Identification of tumor-specific peptides for adjuvant immunotherapy of melanoma patients

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.) DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN DER UNIVERSITÄT REGENSBURG

vorgelegt von
Sandra Huber

aus
Regensburg

im Jahr 2021

Das Promotionsgesuch wurde eingereicht am:
22.04.2021

Die Arbeit wurde angeleitet von:
Prof. Dr. Christoph Klein

Unterschrift:

Table of contents

List of figures

List of tables

Table of abbreviations

Abbreviation	Meaning
AB	Antibody
ASB-PCR	Allele-specific PCR with a blocking reagent
ATP	Adenosine triphosphate
bp	Base pair
BSA	Bovine serum albumin
CD	Cluster of differentiation
cDNA	Complementary DNA
CLND	Complete lymph node dissection
CNA	Copy number alteration
CTC	Circulating tumor cell
DCC	Disseminated cancer cell
DCCD	DCC density; number of DCCs per million mononuclear cells
DSS	Double-strand synthesis
dGTP	Deoxy guanosine triphosphate
dNTP	Deoxynucleotide triphosphate
EDTA	Ethylenediaminetetraacetic acid
FACS	Fluorescence-activated cell sorting
FCS	Fetal calf serum
FFPE	Formalin-fixed paraffin-embedded
gDNA	Genomic deoxyribonucleic acid
GII	Genomic integrity index
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HE	Hematoxylin and eosin
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
H_2O	Water
HS	High sensitivity
IgG	Immunoglobulin G
IHC	Immune histochemistry
LP-Seq	LowPass-Sequencing
MCSP	Melanoma-associated chondroitin sulfate proteoglycan
Met/MET	Metastasis
mRNA	Messenger RNA
MS-Seq	Microsatellite sequencing
NA	Not available
NBT/BCIP	Nitro-blue tetrazolium chloride/5-bromo-4-chloro-3'-indolylphosphate p-tuluidine salt
NEB	New England Biolabs
NGS	Next generation sequencing
NSLN	Non-sentinel lymph node
NTC	Non-tumor cell
OEC	Oral epithelial cell
OPA	One Phor All buffer
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PT	Primary tumor
QC	Quality control
REA	Re-amplification
RNA	Ribonucleic acid
rpm	Rounds per minute
RT	Room temperature
SLN	Sentinel lymph node
TBE (bioinformatic)	Transfer Bootstrapping Expectation
TBE (wet lab)	Tris/Borate/EDTA buffer
TNM	Tumor-Node-Metastasis staging system
T-Vec	Talimogen laherparepvec
UV	Ulraviolet
WES	Whole exome sequencing
WGA	Whole genome amplification
WTA	Whole transcriptome amplification

Summary

Malignant melanoma is an aggressive cancer with increasing incidence, frequent metastasis and poor prognosis. It is also an immunogenic cancer and clinical and experimental data point towards a pivotal role of neoantigens and neoantigen-reactive T cells in controlling melanoma. As mortality from metastatic melanoma is still high despite recent immunotherapy progress, all efforts must be directed to eliminate the metastatic seed - the disseminated cancer cells (DCCs) - early on, before lethal metastasis is established. As dissemination in melanoma starts early during primary tumor (PT) development and DCCs continue their evolution outside of the PT, DCCs can harbor different genetic alterations compared to the PT. However, therapeutic target decisions are currently based on the molecular characteristics of the PT and not on the target cells of adjuvant therapy, the DCCs. DCCs can be detected in sentinel lymph nodes (SLNs), which are routinely extirpated for diagnostic reasons. An ideal adjuvant therapy should eliminate all DCCs, but it is yet unclear whether early SLN-DCCs mirror the molecular characteristics of systemically spread cancer cells and can be used to identify truncal mutations for adjuvant, neoantigen-based immune therapy. The thesis addressed this fundamental question of adjuvant therapy by (i) analysing the phylogenetic relationship of longitudinally sampled DCCs and by (ii) exploring neonantigen conservation across these samples. To this end, more than 500 single DCCs and non-tumor control cells were isolated by micromanipulation from longitudinal samples of three melanoma patients, and subjected to single cell next generation sequencing for 12472 microsatellite (MS) markers and to copy number alteration (CNA) analysis. After successful optimization of sample preparations for MS-sequencing, cells from different patients and different cell lineages (hematopoietic vs. epithelial and tumor vs. non-tumor cells) within one patient could be separated. The phylogenetic trees of all three patients showed that SLN-DCCs clustered with tumor cells isolated at later stages of the disease. This indicated that SLN-DCCs isolated in the early disease phase might indeed serve as good surrogates for systemically spread tumor cells giving rise to metastasis at a later time point. These findings were supported by CNA analyses, showing that SLN-DCCs already harbored aberrations also present in cancer cells from non-sentinel lymph nodes or from cutaneous metastases. Moreover, phylogenetic trees indicated that parallel seeding of lymph nodes and metastases might be more likely than seeding in cascades. *BRAF* mutation analysis of one patient further pointed towards a mixture of cascadic and simultaneous colonization of lymph nodes and the blood stream. To identify truncal neoantigens and test for mechanisms of immune evasion, such as loss of HLA alleles, whole exome sequencing (WES) of DCCs from the three patients was conducted. WES revealed skin cancer-related UV signatures in tumor samples from two patients. For all patients, a high degree of shared exonic mutations and predicted peptides were detected between SLN-DCCs and cells from cutaneous or blood-derived metastases. This further supported the suitability of SLN-DCCs to serve as surrogates for systemic tumor cells detected during disease progression. The number of predicted peptides significantly correlated with the tumor burden of the lymph nodes and significantly increased with decreasing number of HLA alleles in one patient, in whom HLA analysis revealed a potential loss of heterozygosity (LOH), indicating immune pressure and HLA-LOH as a potential mechanism of immune evasion. In order to validate shared, candidate neoantigens and test for their immunogenicity, a test system was successfully established identifying antigen-specific T cells among tumor-reactive T cells by combining (i) expansion of tumor-reactive T cells from peripheral blood cells with (ii) specific recognition of *in vitro* differentiated antigen-presenting cells transfected with antigen-encoding minigenes. Whether identified candidate neoantigens were suitable for adjuvant immunotherapy, remains to be evaluated. In summary, the combination of phylogenetic analyses and WES for neoantigens explored the clinical utility of early disseminated cancer cells in sentinel lymph nodes beyond diagnosis, and pointed towards a pivotal role of SLN-DCCs for therapeutic target identification against early systemic cancer.

1 Introduction

1.1 Malignant melanoma

1.1.1 Epidemiology, aetiology and genetics

Cutaneous malignant melanoma is one of the most aggressive human cancers with a steadily increasing incidence (Raimondi et al. 2020). Melanoma of the skin is the 21st most common cancer worldwide with 287723 new cases and 60712 deaths reported in 2018 (Bray et al. 2018). However, incidences as well as mortality rates vary between different countries due to the diversity in racial skin phenotypes and variations in sun exposures. In contrast to other solid tumors, which are diagnosed in the majority of people over the age of 65, melanoma is diagnosed in more younger patients (Ali et al. 2013, Miller et al. 2019). Malignant melanoma has the highest lethality of all forms of skin cancer and the survival rate decreases with progressing disease to 5-19% once distant metastases have formed (Sandru et al. 2014, Erdei and Torres 2010, Miller et al. 2019). There are four major types of cutaneous melanoma differing in color, localization, genetics and shape. The most common skin cancer is superficial spreading melanoma occurring in around 70% of cases, followed by nodular melanoma, lentigo maligna melanoma and acral lentiginous melanoma (Ward et al. 2017).

The exact aetiology of melanoma is not yet completely understood. However, epidemiological studies revealed several risk factors that play an important role in the development of skin cancer, including environmental and genetic factors. Depending on exogenous and endogenous mutational processes different genetic changes in the cancer genome are generated, resulting in a characteristic mutational signature (Alexandrov and Stratton 2014). Those signatures comprise single base substitutions (SBSs), doublet base substitutions and small insertions/ deletions (available at *https://cancer.sanger.ac.uk/cosmic/signatures/index.tt*). SBS signatures are identified by considering the six possible base substitutions (C>A, C>G, C>T, T>A, T>C, T>G) and the nucleotide context, i.e. bases at the flanking 5'- and 3'-position. The main environmental factor for developing skin cancer is UV light exposure. Therefore, excessive sun exposure is one of the major risk factors for developing melanoma (Gandini et al. 2005). UV radiation can damage the DNA of skin cells leading to pyrimidine dimer formation, one of the main UV-induced DNA photolesions (Ravanat et al. 2001). Those DNA damages are usually repaired by nucleotide excision repair, but can also lead to C>T or CC>TT nucleotide transitions (Budden and Bowden 2013). Those C>T substitutions can occur at different dinucleotides and are characteristic for mutational signature SBS7, also referred to as UV signature (Alexandrov et al. 2013, Hayward et al. 2017, Alexandrov et al. 2019). As not all parts of the body are exposed to UV radiation, melanomas with UV signatures are more often found in upper regions of the body (Craig et al. 2018).

In most of the cases, DNA damage and genetic changes only occur in those cells that give rise to melanomas. Thus, developing melanoma is a sporadic event. However, around 10% of melanomas are family-related showing an autosomal-dominant inheritance of corresponding mutations (Read et al. 2016). In this regard, the most common germline mutations are found in the tumor suppressor gene *CDKN2A* (Goldstein et al. 2007) and the oncogene *CDK4* (Puntervoll et al. 2013). Typical genes playing a role in sporadic melanomas often affect signal transduction pathways and are reciprocally exclusive. Constitutive activation due to mutations in related genes may lead to formation and progression of malignant melanomas. For example, around 50% of cutaneous melanomas show a mutation in the *BRAF* gene involved in the MAP kinase/ERK-signaling pathway. Here, in more than 90% of the cases the most common mutation is BRAF V600E, where a nucleotide exchange leads to an amino acid substitution from valine to glutamic acid (Ascierto et al. 2012). Of note, BRAF V600E does not present the classic C>T nucleotide transition caused by UV radiation (Craig et al. 2018).

1.1.2 Histology, diagnostic and clinic

In cutaneous melanoma DNA damages and mutations are found in melanocytes, the melanin-producing cells of the skin (Uong and Zon 2010). The skin has an outer layer called epidermis and an inner layer, the dermis. Below the skin is the subcutaneous tissue consisting of vessels and fat, which connects the skin to bones and muscles. Melanocytes can be found at the bottom of the epidermis and produce the melanin pigment giving the color to the skin. In general, cancer cells divide more often and more quickly than normal cells leading to tumor formation. Abnormal melanocytes have become cancer cells in cutaneous malignant melanoma due to uncontrolled and extensive proliferation because of genetic changes. As they still produce melanin, the tumors are normally dark in their color.

The first formation of proliferative cancer cells is called the primary tumor. However, cancer cells can spread throughout the body to distant organs, far apart from primary tumors, via blood or lymph vessels and can form so-called distant metastases. If melanoma has spread to a part of the body nearby the primary tumor, this is termed local metastasis. Satellite metastases are located within two centimeters from the primary tumor. Once tumor cells are found within the lymph vessel, one speaks of in-transit metastasis. These differences in metastasis-formation are of high importance for clinical staging of melanoma skin cancer according to the American Joint Committee on Cancer (AJCC) TNM classifications (Gershenwald et al. 2017). Category T is defining the primary tumor, N describes the regional lymph node criteria and M represents the involvement of distant metastases. By combining all TNM scores, the pathological stage group of melanoma is defined ranging from 0 to IV. Stage 0 describes melanoma in situ, where melanoma cells are exclusively found in the epidermis. Stage IA tumors show no ulceration and have a thickness below 1 mm (T1a), in contrast to Stage IB including tumors with or without ulceration depending on its thickness (T1b and T2a). In all stage I melanomas there are no lymph node or distant metastases detected (N0, M0). The same is true for stage II melanomas, which are thicker than stage I tumors and separated in A, B and C depending on its thickness and the presence of ulceration. Once melanoma cells have spread to lymph nodes ($N \geq 1$) one speaks of a stage III tumor. Stage III melanomas are again divided into four subgroups, A to D, based on the different N category criteria. In stage IV tumors, melanoma cells have spread to one or more distant sites ($M \geq 1$) independent on the thickness of the primary tumor (T category) and any involvement of lymph nodes (N category).

Compared to other cancers, melanoma primary tumors can be found relatively easy before manifestation of metastases due to their color. Once detected, melanomas are surgically removed, which is the primary treatment for cutaneous melanoma (Garbe et al. 2008). There are two major types of surgery used for melanoma removal, a wide excision and a lymph node dissection. In order to remove surrounding vessels, a wide excision surgery is performed cutting out the primary tumor and adjacent normal-looking tissue (Ethun and Delman 2016). Depending on the tumor thickness the size of the surgical margin is determined. As soon as the tumor tissue is removed, it is examined by a pathologist searching for cancer within the surgical margin. If there is still tumor detectable in the surrounding normal tissue, more surgery is needed. Depending on the characteristics of the primary tumor, the risk to develop metastases in regional lymph nodes or distant organs may be increased. Nearby lymph nodes are usually located in the nodal basins containing several lymph nodes. If there is a risk for developing lymph node metastasis, a sentinel lymph node (SLN) biopsy is performed (Gershenwald and Ross 2011), recommended for stage IB and stage II melanomas and in some cases for stage IA, suggested by the National Comprehensive Cancer Network guidelines. The SLN or tumor-draining lymph node is the first station after the primary tumor, where cancer cells can be found. For this, a dye is injected into the tissue surrounding the primary tumor. Via afferent lymphatic vessels the dye is transported to the nearest lymph node, thereby staining any draining lymph node blue. Excised sentinel nodes are examined for tumor cells based on histopathology. If the sentinel node is positive

(stage III melanoma), complete lymph node dissection (CLND) may be considered removing all lymph nodes of the nodal basin (Wagner et al. 2000). However, it was shown that CLND did not improve melanoma-specific or overall survival, but provides prognostic information (Faries et al. 2017).

If melanoma is unresectable, chemotherapy might be an adjuvant option for treatment. However, melanomas are not very sensitive to chemotherapy (Eigentler et al. 2003). For this, chemotherapy is much less effective than newer therapies like targeted or immune therapies. Targeted therapies are used as adjuvant treatment by targeting specific proteins that play a role in controlling melanoma cell growth and proliferation. As already mentioned, proliferation is amongst others controlled by the MAP kinase/ERK-signaling pathway including proteins like BRAF. Recent progress in survival and outcome of metastatic melanoma patients with mutated BRAF was achieved due to developing BRAF inhibitors targeting the constitutively activated mutated BRAF protein (Chapman et al. 2011). Furthermore, combining BRAF and MEK inhibitors improved progression free survival rates in patients with metastatic melanoma (Broman et al. 2019). Despite recent progress in the field of targeted therapy, in the last decades immunotherapy for melanoma treatment gained attention ranging from older drugs like Interleukin-2, an immune cell stimulating cytokine, to newer immunotherapeutic drugs like immune checkpoint blockades (Atkins et al. 2000, Lugowska et al. 2018).

1.1.3 Detection of disseminated cancer cells

In order to search for tumor cells in lymph nodes, the lymphatic tissue is analysed by standard histopathology. However, histopathological analysis is often limited to small sections and is not standardized. Hence, tumor cells located at a different part of the tissue might be missed. In contrast, immunocytochemical staining of lymph node single cell suspensions (Figure 1) allows the detection of single disseminated cancer cells (DCCs) in histopathologically negative sentinel lymph nodes and also in a significant number of non-metastatic (M0) patients (Ulmer et al. 2005, Ulmer et al. 2014). However, one can not find DCCs in all melanoma patients, as lymphatic dissemination is only found in 65% of melanoma patients (Werner-Klein et al. 2018). The presence of disseminated cancer cells (DCCs) is determined based on staining for the melanocytic antigen gp100, a transmembrane glycoprotein enriched in melanosomes. This enables to determine a quantitative documentation of the tumor burden by calculating the disseminated cancer cell density (DCCD). This value describes the number of gp100-positive cells per million lymph node cells (Ulmer et al. 2014). Although 60% of patients with gp100-positive sentinel lymph nodes (SLNs) have less than 10 disseminated cancer cells per million lymph node cells (unpublished data of work group and (Ulmer et al. 2014)), cancer cell spread to SLNs is a quantitative risk for melanoma death (Werner-Klein et al. 2018) and increases the risk to develop in-transit, lymph node and distant metastases (Dalal et al. 2007).

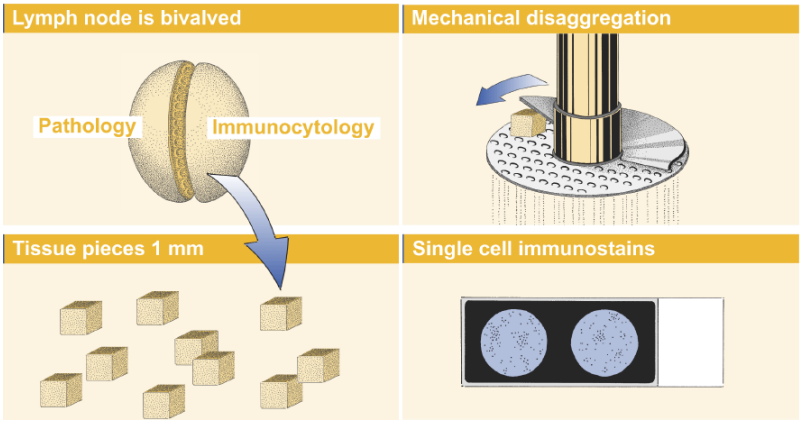


Figure 1: Lymph node disaggregation for detecting melanoma cells
One half of the lymph node is cut into pieces and subjected to mechanical disaggregation. The resulting single-cell suspension is transferred onto slides for immunocytological analysis. The image is adapted from (Ulmer et al. 2014).

Moreover, gp100-positive cells have copy number alterations (CNAs) detected by metaphase comparative genomic hybridization, showing the malignant origin of isolated cells (Ulmer et al. 2014). However, as gp100 is located within the cell, permeabilization of the cell membrane is necessary for successful staining. This makes it impossible to get viable cancer cells for downstream analyses. To solve that problem, another marker for single DCC detection can be used. The melanoma-specific cell surface antigen MCSP, Melanoma-associated-chondroitin-sulfate-proteoglycan, enables the isolation of viable tumor cells, as it is not located inside the cell (Ulmer et al. 2004). However, MCSP is less specific, as it can also be found on different normal cells (Campoli et al. 2004).

1.2 Role of neoantigens in immunotherapy

1.2.1 Immunotherapy in melanoma

Melanoma has a high somatic mutation rate (>10 mutations/ megabase) compared to other cancer types (Alexandrov et al. 2013) and is also a very immunogenic cancer eliciting tumor-specific immune responses, mainly mediated by CD8+ cytotoxic T cells (Blankenstein et al. 2012). Clinical data further point towards a pivotal role of tumor-reactive T cells in the control of melanoma. For example, high objective response rates and long-lasting complete tumor regression are achieved in up to 25% of melanoma patients undergoing adoptive T cell therapy (Rosenberg et al. 2011, Merhavi-Shoham et al. 2017). Moreover, immune checkpoint therapies targeting key regulators of T cell responses like CTLA-4 (Ipilimumab) or PD-1 (Nivolumab) demonstrate high efficacy in melanoma reaching response rates of up to 58% in metastatic patients (Wolchok et al. 2017). Combining different checkpoint inhibitors further improves overall survival (Larkin et al. 2019). However, a long-lasting survival benefit is observed in only about 20% of the patients treated with Ipilimumab and Nivolumab (Wolchok et al. 2017). But, patient response to immune checkpoint therapies is associated with prior infiltration of the tumor tissue with T cells indicating the necessity for an endogenous or vaccine-induced T cell response (Powderly et al. 2013, Huang et al. 2017, Riaz et al. 2017). In this line, a high mutational load correlates with clinical benefit from immune checkpoint blockade therapy (Van Allen et al. 2015, Hugo et al. 2017) and is significantly associated with improved progression-free and overall survival in melanoma (Lauss et al. 2017). Of note, a low level of intratumor heterogeneity is associated with better survival and clinical response is suggested to be associated with higher levels of tumor-infiltrating immune cells in studies of different cancer types (Morris et al. 2016, McGranahan et al. 2016, Poran et al. 2020). Immunotherapy can further activacte CD8+ cytotoxic T cells regulating iron-dependent programmed cell death of cancer cells (Wang et al. 2019). Nevertheless, in addition to the central role of CD8+ cytotoxic T cells in the field of immunotherapy, also tumor-infiltrating CD4+ T cells are important showing to be tumor-reactive (Ahmadzadeh et al. 2019).

1.2.2 Neoantigens and (neo)peptide-presentation

In principle, tumor-reactive T cells can be directed against three categories of antigens: (i) tissue-specific differentiation antigens, (ii) self-antigens overexpressed in the tumor tissue or not expressed in most other adult tissues and (iii) non-self, somatically mutated neoantigens, exclusively found in tumor cells. Differentiation and overexpressed self-antigens usually induce weak immune responses due to the self-tolerant T cell pool, in contrast to tumor-specific neoantigens. Neoantigens derive from non-synonymous protein sequence-changing mutations accumulated during tumor progression and are able to elicit T cell responses. Resulting mutated proteins are either cleaved into 8 to 11 amino acid long fragments, presented on major histocompatibility complex (MHC) class I molecules, or in longer peptides (> 12 amino acids) presented on MHC class II molecules. The human MHC is called human leukocyte antigen (HLA) complex and is a highly polymorphic genetic system (Howell et al. 2010, Rajalingam et al. 2010, Choo 2007). The genetic region is located on chromosome 6p21.3 and includes

around 3.6 megabases DNA encoding for proteins that play an important role in regulating immune responses. MHC molecules can be categorized into three classes. MHC class I melocules comprise HLA-A,B,C molecules and class II include HLA-DR,DP,DQ. Class III molecules are not HLA molecules, but have physiological roles and are proteins of the complement system, cytokines or heat shock proteins. HLA class I molecules are expressed in all nucleated cells and present short peptides (8-11 amino acids) of cytosolic proteins on their surface, recognized by CD8+ T cells. In contrast, HLA class II molecules are expressed in B cells, antigen-presenting cells (monocytes, macrophages, dendritic cells) and activated T cells presenting 13 to 18 amino acid long peptides of proteins from endosomes or lysosomes on their surface. Peptides presented on HLA-II molecules are recognized by CD4+ T cells. The level of polymorphism of HLA genes is very high. In 2019, the HLA system covered up to 25000 HLA alleles with the number of identified alleles steadily increasing over time due to the use of next-generation sequencing (NGS) technologies (Robinson et al. 2020). The HLA nomenclature consists of the HLA gene separated by an asterisk from a four-or-more-digit number describing different factors of the HLA system, like for example HLA-A*68:01. The first field after the HLA gene (two-digit resolution) describes the allele group and the second (four-digit resolution) the specific HLA protein. As every person has a diploid set of chromosomes, every human has two alleles of each HLA molecule. Mutated proteins including neoantigens can be presented on both, HLA-I and HLA-II molecules (Alspach et al. 2019).

1.2.3 Detection of neoantigens and their relevance for immunotherapy

Neoantigens causing tumor rejection were already discovered in 1957 and 1995 in seminal studies of Richmond Prehn and Hans Schreiber, respectively (Beck-Engeser et al. 2001, Prehn and Main 1957). In addition, mutated peptides are predicted to have a higher affinity for HLA molecules than the respective non-mutant peptides (Snyder et al. 2014). However, their applicability as important immunotherapeutic targets for a wide range of patients was certainly neglected due to their highly personalized nature, but has become feasible with the development of next-generation sequencing coupled with computational analyses. With the availability of whole exome sequencing (WES) data of tumor and normal samples from the same patient, tumor-specific non-synonymous mutations, i. e. neoantigens, can be identified bioinformatically. The information from *in silico* typing of the HLA alleles of the patients can than be intersected with algorithms to predict the binding affinities of neopeptides, created from identified neoantigens, to these HLA alleles. The bioinformatic pipeline for *in silico* prediction of neoepitopes from WES data comprises three different computational tasks: (i) HLA typing, (ii) prediction of mutations and (iii) selection of mutated peptides with high HLA binding affinities. There are several HLA typing tools with different accuracies and efficiencies to predict HLA genotypes from WES data, which can be categorized into alignment-based or assembly-based methods (Bauer et al. 2018). In general, if sequencing reads are aligned against reference HLA sequences and true alleles are predicted based on probabilistic models, one speaks of an alignment-based approach. Assembly-based methods are described by assembling reads into contigs and aligning them to reference sequences of known HLA alleles. In this study, the alignment-based OptiType tool (Szolek et al. 2014), which can only predict HLA-I alleles, and the assembly-based HLAminer tool (Warren et al. 2012), able to predict HLA-I and -II alleles, were conducted. For WES data of bulk samples OptiType has a higher accuracy than HLAminer (Bauer et al. 2018). Regarding the prediction of mutations, single nucleotide variants are identified by comparing tumor and non-tumor NGS data. Within the last years more than 40 variant calling tools were developed (Sandmann et al. 2017). In this project two somatic variant callers were conducted: Varscan2 (Koboldt et al. 2012), which relies on heuristic, statistical methods for variant identification, and Platypus (Rimmer et al. 2014), which uses the bayesian inference. For predicting the HLA binding affinities of peptides resulting from identified mutations several algorithms are currently available (Richters et al. 2019). Most tools are trained to predict binding affinities of short peptides to HLA-I molecules. Predicting class II neoantigens is still challenging

and not as accurate as predicting class I neopeptides. Most applied and well-documented algorithms are the NetMHCpan tools used to predict binding affinities to HLA-I molecules (Nielsen et al. 2007, Nielsen and Andreatta 2016).

As less than 1% of all expressed antigens of a cell are presented in the context of HLA molecules on the cell surface (Lamers et al. 2018) and therefore potentially therapeutically relevant due to T cell recognition, functional evaluation of potential neoantigens identified by NGS is mandatory. The immunogenicity of candidate neoantigens can be analysed by detecting neoantigen-specific T cell responses. However, the frequency of neoantigen-specific T cells within blood lymphocytes is very low (Cohen et al. 2015). In order to increase the chance to detect neoantigen-specific T cells among tumor-reactive T cells, tumor-reactive T cells have to be enriched by *in vitro* expansion. This can be done by repeatedly exposing peripheral blood cells to autologous tumor cells, for example cells from patient-derived organoids (Dijkstra et al. 2018). In order to achieve optimal neopeptide presentations on the surface of the tumor cell, tumor cells can be stimulated with Interferon-γ, a potent biological response mediator produced by immune T cells, shown to induce and enhance the expression of HLA class I and class II antigens (Carrel et al. 1985, Hokland et al. 1988). Tumor-reactive T cells can be identified based on the production of effector molecules (Interferon-γ, CD107a) or the expression of certain antigens like CD137, PD-1, CD103 or CD39 (Cafri et al. 2019, Dijkstra et al. 2018, Duhen et al. 2018, Gros et al. 2016). In the end, neoantigen-specific T cells can be identified among tumor-reactive T cells by specifically recognizing antigen-presenting cells transfected with minigenes encoding tumor-specific mutations (Prickett et al. 2016). A different strategy to demonstrate the immunogenicity of candidate neoantigens identified by NGS is described by HLA ligandome analysis (Rammensee and Singh-Jasuja 2013). Here, presented peptides are eluted from HLA molecules of the tumor cells and tested by mass spectrometry whether candidate neoantigens are present among these. However, this approach requires specialized technical expertise and equipment, and only has a detection limit of 30-59% (personal communication Prof. Rammensee).

Using a high-throughput personalized screening approach, neoantigen-specific lymphocytes can be detected in the peripheral blood of melanoma patients (Gros et al. 2016) and combining HLA peptidomics with neoantigen prediction pipelines allows identification of targetable neoantigens in human melanomas (Kalaora et al. 2018). Moreover, within the last four years clinical trials with cancer-specific neoantigens have been reported in metastatic melanoma patients using different vaccination strategies (Ott et al. 2017, Sahin et al. 2017, Hu et al. 2021, Sahin et al. 2020). Here, metastatic patients developed CD4+ and CD8+ T cell responses against multiple vaccine neoepitopes and in combination with immune checkpoint inhibitor anti-PD-1 some patients showed complete tumor regression. However, it is currently unclear which role neoantigens and vaccine-induced neoantigen-specific T cells could play for adjuvant therapy.

1.3 Metastatic seeding

In contrast to the prevailing view that primary tumor and disseminated cancer cells (DCCs) are genetically highly similar, a recent study has shown that these early disseminating melanoma cells acquire genetic alterations important for metastatic colony formation not in the primary tumor, but in the sentinel lymph node (Werner-Klein et al. 2018). As a consequence, primary tumor and early DCCs may share only few common genetic alterations. This genetic disparity between primary tumor and SLN-DCCs might hold true for DCCs at other distant sites as well. Therefore, target identification directly on DCCs and not the primary tumor, as is currently done, may be mandatory for the elimination of the early systemic metastatic seed (Rotman et al. 2019). An ideal, effective adjuvant (immuno-) therapy should eliminate all disseminated cancer cells (Klein 2020). This requires that genetic alterations between DCCs used for target identification, e.g. SLN-DCCs, and those DCCs that have

spread to more distant sites like regional lymph nodes, liver or lung are the targets of therapy and are shared to a high degree (Figure 2 A). Support for this notion comes from a study showing that the sensitivity to PD-1 and CTLA-4 blockades in patients with advanced melanoma was enhanced in tumors enriched for clonal neoantigens (Efremova et al. 2017). To which extent neoantigens are shared between SLN-DCCs and more systemically spread DCCs depends on the order of metastatic spread (Figure 2 B). DCCs may spread simultaneously from the primary tumor directly to the SLN, non-sentinel lymph nodes (NSLNs) and distant metastatic sites or DCCs may disseminate in cascades from the primary tumor to the SLN, then to NSLNs and at last to metastatic sites. However, a mixture of parallel and cascadic seeding might also be possible. Clinical data indirectly suggest a parallel seeding as immediate complete)lymphadenectomy does not increase melanoma-specific survival of patients showing that distant metastases are not formed by melanoma cells that have disseminated to lymph nodes (Faries et al. 2017). In contrast, data from experimental models show that DCCs can use lymph node vessels as exit routes for systemic spread indicating seeding in cascades (Brown et al. 2018, Pereira and Kedrin 2018). In addition, lymph vessels are shown to protect disseminated melanoma cells from iron-dependent programmed cell death thereby increasing their chance of survival during dissemination (Ubellacker et al. 2020).

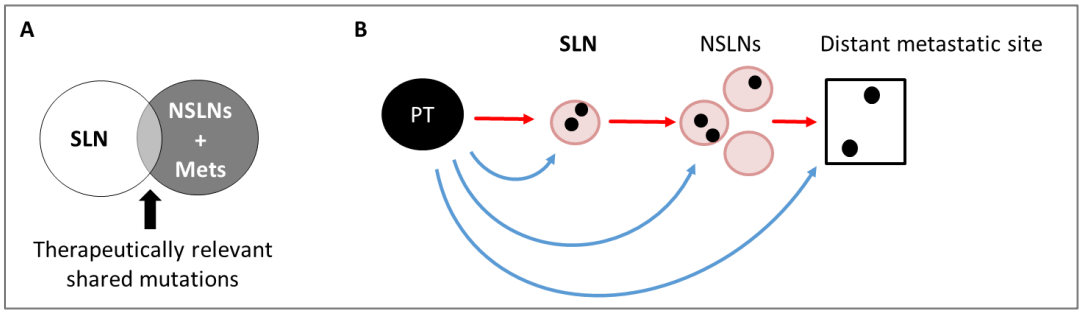


Figure 2: Therapeutically relevant mutations and scenarios of metastatic seeding
A: Degree of shared mutations. Only mutations shared between DCCs from SLN and NSLN/ Mets are expected to be therapeutically relevant. **B:** Possible scenarios of metastatic seeding. On the one hand, DCCs (marked as ●) may spread from the primary tumor directly to the SLN, NSLNs and distant metastatic sites indicating a parallel seeding (blue arrows). On the other hand, DCCs can disseminate in cascades from the primary tumor to the SLN, then to NSLNs and at last to metastatic sites (red arrows). ◍ lymph node; □ cutaneous metastasis; DCC: disseminated cancer cell; PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; Mets: metastasis.

1.4 Methods for phylogenetic analyses

There are different methods for analysing the phylogeny of cancer cells. Microsatellites (MS), also known as short tandem repeats, are highly abundant in the human genome and are among the most polymorphic loci therein (Willems et al. 2014). As they are non-coding sequences, they are not under selection pressure. Moreover, their mutation rate with 10^{-6} to 10^{-2} mutations per locus is much higher than the average rate of point mutations (Eckert and Hile 2009). Hence, MS loci are ideal mutational sources that allow tracking of the number of cell divisions (Baron and van Oudenaarden 2019). Developments in the field of next-generation sequencing enable uncovering the human cell lineage by analysing 12472 MS loci (Tao et al. 2017, Tao et al. 2018). However, the MS-based method requires not only the analysis of tumor cells, but also of non-tumor cells from different cell lineages. Only if cells from different lineages can be separated, the information on the individual localization of single tumor cells will be reliable. In order to address evolutionary aspects of the disease and thus assess the impact of selection pressure on the phylogeny of tumor cells, lineage analyses can be performed based on genetic changes like copy number alterations (CNAs) or somatic mutations. These genomic aberrations, in contrast to non-coding MS mutations, can promote tumor occurrence and progression leading to survival of cancer cells (Hanahan and Weinberg 2011). CNAs are diagnostically considered for the differentiation of non-tumor and tumor cells (Bauer and Bastian 2006). They are found in less

than 5% of non-tumor cells (Jakubek et al. 2020) indicating that CNAs better reflect cancer progression than point mutations, as they are frequently found also in benign lesions (Hafner et al. 2010, Shain et al. 2015, Klein 2013).

1.5 Aim of the thesis

As mortality from metastatic melanoma is still high despite recent therapy progress, all efforts must be directed to support primary surgery in an attempt to prevent manifestation of metastasis and cure the patient early on. This adjuvant therapy setting is characterized by minimal residual disease, i.e. cancer cells that have spread before surgical removal of the primary tumor and that are undetectable by clinical imaging. Clinical and experimental data point towards a pivotal role of neoantigen-reactive T cells in controlling metastatic melanoma. However, adjuvant antigen-specific immunotherapy is a *terra incognita*, because fundamental information is lacking. This study aims to close this gap by addressing the question whether DCCs isolated from sentinel lymph nodes that are routinely extirpated for diagnostic reasons can be used as surrogates for systemic DCCs and enable target identification for adjuvant therapy. To this end, the phylogenetic relationship and neoantigen evolution and conservation between local and systemically spread melanoma cells will be explored over the disease course and various metastatic organs (time and space). Thereby, this project will address fundamental questions of adjuvant immunotherapy against early systemic cancer of melanoma and possibly other cancer types. Specifically the study will assess (i) the phylogenetic relationship between DCCs at different anatomic sites and the order of metastatic seeding, (ii) the degree of neoantigen evolution and conservation between DCCs at different anatomic sites and (iii) potential mechanisms of immune evasion across time and space. These scientific questions will be addressed by: (i) obtaining tumor and non-tumor (control) cells from longitudinally tracked melanoma patients, (ii) isolating DNA from tumor and control cells for single cell molecular analyses, (iii) sequencing for 12472 microsatellites (MS) and copy number alterations (CNA) to evaluate the evolution of cancer cells and (iv) whole exome sequencing (WES) for neoantigen identification to assess the therapeutic relevance of disseminated melanoma cells for treating early systemic cancer.

2 Materials

2.1 Human samples

Cells and bulk DNA were isolated from various fresh or fixed tissues (primary tumor, lymph nodes, metastasis, blood, buccal swabs) sent for analysis to the Chair of Experimental Medicine and Therapy Research, University of Regensburg. The study was approved by the ethics committee of the University of Regensburg (07-079, 18-948-101). Written informed consent was obtained from all patients.

2.2 Patient-derived cell lines

Four patient derived cell lines were used in this project deriving from three melanoma patients which have already been established by Fraunhofer ITEM-Regensburg before the project started (Table 1). Cell line MelDTC_9 was developed from a DCC-derived xenograft transplanting 55 MCSP-positive SLN-DCCs from patient MM15-127 into an immunodeficient mouse. MelDTC_10a and MelDTC_10b derived from spheres generated from DCCs from NSLN3 and NSLN6 from patient MM16-412. MelDTC_11 was also developed via sphere culture from NSLN1 cells of patient MM16-423. Additionally, DCCs from NSLN1 were used for xenograft generation and transplanted into immunodeficient mice. Cell lines were either mutated in *BRAF* or *NRAS* as determined by Sanger sequencing (Sequiserve, Vaterstetten, Germany) and fingerprint analysis (GenePrint® 10 System, Promega) verified their patient origin. Melanoma origin was confirmed by a human pathologist and their aberrant genotype by LP-Seq showing copy number alterations. All cell lines and xenografts were tested for HBV, HCV and HIV showing negative results.

