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
Msc-Biology Mine Özcan
born in: Konak, Turkey
Oral-examination:07.11.17

Models for Immune Response and Immune Evasion in MSI Cancer and Lynch Syndrome

Referees: Prof. Dr. Magnus von Knebel Doeberitz

Prof. Dr. Hans-Reimer Rodewald

ACKNOWLEDGEMENTS

First of all, I would like to thank Prof. Dr. Magnus von Knebel Doeberitz for providing me the opportunity to do my PhD in this lab and all other members of Department of Applied Tumor Biology for their support. In particular I would like to thank PD. Dr. Matthias Kloor for supervising me during my doctoral study, providing me insight into MSI field and answering my questions indulgently.

I wish to acknowledge Dr. Johannes Gebert for responding my questions when I need help and effective scientific discussions.

I would like to offer my special thanks to Andrea Klingmann for her graciousness and helping me with organizational issues, Lena Ehret-Maßholder and Petra Hoefler for being friendly and helping me.

I am very grateful to Prof. Dr. Hans-Reimer Rodewald for agreeing to be my TAC member and my second reviewer; for all kinds of questions and always stimulating the fruitful scientific discussions.

I also would like to thank all of my friends especially Nese Erdem for being there for me any time. And I would like to thank to all of my labmates in both dark side Katharina Urban, Lennard Ganß, Simon Kalteis and bright side Eva-Maria Katzenmeier, Malwina Michalak, Ana-Ligia Gutierrez and Fabia Fricke for their friendship and support throughout studies and more importantly making the time that I spent in the lab more enjoyable.

Last but not the least, I would like to express my gratitude to my family; my mother, Emine Özcan; my father Ibrahim Özcan and my dear sister Lale Özcan for supporting and believing in me all the time and encouraging me to accomplish my goals. I am also highly grateful to Sascha Wahlbrink for being patient and supporting and motivating me all the time.

SUMMARY

Microsatellite-unstable (MSI) cancers occurring in the context of the hereditary Lynch syndrome or as sporadic cancers elicit pronounced tumor-specific immune responses. The pronounced immune response was shown to be closely associated with frameshift peptides (FSP) that are generated as a result of deficiency in DNA mismatch repair system leading to insertion/deletion mutations in coding microsatellites (cMS). FSP neoantigens are long antigenic amino acid stretches that bear multiple epitopes to be presented. There is no central tolerance against FSPs, and shared FSPs derived from driver mutations are promising candidates for vaccination approaches to treat or prevent MSI cancers. In the present thesis, the main goals were to set up a mouse model for the immunology of MSI cancers and to systematically identify immune evasion mechanisms in MSI cancers.

A murine model is essential to characterize alterations of immune responses over time, in all stages of cancer and pre-cancerous stages. In addition, it allows testing an FSP vaccine for efficacy in tumor prevention and treatment, either as a single agent or on combination with other immune-modulatory drugs. To establish such a model, the complete mouse genome was screened and genes bearing cMS were detected. After mutation and expression analysis by using murine Lynch tumors, epitopes of the most promising potential FSP candidates were predicted by using the Syfpeithi and netMHC algorithms. Immunogenicity of the 10 FSPs with the highest ranks was analyzed by vaccinating C57BL/6 mice and analyzing immune responses using IFNg ELISpot. Four FSPs were identified that were highly immunogenic and inducing spot numbers higher than Ova control peptides: Maz (-1) and Senp6 (-1) induced only CD4 T cell responses, Xirp1 (-1) induced only CD8 T cell and Nacad (-1) induced both CD4 and CD8 T cell responses. Peptide-specific IgG Elisa demonstrated that three of the peptides Senp6 (-1), Maz (-1) and Nacad (-1) also induced humoral immune response. Immunogenic regions of the peptides could be mapped to the C-terminus of Senp6 (-1) and Xirp (-1) and to the N-terminus of Nacad (-1), whereas the antigenic region for Maz (-1) spanned almost the entire peptide. These results suggest that the Lynch mouse model is well suitable for evaluating the efficacy of FSP vaccination to treat and even prevent tumors in Lynch syndrome.

The second aim of this thesis was to systematically analyze immune evasion mechanisms in MSI cancer. We first analyzed mutations of genes related to MHC class I antigen presentation in publicly accessible mutation databases. The mutation data of 91 MSI patients in the DFCI cohort showed that 72% of all MSI CRC tumors had defects in MHC class I presentation; displayed by at least one mutation in the corresponding genes (*B2M*, *TAP1*, *TAP2*, *HLA-A*, *HLA-B*, *HLA-C* and *NLRC5*). Mutual exclusivity analysis revealed that mutations affecting *B2M* were negatively related to *HLA-B* mutations, whereas there was a strong positive correlation between HLA class I heavy chain mutations. These results indicate that there is a strong immunoselection in MSI tumorigenesis, leading to immune evasion through mutations of MHC class I-related genes in more than two-thirds of MSI cancers. We identified *NLRC5* mutations as a potential novel immune evasion mechanism in the database analysis; therefore, potential consequences of NLRC5 inactivation were further analyzed in MSI colorectal cancer samples. We detected cMS mutations of *NLRC5* in 4 out of 95 tumor samples (4.2%), three of them being one-basepair deletions and one silent mutation. Importantly, we detected low levels of MHC class I antigen expression in NLRC5-mutated tumors. One tumor showed partial reduction of MHC class I expression, which colocalized with the *NLRC5* mutation. These results suggest *NLRC5* mutations as a novel potential mechanism of immune evasion in MSI cancer.

Taken together, the present thesis led to the establishment of the first model to evaluate the immune biology of MSI cancers and Lynch syndrome in the murine system. Moreover, it has established a comprehensive overview of immune evasion in MSI cancers, thus contributing to the development of better treatment strategies and potentially to the first cancer-preventive vaccine for non-viral human cancers.

ZUSAMMENFASSUNG

Mikrosatelliten-instabile (MSI) Tumoren, die im Kontext des hereditären Lynch-Syndroms oder als sporadische Tumoren entstehen,lösen ausgeprägte Tumor-spezifische Immunantworten aus. Diese Immunantworten sind eng mit Frameshift-Peptiden assoziiert, die als Ergebnis einer DNA-Mismatch-Reparaturdefizienz und in der Folge Insertions-/Deletions-Mutationen an kodierenden Mikrosatelliten (cMS) entstehen. Frameshift-Neoantigene sind lange antigene Aminosäuresequenzen; sie tragen multiple Epitope, die dem Immunsystem gegenüber präsentiert werden können. Gegen FSPs gibt es keine zentrale Toleranz, und FSPs, die von Driver-Mutationen herrühren, sind vielversprechende Kandidaten für Vakzinierungsansätze zur Behandlung oder Prävention von MSI-Tumoren, da sie in der Mehrzahl dieser Tumoren auftreten (shared antigens). Ziele der vorliegenden Arbeit waren die Etablierung eines Maus-Modells für die Immunologie von MSI-Tumoren und die systematische Charakterisierung von Immunevasionsmechanismen in MSI-Tumoren. Ein murines Modell ist essenziell, um Veränderungen der Immunantwort über die Zeit hinweg und im Zusammenhang mit dem Auftreten mit Karzinomen und präkanzerösen Läsionen zu studieren. Zusätzlich erlaubt es ein solches Modell, eine FSP-Vakzine auf ihre Wirksamkeit in der Tumorprävention zu testen, entweder als alleinige Behandlung oder in Kombination mit anderen immunmodulatorischen Wirkstoffen. Um ein solches Mausmodell zu etablieren, wurde zunächst das komplette murine Genom auf Gene hin untersucht, die cMS in ihrer Sequenz enthalten. Nach der Mutations- und Expressionsanalyse in murinen Lynch-Tumoren wurde mit Hilfe der Algorithmen SYFPEITHI und netMHC von den vielversprechendsten Kandidaten Epitopvorhersagen erstellt. Die Immunogenität der 10 FSPs mit den besten Ergebnissen wurde dadurch geprüft, dass C57BL/6-Mäuse mit den Peptiden vakziniert und die Immunantworten über IFN-gamma-ELISpots untersucht wurden. Es wurden vier FSPs identifiziert, die hoch immunogen waren und Spot-Zahlen aufwiesen, welche höher waren als die Ova-Kontroll-Peptide: Maz (-1) and Senp6 (-1) induzierten ausschließlich CD4-T-Zell-Antworten, Xirp1 (-1) induzierte nur CD8-T-Zell-Antworten, und Nacad (-1) induzierte sowohl CD4- als auch CD8-T-Zell-Antworten. Peptid-spezifische IgG-ELISA-Untersuchungen zeigten, dass drei der Peptide, Senp6 (-1), Maz (-1) und Nacad (-1), auch humorale Immunantworten induzierten. Die immunogenen Regionen der Peptide

konnten für Senp6 (-1) und Xirp (-1) dem C-Terminus zugeordnet werden, während Nacad (-1) am N-Terminus immunogen war, und die antigenen Regionen sich bei Maz (-1) über beide Bereiche des Peptids erstreckten. Diese Ergebnisse zeigen, dass das Lynch-Mausmodell vielversprechend ist, um die Effektivität einer FSP-Vakzine zur Behandlung und Prävention von Tumoren beim Lynch-Syndrom zu evaluieren.
Das zweite Ziel der Arbeit war es, Immunevasionsmechanismen bei MSI-Tumoren systematisch zu untersuchen. Zuerst haben wir anhand öffentlich zugänglicher Datenbanken Mutationen in Genen untersucht, die mit der MHC I-Antigenpräsentation in Zusammenhang stehen. Es wurden Mutationsdaten von 91 MSI-Tumoren aus der DFCI-Kohorte analysiert. Dabei zeigte sich, dass 72% aller kolorektalen MSI-Karzinome Veränderungen aufwiesen, die potenziell zu einem Defekt in der MHC I-Antigenpräsentation führen können. Diese Veränderungen betrafen eines oder mehrere der folgenden Gene: *B2M*, *TAP1*, *TAP2*, *HLA-A*, *HLA-B*, *HLA-C* and *NLRC5*. Die Analyse auf gegenseitige Exklusivität der Mutationen zeigte, dass *B2M*-Mutationen negativ korreliert waren mit *HLA-B*-Mutationen. Andererseits gab es eine starke positive Korrelation unter Mutationen der schweren Ketten-Gene von MHC I. Diese Ergebnisse zeigen, dass die MSI-Tumorigenese einer starken Immunselektion unterliegt. Diese Selektion führt in mehr als zwei Dritteln der MSI-Tumoren zur Ausbildung von Immunevasionsphänomenen. Wir haben Mutationen des MHC I-Transaktivators NLRC5 as potenziell neuen Immunevasionsmechanismus in den Datenbanken identifiziert. Daher wurden im Folgenden potenzielle Konsequenzen der NLRC5-Inaktivierung weiter an Gewebeproben von kolorektalen MSI-Tumoren untersucht. Wir konnten in 4 von 95 Tumoren (4.2%) Mutationen nachweisen, wobei drei der Mutationen Ein-Basenpaar-Deletionen waren, eine Mutation war eine silente Mutation.
Alle *NLRC5*-mutierten Tumoren zeigten niedrige Expressionsniveaus von MHC I-Ketten. Ein Tumor wies eine lokale Reduktion von MHC I auf, der mit der *NLRC5*-Mutation kolokalisiert war. Diese Ergebnisse weisen darauf hin, dass *NLRC5*-Mutationen einen neuen Immunevasionsmechanismus in MSI-Tumoren darstellen könnten.
Zusammenfassend hat die vorliegende Arbeit zur Etablierung eines ersten Modells geführt, das es erlaubt, die Immunbiologie von MSI-Tumoren und des Lynch-Syndroms im murinen System zu testen. Weiterhin hat die Arbeit einen umfassenden Überblick über Immunevasionsmechanismen gegeben. Diese Ergebnisse könnten auf dem Weg zu

besseren Behandlungsstrategien bei MSI-Tumoren und möglicherweise zur ersten präventiven Vakzine gegen nicht-virale Tumoren beim Menschen einen wichtigen Beitrag leisten.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

APM	Antigen presentation machinery
B2M	β2-microglobulin
cDNR	Coding dinucleotide repeat
CIMP	CpG island methylator phenotype
CLR	Crohn's like reaction
cMNR	Coding mononucleotide repeats
cMS	Coding microsatellite
CRC	Colorectal cancer
ELISA	Enzyme-linked immunosorbent assay
Elispot	Enzyme linked immuno spot assay
ER	Endoplasmic reticulum
FACS	Fluorescence-activated cell sorting
FBS	Fetal Bovine Serum
FFPE	Formalin-fixed paraffin embedded
FSP	Frameshift peptides
HLA	Human leukocyte antigen
HRP	Horseradish Peroxidase
IgG	Immunoglobulin G
IHC	Immunohistochemistry
MMR	Mismatch repair system
MSI	Microsatellite instability
NLRC5	NOD-like receptor family CARD domain containing 5
PBS	Phosphate buffered saline
RPMI 1640	Roswell Park Memorial Institute 1640 (cell culture medium)
TAP	Transporter associated with antigen processing
TGFBR2	Transforming growth factor beta-receptor 2

1 INTRODUCTION

1.1 Colorectal Cancer

Colorectal cancer (CRC) is the third most common cancer. Reduction of risk factors such as less red meat consumption, advanced screening tests and most importantly current treatment opportunities have lessened the incidence and mortality rate of CRC for decades [1]. However, it still remains the third leading cause of cancer death, illustrating the urgent need for improved treatment and prevention strategies.
In order to develop such strategies, the characterization of subtypes of CRC is of utmost importance. Colon cancer is a heterogeneous tumor type; different subtypes of CRC are described based on genomic and epi-genomic instability of the cancer[2]. Most of CRCs exhibit chromosomal instability and pursue classical adenoma-carcinoma tumor progression. However, around 15% of CRCs are DNA mismatch repair (MMR) deficient and exhibit high microsatellite instability (MSI) phenotype [3].

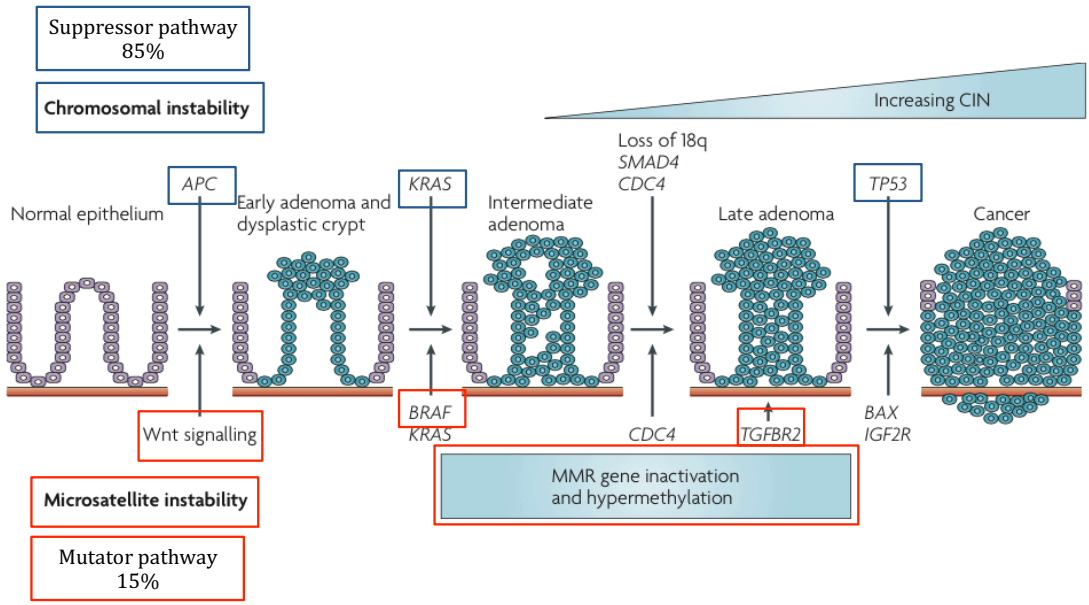


Figure 1 MSI-H CRC tumorigenesis pathway is distinctive compared to chromosomally instable cancers. 85% of all colorectal cancers follow suppressor pathway by chromosomal instability; however, 15% of all CRCs follow a mutator pathway marked by microsatellite instability. Chromosomal instability pathway is marked as blue and microsatellite-unstable pathway is marked with red. Figure adapted from [4].

1.2 Microsatellite Instability

The MMR system is responsible for detection and repair of incorrectly matched bases during DNA replication[5]. Defects in the MMR system result in the accumulation of insertion/deletion mutations in particular in highly repetitive DNA sequences called microsatellites. This in turn leads to MSI and tumor formation [6].

MMR deficient tumors occur either in the context of Lynch syndrome, i.e. in carriers of a germline mutation in one of the MMR genes, or as sporadic MSI cancer. A complex set of heterodimeric proteins; namely MutSα (MSH2 and MSH6), MutSβ (MSH2 and MSH3) and MutLα (MLH1/PMS2) constitute the MMR system, and inactivation of any of the mentioned genes results in MMR deficiency. MSI is scored based on US National Cancer Institute recommended panels of microsatellite markers; if two or more of five markers display mutations; then tumor is called MSI [7,8].

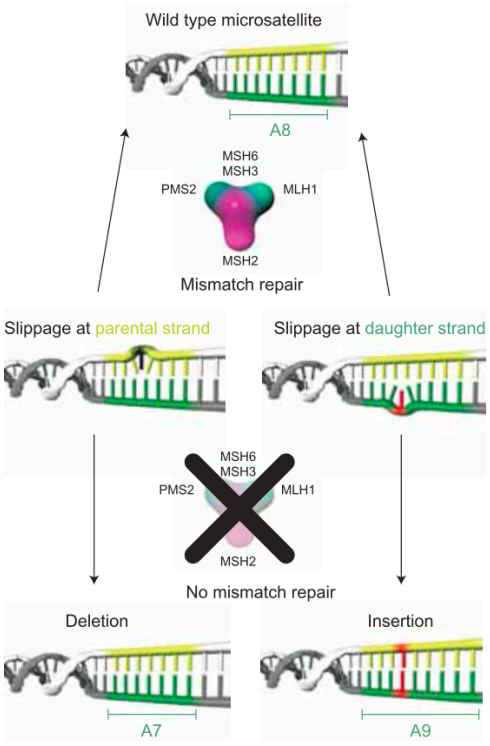


Figure 2 **Molecular mechanisms of microsatellite instability.** Defects in DNA MMR genes result in either slippage at the parental strand and therefore deletion mutations or it results in slippage at daughter strand, thus insertion mutations [6].

In 12 % of all CRCs, epigenetic alterations called CpG island methylator phenotype

(CIMP) cause hypermethylation of *MLH1* promoter and silencing of *MLH1* gene and lead to sporadic MSI cancer. On the other hand, 3% of all CRCs carry monoallelic germline mutations in MMR genes, which together with a second somatic hit, result in Lynch syndrome-associated MMR deficient cancer. Lynch syndrome patients bearing *MLH1* and *MSH2* mutations are more prone to develop MSI CRC when compared to *MSH6* or *PMS2* mutation carriers[9]. Despite Lynch syndrome patients have the risk of developing cancer during their lifetime, some patients never progress to cancer and it was reported that the risk of developing cancer for Lynch syndrome is around 30 to 80% [10–12]. The reason why not all patients develop cancer may be explained by the enhanced immune response that is marked by the high infiltration of immune cells in the tumor in Lynch syndrome patients.

1.3 Characteristics of immune response in MSI cancers

MSI cancers are discerned from MSS tumors by many different characteristics. MSI tumors are predominantly located in right-sided colon, and there may be two or more primary tumors identified in the same patient at the same time or multiple tumors may arise subsequently [13]. In addition, MSI tumors are poorly differentiated and have a mixed differentiation pattern with mucinous, medullar carcinoma and signet ring cells [14]; and tissue invasion is expansive and cohesive in contrast to infiltrative and dissociative invasion that is seen in chromosomally instable cancers [10]. Furthermore, MSI CRC was shown to be closely related with local lymphocyte infiltration with lymphocyte aggregations called Crohn's like reaction (CLR) and low frequency of distant metastases with a good prognosis [15].

According to Smyrk and colleagues, microsatellite-unstable tumors have significantly increased lymphocyte infiltration when compared to MSS tumors [16]. The prognostic impact of tumor infiltrating lymphocytes, which is also evident beyond MSI CRC [17], in particular populations and functions of the cells, were further elaborated in detail in the other studies.

Edin and colleagues in 2012 reported that MSI cases with high infiltration of macrophages have more favorable prognosis [18]. In another study by Prall and colleagues, 11 MSI patients with high infiltration of CD8 T cells displayed an exceptional prognosis, with 9 out of 11 having tumor-free survival [19]. Furthermore,

immunostaining in MSI tumors revealed that MSI tumors have significantly higher numbers of tumor infiltrating activated $CD3^+$ and $CD8^+$ cytotoxic cells expressing high levels of granzyme B and higher percentage of apoptotic tumor cells that are mostly located close to the activated cytotoxic lymphocytes when compared to MSS tumors [20]. Moreover, in another study, while the MSI phenotype was associated with the cytotoxic markers, namely higher expression of granzyme B and perforin, MSS phenotype was correlated with higher Foxp3, IL17 and TGFβ expression [21].

1.4 The role of mutations in cMS in MSI cancer

Highly repetitive DNA sequences, called microsatellites (MS) can be distributed in the genome in intergeneic, intronic, transcribed-nontranslated, or gene encoding regions (figure 3) [10]. Except from the latter, the mutations in those regions are neutral; however, when there is a mutation in gene-coding microsatellite (cMS), then the mutation can potentially induce cancerous transformation of MMR-deficient cells that provides selective growth advantage for that specific cell clone. Mutations in cMSs follow a Darwinian selection pattern; the genes that provide advantages in growth and survival called driver mutations are preserved during the course of cancer progression (typically mutations of tumor suppressor genes), whereas, mutations that do not provide any selective advantage (passenger mutation) or mutations that are growth repressing are lost [22–24].

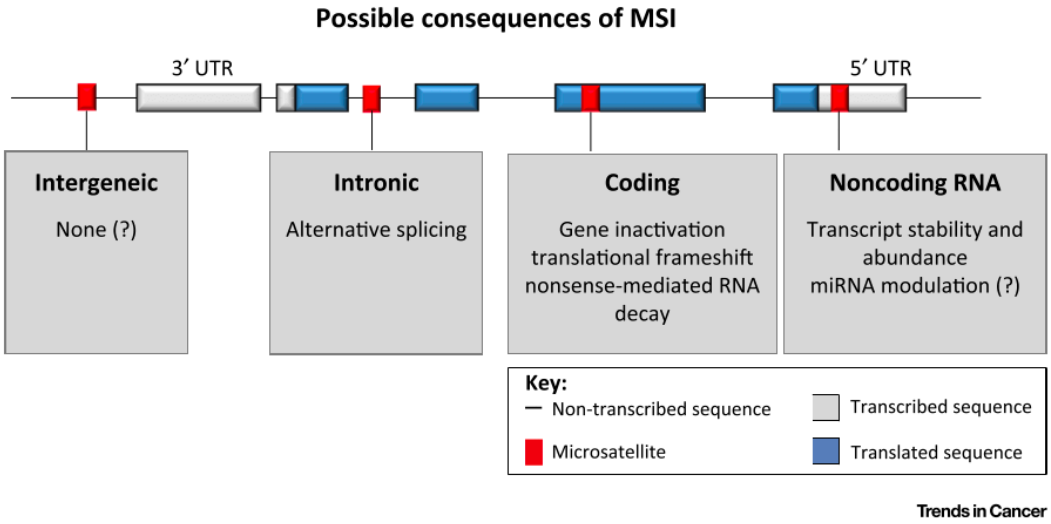


Figure 3 **Distribution of microsatellites in the human genome.** Microsatellites can be located in the intergeneic, intronic, coding or noncoding parts in the genome. While mutations of microsatellites in the noncoding regions namely; intergeneic, intronic and noncoding locations result in neutral consequences; mutations in coding microsatellites change the translational reading frame and result in generation of frameshift peptides.

As a result of insertion and deletion mutations in cMSs and the subsequent shifts in the translational reading frame, neo-antigens called frameshift peptides (FSP) are generated. One of the essential reasons of high TIL infiltration in MSI CRC compared to MSS CRC has been speculated to be due to the generation of FSPs that are recognized by cytotoxic lymphocytes. This hypothesis is supported by the data that declares that the density of tumor infiltrating lymphocytes; specifically $CD8^+$ T cells are positively associated with the total number of frameshift mutations in the tumor in MSI CRCs [25,26]. Moreover, it has been shown that TILs isolated from MSI CRC recognize FSP neoantigens[27].

