

**Dissertation**  
submitted to the  
**Combined Faculties for the Natural Sciences and for**  
**Mathematics**  
**of the Ruperto-Carola University of Heidelberg, Germany**  
for the degree of  
**Doctor of Natural Sciences**

presented by  
M.Sc. Zeynep Koşaloğlu  
born in Göppingen, Germany

Oral examination: June 1st, 2016

# Combining Immunomics and Genomics for the Analysis of the Microenvironment of Colorectal Cancer Liver Metastases

Referees:

Prof. Dr. Benedikt Brors

Prof. Dr. Dirk Jäger

Für meine Eltern,  
Kalender und Saliha Koşalođlu.

**”Acquire knowledge and impart it to the people.”**  
*Prophet Muhammad*

# Erklärung

Hiermit versichere ich, dass ich diese Arbeit selbstständig angefertigt, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und alle wörtlich oder sinngemäß übernommenen Textstellen als solche kenntlich gemacht habe.

Heidelberg, den 07. April 2016

# Acknowledgements

First of all, I want to thank Prof. Dr. Dirk Jäger for funding my work and providing me the opportunity to work in a very exciting field of research, and also for scientific guidance and mentorship.

I would like to acknowledge Prof. Dr. Benedikt Brors for supervising my thesis and providing me bioinformatics support.

I especially want to thank the whole AG Jäger, current as well as former members, for the nice conversations, delicious cakes, and providing me such a pleasant and motivational working atmosphere. Thank you all for these great years.

For performing all the experimental procedures I thank Rosa Eurich, Tina Lerchl, and Iris Kaiser. I want to thank Meggy Suarez-Carmona for her scientific input and inciting discussions.

Special thanks go to Julia Bitzer, who has warmly welcomed me from my first day on and became a dear friend. I feel grateful for having gone through this PhD time together with you.

I want to thank Petra Narrog, Silke Hirsch and Birgit Eberle for administrative support, and Rolf Kabbe and Alex Knurr for IT support.

I thank the DKFZ-Heidelberg Center for Personalized Oncology (DKFZ-HIPO) for technical support and funding through HIPO-034. I also want to thank the DKFZ Sample Processing Laboratory, the DKFZ Genomics Core Facility, and the DKFZ-HIPO Data Management Group for technical support and expertise. Special thanks go to Daniela Richter, Christina Görg, and Katja Beck for assistance and guidance within DKFZ-HIPO.

I want to acknowledge the members from the DKFZ Bioinformatics Departments, for providing such a great infrastructure. Thanks to Matthias Schlesner, Naveed Ishaque, Ivo Buchhalter, Octavio Espinosa, and Tobias Bauer for their expertise and help with different bioinformatics applications. Special thanks go to Barbara Hutter, for patiently answering all my bioinformatics-related questions and providing her expertise on personalized oncology.

I want to thank the members of the Surgery Department, Jürgen Weitz, Alexis Ulrich, Moritz Koch, Christoph Kahlert, and Fee Klupp for their cooperation in providing patient samples.

Special thanks go to Niels Halama, for his scientific groundwork and paving the way for this highly interesting research project, and also for his valuable scientific input and guidance.

I especially want to thank Inka Zörnig, for being a great supervisor. I feel very grateful for your inspiring scientific mentorship and guidance, and also for your personal kindness and friendly companionship.

I want to thank my dear friend Sebastian Boegel for his friendship and his valuable scientific insights.

I would like to acknowledge my fiancé Çağrı Yalçın, who has stepped into my life just recently and has been enriching ever since with his love and kindness.

Finally, and most importantly, I want to thank my parents, Kalender and Saliha Koşaloğlu, this work would not have been possible without you, and my siblings, Zehra, Sara, and Muhammed Koşaloğlu, thank you for your support, patience and love through all the years of my studies. Especially my sister Sara, who has been very motivating and supportive through all my ups and down.

# Summary

Cancer immunotherapies have recently shown outstanding clinical results in a number of patients across various tumor types. However, currently only a fraction of patients responds to immunotherapy, and it is a major concern to understand the underlying mechanisms. The composition of the tumor microenvironment has been shown to have an important impact on tumor growth and progression, as well as on response to therapy. It has been reported that the type and density of tumor-infiltrating immune cells are highly predictive for disease outcome in various cancers. These studies have also suggested that a high density of tumor-infiltrating lymphocytes is strongly correlated with mutational load. One hypothesis in this context is that somatic mutations found in cancer cells may give rise to novel epitopes, so-called neoepitopes, which attract and keep lymphocytes at the tumor site. Neoepitopes have also been suggested to be crucial for the outcome of immune checkpoint therapies, as it was reported that cancers with a high mutational load respond best to checkpoint therapy. An explanation for this is that mutations give rise to neoepitopes that can be targeted by specific T cells following their release from inhibitory signals.

It has now become evident that effective immunotherapies have to be tailored to the specific immune setting of each tumor. The complex interplay between the tumor and the immune system has to be systematically analyzed for characterizing patients and identifying therapies they will most likely benefit from. This highly personalized approach requires the integrated analysis of numerous tumor and host factors. Accordingly, the main aim of this PhD project was the establishment of an integrated analysis pipeline to obtain detailed data about tumor-host interactions, including analysis of the mutational and neoepitope load, the type and densities of tumor-infiltrating immune cells, the expression of immunological markers, and the expression of specific cytokines.

This analysis pipeline combines available genomic and immunomic resources and adds further depth into the analysis by additional computational pipelines. The already well established sequencing and somatic mutation detection pipelines that have been developed in the DKFZ bioinformatics departments (Prof. Roland Eils and Prof. Benedikt Brors) were integrated with the cytokine profiling and histological analysis workflows in Professor Jäger's group (NCT, Medical Oncology). Additional computational pipelines for HLA genotyping from sequencing data, as well as for epitope predictions for HLA class I and class II were implemented and included. Taken together these pipelines provide a broad picture of tumor-host interactions. The established analysis pipeline allows the rapid and systematic analysis of large patient cohorts.

Professor Jäger's group has been collecting colorectal cancer (CRC) liver metastases and systematically characterizing their immune cell infiltration and cytokine profiles, as well as the correlation to clinical outcome. In these studies it was shown that in general, there are at least two patient groups for each CRC stage: patients with high infiltrate density and patients with low infiltrate density, with the latter having a much worse prognosis. A patient cohort including 10 patients with high densities of infiltrating lymphocytes (TIL-high) and 10 patients with low densities (TIL-low) was

assembled and provided for analysis in this PhD project. The described integrated analysis pipeline was developed using this patient cohort.

The established analysis pipeline was then used to systematically investigate TIL-high versus TIL-low CRC metastases in order to assess the correlation of mutational and neoepitope load to lymphocyte infiltration and whether additional factors distinguishing the two groups can be discovered. The results show that the mutational and neoepitope load is not significantly different between patients with high and patients with low lymphocyte density in the analyzed patient cohort. Although a trend can be observed in a way that the TIL-high group seems to be enriched for mutations and neoepitopes, no statistical significance was detectable. Instead, the cytokine expression profiles are clearly distinct between the two subgroups: CXCL12, CXCL9, CCL7, CCL27, IL-17, IL-13, IL-7, IL-4, IFN $\gamma$ , GM-CSF, HGF, and TRAIL are significantly overexpressed in the TIL-high group. Interestingly both, pro-tumorigenic as well as anti-tumorigenic factors are overexpressed in the TIL-high group. Histological analysis additionally revealed that the TIL-high samples are enriched for macrophages. Furthermore, PD-L1, the ligand for the inhibitory immune checkpoint protein PD-1, is overexpressed in the majority of TIL-high samples when compared to the TIL-low samples. These results indicate that the immune contexture at the metastatic lesion seems to be a stronger factor for lymphocyte infiltration than the mutational and neoepitope landscape.

The established integrated analysis pipeline has already been applied in the clinic to conduct case studies with several patients being treated at the NCT. Patients with refractory and rare cancers were extensively analyzed for their genomic and immunomic features, which enabled the exploration of additional immunotherapeutic strategies. In doing so, a working logistics for the clinical setting was established, and the results provided insights into the feasibility of the approach. Based on these findings, clinical studies with neoepitope-based vaccines are currently under development in Professor Jäger's group, and the predictive impact of the newly established integrated analysis pipeline will be evaluated in prospective clinical trials.

# Zusammenfassung

Immuntherapien haben kürzlich in einer Reihe von Patienten und Tumorarten beeindruckende klinische Erfolge erzielt. Allerdings spricht nicht jeder Patient gleichermaßen auf Immuntherapien an, daher ist es derzeit ein großes Anliegen die zugrundeliegenden Mechanismen zu verstehen. Es ist bereits bekannt, dass die Zusammensetzung des immunologischen Mikromilieus einen großen Einfluss auf Tumorwachstum und Prognose, sowie auf Therapieansprechen hat. Es ist in zahlreichen Studien gezeigt worden, dass die Art und Dichte von Immunzellinfiltraten prädiktiv für den Krankheitsverlauf sind. Eine Hypothese in diesem Zusammenhang ist, dass somatische Mutationen in Krebszellen neue Epitope schaffen, sogenannte Neoepitope, die von Lymphozyten erkannt werden. Es wird weiterhin angenommen, dass Neoepitope einen grundlegenden Einfluss auf Immuncheckpoint-Therapien haben, denn es ist bereits mehrfach gezeigt worden, dass Patienten mit vielen somatischen Mutationen besonders gut auf diese Therapien ansprechen. Eine Erklärung hierfür ist, dass T Zellen, die zuvor durch inhibitorische Signale blockiert wurden, durch die Immuncheckpoint-Therapie wieder funktionsfähig werden und die Neoepitope auf den Krebszellen erkennen können.

Es ist nun weithin bekannt, dass effektive Immuntherapien auf das spezifische immunologische Mikromilieu des Patienten angepasst werden müssen. Hierzu muss das komplexe Zusammenspiel zwischen dem Tumor und dem Immunsystem des Patienten systematisch untersucht werden, um Patienten genau zu charakterisieren und personalisierte Therapien zu entwickeln. Dieses Vorgehen erfordert die integrative Analyse von verschiedenen Tumor- und Hostfaktoren. Daher war die Etablierung einer integrativen Analysepipeline, die es ermöglicht detaillierte Daten über Tumor-Host Interaktionen zu erhalten, das Hauptziel dieser Doktorarbeit. Diese integrative Analysepipeline umfasst die Analyse der Mutations- und Neoepitoplast, die Art und Dichte von Immunzellinfiltraten, die Expression von immunologischen Markern, sowie die Expression bestimmter Zytokine.

Die Analysepipeline verknüpft bereits vorhandene Ressourcen zur detaillierten Analyse von genomischen und immunologischen Fragestellungen und ergänzt diese durch weitere computergestützte Analyseverfahren. Die bereits etablierten Pipelines zur Sequenzuntersuchung, die in den DKFZ Bioinformatik Abteilungen (Prof. Roland Eils und Prof. Benedikt Brors) entwickelt worden sind, wurden mit den Methoden zur Untersuchung von Zytokinprofilen und histologischen Analysen in der Arbeitsgruppe von Professor Jäger (NCT, Medizinische Onkologie) verknüpft. Zusätzlich wurden weitere Analysepipelines zur HLA-Genotypisierung aus Sequenzdaten und zur Vorhersage von HLA Klasse I und Klasse II Epitopen entwickelt. Zusammengenommen bietet dieser Ansatz eine umfassende Analyse der Tumor-Host Interaktionen und die entwickelte integrative Analysepipeline ermöglicht die schnelle Verarbeitung von großen Patientenkollektiven.

In der Arbeitsgruppe von Professor Jäger werden verschiedene Studien durchgeführt mit dem Ziel, die Immunzellinfiltration und die Zytokinprofile in Lebermetastasen vom Kolorektalkarzinom systematisch zu untersuchen und mit dem klinischen Verlauf zu korrelieren. Diese Studien haben gezeigt, dass in der Regel mindestens zwischen zwei Patientengruppen unterschieden werden kann: Patienten mit einer hohen

Dichte von Tumor-infiltrierenden Lymphozyten (TIL-high) und Patienten mit niedriger Dichte (TIL-low), wobei letztere eine deutlich schlechtere Prognose haben. Ein Patientenkollektiv mit 10 TIL-high und 10 TIL-low Patienten wurde zusammengestellt und für die Untersuchung in dieser Doktorarbeit bereitgestellt. Die integrative Analysepipeline wurde an diesem Datensatz entwickelt.

Die entwickelte integrative Analysepipeline wurde anschließend dazu verwendet TIL-high und TIL-low Lebermetastasen vom Kolorektalkarzinom systematisch zu untersuchen und gegenüberzustellen, um die Korrelation zwischen Mutations- und Neoepitoplast und Lymphozyteninfiltration zu überprüfen, und um eventuell weitere Faktoren aufzudecken, die die beiden Gruppen unterscheiden. Die Ergebnisse zeigen, dass die Mutations- und Neoepitoplast in Patienten mit hoher Lymphozyteninfiltration tendenziell höher ist als in Patienten mit niedriger Lymphozyteninfiltration, dieser Unterschied ist jedoch nicht statistisch signifikant. Allerdings unterscheiden sich die Zytokinprofile der beiden Untergruppen eindeutig: die Expression von CXCL12, CXCL9, CCL7, CCL27, IL-17, IL-13, IL-7, IL-4, IFN $\gamma$ , GM-CSF, HGF und TRAIL sind in der TIL-high Gruppe signifikant erhöht. Interessanterweise befinden sich sowohl pro- also auch anti-tumorigene Faktoren unter den erhöhten Zytokinen. Histologische Analysen zeigten zudem, dass die TIL-high Proben eine höhere Dichte an Makrophagen aufweisen. Weiterhin ergab die Histologie, dass PD-L1, ein Ligand für das inhibierende Immuncheckpoint-Protein PD-1, in der Mehrzahl der TIL-high Proben überexprimiert ist im Vergleich zu den TIL-low Proben. Zusammenfassend zeigen diese Ergebnisse, dass das immunologische Milieu der Lebermetastasen vom Kolorektalkarzinom einen stärkeren Einfluss auf die Lymphozyteninfiltration zu haben scheint als die Mutations- und Neoepitoplandschaft.

Die etablierte integrierte Analysepipeline wurde bereits in der Klinik angewandt, um Fallstudien mit Patienten, die am NCT behandelt werden, durchzuführen. Patienten mit refraktären und seltenen Tumorkrankheiten wurden detailliert auf ihre genomischen und immunologischen Eigenschaften hin untersucht. Dies ermöglichte die Erforschung zusätzlicher immuntherapeutischer Ansätze. Es wurden dabei logistische und klinische Arbeitsabläufe etabliert, die es nun ermöglichen diese Ansätze klinisch umzusetzen. Basierend auf diesen Erfahrungen werden derzeit in der Arbeitsgruppe von Professor Jäger klinische Studien mit Neoepitop-basierten Impfstoffen entwickelt. Weiterhin sind prospektive Studien in Planung, die den prädiktiven Wert der entwickelten integrativen Analysepipeline evaluieren sollen.

---

# Contents

---

Erklärung . . . . .	v
Acknowledgements . . . . .	vi
Summary . . . . .	viii
Zusammenfassung . . . . .	viii
<b>List of Tables</b>	<b>xv</b>
<b>List of Figures</b>	<b>xvi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The Human Immune System . . . . .	2
1.1.1 The Adaptive Immune System . . . . .	2
1.1.2 Immunoinformatics . . . . .	7
1.2 Cancer . . . . .	8
1.2.1 Cancer Genomics . . . . .	9
1.3 Cancer Immunology . . . . .	10
1.3.1 Cancer Immunoediting . . . . .	10
1.3.2 Cancer Antigens . . . . .	11
1.3.3 Immunomodulatory Pathways . . . . .	15
1.3.4 The Tumor Microenvironment . . . . .	17
1.4 Cancer Immunotherapy . . . . .	20
1.4.1 Neoantigens in Cancer Immunotherapy . . . . .	23

1.5	Colorectal Cancer . . . . .	25
1.5.1	Treatment Approaches for Colorectal Cancer . . . . .	25
1.5.2	The Tumor Microenvironment in Metastatic Colorectal Cancer . . . . .	27
1.6	Motivation . . . . .	28
<b>2</b>	<b>Methods</b>	<b>33</b>
2.1	Samples . . . . .	34
2.1.1	Case Studies . . . . .	35
2.2	Histological Analysis . . . . .	35
2.3	Cytokine Profiling . . . . .	36
2.4	ELISpot Analysis . . . . .	37
2.4.1	Cell Preparation . . . . .	37
2.4.2	Peptides . . . . .	37
2.4.3	IFN $\gamma$ ELISpot assay . . . . .	37
2.5	Sequencing and Sequence Analysis . . . . .	38
2.5.1	DNA Library Preparation and Sequencing . . . . .	38
2.5.2	Mapping and Analysis of DNA Sequence Data . . . . .	38
2.5.3	Mutation Validation . . . . .	39
2.6	HLA Typing Pipeline . . . . .	39
2.7	Immunoinformatics Pipeline . . . . .	40
2.8	Downstream Analysis . . . . .	44
<b>3</b>	<b>Results</b>	<b>45</b>
3.1	Sequencing Data . . . . .	46
3.2	Mutational Landscape . . . . .	46
3.2.1	Single Nucleotide Variations . . . . .	46
3.2.2	Indels . . . . .	53
3.2.3	Copy Number Variations . . . . .	53
3.2.4	Pathway Enrichment Analysis of Mutations . . . . .	54
3.3	HLA Typing . . . . .	59
3.4	Epitope Landscape . . . . .	59
3.4.1	HLA Class I Epitopes . . . . .	59
3.4.2	HLA Class II Epitopes . . . . .	61

3.4.3	Shared Epitopes . . . . .	63
3.4.4	Immunogenic Mutations . . . . .	64
3.5	Histological Analysis . . . . .	65
3.6	Cytokine Analysis . . . . .	65
3.7	Case Studies . . . . .	72
<b>4</b>	<b>Discussion</b>	<b>76</b>
4.1	Neoepitope prediction and prioritization . . . . .	77
4.1.1	HLA class II Neoepitopes . . . . .	80
4.2	Sequencing-based HLA Typing . . . . .	81
4.3	The local microenvironment of CRC liver metastases . . . . .	81
4.3.1	Mutational landscape . . . . .	81
4.3.2	Shared epitopes . . . . .	82
4.3.3	Alterations in immune-related genes . . . . .	83
4.3.4	Pathways enriched for mutations . . . . .	84
4.3.5	Association of HLA allele and neoepitope load . . . . .	84
4.3.6	Microenvironment . . . . .	85
4.4	Further applications of the developed analysis pipeline . . . . .	86
4.4.1	Case studies . . . . .	86
4.5	Conclusion and Outlook . . . . .	88
<b>5</b>	<b>Supplemental Data</b>	<b>89</b>
<b>6</b>	<b>Technical Appendix</b>	<b>92</b>
	<b>Bibliography</b>	<b>95</b>
<b>7</b>	<b>Manuscript</b>	<b>114</b>

---

## List of Tables

---

2.1	Patient IDs and corresponding TIL status . . . . .	34
2.2	Primary antibodies used for immunohistochemical stainings . . . . .	35
2.3	Quantified cytokines . . . . .	36
2.4	List of primers used for PCR and validation sequencing of recurring mutations. . . . .	40
3.1	Mutations in frequently mutated genes. . . . .	51
3.2	Frequency of somatic mutations . . . . .	52
3.3	Indels in frequently mutated genes . . . . .	53
3.4	HLA genotype of patients. . . . .	60
3.5	Immunohistochemical staining for immunomodulatory proteins. . . . .	67
3.6	Selected predicted epitopes patient B8G6 . . . . .	74
3.7	Comparison of expression levels assessed by RNA-Seq and IHC . . . . .	75

---

## List of Figures

---

1.1	Binding grooves of MHC class I and MHC class II molecules. . . . .	4
1.2	Structure of MHC class I and MHC class II molecules. . . . .	4
1.3	The MHC class I and MHC class II antigen presentation pathways. . .	6
1.4	The genetic organization of the major histocompatibility complex (MHC) in human and mouse. . . . .	7
1.5	The hallmarks of cancer. . . . .	10
1.6	The cancer immunoediting concept. . . . .	12
1.7	Tumor antigens recognized by T cells. . . . .	14
1.8	Receptor-ligand pairs between T cells and antigen-presenting cells. . . .	17
1.9	Elements of the tumor microenvironment. . . . .	20
1.10	Immunohistochemical stainings for CD3 in samples with high lymphocyte infiltration . . . . .	30
1.11	Immunohistochemical stainings for CD3 in samples with low lymphocyte infiltration . . . . .	31
1.12	Integrative Analysis Approach . . . . .	32
2.1	Components of the Immunoinformatics pipeline. . . . .	43
3.1	Coverage of sequenced samples . . . . .	46
3.2	Mutational landscape . . . . .	48
3.3	Overview of mutations in frequently mutated genes . . . . .	49
3.4	Frequency of somatic exonic mutations . . . . .	49

3.5	Comparison of frequency of somatic missense mutations between TIL-high and TIL-low samples . . . . .	50
3.6	Copy number variations . . . . .	55
3.7	Comparison of frequency of copy number changes between TIL-high and TIL-low samples . . . . .	56
3.8	Top enriched canonical pathways . . . . .	57
3.9	Comparative analysis of top enriched immune-associated canonical pathways . . . . .	58
3.10	Correlation of number of mutations and number of predicted HLA class I binders . . . . .	61
3.11	Frequency of predicted HLA class I binders . . . . .	61
3.12	Comparison of HLA class I epitope frequencies between TIL-high and TIL-low samples . . . . .	62
3.13	Frequency of peptides predicted to be bound by HLA class II . . . . .	63
3.14	Comparison of HLA class II epitope frequencies between TIL-high and TIL-low samples . . . . .	64
3.15	Frequency of all missense mutations compared to immunogenic mutations	65
3.16	Comparison of immunogenic SNV frequencies between TIL-high and TIL-low samples . . . . .	66
3.17	Hierarchical clustering analysis of cytokine expression . . . . .	69
3.18	Cytokines significantly overexpressed in patients with high lymphocyte infiltration. . . . .	70
3.18	ELISpot results of case studies. . . . .	73

# CHAPTER 1

---

Introduction

---

## 1.1 The Human Immune System

The immune system protects the body against pathogens like bacteria, viruses and parasites. It can be subdivided into the innate immune system and the adaptive immune system. The innate immune system provides a first line of defense against many common pathogens by recognizing typical patterns on the pathogen surface and starts immediately after a pathogen has entered the body. The innate immune system, however, provides only limited protection against fast evolving pathogens. The adaptive immune system in contrast, is able to recognize almost every foreign protein. Main effectors of innate immunity are natural killer cells, mast cells, eosinophils, and basophils.

### 1.1.1 The Adaptive Immune System

B lymphocytes (B cells) and T lymphocytes (T cells) are the main effectors of the adaptive immune system. Receptors on their surface enable them to recognize foreign substances. A substance which can be recognized by surface receptors of the adaptive immune system is called antigen, and the region which is actually bound and recognized by lymphocytes is called epitope (Murphy et al., 2008).

T cells are the main effectors of cellular immunity. There are two major types of effector T cells, CD4<sup>+</sup> T helper cells and CD8<sup>+</sup> cytotoxic T cells. T helper cells have no cytotoxic activity, but they can activate and direct other immune cells, for example B cells, to eliminate the pathogen. Cytotoxic T cells, in contrast, are capable of directly inducing the death of pathogenic or malignant cells.

B cells are main effectors of the humoral immune response or antibody-mediated immune response. This aspect of immunity is mediated through secreted molecules, as opposed to cells mediating the cellular immune response (Murphy et al., 2008). Upon activation, B cells differentiate into plasma cells that secrete large amounts of antibodies. Antibodies can occur in soluble form found free-floating in blood plasma, or attached to the surface of B cells where they function as B cell receptors (BCR). Each antibody binds to a specific antigen and, upon binding, antibodies can either neutralize the bound pathogen or may recruit effector immune cells to eliminate the detected pathogenic or malignant cell.

B cell activation occurs in secondary lymphoid organs, such as the spleen and lymph nodes. There are two modes of B cell activation, either T cell-dependent or T cell-independent. Antigens that activate B cells independent of T cells are often simple, repetitive molecules such as glycoproteins or polysaccharides, as well as whole virus particles or whole bacterial cells that are free-floating in the intercellular space. B cells that differentiated through T cell-independent activation are short-lived, and the produced antibodies are usually weaker in affinity and efficiency. B cells are also capable of processing and presenting antigens which may lead to a T cell-dependent activation. These antigens are usually foreign antigens. Once the antigen is processed and presented on the cell surface (detailed description in section 1.1.1), T helper cells

that were already activated by the same antigen from other antigen presenting cells recognize the presented antigen and express surface proteins as well as cytokines which promote B cell activation and proliferation. These B cells then differentiate to long-lived plasma cells producing antibodies with high affinity and effectivity, and also to memory B cells which provide persistent protection.

The antigen receptors of B cells (BCR) and T cells (TCR) are highly variable. This variability is generated early during development in the primary lymphoid tissues, i.e. the bone marrow for B cells, and the thymus for T cells. The genes encoding the variable regions of BCRs and TCRs are rearranged randomly through somatic recombination which generates a high diversity of antigen binding sites.

T cell receptors can only recognize an antigen if it is bound to specific membrane proteins called major histocompatibility complex (MHC) molecules. CD8<sup>+</sup> cytotoxic T cells recognize antigens which are bound to MHC class I molecules, and CD4<sup>+</sup> T helper cells recognize antigens which are bound to MHC class II molecules.

## Antigen Processing and Presentation

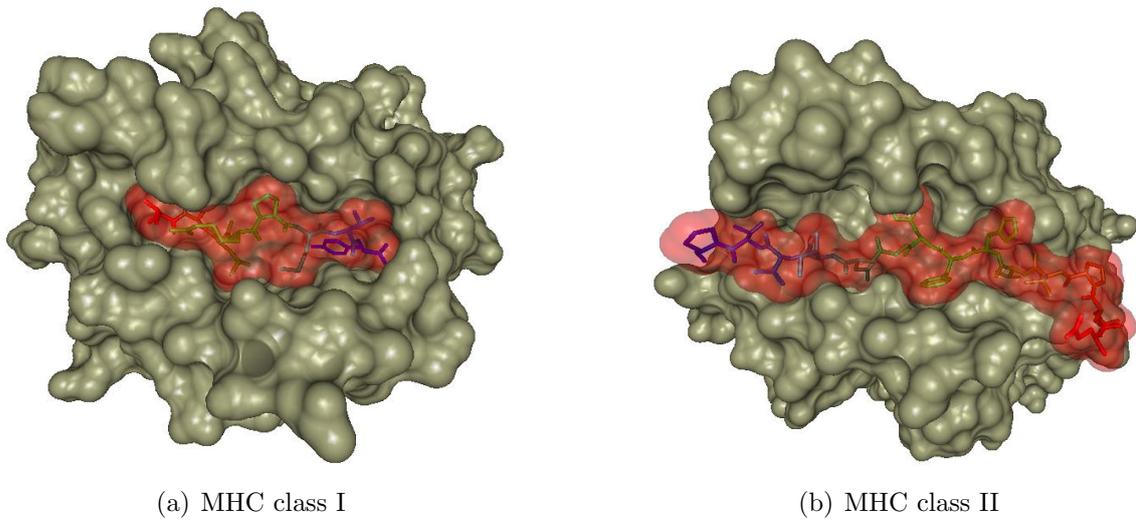
The two classes of MHC molecules, MHC class I and MHC class II are membrane-bound surface molecules. While MHC class I molecules are expressed by all nucleated cells, MHC class II molecules are primarily expressed by professional antigen presenting cells (APCs), such as dendritic cells (DCs), macrophages and B cells.

MHC class I is a heterodimer of a transmembrane  $\alpha$ -chain bound non-covalently to  $\beta_2$ -microglobulin (see Figure 1.2a). The  $\alpha$ -chain folds into three domains,  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$ . The folding of  $\alpha_2$  and  $\alpha_3$  build the antigen binding site of MHC class I. The binding groove of MHC class I is closed, and can bind peptides of length 8 - 10 amino acids (see Figure 1.1a).

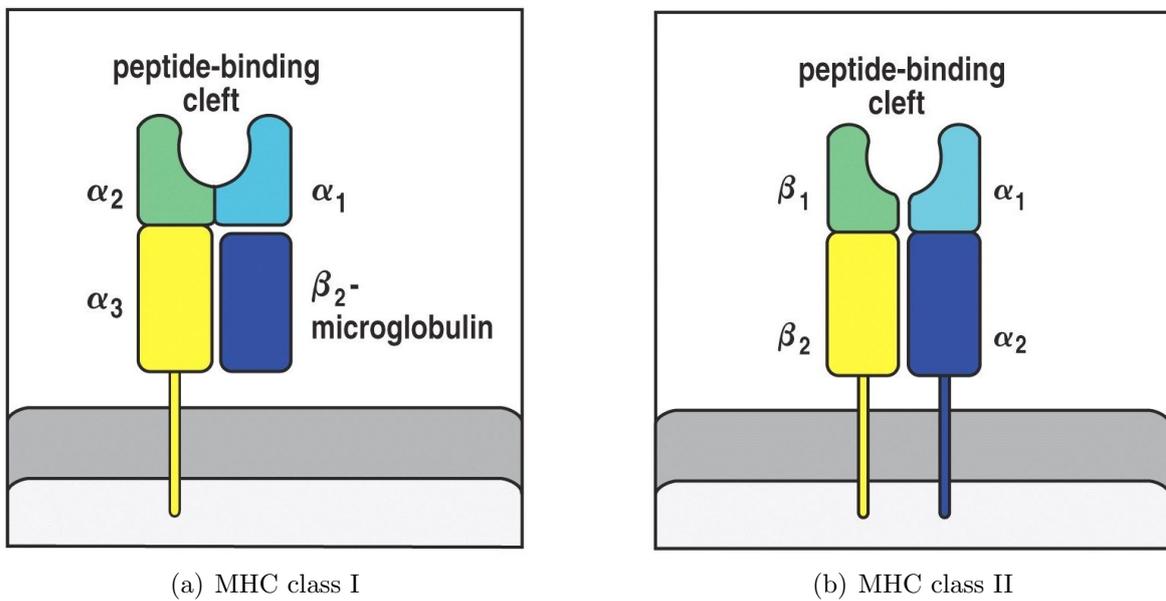
The MHC class II is composed of two transmembrane chains, the  $\alpha$ - and the  $\beta$ -chain (see Figure 1.2b). Each chain has two domains,  $\alpha_1$ ,  $\alpha_2$ , and  $\beta_1$ ,  $\beta_2$ , respectively. The  $\alpha_2$  and  $\beta_2$  domains form the antigen binding site of MHC class II. The binding groove of MHC class II is open, which allows the binding of longer peptides, possibly even whole proteins (see Figure 1.1b) (Nielsen et al., 2010; Sette et al., 1989). Typically, bound peptides are 13 - 18 amino acids long, but the binding groove only interacts with a core region consisting of 9 amino acids (Toussaint and Köhlbacher, 2009).

MHC class I molecules present endogenous peptides to cytotoxic T cells, whereas MHC class II molecules present exogenous peptides to T helper cells. These peptides are obtained via different pathways, and the combined specificities of MHC class I and MHC class II molecules cover antigens from almost all compartments of the cell, as well as antigens derived from intracellular and extracellular pathogens (Neeffjes et al., 2011).

MHC class I molecules present peptides which are the result of protein degradation by a protease complex, called the proteasome. Importantly, the C-terminus of the peptides is commonly determined by proteasomal cleavage, whereas the N-terminus can undergo



**Figure 1.1: Binding grooves of MHC class I and MHC class II molecules.** Panel (a) shows the surface model of the binding groove of MHC class I (gray) together with a bound peptide (red) (PDB-ID 1A1N). The surface model of the binding groove of MHC class II (gray) together with a bound peptide (red) is shown in panel (b) (PDB-ID 1BX2). Models were generated using the software BALLView (Moll et al., 2006)



**Figure 1.2: Structure of MHC class I and MHC class II molecules.** A schematic overview of the MHC class I molecule is shown in panel (a). MHC class I is a heterodimer of a membrane-spanning  $\alpha$ -chain bound covalently to  $\beta_2$ -microglobulin. The  $\alpha$ -chain folds and creates a groove, which is the antigen binding site. A schematic overview of the MHC class II molecule is shown in panel (b). MHC class II is composed of two transmembrane chains,  $\alpha$  and  $\beta$ . The antigen binding site is composed by two different chains. Figures were taken from Murphy et al. (2008)

further trimming by proteases (Backert and Kohlbacher, 2015). After being degraded, the peptides are translocated into the endoplasmic reticulum (ER) by the transporter molecule TAP (= transporter associated with antigen presentation). In the ER, the MHC class I molecule is assembled, and in the absence of a peptide MHC class I is stabilized by chaperone proteins such as calreticulin and ERp57 (see Figure 1.3). When peptides are loaded on MHC class I, the chaperones are released, the peptide:MHC class I complex leaves the ER and is then transported to the cell surface via the Golgi apparatus, where the antigen is presented to CD8<sup>+</sup> cytotoxic T cells. Peptides and MHC class I molecules that fail to associate in the ER are returned to the cytosol for degradation. The MHC class I loading pathway is displayed in Figure 1.3a.

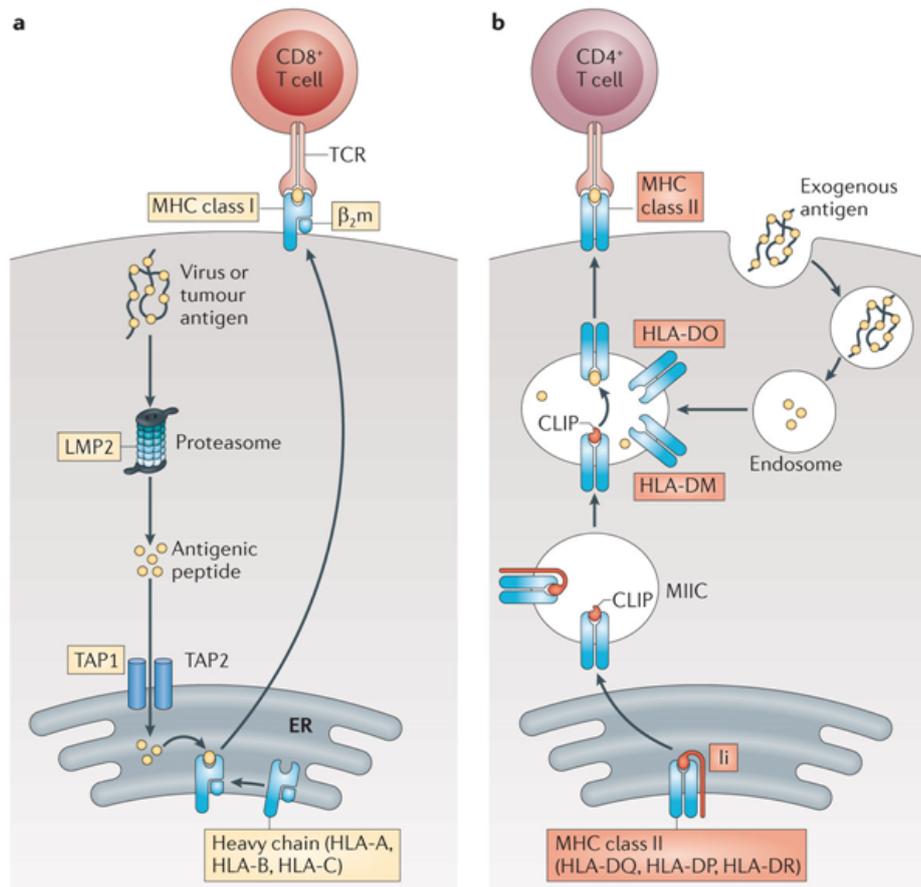
The  $\alpha$ - and  $\beta$ -chains of MHC class II molecules, are assembled in the ER, and form a complex with the invariant chain (Ii). Ii binds MHC class II with part of its polypeptide chain lying within the peptide-binding groove to stabilize the molecule, and also to prevent premature binding of peptides (see Figure 1.3b). The assembled Ii:MHC class II complex is transported to an endosomal compartment, the MHC class II compartment (MIIC), where peptide binding can occur. The invariant chain is then cleaved by acid proteases leaving a truncated form of the invariant chain, the class II associated Ii peptide (CLIP). CLIP blocks the peptide-binding groove of MHC class II. Endocytosed proteins are also degraded into short peptides by resident proteases in MIIC. Before these peptides can be loaded on MHC class II CLIP has to be released. This process is catalyzed by HLA-DM, which also stabilizes the MHC class II complex after CLIP release. Furthermore, HLA-DM catalyzes the loading of peptides to MHC class II. The peptide:MHC class II complex is then transported via the Golgi apparatus to the surface to present peptides to CD4<sup>+</sup> T helper cells. The MHC class II loading pathway is displayed in Figure 1.3a.

## The Genetic Organization of the Major Histocompatibility Complex

The major histocompatibility complex (MHC) is a large gene cluster where the genes of MHC class I and MHC class II are located (Figure 1.4). In humans, this complex is called the human leukocyte antigen (HLA) and is located on the short arm of chromosome 6. The MHC of the mouse is called H-2 and is located on chromosome 17.

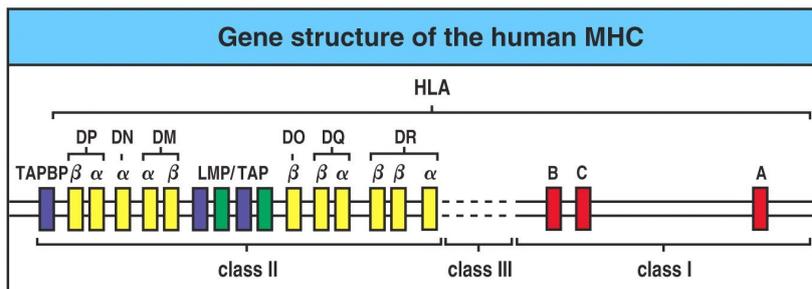
The MHC is a polygenic gene cluster containing several different MHC class I and MHC class II genes. In humans, there are three MHC class I  $\alpha$ -chain genes, HLA-A, HLA-B, HLA-C, and three pairs of genes encoding the MHC class II  $\alpha$ - and  $\beta$ -chains, HLA-DR, HLA-DP, HLA-DQ (see Figure 1.4).

The genes encoding the MHC class I and class II proteins are highly polymorphic (i.e. many alleles exist at each gene locus). Up to date, 5500 MHC class I alleles are known in human (IMGT/HLA database (Robinson and Marsh, 2000; Robinson et al., 2013)). The number of registered DR $\alpha$ , DR $\beta$ , DQ $\alpha$ , DQ $\beta$  and DP $\alpha$ , DP $\beta$  proteins is 2, 637, 26, 77, 16 and 118, respectively. This can potentially generate more than 4000 combinations of HLA class II  $\alpha$  and  $\beta$  subunits (Nielsen et al., 2010). The allelic variant of a particular MHC molecule determines its binding specificity and different MHC molecules bind different sets of peptides. The HLA genotype is variable among



**Figure 1.3: The MHC class I and MHC class II antigen presentation pathways.**

(a) Intracellular antigens, such as virus or tumor antigens, are degraded into peptides by the proteasome. Peptides are translocated via the transporter associated with antigen presentation (TAP) into the endoplasmic reticulum (ER) where they are loaded onto MHC class I molecules. Peptide:MHC class I complexes are released from the ER and transported via the Golgi apparatus to the surface where the antigen is presented to CD8<sup>+</sup> T cells. (b) Extracellular antigens, such as bacterial antigens, are endocytosed and degraded into peptides by resident proteases. The two MHC class II chains are assembled in the ER and form a complex with the invariant chain (Ii). Ii blocks the antigen binding site of the MHC class II molecule in order to prevent premature peptide binding. The Ii:MHC class II complex is transported to the MHC class II compartment (MIIC), where also the degraded peptides are located. In the MIIC, Ii is also degraded, which leaves the class II-associated Ii peptide (CLIP) remaining in the peptide-binding groove of MHC class II. Later, CLIP is exchanged for an antigenic peptide with the help of HLA-DM and the peptide:MHC class II complex is then transported to the surface where antigenic peptides are presented to CD4<sup>+</sup> T helper cell. Figure taken from Kobayashi and van den Elsen (2012) and reprinted with permission.



**Figure 1.4: The genetic organization of the major histocompatibility complex (MHC).** In humans the MHC is called HLA and is located on the short arm of chromosome 6. There are separate clusters of MHC class I and MHC class II genes and some other genes relevant for antigen presentation, like genes for the TAP peptide transporter and HLA-DM. Figure taken from Murphy et al. (2008).

different individuals, while within a population or ethnic group the genotype is more similar. The frequencies with which certain HLA alleles occur in different populations have been determined (Gonzalez-Galarza et al., 2011).

## 1.1.2 Immunoinformatics

Immunoinformatics, or computational immunology, is the application of computational methods to immunological problems (Backert and Kohlbacher, 2015). The most important task of immunoinformatics is to analyze immunological data using computational tools to generate biologically significant and rational interpretations (Rajat K. De, 2014). This includes the development of algorithms for mapping potential B cell and T cell epitopes, and the inference of the HLA genotype from sequencing data, as well as the establishment and curation of immunological databases.

A wealth of immunological information about MHC ligands, epitopes, T cell receptors, and HLA alleles is available in several public databases (reviewed in Backert and Kohlbacher (2015)).

One of the best established branches of immunoinformatics is the development of tools for the prediction of HLA binding, also called epitope prediction. These tools employ so-called machine learning methods, which use a given training dataset to learn a function that maps a given input to its corresponding output (Backert and Kohlbacher, 2015). In the context of epitope prediction, large datasets of experimentally derived epitopes are used as training datasets. The accuracy of these prediction tools increased recently, due to the raising availability of large-scale immunological data (Backert and Kohlbacher, 2015).

There are several tools available for HLA class I epitope prediction, which differ in the used machine learning method and the training datasets (reviewed in Backert and Kohlbacher (2015)). An ongoing continuous benchmark performed by Trolle and colleagues (Trolle et al., 2015) shows that, currently, netMHCpan (Nielsen et al., 2007)

is the best-performing tool, which is why this tool for HLA class I epitope prediction was chosen for this study.

As described above, the first step of the MHC class I antigen processing pathway is the cleavage of proteins by the proteasome. Tools for proteasomal cleavage prediction work analogously to epitope prediction tools and are trained using *in vitro* or *in vivo* data (Backert and Kohlbacher, 2015). The prediction tool NetChop (Kemir et al., 2002) which was developed by the same group that developed netMHCpan, was chosen for this study.

HLA class II epitope prediction is more difficult than predictions for class I, because as described above, the binding groove of HLA class II is open and the position of the binding core within an epitope is not known. Additionally, polymorphism of the HLA locus and the high number of combinations of HLA class II  $\alpha$ - and  $\beta$ -chains complicates the task. For HLA-DR only the beta chain is polymorphic, so only HLA-DRB needs to be considered for epitope predictions. Besides from the higher complexity of HLA class II epitope prediction, the tools work analogously to class I predictions (reviewed in Backert and Kohlbacher (2015)). As netMHCIIpan (Karosiene et al., 2013) is the most recent tool, this tool was chosen for HLA class II epitope prediction for my work.

Development of tools for HLA typing from sequencing data emerged recently together with the emergence of next-generation sequencing (NGS). In many clinical applications sequencing data of a patient is already available, and costs for experimentally determining the HLA genotype can be avoided by using computational tools. Recently, several tools for NGS-based HLA typing emerged (Boegel et al., 2012; Warren et al., 2012; Kim and Pourmand, 2013; Bai et al., 2014). Each tool starts by using an alignment software like Bowtie (Langmead et al., 2009) or BWA (Li and Durbin, 2010) to map the input sequence reads to reference HLA sequences retrieved from the IMGT/HLA (Robinson and Marsh, 2000; Robinson et al., 2013) database. For scoring and the subsequent selection of candidate alleles each tool uses different approaches. For this study, the most recent tool, Phlat, was chosen (Bai et al., 2014).

Immunoinformatics tools are mainly used in the context of epitope-based vaccines, including cancer immunotherapy. The reverse immunology approach starts with the identification of candidate peptides using epitope prediction algorithms. These candidate peptides are then synthesized, tested *in vitro*, and then used to stimulate T cells or used in peptide-based cancer vaccines (see 1.4.1).

## 1.2 Cancer

Cancer is the result of uncontrolled and indefinite cell growth and division. It has become evident that cancer is a genetic disease where somatic cells acquire multiple mutations that overwhelm the barriers that normally restrain their uncontrolled expansion (Vesely et al., 2011). The DNA replication system is prone to mistakes, hence every healthy cell has a wide number of intrinsic defense mechanisms by which damage

to the DNA is identified and corrected. If repairs fail, senescence or apoptosis will be triggered to keep cell proliferation from becoming aberrant.

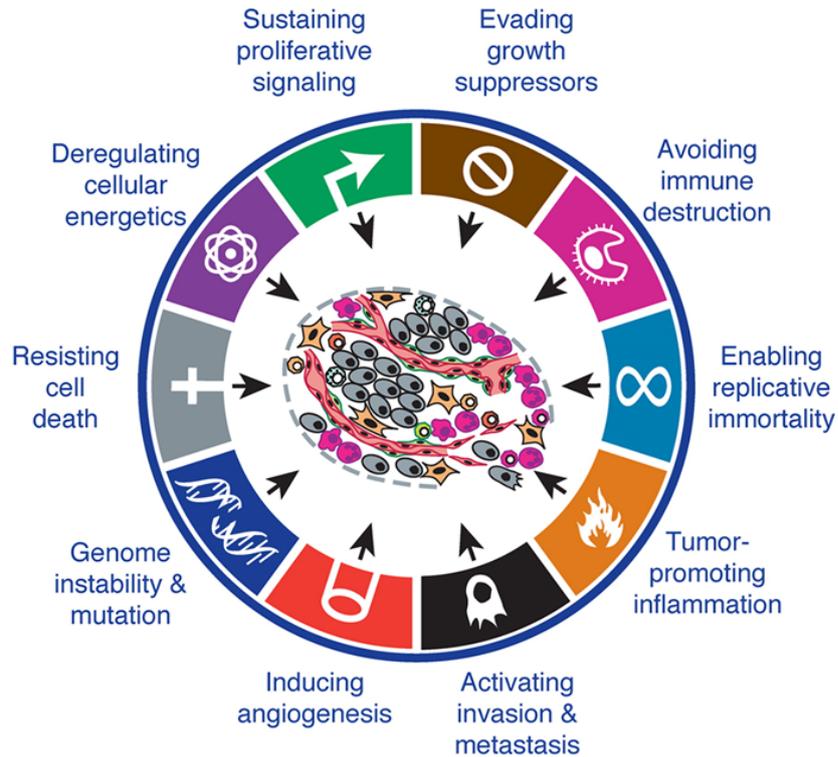
Proto-oncogenes for instance, code for proteins that are involved in the regulation of cell growth and differentiation in normal cells. Due to mutations, a proto-oncogene can be activated and then becomes an oncogene, which contributes positively to the growth of a tumor (Todd and Wong, 1999). Many oncogenes have been identified in human cancer, such as RAS, WNT and MYC.

Tumor suppressor proteins in contrast, contribute negatively to the regulation of the cell cycle (Sherr, 2004). They can additionally sense genomic disturbances caused by mutations, or also sense the activity of oncogenes, and can induce senescence and apoptosis in response (Vesely et al., 2011). Examples of tumor suppressor genes are TP53, APC or PTEN. Mutations and loss of tumor suppressor genes is crucial for the formation of many cancers (Weinberg, 2013).

Multiple changes need to occur, so that intrinsic defense mechanisms of normal cells against becoming cancerous fail and they acquire the capability of indefinite growth. Six fundamental changes and features that drive the transformation of normal cells to cancer were identified and described as the hallmarks of cancer by Hanahan and Weinberg (Hanahan and Weinberg, 2000). These hallmarks were defined as the following: self-sufficient in growth signals, insensitivity to anti-growth signals, evading apoptosis, limitless replicative potential, sustained angiogenesis, tissue invasion and metastasis. Since the publication of these original hallmarks, substantial progress has been made in cancer research, which led to the suggestion of four additional hallmarks: avoiding immune destruction, deregulating cellular energetics, tumor-promoting inflammation, genome instability and mutation (Hanahan and Weinberg, 2011) (Figure 1.9).

### 1.2.1 Cancer Genomics

As cancer is primarily caused by the accumulation of genomic aberrations, it is also referred to as a genomic disease (Meyerson et al., 2010). So-called somatic mutations are alterations in DNA that occur after conception, and are in contrast to germline mutations neither inherited nor passed to offspring (U.S. National Library of Medicine). With next-generation sequencing (NGS) techniques it is today feasible to detect such genomic changes by sequencing the whole genome, or the coding part, the so-called exome, of a tumor sample (Whole-Genome-Sequencing and Whole-Exome-Sequencing). The analysis of the cancer genome or exome, also called cancer genomics, provides important insights for understanding the biology of cancer. The detection of genomic aberrations, including single nucleotide variations (SNVs), small insertions and deletions (Indels), and chromosomal copy number variations (CNVs), can also aid in diagnosis and therapy. The product of each alteration is a potential target for therapy, for example targeted therapies with inhibitors against mutated growth factors, or immunotherapies against mutated antigens.



