JOHANNES GUTENBERG UNIVERSITÄT MAINZ

Antigen specificity of tumor-infiltrating lymphocytes (TILs) in non-small cell lung cancer (NSCLC)

Dissertation

To obtain the doctoral degree (Dr. rer. nat.) of the
Faculty of Biology at the
Johannes Gutenberg University Mainz

Submitted by
Christoph Nico Doppler, M.Sc. in Biomedicine,
born on April 1st, 1993,
in Schwetzingen, Germany

Mainz, March 2025

1st Supervisor / Assessor:	[redacted]
2nd Supervisor:	[redacted]
2nd Assessor:	[redacted]
Faculty:	Faculty 10 – Biology, Johannes Gutenberg University Mainz
Dean:	Univ.-Prof. Dr. Eckhard Thines
Author:	Christoph Nico Doppler
Start of the Doctorate:	05.07.2019 (Application no.: 107037)
Affiliation:	Wölfel Laboratory; Internal Medicine III, University Medical Center of the Johannes Gutenberg University Mainz
Graduate School:	TransMed – Mainz Research School of Translational Biomedicine; University Medical Center of the Johannes Gutenberg University Mainz
Thesis Advisory Committee:	[redacted]
Scientific Cooperation:	HS Diagnomics GmbH; Schloßstraße 110, 12163 Berlin, Germany

Table of Contents

1) Summary

1.1 Summary in English

The direct identification of tumor-specific T cells among polyclonal populations of tumor-infiltrating lymphocytes (TILs) could revolutionize the field of T cell-based cancer immunotherapies. Increasing the total numbers of tumor-specific lymphocytes among T cell populations used for adoptive cell therapies (ACTs) has the potential to dramatically increase the efficiency of tumor cell eradication in patients suffering from various malignancies. Hence, by analyzing T cell receptor (TCR) repertoires in tumor and adjacent normal tissues of individual cancer patients, the biotechnology company HS Diagnomics (HSD) has developed an antigen-agnostic method to identify potentially tumor-specific T cell clonotypes in highly heterogenous TIL populations. The main objective of this doctoral research study was to assess the power of the HSD prediction strategy in three stage III non-small cell lung cancer (NSCLC) patients (P18, P43, P50) by identifying and characterizing the target antigens of supposedly tumor-reactive or -specific TCRs. For this purpose, only TCRs from programmed cell death protein 1 (PD-1)-positive TILs with highest tumor-to-non-tumor ratios were chosen.

TCR-transgenic T cell populations were generated via retroviral transduction of healthy donor lymphocytes after CRISPR/Cas9-mediated knock-out of the endogenous TCR expression. TCR-engineered T cells (TCR-T cells) were then tested for the recognition of known tumor-associated antigens (TAAs) as well as patient-specific mutated neoantigen candidates. The latter were determined for each patient by applying comparative whole exome sequencing (WES) of tumor and adjacent normal tissues, RNA sequencing (RNA-Seq) of tumors, and subsequent human leukocyte antigen (HLA) binding predictions with established algorithms.

While no TAA or neoantigen specificities were verified for the analyzed P43 and P50 TCRs, three TCRs cloned from TILs of patient P18 specifically recognized a mutated neopeptide encoded by $KRAS^{Q61H}$. With the $KRAS^{Q61H}$ mutation creating an immunogenic neoepitope, the corresponding 10-mer peptide ILDTAGHEEY was recognized in the context of HLA-A*01:01. Since the proper processing and presentation of the peptide was also verified, $KRAS^{Q61H}$ was considered an immunogenic recurrent driver mutation that can be targeted in a sizeable number of patients.

Moreover, the three $KRAS^{Q61H}$-specific P18 TCRs formed a specificity cluster consisting of CDR3 sequences with highly homologous antigen recognition motifs. Motivated by this finding, HSD used an in-house developed clustering approach to identify even more TCR clusters in additional patients. Hence, TCRs from two TCR clusters, namely cluster C with six TCRs from four and cluster G with six TCRs from three NSCLC patients, were additionally analyzed. Respective cluster TCRs mediated similar reactivity patterns when tested on partially HLA-matched allogeneic cell lines. However, no target antigen could be identified for the analyzed cluster C TCRs. In contrast, the CDR3 regions of cluster G TCRs were highly related to a cluster previously described by the group of Mark M. Davis (Stanford University, USA). Indeed, cluster G TCRs exhibited identical antigen specificities as the Stanford cluster: They recognized two pathogen-derived 9-mer peptides as well as a peptide originating from the non-mutated TAA TMEM161A in the context of HLA-A*02:01.

In summary, this study demonstrated the power of two *in silico* prediction pipelines resulting in the identification of nine TCRs that had initially been cloned from TIL populations of four different NSCLC patients and either recognized the shared $KRAS^{Q61H}$ neoantigen or the non-mutated TAA TMEM161A. Therefore, the presented data substantially support the efforts to further develop antigen-agnostic, personalized T cell immunotherapies.

1.2 Summary in German

Die Identifizierung tumorspezifischer T-Zellen in polyklonalen Populationen tumorinfiltrierender Lymphozyten (engl.: TILs) könnte das Gebiet der T-Zell-basierten Immuntherapien nachhaltig revolutionieren. So könnte die Wirksamkeit adoptiver Zelltherapien (engl.: ACTs) beispielsweise erheblich gesteigert werden, wenn Krebspatienten mit einer höheren Anzahl tumorspezifischer T-Zellen behandelt würden. Das Biotechnologieunternehmen HS Diagnomics (HSD) hat daher eine Antigen-agnostische Methode zur effizienten Identifizierung tumorspezifischer T-Zell-Kandidaten in heterogenen TIL-Populationen entwickelt, die auf dem dezidierten Vergleich von T-Zell-Rezeptor-Repertoires aus den Tumoren und benachbartem, gesundem Gewebe individueller Krebspatienten basiert. Um die Effizienz dieser Methode zu analysieren, bestand das vorrangige Ziel dieser Dissertation darin, Zielantigene von potenziell tumorreaktiven oder -spezifischen T-Zell-Rezeptoren (engl.: TCRs) in drei Patienten (P18, P43, P50) mit nicht-kleinzelligem Lungenkrebs (engl.: NSCLC) im Stadium III zu identifizieren und zu charakterisierten. Hierfür wurden ausschließlich TCRs aus Programmed cell death protein 1 (PD1)-positiven Klonotypen mit den jeweils höchsten Tumor/Nicht-Tumor-Verhältnissen herangezogen.

Nach einem CRISPR/Cas9-vermitteltem Knock-out der endogenen TCR-Expression in gesunden Spenderlymphozyten, wurden mittels retroviraler Transduktion zunächst TCR-transgene T-Zell-Populationen erzeugt. Die so generierten T-Zellen (TCR-T-Zellen) wurden dann auf die Erkennung von tumor-assoziierten Antigenen (engl.: TAAs) und patientenspezifischen, mutierten Neoantigen-Kandidaten getestet. Letztere wurden hierfür durch eine Kombination aus Whole Exome Sequencing (WES) und RNA-Sequencing (RNA-Seq) von Tumor- und angrenzendem Normalgewebe individuell für jeden Patienten identifiziert. Durch Bindungsvorhersagen an die humanen Leukozytenantigene (engl.: HLAs) der jeweiligen Patienten wurden anschließend die vielversprechendsten Neoantigen-Kandidaten unter Verwendung etablierter Algorithmen identifiziert.

Während für die untersuchten TCRs der Patienten P43 und P50 keinerlei TAA- oder Neoantigen-Spezifitäten nachgewiesen werden konnten, wurden schließlich drei TCRs des Patienten P18 identifiziert, die spezifisch ein durch $KRAS^{Q61H}$ kodiertes, mutiertes Neopeptid erkannten. Das aus der $KRAS^{Q61H}$-Mutation resultierende, immunogene 10-mer-Peptid ILDTAGHEEY wurde dann im Kontext von HLA-A*01:01 erkannt. Da zusätzlich die ordnungsgemäße Prozessierung und Präsentation des Peptides verifiziert wurde, konnte schließlich davon ausgegangen werden, dass $KRAS^{Q61H}$ eine immunogene Treibermutation repräsentiert, die zudem in einer beträchtlichen Anzahl von Patienten auftritt.

Des Weiteren bildeten die drei $KRAS^{Q61H}$-spezifischen TCRs ein Spezifitätscluster, das aus ähnlichen CDR3-Sequenzen mit hochgradig homologen Antigen-Bindestellen bestand. Basierend auf dieser Entdeckung verwendete HSD eine eigens entwickelte Clustering-Methode, um weitere TCR-Cluster in zusätzlichen Patienten zu identifizieren. So wurden TCRs aus zwei weiteren Clustern, nämlich Cluster C und Cluster G, mit jeweils sechs TCRs aus vier bzw. drei NSCLC-Patienten zusätzlich analysiert. Die TCRs der respektiven Cluster wiesen jeweils übereinstimmende Spezifitäten gegen allogene Zelllinien mit teilweise übereinstimmenden HLAs auf. Während für die TCRs des Clusters C kein spezifisches Zielantigen identifiziert werden konnte, waren die CDR3-Regionen der TCRs des Clusters G eng mit einem Cluster verwandt, der zuvor bereits von der Gruppe von Mark M. Davis (Stanford University, USA) beschrieben wurde. Tatsächlich wiesen die TCRs beider Cluster (Cluster G, Stanford Cluster) jeweils identische Antigenspezifitäten auf und erkannten zwei Pathogen-spezifische 9-mer-Peptide sowie ein weiteres Peptid des nicht mutierten TAAs TMEM161A in Verbindung mit HLA-A*02:01.

Zusammenfassend wurden im Rahmen dieser Studie also zwei vielversprechende *in silico* Strategien präsentiert, die schlussendlich zur Identifizierung von neun TCRs führten, welche ursprünglich aus TIL-Populationen von vier verschiedenen Patienten kloniert wurden und entweder das $KRAS^{Q61H}$-Neoantigen oder das nicht-mutierte TAA TMEM161A erkannten. Die hier präsentierten Daten tragen daher erheblich zur Entwicklung Antigen-agnostischer, personalisierter T-Zell-Immuntherapien bei.

2) Introduction

2.1 Scientific purpose and overall aim of the study

Cancer still accounts for millions of premature deaths every year, thereby representing one of the major causes of morbidity and mortality worldwide. Hence, continuous research is an absolute requirement for optimizing existing and developing new cancer therapies to improve the overall global health. Harnessing a patient's immune system to fight tumor progression is one of the most promising treatment strategies that emerged in recent years, thereby revolutionizing the field of modern anti-cancer treatments. Hence, improving T cell-based immunotherapies such as the adoptive transfer of autologous tumor-infiltrating lymphocytes (TILs) could significantly complement the repertoire of pre-existing anti-tumor therapies. However, the efficient transfusion of *ex vivo*-expanded TILs was yet only demonstrated in selected tumor subsets and was only feasible for the treatment of resectable tumors from which enough T cells could be isolated for excessive expansion. The absent tumor control observed in some patients might be explained by relatively low numbers of tumor-reactive T cells within TIL populations used for autologous T cell transfers [1,2]. Identifying and enriching tumor-reactive T cells from bulk TILs could therefore substantially improve overall efficiencies of T cell-based adoptive cellular therapies (ACTs). In addition to the selective expansion and direct application of tumor-reactive TILs, the isolation of their T cell receptors (TCRs) represents another attractive strategy to facilitate the manufacturing of therapeutic TCR-T cells via genetic engineering of fresh lymphocytes.

Due to the impact that high numbers of tumor-reactive T cells might have on the overall efficiency of T cell-based ACTs, our collaboration partner HS Diagnomics (HSD) has developed a method to predict potentially tumor-specific T cells among TIL populations previously isolated from individual cancer patients. Here, tumor-enriched TCR clonotypes are identified by comparing tumor- and adjacent healthy tissue-derived T cell repertoires of cancer patients suffering from various malignancies [3]. However, actual antigen specificities of the identified TCR clonotypes remain elusive since this method mainly focuses on TIL frequency comparisons between a patient's healthy and tumor tissue. Therefore, this collaborative project was designed as a proof-of-concept study to verify tumor-specificities of potentially tumor-reactive TCR clonotypes, thereby demonstrating the power of the HSD-proprietary prediction pipeline. To assess the power of the method, tumor-specificities were analyzed for TCRs cloned from 16 tumor-enriched T cell clonotypes previously isolated from three different patients with stage III non-small cell lung cancer (NSCLC).

Confirming not only tumor-reactivity but also revealing the exact target antigens of those candidate TCRs was highly important to assess both feasibility and safety of the applied prediction pipeline. Confirming the power of the applied pipeline to efficiently predict tumor-specific T cells could revolutionize both personalized and off-the-shelf ACT-based cancer treatments: The existence of shared neoantigens and tumor-associated antigens (TAAs) represents a solid basis for the development of efficient off-the-shelf immunotherapies targeting tumor antigens present in multiple patients and various malignancies [4,5]. In HLA-matched patients, identical peptides can be expected to be processed, presented, and eventually recognized by common TCRs. Thus, antigen-agnostic approaches aiming at the identification of common tumor-reactive T cell clonotypes could eventually pave the way for ACTs using therapeutic T cells genetically engineered to express "off-the-shelf" tumor-specific TCRs.

2.2 T cell-mediated antigen recognition

2.2.1 Antigen presentation via MHC class I and MHC class II

T cell recognition of malignant cells is a multi-step process including both antigen presentation via human leukocyte antigen (HLA) molecules and TCR-mediated detection of the presented antigen peptides. Generally, mature αβ T lymphocytes recognize antigenic oligopeptides either presented by HLA class I or HLA class II proteins, also known as major histocompatibility complex (MHC) class I and MHC class II, thereby initiating complex immune responses to effectively eradicate pathogens and tumor cells [6,7]. Both classes of MHC molecules are highly related but differ in both subunit composition and expression patterns. MHC class I molecules are heterodimers consisting of a highly polymorphic alpha chain that is non-covalently associated with the smaller and non-polymorphic β_2-microglobulin chain. In contrast, MHC class II proteins are composed of a non-covalent complex consisting of polymorphic alpha and beta chains, respectively. With MHC class I and class II alpha chains being different proteins, all classical HLA class I (*HLA-A*, *-B* and *-C*) and class II (*HLA-DP*, *-DQ*, *-DR*) genes encode various allele variants located in the *MHC* gene cluster on human chromosome 6 [8,9]. In contrast, the *B2M* gene encoding the β_2-microglobulin subunit of MHC class I is located on human chromosome 15. B2M expression is a prerequisite for both proper surface expression and function of class I MHC molecules [10–12]. While HLA class I is expressed on all nucleated cells and directly interacts with CD8+ T cells, HLA class II is typically bound by CD4+ T helper cells and is mainly expressed by professional antigen-presenting cells (APCs), such as dendritic cells, macrophages, monocytes, or B cells [6].

The *extended major histocompatibility complex* (*xMHC*) gene cluster was completely revealed by sequencing the entire human chromosome 6 in 2003 and extended far beyond the boundaries defining the classical *MHC* locus at that time. The *xMHC*, in the following equivalent to *MHC*, comprises a total of 421 gene loci (excluding RNA genes) in an overall sequence length of 7.6 Mb [8,13]. Furthermore, the *MHC* locus is both polygenic and highly polymorphic: It contains various MHC class I and class II alleles encoding different MHC molecules with varying peptide-binding specificities. Hence, MHC molecules substantially contribute to the immune system's ability to respond to a multitude of different and rapidly evolving pathogens [9,13]. In fact, *HLA* genes are the most polymorphic genes across the human population accounting for almost 17,000 different HLA class I and over 6,000 HLA class II alleles. Moreover, *HLA-B* is the most polymorphic region of the human genome comprising more than 6,500 different alleles [14]. While MHC class II proteins typically present peptides arising from proteolytical degradation of extracellular antigens, MHC class I proteins usually bind oligopeptides originating from degradation of intracellular proteins [15–17]. Due to bilateral restrictions of class I binding grooves by conserved tyrosine residues, MHC class I proteins preferentially bind smaller oligopeptides of 8-11 amino acids in length. Due to the less restricted molecular composition of class II peptide binding clefts, MHC class II proteins usually accommodate peptides of 13-24 residues in length [18,19]. Efficient oligopeptide presentation by proteins of both MHC classes is highly dependent on molecular interactions, such as the occupation of defined pockets (A-F) within the MHC binding groves by peptide side chains or the formation of hydrogen bonds between MHC side chains and the peptide backbone. Both position and identity of those peptide side chains, the so-called anchor residues, can vary between peptides bound by different MHC molecules. Anchor residues of peptides binding to a particular MHC molecule are not necessarily identical, but they are usually found to exhibit similar biochemical properties [19,20]. Polymorphisms present in different MHC isoforms result in alterations of amino acids lining their respective peptide binding pockets, thereby resulting in altered binding specificities due to differing anchor residues of preferentially bound peptides. Generally, peptides binding to MHC class I molecules were found to contain anchor residues at amino acid positions P2, P5/6, and at the C-terminus. In contrast, amino acid positions P1, P4, P6, and P9 were identified as anchors for oligopeptides binding to MHC class II proteins [19–22].

2.2.2 T cell receptor-mediated binding of MHC-presented target antigens

T lymphocytes specifically recognize HLA-presented antigen peptides by binding peptide-MHC (pMHC) complexes via T cell receptors (TCRs) expressed on their cellular surfaces. Here, the TCRs interact with both the MHC molecule and the bound antigen peptide. Every single T cell expresses approx. 30,000 identical TCRs consisting of two transmembrane glycoprotein chains, namely the alpha and beta chains, that are linked by a disulfide bond, thereby forming αβ heterodimers [23]. While αβ TCRs mediate antigen recognition in most T lymphocytes, a minority of approx. 1-5% of peripheral T cells bear an alternative type of TCR consisting of an γδ heterodimer. Although γδ T lymphocytes share many effector capabilities with conventional αβ T cells, they represent an independent lineage and exhibit highly different antigen recognition properties. For instance, they do not rely on conventional MHC molecules for antigen presentation but recognize alternative ligands, such as non-peptide phospho-antigens, non-classical MHC molecules or lipids, and glycolipids presented by CD1 glycoproteins [24,25].

Both TCR alpha and beta chains are composed of a hydrophobic transmembrane domain, a short stalk segment providing cysteine residues for forming the interchain disulfide bond, constant (*C*) regions, and membrane-distal variable (*V*) regions that are responsible for binding pMHC complexes (see Figure 2.2). During T cell development in the thymus, alpha and beta chain V regions are generated by recombining single *V(D)J* gene segments from a large repertoire of *V* (variable), *D* (diversity), and *J* (joining) gene segments [6,26]. In contrast to the *TCR α* gene locus (human chromosome 14) exclusively containing V_α and J_α gene segments, the *TCR β* locus (human chromosome 7) encodes V_β, J_β, and additional *D* gene segments. Recombined *V(D)J* gene regions are subsequently transcribed and are then spliced to join the variable segments to the C_α or either of two C_β constant gene segments. To form a functional αβ TCR, resulting mRNA transcripts are translated into complete TCR alpha and beta chains paring soon after translation [27,28]. During the process of antigen recognition, TCR alpha and beta chain variable regions contact the HLA molecule and the presented peptide via three hypervariable loops, namely the complementary determining regions (CDRs), which are generated during VDJ recombination [6]. Less variable CDR1 and CDR2 form the periphery of the pMHC binding site and are encoded within the germline *V* gene segments of both TCR chains. The highest TCR diversity resides in its CDR3 hypervariable loops, which are encoded by the V(D)J junctional regions of each TCR chain. Variability results from imprecise joining of the gene segments and is enhanced by a template-independent introduction of P- and N-nucleotides. While CDR1 and CDR2 mainly contact conserved domains of the MHC protein, CDR3 segments form the center of the pMHC-binding site and directly interact with the HLA-presented peptide [26]. Upon canonical antigen recognition, the Vα domain of the TCR is positioned above the N-terminal portion of the MHC-bound peptide while its Vβ domain is located over the C-terminal section of the peptide [15]. However, the structurally diverse TCR αβ heterodimer alone is not sufficient for signal transduction and is therefore non-covalently associated with invariant CD3 (CD3γ, CD3δ, CD3ε) and ζ chains, thereby forming the signaling apparatus of the TCR-CD3 complex (see Fig. 2.2). Since the complete assembly of the TCR-CD3 complex is required for proper cell surface expression, all subunits have to be present to form functional signaling complexes [29–31].

T cell activation is initiated by TCR-mediated binding of the pMHC complex and results in the rapid phosphorylation of immune receptor tyrosine-based activation motifs (ITAMs) present in cytoplasmatic regions of the CD3 chains and the two ζ chains that are additionally associated with the TCR-CD3 complex. The Src family kinase Lck (lymphocyte-specific protein tyrosine kinase) was found to be associated with TCR co-receptors (CD4, CD8) and is thought to be the predominant enzyme involved in ITAM phosphorylation upon co-receptor binding to MHC class I or class II molecules [32,33]. The phosphorylation of both tyrosine residues present in each ITAM facilitates the binding of multiple downstream adaptor proteins, thereby providing a platform for the formation of a multi-protein singling complex eventually resulting in the membrane recruitment of the key signaling protein phospholipase C-γ (PLC-γ) [33,34]. When activated, PLC-γ induces two of the main pathways necessary for efficient T cell activation: The nuclear translocation of NFAT (nuclear factor of activated T cells) transcription

factors and the activation of the Ras-MAPK (mitogen activated protein kinase) pathway [34–36]. However, the full T cell activation requires the engagement of various co-stimulatory receptors, such as lymphocyte function-associated antigen 1 (LFA-1) or CD28. LFA-1 increases calcium signaling crucial for T cell activation and induces IL-2 synthesis by promoting the activity of the activator protein-1 (AP-1) transcription factor [37,38]. On the other hand, CD28 crucially affects T cell activation by inducing PLC-γ activation, prolonging the overall time of NFAT in the nucleus, and being involved in the activation of additional transcription factors like AP-1 and NFκB (nuclear factor kappa-light-chain enhancer of activated B cells). CD28 therefore provides strong synergistic signals for controlling the cytokine production, cell proliferation, cell death, and the differentiation of naïve T cells [39–41].

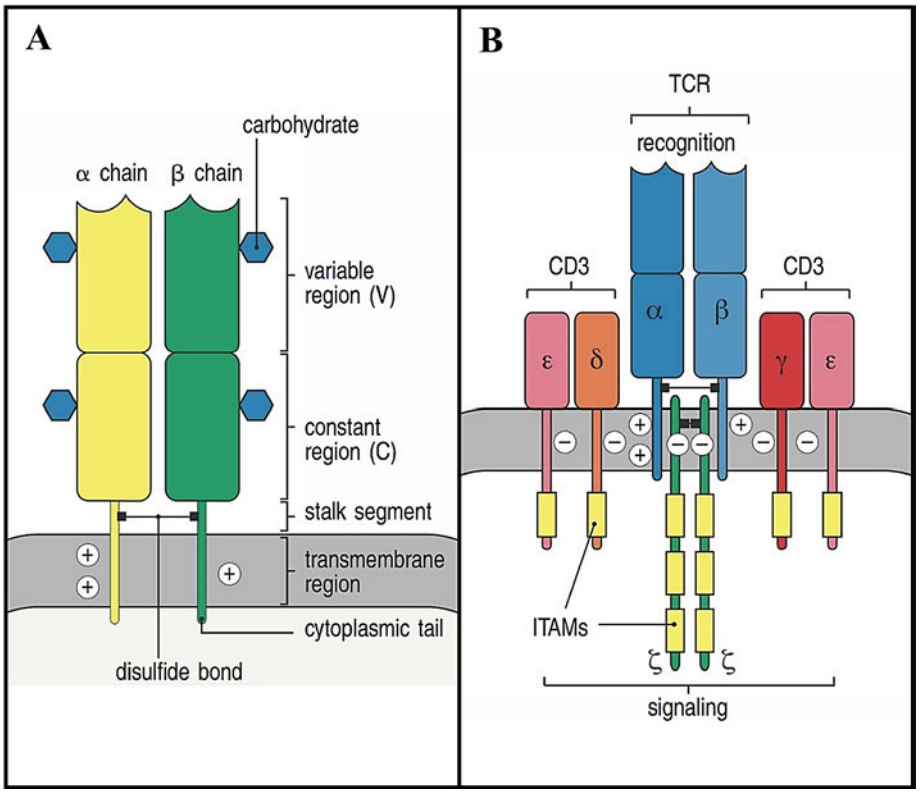


Figure 2.2: Molecular structure of the αβ TCR and its associated CD3 signaling complex

A) *Molecular structure of the αβ T cell receptor (TCR) heterodimer. The TCR heterodimer consists of two transmembrane glycoprotein chains, namely α and β chains. Both chains are linked via a disulfide bond that is formed by cysteine residues present in short stalk segments of either chain, thereby forming the αβ heterodimer. Both carbohydrate peptide chains consist of an antigen-binding variable (V) and a constant (C) region with the latter being proximal to the cell membrane. Both TCR chains are anchored in the cell membrane via transmembrane domains ending up short cytoplasmic tails.* ***B)*** *The antigen-binding αβ TCR heterodimer is associated with 4 signaling chains (one δ, one γ, and two ε chains), collectively designated as CD3 complex, that are required for both proper cell surface expression and signaling. A homodimer of ζ chains is additionally attached to the TCR. Intracellular ITAMs (immunoreceptor tyrosine-based activation motifs) present in each CD3 and ζ chain facilitate the binding of SH2 domain-containing signal proteins for further signal transmission. CD3 complex, ζ chains, and the αβ heterodimer are associated via electrostatic interactions, respectively. Graphics and descriptions were adapted from Janeway's Immunobiology (8th edition)* [23,42].

2.3 The role of CD4+ and CD8+ T cells during tumor cell eradication

The idea that CD4+ and CD8+ T cells specifically recognize tumor cells and are therefore involved in tumor rejection is common for decades [43]. As summarized by Chen and Mellman, a potent anti-cancer immune response requires a stepwise series of events ultimately leading to effective tumor cell killing, referred to as the cancer-immunity cycle. Firstly, tumor cell antigens are released by dying cancer cells and are then captured by dendritic cells (DCs). Next, DCs present the captured antigens via MHC class I and class II to T cells resulting in the priming and activation of effector T cell responses against cancer-released antigens. As a result, activated T cells migrate to the tumor site and infiltrate the tumor bed.

Malignant cells are eventually recognized via specific TCR-pMHC interactions with their cognate antigens being presented by MHC molecules, thereby initiating the killing of target cancer cells. Due to the cancer cell lysis, even more tumor antigens are released to induce subsequent iteration loops of the cycle to eventually increase the breadth and depth of the immune response. However, excessive amplification of the entire cycle may provide anti-cancer activity but simultaneously bears the potential of autoimmunity. A critical balance in the ratio of T effector to inhibitive regulatory T cells (Tregs) is therefore crucial for the performance of immune responses [44].

Tumor-infiltrating immune cells are commonly observed in cancer patients, but the cell compositions involved in innate and adaptive immunity differ across both tumor-types and organ sites [45]. Here, immune cells are frequently organized in cell clusters, referred to as tertiary lymphoid structures (TLSs), which are found within tumor tissues and are scattered throughout the tumor bed and/or the peri-tumoral stroma [46,47]. Those TLSs contain mature DC/T cell zones and B cell follicles with reactive germinal centers and are therefore structurally quite similar to secondary lymphoid organs. The presence of TLSs was suggested to beneficially affect clinical outcomes in cancer patients with colorectal and lung tumors [48–50]. Generally, strong lymphocytic infiltrations seem to be associated with good clinical outcomes for various patients suffering from different tumor types like melanoma, breast, bladder, or lung cancer. High densities of tumor-infiltrating CD3+ T cells, CD8+ cytotoxic T cells, and CD45RO+ memory T cells were shown to be clearly associated with a longer disease-free survival and/or overall survival [51]. Moreover, meta data analysis revealed that high CD4+ T cell infiltration of the tumor stroma might correlate with better overall survival in lung cancer patients. In contrast, high densities of FOXP3+ Treg infiltrations were found to function as a negative prognostic factor for the clinical outcome of patients [52]. Tumor-infiltrating lymphocytes (TILs) had already been described by the Rosenberg group in 1986 and were characterized as a subpopulation of lymphocytes that are able to infiltrate growing cancers [53]. However, cellular compositions of TIL populations are very heterogenous exhibiting diverse proportions of CD4+ and CD8+ T cells, even when comparing independent cultures from the same lesion [54–56]. Therefore, both magnitude and phenotype of infiltrating lymphocytes may affect the clinical outcome and disease progression in cancer patients [57].

CD8+ cytotoxic T cells (CTLs) are considered the most powerful effectors in the anti-cancer immune response [58]. After TCR-mediated recognition of pMHC complexes present on potential target cells, CTLs secrete several toxic proteases and the pore-forming protein perforin into the immunological synapse formed together with the malignant target cell [59]. Perforin is a granule protein that multimerizes to generate pores in the membrane of target cells, thereby enabling the delivery of apoptosis-inducing proteases, known as granzymes, into the their cytoplasm [60]. Although research largely focused on CD8+T cell-mediated anti-tumor cytotoxicity in recent years, both CD8+ and CD4+ lymphocytes are involved in tumor rejection [61]. Absent CD4+ T cell help was shown to impair CD8+ T cell recruitment, proliferation, and effector function *in vivo* [62]. Furthermore, CD4+ T cells provide helper functions for B cell- and CD8+ CTL-mediated immune responses while additionally shaping the character and magnitude of the inflammatory response. In more detail, cytokines like IFN-γ and TNF-α are secreted by type 1 CD4+ T cells, thereby inducing increased overall antigen presentation of APCs and proliferation/activation of tumor cell-lysing CTLs. Additionally, interleukins like IL-5 and IL-6 are secreted by type 2 CD4+ T helper cells to support B cell proliferation and the development of primary antibody responses. Hence, CD4+ T cells are central players in activating and regulating both adaptive and innate immunity [48,57].

2.4 Challenges in T cell-mediated tumor immunity

2.4.1 Immune evasion mechanisms of cancer cells

The process of tumor cells being recognized and eventually killed by immune cells is known as tumor immunosurveillance. Immune evasion describes the capacity of cancer cells to actively prevent the

immune system from attacking. Here, a immuno-suppressive tumor microenvironment (TME) is established by malignant cells to hamper the overall anti-tumor immunity, thereby promoting tumor development, progression, and metastasis [63]. As recently reviewed by Kim and Cho, tumor cells actively induce T cell tolerance, thereby establishing an immunoediting process that facilitates tumor-redistribution and growth. Moreover, effective T cell immune responses are attenuated through various mechanisms, such as decreased absolute lymphocyte numbers, increased absolute counts of regulatory T cells (Tregs), downregulation of antigen expression in tumor-associated cells, or apoptosis of cytotoxic T cells (CTLs). Furthermore, cancer cells directly and indirectly attenuate the function of anti-cancer immunity by reducing Th1 responses necessary for CTL differentiation or by the secretion of immunomodulatory factors like prostaglandin E2, transforming growth factor-β (TGF-β), and IL-10 [64]. Additionally, proper function, proliferation and differentiation of TILs are highly depended on the metabolization of nutrients such as glucose, lipids, and amino acids present in the TME. Due to their enhanced ability to obtain nutrients, cancer cells may outcompete neighboring cells of the immune system for essential macromolecules, thereby restricting both proliferation and activation of tumor-reactive immune cells [65].

In theory, malignant cells can evade elimination literally at every point of the T cell-mediated anti-cancer immune response [44]. Efficient cancer cell recognition, activation of tumor-reactive T cells, and lymphocyte infiltration represent the early stages of the cancer-immunity cycle (see section 2.3), thereby characterizing efficient anti-tumor responses. Immune-inflamed tumors exhibiting high levels of T cell infiltration and IFN-γ signaling are generally designated as immunologically “hot” tumors. While often associated with a favorable disease outcome, even hot tumors can develop immune evasion mechanisms by, for example, interfering with the TCR-mediated pMHC complex recognition or the actual process of tumor cell killing [44,64,66]. This is typically achieved by downregulation or loss of MHC class I antigen presentation, attenuating tumor-infiltrating lymphocyte functions, or by establishing immuno-suppressive TMEs. Here, the crucial components of the antigen presentation machinery are targeted via several molecular mechanisms, such as the accumulation of genetic mutations, the loss of essential transcription factors, epigenetic silencing of regulatory elements, and altered signaling pathways directly affecting MHC class I expression or stability [67]. On the other hand, cancer cells are able to attenuate tumor-infiltrating lymphocyte functions by actively expressing immune checkpoint molecules, thereby resulting in T cell exhaustion mediated by several co-inhibitory receptors like PD-1 (Programmed Cell Death Protein 1), BTLA (B and T lymphocyte attenuator), LAG-3 (lymphocyte activation gene 3 protein), TIM3 (T cell immunoglobulin and mucin domain-containing protein 3), and TIGIT (T cell immunoreceptor with Ig and ITIM domains) [64].

Immunosuppressive TMEs favoring cancer progression and immune escape are typically composed of recruited immunosuppressive cells as well as soluble components, which together are capable to extensively inhibit the anti-cancer immune response. Various types of immunosuppressive cell types, such as myeloid-derived suppressor cells (MDSCs), regulatory T cells (Tregs), or tumor-associated M2 macrophages, play central roles in T effector cell inhibition. They typically inhibit T effector cell functions by either reducing the expression of cytotoxic factors or via the secretion of inhibitory factors, such as TGF-β or IL-10 [64,68]. In contrast to hot tumors, so called “cold” tumors are characterized by absent immune cell infiltration and typically bypass the immune recognition by avoiding full activation, migration, and tumor infiltration of T cells [44,64]. Full T cell activation is typically avoided by reducing the intrinsic immunogenicity of tumors, which either results from the loss of tumor antigen expression/presentation or from early elimination of highly immunogenic cancer cells during tumor formation [69,70]. In contrast, migration and actual tumor infiltration of lymphocytes is mostly inhibited by reducing the expression levels of chemokine receptor ligands like CXCL9/10/11, Fas ligand-induced apoptosis of infiltrating CD8+ T cells, the production of suppressive chemokines like CXCL12, or by physically preventing the interactions between T cells and malignant cells [64].

2.4.2 Central tolerance mechanisms may reduce the diversity of tumor-reactive T cells

Vertebrate adaptive immune responses require highly diverse TCR repertoires to efficiently control viral infections and other pathogens. Random rearrangements of *αβ TCR* gene segments via V(D)J recombination (see section 2.2.2) generate highly diverse T cell repertoires that are essential for the efficiency of the immune system. However, a healthy immune system requires T cells responding to foreign antigens while remaining tolerant to self-antigens [71–73]. Positive and negative selection of thymocytes prevent autoimmunity by ensuring that mature T cells are capable of recognizing foreign peptides in the context of self MHC molecules while not reacting to self-peptide/MHC complexes. Central tolerance mechanisms include the inactivation and deletion of self-reactive thymocytes during T cell maturation, thus being highly important for preventing self-reactive T cells from damaging healthy tissues [74–76].

Several tumor-antigens and CG (cancer germline; also cancer-testis, CT) antigens, such as various members of the MAGE family, NY-ESO-1, MART-1, or tyrosinase, have been demonstrated to be expressed on a mRNA-level in medullary thymic epithelial cells, thereby facilitating negative selection of self-recognizing T cells [77]. Hence, while preventing aberrant T cell responses against self-antigens, immunological tolerance mechanisms may also limit diversity, repertoire, and efficiency of potentially tumor-reactive T lymphocytes. This dilemma eventually resulted in countless studies aiming to identify T cell target antigens abundantly expressed in cancerous cells but not in any non-malignant tissues [76,78]. Nevertheless, cellular and humoral immune responses targeting some of these thymus-expressed antigens can regularly be observed in various cancer patients. Hence, the extent to which central tolerance mechanisms may limit therapy approaches targeting those antigens is largely conjectural [79–81].

2.5 Tumor-specific vs. tumor-associated antigens

Tumor antigens can possess wild-type (self) as well as altered (non-self) amino acid sequences and emerge during all stages of protein synthesis, protein degradation, post-translational modification, or due to alterations in the antigen processing and presentation machinery. Based on whether they are exclusively present on tumor cells or also found in non-malignant tissues, tumor antigens can be classified into either of two subgroups: Tumor-associated antigens (TAAs) and tumor-specific antigens (TSAs) [82].

2.5.1 TAAs: Potent targets for T cell-based immunotherapies

Tumor-associated antigens (TAAs) are normal host proteins that are shared among patients and across various cancers. Furthermore, TAAs usually exhibit distinct expression profiles in both healthy and malignant tissues [78]. The approach of identifying universal antigens capable to trigger broad immune responses against different tumor types is not new. However, targeting TAAs with off-the-shelf anti-tumor treatments has major limitations, the first one being that most TAA epitopes emerge from self-peptides presented by MHC proteins. Hence, T cells recognizing those antigens are often eliminated by the process of central tolerance (see section 2.4.2) [83]. Furthermore, most TAAs are also expressed, albeit at low to intermediate levels, in one or more healthy tissues. Targeting those antigens raises safety concerns due to on-target/off-tumor toxicities with unpredictable severity [84].

Generally, the TAA subgroup is composed of antigens with rather low tumor-specificity, such as (i) differentiation antigens and (ii) overexpressed antigens, as well as antigens with high tumor-specificity like (iii) cancer testis antigens and (iv) endogenous retroelement antigens [78,85]. Differentiation antigens (i) are exclusively present on tumors and their respective tissues of origin [78]. Spontaneous T cell responses were found against peptides derived from GP100, Melan-A/MART-1, and tyrosinase in melanoma patients [80,86–88]. CD19, which is exclusively expressed throughout B cell development and in the context of nearly all B cell malignancies, is another example for a very prominent differentiation

antigen. Interestingly, CD19 can actually be targeted via adoptive cell transfers of autologous lymphocytes expressing anti-CD19 chimeric antigen receptors (CARs) [89]. To circumvent potential off-target antigenicity, epitopes exclusively expressed during early ontogeny could represent the most promising type of differentiation antigens for future treatments [78].

In contrast to differentiation antigens, overexpressed TAAs (ii) are aberrantly expressed proteins conferring growth and/or survival advantages to malignant cells. Prominent examples for overexpressed TAAs are the zink-finger transcription factor encoded by the *Wilms tumor 1* (*WT1*) gene and the Human epidermal growth receptor 2 (HER2). WT1 is thought to be involved in blood cell development of early CD34+ precursors and is additionally found to be highly expressed in most leukemias [90]. HER2 is a transmembrane glycoprotein expressed in many tissues and triggers intracellular signaling cascades leading to cell proliferation, cell differentiation, and many more cellular responses. However, overexpression of the *HER2* gene was found in various cancers, such as breast cancer, metastatic adenocarcinoma of the esophagus, and many more [91,92]. As previously shown by the Rosenberg group, administration of T cells expressing an anti-HER2 CAR induced severe side effects in a patient suffering from metastatic HER2-overexpressing colon cancer and resulted in death only 5 days post treatment. They postulated that the patient's death occurred since the transferred T cells also recognized HER2-expressing normal lung cells, thereby highlighting the risk of on-target/off-tumor toxicities when using therapy approaches targeting TAAs [78,93].

Cancer germline/testis (CG, CT) antigens (iii) are another prominent family of TAAs found to be ectopically expressed across various human cancers. However, they are not present in normal tissues except for testis, placenta, fetal ovaries, or trophoblasts. Due to their highly tumor-restricted expression and their immunogenicity, CG antigens represent very promising targets for tumor-specific immunotherapy approaches [94,95]. Thierry Boon and colleagues identified the first CT antigen in 1991: MAGE-1 (melanoma antigen 1), later renamed MAGE-A1, was found to be specifically recognized by autologous CTLs of a melanoma patient. Additionally, MAGE-A1 was verified to be highly expressed in various tumor types but was absent in healthy tissues other than testis and placenta [96,97]. According to the CG antigen database established by the Ludwig Institute for Cancer Research, there are more than 100 listed gene families that are thought to be upregulated in cancer with their normal expression strongly biased to the germline [98]. Clinical studies using TCR-engineered T cells (TCR-T cells) for targeting CG antigens like NY-ESO-1 (New York esophageal squamous cell carcinoma-1) or CEA (Carcinoembryonic antigen) have shown promising results by inducing tumor regressions in metastatic colorectal cancer as well as melanoma patients [99,100]. However, another trial using affinity-optimized anti-MAGE-A3 TCR-T cells induced critical side effects, such as temporary Parkinson-like symptoms and actual comas, which eventually lead to death in two out of 9 treated cancer patients. Those neurologic toxicities were most likely due to the cross-recognition of MAGE-A12, a TAA expressed by a subset of neurons in the human brain. This study impressively demonstrated that caution needs to be exercised when targeting CG antigens with high avidity TCRs since reported tumor-specificities of target antigens may not always be correct [101].

The concept of human endogenous retroviruses (HERVs) being connected to neoplastic transformation (iv) is not new. Approx. 8% of the human genome is comprised of retroviruses which infected human ancestors and, due to the accumulation of post-insertional mutations, lost the ability to produce replication competent viral particles. Moreover, expression of those HERVs is usually tightly repressed by epigenetic mechanisms, such as DNA methylation and histone modification [102]. Due to epigenetic dysregulation during malignant transformation, cancer cells are especially prone to the reactivation of HERV elements, thus representing a valuable pool of potentially immunogenic TAAs [85]. Although HERV-expression in the thymus is not well characterized, CD8+ T cell responses against HERV-derived antigens had been reported in the past. Hence, the immunological tolerance against retroelements seems incomplete and might be explained by very low expression of retroelements in the thymus. Nevertheless, as previously described for other self-proteins, thymic selection is expected to reduce the repertoire of retroelement-reactive T cells by establishing immunological tolerance [85,103–105].

2.5.2 TSAs: Ideal targets for cancer immunotherapy

2.5.2.1 Neoantigens arising from tumor-specific SNVs, MNVs, InDels, and gene fusions

Neoantigens arising from somatic non-synonymous mutations or aberrantly transcribed/translated gene products are, by definition, exclusively present in tumor cells and account for one of the most common classes of tumor-specific antigens (TSAs) [78,106,107]. In addition to their tumor-specific expression, neoantigens are also considered to bind self-MHC complexes with high affinity, thereby representing attractive targets for precision immunotherapies such as cell-based therapeutics [4,78,108]. However, peptide binding to MHC molecules is highly dependent on interactions between the anchor amino acid residues of a particular peptide and respective MHC binding pockets (see section 2.2.1). Due to potentially increased MHC class I-binding affinities, several studies suggest that high numbers of neoantigens with mutated anchor residues might be a solid determinant for cancer peptide immunogenicity and might also correlate with potentially better prognosis than merely high neoantigen burden [6,109–111]. However, since data on neoantigen-specific TCRs is scarce and since intrinsic TCR affinities may also be affected by respective HLA restrictions, additional studies are needed to further analyze affinity differences among TCRs recognizing neoantigens with mutations in anchor and non-anchor residues [6].

Besides from amino acid-changing single nucleotide variations (SNVs) or multiple nucleotide variations (MNVs), neoantigens can also arise from base pair insertions and deletions (InDels) that induce novel non-synonymous open reading frames. Frameshift mutations were found to be associated with the number of detected tumor-infiltrating lymphocytes (TILs), e.g. in colorectal cancers exhibiting microsatellite instabilities [106,112,113]. Interestingly, InDels were shown to generate three times as many neoantigens when compared to the amount of neoantigens arising from non-synonymous renal cell carcinoma SNVs. Furthermore, as compared to non-synonymous SNVs, neopeptides derived from indel mutations were nine times enriched for mutant-specific MHC class I-binding candidates. Therefore, frameshift mutations represent a highly immunogenic mutational class with the potential to trigger increased abundance of neoantigens with potentially high HLA binding capacity [114].

Moreover, conserved gene fusions generating functional chimeric genes represent another prominent source for the origin of targetable neoantigens and are primarily found in leukemias or some soft-tissue tumors, such as desmoplastic small round cell tumors, several sarcoma subtypes, breast or lung cancers, and many more. Affected genes often encode transcription factors, thereby enabling the emerging fusion products to inappropriately affect gene expression and contribute to malignant cell transformation [115]. The characteristic genetic abnormality of chronic myelogenous leukemia (CML), the Philadelphia chromosome, is the most well-characterized example of a fusion gene involved in tumor generation. It results from a reciprocal translocation involving chromosomes 9 and 22, thereby resulting in the generation of the BCR-ABL fusion protein, a constitutively activated tyrosine kinase present in virtually all CML patients [116]. Besides of being conserved in particular cancer subtypes, gene fusion proteins have also been found to be shared between tumor types [117]. When compared to SNVs and MNVs, out-of-frame gene fusions might, due to their size, provide higher chances of encoding neoepitopes with high levels of dissimilarity to self-antigens [118]. Thus, together with the fact that gene fusions are often the driver mutations for the respective malignancies, they represent promising antigens that are potentially targetable by using universal neoantigen-based therapeutic strategies [106].

2.5.2.2 Alternative TSAs: Oncoviral proteins, TEIPPs, and many more

As described above, the most common neoantigens arise from tumor-specific mutations either inducing amino acid changes or frameshift mutations within cancer genomes. However, there are additional categories of TSAs not originating from tumor-specific mutations: Oncoviral antigens are, for example, derived from viruses that are able to drive oncogenic cell transformation, thereby resulting in various malignancies. Hence, HLA-presented peptides derived from viruses such as Epstein-Barr virus (EBV) and human papillomavirus (HPV) can be targeted via T cell-based treatments. Since oncoviral antigens

lack expression in healthy tissues and since they are regularly shared among tumor types, they are considered tumor- but not patient-specific [119–121].

In addition, there is another class of tumor-specific antigens, referred to as "T cell epitopes associated with impaired peptide processing" (TEIPPs), which are derived from non-mutated housekeeping proteins of tumors exhibiting defects in the conventional TAP-dependent peptide processing pathway. Targeting TEIPPs therefore offers the opportunity to target tumor cells which evade the immune system by acquiring defects in the HLA class I-dependent antigen processing machinery [106,122]. Although data about human TEIPP antigens is scarce, there is an example of an tumor-derived preprocalcitonin signal peptide-derived epitope that is recognized by tumor-specific CTLs and whose surface expression requires the loss/downregulation of TAP [122,123].

Other studies demonstrated autologous CTLs of cancer patients to specifically recognize antigens in the context of point-mutated HLA molecules that are exclusively expressed on patients' tumor cells. However, it is not quite clear whether mutant HLAs just bind different sets of immunogenic peptides or whether T cells are able to directly interacted with the altered HLA molecule in a peptide-independent manner [124]. A more recent study supposed that autologous CTLs from a metastatic melanoma patient might recognize an unknown peptide, or set of peptides, in the context of a mutated HLA-A*11 molecule only expressed in tumor cells [125]. Targeting novel antigens specifically bound by tumor-specific HLA isoforms could therefore represent another promising strategy for future cancer therapies [106].

Another group of TSAs is derived from post-transcriptional modifications, such as splice variant antigens, alternative ribosomal products, or post-translational peptide splicing. Splice variant antigens arise from alternative splicing events, including mRNA splice junction mutations, intron retention, and dysregulation of the splicing machinery during tumorigenesis [106]. Mutations in spliceosome-related genes are, for example, commonly found across the entire spectrum of myeloid malignancies [126,127]. By analyzing alternative splicing events across 32 different cancer types listed in the Cancer Genome Atlas (TCGA), Kahles and colleagues identified thousands of tumor-specific alternative splicing events (data obtained from The Genotype-Tissue Expression project, GTEx) with an average of approx. 930 novel exon-exon junctions, referred to as neojunctions, per tumor sample. Furthermore, they impressively revealed the generation of translatable RNA transcripts arising from tumor-specific alternative splicing events. By also producing peptides that could potentially be presented by MHC molecules, those proteins may represent a suitable source for future immunotherapy target antigens [128]. However, TSAs originating from post-translational peptide splicing or from alternative ribosomal products are difficult to apply in the context of anti-cancer therapies, mainly due to the lack of reliable bioinformatic tools for efficiently predicting potential antigens arising from those categories [106].

The same might be true for cryptic peptides arising from non-canonically transcribed/translated genome regions, also referred to as the "dark matter" of the tumor proteome [129]. While up to 75% of the human genome can be transcribed and potentially translated, tumor antigen discovery approaches have so far largely focused on the immunogenicity of annotated proteins or non-synonymous mutations in coding regions, thereby only accounting for approx. 2% of the genome [130]. Besides of not being extensively investigated, cryptic epitopes originating from non-coding regions or off-frame translation of coding sequences have previously been demonstrated to be recognized by melanoma patient-derived TIL populations and *in vitro*-sensitized peripheral blood lymphocytes (PBLs) [130–133]. Estimates indicate that approx. 10% of all MHC class I-associated peptides originate from allegedly non-coding genomic sequences and that a fraction of those aberrantly translated peptides might be specifically presented on tumor cells, thereby potentially expanding the repertoire or targetable tumor antigens. However, the true quantity, immunogenicity, and expression of cryptic antigens in cancer remains largely unexplored [130,134]. Although tandem mass spectrometry (MS) or RiboSeq, a NGS-based method for determining actively translated RNA molecules, theoretically provide the necessary technical foundations for the detection of cryptic peptides, the efficient identification of such antigens remains challenging [82]. cDNA library screenings were previously demonstrated as potent strategy to identify target epitopes arising from non-coding intronic sequences [132,133]. Additionally, to illuminate the dark side of the

immunopeptidome for ongoing neoepitope discovery, RNA-Seq- and MS-based bioinformatic workflows had recently been established to specifically identify peptides generated from non-canonical mRNA translation [135,136]. Therefore, further exploring the dark matter of the human genome might provide novel target antigens and biomarkers for future immunotherapies, thereby simultaneously impacting future cancer diagnosis and unravelling new mechanisms of tumorigenesis [129].

2.6 Treatment guidelines for non-small cell lung cancer (NSCLC)

2.6.1 Lung cancer subtypes

Lung cancer is the leading cause of cancer-related death worldwide and is broadly divided into the two major groups of small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). SCLC accounts for approx. 15% of all lung cancers, is more aggressive than NSCLC, and remains one of the most lethal forms among malignancies (overall 5-year survival rate <7%) [137,138]. On the other hand, NSCLC accounts for approx. 85% of all lung cancers and can be grouped into three major histologic subtypes (large-cell lung cancer, squamous cell carcinoma, adenocarcinoma) differing in location, cell origin, and growth patterns. Adenocarcinomas and squamous cell carcinoma together account for approx. 80% of NSCLC cases with adenocarcinoma being the most common subtype. Adenocarcinoma generally accounts for roughly half of all lung tumors while squamous cell carcinomas represent about 30% of lung cancers (Fig. 2.6 A). Large-cell lung cancer is a very heterogeneous group of tumors harboring features of adenocarcinoma, squamous cell carcinoma, and SCLC. Moreover, squamous cell carcinoma and SCLC are most strongly associated with smoking. In addition, there are minor lung cancer subtypes with rather unspecified histology or clinical characteristics that are barely distinguishable [139,140].

2.6.2 Treatment guidelines for stage I to stage III NSCLC

Clinical outcomes in lung cancer patients are directly related to the cancer stage at the time of diagnosis. For example, early identification of NSCLC and surgical resection of small, localized, stage I tumors offers favorable prognosis with 5-year overall survival rates of 70-90%. However, patient survival remains poor in patients suffering from advanced NSCLC usually accounting for 5-year overall survival rates between 26% in stage IIIB, about 10% in stage IVA, and even less in stage IVB patients [138,141,142]. Surgical resection followed by optional adjuvant systemic therapy is the standard strategy for the treatment of stage I, stage II, and some stage IIIA non-small cell lung cancers. Today, adjuvant chemotherapy typically consists of 4 cycles of a cisplatin-based doublet regimen and is recommended for the treatment of completely resected stage II to stage IIIA NSCLCs [142,143]. Moreover, recent data indicate that further adjuvant immunotherapy with the anti-PD-L1 monoclonal antibody atezolizumab might represent another promising treatment strategy for patients with resected stage II-IIIA NSCLC [144].

Surgery of stage III tumors is generally considered for NSCLC patients with a single mediastinal lymph node involvement and is regularly combined with both chemotherapy and radiation (trimodality treatment). However, a vast number of stage III NSCLC patients are deemed unresectable, which is either due to tumor burden, patient choice, or because they are considered poor candidates for surgery [142]. Multimodality therapy consisting of concurrently administered chemoradiotherapy (CRT) represents the standard of care for patients with unresectable stage III NSCLC, but it results in rather poor 5-year survival rates of approx. 15 to 30%. Immunotherapy with durvalumab, a selective high-affinity human IgG1 monoclonal antibody specifically blocking programmed death-ligand 1 (PD-L1), was found to substantially improve overall survival and progression-free survival in stage III NSCLC patients already treated with CRT. Durvalumab application was therefore approved for the standard treatment of unresectable stage III NSCLC showing no progression after concurrent CRT, thereby becoming the new

standard of care in this setting [145,146]. Furthermore, when compared to chemotherapy alone, recent data indicate that the combination of perioperative durvalumab administration and neoadjuvant chemotherapy might increase event-free survival in stage IIIA NSCLC patients [147].

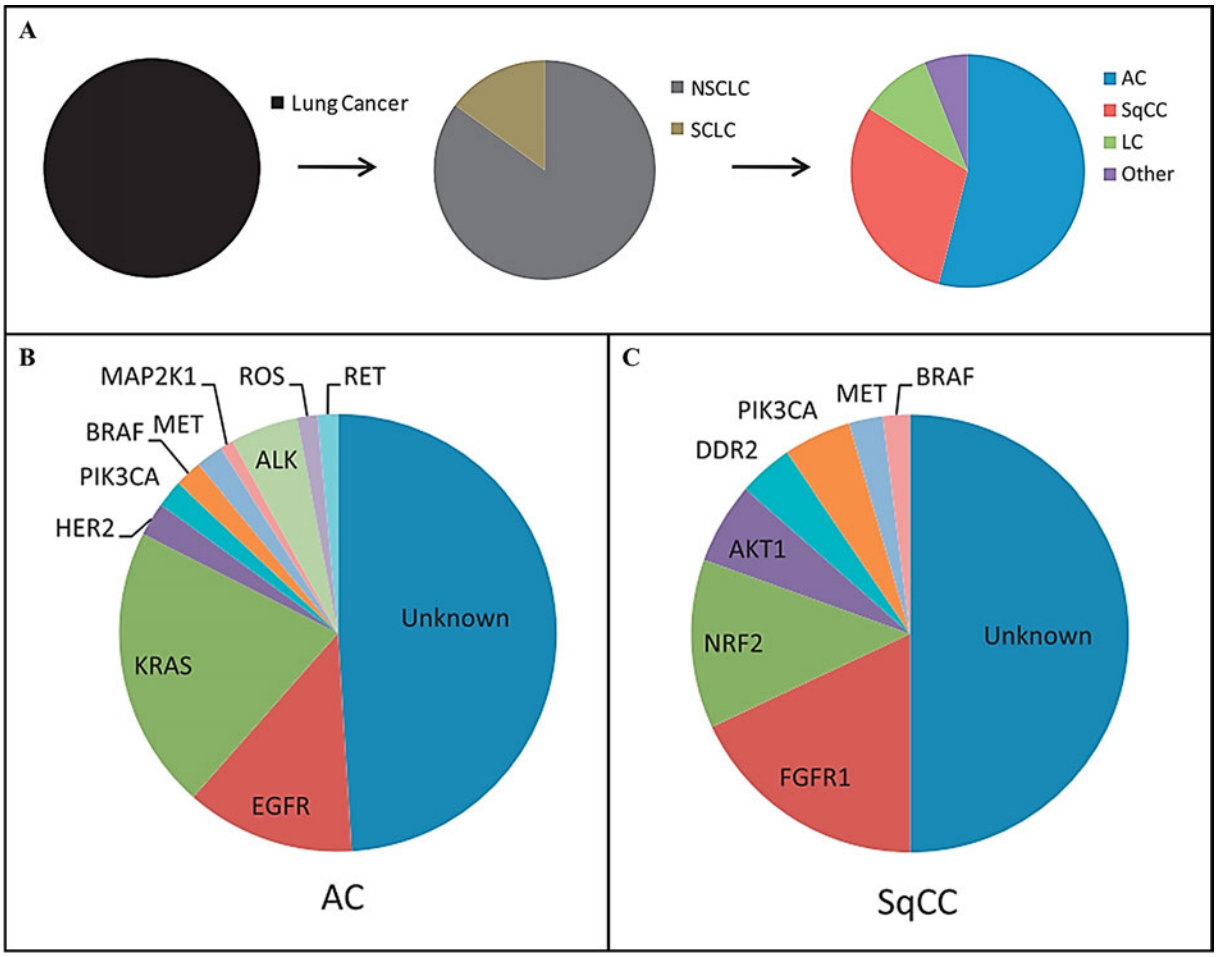


Figure 2.6: Lung cancer subtypes and corresponding oncologic driver genes.
A) *SCLC and NSCLC are the two major subtypes of lung cancer. NSCLC accounts for approx. 85% of all lung cancers while SCLC represents approx. 15% of all cases. NSCLC is divided into the three major histologic subtypes: Large-cell lung cancer, squamous cell carcinoma, and adenocarcinoma. The designation "other" encompasses minor histology types, such as non-specified NSCLC, adenosquamous carcinomas, and sarcomatoid carcinomas.* ***B and C)*** *Molecular subtypes of adenocarcinoma and squamous cell carcinoma. Pie charts show the proportions of patient populations with a particular histological subtype. Only targetable and mutually exclusive hyperactive oncogenes are shown. Alterations in ALK (ca. 4-6%), EGFR (ca. 10-15%), and KRAS (ca. 21%) are the predominant driver mutations found in adenocarcinomas. In contrast, most common targetable driver mutations in squamous cell carcinoma are found in FGFR1 (ca. 16-20%) and NRF2 (ca. 10-15%).* ***Abbr.:*** *AC = adenocarcinoma; ALK = Anaplastic Lymphoma Kinase; EGFR = Epidermal Growth Factor Receptor; FGFR1 = Fibroblast Growth Factor Receptor 1; KRAS = Kirsten Rat Sarcoma viral oncogene; LC = large-cell carcinoma; NRF2 = Nuclear factor erythroid 2-related factor 2; NSCLC = non-small cell lung cancer; SCLC = small cell lung carcinoma; SqCC = squamous cell carcinoma. Graphics and descriptions were adapted from L.A. Pikor and colleagues (Lung Cancer 82 (2013); 179-189)* [139].

2.6.3 Treatment strategies for metastatic stage IV NSCLC

The main therapy goal for the treatment of metastatic NSCLC is to improve or maintain quality of life while simultaneously prolonging overall survival of patients. Despite of platinum-based chemotherapy, patients with advanced NSCLC have very poor prognosis with relatively low 5-year survival rates of about 7% when distant metastases are present. Generally, treatment options include chemotherapy, targeted therapy, and immunotherapy. Targeted therapy can be applied in patients bearing drug-sensitizing oncogenic driver mutations and has the potential to elicit remarkable responses in several cancer subtypes, thereby improving patient outcomes [148,149]. Patients who do not possess any druggable driver mutations are routinely treated with single-agent immunotherapy, combination immunotherapy,

or chemotherapy alone. Furthermore, chemotherapy is the first-line option for patients that may not be suitable for immunotherapy treatments. This can be due to pre-existing autoimmune conditions, concerns about a patient's performance status, or the risk of toxicity. In these settings, platinum doublets are typically used for the treatment of both non-squamous and squamous metastatic NSCLC [142].

2.6.3.1 Targeted therapy for NSCLC

Non-squamous tumors of patients with absent or light smoking history should regularly be tested for specific driver mutations, which can be present in up to 60% of lung cancer patients and are, in case of NSCLC, most frequently found in non-smoking-related adenocarcinomas. When obtained from patients with smoking history, tumors of squamous histology may be considered for driver mutation analysis on a case-by-case basis [142,150,151]. Therapy approaches targeting common NSCLC driver mutations in Epidermal Growth Factor Receptor (EGFR), anaplastic lymphoma kinase (ALK), or ROS Proto-Oncogene 1 (ROS1) provide significant increases in both survival and life quality of patients with advanced lung cancer and rather poor prognosis [152]. By targeting biologically relevant alterations as competitive inhibitors within ATP-binding groves, especially tyrosine kinase inhibitors (TKIs) can provide improved therapeutic responses and clinical benefits for NSCLC patients [139]. Therefore, new guidelines recommend the testing of patient tumors for targetable driver mutations in *EGFR*, *ALK,* and *ROS1* genes. Furthermore, genetic alterations found in other common drivers like *KRAS* could additionally be included in future routine tumor mutation analyses [152]. *EGFR* variations, *KRAS* mutations and *ALK* rearrangements are the three predominant drivers found in adenocarcinomas and occur in approx. 35-40% of all cases (Fig. 2.6 B). In contrast, they are less common in squamous cell carcinomas where they are found with frequencies of 6% and below. Here, targetable driver mutations are most commonly found in *FGFR1* (ca. 16-20%) and *NRF2* (ca. 10-15%), respectively (Fig. 2.6 C) [139].

2.6.3.2 Immunotherapies: Current strategies for the treatment of NSCLC

Immunotherapy has significantly changed the therapeutic landscape for the treatment of patients with metastatic NSCLC. Patient responses to immunotherapy mainly depend on tumor expression of tumor programmed cell death ligand 1 (PD-L1), which can broadly be classified into three subcategories: PD-L1-negative (<1% of tumor cells express PD-L1), PD-L1 low positive (1-49% of tumor cells express PD-L1), and PD-L1-positive (>50% of tumor cells express PD-L1). Together with its corresponding immune checkpoint receptor, namely programmed cell death protein 1 (PD-1), PD-L1 facilitates the activation of immunosuppressive signaling pathways resulting in both induction and maintenance of immune tolerance within the tumor microenvironment (TME). Generally, patient survival and duration of response seem higher in patients with tumors exhibiting increased PD-L1 expression [142,153].

In 2015, a first phase II study demonstrated the clinical efficacy of nivolumab, a fully human IgG4 anti-PD-1 immune checkpoint-inhibiting antibody, in pre-treated patients with refractory squamous NSCLC [154]. Furthermore, it was shown that treating metastatic NSCLC patients with pembrolizumab, another monoclonal antibody targeting the immune checkpoint PD-1, as a first-line therapy induced clinically meaningful long-term overall survival benefits when compared to standard chemotherapy. Both median overall survival and 5-year survival rates of patients with PD-L1-positive tumors were found to be substantially increased upon pembrolizumab administration [149]. Similar results were achieved when comparing the immune checkpoint inhibitors cemiplimab (anti-PD-1) and atezolizumab (anti-PD-L1) with conventional chemotherapy in NSCLC patients with highly PD-L1-positive (>50% of tumor cells) cancers [155,156]. For the first time, these studies indicated that chemotherapy might not be mandatory for all patients with advanced NSCLC.

Combinational treatment strategies with chemotherapy and pembrolizumab were shown to provide clinical benefits for metastatic NSCLC patients with tumors exhibiting PD-L1 expression lower than 50% [157]. Furthermore, therapeutic strategies, including chemotherapy and dual immune checkpoint inhibition with nivolumab and ipilimumab, a fully human anti-cytotoxic T-lymphocyte antigen-4

(CTLA4) antibody, were demonstrated to provide durable benefits in advanced NSCLC patients [158]. Head-to-head comparisons of single and dual immunotherapy approaches could therefore provide new insights for the future treatment of particular patient subgroups [142]. BioNTech's FixVac (fixed vaccine) candidate BNT116 is another very promising approach for treating advanced or metastatic NSCLC patients in the future. BNT116 is an off-the-shelf immunotherapy candidate aiming at the RNA-based induction of immune responses to target 6 pre-defined tumor-associated antigens, thereby, in theory, covering up to 100% of patients included in either of the major histologic NSCLC subtypes [159]. The safety profile and safe dose for BNT116 monotherapy, as well as for BNT116 in combination with immune checkpoint inhibition, is currently tested in a first-in-human phase I study (LuCa-MERIT-1 clinical trial; NCT05142189). An graphical overview listing all FDA-approved immunotherapies for the treatment of NSCLC was recently provided by Shields and colleagues [160].

By using immunotherapy for the treatment of NSCLC patients, monitoring and management of treatment-induced toxicities deservers special attention. In general, immune related adverse events (irAEs) are less predictable than for classic chemotherapy since toxicities can occur at any point, can affect any organ, and are even possible after stopping the immunotherapy treatment. Since immunotherapy agents usually have longer half-lives, commonly used dose-reductions or therapy stops are not routinely prescribed for the management of emerging side effects. In fact, prompt immunosuppression may be crucial to reduce both morbidity and mortality in patients suffering from serious irAEs. Rather common irAEs are rash, pruritus, diarrhea and thyroid dysfunction [142]. However, immune checkpoint inhibition (ICI) can also induce severe toxic events, including colitis, pneumonitis, myocarditis, hepatitis, or neurotoxic effects. Metadata analysis revealed toxicity-related fatality rates of 0.6% for ICI-treated patients with the median time from symptom onset to death being 32 days. Depending on the type of ICI treatment, toxicity-related fatality rates ranged between 0.36% for anti-PD-1 monotherapy and 1.23% for combined (anti-PD-1/PD-L1 and anti-CTLA4) immunotherapy [161]. As revealed by another metadata analysis, the use of ipilimumab might especially be associated with increased risk of treatment-related mortality in cancer patients [162]. Although ICI can induce severe toxic effects, there is also data indicating that the development of irAEs could be used as a positive predictor for survival of NSCLC patients treated with nivolumab monotherapy. For example, Ricciuti and colleagues demonstrated that patients experiencing two or more irAEs had significantly longer progression-free survival and overall survival than patients which developed only one or no irAEs at all [163].

2.7 T cell-based immunotherapies

2.7.1 Classification of T cell-based immunotherapies

Over the past decade, two types of immunotherapies have emerged for the effective treatment of cancer: Firstly, the use of immune checkpoint inhibitors specifically targeting molecules that inhibit T cell activation (e.g. PD-1, PD-L1 or CTLA-4), thereby enhancing natural anti-tumor activity (see section 2.6.3.2). Secondly, the administration of anti-tumor immune cells by using adoptive cell therapy (ACT) to generate robust immune responses via the infusion of *ex vivo*-manipulated T cells. T cell-based ACT strategies include (i) the isolation and application of autologous tumor-infiltrating lymphocytes (TILs) and (ii) the genetic modification and subsequent administration of peripheral T cells to allow tumor recognition [164]. Personalized cell treatments such as adoptive T cell therapies are individually generated for every patient by using autologous cellular starting material. For example, TIL-based ACTs comprise autologous patient TILs that are reinfused after excessive *ex vivo* expansion. On the other hand, TCR-T cell approaches use autologous patient T cells genetically engineered to express tumor-reactive TCRs specifically recognizing MHC-presented tumor antigens. In contrast, genetically modified CAR-T cells express artificial chimeric antigen receptors (CARs) to recognize tumor-expressed cell surface antigens, thereby not depending on the presence of specific MHC alleles [165].

2.7.1.1 TIL-based ACTs

Generating TIL-based adoptive cell therapies (ACTs) is a multistep process comprising TIL isolation, expansion, and subsequent reinfusion (see Fig. 2.7.1 A). Firstly, isolated tumor tissues are dissected into fragments that are either used for further processing, for establishing single tumor cell suspensions, or are grown in the presence of IL-2. The latter will, due to overgrowing and death of included tumor cells, result in relatively pure TIL cultures within two to three weeks. By using autologous tumor cell suspensions for reactivity screening, TIL cultures can be selected based on tumor-reactivity and cytotoxicity. Here, selected TILs are usually subjected to a rapid expansion protocol (REP) comprising the non-specific stimulation via an anti-CD3 monoclonal antibody, the presence of irradiated feeder cells, and high doses of IL-2. Thus, autologous TIL cultures can be expanded to numbers of up to 2E+11 total cells and are eventually reinfused into patients. Typically, those patients had before been subjected to lymphodepleting regimens to increase both the persistence of transferred TILs and overall clinical responses after TIL therapy [164]. Major toxicities during TIL-based ACT are largely related to high dose IL-2 administrations after cell transfer. The applicability of TIL-based ACTs could therefore be improved by using attenuated IL-2 regimens, which were previously demonstrated to be still sufficient for inducing long-lasting responses in, for example, patients with treatment-refractory melanoma [166].

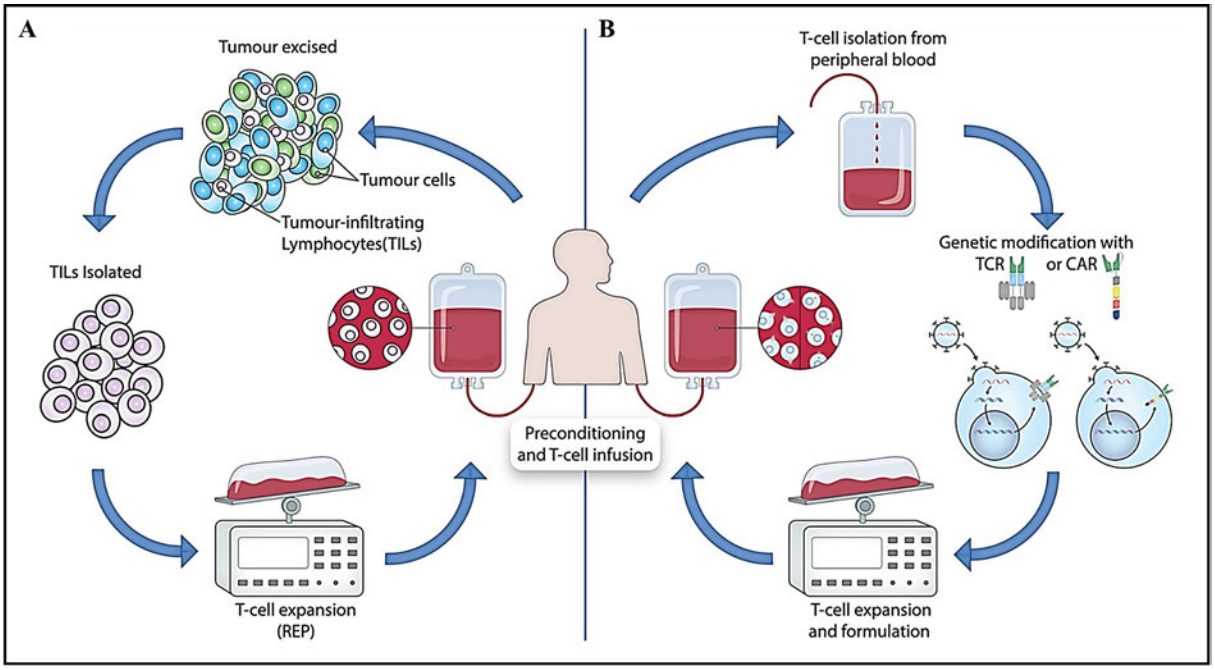


Figure 2.7.1: Adoptive T cell transfer approaches harnessing the human immune system to fight cancer.
A) *Adoptive transfer of tumor-infiltrating lymphocytes (TILs) isolated from a patient's tumor tissue. TILs are isolated from resected patient tumor samples and are then expanded in vitro by using a rapid expansion protocol (REP). Autologous TILs are eventually reinfused into the lymphodepleted patient.* ***B)*** *Adoptive transfer of genetically modified T lymphocytes. T cells are isolated from a patient's peripheral blood, are then expanded in vitro, and are eventually modified to either express a tumor-specific T cell receptor (TCR) or a chimeric antigen receptor (CAR). Expanded and genetically modified T cell populations are eventually reinfused back into the lymphodepleted patient. Graphics and descriptions were adapted from Ö. Met and colleagues (Seminars in Immunopathology (2019); 41:49-58)* [164].

The concept of transferring autologous TILs to treat human cancer is not a novel approach and was initially presented by Rosenberg and colleagues in 1988 [167]. They later demonstrated the power of reinfusing *in vitro*-expanded autologous TILs along with IL-2 into lymphodepleted metastatic melanoma patients, thereby managing to achieve durable complete tumor regressions in 22% and objective responses in 56% of all treated patients [168]. Several studies confirmed that adoptive transfers of autologous TILs can result in durable and complete responses, even after other immunotherapies have failed [166,169]. Furthermore, the potential power of TIL-based ACTs as follow-up treatment of ICI-resistant tumors was highlighted by demonstrating the ability of tumor-reactive T cells to heavily

infiltrate tumor microenvironments of metastatic melanoma patients, even after disease progression during anti-PD-1 or combined anti-PD-1/CTLA-4 treatment [170]. However, ACTs with *in vitro*-expanded TILs induced less effective long-term tumor regressions in other solid tumors such as metastatic ovarian cancer [171]. General efficiencies of TIL-based ACTs could theoretically be increased by previously identifying and enriching tumor-recognizing TILs, thereby reducing the number of transferred non-tumor-reactive "bystander" TILs [2].

2.7.1.2 ACTs based on genetically modified T cells: TCR-engineering and CARs

Genetic modification of autologous T lymphocytes is used to generate tumor-specific T cell therapies with potentially enhanced anti-tumor effects. Here, T cells are genetically modified to specifically target tumor antigens by either expressing a cloned TCR or a synthetic chimeric antigen receptor (CAR). Common methods for establishing those transgenic T cell cultures include transient RNA transfection, retroviral and lentiviral transductions, transposons, or homologous recombination. While there are various protocols available for establishing transgenic T cell cultures, a general outline is depicted in Fig. 2.7.1 B. Briefly, T cells are collected via leukapheresis from a patient's peripheral blood, are then activated, genetically modified, expanded, and eventually reinfused back into the lymphodepleted patient [164].

In a first study, Morgan and colleagues utilized a retrovirus-based transduction approach to generate transgenic T cells expressing a TCR specifically recognizing a MART-1 peptide presented by HLA-A*02:01. This TCR had previously been cloned from a TIL clonotype of a metastatic melanoma patient with nearly complete remission after TIL-based ACT. Hence, Morgan and colleagues managed to induce high sustained levels of circulating transgenic T lymphocytes and full clinical regression in two patients with rapidly progressive metastatic melanoma [172]. Multiple follow-up trials demonstrated significant and prolonged tumor regressions by targeting various tumor antigens in different tumor entities [164], such as gp100 (melanoma), NY-ESO-1 (melanoma, synovial sarcoma), and the carcinoembryonic antigen (colorectal carcinoma) [173–175].

In contrast, the genetic modification of T cells to express CARs combines the specificity of antibody-based antigen recognition with the cytotoxic properties of T lymphocytes. The antigen-binding domain of a CAR is composed of a single-chain fragment variable region derived from an antigen binding domain of antibodies. T cell activation is facilitated by fusing the antigen-binding domain to the CD3ζ transmembrane domains as well as the intracellular signaling domains of the TCR-complex. In case of second and third generation CARs, additional intracellular costimulatory domains, such as CD28, 4-1BB, or OX-40, are added to induce additional activation signals, thereby providing the potential to elicit potent anti-tumor activities and meaningful clinical responses [164]. Although being still exploratory, combination therapies consisting of CAR-T cells and the parallel administration of immune checkpoint inhibition could further improve therapeutic outcomes in cancer patients [176].

Multiple studies demonstrated that ACT with T cells expressing anti-CD19 CARs induced robust and durable responses in patients with chronic lymphocytic leukemia (CLL), diffuse large B cell lymphomas, and B cell acute lymphoblastic leukemia (B-ALL) [177–180]. Another very promising target antigen for treating B cell malignancies with CAR-T cells is the B cell maturation antigen (BCMA). Here, BCMA-specific CAR-T cells induced overall response rates of 81% in multiple myeloma patients [181]. However, patients often developed severe cytokine release syndromes (CRSs) or observed off-tumor toxicities primary resulting in B cell aplasia, a condition that can be clinically managed with prophylactic immunoglobulin infusions [164]. To date, various anti-CD19 (axicabtagen-ciloleucel, tisagenlecleucel, lisocabtagen-maraleucelor, brexucabtagen autoleucel) and anti-BCMA (idecabtagen vicleucel, ciltacabtagene autoleucel) CAR-T cell therapies are approved for the treatment of non-Hodgkin lymphoma, B-ALL, and multiple myeloma [182,183].

In contrast, developing CAR-T cell therapies against solid tumors was significantly less successful so far, with potential challenges being (i) the inefficient T cell migration and infiltration of tumor tissues

due to physical barriers at the tumor site, (ii) difficulties in antigen selection due to high heterogenicity of solid tumors, (iii) the high risk of on-target/off-tumor toxicities due to target antigen expression in healthy tissues, and (iv) the presence of immunosuppressive factors in the TME [164]. However, BioNTech's BNT211 program currently assesses both safety and preliminary efficacy of a CAR-based T cell therapy targeting Claudin-6 (CLDN6), a developmentally regulated tight junction protein, in patients with relapsed or refractory advanced solid tumors (clinical trial ID: NCT04503278). Furthermore, BNT211 is a potential first-in class approach combining a CAR-based T cell therapy approach with nanoparticle RNA vaccination, thereby facilitating the body-wide delivery of the CLDN6 antigen to promote the selective expansion of infused CAR-T cells [184].

Although representing promising strategies for the treatment of multiple cancer types, there are potential safety risks that have to be considered when utilizing genetically altered T cells for ACT approaches. Those safety concerns were observed in multiple studies reporting serious and non-predictable toxicities upon treatment of cancer patients with engineered T lymphocytes, thereby highlighting the necessity of methods to reliably define TCR specificities before administration. Generally, the most severe toxicities are (i) cytokine-release syndromes, (ii) on-target/off-tumor toxicities as a result of tumor antigen expression in healthy tissues, and (iii) off-target toxicities either originating from cross-reactivity towards alternative proteins or from the formation of hybrid TCRs composed of endogenous and exogenous TCR chains [185]. For example, modified T cells expressing high affinity TCRs against MAGE-A3 induced fatal myocardial damages, which were most likely caused by the recognition of an unrelated peptide derived from the muscle-specific self-protein titin [186]. Other studies reported lymphocyte-induced autoimmune colitis or cytokine storms resulting in multiorgan failure and patient death. In both cases, toxicities were most likely due to on-target/off-tumor toxicities induced by target antigen expression in healthy cells [93,173].

2.7.2 Designing TCR-engineered T cell immunotherapies

2.7.2.1 Common strategies for the identification of suitable T cell target antigens

Gene engineering technologies dramatically changed the overall landscape of adoptive cell therapy (ACT) approaches by providing the possibility to specifically redirect T cells towards selected tumor antigens. Hence, research no longer focused on how to harvest sufficient numbers of T cells for subsequent administration, but scientists rather tried to develop new strategies to identify high-avidity tumor-specific TCRs and to efficiently generate genetically engineered T cells [187]. However, the identification of suitable target antigens which are selectively expressed on tumor cells is one major bottleneck limiting the success of ACTs [188]. Besides from tumor-specific expression patterns, promising TCR targets should possess additional key characteristics for being considered ideal target antigen candidates, such as (i) being expressed on cancer stem cells to promote tumor eradication and (ii) being involved in the oncogenic process to reduce the risk of tumor evasion. Furthermore, (iii) they must be able to elicit immune responses and (iv) to bind common HLA alleles [187].

Hence, labor-intensive tumor screenings are necessary to identify immunologically relevant target antigen candidates. Comparative whole exome sequencing and transcriptome sequencing of tumor and healthy patient tissues are commonly used to screen for differently expressed genes and tumor-specific non-synonymous mutations [187,189]. In a next step, potentially immunogenic candidate mutations can be identified by performing MHC-binding predictions with *in silico*-generated polypeptides containing one of the previously identified non-synonymous mutations. The polypeptides predicted to bind the patient's MHC molecules best are eventually synthesized and used for further TIL-based recognition screenings [188]. Moreover, ready-to-transfect tandem minigene (TMG) libraries can be generated to circumvent potential accuracy limitations of peptide binding predictions. Here, recognition screenings can easily be performed for every single patient HLA allele, respectively. Individual TMG cDNA constructs encode

up to 6 candidate mutations and are eventually tested for the recognition by autologous TILs via co-transfection of target cells with relevant patient HLAs and respective TMG libraries [190].

In addition to the methods described above, analysis of the tumor ligandome can be considered a good alternative for the identification of T cell target antigen candidates. This method depends on the mass spectrometry-based analysis of all peptides eluted from the HLA molecules of a particular tumor sample or cancer cell line [187]. However, peptide-specific T cell responses still have to be validated *in vitro* [191]. The screening of tumor cDNA libraries is another approach that is commonly used for T cell antigen discovery. Here, cDNA expression constructs are generated from tumor-obtained total RNA samples and are subsequently amplified in *E. coli* bacteria. Isolated plasmid DNA from several hundred colony pools, which typically consist of 100 bacterial colonies, and relevant patient HLAs are then co-expressed in antigen-presenting cells that are eventually screened for cytotoxic T cell recognition [192]. Finally, recognized cDNA pools are stepwise reduced and tested until the cognate target antigen cDNA can be revealed.

2.7.2.2 Recurrent KRAS mutations as promising target neoantigens

Current immunotherapy approaches mainly target individual neoantigens derived from unique patient mutations. Since cancer driver mutations are recurrent among different patients, immunogenic treatment strategies can, in theory, also be developed against public neoantigens arising from shared variations in oncogenes or tumor suppressor genes [4]. However, in reality, neoantigens arising from tumor-specific perturbations are rarely shared among patients [5]. Pearlman and colleagues recently summarized the future perspectives of immunotherapies targeting recurrent mutations: They demonstrated that the effective identification of recurrent targets is the major challenge when designing immunotherapy approaches for larger patient cohorts. By potentially providing clinical benefits for larger patient cohorts, public neoantigen therapies could circumvent the complex financial and logistic bottlenecks associated with personalized approaches. However, the number of potentially treatable patients highly depends on the frequency of the corresponding HLA allele facilitating the efficient presentation of the neopeptide. Given the fact that several driver gene mutations occur in multiple cancer types, the number of potentially treatable patients can only be estimated. For example, a neoantigen peptide derived from the most frequent KRAS mutation (p.G12D) was found to be presented via HLA-A*03, thereby accounting for a theoretical patient cohort of about 1% of all patients with any of the 12 most common solid tumor types in the US. Therefore, immunotherapeutic agents developed against one of the top 10 public neoantigens could, in theory, be used for the worldwide treatment of more than one million patients per year. However, actual patient numbers are assumed to be lower since the mere existence of a recurrent mutation does not guarantee the presence of a shared neoantigen [4].

The *RAS* gene family encode small enzymes, so-called GTPases, which hydrolyze guanosine triphosphate (GTP), thereby linking the activation of upstream cell surface receptors, such as epidermal growth factor receptor (EGFR) or fibroblast growth factor receptor (FGFR), to downstream signal transduction pathways regulating cell proliferation, survival, and differentiation. Mutant *RAS* family isoforms are among the most frequent oncogenes in cancer and are represented by mutated *KRAS*, *NRAS,* and *HRAS* [193]. Point-mutated RAS proteins significantly contribute to several aspects of malignancy by deregulating tumor-cell growth, programmed cell death, or the ability to induce blood vessel formation [194]. As recently reviewed, *RAS* mutations occur in approx. 9-30% of all sequenced cancers, but the reported variation frequencies substantially differ between databases used for mutation analysis. *KRAS* represents the most frequently mutated (86% of *RAS*-driven cancers) *RAS* isoform, followed by *NRAS* and *HRAS* representing about 11% and 3% of all *RAS* mutations, respectively. The three *RAS* genes encode 4 different proteins sharing 82-90% overall sequence identity. While *HRAS* and *NRAS* encode proteins of 189 amino acids, *KRAS* encodes two different splice variants (188 and 189 amino acids) arising from alternative utilization of exon 4. Generally, frequencies of mutated *RAS* isoforms differ between cancer types with *KRAS* being the predominant mutated isoform in multiple malignancies, such as pancreatic ductal adenocarcinoma (PDAC; mutation frequency of 66.1%), lung adenocarcinoma

(LAC; mutation frequency of 16.5%), or colon and rectal adenocarcinoma (respective mutation frequencies of 30.3% and 34.4%) [195,196].