FSPs are long, highly immunogenic antigenic stretches that bear multiple potential epitopes to be presented. Besides, since there is no tolerance against these neoantigens [28] and the fact that they are completely novel to the immune system of the host, FSPs are truly specific for MMR-deficient tumors and cells that have the potential of developing into tumors.

Additionally, a subset of cMS that can give rise to FSPs are located in tumor suppressor genes [24]. One of the most prominent and well known example of a pertinent cMS is A10 repeat in the coding region of the transforming growth factor beta receptor 2(*TGFBR2*) gene; in 80% of MSI CRCs, cMS of *TGFBR2* gene is mutated [29]. Therefore, particularly such driver mutation-induced FSPs are ideal targets for a preventive and therapeutic vaccination approach for Lynch syndrome and all types of MSI cancers, respectively [10].

Initial studies that showed a pronounced immune response against FSPs [30,31] were further supported by Phase I/IIa clinical trial that has been completed (https://clinicaltrials.gov/show/NCT01461148), demonstrating that a FSP vaccine consisting of three FSPs is safe and induces FSP-specific T cell responses in all patients vaccinated per protocol [10].

1.5 Novel immune therapies in MSI cancer patients

The pronounced immune responses against MSI cancers can lead to immune cell exhaustion [32].It was reported that active tumor microenvironment that is marked by Th1 immune response and cytotoxic T cells, is counterbalanced by up-regulation of

immune checkpoints such as PD-1, PD-L1 and CTLA-4 [32]. Alternatively, tumor cell-associated immune evasion mechanisms may occur, which will be explained in part II of this thesis.

Immune checkpoint upregulation in a subset of MSI cancers explain why some patients with MSI cancer benefit from checkpoint inhibitors, particularly anti-PD-1 therapy [33,34]. The data that was established so far to employ checkpoint inhibitors for MSI cancers are already highly promising, there are multiple other clinical trials with checkpoint inhibitors for MSI CRC that are still ongoing. Despite checkpoint inhibitors are highly effective against a subset of MSI cancers, the combination of checkpoint inhibitors with an FSP vaccine may boost and complement the immune response against MSI cancers and potentially improve the response rate. Moreover, better knowledge of immune evasion phenomena (discussed in part II) would allow better selection of MSI cancer patients likely benefitting from checkpoint blockade.

1.6 From genome to antigen prediction

1.6.1 Human setting

In order to find out the cMSs those are affected during MSI carcinogenesis, Woerner and colleagues in 2001 applied a systematic identification procedure. First, coding sequences were detected from human sequences by search of EMBL database (EMBL Rel. 62, March 2000) by using various filters. More than 90% of the detected coding mononucleotide repeats (cMNR) and coding dinucleotide repeats (cDNR) contained less than 9 and 5 repeat units, respectively. Therefore, the longer the size of the MS is, less frequent it occurs in the genome. As a result of this first database search, 15 novel cMNRs and 4 novel cDNRs were identified and also experimentally verified by specific primers. In the second step, frequencies of cMS instability were determined in MSI and MSS samples; and it was shown that almost all cMNRs were unstable in MSI but not MSS and cDNRs only rarely display MSI. In the third step, expression of genes that contain cMSs were checked by RT-PCR and 10 out of 15 cMNRs were shown to be highly expressed both in MSI and MSS. In the last step, in order to check whether cMNR frameshift mutations in gDNA also occur in cDNA, mutation analysis was done on the cDNA level; half of the genomic cMNR frameshift mutations could be

confirmed to arise also at cDNA level in MSI cell lines [35].

Experimental evidence for the FSP-specific immune response, which marks frameshift mutations of the cMNRs of predicted genes, was presented in various publications. In 2001 by Linnebacher and colleagues, by using FSP pulsed autologous CD40-activated B cells as antigen presenting cells, HLA-A2.1-restricted frameshift peptide specific cytotoxic T cells were generated; and 3 FSPs, including -1 frameshift mutation of *TGFBR2*, used in this study were shown to specifically lyse the target cells [30]. In another study, spontaneous FSP-specific T cell responses (including TAF1B (-1), HT001 (-1), AIM2 (-1) and TGFBR2 (-1)) were detected in the peripheral blood of MSI patients by Elispot assay; furthermore in the same study, tumor-infiltrating T cells were also shown to have cytotoxic activity against MSI cells in a FSP-specific manner [27].

1.6.2 Mouse setting

In order to study the immune consequences of MMR deficiency and MSI in more detail, appropriate mouse models are required, but lacking so far. Initial mouse models of transgenic MMR-deficient mice showed the desired repair defects, but not all characteristics were matching with Lynch syndrome, as mice typically developed lymphomas due to constitutional MMR deficiency in all cells, similar to human CMMRD patients [36].

Since lymphoma development, however, is not common in human Lynch syndrome, the mouse models that developed lymphoma did not meet our needs and do not relate well with human setting. Kucherlapati and colleagues in 2010 established a conditional MMR-knockout mouse model *VCMsh2*$^{\text{Loxp/LoxP}}$ by using the *Msh2*LoxP allele in combination with villin-cre transgene. In this mouse model of Lynch syndrome, tumorigenesis is restricted to the intestinal tract, and 89 % of mice develop tumors in small intestine around 9 months of age spontaneously [37]. This tumor model; *Msh2*$^{\text{Loxp/LoxP}}$ together with *Mlh1*$^{-/-}$ and *Msh2*$^{-/-}$ mice were used by Woerner and colleagues in 2015 to detect the cMS frameshift mutations in the tumors of DNA MMR- deficient mouse models.

A MMR knockout and knock-in mouse lines.

Gene		Molecular Defects				Phenotypic Defects	
Knock-out		MSI	DDR	TRI	SHM/CSR	Cancer	Fertility
MutS	$Msh2^{-/-}$	+++	+++	+++	+++	+++	-
	$Msh3^{-/-}$	+	-	+++	-	+	-
	$Msh6^{-/-}$	+	+++	-	++	++	-
	$Msh4^{-/-}$	-	NA	NA	NA	-	+++
	$Msh5^{-/-}$	-	NA	NA	NA/-	-	+++
MutL	$Mlh1^{-/-}$	+++	+++	+++	+++	+++	+++
	$Mlh3^{-/-}$	++	+++	+++	-	+++	++
	$Pms1^{-/-}$	+	NA	NA	NA	-	++
	$Pms2^{-/-}$	+++	+++	+++	-/+++	+++	+++
Knock-in		MSI	DDR	TRI	SHM/CSR	Cancer	Fertility
MutS	$Msh2^{G674A}$	+++	-	+++	+++/+	+++	-
	$Msh2^{G674D}$	+++	-	NA	NA	+++	-
	$Msh6^{T1217D}$	+++	-	NA	++/-	+++	-
MutL	$Mlh1^{G67R}$	+++	-	NA	-/+++	+++	+++
	$Pms2^{E702K}$	+++	-	NA	-/+++	+++	-

Abbreviations: MSI, microsatellite instability; DDR, DNA damage response; TRI, triplet repeat instability; SHM, somatic hypermutation; CSR, class switch recombination.

B Conditional MMR knockout mouse lines.

Gene	Cre-Transgene	Target Tissue	Observed Phenotype	MSI
$Msh2^{loxP}$	*Villin-Cre*	Small and large intestine	Small intestinal tumors	+++
$Msh2^{loxP}$	*Ella-Cre*	All tissues	T-cell lymphomas Small intestinal tumors	na
$Msh2^{loxP}$	*Lgr5-CreERT2*	Lgr5 expressing stem cells	Small intestinal tumors	na
$Msh2^{loxP}$	*Adenoviral-Cre*	Cre encoding adenovirus	Large intestinal tumors in combination with Apc^{loxP} allele	na
$Msh2^{loxP}$	*D9-Cre*	Medium-spiny GABA-ergic projection neurons (MSNs)	Reduced HTT-CAG expansions and modified huntingtin phenotypes	na
$Mlh1^{loxP}$	*Ella-Cre*	All tissues	T-cell lymphomas Gastrointestinal tumors	+++
$Mlh1^{loxP}$	*Lck-Cre*	Lck expressing early stage thymocytes	T-cell lymphomas	++

Table 1 DNA MMR deficient mouse models. Adapted from [38]

In order to detect frameshift mutations of cMSs in the tumors of the mouse models, a similar approach that was used for detection of human frameshift mutations of cMS was followed. First, to find the cMNRs, the complete mouse genome was screened in the mouse Ensembl database (rel. 75.38); all of the cMNRs with at least four mononucleotide repeat were fetched by excluding the repeat tracts within pseudogenes, vector sequences and homopolymeric nucleotide sequences that are at the very end of 5´and 3´ends of the sequences. Afterwards, specific primers were

designed, and cMNRs were analyzed in gDNA extracted from tumor tissues of the mouse models. In the next step, MSI was determined by using a panel of five long mononucleotide repeats that was defined before to detect MSI in mice [39]. Tumors were annotated as MSI if at least 2 out of five markers showed instability. After the specific primers were designed, cMNR frameshift mutation analysis was carried out by comparing the tumor and healthy tissue by fragment length analysis [40]. All in all by this study, it was shown that the tumors from DNA MMR-deficient mouse models also display cMNR instability that would potentiate the studies on immunogenicity of FSP in DNA MMR-deficient mouse models.

1.7 Immune evasion mechanisms in MSI cancer

MSI colorectal cancers, both those developing as sporadic tumors or in the context of Lynch syndrome, are highly immunogenic. Formation of FSP neoantigens, which is caused by insertion and deletion mutations in cMSs, leads to an activation of the host's immune system, which for example manifests through dense infiltration of MSI tumors with cytotoxic CD8+ T cells [15,20,41]. However, despite immense infiltration by immune cells, due to constant high immune selection pressure, the tumor cells acquire immune evasion mechanisms and grow out to manifest into cancer [41,42]

The immunoselection hypothesis is corroborated by the mechanisms that lead to loss of class I MHC antigen presentation following adoptive immunotherapies. Berger and colleagues in 2004 reported that the melanoma patient treated with vaccinations of autologous peptide-pulsed dendritic cells first developed a remarkably potent antimelanoma cytotoxic T cell responses; however, after the transient response, immune evasion mechanisms related with loss of MHC class I antigens were developed and the patient succumbed to disease progression and died [43]. In another case addressed by Rosenberg and colleagues in 2003, the metastatic melanoma patient received sequential treatments of adoptive cell treatment with minor modifications, yet after all the patient developed resistance mechanisms including the mutation of *β2-microglobulin* gene that resulted in loss of MHC class I antigen presentation by the tumor[44].

The most well-known and characterized immune escape mechanism is the inactivation of both alleles of *β2-microglobulin* gene (B2M), commonly due to mutation in one allele of the gene accompanied by loss of heterozygosity of the other locus resulting in complete loss of MHC class I expression on the cell surface because B2M the essential light chain of MHC class I [45]. 30% of MSI CRCs display *B2M* mutations [46]; however, *B2M* mutations are rarely seen in MSS CRCs [45,47].

The first hint that *B2M* mutations are functionally relevant came from the observation that in contrast to long microsatellites, that are frequently mutant, short microsatellites are very rarely mutant[48,24] (Figure 4). *B2M* gene encompassing three short mononucleotide repeats and one dinucleotide repeat; 5 nucleotide cMS of *B2M* are predicted to be very exceptional. Therefore, the common presence of mutations at such short cMS mutations indicate that there needs to be a strong selection pressure that favors the outgrowth of cells that have such primarily rare mutations.

More evidence regarding MSI tumors being more immunogenic came from Bernal and colleagues in 2012, they declared that MSI tumors show marked difference in the genes related to inflammatory responses against cancer including the overexpression of genes related to immune response intensity and cytotoxic cell activity; namely MSI tumors have higher infiltration of activated $CD8^+$ T cells, T cell-attractant chemokines and cytokines; the response is more favored for M1 macrophages and Th1 rather than M2 and Th2 and NKG2D ligands are more overexpressed in MSI tumors; which would possibly promote the cytotoxic activity of NK cells [49]. Further evidence came from Echterdiek and colleagues in 2016 that there is also a strong immunoselection pressure in the tumors that lost B2M and *B2M* mutations occur particularly in tumors growing in an activated immune cell environment [42].

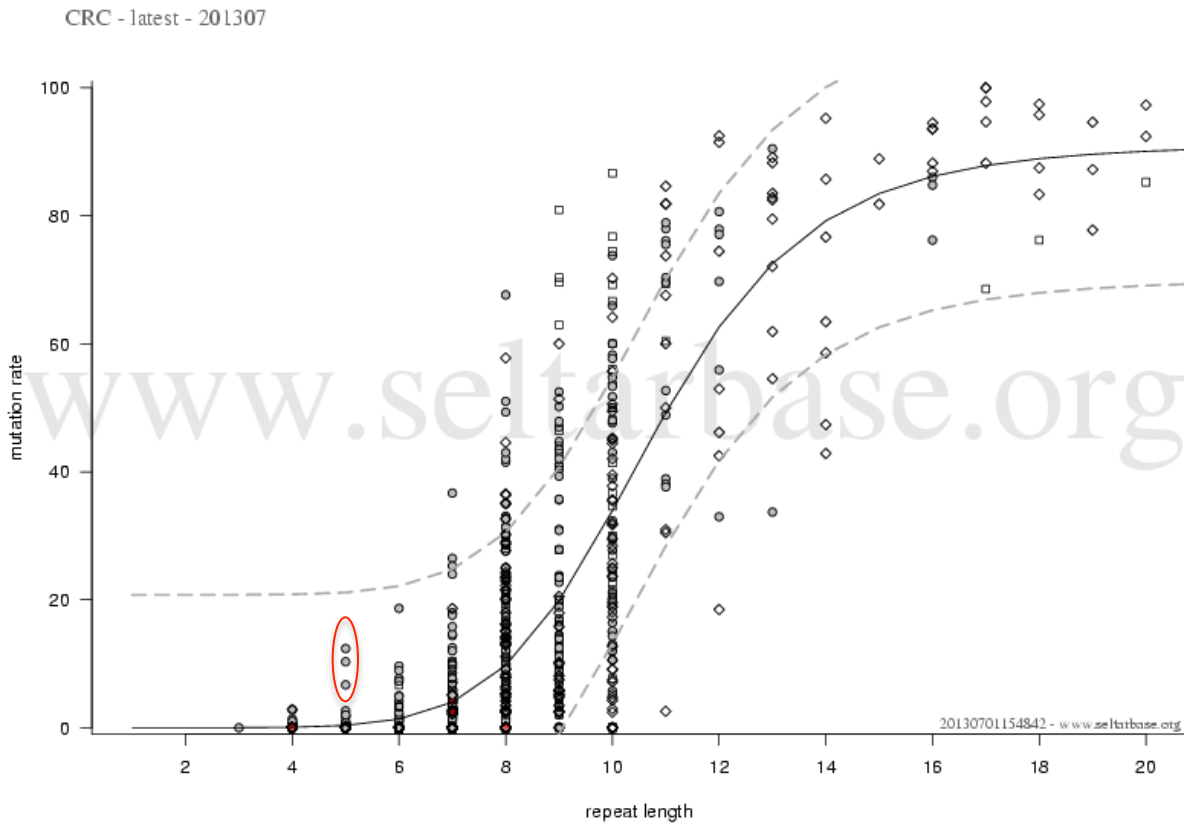


Figure 4 Coding microsatellite mutations in colorectal cancer. Mutation rate of the coding microsatellite increases as the repeat length increases; therefore short repeats are less pronounced compared to the long repeats. ncMNRs are shown by open diamonds, cMNRs by gray filled circles, nMNRs by open squares and red marked MNRs represent new data; the fitted regression line is shown as a black line, upper and lower prediction lines are shown in gray dashed lines and B2M repeats are circled in red.

In summary, *B2M* mutations are the most frequent and best characterized immune evasion mechanism in MSI tumors so far. In 30% of MSI cancers *B2M* is mutated leading to defects in MHC class I machinery [45], even if *B2M* is not mutated in 70% of the MSI tumor, there are still defects in MHC class I antigen presentation pathway. Those pathways still need to be addressed in further studies.

In addition to *B2M* mutations, alternative potential immune evasion mechanisms interfering with MHC class I antigen expression have been described in MSI cancers. Previous studies by Kloor and colleagues in 2005, analyzed the involvement of proteasome subunits and transporters of antigen presentation in immune evasion in MSI CRCs. In other words, in the study it was analyzed if APM components or ER chaperones could be the targets of MS instability and if the corresponding cMS were

mutated. It was found out that there were no mutations in proteasome subunits and ER chaperones but the transporter genes *TAP1* and *TAP2* genes were mutated in MSI colorectal carcinomas [45].

Recently, besides *B2M* and the other APM genes that play roles in MHC Class I presentation, a novel essential transactivator of MHC class I antigens; NOD-like receptor family CARD domain containing 5 (NLRC5) has been described [50]. Following the first study, five other research groups independently generated *Nlrc5* deficient mice and could show that *Nlrc5* deficiency causes a severe loss of MHC Class I and APM gene expression especially in immune cells [51–55].

It has been shown that NLRC5 is the key transcriptional regulator of classical and some non-classical MHC class I genes and APM genes; by directly occupying the promoters of those genes, NLRC5 transactivates the genes in a different extent [56]. For example, in T cells, it was reported that NLRC5 regulates the expression of *H2-Q6/7* genes 100%; whereas, the regulation of *H2-K* gene is 80-90% and *B2M* gene 50-60 % at the transcription level [57]. The transcriptional targets of NLRC5 has been defined extensively for murine cells as shown by table 2; however, the human system still lacks comprehensive information (table 3) [56]. In the murine system, *B2M*, classical and nonclassical MHC class I genes (*H2-D/K/L/M/Q/T*) and *Tap1* and in human; *B2M*, classical and nonclassical MHC class I genes (*HLA A/B/C/D/E/F/G*) and *TAP1* were shown to be the targets of NLRC5, respectively [56].

All in all, the clinical significance of NLRC5 can be discerned by the effect of NLRC5 deficiency in cancer. It has been recently manifested that NLRC5 alterations such as loss of function mutations, copy number loss and promoter hypermethylation are frequently seen in solid cancers; additionally mRNA levels of NLRC5 in tumor compared to healthy tissue is much less. Moreover, *NLRC5* mutations are more often observed as compared to other MHC class I related genes and therefore the global impact is more pronounced on level of MHC class I [58]. It has also been shown that in most of the cancer types analyzed; namely, nonsmall cell lung cancer, malignant melanoma, gastric adenocarcinoma, liver cancer, prostate cancer and rectal cancer, nuclear expression of *NLRC5* is significantly correlated with MHC class I expression[59]

and higher patient survival rates in many cancer types [58]. Furthermore, Rodriguez and colleagues reported that NLRC5 expression that is restored in tumor cell lines could restore the antitumor immunity by boosting the processing and presentation of tumor antigens to CD8 T cells [60]. In the light of all the information mentioned, NLRC5 has a great significance as a tumor evasion mechanism, therefore, any studies in this field would further support and shed light on how the interactions and the mechanism works.

Table 1 Murine NLRC5 Targets

Mouse Targets	Organism Used for Analysis	Methods Used for Validation	Tissue/Cell Type or Line	References
B2m	C57BL/6, *Nlrc5*$^{-/-}$	ChIP, mRNA	Spleen, thymus, T, NK, B	Biswas et al. (2012), Ludigs et al. (2015), Robbins et al. (2012), Staehli et al. (2012), and Yao et al. (2012)
	Cell line (+NLRC5)	mRNA	B16–F10	Rodriguez et al. (2016)
H2-D	C57BL/6, *Nlrc5*$^{-/-}$	ChIP mRNA, protein	Spleen, thymus, LN, liver, T, CD4, CD8, NK, B, NKT, γδT, cDC, Mϕ, BMDM, BMDC, mTEC/cTEC	Biswas et al. (2012), Ludigs et al. (2016, 2015), Robbins et al. (2012), Rota et al. (2016), Staehli et al. (2012), and Yao et al. (2012)
	BALB/c, *Nlrc5*$^{-/-}$	Protein	T	Ludigs et al. (2016)
	Cell line (+NLRC5)	mRNA, protein	B16–F10	Rodriguez et al. (2016)
H2-K	C57BL/6, *Nlrc5*$^{-/-}$	ChIP, mRNA, protein	Spleen, thymus, LN, liver, Ileum, kidney[a], T, CD4, CD8, NK, B, NKT, γδT, cDC, Mϕ, BMDC, BMDM, mTEC/cTEC	Biswas et al. (2012), Ludigs et al. (2016, 2015), Robbins et al. (2012), Rota et al. (2016), Staehli et al. (2012), Tong et al. (2012), and Yao et al. (2012)
	BALB/c, *Nlrc5*$^{-/-}$	Protein	T	Ludigs et al. (2016)
	Cell line (+NLRC5)	mRNA, protein	B16–F10	Neerincx et al. (2014) and Rodriguez et al. (2016)
H2-L	BALB/c, *Nlrc5*$^{-/-}$	Protein	T	Ludigs et al. (2016)
H2-M3	C57BL/6, *Nlrc5*$^{-/-}$	ChIP, mRNA	Spleen, thymus, T, CD8	Biswas et al. (2012), Ludigs et al. (2016, 2015), and Yao et al. (2012)

Continued

Table 1 Murine NLRC5 Targets—cont'd

Mouse Targets	Organism Used for Analysis	Methods Used for Validation	Tissue/Cell Type or Line	References
H2-Q4	C57BL/6, *Nlrc5*$^{-/-}$	ChIP, mRNA	T, CD8	Ludigs et al. (2015)
H2-Q6	C57BL/6, *Nlrc5*$^{-/-}$	ChIP, protein (Qa2 epitope)	T, CD4, CD8, NK, B, NKT, cDC	Ludigs et al. (2016, 2015)
H2-Q7	C57BL/6, *Nlrc5*$^{-/-}$	ChIP, protein (Qa2 epitope)	T, CD4, CD8, NK, B, NKT, cDC	Ludigs et al. (2016, 2015)
H2-T10	C57BL/6, *Nlrc5*$^{-/-}$	ChIP	T	Ludigs et al. (2015)
H2-T18	C57BL/6, *Nlrc5*$^{-/-}$	mRNA	B	Robbins et al. (2012)
H2-T22	C57BL/6, *Nlrc5*$^{-/-}$	ChIP, mRNA	T, CD8	Ludigs et al. (2015)
H2-T23	C57BL/6, *Nlrc5*$^{-/-}$	mRNA	Spleen, thymus	Yao et al. (2012)
Psmb8	Cell line (+NLRC5)	mRNA	B16–F10	Rodriguez et al. (2016)
Psmb9	C57BL/6, *Nlrc5*$^{-/-}$	ChIP, mRNA	Spleen, thymus, T, CD8	Biswas et al. (2012), Ludigs et al. (2015), and Yao et al. (2012)
	Cell line (+NLRC5)	mRNA	B16–F10	Rodriguez et al. (2016)
Tap1	C57BL/6, *Nlrc5*$^{-/-}$	ChIP, mRNA	Spleen, thymus, T, CD8	Biswas et al. (2012), Ludigs et al. (2015), and Yao et al. (2012)
	Cell line (+NLRC5)	mRNA	B16–F10	Rodriguez et al. (2016)

[a]Not confirmed by the study of Ludigs et al. (2016).
+NLRC5 = Analysis performed after NLRC5 overexpression.
ChIP: Chromatin immunoprecipitation of endogenous NLRC5 followed by qPCR and/or deep sequencing; mRNA: qRT–PCR analysis; protein: FACS analysis.
LN: Lymph node; T: T cells; CD4: CD4$^+$ T cells; CD8: CD8$^+$ T cells; NK: natural killer cells; B: B cells; NKT: natural killer T cells; γδT: gamma delta T cells; cDC: conventional dendritic cells; Mϕ: macrophages; BMDC: bone marrow-derived dendritic cells; BMDM: bone marrow-derived macrophages; mTEC: medullary thymic epithelial cells; cTEC: cortical thymic epithelial cells.