**Figure 1.5: The hallmarks of cancer.** Displayed are the original six hallmarks of cancer as proposed by Hanahan and Weinberg in 2000 (Hanahan and Weinberg, 2000), together with the emerging hallmarks added in 2011 (Hanahan and Weinberg, 2011). The original hallmarks are: self-sufficient in growth signals, insensitivity to anti-growth signals, evading apoptosis, limitless replicative potential, sustained angiogenesis, tissue invasion and metastasis, and the four emerging hallmarks are avoiding immune destruction, deregulating cellular energetics, tumor-promoting inflammation, genome instability and mutation. Figure adapted from Hanahan and Weinberg (2011) and reprinted with permission.

## 1.3 Cancer Immunology

### 1.3.1 Cancer Immunoediting

The immune system is capable of specifically identifying transformed cells that have escaped the intrinsic tumor suppressor mechanisms and eliminating them before they can establish malignancy (Vesely et al., 2011). The core principal of this extrinsic tumor suppressor mechanism, also referred to as cancer immunosurveillance, is that cancer cells express antigens that distinguish them from their non-transformed counterparts, thus enabling their recognition by T cells and their eventual destruction by immunological mechanisms (Vesely and Schreiber, 2013). Numerous studies during the past 20 years demonstrated that the immune system can in fact protect mice from outgrowth of many different types of primary and transplanted tumors (reviewed in Vesely et al. (2011)).

In 2001, studies in immunodeficient mice showed that tumors that formed in the absence of a functional immune system are more immunogenic than tumors that arise in immunocompetent hosts (Shankaran et al., 2001; Dunn et al., 2002). These findings led to the cancer immunoediting hypothesis, which postulates that the immune system not only protects the host against tumor formation but also edits tumor immunogenicity. Hence, the immune system is attributed a dual role of host-protective and tumor-promoting behavior. In its most complex form, cancer immunoediting occurs in three sequential phases: elimination, equilibrium, and escape (Figure 1.6) (Schreiber et al., 2011).

The elimination phase can be described as an updated version of immunosurveillance: a developing tumor is detected by the immune system and eliminated before it becomes clinically apparent. For effective cancer elimination, a coordinated interplay of both innate and adaptive immune components is needed. Here, activation and expansion of effector CD4<sup>+</sup> and CD8<sup>+</sup> T cells plays an important role. If the tumor destruction is successful, the elimination phase represents the endpoint of the cancer immunoediting process (Schreiber et al., 2011).

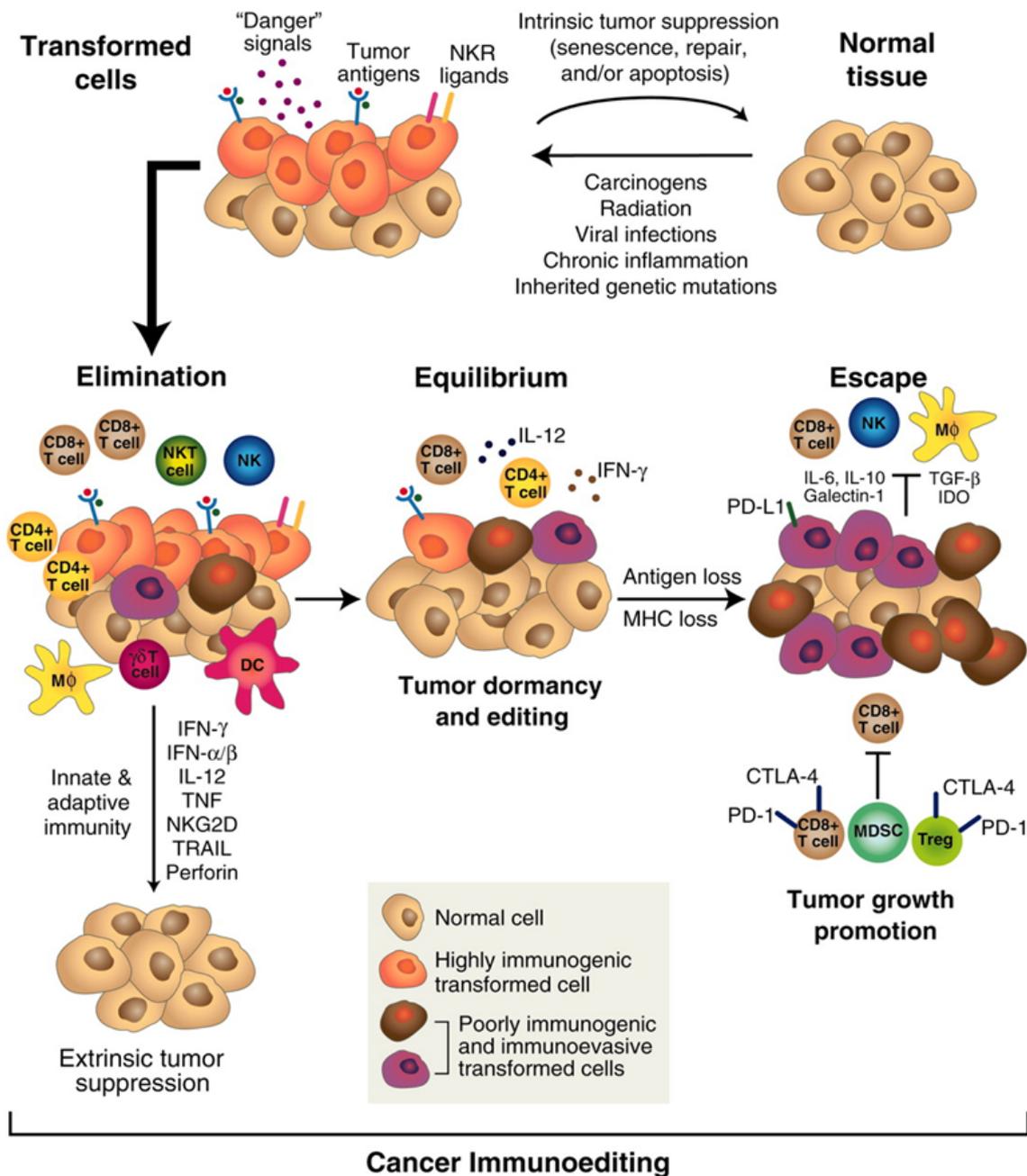
When tumor cells survive the elimination phase, they may enter the equilibrium phase. Here, the adaptive immune system prevents the outgrowth of tumor cells and additionally edits their immunogenicity. Schreiber et al. describe this phase as a dormancy of functional tumor cells, where tumor cells can reside in patients for decades, eventually extending throughout their lives. Thus, the equilibrium phase may also represent a stable endpoint of cancer immunoediting (Schreiber et al., 2011).

Tumor cells may escape the elimination and equilibrium phases by acquiring the ability to circumvent recognition and destruction by the immune system. In this escape phase tumors become clinically apparent. Tumor cell escape can occur through many different mechanisms, either when the cancer establishes an immunosuppressive tumor microenvironment and impairs the function of the immune system (described in section 1.3.4), or when tumor cells change in response to the editing process (Schreiber et al., 2011). The best-studied escape mechanism is the loss of tumor antigens, which is further described in section 1.3.2.

### 1.3.2 Cancer Antigens

At the core of immune recognition of cancer cells lies the fact that cancer cells express antigens that distinguish them from their non-transformed counterparts. These tumor-specific antigens can be recognized by specific T cells which are then capable of destroying the detected tumor cells. There are different classes of tumor-specific antigens: viral antigens, antigens that result from mutations, antigens that are encoded by cancer-germline genes, differentiation antigens, and antigens that are overexpressed in tumor cells (Coulie et al., 2014).

Differentiation antigens are expressed only in the tumor cells and in the normal tissue of origin. Melan-A (also known as MART1) is an example for an antigen which is specific for the melanocyte lineage. Prostatic acid phosphatase (PAP) is an example for a



**Figure 1.6: The cancer immunoediting concept.** The cancer immunoediting hypothesis postulates that the immune system not only protects the host against tumor formation but also edits tumor immunogenicity. In its most complex form, cancer immunoediting occurs in three sequential phases: elimination, equilibrium, and escape. Figure taken from Schreiber et al. (2011) and reprinted with permission.

prostate-specific antigen. Spontaneous T cell responses against differentiation antigens have been reported repeatedly (Cox et al., 1994; Bakker et al., 1995; Kawakami et al., 1994), and it is not known yet, why tolerance is incomplete against these antigens (Coulie et al., 2014).

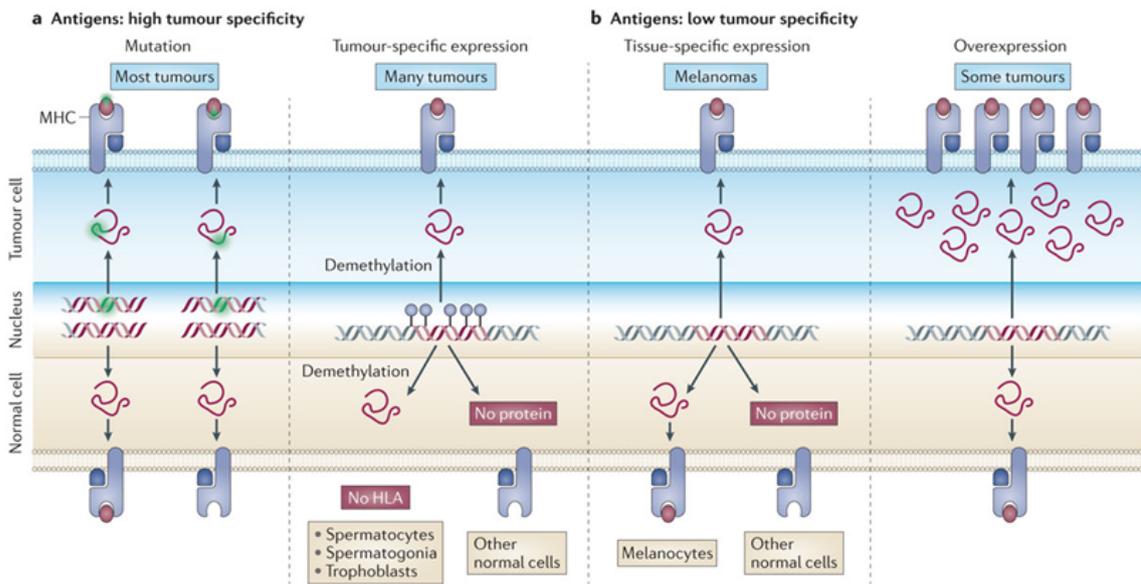
Overexpression of specific proteins in tumor cells can also be used to elicit a tumor-specific immune response, because a threshold level of antigen expression is required for a cell to be recognized by T cells. Several proteins have been detected to be tremendously overexpressed in tumor cells when compared to normal cells. The protein HER2 is such an example, which is overexpressed due to gene amplifications in several epithelial tumors including ovarian and breast cancer. Another example of tumor-specific overexpression is MUC1 which is overexpressed in most adenocarcinomas (Vlad et al., 2004). However it remains difficult to assess an appropriate tumor-specific overexpression, and it has to be ruled out if a high expression occurs even in a small subset of normal cells (Coulie et al., 2014). Viral, mutation-derived, and cancer-germline antigens are highly tumor-specific and can elicit a specific immune response against tumor cells. Viral antigens have been shown to be important for therapy in a subset of tumors, including cervical carcinoma, hepatocarcinoma, and adult T cell leukemia (van der Burg and Melief, 2011; Long et al., 2010; Fujita et al., 2008; Ramos and Lossos, 2011).

Cancer-germline genes were found to be expressed in many human tumors of different histological type, whereas no expression was found in normal tissues, except for male germline cells and trophoblastic cells (Coulie et al., 2014). Hence, cancer-germline genes are an important source of tumor-specific antigens and can be used as targets for immunotherapy. The expression of cancer-germline genes in male germline cells and trophoblastic cells is irrelevant in this context, as these cells do not produce MHC and cannot present antigens to T cells (Fiszler and Kurpisz, 1998). To date, over 60 cancer-germline genes have been discovered (Coulie et al., 2014). The gene MAGEA1 was the first cancer-germline gene that was discovered in melanoma in 1991 (van der Bruggen et al., 1991). Examples for families of cancer-germline genes are MAGE, GAGE, CTAG, and SSX, which are all located on chromosome X. Despite recent progress, the functions of most of these genes remain unclear (Coulie et al., 2014).

Mutated antigens greatly contribute to the immunogenicity of human tumors. A single nucleotide variation can potentially lead to the production of various new antigenic peptides, so-called neoantigens, that can be recognized by autologous T cells (Coulie et al., 2014). About one-half of the recognized tumor-specific antigens in cancer patients are derived by mutations (Vigneron et al., 2013). The contribution of neoantigens to tumor immunogenicity correlates with the mutation rate of the corresponding tumor: entities with high mutation rates like melanoma, lung cancer, and microsatellite instable colorectal cancer are expected to express a high number of neoantigens (Coulie et al., 2014). Most of the neoantigens are caused by passenger mutations and are therefore often patient-specific (Coulie et al., 2014; Schumacher and Schreiber, 2015). Mutation-derived neoantigens are highly attractive as immunotherapeutic targets as they are expected to overcome central tolerance, and their expression is highly tumor-specific (Hacohen et al., 2013; Heemskerk et al., 2013). Usage of neoantigens in personalized

immunotherapies is gaining more momentum and has been implemented in several clinical studies (Schumacher and Schreiber, 2015). This therapy approach will be discussed in more detail in section 1.4.1.

Tumor-specific antigens play an essential role in the establishment of malignancy. Tumor cells can escape the immune system due to a lack of immunological recognition and may remain invisible to the adaptive immune system (Vesely and Schreiber, 2013). There are several possibilities to evade immunological recognition. Tumor cells may stop the expression of strong tumor rejection antigens, or genomic instability within tumor cells may result in the loss of tumor-specific antigens (Coulie et al., 2014). Furthermore, alterations in antigen processing and presentation pathways may facilitate evasion from adaptive immune recognition. Analogous to other antigens, tumor-specific antigens also undergo the MHC class I or MHC class II antigen processing pathways and are then presented on MHC, where they can be detected by specific T cells. Tumor cells can acquire defects in molecules that are involved in antigen processing and presentation, specifically loss of TAP1, MHC class I,  $\beta 2m$ , LMP2, and LMP7 have been reported (Vesely et al., 2011). The result is the emergence of poorly immunogenic tumor cell variants that become invisible to the immune system and can grow unhindered.



**Figure 1.7: Tumor antigens recognized by T cells.** a) Highly tumor-specific antigens. A single nucleotide variation can lead to the production of new antigenic peptides (neoantigen) that can be recognized by autologous T cells. Cancer-germline genes are specifically expressed in many human tumors and on germline cells. As germline cells do not produce HLA, no antigens are presented on these cells. b) Antigens of limited tumor specificity. Differentiation antigens are tissue-specific and are expressed only in the tumor cells and in the normal tissue of origin. Overexpression of specific proteins in tumor cells can also trigger a tumor-specific immune response, because a threshold level of antigen expression is required for a cell to be recognized by T cells. Figure taken from Coulie et al. (2014) and reprinted with permission.

### 1.3.3 Immunomodulatory Pathways

The amplitude and quality of T cell responses is regulated by a balance between co-stimulatory and inhibitory signals, which are commonly referred to as immune-checkpoints. In a non-malignant setting, immune-checkpoints are crucial for the prevention of autoimmunity. Tumors however, dysregulate the expression of immune-checkpoint as an immune-evasion mechanism. Recently, some of these immune-checkpoints are being targeted therapeutically with tremendous clinical advances in numerous cancers. Clinical application of checkpoint therapy is discussed in section 1.4, whereas the biological function and features of selected checkpoints are outlined in the following subsections.

#### Inhibitory checkpoints

Cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) is an inhibitory immune-checkpoint regulator that is expressed by activated T cells and regulatory T cells, where it plays a major role in regulating the amplitude of T cell response (Pardoll, 2012). Primarily, CTLA-4 counteracts the activity of its co-stimulatory homologue CD28 (Schwartz, 1992; Rudd et al., 2009). In addition to the peptide-MHC-TCR interaction, co-stimulatory signals provided by CD28 are required to activate a T cell. CD28 and CTLA-4 share the same ligands: B7-1 (CD80) and B7-2 (CD86) which are expressed on APCs (Hathcock et al., 1993; Freeman et al., 1993). Binding of CTLA-4 to its ligands leads to inhibition of T cells (Krummel and Allison, 1995; Gabriel and Lattime, 2007). CTLA-4 has a much higher affinity than CD28 for both ligands and outcompetes the activity of CD28 (Linsley et al., 1994). Clinical blockade of CTLA-4 is thought to bias this competition for ligands in favor of CD28 and reverse the inhibitory function of CTLA-4 (Melero et al., 2015).

Programmed cell death protein 1 (PD-1) is another inhibitory checkpoint protein. While CTLA-4 predominantly regulates T cell activation, PD-1 predominantly regulates T cell effector activity, hence the function of these two checkpoint proteins is distinct from each other (Sharma and Allison, 2015; Pardoll, 2012). PD-1 expression is induced when T cells become activated to limit the response of the activated T cells in order to avoid autoimmunity (Ishida et al., 1992). PD-1 is more broadly expressed than CTLA-4: it is expressed by activated T cells, B cells, NK cells, and regulatory T cells (Melero et al., 2015). Two ligands are known for PD-1: PD-L1 (CD274) and PD-L2 (CD273). PD-L1 expression is induced after exposure to IFN- $\gamma$  and was found to be expressed on many cell types including T cells, epithelial cells, endothelial cells, and tumor cells (Dong et al., 2002). The PD-1 signalling cascade contributes to T cell exhaustion (Barber et al., 2006) and tumors may exploit this pathway to circumvent elimination. In many different tumor types, PD-1 was found to be highly expressed on tumor-infiltrating lymphocytes such as regulatory T cells and CD8<sup>+</sup> T cells (Sfanos et al., 2009; Ahmadzadeh et al., 2009). Analogously, PD-1 ligands are commonly upregulated on tumor cells, PD-L1 was shown to be highly expressed on many solid tumors, which was associated with poor outcome (Flies et al., 2011; Lipson et al., 2015; Sznol

and Chen, 2013). These expression patterns may support the suppression of immune effector mechanisms, and clinical blockade of these proteins may reverse this state and enhance the activity of effector T cells and NK cells (Pardoll, 2012).

The inhibitory receptor named lymphocyte activation gene 3 protein (LAG-3) is expressed on T cells during exhaustion (Wherry et al., 2007; Sierro et al., 2011). Known ligands for LAG-3 are MHC class II and galectin 3, which was just recently discovered (Kouo et al., 2015). It is anticipated that the blockade of LAG-3 may reverse T cell exhaustion and enhance antitumor activity (Woo et al., 2012; Sierro et al., 2011).

T cell immunoglobulin and mucin domain-containing 3 (TIM-3, also known as HAVCR2) is expressed on T helper 1 cells and cytotoxic T cells, but also on DCs (Anderson, 2012; Snchez-Fueyo et al., 2003). The function of TIM-3 differs depending on the cell type. TIM-3 was found to be commonly expressed by tumor-infiltrating lymphocytes in melanoma (Fourcade et al., 2010) and NSCLC (Gao et al., 2012), where it is thought to keep lymphocyte status inactive or even induce apoptosis upon binding a ligand (Melero et al., 2015). One known ligand of TIM-3 is galectin 9 and it is assumed that there are other ligands which have not been defined yet (Sakuishi et al., 2010).

## Co-stimulatory checkpoints

CD137 (TNFRSF9, also known as 4-1BB) is a T cell and NK cell co-stimulatory receptor, which is expressed on the cell surface upon activation (Wang et al., 2009). Binding of its ligand CD137L activates signalling that improves cytotoxic antitumor responses and T cell survival (Melero et al., 1997; Shuford et al., 1997; Vinay and Kwon, 2012).

Glucocorticoid-induced tumor necrosis factor receptor family-related protein (GITR, also known as TNFRSF18) is a co-stimulatory molecule that is expressed on T cells. Binding of its ligand GITRL induces the reversion of Treg-mediated suppression of T cells and activates proliferation and effector functions in CD4<sup>+</sup> and CD8<sup>+</sup> T cells (Ko et al., 2005; Schaer et al., 2012).

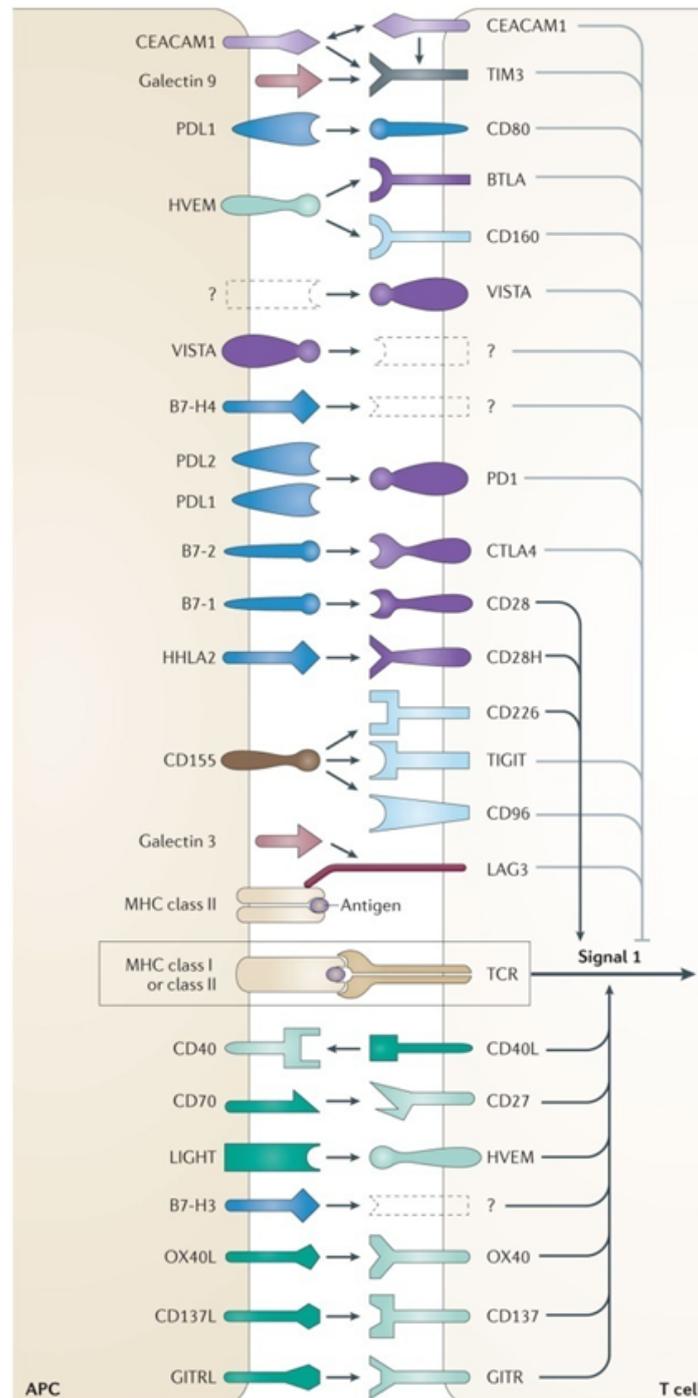
The co-stimulatory receptor OX40 (TNFRSF4) is primarily expressed on activated CD4<sup>+</sup> and CD8<sup>+</sup> T cells where it enhances antitumor immune responses by promoting T cell proliferation and survival (Ruby et al., 2007; Piconese et al., 2008). OX40L is the ligand of OX40.

CD40 (TNFRSF5) is a co-stimulatory receptor that is expressed on APCs where it promotes activation of APCs and enhances their antigen presentation activities, leading to T cell activation (Quezada et al., 2004; Vonderheide and Glennie, 2013). CD40L is the ligand of CD40.

The co-stimulatory receptor CD27 is expressed on resting and naive, but not on fully differentiated T cells. It is also expressed on a subset of NK cells (Denoeud and Moser, 2011). Binding to its ligand CD70 enhances T cell activation effector function, maturation and survival (Hintzen et al., 1995; Hendriks et al., 2003; Hintzen et al., 1994). CD27 was also shown to enhance NK cell proliferation and cytotoxic

### 1.3.4 The Tumor Microenvironment

As tumors grow, they initiate profound molecular, cellular and physical changes on the tissue in which they reside. The so-called tumor microenvironment emerges, a complex milieu that comprises non-malignant cells such as stromal cells (e.g. fibroblasts



**Figure 1.8:** Receptor-ligand pairs between T cells and antigen-presenting cells. Detailed description on following page.

**Figure 1.8:** (Figure on previous page.) An immune synapse between a T cell (right) and an antigen-presenting cell (APC, left) is shown. Various ligand-receptor interactions that regulate T cell response to antigens presented on MHC are highlighted, co-stimulatory interactions with black pointed arrows and inhibitory interactions with gray inhibitory arrows. B7-H3, B7 homologue 3 (also known as CD276); B7-H4, B7 homologue 4 (also known as VCTN1); BTLA, B lymphocyte and T lymphocyte attenuator; CD28H, CD28 homologue; CD40L, CD40 ligand; CD137L, CD137 ligand (also known as TNFSF9); CEACAM1, carcinoembryonic antigen-related cell adhesion molecule 1; CTLA4, cytotoxic T lymphocyte-associated antigen 4; GITR, glucocorticoid-induced TNFR family-related protein; GITRL, GITR ligand; HHLA2, HERV-H LTR-associating 2 (also known as B7-H7); HVEM, herpes virus-entry mediator (also known as TNFRSF14); Ig, immunoglobulin; LAG3, lymphocyte activation gene 3 protein; LIGHT, HVEM ligand (also known as TNFSF14); MHC, major histocompatibility complex; OX40L, OX40 ligand (also known as TNFSF4); PD1, programmed cell death protein 1; PDL, PD1 ligand; TIGIT, T cell immunoreceptor with Ig and ITIM domains; TIM3, T cell Ig mucin domain-containing 3; TNF, tumor necrosis factor; TNFR, TNF receptor; VISTA, V-domain Ig suppressor of T cell activation (also known as PD1 homologue). Figure taken from Melero et al. (2015) and reprinted with permission.

and endothelial cells), mesenchymal cells, and immune cells, along with the numerous mediators these cells secrete (Turley et al., 2015). Tumor cells can closely interact with elements of the tumor microenvironment, which can have effects on tumor cell survival, invasiveness, and responsiveness to therapeutics (Joyce and Pollard, 2009; Polyak et al., 2009).

Tumors are often infiltrated by various innate and adaptive immune cells, which can perform both, protumor and antitumor functions. Establishment of an immunosuppressive microenvironment is a key mechanism of immune escape, as mentioned in section 1.3.1. Tumor cells can directly produce factors to inhibit the function of both, innate and adaptive immunity, for example by upregulating the expression of inhibitory checkpoint proteins as described in section 1.3.3 (Vesely et al., 2011). Also, immunosuppressive leukocytes are deliberately recruited by the tumor, for example: FOXP3<sup>+</sup> regulatory T cells (*T<sub>regs</sub>*), T helper 2 (Th2) cells, myeloid-derived suppressor cells (MDSCs), M2 macrophages, and N2 neutrophils (Senovilla et al., 2012). In contrast, leukocytes with antitumor effects also infiltrate the microenvironment and the tumor as implementers of an ongoing immune response, for example: CD8<sup>+</sup> cytotoxic T lymphocytes (CTLs), T helper 1 (Th1) and T helper 17 (Th17) T cells, M1 macrophages, N1 neutrophils, natural killer (NK) cells, and dendritic cells (DCs) (Fridman et al., 2012; Zitvogel et al., 2011). Moreover, it has been shown for various cancers that type and density of tumor-infiltrating lymphocytes (TILs) correlate with patient survival (Schreiber et al., 2011).

## Tumor-Infiltrating Lymphocytes

The first evidence that TILs are associated with favorable patient prognosis was shown in melanoma (Clark et al., 1989; Clemente et al., 1996), followed by a study in colon cancer, which showed that particularly CD8<sup>+</sup> cytotoxic TILs have important influence

on clinical outcome (Naito et al., 1998). Since then, numerous studies in a wide range of cancers confirmed correlation of TILs and prognosis. It was even found that type and density of TILs can be a more powerful prognostic indicator than previous pathological markers (Schreiber et al., 2011).

Remarkable positive correlation was demonstrated for CD8<sup>+</sup> TILs, for example in: colon cancer (Straeter et al., 2005; Naito et al., 1998; Kondo et al., 2003), esophageal cancer (Yasunaga et al., 2000), breast cancer (Yoshimoto et al., 1993), ovarian cancer (Sato et al., 2005), malignant melanoma (Haanen et al., 2006), and several other cancers (reviewed in Senovilla et al. (2012)).

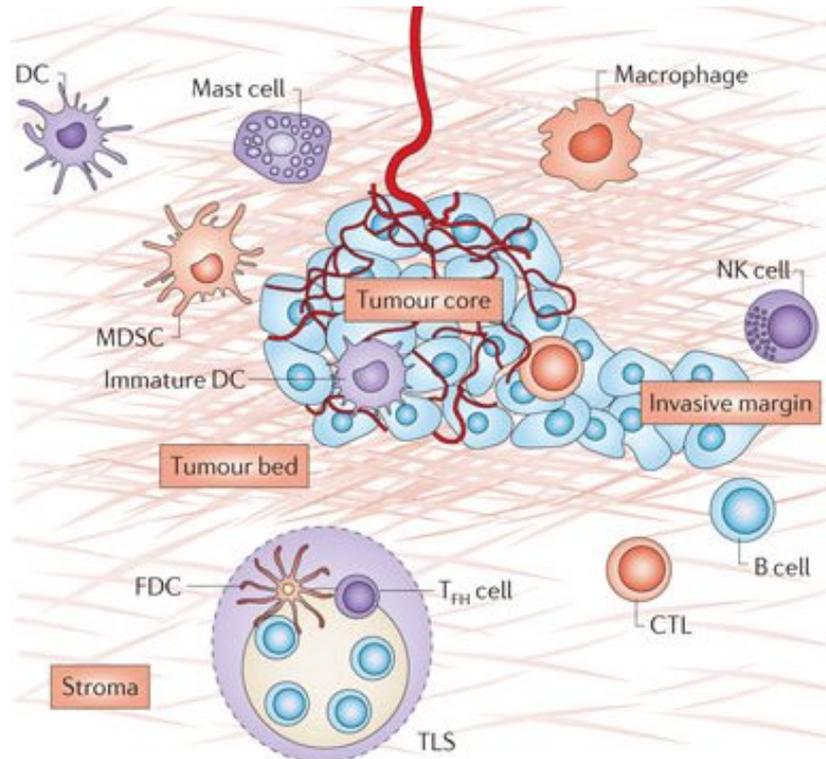
Similar to CD8<sup>+</sup> cytotoxic TILs, CD45RO<sup>+</sup> memory T cells have clearly been associated with better prognosis (Fridman et al., 2012). In contrast, the CD4<sup>+</sup> TILs T<sub>reg</sub> and Th1, Th2, and Th17 T helper cells have been shown to have contradictory effects on clinical outcome among different cancers (Fridman et al., 2012). T<sub>reg</sub> T cells express phenotypic markers such as CD25 and FOXP3. Correlation of intratumoral T<sub>reg</sub> T cells and poor survival was shown for ovarian cancer, breast cancer and hepatocellular carcinoma underlining the suppressive function of T<sub>reg</sub> cells. However, for a number of other cancers no correlation of T<sub>reg</sub> cells and clinical outcome was found, and for other cancers, even an inverse correlation was reported (reviewed in Fridman et al. (2012)). Similar contradictory results were reported for Th17, and Th2 T helper cells. Merely, infiltration with Th1 T helper cells is associated with good clinical outcomes for all cancer types (Fridman et al., 2012). This discrepancies may reflect that tumors that grow in different organs have different phenotypes and that each tumor has a unique microenvironment (Fridman et al., 2012).

However, even when a tumor is highly infiltrated with lymphocytes, their function is often impaired, due to modifications in the complex interplay of immune cells and the expression of cytokines and their receptors (Fridman et al., 2012).

## Cytokines and Chemokines

Cytokines are secreted or membrane-bound proteins that can influence and regulate the function of immune cells (Dranoff, 2004). Chemokines are a subfamily of chemotactic cytokines that regulate leukocyte migration (Cameron and Kelvin, 2003). Cytokines can directly stimulate or suppress immune effector cells in the tumor microenvironment, and cancer cells often initiate changes in local cytokine expression, disrupting the complex interplay between cytokines and immune cells (Fridman et al., 2012).

The different types of CD4<sup>+</sup> T helper cells secrete different combinations of cytokines which mainly drive their effector functions. Th1 cells, which are known to be antitumorigenic, mainly secrete IL-2, interferon- $\gamma$  (IFN $\gamma$ ), and tumor necrosis factor- $\beta$  (TNF $\beta$ ), cytokines which are associated with inflammation and cytotoxicity (Cameron and Kelvin, 2003; Senovilla et al., 2012). Th2 cells mainly secrete IL-4, IL-5, IL-6, IL-10, IL-13, and transforming growth factor- $\beta$  (TGF $\beta$ ), which serve to downregulate inflammatory actions (Cameron and Kelvin, 2003).



**Figure 1.9: Elements of the tumor microenvironment.** The tumor with its invasive margin is shown, together with tertiary lymphoid structures (TLS) and the cells of the tumor microenvironment. CTL, cytotoxic T lymphocyte; DC, dendritic cell; FDC, follicular dendritic cell;  $\text{IFN}\gamma$ , interferon- $\gamma$ ; IL-12, interleukin-12; IM, invasive margin; IRF1, interferon regulatory factor 1; MDSC, myeloid-derived suppressor cell; NK cell, natural killer cell;  $T_H$ , T helper;  $T_{Reg}$  cell, regulatory T cell. Figure adapted from Fridman et al. (2012) and reprinted with permission.

Secretion of cytokines stimulate other immune-cell infiltrates, which in turn release additional cytokines. Hence, the composition of cytokine-secreting cells in the tumor microenvironment is critical. For example, several chemokines are secreted by cancer-associated fibroblasts, such as CCL2, CCL3, CCL4, and CCL5, which influence the macrophage composition in the tumor microenvironment (Turley et al., 2015). Macrophages in turn are key producers of a number of cytokines which are correlated with  $\text{CD8}^+$  cytotoxic T cell infiltration. Other chemokines which have been correlated with a clinical benefit are CXCL1, which may attract Th1 cells, as well as CXCL9 and CXCL10, which may attract memory T cells (Mlecnik et al., 2010).

## 1.4 Cancer Immunotherapy

The idea of harnessing the immune system to fight cancer is not new, but it was considered highly controversial among scientists. Today, there is no doubt that the immune system has the capacity to recognize and eradicate cancer, and the development of methods to induce and manipulate effective antitumor responses is gaining

great momentum. An underlying basis of cancer immunotherapy is the activation of the patient's own immune effector cells to specifically eradicate tumor cells. Extensive research of tumor biology and immunology has led to a better understanding of tumor-host interactions and immunotherapies are now being rationally designed. Immunotherapeutic approaches are versatile, including adoptive T cell therapy, gene-engineered T cells, immune checkpoint blockade, and therapeutic cancer vaccines. These approaches are summarized in this section, with a special emphasis on neoantigens and cancer vaccines.

Adoptive T cell therapy relies on existing tumor-reactive T cells which are isolated from cancer patients, expanded *in vitro*, and reintroduced to the patient. The idea behind this approach is that these TILs are located in high densities within the tumor, but their antitumor effects are impaired by an immunosuppressive microenvironment. Removing these T cell populations from that suppressive environment enables their activation and clonal expansion (Restifo et al., 2012). An additional advantage of adoptive T cell transfer is that the patient can be prepared before the actual transfer to provide a favorable microenvironment for the transferred antitumorigenic T cells (Rosenberg and Restifo, 2015). So-called lymphodepletion can be performed immediately before T cell transfer: the patient's immune system is temporarily depleted using chemotherapy and total-body irradiation, which then leads to an improved persistence and antitumor activity of transferred T cells (Rosenberg and Restifo, 2015). It was demonstrated that this adoptive T cell transfer protocols achieve objective response rates of almost 48% in metastatic melanoma, and a fraction of patients is even considered as 'cured' from metastatic melanoma (Restifo et al., 2012).

It seems however, that the usage of naturally occurring TILs for adoptive T cell therapy is only effective in melanoma, for reasons that are not fully understood (Restifo et al., 2012). Therefore, genetically-engineered T cells are being developed to be used for adoptive T cell therapy in different cancers. T cells can be genetically engineered to express T cell receptors that specifically target tumor antigens. A promising technique for T cell gene-engineering is the usage of chimeric antigen receptors (CARs). CARs are fusion proteins consisting of the antigen-binding domain of an antibody fused with the signalling component of the TCR complex. Retroviruses or lentiviruses encoding the CAR can be used to transduce T cells. Transduced T cells can then recognize the target antigen on the surface of tumor cells with the antibody part of their CAR, and do not depend on a functional antigen processing machinery and antigen presentation on MHC (Restifo et al., 2012). CARs have been used successfully with very promising clinical results for the treatment of hematologic cancers, for example B cell lymphoma (target antigen CD19) (Kochenderfer et al., 2010), the ability to treat epithelial solid cancers however, is limited (Rosenberg and Restifo, 2015).

Immune checkpoint therapy is a treatment approach that radically changed the principals of cancer treatment, as cancer cells are not targeted directly, but molecules regulating T cell activity are targeted (Sharma and Allison, 2015). Tumor cells exploit immune inhibitory pathways to escape T cell-mediated eradication. Checkpoint blockade therapies attempt to disrupt these pathways in order to remove the inhibitory signals that block an effective antitumor T cell response (Sharma and Allison, 2015).

This approach is often referred to as "releasing the brakes" on T cells, enabling them to fight malignancy (Weintraub, 2013), and is also considered one of the most promising approaches for tackling solid tumors (Bordon, 2015). CTLA-4 is the first immune checkpoint that was clinically targeted. Ipilimumab, a fully humanized antibody blocking CTLA-4, entered clinical trials in late 1990s and tumor regression was observed in patients with various tumor types: melanoma, renal cell carcinoma, prostate cancer, urothelial carcinoma, and ovarian cancer (Sharma and Allison, 2015). A phase III clinical trial in patients with advanced melanoma demonstrated improved overall survival, and importantly, durable response was observed in about 20% of the patients (Hodi et al., 2010; Robert et al., 2011; Schadendorf et al., 2015). Ipilimumab was the first checkpoint blocking antibody that was approved by the United States Food and Drug Administration (U.S. FDA). It was approved for the treatment of melanoma in 2011. Tremelimumab is another CTLA-specific antibody that is in development and undergoing clinical trials at the moment (Ribas et al., 2013).

The PD-1 pathway additionally emerged as a promising target for immune checkpoint therapy. Antibodies that inhibit the interaction between the receptor PD-1 and its ligands PD-L1 and PD-L2 have shown promising results in various tumor types. Clinical trials with anti-PD-1 antibodies nivolumab and pembrolizumab demonstrated high clinical response rates for metastatic melanoma, and it was also shown that there is a lower incidence of adverse events compared to treatment with ipilimumab. Clinical efficacy of anti-PD-1 treatments was also shown to be effective in various cancers, including renal cell carcinoma, non-small cell lung cancer (NSCLC) and hematological cancers. Similar to ipilimumab treatment, the clinical response of anti-PD-1 treatments were often shown to be durable. After pembrolizumab treatment yielded an overall response rate of 26% in advanced melanoma (Robert et al., 2014), it was approved by the U.S. FDA for treatment of metastatic melanoma in 2014, and was also approved for NSCLC just recently in October 2015. Nivolumab showed similar clinical benefits (Sharma and Allison, 2015) and received FDA Approval for melanoma, renal cell carcinoma, NSCLC, and squamous NSCLC.

As different checkpoint molecules regulate distinct pathways, combination therapies are of special interest. It was for instance demonstrated in a clinical trial in advanced melanoma that a combination treatment with ipilimumab and nivolumab results in improved clinical benefit when compared to the single agent treatments (Wolchok et al., 2013). Other combination treatments that enable the blockade of multiple inhibitory pathways such as LAG-3 or TIM-3, or combinations of blockade of inhibitory pathways while promoting stimulatory pathways such as OX40, are also under development (reviewed in Melero et al. (2015) and Mahoney et al. (2015)).

Several reports have suggested that cancers with a high mutational load respond best to immune checkpoint therapy (Bordon, 2015). Snyder et al. reported that a high mutational load correlated with an improved response to CTLA-4 blockade (Snyder et al., 2014), and Kvistborg et al. demonstrated that ipilimumab treatment induced a significant number of newly detected T cell responses (Kvistborg et al., 2014). More recently, a similar correlation was reported for the number of neoantigens and the response to pembrolizumab in patients with NSCLC (Rizvi et al., 2015).

These reports highlight that the presence and abundance of mutations and neoantigens they give rise to are crucial to the effect of immunotherapies. An explanation for these observations is that mutations may give rise to neoantigens that can be targeted by T cells following their release from inhibition (Bordon, 2015).

### 1.4.1 Neoantigens in Cancer Immunotherapy

Although checkpoint blockade therapy showed great clinical results and induced durable results in many patients, it only works in a subset of patients. It was suggested that these patients lack preexisting antitumor T cell responses, and that this unfavorable situation can be addressed with therapeutic antitumor vaccination (Delamarre et al., 2015). Neoantigens are of special interest in this context. They are superior to other tumor-associated-antigens, because the immune system was not tolerized against them during development, and their expression is highly tumor-specific (Hacohen et al., 2013; Heemskerk et al., 2013). It has been shown in several studies that neoantigens are not only important targets of checkpoint blockade therapy, but they can also be used for personalized therapeutic vaccines. There is now evidence, showing that CD8<sup>+</sup> T cells that are specific for neoantigens can attack tumor cells in various cancers (Rizvi et al., 2015; Rooney et al., 2015; Snyder et al., 2014). The goal of therapeutic vaccination is to enlarge the pool of tumor-specific T cells from the naïve repertoire, and also to reactivate existing tumor-specific T cells that are in a dormant or anergic state (van der Burg et al., 2016).

Coulie et al. suggested the following scenario for vaccinated patients who show tumor regression (Coulie et al., 2014): upon vaccination, a small number of T cells that are specific against the vaccine-neoepitopes penetrate the tumor, detect their cognate antigen and eventually kill some tumor cells. As a result, these cytotoxic T cells are re-stimulated and produce cytokines that focally reverse the local immunosuppressive environment around. In this now immunostimulatory environment many of the inactive antitumor T cells that are already present in the tumor are reawakened, and new naïve antitumor T cells are additionally stimulated. Of note, the newly stimulated T cells may be directed against various tumor antigen other than the vaccine antigen, which then provides the numbers of T cells that are required to reject a tumor. These active T cells possibly can also migrate to other tumor sites and trigger a response there (Boon et al., 2006). Hence, therapeutic vaccination can not only recruit tumor-specific T cells from the naïve T cell repertoire, but can also reactivate T cells with an existing reactivity against various antigens, which have been in a dormant or anergic state prior to vaccination (van der Burg and Melief, 2011).

However, the development of personalized cancer-specific vaccines for application in patient care still faces several challenges. As generally only few mutations are shared between patients and the bulk of mutations are patient-specific, one major issue is the selection of which neoepitopes to include in a therapeutic vaccine. With recent advances in next-generation-sequencing technologies and the availability of immunoinformatics tools, it has become feasible to detect potential neoantigens in a tumor sample. In brief, after whole-genome, whole-exome, and/or whole-transcriptome sequencing, all

somatic mutations that lead to the formation of novel protein sequences are detected, and immunoinformatics tools are applied to assess their potential immunogenicity. The resulting set of potential neoepitopes can then be used to query T cell reactivity against them. Hence, neoantigen-specific vaccine development requires a personalized approach based on the genomic information of each individual tumor (Schumacher and Schreiber, 2015). Studies have shown that this approach can successfully be exploited in a clinical setting (Schumacher and Schreiber, 2015).

Another question for the development of neoantigen-based therapeutic vaccines is what vaccine platform to use for delivering the neoepitopes to the patient. Several strategies have been proposed including dendritic cells, synthetic peptides, and recombinant DNA and RNA. Dendritic cell vaccines have the superior capacity to induce effective T cell responses, and studies have reported successful application of them (Palucka and Banchereau, 2013; Carreno et al., 2015). However, implantation of dendritic cell vaccines is highly time and cost consuming. A simpler approach is the usage of long synthetic peptides or RNA stretches that encode selected neoantigens, which is also a more feasible approach to implement in the clinic (Delamarre et al., 2015). Peptide and RNA-based vaccination was shown to elicit potent antitumor responses in mouse models (Kreiter et al., 2015; Yadav et al., 2014). Synthetic long peptides may especially be feasible for vaccination purposes, because they mimic the antigenic properties of a whole protein (van den Boorn and Hartmann, 2013) and they are usually easy to produce. Additionally, they can be designed in a way that they contain both, neoepitopes for CD4<sup>+</sup> and CD8<sup>+</sup> T cells. If a therapeutic vaccine is going to be administered, including multiple neoepitopes for both, CD4<sup>+</sup> and CD8<sup>+</sup> T cells, is of advantage. In doing so, the likelihood of generating an immune response against at least some of the neoantigens increases, and the likelihood of the tumor escaping the immune response by immunoediting decreases.

The selected adjuvant included in a vaccine also plays a major role for therapeutic efficacy, as they strongly influence the immunogenicity of the vaccine. Adjuvants stimulate the immune system in a more general way and are supposed to enhance the immune response induced by the administered vaccine. This is of special interest for peptide vaccines, as some immunogenicity might be lost by using only the epitope part of the antigen and not the whole antigen (Banday et al., 2015). Adjuvants used in cancer vaccines include mineral salts such as Alum, oil-in-water emulsions such as Montanide, and portions of pathogens such as Lipid-A or Bacille Calmette-Guerin (BCG). Many of these adjuvants activate cells of both, innate and adaptive immunity, such as DCs and macrophages, by activating toll-like-receptors (TLRs) and enhancing the processing and presentation of antigens to T cells. In addition, some non-specific immune adjuvants such as interferons (IFN- $\alpha$ , IFN- $\beta$  and IFN- $\gamma$ ), interleukins (IL-2, IL-7, IL-12 and IL-21) and GM-CSF have also been used as adjuvants in various cancer vaccines which are in early clinical stages (Banday et al., 2015). Although long-peptide vaccines in montanide adjuvant have been effective in clinical trials (van den Boorn and Hartmann, 2013), it was also reported that oil-in-water emulsions adjuvants cause primed T cells to become sequestered at the vaccination site rather than tumor site, and that the injection site turns into a "graveyard" for T cells (Hailemichael and Overwijk, 2014;

Hailemichael et al., 2013). However this observation was limited to short peptides and usage of long peptides induced minimal T cell sequestration at the vaccination site and superior antitumor activity (Hailemichael and Overwijk, 2014).

Several neoepitope-based anti-cancer vaccines have reached clinical stage. A vaccine based on a single neoepitope from IDH1 has been shown to induce antitumor immunity mouse models by Platten et al. (Schumacher et al., 2014) and the vaccine is currently undergoing a phase I clinical trial for grade III-IV gliomas (NCT02454634). Another concept is to vaccinate with multiple shared neoepitopes known for a distinct tumor entity (NCT01885702). Personalized neoepitope-targeted trials using either mutant RNA or mutant peptides as antigens are currently conducted in patients with melanoma (NCT02035956, NCT01970358) glioblastoma (NCT02149225, NCT02510950, NCT02287428) and triple-negative breast cancer (NCT02348320, NCT02316457) (Vormehr et al., 2015). These vaccine trials will not only provide data about efficacy and safety, but will also provide data for validation and evaluation of *in silico* epitope prediction and selection criteria and may increase the accuracy and efficacy of them.

## 1.5 Colorectal Cancer

Colorectal cancer (CRC) is a major cause of cancer morbidity and mortality worldwide and remains the fourth common cause of death from cancer (Fearon, 2011; Singh et al., 2015). In this section, treatment approaches for CRC, especially immunotherapeutic strategies, are described, as well as the current understanding of the tumor microenvironment of CRC.

### 1.5.1 Treatment Approaches for Colorectal Cancer

The most common treatment for CRC is surgery. In the case of localized tumors which have not spread outside the colon wall itself (stage I), surgery may completely eliminate the cancer. When the cancer has invaded the bowel wall and nearby tissue (stage II) or has invaded nearby lymph nodes (stage III), chemotherapy and radiation therapy may be administered before or after surgically removing all affected areas together with the corresponding lymph nodes. Chemotherapy regimens used for CRC are fluoropyrimidine- and oxaliplatin-based. A diagnosis of stage I-III CRC allows patients to undergo treatment with a curative intent (Ahn and Goldberg, 2016). A great proportion of patients, more than 20%, are diagnosed with distant metastases (stage IV), most commonly in the liver or lung (Brenner et al., 2014). Surgery is often not applicable for these patients with metastatic disease (Singh et al., 2015). Palliative chemotherapy for patients with metastatic CRC can improve survival and downsize metastases in patients with potentially resectable disease (Cunningham et al., 2010). Although significant progress has been made in the treatment of metastatic CRC, the 5-year survival remains <12.5% (Siegel et al., 2014), highlighting the need for the development of new treatment approaches (Ahn and Goldberg, 2016).

Immunotherapy for CRC has been under investigation for many years, but showed limited clinical efficacy. Many recent scientific approaches broadened our understanding of the antitumor immune response, so that immunotherapy may also represent a novel effective treatment approach for metastatic CRC. Current immunotherapies for CRC include: monoclonal antibodies, checkpoint inhibitors, cancer vaccines, adoptive cell therapy and oncolytic virus therapy.

