *440*, 1-11.
- [141] F. Sanger and A. R. Coulson, *Journal of Molecular Biology* **1975**, *94*, 441-448.
- [142] A. Junker, J. Ivanidze, J. Malotka, I. Eiglmeier, H. Lassmann, H. Wekerle, E. Meinl, R. Hohlfeld and K. Dornmair, *Brain* **2007**, *130*, 2789-2799.
- [143] R. Satija and A. K. Shalek, *Trends in Immunology* **2014**, *35*, 219-229.
- [144] A. E. Saliba, A. J. Westermann, S. A. Gorski and J. Vogel, *Nucleic Acids Res* **2014**, *42*, 8845-8860.
- [145] J. Benichou, R. Ben-Hamo, Y. Louzoun and S. Efroni, *Immunology* **2012**, *135*, 183-191.
- [146] M. A. Frohman in [24] *Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: Thermal race, Vol. Volume 218* (Ed. W. Ray), Academic Press, **1993**, pp. 340-356.
- [147] S. Picelli, O. R. Faridani, Å. K. Björklund, G. Winberg, S. Sagasser and R. Sandberg, *Nat. Protocols* **2014**, *9*, 171-181.
- [148] S.-M. Kim, L. Bhonsle, P. Besgen, J. Nickel, A. Backes, K. Held, S. Vollmer, K. Dornmair and J. C. Prinz, *PLoS ONE* **2012**, *7*, e37338.
- [149] E. Beltran, B. Obermeier, M. Moser, F. Coret, M. Simo-Castello, I. Bosca, F. Perez-Miralles, L. M. Villar, M. Senel, H. Tumani, R. Hohlfeld, B. Casanova and K. Dornmair, *Brain* **2014**, *137*, 2703-2714.
- [150] L. A. Gerdes, K. Held, E. Beltrán, C. Berking, J. C. Prinz, A. Junker, J. K. Tietze, B. Ertl-Wagner, A. Straube, T. Kümpfel, K. Dornmair and R. Hohlfeld, *Annals of Neurology* **2016**, *80*, 294-300.
- [151] K. Held, L. Bhonsle-Deeng, K. Siewert, W. Sato, E. Beltran, S. Schmidt, G. Ruhl, J. K. Ng, P. Engerer, M. Moser, W. E. Klinkert, H. Babbe, T. Misgeld, H. Wekerle, D. A. Laplaud, R. Hohlfeld and K. Dornmair, *Neurol Neuroimmunol Neuroinflamm* **2015**, *2*, e107.
- [152] D. Lee-Liu, L. I. Almonacid, F. Faunes, F. Melo and J. Larrain, *Methods Mol Biol* **2012**, *917*, 293-317.
- [153] F. Sanger, S. Nicklen and A. R. Coulson, *Proc Natl Acad Sci U S A* **1977**, *74*, 5463-5467.
- [154] G. Sharma and R. A. Holt, *Hum Immunol* **2014**, *75*, 514-519.
- [155] R. Hohlfeld, K. Dornmair, E. Meinl and H. Wekerle, *Lancet Neurol* **2016**, *15*, 317-331.
- [156] C. X. Chang, L. Dai, Z. W. Tan, J. A. Choo, A. Bertoletti and G. M. Grotenbreg, *Front Biosci (Landmark Ed)* **2011**, *16*, 3014-3035.
- [157] K. Matsui, J. J. Boniface, P. A. Reay, H. Schild, B. Fazekas de St Groth and M. M. Davis, *Science* **1991**, *254*, 1788.

- [158] J. D. Stone, A. S. Chervin and D. M. Kranz, *Immunology* **2009**, *126*, 165-176.
- [159] G. P. Morris and P. M. Allen, *Nature Immunology* **2012**, *13*, 121-128.
- [160] R. N. Germain, *Cell* **1994**, *76*, 287-299.
- [161] J. J. Adams, S. Narayanan, B. Liu, M. E. Birnbaum, A. C. Kruse, N. A. Bowerman, W. Chen, A. M. Levin, J. M. Connolly, C. Zhu, D. M. Kranz and K. C. Garcia, *Immunity* **2011**, *35*, 681-693.
- [162] P. van der Bruggen, C. Traversari, P. Chomez, C. Lurquin, E. De Plaen, B. Van den Eynde, A. Knuth and T. Boon, *Science* **1991**, *254*, 1643-1647.
- [163] C. La Rosa, R. Krishnan, S. Markel, J. P. Schneck, R. Houghten, C. Pinilla and D. J. Diamond, *Blood* **2001**, *97*, 1776-1786.
- [164] B. R. Gundlach, K.-H. Wiesmüller, T. Junt, S. Kienle, G. Jung and P. Walden, *Journal of Immunological Methods* **1996**, *192*, 149-155.
- [165] M. Sospedra, C. Pinilla and R. Martin, *Methods* **2003**, *29*, 236-247.
- [166] Y. Wang, A. Rubtsov, R. Heiser, J. White, F. Crawford, P. Marrack and J. W. Kappler, *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 2476.
- [167] K. Siewert, J. Malotka, N. Kawakami, H. Wekerle, R. Hohlfeld and K. Dornmair, *Nature Medicine* **2012**, *18*, 824-828.
- [168] G. Rühl, A. G. Niedl, A. Patronov, K. Siewert, S. Pinkert, M. Kalemanov, M. A. Friese, K. E. Attfield, I. Antes, R. Hohlfeld and K. Dornmair, *Neurology - Neuroimmunology Neuroinflammation* **2016**, *3*, e241.
- [169] A. Arakawa, K. Siewert, J. Stöhr, P. Besgen, S.-M. Kim, G. Rühl, J. Nickel, S. Vollmer, P. Thomas, S. Krebs, S. Pinkert, M. Spannagl, K. Held, C. Kammerbauer, R. Besch, K. Dornmair and J. C. Prinz, *Journal of Experimental Medicine* **2015**, *212*, 2203-2212.
- [170] W. H. Levering, H. Wind, K. Sintnicolaas, H. Hooijkaas and J. W. Gratama, *Cytometry B Clin Cytom* **2003**, *54*, 28-38.
- [171] J. Hansen, *LMU Munich* **2016**.
- [172] A. Cornish-Bowden, *Nucleic Acids Res* **1985**, *13*, 3021-3030.
- [173] M. Petersen and J. Wengel, *Trends Biotechnol* **2003**, *21*, 74-81.
- [174] Y. Y. Zhu, E. M. Machleder, A. Chenchik, R. Li and P. D. Siebert, *Biotechniques* **2001**, *30*, 892-897.
- [175] S. Seitz, C. K. Schneider, J. Malotka, X. Nong, A. G. Engel, H. Wekerle, R. Hohlfeld and K. Dornmair, *Proc Natl Acad Sci U S A* **2006**, *103*, 12057-12062.
- [176] E. Beltrán, L. A. Gerdes, J. Hansen, A. Flierl-Hecht, S. Krebs, H. Blum, B. Ertl-Wagner, F. Barkhof, T. Kümpfel, R. Hohlfeld and K. Dornmair, *The Journal of Clinical Investigation* **2019**, *129*, 4758-4768.