Furthermore, single nucleotide variations (SNVs) accounting for different cancer-associated RAS mutations are almost exclusively found at amino acid positions G12, G13, and Q61, which are shared among all three RAS isoforms. In PDAC and LAC, RAS SNVs are predominantly found at amino acid position G12, thereby accounting for 94% and 91% of all detected RAS mutations, respectively. Furthermore, G12 mutations are also common in colorectal carcinomas (approx. 76% of all detected RAS mutations), but there is also a considerable frequency of G13 SNVs that are rather rarely detected in PDAC and LAC. Generally, G12 alterations are predominant (approx. 83%) among all KRAS mutations, followed by G13 (approx. 14%) and Q61 (approx. 2%) mutations. In contrast to KRAS, G12 mutations only account for approx. 35% of all HRAS and ca. 23% of all NRAS mutations. Furthermore, approx. 62% of all NRAS and ca. 34% of all HRAS mutations result from SNVs at amino acid position Q61. (Figure 2.7.2) [196]. In NSCLC, KRAS mutations can also be found together with mutations in various tumor suppressors, such as tumor protein 53 (TP53) or liver kinase B1 (LKB1). Recent data suggests that patients with those KRAS co-mutations might have poorer prognosis when compared to patients exhibiting only KRAS mutations [197].

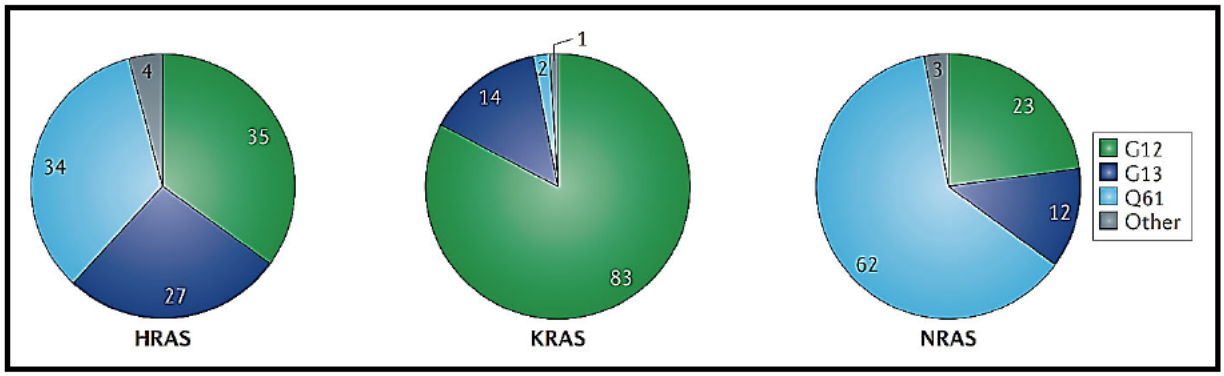


Figure 2.7.2: Hotspot mutation frequencies across RAS isoforms HRAS, KRAS, and NRAS.
While G12 variations are predominant among all KRAS mutations (approx. 83%), they only account for approx. 35% of all HRAS and ca. 23% of all NRAS mutations. In contrast to KRAS, Q61 mutations are substantially more frequent in NRAS (approx. 62%) and HRAS (ca. 34%), respectively. Proportions of RAS hotspot mutations are depicted in percent of all variations occurring at the respective amino acid positions. Graphics and descriptions were adapted from A. Cox et al.: Nature Reviews Drug Discovery (2014); 13: 828-851 [196]. ***Abbr.:*** *G = Glycine; HRAS = Harvey rat sarcoma viral oncogene; KRAS = Kirsten rat sarcoma viral oncogene; NRAS = Neuroblastoma rat sarcoma viral oncogene; Q = Glutamine.*

Despite of 40 years of intensive research, there are no effective treatment strategies targeting mutated KRAS, except for the two selective $KRAS^{G12C}$ inhibitors sotorasib and adagrasib [198,199]. Generally, KRAS mutations are often associated with therapy resistance and poor outcomes in cancer patients. Due to the lack of accessibility and the toxic effects associated with KRAS-targeting, mutated KRAS was deemed "undruggable" when discovered in 1982. However, sotorasib (AMG 510), a small molecule irreversibly targeting the $KRAS^{G12C}$ mutation, was the first KRAS inhibitor to reach clinical testing in humans. Generally, the $KRAS^{G12C}$ oncogenic driver mutation occurs in approx. 13% of all NSCLC and in other solid cancers (>3%). $KRAS^{G12C}$ inhibitors such as sotorasib specifically inhibit the mutated KRAS by locking it in its inactive GDP-bound state, thereby inducing objective response rates (complete or partial responses) of 32.2% and disease control (objective response or stable disease) in 88.1% of treated NSCLC patients. Furthermore, an ongoing clinical trial aims at the comparison of sotorasib and docetaxel chemotherapy in $KRAS^{G12C}$-positive NSCLC patients that had before been treated with platinum-based doublet chemotherapy and checkpoint inhibition [200–202].

Moreover, there is a variety of TCRs originating from naturally occurring patient-derived TILs which are reported to specifically recognize public neoantigens. Those TCRs had been cloned and could, in

theory, be used for genetically engineering other patients' T cell populations to target shared neoantigens [4]. For instance, $KRAS^{G12D}$ is the most frequent mutation in pancreatic cancer (approx. 70-80%), colorectal cancer (approx. 13%), and NSCLC (approx. 7%), thereby representing one of the most promising target antigens for T cell-mediated immunotherapy [203]. Moreover, recently published data revealed that the three most common KRAS mutations at amino acid position 12 (p.G12C, p.G12D, p.G12V) account for up to 70% of all KRAS mutations found across various tumor types. To date, multiple anti-KRAS TCRs specifically targeting those p.G12 mutations were isolated from patients with metastatic cancer [204]. Furthermore, two TCRs had been isolated from immunized HLA-A*11:01-transgenic mice and were shown to be highly reactive towards the KRAS mutations p.G12V and p.G12D [205]. There are now two ongoing clinical trials aiming at analyzing the safety and efficacy of both TCRs in an ACT-based treatment setting (clinical trial IDs: NCT03745326, NCT03190941). Since both studies use murine receptors, it will be important to assess potential adverse events or efficacy-attenuating effects which could potentially be induced by the host's immune system. With the first clinical trials underway, proceeding cancer genome sequencing will probably facilitate the future treatment of patients with new therapies targeting public neoantigens [4].

2.7.2.3 Identification of potentially tumor-reactive T cells among patient-derived TIL populations

Finding the perfect immunogenic epitope is not trivial, but the identification of tumor-reactive T cells from tumor-infiltrating lymphocyte (TIL) populations might be challenging as well [187]. TILs are phenotypically diverse populations consisting of both tumor-reactive T cells as well as bystander T cells, with the latter recognizing a wide range of non-tumor target antigens [206]. Recently, various approaches aimed at the identification of tumor-reactive TILs by either analyzing putative tumor specificity marker (CD39, CD103 and PD-1) expression [207,208] or by *in vitro* analyses of actual tumor-reactivity. The *in vitro* analysis of effector T cell reactivity upon stimulation with tumor antigens or autologous tumor cells represents one of the most common approaches for detecting and characterizing tumor-reactive TILs. However, no standardized pipeline has yet been established to reliably detect tumor-specific TIL repertoires [209].

Several techniques can be used to analyze T cell responses when co-cultured with tumor cells. Those methods include ELISA (Enzyme-linked Immunosorbent Assay) and ELISpot (Enzym-linked Immunospot) assays to determine T cell activation by assessing the levels of secreted pro-inflammatory cytokines, such as IL-2 or IFN-γ. Furthermore, a detailed phenotypic characterization of stimulated T cells may be achieved by intracellular cytokine staining and subsequent flow cytometry analysis [210,211]. The additional staining of functional T cell activation markers for degranulation and cytotoxicity, such as CD107a/b, perforin, or granzyme, may facilitate an even more complete assessment of T cell functionality [212,213]. Another prominent strategy for the identification of tumor-reactive T cell clonotypes is to analyze TCR repertoires of TILs from freshly resected tumors. Rosenberg and colleagues used a high-throughput next generation TCR beta sequencing approach to analyze the TCR diversity in different TIL subsets from human metastatic melanoma patients. They subsequently identified TCR alpha and TCR beta chain pairs via RT-PCR and by using pairSEQ, a statistical model to accurately pair thousands of TCR chain sequences. Thus, they managed to identify potentially tumor-reactive TCRs solely based on PD-1 expression and frequency within tumor tissues, thereby introducing a potent strategy to identify potentially tumor-reactive TCRs without the prior knowledge of their respective antigen specificities [214,215].

2.7.2.4 Identification of potentially tumor-specific TCR candidates by clustering of highly related TCRs with potentially similar target antigen specificity

Some antigen-specific TCR repertoires include TCRs that are frequently observed in multiple individuals, thereby mediating public T cell responses. Shared TCRs have been described in a variety of immune responses, especially in the context of infectious diseases, autoimmunity, and diverse

malignancies [216]. It is assumed that hundreds or even thousands of distinct TCRs could be able to recognize similar immunodominant target antigens, which might be especially true for common viral epitopes [217]. Highly similar subsets of virus-specific TCRs, referred to as public TCRs, were recently verified by comparing related T cell populations initially isolated from 29 different individuals [218].

Hence, a strategy to identify tumor-specific TCRs could also be based on surveying TCR repertoires of multiple cancer patients to uncover highly similar TCRs and their corresponding antigen specificities. By using their GLIPH2 (Grouping of Lymphocyte Interactions with Paratope Hotspots version 2) algorithm, Mark M. Davis and colleagues (Stanford University, USA) recently published a strategy to distill raw TCR sequencing data into many shared specificity groups likely to recognize identical peptide-MHC ligands. By analyzing almost 800,000 TCR beta chain sequences from 178 NSCLC patients, they found a total of 435 shared specificity groups to be clonally expanded and enriched in patients' tumor tissues. They furthermore demonstrated that CD8+ T cells expressing identical TCR alpha chain CDR3 sequences and a highly related CDR3β motif, only differing in one single amino acid, consistently shared similar antigen specificities [219,220]. However, it remained debatable whether the discovered target antigen recognition was either mediated by TCR alpha, TCR beta, or by a combination of both chains. Nevertheless, this study demonstrated that highly related TCRs have the potential to specifically recognize similar targets. Clustering of TCRs based on CDR3 alpha and beta sequence similarities and simultaneously comparing them to TCR specificity databases could therefore represent a promising strategy for identifying plenty of tumor-specific TCRs. Indeed, there are available tools and algorithms designed to identify highly similar TCRs and predict potential antigen specificities based on related TCR sequences [221,222].

3) Results

3.1 Objectives and overall experimental strategy

Harnessing T cell-based immune responses against highly immunogenic tumor-specific antigens (TSAs) or aberrantly expressed tumor-associated antigens (TAAs) represents an attractive strategy to achieve clinical benefits in cancer patients [223,224]. However, both the identification of suitable target epitopes as well as the identification of tumor-specific T cells is not trivial [187]. Hence, the biotechnology company HS Diagnomics (HSD) has recently established a prediction pipeline to effectively identify tumor-specific T cells and their T cell receptors (TCRs) among heterogenous populations of tumor-infiltrating lymphocytes (TILs; EU patent EP 3180433 B1) [3]. Assessing the suitability of this prediction pipeline by attempting to identify the exact target antigens of the TIL-derived TCRs was the major goal of this doctoral research study. To achieve this goal, the following core objectives were defined:

(i) generation of transgenic effector populations (TCR-T cells) expressing TCRs cloned from programmed cell death protein 1 (PD-1)-positive TILs with highest tumor-to-non-tumor (T/nT) ratios in at least one of three analyzed non-small cell lung cancer (NSCLC) patients,
(ii) next generation sequencing (NGS) of the patients' tumor and healthy lung tissues to identify patient-specific HLA class I-binding neoantigen candidates, and
(iii) recognition testing (ELISpot, CD107a degranulation assay) of neoantigen candidates and common TAAs with generated TCR-T cell populations to identify and characterize their targets.

HSD used complete TCR beta chain (TRB) profiling of T cell subsets isolated from tumor or adjacent healthy lung tissues of NSCLC patients P18, P43, and P50 to assess T/nT ratios of TIL clonotypes found in the respective patients. The most promising TCR candidates, i.e. TCRs from PD1+ TIL clonotypes with highest T/nT ratios, were subsequently synthesized by HSD and then cloned into retroviral expression backbones [225] provided by the Wölfel group (University Medical Center, UMC Mainz). Resulting TCR expression constructs were to be used to generate TCR-T cells from buffy coat lymphocytes. To circumvent potential interference of endogenous TCRs with surface expression of introduced TCRs and subsequent specificity testing, endogenous TCR genes had to be knocked out via CRISPR/Cas9. Established TCR-T cell populations should then be used for extensive recognition testing with patient-specific neoantigen candidates and shared TAAs.

To find out if patient-specific mutated neoantigen candidates are targeted by the generated TCR-T cell populations, the following experimental setups were planned:

(i) comparative whole exome sequencing (WES) of genomic DNA (gDNA) isolated from flash frozen tumor and normal lung tissue specimens to identify neoantigen candidate mutations,
(ii) transcriptome sequencing of tumor-derived total RNA to confirm mutated gene expression, and
(iii) extensive HLA class I binding predictions with established algorithms to identify and synthesize individual candidate neopeptides binding to the patients' respective HLA molecules.

Moreover, the following experiments were planned for subsequent recognition testing:

(i) loading of predicted neopeptides on K562 cells transduced with a patient's HLA class I alleles and testing of peptide-pulsed targets for the recognition by generated TCR-T cells, and
(ii) verification that peptide-reactive TCR-T cell responses are also directed against endogenously processed and presented mutated neoantigens. If possible, this should be achieved by applying tumor cell lines exhibiting physiological neoantigen expression, processing, and presentation, or by using HEK 293T target cells co-transfected with cDNAs encoding individual HLA class I molecules and candidate neoantigens.

Furthermore, alternative antigen identification strategies were available in case no reactivity against mutated neoantigen candidates was found. Here, it was planned to test established TCR-T cell

populations for the recognition of an antigen panel consisting of cDNAs encoding 43 shared TAAs by using HEK 293T target cells co-transfected to also express a patient's respective HLA class I molecules. In addition, another option was to apply patient-derived TCRs to an *in silico* TCR clustering approach aiming at the efficient identification of highly related TCRs with potentially shared antigen specificities. Based on the number of identified tumor-specific T cells and their corresponding TCRs, thorough efficiency assessments were planned to evaluate the overall efficiency of the conducted research study.

3.2 Identification of a mutated KRAS neoantigen recognized by three tumor-specific TCRs cloned from TILs of NSCLC patient P18

3.2.1 Identification of three tumor-reactive TIL clonotypes in NSCLC patient P18

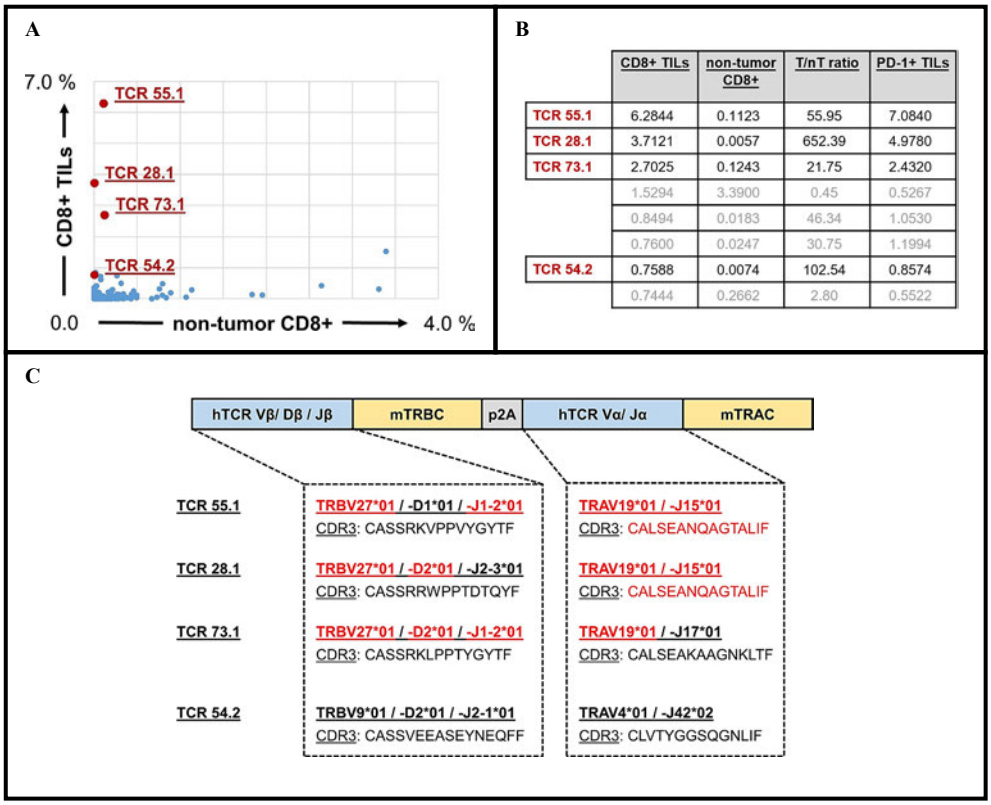
To identify tumor-reactive TIL candidates, complete T cell receptor (TCR) beta chain (TRB) profiling was performed for different T cell fractions (CD3+, CD4+, CD8+, PD-1+) previously isolated from tumor and adjacent healthy lung tissues of NSCLC patient P18. Tumor-to-non-tumor (T/nT) ratios were calculated for the most frequent P18 TILs, thereby facilitating the ranking of T cell clonotypes based on their relative frequencies in tumor and normal lung tissues. Exact TCR alpha and beta chain amino acid sequences of most promising TIL candidates, i.e. PD-1+ clonotypes exhibiting highest T/nT ratios, were subsequently assessed via single cell RNA-sequencing. Respective TCRs were eventually synthesized and used for the generation of retroviral TCR expression constructs.

The TRB profiling approach revealed a total of 33,367 different TCR clonotypes in the tumor tissue of stage III lung adenocarcinoma patient P18 (data not shown). Firstly, the 4779 most common P18 TIL clonotypes were ranked based on their tumor frequencies and were then plotted for their respective T/nT ratios. The three most prominent TIL clonotypes, namely clonotypes 28.1, 55.1, and 73.1, exhibited very high T/nT ratios as well as highest frequencies among PD-1+ TIL populations (Fig. 3.2.1 A and B), thereby representing promising candidates for subsequent tumor specificity assessments. Notably, the three corresponding TCRs shared highly related nucleotide sequences resulting in identical TRA amino acid sequences for TCR 55.1 and TCR 28.1 (Fig. 3.2.1 C). In contrast, the TRA chain of TCR 73.1 differed in its J-segment usage, thereby resulting in 5 amino acid differences within the corresponding CDR3 (complementary determining region 3) sequence. TRB chain variable (V) segments were found to be identical in all three TCRs while similar diversity (D) segments were identified for in TCRs 28.1 and 73.1, respectively. Furthermore, TRB joining (J) segment J1-2*01was shared by TCRs 55.1 and 73.1. In summary, the CDR3 sequences of all three TCRs were similar enough to suspect a shared antigen specificity. As it could be possible that all three TCRs recognize virus-derived epitopes but not actual tumor antigens, another non-related TIL candidate, referred to as clonotype 54.2, exhibiting an exceptionally high T/nT ratio (Fig. 3.2.1 B) was additionally selected for further specificity testing.

Gamma-retroviral TCR expression constructs were generated comprising paired αβ TCR sequences of P18 TIL clonotypes 28.1, 55.1, 73.1, or 54.2, respectively (Fig. 3.2.1 C). Transgenic human T cells (TCR-T cells) were established via retroviral transduction of CD8+ healthy donor lymphocytes which had before been applied to a CRISPR/Cas9-mediated knock-out (KO) of their endogenous (end.) TCR expression. This end.TCR-KO approach was necessary to effectively prevent alloreactive responses of transduced donor lymphocytes, which could either be mediated by residual expression of endogenous TCRs or via the formation of mixed receptor dimers with unpredictable reactivity. After performing the TCR-KO approach, endogenous TCR expression was absent in more than 95% of all treated donor lymphocytes derived from buffy coat (BC) 827 (see Sup. Fig. 1). CRISPR/Cas9-treated CD8+ donor lymphocytes were then used for retroviral transduction with previously generated P18 TCR expression constructs.

After retroviral transduction, established TCR-T cell populations were tested for the specific recognition of P18-derived tumor cell suspensions. Since merely 10% of TCR 73.1-transduced CD8+ T cells

expressed the transgenic TCR (data not shown) and since viable tumor cell yields were limited, TCR 73.1-expressing T cells were excluded from testing. In contrast, TCR-T cell populations either transduced with TCR 28.1, TCR 55.1, or TCR 54.2 ectopically expressed up to 62% of the introduced TCRs (Fig. 3.2.2). In a next step, tumor reactivity was assessed for all three TCR-T cell populations (Fig. 3.2.3) by applying single tumor cell suspensions derived from patient P18. Here, the P18 tumor reactivity was blocked by adding the HLA class I-specific monoclonal antibody W6/32, thereby indicating an HLA class I-dependent target recognition. Due to the low overall quality of P18-derived NSCLC single cell suspensions after thawing, tumor cell-dependent T cell activation was rather low when compared to non-specific T cell activation with OKT-3. Nevertheless, the detected HLA class I-dependent tumor recognition was considered reliable for TCR-T cell cultures either expressing P18 TCRs 28.1, 55.1, or 54.2. This was justified by the facts that (i) IFN-γ secretion was observed for all three effector cultures when co-cultured with P18 tumor cell suspensions, (ii) the observed recognition was specifically blocked via W6/32, and since (iii) no background IFN-γ secretion was observed for TILs potentially included in P18 tumor cell suspensions (tumor only preparation). However, since only one frozen aliquot was available for both P18 tumor and normal lung cell suspensions, it was not possible to confirm the detected tumor reactivity by conducting another experiment. Hence, observed reactivity towards P18 tumor cell suspensions could only be demonstrated in one individual experiment and by using single well formats rather than duplicates.



	CD8+ TILs	non-tumor CD8+	T/nT ratio	PD-1+ TILs
TCR 55.1	6.2844	0.1123	55.95	7.0840
TCR 28.1	3.7121	0.0057	652.39	4.9780
TCR 73.1	2.7025	0.1243	21.75	2.4320
	1.5294	3.3900	0.45	0.5267
	0.8494	0.0183	46.34	1.0530
	0.7600	0.0247	30.75	1.1994
TCR 54.2	0.7588	0.0074	102.54	0.8574
	0.7444	0.2662	2.80	0.5522

Figure 3.2.1: Identification of potentially tumor-specific TCRs in NSCLC patient P18.
A and B) *T/nT frequencies of most prominent PD-1+ P18 TIL clonotypes. Most promising TIL clonotypes were initially identified via comparative TCR beta chain sequencing of T cells isolated from P18 tumor and healthy lung tissues. Clonotypes 28.1, 54.2, 55.1, and 73.1 exhibited high T/nT ratios as well as high frequencies among PD-1+ TILs. Non-relevant clonotypes/TCRs exhibiting relatively low T/nT ratios and/or low frequencies among PD-1+ T cell fractions are shown in blue or grey. Clonotype frequencies are shown as percentages among P18 CD8+ TILs, P18 lung tissue-derived CD8+ T cells, and as percentages among PD-1+ P18 TILs. TCRs used for further antigen screenings are highlighted in red.* ***C)*** *Complete TCR alpha and beta chain nucleotide sequences were revealed via single T cell sequencing. Respective CDR3 amino acid sequences are shown for each of the analyzed receptors. TIL clonotypes 55.1, 28.1, and 73.1, but not clonotype 54.2, expressed highly related TCRs. Similarities between receptors 28.1, 55.1, and 73.1 are shown in red. A schematic overview of the used P18 TCR-encoding pMX expression construct is shown above. Established pMX expression constructs were subsequently used for retroviral transductions of donor lymphocytes.* ***Abbr.:*** *hTCR = human T cell receptor; mTRAC/mTRBC = murine TCR alpha/beta chain constant region; NSCLC = non-small cell lung cancer; TIL = tumor-infiltrating lymphocyte; T/nT = tumor-to-non-tumor; TCR = T cell receptor.*

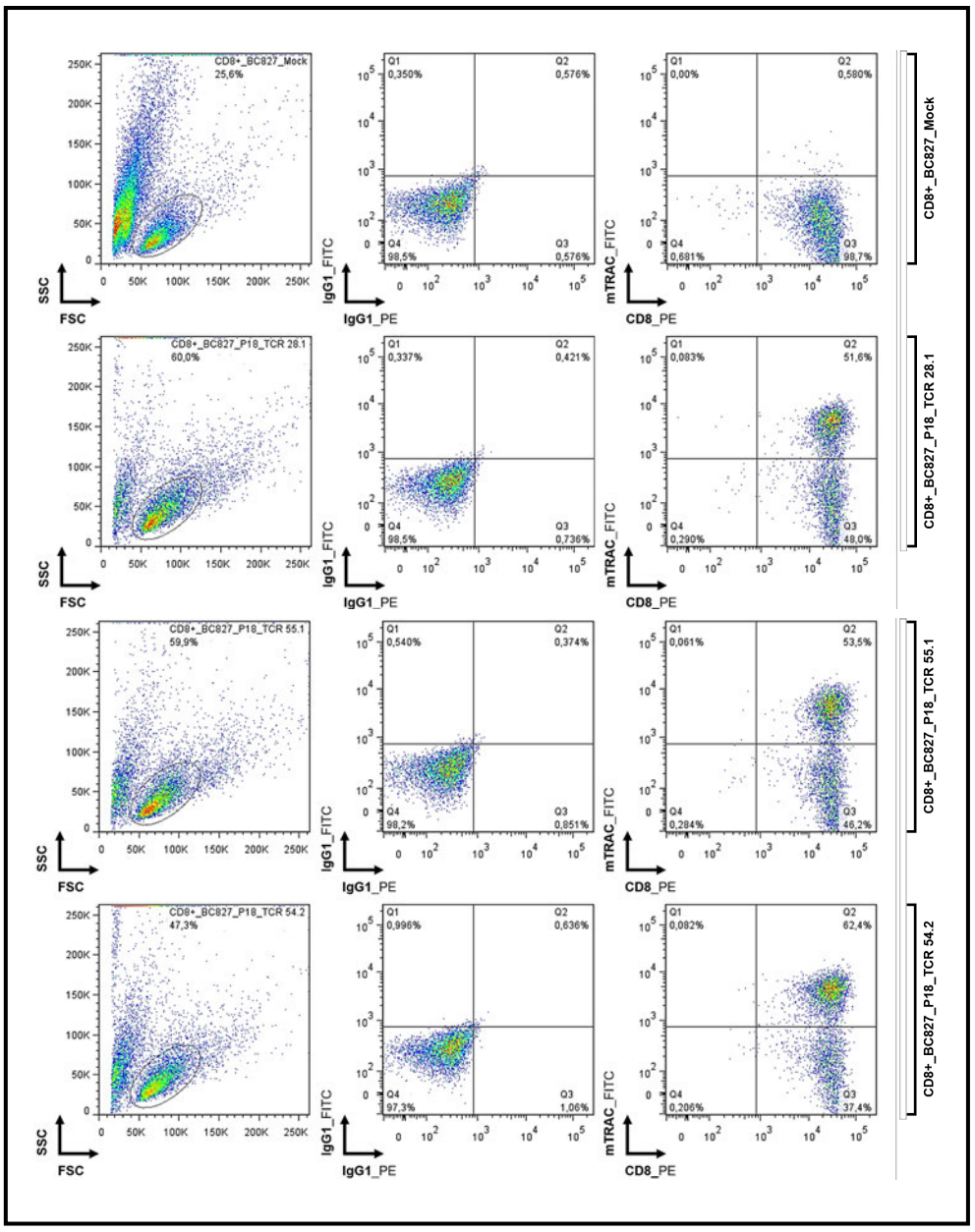


Figure 3.2.2: Generation of TCR-T cell populations expressing three potentially tumor-specific P18 TCRs.
P18 TCR expression in retrovirally transduced BC827-derived T cell populations. Before retroviral transduction with P18 TCR expression constructs, a CRISPR/Cas9-based KO approach was performed for knocking out endogenous TCR expression in BC827-derived T cell populations (Sup. Fig. 1). TCR-T cell populations either transduced with TCRs 28.1, 55.1, or 54.2 ectopically expressed up to 62% of the introduced TCRs. Exogenous TCR expression was determined via antibody-mediated mTRAC staining. Flow cytometry data was gated based on viable T cell populations and non-specific isotype (IgG1) controls. ***Abbr.:*** *BC = buffy coat; FITC = Fluorescein isothiocyanate; FSC = forward scatter; KO = knock-out; mTRAC = murine T cell receptor alpha chain constant region; NSCLC = non-small cell lung cancer; PE = Phycoerythrin; SSC = sideward scatter; TIL = tumor-infiltrating lymphocyte; TCR = T cell receptor.*

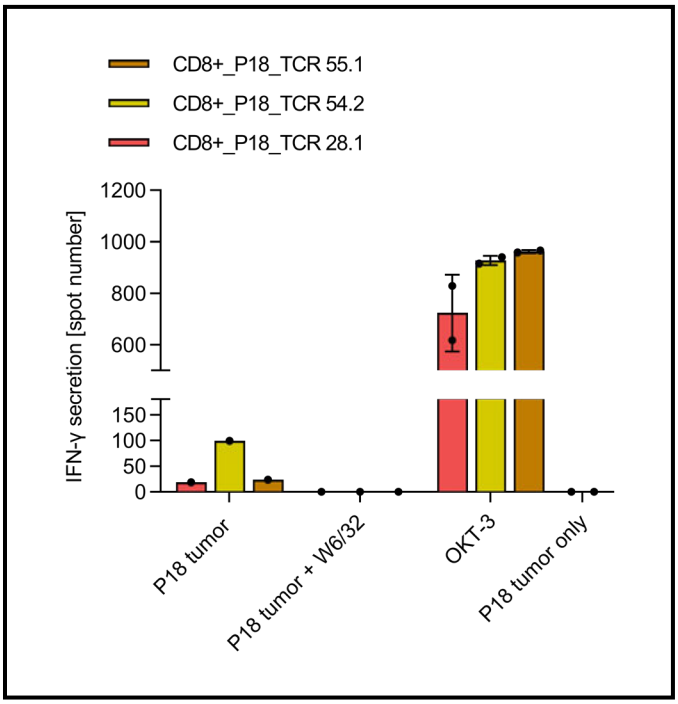


Figure 3.2.3: HLA class I-dependent recognition of P18 tumor cells via TCRs 28.1, 55.1, and 54.2.

TCR-T cell populations either expressing P18 TCRs 28.1, 55.1, or 54.2 were tested for the recognition of P18-derived tumor cell suspensions. MHC class I-dependent recognition of P18 tumor cells was verified for all three TCR-T cell populations by using the HLA class I-specific monoclonal antibody W6/32. T cell activation upon tumor recognition was rather low when compared to non-specific T cell activation via OKT-3. No background IFN-γ secretion was detected for P18 tumor cell suspensions. ELISpot assays were performed 16 days after retroviral transduction (Fig. 3.2.2) of BC827-derived end.TCR-KO T cells (Sup. Fig. 1) with P18 TCRs. Spot numbers either represent mean values resulting from duplicates (OKT-3, P18 tumor only) or single well testing results (P18 tumor, P18 tumor + W6/32). Error bars represent standard deviations of duplicates. Shown spot numbers were normalized for IFN-γ background levels of T cell populations in the absence of target cells. Used effector cells per well: 10,000 TCR-T cells. Used P18 tumor cells per well: 16,000. W6/32 concentration per well: 50 µg/ml. Raw ELISpot data is shown in Sup. Fig. 2. ***Abbr.:*** *autol. = autologous; end. = endogenous; HLA = human leukocyte antigen; IFN = interferon; MHC = Major histocompatibility complex; KO = knock-out; NSCLC = non-small cell lung cancer; TCR = T cell receptor; TIL = tumor-infiltrating lymphocyte.*

3.2.2 Three tumor-specific P18 TCRs specifically recognize a KRAS^{Q61H} neoantigen

Since detected P18 tumor reactivities of tested TCR-T cell populations were rather low (Fig. 3.2.3), revealing the exact target antigens specifically recognized by P18 TCRs 28.1, 55.1, 73.1, and 54.2 was highly important to verify tumor specificities for all three TCRs. Hence, potential neoantigen-encoding mutations were identified in P18-derived NSCLC tissue via next generation sequencing (NGS). Firstly, tumor-specific non-synonymous mutations were identified by using comparative whole exome sequencing (WES) of genomic DNA (gDNA) isolated from tumor and normal lung tissues of non-small cell lung cancer (NSCLC) patient P18. RNA-Seq of tumor-obtained total RNA samples was then used to assess mRNA expression levels of respective candidate mutations.

While mapping paired-end reads (150 nt) resulting from WES and RNA-Seq to the reference genome hg38 (GRCh38), overall read mappings of >99.4% (data not shown) confirmed the high quality of WES and RNA-Seq data used for subsequent variant identification. More than 361,000 patient-specific variants were initially identified when comparing to the hg38 reference. In a next step, 1,898 tumor-specific mutations were determined by excluding all variants also present in the normal lung tissue of patient P18. Tumor-specific variants were subsequently screened for actual gene expression (RPKM ≥0.5), thereby resulting in 738 mRNA-expressed tumor-specific mutations. Finally, 159 high quality non-synonymous mutations were identified by filtering for missense and frameshift mutations. After removing premature stop codon-inducing variants and false InDels via manual curation, a total of 74 non-synonymous tumor-specific candidate variants were finally selected for subsequent *in silico* HLA binding predictions. Those variants included 73 single nucleotide variations (SNVs) and one frameshift-inducing deletion. The overall strategy and corresponding results of the applied P18 neoantigen identification pipeline are again summarized in Fig. 3.2.4 A.

After identifying potential P18 neoantigen mutations via comparative WES and RNA-Seq, extensive *in silico* peptide binding predictions were individually performed for every P18 HLA class I allele, thereby resulting in a total of 581 candidate 9- and 10-mer neopeptides (Fig. 3.2.4 B). Potentially best binding neopeptides (percentile rank ≤6.0 and/or binding affinity IC50 ≤500nM) were determined by using the IEDB MHC-I binding prediction tool (Consensus or NetMHCpan-4.0 prediction methods) and the NetMHC-4.0 binding prediction algorithm (see Table 5.8). The total number of neopeptide candidates identified by both prediction algorithms ranged between 54 and 79 for HLA alleles *A*01:01*, *A*02:01*, *B*08:01*, *B*40:02*, and *C*07:01* (Fig. 3.2.4 B). However, in case of HLA-C*03:04, NetMHCpan-4.0 was the only prediction algorithm available, thereby resulting in 246 different candidate neoantigen peptides. Eventually, the 14 to 16 top scoring 9- and 10-mer neopeptides per HLA-I allele were synthesized for further recognition screenings with TCR-T cells either expressing P18 TCRs 28.1, 55.1, 54.2, or 73.1. In addition, two known neoantigen peptides encoded by mutated genes *MMS22L* and *INSIG1* [226] were synthesized as controls, thereby resulting in a total number of 96 synthesized peptides for subsequent specificity screenings (Sup. Fig. 3). Here, both control peptides were used to assess the quality of the generated peptide screening library (Fast Track Peptide Library, JPT, Berlin). Since the specific recognition of both control antigens (data not shown) by CTL cultures 16C/114 (specifically recognizes mutated MMS22L^{S437F}) and 3A/115 (specifically recognizes mutated INSIG1^{S238F}) could be verified, the quality of the applied peptide library was considered sufficient for further target antigen screenings.

Upon co-incubation with pooled CD8+ T cell populations either overexpressing tumor-reactive P18 TCRs 28.1 or 54.2, peptide immunogenicity was individually analyzed for each of the 94 synthesized neoantigen candidates. For this purpose, K562 cells were modified via retroviral transduction to overexpress one of the P18-derived HLA class I alleles, were then individually pulsed with each of the mutated 9- and 10-mer candidates, and were eventually used as antigen presenting target cells for further recognition screenings (Fig. 3.2.5 A). When presented by HLA-A*01:01-transduced K562 cells, it was found that 6 of the mutated neopeptides substantially induced IFN-γ secretion in TCR 28.1- and TCR 54.2-expressing bulk T cell populations. However, in the context of other P18 HLA-I alleles, K562 target cell populations failed to induce significant T cell activation when pulsed with any of the 94

candidate peptides. Particularly one of the tested neoantigen candidates, a KRAS^{Q61H}-derived 10-mer peptide (ILDTAGHEEY; mutated AA highlighted), was found to be recognized best. By separately analyzing both CD8+ TCR-T cell populations, TCR 28.1-expressing effectors were found to specifically recognize the KRAS^{Q61H} 10-mer peptide (Fig. 3.2.5 B). Concordantly, the previously detected cross-reactivity towards other neopeptide candidates was again demonstrated for TCR 28.1-expressing TCR-T cells. Furthermore, the observed KRAS^{Q61H} recognition could also be verified for effector populations expressing the highly related P18 TCR 73.1 (Fig. 3.2.5 B). In contrast, T cells transduced to express the non-related P18 TCR 54.2 completely failed to recognize any of the 94 tested neopeptides (Sup. Fig. 4).

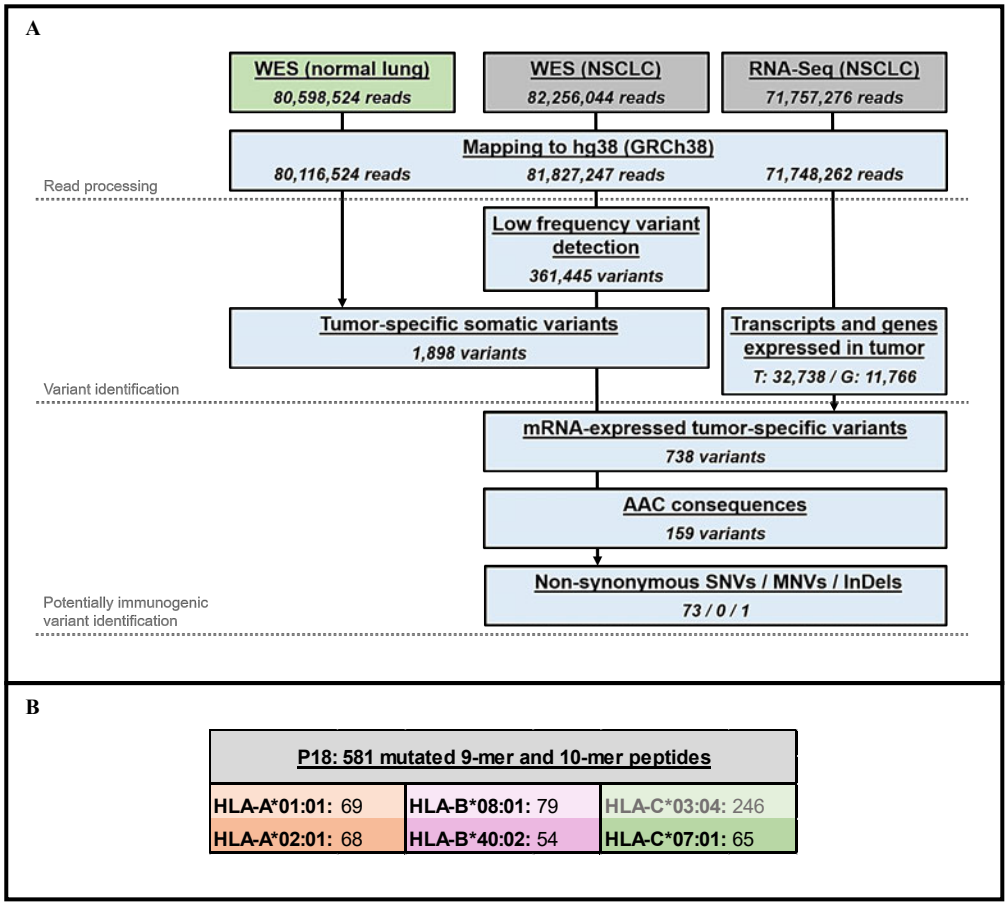


Figure 3.2.4: Antigenic landscape of NSCLC patient P18.

A) *In silico neoantigen prediction pipeline for the identification of non-synonymous tumor-specific mutations in NSCLC patient P18. Comparative WES was performed to identify a total of 1,898 tumor-specific mutations. By filtering for missense mutations, subsequent alignment with corresponding P18 tumor gene expression data, and after manual curation, 74 non-synonymous tumor-specific mutations were considered promising candidates for further peptide binding prediction. Identified mutations included 73 SNVs and one frameshift-inducing deletion. All sequencing reads (paired end reads; 150 nt) resulting from P18 WES and RNA-Seq were mapped to the reference genome hg38 (GRCh38). NGS data was analyzed and processed by using the CLC Genomics Workbench software (versions 20-23).* ***B)*** *HLA class I binding predictions for neoantigen candidates identified in NSCLC patient P18. HLA binding predictions were individually performed for each P18 HLA-I allele by using mutated 19-mer peptides, each of which represented one of the 74 previously identified non-synonymous mutations. Most promising neopeptides were identified as a consensus of the HLA binding algorithms NetMHC-4.0 (binding affinities of ≤500 nM) and IEDB MHC-I binding prediction tool (Consensus or NetMHCpan-4.0 prediction methods; percentile rank of ≤6.0). P18 HLA-I binding predictions resulted in a total of 581 promising 9- /10-mer candidate neopeptides. NetMHCpan-4.0 was the only algorithm available for HLA-C*03:04 binding predictions.* ***Abbr.:*** *AAC = amino acid-changing; G = genes; HLA = human leukocyte antigen; InDels = insertions and deletions; MNV = multiple nucleotide variations; NGS = Next Generation Sequencing; NSCLC = non-small cell lung cancer; nt = nucleotides; RNA-Seq = RNA Sequencing; RPKM = reads per kilobase of transcript per million reads mapped; SNV = single nucleotide variation; T = transcripts; WES = Whole Exome Sequencing.*

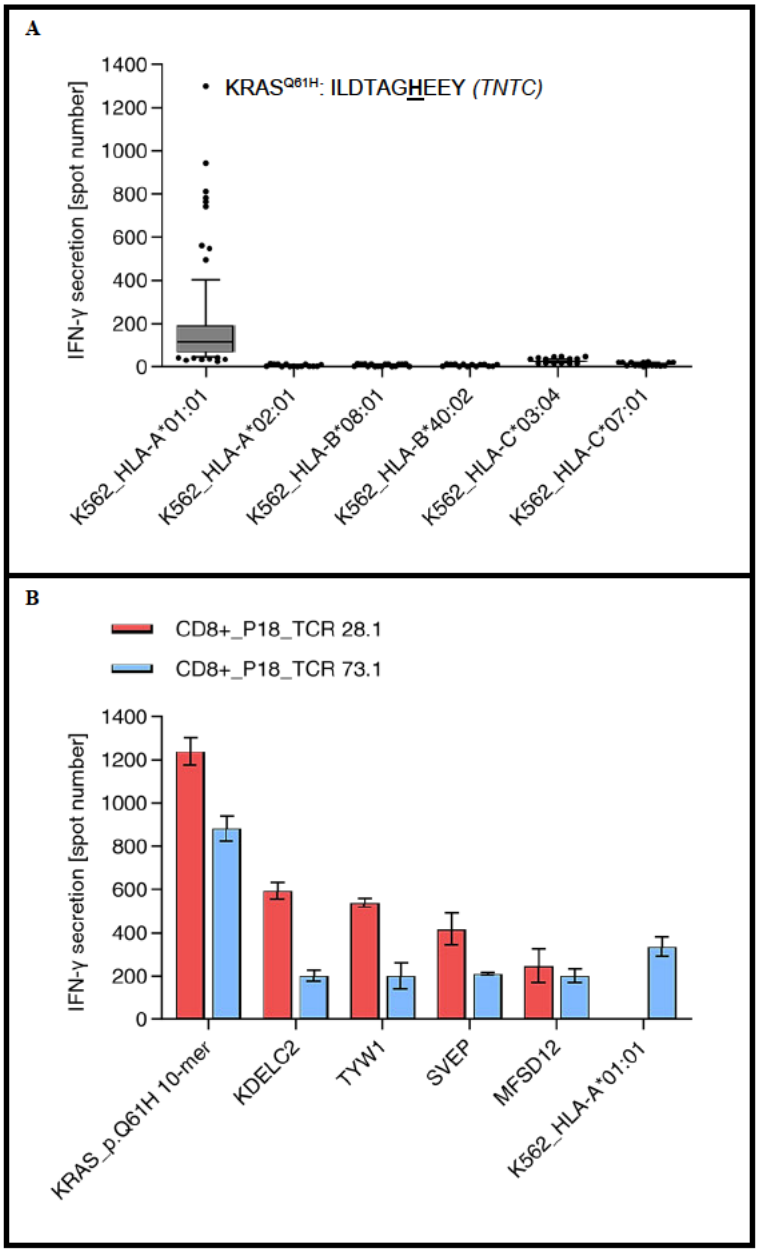


Figure 3.2.5: P18 TCRs 28.1 and 73.1 specifically recognize a KRAS^{Q61H} neopeptide via HLA-A*01:01.

A) *Neopeptide recognition screening with bulk T cell populations expressing P18 TCRs 28.1 and 54.2. IFN-γ secretion was used as a marker for T cell recognition. K562 cells overexpressing one of the P18 HLA-I alleles were separately pulsed with each of the 94 identified 9- and 10-mer neopeptide candidates. Substantial IFN-γ secretion was exclusively observed in the context of HLA-A*01:01. A KRAS^{Q61H}-derived 10-mer peptide (ILDTAGHEEY) was recognized best. Due to excessive IFN-γ release, the detected KRAS^{Q61H} recognition was defined as 1300 spots/well since exact spot numbers could not be assessed (TNTC). The top 5 best recognized neopeptides (from best to least recognized) included mutated 9- and 10-mers derived from KDELC2 (SMLEALRGV), TYW1 (GACYKHTFY), SVEP (ILNGKFSYTY), and MFSD12 (LTYSFHLPK). Spontaneous IFN-γ secretion of TCR-T cell populations was assessed in the absence of target cells (data no shown). Spontaneous IFN-γ secretion was analyzed as duplicates and resulted in mean spot numbers of 3.5 (TCR 28.1) and 6 (TCR 54.2), respectively. IFN-γ background levels were assessed upon co-incubation with non-pulsed K562 target cell pools expressing each of the relevant P18 HLA-I alleles (data not shown). IFN-γ background levels were analyzed as duplicates and resulted in mean spot numbers of 33 (TCR 28.1) and 19.5 (TCR 54.2), respectively. ELISpot assays were performed 29 days after retroviral transduction (Fig. 3.2.2) of BC827-derived end.TCR-KO T cells (Sup. Fig. 1) with P18 TCR expression constructs. Shown spot numbers represent single well testing results. Used effector cells per well: 6000 mTRAC+. Neopeptide concentrations per well: 2 µg/ml. Target cells per well: 50,000. K562 cell pools per well: 16,000 per P18 HLA allele (used for assessing IFN-γ background levels; data not shown). Mutated amino acids are highlighted for each peptide, respectively.* ***B)*** *Neopeptide recognition testing with TCR-T cell populations either expressing P18 TCR 28.1 or TCR 73.1. IFN-γ secretion was used as a marker for T cell recognition. Both TCR-T cell populations were separately tested for the recognition of HLA-A*01:01+ K562 cells pulsed with one of the top 5 neopeptides shown in A). Both T cell populations recognized the p.Q61H-mutated KRAS 10-mer peptide (ILDTAGHEEY). Cross-reactivity towards other neopeptide candidates was detected for TCR 28.1-expressing effectors. Due to substantial IFN-γ release detected upon co-incubation with non-pulsed HLA-A*01:01+ K562 cells, cross-reactivity could not be confirmed for TCR 73.1+ T cells. ELISpot assays were performed 33 days after retroviral transduction (Fig. 3.2.2) of BC827-derived end.TCR-KO T cell cultures (Sup. Fig. 1) with P18 TCR expression constructs. Shown spot numbers represent mean values resulting from duplicates that were additionally normalized for spontaneous IFN-γ secretion in the absence of target cells. Error bars represent standard deviations of duplicates. Used effector cells per well: 2,000 TCR 28.1+ or 3,750 TCR 73.1+. Used target cells per well: 50,000. Neopeptide concentrations per well: 2 µg/ml. Raw ELISpot data is shown in Sup. Fig. 4.* ***Abbr.:*** *BC = buffy coat; end. = endogenous; H = histidine; HLA = human leukocyte antigen; KDELC2 = Protein O-Glucosyltransferase 3; KO = knock-out; KRAS =Kirsten rat sarcoma viral oncogene; Q = glutamine; MFSD12 = Major Facilitator Superfamily Domain-Containing Protein 12; mTRAC = murine T cell receptor alpha chain constant region; SVEP = Sushi, Von Willebrand Factor Type A, EGF and Pentraxin Domain Containing; TCR = T cell receptor; TNTC = too numerous to count; TYW1 = TRNA-YW Synthesizing Protein 1 Homolog.*

3.2.3 Further characterization of three KRAS^{Q61H} neoantigen-specific P18 TCRs

NSCLC patient P18-derived TCRs 28.1 and 73.1 were found to specifically recognize a mutated KRAS 10-mer neopeptide (ILDTAGHEEY; mutated AA highlighted) via HLA-A*01:01 (see section 3.2.2). Due to its sequence similarity to both TCRs (Fig. 3.2.1 C), the potentially tumor-specific P18 TCR 55.1 was also tested for the recognition of the KRAS^{Q61H} neoantigen. To compare the antigen specificities of all three receptors, transgenic CD4+ and CD8+ T cell populations were generated, each of which overexpressed one of the three P18-derived TCRs. For this purpose, healthy donor T cells were isolated from two buffy coats (BC074, BC437), were then applied to a CRISPR/Cas9-based KO of their endogenous TCR expression, and were subsequently transduced with respective retroviral P18 TCR expression constructs. Thus, P18 TCR expression levels of more than 88% (Sup. Fig. 6 A and B) were detected for all generated TCR-T cell populations while simultaneously reducing endogenous TCR expression to less than 9% (Sup. Fig. 5 A and B). Next, the resulting TCR-T cell populations were tested for the recognition of KRAS^{Q61H} neopeptide-pulsed K562 target cells overexpressing HLA-A*01:01 (Fig. 3.2.6). Here, the mutated KRAS neopeptide was recognized by all tested CD4+ and CD8+ TCR-T cell populations. Moreover, highly comparable IFN-γ secretion profiles were found across all effector populations and no major differences were found between CD4+ and CD8+ T cell populations expressing identical P18 TCRs.

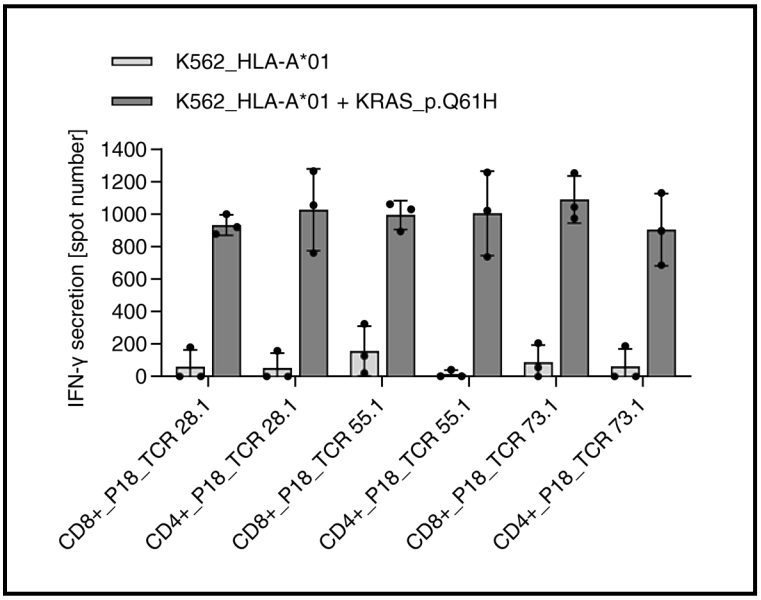


Figure 3.2.6: TCR-T cell populations expressing P18 TCRs 28.1, 55.1, or 73.1 recognize a KRAS^{Q61H} neoantigen via HLA-A*01:01.
*KRAS^{Q61H} neopeptide recognition testing with transgenic T cells either overexpressing P18 TCRs 28.1, 55.1, or 73.1. IFN-γ secretion was used as a marker for T cell recognition. Transgenic CD4+ and CD8+ TCR-T cell populations were individually tested for the recognition of HLA-A*01:01-positive K562 cells pulsed with the mutated KRAS 10-mer peptide (ILDTAGHEEY; mutated AA highlighted). All tested effector populations recognized the mutated KRAS peptide in the context of HLA-A*01:01. Three independent experiments were performed. Shown spot numbers either represent mean values resulting from duplicates or single well testing results of individual experiments. Error bars represent standard deviations of the three experiments. Due to the excessive IFN-γ release detected for some of the TCR-T cell populations, maximal peptide recognition was defined as 1300 spots per well since exact spot numbers could not be assessed (TNTC). Shown spot numbers were normalized for IFN-γ background levels of respective T cell cultures in the absence of target cells. TCR-T cells were generated by using end.TCR-KO T cell populations isolated from BCs 074 and 437, respectively (Sup. Fig. 5). For BC074, ELISpots were performed 20 or 35 days after retroviral transduction with respective P18 TCRs while BC437-derived T cell cultures were used 39 days after TCR transduction (Sup. Fig. 6). Used effector cells per well: 10,000 mTRAC+. Used target cells per well: 50,000. Neopeptide concentrations per well: 2 µg/ml. Raw ELISpot data is shown in Sup. Fig. 7.* ***Abbr.:*** *BC = buffy coat; H = histidine; end. = endogenous; HLA = human leukocyte antigen; KO = knock-out; KRAS = Kirsten rat sarcoma viral oncogene; mTRAC = murine T cell receptor alpha constant region; p. = amino acid position; Q = glutamine; TCR = T cell receptor; TNTC = too numerous to count.*

All generated CD4+ and CD8+ effector populations were additionally exposed to titrated KRAS^{Q61H} 10-mer peptide (ILDTAGHEEY; mutated AA highlighted) concentrations to assess functional avidity differences between TCR-T cells either expressing P18 TCRs 28.1, 55.1, or 73.1 (Fig. 3.2.7 A-C). While neopeptide-pulsed K562_HLA-A*01:01 target cells were recognized by all tested CD4+ and CD8+ TCR-T cell populations with high functional avidity (EC50 <10 nM), TCR 73.1 mediated the strongest T cell activation in both CD4+ and CD8+ TCR-T cell populations. Generally, higher IFN-γ secretion was detected for transgenic CD8+ effector populations than for corresponding CD4+ counterparts. However, at very low peptide doses of approx. 1 nM, higher IFN-γ secretion was observed for TCR 55.1-expressing CD4+ TCR-T cells than for corresponding CD8+ populations. In conclusion, all tested

TCR-T cell populations recognized the neopeptide-pulsed targets with high functional avidity, thereby demonstrating the KRAS^{Q61H} specificity of all three P18 TCRs. However, since previous antigen recognition testing approaches were exclusively performed by artifically pulsing target cells with rather high neopeptide concentrations, the following experiments were designed to also assess the recognition of the endogenously processed and presented KRAS^{Q61H} neopeptide.

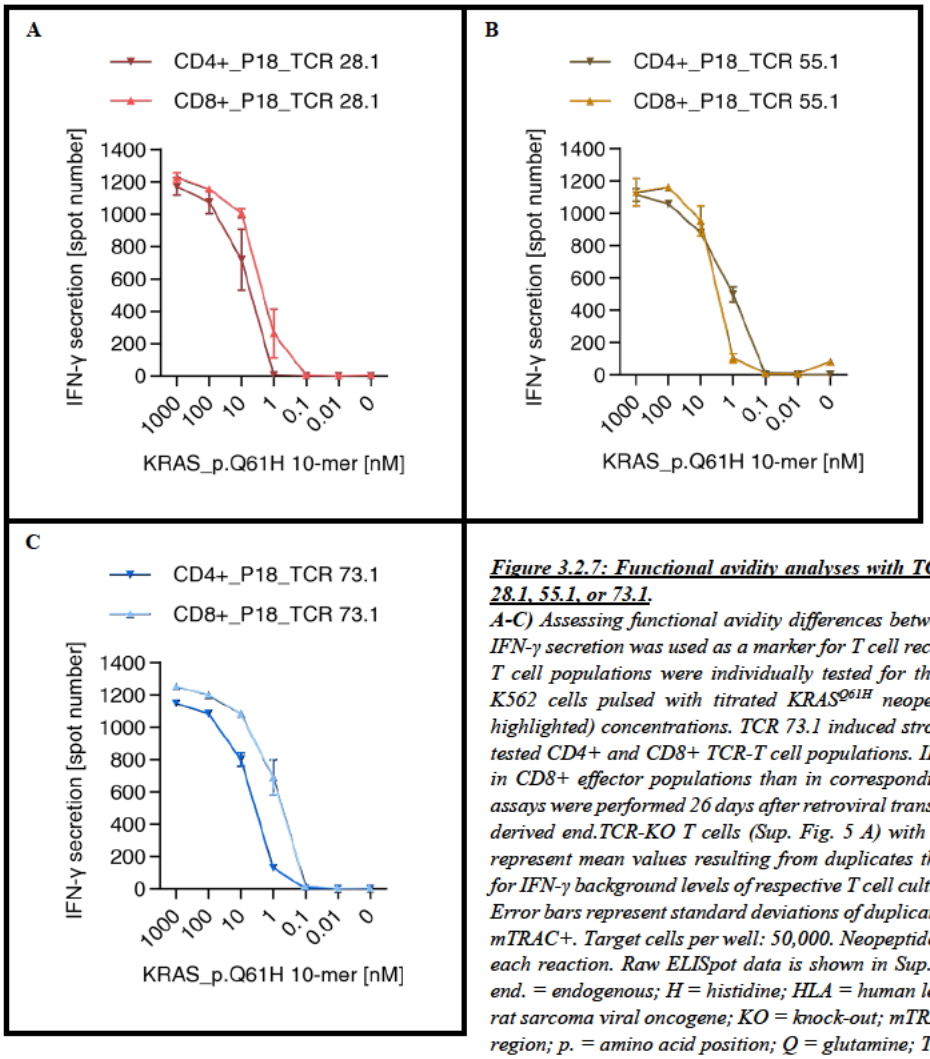


Figure 3.2.7: Functional avidity analyses with TCR-T cells expressing P18 TCRs 28.1, 55.1, or 73.1.

A-C) *Assessing functional avidity differences between P18 TCR-T cell populations. IFN-γ secretion was used as a marker for T cell recognition. CD4+ and CD8+ TCR-T cell populations were individually tested for the recognition of HLA-A*01:01+ K562 cells pulsed with titrated KRAS^{Q61H} neopeptide (ILDTAG**H**EEY; mutation highlighted) concentrations. TCR 73.1 induced strongest T cell activation across all tested CD4+ and CD8+ TCR-T cell populations. IFN-γ secretion was mostly higher in CD8+ effector populations than in corresponding CD4+ counterparts. ELISpot assays were performed 26 days after retroviral transduction (Sup. Fig. 6 A) of BC074-derived end.TCR-KO T cells (Sup. Fig. 5 A) with P18 TCRs. Shown spot numbers represent mean values resulting from duplicates that were additionally normalized for IFN-γ background levels of respective T cell cultures in the absence of target cells. Error bars represent standard deviations of duplicates. Effector cells per well: 10,000 mTRAC+. Target cells per well: 50,000. Neopeptide concentrations are indicated for each reaction. Raw ELISpot data is shown in Sup. Fig. 8.* ***Abbr.:*** *BC = buffy coat; end. = endogenous; H = histidine; HLA = human leukocyte antigen; KRAS =Kirsten rat sarcoma viral oncogene; KO = knock-out; mTRAC = murine TCR alpha constant region; p. = amino acid position; Q = glutamine; TCR = T cell receptor.*

To analyze TCR-T cell responses towards the endogenously processed and presented KRAS^{Q61H} 10-mer peptide, HLA-A*01:01-positive HEK 293T cells were established via retroviral transduction and were then transfected with neoantigen-encoding cDNA constructs. In addition to the previously analyzed p.Q61H mutation, three alternative RAS superfamily alterations that are commonly found at amino acid position 61 of KRAS, NRAS, and HRAS were also tested. Hence, cDNA constructs encoding KRAS^{Q61K}, KRAS^{Q61L}, or KRAS^{Q61R} mutations (Fig. 3.2.9) were generated and subsequently tested for the recognition by P18 TCR 28.1-, 55.1-, or 73.1-transduced effectors. However, the three P18 TCRs exclusively recognized the p.Q61H-mutated KRAS in the context of HLA-A*01:01 but none of the alternative KRAS cDNA constructs. This was observed for all tested CD8+ and CD4+ TCR-T cell populations, each of which expressed one of the three P18-derived TCRs (Fig. 3.2.8 A and B). Generally, CD8+ TCR-T cells expressing TCR 55.1 produced higher IFN-γ background levels than other effector populations (Fig. 3.2.8 A). However, while TCR 55.1-mediated T cell activation was comparable when testing wildtype KRAS cDNA and alternative KRAS mutations, KRAS^{Q61H} was recognized best, thereby emphasizing the KRAS^{Q61H} specificity of TCR 55.1.

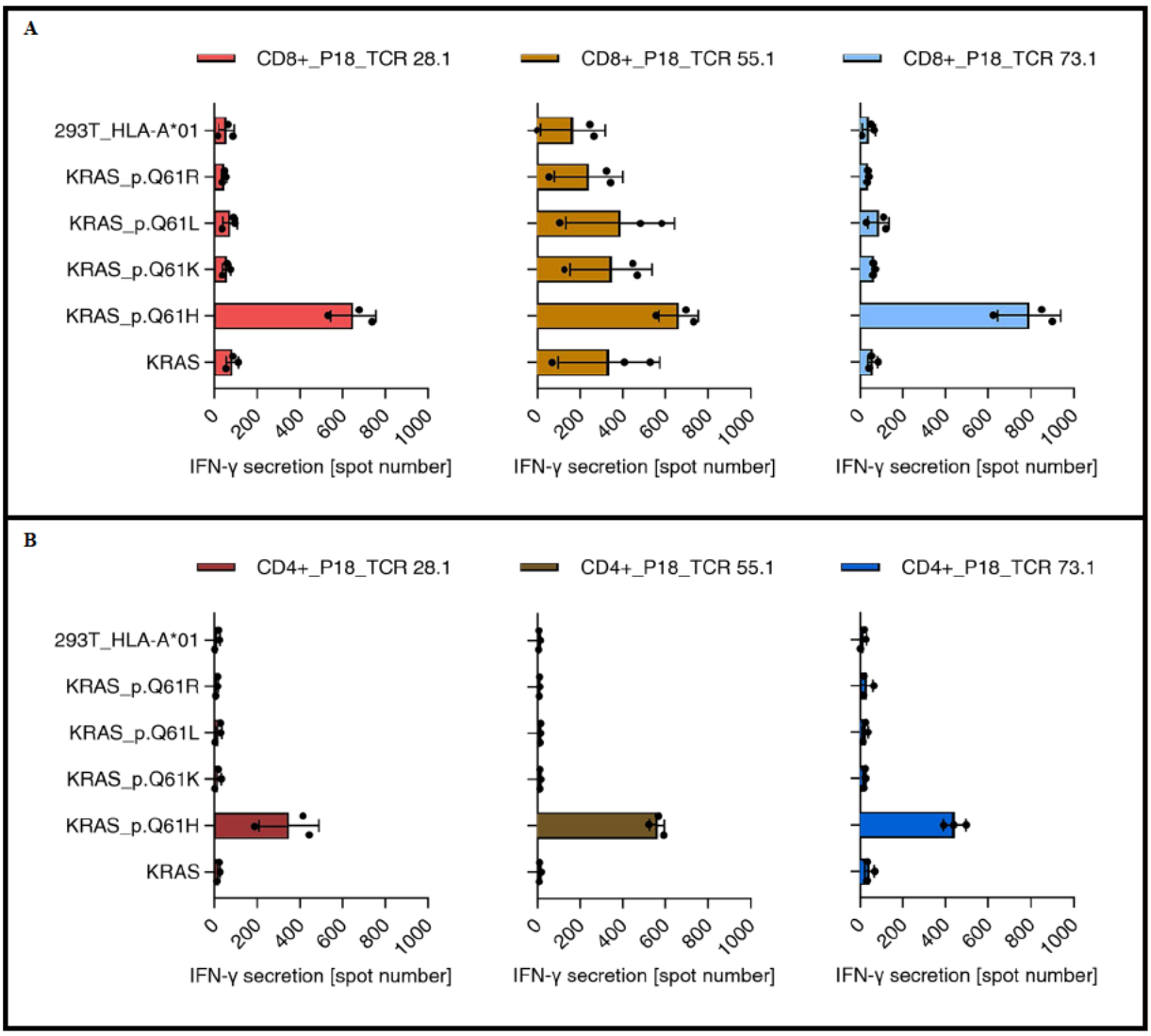


Figure 3.2.8: KRAS cDNA recognition testing with TCR-T cell populations overexpressing P18 TCRs 28.1, 55.1, or 73.1.
A and B) *IFN-γ secretion was used as a marker for T cell recognition. Transgenic CD4+ and CD8+ T cell populations were separately tested for the recognition of HLA-A*01:01+ HEK 293T cells transfected with one of the generated pcDNA3.1 expression constructs shown in Fig. 3.2.9. cDNA recognition analyses with CD8+ T cell populations are shown in A) while corresponding ELISpot results for CD4+ effectors are shown in B). All tested CD4+ and CD8+ effector populations recognized KRAS^{Q61H} cDNA but failed to recognize any of the other KRAS mutations. Shown spot numbers represent mean values calculated from duplicate test wells of three independent experiments. Error bars represent standard deviations of the three experiments. Shown spot numbers were additionally normalized for IFN-γ background levels of respective T cell populations in the absence of target cells. BC437-derived end.TCR-KO T cells (Sup. Fig. 5 B) were used for retroviral transduction (Sup. Fig. 6 B) with P18 TCR expression constructs. For transgenic CD4+ T cell populations, ELISpot assays were either performed 41 days or 55 days after retroviral transduction. For transgenic CD8+ T cell populations, ELISpot assays were either performed 40 days or 54 days after retroviral transduction. KRAS transfection efficiencies were assessed by separately transfecting HEK 293T cells with an EGFP-encoding pcDNA3.1 construct. Assessed transfection efficiencies ranged between 75% and 83% (approx. 24 h post transfection, data not shown). Effector cells per well: 7,500 P18 mTRAC+. Target cells per well: 20,000. KRAS cDNA used for transient transfection per well: 300 ng. Y axis labelling shown on the left applies to all assigned plots. Raw ELISpot data is shown in Sup. Fig. 9.* ***Abbr.:*** *BC = buffy coat; cDNA = complementary DNA; end. = endogenous; H = histidine; HEK = human embryonic kidney; HLA = human leukocyte antigen; K = lysine; KO = knock-out; KRAS = Kirsten rat sarcoma viral oncogene; L = leucine; mTRAC = murine T cell receptor alpha constant region; p. = amino acid position; Q = glutamine; R = arginine; TCR = T cell receptor.*

Generally, observed TCR-T cell recognition was higher in transgenic CD8+ effectors than in corresponding CD4+ counterparts (Fig. 3.2.8 A and B). Furthermore, when compared to the other P18 TCR-T cell populations, strongest T cell activation was again detected for TCR 73.1-expressing CD8+ effectors. Moreover, consistent with results shown in Figure 3.2.7, TCR 55.1 induced highest IFN-γ secretion among tested CD4+ effector populations, thereby indicating that TCR 55.1-mediated T cell activation might be less dependent on CD8 coreceptors. Taken together, the results shown in Figures 3.2.6 to 3.2.8 provide strong evidence that (i) the KRAS^{Q61H} neoantigen recognition is actually

accountable for observed T cell activations in P18 TCR-transgenic effector populations and that (ii) the corresponding 10-mer neopeptide is efficiently processed and presented upon transient transfection of HEK 293T cell cultures. However, further experiments are necessary to assess the extent to which the $KRAS^{Q61H}$ neoantigen is also recognized when being physiologically expressed in non-transfected target cells.

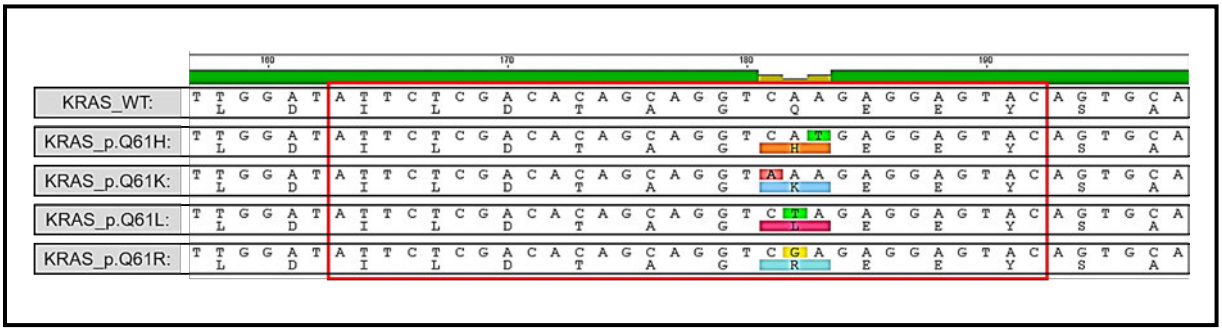


Figure 3.2.9: cDNA constructs encoding 4 mutations commonly found at amino acid position 61 of KRAS, HRAS, and NRAS.
*Ready-to-transfect cDNA expression constructs either encoding p.Q61H-, p.Q61K-, p.Q61L-, or p.Q61R-mutated KRAS full-length open reading frames were established by using site-directed mutagenesis (SDM). Resulting pcDNA3.1 expression constructs were used to transfect HLA-A*01:01+ HEK 293T target cells for ELISpot-based recognition testing. Sanger sequencing confirmed successful SDM reactions for all generated cDNA expression constructs. Neopeptide-encoding cDNA sequences are highlighted in red.* ***Abbr.:*** *cDNA = complementary DNA; H = histidine; HEK = human embryonic kidney; HLA = human leukocyte antigen; K = lysine; KRAS =Kirsten rat sarcoma viral oncogene; L = leucine; p. = amino acid position; Q = glutamine; R = arginine; SDM = site-directed mutagenesis; WT = wildtype.*

3.2.4 Three P18 TCRs recognize KRAS^{Q61H}-positive NCI-H460 tumor cells

To assess whether the mutated KRAS^{Q61H} neoantigen is also recognized when physiologically expressed and processed in non-transfected target cells, TCR-transduced T cell populations, each of which expressed one of the three KRAS^{Q61H} neoantigen-specific P18 TCRs (see section 3.2.3), were tested for the recognition of NCI-H460 lung cancer cells. Interestingly, the NCI-H460 NSCLC cell line naturally carries homozygous KRAS^{Q61H} mutations but lacks HLA-A*01:01 expression. Hence, NCI-H460 cells were firstly used for retroviral transduction approaches to establish transgenic HLA-A*01:01+ NCI-H460 cells, thereby representing suitable targets for further ELISpot-based antigen testing. Resulting NCI-H460_HLA-A*01:01 cells were recognized by CD8+ TCR-T cell populations either expressing P18 TCRs 28.1, 55.1, or 73.1. Furthermore, detected T cell activation levels were comparable across all tested CD8+ TCR-T cell populations (Fig. 3.2.10 A). In contrast, non-transduced HLA-A*01:01-negative NCI-H460 cells were not recognized by any of the tested effector populations. When analyzing CD4+ TCR-T cell populations, only TCR 55.1-expressing effector populations were substantially activated upon stimulation with NCI-H460_HLA-A*01:01 target cells (Fig. 3.2.10 B). This finding was consistent with results of previously conducted KRAS recognition assessments (see section 3.2.3), thereby again indicating that TCR 55.1-mediated KRAS^{Q61H} recognition might be less dependent on CD8 coreceptors.

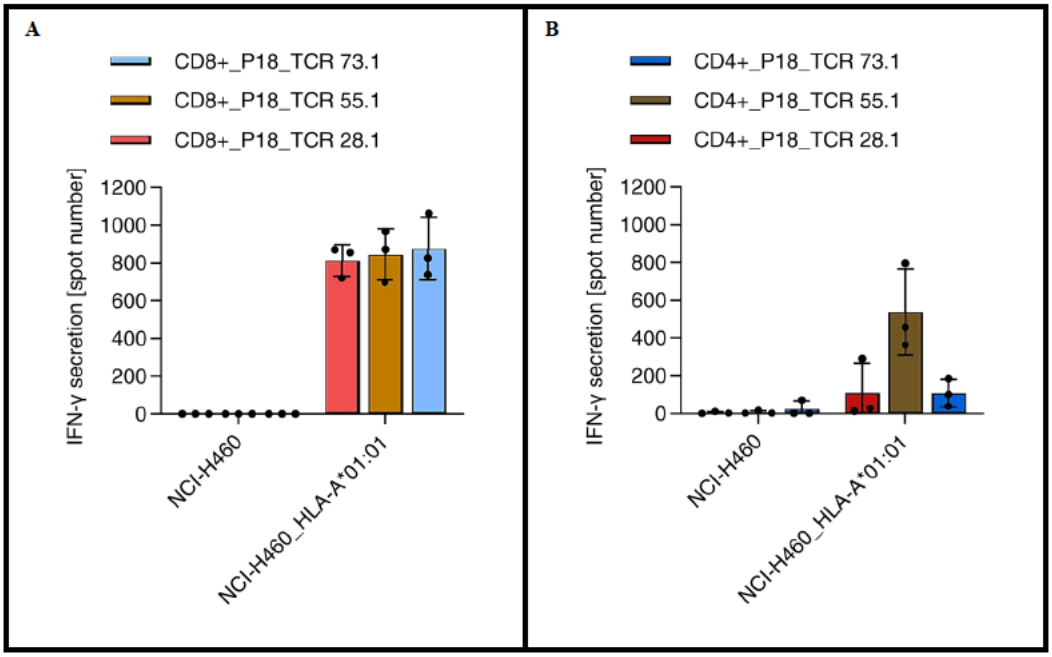


Figure 3.2.10: TCR-T cell populations expressing P18 TCRs 28.1, 55.1, or 73.1 recognize KRAS^{Q61H}-positive NCI-H460 cells.
*NCI-H460_HLA-A*01:01 recognition testing with P18 TCRs 28.1, 55.1, and 73.1. IFN-γ secretion was used as a marker for T cell recognition. CD4+ and CD8+ TCR-T cell populations were individually tested for the recognition of HLA-A*01:01-transduced NCI-H460 cells naturally carrying the p.Q61H KRAS mutation. **A)** All tested CD8+ effector populations recognized NCI-H460_HLA-A*01:01 cells but failed to recognize non-transduced NCI-H460 cells. **B)** When analyzing CD4+ TCR-T cell populations upon co-incubation with NCI-H460_HLA-A*01:01 target cells, substantial IFN-γ secretion was exclusively detected for TCR 55.1-expressing effectors. Non-transduced NCI-H460 cells were not recognized at all. **A/B)** Three independent experiments were performed. Shown spot numbers either represent mean values resulting from duplicates or single well testing results of individual experiments. Error bars represent standard deviations of the three experiments. Due to excessive IFN-γ secretion, maximal NCI-H460_HLA-A*01:01 recognition was defined as 1300 spots per well since exact spot numbers could not be assessed (TNTC). Spot numbers were eventually normalized for IFN-γ background levels of respective T cell cultures in the absence of target cells. ELISpot assays were performed with P18 TCR-T cell populations (Sup. Fig. 6) generated by using end.TCR-KO T cells (Sup. Fig. 5) derived from BCs 074 and 437. For BC074, ELISpots were performed 20 and 35 days after retroviral transduction while BC437-derived T cell cultures were used 39 days after introducing the P18 TCRs. Effector cells per well: 10,000 mTRAC+. Target cells per well: 50,000. Raw ELISpot data is shown in Sup. Fig. 7. **Abbr.:** BC = buffy coat; end. = endogenous; H = histidine; HLA = human leukocyte antigen; KO = knock-out; KRAS =Kirsten rat sarcoma viral oncogene; mTRAC = murine T cell receptor alpha constant region, NSCLC = Non-small cell lung cancer; p. = amino acid position; Q = glutamine; TCR = T cell receptor; TNTC = too numerous to count.*

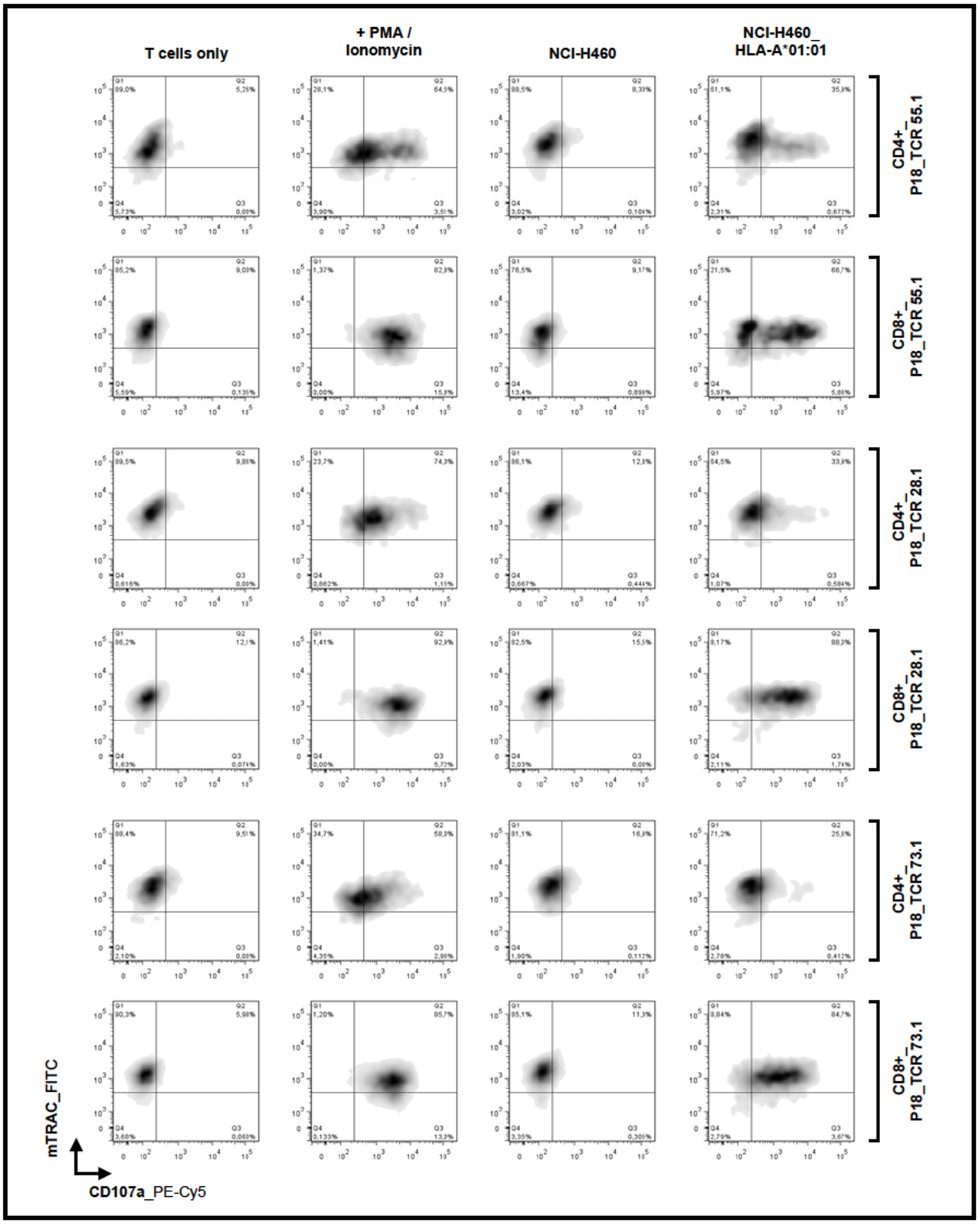


Figure 3.2.11: Cytotoxicity of P18 TCR-transduced T cell populations upon KRAS^{Q61H} recognition.