Table 1: Characteristics of patient-derived tumor cell lines

Cell line	Patient	Generation from	Source	Mutation status
MelDTC_9	MM15-127	55 cells	SLN	BRAF wt; NRAS 50% mut, Q61R (c.182A>G)
MelDTC_10a	MM16-412	Spheres	NSLN3	BRAF wt; NRAS 100% mut, Q61K (c.181C>A)
MelDTC_10b	MM16-412	Spheres	NSLN6	
MelDTC_11	MM16-423	Spheres	NSLN1	BRAF 100% mut, V600E (c.1799T>A); NRAS wt

2.3 Reagents

2.3.1 Chemicals and commercial solutions

Table 2: List of chemicals and commercial solutions

Name	Manufacturer	Catalogue number
1 kb Plus DNA Ladder + Dye	NEB	N3200L
AB serum, human (for blocking unspecific AB binding)	Bio-Rad	805135
AB serum, human (for cell culture)	Valley Biomedical	HP1022
Acetic acid (100%)	Merck	1000631000
Adenosine triphosphate (ATP) 100 mM	Roche Diagnostics	11140965001
Agarose LE	Anprotec	AC-GN-00009
AIM V™ Medium, liquid (research grade)	Thermo Fisher	12055091
Ampicillin	Applichem	A0839
AMPure XP purification beads	Beckman Coultier	A63882
AP-polymer anti-mouse solution	Zytomed Systems	ZUC077-100
Atipamezole	Pharmacy of University Hospital Regensburg	Not available
Bacto™ Peptone	Becton Dickinson	211677
β-Mercaptoethanol 55 mM (for cell culture)	Gibco	21985
β-Mercaptoethanol 14.3 M (for RNA isolation)	Sigma Aldrich	M3148-25ML
Boric acid (H_3BO_3)	Sigma Aldrich	31146-500G
Bovine serum albumie (BSA) (for picking and basal medium)	Sigma Aldrich	B8667-5ml

Bovine serum albumie (BSA) (for PCRs)	Roche Diagnostics	10711454001
Chloroform	Sigma Aldrich	C2432-4x25ml
Cutsmart (10x)	NEB	B7204S
Deparaffinization solution	QIAGEN	19093
dNTP Mix; 10 mM each (for cDNA)	Promega	U1515
dNTP Set; 100 mM each A, C, G, T; 4x 24 µM	GE Healthcare	28-4065-51
DMEM/Ham's F12 Medium	PAN Biotech	P04-41500
DMSO	Sigma Aldrich	D4540-100ml
eBioscience™ Cell Stimulation Cocktail (plus protein transport inhibitors), 500x	Thermo Fisher Scientific	00-4975-93
eBioscience™ Protein Transport Inhibitor, 500x	Thermo Fisher Scientific	00-4980-93
Eosin solution (5%)	Pharmacy of University Hospital Munich	Not available
Eukitt® quick-hardening mounting medium	Sigma Aldrich	03989-100ML
Ethanol absolute ≥99.8%, Mol. Bio.Grade 250 ml	VWR International GmbH	437433T
Ethidium Bromide Solution (10 mg/ml)	Sigma Aldrich	E1510-10ML
Ethylenediaminetetraacetic acid (EDTA)	J.T. Baker	B-1073.1000
Eye ointment	Pharmacy of University Hospital Regensburg	Not available
Expand Long Template Buffer 1	Roche Diagnostics	11759060001
FastStart dNTP mix	Roche Diagnostics	4738420001
FastStart PCR buffer with $MgCl_2$	Roche Diagnostics	4738420001
Fentanyl	Pharmacy of University Hospital Regensburg	Not available
Fetal calf serum (FCS) sera Plus	PAN Biotech	P30-3702
Flumazenile	Pharmacy of University Hospital Regensburg	Not available
FluoroFix™ Buffer	Biolegend	422101
Formaldehyd (<5%)	Pathology, University Hospital Regensburg	605-001-005
Gel loading dye 6x, purple, no SDS	NEB	B7025S
Hank's balanced salt solution 10x	Biochrom AG	L2045
HBSS (10x)	PAN Biotech	P04-49100
Hepes (1 M)	Sigma	H0887
Hematoxylin Solution, Mayer's	Sigma Aldrich GmbH	MHS16-500ml
Igepal CA-630 viscous liquid	Sigma Aldrich	I3021-20ml
Intracellular Staining Permeabilization Wash Buffer (10x)	Biolegend	421002
Iodine	Pharmacy of University Hospital Regensburg	Not available
KAPA Pure Beads	Roche Diagnostics	07983271001
Kodan spray	Pharmacy University Hospital Regensburg	Not available
LB-Agar	Carl Roth	X965.2
LB-Medium	Carl Roth	X968.2
L-Glutamine (200 mM)	PAN Biotech	P04-80100
Low molecular weight DNA ladder	NEB	N3233L
NEBNext® FFPE DNA Repair Mix	NEB	M6630S
Magnesium chloride ($MgCl_2$) solution 1 M	Sigma Aldrich	M1028-100ml
Matrigel® Basement Membrane Matrix HC	BD Bioscience/ Corning	354248
Medetomidine	Pharmacy of University Hospital Regensburg	Not available
Midazolam	Pharmacy of University Hospital Regensburg	Not available
Monopotassium phosphate (KH_2PO_4)	VWR International	1048731000
mTRAP™ Lysis Buffer	Active Motif	29011
Naloxone	Pharmacy of University Hospital Regensburg	Not available
Oligo(dT)	Fermentas	SO132

One Shot™ TOP10 Chemically Competent E. coli	Thermo Fisher	C404010
Opti-MEM™	Life Technologies	11058-021
Penicillin (10.000 U/ml) / Streptomycin (10 mg/ml)	PAN Biotech	P06-07100
Percoll	GE Healthcare	17089101
Phenol/Chloroform/Isoamyl alcohol	Roth	A156.3
Poly(2-hydroxyethyl methacrylate) / Poly-HEMA	Sigma Aldrich	P3932
Potassium Chloride Solution 1M	Sigma Aldrich	60142-100ML-F
RPMI 1640	Anprotech	AC-LM-0060
Sodium azide ReagentPlus®, NaN_3 (for FACS buffer)	Sigma Aldrich	S2002-5G
Sodium acetate 3 M	Ambion	AM9740
Sodium chloride solution 5 M	Sigma Aldrich	71386-1L
Tris ultrapure	AppliChem	A1086,1000
Tris EDTA (TE) buffer 1x pH 8.0 low EDTA	AppliChem	A8569,0500
tRNA from *E. coli* MRE 600	Roche Diagnostics	10109541001
Trypan blue	Sigma Aldrich	T8154-20ml
Trypsin/ EDTA (10x) (0.05% Trypsin/ 0.02% EDTA)	PAN Biotech	P10-024100
Tween20	Sigma Aldrich	P9416-50ml
UltraPure™ DNase/RNase-free water	Life Technologies	10977035
Water for chromatography, LiChrosolv®, LC-MS grade (PCR-water)	Merck	1.15333.1000
Water, aqua ad iniectabilia (NGS-water)	Braun	2351744
Water, demineralized	Taken from tap	Not available
X-Vivo 20	Biozym	881030
Xylol	VWR Chemicals	28973.363

2.3.2 Custom buffers, media and solutions

Table 3: List of custom buffers and solutions

Name	Components
Basal medium	500 ml DMEM/Ham's F12 Medium 5 ml 1 M Hepes 5 ml Penicillin/ Streptomycin 5 g BSA
Blocking solution	5 ml AB serum, human 5 ml 20% Peptone solution 40 ml 1x PBS pH 7.4
Digest medium	500 ml RPMI 5 ml 1 M Hepes 5 ml penicillin/ Streptomycin 5 g BSA
FACS buffer	10 ml FCS 2 ml 0.5 M EDTA 0.5 ml 10% NaN_3 PBS ad 500 ml
LB ampicillin agar solution	500 ml LB-Medium 17.5 g LB-Agar 0.5 ml 100 mg/ml ampicillin
MACS buffer	47.25 ml PBS 2.5 ml 0.5 M EDTA 250 µl AB serum, human
One Phor All (OPA) buffer	5 ml 1 M Tris acetate 5 ml 1 M Magnesium acetate 1 ml 5 M Potassium acetate PCR-water ad 1 L sterile filtered
Peptone solution 20%	100 g Bacto™ Peptone 500 ml 1x PBS pH 7.4 sterile filtered

Percoll 100%	100 ml Percoll stock 9 ml Hank's balanced salt solution sterile filtered
PBS pH 7.4 (10x)	450 g Sodium chloride 71.65 g Disodium phosphate (Na_2HPO_4) 13.35 g Monopotassium phosphate (KH_2PO_4) distilled water ad 5 L
Tris/Borate/EDTA (TBE) buffer (10x)	539 g Tris 275 g Boric acid 37 g EDTA 5 L demineralized water

2.3.3 Enzymes and cytokines

Table 4: List of enzymes and cytokines

Name	Manufacturer	Catalogue number
Alkaline Phosphatase, Calf Intestinal (CIP, 10000 U/ml)	NEB	M0290S
Collagenase	Sigma Aldrich	C9407-1G
DNaseI	Roche	11284932001
EcoRI (20 000 U/ml)	NEB	R0101S
Human GM-CSF (for Fast-DCs)	Bayer	6701801
Human GM-CSF, premium grade (for one-week DCs)	Miltenyi Biotech	130-093-865
Human IL-1β, premium grade (for Fast-DCs)	Miltenyi Biotech	130-093-898
Human IL-4, premium grade (for one-week DCs)	Miltenyi Biotech	130-093-921
Human IL-4, research grade (for Fast-DCs)	Miltenyi Biotech	130-094-117
Human IL-6, premium grade (for Fast-DCs)	Miltenyi Biotech	130-093-932
Human TNF-α (for Fast-DCs)	PromoCell	C-63719
Hyaluronidase	Sigma Aldrich	H4272-30MG
FastStart Taq Polymerase	Roche Diagnostics	4738420001
M-MLV Reverse Transcriptase and buffer	Promega	M5313; M1705
MseI, recombinant, 2500 U; 50000U/ml	NEB	R0525M
NcoI-HF (20000 U/ml)	NEB	R3193S
Pol Mix 5 U/µl (Expand Long Template enzyme mix)	Roche Diagnostics	11759060001
Prostaglandin E_2 /PGE2 (for Fast-DCs)	Sigma-Aldrich	P6532-1MG
Proteinase K, recombinant	Roche	3115828001
Recombinant human Interferon-γ (IFN-γ)	PeproTech	300-02
Recombinant human Interleukin-2 (IL-2), Proleukin®	Clinigen	05060229220264
SpeI (10000 U/µl)	NEB	R0133S
T4 DNA Ligase 400 U/µl (for cloning)	NEB	M0202S
T4 DNA Ligase 500 U, 5 U/µl	Roche Diagnostics	10799009001
XbaI (20000 U/ml)	NEB	R0145S

2.3.4 Antibodies and dyes

Table 5: List of antibodies and dyes

Name	Application	Manufacturer	Catalogue number
Alexa Fluor® 647 anti-human HLA-DR, DP, DQ	DC generation	Biolegend	361703
Alexa Fluor® 700 anti-human CD8a	Marker in co-culture	Biolegend	300920
Anti- human PD-1 / Pembrolizumab	Blocking in co-culture	KEYTRUDA	06G229851A
APC anti-human CD8a	Minigene test	Biolegend	300911
APC anti-human CD103 (Integrin αE)	Marker in co-culture	Biolegend	350215
APC anti-human HLA-A,B,C	Marker cell lines	Biolegend	311409
APC/Cy7 anti-human CD3	Minigene test	Biolegend	344817
APC/Fire™ 750 anti-human IFN-γ	Marker in co-culture	Biolegend	502547
APC/Fire™ 750 Mouse IgG1, κ Isotype Ctrl	Marker in co-culture	Biolegend	400195
APC Mouse IgG1, κ Isotype Ctrl	Marker in co-culture	Biolegend	400121
APC Mouse IgG2a, κ Isotype Ctrl	Marker cell lines	Biolegend	400220

BD™ CompBead Anti-Mouse Ig, κ/Negative Control Compensation Particles Set	Compensation	BD Biosciences	552843
Brilliant Violet 421™ anti-human CD4	Marker in co-culture	Biolegend	317434
Brilliant Violet 421™ anti-human CD80	Marker cell lines + DC generation	Biolegend	305221
Brilliant Violet 421™ anti-human CD86	Marker cell lines + DC generation	Biolegend	305425
Brilliant Violet 421™ Mouse IgG1, κ Isotype Ctrl	Marker cell lines	Biolegend	400157
Brilliant Violet 421™ Mouse IgG2b, κ Isotype Ctrl	Marker cell lines	Biolegend	400341
Brilliant Violet 510™ anti-human CD19	DC generation	Biolegend	302241
Brilliant Violet 650™ anti-human CD137 (4-1BB)	Marker in co-culture	Biolegend	309827
Brilliant Violet 650™ Mouse IgG1, κ Isotype Ctrl	Marker in co-culture	Biolegend	400163
Dynabeads™ CD19 Pan B	DC purification	Thermo Scientific	11143D
Dynabeads™ CD3	DC purification	Thermo Scientific	11151D
Dynabeads™ Pan Mouse IgG	DC purification	Thermo Scientific	11041
FITC anti-human CD3 (clone HIT3a)	Single cell isolation	Biolegend	300305
FITC anti-human CD39	Marker in co-culture	Biolegend	328205
FITC anti-human CD45	Minigene test	Biolegend	304006
FITC anti-human HLA-DR,DP,DQ	Marker cell lines	Biolegend	361705
FITC Mouse IgG1, κ Isotype Ctrl (FC)	Marker in co-culture	Biolegend	400109
FITC Mouse IgG2a, κ Isotype Ctrl	Marker cell lines	Biolegend	400209
Goat anti-mouse-Cy3	Single cell isolation	Jackson	115-166-071
Goat anti-mouse IgG-Alexa Fluor 488	Single cell isolation	Invitrogen	A-11029
Goat anti-rabbit IgG-Alexa Fluor 555	Single cell isolation	Thermo Scientific	A-21428
IgG1κ (clone MOPC 21)	Single cell isolation	Sigma Aldrich	M9269-1MG
IncuCyte® Caspase-3/7 Green Apoptosis Assay Reagent	Killing assay	Essen BioScience	4400
Mouse anti-human Actin [alpha smooth muscle isoform] (clone ASM-1)	Melanoma marker	Chemicon/ Millipore	CBL171
Mouse anti-human CD45 (clones 2B11 + PD7/26)	Staining of PT section	DAKO	M0701
Mouse anti-human CD56	DC purification	Beckman Coulter	IM1844
Mouse anti-human CD68 (clone KP1)	Single cell isolation	Agilent Dako	M081401-2
Mouse anti-human CD68 (clone PG-M1)	Staining of PT section	DAKO	M0876
Mouse anti-human Cytokeratin (clones AE1/AE3)	Staining of PT section	DAKO	M3515
Mouse anti-human Desmin (clone D33)	Melanoma marker	DAKO	M0760
Mouse anti-human Melanosome (clone HMB45)	Single cell isolation	DAKO	M0634
Mouse anti-human MCSP (clone 9.2.27)	Single cell isolation	BD Pharmingen	554275
Mouse anti-human Melan-A (clone A103)	Melanoma marker	DAKO	M7196
Mouse anti-human Microphthalmia Transcription Factor (MiTF) (clones C5/D5)	Melanoma marker	Cell Marque/ Sigma Aldrich	CMC28429030
Mouse anti-human PECAM1/ CD31 (clone JC70A)	Single cell isolation	Agilent Dako	M082329-2
Mouse anti-human S100 (clone 4C4.9)	Melanoma marker	Zytomed Systems	MSK050
Pan Melanoma Cocktail (HMB45, MART-1, Tyrosinase)	Melanoma marker	Biocare Medical	CM165A
PE anti-human CD19 (clone HIB19)	Single cell isolation	Biolegend	302207
PE anti-human CD155	Marker cell lines	Biolegend	337609
PE anti-human CD209	DC generation	Biolegend	330105

PE anti-human CD274 (B7-H1, PD-L1)	Marker cell lines	Biolegend	329705
PE anti-human CD279 (PD-1), clone EH12.2H7	Marker in co-culture	Biolegend	329905
PE anti-human Galectin-9	Marker cell lines	Biolegend	348905
PE anti-human IFN-γ	Minigene test	Biolegend	502509
PE/Cy7 anti-human CD56	DC generation	Biolegend	318317
PE/Cy7 anti-human CD107a (LAMP-1)	Marker in co-culture	Biolegend	328617
PE/Cy7 Mouse IgG1, κ Isotype Ctrl	Marker in co-culture	Biolegend	400125
PE Mouse IgG1, κ Isotype Ctrl	Marker cell lines	Biolegend	400139
PE Mouse IgG2b, κ Isotype Ctrl	Marker cell lines	Biolegend	400313
PerCP/Cyanine5.5 anti-human CD3	Marker in co-culture + DC generation	Biolegend	344808
Purified anti-human CD3 (clone OKT3)	Stimulation of T cells	Biolegend	317301
Purified anti-human CD28 (clone CD28.2)	Stimulation of T cells	Biolegend	302901
Rabbit anti-human CD3	Single cell isolation	Sigma Aldrich	C7930
Rabbit anti-human CD3 (polyclonal)	Staining of PT section	DAKO	A 0452
Rabbit IgG	Single cell isolation	Southern Biotech	0111-01
Viability dye eFluor 780	Marker detection	ebioscience	65-0865-14
Zombie Aqua™ Fixable Viability Kit	Marker in co-culture	Biolegend	423101

2.3.5 Oligonucleotides and primers

Primers for WGA-QC were obtained from Eurofins Germany in HPSF grade. Primers and oligonucleotides for ASB-PCR were purchased from Metabion.

Table 6: List of oligonucleotides and primers

Name	Sequence (5' -> 3')	T_A	Amplicon size (bp)	Application
ACTB for ACTB rev	GCG TGA CAT TAA GGA GAA GCT G CGC TCA GGA GGA GCA ATG AT	58	378	WTA-QC
BRAF V600E mut	GGACCCACTCCATCGAGATTTC**T**	58	147*	ASB-PCR
BRAF V600E block	CGAGATTTCACTGTAGCTAG-*PO_4*	58	-	ASB-PCR
BRAF Exon15_3' BRAF Exon15_5'	CTCTTCATAATGCTTGCTCTG TCCAGACAACTGTTCAAACTG	58	171	WGA-QC bulk + ASB PCR
CKND2A Exon3 for CKND2A Exon3 rev	TGGCTCTGACCATTCTGTTC TGGAAGCTCTCAGGGTACAA	58	380	WGA-QC bulk
D5S2117 for D5S2117 rev	CCA GGT GAG AAC CTA GTC AG ACT GTG TCC TCC AAC CAT GG	58	140	WGA-QC
ddMse11	TAACTGACAG-*ddC*	65	-	WGA
EF1alpha for EF1alpha rev	CTG TGT CGG GGT TGT AGC CA TGC CCC AGG ACA CAG AGA CT	58	290	WTA-QC
EGFR Exon18 for EGFR Exon18 rev	TTGTCCTTCCAAATGAGCTG TGCCTTTGGTCTGTGAATTG	58	496	WGA-QC bulk
EGFR Exon19 for EGFR Exon19 rev	TCCTCGATGTGAGTTTCTGC ATGCCTCCATTTCTTCATCC	58	350	WGA-QC bulk
EGFR Exon20 for EGFR Exon20 rev	AAACGTCCCTGTGCTAGGTC CATGGCAAACTCTTGCTATCC	58	442	WGA-QC bulk
EGFR Exon21 for EGFR Exon21 rev	CAGCGGGTTACATCTTCTTTC AAACAATACAGCTAGTGGGAAGG	58	418	WGA-QC bulk
GAPDH for GAPDH rev	CCA TCT TCC AGG AGC GAG AT CAG TGG GGA CAC GGA AGG	58	489	WTA-QC
KRAS for KRAS rev	ATA AGG CCT GCT GAA AAT GAC CTG AAT TAG CTG TAT CGT CAA GG	58	91	WGA-QC + WGA-QC bulk
KRT19 for KRT19 rev	GAA GAT CCG CGA CTG GTA C TTC ATG CTC AGC TGT GAC TG	58	621	WGA-QC
Lib1	AGTGGGATTCCTGCTGTCAGT	65	-	WGA
NRAS Exon2 for NRAS Exon2 rev	ACACCCCCAGGATTCTTACA TCCGCAAATGACTTGCTATT	58	174	WGA-QC bulk

PIK3CA Exon20 for PIK3CA Exon20 rev	TCTAGCTATTCGACAGCATGC TTGTGTGGAAGATCCAATCCAT	58	221	WGA-QC bulk
Survivin for Survivin rev	CGACCCCATAGAGGAACATAAA GGAATAAACCCTGGAAGTGGTG	59	276	Endpoint PCR
TP53 Exon2/3 for TP53 Exon2/3 rev	GAA GCG TCT CAT GCT GGA TC CAG CCC AAC CCT TGT CCT TA	58	301	WGA-QC
TP53 Exon5/6 for TP53 Exon5/6 rev	ACGCATGTTTGTTTCTTTGC ACCCCTCCTCCCAGAGAC	58	507	WGA-QC bulk
TP53 Exon7 for TP53 Exon7 rev	GAGGCTGAGGAAGGAGAATG AGTATGGAAGAAATCGGTAAGAGG	58	400	WGA-QC bulk
TP53 Exon8 for TP53 Exon8 rev	AGGTAGGACCTGATTTCCTTACTG AGGCATAACTGCACCCTTG	58	245	WGA-QC bulk
TRP2 for TRP2 rev	CCAGCTGGGAAACTGTCTGT AACCCTTCCAAAGCATTCCT	58	223	Melanoma marker PCR
TYR for TYR rev	AGGTCAGCACCCCACAAATC TTTATGCAATGGAACGCCCG	58	122	Melanoma marker PCR

* Used in combination with BRAF Exon15_3‘ primer. T_A: Annealing temperature.

2.3.6 Commercial kits

Table 7: List of commercial kits

Name	Manufacturer	Catalogue number
Ampli1™ LowPass Kit (SET A+Set B) 2 x 48 reactions	Menarini Silicon Biosystems	WGLPAB
AP Conjugate Substrate Kit	BioRad Laboratories GmbH	1706432
Bioanalyzer High Sensitivity DNA Kit	Agilent	5067-4626
DNeasy Blood & Tissue Kit	QIAGEN	69581
Expand Long Template PCR System	Roche Diagnostics	11759060001 (Sigma A.)
FastStart™ Taq DNA Polymerase, dNTPack	Roche Diagnostics	4738420001 (Sigma A.)
GenePrint® 10 System	Promega	B9510
innuPREP RNA Mini Kit 2.0	Analytik Jena	845-KS-2040010
KAPA Library Quantification Kit	Roche Diagnostics	07960298001
Macherey-Nagel™ NucleoSpin™ Gel and PCR Clean-up Kit	Thermo Fisher Scientific	11992242
MiSeq Reagent Nano Kit v2 (300-cycles)	Illumina	MS-103-1001
mMessage mMachine T7 ULTRA Transcription Kit	Life Technologies	AMB1345-5
OptiView DAB IHC Detection Kit	Roche Diagnostics	760-700
QIAamp® DNA FFPE Tissue Kit	QIAGEN	56404
QIAGEN Plasmid Mini Kit	QIAGEN	12123
QIAGEN Plasmid Midi Kit	QIAGEN	12145
QIAGEN Plasmid Maxi Kit	QIAGEN	12162
Qubit® dsDNA HS Assay Kit	Thermo Fisher Scientific	Q32854
RNase-free DNase Set	QIAGEN	79254
RNeasy Mini Kit	QIAGEN	74104
RQ1 RNase-Free DNase (for cDNA)	Promega	M6101
UltraView Universal Alkaline Phosphatase Red Detection Kit	Roche Diagnostics	760-501

2.4 Cell culture media

Table 8: List of cell culture media

Name	Components
Standard cell culture medium	500 ml RPMI 50 ml FCS 5 ml 200 mM L-Glutamine 5 ml Penicillin/ Streptomycin

T cell medium	43.95 ml RPMI 5 ml FCS 0.5 ml Penicillin/ Streptomycin 0.5 ml 1 M Hepes 50 µl 55 mM β-Mercaptoethanol
Co-culture medium	100 ml AIM V 87 ml RPMI 10 ml human AB serum 1 ml Penicillin/ Streptomycin 1 ml 1 M Hepes 1 ml 200 mM L-Glutamine 9.1 µl 55 mM β-Mercaptoethanol

2.5 Consumables

Table 9: List of used consumables

Name	Manufacturer	Catalogue number
6 Well Cell Culture Plate	Greiner Bio-One	657160
12 Well Cell Culture Plate	Greiner Bio-One	657180
96 Well Clear Round Bottom TC-treated Microplate	Corning	3799
96 Well F Tissue Culture Test Plate	TPP	92096
Diagnostic slides 3 x 14 mm	Th. Geyer GmbH & Co KG	9161187
Adhesive clear PCR seal	Biozym	600208
Adhesive sealing sheets	Thermo Fisher Scientific	AB0558
Axygen™ PCR Tubes 0.2 ml (MAXYMum Recovery)	Fisher Scientific	11370145
Biopur® Eppendorf Safe-Lock Tubes™ sterile, 0.5 ml	Eppendorf	0030 121.570
Biopur® Eppendorf Safe-Lock Tubes™ sterile, 1.5 ml	Eppendorf	0030 121.589
BD Luer-Lok™ Syringe 50 ml	BD Plastipak™	300865
Cell culture dish	TPP	93100
Cell culture flask T175	Greiner Bio-One	660175
Cell culture flask T25	Greiner Bio-One	690160
Cell culture flask T75	Greiner Bio-One	658175
Cell strainer (40 µm)	Becton Dickinson GmbH	352340
Cellstar® serological pipette 10 ml	Greiner Bio-One	607180
Cellstar® serological pipette 2 ml	Greiner Bio-One	710180
Cellstar® serological pipette 25 ml	Greiner Bio-One	760180
Cellstar® serological pipette 5 ml	Greiner Bio-One	606180
Centrifuge tube 15 ml	Greiner Bio-One	188271
Centrifuge tube 50 ml	Greiner Bio-One	227261
Combitips advanced 0.1 ml	Eppendorf	0030089618
Combitips advanced 0.5 ml	Eppendorf	0030089634
Combitips advanced 1 ml	Eppendorf	0030089642
Cotton swab	VWR	115-8270
Cryo Tube™ vial	Thermo Fisher Scientific	377267
K2EDTA vacutainer	BD Bioscience	367525
Electropotations cuvettes (4 mm)	Biolab products	75-EP-104
Erlenmeyer flask 250 ml DURAN	Schott	21 216 36
Erlenmeyer flask 250 ml DURAN	Schott	21 216 44
Insulin syringe (Microfine, 29 G, U-100)	BD Biosciences	324824
Micro-hematocrit capillary, non-heparinized length 75 mm x 1.1/1.2 mm	Brand	749321
Microscope slides, ground edges 90°, frosted end	ROTH	H870.1
microTUBE-50 AFA Fiber Screw-Cap	Covaris	PN520166
Multichannel reservoir, 10 ml, sterile	Integra	4331
Nunc™ Lab-Tek™ Chamber Slides; 8 fields	Thermo Fisher Scientific	11367764
PARAFILM® M 4 in x 125 ft	ROTH	H666.1
Pasteur pipettes, long	Brand	747720
PCR SingleCap 8er-Soft Strips 0.2 ml, clear	Biozym	710970

PCR tube 0.2 ml, single tube	4titude Deutschland	4ti-0795
PCR tube 1.5 ml, graduated, non-sterile	Greiner Bio-One	616201
PCR tube 2 ml, graduated, non-sterile	Greiner Bio-One	623201
Petri dish, PS, 100/20 MM, advanced TC, sterile	Greiner Bio-One	664960
Protein LoBind Tube 0.5 ml	Eppendorf	022431064
Protein LoBind Tube 1.5 ml	Eppendorf	022431081
Protein LoBind Tube 2 ml	Eppendorf	022431102
Protein LoBind Tube 5 ml	Eppendorf	0030108302
QIAshredder	QIAGEN	79654
Reagent reservoirs 10 ml	Integra	4331
RNAseZAP™	Sigma Aldrich	R2020-250ML
Rotilabo® syringe filter, PVDF, sterile, 22 µm	ROTH	P666.1
Rotilabo® syringe filter, PVDF, sterile, 45 µm	ROTH	P667.1
SafeSeal Surphob 1250 µl (filter)	Biozym	VT0270
SafeSeal Surphob 20 µl (filter)	Biozym	VT0220
SafeSeal Surphob 200 µl (filter)	Biozym	VT0240
SafeSeal Surphob 10 µl (filter)	Biozym	VT0200
Surgical Disposable Scalpels	Braun	5518091
Terasaki plate, 60 well, PS, 83.3/58 mm, transparent	Greiner Bio-One	653180
Tips LTS 200 µL Filter RT-L200F	RAININ	17002927
Tissue cassettes	A. Hartenstein	EBK
Transparent 96-well PCR plate	Biozym	710884
Tube 5 ml, 75x12 mm, PP (FACS tube)	Sarstedt	55.1578

2.6 Devices

Table 10: List of devices

Name	Manufacturer
Autoclave 3150 EL	Systec
BD FACSAria	BD Bioscience
BD FACSCelesta™	BD Bioscience
BD LSR II	BD Bioscience
Bioanalyzer 2100	Agilent Technologies
CellTram Pump	Eppendorf
Centrifuge 5424R	Eppendorf
Centrifuge 5810R	Eppendorf
Centrifuge Plate Fuge	Benchmark Scientific
Centrifuge Rotina 380R	Hettich
Cryo 1°C Freezing Container	Nalgene
DMZ Universal Puller	Zeitz
DNA Engine Tetrad2 Peltier Thermal Cycler	Bio Rad
DynaMag™-15 Magnet	Thermo Fisher Scientific
Echo® 525 Liquid Handler	Labcyte
Electrophoresis chamber 40-1214	Peqlab
Gene Pulser Xcell™	Bio Rad
GeneTouch Thermal Cycler, Block 96 or Block 48/48	Biozym
HiSeq	Illumina
Incubator Heraeus BB15	Thermo Fisher Scientific
Incuyte® Live-Cell Analysis System	Essen BioScience
Laminar flow bench Her Safe KS18	Thermo Fisher Scientific
NanoDrop 2000c	Thermo Fisher Scientific
Neubauer Counting Chamber	Marienfeld
Manual pipettes (2µl, 10µl, 20µl, 200µl, 1000µl)	Gilson
Medimachine with Medicons	DAKO
Microscope Axiovert 200M (fluorescence)	Zeiss
Microscope IB inverted (cell culture)	Optech
Microscope IX81 (fluorescence)	Olympus

Microvawe	Micromaxx
MiSeq	Illumina
Multipette Stream	Eppendorf
NovaSeq	Illumina
Patchman NP2 micromanipulator	Eppendorf
PCR bench UVT-S-AR	Thermo Fisher Scientific
pH-meter PB-11	Sartorius
P-H Waterbath WB 20	P-D Industriegesellschaft mbH
Pipet-Lite™ XLS+ manual 8-channel pipette, 20-200 µL	RAININ
Power Supply MP-250N	Kisker Biotech
Qubit® 3 fluorometer	Invitrogen/ Thermo Fisher Scientific
Research Pro 8-channel pipette	Eppendorf
Roller Mixer SRT1	Stuart Scientific
Rotary Microtome HM 340E	Thermo Scientific
Scale AVW220D	Shimadzu
Scale PLS 510-3	Kern
UV Illuminator	Intas
Ventana BenchMark Ultra	Roche
Ventana HE 600	Roche
Vortexer	VELP Scientifica

2.7 Softwares and databases

Table 11: List of softwares and databases

Name	Provider/ URL	Application
2100 Expert Software Version B.02.09SIO725 (SR1)	Agilent Technologies	Bioanalyzer
cellSens Dimension 1.9	Olympus	Microscope
Docker Images	Menarini Silicon Biosystems	CNA data
EndNote X7	https://www.endnote.de/	Citation
GraphPad Prism v6.1	https://www.graphpad.com/scientificsoftware/prism/	Plotting of data
FlowJo V10	https://www.flowjo.com/solutions/flowjo/	Flow cytometry
IncuCyte ZOOM® 2018A	Essen BioScience	Killing assay
Microsoft Office 2016	Microsoft	Data management
PubMed	https://www.ncbi.nlm.nih.gov/pubmed	Literature search

3 Methods

3.1 Collaborations

3.1.1 External collaborations

Table 12: List of collaborations

Name	Institution at time of collaboration start	Project part
Prof. Ehud Shapiro; Ofir Raz	Weizmann Institute of Science, Rehovot	Microsatellite sequencing
Dr. Stefan Kirsch	Fraunhofer ITEM-R, Regensburg	Whole exome sequencing
Dr. Jens Warfsmann; Dr. Martin Hoffmann	Fraunhofer ITEM-R, Regensburg	CNA analysis
Prof. Dr. med. Bernd Spriewald; Dr. Christian Bach	Laboratory for Immunogenetics, University Hospital Erlangen	Classical HLA typing
PD Dr. med. Sebastian Haferkamp	Department of Dermatoloy, University Hospital Regensburg	Sample acquisition
Dr. med. Philipp Renner	Department of Surgery, University Hospital Regensburg	Sample acquisition
Dr. med. Florian Weber	Department of Pathology, University Hospital Regensburg	Pathological evaluation
Prof. Dr. med. Philipp Beckhove; Dr. Slava Stamova	RCI Regensburg Center for Interventional Immunology	Functional assays
PD Dr. med. Simone Thomas; Dr. med. Florian Lücke	Clinic and Polyclinic for Internal Medicine III, University Hospital Regensburg	Functional assays
Prof. Dr. Oliver Kohlbacher	Center for Bioinformatics, University of Tübingen	HLA typing

3.1.2 Internal collaborations

Sample collection and preparation for microsatellite sequencing (chapter 4.2.1.1), CNA analysis (chapter 4.1.4) and *BRAF* mutation analysis (chapter 4.2.3) were performed partially as a team within different working groups of the Department of Experimental Medicine and Therapy Research. Parts of the data-set from patient MM15-127 and MM16-423 were also analysed in the thesis of Julia Greindl-Junghans, but from a different perspective. While Julia Greindl-Junghans was interested in the progression of disease based on the cell morphology and mutation status, the focus of this project was the order of metastatic seeding and the suitability of SLN-DCCs as surrogates for systemic cancer cells. Due to the large size of samples and the accompanying enormous effort, the amount of work was shared. Successful establishment of the microsatellite sequencing workflow (chapter 4.2.1.1 and 4.2.1.2) was achieved together with Julia Greindl-Junghans and Manjusha Ghosh as it was strived to a have a large sample collective for optimal protocol establishment. The used samples came from three different projects, comprising three melanoma patients from this thesis and the project of Julia Greindl-Junghans and one breast cancer patient from the project of Manjusha Ghosh.

LP-Seq and CNA analysis were guided and supported by Dr. Zbigniew Czyz, whereas bioinformatic analysis of WES data was performed by Dr. Huiqin Koerkel-Qu.

3.2 Sample acquisition

Patient samples were received between 2013 and 2017 in cooperation with Dr. med. Philipp Renner and PD Dr. med. Sebastian Haferkamp. Primary tumors were located with resident pathologists and therefore sent to Regensburg for analysis. A detailed overview of the samples used in this project is provided in the results in chapter 4.1.1.1.

3.3 Preparation of slides for single cell isolation

3.3.1 Lymph nodes

From lymph nodes single cell suspensions were prepared as previously described (Ulmer et al. 2005, Ulmer et al. 2014). Briefly, after cutting the lymphatic tissue into 1 mm pieces with a sterile scalpel, the tissue pieces were mechanically disaggregated into a cell suspension with the help of the Medimachine using sterile rotating knives. Cell suspension was washed with HBSS. After a centrifugation step (20 min, 1000g, RT) on a density gradient consisting of a 60% Percoll solution mononuclear cells were isolated from the interphase. After washing with PBS (300g, 10 min, RT), cells were resuspended in PBS and counted using a Neubauer counting chamber. Then, cells were diluted to a cell density of 10^6 cells/ml in PBS and transferred onto diagnostic slides at a density of 0.5x 10^6 cells/slide (0.17x 10^6 cells/spot). After one hour, when cells have sedimented, remaining PBS was discarded. On the next day, once diagnostic slides were air-dried (RT), they were stored at -20°C.

3.3.2 Metastasis

After putting the metastatic tissue in a petri dish with basal medium it was cut into small pieces using surgical scalpels. Next, enzymatic digestion was performed with collagenase, hyaluronidase and DNaseI at a final concentration of 0.33 mg/ml, 100 µg/ml and 100 µg/ml, respectively, for 20-40 minutes at 37°C to obtain a single cell suspension. Subsequently, samples were pipetted up and down and put into a 50 ml tube. After washing with PBS (300g, 10 min, RT), cell pellet was resuspended in pre-warmed (37°C) Trypsin 0.05%/EDTA 0.2% in PBS. After five minutes (RT) 5 ml of Basal medium supplemented with 10% FCS was added to stop digestion. Then, filtering the cells using a 40 µm cell strainer they were washed with 14 ml PBS (300g, 10 min, RT). After resuspending the cell pellet in PBS (10^6 cells/ml), the cell suspension was diluted and put onto diagnostic slides at a density of 0.5x 10^6 cells/slide (0.17x 10^6 cells/spot). When cells have sedimented on the slide after one hour at RT, PBS was discarded and stored at -20°C once air-dried overnight at RT.

3.3.3 Buccal swabs

For isolating oral epithelial cells (OECs) a sterile cotton swab was rubbed on the patient's buccal mucosa. After stirring the cotton swab in 1 ml PBS, the resulting cell solution was put onto diagnostic slides. After sedimentation of cells (1 h, RT), residual PBS was removed. Diagnostic slides were air-dried (overnight, RT) and then stored at -20°C.

3.3.4 Blood samples

Peripheral blood (50 ml) was collected in EDTA vacutainers. For one patient 168 ml of diagnostic leukapheresis product were received in a blood bag from the Department of Transfusion Medicine (University Hospital Regensburg). PBMCs from EDTA vacutainers were initially washed with Hank's solution (200g, 10 min, 4°C). Resulting cell pellet as well as leukapheresis product were diluted in Hank's solution followed by centrifugation (1000g, 20 min, 4°C) on a 60% Percoll density gradient. After collecting the interphase containing PBMCs, cell suspension was filled up with PBS and washed (500g, 10 min, 4°C). Supernatant was discarded and cells were resuspended in PBS and counted. Only PBMCs isolated from leukapheresis product were transferred onto diagnostic slides at a density of 0.5x 10^6 cells/slide (0.17x 10^6 cells/spot). After sedimentation of cells (1 h, RT), remaining PBS was discarded and slides were stored at -20°C on the next day, when air-dried overnight at RT. PBMCs from EDTA vacutainers and remaining PBMCs from leukapheresis product were cryo-conserved in 90% FCS and 10% DMSO in liquid nitrogen.

3.3.5 Cell lines / xenografts

Diagnostic slides were prepared from frozen cell line cells or xenograft cells. For this, they were thawed in standard cell culture medium. After washing with PBS, cells were counted and diluted to a density of 10^6 cells/ml in PBS. Cells were transferred onto diagnostic slides at a density of 0.5x 10^6 cells/slide (0.17x 10^6 cells/spot). Following sedimentation of the cells after one hour at RT, PBS was discarded and slides were stored on the next day at -20°C after having air-dried overnight at RT.

3.4 Identification of cells of interest for single cell isolation

3.4.1 Staining for tumor cells

3.4.1.1 Staining for tumor cells on diagnostic slides (gp100)

Tumor cells were isolated from lymph nodes, metastatic tissue, patient-derived cell lines or blood samples from previously prepared diagnostic slides. For this, diagnostic slides were thawed for 30 minutes at RT followed by adding PBS/10% AB serum for 30 minutes at RT to block unspecific binding. For tumor cell staining a primary antibody against gp100 (anti-human melanosome, clone HMB45, 0.7 µg/ml) was used and IgG1κ as isotype control for 45 minutes. After washing with PBS (3 times, 3 min), AP-polymer solution was used for secondary staining for 30 minutes at RT. Following another washing step (3 times, 3 min), AP-polymer solution was developed using AP Conjugate Substrate Kit with BCIP-NBT detection system for 12 minutes yielding a blue reaction product. Slides were finally washed with PBS for three times and analysed under the microscope.

3.4.1.2 Staining for tumor cells in cell suspension (MCSP)

For the isolation of living DCCs, three million lymph node cells were stained with an anti-human MCSP antibody (clone 9.2.27; 25 µg/ml). Before adding the primary antibody, unspecific binding was blocked by incubating cells in 95 µl blocking solution (AB serum 10%/ peptone 2% in PBS) at 4°C on a shaking rotator. After 10 minutes, primary antibody was added and incubated for another 10 minutes at 4°C on a shaking rotator. Following centrifugation (1500 rpm, 5 min, 4°C) supernatant was discarded and cell pellet was washed with 500 µl PBS/peptone (2%). After another centrifugation step (1500 rpm, 5 min, 4°C) and discarding the supernatant, cell pellet was resuspended in 98 µl blocking solution and 2 µl of the secondary antibody goat anti-mouse-Cy3 (15 µg/ml) was added and incubated at 4°C on a shaking rotator. After 10 minutes, cells were washed with 500 µl PBS/peptone (2%) and then with PBS.

3.4.2 Staining for control cells

Staining for control cells from diagnostic slides (T cells, macrophages, endothelial cells)

First, diagnostic slides (from lymph nodes and metastatic tissue) were thawed for 30 minutes at RT, PBS/10% AB serum was added for 30 minutes at RT to block unspecific binding. For isolating CD3-positive T cells and CD68-positive macrophages a double-staining was performed, whereas CD31-positive endothelial cells were stained individually. Slides were stained with anti-CD3 (0.1 mg/ml)/anti-CD68 (1:500) or anti-CD31 (1:250) antibodies for 60 minutes at RT. As isotype control for CD31 and CD68 IgG1κ and for CD3 Rabbit IgG was used. Before secondary staining slides were washed with PBS (3 times, 3 min). CD3/CD68 primary stained cells were stained with goat anti-rabbit IgG-Alexa Fluor 555 and goat anti-mouse IgG-Alexa Fluor 488 (4 µg/ml each) for 30 minutes at RT. For CD31, AP-polymer anti-mouse solution was used for secondary staining for 30 minutes at RT. Following secondary staining, slides were washed with PBS (3 times, 3 min). AP-polymer solution for detecting CD31-positive cells was developed with the BCIP-NBT detection system using AP Conjugate Substrate Kit. Following 12 minutes incubation, slides were washed with PBS (3 times, 3 min).

Staining for control cells from frozen PBMCs (T cells, B cells)

For isolating single cells from frozen PBMCs, cryo-conserved PBMCs were thawed in standard cell culture medium and washed with PBS. After counting, PBS/10% AB serum was added for 5 minutes at 4°C to block unspecific antibody binding. To identify single B and T cells single cell suspension was stained with anti-CD19 PE- and anti-CD3 FITC-conjugated antibody for 15 minutes at 4°C. After washing with FACS buffer, CD19- and CD3-positive cells were separated with FACSAria cell sorter. To enable discrimination of live and dead cells viability dye eFlour 780 was used. Sorted populations were used for isolating single cells.

3.4.3 Identification of oral epithelial cells

Oral epithelial cells from buccal swabs were isolated from diagnostic slides, which were therefore thawed for 30 minutes at RT. Identification of OECs was based on morphologic criteria, diameter of 60-80 µm with a clearly smaller nucleus having an irregular shape.

3.5 Manual isolation of single cells

Single cells were either isolated from diagnostic slides or from cell suspension using an inverted fluorescence microscope (Olympus or Zeiss) equipped with a micromanipulator and a pump. First, a picking field of an 8-chamber slide was prepared. After wettening with BSA, 200 µl of PBS was filled in the picking field. Stained cell suspensions having a maximum density of 10^6 cells per 200 µl PBS were transferred on a free BSA-wettened field of a chamber slide. If using diagnostic slides, 200 µl of PBS was put on the spots of the slide. Positively stained cells or cells of interest were identified, sucked in the FCS-coated glass capillary, which was attached to the micromanipulator, and transferred into the picking field. By using a small pipette the cell was transferred in 1 µl into a 0.2 ml Axygen™ PCR tube containing 2 µl of proteinase K digestion mix, if whole genome amplification was performed (chapter 3.12). For whole transcriptome amplification of the isolated cell, it was transferred in a 0.2 ml Axygen™ PCR tube with 4 µl mTRAP lysis buffer and 0.4 µl *E. coli* tRNA (chapter 3.13), which was stored at -80°C. In addition to single cells, also a pool of cells was isolated by transferring 1 µl of cell suspension to the prepared tubes. In the end, 1 µl of cell-free PBS from the picking field was taken and served as a negative control for subsequent amplification. In addition to single cell isolation for WGA or WTA, cells for xenotransplantation were also isolated manually.