- [177] M. Baym, S. Kryazhimskiy, T. D. Lieberman, H. Chung, M. M. Desai and R. Kishony, *PLOS ONE* **2015**, *10*, e0128036.
- [178] M. M. DeAngelis, D. G. Wang and T. L. Hawkins, *Nucleic Acids Research* **1995**, *23*, 4742-4743.
- [179] S. Andrews, **2010**.
- [180] D. Kim, B. Langmead and S. L. Salzberg, *Nature Methods* **2015**, *12*, 357-360.
- [181] Y. Liao, G. K. Smyth and W. Shi, *Bioinformatics* **2014**, *30*, 923-930.
- [182] M. Shugay, O. V. Britanova, E. M. Merzlyak, M. A. Turchaninova, I. Z. Mamedov, T. R. Tuganbaev, D. A. Bolotin, D. B. Staroverov, E. V. Putintseva, K. Plevova, C. Linnemann, D. Shagin, S. Pospisilova, S. Lukyanov, T. N. Schumacher and D. M. Chudakov, *Nature Methods* **2014**, *11*, 653-655.
- [183] B. Zhou and W. Jin, *Methods Mol Biol* **2020**, *2117*, 159-167.
- [184] X. Brochet, M.-P. Lefranc and V. Giudicelli, *Nucleic Acids Research* **2008**, *36*, W503-W508.
- [185] S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, *Nucleic Acids Research* **1997**, *25*, 3389-3402.
- [186] O. M. Subach, P. J. Cranfill, M. W. Davidson and V. V. Verkhusha, *PLoS One* **2011**, *6*, e28674.
- [187] V. Y. Kiselev, K. Kirschner, M. T. Schaub, T. Andrews, A. Yiu, T. Chandra, K. N. Natarajan, W. Reik, M. Barahona, A. R. Green and M. Hemberg, *Nat Methods* **2017**, *14*, 483-486.
- [188] A. Bar-Or, A. Pachner, F. Menguy-Vacheron, J. Kaplan and H. Wiendl, *Drugs* **2014**, *74*, 659-674.
- [189] E. Padovan, G. Casorati, P. Dellabona, S. Meyer, M. Brockhaus and A. Lanzavecchia, *Science* **1993**, *262*, 422-424.
- [190] G. Rühl, *Dissertation, LMU München* **2016**.
- [191] A.-C. Villani, R. Satija, G. Reynolds, S. Sarkizova, K. Shekhar, J. Fletcher, M. Griesbeck, A. Butler, S. Zheng, S. Lazo, L. Jardine, D. Dixon, E. Stephenson, E. Nilsson, I. Grundberg, D. McDonald, A. Filby, W. Li, P. L. De Jager, O. Rozenblatt-Rosen, A. A. Lane, M. Haniffa, A. Regev and N. Hacohen, *Science* **2017**, *356*, eaah4573.
- [192] D. Grün and A. van Oudenaarden, *Cell* **2015**, *163*, 799-810.
- [193] A. A. Eltahla, S. Rizzetto, M. R. Pirozyan, B. D. Betz-Stablein, V. Venturi, K. Kedzierska, A. R. Lloyd, R. A. Bull and F. Luciani, *Immunol Cell Biol* **2016**, *94*, 604-611.
- [194] S. Han, Y. C. Lin, T. Wu, A. D. Salgado, I. Mexhitaj, S. C. Wuest, E. Romm, J. Ohayon, R. Goldbach-Mansky, A. Vanderver, A. Marques, C. Toro, P. Williamson, I. Cortese and B. Bielekova, *The Journal of Immunology* **2014**, *192*, 2551.

- [195] N. De Stefano, A. Giorgio, M. Tintoré, M. Pia Amato, L. Kappos, J. Palace, T. Yousry, M. A. Rocca, O. Ciccarelli, C. Enzinger, J. Frederiksen, M. Filippi, H. Vrenken and À. Rovira, *Mult Scler* **2018**, *24*, 214-221.
- [196] A. J. Thompson, B. L. Banwell, F. Barkhof, W. M. Carroll, T. Coetzee, G. Comi, J. Correale, F. Fazekas, M. Filippi, M. S. Freedman, K. Fujihara, S. L. Galetta, H. P. Hartung, L. Kappos, F. D. Lublin, R. A. Marrie, A. E. Miller, D. H. Miller, X. Montalban, E. M. Mowry, P. S. Sorensen, M. Tintoré, A. L. Traboulsee, M. Trojano, B. M. J. Uitdehaag, S. Vukusic, E. Waubant, B. G. Weinshenker, S. C. Reingold and J. A. Cohen, *Lancet Neurol* **2018**, *17*, 162-173.
- [197] H. Lassmann, *Front Immunol* **2018**, *9*, 3116.
- [198] C. Lebrun, O. H. Kantarci, A. Siva, D. Pelletier and D. T. Okuda, *Neurol Clin* **2018**, *36*, 59-68.
- [199] B. Obermeier, R. Mentele, J. Malotka, J. Kellermann, T. Kumpfel, H. Wekerle, F. Lottspeich, R. Hohlfeld and K. Dornmair, *Nat Med* **2008**, *14*, 688-693.
- [200] S. Cepok, B. Rosche, V. Grummel, F. Vogel, D. Zhou, J. Sayn, N. Sommer, H.-P. Hartung and B. Hemmer, *Brain* **2005**, *128*, 1667-1676.
- [201] L. Michel, H. Touil, N. B. Pikor, J. L. Gommerman, A. Prat and A. Bar-Or, *Frontiers in immunology* **2015**, *6*, 636-636.
- [202] A. Palanichamy, L. Apeltsin, T. C. Kuo, M. Sirota, S. Wang, S. J. Pitts, P. D. Sundar, D. Telman, L. Z. Zhao, M. Derstine, A. Abounasr, S. L. Hauser and H. C. von Büdingen, *Sci Transl Med* **2014**, *6*, 248ra106.
- [203] J. N. Stern, G. Yaari, J. A. Vander Heiden, G. Church, W. F. Donahue, R. Q. Hintzen, A. J. Huttner, J. D. Laman, R. M. Nagra, A. Nylander, D. Pitt, S. Ramanan, B. A. Siddiqui, F. Vigneault, S. H. Kleinstein, D. A. Hafler and K. C. O'Connor, *Sci Transl Med* **2014**, *6*, 248ra107.
- [204] K. E. Neu, Q. Tang, P. C. Wilson and A. A. Khan, *Trends Immunol* **2017**, *38*, 140-149.
- [205] A. Nguyen, W. H. Khoo, I. Moran, P. I. Croucher and T. G. Phan, *Front Immunol* **2018**, *9*, 1553.
- [206] R. M. Ransohoff and B. Engelhardt, *Nature Reviews Immunology* **2012**, *12*, 623-635.
- [207] J. Smolders, K. M. Heutinck, N. L. Fransen, E. B. M. Remmerswaal, P. Hombrink, I. J. M. ten Berge, R. A. W. van Lier, I. Huitinga and J. Hamann, *Nature Communications* **2018**, *9*, 4593.
- [208] C. Günther, N. Carballido-Perrig, S. Kaesler, J. M. Carballido and T. Biedermann, *J Invest Dermatol* **2012**, *132*, 626-634.
- [209] T. Sato, H. Thorlacios, B. Johnston, T. L. Staton, W. Xiang, D. R. Littman and E. C. Butcher, *The Journal of Immunology* **2005**, *174*, 277.
- [210] R. van der Voort, A. W. van Lieshout, L. W. Toonen, A. W. Slötjes, W. B. van den Berg, C. G. Figdor, T. R. Radstake and G. J. Adema, *Arthritis Rheum* **2005**, *52*, 1381-1391.

