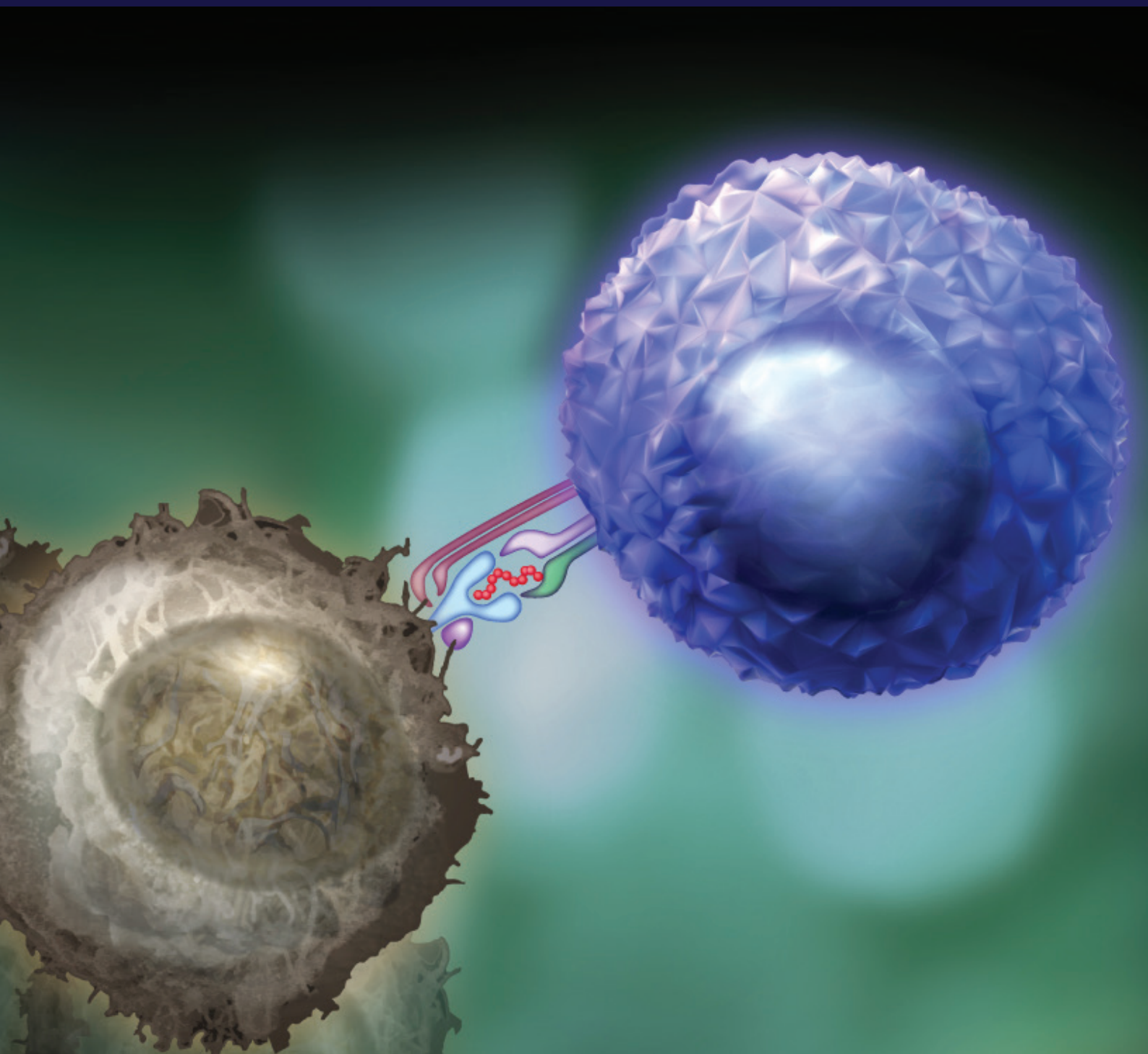
DTU Health Tech
Department of Health Technology

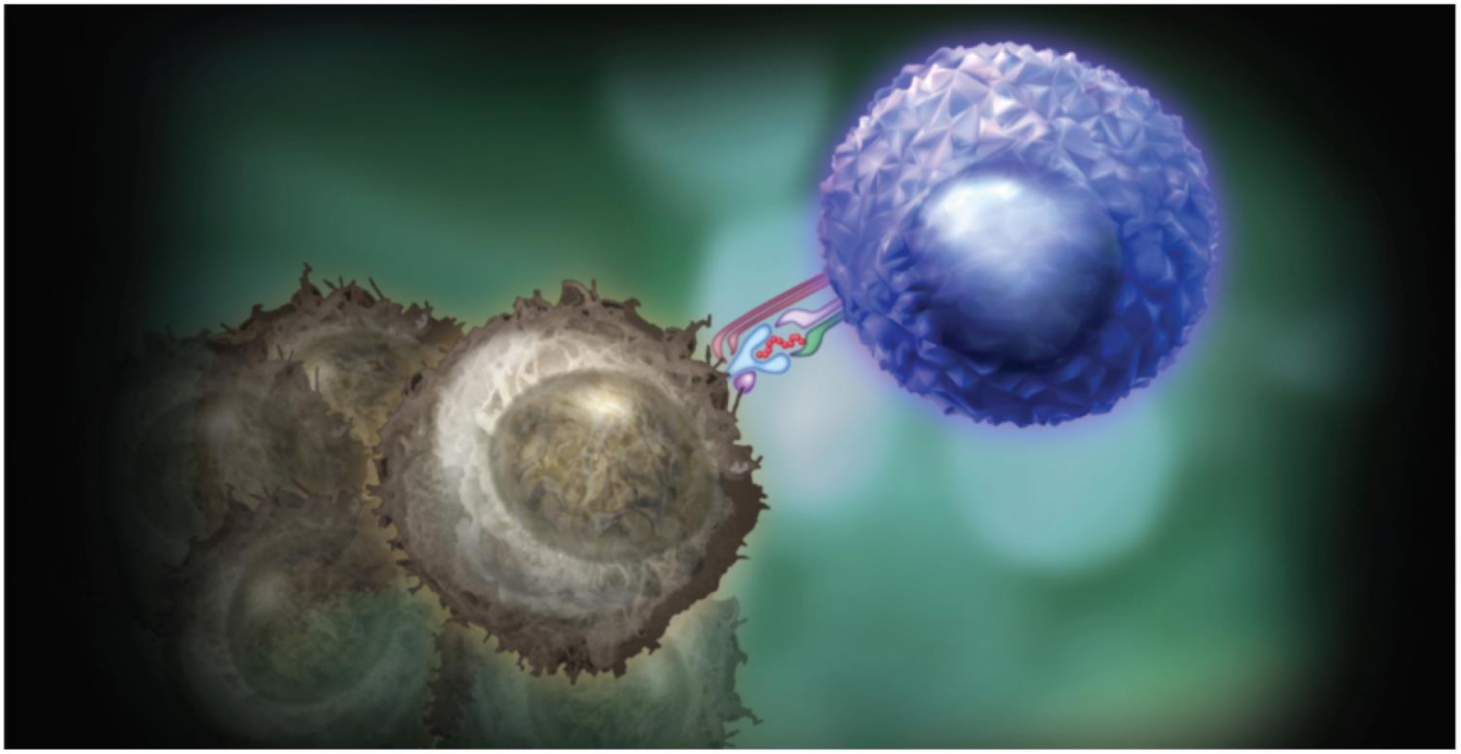
Novel technologies for interrogating T cell recognition

Amalie Kai Bentzen, PhD Thesis January 2019



Novel technologies for interrogating T cell recognition

PhD thesis



Amalie Kai Bentzen
January 2019

Front page art: Nanna Bild

Novel technologies for interrogating T cell recognition

Amalie Kai Bentzen

PhD Thesis
Kongens Lyngby 2019

Preface

This thesis has been submitted to the Technical University of Denmark, the Department of Health Technology, as part of the requirements for obtaining the PhD degree. The research included was conducted first at the Veterinary Institute, later the Institute of Micro- and Nanotechnology and finally at the newly constituted Department of Health Technology. The work was carried out from December 2014 to January 2019 under supervision of Professor Sine Reker Hadrup.

The thesis consists of two parts: first a section that introduces concepts essential for understanding the scope of the thesis and discuss the impact and perspectives of the research included. This is followed by a section composed of the research included in the thesis; presented in the papers or manuscripts listed on the next page. Moreover, a number of manuscripts and patent applications that were prepared during the PhD study are listed but are not included as part of this thesis.

Amalie Kai Bentzen
Kongens Lyngby, January 2019

Research papers

Research included in the thesis

I. **Bentzen AK**, Marquard AM, Lyngaa R, Saini SK, Ramskov S, Donia M, Such L, Furness AJS, McGranahan N, Rosenthal R, thor Straten P, Szallasi Z, Svane IM, Swanton C, Quezada SA, Jakobsen SN, Eklund AC, Hadrup SR. Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nature Biotechnology* **34**, 1037–1045 (2016)

II. **Bentzen AK** and Hadrup SR. Evolution of MHC-based technologies used for detection of antigen-responsive T cells. *Cancer Immunol Immunother* **66**, 657-666 (2017)

III. **Bentzen AK**, Such L, Jensen KK, Marquard AM, Jessen LE, Miller NJ, Church CD, Lyngaa R, Koelle DM, Becker JC, Linnemann C, Schumacher TNM, Marcatili P, Nghiem P, Nielsen M, Hadrup SR. T cell receptor fingerprinting enables in-depth characterization of the interactions governing recognition of peptide–MHC complexes. *Nature Biotechnology* **36**, 1191–1196 (2018)

IV. **Bentzen AK**, Such L, Marquard AM, Kristensen NP, Heeke C, Jakobsen, SN, Eklund AC, Hadrup SR. Interrogating T cell recognition by use of DNA barcode-labeled peptide-MHC multimers. *Manuscript in preparation. Formatted for submission to Nature Protocols*

V. **Bentzen AK**, Marquard AM, Holm JS, Hansen UK, Hadrup SR. Simultaneous identification of antigen-specificity and T cell receptor sequence in single cells. *Data from an ongoing proof-of-concept study*

Papers not included in the thesis

i. Pedersen NW, Holm A, Kristensen NP, Bjerregaard A-M, **Bentzen AK**, Tamhane T, Burgdorf KS, Ullum H, Jennum P, Knudsen S, Hadrup SR, Kornum BR. Broad presence of autoreactive CD8 T cells against narcolepsy type 1-relevant proteins. Accepted for publication in *Nature Communications*, January 2019.

ii. Jurtz VI, Jessen LE, **Bentzen AK**, Jespersen MC, Mahajan S, Vita R, Jensen KK, Marcatili P, Hadrup SR, Peters B, Nielsen M. NetTCR: sequence-based prediction of TCR binding to peptide-MHC complexes using convolutional neural networks. *BioRxiv* (2018) doi.org/10.1101/433706

Patent applications

- **Bentzen AK**, Hadrup SR, Pedersen H, Jakobsen SN. Determining Antigen Recognition through Barcoding of MHC Multimers. International Application No. WO2015185067

- **Bentzen AK**, Hadrup SR, Johansen KH, Pedersen H, Jakobsen SN. Detection and isolation of specific cells by binding of labelled molecules. International Application No. WO2015188839

Abstract

The studies presented in this thesis have all involved the development of novel technologies for interrogating different aspects of antigen-specific T cell recognition.

The immune system's potential for eliminating intruding pathogens or malignancies is largely governed by the ability of T cells to specifically recognize and respond to foreign molecular structures. The activity of T cells is guided by specific interactions of the surface expressed T cell receptor (TCR) with peptides presented in the context of MHC molecules on the surface of target cells. Hence, for broad protection of virtually any encountered pathogen or cellular malignancy, the TCR repertoire of each individual must be able to recognize and adapt to an extensively large and unpredictable range of foreign peptides, while avoiding immunopathology caused by autoreactivity. The scope of this PhD has been to develop novel technologies that enable us in describing peptide-MHC-TCR interactions from a broader and more detailed view, which aims at integrating the complexity of the range of peptide-MHCs potentially recognized by T cells, as well as the cross-recognition potential of the individual T cell.

The implementation of molecular barcodes as labels to trace peptide-MHC-TCR interactions provide the foundation for all the research included in this thesis. In a first study, I show that unique DNA barcodes can be used to label specific peptide-MHC multimer reagents, and subsequently enable tracing of the antigen-specificities of T cells within a heterogeneous sample. With this strategy I show that it is possible to screen for more than 1,000 different peptide-MHC specificities in parallel, which provides completely new possibilities for assessing immune recognition in an individualized manner. In a second study, I discuss the advantages of assessing T cell recognition at such complexities, as well as current shortcomings. In a third study I address the cross-recognition potential of individual T cells through establishment of a technology that assess the relative affinity of a TCR to several MHC multimers carrying related peptide. This in turn enabled the identification of essential molecular interaction-points of TCRs and identification of cross-recognized peptides. The fourth study comprises a protocol which should enable other researchers to understand details of the technologies reported in study one and three, and ultimately assist them in performing the same types of analyses. The data reported in the fifth study comprise an early finding that show the simultaneous identification of peptide-MHC specificity and TCR sequence of many single cells in parallel. This will be implemented in a high-throughput format and hereby facilitate the pairing of numerous peptide-MHC specificities with the corresponding TCR sequences, which in turn can provide a first step towards resolving the relationship between the TCR sequence and the peptide-MHC binding motif. Collectively, the studies reported in this thesis provide a foundation for others to develop new and personalized

therapeutic schemes for immune interventions, and for resolving some of the complex relationships existing between T cells and their targets.

Dansk resumé

Studierne præsenteret i denne afhandling, har alle involveret udviklingen af nye teknologier til at undersøge og forstå forskellige aspekter af antigenspecifik T-cellegenkendelse.

Immunsystemets potentiale til at eliminere patogener eller maligne celler er i høj grad bestemt af T-cellers evne til specifikt at genkende og reagere på fremmede molekylære strukturer. T-cellers aktivitet er styret ud fra specifikke interaktioner mellem de overflade-udtrykte T-celle-receptorer (TCR'er), og peptider præsenteret i konteksten af MHC-molekyler, som selv er udtrykt på overfladen af alle kroppens celler. For at sikre beskyttelse mod ethvert forekommende patogen eller cellulære misdannelser, er det derfor essentielt at TCR-repertoiret i hvert enkelt individ er i stand til at genkende og tilpasse sig et omfattende stort og uforudsigeligt udpluk af fremmede peptider, samtidig med at immunpatologi forårsaget af autoreaktivitet skal undgås. Formålet med denne Ph.d. har været at udvikle nye teknologier, der gør det muligt at beskrive peptid-MHC-TCR-interaktioner bredere og mere detaljeret end tidligere, og som forsøger at integrere kompleksiteten af kombinationen af de peptid-MHC'er, som præsenteres for T celler, såvel som de enkelte T cellers potentiale for krydsgenkendelse.

Implementeringen af molekylære stregkoder, brugt til at opmærke og spore peptid-MHC-TCR-interaktioner, danner grundlaget for al den forskning, der indgår i denne afhandling. I et første studie viser jeg, at unikke DNA-stregkoder kan anvendes til at mærke specifikke peptid-MHC multimer-reagenser, og dermed gøre det muligt at spore antigenspecificiteterne af T-celler i en heterogen prøve. Med denne strategi viser jeg, at det i en enkelt prøve er muligt at screene parallelt for mere end 1000 forskellige peptid-MHC-specificiteter, hvilket giver helt nye muligheder for at evaluere enkelte individers personlige immungenkendelse. I et andet studie diskuterer jeg fordelene ved at udføre så komplekse metoder til at evaluere T-celle-genkendelse, såvel som aktuelle svagheder. I det tredje studie behandler jeg potentialet for kryds-genkendelses af enkelte T-celler ved at etablere en teknologi, der bestemmer den relative affinitet af en TCR mod adskillige MHC multimerer der bærer beslægtede peptider. Denne viden gjorde det muligt at identificere de essentielle molekylære interaktionspunkter for en TCR, som vi benyttede til at identificere kryds-genkendte peptider. Det fjerde studie er en protokol, der skal gøre det muligt for andre forskere at opnå detaljeret forståelse af de teknologier, der er rapporteret i studie et og tre, og i sidste ende hjælpe dem med at udføre lignende analyser. Datarapporteret i det femte studie omfatter tidlige undersøgelser, der viser, at mange enkelte cellers peptid-MHC-specificitet og TCR-sekvens kan identificeres simultant. Denne teknik vil blive implementeret i en ”high-throughput” teknologi og således lette pardannelsen mellem talrige peptid-MHC specificiteter og de sekvenser, der koder for den TCR, der genkender dem. Dette vil potentielt gøre det muligt at tage det første skridt imod en dybere forståelse af forholdet mellem TCR-sekvensen og peptid-

MHC-bindingsmotivet. Samlet set giver de studier, der rapporteres i denne afhandling, grundlag for at andre kan udvikle nye og personlige immunterapeutiske behandlingsmetoder og vil hjælpe os til at forstå nogle af de komplekse relationer, der eksisterer mellem T-celler og de strukturer, de genkender.

Acknowledgements

The research included in this thesis would not have been possible without great collaborations, guidance and invaluable support from colleagues, friends and family, whom I all owe a tremendous thanks to.

My greatest thanks go to Sine; the years of working with you have been more fun, challenging and inspirational than I could ever have imagined. I am truly grateful for all I have learnt from you; for your continuous encouragement, trust and patience. I admire you for your ability to make things happen and for always having the grand overview. The greatest inspiration is your scientific passion and your optimism, which makes it so much fun to work with you. I feel very fortunate that I can look forward to future challenges and inspiring collaboration. Thank you for pushing me to develop as a scientist!

I moreover wish to thank all former and present members of Sines research group; it has been invaluable to have great colleagues to learn from, lift together and to make every day more fun. I especially want to thank Andrea for the best team work, for broadening my scientific insight and essentially for making sense of the hours spent in the lab. Lina, for truly enjoyable and fruitful collaboration, and encouragement when things were more difficult. Sofie and Natasja, for their continuous support, and for fostering the wet-lab knowledge that enabled me to work independently. Staffan, Natasja, Stine and Sofie for feedback on drafts of this thesis. Alexander, Andrea, Anna, Anne-Mette, Annie, Bente, Christina, Dorthe, Esteban, Jeppe, Julien, Keith, Laurin, Lina, Mo, Nadia, Nana, Natasja, Nikolaj, Rikke, Sara, Siri, Sofie, Stine, Sunil, Søren, Trine, Tripti, Ulla and Vibeke, for all types of scientific discussions, cake-club and parties. Thanks to Marianne and Kathrine, for practical help on everything outside the lab. Thanks to colleagues from all groups of the former Section of Immunology and Vaccinology. Collaboration with the groups of Vasileios and Bill were invaluable in establishing the laboratories at Bülowsvej. Particular thanks to Julien and John for trying to point me in the right direction whenever I have been trying something new.

I want to thank Morten, Kamilla and Leon from the former Department of Bio and Health Informatics, for enthusiastic discussions and collaboration on several projects. I look forward to our future collaborations.

I want to thank the team of Dominik Maurer at Immatics Biotechnologies, for generously having me visiting their research group. Especially Andreas Moritz, for his hospitality and technical guidance throughout my stay. Similarly, I wish to thank the lab of Ton Schumacher at the NKI, for also generously having me visiting. Thanks to Wouter Scheper for training me in their state-of-the-art technologies and for coming up with answers to all my questions.

Lastly I want to thank my family and my friends for all the encouragement throughout the years. Thanks to my father for patience in explaining everything that was difficult. My mother, for

endless support. Peer, for being curious about everything. Caroline, for generously helping always. Thank you Janus, for continuously encouraging me to dream and for inspiring me to act. Emil and Janus both, for making every day happier and so much more fun.

Contents

1 Introduction 1

2 Scope of thesis 3

3 Immunological recognition 5

TCR diversity 5

Complex T cell targets 6

T cell cross-reactivity 8

4 Detection of antigen-responsive T cells 11

Probing of disease-associated antigens 12

A need for high-throughput T-cell detection technologies 17

Functional and phenotypic analysis 20

5 Understanding T cell recognition 23

TCR binding degeneracy 23

Dissecting the specificity of TCR interactions 28

6 Molecular encoding 33

DNA-encoded library screenings 34

TCR repertoire tracking 37

Emerging single-cell technologies 38

7 Epilogue 41

Paper I 49

Paper II 77

Paper III 89

Paper IV 135

Paper V 165

Abbreviations 183

Bibliography 185

1 Introduction

Our immune system is comprised of a heterogeneous mixture of cells whose orchestrated actions ensure the repeated elimination of most threats encountered from pathogenic infections or malignant transformation of cells, without damaging healthy tissue. Surveillance is managed by T cells through interaction of their T cell receptor (TCR) with peptide fragments displayed in the Major Histocompatibility Complex (MHC) (Davis et al., 1998) on the surfaces of nucleated or antigen presenting cells. Specific interaction with cognate peptide-MHC (pMHC) activates the T cells, induces proliferation and enables them in exerting essential effector functions; specifically eliminating threats by cytotoxic killing or by potentiating other immune cells, thus initiating an adaptive immune response. The ability of the immune system to react to various threats is therefore largely dependent on the TCR repertoire that is both unique and highly diverse between individuals. Combined with the broad range of presented peptides and the genetic polymorphisms of the Human Leukocyte Antigen (HLA, human encoded MHC) this crucial diversity makes it challenging to accurately describe T cell recognition, both in terms of detecting antigen-responsive T cells and understanding the structural features that determine whether a TCR will recognize a given pMHC (Newell and Davis, 2014). However, due to the key role of T cells in eliciting an immune response, knowledge on these features of TCR recognition is crucial for understanding and monitoring immune mediated diseases, that being virus infections, autoimmunity or cancer, and for developing targeted therapies for immune interventions.

2 Scope of thesis

The scope of this thesis has been to develop technologies that allow extensive interrogations of T cell recognition at a level that better meets the inherent complexity of:

(1) The numerous possible pMHC that can be recognized by the repertoire of TCRs (**Paper I-II**)

Identification of the disease-associated pMHC antigens that are recognized by T cells may help uncover disease etiology, assist therapeutic development, facilitate immune monitoring and help predict response to a given treatment.

(2) The immense number of pMHC targets that one TCR may recognize (**Paper III**)

Knowing the structural patterns that determine a given TCRs ability to recognize different pMHCs will uncover new relationships between the diversity of antigenic targets and specific recognition by a TCR. This in turn may be applied to predict cross-recognized epitopes, relevant for reducing risk of off-target adverse events in vaccine- and precision medicine development, and can reveal potential relationships between infectious disease and autoimmunity.

(3) The many possible pairs of interacting pMHC and TCR (**Paper V**)

Pairing TCR sequence with pMHC recognition at a large scale will open the possibility for computational modeling of TCR interaction with pMHC and thus provide a first step towards resolving the relationship between the sequence of the TCR and its corresponding target(s).

Due to the complexity of T cell recognition I have approached these challenges through the development of high-throughput strategies that allow the parallel interrogation of many cells and molecules at once. The thesis therefore also comprises a protocol (**Paper IV**) that is amenable for high-throughput discovery using the methodologies applied in Paper I and III.

The combined output from such tools will give a more complete description and broader overview of the molecular features of T cell recognition. Knowledge on how the TCR repertoire of individuals work cooperatively will especially help understand how immune-mediated diseases affect each other, and will assist ongoing efforts for developing disease interventions that are tailored to the potential of an individual's TCR repertoire.

3 Immunological recognition

Adaptive immunity relies on the ability to mount an antigen-specific response to virtually all pathogens and cellular malformations. Diversity of the immune receptor repertoire is a prerequisite for broad antigen-specific recognition, crucial for our ability to recognize and specifically eliminate even targets as diverse as viruses that evolve at an incredibly fast pace (Lauring et al., 2013).

The immune systems potential for recognizing such threats is influenced by (1) the diversity of the TCR repertoire, and (2) the potential range and context of the displayed peptides. In turn, diversity of the latter is affected by the variations of the disease-associated peptides and by the HLA polymorphisms of the individual.

TCR diversity

The remarkably high diversity of the TCR repertoire is provided through somatic gene rearrangements and subsequent pairing of α and β chains (or less conventionally γ and δ chains) to produce a heterodimeric TCR with a unique potential for antigen recognition (Davis and Bjorkman, 1988; Marrack and Kappler, 1987). In general terms, it is assumed that every individual T cell express only one unique TCR. During thymocyte development each chain is generated independently through somatic recombination of variable (V), diversity (D), joining (J) and constant (C) regions (Figure 1). The variation provided from almost random combinations of V(D)J segments is further diversified from the imprecise joining of their ends, which result in unique insertions and deletions in each chain (Olaru et al., 2004; Quigley et al., 2010; Schatz, 2004; Tonegawa, 1983). Heterodimeric pairing of these randomly generated αβ chains theoretically exceed 10^{15} unique TCRs (Davis and Bjorkman, 1988). It is however estimated that the actual number in an individual is in the range 10^{6}-10^{8} unique TCRs (Arstila et al., 1999; Robins et al., 2009).

As a result of the somatic recombination, each chain of the TCR produces three complimentary-determining region (CDR) loops that account for main interaction with pMHC. Each CDR1 and CDR2 of the α- and β-chains is generated from the germline encoded V-gene segments, and they are thus conventionally thought to govern most interaction with MHC, hereby determining the MHC restriction of the receptor. Conversely, the CDR3α and CDR3β loops are hypervariable,

generated from the immensely diverse joining regions of the somatically rearranged gene segments, and are supposed to predominantly guide the specific engagement of the TCR with the peptide fragment (Feng et al., 2007; Garboczi et al., 1996; Garcia et al., 1999). However this general view is challenged by several examples of TCRs with CDR3 loops primarily engaging with MHC, as well as examples of CDR1 and CDR2 loops interacting with peptide, which in turn may reflect a conformational plasticity of TCRs that enable interaction with different pMHC epitopes (Burrows et al., 2010; Dendrou et al., 2018; Garcia et al., 1998; Hahn et al., 2005; Holland et al., 2018; Miles et al., 2015; Reiser et al., 2002, 2003; Rossjohn et al., 2015; Song et al., 2017; Yin et al., 2011).

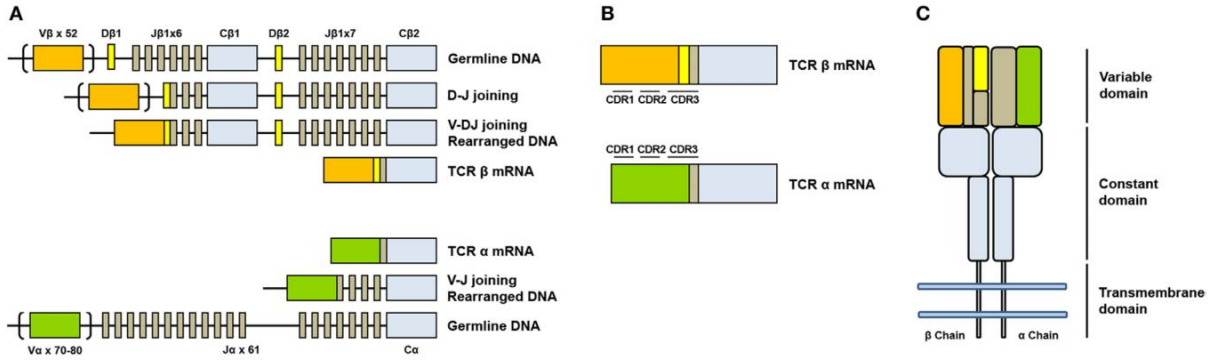


Figure 1 | **Somatic recombination of TCRs.** (**A**) V(D)J and C regions of both the α and the β encoding genes are progressively rearranged during thymic development of T cells (**B**) ultimately producing a set of highly diverse α and β transcripts. (**C**) The surface expressed TCR is composed of a paired α and β chain that is organized with a variable region that largely accounts for the antigen-specificity of that receptor. From De Simone et al., 2018.

Complex T cell targets

The dual nature of the TCR recognition motif, which implies direct contact of the TCR with peptide and MHC simultaneously, adds substantially to the complexity of T cell recognition. On one side there is the potential diversity of the primary peptide sequence, on the other a large repository of different MHC-alleles, of which the TCR must interact specifically with both elements for recognition of antigen (Bevan, 1977; Zinkernagel and Doherty, 1974).