A-C) *Transgenic CD4+ and CD8+ TCR-T cell populations were co-incubated with NCI-H460_HLA-A*01:01 cells to assess CD107a surface expression upon target neoantigen recognition. CD107a was used as a surrogate marker for cytotoxicity. Elevated CD107a expression was detected for all tested T cell populations upon stimulation with HLA-A*01:01-positive NCI-H460 cells, but not when co-incubated with non-transduced NCI-H460 targets. When compared to CD4+ populations, CD107a expression was higher in TCR-transduced CD8+ T cells. CD107a expression levels were comparable across CD4+ and among most CD8+ T cell populations, respectively. However, detected CD107a upregulation seemed slightly lower in TCR 55.1-expressing CD8+ effectors. Non-specific T cell stimulation with Ionomycin and PMA was used as positive control for maximal CD107a expression. Flow cytometry staining was performed after 12 hours of co-incubation with target cells. Canonical CD107a expression levels were individually determined for every T cell culture in the absence of target cells or PMA/Ionomycin (T cells only preparations) and were eventually used for gating. BC074-derived end.TCR-KO T cells (Sup. Fig. 5 A) were tested 21 (CD8+) or 22 (CD4+) days after retroviral transduction with P18 TCRs (Sup. Fig. 6 A). P18 TCR expression was assessed via mTRAC staining.* ***Abbr.:*** *BC = buffy coat; end. = endogenous; FITC = Fluorescein isothiocyanate; H = histidine; HLA = human leukocyte antigen; KO = knock-out; KRAS = Kirsten rat sarcoma viral oncogene; mTRAC = murine T cell receptor alpha chain constant region; NSCLC = non-small cell lung cancer; p. = amino acid position; PE-Cy5 = Phycoerythrin-Cyanine5; PMA = Phorbol 12-myristat 13-acetate; Q = glutamine; TCR = T cell receptor.*

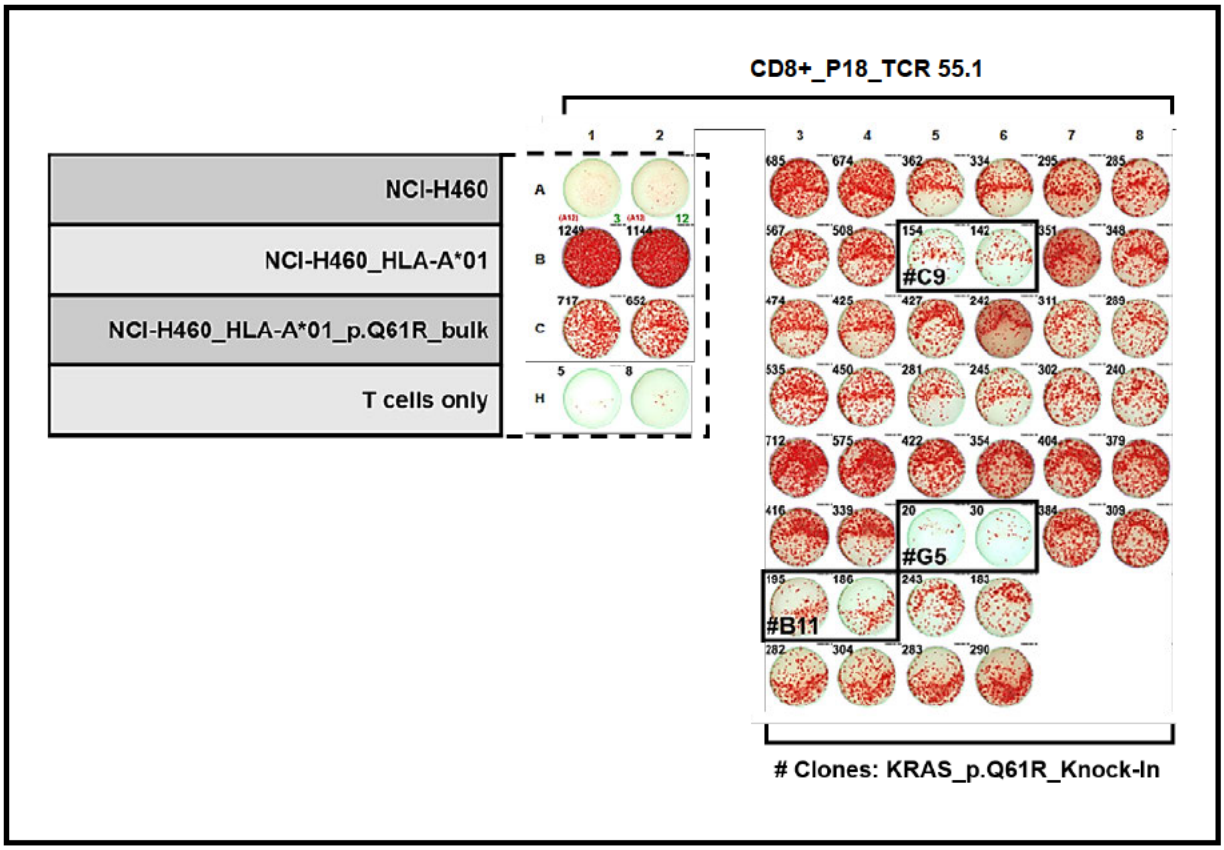


Figure 3.2.12: The HDR-mediated KI of KRAS^{Q61R} results in a reduced recognition of HLA-A*01:01+ NCI-H460 cells via P18 TCR 55.1. *An HDR-mediated knock-in approach was used to substitute biallelic p.Q61H KRAS mutations with corresponding p.Q61R variations in HLA-A*01:01+ NCI-H460 cells. After limiting dilution cloning of KRAS-modified bulk populations, resulting cell clones were tested for the recognition by TCR-T cell populations overexpressing P18 TCR 55.1. IFN-γ secretion was used as a marker for T cell recognition. Recognition testing with the KRAS-modified NCI-H460_HLA-A*01:01 bulk population is shown on the left while the recognition screening with corresponding cell clones is depicted on the right. The KRAS-modified NCI-H460_HLA-A*01:01 bulk population was less recognized than non-treated NCI-H460_HLA-A*01:01 cells (left). When compared to the KRAS-modified NCI-H460_HLA-A*01:01 bulk population (left), reduced T cell activation was clearly detected upon coincubation with HDR-treated clones #B11, #C9, and #G5 (right). Target cell recognition analyses were performed as duplicates. IFN-γ background levels were assessed by analyzing transgenic effector populations in the absence of target cells (T cells only preparations). Transgenic effector populations were generated from BC437-derived end.TCR-KO T cell populations (Sup. Fig. 5 B). ELISpot analysis was performed 25 days after retroviral transduction with the TCR 55.1-encoding expression construct (Sup. Fig. 6 A). Retrospective corrections of inaccurately counted wells are highlighted in green font. Used effector cells per well: 2,500 mTRAC+. Used target cells per well: 50,000. **Abbr.:** BC = buffy coat; end. = endogenous; H = histidine; HDR = homology-directed repair; HLA = human leukocyte antigen; KI = knock-in; KO = knock-out; KRAS = Kirsten rat sarcoma viral oncogene; mTRAC = murine T cell receptor alpha chain constant region; p. = amino acid position; Q = glutamine; R = arginine, TCR = T cell receptor.*

However, INF-γ secretion may not necessarily correlate with cytotoxic capacities of established TCR-T cell cultures. Hence, cytolytic capacities were assessed by analyzing CD107a surface expression of P18 TCR-T cell cultures upon recognition of the KRAS^{Q61H} neoantigen (Fig. 3.2.11). CD107a surface expression, which can be used as a surrogate marker for active degranulation, was detected on all analyzed T cell populations when co-incubated with HLA-A*01:01-transduced NCI-H460 cells. However, CD107a expression was higher in transduced CD8+ effectors than in corresponding CD4+ TCR-T cells. TCR-engineered CD8+ T cell populations expressed comparable CD107a levels upon co-incubation with NCI-H460_HLA-A*01:01 and after non-specific activation with Ionomycin/PMA (Phorbol 12-myristat 13-acetate), respectively. Here, Ionomycin/PMA T cell stimulation was used as a positive control to assess maximal CD107a expression levels. In contrast, when compared to non-specific T cell activation, NCI-H460_HLA-A*01:01 co-incubation induced substantially less CD107a expression in transgenic CD4+ T cells, thereby indicating lower overall cytotoxic capacities. Consistent with previous results, HLA-A*01:01-negative NCI-H460 cells did not induce substantial CD107a upregulations in P18 TCR-T cells when compared to non-stimulated effectors. Interestingly, CD107a expression levels were comparable across CD4+ and among most of the analyzed CD8+ TCR-T cell

populations. However, TCR 55.1-expressing CD8+ TCR-T cells exhibited slightly less CD107a upregulation (approx. 67% vs. 88% and 85%) and, thus, less degranulation than corresponding TCR 28.1+ and TCR 73.1+ cultures. This indicates that similar amounts INF-γ secretion (see Figure 3.2.10 A) may not necessarily correlate with equal cytotoxic properties.

To provide additional evidence that the KRAS^{Q61H} neoantigen is the actual target specifically recognized by all three P18 TCRs, a CRISPR/Cas9-based homology-directed repair (HDR) gene knock-in approach was used to substitute biallelic KRAS^{Q61H} mutations in HLA-A*01:01+ NCI-H460 cells with non-recognized KRAS^{Q61R} variations (Fig. 3.2.8 A and B). When compared to non-treated counterparts, the CRISPR/Cas9-treated NCI-H460_HLA-A*01:01 bulk population was found to be recognized less efficiently by TCR 55.1-expressing CD8+ TCR-T cells (Fig. 3.2.12). Hence, monoclonal tumor cell cultures were established from HDR-treated NCI-H460_HLA-A*01:01 bulk cultures via limiting dilution cloning and were eventually used for recognition screenings with p.Q61H-reactive P18 TCR-T cell populations. Here, the three tumor cell clones #B11, #C9, and #G5 were found to be recognized less efficiently by TCR 55.1-expressing CD8+ T cells (Fig. 3.2.12).

Furthermore, by testing transgenic CD4+ and CD8+ effector populations, each of which expressed one of the three P18-derived TCRs, it was shown that the biallelic substitution of KRAS^{Q61H} with KRAS^{Q61R} in clone #G5 (Fig. 3.2.13) resulted in almost complete loss of target recognition (Fig. 3.2.14 A and B). Since the absence of the KRAS^{Q61H} mutation was the only difference between non-recognized clone #G5 and non-treated NCI-H460_HLA-A*01:01 cells, it was successfully demonstrated that the KRAS^{Q61H}-derived neoepitope is the actual target antigen recognized by all three P18 TCRs. Subsequent sequencing analyses revealed that the reduced recognition detected for clone #B11 was due to a frameshift mutation in one of both *KRAS* alleles (Sup. Fig. 10 B) and that one non-modified *KRAS^{Q61H}* allele was still sufficient to induce substantial T cell activation via TCRs 28.1, 55.1, and 73.1 (Fig. 3.2.14 A and B). Since clone #C9 carried two non-modified *KRAS^{Q61H}* alleles (Sup. Fig. 10 A), no substantial difference was detected when comparing P18 TCR-T cell reactivity towards clone #C9 and non-treated NCI-H460_HLA-A*01:01 cells (Fig. 3.2.14 A and B).

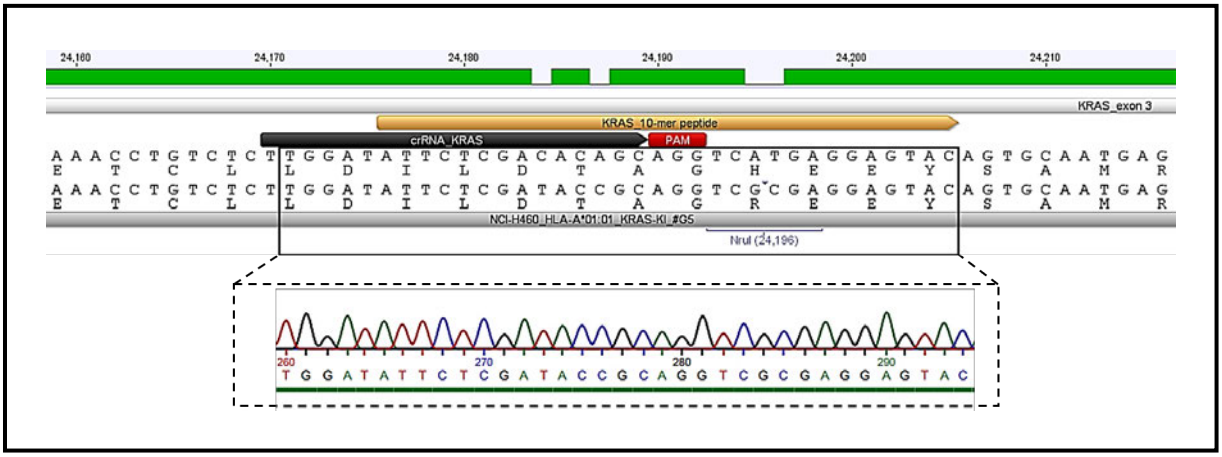


Figure 3.2.13: Sanger Sequencing reveals biallelic KRAS^{Q61R} mutations in HDR-treated NCI-H460_HLA-A*01:01 clone #G5.
*Sanger sequencing data of KRAS^{Q61R}-modified NCI-H460_HLA-A*01:01 clone #G5. Biallelic KRAS^{Q61H} mutations were substituted with non-recognized p.Q61R variations by using a CRISPR/Cas9-based HDR knock-in approach. Substantially reduced IFN-γ secretion was detected upon co-incubation of P18 TCR 55.1-expressing effector cells and KRAS-modified clone #G5 (Fig. 3.2.12). Sanger sequencing data of gDNA and translated amino acid sequences of clone #G5 are shown below corresponding reference sequences of KRAS^{Q61H} exon 3. Related #G5 gDNA sequencing plots are shown in the dashed box. Successful introduction of p.Q61R mutations was verified for both KRAS alleles of clone #G5. The crRNA binding sequence is shown in black while corresponding PAM sequence is highlighted in red. The immunogenic KRAS 10-mer neopeptide sequence is depicted in orange. KRAS reference gDNA bp positions are indicated above.* ***Abbr.:*** *bp = base pair; crRNA = CRISPR-RNA; gDNA = genomic DNA; HDR = homology-directed repair; H = histidine; HLA = human leukocyte antigen; KI = knock-in; KRAS = Kirsten rat sarcoma viral oncogene; p. = amino acid position; PAM = protospacer adjacent motif; Q = glutamine; R = arginine.*

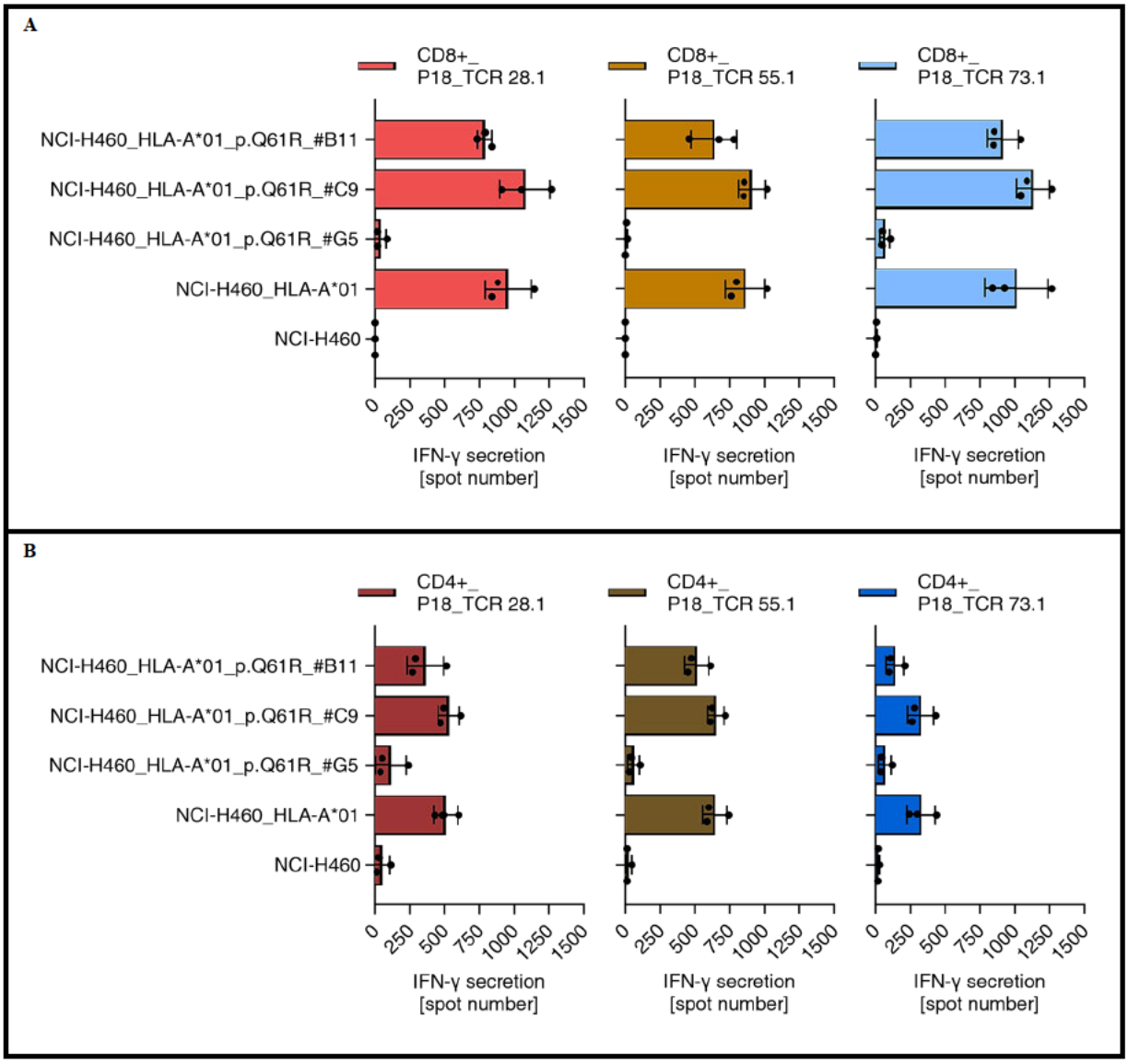


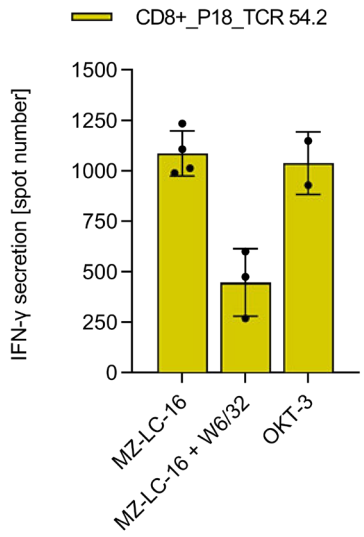
Figure 3.2.14: P18 TCRs 28.1, 55.1, and 73.1 do not recognize KRAS^{Q61H}-negative NCI-H460_HLA-A*01:01 clone #G5.

A and B) *By using a CRISPR/Cas9-based HDR knock-in approach, biallelic KRAS^{Q61H} mutations were substituted with p.Q61R variations in HLA-A*01:01-transduced NCI-H460 clone #G5. Recognition testing was performed with transgenic CD4+ and CD8+ T cell populations expressing P18-derived TCRs 28.1, 55.1, or 73.1. IFN-γ secretion was used as a marker for T cell recognition. Recognition testing of CD8+ T cell populations is shown in A) while ELISpot results detected for CD4+ effectors are shown in B). #G5 recognition was absent across all tested CD4+ and CD8+ effector populations. Due to one non-modified KRAS^{Q61H} allele after HDR treatment (Sup. Fig. 10 B), recognition of clone #B11 was slightly reduced. IFN-γ secretion upon co-incubation with non-modified clone #C9 (Sup. Fig. 10 A) was not reduced at all. T cell activation levels were elevated in transgenic CD8+ effectors when compared to corresponding CD4+ populations. Y axis labelling shown on the left applies to all assigned plots. Shown spot numbers represent mean values calculated from duplicates of three independent experiments. Error bars represent standard deviations of the three experiments. Shown spot numbers were additionally normalized for IFN-γ background levels of T cells in the absence of targets. ELISpot assays were either performed 36 days (CD4+) or 35 days (CD8+) after retroviral transduction of with P18 TCRs. BC437-derived end.TCR-KO T cells (Sup. Fig. 5 B) were used for retroviral transduction (Sup. Fig. 6 B) with P18 TCR expression constructs. Used effectors per well: Either 2,500 or 7,500 P18 TCR+. Used target cells per well: 50,000. Raw ELISpot data is shown in Sup. Fig. 11.* ***Abbr.:*** *BC = buffy coat; end. = endogenous; H = histidine; HDR = homology-directed repair; HLA = human leukocyte antigen; IFN = interferon; KO = knock-out; KRAS = Kirsten rat sarcoma viral oncogene; p. = amino acid position; Q = glutamine; R = arginine; TCR = T cell receptor.*

3.2.5 P18 TCR 54.2 recognizes a non-identified target antigen present in HLA-A*02:01-positive MZ-LC-16 lung cancer cells

Although triggering the strongest T cell activation of all P18 TCRs when co-incubated with P18 tumor cells (Fig. 3.2.3), TCR 54.2 failed to recognize the mutated KRAS^{Q61H} 10-mer neopeptide or any of the other neoantigen candidates tested for patient P18 (see section 3.2.2). However, TCR 54.2-transduced T cells recognized the HLA-A*02:01-positive lung carcinoma cell line MZ-LC-16 in an HLA class I-dependent manner. The latter was repeatedly demonstrated by blocking the target cell recognition of TCR 54.2-expressing effectors via the HLA class I-specific antibody W6/32 (Fig. 3.2.15 A). Since no other HLA-I allele was shared between patient P18 and MZ-LC-16 (Fig. 3.2.15 B), these results indicated that TCR 54.2 might recognize an HLA-A*02:01-presented, presumably non-mutated, antigen expressed in both tumors. Hence, multiple HLA-A*02:01-positive tumor cell lines from various malignancies were tested (Fig. 3.2.15 C), but none of the tested tumors were found to be recognized by TCR 54.2-expressing CD8+ effector populations (data not shown).

A



B

	HLA-A*	HLA-B*	HLA-C*
P18	01:01:01	08:01:01	03:04:01
	02:01:01	40:02:01	07:01:01

	HLA-A*	HLA-B*	HLA-C*
MZ-LC-16	**02:01:01**	07:02:01	07:02:01
	11:01:01	14:02:01	08:02:01

C

Additionally tested HLA-A*02:01-positive tumor cell lines	Tumor type
MZ-LC-1, MZ-LC-6; NW11-BC	Lung carcinoma
Ma-Mel-86a, Ma-Mel-86c, D05-Mel, D28-Mel, D41-Mel	Melanoma
MZ-418, MZ-653, OCI-AML3, HL-60	AML
K562_P18_HLAs (+ CEFT peptide mix)	CML (+ viral antigens)
SW480, SK-CO-1, HT29, LS1034	Colon carcinoma
MZ-PC-2	Pancreatic carcinoma

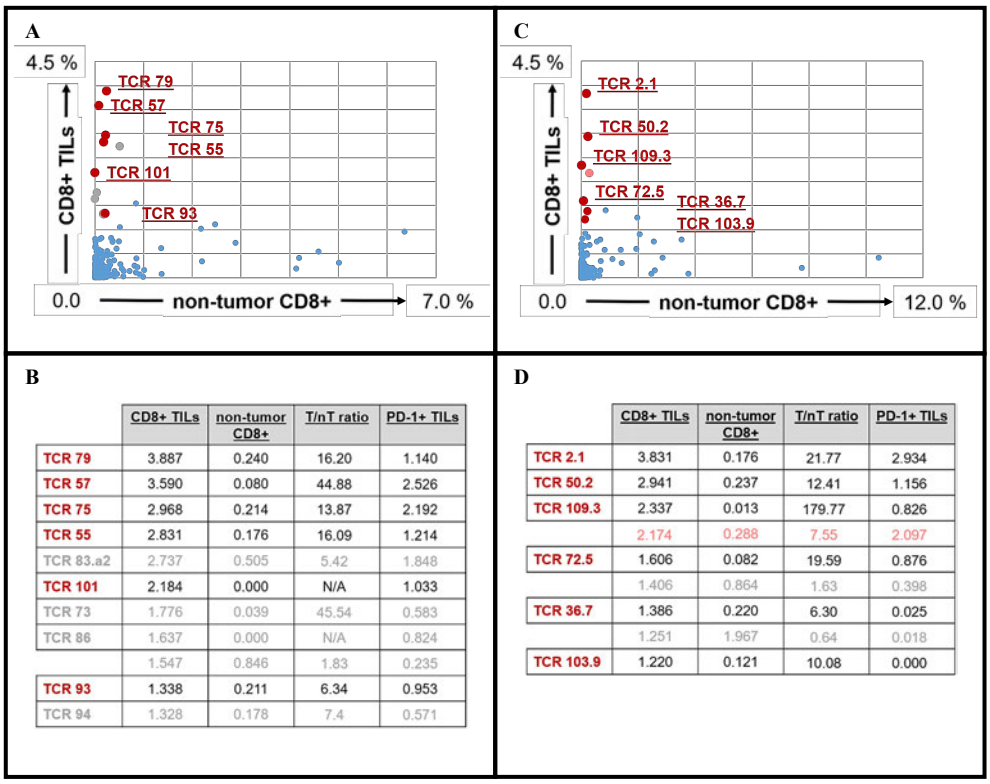
Figure 3.2.15: P18 TCR 54.2-expressing T cells recognize the HLA-A*02:01+ lung cancer cell line MZ-LC-16 via HLA class I.
A) *MZ-LC-16 recognition via P18-derived TCR 54.2. IFN-γ secretion was used as a marker for T cell recognition. MZ-LC-16 recognition by TCR 54.2-expressing CD8+ TCR-T cells was blocked by using the anti-HLA class I antibody W6/32, thereby indicating an HLA class I-dependent antigen recognition. Non-specific stimulation via the anti-CD3 monoclonal antibody OKT-3 was used as positive control for maximal T cell activation. Shown spot numbers represent mean values calculated from duplicate test wells of up to four independent experiments. Error bars represent standard deviations of the respective experiments. Shown spot numbers were additionally normalized for IFN-γ background levels of respective T cell populations in the absence of target cells. BC827-derived end.TCR-KO T cells (Sup. Fig. 1) were used for retroviral transduction (Fig. 3.2.2) with P18 TCR 54.2. ELISpot assays were performed between 16 days and 20 days after retroviral transduction with TCR 54.2-encoding expression constructs. Effector cells per well: Either 5,000 or 10,000 P18 TCR 54.2+ T cells. Target cells per well: 50,000. OKT-3 concentration per well: 400 ng/ml. W6/32 concentration per well: 50 µg/ml. Raw ELISpot data is shown in Sup. Fig. 12.* ***B)*** *Summary of HLA class I alleles expressed by NSCLC patient P18 and lung cancer cell line MZ-LC-16. gDNA sequencing was used for HLA genotyping of P18-derived tumor cells and MZ-LC-16. Only HLA-A*02:01 was shared between both tumors. HLA-A alleles are shown in orange while HLA-B and -C alleles are depicted in purple and green, respectively.* ***C)*** *Additionally tested HLA-A*02:01+ tumor cell lines. Various HLA-A*02:01+ tumor cell lines were tested for the recognition via TCR 54.2. None of the tested tumor cell lines induced substantial IFN-γ secretion upon co-incubation with TCR 54.2-positive CD8+ T cell populations (data not shown).* ***Abbr.:*** *AML = acute myeloid leukemia; BC = buffy coat; CEFT = HLA class I-restricted peptide mix derived from cytomegalovirus, Epstein-Barr virus, Influenza virus, and Tetanus toxin; CML = chronic myeloid leukemia; end. = endogenous; HLA = human leukocyte antigen; KO = knock-out; LC = lung cancer; NSCLC = non-small cell lung cancer; TCR = T cell receptor.*

Furthermore, comparative variant analysis via WES did not reveal any non-synonymous SNVs, MNVs, or InDels shared between P18 NSCLC and MZ-LC-16 (data not shown). Thus, no additional neoantigen candidates were identified for further TCR 54.2 specificity assessments. Moreover, differential gene expression analyses were performed to identify potentially expressed cancer germline (CG) antigens or overexpressed TAAs shared between the P18 tumor tissue and MZ-LC-16 cells (data not shown). Resulting candidate antigen cDNAs (CT83, MAGEA12, XAGE1A) were amplified from tumor-derived RNA, were then cloned into pcDNA3.1 expression constructs, and were eventually added to a pre-existing ready-to-transfect cDNA panel consisting of 41 common TAAs and CG antigens (Sup. Fig. 13). P18 TCR 54.2-positive CD8+ T cells were finally tested for the recognition of transfected HEK 293T target cell populations, each of which expressed one of the panel antigens and HLA-A*02:01. To also assess potential cross-reactivities of TCR 54.2-transduced effector populations towards virus-derived antigens, CD8+ TCR-T cells were additionally tested for their recognition of common virus peptides. Hence, P18 HLA-expressing K562 cells were pulsed with a CEFT peptide mix (JPT Peptide Technologies GmbH, Berlin, Germany) containing immunogenic HLA class I-restricted antigen peptides from cytomegalovirus, Epstein-Barr virus, Influenza virus, and Tetanus toxin. Regardless of expressed HLA-I alleles, none of the tested antigen panel cDNA constructs or virus epitopes were recognized at all (data not shown). Although previous experiments indicated the P18 tumor-reactivity of TCR 54.2-expressing effector populations (see Figure 3.2.3), the actual TCR 54.2 target antigen, which might be present in both P18 tumor as well as lung cancer cell line MZ-LC-16, could not be identified during this study.

3.3 Potentially tumor-specific TCRs cloned from tumor-infiltrating lymphocytes of NSCLC patients P43 and P50

3.3.1 Identification of potentially tumor-specific TCR clonotypes in NSCLC patients P43 and P50

To further assess the efficiency of the HSD-proprietary prediction pipeline aiming at the identification of potentially tumor-specific T lymphocytes, antigen specificities of TCRs cloned from tumor-infiltrating lymphocyte (TIL) clonotypes of two other stage III NSCLC patients, namely patients P43 and P50, were additionally analyzed. As for P18, potentially tumor-specific TIL candidates were identified via TRB profiling of CD3+, CD4+, CD8+, and PD-1+ T cell subsets previously isolated from tumor and adjacent lung tissues of both patients, respectively. Exact TCR alpha and beta chain amino acid sequences of TIL candidates exhibiting very high T/nT ratios were subsequently revealed via RNA-Seq. Most promising TCRs, referring to TCRs expressed by PD-1+ TIL clonotypes with highest T/nT ratios, were finally synthesized and used for the generation of retroviral TCR expression constructs.



B

	CD8+ TILs	non-tumor CD8+	T/nT ratio	PD-1+ TILs
TCR 79	3.887	0.240	16.20	1.140
TCR 57	3.590	0.080	44.88	2.526
TCR 75	2.968	0.214	13.87	2.192
TCR 55	2.831	0.176	16.09	1.214
TCR 83.a2	2.737	0.505	5.42	1.848
TCR 101	2.184	0.000	N/A	1.033
TCR 73	1.776	0.039	45.54	0.583
TCR 86	1.637	0.000	N/A	0.824
	1.547	0.846	1.83	0.235
TCR 93	1.338	0.211	6.34	0.953
TCR 94	1.328	0.178	7.4	0.571

D

	CD8+ TILs	non-tumor CD8+	T/nT ratio	PD-1+ TILs
TCR 2.1	3.831	0.176	21.77	2.934
TCR 50.2	2.941	0.237	12.41	1.156
TCR 109.3	2.337	0.013	179.77	0.826
	2.174	0.288	7.55	2.097
TCR 72.5	1.606	0.082	19.59	0.876
	1.406	0.864	1.63	0.398
TCR 36.7	1.386	0.220	6.30	0.025
	1.251	1.967	0.64	0.018
TCR 103.9	1.220	0.121	10.08	0.000

Figure 3.3.1: Identification of potentially tumor-specific TIL clonotypes in NSCLC patients P43 and P50.
A and B) *T/nT frequencies of most prominent PD-1+ P43-derived TIL clonotypes. Most promising TIL clonotypes were initially identified via comparative TRB sequencing of T cells isolated from P43 tumor and healthy lung tissues. A total of 10 clonotypes exhibited high T/nT ratios as well as high frequencies among PD-1+ TILs. Due to non-viable T cell populations after retroviral transduction, TCRs shown in grey were excluded from subsequent experiments.* ***C and D)*** *T/nT frequencies of most prominent PD-1+ P50-derived TIL clonotypes. As per A/B), most promising P50 TIL clonotypes were initially identified via comparative TRB sequencing. A total of 7 clonotypes exhibited high T/nT ratios as well as high frequencies among PD-1+ TILs. Although exhibiting different TCR-encoding nucleotide sequences, TCR 2.1-derived TRA and TRB amino acid sequences were also expressed by another T cell clonotype. One of those TIL clonotypes (highlighted in light red) was therefore excluded from further analyses.* ***A-D)*** *TCR IDs of clonotypes exhibiting rather low T/nT ratios and/or low frequencies among PD1+ fractions (represented as blue dots) are not shown and were generally excluded from further analyses. TIL-derived TCRs used for further antigen screenings are highlighted in red. Clonotype frequencies are shown as percentages among CD8+ TILs, lung tissue-derived CD8+ T cells, and as percentages among PD-1+ TILs.* ***Abbr.:*** *ID = identifier; NSCLC = non-small cell lung cancer; TIL = tumor-infiltrating lymphocyte; TRA = T cell receptor alpha chain; TRB = T cell receptor beta chain; T/nT = tumor-to-non-tumor; TCR = T cell receptor.*

TRB profiling of TILs isolated from P43 tumor tissue revealed a total of 23,431 different T cell clonotypes while 22,959 clonotypes were identified for P50 (data not shown). Most prominent T cell clonotypes from both patients were ranked based on their tumor frequencies and were then plotted for their respective T/nT ratios (Fig. 3.3.1 A and C). Based on high T/nT ratios and highest frequencies among PD-1+ TILs, 10 P43-derived T cell clonotypes were identified to express potentially tumor-specific TCRs. However, after individually transducing donor lymphocyte cultures with each of the 10 corresponding TCR expression constructs, low TCR expression and scant cell proliferation was observed in 4 of the established P43 TCR-T cell populations. Similar results were achieved when transducing healthy donor CD8+ T lymphocytes previously isolated from an alternative BC (data not shown). Hence, P43 TCRs encoded by those expression constructs were excluded from subsequent antigen recognition screenings, thereby focusing all efforts on only 6 TCRs (Fig. 3.3.1 B).

In case of NSCLC patient P50, 7 TCR clonotypes were found to be highly enriched among TILs as well as PD-1+ T cell fractions (Fig. 3.3.1 D). TCRs expressed by those clonotypes were therefore considered potentially tumor-specific. Although all 7 TILs exhibited differing TCR-encoding gDNA sequences, TCR 2.1 TRA and TRB amino acid sequences were expressed by two of the identified clonotypes. Hence, one of those TIL clonotypes was excluded from subsequent specificity analyses (Fig. 3.3.1 C and D), thereby reducing the number of P50 TCR candidates to a total of 6. In contrast to the KRAS^{Q61H}-specific TCRs cloned from NSCLC patient P18 (see section 3.2.1), potentially tumor-specific TIL candidates of patients P43 and P50 expressed highly variable TCRs (Fig. 3.3.2 A and B).

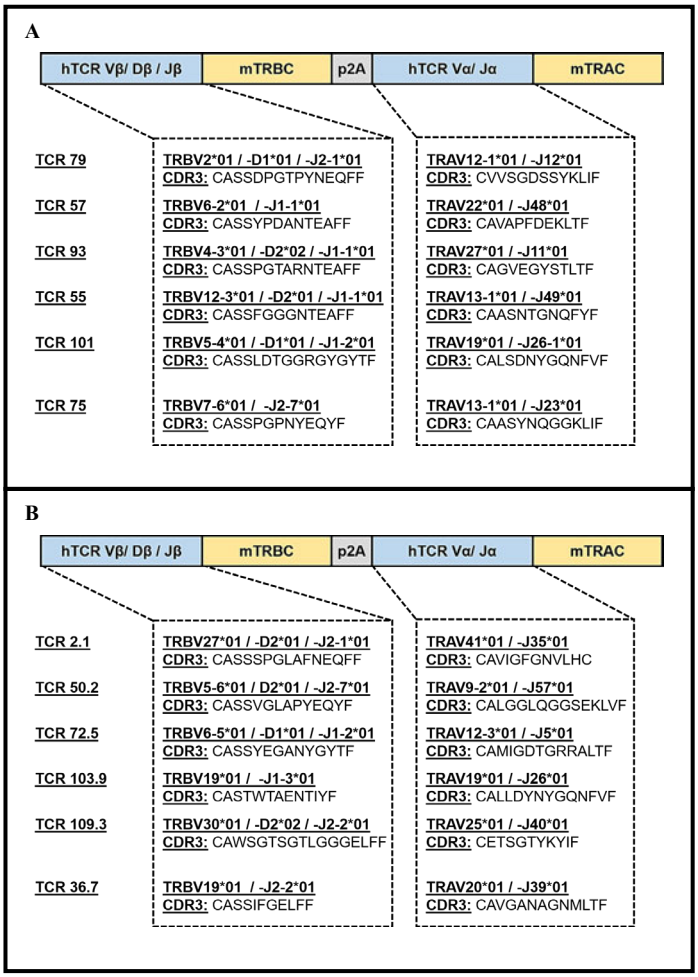


Figure 3.3.2: Potentially tumor-specific P43 and P50 TIL clonotypes express highly variable TCRs.
*Potentially tumor-specific P43 TCRs are shown in **A)** while potentially tumor-specific P50 TCRs are depicted in **B).** Complete TCR alpha and beta chain nucleotide sequences were revealed via single T cell sequencing. Respective CDR3 amino acid sequences are shown for each of the analyzed receptors. All listed P43 and P50 TIL clonotypes expressed highly variable TCRs. Schematic overviews of the applied TCR-encoding expression constructs are shown on top of each figure. Established pMX expression constructs were later used for retroviral transductions of healthy donor lymphocytes. **Abbr.:** CDR3 = Complementarity-determining region 3; hTCR = human T cell receptor; mTRAC = murine TCR alpha chain constant region; mTRBC = murine TCR beta chain constant region; TIL = tumor-infiltrating lymphocyte; TCR = T cell receptor.*

TCR expression constructs were generated comprising paired αβ TCR sequences of the 6 potentially tumor-specific TIL clonotypes either identified in patient P43 or P50, respectively. TCR-T cell populations were established via retroviral transduction of CD8+ donor lymphocytes after performing a

CRISPR/Cas9-based end.TCR-KO approach. Here, utilized T cells had initially been isolated from two different healthy donor buffy coats, namely BC248 and BC652. In both cases, the conducted end.TCR-KO approach resulted in residual TCR expression levels of less than 4% (Sup. Fig. 14 A and B). Furthermore, subsequent retroviral T cell transductions resulted in expression levels ranging between 75.5% and 97.9% for P43 TCRs (Sup. Fig. 15). On the other hand, expression levels between 71.2% and 96.2% were detected for retrovirally transduced P50 TCR-T cell populations (Sup. Fig. 16 and 17).

To further assess potential tumor specificities of the 12 patient-derived TCRs, established TCR-T cell populations were individually tested for the recognition of P43 or P50 tumor cell suspensions, respectively. When compared to P43 and P50 healthy lung cell recognition, tumor reactivity was detected for P43 TCRs 55, 79, and 93 (Fig. 3.3.3 A). However, high IFN-γ background levels were detected when analyzing P43 NSCLC (447 spots) and healthy lung (366 spots) cell suspensions in the absence of effector cells (Sup. Fig. 18 B), thereby suggesting that the gathered tumor reactivity data might be rather inconclusive. Furthermore, no tumor reactivity could be demonstrated for any of the analyzed P50 TCR-T cell populations (Fig. 3.3.3 B). Taken together, tumor reactivity was not clearly demonstrated for any of the tested P43 or P50 TCR-T cell cultures, which was most likely due to the poor quality (low viability after thawing) of analyzed tumor and healthy lung tissue cell suspensions.

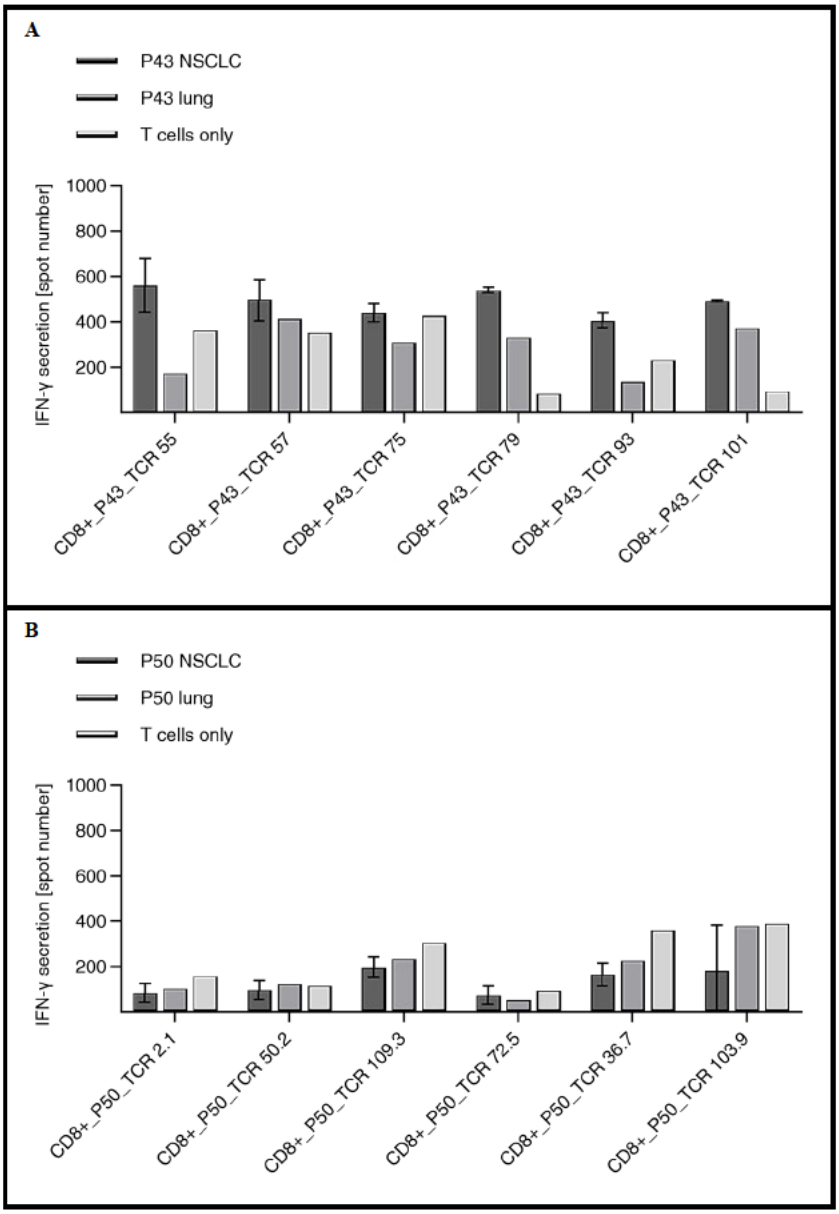


Figure: 3.3.3: Tumor recognition testing of potentially tumor-specific P43 and P50 TCRs.

A and B) *CD8+ TCR-T cell populations were tested for the recognition of P43- and P50-derived NSCLC single cell suspensions. IFN-γ secretion was used as a marker for T cell recognition. Refer to Sup. Fig. 18 A and B for raw ELISpot data and further information about used cell numbers. Spot numbers either represent mean values resulting from duplicates (NSCLC) or single well (lung and T cells only) analyses. Error bars represent standard deviations of duplicates.* ***A)*** *Tumor recognition testing with P43-derived TCRs. Tumor reactivity was detected for P43 TCRs 55, 79, and 93. Poor qualities of thawed P43 tumor cell suspensions prohibited the repetition of the experiment. ELISpot assays were performed between 12 and 69 days after retroviral transduction (Sup. Fig. 15) of BC248- or BC652-derived end.TCR-KO (Sup. Fig. 14 A and B) T cell cultures with P43 TCRs.* ***B)*** *Tumor recognition testing with P50-derived TCRs. Tumor reactivity could not be demonstrated for any of the tested P50 TCR-T cell populations. Poor qualities of thawed P50 tumor cell suspensions prohibited the repetition of the experiment. ELISpot assays were performed between 12 and 19 days after retroviral transduction (Sup. Fig. 16) of BC248-derived end.TCR-KO (Sup. Fig. 14 A) T cell cultures with P50 TCRs.* ***Abbr.:*** *BC = buffy-coat; end. = endogenous; KO = knock-out; NSCLC = non-small cell lung cancer; TCR = T cell receptor; TIL = tumor-infiltrating lymphocyte.*

3.3.2 *In silico* neoantigen candidate prediction for NSCLC patients P43 and P50

To verify the suitability of the HSD-proprietary prediction pipeline aiming at the identification of tumor-specific TIL clonotypes, revealing exact neoantigen epitopes specifically recognized by potentially tumor-specific P43 and P50 TCRs represented one of the major goals of this study. As previously described for patient P18 (see section 3.2), tumor-specific non-synonymous mutations were identified in both patients by using comparative whole exome sequencing (WES) of gDNA samples either isolated from patient-derived tumor or normal lung tissues. RNA-Seq of tumor-obtained total RNA was eventually used to screen for tumor-specific mutations that were actually expressed (mRNA level) in NSCLC tissues of patients P43 and P50, respectively.

Overall read mapping efficiencies of >95.5% to the reference genome hg38 confirmed the high quality of both WES and RNA-Seq data used for the P43 and P50 variant identification (data not shown). More than 14,500 tumor-specific variants were identified for P43 by excluding all variants also found in adjacent lung tissue. In contrast, 10,416 tumor-specific mutations were found by analyzing the respective P50 tissue specimens (Fig. 3.3.4 A and 3.3.5 A). By filtering for tumor expression (RPKM ≥0.5), amino acid (AA) changing consequences, and after manual curation to remove premature stop-codons or false InDels, 187 potential P43 neoantigen candidate mutations were finally identified. Those candidate variants included 178 single nucleotide variants (SNVs), 6 multiple nucleotide variants (MNVs), and 3 insertions/deletions (InDels). In contrast, the same i*n silico* pipeline identified 105 tumor-specific P50 neoantigen candidate mutations, which included 63 SNVs, 15 MNVs, 21 InDels, and 6 gDNA rearrangements.

HLA binding predictions were performed for both patients and, in case of P43, accounted for 849 mutated 9- and 10-mer candidate peptides predicted to effectively bind one of the patient's MHC molecules (Fig. 3.3.4 B). P50 HLA binding predictions resulted in a total of 629 9- and 10-mer neopeptide candidates (Fig. 3.3.5 B). As per patient P18 (see section 3.2), the 94 top scoring 9- and 10-mer neopeptide candidates (percentile rank ≤6.0 and/or binding affinity IC50 ≤500nM) were eventually synthesized for further P43 and P50 neoantigen screening approaches. Two control peptides encoded by mutated *MMS22L* and *INSIG1* genes [226] were additionally synthesized, thereby resulting in a total number of 96 peptides per patient. Hence, patient-specific peptide libraries (Fast Track Peptide Library; JPT, Berlin, Germany) included the 14-16 top scoring 9- and 10-mer neopeptide candidates predicted for each of the P43- or P50-derived HLA alleles, respectively. Synthesized control peptides were used to assess the overall quality of the generated peptide screening libraries. After verifying the recognition of both control antigens by control CTLs 16C/114 and 3A/115 [226] (data not shown), peptide library qualities were considered sufficient for further target antigen screenings.

While verifying the proper functionality of all used CD8+ TCR-T cell populations via non-specific OKT-3 activation, none of the predicted candidate neopeptides was specifically recognized by any of the potentially tumor-specific P43 or P50 TCRs shown in Fig. 3.3.2 (data not shown). In addition, a tumor antigen cDNA panel was tested consisting of 43 common TAAs, cancer germline (CG) antigens, overexpressed tumor antigens, and two CG antigens exclusively found in P50 tumor cells (Sup. Fig. 13). However, no T cell reactivity was detected towards the tested panel antigens. Taken together, no neoantigen specificities or TAA reactivities could be demonstrated for any of the analyzed P43 or P50 TCRs.

Since a total of 187 tumor-specific non-synonymous mutations were identified for P43 (Fig. 3.3.4 A), remaining non-analyzed neoantigen candidate mutations and even more TAAs were additionally tested by using a tandem minigene (TMG) cDNA library screening approach. Each of the 30 analyzed TMG constructs encoded a continuous polypeptide including 4 to 5 candidate mutations or TAA epitopes. Individual mutations were C- and N-terminally flanked by base pairs encoding the directly adjacent amino acids of the respective genes, thereby facilitating the generation of every possible 8- to 12-mer peptide with a particular candidate mutation at any possible amino acid position. As for the previously tested peptide screening libraries, the quality of the synthesized cDNA library (50 ng to 2 µg of clonal

genes; Twist Bioscience, South San Francisco, USA) was verified by demonstrating the control CTL-mediated recognition of two TMG constructs either encoding mutated *MMS22L* or *INSIG1* targets [226]. Due to the rather canonical peptide processing/presentation in target cells and since TMG recognition screenings were performed for every single P43 HLA allele, the TMG approach was independent from previous HLA binding predictions, thereby representing a very promising strategy for the efficient identification of target antigens. However, it could not be demonstrated that any of the tested TMG constructs, together accounting for approx. 120 additional non-synonymous P43 tumor mutations and TAAs, was specifically recognized by analyzed P43 TCR-T cell populations (data not shown).

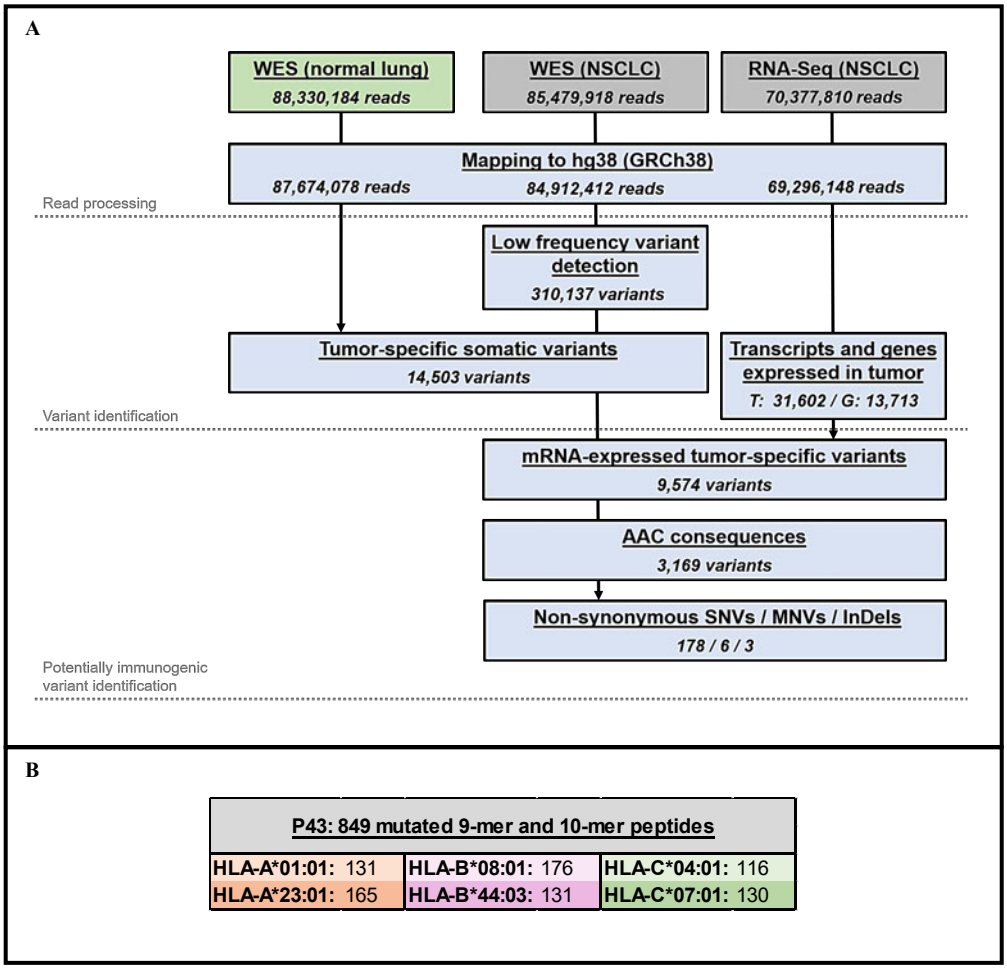


Figure 3.3.4: Antigenic landscape of NSCLC patient P43.

A) *In silico neoantigen prediction pipeline for the identification of non-synonymous tumor-specific mutations in NSCLC patient P43. Comparative WES was performed to identify a total of 14,503 tumor-specific mutations. By filtering for missense mutations, subsequent alignment with corresponding P43 tumor gene expression data, and after manual curation, 187 non-synonymous tumor-specific mutations were considered promising candidates for further peptide binding prediction. Identified mutations included 178 SNVs, 6 MNVs, and 3 InDels. All P43 WES and corresponding RNA-Seq reads (paired end reads; 150 nt) were mapped to the reference genome hg38 (GRCh38). NGS data was analyzed and processed by using the CLC Genomics Workbench software (versions 20-23).* ***B)*** *HLA class I binding predictions for neoantigen candidates identified in NSCLC patient P43. HLA binding predictions were individually performed for each P43 HLA-I allele by using mutated 19-mer peptides, each of which represented one of the 187 non-synonymous mutations. Most promising neopeptides were identified as a consensus of the HLA binding algorithms NetMHC-4.0 (binding affinities of ≤500 nM) and IEDB MHC-I binding prediction tool (Consensus or NetMHCpan-4.0 prediction methods; percentile rank of ≤6.0). P43 HLA-I binding predictions resulted in a total of 849 promising 9-/10-mer candidate neopeptides.* ***Abbr.:*** *AAC = amino acid-changing; G = genes; HLA = human leukocyte antigen; InDels = insertions and deletions; MNV = multiple nucleotide variations; NGS = Next Generation Sequencing; NSCLC = non-small cell lung cancer; nt = nucleotides; RNA-Seq = RNA Sequencing; RPKM = reads per kilobase of transcript per million reads mapped; SNV = single nucleotide variation; T = transcripts; WES = Whole Exome Sequencing.*

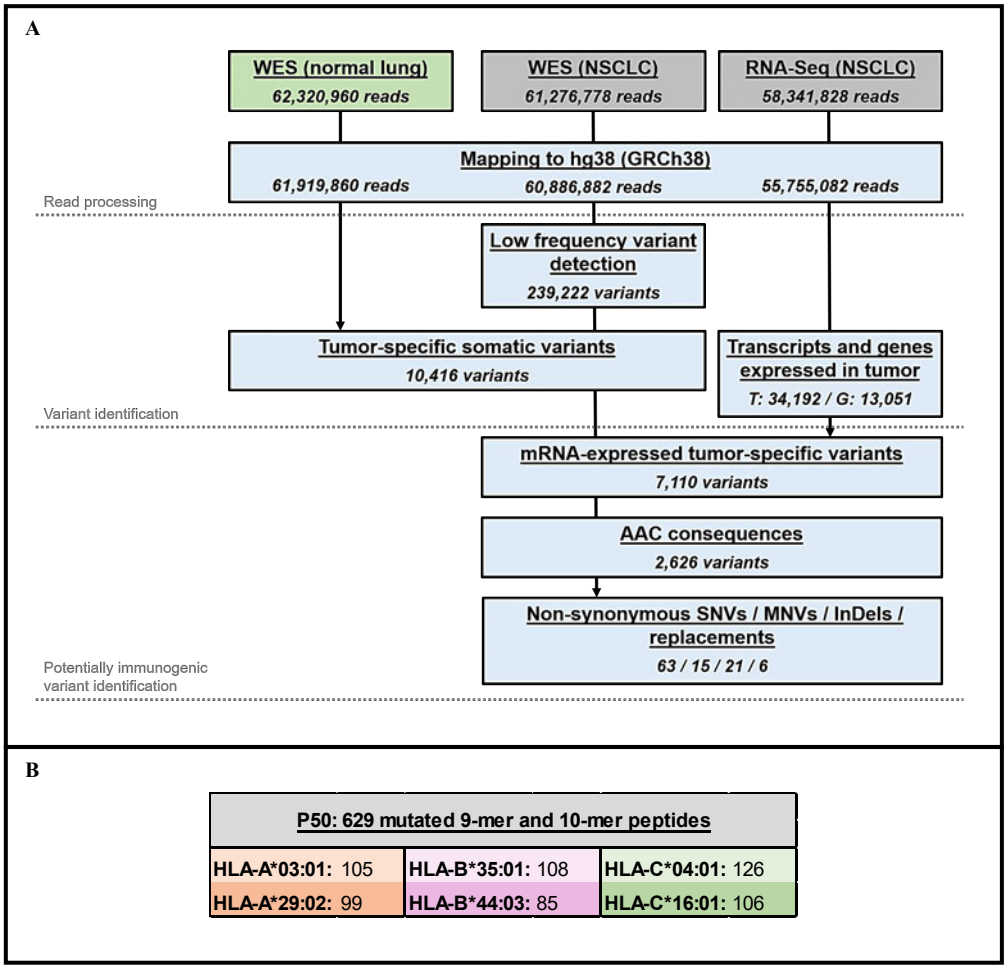


***Figure 3.3.5: Antigenic landscape of NSCLC patient P50*.**

A) *In silico neoantigen prediction pipeline for the identification of non-synonymous tumor-specific mutations in NSCLC patient P50. Comparative WES was performed and identified a total of 10,416 tumor-specific mutations. By filtering for missense mutations, subsequent alignment with corresponding P50 tumor gene expression data, and after manual curation, 105 non-synonymous tumor-specific mutations were considered promising candidates for further peptide binding prediction. Identified mutations included 63 SNVs, 15 MNVs, 21 InDels, and 6 genomic replacements. All P50 WES and corresponding RNA-Seq reads (paired end reads; 150 nt) were mapped to the reference genome hg38 (GRCh38). NGS data was analyzed and processed by using the CLC Genomics Workbench software (versions 20-23).* ***B)*** *HLA class I binding predictions for neoantigen candidates identified in NSCLC patient P50. HLA binding predictions were individually performed for each P50 HLA-I allele by using mutated 19-mer peptides, each of which represented one of the 105 neoantigen candidate mutations. Most promising neopeptides were identified as a consensus of the HLA binding algorithms NetMHC-4.0 (binding affinities of ≤500 nM) and IEDB MHC-I binding prediction tool (Consensus or NetMHCpan-4.0 prediction methods; percentile rank of ≤6.0). P50 HLA-I binding predictions resulted in a total of 629 promising 9-/10-mer candidate peptides.* ***Abbr.:*** *AAC = amino acid-changing; G = genes; HLA = human leukocyte antigen; InDels = insertions and deletions; MNV = multiple nucleotide variations; NGS = Next Generation Sequencing; NSCLC = non-small cell lung cancer; nt = nucleotides; RNA-Seq = RNA Sequencing; RPKM = reads per kilobase of transcript per million reads mapped; SNV = single nucleotide variation; T = transcripts; WES = Whole Exome Sequencing*

3.4 TCR sequence homology as a strong indicator for shared target antigen specificity

Since neoantigen or TAA specificities could not be demonstrated for any of the analyzed P43 or P50 TCRs (see section 3.3), it was decided to apply additional antigen identification strategies for both patients. Those strategies were based on the comparison of P43- and P50-derived TIL sequencing data with TCR repertoires found in additional NSCLC patients and public TCR databases, thereby screening for highly related TCRs of known specificity. To date, HSD has analyzed TCR repertoires of more than 60 NSCLC patients and has access to various TCR databases, such as VDJdb and McPAS-TCR. HSD's bioinformaticians used public algorithms (e.g. GLIPH, GLIPH2, iSMART) to screen all available sequencing data for highly related TCR variable regions commonly found in different patients. To refine the analysis, a proprietary algorithm was applied to effectively group α and β TCR chains with highly related CDR3 amino acid sequences. Hence, HSD was able to reveal dozens of potential specificity clusters consisting of highly similar TRB- and TRA-CDR3 sequences found in multiple individuals (data not shown). Furthermore, all patients within a particular specificity cluster had at least one HLA allele in common. Therefore, further attempts to reveal TCR specificities mainly focused on applying those TCR clustering approaches rather than analyzing patient-specific TCR repertoires, thereby hypothesizing that highly related TCRs of a particular cluster might share similar target antigen specificities.

3.4.1 Using TCR homology screenings to identify a CMV-derived target antigen recognized by P50 TCR 36.7

In contrast to NSCLC patient P18 (see section 3.2), no target antigens were identified for any of the tested P43 and P50 TCRs, respectively (see section 3.3). Since neoantigen predictions and TAA panel screenings were not successful for both patients, multiple TCR clusters were additionally screened for highly related receptors with potentially shared antigen specificities. Here, HSD found that both P50 TCR 36.7-derived CDR3 sequences were part of a TCR specificity cluster consisting of 10 individual TRBV and 5 TRAV chains (Fig. 3.4.1 A), all of which had initially been obtained from VDJdb and McPAS-TCR databases.

Interestingly, all CDR3 sequences grouped in this specificity cluster were apparently related to an HLA-B*44:03-restriced response towards a CMV-derived target epitope. This data immediately suggested that P50 TCR 36.7 might also recognize the associated CMV-IE2 (immediate-early 2) 8-mer target peptide (NEGVKAAW). Hence, to also assess the proper processing and presentation of the epitope, a cDNA expression construct encoding the *CMV-IE2* ORF was generated for subsequent transfection of P50 HLA-expressing HEK 293T target cells. In fact, subsequent antigen testing approaches confirmed that P50 TCR 36.7-expressing CD8+ TCR-T cell populations specifically recognized the CMV-derived IE2 epitope in the context of HLA-B*44:03 (Fig. 3.4.1 B).

However, no highly related TCRs could be identified for any of the other potentially tumor-specific P43 or P50 TCRs, respectively. Hence, these results indicated that TCR database screenings might especially be suitable for identifying common virus-specific TCRs. Nevertheless, the conducted experiments provided evidence that the applied TCR clustering approach is suitable to identify highly related TCRs with potentially shared target antigen specificities.

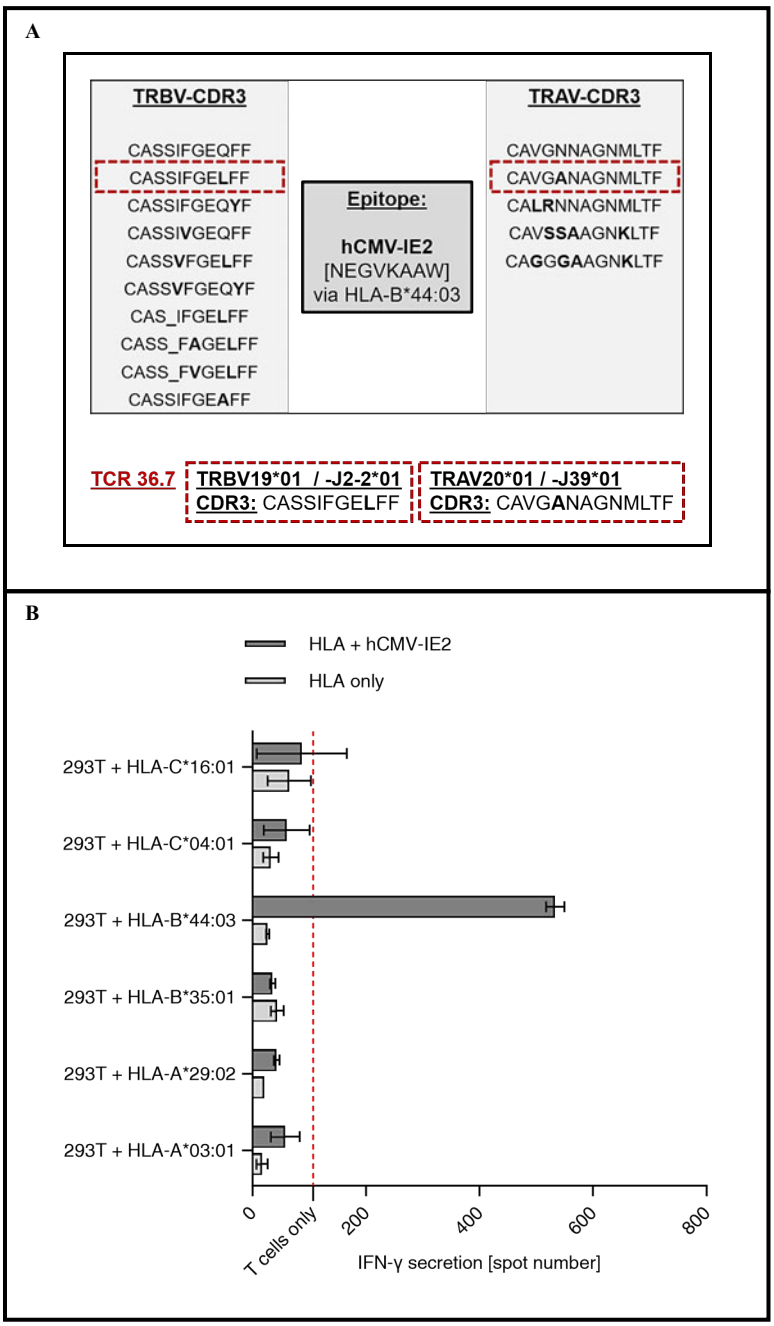


Figure 3.4.1: CD8+ TCR-T cells expressing P50 TCR 36.7 specifically recognize a CMV-derived IE2 target peptide via HLA-B*44:03.
A) *Specificity clustering of P50-derived TCR 36.7 and CDR3 sequences listed in open-source databases VDJdb and McPAS-TCR. P50 TCR 36.7 hTRBV- and hTRAV-CDR3 sequences were grouped with highly related CDR3 regions listed to specifically recognize an hCMV-IE2 8-mer target peptide (NEGVKAAW) in the context of HLA-B*44:03. A total of 10 highly related hTRBV- and 5 hTRAV-CDR3 sequences were identified. When using the topmost sequences as references, hTRBV-CDR3 sequences differed in a maximum of 3 amino acids while hTRAV-CDR3 sequences varied in not more than 4 AAs. Amino acid differences are shown in bold. Underlines indicate missing amino acids in shorter CDR3 sequences.* ***B)*** *hCMV-IE2 recognition testing of CD8+ TCR-T cells expressing P50 TCR 36.7. IFN-γ secretion was used as a marker for T cell recognition. Transgenic effector populations exclusively recognized hCMV-IE2 in the context of HLA-B*44:03+ HEK 293T cells. P50 HLA-transfected HEK 293T cells were not recognized in the absence of hCMV-IE2. Shown spot numbers represent mean values resulting from duplicates or single well testing results (T cells only preparation). Error bars represent standard deviations of duplicates. The IFN-γ background level of tested TCR-T cells was assessed in the absence of target cells and is shown in red. BC248-derived end.TCR-KO T cells (Sup. Fig. 14 A) were used for retroviral transduction with the P50 TCR 36.7 expression construct (Sup. Fig. 16). The ELISpot assay was performed 32 days after retroviral transduction. Effector cells per well: 3,000 TCR 36.7+. Target cells per well: 20,000. Raw ELISpot data is shown in Sup. Fig. 19.* ***Abbr.:*** *AAs = amino acids; BC = buffy coat; CMV = cytomegalovirus; end. = endogenous; HEK = Human embryonic kidney; HLA = human leukocyte antigen; KO = knock-out; hTRAV = human TCR alpha chain variable region; hTRBV = human TCR beta chain variable region; IE2 = immediate early 2; TCR = T cell receptor.*

3.4.2 TCR homology clustering as potent strategy for the identification of highly related TCRs with similar target recognition patterns

Besides screening open-source databases for homologous TCR sequences (see section 3.4.1), HSD patient-derived TIL sequencing data was also used to identify TCRs shared by multiple NSCLC patients. Specificity clustering of TIL sequencing data obtained from more than 60 different NSCLC patients identified 5 additional clonotypes with TCR sequences highly related to P43-derived TCR 75. Hence, the corresponding TCR cluster, referred to as cluster C, included 6 TCRs expressed by TIL clonotypes of 4 different NSCLC patients, namely P43 (TCR 75), P29 (TCRs 1115 and 1946), P55 (TCRs 18 and 1245), and P66 (TCR 10). Generally, all cluster C TRAV- and TRBV-CDR3 sequences did not differ in more than three amino acids, thereby sharing highly related antigen-binding motifs with potentially similar antigen specificity (Fig. 3.4.2 A). Furthermore, all 4 cluster C patients shared *HLA-B*08:01* and *HLA-C*07:01* alleles, respectively. It was therefore suspected that potentially concurring cluster C TCR antigen specificities might depend on one of the two HLA molecules.

While it was not possible to identify a candidate neoantigen mutation expressed (RPKM ≥0.5) in all cluster C tumors, the High Mobility Group Box 3 (HMGB3) protein was found to be overexpressed in in all 4 cluster C patients (data not shown). However, subsequent recognition testing revealed that the HMGB3 TAA was not recognized by CD8+ TCR-T cell populations expressing P43 (cluster C) TCR 75, thereby indicating that HMGB3 is not relevant for cluster C TCR-mediated antigen recognition (data not shown). To assess whether cluster C TCRs (C-TCRs) might be virus-specific, C-TCR 75+ TCR-T cell populations were additionally tested for the recognition of common viral epitopes, such as EBV-derived antigens presented by infected target cells or by applying a CEFT peptide mix (JPT Peptide Technologies GmbH; Berlin, Germany; data not shown) containing common HLA class I-restricted antigen peptides from CMV, EBV, influenza viruses, and tetanus toxin. Regardless of HLA-B*08:01 or HLA-C*07:01 expression (see Table 5.1), C-TCR 75+ effectors specifically recognized various EBV-infected B cell lines derived from melanoma patients D04-Mel, D05-Mel, D10-Mel, or D11-Mel via HLA class I (Fig. 3.4.2 B). In contrast, non-infected D10-Mel tumor cells and the HLA class I-deficient EBV-B cell line L721.221 (see Table 5.1) were not recognized at all. Additionally, HLA-B*08:01+ K562 cells were also recognized while co-incubation with HLA-C*07:01+ K562 cells and non-transduced K562 did not induce substantial IFN-γ secretion. In contrast to HLA-transduced K562, the recognition of EBV-B cells derived from patient D05-Mel (D05-Mel_EBV-B) was not substantially affected by HLA-B*08:01 or HLA-C*07:01 overexpression, thereby indicating that C-TCR 75 might recognize different antigen peptides on HLA-B*08:01+ K562 cells and EBV-infected B cells.

Since similar target antigen specificities were assumed for all 6 C-TCRs, transgenic CD8+ TCR-T cell populations were established, each of which expressed one of the identified cluster C TCRs (Sup. Fig. 21 and 22). Established effector populations were subsequently tested for the reactivity towards EBV-infected B cells and HLA-B*08:01+/HLA-C*07:01+ NSCLC cells. Hence, it was demonstrated that C-TCR+ effector populations recognized EBV-B cells derived from melanoma patient D05-Mel as well as the HLA-B*08:01+/C*07:01+ NSCLC cell line NCI-H1703. In contrast, the HLA-B*08:01/C*07:01-deficient NSCLC cell line NCI-H460 was not substantially recognized (Fig. 3.4.3 A). Generally, when compared to non-specific T cell activation with OKT-3, target cell-induced activation levels seemed rather low. Unfortunately, due to unsuccessful retroviral transductions of CD8+ donor lymphocytes, C-TCR 18 recognition testing could not be performed at this point of the study.

However, NCI-H1703 recognition was verified for all 6 C-TCRs (Fig. 3.4.3 B) by establishing new CD8+ TCR-T cell populations from buffy coat BC652 (Sup. Fig. 22). Consistent with the data previously shown for C-TCR 75 (Figure 3.4.2 B), all 6 cluster C TCR+ effectors specifically recognized HLA-B*08:01-transduced K562 cells but not non-modified K562 (Fig. 3.4.3 B). Detected INF-γ secretion upon target cell recognition was again rather low when compared to non-specific activation with OKT-3. Nevertheless, similar target recognition patterns were demonstrated for all C-TCRs, even when expressed in different healthy donor (BC777, BC652) lymphocytes. Hence, all tested effector populations recognized HLA-B*08:01+ K562, HLA-B*08:01+/C*07:01+ NCI-H1703, and, most likely

(C-TCR 18 not tested), D05-Mel_EBV-B cells (Fig. 3.4.3). Notably, target cell recognition efficiencies slightly differed among tested TCR-T cell populations, thereby indicating varying functional avidities of individual receptors. Taken together, these results suggested that cluster C TCRs might recognize a shared antigen present in K562 (CML) and NCI-H1703 (NSCLC) tumor cell lines cells via HLA-B*08:01. Furthermore, since HLA-B*08:01-negative EBV-B cells from patient D05-Mel were also recognized, C-TCRs might also induce cross-reactivity towards an alternatively presented EBV epitope.

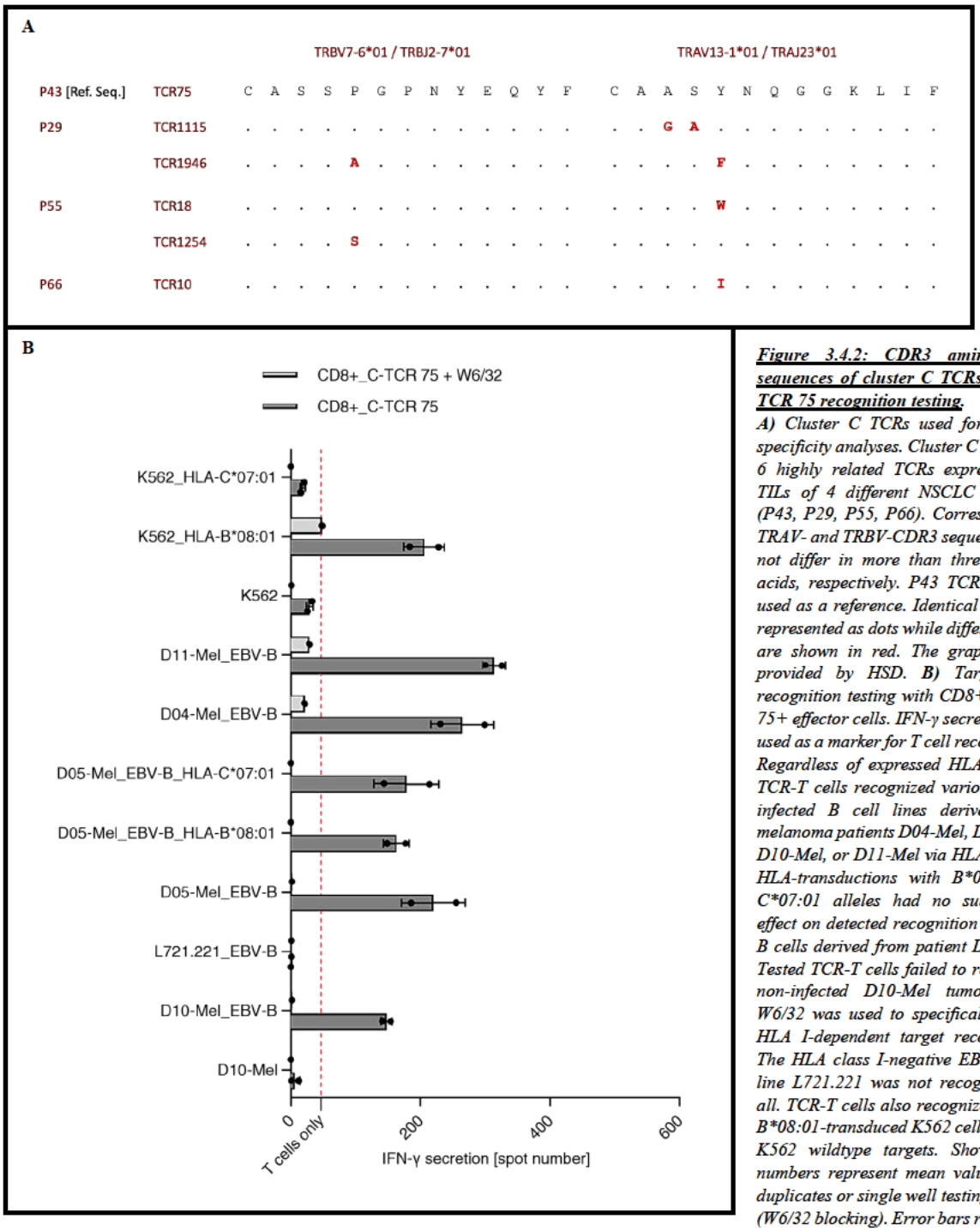


Figure 3.4.2: CDR3 amino acid sequences of cluster C TCRs and C-TCR 75 recognition testing.
A) *Cluster C TCRs used for further specificity analyses. Cluster C includes 6 highly related TCRs expressed by TILs of 4 different NSCLC patients (P43, P29, P55, P66). Corresponding TRAV- and TRBV-CDR3 sequences did not differ in more than three amino acids, respectively. P43 TCR 75 was used as a reference. Identical AAs are represented as dots while differing AAs are shown in red. The graphic was provided by HSD.* ***B)*** *Target cell recognition testing with CD8+/C-TCR 75+ effector cells. IFN-γ secretion was used as a marker for T cell recognition. Regardless of expressed HLA alleles, TCR-T cells recognized various EBV-infected B cell lines derived from melanoma patients D04-Mel, D05-Mel, D10-Mel, or D11-Mel via HLA class I. HLA-transductions with B*08:01 or C*07:01 alleles had no substantial effect on detected recognition of EBV-B cells derived from patient D05-Mel. Tested TCR-T cells failed to recognize non-infected D10-Mel tumor cells. W6/32 was used to specifically block HLA I-dependent target recognition. The HLA class I-negative EBV-B cell line L721.221 was not recognized at all. TCR-T cells also recognized HLA-B*08:01-transduced K562 cells but not K562 wildtype targets. Shown spot numbers represent mean values from duplicates or single well testing results (W6/32 blocking). Error bars represent standard deviations of duplicates. IFN-γ background levels of tested TCR-T cells were assessed in the absence of target cells and are shown in red (mean calculated from duplicate wells). The ELISpot assay was performed 78 days after retroviral transduction (Sup. Fig. 21) of BC777-derived end.TCR-KO T cells (Sup. Fig. 23) with the C-TCR 75 expression construct. Effector cells per well: 15,000 C-TCR 75+. Target cells per well: 50,000. W6/32 concentration per well: 50 µg/ml. Raw ELISpot data is shown in Sup. Fig. 20.* ***Abbr.:*** *AA = amino acid; CDR3 = complementary determining region 3; EBV = Epstein-Barr virus; end. = endogenous; HLA = human leukocyte antigen; KO = knock-out; Mel = melanoma; ref. = reference; Seq. = sequence; TCR = T cell receptor; TRAV = TCR alpha chain variable region; TIL = tumor-infiltrating lymphocyte; TRBV = TCR beta chain variable region.*

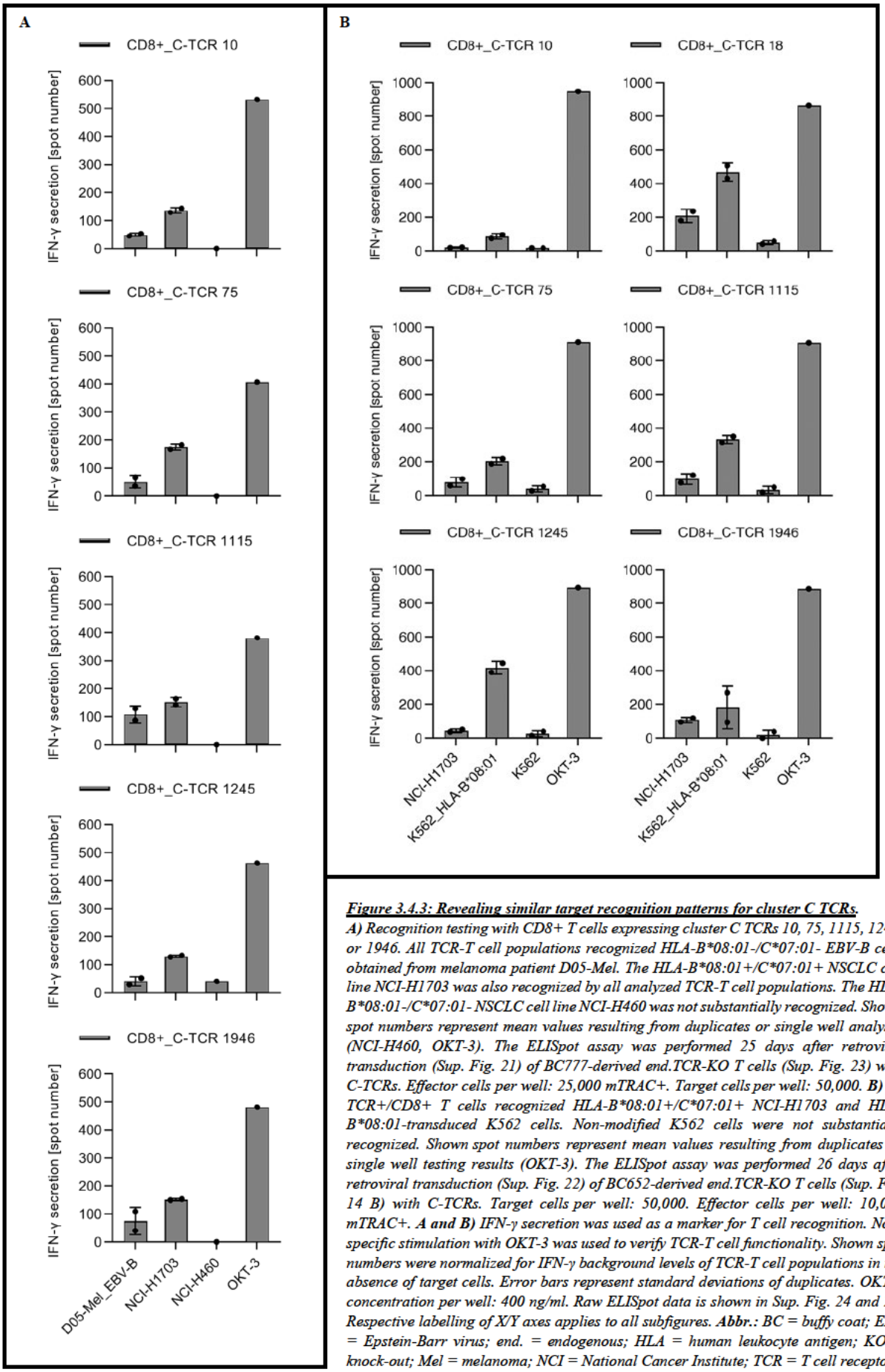


Figure 3.4.3: Revealing similar target recognition patterns for cluster C TCRs.
A) *Recognition testing with CD8+ T cells expressing cluster C TCRs 10, 75, 1115, 1245, or 1946. All TCR-T cell populations recognized HLA-B*08:01-/C*07:01- EBV-B cells obtained from melanoma patient D05-Mel. The HLA-B*08:01+/C*07:01+ NSCLC cell line NCI-H1703 was also recognized by all analyzed TCR-T cell populations. The HLA-B*08:01-/C*07:01- NSCLC cell line NCI-H460 was not substantially recognized. Shown spot numbers represent mean values resulting from duplicates or single well analyses (NCI-H460, OKT-3). The ELISpot assay was performed 25 days after retroviral transduction (Sup. Fig. 21) of BC777-derived end.TCR-KO T cells (Sup. Fig. 23) with C-TCRs. Effector cells per well: 25,000 mTRAC+. Target cells per well: 50,000.* ***B)*** *C-TCR+/CD8+ T cells recognized HLA-B*08:01+/C*07:01+ NCI-H1703 and HLA-B*08:01-transduced K562 cells. Non-modified K562 cells were not substantially recognized. Shown spot numbers represent mean values resulting from duplicates or single well testing results (OKT-3). The ELISpot assay was performed 26 days after retroviral transduction (Sup. Fig. 22) of BC652-derived end.TCR-KO T cells (Sup. Fig. 14 B) with C-TCRs. Target cells per well: 50,000. Effector cells per well: 10,000 mTRAC+.* ***A and B)*** *IFN-γ secretion was used as a marker for T cell recognition. Non-specific stimulation with OKT-3 was used to verify TCR-T cell functionality. Shown spot numbers were normalized for IFN-γ background levels of TCR-T cell populations in the absence of target cells. Error bars represent standard deviations of duplicates. OKT-3 concentration per well: 400 ng/ml. Raw ELISpot data is shown in Sup. Fig. 24 and 25. Respective labelling of X/Y axes applies to all subfigures.* ***Abbr.:*** *BC = buffy coat; EBV = Epstein-Barr virus; end. = endogenous; HLA = human leukocyte antigen; KO = knock-out; Mel = melanoma; NCI = National Cancer Institute; TCR = T cell receptor.*

3.4.3 Cluster C-derived TCRs mediate similar target recognition of transgenic TCR-T cell populations

Similar target recognition patterns were identified for transgenic effector populations expressing one of six highly related cluster C TCRs, respectively. However, when compared to non-specific stimulation with OKT-3, IFN-γ secretion of cluster C TCR-positive CD8+ T cells seemed rather low upon target cell recognition (see Fig. 3.4.3). To verify that the introduced C-TCRs, but not residual endogenous TCR expression, accounted for the observed T cell activation, a CRISPR/Cas9-based knock-out approach was used to specifically target the murine constant regions exclusively present in introduced C-TCRs. Hence, transgenic TCR expression was specifically knocked out in C-TCR 1115+ effectors, thereby serving as a representative for all other cluster C TCR-T cell populations (see section 3.4.2).