- [211] M. Pashenkov, Y. M. Huang, V. Kostulas, M. Haglund, M. Söderström and H. Link, *Brain* **2001**, *124*, 480-492.
- [212] M. T. de Graaf, P. A. Smitt, R. L. Luitwieler, C. van Velzen, P. D. van den Broek, J. Kraan and J. W. Gratama, *Cytometry B Clin Cytom* **2011**, *80*, 43-50.
- [213] M. Konjevic Sabolek, K. Held, E. Beltrán, A. G. Niedl, E. Meinl, R. Hohlfeld, H. Lassmann and K. Dornmair, *Annals of Clinical and Translational Neurology* **2019**, *6*, 1151-1164.
- [214] P. S. Rommer, R. Milo, M. H. Han, S. Satyanarayan, J. Sellner, L. Hauer, Z. Illes, C. Warnke, S. Laurent, M. S. Weber, Y. Zhang and O. Stuve, *Frontiers in immunology* **2019**, *10*, 1564-1564.
- [215] A. E. Coutinho and K. E. Chapman, *Molecular and cellular endocrinology* **2011**, *335*, 2-13.
- [216] A. Molodtsov and M. J. Turk, *Front Immunol* **2018**, *9*, 2810.
- [217] C. O. Park and T. S. Kupper, *Nat Med* **2015**, *21*, 688-697.
- [218] D. Amsen, K. P. J. M. van Gisbergen, P. Hombrink and R. A. W. van Lier, *Nature Immunology* **2018**, *19*, 538-546.
- [219] J. Machado-Santos, E. Saji, A. R. Tröscher, M. Paunovic, R. Liblau, G. Gabriely, C. G. Bien, J. Bauer and H. Lassmann, *Brain* **2018**, *141*, 2066-2082.
- [220] G. P. van Nierop, M. M. van Luijn, S. S. Michels, M. J. Melief, M. Janssen, A. W. Langerak, W. J. D. Ouwendijk, R. Q. Hintzen and G. Verjans, *Acta Neuropathol* **2017**, *134*, 383-401.
- [221] M. Salou, A. Garcia, L. Michel, A. Gainche-Salmon, D. Loussouarn, B. Nicol, F. Guillot, P. Hulin, S. Nedellec, D. Baron, G. Ramstein, J. P. Soulillou, S. Brouard, A. B. Nicot, N. Degauque and D. A. Laplaud, *Ann Clin Transl Neurol* **2015**, *2*, 609-622.
- [222] K. Christopher Garcia, J. J. Adams, D. Feng and L. K. Ely, *Nature Immunology* **2009**, *10*, 143-147.
- [223] N. L. La Gruta, S. Gras, S. R. Daley, P. G. Thomas and J. Rossjohn, *Nature Reviews Immunology* **2018**, *18*, 467-478.
- [224] I. V. Zvyagin, M. V. Pogorelyy, M. E. Ivanova, E. A. Komech, M. Shugay, D. A. Bolotin, A. A. Shelenkov, A. A. Kurnosov, D. B. Staroverov, D. M. Chudakov, Y. B. Lebedev and I. Z. Mamedov, *Proceedings of the National Academy of Sciences* **2014**, *111*, 5980-5985.
- [225] Q. Jiang, Y. Liu, B. Xu, W. Zheng, X. Xiang, X. Tang, H. Dong, Y. Chen, C. Wang, G. Deng, Q. Mao, X. Shang and Y. Wu, *Biochem Biophys Res Commun* **2018**, *497*, 153-159.
- [226] G. M. Edelman and J. A. Gally, *Proceedings of the National Academy of Sciences* **2001**, *98*, 13763.
- [227] H. H. Yiu, L. N. Schoettle, M. Garcia-Neuer, J. N. Blattman and P. L. F. Johnson, *Immunology* **2021**, *162*, 464-475.

- [228] J. I. Elliott and D. M. Altmann, *The Journal of experimental medicine* **1995**, *182*, 953-959.
- [229] R. Planas, I. Metz, R. Martin and M. Sospedra, *Frontiers in immunology* **2018**, *9*, 509-509.
- [230] S. Nozuma, Y. Enose-Akahata, K. R. Johnson, M. C. Monaco, N. Ngouth, A. Elkahloun, J. Ohayon, J. Zhu and S. Jacobson, *JCI Insight* **2021**, *6*.
- [231] V. Venturi, D. A. Price, D. C. Douek and M. P. Davenport, *Nature Reviews Immunology* **2008**, *8*, 231-238.
- [232] S. Dandekar, H. Wijesuriya, T. Geiger, D. Hamm, G. W. Mathern and G. C. Owens, *Frontiers in immunology* **2016**, *7*, 608-608.
- [233] X. Hou, M. Wang, C. Lu, Q. Xie, G. Cui, J. Chen, Y. Du, Y. Dai and H. Diao, *Cellular Physiology and Biochemistry* **2016**, *39*, 651-667.
- [234] B. Serafini, B. Rosicarelli, C. Veroni, G. A. Mazzola and F. Aloisi, *J Virol* **2019**, *93*.
- [235] D. F. Angelini, B. Serafini, E. Piras, M. Severa, E. M. Coccia, B. Rosicarelli, S. Ruggieri, C. Gasperini, F. Buttari, D. Centonze, R. Mechelli, M. Salvetti, G. Borsellino, F. Aloisi and L. Battistini, *PLoS Pathog* **2013**, *9*, e1003220.
- [236] A. Lossius, J. N. Johansen, F. Vartdal, H. Robins, B. Jūratė Šaltytė, T. Holmøy and J. Olweus, *European Journal of Immunology* **2014**, *44*, 3439-3452.
- [237] U. Hershberg and E. T. Luning Prak, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **2015**, *370*, 20140239.
- [238] H. M. Lorenz, M. Herrmann and J. R. Kalden, *Scand J Clin Lab Invest Suppl* **2001**, *235*, 16-26.
- [239] J. Bruder, K. Siewert, B. Obermeier, J. Malotka, P. Scheinert, J. Kellermann, T. Ueda, R. Hohlfeld and K. Dornmair, *J Biol Chem* **2012**, *287*, 20986-20995.
- [240] A. Arakawa, K. Siewert, J. Stöhr, P. Besgen, S. M. Kim, G. Rühl, J. Nickel, S. Vollmer, P. Thomas, S. Krebs, S. Pinkert, M. Spannagl, K. Held, C. Kammerbauer, R. Besch, K. Dornmair and J. C. Prinz, *J Exp Med* **2015**, *212*, 2203-2212.
- [241] M. Rasmussen, M. Harndahl, A. Stryhn, R. Boucherma, L. L. Nielsen, F. A. Lemonnier, M. Nielsen and S. Buus, *The Journal of Immunology* **2014**, 1401689.
- [242] B. Wiltschi, L. Merkel and N. Budisa, *Chembiochem* **2009**, *10*, 217-220.