Table 2 Murine NLRC5 targets retrieved from [56]

Table 2 Human NLRC5 Targets

Human Targets	+NLRC5	−NLRC5	Cell Lines	References
B2M	ChIP, mRNA, protein		Jurkat; 293T	Meissner et al. (2010)
HLA-A	ChIP, mRNA, protein, reporter		Jurkat; 293T; HeLa	Meissner et al. (2010), Meissner, Li, Liu, Gagnon, and Kobayashi (2012), Meissner, Liu, et al. (2012), Neerincx et al. (2014), and Staehli et al. (2012)
		Protein	HeLa	Meissner et al. (2010)
HLA-B	ChIP, mRNA, protein, reporter		Jurkat; 293T; HeLa	Meissner et al. (2010), Meissner, Li, et al. (2012), Meissner, Liu, et al. (2012), Neerincx et al. (2014), and Staehli et al. (2012)
		mRNA, protein	HeLa; THP1	Meissner et al. (2010) and Neerincx et al. (2014)
HLA-C	ChIP, mRNA, protein		Jurkat; 293T	Meissner et al. (2010), Meissner, Li, et al. (2012) and Staehli et al. (2012)
		Protein	HeLa	Meissner et al. (2010)
HLA-E	ChIP, protein		293T	Meissner et al. (2010)
HLA-F	ChIP		293T	Meissner et al. (2010)
HLA-G	ChIP		293T	Meissner et al. (2010)
PSMB9	mRNA, protein		Jurkat	Meissner et al. (2010)
TAP1	ChIP, mRNA, protein		Jurkat; 293T	Meissner et al. (2010)

+NLRC5 = analysis performed after NLRC5 overexpression; −NLRC5 = analysis performed after NLRC5 knockdown by SiRNA.
ChIP: Chromatin immunoprecipitation on tagged NLRC5 followed by qPCR; mRNA: qRT–PCR analysis; protein: FACS and/or Western blot analysis; reporter: luciferase assays.
293T: HEK 293T cells.

Table 3 Human NLRC5 targets retrieved from [56]

1.8 Aims

Lynch syndrome and sporadic MSI cancers represent a unique immunogenic model of cancer where specific FSP neoantigens can be targeted as a vaccination approach. FSPs, being entirely novel to the immune system of the host, are highly immunogenic antigens that specifically occur in MMR-deficient cells. They encompass long antigenic amino acid stretches and contain multiple potential epitopes to be presented to the immune system. Additionally, the fact that FSP-generating mutations affecting cMS in tumor suppressor genes are recurrent due to tumor evolution, FSP neoantigens are shared among the majority of MMR-deficient cancers. Therefore, FSP vaccination approaches can be developed to prevent tumor formation in Lynch syndrome patients and to potentially establish a new therapy against all types of MSI cancers.

In order to translate this concept into clinic, we aim 1) to set up an appropriate mouse model for the evaluation of an FSP vaccine and 2) to provide a comprehensive overview of immune evasion mechanisms in MSI cancers. A mouse model is essential because it allows observing the stages of tumor development including pre-cancerous stages, and because changes of immune responses over time can be related to emerging lesions. Furthermore, FSP vaccine in the mouse model of Lynch syndrome can be evaluated in combination with other treatments such as aspirin or PD1/PD-L1. On the other hand, an extensive overview of immune evasion mechanisms in MSI cancer is required to gain a profound understanding required for better treatment strategies and overcoming therapy resistance. Therefore, there are two parts in this thesis, which will first explain our progress on setting up a mouse model for Lynch syndrome and second the systematic examination of immune evasion mechanisms in MSI cancers both in the cell lines, tumor tissues and publicly available databases.

2 PATIENTS AND MATERIALS

2.1 Patients and tumor samples

Tumor samples were collected from the Institute of Pathology, University Hospital Heidelberg. In compliance with the Declaration of Helsinki and ethical standards, informed written consent was obtained from all the patients included in this study. All tumor specimens were collected as FFPE blocks. Tumor samples were used to characterize the mutations of *NLRC5* gene and MHC class I expression status.

2.2 Cell lines

	Cell Line	Source
1	Co 115	Dr. R. Hamelin, INSERM, Paris, France
2	COGA-1	L.A.Huber, University of Innsbruck, Innsbruck, Austria
3	Colo-60H	CLS Cell lines services, Heidelberg, Germany
4	DLD-1	DKFZ Tumor Bank, Heidelberg, Germany
5	GP2D	W. Bodmer, Cancer Research UK, London, United Kingdom
6	HCT-15	DKFZ Tumor Bank, Heidelberg, Germany
7	HCT 116	ECACC
8	HDC 108	M. Schwab, DKFZ, Heidelberg, Germany
9	HDC 135	M. Schwab, DKFZ, Heidelberg, Germany
10	HDC 143	M. Schwab, DKFZ, Heidelberg, Germany
11	HDC 9	M. Schwab, DKFZ, Heidelberg, Germany
12	HROC 24	M. Linnebacher, University of Rostock, Rostock, Germany
13	HRT-18	CLS Cell line service
14	HUTU-80	DKFZ Tumor Bank, Heidelberg, Germany
15	K073 A	Generated in house
16	KM 12	Dr. I.J.Fidler, MD Anderson Cancer Center, Houston,United States
17	LIM 1215	Dr. R.H. Whitehead, Ludwig Institute of Cancer Research, Melbourne, Australia
18	LIM 2405	Dr. R.H. Whitehead, Ludwig Institute of Cancer Research, Melbourne, Australia
19	LIM 2412	Dr. R.H. Whitehead, Ludwig Institute of Cancer Research, Melbourne, Australia
20	LIM 2537	Dr. R.H. Whitehead, Ludwig Institute of Cancer Research, Melbourne, Australia
21	LIM 2551	Dr. R.H. Whitehead, Ludwig Institute of Cancer Research, Melbourne, Australia
22	LoVo	DKFZ Tumor Bank, Heidelberg, Germany
23	LS 174 T	DKFZ Tumor Bank, Heidelberg, Germany
24	LS 180	DKFZ Tumor Bank, Heidelberg, Germany
25	LS 411	W. Bodmer, Cancer Research UK, London, United Kingdom
26	RKO	Dr. M. Brattain, University of Texas, HealthScience Center, San Antonio, USA
27	SW 48	ECACC
28	TC 7	Dr. R. Hamelin, INSERM, Paris, France
29	TC 71	Dr. R. Hamelin, INSERM, Paris, France
30	Vaco 432	Dr. J. Wilson, Case Western Reserve University, Cleveland, United States
31	Vaco 457	Dr. J. Wilson, Case Western Reserve University, Cleveland, United States
32	Vaco 5	Dr. J. Wilson, Case Western Reserve University, Cleveland, United States
33	Vaco 6	Dr. J. Wilson, Case Western Reserve University, Cleveland, United States

2.3 Antibodies

Antibodies	Dilution	Company (City, Country)
IHC		
Mouse anti-human HLA Class I HCA2	1:150	Acris (Rockville, USA)
Mouse anti-human HLA Class I HC10	1:150	Acris (Rockville, USA)
Biotin coupled horse anti mouse IgG	1:100	Vector Laboratories (Burlingame, USA)
Rat anti-mouse CD4	1:100	Dianova (Hamburg, Germany)
Elispot		
Rat anti-mouse IFNγ	1:200	BD Biosciences (Durham, USA)
Biotin coupled rat anti-mouse IFNY	1:500	BD Biosciences (Durham, USA)
Elisa		
HRP coupled anti-mouse IgG	1:5000	Thermo Fisher Scientific (Waltham, USA)
Flow cytometry		
CD3-PECy7	1:100	BD Biosciences (Durham, USA)
CD4-FITC	1:200	BD Biosciences (Durham, USA)
CD8-PE	1:100	BD Biosciences (Durham, USA)
Aqua Zombie-AM Cyan	1:100	Biolegend (San Diego, USA)
Mouse anti-human HLA-A/B/C TP25.99.8.4	1:25	Kind gift of Soldano Ferrone (Harvard Medical School, Boston, USA)
Rat anti-mouse IgG-FITC	1:20	Thermo Fisher Scientific (California, USA)
Western blot		
Rabbit anti-human B2M	1:250	Sigma Aldrich (Darmstadt, Germany)

2.4 Oligonucleotides

Oligonucleotides	Sequence	Company (City, Country)
CpG ODN 1628	TsCsCsAsTsgsAsCsgsTsTsCsCsTsgsAsCsgsTsT	TIB MolBiol (Berlin, Germany)

2.5 Primers

Primers	Sequence 5´to 3´	Annealing temperature (degree celcius)	Company (City, Country)
B2M Exon 1 F	GGCATTCCTGAAGCTGACA	59	Life Technologies (Darmstadt, Germany)
B2M Exon 1 R	AGAGCGGGAGAGGAAGGAC	59	Life Technologies (Darmstadt, Germany)
B2M Exon 2a F	TTTTCCCGATATTCCTCAGGTA	57	Life Technologies (Darmstadt, Germany)
B2M Exon 2a R	AATTCAGTGTAGTACAAGAG	57	Life Technologies (Darmstadt, Germany)
B2M Exon 2b F	CATTCAGACTTGTCTTTCAG	64	Life Technologies (Darmstadt, Germany)
B2M Exon 2b R	TTTCAGCAGCTTACAA	64	Life Technologies (Darmstadt, Germany)
NLRC5 F	CTCCTCTCACCCTCTCCTCT	60	biomers.net (Ulm, Germany)
NLRC5 R	GCAGCCCCTACTTACCTGAT	60	biomers.net (Ulm, Germany)

2.6 Consumables

Consumables	Company (City, Country)
Cell culture disposables (sterile, pyrogen-free)	Greiner Bio-One (Frickenhausen,Germany)
Dako pen	DAKO (Hamburg, Germany)
ELISA plates	Fisher Scientific (Roskilde, Denmark)
Elispot plates	Merck Milipore (Darmstadt, Germany)
Microscope cover glasses	Marienfeld (Lauda-Königshofen, Germany)
Microscope slides	Menzel (Braunschweig, Germany)
NuPAGE Novex 4-12% Bis-Tris	Thermo Fisher Scientific (Carlsbad, USA)
PCR tubes	Steinbrenner (Wiesenbach, Germany)
Pipette tips	Corning Life Science (New York, USA)
Reaction tubes	Greiner Bio-One (Frickenhausen,Germany)
Serological pipettes	Sarstedt (Nümbrecht, Germany)

2.7 Kits

Kits	Company (City, Country)
BCTP/NBT substrate	Sigma Aldrich (Steinheim, Germany)
Bio-rad Nupage Electrophoresis system	Life Technologies (Darmstadt, Germany)
CD4 T cell isolation kit	Miltenyi Biotec (Bergisch Gladbach, Germany)
CD8 T cell isolation kit	Miltenyi Biotec (Bergisch Gladbach, Germany)
DAB Chromogen	DAKO (Hamburg, Germany)
Fixation/Permeabilization kit	BD Biosciences (Durham, USA)
ImmPress-Reagent anti rat IgG HRP-conjugated; mouse adsorbed	Vector labs (Burlingame, USA)
Platinum Taq DNA Polymerase kit	Thermo Fisher Scientific (Waltham, USA)
QIAamp DNA FFPE tissue kit, Qiagen	Qiagen (Hilden, Germany)
Qiagen DNeasy Blood and Tissue Kit	Qiagen (Hilden, Germany)
QIAquick PCR purification kit	Qiagen (Hilden, Germany)
Vectastain Elite ABC kit	Vector (Burligame, USA)
Western Breeze Chemoluminescent immunodetection kit	Thermo Fisher Scientific (Waltham, USA)

2.8 Peptides

Peptides	Sequence	Length	Company (City, Country)
Nacad (-1)	VIYAPPPPAEGRWPCWLLRAH	21	Genaxxon Bioscience GmbH (Ulm, Germany)
Nacad wt	VIYAPPPPSRGAVAVLATPS	20	Genaxxon Bioscience GmbH (Ulm, Germany)
Nacad (-1) N	VIYAPPPPAEGRW	13	Genaxxon Bioscience GmbH (Ulm, Germany)
Nacad (-1) C	AEGRWPCWLLRAH	13	Genaxxon Bioscience GmbH (Ulm, Germany)
Xirp (-1)	GKGPGGPPLSSPKRVMYRLSVGCLRPTL	28	Genaxxon Bioscience GmbH (Ulm, Germany)
Xirp1 wt	GKGPGGPPPELPKKGDVQTI	20	Genaxxon Bioscience GmbH (Ulm, Germany)
Xirp1 (-1) N	GKGPGGPPLSSPKRVMYRLS	20	Genaxxon Bioscience GmbH (Ulm, Germany)
Xirp1 (-1) C	VMYRLSVGCLRPTL	14	Genaxxon Bioscience GmbH (Ulm, Germany)
Maz (-1)	PCTLLAPPSPCWAWTPGGWAAS	22	Genaxxon Bioscience GmbH (Ulm, Germany)
Maz wt	PCTLLAPPFPVLGLDSRGVG	20	Genaxxon Bioscience GmbH (Ulm, Germany)
Maz (-1) N	PCTLLAPPSPCWAW	14	Genaxxon Bioscience GmbH (Ulm, Germany)
Maz (-1) C	SPCWAWTPGGWAAS	14	Genaxxon Bioscience GmbH (Ulm, Germany)
Senp6 (-1)	VKCSMKKKIMLSMKMKNQVTENLRARTFVIEPKVRMASGMNASVLYIIQMP	51	Genaxxon Bioscience GmbH (Ulm, Germany)
Senp6 wt	VKCSMKKKNHAINENEEPSN	20	Genaxxon Bioscience GmbH (Ulm, Germany)
Senp6 (-1) N	VKCSMKKKIMLSMKMKNQVTENLRA	25	Genaxxon Bioscience GmbH (Ulm, Germany)
Senp6 (-1) C	NLRARTFVIEPKVRMASGMNASVLYIIQMP	30	Genaxxon Bioscience GmbH (Ulm, Germany)

2.9 Chemicals and reagents

Chemicals and reagents	Company (City, Country)
ABI Prism Big Dye Terminal v1.1 Ready Reaction Mix	Thermo Fisher Scientific (Waltham, USA)
ACK buffer	Gibco (New York, USA)
Bio-rad precision plus protein standard protein ladde	Bio-rad (California, USA)
Bovine Serum Albumin (BSA)	Sigma Aldrich (Taufkirchen, Germany)
Bromophenol blue	Schmid (Köngen, Germany)
Citric acid	Merck (Darmstadt, Germany)
Di-Aminobenzidine (Liquid DAB+substrate)	DAKO (Hamburg, Germany)
Dimethyl sulfoxide	Merck (Darmstadt, Germany)
Disodium Hydrogenphosphate	VWR International (Bruchsal, Germany)
Dulbecco's PBS without Ca and Mg	Gibco (New York, USA)
Ethanol absolute	VWR International (Bruchsal, Germany)
Ethylenediaminetetraacetic acid	Merck (Darmstadt, Germany)
Fetal Bovine Serum	Gibco (New York, USA)
Hi-Di formamide	Applied Biosystems (Darmstadt, Germany)
Hydrochloric acid	VWR International (Bruchsal, Germany)
Isopropanol	Sigma Aldrich (Taufkirchen, Germany)
Midori Green	Biozym (Vienna, Austria)
NuPAGE Sample reducing agent, DTT	Thermo Fisher Scientific (Carlsbad, USA)
Penicillin-Streptomycin	Gibco (New York, USA)
Potassium Chloride	J. T. Baker (Deventer, Netherlands)
Potassium dihydrogenphosphate	Gerbu Biochemicals (Gaisberg, Germany)
Proteinase K	Qiagen (Hilden, Germany)
RPMI 1640	Gibco (New York, USA)
Serum free protein block	DAKO (Hamburg, Germany)
Sodium Chloride	AppliChem (Darmstadt, Germany)
Sodium hydroxide	Carl Roth (Karlsruhe, Germany)
TMB substrate	Sigma (Steinheim, Germany)
Tris base	AppliChem (Darmstadt, Germany)
Trypsin	Gibco (New York, USA)
Tween 20	Sigma Aldrich (Taufkirchen, Germany)
Water, DNase RNase free	MP Biomedicals (Solon, USA)

2.10 Buffers

Buffers	Formulations
MACS buffer	0.5% bovine serum albumin (BSA) and 2 mM EDTA in PBS pH 7.2
Citrate buffer (10X)	21 g citric acid dissolved in distilled water, pH 6.0
TBS (1X)	50 mM Tris base and 150 mM NaCl dissolved in distilled water, pH 7.6
TBST	0.05% Tween 20 in TBS
PBS (1X)	140 mM NaCl, 2.7 mM KCl, 8.1 mM Na2HPO4 and 1.5 mM KH2PO4 dissolved in distilled water pH 7.4
RIPA buffer	50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% Triton x-100, 1% Socium deoxycholate, 0.1% SDS, 0.1 mM CaCl2 and 0.01 mM MgCl2

2.11 Instruments

Instruments	Company (City, Country)
Analytical scale (BP 210 D)	Sartorius (Göttingen, Germany)
Camera	Olympus U CMAD3 (Tokyo, Japan)
Camera (Gel Doc 2000)	Bio-Rad (Munich, Germany)
Centrifuge (Biofuge 13)	Heraus Holding (Hanau, Germany)
Centrifuge (Microcentrifuge 1-14)	Sigma (Osterode, Germany)
Chemidoc MP System	Bio-Rad (Vienna, Austria)
DNA Speed Vac Vacuum dryer (Savant)	Thermo Fisher Scientific (Waltham, USA)
ELISA reader (GENios)	GENios Tecan (Crailsheim, Germany)
Elispot reader	Immunospot (Ohio, USA)
FACS Calibur	Becton Dickinson (Franklin Lakes, USA)
FACS Canto II	BD Biosciences (Durham, USA)
Gel chamber (Sub-Cell GT)	Bio-Rad (Munich, Germany)
Genetic analyzer (ABI3130xl Sequencer)	Applied Biosystems (Darmstadt, Germany)
Incubator (CO2)	Heraus Holding (Hanau, Germany)
Magnetic stirrer (IKAMAC RCT)	IKA (Staufen, Germany)
Microscope	Zeiss (Jena, Germany)
Microwave	Siemens (Munich, Germany)
PCR system (GeneTouch)	Bulldog Bio (Portsmouth, USA)
pH meter (PB-11)	Sartorius (Göttingen, Germany)
Pipettes (10, 20, 200, 1000 µl)	Gilson (Limburg-Offheim, Germany)
Power supply (Consort e835)	Peqlab (Erlangen, Germany)
Protein Electrophoresis system (NuPAGE)	Life Technologies (Darmstadt, Germany)
Sonicator (Transsonic T310)	Elma Hans Schmidbauer (Singen, Germany)
Spectrophotometer Nanodrop 2000	Thermo Fisher Scientific (Massachussets, USA)
Stereomicroscope	Leica (Bensheim, Germany)
Vortex (MS1 Minishaker)	IKA (Staufen, Germany)
Waterbath (SW20)	Julabo Labortechnik (Seelbach, Germany)

2.12 Softwares

Softwares	Company (City, Country)
Graphpad Prism 7	Graphpad Software (San Diego, USA)
ImmunoSpot	Cellular Technology Limited (Cleveland, USA)
Magellan	Tecan (Männedorf, Switzerland)
Mendeley 1.17.9	Elsevier Inc. (New York, USA)
Sequencing Analysis Software 6	Applied Biosystems (Foster City, USA)

3 METHODS

3.1 Cell culture and maintenance

All cell culture experiments were performed under sterile conditions in laminar flow hood and the cell lines were cultured in RPMI media with 10% heat-inactivated FBS and 1% penicillin-streptomycin in the incubator at 37^0C and 5% CO_2.
Cells were grown in flasks until 80% of confluency, then after washing the cells with sterile PBS, cells were trypsinized and splitted according to need.
In order to freeze the cells, cells were counted by using a hematocytometer and $5x10^6$ cells per vial was resuspended in 1 ml of cold freezing medium containing FBS with 10% DMSO. For long-term storage, cells were frozen in liquid nitrogen tanks.
In order to thaw the cells, a vial of frozen cells were hold in 37 ^{0}C water bath until the ice started melting, then cells were washed once with the culture media, centrifuged and resuspended in 1 ml of media and transferred to a cell culture flask for further experiments.

3.2 DNA extraction from cell lines

Cell lines were grown until 80% of confluency and $5x10^6$ cells were used in DNA extraction protocol.
After cells were transferred into 15 ml tubes, cells were washed with PBS to get rid of media and cell pellets were frozen at -20^0C until DNA isolation.
DNA isolation was carried out by using Qiagen DNeasy Blood and Tissue Kit according to manufacturer´s protocol. Briefly, cells were thawed and centrifuged at 1200 rpm for 5 minutes, resuspended in 200µl PBS. After addition of 20µl proteinase K and 200µl of buffer AL, the samples were mixed thoroughly and incubated at 70 ^{0}C for 10 minutes. After incubation, 200µl ethanol (96-100%) was added on the samples and vortexed. Next, the mixture was pipetted into DNeasy mini spin column in a 2 ml collection tube and centrifuged at 6,000g for 1 minute. Following the discarding of the flow-through in the collection tube, 500µl buffer AW1 was added centrifuged at 6,000g for 1 minute and the flow through was discarded. Then, 500µl buffer AW2 was added centrifuged at 20,000g for 3 minutes to dry the membrane and flow through

was discarded one more time. In the next step, a clean 1.5 ml tube was placed under the column and 200µl DNase-RNase free water was added in DNeasy membrane, incubated at 70 ^{0}C for 10 minutes and centrifuged for 1 minute at 6,000 g to elute the DNA. The concentration and quality of DNA was measured in NanoDrop spectrophotometer by 260/280 (around 1.8 for DNA) and 260/230 (greater than 1.5) ratios.

3.3 DNA isolation from formalin fixed paraffin embedded tissues

Genomic DNA was isolated from hematoxylin/eosin stained FFPE tissue sections. By using a stereomicroscope, tumor and normal tissue of the sample was microdissected. Microdissection was done by wetting the 20 G needle with ATL buffer (QIAamp DNA FFPE tissue kit, Qiagen) and scratching the corresponding part of the slide into the 1.5 ml tubes containing 180 µl ATL buffer and 12 µl Proteinase K. After all the samples were collected into tubes, the samples were incubated overnight at 56 ^{0}C. The next day, DNA was isolated according to manufacturer´s protocol for QIAamp DNA FFPE tissue kit. Shortly, 200 µl Buffer AL was added on the tubes and vortexed, following incubation at 70 ^{0}C for 10 min, 200 µl ethanol was added, samples were mixed and transferred to columns for centrifugation at 6,000 g for 2 minutes. Afterwards, columns were washed twice with buffer AW1 and AW2 and flow through was discarded. Lastly, to elute the samples 35 µl of buffer AE was added on the columns, incubated at 70 ^{0}C for 15 minutes and centrifuged at 6,000 g for 3 minutes. The eluted DNA concentration and quality was measured by NanoDrop spectrophotometer.

3.4 Sequencing

After DNA isolation, in order to sequence DNA, several steps were performed;

a) **PCR amplification:** Prepare 50 ng/µl DNA stock from each sample and prepare the PCR mix

	Initial concentration	Final concentration	Volume (µl)
DNase-RNase free H_2O			14,5
PCR buffer*	10X	1X	2,5
MgCl2*	50 mM	3 mM	0,75
dNTPs	1,25 mM	0,2 mM	4
Primer forward	15 µM	0,3 µM	0,5
Primer reverse	15 µM	0,3 µM	0,5
Platinum Taq* Polymerase	5 U/µl	0,02 U/µl	0,2
Template DNA	50 ng/µl	150 ng/well	3
			In total 25 µl/sample

* Provided with Platinum Taq polymerase

PCR conditions are described below; for the elongation 38 cycles were performed.