3.6 Generation of patient-derived xenograft

Xenotransplantation

For xenograft generation, spheres or cells of interest were transplanted into 6-8 weeks old NOD.Cg-*Prkdc*^{scid} *IL2rγ*$^{tmWjl/Sz}$ (NSG) mice purchased from the Jackson Laboratory. Cells or spheres were collected from disaggregated lymph nodes or metastatic tissue with the help of a micromanipulator and transferred into a microwell of a terasaki plate in a total volume of 10 µl. To prevent cell adhesion to the wells, microwells were pre-coated over night with 12 mg/ml poly-HEMA at RT. Immunodeficient mice were injected intraperitoneal with the fully antagonisable anaesthetic consisting of midazolam (5 mg/kg), fentanyl (0.05 mg/kg), medetomidine (0.5 mg/kg) and placed on a heating plate (37°C) until the anesthetic set in. For mice weighing 25 grams injection volume of anesthetic was 160 µl. Preventing drying out, eye ointment was put on the eyes of anesthetized mice using a sterile swab. In the middle of the mice's back hair were removed with a razor on an area of 1.5 cm x 1 cm. The shaved skin area was desinfected with iodine using sterile swabs. The iodine was removed with a sterile swab and Kodan spray. Then, cells were injected with an insulin syringe subcutaneously in a final volume of 50 µl and 30% high-concentration matrigel. After transplantation, mice were injected subcutaneously with flumazenil (0.5 mg/kg), atipamezole (2.5 mg/kg), naloxone (1.2 mg/kg) having an injection volume half

of anesthetic volume. Mice were maintained in the research animal facilities of the University of Regensburg. They were kept under specific-pathogen-free conditions using acidified water and food ad libitum. Approved experimental animal procedures were transacted by German federal and state regulations (Government of Upper Palatinate or Lower Franconia). Every week transplanted mice were palpated at the injection site.

Xenograft removal and tumor digest

Once the tumor reached a maximum diameter of around 1.5 cm, it was removed. After sacrifying, the mouse was covered with ethanol. Using sterile knives and tweezers, the skin on the back was cut generously around the tumor and pulled over the head. The tumor was removed and put into a petri dish for weighing and size measurement. Then, a small piece of tumor was cut, put into a tissue cassette and transferred in formaldehyde (< 5%) for fixation, while the rest of the tumor was digested as follows. The tumor tissue was transferred into a petri dish and covered with digest medium. After cutting the tumor into small pieces with sterile scalpels, enzymatic digestion was done with collagenase and DNaseI at a final concentration of 0.33 mg/ml and 100 µg/ml, respectively. Following incubation for 20-40 minutes at 37°C, samples were pipetted up and down to obtain single cell suspension and transferred into a 50 ml tube. Then, cells were washed with PBS (300g, 10 min, RT) and resuspended in pre-warmed (37°C) Trypsin 0.05%/ EDTA 0.2% in PBS. Then, 5 ml of digest medium supplemented with 10% FCS was added after five minutes to stop digestion. After filtering through a 40 µm cell strainer, cells were washed with 4 ml PBS (300g, 5 min, RT). In the end, the cell pellet was resuspended in 500 µl PBS. After counting, cells were either used for surface marker analysis using flow cytometry or frozen in 70% FCS/ 20% RPMI/ 10% DMSO. Formalin-fixed tissue was given to the Department of Pathology for preparing FFPE blocks, later used for cutting sections for marker staining.

3.7 Analysis of primary tumor material

Isolation of genomic DNA from primary tumors

Genomic DNA was isolated from primary tumor FFPE blocks using QIAGEN supplementary Protocol "Purification of genomic DNA from FFPE tissue using the QIAamp® DNA FFPE Tissue Kit and Deparaffinization Solution" according to the manufacturer's instructions. For this, either sections or punches from the FFPE tumor block were used. Afterwards, DNA concentration was measured using NanoDrop device and success of gDNA isolation was visualized by agarose gel electrophoresis.

Staining of primary tumor sections

For analysis of primary tumors, 2-3 µm sections were cut from the FFPE block with the help of a rotary microtome and put onto microscope slides. Tumor sections were then stained for hematoxylin and eosin (HE) or CD molecules (surface markers characteristic for specific immune cell populations) using full-automated staining systems. HE staining was performed with Ventana HE 600. For immunophenotypic characterization, the alkaline phosphatase detection system was used leading to a red end product. For this, CD molecule stainings were performed with the UltraView Universal Alkaline Phosphatase Red Detection Kit according to the manufacturer's instructions using Ventana BenchMark Ultra. Antibodies against CD45 (1:800), CD3 (1:200) and CD68 (1:200) were used.

3.8 HE and melanoma marker staining of metastatic cells and xenograft sections

For HE and melanoma staining of single cells, frozen diagnostic slides with single cell suspensions were thawed. For HE staining, slides were put into demineralized water for a short time before performing nuclear staining. For this, slides were put in hematoxylin solution for 10 minutes and again shortly washed with demineralized water. Afterwards, slides were transferred in tap water for 30 minutes to

complete nuclear stain by 'blueing'. Before eosin counterstain, slides were put into demineralized water for a short time. Coloring non-nuclear elements slides were transferred in 0.5% eosin solution including 5 drops of acetic acid and incubated for 10 minutes. Then, slides were shortly washed again in demineralized water and slide was rinsed in several baths of alcohol in order to remove all traces of water. For this, slides were put for each 1 minute in 70% and 80% ethanol and 1-2 minutes in 100% ethanol. After bathing in ethanol, the slides were transferred for 1 minute in xylol. In the end, a thin layer of Eukitt® quick-hardening mounting medium was used to cover up the slides with a glass cover slip. In order to stain single cells on diagnostic slides for additional melanoma markers, Melan-A staining was performed. For this, as primary antibody mouse anti-human Melan-A (4.4 µg/ml) was used to detect melanoma antigen recognized by T cells-1 (MART1). The staining procedure was done in the same way than gp100 staining (chapter 3.4.1.1) using IgG1κ as isotype control. In addition to single cell staining from metastatic tissue, sections from metastasis-derived xenograft FFPE blocks were stained for several markers. For this, antibodies against S100 (1:2400), MiTF (1:100), desmin (1:100), actin (1:500) and cytokeratin (1:200) were used. For detecting gp100, MART-1 and tyrosinase, a pan melanoma cocktail (1:400) was taken. Stainings were performed with the OptiView DAB IHC Detection Kit using full-automated staining system Ventana BenchMark Ultra.

3.9 Isolation of genomic DNA from cell suspensions

Genomic DNA from a maximum of 5x 10^6 PBMCs, lymph node cells, metastasis cells or cell line cells was isolated using DNeasy Blood & Tissue Kit following the manufacturer's protocol. DNA was eluted in 80-100 µl of H_2O and concentration was measured with NanoDrop.

3.10 Isolation of RNA

RNA was isolated from 1-3x 10^6 cell line cells or expanded T cells using RNeasy Mini kit according to the manufacturer's instructions. For isolating RNA from PBMCs QIAGEN Supplementary Protocol "Isolation of Peripheral Blood Mononuclear Cells (PBMC) and Purification of Total RNA from PBMC Using the RNeasy® Micro or Mini Kit" was conducted starting from the section "Total RNA purification". RNA was eluted in 30 µl RNase-free water. RNA isolation from dendritic cells war performed with innuPREP RNA Mini Kit 2.0 according to the manufacturer's instructions.

3.11 cDNA synthesis

For cDNA synthesis 1 µg of RNA was filled up with H_2O to a total volume of 8 µl. Adding 1 µl of DNAase buffer and 1 µl of DNase, samples were incubated for 30 minutes at 37°C. Then, 1 µl of DNase stop solution was added and incubated for 10 minutes at 65°C. Next, 1 µl of Oligo d(T)s were transferred to the samples. After 5 minutes at 70°C, samples were put on ice for 5 minutes. A mix consisting of 5 µl M-MLV buffer, 1.25 µl dNTPs, 1 µl of M-MLV Reverse Transcriptase and 5.75 µl of H_2O was prepared and added to the samples resulting in 25 µl end volume. cDNA was synthesized keeping the samples for 1 hour at 39°C and 5 minutes at 70°C.

3.12 Whole genome amplification and quality control

3.12.1 Whole genome amplification and quality control of single cells

3.12.1.1 Whole genome amplification of single cells

After single cell isolation, whole genome amplification (WGA) was performed as previously described (Klein et al. 1999, Klein et al. 2002a). Briefly, protease digestion was performed at the beginning removing residual proteins and exposing the DNA by cell lysis. This was followed by MseI digestion of the cell's genome cutting the DNA at -*TTAA*- nucleotide motifs thereby creating fragments of 150 -

1500 bp in length. Next, a double-stranded adaptor oligonucleotide was constructed by pre-annealing Lib1 and ddMse11 primers and ligated to the DNA fragments. After removing the non-ligated adapter strand an overhang of the Lib1 adaptor was created. In the end, MseI fragments were amplified in a so-called primary PCR by using Lib1 as primer creating consistent amplification of the entire genome. Whole genome amplification is also commercially available as *Ampli1™* WGA kit.

Detailed protocol

The composition of the different reactions mixes are described in Table 13. After single cell isolation, the cell was transferred in 1 µl into a 0.2 ml Axygen™ PCR tube containing 2 µl of the proteinase K digestion mix. The sample was then incubated in a thermal cycler (42°C, 10 h). Afterwards, enzyme inactivation was carried out at 80°C for 10 minutes and the samples were cooled to 4°C. Following proteinase K digestion, 2 µl of MseI digestion mix was added to the sample incubating for 3 hours at 37°C. In the meantime, the pre-annealing mix was prepared and a double-stranded adapter complex was formed by decreasing the temperature from 65°C down to 15°C with 1°C per minute. After 3 hours of MseI digestion, the enzymatic reaction was inactivated at 65°C for 5 minutes. Adding 1 µl ATP (10 mM) and 1 µl T4 ligase (5 U/µl) to the pre-annealing mix, the resulting 5 µl of pre-annealing/ligation mix were transferred to the MseI-digested sample and incubated at 15°C overnight. After adaptor ligation, 40 µl of primary PCR mix were added and the primary PCR program was started (Table 14).

Table 13: Mastermix compositions for WGA

Reaction mix		Amount for one reaction [µl]
Proteinase K digestion mix:	OPA buffer	0.2
	Tween 10%	0.13
	Igepal 10%	0.13
	Proteinase K (10 mg/ml)	0.26
	PCR-H_2O	1.27
MseI digestion mix:	OPA buffer	0.2
	Mse I (50.000 U/µl)	0.2
	PCR-H_2O	1.6
Pre-annealing mix:	OPA buffer	0.5
	Lib1 100 µM	0.5
	ddMse11 100 µM	0.5
	PCR-H_2O	1.5
Primary PCR mix:	Expand Long Template Buffer 1	3
	dNTPs (10 mM)	2
	Expand Long Template Pol Mix	1
	PCR-H_2O	34

Table 14: Cycler program for primary PCR of WGA

Step	Temperature	Time	Cycles
1)	68°C	3 min	
2)	94°C	40 s	
3)	57°C	30 s	15
4)	68°C	1 min 30 s + 1 s/cycle	
5)	94°C	40 s	
6)	57°C + 1°C/cycle	30 s	9
7)	68°C	1 min 45 s + 1 s/cycle	
8)	94°C	40 s	
9)	65°C	30 s	23
10)	68°C	1 min 53 s + 1s/cycle	
11)	68°C	3 min 40 s	
12)	4°C	forever	

3.12.1.2 Quality control of single-cell WGA products

Subsequently, to control the quality of the resulting products a dedicated endpoint PCR (WGA-QC) assay was done (Polzer et al. 2014). With a four marker multiplex PCR, the presence of *KRAS, KRT19*, and *TP53 Exon2/3*, as well as a polymorphic DNA area on chromosome 5 using the *D5S2117* primers were analysed. For this, 9 µl of mastermix (Table 15) were mixed with 1 µl of primary WGA sample in a PCR tube. Controlling the purity of all reagents a negative and a positive control were prepared before running the PCR program (Table 15) using a thermal cycler. In the end, amplified products were analysed by agarose gel electrophoresis (chapter 3.18).

Table 15: Composition of mastermix and cycler program for WGA-QC

Reagent	Amount per reaction (µl)
10x FastStart PCR Buffer (20 mM $MgCl_2$)	1.0
Primer mix (8 µM per primer)	1.0
dNTPs (10 mM)	0.2
BSA (20 mg/ml)	0.2
FastStart Taq Polymerase (5 U/µl)	0.1
PCR-H_2O	6.5

Step	Temperature	Time	Cycles
1)	95°C	4 min	
2)	95°C	30 s	
3)	58°C	30 s	32
4)	72°C	90 s	
5)	72°C	7 min	
6)	4°C	forever	

3.12.2 Whole genome amplification and quality control of bulk DNA from FFPE samples

3.12.2.1 Whole genome amplification of bulk DNA from FFPE samples

In order to amplify bulk DNA WGA for single cells (chapter 3.12.1.1) was slightly modified. After digesting the remaining proteins using proteinase K for 15 hours an optional DNArepair step using NEBNext® FFPE DNA Repair Mix was done. Here, 1 µl consisting of 0.4 µl repair buffer, 0.1 µl repair mix and 0.5 µl H_2O were added and incubated for 20 minutes at 37°C. After inactivation (20 min at 80°C) MseI digestion followed as well as ligation and primary PCR as mentioned before. For MseI 0.6 µl of water was used instead of 1.6 µl. Controlling the quality of amplified DNA a particular WGA-QC was performed specifically designed for WGA samples generated from FFPE samples.

3.12.2.2 Quality control of bulk WGA products

For assessing the quality of WGA products generated from FFPE samples three multiplex PCRs were designed (WGA-QC bulk). Multiplex #1 addressed five regions, multiplex #2 analysed the presence of four genes and multiplex #3 of three genes (Table 16). Before preparing the mastermixes (Table 17), four primer sets consisting of 8 µM per forward and reverse primer were made facilitating the procedure. Then, 1 µl of template DNA was added to each 9 µl of mastermix for multiplex PCR #1-3. Including a positive and negative control, the same PCR program as for WGA-QC for single cells was used (Table 15). Success of PCRs were analysed by agarose gel electrophoresis.

Table 16: Overview multiplex PCRs for WGA-QC bulk

Multiplex	Primer	Primer set
Multiplex #1 (5-plex)	KRAS	A
	BRAF Exon15	B
	PIK3CA Exon20	A
	EGFR Exon19	A
	TP53 Exon	B
Multiplex #2 (4-plex)	NRAS Exon2	C
	TP53 Exon8	C
	EGFR Exon21	C
	EGFR Exon18	C
Multiplex #3 (3-plex)	TP53 Exon5/6	D
	EGFR Exon20	D
	CKND2A Exon3	D

Table 17: Composition of mastermix for WGA-QC bulk

	Amount per reaction (µl)		
Reagent	**Multiplex #1**	**Multiplex #2**	**Multiplex #3**
10x FastStart PCR Buffer (20 mM $MgCl_2$)	1.0	1.0	1.0
Primer set A (8 µM per primer)	3.0		
Primer set B (8 µM per primer)	4.0		
Primer set C (8 µM per primer)		4.0	
Primer set D (8 µM per primer)			3.0
dNTPs (10 mM)	0.2	0.2	0.2
BSA (20 mg/ml)	0.2	0.2	0.2
FastStart Taq Polymerase (5 U/µl)	0.1	0.1	0.1
PCR-H_2O	0.5	3.5	4.5

3.13 Whole transcriptome amplification and quality control

Whole transcriptome amplification (WTA) describes a multistep process designed to separate mRNA from the gDNA of one single cell before amplification of the transcriptome and genome (Klein et al. 2002b, Hartmann and Klein 2006). The method was made commercially available as a kit, *Ampli1™* WTA kit. Cell lysates (single cell in mTRAP lysis buffer and *E.coli* tRNA) were thawed and WTA was performed as previouly described (Klein et al. 2002b). After primary WTA, success of amplification was analysed by multiplex endpoint PCR (WTA-QC) for three transcripts (ACTB, EEF1A1, GAPDH). For this, 9.5 µl of the mastermix (Table 18) were transferred in a 0.2 ml PCR tube and 0.5 µl of primary WTA product were added. The PCR was performed in PCR cyclers using the program of WGA-QC (Table 15) including a negative and positive control. Amplified cDNA was visualized by gel electrophoresis.

Table 18: Composition of mastermix for WTA-QC

Reagent	**Amount per reaction (µl)**
10x FastStart PCR Buffer (20 mM $MgCl_2$)	1.0
Primer mix (8 µM per primer)	1.0
dNTPs (10 mM)	0.2
BSA (20 mg/ml)	0.2
FastStart Taq Polymerase (5 U/µl)	0.1
PCR-H_2O	7

3.14 Fingerprint analysis

Fingerprint analysis was performed with 20 ng/µl of bulk DNA or 1 µl of WGA product. GenePrint® 10 System was used for patient authentication considering 10 STR loci.

3.15 Melanoma marker PCR

For detecting melanoma markers on single cell level by PCR, standard endpoint PCRs for tyrosinase (TYR), an enzyme in melanogenesis, and for Tyrosinase-related protein 2 (TRP2) were performed using cDNA from WTA samples as template. For each reaction 9 µl of mastermix (Table 19) were prepared and transferred in a 0.2 ml PCR reaction tube. Then, 1 µl of template DNA were added and PCR was carried out in a thermal cycler using the corresponding PCR program (Table 19). As positive control, DNA from melanoma cell line MelHo (MEL-HO, ACC-62, Leibniz Institute DSMZ) was used and PCR-H_2O served as negative control.

Table 19: Composition of mastermix and cycler program for melanoma marker PCR

Reagent	Amount per reaction (µl)
10x FastStart PCR Buffer (20 mM $MgCl_2$)	1.0
Primer for (8 µM per primer)	0.5
Primer rev (8 µM per primer)	0.5
dNTPs (10 mM)	0.2
BSA (20 mg/ml)	0.2
FastStart Taq Polymerase (5 U/µl)	0.1
PCR-H_2O	6.5

Step	Temperature	Time	Cycles
1)	94°C	2 min	
2)	58°C	30 s	
3)	72°C	2 min	
4)	94°C	15 s	
5)	58°C	30 s	15
6)	72°C	20 s	
7)	94°C	15 s	
8)	58°C	30 s	25
9)	72°C	30 s	
10)	72°C	2 min	
11)	4°C	forever	

3.16 Allele-specific PCR with a blocking reagent for *BRAF* hotspot mutation

In order to analyse *BRAF* hotspot mutation BRAF V600E (c.1799T>A; c.1798GT>AA) located on *BRAF* exon 15 an allele-specific PCR with a blocking reagent (ASB-PCR) was designed (Morlan et al. 2009). A multiplex endpoint PCR was performed for detecting *BRAF* mutation in single-cell WGA products combined with amplification of allele-independent *BRAF* gene. The method was established together with Julia Greindl-Junghans and was therefore also partly described in her thesis. By using mutation independent primers for *BRAF* exon 15, the presence of the *BRAF* gene was proved resulting in a 171 bp amplicon after amplification (Table 20). The mutation-specific primer (BRAF V600E mut) was designed with a complementary base to the hotspot BRAF mutation c.1799T>A at its 3' end. If the mutation was present, a 147 bp amplicon was created during amplification by combining with BRAF Exon15_3' primer. To suppress non-specific amplification of the wild type allele by the mutant-specific primer a blocking oligonucleotide (BRAF V600E block) was designed, which was fully complementary to the wild type sequence of the *BRAF* gene. The base complementary to the mutation site was located in the middle of the oligonucleotide having a 3' phosphate group to prevent elongation during PCR. The reaction was performed by adding 1 µl of template DNA to 9 µl of mastermix (Table 21) in 0.2 ml PCR tubes. Including a positive and negative control, amplification was carried out in a thermal cycler using a touch-down PCR program (Table 21) in order to increase specificity and selectivity. As positive control, DNA from melanoma cell line MelHo was used carrying a heterozygous BRAF c.1799T>A mutation. In addition, a negative control was included. Success of ASB-PCR was analysed by electrophoresis using 3% agarose gels.

Table 20: Primer composition of ASB-PCR for BRAF V600E

Primer combination	Amplicon length
BRAF V600E mut + BRAF Exon15_ 3'	147 bp
BRAF Exon15_ 5'+ BRAF Exon15_ 3'	171 bp

Table 21: Composition of mastermix and cycler program for ASB-PCR for BRAF V600E

Reagent	Amount per reaction (µl)
10x FastStart PCR Buffer (20 mM $MgCl_2$)	1.0
BRAF V600E block (32 µM)	0.5
BRAF V600E mut (8 µM)	0.5
BRAF Exon15_ 5' (8 µM)	0.5
BRAF Exon15_ 3' (8 µM)	1.0
dNTPs (10 mM)	0.2
BSA (20 mg/ml)	0.2
FastStart Taq Polymerase (5 U/µl)	0.1
PCR-H_2O	5.0

Step	Temperature	Time	Cycles
1)	95°C	4 min	
2)	95°C	30 s	
3)	68 - 58°C	30 s	10
4)	72°C	90 s	
5)	95°C	30 s	
6)	58°C	30 s	29
7)	72°C	90 s	
8)	72°C	7 min	
9)	4°C	forever	

3.17 Endpoint PCR on cDNA

Endpoint PCR was performed on cDNA to check for expression of Survivin in dendritic cells. For this, 44 µl of mastermix were prepared and transferred in a 0.2 ml PCR tube (Table 22). Then, 1 µl of template DNA consisting of 40 ng of cDNA were added and PCR was carried out in a thermal cycler using corresponding PCR program (Table 22). Success of endpoint PCR was analysed by electrophoresis using 3% agarose gels. Enzymatic digestion of the PCR endproduct was performed to check for amplification of the correct sequence. For this, 17 µl of PCR product were digested using 1 µl of enzyme NcoI and 2 µl of cutsmart buffer in a total volume of 20 µl at 37°C. After one hour, success of digestions was analysed by electrophoresis using 3% agarose gels. NcoI was cutting the 276 bp Survivin amplicon in a way creating 207 bp and 69 bp long amplicons.

Table 22: Composition of mastermix and cycler program for endpoint PCR on cDNA

Reagent	Amount per reaction (µl)
10x FastStart PCR Buffer (20 mM $MgCl_2$)	4.5
Primer for (10 µM)	2.25
Primer rev (10 µM)	2.25
dNTPs (10 mM)	0.9
BSA (20 mg/ml)	0.9
FastStart Taq Polymerase (5 U/µl)	0.4
PCR-H_2O	32.8

Step	Temperature	Time	Cycles
1)	94°C	2 min	
2)	94°C	30 s	
3)	59°C	30 s	32
4)	72°C	60 s	
5)	72°C	7 min	
6)	4°C	forever	

3.18 Agarose gel electrophoresis

Success of PCRs and fragment size distribution of DNA were analysed by agarose gel electrophoresis. Depending on analysed samples 1.5-3% agarose gels were prepared. Corresponding amount of agarose was dissolved in 100 ml 1x TBE buffer in an Erlenmeyer flask by heating in a microwave. After adding 4 µl of 10 mg/ml ethidium bromide solution and shaking, the liquid gel was transferred into a gel tray with combs for the pockets. Once polymerized (30 min, RT) the gels were ready to use. DNA samples were mixed with 3 µl of loading dye and put onto the gels. DNA fragments were size-separated at 4.57 V/cm (160 V/ 35 cm) for 45-60 minutes. Bands were visualized on a UV transilluminator and size-evaluated using 8 µl of 1 kb DNA ladder or low molecular weight DNA ladder as reference.

3.19 Re-amplification of WGA products and quality control

To re-amplify the primary WGA product, different protocols were used. Standard re-amplification (REA) prior to library preparation for LowPass-sequencing was performed at 68°C (REA68) annealing temperature in a total volume of 50 µl (Table 23). For preparing DNA for microsatellite marker sequencing, a re-amplification mixture was used (REA65/68). For this, two independent REAs were performed, one at 65°C, one at 68°C using half of standard input amount of all components. REAs were carried out in 0.2 ml tubes or 96-well PCR plates in thermal cycler using corresponding PCR program (Table 23). For controlling the success of the re-amplification another WGA-QC was performed.

Table 23: Composition of mastermix and cycler program for re-amplification

Reagent	Amount per reaction (µl)
Expand Long Template Buffer 1	5.0
Lib1 (10 µM)	5.0
dNTPs (10 mM)	1.75
BSA (20 mg/ml)	1.25
Pol. Mix	0.5
PCR-H_2O	35.5
Template DNA	1

Step	Temperature	Time	Cycles
1)	94°C	1 min	
2)	60°C	30 s	
3)	65 or 68 °C	2 min	
4)	94°C	30 s	
5)	60°C	30 s	10
6)	65 or 68°C	2 min + 20 s/cycle	
7)	4°C	forever	

3.20 Double-strand synthesis of WGA products

Double-strand DNA was reconstituted by adding 2 µl of mastermix (Table 24) to 10 µl of original WGA product and incubating the sample for 2 hours at 68°C followed by keeping at 12°C. Reconstitution of double-strand DNA was performed for initial sample preparation tests for microsatellite marker sequencing and library preparation for whole exome sequencing.

Table 24: Composition of mastermix and cycler program for double-strand synthesis

Reagent	Amount per reaction (µl)
Expand Long Template Buffer 2	0.2
dNTPs (10 mM)	0.2
Lib1 (100 µM)	0.2
Pol. Mix	0.1
PCR-H_2O	1.3

Step	Temperature	Time
1)	68°C	2 h
2)	12°C	forever

3.21 DNA purification

To clean up amplified DNA samples remaining oligonucleotides and primers were removed. For this, DNA samples were purified using AMPure XP beads with bead to DNA ratio of 1.8x following the protocol of the manufacturer.

3.22 Measurement of DNA concentration and DNA quality

Qubit and NanoDrop concentration measurement

Concentration of sequencing libraries was measured with Qubit® 3 fluorometer using Qubit® dsDNA HS Assay Kit according to the manufacturer's instructions. For this, a working solution was prepared consisting of 199 µl of Qubit® dsDNA HS buffer and 1 µl of HS dye. Afterwards, 2 µl of DNA sample were mixed with 198 µl of working solution. Prior to measurement two standards were prepared by mixing 190 µl of working solution with 10 µl of standard 1 and 2. Following incubation (2 min, RT), the standards were measured in order to calibrate the fluorometer. Then, the concentration of the samples was determined. Due to the specificity of the fluorescent dye, the result concentration reflected the amount of double-stranded DNA, whereas concentration of RNA and single stranded DNA was not assessed. Nucleic acid concentration of genomic DNA from cells or primary tumor blocks was determined using NanoDrop 2000c spectrophotometer. For this, 1 µl of the DNA sample was taken. Before measuring the samples, the device was calibrated by using 1 µl of water as 'blank'. Pure DNA was determined by values of 1.8-2.0 for A260/280 and A260/230 larger than 2.0. In contrast to Qubit® 3 fluorometer, the spectrophotometer measures the absorbance of all types of nucleic acids.

Bioanalyzer analysis

The DNA fragment size distribution and quality of DNA sequencing libraries were visualized using the Bioanalyzer device. The analysis was performed using Bioanalyzer High Sensitivity (HS) DNA Kit and corresponding reagents following the manufacturer's instructions with an input amount of 1 µl of the samples (at least 1 ng/µl). For the analysis 300 bp was set as lower border and 3000 bp as upper border.

3.23 Microsatellite sequencing

3.23.1 Sample preparation for microsatellite sequencing

At the beginning of the project several protocols and protocol combinations were tested for microsatellite sequencing (MS-Seq), like adding a step of double-strand synthesis (DSS) after WGA of original or re-amplified WGA products combined with or without DNA purification. In the end, all samples were re-amplified and subjected to DNA purification (chapter 4.2.1.1) as follows:

Single cells

Every single-cell WGA product was re-amplified in two technical replicates. One of the two re-amplification reaction was conducted with the temperature of the elongation step set to 65°C, while the other reaction was run with a corresponding setting set to 68°C. Afterwards, both technical replicates were mixed in 1:1 ratio. Eliminating residual oligonucleotides and reagents, pooled samples were purified with AMPure XP beads with 1.8x bead to DNA ratio and eluted in PCR-water.

Germline samples

Germline samples were prepared from bulk gDNA from lymph node cells or PBMCs (Table 25). On the one hand, 200 ng of unamplified bulk gDNA was used as template for library preparation. On the other hand, 1 ng and 10 ng of gDNA were subjected to WGA using five technical replicates. WGA products were subjected to REA65/68 like single cell WGA products and purified using AMPure XP beads.

Table 25: Overview of germline samples for MS-Seq

	MM15-127	MM16-423	MM16-412
Source	Bulk from SLN	PBMCs (leukapheresis)	PBMCs (EDTA blood)
Unamplified	1x 200 ng	2x 200 ng	2x 200 ng
WGA	5x 1 ng and 5x 10 ng for WGA → REA + XP		

SLN: sentinel lymph node; WGA: Whole-genome amplification; REA: Re-amplification mixture of once using an annealing temperature of 65° and once at 68°; XP: DNA purification using AMPureXP beads.

Bulk DNA from primary tumors

For library preparation bulk DNA isolated from primary tumors was either sent unamplified or was subjected to WGA (Table 26). 600-1300 ng of unamplified DNA was left untreated or was repaired using NEBNext® FFPE DNA Repair Mix. Regarding amplified samples, WGA was performed on 10-50 ng of bulk DNA with or without repair mix using three or four technical replicates. WGA samples from primary tumors were also subjected to REA65/68 as already described followed by DNA purification.

Table 26: Overview of primary tumor samples for MS-Seq

	MM15-127 PT1-3	MM16-412 PT
Unamplified	1x 1000 ng	2x 1000 ng + 2x 1300 ng + repair
WGA	4x 50 ng +/- repair → REA + XP	3x 10 ng +/- repair → +/- REA + XP

PT: primary tumor; WGA: Whole-genome amplification; REA: Re-amplification mixture of once using an annealing temperature of 65° and once at 68°; XP: DNA purification using AMPureXP beads.

3.23.2 Target enrichment and sequencing

Enrichment of microsatellites with molecular inversion probes and sequencing of the samples was performed at the Weizmann Institute of Science in Israel as previously described (Tao et al. 2018).

3.24 Ampli1™ LowPass sequencing

For analysing copy number alterations (CNA) LowPass sequencing (LP-Seq) was performed as previously described (Buson et al. 2016, Ferrarini et al. 2017, Ferrarini et al. 2018). By using *Ampli*1™ LowPass Kits barcoded sequencing libraries were produced for genome-wide copy-number profiling.

Library preparation for LP-Seq

DNA from high-quality WGA products were subjected to re-amplification (REA68, chapter 3.19) compatible for preparing sequencing libraries. First, 5 µl of re-amplified WGA product were purifed with 1.8x AMPure XP Beads according to manufacturer's instructions. After eluting in 22 µl of nuclease free water, 20 µl were transferred into a new tube avoiding bead aspiration. The libraries for LP-Seq were prepared using *Ampli1*™ LowPass kit for Illumina® platforms. Starting from 3 µl of purified re-amplified WGA product, barcoded libraries compatible with Illumina® systems were prepared according to the manufacturer's instructions using TruSeq HT-compatible dual-indexes.

Library quantification and pooling

Before pooling, libraries were quantified by combining Qubit concentration measurement and Bioanalyzer fragment size analysis using 1 µl each (chapter 3.22). Based on this, libraries were pooled in equimolar concentrations obtaining 4-10 nM pools. In order to reduce costs and time, library quantification was alternatively done by quantifying an unbalanced pool consisting of 0.5 µl of each library. Unbalanced library pools were purified using AMpure XP beads and eluted in 20 µl TE buffer low EDTA. Purified unbalanced pools were quantified using Qubit and Bioanaylzer analysis and sequenced using MiSeq Reagent Nano Kit v2 (300 cycles) on a MiSeq device at Fraunhofer ITEM-R. Following quality control (chapter 3.26.2), new equimolar pools were generated based on calculations from MiSeq normalization. Pooling of libraries was done manually or with Echo® 525 Liquid Handler. For MiSeq run 32 samples were pooled, whereas for HiSeq or NovaSeq runs a pool contained 96 samples. Before sequencing, equimolar pools were purified using KAPA Pure Beads (1.8x) eluted in 38 µl TE buffer low EDTA and quantified using Qubit, Bioanaylzer and KAPA Library Quantification Kit for Illumina® Platforms according to manufacturer's instructions.

LP-Seq of equimolar library pools

Equimolar pools were sequenced in 150 bp single end mode using a MiSeq device at Fraunhofer ITEM-R or in 100 bp single end mode by CeGaT GmbH (Tübingen) on HiSeq2500 or NovaSeq6000 platforms using a custom read primer provided by *Ampli1*™ LowPass kit.

3.25 Whole exome sequencing

Library preparation for whole exome sequencing (WES) using Agilent SureSelect^{XT} Target Enrichment System for Illumina paired-end sequencing was done together with Dr. Stefan Kirsch (Fraunhofer ITEM-Regensburg). For this, WGA products were subjected to double-strand synthesis (chapter 3.20) and sent to the Fraunhofer Institute where libraries were made based on the Institute's protocol. In brief, this comprised two parts, sample preparation and enrichment including additional purification steps and quality controls. After quantification, libraries were pooled and sent for paired-end sequencing (2x 150 bp) using Illumina NovaSeq platform provided by Genewiz.

3.26 Bioinformatics

3.26.1 Microsatellite marker sequencing

Bioinformatic analysis of MS-Seq and lineage tree reconstruction was performed by bioinformaticians at the Department of Computer Science and Applied Mathematics from the Weizmann Institute of Science, Israel. In the end, FastTree2 algorithm was applied for final tree reconstructions (Price et al. 2010). For evaluating the lineage tree robustness, bootstrapping analysis was performed qith 1000 iterations using a threshold of transfer bootstrapping expectation (TBE) index larger than 70% (Lemoine et al. 2018).

3.26.2 *Ampli1*™ LowPass sequencing data - CNA analysis using Docker Images

Quality control (QC) of raw data

Having received sequencing raw data (fastq files) from the sequencing service provider, a quality control was performed. For this, an established QC-workflow was used. Briefly, reads were identified clearly mapping to the human genome. Those reads were then stored in separate fastq files which were later used for producing CNA profiles. Remaining reads, which did not map to the human genome, were not being assessed for generating CNA profiles.

Processing of QC-data using Docker Images

After QC, the newly created fastq files containing only human sequences were processed using Docker Images, a modified version of the open source software Control-FREEC (Control-Free Copy number caller). The results of the analysis included graphical and tabular summaries of identified copy number alterations. Scatter plots visualized CNA profiles in which genomic gains were represented in red and genomic losses in blue. A low level of noise was represented by a derivative log ratio spread (DLRS) smaller than 0.35 used for quality evaluation of the profiles.

3.26.3 WES data - HLA analysis

In silico HLA typing from WES data was performed together with Dr. Koerkel-Qu using different typing tools originally tested on sequencing data of bulk samples (Bauer et al. 2018). At first, raw data from WES were checked for their quality using fastqc and multiqc. Afterwards, tile filtering and trimming with BBduk according to the quality of fastq were performed. For HLA genotyping the alignment-based OptiType tool, provided by Prof. Kohlbacher from Tübingen (Szolek et al. 2014), and the assembly-based HLAminer tool were conducted.

3.26.4 WES data - mutation detection

For mutation detection fastq files from WES were used. After trimming files to remove adaptor and low quality of reads, they were mapped to NCBI GRCh38. Next, duplicated reads were marked and removed, before base quality scores were recalculated. Output bam files were then used as input for mutation calling. This was analysed by applying two different algorithms, Varscan and Platypus. Varscan2 (version 2.4.4) algorithm was applied with min var freq set to be 0.35 for single cells, 0.05 for pool of single cells and 0.01 for bulk samples and coverage set to at least 8 for normal samples and 6 for tumor samples, whereas Platypus (version 0.1.5) was applied with minimal mutation call of reads set to be 4. To eliminate false positive mutations, mutations were called when they appeared in at least two non-pool tumor samples. Tumor samples being a pool of single cells were excluded from mutation calling, as more bioinformatic expertise was neccessary to discriminate true-positive from false-positive mutations within those samples. Checking for the presence of ultraviolet (UV) signatures of the tumor samples, mutational signatures from cosmic database were considered focussing on UV-associated signature 7 (*https://cancer.sanger.ac.uk/cosmic/signatures/SBS/index.tt*). For mutational overlap between tumor samples only exonic mutations changing the protein sequence were considered, whereas for UV signature analysis all mutations were used. Somatic mutations for neopeptide prediction were annotated with ANNOVAR (version 20191024) and HLA-I binding affinity of neopeptides created by detected mutations was predicted with NetMHCpan 4.0 (Jurtz et al. 2017) with given HLA type. Class I neopeptides with %Rank smaller than 0.5 were defined as strong binders, whereas weak binders showed %Rank smaller than 2.

3.27 Classical HLA typing

For classical HLA typing 3 µg of bulk DNA having a concentration of 100 ng/µl were sent to the Laboratory for Immunogenetics of the University Hospital Erlangen for analysis.

3.28 Cell culture

3.28.1 Standard cell culture

For thawing frozen cells, the cryo vial was put into a waterbath (37°C) for around 2 minutes. If only a small piece of ice was visible, the cryo vial was wiped with 70% ethanol and the cell suspension was transferred into a new 50 ml tube. Then, 10 ml of pre-warmed standard cell culture medium (Table 8)

was added dropwise. If primary cells were thawed, 100 µl DNase I (10 mg/ml) were added to the 10 ml of pre-warmed medium beforehand. After centrifuging (300g, 5 min), the supernatant containing cell-toxic DMSO was removed and cells were resuspended in corresponding culture medium. When thawing PBMCs or lymph node cells, they were counted using a Neubauer counting chamber. Therefore, 10 µl of cells suspension were diluted 1:10 in Trypan blue solution for discriminating dead cells. Cells were resuspended in culture medium or T cell medium having a density of 10x 10^6 cell/ml and seeded in 200 µl per well of a 96-U-bottom plate to let them rest overnight. On the next day, cells were harvested, washed with PBS and counted. Regarding MelDTC cancer cell line cells, 1x 10^6 cells were thawed and put in a T75 flask in 13 ml of cell culture medium. After several days, when reaching a confluency of 80-90%, cells were transferred into a T175 flask. For this, old medium was removed and cells were washed with 1x PBS. Then, 1 ml 0.05% trypsin/ ETDA solution was added and incubated at 37 °C. After 3 minutes, trypsin was inactivated by adding 10 ml of fresh pre-warmed medium. Cell suspension was transferred into a 50 ml tube and centrifuged for 5 minutes at 300g. Discarding the supernatant, the cells were resuspended in 25 ml of cell culture medium and put into a T175 flask. Cells were passaged having a confluency of around 90%. After washing with PBS, cells were trypsinized using 2 ml 0.05% trypsin/ ETDA solution as already described. After centrifugation, cells were resuspended in 5 ml of fresh medium and counted. Depending on the experiments 1-2x 10^6 cells were resuspended in 25 ml of fresh medium and seeded into a new T175 flask. All cells were cultured at 37°C and 5% CO_2. In order to freeze cells in cryo vials, 1-10x 10^6 cells were frozen in 70% FCS/ 20% RPMI/ 10% DMSO when freezing cell line cells, whereas primary cells were frozen in 90% FCS/10% DMSO. Cryo vials were put into a Cryo 1°C Freezing Container and stored at -80°C overnight. Cells were either kept at -80°C or transferred to liquid nitrogen tanks.

3.28.2 *In vitro* stimulation of T cells

For *in vitro* stimulation of T cells, soluble purified antibodies against human CD28 were used. First, antibodies were diluted to a concentration of 2 µg/ml in PBS. Then, 200 µl of 2 µg/ml CD28 solution were transferred to a well of a 96-well round bottom plate. Sealing the plate with parafilm, it was incubated over night at 4°C. On the next day, the solution was sucked out with a pasteur pipette and 200 µl of PBS were added to the wells to wash away unbound antibodies. After sucking out the PBS again with a pasteur pipette, the CD28-coated wells were ready to be seeded with desired cells at a density of 1x 10^5 cells per well.

3.28.3 *In vitro* expansion of tumor-reactive T cells

In vitro expansion of tumor-reactive T cells, was performed by co-culturing blood cells together with autologous DCC-derived cell line cells as previously described (Dijkstra et al. 2018). In brief, co-cultures of DCC-derived cell line cells and blood cells were performed in anti-CD28-coated round bottom plates (2 µg/ml) in the presence of IL-2 (150 U/ml) in a 20:1 effector:target ratio. Half of the co-culture medium was exchanged every 2 or 3 days. Every week, T cells were harvested, counted and re-stimulated with fresh DCC cell line cells. Every two weeks, cells were analysed by flow cytometry for production of effector molecules (Interferon-γ, CD107a) or the expression of antigens that have been published to be expressed on tumor-specific T cells (CD137, PD-1, CD103, CD39 (Dijkstra et al. 2018, Cafri et al. 2019, Duhen et al. 2018, Gros et al. 2016)). In total, the co-culture was performed over four weeks. However, before DCC-derived cell lines were co-cultured with blood cells, they were stimulated with Interferon-γ (IFN-γ; 100 U/ml; 48 h) to induce upregulation and optimal presentation of peptide-presenting HLA complexes. As this might result in an increased expression of immune inhibitory checkpoint ligands possibly interfering with the *in vitro* expansion of T cells, the necessity of adding blocking antibody anti-PD-1 (5 µg/ml) to the co-culture was tested.