The MHC restriction is remnant of the sculpting of each individual's TCR repertoire (Rangarajan and Mariuzza, 2014); a thymic developmental process of positive and negative selection that serves to rescue only TCRs that are able to bind an MHC belonging to that individual, while disposing T cells with receptors that are (too) reactive to self (Hong et al., 2018; Ignatowicz et al., 1996a, 1996b; Starr et al., 2003a, 2003b). Broadly stated a TCR will interact with peptides bound by the same MHC haplotype that they were originally selected for

(Dendrou et al., 2018), but each haplotype has the ability to bind a large range of diverse peptide ligands ensuring presentation of a sufficient proportion of peptides for displaying the health status of a cell (Mason, 1998; Rudensky AYu et al., 1991). Peptide binding to a given MHC haplotype is primarily dictated by an invariant preference for a few amino acids at the anchor positions, most often at positions 2, (3) and ∞ (Bouvier and Wiley, 1994). Apart from the anchors, a peptide may be composed of almost any sequence that can adapt to an overall linear conformation in the MHC binding groove (Cohn, 2016). Correspondingly, T cell specificity is expected to be particularly influenced by the length of the peptide, the varying amino acids at the central positions and the MHC-haplotype (Ekeruche-Makinde et al., 2013). Moreover, a potential response will be affected by different affinities between peptide and MHCs (Engels et al., 2013). There are however multiple examples of more unconventional peptide presentations that will change the linearity and conformation of the peptide, which in turn will affect the overall recognition motif and the concomitant response (Beringer et al., 2015; Dendrou et al., 2018; Gras et al., 2016; Motozono et al., 2015; Pymm et al., 2017; Remesh et al., 2017; Riley et al., 2018).

HLA polymorphisms

From a population-wide perspective the variations in MHC expression amongst individuals are of uttermost importance for broad immunological protection of any species. In human, every individual express six HLA class I and HLA class II molecules respectively (two of each HLA-A, -B, -C and HLA-DR, -DQ, -DP), and because the HLA locus is the most polymorphic region of the human genome, with more than 15,000 different molecules identified (Beck and Trowsdale, 2000; Robinson et al., 2003, 2016), most individuals are heterozygous for each pair. Each individual benefits from this heterozygosity by having multiple chances of presenting an antigenic peptide, and most importantly the overall HLA variation help ensure presentation of different peptides across individuals of a population (Parham and Ohta, 1996), which, in turn, reduces the risk that a pathogen would evolve strategies to avoid overall HLA presentation, which could extinguish an entire species. Pathogen-mediated selection is proposed as a mechanisms that maintain HLA heterozygosity (Pierini and Lenz, 2018; Trowsdale and Knight, 2013) and its impact is exemplified from monitoring disease progression of Human Immunodeficiency Virus (HIV) infected individuals. In HIV patients, a maximal class I heterozygosity is associated with a delayed onset of acquired immunodeficiency syndrome (AIDS) compared to homozygous counterparts (Carrington et al., 1999). Another study has found that a maximal class I heterozygosity also corresponds with an improved overall survival in cancer patients receiving immune checkpoint therapy (Chowell et al., 2018). More evidence to the importance of HLA heterozygosity for enabling presentation of different peptides comes from studies that show how most mutations in these molecules are at locations that alter the peptide binding properties of the MHC (Kumánovics et al., 2003).

Peptides

The peptides presented in HLA-molecules at the surface of cells are derived from both the healthy host proteome and from alternative proteins introduced through evolving viral infections or random mutations from cancerous cells. Inherent to the purpose of T cell recognition, which is to detect any type of cellular danger (Matzinger, 1994, 2002), the TCR repertoire must therefore be able to interact with targets that are as randomly formed as the range of pathogenic peptides, while not responding inappropriately to self. Reflecting only on the peptide targets of CD8 T cells, a sequence of 8-14 amino acids, (Gfeller et al., 2018; Guo et al., 1992) in terms of potential diversity, the total number of different peptides that could theoretically be generated from the 20 naturally occurring amino acids would exceed 10^{18} (Sewell, 2012). Moreover, this does not account for post-translational modifications, which are believed to alter peptides sufficiently to generate an even greater number of T cell targets (Chen et al., 1999; Green et al., 2007; Hill et al., 2008; Meadows et al., 1997; Mohammed et al., 2008; Omenn et al., 2016; Petersen et al., 2009; Ramarathinam et al., 2018; Scally et al., 2013). Although the imaginable number of T cell targets are decreased to only a few percentages (1-3%) when accounting for peptide processing and the ability to bind to a given HLA (Mason, 1998; Nielsen et al., 2007; Yewdell and Bennink, 1999), it still amounts to an astounding number of potential targets, and corroborate the complicated task assigned to the TCR repertoire of covering all imaginable pathogen-derived peptides that bind to MHC, while not targeting self. A now abandoned theory proposed that the solution comprised an extremely diverse but highly selective TCR repertoire, in which each clonal T cell could only respond to one target (Jerne, 1955, 1971). However, to comply with this theory, the number of required single naïve T cells carrying a different TCR would exceed by far the number of cells of the human body (Bianconi et al., 2013). It is instead now accepted that T cells are in fact highly cross-reactive (Mason, 1998; Sewell, 2012; Wucherpfennig et al., 2007), with extrapolations from experimental data estimating that a single TCR can bind at least 10^6 different MHC-bound peptides (Wooldridge et al., 2012). Additional mechanisms have been proposed, which indicate that this number may indeed be much higher (Riley et al., 2018).

T cell cross-reactivity

Experiments providing the initial evidence of cross-reactive T cells found that CD8 T cell responses of mice that were prior vaccinated with one virus, could be boosted through stimulation with another virus (Selin et al., 1994). Thus, several studies preceded the assertion that T cells are inherently cross-reactive (Ashton-Rickardt et al., 1993; Evavold et al., 1995; Hagerty and Allen, 1995; Ignatowicz et al., 1996a, 1996b; Kaliyaperumal et al., 1996; Kersh and Allen, 1996), and that this is both necessary for TCRs to cover all imaginable pMHC epitopes and for many reasons also advantageous (Mason, 1998; Sewell, 2012). The largest hesitancy of this idea relates to the paradox of how T cell recognition appears to be so specific, if in fact a

single TCR can recognize 10^6 different targets. Both Mason, who proposed the theory, and later Sewell applied the same reasoning to explain why this is not a contradiction in relation to the complexity of our immune system as a whole: imagining a conservative number of 10^{13} possible 10-mer peptides (from Sewell, 2012), one TCR recognizing 10^6 of these would still only correspond to the recognition of 1 in 10^7 peptides. This number may in reality be different when accounting only for the fraction of peptides presented, but still the enormity of all potentially presented peptides allows a TCR to be very cross-reactive, while in reality only being able to interact specifically with a small proportion of the targets that it is exposed to (Mason, 1998; Sewell, 2012).

The immediate advantages of having a confined repertoire of highly cross-reactive TCRs is that various clonal T cells then can respond to a given pathogenic pMHC. This means that there is a greater chance that a T cell bearing a suitable TCR is able to respond appropriately in terms of spatial and temporal requirements. It moreover reduces the risk that pathogens will produce escape-variants, since a single mutation in a peptide sequence is unlikely to bypass recognition by several different TCRs that could recognize the original pMHC (Figure 2) (Chan et al., 2018; Song et al., 2017).

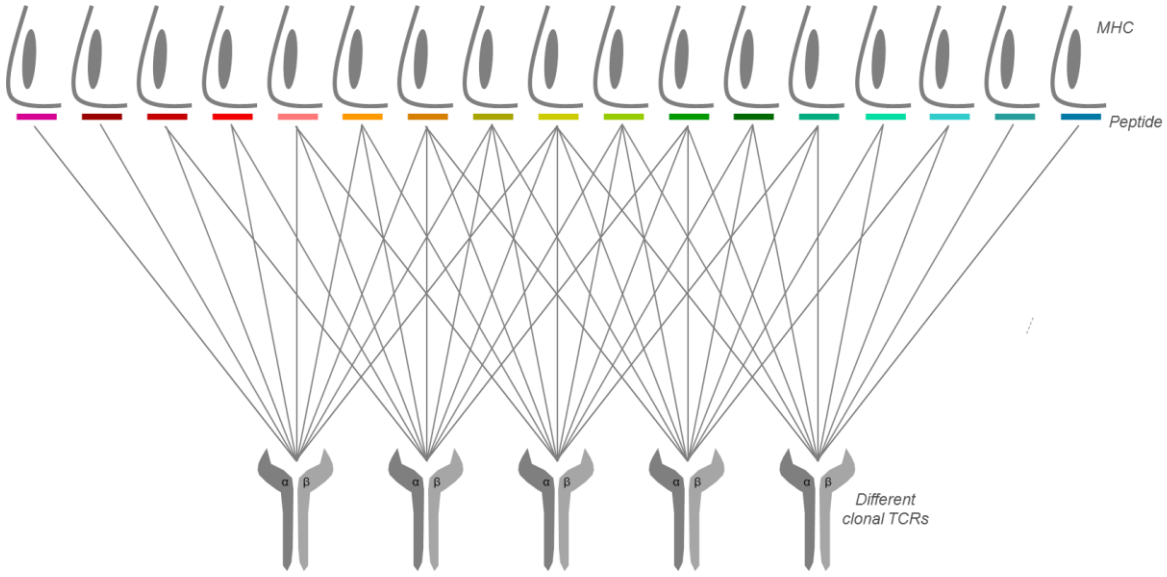


Figure 2 | **Broad antigenic coverage by a confined cross-reactive TCR repertoire.** Here five TCRs are shown to each recognize nine pMHCs providing a broader coverage than if each TCR would only recognize one epitope. Moreover, each pMHC is recognized by several TCRs which reduces the risk that a virus or malignancy will produce escape variants. Figure adapted from Mason, 1998.

In terms of the disadvantages related to having an extensively cross-reactive TCR repertoire, Mason summarized the evidence that clonal deletion of developing thymocytes seemed to have evolved to rather risk autoimmunity than leaving space for pathogens to escape discovery by the immune system (Mason, 1998; Vafiadis et al., 1996; Yu et al., 2015). Thus, both pathogenic and protective effects of cross-reactivity has been reported in a number of diseases (Benoist and Mathis, 2001; Clute et al., 2010; Partinen et al., 2014; Shann et al., 2010; Welsh et al., 2010; Wucherpfennig and Strominger, 1995). It still remains unresolved exactly how most individuals are exempt from broad immune reactions that target self. Several peripheral mechanisms have been shown to at least play a role in maintaining self-tolerance in spite of extensive cross-recognition; this includes immune suppression by regulatory T cells, or surface expressed and secreted molecules (Mueller, 2010; Nishimura et al., 1999; Ohkura et al., 2013; Tivol et al., 1995; Waterhouse et al., 1995). Moreover, low affinity of peptide-MHC complexes, and a requirement for reaching a certain threshold of pMHC-TCR interactions to activate a T cell, means that T cell recognition can only be translated into a functional response if the peptide is derived from a protein expressed above a certain level (Zhong et al., 2013). Thus, because TCR interaction with a given pMHC does not necessarily correspond to T cell reactivity, it is important to distinguish between biochemical and immunological recognition. The occurrences of autoimmune diseases are proof that the self-tolerating mechanisms frequently break, which is moreover exploited in immune therapy of cancer (Couzin-Frankel, 2013), where it can be desirable to have a TCR repertoire that is able to respond to self-derived targets or targets that resemble self. Thus, several strategies of immune therapy of cancer work to deliberately break some level of tolerance (Ribas and Wolchok, 2018; Rosenberg and Restifo, 2015). This is conceptually important for the scope of this thesis, since it supports the rationale that T cell responsiveness can be modeled within an individual, which is the great potential of immune therapeutic strategies and worth much further discussion. It however also highlights the major caveats associated with such interference; since T cell recognition is so immensely complex, and hence difficult to describe accurately, it makes it extremely challenging to predict the extent and final outcome of any immune intervention that affect T cells.

4 Detection of antigen-responsive T cells

Knowledge on the exact antigenic pMHCs that are recognized by T cells can instruct disease management and therapeutic development across immune related disorders such as infectious disease, cancer and autoimmunity. Within these disorders the advantages of knowing the landscape of immunogenic pMHCs, i.e. the epitopes that are capable of mounting a functional T cell response, are particular evident when monitoring disease progression or response to therapy, and for use in adoptive T cell therapy or vaccine developmental studies. However, direct *ex vivo* detection of the specific pMHCs that are recognized by T cells is not trivial. This is in part due to the inherent complexity of T cell recognition accounted for in the previous chapter, and moreover, because the interaction between pMHC and TCR is of very low avidity.

MHC multimers. Stable interaction that enables direct probing of T cell recognition is achieved by multimerization of single pMHC complexes on a fluorescent-labeled streptavidin molecule (Altman et al., 1996). Peptide-MHC interacting T cells are thus traced through the associated fluorescent label, which is typically visualized by flow cytometry, or less frequently by microscopy. The associated label can moreover be used to isolate the antigen-specific T cells through for example fluorescence activated cell sorting (FACS) (Andersen et al., 2013) or magnetic bead-based separation (Hombrink et al., 2011). In the first use of multimerized pMHCs Altman and colleagues applied MHC tetramers to detect CD8 T cells recognizing HLA-A0201 complexed with HIV and influenza- (FLU-) derived peptides, respectively (Altman et al., 1996). Besides such tetramers, MHC multimers variants ranging from dimers to octamers, or dextramers and dodecamers, have been applied for detection of pMHC engaging T cells (Bakker and Schumacher, 2005; Dolton et al., 2014; Huang et al., 2016). The higher order of multimerization should theoretically increase the stability of the TCR-pMHC interaction, and thus result in improved detection of antigen-specific T cells. This is valuable for discovery of rare T cell populations or when the pMHC-TCR interaction is of particularly low avidity, as is the case of many autoantigens and some cancer-associated epitopes (Dolton et al., 2018). *In vitro* production of different pMHCs is difficult because MHC molecules are unstable if not complexed with a peptide. Therefore the first applied MHC multimers were generated from separate recombinant expressions, refoldings and purifications of each new peptide complexed with MHC, and were thus not amenable for high-throughput production of individual pMHCs (Bakker and Schumacher, 2005; Garboczi et al., 1992). Several strategies have now been

developed for parallel generation of individual peptides complexed with MHC class I; these include UV-mediated exchange of conditional ligands, temperature-induced peptide exchange and stabilization of empty class I MHC molecules through gene-modifications (Bakker et al., 2008; Hadrup et al., 2009a; Luimstra et al., 2018; Rodenko et al., 2006; Saini et al., 2018). This has in turn enabled the development of technologies for high-throughput investigations of CD8 T cell recognition across various diseases.

Probing of disease-associated antigens

Despite these extensive efforts to increase the complexity of T cell detection techniques there is still a large gap between the actual number of pMHCs that can be screened simultaneously and the numbers that are required for genome-wide analyses of T cell recognition. The recent improvements in technologies for probing pMHC specificities have made it possible to study T cell recognition in a less biased and highly parallel manner. In turn, this has broadened the scope of studying T cell recognition; from only analyzing the T cell responses towards a few model antigens to now, studying broad T cell recognition of personal disease-associated antigens within patient material. The following will describe motives for probing disease-related pMHCs in infections, cancer and autoimmunity, and will moreover recount characteristics that may affect the ability to detect such antigen-specific T cells.

T cell responses in infectious disease

Broadly, virus-responsive T cells are more easily detected compared to T cells responding to cancer or autoantigens; in part because they are often present in blood of infected individuals at high frequencies, and because antigenic peptides are encoded in the viral genome they may be recognized across infected individuals expressing the same HLA-haplotype. Moreover, the TCRs targeting infected cells are often able to bind their virus-derived epitopes at relative high affinities (Hebeisen et al., 2013a, 2013b), reflecting that these epitopes, which are dissimilar to self, does not take part in the negative selection of developing thymocytes. Therefore, virus-derived responses have been particularly amenable to early T cell detection technologies, and collectively these features have made it possible to study T cell recognition in virus-infected individuals by utilizing are range of commonly observed virus-associated peptide-antigens.

In terms of T cell recognition, cytomegalovirus (CMV), Epstein-Barr virus (EBV) and FLU are amongst the most widely probed viral infections and have provided essential insight to general features of T cell immunology. CMV has particularly been investigated to shed light on the consequences of chronic viral infections (Colugnati et al., 2007; Griffiths et al., 2015; Sylwester et al., 2005), and provided an example for investigating the evolutionary paths of the TCR repertoire upon latent infection (Schober et al., 2018). Unfortunately, early high-throughput

MHC multimer technologies have been restricted in terms of studying T cell recognition of structurally related peptides (Andersen et al., 2012a; Hadrup et al., 2009b; Newell et al., 2013), and comprehensive studies of the sequence variants that arise during infections with pathogens such as HIV or hepatitis C virus have therefore been limited.

T cell responses in cancer

Detection of cancer-responsive T cells is particularly difficult. This is in part due to the fact that some cancer-associated antigens are of self-origin, which implies that the T cells that are able to recognize such antigens have escaped clonal deletion during thymocyte development, either because of low affinity of the expressed TCR with cognate pMHC or of the corresponding peptide-MHC complex. Such low affinity interactions makes it difficult to obtain sufficient avidity using pMHC reagents for binding to T cells, or complicates the process of *ex vivo* generation of stable pMHC complexes. Due to cancer immunoediting (Dunn et al., 2004a, 2004b), the T cells with potential for recognizing cancer-derived epitopes are often suppressed in the tumor environment and therefore also difficult to culture and stimulate *ex vivo*. Cancer immunoediting also result in high heterogeneity of cancer cells within a single tumor lesion and across individuals with the same cancer type (Gerlinger et al., 2012). Thus the range of potential cancer-derived epitopes is enormous, and to a large extent unique for each patient. This in turn set high requirements for technologies that allow parallel assessment of T cell recognition of many potential pMHC antigens. Besides, cancer-responsive T cells are often only present at low frequencies, which make detection of these cell populations even more challenging.

Recently, the clinical potential of immunotherapy of cancer (Couzin-Frankel, 2013) has provided strong motivation for advancing T cell detection technologies to accommodate some of these hindrances. Particularly, correlations of clinical outcome from cancer-immune-therapeutic treatments with tumor mutation burden have indicated that formation of the corresponding mutation-derived neoepitopes play an important role in the immune systems' ability to eradicate cancer (Van Allen et al., 2015; Hugo et al., 2016; Rizvi et al., 2015; Rooney et al., 2015; Snyder et al., 2014). Theoretically such neoepitopes represent optimal targets for cancer immunotherapy since they are both tumor-specific and immunologically dissimilar to self. It is therefore anticipated that some TCRs will have high affinity for neoantigen-derived pMHCs, and thereby be tailored for immune therapeutic strategies. Neoepitope-responsive T cells can be identified following a pipeline of first mapping the cancer-specific mutations within an individual using high-throughput sequencing technologies, then applying a number of *in silico* filtering steps for prediction of which of such genetic alterations are most likely to be presented at the level of the MHC (Bais et al., 2017; Bjerregaard et al., 2017a; Bulik-Sullivan et al., 2018; Hundal et al., 2016; Rubinsteyn et al., 2017; Wood et al., 2018), and lastly screening patient material *in vitro* for T cell recognition of any of the predicted neoepitopes (Figure 3). In this way examples of T cells recognizing cancer-derived neoepitopes have been detected in a number of cancer types

(Bassani-Sternberg et al., 2016; Bentzen et al., 2016; Gros et al., 2016; Gubin et al., 2014; McGranahan et al., 2016; Rajasagi et al., 2014; Snyder et al., 2014; van Rooij et al., 2013; Verdegaal et al., 2016; Wick et al., 2014). Substantial ongoing efforts in clinical translation aim at developing strategies for specific activation of neoepitope reactive T cells (Keskin et al., 2018; Ott et al., 2017; Sahin et al., 2017). However, it has still not been possible to realize the therapeutic potential of neoepitope-specific T cell recognition due to lack of methodological frameworks, that 1) fully consider the heterogeneity within tumor lesions (Jia et al., 2018), 2) support the computational predictions of the immunogenic neoepitopes (Chowell et al., 2015; Łuksza et al., 2017), and 3) sufficiently enable the *in vitro* screening platforms to accommodate the unique and variable character of neoepitope-responsive T cells within individuals.

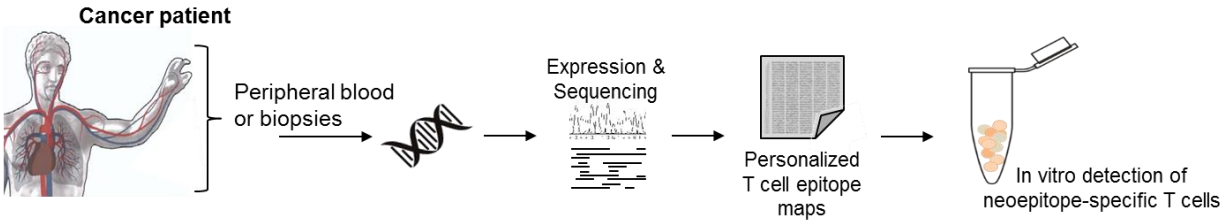


Figure 3 | **Pipeline for identification of neoepitope-responsive T cells.** Whole exome- and RNA sequencing of tumor sample and normal tissue are applied to identify mutations that will potentially give rise to neoepitopes. The potential neoepitopes can be synthesized, and true T cell epitopes can be identified through *in vitro* investigations.

An early approach has been to study T cell recognition of the antigens derived from *de novo* or overexpressed proteins that occur in cancer. While detection of this type of cancer-antigens is particular challenging due to their self-origin, they are advantageous from a therapeutic perspective, because they may be shared between individuals with the same cancer type or even across cancers. Thus, adoptive transfer of T cells specific to the cancer-testis antigen NY-ESO-1 has been reported (Rapoport et al., 2015; Robbins et al., 2011, 2015), and the melanoma-associated tumor antigens (tyrosinase, gp100 and MART-1) have been used to examine the infusion product applied in adoptive cell therapy (Andersen et al., 2012b; Kvistborg et al., 2012). These studies have revealed that those shared cancer-associated antigens that have been identified only accounts for a fraction of the total T cell reactivity encountered in the infusion products. This in turn has provided a motive for further interrogation of the cancer-responsive T cell repertoire.

Virus-associated cancer-antigens are an attractive group of targets of immune therapy; since these are virally encoded they are at once foreign to the immune system and can be recognized across individual patients, thus providing an optimal setting for development of efficient targeted treatments that can be used broadly in these virus-associated malignancies. As such, broad application of the human papilloma virus (HPV) vaccine illustrates the huge

potential of inducing prophylactic immunological protection of the viral infection that may lead to most cervical cancers (Schiller and Müller, 2015), and moreover provides motivation for also developing therapeutic vaccines that target oncoviral peptides (Çuburu and Schiller, 2016). Another oncovirus, Merkel cell polyomavirus (MCP*y*V), is associated with at least 80% of cases with Merkel cell carcinoma (MCC) (Feng et al., 2008; Rodig et al., 2012; Shuda et al., 2008). MCC patients respond well to checkpoint inhibition, but more than 40% will not experience long-term benefit (Nghiem et al., 2016); therefore therapeutic targeting of MCP*y*V-derived epitopes would seem an ideal strategy for treating these patients, for example by leveraging on adoptive transfer of T cells transduced with a TCR responsive to a MCP*y*V-derived epitopes. Moreover, MCPyV-derived epitopes provides a model system for monitoring the dynamics of tumor-specific T cell responses upon treatment with immune checkpoint inhibition (Miller et al., 2018).

T cell responses in autoimmune diseases

The risk of individual cases of autoimmunity represents a compromise of having a relatively confined TCR repertoire that must be able to recognize the vast majority of disease-associated peptide-antigens (Mason, 1998). Peripheral tolerance mechanisms will in most cases ensure that T cells expressing a TCR that recognize endogenously derived pMHCs are not activated inappropriately (Mueller, 2010). However, immune tolerance regularly fails, thereby causing autoimmune diseases. Breakage of tolerance may be caused by a failure of the peripheral mechanism, induced by e.g. TCR cross-recognition of a viral-epitope (Wucherpfennig and Strominger, 1995), overexpression of endogenous proteins (Negro et al., 2012), or defects of regulatory T cell mechanisms (Ehrenstein et al., 2004; Ferraro et al., 2011; Lindley et al., 2005; Singh et al., 2018; Viglietta et al., 2004). Moreover, post-translational modifications, such as glycosylation, phosphorylation, deamidation and citrullination, does not occur during thymic selection (Raposo et al., 2018), and can thus produce neo-epitopes (Mohammed et al., 2008; Molberg et al., 1998) that has the potential for inducing a potent T cell response (Babon et al., 2016; Koning et al., 2015; Kracht et al., 2017; Liepe et al., 2016; Mannering et al., 2005; Salou et al., 2015; Scally et al., 2013; Suárez-Fueyo et al., 2016). Because the mechanisms that lead to failure of tolerance are so unpredictable, autoimmune reactions can target virtually any peptide and occur in any tissue. This lack of defined antigens has greatly complicated the means of detecting autoreactive T cells, and direct detection using pMHC multimers is moreover often impeded since a substantial fraction of the expected autoantigens comprise low pMHC-TCR avidity due to central tolerance mechanisms (Dolton et al., 2018). Many autoimmune disorders, such as type-1-diabetes, multiple sclerosis and rheumatoid arthritis have particular genetic association with certain HLA class II alleles (Dendrou et al., 2018), however both CD4 and CD8 autoreactive T cells have been reported. For example in several recent studies that investigate the role of T cells in narcolepsy (Latorre et al., 2018; Luo et al., 2018; Pedersen et al., 2019). In

general, the exact CD4 T cell epitopes of importance in most autoimmune disease are poorly described because class II MHC multimers are notably difficult to produce, and especially in a high-throughput format (Vollers and Stern, 2008). Collectively, these obstacles have delayed studies that investigate the role of autoreactive T cells, but with the advances driven largely by the field of cancer immunotherapy, technologies are slowly also being applied in the field of autoimmune diseases.

Probing of CD4 T cell specificity. Analysis of CD4 T cell recognition as a whole has been delayed largely due to features of the MHC class II molecules that make it particularly difficult to predict and produce the individual class II molecules in complex with peptide. The practical features that complicate detection include: **1)** Inefficient refolding of MHC class II proteins with peptide ligands. Opposed to MHC class I that can be expressed in bacteria and subsequently refolded with protein (Toebes et al., 2006), the class II molecules require expression and refolding in insect (Kozono et al., 1994) or mammalian (Day et al., 2003) expression systems, which are difficult to optimize. **2)** Due to the more open structure of the binding groove of MHC class II, and hence less restricted peptide binding properties, one peptide ligand may bind to the same haplotype differently, i.e. with an altered register (Bankovich et al., 2004; Landais et al., 2009). This makes it more difficult to both predict the complex binding of peptide ligands (Andreatta et al., 2015), and in practice impedes the generation of specific pMHC complexes, since refolding must be carefully controlled, and is therefore not suitable for high throughput generation of individual pMHCs. However, in some cases peptide exchange is feasible for MHC class II molecules (shown for HLA-DR alleles and mouse I-E) (Day et al., 2003; Landais et al., 2009), which provides a first step towards more parallel production of individual pMHC complexes. **3)** A higher diversity of the MHC class II alleles in human makes it is unlikely that a few HLA-haplotypes will cover a substantial fraction of a population. It is therefore less convenient to study CD4 T cell responses than CD8 responses, where e.g. HLA-A0201 provides a model MHC class I expressed by ~50% of the Caucasian population (Ellis et al., 2000). **4)** It is commonly reported that TCRs have low affinity for MHC class II, which will impede direct detection using MHC multimers, and in some cases make such probing impossible (Lissina et al., 2009; Tan et al., 2017; Xie et al., 2012). **5)** From studies investigating antigen-specific CD4 T cells with MHC tetramers in allergy and infectious disease, it is suggested that low magnitudes of CD4 T cell frequencies after initial exposure and formation of memory cells, reflects that the CD4 T cell subsets in general target a broader range of peptide antigens (Su et al., 2013; Uchtenhagen et al., 2016). The low frequency of CD4 T cells however also complicates the task of detection antigen-specific CD4 T cells. In spite of these limitations, strategies for parallel generation of individual pMHC class II molecules are intensively sought, and upon establishment of such, most high-throughput technologies for detection of antigen-responsive T cells will be amenable for the class II molecules also (Blahnik et al., 2018).

A need for high-throughput T-cell detection technologies

Patient material is often of limited supply and since *in vitro* expansion of T cells has been shown to skew the composition of antigen-specific T cells, by providing some cells with a growth advantage compared to others (Andersen et al., 2012b), cell culturing prior to MHC multimer staining should be kept at a minimum. Moreover, antigen-specific T cell populations are in general low in frequency, which jointly makes it undesirable to split samples for interrogation of T cell recognition since this comes at a loss of sensitivity (Bentzen et al., 2016; Cossarizza et al., 2017). Therefore a crucial obstacle for probing any heterogeneous T cell sample for recognition of numerous pMHCs, relates to limitations in terms of the number of different labels that can be used in parallel. The heterogeneity of T cell recognition greatly outnumbers the available single-color labels for flow cytometry, and even though the complexity has increased substantially by using combinatorial-labeling schemes (Figure 4) (Andersen et al., 2012a; Hadrup et al., 2009b; Klinger et al., 2015; Newell et al., 2009, 2013) there is still a great gap between the number of pMHCs that can be probed simultaneously and the millions of potential immunogenic pMHCs of an individual.