B2M Exon1	B2M Exon 2a	B2M Exon 2b	NLRC5
94 ^{0}C 5 min	94 ^{0}C 5 min	94 ^{0}C 5 min	95 ^{0}C 10 min
94 ^{0}C 30sec 59 ^{0}C 15 sec	94 ^{0}C 52 sec 57 ^{0}C 66 sec 72 ^{0}C 44 sec	94 ^{0}C 52 sec 64 ^{0}C 66 sec 72 ^{0}C 44 sec	95 ^{0}C 30 sec 60 ^{0}C 30 sec 72 ^{0}C 30 sec
72 ^{0}C 6 min	72 ^{0}C 7 min	72 ^{0}C 7 min	72 ^{0}C 7 min
4 ^{0}C ∞	4 ^{0}C ∞	4 ^{0}C ∞	4 ^{0}C ∞

The PCR samples were run on 2 % agarose gel in 1 X TBE, midori green was used as a nucleic acid staining solution. 2 µl of loading dye was mixed with 5 µl PCR product or 1 µl 100 bp DNA ladder, and ran on gel at 140 V for 50 minutes. The bands were checked on the agarose gel image and the samples were stored at 4^{0}C for further analysis.

b) PCR purification: QIAquick PCR purification kit was used to get rid of impurities due to PCR reagents. 20 µl PCR product was mixed with 100 µl binding buffer and loaded on a spin column that is provided with the kit and centrifuged for 3 minutes at 13,200 rpm. Then, supernatant was discarded, washed by 700 µl PE buffer, centrifuged once more and flow through was discarded. To dry the

membrane, with a clean tube, another centrifugation was done for 1 minutes, columns were transferred to clean 1,5 ml tubes, depending on the intensity of the band on the agarose gel image, 30-60 µl of elution buffer was added onto columns, incubated for 20 minutes at room temperature and centrifuged for 3 minutes at 13,200 rpm. Following completion of PCR purification, samples were stored either at 4 ^{0}C or -20 ^{0}C.

c) **Single stranded linear amplification:** 4 µl of the purified PCR product was mixed with 2 µl of forward or reverse primer (1,5 µM) and 4 µl of ABI Prism Big Dye Terminal v1.1 Ready Reaction Mix and the following PCR conditions were used by GeneTouch PCR system; elongations were 25 cycles for all conditions and the samples were stored at 4 ^{0}C for further use.

B2M Exon 1 and 2a	B2M Exon 2b	NLRC5
96 ^{0}C 5 min	96 ^{0}C 5 min	96 ^{0}C 5 min
96 ^{0}C 10 sec	96 ^{0}C 10 sec	96 ^{0}C 10 sec
58 ^{0}C 10 sec	60 ^{0}C 10 sec	60 ^{0}C 10 sec
60 ^{0}C 4 min	60 ^{0}C 4 min	60 ^{0}C 4 min
4 ^{0}C ∞	4 ^{0}C ∞	4 ^{0}C ∞

d) **Ethanol precipitation:** Following addition of 90 µl of DNase-RNase free water on the PCR products, samples were mixed with 250 ul absolute ethanol and 10 µl sodium acetate pH 4,6 in a 1,5 ml tube and mixed thoroughly. After mixing, the samples were centrifuged at 13,000 rpm for 15 minutes, the supernatant was discarded, and the tubes were dried upside down. And the pellet was washed with 250 µl 75 % ethanol and centrifuged and the supernatant was removed. In order to get rid of remaining liquid, the pellets were vacuum dried for 6-8 minutes. Then the tubes containing the DNA were stored at 4 ^{0}C for further use.

e) **Sanger sequencing:** DNA pellets were resuspended in 12 µl Hi-Di Formamide and incubated at room temperature for 30 minutes. Afterwards, the samples were transferred into 96 well of ABI PCR plate and put into ABI prism 3100 Genetic Analyzer. The sequencing was analyzed by Sequencing Analysis software v.6.

3.5 Immunohistochemistry (IHC)

3.5.1 IHC staining for MHC class I heavy chains (HCA2 and HC10)

Thin sections of FFPE tissue of tumor specimens were deparaffinized and rehydrated by treatment of xylene three times for 5 minute, and graded alcohol series by 100 % of ethanol for two times for 5 minutes, 96 % ethanol for 5 minutes and 70 % ethanol for 5 minutes. At the end of these steps, the slides were rinsed with deionized water four to five times. Epitope retrieval was done by using 200 ml of 10 mM citric acid monohydrate solution for three times for 5 minutes at 560 watt in the microwave. The slides were cooled down for 20 minutes at room temperature and rinsed in deionized water. Afterwards, the endogenous peroxidase was quenched by incubation of slides for 20 minutes with 2 % of hydrogen peroxide solution in methanol. Then the slides were again rinsed in deionized water for four to five times. Water repelling Dako pen was used edge the tissue and the slides were washed for 5 minutes in PBS/0.05 % tween buffer. To block nonspecific antibody binding, slides were incubated with 10 % horse serum in PBS, 150-200 µl/slide for 30 minutes at room temperature. Primary antibodies were diluted 1:150 in PBS with 1 % serum, and the slides were incubated at 4 ^{0}C overnight. The slides were washed with PBS 0.05 % tween. Following washing, the secondary antibody biotin coupled horse anti-mouse IgG was diluted 1:100 in PBS with 1 % horse serum and incubated for 30 minutes at room temperature. The slides were washed two times in PBS/tween for 3 minutes. For the next step, reagent A/B from Vectastain Elite ABC kit was prepared (should be prepared 30 minutes before use) by dilution of both reagents 1:50 in PBS. The slides were incubated with reagent A/B for 30 minutes at room temperature and washed with PBS/tween. To develop the slides, chromogen system was used to stain; 1 drop of DAB and 1 ml of chromogen peroxidase substrate buffer was mixed and applied on slides 150- 200 µl. The color formation was observed and rinsing the slides in deionized water stopped the reaction. Hematoxylin/Eosin staining was used as a counterstain to stain the nucleus of the cells, the slides were exposed for 10-second intervals and washed under tap water to get rid of the excess dye and stop the reaction. The slides were covered with 2 drops of aquatex and cover glasses.

3.5.2 IHC for mouse CD4 T cells

Sections of mouse FFPE tumor specimens were deparaffinized and hydrated as explained above, then the antigen retrieval was done by using 200 ml of 10 mM citric acid monohydrate solution for 35 minutes in water bath at 97 ^{0}C and the slides were cooled down at room temperature for 20 minutes. The slides were washed three times for 1 minute, incubating the sections with 3 % hydrogen peroxide blocked the endogenous peroxidase activity, then the slides were washed for two times for 1 minute with PBS and one time with TBST for 2 minutes, respectively. To block nonspecific binding, serum-free protein block (Dako) was applied on the slides and incubated 10 minutes. The primary antibody rat monoclonal anti-mouse CD4 was diluted 1:100 in PBS, applied on the slides and incubated overnight. Next day slides were washed with TBST for three times for 2 minutes and secondary antibody, Impress anti-rat IgG was dropped on the slides, incubated 30 minutes at room temperature. After washing the slides with TBST for three times for 2 minutes, DAB substrate (DAKO) was added on the slides and color formation was observed and stopped by rinsing with distilled water. As a last step, counterstain was applied and the slides were covered as explained in IHC for HCA2 and HC10.

3.6 Immunization experiments

3.6.1 Animal maintenance

For immunization studies 6-8 weeks old female C57/BL6J mice were used. All the animals were housed at Animal House Facility of German Cancer Research Center (DKFZ, Heidelberg, Germany) maintained under controlled ambient conditions and unlimited access to food and water. Experimental protocols and animal handling were approved by the EU and national regulations for animal experimentation. Mice were purchased from Charles River WIGA Laboratories (Sulzfeld, Germany).

3.6.2 Prediction of the peptides

After the selection of 13 cMS genes, which have more than 15% of mutation frequency and expressed both in the tumor and normal tissue, potential 26 FSPs were assessed. All the peptide sequences shared 8 N terminal amino acids with the wild type peptide and 2 FSPs; one base deletion (FSP (-1)), and two base deletion (FSP (-2) after the wild type sequence were determined as a sequence for all the candidate 13 cMS genes. Therefore, having 26 FSPs out of 13 cMS genes were assessed in epitope prediction databases; SYFPEITHI database of MHC ligands and peptide motifs version 1.0 and netMHC version 4.0.

SYFPEITHI database scores the peptides based on published motifs; the frequency of the corresponding amino acid in natural ligands, binding peptides and T cell epitopes are estimated and each amino acid within a peptide sequence is given 10-15 points if it is an ideal anchor, 6-8 points if it is an unusual anchor, 4-6 if auxiliary anchors and 1-4 to the preferred residues. Some of the amino acids can also be scored as negative if it is disadvantageous for the peptide's binding capacity at a specific position[61]. On the other hand, netMHC provides two significant outputs; Affinity (nM) that shows the predicted binding affinity in nanomolar units and %Rank which represents the rank of the estimated affinity that is compared to 400,000 random natural peptides; strong binders are scored as less than 0.5% and weak binders are less than 2%[62,63]. In our experimental setup, epitope prediction was done by assessing binding capacity of the peptides to most common classical MHC class I alleles H2-Kb and H2-Db; that are also present in C57/BL6 mice. Furthermore, the FSPs that have a SYFPEITHI score of more than 15 and netMHC score of less than 5 were selected as candidates in order to analyze the immunogenicity.

3.6.3 Immunization with FSPs

All frameshift peptides were dissolved in DMSO to have a final concentration of 25 mg/ml. The peptides, which did not dissolve in DMSO, were further diluted in DNase-RNase free PBS to 10 mg/ml. The peptides were diluted to 5 mg/ml in PBS to further

use in the experiments and stored at -20 ^{0}C for storage. OVA SIINFEKL (257-264) and OVA (323-339) were dissolved in DNase-RNase free water to final concentration of 4 mg/ml and mixed 1:1 in the experiments to have a final concentration of 2 mg/ml and stored at -20 ^{0}C. CpG ODN 1826 was resuspended in DNase-RNase free water to 2mg/ml solution (both stock and working concentration) and stored at -20 ^{0}C.

The vaccination formula included 50 µg of each peptide mixed with 20 µg CpG ODN 1628. For vaccination, C57/BL6 mice injected biweekly for 4 times and at the end of the 7th week, the mice were sacrificed and Elispot and Elisa were carried out.

3.7 Preparation of sera

Blood samples were taken from heart blood of the mice into 1.5 ml reaction tubes and incubated at room temperature until the blood clots, then the samples were centrifuged for 10 minutes at 3,000 rpm at 4 ^{0}C, and the supernatants were collected and stored at -80 ^{0}C for further analysis.

3.8 Elispot Protocol

Day 1

Elispot plates (Milipore MSIPS4510) were activated by adding 70 µl of 70 % sterile ethanol per well and incubating for 5 minutes. Afterwards, the plates were washed 5 times with 200 µl sterile PBS. And the wells were coated with 100µl rat anti-mouse IFNγ (BD 551216) (1:200 dilution in PBS) and the plates were incubated overnight at 4 ^{0}C.

Day 2

The primary antibody was discarded from the plates, and washed 4 times with 200 µl sterile PBS. PBS was discarded and plates were blocked with 200 µl RPMI 1640 with 10% FBS for 1-2 hours at 37 ^{0}C. The spleens from the mice were taken out and

mashed on 6 well plates with cold 3 ml RPMI 1640 with 10%FBS. The cells were collected from the 6 well plates into 15 ml falcon tubes and centrifuged at 1200 rpm for 5 min. Then, the splenocytes were washed 2 times with media, resuspended with 1 ml ACK buffer and incubated for 10-15 min at room temperature. Following red blood cell lysis, the cells were centrifuged one more time and the cells were counted.

From this step on depending on the experiment the following procedures was followed 1) IFNγ Elispot Protocol or 2) CD4 and CD8 IFNγ Elispot Protocol.

3.8.1 IFNγ Elispot Protocol

After discarding the medium from the plates, 100µl cell suspension-containing $1,5x10^6$ cells were added to each well of a 96 well plate. Following addition of cells into 96 well plates, peptide dilutions were prepared as 2 µg/ well, dissolved in 100 µl media per well and was also added on the wells. The plates were incubated for 16-20 hours at 37 ^{0}C in CO_2 incubator.

3.8.2 CD4 and CD8 IFNγ Elispot Protocol

After discarding the medium from the plates, the cell number of the naive mice was adjusted to $4x10^5$ cells/well of a 96 well plate (these cells will be used as antigen presenting cells) and $4x10^5$ cells per well in 80 µl and 2 µg of peptide/well in 20 µl were added to the wells. The plates were incubated at 37 ^{0}C in incubator for 4 hours for antigen processing. The spleens from the vaccinated mice were taken out, and followed by isolation protocols with CD4 and CD8 T cell isolation kits (MACS 130-104-454 and 130-104-075, respectively) The spleens were mashed on ice and precooled solutions were used all the time for isolation. The cells were washed 2 times with sterile PBS and the cell pellet was resuspended in 40µl of MACS buffer for 10^7 cells. (0.5% BSA+ 2mM EDTA in PBS) 10µl of Biotin-Antibody cocktail were added per 10^7 cells and mixed well and incubated 5 min at 4^0C. On the top of this mix, 30µl of MACS buffer per 10^7cells were added, mixed again and incubated for 10 minutes at 4^0C. The magnetic separation was carried out by placing the LS column in the magnetic field of a magnetic separator and preparing the column by rinsing with 3 ml of MACS buffer. The cell suspension was placed into the column and the unlabeled $CD4^+$ or $CD8^+$ T cells were collected in 15 ml tubes. The isolated T cells were counted and $1x10^5$ of isolated CD4 or CD8 T cell per well (in 50µl) from the vaccinated mice into the corresponding wells with the peptide of interest were added into wells and incubated for 16 hours at 37 ^{0}C in the incubator.

The protocol from this step on continued the same for both the general IFNγ Elispot and CD4 and CD8 IFNγ Elispot Protocol.

Day 3

The cells were removed and the plates were washed 4 times with 200 µl of PBS 0.01% Tween and 1 time with 200 µl PBS. 100µl biotin coupled rat anti-mouse IFNϒ per well was added into wells (BD 554410) (1:500 dilution in PBS) and incubated for 1-2 hours at room temperature. The secondary antibody was removed and the plates were washed 6 times with 200 µl PBS 0.01% Tween and 1 time with PBS. 100µl streptavidin-ALP per well (BD 554065) (1:500 dilution in PBS) was added into wells and incubated 30 min at room temperature in dark. The plates were washed 3 times with 200 µl PBS 0.01% Tween and 3 times with PBS. 100 µl per well BCTP/NBT substrate (Sigma B1911) was added into wells and incubated in dark at room temperature until the plates are developed (15 min-30 min). The reaction was stopped under running distilled water. The spots were counted using an Elispot counter Immunospot.

3.9 Peptide specific total IgG Elisa

96 well (Nunc) Elisa plates were coated with 50 µl peptide solution containing 2 µg of an immunogenic FSP in PBS and incubated overnight at 4^0C. Following discarding the peptide solution from the wells, the plates were washed with 200 µl PBS/0,05% Tween for 3 times and nonspecific binding was blocked by addition of 100 µl per well 0,5% casein in PBS/tween into wells and incubation for 1 hour at room temperature. Following blocking, 50 µl sera of the immunized mice were added into wells in different dilutions and incubated for 1 hour at room temperature. Then, the plates were washed 4 times with PBS/tween, dried, 50 µl/well (1:5000) secondary antibody HRP coupled anti-mouse IgG antibody was added into wells and incubated for 1 hour at room temperature. The plates were washed 4 times with PBS/tween again. And following drying of the plates, TMB substrate was added 50 µl per well and incubated for 30 minutes at room temperature. Lastly, in order to stop the reaction 50 µl 1M H_2SO_4 was added into wells. The color formation was measured at 405 nm in Elisa reader with a reference wavelength 620 nm.

3.10 Flow cytometry

3.10.1 Staining of MACS isolated CD4 and CD8 T cells

After magnetic isolation of CD4 and CD8 T cells, in order to confirm the populations, the isolated cells were washed with FACS buffer and stained with 100 µl per well cocktail of antibodies dissolved in FACS buffer; for anti-mouse CD3 PECy7, anti-mouse CD4 FITC, anti-mouse CD8 PE and aqua zombie AM Cyan for live/dead staining. And incubated for 30 minutes in dark at room temperature. Following staining, samples were washed 2 times with FACS buffer and fixed with Cytofix/Cytoperm solution by resuspending cells in 100 µl and incubation for 20 min at 4 ^{0}C. After fixation, cells were washed with perm/wash buffer, resuspended and analyzed by FACSCanto II.

3.10.2 MHC class I cell surface staining

The cell lines were grown until 80% of confluency and the cells were harvested from flasks. The cell lines that were treated with IFNγ were treated before for 48 hours with 500 U/ml IFNγ. The cells were counted and 0.5x10^{6} cells were transferred into 1.5 ml tubes. Three vials were prepared for each cell line that contains unstained, secondary antibody isotype mouse anti-human IgG-FITC and unlabeled MHC class I primary antibody labeled with secondary antibody mouse anti-human IgG-FITC. FACS buffer containing 1% FCS in sterile PBS were prepared and used in all of the washings and stainings. For staining, after the cells were collected in 1.5 ml tubes, the cells were centrifuged, washed twice with FACS buffer, stained with primary antibody for 30 minutes on ice, washed 2 times with FACS buffer and labeled with secondary antibody for 30 minutes on ice. Following two times washing, the pellets were resuspended in FACS buffer and analyzed by using FACS Calibur.

3.11 Protein Techniques

3.11.1 Cell lysis

Cell pellets were washed with PBS and centrifuged at 1200 rpm for 5 minutes. Following resuspension of the cell pellet in 50-200 µl RIPA buffer with 1% protease inhibitor (according to the number of cells), the samples were sonicated for 10 seconds at 30 % on ice and incubated on ice for 30 minutes. After incubation, the samples were centrifuged at 13000 rpm for 15 minutes at 4 ^{0}C and the supernatant containing the proteins from the cell lysate was transferred into 1.5 ml tubes and stored at -20 ^{0}C for further use.

3.11.2 Measurement of protein concentration by Bradford assay

The samples were diluted 1:10-1:20 depending on the cell pellets and 20 µl of DNase RNase free water; BSA standards (0, 0.125, 0.25, 0.5, 0.75, 1.0 and 1.5 mg/ml) and the samples were all prepared in duplicate. Following addition of 1ml Bradford solution (1:5 in DNase RNase free water) to all standard and samples and 5 minutes of incubation, the Optical density was measured at 595 nm. According to the curve and the equation from the measurements, the concentration of each sample was calculated.

3.11.3 SDS-PAGE

Samples were prepared to contain 65% of the sample (containing 20 µg), 25% of Laemmli buffer and 10% DTT as reducing agent. After they were prepared as specified, the samples were denatured for 5 minutes at 99 ^{0}C in a PCR cycler with a heated lid. While they were denatured, Bio-rad SDS page system was prepared. First, 10x SDS running buffer was diluted with ddH_2O to 1x and ready to use NuPAGE Novex 4-12% Bis-Tris gel was placed in the gel chamber and the inner part was filled with 1x SDS buffer, the combs were removed and the gel pockets were cleaned with a syringe. Next, samples (20-50 µl) and the Bio-rad precision plus protein standard protein ladder (7 µl) were loaded in the wells and the outer part of the chamber was

also filled with SDS buffer. The gels were run at 35 mA per gel for 30 minutes.

3.11.4 Semidry blotting

The three buffers that are required for blotting were prepared in the beginning. Buffer 1 contains 36.33 g of Tris base dissolved in 600 ml of ddH2O and adjusted to pH 10.4, after addition of 200 ml methanol, it was completed up to 1 L with ddH_2O. The buffer 2 consists of 3.03 g Tris base dissolved in 600 ml ddH2O and adjusted to pH 10.4, after addition of 200 ml of methanol, ddH_2O was added up to 1 L. Buffer 3 includes 5.25 g Norleucin and 3.03 g Tris base, dissolved together in 600 ml ddH_2O and adjusted to pH 9.4, after addition of 200 ml methanol, ddH_2O was added up to 1 L. 10 pieces per gel whatman papers were cut according to the size of the gels and 2 papers were incubated in buffer 1, 3 papers were incubated in buffer 2 and 5 papers were incubated in buffer 3, respectively for 10 minutes. Nitrocellulose membrane was cut according to gel and incubated for 1 minute in methanol, 5 minutes in ddH2O and 10 minutes in buffer 2. The blots were prepared as follows, avoiding making bubbles in between, 2 whatman papers soaked in buffer 1 on the bottom, 3 whatman papers soaked in buffer 2 on the top, nitrocellulose membrane soaked in buffer 2 onto the pile, the gel in buffer 3 and 5 whatman buffer soaked in buffer 3 at the very top. Then the system was run for 1 hour at 230 mA (12-17V). After the run was completed, the nitrocellulose membrane was stained with ponceau S staining for 20 seconds and the excess dye was washed away with ddH_2O.

3.11.5 Western blot

Western blot was carried out using Western Breeze Chemiluminescent immunodetection kit. First, the membrane stripping solution was heated at 37 ^{0}C. The membrane was washed 2 times for 5 minutes with 10 ml PBS. The membrane was incubated with preheated stripping solution for 30 minutes at rolling shaker. The membrane was washed again and western breeze chemiluminescent immunodetection reagent was used according to manufacturer´s instructions to stain the membrane with the antibody of interest. The chemiluminescent substrate was detected using Bio-rad ChemiDoc MP system.

4 RESULTS

4.1 Setting up a mouse model and characterization of immune phenotype

4.1.1 Selection of frameshift peptides

In order to select possible FSP-encoding genes containing cMSs, the protocol that was explained in [40] was applied. After genome-wide screen of mouse cMS database, the candidate 53 cMS-bearing genes were analyzed to check the mutation frequency of the genes; the genes which have more than 15% of mutation frequency and monomorphic in normal tissue were selected Table 4.

Repeat	Mutation frequency
Ncad	75%
Xirp1	37.50%
5730596B20Rik	37.50%
Rif1	33.30%
Maz	33.30%
Hic1	31.25%
Sdccag1	25%
Tmem107	25%
Srcin1	25%
Marcks	20%
Senp6	18.75%
Phactr4	18.75%
Chrnb2	18.75%

Table 4 Mutation frequency of the selected cMS bearing genes. The genes that contain coding microsatellites were analyzed by cMNR length analysis utilizing the paraffin blocks of mouse tumors, and the genes that have more than 15% of mutation frequency and monomorphic in normal tissue were sorted out. The first column corresponds the names of the cMS containing genes and the second column corresponds to the mutation frequency that was calculated from mouse tumors.

The 13 selected genes that have more than 15% of mutation frequency and monomorphic in normal tissue were further analyzed if the genes are expressed in both normal tissue and tumor. All cMS-encompassing genes were found to be expressed in MSI tumors from Lynch mice and corresponding normal mucosa. One

representative expression analysis result is shown in figure 5.

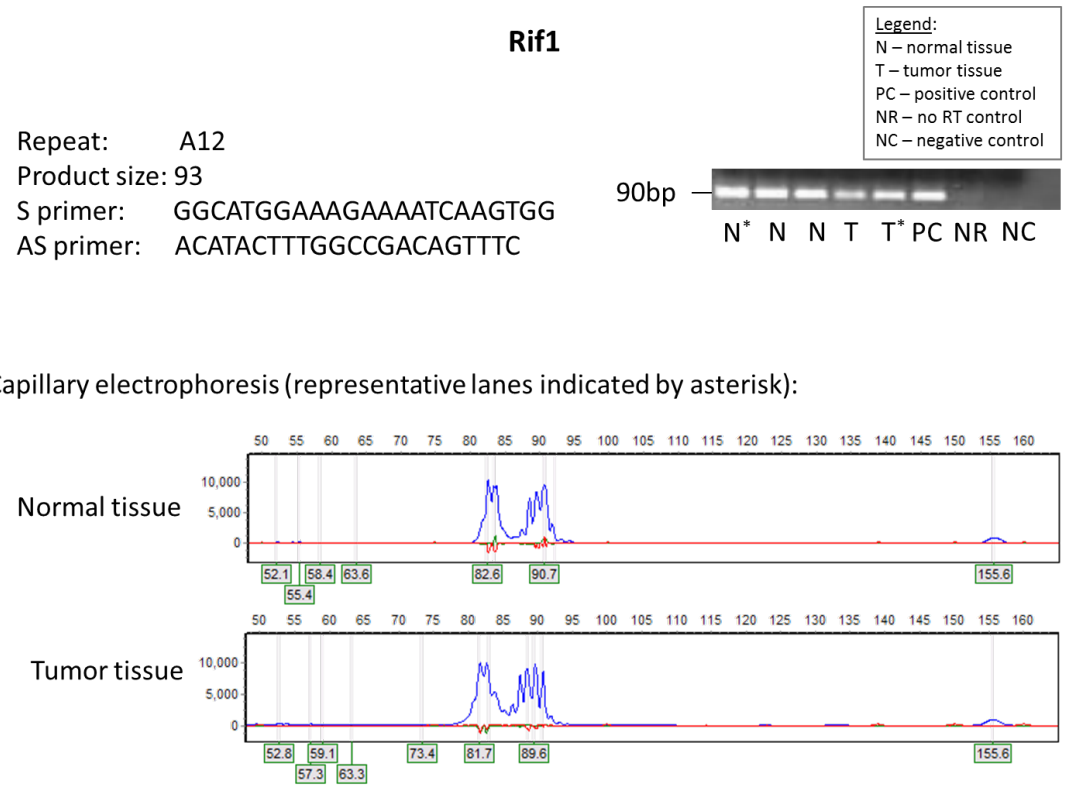


Figure 5 Representative gene expression analysis of the cMS-bearing genes in healthy and tumor tissue by RT-PCR and capillary electrophoresis. The expression of the genes that were sorted out in the first step was determined by RT-PCR in normal and tumor tissue. Rif1 is one of the genes analyzed; S primer is sense primer and AS primer is anti-sense primer that was used for RT-PCR; the expression of the genes are shown in the agarose gel image; legend for the gel is depicted next to the gel image.