3.28.4 Immune cell-mediated killing assay

In order to analyse immune cell cytotoxicity *in vitro*, immune cell-mediated killing of tumor cells was performed. For this, target tumor cells were co-cultured with autologous immune cells in the presence of a DNA intercalating dye enabling quantification of apoptosis. In every experiment frozen expanded or naiive immune cells were used. One day before the assay, tumor cells were seeded in a 96-well flat plate at 20000 cells/well (50 µl/well). Tumor cells were cultured in co-culture medium overnight to let them attach reaching a cell density of around 80-90% after 24 hours. On the next day, effector cells, thawed the day before (chapter 3.28.1), were counted and added to the tumor cells (25 µl/well) in effector:target ratio of 5:1. In the end, Incucyte® Caspase-3/7 Green Apoptosis Assay Reagent (25 µl/well) was added achieving a final concentration of 1:1000. Several controls were used in the assay. Effector and target cells were cultured alone, but in prescence of the dye. Additionally, the co-cultures were cultured without the dye. The killing assay was performed in triplicates for different treatment options, whereas duplicates were used for controls. Cytotoxicity was measured over 24 hours using Incuyte® Live-Cell Analysis System. Automated image analysis with the IncuCyte ZOOM® 2018A Software enabled selective quantification of tumor cell death. For this, four images taken by the system were used for the analysis and evaluated for their Total Green Object Area (µm²/well).

3.28.5 Generation of dendritic cells from frozen PBMCs

Fast-DC protocol

PBMCs were thawed as described in chapter 3.28.1 with 10 ml of pre-warmed standard cell culture medium and 100 µl DNase I (10 mg/ml). The tube was filled up with pre-warmed standard cell culture medium to 40 ml. After centrifugation (1400 rpm, 10 min), the supernatant was discarded and the cell pellet was resuspended in 5 ml AIM V medium/ 1% human serum. After counting, the cell suspension was diluted to a concentration of 0.5x 10^7/ml. Cells were incubated over night (37 °C, 5% CO_2). For this, the tube was put into a diagonal position in the incubator and the lid of the tube was loosend enabling gas exchange. On the next day, the settled cells were resuspended and the cell suspension was filled up with PBS to 40 ml. Following centrifugation (1400 rpm, 10 min), the supernatant was discarded and the cell pellet was resuspended in AIM V medium/ 1% human serum and counted. 15-20x 10^6 cells were seeded per well of a 6-well plate in total 3 ml per well. Cells were incubated 1.5 hours (37 °C, 5% CO_2) enabling adherence. Then, the 6-well plate was swiveld carefully to merge the white streak from non-adherent cells. Medium was removed and non-adherent cells were washed off in two washing steps with 2 ml pre-warmed PBS. PBS was rinsed over the cells with a gentle stream and washing progress was checked under the microscope. After washing and removing the supernatant, 3 ml AIM V medium/ 1% human serum were added as well as differentiation cytokines GM-CSF (1000 U/ml) and IL-4 (500 U/ml). Cells were incubated 24 hours (37°C, 5% CO_2). Then, maturation cytokines IL-1β (10 ng/ml), TNF-α (10 ng/ml), IL-6 (1000 U/ml) and PGE2 (1 µg/ml) were added and cells were incubated for 24 hours (37°C, 5% CO_2). Mature dendritic cells (DCs) were harvested 3 days after thawing. For this, 6-well plates were put on ice and cells were detached using cold PBS. The supernatant was collected together with the cells and centrifuged (1400 rpm, 10 min). As cells were used for electroporation, they were resuspended in Opti-MEM medium prior to counting. The success of DC generation was analysed by flow cytometry using antibodies against CD3, CD19, CD209, HLA-II and CD80/86.

One-week DC protocol

PBMCs were thawed and rested over night as already described in the Fast-DC protocol. Cells were then resuspended in 10 ml X-Vivo 20 medium, transferred into a culture dish and incubated for 2 hours (37°C, 5% CO_2). After incubation, supernatant was taken with a 10 ml pipette and discarded. 10 ml DC medium (X-Vivo 20 with 560 U/ml GM-CSF an 500 U/ml IL-4) were added to the attached cells and

incubated (37°C, 5% CO_2). After two days, 2 ml of fresh DC medium were added. On the next day, again 3 ml of fresh DC medium were added. Three days later i.e. 7 days after thawing, DCs were harvested for culturing cytokine-free. For this, DCs were flushed from the dish and transferred into a 50 ml tube. Dish was rinsed with 5 ml X-Vivo 20 medium and collected in the tube. After centrifugation (1400 rpm, 10 min), supernatant was discarded. The cell pellet was then resuspended in 4 ml X-Vivo 20 medium (without cytokines) and transferred back into the cell culture dish. The tube was washed with 4 ml X-Vivo medium and cell suspension was added to the dish and incubated for another day. On the next day, 100 µl (per 1x 10^7 DCs) of Dynabeads Pan mouse IgG were transferred into a 15 ml tube and 20 µl of CD56 antibody were added per 100 µl beads before incubating the tube for 20 minutes at 4°C under rotation for later DC purification. DCs were rinsed from the culture dish and transferred into a 15 ml tube washing the dish with 5 ml PBS. After centrifugation (1400 rpm, 10 min), DCs were resuspended in 300 µl MACS buffer. Each 100 µl of CD3 and CD19 Dynabeads were added to antibody-coupeld beads (100 µl/ 1x 10^7 cells). 5 ml of PBS were added to the beads and put onto the magnet. After 1 minute, PBS was removed with a pasteur pipette without touching the beads. Removing the tube from the magnet, beads were again washed with 5 ml PBS and put onto the magnet for 1 min. PBS was removed and beads were resuspended in 300 µl of MACS buffer (per 1x 10^7 cells). Then, 300 µl of bead mixture was added to the DCs and incubated for 20 minutes at 4°C unter rotation. Afterwards, 4 ml MACS buffer were added and mixed with the cell suspension. The tube was transferred to the magent for 2 min. For negative selection, the supernatant was removed and transferred into a new tube. Washing with MACS buffer was repeated and thesupernatant was again transferred into the tube. After centrifugation (1400 rpm, 10 min, 4°C), the cell pellet was resuspended in 1 ml X-Vivo 20 medium and counted using 1:10 Trypan blue solution. In the end, flow cytometric analysis was performed to check for success of DC generation using antibodies against CD3, CD19, CD209, HLA-II and CD80/86.

3.28.6 Electroporation of dendritic cells

Before electroporation, 4 ml AIM V medium including 10% human serum were put into a well of a 6-well plate and incubated at 37°C and 5% CO_2. Dendritic cells (DCs) were counted and 5-10x 10^6 cells were centrifuged (10 min, 1000 rpm). Medium was poured off and cell pellet was washed with 10 ml Opti-MEM medium. Washing and centrifugation step with 10 ml Opti-MEM medium was repeated for two times. Afterwards, the medium was poured off and the pellet was remaining liquid was removed. Cells were resuspended in 190 µl Opti-MEM while electroporator (Gene Pulser Xcell™) was turned on. Before collecting the RNA, the work space was cleaned with RNaseZAP. RNA was put on ice before transferring 5-10 µg RNA into the electroporation cuvette. Then, the cell suspension was added to the cuvette and knocked onto the bench to mix RNA and cell suspension. The cuvette was put into the electroporator and electroporation was performed using the following setting: 3 square waves, 5 ms, interval 0, 400 V and 1 pulse. Afterwards, the cuvette was taken out and cells were transferred carefully into a well of a 6-well plate already filled with 4 ml pre-warmed AIM V medium including 10% human serum. Electroporated cells were incubated at 37°C and 5% CO_2 for at least 10 hours.

3.28.7 Dendritic cell / T cell co-culture for minigene test

For testing the minigene and checking for minigene-specific T cell responses half of DCs were electroporated with minigene-encoding RNA, whereas remaining cells were left unelectroporated as negative control. Both, electroporated and unelectroporated DCs were incubated over night in 4 ml AIM V including 10% human serum per 6 well. On the day of electroporation, peptide-specific T cells were thawed and incubated in 10 ml RPMI inlcuding 5% human serum in a 50 ml tube in diagonal position at 37°C and 5% CO_2 over night. On the next day, cells were harvested from 6-well plates putting 30 min on ice. After washing with PBS, cells were counted. 50000 to 100000 peptide-specific T cells, kindly provided by Dr. Slava Stamova, were co-cultured with DCs in 1:1 ratio in well of a 96-well

round bottom plate and incubated for 4 hours in 200 µl RPMI including 5% human serum in presence of eBioscience™ Protein Transport Inhibitor. To check for reactive T cells, flow cytometry staining was performed using antibodies against CD3, CD8a, CD45 and intracellular IFN-γ.

3.29 Construction of tandem DNA minigene

For each test peptide a mingene was generated, consisting of DNA sequences coding for 25 amino acids (AA) with the sequence of test peptides located in the middle of the each Minigene (Table 27).

Table 27: Amino acid sequences of test peptides for minigene construction

Test peptide	AA sequence	25 AA sequence for minigene
MART-1 (26-35, modified)	ELAGIGILTV	GHSYTTAE**ELAGIGILTV**ILGVLLL
Survivin (95-104)	ELTLGEFLKL	LSVKKQFE**ELTLGEFLKL**DRERAKN
Influenza A MP (58-66)	GILGFVFTL	PILSPLTK**GILGFVFTL**TVPSERGL

AA: amino acid

The DNA sequence of each minigene was codon optimized for homo sapiens (Table 28). For this, gene synthesis tool of Invitrogen GeneArt was used *(https://www.thermofisher.com/de/de/home/life-science/cloning/gene-synthesis/geneart-gene-synthesis.html)*.

Table 28: Codon optimized DNA sequences of test minigenes

Test peptide	DNA sequence
MART-1 (26-35, modified)	GGCCACAGCTATACCACAGCCGAAGAACTGGCCGGC ATCGGCATCCTGACAGTGATTCTGGGAGTCCTGCTGCTG
Survivin (95-104)	CTGAGCGTGAAGAAGCAGTTCGAGGAACTGACCCTGGGC GAGTTCCTGAAGCTGGACAGAGAGAGAGCCAAGAAT
Influenza A MP (58-66)	CCCATTCTGAGCCCTCTGACCAAGGGCATCCTGGG CTTCGTGTTCACACTGACCGTGCCTAGCGAGAGAGGACTG

To generate a tandem minigene consisting of three minigenes, they were connected with Linker_1 (AA-sequence of Linker_1: GGSGGGGSGG). At the beginning of the tandem mingene sequence an additional Linker_1 was inserted, whereas after the last minigene sequence an end linker, Linker_2 was added (AA-sequence of Linker_2: GGSLGGGGSG). Prior to the start linker a cloning site for enzyme BamHI was inserted and after the end linker a cutting site for XhoI was added. A Kozak-sequence including a start codon, as well as sequences for MHC-I signal and trafficking peptides (SP and MITD) were additionally used (AA-sequence for SP: MRVTAPRTLILLLSGALALTETWAGS; AA-sequence for MITD: IVGIVAGLAVLAVVVIGAVVATVMCRRKSSGGKGGSYSQAASSDSAQGSDVSLTA (Sahin et al. 2017)). After the end linker a His-tag and a stop codon were inserted. As the cloning vector had cutting sites for XbaI and EcoRI, sequences for those enzymes were put at the edges of the tandem minigene. The tandem minigene looked as follows: - **XbaI**- Kozak- SP- **BamHI**- *Linker1*- **MART-1**- *Linker1*- **Survivin**- *Linker1*- **Influenza A**- *Linker2*- **XhoI**- MITD- *Linker1*- **His tag**- Stop***- EcoRI- **.** It was constructed by Invitrogen GeneArt Gene Synthesis Regensburg (Thermo Fisher Scientific) with the following 673 bp DNA sequence cloned into vector pMA-RQ:

GGG**TCTAGA**CACCATGAGAGTGACAGCCCCTCGGACACTGATCCTGCTGCTTTCTGGTGCCCTGGCTCTGACA
GAAACATGGGCCGGATCT**GGATCC***GGAGGATCTGGTGGCGGAGGAAGCGGAGGA***GGCCACAGCTATACCAC**
AGCCGAAGAACTGGCCGGCATCGGCATCCTGACAGTGATTCTGGGAGTCCTGCTGCTG*GGAGGATCTGGTG*
*GCGGAGGAAGCGGAGGA***CTGAGCGTGAAGAAGCAGTTCGAGGAACTGACCCTGGGCGAGTTCCTGAAGCT**
GGACAGAGAGAGAGCCAAGAAT*GGAGGATCTGGTGGCGGAGGAAGCGGAGGA***CCCATTCTGAGCCCTCTG**
ACCAAGGGCATCCTGGGCTTCGTGTTCACACTGACCGTGCCTAGCGAGAGAGGACTG*GGAGGATCTCTTGGA*
*GGCGGAGGATCTGGA***CTCGAG**ATCGTGGGAATTGTGGCCGGACTGGCTGTGCTGGCCGTGGTGGTTATTGGA
GCTGTGGTGGCCACAGTGATGTGCAGAAGAAAGAGCAGCGGCGGCAAAGGCGGCAGCTATTCTCAGGCCGC
CTCTAGCGATTCTGCCCAGGGCTCTGATGTGTCTCTGACAGCT*GGAGGATCTGGTGGCGGAGGAAGCGGAGG*
*A***CATCATCACCATCACCAT**TAG**GAATTC**GGG

3.30 Cloning of tandem minigene

The tandem minigene was cloned from insert vector (pMA-RQ) into target vector pGEM4Z having a T7 promotor, for later RNA *in vitro* transcription. The target vector was kindly provided by Prof. Dr. Simone Thomas. At first, 5 µg of both vectors were digested with each 1 µl of XbaI and EcoRI (each 20 U/µl), 5 µl cutsmart buffer and H_2O in a total volume of 50 µl at 37°C. After two hours, target vector was dephosphorylated using 1 µl of CIP and incubating for another hour at 37°C. Heat inactivation was performed putting the sample for 10 minutes at 75°C. Then, 10 µl of loading dye was added to both samples and transferred into big wells of a 2% agarose gel. DNA fragments were size-separated at 120 V for 40 minutes. Afterwards, the big fragment of the target vector digested sample and the small fragment of the insert vector digested sample were excised using 70% UV light. Cutted samples were purified using Macherey-Nagel™ NucleoSpin™ Gel and PCR Clean-up Kit and concentration was measured. Samples were ligated using 70 ng of cipped target vector and 210 ng of insert DNA together with 1 µl of T4 Ligase (400 U/µl), ligase buffer and H_2O in a total volume of 15 µl. Ligation was performed over night at RT. On the next day, transformation was done using One Shot™ TOP10 Chemically Competent E. coli according to the manufacturer's instructions. Cell suspension was spread in LB ampicillin (100 µg/ml) agar plates and incubated in inverted position over night at 37°C. Ten clones were picked and incubated in 5 ml LB ampicillin medium over night at 37°C shaking. Then, plasmid DNA was purified using QIAGEN Plasmid Mini Kit and eluted in 50 µl H_2O. Concentration was measured and control digestion with XbaI and EcoRI was performed using 10 µl of DNA, 1 µl of each enzyme, cutsmart buffer and H_2O in a total volume of 20 µl. Samples were digested at 37°C for one hour. Success of digestion was analysed by electrophoresis (40 min, 120 V) using 1.5% agarose gels. Samples were sent for sequencing to proof for correct plasmid DNA. Samples were sequenced with T7 primer starting at plasmid's T7 promotor using DNA-Sequencing Service of Thermo Fisher Scientific.

3.31 *In vitro* transcription of RNA

Plasmid DNA (20 µg) containing the tandem minigene was linearized using 2 µl of enzyme SpeI, cutsmart buffer and H_2O in a total volume of 45 µl. Linearization was performed at 37°C over night. Then, DNA was purified using phenol/ chloroform/ isoamyl alcohol. Linearized DNA (20 µg) was filled up with H_2O to 400 µl. After resuspension, 400 µl of phenol/ chloroform/ isoamyl alcohol were added and vortexed. After centrifugation (14000 rpm, 5 min), aqueous upper phase containing the DNA was removed carefully and transferred into a new 1.5 ml tube. This was repeated for four times. Then, 400 µl of chloroform were added and samples were vortexed. After centrifugation (14000 rpm, 5 min), aqueous upper phase containing DNA was removed again carefully and transferred into a new 1.5 ml tube. This was also repeated for four times. After the last repetition, the aqueous phase was transferred into a RNase free 1.5 ml Eppendorf Safe-Lock Tube. 40 µl of sodium acetate were added and mixed before adding 1000 µl of 96-100% ethanol. Samples were incubated for 2 hours at -20°C. DNA was pelleted during centrifugation (14000 rpm, 30 min, 4°C). The supernatant was discarded without peeling off the DNA pellet from the tube rim. The DNA pellet was washed carfully with 1400 µl of cold 70% ethanol. Following centrifugation (14000 rpm, 30 min, 4°C), the supernatant was discarded and the DNA pellet was air-dried for 5 minutes. Adding 20 µl of pre-warmed RNase-free H_2O, the DNA was dissolved keeping samples at 40°C rotating. After concentration measurement, purified linearized DNA was stored at -80°C. *In vitro* transcription of linearized DNA was performed using mMessage mMachine T7 Ultra Kit. For this, 2.5 µg of DNA were transcribed based on manufacturer's instructions. Transcription reaction was performed at 37°C (2h). For poly(A) tailing, samples were incubated for 1 hour at 37°C. After *in vitro* transcription, RNA was purified using RNeasy Mini Kit according to manufacturer's instructions. RNA was eluted in 0.5 ml Eppendorf Safe-Lock Tubes using 50 µl nuclease-free H_2O. After concentration measurement, RNA was aliquoted and stored at -80°C.

3.32 Flow cytometric analysis for marker detection

Flow cytometric analysis was performed for detecting cell surface or intracellular markers. For this, cells were stained in FACS tubes or 96-well round bottom plates, depending on the assay and amount of samples (Table 29). Depending on the marker staining panel live/dead staining was done with viability dye eFluor 780 before surface staining or with Zombie Aqua™ Fixable Viability dye (1:500) together with surface marker staining. When using viability dye eFluor 780, 200 µl of 1:1000 in PBS diluted dye were put on the cells present in 50 µl of PBS. After incubation (20 min, RT, dark), FACS buffer was added to wash the cells. Following centrifugation, the supernatant was discarded and F_c receptors were blocked adding 2.5 µl of AB serum. Incubating for 10 minutes at RT, surface marker staining was performed by transferring 10 µl of mastermix including all antibodies against marker of interest to the cells. After mixing cell suspensions, surface markers were stained by incubating for 20 minutes at RT in the dark. Afterwards, cells were fixed by adding 200 µl of FluoroFix™ Buffer and incubating for 20 minutes at RT in the dark. Following another centrifugation step, cells were washed and permeabilized by using Intracellular Staining Permeabilization Wash Buffer (Perm-Wash) for two times. Discarding the supernatant, unspecific binding was again blocked by adding 2.5 µl of AB serum for 10 minutes (RT, dark). Intracellular staining was performed by adding antibodies and incubating 30 minutes at RT in the dark. In the end, cells were washed two times with Perm-Wash and once with FACS buffer. After adding 200 µl of FACS buffer, stained and fixed cells were measured directly or kept at 4°C in the dark until measurement using BD FACS LSR II or BD FACS Celesta™. When intracellular staining was performed eBioscience™ Protein Transport Inhibitor was added prior to staining. As positive control, T cells were activated with eBioscience™ Cell Stimulation Cocktail (plus protein transport inhibitors) prior to flow cytometry staining.

Table 29: Details for different staining formats

	FACS tube	96-well plate
Staining volume	50 µl	50 µl
Centrifugation	300g, 5 min, 4°C	1000g, 2 min, 4°C
Mixing	vortexing	pipetting up and down
Volume of FACS buffer	2 ml	200 µl
Volume of Intracellular Staining Permeabilization Wash Buffer (Perm-Wash)	2 ml	200 µl

To adjust the flow cytometry device and compensate spectral overlays of the fluorochromes compensation beads and controls including unstained cells and cells stained only with viability dye were prepared. One drop of vortexed negative beads and one drop of vortexed anti-mouse Ig beads were transferred in a 1.5 ml cup and X+1 times 50 µl of PBS (X = amount of used antibodies) was added. After vortexing, 50 µl of bead suspension was put into FACS tubes. One sample was left unstained, whereas remaining bead samples were stained with each used antibody and incubated at RT in the dark for 10 minutes. After washing with FACS buffer beads and unstained samples were also fixed as previously described. Staining for flow cytometric analysis was performed in tubes, when MelDTC cell lines were analysed for their marker expression, for isolating sorted single cells, when dendritic cells were generated and for minigene test. In 96-well format cells were stained for marker analysis of T cell expansion and killing assays.

3.33 Statistics

Statistical tests, like unpaired t test, one-way ANOVA with multiple comparisons or Pearson correlation were performed using *GraphPad Prism* version 6.1. P-values < 0.05 were accepted as significant.

4 Results

4.1 Sample acquisition and sample preparation

This study explores the clinical utility of disseminated cancer cells from sentinel lymph nodes (SLN-DCCs) of melanoma patients for adjuvant therapeutic target identification (Figure 3). First, tumor and non-tumor (control) cells are isolated from melanoma patients over time and space followed by DNA isolation for molecular analyses. Tumor cells are further used for the generation of DCC-derived cell lines and remaining patient samples are used for immune cell biobanking for later functional assays. Next, single-cell sequencing libraries of tumor and control cells are prepared for sequencing of copy number alterations (CNAs), microsatellite markers (MS) or whole exome. Based on MS and CNAs the phylogenetic relationship of disseminated melanoma cells is investigated, whereas whole exome sequencing (WES) is used for neoantigen identification. Finally, identified neoantigens are functionally evaluated using patient-derived cell lines and immune cells. Thus, fundamental questions of adjuvant immunotherapy against early systemic cancer of melanoma are explored.

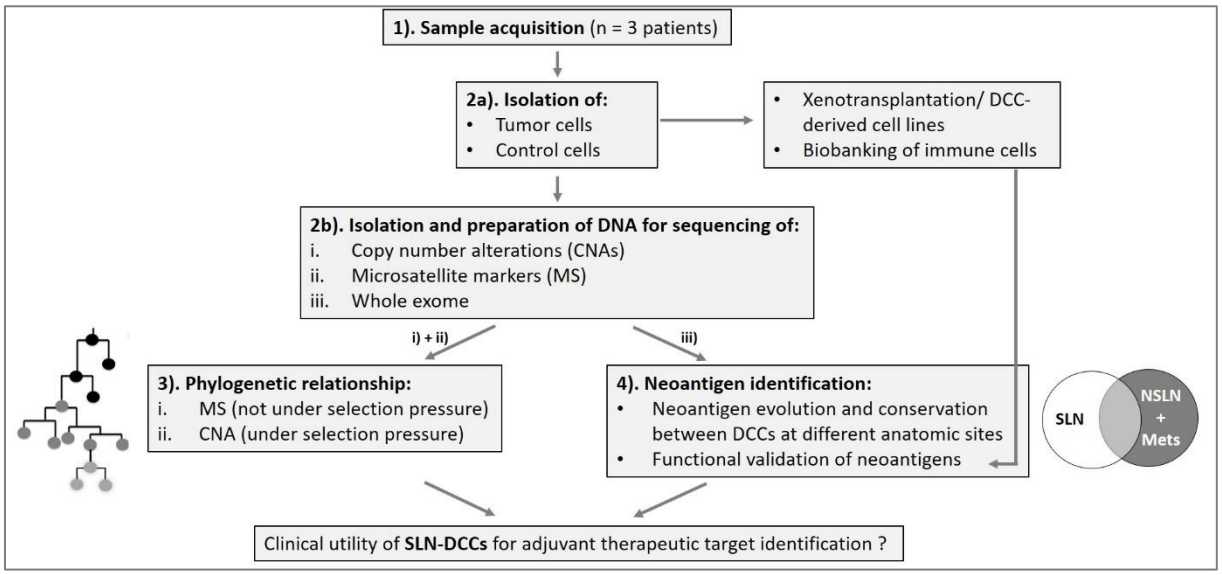


Figure 3: Project overview
From acquired samples, single tumor and control cells are isolated for DNA isolation and tumor cells for xenotransplantations to generate DCC-derived cell lines. Sample leftovers are used for immune cell biobanking for later functional assays. Single-cell sequencing libraries of tumor and control cells are prepared for sequencing of copy number alterations (CNAs), microsatellite markers (MS), or whole exome. Based on MS and CNAs the phylogenetic relationship of disseminated melanoma cells are investigated, whereas whole exome sequencing (WES) is used for neoantigen identification. Identified neoantigens are functionally evaluated using patient-derived cell lines and immune cells. With this approach the clinical utility of SLN-DCCs for adjuvant therapeutic target identification will be explored.

For analysing the phylogenetic relationship of disseminated cancer cells, several samples from different patients were collected, processed and prepared. Chapter 4.1.1 covers the criteria for patient and sample recruitment and provides details on the medical histories of the patients. Chapters 4.1.2 and 4.1.3 describe, how individual cells and DNA were isolated and prepared for molecular analyses. Copy number alteration (CNA) analysis of isolated samples is shown in chapter 4.1.4.

4.1.1 Sample acquisition and courses of disease

4.1.1.1 Sample acquisition

To determine the phylogeny and evolution of disseminated cancer cells (DCCs) and neoantigens present at different anatomic sites and stages of disease, melanoma patients were recruited in cooperation with Dr. med. Philipp Renner and PD Dr. med. Sebastian Haferkamp according to the following criteria:

1. Patients from whom DCCs from sentinel lymph nodes (SLNs) and/or non-sentinel lymph nodes (NSLNs) were available;
2. Patients for whom the archived primary tumor was accessible;
3. Patients who developed metastases;
4. Patients from whom a DCC-derived xenograft or cell line was established to allow functional evaluations of candidate neoantigens;
5. Patients who gave informed consent;

For downstream analyses, complete data-sets of longitudinally tracked patients were neccessary. More than 300 melanoma patients undergoing routine sentinel lymph node extirpation had to be screened for SLN-DCCs. Finally, three patients matched the selection criteria (MM15-127, MM16-412, MM16-423) and were analysed in this study. Reasons for patient drop-out included (i) no informed consent, (ii) DCC-negative SLN or NSLNs, (iii) no metastasis or (iv) no established cell line. Different tissues and samples including the primary tumor, lymph nodes, metastases, buccal swabs and blood were acquired from selected patients. However, not all samples and tissues that were resected during surgeries and were mentioned in the medical histories of the patients were accessible for the analysis. An overview of available samples from each patient is provided in Figure 4.

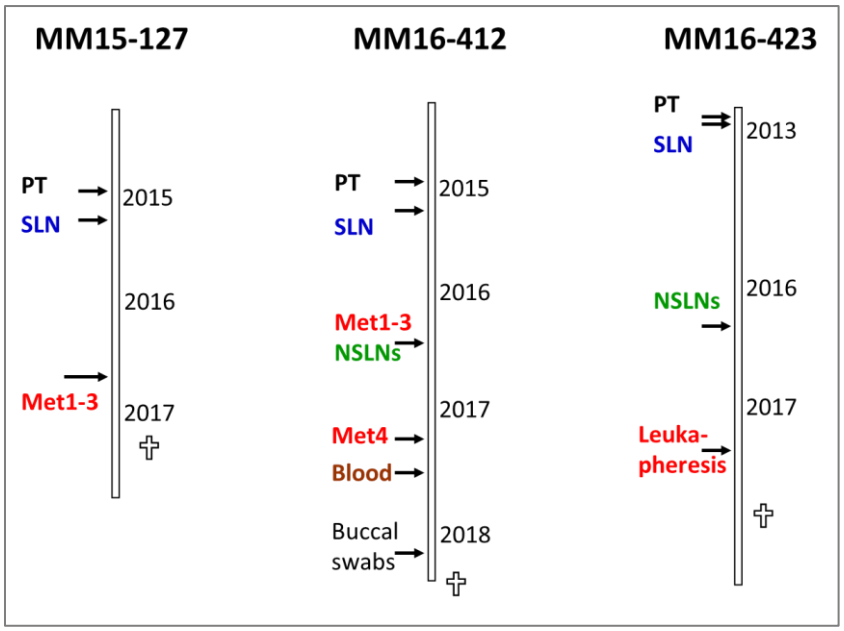


Figure 4: Overview of melanoma patient samples
From three melanoma patients, samples were acquired longitudinally. Tumor cells and control cells were isolated from the different samples. PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; Met: metastasis; ✞: death.

For all three patients, sentinel lymph nodes (SLNs) were extirpated shortly after primary tumor removal. Additionally, from two of the three patients up to six non-sentinel lymph nodes (NSLNs) excised during complete lymph node dissection (CLND) were accessible. All three patients developed

metastases and from MM15-127 and MM16-412 cutaneous metastases were available. As MM16-423 developed not-accessible liver metastases, a leukapheresis was performed for isolating circulating tumor cells (CTCs). For all patients, tumor cells and control cells (non-tumor cells) were isolated depending on the sample availability from either lymph nodes, metastases, blood or buccal swabs. In the next chapter, the medical histories of the patients are delineated in detail.

4.1.1.2 Patient histories

MM15-127

The first patient to be described is MM15-127, a woman who was born in 1936 (Figure 5). In February 2015, she was diagnosed with nodular malignant melanoma located at her right lower leg. The primary tumor had a thickness of 6 mm, showed ulceration and invaded into the subcutaneous tissue. Two months later she underwent sentinel lymph node biopsy with three inguinal lymph nodes being removed, out of which one was found to harbor tumor cells (DCCD = 79). As there was one tumor-involved node and in-transit, satellite, and/or microsatellite metastasis were present without evidence of distant metastasis, the patient was diagnosed with melanoma stage IIIB (T4b N2c M0) based on AJCC 2009. SLN-DCCs were transplanted into an immunodeficient mouse for generating a xenograft-derived cell line (MelDTC_9). In June 2015, complete inguinal lymph node dissection was performed removing non-sentinel lymph nodes (NSLNs) not available for this project. The first lymph node recurrence occurred in August 2015, 131 days after SLN biopsy and the second recurrence, 212 days after SLN biopsy in November 2015. In March, May and August 2016, the patient was treated with palliative electrochemotherapy using bleomycin. In October 2016, she was enrolled in the MASTERKEY 265 study receiving intravenously anti-PD-1 (Pembrolizumab) combined with intralesional injection of the oncolytic virus T-VEC (Talimogen laherparepvec, IMLYGIC/ Amgen) or the T-VEC placebo. At the end of 2016, several cutaneous metastases (Met1-3) were found and removed. Finally, the patient succumbed to melanoma in January 2017. From MM15-127 the existence of any factors increasing the risk of developing melanoma is not reported.

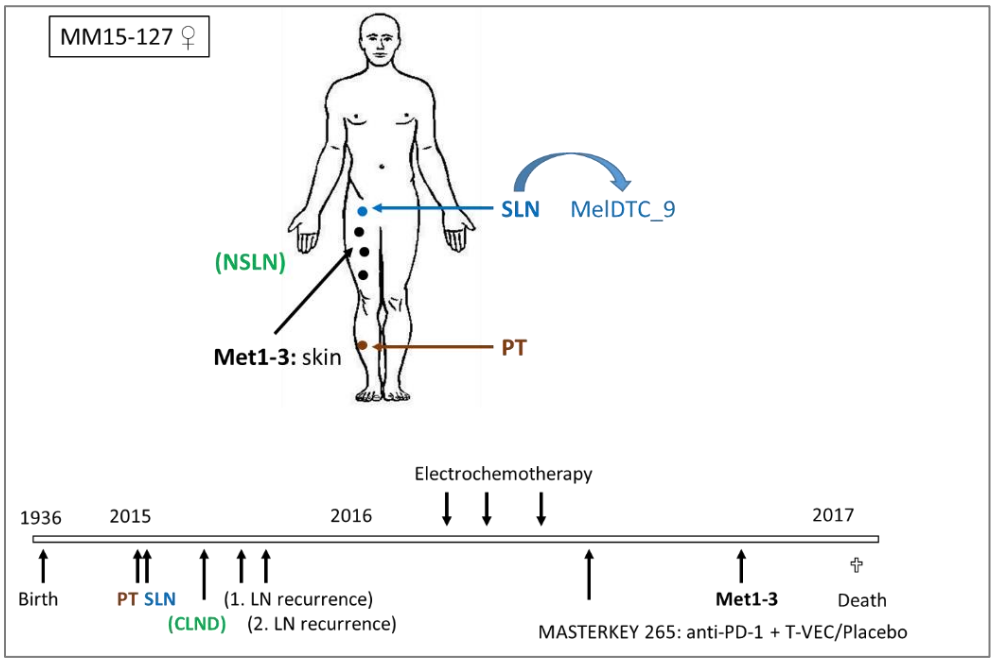


Figure 5: Medical history of MM15-127
Timeline and body map of MM15-127's tumor development, progression and treatment. Tumor cells from the SLN were used for autologous MelDTC_9 cell line generation. PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; Mets: metastases; CLND: complete lymph node dissection. Samples in brackets were not available for this project.

MM16-412

The second patient, MM16-412, was male and born in 1936 (Figure 6). In January 2015, a 2-3 mm thick superficial spreading melanoma was removed at his left abdomen showing no ulceration but invasion into the reticular dermis, the lower layer of the dermis. One month later, a SLN biopsy was performed (DCCD = 9.5). As there were no distant metastases detected, MM16-412 was initially diagnosed with melanoma stage IIIA (T2a N1a M0) based on AJCC 2017. In March 2015, the patient underwent regional lymph adenectomy finding 12 tumor-free nodes. After one month, the patient was treated with adjuvant radiation therapy. In January 2016, one year after primary tumor removal, the first lymph node recurrence was diagnosed. One month later, multiple inguinal lymph node metastases were found with additional multiple cutaneous metastases, but were not accessible for this project. A second complete lymph node dissection (CLND) followed in May 2016. In 21 from 33 nodes tumor cells were found. Six of those lymph nodes were available for this study (NSLN1-6) having different tumor burden reflected by the DCCD (Table in Figure 6). Tumor cells from NSLN3 and 6 were used to generate two patient-derived cell lines via spheres (MelDTC_10a/b). Simultaneously with CLND, cutaneous metastases (Met1-3) were removed. Similar to MM15-127, the patient was recruited for the MASTERKEY 265 study in September 2016 treated with anti-PD-1 (Pembrolizumab) and the oncolytic virus T-VEC (Talimogen laherparepvec, IMLYGIC/ Amgen) or the T-VEC placebo. In March 2017, another immunotherapy was initiated receiving 3 mg/KG anti-PD-1. At the same time, a cutaneous metastasis (Met4) was removed from the upper leg from the left groin. Few days later, 50 ml blood was taken from the patient for isolating non-tumor control cells needed within this project. At the end of April, another metastasis (Met5) was excised from the left abdominal wall, but was not analysed in this study. Two months later, new lymph node metastases were found at the middle abdomen and in August 2017 he was diagnosed with lung metastases. In April 2018, buccal swabs were taken from MM16-412 for the isolation of oral epithelial cells. The disease progressed intra-abdominally and several mesenterial lymph node metastases were diagnosed in May 2018. Finally, the patient died to melanoma mid of the year 2018. With regard to melanoma-related risk factors, MM16-412 was frequently exposed to UV radiation leading to several sunburns during his lifetime.

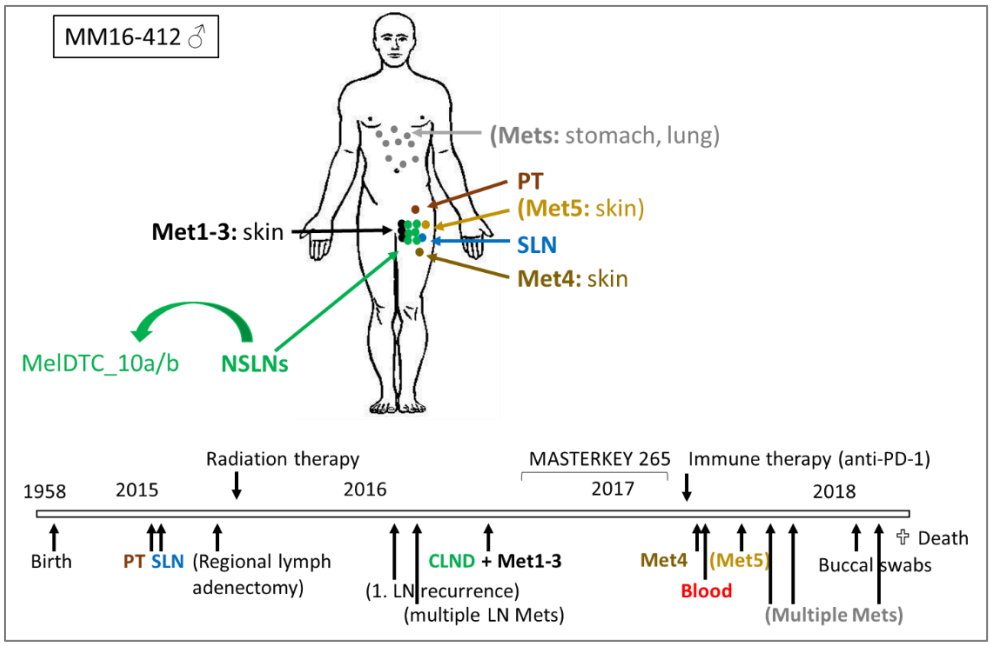


Figure 6: Medical history of MM16-412
Timeline and body map of MM16-412's tumor development, progression and treatment. Tumor cells from NSLNs were used for cell line generation (MelDTC_10a/b). Disseminated cancer cell densities (DCCDs) of the different lymph nodes are shown in the table. PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; Mets: metastases; CLND: complete lymph node dissection. Samples in brackets were not available for this project.

MM16-423

The third patient, MM16-423, a man born in 1958 (Figure 7) got the diagnosis of nodular malignant melanoma of the lower extremity, which was removed from his left heel in January 2013, at the age of 54. The primary tumor had a thickness of 4.4 mm with ulcerations and invaded into the reticular dermis. Three weeks later, he underwent inguinal sentinel lymph node biopsy finding one lymph node with tumor cells (DCCD = 2). Hence, he was diagnosed with stage IIIB melanoma (T4b N1a M0) based on AJCC 2017 criteria as no distant metastases were found. In May 2013, he received adjuvant chemotherapy and was treated with Interferon-α for 18 months. Two and a half years later, the first lymph node recurrence occurred in June 2016 leading to complete lymph node dissection. From nine inguinal lymph nodes six nodes (NSLN1-6) containing tumor cells were available for this project with varying DCCDs (Table in Figure 7). DCCs from NSLN1 were used for autologous cell line generation via spheres (MelDTC_11). Over the year 2017, he developed multiple non-accessible liver metastases. Thus, a diagnostic leukapheresis was performed for isolation of circulating tumor cells (CTCs) in September 2017. He finally succumbed to melanoma in December 2017. During his life, MM16-423 had multiple severe sunburns caused by enhanced UV exposition increasing the risk for developing melanoma.

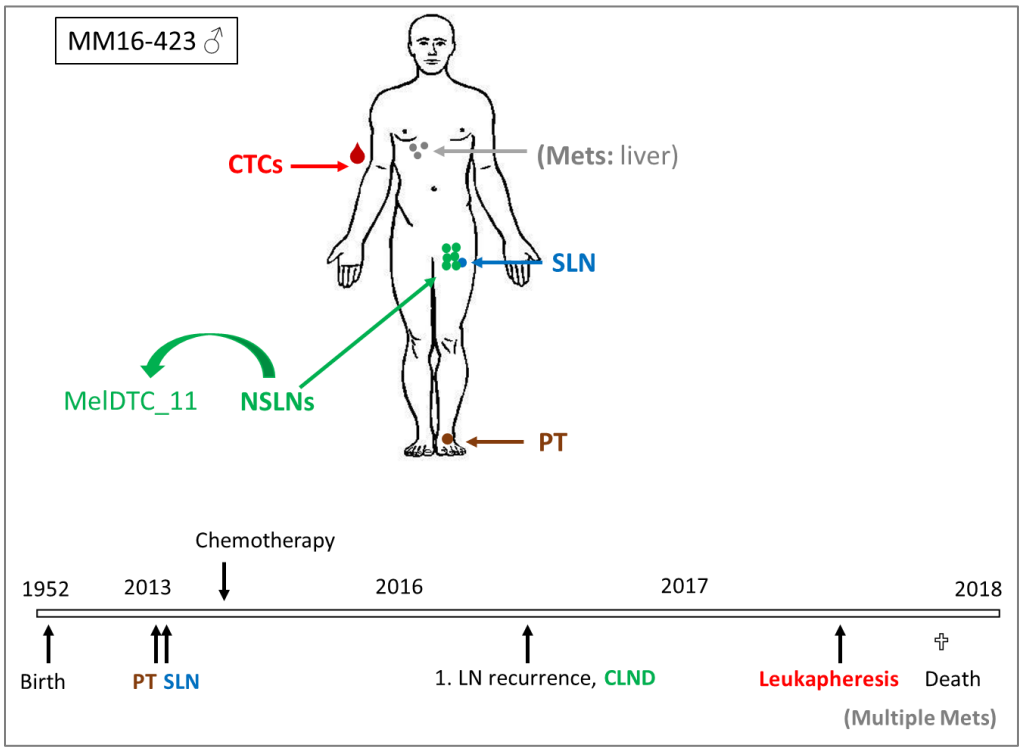


Figure 7: Medical history of MM16-423
Timeline and body map of MM16-423's tumor development, progression and treatment. Tumor cells from NSLN1 were used for autologous MelDTC_11 cell line generation. Disseminated cancer cell densities (DCCDs) of the different lymph nodes are shown in the table. PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; CTCs: Circulating tumor cells; Mets: metastases; CLND: complete lymph node dissection. Samples in brackets were not available for this project.