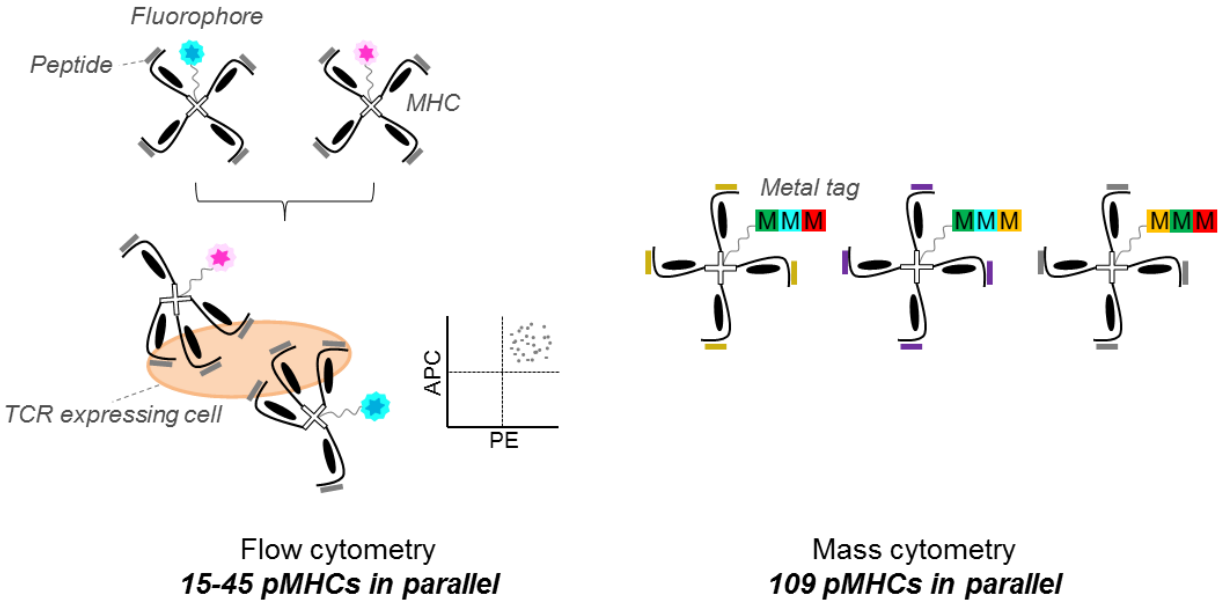


Figure 4 | **Cytometry-based combinatorial approaches.** MHC multimer-based detection of antigen-specific T cells has been expanded by use of combinatorial-labeling strategies. Hence, applying two or three-dimensional encoding schemes, fluorophores or metal tags respectively, the complexity has been increased from *n* to $\frac{n(n-1)}{2!}$ or $\frac{n(n-1)(n-2)}{3!}$. Reported complexities of each strategy are indicated in bold.

Paper I

This study provides proof-of-concept for a novel technology that enables parallel assessment of T cell recognition of more than 1000 pMHCs simultaneously. The high complexity of the assay is reached by labeling individual MHC multimers with each their unique DNA barcode that can be recovered through amplification and high-throughput sequencing. The DNA barcodes are synthetic nucleotide sequences that are generated through a combinatorial scheme; comprised of two partially complementary oligonucleotides with each their unique 25mer sequence (Xu et al., 2009). This design can produce 10^{10} unique DNA barcodes, and thus redefines the limitations of high-throughput T cell detection techniques. We anticipate that the technology can by applied for parallel assessment of T cell recognition of at least 10,000-100,000 pMHC specificities, and expect that new limitations will relate to unspecific interactions imposed by the high concentrations of pMHC applied in a combined assay, rather than the number of available labels. We show that the technology can be applied to detect T cells responsive to virus- and cancer-derived pMHCs in a highly parallel manner, and that the complexity of the assay enables interrogation of uncultured tumor material. We moreover detect several neoantigen-responsive T cells in blood and tumor infiltrating lymphocytes from non-small cell lung carcinoma patients. Since the technology requires only a few fluorescent labels it integrates the advantage of assessing other cellular functions in parallel. We specifically show the feasibility of combining the knowledge of pMHC specificity with the functional profile of both virus- and cancer-specific T cells. The technology thus provides an attractive tool for ongoing efforts to map immunogenic pMHCs across human disorders such as infections, cancers and autoimmune diseases.

The use of DNA barcodes as labels for antigen-specificity has provided a quantitative leap in terms of the number of T cell specificities that can be probed in parallel; in a first embodiment we manage to advance the previous limit of a two digit complexity to more than 1000 MHC multimers. However, a limitation of the technique, as applied in paper I, relates to the sorting of bulk populations of MHC multimer interacting T cells. This comes at a loss of single-cell resolution since the number of individual DNA barcode reads will reflect an average number of interactions within the sorted population rather than an exact number of MHC multimer-interacting cells per specificity. Thus, there are no means of differentiating between small populations of high-avidity interactions, and large populations of low-avidity interactions, which could incidentally result in the same number of DNA barcode reads for a given specificity, despite a completely different number of antigen-specific cells. To investigate the practical implications of this limitation, we estimate the frequencies of individual antigen-specific T cell

populations within a sample based on the number of DNA barcode reads associated with a given pMHC relative to the total number of reads, and relate this to the total frequency of MHC multimer binding T cells. When comparing these estimates with the frequencies obtained from fluorescent-based assessments of the same samples, we obtain correlations r^2>0.90. This in turn tells us that the relative frequencies determined within a sample are reasonably accurate and that our concern may mostly be of theoretical importance. However, the impact is less obvious when acknowledging that biochemical and immunological recognition is not the same, since differences in the avidity may reflect on the functional outcome of an interaction. Although, assuming that other MHC multimer technologies are limited in their ability to trace low-avidity interactions that would induce a T cell response (Rius et al., 2018; Tungatt et al., 2015), we could speculate that the resolution obtained from estimating the frequencies of antigen-specific T cells is within an acceptable range.

Due to the central role of T cell recognition for eliciting an immune response, several alternative approaches aiming at increasing the complexity of MHC-based assays have been proposed (Hadrup and Schumacher, 2010). However, apart from soluble MHC multimer strategies (Andersen et al., 2012a; Hadrup et al., 2009b; Klinger et al., 2015; Newell et al., 2009, 2013), several attempts at developing a reliable MHC microarray have in general produced unsatisfactory results in terms of sensitivity and reproducibility (Brooks et al., 2015; Chen et al., 2005; Deviren et al., 2007; Kwong et al., 2009; Soen et al., 2003; Stone et al., 2005). The referenced soluble and solid phase MHC-based technologies are described in more detail in paper II. Methodologies that apply such spatial separation of individual pMHCs have therefore not been implemented in high-throughput schemes.

Paper II

This review article recounts the opportunities and challenges associated with different MHC-based approaches that have arisen in the wake of the first direct detection of antigen-specific T cells (Altman et al., 1996). Emphasis lies on the ability of evolving T cell detection technologies to accommodate numerous potential pMHC-targets in a convenient and high-throughput screening format. With examples from immune therapy of cancer, it is moreover discussed how emerging highly parallel techniques have the potential for moving immune therapeutic approaches into the era of personalized medicine.

Most of the technologies recounted for direct detection of antigen-specific T cells assumes only one pMHC target per T cell specificity, which imposes a challenge since assessment of antigen-recognition may be completely impeded if a T cell in fact recognize two or more of the pMHC epitopes probed in parallel (Andersen et al., 2012a). With lower-throughput technologies this

will in most cases only represent a theoretical problem since it is unlikely that a TCR will recognize several unrelated epitopes in a pool of only ~30-100 pMHCs. When related peptide sequences are included in a T cell epitope screen, it however set high requirements for careful design of individual MHC multimer libraries. In such cases the related pMHC should be divided into separate T cell incubations to avoid cross-recognition of different MHC multimers, which would otherwise be interpreted as background. However, with the growing complexity of T cell detection technologies that enables assessment of a greater number of pMHCs in a single reaction, there is an increased requirement that such technologies can also accommodate cross-recognized pMHCs. In that context, the DNA barcode-labeled MHC multimer technology provides an advantage since T cell recognition is determined from absolute numbers of pMHC interactions, rather than the number of cells. Thus, irrespective of the number of different pMHCs that a given TCR recognize, all pMHC-TCR interactions that occur will be included in the analysis. However, as far as the experiments are conducted on bulk isolated MHC multimer-interacting T cells, the analysis provides only an average of the number of pMHC-TCR interaction within the total pool of isolated cells, and no way of discriminating whether the MHC-associated signals are derived from one or more cells (Bentzen et al., 2016).

A more recent publication describes another DNA barcode-labeled MHC tetramer approach related to the one developed in paper I. This approach employs an in vitro transcription and translation (IVTT) step that facilitate direct peptide synthesis from a nucleotide template that itself constitute the DNA barcode for that given peptide-specificity (Zhang et al., 2018b). This is advantageous since chemical peptide synthesis takes several weeks (opposed to a few days for IVTT) and represents a substantial part of the full financial costs related to the MHC multimer assay. The IVTT strategy however loses the advantage of the combinatorial DNA barcode design, and will have to purchase new barcodes for every new peptide. This in turn makes it less straightforward to account for the actual differences in costs between the IVTT strategy and the one reported in paper I. In the study by Zhang and colleagues they moreover sort the MHC multimer-binding T cells in a single-cell manner, which; 1) provides direct information of T cells interacting with several pMHCs, 2) ameliorates the limitations related to the necessity of estimating the frequencies of antigen-specific T cells, and 3) enables correlations of pMHC recognition with transcriptomic analyses.

Functional and phenotypic analysis

The functional capacity or proliferative potential of T cells upon stimulation with peptide targets can also be used to map T cell specificities (Lyons and Parish, 1994). The readout from the enzyme-linked immunospot (ELISPOT) assay is for example based on cytokine release from heterogeneous populations of cells after antigen stimulation (Sedgwick and Holt, 1983). Such assay thus gives a measure of the average responses within a sample, but does not account for

the heterogeneous effector functions of T cells, nor does it provide information about the exact immunogenic pMHCs that T cells responded to. Fluorescence-labeled antibodies used in multicolor flow cytometry are better equipped to accommodate the heterogeneity of T cell populations, since it allows simultaneous detection of multiple cellular markers at single cell resolution (Chattopadhyay et al., 2014). This in turn enables functional interrogation of T cell recognition through tracking of cytokine secreting cells (Maecker and Maino, 2004) or cells expressing surface markers that are differently regulated in response to antigen-recognition. CD39 expression has for instance been associated with chronic antigen-stimulation of CD8 T cells, and its expression can thus be applied for discriminating tumor-specific CD8 T cells from bystander CD8 T cells in tissue from cancer patients (Canale et al., 2018; Duhen et al., 2018; Simoni et al., 2018). However efficient, this still lacks direct coupling of antigen-specificity with exact effector functions. In theory staining of antigen-specific T cells with MHC multimers could be combined with functional analysis; however, in practice such dual interrogation is impeded by TCR internalization after T cell activation (Lissina et al., 2009; Valitutti et al., 1997; von Essen et al., 2004). When using DNA barcode-labeled MHC multimers, the threshold of interactions required for defining T cell recognition seems to be less strict than other MHC multimer assays, which is likely because the specificity of the T cells are assessed in a fluorescent-independent manner (Bentzen et al., 2016). Thus, when used in combination with T-cell stimulation assays it has shown to be less sensitive to the problematic feature of TCR internalization. For other MHC-based approaches, the internalization of TCRs represents an obstacle for sufficient interrogation of T cell recognition in a disease setting, since it has been reported, for example in cancer, that T cells may in fact both recognize a mutation-derived neoepitope and the corresponding wildtype pMHC (McGranahan et al., 2016, Zhang et al., 2018b). Thus, technologies that aim at facilitating high-throughput screening of the antigen-specificities of T cells should also focus on the ability of combining these with functional, phenotypic and transcriptomic analysis of the interrogated T cells.

5 Understanding T cell recognition

The ability of a single TCR to recognize an average of 1 million different pMHCs is corroborated through studies providing evidence of structural flexibility of TCR-pMHC interactions and the continuous emergence of new examples of clonal T cells recognizing peptides of varying sequence homology (Adams et al., 2016; Bentzen et al., 2018; Borbulevych et al., 2009; Cameron et al., 2013; Chen et al., 2012; Cole et al., 2017; Cornberg et al., 2010; Degauque et al., 2016; Holler et al., 2003; Wooldridge et al., 2012). These studies emphasize that although TCR cross-reactivity is essential for providing effective immunological surveillance (Mason, 1998), it is also implicated in the development of autoimmune diseases (Cole et al., 2016; Oldstone, 1998; Rist et al., 2015; Yeo et al., 2018; Zehn and Bevan, 2006), and complicates the development of new immune therapeutic treatments. Particularly the clinical potential of cancer immunotherapy acts as a strong driver for understanding the structural features that determine a given TCRs ability to recognize different pMHCs, as such knowledge would be valuable for evaluating the safety of TCRs prior to clinical development. Technologies that enable an extensive characterization of the patterns decisive for TCR engagement are emerging (Bentzen et al., 2018; Birnbaum et al., 2014; Schaubert et al., 2010; Wooldridge et al., 2010), along with the prospects of describing and understanding some of the elementary relationships between antigen-specificity and the cross-recognition potential of a TCR (Mendes et al., 2015). This will in turn provide a foundation for understanding T cell responsiveness in terms of the delicate balance between quiescence to self-antigens, and active protective immunity towards pathogens or malignancies.

TCR binding degeneracy

The dual nature of the TCR recognition motif, comprised of both peptide and MHC, represents an inherent challenge for specific recognition since only a small area of peptide is exposed for direct contact with a given TCR. Thus, much of the T cell specificity is conferred by the more invariant interaction with MHC, which also explains how TCRs can be highly promiscuous to different targets presented by the same MHC haplotype (Adams et al., 2016; Riley and Baker, 2018). It has been reported that TCRs that require a lower number of direct peptide interactions, so-called ‘hotspots’, will have a greater binding degeneracy in terms of the number of different

pMHCs that they may recognize (Adams et al., 2016; Cole et al., 2016). Such promiscuity enables T cells to respond to a large variety of target peptides, and thus favors heterologous immunity, which describes how the exposure and immunological recognition of one pathogen can result in full- or partial immunity against a second pathogen (Welsh and Selin, 2002). There are many documented examples of heterologous immunity between related pathogens with high sequence similarity, such as the immunological protection against smallpox provided through exposure to cowpox (Stewart and Devlin, 2006)), or the Bacillus Calmette–Guérin (BCG) vaccine that protects against both tuberculosis and to some degree also leprosy (Setia et al., 2006). However, TCR promiscuity also allows recognition of peptides with a minimal sequence-overlap derived from very different antigens, such as HLA-A0201 restricted TCRs simultaneously recognizing an immunodominant FLU- and an EBV-derived peptide (GILGFVFTL and GLCTLVAML respectively) (Clute et al., 2010) or an HIV-derived (SLYNTVATL) peptide (Acierno et al., 2003). The full extent and consequences of TCR binding degeneracy is not known, however both positive and negative effects of heterologous immunity have been described (Chen et al., 2012; Su et al., 2013; Welsh et al., 2010), and the accumulating evidence points to the individuals' personal TCR repertoire and immunological history being important for the final outcome of antigen exposure, which may either confer protective immunity or damaging immunopathology (Cornberg et al., 2006, 2013; Kim et al., 2005; Nie et al., 2010). Detrimental effects of TCR promiscuity may include narrowing of the TCR repertoire and subsequent generation of viral escape variants (Cornberg et al., 2006), or the induction of an unfavorably polarized response to a pathogen or vaccination, caused by prior exposure to a cross-recognized pMHC (Chen et al., 2012; Wlodarczyk et al., 2013). Moreover, each new antigen exposure seems to impact the activities as well as the frequencies and distributions of memory T cells related to previous infections (Cornberg and Wedemeyer, 2016; Selin et al., 1996, 1999; Welsh and Selin, 2002).

TCR binding degeneracy and concomitant occurrences of T cell cross-reactivity poses an inherent risk of autoimmune disease. T cells expressing a TCR with potential for responding to self-antigens will typically have survived negative selection in the thymus if; 1) the interaction between the TCR and self-derived pMHCs are of a tolerable (low) avidity (Starr et al., 2003a, 2003b), or 2) if the peptide is derived from a low-expressed self-protein (Vafiadis et al., 1997). Tolerance to the latter reflects that there seems to be a requirement for a minimum of 60-200 high avidity pMHC-TCR interactions to activate a T cell (Demotz et al., 1990; Harding and Unanue, 1990), why peptides derived from low-expression proteins will most often be ignored in the periphery. However, tolerance to such types of self-derived antigens may be broken if T cells are activated through cross-recognition of a pathogenic peptide (Figure 5). The resulting memory T cells will have the potential of being stimulated by 50-fold lower peptide concentrations, which means that they may then become responsive to some cross-recognized self-peptides that they were previously ignorant to (Curtsinger et al., 1998; Veiga-Fernandes et al., 2000), a mechanism known as molecular mimicry (Rose, 2017). Such mechanism is supported by

observations of infections triggering autoimmune reactions (Albert and Inman, 1999; Ang et al., 2004; Chakravarty et al., 2014; Oldstone, 1998; Partinen et al., 2014; Pender et al., 2017; Rose, 2001; Wucherpfennig and Strominger; Zhao et al., 1998). It is moreover conceivable that a high collective level of TCR interactions imposed by several low-level cross-recognized pMHCs may exceed an activation threshold of a given T cell, and thus result in break of tolerance.

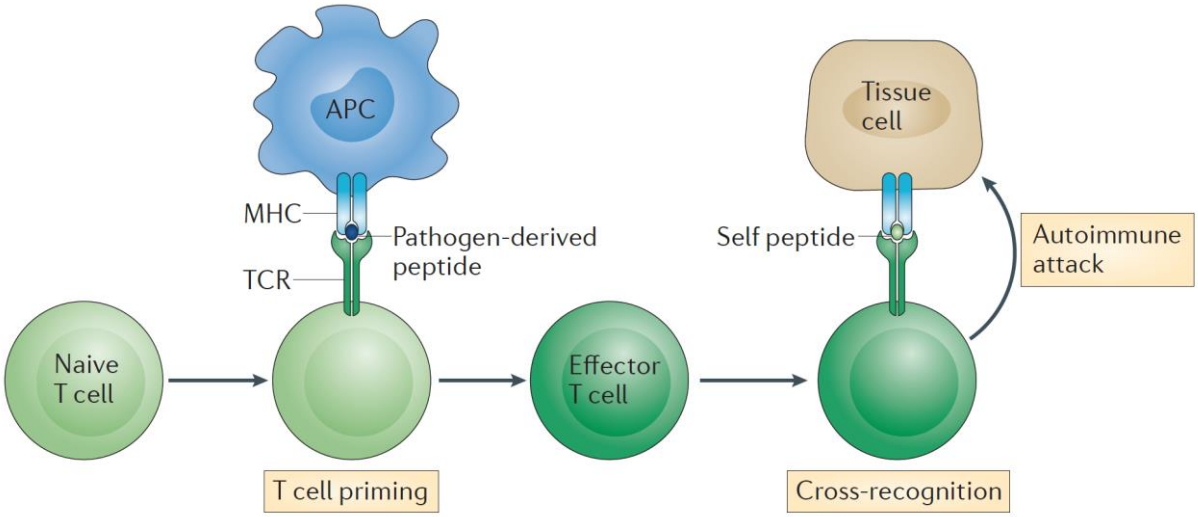


Figure 5 | **T cell cross-recognition leads to breakage of self-tolerance.** If naïve T cells are activated through recognition of a pathogenic peptide, they may become more sensitive to cross-recognized peptides derived from low-expression proteins or low avidity interactions. Self-derived peptides that were previously tolerated may thus be cross-recognized and attacked by effector T cells leading to autoimmune disease. APC, antigen presenting cell. Figure from Sewell, 2012.

Collectively this points to a need for evaluating immune responses and potential therapeutic utilizations of the immune system from a complex view that incorporates the dynamics of all the potential interactions of T cells with various pMHC epitopes derived from pathogens, malignancies and self. A first step would thus be to account for the cross-recognition potential of individual TCRs and potentially extrapolate such knowledge to help understand the extent of cross-reactivity of immune responses.

Implications for immune therapy

The advantages of utilizing heterologous immunity for therapeutic applications are clear from the success of both the smallpox- and the BCG vaccines (Setia et al., 2006; Stewart and Devlin, 2006). However, due to lack of knowledge of the exact molecular principles that determine TCR binding degeneracy, along with the far-reaching and often personal effects of T cell cross-reactivity, it is difficult to predict the final outcome of any therapy that aims at redirecting immunological recognition and responsiveness. With multiple studies pointing at pathogen-induced molecular mimicry leading to autoimmunity (Albert and Inman, 1999), it is reasonable to expect that the same detrimental mechanism may apply to vaccine-induced immunogenicity. It is for example speculated that molecular mimicry is accountable for an increase in the incidences of narcolepsy following the 2009/2010 H1N1 FLU Pandemrix (Dauvilliers et al., 2013; Heier et al., 2013; Partinen et al., 2012; Sarkanen et al., 2018). This is supported by the finding of antibodies that are cross-reactive to a peptide from the H1N1 influenza nucleoprotein A and a peptide from the hypocretin receptor 2 (Ahmed et al., 2015), a protein whose degradation is implicated in the clinical manifestation of narcolepsy (Liblau et al., 2015). Additional support relates to a seasonal onset of narcolepsy, which has been correlated with FLU season (Han et al., 2011). These occurrences may reflect that auto-reactivity can arise from molecular mimicry induced by both therapeutic interventions and pathogenic infections. Even though the exact relationship between narcolepsy and the H1N1 FLU vaccine is not explicitly known (Vassalli et al., 2015), current evidence; that autoimmune disease can be induced by molecular mimicry, is adequate to suggest that potential cross-recognition should be evaluated carefully in vaccine-developmental studies (Albert and Inman, 1999).

Adoptive T cell transfer. The renewed scientific interest in TCR binding degeneracy comes especially from cancer immunology research. In 1988 the field of adoptive T cell transfer was established with the therapeutic application of *ex vivo* expanded tumor infiltrating lymphocytes that effectively mediate tumor regression when re-infused into cancer patients (Rosenberg et al., 1988). This effect has since been widely demonstrated with impressive response rates of 20-50% (Besser et al., 2010; Ellebaek et al., 2012; Rosenberg, 2011). Despite its promise, broad application has been impeded due to tedious and unreliable production of individual T cell transfer products, intensive patient care during treatment, limited commercial potential, and a lack of tools for predicting the outcome of treatment in individual patients. This has instead driven the interest for developing T cell therapies that are amenable for a broader group of patients; the transfer of T cells transduced with a chimeric antigen receptor (CAR) (Gross et al., 1989; June et al., 2018) or a TCR that recognize known cancer-associated targets (Kessels et al., 2001; Rosenberg et al., 1990). The proposal that TCR-based gene therapy of cancer would be feasible came from early studies showing the effective redirection of T cell specificity through transduction with αβTCR genes that could provide anti-viral and anti-tumor immunity (Chamoto et al., 2004; Clay et al., 1999; Cooper et al., 2000; Dembić et al., 1986; Hughes et al., 2005;

Kessels et al., 2001; Rosenberg et al., 1990; Tahara et al., 2003). Adoptive transfer of T cells transduced with the DMF4 TCR that target an HLA-A0201 MART-1 melanoma peptide became the first applied in a clinical trial for treatment of melanoma patients (Morgan et al., 2006). In a second trial the same epitope was targeted with the DMF5 TCR that seemed to have a greater promise for treatment of melanoma but, which also caused autotoxicities (Johnson et al., 2009). *In vitro* investigations of the two TCRs showed that the DMF5 possessed a greater functional avidity and a higher affinity towards the HLA-A0201-MART-1 target than the DMF4 (Borbulevych et al., 2011; Johnson et al., 2006), and corresponding findings were obtained from a trial that applied a TCR targeting another melanoma-associated peptide (gp100) (Johnson et al., 2009). These studies supports that successful TCR-based gene therapy of cancer require transfer of T cells expressing high-affinity TCRs, which is also corroborated by early studies correlating TCR binding affinity with functional responses (Bowerman et al., 2009; Chervin et al., 2009; Irving et al., 2012; Manning and Kranz, 1999; Robbins et al., 2008; Stone and Kranz, 2013; Zhong et al., 2013). This in turn has encouraged strategies for design of high-affinity TCRs intended for adoptive cell transfer schemes (Harris et al., 2016; Ohta et al., 2018; Pierce et al., 2014; Robbins et al., 2008).

Several trials of using adoptive transfer of TCR transduced T cells have followed (Johnson and June, 2017; Johnson et al., 2009; Kageyama et al., 2015; Morgan et al., 2006; Parkhurst et al., 2011; Rapoport et al., 2015; Robbins et al., 2011, 2015) and from a general perspective, the treatment represents an extremely attractive strategy for effectively treating various cancers with an ‘off-the-shelf’ T cell product that can be used across patients with the same HLA-haplotype. However, particularly two clinical trials highlight the need to approach the use of high-affinity TCRs with extreme care. The first trial applied a high avidity TCR generated from vaccinated HLA-A0201 transgenic mice that targeted a melanoma antigen-encoding-A3 (MAGE-A3)-derived peptide bound to HLA-A0201. Of nine enrolled patients, five had objective clinical responses, however, three experienced neurological toxicity, and two eventually died (Morgan et al., 2013). The observed toxicities were subsequently assigned to the TCR also recognizing and responding to a related MAGE-A12-derived peptide, which is expressed in the human brain. This peptide in question only varied at the position 2 anchor residue, and may therefore reflect that although natural TCRs should theoretically not target healthy tissue, the strategies applied in adoptive cell therapy may override peripheral tolerance mechanisms, and thus induce autotoxicities (van den Berg et al., 2015). The second trial represents an example of molecular mimicry between a cancer antigen and an endogenous-derived peptide. This trial applied a TCR targeting a MAGE-A3-derived peptide that binds HLA-A0101 (Cameron et al., 2013). The TCR had been obtained from a vaccinated patient (Connerotte et al., 2008; Karanikas et al., 2003) and engineered to obtain a higher affinity (applying a strategy similar to Robbins et al., 2008). Only two patients were treated in this trial, however both developed carcinogenic shock and died short after T-cell infusion. Extensive investigations revealed that, despite a restricted sequence

similarity with the MAGE-A3 peptide (55% sequence overlap), the TCR had cross-reacted towards a muscle-specific protein, titin, expressed in beating cardiomyocytes.

As a natural consequence of these trials a new field concerned with rational design of TCRs intended for immune therapy has emerged (Hebeisen et al., 2015; Riley and Baker, 2018; Zoete et al., 2013). The field asks questions revolving on how changing of TCR affinity also changes T cell (poly)specificity, and how TCR responsiveness is affected by the whole pool of cross-recognized pMHCs? As a prerequisite for answering these questions there is a need for tools that enable detailed analyses of the cross-recognition potential of TCRs (Singh et al., 2017).

Dissecting the specificity of TCR interactions

The affinity-optimized MAGE-A3/titin cross-reactive TCR (a3a) represents an illustrative case of the complex task of assessing the cross-recognition potential of TCRs. Prior to the clinical trial that ended in cardiac toxicities, two affinity optimized TCRs were ‘specificity-tested’ against primary cells from a broad range of normal tissue using ELISPOT. Incidentally, this lead to exclusion of one affinity optimized TCR that targeted colorectal tissue; however the reactivity of the a3a TCR did not cause any concern and was therefore included in the clinical trial. It was only after the observed fatalities, which resulted in a T cell infiltrate of the myocardium that actively beating cardiomyocyte cell cultures were investigated more extensively and reactivity was detected, though without the immediate discovery of the cross-recognized target due to a limited sequence overlap. When the analyses finally focused in on the TCRs ability to recognize more diverse targets, with emphasis on the specific amino acid requirements at certain peptide positions, the authors identified the cross-recognized titin-derived peptide-HLA-A0101-target. The probing had required extensive analyses, applying synthetic peptides generated from sequentially substituting each position of the MAGE-A3 target peptide with alanine or glycine, and measuring the functional responsiveness of the a3a TCR separately to each peptide (Cameron et al., 2013). Ideally, their initial approach of assessing T cell reactivity towards numerous cells and tissue types would be the most optimal way of determining biologically relevant cross-reactivity, since this also accounts for antigen presentation and post-translational modifications. However, it only represents a valid approach if all tissue types within all stages of development could be tested, which is currently not feasible. The alternative of dissecting the TCR specificity more directly is currently the most attractive choice, since this help identify TCR interaction ‘hot-spots’ (Adams et al., 2016) and potential pMHC cross-reactivity despite limited sequence overlap. Such approaches aim at determining the exact amino acids and peptide positions that are required for TCR recognition, which, when correlated with knowledge of the entire human proteome and potential HLA presentation, provides a platform for predicting cross-reactivity.