## 7 Appendix

### 7.1 Methodical part: Nucleotide sequences for cloning reactions

Nucleotide sequence were listed in 5'→3' direction and start codons (ATG) as well as stop codons (TGA) of plasmid constructs marked by bold letters. Important sequence sections such as restriction sites or primer annealing sites were coloured and/or underlined as indicated. Used plasmids, primers and restriction enzymes were listed in Table 5-6. TCR $\alpha$  and  $\beta$  chain nomenclature was used as defined by IMGT.

#### 7.1.1 TRAV17\*01 TRAJ23\*01

>cloned into pRSV5hygro (Sall/ PvuII restr. digest)

Sall, PvuII, BamHI

TRAPPIST-161 TRA-for, C $\alpha$  rev IN

CTTGGGCTGCAGGTCGACATGGAAACTCTCCTGGGAGTGTCTTTGGTGATTCTATGGCTTC  
 AACTGGCTAGGGTGAACAGTCAACAGGGAGAAGAGGATCCTCAGGCCTTGAGCATCCAG  
 GAGGGTGA AAAATGCCACCATGAACTGCAGTTACAAAAGTATAAACAATTTACAGTG  
 GTATAGACAAAATTCAGGTAGAGGCCTTGTCACCTAATTTTAATACGTTCAAATGAAAG  
 AGAGAAACACAGTGGAAAGATTAAGAGTACGCTTGACACTTCCAAGAAAAGCAGTTCCT  
 TGTTGATCACGGCTTCCCGGGCAGCAGACACTGCTTCTTACTTCTGTGCTACGGACTACAA  
 CCAGGGAGGAAAGCTTATCTTCGGACAGGGAACGGAGTTATCTGGAAACCCAAATATCC  
 AGAACCCTGACCCTGCCGTGTACAGCTGAGAGACTCTAAATCCAGTGACAAGTCTGTC  
 TGCCTATTCACCGATTTTGATTCTCAAACAAATGTGTCCAAAGTAAGGATTCTGATGTGT  
 ATATCACAGACAAAAGTGTGCTAGACATGAGGTCTATGGACTTCAAGAGCAACAGTGCTG  
 TGGCCTGGAGCAACAAATCTGACTTTGTCATGTGCAAACGCCTTCAACAACAGCATTATTC  
 CAGAAGACACCTTCTTCCCCAGCCAGAAAAGTTCTGTGATGTCAAGCTGGTCCGAGAAAA  
 GCTTTGAAACAGATACGAACCTAAACTTTCAAACCTGTGAGTATTGGGTTCCGAATCC  
 TCCTCCTGAAAGTGGCCGGGTTAATCTGCTCATGACGCTGCGGCTGTGGTCCAGTTGAG  
 GATCCAGACATGATAAGATAATTGATGAGTTTGGACAAACCACAAGTGAATGCAGTG  
 AAAAAATGCTTTATTGTGAAATTTGTGATG

#### 7.1.2 TRBV4-1\*01

>cloned into pRSV5-neo (Sall/AvaI restr. digest)

Sall, AvaI, BamHI

TRAPPIST-161 TRB-for, C $\beta$  rev IN

CTTGGGCTGCAGGTCGACATGGGCTGCAGGCTGCTCTGCTGTGCGGTTCTCTGTCTCCTGG  
 GAGCAGTTCCCATAGACACTGAAGTTACCCAGACACAAAACACCTGGTCATGGGAATG  
 ACAAATAAGAAGTCTTTGAAATGTGAACAACATATGGGGCACAGGGCTATGTATTGGTAC  
 AAGCAGAAAAGCTAAGAAGCCACCGGAGCTCATGTTTGTCTACAGCTATGAGAACTCTCT  
 ATAAATGAAAGTGTGCCAAGTCGCTTCTCACCTGAATGCCCAACAGCTCTCTCTTAAAC

CTTCACCTACACGCCCTGCAGCCAGAAGACTCAGCCCTGTATCTCTGCGCCAGCAGCCAA  
 GATCCGGGGGCTGGGAGACCCAGTACTTCGGGCCAGGCACGCGGCTCCTGGTGCTCGA  
 GGACCTGAACAAGGTGTTCCCA<sup>CCCGAG</sup>GTC<sup>GCTGTGTTGAGCCATCAGA</sup>AAGCAGAG  
 ATCTCCCACACCCAAAAGGCCACACTGGTGTGCCTGGCCACAGGTATCTTCCCTGACCAC  
 GTGGAGCTGAGCTGGTGGGTGAATGGGAAGGAGGTGCACAGTGGGGTCAGCACGGACCC  
 GCAGCCCCTCAAGGAGCAGCCC GCCCTCAATGACTCCAGATACTGCCTGAGCAGCCGCT  
 GAGGGTCTCGGCCACCTTCTGGCAGAACCCCCGCAACCACTTCCGCTGTCAAGTCCAGTT  
 CTACGGGCTCTCGGAGAATGACGAGTGGACCCAGGATAGGGCCAAACCCGTCACCCAGA  
 TCGTCAGCGCCGAGGCCTGGGGTAGAGCAGACTGTGGCTTTACCTCGGTGTCTACCAGC  
 AAGGGGTCTGTCTGCCACCATCTCTATGAGATCCTGCTAGGGAAGGCCACCATGTATG  
 CTGTGCTGGTCAGCGCCCTTGTGTTGATGGCCATGGTCAAGAGAAAGGATTTCTGA<sup>GGAT</sup>  
<sup>CC</sup>AGACATGATAAGATA

### 7.1.3 TRDV\*01

>cloned into pRSV5hygro (Sall/ PvuII restr. digest)