The earliest attempts of CRC immunotherapy were in the form of whole-tumor-cell vaccination, where autologous tumor cells are lysed and re-infused into the patient together with an immune adjuvant. Several clinical trials have been conducted which however, showed limited clinical efficacy (reviewed in Markman and Shiao (2015)). Whole-tumor vaccines are typically poorly immunogenic as, only a small number of the present antigens are tumor-specific, while a vast majority of antigens in the vaccine are shared with normal cells, thus diluting the amount of tumor-specific antigens (Markman and Shiao, 2015). To circumvent this problem, peptide vaccines were developed in order to generate an immune response against a specific known tumor antigen. Peptide vaccines incorporate whole proteins or fragments of proteins that are administered together with adjuvants. In CRC, multiple tumor-associated antigens have been identified and utilized for vaccination, including carcino-embryonic antigen (CEA), MUC-1, and survivin (Markman and Shiao, 2015). While early trials with peptide vaccines demonstrated low efficacy due to limited immunogenicity and small numbers of responding T cells, new strategies to simultaneously boost the immune response show promise (Markman and Shiao, 2015). One of the more promising approaches to boost the immune activation is packing the tumor antigens plus co-stimulatory whole proteins or fragments of proteins that are administered together with adjuvants. Clinical trials have shown that these viral vaccines produce significantly more effective responses compared to peptide vaccines, however toxicity in form of a cytokine storm was also reported in several cases (Markman and Shiao, 2015). Another vaccination approach for CRC is the usage of dendritic cells. Several clinical trials with DCs that have been pulsed with tumor-specific antigens have been conducted. Since CEA is expressed in most CRCs, this antigen was utilized for DC vaccines for CRC (Lesterhuis et al., 2006; Morse et al., 2013). Although CEA-specific T cell responses could be induced with these DC vaccines, survival was not improved significantly (Markman and Shiao, 2015).

Adoptive cell therapies have also shown some activity in CRC immunotherapy. For instance, autologous T cells genetically engineered to express a high-affinity murine T cell receptor against human CEA were administered to refractory CRC patients, and objective regression was observed in one patient. However, because patients developed immune colitis, further enrollment of patients was stopped (Parkhurst et al., 2011). Various trials are in progress using T cells specific for the tumor-associated antigens NY-ESO-1, MAGEA4, PRAME, Survivin, and SSX in refractory solid tumors, including CRC (NCT02239861).

The main advance in the treatment of metastatic CRC has been the addition of targeted therapies (Cunningham et al., 2010). The monoclonal antibodies bevacizumab (anti-VEGF), panitumumab and cetuximab (anti-EGFR), and regorafenib, dabrafenib,

and neratenib (kinase inhibitors) showed some clinical efficacy as combination therapies with chemotherapy (Brenner et al., 2014; Ahn and Goldberg, 2016). Recently, the monoclonal antibody against PD-1, pembrolizumab, created some excitement in the field of CRC immunotherapy. As mentioned in section 1.4, immune checkpoint therapy showed great success in a number of patients and cancers. For a long time, these checkpoint blockade therapies seemed useless in CRC. Early data suggested limited efficacy for anti-CTLA-4 treatment in CRC (Chung et al., 2010), and because low expression of PD-1 was demonstrated in CRC, it was also suggested that anti-PD-1 or anti-PD-L1 therapy might also be inefficient in CRC (Taube et al., 2014). A large phase I trial evaluating safety and activity of an anti-PD-1 antibody enrolled patients with NSCLC, prostate, renal-cell and colorectal cancer (Topalian et al., 2012). In this trial, only one patient with metastatic CRC showed response after treatment. Further investigation of this patient revealed that this patient was mismatch-repair deficient. Mismatch-repair deficiency leads to the instability of the microsatellites, and as a result, these tumors contain 10-100 times more somatic mutations than mismatch-repair proficient tumors (Topalian et al., 2012). It was hypothesized that this high mutational load triggers response from T cells that have been blocked by PD-1 prior to therapy. To test this hypothesis, Le et al. conducted a phase II trial to assess the efficacy of the anti-PD-1 antibody pembrolizumab in mismatch-repair deficient and proficient metastatic CRC patients (Le et al., 2015). The hypothesis was clearly confirmed in this trial: patients with mismatch-repair deficiency demonstrated an objective response rate of 40%, whereas no objective response was seen in mismatch-repair proficient patients.

Based on the promising clinical activity of pembrolizumab, several ongoing studies are investigating various immunotherapeutic agents in several subsets of CRC (reviewed in Ahn and Goldberg (2016)). It remains to be elucidated, which subsets of CRC patients respond to certain immunotherapies and predictive biomarkers have to be established.

## 1.5.2 The Tumor Microenvironment in Metastatic Colorectal Cancer

The first evidence for a correlation of immune infiltration and survival in CRC patients was reported in 1987 (Jass et al., 1987). The prognostic value of TILs in CRC was confirmed by Ropponen et al. (Ropponen et al., 1997): they quantified the TILs in the tumor stroma and along the invasive margin and demonstrated that the grade of infiltrating lymphocytes was a predictive factor for disease free and overall survival. Another study in colon cancer, showed that particularly CD8<sup>+</sup> cytotoxic TILs have an important influence on clinical outcome (Naito et al., 1998), and it became evident that the specific TIL composition has a crucial role in clinical evolution of CRC (de la Cruz-Merino et al., 2011). With the emergence of more sophisticated methodology to delineate the effects of different immune cell populations and their spatial distribution within and around cancer lesions, large studies have been conducted and identified the prognostic role of T cell infiltrates in the center and the invasive margin of the primary tumor (Halama et al., 2012; Galon et al., 2006). It was furthermore shown that infiltrating immune cells are heterogeneous and that they accumulate in distinct

regions and compartments, namely core of the tumor, invasive margin, and adjacent healthy tissue (Halama et al., 2009a, 2012).

## 1.6 Motivation

It is now commonly known that the composition of the tumor microenvironment has a major impact on tumor growth and progression, as well as on response to therapy. It has been shown that the type and density of tumor-infiltrating immune cells are highly predictive for disease outcome in various cancers. These studies have also suggested that a high density of tumor-infiltrating lymphocytes is correlated with mutational load. One hypothesis in this context is that somatic mutations found in cancer cells may give rise to neoepitopes that attract and keep lymphocytes at the tumor site (Schumacher and Schreiber, 2015). Neoepitopes have also been suggested to be crucial for the outcome of immune checkpoint therapies, as it was reported that cancers with a high mutational load respond best to checkpoint therapy (Delamarre et al., 2015). An explanation for this is that mutations give rise to neoepitopes that can be targeted by T cells once they are released from inhibitory signals.

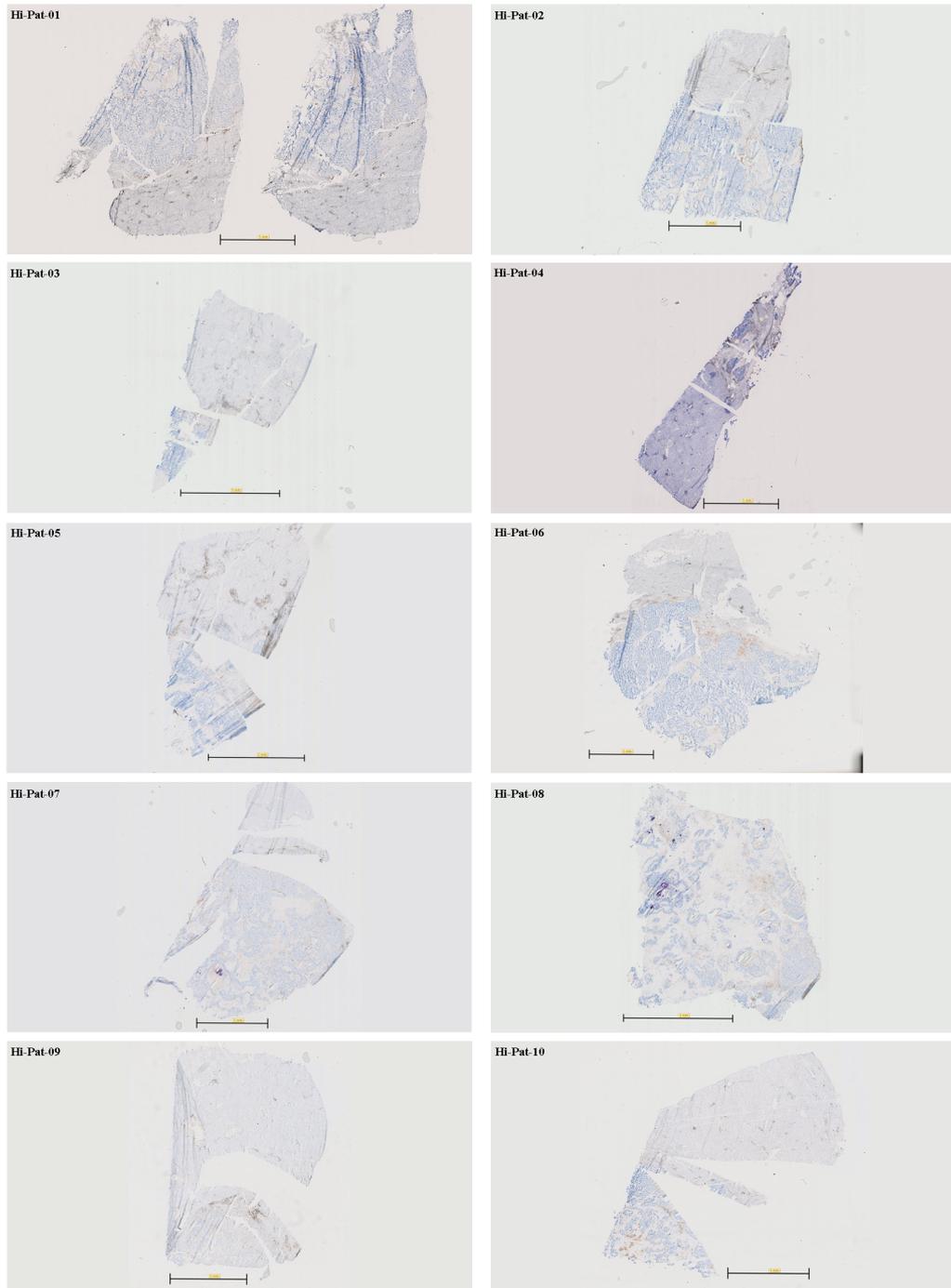
It has now become evident that effective immunotherapies have to be tailored to the specific immune setting of each tumor (Smyth et al., 2016). The complex interplay between the tumor and the immune system has to be systematically analyzed for characterizing patients and identifying therapies they will most likely benefit from. This highly personalized approach requires the integrated analysis of numerous tumor and host factors. Accordingly, the main aim of this PhD project was the establishment of an integrated analysis pipeline to obtain detailed data about tumor-host interactions, including analysis of the mutational and neoepitope load, the type and densities of tumor infiltrating immune cells, the expression of immunological markers, and the expression of specific cytokines.

This analysis pipeline combines already available genomic and immunomic resources and adds further depth into the analysis by additional computational pipelines. The already well established sequencing and somatic mutation detection pipelines that have been developed in the DKFZ bioinformatics departments (Prof. Roland Eils and Prof. Benedikt Brors) were integrated with the cytokine profiling and histological analysis workflows in Professor Jäger's group (NCT, Medical Oncology). Additional computational pipelines for HLA genotyping from sequencing data, as well as for epitope predictions for HLA class I and class II were implemented and included. Taken together these analysis pipelines provide a broad picture of tumor-host interactions. The established analysis pipeline allows the rapid and systematic analysis of large patient cohorts.

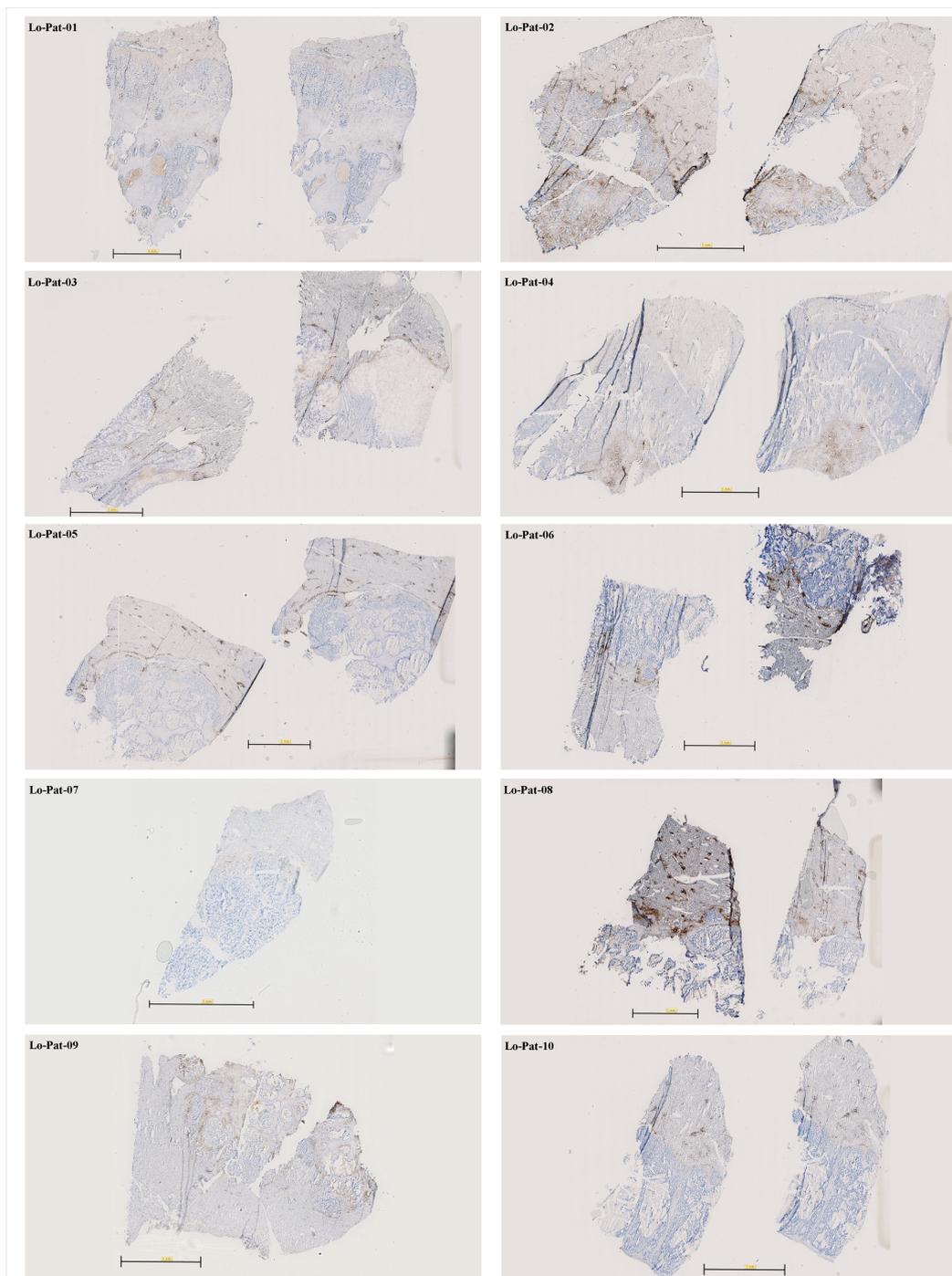
Professor Jäger's group has been collecting CRC liver metastases and systematically characterizing their immune infiltration and cytokine profiles, as well as the correlation to clinical outcome (Halama et al., 2011a,b, 2009a, 2010, 2009b). In these studies it was shown that in general, there are at least two patient groups for each CRC stage: patients with high infiltrate density and patients with low infiltrate density,

with the latter having a much worse prognosis (Halama et al., 2012). A patient cohort including 10 patients with high densities of infiltrating lymphocytes (TIL-high) and 10 patients with low densities (TIL-low) was assembled and provided by Niels Halama from Professor Jäger's group for analysis in this study (section 2.1, Figures 1.10 and 1.11). The described integrated analysis pipeline was developed using this patient cohort.

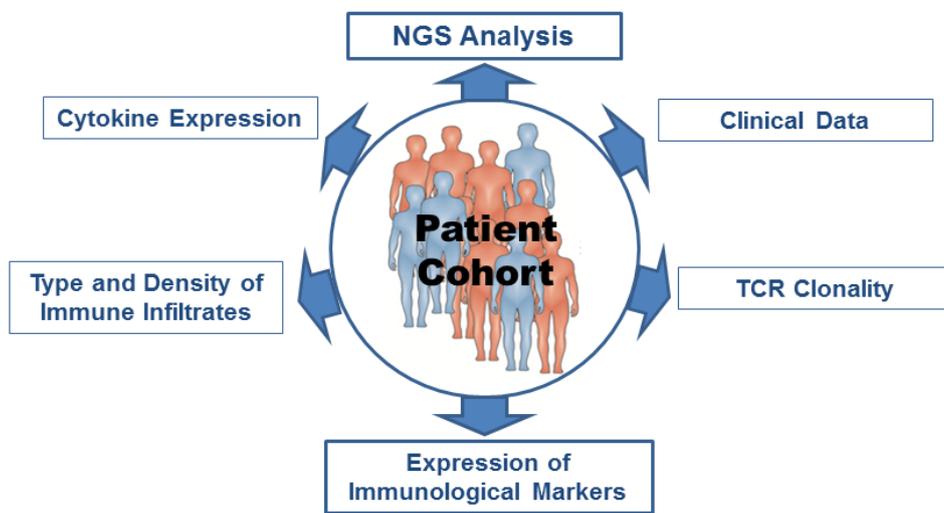
The analysis pipeline was then used to systematically investigate TIL-high versus TIL-low CRC metastases in order to assess the correlation of mutational and neoepitope load to lymphocyte infiltration and whether additional factors distinguishing the two groups can be discovered. The results of this analysis provide valuable detailed insights to the immune setting of CRC liver metastases, which are presented and discussed in this thesis. The pipeline was also applied in the clinic to conduct case studies with several patients being treated at the NCT. Two case studies are also exemplarily presented in this thesis.



**Figure 1.10: Immunohistochemical stainings for CD3 in samples with high lymphocyte infiltration.** Tissue sections of all 20 patients were immunohistochemically analyzed for their infiltration with T cells by staining the T cell marker CD3. The stainings were evaluated semi-automatically and the samples were assigned to be highly or lowly infiltrated accordingly. Each panel shows the CD3 staining of tissue derived from a TIL-high patient, and the scale bar indicates 5mm. Samples were provided by Niels Halama (NCT, Medical Oncology).



**Figure 1.11: Immunohistochemical stainings for CD3 in samples with low lymphocyte infiltration.** Tissue sections of all 20 patients were immunohistochemically analyzed for their infiltration with T cells by staining the T cell marker CD3. The stainings were evaluated semi-automatically and the samples were assigned to be highly or lowly infiltrated accordingly. Each panel shows the CD3 staining of tissue derived from a TIL-low patient, and the scale bar indicates 5mm. Samples were provided by Niels Halama (NCT, Medical Oncology).



**Figure 1.12: Integrative Analysis Approach.** The established integrated analysis workflow provides a broad picture of the tumor-host interactions by combining data from analysis of the mutational and neoepitope load, the type and densities of tumor infiltrating immune cells, the expression of immunological markers, and the expression of specific cytokines.

## CHAPTER 2

---

### Methods

---

Samples were collected and provided for this study by Niels Halama from Professor Jäger’s group (NCT, Medical Oncology) in collaboration with the Surgery Department (University Hospital Heidelberg). All experimental procedures were conducted by members of Professor Jäger’s group (NCT, Medical Oncology). I implemented the HLA Typing and Immunoinformatics pipelines, as well as several analysis scripts to evaluate the obtained data.

## 2.1 Samples

**Table 2.1:** Patient IDs and corresponding TIL status

Patient ID	TIL status	ID	Gender	Age
Hi-Pat-01	high	HLM 64	m	64
Hi-Pat-02	high	HLM 61	m	61
Hi-Pat-03	high	HLM 39	m	81
Hi-Pat-04	high	HLM 50	w	67
Hi-Pat-05	high	HLM 54	w	70
Hi-Pat-06	high	HLM 30	m	72
Hi-Pat-07	high	HLM 55	m	54
Hi-Pat-08	high	HLM 56	m	71
Hi-Pat-09	high	HLM 31	w	68
Hi-Pat-10	high	HLM 71	m	48
Lo-Pat-01	low	HLM 37_1	m	72
Lo-Pat-02	low	HLM 69	m	72
Lo-Pat-03	low	HLM 52	m	68
Lo-Pat-04	low	HLM 84	w	50
Lo-Pat-05	low	HLM 8-3	w	49
Lo-Pat-06	low	HLM 51	m	69
Lo-Pat-07	low	HLM 32B	m	63
Lo-Pat-08	low	HLM 33	m	65
Lo-Pat-09	low	HLM 67	w	63
Lo-Pat-10	low	HLM 73B	m	79

TIL = tumor-infiltrating lymphocyte

Samples from colorectal cancer liver metastases from a cohort of patients with incomplete resection following palliative chemotherapy was collected previously by Niels Halama et al. (NCT, Medical Oncology, Jäger) in collaboration with the Surgery Department (University Hospital Heidelberg). Tumor tissue from the metastases, as well as tissue from the adjacent healthy liver were provided as fresh frozen tissue blocks.

Lymphocyte infiltration densities of the provided samples have already been determined using immunohistochemistry and virtual microscopy for quantification, and 10 patients with high and low TIL density were selected, respectively. An overview of the

patient IDs together with corresponding TIL status is presented in Table 2.1. Complete clinical data for this patient cohort is available.

### 2.1.1 Case Studies

Two patients with metastatic CRC, who are currently being treated at the NCT and enrolled in the NCT Precision Oncology Program (NCT POP), were analyzed additionally to explore further immunotherapeutic options.

## 2.2 Histological Analysis

**Table 2.2:** Primary antibodies used for immunohistochemical stainings

<b>Mouse monoclonal antibodies</b>			
<b>Protein</b>	<b>Lot number</b>	<b>Supplier</b>	<b>Clone</b>
PDL2	#345503	BioLegend	MIH18
PDL1	#329710	BioLegend	29E. 2A3
PD1	#ab52587	Abcam	NAT 105
FoxP3	#14-4777	Bioscience	236A/E7
CTLA4	#70-1529	Tonbo Biosciences	BNIS
CD3	#NCL-L-CD3-565	Novocastra	LN10
CD8	#NCL-L-CD8-4B11	Novocastra	4B11
CD20	#NCL-L-CD20-L26	Novocastra	L26
CD68	#ab955	Abcam	KP1
CD163	#MCA 1853	AbD Serotec	EDHu-1
Ki67	#M7240	Dako	MIB-1
HLA I	#D226-3	MBL	EMR8-5
NKP46	#MAB1850	R&D	195314
<b>Rabbit polyclonal antibodies</b>			
<b>Protein</b>	<b>Lot number</b>	<b>Supplier</b>	<b>Clone</b>
aCasp3	#ab2302	Abcam	-

Immunohistochemical stainings were performed by Rosa Eurich in Professor Jäger’s group (NCT, Medical Oncology), as previously described (Halama et al., 2011a).

Tissue sections of all 20 patients were immunohistochemically analyzed for their infiltration with T cells (CD3 and CD8), regulatory T cells (Foxp3), B cells (CD20), NK cells (NKP46), and macrophages (CD68, CD163). The expression of HLA class I and the immunomodulatory proteins PD-1, PD-L1, PD-L2, CTLA-4 was also evaluated immunohistochemically. The used antibodies are listed in Table 2.2. Nuclei were counterstained with hematoxylin. All processing steps were performed with a BOND-II autostainer (Leica) according to manufacturers recommendations. Antigen

detection was performed by a color reaction with 3,3-diamino-benzidine (DAB chromogen, Menarini).

High-precision quantification of cell T cell densities (CD3 and CD8 stainings) was performed as described previously (Halama et al., 2010, 2009b) using the VIS software package (Visiopharm). All other stainings were assessed manually .

## 2.3 Cytokine Profiling

**Table 2.3:** Quantified cytokines

<b>Protein</b>	<b>Gene Symbol</b>	<b>Protein</b>	<b>Gene Symbol</b>
Eotaxin	CCL11	IL-16	IL16
MCP-1	CCL2	IL-17	IL17A
CTACK	CCL27	IL-18	IL18
MIP-1a	CCL3	IL-1a	IL1A
MIP-1b	CCL4	IL-1b	IL1B
RANTES	CCL5	IL-2Ra	IL1RN
MCP-3	CCL7	IL-1ra	IL2RA
GROalpha	CXCL1	IL-2	IL2
IP-10	CXCL10	IL-3	IL3
SDF-1a	CXCL12	IL-4	IL4
MIG	CXCL9	IL-5	IL5
SCGF-b	CLEC11A	IL-6	IL6
M-CSF	CSF1	IL-7	IL7
GM-CSF	CSF2	IL-8	IL8
G-CSF	CSF3	IL-9	IL9
FGF basic	FGF2	SCF	KITLG
HGF	HGF	LIF	LIF
VEGF	VEGFA	TNF-beta	LTA
IFN-alpha2	IFNA2	MIF	MIF
IFN-g	IFNG	b-NGF	NGF
IL-10	IL10	PDGF bb	PDGFB
IL-12p70	IL12A	TNF-a	TNF
IL-12p40	IL12B	TRAIL	TNFSF10
IL-13	IL13		
IL-15	IL15		

Chemokine and cytokine detection in tissue lysates were performed by Tina Lerchl in Professor Jäger’s group (NCT, Medical Oncology), and the three compartments adjacent liver (AM), invasive margin (IM), and liver metastasis (LM) were considered separately.

Tissue lysates were prepared from frozen material according to the manufacturers instructions (BioRad Laboratories). A Luminex100 reader was used to simultaneously quantify a panel of 48 cytokines (BioRad Laboratories, Bio-Plex Pro Human Cytokine 27-plex (#M500KCAF0Y) and 21-plex (#MF0005KMII)), according to manufacturers

protocol and the raw data was analyzed using the software BioPlex Manager Version 6.0. The analyzed cytokines are listed in Table 2.3.

## 2.4 ELISpot Analysis

ELISpot analysis and all preparational steps for the measurement of T cell reactivity was conducted by Iris Kaiser in Professor Jäger’s group (NCT, Medical Oncology).

### 2.4.1 Cell Preparation

Ficoll density gradient centrifugation was performed to isolate PBMC. Thereof, T cells and dendritic cells (DCs) were purified as described previously (Bonertz et al., 2009; Horn et al., 2013). Briefly, T cells were cultured for 7 days in X-VIVO 20 medium containing 100 U/ml human rIL-2 (Proleukin, Chiron, Ratingen, Germany), and 60 U/ml human rIL-4 (Promokine, PromoCell, Heidelberg). Afterward, cells were kept in cytokine-free medium for 12 h and human CD3 T cells were purified using the Dynabeads untouched human T cell kit (Invitrogen, Darmstadt, Germany). For DC maturation, adherent cells were cultured for 7 days in X-VIVO 20 medium containing 560 U/ml human rGM-CSF (Leukine, Berlex, Bayer, Leverkusen, Germany), and 500 U/ml human rIL-4. DCs were enriched using anti-CD56 coupled magnetic beads (C218, Beckman Coulter, Krefeld, Germany), and anti-CD3 and anti-CD19 Dynabeads (Invitrogen, Darmstadt, Germany), and pulsed for 18 h with 0.8 g/l test peptides or IgG. As positive control, 0.1 g/l staphylococcal enterotoxin B (SEB) was used.

### 2.4.2 Peptides

Peptides were produced by the Peptide Synthesis Facility of the DKFZ. Lyophilized synthetic peptides were solved in distilled water containing 10 % DMSO. Peptide purity was >98 %. Peptides were designed to contain the identified immunogenic HLA restricted T cell epitope. Synthesized human IgG peptides as well as IgG (Kiovig, Baxalta, Unterschleißheim, Germany) were used as negative control antigens.

### 2.4.3 IFN $\gamma$ ELISpot assay

ELISpot assays were done as described previously (Bonertz et al., 2009; Horn et al., 2013) with modifications. Briefly, peptide-pulsed DCs ( $2 \times 10^4$ ) were incubated with autologous T cells ( $1 \times 10^5$ ) at a 1:5 ratio for 40 h in ELISpot plates (MAHA S45, Millipore, Eschborn, Germany). All tests were performed in triplicate wells. IFN $\gamma$  spots were measured using the automated system CTL ImmunoSpot analyzer (CTLEurope, Bonn Germany). Spots in the IgG control wells were considered background reactivity. A reaction against a test peptide was considered a positive response if the spot counts were significantly higher than the IgG counts ( $p < 0.05$  two-sided student’s t-test).

## 2.5 Sequencing and Sequence Analysis

Library preparation and sequencing were performed by the DKFZ Genomics and Proteomics Core Facility. Mapping and analysis of DNA sequence data, i.e. detection of SNVs, Indels and CNVs, were performed using the well-established analysis pipelines at the DKFZ Department of Theoretical Bioinformatics (Prof. Roland Eils) and Applied Bioinformatics (Prof. Benedikt Brors). Sequencing was funded by the DKFZ-Heidelberg Center for Personalized Oncology (DKFZ-HIPO) through HIPO-034. Sequencing of the two patients enrolled in NCT POP was funded through HIPO-021.

The methods for sample preparation and sequencing are described briefly in this section for completeness.

### 2.5.1 DNA Library Preparation and Sequencing

DNA for the whole-exome sequencing were isolated using Qiagen AllPrep DNA/RNA/Protein Mini Kit.

Library preparation was performed on the Agilent NGS Workstation (Version F.0, November 2013) using Agilent SureSelectXT Automation Reagent Kit (G9641B) and Human All Exon v5 + UTRs.

Paired-end whole-exome sequencing with a read length of 101 bp was performed on Illumina HiSeq 2000 according to the manufacturer's protocol. According to the HIPO guidelines, the sequencing aimed for a coverage of 60x for control and 80x for tumor samples.

### 2.5.2 Mapping and Analysis of DNA Sequence Data

Paired-end DNA sequencing reads were mapped to the 1000 genomes phase 2 reference genome hs37d5 as previously described (Jones et al., 2012), using Burrows-Wheeler-Aligner (BWA) (version 0.6.2) (Li and Durbin, 2010), and were further processed with SAMtools (version 0.1.17) (Li et al., 2009) and Picard tools (version 1.61) (<http://broadinstitute.github.io/picard>). For a more detailed description of the mapping pipeline, refer to (Jones et al., 2012).

The SNV calling pipeline is based on SAMtools mpileup and bcftools (Li et al., 2009). In addition to previously described filters to remove artifacts (Jones et al., 2012), additional filters have been implemented, such as exclusion of variants located in regions of low mappability. To ensure high confidence for somatic SNVs, several criteria are being considered. For a more detailed description of used filters and quality criteria, refer to (Jones et al., 2013). The output is a file in Variant Call Format (VCF), with some additional columns. One VCF file for each patient was stored for further usage in the Immunoinformatics Pipeline.

The pipeline for the detection of Indels is based on the tool Platypus (0.5.2) (Rimmer et al., 2014). Since Platypus was developed to detect variants in normal genomes,

additional custom filters were added to reliably detect somatic indels in tumor normal pairs. These filters integrate the genotype likelihood as well as other filter criteria originally generated by Platypus.

For the detection of copy number variations the according pipeline is based on VarScan2 (Koboldt et al., 2012). As part of the pipeline, usage of VarScan2 was optimized for large datasets, and a workflow for the automated annotation of genes in the identified regions is included.

### 2.5.3 Mutation Validation

Selected SNVs that were detected in the whole-exome sequencing were validated with PCR amplification and Sanger Sequencing. The procedure was conducted by Claudia Ziegelmeier and Jin-Ho Lee in Professor Jäger’s group (NCT, Medical Oncology).

Primers were designed using online tool primer-blast (<http://www.ncbi.nlm.nih.gov>, 11 February 2016, date last accessed) (Ye et al., 2012) and SeqAnalyzer (LaserGene, DNASTAR, Madison, WI). Primers were designed to span multiple SNVs in the same PCR amplicon, if possible. PCR primer oligonucleotides were synthesized by and purchased from Sigma Aldrich (Germany). A list of all used primers are shown in table 2.4. All PCRs were performed with OptiTaQ DNA Polymerase (Roboklon GmbH, Berlin, Germany) using the standard 20 l mix (for buffer B). Same PCR reaction conditions were used for all experiments, in exceptional cases 10% DMSO was added to the PCR mix and extension duration was adjusted. Specificities of PCRs were checked on a 1.5% Agarose gel and PCRs were purified via QIAquick PCR purification Kit (Qiagen GmbH, Hilden, Germany) and DNA concentrations were measured on the NanoDrop 1000 Spectrophotometer (Peqlab Biotechnologie GmbH, Erlangen, Germany). Samples with purified DNA concentrations between 10 and 50 ng/l were prepared and sent in with either the corresponding forward or reversed PCR primers or both for Sanger sequencing by the SUPREMERun™ sequencing service (GATC Biotech AG, Konstanz, Germany).

## 2.6 HLA Typing Pipeline

The sequence-based HLA typing tool Phlat (Bai et al., 2014) was used to infer the HLA genotype of each patient from the corresponding WXS sequences.

Here, an analysis pipeline was implemented which runs Phlat for multiple input sample in parallel and then summarizes the results of all input samples in one combined result table. This HLA typing pipeline is a collection of custom shell and python scripts. Required input parameters are the name and location of sequencing read files, sample ID, and output directory.

First, the FASTQ files which are usually provided in gz or bz2 archives are unpacked and stored in the output directory (script `run_gunzip.sh`). Next, computing jobs for each of the input samples are submitted and run in parallel on a computing cluster

**Table 2.4:** List of primers used for PCR and validation sequencing of recurring mutations.

Oligo Name	Sequence
APC-FW-nc3296-3316	TTTCTCCATACAGGTCACGGG
APC-RV-nc4149-4129	CATGAGTGGGGTCTCCTGAAC
TP53-RV-nc882-863	CTCCCCTTTCTTGCGGAGAT
APC-FW-in-ex1409-1425	TAGGGGGACTACAGGCCATT
APC-RV-intron19	GCGAATGTGAAGCACAGGTT
APC-FW-nc2339-2358	GTCCCAAGGCATCTCATCGT
APC-RV-nc3333-3314	TTCTGAACCATTGGCTCCCC
APC-FW-nc4005-4027	CAGACTGCAGGGTTCTAGTTTAT
APC-RV-nc4360-4339	TAGGTACTTCTCGCTTGTTTG
APC-FW-intron12	GATAGTCGACCGCCAATCGT
APC-RV-intron13	CAGCACATTGGTACTGAATGCTT
APC-FW-intron13	CAGACACTTCATTTGGAGTACCTTA
APC-RV-intron15	TGGCATTAGTGACCAGGGTT
TP53_FW_exon5	TTTGCTGCCGTCTTCCAGTT
TP53_RV_exon6	GAGGCCCTTAGCCTCTGTAAG
TP53_FW_exon6	AACCCCATGAGATGTGCAAAGT
TP53_RV_intron8	GCCGGGGATGTGATGAGAG
TP53_FW_intron8.1	CTTCATCACATCCCCGGC
TP53_FW_intron8.2	GGCGGGGAATCTCCTTACTG
APC-intron-20-fw1	TTTGGCACTGTAGTAGCATTTAGG
APC-intron-20-fw2	GAGTGCAGTGGTGCGATTAT
APC-EXON20-RV	GCAGATCACAAAGTCAGGAGTAG
KRAS-intron2-fw	GCGTCGATGGAGGAGTTTGT
KRAS-intron3-RV	GGTCCTGCACCAGTAATATGC

(script `run_phlat.sh`). After all jobs are completed, a script to summarize the results is invoked (`run_summary.py`). Finally, all temporary files, such as unpacked FASTQ and HLA alignment files are removed using the script `clean_up.sh`.

The output of the pipeline is a summary table, where the result of all input samples is listed. In doing so, for each patient the HLA genotype for class I alleles HLA-A, -B, -C, and class-II allele -DRB1 was compiled. As described in 1.1.2, only HLA-DRB1 needs to be considered for HLA class II.

## 2.7 Immunoinformatics Pipeline

In addition to the HLA typing pipeline, an analysis pipeline to rapidly perform epitope predictions for large datasets in an automated way was implemented. As part of the pipeline, the HLA binding prediction tools NetMHCpan (2.8) (Hoof et al., 2009) and NetMHCIIpan (Andreatta et al., 2015) (3.0) (for HLA class I and II, respectively), as well as the proteasomal cleavage prediction tool NetChop (3.1) (Kemir et al., 2002) are used. These tools are publicly available and can be downloaded under academic license (<http://www.cbs.dtu.dk/services/software>). Similar to the HLA typing pipeline, this

Immunoinformatics pipeline is also a collection of custom shell, python and R scripts. The input files are the VCF file from the SNV calling pipeline and a table with the HLA genotype. Additional input parameters are sample ID and output directory. An overview of the pipeline components is illustrated in Figure 2.1.

First, using a Python script (`missense.to.table.py`), only missense mutations are extracted from the VCF file and stored in a tab-delimited table. This table is a condensed version of the VCF file, only storing relevant information from the VCF, such as the exact location of the mutation and the gene and protein annotations.

The second part is the retrieval of peptide sequences. Using an R script (`get_peptides.R`) that utilizes the packages `biomaRt` (Durinck et al., 2009, 2005), `annotate` (Gentleman, 2015), and `AnnotationDbi` (Pages et al., 2015), for each missense SNV the corresponding peptide sequence is retrieved. For this, the RefSeq ID of the mutated gene is used, as well as the information about the exact amino acid position and consequence of the mutation, which already have been annotated in the SNV calling pipeline (using `AnnoVar` (Wang et al., 2010)). 29mer peptides are generated, in a way that 14 amino acids flank the mutated residue at each side. Both peptides, the mutated and the wildtype counterpart, are stored. Here, all transcripts of the mutated gene are considered separately, thereby often multiple peptide sequences for one SNV are produced. As different transcript sequences can result in different protein sequences, generated 29mer peptides can also be different, depending on the position of the mutation. Another R script (`get_unique_peptides.R`) is used, to only extract unique peptides. The output of this part is another tab-delimited table containing all distinct peptides together with all the transcript and mutation information.

The next part is the HLA binding prediction using the tools `NetMHCpan` and `NetMHCIIpan`, for HLA class I and class II predictions respectively. This part consists of several steps. As a first step using the peptide table generated during the previous part, for each sample two FASTA files are generated containing all mutated and wildtype peptides, respectively (`script create.fasta.files.R`). Next, using the script `create_netmhc_commands.R`, the `NetMHCpan` and `NetMHCIIpan` commands are created: for each sample one shell script containing the commands for all predictions is automatically generated. Two predictions are run for each allele, one for the mutated peptides and one for the wildtype ones. A `netMHCpan/netMHCIIpan` command has the following format:

```
netMHCpan -a <HLA allele> -f <input FASTA file> >> <output file>
```

The created scripts containing the `netMHCpan/netMHCIIpan` commands are then submitted to the computing cluster. After the computing jobs are finished, the `netMHCpan/netMHCIIpan` files are rearranged and stored as tab-delimited files (`script format_netmhc_files.sh`). Finally, for every sample one large epitope table is created (`script epitope_table.R`) which contains all predicted mutated HLA binders matched with their wildtype counterpart and all information about the prediction, such as the affinity.

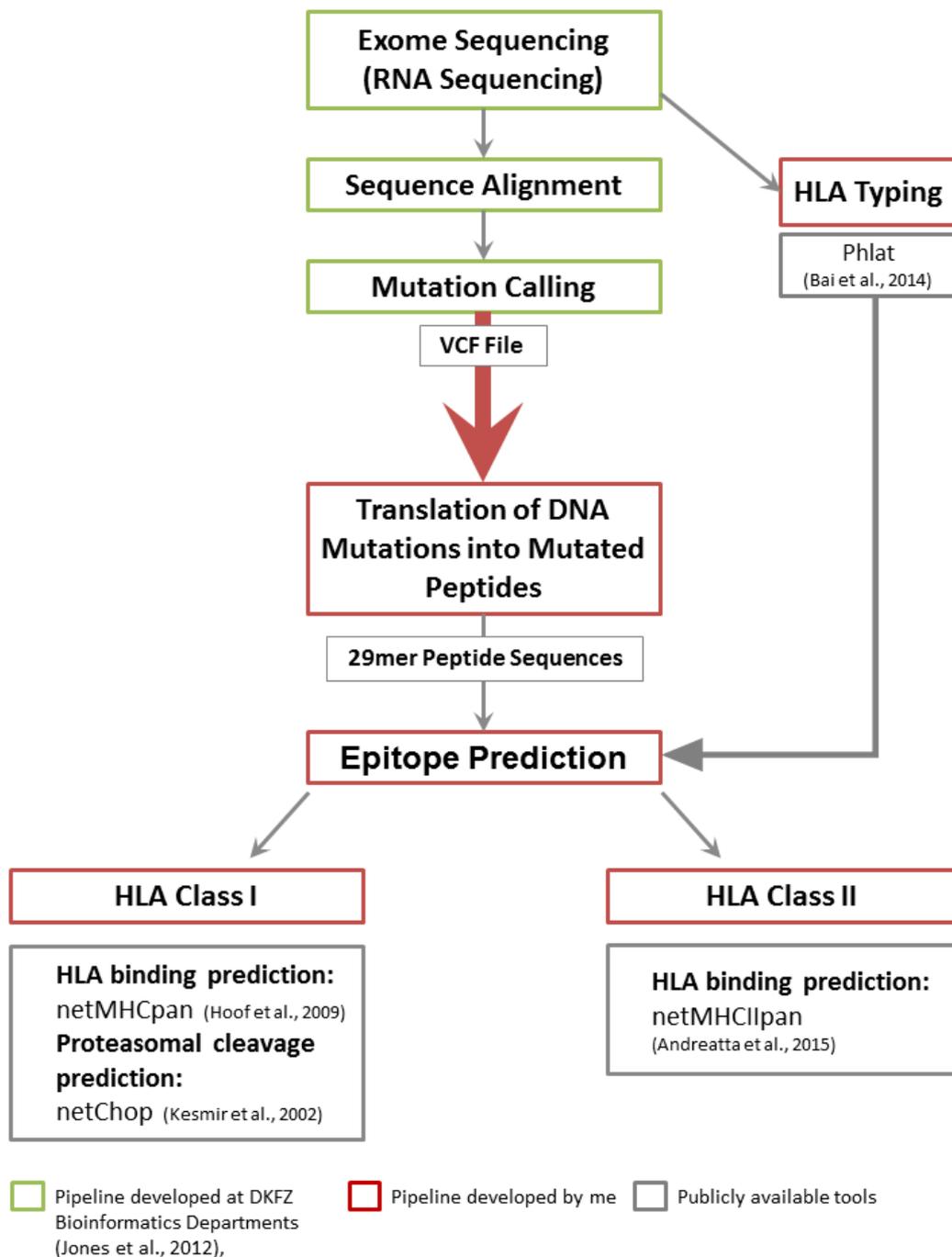
For the predicted HLA class I binders, additionally a proteasomal cleavage prediction is performed using the tool `NetChop`. The script `run_netchop.sh` generates a `netChop` command, directly runs it in the command line and creates a tab-delimited table con-

taining the prediction scores for each peptide. These prediction scores are then added to the epitope table from the previous part using the script `add_netchop_to_table.R`. This table `epitope_table_with_netchop.csv` contains all predicted binders together with the HLA binding and proteasomal cleavage prediction scores.

The final part of the pipeline is the prioritization and selection of epitopes. After extensive literature research, three criteria for the selection of epitopes were defined (also discussed in chapter 4):

- predicted binding affinity of  $\leq 500\text{nM}$
- predicted proteasomal cleavage score of  $\geq 0.5$
- predicted affinity of mutated peptide is higher than corresponding wild-type peptide

The script `rank_epitopes.R` applies these criteria and generates a new table `ranked_epitopes.csv` which only contains the peptides which meet the defined criteria.



**Figure 2.1: Components of the Immunoinformatics pipeline.** The Immunoinformatics pipeline is a collection of custom shell, python and R scripts. The input data are a file with the somatic point mutations in Variant Call Format (VCF) from the Single Nucleotide Variation (SNV) calling pipeline, and a table with the HLA genotype for the corresponding samples from the HLA Typing pipeline. Parts of the pipeline that have been implemented by me are shown in red boxes, in green boxes the parts that have been established at the DKFZ Bioinformatics departments Theoretical Bioinformatics (Prof. Roland Eils) and Applied Bioinformatics (Prof. Benedikt Brors), are shown, and the gray boxes indicate the publicly available tools.

## 2.8 Downstream Analysis

Custom scripts were used for downstream analysis and plotting of figures using R programming language (R Core Team, 2014).

The R package ggplot (Wickham, 2009) was used to create all plots.

For the clustering analysis and creation of heatmaps for the cytokine data, the R package heatmap3 (Zhao et al., 2015) was used. Cytokine expression values were z-score transformed prior to analysis.

For the pathway enrichment analysis of mutations the Ingenuity Pathway Analysis (IPA) software was used (<http://www.ingenuity.com>). IPA performs a gene set enrichment analysis on an input list of genes and generates a p-value for each canonical pathway or functional category based on Fisher's exact test. The p-value reflects the significance of the enrichment of each pathway for the list of input genes. For each canonical pathway, IPA also calculates the ratio of the number of genes from the input list that are annotated to the pathway to the total number of genes annotated to the pathway.

## CHAPTER 3

---

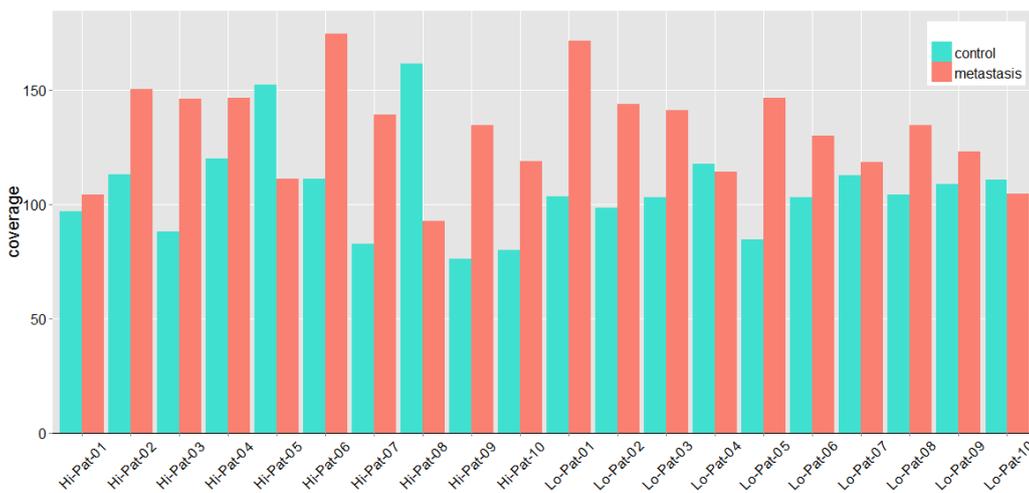
Results

---

As part of this PhD project, an integrated analysis pipeline to obtain detailed data about tumor-host interactions was implemented and applied to a dataset of CRC liver metastases. Samples and experimental procedures were provided by members of Professor Jäger’s group (NCT, Medical Oncology), and sequencing was conducted by the DKFZ Genomics and Proteomics Core Facility.

### 3.1 Sequencing Data

The minimum sequencing coverage of 60x for control and 80x for tumor samples required by Hipo was achieved for all samples (Figure 3.1). For some samples, a second sequencing run had to be performed.



**Figure 3.1: Coverage of sequenced samples.** For each patient, the sequencing coverage, i.e. read depth, of metastasis (red) and control (turquoise) samples are shown.

### 3.2 Mutational Landscape

Analysis of the somatic mutation profile revealed a heterogeneous landscape. Only few SNVs and Indels are shared between patients. Several well-known arm-level CNVs were found in a number of patients.

#### 3.2.1 Single Nucleotide Variations

The number of somatic SNVs for each patient ranges between 62 and 643, which corresponds to a mutation frequency of 0.83 - 8.62 mutations per megabase. When only somatic SNVs in coding regions (exonic SNVs) are considered, the numbers are lower,

ranging between 11 and 183. The bulk of the somatic mutations are located in non-coding regions which are: splicing, ncRNA, intergenic, intronic, upstream/downstream, and UTR (Table 3.2). Considering exonic SNVs there are again different types, which are: stopgain, stoploss, missense (nonsynonymous), and synonymous. Here, the majority of exonic SNVs are the missense ones (Figure 3.4). There are up to 13 stopgain mutations and only one stoploss mutation occurs in Lo-Pat-10.