One of such approaches, which have been adapted to several laboratories, utilizes large combinatorial peptide libraries (CPLs) (Bijen et al., 2018; Ekeruche-Makinde et al., 2012; Wooldridge et al., 2010, 2012). The peptides are synthesized such that one position of a 9mer peptide is fixed, while the rest is composed of a stochastic equimolar distribution of all naturally occurring amino acids, except cysteine (corresponding to 19^8 different 9mer peptides) (see Figure 6). The fixed positions are sequentially exchanged so that the complete pool of peptides contains all 20 naturally occurring amino acids at all fixed positions (i.e. $(9+19)\times19^8 = 5\times10^{11}$ individual peptides divided into 180 different peptide pools), and T cell recognition is assessed based on functional interrogation of T cells expressing a clonal TCR after individual peptide pool incubations. The absence of T cell reactivity will thus reflect the requirement for a different amino acid at the fixed position, and collectively the data can inform on those amino acids and positions that are essential for TCR interaction. The assay moreover enables interrogation of how changing variables such as presence or absence of the CD8 co-receptor will affect functionality (Wooldridge et al., 2010).

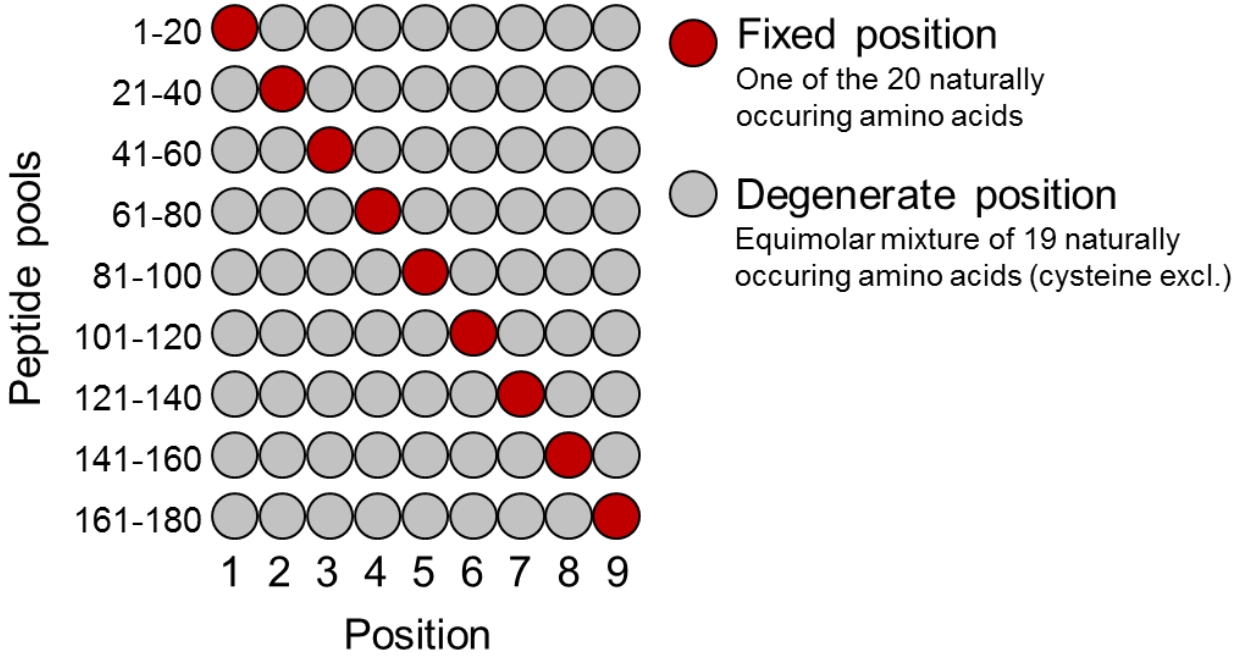


Figure 6 | **Schematic representation of a 9mer combinatorial peptide library (CPL).** The 5×10^{11} peptides of the CPL are distributed in 180 different pools. Each pool contain one specific amino acid at a fixed position (red circles) while the remaining eight positions comprise a random equimolar composition of all other naturally occurring amino acids except cysteine (grey circles). The CPL comprise a new peptide pool for each of the 20 naturally occurring amino acids at each position. Figure adapted from Wooldridge et al., 2010.

In a recent study a variant of this screening platform was applied to identify a peptide mimic that could induce an *in vivo* FLU specific T cell response in mice that were vaccinated with this mimetic, also after oral administration (Miles et al., 2018). The CPL platform was modified to comprise synthetic peptides that were composed of the corresponding D-enantiomer amino acids. CPL-based screening of a known HLA-A0201-GILGFVFTL responsive TCR revealed very different recognition profiles derived from the L- and D-amino acid-scan respectively, which indicates that the sequence of a known biological immunogenic peptide cannot be used to predict D-amino acid-based counterparts. Moreover, they found that the clonal T cells responded to an equal large number of D- and L-enantiomer peptides, corroborating the extensive cross-recognition potential of T cells. This study illustrates how knowledge on TCR binding degeneracy might foster clinical advantages within vaccine design and delivery by enabling antigen-specific recognition of synthetic molecules that are more stable *in vivo*. However, the CPL strategy in general requires large amounts of TCR-expressing cells to include in each individual pool. Moreover, because the interrogation is based on individual incubations and due to the inability to directly quantify functionality, the readout does not provide a relative hierarchy of the interactions, but rather a yes or no answer to the importance of different amino acid positions.

Another approach that provides a more accurate description of TCR binding degeneracy because it relates to direct pMHC-TCR interactions, applies a yeast display system where the pMHCs is encoded in gene constructs with a linker between the peptide and the MHC molecule (Birnbaum et al., 2014). This enables the generation of a random peptide sequence through use of mutagenic primers that produce degenerate codons allowing all 20 naturally occurring amino acids. Peptide-MHC displaying yeast cells are probed with tetramerized TCRs of interest, and interacting cells are isolated through rounds of magnetic bead enrichment and FACS. Subsequent sequencing informs about the peptide sequence expressed by the enriched yeast-displaying cells, and thus reveals the TCRs amino acid requirements for interaction with pMHC. The great potential of this strategy relates to the ability of performing largely unbiased screens, and is illustrated from the identification of pMHC targets of TCRs with otherwise unknown specificities (Gee et al., 2018a). Moreover, a recent study interrogates the a3a TCR using the methodology on an HLA-0101 yeast display library. From the TCR recognition pattern it is evident that the amino acids on positions 3, 4 and 9 comprise the main interaction points with the receptor. This can thus help explain the observed cross-recognition between the MAGE-A3- and titin-derived peptide, since these positions (and more) are all shared between the peptides (Gee et al., 2018b). Disadvantages of the technique relates to the inability to characterize all TCRs, and the display of peptides, which does not represent all possible positions equally well. Moreover, the system has so far only been developed for a few MHC molecules, and is restricted to specialized laboratories.

Paper III

This study provides proof-of-concept of a novel "one-pot" tool for determining the amino acids requirements at specific peptide positions for TCR recognition of pMHC. It leverage on DNA barcode-labeled MHC multimers, which allows determination of the relative affinity of a clonal TCR to libraries of MHCs that carries peptides with substantial sequence overlap. Peptide libraries are constructed from the originally identified targets, and include all variants generated from sequentially substituting each position with the 20 naturally occurring amino acids. This generates libraries of (n×19)+1 peptides where n denominates the number of amino acid residues of the peptide sequence (and includes the originally identified target). Individual DNA barcode-labeled MHC multimers are generated from each peptide, and the multimers are pooled and incubated with clonal TCRs. Sorting of all MHC multimer-binding T cells and subsequent sequencing reveals a distribution of DNA barcode reads that reflects the pMHC-TCR binding-hierarchy. The experimental design, which include all amino acids at all peptide positions, enables translation of the binding hierarchy into a positional scoring matrix for each TCR. This in turn is used to determine the relative amino acid preferences of the TCR at individual positions of the peptide sequence, which we call the TCR recognition pattern or fingerprint. In this study we show the feasibility of this strategy by investigating 16 different TCRs. Importantly, we assess the recognition pattern of 12 TCRs all recognizing the same MCC-derived epitope, and apply this knowledge to predict the peptides from the human proteome that are most likely to be cross-recognized by any of these TCRs. By assessing the TCRs' ability to recognize and respond to these peptides we moreover identify cross-recognized peptides derived from proteins expressed in healthy human tissue, and suggest that this could be used as a tool prior to selecting TCRs intended for adoptive cell therapy. Additionally, we contemplate that the binding degeneracy is linked to the functional avidity of TCRs, and that the recognition patterns thus can indicate the potency of a TCR.

An immediate disadvantage of the strategy presented in paper III relates to the number of peptides included in a parallel screen. The strategy of substituting only one amino acid position at a time underscores the possibility that several simultaneous amino acid substitutions might promote TCR interaction, or that a TCR may even be able to recognize a largely different peptide (Miles et al., 2018; Riley et al., 2018). Thus an optimal assay would include many more peptide variants to reflect the great variety of pMHC targets that a TCR may potentially meet *in vivo*. However, the strategy requires a new peptide library for every new TCR that is interrogated, and because peptide synthesis comprise a considerable financial cost of the total MHC multimer analysis this is not feasible. To accommodate this drawback we interrogated one

of the TCRs using a larger peptide library that included peptides with two simultaneous amino acid substitutions, and saw that the results were comparable with the results obtained from only substituting one position at a time. This result does not imply that simultaneous substitutions will never generate new possibilities for a TCR to interact, but indicates that, for peptides with some sequence overlap, single-substitutions can describe most of the amino acids required for TCR interaction. A related drawback relates to the necessity of having a pre-established peptide target of the interrogated TCR to construct the substitution library from. Thus, in the current embodiment of the strategy it would not be possible to identify new pMHC targets of TCRs with unknown specificities. Advantages includes flexibility in terms of the possibility of interrogating TCR recognition restricted to all foldable MHC molecules, and a relatively straightforward methodology which should make it possible to implement the technology in most laboratories familiar with molecular biology and MHC multimers.

Even though all three strategies; combinatorial peptide libraries, yeast display, and DNA barcode-labeled MHC multimers, represents a leap compared to previous methods that applied single position alanine substitutions, they are still experimentally limited compared to the number of possible ligands that may be encountered. Therefore, interpretations of the results should reflect that it is a merely a window of estimated cross-reactivities, not an inexhaustible assessment. Structural information and *in silico* modeling (Fodor et al., 2018; Reinherz et al., 1999) may in turn enhance the utility of the experimental approaches for assessing TCR binding degeneracy.

6 Molecular encoding

This thesis builds upon the well-established field of tracing protein interactions through fluorescent-based staining and flow cytometry (Fulwyler, 1965; Perfetto et al., 2004) that, when combined with the advantages of the more recently established field of molecular encoding, opens new possibilities for developing high-throughput technologies for investigation of the heterogeneity of T cell recognition.

Molecular encoding applies unique nucleotide sequences, termed barcodes, to label individual molecules, which allow the identification of interaction partners, or temporal and spatial tracking, depending on the application. Accordingly, the use of unique nucleotide sequences to label molecules or structures was borne out of the effort to screen large displays of chemical drug candidates in parallel (Brenner and Lerner, 1992), and it has since also been applied widely as cellular barcodes in linage tracing studies (Naik et al., 2013; Zhang et al., 2018a). Nucleotide sequences provide an advantage over other chemical structures since they are readily amplified, which enable facile recovery from heterogeneous mixtures. This in turn provides a seemingly endless range of unique labels, theoretically only restricted by the length of the nucleotide sequence and the technology applied to extract the barcode information.

Sequencing. Improvements of the variety of sequencing platforms and lowering of sequencing-costs has driven advances of the application of molecular barcodes, which are continuously seeing broader and more comprehensive uses (Favalli et al., 2018; Kebschull and Zador, 2018; Mimitou et al., 2018; Stoeckius et al., 2017). Thus, from initial extraction of barcode information through amplification and separation according to size (Cochet et al., 1992), or applying microarray detection (Melkko et al., 2007; Scheuermann et al., 2008) or Sanger sequencing (Needels et al., 1993; Nielsen et al., 1993); high-throughput sequencing is now a preferred read-out from molecular encoded screenings (Mannocci et al., 2008). High-throughput sequencing approaches rely on the construction of libraries of genomic material (typically genomic- or complementary DNA), flanked by synthetic DNA adaptor sequences. The adaptors are introduced through ligation, tagmentation (Picelli et al., 2014), or may be embedded in amplification primers. They have a dual role of immobilizing the DNA onto a solid surface, and provide the primer template for initiating a monitored amplification that will convey the sequence of the DNA. The amplification of immobilized DNA produces clusters of clonal sequences that collectively augment the signal from a single originating template to reach above

the detection limit. For each sequencing platform, universal adaptors enable the process to be carried out in a massively parallel manner, where the sequence of each individual DNA stretch is read simultaneously through fluorescence emission (Illumina) or changes in pH (Ion Torrent) (Mardis, 2013).

High-throughput sequencing is a dynamic tool where platforms are continuously being optimized and new techniques are emerging. This provides still more options for increased quality, longer reads, and a greater overall throughput-per-run. Moreover, massively parallel sequencing can now be performed within hours, and the associated costs are continuously dropping. To get full advantage of this progress, and to reduce costs, it becomes attractive to combine multiple samples in one parallel sequencing reaction. Sample multiplexing is managed by encoding all DNA derived from the same sample with a sample-specific DNA barcode (sample ID). The sample ID is introduced prior to mixing samples, commonly along with the sequencing adaptor, and will be read in the sequencing reaction. It can thus be used for subsequent demultiplexing of reads derived from individual samples.

DNA-encoded library screenings

Synthetic DNA barcodes have been extensively applied in the search for new small molecule pharmaceuticals, where they have facilitated in vitro affinity screenings of large arrays of chemical compounds. The technique has reached an impressive complexity (Mannocci et al., 2008) making parallel screenings of up to 1 billion molecules a reality (Nuevolution, 2015). Moreover, display of different peptides on the surface of phages or bacterial cells have also been applied in parallel screenings of peptide ligands, and hence also represent a DNA-encoded screening, since interactions will be revealed from the sequence of the DNA-encoded peptide after target incubation, isolation and sequencing (Birnbaum et al., 2014; Liu et al., 2017a; Scott and Smith, 1990).

The most elementary components of a synthetic DNA barcode comprise a defined nucleotide sequence flanked by a forward and a reverse primer region. Optimally, a collection of DNA barcodes used in a parallel screen should have shared amplification properties, i.e. common length, melting temperature and identical primer regions, while carrying a unique sequence identity (Xu et al., 2009). They should moreover comprise a means of attachment to the probed molecule or compound.

DNA barcode design. The DNA barcodes applied in our studies are generated from a combinatorial construction scheme comprising two partially complementary nucleotide sequences; A and B (Figure 7A). This design is adapted from molecular encoding in drug discovery, which are amongst the fields that required the greatest possible complexity of labels from an early time point (Liu et al., 2017b; Mannocci et al., 2008). As such the combinatorial

approach provides a measure for reducing the purchasing costs by reducing the total number of individually synthesized oligonucleotides, while maintaining a high complexity of the final library of DNA barcodes. Thus, we are currently able to produce close to 5000 unique AxBy DNA barcodes from combining a confined number of 30 and 195 individually synthesized A and B oligonucleotides, respectively. Moreover, by combining the "old" set of A oligos with a new set of B oligos containing a different primer region, the combinatorial approach provides a flexible platform for producing new DNA barcodes that will not be amplified in parallel with previous libraries. This adaptability contributes a tool for elimination of carryover cross-contamination from previous amplifications.

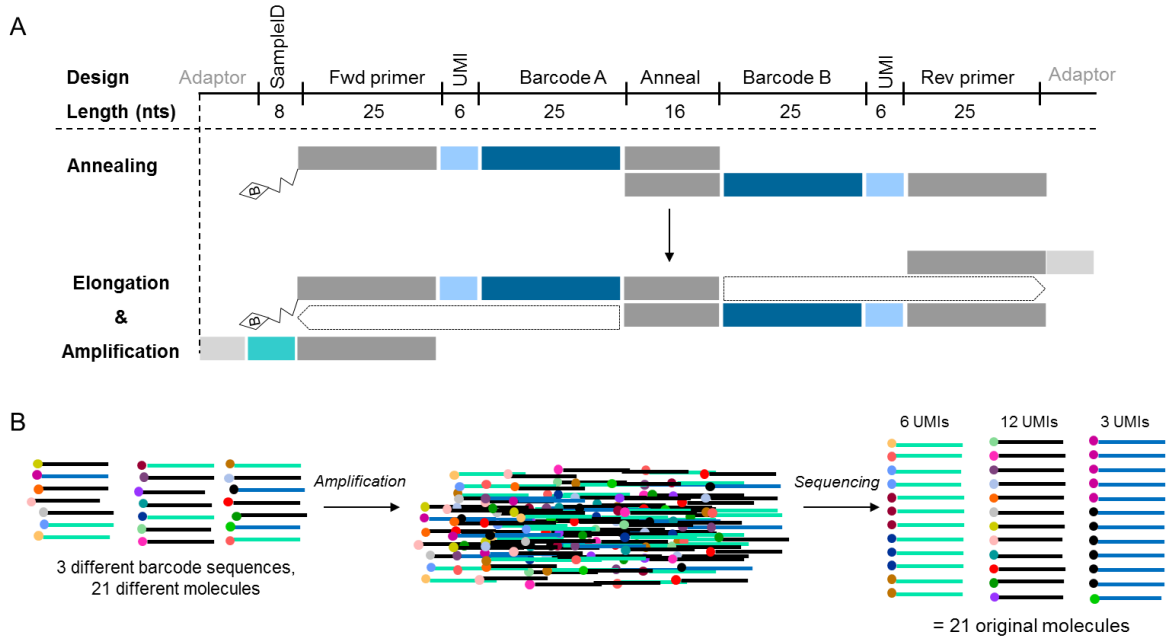


Figure 7 | **Combinatorial DNA barcodes with UMIs for accurate quantification of original molecules.** (**A**) The combinatorial DNA barcode design that provides a high complexity for labeling MHC multimers. Oligos are annealed and elongated prior to labeling MHC multimers. In a final experimental step a sample ID sequence is incorporated via amplification primers to allow multiplexing of several samples. Grey scaling denotes the parts that are shared between oligos, while the colored denotes unique sequences. Oligo A carries a biotin (B) for attachment of the AxBy DNA barcode to MHC multimers (**B**) Schematic of three different DNA barcodes (colored lines), all synthesized with individual UMIs (colored dots). Amplification may lead to a bias in the number of individual barcodes. However, after sequencing, the original number of individual molecules can be accounted for by only counting the same DNA barcode-UMI combination once. Figure B adapted from Kivioja et al., 2011.

To enable accurate quantification of DNA most modern molecular encoding schemes utilize short stretches of random nucleotides, denoted unique molecular identifiers (UMIs) (Kivioja et al., 2011), which are incorporated either directly into the synthetic nucleotide sequence or

integrated prior to the first amplification, e.g. within the first set of primers or on a template switch adaptor. The random nucleotides are incorporated individually in the synthesis of each new molecule, resulting in different UMIs across oligonucleotides with the same sequence identity. This in turn provides a tool to correct for potential amplification biases, and is applied after sequencing. Identical DNA barcodes with a common UMI will denote the clonal reads that originated from the same molecule. By only counting these once, it will provide an absolute count for the number of molecules present before amplification (see Figure 7B). We use the term clonality reduction to describe this process. The UMIs moreover provide a tool for tracking potential DNA barcode contaminations over time, which may be introduced from amplifications in previous experiments.

Paper IV

This manuscript comprise a detailed protocol for generating and applying DNA barcode-labeled MHC multimers to detect antigen-responsive T cells or to determine the recognition pattern of individual TCRs, as described in paper I and III respectively. It moreover walks the reader through applications of the method, data processessing, expected outcomes, and the advantages and disadvantages compared to other technologies.

Analyses. A few conceptual features related to the analysis of sequencing data derived from our DNA barcode design are worth more attention. Firstly, since we apply synthetic DNA barcodes, the sequences are pre-established. They would therefore map perfectly to a barcode if not for sequencing or amplification errors. Secondly, the DNA barcodes comprise different parts that are not equally important when aligning the sequences, i.e. the common primer and annealing regions are less important than the unique A and B sequences that encode a given pMHC. Considering this, sequencing errors that lie within the common parts may still enable alignment to a correct unique DNA sequence, whereas to many errors within the DNA-encoding part would result in a sequence that cannot be confidently assigned to a given pMHC. Therefore our analyses program, Barracoda (http://www.cbs.dtu.dk/services/barracoda/) (Bentzen et al., 2016), has been developed to align the different parts of DNA barcode sequence independently. First the common parts (forward primer, annealing region and reverse primer) are aligned, and only those reads that map to at least two out of three of these regions are aligned to the database of unique A and B sequences. Once the individual AxBy DNA barcodes have been identified within a read, the individual sample identification sequences and UMIs are also determined. This strategy provides the advantage that, because the sequences are synthetic and of a known structural design, the location of the common parts are used as a guide for the location of the variable parts, and can therefore yield a greater amount of useful reads from each sequencing reaction.

TCR repertoire tracking

The diverse TCR repertoire, defined from the somatically rearranged genes encoding individual α and β chains, provide an option of tracking T cells of the same clonal origin through molecular encoding. The individual TCR transcripts form an endogenous cellular barcode that is inherited by the T cell progeny and can be recovered by the same means as synthetically derived DNA barcodes. This has been utilized for providing quantitative insight into TCR diversity (Robins et al., 2009), to trace lineage decisions of T cells (Gerlach et al., 2013) and to monitor the dynamics of T cells across immune-related disorders, such as infectious disease (Dziubianau et al., 2013; Hou et al., 2016), cancer (Kirsch et al., 2015; Sherwood et al., 2013; Zhang et al., 2018a) and autoimmunity (Acha-Orbea et al., 1988; Madi et al., 2014). Most of such TCR repertoire studies have been confined to bulk experiments, where tracing the TCR β chain has been the most attractive option because it is the more diverse of the two chains and also less ambiguous due to allelic exclusion (Bergman, 1999). However, accurate pairing of the variable TCR α and β regions is valuable for uncovering the biological function of a T cell, and is generally lost in bulk experiments since the transcripts are separately encoded. Assuming that naturally pairing chains derived from the same cell will be found in the same relative frequency, bioinformatics approaches has been developed that pair α and β sequences from the relative number of transcriptomic reads (Greiff et al., 2015; Linnemann et al., 2013). This straightforward approach mainly enables pairing of chains from clones present at high-frequencies, but has been advanced by subdividing T cells into multiple aliquots prior to sequencing (Howie et al., 2015). Treating the subsequent analyses from a combinatorial view, this methodology enables identification of α and β sequences that only occur in combination in the same aliquots, and thus allow pairing of TCR chains derived from less frequent clones.

Moreover, single T cell sorting has been applied widely for capture of paired TCR α and β sequences (Dash et al., 2015; Kim et al., 2012). These techniques have utilized a range of approaches to ensure targeted amplification of each chain and efficient capture of the 5'-encoded CDR3 region. An advanced multiplex PCR approach reported by Han and colleagues (Han et al., 2014) combines single-cell capture of individual αβ TCR transcripts with a selection of mRNA transcripts, encoding functional and phenotypic markers. The individual transcripts are encoded with a synthetic cell-specific DNA barcode in a final amplification, and can therefore be mixed in a single high-throughput sequencing reaction, and subsequently be assigned to its cell-of-origin. Moreover, in the study by Zhang and colleagues discussed previously, single-cell TCR transcripts are captured in parallel with DNA barcodes encoding the T cell specificity, and thus provide a direct link between pMHC recognition and TCR sequence (Zhang et al., 2018b). These single-cell technologies have provided high-quality paired TCR sequences, but are limited in throughput due to manual handling and high costs associated with individual reactions for every single cell.

Emerging single-cell technologies

With contributions from molecular encoding, the development of highly parallel single-cell technologies has accelerated in recent years. In 2015 two independent studies were published (Klein et al., 2015; Macosko et al., 2015) that described related transcriptomic approaches for capturing single-cells in individual droplets. In these approaches each droplet carries uniquely encoded poly-dT primers that ensure that all captured mRNA from one cell is labeled with the same synthetic DNA barcode. Hence, after the initial annealing of primers (and for the Klein approach also reverse transcription), the droplets can be broken, and all subsequent processing is performed in one bulk reaction. This enables capture of mRNA from many more cells in parallel than what is possible with manual single-cell handling. The droplet-specific DNA barcodes are revealed after high-throughput sequencing, and enable mapping of each mRNA to its cell-of-origin. To generate a large library of diverse DNA barcodes, Macosko and colleagues applied a split-and-pool synthesis, which is a concept used in combinatorial chemistry (Mannocci et al., 2008).

The two strategies provided a turning point for single-cell transcriptomic analyses, and multiple techniques are now emerging that utilize highly parallel DNA encoding of mRNA, either in droplets (Klein et al., 2015; Macosko et al., 2015; Zheng et al., 2017), micro-wells (Gierahn et al., 2017; Han et al., 2018) or within cell nuclei using *in situ* split-and-pool approaches (Cao et al., 2017, 2018; Rosenberg et al., 2017), to allow subsequent bulk processing. The technologies have been applied mostly within neuroscience and immunology, where there is an established need for describing heterogeneous populations of cells at a single-cell resolution. Although these are powerful technique for capturing numerous single cells in parallel, a general limitation for many of these strategies relates to a poor capture rate, i.e. a low percentage of the cells that are introduced to the system will be captured, reverse transcribed and sequenced. This is problematic when working with low frequent cell populations especially in patient material. Additionally, these protocols do not provide full-length coverage of transcripts and has a lower sensitivity than the manual techniques. Currently most of the platforms are not flexible in terms of variations of the obtained sequences, and provides only the option of polyA-capture, which means that all information derived from the 5’ end of transcripts is lost in fragmentation steps prior to sequencing, including the CDR3 region of TCR transcripts. 10x Genomics has been amongst the first to provide a commercial platform with the option of encoding nucleotide sequences with cell-specific DNA barcodes at the 5’end, thus amending this particular limitation. They moreover have a reasonable capture rate of 40-60%.

Several techniques have been developed that utilizes the power of capturing single-cell mRNA in a bulk processing scheme. Thus DNA barcode-labeled antibodies that bind cell-expressed proteins are being used more commonly in combination with the DNA-encoded single-cell techniques (Mimitou et al., 2018; Peterson et al., 2017; Shahi et al., 2017; Stoeckius et al., 2017). In these approaches the antibody-derived DNA barcode-labels are captured in

parallel with all mRNA from single-cells, and are thus encoded with the same synthetic cell-barcode. Hence, DNA labels and mRNA can be mapped simultaneously to their cell-of-origin, providing correlates between protein expression and transcriptomic information. Moreover, in comparison with fluorescent-labeled antibodies, the molecular labeling of antibodies allows a greater number of cell markers to be probed in parallel.

Paper V

This manuscript comprises data from a proof-of-concept study for integrating knowledge of pMHC specificity with high-throughput retrieval of TCR sequences from single-cells. We apply a large library of >100 DNA barcode-labeled MHC multimers to stain antigen-specific T cells, and isolate all MHC multimer binding T cells by FACS. A droplet-based single-cell methodology is applied on the isolated cells, which enables parallel capture of αβTCR transcripts and the MHC-associated DNA barcodes. This strategy required adjustment of the DNA barcode design used previously and minor modifications of existing droplet-based protocols; both to enable capture of the MHC-associated DNA barcodes along with the TCR transcripts. The data shown is from a pilot study, whose purpose was to ensure that the new DNA barcode design was amenable for capture in the single-cell platform, and that our modification did not impede recovery of TCR sequences. Accordingly, we could extract TCR sequences from 5,686 T cells (~40 % of injected cells), and for a small fraction of these (128 cells) we could extract paired information of pMHC specificity and TCR sequence, which matched the expected characteristics from healthy donor blood or a MCP*y*V-responsive T cell clone. This served to verify that our chosen strategy is feasible; however additional improvements of the protocol are needed to reach a more efficient capture of DNA barcodes.

7 Epilogue

The research projects presented in this thesis are all concerned with developing technologies that enable us in describing T cell recognition more accurately. Through interrogation of T cell recognition within individuals and across different immune-related diseases, this has the potential for providing more comprehensive knowledge of how antigen-responsive T cells influence disease etiology, affects the response to treatment, or how it may be harnessed specifically (and individually) for disease interventions.