<sup>Sall</sup>, <sup>PvuII</sup>, <sup>BamHI</sup>

TRAPPIST-161 aII-for, **Ca rev IN**

CTTGGGCTGCAG<sup>GTCGAC</sup>ATGCTGTTCTCCAGCCTGCTGTGTGTATTTGTGGCCTTCAGCT  
 ACTCTGGATCAAGTGTGGCCAGAAGGTTACTCAAGCCCAGTCATCAGTATCCATGCCAG  
 TGAGGAAAGCAGTCACCCTGAACTGCCTGTATGAAACAAGTTGGTGGTCATATTATATTT  
 TTTGGTACAAGCAACTTCCCAGCAAAGAGATGATTTTCCTTATTCGCCAGGGTTCTGATGA  
 ACAGAATGCAAAAAGTGGTCGCTATTCTGTCAACTTCAAGAAAGCAGCGAAATCCGTCGC  
 CTTAACCATTTTCAGCCTTACAGCTAGAAGATTCAGCAAAGTACTTTTTGTGCTCTTGGGGAA  
 CTGTTTTTCTACAGCAGTGCTTCCAAGATAATCTTTGGATCAGGGACCAGACTCAGCATCC  
 GGCCAAATATCCAGAACCCTGACCCTGC<sup>CCGTGTAC</sup><sup>CAGCTGAGAGACT</sup>CTAAATCCAGT  
 GACAAGTCTGTCTGCCTATTCACCGATTTTGGATTCTCAAACAAATGTGTCAAAAGTAAG  
 GATTCTGATGTGTATATCACAGACAAAACCTGTGCTAGACATGAGGTCTATGGACTTCAAG  
 AGCAACAGTGCTGTGGCCTGGAGCAACAAATCTGACTTTGCATGTGCAAACGCCTTCAAC  
 AACAGCATTATTCCAGAAGACACCTTCTTCCCCAGCCCAGAAAGTTCCTGTGATGTCAAG  
 CTGGTCGAGAAAAGCTTTGAAACAGATACGAACCTAAACTTTCAAACCTGTCAGTGATT  
 GGGTTCCGAATCCTCCTCCTGAAAGTGGCCGGGTTTAATCTGCTCATGACGCTGCGGCTGT  
 GGTCCAGTTGA<sup>GGATCC</sup>AGACATGATAAGATAATTGATGAGTTTGGACAAACCACAAC  
 TAGAATGCAGTGAAAAAATGCTTTATTGTGAAATTTGTGATG

### 7.1.4 pHSE3' section (full sequence: 8072 bp)

<sup>Sall</sup>, <sup>BamHI</sup>, <sup>pHSE rev</sup>

CTAAGCGCTTGTGTCGCCATTGTATTCCCGGAAGAGGCTTTTCTTCTAGAAGACTCCAGGG  
 TCTGACTTCTGAAGAGAAGAAGAAAGAGGAAGAGTGGAAAGAGAGGACACAGAGAGTCT  
 GGCCTGCGGGTCTCTCCTGGTGTTTTGGAGAGTTTCTGGATCAGAAGTCCGGAGACGACAGC  
 ACAGGGTTTCAGGCAAAGTCTTAGTCGCCAGGCAGTGAGGTTCAGGGGTGGGGAAAGCCCAG  
 GGCTGGGGATTCCCCATCTCCACAGTTTCACTTCTGCACCTAACCTGGGGTCAGGTCCTTCT  
 GTCCGGACACTGTTGACGCGCAGTCAGCTCTTACCCCCATTGGGTGGCGCGATACCAAG  
 AACCAATCAGTGTGCGCCGCGGACGCTGGATATAAAGTCCACGCAGCCCCGAGAAGTCTCAG

AAGTCGTG**GTCGACT**CTAGA**GGATCC**TGAGAACTTCAGGGTGAGTC**FATGGGACCCTTGA**  
**TGTTT**

### 7.1.5 HLA A\*24:02:01:01 (1098 bp)

**A101\_lead fwd**, **A23\_NotI rev**

**ATGGCCGTCATGGCGCCC**CGAACCCCTCGTCCTGCTACTCTCGGGGGCCCTGGCCCTGACC  
 CAGACCTGGGCAGGCTCCCACTCCATGAGGTATTTCTCCACATCCGTGTCCCGGCCCGGC  
 CGCGGGGAGCCCCGCTTCATCGCCGTGGGCTACGTGGACGACACGCAGTTCGTGCGGTTT  
 GACAGCGACGCCGCGAGCCAGAGGATGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGG  
 GGCCGGAGTATTGGGACGAGGAGACAGGGAAAGTGAAGGCCCACTCACAGACTGACCGA  
 GAGAACCTGCGGATCGCGCTCCGCTACTACAACCAGAGCGAGGCCGGTTCTCACACCCTC  
 CAGATGATGTTTGGCTGCGACGTGGGGTTCGGACGGGCGCTTCCTCCGCGGGTACCACCAG  
 TACGCCTACGACGGCAAGGATTACATCGCCCTGAAAGAGGACCTGCGCTCTTGGACCGCG  
 GCGGACATGGCGGCTCAGATCACCAAGCGCAAGTGGGAGGCGGCCCATGTGGCGGAGCA  
 GCAGAGAGCCTACCTGGAGGGCACGTGCGTGGACGGGCTCCGCAGATACCTGGAGAACG  
 GGAAGGAGACGCTGCAGCGCACGGACCCCCCAAGACACATATGACCCACCACCCCATC  
 TCTGACCATGAGGCCACTCTGAGATGCTGGGCCCTGGGCTTCTACCCTGCGGAGATCACA  
 CTGACCTGGCAGCGGGATGGGGAGGACCAGACCCAGGACACGGAGCTTGTGGAGACCAG  
 GCCTGCAGGGGATGGAACCTTCCAGAAGTGGGCAGCTGTGGTGGTACCTTCTGGAGAGG  
 AGCAGAGATACACCTGCCATGTGCAGCATGAGGGTCTGCCCAAGCCCTCACCTGAGAT  
 GGGAGCCATCTTCCAGCCACCGTCCCATCGTGGGCATCATTGCTGGCCTGGTTCTCCT  
 TGGAGCTGTGATCACTGGAGCTGTGGTTCGCTGCTGTGATGTGGAGGAGGAACAGCTCAGA  
 TAGAAAAGGAGGGAGCTACTCTCAGGCTGCAAGCAGTGACAGTGCCCAGGGCTCTGATG  
 T**GTCTCTCACAGCTTGTAAGTC**TGA