4.1.2 Isolation of single tumor cells and non-tumor control cells

For analysing the phylogeny of DCCs, single tumor and non-tumor cells as well as bulk DNA needed to be isolated from different anatomical sites and disease stages. To this end, genomic DNA of single tumor and control cells as well as from bulk samples were isolated from primary tumors, lymph nodes, metastases, blood and buccal swabs from available tissues of the three longitudinally sampled melanoma patients (Figure 4). All single cells were identified based on specific protein marker-expressions or morphologic criteria and manually isolated using a micromanipulator.

4.1.2.1 Isolation of tumor cells

All tumor cells from lymph nodes (Figure 8 A), patient-derived cell lines/ xenograft (Figure 8 C), cutaneous metastases (Figure 8 D and F) or blood (Figure 8 E) were isolated using antibody staining against the highly specific melanoma marker gp100. From one SLN sufficient material was available enabling the isolation of additional living DCCs based on MCSP staining (Figure 8 B).

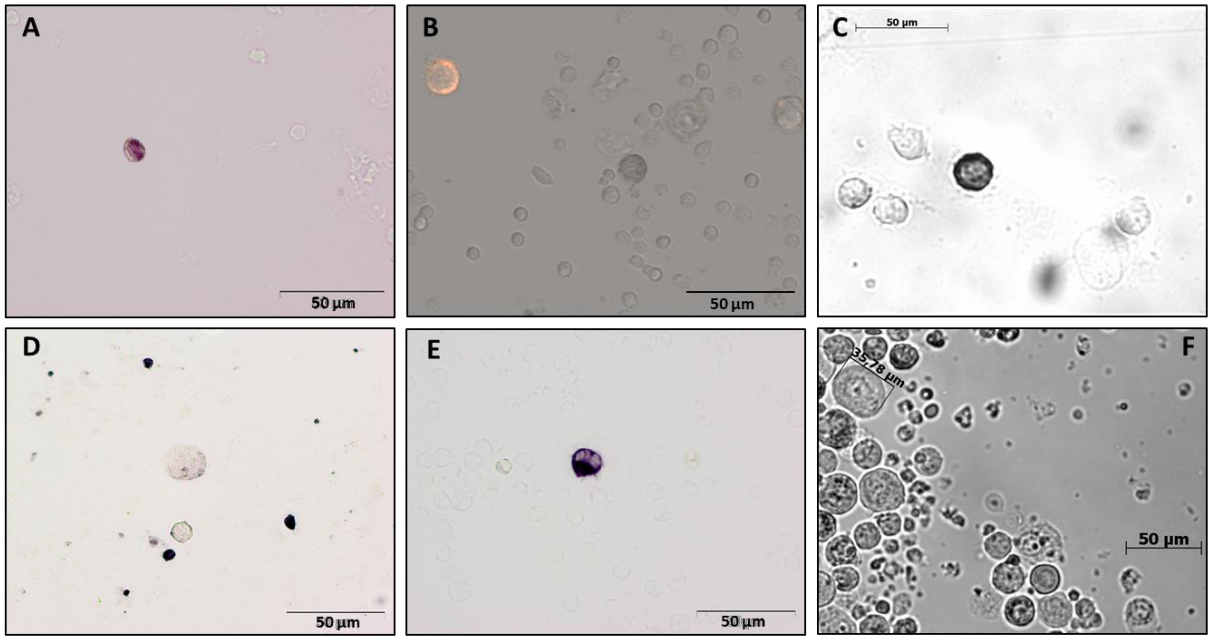


Figure 8: Pictures of tumor cells isolated from different tissues
Pictures of gp100-stained **(A)** or MCSP-stained **(B)** tumor cells from lymph node samples. Gp100-stained slides from an autologous cell line of patient MM16-412 **(C)** and cutaneous metastases of patient MM15-127 **(D)** are shown. **E:** Gp100-positive circulating tumor cell from blood of MM16-423. **F:** Gp100-negative staining of Met4 from MM16-412.

A summary of all isolated single tumor cells is provided in Table 30. Between 4 and 25 gp100-positive cells were isolated from the SLNs and from each of the six NSLNs, if available. In total, 280 gp100-positive cells were isolated from lymph nodes with different DCCDs (Table 31).

Table 30: Overview of isolated single tumor cells from different patients

Source	Marker for cell selection	MM15-127	MM16-412	MM16-423
SLN	gp100	20	20	4
	MCSP	59	--	--
NSLN1		--	20	20
NSLN2		--	20	20
NSLN3		--	20	13
NSLN4		--	20	20
NSLN5		--	19	19
NSLN6		--	25	20
DCC-**PDX /-cell line**	gp100	26	44	35
Cutaneous Met1	gp100	44	21	--
Cutaneous Met2	gp100	--	21	--
Cutaneous Met3	gp100	44	--	--
Cutaneous Met4	Size	--	20	--
Blood (CTCs)	gp100	--	--	19
Sum of isolated single cells		**193**	**250**	**170**
			∑ 613	

Number of isolated single tumor cells from different tissues of the three melanoma patients.

Table 31: Overview of DCCDs for the different lymph nodes

Source	MM15-127	MM16-412	MM16-423
SLN	79	9.5	2
NSLN1	--	750 000	900 000
NSLN2	--	500 000	50 000
NSLN3	--	750 000	10
NSLN4	--	750 000	58
NSLN5	--	258	42
NSLN6	--	50 000	138

DCCD: number of gp100-positive cells per million lymph node cells.

Additional 59 living DCCs were isolated from the SLN from MM15-127 based on MCSP expression (Figure 8 B). For MM15-127 two of three available cutaneous metastases were analysed, whereas for MM16-412 three of received four skin metastases were included in the analysis. For Met1 and Met3 of MM15-127, having 61% and 64% gp100-positive cells, 44 single cells were isolated each. For cutaneous metastases of MM16-412, 21 gp100-positive cells were isolated from each Met1 and Met2. For Met4 no gp100-positive cell could be found (Figure 8 F). However, very large cells were noted (Figure 9 A). These cells were also negative for melanoma marker Melan-A (Figure 9 B). As the pathologist rated these cells as most likely tumor cells, however with a de-differentiated phenotype due to the missing expression of melanoma markers, 20 of these large cells were isolated for whole genome amplification (WGA).

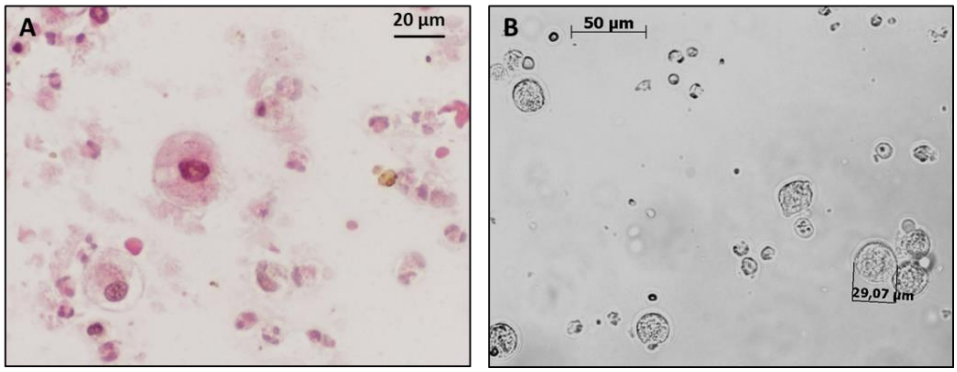


Figure 9: Pictures of melanoma marker-negative cutaneous Met4 of MM16-412
HE staining **(A)** and Melan-A staining **(B)** of Met4 showing very large cells.

Moreover, whole transcriptome amplification (WTA) was performed for additional 20 single cells to assess the transcriptome of these large cells for analysing the expression of additional melanoma markers. After WTA, cDNA of those cells was used for a melanoma marker PCR. Here, the expression of tyrosinase (TYR), an enzyme playing a role in melanogenesis, and tyrosinase-related protein 2 (TRP2) were evaluated on single cell level showing negative results (Figure 10 A and B).

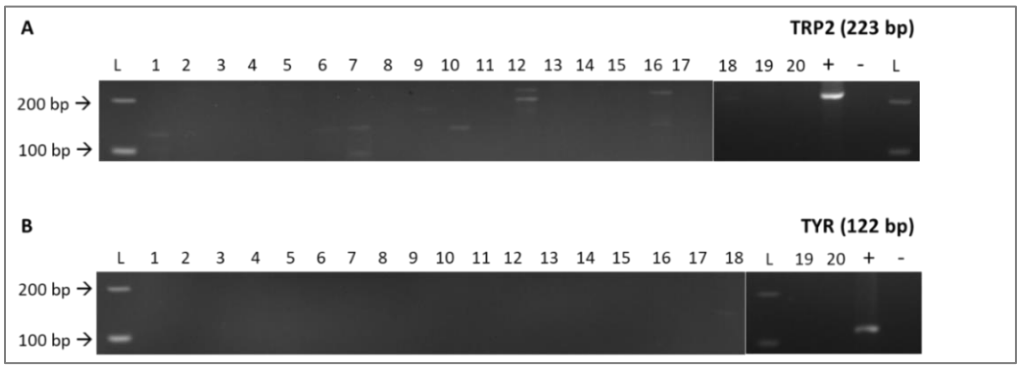


Figure 10: Agarose gels of melanoma marker PCRs of MCSP-negative cells of MM16-412 Met4
Results of PCRs for TRP2 **(A)** and TYR **(B)** on cDNA of large MCSP-negative cells isolated for WTA (1-20). As positive control (+) gDNA of melanoma cell line MelHo was used showing bands at 223 bp for TRP2 and at 122 bp for TYR. L: 1 kb Plus DNA ladder.

To test whether large MCSP-negative cells were able to form tumors, cells from digested Met4 were transplanted into NSG-mice (Table 32). Formed tumors were negative for melanoma markers. In addition, the tumor section of one sample was negative for cytokeratin, desmin and sm-actin excluding that formed tumor was a carcinoma or a muscle tumor (Table 32).

Table 32: Xenograft generation and analysis derived from cells of Met4 from MM16-412

Mouse ID	Injected cells	Injection - palpable tumor/ analysis	Stainings of FFPE sections	Staining result
70-240	1.2x 10^6	21 days / 35 days	HE, pan melanoma, S100, MiTF, desmin, sm-actin, cytokeratin	Melanoma marker-negative
70-241	100 large MCSP-negative	27 days / 130 days	HE, pan melanoma	Melanoma marker-negative

In summary, large cells from MM16-412 Met4 were able to form tumors in NSG-mice, indicating that they were really tumor cells without the expression of melanoma markers. Due to this, those cells were included in the analysis.

4.1.2.2 Isolation of non-tumor control cells

For isolating non-tumor (control) cells, hematopoietic cells were isolated from blood, lymph nodes or metastases based on staining for CD3 (T cells), CD19 (B cells) or CD68 (macrophages). Endothelial cells were isolated from lymph nodes based on CD31 expression (Figure 11 A). Oral epithelial cells (OECs) were isolated from oral mucosa using buccal swabs based on morphologic criteria (Figure 11 B). An overview of isolated control cells from the different tissues of the melanoma patients is shown in Table 33. In total, 376 single non-tumor control cells were isolated from the three patients.

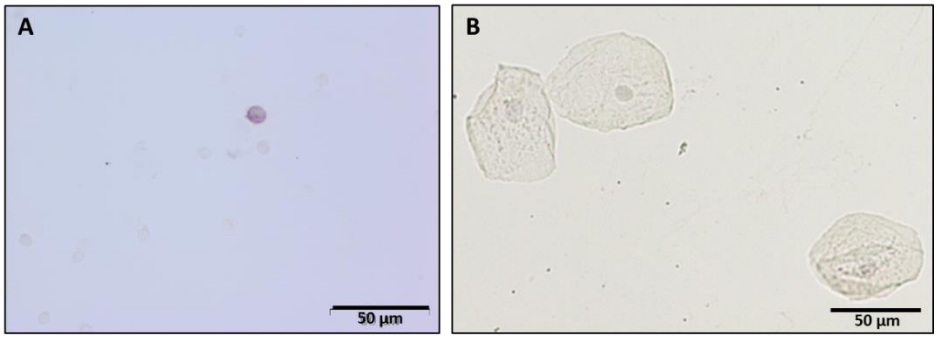


Figure 11: Pictures of CD31-positive cell from lymph node and OECs from oral mucosa
Endothelial cells **(A)** were isolated based on CD31 expression and oral epithelial cells (OECs) using morphologic criteria **(B)**.

Table 33: Overview of isolated single non-tumor cells from different patients

	Marker for cell selection	MM15-127	MM16-412	MM16-423
T cells	CD3	**40**	**40; 18; 13**	**41**
Source		SLN	NSLN2; Met1; Met4	NSLN2
Macrophages	CD68	--	**40; 18; 14**	**41**
Source			NSLN2; Met1; Met4	NSLN2
B cells	CD19	--	**28**	--
Source			PB	
Endothelial cells	CD31	**16**	--	**34**
Source		SLN		NSLN1-5
OECs	Morphology	--	**33**	--
Source			OM	
Sum of isolated single cells		**56**	**204**	**116**
			Σ 376	

Number of isolated single non-tumor cells from different tissues of the three melanoma patients. OEC: oral epithelial cell; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; Met: metastasis; PB: peripheral blood; OM: oral mucosa.

4.1.3 Isolation of DNA from single cells and bulk

For molecular analyses, high-quality DNA was obtained from either single cells or bulk samples. All single cells were subjected to whole genome amplification (WGA). Quality was controlled for each individual WGA product using a four marker multiplex PCR assay (WGA-QC) informing about the genome integrity index (Figure 12). The genome integrity index (GII) values reflect the quality of the DNA and range from 0 to 4 with 4 representing the highest quality (Polzer et al. 2014). For this study 'high-quality' DNA is determined by a GII larger than 2 for control and larger than 1 for tumor cells.

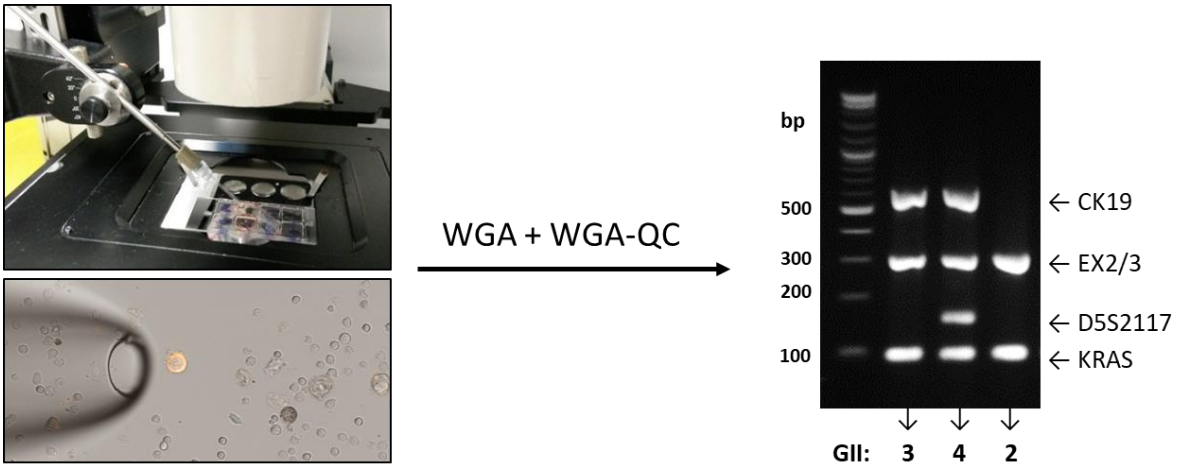


Figure 12: DNA isolation of single cells
Cells of interest were identified and manually isolated with the help of a micromanipulator. After whole genome amplification (WGA) and quality control (WGA-QC) high-quality cells were determined using the genome integrity index (GII).

4.1.3.1 DNA isolation from single cells

Table 34 shows an overview of high-quality DNA isolation from single cells. The highest percentage of high-quality DNA (86%) was obtained from single cells of MM16-412. From lymph node samples, the lowest DNA isolation efficiency was found for patient MM16-423. Of its NSLN3 and NSLN5, only 31% and 37% of cells showed high-quality DNA isolating mainly cells with degraded DNA. Differences in high-quality DNA isolation efficiencies were seen in cells from cutaneous metastases with 48% for samples from MM15-127 and 85% for samples of MM16-412, thus showing that single cells from metastatic tissues of MM16-412 had higher DNA qualities. Regarding control cells, variations in DNA isolation efficiencies (65-100%) were found for each patient between different cell types and for patient MM16-412 within each cell type between different tissues. In general, oral epithelial cells had lower DNA qualities than T cells or macrophages. However, similar efficiencies (65-69%) to that from OECs (70%) were seen for endothelial cells from NSLN1-5 of MM16-423 and T cells from Met4 of MM16-412. In summary, high-quality DNA was successfully isolated from in total 792 of 989 single cells (80%).

Table 34: Overview of isolated high-quality DNA from single cells of different patients

	MM15-127	MM16-412	MM16-423
SLN-DCCs (all)	**89% [70/79]**	**75% [15/20]**	**75% [3/4]**
Gp100	85% [17/20]	75% [15/20]	75% [3/4]
MCSP	90% [53/59]	--	--
NSLN-DCCs (all)	**--**	**86% [107/124]**	**67% [75/112]**
NSLN1	--	90% [18/20]	80% [16/20]
NSLN2	--	100% [20/20]	70% [14/20]
NSLN3	--	80% [16/20]	31% [4/13]
NSLN4	--	75% [15/20]	100% [20/20]
NSLN5	--	74% [14/19]	37% [7/19]
NSLN6	--	96% [24/25]	70% [14/20]
DCC-PDX /-cell line cells	**92% [24/26]**	**93% [41/44]**	**94% [33/35]**
Cells from metastasis (all)	**48% [42/88]**	**85% [53/62]**	**79% [15/19]**
Cutaneous Met1	41% [18/44]	67% [14/21]	--
Cutaneous Met2		95% [20/21]	--
Cutaneous Met3	55% [24/44]	--	--
Cutaneous Met4	--	95%[19/20]	--
Cutaneous Met5	--	--	--
CTCs from blood	--	--	79% [15/19]
Control cells (all)	**88% [49/56]**	**85% [174/204]**	**78% [91/116]**
T cells	SLN: 85% [34/40]	NSLN2: 98% [39/40] Met1: 100% [18/18] Met4: 69% [9/13]	NSLN2: 80% [33/41]
Macrophages	--	NSLN2: 78% [31/40] Met1: 78% [14/18] Met4: 93% [13/14]	NSLN2: 88% [36/41]
B cells	--	PB: 96% [27/28]	--
Endothelial cells	SLN: 94% [15/16]	--	NSLN1-5: 65% [22/34]
OECs	--	OM: 70% [23/33]	--
Sum of isolated cells	**74% [185/249]**	**86% [390/454]**	**76% [217/286]**
		Σ 80% [792/989]	

Efficiency (percentage and absolute numbers) of high-quality DNA isolation from single cells derived from different tissues for each melanoma patient. SLN: sentinel lymph node; NSLN: non-sentinel lymph node, Met: metastasis; PDX: patient-derived xenograft; CTC: circulating tumor cell; OEC: oral epithelial cell; PB: peripheral blood; OM: oral mucosa.

4.1.3.2 DNA isolation from bulk samples

Isolation of bulk DNA from primary tumors

In addition to single DCCs, the primary tumors (PTs) of the patients were analysed. For this, FFPE blocks differing in their block quality were collected from resident pathologists (Figure 13). The PTs of MM15-127 and MM16-412 were well preserved, whereas MM16-423 PT was not fixed optimally.

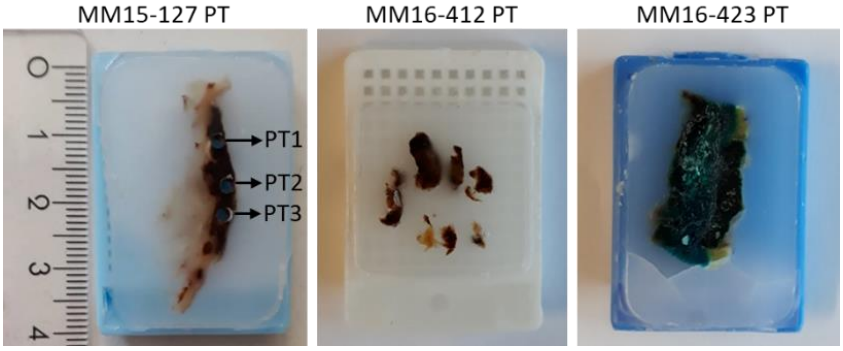


Figure 13: FFPE blocks of primary tumors from melanoma patients
PT: primary tumor; PT1-3: punches of MM15-127 PT used for DNA isolation.

For bulk DNA isolation either punches or sections were used (Table 35). From the primary tumor of MM15-127 three individual punches were taken (PT1-3), from MM16-412 PT one punch and from MM16-423 PT a section of the FFPE block were used in the analysis. From punches high amounts of DNA could be obtained, in contrast to the section of MM16-423 PT (Table 35). DNA was quantified using ratios of absorbance A260/280 and A260/230 to assess purity of the samples. From the PTs of MM15-127 and MM16-412 pure DNA was isolated, whereas DNA from MM16-423 PT showed a low quality and potential contaminations with organic substances (A260/230 < 2.0).

Table 35: Overview of DNA isolation from primary tumors

	Input material	DNA [ng/µl]	Total amount [µg]	A260/280	A260/230
MM15-127 PT1	Punch	603	30.2	1.984	2.033
MM15-127 PT2	Punch	797	39.9	1.970	1.901
MM15-127 PT3	Punch	1103	55.2	1.965	2.133
MM16-412 PT	Punch	1050	52.5	1.985	2.261
MM16-423 PT	Section	66.7	3.3	1.861	1.186

First, unamplified bulk DNA isolated from primary tumors was quality-controlled by performing standard WGA-QC and WGA-QC bulk, which was specifically designed for DNA from FFPE samples (Figure 14 A and B 'no WGA'). Primary tumor DNA from MM16-412 and MM16-423 had a GII of 4 and MM15-127 PT had a GII of 2 after standard WGA-QC. However, in the quality control for FFPE samples (WGA-QC bulk) only one amplicon was present for DNA from MM15-127 PT (Figure 14 A and B 'no WGA'). In contrast, MM16-412 PT was positive for 11 and MM16-423 PT for 10 amplicons showing DNA of high quality.

A

Quality control	Amplicon	MM15-127 PT1-3	MM16-412 PT	MM16-423 PT	+	-
WGA-QC bulk Multiplex #1	KRAS					
	BRAF Exon15					
	PIK3CA Exon20					
	EGFR Exon19					
	TP53 Exon					
WGA-QC bulk Multiplex #2	NRAS Exon2					
	TP53 Exon8*					
	EGFR Exon21					
	EGFR Exon18					
WGA-QC bulk Multiplex #3	TP53 Exon5/6					
	EGFR Exon20					
	CKND2A Exon3					
WGA-QC	GII	2	4	4	4	0

B

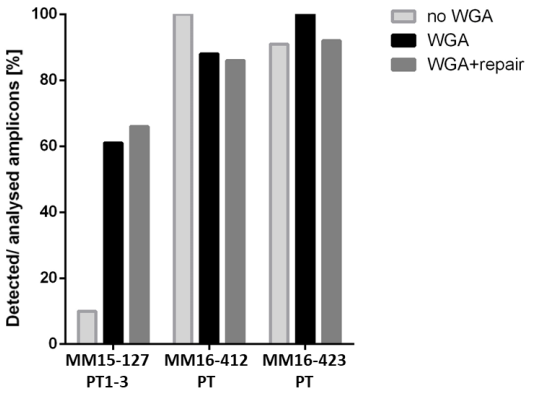


Figure 14: Evaluation of quality controls of primary tumor bulk DNA
A: Result of quality controls of unamplified bulk DNA isolated from PTs. WGA: whole genome amplification; QC: quality control; PT: primary tumor; *: primers failed; GII: genome integrity index; +/-: positive/negative control; grey color: amplicon was present. **B:** Comparison of WGA-QC bulk results of different PT samples. No WGA: unamplified; WGA+repair: amplified with additional DNA repair step. The more amplicons detected in WGA-QC bulk, the higher the DNA quality.

As sample preparations for downstream analyses were established for WGA samples, bulk DNA isolated from FFPE primary tumor blocks was additionally amplified using WGA. However, crosslinking fixative formaldehyde used for tissue fixation may also react directly with nucleic acids creating DNA-DNA or DNA-protein crosslinks or degraded DNA (Campos and Gilbert 2012), which can interfere with WGA. NEBNext® FFPE DNA Repair Mix can reverse chemical DNA modifications and was added before MseI DNA digestion of the WGA protocol. In order to see, whether DNA quality could be improved after WGA, the relevance of an additional DNA repair step using NEBNext® FFPE DNA Repair Mix was tested. WGA improved the DNA quality of MM15-127 and MM16-423 after WGA-QC bulk (Figure 14 B and appendix Table 54 and Table 55). When adding the DNA repair step during WGA, a furhter improvement was only seen for MM15-127 PT samples. The highest DNA quality was seen in case of MM16-412 for unamplified PT bulk samples and in case of MM16-423 for amplified PT bulk samples

(Figure 14 B and appendix Table 55). In summary, performing WGA on bulk samples improved the DNA quality from MM15-127 remarkably. However, the additional repair step could not improve PT DNA qualities for all patients. Nevertheless, as WGA resulted in high-quality DNA, amplified bulk DNA from primary tumors was suitable for copy number alteration analysis.

Isolation of bulk DNA from germline samples

For isolating germline DNA of the melanoma patients, PBMCs from blood samples were used. However, blood samples were only accessible from two patients. For MM16-412 EDTA blood was available and from MM16-423 a leukapheresis was performed. In total, 15 and 19.5 µg of genomic DNA were isolated from PBMCs of MM16-412 and MM16-423, respectively. For MM15-127 only bulk DNA from SLN cells was available. As the SLN had a relatively low DCCD of 79, and thus consisted of mainly non-tumor cells, bulk DNA thereof was representative for germline DNA and used for microsatellite sequencing. For whole exome sequencing, a pool of CD31-positive endothelial cells were used as corresponding 'germline' sample (chapter 4.3.1).

4.1.3.3 Summary and overview of isolated DNA from single cells and bulk samples

From the three longitudinally sampled melanoma patients genomic DNA of 989 single DCCs and control cells from lymph nodes, metastases, blood and oral mucosa were manually isolated and subjected to WGA. Of the 989 isolated single cells, high-quality DNA could be obtained from 792 cells (Table 36). The remaining 197 WGA products had a low GII indicating isolation of degraded or no genomic DNA and were therefore excluded from further analyses. In addition to single cells, bulk DNA from primary tumors and germline samples were obtained for all patients and used for downstream analyses.

Table 36: Summary of isolated high-quality DNA from single cells and bulk samples

	Marker for cell selection	MM15-127	MM16-412	MM16-423
PT		✓	✓	✓
Germline		✓	✓	✓
SLN	gp100	17	15	3
	MCSP	53	--	--
NSLNs (6 LN/patient)	gp100	--	107	75
Metastasis (skin or CTCs from blood)	gp100	42	53	15
DCC-**PDX** and DCC-**cell line**	gp100	24	41	33
Control cells: T cells, macrophages, B cells, endothelial cells, OECs	CD3, CD68, CD19, CD31	49	174	91
Sum of cells with high-quality DNA		**185**	**390**	**217**
			∑ 792	

4.1.4 Copy number alteration analysis

Copy number alteration (CNA) analysis was performed for all single cells and primary tumor bulk samples (i) to control for copy number variations of protein marker-based isolated single cells and tumor DNA and (ii) to generate CNA-based phylogenetic trees. Tumor cells were expected to show copy number alterations, whereas non-tumor control cells not (Figure 15). All single tumor and non-tumor cells were so far isolated based on the expression of a single protein marker thought to be representative of a certain cell lineage, like CD68 for macrophages or CD3 for T cells. But, as molecular mimicry has been reported for melanoma, for example melanoma cells expressing CD31 (Pisacane et al. 2007) or CD68 (Pernick et al. 1999, Banerjee and Harris 2000), the group assignment might be error-prone. Therefore, the correct assignment of individual cells to the group of control and tumor cells was decided to be confirmed on the molecular level by determining CNAs of putative tumor cells and also for all control cells. For this, samples were subjected to genome-wide copy number profiling with the *Ampli1*™ LowPass Kit for Illumina® resulting in scatter plots visualized CNA profiles with genomic gains

marked in red and genomic losses in blue. The level of noise is measured by the derivative log ratio spread (DLRS). DLRS values smaller than 0.35 are describing low noise levels. A summary of all analysed single cells is provided at the end of this chapter (Table 38).

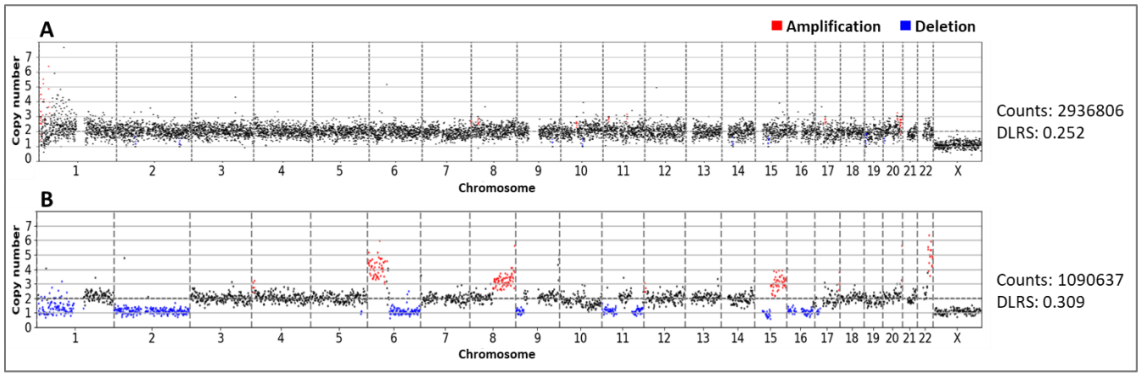


Figure 15: Representative CNA profiles of a control and a tumor cell of MM16-423
CNA profiles of a control **(A)** and a tumor **(B)** cell isolated from lymph nodes from MM16-423. Amplifications are marked with red, deletions with blue color. DLRS: derivative log ratio spread.

At first, the quality of the CNA profiles was assessed by visual analysis. The samples were categorized into four groups ranging from 0 to 3 (category 0: low; 1: sufficient; 2: good; 3: very good). Group 0 was representing cells with CNA profiles of low quality (Figure 16 D), whereas cells of category 1 to 3 were defined as cells with high-quality CNA profiles (Figure 16 A-C).

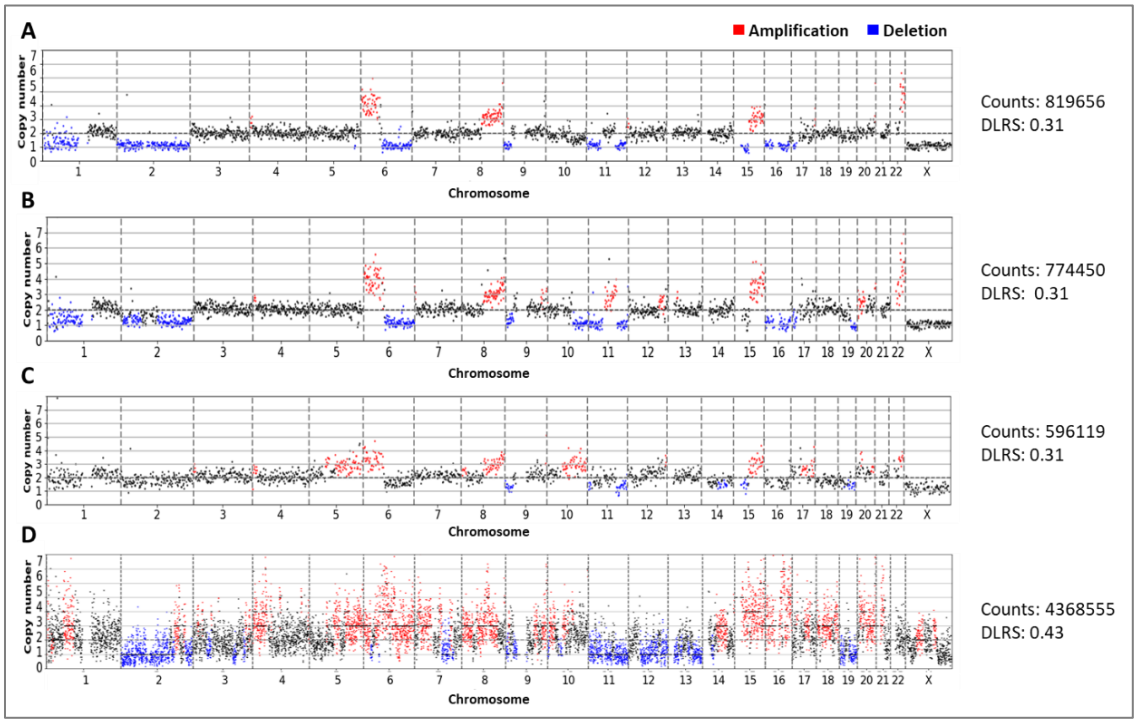


Figure 16: CNA profiles of different qualities
Representative CNA profiles of tumor cells from MM16-423 with different qualities. **A:** CNA profile of high quality: category 3; **B:** CNA profile of high quality: category 2; **C:** CNA profile of high quality: category 1; **D:** CNA profile of low quality: category 0; Amplifications are marked with red, deletions with blue color. DLRS: derivative log ratio spread.

More than 90% of all analysed single cells from each patient showed profiles of high quality (Figure 17 A). Around 70% of analysed single cells fell into category 2 having CNA profiles of good quality. Notably, 4.87% of cells from MM15-127, 5.1% of cells from MM16-412 and 8.45% of cells from MM16-423 had CNA profiles of low quality and were therefore excluded from further analyses.

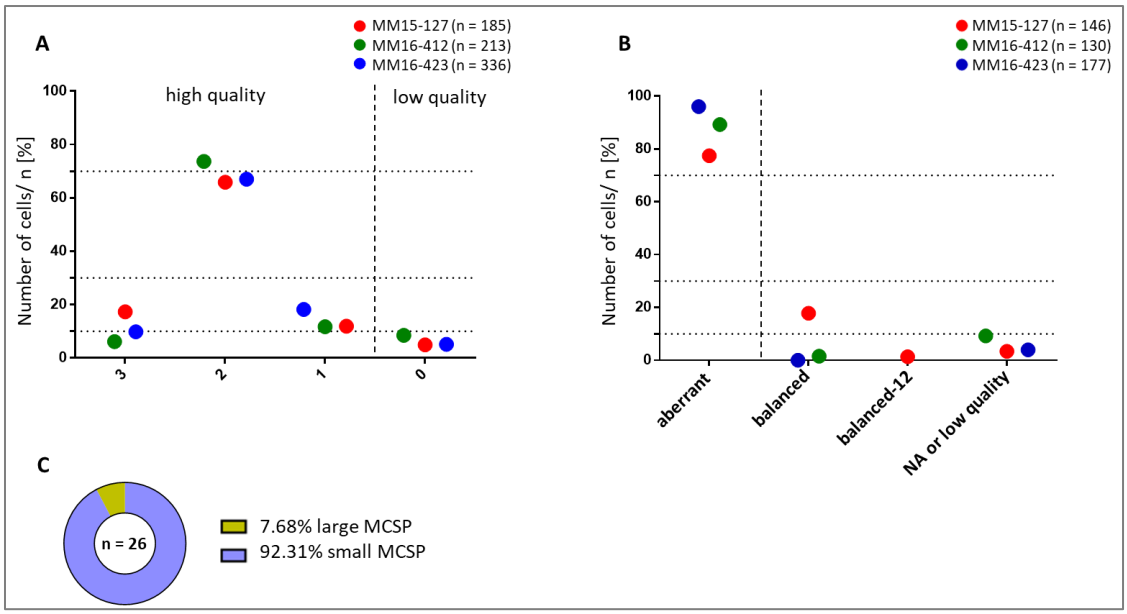


Figure 17: Quality of CNA profiles and evaluation of tumor cells
A: CNA profile quality of single cells from the three patients (different colors). Quality of CNA profiles ranged from 0 to 3 with 0=low, 1=sufficient, 2=good, 3=very good quality. Categories 1-3 were determined as high quality. Low-quality cells were excluded. **B:** CNA profiles of tumor cells from three patients (different colors). Cells were categorized into four groups: aberrant tumor cells, balanced tumor cells, tumor cells having a balanced profile besides aberrations on chromosome 12 (balanced-12) and cells whose CNA profiles were either not available (NA) or of low quality. Only aberrant tumor cells were used further. Remaining cells were excluded (separated by dotted line). **C:** Tumor cells of MM15-127 with balanced CNA profiles. Of balanced tumor cells, 7.69% [2/26] were large, 92.31% [24/26] were small MCSP-positive cells.

4.1.4.1 CNA analysis of tumor samples

CNA evaluation of single tumor cells

From the melanoma patients, 75-96% of single tumor cells had aberrant profiles with several amplifications or deletions (Figure 17 B, Figure 18 A). However, in some cases no CNAs were detected in tumor cells. This was mainly true for patient MM15-127. Of note, 26 of 146 tumor cells (17.8%) from MM15-127 had a balanced CNA profile (Figure 17 B, Figure 18 B). All these cells were isolated based on the expression of the MCSP marker. However, MCSP has been reported to be expressed on a variety of normal cells (Campoli et al. 2004). In line with the usually smaller size of non-tumor cells compared to melanoma cells, 24 of those 26 MCSP-positive tumor cells were small in size with a diameter of around 10 µm, whereas only two samples were large MCSP-positive cells with a diameter of around 20 µm (Figure 17 C). As these cells might be normal non-tumor cells due to their balanced genome, those cells were excluded from lineage analyses. The difference between small and large MCSP-positive tumor cells was further investigated in another project. In addition to MCSP-positive cells with complete balanced CNA profiles, two small MCSP-positive cells from MM15-127 showed a balanced CNA profile with the exception of an amplification of chromosome 12 named as category ‘balanced-12’ (Figure 17 B, Figure 18 C). All tumor cells with balanced, balanced-12 or low-quality profiles were excluded (Figure 17 B). Balanced-12 samples were temporarily removed from lineage analyses of this project, as the amplification on chromosome 12 was exclusively found in those two cells and not in any other tumor cell of MM15-127 (Figure 35). However, further analyses are needed to study those ‘balanced-12’ cells in more detail also including them in lineage analyses that are performed beyond this project.

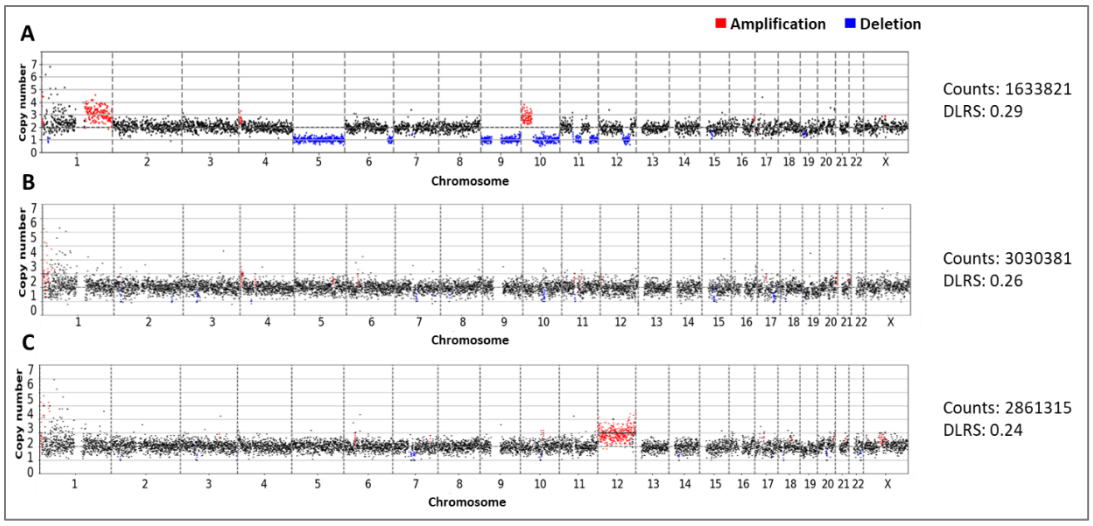


Figure 18: Representative CNA profiles of tumor cells
Representative CNA profiles of tumor cells from MM15-127 having an aberrant **(A)**, a balanced **(B)** or a balanced-12 **(C)** CNA profile. Balanced-12: balanced profile besides aberrations on chromosome 12. Amplifications are marked with red, deletions with blue color. DLRS: derivative log ratio spread.