After it was confirmed that all candidate cMS bearing genes were expressed in both normal tissue and tumor, I took over the project for the immunogenicity testing of the 13 genes and 26 candidate FSPs and started with the epitope prediction for the FSPs.

Epitope prediction was performed using SYFPEITHI and netMHC, being independent tools available publicly [61,63]. For this purpose, whole FSP sequences encompassing 8 N-terminal wild type-derived amino acids were screened in order to cover the potential immunogenic epitopes that might be located at the junction site between the wild type and mutated regions. Furthermore, since Lynch mice have C57/BL6 background, the presences of immunogenic epitopes that can bind to H2-Kb/Db were evaluated. The criteria for epitope prediction for either one base (-1) or two base frameshift (-2) mutation of the genes were having a SYFPEITHI score of more than

12 and netMHC score of less than 5% (Table 5). These non-stringent cutoffs were selected to ensure that at least 10 candidate FSP sequences fulfilled the criteria. As a result of the database prediction and the mutation frequency data, 10 most promising candidate FSPs (shown in yellow in table 5) were selected for immunogenicity testing.

FSP (-1)

Repeat	FSP (-1)	FSP (-1) freq	SYFPEITHI score	netMHC Db	netMHC Kb	netMHC % rank
Ncad	VIYAPPPPAEGRWPCWLLRAH	56.25%	15	s0 w0	s0 w0	3
Xirp1	GKGPGGPPLSSPKRVMYRLSVGCLRPTL	31.25%	19	s0 w2	s1 w4	0.4
5730596B20Rik	GTLPPPPPTQH	37.50%	6	s0 w0	s0 w0	39
Rif1	–	20.00%	–	–	–	–
Maz	PCTLLAPPSPCWAWTPGGWAAS	20.00%	7	s0 w0	s0 w0	9.5
Hic1	DRTFPSPPRIGAI	18.75%	14	s0 w0	s0 w0	5
Sdccag1	EAPKGKKKSKRTSSCRSRRRTSRCL	25.00%	12	s0 w0	s0 w0	8
Tmem107	TQYFGMGGVVENRSQI	25.00%	24	s2 w1	s0 w0	0.5
Srcin1	DEGMWPPPTTS	25.00%	6	s0 w0	s0 w0	31
Marcks	SSETPKKKRSAFPSRSPSS	20.00%	12	s0 w0	s0 w0	5.5
Senp6	VKCSMKKKIMLSMKMKNQVTENLRARTFVIEPKVRMASGMNASVLYIIQMP	18.75%	27	s10 w12	s0 w2	0.03
Phactr4	PWKWRKKKAVISSKRHQKF	18.75%	12	s0 w0	s0 w0	3.5
Chrnb2	VRTRPSPPHLSPASWVLKPFAINAKGIFLILCGNWQQGCLCHLGMHLRHRQVGL	12.50%	21	s1 w3	s0 w1	0.5

FSP (-2)

Repeat	FSP (-2)	FSP (-2) freq	SYFPEITHI score	netMHC Db	netMHC Kb	netMHC % ra
Ncad	DVIYAPPPQQRGGGRAGYSERIDGQRDRETGVSAGTRPGHARGGCGR	37.50%	14	s0 w0	s0 w0	6
Xirp1	–	6.25%	–	–	–	–
5730596B20Rik	LGTLPPPPQPSTEQSGWKHHQ	0.00%	7	s0 w0	s0 w0	55
Rif1	AHTKDKKKSETVGQTETRIFISKNKEW	13.33%	15	s0 w0	s0 w0	4.5
Maz	FPCTLLAPLPRAGPGLPGGGRPHELLPATSGSRPEPPAGRG	13.33%	17	s0 w0	s0 w0	3.5
Hic1	DRTFPSPPELARYNI	12.50%	18	s0 w0	s0 w0	4.5
Sdccag1	EAPKGKKKAKEQAAAEAAEEQAAACRCGSQPVSLCQCQKIL	6.25%	16	s0 w0	s0 w0	2.5
Tmem107	TQYFGMGGWWKIDPKSEGFPHLDLSCTCEIGRVLKSHTHPNPPAYKVWRFKCLG	6.25%	21	s0 w0	s1 w2	0.25
Srcin1	VDEGMWPPQQPPEPVPQEGGS	0.00%	12	s0 w0	s0 w0	22
Marcks	SSETPKKKEALFLQEVLQAERLLLQEEQEGVGRGR	0.00%	16	s0 w3	s0 w0	0.8
Senp6	VKCSMKKKSCYQ	0.00%	1	s0 w0	s0 w0	20
Phactr4	PWKWRKKKQ	0.00%	2	s0 w0	s0 w0	75
Chrnb2	VRTRPSPPISLQPHGS	6.25%	16	s0 w0	s0 w0	11

Table 5 Epitope prediction of the selected FSPs by using Syfpeithi and netMHC as tools. The sorted genes that bear cMSs were assessed for one and two base frameshift mutations called FSP (-1) and FSP (-2), respectively by using Syfpeithi and netMHC as prediction tools. Syfpeithi score of more than 12 and netMHC score of less than 5 filtered out and the genes selected are marked with yellow in the table.

The results of the cMS mutation analysis described above suggested that these ten FSPs could potentially cover 12 out of 16 tumors analyzed, corresponding to 75% vaccine coverage.

4.1.2 Immunogenicity of the FSPs

The immunogenicity of the ten FSPs was evaluated by biweekly immunization of C57BL/6 mice with FSP antigens for four times, using CpG ODN 1826 (TIB MolBiol, Berlin, Germany) as an adjuvant. At the end of the seventh week, mice were sacrificed and IFNγ Elispot analysis was done.

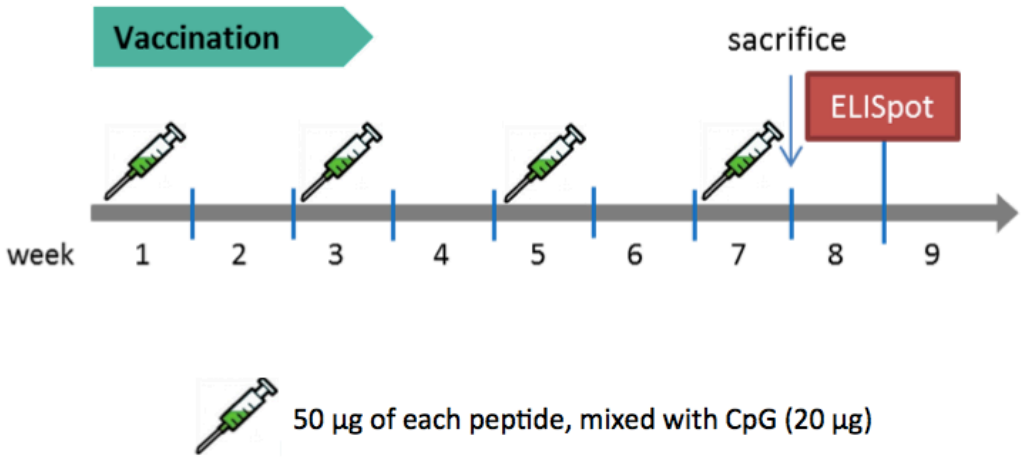


Figure 6 FSP vaccination protocol. 50 µg of each peptide was mixed with 20 µg of CpG ODN 1826, and subcutaneous injection was done 4 times every two weeks. At the end of the seventh week C57/BL6 mice were sacrificied and ELISPOT analysis was carried out.

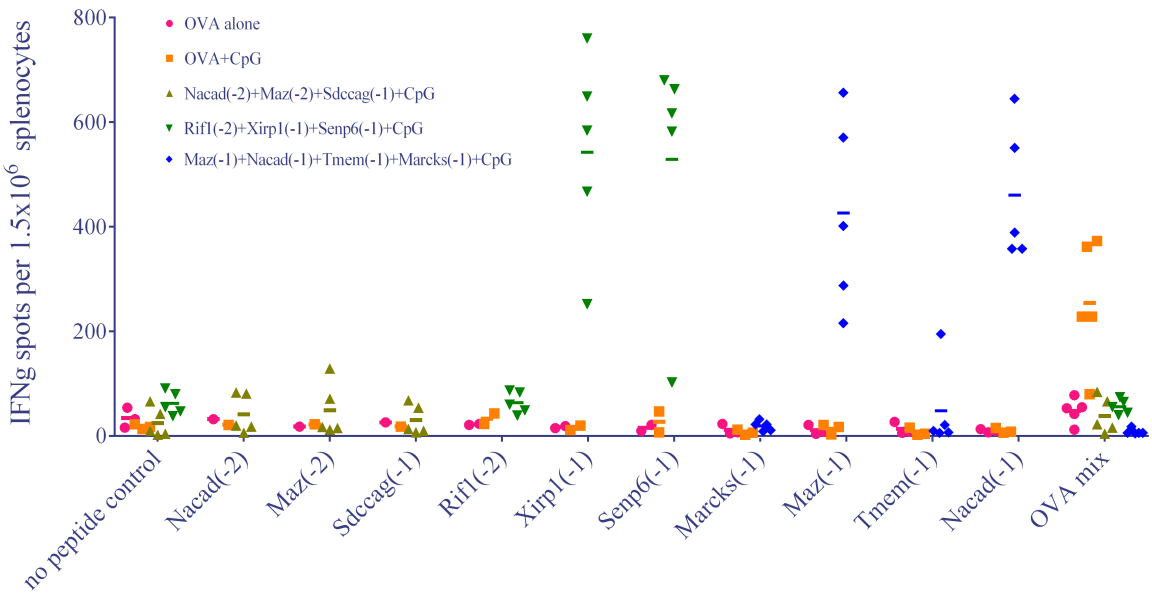


Figure 7 Immune responses against the ten selected candidate FSPs. C57BL/6 mice were vaccinated with the FSPs biweekly using CpG ODN 1826 as an adjuvant for four times. At the end of the seventh week, mice were sacrificied and IFNγ Elispot analysis was performed. Elispot plates were coated with recombinant IFNγ, after blocking step with 10%FBS supplemented culture medium, $1.5x10^6$ splenocytes were added per well together with the peptide of interest and incubated at 37°C for 16-20 hours. After the incubation period, biotinylated-IFNγ was added on the wells and the reaction was developed by the substrate addition. The spots were counted by Elispot counter (Immunospot). The average number of spots per each mouse is shown for each peptide. OVA mix corresponds to the mix of the OVA peptide of CD4 and CD8 epitope, which was used as a control in this experimental setup. The x-axis shows the names of the peptides that were used and the y-axis shows the number of cells, which secrete IFNγ per $1.5X10^6$ splenocytes.

Four out of ten FSPs (Xirp(-1), Senp(-1), Maz(-1) Nacad(-1)), that were applied together with CpG ODN 1628, induced immune response in C57BL/6 mice (figure 7). In the next step, our aim was to verify if the characterized immunogenic peptides would also induce immune response when applied together in one vaccine. When the four immunogenic FSPs were applied together, they retained the capability to induce immune response confirmed by IFNγ Elispot (figure 8).

	Size (aa)	Mutation freq of the gene(%)	FSP(-1) freq (%)	Sypeithi score	net MHC% rank
Maz(-1)	22	33,3	20	7	9,5
Nacad(-1)	21	75	56	15	3
Xirp(-1)	28	37,5	31	19	0,4
Senp6(-1)	52	18,75	19	27	0,03

Table 6 Characteristics of the four immunogenic peptides. Size of the peptides are shown in the number of amino acids in the peptide sequence in the second column, mutation frequency of the corresponding genes of the frameshift peptides are shown in the third column, the frequency of the frameshift peptides are followed by the SYFPEITHI and netMHC scores that are represented in the last two columns color coded with predicted high immunogenicity as orange.

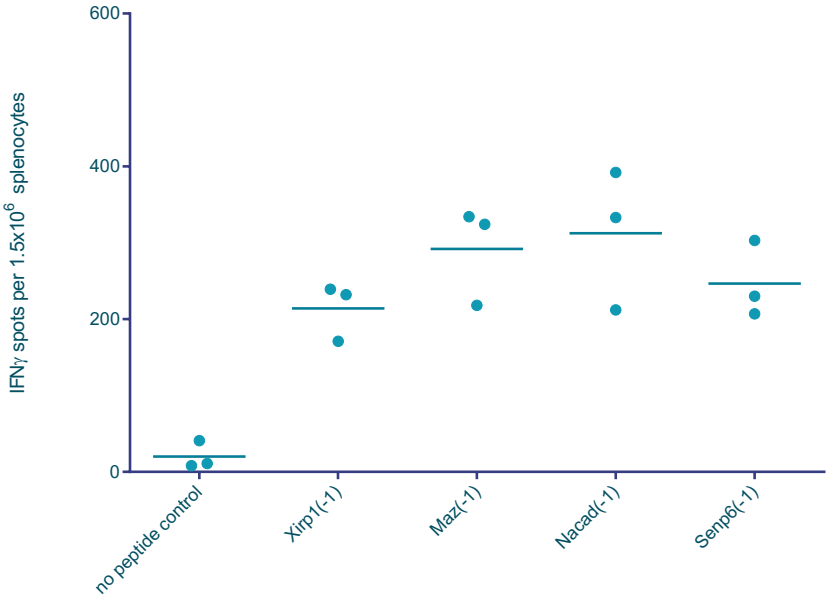


Figure 8 Immune responses against four immunogenic FSPs applied altogether in one vaccine. Following the four times biweekly immunization protocol of the mice, IFNγ Elispot analysis was done to detect immune response against FSPs when applied together in one vaccine; maintaining high immune response. The x-axis corresponds to the name of the peptides and the y-axis corresponds to the number of splenocytes that secrete IFNγ upon antigen specific stimulation per 1.5X10^6 splenocytes. The average number of spots is shown for every mouse.

Due to the fact that all of the FSPs used in this study share 8 N-terminal amino acids with the wild type peptide, our purpose in the next step was to confirm that there was no cross-reaction with the wild type peptide sequence. In order to accomplish this goal, wild type peptides were synthesized for the corresponding peptides and immunogenicity was tested. Figure 9 presents that none of the FSPs tested had cross-reaction with the wild type and the induced immune response is FSP-specific.

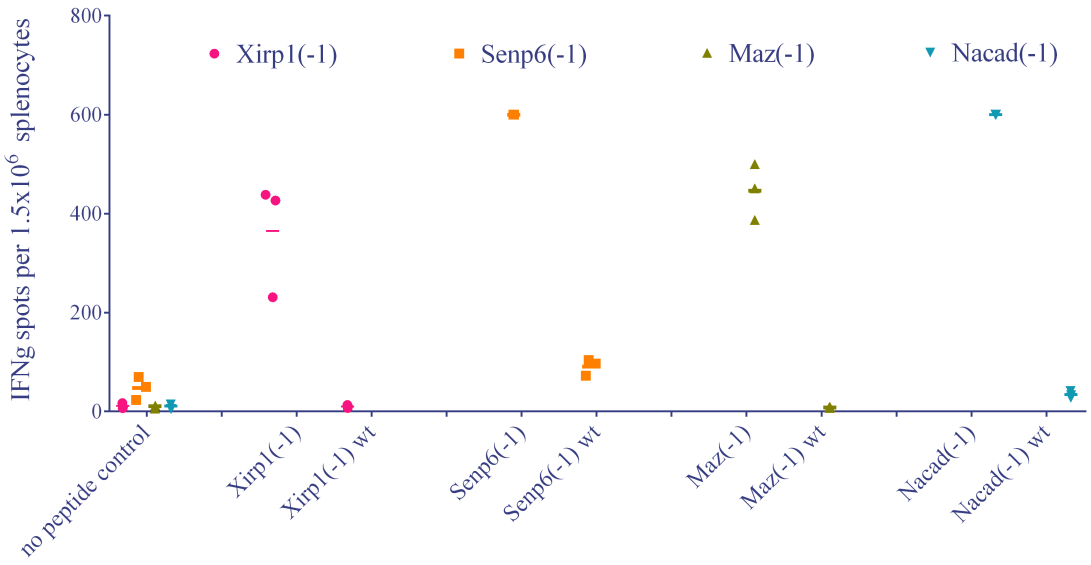


Figure 9 Absence of cross-reaction with wild-type peptides. After immunization of the mice as previously described. IFNγ Elispot was carried out with the wild type peptide and the corresponding FSP of interest. The average number of cell numbers that are secreting IFN**γ** per 1.5X10^6 cells per mice is shown in the y-axis and the FSPs and wild-type peptides are shown in x-axis.

In order to further characterize the immune response induced by FSPs in detail, another round of immunization was carried out with the same experimental setup and induction of CD4 and CD8 T cell response was analyzed by Elispot. As it is seen in figure 10, two of the FSPs namely Maz (-1) and Senp (-1) only induced CD4 T cell respose; whereas, Xirp (-1) induced mostly CD8 T cell response and Nacad (-1) induced both CD4 and CD8 T cell response following immunization of C57BL/6 mice with FSPs of interest with CpG ODN 1628.

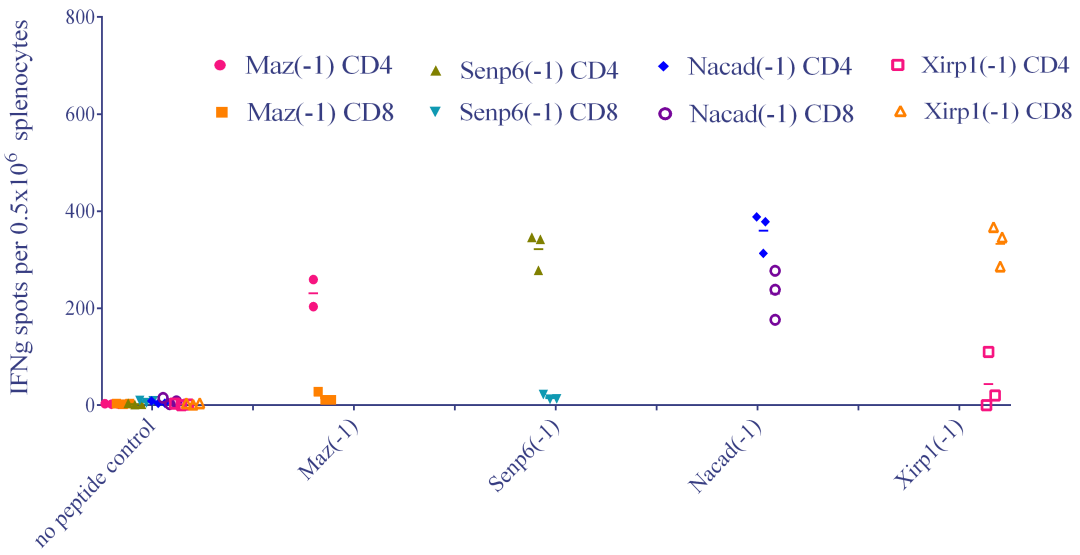


Figure 10 CD4 and CD8 T cell responses against four FSPs. Elispot plates were coated with recombinant IFNγ, and then the peptides were incubated for 4 hours at 37^{0}C together with 4x10^5 splenocytes per well for antigen processing, at the end of incubation, 1x10^5 magnetic separated (from spleen of immunized mice) CD4 or CD8 T cell were added per well. Following addition of biotinylated-IFNγ antibody, substrate was added and spot numbers were scored by Elispot counter (Immunospot). Average spot numbers are presented for each peptide.

The regions, in which the immunogenic epitopes might be located, were described by synthesis of the shorter peptides of interest from N- and C-terminus of the FSPs; called FSP (-1) N and FSP (-1) C. The sequences are shown in the figure 11A and the immune response against (-1) N and (-1) C is shown in figure 11B.

A

peptide ID	sequence	length
Nacad wt	VIYAPPPPSRGAVAVLATPS	20
Nacad (-1) N	VIYAPPPPAEGRW	13
Nacad (-1) C	AEGRWPCWLLRAH	13
Xirp1 wt	GKGPGGPPPELPKKGDVQTI	20
Xirp1 (-1) N	GKGPGGPPLSSPKRVMYRLS	20
Xirp1 (-1) C	VMYRLSVGCLRPTL	14
Maz wt	PCTLLAPPFPVLGLDSRGVG	20
Maz (-1) N	PCTLLAPPSPCWAW	14
Maz (-1) C	SPCWAWTPGGWAAS	14
Senp6 wt	VKCSMKKKNHAINENEEPSN	20
Senp6 (-1) N	VKCSMKKKIMLSMKMKNQVTENLRA	25
Senp6 (-1) C	NLRARTFVIEPKVRMASGMNASVLYIIQMP	30

B

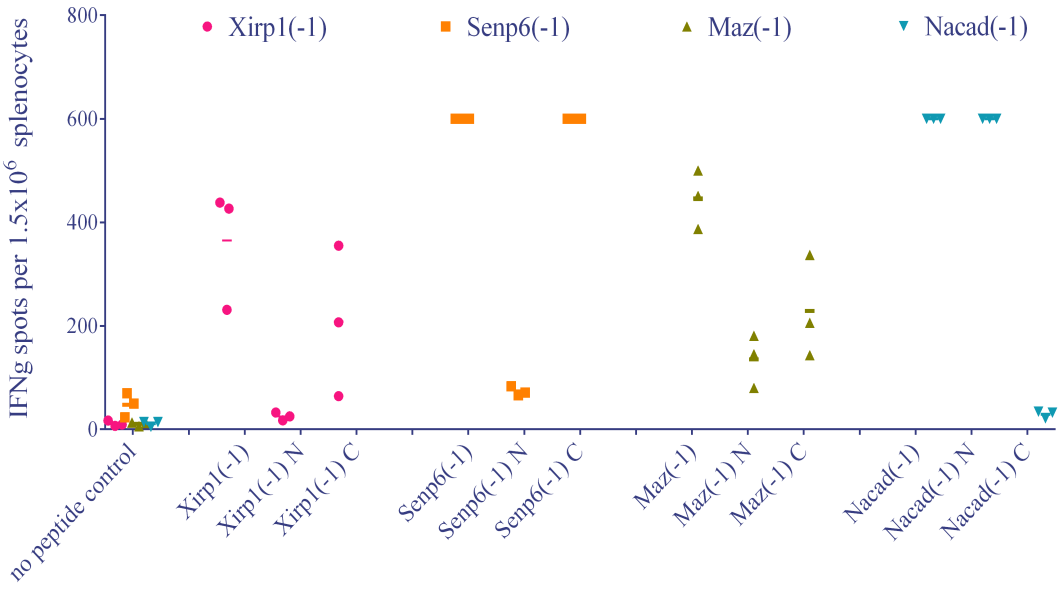


Figure 11 Orienting epitope mapping of the immunogenic peptides. The peptides shown in part A were synthesized for each peptide and the immunogenicity for the regions where epitopes might be located was determined by IFNγ Elispot. Elispot plates were coated with recombinant IFNγ, then after blocking step with 10%FBS supplemented culture medium, 1.5x106 splenocytes(from vaccinated mice) per well was added into wells together with the peptide of interest. After addition of secondary biotinylated-IFNγ antibody and substrate addition, the cell numbers for every well was counted with Elispot reader (Immunospot). The average number of spots per mice are shown in the graph(B).

Characterization of the immunogenic regions of the FSPs by utilizing the shorter peptides that are synthesized from N- and C-terminus of the FSPs manifested that the immunogenic part of Xirp(-1) and Senp(-1) are in the C-terminus of the FSP; whereas, Nacad(-1) is immunogenic in the N-terminus of the FSP and Maz(-1) is immunogenic as whole in both parts of the FSP.

4.1.3 Humoral response

In order to characterize the induction of humoral immune response by FSPs, peptide specific IgG Elisa was carried out by using the serum from the immunized mice. All of the FSPs except from Xirp(-1) induced humoral immune response(figure 12): this response also might relate with the data before that Xirp(-1) induced only CD8 T cell but no CD4 T cell response (figure 10).