### 7.1.6 HLA B\*07:02:01:01 (1089 bp)

**B0801\_lead fwd**, **B27:02\_Not rev**

**ATGCTGGTCATGGCGCCC**CGAACCGTCCTCCTGCTGCTCTCGGCGGCCCTGGCCCTGACC  
 GAGACCTGGGCGGGTCCCACTCCATGAGGTATTTCTACACCTCCGTGTCCCGGCCCGGC  
 CGCGGGGAGCCCCGCTTCATCTCAGTGGGCTACGTGGACGACACCCAGTTCGTGAGGTTT  
 GACAGCGACGCCGCGAGTCCGAGAGAGGAGCCGCGGGCGCCGTGGATAGAGCAGGAGG  
 GGCCGGAGTATTGGGACCGGAACACACAGATCTACAAGGCCAGGCACAGACTGACCGA  
 GAGAGCCTGCGGAACCTGCGCGGCTACTACAACCAGAGCGAGGCCGGGTCTCACACCCT  
 CCAGAGCATGTACGGCTGCGACGTGGGGCCGACGGGCGCCTCCTCCGCGGGCATGACC  
 AGTACGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCG  
 CCGCGGACACGGCGGCTCAGATCACCCAGCGCAAGTGGGAGGCGGCCCGTGAGGCGGAG  
 CAGCGGAGAGCCTACCTGGAGGGCGAGTGCCTGGAGTGGCTCCGCAGATACCTGGAGAA  
 CGGGAAGGACAAGCTGGAGCGCGCTGACCCCCAAAGACACACGTGACCCACCACCCCA  
 TCTCTGACCATGAGGCCACCCTGAGGTGCTGGGCCCTGGGTTTCTACCCTGCGGAGATCA  
 CACTGACCTGGCAGCGGGATGGCGAGGACCAAACCTCAGGACACTGAGCTTGTGGAGACC  
 AGACCAGCAGGAGATAGAACCTTCCAGAAGTGGGCAGCTGTGGTGGTGCCTTCTGGAGA  
 AGAGCAGAGATACACATGCCATGTACAGCATGAGGGGCTGCCGAAGCCCTCACCTGA  
 GATGGGAGCCGTCTTCCCAGTCCACCGTCCCATCGTGGGCATTGTTGCTGGCCTGGCTGT  
 CCTAGCAGTTGTGGTTCATCGGAGCTGTGGTTCGCTGCTGTGATGTGTAGGAGGAAGAGTTC  
 AGGTGGAAAAGGAGGGAGCTACTCTCAGGCTGCGTGCAGCGACAGTGCCCAGGGCT**CTG**  
**ATGTGTCTCTCACAGCT**TGA

**7.1.7 HLA C\*07:02:01:01 (1101 bp)**

**HLA\_C\_for\_SaII\_for\_(KS), HLA\_C\*07:02:01\_NotI\_rev**

**ATGCGGGTCATGGCGCC**CCGAGCCCTCCTCCTGCTGCTCTCGGGAGGCCTGGCCCTGACC  
GAGACCTGGGCCTGCTCCCACTCCATGAGGTATTTTCGACACCGCCGTGTCCCGGCCCGGC  
CGCGGAGAGCCCCGCTTCATCTCAGTGGGCTACGTGGACGACACGCAGTTTCGTGCGGTTC  
GACAGCGACGCCGCGAGTCCGAGAGGGGAGCCGCGGGCGCCGTGGGTGGAGCAGGAGG  
GGCCGGAGTATTGGGACCGGGAGACACAGAAGTACAAGCGCCAGGCACAGGCTGACCGA  
GTGAGCCTGCGGAACCTGCGCGGCTACTACAACCAGAGCGAGGACGGGTCTCACACCCT  
CCAGAGGATGTCTGGCTGCGACCTGGGGCCCGACGGGCGCCTCCTCCGCGGGTATGACCA  
GTCCGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGC  
CGCGGACACCGCGGCTCAGATACCCAGCGCAAGTTGGAGGCGGCCCGTGCGGCGGAGC  
AGCTGAGAGCCTACCTGGAGGGCACGTGCGTGGAGTGGCTCCGCAGATACCTGGAGAAC  
GGGAAGGAGACGCTGCAGCGCGCAGAACCCCCAAAGACACACGTGACCCACCACCCCT  
CTCTGACCATGAGGCCACCCTGAGGTGCTGGGCCCTGGGCTTCTACCCTGCGGAGATCAC  
ACTGACCTGGCAGCGGGATGGGGAGGACCAGACCCAGGACACCGAGCTTGTGGAGACCA  
GGCCAGCAGGAGATGGAACCTTCCAGAAGTGGGCAGCTGTGGTGGTGCCTTCTGGACAA  
GAGCAGAGATACACGTGCCATATGCAGCACGAGGGGCTGCAAGAGCCCTCACCTGAG  
CTGGGAGCCATCTTCCAGCCACCATCCCATCATGGGCATCGTTGCTGGCCTGGCTGT  
CTGTTTGTCTAGCTGTCTTGGAGCTGTGGTCACCGCTATGATGTGTAGGAGGAAGAGC  
TCAGGTGGAAAAGGAGGGAGCTGCTCTCAGGCTGCGTGCAGCAACAGTGCCAGG**GCTC**  
**TGATGAGTCTCTCATC**ACTTGTAAGCCTGA

**7.1.8 mtagBFP2 (702 bp)**

**mtagBFP2\_Not\_fwd, mtagBFP2\_stop\_BamHI\_rev**

**ATGAGCGAGCTGATTAAGGAGAAC**ATGCACATGAAGCTGTACATGGAGGGCACCGTGGA  
CAACCATCACTTCAAGTGCACATCCGAGGGCGAAGGCAAGCCCTACGAGGGCACCCAGA  
CCATGAGAATCAAGGTGGTCGAGGGCGGCCCTCTCCCTTCGCCTTCGACATCCTGGCTA  
CTAGCTTCTCTACGGCAGCAAGACCTTCATCAACCACACCCAGGGCATCCCCGACTTCTT  
CAAGCAGTCCTTCCCTGAGGGCTTCACATGGGAGAGAGTCACCACATACGAAGACGGGG  
GCGTGCTGACCGCTACCCAGGACACCAGCCTCCAGGACGGCTGCCTCATCTACAACGTCA  
AGATCAGAGGGGTGAACTTCACATCCAACGGCCCTGTGATGCAGAAGAAAACACTCGGC  
TGGGAGGCCTTACCGAGACGCTGTACCCCGCTGACGGCGGCCCTGGAAGGCAGAAACGA  
CATGGCCCTGAAGCTCGTGGGCGGGAGCCATCTGATCGCAAACGCCAAGACCACATATA  
GATCCAAGAAACCCGCTAAGAACCTCAAGATGCCTGGCGTCTACTATGTGGACTACAGAC  
TGGAAGAATCAAGGAGGCCAACACGAGACCTACGTCGAGCAGCACGAGGTGGCAGTG  
GCCAGATACTGCGACCTCCCTAGCAAA**CTGGGGCACAAAGCTTAATTAA**