Next, the large melanoma marker-negative cells from MM16-412 Met4, which formed tumors in mice, were analysed for CNAs as changes in the chromosomal structure may indicate tumorigenic potential. Here, all large single cells isolated from Met4 showed aberrant CNA profiles with many amplifications and deletions on several chromosomes shared with other tumor cells of this patient (Figure 19). This indicated that these were really tumor cells although they did not express any melanoma marker.

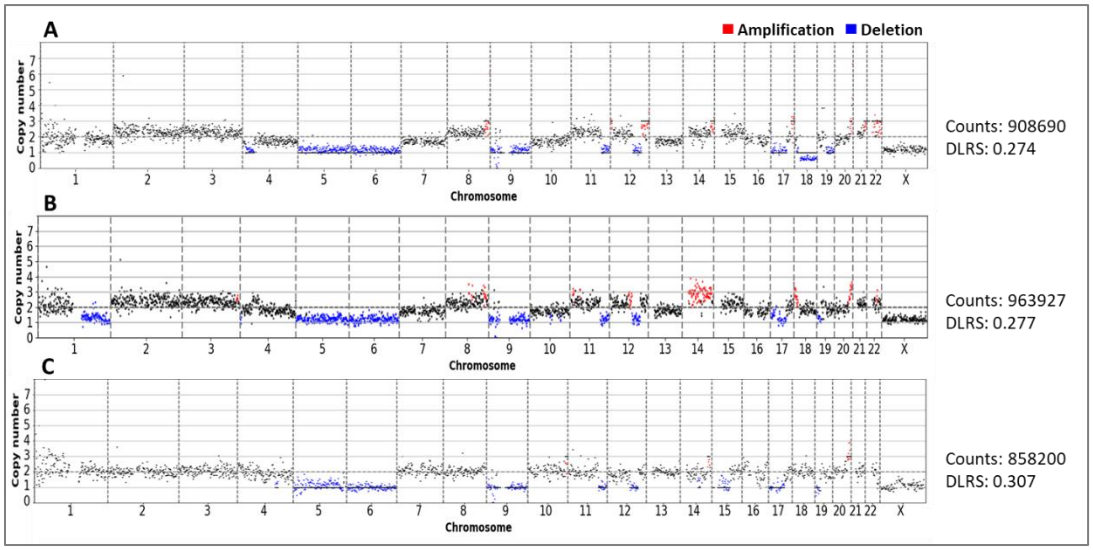


Figure 19: CNA profiles of cells from Met4 and of a tumor cell of NSLN2 of MM16-412
CNA profiles of two melanoma marker-negative cells from Met4 **(A** and **B)**, as well as a CNA profile of a tumor cells from NSLN2 of MM16-412 **(C)**. Amplifications are marked with red, deletions with blue color. DLRS: derivative log ratio spread.

CNA analysis of bulk DNA from primary tumors

Not only single tumor cells, but also bulk DNA from the primary tumors of the melanoma patients were analysed for CNAs. This was possible as WGA resulted in high-quality bulk DNA. Adding the DNA repair step did not have an influence on the CNA profile quality measured by the DLSR value (see appendix Figure 74 and Figure 75). In Figure 20 CNA profiles of the primary tumors are depicted that were generated from repaired WGA samples. The CNA profiles of DNA isolated from three different regions of MM15-127's primary tumor (PT1-3) showed a similar alteration pattern sharing amplifications on chromosome 1q, 5p, 7q, 11q, 16q and 20 and deletions on chromosome 6q, 10p and 15q. Moreover,

some alterations of the primary tumor were shared with DCCs from the SLN or skin metastases from this patient including amplifications on chromosome 1q, 5p, 16q, and 20 and deletions on chromosome 6q, 9, 10, 14 and 15 (see appendix Figure 76). For the primary tumor of MM16-412 deletions were detected on chromosome 5, 6, 9, 11q, 17p and 19 (Figure 20), which were also found in DCCs from the SLN and Met1 of this patient (see appendix Figure 77). The primary tumor of MM16-423 showed a balanced CNA profile without alterations indicating isolation of DNA from mostly non-tumor cells (Figure 20). However, for this male patient a potential gain of chromosome X was detected.

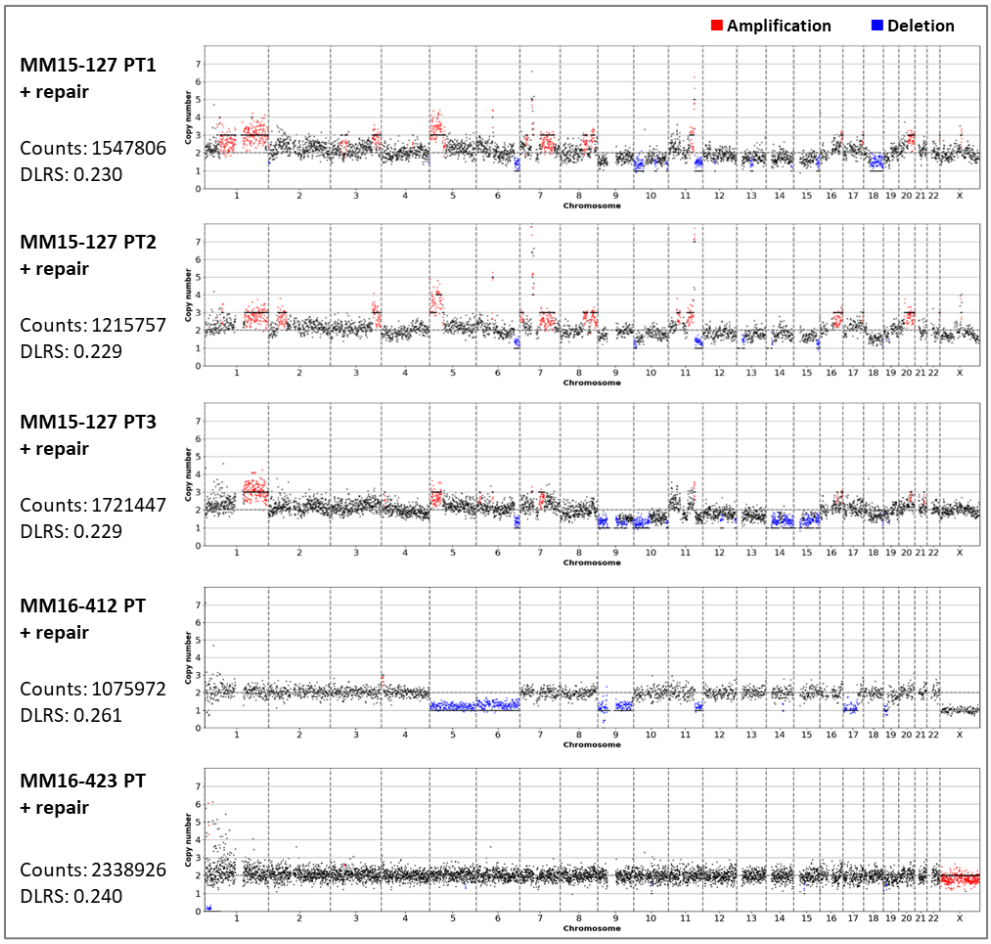


Figure 20: CNA profiles of DNA from primary tumors
CNA profiles of DNA isolated from primary tumors of MM15-127, MM16-412 and MM16-423. WGA was performed with (+) DNA repair step. Amplifications are marked with red, deletions with blue color. DLRS: derivative log ratio spread.

Next, the primary tumors were compared for the presence of non-tumor cells that could have led to potential isolation and amplification of non-tumor DNA, explaining the balanced CNA profile of MM16-423 PT. For this, sections of the primary tumors were stained for markers specific for hematopoietic cells using lymphocyte marker CD45, T cell marker CD3 and macrophage marker CD68. However, this was only possible for PTs of MM15-127 and MM16-423, as the PT block from MM16-412 was not available for the analysis anymore. For MM15-127 PT, there was a clear difference between the brown melanin pigment and the positive red staining (Figure 21, upper row). Indeed, some single hematopoietic cells were detected within the tissue. But, as 95% of the cells from the tissue were tumor cells according to the pathologist, the DNA isolated from MM15-127 PT reflected mainly tumor DNA in line with the aberrant CNA profile. The staining of sections from low-quality FFPE block of MM16-423 PT mainly showed a mixture of positive red staining and brown melanin (Figure 21, lower row). Consequently, it remained unclear whether red dots detected for MM16-423 PT were really hematopoietic cells or rather tumor cells due to their large size. Thus, immunostainings of MM16-423 PT did not help to explain the balanced CNA profile.

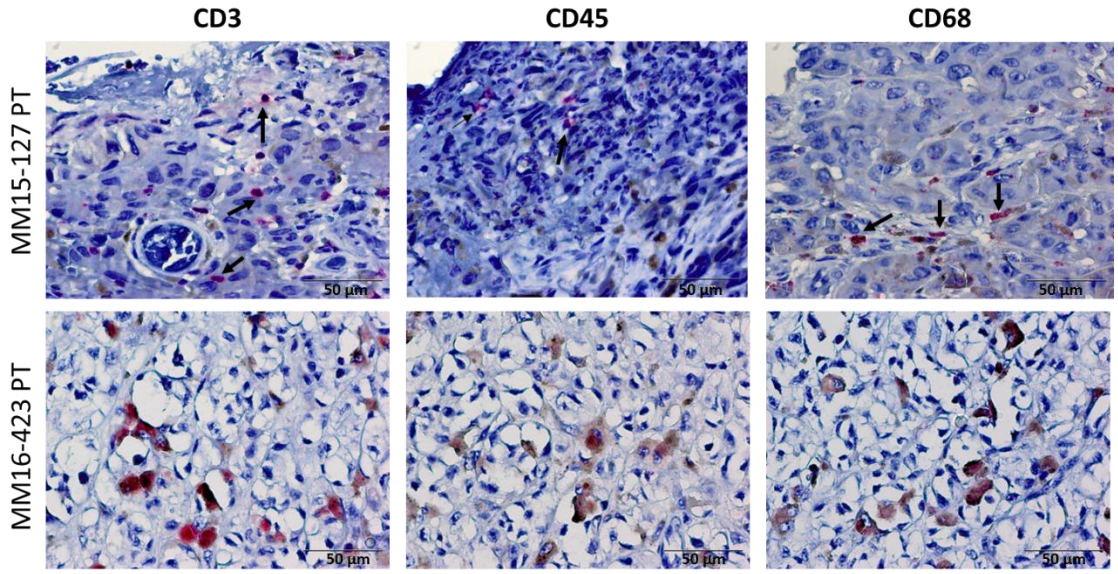


Figure 21: Immunostainings of primary tumors
Immunostainings for CD3, CD45 and CD68 of primary tumor sections from MM15-127 and MM16-423 using AP detection system resulting in a red end product. Positive stained cells are marked with arrows. PT: primary tumor.

Fingerprint analysis of primary tumor from MM16-423

In addition, as the PT DNA from male patient MM16-423 showed a potential gain of chromosome X indicating female DNA, fingerprint analysis was performed to verify the correct patient origin of the PT block. For this, bulk DNA from the primary tumor FFPE block and from an additional FFPE block consisting of only normal tissue with the same block ID were used. As separate control, DNA isolated from MM16-423's PBMCs were considered. For fingerprinting, 10 short tandem repeat (STR) loci were analysed and their detected alleles, varrying in the number of copies of the repeat sequence of those STRs, were compared to those of already 'verified' control samples of MM16-423 (DCC, PDX, cell line). Fingerprint analysis revealed that DNA isolated from the PT block did not show the same alleles like the controls (Table 37). The same was true for DNA isolated from the second FFPE block with normal tissue with the same patient ID. Moreover, fingerprint analysis was difficult to evaluate for all FFPE block samples probably due to the low block quality. However, all alleles detected in the controls from MM16-423 were found in DNA isolated from its PBMCs. Still, as from this analysis it was not 100% clear whether the PT block was really from MM16-423, PT DNA was excluded from further analysis.

Table 37: Fingerprint analysis of MM16-423 PT

	STR loci									
Sample	**TH01**	**D21S11**	**D5S818**	**D13S317**	**D7S820**	**D16S538**	**CSF1PO**	**AMEL**	**vWA**	**TPOX**
DCC	**9**	**30, 31**	**11, 12**	**14**		**11**		**X, Y**		
PDX	**9**	**30, 31**	**11, 12**	**12, 14**	**8**	**11**	**11, 12**	**X, Y**	**15, 18**	**12**
Cell line	**9**	**30, 31**	**11, 12**	**12, 14**	**8**	**11**	**11, 12**	**X, Y**	**15, 18**	**12**
PBMCs	9, 9.3	30, 31	11,12	12, 14	8, 11	11, 13	11, 12	X,Y	15, 18	11, 12
PT block WGA	6, 9, 9.3	28, 29, 30, 31.2	11,12	11, 12		9, 11, 12, 13	7, 10?	X,?	12, 13	
PT block bulk	6, 7, 9, 9.3	30, 31.2	11,12	11, 12	8, 10, 11, 12	9, 11	10, 11, 12	X,?	15, 16	8, 11
NT block WGA	6, 7, 9.3	29, 30, 31.2	11, 12, 13	9, 11		9, 11, 12, 14	10?, 11?	X, Y?	13, 20	
NT block bulk	9, 9.3	28, 29, 30, 31, 31.2	11, 12	9, 11, 12, 14	8, 11	9, 11, 12, 13	10, 11?, 12	X, Y	15, 18	7, 8, 11, 12

STR: short tandem repeat; DCC: disseminated cancer cell; PDX: patient-derived xenograft; PT: primary tumor; NT: normal tissue; WGA: whole-genome amplification. Alleles different from those detected in controls (DCC, PDX, Cell line) are marked in red. ? indicated difficulties in the evaluation.

4.1.4.2 CNA anaylsis of non-tumor control cells

Next, the CNA profiles of all control cells were evaluated (Figure 23 A). Between 57% and 87% of control cells had a balanced genome (Figure 22 A, Figure 23 A). Moreover, less than 10% of control cells showed a balanced CNA profile with small aberrations on one chromosome determined as 'balanced-Chr' (Figure 22 B, Figure 23 A). As those aberrations were very small and most likely artefacts, those cells were not excluded. However, several control cells did not have a balanced genome. Those samples either carried tumor-specific aberrations, which were aberrations shared between tumor samples of the same patient or non-tumor-specific aberrations. Control cells with tumor-specific aberrations were found for MM15-127 and MM16-412 (Figure 23 A). From MM15-127, only one CD3-positive control cell fell into this category. However, 45 of 159 control cells (28.3%) from MM16-412 showed aberrations that were characteristic for tumor cells isolated from the same lymph node (Figure 22 C and D). From these 45 control cells with tumor-specific aberrations, 7 cells were isolated as CD3-positive "T cells", whereas the remaining cells were isolated as CD68-positive "macrophages" (Figure 23 B). Hence, these cells were not hematopoietic cells, but most likely disseminated cancer cells. Remarkably, this phenomenon was patient- and lymph node-specific (MM16-412 NSLN2). As those cells harbored CNAs, they were relabeled as DCCs. In addition, 2.4 - 5.1% of control cells harbored non-tumor specific aberrations (Figure 22 E, Figure 23 A). Those control cells and control cells without profiles or with low-quality profiles were excluded from further analysis.

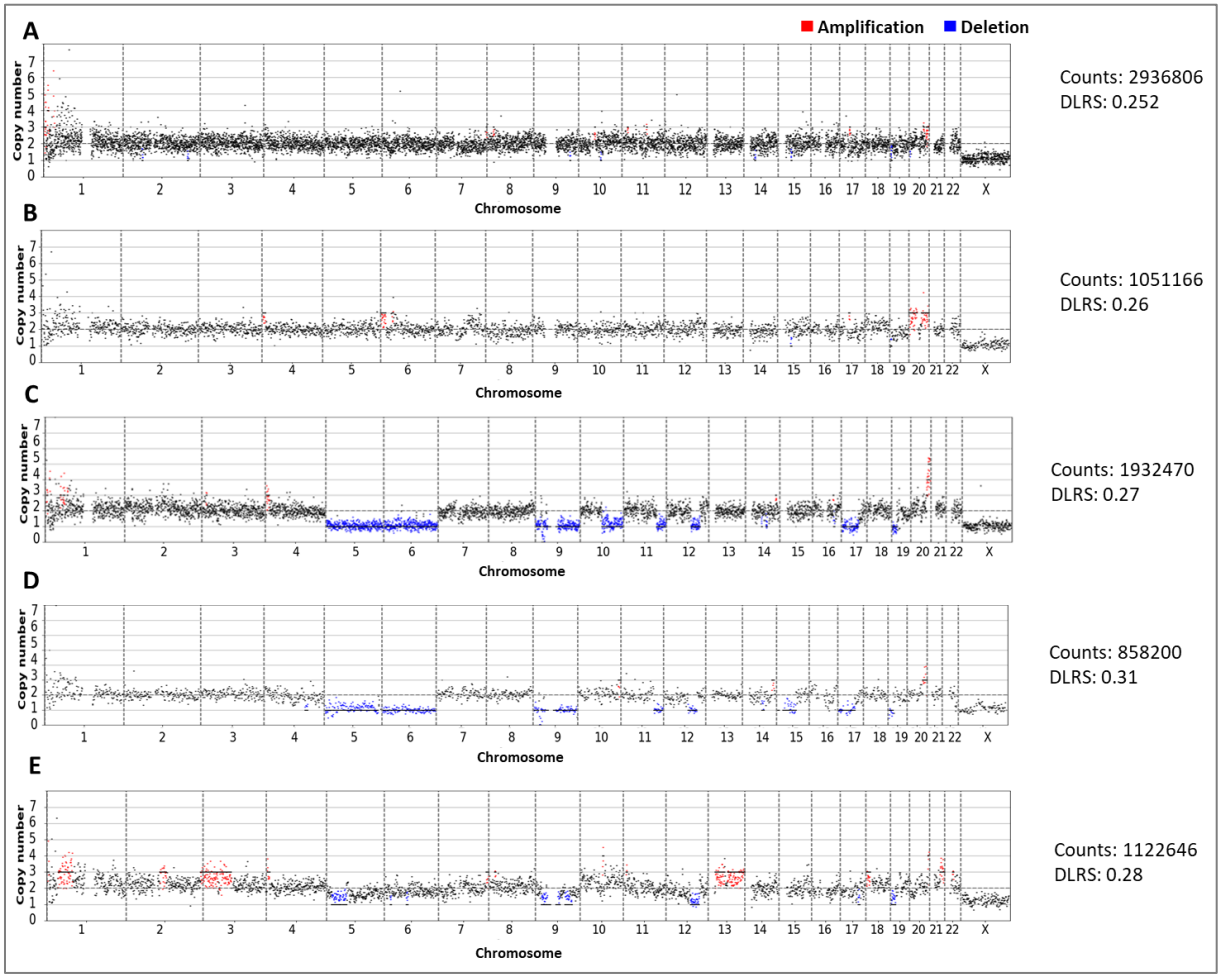


Figure 22: Representative CNA profiles of control cells
A: Control cell with a balanced profile; **B:** Control cell with a balanced profile besides small aberrations on single chromosomes (balanced-Chr). **C+E:** Control cells with aberrant profiles. **C:** Aberrant control cell with tumor-specific aberrations shared with other tumor cells of the same patient **(D)**. **E:** Control cell with non-tumor-specific aberrations. Amplifications are marked with red, deletions with blue color. DLRS: derivative log ratio spread.

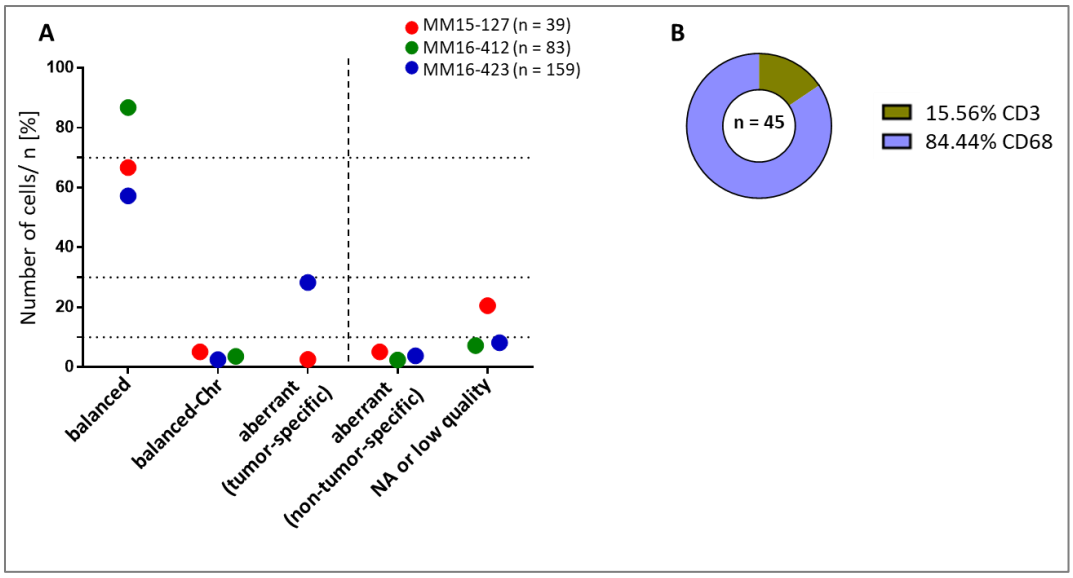


Figure 23: CNA evaluation of control cells
A: Evaluation of CNA profiles of control cells from the three patients (different colors). Control cells were categorized into five groups: balanced cells, cells having a balanced profile besides aberrations on single chromosomes (balanced-Chr), aberrant cells with either characteristic aberrations shared between tumor samples of the same patient (tumor-specfic) or non-tumor-specific aberrations and cells whose CNA profiles were either not available (NA) or of low quality. Control cells with tumor-specific aberrations were relabelled as tumor cells. Control cells with non-tumor-specific aberrations, without profiles (NA) or with low-quality profiles were excluded (separated by dotted line). **B:** Control cells of MM16-412 with tumor-specfic aberrations. Of control cells with tumor-specific aberrations, 15.56% [7/45] were positive for CD3, 84.44% [38/45] were positive for CD68 isolated from the same lymph node.

To summarize, tumor cells without clear aberrations, control cells with non-tumor-specific aberrations and single cells with low-quality profiles were exclduded from lineage reconstructions (Table 38). Control cells with tumor-specific aberrations characteristic for tumor cells of this patient were relabeled as tumor cells. These results pointed to the importance of additionally checking for CNAs of marker-based isolated cells.

Table 38: Summary of CNA analysis of single cells from melanoma patients

	MM15-127	MM16-412	MM16-423	Status
Number of single cells	**185**	**336**	**213**	
High quality	95.13% [176/185]	94.9% [319/336]	91.55% [195/213]	Included
Low quality	4.87% [9/185]	5.1% [17/336]	8.45% [18/213]	Excluded
Number of tumor cells	**146**	**177**	**130**	
Aberrant	77.40% [113/146]	96.05% [170/177]	89.23% [116/130]	Included
Balanced	17.81% [26/146]	0% [0/177]	1.54% [2/130]	Excluded
Balanced-12	1.37% [2/146]	0% [0/177]	0% [0/130]	Excluded
NA or low quality	3.42% [5/146]	3.95% [7/177]	9.23% [12/130]	Excluded
Number of control cells	**39**	**159**	**83**	
Balanced	66.67% [26/39]	57.23% [91/159]	86.75% [72/83]	Included
Balanced-Chr	5.13% [2/39]	2.52% [4/159]	3.61% [3/83]	Included
Aberrant (tumor-specific)	2.56% [1/39]	28.3% [45/159]	0% [0/83]	Relabeling
Aberrant (non-tumor-specific)	5.13% [2/39]	3.77% [6/159]	2.41% [2/83]	Excluded
NA or low quality	20.51% [8/39]	8.18% [13/159]	7.22% [6/83]	Excluded

Note: For MM16-412 and MM16-423 not all single cells with high quality (390 for MM16-412 and 217 for MM16-423, Table 36) were analysed for CNAs, as they did not pass first microsatellite-based quality control.

4.2 Phylogenetic relationship of disseminated melanoma cells

To determine the phylogenetic relationship between DCCs present at different anatomic sites and disease stages, several approaches were employed. The first approach was performed in collaboration with Ehud Shapiro from the Weizmann Institute of Science in Israel and was based on the analysis of microsatellites (MS). The second approach in this study used copy number alterations (CNAs) for analysing the phylogeny of melanoma cells. For one *BRAF* mutated patient, a third approach was applied using the *BRAF* mutation status of disseminated cancer cells. In contrast to non-coding MS mutations, which are not influenced by selective pressures, copy number alterations and somatic mutations can lead to survival benefit of cancer cells and thus are under selection pressure.

4.2.1 Lineage tree analysis based on microsatellites

Regarding the general workflow for sequencing of microsatellites (MS-Seq), original DNA samples were prepared as described in the following section and shipped to the Weizmann Institute for target enrichment, library preparation and sequencing (Figure 24). All samples were first subjected to a quantification sequencing run on a MiSeq instrument allowing the preparation of balanced sequencing pools. These pools were subsequently sequenced on a HiSeq or NextSeq instrument. In the end, data were analysed bioinformatically and lineage trees were reconstructed. In the next chapters, optimization steps that were performed during the project are described starting with optimizing the protocol for sample preparation.

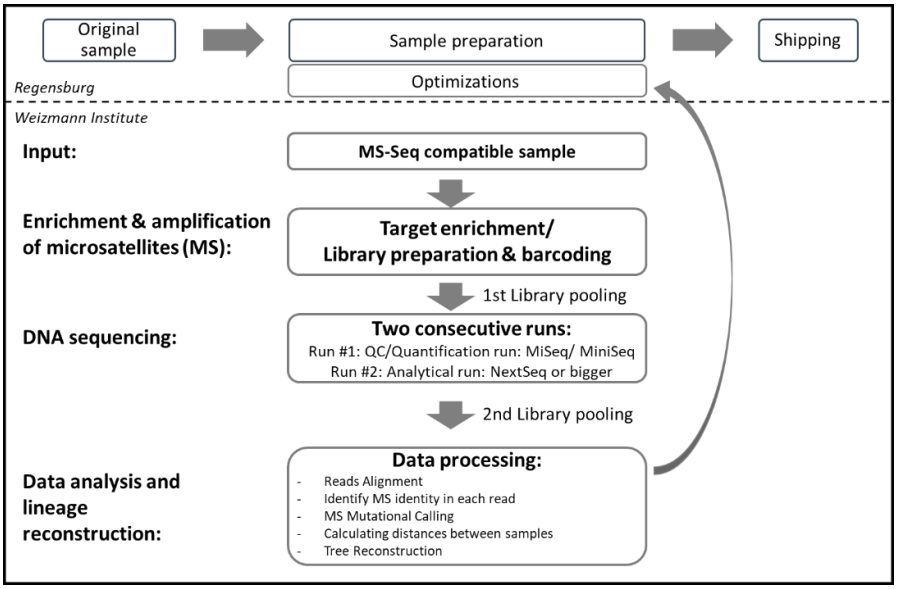


Figure 24: Workflow of sample and library preparation for MS-Seq based lineage reconstruction
Quality controlled DNA was obtained from original samples and prepared in Regensburg. After shipping to the Weizmann Institute the samples were used for target enrichment and library preparation before sequencing. Following data processing, lineage trees were reconstructed. After data evaluation, protocol optimizations were performed in Regensburg and samples were sent again for MS marker sequencing. MS: microsatellites.

4.2.1.1 Protocol optimization for sample preparation

At the beginning, original WGA samples subjected to double-strand synthesis (WGA + DSS) were used. Here, samples showed mapping rates (proportion of reads mapping to the targeted MS markers of interest) of almost 90% revealing compatibility of WGA samples subjected to DSS with the MS-Seq protocol of the collaborators (Figure 25). However, as the collaborators changed their MS-Seq library protocol during the project to increase the throughput, first experiments exhibited low mapping rates

(40-70%) of WGA samples subjected to DSS (Figure 25). Additionally, re-amplified WGAs (WGA + REA) instead of original WGAs were considered for MS-Seq analysis in order to save material. Thus, the protocol for sample preparation had to be re-evaluated and optimized. In a joined experiment together with Manjusha Ghosh and Julia Greindl-Junghans, different parameters were tested on 20 single control cells with high-quality DNA. In addition to samples of the three melanoma patients, control cells from a breast cancer patient of Manjusha Ghosh's project were included in the analysis. For this, several preparation steps possibly interfering with the MS-Seq downstream application were considered, like adding a step of double strand synthesis (DSS) after WGA of original or re-amplified WGA samples (REAs) to complete thereafter present single strand products. This combinations were further tested with or without DNA purification (XP) to remove reagents like primers or adaptors. The highest mapping rates of samples with the new library preparation protocol were found for re-amplified WGA products followed by DNA purification independent of performing DSS (Figure 25, WGA+REA+DSS+XP and WGA+REA+XP). These protocol combinations showed even higher mapping rates than samples prepared with the old library preaparation protocol highlighting the importance of a purification step. As performing DSS after re-amplification did not show any improvements, all single cells with high-quality DNA were re-amplified and subjected to DNA purification prior to shipping.

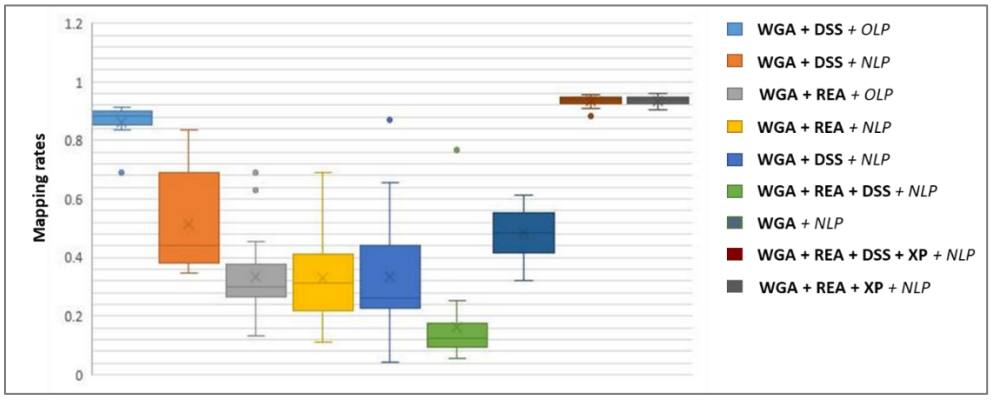


Figure 25: Mapping rates of samples prepared with different protocols
Samples with different preparation protocols indicated by colors were compared with regard to their mapping rates. Libraries were prepared either with the old library protocol (OLP) or with the new libraray protocol (NLP) developed by the Weizmann group. Samples of the three melanoma and one breast cancer patients were prepared together with Julia Greindl-Junghans and Manjusha Ghosh before sending to the collaboration partners for library preparation. This figure was provided by the collaboration partners from the Weizmann Institute. WGA: whole genome amplification; DSS: double strand synthesis; REA: re-amplification; XP: purification using Ampure XP beads; lib prep: library preparation.

In addition to single cells also bulk DNA from primary tumors and germline samples were used for MS-based analysis. As the sample preparation protocol was optimized for DNA from single cells and as there was no information of its compatibility with bulk samples, bulk DNA samples were prepared with different protocols and sent for analysis. In addition to unamplified DNA, also amplified bulk DNA from primary tumors was used testing again the relevance of the additional DNA repair step. WGA samples were directly purified or purified after additional re-amplification. For germline samples, bulk DNA was either left unamplified or was amplified using the WGA protocol with 1 or 10 ng as input followed by re-amplification and DNA purification.

4.2.1.2 Separation of non-tumor cells from different individuals

Using the optimized sample preparation protocol, the resolution of the method was tested asking whether control cells from different individuals could reliably be separated. For this, T cells from four different cancer patients were considered. Here, in addition to three melanoma patients, also T cells from a breast cancer patient of Manjusha Ghosh's project were included to increase the sample size. For the phylogenetic tree reconstruction, the Triplet MaxCut (TMC) algorithm was used putting all cells

at an equal depth (Sevillya et al. 2016). Therefore, the depth of the tree was additionally calculated comparing the number of microsatellite mutations from the patients with an *ex vivo* tree of the DU-145 cancer cell line (Biezuner et al. 2016). Thus, the depth of the lineage tree is not an absolute number but reflects the relative depth. However, the depth of samples and with this the number of cell divisions can not be compared between different individuals, but can only be analysed within the same patient. As this experiment was conducted at the beginning of the project, the cells were sequenced with one million reads per sample and not analysed for CNAs. However, even with this sequencing depth without controlling for CNAs a perfect separation between T cells from different individuals was achieved (Figure 26).

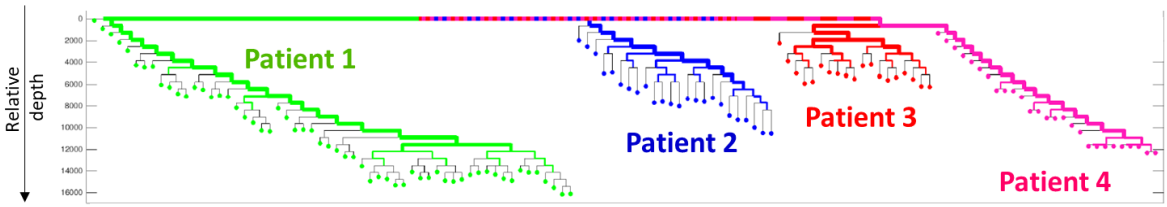


Figure 26: Cell lineage reconstruction of T cells from four different cancer patients
Each dot represents a T cell isolated from four cancer patients differing in their color. Green: T cells from patient MM16-412; blue: T cells from a breast cancer patient; red: T cells from a different melanoma patient not included in this project; pink: T cells from MM16-423. The tree was reconstructed using Triplet MaxCut (TMC) algorithm with around 1x 10^6 reads per sample. Y-axis represents the relative depth. The relative depth between different patients can not be compared. This figure was also shown in the thesis of Julia Greindl-Junghans.

4.2.1.3 Influence of sequencing depth on the tree accuracy

In the first sequencing round, a median success rate of 85% was achieved for samples that had passed the initial quality control for MS-library preparation (MS-QC). Around one million total reads and between 0.71 and 0.99 million successful reads mapping to the human genome were obtained per sample and patient (Table 39, Figure 27 A). However, from simulation studies of the cooperation partners it was deduced that at least three million reads per sample were needed to retrieve an accurate tree separating different cells from one indiviual (Figure 27 B). Therefore, a second sequencing round was performend using the NovaSeq technology achieving in median between 3.60 and 5.19 million total reads per sample and patient (Table 39, Figure 27 A). With this, a 1.5-fold increase of the number of loci covered by more than 10 reads was achieved for all patients (Table 39, Figure 27 C). Thus, the higher the number of total reads, the higher the number of loci were covered with more than 10 reads reaching a plateau at around four million total reads (Figure 27 D and F, brown line). After the first sequencing round, in median 5700 loci and after the second sequencing round about 8700 loci were covered by more than 10 reads for all three patients (Figure 27 E and F). A further increase in sequencing depth would have not brought profit compared to the accompanying costs.

Table 39: Statistics of sequencing parameters

	MM15-127		MM16-412		MM16-423	
Sequencing round	1	2	1	2	1	2
Number of samples	120	246	371	284	208	196
Total reads [x 10^6]	1.10	3.60	1.14	5.19	0.91	3.74
Fold increase		3.27		4.55		4.11
Successful reads [x 10^6]	0.84	2.94	0.99	4.47	0.71	3.12
Fold increase		3.5		4.51		4.39
Loci covered by > 10 reads	5735	8573	6228	8807	5341	8739
Fold increase		1.49		1.41		1.64

Note: Number of samples differ in sequencing rounds. Some samples were added after first sequencing round (for MM15-127) or some samples dropped out from the analysis (MM16-412; MM16-423). The median of analysed parameters is shown.

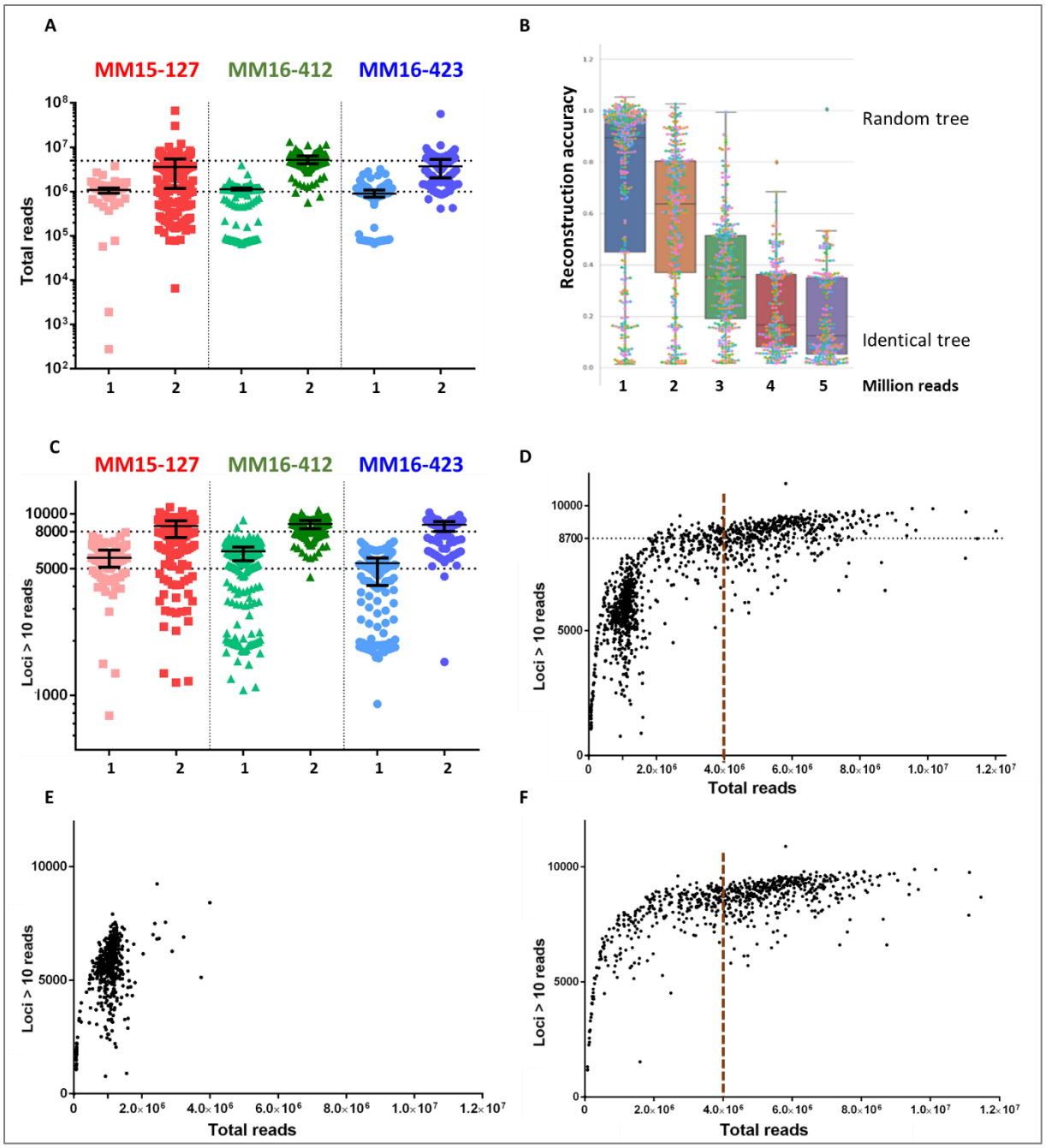


Figure 27: Sequencing depth and reconstruction accuracy
A: Number of total reads. Total reads after the first round of sequencing using HiSeq/NextSeq (1) and after the second round of sequencing using NovaSeq (2) are depicted for each patient. Dotted lines at 1 and 4x 10^6 total reads. **B:** Simulation study. The tree reconstruction accuracy depended on the number of successful reads. Values above 0.5 showed a random tree reconstruction. The smaller the value, the more accurate the tree reconstruction got. The figure was provided by the collaboration partners. **C:** Number of loci covered with more the 10 reads. Number of loci covered with more than 10 reads after first (1) and second (2) sequencing round for each patient. **D** to **F:** Total reads versus loci covered with more than 10 reads. Total reads are plotted against loci covered with more than 10 reads for all samples after both sequencing rounds **(D)**, after first **(E)** and second **(F)** sequencing round. Brown line describes the number of total reads reaching a plateau at around 8700 (dotted line in **D**). In **A** and **C** median with interquartile range is shown. In **A**, **C**, **D**, **E**, **F** one dot represents one sample.

4.2.1.4 Summary of sample selection criteria

Lineage tree analyses were performed from in summary 624 samples including 578 single cells, 19 primary tumor bulk samples and 27 germline bulk samples for the three melanoma patients (Figure 28). The samples were selected if they showed high-quality DNA after WGA, had confirmed CNA profiles and showed sufficient MS coverage by passing quality controls of MS sequencing performed by the collaborators (MS-QCs).