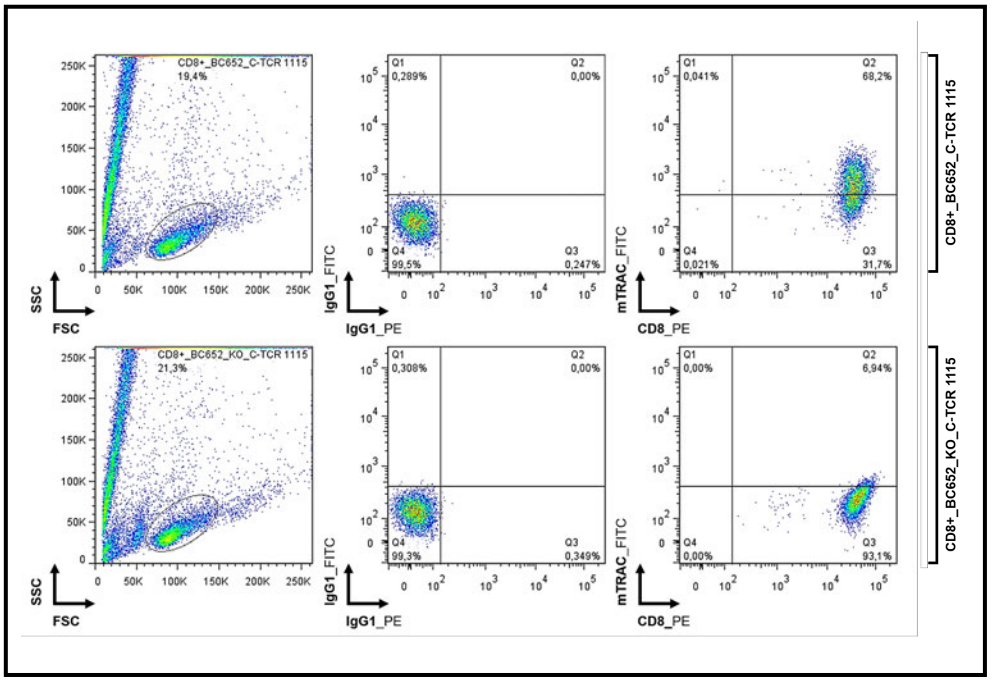


Figure 3.4.4: CRISPR/Cas9-based KO of C-TCR 1115 expression in transgenic CD8+ effector populations.
Transgenic CD8+ effector populations (BC652; Sup. Fig. 22) were used for a CRISPR/Cas9-based KO of C-TCR 1115 by simultaneously targeting murinized αβ TCR constant regions. C-TCR 1115 expression was determined via mTRAC staining. Non-treated TCR-T cell populations were used as positive control for mTRAC staining. C-TCR 1115 expression was reduced to less than 7% in CRISPR/Cas9-treated TCR-T cell populations. Flow cytometry was performed 6 days after CRISPR/Cas9 treatment. Flow cytometry data was gated based on viable T cell populations and non-specific isotype (IgG1) controls. ***Abbr.:*** *BC = buffy coat; FSC = forward scatter; mTRAC = murinized T cell receptor alpha chain constant region; KO = knock-out; PE = Phycoerythrin; SSC =sideward scatter; TCR = T cell receptor.*

By using crRNAs exclusively targeting the murinized cluster C TCR constant regions (Sup. Fig. 26), C-TCR 1115 surface expression was reduced to less than 7% in CRISPR/Cas9-treated effector populations (Fig. 3.4.4). Target recognition patterns of resulting CD8+_KO_C-TCR 1115 (TCR 1115-KO) effector populations and non-treated C-TCR 1115+ effectors were compared to assess the role of the cluster TCR during T cell activation. The conducted KO approach resulted in a substantial decrease of T cell activation detected upon co-incubation of TCR 1115-KO effectors with NCI-H1703 and EBV-infected B cells obtained from melanoma patients D10-Mel or D11-Mel, respectively (Fig. 3.4.5 A and B). In fact, IFN-γ secretion of TCR 1115-KO effectors were decreased by approx. 50% when compared to non-treated control TCR-T cells. The observed effect was even more evident when analyzing HLA-B*08:01-transduced K562 target cell recognition where IFN-γ secretion of TCR 1115-KO T cells was reduced by up to 90% (Fig. 3.4.5 B). Consistent with previous results (see section 3.4.2), both non-transduced and HLA-C*07:01+ K562 targets were not recognized at all. These findings again indicated an HLA-B*08:01-restriced recognition of K562 cells via C-TCR 1115. Furthermore, OKT-3-mediated

T cell activation was found to be reduced by up to 70% in TCR 1115-KO T lymphocytes, thereby indicating reduced overall activation capacities of effector populations after CRISPR/Cas9 treatment.

By demonstrating reduced target cell recognition capacities of TCR 1115-KO effector populations, it was clearly demonstrated that C-TCR 1115 was responsible for mediating the recognition of NSCLC target cell line NCI-H1703, two EBV-infected B cell lines obtained from melanoma patients D10-Mel or D11-Mel, and HLA-B*08:01+ K562 cells. Furthermore, residual TCR 1115 expression levels of less than 7% seemed sufficient to induce up to 50% of T cell activation previously detected for non-treated effector cultures (approx. 90% mTRAC+), thereby highlighting the high functional avidity of cluster C TCR 1115-transduced T cells. Taken together, similar target recognition patterns were revealed for all C-TCRs (see section 3.4.2) and, in addition, it was verified that the observed target cell recognition was actually mediated by the introduced cluster C TCR 1115. Hence, although not being able to identify a definitive target antigen recognized by all C-TCRs, the conducted TCR clustering approach was demonstrated to be suitable to identify highly related TCRs with similar antigen recognition patterns.

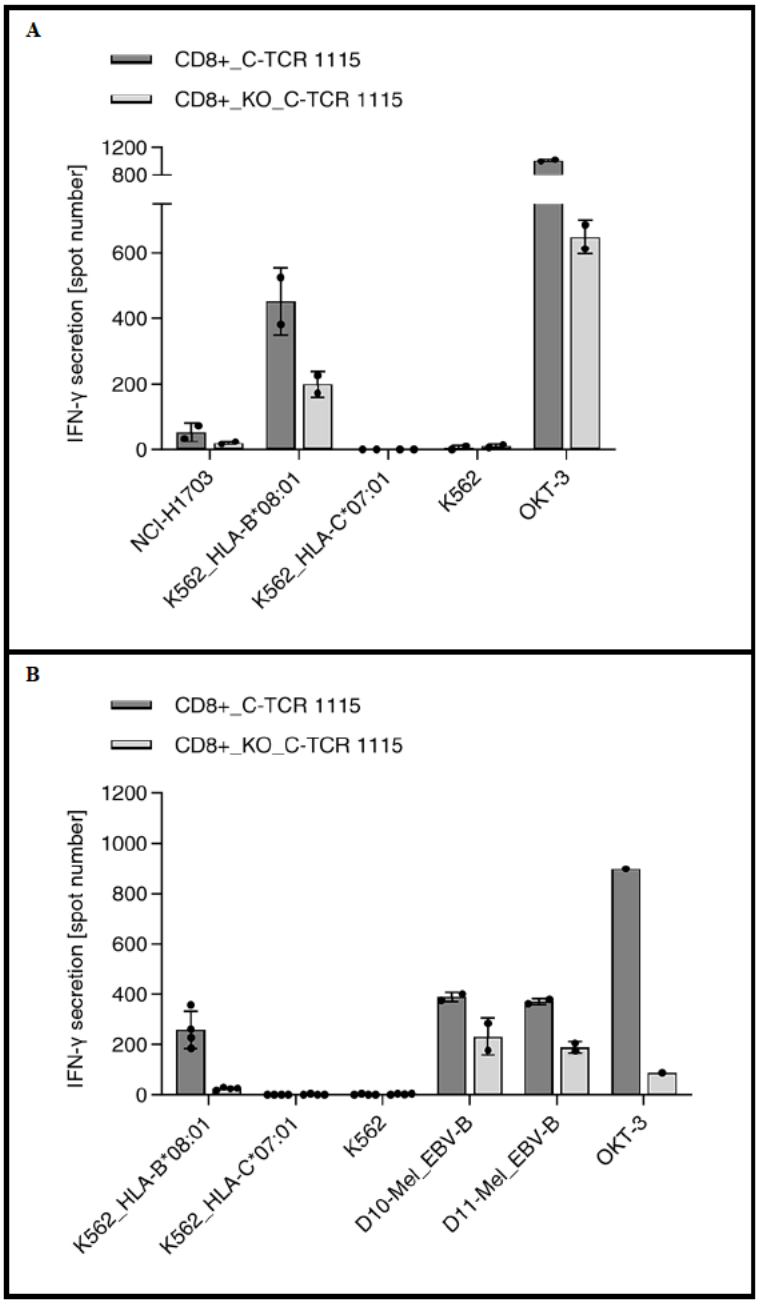


Figure 3.4.5: Recognition testing of cluster C TCR 1115-KO T cells.

A and B) *Target cell recognition of CD8+ TCR-T cells after CRISPR/Cas9-based KO of transduced C-TCR 1115. IFN-γ secretion was used as a marker for T cell recognition. Non-specific stimulation with OKT-3 was used to verify T cell functionality. Shown spot numbers were normalized for IFN-γ background levels of effectors in the absence of target cells.* ***A)*** *NCI-H1703 and K562_HLA-B*08:01 target cell recognition testing. When compared to non-treated C-TCR 1115+ T cells, NCI-H1703 and K562_HLA-B*08:01 recognition was reduced by approx. 50% after CRISPR/Cas9 treatment. T cell activation via OKT-3 was reduced by almost 40%. Shown spot numbers represent mean values resulting from duplicate test wells. Error bars represent standard deviations of duplicates. The ELISpot assay was performed 24 days after retroviral transduction (Sup. Fig. 22) of BC652-derived end.TCR-KO T cells (Sup. Fig. 14 B) and 7 days after performing the KO of C-TCR 1115 (Fig. 3.4.4). Effector cells per well: Either 10,000 mTRAC+ or 10,000 TCR 1115-KO T cells. Target cells per well: 50,000. OKT-3 concentration per well: 400 ng/ml. Raw ELISpot data is shown in Sup. Fig. 27.* ***B)*** *EBV-B cell recognition testing. When compared to non-treated C-TCR 1115+ T cells, recognition of EBV-infected B cells derived from melanoma patients D10-Mel and D11-Mel was reduced by almost 50% after CRISPR/Cas9 treatment. T cell activation via OKT-3 was reduced by approx. 90%. Shown spot numbers represent mean values either resulting from quadruplicates (K562), duplicates (EBV-B cell), or single well (OKT-3) analysis. Error bars represent standard deviations of duplicates or quadruplicates. The ELISpot assay was performed 28 days after retroviral transduction (Sup. Fig. 22) of BC652-derived end.TCR-KO T cells (Sup. Fig. 14 B) and 11 days after performing the KO of C-TCR 1115 (Fig. 3.4.4). Effector cells per well: Either 10,000 mTRAC+ or 10,000 TCR 1115-KO T cells. Target cells per well: 50,000. OKT-3 concentration per well: 400 ng/ml. Raw ELISpot data is shown in Sup. Fig. 28.* ***Abbr.:*** *BC = buffy coat; EBV = Epstein-Barr virus; end. = endogenous; HLA = human leukocyte antigen; KO = knock-out; Mel = melanoma; mTRAC = murinized T cell receptor alpha chain constant region; NCI = National Cancer Institute; TCR = T cell receptor.*

3.4.4 Ongoing projects: Recognition testing of cluster G-derived TCRs

The power of the HSD-proprietary prediction pipeline for the efficient identification of tumor-specific TIL clonotypes was previously demonstrated in the context of NSCLC patient P18 (see section 3.2). However, when analyzing P43 and P50, it was also found that the efficiency of the pipeline might vary between individual patients (see section 3.3). Hence, the TCR clustering approach represented a promising strategy to complement the HSD prediction pipeline by providing the chance to identify potentially tumor-specific TCRs based on highly related CDR3 sequences and shared target antigen specificities. In addition to TCR cluster C (see sections 3.4.2 and 3.4.3), HSD identified another promising TCR cluster, referred to as cluster G, consisting of 6 highly related TCRs obtained from 3 different NSCLC patients (P44, P70, P82; see Figure 3.4.6). Independent of the clustering approach performed by HSD, CDR3 regions expressed by cluster G TCRs 1702 and 5094 had previously been identified by Mark M. Davis and colleagues (Stanford University, USA) when using their GLIPH2 (Grouping of Lymphocyte Interactions with Paratope Hotspots version 2) algorithm for the identification of TCRs with shared antigen specificity. More precisely, identical TRAV/TRBV-CDR3 sequences were found in Stanford TCR 18 and G-TCR 5094, as well as in Stanford TCR 19 and G-TCR 1702, respectively. Interestingly both Stanford TCRs were demonstrated to specifically recognize a TMEM161A-derived TAA epitope as well as two pathogen-derived 9-mer peptides via HLA-A*02:01 [220], thereby indicating similar antigen specificities not only for TCRs 1702 and 5094 but for all TCRs included in cluster G.

Patient	TIL clonotype	TRBV-CDR3	TRAV-CDR3
P44 [ref.]	5094	CASSEDGMNTEAFF	CAVLMDSNYQLIW
P44	697	CASS**P**DGMNTEAFF	CA**L**LMDSNYQLIW
P70	1702	CASS**S**DGMNTEAFF	CAVLMDSNYQLIW
P82	159	CASS**N**DGMNTEAFF	CAVLMDSNYQLIW
P82	159_v1	CASS**T**DGMNTEAFF	[CAVLMDSNYQLIW]
P82	159_v2	CASS**D**DGMNTEAFF	[CAVLMDSNYQLIW]

Figure 3.4.6: Cluster G TCRs used for further specificity analyses.
Cluster G comprises 6 highly related TCRs expressed by TIL clonotypes obtained from three different NSCLC patients (P44, P70, P82). Corresponding TRAV- and TRBV-CDR3 sequences did not differ in more than one amino acid, respectively. When using TCR 5094 as a reference, differing AAs are shown in red. Due to non-available TCR 159_v1 and TCR 159_v2 TRAV sequencing data, corresponding TRBV-CDR3 sequences were artificially paired with the most common cluster G TRAV-CDR3 region found in G-TCRs 5094, 1702, and 159. TRAV-CDR3 sequences used for artificial pairing are shown in squared brackets. Graphic was provided by HS Diagnomics. ***Abbr.:*** *AA = amino acid; CDR3 = complementary determining region 3; ref. = reference; TCR = T cell receptor; TRAV = TCR alpha chain variable region; TIL = tumor-infiltrating lymphocyte; TRBV = TCR beta chain variable region.*

Hence, transgenic TCR-T cell populations were generated, each of which expressed one of the 6 cluster G TCRs, respectively (Sup. Fig. 29). By analyzing the target recognition patterns of established effector populations, similar target cell specificities were verified for all tested cluster G TCR-transduced effectors. All analyzed TCR-T cell populations recognized three HLA-A*02:01+ NSCLC cell lines (MZ-LC-16, NCI-H1703, NCI-H1792; see Table 5.1 for respective HLA-I genotypes), HLA-A*02:01-expressing K562 cells, and TAP-deficient T2 EBV-B cells. In contrast, HLA-A*02:01+ NSCLC cell lines MOR/CPR and NCI-H661 were not recognized (Fig. 3.4.7). Furthermore, pathogen-derived antigen specificities were verified for 5 of 6 cluster G TCR-expressing T cell cultures (Fig. 3.4.8). Consistent with data previously published for TRAV/TRBV-CDR3 regions of G-TCRs 5094 and 1702 [220], established effector populations specifically recognized HLA-A*02:01+ MOR/CPR lung cancer cells when pulsed with immunogenic 9-mer peptides either originating from the non-mutated TAA

TMEM161A (Transmembrane Protein 161A), EBV-derived LMP2A (latent membrane protein 2A), or the *E. coli*-derived enterobactin exporter EntS, respectively. Additionally, the observed recognition of the HLA-A*02:01+ lung cancer cell line MZ-LC-16 was specifically blocked via the HLA-A*02-specific antibody MA2.1, thereby confirming an HLA-A*02:01-restricted target cell recognition (Fig. 3.4.8). Unfortunately, due to impaired proliferation capacities after retroviral transduction, G-TCR 1702-expressing effectors could not be used for the peptide recognition testing shown in Figure 3.4.8.

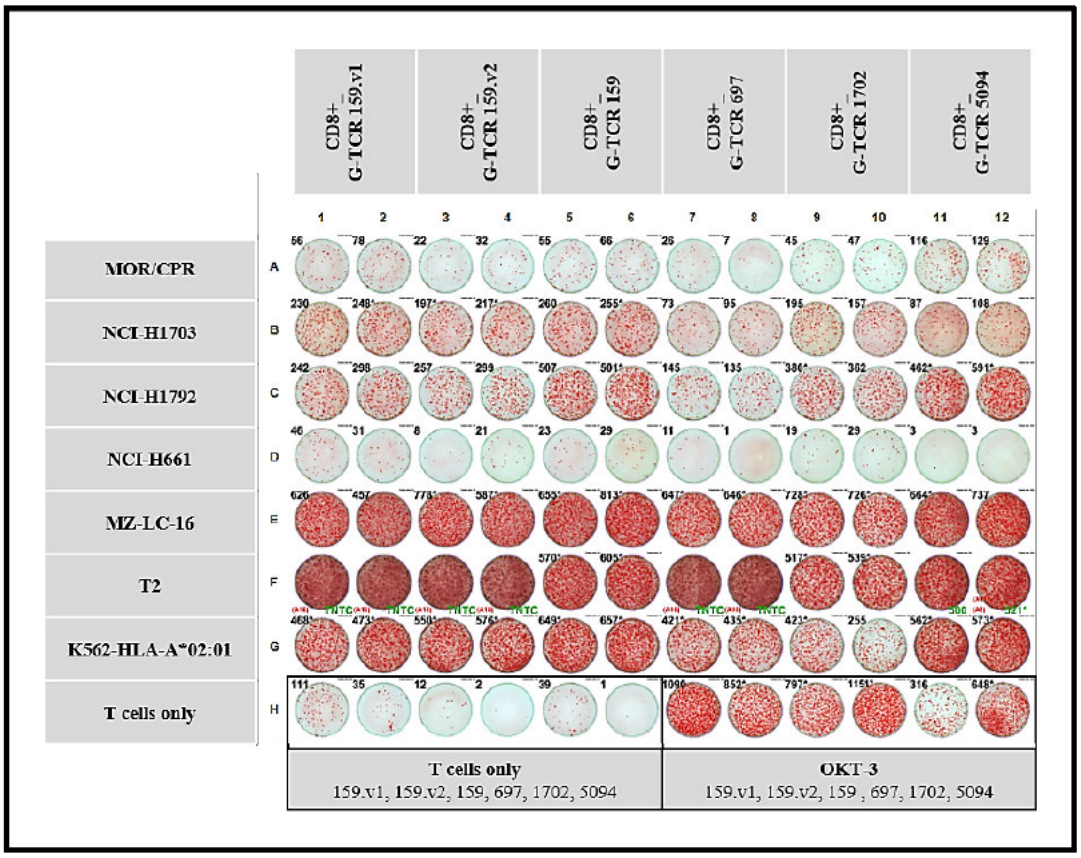


Figure 3.4.7: Revealing similar target recognition patterns for cluster G TCRs.
*Recognition testing with CD8+ TCR-T cell populations, each of which expressed one of the previously identified cluster G TCRs. All tested effector populations recognized the three HLA-A*02:01+ lung cancer cell lines NCI-H1703, NCI-H1792, and MZ-LC-16. HLA-A*02:01+ K562 cells and the TAP-deficient T2 cell line were also recognized by all tested effector populations. The ELISpot assay was performed 29 days after retroviral transduction (Sup. Fig. 29) of BC652-derived end.TCR-KO T cells (Sup. Fig. 14 B) with cluster G TCR expression constructs. IFN-γ secretion was used as a marker for T cell recognition. IFN-γ background levels were assessed by analyzing effector populations in the absence of target cells (T cells only preparations). Non-specific stimulation with OKT-3 was used to verify the general functionality of established effectors. Cluster G TCR+ effector numbers per well were calculated based on mTRAC staining performed 1-2 days before ELISpot analysis. Retrospective corrections of inaccurately counted wells are highlighted in green font. Effector cells per well: 10,000 mTRAC+. Target cells per well: 50,000. OKT-3 concentration per well: 400 ng/ml. **Abbr.:** BC = buffy coat; end. = endogenous; HLA = human leukocyte antigen; KO = knock-out; LC = lung cancer; mTRAC = murine TCR alpha chain constant region; NCI = National Cancer Institute; TAP = transporter associated with antigen processing; TCR = T cell receptor; TNTC= too numerous to count.*

The recognition of various HLA-A*02:01+ NSCLC cell lines and HLA-expressing K562 indicated a broad expression of the non-mutated TAA TMEM161A across multiple cancers. Moreover, the specific recognition of TAP-deficient T2 cells, a T/B-hybridoma with EBV-transformed B cells, suggested that the recognized EBV-peptide might be presented independently form TAP peptide transporters. Together with revealing similar target antigen recognition patterns for cluster C TCR-expressing effectors (see sections 3.4.2 and 3.4.3), these results again demonstrated the power of the performed TCR clustering approach using sequence homologies for subsequent target antigen identification. Furthermore, the performed G-TCR recognition analyses clearly demonstrated that the applied clustering approach can not only be used to successfully identify rather common virus-specific TCRs, but it might also be

suitable to identify highly related TCRs specifically recognizing shared TAAs. Currently ongoing projects are now focusing on the identification of even more TCR clusters and their respective target antigens. While exceeding the scope of this dissertation, one of the currently most relevant objectives is to reveal the exact target antigens recognized by cluster C TCR-transduced effector populations (see sections 3.4.2 and 3.4.3).

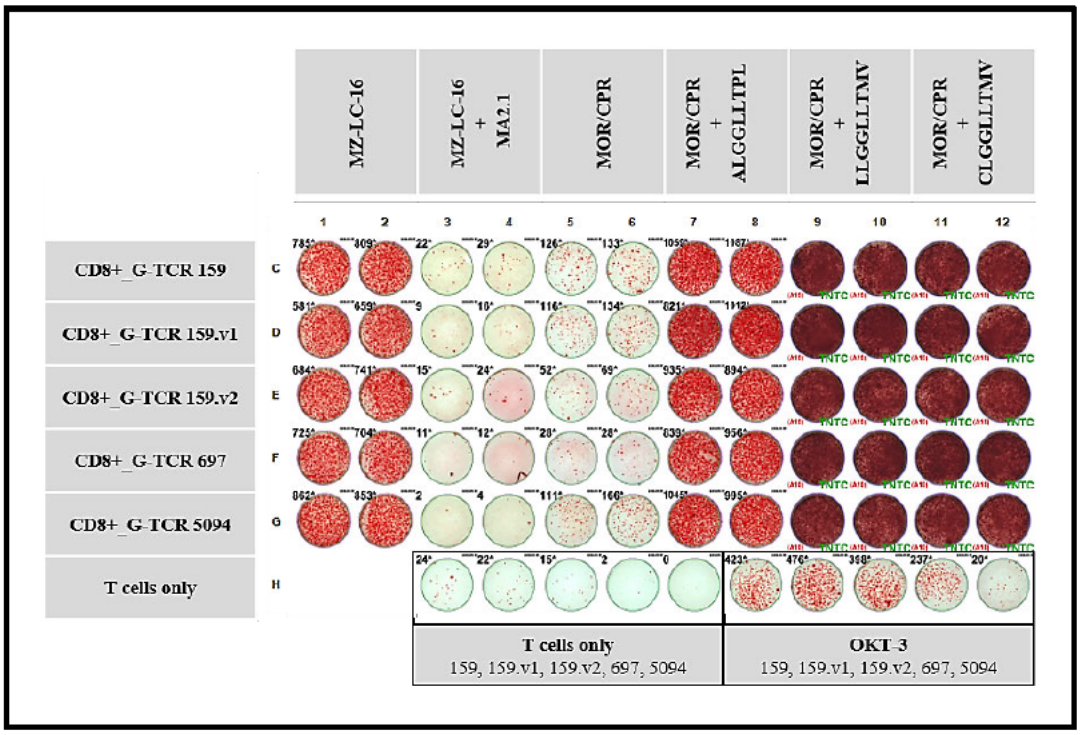


Figure 3.4.8: G-TCRs recognize two pathogen-derived target epitopes and a TMEM161A-derived 9-mer peptide via HLA-A*02:01.
*Recognition testing with CD8+ TCR-T cells either expressing cluster G TCRs 159, 159.v1, 159.v2, 697, or 5094. All tested effector populations recognized the MZ-LC-16 lung cancer cell line via HLA-A*02:01. HLA-A*02:01+ MOR/CPR lung cancer cells were only recognized when pulsed with one of three immunogenic 9-mer peptides: ALGGLLTPL (TMEM161A), LLGGLLTMV (EntS), or CLGGLLTMV (LMP2A). The ELISpot assay was performed 34 days after retroviral transduction (Sup. Fig. 29) of BC652-derived end.TCR-KO T cells (Sup. Fig. 14 B) with cluster G TCR expression constructs. IFN-γ secretion was used as a marker for T cell recognition. IFN-γ background levels were assessed by analyzing effector populations in the absence of target cells (T cells only preparations). The HLA-A*02-specific antibody MA2.1 was used to block HLA-A*02-dependent target recognition. Non-specific stimulation with OKT-3 was used to verify the general functionality of established effectors. Non-used wells and wells containing non-relevant antigen testing information were excluded for final evaluation. G-TCR+ effector numbers per well were calculated based on mTRAC staining performed 1-2 days before ELISpot analysis. Retrospective corrections of inaccurately counted wells are highlighted in green font. Effector cells per well: 10,000 mTRAC+. Target cells per well: 50,000. OKT-3 concentration per well: 400 ng/ml. MA2.1 concentration per well: 400 µg/ml.* ***Abbr.:*** *BC = buffy coat; end. = endogenous; HLA = human leukocyte antigen; KO = knock-out; LC = lung cancer; mTRAC = murine TCR alpha chain constant region; TCR = T cell receptor; TNTC = too numerous to count.*

4) Discussion

The conducted proof-of-concept study successfully demonstrated the power of a novel *in silico* pipeline to predict tumor-specific T cell clonotypes among patient-derived tumor-infiltrating lymphocyte (TIL) populations. T cell receptor (TCR) beta chain profiling was used to identify potentially tumor-reactive TIL clonotypes in three stage III lung cancer patients (P18, P43, P50) previously subjected to surgery with curative intent. Tumor-specific non-synonymous mutations were identified for each patient by using comparative WES of respective lung and tumor tissues. Furthermore, mRNA expression levels of variant alleles and aberrantly expressed genes were assessed by tumor transcriptome sequencing. The conducted mutation screenings were followed by MHC class I binding predictions of mutated peptides to eventually identify the most promising neoantigen candidates for each patient, respectively. Three P18 TIL-derived TCRs were finally demonstrated to specifically recognize autologous tumor tissue via an HLA-A*01:01-presented $KRAS^{Q61H}$ neoantigen shared by various patients and cancers, thereby providing three promising candidate TCRs for future T cell-based immunotherapies.

Motivated by the observation that all three P18 TCRs exhibited highly similar αβ TCR amino acid sequences, an additional *in silico* method was applied to facilitate the clustering of sequence-related TCRs with potentially shared antigen specificities. By applying this TCR clustering approach, tumor cell line reactivities were demonstrated for two different TCR clusters containing a total of 12 TCRs from 7 different patients. Here, similar antigen specificities towards a TMEM161-derived TAA epitope were successfully verified for at least 5 of the 6 highly related cluster G TCRs only differing in 1-2 amino acids within their respective α- and β-CDR3 regions. Taken together, this study successfully demonstrated the power of two independent *in silico* prediction pipelines aiming at the efficient identification of potentially tumor-specific TCRs, which was either achieved by assessing personalized TIL clonotype frequencies or via clustering of common TCRs with potentially shared tumor antigen specificities. The following paragraphs will critically assess the efficiency of all experiments conducted in the course of this study while also discussing the overall relevance of the achieved results.

4.1 Direct identification of potentially tumor-specific T cell clonotypes from patient-derived CD8+ TIL populations

A novel *in silico* prediction pipeline was applied to identify potentially tumor-specific TCRs obtained from TIL populations of the three stage III NSCLC patients P18, P43, and P50. Here, transgenic effector populations were established by cloning TCRs from tumor-enriched T cell clonotypes into bicistronic vector backbones suitable for the subsequent retroviral transduction of healthy donor T lymphocytes [225]. Tumor-specific candidate clonotypes were selected based on overall frequencies in a patient's tumor tissue, high tumor-to-non-tumor (T/nT) ratios, and the relative occurrence in PD-1+ TIL subsets. Notably, patient-derived TIL clonotype frequencies were assessed *ex vivo* and without any *in vitro* culturing of T cells, thereby circumventing altered TIL compositions and non-representative frequencies of tumor-reactive clonotypes [215].

Despite of blood being considered a suitable source for the identification of neoantigen-specific T cells [226–228], TIL populations were preferred over patient-derived PBMCs (peripheral blood mononuclear cells) for analyzing TCR repertoire frequencies during this research project. By using PD-1 expression as a suitable marker for the efficient identification of clonally expanded anti-tumor T cells, other studies demonstrated tumor-reactive T cell clonotypes to be highly enriched in CD8+ TIL populations of melanoma patients [208,229]. Furthermore, when comparing PD-1 expression levels between tumor-resident T cells and circulating peripheral blood lymphocytes (PBLs) of melanoma patients, CD8+ TIL populations included significantly higher overall levels of PD-1+ T cells than corresponding CD8+ PBLs (approx. 36% in TILs vs. approx. 4.1% in PBLs) [230]. Moreover, levels of tumor mutation-reactive T cells are presumably higher in TIL populations than in circulating PBLs, with the latter being supposed

to include neoantigen-specific T cells at frequencies estimated to be as low as 0.002% to 0.4% [227,231]. Due to the oligoclonal expansion after antigen contact within the tumor microenvironment (TME), tumor-specific T cell frequencies are assumed to be substantially higher among TIL populations, which might include up to 10% of intra-tumoral CD8+ T cells that are capable of recognizing autologous tumors. Pasetto and colleagues recently found that many of the most frequent PD-1+/CD8+ TIL clonotypes from metastatic melanoma patients recognized autologous tumor cells and, more precisely, non-mutated TAAs or mutated neoepitopes [215,232]. They thereby demonstrated the potential of identifying tumor-reactive T cells based on relative frequencies in patient-derived TIL populations, even without having prior knowledge about the exact antigen specificities. Hence, time-consuming antigen identification experiments could, in theory, be circumvented for future T cell-based cancer therapies. However, this strategy might not be suitable for every other tumor type and may particularly be challenging in unresectable cancers, poorly immunogenic malignancies, and for tumors in which T cell infiltrates are scant [187]. Moreover, ratios of tumor-reactive T cells might be highly diverse among TIL populations obtained from different cancer patients and tumor specificity might be restricted, if present at all, to only a minority of TILs [232].

Reducing the error rate while simultaneously improving the accuracy of current *in silico* pipelines is absolutely necessary to predict tumor-specific T lymphocytes more efficiently in the future. One major bottleneck of using patient-derived TILs for this purpose is that individual TIL populations were shown to include high numbers of CD39-negative bystander T cells specifically recognizing target antigens unrelated to cancer [206]. Generally, CD8+ TIL populations co-expressing CD39 (Ectonucleoside triphosphate diphospho-hydrolase-1; ENTPD1) and CD103 (Integrin alpha E; ITGAE) were demonstrated to be enriched for tumor-reactive T cells in both primary and metastatic tumors. Although being partially expressed by some regulatory T cells, CD39 and CD103 co-expression was primarily found on exhausted tumor-resident memory T cells present within the TME. While not being enriched for EBV- or CMV-specific T cells, those TIL populations were shown to specifically recognize and kill autologous tumor cells of melanoma patients in an MHC class I-dependent manner [207]. In addition to CD39 and CD103 co-expression, immune checkpoints LAG-3, TIM-3, and 4-1BB were also shown to be expressed or co-expressed by tumor-reactive CD8+ TIL populations [208]. However, while surface marker analyses might be useful to provide subsidiary information referring to the potential tumor specificity of particular T cell clonotypes, those assessments alone are probably not suitable to reliably identify tumor-specific TCRs for future immunotherapies. Nevertheless, simultaneously assessing T cell surface marker expression levels might represent a promising strategy to further improve the efficacy of the potentially tumor-specific TIL prediction pipeline applied in this study.

Due to their exceptional role in cancer eradication, this study exclusively focused on analyzing tumor specificities of TCRs obtained from CD8+ TIL clonotypes. CD8+ cytotoxic T lymphocytes (CTLs) are considered major killers of neoplastic cells, which is due to their ability to directly recognize tumor cells via MHC class I-presented peptides and immediately induce target cell killing [58]. However, CD4+ T cells also play crucial roles in the anti-tumor immunity since they are involved in CD8+ CTL recruiting, pro-inflammatory cytokine secretion, and the maintenance of CD8+ immune responses [57,62]. Furthermore, various studies demonstrated that CD4+ T cells are also able to recognize tumor-specific-antigens, such as neoantigens derived from mutated ERBB2IP (erbb2 interacting protein) or LEMD2 (LEM domain nuclear envelope protein 2) [233–235]. Similar results were recently found for tumor-resident FOXP3+/CD4+ regulatory T cell (Treg) clonotypes that were shown to recognize cancer neoantigens including mutated CCL-5 (CC chemokine ligand 5, also known as RANTES) [236]. However, CD4+ lymphocytes are restricted by MHC class II molecules that are mainly expressed by professional antigen presenting cells (APCs). Although there is evidence that tumor cells are able to express HLA class II, neoantigen recognition of MHC class II-presented neoantigens might be largely mediated by tumor-resident professional APCs [6]. Hence, CD4+ T cell clonotypes were excluded from tumor specificity assessments performed during this study, thereby exclusively focusing on the identification of MHC class I-restricted CD8+ TILs with the potential to directly recognize and kill tumor cells.

4.2 Comparative WES and RNA-Seq as a potent strategy for the efficient identification of tumor-specific mutations in NSCLC patients

By combining a potent next-generation sequencing (NGS) approach with efficient HLA binding predictions, this study identified a KRAS-derived neoantigen peptide recognized by TCRs cloned from TIL populations of NSCLC patient P18. As previously described by the Wölfel group [226], comparative whole exome (WES) and transcriptome sequencing (RNA-Seq) were used to identify somatic nucleotide substitutions (SNSs, MNVs) and genetic insertions/deletions (InDels) exclusively expressed in the patient's tumor but not in corresponding normal lung tissue samples. Next, mutated 19-mer peptides, each of which contained one of the previously identified mutations, were used for *in silico* binding predictions to identify potential neoantigen candidate peptides. Most promising neopeptide candidates, referring to those peptides with highest binding affinities to the patient's HLA class I alleles, were eventually synthesized for extensive *in vitro* neoantigen screenings.

Generally, various sequencing strategies can be used for the prediction of neoantigen candidates, thereby complicating the interpretation of neoantigen loads reported in different studies [237]. The current study integrated both DNA and RNA sequencing data to assess variant allele read frequencies and corresponding gene expression levels for every single mutation. Furthermore, as recently shown by Barnell and colleagues, manual revision and curation of most promising candidate mutations was crucial for detecting artifacts, sequencing errors, and other findings potentially resulting in false-positive neoantigen prediction [238]. Combining comparative WES and RNA-Seq to (i) detect tumor-specific non-synonymous mutations and (ii) assess corresponding variant allele expression is a common approach to identify putative neoantigen candidates [230,239,240]. Exclusively relying on transcriptome data for the identification of expressed non-synonymous tumor mutations represents an alternative and seemingly more straightforward approach that might generally result in higher numbers of identified somatic mutations [226]. However, transcriptome-based variant detection can be challenging, especially due to dynamic gene expression, RNA-editing, or splicing processes [241]. Hence, variant calling from RNA-Seq data greatly depends on the transcript expression status in the reference sample and was shown to regularly result in high false-positive as well as high false-negative error rates [240]. This approach was therefore considered less suitable for the identification of potential neoantigens during this study. Moreover, other studies demonstrated that a majority (approx. 86%) of somatic variants were missed when using tumor RNA-Seq and matched normal tissue-derived WES data for the variant identification. Hashimoto *et al.* considered low variant transcript expression as the fundamental factor for missing various mutations on an RNA level [240].

Hence, the current study followed the strategy of exclusively using transcriptome sequencing data to assess expression levels of mutations previously detected via comparative WES, thereby aligning with data demonstrating that transcriptome sequencing represents a crucial factor to efficiently identify putative neoantigen candidates. Since tumor transcriptome sequencing facilitates the direct identification of reads referring to mutated RNA transcript expression, RNA-Seq was considered mandatory to efficiently narrow down the list of somatic mutations detected via WES [237]. For this study, variant expression levels were initially assessed and then confirmed via manual curation to individually verify mutated RNA transcripts. Generally, transcript levels can be quantified by using the reads per kilobase per million mapped reads (RPKM) measure of read density, thereby reflecting the frequency of a given transcript in the starting sample by normalizing for both RNA length and total read numbers. Therefore, normalized RPKM values can facilitate the transparent comparison of transcript levels within as well as between samples [242]. Here, RPKM values between 0.3 and 1.0 are commonly used for defining individual gene expression cutoffs [243]. Therefore, when also considering neoantigen prediction studies previously conducted by the Wölfel group [226], a cutoff value of RPKM $\geq$0.5 was used for assessing variant allele expression levels in respective patient-derived tumor tissues.

Screening for tumor mutations with a rather high variant allele frequency (VAF), which is defined as the percentage of detected sequencing reads that match a specific variant divided by the overall coverage

at the respective locus [244], is another common approach used for the efficient prediction of neoantigen candidates. However, defining universal VAF recommendations for the efficient identification of somatic variants is not trivial as they are highly dependent on sequencing depth, tumor purity, genetic heterogeneity, and the fact that not all genes can be screened properly during sequencing [245]. Hence, excluding somatic variations based on low VAF values might bear the risk of missing promising tumor mutations potentially resulting in neoantigen occurrence. In contrast, the conducted study identified most promising tumor mutations and corresponding transcript expression based on actual/normalized read counts (WES: ≥4 variant reads, RNA-Seq: RPKM ≥0.5) but independent from overall coverages detected for individual gene loci. A very similar approach was previously applied by the Wölfel group and was shown to represent a potent strategy for the effective identification of tumor-specific target antigens in a human melanoma model [226]. Therefore, the conducted strategy was also considered suitable for the efficient identification of NSCLC-derived neoantigen candidates eventually used for further *in vitro* recognition screenings.

Generally, efficient neoantigen prediction via comparative WES and transcriptome sequencing is highly dependent on the quality of applied nucleic acids, number and quality of total sequencing reads, and the efficiency of the exome enrichment protocol during WES library preparation [226,237]. In this study, high WES and transcriptome sequencing qualities were confirmed for all three patients (P18, P43, P50) by achieving very high overall read mappings (>99.4%) to the reference genome hg38. When compared to quality filter-passing read numbers (51-79.5 M), referring to total reads after adapter trimming, resulting from previously conducted tumor WES of melanoma patient Ma-Mel-86 [226], rather high yields of mapped tumor- and normal tissue-derived WES reads were detected for P18 (≥80.1 M), P43 (≥84.9 M), and P50 (≥60.9 M), respectively. Transcriptome sequencing of tumor-derived RNA samples resulted in total mapped read counts of approx. 71.7 M for patient P18, 69.3 M for P43, and 55.8 M in case of P50. While detection rates of mutated mRNA transcripts decrease with fewer RNA-Seq reads, total read counts of ≥50 M were shown to be ideal for the effective identification of neoantigens [237]. Therefore, all conducted NGS approaches were considered suitable for the intended purpose of detecting neoantigen candidates as efficiently as possible.

4.3 Using *in silico* HLA binding predictions for the selection of potentially immunogenic 9- and 10-mer neopeptide candidates

HLA class I peptide binding predictions were performed after identifying non-synonymous P18, P43, or P50 tumor mutations to rank resulting 9- and 10-mer neopeptide candidates based on their potential capability to bind to a patient's individual MHC class I molecules. Prediction tools typically report peptide binding strengths in units of predicted affinity, referring to half maximal inhibitory concentration IC_{50}, or as percentile score reflecting the relative affinity of a selected peptide compared to a universe of random sequences [246]. While high affinity epitopes are typically defined by very low IC_{50} values of ≤50 nM, intermediate and poor binders are regularly found to have affinities in the 50 nM to 500 nM range and above [247]. In this study, most promising candidate peptides had predicted binding affinities of ≤500 nM (NetMHC-4.0 prediction algorithm) and/or a percentile rank of ≤6.0 (IEDB MHC-I binding prediction tool; Consensus or NetMHCpan4.0 prediction methods).

The top scoring 94 neopeptide candidates (9- and 10-mers) were predicted for each patient and were then synthesized for further *in vitro* neoantigen recognition screenings. In case of NSCLC patient P18, 73 SNVs and one frameshift-inducing deletion accounted for a total of 581 potentially well-binding 9-/10-mer peptide candidates (Figure 3.2.4). Notably, NetMHCpan-4.0 was the only prediction algorithm available for HLA-C*03:04 binding assessments, thereby resulting in three times more high scoring peptide candidates than for any of the other HLA alleles. In comparison, 184 tumor-specific variations (SNVs or MNVs) and 3 indel mutations were identified for P43 while NGS analysis of P50 reveled 21 tumor-specific indel mutations, 6 replacements, 15 MNVs, and 63 SNVs. Subsequent MHC class I

binding predictions resulted in 849 and 629 potential well-binding neopeptide candidates for P43 and P50, respectively (Figures 3.3.4 and 3.3.5).

It is known that most HLA class I alleles preferentially bind 9- and 10-mer peptides and that peptides with high predicted binding affinities contain the vast majority of T cell epitopes. Furthermore, data suggests that 9-mer peptides are the most common peptides available for MHC-I binding with 8-mers, 10-mers, and longer peptides being far less frequent [18]. However, exclusively focusing on peptide binding affinity assessments of 9-/10-mer neopeptide candidates carries the risk of missing peptides of non-canonical lengths, thereby neglecting up to 8% of all possible T cell epitopes [248]. In theory, binding predictions could be conducted for peptides of any length to increase the overall chances of antigen identification. But even when peptides of non-canonical lengths might have decent binding predictions, they usually end up being underrepresented among the naturally presented ligands and are consequently less frequently recognized by T cells [18]. Hence, for the current study, it was decided to focus all neoantigen prediction efforts solely on the identification of potentially well-binding 9- and 10-mer neopeptide candidates.

Generally, the T cell-mediated recognition of peptide-MHC complexes is a highly sophisticated process encompassing many crucial steps and requirements. Hence, various antigen characterization pipelines have been established to specifically predict neoantigen binding to a patient's unique pattern of HLA molecules [245]. The peptide-binding sites of MHC class I molecules are encoded by highly variable *HLA* exons 2 and 3, which substantially contribute to the overall polymorphism of HLA alleles [249]. Hence, different MHC class I molecules bind distinct sets of peptides and are therefore capable of eliciting highly specific CD8+ T cell responses [250]. Different factors such as the pre-processing of peptides, the transport into the endoplasmatic reticulum (ER), MHC-loading, and the stability of assembled peptide-MHC complexes have the potential to crucially affect neoantigen presentation and, thus, also immunogenicity [251–253]. Hence, strong binding predictions may be meaningless if upstream processing events prevent the actual loading of the peptides or when predicted peptides cannot be generated by the immune proteasome *in vivo* [245]. Consistent with the data presented in the current project, former studies found that only a very small portion (approx. 5-6%) of all identified mutations might be associated with neoantigen-specific T cell responses. Therefore, accuracies of peptide binding predictions are limited, which is particularly true for some *HLA-C* alleles. [226,254,255].

As reviewed by Richters and colleagues, most of the more recent prediction algorithms are trained on data either generated from mass spectrometry-based approaches or from *in vitro* binding assays involving specific MHC molecules and peptide libraries. In contrast, early prediction algorithms mostly relied on linear regression models, which assumed a linear contribution of individual residues to the overall binding affinity. Moreover, modern prediction algorithms almost exclusively rely on artificial neural networks to capture the nonlinear relationship between peptide sequences and respective MHC-binding affinities [245]. Due to the growing number of available training data sets, artificial neural network-based algorithms usually achieve higher accuracy than linear regression models [256]. In the present study, a combination of two prediction tools, namely the NetMHC-4.0 algorithm and the IEDB MHC-I binding prediction tool (see Table 5.8), were used to predict the MHC class I binding of neoantigen candidates as accurately as possible. NetMHC-4.0 [257] is an artificial neural network-based algorithm that has been trained on data from 81 different human and 41 non-human MHC alleles to accurately predict peptide binding affinities to HLA class I molecules. Depending on the HLA alleles analyzed during this study, two different prediction methods were applied when using the IEDB MHC-I binding prediction tool, namely the Consensus [258] and NetMHCpan-4.0 [259] algorithms. The Consensus prediction method leverages combined data from of up to four different scoring matrices to predict 8-, 9-, and 10-mer peptide affinities to MHC-I molecules. On the other hand, NetMHCpan-4.0 is another artificial neural network-based algorithm that has been trained by using (i) naturally eluted ligand data covering 55 HLA and mouse alleles and (ii) binding affinity data of 172 human and animal MHC molecules, thus being able to predict peptide binding affinities to any MHC molecule of known sequence [245].

However, additional artificial neural network-based algorithms could be used to optimize the peptide binding predictions performed during this study, which might especially be useful for P43 predictions resulting in the highest number of potentially well-binding candidate peptides (849 different 9-/10-mers) among all analyzed patients. On the one hand, NetChop, the most cited proteasomal cleavage tool, could be used to predict intracellular processing of peptide candidates for later MHC class I-dependent antigen presentation [260]. NetChop is a neural network-based prediction algorithm exhibiting the best performance for the prediction of both *in vivo* (NetChop-3.0 20S) and *in vitro* (NetChop-3.0 C-term) cleavage patterns [261]. On the other hand, NetMHCstabpan-1.0 [262] could be used for assessing the binding stabilities of peptide-MHC complexes while NetMHCcons-1.1 [263] combines three state-of-the-art prediction methods (NetMHC, NetMHCpan, PickPocket [264]) to potentially achieve the best results possible. However, since prediction pipelines vary in their choices of how to rank putative neopeptides by differentially weighting each component of the antigen presentation pathway, the common Immune Epitope Database (IEDB) could eventually be used to optimize and rank the conducted predictions by leveraging immune epitope data from more than 1.6 million experiments [245,265]. However, different MHC class I molecules are associated with different peptide-binding-repertoires, thereby resulting in varying epitope numbers being bound with accordingly variable median affinities. Thus, including dedicated binding thresholds for individual HLA alleles could be beneficial to increase the efficacy of future antigen prediction pipelines [246]. With only a few gold-standard data sets being available for evaluating the precisions of divergent predictions and although best practices for determining consensus neoantigen candidates are poorly articulated [245], combining various prediction tools might provide the opportunity to efficiently streamline bioinformatic prediction pipelines while simultaneously compensating variabilities present among different algorithms.

However, instead of optimizing the peptide binding predictions performed during this study by using additional algorithms, an alternative and rather generalized neoantigen identification strategy was applied to completely circumvent the necessity of HLA-peptide binding predictions and dedicated peptide synthesis for subsequent candidate neoantigen identification. The screening of tandem minigene (TMG) constructs, each of which encoded multiple 25-mer peptides with one tumor-specific mutation or TAA-derived epitope in their respective centers, offered the opportunity to analyze candidate antigen processing, presentation, and recognition in the context of all patient HLAs, respectively. Hence, as already described in other studies [5,190,266], target cell lines were generated to express one of the patient's (P43) respective HLA alleles and were subsequently transfected with patient-specific TMG libraries encoding a total of approx. 120 TAAs and tumor-specific mutations that were previously identified via WES and RNA-Seq. However, consistent with the data obtained from P43 TMG analyses, applying TMG screening approaches might not necessarily result in improved overall neoantigen detection efficiencies. While other large scale *in vitro* screenings comprising hundreds of TMG-encoded neoantigen candidates previously resulted in the identification of novel targets specifically recognized by patient-derived TIL TCRs [190,267], not a single TCR-T cell reactivity was observed when screening for P43 TMG library recognition (data not shown). Hence, these results indicate that (i) extensive TMG screening approaches might be necessary for efficiently identifying patient-specific neoantigens and that (ii) HLA-I binding prediction is just one of many bottlenecks that must be mastered to successfully identify promising neoantigen candidates.

4.4 Establishing transgenic TCR-T cells for further neoantigen recognition testing

After the identification of potentially tumor-specific TCRs from patient-derived TIL populations, most promising candidate TCRs were cloned into bicistronic vector backbones eventually used for retroviral transductions of healthy donor T lymphocytes. To facilitate the efficient surface expression of the introduced TCRs, chimeric TCR expression constructs were used comprising codon-optimized murine αβ TCR constant regions, a second interchain disulfide bond, rearranged TCR chains (β-P2A-α), and human αβ chain variable domains [225]. Using chimeric TCRs for retroviral transduction of donor

lymphocytes represents an effective strategy to reduce mispairing between transgenic and endogenous TCR chains and, due to improved pairing and TCR/CD3 stability, might even result in genetically modified T cells with increased reactivity [268]. Since mispairing might result in the formation of mixed TCR dimers consisting of transgenic and endogenous αβ chains, minimizing the pairing of endogenous and transduced TCR chains was an absolute prerequisite to avoid unpredictable target specificities during TCR antigen screenings [269].

Before being retrovirally transduced to express potentially tumor-specific TCRs, purified healthy donor lymphocytes were additionally used for a CRISPR/Cas9-based KO approach simultaneously targeting both endogenous TCR (end.TCR) chains. Since transgenic and endogenous TCRs compete for CD3 complex recruitment, the conducted end.TCR-KO approach facilitated the efficient binding of transgenic TCRs to CD3, thereby generating TCR-T cells with high functional avidity [269,270]. Single alpha chain end.TCR-KO approaches previously achieved expression reductions of up to 81% (± 11%) while declines of 87% (± 4.5%) were observed when exclusively targeting the TCR beta chain. However, highest TCR-KO efficiencies (91.6% ± 4.2%) were achieved when simultaneously targeting both TCR chains [271]. Concordantly, end.TCR expression was regularly reduced to less than 5% in purified donor lymphocyte populations used during the current study. In some cases, residual end.TCR expression was actually absent when analyzed via flow cytometry. Hence, the overall risk that the formation of mixed TCR dimers may affect the efficiency of conducted antigen screenings was dramatically reduced.

The generated end.TCR-KO T cells were eventually used for the retroviral transduction to introduce the ectopic expression of potentially tumor-specific TCRs initially cloned from TILs of multiple NSCLC patients. Firstly described by Steven Rosenberg *et al.* [272], retroviral transduction of donor lymphocytes with TCR expression constructs has become a common strategy to generate transgenic TCR-T cell populations for basic research as well as clinical studies [273]. To achieve maximal yields of end.TCR-KO lymphocytes prior to retroviral transduction, the pre-existing T cell transduction protocol was optimized by applying a common method, referred to as “rapid expansion protocol” (REP), aiming at quickly generating large numbers of T lymphocytes suitable for following transduction [274]. Here, T lymphocytes were excessively expanded by non-specific activation with the anti-CD3 monoclonal antibody OKT3, cultivation in the presence of high IL-2 doses, and the addition of irradiated feeder cells for co-stimulation [275]. The addition of auxiliary cells expressing costimulatory molecules like CD80, CD86, or CD137 was previously demonstrated to induce more effective cell proliferation than anti-CD3 antibodies alone. Furthermore, when compared to anti-CD3/CD28 bead stimulation, this protocol was previously demonstrated to result in higher overall T cell yields, which might be mediated by inducing significantly less activation-induced cell death (AICD) [276].

In contrast to lentiviruses, which are regularly used for the infection and genomic engineering of proliferating as well as non-proliferating cells, most other retroviruses rely on cell cycle progression for productive infection and transduction [277]. Hence, during this study, highest yields of TCR-T cells were established by combining two stimulation strategies: The REP was used to activate T lymphocytes directly after purification from healthy donor blood samples and before performing the end.TCR-KO, thereby enhancing the efficiency of following retroviral transduction approaches. In a second step, anti-CD3/CD28 beads were used to re-stimulate transgenic T cell populations directly after retroviral transduction with exogenous TCR (exo.TCR) expression constructs. By combining the adapted REP and subsequent puromycin selection for TCR-T cell manufacturing, exo.TCR expression of >90% was regularly achieved in established effector populations. In some cases, especially upon transduction of patient P43-derived TCRs, transgenic T cell populations quickly perished upon puromycin selection (see section 3.3.1) while not expressing detectable exo.TCR levels at all (data not shown). However, Sanger sequencing of the respective TCR expression constructs revealed the integrity of included puromycin resistance genes (data not shown). Hence, it was suspected that bacterial amplification might have introduced crucial mutations within long terminal repeat (LTR) sequences of used expression constructs, thereby impairing gene expression, polyadenylation, as well as retrotransposition of transduced exo.TCRs and corresponding puromycin resistance genes [278].

Today, CRISPR/Cas9-based orthotopic TCR replacements (OTRs) facilitate the controlled editing of TCR expression via the targeted knock-in of transgenic αβ chains into the endogenous *TCR* locus of a target T cell. Here, full transgenic αβ TCRs are usually inserted into the endogenous *TCR alpha* locus, thereby introducing transgenic TCR expression under the control of physiological promotors [273,279]. Although resulting in effector populations with safety profiles quite similar to natural T cells, CRISPR/Cas9-based OTRs might induce numerous off-target effects, such as off-target integration, double strand breaks, or chromosomal translocations. Furthermore, due to cargo delivery via electroporation, CRISPR/Cas9-based OTRs often result in the excessive loss of target T cells after genetic modification [279–281]. When compared to OTRs typically resulting in very homogenous TCR expression that is quite similar to native T cells, varying transgene numbers and untargeted transgene integration may result in more variable TCR expression as well as functionality when performing conventional viral transduction [273]. Hence, CRISPR/Cas9-mediated OTRs could represent a potent alternative for generating transgenic T cell populations for future clinical trials, especially when overcoming current challenges such as the rather low T cell engineering efficiency compared to conventional transduction approaches [279].

4.5 Efficiency assessment: Identification of tumor-specific T cells in patient-derived TIL populations

The overall aim of the conducted proof-of-concept study was to verify the suitability of a novel prediction pipeline aiming at the identification of tumor-specific T cells among TIL populations of NSCLC patients, thereby providing a promising starting point for future T cell-based immunotherapy approaches. Hence, tumor-specificities of 16 TCRs were analyzed after being cloned from tumor-enriched TIL populations of the three NSCLC patients P18, P43, and P50. By using complex recognition screenings, dedicated neoantigen specificities were revealed for three of those TCRs while simultaneously demonstrating that 7 of the 16 analyzed TCRs were capable of recognizing autologous tumor cells. More precisely, while tumor reactivity was observed for TCRs cloned from two (P18, P43) of three NSCLC patients, it was demonstrated that the three P18-derived TCRs 55.1, 28.1, and 73.1 specifically recognized a prominent $KRAS^{Q61H}$ neoantigen via HLA-A*01:01 (see section 3.2).

When analyzing all four P18 TCR-T cell populations, P18 TCR 54.2 was the only receptor shown to recognize P18-derived tumor cell suspensions (Figure 3.2.3) as well as the HLA-A*02-matched lung cancer cell line MZ-LC-16 (Figure 3.2.15), thereby suggesting the recognition of an antigen shared by both tumors. However, to exclude the possibility that the observed MZ-LC-16 recognition might be due to alloreactive T cell responses against foreign peptide-MHC complexes, additional experiments must be performed using monoclonal antibodies to specially block the HLA-A*02-mediated antigen recognition.

In contrast to P18, no tumor reactivity could be demonstrated for any of the tested P50 TCRs (Figure 3.3.3 B) while P43 tumor reactivity was only observed for TCR-T cells either expressing P43-derived TCRs 55, 79, or 93 (Figure 3.3.3 A). Generally, both P43 and P50 tumor recognition testing experiments lead to rather ambiguous results, which was most likely due to the poor qualities of patient-derived tumor cell suspension accounting for high numbers of non-viable cells after thawing. In addition, except for P50 TCR 36.7 that specifically recognized a CMV-derived IE2 target peptide in the context of HLA-B*44:03 (Figure 3.4.1 B), no specific target antigens could be identified for any of the remaining P43 or P50 TCRs analyzed during this study. However, those results do not necessarily imply the ineligibility of the HSD-presented *in silico* pipeline applied for the identification of potentially tumor-reactive T cell clonotypes. For example, the rather inefficient recognition testing experiments performed for patient P50 might be explained by the fact that only a small fraction (94 peptides) of all identified P50 neopeptide candidates (629 9-/10-mer peptides in total; see Figure 3.3.5) were tested in the course of this study. When additionally considering the challenges that the applied and highly complex antigen prediction pipeline entails, such as assessing potential HLA binding affinities as well as intracellular

peptide processing (see section 4.3), even more extensive neoantigen screenings might be necessary to efficiently identify tumor-specific target antigens across broad patient cohorts.

Predicting tumor-specific clonotypes from highly heterogenous TIL populations still remains challenging, which is also true for the attempt to compare pipeline efficiencies using different prediction strategies. However, Pasetto and colleagues published an approach quite similar to the strategy presented in the current study. While exclusively using patient-derived tumor tissues for TCR beta chain sequencing and following TIL clonotype frequency assessments, they also assessed tumor-reactivities of most dominant CD8+/PD-1+ TIL clonotypes found in freshly resected tumors of 12 metastatic melanoma patients. By analyzing 4-1BB (CD137) expression of TCR-T cell populations upon co-incubation with autologous tumor cells, they managed to identify 36 paired TCRs of mostly undefined specificity, each of which recognized autologous tumor cells, respectively. Hence, they managed to verify tumor reactivities for 43.37% of all analyzed TCRs (83 receptors were analyzed in total) [215]. Consistently, comparable efficiencies were demonstrated for the HSD-presented prediction pipeline by demonstrating autologous tumor reactivities for 43.75% (7 of 16 analyzed TCRs) of all TCRs that had initially been predicted to be potentially tumor-specific.

Moreover, Pasetto *et al*. detected 11 neoantigen-specific TCRs in TIL populations obtained from 5 patients with metastatic melanoma, thereby accounting for 13.25% of all analyzed TCRs [215]. In contrast, the current study identified three neoantigen-specific TCRs together accounting for 18.75% of all analyzed TCRs, thereby again highlighting the suitability of the applied tumor-specific TCR identification pipeline. However, while Pasetto and colleagues verified target neoantigen recognition in 5 different melanoma patients, the current study demonstrated neoantigen specificities towards a single KRAS mutation (p.Q61H) present in only one NSCLC patient. Furthermore, the Pasetto study verified neoantigen recognition in 5 of the 10 patients whose TCRs were tested against autologous mutations [215], thereby being more efficient when compared to the approach applied for the current study that demonstrated neoantigen-specific T cells in one of three NSCLC patients. Taken together, both studies demonstrated that patient-derived TIL populations are highly suitable for the efficient identification of tumor-specific T lymphocytes, but they also indicate that identification efficiencies might be highly variable among different patients and/or tumor types.

Since the efficiency to identify tumor-specific T cells might be highly dependent on analyzed tumor specimens, it might not be representative to compare prediction pipelines analyzing different tumor types. This might particularly apply to tumors with highly different tumor mutational burden (TMB) levels, which are, in turn, supposed to directly correlate with detectable TIL numbers [282]. Generally, high-TMB tumors are found across nearly all malignancies but corresponding frequencies might be highly variable among both patient cohorts and cancer types. Cancer types known to have significant mutagen exposure, such as lung cancers (median TMB 7.2 mutations/Mb) or melanomas (median 13.5 mutations/Mb), typically exhibit higher TMB levels than others, including pediatric malignancies (median 1.7 mutations/Mb) or bone marrow myelodysplastic syndromes (median 0.8 mutations/Mb) [283]. Hence, NSCLCs analyzed during the current study, as well as metastatic melanomas analyzed by Pasetto and colleagues, can certainly be considered as high-TMB malignancies, thereby potentially accounting for high numbers of tumor-specific mutations.

When comparing the three patients analyzed during this study, combined WES and RNA-Seq of P18 tumor and normal lung tissues revealed 159 non-synonymous tumor-specific variants (before manual curation) also present (RPKM ≥0.5) in the corresponding tumor transcriptome data (Figure 3.2.4). In contrast, 3,169 (P43) and 2,626 (P50) non-synonymous tumor-specific mutations were identified when analyzing respective P43- and P50-derived exome and transcriptome data (Figures 3.3.4 and 3.3.5), thereby indicating substantial TMB differences between the three analyzed tumors. Although using different parameters and predictions strategies for the identification of those mutations, this data is consistent with mutational burdens detected in the Pasetto study, where the analysis of 11 metastatic melanoma exomes revealed non-synonymous tumor mutation levels (3 exome variant reads, 8% variant allele fraction in the exome, 10 reads in the matched normal sample) ranging between 159 and 3,976

variants per patient [215]. Furthermore, the results presented in the current study are highly comparable to data generated in a previous project conducted by the Wölfel group, which identified 2,099 to 3,108 tumor-specific mutations in exomes of 4 different melanoma cell lines by applying a very similar WES approach [226]. However, in contrast to the other studies, non-synonymous tumor-specific mutation levels reported for P18, P43, and P50 exclusively represented variations that were also found (RPKM ≥0.5) in the corresponding donor-matched tumor transcriptome data. When also considering that melanomas typically tend to have higher TMBs than NSCLCs [283], it can be assumed that the NGS pipeline used in the current study represents a very promising approach for the efficient identification of mRNA-expressed non-synonymous tumor mutations in individual cancer patients.

Furthermore, high numbers of MHC class I-restricted neoantigens are typically associated with increased anti-tumor activity and are therefore considered potent drivers for cytotoxic T cell responses [284]. Since neoantigens arise from tumor-specific mutations, high-TMB cancers generally provide greater chances to exhibit immunogenic neoantigens resulting in T cell-mediated cancer recognition and eradication [285]. Therefore, the neoepitope prediction pipeline applied during this study might particularly be suitable for analyzing hot tumors with increased TIL numbers and high levels of non-synonymous mutations. Since tumors of NSCLC patients with smoking history typically exhibit rather high mutation rates [286], they were perfectly suitable for conducting this proof-of-concept study. Moreover, immune checkpoint blockade with monoclonal antibodies seems especially effective for the treatment of high-TMB cancers with elevated neoantigen loads [287]. Hence, a combination of tumor-specific cell therapies and immune checkpoint inhibition (ICI) could eventually represent a promising treatment strategy for all three patients, but especially for P43 and P50 exhibiting even higher TMB levels and potential neoantigen loads than P18. By facilitating the activation of neoantigen-specific T cells, ICI could circumvent T cell suppression within a patient's TME, thereby promoting immune cell-mediated tumor eradication [245].

In addition to the prediction of personalized neoantigen candidates and the identification of potentially tumor-specific TCRs, applying a reliable *in vitro* readout test was essential to finally verify TAA and neoantigen reactivities of generated TCR-T cells. While the current study used IFN-γ secretion as marker for T cell recognition, Passetto and colleagues assessed CD137 (4-1BB) expression to measure target recognition by transduced TCR-T cells [215]. Regardless of produced effector cytokines or the differentiation stage of the cell, CD137 expression is uniformly up-regulated in CD8+ T cells approx. 24 h after stimulation but is absent in non-stimulated CD8+ lymphocytes [215,288]. Therefore, CD137 expression can be used to identify CD8+ T cells that have recently been activated via TCR engagement and signaling, thereby representing a suitable marker for the identification of tumor-reactive TILs and neoantigen-specific T cell populations [231]. Generally, CD137 expression seems to correlate with IFN-γ and TNF-α cytokine production in activated T cells [288]. However, since there are T cells that upregulate CD137 expression but do not secrete IFN-γ [289], a significant portion of tumor-reactive T cells might be missed when exclusively assessing IFN-γ release [288]. Hence, also considering expression levels of additional T cell activation markers such as CD137 might offer the opportunity to further increase the efficiency of the applied strategy aiming at the reliable verification of tumor-reactive T cells *in vitro*. Ultimately, steady optimization of the conducted *in vitro* readout assays could significantly contribute to gradually reduce the overall error-proneness of the applied pipeline to a level where the necessity of identifying actual target neoantigens before adoptive cell therapy (ACT) might be obsolete, thereby also bypassing labor-intensive HLA binding predictions and extensive neoantigen candidate screening experiments [215,231].

4.6 KRAS^{Q61H} as a potent target for T cell-based immunotherapies

The RAS family is a group of small GTPases hydrolyzing guanosine triphosphate (GTP) to convert cell surface receptor stimulation into downstream signaling necessary for cell proliferation and survival [193]. Due to alternative exon 4 splicing, the three genes of the *RAS* family (*KRAS*, *NRAS*, *HRAS*) encode 4

highly similar RAS protein isoforms with highly related structural characteristics, biochemical properties, and shared overall amino acid identities of 82-90%. More precisely, amino acid sequences of all isoforms are identical at the N-terminus of their G domain (AA 1-86) while exhibiting 82% sequence similarity within residues 87-166 and rather low conformity in C-terminal hypervariable regions (AA 167-188/189) [290]. Members of the *RAS* family are amongst the most frequent oncogenes with *RAS* point mutations occurring in up to 30% of all sequenced cancers. Regarding protein variability, KRAS is the most frequently mutated protein in human cancers, followed by NRAS and HRAS [291]. While KRAS mutations account for up to 86% of all found RAS mutations, NRAS and HRAS represent about 11% and 3% of all detected variants, respectively [196,290]. Moreover, there are cancer-specific mutational profiles of RAS isoforms, thereby suggesting tissue-distinct roles of RAS proteins in driving oncogenesis. For example, KRAS is the only isoform to be mutated in pancreatic ductal adenocarcinoma while also accounting for up to 86% and 96% of RAS mutations found in colorectal adenocarcinoma and lung adenocarcinoma, respectively. In contrast, NRAS is the predominant isoform mutated in cutaneous melanoma while mutated HRAS is most common in head and neck squamous cell carcinomas [290].

During this project, three different TCRs (TCRs 28.1, 55.1, 73.1), which had initially been cloned from heterogenous TIL populations of lung adenocarcinoma patient P18, were demonstrated to recognize a p.Q61H-mutated KRAS 10-mer peptide (ILDTAG<u>H</u>EEY; mutated AA highlighted) via HLA-A*01:01 with no substantial reactivity against wild-type KRAS. Here, the $KRAS^{Q61H}$ recognition was demonstrated by using (i) K562 targets pulsed with mutated 10-mer peptides (see Figure 3.2.6), (ii) mutant cDNA-transfected HEK 293T cells (see Figure 3.2.8), and (iii) via the $KRAS^{Q61H}$-positive lung cancer cell line NCI-H460 (see Figure 3.2.10). However, since allogenic T cells were used for target recognition testing and although a CRISPR/Cas9-based knock-out of endogenous TCR expression was performed before introducing the P18 TCRs, the final proof that the observed T cell activation was actually due to the specific recognition of the $KRAS^{Q61H}$ neoantigen was still pending. Hence, a CRISPR/Cas9-based knock-in (KI) approach was used to replace the p.Q61H mutation with one (p.Q61R) of three (p.Q61K; p.Q61L; p.Q61R) non-recognized KRAS variations (see Figure 3.2.8) in HLA-A*01:01+ NCI-H460 lung cancer cells. The lacking recognition of CRISPR/Cas9-treated NCI-H460 cells by P18 TCR-T cell populations (Figure 3.2.14) clearly verified the $KRAS^{Q61H}$ neopeptide as the specific target antigen recognized by all three TCRs. Moreover, by assessing the functional avidity of P18 TCR-transduced CD4+ and CD8+ effector populations, it was found that, depending on the introduced TCR, both T cell subsets specifically recognized the $KRAS^{Q61H}$ neopeptide (see section 3.2).

The target recognition mediated by P18 TCR 55.1 seemed less dependent on CD8+ coreceptor expression than T cell activations induced by TCRs 28.1 and 73.1, which might be suitable in the context of future immunotherapies: As target antigen-specific CD4+ T cells can either execute direct effector functions or support the proliferation and survival of CD8+ effector cells, injection of CD4+ and CD8+ TCR-T cells targeting the same antigen might produce synergistic effects potentially resulting in the complete eradication of malignant cells [292]. However, when compared to CD8+ TCR-T cells, less IFN-γ secretion was generally observed in transgenic CD4+ effectors. In addition, when co-incubated with $KRAS^{Q61H}$-positive target cells, upregulation of the degranulation marker CD107a was weaker in CD4+ TCR-T cells than in corresponding CD8+ effector populations. The overall degranulation capacity seemed elevated in transgenic CD8+ lymphocytes expressing TCRs 28.1 and 73.1, respectively (Figure 3.2.11). Also, overall T cell activation of TCR 73.1-expressing effectors seemed slightly increased when comparing IFN-γ secretion capacities of P18 TCR-transduced T cell populations generated from identical healthy donor lymphocytes, thereby indicating a particularly high functional avidity of TCR 73.1-transduced effectors. Nevertheless, whether and to which extent minor differences in CD8 coreceptor dependence or functional avidity of transgenic T cells might actually affect the suitability of P18 TCR candidates for future immunotherapies is highly speculative and requires dedicated studies to assess TCR-T cell performances *in vivo*.

In general, KRAS mutations are commonly found (ca. 98% of all KRAS mutations) at three hotspot positions (codons 12, 13 and 61) with multiple substitutions occurring at each of the corresponding amino acids. With variations at amino acid positions 12 and 13 being predominant across all KRAS mutations [4,293,294], the p.Q61H mutation is less frequent and only accounts for approx. 2% of all detected KRAS mutations [290,293]. Assessing overall *RAS* codon variation frequencies revealed distinctive mutation signatures for every isoform, respectively. For example, the *KRAS* gene is commonly mutated at codon 12 (ca. 80% of all mutations) whereas very few mutations are observed at codon 61. In contrast, up to 60% of *NRAS*-mutated tumors harbor mutations at codon 61 and only approx. 35% exhibit variations at codons 12 or 13. Moreover, *HRAS*-mutated tumors show a rather balanced ratio between variations found at codons 12 and 61, respectively. Depending on the reference used, codon 12 mutations are thought to account for approx. 35-50% of all *HRAS* variants while codon 61 variation frequencies are assumed to range between 34% and 40%. In contrast to *RAS* codons 12 and 61, codon 13 mutation frequencies (roughly 12-27%; depending on the reference used) are more comparable across all three isoforms [196,295].

According to the AACR Project GENIE Consortium, an international data-sharing consortium focused on the integration of genomic data from thousands of cancer patients, $KRAS^{Q61H}$ is the most prevalent mutation occurring at codon 61, thus being found in approx. 1% of all colorectal adenocarcinomas, approx. 1.2% of all non-small cell lung carcinomas, ca. 5% of pancreatic ductal adenocarcinomas, and roughly 0.6% of all malignant solid tumors [296,297]. Here, somatic missense mutations at KRAS amino acid 61 facilitate the perturbation of the intrinsic GTPase activity of the KRAS protein, thereby resulting in reduced GTP hydrolysis capacities and permanent activation of the RAS molecule. Moreover, oncogenic KRAS mutations were found to perturb multiple biochemical cell pathways, such as MAPK, PIK3K/AKT, NORE1/RASSF1, JAK/STAT3, and many more. Therefore, mutated KRAS proteins potentially increase cell proliferation and survival while simultaneously deregulating apoptosis, growth arrest, and cell differentiation [298]. Generally, there are 6 single-base missense mutations at *KRAS* codon 61, with p.Q61H (approx. 57%) and p.Q61K/L/R (collectively approx. 40%) being most frequent across all human cancers. In contrast, p.Q61P and p.Q61E mutations are the rarest and account for approx. 1-2% of codon 61 variations, respectively. While p.Q61H is predominant at KRAS residue 61, it is rather rare in both NRAS (ca. 6%) and HRAS (ca. 5%). Moreover, p.Q61R is most common among NRAS and HRAS mutations observed at amino acid position 61 and accounts for approx. 47% (NRAS) and 43% (HRAS) of all codon 61 mutations, respectively [290,299,300]. To be able to assess the number of potentially treatable cancer patients as accurately as possible, all P18 TCRs were additionally tested for potential cross-reactivities against p.Q61K-, p.Q61L-, and p.Q61R-derived neopeptides. However, none of the three P18 TCRs recognized one of the alternative KRAS mutations at all (Figure 3.2.8).

As previously reported by Pearlman and colleagues, the p.Q61R mutation can be found across all three RAS isoforms and encodes an HLA-A*01:01-presented 10-mer peptide (ILDTAGREEY; mutated AA highlighted) representing one of the top 10 public target neoantigens for future cancer immunotherapy approaches. Furthermore, the presenting HLA allele, *HLA-A*01:01*, is among the top 10 most frequent HLA class I alleles in the US (approx. 23.6%). By considering both HLA-A*01:01 expression and p.Q61R frequency, Pearlman *et al*. estimated the theoretical frequency of patients for whom neoantigen-specific T cell immunotherapy could be considered: Depending on HLA-A*01:01 expression, the p.Q61R-derived neoantigen is assumed to be found in approx. 2.9% of melanoma patients and in approx. 0.4% of patients suffering from one of the 12 most common solid tumors in the US. In turn, this estimation results in up to 2900 HLA-matched melanoma patients and, when only referring to the 12 most common solid tumor types, a total of approx. 5000 new US patients per year that carry the p.Q61R-mutated RAS neoantigen [4]. Since the top 5 RAS variations (p.G12D, p.G12V, p.G12C, p.G13D, p.Q61R) account for approx. 70% of all RAS mutations [301], an even lower frequency can be assumed for HLA-A*01:01+ cancer patients carrying the p.Q61H neoantigen mutation.

While overall mutation frequency and HLA occurrence define the number of potentially treatable patients, efficient peptide processing is essential for proper neoantigen presentation and subsequent T

cell recognition. However, direct experimental demonstration of proper processing and HLA binding of cancer-derived peptides is often challenging. Wang and colleagues recently analyzed common mutation-associated neoantigens (MANAs) by co-transfecting COS-7 cells with HLA alleles as well as oncogenes predicted to generate MANAs with the potential to be presented by the respective HLAs. By inducing heavy isotype labelling, they were able to quantify the MANA abundance (femtomole/ml) and corresponding MANA copy numbers per transfected cell. When compared to respective presentations of MANA peptides from rather common KRAS mutations p.G12C, p.G12D, p.G13D, p.Q61L, or p.Q61R, they found that the p.Q61H-derived RAS 10-mer peptide (ILDTAGHEEY; mutated AA highlighted) was most abundant on transfected cells. Co-transfection with HLA-A*01:01 and KRAS oncogenes resulted in substantially higher levels of presented p.Q61H 10-mer peptides than p.Q61R neopeptides (583 vs. 127 copies per transfected cell) [302]. With similar amounts of transfected cDNA, those experiments indicated that the processing and/or presentation of the p.Q61H-derived neopeptide might be more efficient than for other KRAS mutations. However, since peptide processing and presentation may be different among target cells, the extent to which those results also correlate with actual peptide processing and presentation efficiencies in human cancer remains a matter of debate.

By equipping cancerous cells with highly transforming potentials, activating KRAS hotspot mutations are considered major genetic drivers of aggressive tumors [299,303]. Hence, since genetic drivers are essential for tumor establishment, they are likely to be clonally expressed in all tumor cells of a particular patient [304]. In addition, $KRAS^{Q61H}$ was shown to provide resistance to therapeutic strategies aiming at the treatment of cancers driven by major oncogenic KRAS mutations at codon 12, such as SHP2 inhibitors and p.G12C inhibitors (sotorasib, adagrasib) [198,297]. Targeting $KRAS^{Q61H}$ with T cell-based immunotherapies could therefore represent a unique chance to simultaneously target virtually all tumor cells of malignancies with intact MHC class I antigen presentation, which might especially be beneficial for patients that are already excluded from pre-existing treatment strategies. When also considering the severe oncogenic properties of *RAS* gene mutants in various cancer types [298], $KRAS^{Q61H}$ represents a very attractive target neoantigen for the development of novel TCR-T cell treatment strategies in both fundamental research and pharmaceutical industry. Moreover, targeting public neoantigens, including mutated RAS isoform-derived peptides, generally provides the chance to circumvent many limitations inherent to personalized immunotherapy approaches, such as the logistical and financial challenges of producing individual therapies for every patient [4].

When compared to CAR-T cell therapy, TCR-T cell approaches enable the targeting of HLA-presented neoantigens or cancer germline antigens derived from any cellular compartment. Furthermore, they might generally induce less side effects while being considerably more sensitive to low target antigen concentrations [305,306]. Hence, TCRs can not only target membrane associated proteins, which collectively represent approx. 27% of the human proteome, but also intracellular targets like cytoplasmic peptides or nuclear proteins. Thus, TCR-T cells detect a broader universe of targets with high sensitivity, including neoantigens, cancer germline antigens, or viral oncoproteins [307]. By combining lower overall target binding affinity with high functional avidity, TCR-transduced T cells are capable to serially scan and efficiently kill multiple tumor cells while maintaining persistence and exhibiting physiological signaling [308]. Unlike CAR-T cells, TCR-immunotherapy employs the natural TCR to target tumor antigens, thereby underlying a more selective and regulated recognition process with less risk of excessive activation and cytokine release [309]. Moreover, using natural TCRs for therapeutic applications can be considered rather safe, especially when cloning tumor-specific TCRs from autologous TIL clonotypes which have previously passed thymic selection without inducing apparent adverse effects *in vivo*.

Douglass and colleagues recently developed a special type of bispecific antibodies, the so called single-chain diabodies, that are capable to induce T cell activation as well as killing of cancer cells when exhibiting endogenous levels of recurrent RAS mutations (p.G12V, p.Q61H/L/R) together with HLA-A*03:01 or HLA-A*01:01 expression [310]. However, to our knowledge, a patient-derived TCR specifically recognizing the $KRAS^{Q61H}$-derived target peptide (ILDTAGHEEY; mutated AA

highlighted) in the context of HLA-A*01:01 has not yet been identified. However, Rosenberg and colleagues identified various KRAS neoantigen-specific TCRs and their cognate HLA restrictions: By using TIL screenings and *in vitro* stimulation approaches, they verified 10 TCRs in 7 different cancer patients, each of which recognized one of two KRAS hotspot mutations (p.G12V, p.G12D) via differing HLA class I alleles. Furthermore, they identified 13 HLA II-restricted TCRs that recognized KRAS neoepitopes either derived from codon 12 or codon 13 hotspot mutations. When considering overall RAS mutation frequencies and corresponding HLA allele occurrence, they claim that approx. 45% of the Caucasian and about 60% of the Asian US population could potentially benefit from future therapies designed to harness RAS mutation-specific TCRs identified by the Rosenberg group [304]. Since they are currently attempting to identify even more of those TCRs, the p.Q61H-reactive TCRs identified in the current study could substantially contribute to enlarge the population of cancer patients that could, in theory, be treated with RAS neoantigen-specific TCRs.

4.7 *In silico* TCR clustering: A potent strategy for the identification of tumor-reactive TCRs with shared antigen specificities

It is assumed that hundreds or even more distinct TCRs could be able to specifically recognize common immunodominant target antigens such as viral epitopes [217]. Common virus-specific TCRs with highly related CDR3 sequences were previously found in patient-derived T cell populations exhibiting shared antigen specificities [218]. There is some evidence that this might also be true for T cell clonotypes specifically recognizing identical TSAs or TAAs, thereby potentially driving cancer immunity [304,311,312]. For example, the Rosenberg group recently identified distinct T cell clonotypes with consistent CDR3 alpha or beta motif sequences in different cancer patients, thereby revealing highly similar TCRs specifically recognizing neoantigen peptides derived from different KRAS mutations at amino acid positions 12 and 13 [223,304,313]. This is convergent with the data generated during the current study where three neoantigen-specific TCRs were identified that specifically recognized a p.Q61H-mutated KRAS 10-mer epitope via HLA-A*01:01. All three TCRs were cloned from tumor-enriched TIL populations of a single NSCLC patient (P18) and were found to share very similar TCR alpha and beta chain amino acid sequences, with two even sharing identical alpha chains (Figure 3.2.1 C).

The TCR-mediated recognition of MHC-presented peptides is mainly facilitated by three complementarity-determining regions (CDRs) present in each of the two TCR chains. While respective CDR1 and CDR2 regions of the TCR contact the MHC molecule, the hypervariable CDR3 domains mainly contact the MHC-bound peptide [314]. Hence, minimal changes in CDR3 regions can dramatically alter the antigen specificity pattern of a particular TCR, even to the point of entirely abrogated recognition [315]. When pre-selected against a single peptide, Glanville and colleagues found that antigen-specific TCR pools were enriched for very similar CDR3 sequences not differing in more than 2-4 AAs on average. Based on their findings, they established the GLIPH (grouping lymphocyte interactions by paratope hotspots) algorithm to organize TCR sequences into distinct clusters with shared specificities across a group of individuals. Here, they assessed additional attributes, such as global TCR sequence similarity, structural peptide antigen contact propensity, CDR3 length, and shared HLA alleles, among TCR contributors to analyze T cell responses independently of knowing the actual epitope specificities or MHC restrictions [316]. However, GLIPH and comparable public algorithms exclusively focus on TRBV-CDR3 sequences. Hence, HSD established a bioinformatic pipeline to detect potential CDR3 sequence similarities between paired TRAV/TRBV-CDR3 domains included in the HSD sequencing database comprising >600,000 TCRs from approx. 60 NSCLC patients. To identify target antigen candidates potentially recognized by highly related TCRs, referred to as TCR clusters, HSD also included TCRs of known antigen specificity that had previously been obtained from databases like VDJdb and McPAS-TCR (see Table 5.8).

Hence, HSD identified a promising TCR cluster, referred to as cluster C, consisting of 6 TCRs (C-TCRs) with shared TRAV- and TRBV-CDR3 regions only differing in a maximum of 3 amino acids,

respectively. Interestingly, one of those C-TCRs (TCR 75) was previously selected as potentially tumor-specific in patient P43 and had therefore already been used for extensive neoantigen screenings (see section 3.3). Since previous TCR 75 recognition testing approaches were not successful, the TCR clustering approach represented an alternative strategy to identify potential target antigen candidates for this receptor. Moreover, since all 4 TCR cluster C patients shared *HLA-B*08:01* and *HLA-C*07:01* alleles, all upcoming antigen screening approaches exclusively focused on HLA-B*08:01- and HLA-C*07:01-presented peptide candidates. By testing various lung cancer cell lines, similar recognition patterns were verified for all C-TCRs, each of which mediated the recognition of the HLA-B*08:01+/C*07:01+ squamous cell carcinoma cell line NCI-H1703 (Figure 3.4.3). Moreover, C-TCR 75 specifically recognized various EBV+ melanoma patient-derived B cell lines but not the corresponding non-infected tumor cells (Figure 3.4.2 B). Since HLA-B*08:01+ K562 but not HLA-C*07:01+ or non-transduced K562 cells were recognized by all C-TCR+ effectors (Figure 3.4.3), it was concluded that the observed tumor cell recognition was most likely dependent on HLA-B*08:01 expression. However, no HLA allele, not even *HLA-B*08:01*, was shared across all recognized target cells (see Table 5.1 for respective HLA-I genotypes).

It was therefore assumed that the corresponding target peptide might either be presented by more than one MHC molecule or by non-classical HLAs shared by all recognized cell lines, such as HLA-E, HLA-F, or HLA-G. In contrast to other non-classical HLA class I molecules, HLA-E is transcribed in most human tissues and cell lines, although at lower levels than highly polymorphic classical MHC molecules [317,318]. A total of 27 HLA-E alleles have been reported to date, most of which are rather rare or encode non-functional proteins. However, two HLA-E alleles, namely *HLA-E*01:01* and *HLA-E*01:03*, are predominant in the human population while only differing in a single amino acid at position 107 [319]. When compared to HLA-E*01:01, normal cells usually express significantly higher levels of HLA-E*01:03, which binds available peptides with higher affinity and tends to form more stable HLA-peptide complexes [318]. Under normal conditions, HLA-E proteins almost exclusively bind a limited set of 9-mer peptides derived from leader peptides of classical HLA proteins. However, HLA-E expression can often be detected under pathologic conditions, such as lymphomas, colon cancers, melanomas, and autoimmune diseases. Besides of binding heat shock protein-derived peptides under cell stress, HLA-E is also able to present pathogen-specific peptides either arising from bacteria or viruses such as EBV [318,320].

Potent CD8+ T cell responses against an HLA-E-presented target peptide derived from the EBV transcription factor BZLF1 were previously verified across large patient cohorts [321]. When compared to other non-classical HLA class I molecules, databases like IEDB (see Table 5.8) list plenty of EBV-derived peptides predicted to bind HLA-E*01:01 or HLA-E*01:03, thereby indicating HLA-E as a potent contributor to EBV-directed immune responses [321]. Hence, the non-recognized (Figure 3.4.2 B), EBV-transformed, and HLA class I-negative B-lymphoblastoid cell line L721.221 [322] was individually transfected with one of both HLA-E alleles to further analyze the MHC restriction of C-TCRs. However, HLA-E transfection did not affect the observed target cell recognition at all (data not shown), thereby indicating that cluster C TCR-mediated target recognition is not dependent on the expression of HLA-E*01:01 or HLA-E*01:03, respectively. According to databases like IEDB, HLA-E was the only MHC class Ib molecule predicted to bind any EBV-derived peptides. While HLA-F appears to be strongly involved in the HIV immune response [323], HLA-G isoforms are supposed to have immunosuppressive properties by interacting with various inhibitory receptors, thereby hampering activities of cytolytic NK cells, CD8+ cytotoxic T cells, cytotoxic macrophages, allo-proliferative CD4+ T cells, and DCs [324]. It was therefore decided to not further assess the roles of non-classical HLA alleles other than *HLA-E*01:01* and *E*01:03*.