### 7.1.9 HLA C\*07:02:01:01-BFP construct

HLA C\*07:02:01:01-BFP construct: HLA C\*07:02:01:01plus NotI site plus mtagBFP2

SalI/BamHI/NotI restr. sites

HLA\_C\_for\_SalI\_for\_(KS), HLA\_C\*07:02:01 NotI rev

mtagBFP2\_Not\_fwd, mtagBFP2\_stop\_BamHI rev

GTTCGACATGCGGGTTCATGGCGCCCGAGCCCTCCTCCTGCTGCTCTCGGGAGGCCTGGCC  
 CTGACCGAGACCTGGGCCTGCTCCCACTCCATGAGGTATTTTCGACACCGCCGTGTCCCGG  
 CCCGGCCGCGGAGAGCCCCGCTTCATCTCAGTGGGCTACGTGGACGACACGCAGTTCGTG  
 CGGTTTCGACAGCGACGCCGCGAGTCCGAGAGGGGAGCCGCGGGCGCCGTGGGTGGAGCA  
 GGAGGGGCCGGAGTATTGGGACCGGGAGACACAGAAGTACAAGCGCCAGGCACAGGCT  
 GACCGAGTGAGCCTGCGGAACCTGCGCGGCTACTACAACCAGAGCGAGGACGGGTCTCA  
 CACCCTCCAGAGGATGTCTGGCTGCGACCTGGGGCCCCGACGGGCGCCTCCTCCGCGGGTA  
 TGACCAGTCCGCCTACGACGGCAAGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTG  
 GACCGCCGCGGACACCGCGGCTCAGATCACCCAGCGCAAGTTGGAGGCGGCCCGTGC GG  
 CGGAGCAGCTGAGAGCCTACCTGGAGGGCACGTGCGTGGAGTGGCTCCGCAGATACCTG  
 GAGAACGGGAAGGAGACGCTGCAGCGCGCAGAACCCCAAAGACACACGTGACCCACC  
 ACCCCCTCTCTGACCATGAGGCCACCCTGAGGTGCTGGGCCCTGGGCTTCTACCCTGCGG  
 AGATCACACTGACCTGGCAGCGGGATGGGGAGGACCAGACCCAGGACACCGAGCTTGTG  
 GAGACCAGGCCAGCAGGAGATGGAACCTTCCAGAAGTGGGCAGCTGTGGTGGTGCCTTC  
 TGGACAAGAGCAGAGATACACGTGCCATATGCAGCACGAGGGGCTGCAAGAGCCCCTCA  
 CCCTGAGCTGGGAGCCATCTTCCCAGCCCACCATCCCCATCATGGGCATCGTTGCTGGCCT  
 GGCTGTCCTGGTTGTCCTAGCTGTCCTTGGAGCTGTGGTCACCGCTATGATGTGTAGGAGG  
 AAGAGCTCAGGTGGAAAAGGAGGGAGCTGCTCTCAGGCTGCGTGCAGCAACAGTGCCCA  
 GGCTCTGATGAGTCTCTCATCGCGGCCGAAAGCGAGCTGATTAAGGAGAACATGCACAT  
 GAAGCTGTACATGGAGGGCACCGTGGACAACCATCACTTCAAGTGCACATCCGAGGGCG  
 AAGGCAAGCCCTACGAGGGCACCCAGACCATGAGAATCAAGGTGGTTCGAGGGCGGCCCT  
 CTCCCCTTCGCCTTCGACATCCTGGCTACTAGCTTCTCTACGGCAGCAAGACCTTCATCA  
 ACCACACCCAGGGCATCCCCGACTTCTTCAAGCAGTCCTTCCCTGAGGGCTTCACATGGG  
 AGAGAGTCACCACATACGAAGACGGGGGCGTGTGACCGCTACCCAGGACACCAGCCTC  
 CAGGACGGCTGCCTCATCTACAACGTCAAGATCAGAGGGGTGAACTTCACATCCAACGGC  
 CCTGTGATGCAGAAGAAAACACTCGGCTGGGAGGCCTTACCGAGACGCTGTACCCCGCT  
 GACGGCGGCCCTGGAAGGCAGAAACGACATGGCCCTGAAGCTCGTGGGCGGGAGCCATCT  
 GATCGCAAACGCCAAGACCACATATAGATCCAAGAAACCCGCTAAGAACCTCAAGATGC  
 CTGGCGTCTACTATGTGACTACAGACTGGAAAGAATCAAGGAGGCCAACACGAGACC  
 TACGTCGAGCAGCACGAGGTGGCAGTGGCCAGATACTGCGACCTCCCTAGCAAACTGGG  
 GCACAAGCTTAATTAAGGATCC

### 7.1.10 pcDNA6-V5 segment of N24 PECPL library (146 bp)

pcDNA-? pcDNA-?

CACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGAT  
 CCGGCGCGCCACCATGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGAGCGGCCGCTCGAGTCTA  
 GAGGGCCCTTCGAAGGTAAGCCTATC

## 7.2 Results part: T cell antigen search

### 7.2.1 Co-culture experiments: Single cell isolation and reactivation assays

Single cell isolation experiments were performed in 3.5 cm cell culture dishes 24, 48 and 72 hours after FuGene transfection of COS-7 cells with 1 µg HLA and 1 µg plasmid DNA of different PECPL libraries. COS-7 cells underneath green fluorescing clusters (singlets, doublets or triplets) of T hybridoma cells were isolated with a micromanipulator and listed in Table 18 (upper and central section). Samples were named according to their HLA restriction and the date of the performed experiment.

PECPL plasmid DNA from COS cells was enriched by PCR amplification with pcDNA-for1/pcDNA-rev1 (product size: 280 bp) primers and all 37 samples pooled in equimolar concentrations into six different pools (pool#1-pool#6) as assigned in Table 18. Pools were reconstituted in pcDNA6 vectors and used for another round of single cell isolation experiments (named as “re-isolation”, lower section in Table 18). Re-isolated single cells were mainly found in pool#3 and pool#5 and the respective original samples of those pools highlighted in bold letters. The nomenclature of the 11x original samples from pool#3 and pool#5 plus the 11x re-isolated samples that were individually tested for reactivation in following subpool experiments (Table 19) are given in green bold letters.

**Table 18: T cell antigen search single cell isolation experiments**

HLA	libraries	surface screened	Detected T cell clusters	Nomenclature of isolated single cells								
B07:02	8L/9L	57.6 cm <sup>2</sup>	Single: 2x Double: 4x Triple: 1x	Pool #1	09.08.17	B1						
					16.08.17:	B1						
						B2						
					24.08.17:	B1						
						B2						
					25.08.17:	B1						
						B2						
					C07:02	8L/9L	57.6 cm <sup>2</sup>	Single: 3x Double: 5x Triple: 6x	Pool #2	09.08.17	C1	
											C2	
											C3	
											C4	
											C5	
16.08.17	C3											
24.08.17	C1											
	C5											
25.08.17	C1											
	C2											
<b>Pool #3</b>	<b>16.08.17</b>	<b>C1</b>	<b>#1</b>									
		<b>C2</b>	<b>#15</b>									
		<b>24.08.17</b>	<b>C2</b>	<b>#2</b>								
		<b>C3</b>	<b>#3</b>									
		<b>C4</b>	<b>#16</b>									
		<b>C6</b>	<b>#17</b>									