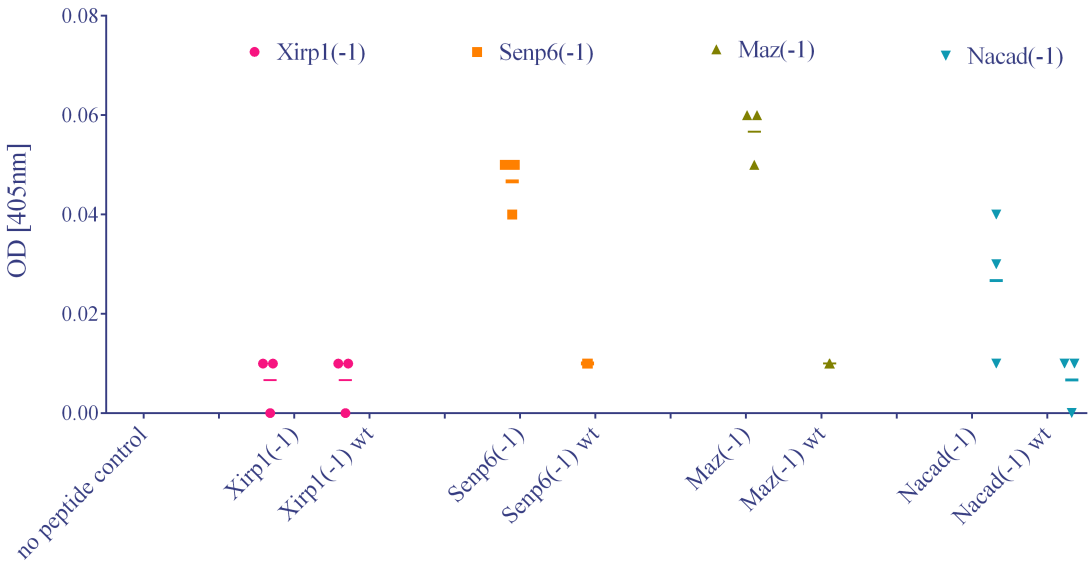


Figure 12 Humoral immune responses induced by FSPs. After immunization of the mice with FSPs for four times biweekly, mice were sacrificed and the heart blood was taken out. After centrifugation of the blood, serum was collected. 96 well Elisa plates were coated with the peptide of interest and blocked with casein solution, afterwards sera of the corresponding mice was added into wells. Following secondary antibody HRP coupled anti-mouse IgG was added into wells and developed by TMB substrate. Lastly, H_2SO_4 was added to stop the reaction and the optical density was measured at 405nm in Elisa reader with a reference wavelength of 620 nm. The average OD value is shown for every mouse.

4.1.4 CD4 T cell infiltration in the tumors of Lynch mice

To detect the tumor-infiltrating lymphocytes in the tumor tissue of mouse model of Lynch syndrome, a staining protocol was established allowing the quantification of CD4-positive T cells in murine tissue sections. A rat monoclonal anti-mouse CD4 antibody (clone: GHH4) was used as a primary antibody in IHC. As it can be seen from the stainings in figure 13, the number of CD4-positive T cells infiltrating the tumor tissue was variable; in some of the tumors infiltration was quite high, in others, there

was a low infiltration of CD4-positive T cells.

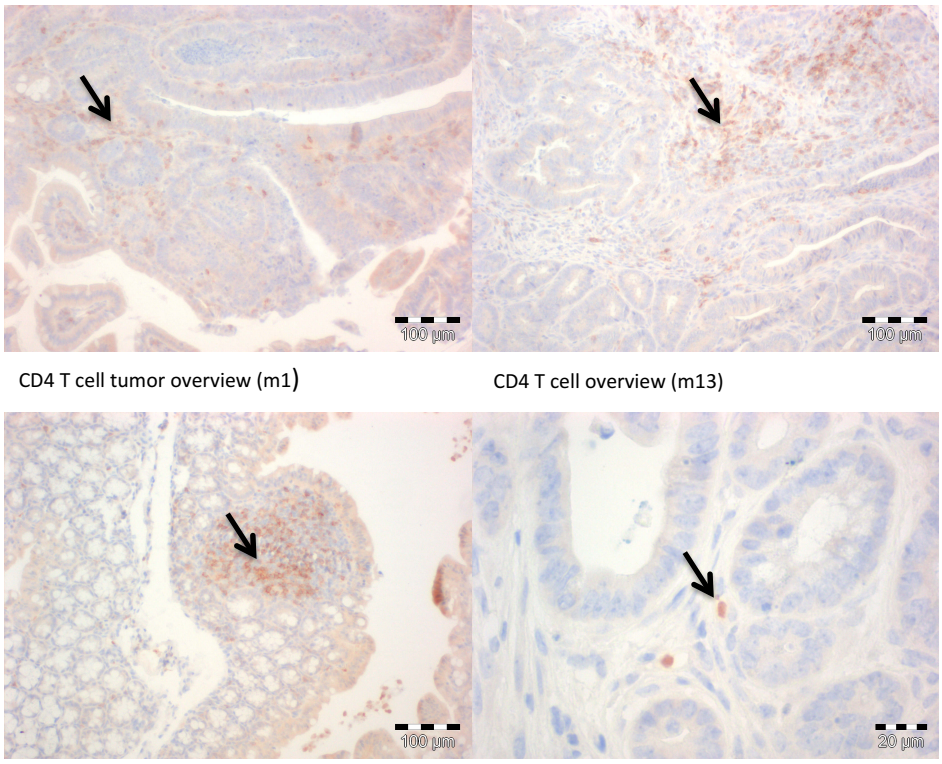


Figure 13CD4 T cell infiltration in tumors of mouse model of Lynch syndrome. Sections of mouse tumor specimens were deparaffinized and hydrated. Then antigen retrieval was done by 10mM citric acid monohydrate solution for 35 minutes at 97 ^{0}C and the slides were cooled down at room temperature for 20 minutes. Following incubation of sections with 3% hydrogen peroxide, the slides were washed and nonspecific binding was blocked by serum-free protein block for 10 minutes and primary antibody rat monoclonal anti-mouse CD4 was applied on the slides. After incubation with the secondary antibody Impress anti-rat IgG for 30 minutes, DAB substrate was used to develop the slides. The black arrows show the CD4 T cells that are stained.

4.2 Comprehensive analysis of MHC-related immune evasion in MSI Cancer

In order to evaluate the immune evasion mechanisms of MSI cancers, we started by characterization of all the MSI CRC cell lines that are present in the lab.

The characterization started with typing the cells lines for *B2M* cMS mutations by Sanger sequencing; briefly, after the DNA was isolated from the cell lines as described in 3.3 and DNA sequencing was carried out for all the cell lines as explained in 3.4. 9 out of 33 cell lines identified have either homozygous or heterozygous *B2M* mutations in their cMSs; namely, 3 of the cell lines (GP2D, LoVo and K073A) have biallelic frameshift deletion; 3 of the cell lines has mutations in one copy of their *B2M*

allele and the rest of the 3 cell lines have heterozygous point mutations (Figure 14).

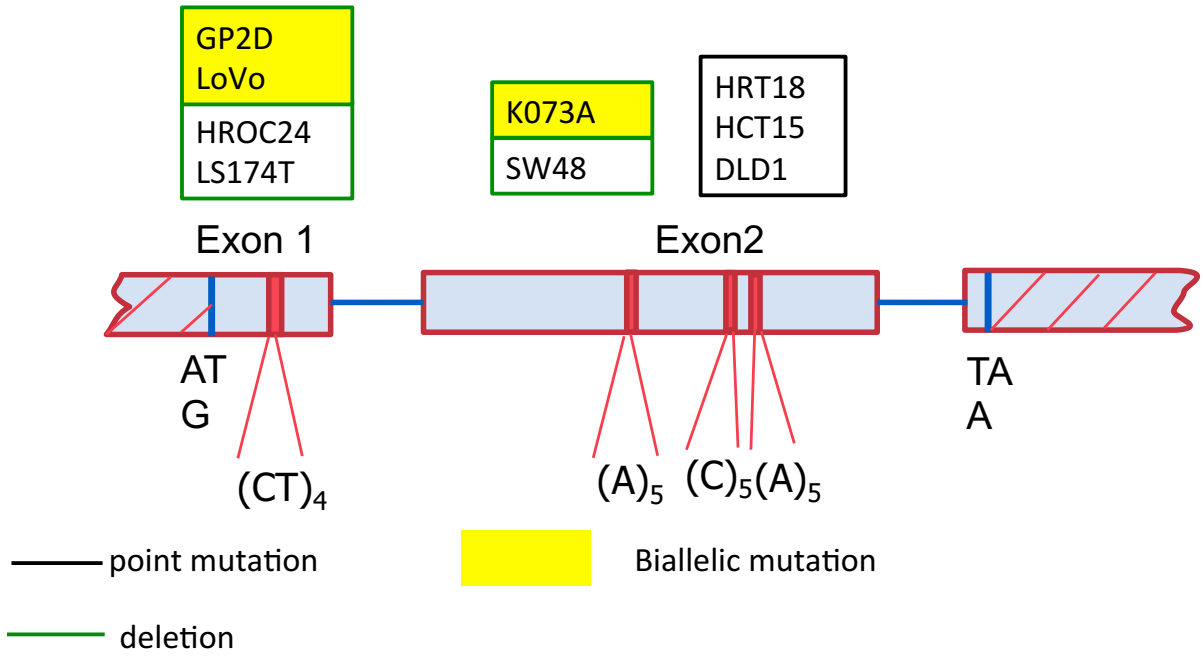


Figure 14 The characterized cell lines that have either homozygous or heterozygous B2M mutations. Following harvesting of $5x10^6$ cells from each cell line, DNA was isolated from cell lines by using DNeasy blood and tissue kit according to manufacturer's protocol, then the samples were sequenced by Sanger sequencing. The coding microsatellites of B2M was analyzed for all the cell lines; 9 out of 33 cell lines were characterized as mutant; 3 of which have mutations in both their alleles of B2M and the rest have either point mutation or deletion in their cMS in a heterozygous setting.

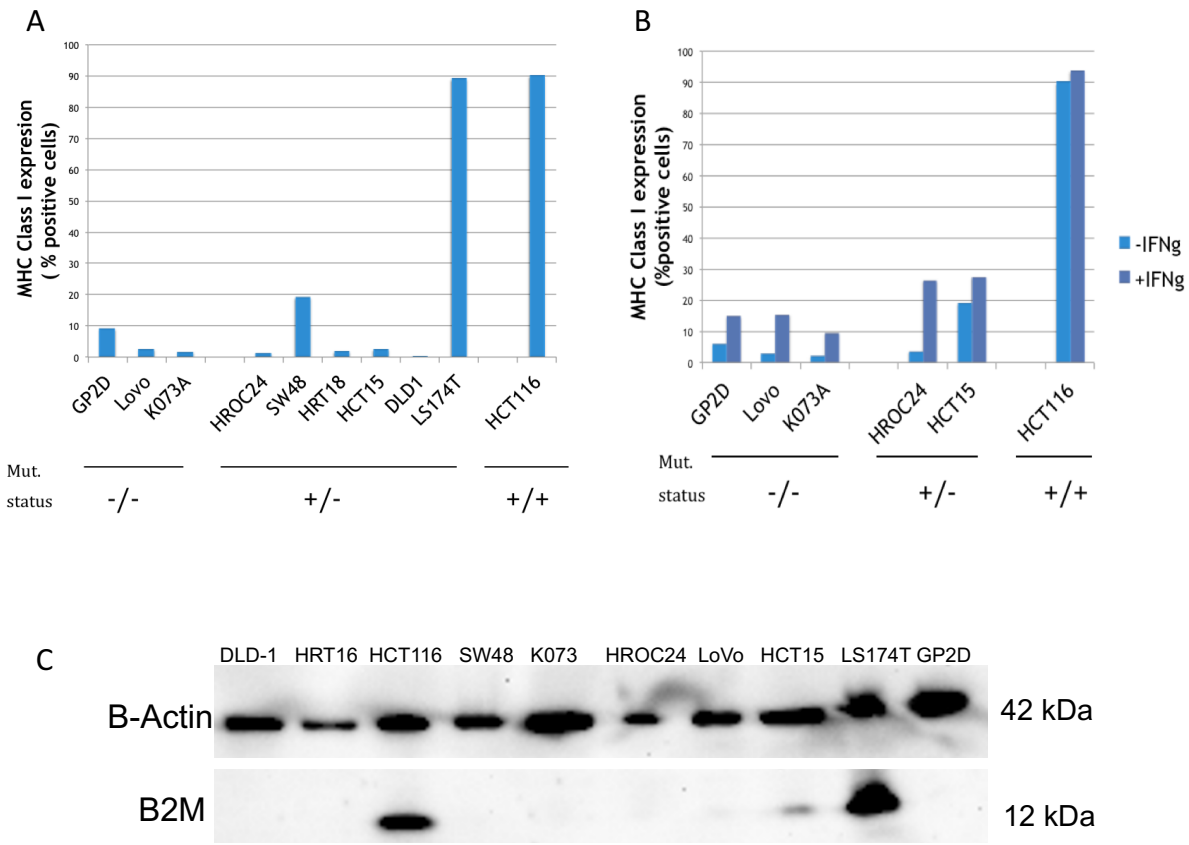


Figure 15 Flow cytometric analysis of MHC Class I expression from MSI-H CRC B2M mutant cell lines with and without IFNγ treatment (500IU/ml). (A) 0.5x10^6 cells were stained with mouse anti-humanTP25.99.8.4. (B) 0.35x10^6 cells were either untreated or treated with 500 IU/ml IFNγ for 48 hours, cells were harvested and stained with mouse anti-human TP25.99.8.4. (C) Western blot of B-Actin and B2M protein from MSI-H CRC B2M mutant cell lines.

In figure 15 A, MHC Class I staining is shown; except from 1 cell line, all the cell lines express low levels of MHC Class I. In the next step, homozygous and low level of MHC Class I expressing cell lines were chosen to check if the expression increases upon stimulation with IFNγ (Figure 15 B). After checking MHC Class I expression, western blot was done to check *B2M* in the protein level (Figure 15 C). It is seen in the blot result that only 1 cell line expresses high level of B2M, and the other cell lines have either low levels or no expression of B2M. This result correlates with the MHC Class I expression since when cells are not treated with IFNγ, only LS174T expresses high levels of MHC class I. Therefore, sequencing data correlates with protein expression data that all *B2M* mutant the cell lines except from LS174T, which has *B2M* mutation in one allele, expresses very low levels of B2M and MHC class I. Additionally, there are heterozygously mutant cell lines such as LS174T that still retained B2M and MHC class I expression.

These analyses were complemented by evaluating additional mechanisms of immune evasion in MSI cancers, interestingly only about 30% of MSI cancers have *B2M* mutation-induced loss of MHC class I antigen expression, among the remaining patients there are some showing other alterations that interfere with MHC class I and II antigen expression (i.e. *RFX5*, *CIITA*, *TAP1* and *TAP2* mutations). However, in 30% of tumors, no alterations affecting antigen presentation and recognition by immune cells have been discovered so far.

In order to obtain a more comprehensive overview of the different immune evasion mechanisms and their relative contribution in MSI colorectal cancer, we decided to examine MSI colorectal cancers contained in the DFCI database for mutations in genes related to antigen presentation [64,65].Colorectal adenocarcinoma (DFCI, Cell reports 2016) study report that contains 91 MSI-H samples was used in the analysis. Our approach was to check the mutation of the genes that are related with MHC class I antigen presentation; namely *NLRC5*, *B2M*, *TAP1*, *TAP2*, *HLA-A*, *HLA-B* and *HLA-C* in these 91 MSI patient samples. The result can be seen in the table 7. It was observed that around 72 % of MSI tumors have defects in MHC class I antigen presentation; proven by at least one mutation in the corresponding genes (Table 7). Most of the mutations observed were truncating followed by missense mutations. Moreover, *B2M* and *HLA-B* gene was significantly mutually exclusive and within all of the MHC class I heavy chain genes co-occurance of the respective genes (*HLA-A* vs *HLA-B*; *HLA-B* vs *HLA-C* and *HLA-A* vs *HLA-C*) was determined.

DFCI patient ID	B2M	NLRC5	HLA-A	HLA-B	HLA-C	TAP1	TAP2
coadread_dfci_2016_1212	Y46Cfs*10					V94A	C213*
coadread_dfci_2016_1202	X23_splice	L264P					
coadread_dfci_2016_3643	X23_splice	Q1710H					
coadread_dfci_2016_341	X23_splice						
coadread_dfci_2016_3676	X23_splice						
coadread_dfci_2016_2936	W115fs,X23_splice,L12P						
coadread_dfci_2016_3094	V69Wfs*34,L15Ffs*41,Q109*	Y281*					
coadread_dfci_2016_261	V69Wfs*34,L15Ffs*41						
coadread_dfci_2016_3152	V69Wfs*34,L15Ffs*41						
coadread_dfci_2016_2448	V69Wfs*34						
coadread_dfci_2016_55	T93Lfs*10,A8D						
coadread_dfci_2016_111	T93Lfs*10	A1769D				A263V	
coadread_dfci_2016_2944	T93Lfs*10						
coadread_dfci_2016_3263	T93Lfs*10	S1383I					
coadread_dfci_2016_1225	T24I	T224A					
coadread_dfci_2016_3080	S16Afs*27						
coadread_dfci_2016_1230	Q22*,L15Ffs*41,W115R,D96Mfs*7		T88I		G42*		
coadread_dfci_2016_3721	M1?,S16Afs*27						
coadread_dfci_2016_2379	M1?						
coadread_dfci_2016_3535	L60P,L15Ffs*41			X337_splice		V316M	
coadread_dfci_2016_407	L43Pfs*14						
coadread_dfci_2016_3181	L15Ffs*41,S16Afs*27						
coadread_dfci_2016_3024	L15Ffs*41						
coadread_dfci_2016_60	L15Ffs*41						
coadread_dfci_2016_2641	L15Ffs*41						
coadread_dfci_2016_354	L15Ffs*41						
coadread_dfci_2016_3690	L15Ffs*41						
coadread_dfci_2016_197	D96Mfs*7						
coadread_dfci_2016_1849		L865P,Q1581*					
coadread_dfci_2016_102		R371W,L1844R					
coadread_dfci_2016_1241		P1696S,Q1352H					
coadread_dfci_2016_2197		L1674P					
coadread_dfci_2016_3503		R550W,Q1360*					
coadread_dfci_2016_3701		V1652Sfs*100,S385P					
coadread_dfci_2016_649		Q1472R					
coadread_dfci_2016_3262		D479Ifs*6,R262C					
coadread_dfci_2016_1791		V499I					
coadread_dfci_2016_2257		P1740S	Y281*			C133G	
coadread_dfci_2016_1748		R1300C				A469T	
coadread_dfci_2016_3640		L627P,C648G			L102P		
coadread_dfci_2016_3064		A566V,L889P		T211Dfs*10		A2T	
coadread_dfci_2016_1762		Q1269*	C283G	P209Qfs*5	E79D		
coadread_dfci_2016_1231		L390M		L239P			
coadread_dfci_2016_3225		R1797W	Q286*	Y233C	Y233C		
coadread_dfci_2016_4500		V800I		X298_splice			

Table 7 continued

coadread_dfci_2016_3646		L351M,G1681W		T97A,T211Dfs*10		*809Rext*32	
coadread_dfci_2016_3729			L196P	T211Dfs*10	G69W		
coadread_dfci_2016_207430			Y108C,L9R	W241Gfs*56			
coadread_dfci_2016_2564			X349_splice	A141D	A141D		
coadread_dfci_2016_293			A16Gfs*83	T249Lfs*48,*363Lfs*37	G80V		
coadread_dfci_2016_3321			W191*	V317*		X694_splice,K407Nfs*51	
coadread_dfci_2016_3658				L105P,Q89Dfs*9			
coadread_dfci_2016_2762				C283R	W228C		
coadread_dfci_2016_3025				R258I			
coadread_dfci_2016_3091				R45Pfs*54			
coadread_dfci_2016_4430				P209Qfs*5			
coadread_dfci_2016_2354				X114_splice,I76T,L154P			
coadread_dfci_2016_2227				W241R			
coadread_dfci_2016_3641					Y195C		
coadread_dfci_2016_390						R635H	
coadread_dfci_2016_92						R438*,L83P	
coadread_dfci_2016_3319						L290F	
coadread_dfci_2016_2932							A585T
coadread_dfci_2016_1161							R220Q
coadread_dfci_2016_1222							L75*
coadread_dfci_2016_345							X165_splice
coadread_dfci_2016_2281							
coadread_dfci_2016_2624							
coadread_dfci_2016_2946							
coadread_dfci_2016_3005							
coadread_dfci_2016_3202							
coadread_dfci_2016_3269							
coadread_dfci_2016_3422							
coadread_dfci_2016_3704							
coadread_dfci_2016_3724							
coadread_dfci_2016_523							
coadread_dfci_2016_3021							
coadread_dfci_2016_3669							
coadread_dfci_2016_458							
coadread_dfci_2016_613							
coadread_dfci_2016_3174							
coadread_dfci_2016_492							
coadread_dfci_2016_546							
coadread_dfci_2016_304							
coadread_dfci_2016_3049							
coadread_dfci_2016_3249							
coadread_dfci_2016_3451							
coadread_dfci_2016_3498							
coadread_dfci_2016_3653							
coadread_dfci_2016_605							
coadread_dfci_2016_68							

Table 7 Colorectal adenocarcinoma patient cohort retrieved from TCGA database (DFCI, Cell Reports 2016) showing the mutations of MSI CRC patients in the MHC class I related genes. The table is sorted based on tumor stage of the patients; the patients were annotated with numbers as GENE_ID in the second column, analyzed genes are depicted with different colors in the columns next to each other, the colored cells show that patient has mutation in that

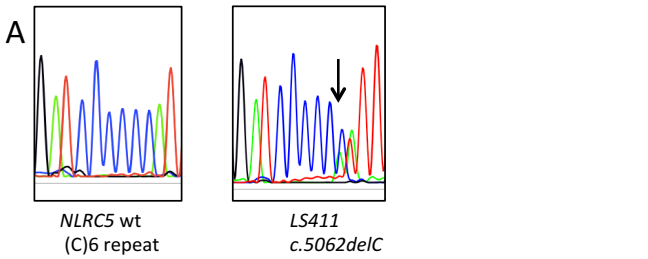
corresponding gene and the mutations are written in the cells; grey cells show no mutation.

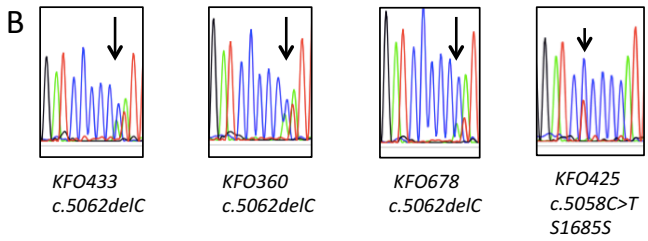
The overall view of mutations of MHC class I-related genes as an immune evasion mechanism in MSI cancer from DFCI cohort via cbioportal led us to discover NLRC5 as a potential immune evasion mechanism. Recently, NLRC5 has been identified as a transcriptional coactivator of MHC class I gene expression, and it was shown that NLRC5 expression is highly correlated with MHC class I expression [58].So far, nothing has been known about the role of NLRC5 in the MSI cancers, I examined for the first time whether *NLRC5* mutations also occurred in MSI cancer. Following the data that we gathered from DFCI cohort, we continued further to analyze mutations of NLRC5 in our cohort of MSI CRC tumor samples.

By a systematic search for cMS in the *NLRC5* gene using COSMIC database, we detected a C_6 microsatellite, designed primers and sequenced the respective region to unravel whether MSI CRCs had mutations in that region.

NLRC5 mutation analysis was first done on the same MSI CRC cell lines in order to detect if we can observe any *NLRC5* mutations in those cell lines in that specific cMS. In one of the cell line sequenced; LS411, we detected a frameshift deletion in the corresponding cMS of *NLRC5* gene (Figure 16 A); that was previously identified as *B2M* wild type.

In order to analyze *NLRC5* more extensively, a series of 95 MSI tumor specimens were analyzed by Sanger sequencing. It was found that 4 (4.2%) of 95 tumors had mutations in the respective region of the *NLRC5* gene (Figure 16 B); being 3 inactivating and one silent mutation. The mutation data (Figure 16 C) proposes that except from one exception, B2M expression could be inversely related with NLRC5.