Another possible explanation for the observed C-TCR recognition pattern is that those receptors might recognize a target peptide which is canonically bound and presented by different MHC proteins. The phenomenon of overlapping binding specificities between HLA molecules was mainly reported for MHC proteins sharing similar binding motives, thereby being assigned to the same HLA supertype [325–

[328]. However, few studies also indicate shared antigen binding specificities across HLA supertypes and even across HLA loci [329,330]. For instance, Frahm and colleagues demonstrated a total of 117 different HIV- and 49 EBV-derived CTL epitopes to be recognized in the context of at least three different HLA class I alleles. They thereby demonstrated that promiscuous presentation of epitopes is not only restricted to few highly related HLA alleles but might rather be a widespread phenomenon [331]. While a more recent study conducted by Rao *et al.* indicates promiscuous MHC binding for more than 60% of all HIV- and EBV-derived HLA ligands [332], the extent to which this phenomenon also applies to MHC-dependent neopeptide presentation remains elusive. Nevertheless, it is widely accepted that TCRs can bind more than one peptide and that some tumor-enriched TCRs cross-react with pathogen-derived antigens, such as EBV peptides [220]. It could therefore be possible that C-TCRs specifically recognize an HLA-B*08:01-restricted tumor peptide in NCI-H1703 and K562 cells while, due to cross-reactivity, simultaneously binding an EBV-specific peptide that is presented by different MHC proteins. However, this hypothesis is highly speculative since none of those potentially recognized antigen peptides were identified during this study. Future cDNA library screenings could be used to further analyze the antigen specificity of C-TCRs and could finally verify that the applied clustering approach successfully grouped those TCRs based on similar antigen specificities.

Besides of the TCRs included in cluster C, an additional specificity cluster, referred to as TCR cluster G, of highly related TCRs was analyzed during this study. Cluster G comprised three NSCLC patients and a total 6 TCRs (G-TCRs), each of which shared TRAV- as well as TRBV-CDR3 regions not differing in more than one amino acid, respectively (Figure 3.4.6). Interestingly, CDR3 regions identical to two G-TCRs (TCRs 5094 and 1702) had previously been found by the group of Mark M. Davis (Stanford University, USA) via their GLIPH2 (Grouping of Lymphocyte Interactions with Paratope Hotspots version 2) algorithm aiming at the identification of TCRs with a shared antigen specificity. While analyzing 435 tumor-enriched specificity groups, they identified three different TCRs (TCRs 2, 18, 19) exhibiting a dedicated S%DGMNTE (% symbol referring to any possible AA) TRBV-CDR3 motif as well as a shared TRAV-CDR3 region (CAVLMDSNYQLIW), both of which are also present in most cluster G TCRs (Figure 3.4.6). Notably, while no identical G-TCR counterpart was identified for Stanford TCR 2, identical TRAV/TRBV-CDR3 sequences were found in Stanford TCR 18 and G-TCR 5094, as well as in Stanford TCR 19 and G-TCR 1702. Moreover, they demonstrated the three TCRs to bind an HLA-A*02:01-presented human antigen as well as epitopes arising from two common pathogens. More precisely, the three Stanford TCRs specifically recognized 9-mer peptides derived from the *E. coli* enterobactin exporter EntS9 (LLGGLLTMV), the latent membrane protein 2a from EBV (CLGGLLTMV), and the mammalian NSCLC TAA TMEM161A (ALGGLLTPL) [220].

The discovery made by the Davis group provided the unique chance of verifying the efficiency of the TCR clustering pipeline applied by HSD. Hence, TCR cluster G was used as a control for the performed *in silico* approach aiming at the identification of TCRs with shared antigen specificities. For this purpose, all G-TCRs were analyzed for the recognition of the 9-mer epitopes previously published by the Davis group. As a result, it was demonstrated that all G-TCRs shared identical peptide specificities and recognized the exact antigen epitopes previously reported for Stanford TCRs 2, 18, and 19 [220]. Consistent with their data, TMEM161A was found to be the weakest stimulator when compared to pathogen-derived EntS and LMP2 9-mer peptides (Figure 3.4.8). Moreover, G-TCR+ effector populations simultaneously recognized HLA-A*02:01+ K562 cells as well as several, but not all, HLA-A*02:01+ NSCLC cell lines (MZ-LC-16, NCI-H1703, NCI-H1792) with differing levels of reactivity (Figure 3.4.7). Since K562 is an EBV-negative human malignant myeloid cell line [333] and *E. coli* is not considered a classical respiratory antigen [334,335], these results indicate a broad expression of the non-mutated TAA TMEM161A not only in NSCLC but across various cancer cell lines. The latter is also consistent with the TMEM161A transcript expression data available in public database like *The human Protein Atlas* (see Table 5.8). Furthermore, RNA-Seq analyses of the recognized NSCLC cell lines did not reveal any EBV-transcripts at all (data not shown), thereby excluding the possibility that the EBV-peptide is responsible for the detected recognition of tumor cells.

Interestingly, the recognition of TAP-deficient T2 (HLA-A*02:01+) cells (Figure 3.4.7) suggested that at least one of the three recognized antigens might be presented independently from TAP peptide transporters. In fact, it was already shown in 1996 that the EBV-derived LMP2A 9-mer peptide (CLGGLLTMV) can be processed and presented through an alternative TAP-independent pathway and that LMP2A-specific CTL clones were able to efficiently recognize this epitope on TAP-deficient target cells in the context of HLA-A*02 [336]. It was therefore assumed that G-TCR+ effectors specifically recognized EBV-positive T2 cells due to the TAP-independent presentation of the LMP2A-derived 9-mer target peptide (CLGGLLTMV). In contrast, as described the previous paragraph, the recognition of all other tumor cell lines might rather be dependent on the presentation of the TMEM161A-derived target epitope (ALGGLLTPL). Therefore, the conducted study not only demonstrated the power of the applied approach to cluster TCRs of concurring antigen specificity based on shared CDR3 amino acid sequences, but also identified additional TCRs specifically recognizing the TAA TMEM161A.

By validating the antigen specificities of analyzed cluster G TCRs, this study facilitated the identification of additional TAA (TMEM161A) epitope-specific TCRs that are cross-reactive towards two pathogen-derived antigens. Together with the data published by the Davis group, these results support the assumption that T cell specificities for TAAs and pathogen-derived antigens are not mutually exclusive [220]. While often being referred to as pathogen-primed and tumor-unreactive bystander T cells, it is still unclear what roles these cross-reactive T cells might play in anti-tumoral immune responses and to which extent they are able to control tumor growth in humans [337,338]. Recent studies indicate that pathogens like EBV and influenza viruses can induce anti-tumor T cell responses against abnormally expressed cellular antigens functioning as TAAs [339,340]. Hence, despite having rather weak reactivities towards non-mutated TAAs, those cross-reactive T cells might play a crucial role in controlling cancer progression. This might especially be true for settings in which the anti-tumor response is unleashed due to checkpoint inhibition. Nevertheless, additional studies are needed to assess the causal relationships between microbe/TAA-cross-reactive T cells and clinical benefits potentially achieved upon checkpoint inhibition [220].

5) Materials

5.1 Cells and bacteria

K562

The K562 cell line was initially obtained from a patient suffering from chronic myelogenous leukemia (CML) in blast crisis. Due to their low levels of differentiation, the nature of K562 cells cannot be stated certainly. However, some characteristics indicate a non-lymphoid origin. In contrast, the cell line exhibits some T cell properties and basically no B lymphocyte features. K562 cells were found to be rounded, smooth, and apparently exhibit small numbers of short microvilli resembling to undifferentiated leukemic cells [341]. In the current study, K562 was used as a target cell line in the context of IFN-γ ELISpot assays. Hence, K562 cells were used for retroviral transductions with patient-specific HLA expression constructs and were subsequently pulsed with synthesized neoantigen peptides. Inherent K562 HLA haplotypes are listed in Table 5.1.

HEK 293T

HEK (Human embryonic kidney) 293 cells were established by transforming HEK cells via exposure to sheared fragments of adenovirus type 5 DNA. The resulting and loosely adherent HEK 293 cell line grossly exhibited epithelioid characteristics, swollen nuclei, and tended to grow in clumps after reaching confluency [342]. In contrast, the HEK 293T cell line was established by transfecting HEK 293 cells with the SV40 large T-antigen required for the replication of SV40 origin-containing plasmids [343]. The HEK 293T cells used for the current study had kindly been provided by [redacted] (The University of California, Berkeley, USA) and were used to assess TCR target antigen specificities. Here, HEK 293T cells were used as target cells in IFN-γ ELISpot assays. Hence, HEK 293T were transiently co-transfected with cDNA expression constructs encoding patient-specific HLA alleles, cancer germline antigens, and other putative neoantigens. Inherent HEK 293T HLA haplotypes are listed in Table 5.1.

Phoenix-Ampho

The Phoenix-Ampho retroviral packing cell line was used for the rapid generation of recombinant retroviral particles utilized for the transduction of dividing target cells. The Phoenix-Ampho cell line had initially been generated by stably transfecting HEK 293T cells with expression constructs encoding retroviral gag (group-specific antigen), pol (polymerase), and amphotropic env (envelope) proteins [344]. For this study, a transient three-plasmid expression system was used to optimize the production of retroviral particles by Phoenix-Ampho cells. Hence, Phoenix-Ampho cells were co-transfected with two vectors encoding *gag-pol* genes of the murine leukemia virus (MLV) and gibbon ape leukemia virus-derived envelope proteins [345,346]. Phoenix-Ampho cells had kindly been provided by [redacted] (University Medical Center Mainz, Mainz, Germany).

Human T cells

Primary human T cells were isolated from healthy donor buffy coat (BCs) samples kindly provided by [redacted] (Transfusion Center of the University Medical Center Mainz, Mainz, Germany). BCs are leukocyte concentrates generated as a byproduct during the manufacturing process of red blood cell and platelet concentrates from whole blood donations. Magnetic activated cell sorting (MACS) was used to isolate CD4+ and CD8+ T cells from buffy coat samples. Purified T cells were eventually used for the CRISPR/Cas9-mediated KO of their endogenous TCR expression and subsequent retroviral transduction with exogenous TCR expression constructs, thereby establishing TCR-transgenic T cell populations expressing potentially tumor-reactive TCRs.

NCI-H460

NCI-H460 is a human large cell lung cancer cell line initially isolated from the pleural fluid of a male patient in 1982 [347]. NCI-H460 cells exhibit homozygous $KRAS^{Q61H}$ mutations [348] and were therefore used as target cells in IFN-γ ELISpot and CD107a degranulation assays to verify antigen specificities of potentially $KRAS^{Q61H}$-specific TCRs. Since TCR-mediated recognition of mutated KRAS cDNA and corresponding 10-mer peptides was restricted to HLA-A*01:01, NCI-H460 cells were retrovirally transduced to overexpress the *HLA-A*01:01* allele. Additionally, NCI-H460 cells were used for the knock-in of the $KRAS^{Q61R}$ mutation via a CRISPR/Cas9-based homology-directed repair (HDR) approach. Furthermore, total RNA obtained from NCI-H460 cells was used for Gateway cloning of the p.Q61H-mutated KRAS open reading frame (ORF). NCI-H460 cells were kindly provided by [redacted] (University Medical Center Mainz, Mainz, Germany) and corresponding HLA haplotypes are listed in Table 5.1.

Bacteria

NEB 10-beta Competent *E. coli* (High Efficiency) bacteria (NEB, Ipswich, USA) were used for the amplification of plasmid cDNA. Hence, bacteria were transformed as described in 6.4.13. The used NEB 10-beta competent *E. coli* exhibited the following genotype: *Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14- ϕ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL* (Str^R) *rph spoT1 Δ(mrr-hsdRMS-mcrBC)*.

Table 5.1: HLA class I genotyping of target cell lines used for IFN-γ ELISpot assays

Multiple cancer cell lines, HEK 293T, and EBV-infected B cells were used as targets in IFN-γ ELISpot assays to assess the antigen specificities of established TCR-T cell cultures. Origins and HLA class I haplotypes are listed below for all target cells used during this study. Respective HLA class I alleles were obtained from the TRON (Translational Oncology, Mainz, Germany) Cell Line Portal (http://celllines.tron-mainz.de/) and from the ATCC database (www.atcc.org). HLA alleles listed in brackets were obtained from the TRON Cell Line Portal and differed from those listed in the ATCC database. HLA alleles revealed by the *Institut für Immunologie und Genetik* (Kaiserslautern, Germany) are labelled accordingly.

Cell line	Origin	HLA-A		HLA-B		HLA-C		Provided by
D04-Mel [a]	Mel	A*11:01	A*03:01	B*07:02	B*51:01	C*15:02	C*07:02	[redacted]
D05-Mel [a]	Mel	A*02 [nr]		B*27 [nr]	B*44 [nr]	C*02 [nr]	C*05 [nr]	[redacted] Q
D10-Mel [a]	Mel	A*01:01	A*31:01	B*08:01	B*40:01	C*03:04	C*07:01	[redacted] Q
D11-Mel [a]	Mel	A*34:01	A*32:02	B*15:21	B*44:02	C*04:03	C*05:01	[redacted] Q
D17-Mel [a]	Mel	A*01:01	A*32:01	B*08:01	B*15:01	C*03:03	C*07:01	[redacted] Q
D28-Mel [a]	Mel	A*01:01	A*02:01	B*08:01	B*51:01	C*07:01	C*14:02	[redacted] Q
D41-Mel [a]	Mel	A*01:01	A*02:01	B*08:01		C*07:01		[redacted] Q
HEK 293T [a]	HEK	A*02:01	A*03:01	B*07:02		C*07:02		[redacted] U
K562 [a]	CML	A*11 [nr]	A*31 [nr]	B*18 [nr]	B*40 [nr]	C*03 [nr]	C*05 [nr]	[redacted] M
MZ-LC-16 [a]	LC	A*02:01	A*11:01	B*07:02	B*14:02	C*07:02	C*08:02	[redacted] M

Cell line	Origin	HLA-A		HLA-B		HLA-C		Provided by
L721.221 [b]	EBV-B	Deficient for HLA-A, -B, and -C expression due to a deletion on chromosome 6						[redacted] 1 [M]
Ma-Mel-86a [a]	Mel	A*01:01	A*24:02	B*08:01	B*15:01	C*03:03	C*07:01	[redacted] [E]
MOR/ CPR	NSCLC	A*02:01		B*08:01	B*44:02	C*05:01	C*07:01	ECACC: 96042333
NCI-H1703	NSCLC (LUSC)	A*01:01	A*02:01	B*07:02	B*08:01	C*07:01 (C*07:06)	C*07:02 (C*07:06)	ATCC: CRL-5889
NCI-H1792	NSCLC (LUAD)	A*02:01	A*26:08	B*08:01	B*15:01	C*04:01	C*07:01	ATCC: CRL-5895
NCI-H1869	NSCLC	A*03:01	A*11:01	B*07:02	B*18:01	C*07:01	C*07:02	ATCC: CRL-5900
NCI-H2023	NSCLC (LUAD)	A*02:01		B*07:02	B*08:01	C*07:01 (C*07:06)	C*07:02 (C*07:06)	ATCC: CRL-5912
NCI-H2030	NSCLC	A*11:01	A*24:02	B*44:03	B*51:09	C*01:02	C*04:01	ATCC: CRL-5914
NCI-H226	NSCLC	A*01:01	A*24:02	B*07:02	B*39:01	C*07:02	C*12:03	ATCC: CRL-5826
NCI-H2342	NSCLC (LUAD)	A*02:01		B*08:01	B*18:01	C*05:01	C*07:01	ATCC: CRL-5941
NCI-H358	NSCLC	A*03:01		B*15:01	B*35:01	C*03:04	C*04:01	ATCC: CRL-5807
NCI-H460	NSCLC	A*24:02	A*68:01	B*35:01	B*51:01	C*03:03	C*15:02	[redacted] [M]
NCI-H661	NSCLC	A*02:01 (A*24:02)		B*08:01		C*07:01		ATCC: HTB-183
P18 lung/tumor [a]	NSCLC	A*01:01	A*02:01	B*08:01	B*40:02	C*03:04	C*07:01	HSD [H]
P43 lung/tumor [a]	NSCLC	A*01:01	A*23:01	B*08:01	B*44:03	C*04:01	C*07:01	HSD [H]
P50 lung/tumor [a]	NSCLC	A*03:01	A*29:02	B*35:01	B*44:03	C*04:01	C*16:01	HSD [H]

Abbr.: ATCC = American Type Culture Collection; CML = chronic myelogenous leukemia; EBV-B = EBV-infected B cells; ECACC = European Collection of Authenticated Cell Cultures; HEK = Human embryonic kidney; HSD = HS Diagnomics GmbH (Berlin, Germany); Mel = melanoma; LC= lung cancer; LUAD = lung adenocarcinoma; LUSC = lung squamous cell carcinoma; NSCLC = non-small cell lung cancer

Footnotes: [a] = HLA class I alleles were revealed by the Institut für Immunologie und Genetik (Kaiserslautern, Germany); [b] = HLA class I haplotypes were adapted from Shimizu and DeMars [322]; [E] = Department of Dermatology, University Medical Center Essen (Essen, Germany); [H] = HS Diagnomics (Berlin, Germany); [M] = University Medical Center Mainz (Mainz, Germany); [nr] = HLA class I alleles not completely resolved; [Q] = Queensland Institute of Medical Research (Brisbane, Australia); [U] = University of California (Berkeley, USA).

5.2 Materials and consumables

Table 5.2.1: Laboratory equipment

Laboratory equipment:	Manufacturer:
Agarose gel electrophoresis tank	PEQLAB (Erlangen, Germany)
Autoclave (VX-150)	Systec (Wettenberg, Germany)
Autoclave (DSL-20/9/9-2-FDGD)	Holzner (Nußloch, Germany)
Biological safety cabinet (HERAsafe EN12469)	Heraeus (Hanau, Germany)
Cell electroporation (Gene Pulser Xcell)	Bio-Rad GmbH (München, Germany)
Cell nucleofection (Nucleofector 2b)	Lonza (Basel, Switzerland)
Cell radiation (Gammacell 2000)	Molsgaarg Medical (Gansloe, Denmark)
Centrifuge (Multifuge Rotina 420)	Hettich (Tuttlingen, Germany)
Centrifuge (5424; 5415 R)	Eppendorf (Hamburg, Germany)
CTL ImmunoSpot S5 Micro Analyzer	C.T.L. Europe GmbH (Bonn, Germany)
CO_2 Incubator (BBD6220)	Heraeus (Hanau, Germany)
Deep freezer -80°C (HERAfreeze HLE series)	Heraeus (Hanau, Germany)
Digital Graphic Printer UP-D890	Biometra, Sony (Göttingen, Germany)
Drying cupboard (UT6420)	Heraeus (Hanau, Germany)
Electrophoresis power supply (GenePower Supply GPS200/400)	Pharmacia (Uppsala, Schweden)
ErgoOne multichannel pipettes	Starlab (Ahrensburg, Germany)
ErgoOne pipettes	Starlab (Ahrensburg, Germany)
FACS (BD Canto II)	BD Biosciences (Heidelberg, Germany)
Thermomixer comfort 5436	Eppendorf (Hamburg, Germany)
Laboratory balance	Sartorius (Göttingen, Germany)
Magnetic particle concentrator (magnetic rack)	Dynal A.S. (Oslo, Norway)
Microscope (Eclipse TS100)	Nikon (Alzenau, Germany)
Microwave (MWS 2819)	Bauknecht (Stuttgart, Germany)
Mini table centrifuge	Thermo Fisher Scientific (Waltham, USA)
Mr. Frosty Freezing Container	Thermo Fisher Scientific (Waltham, USA)
mLINE multichannel pipette	BIOHIT (Helsinki, Finnland)
Pipetman G piston pipettes	Gilson (Middleton, USA)
Purelab Classic ELGA (water purification system)	Elga Labwater (Celle, Germany)
Qubit 2.0 Fluorometer	Thermo Fisher Scientific (Waltham, USA)
Shaking incubator (CERTOMAT H und SII)	Sartorius (Göttingen, Germany)
Spectrophotometer (Ultrospec 3000 UV)	Pharmacia (Freiburg, Germany)
TGradient Thermocycler	Biometra (Göttingen, Germany)
UVsolo TS Gel Imaging System	Biometra (Göttingen, Germany)
VF2 vortex	Janke & Kunkel GmbH & Co. KG IKA Labortechnik (Staufen, Germany)
Waterbath (UT 6420)	Heraeus (Hanau, Germany)
XL-180 liquid nitrogen container	Taylor-Wharton (Husum, Germany)

Table 5.2.2: Consumption items

Consumption items:	Manufacturer:
Disposable inoculation loop 10 µl	Thermo Fisher Scientific (Schwerte, Germany)
Cell culture flasks (Cellstar 175, 75, 25 cm^2)	Greiner Bio-One GmbH (Frickenhausen, Germany)
Cell culture plates (Cellstar 96-, 48-, 24-, 6-well)	Greiner Bio-One GmbH (Frickenhausen, Germany)
Cryovials (Cellstar Cryo-S™ 2 ml)	Greiner Bio-One GmbH (Frickenhausen, Germany)
Electroporation cuvettes (2 mm)	VWR International GmbH (Darmstadt, Germany)
ELISpot plates (MultiScreen Filterplates)	Merck Millipore (Burlington, USA)
Eppendorf tubes	Eppendorf (Wesseling, Germany)
Erlenmeyer flasks (500-1000 ml)	Schott AG (Mainz, Germany)
Falcon tubes (15 ml, 50 ml)	Greiner Bio-One GmbH (Frickenhausen, Germany)
Filter attachments (0.22 µm, 0.45 µm)	Whatman GmbH (Dassel, Germany)
Flow cytometry tubes	Sarstedt (Nümbrecht, Germany)
Fuchs-Rosenthal counting chamber (0.2 mm)	Paul Marienfeld GmbH & Co. KG (Lauda-Königshofen, Germany)
Glass beakers	Schott AG (Mainz, Germany)
Serological pipettes (Cellstar 2-50 ml)	Greiner Bio-One GmbH (Frickenhausen, Germany)
MACS Magnetic MiniMACS Separator	Miltenyi Biotec GmbH (Bergisch Gladbach, Germany)
MACS separation columns with attached SmartStrainers (30 µm)	Miltenyi Biotec GmbH (Bergisch Gladbach, Germany)
Nucleofection cuvettes	Molecular Bioproducts (San Diego, USA)
Petri Dishs (30 mm, 90mm)	Greiner Bio-One GmbH (Frickenhausen, Germany)
Plastic beakers	Vitlab (Grossostheim, Germany)
Quartz glass cuvettes SUPRASIL	Hellma GmbH & Co. KG (Müllheim, Germany)
Scalpel	pfm medical (Köln, Germany)
Sempercare® Latex premium gloves	Semperit Technische Produkte GmbH (Wien, Österreich)
TipOne pipette tips	Starlab GmbH (Hamburg, Germany)

Table 5.2.3: Buffers, chemicals, cell culture media, and molecular biology reagents

Buffers, chemicals, cell culture media, and molecular biology reagents:	Manufacturer:
1 kb DNA Ladder	NEB (Frankfurt am Main, Germany)
100 bp DNA Ladder	NEB (Frankfurt am Main, Germany)
ABC Kit	Vectastain (San Francisco, USA)
AEC (3-Amino-9-Ethylcarbazole)	Carl Roth AG (Karlsruhe, Germany)
Agarose	Serva Serva Electrophoresis GmbH (Berlin, Germany)
Aqua ad injectabilia	B. Braun Melsungen AG (Melsungen, Germany)
Bacto Agar	BD Biosciences (Heidelberg, Germany)
Bovine Serum Albumin (BSA)	Sigma-Aldrich Chemie GmbH (Steinheim, Germany)
Chloroform	Riedel-de Haen (Selze, Germany)
Deoxynucleotide triphosphates Solution Mix (10 mM)	Carl Roth AG (Karlsruhe, Germany)

Buffers, chemicals, cell culture media, and molecular biology reagents:	Manufacturer:
Dimethylformamide	Merck KGaA (Darmstadt, Germany)
Dimethyl sulfoxide (DMSO)	Merck KGaA (Darmstadt, Germany)
Dulbecco's Modified Eagle's Medium (DMEM)	Gibco, Life Technologies Limited (Paisley, United Kingdom)
Dulbecco´s Phosphate Buffered Saline (PBS)	Sigma-Aldrich Chemie GmbH (Steinheim, Germany)
Ethanol (99%)	AppliChem GmbH (Darmstadt, Germany)
Ethylenediaminetetraacetic acid (EDTA)	Merck KGaA (Darmstadt, Germany)
RiboRuler High Range RNA Ladder	Thermo Fisher Scientific (Waltham, USA)
FACSClean, FACSFlow, FACSRinse	BD Biosciences (Heidelberg, Germany)
Fetal calf serum (FCS)	Sigma-Aldrich Chemie GmbH (Steinheim, Germany)
Formaldehyde (37%)	Merck KGaA (Darmstadt, Germany)
FuGENE 6 Transfection Reagent	Promega (Mannheim, Germany)
Gel Loading Dye Purple (6x)	NEB (Frankfurt am Main, Germany)
GelRed Nucleic Acid Gel Stain (10000x)	Biotium (Hayward, USA)
Glycerol	Sigma-Aldrich Chemie GmbH (Steinheim, Germany)
Histopaque-1077 (Ficoll)	Sigma-Aldrich Chemie GmbH (Steinheim, Germany)
Human serum (HS)	Transfusion Center (UMC, Mainz, Germany)
Hexadimethrine bromide (Polybrene)	Sigma-Aldrich Chemie GmbH (Steinheim, Germany)
HEPES	Sigma-Aldrich Chemie GmbH (Steinheim, Germany)
Hydrogen peroxide (30%)	Carl Roth AG (Karlsruhe, Germany)
Isopropanol (99%)	Thermo Fisher Scientific (Waltham, USA)
L-Glutamine	Life Technologies (Carlsbad, USA)
Lipofectamine 2000 Transfection Reagent	Life Technologies (Carlsbad, USA)
Panserin 413	PAN-Biotech GmbH (Aidenbach, Germany)
PBS powder	InstaMed (Philadelphia, USA)
peqGREEN (20000x)	PEQLAB (Erlangen, Germany)
RPMI 1640	Gibco, Life Technologies Limited (Paisley, United Kingdom)
SOC Outgrowth Medium	Life Technologies (Carlsbad, USA)
Sodium acetate	Merck KGaA (Darmstadt, Germany)
Sodium azide	Merck KGaA (Darmstadt, Germany)
Sodium bicarbonate	Merck KGaA (Darmstadt, Germany)
Sodium chloride	Carl Roth AG (Karlsruhe, Germany)
Sodium hydroxide	Merck KGaA (Darmstadt, Germany)
TAE Buffer (50x)	Serva Electrophoresis GmbH (Berlin, Germany)
TBE Buffer (10x)	Serva Electrophoresis GmbH (Berlin, Germany)
Terralineliquid disinfectant	Schülke&Mayr GmbH (Norderstedt, Germany)
Trypan blue	Merck KGaA (Darmstadt, Germany)
Trypsin-EDTA	Life Technologies (Carlsbad, USA)
Tween 20	Sigma-Aldrich Chemie GmbH (Steinheim, Germany)

Table 5.2.4: Antibiotics

Antibiotics	Used concentration	Manufacturer
Ampicillin	100 µg/ml	Sigma-Aldrich Chemie (Steinheim, Germany)
Kanamycin	50 µg/ml	Sigma-Aldrich Chemie (Steinheim, Germany)
Penicillin-Streptomycin	500 µg/ml	Life Technologies (Carlsbad, USA)
Puromycin	2 µg/ml	Carl Roth AG (Karlsruhe, Germany)

Table 5.2.5: Cytokines

Cytokines	Used concentration	Manufacturer
IL-2	Cell culture: 600 U/ml IFN-γ ELISpot assay: 250 U/ml	R&D Systems (Minneapolis, USA)

5.3 Molecular compositions of buffers, chemical solutions, and cell culture media

Acetate buffer

Acetic acid (0.2 N)	4.6 ml
Sodium acetate (0.2 N)	11 ml
d.H_2O	46.9 ml

Acetic acid (0.2 N)

Acetic acid (96%)	11.3 g
d.H_2O	ad. 1 l

Panserin 413 complete T cell medium
10% Human serum (heat-inactivated, pool obtained from healthy donors)
1% Penicillin-Streptomycin

Avidin-peroxidase complex (10 ml is sufficient for 1 MultiScreen Filterplate)
10 ml PBS (1x)
0.1% Tween® 20
1 drop of solution A: ABC Kit
1 drop of solution B: ABC Kit

DMEM complete Phoenix-Ampho medium
2 mM L-Glutamine
25 mM HEPES
1% Penicillin-Streptomycin
10% FCS (heat-inactivated)

ELISpot buffer

$NaHCO_3$	2.9 g
Na_2CO_3	1.6 g
NaN_3	0.2% (w/v)
d.H_2O	ad. 1 l

ELISpot wash solution
0.05% Tween 20
in PBS (1x)

Erythrocyte lysis buffer (pH = 7.3)

NH_4Cl	8.29 g
$KHCO_3$	1 g
EDTA	0.0372 g
H_2O	ad. 1 l

FACS buffer
0.1% BSA
2 mM EDTA
in PBS (1x)

FACS fixation buffer
1% Formalin
2 mM EDTA
in PBS (1x)

Freezing medium (bacteria): 15% glycerol solution

Freezing medium (eukaryotic cells): 10% DMSO in FCS (heat-inactivated)

TAE buffer (1%)

TAE (50x)	2 ml
d.H_2O	ad. 1 l

TBE buffer (0.5%)

TBE (10x)	50 ml
d.H_2O	ad. 1 l

DEPC-treated water

DEPC (Diethylpyrocarbonate)	1 ml
d.H_2O	1 l

LB medium (pH = 7, autoclaved)

LB medium was supplemented 1/1000 with kanamycin or ampicillin

Bacto Trypton	10g
yeast extract	5g
NaCl	10
d.H_2O	ad. 1 l

MACS buffer (sterile filtered with 0.22 µm filter attachments)

2 mM EDTA
0.5% BSA
in PBS (1x)

MOPS buffer (10x)

200 mM MOPS
50 mM Na-Acetate
5 mM EDTA

PBS (1x)

PBS powder	95.5 g
d.H_2O	ad. 10 l

Peroxidase substrate solution (filtered with 0.45 µm filter-attachments)

1 pellet AEC dissolved in:

Dimethyformamide	2.5 ml
Acetate buffer	ad. 50 ml

RPMI medium

no supplementation

RPMI complete cell culture medium

10% FCS (heat-inactivated)
1% Penicillin-Streptomycin

RPMI recovery cell culture medium

10% FCS (heat-inactivated)
1% Penicillin-Streptomycin
25 mM HEPES (pH 7.2)
50 mM 2-Mercaptoethanol

Trypan Blue

3x Solution A	0.2% Trypan blue solution (1 g Trypan blue in 500 ml d.H_2O)
1x Solution B	4.5% NaCl (22.5 g NaCl in 500 ml d.H_2O)

5.4 Enzymes and buffers

Table 5.4: Used enzymes, enzyme buffers, and corresponding experimental approaches

Enzyme	Concentration	Buffer	Experiment	Manufacturer (enzyme and buffer)
BP Clonase II enzyme mix	5x	TE buffer	Gateway Cloning	Invitrogen (Carlsbad, USA)
DNase I	2727 U/ml (stock solution)	RDD buffer	total RNA isolation	Thermo Fisher Scientific (Waltham, USA)
DpnI	20000 U/ml	T4 ligase buffer	KLD-reaction	NEB (Frankfurt am Main, Germany)
LR Clonase II enzyme mix	5x	TE buffer	Gateway Cloning	Invitrogen (Carlsbad, USA)
One Taq Master Mix	2x	Standard buffer	Colony PCR	NEB (Frankfurt am Main, Germany)
Q5 High Fidelity DNA Polymerase	2000 U/ml	Q5 Reaction Buffer	PCR	NEB (Frankfurt am Main, Germany)
PrimeScript RTase	N/A	5X PrimeScript Buffer	cDNA synthesis	Takara BIO INC. (Kusatsu, Japan)
Proteinase K	2 µg/µl	N/A	Gateway Cloning	Invitrogen (Carlsbad, USA)
T4 DNA polynucleotide Kinase	10000 U/ml	T4 ligase buffer	KLD-reaction	NEB (Frankfurt am Main, Germany)
T4 DNA Ligase	400000 U/ml	T4 ligase buffer	KLD-reaction	NEB (Frankfurt am Main, Germany)

5.5 Vectors

Table 5.5: Vectors used for Gateway cloning, retroviral transduction, and transient transfection

Vectors	Sequencing primer	Antibiotic resistance	Experiment	Included ORFs
pcDNA3.1_Dest#6 [L] (abbr.: pcDNA3.1)	T7.for and BGH.rev	Bacteria: Ampicillin Cell culture: Neomycin	Destination vector for Gateway Cloning comprising expression constructs for transient cell transfection	TSAs, TAAs, EGFP
pCOLT-GALV [T]	N/A	N/A	Helper plasmid for virion assembly used during retroviral cell transduction	GALV-derived envelope protein
pDONR-221-P1-P2 [E] (abbr.: pDONR)	M13.for. and M13.rev.	Bacteria: Kanamycin	Donor vector for Gateway Cloning and template plasmid for SDM reactions	cDNA fragments to be shuttled into destination vectors
pHIT60 [T]	N/A	N/A	Helper plasmid for virion assembly used during retroviral cell transduction	MLV-derived *gag-pol* genes
pMX-puro.Dest#35 [E] (abbr.: pMX_puro)	Seq.206.for. and pMX_puro-p2086.rev.	Bacteria: Ampicillin Cell culture: Puromycin	Destination vector for Gateway Cloning comprising expression constructs for stable cell transduction	Patient-derived HLAs, TCR expression constructs
pTwist CMV BetaGlobin WPRE Neo [S] (abbr.: pTwist_neo)	pTwist_Seq.for	Bacteria: Ampicillin Cell culture: Neomycin	Includes TMG expression constructs for transient cell transfection	TMG constructs

Abbr.: cDNA = complementary DNA, CMV = cytomegalovirus, Dest = destination, EGFP = enhanced green fluorescent protein, for = forward primer, gag = group antigens, GALV = gibbon ape leukemia virus, HLA = human leukocyte antigen, MLV = murine leukemia virus, Neo = neomycin, pol = polymerase, puro = puromycin, rev = reverse primer, Seq = sequencing, TAA = tumor-associated antigen, TCR = T cell receptor, TMG = tandem minigene, TSA = tumor-specific antigen, T7 = T7 promotor, WPRE = woodchuck hepatitis virus post-transcriptional regulatory element.

Footnotes: [E] = original vectors were obtained from [redacted] (University Medical Center, Mainz, Germany) and had then been modified by [redacted] (University Medical Center, Mainz, Germany), [L] = original vectors were obtained from Life Technologies (Carlsbad, USA) and had then been modified by [redacted] (University Medical Center, Mainz, Germany), [S] = vectors were obtained from Twist Bioscience (South San Francisco, USA), [T] = kindly provided by [redacted] (University Medical Center, Mainz, Germany).

5.6 Oligonucleotides

Table 5.6.1: Sequencing primers

Primer	Nucleotide sequence (5' → 3')	Melting temperature	Experiment
BGH.rev	TAGAAGGCACAGTCCAGG	53.7 °C	Sequencing of pcDNA3.1
M13.for	GTAAAACGACGGCCAGT	52.6 °C	Sequencing of pDONR
M13.rev	CAGGAAACAGCTATGACCATG	53.5 °C	Sequencing of pDONR
pTwist_Seq.for	CCCTTTTGCTAATCATGTTCATACC	54.3 °C	Sequencing of pTwist_Neo
Seq.206.for	TTACACAGTCCTGCTGACCACC	59.0 °C	Sequencing of pMX_puro
Seq.2086.rev	GCCCTCACATTGCCAAAAGACG	59.3 °C	Sequencing of pMX_puro
T7.for	TAATACGACTCACTATAGGG	47.5 °C	Sequencing of pcDNA3.1

Abbr.: for = forward primer, rev = reverse primer, Seq = sequencing, puro = puromycin, Neo = neomycin.

Table 5.6.2: Gateway cloning primers

AttB1/attB2 recombination site sequences are italicized while Kozak consensus sequences are underlined.

Primer	Nucleotide sequence (5' → 3')	Melting temperatures	Used for cloning of:
attB1-Koz-START-CMV_IE2.for	*GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC* GCC ACC ATG GAG CAC ACC ATG CCC	TM_1 = 58.4 °C TM_2 = 72.4 °C	CMV-IE2
attB1-Koz-START-HLA-A*01:01:01.for	*GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC* GCC ACC ATG GCC GTC ATG GCG CCC	TM_1 = 64.5 °C TM_2 = 73.8 °C	HLA-A*01:01:01, HLA-A*03:01:01, HLA-A*23:01:01, HLA-A*29:02:01
attB1-Koz-START-HLA-B*08:01:01.for	*GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC* GCC ACC ATG CTG GTC ATG GCG CCC CGA	TM_1 = 66.6 °C TM_2 = 74.1 °C	HLA-B*08:01:01
attB1-Koz-START-HLA-B*40:02:01.for	*GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC* GCC ACC ATG CGG GTC ACG GCG CCC	TM_1 = 67.1 °C TM_2 = 74.4 °C	HLA-B*40:02:01, HLA-B*44:03:01
attB1-Koz-START-HLA-Cw*03:04:01.for	*GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC* GCC ACC ATG CGG GTC ATG GCG CCC	TM_1 = 64.5 °C TM_2 = 73.8 °C	HLA-C*03:04:01, HLA-C*04:01:01, HLA-C*07:01:01
attB1-Koz-START-HLA-E*01:03.for	*GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC* GCC ACC ATG GTA GAT GGA ACC CTC CTT TTA CTC C	TM_1 = 59.4 °C TM_2 = 70.8 °C	HLA-E*01:03
attB1-Kos.KRAS.CDSp1.for	*GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC* GCC ACC ATG ACT GAA TAT AAA CTT GTG GTA GTT GG	TM_1 = 55.2 °C TM_2 = 69.2 °C	KRAS_WT, KRAS_p.Q61H
attB2-STOP-CMV_IE2.rev	*GGGG AC CAC TTT GTA CAA GAA AGC TGG GTA* TTA CTG AGA CTT GTT TCT CAG GTC CTG G	TM_1 = 59.5 °C TM_2 = 69.0 °C	CMV-IE2
attB2-STOP-HLA-A*01:01:01.rev	*GGGG AC CAC TTT GTA CAA GAA AGC TGG GTA* TCA CAC TTT ACA AGC TGT GAG AGA CAC ATC AGA GC	TM_1 = 63.3 °C TM_2 = 69.7 °C	HLA-A*01:01:01, HLA-A*23:01:01
attB2-STOP-HLA-A*03:01:01.rev	*GGG GAC CAC TTT GTA CAA GAA AGC TGG GTA* TCA CAC TTT ACA AGC TGT GAG GGA CAC ATC AGA GC	TM_1 = 64.7 °C TM_2 = 70.3 °C	HLA-A*03:01:01

Primer	Nucleotide sequence (5' → 3')	Melting temperatures	Used for cloning of:
attB2-STOP-HLA-A*29:02:01.rev	*GGG GAC CAC TTT GTA CAA GAA AGC TGG GTA* TCA CAC TTT ACA AGC TGT GAG AGA CAT ATC AGA GCC C	TM_1 = 63.9 °C TM_2 = 69.8 °C	HLA-A*29:02:01
attB2-STOP-HLA-B*08:01:01.rev	*GGGG AC CAC TTT GTA CAA GAA AGC TGG GTA* TCA AGC TGT GAG AGA CAC ATC AGA GCC CTG G	TM_1 = 64.4 °C TM_2 = 70.5 °C	HLA-B*08:01:01
attB2-STOP-HLA-B*40:02:01.rev	*GGGG AC CAC TTT GTA CAA GAA AGC TGG GTA* TCA AGC TGT GAG AGA CAC ATC AGA GCC CTG GG	TM_1 = 66.6 °C TM_2 = 71.3 °C	HLA-B*40:02:01, HLA-B*44:03:01
attB2-STOP-HLA-Cw*03:04:01.rev	*GGGG AC CAC TTT GTA CAA GAA AGC TGG GTA* TCA GGC TTT ACA AGC GAT GAG AGA CTC ATC AGA GC	TM_1 = 64.3 °C TM_2 = 70.1 °C	HLA-C*03:04:01, HLA-C*04:01:01, HLA-C*07:01:01
attB2-STOP- HLA-E*01:03.rev	*GGGG AC CAC TTT GTA CAA GAA AGC TGG GTA* TTA CAA GCT GTG AGA CTC AGA CCC C	TM_1 = 60.1 °C TM_2 = 69.6 °C	HLA-E*01:03
attB2-KRAS.CDS_TAA.p570.rev	*GGGG AC CAC TTT GTA CAA GAA AGC TGG GTA* TTA CAT TAT AAT GCA TTT TTT AAT TTT CAC ACA GC	TM_1 = 55.0 °C TM_2 = 66.4 °C	KRAS_WT, KRAS_p.Q61H

Abbr.: CMV = cytomegalovirus, for = forward primer, p. = amino acid position; pep = peptide, H = histidine, HLA = human leukocyte antigen, Q = glutamine; rev = reverse primer, TM_1 = melting temperature for sequence-specific flanking with attB1/B2 recombination sites, TM_2 = melting temperature for full-length primer binding, WT = wild-type.

Table 5.6.3: SDM primers

Primer	Nucleotide sequence (5' → 3')	Melting temperature	SDM was performed to establish:
KRAS_SDM_p.183K.for	CGACACAGCAGGTAAAGAGGAGTACAG TGC	63.1 °C	KRAS_p.Q61K
KRAS_SDM_p.183L.for	CGACACAGCAGGTCTAGAGGAGTACAG TGC	64.1 °C	KRAS_p.Q61L
KRAS_SDM_p.183R.for	CGACACAGCAGGTCGAGAGGAGTACAG TGC	65.9 °C	KRAS_p.Q61R
KRAS_SDM.rev	AGAATATCCAAGAGACAGGTTTCTCCA TCAATTACTAC	60.0 °C	KRAS_p.Q61K; KRAS_p.Q61L; KRAS_p.Q61R
SDM_HLA-E*01:01.for	GGCCCGACAGGCGCTTCC	64.9 °C	HLA-E*01:01
SDM_HLA-E*01:01.rev	CCAGCTCGCAGCCATGC	60.3 °C	HLA-E*01:01

Abbr.: for = forward primer, HLA = human leukocyte antigen, K = lysine, L = leucine, R = arginine, rev = reverse primer, SDM = site-directed mutagenesis.

Table 5.6.4: CRISPR/Cas9 oligonucleotides

Oligo	Oligo type	Nucleotide sequence (5' → 3')	Experiment
crRNA#4	crRNA	AltR1 UCAGGGUUCUGGAUAUCUGUGU UUUAGAGCUAUGCU AltR2	Endogenous TCR-KO: crRNA targeting human TRAC
crRNA.TRBC1/2.KNK	crRNA	AltR1 CAAACACAGCGACCUCGGGUGUU UUAGAGCUAUGCU AltR2	Endogenous TCR-KO: crRNA targeting human TRBC

Oligo	Oligo type	Nucleotide sequence (5' → 3')	Experiment
Hs.Cas9.KRAS.1.AA	crRNA	AltR1 UUGGAUAUUCUCGACACAGCGUU UUAGAGCUAUGCU AltR2	crRNA targeting the KRAS ORF for the knock-in of KRAS_p.Q61R in NCI-H460 cells
KRAS-KI_ssODN_p.Q61R	ssODN for HDR	Alt-R1-HDR1 T*A*ATTGATGGAGAAACCTGTCT CTTAGATATTCTCGATACCGCAGG TCGCGAGGAGTACAGTGCAATGA GGGACCAGTACATGAGGAC*T*G HDR2-Alt-R2	ssODN used for the knock-in of KRAS_p.Q61R in NCI-H460 cells

Abbr.: Alt-R1/2 = IDT-proprietary end-blocking groups for reduced degradation, crRNA = CRISPR-RNA, H = histidine, HDR = homology directed repair, HDR1/2 = homology arms (40-46 nucleotides), K = lysine, KI = knock-in, KNK = designed by Korbinian N. Kropp, KO = knock-out, L = leucine, ODN = oligodeoxynucleotides, oligo = oligonucleotide; ORF = open reading frame, Q = glutamine, R = arginine, RNA = ribonucleic acid, ss = single strand, TCR= T cell receptor, TRAC = TCR alpha constant region, TRBC = TCR beta constant region.

*Footnotes: * = phosphorothioate bond.*

5.7 Antibodies

Table 5.7.1: Non-labelled ELISpot antibodies

Antibody/specificity	Used concentration	Manufacturer	Article number	Experiment
anti-human IFN-γ mAb 1-D1K	10 µg/ml	Mabtech AB (Nacka, Sweden)	#3420-3-1000	ELISpot primary Ab
anti-human IFN-γ mAb 7-B6-1; biotinylated	2 µg/ml	Mabtech AB (Nacka, Sweden)	#3420-6-1000	ELISpot secondary Ab
MA2.1 (anti-HLA-A*02)	400 µg/ml	ATCC (Wesel, Germany)	purified hybridoma supernatant	Blocking of target cell recognition
OKT-3 (anti-CD3)	400 ng/ml	ATCC (Wesel, Germany)	purified hybridoma supernatant	Non-specific T cell activation
W6/32 (anti-HLA class I)	200 µg/ml	Provided by [redacted] [1]	purified hybridoma supernatant	Blocking of target cell recognition

Abbr.: Ab = antibody, HLA = human leukocyte antigen, mAb = monoclonal antibody.
Footnotes: [1] = Stanford University School of Medicine (Stanford, USA).

Table 5.7.2: Flow cytometry antibodies

All listed antibodies were directly added to cell pellets and were afterwards incubated at 4°C for 20-25 min.

Antibody/specificity	Dye	Used volume/dilution per staining	Manufacturer	Article number
Goat Anti-Mouse (AffiniPure F(ab)2 Fragment) [b]	FITC	1/100 [1]	Jackson ImmunoResearch (Ely, UK)	#115-096-003
anti-human HLA class I: A*01, A*11, A*26 [a]	N/A	1/50 [1]	One Lambda Inc. (Los Angeles, USA)	#0544HA
anti-human CD3	FITC	4 µl	Beckmann Coulter (Brea, USA)	#A07746
anti-human CD4	PE	4 µl	Beckmann Coulter (Brea, USA)	#A07750

Antibody/specificity	Dye	Used volume/dilution per staining	Manufacturer	Article number
anti-human CD8	PE	5 µl	BD Biosciences (Heidelberg, Germany)	#340046
anti-human CD16	FITC	4 µl	Beckmann Coulter (Brea, USA)	#B49215
anti-human CD56	PE	4 µl	Beckmann Coulter (Brea, USA)	#A07788
anti-human CD107a	PE-Cy5	4 µ	BD Biosciences (Heidelberg, Germany)	#555802
TCR PAN αβ (anti-human TRAC)	APC	1 µl	Beckmann Coulter (Brea, USA)	#B13981
IgG1 isotype control	APC	1 µl	BD Biosciences (Heidelberg, Germany)	#555751
IgG1 isotype control	FITC/PE	4 µl	Beckmann Coulter (Brea, USA)	#A07794
TCR α/β (anti-murine TRAC)	FITC	1 µl	OriGene Technologies, Inc. (Rockville, USA)	#CL075F
W6/32 [a] (anti-HLA class I)	N/A	100 µl	provided by [redacted] [2]	hybridoma SN

Abbr.: APC = Allophycocyanin, FITC = Fluorescein isothiocyanate, HLA = human leukocyte antigen, PE = Phycoerythrin, SN = supernatant, TRAC = TCR alpha chain constant region, TRBC = TCR beta chain constant region.

Footnotes: [a] = non-labelled primary antibody used together with dye-labelled secondary antibodies, [b] = dye-labelled secondary antibody, [1] = 1 µl was diluted in RPMI (10% FCS) before adding onto cell pellets, [2] = Stanford University School of Medicine (Stanford, USA).

5.8 Software

Table 5.8: Software and web tools

Software and web tools	Publishers and internet domains
Alt-R CRISPR HDR Design Tool	Integrated DNA Technologies (Coralville, USA): https://eu.idtdna.com/pages/tools/alt-r-crispr-hdr-design-tool
Alt-R Custom Cas9 crRNA Design Tool	Integrated DNA Technologies (Coralville, USA): https://eu.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM
CLC Genomics Workbench; versions 20-23	QIAGEN (Hilden, Germany)
COSMIC	Wellcome Sanger Institute (Hinxton, UK): https://cancer.sanger.ac.uk/cosmic
Ensembl genome browser 105	EMBL-EBI (Hinxton, UK): https://www.ensembl.org/index.html
FACSDiva version 9.0	BD Biosciences (Heidelberg Germany)
FlowJo; version 7.6.5	FlowJo LLC (Ashland, USA)
Geneious; version 8.1.9	Biomatters Inc. (Auckland, New Zealand)
Graphpad Prism; version 9.3.1	Graphpad Holdings, LLC (San Diego, USA)
Immune Epitope Database (IEDB)	NIAID (Maryland, USA): https://www.iedb.org/home_v3.php
IEDB MHC-I binding prediction tool; version 2.24	NIAID (Maryland, USA): http://tools.iedb.org/mhci/
ImmunoSpot Software; version 7.0.15.1	Cellular Technology Limited (Cleveland, USA)
McPAS-TCR	Friedmann Lab at the Weizmann Institute of Science (Rechovot, Israel): http://friedmanlab.weizmann.ac.il/McPAS-TCR/
Microsoft Office Professional Plus 2016	Microsoft (Redmond, USA)

Software and web tools	Publishers and internet domains
NCBI BLAST	NCBI (Bethesda, USA): https://blast.ncbi.nlm.nih.gov/Blast.cgi
NCBI Database	NCBI (Bethesda, USA): https://www.ncbi.nlm.nih.gov/gene
NetChop-3.1	CBS Technical University of Denmark (Copenhagen, Denmark): https://services.healthtech.dtu.dk/service.php?NetChop-3.1
NetMHC-4.0	CBS Technical University of Denmark (Copenhagen, Denmark) https://services.healthtech.dtu.dk/service.php?NetMHC-4.0
NetMHCcons-1.1	CBS Technical University of Denmark (Copenhagen, Denmark): https://services.healthtech.dtu.dk/service.php?NetMHCcons-1.1
NetMHCpan-4.0	CBS Technical University of Denmark (Copenhagen, Denmark): https://services.healthtech.dtu.dk/services/NetMHCpan-4.0/
NetMHCstabpan-1.0	CBS Technical University of Denmark (Copenhagen, Denmark): https://services.healthtech.dtu.dk/services/NetMHCstabpan-1.0/
Oligo Analysis Tool	Integrated DNA Technologies (Coralville, USA): https://eu.idtdna.com/calc/analyzer
TCGA	National Cancer institute (Bethesda USA): https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga
The Human Protein Atlas	proteinatlas.org [349] *TM161A* expression analysis (02.12.2023): https://www.proteinatlas.org/ENSG00000064545-TMEM161A/cell+line#lung_cancer
UVsolo TS Imaging System image acquisition software	Biometra (Göttingen, Germany)
VDJdb	https://vdjdb.cdr3.net/ [350] :

Abbr.: BLAST = Basic Local Alignment Search Tool, CBS = Center for Biological Sequence Analysis, COSMIC = Catalogue of Somatic Mutations in Cancer, CRISPR = Clustered regularly interspaced short palindromic repeat, crRNA = CRISPR-RNA, EMBL-EBI = European Molecular Biology Laboratory's European Bioinformatics Institute, FACS = Fluorescence activated cell sorting, HDR = homology-directed repair, IDT = Integrated DNA Technologies, IEDB = The Immune Epitope Database, MHC = Major histocompatibility complex, NCBI = National Center for Biotechnology Information, NIAID = National Institute of Allergy and Infectious Diseases, RNA = ribonucleic acid, TCGA = The Cancer Genome Atlas, TS = Touch screen, UV = ultraviolet.

6) Methods

6.1 Cell culture and bacteria

6.1.1 Tumor cell culture

6.1.1.1 Non-adherent tumor cells

Non-adherent tumor cells were maintained in an incubator at 37 °C, 5% CO_2, 95% air humidity, and were always kept in a sterile environment. Depending on cell numbers, tumor cell lines were cultivated in 25 cm^2, 75 cm^2, or 175 cm^2 cell culture flasks containing 5-60 ml of RPMI complete medium and were, if necessary, grown under antibiotic selection (see Table 5.2.4). Tumor cells were passaged and split 1:10 as soon as cell concentrations exceeded 8E+05 cells/ml. For this purpose, cells were centrifuged for 5 min at 1500 g (RT), counted (see section 6.1.4), and subsequently cultivated at concentrations of 2E+05 cells/ml in RMPI complete medium.

6.1.1.2 Adherent tumor cells

Adherent tumor cells were maintained in an incubator at 37 °C, 5% CO_2, 95% air humidity, and were always kept in a sterile environment. Depending on cell numbers, adherent tumor cells were cultivated in 25 cm^2, 75 cm^2, or 175 cm^2 cell culture flasks containing 5-25 ml of RPMI complete medium and were grown under antibiotic selection if necessary (see Table 5.2.4). Adherent tumor cells were routinely passaged at approx. 80% confluence. For this purpose, the RPMI complete medium was discarded and adherent cells were washed with PBS. Tumor cells were then detached by adding 2-5 ml of Trypsin-EDTA for 3-5 minutes. After adding another 10 ml of RMPI complete medium, detached cells were centrifuged at 1500 g for 5 min (RT). After resuspending cell pellets in RPMI complete medium, adherent cells were counted (see section 6.1.4) and subsequently cultivated at concentrations of 1E+05-1E+06 cells in 5 ml (25 cm^2 cell culture flask), 0.5-1E+06 cells in 12 ml (75 cm^2 cell culture flask), or 0.8-4E+06 cells in 25 ml RPMI complete medium (175 cm^2 cell culture flask). The exact number of seeded tumor cells was individually determined based on respective proliferation rates of cultivated cell lines.

6.1.2 T cell culture

Human primary T cells were maintained in an incubator at 37 °C, 5% CO_2, 95% air humidity, and were always kept in a sterile environment. Depending on cell numbers, T cells were cultivated in 48-well plates (1 ml final volume per well) or 24-well plates (2 ml final volume per well), respectively. Generally, T cells were seeded at concentrations of 0.5-1E+06 cells per ml Panserin complete medium additionally supplemented with 600 U/ml IL-2. Transgenic T cells were cultivated in the presence of puromycin (see Table 5.2.4) to enrich T cell populations overexpressing transgenic TCRs after retroviral transduction (see section 6.4.10). T cells were routinely passaged when concentrations exceeded 1.5E+06 T cells per ml. For this purpose, cells were centrifuged for 5 min at 1500 g (RT), counted (see section 6.1.4), and subsequently cultivated at concentrations described at the beginning of this paragraph.

6.1.3 Phoenix-Ampho cell culture

The retrovirus producer cell line Phoenix-Ampho was used for producing infectious virions necessary for subsequent retroviral transduction of multiple tumor cell lines or human T cells (see section 6.4.10). Phoenix-Ampho cells were cultivated in an incubator at 37 °C, 5% CO_2, 95% air humidity, and were

always kept in a sterile environment. Depending on cell numbers, Phoenix-Ampho cells were either cultivated in 75 cm^2 or 175 cm^2 cell culture flasks containing 12-25 ml of DMEM complete cell culture medium. Poenix-Ampho cells were routinely passaged at approx. 80% confluence. For this purpose, the cell DMEM complete medium was discarded, adherent cells were washed with PBS, and were then detached by adding 5 ml TripleX solution for approx. 2 minutes. After adding another 10 ml of RMPI complete medium, detached cells were centrifuged at 1500 g for 5 min (RT). After resuspending cell pellets in DMEM complete medium, cells were counted (see section 6.1.4) and subsequently cultivated at concentrations of 0.5-1E+06 cells in 12 ml DMEM complete (75 cm^2 cell culture flask) or 0.8-4E+06 cells in 25 ml DMEM complete medium (175 cm^2 cell culture flask). For each retroviral transduction approach, Phoenix-Ampho cells were firstly seeded in 10 cm petri dishes at concentrations of 1.2E+06 cells per 8 ml DMEM medium and were then cultivated overnight. Cultivated Phoenix-Ampho cells were subsequently used for FuGENE-based transfection (see section 6.4.11) with respective expression constructs and helper plasmids (see Table 5.5).

6.1.4 Cell counting and determination of cell viability

Cell numbers were determined by using a 0.2 mm Fuchs-Rosenthal counting chamber. To stain all non-viable cells, 30 µl of cell suspension were diluted (1:2) with trypan blue. Since viable cells usually exhibit intact cell membranes, they typically do not intracellularly accumulate the trypan blue dye and can therefore be distinguished from non-viable cells with perforated cell membranes. Finally, cells were counted according to the manufacturer's instruction by using an optical microscope and respective cell numbers were calculated by using the following equation:

$$\frac{\textit{mean value of counted vital cells in 2 squares}}{\textit{0.2 mm}}\ \textit{x 1000 x 2 = cell concentration per ml}$$

6.1.5 Cryopreservation of tumor cell lines and human T cells

Tumor cell lines and human T cells were cryopreserved in liquid nitrogen to establish working stocks for ongoing and future experiments. To prevent the release of infectious virions after thawing, all retrovirally transduced tumor and T cells were cultured for two additional weeks before being frozen. Firstly, all cells intend for freezing were centrifuged at 1500 g for 5 min (RT), washed once in 10 ml PBS, and were subsequently counted for determining cell viabilities and cell numbers (see section 6.1.4). In case of adherent cells, detaching of respective cell lines was performed as described in sections 6.1.1.2 and 6.1.3, respectively. After washing in PBS, cell pellets were resuspended in 1 ml of freezing medium and were finally transferred into cryogenic tubes suitable for freezing in liquid nitrogen. Cell numbers varied among cryovials and ranged between 2E+06 and 10E+06 cells per ml freezing medium. Cryovials were eventually stored in pre-chilled Mr. Frosty freezing containers filled with isopropyl alcohol, thereby facilitating steady freezing of cells with approx. -1 °C per minute. Since DMSO is toxic for cells at room temperature, loaded freezing containers were transferred to -80 °C not later than 10 minutes after adding the freezing medium to respective cells. Mr. Frosty freezing containers with cryopreserved cells were subsequently kept at -80 °C for 24 hours before cryovials were stored in the vapor phase of liquid nitrogen.

6.1.6 Thawing of cryopreserved tumor cells and human T cells

Since DMSO is toxic for cells at room temperature, thawing of cryopreserved cells was performed as quickly as possible to maintain cell viability. Firstly, frozen cells were thawed in a water bath at 37 °C to a point where small lumps of ice were still visible. Thawed cells were immediately transferred into a 15 ml falcon tube containing 10 ml PBS and were then centrifuged at 1500 RPM for 5 minutes (RT).

Centrifuged cells were afterwards washed again with 10 ml PBS, then centrifuged at 1500 RPM for 5 minutes (RT), counted (see section 6.1.4), and eventually seeded in appropriate volumes of cell culture medium (see sections 6.1.1 to 6.1.3).

6.1.7 Glycerol storage of *E. coli* bacteria cultures

For establishing transgenic bacteria working stocks for later plasmid purifications (see section 6.4.1), transgenic *E. coli* bacteria were frozen in a 15% glycerol solution. Here, 750 µl of the respective bacteria cultures were supplemented with 250 µl of a 60% glycerol stock solution. Glycerol cultures were then transferred into cryovials and subsequently stored at -80 °C for 24 hours in pre-chilled Mr. Frosty freezing containers filled with isopropyl alcohol. Storing bacterial cultures in pre-chilled isopropyl alcohol facilitated steady freezing with approx. -1 °C per minute. Frozen bacterial cultures were finally removed from isopropyl alcohol-containing containers after 24 hours and were transferred into the vapor phase of liquid nitrogen for long-term storage.

6.1.8 Thawing of *E. coli* bacteria cultures

A starter culture of *E. coli* bacteria was inoculated by using frozen *E. coli* cultures (see section 6.1.7) and a 10 µl disposable inoculation loop. Here, frozen bacteria were transferred into 1 ml of LB bacterial culture medium (starter culture) and were then incubated for approx. 5 hours at 32 °C on a shaker before being transferred into conical tubes either containing 50 ml or 6 ml of LB medium. Resulting bacteria cultures were expanded overnight at 32 °C on a shaker. 50 ml cultures were subsequently used for Midipreps while 6 ml suspensions were utilized for Miniprep-based plasmid purification (see section 6.4.1). Throughout the thawing process, bacteria cultures were either supplemented with kanamycin (50 µg/ml) or ampicillin (100 µg/ml) to facilitate antibiotic selection based on used vector backbones (see Table 5.5).

6.2 Processing of NSCLC patient material

6.2.1 Resection of human lung and NSCLC tissue samples

Flash-frozen normal lung and tumor tissues from stage III NSCLC patients were provided by HS Diagnomics (HSD, Berlin, Germany). All analyzed patients underwent surgery with curative intent and provided written informed consent for the scientific analysis of all resected biospecimens. The study was performed in accordance with the declaration of Helsinki. Sample acquisition from NSCLC patients was approved by the ethics committee of the *Ärztekammer Berlin* (Eth-08/18). All biospecimens used during this project had initially been resected by the Institute of Pathology (Charité, University Medical Center Berlin, Germany). All provided tissue specimens were either classified as NSCLC or healthy lung tissue by the surgeons and pathologists of the Charité. Therefore, both HSD and the University Medical Center Mainz (UMC) had no influence on any of the performed tissue classifications. Except for specimens isolated from lung adenocarcinoma patient P18, all tissue samples were embedded in an RNA stabilizing reagent (RNAlater; Thermo Fisher Scientific, Waltham, USA) immediately after surgery and before freezing. Patient's flash-frozen lung and tumor tissue specimens were eventually sent on dry ice to the UMC for further comparative WES and RNA-Seq (see section 6.5.4) approaches. Blood, urine, and FFPE (formalin-fixed paraffin-embedded) tissue samples were additionally archived for each patient but were not transferred to the UMC. Patient-derived T cells were isolated from each blood, tumor and normal lung tissue sample and were eventually used for high-throughput TCR-beta profiling (see section 6.5.1).

6.2.2 Disruption of human lung and NSCLC tissues

Snap-frozen NSCLC and normal lung tissue specimens were used for the separate extraction of high molecular weight genomic DNA and total RNA (see sections 6.4.2 and 6.4.3). Patient-matched total RNA and genomic DNA samples were subsequently used for comparative WES and RNA-Seq (see section 6.5.4). Those NGS-based analyses were finally used for the identification of non-synonymous tumor-specific mutations in each patient. At the UMC, tumor and healthy lung tissue blocks (ca. 1 g) were cut into chunks of approx. 25 mg by using pre-chilled mortars and RNase-free scalpels. To avoid RNase-mediated degradation, tissue samples were constantly processed in the presence of liquid nitrogen. Additionally, mortars were baked for 6 h at 180 °C before tissue sample processing. Due to tumor heterogeneity, three individual tissue chunks from different tissue block areas were used for every gDNA isolation and total RNA purification. Hence, prepared tissue portions were dissected into pieces of approx. 16 mg and 9 mg, respectively. Three 16 mg tissue slices were pooled (approx. 50 mg in total) and then used as starting material for total RNA purification. In parallel, three tissue chunks (ca. 25 mg in total) were used for the isolation of high molecular weight gDNA. All tissues were homogenized by using a TissueLyser LT bead mill (Qiagen, Hilden, Germany) according to the manufacturer`s instructions (see: section 6.2.3). Due to tumor heterogenicity, performing RNA-Seq and WES from similar tissue block areas was crucial to effectively identify mRNA-expressed candidate mutations.

6.2.3 Homogenization of human lung and NSCLC tissues

Approx. 50 mg of normal lung or NSCLC tissues were used for the isolation of total RNA for subsequent RNA-Seq (see 6.5.4). For this purpose, respective tissues were homogenized in QIAzol Lysis Reagent (Qiagen, Hilden Germany) by using a TissueLyser LT bead mill (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Hence, pooled tissue samples (see: section 6.2.2) were transferred into 2 ml microfuge tubes containing one stainless steel bead. Microfuge tubes and steel beads had previously been pre-chilled on dry ice for 15 min, followed by another 15 min incubation after adding the tissue portions to the tubes. In case of homogenizing tissue samples stabilized with the RNAlater, chilling on dry ice was not performed and samples were immediately transferred into 2 ml microfuge tubes on ice. According to the manufacturer's protocol, tissue-containing microfuge tubes were subsequently transferred into the inserts of the TissueLyser LT Adapter and were kept on RT for 2 min. This was done to avoid freezing of the lysis buffer (700 µl QIAzol Lysis Reagent) that was immediately added to the tubes after RT incubation. Tissue specimens were homogenized twice at 50 Hz for 2.5 minutes. The homogenization was stopped for 30 s between both homogenization steps, thereby avoiding the potential heating of the processed samples. Homogenized samples were immediately used for total RNA purification (see section 6.4.3).

The homogenization of patient-derived tissue samples for subsequent purification of high molecular weight genomic DNA (see section 6.4.2) was also performed by using the TissueLyser LT bead mill (Qiagen, Hilden, Germany). However, only ca. 25 mg of pooled tissue samples (see section 6.2.2) were used for the gDNA isolation. Furthermore, tissue homogenizations were performed according to the instructions specified in the QIAamp Fast DNA Tissue Kit (Qiagen, Hilden Germany). Hence, tissue homogenizations were performed at 45 Hz for 2 min. Homogenized samples were immediately used for further processing and gDNA isolation.

6.2.4 Establishing single cell suspensions from human lung and NSCLC tissues

For assessing the recognition of patient-derived NSCLC cells via potentially tumor-specific TCRs, HSD generated single cell suspensions from normal lung and tumor tissues of stage III NSCLC patients P18, P43, and P50. Hence, tissue portions of approx. 1 g were manually disrupted by using sharp scalpels and petri dishes on ice. Due to tumor heterogenicity, multiple tissue chunks isolated from different tumor and lung regions were subjected to the GentleMACS tissue dissociation system (Miltenyi Biotec,

Bergisch Gladbach, Germany) according to the manufacturer's instructions. After passing through 70 µm strainers, dissociated single cell suspensions were cryopreserved in freezing medium (90% FCS / 10% DMSO) and small fractions were used for further Percoll-based density gradient centrifugation. Cells from Percoll-interphases were collected and rested overnight at concentrations of 0.5E+06 cells/ml in TexMACS medium (Miltenyi Biotec, Bergisch Gladbach, Germany) additionally supplemented with L-glutamine (Lonza, Basel, Switzerland), 25 mM HEPES (pH 7.2), 50 mM 2-mercaptoethanol (ThermoFisher Scientific, Waltham, USA), and 10% autologous serum. After harvesting and washing, CD3+, CD4+, CD8+ and PD-1+ cells were individually isolated from TIL- and normal lung-derived leukocytes by using magnetic bead separation according to the manufacturer's (Miltenyi Biotec, Bergisch Gladbach, Germany) instructions. For the identification of potentially tumor-reactive T cell clonotypes (see section 6.5.2), separated cell fractions were eventually shock frozen in liquid nitrogen.

6.2.5 Limited dilution cloning

Limited dilution cloning was used to establish monoclonal cell populations from polyclonal bulk cell cultures. Hence, bulk cell cultures of HLA-transduced or KRAS-modified tumor cell lines were plated in 96-well flat bottom cell culture plates at concentrations of 0.3, 1, and 3 cells/well, respectively. After cells were cultivated in RPMI complete medium supplemented with 2 µg/ml puromycin for 12-16 days, clonal cell populations were picked and transferred to 48-well plates for further expansion. Non-clonal cell populations were visually identified by using an optical microscope and were discarded. HLA-transduced monoclonal cell cultures were analyzed via W6/32 staining (see section 6.3.5) when cell numbers exceeded 1E+06 total cells. In contrast, successfully introduced KRAS-modifications were identified via IFN-γ ELISpot. Here, $KRAS^{Q61H}$-specific TCR-T cell populations were tested for the recognition of genetically modified NCI-H460 cells expressing the $KRAS^{Q61R}$ mutation (see section 6.4.14). $KRAS^{Q61R}$ modifications of non-recognized clones were eventually verified via Sanger Sequencing (see section 6.5.6).

6.3 Immunobiology

6.3.1 Purification of PBMCs from human leukocyte concentrates

Human T cells were isolated from healthy donor leukocyte concentrates (buffy coats) via density gradient centrifugation. Firstly, blood samples were diluted 1:3 with PBS before overlaying 15 ml Ficoll (Sigma-Aldrich Chemie GmbH, Steinheim am Albuch, Germany) with up to 35 ml of the resulting blood/PBS mixture. After being centrifuged for 20 min at RT (800 RCF, no brakes), PBMCs accumulated in a thin interphase between blood plasma and the high-density erythrocyte phase. This interphase was isolated by using electric pipette aids and 10 ml serological pipettes. After washing with 45 ml PBS (1800 rpm, RT, 10 min), remaining erythrocytes were lysed by carefully resuspending the cell pellet in 25 ml erythrocyte lysis buffer and subsequent incubation for 5 min at 33 °C. After another centrifugation (5 min, 1800 RPM, RT), the cell pellet was washed twice with 45 ml PBS. Purified human PBMCs were eventually used for MACS-mediated T cell isolation (see section 6.3.2).

6.3.2 MACS: Isolation of human T cells from purified PBMCs

Human CD4+ and CD8+ T cells were purified from PBMCs initially isolated from human leukocyte concentrates (see section 6.3.1). Hence, purified PBMCs were used for Magnetic Activated Cell Sorting (MACS) according to the manufacturer's (Miltenyi Biotec, Bergisch Gladbach, Germany) instructions. Firstly, purified PBMCs were counted (see section 6.1.4) and were then washed twice with 10 ml MACS buffer (1500 RPM, RT, 5 min). The supernatant was completely removed from cell pellets before adding 20 µl of antibody-labelled MicroBeads (Miltenyi Biotec, Bergisch Gladbach, Germany) in 500 µl

MACS buffer per 1E+06 PBMCs. PBMCs and magnetic MicroBeads were incubated for 15 min at 4 °C and were gently mixed every 5 minutes. To facilitate an optimal attachment, not more than 1E+08 cells were incubated with the antibody-conjugated magnetic bead mixtures per 500 µl MACS buffer. After incubation, cells were resuspended in 3 ml MACS buffer and were then transferred in 1 ml portions onto MACS MS columns with attached MACS SmartStrainer filters (both: Miltenyi Biotec, Bergisch Gladbach, Germany). Separation columns had before been placed into magnetic MiniMACS Separators and were equilibrated with 500 µl MACS buffer at RT. After separation, columns were washed three times with 500 µl MACS buffer to remove unbound cells. For elution, separation columns were eventually removed from magnetic MiniMACS Separators and bead-labelled human T cells were eluted by applying 2 ml Panserin complete medium per separation column and by using the provided plungers as described by the manufacturer. Finally, cell concentrations were determined (see section 6.1.4) and purification efficiencies were assessed via Flow cytometry (see section 6.3.5). All purified human T cells were subsequently used for OKT3-mediated T cell expansion (see section 6.3.3).

6.3.3 OKT3-mediated expansion of human T lymphocytes

Human T lymphocytes were purified from human buffy coat samples (see section 6.3.1) and were then expanded for following CRISPR/Cas9-mediated KOs of endogenous TCR (end.TCR) expression (see section 6.4.9). Non-specific activation with the anti-human CD3 monoclonal antibody OKT3 (30 ng/ml) was used to expand isolated T cell populations prior to the performed KO approach. For CD8+ T cell expansion, autologous CD8-negative PBMC fractions were irradiated (10,000 Rad) and additionally added to expanding T cell cultures, thereby serving as feeder cells. Hence, 0.5-1E+06 purified CD8+ T cells were cultivated in Panserin complete medium additionally supplemented with IL-2 (600 U/ml), OKT3 mAb (30 ng/ml), and autologous feeder cells at ratios of 1:1. OKT3-stimulated human T cells were used for CRISPR/Cas9-based KO approaches after being maintained for 2-4 days in an incubator (37 °C, 5% CO_2, 95% air humidity). CD4+ human T lymphocytes were expanded as described above for CD8+ T cells, but they were cultivated in the presence of irradiated CD4-negative PBMC fractions. End.TCR-KO T cells were eventually used for the retroviral transduction (see section 6.4.10) with expression constructs encoding potentially tumor-reactive TCRs (see section 6.5.2).

6.3.4 IFN-γ ELISpot assay

IFN-γ ELISpot assays were used to analyze the antigen-specific activation of TCR-transgenic CD4+ and CD8+ human T lymphocytes. Here, TCR-transduced effector T cells were co-incubated with target cells presenting potentially immunogenic antigen peptides on corresponding MHC class I molecules. Used targets were either HEK 293T cells simultaneously co-transfected with HLA alleles and antigen expression constructs (see section 6.4.12), patient-derived NSCLC cells, patient-derived normal lung cell suspensions, EBV-infected B cells, or HLA-transduced / non-modified tumor cell lines (see Table 5.1). Secreted IFN-γ was used as a marker for T cell activation upon antigen recognition and was quantified by using two anti-human IFN-γ monoclonal antibodies. Firstly, the released IFN-γ was bound by a primary antibody previously immobilized on PVDF membranes of 96-well ELISpot filter plates (Merck Millipore, Burlington, USA). A biotinylated secondary antibody was then used for detecting the immobilized IFN-γ by specifically binding to an alternative epitope. Peroxidase-conjugated avidin was eventually added, thereby facilitating the IFN-γ quantification after addition of a peroxidase-substrate solution. Due to the peroxidase enzyme activity, colored and insoluble precipitates (red spots) emerged on PVDF membranes of utilized ELISpot filter plates, thereby allowing the quantification of cytokine secretion upon antigen contact.

Briefly, PVDF membranes of MultiScreen filter plates were initially activated by adding 15 µl of 35% ethanol per well. After washing three times with 150 µl PBS per well, 50 µl PBS containing 500 ng of the primary antibody (anti-human IFN-γ mAb 1-D1K) were added per well and plates were incubated over night at 4 °C. For some wells, the monoclonal anti-human CD3 antibody OKT3 was additionally

added at a concentration of 400 ng/ml. Those wells were used as positive controls for maximal T cell activation. After incubation, primary antibody solutions were discarded and wells were washed three times with 150 µl PBS. PVDF membranes were afterwards blocked with 100 µl RPMI (10% FCS) per well for 30 minutes at 37 °C. Effector and target cells were then plated at desired concentrations in final volumes of 100 µl RPMI (10% FCS, 250 U/ml IL-2) per well. Exact target and effector cell quantities are individually depicted for each ELISpot assay (see section 7.2). Co-transfected HEK 293T cells were plated as described in section 6.4.12. All antibodies either used for IFN-γ detection or antigen recognition blocking are listed in Table 5.7.1.

Plated effector and target cells were eventually co-incubated for 20-24 h at 37 °C, 5% CO_2, and 95% air humidity. In contrast, spontaneous IFN-γ secretion was assessed by exclusively incubating T cells in the absence of OKT3 and target cells. Afterwards, PVDF membranes were washed six times with PBS (0.05% Tween20) before adding 60 µl of the secondary antibody (anti-human IFN-γ mAb 7-B6-1) at concentrations of 2 µg/ml in PBS (0.5% BSA). After incubating MultiScreen filter plates for 2 hours at 37 °C, antibody solutions were discarded and PVDF membranes were again washed six times with PBS (0.05% Tween20). In parallel, 100 µl of the avidin-peroxidase complex solution (see section 5.3) were prepared as described by the manufacturer, then incubated for 30 min at RT, and subsequently added to each well, respectively. After keeping the plates at RT and in the absence of light for another hour, PVDF membranes were washed three times with PBS (0.05% Tween20) and pure PBS for a total of 6 washings. Finally, 100 µl of H_2O_2-supplemented peroxidase-substrate solution were added and ELISpot plates were kept at RT for 10 min or until spot formation was clearly detectable. The precipitate-forming reaction was stopped by washing the MultiScreen filter plates under tap water for approx. 1 min. PVDF membranes were eventually dried at 50 °C and IFN-γ secretion was quantified by using the ImmunoSpot S5 Micro Analyzer (CTL Europe GmbH, Bonn, Germany) according to the manufacturer's instructions.

6.3.5 Flow cytometry

Flow cytometry was used to analyze protein surface expression on human primary T lymphocytes, EBV-infected B cells, HEK 293T cells, and multiple cancer cell lines. For analyzing endogenous and transgenic TCR expression, T lymphocyte subpopulations were either stained with the anti-human TCR constant region monoclonal antibody hTRAC-APC (Beckman Coulter, Krefeld, Germany) or with the anti-murine TCR constant region monoclonal antibody mTRAC-FITC (Origene, Rockville, MD, USA), respectively. T cell co-receptors were additionally stained with dye-labelled anti-CD3 (Beckman Coulter, Krefeld, Germany), anti-CD4 (Beckman Coulter, Krefeld, Germany), or anti-CD8 (BD Biosciences Franklin Lakes, USA) monoclonal antibodies. HLA expression of EBV-B cells, various cancer cell lines, and HEK 293T cells was analyzed by using a non-fluorescent HLA-A*01:01-specific antibody (One Lambda, Los Angeles, USA) and the pan-HLA class I antibody W6/32 (purified from hybridoma supernatants). Here, a fluorophore-labelled GaM-FITC secondary antibody was subsequently used for staining. All used flow cytometry antibodies and their corresponding specificities are listed in Table 5.7.2.

Briefly, 1E+05 cells were transferred into FACS tubes containing 2 ml FACS buffer (see section 5.3) and were then centrifuged for 5 minutes at 1500 rpm (RT). After supernatant removal, cell pellets were vortexed and incubated with flow cytometry antibodies for 20 minutes at 4 °C and in the absence of light. After incubation, cells were washed with 2 ml pre-chilled FACS buffer and were then centrifuged for 5 minutes at 1500 rpm (RT). After supernatant removal, stained cells were vortexed and either fixed with 200-500 µl FACS fixation buffer (see section 5.3) or, as previously described for dye-labelled antibodies, were stained with another GaM-FITC secondary antibody. Non-specific isotype controls were performed for every flow cytometry analysis to assess non-specific binding of used antibodies. Fixed cells were eventually stored at 4 °C and in the absence of light for a maximum of 5 days. Stained cells were analyzed by using a FACS Canto II (BD Biosciences, Heidelberg, Germany) device and the corresponding FACS Diva / FlowJo (both: BD Biosciences, Heidelberg, Germany) software.

6.4 Molecular Biology

6.4.1 Plasmid DNA purification

Plasmid DNA was purified from transformed *E. coli* bacteria by either using a QIAprep Spin Miniprep or a Plasmid Plus Midi Kit in combination with a QIAvac 24 Plus Vacuum Manifold (all: Qiagen, Hilden, Germany). DNA Minipreps were performed with 5 ml overnight bacteria cultures while 50 ml overnight cultures were used for DNA Midipreps. Glycerol stocks were prepared for each overnight culture as described in section 6.1.7. Briefly, bacteria were harvested by centrifugation (4000 g, 10 minutes, 4 °C) and subsequent DNA purifications were performed according to the manufacturer's instructions. Plasmid DNA was eventually eluted in 30 µl (Miniprep) or 100 µl (Midiprep) of the provided elution buffer, respectively. Hence, the buffer was added onto the spin columns and was incubated for 1 min at RT before plasmid DNA was eluted via centrifugation at 13000 rpm for 1min (RT). The elution was repeated once by using the flow-through. DNA concentrations were eventually determined by using the Qubit dsDNA BR-Assay-Kit (Thermo Fisher Scientific, Waltham, USA) according to the manufacturer's instructions.

6.4.2 Genomic DNA isolation

For the identification of non-synonymous tumor-specific mutations (see section 6.5.4), homogenized NSCLC and normal lung tissue samples were used for the extraction of high molecular weight genomic DNA (gDNA). Hence, the QIAamp Fast DNA Tissue Kit (Qiagen, Hilden, Germany) was used according to the manufacturer's instructions. Briefly, tissue samples were homogenized by using a TissueLyser LT device (see section 6.2.3) and gDNA was eventually eluted in 100 µl nuclease-free water. Here, the water was added onto the spin columns and incubated for 1 minute at RT before gDNA was eluted by centrifugation at full speed for 1min (RT). The elution was repeated once and resulting eluates were pooled to a final volume of 200 µl. Eventually, gDNA concentrations were determined by using the Qubit dsDNA BR-Assay-Kit (Thermo Fisher Scientific, Waltham, USA) according to the manufacturer's instructions. After isolation, gDNA integrities were checked via agarose gel electrophoresis (see section 6.4.5).

6.4.3 Total RNA purification

Homogenized NSCLC samples were used for the isolation of total RNA required for later gene expression analyses of tumor-specific patient mutations via RNA-Seq (StarSEQ GmbH, Mainz, Germany; see section 6.5.4). To also facilitate the efficient purification of rather small RNA molecules (≥18 nucleotides), total RNA samples were purified by using the miRNeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Briefly, tissue samples were initially homogenized by using a TissueLyser LT device (see section 6.2.3) and were then used for another homogenization via QIAshredder spin columns (Qiagen, Hilden Germany) as per the standard protocol. After DNase digestion, total RNA was eventually eluted in 30 µl nuclease-free water. Here, the water was added onto the RNeasy Mini spin columns and was then incubated for 1 minute at RT before elution via centrifugation at >10,000 RPM for 1min (RT). The elution was repeated once by using the flow-through. Eventually, RNA concentrations were determined by using the Qubit RNA BR Assay Kit (Thermo Fisher Scientific, Waltham, USA) as per the manufacturer's instructions. RNA integrities were finally analyzed via denaturing RNA gel electrophoresis (see section 6.4.6) or by assessing the corresponding RNA integrity numbers (RINs) via Bioanalyzer-based quality control (StarSEQ, Mainz, Germany).

6.4.4 Polymerase chain reaction (PCR)

6.4.4.1 Colony PCR

Colony PCRs were performed to verify the presence of DNA inserts in plasmid constructs of individual *E. coli* transformants (see section 6.4.13) after Gateway cloning (see section 6.4.8). Insert DNA orientation was additionally analyzed by pairing insert- and vector-specific primers, respectively. All primers had previously been designed by using the OligoAnalyzer Tool (IDT, Coralville, USA) and are listed together with their corresponding annealing temperatures (TMs) in Table 5.6.2. Briefly, single colonies were picked from agar plates and were then transferred into 100 µl LB medium either supplemented with 50 µg/ml kanamycin or 100 µg/ml ampicillin. 20 µl of the resulting pre-cultures were centrifuged (10 min, 10,000 RPM, RT) and cell pellets were subsequently resuspended in 20 µl PCR mix. PCR programs were set to 94 °C for initial denaturation (3 minutes), followed by 30 cycles of denaturation (94 °C, 30 s), primer annealing (20 s, TM), and elongation at 68 °C. Samples were kept at 4 °C after a final elongation of 3 minutes at 68 °C. PCR amplicons were eventually analyzed via agarose gel electrophoresis (see section 6.4.5). Remaining 80 µl of inoculated pre-cultures were incubated for approx. 1.5 h at 31 °C and were then used for establishing overnight cultures for later plasmid DNA purification (see section 6.4.1). Exact compositions of used colony PCR mixes are shown below.