C07:02	8L	67.2 cm <sup>2</sup>	Single: 1x Double: 5x Triple: 2x	Pool #4	14.09.17	C_8L-3				
					25.10.17	C_8L-1				
					<b>Pool #5</b>	<b>13.09.17</b>	<b>C_8L</b>	<b>#4</b>		
						<b>14.09.17</b>	<b>C_8L-1</b>	<b>#5</b>		
							<b>C_8L-2</b>	<b>#6</b>		
				<b>10.10.17</b>	<b>C_8L</b>	<b>#18</b>				
				<b>24.10.17</b>	<b>C_8L</b>	<b>#7</b>				
				Pool #6	25.10.17	C_8L-2				
				9L	57.6 cm <sup>2</sup>	Single: 1x	Pool #4	14.09.17	C_9L	
				10L	38.4 cm <sup>2</sup>	Single: 1x Double: 2x	Pool #4	10.10.17	C_10L-1	
24.10.17	C_10L									
Pool #6	10.10.17	C_10L-2								
N24	28.8 cm <sup>2</sup>	Ø								
N27	28.8 cm <sup>2</sup>	Double: 1x	Pool #6	27.10.17	C_N27					
N30	19.2 cm <sup>2</sup>	Double: 1x	Pool #6	26.10.17	C_N30					
Re-isolated from pools	Pool #1	9.6 cm <sup>2</sup>	Ø							
	Pool #2	9.6 cm <sup>2</sup>	Triple: 1x							
	<b>Pool #3</b>	57.6 cm <sup>2</sup>	Double: 5x	<b>15.02.18</b>	<b>#3-1</b>	<b>#20</b>				
					<b>#3-2</b>	<b>#9</b>				
					<b>16.02.18</b>	<b>#3-1</b>	<b>#10</b>			
						<b>#3-2</b>	<b>#11</b>			
						<b>#3-3</b>	<b>#21</b>			
					<b>12.07.18</b>	<b>#3-1</b>	<b>#22</b>			
	<b>#3-2</b>	<b>#12</b>								
	Pool #4	57.6 cm <sup>2</sup>	Triple: 2x							
	<b>Pool #5</b>	57.6 cm <sup>2</sup>	Double: 4x	<b>08.02.18</b>	<b>#5-1</b>	<b>#8</b>				
					<b>#5-2</b>	<b>#19</b>				
				<b>12.07.18</b>	<b>#5-1</b>	<b>#13</b>				
					<b>#5-2</b>	<b>#14</b>				
Pool #6	9.6 cm <sup>2</sup>	Ø								

T cell reactivation from isolated samples were validated in subpool experiments. The 22 samples highlighted in bold, green numbers in Table 18 were chosen for reactivation experiments. Assays were performed in 24-well cell culture plates 24, 48 and 72 hours after FuGene transfection of COS cells with 0.5 µg HLA~C07:02 and 0.5 µg recloned plasmid DNA. Green fluorescing clusters (singlets, doublets or triplets) of T hybridoma cells of each transfected well were counted and listed below (Table 19).

Pool #4 with the highest numbers of T cell reactivation was divided into 30 subpools and each subpool tested for individual reactivation. In total, three rounds of subpool experiments were performed.

Green, bold letters were used in accordance with the nomenclature of Table 18. Green shaded rows indicate the pool with the most reactivated green T cell clusters of each subpool round that was then used for the next round of subpool experiments.

**Table 19: T-cell reactivation assays**

<b>Isolated single cells</b>	<b>24 hours</b>	<b>48 hours</b>
#1		
#2		
#3	Single: 1x	Single: 1x
#4	Single: 5x	Single: 2x
#5		Double: 1x
#6		
#7	Single: 3x	Single: 1x
#8		Single: 2x
#9		
#10	Single: 1x	
#11		
#12		
#13		Single: 1x
#14		Single: 1x
#15		
#16		Single: 1x
#17		
#18		
#19		
#20	Single: 1x	Single: 1x
#21		Single: 1x
<b>Subpool #4</b>	<b>24 hours</b>	<b>48 hours</b>
<b>Round 1</b>		
<b>120.000 clones</b>		
#1		Single: 1x
#2		Single: 2x
		Double: 1x
#3	Single: 1x	Single: 1x
#4		
#5		Single: 1x
#6		
#7		
#8		
#9		
#10		
#11		
#12	Single: 1x	Single: 1x
#13	Single: 1x	
#14		
#15		Single: 1x
#16		
#17	Single: 2x	
#18		
#19		Single: 1x
#20	Single: 1x	
#21		Single: 1x
#22		
#23		Double: 1x
#24		

#25			
#26			
#27		Single: 2x	
#28		Single: 1x	
#29			Double: 1x
#30	Single: 2x	Triple: 1x	Single: 3x
<b>Subpool #4-30</b>			
<b>Round 2</b>			
<b>14.000 clones</b>			
#1	Single: 5x	Single: 1x	Single: 1x Double: 1x
#2	Single: 1x	Single: 2x	Single: 2x
#3	Single: 1x	Single: 1x	Single: 1x
#4	Single: 1x	Single: 1x	Single: 1
#5	Single: 2x	Single: 2x	
#6		Single: 1x	Single: 2x
#7		Single: 2x	
#8	Single: 2x	Single: 3x	
#9	Single: 2x	Single: 1x	Single: 2x
#10	Single: 1x	Single: 1x	Single: 1x
#11	Single: 1x		
#12		Triple: 1x Single: 2x	Single: 1x
#13		Double: 1x Single: 1x	
#14	Single: 2x		
#15	Single: 2x	Single: 3x	Single: 1x
#16	Single: 5x	Single: 1x	
#17	Single: 1x	Single: 2x	
#18	Single: 2x	Single: 2x	
#19		Double: 1x	
#20	Single: 1x	Single: 2x	
#21			Single: 2x
#22	Single: 4x	Single: 1x	Single: 2x
#23	Single: 2x	Single: 2x	Single: 2x
#24		Single: 1x	Single: 1x
#25	Single: 2x	Single: 3x	Single: 2x
#26	Single: 1x	Single: 1x	Single: 3x
#27	Single: 2x	Single: 1x	
#28	Single: 2x	Double: 1x	
#28	Single: 2x	Single: 4x	Single: 1x
#29	Single: 2x	Single: 1x	Single: 2x
#30	Single: 4x	Single: 1x	
<b>Subpool #4-30- 1</b>			
<b>Round 3</b>			
<b>1.400 clones</b>			
#1	Single: 1x		
#2	Single: 1x	Single: 2x	
#3	Single: 2x	Single: 2x	
#4		Single: 2x	
#5		Single: 4x	
#6	Single: 3x	Single: 1x	
#7			
#8			
#9			
#10	Single: 1x		
#11	Single: 2x		
#12	Single: 1x		
#13	Single: 1x	Single: 3x	

---

#14		
#15		Single: 1x
#16		Single: 3x
#17		Single: 1x
#18		
#19		Single: 2x
#20	Single: 1x Double: 1x	Single: 2x
#21	Single: 1x	Single: 1x
#22	Single: 2x	
#23		Single: 2x
#24	Single: 1x	Single: 1x
#25	Single: 1x	Single: 1x
#26	Single: 2x	
#27	Single: 2x	
#28	Single: 1x	Single: 4x
#29		Single: 1x
#30	Single: 4 Double: 1x	

---

## 8 List of Publications

E. Beltrán\*, L. A. Gerdes\*, **J. Hansen**\*, A. Flierl-Hecht, S. Krebs, H. Blum, B. Ertl-Wagner, F. Barkhof, T. Kümpfel, R. Hohlfeld\* and K. Dornmair\*, “Early adaptive immune activation detected in monozygotic twins with prodromal multiple sclerosis.”, *The Journal of Clinical Investigation* 2019, 129, 4758-4768.

\*These authors contributed equally to the work.