Identification of immunogenic targets. Existing technologies for probing antigen-specificity are insufficient in terms of matching the diversity of T cell recognition, and hence limits our ability to map individual T cell epitopes across diseases. Knowledge of the disease-associated targets that induce an immune response are valuable for enhancing our ability to control viral pathogenesis, tumor malignancies and auto-reactive cells, but is largely confined to model antigens within these diseases. Efficient *in silico* tools for predicting the peptides that will be presented by MHC are valuable for narrowing down the number of potential immunogenic epitopes associated with a given disease (Jurtz et al., 2017; van Rooij et al., 2013), but only describes a first prerequisite of an immunogenic pMHC, namely peptide presentation. Thus, there is currently a great gap between the hundreds or thousands of potential immunogenic peptides that may be predicted to bind MHC, and those relatively few that are most likely to induce a potent immune response (Bjerregaard et al., 2017b). Hence, to broaden our current understanding of immunogenicity, technologies are needed that enable unbiased analysis of the range of immunogenic pMHCs in any given individual. The technology developed in **paper I** has provided a leap in the number of T cell specificities that can be probed in parallel. In current implementations of probing ~1,000 pMHC epitopes it is however still restricted in terms of matching the diversity of T cell recognition, but does provide a first step towards broadly identifying immunogenic pMHCs. This may contribute to novel understanding of features that are important for inducing an effective T cell response. Such knowledge may in turn feed back into prediction algorithms (Bjerregaard et al., 2017a; Łuksza et al., 2017), and for example assist ongoing efforts to accurately identify the patient-specific neoepitopes that are most likely to induce a potent anti-tumor immune response (Chowell et al., 2015; Strønen et al., 2016; Yadav et al., 2014). This is particularly valuable for the development of personalized neoepitope vaccines (Carreno et al., 2015; Keskin et al., 2018; Ott et al., 2017; Sahin et al., 2017), which are

restricted to include only a few different epitopes and therefore require precise identification of the pMHCs that are most likely to be immunogenic.

The readout of the DNA barcode-based MHC multimer technology is conceptually different from fluorescent-based analyses of antigen-specific T cells. This relates mostly to the difference between assessing the total number of pMHC-TCR interactions, and assessing the total number of pMHC interacting cells, as done in the barcode- and fluorescent-based approaches respectively. Each type of readout provides a number of advantages and limitations as touched upon in chapter 4. The greatest limitation of the bulk-derived barcode-based readout is the lack of direct assessment of the frequencies of each antigen-specificity, and the lack of visual conformation of a response. The latter set high requirements for data analyses and demands effort and experience from users to confidently interpret the results. Yet, it seems that the independence of a fluorescent signal offers the advantage of detection of more low-avidity interacting T cells, since we can also recover DNA barcodes associated with MHC multimers bound in such low numbers, on a per cell basis, that individual T cells would not be distinguishable by flow cytometry alone. This is valuable for detection of self-reactive T cells of interest in cancer or autoimmune diseases, and particularly important with the notion that many responses are not detected with fluorescent-based MHC multimers (Rius et al., 2018).

An important characteristic of T cell detection technologies is the ability to detect low frequent antigen-specific responses, which are common especially among cancer- and autoreactive T cells, and in particular for CD4 T cells. Thus, several approaches that could otherwise enable probing of an immense number of T cell specificities in parallel, has not seen their potential due to low sensitivity for detecting disease-associated T cell responses (**paper II**). The DNA barcode-based MHC multimer technology enables detection of antigen-specific T cell responses down to 0.005% of CD8 T cells when staining a total of two million cells. However, we argue that one of the advantages of the technology relates to the sensitivity depending on absolute rather than relative numbers of cells, and have therefore defined a minimum threshold of approximately 20 cells to allow detection of a given antigen-specificity. This however, may depend on the affinity of the interaction and the expression-level of the TCR. By increasing the total number of cells included in the initial staining reaction, the number of T cells recognizing any pMHC will increase correspondingly and thus lead to a greater overall sensitivity. The usage of a greater number of cells is reasonable to propose, since the technology provides the opportunity of screening many more pMHCs in parallel than what was previously possible, and it therefore largely eliminate the need of splitting samples into several incubations. This is also valuable for probing material of limited access, and enables screening of unmanipulated material that will more accurately reflect the *in vivo* composition of T cell specificities (Andersen et al., 2012b). By combining with functional and phenotypic analyses of the antigen-specific T cells, this tool can thus provide insight to the breadth and the dynamics of T cell responses, across individuals and diseases.

To truly account for the dynamics of T cell responses, there is a need to also look into the naïve repertoire of antigen-specific T cells, since this could provide more extensive knowledge of immunogenic pMHCs, and the levels or states of T cells needed to mount an antigen-specific immune response. Therapeutic neoepitope vaccine strategies are particularly occupied with precursor levels of cancer-specific T cells, since knowledge of these could help point to the patient-specific peptides most likely of inducing immunogenicity. However, probing of the naïve repertoire set even greater requirements for the sensitivity and the extent of T cell interrogations, and such comprehensive analyses are currently not feasible at the level of the individual patient. We suggest that DNA barcode-labeled MHC multimers could provide an improved strategy for analyzing the naïve repertoire, with the specific advantages of a greater sensitivity (if enough cells are probed), and the option of screening for many specificities in parallel. This would require a pre-enrichment step of the T cells expressing naïve markers and will moreover require practical validation. If the expectations are met, it may assist ongoing efforts to interrogate the naïve repertoire of individuals (Egorov et al., 2018). There is evidence that tumor immunogenicity may be underscored from looking only into tumor infiltrating lymphocytes of patients, since a greater proportion of T cells from blood of multiple healthy donors have been shown to directly recognize and respond to patient-specific neoepitopes (Strønen et al., 2016). Such investigations of healthy donor blood may thus provide a basis for identifying immunogenic targets for neoepitope vaccines, or for isolating donor-derived cancer-specific TCRs that could be used in transgenic adoptive T cell therapy. However, like any other cancer vaccine or transgenic TCR therapy, this comes at risk of inducing auto-reactivity in patients by breaking endogenous tolerance mechanisms through introduction of potent T cells that respond to a target very similar to self. To accommodate this, Zhang and colleagues (Zhang et al., 2018b) reported a straightforward approach to distinguish between the donor-derived TCRs that recognize only neopeptides, and those that also cross-react with the wildtype peptides. Applying DNA barcode-labeled MHC multimers they used one fluorescent label on all MHCs holding a neopeptide and another fluorescent label on the corresponding wildtype peptide. T cells that cross-recognized both types of peptides would thus stain in two colors. By single-cell sorting and capturing the TCR transcripts associated with the single-color neoepitope-responsive T cells, the TCRs with a lower risk of cross-recognizing wildtype peptides could be isolated, while the simultaneously captured DNA barcode would inform on the exact antigen-specificities. A completely analogous approach could be implemented using the DNA barcode-labeled MHC multimer technology developed during this PhD. It would moreover be interesting to corroborate the findings with a functional, phenotypic and transcriptomic readout of the individual T cells. The TCRs identified with such methodology would still require comprehensive analysis of the complete cross-recognition potential before seeing therapeutic uses, but the strategy provides a clever first selection criterion of neoepitope-responsive TCRs intended for the clinic.

Predicting TCR cross-reactivity. Turning to analyses of potential TCR binding degeneracy and the associated cross-recognition potential of T cells, this thesis has provided a tool for systematic and detailed investigation of TCR recognition patterns (**paper III**). By identifying the essential interaction-points of a TCR with pMHC, this knowledge can be applied to predict potential cross-reactivity, which can inform on some general characteristics of polyclonal immunity. This will be particular valuable for understanding the development, and maybe even the triggering events leading to autoimmunity, which could most likely foster new therapeutic interventions. It would however be exceptionally challenging to account for the whole available TCR repertoire and the dynamic influences of existing T cell responses in each individual. Thus, in the context of harnessing endogenous immunity, it is very difficult to predict final outcomes of potential cross-reactivity. Contrarily, immunotherapeutic strategies of transgenic TCR transfer provides a more narrow scope of the cross-reactivity analyses. These strategies provide a window for controlling the dominating T cell population that expresses a known TCR. Thus, even though it is still an enormous challenge, zooming in on only one or a few TCRs will enable a thorough characterization of the molecular patterns or 'hot-spots' that will be decisive for any potential off-target cross-recognition, which in turn will be valuable for evaluating TCRs prior to clinical development. For clinical translation of the findings described in **paper III,** all the interrogated MCP*y*V-responsive TCRs could have potential in transgenic T cell strategies. The data provided give an in-depth understanding of the recognition profile of specific oncovirus-targeting TCRs restricted to three different HLAs, namely HLA-B0702, -A2402, or -A0201, and hence will provide coverage for a large fraction of patients with MCC. Among the 12 HLA-A0201-restricted TCRs tested, we discovered several TCRs that would potentially cross-recognize off-target peptides that might also be naturally presented. These may therefore be excluded from further clinical evaluation.

Knowing the landscape of the cross-recognized pMHCs of individual receptors has proven particularly important for the *in vitro* affinity-optimized TCRs developed for clinical applications. Due to the severe adverse effects observed with such genetically modified TCRs, novel structure-guided approaches are concerned with the effect of modifying TCRs in a way that finds an optimal TCR affinity-range (Hebeisen et al., 2013a; Slansky and Jordan, 2010; Stone and Kranz, 2013; Zhong et al., 2013), and propose to incorporate changes that simultaneously enhance and decrease the affinity of the pMHC-TCR interaction (Riley and Baker, 2018). Basically, it is suggested that a TCR with optimal on-target(-only) affinity may be generated by modifying the TCR in a way that promote a stronger binding to the peptide-part of the TCR recognition motif, while decreasing the overall binding to MHC. Support of this theory is slowly emerging (Hellman et al., 2018), and provides an increased motivation for describing the molecular recognition patterns of a TCR to assist the process of "rational-design" of genetically modified TCRs intended for the clinic. Moreover, knowledge of the molecular interaction-points of TCRs can feed into developing *in silico* platforms, increasing our overall capacity to understand TCR interactions and predict T cell cross-recognition.

Predicting TCR targets. Together with continuously emerging techniques that enable comprehensive interrogation of T cell immunology, the technologies developed in this thesis are allowing us to get insight into the enormous complexity of T cell interactions. In particular any links to the relationship between the TCR sequence and pMHC recognition, which is highly anticipated, has proven difficult to predict because of the complex nature of these interactions. Some progresses have been made in this area (Yu et al., 2015), but have so far focused on a few frequently observed viral antigens, and the most advanced algorithms are currently able to predict only the similarity among TCRs recognizing the same pMHC (Dash et al., 2017; Glanville et al., 2017; Jurtz et al., 2018). Progress in describing such TCR recognition has largely been impeded due to lack of technologies that can accommodate the simultaneous identification of antigen-specificity and TCR sequence in a high-throughput format and across multiple pMHC specificities. By tracing MHC multimer specificities with DNA barcodes we have provided a high-complexity label that can be amplified and, in theory, sequenced along with any transcriptomic information from cells. Thus, by implementing this in increasingly high-throughput single-cell platforms, the technology provides the means of capturing large numbers of paired TCRs together with DNA barcodes encoding the antigen-specificity (Zhang et al., 2018b and **Paper V**). This provides the opportunity for generating a comprehensive data set of full TCR sequences matched with information of cognate pMHC specificity, which in turn can be fed into computational models. Combined with the improvements in technologies that enable us in describing the TCR recognition pattern (Bentzen et al., 2018; Birnbaum et al., 2014; Wooldridge et al., 2010), and structural knowledge of pMHC interactions (Fodor et al., 2018; Reinherz et al., 1999), this may ultimately assist the prediction of specific pMHC-interactions based on the TCR sequence (Gee et al., 2018a).

Integrated T cell profiling. By fully utilizing the molecular-encoded single-cell technologies in combination with our DNA barcode-labeled MHC multimers, we will be able to further increase the breadth of our T cell analyses through integration of full transcriptomic profiling of cells and protein expression by use of DNA barcode-labeled antibodies (Peterson et al., 2017; Stoeckius et al., 2017). Such analysis will simultaneously inform on 1) the pMHC specificitie(s) across a large library of potential pMHC epitopes, 2) the paired TCR sequence, 3) the full transcriptomic quantities, and 4) the broad proteomic phenotype of numerous, possibly 10,000s, of single T cells profiled in parallel. It moreover offers the possibility of simultaneously performing the same extensive phenotypic and transcriptomic analysis of non-MHC multimer binding cells (see proposed workflow in Figure 8). Such extensive single-cell analyses will at one hand enable us in describing T cell recognition at a complexity beginning to match the number of pMHCs of interest in a given disease (especially if we can increase the complexity to the proposed 10,000-100,000), and at the other hand enable an in-depth analysis of the functional, phenotypic or exhaustion state of every single of the pMHC interacting and non-interacting immune cells. By additionally capturing the sequence of the TCR, we will have the option of modeling the

interaction and investigating the structural relationship between the TCR and the recognized pMHC, and maybe uncover patterns decisive for the type of antigen recognized and the concomitant type of response.

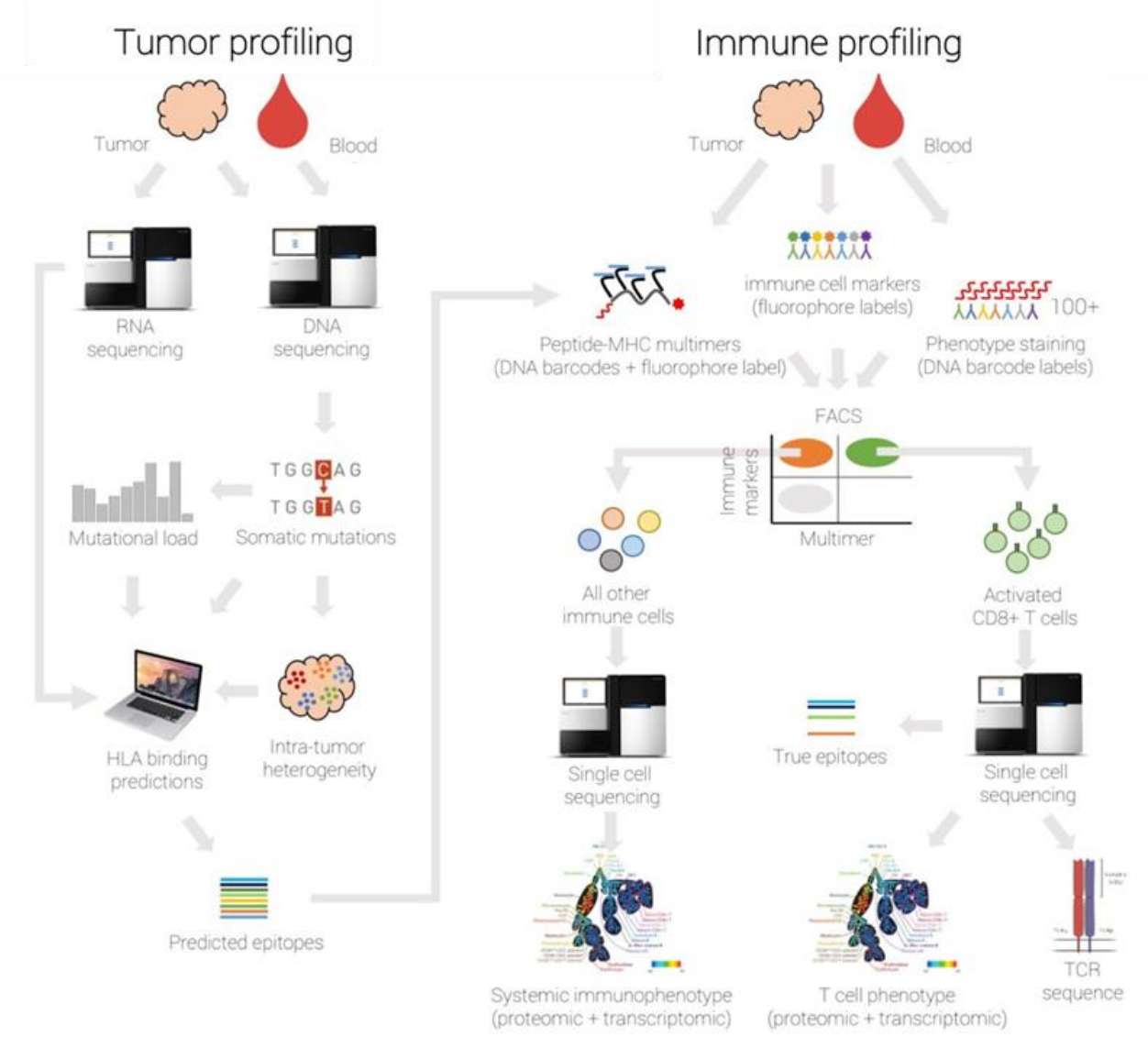


Figure 8 | **The workflow of the proposed integrated T cell profiling.** Here applied in a cancer setting where tumor profiling leads to the prediction of a large number of potential neoepitopes that forms the basis of the MHC multimer part of the integrated analysis. Details of the full integrated immune profiling are described in the main text.

The analysis will still not be able to exhaustively capture information of all potentially cross-recognized targets at a personal level, thus further investigations of the binding degeneracy of the individual TCRs will be necessary. As these investigations become more thorough, convenient, and in the future potentially also achievable in a high-throughput manner, there is ample scope to develop more complete maps of the extended recognition-potential of TCR repertoires. This will provide whole new opportunities for describing the relationships between immune cells; how the dynamics of T cell responses may be affected by other immune cells, which is indicated by dynamic changes of antigen-specificities after adoptive transfer of tumor infiltrating lymphocytes (Zacharakis et al., 2018), or how T cell responses that target one pathogen or malignancy might impact the response to several other antigens, potentially providing protection against a completely different antigen or inducing an autoimmune reaction (Ang et al., 2004; Rose, 2017), and finally how this may vary between individuals with each their private repertoire and immunological history. Such extensive understanding of the dynamics of T cell responses will enable the development of more personalized treatments, which directly accounts for the individual's endogenous ability to target a given disease-antigen, and at the same time appropriately considers the personal risk of inducing an autoimmune reaction.

Paper I

Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes

Paper II

Evolution of MHC-based technologies
used for detection of antigen-responsive T cells

Paper III

T cell receptor fingerprinting enables in-depth characterization of the interactions governing recognition of peptide-MHC complexes

Paper IV

Interrogating T cell recognition by use of DNA barcode-labeled peptide-MHC multimers

Manuscript in preparation. It is formatted with the intention of submitting to Nature Protocols. Manuscript preparation is ongoing and is in the current form particularly lacking details on data analyses.

Paper V

Simultaneous identification of antigen-specificity and T cell receptor sequence in single cells

The following research comprises early findings from a proof-of-concept study for developing a high-throughput single-cell platform for simultaneous capture of pMHC specificity and TCR sequence. The manuscript is a summary of current results from an ongoing study, and further analyses will be included once the platform has been refined. Currently the manuscript comprises data from a pilot study with details of the experimental setup and schematic figures explaining the overall concept.

Abbreviations

AIDS	acquired immunodeficiency disease
APC	antigen presenting cell
B	biotin
C	constant
CAR	chimeric antigen receptor
CDR	complimentary-determining region
CMV	cytomegalovirus
CPL	combinatorial peptide library
D	diversity
EBV	Ebstein-Barr virus
ELISPOT	enzyme-linked immunospot
FACS	fluorescence-activated cell sorting
FLU	influenza
HLA	Human Leukocyte Antigen
HIV	human immunodeficiency virus
HPV	human papilloma virus
IVTT	in vitro transcription and translation
J	joining
MCC	Merkel cell carcinoma
MCP*y*V	Merkel cell polyomavirus
MHC	major histocompatibility complex
p	peptide
PE	phycoerythin
Sample ID	DNA barcode denoting all molecules derived from one sample
TCR	T cell receptor
UMI	unique molecular identifier
V	variable

Bibliography

Acha-Orbea, H., Mitchell, D.J., Timmermann, L., Wraith, D.C., Tausch, G.S., Waldor, M.K., Zamvil, S.S., McDevitt, H.O., and Steinman, L. (1988). Limited heterogeneity of T cell receptors from lymphocytes mediating autoimmune encephalomyelitis allows specific immune intervention. Cell *54*, 263–273.

Acierno, P.M., Newton, D.A., Brown, E.A., Maes, L.A., Baatz, J.E., and Gattoni-Celli, S. (2003). Cross-reactivity between HLA-A2-restricted FLU-M1:58-66 and HIV p17 GAG:77-85 epitopes in HIV-infected and uninfected individuals. J. Transl. Med. *1*, 3.

Adams, J.J., Narayanan, S., Birnbaum, M.E., Sidhu, S.S., Blevins, S.J., Gee, M.H., Sibener, L.V., Baker, B.M., Kranz, D.M., and Garcia, K.C. (2016). Structural interplay between germline interactions and adaptive recognition determines the bandwidth of TCR-peptide-MHC cross-reactivity. Nat. Immunol. *17*, 87–94.

Ahmed, S.S., Volkmuth, W., Duca, J., Corti, L., Pallaoro, M., Pezzicoli, A., Karle, A., Rigat, F., Rappuoli, R., Narasimhan, V., et al. (2015). Antibodies to influenza nucleoprotein cross-react with human hypocretin receptor 2. Sci. Transl. Med. *7*, 294ra105.

Albert, L.J., and Inman, R.D. (1999). Molecular mimicry and autoimmunity. N. Engl. J. Med. *341*, 2068–2074.

Van Allen, E.M., Miao, D., Schilling, B., Shukla, S.A., Blank, C., Zimmer, L., Sucker, A., Hillen, U., Foppen, M.H.G., Goldinger, S.M., et al. (2015). Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. Science *350*, 207–211.

Altman, J.D., Moss, P.A., Goulder, P.J., Barouch, D.H., McHeyzer-Williams, M.G., Bell, J.I., McMichael, A.J., and Davis, M.M. (1996). Phenotypic analysis of antigen-specific T lymphocytes. Science *274*, 94–96.

Andersen, R.S., Kvistborg, P., Frøsig, T.M., Pedersen, N.W., Lyngaa, R., Bakker, A.H., Shu, C.J., Straten, P. thor, Schumacher, T.N., and Hadrup, S.R. (2012a). Parallel detection of antigen-specific T cell responses by combinatorial encoding of MHC multimers. Nat. Protoc. *7*, 891–902.

Andersen, R.S., Thrue, C.A., Junker, N., Lyngaa, R., Donia, M., Ellebæk, E., Svane, I.M., Schumacher, T.N., Thor Straten, P., and Hadrup, S.R. (2012b). Dissection of T-cell antigen specificity in human melanoma. Cancer Res. *72*, 1642–1650.

Andersen, R.S., Andersen, S.R., Hjortsø, M.D., Lyngaa, R., Idorn, M., Køllgård, T.M., Met, O., Thor Straten, P., and Hadrup, S.R. (2013). High frequency of T cells specific for cryptic epitopes in melanoma patients. Oncoimmunology *2*, e25374.

Andreatta, M., Karosiene, E., Rasmussen, M., Stryhn, A., Buus, S., and Nielsen, M. (2015). Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. Immunogenetics *67*, 641–650.

Ang, C.W., Jacobs, B.C., and Laman, J.D. (2004). The Guillain-Barré syndrome: a true case of molecular mimicry. Trends Immunol. *25*, 61–66.

Arstila, T.P., Casrouge, A., Baron, V., Even, J., Kanellopoulos, J., and Kourilsky, P. (1999). A direct estimate of the human alphabeta T cell receptor diversity. Science *286*, 958–961.

Ashton-Rickardt, P.G., Van Kaer, L., Schumacher, T.N., Ploegh, H.L., and Tonegawa, S. (1993). Peptide contributes to the specificity of positive selection of CD8+ T cells in the thymus. Cell *73*, 1041–1049.

Babon, J.A.B., DeNicola, M.E., Blodgett, D.M., Crèvecoeur, I., Buttrick, T.S., Maehr, R., Bottino, R., Naji, A., Kaddis, J., Elyaman, W., et al. (2016). Analysis of self-antigen specificity of islet-infiltrating T cells from human donors with type 1 diabetes. Nat. Med. *22*, 1482–1487.

Bais, P., Namburi, S., Gatti, D.M., Zhang, X., and Chuang, J.H. (2017). CloudNeo: a cloud pipeline for identifying patient-specific tumor neoantigens. Bioinformatics *33*, 3110–3112.

Bakker, A.H., and Schumacher, T.N.M. (2005). MHC multimer technology: current status and future prospects. Curr. Opin. Immunol. *17*, 428–433.

Bakker, A.H., Hoppes, R., Linnemann, C., Toebes, M., Rodenko, B., Berkers, C.R., Hadrup, S.R., van Esch, W.J.E., Heemskerk, M.H.M., Ovaa, H., et al. (2008). Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7. Proc. Natl. Acad. Sci. USA *105*, 3825–3830.

Bankovich, A.J., Girvin, A.T., Moesta, A.K., and Garcia, K.C. (2004). Peptide register shifting within the MHC groove: theory becomes reality. Mol. Immunol. *40*, 1033–1039.

Bassani-Sternberg, M., Bräunlein, E., Klar, R., Engleitner, T., Sinitcyn, P., Audehm, S., Straub, M., Weber, J., Slotta-Huspenina, J., Specht, K., et al. (2016). Direct identification of clinically

relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. Nat. Commun. *7*, 13404.

Beck, S., and Trowsdale, J. (2000). The human major histocompatability complex: lessons from the DNA sequence. Annu. Rev. Genomics. Hum. Genet. *1*, 117–137.

Benoist, C., and Mathis, D. (2001). Autoimmunity provoked by infection: how good is the case for T cell epitope mimicry? Nat. Immunol. *2*, 797–801.

Bentzen, A.K., Marquard, A.M., Lyngaa, R., Saini, S.K., Ramskov, S., Donia, M., Such, L., Furness, A.J.S., McGranahan, N., Rosenthal, R., et al. (2016). Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. Nat. Biotechnol. *34*, 1037–1045.

Bentzen, A.K., and Hadrup, S.R. (2017). Evolution of MHC-based technologies used for detection of antigen-responsive T cells. Cancer Immunol Immunother. *66*, 657–666.

Bentzen, A.K., Such, L., Jensen, K.K., Marquard, A.M., Jessen, L.E., Miller, N.J., Church, C.D., Lyngaa, R., Koelle, D.M., Becker, J.C., et al. (2018). T cell receptor fingerprinting enables in-depth characterization of the interactions governing recognition of peptide-MHC complexes. Nat. Biotechnol. *36*, 1191–1196.

Bergman, Y. (1999). Allelic exclusion in B and T lymphopoiesis. Semin. Immunol. *11*, 319–328.

Beringer, D.X., Kleijwegt, F.S., Wiede, F., van der Slik, A.R., Loh, K.L., Petersen, J., Dudek, N.L., Duinkerken, G., Laban, S., Joosten, A., et al. (2015). T cell receptor reversed polarity recognition of a self-antigen major histocompatibility complex. Nat. Immunol. *16*, 1153–1161.

Besser, M.J., Shapira-Frommer, R., Treves, A.J., Zippel, D., Itzhaki, O., Hershkovitz, L., Levy, D., Kubi, A., Hovav, E., Chermoshniuk, N., et al. (2010). Clinical responses in a phase II study using adoptive transfer of short-term cultured tumor infiltration lymphocytes in metastatic melanoma patients. Clin. Cancer Res. *16*, 2646–2655.

Bevan, M.J. (1977). In a radiation chimaera, host H-2 antigens determine immune responsiveness of donor cytotoxic cells. Nature *269*, 417–418.

Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R., Frabetti, F., Vitale, L., Pelleri, M.C., Tassani, S., Piva, F., et al. (2013). An estimation of the number of cells in the human body. Ann. Hum. Biol. *40*, 463–471.

Bijen, H.M., van der Steen, D.M., Hagedoorn, R.S., Wouters, A.K., Wooldridge, L., Falkenburg, J.H.F., and Heemskerk, M.H.M. (2018). Preclinical Strategies to Identify Off-Target Toxicity of High-Affinity TCRs. Mol. Ther. *26*, 1206–1214.

Birnbaum, M.E., Mendoza, J.L., Sethi, D.K., Dong, S., Glanville, J., Dobbins, J., Ozkan, E., Davis, M.M., Wucherpfennig, K.W., and Garcia, K.C. (2014). Deconstructing the peptide-MHC specificity of T cell recognition. Cell *157*, 1073–1087.

Bjerregaard, A.-M., Nielsen, M., Hadrup, S.R., Szallasi, Z., and Eklund, A.C. (2017a). MuPeXI: prediction of neo-epitopes from tumor sequencing data. Cancer Immunol. Immunother. *66*, 1123–1130.

Bjerregaard, A.-M., Nielsen, M., Jurtz, V., Barra, C.M., Hadrup, S.R., Szallasi, Z., and Eklund, A.C. (2017b). An analysis of natural T cell responses to predicted tumor neoepitopes. Front. Immunol. *8*, 1566.

Borbulevych, O.Y., Piepenbrink, K.H., Gloor, B.E., Scott, D.R., Sommese, R.F., Cole, D.K., Sewell, A.K., and Baker, B.M. (2009). T cell receptor cross-reactivity directed by antigen-dependent tuning of peptide-MHC molecular flexibility. Immunity *31*, 885–896.

Borbulevych, O.Y., Santhanagopolan, S.M., Hossain, M., and Baker, B.M. (2011). TCRs used in cancer gene therapy cross-react with MART-1/Melan-A tumor antigens via distinct mechanisms. J. Immunol. *187*, 2453–2463.

Bouvier, M., and Wiley, D.C. (1994). Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. Science *265*, 398–402.