DNA	N/A	Bacterial pellets
One Taq Master Mix with Standard Buffer (NEB, Frankfurt, Germany)	2x	10 µl
Primer sense	[10 µM]	1 µl
Primer reverse	[10 µM]	1 µl
Total volume		+ d.H_2O ad. 20 µl

6.4.4.2 cDNA flanking with attB1/attB2 recombination sites

cDNA fragments either encoding human HLA molecules or potential TCR antigens were flanked with attB1 and attB2 recombination sites to facilitate recombinase-mediated shuttling during Gateway cloning (see section 6.4.8). Hence, specific primers were designed by using the OligoAnalyzer Tool (IDT, Coralville, USA) and either included attB1 or attB2 recombination sites, respectively (see Table 5.6.2). Generally, PCRs were performed by using two different primer annealing temperatures (TMs) with (i) TM_1 facilitating sequence-specific flanking of the target cDNA and (ii) TM_2 enabling full-length primer binding for subsequent amplification. PCRs were performed by using the Q5 High-Fidelity DNA Polymerase (NEB, Frankfurt, Germany). PCR programs were set to 98 °C (3 min) for initial denaturation. For attB1/B2-flanking, 5 cycles of denaturation (98 °C, 10 s), primer annealing (20 s, TM_1), and elongation at 72 °C were performed. The amplification of flanked cDNA was facilitated by conducting 20 additional cycles comprising denaturation (98 °C, 10 seconds), primer-annealing (20 seconds at TM_2), and elongation at 72 °C. Finally, samples were kept at 4 °C after a final elongation of 3 minutes at 72 °C. Exact compositions of used PCR mixes are shown below.

cDNA template	N/A	5 µl
Q5 High GC Enhancer (NEB, Frankfurt, Germany)	5x	10 µl
Q5 Reaction buffer (NEB, Frankfurt, Germany)	5x	10 µl
dNTPs [10 mM]	[10 mM]	2 µl
Primer sense	[10 µM]	2 µl
Primer reverse	[10 µM]	2 µl
Q5 High-Fidelity DNA Polymerase (NEB, Frankfurt, Germany)	[2000 U/ml]	0.8 µl
Total volume		+ d.H_2O ad. 50 µl

6.4.4.3 Site-directed Mutagenesis

Amino acid substitutions were introduced in KRAS-encoding cDNA sequences by using the Q5 Site-Directed Mutagenesis Kit (NEB, Frankfurt, Germany). Here, site-directed mutagenesis (SDM) was used to introduce p.Q61K, p.Q61L, or p.Q61R KRAS amino acid substitutions and for establishing a non-mutated (wild-type) KRAS expression construct. Briefly, SDM reactions were conducted by using mutagenic back-to-back oligonucleotide primes, thereby introducing desired nucleotide mutations in plasmid cDNA inserts. All primers were custom designed by using the OligoAnalyzer Tool (IDT, Coralville, USA) and are listed, together with their corresponding annealing temperatures (TMs), in Table 5.6.3. All SDM reactions were performed in recombinant donor vectors containing $KRAS^{Q61H}$ cDNA constructs. After amplification, PCR products were purified via column purification (see section 6.4.4.4) and were then analyzed via agarose gel electrophoresis (see section 6.4.5). For amplification, PCR programs were set to 98 °C for initial denaturation (3 minutes), followed by 25 cycles of denaturation (98 °C, 10 s), primer annealing (20 s, TM), and elongation at 72 °C. Samples were kept at 4 °C after a final elongation of 3 minutes at 72 °C. The following Kinase-Ligase-DpnI (KLD) reaction simultaneously mediated both plasmid circulation and template plasmid removal. KLD reactions were performed for 10 min at RT before 0.5 µl of T4-DNA ligase (NEB, Frankfurt, Germany) was added and the reaction was incubated for another hour on ice. Finally, NEB 10-beta *E. coli* bacteria were transformed (see section 6.4.13) with the resulting donor vectors and were then cultivated overnight on kanamycin-containing LB agar plates. Exact compositions of used PCR and KLD reaction mixes are listed below.

PCR mix:

plasmid DNA template	5 ng	X µl
Q5 High GC Enhancer (NEB, Frankfurt, Germany)	5x	5 µl
Q5 Reaction buffer (NEB, Frankfurt, Germany)	5x	5 µl
dNTPs [10 mM]	[10 mM]	1 µl
Primer sense	[10 µM]	1 µl
Primer reverse	[10 µM]	1 µl
Q5 High-Fidelity DNA Polymerase (NEB, Frankfurt, Germany)	[2000 U/ml]	0.8 µl
Total volume		+ $d.H_2O$ ad. 25 µl

KLD reaction mix:

PCR-product	[100 ng/µl]	1 µl
T4 ligase buffer (NEB, Frankfurt, Germany)	10x	0.5 µl
BSA	[1 mg/ml]	0.5 µl
DpnI (NEB, Frankfurt, Germany)	[20000 U/ml]	0.5 µl
T4 DNA polynucleotide kinase (NEB, Frankfurt, Germany)	[10000 U/ml]	0.5 µl
$d.H_2O$	N/A	1.5 µl
T4 ligase (NEB, Frankfurt, Germany)	[400000 U/ml]	0.5 µl
Total volume		5 µl

6.4.4.4 Column Purification of PCR products

Amplified PCR products were purified by using the NuceloSpin Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions. Purified PCR products were eluted in 20 µl elution buffer (Qiagen, Hilden, Germany). Hence, the buffer was added onto the spin columns and was then incubated for 1 minute at RT before amplicons were eluted via centrifugation at 13,000 rpm for 1 min (RT). The elution was repeated once by using the flow-through. Finally, DNA concentrations were determined by using the Qubit dsDNA BR-Assay-Kit (Thermo Fisher Scientific, Waltham, USA) as per the manufacturer's instructions.

6.4.5 DNA gel electrophoresis

Amplified PCR product and gDNA integrities were analyzed via agarose gel electrophoresis. Hence, amplicons were either separated on 1% agarose (Serva Electrophoresis GmbH, Heidelberg Germany) TBE/TAE (both: Serva Electrophoresis GmbH, Heidelberg Germany) gels supplemented with 2 µl peqGREEN (Peqlab, Erlangen, Germany), thereby facilitating the detection via the UVsolo TS system (Analytic Jena, Jena, Germany). All samples were supplemented with 6x gel loading dye (NEB, Frankfurt, Germany) and were then loaded on agarose gels either containing 100 bp or 1000 bp DNA ladders (both: NEB, Frankfurt, Germany). Gel electrophoreses were performed for 30 to 45 min at 90-120 V by using 0.5x TBE or 1x TAE running buffers, respectively.

6.4.6 Denaturing RNA gel electrophoresis

Total RNA samples were analyzed via denaturing agarose gel electrophoresis. Hence, RNA samples were separated on 1% agarose TAE (Serva Electrophoresis GmbH, Heidelberg, Germany) gels supplemented with 2 µl GelRed (Biotium, San Francisco, USA), thereby facilitating the detection via the UVsolo TS system (Analytic Jena, Jena, Germany). RNA samples and RiboRuler High Range RNA Ladder (Thermo Fisher Scientific, Waltham, USA) were supplemented with formamide-containing 2x RNA loading dye (Thermo Fisher Scientific, Waltham, USA) before analysis. RNA samples and ladders were loaded on 1% TAE gels after they were incubated at 70 °C for 10 min and eventually chilled for 2 min on ice. Gel electrophoreses were performed for 30 to 45 min at 90 V by using 1x TAE running buffer. All buffers and agarose gels were prepared by using autoclaved, deionized, DEPC-treated water.

6.4.7 Reverse transcription cDNA synthesis

Purified total RNA from a patient's tumor tissue (see section 6.4.3) was used for reverse transcription cDNA synthesis. Established cDNA samples were eventually used for the cloning of potentially immunogenic target antigen open reading frames (ORFs). Those ORFs were eventually transferred into expression vector constructs for subsequent retroviral transduction (see section 6.4.10) or transient transfection (see section 6.4.12). Briefly, the PrimeScript RT-PCR Kit (Takara Bio Inc, Kusatsu, Japan) was used according to the manufacturer's instructions. Hence, 1-1.5 µg of total RNA was used together with oligo dT primers for reverse transcription (41 °C, 30 min). Target regions encoding potential TCR target antigens were eventually amplified via PCR, thereby resulting in ORFs flanked with attB1/B2 recombination sites (see section 6.4.4.2). Resulting PCR products were firstly analyzed on 1% agarose TBE gels (see section 6.4.5), were then purified via column-purification (see section 6.4.4.4), and finally used for Gateway cloning (see section 6.4.8).

6.4.8 Gateway Cloning

The Gateway Cloning system was used for shuttling recombination site-flanked cDNA constructs (see section 6.4.4.2) into expression vectors suitable for retroviral transduction (see section 6.4.10) or transient transfection (see section 6.4.12) of various target cells. Directional cloning of cDNA sequences into donor and destination vectors was mediated by recombinases of the bacteriophage λ integrase family. Briefly, the BP Clonase II Enzyme Mix (Invitrogen, Carlsbad, USA) facilitated the recombination of attB1/B2-flanked cDNA constructs and donor vectors containing attP1/attP2 recombination site-flaked *ccdB* suicide gene expression cassettes [351]. Due to integrase-mediated recombination, resulting donor vectors, now referred to as "entry clones", contained desired cDNA fragments flanked by attL1/L2 recombination sites. In contrast, undesired BP reaction byproducts contained the attR1/R2-flanked *ccdB* suicide gene. BP reactions were performed overnight at 25 °C in final volumes of 5 µl. Every reaction included 1-2 µl of TE buffer (10 mM), 1 µl BP Clonase II Enzyme Mix and similar amounts (20-100 ng) of donor vectors and attB1/B2-flanked PCR products. NEB 10-

beta *E. coli* bacteria were transformed (see section 6.4.13) with established entry clones after Protease-K treatments and were subsequently cultivated overnight on kanamycin-containing LB agar plates. Protease-K treatments were performed by adding 1 µl Protease-K (2 µg/µl), following incubation at 37 °C for 10 min, and subsequent chilling on ice for 2 min. The next day, individual bacteria colonies were picked to verify the presence of inserted cDNA sequences via colony-PCR (see section 6.4.4.1). Amplified plasmid DNA of positive colonies was subsequently isolated (see section 6.4.1) for the following LR reaction. LR reactions were performed as described above for BP reactions. However, they included 1 µl of the LR Clonase II Enzyme Mix (Invitrogen, Carlsbad, USA) catalyzing the site-directed recombination of attL1/L2-flanked entry clone inserts and destination vectors containing attR1/R2-flaked *ccdB* suicide gene expression cassettes. Hence, undesired LR reaction byproducts contained the *ccdB* suicide genes, thereby preventing the amplification in transformed bacteria. Resulting "expression clones" included attB1/B2-flanked cDNA fragments and were eventually used for Protease-K treatment and subsequent *E. coli* transformation (see section 6.4.13). The next day, colony-PCRs were performed from individual colonies after overnight cultivation on ampicillin-containing (100 µg/ml) LB agar plates. Amplified plasmid DNA of positive colonies was subsequently isolated (see section 6.4.1) for further transient transfections and retroviral transductions of various target cells (see sections 6.4.10 to 6.4.12).

6.4.9 CRISPR/Cas9-based KO of endogenous TCR expression in human T cells

Primary human T cells were isolated from human buffy coat samples of healthy donors by using density gradient centrifugation (see section 6.3.1) and subsequent MACS purification (see section 6.3.2) as per the manufacturer's instructions. After 2-4 days of non-specific activation with OKT3 (see section 6.3.3), purified T cells were used for a CRISPR/Cas9-mediated KO of their endogenous TCR (end.TCR) expression. The CRISPR/Cas9 genome editing system was delivered as ribonucleoprotein (RNP) complexes by applying the Amaxa Human T Cell Nucleofector Kit (Lonza, Basel, Switzerland) according to the manufacturer's instructions. For achieving maximal KO efficiencies, both *TCR* chain gene loci were targeted by simultaneously using two different crRNAs. The crRNA targeting the human *TCR beta* chain locus was synthesized as previously described by Roth et al.[352]. The *TCR alpha* chain-specific crRNA had previously been designed by using the Alt-R Custom Cas9 crRNA Design Tool (IDT, Coralville, USA). Both crRNAs (see Table 5.6.4) were combined at ratios of 1:1 with the Alt-R CRISPR-Cas9 tracrRNA (Integrated DNA Technologies, Coralville, USA) to form two different guide RNA (gRNA) complexes. Resulting gRNA complexes were used together with the recombinant Cas9-NLS nuclease (QB_3 Macrolab, Berkeley, USA) for generating two RNP complexes either targeting human *TCR-alpha* or *TCR-beta* chain loci. After RNP formation (20 min, RT), 3-4E+06 T cells were transfected by using the provided nucleofection mix comprising the Amaxa Nucleofector Solution additionally supplemented with 1 µM Alt-R Cas9 Electroporation enhancer (IDT, Coralville, USA) and 4 µM of both RNPs, respectively. T cells were electroporated by using the pre-set program T-023 on a Nucleofector 2b Device (Lonza, Basel, Switzerland). Nucleofected T cells were eventually cultivated at concentrations of 1E+06 cells/ml in Panserin complete medium supplemented with 600 U/ml IL-2. TCR-KO efficiencies were assessed via Flow cytometry analysis (see section 6.3.5) after expanding CRISPR/Cas9-treated cells for 4-6 days at 37 °C, 5% CO_2, and 95% humidity.

6.4.10 Retroviral transduction

Purified human T lymphocytes were used for retroviral transductions with TCR expression constructs (see Table 5.5) encoding potentially tumor-reactive TCRs after being used for a CRISPR/Cas9-based end.TCR-KO approach (see section 6.4.9). In contrast, other human cell lines were used for retroviral transductions to facilitate long-term overexpression of exogenous HLA alleles. TCR-transduced T cells were eventually used as effector cells in IFN-γ ELISpot assays while HLA-transduced cells functioned as targets (see section 6.3.4). All transduction approaches were conducted by using the Phoenix-Ampho virus producer cell line (see section 6.4.11). Approx. 6 ml of infectious gamma retroviral supernatant

were collected from transfected Phoenix-Ampho cells and were then centrifuged (2000 RPM, 10 min, 32 °C) to remove residual retroviral producer cells and debris. For retroviral transduction of human T lymphocytes, 3-5E+06 T cells were resuspended in 6 ml retroviral supernatant (4 µg/ml Polybrene) and were then transferred into 24-well plates (2 ml per well). For transducing other human cell lines with exogenous HLA expression constructs, used cell numbers differed between adherent and non-adherent tumor cells. In case of cell suspensions, 2-3E+06 cells were resuspended in 6 ml of the infectious supernatant (4 µg/ml Polybrene) and were subsequently transferred into 6-well plates (3 ml per well). For transducing adherent cells, 8E+06 cells were plated in two wells of a 6-well plate and 3 ml retroviral supernatant (4 µg/ml Polybrene) were added after approx. 20 h to both wells, respectively. To optimize transduction efficiencies, all prepared 6- and 24-well plates were eventually centrifuged (2000 RPM, 90 min, no breaks, 32 °C) and were then cultivated for approx. 20 h at 37 °C, 5% CO_2, and 95% humidity. Infectious viral supernatants were removed by washing the cells with 10 ml PBS (1500 RPM, 5 min, RT) approx. 24 h after transduction. Depending on cell types, transduced cells were cultivated in appropriate volumes of RPMI complete or Panserin complete medium additionally supplemented with 1 µg/ml of puromycin, respectively. Transduced adherent cells were detached from 6 well plates before cultivation in RPMI complete medium. Transduced human T cells were additionally stimulated with Human T-Activator CD3/CD28 Dynabeads (Thermo Fisher Scientific, Waltham, USA) at ratios of 3 µl per 1E+06 T cells. Anti-CD3/28 beads were eventually removed after 1 week of non-specific T cell expansion by using a magnetic particle separator (Dynal, Hamburg, Germany) according to the manufacturer's instructions.

6.4.11 Transient transfection of Phoenix-Ampho cells

The Phoenix-Ampho virus producer cell line was utilized for generating infectious retroviral particles for subsequent transduction of human T lymphocytes and various tumor cell lines (see section 6.4.10). Hence, Phoenix-Ampho cells were co-transfected with three different vectors: While cDNA expression constructs were provided in pMX destination vectors (see Table 5.5), two additional helper plasmids (pHIT60-GagPol, pCOLT-GALV) were used to encode essential parts for later virion assembly and function, such as viral enzymes, envelope proteins, and viral core structure proteins [345,346]. Phoenix-Ampho cells were plated in 10 cm petri dishes at concentrations of 1.2E+06 cells in 8 ml DMEM complete medium. The medium was replaced with 6 ml of fresh DMEM complete medium after approx. 24 h of incubation (37 °C, 5% CO_2, 95% humidity). Cells were then transfected by using the non-liposomal multicomponent transfection reagent FuGENE 6 (Promega, Madison, USA). Hence, 35 µl FuGENE 6 were added to 800 µl of non-supplemented DMEM medium before the mixture was incubated for 5 min at RT. Afterwards, 5 µg of both helper plasmids (pHIT60, pCOLT-GALV) were added together with 10 µg of the respective pMX plasmids either encoding TCR or HLA expression cDNA constructs (see Table 5.5). Subsequently, the transfection mix was dropped onto the plated Phoenix-Ampho cells after a final incubation period of 15 min at RT. After approx. 24 hours of incubation (37 °C, 5% CO_2, 95% humidity), the DMEM medium was removed and 8 ml of fresh RPMI complete medium were carefully transferred onto plated Phoenix-Ampho cells. After a final incubation of 19-20 hours (37 °C, 5% CO_2, 95% humidity), retrovirus-containing supernatants were harvested and used for subsequent transductions of human T lymphocytes and tumor cells.

6.4.12 Transient cDNA transfection of HEK 293T cells

For the transient expression of HLA/antigen combinations, HEK 293T cells were co-transfected with patient-specific HLA alleles and expression constructs either encoding tumor-specific antigens (TSAs) or tumor-associated antigens (TAAs). High transfection efficiencies were achieved by using the cationic transfection reagent Lipofectamine 2000 (Invitrogen, Waltham, USA) to transiently modify HEK 293T cell cultures exhibiting maximal cell confluences of 60-80%. Transfected HEK 293T were eventually used as target cells in IFN-γ ELISpot Assays (see section 6.3.4). Since ELISpot assays were performed in 96-well filter plates (Merck Millipore, Burlington, USA), respective transfection mixes were prepared

in conventional 96-well U-bottom plates and were then transferred onto 2E+04 HEK 293T cells previously plated in 120 µl RPMI medium per well (ELISpot filter plate). Lipofectamine Master mixes were prepared for each transfection approach. For each well, 0.5 µl of the transfection reagent were mixed with 20.5 µl RPMI medium, followed by incubation for 5 min at RT. In parallel, 300 ng TAA/TSA expression constructs were mixed with 100 ng of HLA-encoding pcDNA3.1 vector (see Table 5.5) before being resuspended in 12 µl RPMI and subsequently transferred into a 96-well U-bottom plate. 21 µl of the previously prepared Lipofectamine master mix were added per well before mixing wells by pipetting and subsequent incubation for 20 min at RT. Finally, 30 µl of the resulting transfection mix were added to each well of HEK 293T cells (20,000 per well) that had previously been seeded in ELISpot filter plates. After mixing cells and transfection mixes by pipetting, HEK 293T cells were incubated for 20-24 h at 37 °C (5%, CO_2, 95% humidity). After incubation, 100 µl transfection mix were discarded and effector cells were added at desired cell concentrations in 50 µl RPMI medium per well. Transfection efficiencies were assessed via EGFP control transfections. Here, HEK 293T cells were transfected with 100 ng of an pcDNA3.1-encoded EGFP construct (see Table 5.5), thereby facilitating flow cytometry-based expression analysis approx. 24 h after transfection.

6.4.13 Transformation of NEB 10-beta competent *E. coli* bacteria

Chemically competent NEB 10-beta *E. coli* bacteria were used to amplify plasmid cDNA constructs during Gateway cloning (see section 6.4.8) and site-directed mutagenesis (see section 6.4.4.3). For transformation, 3 µl of plasmid cDNA constructs were mixed with 25 µl of NEB 10-beta *E. coli* bacteria and were afterwards incubated for 30 min on ice. After heat shocking for 30 s at 42 °C (water bath), bacteria were incubated on ice for 2 min before being transferred into 125 µl SOC Outgrowth Medium (New England Biolabs, Ipswich, USA) and subsequent incubation for 2 h at 37 °C (shaker). Approx. 50-70 µl of transformed bacteria were eventually streaked on LB agar plates either containing 100 µg/ml ampicillin or 50 µg/ml kanamycin. Finally, LB agar plates were incubated over night at 32 °C and individual bacteria colonies were picked for later Colony-PCR (see section 6.4.4.1).

6.4.14 CRISPR/Cas9-based knock-in of the $KRAS^{Q61R}$ mutation in NCI-H460 cells

The CRISPR/Cas9 genome editing system was used to introduce a glutamine to arginine substitution at KRAS amino acid position 61 (p.Q61R) in HLA-A*01:01-expressing NCI-H460 lung carcinoma cells. A KRAS-specific crRNA was designed by utilizing the Alt-R Custom Cas9 crRNA Design Tool (IDT, Coralville, USA) and was then combined with the Alt-R CRISPR-Cas9 tracrRNA (IDT, Coralville, USA) to generate highly specific ribonucleoproteins (RNPs). Single-stranded oligodeoxynucleotide (ssODN) constructs were designed by using the Alt-R CRISPR HDR Design Tool (IDT, Coralville, USA) and were then used to introduce the p.Q61R mutation via homology-directed repair (HDR).These ssODN constructs encoded the $KRAS^{Q61R}$ mutation flanked by homology arms of 40-46 nucleotides. To improve donor oligo stabilities, ssODN constructs additionally included IDT-proprietary end-blocking groups and two phosphorothioate bonds between their first and last three bases, respectively. Incorporating three additional silent mutations at ssODN nucleotide positions 48, 60, and 63 prevented the Cas9 enzyme from re-cutting target sequences after successful HDR reaction. The CRISPR/Cas9 system was delivered by using the preset electroporation program X-001 on a Nucleofector 2b device (Lonza, Basel, Switzerland). Briefly, two-part gRNA complexes were prepared at final concentrations of 100 µM and were then combined with recombinant Cas9-NLS nucleases (QB_3 Macrolab, Berkeley, USA) for the generation of *KRAS*-specific RNPs. After RNP formation (20 min, RT), 3E+06 HLA-A*01:01-positive NCI-H460 cells were transfected in 110 µl OptiMeM medium containing 4 µM ssODN templates, 1 µM Alt-R Cas9 Electroporation Enhancer (IDT, Coralville, USA), and 4 µM RNPs. Nucleofected cells were finally cultivated (37 °C, 5% CO_2, 95% humidity) in 6-well plates by using 2 ml RPMI complete medium supplemented with 30 µM Alt-R HDR Enhancer V2 (IDT, Coralville, USA) per well. The medium mix was replaced with RPMI complete medium after 18-20 h of cultivation.

Single cell clones were established via limiting dilution cloning (see section 6.2.5) after 3 days of cultivation (37 °C, 5% CO_2, 95% humidity).

6.4.15 Preparation of peptide libraries

Most promising neoantigen candidates were identified via comparative WES and RNA-Seq (see section 6.5.4) and were then synthesized (JPT, Berlin, Germany) as patient-specific 9- and 10-mer neopeptide libraries (purity >90%). Hence, best HLA-binding 9- and 10-mer neopeptide sequences were identified (see section 6.5.8) for each patient, respectively. Lyophilized peptides (31 µg/peptide) were subsequently dissolved in 60 µl DMSO to establish concentrations of approx. 516 µg/ml and were then diluted with RPMI complete medium to final concentrations of 16 µg/ml. Eventually, peptide library master plates (96-well) containing 200 µl/well of each peptide (16 µg/ml) were prepared and subsequently frozen at -20 °C. Individual 9- and 10-mer peptides were finally used at concentrations of 2 µg/ml for following ELISpot assays (see section 6.3.4).

6.4.16 Preparation of tandem minigene libraries

Most promising P43 neoantigen candidates were identified via comparative WES and RNA-Seq (see section 6.5.4) and were subsequently synthesized as peptide (see section 6.4.15) or tandem minigene (TMG) libraries. According to Lu *et al.* [190], individual TMGs were designed as continuous ORFs with no spacer sequences and contained up to 5 patient-specific amino acid mutations or TAA epitopes. Every single candidate mutation was adjacently flanked by the 12 amino acids of the corresponding non-mutated protein, thereby facilitating the generation of every possible 8- to 12-mer peptide exhibiting the respective mutation at any possible position. TMGs were synthesized and cloned into ready-to-transfect pTwist_CMV_BetaGlobin_WPRE_Neo vectors (Twist Biosciences, South San Francisco, USA). Lyophilized DNA vectors were resuspended in 100 µl TE buffer, thereby establishing individual TMG concentrations of 20 ng/µl, and were subsequently stored at -20 °C. HEK 293T cells were eventually transfected (see section 6.4.12) by using 150 ng TMG cDNA and 100 ng of a patient-specific HLA expression construct. Finally, co-transfected HEK 293T cells were used as targets in subsequent IFN-γ ELISpot assays (see section 6.3.4). Transfection efficiencies were assessed as described in section 6.4.12.

6.5 Sequencing and bioinformatics

6.5.1 TCR beta profiling

HSD analyzed TCR beta repertoires of stage III NSCLC patients by isolating high molecular weight genomic DNA (gDNA) from a patient's tumor- and healthy lung tissue-resident T cells. Prior to gDNA purification, single cell suspensions were created from excised lung and tumor tissues before being used for Percoll-based density gradient centrifugation (see section 6.2.4). Here, CD3+, CD4+, CD8+, and PD-1+ T lymphocytes were purified from isolated leukocyte fractions by using MACS cell separation according to the manufacturer's instructions. Eventually, gDNA was isolated from respective T cell fractions by using the QIAamp DNA Blood Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Isolated gDNA samples were finally used for tissue-specific TCR beta profiling via NGS. Hence, *TCR beta* chain-encoding gene regions were analyzed by using the Illumina MiSeq System (Illumina, San Diego, USA) by applying TRBV/J-specific primer sets and a proprietary two-step contamination protection method for NGS-library preparation [353,354]. TCR repertoire sequencing results (TRBV chains, 2x 150 bp paired Illumina reads) were processed by using an HSD-developed pipeline to establish CDR3 consensus sequences while simultaneously removing inconsistent or non-overlapping read-pairs. TCR clonotypes were identified by clustering of high-quality consensus

CDR3 sequences. Clonotype frequencies were calculated as percentages of all clonotypes found in a given sample. Revealed TCR beta repertoires of analyzed T cell subpopulations were ultimately used for the identification of potentially tumor-reactive TCR clonotypes (see section 6.5.2).

6.5.2 Identification of potentially tumor-specific TCR clonotypes

For identifying potentially tumor-specific T cell clonotypes, HSD compared TCR beta repertoires in healthy lung and tumor tissues of NSCLC patients P18, P43, and P50 (see section 6.5.1). Tumor-to-non-tumor (T/nT) ratios of individual clonotypes were finally determined with high T/nT ratios indicating both tumor enrichment and potential tumor reactivity. Here, promising TCR clonotypes exhibited T/nT ratios of >5 while T/nT ratios of >20 characterized the most dominant clonotype candidates. By also analyzing PD-1+ cell fractions from a patient's normal lung and tumor tissues, TCR clonotypes were eventually ranked based on both T/nT ratios as well as PD-1 expression. Thus, only PD-1+ TCR clonotypes exhibiting the highest T/nT ratios were considered potential tumor-reactive [208] and were therefore chosen for further antigen screenings. Since T/nT ratios were exclusively determined via CDR3 beta sequencing (see section 6.5.1), RNA-Seq (10x Genomics, Pleasanton, USA) was eventually used to reveal complete αβ sequences of most promising TCR candidates (see section 6.5.3). Ready-to-transfect expression constructs encoding paired αβ TCR sequences were finally synthesized for the generation of transgenic T cell populations (see section 6.4.10).

6.5.3 Establishing TCR expression constructs

Single T cell sequencing was used for revealing paired αβ TCR sequences of potentially tumor-reactive TCR clonotypes. Hence, HSD used the 10x Genomics GemCode Technology (10x Genomics, Pleasanton, USA) for generating single cell cDNA-libraries from CD3+ TIL suspensions. Briefly, TCR expression of single T lymphocytes were analyzed on mRNA level by applying the Chromium Next GEM Single Cell V(D)J Reagent Kit in combination with the Chromium Single Cell V(D)J Enrichment Kit (both: 10x Genomics, Pleasanton, USA) according to the manufacturer's instructions. Here, cDNA was generated from individual T cells by using constant region-specific primers for TCR alpha and beta chains, respectively. Complete TCR alpha and TCR beta segments were subsequently amplified by using a full-nested PCR approach including primers for every possible combination of V_α and C_α segments. Since VDJ regions of most dominant TCR clonotypes had previously been revealed (see section 6.5.2), the PCR-mediated amplification of corresponding TCR beta chains was performed by exclusively using the respective V_β primers. Adding plate- and well-specific barcodes during library preparation enabled the pooling of up to 6 different 96-well plates for subsequent NGS analysis. Single cell sequencing reads were eventually processed by using the 10x Genomics Cell Ranger pipeline (v6.0.2) with default parameters to generate unique molecular identifier (UMI) matrices. After further downstream analysis and quality control, cells with less than 400 UMI, fewer than 250 genes, and greater than 20% UMI in mitochondrial genes were removed.

Paired αβ TCR sequences of potentially tumor-reactive T cell clonotypes were finally cloned into expression constructs suitable for subsequent retroviral transduction of human donor lymphocytes (see Table 5.5). Hence, VDJ regions of respective αβ TCRs were synthesized as gBlock (IDT, Coralville, USA) gene fragments, which were eventually used for generating bicistronic cDNA constructs encoding human variable domains together with murine TCR constant regions. As previously described by Kropp *et al.* [225], TCR sequences were cloned into pMX-puro retroviral expression vectors containing single P2A segments to facilitate the separation of respective TCR chains during translation. While HSD performed all TCR sequencing analyses and corresponding cloning approaches (see also section 6.5.2), pMX cDNA backbones were initially designed and provided by the Wölfel group (University Medical Center Mainz, Mainz, Germany).

6.5.4 Comparative WES and RNA-Seq

Comparative whole exome sequencing (WES) and RNA-Seq were used to identify mRNA-expressed tumor-specific nucleotide variations either causing amino acid substitutions or aberrant reading frames in NCSLC tissues of individual patients. Therefore, high molecular weight gDNA (see section 6.4.2) and total RNA (see section 6.4.3) were isolated for subsequent NGS from NSCLC and normal lung tissues of patients P18, P43, and P50. Comparative WES of tumor and normal tissue-derived gDNA was used to identify potential neoantigen-encoding mutations, such as single nucleotide variations (SNVs), multiple nucleotide variations (MNV), and insertions/deletions (InDels). In contrast, RNA-Seq was exclusively performed for tumor-obtained total RNA samples to verify actual mRNA expression of neoantigen candidates previously identified via WES. For library preparation and subsequent transcriptome sequencing, 6 µg of NSCLC-derived total RNA were sent to StarSEQ GmbH (Mainz, Germany) on dry ice. In contrast, comparative WES was performed (StarSEQ GmbH, Mainz, Germany) by either using 1 µg of tumor- or normal lung-derived gDNA. While RNA sample qualities were generally assessed by using a Bioanalyzer (Agilent Technologies, Santa Clara, USA) device, gDNA qualities were analyzed via gel electrophoresis (see section 6.4.5) and NanoDrop (Thermo Fisher Scientific, Waltham, USA) analysis, respectively. The NextSeq500 (Illumina, San Diego, USA) system was used for all NGS approaches, thereby resulting in approx. 60 million paired-end WES reads (150 nt) and an estimated coverage of 100x per sample. Furthermore, the SureSelectXT Human All Exon Kit (Agilent Technologies, Santa Clara, USA) was applied for all exome enrichments conducted by StarSEQ (StarSEQ GmbH, Mainz, Germany). Transcriptome sequencing of tumor-obtained total RNA samples resulted in approx. 80 million paired-end reads (150 nt) and an estimated coverage of at least 30x per sample. The resulting sequencing data was further processed as described in section 6.5.7.

6.5.5 HLA genotyping

HLA genotyping was performed for NSCLC patients P18, P43, and P50. Hence, gDNA was isolated (see section 6.4.2) from NSCLC tissues of individual patients and was then sent to the *Institut für Immunologie und Genetik* (Kaiserslautern, Germany) for further analysis. Due to potential HLA loss in patient-derived tumors, only tumor-derived gDNA (approx. 1 µg) was used for HLA genotyping.

6.5.6 Sanger Sequencing

Sanger Sequencing was used to verify the integrities of all established cDNA expression constructs (see Table 5.5). Hence, Mix2Seq Sequencing Kits (Eurofins Genomics, Ebersberg, Germany) were used according to the manufacturer's instructions after expression constructs were amplified in previously transformed *E. coli* bacteria (see section 6.4.13). Depending on cDNA inserts and vector backbones, matching primers were chosen for subsequent sequencing (see Table 5.6.1). Obtained sequencing data was subsequently verified by using the Geneious 8.1.9 (Biomatters Inc, Auckland, New Zealand) sequencing data analysis software and intact expression constructs were eventually used for retroviral transduction (see section 6.4.10) or transient transfection (see section 6.4.12) approaches, respectively.

6.5.7 Processing of WES and RNA-Seq data

Comparative WES and RNA-Seq was used for the identification of potential neoantigen candidates that, in theory, could represent new target antigens for future immunotherapies (see section 6.5.4). Hence, WES and RNA-Seq data was further analyzed by using the CLC Genomics Workbench (Qiagen, Hilden, Germany) software (versions 20-23; see corresponding tools depicted below in italics). NGS data analyses were mostly performed by [redacted] (UMC Mainz, Mainz, Germany).

Briefly, all sequencing reads were used for adaptor trimming before they were mapped to the human reference genome assembly hg38 (GCA_000001405.28, GCF_000001405.39). The *Local Realignment*

tool was then used to optimize overall read mappings to the reference genome. Next, tumor tissue-derived WES data was used for *Low frequency Variant Detection* to identify all patient-specific mutations not present in hg38. Tumor-specific mutations were then identified by excluding all variants also present in the normal lung tissue of the respective patient via the *Remove Variants Present in Control Reads* function. Tumor-specific variants resulting from poor sequencing quality or few sequencing reads were immediately excluded by applying the *Remove Marginal Variants* algorithm. Remaining mutations were afterwards annotated based on mappings with hg38 (*Annotate with Overlap Information*). In parallel, analysis and annotation of tumor-specific gene transcript expression was facilitated by applying the *RNA-Seq Analysis* tool and by using adapter-trimmed patient tumor RNA-Seq data. Detected RNA transcripts were considered to be actually expressed in a patient's tumor tissue when RPKM scores were ≥0.5. The *in silico* screening for RNA transcripts was subsequently optimized by applying the *Local Realignment_RNASeq* algorithm, thereby removing false positive results. Potentially immunogenic candidate neoantigens were eventually identified by filtering for mRNA-expressed (RPKM ≥0.5) tumor-specific mutations present in a patient's tumor tissue (*Filter Based on Overlap*) and subsequent filtering for amino acid changing consequences (*Amino Acid Changes*). Finally, most prominent candidate neoantigens (average quality >34, variant read counts >4, quality >100) were manually curated and used for following peptide binding predictions (see section 6.5.8).

6.5.8 Candidate neoantigen peptide binding predictions

To screen for potentially immunogenic peptides that are most likely to be presented by a patient's HLA molecules, annotated candidate neoantigen lists resulting from NGS data processing (see section 6.5.7) were further used for patient-specific peptide binding predictions. Peptide binding predictions were performed by [redacted] (UMC Mainz, Mainz, Germany). Nucleotide and corresponding amino acid sequences of mRNA-expressed neoantigen candidates (see section 6.5.4) were obtained from the Ensemble Genome Browser 105 database (see Table 5.8).

Since HLA class I molecules preferentially bind short peptides of 9 to 10 amino acids [18], all possible 9- and 10-mer peptides comprising one of the identified candidate mutations were considered to be potentially immunogenic. Hence, peptide binding predictions were performed by utilizing 19-mer peptides comprising one of the respective candidate mutations flanked by the 9 adjacent amino acids of the corresponding canonical protein. Peptide binding analyses were performed for each of a patient's HLA alleles by applying the *in silico* generated 19-mer peptides to different prediction tools, namely the IEDB MHC-I Binding Prediction tool and the NetMHC-4.0 [257] algorithm (see Table 5.8). Depending on the HLA alleles analyzed during this study, two different prediction methods were applied when using the IEDB MHC-I binding prediction tool, namely the Consensus [258] and the NetMHCpan-4.0 [259] algorithms. The 9- and 10-mer peptides predicted to bind a patient's HLA molecules best (IEDB percentile rank ≤6.0; NetMHC-4.0 binding affinity IC50 ≤500nM) were, if possible, determined as consensus candidates identified by both prediction tools. Most promising candidate neopeptides with highest predicted binding affinities were eventually synthesized (JPT Peptide Technlogies GmbH, Berlin, Germany) and used for subsequent IFN-γ ELISpot-based antigen screenings (see section 6.3.4). In case of tumor-specific frameshift mutations, peptide binding predictions were performed as described above for non-synonymous SMVs and MNVs. In contrast, 19-mer peptides were generated for every downstream mutation that might be induced by the respective nucleotide insertions or deletions.

6.5.9 *In silico* identification of recurrent diver mutations

One major goal of this study was to reveal patient-specific target neoantigens for future immunotherapies. Hence, comparative WES and RNA-Seq were used to identify mRNA-expressed candidate mutations in tumor tissues of NSCLC patients P18, P43, and P50 (see sections 6.5.4 and 6.5.7). To be able to identify recurrent somatic mutations shared among patients and multiple tumor types, all identified variations were screened for concordances with public databases listing hundreds of

recurrent driver mutations, such as COSMIC, Intogen, and TCGA (see Table 5.8). Whenever one of the variations was found to be listed as a recurrent driver mutation, corresponding cDNA was synthesized (see section 6.4.7) from patient-specific tumor total RNA samples and was then cloned into a ready-to-transfect pcDNA3.1 expression construct (see section 6.4.8). Generated expression constructs were finally added to the pre-existing TAA panel (Sup. Fig. 13) used for subsequent antigen screenings.

6.5.10 *In silico* identification of recurrent TCRs

One major goal of this study was to verify the tumor specificity of TCRs cloned from potentially tumor-reactive T cell clonotypes (see section 6.5.2) derived from NSCLC patients P18, P43, and P50. Tumor specificities were verified by identifying the exact tumor antigens specifically recognized by those TCRs. Hence, amino acid sequences of analyzed TCRs were checked against TAA/TSA specificity databases, thereby potentially circumventing highly extensive antigen screening approaches for some of the tested TCRs. Recurrent TCR analyses were performed by [redacted]. Here, individual α- and β-chain sequences, CDR3 sequences, and paired αβ TCR sequences were used to query databases like VDJdb (see Table 5.8) listing thousands common TCRs. Since pathogen-reactive T cells were previously also found to accumulate in tumor tissues of NSCLC patients [220], those database screenings were also performed to exclude pathogen-specific TCRs from following antigen screening approaches. By comparing HSD-derived TCR sequences with public receptor databases, it was possible to reveal antigen specificities towards both pathogen-derived antigens and TAAs (see section 3.4.4). ORFs encoding respective antigen epitopes were subsequently cloned (see section 6.4.8) in ready-to-transfect pcDNA3.1 vector backbones and were eventually used for further antigen screenings.

6.5.11 *In silico* identification of TCR clusters

TCR repertoire data obtained from different cancer patients were compared by means of sequence similarities in their respective CDR3 regions. The applied algorithm worked in two rounds, with round one generating a number of seed clusters identified by grouping of all paired TCRs (associated alpha and beta chain were treated as one connected sequence) exhibiting sequence similarities of at least 89% in both TRAV- and TRBV-CDR3 regions, respectively. For conducting efficient sequence comparisons between thousands of TCRs, several algorithms are publicly available, such as CDHIT, iSMART, and many more. Once the seed clustering was completed, consensus amino acid sequences covering the respective TRBV-CDR3 regions were calculated for each cluster. In a second round, generated consensus sequences and non-paired TCR beta chain sequencing data from additional patients were subjected to another clustering by applying similar parameters as in the first round. Finally, most promising cluster TCRs were determined when they were, according to the HS Diagnomics definition (see section 6.5.2), considered potentially tumor-specific in at least a subgroup of patients included in each of the respective clusters. All TCR clusters were identified by HS Diagnomics (Berlin, Germany).

6.5.12 Figure design

Figures were designed by using the PowerPoint 2019 (Microsoft, Redmond, USA) or the GraphPad Prism 10.0.3 (GraphPad Software, Boston, USA) software. Flow cytometry plots were designed by using the FACS DIVA 9.0 (BD Biosciences, Heidelberg, Germany) and the FlowJo 7.6.5 (FlowJo, Ashland, USA) software. All sequence analyses were conducted by using the Geneious 8.1.9 (Biomatters Inc., Auckland, New Zealand) software (see summarizing Table 5.8).

7) Appendix

7.1 List of References

1. Westergaard, M. C. W. *et al.* Tumour-reactive T cell subsets in the microenvironment of ovarian cancer. *Br. J. Cancer* **120**, 424–434 (2019).
2. Chruściel, E. *et al.* Adoptive Cell Therapy - Harnessing antigen-specific T cells to target solid tumours. *Cancers (Basel).* **12**, 683 (2020).
3. Hammer, R. & Steffen, H. European Patent EP3180433B1: Method for providing tumor-specific T cells. (2017).
4. Pearlman, A. H. *et al.* Targeting public neoantigens for cancer immunotherapy. *Nat. Cancer* **2**, 487–497 (2021).
5. Leko, V. & Rosenberg, S. A. Identifying and Targeting Human Tumor Antigens for T Cell-Based Immunotherapy of Solid Tumors. *Cancer Cell* **38**, 454–472 (2020).
6. Sim, M. J. W. & Sun, P. D. T Cell Recognition of Tumor Neoantigens and Insights Into T Cell Immunotherapy. *Front. Immunol.* **13**, 833017 (2022).
7. Matzaraki, V., Kumar, V., Wijmenga, C. & Zhernakova, A. The MHC locus and genetic susceptibility to autoimmune and infectious diseases. *Genome Biol.* **18**, 76 (2017).
8. Horton, R. *et al.* Gene map of the extended human MHC. *Nat. Rev. Genet.* **5**, 889–899 (2004).
9. Murphy, K. P., Travers, P., Walport, M. & Janeway, C. Chapter 6: Antigen Presentation to T Lymphocytes. in *Janeway's Immunobiology, 8th Edition,* 217–220 (2012).
10. Goodfellow, P. N. *et al.* The β2-microglobulin gene is on chromosome 15 and not in the HL-A region. *Nature* **254**, 267–269 (1975).
11. Ploegh, H. L., Cannont, L. E. & Strominger, J. L. Cell-free translation of the mRNAs for the heavy and light chains of HLA-A and HLA-B antigens. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 2273–2277 (1979).
12. Güssow, D. *et al.* The human beta 2-microglobulin gene. Primary structure and definition of the transcriptional unit. *J. Immunol.* **139**, 3132–3138 (1987).
13. Mungall, A. J., Palmer, S. A., Sims, S. K. & Edwards, C. A. DNA sequence and analysis of human chromosome 6. *Nature* **425**, 805–811 (2004).
14. Robinson, J. *et al.* IPD-IMGT/HLA Database. *Nucleic Acids Res.* **48**, 948–955 (2020).
15. Rudolph, M. G., Stanfield, R. L. & Wilson, I. A. How TCRs bind MHCs, Peptides, and Coreceptors. *Annu. Rev. Immunol.* **24**, 419–466 (2006).
16. Germain, R. N. MHC-dependent antigen processing and peptide presentation: Providing ligands for T lymphocyte activation. *Cell* **76**, 287–299 (1994).
17. Ioannides, C. G. & Whiteside, T. L. T Cell Recognition of Human Tumors: Implications for Molecular Immunotherapy of Cancer. *Clin. Immunol. Immunopathol.* **66**, 91–106 (1993).
18. Trolle, T. *et al.* The Length Distribution of Class I–Restricted T Cell Epitopes Is Determined by Both Peptide Supply and MHC Allele–Specific Binding Preference. *J. Immunol.* **196**, 1480–1487 (2016).
19. Wieczorek, M. *et al.* Major histocompatibility complex (MHC) class I and MHC class II proteins: Conformational plasticity in antigen presentation. *Front. Immunol.* **8**, 292 (2017).
20. Murphy, K. P., Travers, P., Walport, M. & Janeway, C. Chapter 4: Antigen Recognition by B-cell and T-cell Receptors. in *Janeway's Immunobiology, 8th Edition,* 145–146 (2012).
21. Rammensee, H. G. Chemistry of peptides associated with MHC class I and class II molecules. *Curr. Opin. Immunol.* **7**, 85–96 (1995).
22. Falk, K., Rotzschke, O., Stevanovic, S., Jung, G. & Rammenseet, H.-G. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* **351**, 290–296 (1991).
23. Murphy, K. P., Travers, P., Walport, M. & Janeway, C. Chapter 4: Antigen Recognition by B-cell and T-cell Receptors. in *Janeway's Immunobiology, 8th Edition,* 139–141 (2012).
24. Gogoi, D. & Chiplunkar, S. V. Targeting gamma delta T cells for cancer immunotherapy: Bench to bedside. *Indian J Med Res* **138**, 755–761 (2013).
25. Wo, J. *et al.* The Role of Gamma-Delta T Cells in Diseases of the Central Nervous System. *Front. Immunol.* **11**, 580304 (2020).
26. Murphy, K. P., Travers, P., Walport, M. & Janeway, C. Chapter 5: The Generation of Lymphocyte Antigen Receptors. in *Janeway's Immunobiology, 8th Edition,* 169–171 (2012).
27. Janeway, C. J., Travers, P. & Walport, M. Chapter 4: The Generation of Lymphocyte Antigen Receptors. in *Immunobiology: The Immune System in Health and Disease, 5th edition,* 166 (2001).
28. Rowen, L., Koop, B. F. & Hood, L. The complete 685-kilobase DNA sequence of the human β T cell receptor locus. *Science* **272**, 1755–1762 (1996).
29. Koning, F., Maloyo, W. L. & Coligano, J. E. The implications of subunit interactions for the structure of the T cell receptor-CD3 complex. *Eur. J. Immunol* **20**, 299–305 (1990).

30. Call, M. E., Pyrdol, J., Wiedmann, M. & Wucherpfennig, K. W. The organizing principle in the formation of the T cell receptor-CD3 complex. *Cell* **111**, 967–979 (2002).

31. Irving, B. A. & Weiss, A. The cytoplasmic domain of the T cell receptor zeta chain is sufficeint to couple to receptor-associated signal transduction pathways. *Cell* **64**, 891–901 (1991).

32. Barber, E. K., Dasgupta, J., Schlossman, S. F., Trevillyan, J. M. & Rudd, C. E. The CD4 and CD8 antigens are coupled to a protein-tyrosine kinase (p561ck) that phosphorylates the CD3 complex. *Proc. Nadl. Acad. Sci. USA* **86**, 3277–3281 (1989).

33. Samelson, L. E. Signal transduction mediated by the T cell antigen receptor: The Role of Adapter Proteins. *Annu. Rev. Immunol* **20**, 371–94 (2002).

34. Berg, L. J., Finkelstein, L. D., Lucas, J. A. & Schwartzberg, P. L. Tec family kinases in T lymphocyte development and function. *Annu. Rev. Immunol* **23**, 549–600 (2005).

35. Macian, F. NFAT proteins: Key regulators of T-cell development and function. *Nat. Rev. Immunol.* **5**, 472–484 (2005).

36. Chang, L. & Karin, M. Mammalian MAP kinase signalling cascades. *Nature* **410**, 37–40 (2001).

37. Perez, O. D. *et al.* Leukocyte functional antigen 1 lowers T cell activation thresholds and signaling through cytohesin-1 and Jun-activating binding protein 1. *Nat. Immunol.* **4**, 1083–1092 (2003).

38. Wülfing, C., Sjaastad, M. D. & Davis, M. M. Visualizing the dynamics of T cell activation: Intracellular adhesion molecule 1 migrates rapidly to the T cell/B cell interface and acts to sustain calcium levels. *Proc. Nadl. Acad. Sci. USA* **95**, 6302–6307 (1998).

39. Acuto, O. & Michel, F. CD28-mediated co-stimulation: A quantitative support for TCR signalling. *Nat. Rev. Immunol.* **3**, 939–951 (2003).

40. Diehn, M. *et al.* Genomic expression programs and the integration of the CD28 costimulatory signal in T cell activation. *Proc. Natl. Acad. Sci.* **99**, 11796–11801 (2002).

41. Wang, D. *et al.* CD3/CD28 costimulation-induced NF-κB activation is mediated by recruitment of protein kinase C-theta, Bcl10, and IκB kinase β to the immunological synapse through CARMA1. *Mol. Cell. Biol.* **24**, 164–171 (2004).

42. Murphy, K. P., Travers, P., Walport, M. & Janeway, C. Chapter 7: Signaling Through Immune-System Receptors. in *Janeway's Immunobiology, 8th Edition,* 247–249 (2012).

43. Boon, T., Van Den Eynde, B., Van den Bruggen, P. & Van Pel, A. Tumor antigens recognized by T lymphocytes. *Annual review of immunology* **12**, 337–365 (1997).

44. Chen, D. S. & Mellman, I. Oncology meets immunology: The cancer-immunity cycle. *Immunity* **39**, 1–10 (2013).

45. Salgado, R. *et al.* The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: Recommendations by an International TILs Working Group 2014. *Ann. Oncol.* **26**, 259–271 (2015).

46. Buisseret, L. *et al.* Tumor-infiltrating lymphocyte composition, organization and PD-1/PD-L1 expression are linked in breast cancer. *Oncoimmunology* **6**, e1257452 (2017).

47. De Chaisemartin, L. *et al.* Characterization of chemokines and adhesion molecules associated with T cell presence in tertiary lymphoid structures in human lung cancer. *Cancer Res.* **71**, 6391–6399 (2011).

48. Gu-trantien, C. *et al.* CD4 + follicular helper T cell infiltration predicts breast cancer survival. *J. Clin. Invest.* **123**, 2873–2892 (2013).

49. Dieu-Nosjean, M. C. *et al.* Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. *J. Clin. Oncol.* **26**, 4410–4417 (2008).

50. Coppola, D. *et al.* Unique ectopic lymph node-like structures present in human primary colorectal carcinoma are identified by immune gene array profiling. *Am. J. Pathol. J. Pathol.* **179**, 37–45 (2011).

51. Fridman, W. H., Pagès, F., Sauts-Fridman, C. & Galon, J. The immune contexture in human tumours: Impact on clinical outcome. *Nat. Rev. Cancer* **12**, 298–306 (2012).

52. Geng, Y. *et al.* Prognostic role of tumor-infiltrating lymphocytes in lung cancer: A meta-analysis. *Cell. Physiol. Biochem.* **37**, 1560–1571 (2015).

53. Rosenberg, S. A., Spiess, P. & Lafreniere, R. A new approach to the adoptive immunotherapy of cancer with tumor-infiltrating lymphocytes. *Science* **233**, 1318–1321 (1986).

54. Lin, B. *et al.* Tumor-infiltrating lymphocytes: Warriors fight against tumors powerfully. *Biomed. Pharmacother.* **132**, 110873 (2020).

55. Becker A Schwinn R Dummer, J. C., Burg, G. & Brocker, E. Tumour-infiltrating lymphocytes in primary melanoma: functional consequences of differential IL-2 receptor expression. *Clin Exp Immunol* **91**, 121–125 (1993).

56. Dudley, M. E., Wunderlich, J. R., Shelton, T. E., Even, J. & Rosenberg, S. A. Generation of Tumor-Infiltrating Lymphocyte Cultures for Use in Adoptive Transfer Therapy for Melanoma Patients. *J. Immunother.* **26**, 332–342 (2003).

57. Stanton, S. E., Adams, S. & Disis, M. L. Variation in the Incidence and Magnitude of Tumor-Infiltrating Lymphocytes in Breast Cancer Subtypes: A Systematic Review. *JAMA Oncol.* **2**, 1354–1360 (2016).

58. Raskov, H., Orhan, A., Christensen, J. P. & Gögenur, I. Cytotoxic CD8+ T cells in cancer and cancer immunotherapy. *Br. J. Cancer* **124**, 359–367 (2021).

59. Basu, R. *et al.* Cytotoxic T Cells Use Mechanical Force to Potentiate Target Cell Killing. *Cell* **165**, 100–110 (2016).

60. Thiery, J. *et al.* Perforin pores in the endosomal membrane trigger the release of endocytosed granzyme B into the cytosol of target cells. *Nat. Immunol.* **12**, 770–777 (2011).

61. Kumar, A., Watkins, R. & Vilgelm, A. E. Cell Therapy With TILs: Training and Taming T Cells to Fight Cancer. *Front. Immunol.* **12**, 690499 (2021).

62. Bos, R. & Sherman, L. A. CD4+ T-cell help in the tumor milieu is required for recruitment and cytolytic function of CD8+ T lymphocytes. *Cancer Res.* **70**, 8368–8377 (2010).

63. Labani-Motlagh, A., Ashja-Mahdavi, M. & Loskog, A. The Tumor Microenvironment: A Milieu Hindering and Obstructing Antitumor Immune Responses. *Front. Immunol.* **11**, 940 (2020).

64. Kim, S. K. & Cho, S. W. The Evasion Mechanisms of Cancer Immunity and Drug Intervention in the Tumor Microenvironment. *Front. Pharmacol.* **13**, 868695 (2022).

65. Wu, Q., Yu, X., Li, J., Sun, S. & Tu, Y. Metabolic regulation in the immune response to cancer. *Cancer Commun.* **41**, 661–694 (2021).

66. Liu, Y. T. & Sun, Z. J. Turning cold tumors into hot tumors by improving T-cell infiltration. *Theranostics* **11**, 5265–5286 (2021).

67. Dhatchinamoorthy, K., Colbert, J. D. & Rock, K. L. Cancer Immune Evasion Through Loss of MHC Class I Antigen Presentation. *Front. Immunol.* **12**, 636568 (2021).

68. Tie, Y., Tang, F., Wei, Y. & Wei, X. Immunosuppressive cells in cancer: Mechanisms and potential therapeutic targets. *J. Hematol. Oncol.* **15**, 61 (2022).

69. Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J. & Schreiber, R. D. Cancer immunoediting: From immunosurveillance to tumor escape. *Nat. Immunol.* **3**, 991–998 (2002).

70. Dunn, G. P., Old, L. J. & Schreiber, R. D. The three Es of cancer immunoediting. *Annu. Rev. Immunol.* **22**, 329–360 (2004).

71. Laydon, D. J., Bangham, C. R. M. & Asquith, B. Estimating T-cell repertoire diversity: Limitations of classical estimators and a new approach. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **370**, 201402 (2015).

72. Schatz, D. G. & Ji, Y. Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* **11**, 251–263 (2011).

73. Hu, Q., Nicol, S. A., Suen, A. Y. W. & Baldwin, T. A. Examination of Thymic Positive and Negative Selection by Flow Cytometry. *J. Vis. Exp.* **68**, 4269 (2012).

74. Venanzi, E. S., Benoist, C. & Mathis, D. Good riddance: Thymocyte clonal deletion prevents autoimmunity. *Curr. Opin. Immunol.* **16**, 197–202 (2004).

75. Xing, Y. & Hogquist, K. A. T-Cell tolerance: Central and peripheral. *Cold Spring Harb. Perspect. Biol.* **4**, a006957 (2012).

76. Mathis, D. & Benoist, C. Back to Central Tolerance. *Immunity* **20**, 509–516 (2004).

77. Gotter, J., Brors, B., Hergenhahn, M. & Kyewski, B. Medullary Epithelial Cells of the Human Thymus Express a Highly Diverse Selection of Tissue-specific Genes Colocalized in Chromosomal Clusters. *J. Exp. Med.* **199**, 155–166 (2004).

78. Zamora, A. E., Crawford, J. C. & Thomas, P. G. Hitting the target: How T cells detect and eliminate tumors. *J. Immunol.* **200**, 392–399 (2018).

79. Akcakanat, A. *et al.* NY-ESO-1 expression and its serum immunoreactivity in esophageal cancer. *Cancer Chemother. Pharmacol.* **54**, 95–100 (2004).

80. Kawakami, Y. *et al.* Identification of the immunodominant peptides of the MART-1 human melanoma antigen recognized by the majority of HLA-A2-restricted tumor infiltrating lymphocytes. *J. Exp. Med.* **180**, 347–352 (1994).

81. Tsuji, T., Altorki, N. K., Ritter, G., Old, L. J. & Gnjatic, S. Characterization of preexisting MAGE-A3-specific CD4 + T Cells in cancer patients and healthy individuals and their activation by protein vaccination. *J. Immunol.* **183**, 4800–4808 (2009).

82. Haen, S. P., Löffler, M. W., Rammensee, H. G. & Brossart, P. Towards new horizons: characterization, classification and implications of the tumour antigenic repertoire. *Nat. Rev. Clin. Oncol.* **17**, 595–610 (2020).

83. Stone, J. D., Harris, D. T. & Kranz, D. M. TCR affinity for p/MHC formed by tumor antigens that are self-proteins: Impact on efficacy and toxicity. *Curr. Opin. Immunol.* **33**, 16–22 (2015).

84. Drent, E. *et al.* A Rational Strategy for Reducing On-Target Off-Tumor Effects of CD38-Chimeric Antigen Receptors by Affinity Optimization. *Mol. Ther.* **25**, 1946–1958 (2017).

85. Attermann, A. S., Bjerregaard, A. M., Saini, S. K., Grønbæk, K. & Hadrup, S. R. Human endogenous retroviruses and their implication for immunotherapeutics of cancer. *Ann. Oncol.* **29**, 2183–2191 (2018).

86. Bakker, A. B. *et al.* Identification of a novel peptide derived from the melanocyte-specific gp100 antigen as the dominant epitope recognized by an HLA-A2.1-restricted anti-melanoma CTL line. *Int. J. Cancer* **62**, 97–102 (1995).

87. Salgaller, M. L. *et al.* Recognition of multiple epitopes in the human melanoma antigen gp100 by peripheral blood lymphocytes stimulated in vitro with synthetic peptides. *Cancer Res.* **55**, 4972–4979 (1995).

88. Wölfel, T. *et al.* Two tyrosinase nonapeptides recognized on HLA-A2 melanomas by autologous cytolytic T lymphocytes. *Eur. J. Immunol* **24**, 759–764 (1994).

89. Maude, S. L., Teachey, D. T., Porter, D. L. & Grupp, S. A. CD19-targeted chimeric antigen receptor T-cell therapy for acute lymphoblastic leukemia. *Blood* **125**, 4017–4023 (2015).

90. Rodrigues, P. C. *et al.* Prognostic Significance of WT1 Gene Expression in Pediatric Acute Myeloid Leukemia. *Pediatr Blood Cancer* **49**, 133–138 (2007).

91. Iqbal, N. & Iqbal, N. Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications. *Mol. Biol. Int.* **2014**, 852748 (2014).

92. Reichelt, U. *et al.* Frequent homogeneous HER-2 amplification in primary and metastatic adenocarcinoma of the esophagus. *Mod. Pathol.* **20**, 120–129 (2007).

93. Morgan, R. A. *et al.* Case report of a serious adverse event following the administration of T cells transduced with a chimeric antigen receptor recognizing ERBB2. *Mol. Ther.* **18**, 843–851 (2010).

94. Gjerstorff, M. F., Andersen, M. H. & Ditzel, H. J. Oncogenic cancer/testis antigens: Prime candidates for immunotherapy. *Oncotarget* **6**, 15772–87 (2015).

95. Fratta, E. *et al.* The biology of cancer testis antigens: Putative function, regulation and therapeutic potential. *Mol. Oncol.* **5**, 164–182 (2011).

96. Van Der Bruggen, P. *et al.* A Gene Encoding an Antigen Recognized by Cytolytic T Lymphocytes on a Human Melanoma. *Science* **254**, 1643–1647 (1991).

97. De Plaen, E. *et al.* Structure, chromosomal localization, and expression of 12 genes of the MAGE family. *Immunogenetics* **40**, 360–369 (1994).

98. Gonzaga Almeida, L. *et al.* CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens. *Nucleic Acids Res.* **37**, 816–819 (2009).

99. Robbins, P. F. *et al.* Tumor regression in patients with metastatic synovial cell sarcoma and melanoma using genetically engineered lymphocytes reactive with NY-ESO-1. *J. Clin. Oncol.* **29**, 917–924 (2011).

100. Parkhurst, M. R. *et al.* T cells targeting carcinoembryonic antigen can mediate regression of metastatic colorectal cancer but induce severe transient colitis. *Mol. Ther.* **19**, 620–626 (2011).

101. Morgan, R. A. *et al.* Cancer regression and neurologic toxicity following anti-MAGE-A3 TCR gene therapy. *J Immunother* **36**, 133–151 (2013).

102. Bannert, N., Hofmann, H., Block, A. & Hohn, O. HERVs new role in cancer: From accused perpetrators to cheerful protectors. *Front. Microbiol.* **9**, 178 (2018).

103. Kassiotis, G. & Stoye, J. P. Immune responses to endogenous retroelements: Taking the bad with the good. *Nat. Rev. Immunol.* **16**, 207–219 (2016).

104. Young, G. R. *et al.* Negative selection by an endogenous retrovirus promotes a higher-avidity CD4+ T cell response to retroviral infection. *PLoS Pathog.* **8**, e1002709 (2012).

105. Seifarth, W. *et al.* Comprehensive analysis of human endogenous retrovirus transcriptional activity in human tissues with a retrovirus-specific microarray. *J. Virol.* **79**, 341–352 (2005).

106. Smith, C. C. *et al.* Alternative tumour-specific antigens. *Nat. Rev. Cancer* **19**, 465–478 (2019).

107. Lang, F., Schrörs, B., Löwer, M., Türeci, Ö. & Sahin, U. Identification of neoantigens for individualized therapeutic cancer vaccines. *Nat. Rev. Drug Discov.* **21**, 261–282 (2022).

108. Fritsch, E. F. *et al.* HLA-binding properties of tumor neoepitopes in humans. *Cancer Immunol. Res.* **2**, 522–529 (2014).

109. Rech, A. J. *et al.* Tumor immunity and survival as a function of alternative neopeptides in human cancer. *Cancer Immunol. Res.* **6**, 276–287 (2018).

110. Duan, F. *et al.* Genomic and bioinformatic profiling of mutational neoepitopes reveals new rules to predict anticancer immunogenicity. *J. Exp. Med.* **211**, 2231–2248 (2014).

111. Ghorani, E. *et al.* Differential binding affinity of mutated peptides for MHC class I is a predictor of survival in advanced lung cancer and melanoma. *Ann. Oncol.* **29**, 271–279 (2018).

112. Maby, P. *et al.* Correlation between density of CD8+ T-cell infiltrate in microsatellite unstable colorectal cancers and frameshift mutations: A rationale for personalized immunotherapy. *Cancer Res.* **75**, 3446–3455 (2015).

113. Tougeron, D. *et al.* Tumor-infiltrating lymphocytes in colorectal cancers with microsatellite instability are correlated with the number and spectrum of frameshift mutations. *Mod. Pathol.* **22**, 1186–1195 (2009).

114. Turajlic, S. *et al.* Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. *Lancet Oncol.* **18**, 1009–1021 (2017).

115. Worley, B. S. *et al.* Antigenicity of fusion proteins from sarcoma-associated chromosomal translocations. *Cancer Res.* **61**, 6868–6875 (2001).

116. Druker, B. J. *et al.* Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *The New England Journal of Medicine* **344**, 1031–1037 (2001).

117. Yu, Y. P. *et al.* Identification of recurrent fusion genes across multiple cancer types. *Sci. Rep.* **9**, 1074 (2019).

118. Weber, D. *et al.* Accurate detection of tumor-specific gene fusions reveals strongly immunogenic personal neo-antigens. *Nat. Biotechnol.* **40**, 1276–1284 (2022).

119. Feola, S., Chiaro, J., Martins, B. & Cerullo, V. Uncovering the tumor antigen landscape: What to know about the discovery process. *Cancers (Basel).* **12**, 1660 (2020).

120. Thompson, M. P. & Kurzrock, R. Epstein-Barr Virus and Cancer. *Clin. Cancer Res.* **10**, 803–821 (2004).

121. Draper, L. M. *et al.* Targeting of HPV-16+ epithelial cancer cells by TCR gene engineered T cells directed against E6. *Clin. Cancer Res.* **21**, 4431–4439 (2015).

122. Marijt, K. A., Doorduijn, E. M. & van Hall, T. TEIPP antigens for T-cell based immunotherapy of immune-edited HLA class I-low cancers. *Mol. Immunol.* **113**, 43–49 (2019).

123. Durgeau, A. *et al.* Different Expression Levels of the TAP Peptide Transporter Lead to Recognition of Different Antigenic Peptides by Tumor-Specific CTL. *J. Immunol.* **187**, 5532–5539 (2011).

124. Brändle, D., Brasseur, F., Weynants, P., Boon, T. & Van Den Eynde, B. J. A mutated HLA-A2 molecule recognized by autologous cytotoxic T lymphocytes on a human renal cell carcinoma. *J. Exp. Med.* **183**, 2501–2508 (1996).

125. Huang, J. *et al.* T Cells Associated with Tumor Regression Recognize Frameshifted Products of the CDKN2A Tumor Suppressor Gene Locus and a Mutated HLA Class I Gene Product. *J. Immunol.* **172**, 6057–6064 (2004).

126. Lee, S. C. W. *et al.* Modulation of splicing catalysis for therapeutic targeting of leukemia with mutations in genes encoding spliceosomal proteins. *Nat. Med.* **22**, 672–678 (2016).

127. Visconte, V., Makishima, H., MacIejewski, J. P. & Tiu, R. V. Emerging roles of the spliceosomal machinery in myelodysplastic syndromes and other hematological disorders. *Leukemia* **26**, 2447–2454 (2012).

128. Kahles, A. *et al.* Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer Cell* **34**, 211–224 (2018).

129. Diederichs, S. *et al.* The dark matter of the cancer genome: aberrations in regulatory elements, untranslated regions, splice sites, non-coding RNA and synonymous mutations. *EMBO Mol. Med.* **8**, 442–457 (2016).

130. Lozano-Rabella, M. *et al.* Exploring the Immunogenicity of Noncanonical HLA-I Tumor Ligands Identified through Proteogenomics. *Clin. Cancer Res.* **29**, 2250–2265 (2023).

131. Andersen, R. S. *et al.* High frequency of T cells specific for cryptic epitopes in melanoma patients. *Oncoimmunology* **2**, e25374 (2013).

132. Guilloux, Y. *et al.* A peptide recognized by human cytolytic T lymphocytes on HLA-A2 melanomas is encoded by an intron sequence of the N-acetylglucosaminyltransferase V gene. *J. Exp. Med.* **183**, 1173–1183 (1996).

133. Harada, Mamoru; Li, Yong F.; El-Gamil, Mona; Ohnmacht, Galen A.; Rosenberg, Steven A.; Robbins, P. F. Melanoma-reactive CD8+ T Cells recognize a novel tumor antigen expressed in a wide variety of tumor types. *J. Immunother.* **24**, 323–333 (2001).

134. Laumont, C. M. *et al.* Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat. Commun.* **7**, 10238 (2016).

135. Scull, K. E., Pandey, K., Ramarathinam, S. H. & Purcell, A. W. Immunopeptidogenomics: Harnessing RNA-Seq to illuminate the dark immunopeptidome. *Mol. Cell. Proteomics* **20**, 100143 (2021).

136. Erhard, F., Dölken, L., Schilling, B. & Schlosser, A. Identification of the cryptic HLA-I immunopeptidome. *Cancer Immunol. Res.* **8**, 1018–1026 (2020).

137. Schwendenwein, A. *et al.* Molecular profiles of small cell lung cancer subtypes: therapeutic implications. *Mol. Ther. - Oncolytics* **20**, 470–483 (2021).

138. Knight, S. B. *et al.* Progress and prospects of early detection in lung cancer. *Open Biol.* **7**, 170070 (2017).

139. Pikor, L. A., Ramnarine, V. R., Lam, S. & Lam, W. L. Genetic alterations defining NSCLC subtypes and their therapeutic implications. *Lung Cancer* **82**, 179–189 (2013).

140. Roy S. Herbst, John V. Heymach, S. M. L. Lung cancer. *N. Engl. J. Med.* **359**, 1367–1380 (2008).

141. Goldstraw, P. *et al.* The IASLC lung cancer staging project: Proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM Classification for lung cancer. *J. Thorac. Oncol.* **11**, 39–51 (2016).

142. Mithoowani, H. & Febbraro, M. Non-Small-Cell Lung Cancer in 2022: A Review for General Practitioners in Oncology. *Curr. Oncol. 2022* **29**, 1828–1839 (2022).

143. Pisters, K. M. W. *et al.* Cancer Care Ontario and American Society of clinical oncology adjuvant chemotherapy and adjuvant radiation therapy for stages I-IIIA resectable non-small-cell lung cancer guideline. *J. Clin. Oncol.* **25**, 5506–5518 (2007).

144. Felip, E. *et al.* Adjuvant atezolizumab after adjuvant chemotherapy in resected stage IB–IIIA non-small-cell lung cancer (IMpower010): a randomised, multicentre, open-label, phase 3 trial. *Lancet* **398**, 1344–1357 (2021).

145. Faivre-finn, C. *et al.* Four-year survival with Durvalumab after chemoradiotherapy in stage III NSCLC - an update from the PACIFIC Trial. *J. Thorac. Oncol.* **16**, 860–867 (2021).

146. Antonia, S. J. *et al.* Durvalumab after Chemoradiotherapy in Stage III Non–Small-Cell Lung Cancer. *N. Engl. J. Med.* **377**, 1919–1929 (2017).

147. Rothschild, S. I. *et al.* SAKK 16/14: Durvalumab in addition to neoadjuvant chemotherapy in patients with stage IIIA(N2) non–small-cell lung cancer - A multicenter single-arm phase II trial. *J. Clin. Oncol.* **39**, 2872–2880 (2021).

148. Mayekar, M. K. & Bivona, T. G. Current Landscape of Targeted Therapy in Lung Cancer. *Clin. Pharmacol. Ther.* **102**, 757–764 (2017).

149. Reck, M. *et al.* Five-year outcomes with pembrolizumab versus chemotherapy for metastatic non-small-cell lung cancer with PD-L1 tumor proportion score ≥ 50%. *J Clin Oncol* **39**, 2339–2349 (2021).

150. Gou, L. Y. & Wu, Y. L. Prevalence of driver mutations in non-small-cell lung cancers in the People's Republic of China. *Lung Cancer Targets Ther.* **5**, 1–9 (2014).

151. Hanna, N. H. *et al.* Therapy for Stage IV Non-Small-Cell Lung Cancer With Driver Alterations: ASCO and OH (CCO) Joint Guideline Update. *J Clin Oncol* **39**, 1040–1091 (2021).

152. Lindeman, N. I. *et al.* Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: Guideline from the College of American Pathologists, the International Association for the Study of Lung Cancer, and the . *Arch. Pathol. Lab. Med.* **142**, 321–346 (2018).

153. Filippone, A. *et al.* PD1/PD-L1 immune checkpoint as a potential target for preventing brain tumor progression. *Cancer Immunol. Immunother.* **71**, 2067–2075 (2022).

154. Rizvi, N. A. *et al.* Activity and safety of nivolumab, an anti-PD-1 immune checkpoint inhibitor, for patients with advanced, refractory squamous non-small-cell lung cancer (CheckMate 063): A phase 2, single-arm trial. *Lancet Oncol.* **16**, 257–265 (2015).

155. Herbst, R. S. *et al.* Atezolizumab for First-Line Treatment of PD-L1–Selected Patients with NSCLC. *N. Engl. J. Med.* **383**, 1328–1339 (2020).

156. Sezer, A. *et al.* Cemiplimab monotherapy for first-line treatment of advanced non-small-cell lung cancer with PD-L1 of at least 50%: A multicentre, open-label, global, phase 3, randomised, controlled trial. *Lancet* **397**, 592–604 (2021).

157. Gadgeel, S., Rodriguez-Abreu, D., Speranza, G., Esteban, E. & Felip, E. Updated Analysis From KEYNOTE-189: Pembrolizumab or Placebo Plus Pemetrexed and Platinum for Previously Untreated Metastatic Nonsquamous Non-Small-Cell Lung Cancer. *J. Clin. Oncol.* **38**, 1505–1517 (2020).

158. Reck, M. *et al.* First-line nivolumab plus ipilimumab with two cycles of chemotherapy versus chemotherapy alone (four cycles) in advanced non-small-cell lung cancer: CheckMate 9LA 2-year update. *ESMO Open* **6**, 100273 (2021).

159. BioNTech. mRNA Programs: BNT116 (FixVac). *Second Quarter Financial Results and Corporate Update* (2022). Available at: https://investors.biontech.de/node/13401/pdf.

160. Shields, M. D., Marin-Acevedo, J. A. & Pellini, B. Immunotherapy for Advanced Non–Small Cell Lung Cancer: A Decade of Progress. *Am. Soc. Clin. Oncol. Educ. B.* **41**, 105–127 (2021).

161. Wang, D. Y. *et al.* Fatal Toxic Effects Associated With Immune Checkpoint Inhibitors: A Systematic Review and Meta-analysis. *JAMA Oncol.* **4**, 1721–1728 (2018).

162. Zhang, S., Liang, F., Li, W. & Wang, Q. Risk of treatment-related mortality in cancer patients treated with ipilimumab: A systematic review and meta-analysis. *Eur. J. Cancer* **83**, 71–79 (2017).

163. Ricciuti, B. *et al.* Impact of immune-related adverse events on survival in patients with advanced non-small cell lung cancer treated with nivolumab: Long-term outcomes from a multi-institutional analysis. *J. Cancer Res. Clin. Oncol.* **145**, 479–485 (2019).

164. Met, Ö., Jensen, K. M., Chamberlain, C. A., Donia, M. & Svane, I. M. Principles of adoptive T cell therapy in cancer. *Semin. Immunopathol.* **41**, 49–58 (2019).

165. Morotti, M. *et al.* Promises and challenges of adoptive T-cell therapies for solid tumours. *Br. J. Cancer* **124**, 1759–1776 (2021).

166. Andersen, R. *et al.* Long-Lasting complete responses in patients with metastatic melanoma after adoptive cell therapy with tumor-infiltrating lymphocytes and an attenuated IL2 regimen. *Clin. Cancer Res.* **22**, 3734–3745 (2016).

167. Rosenberg, Steven A. Packard, B. S. *et al.* Use of Tumor-Infiltrating Lymphocytes and Interleukin-2 in the Immunotherapy of Patients with Metastatic Melanoma. *N. Engl. J. Med.* **319**, 1676–1680 (1988).

168. Rosenberg, S. A. *et al.* Durable Complete Responses in Heavily Pretreated Patients with Metastatic Melanoma Using T Cell Transfer Immunotherapy. *Clin. Cancer Res.* **17**, 4550–4557 (2012).

169. Besser, M. J. *et al.* Adoptive transfer of tumor-infiltrating lymphocytes in patients with metastatic melanoma: Intent-to-treat analysis and efficacy after failure to prior immunotherapies. *Clin. Cancer Res.* **19**, 4792–4800 (2013).

170. Andersen, R. *et al.* T cells isolated from patients with checkpoint inhibitor-resistant melanoma are functional and can mediate tumor regression. *Ann. Oncol.* **29**, 1575–1581 (2018).

171. Pedersen, M. *et al.* Adoptive cell therapy with tumor-infiltrating lymphocytes in patients with metastatic ovarian cancer: A pilot study. *Oncoimmunology* **7**, e1502905 (2018).

172. Morgan, R. A. *et al.* Cancer regression in patients after transfer of genetically engineered lymphocytes. *Science* **314**, 126–129 (2006).

173. Parkhurst, M. R. *et al.* T cells targeting carcinoembryonic antigen can mediate regression of metastatic colorectal cancer but induce severe transient colitis. *Mol. Ther.* **19**, 620–626 (2011).

174. Johnson, L. A. *et al.* Gene therapy with human and mouse T-cell receptors mediates cancer regression and targets normal tissues expressing cognate antigen. *Blood* **114**, 535–546 (2009).

175. Rapoport, A. P. *et al.* NY-ESO-1-specific TCR-engineered T cells mediate sustained antigen-specific antitumor effects in myeloma. *Nat. Med.* **21**, 914–921 (2015).

176. John, L. B. *et al.* Anti-PD-1 antibody therapy potently enhances the eradication of established tumors by gene-modified T cells. *Clin. Cancer Res.* **19**, 5636–5646 (2013).

177. Turtle, C. J. *et al.* Durable molecular remissions in chronic lymphocytic leukemia treated with CD19-Specific chimeric antigen Receptor-modified T cells after failure of ibrutinib. *J. Clin. Oncol.* **35**, 3010–3020 (2017).

178. Kochenderfer, J. N. *et al.* Chemotherapy-refractory diffuse large B-cell lymphoma and indolent B-cell malignancies can be effectively treated with autologous T cells expressing an anti-CD19 chimeric antigen receptor. *J. Clin. Oncol.* **33**, 540–549 (2015).

179. Neelapu, S. S. *et al.* Axicabtagene Ciloleucel CAR T-Cell Therapy in Refractory Large B-Cell Lymphoma. *N. Engl. J. Med.* **377**, 2531–2544 (2017).

180. Park, J. H. *et al.* Long-Term Follow-up of CD19 CAR Therapy in Acute Lymphoblastic Leukemia. *N. Engl. J. Med.* **378**, 449–459 (2018).

181. Brudno, J. N. *et al.* T cells genetically modified to express an anti–B-Cell maturation antigen chimeric antigen receptor cause remissions of poor-prognosis relapsed multiple myeloma. *J. Clin. Oncol.* **36**, 2267–2280 (2018).

182. Sengsayadeth, S., Savani, B. N., Oluwole, O. & Dholaria, B. Overview of approved CAR-T therapies, ongoing clinical trials, and its impact on clinical practice. *EJHaem* **3 (Sup. 1)**, 6–10 (2021).

183. U.S. Food & Drug Administration (FDA). FDA approves ciltacabtagene autoleucel for relapsed or refractory multiple myeloma. (2022). Available at: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-ciltacabtagene-autoleucel-relapsed-or-refractory-multiple-myeloma.

184. Reinhard, K. *et al.* An RNA vaccine drives expansion and efficacy of claudin-CAR-T cells against solid tumors. *Science* **367**, 446–453 (2020).

185. Casucci, M., Hawkins, R. E., Dotti, G. & Bondanza, A. Overcoming the toxicity hurdles of genetically targeted T cells. *Cancer Immunol Immunother* **64**, 123–130 (2015).

186. Linette, G. P. *et al.* Cardiovascular toxicity and titin cross-reactivity of affinity-enhanced T cells in myeloma and melanoma. *Blood* **122**, 863–871 (2013).

187. Manfredi, F. *et al.* TCR Redirected T Cells for Cancer Treatment: Achievements, Hurdles, and Goals. *Front. Immunol.* **11**, 1689 (2020).

188. Rosenberg, S. A. & Restifo, N. P. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science* **348**, 62–68 (2015).

189. Robbins, P. F. *et al.* Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells. *Nat. Med.* **19**, 747–752 (2013).

190. Lu, Y. C. *et al.* Efficient identification of mutated cancer antigens recognized by T cells associated with durable tumor regressions. *Clin. Cancer Res.* **20**, 3401–3410 (2014).

191. Ramakrishna, V. *et al.* Naturally occuring peptides associated with HLA-A2 in ovarian cancer cell lines identified by mass spectrometry are targets of HLA-A2-restricted cytotoxic T cells. *Int. Immunol.* **15**, 751–763 (2003).

192. Coulie, P. G. *et al.* A new gene coding for a differentiation antigen recognized by autologous cytolytic T lymphocytes on HLA-A2 melanomas. *J. Exp. Med.* **180**, 35–42 (1994).

193. Adderley, H., Blackhall, F. H. & Lindsay, C. R. KRAS-mutant non-small cell lung cancer: Converging small molecules and immune checkpoint inhibition. *EBioMedicine* **41**, 711–716 (2019).

194. Downward, J. Targeting RAS signalling pathways in cancer therapy. *Nat. Rev. Cancer* **3**, 11–22 (2003).

195. Chen, K., Zhang, Y., Qian, L. & Wang, P. Emerging strategies to target RAS signaling in human cancer therapy. *J. Hematol. Oncol.* **14**, 116 (2021).

196. Cox, A. D., Fesik, S. W., Kimmelman, A. C., Luo, J. & Der, C. J. Drugging the undruggable RAS: Mission Possible? *Nat. Rev. Drug Discov.* **13**, 828–851 (2014).

197. Cai, D. *et al.* The prevalence and prognostic value of KRAS co-mutation subtypes in Chinese advanced non-small cell lung cancer patients. *Cancer Med.* **9**, 84–93 (2020).

198. Huang, L., Guo, Z., Wang, F. & Fu, L. KRAS mutation: From undruggable to druggable in cancer. *Signal Transduct. Target. Ther.* **6**, 386 (2021).

199. Jänne, P. A. *et al.* Adagrasib in Non–Small-Cell Lung Cancer Harboring a KRAS G12C Mutation. *N. Engl. J. Med.* **387**, 120–131 (2022).

200. Hong, D. S. *et al.* KRAS G12C Inhibition with Sotorasib in Advanced Solid Tumors. *N. Engl. J. Med.* **383**, 1207–1217 (2020).

201. Canon, J. *et al.* The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature* **575**, 217–223 (2019).

202. Camidge, D. R. *et al.* A phase III multicenter study of sotorasib (AMG 510), a KRAS(G12C) inhibitor, versus docetaxel in patients with previously treated advanced non-small cell lung cancer (NSCLC) harboring KRAS p.G12C mutation. *Ann. Oncol.* **31**, 894–895 (2020).

203. Nagasaka, M. *et al.* KRAS Inhibitors - yes but what next? Direct targeting of KRAS - vaccines, adoptive T cell therapy and beyond. *Cancer Treat. Rev.* **101**, 102309 (2021).

204. Chatani, P. D. & Yang, J. C. Mutated Ras: Targeting the “untargetable” with T cells. *Clin. Cancer Res.* **26**, 537–544 (2020).

205. Wang, Q. J. *et al.* Identification of T-cell receptors targeting KRAS-mutated human tumors. *Cancer Immunol. Res.* **4**, 204–214 (2016).

206. Simoni, Y. *et al.* Bystander CD8+ T cells are abundant and phenotypically distinct in human tumour infiltrates. *Nature* **557**, 575–579 (2018).

207. Duhen, T. *et al.* Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat. Commun.* **9**, 2724 (2018).

208. Gros, A. *et al.* PD-1 identifies the patient-specific CD8+ tumor-reactive repertoire infiltrating human tumors. *J. Clin. Invest.* **124**, 2246–2259 (2014).

209. Draghi, A. *et al.* Rapid identification of the tumor-specific reactive TIL repertoire via combined detection of CD137, TNF, and IFNγ, following recognition of autologous tumor-antigens. *Front. Immunol.* **12**, 705422 (2021).

210. Melief, J., Wickström, S., Kiessling, R. & Pico De Coañ, Y. Assessment of Antitumor T-Cell Responses by Flow Cytometry After Coculture of Tumor Cells with Autologous Tumor-Infiltrating Lymphocytes. *Methods Mol. Biol.* **1913**, 133–140 (2019).

211. Schmittel, A., Keilholz, U., Thiel, E. & Scheibenbogen, C. Quantification of Tumor-Specific T Lymphocytes with the ELISPOT Assay. *J. Immunother.* **23**, 289–295 (2000).

212. Mellor-Heineke, S. *et al.* Elevated granzyme B in cytotoxic lymphocytes is a signature of immune activation in hemophagocytic lymphohistiocytosis. *Front. Immunol.* **4**, 72 (2013).