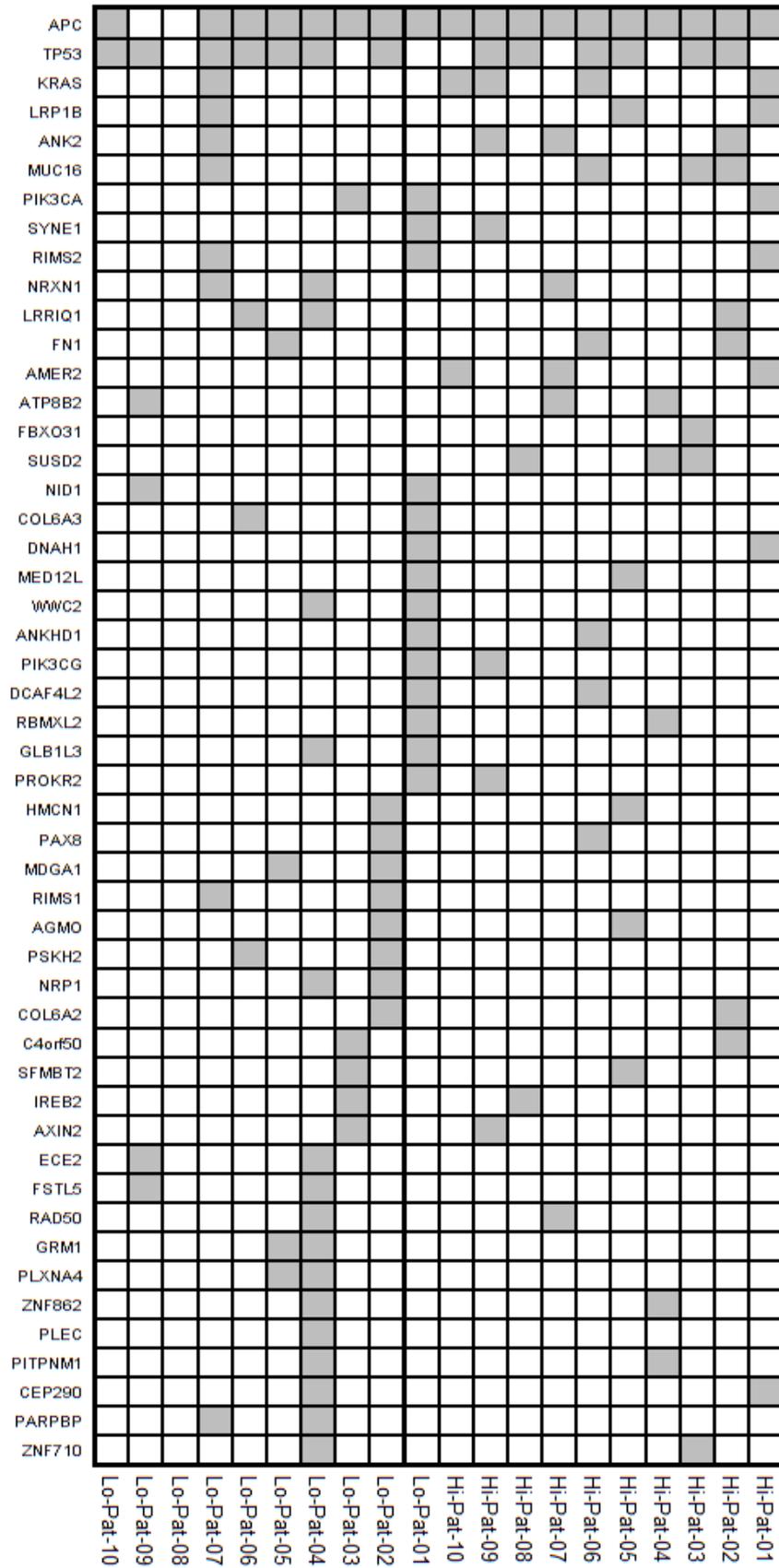
The well-known tumor suppressor genes APC and TP53, as well as the oncogene KRAS are frequently mutated in our patient cohort. Besides from these frequent mutations, the somatic mutation profiles are quite heterogeneous, with low overlap in mutations between single patients (Figure 3.2). 1445 different somatic nonsynonymous coding SNVs were identified, and a total of 1262 genes were found to be somatically mutated. Three genes are mutated in at least four, 10 genes in at least three, and 122 in at least two patients.

The gene APC is mutated in 18 of the 20 patients, and TP53 in 13 patients. Of note, the genes are not always mutated at the same position in each patient. APC, for instance, is mutated at 15 different positions and TP53 at 10 different positions. Also, almost all mutations in APC are stopgain mutations. The oncogene KRAS is mutated in 5 patients and only at the protein-position 12 (Table 3.1 and Figure 3.3).

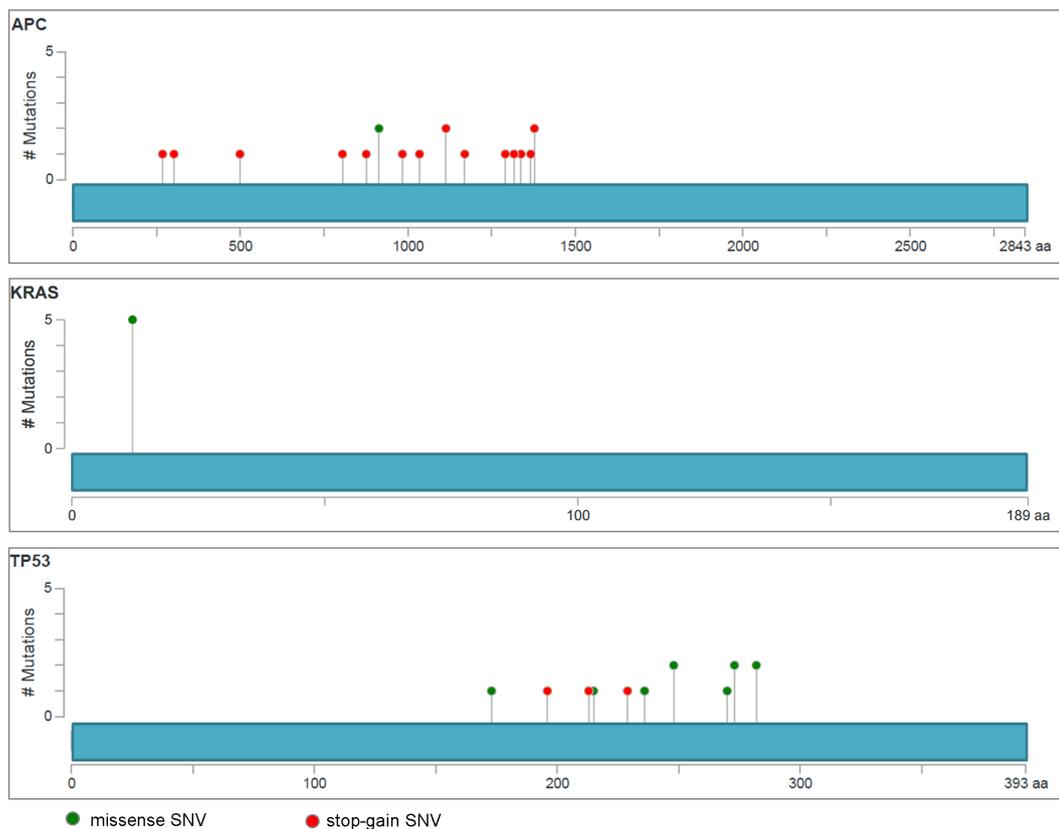
When comparing the frequencies of missense mutations between the two patient groups TIL-high and TIL-low, there is no significant difference between the two groups ( $p = 0.2959$ , two-sided student's t-test) (Figure 3.5). The median of the number of missense mutations is 53 for TIL-high and 52.5 for TIL-low samples.

A set of 57 immune-related genes (Angelova et al., 2015) was further analyzed for accumulation of mutations. Only two genes were found to harbor missense mutations and these are individual cases: CD40LG is mutated in Hi-Pat-05 and CD274 (PD-L1) in Lo-Pat-10. Also, all cytokine and cytokine receptor genes were analyzed for accumulation of mutations and no recurrent mutations were found.

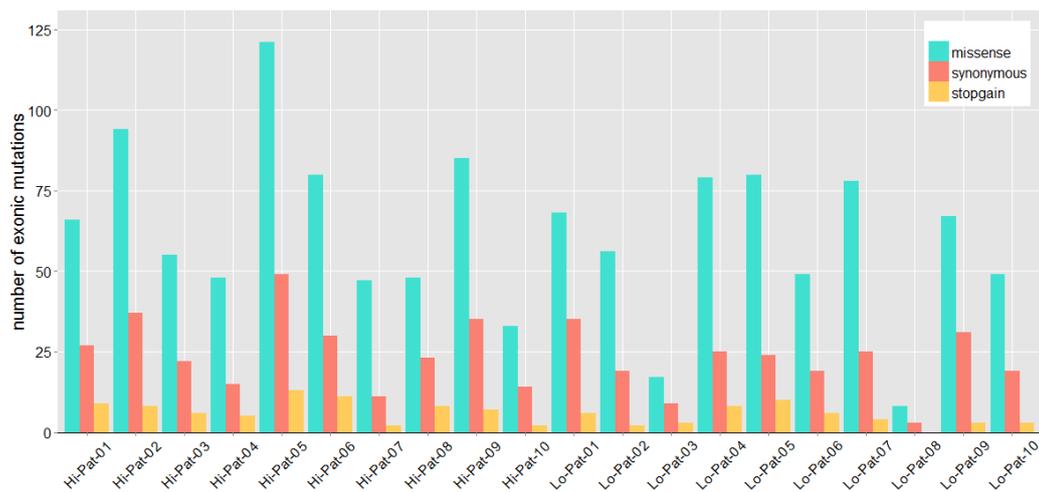
The mutations in the frequently mutated genes APC, KRAS and TP53 were also validated via PCR amplification and Sanger Sequencing. In doing so, almost all mutations were validated, except for the APC mutations in Hi-Pat-03 and Hi-Pat-04. Also when analyzing the KRAS mutations, the mutations G12D and G12S were detected in Lo-Pat-03 and Lo-Pat-08, respectively. These mutations have not been detected through SNV calling on the whole-exome-sequencing data.



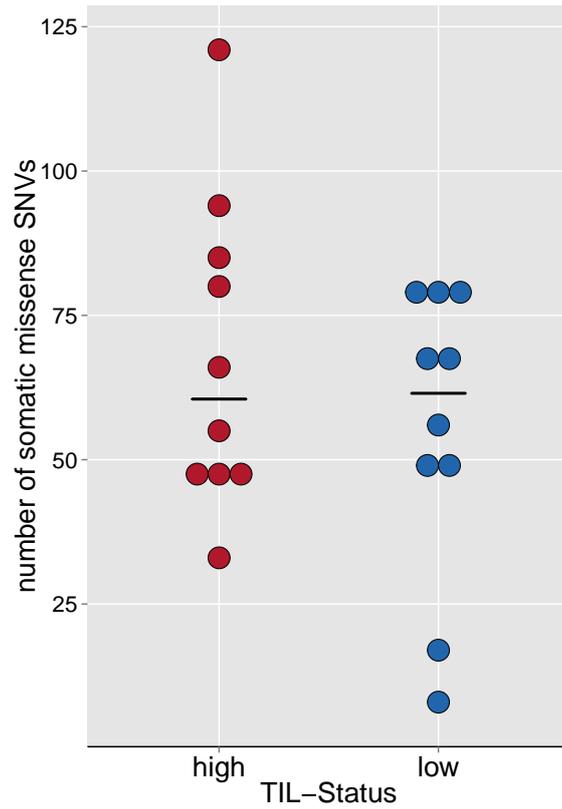
**Figure 3.2: Mutational landscape.** The top 50 genes which are mutated in more than one patient are shown. A gray cell indicates that an exonic mutation was detected in the corresponding patient and gene.



**Figure 3.3: Overview of mutations in frequently mutated genes.** Distribution, frequency and type of mutations in the frequently mutated genes APC, KRAS, and TP53 are illustrated. The initial diagram was created using the tool MutationMapper (v.1.0.1) ([http://www.cbioportal.org/mutation\\_mapper](http://www.cbioportal.org/mutation_mapper)) and modified manually.



**Figure 3.4: Frequency of somatic exonic mutations.** For each sample, the frequency of the different types of somatic exonic mutations missense (turquoise), synonymous (red) and stopgain (yellow) are shown. One stoploss mutation occurs in patient Lo-Pat-10 (not shown).



**Figure 3.5: Comparison of frequency of somatic missense mutations between TIL-high and TIL-low samples.** Each dot represents the number of somatic missense mutations in one patient sample. The bars represent the median for the corresponding patient groups. The difference in the number of somatic missense mutations between TIL-high and TIL-low samples is not significant in the analyzed patient cohort ( $p = 0.2959$ , two-sided student's t-test).

**Table 3.1:** Mutations in frequently mutated genes.

Gene	PID	Protein Change	Mutation Type	Chrom Position	Ref Allele	Var Allele
APC	Lo-Pat-04	E268X	stopgain SNV	5:112137048	G	T
APC	Lo-Pat-01	R302X	stopgain SNV	5:112151261	C	T
APC	Lo-Pat-02	R499X	stopgain SNV	5:112162891	C	T
APC	Hi-Pat-01	R805X	stopgain SNV	5:112173704	C	T
APC	Hi-Pat-03	R876X	stopgain SNV	5:112173917	C	T
APC	Hi-Pat-02	C914Y	missense SNV	5:112174032	G	A
APC	Hi-Pat-02	C914W	missense SNV	5:112174033	T	G
APC	Hi-Pat-09	E984X	stopgain SNV	5:112174241	G	T
APC	Hi-Pat-06	Q1035X	stopgain SNV	5:112174394	C	T
APC	Hi-Pat-08, Hi-Pat-04	R1114X	stopgain SNV	5:112174631	C	T
APC	Lo-Pat-01	K1170X	stopgain SNV	5:112174799	A	T
APC	Lo-Pat-06	Q1291X	stopgain SNV	5:112175162	C	T
APC	Hi-Pat-10	E1317X	stopgain SNV	5:112175240	G	T
APC	Hi-Pat-03	Q1338X	stopgain SNV	5:112175303	C	T
APC	Hi-Pat-05	Q1367X	stopgain SNV	5:112175390	C	T
APC	Lo-Pat-07, Lo-Pat-03	Q1378X	stopgain SNV	5:112175423	C	T
KRAS	Hi-Pat-01	G12A	missense SNV	12:25398284	C	G
KRAS	Lo-Pat-07	G12V	missense SNV	12:25398284	C	A
KRAS	Hi-Pat-06, Hi-Pat-09	G12D	missense SNV	12:25398284	C	T
KRAS	Hi-Pat-10	G12S	missense SNV	12:25398285	C	T
TP53	Lo-Pat-05, Hi-Pat-09	R282W	missense SNV	17:7577094	G	A
TP53	Lo-Pat-09, Hi-Pat-06	R273C	missense SNV	17:7577121	G	A
TP53	Hi-Pat-02	F270L	missense SNV	17:7577130	A	G
TP53	Lo-Pat-04, Hi-Pat-05	R248Q	missense SNV	17:7577538	C	T
TP53	Lo-Pat-07	Y236C	missense SNV	17:7577574	T	C
TP53	Hi-Pat-08	C229X	stopgain SNV	17:7577594	A	T
TP53	Lo-Pat-02	S215R	missense SNV	17:7578206	T	G
TP53	Lo-Pat-06	R213X	stopgain SNV	17:7578212	G	A
TP53	Lo-Pat-10	R196X	stopgain SNV	17:7578263	G	A
TP53	Hi-Pat-03	V173M	missense SNV	17:7578413	C	T

**Table 3.2:** Frequency of somatic mutations

PID	somatic all	mutation frequency	somatic splicing	somatic ncRNA	somatic UTR	somatic up/down stream	somatic intronic	somatic intergenic	somatic all exonic	somatic exonic stopgain	somatic exonic stoploss	somatic exonic missense	somatic exonic synonymous
Hi-Pat-01	287	3.84	0	23	64	9	70	22	102	9	0	66	27
Hi-Pat-02	643	8.62	5	58	125	11	233	77	139	8	0	94	37
Hi-Pat-03	191	2.56	1	10	41	0	48	9	83	6	0	55	22
Hi-Pat-04	169	2.26	0	6	45	5	34	10	68	5	0	48	15
Hi-Pat-05	588	7.88	4	42	140	13	158	53	183	13	0	121	49
Hi-Pat-06	498	6.67	4	37	80	14	159	85	121	11	0	80	30
Hi-Pat-07	149	1.99	1	7	21	5	41	14	60	2	0	47	11
Hi-Pat-08	215	2.88	0	13	51	1	56	19	79	8	0	48	23
Hi-Pat-09	380	5.09	1	16	85	10	105	39	127	7	0	85	35
Hi-Pat-10	140	1.87	1	13	36	3	34	10	49	2	0	33	14
Lo-Pat-01	411	5.51	2	29	84	3	138	53	109	6	0	68	35
Lo-Pat-02	254	3.40	0	23	45	4	90	19	77	2	0	56	19
Lo-Pat-03	113	1.51	1	6	31	3	34	10	29	3	0	17	9
Lo-Pat-04	358	4.80	4	27	70	10	109	27	112	109	0	79	25
Lo-Pat-05	326	4.37	3	22	68	10	101	12	114	10	0	80	24
Lo-Pat-06	239	3.20	0	17	53	9	65	25	74	6	0	49	19
Lo-Pat-07	326	4.37	0	22	69	4	89	37	107	4	0	78	25
Lo-Pat-08	62	0.83	0	5	5	0	23	18	11	0	0	8	3
Lo-Pat-09	288	3.86	1	23	74	4	70	20	101	3	0	67	31
Lo-Pat-10	230	3.08	0	14	51	4	65	26	72	3	1	49	19

ncRNA = non-coding RNA

UTR = untranslated region

### 3.2.2 Indels

91 different somatic coding indels in 77 different genes were detected in the analyzed samples. Of these indels, 69 are frameshift mutations (35 deletions and 34 insertions), and 17 are non-frameshift mutations (12 deletions and 5 insertions). Again, the onco-gene APC is frequently affected in 11 patients, and TP53 is affected in two patients (Table 3.3). The remaining 77 genes are only affected in single patients.

**Table 3.3:** Indels in frequently mutated genes

Gene	PID	Protein Change	Mutation Type	Chrom Position	Ref Allele	Var Allele
APC	Lo-Pat-01	D1297fs	fs del	5:112175180	GA	G
APC	Lo-Pat-02	1397del	fs del	5:112175479	TGA	T
APC	Lo-Pat-04	S1411fs	fs del	5:112175523	GT	G
APC	Hi-Pat-01	E1554fs	fs ins	5:112175951	G	GA
APC	Lo-Pat-09	S1395fs	fs ins	5:112175475	G	GT
APC	Hi-Pat-04	1327del	fs del	5:112175270	TCA	T
APC	Hi-Pat-05	L1302fs	fs del	5:112175195	CT	C
APC	Hi-Pat-06, Hi-Pat-09	P1439fs	fs del	5:112175605	AC	A
APC	Hi-Pat-07	S1426fs	fs del	5:112175568	GC	G
APC	Hi-Pat-08	A1485fs	fs del	5:112175745	CT	C
TP53	Lo-Pat-01	131-132del	non-fs del	17:7578534	CTTG	C
TP53	Lo-Pat-03	L62fs	fs del	17:7578266	TA	T

### 3.2.3 Copy Number Variations

Detailed analysis of CNVs was reported comprehensively elsewhere (Network, 2012; Xie et al., 2012) and was not in the scope of this study. Instead, here the focus was on the comparative analysis of the TIL-high and TIL-low groups, and on deletions and amplifications of cytokine and cytokine-receptor genes (Consortium), and immunomodulatory genes (Angelova et al., 2015).

Several arm-level changes were detected, which are known to be frequent in CRC patients (Network, 2012). These include amplifications of 7p and q, 8q, 13q, 20p and q, and deletion of arms 8p, 17p, 18p and q (Figure 3.6). Additionally, several focal amplifications and deletions in oncogenes as well as in tumor suppressor genes were found. According to our data, there are no copy number changes, which are distinct in the two patient groups TIL-high and TIL-low. However, if the number of events are considered, some differences between the two groups TIL-high and TIL-low become apparent (Figure 3.7). While the difference in the number of arm-level aberrations between TIL-high and TIL-low samples is not significant ( $p = 0.16$ , two-sided student's t-test) the difference in the number of focal aberrations between TIL-high and TIL-low samples was found to be significant ( $p = 0.03$ , two-sided student's t-test).

Next, chromosomal alterations on the gene level were analyzed, i.e. it was investigated which genes are frequently affected by amplifications or deletions and if they are over-represented in one of the two patient groups. About 700 were found to be significantly more amplified in the TIL-high group ( $p \leq 0.01$ , two-sided student's t-test). Interestingly, no genes were found to be significantly more deleted. Analysis of the 702 overrepresented genes for pathway enrichment using the ClueGo tool, showed that no pathway is significantly enriched.

Again, the focus was on the set of immunomodulatory genes and investigated, whether any of them is among the genes that are significantly more often amplified in the TIL-high group. As a result, three immunomodulatory genes were found to be more often amplified in the TIL-high group:

- FOXP3 (amplified in 7 TIL-high and 0 TIL-low patients)
- CYBB (amplified in 6 TIL-high and 0 TIL-low patients)
- Phex (amplified in 8 TIL-high and 2 TIL-low patients)

Also, all cytokine and cytokine receptor genes were analyzed for accumulation of amplifications and no gene was found to be significantly more amplified.

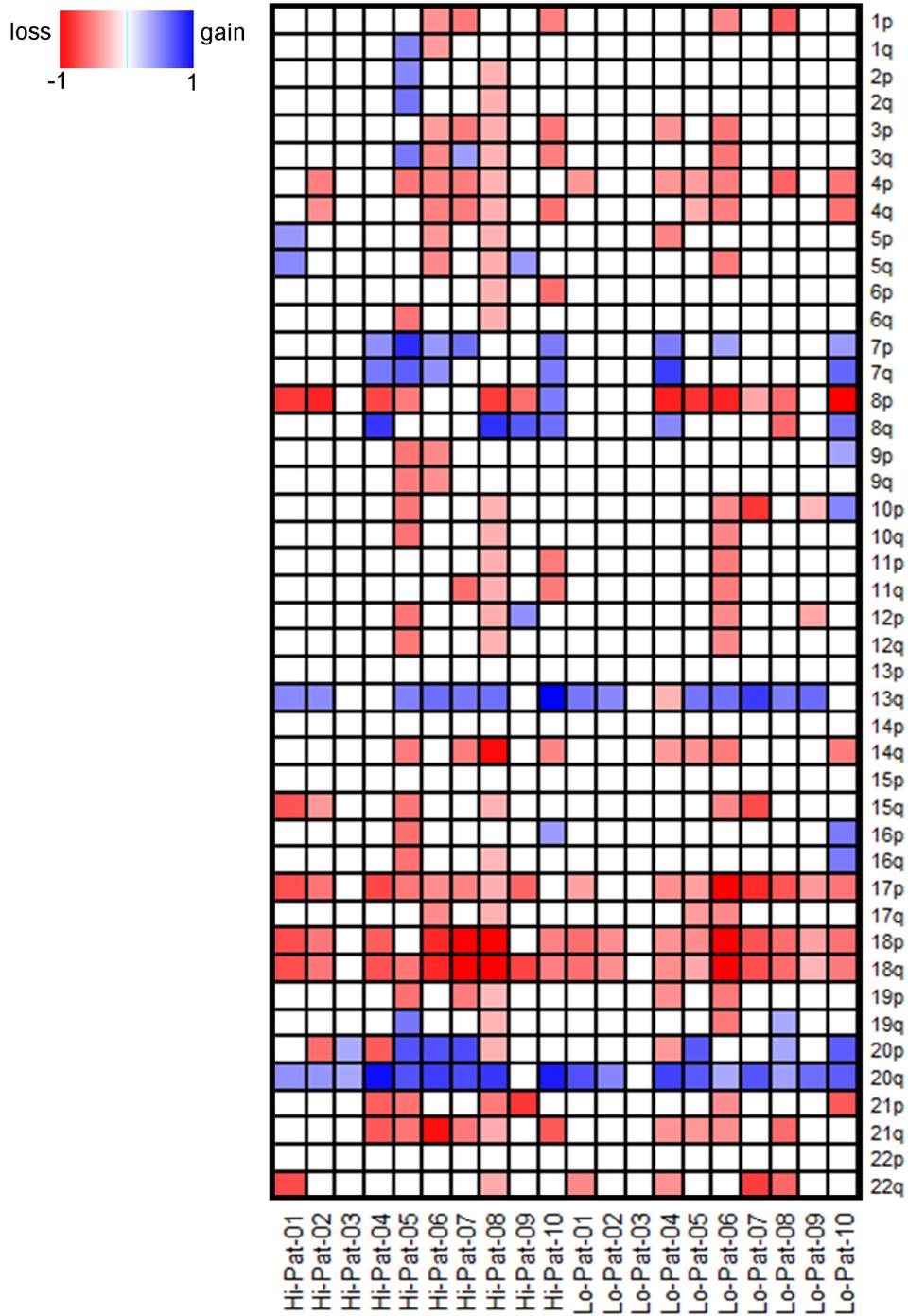
### 3.2.4 Pathway Enrichment Analysis of Mutations

For the pathway enrichment analysis of mutations the Ingenuity Pathway Analysis (IPA) software was used.

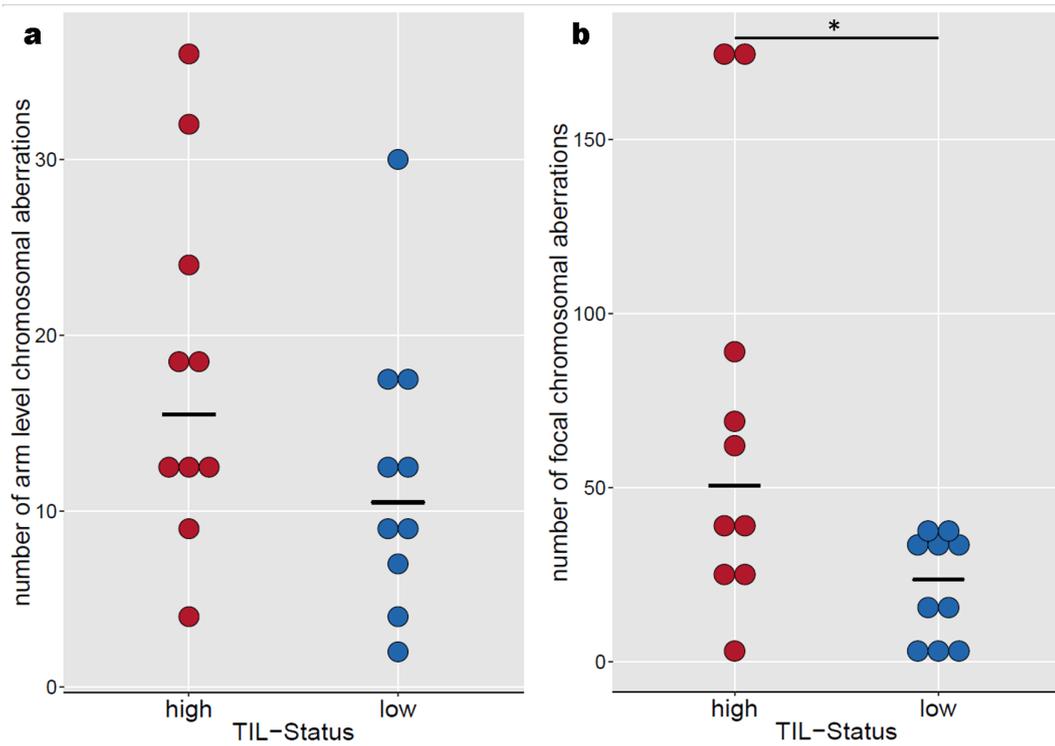
In both groups, several cancer-related pathways are in the top 10 enriched pathways, non of which correspond to colorectal cancer, but to other cancers instead.

In the TIL-low group, two immune-associated pathways are in the top 10 enriched pathways, namely "acute phase response signaling" and "PI3K signaling in B lymphocytes". These pathways are highly enriched in the TIL-low group, when compared to the TIL-high group: the  $-\log(p\text{-value})$  of "acute phase response signaling" is 3.11 in TIL-low, compared to 0.78 in TIL-high, and for "PI3K signaling in B lymphocytes", 2.33 in TIL-low, compared to 0.77 in TIL-high.

Next, only immune-associated pathways were considered for enrichment analysis and the results were visualized in a heatmap using IPA's comparison analysis feature (Figure 3.9). In this analysis, several other immune pathways become apparent that are more significantly effected in one of the two patient groups. The pathways "NF- $\kappa$ B activation by viruses" and "IL-2 signaling" for instance are more affected in the TIL-high group.

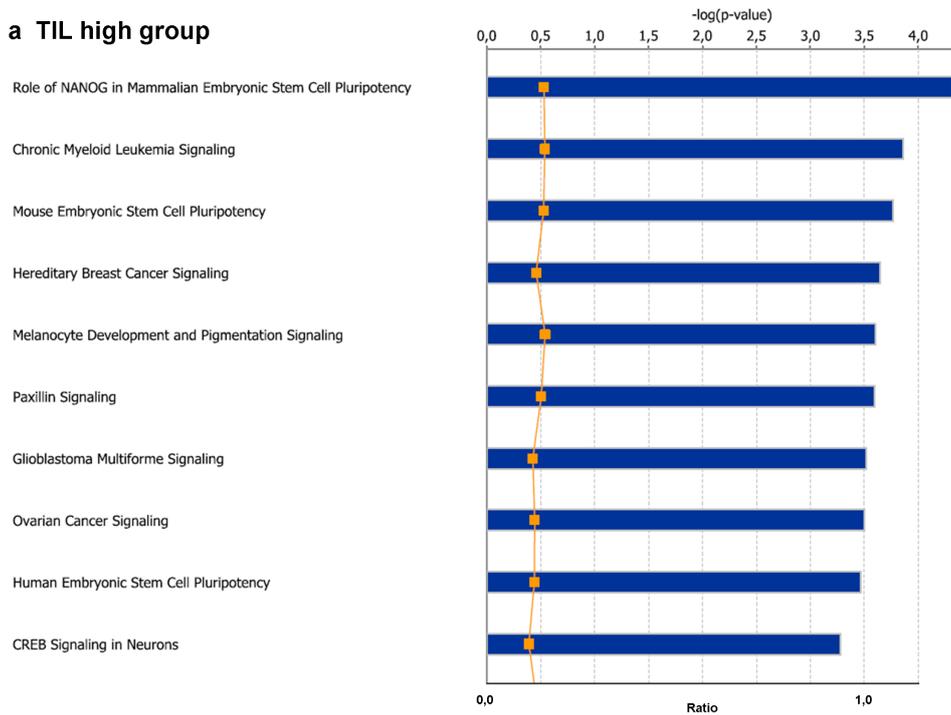


**Figure 3.6: Copy number variations.** Arm-level copy number changes for the 20 patients are shown. A red cell indicates a loss (deletion) whereas a blue cell indicates a gain (amplification) of the corresponding chromosome arm in the corresponding patient sample.

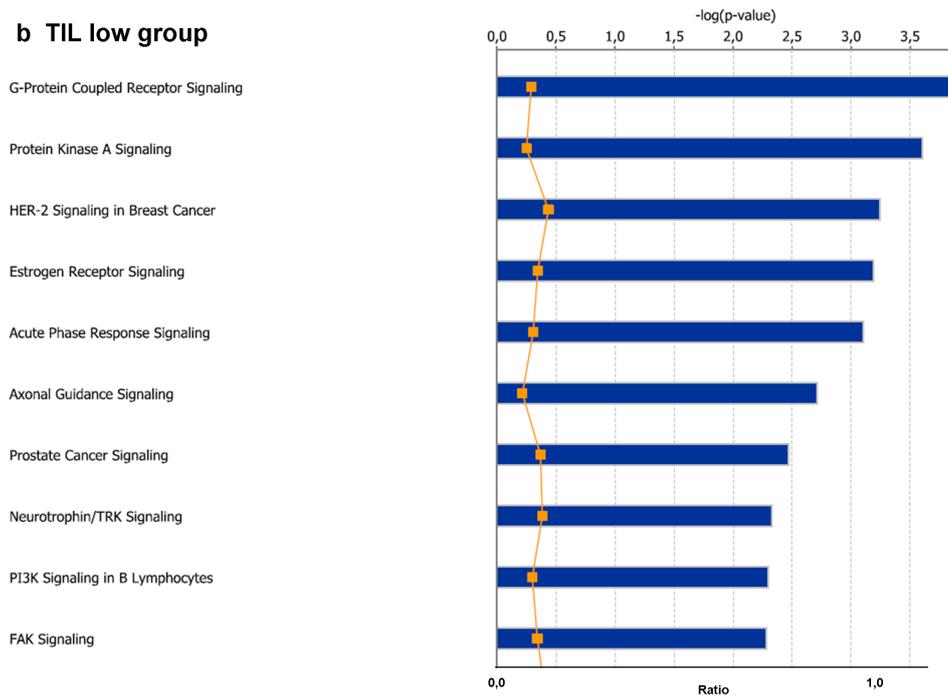


**Figure 3.7: Comparison of frequency of copy number changes between TIL-high and TIL-low samples.** Each dot represents the number of copy number changes in one patient sample. The bars represent the median for the corresponding patient groups. (a) number of chromosomal aberrations affecting a whole chromosome arm (b) number of focal chromosomal aberrations. The difference in the number of focal aberrations between TIL-high and TIL-low samples was found to be significant ( $p = 0.03$ , two-sided student's t-test)

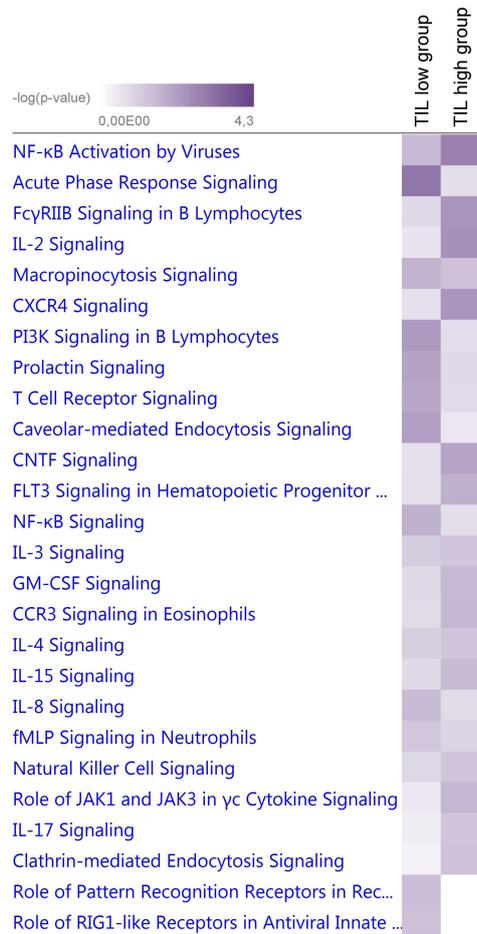
### a TIL high group



### b TIL low group



**Figure 3.8: Top enriched canonical pathways.** Ingenuity Pathway Analysis (IPA) software was used to characterize the lists of mutated genes for the TIL-high and TIL-low patients. Shown are the top 10 canonical pathways that are enriched for mutations, (a) in the TIL-high group, (b) in the TIL-low group. The blue bars represent the p-value, which indicates the significance of the enrichment of input genes in each pathway. The orange squares represent the the ratio of the number of genes from the input list that were annotated to the pathway, to the total number of genes annotated to the pathway.



**Figure 3.9: Comparative analysis of top enriched immune-associated canonical pathways.** Ingenuity Pathway Analysis (IPA) software was used to detect pathways that are significantly enriched for mutations in the TIL-high and TIL-low patients. For both patient groups, enrichment analysis was first performed separately, and IPA calculated a p-value for each pathway, which indicates the significance of the enrichment of input genes. For this comparative analysis, only immune-associated canonical pathways were considered. The heatmap shows the pathways, where a significant difference in enrichment was detected between the two patient groups.

### 3.3 HLA Typing

Sequence-based HLA typing was performed to retrieve class I alleles HLA-A,-B,-C, and class II allele HLA-DR. For all patients, sufficient coverage of the HLA loci was provided. An overview of the HLA genotype for each patient is shown in Table 3.4. Of note, several patients have some homozygous HLA alleles. A total of 54 different class I alleles are present in the analyzed patient cohort: 16 HLA-A, 22 HLA-B, and 16 HLA-C. For HLA class II, 18 different HLA-DRB1 alleles are present in the analyzed patient cohort.

The HLA typing accuracy using Phlat was also validated using sequencing data from two patients who are being treated at the NCT and HLA genotyping data is available from conventional PCR-based typing techniques. For all six HLA class I alleles, the genotypes obtained with Phlat were consistent with the experimentally derived genotypes in both patients. For the two HLA class II alleles, in one patient the results were consistent and in the other patient one allele was typed differently. In summary, for HLA class I Phlat achieved 100% accuracy, and 75% for HLA class II in the analyzed two patients

### 3.4 Epitope Landscape

All 1445 somatic coding SNVs were considered for allele-specific HLA Class I and HLA Class II epitope predictions as described in Section 2.

#### 3.4.1 HLA Class I Epitopes

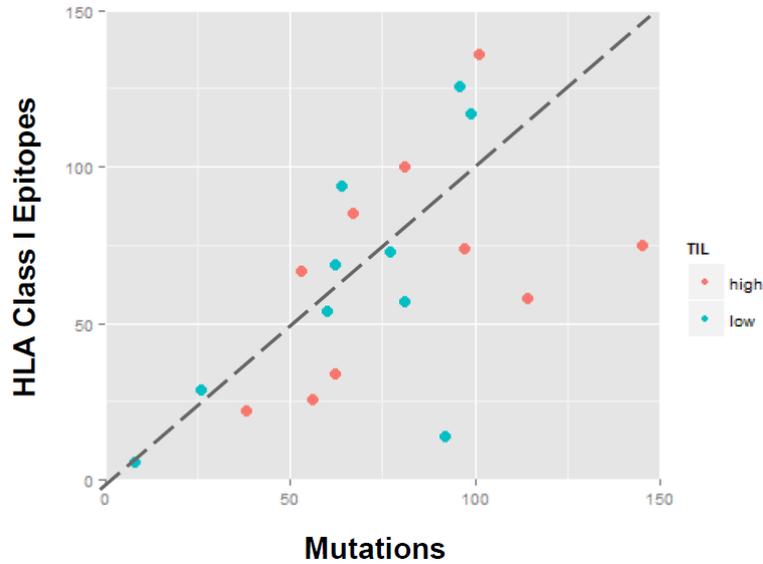
A total of 1962 peptides of length 9-11aa were predicted to bind with high to intermediate affinity ( $IC_{50} \leq 500nM$ ) to one of the analyzed HLA class I allele. The number of predicted binders for each patient ranges between 7 and 229. When only predicted binders, which also were predicted to be cleaved by the proteasome are considered, these numbers drop for about 10% - 50% and range between 4 and 158. Comparing the predicted binders with the corresponding wild-type peptides shows that in about 25% - 40% of the cases the wild-type peptide was also a predicted binder with a higher affinity than the mutated peptide. Hence, after filtering for cleaved and stronger binding mutated peptides, about 50% of the initial predicted binders remain and are being considered as epitopes here (Figure 3.11).

Comparing the HLA class I epitope frequencies between TIL-high and TIL-low samples shows that the numbers are scattered in a range between 22 and 106 for the TIL-high samples, and between four and 107 for the TIL-low samples (Figure 3.12). The median is 47 for the TIL-high samples and 60.5 for the TIL-low samples. Statistically, there is no significant difference ( $p = 0.925$ , two-sided student's t-test).

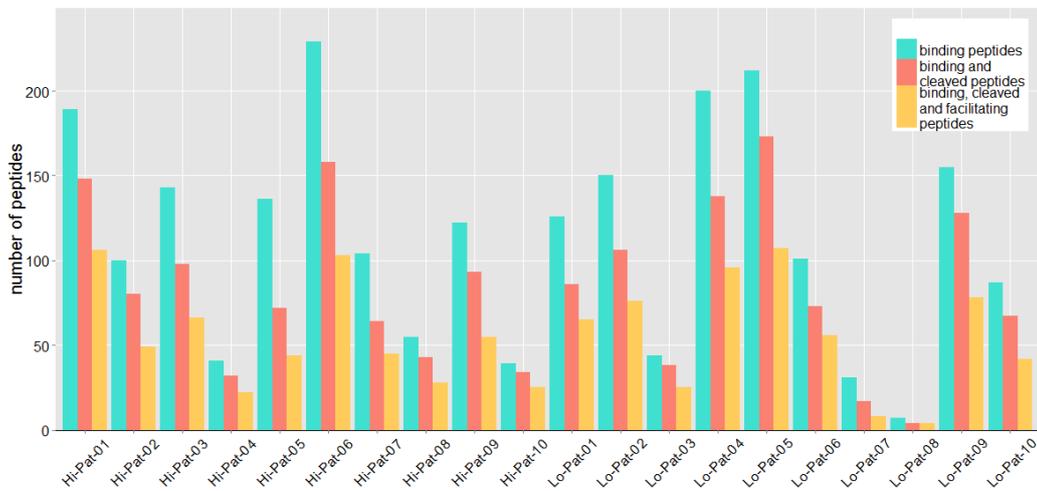
**Table 3.4:** HLA genotype of patients.

PID	HLA Class I Alleles			HLA Class II Alleles
Lo-Pat-01	A*01:81	B*52:01	C*15:05	DRB1*04:05
	A*02:01	B*73:01		DRB1*15:02
Lo-Pat-02	A*02:01	B*41:02	C*15:02	DRB1*04:04
	A*66:01	B*51:01	C*17:01	DRB1*03:03
Lo-Pat-03	A*02:01	B*08:01	C*07:02	DRB1*01:01
	A*24:02	B*35:02	C*04:01	
Lo-Pat-04	A*02:01	B*41:01	C*17:01	DRB1*03:01
	A*32:01	B*38:01	C*12:03	DRB1*13:01
Lo-Pat-05	A*01:81	B*40:01	C*07:02	DRB1*11:01
	A*30:02	B*07:02	C*03:04	DRB1*15:01
Lo-Pat-06	A*23:01	B*44:03	C*04:01	DRB1*14:54
	A*29:02	B*35:01		DRB1*15:02
Lo-Pat-07	A*24:02	B*35:03	C*04:01	DRB1*04:03
	A*23:01			DRB1*14:54
Lo-Pat-08	A*25:01	B*14:01	C*07:02	DRB1*07:01
	A*32:01	B*39:01	C*08:02	DRB1*14:54
Lo-Pat-09	A*24:02	B*39:01	C*12:03	DRB1*11:04
		B*44:02	C*16:04	DRB1*11:01
Lo-Pat-10	A*01:01	B*35:01	C*07:01	DRB1*01:01
	A*03:26	B*08:01	C*04:01	DRB1*03:01
Hi-Pat-01	A*25:01	B*38:01	C*03:04	DRB1*04:01
	A*24:02	B*15:01	C*12:03	DRB1*01:01
Hi-Pat-02	A*24:02	B*13:02	C*06:02	DRB1*07:01
	A*03:01	B*07:02	C*07:02	DRB1*15:01
Hi-Pat-03	A*01:01	B*08:01	C*04:01	DRB1*04:04
	A*02:01	B*15:01	C*07:01	DRB1*03:01
Hi-Pat-04	A*26:01	B*57:01	C*06:02	DRB1*01:01
	A*01:22	B*27:05	C*01:02	DRB1*04:02
Hi-Pat-05	A*33:01	B*08:01	C*08:02	DRB1*13:02
	A*01:01	B*14:02	C*07:01	DRB1*01:02
Hi-Pat-06	A*02:01	B*07:02	C*01:02	DRB1*07:01
	A*11:01	B*15:01	C*07:02	DRB1*13:01
Hi-Pat-07	A*31:01	B*49:01	C*08:01	DRB1*01:01
	A*30:04	B:35:01	C:07:01	DRB1*13:02
Hi-Pat-08	A*26:01	B*07:02	C*07:02	DRB1*11:42
	A*24:02	B*35:02	C*04:01	DRB1*07:01
Hi-Pat-09	A*23:01	B*33:03	C*07:01	DRB1*03:01
	A*01:01	B*08:01	C*02:02	DRB1*13:02
Hi-Pat-10	A*24:02	B*13:02	C*05:01	DRB1*04:01
	A*32:01	B*44:02	C*08:02	DRB1*01:02

Correlation analysis of the number of predicted HLA class I binders and the number of mutations showed a significant correlation (Spearman's rho = 0.634, p = 0.002) (Figure 3.10).



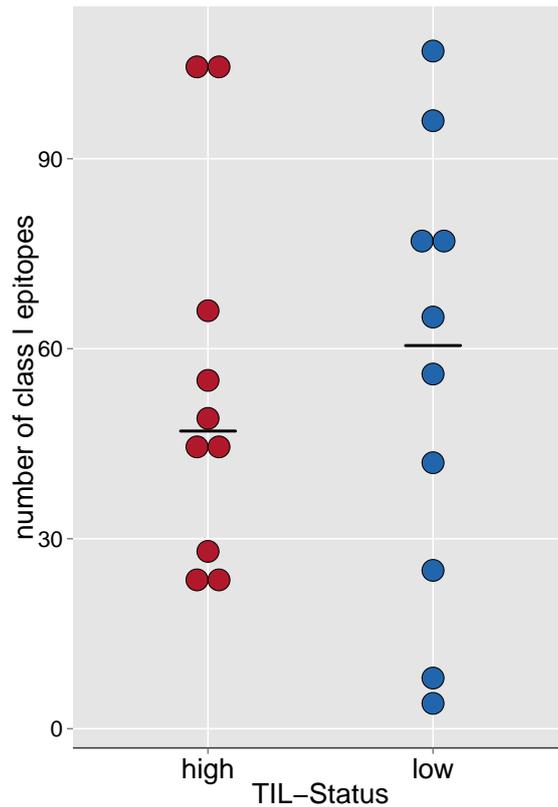
**Figure 3.10: Correlation of number of mutations and number of predicted HLA class I binders.** For each patient, the frequency of predicted HLA class I binders is plotted against the number of mutations. TIL-high samples are indicated in red, TIL-low samples in turquoise. The correlation was found to be significant (Spearman's rho = 0.634,  $p = 0.002$ ).



**Figure 3.11: Frequency of predicted HLA class I binders.** For each sample, the frequency of peptides predicted to be bound by one of the patients' HLA Class I molecules (turquoise), the frequency of peptides that are also predicted to be cleaved (red), and the frequency of those peptides, where the mutated peptide was predicted to bind stronger than the wildtype (yellow), are illustrated.

### 3.4.2 HLA Class II Epitopes

For HLA Class II only the DRB1 alleles were considered (see Section 2). As a result, a total of 7848 peptides of length 9-15aa were predicted to bind with high to intermediate affinity ( $IC_{50} \leq 500nM$ ) to one of the analyzed HLA class II allele. As described in

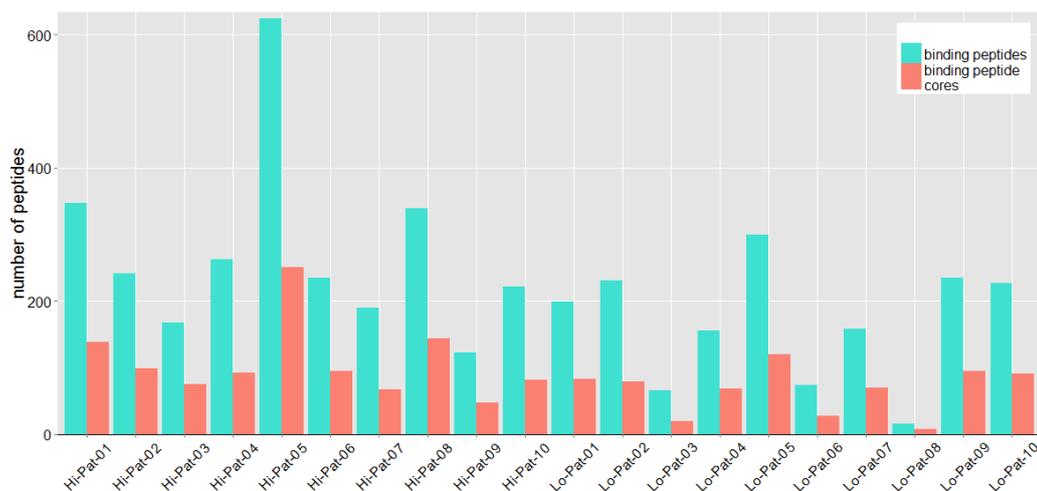


**Figure 3.12: Comparison of HLA class I epitope frequencies between TIL-high and TIL-low samples.** Each dot represents the number of predicted epitopes, as predicted by netMHCpan and netChop for HLA class I binding prediction and proteasomal cleavage prediction, respectively, in one patient sample. The bars represent the median for the corresponding patient groups. The difference in the number of HLA class I epitopes between TIL-high and TIL-low samples is not significant in the analyzed patient cohort ( $p = 0.925$ , two-sided student's t-test).

Section 1, one HLA class II binding core can have different flanking residues, resulting in multiple HLA class II binding peptides. The total number of distinct binding cores is 2986, hence much lower than the number of binding peptides. A detailed overview of comparing the total number of predicted binding peptides and binding cores for each analyzed patient is shown in Figure 3.13.

Comparing the numbers of HLA class I binders versus HLA class II binders shows that the analyzed 18 HLA class II alleles bind 4-fold more peptides than the analyzed 54 class I alleles (1.5-fold, when only binding cores are considered).

Comparing the HLA class II epitope frequencies between TIL-high and TIL-low samples shows that the numbers are scattered in a range between 15 and 299 for the TIL-low samples, and between 122 and 624 for the TIL-high samples (Figure 3.14). The median is 238 for the TIL-high samples and 178.5 for the TIL-low samples. Statistically, there is no significant difference ( $p = 0.0561$ , two-sided student's t-test).



**Figure 3.13: Frequency of peptides predicted to be bound by HLA class II.** For each sample, the frequency of peptides predicted to be bound by one of the patient’s HLA Class II molecule (turquoise) and the corresponding frequency of distinct binding cores (red) are illustrated.