Bowerman, N.A., Crofts, T.S., Chlewicki, L., Do, P., Baker, B.M., Christopher Garcia, K., and Kranz, D.M. (2009). Engineering the binding properties of the T cell receptor:peptide:MHC ternary complex that governs T cell activity. Mol. Immunol. *46*, 3000–3008.

Brenner, S., and Lerner, R.A. (1992). Encoded combinatorial chemistry. Proc. Natl. Acad. Sci. USA *89*, 5381–5383.

Brooks, S.E., Bonney, S.A., Lee, C., Publicover, A., Khan, G., Smits, E.L., Sigurdardottir, D., Arno, M., Li, D., Mills, K.I., et al. (2015). Application of the pMHC Array to Characterise Tumour Antigen Specific T Cell Populations in Leukaemia Patients at Disease Diagnosis. PLoS One *10*, e0140483.

Bulik-Sullivan, B., Busby, J., Palmer, C.D., Davis, M.J., Murphy, T., Clark, A., Busby, M., Duke, F., Yang, A., Young, L., et al. (2019). Deep learning using tumor HLA peptide mass spectrometry datasets improves neoantigen identification. Nat. Biotechnol. 37, 55–63.

Burrows, S.R., Chen, Z., Archbold, J.K., Tynan, F.E., Beddoe, T., Kjer-Nielsen, L., Miles, J.J., Khanna, R., Moss, D.J., Liu, Y.C., et al. (2010). Hard wiring of T cell receptor specificity for the major histocompatibility complex is underpinned by TCR adaptability. Proc. Natl. Acad. Sci. USA *107*, 10608–10613.

Cameron, B.J., Gerry, A.B., Dukes, J., Harper, J.V., Kannan, V., Bianchi, F.C., Grand, F., Brewer, J.E., Gupta, M., Plesa, G., et al. (2013). Identification of a Titin-derived HLA-A1-presented peptide as a cross-reactive target for engineered MAGE A3-directed T cells. Sci. Transl. Med. *5*, 197ra103.

Canale, F.P., Ramello, M.C., Núñez, N., Araujo Furlan, C.L., Bossio, S.N., Gorosito Serrán, M., Tosello Boari, J., Del Castillo, A., Ledesma, M., Sedlik, C., et al. (2018). CD39 Expression Defines Cell Exhaustion in Tumor-Infiltrating CD8+ T Cells. Cancer Res. *78*, 115–128.

Cao, J., Packer, J.S., Ramani, V., Cusanovich, D.A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S.N., Steemers, F.J., et al. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. Science *357*, 661–667.

Cao, J., Cusanovich, D.A., Ramani, V., Aghamirzaie, D., Pliner, H.A., Hill, A.J., Daza, R.M., McFaline-Figueroa, J.L., Packer, J.S., Christiansen, L., et al. (2018). Joint profiling of chromatin accessibility and gene expression in thousands of single cells. Science *361*, 1380–1385.

Carreno, B.M., Magrini, V., Becker-Hapak, M., Kaabinejadian, S., Hundal, J., Petti, A.A., Ly, A., Lie, W.-R., Hildebrand, W.H., Mardis, E.R., et al. (2015). Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. Science *348*, 803–808.

Carrington, M., Nelson, G.W., Martin, M.P., Kissner, T., Vlahov, D., Goedert, J.J., Kaslow, R., Buchbinder, S., Hoots, K., and O'Brien, S.J. (1999). HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. Science *283*, 1748–1752.

Chakravarty, S.D., Zabriskie, J.B., and Gibofsky, A. (2014). Acute rheumatic fever and streptococci: the quintessential pathogenic trigger of autoimmunity. Clin. Rheumatol. *33*, 893–901.

Chamoto, K., Tsuji, T., Funamoto, H., Kosaka, A., Matsuzaki, J., Sato, T., Abe, H., Fujio, K., Yamamoto, K., Kitamura, T., et al. (2004). Potentiation of tumor eradication by adoptive immunotherapy with T-cell receptor gene-transduced T-helper type 1 cells. Cancer Res. *64*, 386–390.

Chan, K.F., Gully, B.S., Gras, S., Beringer, D.X., Kjer-Nielsen, L., Cebon, J., McCluskey, J., Chen, W., and Rossjohn, J. (2018). Divergent T-cell receptor recognition modes of a HLA-I restricted extended tumour-associated peptide. Nat. Commun. *9*, 1026.

Chattopadhyay, P.K., Gierahn, T.M., Roederer, M., and Love, J.C. (2014). Single-cell technologies for monitoring immune systems. Nat. Immunol. *15*, 128–135.

Chen, A.T., Cornberg, M., Gras, S., Guillonneau, C., Rossjohn, J., Trees, A., Emonet, S., de la Torre, J.C., Welsh, R.M., and Selin, L.K. (2012). Loss of anti-viral immunity by infection with a virus encoding a cross-reactive pathogenic epitope. PLoS Pathog. *8*, e1002633.

Chen, D.S., Soen, Y., Stuge, T.B., Lee, P.P., Weber, J.S., Brown, P.O., and Davis, M.M. (2005). Marked differences in human melanoma antigen-specific T cell responsiveness after vaccination using a functional microarray. PLoS Med. *2*, e265.

Chen, W., Yewdell, J.W., Levine, R.L., and Bennink, J.R. (1999). Modification of cysteine residues in vitro and in vivo affects the immunogenicity and antigenicity of major histocompatibility complex class I-restricted viral determinants. J. Exp. Med. *189*, 1757–1764.

Chervin, A.S., Stone, J.D., Holler, P.D., Bai, A., Chen, J., Eisen, H.N., and Kranz, D.M. (2009). The impact of TCR-binding properties and antigen presentation format on T cell responsiveness. J. Immunol. *183*, 1166–1178.

Chowell, D., Krishna, S., Becker, P.D., Cocita, C., Shu, J., Tan, X., Greenberg, P.D., Klavinskis, L.S., Blattman, J.N., and Anderson, K.S. (2015). TCR contact residue hydrophobicity is a hallmark of immunogenic CD8+ T cell epitopes. Proc. Natl. Acad. Sci. USA *112*, E1754–62.

Chowell, D., Morris, L.G.T., Grigg, C.M., Weber, J.K., Samstein, R.M., Makarov, V., Kuo, F., Kendall, S.M., Requena, D., Riaz, N., et al. (2018). Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. Science *359*, 582–587.

Clay, T.M., Custer, M.C., Sachs, J., Hwu, P., Rosenberg, S.A., and Nishimura, M.I. (1999). Efficient transfer of a tumor antigen-reactive TCR to human peripheral blood lymphocytes confers anti-tumor reactivity. J. Immunol. *163*, 507–513.

Clute, S.C., Naumov, Y.N., Watkin, L.B., Aslan, N., Sullivan, J.L., Thorley-Lawson, D.A., Luzuriaga, K., Welsh, R.M., Puzone, R., Celada, F., et al. (2010). Broad cross-reactive TCR repertoires recognizing dissimilar Epstein-Barr and influenza A virus epitopes. J. Immunol. *185*, 6753–6764.

Cochet, M., Pannetier, C., Regnault, A., Darche, S., Leclerc, C., and Kourilsky, P. (1992). Molecular detection and in vivo analysis of the specific T cell response to a protein antigen. Eur. J. Immunol. *22*, 2639–2647.

Cohn, M. (2016). Dissecting the two models of TCR structure-function relationships. Immunol. Res. *64*, 795–803.

Cole, D.K., Bulek, A.M., Dolton, G., Schauenberg, A.J., Szomolay, B., Rittase, W., Trimby, A., Jothikumar, P., Fuller, A., Skowera, A., et al. (2016). Hotspot autoimmune T cell receptor binding underlies pathogen and insulin peptide cross-reactivity. J. Clin. Invest. *126*, 2191–2204.

Cole, D.K., van den Berg, H.A., Lloyd, A., Crowther, M.D., Beck, K., Ekeruche-Makinde, J., Miles, J.J., Bulek, A.M., Dolton, G., Schauenburg, A.J., et al. (2017). Structural Mechanism Underpinning Cross-reactivity of a CD8+ T-cell Clone That Recognizes a Peptide Derived from Human Telomerase Reverse Transcriptase. J. Biol. Chem. *292*, 802–813.

Colugnati, F.A.B., Staras, S.A.S., Dollard, S.C., and Cannon, M.J. (2007). Incidence of cytomegalovirus infection among the general population and pregnant women in the United States. BMC Infect. Dis. *7*, 71.

Connerotte, T., Van Pel, A., Godelaine, D., Tartour, E., Schuler-Thurner, B., Lucas, S., Thielemans, K., Schuler, G., and Coulie, P.G. (2008). Functions of Anti-MAGE T-cells induced in melanoma patients under different vaccination modalities. Cancer Res. *68*, 3931–3940.

Cooper, L.J., Kalos, M., Lewinsohn, D.A., Riddell, S.R., and Greenberg, P.D. (2000). Transfer of specificity for human immunodeficiency virus type 1 into primary human T lymphocytes by introduction of T-cell receptor genes. J. Virol. *74*, 8207–8212.

Cornberg, M., and Wedemeyer, H. (2016). Hepatitis C virus infection from the perspective of heterologous immunity. Curr Opin Virol *16*, 41–48.

Cornberg, M., Chen, A.T., Wilkinson, L.A., Brehm, M.A., Kim, S.-K., Calcagno, C., Ghersi, D., Puzone, R., Celada, F., Welsh, R.M., et al. (2006). Narrowed TCR repertoire and viral escape as a consequence of heterologous immunity. J. Clin. Invest. *116*, 1443–1456.

Cornberg, M., Clute, S.C., Watkin, L.B., Saccoccio, F.M., Kim, S.-K., Naumov, Y.N., Brehm, M.A., Aslan, N., Welsh, R.M., and Selin, L.K. (2010). CD8 T cell cross-reactivity networks mediate heterologous immunity in human EBV and murine vaccinia virus infections. J. Immunol. *184*, 2825–2838.

Cornberg, M., Kenney, L.L., Chen, A.T., Waggoner, S.N., Kim, S.-K., Dienes, H.P., Welsh, R.M., and Selin, L.K. (2013). Clonal exhaustion as a mechanism to protect against severe immunopathology and death from an overwhelming CD8 T cell response. Front. Immunol. *4*, 475.

Cossarizza, A., Chang, H.-D., Radbruch, A., Akdis, M., Andrä, I., Annunziato, F., Bacher, P., Barnaba, V., Battistini, L., Bauer, W.M., et al. (2017). Guidelines for the use of flow cytometry and cell sorting in immunological studies. Eur. J. Immunol. *47*, 1584–1797.

Couzin-Frankel, J. (2013). Breakthrough of the year 2013. Cancer immunotherapy. Science *342*, 1432–1433.

Çuburu, N., and Schiller, J.T. (2016). Moving forward with human papillomavirus immunotherapies. Hum. Vaccin. Immunother. *12*, 2875–2880.

Curtsinger, J.M., Lins, D.C., and Mescher, M.F. (1998). CD8+ memory T cells (CD44high, Ly-6C+) are more sensitive than naive cells to (CD44low, Ly-6C-) to TCR/CD8 signaling in response to antigen. J. Immunol. *160*, 3236–3243.

Dash, P., Wang, G.C., and Thomas, P.G. (2015). Single-Cell Analysis of T-Cell Receptor αβ Repertoire. Methods Mol. Biol. *1343*, 181–197.

Dash, P., Fiore-Gartland, A.J., Hertz, T., Wang, G.C., Sharma, S., Souquette, A., Crawford, J.C., Clemens, E.B., Nguyen, T.H.O., Kedzierska, K., et al. (2017). Quantifiable predictive features define epitope-specific T cell receptor repertoires. Nature *547*, 89–93.

Dauvilliers, Y., Arnulf, I., Lecendreux, M., Monaca Charley, C., Franco, P., Drouot, X., d Ortho, M.-P., Launois, S., Lignot, S., Bourgin, P., et al. (2013). Increased risk of narcolepsy in children and adults after pandemic H1N1 vaccination in France. Brain *136*, 2486–2496.

Davis, M.M., and Bjorkman, P.J. (1988). T-cell antigen receptor genes and T-cell recognition. Nature *334*, 395–402.

Davis, M.M., Boniface, J.J., Reich, Z., Lyons, D., Hampl, J., Arden, B., and Chien, Y. (1998). Ligand recognition by alpha beta T cell receptors. Annu. Rev. Immunol. *16*, 523–544.

Day, C.L., Seth, N.P., Lucas, M., Appel, H., Gauthier, L., Lauer, G.M., Robbins, G.K., Szczepiorkowski, Z.M., Casson, D.R., Chung, R.T., et al. (2003). Ex vivo analysis of human memory CD4 T cells specific for hepatitis C virus using MHC class II tetramers. J. Clin. Invest. *112*, 831–842.

Degauque, N., Brouard, S., and Soulillou, J.-P. (2016). Cross-Reactivity of TCR Repertoire: Current Concepts, Challenges, and Implication for Allotransplantation. Front. Immunol. *7*, 89.

Dembić, Z., Haas, W., Weiss, S., McCubrey, J., Kiefer, H., von Boehmer, H., and Steinmetz, M. (1986). Transfer of specificity by murine alpha and beta T-cell receptor genes. Nature *320*, 232–238.

Demotz, S., Grey, H.M., and Sette, A. (1990). The minimal number of class II MHC-antigen complexes needed for T cell activation. Science *249*, 1028–1030.

Dendrou, C.A., Petersen, J., Rossjohn, J., and Fugger, L. (2018). HLA variation and disease. Nat. Rev. Immunol. *18*, 325–339.

Deviren, G., Gupta, K., Paulaitis, M.E., and Schneck, J.P. (2007). Detection of antigen-specific T cells on p/MHC microarrays. J Mol Recognit *20*, 32–38.

Dolton, G., Lissina, A., Skowera, A., Ladell, K., Tungatt, K., Jones, E., Kronenberg-Versteeg, D., Akpovwa, H., Pentier, J.M., Holland, C.J., et al. (2014). Comparison of peptide-major

histocompatibility complex tetramers and dextramers for the identification of antigen-specific T cells. Clin. Exp. Immunol. *177*, 47–63.

Dolton, G., Zervoudi, E., Rius, C., Wall, A., Thomas, H.L., Fuller, A., Yeo, L., Legut, M., Wheeler, S., Attaf, M., et al. (2018). Optimized Peptide-MHC Multimer Protocols for Detection and Isolation of Autoimmune T-Cells. Front. Immunol. *9*, 1378.

Duhen, T., Duhen, R., Montler, R., Moses, J., Moudgil, T., de Miranda, N.F., Goodall, C.P., Blair, T.C., Fox, B.A., McDermott, J.E., et al. (2018). Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. Nat. Commun. *9*, 2724.

Dunn, G.P., Old, L.J., and Schreiber, R.D. (2004a). The immunobiology of cancer immunosurveillance and immunoediting. Immunity *21*, 137–148.

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

Dziubianau, M., Hecht, J., Kuchenbecker, L., Sattler, A., Stervbo, U., Rödelsperger, C., Nickel, P., Neumann, A.U., Robinson, P.N., Mundlos, S., et al. (2013). TCR repertoire analysis by next generation sequencing allows complex differential diagnosis of T cell-related pathology. Am. J. Transplant. *13*, 2842–2854.

Egorov, E.S., Kasatskaya, S.A., Zubov, V.N., Izraelson, M., Nakonechnaya, T.O., Staroverov, D.B., Angius, A., Cucca, F., Mamedov, I.Z., Rosati, E., et al. (2018). The changing landscape of naive T cell receptor repertoire with human aging. Front. Immunol. *9*, 1618.

Ehrenstein, M.R., Evans, J.G., and Singh, A. (2004). Compromised function of regulatory T cells in rheumatoid arthritis and reversal by anti-TNFα therapy. Journal Of ….

Ekeruche-Makinde, J., Clement, M., Cole, D.K., Edwards, E.S.J., Ladell, K., Miles, J.J., Matthews, K.K., Fuller, A., Lloyd, K.A., Madura, F., et al. (2012). T-cell receptor-optimized peptide skewing of the T-cell repertoire can enhance antigen targeting. J. Biol. Chem. *287*, 37269–37281.

Ekeruche-Makinde, J., Miles, J.J., van den Berg, H.A., Skowera, A., Cole, D.K., Dolton, G., Schauenburg, A.J.A., Tan, M.P., Pentier, J.M., Llewellyn-Lacey, S., et al. (2013). Peptide length determines the outcome of TCR/peptide-MHCI engagement. Blood *121*, 1112–1123.

Ellebaek, E., Iversen, T.Z., Junker, N., Donia, M., Engell-Noerregaard, L., Met, Ö., Hölmich, L.R., Andersen, R.S., Hadrup, S.R., Andersen, M.H., et al. (2012). Adoptive cell therapy with autologous tumor infiltrating lymphocytes and low-dose Interleukin-2 in metastatic melanoma patients. J. Transl. Med. *10*, 169.

Ellis, J.M., Henson, V., Slack, R., Ng, J., Hartzman, R.J., and Katovich Hurley, C. (2000). Frequencies of HLA-A2 alleles in five U.S. population groups. Hum. Immunol. *61*, 334–340.

Engels, B., Engelhard, V.H., Sidney, J., Sette, A., Binder, D.C., Liu, R.B., Kranz, D.M., Meredith, S.C., Rowley, D.A., and Schreiber, H. (2013). Relapse or eradication of cancer is predicted by peptide-major histocompatibility complex affinity. Cancer Cell *23*, 516–526.

Evavold, B.D., Sloan-Lancaster, J., Wilson, K.J., Rothbard, J.B., and Allen, P.M. (1995). Specific T cell recognition of minimally homologous peptides: evidence for multiple endogenous ligands. Immunity *2*, 655–663.

Favalli, N., Bassi, G., Scheuermann, J., and Neri, D. (2018). DNA-encoded chemical libraries - achievements and remaining challenges. FEBS Lett. *592*, 2168–2180.

Feng, D., Bond, C.J., Ely, L.K., Maynard, J., and Garcia, K.C. (2007). Structural evidence for a germline-encoded T cell receptor-major histocompatibility complex interaction “codon”. Nat. Immunol. *8*, 975–983.

Feng, H., Shuda, M., Chang, Y., and Moore, P.S. (2008). Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science *319*, 1096–1100.

Ferraro, A., Socci, C., Stabilini, A., Valle, A., Monti, P., Piemonti, L., Nano, R., Olek, S., Maffi, P., Scavini, M., et al. (2011). Expansion of Th17 cells and functional defects in T regulatory cells are key features of the pancreatic lymph nodes in patients with type 1 diabetes. Diabetes *60*, 2903–2913.

Fodor, J., Riley, B.T., Borg, N.A., and Buckle, A.M. (2018). Previously Hidden Dynamics at the TCR-Peptide-MHC Interface Revealed. J. Immunol. *200*, 4134–4145.

Fulwyler, M.J. (1965). Electronic separation of biological cells by volume. Science *150*, 910–911.

Garboczi, D.N., Hung, D.T., and Wiley, D.C. (1992). HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in Escherichia coli and complexed with single antigenic peptides. Proc. Natl. Acad. Sci. USA *89*, 3429–3433.

Garboczi, D.N., Ghosh, P., Utz, U., Fan, Q.R., Biddison, W.E., and Wiley, D.C. (1996). Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. Nature *384*, 134–141.

Garcia, K.C., Degano, M., Pease, L.R., Huang, M., Peterson, P.A., Teyton, L., and Wilson, I.A. (1998). Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. Science *279*, 1166–1172.

Garcia, K.C., Teyton, L., and Wilson, I.A. (1999). Structural basis of T cell recognition. Annu. Rev. Immunol. *17*, 369–397.

Gee, M.H., Han, A., Lofgren, S.M., Beausang, J.F., Mendoza, J.L., Birnbaum, M.E., Bethune, M.T., Fischer, S., Yang, X., Gomez-Eerland, R., et al. (2018a). Antigen Identification for Orphan T Cell Receptors Expressed on Tumor-Infiltrating Lymphocytes. Cell *172*, 549–563.e16.

Gee, M.H., Yang, X., and Garcia, K.C. (2018b). Facile method for screening clinical T cell receptors for off-target peptide-HLA reactivity. BioRxiv. doi.org/10.1101/472480

Gerlach, C., Rohr, J.C., Perié, L., van Rooij, N., van Heijst, J.W.J., Velds, A., Urbanus, J., Naik, S.H., Jacobs, H., Beltman, J.B., et al. (2013). Heterogeneous differentiation patterns of individual CD8+ T cells. Science *340*, 635–639.

Gerlinger, M., Rowan, A.J., Horswell, S., Math, M., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., et al. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N. Engl. J. Med. *366*, 883–892.

Gfeller, D., Guillaume, P., Michaux, J., Pak, H.-S., Daniel, R.T., Racle, J., Coukos, G., and Bassani-Sternberg, M. (2018). The Length Distribution and Multiple Specificity of Naturally Presented HLA-I Ligands. J. Immunol. *201*, 3705–3716.

Gierahn, T.M., Wadsworth, M.H., Hughes, T.K., Bryson, B.D., Butler, A., Satija, R., Fortune, S., Love, J.C., and Shalek, A.K. (2017). Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. Nat. Methods *14*, 395–398.

Glanville, J., Huang, H., Nau, A., Hatton, O., Wagar, L.E., Rubelt, F., Ji, X., Han, A., Krams, S.M., Pettus, C., et al. (2017). Identifying specificity groups in the T cell receptor repertoire. Nature *547*, 94–98.

Gras, S., Chadderton, J., Del Campo, C.M., Farenc, C., Wiede, F., Josephs, T.M., Sng, X.Y.X., Mirams, M., Watson, K.A., Tiganis, T., et al. (2016). Reversed T cell receptor docking on a major histocompatibility class I complex limits involvement in the immune response. Immunity *45*, 749–760.

Green, R.S., Stone, E.L., Tenno, M., Lehtonen, E., Farquhar, M.G., and Marth, J.D. (2007). Mammalian N-glycan branching protects against innate immune self-recognition and inflammation in autoimmune disease pathogenesis. Immunity *27*, 308–320.

Greiff, V., Bhat, P., Cook, S.C., Menzel, U., Kang, W., and Reddy, S.T. (2015). A bioinformatic framework for immune repertoire diversity profiling enables detection of immunological status. Genome Med. *7*, 49.

Griffiths, P., Baraniak, I., and Reeves, M. (2015). The pathogenesis of human cytomegalovirus. J. Pathol. *235*, 288–297.

Gros, A., Parkhurst, M.R., Tran, E., Pasetto, A., Robbins, P.F., Ilyas, S., Prickett, T.D., Gartner, J.J., Crystal, J.S., Roberts, I.M., et al. (2016). Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients. Nat. Med. *22*, 433–438.

Gross, G., Waks, T., and Eshhar, Z. (1989). Expression of immunoglobulin-T-cell receptor chimeric molecules as functional receptors with antibody-type specificity. Proc. Natl. Acad. Sci. USA *86*, 10024–10028.

Gubin, M.M., Zhang, X., Schuster, H., Caron, E., Ward, J.P., Noguchi, T., Ivanova, Y., Hundal, J., Arthur, C.D., Krebber, W.-J., et al. (2014). Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. Nature *515*, 577–581.

Guo, H.C., Jardetzky, T.S., Garrett, T.P., Lane, W.S., Strominger, J.L., and Wiley, D.C. (1992). Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. Nature *360*, 364–366.

Hadrup, S.R., and Schumacher, T.N. (2010). MHC-based detection of antigen-specific CD8+ T cell responses. Cancer Immunol. Immunother. *59*, 1425–1433.

Hadrup, S.R., Toebes, M., Rodenko, B., Bakker, A.H., Egan, D.A., Ovaa, H., and Schumacher, T.N.M. (2009a). High-throughput T-cell epitope discovery through MHC peptide exchange. Methods Mol. Biol. *524*, 383–405.

Hadrup, S.R., Bakker, A.H., Shu, C.J., Andersen, R.S., van Veluw, J., Hombrink, P., Castermans, E., Thor Straten, P., Blank, C., Haanen, J.B., et al. (2009b). Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. Nat. Methods *6*, 520–526.

Hagerty, D.T., and Allen, P.M. (1995). Intramolecular mimicry. Identification and analysis of two cross-reactive T cell epitopes within a single protein. J. Immunol. *155*, 2993–3001.

Hahn, M., Nicholson, M.J., Pyrdol, J., and Wucherpfennig, K.W. (2005). Unconventional topology of self peptide-major histocompatibility complex binding by a human autoimmune T cell receptor. Nat. Immunol. *6*, 490–496.

Han, A., Glanville, J., Hansmann, L., and Davis, M.M. (2014). Linking T-cell receptor sequence to functional phenotype at the single-cell level. Nat. Biotechnol. *32*, 684–692.

Han, F., Lin, L., Warby, S.C., Faraco, J., Li, J., Dong, S.X., An, P., Zhao, L., Wang, L.H., Li, Q.Y., et al. (2011). Narcolepsy onset is seasonal and increased following the 2009 H1N1 pandemic in China. Ann. Neurol. *70*, 410–417.

Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., Saadatpour, A., Zhou, Z., Chen, H., Ye, F., et al. (2018). Mapping the Mouse Cell Atlas by Microwell-Seq. Cell *172*, 1091–1107.e17.

Harding, C.V., and Unanue, E.R. (1990). Quantitation of antigen-presenting cell MHC class II/peptide complexes necessary for T-cell stimulation. Nature *346*, 574–576.

Harris, D.T., Wang, N., Riley, T.P., Anderson, S.D., Singh, N.K., Procko, E., Baker, B.M., and Kranz, D.M. (2016). Deep Mutational Scans as a Guide to Engineering High Affinity T Cell Receptor Interactions with Peptide-bound Major Histocompatibility Complex. J. Biol. Chem. *291*, 24566–24578.

Hebeisen, M., Oberle, S.G., Presotto, D., Speiser, D.E., Zehn, D., and Rufer, N. (2013a). Molecular insights for optimizing T cell receptor specificity against cancer. Front. Immunol. *4*, 154.

Hebeisen, M., Baitsch, L., Presotto, D., Baumgaertner, P., Romero, P., Michielin, O., Speiser, D.E., and Rufer, N. (2013b). SHP-1 phosphatase activity counteracts increased T cell receptor affinity. J. Clin. Invest. *123*, 1044–1056.

Hebeisen, M., Allard, M., Gannon, P.O., Schmidt, J., Speiser, D.E., and Rufer, N. (2015). Identifying individual T cell receptors of optimal avidity for tumor antigens. Front. Immunol. *6*, 582.

Heier, M.S., Gautvik, K.M., Wannag, E., Bronder, K.H., Midtlyng, E., Kamaleri, Y., and Storsaeter, J. (2013). Incidence of narcolepsy in Norwegian children and adolescents after vaccination against H1N1 influenza A. Sleep Med. *14*, 867–871.

Hellman, L.M., Foley, K.C., Singh, N.K., Alonso, J.A., Riley, T.P., Devlin, J.R., Ayres, C.M., Keller, G.L.J., Zhang, Y., Vander Kooi, C.W., et al. (2018). Improving T cell receptor on-target specificity via structure-guided design. *In Press* Mol. Ther.

Hill, J.A., Bell, D.A., Brintnell, W., Yue, D., Wehrli, B., Jevnikar, A.M., Lee, D.M., Hueber, W., Robinson, W.H., and Cairns, E. (2008). Arthritis induced by posttranslationally modified (citrullinated) fibrinogen in DR4-IE transgenic mice. J. Exp. Med. *205*, 967–979.

Holland, C.J., MacLachlan, B.J., Bianchi, V., Hesketh, S.J., Morgan, R., Vickery, O., Bulek, A.M., Fuller, A., Godkin, A., Sewell, A.K., et al. (2018). In silico and structural analyses demonstrate that intrinsic protein motions guide T cell receptor complementarity determining region loop flexibility. Front. Immunol. *9*, 674.

Holler, P.D., Chlewicki, L.K., and Kranz, D.M. (2003). TCRs with high affinity for foreign pMHC show self-reactivity. Nat. Immunol. *4*, 55–62.

Hombrink, P., Hadrup, S.R., Bakker, A., Kester, M.G.D., Falkenburg, J.H.F., von dem Borne, P.A., Schumacher, T.N.M., and Heemskerk, M.H.M. (2011). High-throughput identification of potential minor histocompatibility antigens by MHC tetramer-based screening: feasibility and limitations. PLoS One *6*, e22523.

Hong, J., Ge, C., Jothikumar, P., Yuan, Z., Liu, B., Bai, K., Li, K., Rittase, W., Shinzawa, M., Zhang, Y., et al. (2018). A TCR mechanotransduction signaling loop induces negative selection in the thymus. Nat. Immunol. *19*, 1379–1390.

Hou, D., Chen, C., Seely, E.J., Chen, S., and Song, Y. (2016). High-Throughput Sequencing-Based Immune Repertoire Study during Infectious Disease. Front. Immunol. *7*, 336.

Howie, B., Sherwood, A.M., Berkebile, A.D., Berka, J., Emerson, R.O., Williamson, D.W., Kirsch, I., Vignali, M., Rieder, M.J., Carlson, C.S., et al. (2015). High-throughput pairing of T cell receptor α and β sequences. Sci. Transl. Med. *7*, 301ra131.