213. Betts, M. R. *et al.* Sensitive and viable identification of antigen-specific CD8+ T cells by a flow cytometric assay for degranulation. *J. Immunol. Methods* **281**, 65–78 (2003).

214. Howie, B. *et al.* High-throughput pairing of T cell receptor α and β sequences. *Sci. Transl. Med.* **7**, 301ra131 (2015).

215. Pasetto, A. *et al.* Tumor- and neoantigen-reactive T-cell receptors can be identified based on their frequency in fresh tumor. *Cancer Immunol. Res.* **4**, 734–743 (2017).

216. Li, H., Ye, C., Ji, G. & Han, J. Determinants of public T cell responses. *Cell Res.* **22**, 33–42 (2012).

217. Song, I. *et al.* Broad TCR Repertoire And Diverse Structural Solutions To Recognition Of An Immunodominant CD8 T Cell Epitope. *Nat Struct Mol Biol* **24**, 395–406 (2017).

218. Huisman, W. *et al.* Public T-Cell Receptors (TCRs) Revisited by Analysis of the Magnitude of Identical and Highly-Similar TCRs in Virus-Specific T-Cell Repertoires of Healthy Individuals. *Front. Immunol.* **13**, 851868 (2022).

219. Huang, H., Wang, C., Rubelt, F., Scriba, T. J. & Davis, M. M. Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat. Biotechnol.* **38**, 1194–1202 (2020).

220. Chiou, S. *et al.* Global analysis of shared T cell specificities in human non-small cell lung cancer enables HLA inference and antigen discovery. *Immunity* **54**, 586–602 (2021).

221. Chronister, W. D. *et al.* TCRMatch: Predicting T-Cell Receptor Specificity Based on Sequence Similarity to Previously Characterized Receptors. *Front. Immunol.* **12**, 640725 (2021).

222. Jokinen, E., Huuhtanen, J., Mustjoki, S., Heinonen, M. & Lähdesmäki, H. Predicting recognition between T cell receptors and epitopes with TCRGP. *PLoS Comput. Biol.* **17**, e1008814 (2021).

223. Tran, E. *et al.* T-Cell Transfer Targeting Mutant KRAS in Cancer. *N. Engl. J. Med.* **375**, 2255–2262 (2016).

224. Jiang, C. *et al.* Influence of T Cell-Mediated Immune Surveillance on Somatic Mutation Occurrences in Melanoma. *Front. Immunol.* **12**, 703821 (2022).

225. Kropp, K. N. *et al.* A bicistronic vector backbone for rapid seamless cloning and chimerization of αβT-cell receptor sequences. *PLoS One* **15**, e0238875 (2020).

226. Schrörs, B. *et al.* HLA class I loss in metachronous metastases prevents continuous T cell recognition of mutated neoantigens in a human melanoma model. *Oncotarget* **8**, 28312–28327 (2017).

227. Cohen, C. J. *et al.* Isolation of neoantigen-specific T cells from tumor and peripheral lymphocytes. *J. Clin. Invest.* **125**, 3981–3991 (2015).

228. Lennerz, V. *et al.* The response of autologous T cells to a human melanoma is dominated by mutated neoantigens. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 16013–16018 (2005).

229. Inozume, T. *et al.* Selection of CD8+PD-1+ lymphocytes in fresh human melanomas enriches for tumor-reactive T cells. *J. Immunother.* **33**, 956–964 (2010).

230. Gros, A. *et al.* Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients. *Nat. Med.* **22**, 433–438 (2016).

231. Parkhurst, M. *et al.* Isolation of T-cell receptors specifically reactive with mutated tumor-associated antigens from tumor-infiltrating lymphocytes based on CD137 expression. *Clin. Cancer Res.* **23**, 2491–2505 (2017).

232. Scheper, W. *et al.* Low and variable tumor reactivity of the intratumoral TCR repertoire in human cancers. *Nat. Med.* **25**, 89–94 (2019).

233. Kreiter, S. *et al.* Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature* **520**, 692–696 (2015).

234. Linnemann, C. *et al.* High-throughput epitope discovery reveals frequent recognition of neo-antigens by CD4+ T cells in human melanoma. *Nat. Med.* **21**, 81–85 (2015).

235. Tran, E. *et al.* Cancer Immunotherapy Based on Mutation-Specific CD4+ T Cells in a Patient with Epithelial Cancer. *Science* **344**, 641–645 (2014).

236. Ahmadzadeh, M. *et al.* Tumor-infiltrating human CD4+ regulatory T cells display a distinct TCR repertoire and exhibit tumor and neoantigen reactivity. *Sci. Immunol.* **4**, eaao4310 (2019).

237. Karasaki, T. *et al.* Prediction and prioritization of neoantigens: integration of RNA sequencing data with whole-exome sequencing. *Cancer Sci.* **108**, 170–177 (2017).

238. Barnell, E. K. *et al.* Standard operating procedure for somatic variant refinement of sequencing data with paired tumor and normal samples. *Genet. Med.* **21**, 972–981 (2019).

239. Wang, C., Xue, W., Zhang, H. & Fu, Y. Identification of candidate genes encoding tumor-specific neoantigens in early- and late-stage colon adenocarcinoma. *Aging (Albany. NY).* **13**, 4024–4044 (2021).

240. Hashimoto, S. *et al.* Neoantigen prediction in human breast cancer using RNA sequencing data. *Cancer Sci.* **112**, 465–475 (2021).

241. Tang, X. *et al.* The eSNV-detect: A computational system to identify expressed single nucleotide variants from transcriptome sequencing data. *Nucleic Acids Res.* **42**, e172 (2014).

242. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).

243. Hart, T., Komori, H. K., LaMere, S., Podshivalova, K. & Salomon, D. R. Finding the active genes in deep RNA-seq gene expression studies. *BMC Genomics* **14**, 778 (2013).

244. Strom, S. P. Current practices and guidelines for clinical next-generation sequencing oncology testing. *Cancer Biol. Med.* **13**, 3–11 (2016).

245. Richters, M. M. *et al.* Best practices for bioinformatic characterization of neoantigens for clinical utility. *Genome Med.* **11**, 56 (2019).

246. Paul, S. *et al.* HLA class I alleles are associated with peptide-binding repertoires of different size, affinity, and immunogenicity. *J. Immunol.* **191**, 5831–5839 (2013).

247. Sette, A. *et al.* The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes. *J. Immunol.* **153**, 5586–5592 (1994).

248. Kotturi, M. F. *et al.* The CD8 + T-Cell Response to Lymphocytic Choriomeningitis Virus Involves the L Antigen: Uncovering New Tricks for an Old Virus. *J. Virol.* **81**, 4928–4940 (2007).

249. Boegel, S. *et al.* HLA typing from RNA-Seq sequence reads. *Genome Med.* **4**, 102 (2012).

250. Ilca, F. T., Drexhage, L. Z., Brewin, G., Peacock, S. & Boyle, L. H. Distinct Polymorphisms in HLA Class I Molecules Govern their Susceptibility to Peptide Editing by TAPBPR. *Cell Rep.* **29**, 1621–1632 (2019).

251. Harndahl, M. *et al.* Peptide-MHC class I stability is a better predictor than peptide affinity of CTL immunogenicity. *Eur. J. Immunol.* **42**, 1405–1416 (2012).

252. Rasmussen, M. *et al.* Pan-Specific Prediction of Peptide–MHC Class I Complex Stability, a Correlate of T Cell Immunogenicity. *J. Immunol.* **197**, 1517–1524 (2016).

253. Calis, J. J. A. *et al.* Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. *PLoS Comput. Biol.* **9**, e1003266 (2013).

254. Rajasagi, M. *et al.* Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. *Blood* **124**, 453–462 (2014).

255. Wick, D. A. *et al.* Surveillance of the tumor mutanome by T cells during progression from primary to recurrent ovarian cancer. *Clin. Cancer Res.* **20**, 1125–1134 (2014).

256. Zhao, W. & Sher, X. Systematically benchmarking peptide-MHC binding predictors: From synthetic to naturally processed epitopes. *PLOS Comput. Biol.* **14**, e1006457 (2018).

257. Andreatta, M. & Nielsen, M. Gapped sequence alignment using artificial neural networks: Application to the MHC class I system. *Bioinformatics* **32**, 511–517 (2016).

258. Moutaftsi, M. *et al.* A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus. *Nat. Biotechnol.* **24**, 817–819 (2006).

259. Jurtz, V. *et al.* NetMHCpan 4.0: Improved peptide-MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *Physiol. Behav.* **199**, 3360–3368 (2017).

260. Weeder, B. R., Wood, M. A., Li, E., Nellore, A. & Thompson, R. F. Pepsickle Rapidly and Accurately Predicts Proteasomal Cleavage Sites for Improved Neoantigen Identification. *Bioinformatics* **37**, 3723–3733 (2021).

261. Calis, J. J. A., Reinink, P., Keller, C., Kloetzel, P. M. & Keşmir, C. Role of peptide processing predictions in T cell epitope identification: contribution of different prediction programs. *Immunogenetics* **67**, 85–93 (2015).

262. Rasmussen, M. *et al.* Pan-Specific Prediction of Peptide–MHC Class I Complex Stability, a Correlate of T Cell Immunogenicity. *J. Immunol.* **197**, 1517–1524 (2016).

263. Karosiene, E., Lundegaard, C., Lund, O. & Nielsen, M. NetMHCcons: A consensus method for the major histocompatibility complex class I predictions. *Immunogenetics* **64**, 177–186 (2012).

264. Zhang, H., Lund, O. & Nielsen, M. The PickPocket method for predicting binding specificities for receptors based on receptor pocket similarities: Application to MHC-peptide binding. *Bioinformatics* **25**, 1293–1299 (2009).

265. Vita, R. *et al.* The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* **47**, 339–343 (2019).

266. Li, Q. & Ding, Z. Y. The Ways of Isolating Neoantigen-Specific T Cells. *Front. Oncol.* **10**, 1347 (2020).

267. Lu, Y. C. *et al.* An Efficient Single-Cell RNA-Seq Approach to Identify Neoantigen-Specific T Cell Receptors. *Mol. Ther.* **26**, 379–389 (2018).

268. Cohen, C. J., Zhao, Y., Zheng, Z., Rosenberg, S. A. & Morgan, R. A. Enhanced Antitumor Activity of Murine-Human Hybrid T-Cell Receptor (TCR) in Human Lymphocytes is Associated with Improved Pairing and TCR/CD3 Stability. *Cancer Res.* **66**, 8878–8886 (2006).

269. Legut, M., Dolton, G., Mian, A. A., Ottmann, O. G. & Sewell, A. K. CRISPR-mediated TCR replacement generates superior anticancer transgenic T cells. *Blood* **131**, 311–322 (2018).

270. Ahmadi, M. *et al.* CD3 limits the efficacy of TCR gene therapy in vivo. *Blood* **118**, 3528–3537 (2011).

271. Morton, L. T. *et al.* Simultaneous Deletion of Endogenous TCRαβ for TCR Gene Therapy Creates an Improved and Safe Cellular Therapeutic. *Mol. Ther.* **28**, 64–74 (2020).

272. Rosenberg, S. A. *et al.* Gene Transfer into Humans - Immunotherapy of Patients with Advanced Melanoma, Using Tumor-Infiltrating Lymphocytes Modified by Retroviral Gene Transduction. *N. Engl. J. Med.* **323**, 570–578 (1990).

273. Müller, T. R. *et al.* Targeted T cell receptor gene editing provides predictable T cell product function for immunotherapy. *Cell Reports Med.* **2**, 100374 (2021).

274. Riddell, S. R. & Greenberg, P. D. United States Patent 6040177: High efficiency transduction of T lymphocytes using rapid expansion methods ('REM'). (2000).

275. Weinstein-Marom, H. *et al.* Genetic Modification of Tumor-Infiltrating Lymphocytes via Retroviral Transduction. *Front. Immunol.* **11**, 584148 (2021).

276. Li, Y. & Kurlander, R. J. Comparison of anti-CD3 and anti-CD28-coated beads with soluble anti-CD3 for expanding human T cells: Differing impact on CD8 T cell phenotype and responsiveness to restimulation. *J. Transl. Med.* **8**, 7126–7140 (2010).

277. Yamashita, M. & Emerman, M. Retroviral infection of non-dividing cells: Old and new perspectives. *Virology* **344**, 88–93 (2006).

278. Liu, M. & Eiden, M. V. Role of human endogenous retroviral long terminal repeats (LTRs) in maintaining the integrity of the human germ line. *Viruses* **3**, 901–905 (2011).

279. Schober, K., Müller, T. R. & Busch, D. H. Orthotopic T-Cell Receptor Replacement - An "Enabler" for TCR-Based Therapies. *Cells* **9**, 1367 (2020).

280. Wei, F., Cheng, X. X., Xue, J. Z. & Xue, S. A. Emerging Strategies in TCR-Engineered T Cells. *Front. Immunol.* **13**, 850358 (2022).

281. Rezalotfi, A., Fritz, L., Förster, R. & Bošnjak, B. Challenges of CRISPR-Based Gene Editing in Primary T Cells. *Int. J. Mol. Sci.* **23**, 1689 (2022).

282. Loupakis, F. *et al.* Prediction of Benefit from Checkpoint Inhibitors in Mismatch Repair Deficient Metastatic Colorectal Cancer: Role of Tumor Infiltrating Lymphocytes. *Oncologist* **25**, 481–487 (2020).

283. Chalmers, Z. R. *et al.* Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* **9**, 34 (2017).

284. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015).

285. Jardim, D. L., Goodman, A., Gagliato, D. de M. & Kurzrock, R. The Challenges of Tumor Mutational Burden as an Immunotherapy Biomarker. *Cancer Cell* **39**, 154–173 (2021).

286. Govindan, R. *et al.* Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* **150**, 1121–1134 (2012).

287. Van Allen, E. M. *et al.* Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–212 (2016).

288. Wolfl, M. *et al.* Activation-induced expression of CD137 permits detection, isolation, and expansion of the full repertoire of CD8+ T cells responding to antigen without requiring knowledge of epitope specificities. *Blood* **110**, 201–210 (2007).

289. Turcotte, S. *et al.* Tumor-Reactive CD8+ T cells in metastatic gastrointestinal cancer refractory to chemotherapy. *Clin. Cancer Res.* **20**, 331–343 (2014).

290. Hobbs, G. A., Der, C. J. & Rossman, K. L. RAS isoforms and mutations in cancer at a glance. *J. Cell Sci.* **129**, 1287–1292 (2016).

291. Muñoz-Maldonado, C., Zimmer, Y. & Medová, M. A comparative analysis of individual RAS mutations in cancer biology. *Front. Oncol.* **9**, 1088 (2019).

292. Davari, K. *et al.* Development of a CD8 co-receptor independent T-cell receptor specific for tumor-associated antigen MAGE-A4 for next generation T-cell-based immunotherapy. *J. Immunother. cancer* **9**, e002035 (2021).

293. Stolze, B., Reinhart, S., Bulllinger, L., Fröhling, S. & Scholl, C. Comparative analysis of KRAS codon 12, 13, 18, 61, and 117 mutations using human MCF10A isogenic cell lines. *Sci. Rep.* **5**, 8535 (2014).

294. Martinov, T. & Greenberg, P. D. Targeting Driver Oncogenes and Other Public Neoantigens Using T Cell Receptor-Based Cellular Therapy. *Annu. Rev. Cancer Biol.* **7**, 331–351 (2023).

295. Prior, I. A., Lewis, P. D. & Mattos, C. A comprehensive Survey of Ras mutations in Cancer. *Cancer Res.* **72**, 2457–2467 (2012).

296. Sweeney, S. M. *et al.* AACR Project Genie: Powering Precision Medicine through an International Consortium. *Cancer Discov.* **7**, 818–831 (2017).

297. Gebregiworgis, T. *et al.* The Q61H mutation decouples KRAS from upstream regulation and renders cancer cells resistant to SHP2 inhibitors. *Nat. Commun.* **12**, 6274 (2021).

298. Jančík, S., Drábek, J., Radzioch, D. & Hajdúch, M. Clinical relevance of KRAS in human cancers. *J. Biomed. Biotechnol.* **2010**, 150960 (2010).

299. Huynh, M. V. *et al.* Functional and biological heterogeneity of KRASQ61 mutations. *Sci. Signal.* **15**, eabn2694 (2022).

300. Linard, B. *et al.* A ras-Mutated Peptide Targeted by CTL Infiltrating a Human Melanoma Lesion. *J. Immunol.* **168**, 4802–4808 (2002).

301. Prior, I. A., Hood, F. E. & Hartley, J. L. The frequency of RAS mutations in cancer. *Cancer Res.* **80**, 2669–2974 (2020).

302. Wang, Q. *et al.* Direct detection and quantification of neoantigens. *Cancer Immunol. Res.* **7**, 1748–1754 (2019).

303. Zhou, Z. W. *et al.* KRASQ61H preferentially signals through MAPK in a RAF dimer-dependent manner in non-small cell lung cancer. *Cancer Res.* **80**, 3719–3731 (2020).

304. Levin, N. *et al.* Identification and validation of T-cell receptors targeting RAS hotspot mutations in human cancers for use in cell-based immunotherapy. *Clin. Cancer Res.* **27**, 5084–5095 (2021).

305. Li, D. *et al.* Genetically engineered T cells for cancer immunotherapy. *Signal Transduct. Target. Ther.* **4**, 35 (2019).

306. Shafer, P., Kelly, L. M. & Hoyos, V. Cancer Therapy With TCR-Engineered T Cells: Current Strategies, Challenges, and Prospects. *Front. Immunol.* **13**, 835762 (2022).

307. Chandran, S. S. & Klebanoff, C. A. T cell receptor-based cancer immunotherapy: Emerging efficacy and pathways of resistance. *Immunol. Rev.* **290**, 127–147 (2019).

308. Baulu, E., Gardet, C., Chuvin, N. & Depil, S. TCR-engineered T cell therapy in solid tumors: State of the art and perspectives. *Sci. Adv.* **9**, eadf3700 (2023).

309. Pang, Z. *et al.* Neoantigen-targeted TCR-engineered T cell immunotherapy: current advances and challenges. *Biomark. Res.* **11**, 104 (2023).

310. Douglass, J. *et al.* Bispecific antibodies targeting mutant RAS neoantigens. *Sci. Immunol.* **6**, eabd5515 (2021).

311. Cole, D. *et al.* Identification of MART-1-specific T-cell receptors: T cells utilizing distinct T-cell receptor variable and joining regions recognize the same tumor epitope. *Cancer Res.* **54**, 5265–8 (1994).

312. Derré, L. *et al.* Distinct sets of αβ TCRs confer similar recognition of tumor antigen NY-ESO-1157-165 by interacting with its central Met/Trp residues. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15010–15015 (2008).

313. Sim, M. J. W. *et al.* High-affinity oligoclonal TCRs define effective adoptive T cell therapy targeting mutant KRAS-G12D. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 12826–12835 (2020).

314. Borg, N. A. *et al.* The CDR3 regions of an immunodominant T cell receptor dictate the 'energetic landscape' of peptide-MHC recognition. *Nat. Immunol.* **6**, 171–180 (2005).

315. Goyarts, E. *et al.* Point mutations in the beta chain CDR3 can alter the T cell receptor recognition pattern on an MHC class I/peptide complex over a broad interface area. *Mol. Immunol.* **35**, 593–607 (1998).

316. Glanville, J. *et al.* Identifying specificity groups in the T cell receptor repertoire. *Nature* **547**, 94–98 (2017).

317. Ulbrecht, M. *et al.* Cell surface expression of HLA-E: Interaction with human β2-microglobulin and allelic differences. *Eur. J. Immunol.* **29**, 537–547 (1999).

318. Pietra, G., Romagnani, C., Manzini, C., Moretta, L. & Mingari, M. C. The emerging role of HLA-E-restricted CD8+ T lymphocytes in the adaptive immune response to pathogens and tumors. *J. Biomed. Biotechnol.* **2010**, 907092 (2010).

319. Sharpe, H. R., Bowyer, G., Brackenridge, S. & Lambe, T. HLA-E: exploiting pathogen-host interactions for vaccine development. *Clin. Exp. Immunol.* **196**, 167–177 (2019).

320. Kanevskiy, L. *et al.* Dimorphism of HLA-E and its disease association. *Int. J. Mol. Sci.* **20**, 5496 (2019).

321. Vietzen, H. *et al.* HLA-E-Restricted Immune Responses are Crucial for the Control of Epstein-Barr-Virus Infections and the Prevention of Post-Transplant Lymphoproliferative Disorders. *Blood* **41**, 1560–1573 (2022).

322. Shimizu, Y. & DeMars, R. Production of human cells expressing individual transferred HLA-A,-B,-C genes using an HLA-A,-B,-C null human cell line. *J. Immunol.* **142**, 3320–3328 (1989).

323. Hò, G. G. T., Heinen, F. J., Huyton, T., Blasczyk, R. & Bade-Döding, C. HLA-F*01:01 presents peptides with N-terminal flexibility and a preferred length of 16 residues. *Immunogenetics* **71**, 353–360 (2019).

324. Jasinski-Bergner, S., Schmiedel, D., Mandelboim, O. & Seliger, B. Role of HLA-G in Viral Infections. *Front. Immunol.* **13**, 826074 (2022).

325. Brusic, V., Petrovsky, N., Zhang, G. & Bajic, V. B. Prediction of promiscuous peptides that bind HLA class I molecules. *Immunol. Cell Biol.* **80**, 280–285 (2002).

326. Leslie, A. *et al.* Differential Selection Pressure Exerted on HIV by CTL Targeting Identical Epitopes but Restricted by Distinct HLA Alleles from the Same HLA Supertype. *J. Immunol.* **177**, 4699–4708 (2006).

327. Ueno, T., Tomiyama, H. & Takiguchi, M. Single T Cell Receptor-Mediated Recognition of an Identical HIV-Derived Peptide Presented by Multiple HLA Class I Molecules. *J. Immunol.* **169**, 4961–4969 (2002).

328. Burrows, S. R. *et al.* Promiscuous CTL Recognition of Viral Epitopes on Multiple Human Leukocyte Antigens: Biological Validation of the Proposed HLA A24 Supertype. *J. Immunol.* **171**, 1407–1412 (2003).

329. Sabbaj, S. *et al.* Cross-Reactive CD8+ T Cell Epitopes Identified in US Adolescent Minorities. *J. Acquir. Immune Defficiency Syndr.* **33**, 426–438 (2003).

330. Masemola, A. M. *et al.* Novel and Promiscuous CTL Epitopes in Conserved Regions of Gag Targeted by Individuals with Early Subtype C HIV Type 1 Infection from Southern Africa. *J. Immunol.* **173**, 4607–4617 (2004).

331. Frahm, N. *et al.* Extensive HLA class I allele promiscuity among viral CTL epitopes. *Eur. J. Immunol.* **37**, 2419–2433 (2007).

332. Rao, X., Hoof, I., Fontaine Costa, A. I. C. A., Van Baarle, D. & Keşmir, C. HLA class I allele promiscuity revisited. *Immunogenetics* **63**, 691–701 (2011).

333. Herrod, G., Dow, L. W. & Sullivan, L. Persistent Epstein-Barr Virus Infection Mimicking Juvenile Chronic Myelogenous Leukemia: Immunologic and Hematologic Studies. *Blood* **61**, 1098–1104 (1983).

334. Edwards, B. D. *et al.* Clinical outcomes associated with Escherichia coli infections in adults with cystic fibrosis: A cohort study. *Open Forum Infect. Dis.* **7**, ofz476 (2020).

335. Wang, J. Y. *et al.* Recurrent infections and chronic colonization by an Escherichia coli clone in the respiratory tract of a patient with severe cystic bronchiectasis. *J. Clin. Microbiol.* **38**, 2766–2767 (2000).

336. Khanna, R., Burrows, S. R., Moss, D. J. & Silins, S. L. Peptide transporter (TAP-1 and TAP-2)-independent endogenous processing of Epstein-Barr virus (EBV) latent membrane protein 2A: implications for cytotoxic T-lymphocyte control of EBV-associated malignancies. *J. Virol.* **70**, 5357–5362 (1996).

337. Bessell, C. A. *et al.* Commensal bacteria stimulate antitumor responses via T cell cross-reactivity. *JCI Insight* **5**, e135597 (2020).

338. Voss, F. K. *et al.* Cross-reactivity between tumor MHC class I – restricted antigens and an enterococcal bacteriophage. *Science* **369**, 936–942 (2020).

339. Choi, I. K. *et al.* Mechanism of EBV inducing anti-tumour immunity and its therapeutic use. *Nature* **590**, 157–162 (2021).

340. Iheagwara, U. K. *et al.* Influenza virus infection elicits protective antibodies and T cells specific for host cell antigens also expressed as tumor-associated antigens: a new view of cancer immunosurveillance. *Cancer Immunol. Res.* **2**, 263–273 (2014).

341. Klein, E. *et al.* Properties of the K562 cell line, derived from a patient with chronic myeloid leukemia. *Int. J. Cancer* **18**, 421–31 (1976).

342. Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* **36**, 59–74 (1977).

343. Gerard, R. D. & Gluzman, Y. New host cell system for regulated simian virus 40 DNA replication. *Mol Cell Biol* **5**, 3231–3240 (1985).

344. Swift, S., Lorens, J., Achacoso, P. & Nolan, G. P. Rapid production of retroviruses for efficient gene delivery to mammalian cells using 293T cell-based systems. *Curr. Protoc. Immunol.* **Chapter 10**, Unit 10.17C (2001).

345. Soneoka, Y. *et al.* A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Res.* **23**, 628–33 (1995).

346. Tan, P. H. *et al.* Changing viral tropism using immunoliposomes alters the stability of gene expression: implications for viral vector design. *Mol. Med. (Cambridge, Mass)* **13**, 216–226 (2007).

347. Gazdar, A. F. & Oie, H. K. Growth of cell lines and clinical specimens of human non-small cell lung cancer in a serum-free defined medium. *Cancer Res.* **46**, 6011–6012 (1986).

348. Ikediobi, O. N. *et al.* Mutation analysis of 24 known cancer genes in the NCI-60 cell line set. *Mol. Cancer Ther.* **5**, 2606–2612 (2006).

349. Pontén, F., Jirström, K. & Uhlen, M. The Human Protein Atlas - A tool for pathology. *J. Pathol.* **216**, 387–393 (2008).

350. Bagaev, D. V. *et al.* VDJdb in 2019: Database extension, new analysis infrastructure and a T-cell receptor motif compendium. *Nucleic Acids Res.* **48**, 1057–1062 (2020).

351. Bernard, P. & Couturier, M. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J. Mol. Biol.* **226**, 735–745 (1992).

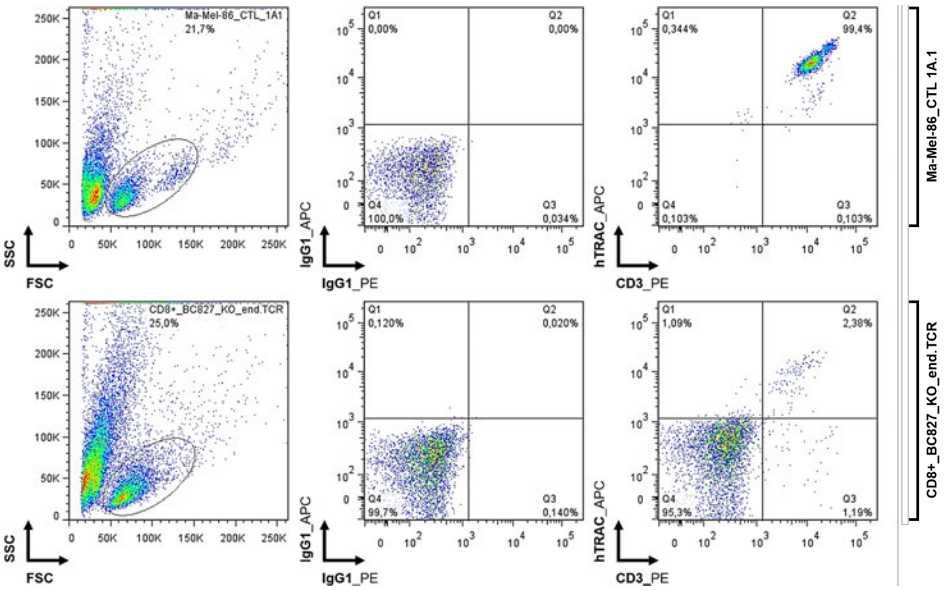
352. Roth, T. L. *et al.* Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature* **559**, 405–409 (2018).

353. Seitz, V. *et al.* A new method to prevent carry-over contaminations in two-step PCR NGS library preparations. *Nucleic Acids Res.* **43**, e135 (2015).

354. Seitz, V. *et al.* Evidence for a role of RUNX1 as recombinase cofactor for TCRβ rearrangements and pathological deletions in ETV6-RUNX1 ALL. *Sci. Rep.* **10**, 10024 (2020).

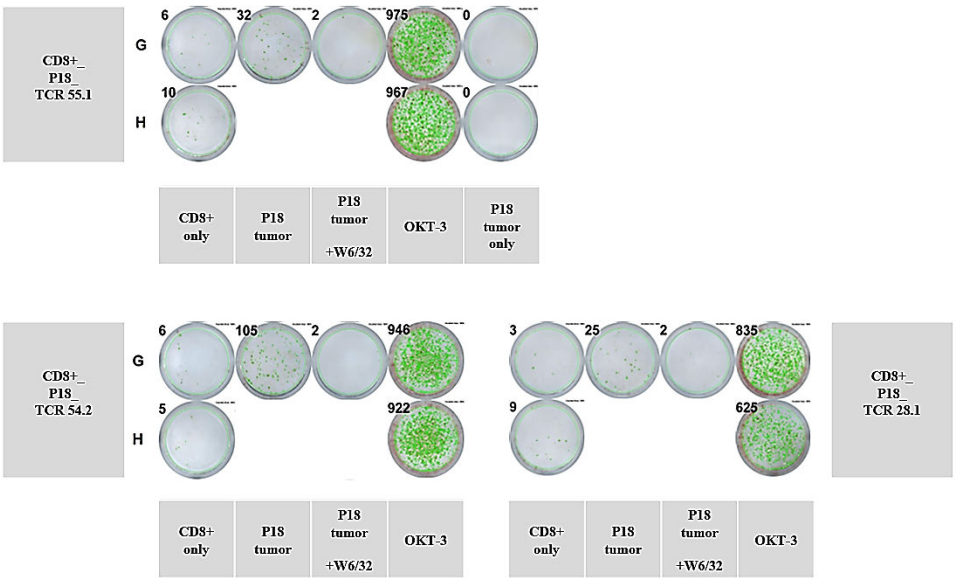
7.2 Supplementary Figures

Supplementary Figure 1:



Sup. Fig. 1: CRISPR/Cas9-based KO of endogenous TCR expression in CD8+ donor lymphocytes derived from BC827.
Endogenously expressed TCRs of BC827-derived T cell populations were knocked out before retroviral transduction with P18 TCR expression constructs. When compared to a control CTL clone (CTL 1A.1) derived from the peripheral blood of melanoma patient Ma-Mel-86 (provided by the Wölfel lab), endogenous TCR expression was absent in more than 95% of all treated BC827 donor lymphocytes. TCR expression was determined via hTRAC staining. Flow cytometry was performed 18 days after CRISPR/Cas9 treatment. Flow cytometry data was gated based on viable T cell populations and non-specific isotype (IgG1) controls. ***Abbr.:*** *APC = Allophycocyanin; BC = buffy coat; CTL = cytotoxic T lymphocyte; FSC = forward scatter; hTRAC = human T cell receptor alpha chain constant region; KO = knock-out; PE = Phycoerythrin; SSC = sideward scatter; TCR = T cell receptor.*

Supplementary Figure 2:



Sup. Fig. 2: Raw ELISpot data for P18 tumor recognition shown in Figure 3.2.3.
IFN-γ secretion was used as a marker for T cell recognition. ELISpot assays were performed 16 days after retroviral transduction of BC827-derived CD8+ T cell cultures with P18 TCR expression constructs. BC827-derived T cell populations had before been used for a CRISPR/Cas9-based KO of their endogenous TCR expression. Effector numbers per well were calculated based on mTRAC staining performed on day 13 after retroviral transduction (Fig. 3.2.2). Potential IFN-γ secretions of P18 tumor cell suspensions were assessed by analyzing P18 tumor cells in the absence of effector cells (P18 tumor only preparation). IFN-γ background levels were assessed by analyzing transgenic T cell populations in the absence of any target cells (CD8+only preparations) and were then used for data normalization. General T cell functionality was assessed by analyzing IFN-γ secretion capacities of transgenic T cell populations upon non-specific activation with OKT-3. For final evaluation, non-used wells and wells containing irrelevant antigen testing information were excluded from plate scans. Due to the poor quality of autologous tumor cell suspensions, P18 recognition analyses could only be performed in single wells. All performed control analyses were conducted as duplicates. Used effector cells per well: 10,000 mTRAC+ TCR-T cells. Used P18 tumor cells per well: 16,000. W6/32 concentration per well: 50 µg/ml. ***Abbr.:*** *BC = buffy coat; KO = knock-out; mTRAC = murine T cell receptor alpha constant region; TCR = T cell receptor.*

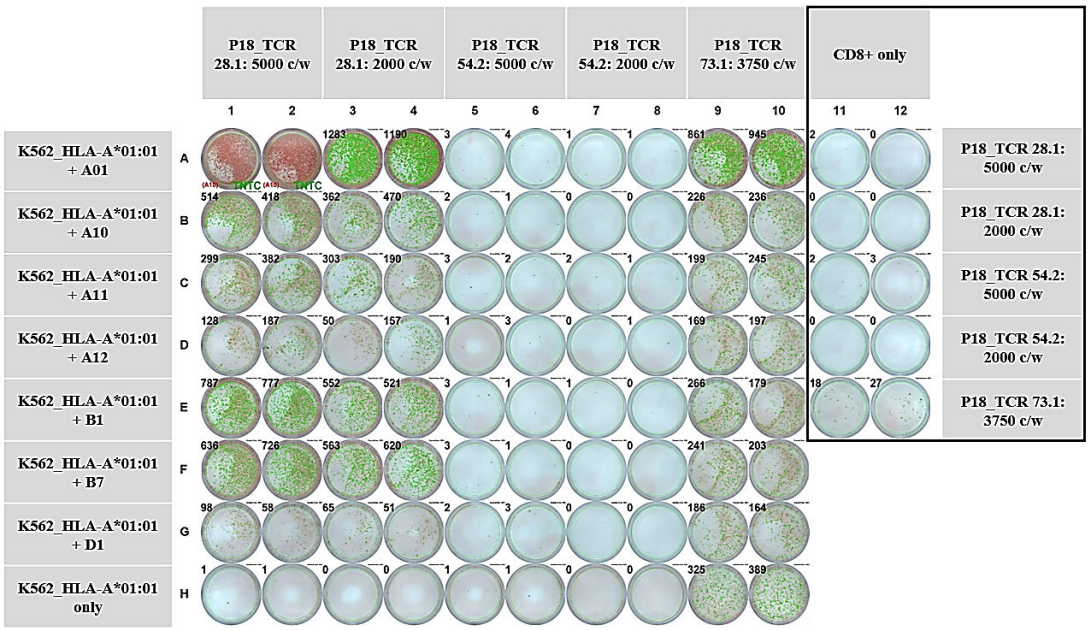
Supplementary Figure 3:

HLA-A*01:01	HLA-A*02:01	HLA-B*08:01	HLA-B*40:02	HLA-C*03:04	HLA-C*07:01	Controls
ILDTAGHEEY *KRAS*	AMYLTYSFHL *MFSD12*	CRLRRRRARA *GP1BB*	FETLLYRFSF *LARGE2*	FLSSATTTL *BRD9*	TRSNRAVYW *DIS3L2*	YYSKNLNSFF *MMS22L*
LSELQPGVQY *FN1*	FVLAPICLV *SMO*	FPLLRGSPF *LBHD1*	REHTNLYHV *DNMBP*	KAFNRSSTL *ZFN66*	SSSSYQRLW *FNIP2*	VYQYTFPDF *INSIG1*
FTVKRNWYPY *HERC4*	SMLEALRGV *KDELC2*	RLRRRRARA *GP1BB*	YEMELDELL *PTPRD*	LSMQNNSAF *BRD9*	RRFVNVVTT *FAU*	
VVPAFETLLY *LARGE2*	TLSELQPGV *FN1*	FSKNHQTQSL *MUC4*	KEASPSWSA *TICRR*	MAMYLTYSF *MFSD12*	FSYTYLHYG *SVEP1*	
LSDILGVHL *ZNF528*	GMAESLLGL *SLC25A22*	LWLRSQTTSL *FNIP2*	AEKFGGSLIV *DSTN*	YAAQNLLSHL *CCNT1*	LRGVTNDLL *KDELC2*	
WWALGCLPY *GRK6*	HSMLEALRGV *KDELC2*	CIAEKFGGSL *DSTN*	REHTNLYHVP *DNMBP*	SSATTTLSM *BRD9*	DRNTFQHSV *TP53*	
NTSFFPLLR *LBHD1*	LGMAESLLGL *SLC25A22*	LRRRRARARA *GP1BB*	LEDTASIFFL *CDK18*	HSIHAAYQL *TSHZ3*	ARAAARALV *HGH1*	
YLHYGQTVTY *SVEP1*	HLLGHQLPPA *ZNF528*	TMKETSLGPL *TMEM94*	WESPYEMEL *PTPRD*	AAARALVNL *HGH1*	ERSPHSPCE *ZSWIM8*	
LSDILGVHLL *ZNF528*	FVLAPICLVL *SMO*	MAKLHDINA *AP2B1*	FEERENSNPF *MKRN1*	ITHSMLEAL *KDELC2*	SSYQRLWLR *FNIP2*	
ILNGKFSYTY *SVEP1*	VLAPICLVLI *SMO*	LCRLRRRRA *GP1BB*	SEEEAVTII *VCAN*	SASDVSDVI *HTRA1*	FYCREHTNLY *DNMBP*	
LTYSFHLPK *MFSD12*	FLGMAESLL *SLC25A22*	MPAKLFRFA *RANBP1*	KETSLGPLSC *TMEM94*	YTGQMINVL *TAF5L*	NRSSTLTKH *ZNF66*	
VPAFETLLY *LARGE2*	GMAESLLGLL *SLC25A22*	MLRGRGACY *TYW1*	AEKFGGSLI *DSTN*	IAEKFGGSL *DSTN*	WRTPPPRSL *SEMA6C*	
GACYKHTFY *TYW1*	VTLSELQPGV *FN1*	TLLYRFSFP *LARGE2*	EELFKMPAKL *RANBP1*	VASLHGPIF *XPO4*	RRNHFWELF *MKRN1*	
FIKPWESPY *PTPRD*	KQDHYAWGL *DNAH17*	LRRRRARAR *GP1BB*	IEVLRINNT *LRP2*	NTFQHSVVV *TP53*	IRNCYLPHC *RPAP1*	
ASDVSDVIK *HTRA1*	KMNSTLGGV *ADCY5*	NFYCREHTNL *DNMBP*	QEGATSVVL *AC124319.1*	AGHEEYSAM *KRAS*	RRRARARAR *GP1BB*	
ETSLGPLSCV *TMEM94*	YMAMYLTYSF* *MFSD12*			TAGHEEYSAM *KRAS*	GRLVLPQLL *SPOCK2*	

Sup. Fig. 3: HLA class I peptide binding prediction results for NSCLC patient P18.

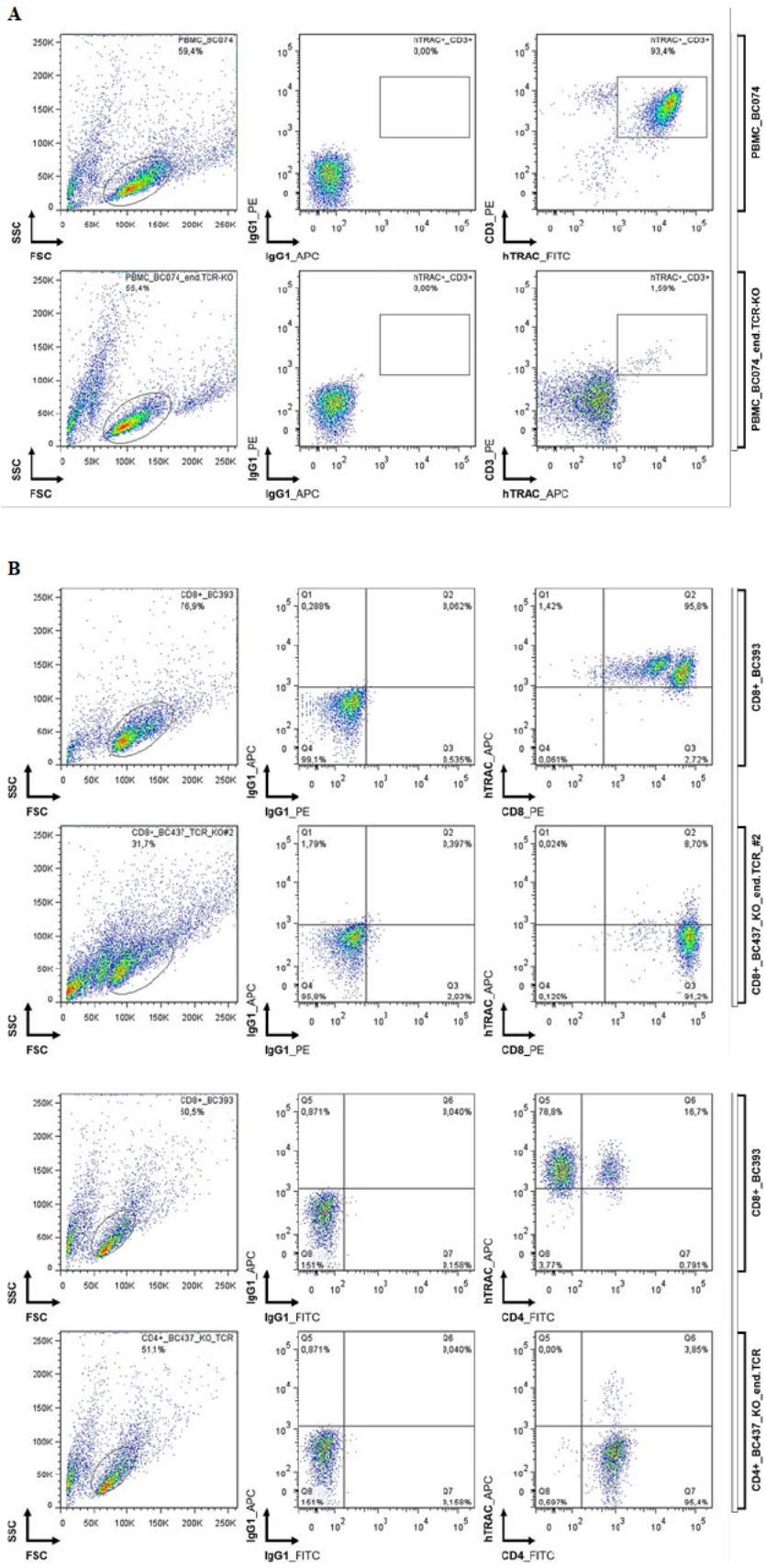
*In silico HLA binding predictions were individually performed for each P18 HLA-I allele by using mutated 19-mer peptides, each of which represented one of the 74 identified non-synonymous variations. P18 HLA class I binding predictions resulted in a total of 581 candidate neoantigen 9- and 10-mer peptides. The 14-16 potentially best binding neopeptide candidates were synthesized (JPT Peptide Technologies, Berlin, Germany) for each P18 HLA-I allele. 9-/10-mer sequences are shown in bold letters with mutated amino acids being highlighted in red. Corresponding gene annotations are depicted in italic font. Neopeptides predicted for all HLA-I alleles except for HLA-C*07 and HLA-B*40 are additionally marked with an asterisk. In addition, two control peptides were synthesized to assess the overall quality of the peptide library. Both control peptides were previously shown to be specifically recognized by CTL cultures initially obtained from the melanoma patient Ma-Mel-86 [226]: CTL 16C/114 specifically recognized YYSKNLNSFF while CTL 3A/115 recognized VYQYTFPDF. Hence, the overall quality of the peptide library was verified before starting any neopeptide screenings.* ***Abbr.:*** *CTL = cytotoxic T lymphocyte; HLA = human leukocyte antigen; Mel = melanoma; NSCLC = non-small cell lung cancer.*

Supplementary Figure 4:



Sup. Fig. 4: Raw ELISpot data for neopeptide recognition screening shown in Fig. 3.2.5 B.
*IFN-γ secretion was used as a marker for T cell recognition. ELISpot assays were performed 33 days after retroviral transduction of BC827-derived CD8+ T cells with P18 TCR expression constructs. Used T cell populations had before been used for a CRISPR/Cas9-based KO of their endogenous TCR expression (Sup. Fig. 1). Effector numbers per well were calculated based on previously performed mTRAC staining. Different concentrations of transgenic CD8+ T cell populations were tested per well. Target recognition analyses were performed as duplicates and used effector cell numbers are stated for each well. IFN-γ background levels were assessed by analyzing TCR-T cell populations in the absence of target cells (CD8+only preparations) and were subsequently used for data normalization. Non-used wells and wells containing irrelevant antigen screening information were excluded from plate scans for final evaluation. Used target cells per well: 50,000. Neopeptide concentrations per well: 2 µg/ml. Tested peptides (mutated AA highlighted): A01 = KRAS (ILDTAG**H**EEY); A10 = SVEP (ILNGKFSYT**Y**); A11 = MFSD12 (LTYS**F**HLPK); A12 = LARGE2 (VPAFETL**L**Y); B1 = TYW1 (G**A**CYKHTFY); B7 = KDELC2 (SMLEA**L**RGV); D1 =FNIP2 (**L**WLRSQTTSL). Retrospective corrections of inaccurately counted wells are highlighted in green font.* ***Abbr.:*** *BC = buffy coat; c/w = cells/well, FNIP2 = Folliculin Interacting Protein 2; H = histidine; HLA = human leukocyte antigen; KDELC2 = Protein O-Glucosyltransferase 3; KRAS = Kirsten rat sarcoma viral oncogene; Q = glutamine; LARGE2 = LARGE Xylosyl- and Glucuronyltransferase 2; MFSD12 = Major Facilitator Superfamily Domain-Containing Protein 12; mTRAC = murine T cell receptor alpha constant region; SVEP = Sushi, Von Willebrand Factor Type A, EGF And Pentraxin Domain Containing; TCR = T cell receptor; TYW1 = TRNA-YW Synthesizing Protein 1 Homolog.*

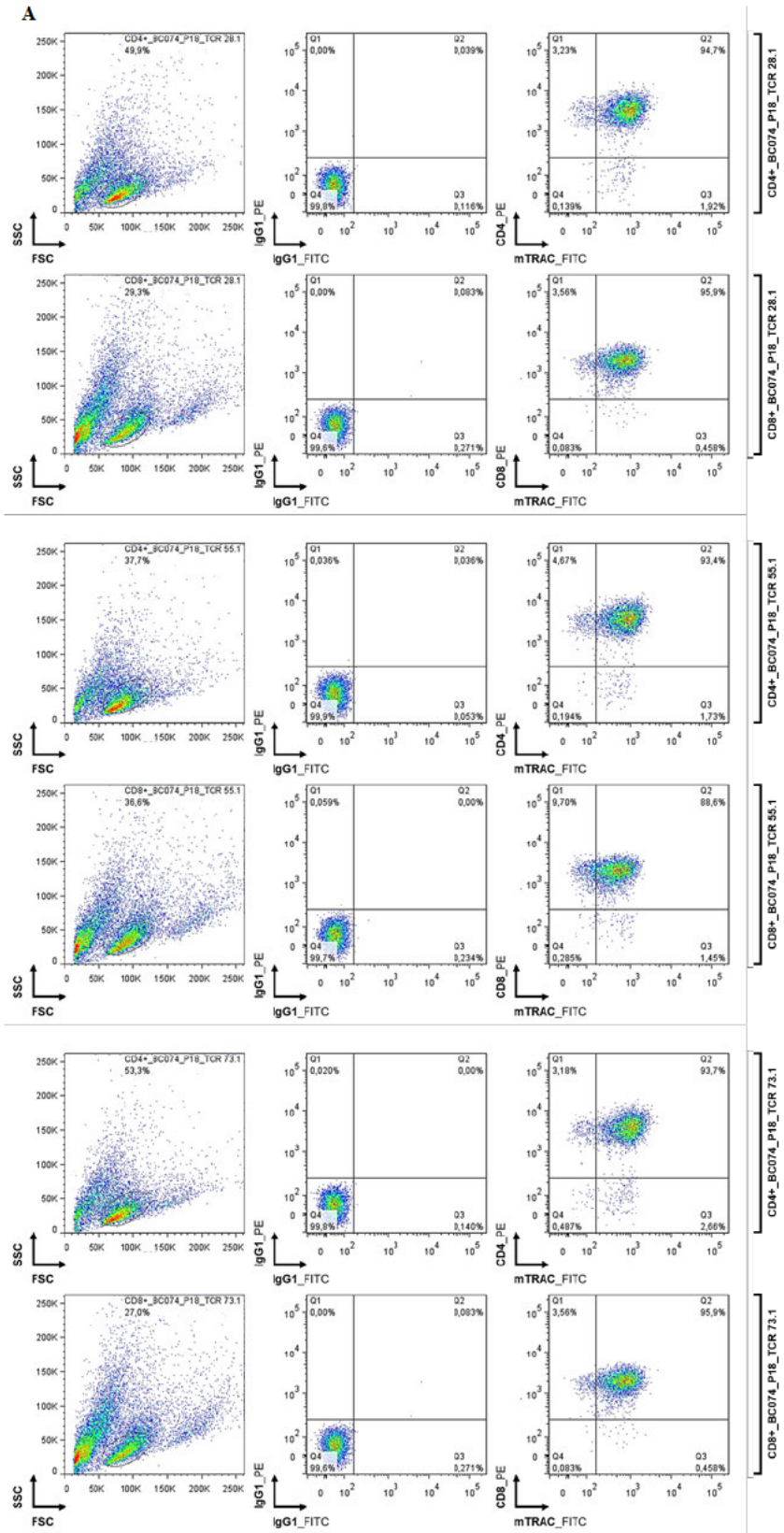
Supplementary Figure 5:



Sup. Fig. 5: CRISPR/Cas9-based KO of endogenous TCR expression in donor lymphocyte populations from BCs 074 and 437.

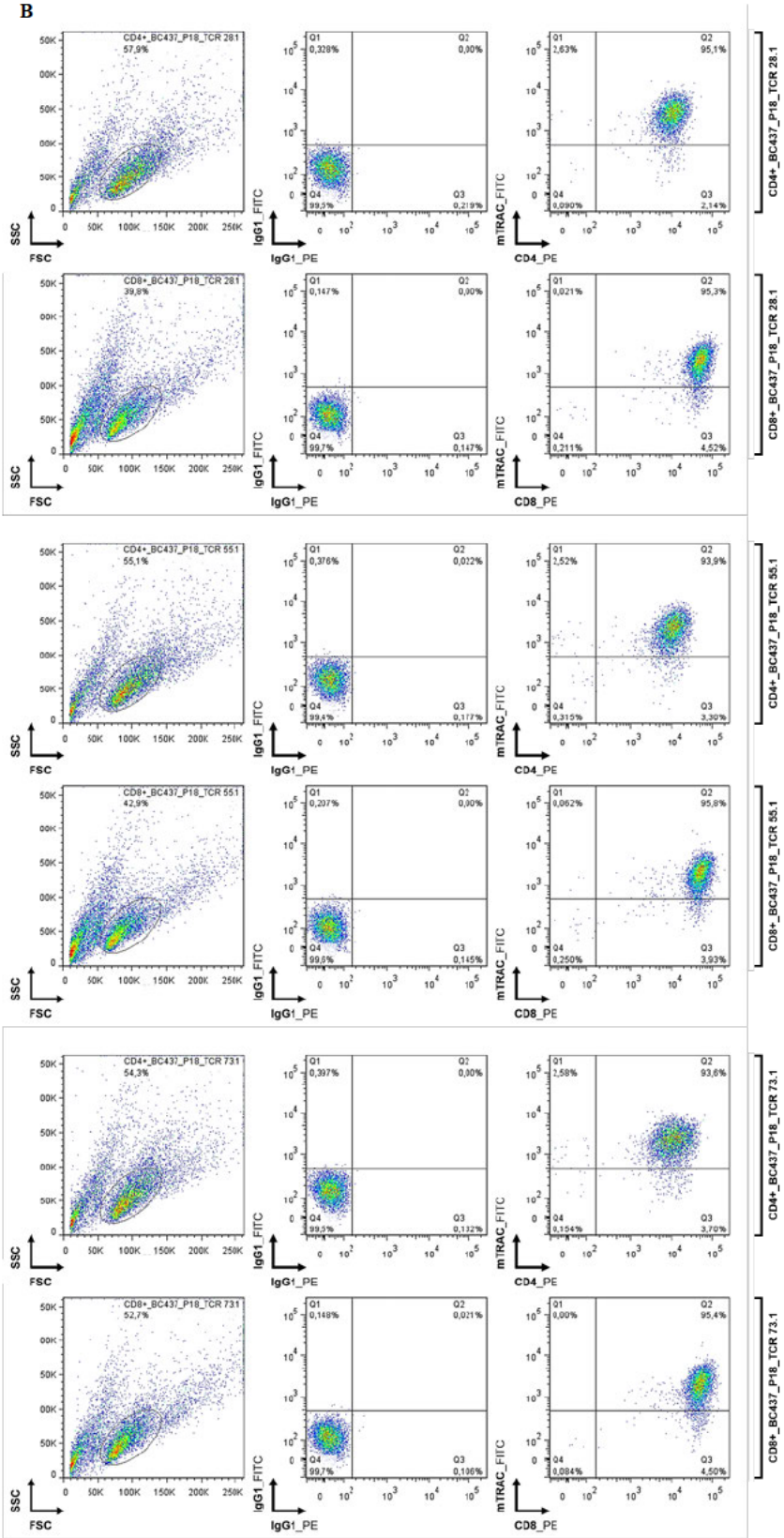
Endogenous TCR expression was knocked out in human T cell populations before retroviral transduction with P18 TCR expression constructs. When compared to non-treated control populations exhibiting >93% of human TCR expression, the performed KO approaches dramatically reduced endogenous TCR expression in both BCs: TCR expression in BC437 CD4+ populations was reduced to less than 4% while approx. 9% of cells in corresponding CD8+ populations expressed the residual human TCR. Two consecutive KO approaches were performed for BC437 CD8+ T cells since the first CRISPR/Cas9 treatment resulted in residual endogenous TCR expression of more than 14% (data not shown). In contrast, CRISPR/Cas9-treated BC074 PBMCs exhibited endogenous TCR expression of less than 2% when compared to non-treated control populations. Residual human TCR expression was determined via hTRAC staining by using a monoclonal antibody specifically binding the human TCR constant region. Flow cytometry staining was performed 10 days (CD4+ BC437), 14 days (CD8+ CB437) and 8 days (BC074) after CRISPR/Cas9 treatment, respectively. Flow cytometry data was gated based on viable T cell populations, non-specific isotype (IgG1) controls, and hTRAC staining of non-treated control T cell populations. ***Abbr.:*** *APC = Allophycocyanin; BC = buffy coat; end. = endogenous; FITC = Fluorescein isothiocyanate; FSC = forward scatter; hTRAC = human T cell receptor alpha chain constant region; KO = knock-out; NSCLC = non-small cell lung cancer; PBMC = peripheral blood mononuclear cell; SSC = sideward scatter; TCR = T cell receptor; TIL = tumor-infiltrating lymphocyte.*

Supplementary Figure 6:

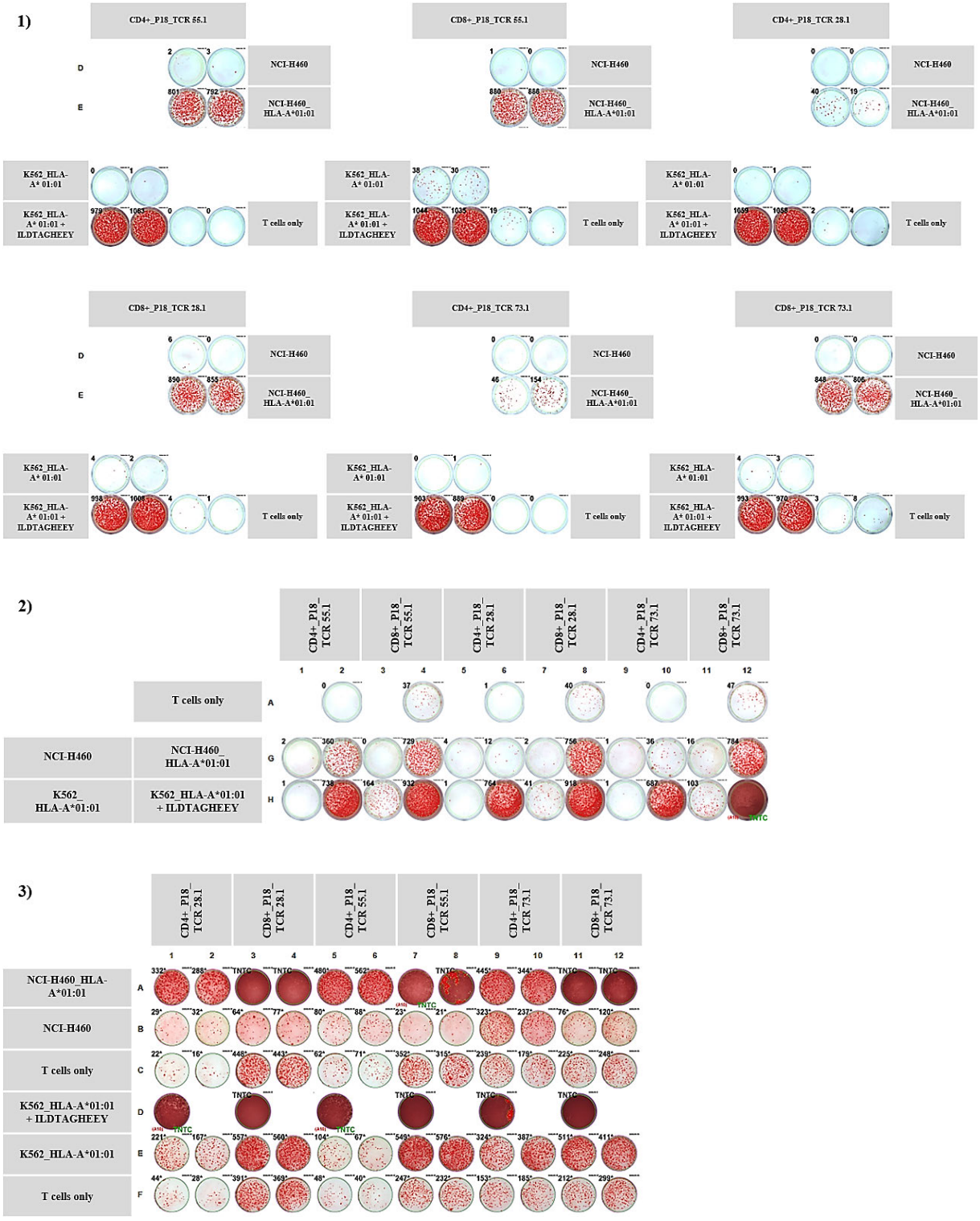


Sup. Fig. 6: Generation of P18 TCR-T cell populations from BC074- and BC437-derived donor lymphocytes.
*A CRISPR/Cas9-based KO approach was used for knocking out endogenous TCR expression in BC074- and BC437-derived T cell populations (see Sup. Fig. 5 A and B) before retroviral transduction with P18 TCR 28.1, 55.1, or 73.1 expression constructs. Resulting BC074-derived effector populations exhibited exogenous TCR expression ranging between 88.6% and 95.9% **(A)** while slightly higher (93.6% to 95.8%) overall P18 TCR expression was detected in transduced BC437 T cell cultures **(B)**. P18 TCR expression was determined by using a monoclonal antibody specifically binding the murinzed TCR constant region (mTRAC) of the introduced TCRs. Flow cytometry staining was performed 11 days (BC074) and 14 days (BC437) after transduction, respectively. Flow cytometry data was gated based on viable T cell populations and non-specific isotype (IgG1) control staining. **Abbr.**: BC = buffy coat; FITC = Fluorescein isothiocyanate; FSC = forward scatter; KO = knock-out; mTRAC = murine T cell receptor alpha chain constant region; NSCLC = non-small cell lung cancer; PE = Phyco-erythrin; SSC = sideward scatter; TCR = T cell receptor.*

B


CD4+_BC437_P18_TCR 28.1
CD8+_BC437_P18_TCR 28.1
CD4+_BC437_P18_TCR 55.1
CD8+_BC437_P18_TCR 55.1
CD4+_BC437_P18_TCR 73.1
CD8+_BC437_P18_TCR 73.1
SSC
FSC
IgG1_FITC
IgG1_PE
mTRAC_FITC
CD4_PE
CD8_PE

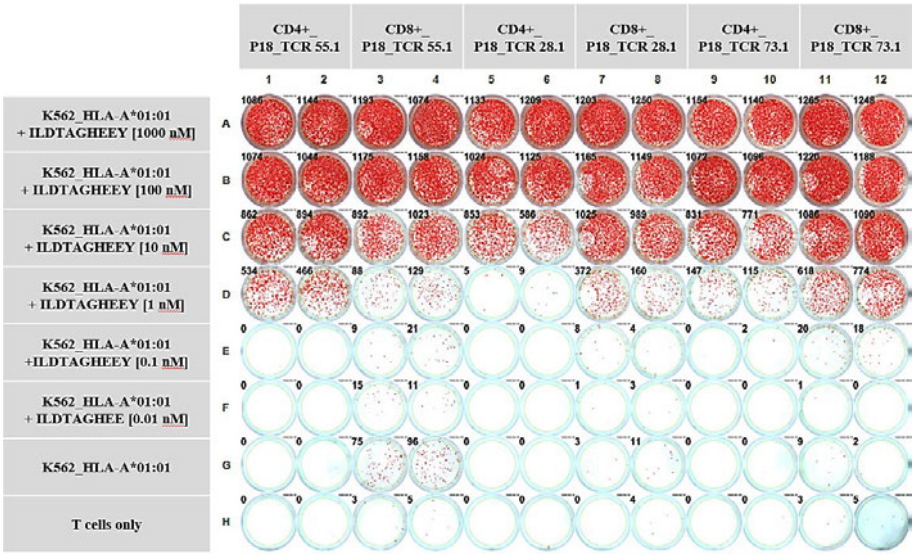
Supplementary Figure 7:



Sup. Fig. 7: Raw ELISpot data for antigen recognition testing shown in Figures 3.2.6 and 3.2.10 A/B.

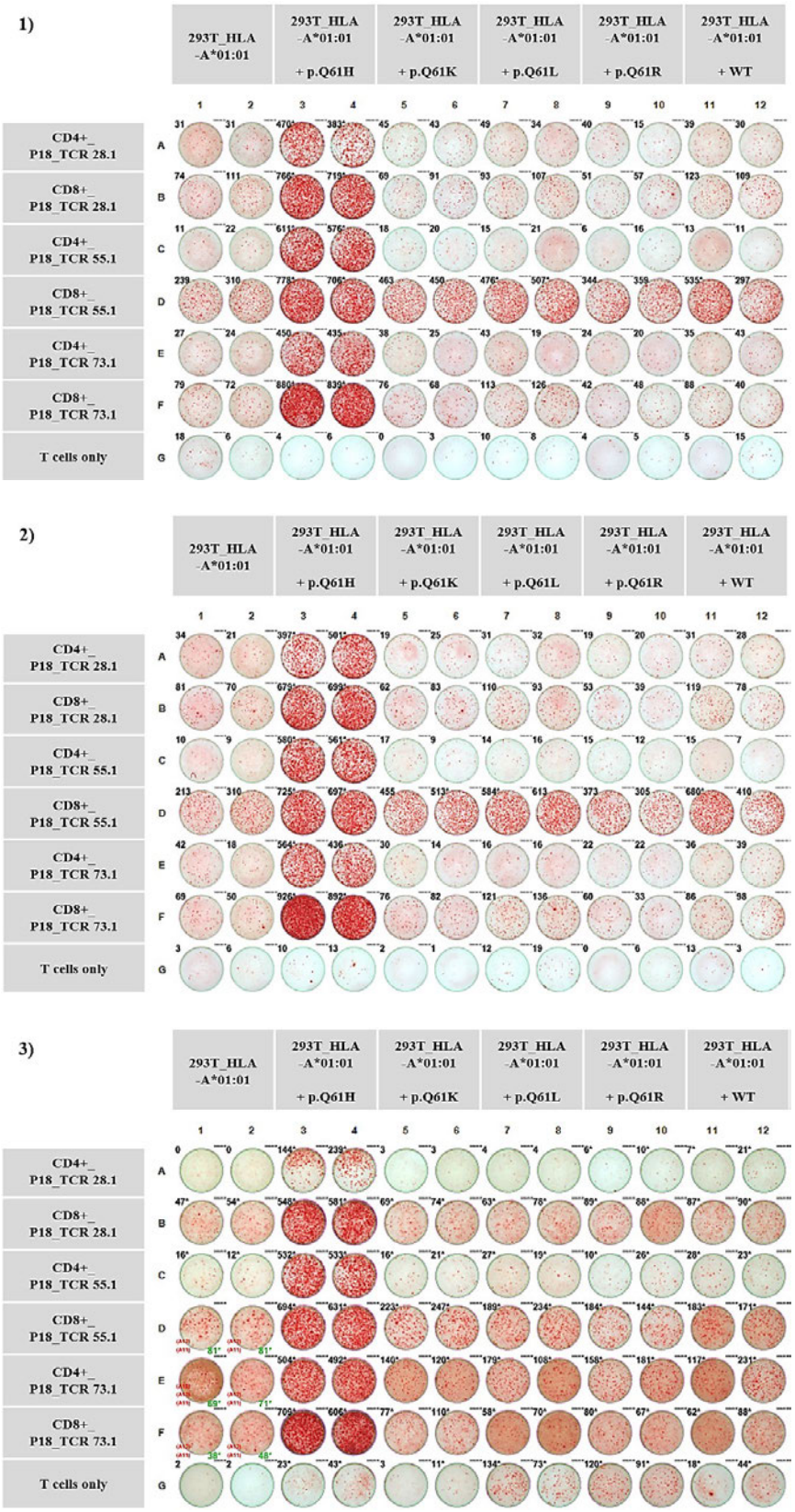
IFN-γ secretion was used as a marker for T cell recognition. ELISpot assays 1-3) were performed by using TCR-T cell populations either generated from BC074- or BC437-derived T lymphocytes that had before been applied to a CRISPR/Cas9-based end.TCR-KO approach (Sup. Fig. 5 A and B). For BC074, ELISpots were performed 20 days (1) and 35 days (2) after retroviral transduction. BC437-derived T cells were used 39 days (3) after TCR transduction. Non-used wells were excluded for final evaluation. Target recognition analyses were performed with 10,000 mTRAC+ T cells, 50,000 target cells and final neopeptide concentrations of 2 µg/ml per well, respectively. Effector numbers per well were calculated based on previously performed mTRAC staining. IFN-γ background levels were assessed by analyzing effector populations in the absence of any target cells (T cells only preparations) and were subsequently used for data normalization. Retrospective corrections of inaccurately counted wells are highlighted in green font. ***Abbr.:*** *BC = buffy coat; end. = endogenous; H = histidine; HLA = human leukocyte antigen; KO = knock-out; KRAS = Kirsten rat sarcoma viral oncogene; mTRAC = murine T cell receptor alpha constant region; p. = amino acid position; Q = glutamine; TCR = T cell receptor; TNTC = too numerous to count.*

Supplementary Figure 8:



Sup. Fig. 8: Raw ELISpot data for KRAS^{Q61H} peptide titrations shown in Figure 3.2.7. *IFN-γ secretion was used as a marker for T cell recognition. ELISpot assays were performed 26 days after retroviral transduction of BC074-derived T cell cultures with P18 TCR expression constructs. BC074-derived T cell populations had before been applied to a CRISPR/Cas9-based KO of their endogenous TCR expression (Sup. Fig. 5 A). Recognition testing was performed as duplicates with 10,000 mTRAC-positive T cells, 50,000 target cells and differing peptide concentrations for each well. Effector numbers per well were calculated based on previously performed mTRAC staining. IFN-γ background levels were assessed by analyzing transgenic T cell populations in the absence of target cells (T cells only preparations) and were subsequently used for data normalization.* ***Abbr.:*** *BC = buffy coat; H = histidine; HLA = human leukocyte antigen; KO = knock-out; KRAS = Kirsten rat sarcoma viral oncogene; mTRAC = murine T cell receptor alpha constant region; p. = amino acid position; Q = glutamine; TCR = T cell receptor.*

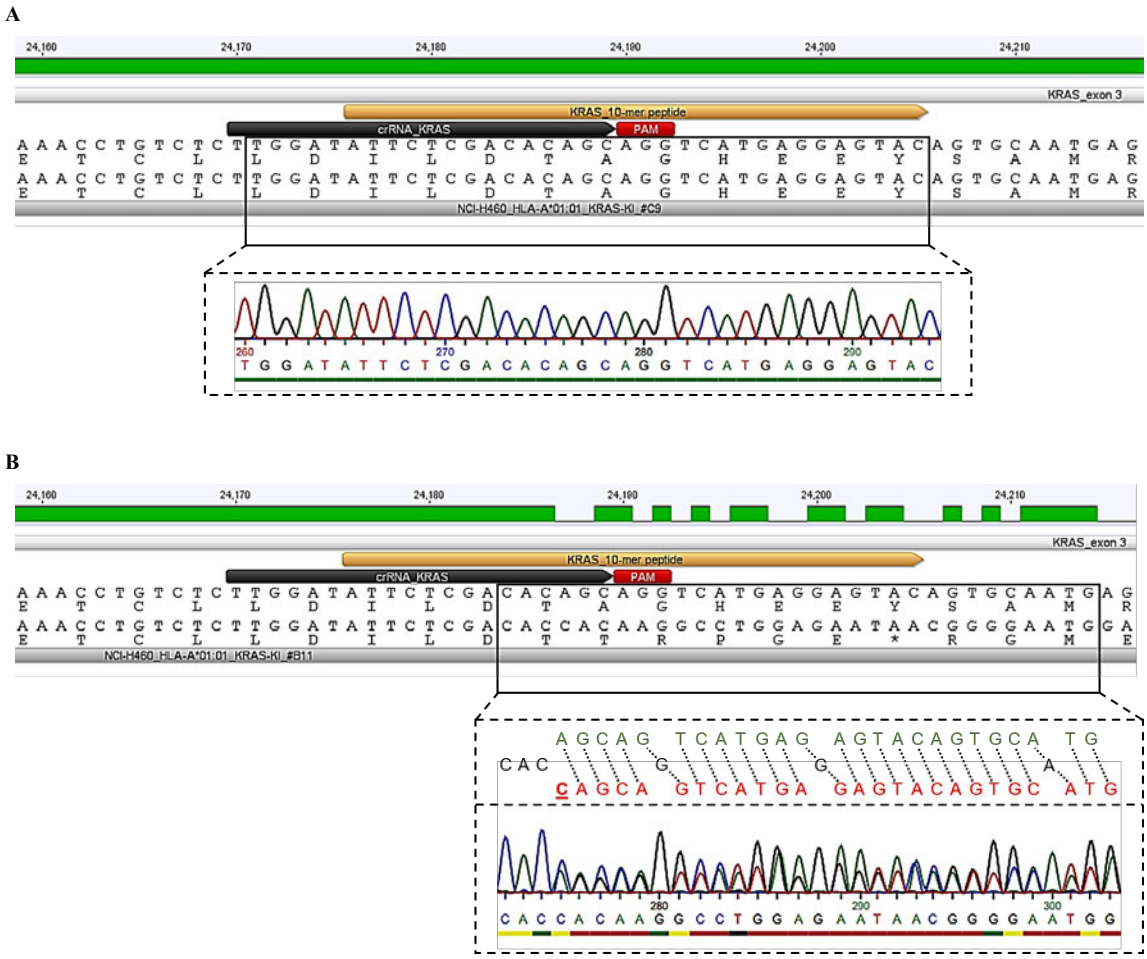
Supplementary Figure 9:



Sup. Fig. 9: Raw ELISpot data for KRAS mutation testing shown in Figure 3.2.8.

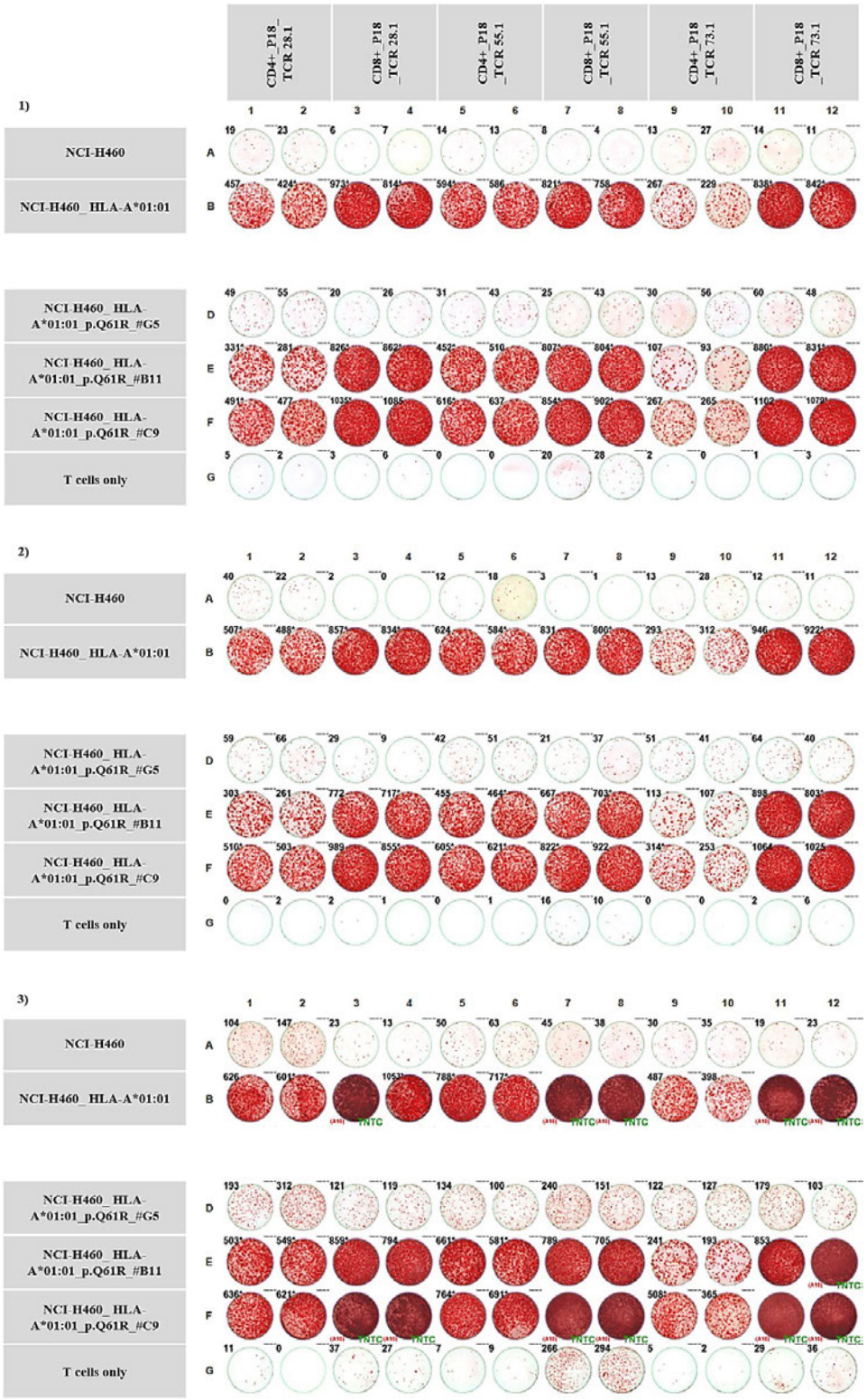
IFN-γ secretion was used as a marker for T cell recognition. Target recognition analyses were performed as duplicates with 7,500 mTRAC+ T cells and 20,000 transfected target cells per well. 300 ng of KRAS cDNA were used per well. IFN-γ background levels were assessed by analyzing transgenic T cell populations in the absence of target cells (T cells only preparations) and were subsequently used for data normalization. ELISpot assays 1) and 2) were performed 41 days (CD4+) and 40 days (CD8+) after retroviral transduction of BC437-derived T cells with TCR expression constructs. ELISpot assay 3) was performed 55 days (CD4+) and 54 days (CD8+) after retroviral transduction of BC437-derived T cells with TCR expression constructs. BC437-derived T cell populations had before been used for a CRISPR/Cas9-based end.TCR-KO approach (Sup. Fig. 5 B). Retrospective corrections of inaccurately counted wells are shown in green font. ***Abbr.:*** *BC = buffy coat; cDNA = complementary DNA; end. = endogenous. H = histidine; HEK = human embryonic kidney; HLA = human leukocyte antigen; K = lysine; KO = knock-out; KRAS = Kirsten rat sarcoma viral oncogene; L = leucine; mTRAC = murine T cell receptor alpha constant region; p. = amino acid position; Q = glutamine; R = arginine; TCR = T cell receptor; WT = wildtype.*

Supplementary Figure 10:



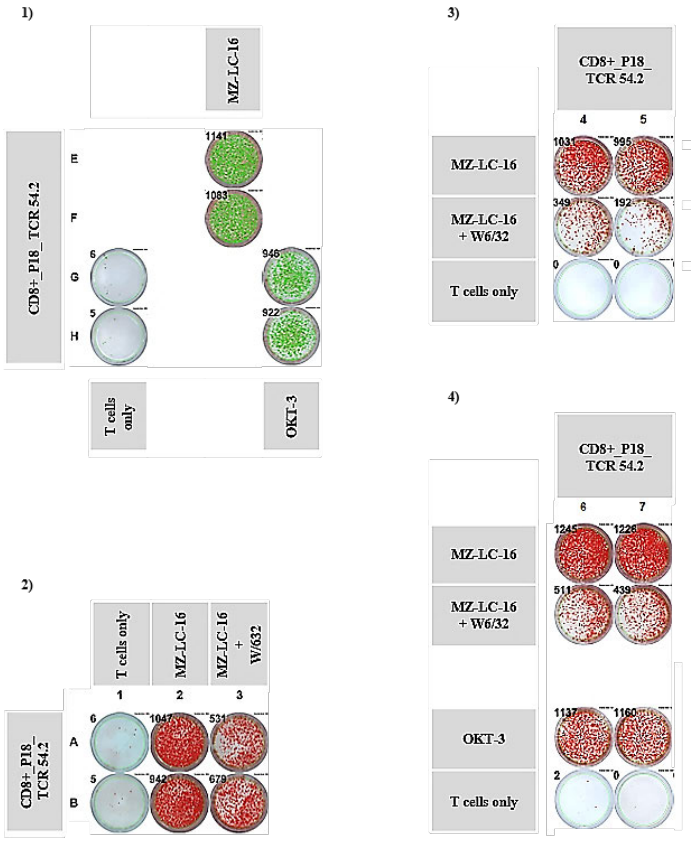
Sup. Fig. 10: Sanger sequencing of NCI-H460_HLA-A*01:01 clones #B11 and #C9 after CRISPR/Cas9-based HDR treatment.
*Sanger sequencing data of HDR-treated NCI-H460_HLA-A*01:01 clones #C9 and #B11. A CRISPR/Cas9-based HDR knock-in approach was performed to substitute biallelic KRAS^{Q61H} mutations with p.Q61R variants. Sequencing data and translated amino acid sequences of both clones are shown below the corresponding reference sequence of KRAS^{Q61H} exon 3, respectively. Related original sequencing plots are shown in dashed boxes. crRNA binding sequences are shown in black while corresponding PAM sequences are highlighted in red. Immunogenic KRAS 10-mer neopeptide sequences are depicted in orange. KRAS reference gDNA bp positions are indicated above.* ***A)*** *Both KRAS^{Q61H} alleles in #C9 were non-modified.* ***B)*** *#B11 exhibited an artificial cytosine insertion in one of both KRAS^{Q61H} alleles (nucleotide position 173), thereby resulting in two different KRAS reading frames. Both #B11 KRAS reading frames are shown in green and red with the inserted cytosine being highlighted accordingly. Concordant amino acids are depicted in black.* ***Abbr.:*** *bp = base pair; crRNA = CRISPR-RNA; gDNA = genomic DNA; HDR = homology-directed repair; H = histidine; HLA = human leukocyte antigen; KI = knock-in; KRAS =Kirsten rat sarcoma viral oncogene; p. = amino acid position; PAM = protospacer adjacent motif; Q = glutamine; R = arginine.*

Supplementary Figure 11:



Sup. Fig. 11: Raw ELISpot data for recognition testing of KRAS-modified cell clones (#B11, #C9, #G5) shown in Figure 3.2.14.
*KRAS-modified NCI-H460_HLA-A*01:01 clones #B11, #C9, and #G5 were screened for the recognition by P18 TCR-T cell populations. IFN-γ secretion was used as a marker for T cell recognition. Target recognition analyses were performed as duplicates with 50,000 target cells and either 2,500 (ES 1&2) or 7,500 (ES 3) effectors per well. ELISpot assays were either performed 36 days (CD4+) or 35 days (CD8+) after retroviral transduction of T cells (BC437) with P18 TCR expression constructs. IFN-γ background levels were assessed in the absence of target cells (T cells only preparations) and were subsequently used for data normalization. Effectors had before been used for CRISPR/Cas9-based end.TCR-KOs (Sup. Fig. 5 B). Retrospective corrections of inaccurately counted wells are highlighted in green font. Non-used wells and wells containing irrelevant antigen screening information were excluded from plate scans for final evaluation* ***Abbr.:*** *BC = buffy coat; ES = ELISpot; end. = endogenous; H = histidine; HDR = homology-directed repair; HLA = human leukocyte antigen; KO = knock-out; KRAS =Kirsten rat sarcoma viral oncogene; p. = amino acid position; Q = glutamine; R = arginine; TCR = T cell receptor.*

Supplementary Figure 12:



Sup. Fig. 12: Raw ELISpot data for MZ-LC-16 recognition testing shown in Fig. 3.2.15 A.
IFN-γ secretion was used as a marker for T cell recognition. ELISpot assays 1-4 were performed by using P18 TCR 54.2-expressing CD8+ TCR-T cell populations generated from BC827. ELISpots were performed 16 days (ES 1), 18 days (ES 2), and 20 days (ES 3&4) after retroviral transduction, respectively. BC827-derived T cell populations had before been used for a CRISPR/Cas9-based KO of their endogenous TCR expression (Sup. Fig. 1). For final evaluation, non-used wells and wells containing irrelevant antigen testing information were excluded from plate scans. Target recognition analyses were either performed with 10,000 (ES 1,2,4) or 5,000 (ES 3) mTRAC+ T cells per well. 50,000 target cells were used per well and antibodies were utilized at concentrations of 400 ng/ml (OKT-3) and 50µg/ml (W6/32), respectively. Effector numbers per well were calculated based on previously performed mTRAC staining. IFN-γ background levels were assessed by analyzing transgenic T cell populations in the absence of target cells (T cells only preparations) and were subsequently used for data normalization. ***Abbr.:*** *BC = buffy coat; ES = ELISpot; HLA = human leukocyte antigen; KO = knock-out; mTRAC = murine T cell receptor alpha constant region; TCR = T cell receptor.*

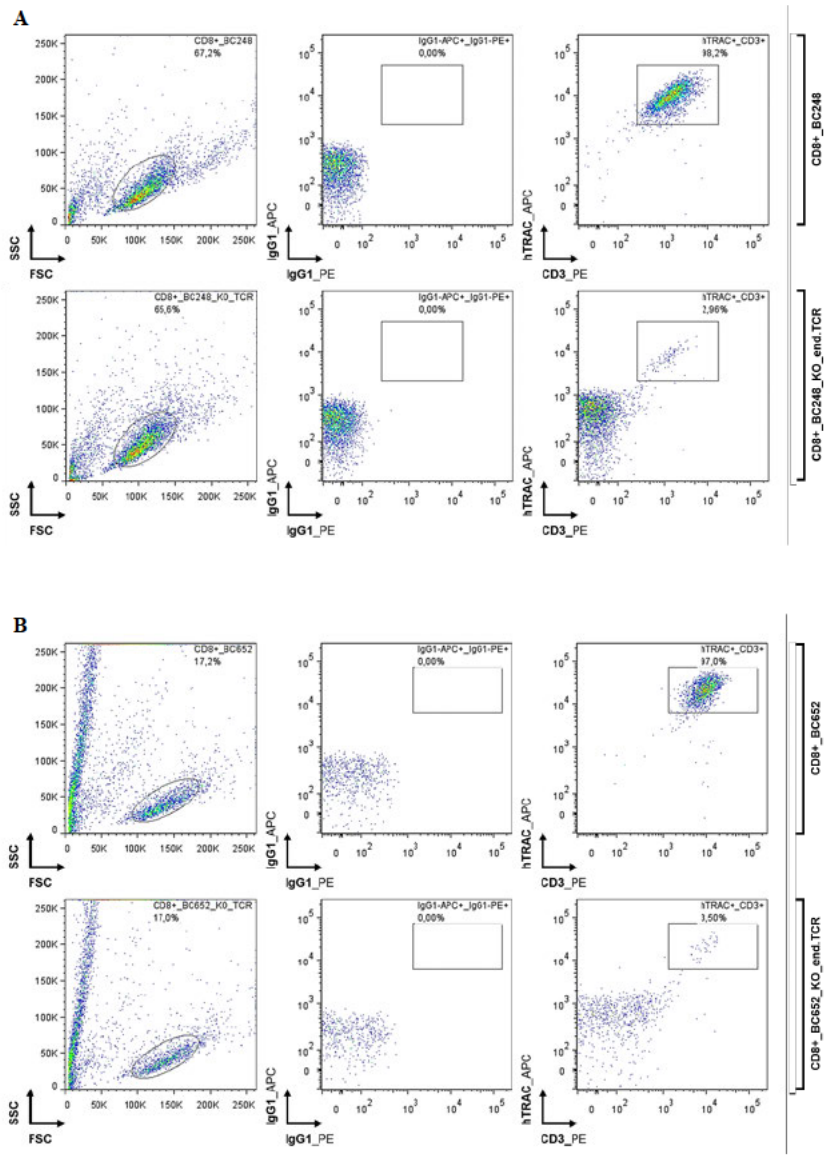
Supplementary Figure 13:

	1	2	3	4	5	6	7	8	9	10	11	12
A	MAGEA1	MAGEA2	MAGEA3	MAGEA4	MAGEA6	MAGEA8	MAGEA9	MAGEA10	MAGEA11	MAGEA12	MAGEB2	MAGEC1
B												
C	MAGEC2	GAGE1	GAGE2	GAGE3	GAGE4	GAGE5	GAGE6	GAGE7B	GAGE8	RAGE1	RAGE2	RAGE3
D												
E	RAGE-4	BAGE-1	PRAME	WT1	CTAG1B (NY-ESO-1)	BIRC5 (Surv.-1)	BIRC5 (Surv.-2)	ERBB2 (HER2/neu)	BRAF-V600E	TSPY1	Meloe [1]	CSPG4
F												
G	CSFR2A	XAGE1 A/B	CT83	pp65	PLAUR (uPAR)	Controls						
H												

Sup. Fig. 13: Ready-to-transfect TAA/CG antigen panel.