### 3.4.3 Shared Epitopes

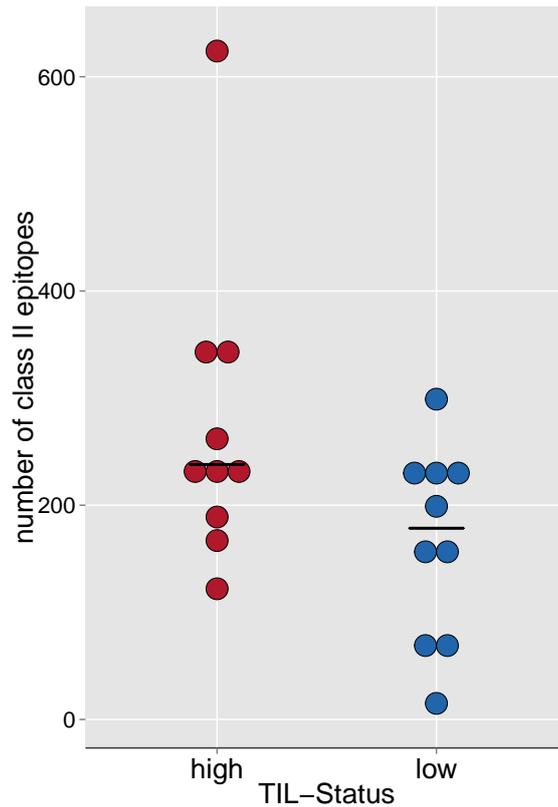
Analogous to the mutational landscape, the epitope landscape is also very heterogeneous and there is no epitope which is frequently shared between patients. Even if a subset of patients have the same mutation, they do mostly not result in the same epitopes because of different HLA alleles.

APC for instance is mutated in 18 patients, however except for one mutation, all mutations are stopgain mutations which do not generate new epitopes.

TP53 in contrast, has some missense mutations that are shared between two patients. The mutation R282W for instance is shared between patients Lo-Pat-05 and Hi-Pat-09, however the HLA genotype of these patients is completely different (Table 3.4) and the mutation does not generate a shared neopeptide. The mutation R248Q is also shared between two patients: Lo-Pat-04 and Hi-Pat-05. These patients also have a different HLA genotype, nevertheless, the mutation generates several shared neopeptides which are bound by different HLA alleles. The peptide MNQRPIITI for instance, is bound by HLA alleles DRB1\*13:01, DRB1\*01:02, and HLA-B\*14:02.

Similarly, the KRAS mutation G12A is shared between Hi-Pat-01 and Hi-Pat-10, and a shared neopeptide is generated for HLA alleles DRB1\*01:01 and DRB1\*01:02.

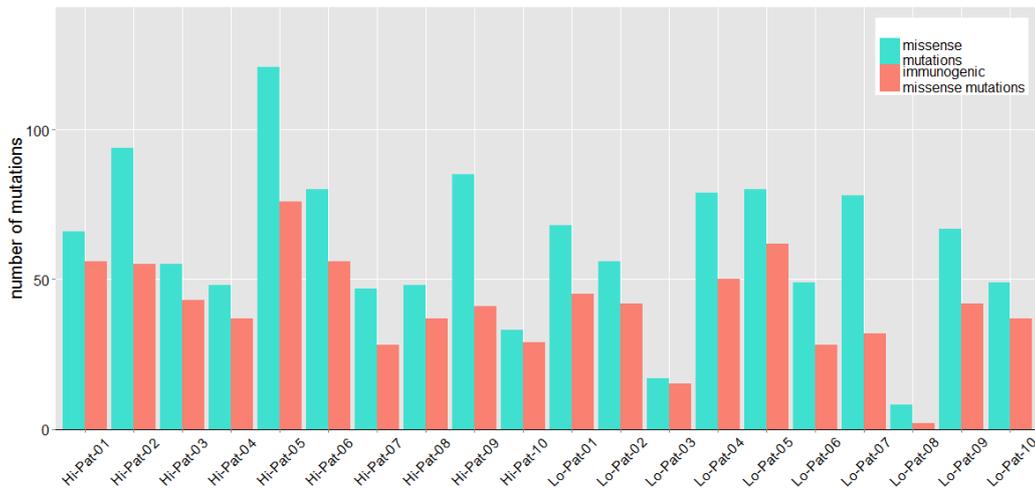
In fact, it occurs quite frequently that a peptide is bound by multiple HLA alleles. More than 1400 peptides are bound by two or more HLA alleles, and more than 300 by three or more alleles. The peptide YLSLGLAF for instance, which is generated by the mutation L157F in the gene GCGR, was predicted to be bound seven different HLA alleles. This mutation however occurs only in patient Hi-Pat-10 and all seven HLA alleles correspond to this patient.



**Figure 3.14: Comparison of HLA class II epitope frequencies between TIL-high and TIL-low samples.** Each dot represents the number of predicted HLA class II epitopes, as predicted by netMHCIIpan for HLA class II binding prediction, in one patient sample. The bars represent the median for the corresponding patient groups. The difference in the number of HLA class II epitopes between TIL-high and TIL-low samples is not significant in the analyzed patient cohort ( $p = 0.0561$ , two-sided student's t-test).

### 3.4.4 Immunogenic Mutations

Not all mutations generate a peptide which is bound by HLA. Mutations that result in at least one neoepitope are referred to as immunogenic mutations. In our patient cohort, the number of immunogenic mutations ranges between two and 62 for TIL-low patients and between 28 and 76 for TIL-high patients (Figure 3.16). In TIL-low patients and average of 63% of mutations are immunogenic, while in TIL-high samples an average of 70% of mutations are immunogenic. Comparing the frequencies of immunogenic mutations between the two patient groups showed no significant difference ( $p = 0.3243$ , two-sided student's t-test).



**Figure 3.15: Frequency of all missense mutations compared to immunogenic mutations.** For each sample, the frequency of all missense mutations (turquoise) are shown and compared to those mutations, which generate at least one epitope (immunogenic mutation, in red).

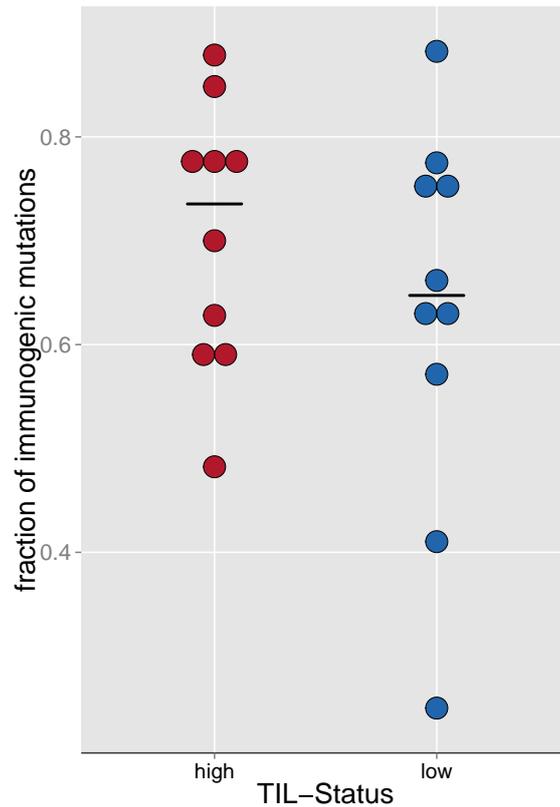
### 3.5 Histological Analysis

Immunohistochemical stainings were performed for selected immunomodulatory proteins. The qualitative evaluation of the stainings is shown in Table 3.5. All samples were infiltrated by macrophages, as indicated by the CD68 stainings. The HLA stainings were almost all found to be heterogeneous, indicating that not all cells express HLA. Similarly, the expression of FOXP3, PD-1 and PD-L2 were found to be mainly weak and heterogeneous. PD-L2 in contrast is expressed strongly in several samples.

### 3.6 Cytokine Analysis

To examine the associations between cytokine expression, TIL-status and neoepitope load, unsupervised hierarchical clustering analysis was performed separately for each compartment. In the heatmaps for adjacent liver and invasive margin the patients are divided into two main clusters of 10 patients each, which mainly correspond to the TIL-status. Only Hi-Pat-01 and Lo-Pat-07 do not cluster according to their TIL-status, as Hi-Pat-01 is clustered together with TIL-low patients, and Lo-Pat-07 with TIL-high patients. The heatmap for the liver metastasis looks slightly different; although the patients are still clustered in two main clusters, the clusters are not of the same size and the correspondence to the TIL-status is not given as clearly as for the other two compartments (Figure 3.17). No clustering of patients with high neoepitope load was detected.

When clustering of the cytokines is considered, various smaller clusters of cytokines are apparent, and it is noticeable that there are several cytokines which are expressed at a



**Figure 3.16: Comparison of immunogenic SNV frequencies between TIL-high and TIL-low samples.** Each dot represents the number of mutations, which generate at least one epitope (immunogenic mutation), in one patient sample. The bars represent the median for the corresponding patient groups. The difference in the number immunogenic mutations between TIL-high and TIL-low samples is not significant in the analyzed patient cohort ( $p = 0.3242$ , two-sided student's t-test).

higher level in patients from the TIL-high group. To further investigate the variance of cytokine expression between the two groups TIL-high and TIL-low, statistical testing was performed: for each cytokine and each compartment, data of all TIL-high and TIL-low patients were pooled, respectively, and significant differential expression between the two groups was assessed using a two-sided student's t-test. 12 cytokines were found to be overexpressed in the TIL-high group in at least two compartments: SDF-1a (CXCL12), MIG (CXCL9), MCP-3 (CCL7), IL-17, IFN $\gamma$ , IL-13, IL-7, IL-4, GM-CSF, HGF, CCL27, and TRAIL (Figure 3.18). Additionally, the following cytokines were found to be overexpressed in the TIL-high group in one compartment:

- IP-10 (CXCL10) in AL ( $p = 0.0326$ )
- b-NGF in IM ( $p = 0.0174$ )
- GRO $\alpha$  (CXCL1) in AL ( $p = 0.0264$ )
- IL-2 in AL ( $p = 0.0161$ )

**Table 3.5:** Immunohistochemical staining for immunomodulatory proteins.

PID	CD68	HLA	FOXP3	PD-1	PD-L1	PD-L2
Hi-Pat-01	+++	+/-	+	+/-	+	+
Hi-Pat-02	+++	++/-	+	+/-	+	+
Hi-Pat-03	++	+/-	+	+	+++	(+)
Hi-Pat-04	+++	+/-	(+)	+ (intra)	-	+
Hi-Pat-05	+++	++/-	(+)	+	+++	+/-
Hi-Pat-06	+++/++++	+/-	+	+	+++	(+)
Hi-Pat-07	+++	++/-	+	+	+	+/-
Hi-Pat-08	+++/++++	++/-	+	+	+++	+/-
Hi-Pat-09	NA	NA	NA	+	+	(+)
Hi-Pat-10	++/-	+	+	+/-	(+)	+
Lo-Pat-01	++	+/-	+	+	(+)	+ (intra)
Lo-Pat-02	++	+/-	+	+ (intra)	+	+/-
Lo-Pat-03	++	++	+	+	+	+
Lo-Pat-04	++	++/-	(+)	(+)	+	(+)
Lo-Pat-05	+++/++++	++/-	+	+/-	+	+/-
Lo-Pat-06	++	+/-	+	(+)	+	(+)
Lo-Pat-07	++	+/-	+	+	+	(+)
Lo-Pat-08	+	+/-	+	+	-	+/-
Lo-Pat-09	+++	++/-	+	+ (intra)	+	+
Lo-Pat-10	++	++/-	(+)	+	+	+ (intra)

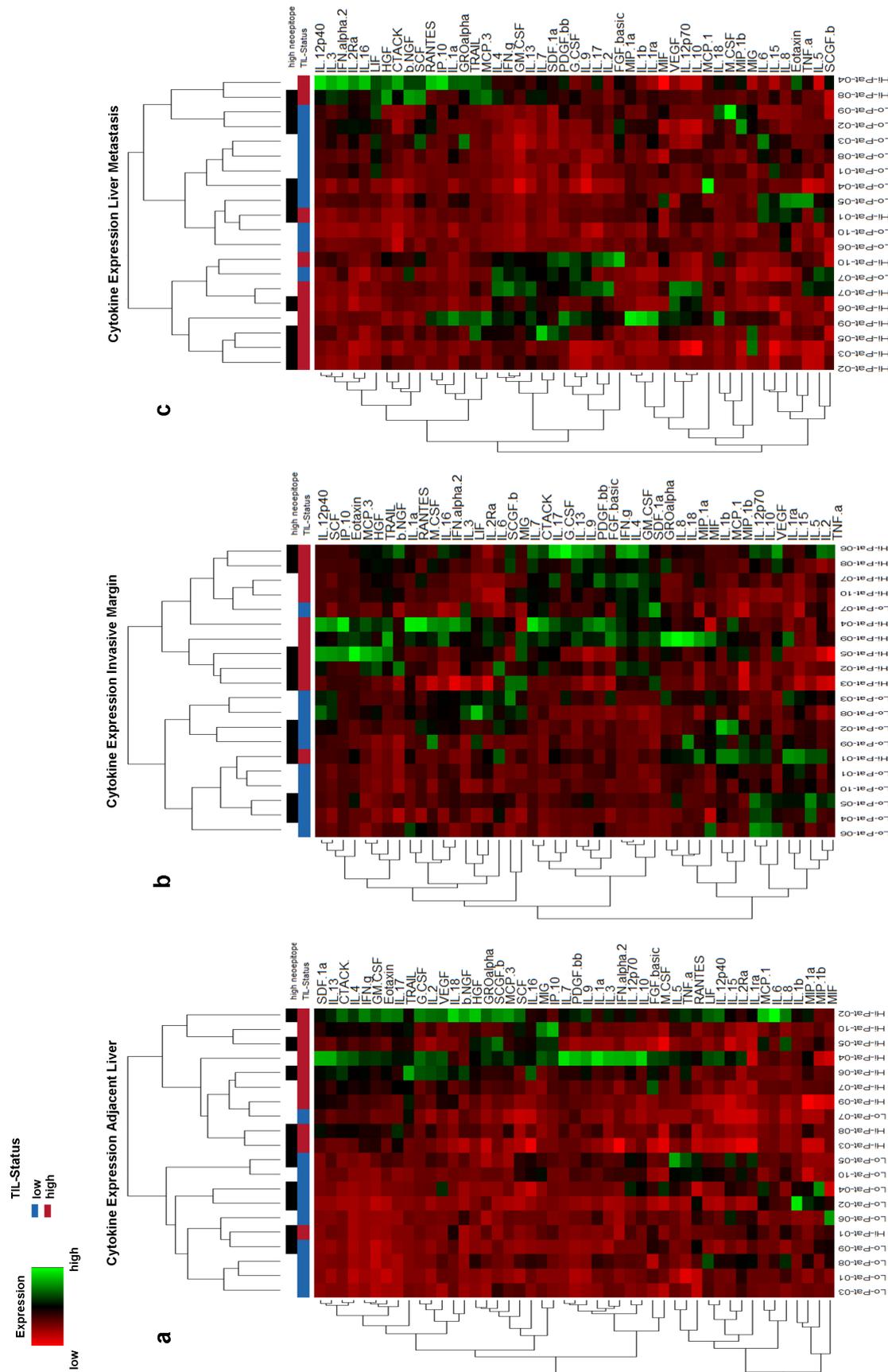
(+) very weak staining; +, weak staining; ++, intermediate staining; ++++, strong staining; +/- heterogeneous staining; intra, intratumoral staining; NA, no staining available

- IL-9 in AL (p = 0.0476)
- Eotaxin (CCL11) in AL (p = 0.0036)
- G-CSF in AL (p = 0.0429)
- PDGFbb in LM (p = 0.0446)

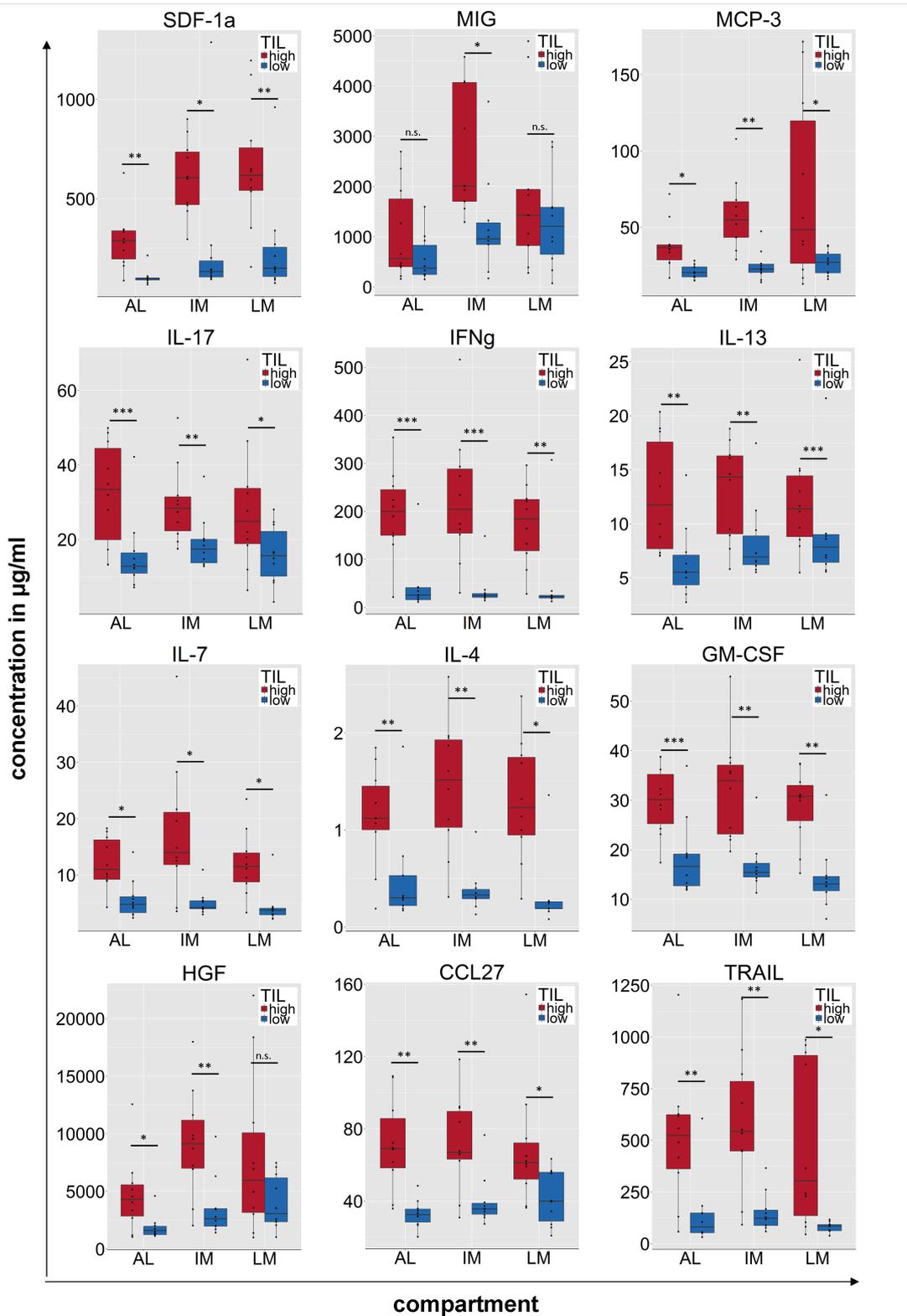
IL-5 is the only cytokine that is overexpressed in the TIL-low group at the invasive margin (p = 0.04).

As it has been reported previously that mutations in KRAS and TP53 can have an impact on cytokine expression, this was also investigated and the cytokine expression in patients with mutations were compared to patients without mutations in KRAS and TP53. Of note, TP53 is mutated in 13 patients and KRAS in 5 patients, so the groups with and without mutation are not exactly of comparable size. Cytokine expression at the invasive margin was not found to be significantly altered in patients with mutations in KRAS or TP53. Some differences were detected in cytokine expression levels in the liver metastasis: in patients with KRAS mutations, the expression of the cytokines

RANTES, MIP-1b, LIF, and IL-16 is significantly reduced ( $p < 0.05$  two-sided student's t-test).



**Figure 3.17: Hierarchical clustering analysis of cytokine expression.** Cytokine detection was performed on tissue lysates of the 20 patients for a panel of 48 cytokines. To examine the associations between cytokine expression, TIL-status and neopeptide load, unsupervised hierarchical clustering analysis was performed on z-score transformed expression values, separately for each compartment. TIL, tumor-infiltrating lymphocyte. High neopeptide load, number of neopeptides in patient > median of all neopeptide numbers.



**Figure 3.18: Cytokines significantly overexpressed in patients with high lymphocyte infiltration.** Detailed description on following page.

**Figure 3.18:** (Figure on previous page.) Cytokine detection was performed on tissue lysates of the 20 patients for a panel of 48 cytokines. For each cytokine and each compartment, data of all TIL-high and TIL-low patients were pooled, respectively, and significant overexpression between the two groups was assessed using a two-sided student's t-test. Shown are the 12 cytokines, where a significant overexpression was detected in at least two compartments. AL, adjacent liver; IM, invasive margin; LM, liver metastasis; TIL, tumor-infiltrating lymphocyte. \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ; two-sided student's t-test.

### 3.7 Case Studies

Two patients with metastatic CRC, which are currently being treated at the NCT could be analyzed in more detail due to continued availability of tumor sample and fresh blood.

HLA typing from the corresponding WXS sequences revealed an HLA genotype of A\*02:01, A\*29:05, B\*07:02, B\*18:01, C\*07:01 (homozygous) for patient B8G6 and A\*03:01, A\*68:01, B\*27:05, B\*44:02, C\*02:02, C\*07:04 for patient VF77.

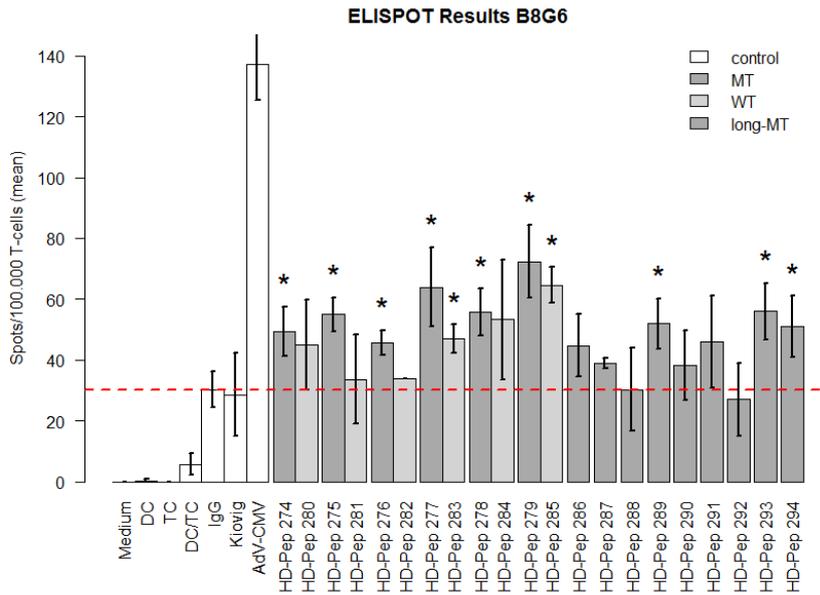
64 somatic-coding SNVs were detected in patient B8G6, 54 of them being somatic exonic missense mutations. In patient VF77 109 somatic coding mutations were found, 100 of them being somatic exonic missense mutations. Of note, driver mutations for CRC were detected in both patients: a stopgain mutation in APC and a missense mutation in TP53 were detected in patient B8G6, whereas a KRAS missense mutation was detected in patient VF77.

HLA class I and class II epitope predictions were conducted as described above. For patient B8G6 35 class I and 61 class II epitopes, for patient VF77 128 class I and 346 class II epitopes were predicted. As RNA sequencing data was available for these patients, I filtered for mutations that were also detected in the RNA sequences. This filtering left 17 class I and 157 class II epitopes for patient B8G6, and 29 class I and 185 class II epitopes for patient VF77.

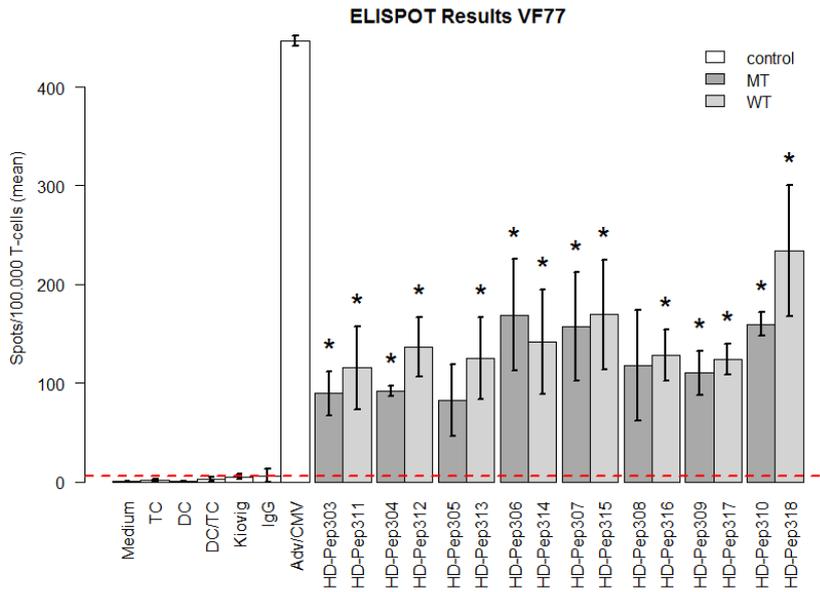
For patient B8G6 six short peptides of 9-11 amino acids containing class I epitopes (HD-Pep-274 - HD-Pep-279) were selected, together with the corresponding wildtype peptides (HD-Pep-280 - HD-Pep-285), as well as long peptides of length 29 which also contain those epitopes together with several flanking amino acids (HD-Pep-286 - HD-Pep-291), and also three long peptides of length 29 containing class II epitopes (HD-Pep-292 - HD-Pep-294) were selected (Table 3.6). For patient VF77 eight long peptides of length 29 containing both, class I and class II epitopes (HD-Pep-303, HD-Pep-304, HD-Pep-307, HD-Pep-309) were selected, one class I epitope (HD-Pep-306), and three class II epitopes (HD-Pep-305, HD-Pep-308, HD-Pep-310), together with the corresponding wildtype peptides (HD-Pep-312 - HD-Pep-319). These peptides were synthesized and analyzed for existing T cell reactivity via ELISpot assays with blood from the corresponding patient.

Quantitative data of the ELISpot assays are shown in Figure 3.18. In both patients, most of the peptides showed significantly higher spot counts than the IgG control and it can be considered that these peptides elicit a T cell response. In patient B8G6, all mutated peptides had higher spot counts compared to the corresponding wildtype peptide, whereas in patient VF77, only one mutated peptide had higher spot counts than its corresponding wildtype counterpart.

The expression of immunological markers was analyzed using the RNA sequencing data and compared to IHC stainings of the respective proteins (Table 3.7).



(a) ELISpot results of patient B8G6



(b) ELISpot results of patient VF77

**Figure 3.18: ELISpot results of case studies.** Peripheral blood dendritic cells were incubated without peptides (DC), and also pulsed with mutated and corresponding wild-type peptides (HD-Pep), as well as with a negative control antigen (IgG), and incubated with autologous T cells. As positive control AdV-CMV and SEB were used (data not shown). ELISpot assays were performed in triplicates. Shown are the quantitative ELISpot results, indicating mean spot numbers  $\pm$  SD for each tested peptide. (a) patient B8G6, (b) patient VF77. Red line shows spot numbers of negative control antigen IgG (considered background). \*, the spot numbers are significantly higher in test wells compared with spot numbers of the negative control antigen IgG ( $p \leq 0.05$ , two-sided student's t-test).

**Table 3.6:** Selected predicted epitopes patient B8G6

HLA	MT peptide ID	MT epitope	MT aff.	NetChop score	WT peptide ID	WT epitope	WT aff.	pos of mut.in epitope	gene	mutation	VAF	TPM
A*02:01	HD-Pep-274, HD-Pep-286	YLVTYDTGMEV	6.1	0.916351	HD-Pep-280	YLGTYDTGMEV	9.06	3	POC1B-GALNT4	NM_001199781: I491V	0.17	25.15
B*07:02	HD-Pep-275, HD-Pep-287	SPQKGRWTEL	6.77	0.934501	HD-Pep-281	SPEKGRWTEL	15.45	3	SNRNP200	NM_014014: E837Q	0.32	184.08
B*07:02	HD-Pep-276, HD-Pep-288	KPRFGLREEQL	16.72	0.855135	HD-Pep-282	KPRPGLREEKL	28.16	10	UBN2	NM_173569: K805Q	0.26	9.53
A*02:01	HD-Pep-277, HD-Pep-289	HLHFNLPPQV	30.97	0.952428	HD-Pep-283	HLRFNLPQV	267.93	3	GAL3ST2	NM_022134: R111H	0.22	8.41
B*07:02	HD-Pep-278, HD-Pep-290	LPRVQVKSF	39.25	0.90818	HD-Pep-284	LPRVQVESF	46.35	7	NOD1	NM_006092: E641K	0.18	15.28
A*29:05	HD-Pep-279, HD-Pep-291	QMYNNLLKLY	47.1	0.952876	HD-Pep-285	QMYNNLLKLY	61.34	4	LRPPRC	NM_133259: Y959N	0.29	184.96
DRB1*11:01	HD-Pep-292	RIYFVTRKNPFRFL	9.76	-	-	LIYFVTRKNPFRFL	10.28	1	SLC1A4	NM_003038: L315R	0.2	23.03
DRB1*11:01	HD-Pep-293	QSAINKLMQMLRSG	27.41	-	-	QSAINELMQMLRSG	107.56	6	SRP68	NM_01423002: E221K	0.27	91.81
DRB1*11:01	HD-Pep-294	LCYTHTFNPKVILNA	86.66	-	-	LCYAHTFNPKVILNA	87.41	4	LAP3	NM_015907: A378T	0.42	121.12

MT, mutated; WT, wildtype; VAF, variant allele frequency; TPM, transcripts per million

**Table 3.7:** Comparison of expression levels assessed by RNA-Seq and IHC

Patient VF77				Patient B8G6			
Gene	RPKM	TPM	Staining	Gene	RPKM	TPM	Staining
CD3	2.61	6.18	+/-	CD3	5.57	13.28	+/-
CD8	1.26	2.98	+	CD8	0.9	2.15	+/-
CD20	0.06	0.13	+	PDL1	0.45	1.07	(+)
PDL1	0.70	1.67	-	CTLA4	0.51	1.21	+/-
CTLA4	0.51	1.2	++	PD1	0.44	1.05	+/-
MHC1	895.46	2117.88	+/-	PDL2	1.72	4.1	(+)
PD1	0.48	1.14	+/-				
PDL2	1.23	2.91	+/-				

(+) very weak staining; +, weak staining; ++, intermediate staining;  
+++, strong staining; +/- heterogeneous staining

## CHAPTER 4

---

Discussion

---

Cancer immunotherapies have recently shown outstanding clinical results in a number of patients across various tumor types. However, currently only a fraction of patients responds to immunotherapy and it is a major concern to understand why. The composition of the tumor microenvironment was shown to have an important impact on tumor growth and progression, as well as on response to therapy. Integrated analysis of multiple tumor and host factors is a promising approach which may help to identify patients who are most likely to benefit from immunotherapy.

Factors that should be considered in an integrated analysis include the mutational and neoepitope landscape, the type and densities of tumor infiltrating immune cells, the expression of immunological markers, and the expression of specific cytokines. In order to efficiently analyze the complex data, as part of this PhD project an integrated analysis pipeline was established that combines already available genomic and immunomic resources and adds further depth into the analysis by additional computational pipelines.

The established analysis pipeline was then used to investigate how lymphocyte infiltration is correlated to mutational and neoepitope load in metastatic lesions of CRC and also to uncover additional factors that might drive lymphocyte infiltration. The pipeline was additionally applied in the clinic to conduct case studies, which paved the way for the development of clinical studies.

## 4.1 Neoepitope prediction and prioritization

The developed computational pipelines employ publicly available immunoinformatics tools for sequencing-based HLA typing and HLA class I and II binding prediction. Immunoinformatics has established itself as a cornerstone in the research and development of cancer immunotherapies. Immunoinformatics tools are mainly used in the context of neoepitope research, for instance to analyze the abundance and effect of neoepitopes in a certain setting of tumor diseases, or also for the development of neoepitope-based vaccines.

Thanks to the advancements in high-throughput sequencing of cancer genomes and established downstream analysis pipelines, it is now possible to rapidly detect and analyze neoepitopes in large patient cohorts. Immunoinformatics tools were used in many studies to analyze large-scale cancer sequencing data in order to predict neoepitopes and identify epitopes with clinical relevance or to select candidates for clinical application (Rizvi et al., 2015; Rooney et al., 2015; Khalili et al., 2012; Gubin et al., 2014).

These studies have used different combinations of tools and applied different criteria to prioritize the set of neoepitopes, including gene expression levels, predicted binding affinity, comparison to wildtype binding affinities, structural properties, and the position of the mutated residue within the HLA binding groove. The ideal criteria for neoepitope prioritization remain to be controversial, as it is not known yet which combinations of criteria are most suitable to detect effective neoepitopes.

The selection of the epitope prediction tool and the definition of criteria for neoepitope prioritization were also one of the first issues discussed for our study. As netMHCpan (Nielsen et al., 2007) was repeatedly shown to be the best-performing tool in several benchmarking studies (Trolle et al., 2015), this tool was chosen for HLA class I binding prediction for our study.

Additionally, we decided to use three criteria for epitope prioritization:

- predicted binding affinity of  $\leq 500\text{nM}$
- predicted proteasomal cleavage score of  $\geq 0.5$
- predicted affinity of mutated peptide is higher than corresponding wild-type peptide

The only criteria common to all studies is the prioritization of neoepitopes with strong binding affinities (Rizvi et al., 2015; Rooney et al., 2015; Gubin et al., 2014; Yadav et al., 2014). For our study, the cut-off of 500nM was used, which is commonly used across various studies. It was also suggested previously that different HLA alleles bind their peptides with different affinities and that each allele has a unique threshold. However, a large-scale study on the peptide binding repertoire of various HLA class I alleles showed that 500nM is a suitable universal threshold for HLA-binding prediction (Paul et al., 2013).

The second prioritization is to only use neoepitopes that were predicted to be cleaved by the proteasome. Although presentation of peptides on HLA is the bottleneck in antigen processing and presentation, previous steps in the antigen processing machinery also have a major impact on the peptide repertoire that is presented on HLA and should be considered for neoepitope prioritization.

The third prioritization criteria used in our study is that only facilitating epitopes were considered, i.e. mutated peptides which have stronger binding affinities than their corresponding wildtype counterpart. This filter reduced the number of potential neoepitope candidates drastically (Figure 3.11). The peptide binding affinity mainly depends on the properties of the amino acids located at the so-called anchor positions of the HLA peptide binding groove (Fruci et al., 1993). Hence, the binding affinity of neoepitopes with mutations in anchor positions will most likely differ from the binding affinity of the wildtype peptide, while mutations in non-anchor positions most likely have no effect on binding affinity. Mutations in anchor positions can increase the binding affinity of a peptide, but, depending on the specific amino acid change, can also decrease the affinity. Interestingly, it has been indicated in several studies that facilitating neoepitopes are more effective in eliciting specific T cell responses (Matsushita et al., 2012; Khalili et al., 2012). The theory behind this is that if HLA binding affinity of a peptide is increased due to a mutation, this may translate into increased presentation of the mutant peptide, because during the loading of peptides on HLA, competition between the mutated and wildtype peptides can occur and the peptide with the stronger affinity is advantageous in this scenario (van der Burg et al., 1996; Busch and Pamer, 1998; Khalili et al., 2012).

Another useful filter is to prioritize neoepitopes which originated from a highly expressed gene, as it was shown that peptides derived from abundant transcripts are more likely to be presented by HLA class I (Fortier et al., 2008; Granados et al., 2012; Yadav et al., 2014). Unfortunately, no expression data was available for the CRC liver metastases dataset used in this study. This filter was included into the immunoinformatics analysis pipeline as an additional option. Mutations at the DNA level are not always transcribed to RNA, and it also occurs that the mutated gene is not expressed at all. If a mutation was detected in DNA sequences, it does not necessarily occur in the RNA sequences, as demonstrated in the data of the two case studies. Filtering out mutations that were not detected in the RNA-Seq data removed more than half of the potential neoepitopes.

Another point to take into account is that many peptides which elicit a T cell response are not produced by the antigen-processing machinery, but by peptidases and other enzymes (Parmentier et al., 2010; Kessler et al., 2011). Also, the proteasome composition of the cell is crucial for antigen processing: different proteasomes can result from different combinations of subunits, and it has been reported that several tumor antigens are produced by only one specific type of proteasome (Morel et al., 2000; Schultz et al., 2002; Chapiro et al., 2006; Guillaume et al., 2012). Additionally, proteasomes can splice peptide fragments that are located in distance in the parental protein, and generate novel stretches of amino acids, which might be highly tumor-specific (Hanada et al., 2004; Vigneron et al., 2004). The contribution of these alternatively produced antigens to tumor immunogenicity has not been analyzed yet in a comprehensive manner and remains to be elucidated.

As all epitope prediction algorithms were developed using data on pathogen-derived epitopes and not with tumor-specific data, it was also suggested that they might not be suitable to analyze tumor neoepitopes. Duan et al. analyzed neoantigens in a transplantable mouse tumor model and argued that those neoantigens display very different properties from viral antigens, and that a specific approach is required to successfully detect tumor neoepitopes (Duan et al., 2014). However, this approach was not able to identify neoantigens that have previously been proven to induce tumor rejection (Gubin et al., 2014), and the proposed approach by Duan et al. is not widely accepted (Schumacher and Schreiber, 2015).

Another interesting hypothesis was suggested by Snyder et al. (Snyder et al., 2014). They analyzed the neoantigen repertoire in melanoma patients receiving the anti-CTLA-4 antibody ipilimumab and showed that the neoantigens in patients with a long-term clinical benefit were enriched for a set of tetrapeptide motives. The hypothesis behind this observation was that neoepitope-specific T cells preferentially recognize a subset of mutated peptides. However, this observation could not be confirmed by others as the neoantigens identified in other studies did not show the bias toward the tetrapeptide signature (Schumacher et al., 2015).

In conclusion, the improvement of the efficiency of neoepitope identification and selection is still an ongoing challenge, and researches have suggested several parameters for prioritization of neoepitopes. Additionally, there are still some unknown aspects about the antigen processing and presentation in tumor cells. The availability of more

experimental cancer-specific data and also a better understanding of the biochemical properties of effective neoepitopes will improve the prediction and selection of appropriate neoepitopes (Delamarre et al., 2015).

#### 4.1.1 HLA class II Neoepitopes

For a long time, researchers focused mainly on HLA class I epitopes recognized by CD8<sup>+</sup> cytotoxic T cells when analyzing neoepitopes for cancer immunotherapy. However, there is increasing evidence that recognition of HLA class II neoepitopes by CD4<sup>+</sup> T helper cells occurs in cancer, and that their interaction is required to further boost the activity of CD8<sup>+</sup> T cells targeting neoantigens (Delamarre et al., 2015; Kreiter et al., 2015; Linnemann et al., 2015).

Tran et al. published a case report where measurable tumor regression was demonstrated in a patient with metastatic cholangiocarcinoma (Tran et al., 2014). In this study, CD4<sup>+</sup> T cells that were specific for a HLA class II neoepitope were expanded ex vivo and re-infused to the patient. The success of this approach underscores the potential clinical relevance of HLA class II neoantigens.

Sahin et al. conducted a study in mouse models of melanoma, lung, and colon cancer and identified more than 500 mutations in each tumor (Kreiter et al., 2015). They chose to focus on 50 random mutations and the corresponding HLA class I neoepitopes. Surprisingly, they found that about 20% of the mutations triggered an immune response which is most frequently mediated by CD4<sup>+</sup> T cells. Apparently, HLA class II molecules on antigen-presenting cells present these neoepitopes, and CD4<sup>+</sup> T cells recognize them and induce an antitumor response. They then incorporated 10 neoepitopes in mRNA vaccines and vaccinated the mice. It was demonstrated that the T cells were capable of eliminating even established, growing tumors, causing the mice to live longer. To test the efficacy of vaccines based on HLA class II neoepitopes, Sahin and colleagues are currently running two phase I clinical trials, in patients with melanoma (NCT02035956) and triple-negative breast cancer (NCT02316457).

In fact, neoantigen-specific CD4<sup>+</sup> T cells were shown to augment the clonal expansion of neoantigen-specific CD8<sup>+</sup> T cells (Wong et al., 2008). Moreover, the CD4<sup>+</sup> T cells in the tumor microenvironment facilitate infiltration by tumor-specific CD8<sup>+</sup> T cells, augmenting their antitumor activity (Bos and Sherman, 2010). Thus, CD4<sup>+</sup> T cell activity helps to amplify the CD8<sup>+</sup> T cell response in both initiation and effector phase of an immune response following vaccination, and enhancing CD4<sup>+</sup> T cell activity should not be underestimated when designing therapeutic cancer vaccines (van den Boorn and Hartmann, 2013).

Due to the open binding groove and the resulting higher diversity of HLA class II epitopes, prediction algorithms are not as accurate as for HLA class I epitopes (Backert and Kohlbacher, 2015). Nevertheless, using the HLA class II epitope prediction algorithm NetMHCIIpan, neoepitopes could be detected that augmented T cell response in mice (Kreiter et al., 2015). HLA class II binding predictions using NetMHCIIpan was

also included in our study and is included as part of the immunoinformatics analysis pipeline.

## 4.2 Sequencing-based HLA Typing

To apply epitope prediction tools, knowledge of the HLA genotype is required. As in many clinical applications sequencing data of a patient is already available, costs for experimentally determining the HLA genotype can be avoided by using computational tools. Several tools for NGS-based HLA typing emerged (Boegel et al., 2012; Warren et al., 2012; Kim and Pourmand, 2013; Bai et al., 2014). The accuracy of these sequencing-based HLA typing tools is still being critically discussed and is not yet fully convincing (Backert and Kohlbacher, 2015). One of the most recent tools, Phlat (Bai et al., 2014), was chosen for HLA typing from the exome-sequencing data of our patient cohort. Phlat was shown over a comprehensive set of benchmarking data from both RNA and DNA to achieve a high accuracy of 92%-95% at four-digit resolution, outcompeting most of the existing methods (Bai et al., 2014). We also assessed the accuracy of Phlat on two patients where HLA genotyping data was available from conventional PCR-based typing techniques. For these two patients, Phlat achieved 100% accuracy for HLA class I typing, and 75% for HLA class II.

More recently, Shukla and colleagues performed a comparative analysis of various tools on over 250 samples, and found their own tool PolySolver together with the tool OptiType (Szolek et al., 2014) to be the most accurate tools for HLA class I inference, out competing Phlat and other previous tools in sensitivity, precision, and accuracy (Shukla et al., 2015). The used dataset for benchmarking consists of 253 samples with known HLA genotypes from the HapMap projects (Consortium, 2003). As a matter of fact, many of the sequencing-based HLA typing tools have been trained and developed using this exact dataset. Due to the time and cost consuming nature of conventional PCR-based HLA typing methods, there is a lack of appropriate training and benchmarking datasets. Availability of additional large-scale datasets will reveal the true accuracy and value of sequencing-based HLA typing tools. Another issue is that due to the higher complexity of the genomic locus of HLA class II, the inference of the HLA class II genotype with sequencing-based HLA typing tools remains to be limited (Backert and Kohlbacher, 2015).

## 4.3 The local microenvironment of CRC liver metastases

### 4.3.1 Mutational landscape

Analysis of somatic mutations and chromosomal aberrations in our patient cohort showed alterations that are well-known and have been reported repeatedly for colorectal

cancer (Network, 2012) and therefore will not be discussed here. Importantly, no mutation was found to be specific for one of the groups TIL-high and TIL-low.

One aim of this study was to analyze the correlation of the number of mutations with lymphocyte infiltration, as several studies have reported a strong correlation in various tumor types (Lu and Robbins, 2015; Brown et al., 2014; Rooney et al., 2015). The presence of CD8<sup>+</sup> T cells in cancer lesions was shown to be higher in tumors with a high mutational burden, and it was also shown that treatment efficacy of immunotherapies is associated with a higher number of somatic mutations (Rizvi et al., 2015; Snyder et al., 2014).

In our study, the previously reported strong correlation between mutational load and lymphocyte infiltration could not be confirmed. As shown in Figure 3.5, the median number of missense mutations is equal in the two groups. Nevertheless, when the mutational load of the samples from the two groups are considered in detail, a trend can be detected. In fact, the samples with the highest mutational burden (Hi-Pat-02, Hi-Pat-05, Hi-Pat-08) are from the TIL-high group, and the samples with the lowest mutational burden (Lo-Pat-03, Lo-Pat-08) are from the TIL-low group. We observed a similar picture when predicted neoepitopes or the number of immunogenic mutations were considered (Figures 3.12, 3.14, 3.16). Although no significance was detected, a trend can be observed, mainly because the bottom outliers are from the TIL-low group and the top outliers are from the TIL-high group.

Considering our results, the obvious question is how the discordance between our results and previous reports on a strong correlation between mutational/neoepitope load and lymphocyte infiltration can be explained.

Numerous studies that have reported a positive correlation of mutational and/or neoepitope load with lymphocyte infiltration and/or prognosis in various cancer patients, were mainly based on assessing the lymphocyte infiltration from RNA-Seq data. Brown et al. defined the CD8<sup>+</sup> infiltration defined as the abundance of CD8A transcripts (Brown et al., 2014), and Rooney et al. assessed the activity of lymphocytes by combining the mRNA expression of granzyme B and perforin (Rooney et al., 2015). Although these types of studies allow the analysis of huge amounts of data, they should be handled with some caution, as high abundance of transcripts may not translate to high abundance of protein. Recently, a large-scale study on proteogenomic characterization of colorectal cancer revealed that mRNA level is not always correlated to protein level (Zhang et al., 2014).

### 4.3.2 Shared epitopes

The peptide repertoire presented on HLA is largely determined by the structure of the peptide binding groove 1.1. It is expected that the HLA molecules having similar grooves might present similar/overlapping peptides. These peptides are referred to as promiscuous binders (Sturniolo et al., 1999). In a large-scale study the HLA peptide binding data from the IEDB was analyzed which revealed that an unexpectedly large

fraction of HLA ligands (>50%) bind two or more HLA molecules, often across HLA supertypes or even loci (Rao et al., 2011).

In fact, we made a similar observation in our patient cohort as it occurs quite frequently that a peptide is bound by multiple HLA alleles. More than 1400 peptides are bound by two or more HLA alleles, and more than 300 by three or more alleles. The peptide YLSLGLGF for instance, which is generated by the mutation L157F in the gene GCGR, was predicted to be bound by seven different HLA alleles. This mutation however occurs only in patient Hi-Pat-10 and all seven HLA alleles that bind the epitope correspond to this patient. Accordingly, the occurrence of shared epitopes is a combinatorial question, involving the patient's HLA genotype and the mutation. However, as the frequency of shared mutations is quite low to begin with, the frequency of shared epitopes is expected to be much lower.

These findings underline the necessity of a personalized approach for neoepitope-based immunotherapies.

### 4.3.3 Alterations in immune-related genes

It was reported previously that inactivation of HLA class I molecules is a frequent mechanism, of immune evasion in several cancers (Schumacher and Schreiber, 2015). Loss of HLA reduces the ability of cancer cells to present antigens to T cells and therefore leads to failure of peptide-based cancer vaccination, for instance. Loss of HLA can occur through mutations in HLA genes or through downregulation of those genes. Abnormal expression of HLA class I molecules in tumor cells is a frequent event which can occur as total loss of class I molecules or partial loss of specific haplotypes or alleles (Cabrera et al., 2007). In colorectal cancer beta2-microglobulin seems to be especially affected in this context, as Cabrera et al. reported frequent downregulation of this molecule in microsatellite instable CRC (Cabrera et al., 2003), and Tikidzhieva et al. reported that beta2-microglobulin is mutated in about 30% of microsatellite instable CRC patients (Tikidzhieva et al., 2012). We analyzed the expression of HLA class I molecules in our patient cohort and found that HLA is expressed in all patients, as weak to intermediate stainings were observed in the immunohistochemical analysis 3.5. Of note, the stainings were found to be heterogeneous, indicating that not all cells express HLA uniformly.

Immunomodulatory genes, as well as cytokines and cytokine receptors were additionally analyzed for accumulation of mutations, but no recurrent mutations were found. A PD-L1 mutation was detected in Lo-Pat-10; nevertheless, according to the immunohistochemical analysis a weak PD-L1 expression is evident in this patient. Similarly, we also analyzed if these genes of interest are frequently affected by chromosomal instability, i.e. amplifications or deletions. The regulatory T cells marker FOXP3 was found to be amplified in seven TIL-high but no TIL-low patients. However, this amplification does not seem to translate in elevated protein expression, as very weak to weak stainings were observed in all patients, regardless if they harbor a FOXP3 amplification or not.

Taken together, the immunomodulatory pathways do not seem to be affected by single alterations on the genomic level. Furthermore, the few detected genomic changes do not seem to translate into altered protein expression, as indicated by immunohistochemical staining analysis.

#### 4.3.4 Pathways enriched for mutations

As a next step, pathway enrichment analysis was performed with the set of mutated genes for the TIL-high and low groups separately (section 3.2.4). The results clearly show that there are differences in affected pathways between the two groups. Interestingly, in the TIL low group two immune-associated pathways are in the top 10 enriched pathways, namely "acute phase response signaling" and "PI3K signaling in B lymphocytes", and according to the Ingenuity Pathway Analysis (IPA) tool, these pathways are enriched in the TIL-low group when compared to the TIL-high group.