Huang, J., Zeng, X., Sigal, N., Lund, P.J., Su, L.F., Huang, H., Chien, Y., and Davis, M.M. (2016). Detection, phenotyping, and quantification of antigen-specific T cells using a peptide-MHC dodecamer. Proc. Natl. Acad. Sci. USA *113*, E1890–7.

Hughes, M.S., Yu, Y.Y.L., Dudley, M.E., Zheng, Z., Robbins, P.F., Li, Y., Wunderlich, J., Hawley, R.G., Moayeri, M., Rosenberg, S.A., et al. (2005). Transfer of a TCR gene derived from a patient with a marked antitumor response conveys highly active T-cell effector functions. Hum. Gene Ther. *16*, 457–472.

Hugo, W., Zaretsky, J.M., Sun, L., Song, C., Moreno, B.H., Hu-Lieskovan, S., Berent-Maoz, B., Pang, J., Chmielowski, B., Cherry, G., et al. (2016). Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. Cell *165*, 35–44.

Hundal, J., Carreno, B.M., Petti, A.A., Linette, G.P., Griffith, O.L., Mardis, E.R., and Griffith, M. (2016). pVAC-Seq: A genome-guided in silico approach to identifying tumor neoantigens. Genome Med. *8*, 11.

Ignatowicz, L., Kappler, J., and Marrack, P. (1996a). The repertoire of T cells shaped by a single MHC/peptide ligand. Cell *84*, 521–529.

Ignatowicz, L., Kappler, J., Parker, D.C., and Marrack, P. (1996b). The responses of mature T cells are not necessarily antagonized by their positively selecting peptide. J. Immunol. *157*, 1827–1831.

Irving, M., Zoete, V., Hebeisen, M., Schmid, D., Baumgartner, P., Guillaume, P., Romero, P., Speiser, D., Luescher, I., Rufer, N., et al. (2012). Interplay between T cell receptor binding kinetics and the level of cognate peptide presented by major histocompatibility complexes governs CD8+ T cell responsiveness. J. Biol. Chem. *287*, 23068–23078.

Jerne, N.K. (1955). The natural-selection theory of antibody formation. Proc. Natl. Acad. Sci. USA *41*, 849–857.

Jerne, N.K. (1971). The somatic generation of immune recognition. Eur. J. Immunol. *1*, 1–9.

Jia, Q., Wu, W., Wang, Y., Alexander, P.B., Sun, C., Gong, Z., Cheng, J.-N., Sun, H., Guan, Y., Xia, X., et al. (2018). Local mutational diversity drives intratumoral immune heterogeneity in non-small cell lung cancer. Nat. Commun. *9*, 5361.

Johnson, L.A., and June, C.H. (2017). Driving gene-engineered T cell immunotherapy of cancer. Cell Res. *27*, 38–58.

Johnson, L.A., Heemskerk, B., Powell, D.J., Cohen, C.J., Morgan, R.A., Dudley, M.E., Robbins, P.F., and Rosenberg, S.A. (2006). Gene transfer of tumor-reactive TCR confers both high avidity and tumor reactivity to nonreactive peripheral blood mononuclear cells and tumor-infiltrating lymphocytes. J. Immunol. *177*, 6548–6559.

Johnson, L.A., Morgan, R.A., Dudley, M.E., Cassard, L., Yang, J.C., Hughes, M.S., Kammula, U.S., Royal, R.E., Sherry, R.M., Wunderlich, J.R., et al. (2009). Gene therapy with human and mouse T-cell receptors mediates cancer regression and targets normal tissues expressing cognate antigen. Blood *114*, 535–546.

June, C.H., O'Connor, R.S., Kawalekar, O.U., Ghassemi, S., and Milone, M.C. (2018). CAR T cell immunotherapy for human cancer. Science *359*, 1361–1365.

Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J. Immunol. *199*, 3360–3368.

Jurtz, V.I., Jessen, L.E., Bentzen, A.K., Jespersen, M.C., Mahajan, S., Vita, R., Jensen, K.K., Marcatili, P., Hadrup, S.R., Peters, B., et al. (2018). NetTCR: sequence-based prediction of TCR binding to peptide-MHC complexes using convolutional neural networks. BioRxiv. doi.org/10.1101/433706

Kageyama, S., Ikeda, H., Miyahara, Y., Imai, N., Ishihara, M., Saito, K., Sugino, S., Ueda, S., Ishikawa, T., Kokura, S., et al. (2015). Adoptive Transfer of MAGE-A4 T-cell Receptor Gene-Transduced Lymphocytes in Patients with Recurrent Esophageal Cancer. Clin. Cancer Res. *21*, 2268–2277.

Kaliyaperumal, A., Mohan, C., Wu, W., and Datta, S.K. (1996). Nucleosomal peptide epitopes for nephritis-inducing T helper cells of murine lupus. J. Exp. Med. *183*, 2459–2469.

Karanikas, V., Lurquin, C., Colau, D., van Baren, N., De Smet, C., Lethé, B., Connerotte, T., Corbière, V., Demoitié, M.-A., Liénard, D., et al. (2003). Monoclonal anti-MAGE-3 CTL responses in melanoma patients displaying tumor regression after vaccination with a recombinant canarypox virus. J. Immunol. *171*, 4898–4904.

Kebschull, J.M., and Zador, A.M. (2018). Cellular barcoding: lineage tracing, screening and beyond. Nat. Methods *15*, 871–879.

Kersh, G.J., and Allen, P.M. (1996). Structural basis for T cell recognition of altered peptide ligands: a single T cell receptor can productively recognize a large continuum of related ligands. J. Exp. Med. *184*, 1259–1268.

Keskin, D.B., Anandappa, A.J., Sun, J., Tirosh, I., Mathewson, N.D., Li, S., Oliveira, G., Giobbie-Hurder, A., Felt, K., Gjini, E., et al. (2018). Neoantigen vaccine generates intratumoral T cell responses in phase Ib glioblastoma trial. Nature.

Kessels, H.W., Wolkers, M.C., van den Boom, M.D., van der Valk, M.A., and Schumacher, T.N. (2001). Immunotherapy through TCR gene transfer. Nat. Immunol. *2*, 957–961.

Kim, S.-K., Cornberg, M., Wang, X.Z., Chen, H.D., Selin, L.K., and Welsh, R.M. (2005). Private specificities of CD8 T cell responses control patterns of heterologous immunity. J. Exp. Med. *201*, 523–533.

Kim, S.-M., Bhonsle, L., Besgen, P., Nickel, J., Backes, A., Held, K., Vollmer, S., Dornmair, K., and Prinz, J.C. (2012). Analysis of the paired TCR α- and β-chains of single human T cells. PLoS One *7*, e37338.

Kirsch, I.R., Watanabe, R., O'Malley, J.T., Williamson, D.W., Scott, L.-L., Elco, C.P., Teague, J.E., Gehad, A., Lowry, E.L., LeBoeuf, N.R., et al. (2015). TCR sequencing facilitates diagnosis and identifies mature T cells as the cell of origin in CTCL. Sci. Transl. Med. *7*, 308ra158.

Kivioja, T., Vähärautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S., and Taipale, J. (2011). Counting absolute numbers of molecules using unique molecular identifiers. Nat. Methods *9*, 72–74.

Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell *161*, 1187–1201.

Klinger, M., Pepin, F., Wilkins, J., Asbury, T., Wittkop, T., Zheng, J., Moorhead, M., and Faham, M. (2015). Multiplex Identification of Antigen-Specific T Cell Receptors Using a Combination of Immune Assays and Immune Receptor Sequencing. PLoS One *10*, e0141561.

Koning, F., Thomas, R., Rossjohn, J., and Toes, R.E. (2015). Coeliac disease and rheumatoid arthritis: similar mechanisms, different antigens. Nat. Rev. Rheumatol. *11*, 450–461.

Kozono, H., White, J., Clements, J., Marrack, P., and Kappler, J. (1994). Production of soluble MHC class II proteins with covalently bound single peptides. Nature *369*, 151–154.

Kracht, M.J.L., van Lummel, M., Nikolic, T., Joosten, A.M., Laban, S., van der Slik, A.R., van Veelen, P.A., Carlotti, F., de Koning, E.J.P., Hoeben, R.C., et al. (2017). Autoimmunity against a defective ribosomal insulin gene product in type 1 diabetes. Nat. Med. *23*, 501–507.

Kumánovics, A., Takada, T., and Lindahl, K.F. (2003). Genomic organization of the mammalian MHC. Annu. Rev. Immunol. *21*, 629–657.

Kvistborg, P., Shu, C.J., Heemskerk, B., Fankhauser, M., Thrue, C.A., Toebes, M., van Rooij, N., Linnemann, C., van Buuren, M.M., Urbanus, J.H.M., et al. (2012). TIL therapy broadens the tumor-reactive CD8(+) T cell compartment in melanoma patients. Oncoimmunology *1*, 409–418.

Kwong, G.A., Radu, C.G., Hwang, K., Shu, C.J., Ma, C., Koya, R.C., Comin-Anduix, B., Hadrup, S.R., Bailey, R.C., Witte, O.N., et al. (2009). Modular nucleic acid assembled p/MHC microarrays for multiplexed sorting of antigen-specific T cells. J. Am. Chem. Soc. *131*, 9695–9703.

Landais, E., Romagnoli, P.A., Corper, A.L., Shires, J., Altman, J.D., Wilson, I.A., Garcia, K.C., and Teyton, L. (2009). New design of MHC class II tetramers to accommodate fundamental principles of antigen presentation. J. Immunol. *183*, 7949–7957.

Latorre, D., Kallweit, U., Armentani, E., Foglierini, M., Mele, F., Cassotta, A., Jovic, S., Jarrossay, D., Mathis, J., Zellini, F., et al. (2018). T cells in patients with narcolepsy target self-antigens of hypocretin neurons. Nature *562*, 63–68.

Lauring, A.S., Frydman, J., and Andino, R. (2013). The role of mutational robustness in RNA virus evolution. Nat. Rev. Microbiol. *11*, 327–336.

Liblau, R.S., Vassalli, A., Seifinejad, A., and Tafti, M. (2015). Hypocretin (orexin) biology and the pathophysiology of narcolepsy with cataplexy. Lancet Neurol. *14*, 318–328.

Liepe, J., Marino, F., Sidney, J., Jeko, A., Bunting, D.E., Sette, A., Kloetzel, P.M., Stumpf, M.P.H., Heck, A.J.R., and Mishto, M. (2016). A large fraction of HLA class I ligands are proteasome-generated spliced peptides. Science *354*, 354–358.

Lindley, S., Dayan, C.M., Bishop, A., Roep, B.O., Peakman, M., and Tree, T.I.M. (2005). Defective suppressor function in CD4(+)CD25(+) T-cells from patients with type 1 diabetes. Diabetes *54*, 92–99.

Linnemann, C., Heemskerk, B., Kvistborg, P., Kluin, R.J., Bolotin, D.A., Chen, X., Bresser, K., Nieuwland, M., Schotte, R., Michels, S., et al. (2013). High-throughput identification of antigen-specific TCRs by TCR gene capture. Nat. Med. *19*, 1534–1541.

Lissina, A., Ladell, K., Skowera, A., Clement, M., Edwards, E., Seggewiss, R., van den Berg, H.A., Gostick, E., Gallagher, K., Jones, E., et al. (2009). Protein kinase inhibitors substantially improve the physical detection of T-cells with peptide-MHC tetramers. J. Immunol. Methods *340*, 11–24.

Liu, R., Li, X., Xiao, W., and Lam, K.S. (2017a). Tumor-targeting peptides from combinatorial libraries. Adv. Drug Deliv. Rev. *110-111*, 13–37.

Liu, R., Li, X., and Lam, K.S. (2017b). Combinatorial chemistry in drug discovery. Curr. Opin. Chem. Biol. *38*, 117–126.

Luimstra, J.J., Garstka, M.A., Roex, M.C.J., Redeker, A., Janssen, G.M.C., van Veelen, P.A., Arens, R., Falkenburg, J.H.F., Neefjes, J., and Ovaa, H. (2018). A flexible MHC class I multimer loading system for large-scale detection of antigen-specific T cells. J. Exp. Med. *215*, 1493–1504.

Łuksza, M., Riaz, N., Makarov, V., Balachandran, V.P., Hellmann, M.D., Solovyov, A., Rizvi, N.A., Merghoub, T., Levine, A.J., Chan, T.A., et al. (2017). A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. Nature *551*, 517–520.

Luo, G., Ambati, A., Lin, L., Bonvalet, M., Partinen, M., Ji, X., Maecker, H.T., and Mignot, E.J.-M. (2018). Autoimmunity to hypocretin and molecular mimicry to flu in type 1 narcolepsy. Proc. Natl. Acad. Sci. USA.

Lyons, A.B., and Parish, C.R. (1994). Determination of lymphocyte division by flow cytometry. J. Immunol. Methods *171*, 131–137.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–1214.

Madi, A., Shifrut, E., Reich-Zeliger, S., Gal, H., Best, K., Ndifon, W., Chain, B., Cohen, I.R., and Friedman, N. (2014). T-cell receptor repertoires share a restricted set of public and abundant CDR3 sequences that are associated with self-related immunity. Genome Res. *24*, 1603–1612.

Maecker, H.T., and Maino, V.C. (2004). Analyzing T-cell responses to cytomegalovirus by cytokine flow cytometry. Hum. Immunol. *65*, 493–499.

Mannering, S.I., Harrison, L.C., Williamson, N.A., Morris, J.S., Thearle, D.J., Jensen, K.P., Kay, T.W.H., Rossjohn, J., Falk, B.A., Nepom, G.T., et al. (2005). The insulin A-chain epitope recognized by human T cells is posttranslationally modified. J. Exp. Med. *202*, 1191–1197.

Manning, T.C., and Kranz, D.M. (1999). Binding energetics of T-cell receptors: correlation with immunological consequences. Immunol. Today *20*, 417–422.

Mannocci, L., Zhang, Y., Scheuermann, J., Leimbacher, M., De Bellis, G., Rizzi, E., Dumelin, C., Melkko, S., and Neri, D. (2008). High-throughput sequencing allows the identification of binding molecules isolated from DNA-encoded chemical libraries. Proc. Natl. Acad. Sci. USA *105*, 17670–17675.

Mardis, E.R. (2013). Next-generation sequencing platforms. Annu. Rev. Anal. Chem. (Palo Alto, Calif.) *6*, 287–303.

Marrack, P., and Kappler, J. (1987). The T cell receptor. Science *238*, 1073–1079.

Mason, D. (1998). A very high level of crossreactivity is an essential feature of the T-cell receptor. Immunol. Today *19*, 395–404.

Matzinger, P. (1994). Tolerance, danger, and the extended family. Annu. Rev. Immunol. *12*, 991–1045.

Matzinger, P. (2002). The danger model: a renewed sense of self. Science *296*, 301–305.

McGranahan, N., Furness, A.J.S., Rosenthal, R., Ramskov, S., Lyngaa, R., Saini, S.K., Jamal-Hanjani, M., Wilson, G.A., Birkbak, N.J., Hiley, C.T., et al. (2016). Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science *351*, 1463–1469.

Meadows, L., Wang, W., den Haan, J.M., Blokland, E., Reinhardus, C., Drijfhout, J.W., Shabanowitz, J., Pierce, R., Agulnik, A.I., Bishop, C.E., et al. (1997). The HLA-A*0201-restricted H-Y antigen contains a posttranslationally modified cysteine that significantly affects T cell recognition. Immunity *6*, 273–281.

Melkko, S., Zhang, Y., Dumelin, C.E., Scheuermann, J., and Neri, D. (2007). Isolation of high-affinity trypsin inhibitors from a DNA-encoded chemical library. Angew. Chem. Int. Ed. Engl. *46*, 4671–4674.

Mendes, M.F.A., Antunes, D.A., Rigo, M.M., Sinigaglia, M., and Vieira, G.F. (2015). Improved structural method for T-cell cross-reactivity prediction. Mol. Immunol. *67*, 303–310.

Miles, J.J., McCluskey, J., Rossjohn, J., and Gras, S. (2015). Understanding the complexity and malleability of T-cell recognition. Immunol. Cell Biol. *93*, 433–441.

Miles, J.J., Tan, M.P., Dolton, G., Edwards, E.S., Galloway, S.A., Laugel, B., Clement, M., Makinde, J., Ladell, K., Matthews, K.K., et al. (2018). Peptide mimic for influenza vaccination using nonnatural combinatorial chemistry. J. Clin. Invest. *128*, 1569–1580.

Miller, N.J., Church, C.D., Fling, S.P., Kulikauskas, R., Ramchurren, N., Shinohara, M.M., Kluger, H.M., Bhatia, S., Lundgren, L., Cheever, M.A., et al. (2018). Merkel cell polyomavirus-specific immune responses in patients with Merkel cell carcinoma receiving anti-PD-1 therapy. J. Immunother. Cancer *6*, 131.

Mimitou, E., Cheng, A., Montalbano, A., Hao, S., Stoeckius, M., Legut, M., Roush, T., Herrera, A., Papalexi, E., Ouyang, Z., et al. (2018). Expanding the CITE-seq tool-kit: Detection of proteins, transcriptomes, clonotypes and CRISPR perturbations with multiplexing, in a single assay. BioRxiv. doi.org/10.1101/466466

Mohammed, F., Cobbold, M., Zarling, A.L., Salim, M., Barrett-Wilt, G.A., Shabanowitz, J., Hunt, D.F., Engelhard, V.H., and Willcox, B.E. (2008). Phosphorylation-dependent interaction between antigenic peptides and MHC class I: a molecular basis for the presentation of transformed self. Nat. Immunol. *9*, 1236–1243.

Molberg, O., Mcadam, S.N., Körner, R., Quarsten, H., Kristiansen, C., Madsen, L., Fugger, L., Scott, H., Norén, O., Roepstorff, P., et al. (1998). Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. Nat. Med. *4*, 713–717.

Morgan, R.A., Dudley, M.E., Wunderlich, J.R., Hughes, M.S., Yang, J.C., Sherry, R.M., Royal, R.E., Topalian, S.L., Kammula, U.S., Restifo, N.P., et al. (2006). Cancer regression in patients after transfer of genetically engineered lymphocytes. Science *314*, 126–129.

Morgan, R.A., Chinnasamy, N., Abate-Daga, D., Gros, A., Robbins, P.F., Zheng, Z., Dudley, M.E., Feldman, S.A., Yang, J.C., Sherry, R.M., et al. (2013). Cancer regression and neurological toxicity following anti-MAGE-A3 TCR gene therapy. J Immunother *36*, 133–151.

Motozono, C., Pearson, J.A., De Leenheer, E., Rizkallah, P.J., Beck, K., Trimby, A., Sewell, A.K., Wong, F.S., and Cole, D.K. (2015). Distortion of the Major Histocompatibility Complex Class I Binding Groove to Accommodate an Insulin-derived 10-Mer Peptide. J. Biol. Chem. *290*, 18924–18933.

Mueller, D.L. (2010). Mechanisms maintaining peripheral tolerance. Nat. Immunol. *11*, 21–27.

Naik, S.H., Perié, L., Swart, E., Gerlach, C., van Rooij, N., de Boer, R.J., and Schumacher, T.N. (2013). Diverse and heritable lineage imprinting of early haematopoietic progenitors. Nature *496*, 229–232.

Needels, M.C., Jones, D.G., Tate, E.H., Heinkel, G.L., Kochersperger, L.M., Dower, W.J., Barrett, R.W., and Gallop, M.A. (1993). Generation and screening of an oligonucleotide-encoded synthetic peptide library. Proc. Natl. Acad. Sci. USA *90*, 10700–10704.

Negro, R., Gobessi, S., Longo, P.G., He, Y., Zhang, Z.-Y., Laurenti, L., and Efremov, D.G. (2012). Overexpression of the autoimmunity-associated phosphatase PTPN22 promotes survival of antigen-stimulated CLL cells by selectively activating AKT. Blood *119*, 6278–6287.

Newell, E.W., and Davis, M.M. (2014). Beyond model antigens: high-dimensional methods for the analysis of antigen-specific T cells. Nat. Biotechnol. *32*, 149–157.

Newell, E.W., Klein, L.O., Yu, W., and Davis, M.M. (2009). Simultaneous detection of many T-cell specificities using combinatorial tetramer staining. Nat. Methods *6*, 497–499.

Newell, E.W., Sigal, N., Nair, N., Kidd, B.A., Greenberg, H.B., and Davis, M.M. (2013). Combinatorial tetramer staining and mass cytometry analysis facilitate T-cell epitope mapping and characterization. Nat. Biotechnol. *31*, 623–629.

Nghiem, P.T., Bhatia, S., Lipson, E.J., Kudchadkar, R.R., Miller, N.J., Annamalai, L., Berry, S., Chartash, E.K., Daud, A., Fling, S.P., et al. (2016). PD-1 Blockade with Pembrolizumab in Advanced Merkel-Cell Carcinoma. N. Engl. J. Med. *374*, 2542–2552.

Nie, S., Lin, S.-J., Kim, S., Welsh, R.M., and Selin, L.K. (2010). Pathological features of heterologous immunity are regulated by the private specificities of the immune repertoire. Am. J. Pathol. *176*, 2107–2112.

Nielsen, J., Brenner, S., and Janda, K.D. (1993). Synthetic methods for the implementation of encoded combinatorial chemistry. J. Am. Chem. Soc. *115*, 9812–9813.

Nielsen, M., Lundegaard, C., Blicher, T., Lamberth, K., Harndahl, M., Justesen, S., Røder, G., Peters, B., Sette, A., Lund, O., et al. (2007). NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. PLoS One *2*, e796.

Nishimura, H., Nose, M., Hiai, H., Minato, N., and Honjo, T. (1999). Development of lupus-like autoimmune diseases by disruption of the PD-1 gene encoding an ITIM motif-carrying immunoreceptor. Immunity *11*, 141–151.

Nuevolution (2015). Combinatorial chemistry in a historical perspective. NUEsFEATURE *1*, nuevolution.com/features.

Ohkura, N., Kitagawa, Y., and Sakaguchi, S. (2013). Development and maintenance of regulatory T cells. Immunity *38*, 414–423.

Ohta, R., Demachi-Okamura, A., Akatsuka, Y., Fujiwara, H., and Kuzushima, K. (2018). Improving TCR affinity on 293 T cells. J. Immunol. Methods.

Olaru, A., Patterson, D.N., Cai, H., and Livák, F. (2004). Recombination signal sequence variations and the mechanism of patterned T-cell receptor-beta locus rearrangement. Mol. Immunol. *40*, 1189–1201.

Oldstone, M.B. (1998). Molecular mimicry and immune-mediated diseases. FASEB J. *12*, 1255–1265.

Omenn, G.S., Lane, L., Lundberg, E.K., Beavis, R.C., Overall, C.M., and Deutsch, E.W. (2016). Metrics for the Human Proteome Project 2016: Progress on Identifying and Characterizing the Human Proteome, Including Post-Translational Modifications. J. Proteome Res. *15*, 3951–3960.

Ott, P.A., Hu, Z., Keskin, D.B., Shukla, S.A., Sun, J., Bozym, D.J., Zhang, W., Luoma, A., Giobbie-Hurder, A., Peter, L., et al. (2017). An immunogenic personal neoantigen vaccine for patients with melanoma. Nature *547*, 217–221.

Parham, P., and Ohta, T. (1996). Population biology of antigen presentation by MHC class I molecules. Science *272*, 67–74.

Parkhurst, M.R., Yang, J.C., Langan, R.C., Dudley, M.E., Nathan, D.-A.N., Feldman, S.A., Davis, J.L., Morgan, R.A., Merino, M.J., Sherry, R.M., et al. (2011). T cells targeting carcinoembryonic antigen can mediate regression of metastatic colorectal cancer but induce severe transient colitis. Mol. Ther. *19*, 620–626.

Partinen, M., Saarenpää-Heikkilä, O., Ilveskoski, I., Hublin, C., Linna, M., Olsén, P., Nokelainen, P., Alén, R., Wallden, T., Espo, M., et al. (2012). Increased incidence and clinical picture of childhood narcolepsy following the 2009 H1N1 pandemic vaccination campaign in Finland. PLoS One *7*, e33723.

Partinen, M., Kornum, B.R., Plazzi, G., Jennum, P., Julkunen, I., and Vaarala, O. (2014). Narcolepsy as an autoimmune disease: the role of H1N1 infection and vaccination. Lancet Neurol. *13*, 600–613.

Pedersen, N.W., Holm, A., Kristensen, N.P., Anne-Mette Bjerregaard, Bentzen, A.K., Marquard, A.M., Tamhane, T., Burgdorf, K.S., Ullum, H., Knudsen, S., et al. (2019). Broad presence of autoreactive CD8 T cells against 1 narcolepsy type 1-relevant proteins. Nat. Commun.

Pender, M.P., Csurhes, P.A., Burrows, J.M., and Burrows, S.R. (2017). Defective T-cell control of Epstein-Barr virus infection in multiple sclerosis. Clin. Transl. Immunology *6*, e126.

Perfetto, S.P., Chattopadhyay, P.K., and Roederer, M. (2004). Seventeen-colour flow cytometry: unravelling the immune system. Nat. Rev. Immunol. *4*, 648–655.

Petersen, J., Wurzbacher, S.J., Williamson, N.A., Ramarathinam, S.H., Reid, H.H., Nair, A.K.N., Zhao, A.Y., Nastovska, R., Rudge, G., Rossjohn, J., et al. (2009). Phosphorylated self-peptides alter human leukocyte antigen class I-restricted antigen presentation and generate tumor-specific epitopes. Proc. Natl. Acad. Sci. USA *106*, 2776–2781.

Peterson, V.M., Zhang, K.X., Kumar, N., Wong, J., Li, L., Wilson, D.C., Moore, R., McClanahan, T.K., Sadekova, S., and Klappenbach, J.A. (2017). Multiplexed quantification of proteins and transcripts in single cells. Nat. Biotechnol. *35*, 936–939.

Picelli, S., Björklund, A.K., Reinius, B., Sagasser, S., Winberg, G., and Sandberg, R. (2014). Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. Genome Res. *24*, 2033–2040.

Pierce, B.G., Hellman, L.M., Hossain, M., Singh, N.K., Vander Kooi, C.W., Weng, Z., and Baker, B.M. (2014). Computational design of the affinity and specificity of a therapeutic T cell receptor. PLoS Comput. Biol. *10*, e1003478.

Pierini, F., and Lenz, T.L. (2018). Divergent allele advantage at human MHC genes: signatures of past and ongoing selection. Mol. Biol. Evol.

Pymm, P., Illing, P.T., Ramarathinam, S.H., O'Connor, G.M., Hughes, V.A., Hitchen, C., Price, D.A., Ho, B.K., McVicar, D.W., Brooks, A.G., et al. (2017). MHC-I peptides get out of the groove and enable a novel mechanism of HIV-1 escape. Nat. Struct. Mol. Biol. *24*, 387–394.

Quigley, M.F., Greenaway, H.Y., Venturi, V., Lindsay, R., Quinn, K.M., Seder, R.A., Douek, D.C., Davenport, M.P., and Price, D.A. (2010). Convergent recombination shapes the clonotypic landscape of the naive T-cell repertoire. Proc. Natl. Acad. Sci. USA *107*, 19414–19419.

Rajasagi, M., Shukla, S.A., Fritsch, E.F., Keskin, D.B., DeLuca, D., Carmona, E., Zhang, W., Sougnez, C., Cibulskis, K., Sidney, J., et al. (2014). Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. Blood *124*, 453–462.

Ramarathinam, S.H., Gras, S., Alcantara, S., Yeung, A.W.S., Mifsud, N.A., Sonza, S., Illing, P.T., Glaros, E.N., Center, R.J., Thomas, S.R., et al. (2018). Identification of Native and Posttranslationally Modified HLA-B*57:01-Restricted HIV Envelope Derived Epitopes Using Immunoproteomics. Proteomics *18*, e1700253.

Rangarajan, S., and Mariuzza, R.A. (2014). T cell receptor bias for MHC: co-evolution or co-receptors? Cell Mol. Life Sci. *71*, 3059–3068.

Rapoport, A.P., Stadtmauer, E.A., Binder-Scholl, G.K., Goloubeva, O., Vogl, D.T., Lacey, S.F., Badros, A.Z., Garfall, A., Weiss, B., Finklestein, J., et al. (2015). NY-ESO-1-specific TCR-engineered T cells mediate sustained antigen-specific antitumor effects in myeloma. Nat. Med. *21*, 914–921.

Raposo, B., Merky, P., Lundqvist, C., Yamada, H., Urbonaviciute, V., Niaudet, C., Viljanen, J., Kihlberg, J., Kyewski, B., Ekwall, O., et al. (2018). T cells specific for post-translational modifications escape intrathymic tolerance induction. Nat. Commun. *9*, 353.

Reinherz, E.L., Tan, K., Tang, L., Kern, P., Liu, J., Xiong, Y., Hussey, R.E., Smolyar, A., Hare, B., Zhang, R., et al. (1999). The crystal structure of a T cell receptor in complex with peptide and MHC class II. Science *286*, 1913–1921.