C

Patient number	MSI-Status	hereditary	B2M (Staining)	B2M (Mutation analysis)	T	N	M	UICC	Age	proximal/distal	Localisation
KFO 433	MSI	0	2	1	4	2	1	4	67	proximal	ascending
KFO 360	MSI	0	2	0	3	2	x	3	80	proximal	coecum
KFO 678	MSI	0	2	0	2	0	-	1	67	proximal	ascending
KFO 425	MSI	0	1	1	3	1	0	3	58	proximal	ascending

0 = sporadic
1 = hereditary

0 = no positive tumor cells
1 = some areas with positive tumor cells
2 = all tumor cells positive

0 = wt
1 = mut

Figure 16 Detected mutations in the cell line or tumor samples and the corresponding patient information. Mutations were detected by Sanger sequencing in the respective C6 repeat of the coding microsatellite of NLRC5; the wild type C6 repeat and the mutated cell line LS411 is shown in A, and the mutated patients are shown in B. The patients' information with cMS mutations of NLRC5 regarding the MSI status, B2M expression and mutation, MHC class II expression, TNM stage, UICC, age and tumor location is shown in C.

In the next step, tissue sections from the *NLRC5*-mutant primary tumors and lymph node metastasis (when available) were stained by using antibodies against HLA class I heavy chains (HCA2 (against all HLA-A chains) and HC-10 (predominantly against HLA-B and HLA-C heavy chains). Illustrative staining can be seen in figure 17. We observed that the primary tumors of the patients with the FSP (-1) mutations and some of the metastasis sites express low MHC Class I on the membrane (Figure 17). We conclude that HLA class I expression may in fact be diminished as a result of *NLRC5* mutation.

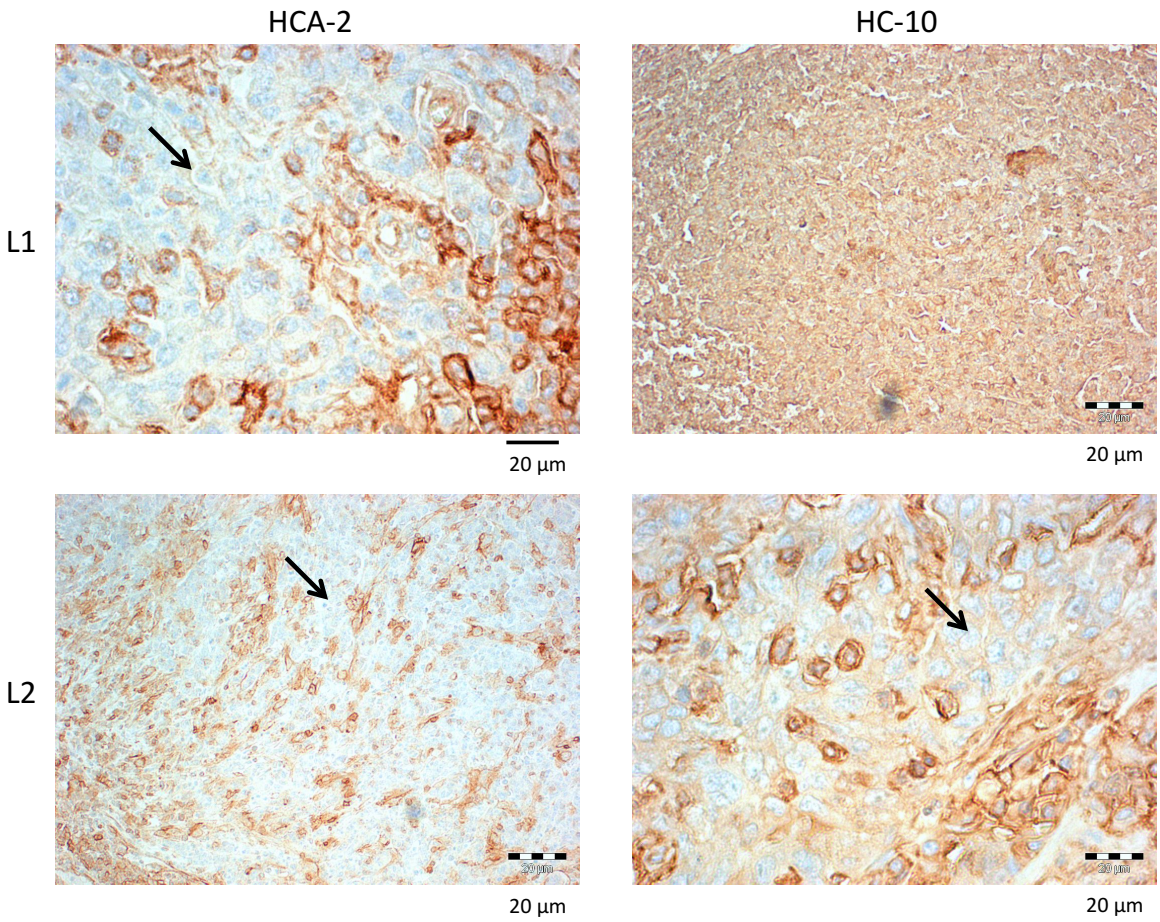


Figure 17 KFO 360 tumor samples from lymph nodes stained with HC-A2 and HC-10 antibody for heavy chains of MHC class I by immunohistochemistry. Sections of tumor samples were deparaffinized and rehydrated by xylene and sequential treatment with decreasing ethanol concentration (100%, 96% and 70%). Then epitope retrieval was done by at high temperature in the microwave for 15 minutes and endogenous peroxidase was quenched by 2% hydrogen peroxide solution in methanol for 20 minutes, after washing and edging the slides, nonspecific binding was blocked with 10% horse serum in PBS, and incubated at room temperature for 30 minutes. The slides were incubated in 1:150 diluted primary antibodies overnight, washed and secondary antibody biotin coupled horse anti-mouse IgG (1:100 diluted) was applied onto slides and incubated for 30 minutes. The before hand prepared A/B from Vectastain Elite ABC kit (1:50 diluted) was added onto slides and incubated 30 minutes and the slides were developed by DAB solution, color formation was observed and reaction was stopped by running tap water.

In order to analyze expression correlation in a tumor with heterogeneous distribution of MHC class I antigen expression, one of the HCA2 stained tumor sample that showed decreased MHC class I expression in a region was chosen and the regions which express MHC class I low and high were microdissected and those regions were examined by sequencing of cMS of *NLRC5*. Strikingly, in the 'MHC class I low' region, cMS of *NLRC5* had a one base frameshift deletion; however, in the 'MHC class I high' region, *NLRC5* was still wild type (Figure 18).

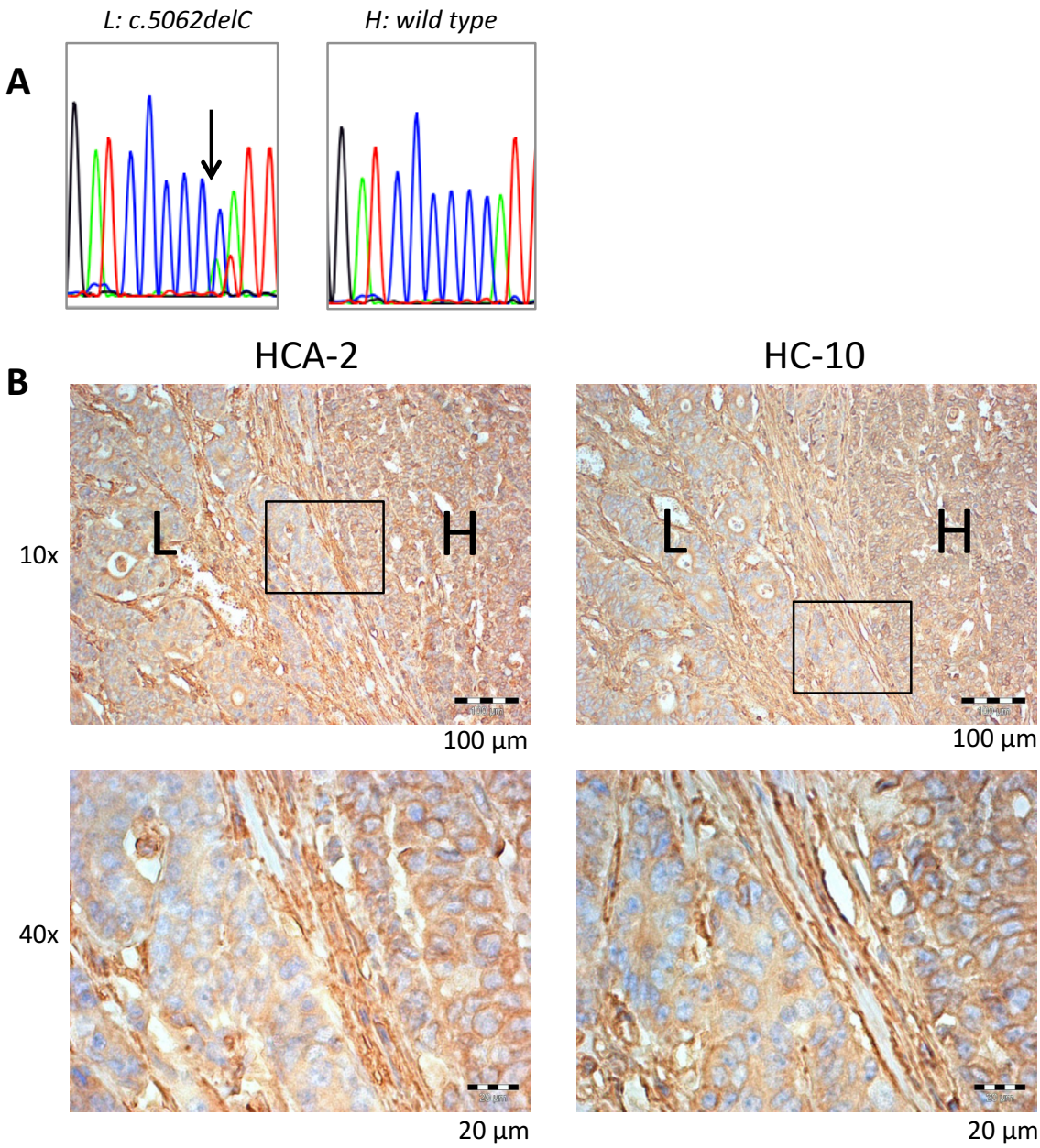


Figure 18 KFO 433 tumor sample staining with HC-A2 antibody by immunohistochemistry. MHC class I low region has a deletion in the C6 cMS of *NLRC5* whereas, MHC class I high region has still wild type copies of C6 repeat of cMS of *NLRC5*. H corresponds to MHC class I expression high region; L corresponds to MHC class I low region. The microdissection was done from the shown windows of corresponding fields and the sequencing of the C6 repeat of *NLRC5* gene is shown in the figure.

5 DISCUSSION

Part 1: Setting up a mouse model and characterization of immune phenotype

One of the most suitable MMR deficient mouse model that reflects human Lynch syndrome has been shown to be the conditional tissue specific knockout mouse model in which *Msh2*^{loxP} allele was inactivated by intestine specific *Villin-Cre* recombinase transgene and led to loss of *Msh2* in the intestinal epithelium (termed as *VCMsh2*^{loxP}): this mice spontaneously develop 1-2 intestinal adenocarcinomas and/or carcinomas that were pathologically similar to tumors from Lynch syndrome around 9 months of age [37].

In our experiments to characterize immunogenic FSPs and T cell infiltration into the tumor, we used this specific mouse model. Although this model has been described to be the most appropriate model to be used, there are still observed differences and similarities between MMR-deficient human and mouse tumors and their MSI phenotypes. The tumors from the human Lynch syndrome occur mostly in the colon [9,66],whereas in the mouse model, tumors arise predominantly in the small intestine [37]. This difference could lead to absence or disruption of mutations of the genes bearing cMS that are observed in the human setting. However, despite the fact that the tumors manifest in different organs, there are still some mouse orthologs that displays cMS mutations such as *Rfc3* and *Elavl3* that are conserved between mouse and humans and reported to be mutated in human MSI colorectal, endometrial and gastric tumors [40]. Furthermore, in mouse model of MMR deficiency, cMSs have a lower frequency of mutations as compared to human [40]; this might be due to a shorter life span of the mice that leads to less cell divisions during tumor development and consequently less mutational events. The lower mutation frequency of cMSs is also reflected in the data of the present thesis, as can be seen in table 4: the cMS located in the *Nacad* gene, with a mutation frequency of 75%, was the most commonly mutated cMS , whereas all other cMS had percentages of mutation lower than 50%.

In order to evaluate immune responses and a potential tumor-preventive effect of a

vaccine, however, the most important criterion for the suitability of a mouse model is the immune phenotype of the mice. Importantly, our IHC staining of Lynch mouse tumors revealed that CD4 T cell infiltration in a subset of the tumors was high, according to the human situation. In a subset of the tumors, however, the CD4 T cell infiltration was low, demonstrating that immune infiltration in the mouse tumors varies. This data are also in line with human Lynch syndrome cancers that show pronounced differences in immune infiltration across tumors and different individuals [42]. Therefore, we can assume that varying degrees of anti-tumoral immune responses and potentially immune evasion mechanisms occur in murine Lynch syndrome as it does in human Lynch syndrome. Accordingly, the mouse model seems apt to evaluate tumor-preventive effects of a vaccine, as stimulation of anti-tumoral immune responses may contribute to tumor prevention in some individuals, i.e. mice, more than in others.

In addition, histology of the mouse tumors were mixed with mucinous compartments, which is also resembling the human MMR deficiency [3,14]. Furthermore, cMS mutations have also been detected in MMR-deficient mouse tumors in this mouse model[40]. Therefore, all in all, even if there are differences between the mouse and human MMR deficiency, there is also considerable evidence that the mouse model has sufficient similarity and reflects the human situation well enough to at least show the feasibility of the approach. Due to the lower number of somatic mutations in Lynch mice compared to humans, one may anticipate that if the proposed FSP vaccine were effective in the mouse model, it might even be more effective in humans.

In the first part of this thesis, frameshift peptides (FSP) were selected based on a systematic search with various criteria. Following a genome wide screen of the mouse genome, mouse cMS-bearing genes were selected. Afterwards, by using the tumors from mouse model of Lynch syndrome, the genes that displayed more than 15% of mutation frequency and monomorphic in normal tissue were elected. This step was particularly crucial since setting up the cutoff for mutation frequency at 15% targeted the cMSs that are frequently mutated. Targeting only cMS bearing genes that are lacking length variants in the normal tissue ensured that the consequence of mutations would be prominent, and that mutation-induced neoantigens would not

occur in normal tissue. After mutation analysis, expression of the all cMS bearing genes was also confirmed both in normal tissue and tumor; assuring that frameshift mutations leading to FSPs both in the tumor and normal tissue could be detected by immune cells, as transcriptionally silent genes would not give rise to measurable amounts of neoantigens.

After the cMS-bearing genes were filtered, epitope prediction was carried out based on netMHC and SYFPEITHI scores [61,63].

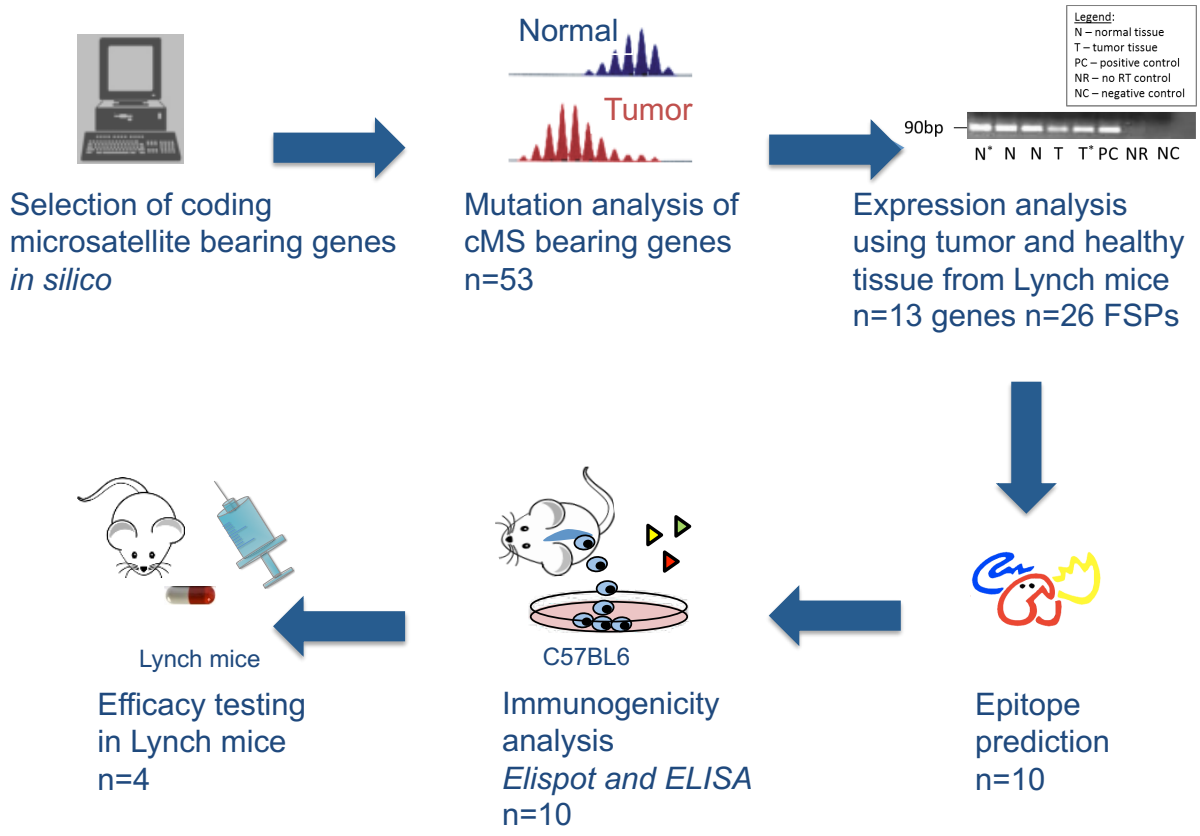


Figure 19 Schematic representation of experimental outline in the first part of the thesis.

Since the mouse model of Lynch syndrome is based on a C57/BL6 background, the epitopes that would be presented on H2-Kb/Db were scored and threshold values for SYFPETHI and netMHC were 12 and 8, respectively. SYFPEITHI scores are calculated by every amino acid of a peptide; a random value of 1 is annotated for the amino acids that are preferred only with a low probability in the respective position; however, optimal anchor residues are scored, as 15 and any value between 1 and 15 is possible for any amino acid. The data and the scores in SYFPEITHI database is

based on published motifs and it was reported that the scores are assigned according to the frequency of the corresponding amino acid in natural ligands, T cell epitopes or binding peptides[61]. In addition, the maximum scores of SYFPEITHI differs between various MHC alleles, for example, the highest score for HLA-A*0201 peptides is 36, however this score may change according to MHC alleles of interest. On the other hand, netMHC scores were the % rank of the predicted affinity of the respective peptide compared to a set of 400.000 random natural peptides; strong binders according to the website rank among the 0.5% of candidates with highest affinity, and weak binders are among the best 2% [62,63]. In order to have more candidate FSPs for the immunogenicity analysis, we used relaxed cutoffs (12 for SYFPEITHI and 5 for netMHC, not to miss any potentially immunogenic FSP derived from murine cMS mutations. As a result of the database epitope prediction, 10 candidate FSPs were evaluated for their immunogenicity.

Immunization of the C57/BL6 mice was performed as shown in figure 6. In this experimental setup, we used CpG DNA as a vaccine adjuvant. Using any other adjuvants such as TLR ligands would also be possible; however, it might have changed the type and intensity of the induced immune response. It has been described by Knudsen and colleagues in 2015 that independent of the antigen in the vaccine formulation, each adjuvant promotes distinct immune responses; for example in this study; Alum elevated humoral immune response, however, GLA-SE triggered antibody production and Th1 immune response [67]. Therefore, the adjuvants for a specific type of disease target are to be chosen based on the type of immune response that is favored.
CpG DNA has been reported to trigger an innate immune response by production of Th1 and proinflammatory cytokines and as a vaccine adjuvant, it has been shown that CpG ODNs enhance the function of antigen presenting cells and boost antigen specific humoral and cellular immune responses [68]. Therefore, CpG ODN met our needs in this vaccine formulation that it can induce both the humoral and cellular immune response.

In addition, in the immunization studies, a combination of CD4 and CD8 epitopes of the OVA peptide was used as a control either alone or in combination with CpG ODN

as an adjuvant. Having a well-characterized control peptide both with CD4 and CD8 epitopes aided us to compare the efficacy of the induction of the immune response for all of the frameshift peptides effectively, and along with the FACS stainings for CD4 and CD8 T populations, we could confirm the induction of CD4 and CD8 T cell responses against FSPs more thoroughly.

Following immunization, Elispot analysis was performed and, 4 out of 10 FSPs were detected to induce immunogenicity. Despite the fact that the scores were variable for the candidate FSPs with a SYFPEITHI score of 7 the least and 27 the most and netMHC score of 0,03 and 9,5 the most, the detected 4 FSPs induced a prominent immune response. Therefore, having a wider range of threshold for the database prediction provided us a better list for candidates of FSPs and helped us to identify the highly immunogenic FSP antigens. Our next step was to examine the immunogenicity of the 4 FSPs in one vaccine combination to observe if the combination retains the induction of prominent immune response. Application of FSPs in one vaccine formulation manifested that there was no unforeseen reaction between the FSPs and the immunogenicity of the peptides are maintained in a formulation with all FSPs.

Since all of the FSPs included in this study share 8 N-terminal amino acid sequences with the wild type of the corresponding peptide, it was also of our interest to examine if there is any effect due to cross-reactivity of the FSPs with the wild type peptide. The results showed us that the wild type peptide was not immunogenic by itself and the immune response induced by FSPs was peptide-specific. This is of importance, because cross-reactivity might lead to the activation of T cells against epitopes present not only on tumor cells, but also healthy organs. The lack of immune reactivity against wild type peptide stretches suggests that autoimmune side effects are unlikely induced by vaccination with the four FSPs.

In order to have a detailed overview of immunogenicity of the four FSPs, the induction of CD4 and CD8 T cell response by the FSPs were evaluated by IFNγ Elispot that was carried out by magnetic separated populations of FSP specific CD4 and CD8 T cells. This experiment showed us that two of the peptides; Maz (-1) and Senp6 (-1) induce only CD4 T cell responses; Xirp1 (-1) induces only CD8 T cell response;

whereas, Nacad (-1) induce both CD4 and CD8 T cell response. The predictions for the FSPs were based on mouse MHC class I molecules H2-Kb/Db, therefore, it would be plausible that the peptides that we have predicted would induce CD8 T cell response; however, two of the peptides; namely Maz (-1) and Senp6 (-1) induce only CD4 T cell response; presumably being presented on MHC class II molecules. Epitope predictions for MHC class II alleles are much less reliable due to their increased length and less stringent binding characteristics in the MHC pocket; therefore, *in silico* predictions of MHC class II antigen presentation and CD4-positive T cell responses are difficult [69]. In summary, the comparison of epitope prediction with Elispot results demonstrates a divergence between the two, underlining that prediction algorithms are useful tools for rough selection of candidates, but unable to precisely rank the best candidates without further experimental validation.

As it is shown in table 6, the size of the FSPs were ranging between 21 to 52 amino acids and for mouse MHC class I molecules H2-Kb/Db. It was reported by Stevens and colleagues in 1998 that H2-Kb have a length preference of 8-13mer and H2-Db molecules have a length preference of 9-15mer peptides [70]. Thus, it was of our interest to find the immunogenic regions of the FSPs that would be the potential epitopes presented on MHC class I and II molecules. By synthesizing shorter peptides from the N and C terminus of the corresponding FSPs, immunogenicity was evaluated against those peptides. It was shown that only C terminus of the Xirp1 (-1) and Senp6 (-1) was highly immunogenic; whereas, for Nacad (-1) only N terminus and for Maz (-1) the complete FSP; both the C and N terminus were immunogenic. This result suggests that 14 and 30 amino acids that were in the C terminus of Xirp (-1) and Senp6 (-1) and the 13 amino acids that were in the N terminus of Nacad (-1) have potential epitopes; and Maz (-1) as complete peptide bears multiple potential epitopes to be presented. Precise mapping of the relevant epitopes, which would for example be required for tetramer staining and epitope-specific immune monitoring, is now feasible on the basis of these data.