Although IPA suggests a higher enrichment of mutations relevant for these pathways in the TIL-low group compared to the TIL-high group, a closer look at the mutated genes revealed a different picture. For instance, in the TIL-low group 14 genes associated with acute phase response are mutated, whereas in the TIL-high group nine genes are mutated, with the genes FN1, PIK3CA and KRAS being mutated in both groups. Hence, the enrichment was found to be not very consistent with the pathway models and was not further considered.

Of note, from the sequencing data alone it is not possible to sufficiently assess the effect of a somatic mutation in cancer yet. It remains thus unclear whether the mutation activates or impairs the protein function, or has any function at all. There are several tools developed to predict the functional effects of single nucleotide polymorphisms; however, a benchmark study assessing the performance of 15 predictors revealed that no algorithm is able to sufficiently accurately predict the effects of missense mutations in cancer (Martelotto et al., 2014).

#### 4.3.5 Association of HLA allele and neoepitope load

High mutational load does not necessarily lead to a high neoepitope load as neoepitope presentation also depends on the specific HLA genotype. It is known that because of the different binding preferences of different HLA molecules, there is variation in peptide-repertoire diversity, meaning that some HLA alleles can bind a more diverse set of peptides compared to other HLA alleles (Yanover and Bradley, 2011). We have also made this observation, as for instance the patient with the highest amount of mutations, patient Hi-Pat-05, does not have the highest amount of predicted neoepitopes. Additionally, patient Lo-Pat-07 has the second lowest number of predicted epitopes, despite having a relatively high number of mutations. Lo-Pat-07 is the only patient who has homozygous HLA-B and HLA-C alleles; hence it seems obvious that the peptide diversity is smaller for this patient.

In fact, the role of HLA in resistance or susceptibility to cancer is not well understood yet. Only few studies report association between HLA genotype and prognosis in cancer patients. In the US two HLA Class II genes, HLA DQB\*03032 and HLA DRB1\*11, have been identified that may have a protective role in human breast cancer in Caucasians (Chaudhuri et al., 2000). Also in cervical cancer, DQB1\*0602 and DRB1\*1501, are associated with the presence of cancer in Hispanic patients, whereas HLA alleles, DRB1\*1301 and \*1302, were reported to protect against cancer induced by HPV infection (Garca-Corona et al., 2004).

With the recently appreciated relevance of neoepitopes for cancer immunotherapy, HLA genotyping has become a routine step in several clinical applications. The availability of large scale HLA genotyping and clinical data of cancer patients will allow the analysis of the correlation between HLA genotype and prognosis.

### 4.3.6 Microenvironment

Data from primary tumors identified the cytokines CXCL10, CXCL9, and CX3CL1 to be the key factors for lymphocyte recruitment, but it was reported that this cytokine regulation is weakened in the metastatic setting (Halama et al., 2011a). These reports are in line with our findings: MIG (CXCL9) is indeed significantly elevated at the invasive margin in the TIL-high group ( $p = 0.02$ ), and IP-10 (CXCL10) is not significantly elevated at the invasive margin. Instead, a set of 12 cytokines were found to be significantly higher expressed in the TIL-high group when compared to the TIL-low group: SDF-1a (CXCL12), MIG (CXCL9), MCP-3 (CCL7), IL-17, IFN $\gamma$ , IL-13, IL-7, IL-4, GM-CSF, HGF, CCL27, and TRAIL.

These elevated cytokines are mainly associated with lymphocyte recruitment, however they do not show a clear pattern specific for any subpopulation, as cytokines produced by Th1 T cells, as well as cytokines produced by Th2 T cells are overexpressed: while IFN $\gamma$  is produced by Th1 T cells, IL-4 and IL-13 are produced by Th2 T cells. Hence, interestingly both, protumorigenic as well as antitumorigenic factors are overexpressed in the TIL-high group.

Several of the elevated cytokines, such as CXCL12, CXCL9, CCL7, IL-13, and GM-CSF are mainly produced by macrophages and are associated with leukocyte and lymphocyte recruitment. In our analyzed patient cohort, immunohistochemical stainings for the macrophage marker CD68 revealed the presence of macrophages at the invasive margin of all analyzed samples. However, the density of macrophages in the TIL-high group seems to be higher than in the TIL-low group, as almost all CD68 stainings were strong in the TIL-high group, whereas only intermediate staining was observed in the TIL-low group (Table 3.5). Hence, the presence of more macrophages in the TIL-high group could explain the overexpression of several cytokines. Macrophages that infiltrate a tumor are often called tumor-associated macrophages (TAMs). The function of TAMs is controversial as there is growing evidence showing that they have both, protumorigenic as well as antitumorigenic properties (Bingle et al., 2002). Professor Jäger's group just recently conducted a study showing that in metastatic CRC

macrophages play a major role in producing tumor-promoting factors (Niels Halama et al., in press).

The cytokines IFN $\gamma$  and TRAIL are both significantly elevated in all compartments in the TIL-high group, indicating functional activity of the infiltrating T cells. Additionally, according to the immunohistochemical analysis, PD-L1 seems to be expressed in all samples. Interestingly, four samples were found to express PD-L1 strongly, and these are all TIL-high samples: Hi-Pat-03, Hi-Pat-05, Hi-Pat-06, and Hi-Pat-08. As the PD-1 signaling cascade contributes to T cell exhaustion, a strong prevalence of PD-L1 might indicate that the present lymphocytes in these highly infiltrated samples are already exhausted and their antitumor activity is weakened or even absent. These patients might especially benefit from PD-1 and PD-L1 blockade therapy as it was reported that presence of PD-L1 is correlated with response to therapy (Patel and Kurzrock, 2015).

## 4.4 Further applications of the developed analysis pipeline

The analysis pipeline developed in this study adds valuable information to the analysis of cancer patient data and the exploration of additional treatment options. The pipeline is already being used to assess the immunogenicity of tumor samples from several large cohorts of patients with rare cancers, which have not been analyzed in this manner previously. These studies will reveal the potential immunogenicity of those tumor entities and provide new insights on whether immunotherapy would be feasible for affected patients. Additionally, the pipeline is being applied to samples from patients with refractory or rare cancers to explore new therapeutic strategies. The obtained detailed data about tumor-host interactions may provide information on clinical outcome and help identify patients who are most likely to benefit from immunotherapy. The developed pipelines nicely complement the available genomics pipelines. The pipelines are fast and robust and have been nicely integrated into the clinical setting. From the point where the whole-exome-sequencing data is available, prioritized epitope lists can be generated within 48 hours.

### 4.4.1 Case studies

We applied our analysis workflow and performed immunomic profiling of an extremely rare metastatic neuroendocrine tumor of the rectum. This analysis, for the first time, provides comprehensive details about the immunogenicity and microenvironment of a neuroendocrine tumor. A manuscript describing this case report is currently in revision for publication in *Oncoimmunology*.

In this thesis, two case studies of patients with metastatic colorectal cancer have been included to demonstrate and discuss the application of the analysis pipeline (section 3.7). The HLA typing pipeline was used with the whole-exome-sequencing data of

the patients to infer the corresponding HLA genotypes. Subsequently, the epitope prediction and selection pipeline was used to select suitable candidates to include in a potential therapeutic vaccine. We then synthesized selected peptides containing HLA class I or class II epitopes, or both, and analyzed them for existing T cell reactivity via ELISpot assays with blood from the corresponding patient. For patient B8G6 short peptides (9mers) only containing the predicted neoepitope, as well as long peptides (29mers), which have the neoepitope centered and additional flanking amino acids on each side, were analyzed in the ELISpot assay. All mutated short peptides showed significantly higher spot counts than the IgG control and it can be considered that these peptides elicit a T cell response. Also, all mutated peptides had higher spot counts compared to the corresponding wildtype peptide. These results indicate that there is a preexisting T cell response against all of the analyzed neoepitopes, and that they elicit stronger T cell responses than their wildtype counterparts. These peptides would be ideal candidates for therapeutic vaccination aiming to boost and reactivate T cells specific for them. Interestingly, when the spot counts of the short mutated peptides are compared to spot counts of the corresponding long peptides, it is striking that the long peptides have lower counts. Surprisingly, the long peptides seem less reactive when compared to the short peptides, although they contain the exact same predicted neoepitope. An explanation for this might be that the long peptides are not processed as expected, and that the predicted neoepitope is not presented properly on HLA. For patient VF77 only long peptides, containing mainly both, predicted HLA class I and class II epitopes, were synthesized and tested. Here, the spot counts of all tested peptides, mutated and wildtype, strongly exceeded the counts of the IgG control, indicating strong T cell reactivity. Also, almost all wildtype peptides had stronger spot counts than the corresponding mutated peptides, although the predicted affinities of the mutated peptides were all stronger. However, a vaccination with mutated peptides would lead to an oversaturation with them, so that they may be presented more frequently increasing the chance of specific T cells detecting the presented neoepitopes, boosting the neoepitope-specific T cell response.

For a long time, it was controversially discussed among researchers whether enhancement of an existing T cell response or generation of de novo responses is clinically relevant for an effective tumor vaccine (Fritsch et al., 2014). Carreno et al. recently reported that neoepitope vaccination not only can amplify existing CD8<sup>+</sup> T cell responses but also can produce responses that might have been silent prior to vaccination (Carreno et al., 2015; Delamarre et al., 2015). Thus, even if no T cell reactivity was seen for certain neoepitopes in the ELISpot assays, vaccination with them might still generate potent antitumor T cell responses.

If a therapeutic vaccine is going to be administered, including multiple neoepitopes, for both CD4<sup>+</sup> and CD8<sup>+</sup> T cells, is of advantage. In doing so, the likelihood of generating an immune response against at least some of the neoantigens increases, and the likelihood of the tumor escaping the immune response by immunoediting decreases.

Case studies like these were conducted with several patients. In doing so, a working logistics for the clinical setting was established, and the results provided insights into

the feasibility of the approach. Based on these findings, clinical studies with neoepitope-based vaccines are already in development in Professor Jäger's group.

## 4.5 Conclusion and Outlook

We analyzed the correlation of the number of mutations and the number of neoepitopes with lymphocyte infiltration. We could not confirm the strong correlation that was reported previously numerous times. However, a trend can be observed in a way that the TIL-high group seems to be enriched for mutations and neoepitopes, but no statistical significance was detectable. Instead, the cytokine expression profiles as well as the expression of some immunomodulatory proteins are clearly distinct. These results indicate that the immune contexture at the metastatic lesion seems to be the driving force for lymphocyte infiltration instead of mutations and neoepitope landscape.

It might be a concern that the analyzed cohort sizes are too small to yield a significant difference. However, if the relationship between mutation/neoepitope load and lymphocyte infiltration was pivotal, it should also be evident even in a small patient cohort. Differences in the microenvironment in contrast, are clearly evident and statistically significant even in the analyzed small patient cohort. Of note, it is a highly challenging endeavor to conduct all the different analyses as we did in this study. Apart from the technical and logistical challenge, acquisition of an appropriate amount of patient samples with sufficient quality is another big hurdle. Hence, assembly of a bigger patient cohort was not easily possible but is being considered to further validate and underline the results presented here. Here, we performed an explorative study which provides valuable detailed insights to the immune setting of CRC liver metastases and builds the basis for a prospective study.

In the past few years, the microenvironment has received much attention, as a wealth of evidence was reported to support the critical role of the composition of the microenvironment on tumor growth and progression (Smyth et al., 2016). A recent study by Mlecnik and colleagues even postulated that the microenvironment seems to be the cause for metastasis formation (Mlecnik et al., 2016). Our results also clearly highlight the importance of the microenvironment. It has to be kept in mind that the presence of neoantigens does not necessarily induce T cell reactivity. It is known that human tumors vary substantially in the composition of their microenvironment, and this is most likely to influence the ability of T cells to respond to mutated antigens (Schumacher and Schreiber, 2015).

It has now been appreciated that cancer immunotherapies have to be tailored to the microenvironment (Smyth et al., 2016). Here we have established an integrated analysis workflow which provides detailed information about tumor-host interactions. The obtained data about the genomic and immunomic setting will be highly valuable for characterizing patients systematically and identifying therapies they will most likely benefit from.

## CHAPTER 5

---

### Supplemental Data

---

## Supplemental Tables

### Cytokine Data

Tables with raw data of cytokine measurements.

CSV tables of cytokine measurements, formatted according to the different compartments and sample IDs used in this work.

Table with calculates p-values (two-sided student's t-test) for all cytokines, assessing the significance between TIL-high and TIL-low samples.

### Epitope Prediction

Epitope tables for all patients as produced by the immunoinformatics pipeline.

### Other

Selected epitopes of the two patients from the case studies, that have been tested in ELISpot assays.

ID Mapping, Overview of all used sample IDs in this work.

## Supplemental Figures

### Cytokine Plots

Scatterplots for each cytokine, comparing concentrations in TIL-high and TIL-low, separately for each compartment adjacent liver (AL), invasive margin (IM), and liver metastasis (LM).

### Stainings

Images of all performed immunohistochemical stainings of CD3, CD8, CD68, FOXP3, HLA, PD-1, PD-L1, PD-L2

## **ELISpot Data**

Overview of ELISpot data of the two patients from the case studies, including images of ELISpot plates and measured spot counts of each well.

## CHAPTER 6

---

### Technical Appendix

---

## HLA Typing Pipeline

Collection of scripts that are required to run the tool Phlat for sequencing-based HLA typing.

**run\_gunzip.sh:** uncompress the fastq files, as required by Phlat. Required parameters are the locations of the compressed fastq files and the output directory.

**run\_phlat.sh:** run Phlat for sequencing-based HLA typing. Required parameters are the locations of the fastq files and the output directory.

**run\_summary.sh:** format the output of Phlat into a csv table. Required parameter is the output directory, where Phlat output files are located.

**clean\_up.sh:** delete all temporary files. Required parameter is the output directory, where Phlat output files are located.

**run\_hlatyping:** run the whole pipeline combining all scripts above. Parameters for input and output files have to be defined in this file.

The pipeline is run for a number as samples (PIDs) as follows:

```
sh run_hlatyping PID1 PID2 ...
```

## Immunoinformatics Pipeline

Collection of scripts that are required to run the Immunoinformatics Pipeline for rapidly performing HLA class I and class II epitope predictions, proteasomal cleavage predictions, and filtering and prioritizing the predicted epitopes.

### 1. Extract and format missense mutations

As input, a VCF with single nucleotide variations (SNVs) is required. From this VCF, all missense SNVs are extracted and formatted in a CSV table.

**Scripts:** 01-1\_missense\_to\_table.py, 01-2\_format\_missense.R

### 2. Retrieve peptide sequences

For each missense mutation extracted in step 1, all peptide sequences are retrieved, and mutated and wildtype peptides are stored in a CSV table.

**Scripts:** 02-1\_get\_peptides.R, 02-2\_get\_unique\_peptides.R

### 3. Perform epitope predictions

Fasta files that are required as input for netMHCpan and netMHCIIpan are generated for all peptides from step 2 and epitope predictions are performed. Subsequently, the output files are formatted as CSV tables.

**Scripts:** 03\_run\_epitope\_prediction.sh, 03-1\_create\_fasta\_files.R,  
03-2\_create\_netmhc\_commands.R, 03-3\_format\_netmhc\_files.sh

### 4. Create epitope table combining predictions for all alleles

For each patient, all predicted HLA class I and class II epitopes are combined in one table, respectively, and predictions for mutated and wildtype peptides are matched.

**Scripts:** 04-1\_epitope\_table\_classI.R 04-2\_epitope\_table\_classII.R

### 5. Perform proteasomal cleavage predictions

Proteasomal cleavage prediction is performed using NetChop for each predicted HLA class I binder, and the prediction score is added to the epitope table created in step 4.

**Scripts:** 05-1\_run\_netchop.sh, 05-2\_add\_netchop\_to\_table.R

### 6. Filter and prioritize predicted epitopes

Epitopes are filtered and ranked according to their prediction scores, and the final output table with ranked epitopes is generated.

**Scripts:** 06-1\_rank\_epitopes\_classI.R, 06-2\_rank\_epitopes\_classII.R

**run\_immuno\_pipeline.sh** runs the whole pipeline with all the steps above for a number of samples (PIDs). Parameters for input and output files have to be defined in this file.

The command to run the pipeline is as follows:

```
sh run_immuno_pipeline PID1 PID2 ...
```

---

## Bibliography

---

- M. Ahmadzadeh, L. A. Johnson, B. Heemskerk, J. R. Wunderlich, M. E. Dudley, D. E. White, and S. A. Rosenberg. Tumor antigen-specific CD8 T cells infiltrating the tumor express high levels of PD-1 and are functionally impaired. *Blood*, 114(8):1537–1544, Aug 2009. doi: 10.1182/blood-2008-12-195792. URL <http://dx.doi.org/10.1182/blood-2008-12-195792>.
- D. H. Ahn and R. M. Goldberg. Colorectal clinical trials: what is on the horizon? *Future Oncol*, 12(4):525–531, Feb 2016. doi: 10.2217/fon.15.327. URL <http://dx.doi.org/10.2217/fon.15.327>.
- A. C. Anderson. Tim-3, a negative regulator of anti-tumor immunity. *Curr Opin Immunol*, 24(2):213–216, Apr 2012. doi: 10.1016/j.coi.2011.12.005. URL <http://dx.doi.org/10.1016/j.coi.2011.12.005>.
- M. Andreatta, E. Karosiene, M. Rasmussen, A. Stryhn, S. Buus, and M. Nielsen. Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. *Immunogenetics*, Sep 2015. doi: 10.1007/s00251-015-0873-y. URL <http://dx.doi.org/10.1007/s00251-015-0873-y>.
- M. Angelova, P. Charoentong, H. Hackl, M. L. Fischer, R. Snajder, A. M. Krogsdam, M. J. Waldner, G. Bindea, B. Mlecnik, J. Galon, and Z. Trajanoski. Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy. *Genome Biol*, 16:64, 2015. doi: 10.1186/s13059-015-0620-6. URL <http://dx.doi.org/10.1186/s13059-015-0620-6>.
- L. Backert and O. Kohlbacher. Immunoinformatics and epitope prediction in the age of genomic medicine. *Genome Med*, 7(1):119, 2015. doi: 10.1186/s13073-015-0245-0. URL <http://dx.doi.org/10.1186/s13073-015-0245-0>.
- Y. Bai, M. Ni, B. Cooper, Y. Wei, and W. Fury. Inference of high resolution HLA types using genome-wide RNA or DNA sequencing reads. *BMC Genomics*, 15:325, 2014. doi: 10.1186/1471-2164-15-325. URL <http://dx.doi.org/10.1186/1471-2164-15-325>.
- A. B. Bakker, M. W. Schreurs, G. Tafazzul, A. J. de Boer, Y. Kawakami, G. J. Adema, and C. G. Figdor. Identification of a novel peptide derived from the melanocyte-specific gp100 antigen as the dominant epitope recognized by an HLA-A2.1-restricted anti-melanoma CTL line. *Int J Cancer*, 62(1):97–102, Jul 1995.

- A. H. Banday, S. Jeelani, and V. J. Hruby. Cancer vaccine adjuvants—recent clinical progress and future perspectives. *Immunopharmacol Immunotoxicol*, 37(1):1–11, Feb 2015. doi: 10.3109/08923973.2014.971963. URL <http://dx.doi.org/10.3109/08923973.2014.971963>.
- D. L. Barber, E. J. Wherry, D. Masopust, B. Zhu, J. P. Allison, A. H. Sharpe, G. J. Freeman, and R. Ahmed. Restoring function in exhausted CD8 T cells during chronic viral infection. *Nature*, 439(7077):682–687, Feb 2006. doi: 10.1038/nature04444. URL <http://dx.doi.org/10.1038/nature04444>.
- L. Bingle, N. J. Brown, and C. E. Lewis. The role of tumour-associated macrophages in tumour progression: implications for new anticancer therapies. *J Pathol*, 196(3):254–265, Mar 2002. doi: 10.1002/path.1027. URL <http://dx.doi.org/10.1002/path.1027>.
- S. Boegel, M. Löwer, M. Schäfer, T. Bukur, J. de Graaf, V. Boisguérin, O. Türeci, M. Diken, J. C. Castle, and U. Sahin. HLA typing from RNA-Seq sequence reads. *Genome Med*, 4(12):102, 2012. doi: 10.1186/gm403. URL <http://dx.doi.org/10.1186/gm403>.
- A. Bonertz, J. Weitz, D.-H. K. Pietsch, N. N. Rahbari, C. Schlude, Y. Ge, S. Juenger, I. Vlodavsky, K. Khazaie, D. Jaeger, C. Reissfelder, D. Antolovic, M. Aigner, M. Koch, and P. Beckhove. Antigen-specific Tregs control T cell responses against a limited repertoire of tumor antigens in patients with colorectal carcinoma. *J Clin Invest*, 119(11):3311–3321, Nov 2009. doi: 10.1172/JCI39608. URL <http://dx.doi.org/10.1172/JCI39608>.
- T. Boon, P. G. Coulie, B. J. Van den Eynde, and P. van der Bruggen. Human T cell responses against melanoma. *Annu Rev Immunol*, 24:175–208, 2006. doi: 10.1146/annurev.immunol.24.021605.090733. URL <http://dx.doi.org/10.1146/annurev.immunol.24.021605.090733>.
- Y. Bordon. Tumour immunology: Checkpoint parley. *Nat Rev Immunol*, 15(1):5, Jan 2015. doi: 10.1038/nri3792. URL <http://dx.doi.org/10.1038/nri3792>.
- R. Bos and L. A. Sherman. CD4+ T-cell help in the tumor milieu is required for recruitment and cytolytic function of CD8+ T lymphocytes. *Cancer Res*, 70(21):8368–8377, Nov 2010. doi: 10.1158/0008-5472.CAN-10-1322. URL <http://dx.doi.org/10.1158/0008-5472.CAN-10-1322>.
- H. Brenner, M. Kloor, and C. P. Pox. Colorectal cancer. *Lancet*, 383(9927):1490–1502, Apr 2014. doi: 10.1016/S0140-6736(13)61649-9. URL [http://dx.doi.org/10.1016/S0140-6736\(13\)61649-9](http://dx.doi.org/10.1016/S0140-6736(13)61649-9).
- S. D. Brown, R. L. Warren, E. A. Gibb, S. D. Martin, J. J. Spinelli, B. H. Nelson, and R. A. Holt. Neo-antigens predicted by tumor genome meta-analysis correlate with increased patient survival. *Genome Res*, 24(5):743–750, May 2014. doi: 10.1101/gr.165985.113. URL <http://dx.doi.org/10.1101/gr.165985.113>.
- D. H. Busch and E. G. Pamer. MHC class I/peptide stability: implications for immunodominance, in vitro proliferation, and diversity of responding CTL. *J Immunol*, 160(9):4441–4448, May 1998.
- C. M. Cabrera, P. Jimnez, T. Cabrera, C. Esparza, F. Ruiz-Cabello, and F. Garrido. Total loss of MHC class I in colorectal tumors can be explained by two molecular pathways: beta2-microglobulin inactivation in MSI-positive tumors and LMP7/TAP2 downregulation in MSI-negative tumors. *Tissue Antigens*, 61(3):211–219, Mar 2003.
- T. Cabrera, I. Maleno, A. Collado, M. A. Lopez Nevot, B. D. Tait, and F. Garrido. Analysis of HLA class I alterations in tumors: choosing a strategy based on known patterns of underlying molecular mechanisms. *Tissue Antigens*, 69 Suppl 1:264–268, Apr 2007. doi: 10.1111/j.1399-0039.2006.00777.x. URL <http://dx.doi.org/10.1111/j.1399-0039.2006.00777.x>.
- M. J. Cameron and D. J. Kelvin. Cytokines and chemokines—their receptors and their genes: an overview. *Adv Exp Med Biol*, 520:8–32, 2003.

- B. M. Carreno, V. Magrini, M. Becker-Hapak, S. Kaabinejadian, J. Hundal, A. A. Petti, A. Ly, W.-R. Lie, W. H. Hildebrand, E. R. Mardis, and G. P. Linette. Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. *Science*, 348(6236):803–808, May 2015. doi: 10.1126/science.aaa3828. URL <http://dx.doi.org/10.1126/science.aaa3828>.
- J. Chapiro, S. Claverol, F. Piette, W. Ma, V. Stroobant, B. Guillaume, J.-E. Gairin, S. Morel, O. Burlet-Schiltz, B. Monsarrat, T. Boon, and B. J. Van den Eynde. Destructive cleavage of antigenic peptides either by the immunoproteasome or by the standard proteasome results in differential antigen presentation. *J Immunol*, 176(2):1053–1061, Jan 2006.
- S. Chaudhuri, A. Cariappa, M. Tang, D. Bell, D. A. Haber, K. J. Isselbacher, D. Finkelstein, D. Forcione, and S. Pillai. Genetic susceptibility to breast cancer: HLA DQB\*03032 and HLA DRB1\*11 may represent protective alleles. *Proc Natl Acad Sci U S A*, 97(21):11451–11454, Oct 2000. doi: 10.1073/pnas.97.21.11451. URL <http://dx.doi.org/10.1073/pnas.97.21.11451>.
- K. Y. Chung, I. Gore, L. Fong, A. Venook, S. B. Beck, P. Dorazio, P. J. Criscitiello, D. I. Healey, B. Huang, J. Gomez-Navarro, and L. B. Saltz. Phase II study of the anti-cytotoxic T-lymphocyte-associated antigen 4 monoclonal antibody, tremelimumab, in patients with refractory metastatic colorectal cancer. *J Clin Oncol*, 28(21):3485–3490, Jul 2010. doi: 10.1200/JCO.2010.28.3994. URL <http://dx.doi.org/10.1200/JCO.2010.28.3994>.
- W. Clark, Jr, D. E. Elder, D. Guerry, 4th, L. E. Braitman, B. J. Trock, D. Schultz, M. Synnestvedt, and A. C. Halpern. Model predicting survival in stage I melanoma based on tumor progression. *J Natl Cancer Inst*, 81(24):1893–1904, Dec 1989.
- C. G. Clemente, M. Mihm, Jr, R. Bufalino, S. Zurrida, P. Collini, and N. Cascinelli. Prognostic value of tumor infiltrating lymphocytes in the vertical growth phase of primary cutaneous melanoma. *Cancer*, 77(7):1303–1310, Apr 1996. doi: 3.0.CO;2-5. URL <http://dx.doi.org/3.0.CO;2-5>.
- H. I. P. Consortium. URL [www.immuneprofiling.org](http://www.immuneprofiling.org).
- I. H. Consortium. The International HapMap Project. *Nature*, 426(6968):789–796, Dec 2003.
- P. G. Coulie, B. J. Van den Eynde, P. van der Bruggen, and T. Boon. Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. *Nat Rev Cancer*, 14(2):135–146, Feb 2014. doi: 10.1038/nrc3670. URL <http://dx.doi.org/10.1038/nrc3670>.
- A. L. Cox, J. Skipper, Y. Chen, R. A. Henderson, T. L. Darrow, J. Shabanowitz, V. H. Engelhard, D. F. Hunt, and C. Slingsluff, Jr. Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines. *Science*, 264(5159):716–719, Apr 1994.
- D. Cunningham, W. Atkin, H.-J. Lenz, H. T. Lynch, B. Minsky, B. Nordlinger, and N. Starling. Colorectal cancer. *Lancet*, 375(9719):1030–1047, Mar 2010. doi: 10.1016/S0140-6736(10)60353-4. URL [http://dx.doi.org/10.1016/S0140-6736\(10\)60353-4](http://dx.doi.org/10.1016/S0140-6736(10)60353-4).
- L. de la Cruz-Merino, F. Henao Carrasco, D. Vicente Baz, E. Nogales Fernandez, J. J. Reina Zoilo, M. Codes Manuel de Villena, and E. G. Pulido. Immune microenvironment in colorectal cancer: a new hallmark to change old paradigms. *Clin Dev Immunol*, 2011:174149, 2011. doi: 10.1155/2011/174149. URL <http://dx.doi.org/10.1155/2011/174149>.
- L. Delamarre, I. Mellman, and M. Yadav. Cancer immunotherapy. Neo approaches to cancer vaccines. *Science*, 348(6236):760–761, May 2015. doi: 10.1126/science.aab3465. URL <http://dx.doi.org/10.1126/science.aab3465>.
- J. Denoeud and M. Moser. Role of CD27/CD70 pathway of activation in immunity and tolerance. *J Leukoc Biol*, 89(2):195–203, Feb 2011. doi: 10.1189/jlb.0610351. URL <http://dx.doi.org/10.1189/jlb.0610351>.

- H. Dong, S. E. Strome, D. R. Salomao, H. Tamura, F. Hirano, D. B. Flies, P. C. Roche, J. Lu, G. Zhu, K. Tamada, V. A. Lennon, E. Celis, and L. Chen. Tumor-associated B7-H1 promotes T-cell apoptosis: a potential mechanism of immune evasion. *Nat Med*, 8(8):793–800, Aug 2002. doi: 10.1038/nm730. URL <http://dx.doi.org/10.1038/nm730>.
- G. Dranoff. Cytokines in cancer pathogenesis and cancer therapy. *Nat Rev Cancer*, 4(1):11–22, Jan 2004. doi: 10.1038/nrc1252. URL <http://dx.doi.org/10.1038/nrc1252>.
- F. Duan, J. Duitama, S. Al Seesi, C. M. Ayres, S. A. Corcelli, A. P. Pawashe, T. Blanchard, D. McMahon, J. Sidney, A. Sette, B. M. Baker, I. I. Mandoiu, and P. K. Srivastava. Genomic and bioinformatic profiling of mutational neoepitopes reveals new rules to predict anticancer immunogenicity. *J Exp Med*, 211(11):2231–2248, Oct 2014. doi: 10.1084/jem.20141308. URL <http://dx.doi.org/10.1084/jem.20141308>.
- G. P. Dunn, A. T. Bruce, H. Ikeda, L. J. Old, and R. D. Schreiber. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat Immunol*, 3(11):991–998, Nov 2002. doi: 10.1038/ni1102-991. URL <http://dx.doi.org/10.1038/ni1102-991>.
- S. Durinck, Y. Moreau, A. Kasprzyk, S. Davis, B. De Moor, A. Brazma, and W. Huber. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*, 21(16):3439–3440, Aug 2005. doi: 10.1093/bioinformatics/bti525. URL <http://dx.doi.org/10.1093/bioinformatics/bti525>.
- S. Durinck, P. T. Spellman, E. Birney, and W. Huber. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*, 4(8):1184–1191, 2009. doi: 10.1038/nprot.2009.97. URL <http://dx.doi.org/10.1038/nprot.2009.97>.
- E. R. Fearon. Molecular genetics of colorectal cancer. *Annu Rev Pathol*, 6:479–507, 2011. doi: 10.1146/annurev-pathol-011110-130235. URL <http://dx.doi.org/10.1146/annurev-pathol-011110-130235>.
- D. Fiszer and M. Kurpisz. Major histocompatibility complex expression on human, male germ cells: a review. *Am J Reprod Immunol*, 40(3):172–176, Sep 1998.
- D. B. Flies, B. J. Sandler, M. Sznol, and L. Chen. Blockade of the B7-H1/PD-1 pathway for cancer immunotherapy. *Yale J Biol Med*, 84(4):409–421, Dec 2011.
- M.-H. Fortier, E. Caron, M.-P. Hardy, G. Voisin, S. Lemieux, C. Perreault, and P. Thibault. The MHC class I peptide repertoire is molded by the transcriptome. *J Exp Med*, 205(3):595–610, Mar 2008. doi: 10.1084/jem.20071985. URL <http://dx.doi.org/10.1084/jem.20071985>.
- J. Fourcade, Z. Sun, M. Benallaoua, P. Guillaume, I. F. Luescher, C. Sander, J. M. Kirkwood, V. Kuchroo, and H. M. Zarour. Upregulation of Tim-3 and PD-1 expression is associated with tumor antigen-specific CD8+ T cell dysfunction in melanoma patients. *J Exp Med*, 207(10):2175–2186, Sep 2010. doi: 10.1084/jem.20100637. URL <http://dx.doi.org/10.1084/jem.20100637>.
- G. J. Freeman, J. G. Gribben, V. A. Boussiotis, J. W. Ng, V. Restivo, Jr, L. A. Lombard, G. S. Gray, and L. M. Nadler. Cloning of B7-2: a CTLA-4 counter-receptor that costimulates human T cell proliferation. *Science*, 262(5135):909–911, Nov 1993.
- W. H. Fridman, F. Pags, C. Sauts-Fridman, and J. Galon. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer*, 12(4):298–306, Apr 2012. doi: 10.1038/nrc3245. URL <http://dx.doi.org/10.1038/nrc3245>.
- E. F. Fritsch, M. Rajasagi, P. A. Ott, V. Brusic, N. Hacohen, and C. J. Wu. HLA-binding properties of tumor neoepitopes in humans. *Cancer Immunol Res*, 2(6):522–529, Jun 2014. doi: 10.1158/2326-6066.CIR-13-0227. URL <http://dx.doi.org/10.1158/2326-6066.CIR-13-0227>.

- D. Fruci, P. Rovero, G. Falasca, A. Chersi, R. Sorrentino, R. Butler, N. Tanigaki, and R. Tosi. Anchor residue motifs of HLA class-I-binding peptides analyzed by the direct binding of synthetic peptides to HLA class I alpha chains. *Hum Immunol*, 38(3):187–192, Nov 1993.
- Y. Fujita, C. M. Rooney, and H. E. Heslop. Adoptive cellular immunotherapy for viral diseases. *Bone Marrow Transplant*, 41(2):193–198, Jan 2008. doi: 10.1038/sj.bmt.1705906. URL <http://dx.doi.org/10.1038/sj.bmt.1705906>.
- E. M. Gabriel and E. C. Lattime. Anti-CTL-associated antigen 4: are regulatory T cells a target? *Clin Cancer Res*, 13(3):785–788, Feb 2007. doi: 10.1158/1078-0432.CCR-06-2820. URL <http://dx.doi.org/10.1158/1078-0432.CCR-06-2820>.
- J. Galon, A. Costes, F. Sanchez-Cabo, A. Kirilovsky, B. Mlecnik, C. Lagorce-Pags, M. Tosolini, M. Camus, A. Berger, P. Wind, F. Zinzindohou, P. Bruneval, P.-H. Cugnenc, Z. Trajanoski, W.-H. Fridman, and F. Pags. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*, 313(5795):1960–1964, Sep 2006. doi: 10.1126/science.1129139. URL <http://dx.doi.org/10.1126/science.1129139>.
- X. Gao, Y. Zhu, G. Li, H. Huang, G. Zhang, F. Wang, J. Sun, Q. Yang, X. Zhang, and B. Lu. TIM-3 expression characterizes regulatory T cells in tumor tissues and is associated with lung cancer progression. *PLoS One*, 7(2):e30676, 2012. doi: 10.1371/journal.pone.0030676. URL <http://dx.doi.org/10.1371/journal.pone.0030676>.
- C. Garca-Corona, E. Vega-Memije, A. Mosqueda-Taylor, J. K. Yamamoto-Furusho, A. A. Rodriguez-Carren, J. A. Ruiz-Morales, N. Salgado, and J. Granados. Association of HLA-DR4 (DRB1\*0404) with human papillomavirus infection in patients with focal epithelial hyperplasia. *Arch Dermatol*, 140(10):1227–1231, Oct 2004. doi: 10.1001/archderm.140.10.1227. URL <http://dx.doi.org/10.1001/archderm.140.10.1227>.
- R. Gentleman. *annotate: Annotation for microarrays*, 2015. R package version 1.40.1.
- F. F. Gonzalez-Galarza, S. Christmas, D. Middleton, and A. R. Jones. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res*, 39(Database issue):D913–D919, Jan 2011. doi: 10.1093/nar/gkq1128. URL <http://dx.doi.org/10.1093/nar/gkq1128>.
- D. P. Granados, W. Yahyaoui, C. M. Laumont, T. Daouda, T. L. Muratore-Schroeder, C. Ct, J.-P. Laverdure, S. Lemieux, P. Thibault, and C. Perreault. MHC I-associated peptides preferentially derive from transcripts bearing miRNA response elements. *Blood*, 119(26):e181–e191, Jun 2012. doi: 10.1182/blood-2012-02-412593. URL <http://dx.doi.org/10.1182/blood-2012-02-412593>.
- M. M. Gubin, X. Zhang, H. Schuster, E. Caron, J. P. Ward, T. Noguchi, Y. Ivanova, J. Hundal, C. D. Arthur, W.-J. Krebber, G. E. Mulder, M. Toebes, M. D. Vesely, S. S. K. Lam, A. J. Korman, J. P. Allison, G. J. Freeman, A. H. Sharpe, E. L. Pearce, T. N. Schumacher, R. Aebersold, H.-G. Rammensee, C. J. M. Melief, E. R. Mardis, W. E. Gillanders, M. N. Artyomov, and R. D. Schreiber. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature*, 515(7528):577–581, Nov 2014. doi: 10.1038/nature13988. URL <http://dx.doi.org/10.1038/nature13988>.
- B. Guillaume, V. Stroobant, M.-P. Bousquet-Dubouch, D. Colau, J. Chapiro, N. Parmentier, A. Dalet, and B. J. Van den Eynde. Analysis of the processing of seven human tumor antigens by intermediate proteasomes. *J Immunol*, 189(7):3538–3547, Oct 2012. doi: 10.4049/jimmunol.1103213. URL <http://dx.doi.org/10.4049/jimmunol.1103213>.
- J. B. A. G. Haanen, A. Baars, R. Gomez, P. Weder, M. Smits, T. D. de Gruijl, B. M. E. von Blomberg, E. Bloemena, R. J. Scheper, S. M. van Ham, H. M. Pinedo, and A. J. M. van den Eertwegh. Melanoma-specific tumor-infiltrating lymphocytes but not circulating melanoma-specific T cells

- may predict survival in resected advanced-stage melanoma patients. *Cancer Immunol Immunother*, 55(4):451–458, Apr 2006. doi: 10.1007/s00262-005-0018-5. URL <http://dx.doi.org/10.1007/s00262-005-0018-5>.
- N. Hacohen, E. F. Fritsch, T. A. Carter, E. S. Lander, and C. J. Wu. Getting personal with neoantigen-based therapeutic cancer vaccines. *Cancer Immunol Res*, 1(1):11–15, Jul 2013. doi: 10.1158/2326-6066.CIR-13-0022. URL <http://dx.doi.org/10.1158/2326-6066.CIR-13-0022>.
- Y. Hailemichael and W. W. Overwijk. Cancer vaccines: Trafficking of tumor-specific T cells to tumor after therapeutic vaccination. *Int J Biochem Cell Biol*, 53:46–50, Aug 2014. doi: 10.1016/j.biocel.2014.04.019. URL <http://dx.doi.org/10.1016/j.biocel.2014.04.019>.
- Y. Hailemichael, Z. Dai, N. Jaffarad, Y. Ye, M. A. Medina, X.-F. Huang, S. M. Dorta-Estremera, N. R. Greeley, G. Nitti, W. Peng, C. Liu, Y. Lou, Z. Wang, W. Ma, B. Rabinovich, R. T. Sowell, K. S. Schluns, R. E. Davis, P. Hwu, and W. W. Overwijk. Persistent antigen at vaccination sites induces tumor-specific CD8 T cell sequestration, dysfunction and deletion. *Nat Med*, 19(4):465–472, Apr 2013. doi: 10.1038/nm.3105. URL <http://dx.doi.org/10.1038/nm.3105>.
- N. Halama, S. Michel, M. Kloor, I. Zoernig, T. Pommerencke, M. von Knebel Doeberitz, P. Schirmacher, J. Weitz, N. Grabe, and D. Jger. The localization and density of immune cells in primary tumors of human metastatic colorectal cancer shows an association with response to chemotherapy. *Cancer Immunol*, 9:1, 2009a.
- N. Halama, I. Zoernig, A. Spille, K. Westphal, P. Schirmacher, D. Jaeger, and N. Grabe. Estimation of immune cell densities in immune cell conglomerates: an approach for high-throughput quantification. *PLoS One*, 4(11):e7847, 2009b. doi: 10.1371/journal.pone.0007847. URL <http://dx.doi.org/10.1371/journal.pone.0007847>.
- N. Halama, I. Zoernig, A. Spille, S. Michel, M. Kloor, S. Grauling-Halama, K. Westphal, P. Schirmacher, D. Jger, and N. Grabe. Quantification of prognostic immune cell markers in colorectal cancer using whole slide imaging tumor maps. *Anal Quant Cytol Histol*, 32(6):333–340, Dec 2010.
- N. Halama, M. Braun, C. Kahlert, A. Spille, C. Quack, N. Rahbari, M. Koch, J. Weitz, M. Kloor, I. Zoernig, P. Schirmacher, K. Brand, N. Grabe, and C. S. Falk. Natural killer cells are scarce in colorectal carcinoma tissue despite high levels of chemokines and cytokines. *Clin Cancer Res*, 17(4):678–689, Feb 2011a. doi: 10.1158/1078-0432.CCR-10-2173. URL <http://dx.doi.org/10.1158/1078-0432.CCR-10-2173>.
- N. Halama, S. Michel, M. Kloor, I. Zoernig, A. Benner, A. Spille, T. Pommerencke, D. M. von Knebel, G. Folprecht, B. Lubber, N. Feyen, U. M. Martens, P. Beckhove, S. Gnjatic, P. Schirmacher, E. Herpel, J. Weitz, N. Grabe, and D. Jaeger. Localization and density of immune cells in the invasive margin of human colorectal cancer liver metastases are prognostic for response to chemotherapy. *Cancer Res*, 71(17):5670–5677, Sep 2011b. doi: 10.1158/0008-5472.CAN-11-0268. URL <http://dx.doi.org/10.1158/0008-5472.CAN-11-0268>.
- N. Halama, I. Zoernig, N. Grabe, and D. Jaeger. The local immunological microenvironment in colorectal cancer as a prognostic factor for treatment decisions in the clinic: The way ahead. *Oncoimmunology*, 1(1):62–66, Jan 2012. doi: 10.4161/onci.1.1.18460. URL <http://dx.doi.org/10.4161/onci.1.1.18460>.
- K.-I. Hanada, J. W. Yewdell, and J. C. Yang. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature*, 427(6971):252–256, Jan 2004. doi: 10.1038/nature02240. URL <http://dx.doi.org/10.1038/nature02240>.
- D. Hanahan and R. A. Weinberg. The hallmarks of cancer. *Cell*, 100(1):57–70, Jan 2000.

- D. Hanahan and R. A. Weinberg. Hallmarks of cancer: the next generation. *Cell*, 144(5):646–674, Mar 2011. doi: 10.1016/j.cell.2011.02.013. URL <http://dx.doi.org/10.1016/j.cell.2011.02.013>.
- K. S. Hathcock, G. Laszlo, H. B. Dickler, J. Bradshaw, P. Linsley, and R. J. Hodes. Identification of an alternative CTLA-4 ligand costimulatory for T cell activation. *Science*, 262(5135):905–907, Nov 1993.
- B. Heemskerk, P. Kvistborg, and T. N. M. Schumacher. The cancer antigenome. *EMBO J*, 32(2):194–203, Jan 2013. doi: 10.1038/emboj.2012.333. URL <http://dx.doi.org/10.1038/emboj.2012.333>.
- J. Hendriks, Y. Xiao, and J. Borst. CD27 promotes survival of activated T cells and complements CD28 in generation and establishment of the effector T cell pool. *J Exp Med*, 198(9):1369–1380, Nov 2003. doi: 10.1084/jem.20030916. URL <http://dx.doi.org/10.1084/jem.20030916>.
- R. Q. Hintzen, S. M. Lens, M. P. Beckmann, R. G. Goodwin, D. Lynch, and R. A. van Lier. Characterization of the human CD27 ligand, a novel member of the TNF gene family. *J Immunol*, 152(4):1762–1773, Feb 1994.
- R. Q. Hintzen, S. M. Lens, K. Lammers, H. Kuiper, M. P. Beckmann, and R. A. van Lier. Engagement of CD27 with its ligand CD70 provides a second signal for T cell activation. *J Immunol*, 154(6):2612–2623, Mar 1995.
- F. S. Hodi, S. J. O’Day, D. F. McDermott, R. W. Weber, J. A. Sosman, J. B. Haanen, R. Gonzalez, C. Robert, D. Schadendorf, J. C. Hassel, W. Akerley, A. J. M. van den Eertwegh, J. Lutzky, P. Lorigan, J. M. Vaubel, G. P. Linette, D. Hogg, C. H. Ottensmeier, C. Lebb, C. Peschel, I. Quirt, J. I. Clark, J. D. Wolchok, J. S. Weber, J. Tian, M. J. Yellin, G. M. Nichol, A. Hoos, and W. J. Urba. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med*, 363(8):711–723, Aug 2010. doi: 10.1056/NEJMoa1003466. URL <http://dx.doi.org/10.1056/NEJMoa1003466>.
- I. Hoof, B. Peters, J. Sidney, L. E. Pedersen, A. Sette, O. Lund, S. Buus, and M. Nielsen. NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics*, 61(1):1–13, Jan 2009. doi: 10.1007/s00251-008-0341-z. URL <http://dx.doi.org/10.1007/s00251-008-0341-z>.
- T. Horn, J. Grab, J. Schusdziarra, S. Schmid, T. Maurer, R. Nawroth, P. Wolf, M. Pritsch, J. E. Gschwend, H. R. Kbler, and P. Beckhove. Antitumor T cell responses in bladder cancer are directed against a limited set of antigens and are modulated by regulatory T cells and routine treatment approaches. *Int J Cancer*, 133(9):2145–2156, Nov 2013. doi: 10.1002/ijc.28233. URL <http://dx.doi.org/10.1002/ijc.28233>.
- Y. Ishida, Y. Agata, K. Shibahara, and T. Honjo. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *EMBO J*, 11(11):3887–3895, Nov 1992.
- J. R. Jass, S. B. Love, and J. M. Northover. A new prognostic classification of rectal cancer. *Lancet*, 1(8545):1303–1306, Jun 1987.
- D. T. W. Jones, N. Jger, M. Kool, T. Zichner, B. Hutter, M. Sultan, Y.-J. Cho, T. J. Pugh, V. Hovestadt, A. M. Sttz, T. Rausch, H.-J. Warnatz, M. Ryzhova, S. Bender, D. Sturm, S. Pleier, H. Cin, E. Pfaff, L. Sieber, A. Wittmann, M. Remke, H. Witt, S. Hutter, T. Tzaridis, J. Weischenfeldt, B. Raeder, M. Avci, V. Amstislavskiy, M. Zapatka, U. D. Weber, Q. Wang, B. Lasitschka, C. C. Bartholomae, M. Schmidt, C. von Kalle, V. Ast, C. Lawerenz, J. Eils, R. Kabbe, V. Benes, P. van Sluis, J. Koster, R. Volckmann, D. Shih, M. J. Betts, R. B. Russell, S. Coco, G. P. Tonini, U. Schller, V. Hans, N. Graf, Y.-J. Kim, C. Monoranu, W. Roggendorf, A. Unterberg, C. Herold-Mende, T. Milde, A. E. Kulozik, A. von Deimling, O. Witt, E. Maass, J. Rssler, M. Ebinger, M. U. Schuhmann, M. C. Frhwald, M. Hasselblatt, N. Jabado, S. Rutkowski, A. O. von Bueren,

- D. Williamson, S. C. Clifford, M. G. McCabe, V. P. Collins, S. Wolf, S. Wiemann, H. Lehrach, B. Brors, W. Scheurlen, J. Felsberg, G. Reifenberger, P. A. Northcott, M. D. Taylor, M. Meyerson, S. L. Pomeroy, M.-L. Yaspo, J. O. Korbel, A. Korshunov, R. Eils, S. M. Pfister, and P. Lichter. Dissecting the genomic complexity underlying medulloblastoma. *Nature*, 488(7409):100–105, Aug 2012. doi: 10.1038/nature11284. URL <http://dx.doi.org/10.1038/nature11284>.
- D. T. W. Jones, B. Hutter, N. Jäger, A. Korshunov, M. Kool, H.-J. Warnatz, T. Zichner, S. R. Lambert, M. Ryzhova, D. A. K. Quang, A. M. Fontebasso, A. M. Stütz, S. Hutter, M. Zuckermann, D. Sturm, J. Gronych, B. Lasitschka, S. Schmidt, H. Seker-Cin, H. Witt, M. Sultan, M. Ralser, P. A. Northcott, V. Hovestadt, S. Bender, E. Pfaff, S. Stark, D. Faury, J. Schwartzentruber, J. Majewski, U. D. Weber, M. Zapatka, B. Raeder, M. Schlesner, C. L. Worth, C. C. Bartholomae, C. von Kalle, C. D. Imbusch, S. Radomski, C. Lawerenz, P. van Sluis, J. Koster, R. Volckmann, R. Versteeg, H. Lehrach, C. Monoranu, B. Winkler, A. Unterberg, C. Herold-Mende, T. Milde, A. E. Kulozik, M. Ebinger, M. U. Schuhmann, Y.-J. Cho, S. L. Pomeroy, A. von Deimling, O. Witt, M. D. Taylor, S. Wolf, M. A. Karajannis, C. G. Eberhart, W. Scheurlen, M. Hasselblatt, K. L. Ligon, M. W. Kieran, J. O. Korbel, M.-L. Yaspo, B. Brors, J. Felsberg, G. Reifenberger, V. P. Collins, N. Jabado, R. Eils, P. Lichter, S. M. Pfister, and I. C. G. C. P. T. P. . Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nat Genet*, 45(8):927–932, Aug 2013. doi: 10.1038/ng.2682. URL <http://dx.doi.org/10.1038/ng.2682>.
- J. A. Joyce and J. W. Pollard. Microenvironmental regulation of metastasis. *Nat Rev Cancer*, 9(4): 239–252, Apr 2009. doi: 10.1038/nrc2618. URL <http://dx.doi.org/10.1038/nrc2618>.
- E. Karosiene, M. Rasmussen, T. Blicher, O. Lund, S. Buus, and M. Nielsen. NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. *Immunogenetics*, 65(10):711–724, Oct 2013. doi: 10.1007/s00251-013-0720-y. URL <http://dx.doi.org/10.1007/s00251-013-0720-y>.
- Y. Kawakami, S. Eliyahu, C. H. Delgado, P. F. Robbins, K. Sakaguchi, E. Appella, J. R. Yannelli, G. J. Adema, T. Miki, and S. A. Rosenberg. Identification of a human melanoma antigen recognized by tumor-infiltrating lymphocytes associated with in vivo tumor rejection. *Proc Natl Acad Sci U S A*, 91(14):6458–6462, Jul 1994.
- C. Kemir, A. K. Nussbaum, H. Schild, V. Detours, and S. Brunak. Prediction of proteasome cleavage motifs by neural networks. *Protein Eng*, 15(4):287–296, Apr 2002.
- J. H. Kessler, S. Khan, U. Seifert, S. Le Gall, K. M. Chow, A. Paschen, S. A. Bres-Vloemans, A. de Ru, N. van Montfoort, K. L. M. C. Franken, W. E. Benckhuijsen, J. M. Brooks, T. van Hall, K. Ray, A. Mulder, I. I. N. Doxiadis, P. F. van Swieten, H. S. Overkleeft, A. Prat, B. Tomkinson, J. Neefjes, P. M. Kloetzel, D. W. Rodgers, L. B. Hersch, J. W. Drijfhout, P. A. van Veelen, F. Ossendorp, and C. J. M. Melief. Antigen processing by nardilysin and thimet oligopeptidase generates cytotoxic T cell epitopes. *Nat Immunol*, 12(1):45–53, Jan 2011. doi: 10.1038/ni.1974. URL <http://dx.doi.org/10.1038/ni.1974>.
- J. S. Khalili, R. W. Hanson, and Z. Szallasi. In silico prediction of tumor antigens derived from functional missense mutations of the cancer gene census. *Oncoimmunology*, 1(8):1281–1289, Nov 2012. doi: 10.4161/onci.21511. URL <http://dx.doi.org/10.4161/onci.21511>.
- H. J. Kim and N. Pourmand. HLA typing from RNA-seq data using hierarchical read weighting [corrected]. *PLoS One*, 8(6):e67885, 2013. doi: 10.1371/journal.pone.0067885. URL <http://dx.doi.org/10.1371/journal.pone.0067885>.
- K. Ko, S. Yamazaki, K. Nakamura, T. Nishioka, K. Hirota, T. Yamaguchi, J. Shimizu, T. Nomura, T. Chiba, and S. Sakaguchi. Treatment of advanced tumors with agonistic anti-GITR mAb and its effects on tumor-infiltrating Foxp3+CD25+CD4+ regulatory T cells. *J Exp Med*, 202(7):885–891, Oct 2005. doi: 10.1084/jem.20050940. URL <http://dx.doi.org/10.1084/jem.20050940>.