Reiser, J.B., Grégoire, C., Darnault, C., Mosser, T., Guimezanes, A., Schmitt-Verhulst, A.M., Fontecilla-Camps, J.C., Mazza, G., Malissen, B., and Housset, D. (2002). A T cell receptor CDR3beta loop undergoes conformational changes of unprecedented magnitude upon binding to a peptide/MHC class I complex. Immunity *16*, 345–354.

Reiser, J.-B., Darnault, C., Grégoire, C., Mosser, T., Mazza, G., Kearney, A., van der Merwe, P.A., Fontecilla-Camps, J.C., Housset, D., and Malissen, B. (2003). CDR3 loop flexibility contributes to the degeneracy of TCR recognition. Nat. Immunol. *4*, 241–247.

Remesh, S.G., Andreatta, M., Ying, G., Kaever, T., Nielsen, M., McMurtrey, C., Hildebrand, W., Peters, B., and Zajonc, D.M. (2017). Unconventional Peptide Presentation by Major Histocompatibility Complex (MHC) Class I Allele HLA-A*02:01: BREAKING CONFINEMENT. J. Biol. Chem. *292*, 5262–5270.

Ribas, A., and Wolchok, J.D. (2018). Cancer immunotherapy using checkpoint blockade. Science *359*, 1350–1355.

Riley, T.P., and Baker, B.M. (2018). The intersection of affinity and specificity in the development and optimization of T cell receptor based therapeutics. Semin. Cell Dev. Biol. *84*, 30–41.

Riley, T.P., Hellman, L.M., Gee, M.H., Mendoza, J.L., Alonso, J.A., Foley, K.C., Nishimura, M.I., Vander Kooi, C.W., Garcia, K.C., and Baker, B.M. (2018). T cell receptor cross-reactivity expanded by dramatic peptide-MHC adaptability. Nat. Chem. Biol. *14*, 934–942.

Rist, M.J., Hibbert, K.M., Croft, N.P., Smith, C., Neller, M.A., Burrows, J.M., Miles, J.J., Purcell, A.W., Rossjohn, J., Gras, S., et al. (2015). T Cell Cross-Reactivity between a Highly

Immunogenic EBV Epitope and a Self-Peptide Naturally Presented by HLA-B*18:01+ Cells. J. Immunol. *194*, 4668–4675.

Rius, C., Attaf, M., Tungatt, K., Bianchi, V., Legut, M., Bovay, A., Donia, M., Thor Straten, P., Peakman, M., Svane, I.M., et al. (2018). Peptide-MHC Class I Tetramers Can Fail To Detect Relevant Functional T Cell Clonotypes and Underestimate Antigen-Reactive T Cell Populations. J. Immunol. *200*, 2263–2279.

Rizvi, N.A., Hellmann, M.D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J.J., Lee, W., Yuan, J., Wong, P., Ho, T.S., et al. (2015). Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science *348*, 124–128.

Robbins, P.F., Li, Y.F., El-Gamil, M., Zhao, Y., Wargo, J.A., Zheng, Z., Xu, H., Morgan, R.A., Feldman, S.A., Johnson, L.A., et al. (2008). Single and dual amino acid substitutions in TCR CDRs can enhance antigen-specific T cell functions. J. Immunol. *180*, 6116–6131.

Robbins, P.F., Morgan, R.A., Feldman, S.A., Yang, J.C., Sherry, R.M., Dudley, M.E., Wunderlich, J.R., Nahvi, A.V., Helman, L.J., Mackall, C.L., et al. (2011). 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.

Robbins, P.F., Kassim, S.H., Tran, T.L.N., Crystal, J.S., Morgan, R.A., Feldman, S.A., Yang, J.C., Dudley, M.E., Wunderlich, J.R., Sherry, R.M., et al. (2015). A pilot trial using lymphocytes genetically engineered with an NY-ESO-1-reactive T-cell receptor: long-term follow-up and correlates with response. Clin. Cancer Res. *21*, 1019–1027.

Robins, H.S., Campregher, P.V., Srivastava, S.K., Wacher, A., Turtle, C.J., Kahsai, O., Riddell, S.R., Warren, E.H., and Carlson, C.S. (2009). Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. Blood *114*, 4099–4107.

Robinson, J., Waller, M.J., Parham, P., de Groot, N., Bontrop, R., Kennedy, L.J., Stoehr, P., and Marsh, S.G.E. (2003). IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. Nucleic Acids Res. *31*, 311–314.

Robinson, J., Soormally, A.R., Hayhurst, J.D., and Marsh, S.G.E. (2016). The IPD-IMGT/HLA Database - New developments in reporting HLA variation. Hum. Immunol. *77*, 233–237.

Rodenko, B., Toebes, M., Hadrup, S.R., van Esch, W.J.E., Molenaar, A.M., Schumacher, T.N.M., and Ovaa, H. (2006). Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. Nat. Protoc. *1*, 1120–1132.

Rodig, S.J., Cheng, J., Wardzala, J., DoRosario, A., Scanlon, J.J., Laga, A.C., Martinez-Fernandez, A., Barletta, J.A., Bellizzi, A.M., Sadasivam, S., et al. (2012). Improved detection

suggests all Merkel cell carcinomas harbor Merkel polyomavirus. J. Clin. Invest. *122*, 4645–4653.

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

Rose, N.R. (2001). Infection, mimics, and autoimmune disease. J. Clin. Invest. *107*, 943–944.

Rose, N.R. (2017). Negative selection, epitope mimicry and autoimmunity. Curr. Opin. Immunol. *49*, 51–55.

Rosenberg, S.A. (2011). Cell transfer immunotherapy for metastatic solid cancer--what clinicians need to know. Nat. Rev. Clin. Oncol. *8*, 577–585.

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

Rosenberg, A.B., Roco, C., Muscat, R.A., Kuchina, A., Sample, P., Yao, Z., Graybuck, L.T., Peeler, D.J., Mukherjee, S., Chen, W., et al. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science *360*, 176–182.

Rosenberg, S.A., Packard, B.S., Aebersold, P.M., Solomon, D., Topalian, S.L., Toy, S.T., Simon, P., Lotze, M.T., Yang, J.C., and Seipp, C.A. (1988). Use of tumor-infiltrating lymphocytes and interleukin-2 in the immunotherapy of patients with metastatic melanoma. A preliminary report. N. Engl. J. Med. *319*, 1676–1680.

Rosenberg, S.A., Aebersold, P., Cornetta, K., Kasid, A., Morgan, R.A., Moen, R., Karson, E.M., Lotze, M.T., Yang, J.C., and Topalian, S.L. (1990). 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.

Rossjohn, J., Gras, S., Miles, J.J., Turner, S.J., Godfrey, D.I., and McCluskey, J. (2015). T cell antigen receptor recognition of antigen-presenting molecules. Annu. Rev. Immunol. *33*, 169–200.

Rubinsteyn, A., Kodysh, J., Hodes, I., Mondet, S., Aksoy, B.A., Finnigan, J.P., Bhardwaj, N., and Hammerbacher, J. (2017). Computational Pipeline for the PGV-001 Neoantigen Vaccine Trial. Front. Immunol. *8*, 1807.

Rudensky AYu, Preston-Hurlburt, P., Hong, S.C., Barlow, A., and Janeway, C.A. (1991). Sequence analysis of peptides bound to MHC class II molecules. Nature *353*, 622–627.

Sahin, U., Derhovanessian, E., Miller, M., Kloke, B.-P., Simon, P., Löwer, M., Bukur, V., Tadmor, A.D., Luxemburger, U., Schrörs, B., et al. (2017). Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. Nature *547*, 222–226.

Saini, S.K., Tamhane, T., Anjanappa, R., Saikia, A., Ramskov, S., Donia, M., Svane, I.M., Jakobsen, S.N., Garcia-Alai, M., Zacharias, M., et al. (2018). Empty peptide-receptive MHC class I moleculesfor efficient detection of antigen-specific T cells. Science Immunology.

Salou, M., Nicol, B., Garcia, A., and Laplaud, D.-A. (2015). Involvement of CD8(+) T cells in multiple sclerosis. Front. Immunol. *6*, 604.

Sarkanen, T., Alakuijala, A., Julkunen, I., and Partinen, M. (2018). Narcolepsy Associated with Pandemrix Vaccine. Curr Neurol Neurosci Rep *18*, 43.

Scally, S.W., Petersen, J., Law, S.C., Dudek, N.L., Nel, H.J., Loh, K.L., Wijeyewickrema, L.C., Eckle, S.B.G., van Heemst, J., Pike, R.N., et al. (2013). A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. J. Exp. Med. *210*, 2569–2582.

Schatz, D.G. (2004). Antigen receptor genes and the evolution of a recombinase. Semin. Immunol. *16*, 245–256.

Schaubert, K.L., Price, D.A., Salkowitz, J.R., Sewell, A.K., Sidney, J., Asher, T.E., Blondelle, S.E., Adams, S., Marincola, F.M., Joseph, A., et al. (2010). Generation of robust CD8+ T-cell responses against subdominant epitopes in conserved regions of HIV-1 by repertoire mining with mimotopes. Eur. J. Immunol. *40*, 1950–1962.

Scheuermann, J., Dumelin, C.E., Melkko, S., Zhang, Y., Mannocci, L., Jaggi, M., Sobek, J., and Neri, D. (2008). DNA-encoded chemical libraries for the discovery of MMP-3 inhibitors. Bioconjug. Chem. *19*, 778–785.

Schiller, J.T., and Müller, M. (2015). Next generation prophylactic human papillomavirus vaccines. Lancet Oncol. *16*, e217–25.

Schober, K., Buchholz, V.R., and Busch, D.H. (2018). TCR repertoire evolution during maintenance of CMV-specific T-cell populations. Immunol. Rev. *283*, 113–128.

Scott, J.K., and Smith, G.P. (1990). Searching for peptide ligands with an epitope library. Science *249*, 386–390.

Sedgwick, J.D., and Holt, P.G. (1983). A solid-phase immunoenzymatic technique for the enumeration of specific antibody-secreting cells. J. Immunol. Methods *57*, 301–309.

Selin, L.K., Nahill, S.R., and Welsh, R.M. (1994). Cross-reactivities in memory cytotoxic T lymphocyte recognition of heterologous viruses. J. Exp. Med. *179*, 1933–1943.

Selin, L.K., Vergilis, K., Welsh, R.M., and Nahill, S.R. (1996). Reduction of otherwise remarkably stable virus-specific cytotoxic T lymphocyte memory by heterologous viral infections. J. Exp. Med. *183*, 2489–2499.

Selin, L.K., Lin, M.Y., Kraemer, K.A., Pardoll, D.M., Schneck, J.P., Varga, S.M., Santolucito, P.A., Pinto, A.K., and Welsh, R.M. (1999). Attrition of T cell memory: selective loss of LCMV epitope-specific memory CD8 T cells following infections with heterologous viruses. Immunity *11*, 733–742.

Setia, M.S., Steinmaus, C., Ho, C.S., and Rutherford, G.W. (2006). The role of BCG in prevention of leprosy: a meta-analysis. Lancet Infect. Dis. *6*, 162–170.

Sewell, A.K. (2012). Why must T cells be cross-reactive? Nat. Rev. Immunol. *12*, 669–677.

Shahi, P., Kim, S.C., Haliburton, J.R., Gartner, Z.J., and Abate, A.R. (2017). Abseq: Ultrahigh-throughput single cell protein profiling with droplet microfluidic barcoding. Sci. Rep. *7*, 44447.

Shann, F., Nohynek, H., Scott, J.A., Hesseling, A., Flanagan, K.L., and Working Group on Nonspecific Effects of Vaccines (2010). Randomized trials to study the nonspecific effects of vaccines in children in low-income countries. Pediatr. Infect. Dis. J. *29*, 457–461.

Sherwood, A.M., Emerson, R.O., Scherer, D., Habermann, N., Buck, K., Staffa, J., Desmarais, C., Halama, N., Jaeger, D., Schirmacher, P., et al. (2013). Tumor-infiltrating lymphocytes in colorectal tumors display a diversity of T cell receptor sequences that differ from the T cells in adjacent mucosal tissue. Cancer Immunol. Immunother. *62*, 1453–1461.

Shuda, M., Feng, H., Kwun, H.J., Rosen, S.T., Gjoerup, O., Moore, P.S., and Chang, Y. (2008). T antigen mutations are a human tumor-specific signature for Merkel cell polyomavirus. Proc. Natl. Acad. Sci. USA *105*, 16272–16277.

De Simone, M., Rossetti, G., and Pagani, M. (2018). Single cell T cell receptor sequencing: techniques and future challenges. Front. Immunol. *9*, 1638.

Simoni, Y., Becht, E., Fehlings, M., Loh, C.Y., Koo, S.-L., Teng, K.W.W., Yeong, J.P.S., Nahar, R., Zhang, T., Kared, H., et al. (2018). Bystander CD8+ T cells are abundant and phenotypically distinct in human tumour infiltrates. Nature *557*, 575–579.

Singh, N.K., Riley, T.P., Baker, S.C.B., Borrman, T., Weng, Z., and Baker, B.M. (2017). Emerging concepts in TCR specificity: rationalizing and (maybe) predicting outcomes. J. Immunol. *199*, 2203–2213.

Singh, S.K., Larsson, M., Schön, T., Stendahl, O., and Blomgran, R. (2018). HIV Interferes with the Dendritic Cell-T Cell Axis of Macrophage Activation by Shifting Mycobacterium tuberculosis-Specific CD4 T Cells into a Dysfunctional Phenotype. J. Immunol.

Slansky, J.E., and Jordan, K.R. (2010). The Goldilocks model for TCR-too much attraction might not be best for vaccine design. PLoS Biol. *8*.

Snyder, A., Makarov, V., Merghoub, T., Yuan, J., Zaretsky, J.M., Desrichard, A., Walsh, L.A., Postow, M.A., Wong, P., Ho, T.S., et al. (2014). Genetic basis for clinical response to CTLA-4 blockade in melanoma. N. Engl. J. Med. *371*, 2189–2199.

Soen, Y., Chen, D.S., Kraft, D.L., Davis, M.M., and Brown, P.O. (2003). Detection and characterization of cellular immune responses using peptide-MHC microarrays. PLoS Biol. *1*, E65.

Song, I., Gil, A., Mishra, R., Ghersi, D., Selin, L.K., and Stern, L.J. (2017). Broad TCR repertoire and diverse structural solutions for recognition of an immunodominant CD8+ T cell epitope. Nat. Struct. Mol. Biol. *24*, 395–406.

Starr, T.K., Daniels, M.A., Lucido, M.M., Jameson, S.C., and Hogquist, K.A. (2003a). Thymocyte sensitivity and supramolecular activation cluster formation are developmentally regulated: a partial role for sialylation. J. Immunol. *171*, 4512–4520.

Starr, T.K., Jameson, S.C., and Hogquist, K.A. (2003b). Positive and negative selection of T cells. Annu. Rev. Immunol. *21*, 139–176.

Stewart, A.J., and Devlin, P.M. (2006). The history of the smallpox vaccine. J. Infect. *52*, 329–334.

Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. Nat. Methods *14*, 865–868.

Stone, J.D., and Kranz, D.M. (2013). Role of T cell receptor affinity in the efficacy and specificity of adoptive T cell therapies. Front. Immunol. *4*, 244.

Stone, J.D., Demkowicz, W.E., and Stern, L.J. (2005). HLA-restricted epitope identification and detection of functional T cell responses by using MHC-peptide and costimulatory microarrays. Proc. Natl. Acad. Sci. USA *102*, 3744–3749.

Strønen, E., Toebes, M., Kelderman, S., van Buuren, M.M., Yang, W., van Rooij, N., Donia, M., Böschen, M.-L., Lund-Johansen, F., Olweus, J., et al. (2016). Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. Science *352*, 1337–1341.

Su, L.F., Kidd, B.A., Han, A., Kotzin, J.J., and Davis, M.M. (2013). Virus-specific CD4(+) memory-phenotype T cells are abundant in unexposed adults. Immunity *38*, 373–383.

Suárez-Fueyo, A., Bradley, S.J., and Tsokos, G.C. (2016). T cells in Systemic Lupus Erythematosus. Curr. Opin. Immunol. *43*, 32–38.

Sylwester, A.W., Mitchell, B.L., Edgar, J.B., Taormina, C., Pelte, C., Ruchti, F., Sleath, P.R., Grabstein, K.H., Hosken, N.A., Kern, F., et al. (2005). Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects. J. Exp. Med. *202*, 673–685.

Tahara, H., Fujio, K., Araki, Y., Setoguchi, K., Misaki, Y., Kitamura, T., and Yamamoto, K. (2003). Reconstitution of CD8+ T cells by retroviral transfer of the TCR alpha beta-chain genes isolated from a clonally expanded P815-infiltrating lymphocyte. J. Immunol. *171*, 2154–2160.

Tan, M.P., Dolton, G.M., Gerry, A.B., Brewer, J.E., Bennett, A.D., Pumphrey, N.J., Jakobsen, B.K., and Sewell, A.K. (2017). Human leucocyte antigen class I-redirected anti-tumour CD4+ T cells require a higher T cell receptor binding affinity for optimal activity than CD8+ T cells. Clin. Exp. Immunol. *187*, 124–137.

Tivol, E.A., Borriello, F., Schweitzer, A.N., Lynch, W.P., Bluestone, J.A., and Sharpe, A.H. (1995). Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. Immunity *3*, 541–547.

Toebes, M., Coccoris, M., Bins, A., Rodenko, B., Gomez, R., Nieuwkoop, N.J., van de Kasteele, W., Rimmelzwaan, G.F., Haanen, J.B.A.G., Ovaa, H., et al. (2006). Design and use of conditional MHC class I ligands. Nat. Med. *12*, 246–251.

Tonegawa, S. (1983). Somatic generation of antibody diversity. Nature *302*, 575–581.

Trowsdale, J., and Knight, J.C. (2013). Major histocompatibility complex genomics and human disease. Annu. Rev. Genomics. Hum. Genet. *14*, 301–323.

Tungatt, K., Bianchi, V., Crowther, M.D., Powell, W.E., Schauenburg, A.J., Trimby, A., Donia, M., Miles, J.J., Holland, C.J., Cole, D.K., et al. (2015). Antibody stabilization of peptide-MHC multimers reveals functional T cells bearing extremely low-affinity TCRs. J. Immunol. *194*, 463–474.

Uchtenhagen, H., Rims, C., Blahnik, G., Chow, I.-T., Kwok, W.W., Buckner, J.H., and James, E.A. (2016). Efficient ex vivo analysis of CD4+ T-cell responses using combinatorial HLA class II tetramer staining. Nat. Commun. *7*, 12614.

Vafiadis, P., Bennett, S.T., Colle, E., Grabs, R., Goodyer, C.G., and Polychronakos, C. (1996). Imprinted and genotype-specific expression of genes at the IDDM2 locus in pancreas and leucocytes. J. Autoimmun. *9*, 397–403.

Vafiadis, P., Bennett, S.T., Todd, J.A., Nadeau, J., Grabs, R., Goodyer, C.G., Wickramasinghe, S., Colle, E., and Polychronakos, C. (1997). Insulin expression in human thymus is modulated by INS VNTR alleles at the IDDM2 locus. Nat. Genet. *15*, 289–292.

Valitutti, S., Müller, S., Salio, M., and Lanzavecchia, A. (1997). Degradation of T cell receptor (TCR)-CD3-zeta complexes after antigenic stimulation. J. Exp. Med. *185*, 1859–1864.

van den Berg, J.H., Gomez-Eerland, R., van de Wiel, B., Hulshoff, L., van den Broek, D., Bins, A., Tan, H.L., Harper, J.V., Hassan, N.J., Jakobsen, B.K., et al. (2015). Case Report of a Fatal Serious Adverse Event Upon Administration of T Cells Transduced With a MART-1-specific T-cell Receptor. Mol. Ther. *23*, 1541–1550.

van Rooij, N., van Buuren, M.M., Philips, D., Velds, A., Toebes, M., Heemskerk, B., van Dijk, L.J.A., Behjati, S., Hilkmann, H., El Atmioui, D., et al. (2013). Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma. J. Clin. Oncol. *31*, e439–42.

Vassalli, A., Li, S., and Tafti, M. (2015). Comment on "Antibodies to influenza nucleoprotein cross-react with human hypocretin receptor 2". Sci. Transl. Med. *7*, 314le2.

Veiga-Fernandes, H., Walter, U., Bourgeois, C., McLean, A., and Rocha, B. (2000). Response of naïve and memory CD8+ T cells to antigen stimulation in vivo. Nat. Immunol. *1*, 47–53.

Verdegaal, E.M.E., de Miranda, N.F.C.C., Visser, M., Harryvan, T., van Buuren, M.M., Andersen, R.S., Hadrup, S.R., van der Minne, C.E., Schotte, R., Spits, H., et al. (2016). Neoantigen landscape dynamics during human melanoma-T cell interactions. Nature *536*, 91–95.

Viglietta, V., Baecher-Allan, C., Weiner, H.L., and Hafler, D.A. (2004). Loss of functional suppression by CD4+CD25+ regulatory T cells in patients with multiple sclerosis. J. Exp. Med. *199*, 971–979.

Vollers, S.S., and Stern, L.J. (2008). Class II major histocompatibility complex tetramer staining: progress, problems, and prospects. Immunology *123*, 305–313.

von Essen, M., Bonefeld, C.M., Siersma, V., Rasmussen, A.B., Lauritsen, J.P.H., Nielsen, B.L., and Geisler, C. (2004). Constitutive and ligand-induced TCR degradation. J. Immunol. *173*, 384–393.

Waterhouse, P., Penninger, J.M., Timms, E., Wakeham, A., Shahinian, A., Lee, K.P., Thompson, C.B., Griesser, H., and Mak, T.W. (1995). Lymphoproliferative disorders with early lethality in mice deficient in Ctla-4. Science *270*, 985–988.

Welsh, R.M., and Selin, L.K. (2002). No one is naive: the significance of heterologous T-cell immunity. Nat. Rev. Immunol. *2*, 417–426.

Welsh, R.M., Che, J.W., Brehm, M.A., and Selin, L.K. (2010). Heterologous immunity between viruses. Immunol. Rev. *235*, 244–266.

Wick, D.A., Webb, J.R., Nielsen, J.S., Martin, S.D., Kroeger, D.R., Milne, K., Castellarin, M., Twumasi-Boateng, K., Watson, P.H., Holt, R.A., et al. (2014). Surveillance of the tumor mutanome by T cells during progression from primary to recurrent ovarian cancer. Clin. Cancer Res. *20*, 1125–1134.

Wlodarczyk, M.F., Kraft, A.R., Chen, H.D., Kenney, L.L., and Selin, L.K. (2013). Anti-IFN-γ and peptide-tolerization therapies inhibit acute lung injury induced by cross-reactive influenza A-specific memory T cells. J. Immunol. *190*, 2736–2746.

Wood, M.A., Paralkar, M., Paralkar, M.P., Nguyen, A., Struck, A.J., Ellrott, K., Margolin, A., Nellore, A., and Thompson, R.F. (2018). Population-level distribution and putative immunogenicity of cancer neoepitopes. BMC Cancer *18*, 414.

Wooldridge, L., Laugel, B., Ekeruche, J., Clement, M., van den Berg, H.A., Price, D.A., and Sewell, A.K. (2010). CD8 controls T cell cross-reactivity. J. Immunol. *185*, 4625–4632.

Wooldridge, L., Ekeruche-Makinde, J., van den Berg, H.A., Skowera, A., Miles, J.J., Tan, M.P., Dolton, G., Clement, M., Llewellyn-Lacey, S., Price, D.A., et al. (2012). A single autoimmune T cell receptor recognizes more than a million different peptides. J. Biol. Chem. *287*, 1168–1177.

Wucherpfennig, K.W., and Strominger, J.L. (1995). Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. Cell *80*, 695–705.

Wucherpfennig, K.W., Allen, P.M., Celada, F., Cohen, I.R., De Boer, R., Garcia, K.C., Goldstein, B., Greenspan, R., Hafler, D., Hodgkin, P., et al. (2007). Polyspecificity of T cell and B cell receptor recognition. Semin. Immunol. *19*, 216–224.

Xie, J., Huppa, J.B., Newell, E.W., Huang, J., Ebert, P.J.R., Li, Q.-J., and Davis, M.M. (2012). Photocrosslinkable pMHC monomers stain T cells specifically and cause ligand-bound TCRs to be “preferentially” transported to the cSMAC. Nat. Immunol. *13*, 674–680.

Xu, Q., Schlabach, M.R., Hannon, G.J., and Elledge, S.J. (2009). Design of 240,000 orthogonal 25mer DNA barcode probes. Proc. Natl. Acad. Sci. USA *106*, 2289–2294.

Yadav, M., Jhunjhunwala, S., Phung, Q.T., Lupardus, P., Tanguay, J., Bumbaca, S., Franci, C., Cheung, T.K., Fritsche, J., Weinschenk, T., et al. (2014). Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. Nature *515*, 572–576.

Yeo, L., Woodwyk, A., Sood, S., Lorenc, A., Eichmann, M., Pujol-Autonell, I., Melchiotti, R., Skowera, A., Fidanis, E., Dolton, G.M., et al. (2018). Autoreactive T effector memory differentiation mirrors β cell function in type 1 diabetes. J. Clin. Invest. *128*, 3460–3474.

Yewdell, J.W., and Bennink, J.R. (1999). Mechanisms of viral interference with MHC class I antigen processing and presentation. Annu. Rev. Cell Dev. Biol. *15*, 579–606.

Yin, Y., Li, Y., Kerzic, M.C., Martin, R., and Mariuzza, R.A. (2011). Structure of a TCR with high affinity for self-antigen reveals basis for escape from negative selection. EMBO J. *30*, 1137–1148.

Yu, W., Jiang, N., Ebert, P.J.R., Kidd, B.A., Müller, S., Lund, P.J., Juang, J., Adachi, K., Tse, T., Birnbaum, M.E., et al. (2015). Clonal Deletion Prunes but Does Not Eliminate Self-Specific αβ CD8(+) T Lymphocytes. Immunity *42*, 929–941.

Zacharakis, N., Chinnasamy, H., Black, M., Xu, H., Lu, Y.-C., Zheng, Z., Pasetto, A., Langhan, M., Shelton, T., Prickett, T., et al. (2018). Immune recognition of somatic mutations leading to complete durable regression in metastatic breast cancer. Nat. Med. *24*, 724–730.

Zehn, D., and Bevan, M.J. (2006). T cells with low avidity for a tissue-restricted antigen routinely evade central and peripheral tolerance and cause autoimmunity. Immunity *25*, 261–270.

Zhang, L., Yu, X., Zheng, L., Zhang, Y., Li, Y., Fang, Q., Gao, R., Kang, B., Zhang, Q., Huang, J.Y., et al. (2018a). Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. Nature *564*, 268–272.

Zhang, S.-Q., Ma, K.-Y., Schonnesen, A.A., Zhang, M., He, C., Sun, E., Williams, C.M., Jia, W., and Jiang, N. (2018b). High-throughput determination of the antigen specificities of T cell receptors in single cells. Nat. Biotechnol. *36*, 1156–1159.

Zhao, Z.S., Granucci, F., Yeh, L., Schaffer, P.A., and Cantor, H. (1998). Molecular mimicry by herpes simplex virus-type 1: autoimmune disease after viral infection. Science *279*, 1344–1347.

Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. *8*, 14049.

Zhong, S., Malecek, K., Johnson, L.A., Yu, Z., Vega-Saenz de Miera, E., Darvishian, F., McGary, K., Huang, K., Boyer, J., Corse, E., et al. (2013). T-cell receptor affinity and avidity defines antitumor response and autoimmunity in T-cell immunotherapy. Proc. Natl. Acad. Sci. USA *110*, 6973–6978.

Zinkernagel, R.M., and Doherty, P.C. (1974). Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. Nature *248*, 701–702.

Zoete, V., Irving, M., Ferber, M., Cuendet, M.A., and Michielin, O. (2013). Structure-Based, Rational Design of T Cell Receptors. Front. Immunol. *4*, 268.

TEXT FOR WWW-VERSION (without papers)

In this online version of the thesis, paper I-V are not included but can be obtained from electronic article databases or on request from:

DTU Health Tech
Technical University of Denmark
Kemitorvet, Building 202
2800 Kongens Lyngby
Denmark

healthtech-info@dtu.dk

Online papers can be obtained through the following links:
Paper I: https://www.nature.com/articles/nbt.3662
Paper II: https://link.springer.com/article/10.1007%2Fs00262-017-1971-5
Paper III: https://www.nature.com/articles/nbt.4303

Technical University of Denmark
Department of Health Technology

Kemitorvet, building 202 2800 Kongens Lyngby

www.healthtech.dtu.dk

Danmarks Tekniske Universitet

DTU Health Tech
Department of Health Technology
Ørsteds Plads, Building 345C
DK-2800 Kgs. Lyngby
Denmark

Email: healthtech-info@dtu.dk
www.healthtech.dtu.dk