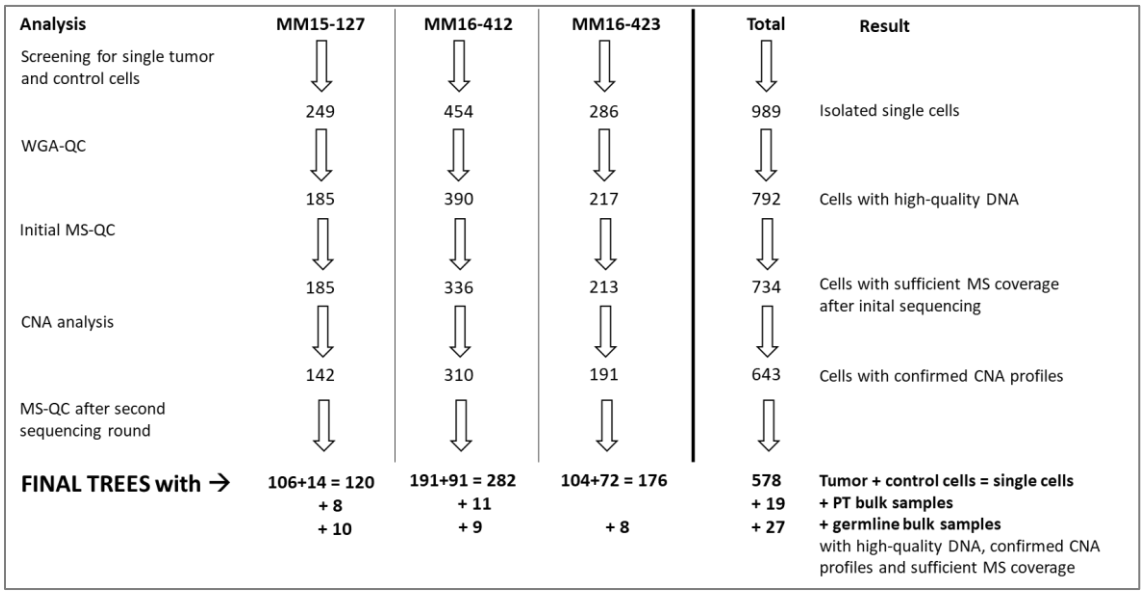


Figure 28: Summary of sample selection criteria for tree reconstructions
Samples were selected if the showed high-quality DNA after WGA, had confirmed CNA profiles and passed quality controls of microsatellite sequencing performed by the collaborators (MS-QCs) showing sufficient MS coverage.

4.2.1.5 Separation of control cells of one patient

First, to see whether non-tumor cells from different lineages could be separated within one patient, trees were generated only with control samples using a threshold of 10 reads per locus. In addition to hematopoietic cells, oral epithelial cells and endothelial cells, bulk samples from peripheral blood or lymph node cells representing germline DNA were used for root approximations. Only amplified germline bulk samples passed MS-QCs (Figure 29 A). Amplified samples did not show a significant difference in the number of loci covered with more than 10 reads when using 1 or 10 ng of DNA as WGA input (Figure 29 B).

A

	# Input samples	# Samples passing MS-QCs	Success rate
Unamplified (no WGA)	**5**	**0**	**0% [0/5]**
1 ng WGA + REA (all)	**15**	**13**	**87% [13/15]**
MM15-127	5	5	100% [5/5]
MM16-412	5	5	100% [5/5]
MM16-423	5	3	60% [3/5]
10 ng WGA + REA (all)	**15**	**14**	**93% [14/15]**
MM15-127	5	5	100% [5/5]
MM16-412	5	4	80% [4/5]
MM16-423	5	5	100% [5/5]

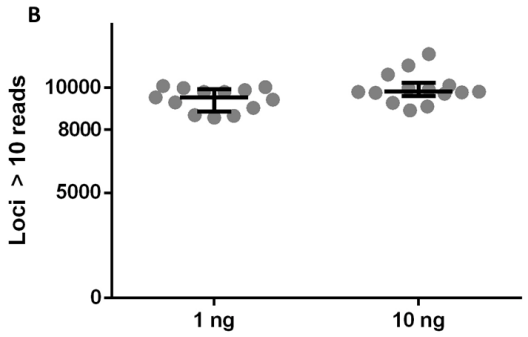


Figure 29: Analysis of germline bulk samples of all melanoma patients
A: Overview of analysed germline samples. All amplified samples were purified prior to library preparation. MS-QC: quality control of microsatellte sequencing performed by collaborators; #: number; WGA: whole genome amplification; REA: re-amplification. **B:** Number of loci covered with more the 10 reads of germline samples. Number of loci covered with more than 10 reads after final lineage tree reconstruction for germline samples using either 1 or 10 ng as input DNA. Median with interquartile range is shown.

In the meantime, the collaboration partners improved the reconstruction algorithm replacing the TMC algorithm for the FastTree2 algorithm, an approach allowing the inference of maximum-likelihood phylogenies (Price et al. 2010). As all lineage trees generated with the FastTree2 algorithm were rooted against the median genotype of germline samples, the Y-axis indicated the relative depth, which is proportional to the number of cell divisions. Bootstrapping analysis was performed to validate the robustness of the clusters and to see whether a particular branching was significant (Lemoine et al. 2018): after iteratively resampling with replacement from the original dataset, replications of the original estimates were provided enabling the calculation of the variance and distribution of the estimate. Here, 1000 iterations were used to evaluate whether the proximity of samples within a certain branch was essentially closer than within a randomly sampled branch. Next, the transfer bootstrapping expectation (TBE) index was determined ranging from 0 to 1 (Lemoine et al. 2018). A TBE of 0 indicated a random clustering within a branch, whereas a TBE of 1 meant that clustering of samples was present in all bootstrap trees. Usually a threshold of 70% is used for a true branch (Hillis and Bull 1993). In this study, branches supported with a TBE larger than 70% were marked with (*) indicating that corresponding clades were robust. The following tree reconstructions used a threshold of 10 reads per locus.

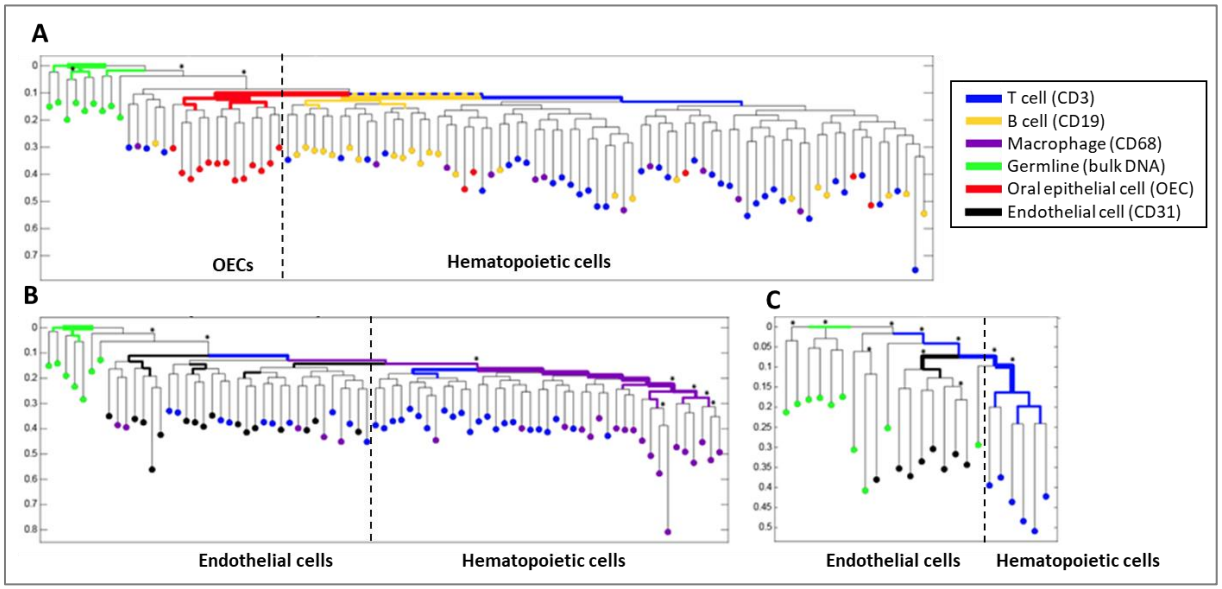


Figure 30: Cell lineage reconstructions of control samples for each patient
Lineage trees of non-tumor samples of MM16-412 **(A)**, MM16-423 **(B)** and MM15-127 **(C)**. Samples were sequenced at around 4x 10^6 reads represented by single dots. The trees were reconstructed using the FastTree2 algorithm and were rooted at the median genotype of the germline samples. Edges reproduced with significant support (TBE>70%) in bootstrap analysis were marked with *. Y-axis represents the relative depth. A clear separation of a cluster containing either oral epithelial cells (OECs) or endothelial cells from the hematopoietic cells can be seen. The figure was also shown in (Tao et al. 2020).

The reconstructed lineage tree of MM16-412's control samples (Figure 30 A) showed a clear separation of a cluster containing most of the epidermal stem cell-derived oral epithelial cells (OECs) from the hematopoietic cells deriving from the mesoderm. This showed that in line with the current understanding of embryonic development cells from different embryonic germ layers (mesoderm and ectoderm) could be separated. In addition, a separation of hematopoietic and endothelial cells was seen for patient MM15-127 and MM16-423 (Figure 30 B and C). For patient MM16-423 a clear separation of T cells and macrophages even within the hematopoietic cluster was found. In summary, clustering of hematopoietic cells versus either oral epithelial cells or versus endothelial cells was obtained for all patients reflecting the early developmental separation of these lineages.

4.2.1.6 Cell lineage reconstructions of melanoma patients

As T cells from different individuals and control cells from different lineages within one patient could be separated, the next step was performed adding single cancer cells and bulk samples from primary tumors (PTs) to the lineage reconstructions (Table 40).

Table 40: Summary of high-quality samples for final lineage tree reconstructions

	MM15-127	MM16-412	MM16-423
Primary tumor	8	11	--
Germline	10	9	8
Single cells	120	282	176
Sum of samples	**138**	**302**	**184**

PT bulk samples were prepared with different protocols, as there was no information on the compatibility of bulk samples with the new library preparation protocol of the collaborators. Here, the aim was to see whether there was a difference in success rates or sequencing parameters and whether replicates clustered together in the trees serving as an internal control of the method. From initial 44 PT bulk samples of two melanoma patients only 19 samples passed the MS-QCs (Figure 31 A). The highest success rates (100%) were seen for samples subjected to WGA without re-amplification independent on an additional repair step. Then, differences between the groups regarding their number of loci covered with more than 10 reads were analysed (Figure 31 B). The highest number of loci covered with more than 10 reads were found for WGA samples repaired before library preparation. A significant difference between the groups was only seen between repaired amplified samples (WGA+repair) and re-amplified WGA samples (WGA+REA). Nevertheless, despite different sample preparations, all primary tumor samples passing MS-QCs clustered together in final lineage reconstructions of the individual patients (Figure 33, Figure 36). These findings indicated that the MS-Seq method yielded reliable results as technical replicates were found within the same cluster.

A

	# Input samples	# Samples passing MS-QCs	Success rate
Unamplified (no WGA)	6	0	0% [0/6]
Unamplified (no WGA) + repair	2	1	50% [1/2]
WGA	3	3	100% [3/3]
WGA + repair	3	3	100% [3/3]
WGA + REA	15	4	27% [4/15]
WGA + REA + repair	15	8	53% [8/15]
Sum of samples	**44**	**19**	**43% [19/44]**

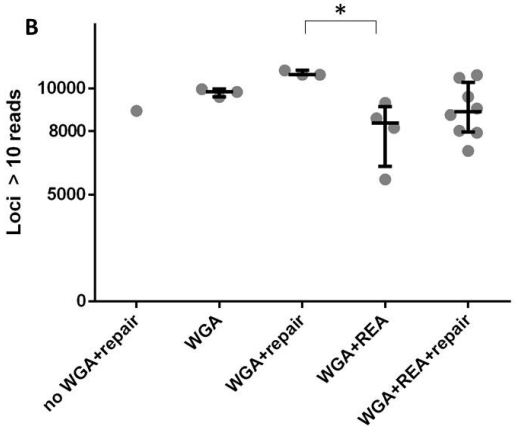


Figure 31: Analysis of primary tumor bulk samples of all melanoma patients
A: Overview of analysed PT samples. All amplified samples were purified prior to library preparation. PT: primary tumor; MS-QC: quality control of microsatellite sequencing performed by collaborators; #: number; WGA: whole genome amplification; REA: re-amplification. **B:** Number of loci covered with more the 10 reads of PT samples after final lineage tree reconstruction. In total, 12472 MS loci were analysed. * indicated statistically significant difference using one-way ANOVA and Tukey's multiple comparisons test. Median with interquartile range is shown.

After bioinformatic analyses, lineage trees were generated with 138 samples for MM15-127, 302 samples for MM16-412 and 184 samples for MM16-423 using a threshold of 10 reads per locus. In median between 4.01 to 5.10 million total reads were obtained for analysed samples of the three melanoma patients (Figure 32 A). With achieved sequencing depth, in median more than 77% (9600/ 12472) of all sequenced loci were covered by more than 10 reads for each sample (Figure 32 B).

A

	MM15-127	MM16-412	MM16-423
Number of samples	138	302	184
Total reads [x 10^6]	4.55	5.10	4.01
Loci covered by > 10 reads	9653	9675	9678

B

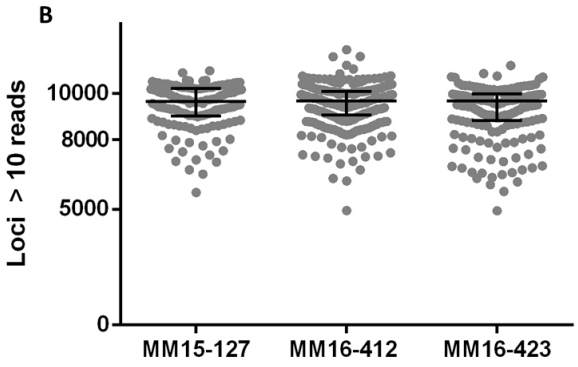


Figure 32: Sequencing parameters of samples for final lineage trees
A: Table of sequencing parameters of samples from each patient. Median of each analysed parameter is shown. **B:** Number of loci covered with more the 10 reads of samples from the melanoma patients after final lineage tree reconstructions. One dot represents one sample from the corresponding patient. In total, 12472 MS loci were analysed. Median with interquartile range is shown.

4.2.1.6.1 Phylogenetic analysis of MM15-127

At first, the cell lineage reconstruction of patient MM15-127 was evaluated comprising 138 samples. The tree separated into a control and a tumor cell cluster (Figure 33). However, there were still four SLN-DCCs found within the control cell cluster and seven control cells within the tumor cell cluster (Figure 33 and Figure 44). Strikingly, all seven control cells were T cells isolated from the SLN and were located between SLN-DCCs and primary tumor samples (Figure 33). However, the evaluation of the CNA profiles of the T cells did not help to understand their location as all T cells showed balanced genomes (Figure 34). Thus, more analyses are needed to explain the potential mislocation of the T cell cluster within the lineage tree. The tumor cell cluster could further be separated into five groups. The first group branching off included DCCs isolated from the SLN and cells from the DCC-derived cell line. Next, one could find a big branch containing cells isolated from two cutaneous metastases. The third group consisted of SLN-DCCs and cell line cells like the first group of the tumor cluster. At the end of the tree, the T cells cluster as well as the primary tumor cluster could be found. The tree was reliably reflecting the ancestry of cell line cells from SLN-DCCs as they were found within the same groups. As the first DCC cluster branched off from the metastases branch, SLN-DCCs contained therein, shared more MS loci with cells from skin metastases than for example the primary tumor. The tree additionally revealed that PT samples showed more depth than other tumor cells.

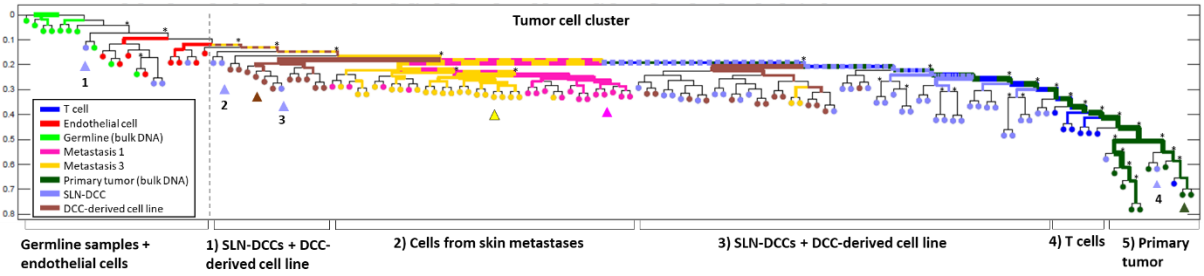


Figure 33: Cell lineage reconstruction of MM15-127
The phylogenetic tree of MM15-127 included control and tumor cells as well as bulk DNA from germline samples and the primary tumor. All samples were sequenced in median with 4.55x 10^6 reads. The tree was reconstructed using the FastTree2 algorithm and was rooted at the median genotype of the germline samples. Edges reproduced with significant support (TBE>70%) in bootstrap analysis were marked with *. Y-axis represents the relative depth. Cells whose CNA profiles are shown in Figure 35 are marked with ▲.

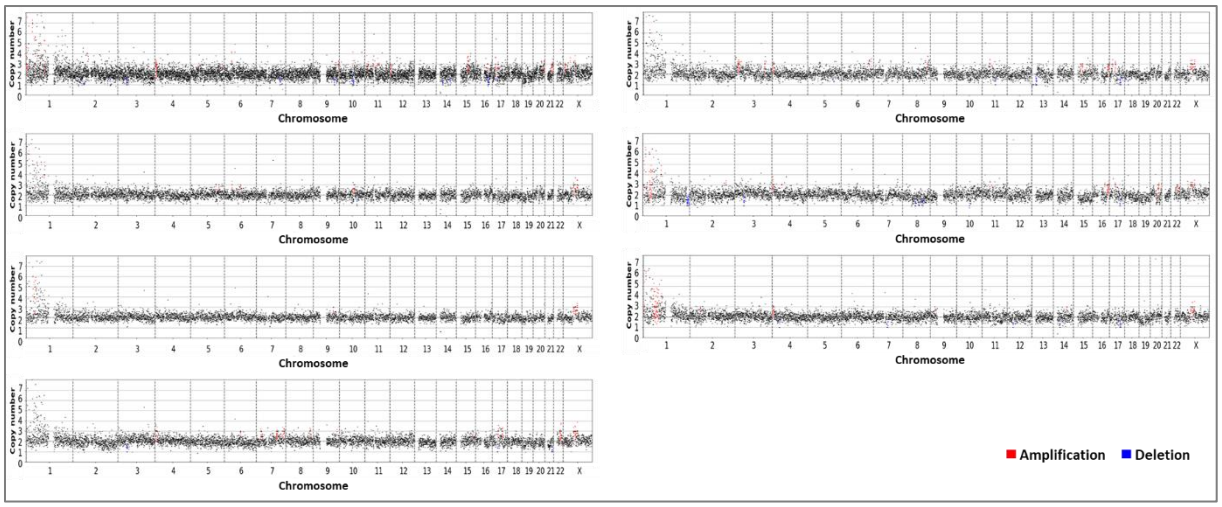


Figure 34: CNA profiles of T cells from MM15-127
Balanced CNA profiles of T cells located within the tumor cell cluster. Amplifications are marked in red, deletions in blue.

As the project focussed on the role of SLN-DCCs in the context of metastatic seeding, the locations of SLN-DCCs within the tree were analysed in more detail. To illuminate their part within the phylogeny of the tumor cells, the CNA profiles of the corresponding cells were evaluated in addition (Figure 35). One could see that SLN-DCCs were distributed among the whole tree and shared CNAs on chromosome 6, 9, 10, 11 and 12 with each other (Figure 35). As SLN-DCCs were found in nearly every branch of the tree, this indicated that SLN-DCCs indeed might represent the whole spectrum of the tumor burden of the patient. Going more into detail, some SLN-DCCs were found within the control cell cluster and within small subbranches splitting from the tumor cluster right at the beginning (Figure 33). Interestingly, those early SLN-DCCs (represented by SLN-DCC1 and 2) shared a characterstic deletion on chromosome 5 that was not found in the remaining tumor cells (Figure 35). The remaining SLN-DCCs (representad by SLN-DCC3 and 4) were mostly found together with cell line cells sharing the same aberrations on chromosome 6, 9, 10, 11 and 12 (Figure 35). With cells from cutaneous metastases, SLN-DCCs only shared deletions on chromosome 6, 9 and 10. Deletions of chromosome 11 and 12 were not seen in cells from skin metastases (Figure 35). However, cells from skin metastases carried deletions on chromosome 14 and 15 indicating that those cells progressed (Figure 35). Two SLN-DCCs were also found within the primary tumor branch sharing again deletions on chromosome 6, 9, 10 and 11 (Figure 33 and Figure 35). Notably, the primary tumor DNA shared more copy number alterations (CNAs) with cells from cutaneous metastases including deletions on chromosome 14 and 15 (Figure 35) although they were clearly separated in the lineage tree. Thus, the evaluation of CNA data of single cells could not explain the position of those cells or cell groups within the tree. This further highlighted the importance of generating a CNA-based tree as macroscopic evaluation of single CNA profiles did not allow clear statements when cells phylogenetically separated from each other. The CNA data were rather helpful to assess the plausibility of MS-based trees and to point towards potential mislocations, like for example the position of SLN-DCC1 within control cells although showing clear CNAs.

In general, the MS-based analysis showed a separation of tumor from control samples and a clustering of cells from different tumor tissues. The plausibility of the MS-based analysis with disease course and sample acquisition was shown as cells from both metastases nicely clustered together as well as SLN-DCCs and SLN-DCC derived cell line cells. However, some SLN-DCCs clustered together with control cells and T cells were located within the tumor branch next to the primary tumor, which had an increased depth. The evaluation of CNA profiles could not help explaining those findings. Nevertheless, SLN-DCCs were found within all clusters of the phylogenetic lineage tree suggesting that they might enable therapeutic target identification.

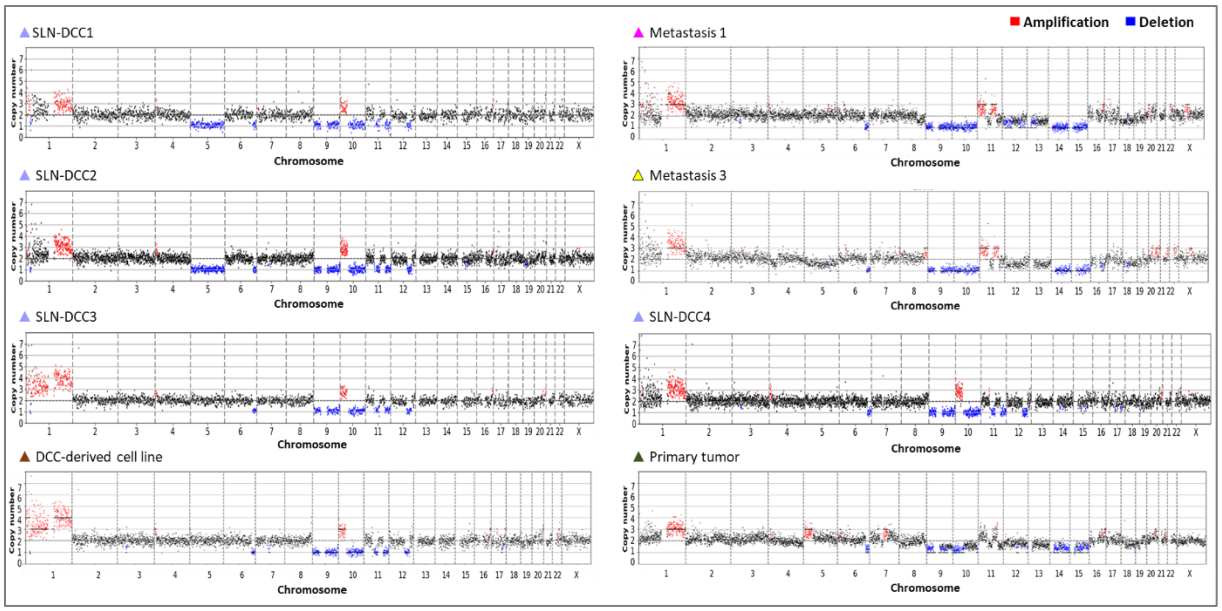


Figure 35: CNA profiles of selected tumor cells from MM15-127's lineage tree
CNA profiles of different tumor cell types from MM15-127. Amplifications are marked with red, deletions with blue color.

4.2.1.6.2 Phylogenetic analysis of MM16-412

The lineage tree of MM16-412 comprised 302 samples and could be separarted into a control cell cluster and a tumor cell cluster (Figure 36). As previously described, some CD3- and CD68-positive cells isolated from the non-sentinel lymph node 2 (NSLN2) showed tumor-specific aberrations and were thus relabeled as NSLN-DCCs (chapter 4.1.4.2). Within the lineage tree, those cells were marked as 'CD3/ CD68 cancer-like cells' and colored like NSLN-DCCs. All CD68 cancer-like cells (28/28) were indeed found within the tumor cell cluster (Figure 36) showing reliability of MS-based method. Moreover, 97% of control samples [96/99] were located within the control cell cluster and technical replicates from PT bulk samples nicely clustered together. However, there were still some mislocated cells (Figure 44). Three from seven CD3 cancer-like cells were located within the control cluster and three T cells isolated from NSLN2 and metastasis 1 with balanced genomes were detected within the tumor cell cluster (Figure 36 and Figure 37).

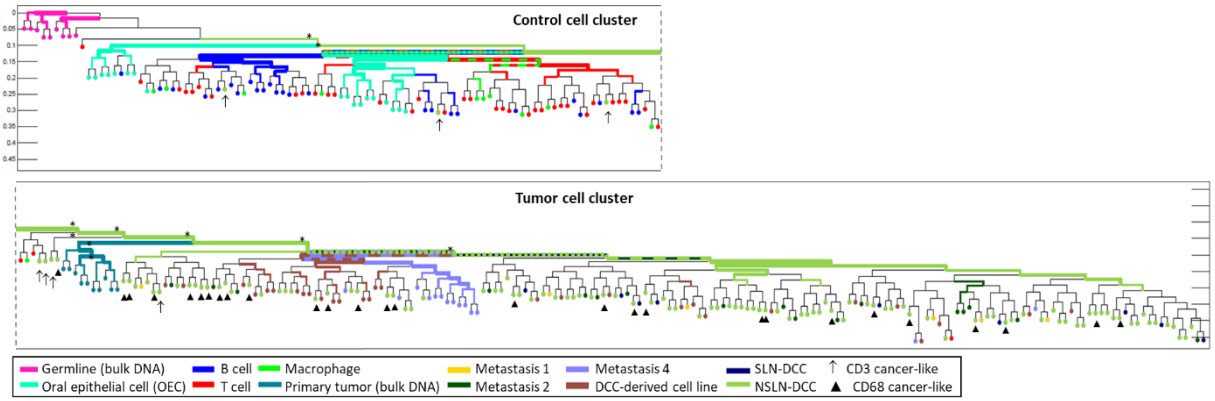


Figure 36: Cell lineage reconstruction of MM16-412
All high-quality samples were sequenced in median with 5.10x 10^6 reads. For better vizualiation, the complete tree is separated into two parts by the dashed line and the resulting two clusters are shown below each other. The tree was reconstructed using the FastTree2 algorithm and rooted at the median genotype of the germline samples. Edges reproduced with significant support (TBE>70%) in bootstrap analysis were marked with *. Y-axis represents the relative depth. CD3-positive cells showing tumor-specific aberrations (CD3 cancer-like cells) are marked with ↑ and CD68-positive cells with tumor-specific aberrations (CD68 cancer-like cells) are labeled with▲.

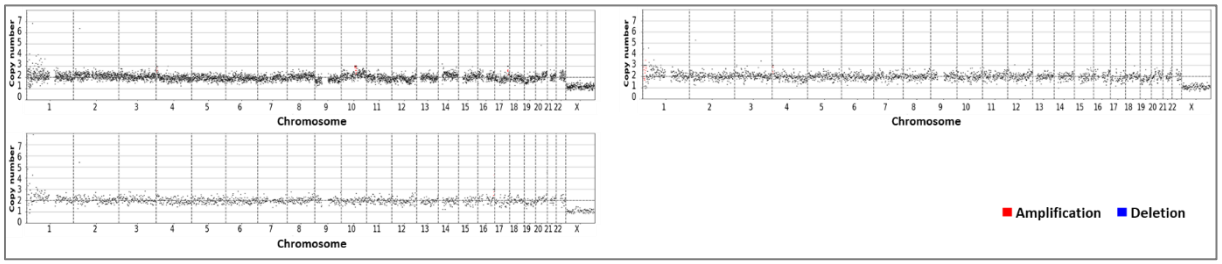


Figure 37: CNA profiles of T cells from MM16-412's tumor cell cluster
Balanced CNA profiles of T cells located within the tumor cell cluster. Amplifications are marked in red, deletions in blue.

Next, the phylogeny of the tumor cells was analysed in more detail (Figure 38). The tumor cell cluster could further be separated into several groups (Figure 38 A). At first, there were some NSLN-DCCs branching off at the beginning of the tumor cell cluster. The next branch included bulk DNA from the primary tumor with all samples and replicates nicely clustering together. The remaining part of the phylogenetic tree could be separted into a small and a large mixed DCC cluster, a group mainly consisting of NSLN-DCCs and NSLN-DCC derived cell line cells and a branch of tumor cells from metastasis 4. The mixed DCC clusters mainly contained NSLN-DCCs, some NSLN-DCC derived cell line cells and gp100-positive cells from metastasis 1 and 2. In contrast, large melanoma marker-negative cells from metastasis 4 clinically removed more than half a year later built a separate branch indicating their diversity and different nature compared to the other DCCs. Moreover, the majority of cells from the cell lines were found within one branch. As single cells from metastases 4 and from the cell line were clustering each more closely compared to SLN-DCCs, which were scattered within the tree, a more clonal picture of tumor cells from metastases 4 and cell lines was observed. However, this could not be explained by macroscopic evaluation of their CNA profiles as all tumor cells shared characteristic deletions on chromosome 5, 6, 9, 11q and 17 (Figure 38 B). Also all nine scattered SLN-DCCs shared these characteristic CNAs except SLN-DCC4 (see appendix Figure 78). An additional deletion on chromosome 12 could only be found for the NSLN-DCCs, NSLN-DCC derived cell line cells and cells from metastasis 1 and 4. Regarding the order of metastatic seeding there was no clear separation of cells from SLN, NSLN and metastases 1 and 2. Primary tumor samples could be found at the beginning of the tumor cell cluster (Figure 38 A). All six NSLNs with different tumor burden reflected by their DCC density (DCCD) were removed at the same time. However, within the tumor cell cluster cells from the different NSLNs did not separate clearly from each other (Figure 39). Thus, based on the lineage tree one could also not recognize that the cell lines were derived from NSLN3 and 6 as those DCCs were spread across the whole tumor cell cluster. Notably, the SLN-DCCs were found together with most of the NSLN-DCCs and metastases 1 and 2. The last-occurring metastasis 4, removed a year after metastases 1 and 2, clustered with NSLN-DCCs and the cell line.

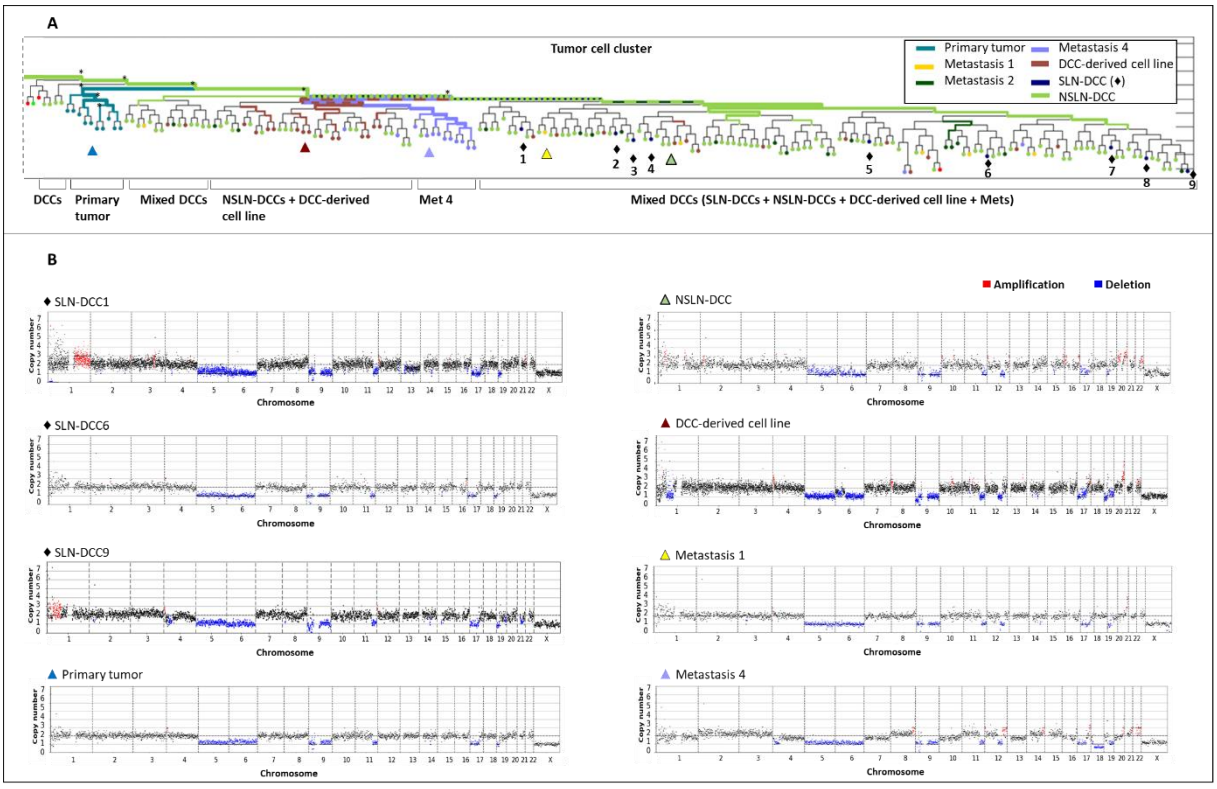


Figure 38: Tumor cell cluster of MM16-412's cell lineage reconstruction and selected CNA profiles
A: Tumor cell cluster of MM16-412 from MM16-412's lineage tree. Tumor cell cluster of Figure 36 could be splitted into different clades of tumor cells. All nine SLN-DCCs were labelled with ♦. Cells whose CNA profiles are shown in **B** are marked with▲. **B:** CNA profiles of selected tumor cells from MM16-412's tumor cell cluster. Representative CNA profiles of different types of tumor cells from MM16-412 were shown. Amplifications were marked in red, deletions in blue.

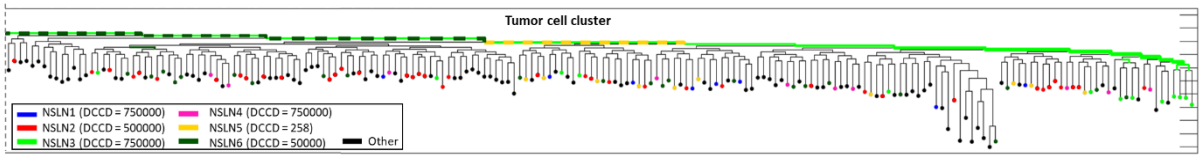


Figure 39: Tumor cell cluster of MM16-412 with coloring of different NSLNs
The tumor cell cluster is shown focussing on the six different non-sentinel lymph nodes (NSLNs) separated by colors. DCCD: DCC density (number of DCCs per million lymph node cells).

4.2.1.6.3 Phylogenetic analysis of MM16-423

The cell lineage reconstruction of MM16-423 contained 184 tumor and non-tumor control samples. The phylogenetic tree showed that samples clustered in two main groups, the control and tumor cell cluster (Figure 40). Indeed, more than 96% of tumor cells were located in the tumor cell cluster and 97.5% of control samples were detected within the control cell cluster showing reliability of the method (Figure 44). However, there were still three NSLN2-derived macrophages and one endothelial cell isolated from NSLN1-5, found within the tumor cell cluster although having balanced CNA profiles (Figure 41). The tumor arm further splitted into three clusters. The first two contained a mixture of tumor cells coming from blood and lymph nodes and the third cluster consisted of DCC-derived cell line/ xenograft cells. The cell line/ xenograft cells showed a clonal pattern with all single cells clustering together and branched off from thereof derived NSLN-DCCs showing reliability of the method (Figure 40). Interestingly, circulating tumor cells (CTCs) were scattered within the mixed DCC clusters indicating heterogeneity among these samples. One group of two CTCs (represented by CTC-A) was located at the beginning of the tumor cell cluster, whereas another group of two CTCs (represented by CTC-B) was found within the second cluster of mixed DCCs (Figure 40). Three additional single CTCs

were found within the first cluster of mixed DCCs within NSLN-DCCs and SLN-DCCs. Having a look at the CNA profiles of representative cells of the different CTC groups, CTC-A and CTC-B shared aberrations on chromosome 2, 6, 11, 14 and 15 that were also found in NSLN-DCCs and xenograft samples (Figure 42). Moreover, CTC-B of the second group showed an additional deletion on chromosome 16 shared with NSLN-DCCs and xenograft samples and harbored unique alterations on chromosome 5 and 12 confirming heterogeneity among analysed CTCs (Figure 42).

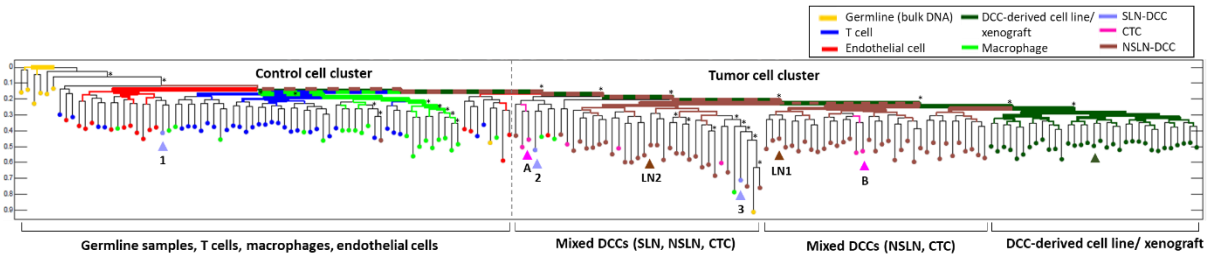


Figure 40: Cell lineage reconstruction of MM16-423
The phylogenetic tree of MM16-423 included control and tumor cells as well as bulk DNA from germline samples and the primary tumor. All samples were sequenced in median with 4.01x 10^6 reads. The tree was reconstructed using the FastTree2 algorithm and was rooted at the median genotype of the germline samples. Edges reproduced with significant support (TBE>70%) in bootstrap analysis are marked with *. Y-axis represents the relative depth. Cells whose CNA profiles are shown in Figure 42 are marked with▲.

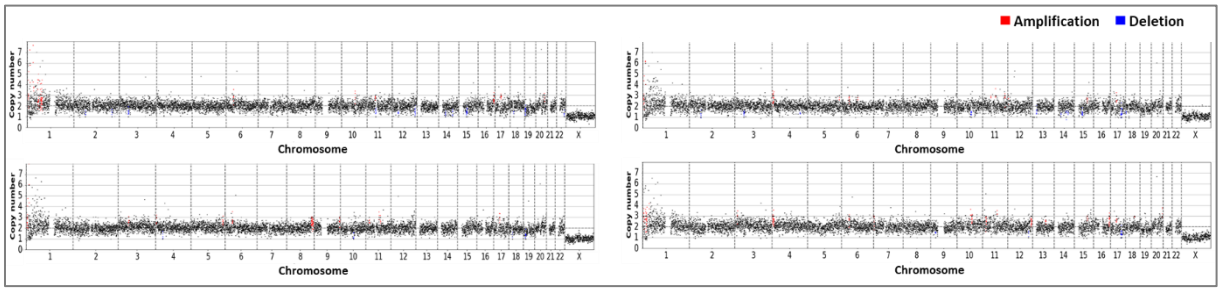


Figure 41: CNA profiles of control cells from MM16-423's tumor cell cluster
Balanced CNA profiles of three macrophages and one endothelial cell located within the tumor cell cluster. Amplifications are marked with red, deletions with blue color.

Notably, the SLN-DCCs isolated years before the CTCs were also scattered and mixed with other tumor cells. One SLN-DCC was detected within the control cell cluster and two SLN-DCCs at the beginning of the tumor cell cluster between NSLN-DCCs and CTCs (Figure 40). As the first SLN-DCC showed only few CNAs and was located within control cells, this might be the earliest DCC (Figure 42). However, in order to confirm this more analyses are needed. Similar to SLN-DCC1, SLN-DCC2 found together with CTCs and NSLN-DCCs had only few CNAs. However, the amplification on chromosome 6 and deletion on chromosome 15 were shared with NSLN-DCCs (Figure 42). The most progressed SLN-DCC was SLN-DCC3 that was located between NSLN-DCCs sharing CNAs on chromosome 6, 11 and 15. Aberrations on chromosome 4 were unique to that cell (Figure 42). For this patient, the macroscopic evaluation of CNA profiles helped to understand the location of some tumor cells as SLN-DCC3 showed more aberrations than SLN-DCC1 and SLN-DCC2 and CTC-B more than CTC-A. However, from CNA analysis one could not explain, why CTC-A and SLN-DCC2 clustered together (Figure 40, Figure 42).