- K. S. Kobayashi and P. J. van den Elsen. NLRC5: a key regulator of MHC class I-dependent immune responses. *Nat Rev Immunol*, 12(12):813–820, Dec 2012. doi: 10.1038/nri3339. URL <http://dx.doi.org/10.1038/nri3339>.
- D. C. Koboldt, Q. Zhang, D. E. Larson, D. Shen, M. D. McLellan, L. Lin, C. A. Miller, E. R. Mardis, L. Ding, and R. K. Wilson. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*, 22(3):568–576, Mar 2012. doi: 10.1101/gr.129684.111. URL <http://dx.doi.org/10.1101/gr.129684.111>.
- J. N. Kochenderfer, Z. Yu, D. Frasheri, N. P. Restifo, and S. A. Rosenberg. Adoptive transfer of syngeneic T cells transduced with a chimeric antigen receptor that recognizes murine CD19 can eradicate lymphoma and normal B cells. *Blood*, 116(19):3875–3886, Nov 2010. doi: 10.1182/blood-2010-01-265041. URL <http://dx.doi.org/10.1182/blood-2010-01-265041>.
- E. Kondo, K. Koda, N. Takiguchi, K. Oda, K. Seike, M. Ishizuka, and M. Miyazaki. Preoperative natural killer cell activity as a prognostic factor for distant metastasis following surgery for colon cancer. *Dig Surg*, 20(5):445–451, 2003. doi: 72714. URL <http://dx.doi.org/72714>.
- T. Kouo, L. Huang, A. B. Pucsek, M. Cao, S. Solt, T. Armstrong, and E. Jaffee. Galectin-3 Shapes Antitumor Immune Responses by Suppressing CD8+ T Cells via LAG-3 and Inhibiting Expansion of Plasmacytoid Dendritic Cells. *Cancer Immunol Res*, 3(4):412–423, Apr 2015. doi: 10.1158/2326-6066.CIR-14-0150. URL <http://dx.doi.org/10.1158/2326-6066.CIR-14-0150>.
- S. Kreiter, M. Vormehr, N. van de Roemer, M. Diken, M. Lwer, J. Diekmann, S. Boegel, B. Schrrs, F. Vascotto, J. C. Castle, A. D. Tadmor, S. P. Schoenberger, C. Huber, z. Treci, and U. Sahin. Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature*, 520(7549):692–696, Apr 2015. doi: 10.1038/nature14426. URL <http://dx.doi.org/10.1038/nature14426>.
- M. F. Krummel and J. P. Allison. CD28 and CTLA-4 have opposing effects on the response of T cells to stimulation. *J Exp Med*, 182(2):459–465, Aug 1995.
- P. Kvistborg, D. Philips, S. Kelderman, L. Hageman, C. Ottensmeier, D. Joseph-Pietras, M. J. P. Welters, S. van der Burg, E. Kapiteijn, O. Michielin, E. Romano, C. Linnemann, D. Speiser, C. Blank, J. B. Haanen, and T. N. Schumacher. Anti-CTLA-4 therapy broadens the melanoma-reactive CD8+ T cell response. *Sci Transl Med*, 6(254):254ra128, Sep 2014. doi: 10.1126/scitranslmed.3008918. URL <http://dx.doi.org/10.1126/scitranslmed.3008918>.
- B. Langmead, C. Trapnell, M. Pop, and S. L. Salzberg. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*, 10(3):R25, 2009. doi: 10.1186/gb-2009-10-3-r25. URL <http://dx.doi.org/10.1186/gb-2009-10-3-r25>.
- D. T. Le, J. N. Uram, H. Wang, B. R. Bartlett, H. Kemberling, A. D. Eyring, A. D. Skora, B. S. Lubner, N. S. Azad, D. Laheru, B. Biedrzycki, R. C. Donehower, A. Zaheer, G. A. Fisher, T. S. Crocenzi, J. J. Lee, S. M. Duffy, R. M. Goldberg, A. de la Chapelle, M. Koshiji, F. Bhajee, T. Huebner, R. H. Hruban, L. D. Wood, N. Cuka, D. M. Pardoll, N. Papadopoulos, K. W. Kinzler, S. Zhou, T. C. Cornish, J. M. Taube, R. A. Anders, J. R. Eshleman, B. Vogelstein, and L. A. Diaz, Jr. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med*, 372(26):2509–2520, Jun 2015. doi: 10.1056/NEJMoa1500596. URL <http://dx.doi.org/10.1056/NEJMoa1500596>.
- W. J. Lesterhuis, I. J. M. de Vries, D. H. Schuurhuis, A. C. I. Boullart, J. F. M. Jacobs, A. J. de Boer, N. M. Scharenborg, H. M. H. Brouwer, M. W. M. M. van de Rakt, C. G. Figdor, T. J. Ruers, G. J. Adema, and C. J. A. Punt. Vaccination of colorectal cancer patients with CEA-loaded dendritic cells: antigen-specific T cell responses in DTH skin tests. *Ann Oncol*, 17(6):974–980, Jun 2006. doi: 10.1093/annonc/mdl072. URL <http://dx.doi.org/10.1093/annonc/mdl072>.
- H. Li and R. Durbin. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5):589–595, Mar 2010. doi: 10.1093/bioinformatics/btp698. URL <http://dx.doi.org/10.1093/bioinformatics/btp698>.

- H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and . G. P. D. P. S. . The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16):2078–2079, Aug 2009.
- C. Linnemann, M. M. van Buuren, L. Bies, E. M. E. Verdegaal, R. Schotte, J. J. A. Calis, S. Behjati, A. Velds, H. Hilkmann, D. E. Atmioui, M. Visser, M. R. Stratton, J. B. A. G. Haanen, H. Spits, S. H. van der Burg, and T. N. M. Schumacher. High-throughput epitope discovery reveals frequent recognition of neo-antigens by CD4+ T cells in human melanoma. *Nat Med*, 21(1):81–85, Jan 2015. doi: 10.1038/nm.3773. URL <http://dx.doi.org/10.1038/nm.3773>.
- P. S. Linsley, J. L. Greene, W. Brady, J. Bajorath, J. A. Ledbetter, and R. Peach. Human B7-1 (CD80) and B7-2 (CD86) bind with similar avidities but distinct kinetics to CD28 and CTLA-4 receptors. *Immunity*, 1(9):793–801, Dec 1994.
- E. J. Lipson, W. H. Sharfman, S. Chen, T. L. McMiller, T. S. Pritchard, J. T. Salas, S. Sartorius-Mergenthaler, I. Freed, S. Ravi, H. Wang, B. Lubner, J. D. Sproul, J. M. Taube, D. M. Pardoll, and S. L. Topalian. Safety and immunologic correlates of Melanoma GVAX, a GM-CSF secreting allogeneic melanoma cell vaccine administered in the adjuvant setting. *J Transl Med*, 13:214, 2015. doi: 10.1186/s12967-015-0572-3. URL <http://dx.doi.org/10.1186/s12967-015-0572-3>.
- H. M. Long, G. Parsonage, C. P. Fox, and S. P. Lee. Immunotherapy for Epstein-Barr virus-associated malignancies. *Drug News Perspect*, 23(4):221–228, May 2010. doi: 10.1358/dnp.2010.23.4.1439500. URL <http://dx.doi.org/10.1358/dnp.2010.23.4.1439500>.
- Y.-C. Lu and P. F. Robbins. Cancer immunotherapy targeting neoantigens. *Semin Immunol*, Nov 2015. doi: 10.1016/j.smim.2015.11.002. URL <http://dx.doi.org/10.1016/j.smim.2015.11.002>.
- K. M. Mahoney, P. D. Rennert, and G. J. Freeman. Combination cancer immunotherapy and new immunomodulatory targets. *Nat Rev Drug Discov*, 14(8):561–584, Aug 2015. doi: 10.1038/nrd4591. URL <http://dx.doi.org/10.1038/nrd4591>.
- J. L. Markman and S. L. Shiao. Impact of the immune system and immunotherapy in colorectal cancer. *J Gastrointest Oncol*, 6(2):208–223, Apr 2015. doi: 10.3978/j.issn.2078-6891.2014.077. URL <http://dx.doi.org/10.3978/j.issn.2078-6891.2014.077>.
- L. G. Martelotto, C. K. Ng, M. R. De Filippo, Y. Zhang, S. Piscuoglio, R. S. Lim, R. Shen, L. Norton, J. S. Reis-Filho, and B. Weigelt. Benchmarking mutation effect prediction algorithms using functionally validated cancer-related missense mutations. *Genome Biol*, 15(10):484, 2014. doi: 10.1186/s13059-014-0484-1. URL <http://dx.doi.org/10.1186/s13059-014-0484-1>.
- H. Matsushita, M. D. Vesely, D. C. Koboldt, C. G. Rickert, R. Uppaluri, V. J. Magrini, C. D. Arthur, J. M. White, Y.-S. Chen, L. K. Shea, J. Hundal, M. C. Wendl, R. Demeter, T. Wylie, J. P. Allison, M. J. Smyth, L. J. Old, E. R. Mardis, and R. D. Schreiber. Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature*, 482(7385):400–404, Feb 2012. doi: 10.1038/nature10755. URL <http://dx.doi.org/10.1038/nature10755>.
- I. Melero, W. W. Shuford, S. A. Newby, A. Aruffo, J. A. Ledbetter, K. E. Hellstrm, R. S. Mittler, and L. Chen. Monoclonal antibodies against the 4-1BB T-cell activation molecule eradicate established tumors. *Nat Med*, 3(6):682–685, Jun 1997.
- I. Melero, D. M. Berman, M. A. Aznar, A. J. Korman, J. L. Prez Gracia, and J. Haanen. Evolving synergistic combinations of targeted immunotherapies to combat cancer. *Nat Rev Cancer*, 15(8):457–472, Aug 2015. doi: 10.1038/nrc3973. URL <http://dx.doi.org/10.1038/nrc3973>.
- M. Meyerson, S. Gabriel, and G. Getz. Advances in understanding cancer genomes through second-generation sequencing. *Nat Rev Genet*, 11(10):685–696, Oct 2010. doi: 10.1038/nrg2841. URL <http://dx.doi.org/10.1038/nrg2841>.

- B. Mlecnik, M. Tosolini, P. Charoentong, A. Kirilovsky, G. Bindea, A. Berger, M. Camus, M. Gillard, P. Bruneval, W.-H. Fridman, F. Pags, Z. Trajanoski, and J. Galon. Biomolecular network reconstruction identifies T-cell homing factors associated with survival in colorectal cancer. *Gastroenterology*, 138(4):1429–1440, Apr 2010. doi: 10.1053/j.gastro.2009.10.057. URL <http://dx.doi.org/10.1053/j.gastro.2009.10.057>.
- B. Mlecnik, G. Bindea, A. Kirilovsky, H. K. Angell, A. C. Obenauf, M. Tosolini, S. E. Church, P. Maby, A. Vasaturo, M. Angelova, T. Fredriksen, S. Mauger, M. Waldner, A. Berger, M. R. Speicher, F. Pags, V. Valge-Archer, and J. Galon. The tumor microenvironment and Immunoscore are critical determinants of dissemination to distant metastasis. *Sci Transl Med*, 8(327):327ra26, Feb 2016. doi: 10.1126/scitranslmed.aad6352. URL <http://dx.doi.org/10.1126/scitranslmed.aad6352>.
- A. Moll, A. Hildebrandt, H.-P. Lenhof, and O. Kohlbacher. BALLView: a tool for research and education in molecular modeling. *Bioinformatics*, 22(3):365–366, Feb 2006. doi: 10.1093/bioinformatics/bti818. URL <http://dx.doi.org/10.1093/bioinformatics/bti818>.
- S. Morel, F. Lvy, O. Burlet-Schiltz, F. Brasseur, M. Probst-Kepper, A. L. Peitrequin, B. Monsarrat, R. Van Velthoven, J. C. Cerottini, T. Boon, J. E. Gairin, and B. J. Van den Eynde. Processing of some antigens by the standard proteasome but not by the immunoproteasome results in poor presentation by dendritic cells. *Immunity*, 12(1):107–117, Jan 2000.
- M. A. Morse, D. Niedzwiecki, J. L. Marshall, C. Garrett, D. Z. Chang, M. Aklilu, T. S. Crocenzi, D. J. Cole, S. Dessureault, A. C. Hobeika, T. Osada, M. Onaitis, B. M. Clary, D. Hsu, G. R. Devi, A. Bulusu, R. P. Annechiarico, V. Chadaram, T. M. Clay, and H. K. Lyerly. A randomized phase II study of immunization with dendritic cells modified with poxvectors encoding CEA and MUC1 compared with the same poxvectors plus GM-CSF for resected metastatic colorectal cancer. *Ann Surg*, 258(6):879–886, Dec 2013. doi: 10.1097/SLA.0b013e318292919e. URL <http://dx.doi.org/10.1097/SLA.0b013e318292919e>.
- K. Murphy, P. Travers, M. Walport, and C. Janeway. *Janeway’s Immunobiology*. Number Bd. 978,Nr. 0-4129 in Janeway’s Immunobiology. Garland Science, 2008. ISBN 9780815341239. URL <https://books.google.de/books?id=-SNrAAAAMAAJ>.
- Y. Naito, K. Saito, K. Shiiba, A. Ohuchi, K. Saigenji, H. Nagura, and H. Ohtani. CD8+ T cells infiltrated within cancer cell nests as a prognostic factor in human colorectal cancer. *Cancer Res*, 58(16):3491–3494, Aug 1998.
- J. Neefjes, M. L. M. Jongsma, P. Paul, and O. Bakke. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol*, 11(12):823–836, 2011. doi: 10.1038/nri3084. URL <http://dx.doi.org/10.1038/nri3084>.
- C. G. A. Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, 487(7407):330–337, Jul 2012. doi: 10.1038/nature11252. URL <http://dx.doi.org/10.1038/nature11252>.
- M. Nielsen, C. Lundegaard, T. Blicher, K. Lamberth, M. Harndahl, S. Justesen, G. Rder, B. Peters, A. Sette, O. Lund, and S. Buus. NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS One*, 2(8):e796, 2007. doi: 10.1371/journal.pone.0000796. URL <http://dx.doi.org/10.1371/journal.pone.0000796>.
- M. Nielsen, O. Lund, S. Buus, and C. Lundegaard. MHC class II epitope predictive algorithms. *Immunology*, 130(3):319–328, Jul 2010. doi: 10.1111/j.1365-2567.2010.03268.x. URL <http://dx.doi.org/10.1111/j.1365-2567.2010.03268.x>.
- H. Pages, M. Carlson, S. Falcon, and N. Li. *AnnotationDbi: Annotation Database Interface*, 2015. R package version 1.24.0.

- K. Palucka and J. Banchereau. Dendritic-cell-based therapeutic cancer vaccines. *Immunity*, 39(1):38–48, Jul 2013. doi: 10.1016/j.immuni.2013.07.004. URL <http://dx.doi.org/10.1016/j.immuni.2013.07.004>.
- D. M. Pardoll. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*, 12(4):252–264, Apr 2012. doi: 10.1038/nrc3239. URL <http://dx.doi.org/10.1038/nrc3239>.
- M. R. Parkhurst, J. C. Yang, R. C. Langan, M. E. Dudley, D.-A. N. Nathan, S. A. Feldman, J. L. Davis, R. A. Morgan, M. J. Merino, R. M. Sherry, M. S. Hughes, U. S. Kammula, G. Q. Phan, R. M. Lim, S. A. Wank, N. P. Restifo, P. F. Robbins, C. M. Laurencot, and S. A. Rosenberg. T cells targeting carcinoembryonic antigen can mediate regression of metastatic colorectal cancer but induce severe transient colitis. *Mol Ther*, 19(3):620–626, Mar 2011. doi: 10.1038/mt.2010.272. URL <http://dx.doi.org/10.1038/mt.2010.272>.
- N. Parmentier, V. Stroobant, D. Colau, P. de Diesbach, S. Morel, J. Chapiro, P. van Endert, and B. J. Van den Eynde. Production of an antigenic peptide by insulin-degrading enzyme. *Nat Immunol*, 11(5):449–454, May 2010. doi: 10.1038/ni.1862. URL <http://dx.doi.org/10.1038/ni.1862>.
- S. P. Patel and R. Kurzrock. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. *Mol Cancer Ther*, 14(4):847–856, Apr 2015. doi: 10.1158/1535-7163.MCT-14-0983. URL <http://dx.doi.org/10.1158/1535-7163.MCT-14-0983>.
- S. Paul, D. Weiskopf, M. A. Angelo, J. Sidney, B. Peters, and A. Sette. HLA class I alleles are associated with peptide-binding repertoires of different size, affinity, and immunogenicity. *J Immunol*, 191(12):5831–5839, Dec 2013. doi: 10.4049/jimmunol.1302101. URL <http://dx.doi.org/10.4049/jimmunol.1302101>.
- S. Piconese, B. Valzasina, and M. P. Colombo. OX40 triggering blocks suppression by regulatory T cells and facilitates tumor rejection. *J Exp Med*, 205(4):825–839, Apr 2008. doi: 10.1084/jem.20071341. URL <http://dx.doi.org/10.1084/jem.20071341>.
- K. Polyak, I. Haviv, and I. G. Campbell. Co-evolution of tumor cells and their microenvironment. *Trends Genet*, 25(1):30–38, Jan 2009. doi: 10.1016/j.tig.2008.10.012. URL <http://dx.doi.org/10.1016/j.tig.2008.10.012>.
- S. A. Quezada, L. Z. Jarvinen, E. F. Lind, and R. J. Noelle. CD40/CD154 interactions at the interface of tolerance and immunity. *Annu Rev Immunol*, 22:307–328, 2004. doi: 10.1146/annurev.immunol.22.012703.104533. URL <http://dx.doi.org/10.1146/annurev.immunol.22.012703.104533>.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014. URL <http://www.R-project.org/>.
- N. T. Rajat K. De, editor. *Immunoinformatics (Methods in Molecular Biology)*. Humana Press, 2014. ISBN 1493911147. URL 10.1007/978-1-4939-1115-8.
- J. C. Ramos and I. S. Lossos. Newly emerging therapies targeting viral-related lymphomas. *Curr Oncol Rep*, 13(5):416–426, Oct 2011. doi: 10.1007/s11912-011-0186-8. URL <http://dx.doi.org/10.1007/s11912-011-0186-8>.
- X. Rao, I. Hoof, A. I. C. A. F. Costa, D. van Baarle, and C. Kemir. HLA class I allele promiscuity revisited. *Immunogenetics*, 63(11):691–701, Nov 2011. doi: 10.1007/s00251-011-0552-6. URL <http://dx.doi.org/10.1007/s00251-011-0552-6>.
- N. P. Restifo, M. E. Dudley, and S. A. Rosenberg. Adoptive immunotherapy for cancer: harnessing the T cell response. *Nat Rev Immunol*, 12(4):269–281, Apr 2012. doi: 10.1038/nri3191. URL <http://dx.doi.org/10.1038/nri3191>.

- A. Ribas, R. Kefford, M. A. Marshall, C. J. A. Punt, J. B. Haanen, M. Marmol, C. Garbe, H. Gogas, J. Schachter, G. Linette, P. Lorigan, K. L. Kendra, M. Maio, U. Trefzer, M. Smylie, G. A. McArthur, B. Dreno, P. D. Nathan, J. Mackiewicz, J. M. Kirkwood, J. Gomez-Navarro, B. Huang, D. Pavlov, and A. Hauschild. Phase III randomized clinical trial comparing tremelimumab with standard-of-care chemotherapy in patients with advanced melanoma. *J Clin Oncol*, 31(5):616–622, Feb 2013. doi: 10.1200/JCO.2012.44.6112. URL <http://dx.doi.org/10.1200/JCO.2012.44.6112>.
- A. Rimmer, H. Phan, I. Mathieson, Z. Iqbal, S. R. F. Twigg, W. G. S. C. , A. O. M. Wilkie, G. McVean, and G. Lunter. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet*, 46(8):912–918, Aug 2014. doi: 10.1038/ng.3036. URL <http://dx.doi.org/10.1038/ng.3036>.
- N. A. Rizvi, M. D. Hellmann, A. Snyder, P. Kvistborg, V. Makarov, J. J. Havel, W. Lee, J. Yuan, P. Wong, T. S. Ho, M. L. Miller, N. Rekhtman, A. L. Moreira, F. Ibrahim, C. Bruggeman, B. Gasmir, R. Zappasodi, Y. Maeda, C. Sander, E. B. Garon, T. Merghoub, J. D. Wolchok, T. N. Schumacher, and T. A. Chan. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*, 348(6230):124–128, Apr 2015. doi: 10.1126/science.aaa1348. URL <http://dx.doi.org/10.1126/science.aaa1348>.
- C. Robert, L. Thomas, I. Bondarenko, S. O’Day, J. Weber, C. Garbe, C. Lebbe, J.-F. Baurain, A. Testori, J.-J. Grob, N. Davidson, J. Richards, M. Maio, A. Hauschild, W. H. Miller, Jr, P. Gascon, M. Lotem, K. Harmankaya, R. Ibrahim, S. Francis, T.-T. Chen, R. Humphrey, A. Hoos, and J. D. Wolchok. Ipilimumab plus dacarbazine for previously untreated metastatic melanoma. *N Engl J Med*, 364(26):2517–2526, Jun 2011. doi: 10.1056/NEJMoa1104621. URL <http://dx.doi.org/10.1056/NEJMoa1104621>.
- C. Robert, A. Ribas, J. D. Wolchok, F. S. Hodi, O. Hamid, R. Kefford, J. S. Weber, A. M. Joshua, W.-J. Hwu, T. C. Gangadhar, A. Patnaik, R. Dronca, H. Zarour, R. W. Joseph, P. Boasberg, B. Chmielowski, C. Mateus, M. A. Postow, K. Gergich, J. Ellassaiss-Schaap, X. N. Li, R. Iannone, S. W. Ebbinghaus, S. P. Kang, and A. Daud. Anti-programmed-death-receptor-1 treatment with pembrolizumab in ipilimumab-refractory advanced melanoma: a randomised dose-comparison cohort of a phase 1 trial. *Lancet*, 384(9948):1109–1117, Sep 2014. doi: 10.1016/S0140-6736(14)60958-2. URL [http://dx.doi.org/10.1016/S0140-6736\(14\)60958-2](http://dx.doi.org/10.1016/S0140-6736(14)60958-2).
- J. Robinson and S. G. Marsh. The IMGT/HLA sequence database. *Rev Immunogenet*, 2(4):518–531, 2000.
- J. Robinson, J. A. Halliwell, H. McWilliam, R. Lopez, P. Parham, and S. G. E. Marsh. The IMGT/HLA database. *Nucleic Acids Res*, 41(Database issue):D1222–D1227, Jan 2013. doi: 10.1093/nar/gks949. URL <http://dx.doi.org/10.1093/nar/gks949>.
- M. S. Rooney, S. A. Shukla, C. J. Wu, G. Getz, and N. Hacohen. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*, 160(1-2):48–61, Jan 2015. doi: 10.1016/j.cell.2014.12.033. URL <http://dx.doi.org/10.1016/j.cell.2014.12.033>.
- K. M. Ropponen, M. J. Eskelinen, P. K. Lipponen, E. Alhava, and V. M. Kosma. Prognostic value of tumour-infiltrating lymphocytes (TILs) in colorectal cancer. *J Pathol*, 182(3):318–324, Jul 1997. doi: 3.0.CO;2-6. URL <http://dx.doi.org/3.0.CO;2-6>.
- S. A. Rosenberg and N. P. Restifo. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science*, 348(6230):62–68, Apr 2015. doi: 10.1126/science.aaa4967. URL <http://dx.doi.org/10.1126/science.aaa4967>.
- C. E. Ruby, W. L. Redmond, D. Haley, and A. D. Weinberg. Anti-OX40 stimulation in vivo enhances CD8+ memory T cell survival and significantly increases recall responses. *Eur J Immunol*, 37(1):157–166, Jan 2007. doi: 10.1002/eji.200636428. URL <http://dx.doi.org/10.1002/eji.200636428>.

- C. E. Rudd, A. Taylor, and H. Schneider. CD28 and CTLA-4 coreceptor expression and signal transduction. *Immunol Rev*, 229(1):12–26, May 2009. doi: 10.1111/j.1600-065X.2009.00770.x. URL <http://dx.doi.org/10.1111/j.1600-065X.2009.00770.x>.
- K. Sakuishi, L. Apetoh, J. M. Sullivan, B. R. Blazar, V. K. Kuchroo, and A. C. Anderson. Targeting Tim-3 and PD-1 pathways to reverse T cell exhaustion and restore anti-tumor immunity. *J Exp Med*, 207(10):2187–2194, Sep 2010. doi: 10.1084/jem.20100643. URL <http://dx.doi.org/10.1084/jem.20100643>.
- E. Sato, S. H. Olson, J. Ahn, B. Bundy, H. Nishikawa, F. Qian, A. A. Jungbluth, D. Frosina, S. Gnjatic, C. Ambrosone, J. Kepner, T. Odunsi, G. Ritter, S. Lele, Y.-T. Chen, H. Ohtani, L. J. Old, and K. Odunsi. Intraepithelial CD8+ tumor-infiltrating lymphocytes and a high CD8+/regulatory T cell ratio are associated with favorable prognosis in ovarian cancer. *Proc Natl Acad Sci U S A*, 102(51):18538–18543, Dec 2005. doi: 10.1073/pnas.0509182102. URL <http://dx.doi.org/10.1073/pnas.0509182102>.
- D. Schadendorf, F. S. Hodi, C. Robert, J. S. Weber, K. Margolin, O. Hamid, D. Patt, T.-T. Chen, D. M. Berman, and J. D. Wolchok. Pooled Analysis of Long-Term Survival Data From Phase II and Phase III Trials of Ipilimumab in Unresectable or Metastatic Melanoma. *J Clin Oncol*, 33(17):1889–1894, Jun 2015. doi: 10.1200/JCO.2014.56.2736. URL <http://dx.doi.org/10.1200/JCO.2014.56.2736>.
- D. A. Schaer, J. T. Murphy, and J. D. Wolchok. Modulation of GITR for cancer immunotherapy. *Curr Opin Immunol*, 24(2):217–224, Apr 2012. doi: 10.1016/j.coi.2011.12.011. URL <http://dx.doi.org/10.1016/j.coi.2011.12.011>.
- R. D. Schreiber, L. J. Old, and M. J. Smyth. Cancer immunoediting: integrating immunity’s roles in cancer suppression and promotion. *Science*, 331(6024):1565–1570, Mar 2011. doi: 10.1126/science.1203486. URL <http://dx.doi.org/10.1126/science.1203486>.
- E. S. Schultz, J. Chapiro, C. Lurquin, S. Claverol, O. Burlet-Schiltz, G. Warnier, V. Russo, S. Morel, F. Lvy, T. Boon, B. J. Van den Eynde, and P. van der Bruggen. The production of a new MAGE-3 peptide presented to cytolytic T lymphocytes by HLA-B40 requires the immunoproteasome. *J Exp Med*, 195(4):391–399, Feb 2002.
- T. Schumacher, L. Bunse, S. Pusch, F. Sahm, B. Wiestler, J. Quandt, O. Menn, M. Osswald, I. Oezen, M. Ott, M. Keil, J. Bal, K. Rauschenbach, A. K. Grabowska, I. Vogler, J. Diekmann, N. Trautwein, S. B. Eichmller, J. Okun, S. Stevanovi, A. B. Riemer, U. Sahin, M. A. Friese, P. Beckhove, A. von Deimling, W. Wick, and M. Platten. A vaccine targeting mutant IDH1 induces antitumour immunity. *Nature*, 512(7514):324–327, Aug 2014. doi: 10.1038/nature13387. URL <http://dx.doi.org/10.1038/nature13387>.
- T. N. Schumacher and R. D. Schreiber. Neoantigens in cancer immunotherapy. *Science*, 348(6230):69–74, Apr 2015. doi: 10.1126/science.aaa4971. URL <http://dx.doi.org/10.1126/science.aaa4971>.
- T. N. Schumacher, C. Kesmir, and M. M. van Buuren. Biomarkers in cancer immunotherapy. *Cancer Cell*, 27(1):12–14, Jan 2015. doi: 10.1016/j.ccell.2014.12.004. URL <http://dx.doi.org/10.1016/j.ccell.2014.12.004>.
- R. H. Schwartz. Costimulation of T lymphocytes: the role of CD28, CTLA-4, and B7/BB1 in interleukin-2 production and immunotherapy. *Cell*, 71(7):1065–1068, Dec 1992.
- L. Senovilla, E. Vacchelli, J. Galon, S. Adjemian, A. Eggermont, W. H. Fridman, C. Sauts-Fridman, Y. Ma, E. Tartour, L. Zitvogel, G. Kroemer, and L. Galluzzi. Trial watch: Prognostic and predictive value of the immune infiltrate in cancer. *Oncoimmunology*, 1(8):1323–1343, Nov 2012. doi: 10.4161/onci.22009. URL <http://dx.doi.org/10.4161/onci.22009>.

- A. Sette, L. Adorini, S. M. Colon, S. Buus, and H. M. Grey. Capacity of intact proteins to bind to MHC class II molecules. *J Immunol*, 143(4):1265–1267, Aug 1989.
- K. S. Sfanos, T. C. Bruno, A. K. Meeker, A. M. De Marzo, W. B. Isaacs, and C. G. Drake. Human prostate-infiltrating CD8+ T lymphocytes are oligoclonal and PD-1+. *Prostate*, 69(15):1694–1703, Nov 2009. doi: 10.1002/pros.21020. URL <http://dx.doi.org/10.1002/pros.21020>.
- V. Shankaran, H. Ikeda, A. T. Bruce, J. M. White, P. E. Swanson, L. J. Old, and R. D. Schreiber. IFN $\gamma$  and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature*, 410(6832):1107–1111, Apr 2001. doi: 10.1038/35074122. URL <http://dx.doi.org/10.1038/35074122>.
- P. Sharma and J. P. Allison. The future of immune checkpoint therapy. *Science*, 348(6230):56–61, Apr 2015. doi: 10.1126/science.aaa8172. URL <http://dx.doi.org/10.1126/science.aaa8172>.
- C. J. Sherr. Principles of tumor suppression. *Cell*, 116(2):235–246, Jan 2004.
- W. W. Shuford, K. Klussman, D. D. Tritchler, D. T. Loo, J. Chalupny, A. W. Siadak, T. J. Brown, J. Emswiler, H. Raecho, C. P. Larsen, T. C. Pearson, J. A. Ledbetter, A. Aruffo, and R. S. Mittler. 4-1BB costimulatory signals preferentially induce CD8+ T cell proliferation and lead to the amplification in vivo of cytotoxic T cell responses. *J Exp Med*, 186(1):47–55, Jul 1997.
- S. A. Shukla, M. S. Rooney, M. Rajasagi, G. Tiao, P. M. Dixon, M. S. Lawrence, J. Stevens, W. J. Lane, J. L. Dellagatta, S. Steelman, C. Sougnez, K. Cibulskis, A. Kiezun, N. Hacohen, V. Brusic, C. J. Wu, and G. Getz. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. *Nat Biotechnol*, 33(11):1152–1158, Nov 2015. doi: 10.1038/nbt.3344. URL <http://dx.doi.org/10.1038/nbt.3344>.
- R. Siegel, C. Desantis, and A. Jemal. Colorectal cancer statistics, 2014. *CA Cancer J Clin*, 64(2):104–117, 2014. doi: 10.3322/caac.21220. URL <http://dx.doi.org/10.3322/caac.21220>.
- S. Sierro, P. Romero, and D. E. Speiser. The CD4-like molecule LAG-3, biology and therapeutic applications. *Expert Opin Ther Targets*, 15(1):91–101, Jan 2011. doi: 10.1517/14712598.2011.540563. URL <http://dx.doi.org/10.1517/14712598.2011.540563>.
- P. P. Singh, P. K. Sharma, G. Krishnan, and A. C. Lockhart. Immune checkpoints and immunotherapy for colorectal cancer. *Gastroenterol Rep (Oxf)*, 3(4):289–297, Nov 2015. doi: 10.1093/gastro/gov053. URL <http://dx.doi.org/10.1093/gastro/gov053>.
- M. J. Smyth, S. F. Ngiew, A. Ribas, and M. W. L. Teng. Combination cancer immunotherapies tailored to the tumour microenvironment. *Nat Rev Clin Oncol*, 13(3):143–158, Mar 2016. doi: 10.1038/nrclinonc.2015.209. URL <http://dx.doi.org/10.1038/nrclinonc.2015.209>.
- A. Sanchez-Fueyo, J. Tian, D. Picarella, C. Domenig, X. X. Zheng, C. A. Sabatos, N. Manlongat, O. Bender, T. Kamradt, V. K. Kuchroo, J.-C. Gutierrez-Ramos, A. J. Coyle, and T. B. Strom. Tim-3 inhibits T helper type 1-mediated auto- and alloimmune responses and promotes immunological tolerance. *Nat Immunol*, 4(11):1093–1101, Nov 2003. doi: 10.1038/ni987. URL <http://dx.doi.org/10.1038/ni987>.
- A. Snyder, V. Makarov, T. Merghoub, J. Yuan, J. M. Zaretsky, A. Desrichard, L. A. Walsh, M. A. Postow, P. Wong, T. S. Ho, T. J. Hollmann, C. Bruggeman, K. Kannan, Y. Li, C. Elipenahli, C. Liu, C. T. Harbison, L. Wang, A. Ribas, J. D. Wolchok, and T. A. Chan. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med*, 371(23):2189–2199, Dec 2014. doi: 10.1056/NEJMoa1406498. URL <http://dx.doi.org/10.1056/NEJMoa1406498>.

- J. Straeter, U. Hinz, C. Hasel, U. Bhanot, G. Mechttersheimer, T. Lehnert, and P. Mueller. Impaired CD95 expression predisposes for recurrence in curatively resected colon carcinoma: clinical evidence for immunoselection and CD95L mediated control of minimal residual disease. *Gut*, 54(5):661–665, May 2005. doi: 10.1136/gut.2004.052696. URL <http://dx.doi.org/10.1136/gut.2004.052696>.
- T. Sturniolo, E. Bono, J. Ding, L. Radrizzani, O. Tuereci, U. Sahin, M. Braxenthaler, F. Gallazzi, M. P. Protti, F. Sinigaglia, and J. Hammer. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat Biotechnol*, 17(6): 555–561, Jun 1999. doi: 10.1038/9858. URL <http://dx.doi.org/10.1038/9858>.
- M. Sznol and L. Chen. Antagonist antibodies to PD-1 and B7-H1 (PD-L1) in the treatment of advanced human cancer–response. *Clin Cancer Res*, 19(19):5542, Oct 2013. doi: 10.1158/1078-0432.CCR-13-2234. URL <http://dx.doi.org/10.1158/1078-0432.CCR-13-2234>.
- A. Szolek, B. Schubert, C. Mohr, M. Sturm, M. Feldhahn, and O. Kohlbacher. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics*, 30(23):3310–3316, Dec 2014. doi: 10.1093/bioinformatics/btu548. URL <http://dx.doi.org/10.1093/bioinformatics/btu548>.
- J. M. Taube, A. Klein, J. R. Brahmer, H. Xu, X. Pan, J. H. Kim, L. Chen, D. M. Pardoll, S. L. Topalian, and R. A. Anders. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res*, 20(19):5064–5074, Oct 2014. doi: 10.1158/1078-0432.CCR-13-3271. URL <http://dx.doi.org/10.1158/1078-0432.CCR-13-3271>.
- A. Tikidzhieva, A. Benner, S. Michel, A. Formentini, K.-H. Link, W. Dippold, M. von Knebel Doeberitz, M. Kornmann, and M. Kloor. Microsatellite instability and Beta2-Microglobulin mutations as prognostic markers in colon cancer: results of the FOGT-4 trial. *Br J Cancer*, 106(6):1239–1245, Mar 2012. doi: 10.1038/bjc.2012.53. URL <http://dx.doi.org/10.1038/bjc.2012.53>.
- R. Todd and D. T. Wong. Oncogenes. *Anticancer Res*, 19(6A):4729–4746, 1999.
- S. L. Topalian, F. S. Hodi, J. R. Brahmer, S. N. Gettinger, D. C. Smith, D. F. McDermott, J. D. Powderly, R. D. Carvajal, J. A. Sosman, M. B. Atkins, P. D. Leming, D. R. Spigel, S. J. Antonia, L. Horn, C. G. Drake, D. M. Pardoll, L. Chen, W. H. Sharfman, R. A. Anders, J. M. Taube, T. L. McMiller, H. Xu, A. J. Korman, M. Jure-Kunkel, S. Agrawal, D. McDonald, G. D. Kollia, A. Gupta, J. M. Wigginton, and M. Sznol. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med*, 366(26):2443–2454, Jun 2012. doi: 10.1056/NEJMoa1200690. URL <http://dx.doi.org/10.1056/NEJMoa1200690>.
- N. Toussaint and O. Kohlbacher. Towards in silico design of epitope-based vaccines. *Expert Opin Drug Discovery*, 4(10):1047–1060, 2009.
- E. Tran, S. Turcotte, A. Gros, P. F. Robbins, Y.-C. Lu, M. E. Dudley, J. R. Wunderlich, R. P. Somerville, K. Hogan, C. S. Hinrichs, M. R. Parkhurst, J. C. Yang, and S. A. Rosenberg. Cancer immunotherapy based on mutation-specific CD4+ T cells in a patient with epithelial cancer. *Science*, 344(6184):641–645, May 2014. doi: 10.1126/science.1251102. URL <http://dx.doi.org/10.1126/science.1251102>.
- T. Trolle, I. G. Metushi, J. A. Greenbaum, Y. Kim, J. Sidney, O. Lund, A. Sette, B. Peters, and M. Nielsen. Automated benchmarking of peptide-MHC class I binding predictions. *Bioinformatics*, 31(13):2174–2181, Jul 2015. doi: 10.1093/bioinformatics/btv123. URL <http://dx.doi.org/10.1093/bioinformatics/btv123>.
- S. J. Turley, V. Cremasco, and J. L. Astarita. Immunological hallmarks of stromal cells in the tumour microenvironment. *Nat Rev Immunol*, 15(11):669–682, Nov 2015. doi: 10.1038/nri3902. URL <http://dx.doi.org/10.1038/nri3902>.

- J. G. van den Boorn and G. Hartmann. Turning tumors into vaccines: co-opting the innate immune system. *Immunity*, 39(1):27–37, Jul 2013. doi: 10.1016/j.immuni.2013.07.011. URL <http://dx.doi.org/10.1016/j.immuni.2013.07.011>.
- P. van der Bruggen, C. Traversari, P. Chomez, C. Lurquin, E. De Plaen, B. Van den Eynde, A. Knuth, and T. Boon. A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science*, 254(5038):1643–1647, Dec 1991.
- S. H. van der Burg and C. J. M. Melief. Therapeutic vaccination against human papilloma virus induced malignancies. *Curr Opin Immunol*, 23(2):252–257, Apr 2011. doi: 10.1016/j.coi.2010.12.010. URL <http://dx.doi.org/10.1016/j.coi.2010.12.010>.
- S. H. van der Burg, M. J. Visseren, R. M. Brandt, W. M. Kast, and C. J. Melief. Immunogenicity of peptides bound to MHC class I molecules depends on the MHC-peptide complex stability. *J Immunol*, 156(9):3308–3314, May 1996.
- S. H. van der Burg, R. Arens, F. Ossendorp, T. van Hall, and C. J. M. Melief. Vaccines for established cancer: overcoming the challenges posed by immune evasion. *Nat Rev Cancer*, Mar 2016. doi: 10.1038/nrc.2016.16. URL <http://dx.doi.org/10.1038/nrc.2016.16>.
- M. D. Vesely and R. D. Schreiber. Cancer immunoediting: antigens, mechanisms, and implications to cancer immunotherapy. *Ann N Y Acad Sci*, 1284:1–5, May 2013. doi: 10.1111/nyas.12105. URL <http://dx.doi.org/10.1111/nyas.12105>.
- M. D. Vesely, M. H. Kershaw, R. D. Schreiber, and M. J. Smyth. Natural innate and adaptive immunity to cancer. *Annu Rev Immunol*, 29:235–271, 2011. doi: 10.1146/annurev-immunol-031210-101324. URL <http://dx.doi.org/10.1146/annurev-immunol-031210-101324>.
- N. Vigneron, V. Stroobant, J. Chapiro, A. Ooms, G. Degiovanni, S. Morel, P. van der Bruggen, T. Boon, and B. J. Van den Eynde. An antigenic peptide produced by peptide splicing in the proteasome. *Science*, 304(5670):587–590, Apr 2004. doi: 10.1126/science.1095522. URL <http://dx.doi.org/10.1126/science.1095522>.
- N. Vigneron, V. Stroobant, B. J. Van den Eynde, and P. van der Bruggen. Database of T cell-defined human tumor antigens: the 2013 update. *Cancer Immunol*, 13:15, 2013.
- D. S. Vinay and B. S. Kwon. Immunotherapy of cancer with 4-1BB. *Mol Cancer Ther*, 11(5):1062–1070, May 2012. doi: 10.1158/1535-7163.MCT-11-0677. URL <http://dx.doi.org/10.1158/1535-7163.MCT-11-0677>.
- A. M. Vlad, J. C. Kettel, N. M. Alajez, C. A. Carlos, and O. J. Finn. MUC1 immunobiology: from discovery to clinical applications. *Adv Immunol*, 82:249–293, 2004. doi: 10.1016/S0065-2776(04)82006-6. URL [http://dx.doi.org/10.1016/S0065-2776\(04\)82006-6](http://dx.doi.org/10.1016/S0065-2776(04)82006-6).
- R. H. Vonderheide and M. J. Glennie. Agonistic CD40 antibodies and cancer therapy. *Clin Cancer Res*, 19(5):1035–1043, Mar 2013. doi: 10.1158/1078-0432.CCR-12-2064. URL <http://dx.doi.org/10.1158/1078-0432.CCR-12-2064>.
- M. Vormehr, M. Diken, S. Boegel, S. Kreiter, z. Treci, and U. Sahin. Mutanome directed cancer immunotherapy. *Curr Opin Immunol*, 39:14–22, Dec 2015. doi: 10.1016/j.coi.2015.12.001. URL <http://dx.doi.org/10.1016/j.coi.2015.12.001>.
- C. Wang, G. H. Y. Lin, A. J. McPherson, and T. H. Watts. Immune regulation by 4-1BB and 4-1BBL: complexities and challenges. *Immunol Rev*, 229(1):192–215, May 2009. doi: 10.1111/j.1600-065X.2009.00765.x. URL <http://dx.doi.org/10.1111/j.1600-065X.2009.00765.x>.

- K. Wang, M. Li, and H. Hakonarson. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*, 38(16):e164, Sep 2010. doi: 10.1093/nar/gkq603. URL <http://dx.doi.org/10.1093/nar/gkq603>.
- R. L. Warren, G. Choe, D. J. Freeman, M. Castellarin, S. Munro, R. Moore, and R. A. Holt. Derivation of HLA types from shotgun sequence datasets. *Genome Med*, 4(12):95, 2012. doi: 10.1186/gm396. URL <http://dx.doi.org/10.1186/gm396>.
- R. A. Weinberg. *The Biology of Cancer, 2nd Edition*. Garland Science, 2013. ISBN 0815342209.
- K. Weintraub. Drug development: Releasing the brakes. *Nature*, 504(7480):S6–S8, Dec 2013. doi: 10.1038/504S6a. URL <http://dx.doi.org/10.1038/504S6a>.
- E. J. Wherry, S.-J. Ha, S. M. Kaeck, W. N. Haining, S. Sarkar, V. Kalia, S. Subramaniam, J. N. Blattman, D. L. Barber, and R. Ahmed. Molecular signature of CD8+ T cell exhaustion during chronic viral infection. *Immunity*, 27(4):670–684, Oct 2007. doi: 10.1016/j.immuni.2007.09.006. URL <http://dx.doi.org/10.1016/j.immuni.2007.09.006>.
- H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009. ISBN 978-0-387-98140-6. URL <http://had.co.nz/ggplot2/book>.
- J. D. Wolchok, H. Kluger, M. K. Callahan, M. A. Postow, N. A. Rizvi, A. M. Lesokhin, N. H. Segal, C. E. Ariyan, R.-A. Gordon, K. Reed, M. M. Burke, A. Caldwell, S. A. Kronenberg, B. U. Agunwamba, X. Zhang, I. Lowy, H. D. Inzunza, W. Feely, C. E. Horak, Q. Hong, A. J. Korman, J. M. Wigginton, A. Gupta, and M. Sznol. Nivolumab plus ipilimumab in advanced melanoma. *N Engl J Med*, 369(2):122–133, Jul 2013. doi: 10.1056/NEJMoa1302369. URL <http://dx.doi.org/10.1056/NEJMoa1302369>.
- S. B. J. Wong, R. Bos, and L. A. Sherman. Tumor-specific CD4+ T cells render the tumor environment permissive for infiltration by low-avidity CD8+ T cells. *J Immunol*, 180(5):3122–3131, Mar 2008.
- S.-R. Woo, M. E. Turnis, M. V. Goldberg, J. Bankoti, M. Selby, C. J. Nirschl, M. L. Bettini, D. M. Gravano, P. Vogel, C. L. Liu, S. Tansombatvisit, J. F. Grosso, G. Netto, M. P. Smeltzer, A. Chaux, P. J. Utz, C. J. Workman, D. M. Pardoll, A. J. Korman, C. G. Drake, and D. A. A. Vignali. Immune inhibitory molecules LAG-3 and PD-1 synergistically regulate T-cell function to promote tumoral immune escape. *Cancer Res*, 72(4):917–927, Feb 2012. doi: 10.1158/0008-5472.CAN-11-1620. URL <http://dx.doi.org/10.1158/0008-5472.CAN-11-1620>.
- T. Xie, G. D’Ario, J. R. Lamb, E. Martin, K. Wang, S. Tejpar, M. Delorenzi, F. T. Bosman, A. D. Roth, P. Yan, S. Bougel, A. F. Di Narzo, V. Popovici, E. Budinsk, M. Mao, S. L. Weinrich, P. A. Rejto, and J. G. Hodgson. A comprehensive characterization of genome-wide copy number aberrations in colorectal cancer reveals novel oncogenes and patterns of alterations. *PLoS One*, 7(7):e42001, 2012. doi: 10.1371/journal.pone.0042001. URL <http://dx.doi.org/10.1371/journal.pone.0042001>.
- M. Yadav, S. Jhunjhunwala, Q. T. Phung, P. Lupardus, J. Tanguay, S. Bumbaca, C. Franci, T. K. Cheung, J. Fritsche, T. Weinschenk, Z. Modrusan, I. Mellman, J. R. Lill, and L. Delamarre. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature*, 515(7528):572–576, Nov 2014. doi: 10.1038/nature14001. URL <http://dx.doi.org/10.1038/nature14001>.
- C. Yanover and P. Bradley. Large-scale characterization of peptide-MHC binding landscapes with structural simulations. *Proc Natl Acad Sci U S A*, 108(17):6981–6986, Apr 2011. doi: 10.1073/pnas.1018165108. URL <http://dx.doi.org/10.1073/pnas.1018165108>.
- M. Yasunaga, Y. Tabira, K. Nakano, S. Iida, N. Ichimaru, N. Nagamoto, and T. Sakaguchi. Accelerated growth signals and low tumor-infiltrating lymphocyte levels predict poor outcome in T4 esophageal squamous cell carcinoma. *Ann Thorac Surg*, 70(5):1634–1640, Nov 2000.