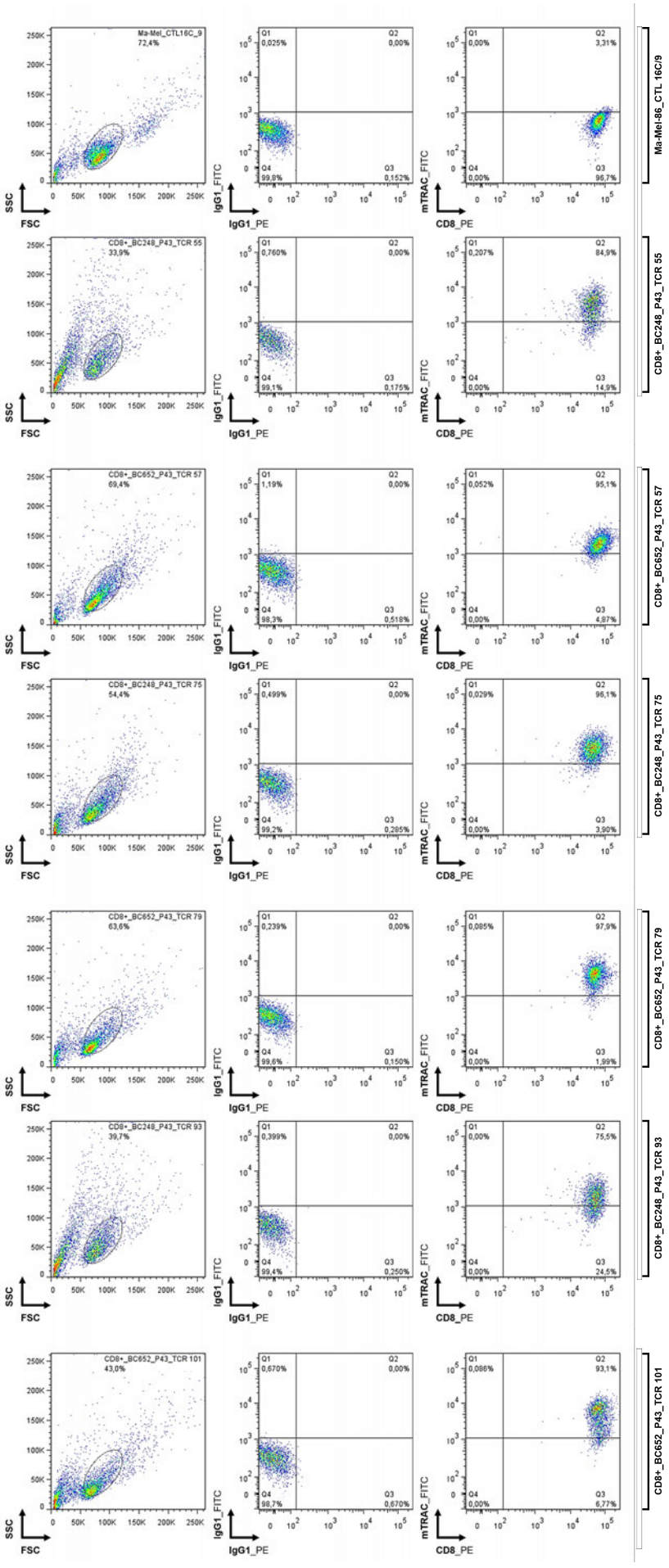
A ready-to-transfect cDNA panel consisting of common TAAs and CG antigens was established for subsequent T cell recognition testing. To establish pcDNA3.1 expression constructs, ORFs of respective antigens were cloned from patient-derived tumor cells or tumor cell lines. The panel was prepared by transferring relevant antigen constructs into 96-well plates according to the scheme shown above. Two wells were prepared for every single antigen with each of them containing 9 µg cDNA in 210 µl RPMI medium. 300 ng (7 µl) were then used for transient transfection of HEK 293T target cells. None of the shown antigens was recognized by any of the tested P18 TCRs (data now shown). Based on P43 and P50 NSCLC gene expression analyses, the panel was successively complemented by adding SSX-1 and HMGB3 TAAs for further antigen testing. ***Abbr.:*** *cDNA = complementary DNA; CG = cancer germline; NSCLC = non-small cell lung cancer; TAA = tumor associated antigen; TCR = T cell receptor.* ***Legend:*** *Green = CG antigens; blue = overexpressed TAAs; red = mutated TAAs; orange = viral antigens.* [1] *= Meloe is classified as both differentiation antigen and aberrantly expressed TAA.*

Supplementary Figure 14:



Sup. Fig. 14: CRISPR/Cas9-based KO of endogenous TCR expression in CD8+ donor lymphocytes derived from BC248 and BC652.

Endogenous TCR expression of human T cell populations was knocked out before retroviral transduction with P43 and P50 TCR expression constructs. When compared to corresponding non-treated T cell populations exhibiting more than 95% of endogenous TCR expression, the performed KO approach dramatically reduced human TCR expression to less than 4% in both BCs. Residual human TCR expression was determined by using a monoclonal antibody specifically binding the human TCR constant region (hTRAC). Flow cytometry staining was performed 4 days after CRISPR/Cas9 treatment. Flow cytometry data was gated based on viable T cell populations and hTRAC staining of non-treated CD8+ T cell populations. Isotype (IgG1) controls were performed to assess non-specific IgG1 binding of utilized antibodies. ***Abbr.:*** *APC = Allophycocyanin; BC = buffy coat; end. = endogenous; FITC = Fluorescein isothiocyanate; FSC = forward scatter; hTRAC = human T cell receptor alpha chain constant region; KO = knock-out; NSCLC = non-small cell lung cancer; SSC = sideward scatter; TCR = T cell receptor; TIL = tumor-infiltrating lymphocyte.*

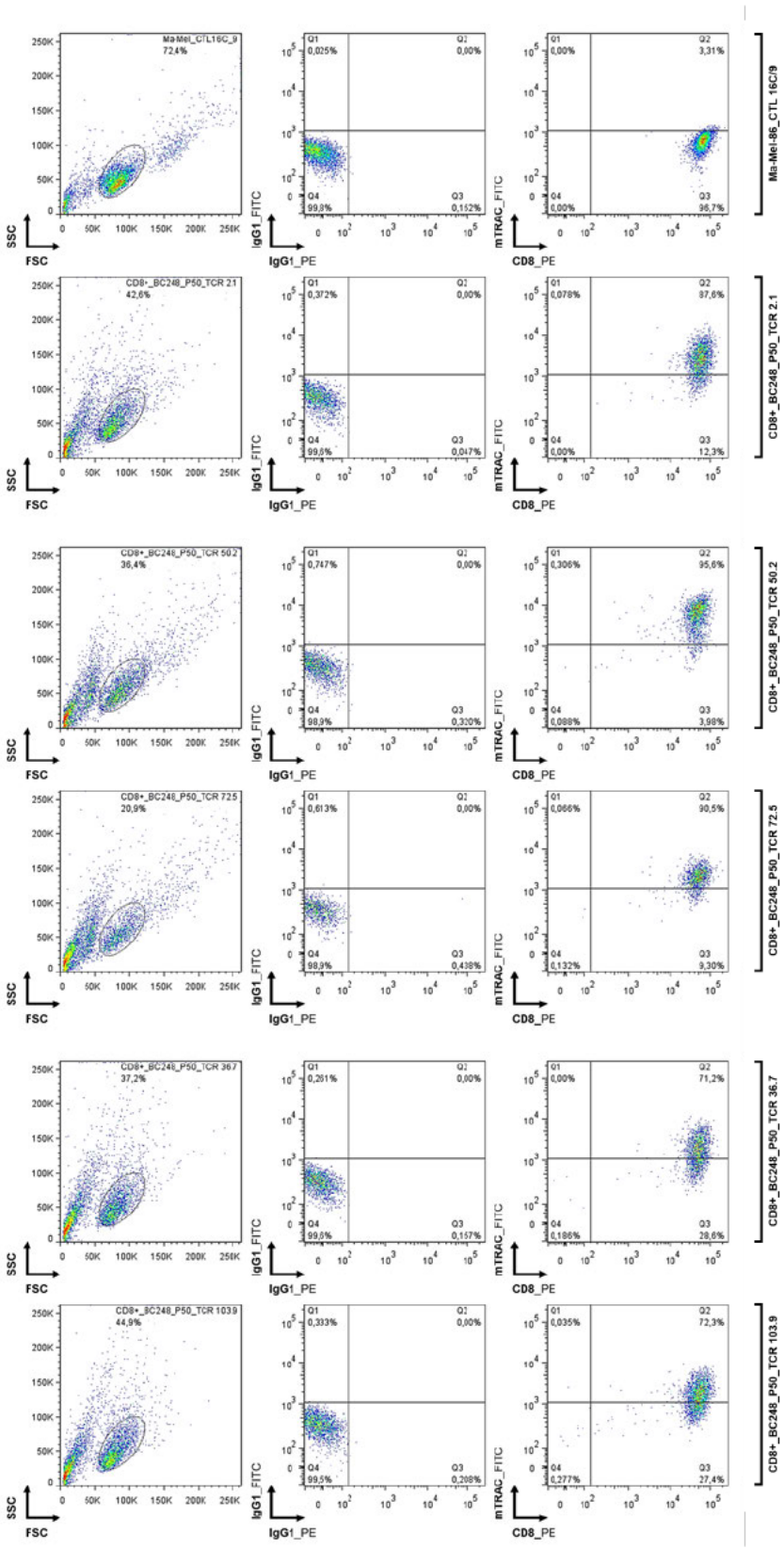


Supplementary Figure 15:

Sup. Fig. 15: Generation of P43 TCR-T cell populations derived from BC248 and BC652.

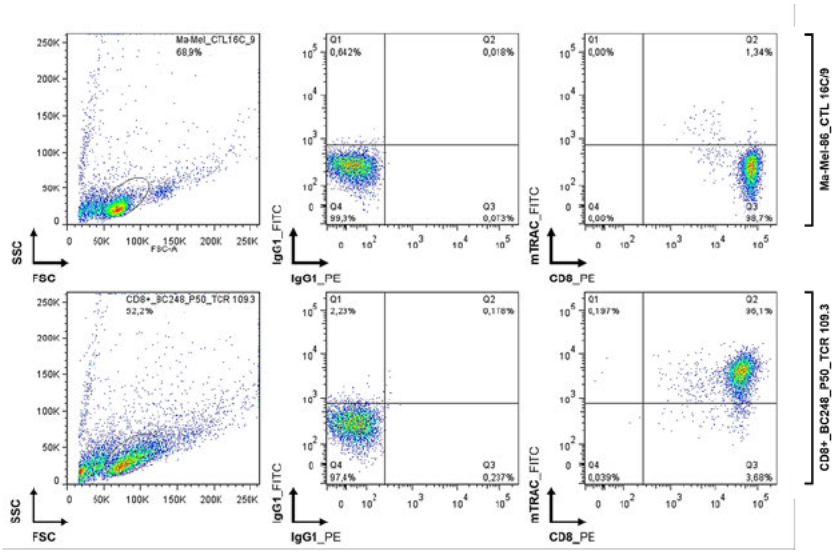
P43 TCR expression of retrovirally transduced T cell populations either obtained from BC248 or BC652. A CRISPR/Cas9-based approach was used to knock out endogenous TCR expression (Sup. Figure 14) before retroviral transduction of T cell populations with P43 TCR expression constructs. Depending on generated effector populations, between 75.5% and 97.9% of T cells expressed the introduced P43 TCRs, respectively. Exogenous TCR expression was assessed by using a monoclonal antibody that specifically binds the murinzed TCR constant region (mTRAC) of the introduced TCRs. Flow cytometry staining was performed 12-26 days (BC248) and 53-69 days (BC652) after retroviral transduction. Flow cytometry data was gated based on viable T cell populations and non-specific isotype (IgG1) controls. Non-transduced CTL populations derived from melanoma patient Ma-Mel-86 were used as a negative control for mTRAC staining. ***Abbr.:*** *BC = buffy coat; CTL = cytotoxic T lymphocyte; FITC = Fluorescein isothiocyanate; FSC= forward scatter; KO = knock-out; mTRAC = murine T cell receptor alpha chain constant region; NSCLC = non-small cell lung cancer; PE = Phycoerythrin; SSC = sideward scatter; TCR = T cell receptor.*

Supplementary Figure 16:



Sup. Fig. 16: Generation of P50 TCR-T cell populations derived from BC248. *P50 TCR expression of retrovirally transduced T cell populations obtained from BC248. A CRISPR/Cas9-based approach was used to knock out endogenous TCR expression (Sup. Fig. 14) before retroviral transduction of T cell populations with created P50 TCR expression constructs. Depending on generated effector populations, between 71.2% and 95.6% of T cells expressed the introduced P50 TCRs, respectively. Exogenous TCR expression was assessed by using a monoclonal antibody that specifically binds the murinzed TCR constant region (mTRAC) of the introduced TCRs. Flow cytometry staining was performed 12-26 days after retroviral transduction. Flow cytometry data was gated based on viable T cell populations and non-specific isotype (IgG1) controls. Non-transduced CTLs derived from melanoma patient Ma-Mel-86 were used as a negative control for mTRAC staining.* ***Abbr.:*** *BC = buffy coat; CTL = cytotoxic T lymphocyte; FITC = Fluorescein isothiocyanate; FSC = forward scatter; KO = knock-out; mTRAC = murine T cell receptor alpha chain constant region; NSCLC = non-small cell lung cancer; PE = Phyco-erythrin; SSC = sideward scatter; TCR = T cell receptor.*

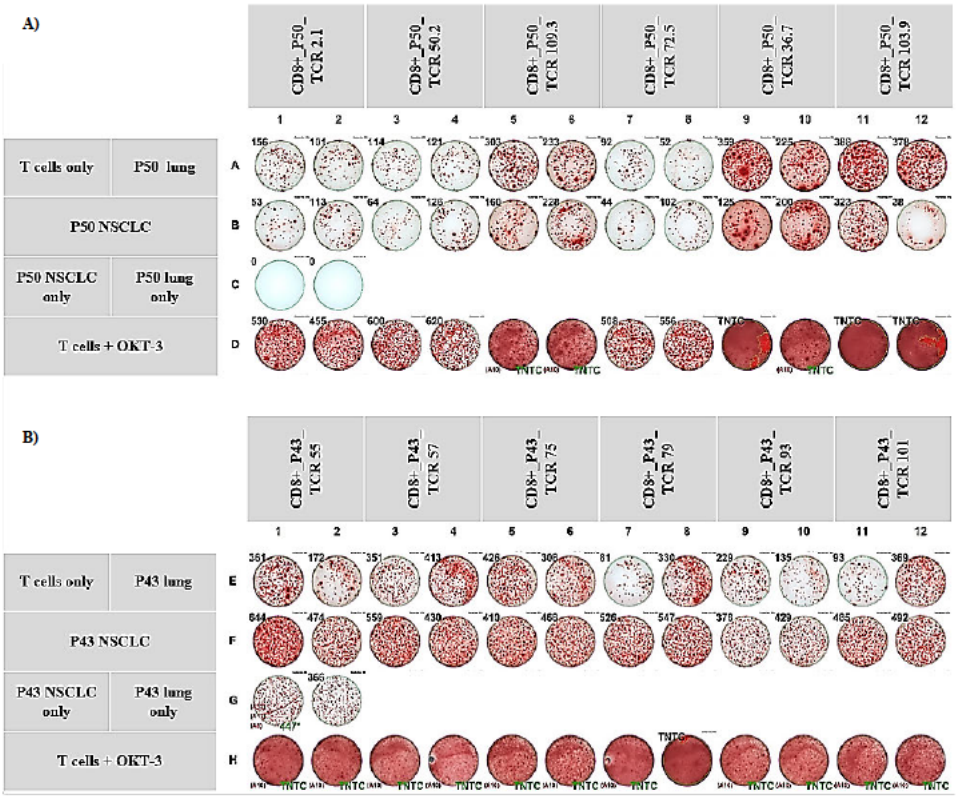
Supplementary Figure 17:



Sup. Fig. 17: Generation of P50 TCR 109.3-expressing T cells derived from BC248.

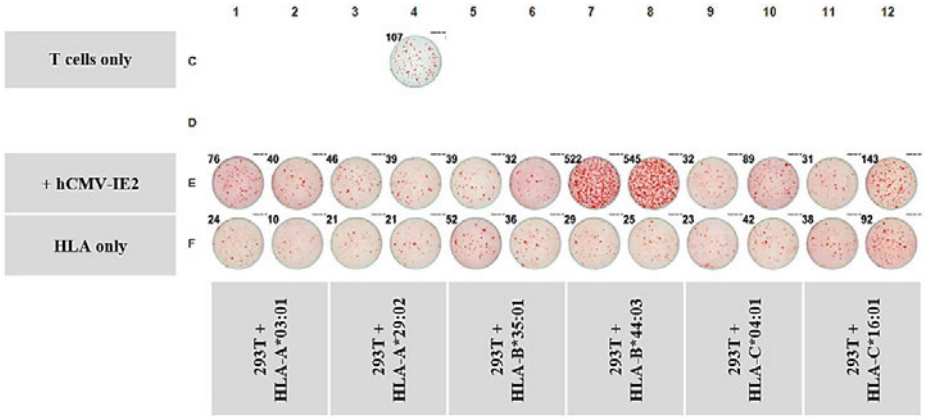
P50 TCR 109.3 expression of retrovirally transduced T cells obtained from BC248. A CRISPR/Cas9-based approach was used to knock out endogenous TCR expression (Sup. Figure 14) before retroviral transduction of T cell populations with the P50 TCR-109.3 expression construct. When compared to non-modified Ma-Mel-86 CTL controls, 96.1% of the generated effector population expressed the introduced TCR. Exogenous TCR expression was assessed by using a monoclonal antibody that specifically binds the murinzed TCR constant region (mTRAC) of the introduced TCR. Flow cytometry staining was performed 22 days after retroviral transduction. Flow cytometry data was gated based on viable T cell populations and non-specific isotype (IgG1) controls. ***Abbr.:*** *BC = buffy coat; CTL = cytotoxic T lymphocyte; FITC = Fluorescein isothiocyanate; FSC = forward scatter; KO = knock-out; mTRAC = murine T cell receptor alpha chain constant region; NSCLC = non-small cell lung cancer; PE = Phycoerythrin; SSC = sideward scatter; TCR = T cell receptor.*

Supplementary Figure 18:



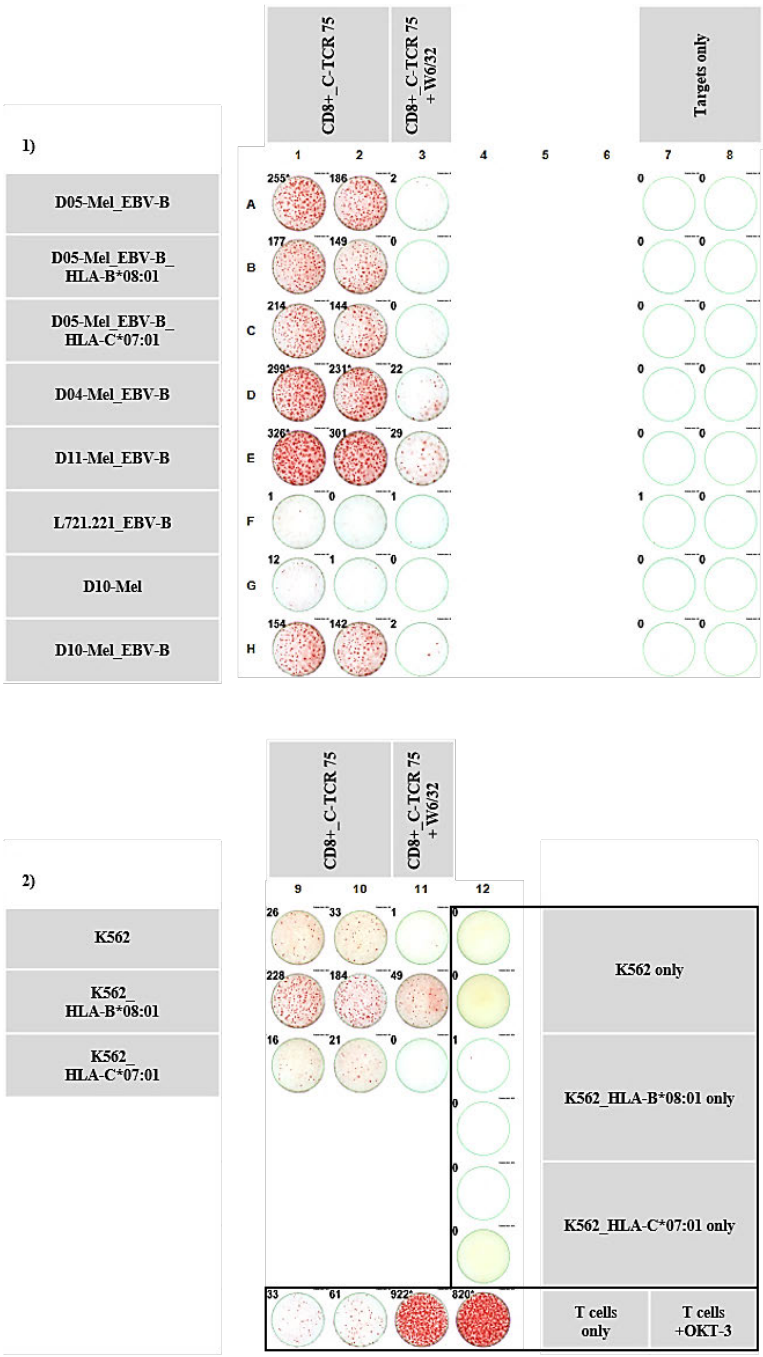
Sup. Fig. 18: Raw ELI-Spot data for P43 and P50 tumor recognition testing shown in Figure 3.3.3. *A/B) IFN-γ secretion was used as a marker for T cell recognition. Spontaneous IFN-γ secretion of target cells was assessed by analyzing patient-derived tumor and normal lung cell suspensions in the absence of effector cells. IFN-γ background levels were assessed by analyzing TCR-T cell populations in the absence of target cells. Effector functionality was assessed via non-specific activation with OKT-3. Non-used wells and wells containing non-relevant antigen testing information were excluded for final evaluation. BC652- and BC248-derived T cells had been used for a CRISPR/Cas9-based end.TCR-KO (Sup. Fig. 14) before being used for the retroviral transduction with P43 and P50 TCR expression constructs (Sup. Fig. 15-17). Effector numbers per well were calculated based on mTRAC staining performed 1-2 days before ELISpot analysis. Utilized patient target cell suspensions were mostly non-viable. Due to limiting numbers of viable target cells, some recognition analyses were performed in a single well format. Retrospective corrections of inaccurately counted wells are highlighted in green font. OKT-3 concentration: 400 ng/ml.* ***A)*** *The P50 recognition testing was either performed 12 days (TCRs 36.7, 103.9), 19 days (TCRs 50.2, 72.5, 109.3), or 26 days (TCR 2.1) after retroviral transduction. All effector populations had initially been generated from BC248-derived T cells. P50 lung and NSCLC cells per well: 14,000. TCR 2.1- / 36.7- / 103.9-expressing effectors per well: 30,000. Other effectors per well: 10,000 mTRAC+.* ***B)*** *The P43 recognition testing was either performed 12 days (TCRs 55, 93), 26 days (TCR 75), 53 days (TCRs 57, 101), or 69 days (TCR 79) after retroviral transduction. Effector populations expressing TCRs 55, 75, and 93 had initially been generated from BC248-derived T cells while others were generated from BC652. P43 lung cells per well: 37,000. P43 NSCLC cells per well: 57,000. TCR 55- and TCR 79-expressing effectors per well: 30,000. Other effectors per well: 10,000 mTRAC+.* ***Abbr.:*** *autol. = autologous; BC = buffy-coat; end. = endogenous; KO = knock-out; mTRAC = murine T cell receptor alpha chain constant region; NSCLC = non-small cell lung cancer; TCR = T cell receptor; TNTC = too numerous to count.*

Supplementary Figure 19:



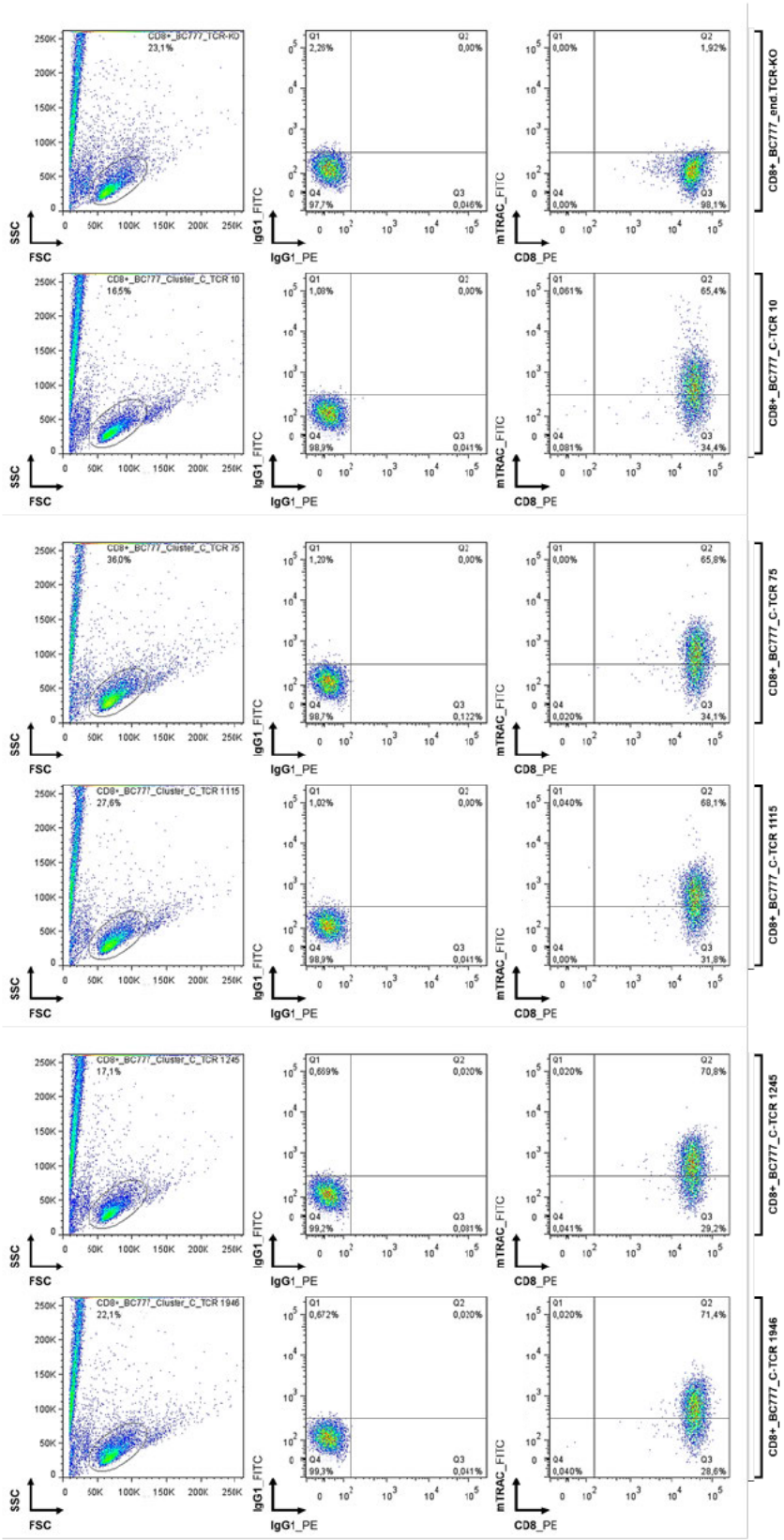
Sup. Fig. 19: Raw ELISpot data for P50 TCR 36.7 recognition testing shown in Figure 3.4.1 B.
IFN-γ secretion was used as a marker for T cell recognition. The corresponding IFN-γ background level was assessed by analyzing P50 TCR 36.7+ T cell populations in the absence of target cells (T cells only preparations). Non-used wells and wells containing non-relevant testing information were excluded for subsequent evaluation. BC248-derived T cell populations had been used for a CRISPR/Cas9-based end.TCR-KO approach (Sup. Fig. 14 A) before being used for retroviral transduction with the TCR 36.7 expression construct (Sup. Fig. 16). Effector numbers per well were calculated based on mTRAC staining performed 1-2 days before ELISpot analysis. The ELISpot assay was performed 32 days after retroviral transduction. HEK 293T target cells were co-transfected with hCMV-IE2 (300 ng/well) and P50 HLA (100 ng/well) pCDNA3.1 expression constructs. P50 TCR 36.7+ T cells per well: 3,000. HEK 293T targets per well: 20,000. ***Abbr.:*** *AA = amino acids; BC = buffy coat; CMV = cytomegalovirus; end. = endogenous; HEK = Human embryonic kidney; HLA = human leukocyte antigen; hTRAV = human TCR alpha chain variable region; hTRBV = human TCR beta chain variable region; IE2 = immediate early 2; KO = knock-out; TCR = T cell receptor.*

Supplementary Figure 20:

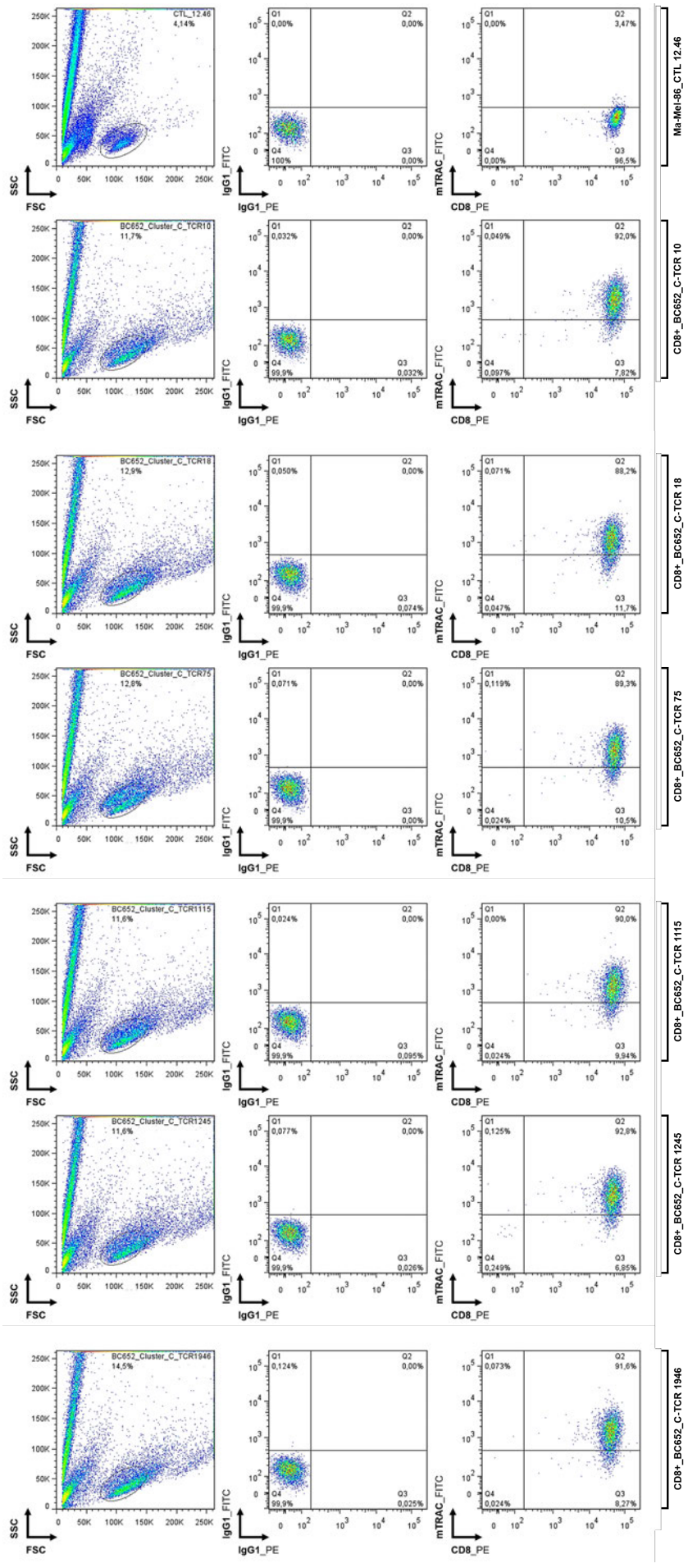


Sup. Fig. 20: Raw ELISpot data for recognition testing with C-TCR 75 shown in Fig. 3.4.2 B.
IFN-γ secretion was used as a marker for T cell recognition. IFN-γ background levels were assessed by analyzing TCR-T cells in the absence of target cells (T cells only preparations). Spontaneous IFN-γ secretion of target cells was assessed in the absence of effectors (targets only preparations). General effector functionality was assessed via non-specific activation with OKT-3. W6/32 was used to block HLA class I-dependent antigen recognition. Recognition testing of melanoma patient-derived tumor cells and EBV-infected B cell lines is shown in plate scan 1) while K562 recognition testing is shown in scan 2). Non-used wells and wells containing non-relevant antigen testing information were excluded for final evaluation. BC777-derived T cells had been used for a CRISPR/Cas9-based end.TCR-KO approach (Sup. Fig. 23) before being used for retroviral transduction with the C-TCR 75 expression construct. Effector numbers per well were calculated based on mTRAC staining performed 1-2 days before ELISpot analysis. The ELISpot assay was performed 78 days after retroviral transduction. HLA alleles of respective target cell lines are listed in Table 5.1. Effectors per well: 15,000 mTRAC+. Targets per well: 50,000. W6/32 concentration per well: 50 µg/ml. OKT-3 concentration per well: 400 ng/ml.
Abbr.: *BC = buy coat; EBV = Epstein-Barr virus; end = endogenous; HLA = human leukocyte antigen; KO = knock-out; Mel = melanoma; mTRAC = murinized TCR alpha chain constant region; NCI = National Cancer Institute; TCR = T cell receptor.*

Supplementary Figure 21:



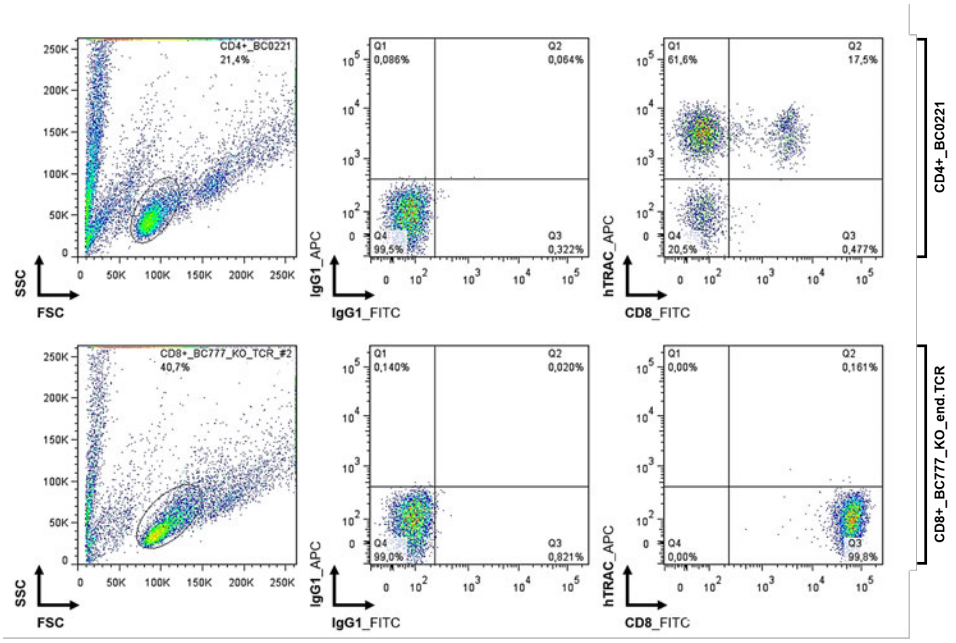
Sup. Fig. 21: Generation of C-TCR+ T cell populations from BC777.
*CD8+ lymphocytes obtained from BC777 were used for retroviral transductions with C-TCR expression constructs after being applied to an end.TCR-KO approach (Sup. Fig. 23). C-TCR expression was determined via mTRAC staining. Non-transduced BC777 CD8+ T cells were used as negative control for mTRAC staining. Depending on the generated effector populations, between 65.4% and 71.4% of T cells expressed the introduced C-TCRs, respectively. Flow cytometry staining was performed 14 days after retroviral transduction. Flow cytometry data was gated based on viable T cell populations and non-specific isotype (IgG1) controls. **Abbr.**: BC = buffy coat; CTL = cytotoxic T lymphocyte; end. = endogenous; FITC = Fluorescein isothiocyanate; FSC = forward scatter; KO = knock-out; Mel = melanoma; mTRAC = murine TCR alpha chain constant region; PE = Phycoerythrin; SSC = sideward scatter; TCR = T cell receptor.*



Supplementary Figure 22:

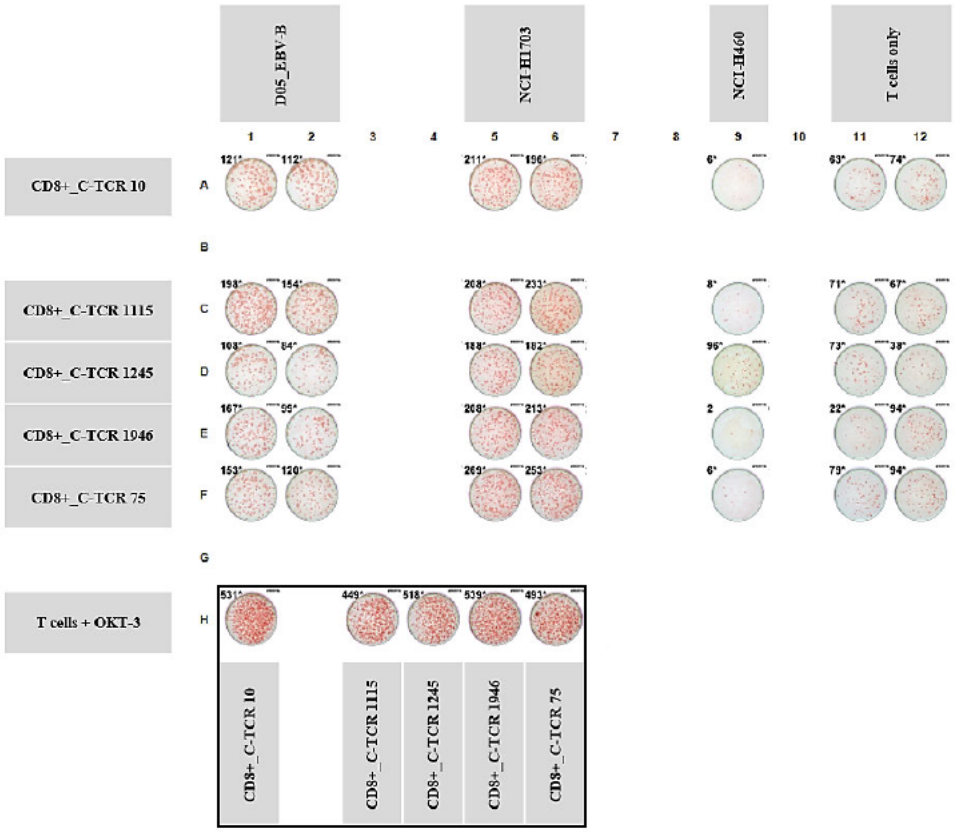
Sup. Fig. 22: Generation of C-TCR+ T cell populations from BC652.
CD8+ lymphocytes obtained from BC652 were used for retroviral transductions with C-TCR expression constructs after being applied to an end.TCR-KO approach (Sup. Fig. 14 B). C-TCR expression was determined via mTRAC staining. CTLs obtained from melanoma patient Ma-Mel-86 were used as negative control for mTRAC staining. Depending on the generated effector populations, between 88.2% and 92.8% of T cells expressed the introduced C-TCRs, respectively. Flow cytometry staining was performed 13 days after retroviral transduction. Flow cytometry data was gated based on viable T cell populations and non-specific isotype (IgG1) controls. ***Abbr.:*** *BC = buffy coat; CTL = cytotoxic T lymphocyte; end. = endogenous; FITC = Fluorescein isothiocyanate; FSC = forward scatter; KO = knock-out; Mel = melanoma; mTRAC = murine T cell receptor alpha chain constant region; PE = Phycoerythrin; SSC = sideward scatter; TCR = T cell receptor.*

Supplementary Figure 23:



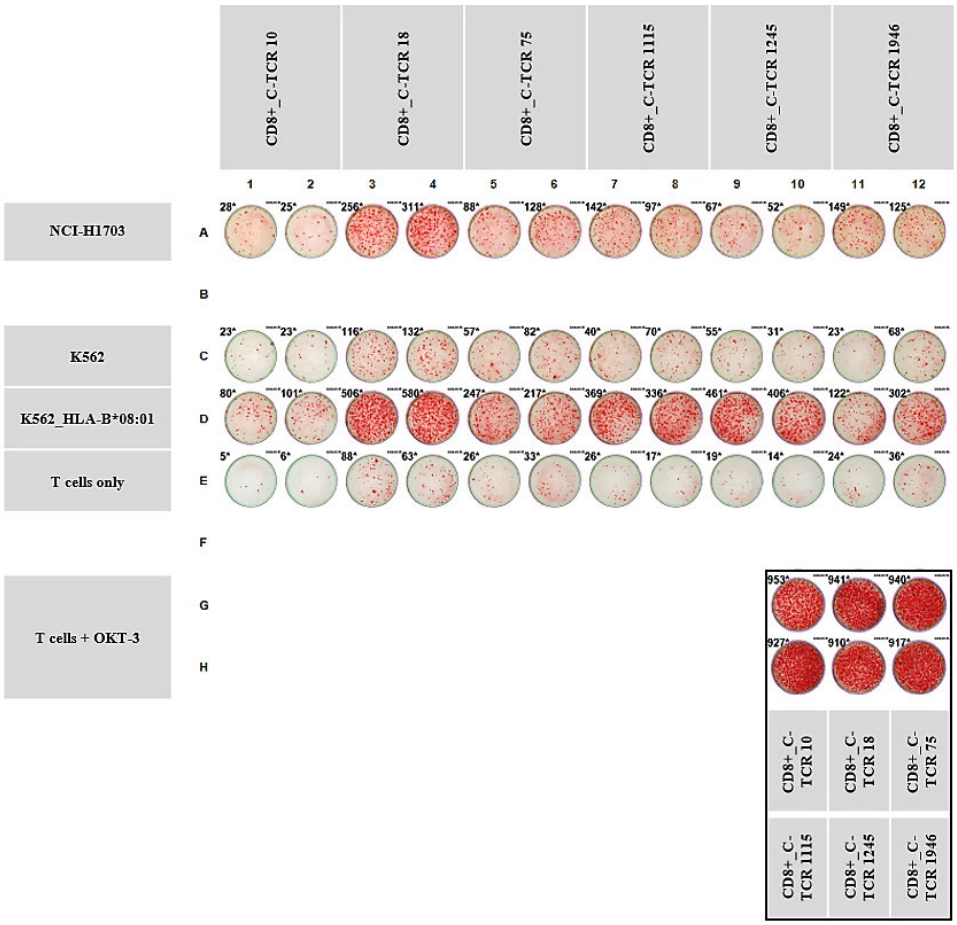
Sup. Fig. 23: CRISPR/Cas9-based KO of endogenous TCR expression in CD8+ lymphocytes derived from BC777.
BC777-derived CD8+ donor lymphocytes were used for a CRISPR/Cas9-mediated end.TCR-KO approach before retroviral transduction with cluster C TCRs (Sup. Fig. 21). Endogenous TCR expression was determined via hTRAC staining. Non-treated BC0221 CD4+ T cells were used as positive control for hTRAC staining. When compared to control populations containing more than 79% hTRAC+ T cells, endogenous TCR expression was reduced to less than 1% in BC777-derived T cell populations. Flow cytometry was performed 49 days after CRISPR/Cas9 treatment. Two consecutive CRISPR/Cas9 treatments were performed to efficiently reduce end.TCR expression in BC777 donor lymphocytes. Flow cytometry data was gated based on viable T cell populations and non-specific isotype (IgG1) controls. ***Abbr.:*** *APC = Allophycocyanin; BC = buffy coat; CTL = cytotoxic T lymphocyte; end. = endogenous; FITC = Fluorescein isothiocyanate; FSC = forward scatter; hTRAC = human T cell receptor alpha chain constant region; KO = knock-out; PE = Phycoerythrin; SSC = sideward scatter; TCR = T cell receptor.*

Supplementary Figure 24:



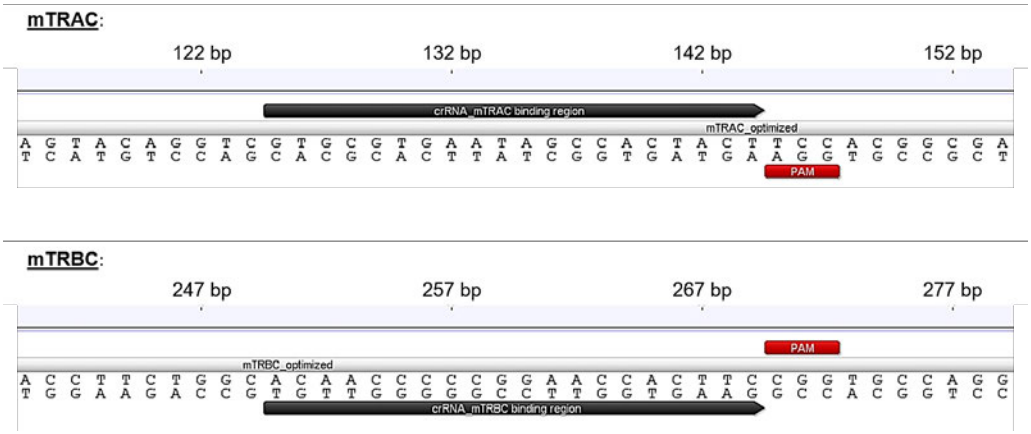
Sup. Fig. 24: Raw ELISpot data for C-TCR recognition testing shown in Figure 3.4.3 A.
IFN-γ secretion was used as a marker for T cell recognition. IFN-γ background levels were assessed by analyzing TCR-T cell populations in the absence of target cells (T cells only preparations). General effector functionality was assessed via non-specific activation with OKT-3. Non-used wells and wells containing non-relevant antigen testing information were excluded for final evaluation. BC777-derived T cell populations had been used for a CRISPR/Cas9-based end.TCR-KO (Sup. Fig. 23) approach before being used for retroviral transduction with C-TCR expression constructs (Sup. Fig. 21). Effector numbers per well were calculated based on mTRAC staining performed 1-2 days before ELISpot analysis. The ELISpot assay was performed 25 days after retroviral transduction. HLA alleles of respective target cell lines are listed in Table 5.1. Effectors per well: 25,000 mTRAC+. Targets per well: 50,000. OKT-3 concentration per well: 400 ng/ml. ***Abbr.:*** *BC = buffy coat; D05 = melanoma patient D05-Mel; EBV = Epstein-Barr virus; end. = endogenous; HLA = human leukocyte antigen; KO = knock-out; Mel = melanoma; mTRAC = murinized TCR alpha chain constant region; NCI = National Cancer Institute; TCR = T cell receptor.*

Supplementary Figure 25



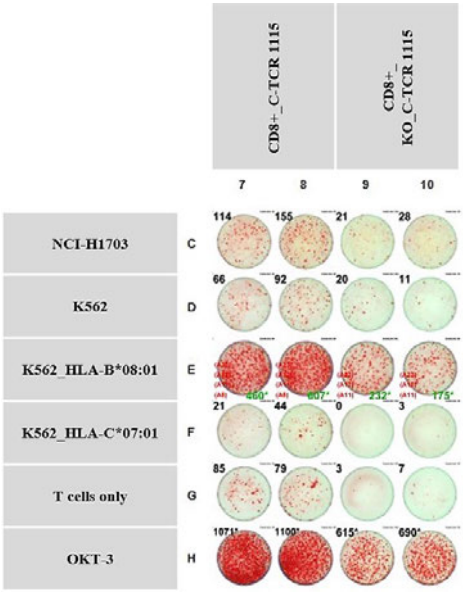
Sup. Fig. 25: Raw ELISpot data for C-TCR recognition testing shown in Figure 3.4.3 B.
IFN-γ secretion was used as a marker for T cell recognition. IFN-γ background levels were assessed by analyzing TCR-T cell populations in the absence of target cells (T cells only preparations). General effector functionality was assessed via non-specific activation with OKT-3. Non-used wells and wells containing non-relevant antigen testing information were excluded for final evaluation. BC652-derived T cell populations had been used for a CRISPR/Cas9-based end.TCR-KO (Sup. Fig. 14 B) approach before retroviral transduction with C-TCR expression constructs (Sup. Fig. 22). Effector numbers per well were calculated based on mTRAC staining performed 1-2 days before ELISpot analysis. The ELISpot assay was performed 26 days after retroviral transduction. HLA alleles of respective target cell lines are listed in Table 5.1. Effectors per well: 10,000 mTRAC+. Targets per well: 50,000. OKT-3 concentration per well: 400 ng/ml. ***Abbr.:*** *BC = buffy coat; end. = endogenous; HLA = human leukocyte antigen; KO = knock-out; mTRAC = murinized TCR alpha chain constant region; NCI = National Cancer Institute; TCR = T cell receptor.*

Supplementary Figure 26:



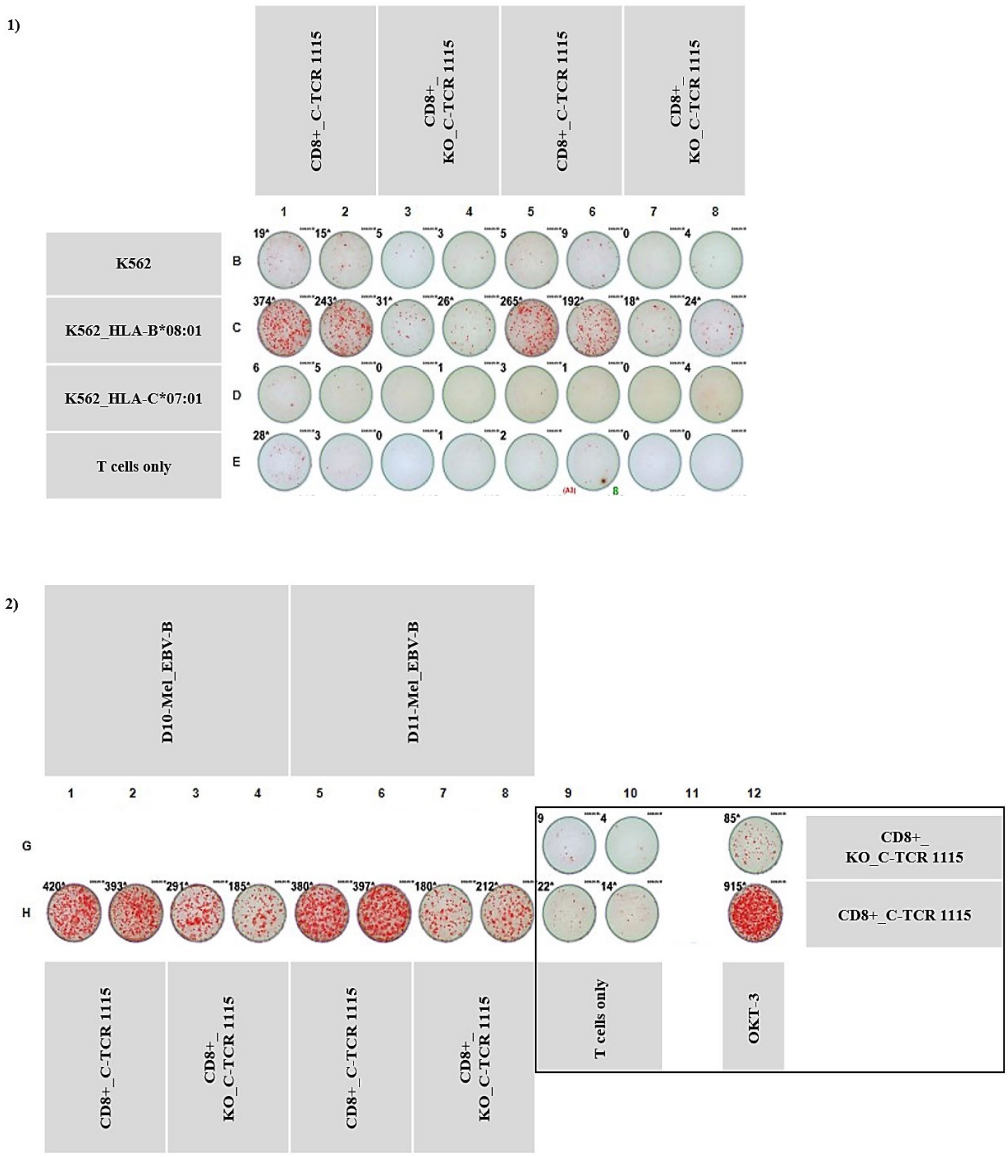
Sup. Fig. 26: Scheme of crRNA binding regions in mTRAC/mTRBC sequences.
A CRISPR/Cas9-based approach was used to knock out cluster C TCR 1115 expression in transgenic CD8+ T cell populations. Two different crRNAs were used to specifically target murinized αβ TCR constant regions of the introduced receptor. Codon-optimized mTRAC and mTRBC nucleotide sequences of C-TCR 1115 are schematically shown on top of each other. Binding sequences of both crRNAs are shown in black while corresponding PAM sequences are highlighted in red. ***Abbr.:*** *bp = base pair; crRNA = CRISPR-RNA; KO = knock-out; mTRAC = murine TCR alpha chain constant region; mTRBC = murine TCR beta chain constant region; PAM = protospacer adjacent motif; TCR = T cell receptor.*

Supplementary Figure 27:

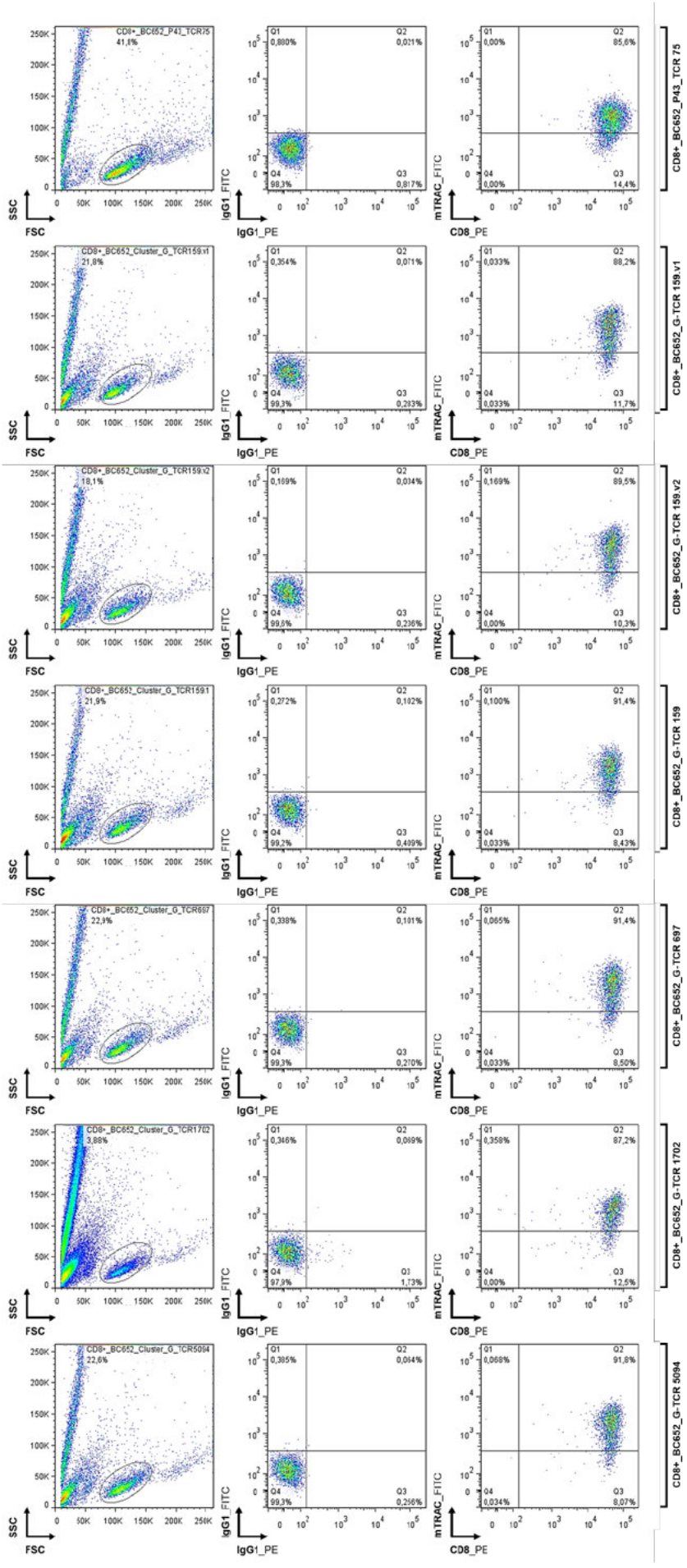


Sup. Fig. 27: Raw ELISpot data for NCI-H1703 and K562 recognition testing with TCR 1115-KO effector populations shown in Fig. 3.4.5 A.
IFN-γ secretion was used as a marker for T cell recognition. IFN-γ background levels were assessed by analyzing effector populations in the absence of target cells (T cells only preparations). General effector functionality was assessed via non-specific activation with OKT-3. Non-used wells and wells containing non-relevant antigen testing information were excluded for final evaluation. BC652-derived T cell populations had been used for a CRISPR/Cas9-based end.TCR-KO approach (Sup. Fig. 14 B) before retroviral transduction with the C-TCR 1115 expression construct (Sup. Fig. 22). TCR 1115-KO effector populations were established from transgenic T cell populations by using another CRISPR/Cas9-based KO approach specifically targeting murinized αβ TCR constant regions (Sup. Fig. 26). TCR 1115-positive effector numbers per well were calculated based on mTRAC staining performed 1-2 days before ELISpot analysis. The ELISpot assay was performed 24 days after retroviral transduction. Cluster C TCR 1115-KO was performed 7 days before conducting the ELISpot assay. Effectors per well: 10,000 mTRAC+ or 10,000 TCR 1115-KO T cells. Targets per well: 50,000. OKT-3 concentration per well: 400 ng/ml. Retrospective corrections of inaccurately counted wells are highlighted in green font. ***Abbr.:*** *BC = buffy coat; end. = endogenous; HLA = human leukocyte antigen; KO = knock-out; Mel = melanoma; NCI = National Cancer Institute; TCR = T cell receptor.*

Supplementary Figure 28:



Sup. Fig. 28: Raw ELISpot data for K562 and EBV-B cell recognition testing with TCR 1115-KO T cells shown in Fig. 3.4.5 B.
IFN-γ secretion was used as a marker for T cell recognition. IFN-γ background levels were assessed by analyzing effector populations in the absence of target cells (T cells only preparations). General effector functionality was assessed via non-specific activation with OKT-3. Non-used wells and wells containing non-relevant antigen testing information were excluded for final evaluation. BC652-derived T cell populations had been used for a CRISPR/Cas9-based end.TCR-KO approach (Sup. Fig. 14 B) before retroviral transduction with the C-TCR 1115 expression construct (Sup. Fig. 22). TCR 1115-KO effector populations were established from transgenic T cell populations by using another CRISPR/Cas9-based KO approach specifically targeting murinized αβ TCR constant regions (Sup. Fig. 26). ELISpots 1) and 2) were simultaneously performed on two plates by using identical effector populations. TCR 1115+ effector numbers per well were calculated based on mTRAC staining performed 1-2 days before ELISpot analysis. The ELISpot assay was performed 28 days after retroviral transduction. The cluster C TCR 1115-KO was performed 11 days before conducting the ELISpot assay. Due to perforated PVDF membranes, OKT-3 stimulations had to be performed in a single well format. Effectors per well: 10,000 mTRAC+ or 10,000 TCR 1115-KO T cells. Targets per well: 50,000. OKT-3 concentration per well: 400 ng/ml. Retrospective corrections of inaccurately counted wells are highlighted in green font.
Abbr.: *BC = buffy coat; EBV = Epstein-Barr virus; end. = endogenous; HLA = human leukocyte antigen; KO = knock-out; Mel = melanoma; mTRAC = murinized TCR alpha chain constant region; NCI = National Cancer Institute; PVDF = Polyvinylidene fluoride; TCR = T cell receptor.*



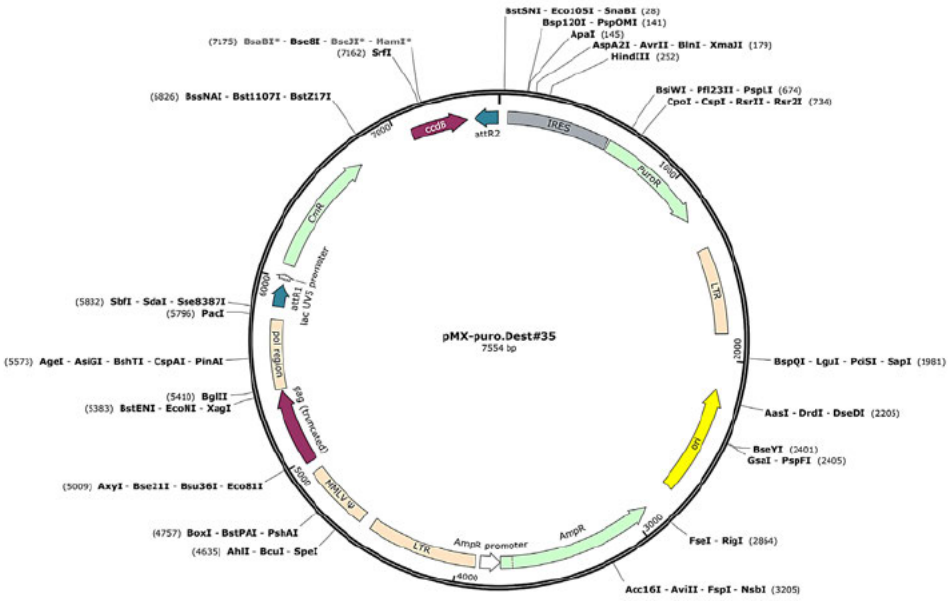
Supplementary Figure 29:

Sup. Fig. 29: Generation of G-TCR+ T cell populations from BC652.
BC652-derived CD8+ T cells were used for retroviral transductions with cluster G TCR expression constructs after being applied to an end.TCR-KO approach (Sup. Fig. 14 B). G-TCR expression was determined via mTRAC staining. P43 TCR 75-transduced effector populations were used as a positive control for mTRAC staining. Depending on the generated effector populations, between 87.2% and 91.8% of T cells expressed the introduced TCRs, respectively. Flow cytometry staining was performed 16 days after retroviral transduction. Flow cytometry data was gated based on viable T cell populations and non-specific isotype (IgG1) controls. ***Abbr.:*** *BC = buffy coat; end. = endogenous; FITC = Fluorescein isothiocyanate; FSC = forward scatter; KO = knock-out; mTRAC = murine T cell receptor alpha chain constant region; PE = Phyco-erythrin; SSC = sideward scatter; TCR = T cell receptor.*

7.3 Vector Maps

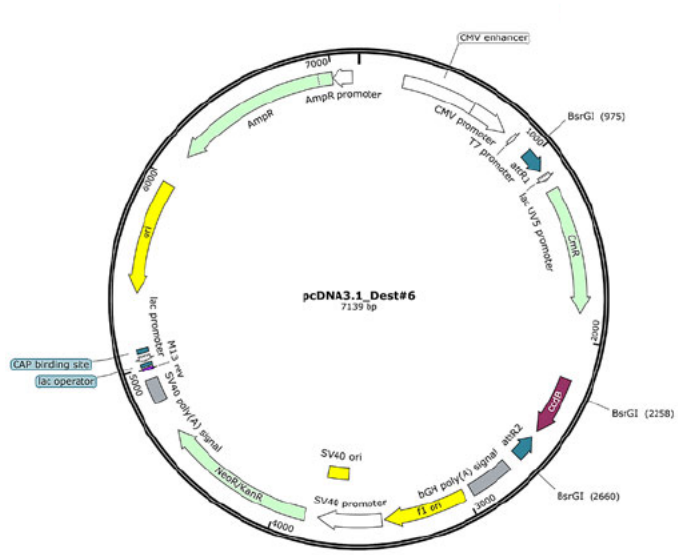
As per Table 5.5, three different expression vector backbones were used during this study: While *pMX-puro.Dest#35* was used for retrovirally transducing human healthy donor T cells and other cell lines (see section 6.4.10), *pTwist CMV BetaGlobin WPRE Neo* and *pcDNA3.1_Dest#6* were used for the transient transfection of HEK 293T targets (see section 6.4.12) used for further antigen recognition analyses. Original *pMX-puro.Dest#35* and *pcDNA3.1_Dest#6* vectors were further modified by [redacted] (UMC Mainz, Mainz, Germany) and resulting vector maps are shown below. All vector maps were obtained from [redacted] and were designed by using the SnapGene7.1 software (GSL Biotech LLC, Boston, USA). *pTwist CMV BetaGlobin WPRE Neo* vectors were purchased form Twist Bioscience (South San Francisco, USA) and were not modified at all. The corresponding vector map can be found on the Twist Bioscience website (https://www.twistbioscience.com).

pMX-puro.Dest#35



Annotations: *AmpR = ampicillin resistance gene; attR1/R2 = recombination sites for Gateway cloning; ccdB = ccdB suicide gene; CmR = Chloramphenicol resistance gene; gag = group antigens, IRES = internal ribosome entry site; pol = polymerase; PuroR = puromycin resistance gene; LTR = long terminal repeat element;* MMLVψ *= Moloney murine leukemia virus packaging signal; ori = origin of replication.*

pcDNA3.1_Dest#6



Annotations: *AmpR = ampicillin resistance gene; attR1/R2 = recombination sites for Gateway cloning; ccdB = ccdB suicide gene; CmR = Chloramphenicol resistance gene, CMV = cytomegalovirus; lac = lactose operon; M13.rev = binding site for M13 rev. sequencing primer; NeoR/KanR = kanamycin/neomycin resistance gene; ori = origin of replication.*

7.4 List of Abbreviations

AA	amino acid
Ab	antibody
ABL1	Abelson murine leukemia viral oncogene homolog 1
abbr.	abbreviation
ACT	adoptive cell therapy
AEC	3-Amino-9-Ethylcarbazole
ALL	acute lymphoblastic leukemia
Ampho	amphotropic
AP-1	Activator protein-1
APC	professional antigen-presenting cell
approx.	approximately
ATCC	American Type Culture Collection
BC	buffy coat
BCR	breakpoint cluster region
BLAST	Basic Local Alignment Search Tool
bp	base pairs
BR	broad range
BSA	bovine serum albumin
C_α/C_β	T cell receptor alpha/beta chain constant gene segment
ca.	circa
CAR	chimeric antigen receptor
CBS	Center for Biological Sequence Analysis
cDNA	complementary DNA
CDR	complementarity-determining region
CDS	coding DNA sequence
CEA	carcinoembryonic antigen
CG	cancer germline
CLL	chronic lymphocytic leukemia
CML	chronic myelogenous leukemia
CMV	cytomegalovirus
compl.	complete
COSMIC	Catalogue of Somatic Mutations in Cancer
CRC	colorectal carcinoma
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeat
crRNA	CRISPR-RNA
CRS	cytokine release syndrome
CRT	chemoradiotherapy
CT antigen	cancer testis antigen
CTL	cytotoxic T lymphocyte
DAG	diacylglycerol
D_β	T cell receptor beta chain diversity gene segment
DC	dendritic cell
DEPC	diethyl pyrocarbonate
Dest	destination
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
$d.H_2O$	deionized Water
EB	elution buffer
EBV	Epstein-Barr virus
EBV-B	EBV-infected B cells
ECACC	European Collection of Authenticated Cell Cultures
EC50	median effective concentration
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
ELISpot	enzyme-linked immunospot
EMBL-EBI	European Molecular Biology Laboratory's European Bioinformatics Institute
Emv	ecotropic murine leukemia virus

end.	endogenous
EntS	Enterobactin exporter S
env	envelope
ER	endoplasmatic reticulum
ERBB2	Erythroblastic oncogene B
Exo.	exogenous
FACS	fluorescence activated cell sorting
FCS	fetal calf serum
FFPE	formalin-fixed, paraffin-embedded
Fig.	figure
FITC	fluorescein isothiocyanate
for.	forward
gag	group antigens
GALV	gibbon ape leukemia virus
GaM	goat anti-mouse
gDNA	genomic DNA
GLIPH	grouping lymphocyte interactions by paratope hotspots
GP100	Glycoprotein of 100 kDA
gRNA	guide RNA
GTEx	The Genotype-Tissue Expression Project
GTP	guanosine-5'-triphosphate
HDR	homology directed repair
HEK	human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HERV	human endogenized retroviruses
HER2	Human epidermal growth receptor 2
HLA	human leukocyte antigen
HPV	human papillomavirus
HRAS	Harvey rat sarcoma viral oncogene
HS	human serum
HSD	HS Diagnomics GmbH
hTRAC	human T cell receptor constant region
Hz	hertz
ICI	immune checkpoint inhibition
ID	identifier
IEDB	The Immune Epitope Database
IE2	Immediate early protein 2
IFN-γ	interferon gamma
IL	interleukin
InDels	insertions and deletions
Intogen	Integrative onco genomics
IP_3	inositol tripohosphate
irAE	immune related adverse event
IRES	internal ribosomal entry site
ITAM	immune receptor tyrosine-based activation motifs
Itk	IL-2-inducible T cell kinase
J_α/J_β	TCR alpha/beta chain joining gene segment
kb	kilobase
KI	knock-in
KLD	kinase-ligase-DpnI
KO	knock-out
KRAS	Kirsten rat sarcoma viral oncogene
LAC/LUAD	lung adenocarcinoma
LAT	Linker of activated T cells
LB	lysogeny broth
LC	lung cancer
Lck	Lymphocyte-specific protein tyrosine kinase
LFA-1	Llymphocyte function-associated antigen 1
LKB1	Liver kinase B1
LMP2A	Latent membrane protein 2A
LUSC	lung squamous cell carcinoma
mAb	monoclonal antibody

MACS	magnetic activated cell sorting
MAGE	Melanoma antigen
MANA	mutation-associated neoantigens
MAPK	mitogen-activated protein kinase
MDSC	myeloid-derived suppressor cell
Mel	melanoma
Melan-A/MART-1	Melanoma antigen recognized by T cells 1
MHC	major histocompatibility complex
ml	milliliter
MLV	murine leukemia virus
mm	millimeter
MNV	multiple nucleotide variation
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	messenger RNA
MS	mass spectrometry
NCBI	National Center for Biotechnology Information
NEB	New England Biolabs
Neo	neomycin
NFAT	Nuclear factor of activated T cells
NFκB	Nuclear factor kappa-light-chain enhancer of activated B cells
NGS	next-generation sequencing
NIAID	National Institute of Allergy and Infectious Diseases
NLS	nuclear localization signal
NRAS	Neuroblastoma rat sarcoma viral oncogene
ns	non-synonymous
NSCLC	non-small cell lung cancer
nt	nucleotide
NY-ESO-1	New York esophageal squamous cell carcinoma-1
ODN	oligodeoxynucleotide
ORF	open reading frame
OTR	orthotopic TCR replacement
PBMC	peripheral blood mononuclear cell
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDAC	pancreatic ductal adeno carcinoma
PD-L1	Programmed death-ligand 1
PD-1	Programmed cell death protein 1
PE	phycoerythrin
PI3K	Phosphatidylinositol 3-kinase
PLC-γ	Protein phospholipase C-γ
PMA	phorbol myristate acetate
pol	polymerase
PS	penicillin-streptomycin
Puro	puromycin
PVDF	polyvinylidene fluoride
p.Q61E	glutamine to glutamic acid substitution at amino acid position 61
p.Q61H	glutamine to histidine substitution at amino acid position 61
p.Q61K	glutamine to lysine substitution at amino acid position 61
p.Q61L	glutamine to leucine substitution at amino acid position 61
p.Q61P	glutamine to proline substitution at amino acid position 61
p.Q61R	glutamine to arginine substitution at amino acid position 61
p53	Tprotein p53
REP	rapid expansion protocol
RCF	relative centrifugal force
rev.	reverse
RIN	RNA integrity number
RNA	ribonucleic acid
RNA-Seq	RNA sequencing
RNP	ribonucleoprotein
RPKM	reads per kilobase of transcript per million mapped reads
rpm	rotations per minute
RPMI	Roswell Park Memorial Institute

rRNA	ribosomal RNA
RT	room temperature
s	seconds
SC	single cell
SCLC	small cell lung cancer
SDM	site-directed mutagenesis
Seq	sequencing
SH2	Src Homology 2
SLP-76	SH2 domain-containing leukocyte protein of 76 kDa
SNV	single nucleotide variation
SOC	super optimal broth with catabolite repression
ss	single-stranded
ssODN	single-stranded oligodeoxynucleotide
sup	supplementary
S2	biosafety level 2
TAA	tumor-associated antigen
TAE	tris-acetat-EDTA
TAP	Transporter associated with antigen processing
TBE	tris-borat-EDTA
TCGA	The Cancer Genome Atlas
TCR	T cell receptor
TEIPP	T cell epitopes associated with impaired peptide processing
TIL	tumor-infiltrating lymphocyte
TKI	tyrosine kinase inhibitor
TLS	tertiary lymphoid structure
TM	primer annealing temperature
TME	tumor microenvironment
TMEM161A	Transmembrane protein 161A
TMG	tandem minigene
TRA	TCR alpha chain
TRAC	TCR alpha chain constant region
tracrRNA	trans-activating crRNA
TRAV	TCR alpha chain variable region
TRB	TCR beta chain
TRBC	TCR beta chain constant region
TRBV	TCR beta chain variable region
Treg	regulatory T cell
TS	touch screen
TSA	tumor-specific antigen
T/nT ratio	tumor-to-non-tumor ratio
U	units
UK	United Kingdom
UMC	University Medical Center
US/USA	United States of America
UV	ultraviolet
V_α/V_β	T cell receptor alpha/beta chain variable gene segment
VDJdb	VDJ TCR database
WES	whole exome sequencing
WPRE	woodchuck hepatitis virus post-transcriptional regulatory element
WT	wild-type
WT1	Wilms tumor 1 gene
w/v	weight per volume
xMHC	extended MHC gene cluster
ZAP-70	ζ-associated protein of 70 kDa

7.5 List of Figures

7.6 List of Supplementary Figures

7.7 List of Tables

7.8 Contributions

The study summarized in this document was collaboratively designed and executed by representatives of (i) the Internal Medicine III department of the University Medical Center Mainz [M] (Mainz, Germany), (ii) HS Diagnomics GmbH [H] (Berlin, Germany), (iii) TheryCell GmbH [T] (Berlin, Germany), (iv) the Institute of Pathology - Charité [C] (University Medical Center Berlin, Berlin, Germany), (v) the Thorax Surgery department - Vivantes Clinic Neukölln [V] (Berlin Germany), and (vi) the Cancer Immunology & Immune Modulation department at the Boehringer Ingelheim Regional Center Vienna [B] (Vienna, Austria). In particular, [redacted] [M, H, T], [redacted] [H, T], [redacted] [M], [redacted] [H, C], [redacted] [H, T], [redacted] [B] and myself [M] substantially contributed to the study design. Furthermore, all patient materials were provided by [redacted] [V] and [redacted] [V] while funding was acquired by [redacted] [B], [redacted] [M], and [redacted] [H, T]. Supervision of the study was provided by [redacted] [M, H, T], [redacted] [M], [redacted] [M], [redacted] [B], [redacted] [B], [redacted] [H, T], and [redacted] [H, T].

TCR cloning, T cell profiling, 10X Genomics single cell RNA sequencing, and corresponding data analyses were performed by representatives of HS Diagnomics GmbH and TheryCell GmbH. *In silico* neoantigen identifications via WES and RNA-Seq, peptide-binding predictions, and gene expression analyses were conducted and evaluated by [redacted] [M, H, T]. All shown antigen screenings and corresponding experiments, such as molecular cloning, genome modification, transient transfection, cell maintenance, tissue processing, or flow cytometry, were performed in laboratories of the Wölfel research group at the University Medical Center (UMC) in Mainz. While most of the hands-on laboratory work was carried out by myself, also other members of the Wölfel group were substantially involved in both planning and execution of the conducted experiments. First of all, [redacted] [M] significantly contributed to the implementation of various genetic engineering approaches as well as sequencing data analysis. Furthermore, [redacted] [M] and [redacted] [M] actively supported my work when performing extensive antigen screening experiments, corresponding cell maintenance, or flow cytometry analyses.

7.9 Acknowledgements

At first and foremost, I would like to sincerely thank my doctoral advisor, [redacted] for offering me the opportunity to work under his direct supervision. I greatly appreciate the chance of being able to work in such a prestigious and well-established research group. Furthermore, I would like to thank [redacted] for critically reading and grading this thesis as well as for providing his scientific guidance throughout this project.

Equally, I want to thank my graduate school *TransMed* (Mainz Research School of Translational Biomedicine) and all additional members of my thesis advisory committee, namely [redacted], [redacted], [redacted], and [redacted], who provided consistent scientific guidance and supported this study with highly valuable input throughout the years. I am especially grateful to [redacted] for agreeing to be my second supervisor, thus also being responsible for the critical evaluation of my work. I would also like to thank all future members of my examination board for participating in my oral examination, especially [redacted] who agreed to serve as my second assessor.

I greatly thank HS Diagnomics (Berlin, Germany) for planning and financing this project, providing TCR expression constructs, and conducting most of the bioinformatical analyses, thereby making this project even possible in the first place. Here, I would especially like to thank [redacted] and [redacted], who both brought in all their scientific experience to drive the progress of this project. However, most of all, I would like to thank [redacted] for directly supervising this project and my practical work in the lab. I very much appreciated the critical discussions, which really helped me to better understand the scientific background of my work and the overall field of cancer immunotherapies. Thank you for your dedicated supervision, I have no doubt that the success of this project would not have been possible without your scientific expertise in both immunology and bioinformatics!

My sincere thanks also go to all members of the Wölfel group: Thank you for the general support in the lab and for critically discussing all results, ideas, and problems that occurred during this project. Especially, I want to thank [redacted] for her long-standing support in solving both immunology- as well as molecular biology-related issues. You really managed to make cDNA cloning look easy, even in the most complex experimental setups. Furthermore, my heartfelt thanks to [redacted] for her dedicated support (technically and mentally) during highly complex and extensive experimental setups. I certainly couldn't have managed the thousands of ELISpot wells on my own! My sincere thanks also go to all members of the contiguous research groups, who always offered their kind help, technical support, and shared materials or protocols whenever needed. First and foremost, I would really like to thank [redacted] for thoroughly supervising various flow cytometry experiments and for making the office work a blast.

My most special thank-you and infinite gratitude is dedicated to my friends and family, above all my parents and my sister, [redacted]. Thank you for the constant motivation that, without a doubt, substantially contributed to the success of my doctorate. Last but most importantly, I would like to thank my wonderful partner, [redacted], for always having my back during this journey and for providing a level of mental support that I didn't know existed before. The mere fact that I met you during my doctorate made it all worth it!

7.10 Statement in Lieu of an Oath

I herewith declare that I wrote this dissertation all by myself and that I have not sought any paid assistance from third parties to author this document. Furthermore, I declare that I did not use any other media, materials, or resources than the ones referenced in this document. All text passages referring to any of the used resources or literature were cited correctly and to the best of my knowledge. Additionally, I declare that this dissertation has not yet been submitted to any other university, examination board, or comparable institution for the award of an academic degree. Moreover, I assure that I have not yet unsuccessfully or successfully completed a doctoral, Ph.D., or comparable graduation program of any kind.

At the time this thesis was finalized, some of the results summarized in this document were already submitted for publication to the *Frontiers in Immunology (Avenue du Tribunal Fédéral 34, 1005 Lausanne, Switzerland)* editorial board. Author contributions will be clearly specified in the corresponding article. However, the peer review process of the submitted manuscript (title: T-cell receptors identified by a personalized antigen-agnostic screening approach target shared neoantigen KRAS Q61H; assigned manuscript ID: 1509855) is currently still ongoing.

Christoph Doppler Digital unterschrieben von Christoph Doppler
Datum: 2025.06.24 18:21:09 +02'00'

Date and Signature

8) Curriculum Vitae

[redacted]