The experiments that were performed until this part confirmed that FSPs induce cellular immune response; however, in order for FSPs to induce a prominent and sustainable immune response as a vaccine, the induction of peptide specific humoral

response may also be of significance[71]. Therefore, we analyzed the induction of peptide specific total IgG by Elisa. Three out of four FSPs induced humoral immune response; only Xirp1 (-1) does not induce peptide specific IgG; and this result correlates with cellular immune response that Xirp1 (-1) does not induce CD4 T cell response but only CD8 T cell response, so there is no helper CD4 T cells that are required for antibody responses[72].
As we are planning to use a combination of FSPs in our vaccine combination, we would have various FSPs that have multiple potential epitopes to be presented, inducing both humoral and cellular immune response effectively. Moreover, the vaccine coverage of the FSP vaccine with combination of these four FSPs is expected to be about 75% in the Lynch mouse model. Even if the vaccine coverage for this mouse FSP vaccine could be improved even more by selection of different and/or more FSPs, with these four FSPs, we could already show the immense immunogenicity in one vaccine formulation and based on our immunogenicity experiments, we suggest that it is already highly promising that is yet to be tested in the mouse model.

In conclusion, the four immunogenic FSPs that we characterized the immune response so far is very promising to be tested as FSP vaccination in Lynch syndrome mouse model both in a therapeutic and preventive setting to prevent formation of cancer. Following testing the efficacy of the FSP vaccination in the mouse model, different delivery methods is to be assessed. After the efficacy of FSP vaccination is shown in mouse model, our aim is to employ this approach for human setting and develop a FSP vaccine in a therapeutic setting for all types of MSI cancer and in a preventive setting for Lynch syndrome. Furthermore, considering the success of PD1 immunotherapy in MSI cancer[73], FSP vaccination could be combined as a therapy to target the patients that are resistant to checkpoint inhibitors and immune response might be more effectively boosted by the combination therapy.

Part 2: Comprehensive analysis of MHC-related immune evasion in MSI Cancer

In order to determine potential immune evasion mechanisms that the tumor might develop in spite of a vaccine, we first characterized mutations of *B2M* both in the gene

and protein level and verified if mutations of *B2M* are functional in MSI CRC cell lines. We detected MSI CRC cell lines exhibiting either homozygous or heterozygous *B2M* frameshift and point mutations in their corresponding cMSs. However, expression data showed us that one of the cell line which had a frameshift deletion retained high B2M and MHC class I expression; whereas all the other cell lines having *B2M* mutations do not express B2M and display low or no MHC class I expression even after IFNγ treatment. Therefore, except from one example that maintained MHC class I expression in spite of a heterozygous *B2M* mutation, all the other mutations in *B2M* in other MSI cell lines resulted in loss of MHC class I expression as in line with the previous publications that correlates *B2M* mutations with loss of MHC class I expression[45].

B2M has already been shown to be mutated in 30% in MSI CRCs [46]; whereas *B2M* mutations in MSS CRCs are shown to be rare [45,47]. Accordingly, mutations of *B2M* are an important immune evasion mechanism in MSI cancer that leads to loss of MHC class I. Considering that mutations of *B2M* are only displayed in 30% of the MSI cancers, for the remaining 70% it is still yet to be discovered if they also show defects in the MHC class I machinery that were to be described or if they were developing without immune evasion, potentially due to a weaker local immune response, as has been postulated by Echterdiek et al. [42].

In order to address this question, in this part of the thesis, a more comprehensive overview of the immune evasion mechanisms were implemented by analyzing the genes that are related to MHC class I antigen presentation (*NLRC5*, *B2M*, *TAP1*, *TAP2*, *HLA-A*, *HLA-B* and *HLA-C*) in 91 MSI CRC patients in DFCI cohort of colorectal carcinoma [74] via cBioportal. DFCI cohort data showed us that 72% of MSI tumors have defects in MHC class I machinery; demonstrated by at least one mutation in the respective genes. In the light of this data, here we could unravel other types of MHC class I related mutations than *B2M* that led to defects in MHC class I machinery in another 42% of the MSI cancers. The high prevalence of the MHC class I related immune evasion mechanisms in MSI cancers may be explained by high mutational load of MSI cancer due to the FSPs and high TILs [75] that leads to high immune selection pressure. As a result of stringent immunoselection, tumor cells gain immune

evasion mechanisms, disrupting the equilibrium between immune and tumor cells; resulting in immunoediting by the escape of tumor cells from immune system of the host [76,77].

Although we detected 72% of MSI tumors with MHC class I related alterations on the DNA level, one has to consider that not all of the mutations were stop or frameshift mutations. Therefore, the percentage of tumors showing functionally relevant mutations affecting the MHC class I pathway might be lower. On the other hand, the remaining 28% of the MSI CRC tumors that did not represent any mutations in the respective genes might have mutations in any other genes that we didn't analyze but still may be associated with antigen presentation. Alternatively, it is plausible that these tumors were less immunogenic, and immune selection was not as strong as in the other tumors that displayed alterations associated with immune evasion related with MHC class I machinery. In order to better characterize the correlation between immune responses and immune evasion, infiltration of tumors with immune cells needs to be assessed in future immune evasion studies, according to what has already been evaluated in tumor-adjacent mucosa before [42] .
In line with the data that was reported before by Kloor and colleagues in 2007 [46] that *B2M* mutations are observed in 30% of MSI tumors, we detected *B2M* mutations in 28 out of 91 MSI tumor samples leading to loss of MHC class I expression. Furthermore, mutations in MHC class I heavy chains; *HLA-A, -B* and *–C*, were observed in 23 out of 91 MSI cancers; most of the mutations resulted in nonfunctional protein products due to truncating mutations. In addition, correlation coefficients that was calculated from the software of cbioportal showed us that mutations of the genes *HLA –A, -B* and *–C* were positively correlated and those specific mutations arise less frequently together with *B2M* mutations; the negative correlation coefficiency for *B2M* and *HLA-B* is significant with log odds ratio of 2.359, p=0.005. These observations provide very strong support to the hypothesis that the observed mutations are in fact functionally relevant, as the observed positive and negative correlations are very hard to explain without the existence of immune selection that shapes and 'edits' the phenotype of the tumors.

Thus, the results of the present thesis clearly support the concept of stringent

immunoselection in the tumor microenvironment in Lynch syndrome. This force drives immune evasion mechanisms leading to mutations in MHC class I related genes.

Apart from mutations in MHC class I antigen encoding genes, as is in line with Kloor *et al*. 2005, mutations in antigen processing (*TAP1* and *TAP2*) were detected in MSI tumors [45].

Additionally, besides *TAP1* and *TAP2* mutations, another novel potential immune evasion mechanism was discovered in MSI cancer by mutations of *NLRC5*. Being recently described as the transcriptional coactivator of MHC class I gene expression [50], NLRC5 was shown to transactivate classical and non-classical MHC class I genes including *B2M*, *TAP1*, *PSMB9* and *HLA A/B/C/D/E/F/G* [56]. On the other hand, in contrast to MHC class I deficiency, *Nlrc5* knockout mice exhibit mosaic MHC class I expression in various cells and tissues[56]; therefore, it is plausible that NLRC5 deficiency would result only in partial loss of MHC class I rather than a complete loss of expression.

In DFCI cohort, we detected *NLRC5* sequence variants in 24 out of 91 patients (26%), with 6 out of 91 tumors (6.5%) showing functionally relevant nonsense or frameshift mutations. This suggests that NLRC5 mutations may represent a potential novel immune evasion mechanism in MSI cancer. To examine this further, we analyzed a specific cMS of NLRC5 to determine if we observe mutations in this region that might lead to an immune evasion mechanism by altering MHC class I expression.

In our MSI CRC tumor samples, we detected *NLRC5* mutations in 4 out of 95 MSI cancers. Considering that we analyzed mutations only in the region where the C6 repeat is, and sorting the stop mutations and frameshift mutations in DFCI cohort, it conceivable that we detected 3.15% of NLRC5-inactivating mutations in MSI tumor samples; corresponding roughly to the frequency in DFCI cohort. Furthermore, in one of the tumor sample that was stained for MHC class I, we microdissected the parts with high and low MHC class I expression and sequenced those specific regions; we determined that only the region with the low MHC class I expression has a frameshift

deletion mutation; whereas in the MHC class I high region, the specific cMS of NLRC5 was still wild type. This data led us to conclude that NLRC5 mutations might induce partial loss of MHC class I expression.

Moreover, the functional significance of NLRC5 for MHC class I expression is suggested by many factors: first, all tumors displaying *NLRC5* mutations in our collection maintained *B2M* expression. And second, NLRC5 inactivation results in lower MHC class I expression, and mutations in NLRC5 lead to partial loss of MHC class I expression correlated with *NLRC5* mutation in the corresponding region.

All in all, we for the first time describe *NLRC5* mutations in MSI cancer as a novel mechanism likely leading to immune evasion. Our results support the concept that MSI cancers follow a stringent process of immunoselection, as is evidenced by the occurrence of mutations in MHC class I related genes occurring in more than two-thirds of MSI colorectal cancers. The determination of comprehensive overview of immune evasion mechanisms in MSI cancer and discovery of potential new immune evasion mechanism by *NLRC5* mutations provided us new insights into MSI cancer pathogenesis. With the help of this information, possible immune evasion mechanisms can be more efficiently characterized. Clinically, comprehensive typing of immune evasion phenomena may serve as a marker to predict responsiveness towards immunotherapy, namely immune checkpoint blockade, and ultimately to develop more effective immunotherapies.

6 REFERENCES

1. Siegel, R. L. *et al.* Colorectal Cancer Statistics , 2017. **67,** 177–193 (2017).
2. Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21,** 1350–1356 (2015).
3. Kloor, M., Staffa, L., Ahadova, A. & Von Knebel Doeberitz, M. Clinical significance of microsatellite instability in colorectal cancer. *Langenbeck's Arch. Surg.* **399,** 23–31 (2014).
4. Walther, A. *et al.* Genetic prognostic and predictive markers in colorectal cancer. *Nat. Rev. Cancer* **9,** 489–499 (2009).
5. Gelsomino, F., Barbolini, M., Spallanzani, A., Pugliese, G. & Cascinu, S. The evolving role of microsatellite instability in colorectal cancer: A review. *Cancer Treat. Rev.* **51,** 19–26 (2016).
6. Kloor, M., von Knebel Doeberitz, M. & Gebert, J. F. Molecular testing for microsatellite instability and its value in tumor characterization. *Expert Rev. Mol. Diagn.* **5,** 599–611 (2005).
7. Lynch, H. T., Snyder, C. L., Shaw, T. G., Heinen, C. D. & Hitchins, M. P. Milestones of Lynch syndrome: 1895–2015. *Nat. Rev. Cancer* **15,** 181–194 (2015).
8. Boland, C. R. *et al.* A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* **58,** 5248–57 (1998).
9. Jasperson, K. W., Tuohy, T. M., Neklason, D. W. & Burt, R. W. Hereditary and Familial Colon Cancer. *Gastroenterology* **138,** 2044–2058 (2010).
10. Kloor, M. & Von Knebel Doeberitz, M. The immune biology of microsatellite-unstable cancer. *Trends in Cancer* **2,** 121–133 (2016).
11. Hampel, H. *et al.* Cancer risk in hereditary nonpolyposis colorectal cancer syndrome: later age of onset. *Gastroenterology* **129,** 415–21 (2005).
12. Quehenberger, F., Vasen, H. F. A. & van Houwelingen, H. C. Risk of colorectal and endometrial cancer for carriers of mutations of the hMLH1 and hMSH2 gene: correction for ascertainment. *J. Med. Genet.* **42,** 491–6 (2005).
13. Umar, A. *et al.* Revised Bethesda Guidelines for Hereditary Nonpolyposis Colorectal Cancer (Lynch Syndrome) and Microsatellite Instability. *JNCI J. Natl.*

Cancer Inst. **96,** 261–268 (2004).

14. Shia, J., Holck, S., Depetris, G., Greenson, J. K. & Klimstra, D. S. Lynch syndrome-associated neoplasms: A discussion on histopathology and immunohistochemistry. *Fam. Cancer* **12,** 241–260 (2013).
15. Buckowitz, A. *et al.* Microsatellite instability in colorectal cancer is associated with local lymphocyte infiltration and low frequency of distant metastases. *Br. J. Cancer* **92,** 1746–1753 (2005).
16. Smyrk, T. C., Watson, P., Kaul, K. & Lynch, H. T. Tumor-infiltrating lymphocytes are a marker for microsatellite instability in colorectal carcinoma. *Cancer* **91,** 2417–2422 (2001).
17. Galon, J. Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome. *Science (80-.).* **313,** 1960–1964 (2006).
18. Edin, S. *et al.* The Distribution of Macrophages with a M1 or M2 Phenotype in Relation to Prognosis and the Molecular Characteristics of Colorectal Cancer. *PLoS One* **7,** (2012).
19. Prall, F. *et al.* Prognostic role of CD8+ tumor-infiltrating lymphocytes in stage III colorectal cancer with and without microsatellite instability. *Hum. Pathol.* **35,** 808–816 (2004).
20. Dolcetti, R. *et al.* High Prevalence of Activated Intraepithelial Cytotoxic T Lymphocytes and Increased Neoplastic Cell Apoptosis in Colorectal Carcinomas with Microsatellite Instability. *Am. J. Pathol.* **154,** 1805–1813 (1999).
21. Le Gouvello, S. *et al.* High prevalence of Foxp3 and IL17 in MMR-proficient colorectal carcinomas. *Gut* **57,** 772–779 (2008).
22. Duval, a *et al.* Evolution of instability at coding and non-coding repeat sequences in human MSI-H colorectal cancers. *Hum. Mol. Genet.* **10,** 513–518 (2001).
23. Alhopuro, P. *et al.* Candidate driver genes in microsatellite-unstable colorectal cancer. *Int. J. Cancer* **130,** 1558–1566 (2012).
24. Woerner, S. M. *et al.* Pathogenesis of DNA repair-deficient cancers: a statistical meta-analysis of putative Real Common Target genes. *Oncogene* **22,** 2226–2235 (2003).
25. Tougeron, D. *et al.* Tumor-infiltrating lymphocytes in colorectal cancers with microsatellite instability are correlated with the number and spectrum of

frameshift mutations. *Mod. Pathol.* **22,** 1186–1195 (2009).

26. Maby, P. *et al.* Correlation between density of CD8+ T-cell infiltrate in microsatellite unstable colorectal cancers and frameshift mutations: A rationale for personalized immunotherapy. *Cancer Res.* **75,** 3446–3455 (2015).
27. Schwitalle, Y. *et al.* Immune Response Against Frameshift-Induced Neopeptides in HNPCC Patients and Healthy HNPCC Mutation Carriers. *Gastroenterology* **134,** 988–997 (2008).
28. Bauer, K. *et al.* T cell responses against microsatellite instability-induced frameshift peptides and influence of regulatory T cells in colorectal cancer. *Cancer Immunol. Immunother.* **62,** 27–37 (2013).
29. Markowitz, S. *et al.* Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science* **268,** 1336–8 (1995).
30. Linnebacher, M. *et al.* Frameshift peptide-derived T-cell epitopes: A source of novel tumor-specific antigens. *Int. J. Cancer* **93,** 6–11 (2001).
31. Schwitalle, Y., Linnebacher, M., Ripberger, E., Gebert, J. & von Knebel Doeberitz, M. Immunogenic peptides generated by frameshift mutations in DNA mismatch repair-deficient cancer cells. *Cancer Immun. a J. Acad. Cancer Immunol.* **4,** 14 (2004).
32. Llosa, N. J. *et al.* The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer Discov.* **5,** 43–51 (2015).
33. Le, D. T. *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* **372,** 2509–2520 (2015).
34. Lee, V., Murphy, A., Le, D. T. & Diaz, L. A. Mismatch Repair Deficiency and Response to Immune Checkpoint Blockade. *Oncologist* **21,** 1200–1211 (2016).
35. Woerner, S. M. *et al.* Systematic identification of genes with coding microsatellites mutated in DNA mismatch repair-deficient cancer cells. *Int. J. Cancer* **93,** 12–19 (2001).
36. Wimmer, K. & Etzler, J. Constitutional mismatch repair-deficiency syndrome: have we so far seen only the tip of an iceberg? *Hum. Genet.* **124,** 105–122 (2008).
37. Kucherlapati, M. H. *et al.* An Msh2 Conditional Knockout Mouse for Studying Intestinal Cancer and Testing Anticancer Agents. *Gastroenterology* **138,** 993–

1002.e1 (2010).

38. Lee, K., Tosti, E. & Edelmann, W. Mouse models of DNA mismatch repair in cancer research. *DNA Repair (Amst).* **38,** 140–146 (2016).
39. Bacher, J. W., Abdel Megid, W. M., Kent-First, M. G. & Halberg, R. B. Use of mononucleotide repeat markers for detection of microsatellite instability in mouse tumors. *Mol. Carcinog.* **44,** 285–292 (2005).
40. Woerner, S. M. *et al.* Detection of coding microsatellite frameshift mutations in DNA mismatch repair-deficient mouse intestinal tumors. *Mol. Carcinog.* **54,** 1376–1386 (2015).
41. Kloor, M., Michel, S. & von Knebel Doeberitz, M. Immune evasion of microsatellite unstable colorectal cancers. *Int. J. Cancer* **127,** 1001–1010 (2010).
42. Echterdiek, F. *et al.* Low density of FOXP3-positive T cells in normal colonic mucosa is related to the presence of beta2-microglobulin mutations in Lynch syndrome-associated colorectal cancer. *Oncoimmunology* **5,** e1075692 (2016).
43. Berger, T. G. *et al.* Circulation and homing of melanoma-reactive T cells to both cutaneous and visceral metastases after vaccination with monocyte-derived dendritic cells. *Int. J. Cancer* **111,** 229–237 (2004).
44. Rosenberg, S. A. *et al.* Cell transfer therapy for cancer: lessons from sequential treatments of a patient with metastatic melanoma. *J Immunother* **26,** 385–393 (2003).
45. Kloor, M. Immunoselective Pressure and Human Leukocyte Antigen Class I Antigen Machinery Defects in Microsatellite Unstable Colorectal Cancers. *Cancer Res.* **65,** 6418–6424 (2005).
46. Kloor, M. *et al.* Beta2-microglobulin mutations in microsatellite unstable colorectal tumors. *Int. J. Cancer* **121,** 454–458 (2007).
47. Bicknell, D. C., Kaklamanis, L., Hampson, R., Bodmer, W. F. & Karran, P. Selection for beta 2-microglobulin mutation in mismatch repair-defective colorectal carcinomas. *Curr Biol* **6,** 1695–1697 (1996).
48. Woerner, S. M. *et al.* SelTarbase, a database of human mononucleotide-microsatellite mutations and their potential impact to tumorigenesis and immunology. *Nucleic Acids Res.* **38,** 682–689 (2009).
49. Bernal, M. *et al.* Genome-wide differential genetic profiling characterizes

colorectal cancers with genetic instability and specific routes to HLA class I loss and immune escape. *Cancer Immunol. Immunother.* **61,** 803–816 (2012).

50. Meissner, T. B. *et al.* NLR family member NLRC5 is a transcriptional regulator of MHC class I genes. *Proc. Natl. Acad. Sci.* **107,** 13794–13799 (2010).
51. Biswas, A., Meissner, T. B., Kawai, T. & Kobayashi, K. S. Cutting Edge: Impaired MHC Class I Expression in Mice Deficient for Nlrc5/Class I Transactivator. *J. Immunol.* **189,** 516–520 (2012).
52. Robbins, G. R. *et al.* Regulation of Class I Major Histocompatibility Complex (MHC) by Nucleotide-binding Domain, Leucine-rich Repeat-containing (NLR) Proteins. *J. Biol. Chem.* **287,** 24294–24303 (2012).
53. Staehli, F. *et al.* NLRC5 Deficiency Selectively Impairs MHC Class I- Dependent Lymphocyte Killing by Cytotoxic T Cells. *J. Immunol.* **188,** 3820–3828 (2012).
54. Tong, Y. *et al.* Enhanced TLR-induced NF-κB signaling and type I interferon responses in NLRC5 deficient mice. *Cell Res.* **22,** 822–835 (2012).
55. Yao, Y. *et al.* NLRC5 regulates MHC class I antigen presentation in host defense against intracellular pathogens. *Cell Res.* **22,** 836–847 (2012).
56. Chelbi, S. T., Dang, A. T. & Guarda, G. *Emerging Major Histocompatibility Complex Class I-Related Functions of NLRC5*. *Advances in Immunology* **133,** (Elsevier Inc., 2016).
57. Ludigs, K. *et al.* NLRC5 Exclusively Transactivates MHC Class I and Related Genes through a Distinctive SXY Module. *PLoS Genet.* **11,** 1–22 (2015).
58. Yoshihama, S. *et al.* NLRC5/MHC class I transactivator is a target for immune evasion in cancer. *Proc. Natl. Acad. Sci.* **113,** 5999–6004 (2016).
59. Li, X. *et al.* NLRC5 expression in tumors and its role as a negative prognostic indicator in stage {III} non-small-cell lung cancer patients. *Oncol. Lett.* **10,** 1533–1540 (2015).
60. Rodriguez, G. M. *et al.* NLRC5 elicits antitumor immunity by enhancing processing and presentation of tumor antigens to CD8 + T lymphocytes. *Oncoimmunology* **5,** e1151593 (2016).
61. Rammensee, H.-G., Bachmann, J., Emmerich, N. P. N., Bachor, O. A. & Stevanović, S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* **50,** 213–219 (1999).
62. Nielsen, M. *et al.* Reliable prediction of T-cell epitopes using neural networks

with novel sequence representations. *Protein Sci.* **12,** 1007–1017 (2003).

63. Andreatta, M. & Nielsen, M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* **32,** 511–517 (2015).
64. Cerami, E. *et al.* The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2,** 401–404 (2012).
65. Gao, J. *et al.* Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* **6,** pl1-pl1 (2013).
66. Goel, A. *et al.* Low frequency of Lynch syndrome among young patients with non-familial colorectal cancer. *Clin. Gastroenterol. Hepatol.* **8,** 966–71 (2010).
67. Knudsen, N. P. H. *et al.* Different human vaccine adjuvants promote distinct antigen-independent immunological signatures tailored to different pathogens. *Sci. Rep.* **6,** 19570 (2016).
68. Bode, C., Zhao, G., Steinhagen, F., Kinjo, T. & Klinman, D. M. CpG DNA as a vaccine adjuvant. *Expert Rev. Vaccines* **10,** 499–511 (2011).
69. Nielsen, M., Lund, O., Buus, S. & Lundegaard, C. MHC Class II epitope predictive algorithms. *Immunology* **130,** 319–328 (2010).
70. Stevens, J., Wiesmüller, K. H., Walden, P. & Joly, E. Peptide length preferences for rat and mouse MHC class I molecules using random peptide libraries. *Eur. J. Immunol.* **28,** 1272–9 (1998).
71. Reuschenbach, M. & Doeberitz, M. V. K. A systematic review of humoral immune responses against tumor antigens. *Cancer Immunol.* **58,** 1535–1544 (2009).
72. Kenneth Murphy, Paul Travers, Mark Walport, Michael Ehrenstein, Claudia Mauri, Allan Mowat, A. S., Hsu, D. C., Murphy, K. M., Travers, P. & Walport, M. *Janeway's Immunobiology*. *Garland Science* **7,** (2008).
73. Link, J. T. & Overman, M. J. Immunotherapy Progress in Mismatch Repair–Deficient Colorectal Cancer and Future Therapeutic Challenges. *Cancer J.* **22,** 190–195 (2016).
74. Giannakis, M. *et al.* Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. *Cell Rep.* **15,** 857–865 (2016).
75. Drescher, K. M., Sharma, P. & Lynch, H. T. Current hypotheses on how

microsatellite instability leads to enhanced survival of lynch syndrome patients. *Clinical and Developmental Immunology* **2010,** (2010).

76. Ribatti, D. The concept of immune surveillance against tumors. The first theories. *Oncotarget* **8,** 7175–7180 (2017).

77. Ward, J. P., Gubin, M. M. & Schreiber, R. D. *The Role of Neoantigens in Naturally Occurring and Therapeutically Induced Immune Responses to Cancer*. *Advances in Immunology* **130,** (Elsevier Inc., 2016).