- J. Ye, G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, and T. L. Madden. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, 13:134, 2012. doi: 10.1186/1471-2105-13-134. URL <http://dx.doi.org/10.1186/1471-2105-13-134>.
- M. Yoshimoto, G. Sakamoto, and Y. Ohashi. Time dependency of the influence of prognostic factors on relapse in breast cancer. *Cancer*, 72(10):2993–3001, Nov 1993.
- B. Zhang, J. Wang, X. Wang, J. Zhu, Q. Liu, Z. Shi, M. C. Chambers, L. J. Zimmerman, K. F. Shaddox, S. Kim, S. R. Davies, S. Wang, P. Wang, C. R. Kinsinger, R. C. Rivers, H. Rodriguez, R. R. Townsend, M. J. C. Ellis, S. A. Carr, D. L. Tabb, R. J. Coffey, R. J. C. Slebos, D. C. Liebler, and N. C. I. C. P. T. A. C. . Proteogenomic characterization of human colon and rectal cancer. *Nature*, 513(7518):382–387, Sep 2014. doi: 10.1038/nature13438. URL <http://dx.doi.org/10.1038/nature13438>.
- S. Zhao, Y. Guo, Q. Sheng, and Y. Shyr. *heatmap3: An Improved Heatmap Package*, 2015. URL <http://CRAN.R-project.org/package=heatmap3>. R package version 1.1.1.
- L. Zitvogel, O. Kepp, and G. Kroemer. Immune parameters affecting the efficacy of chemotherapeutic regimens. *Nat Rev Clin Oncol*, 8(3):151–160, Mar 2011. doi: 10.1038/nrclinonc.2010.223. URL <http://dx.doi.org/10.1038/nrclinonc.2010.223>.

## CHAPTER 7

---

Manuscript

---

# Identification of immunotherapeutic targets by genomic profiling of rectal NET metastases

Zeynep Kosaloglu<sup>1,2</sup>, Inka Zörnig<sup>2</sup>, Niels Halama<sup>2</sup>, Iris Kaiser<sup>2</sup>, Ivo Buchhalter<sup>3</sup>, Niels Grabe<sup>2</sup>, Roland Eils<sup>3,4</sup>, Matthias Schlesner<sup>3</sup>, Andrea Califano<sup>5</sup>, Dirk Jäger<sup>1,2</sup>

1 Clinical Cooperation Unit “Applied Tumor Immunity”, National Center for Tumor Diseases (NCT) and German Cancer Research Center (DKFZ), Heidelberg, Germany

2 Department of Medical Oncology, National Center for Tumor Diseases (NCT) and University Hospital Heidelberg, Heidelberg, Germany

3 Division of Theoretical Bioinformatics (B080), German Cancer Research Center (DKFZ), Heidelberg, Germany

4 Department for Bioinformatics and Functional Genomics, Institute for Pharmacy and Molecular Biotechnology (IPMB) and BioQuant, Heidelberg University, Heidelberg, Germany

5 Department of Biomedical Informatics, Department of Systems Biology, Center for Computational Biology and Bioinformatics, Herbert Irving Comprehensive Cancer Center, Columbia University, New York, USA

## Abstract

Neuroendocrine tumors (NETs) of the gastrointestinal tract are a rare and heterogeneous group of neoplasms with unique tumor biology and clinical management issues. While surgery is the only curative treatment option in patients with early stage NETs, the optimal management strategy for patients with advanced metastatic NETs is unknown. Based on the tremendous success of immunotherapeutic approaches, we sought to investigate such approaches in a case of metastatic rectal NET. Here, we apply an integrative approach using various computational and experimental methods to explore several aspects of the tumor-host immune interactions for immunotherapeutic options. Sequencing of six different liver metastases revealed a quite homogenous set of mutations, and further analysis of these mutations for immunogenicity revealed few neo-epitopes with preexisting T cell reactivity, which can be used in therapeutic vaccines. Staining for immunomodulatory proteins and cytokine profiling showed that the immune setting is surprisingly different, when compared to liver metastases of colorectal cancer for instance. Taken together, our results highlight the broad range and complexity of tumor-host immune interaction and underline the value of an integrative approach.

**Keywords:** immunotherapy, neoantigens, neo-epitopes, bioinformatics, NGS, IHC, cytokines, tumor microenvironment, vaccination, tumor immunology, integrated analysis

**Abbreviations:** NET, neuroendocrine tumor; WGS, whole-genome-sequencing; RNA-Seq, RNA-sequencing; SNV, single nucleotide variation; indel, short insertion or deletion; HLA, human leukocyte antigen; RPKM, reads per kilobase per million reads

## Introduction

Neuroendocrine tumors (NETs) represent an extremely heterogeneous group of tumors. NETs develop from neuroendocrine cells at various primary organ sites and are grouped according to their origin: the foregut, the midgut, and hindgut <sup>1</sup>. The majority of NETs are found throughout the intestinal tract, most commonly in the pancreas. NETs of the colon and rectum are rare, comprising less than 1 percent of colon and rectal cancers, and rectal NETs constitute about 19 percent of all gastrointestinal NETs <sup>2,3</sup>.

Despite the considerable advances made in recent decades, the genetic and molecular determinants of NET tumor biology remain poorly characterized. Pathological features found in many NETs are poor differentiation, distinctive expression of neuroendocrine markers and the ability to secrete bioactive peptides <sup>4</sup>.

NETs present a wide spectrum of malignancies from relatively indolent to highly aggressive variants. The prognosis for high-grade poorly differentiated neuroendocrine carcinomas however is generally poor, as most patients have metastatic disease at the time of diagnosis, whereas patients with low-grade NETs have a rather favorable prognosis <sup>2</sup>. The low incidence of neuroendocrine carcinomas coupled with their variable clinical manifestation was a major barrier to investigation, and advances in the treatment of NETs have been slow. Optimal treatment of NETs has been widely debated and remains to be controversial. Response to chemotherapy has been traditionally poor for low grade NETs with an unfavorable benefit-to-risk ratio. When feasible, surgical removal of malignant tissue is the primary treatment option for NETs and offers the best prognosis <sup>5,6</sup>.

Due to the highly heterogeneous nature of most NETs, a personalized treatment strategy might be an appropriate approach and it has been suggested, that genome-wide screening for mutations may reveal new data that can be used for a more appropriate treatment selection <sup>7</sup>.

Furthermore, immunotherapies such as vaccinations against neoantigens and checkpoint blockade therapies have shown dramatic success in a number of tumor entities (for a review see refs. 8 and 9), and it has been suggested, that immunotherapy in NETs provides opportunities for future advances <sup>1</sup>. However it is largely unclear how the setup of the immunological microenvironment in metastatic NET is and especially how diverse these mutations and immunological setups can be between different metastatic lesions.

In this study, we characterized the genomic features of an individual metastatic NET of the rectum, by whole-genome and RNA-sequencing analysis (WGS and RNA-Seq) of six metastatic lesions. It has been repeatedly shown, that individual, tumor-specific mutations found in the genome of cancer patients provide a superior source of immunogenic targets. Mutation-derived epitopes, so called neo-epitopes, are expressed in a highly tumor-specific manner, and are also expected to overcome central tolerance. We identified several neo-epitopes in the tumor samples, which are suitable candidates for peptide- or RNA- based vaccines. We furthermore characterized the immunological features of the tumor by assessing the density of lymphocyte infiltration into the metastatic site, as well as the expression of immunomodulatory proteins, as it is known, that these can be indicators for disease progression and immunotherapy outcome. We also measured existing immune responses against several predicted neo-epitopes from the peripheral blood of the same patient.

Here, we present a unique dataset covering different aspects of tumor immunology. Our data gives a comprehensive insight into the immune setting of a metastatic NET of the rectum, and nicely demonstrates our integrative analysis-workflow for the prediction of mutation-derived neo-epitopes and profiling of the immunological tumor microenvironment leading to the identification of suitable immunotherapeutic target molecules.

## Results

### 3.1 Somatic mutations

---

A total of 15 non-synonymous somatic coding SNVs and 4 indels were detected in the WGS data, which are also expressed according to the RNA-Seq data. Mutations in 11 genes were found in all of the six liver metastases, and additional two mutations were found in the majority of the liver metastases. One mutation was detected in two samples and the remaining five mutations were unique to one sample (Figure 1 and Table 3).

### 3.3 Epitopes

---

For the prediction of neo-epitopes we only considered expressed SNVs and indels that were present in the majority of samples. Hence, 13 mutations were considered. As a result, nineteen peptides from eight mutated genes were predicted to bind with high to intermediate affinity to one of the patient's HLA alleles (Table 4). Seven peptides were predicted to bind to HLA-A, fourteen to HLA-C, and no peptides were predicted to bind to HLA-B.

Thirteen of the mutated peptides have better predicted binding affinities compared to the cognate wild type peptide. Seven mutated/wild type peptide pairs have comparable predicted binding affinities, and for one peptide, the wild type peptide has a better affinity than the mutated counterpart. Two mutated peptides were predicted to bind by both HLA-C08 and HLA-C03.

A frameshift deletion in the gene OBSL1 generates a novel peptide of 28 amino acids. In these peptide, nine 8-11mer peptides were predicted to bind with high to intermediate affinity to one of the patient's HLA alleles. Three peptides were predicted to bind both HLA-A03 and HLA-A33.

For the mutated genes with predicted epitopes, we also calculated the RPKM values from the normal and tumor RNA-Seq data to assess change in gene expression. For six of these eight genes, namely ADAM9, MAT2A, RABEP1, SLC11A2, SMARCA1, and OBSL1, gene expression is increased in the tumor, with ADAM9 and OBSL1 being highly over-expressed in the tumor with a fold change of over 2.

### **3.4 ELISPOT**

To assess whether there are T cells in the patient's peripheral blood that are reactive against the predicted neo-epitopes, we conducted IFN $\gamma$  ELISpot assays. As negative control, DCs were loaded with human IgG and antigen-specific T cell reactivity was assumed if spot numbers in triplicate test wells significantly exceeded those of control wells. An overview image of the IFN $\gamma$  ELISpot plate with patient-derived cells and quantitative data of this plate presented in Figure 2.

The peptides MAT2A-mt3, MAT2A-wt3, RAPEP1-mt, RABEP1-wt, SERPINF2-wt3, OBSL1-neoORF3 showed higher spot numbers when compared to the spot counts of control wells, indicating that there are reactive T cells in the peripheral blood against the tested peptides.

### **3.5 Immunohistochemistry and Cytokine analysis**

---

In order to assess, whether epitope presentation is altered, we analyzed HLA class I expression in all samples. Interestingly, a uniform HLA class I expression among all samples could be detected in our immunohistochemical stainings. T cell and B cell numbers at the metastatic sites are generally very low, and even absent in many parts of the analyzed metastases. Not surprisingly, cell numbers for NK cells are also generally low. Concordant with the low infiltration of effector T cells, PD-1 and PD-L1 are also only found in low levels in the analyzed tissues.

The metastatic tissue LM6 was used for the analysis of 50 cytokines and chemokines. Protein concentrations of the analyzed cytokines were found to be quite low and the measured cytokines do not reflect a Th1 or Th2 cytokine profile. Only a small subset of cytokines, namely MIF, VCAM-1 and ICAM-1, are expressed in a significant concentration, compared to normal adjacent liver (Figure 3).

## Discussion

NETs of the gastrointestinal tract are a rare and heterogeneous group of neoplasms with a unique tumor biology and clinical management issues. While surgery is the only curative treatment option in patients with early stage NETs, the optimal management strategy for patients with advanced metastatic NETs is unknown. Based on the tremendous success of immunotherapeutic approaches, we sought to investigate such approaches in a case of metastatic rectal NET.

As known from studies on other NET types, the total number of somatic mutations is low compared to other tumor entities. In our case, between thirteen and fifteen expressed mutations were detected in each sample and 19 distinct mutations were found in all samples together.

Comparing the mutated genes from all liver metastases, it is striking that the overlap between them is quite large. A set of 13 mutations can be found nearly in all analyzed liver metastasis samples, regardless of their spatial and temporal differences. In cancer it is usually observed that the tumors acquire more mutations with disease progression, and emerging evidence suggests that genetic abnormalities vary substantially between metastases or even within a single tumor mass, indicating intratumor heterogeneity<sup>10-12</sup>. Determining the impact of such clonal structures still requires future studies. Intratumoral heterogeneity has important consequences for personalized-medicine approaches. Several clinical observations of different tumor types have shown variability in response to therapy that can occur between metastases

or within a single tumor mass<sup>13-16</sup>. This variability may be explained by the emergence of genomically distinct clones of malignant cells.

We could not observe any major clonal variation in the metastases of this rectal NET patient, as there is a striking overlap of mutated genes. This may be due to the indolent nature of neuroendocrine tumors compared to their adenocarcinomatous counterparts. Thirteen mutations are shared among nearly all analyzed metastases. This fact alone indicates a central role of these genes in disease progression. Exploring the molecular functions of these mutated genes revealed some insights which may indicate a role in tumorigenesis.

HSPG2 encodes the protein perlecan which is a large multidomain proteoglycan that binds and cross-links cell surface molecules. Due to its many interactions, perlecan plays essential roles in multiple biological activities, such as vascularization and angiogenesis, and might hence contribute to disease progression in this patient<sup>17</sup>. It was suggested, that the protein encoded by SMARCA1, also called SNF2L, may play a role in DNA damage, growth inhibition and apoptosis of cancer cells<sup>18,19</sup>. Thus, we can assume that a non-synonymous mutation in SMARCA1 could contribute positively to tumorigenesis. Another interesting mutated gene is SERPINF 2. This gene encodes a member of the serpin family of serine proteases and the proper function of this gene has a major role in regulating the blood clotting pathway. Mutations of SERPINF2 are characterized by severe hemorrhage, which was also observed in metastasis LM6 in our case.

Mutated antigens greatly contribute to the immunogenicity of human tumors. A single nucleotide variation can potentially lead to the production of various new antigenic peptides (neoantigens) that can be recognized by autologous T cells. Additionally, so called neoORF antigens can be generated by frameshifting insertions or deletions, which are supposed to be completely novel and were also shown to be recognized by autologous T cells<sup>20-22</sup>.

These neoantigens are highly attractive as immunotherapeutic targets as they are expected to overcome central tolerance, and their expression is tumor-specific<sup>23,24</sup>. On the basis of animal model data, proof-of-principle for the feasibility to identify T cell reactivity against patient-specific neoantigens through the exploitation of genome data was obtained<sup>25,26</sup>. It was also shown in numerous studies in mouse models that vaccination with predicted neoantigens results in increased tumor control<sup>25,27-29</sup>. In subsequent studies in humans it has then been demonstrated that tumor sequencing data can also be exploited in a clinical setting, and that neoantigens serve as tumor rejection antigens<sup>28,30-33</sup>.

As no standard treatment for metastatic NETs was established yet, and because NETs are known to be highly heterogeneous and unique regarding the genomic landscape, we suggest, that a personalized treatment strategy, targeting patient-specific neoantigens, might be an appropriate approach.

The generation of such personalized immunotherapeutics requires time, as several steps like sequencing, data analysis, in vitro analysis, and agent production are very time consuming, which makes it inapplicable for some patients with end-stage tumors. Due to the slow progression rate of most NETs, this problem might not apply for NET patients. We found that intratumoral heterogeneity is very limited in our case, which is also favorable for a personalized therapy.

As reconfirmed in this work, NETs acquire a relatively small number of mutations, in our case only 15 expressed non-synonymous somatic coding SNVs and 4 indels were detected. Epitope prediction on peptides containing these SNVs and on neoORF peptides generated by frameshift indels, revealed 28 epitopes that bind with high to intermediate affinity to one of the patient's HLA allele. We then synthesized these predicted neoantigens together with the corresponding wildtype peptides and conducted IFN $\gamma$  ELISpot assays using the patient's peripheral blood to assess if there are T cells that are reactive against them. Our results showed

preexisting T cell responses against some of the tested peptides. Interestingly, peptides with predicted high binding affinities did not show T cell reactivity in our ELISpot analysis. Only the RABEP1 mt and wt peptides which are considered strong binders with a binding affinity of < 50nM, elicited a T cell response in the ELISpot. The wildtype peptide was predicted to bind slightly stronger, and this is also reflected in the ELISpot results, with the wildtype peptide having higher spot counts. The fact that the RABEP1 wildtype peptide also triggers T cell reactivity makes the RABEP1 mutated neo-epitope not suitable for therapeutic vaccination as it might cause auto-reactivity. MAT2A-mt3 is another epitope that showed T cell reactivity in the ELISpot. This neo-epitope was predicted to bind weakly with a binding score of 270nM, whereas the wildtype counterpart MAT2A-wt3 was predicted to be a non-binder. The spot numbers in the ELISpot for MAT2A-wt3 were also not high enough to be considered significant, however the difference to the spot counts of MAT2A-mt3 is not as obvious as the predicted binding affinities. Nevertheless, MAT2A-mt3 can be considered as a suitable neo-epitope for therapeutic vaccination. When looking at the neoORF epitopes generated through a frameshift-mutation in OBSL1, there are several neo-epitopes with high binding affinities. In the ELISpot however, only one neo-epitope, OBSL1-neoORF3 which has an intermediate predicted binding affinity of 188nM, showed T cell reactivity. NeoORF epitopes are ideal candidates for therapeutic vaccination as there is no wildtype counterpart to those epitopes. In our case the OBSL1 gene is additionally overexpressed in the tumor which makes the neo-epitopes harbored in this gene even better candidates.

Using the predicted neoantigens, synthetic vaccines may be produced and administered to the patient. Such therapeutic vaccines are supposed to aid in tumor control. ELISpot analysis for immunogenicity of the predicted neoantigens revealed two neo-epitopes as suitable candidates for therapeutic vaccination, MAT2A-mt3 and OBSL1-neoORF3. Although, there seem to be no reactive T cells against the majority of the predicted neoantigens in the patient's peripheral

blood, antitumor immunity could be induced/boosted with therapeutic vaccination. It is currently still unknown, whether enhancement of an existing T cell response or generation of de novo responses is clinically relevant for an effective tumor vaccine <sup>33</sup>. If a therapeutic vaccine is going to be administered, immunization with multiple neoantigens is of advantage. In doing so, the likelihood of generating an immune response against at least some of the neoantigens increases, and the likelihood of the tumor escaping the immune response by immunoediting decreases. Another option for the therapeutic usage of the patient-specific neoantigen repertoire is to create neoantigen-specific lymphocyte products in vitro, like for example, the adoptive transfer of ex-vivo-activated autologous T cells and natural killer cells. The immunohistochemical stainings for CD3, CD20 and NKp46 showed that the lymphocyte numbers at the metastatic lesions are generally low. In line with this, staining for the modulatory molecules PD-1 and the corresponding ligand PD-L1 showed that these molecules are not prominent as well. Tumeh and colleagues have recently shown in melanoma patients, that the success of checkpoint blockade therapy against PD-1 depends on existing CD8+ T cells that are negatively regulated by PD-1/PD-L1 <sup>34</sup>. Based on these findings, immune checkpoint therapy targeting PD-1 and its ligands seems not to be a suitable therapy option for our NET patient.

Compared to data from colorectal cancer liver metastases <sup>35,36</sup> the cell densities of effector cells are dramatically low. In line with the clinical findings for low-proliferative NETs this low infiltrating indicates a low chance of responding to chemotherapy. The ELISpot data suggest that there is preexisting T cell reactivity against some neoantigens, showing interaction between tumor cells and the immune system, but apparently this interaction is not sufficient to allow higher infiltration of effector cells. An alternative explanation is the existence of other inhibitory signals that prevent effector T cell influx at the metastatic site.

The cytokine landscape of the metastatic sites shows a general absence of classical immunological effects within the microenvironment. Comparing the data to findings from metastatic colorectal cancer<sup>37</sup> it is clear that there is not much immune activation present in this metastatic NET. MIF, VCAM-1 and ICAM-I appear to be elevated due to the presence of tumor cells and are most likely observed in this context. It has been shown previously, that MIF expression is dramatically increased in hepatic metastases of colorectal cancer patients, and there are numerous studies suggesting a direct role of MIF in tumor pathogenesis and progression in different tumor entities<sup>38</sup>. Furthermore, anti-MIF therapeutics such as neutralizing anti-MIF antibodies have been shown to have a significant effect on tumor growth in colorectal cancer<sup>38</sup>, an option which might also apply to this case of metastatic NET.

Together, our results highlight the broad range and complexity of different levels of tumor-host immune interactions in a case of metastatic NET of the rectum. Using our integrative workflow, we have investigated various immunotherapeutic options and can suggest promising approaches for this patient. Given the tremendous success of immunotherapy, such an integrative approach incorporating many aspects of the tumor-host immune setting, can be a valuable option for complex cases where no standard treatment is available.

## Materials and Methods

### 2.1 Patient sample

---

A 51-year-old Caucasian male was diagnosed with a rectal tumor in 2010. Pathology evaluation revealed a well differentiated neuroendocrine tumor of the rectum. The patient underwent surgery in 2010 where the primary tumor and some liver metastases were resected, and subsequently most of the liver metastases were removed in 2012.

Several OCT embedded fresh-frozen samples from a liver metastasis resected in 2010 (LM1) were analyzed using whole-genome and RNA-Sequencing (WGS and RNA-Seq).

Additionally, OCT embedded fresh-frozen samples from five additional, distinct liver metastases resected in 2012 (LM2-LM6) were analyzed using WGS and RNA-Seq (Table 1).

Whole-blood and healthy liver samples were used as germline control sequences.

Using the patient's blood, PCR-based HLA typing was performed (Table 2).

The patient underwent leukapheresis and 10 mL of the leukapheresis product was used for ELISpot analysis.

### 2.2 Sequence data analysis

---

Sequence read pairs from whole-genome sequencing were mapped and aligned to the 1000 genomes phase 2 reference genome hs37d5 as previously described<sup>39</sup>, using Burrows-Wheeler-Aligner (BWA) (version 0.6.2), and were processed with SAMtools (version 0.1.17) and Picard tools (version 1.61).

Somatic single nucleotide variants (SNVs) were identified in the aligned sequences using an in-house analysis pipeline based on SAMtools mpileup and bcftools, as described previously

<sup>39</sup>.

Short insertions and deletions were called using an in-house pipeline based on platypus (0.5.2)<sup>40</sup>. Since platypus was developed to detect variants in normal genomes, additional custom filters were added to reliably detect somatic indels in tumor normal pairs. These filters integrate the genotype likelihood as well as other filter criteria originally generated by platypus. All calls were annotated with annovar<sup>41</sup> using the gencode reference (v17). All somatic high confidence indels that fall into a coding gene or a splice site were extracted and visually inspected.

RNA sequencing read pairs were mapped to the NCBI human reference genome build 37.2 using Tophat (version 2.0.4). For tumor samples where both DNA and RNA sequencing data were available, candidate DNA variant positions were annotated with RNA information as described previously<sup>42</sup>. Briefly, a candidate DNA variant was called expressed if one high-quality RNA read containing the same variant was present. RNA-Seq data was also used to assess gene expression levels in means of the reads per kilobase per million reads (RPKM) measure.

### **2.3 Epitope Prediction**

---

We used computational methods to predict the immunogenicity of non-synonymous tumor-specific mutations which were detected in the sequencing data. More precisely, HLA class I binding prediction was performed on amino acid peptides containing the non-synonymous mutations, and on peptides containing the corresponding reference residue. As T-cell recognition of peptides is crucially dependent on the ability of the HLA molecule to effectively bind the peptide, comparison of predicted binding affinities of the mutated and reference peptides can be used to assess the immunogenicity of the mutation and to detect neo-epitopes.

For peptide-HLA binding affinity prediction, we chose the NetMHCcons method which was developed at the Technical University of Denmark<sup>43</sup>. NetMHCcons is a consensus method for MHC class I predictions, integrating the three state-of-the-art methods NetMHC, NetMHCpan and PickPocket<sup>44-46</sup>. Depending on how much training data is available for the HLA allele of interest, one of the three methods or a combination of them is used to give the most accurate prediction.

For each non-synonymous mutation, peptides of length 8-11 that contain the mutated residue were extracted from the corresponding protein sequence, and stored in Fasta files to be used as input files for the prediction algorithm. All possible combinations of peptide length and mutation position were considered. Analogously, Fasta files containing the reference peptides were generated. Novel tumor-specific peptides which were generated by frameshift insertions or deletions were also analyzed with NetMHCcons. These so called neoORF antigens are supposed to be highly immunogenic, as they provide longer, completely novel stretches of tumor-specific antigens.

Both, the mutated and reference sequences were submitted to the NetMHCcons 1.0 Server for each HLA allele of the patient, in order to perform allele-specific HLA class I binding predictions. The results of the prediction are IC50 values given in nanomolar (nM) affinity values. Peptides with less than 50 nM are considered as strong-binding, and those with less than 500 nM as weak-binding. Peptides with more than 500 nM are considered as non-binders.

## **2.4 ELISPOT**

---

### **Cell preparation**

10 mL of leukapheresis product was freshly used for the analysis. Ficoll density gradient centrifugation was performed using 50 mL Leucosep tubes (Greiner Bio-One, Kremsmünster, Austria) to isolate PBMC. The median PBMC obtained was  $1 \times 10^8$ . Thereof, T cells and dendritic cells (DCs) were purified as described previously<sup>47, 48</sup>. Briefly, T cells were cultured for 7 days in X-VIVO 20 medium containing 100 U/mL human rIL-2 (Proleukin, Chiron, Ratingen, Germany), and 60 U/mL human rIL-4 (Promokine, PromoCell, Heidelberg). Afterward, cells were kept in cytokine-free medium for 12 h and human CD3 T cells were purified using the Dynabeads untouched human T cell kit (Invitrogen, Darmstadt, Germany).

For DC maturation, adherent cells were cultured for 7 days in X-VIVO 20 medium containing 560 U/ml human rGM-CSF (Leukine, Berlex, Bayer, Leverkusen, Germany), and 500 U/mL human rIL-4. DCs were enriched using anti-CD56 coupled magnetic beads (C218, Beckman Coulter, Krefeld, Germany), and anti-CD3– and anti-CD19– Dynabeads (Invitrogen, Darmstadt, Germany), and pulsed for 18 h with 0.8  $\mu\text{g}/\mu\text{L}$  test peptides or IgG. As positive control, 0.1  $\mu\text{g}/\mu\text{L}$  staphylococcal enterotoxin B (SEB) was used.

## **Peptides**

Peptides were produced by the Peptide Synthesis Facility of the DKFZ. Lyophilized synthetic peptides were solved in distilled water containing 10 % DMSO. Peptide purity was >98 %. Peptides were designed to contain the identified immunogenic HLA restricted T cell epitope. Synthesized human IgG peptides as well as IgG (Kiovig, Baxalta, Unterschleißheim, Germany) were used as negative control antigens.

## **IFN $\gamma$ ELISpot assay**

ELISpot assays were done as described previously<sup>47, 48</sup> with modifications. The assay was carried out in X-VIVO 20 medium, which has been pretested for ELISpot performance in

comparative testing with other media. On day 1, ELISpot plates (MAHA S45, Millipore, Eschborn, Germany) were washed with PBS and coated with 1  $\mu$ L per well of anti- IFN $\gamma$  antibody. The plate was stored overnight at 4°C.

Peptide-pulsed DCs ( $2 \times 10^4$ ) were incubated with autologous T cells ( $1 \times 10^5$ ) at a 1:5 ratio for 40 h in ELISpot plates. All tests were performed in triplicate wells. Internal established operating procedures of an exploratory research laboratory were used for ELISpot testing.

IFN $\gamma$  spots were measured using the automated system CTL ImmunoSpot analyzer (CTLEurope, Bonn, Germany). Each well was subjected to a manual quality control and was reviewed by an independent scientist. Counting parameters were established using the IgG control wells obtaining a high signal-to-noise ratio. Spots induced by the control peptide (human IgG) were considered as background reactivity. A reaction against a test peptide was considered a positive response according to predefined criteria if the spot counts were significantly ( $p < 0.05$ ) higher than the IgG counts according to a permutation test using the difference in means as the test statistic. Raw data is accessible upon request.

## **2.5 Immunohistochemistry and Cytokine Detection**

---

Multiple FFPE blocks of metastatic tissue were immunohistochemically analyzed for their infiltration with T cells (CD3e), B cells (CD20) and NK cells (NKp46), and evaluated as previously described<sup>37</sup>. HLA class I expression was also evaluated by immunohistochemical staining.

Tissue lysates were prepared from frozen material and were used for cytokine detection, as previously described<sup>37</sup>.

## References

1. Dong M, Phan AT, Yao JC. New strategies for advanced neuroendocrine tumors in the era of targeted therapy. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2012; 18:1830-6.
2. Bernick PE, Klimstra DS, Shia J, Minsky B, Saltz L, Shi W, Thaler H, Guillem J, Paty P, Cohen AM, et al. Neuroendocrine carcinomas of the colon and rectum. *Diseases of the colon and rectum* 2004; 47:163-9.
3. Moore JR, Greenwell B, Nuckolls K, Schammel D, Schisler N, Schammel C, Culumovic P, McKinley BP, Trocha SD. Neuroendocrine tumors of the rectum: a 10-year review of management. *The American surgeon* 2011; 77:198-200.
4. Bergsland EK. Introduction: recent advances in the genetics, diagnosis, and treatment of neuroendocrine tumors. *Seminars in oncology* 2013; 40:1-3.
5. Bergsland EK. The evolving landscape of neuroendocrine tumors. *Seminars in oncology* 2013; 40:4-22.
6. Eggenberger JC. Carcinoid and other neuroendocrine tumors of the colon and rectum. *Clinics in colon and rectal surgery* 2011; 24:129-34.
7. Oberg K. The genetics of neuroendocrine tumors. *Seminars in oncology* 2013; 40:37-44.
8. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science* 2015; 348:69-74.
9. Sharma P, Allison JP. The future of immune checkpoint therapy. *Science* 2015; 348:56-61.
10. Aparicio S, Caldas C. The implications of clonal genome evolution for cancer medicine. *The New England journal of medicine* 2013; 368:842-51.
11. Yap TA, Gerlinger M, Futreal PA, Pusztai L, Swanton C. Intratumor heterogeneity: seeing the wood for the trees. *Science translational medicine* 2012; 4:127ps10.
12. Caldas C. Cancer sequencing unravels clonal evolution. *Nature biotechnology* 2012; 30:408-10.
13. Baldus SE, Schaefer KL, Engers R, Hartleb D, Stoecklein NH, Gabbert HE. Prevalence and heterogeneity of KRAS, BRAF, and PIK3CA mutations in primary colorectal adenocarcinomas and their corresponding metastases. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2010; 16:790-9.
14. Huyge V, Garcia C, Alexiou J, Ameye L, Vanderlinden B, Lemort M, Bergmann P, Awada A, Body JJ, Flamen P. Heterogeneity of metabolic response to systemic therapy in metastatic breast cancer patients. *Clinical oncology* 2010; 22:818-27.
15. Kidd EA, Grigsby PW. Intratumoral metabolic heterogeneity of cervical cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2008; 14:5236-41.
16. Folprecht G, Gruenberger T, Bechstein WO, Raab HR, Lordick F, Hartmann JT, Lang H, Frilling A, Stoeckelmacher J, Weitz J, et al. Tumour response and secondary resectability of colorectal liver metastases following neoadjuvant chemotherapy with cetuximab: the CELIM randomised phase 2 trial. *The lancet oncology* 2010; 11:38-47.
17. Whitelock JM, Melrose J, Iozzo RV. Diverse cell signaling events modulated by perlecan. *Biochemistry* 2008; 47:11174-83.
18. Euskirchen G, Auerbach RK, Snyder M. SWI/SNF chromatin-remodeling factors: multiscale analyses and diverse functions. *The Journal of biological chemistry* 2012; 287:30897-905.
19. Ye Y, Xiao Y, Wang W, Gao JX, Yearsley K, Yan Q, Barsky SH. Singular v dual inhibition of SNF2L and its isoform, SNF2LT, have similar effects on DNA damage but opposite effects on the DNA damage response, cancer cell growth arrest and apoptosis. *Oncotarget* 2012; 3:475-89.
20. Huang J, El-Gamil M, Dudley ME, Li YF, Rosenberg SA, Robbins PF. T cells associated with tumor regression recognize frameshifted products of the CDKN2A tumor suppressor gene locus and a mutated HLA class I gene product. *Journal of immunology* 2004; 172:6057-64.
21. Linnebacher M, Gebert J, Rudy W, Woerner S, Yuan YP, Bork P, von Knebel Doeberitz M. Frameshift peptide-derived T-cell epitopes: a source of novel tumor-specific antigens. *International journal of cancer Journal international du cancer* 2001; 93:6-11.

22. Saeterdal I, Bjorheim J, Lislrud K, Gjertsen MK, Bukholm IK, Olsen OC, Nesland JM, Eriksen JA, Moller M, Lindblom A, et al. Frameshift-mutation-derived peptides as tumor-specific antigens in inherited and spontaneous colorectal cancer. *Proceedings of the National Academy of Sciences of the United States of America* 2001; 98:13255-60.
23. Hacohen N, Fritsch EF, Carter TA, Lander ES, Wu CJ. Getting personal with neoantigen-based therapeutic cancer vaccines. *Cancer immunology research* 2013; 1:11-5.
24. Heemskerk B, Kvistborg P, Schumacher TN. The cancer antigenome. *The EMBO journal* 2013; 32:194-203.
25. Castle JC, Kreiter S, Diekmann J, Lower M, van de Roemer N, de Graaf J, Selmi A, Diken M, Boegel S, Paret C, et al. Exploiting the mutanome for tumor vaccination. *Cancer research* 2012; 72:1081-91.
26. Matsushita H, Vesely MD, Koboldt DC, Rickert CG, Uppaluri R, Magrini VJ, Arthur CD, White JM, Chen YS, Shea LK, et al. Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* 2012; 482:400-4.
27. Duan F, Duitama J, Al Seesi S, Ayres CM, Corcelli SA, Pawashe AP, Blanchard T, McMahon D, Sidney J, Sette A, et al. Genomic and bioinformatic profiling of mutational neoepitopes reveals new rules to predict anticancer immunogenicity. *The Journal of experimental medicine* 2014; 211:2231-48.
28. Gubin MM, Zhang X, Schuster H, Caron E, Ward JP, Noguchi T, Ivanova Y, Hundal J, Arthur CD, Krebber WJ, et al. Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* 2014; 515:577-81.
29. Yadav M, Jhunjhunwala S, Phung QT, Lupardus P, Tanguay J, Bumbaca S, Franci C, Cheung TK, Fritsche J, Weinschenk T, et al. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* 2014; 515:572-6.
30. Robbins PF, Lu YC, El-Gamil M, Li YF, Gross C, Gartner J, Lin JC, Teer JK, Cliften P, Tycksen E, et al. Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells. *Nature medicine* 2013; 19:747-52.
31. van Rooij N, van Buuren MM, Philips D, Velds A, Toebe M, Heemskerk B, van Dijk LJ, Behjati S, Hilkmann H, El Atmioui D, et al. Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2013; 31:e439-42.
32. van Buuren MM, Calis JJ, Schumacher TN. High sensitivity of cancer exome-based CD8 T cell neo-antigen identification. *Oncoimmunology* 2014; 3:e28836.
33. Edward F. Fritsch, Mohini Rajasagi, Patrick A. Ott, Vladimir Brusic, Hacohen N, Wu CJ, . HLA-Binding Properties of Tumor Neoepitopes in Humans. *Cancer immunology research* 2014; 2:1-8.
34. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, Chmielowski B, Spasic M, Henry G, Ciobanu V, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 2014; 515:568-71.
35. Halama N, Michel S, Kloor M, Zoernig I, Benner A, Spille A, Pommerencke T, von Knebel DM, Folprecht G, Lubber B, et al. Localization and density of immune cells in the invasive margin of human colorectal cancer liver metastases are prognostic for response to chemotherapy. *Cancer research* 2011; 71:5670-7.
36. Halama N, Michel S, Kloor M, Zoernig I, Pommerencke T, von Knebel Doeberitz M, Schirmacher P, Weitz J, Grabe N, Jager D. The localization and density of immune cells in primary tumors of human metastatic colorectal cancer shows an association with response to chemotherapy. *Cancer immunity* 2009; 9:1.
37. Halama N, Braun M, Kahlert C, Spille A, Quack C, Rahbari N, Koch M, Weitz J, Kloor M, Zoernig I, et al. Natural killer cells are scarce in colorectal carcinoma tissue despite high levels of chemokines and cytokines. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2011; 17:678-89.
38. He XX, Chen K, Yang J, Li XY, Gan HY, Liu CY, Coleman TR, Al-Abed Y. Macrophage migration inhibitory factor promotes colorectal cancer. *Molecular medicine* 2009; 15:1-10.

39. Jones DT, Hutter B, Jager N, Korshunov A, Kool M, Warnatz HJ, Zichner T, Lambert SR, Ryzhova M, Quang DA, et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nature genetics* 2013; 45:927-32.
40. Rimmer A, Phan H, Mathieson I, Iqbal Z, Twigg SR, Consortium WGS, Wilkie AO, McVean G, Lunter G. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nature genetics* 2014; 46:912-8.
41. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research* 2010; 38:e164.
42. Jones DT, Jager N, Kool M, Zichner T, Hutter B, Sultan M, Cho YJ, Pugh TJ, Hovestadt V, Stutz AM, et al. Dissecting the genomic complexity underlying medulloblastoma. *Nature* 2012; 488:100-5.
43. Karosiene E, Lundegaard C, Lund O, Nielsen M. NetMHCcons: a consensus method for the major histocompatibility complex class I predictions. *Immunogenetics* 2012; 64:177-86.
44. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. *Nucleic acids research* 2008; 36:W509-12.
45. Zhang H, Lund O, Nielsen M. The PickPocket method for predicting binding specificities for receptors based on receptor pocket similarities: application to MHC-peptide binding. *Bioinformatics* 2009; 25:1293-9.
46. Zhang H, Lundegaard C, Nielsen M. Pan-specific MHC class I predictors: a benchmark of HLA class I pan-specific prediction methods. *Bioinformatics* 2009; 25:83-9.
47. Bonertz A, Weitz J, Pietsch DH, Rahbari NN, Schlude C, Ge Y, Juenger S, Vlodaysky I, Khazaie K, Jaeger D, et al. Antigen-specific Tregs control T cell responses against a limited repertoire of tumor antigens in patients with colorectal carcinoma. *The Journal of clinical investigation* 2009; 119:3311-21.
48. Horn T, Grab J, Schusdziarra J, Schmid S, Maurer T, Nawroth R, Wolf P, Pritsch M, Gschwend JE, Kubler HR, et al. Antitumor T cell responses in bladder cancer are directed against a limited set of antigens and are modulated by regulatory T cells and routine treatment approaches. *International journal of cancer Journal international du cancer* 2013; 133:2145-56.

### Figure 1

**Mutational landscape of all analyzed liver metastases.** Whole-Genome-Sequencing and Whole-Transcriptome-Sequencing was performed on six different liver metastases as well as on healthy liver tissue and whole-blood of the same patient. Single nucleotide variations (SNVs) as well as indels (short insertions or deletions) were detected in each of the sequences. A grey cell indicates a somatic mutation in the corresponding gene and metastasis, an asterisk indicates an indel.

### Figure 2

**Overview of IFN- $\gamma$  ELISpot data for all tested neo-epitopes and controls with patient-derived cells.** Peripheral blood dendritic cells (DCs) were pulsed with mutated and corresponding wildtype peptides as well as negative control antigen (human IgG), and incubated with autologous T cells. ELISpot assays were performed in triplicates.

(A) Triplicate wells from IFN- $\gamma$  ELISpot analysis of all tested peptides.

(B) Summary of ELISpot data showing mean spot number + SD for each tested peptide.

An asterisk indicates significantly higher spot numbers in test wells compared with spot numbers of negative control antigen.

### Figure 3

**Cytokine Profile.** Cytokines concentrations were measured in tissue lysates of a liver metastasis.

Figure 1

	LM 1	LM 2	LM 3	LM 4	LM 5	LM 6
SLC11A2						
HSPG2						
CCSER1						
BAZA						
DET1						
SERPINF2						
RABEP1						
SMARCA1						
MAT2A						
ADAM9						
OBSL1*						
PRKDC*						
ENDOG						
POLR1B*						
SPAG6						
CDKN1C						
SF3B3						
ZC3H1B						
RMBP2*						

\* Indel

Figure 2

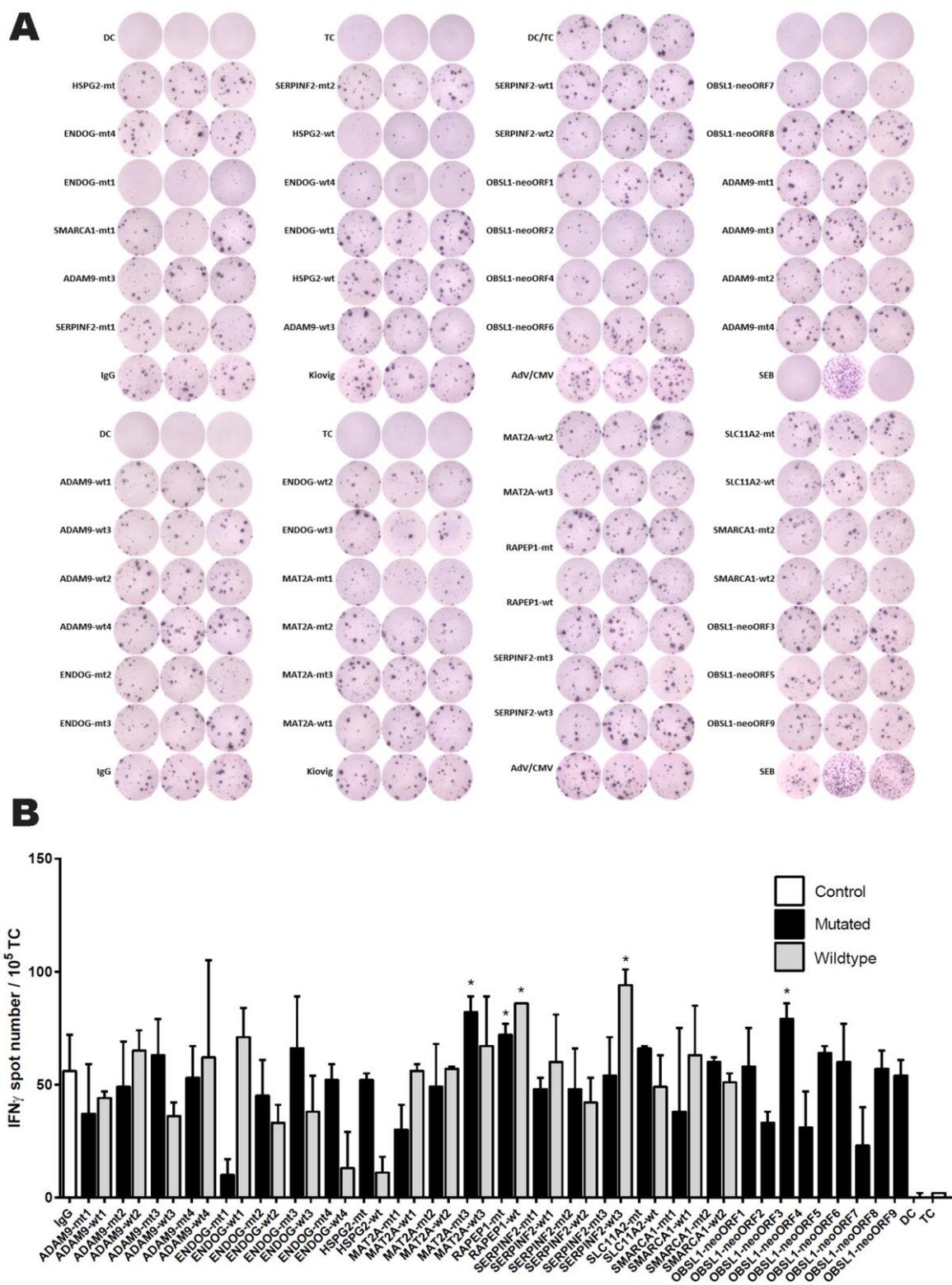




Table 1 Overview of analyzed samples

Sample	Year of surgery	Information	Detected mutations
LM1	2010	NA	14
LM2	2012	piece of tumor from a 2cm lesion from posterior segment 2, left lobe	13
LM3	2012	piece of tumor from a 2cm lesion from the left lobe of the liver, segment 4	14
LM4	2012	piece of tumor from a 3cm lesion from the right lobe of the liver, segment 4	13
LM5	2012	piece of tumor from a 1cm lesion from the right lobe of the liver (i.e. close to the edge of the liver).	15
LM6	2012	piece of tumor from the 8cm major lesion from the right lobe of the liver (i.e. close to the middle of the liver) note: this lesion was largely hemorrhagic	14

Table 2 HLA genotype of the patient

<b>HLA</b>	<b>Allele 1</b>	<b>Allele 2</b>
HLA-A	*03:02	*33:01
HLA-B	*14:02	*55:01
HLA-C	*03:03	*08:02

Table 3 Overview of detected mutations

Gene	Chr	Pos	Ref	Alt	classification	#of affected samples	Affected Samples	Epitopes Predicted
SLC11A2	12	51382161	T	C	nonsynonymous SNV	6	LM1;LM2;LM3;LM4;LM5;LM6	1
HSPG2	1	22150143	G	T	nonsynonymous SNV	6	LM1;LM2;LM3;LM4;LM5;LM6	1
CCSER1	4	92519846	A	G	nonsynonymous SNV	6	LM1;LM2;LM3;LM4;LM5;LM6	0
BAZ2A	12	57003963	C	G	nonsynonymous SNV	6	LM1;LM2;LM3;LM4;LM5;LM6	0
DET1	15	89074654	C	A	nonsynonymous SNV	6	LM1;LM2;LM3;LM4;LM5;LM6	0
SERPINF2	17	1655979	G	A	nonsynonymous SNV	6	LM1;LM2;LM3;LM4;LM5;LM6	4
RABEP1	17	5264636	G	T	nonsynonymous SNV	6	LM1;LM2;LM3;LM4;LM5;LM6	1
SMARCA1	X	128599617	C	G	nonsynonymous SNV	6	LM1;LM2;LM3;LM4;LM5;LM6	2
MAT2A	2	85770808	G	T	nonsynonymous SNV	6	LM1;LM2;LM3;LM4;LM5;LM6	3
ADAM9	8	38880677	G	A	nonsynonymous SNV	6	LM1;LM2;LM3;LM4;LM5;LM6	5
OBSL1	2	220416342	GC	G	frameshift deletion	6	LM1;LM2;LM3;LM4;LM5;LM6	9
PRKDC	8	48701554	G	GT	frameshift insertion	5	LM1;LM2;LM3;LM4;LM6	0
ENDOG	9	131581121	C	T	nonsynonymous SNV	5	LM1;LM3;LM4;LM5;LM6	5
POLR1B	2	113333084	A	AGATCG	frameshift insertion	2	LM1;LM6	NA
SPAG6	10	22705558	C	T	nonsynonymous SNV	1	LM2	NA
CDKN1C	11	2905268	G	A	nonsynonymous SNV	1	LM3	NA
SF3B3	16	70582330	C	T	nonsynonymous SNV	1	LM5	NA
ZC3H18	16	88690471	G	A	splicing	1	LM5	NA
RIMBP2	12	130926721	GC	G	frameshift deletion	1	LM5	NA

Table 4 Overview of predicted neo-epitopes

Gene: Mutation	Mutated		Wildtype		HLA	Gene expression fold change
	Peptide ID	Affinity (mt)	Peptide ID	Affinity (wt)		
ADAM9: M249I	ADAM9-mt1	37.92	ADAM9-wt1	33630.31	A33	2.21
	ADAM9-mt2	16.3	ADAM9-wt2	14.56	C08	
	ADAM9-mt3	461.65	ADAM9-wt3	13.56	C03	
		12.61		13.56	C08	
ADAM9-mt4	37.81	ADAM9-wt4	37.61	C08		
ENDOG: P53L	ENDOG-mt1	311.03	ENDOG-wt1	25109.54	C03	-0.16
	ENDOG-mt2	479.1	ENDOG-wt2	663.31	C08	
	ENDOG-mt3	377.91	ENDOG-wt3	28319.21	A33	
	ENDOG-mt4	179.12	ENDOG-wt4	19563.17	C03	
HSPG2: N4323K	HSPG2-mt	443.88	HSPG2-wt	36429.09	A03	-0,1
MAT2A K367N	MAT2A-mt1	261.59	MAT2A-wt1	39619.11	A33	1.44
	MAT2A-mt2	18.67	MAT2A-wt2	44256.9	A33	
	MAT2A-mt3	270.22	MAT2A-wt3	47740.48	A33	
RABEP1: S410I	RAPEP1-mt	49.63	RAPEP1-wt	47.75	C08	0.59
SERPINF2: D204N	SERPINF2-mt1	279.13	SERPINF2-wt1	533.92	C03	-2.48
		338.95		533.92	C08	
	SERPINF2-mt2	412.08	SERPINF2-wt2	843.61	C03	