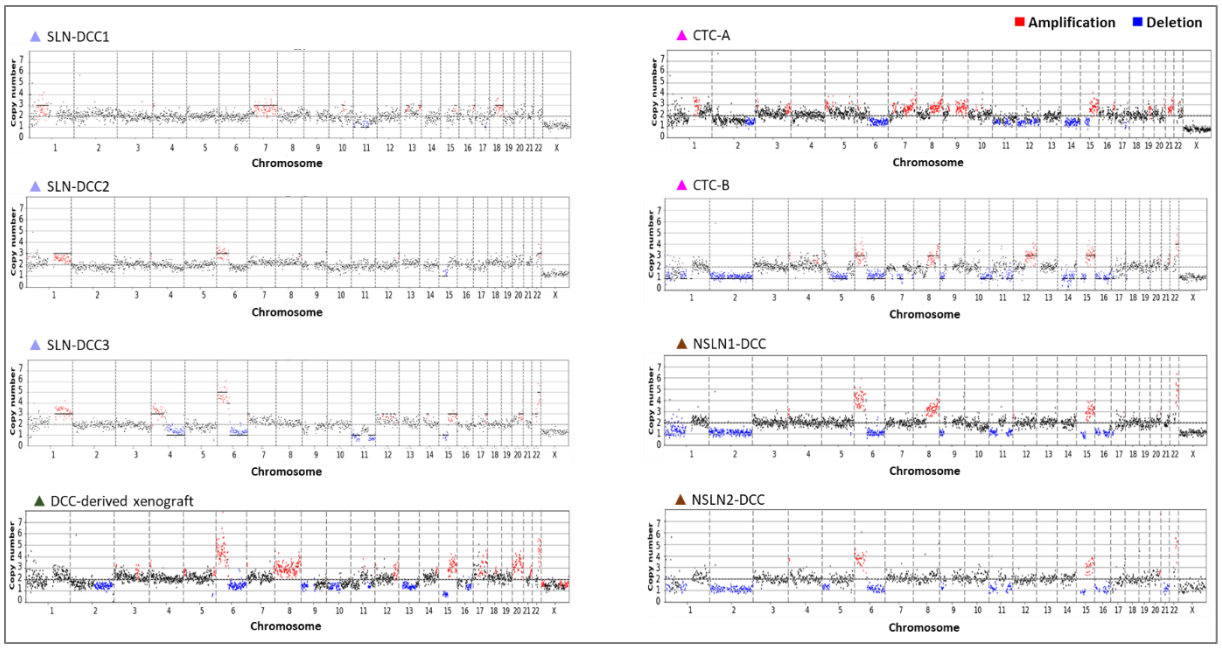


Figure 42: CNA profiles of selected tumor cells from MM16-423's lineage tree
Representative CNA profiles of different types of tumor cells from MM16-423. Amplifications are marked with red, deletions with blue color.

Next, the colonization of the different NSLNs was evaluated as six NSLNs with different tumor cell burden were removed from the patient at the same time (Figure 43). DCCs from NSLN2 and 4 were mainly found in one mixed DCC cluster whereas the other mixed DCC cluster contained the majority of DCCs from NSLN1 and 6. However, analogous to MM16-412, there was no clear separation of the DCCs from the different NSLNs.

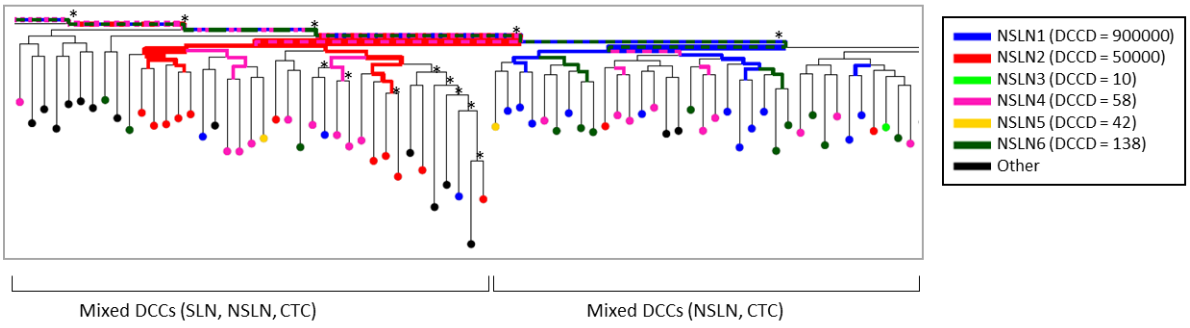


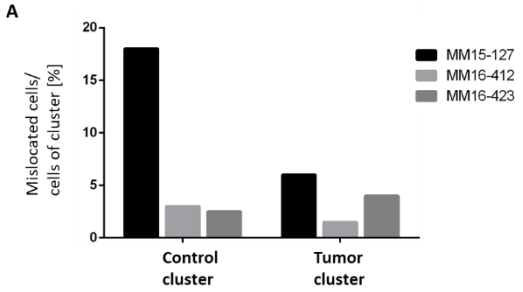
Figure 43: Part of tumor cell cluster of MM16-423 with coloring of different NSLNs
Relevant part of the tumor cell cluster was shown focussing on the six different non-sentinel lymph nodes (NSLNs) separated by colors. Edges reproduced with significant support (TBE>70%) in bootstrap analysis are marked with *. DCCD: DCC density (number of DCCs per million lymph node cells).

4.2.1.7 Summary of phylogenetic analysis based on microsatellite sequencing

After protocol and algorithm optimizations, the compatibility of the established library preparation protocol with manually isolated single cells and tumor bulk DNA was tested. The sequencing depth was increased to improve the tree reconstruction accuracy. The consideration of the copy number alteration (CNA) profiles confirmed correct group assignment of cells belonging to either control or tumor cell group. In addition, with the help of the CNA analysis low-quality cells were removed from lineage reconstructions. It was shown that the method was realible as technical replicates of bulk samples indeed clustered together and tumor cells could be separated from non-tumor control cells. However, in all patients potentially mislocated cells were detected, i.e. tumor cells within the control cell cluster or control cells within the tumor cell cluster. The highest percentages of potentially

mislocated cells were detected in MS-based trees from MM15-127 (Figure 44 A). For MM16-412 and MM16-423, less than 5% of cells were mislocated. T cells were found within the tumor cell cluster of MM15-127 and MM16-412, whereas macrophages and an endothelial cell were detected in the tumor cell cluster of MM16-423 (Figure 44 B). However, all control cells found in the tumor cell cluster showed balanced profiles. Tumor cells detected within the control cell cluster showed aberrant genomes and were derived from SLN or NSLNs (Figure 44 B). However, CNA analysis could not explain these mislocations remaining unclear whether tumor cells detected in the control cell cluster were indeed the earliest DCCs or misolacted due to technical reasons.

A



B

	MM15-127	MM16-412	MM16-423
# of cells in control cluster	**22**	**99**	**77**
Mislocated cells	**18% [4/22]**	**3% [3/99]**	**2.5% [2/77]**
Cell type	SLN-DCC	NSLN-DCC	SLN-DCC + NSLN-DCC
# of cells in tumor cluster	**116**	**203**	**199**
Mislocated cells	**6% [7/116]**	**1.5% [3/203]**	**4% [4/199]**
Cell type	T cell	T cell	3 MO + endothelial

Figure 44: Potentially mislocated cells in MS-based trees
A: Mislocation of cells in the control or tumor cluster for each patient. The percentage of misloacated cells either in the control or tumor cluster are shown. **B:** Table of mislocated cells for each patient showing the number and cell type of those cells. SLN: sentinel lymph node; NSLN: non-sentinel lymph node; DCC: disseminated cancer cell; MO: macrophage.

In summary, in the phylogenetic trees one could not see a clear grading or separation of the different DCCs for all patients indicating that a seeding of lymph nodes and metastases in parallel is more likely than in cascades. Regarding the colonization of the different non-sentinel lymph nodes with varrying tumor burden, there was also no order of seeding for patient MM16-412 and MM16-423 suggesting a simultaneous colonization of their NSLNs. This further indicated that lymph node DCCs did not form a separate entity but were in close phylogenetic relationship with systemically spread cancer cells. SLN-DCCs isolated at an early time point of disease were found in almost every branch or cluster of the lineage trees suggesting that SLN-DCCs might cover the tumorigenic potential of systemic cancer cells and be therapeutically relevant. However, as macroscopic evaluation of single CNA profiles did not help to understand individual locations of DCCs from SLN, NSLNs and metastases more analyses are needed to confirm the plausibility of MS-based trees and to validate drawn conclusions. In order to address above mentioned findings from a different point of view phylogenetic trees should be reconstructed from CNA data.

4.2.2 Lineage tree analysis based on copy number alterations

Tracking the phylogeny of disseminated melanoma cells on the basis of copy number alterations (CNAs) was assessed in cooperation with Dr. Martin Hoffmann (Fraunhofer ITEM-R).

4.2.2.1 Results of melanoma patient MM16-423

As part of a pilot experiment it was attempted to generate a phylogenetic tree for patient MM16-423 based on existing CNA data from 91 tumor cells including two SLN-DCCs, 64 NSLN-DCCs, 10 CTCs and 15 single cell line cells derived from the NSLN1. At first, CNA data of single cells representative for the different tumor samples were macroscopically analysed. As already described, there were many shared amplifications or deletions, for example on chromosome 6, 11 or 15 between the single tumor cells of the different tissues (Figure 45). The same was seen in the CNA-based heatmap comparing all analysed tumor cells (Figure 46 A). In addition, the two SLN-DCCs also shared amplifications and deletions among others on chromosome 6, 11 and 15 (Figure 45 and Figure 46 A). However, not all alterations found within the remaining tumor cells were detected within SLN-DCC2 indicating that this

SLN-DCC represented a very early DCC. Nevertheless, as CNAs shared between many tumor cells were already found in DCCs from the SLN, SLN-DCCs might have the potential to serve as surrogates for systemically spread cancer cells. Within the reconstructed phylogenetic tree from CNA data, two big branches were found (Figure 46 B). The majority of cell line cells clustered together forming one branch, whereas the other branch consisted of CTCs and lymph node-derived DCCs. Comparing the preliminary CNA-based tree with the MS-based tree one could see that with both methods cell line cells were clearly separated from the remaining tumor cells (Figure 46 B and Figure 40). Moreover, in the CNA-based tree SLN-DCC2 and SLN-DCC3 were located within two different clusters (Figure 46 B). SLN-DCC3 was mixed with NSLN-DCCs and few cell line cells, whereas SLN-DCC2 was found together with NSLN-DCCs and CTCs including CTC-A. This was consistent with MS-based findings as SLN-DCC2 closely clustered together with CTC-A. In the CNA-based tree CTC-B was detected within NSLN-DCCs in a different branch separated from CTC-A and SLN-DCCs in line with the MS-based analysis.

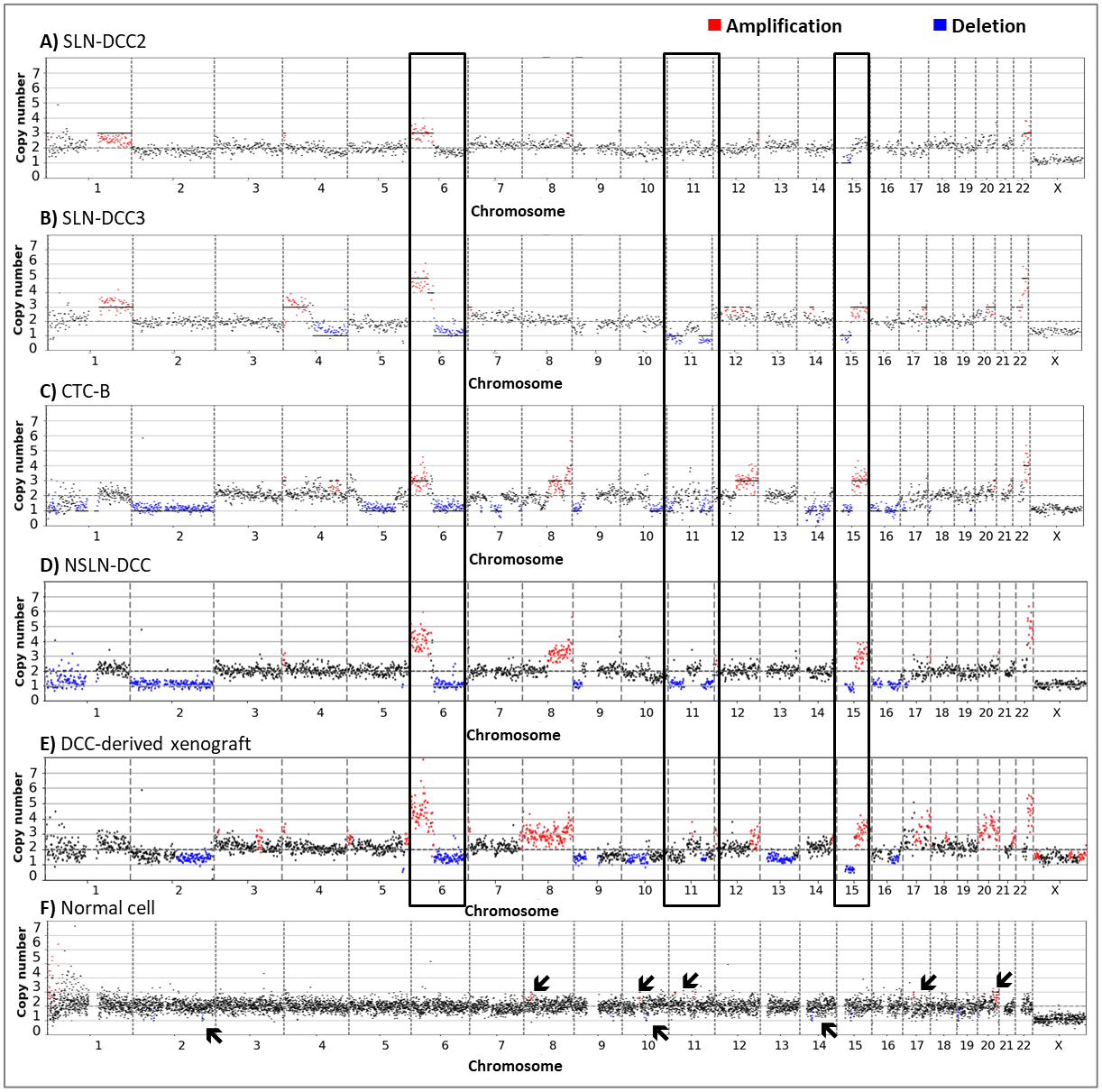


Figure 45: CNA profiles of single cells from MM16-423
Representative CNA profiles of tumor cells (**A** to **E**, from Figure 42) and a normal cell **(F)** from MM16-423 are shown. Amplifications are marked with red, deletions with blue color. Examples of shared regions are marked with black frames. Arrows in **(F)** indicate small putatively false-positive CNAs found in normal cells that were called by the software and therefore marked in colors.

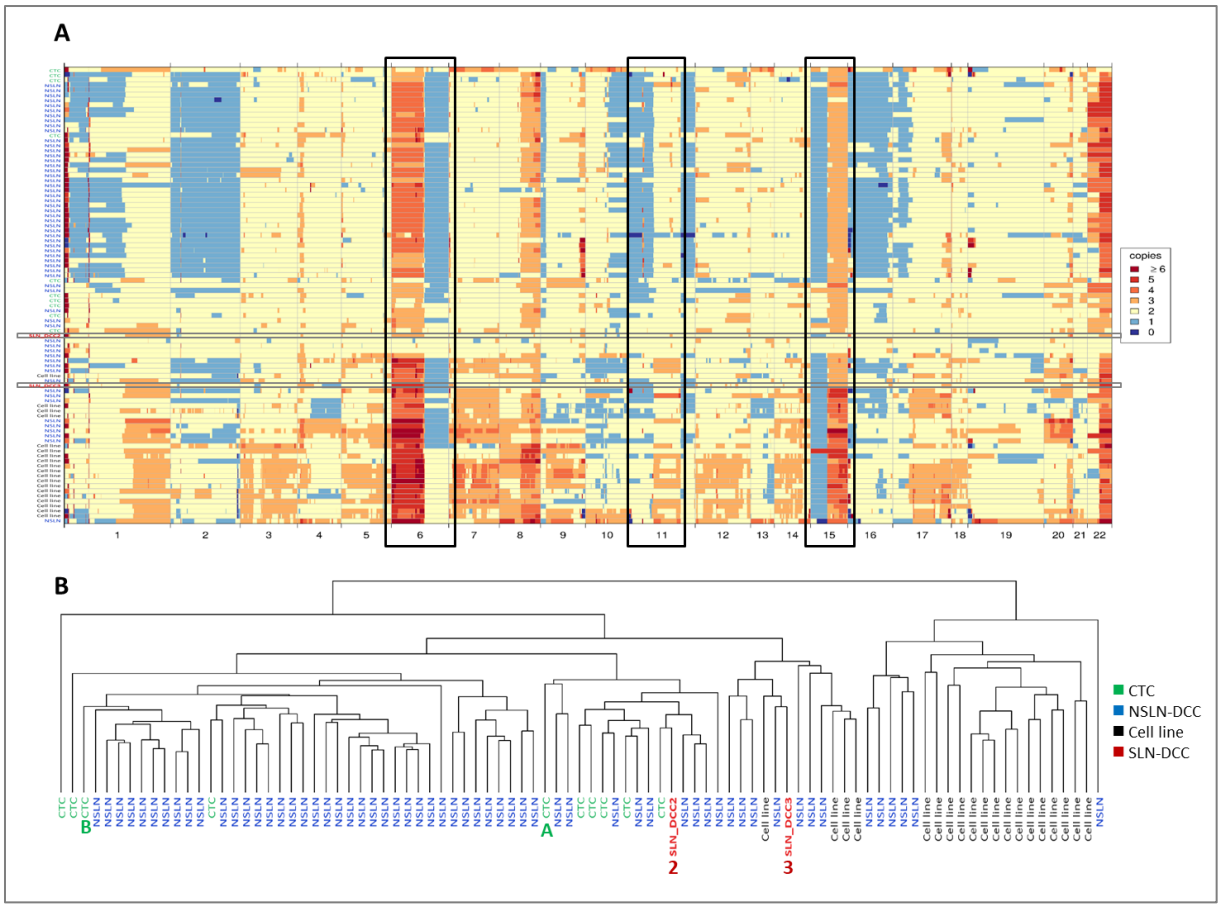


Figure 46: CNA-based heatmap and phylogenetic tree of MM16-423
Heatmap **(A)** and corresponding clustering-based phylogenetic tree **(B)** of CNA profiles of 91 samples (10 CTCs, 64 NSLN-DCCs, 2 SLN-DCCs, 15 cell line cells) from MM16-423. Amplifications are marked with red, deletions with blue color. In **(A)** two SLN-DCCs are marked with a grey thin frame and prominent shared regions on chromosome 6, 11 and 15 with black frames.

4.2.2.2 Challenges and outlook of CNA-based phylogenetic analysis

Although the currently used software tool for the assessment of CNAs, provided by the company selling *Ampli1™* low pass kit, allowed discrimination of tumor cells and non-tumor cells, repeatedly small putatively false-positive CNAs in control cells were detected by the software (Figure 45 F; red dots on chromosome 8, 10, 11, 17, 20 and blue dots on chromosome 2, 10, 14 in the normal cell marked with arrows). However, these small putatively false-positive CNAs represented artefacts that could confound cells in the analyses. Those artefacts might be caused by the different sequencing technologies that were used to originally develop the analysis tool (Ion torrent) and that was used within this project (Illumina). As these artefacts did not occur in 100% of cells, they may affect the construction of the CNA-based trees. Aiming to reduce and eliminate these artefacts an alternative software developed by a group from the Department of Applied Research and Technical Development, Fondazione IRCCS Instituto Nazionale dei Tumori (Milano, Italy) was currently being tested. However, this bioinformatic analysis exceeded the frame of the thesis and required more time and effort. Nevertheless, this software will then allow normalization against a specific determined pool of control cells and will therefore significantly reduce the confounding artefacts. Once this is finished, phylogenetic trees based on CNAs will be performed for all three melanoma patients. Finally, trees based on MS-analysis and CNA-analysis can be compared to assess the impact of selection pressure on the tree reconstructions and will inform on the necessity to implement methods that are independent from selection pressure in order to address evolutionary aspects of the disease.

4.2.3 Phylogeny based on *BRAF* mutation rate

Around 65% of melanoma patients carry a mutation either in the *NRAS* or *BRAF* gene (Muñoz-Couselo et al. 2017, Ascierto et al. 2012). The analysis of the cell lines derived from the three patients of this project revealed *NRAS* mutations for MM15-127 and MM16-412 and a mutation in the *BRAF* hotspot region (V600E) for patient MM16-423 (chapter 2.2). The allele-specific PCR with a blocking reagent (ASB-PCR) previously developed in our laboratory enabeled testing for this hotspot mutation of DCCs from MM16-423 at different anatomic sites and time points of disease progression. In total 99 single tumor cells including DCCs from the SLN, NSLNs, cell line and CTCs were tested. However, 10 tumor cells dropped out from ASB-PCR analysis and four tumor cells because of showing either balanced or low-quality CNA profiles (see appendix Table 56). Of the remaining 85 tumor cells, all SLN-DCCs were wild type whereas all tumor cells that were removed at a later time point of disease progression (CTCs, DCCs from the cell line and from all NSLNs, except NSLN4) were mutated (Table 41). DCCs from NSLN4 showed a mixture of *BRAF* wild type and mutated cells. In summary, 92.9% of analysed tumor cells carried the *BRAF* hotspot mutation.

Table 41: *BRAF* mutation analysis of single cells from MM16-423 using ASB-PCR

MM16-423	DCCD	Evaluated cells	BRAFmut	BRAFwt	BRAFmut [%]	BRAFwt [%]
SLN	2	3	0	3	0	100
NSLN4	58	15	12	3	80	20
NSLN3	10	1	1	0	100	0
NSLN5	42	1	1	0	100	0
NSLN6	138	12	12	0	100	0
NSLN2	50 000	10	10	0	100	0
NSLN1	900 000	13	13	0	100	0
Cell line		18	18	0	100	0
CTC		12	12	0	100	0
		85	79	6	92.9	7.1

Analysis of single cells from MM16-423 for BRAF V600E mutation using ASB-PCR. SLN: sentinel lymph node; NSLN: non-sentinel lymph node; CTC: circulating tumor cell; mut: mutated (BRAF V600E); wt: wild type; red: 100% BRAFmut; green: 100% BRAFwt. This data-set was also used in the thesis of Julia Greindl-Junghans due to an internal collaboration.

The fact that *BRAF* mutations were not detected in SLN-DCCs but in 80% of DCCs from NSLN4 and in 100% of remaining NSLN-DCCs and CTCs indicated that SLN and NSLN4 have been seeded first. It further showed that tumor cells isolated early during disease progression were *BRAF* wild type. Notably, without the information of the *BRAF* status of the primary tumor it remained unclear whether only *BRAF* wild type cells or *BRAF* wild type and mutated cells disseminated. However, when combining the *BRAF* status and MS-based phylogenentic trees, *BRAF* wild type samples were indeed found early in the tumor cell cluster, thus suggesting the tissues with *BRAF* wild type cells were one of the first stations within dissemination (Figure 47).

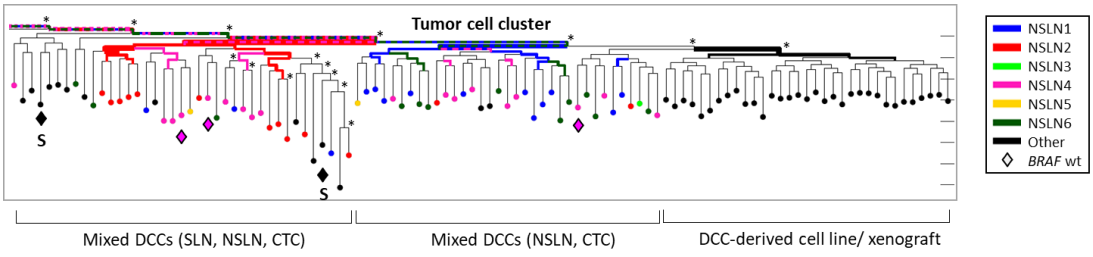


Figure 47: Tumor cell cluster of MM16-423 including *BRAF* mutation status
The tumor cell cluster is shown with non-sentinel lymph nodes (NSLNs) separated by colors. In addition, the *BRAF* mutation status is depicted. *BRAF* wt cells were marked with ◊. The remaining samples were *BRAF* mutated. S: SLN-DCC; wt: wild type. Note: *BRAF* wt SLN-DCC1 was located in the control cluster and therefore not shown here.

4.3 Identification of shared tumor-specific coding mutations

Phylogenetic analyses indicated a parallel seeding of lymph nodes and metastases and revealed that SLN-DCCs were spread across the tumor clusters of MS-based lineage trees. Moreover, CNA analysis showed that SLN-DCCs already contained aberrations shared with other tumor cells from non-sentinel lymph nodes or metastases. This suggested that SLN-DCCs indeed could comprise metastasis-initiating cells and pointed towards the suitability of SLN-DCCs to serve as surrogates for systemically spread cancer cells. Whether early DCCs from routinely removed SLNs were indeed surrogates for systemic DCCs, it was evaluated whether SLN-DCCs could enable target identification in context of shared neoantigens for adjuvant therapy. To identify truncal tumor-specific mutations shared between cancer cells present at different anatomical sites and disease stages, cells of interest were subjected to whole exome sequencing (WES).

4.3.1 Overview of sample preparation for WES

High-quality samples (246 single cells, 9 bulk samples) were prepared for WES in cooperation with the Fraunhofer ITEM-R and sequenced either as single cells, bulk or pool of single cells (Table 42 and appendix Figure 79). Sequencing of early DCCs from SLNs as single cells was of high importance (i) to assess whether these cells allowed the identification of truncal neoantigens shared with systemic disseminated cancer cells and (ii) to get insight into the heterogeneity among early DCCs. All bioinformatic analyses from WES data (i.e. HLA typing, mutational signature analysis, mutation detection and peptide prediction) were performed by Dr. Huiqin Koerkel-Qu.

Table 42: Overview of high-quality samples for WES

	MM15-127	MM16-412	MM16-423
Germline	Pool of 15 CD31+ SCs	Bulk	Bulk
Non-tumor cells	--	2 single B cells	2 gp100-negative SCs
PT	*Low quality*	Bulk	--
SLN (gp100)	8 SCs	10 SCs	3 SCs
SLN (large MCSP)	7 SCs	--	--
NSLN1	--	Pool of 11 SCs	Pool of 11 SCs
NSLN2	--	Pool of 14 SCs	Pool of 10 SCs
NSLN3	--	Pool of 10 SCs	Pool of 4 SCs
NSLN4	--	Pool of 10 SCs	Pool of 16 SCs
NSLN5	--	Pool of 6 SCs	Pool of 5 SCs
NSLN6	--	Pool of 21 SCs	Pool of 8 SCs
Met1	Pool of 16 SCs	Bulk + pool of 9 SCs	--
Met2	--	Pool of 13 SCs	--
Met3	Pool of 20 SCs	--	--
Met4	--	Bulk + pool of 10 SCs	--
CTCs	--	--	5 SCs
Cell line from LN	Bulk	2x bulk	Bulk
		∑ 64 reactions	

Reactions were sequenced on a NovaSeq 6000 (Illumina) in 6 different pools consisting of 350 bp libraries (see appendix Table 57, Table 58, Table 59). Sequencing coverage: 500x for bulk samples, 200x for single tumor cells and pools of single cells (SCs).

4.3.2 HLA-I analysis

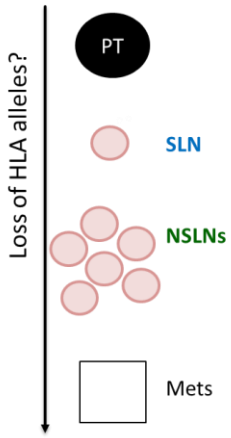
First, class I HLA alleles were analysed for all patients as this information was necessary for later class I neopeptide prediction where the HLA information was intersected with the binding affinity of identified neoantigens. Another aim was to assess mechanisms of immune evasion operating at different time points of the disease by investigating a potential loss of heterozygosity (LOH) for HLA alleles (Figure 48 A). In order to retrieve information on the HLA alleles of the patients, standard HLA typing (4-digit resolution) was performed using PCR-based amplifications of HLA loci, determined as ‘classical HLA typing’. However, as this needed up to 3 µg of DNA, classical HLA typing was only

performed for bulk samples from peripheral blood or lymph nodes and DCC-derived cell lines. HLA information for single cells had to be retrieved from WES data. As HLA retrival from WES data was not yet tested for single cells, the correct HLA alleles identified in classical HLA typing served as positive controls for *in silico* HLA typing.

4.3.2.1 Classical HLA-I typing

Classical HLA typing performed in cooperation with Dr. Bach (University Hospital Erlangen) revealed a LOH in HLA-I alleles in the DCC-derived cell line samples from MM16-412 (Figure 48 B). For the other two patients, all six HLA class I alleles were detected in germline and cell line samples.

A



B

	MM15-127	MM16-412	MM16-423
Germline	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	HLA-A*24:02, ***66:01** HLA-B*27:02, ***41:02** HLA-C*02:02, ***17:01**	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01
DCC line from SLN	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	?	--
DCC line from NSLN	--	HLA-A*24:02 HLA-B*27:02 HLA-C*02:02	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01

Figure 48: Overview and results of classical HLA-I typing
A: Sample overview for HLA anlaysis; ◉ lymph node; □ cutaneous metastasis; **B:** Classical 4-digit HLA-I typing results of germline and cell line samples. Alleles for which a loss of heterozygosity (LOH) in the DCC-derived cell lines was found are indicated in red in the germline sample.

4.3.2.2 In silico HLA-I typing using WES data

In order to retrieve HLA information from WES data, two commonly used HLA typing tools (HLA-miner and OptiType) were applied. As those HLA typing tools were developed and tested only on sequencing data of bulk samples, it had to be analysed whether they were applicable also to sequencing data of single cells. Both typing tools were first applied on sequencing data of germline bulk samples as from these samples the true HLA-I alleles were known due to classical HLA typing (chapter 4.3.2.1) serving as positive controls for bulk samples. As HLA-I alleles different from those detected by classical HLA typing were found in all germline samples using HLAminer, this tool failed already in its applicability to bulk samples and was therefore not tested on sequencing data of single cells (Table 43).

Table 43: Results of HLA-I typing of WES data from germline bulk using HLAminer

	MM15-127		MM16-412		MM16-423	
Tool	Classical typing	HLA-miner	Classical typing	HLA-miner	Classical typing	HLA-miner
Germline	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	HLA-A*24:02, *23:01 HLA-B*07:02, *48:0 HLA-C*01:02, *07:02	HLA-A*24:02, *66:01 HLA-B*27:02, *41:02 HLA-C*02:02, *17:01	HLA-A*26:01, *66:01 HLA-B*27:05, *41:02 HLA-C*02:02, *17:01	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01	HLA-A*43:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:13, *04:01

HLA-I typing results from WES data of germline samples using HLAminer tool compared to correct HLA alleles from classical HLA typing. red: alleles not detected by classical HLA typing.

When using OptiType, all correct HLA-I alleles, i.e. alleles determined by classical HLA typing of germline samples (Figure 48 B), were found in WES germline samples of each patient (Table 44). Next, the applicability of OptiType on single cell level was evaluated, analysing sequencing data of single non-tumor cells, either B cells or gp100-negative cells from patient MM16-412 and MM16-423, respectively. These cells were expected to have the same six HLA-I alleles as found in the germline samples. Due to limited sample availability of patient MM15-127, no single non-tumor cell was

applicable for WES. In case of non-tumor cells from MM16-412 and MM16-423, less than six alleles and also alleles not detected by classical HLA typing were detected (Table 44). These data indicated that some HLA-I sequences might not be there at all. As not all correct six alleles were found in non-tumor cells, in the next step the OptiType tool was used specifically searching for only the correct six HLA-I alleles from classical typing ('OptiType_new'). This increased the number of correct alleles for all non-tumor cells (Table 44). In summary, on bulk level all correct six alleles were detected for germline WES samples using OptiType (Table 44). However, this was not the case for all single non-tumor control cells, even when specifically searching for the correct alleles indicating technical challenges. The correct HLA-I alleles could be faithfully retrieved from WES data, if the correct alleles were known, but only if the alleles were present. On single cell level not all HLA-I alleles could be faithfully retrieved from control cells suggesting that some HLA sequences were missing in WES data from single cells. The reason for this might be unbalanced amplification of HLA alleles of single cells during WGA leading to amplification bias that was interfering with *in silico* typing using single cell WES data.

Table 44: Results of HLA-I typing of positive controls from WES data using OptiType

	MM15-127		MM16-412		MM16-423	
Tool	OptiType	OptiType_new	OptiType	OptiType_new	OptiType	OptiType_new
Germline	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	HLA-A*24:02, *66:01 HLA-B*27:02, *41:02 HLA-C*02:02, *17:01	HLA-A*24:02, *66:01 HLA-B*27:02, *41:02 HLA-C*02:02, *17:01	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01
Non-tumor cell 1	NA	NA	HLA-A*24:02, *66:01 HLA-B*41:02 HLA-C*02:02, *17:01	HLA-A*24:02, *66:01 HLA-B*27:02, *41:02 HLA-C*02:02, *17:01	HLA-A*02:01, *32:02 HLA-B*15:01 HLA-C *03:04	HLA-A*02:01, *26:01 HLA-B*15:01 HLA-C*03:03
Non-tumor cell 2	NA	NA	HLA-A*24:02 HLA-B*41:02 HLA-C*02:02	HLA-A*24:02 HLA-B*41:02 HLA-C*02:02, *17:01	HLA-A*02:01, *32:12 HLA-B*35:01 HLA-C	HLA-A*02:01, *26:01 HLA-B*35:01 HLA-C

HLA-I typing results from WES data of germline samples and two non-tumor cells per patient using OptiType. NA: not available; red: alleles not detected by classical HLA typing.

After testing the OptiType tool on sequencing data of non-tumor control samples, HLA-I typing of tumor samples from melanoma patients was performed specifically searching for the correct HLA-I alleles confirmed by classical HLA typing.

All correct alleles were found in pools of single non-tumor cells, pools of single cells from metastasis 1 and 3 and in the bulk sample of the autologous cell line of MM15-127 (Figure 49, Table 45). In nine from 15 single SLN-DCCs of MM15-127 all correct six HLA-I alleles were detected. In the remaining six single SLN-DCCs showing less than 6 HLA-I alleles not always the same alleles were missing (Table 45). For these samples one could not confirm a potential HLA-LOH as for MM15-127 no single non-tumor cells with all correct alleles were available and positive controls for single cells were lacking. Whether missing HLA-I alleles came from LOH or due to technical reasons remains unclear.

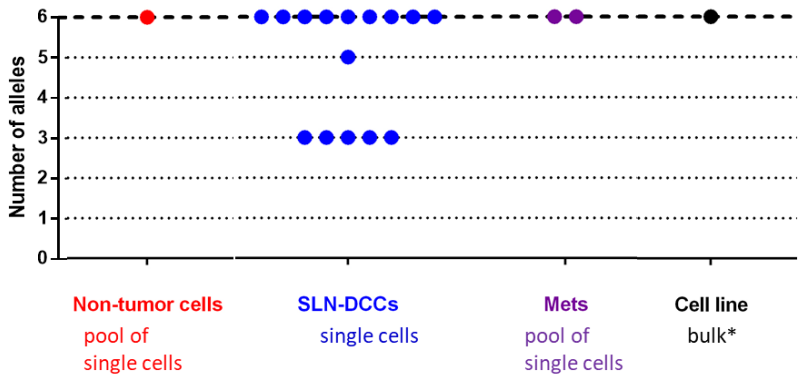


Figure 49: Number of detected HLA-I alleles of samples from MM15-127 using OptiType
SLN: sentinel lymph node; DCC: disseminated cancer cell; Met: metastasis; (*) confirmed by classical HLA typing.

Table 45: Results of HLA-I typing of all samples from WES data using OptiType

MM15-127		MM16-412		MM16-423	
Sample	HLA result	Sample	HLA result	Sample	HLA result
Pool of NTCs	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	Germline	HLA-A*24:02, *66:01 HLA-B*27:02, *41:02 HLA-C*02:02, *17:01	Germline	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01
SLN-DCC 10	HLA-A*24:02 HLA-B*27:05 HLA-C*01:02	NTC1	HLA-A*24:02, *66:01 HLA-B*27:02, *41:02 HLA-C*02:02, *17:01	NTC1	HLA-A*02:01, *26:01 HLA-B*15:01 HLA-C*03:03
SLN-DCC 6	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	NTC2	HLA-A*24:02 HLA-B*27:02 HLA-C*02:02, *17:01	NTC2	HLA-A*02:01, *26:01 HLA-B*15:01 HLA-C
SLN-DCC 9	HLA-A*24:02 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	PT	HLA-A*24:02, *66:01 HLA-B*27:02, *41:02 HLA-C*02:02, *17:01	SLN-DCC 1	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01
SLN-DCC 3	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	SLN-DCC 12	HLA-A*24:02 HLA-B*27:02, *41:02 HLA-C*02:02, *17:01	SLN-DCC 2	HLA-A*26:01 HLA-B*35:01 HLA-C*04:01
SLN-DCC 11	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	SLN-DCC 3	HLA-A*24:02 HLA-B HLA-C*02:02	SLN-DCC 3	HLA-A HLA-B*15:01 HLA-C*04:01
SLN-DCC 16	HLA-A HLA-B*07:02 HLA-C*01:02 *07:02	SLN-DCC 1	HLA-A*24:02 HLA-B*27:02 HLA-C*02:02, *17:01	NSLN1-3,5, 6	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01
SLN-DCC 8	HLA-A*68:01 HLA-B*07:02 HLA-C*01:02	SLN-DCC 6	HLA-A HLA-B HLA-C*02:02	NSLN4	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*04:01
SLN-DCC 17	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	SLN-DCC 5	HLA-A*24:02 HLA-B*27:02 HLA-C*02:02	CTC 3	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01
SLN-DCC 7	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	SLN-DCC 7	HLA-A HLA-B HLA-C	CTC 4	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01
SLN-DCC 4	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	SLN-DCC 13	HLA-A*24:02, *66:01 HLA-B*27:02 HLA-C*02:02, *17:01	CTC 5	HLA-A*02:01 HLA-B*35:01 HLA-C*03:03, *04:01
SLN-DCC 12	HLA-A*68:01 HLA-B*07:02 HLA-C*01:02	SLN-DCC 11	HLA-A*24:02, *66:01 HLA-B*27:02 HLA-C*02:02, *17:01	CTC 6	HLA-A*02:01, *26:01 HLA-B*35:01 HLA-C*03:03, *04:01
SLN-DCC 13	HLA-A*68:01 HLA-B HLA-C*01:02, *07:02	SLN-DCC 4	HLA-A*24:02 HLA-B*27:02 HLA-C*02:02	CTC 7	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01
SLN-DCC 14	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	SLN-DCC 10	HLA-A*24:02 HLA-B*27:02, *41:02 HLA-C*02:02, *17:01	Cell line	HLA-A*02:01, *26:01 HLA-B*15:01, *35:01 HLA-C*03:03, *04:01
SLN-DCC 6	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	NSLN1-6	HLA-A*24:02 HLA-B*27:02 HLA-C*02:02		
SLN-DCC 15	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	Met1, Met2, Met4 (pool)	HLA-A*24:02 HLA-B*27:02 HLA-C*02:02		
Met1, Met3, Cell line	HLA-A*24:02, *68:01 HLA-B*07:02, *27:05 HLA-C*01:02, *07:02	Met1, Met4 (bulk)	HLA-A*24:02, *66:01 HLA-B*27:02, *41:02 HLA-C*02:02, *17:01		
		Cell line a/b	HLA-A*24:02 HLA-B*27:02 HLA-C*02:02		

SLN: sentinel lymph node, NSLN: non-sentinel lymph node, PT: primary tumor, NTC: non-tumor cell, Met: metastasis, CTC: circulating tumor cell.

Regarding patient MM16-423 all correct six alleles were detected from the bulk sample of the autologous cell line (Figure 50, Table 45). In pools of single cells from NSLN-DCCs, for five of six pools all alleles were detected. On single cell level, not all six correct alleles were detected in both non-tumor cells. In single tumor cells all correct six alleles were found for one of the three SLN-DCCs and for three of five CTCs. From the remaining single tumor cells the absence of some different HLA-I alleles might be a technical problem or a potential LOH. However, HLA-LOH could not be confirmed as all correct alleles could not be detected in single non-tumor cells.

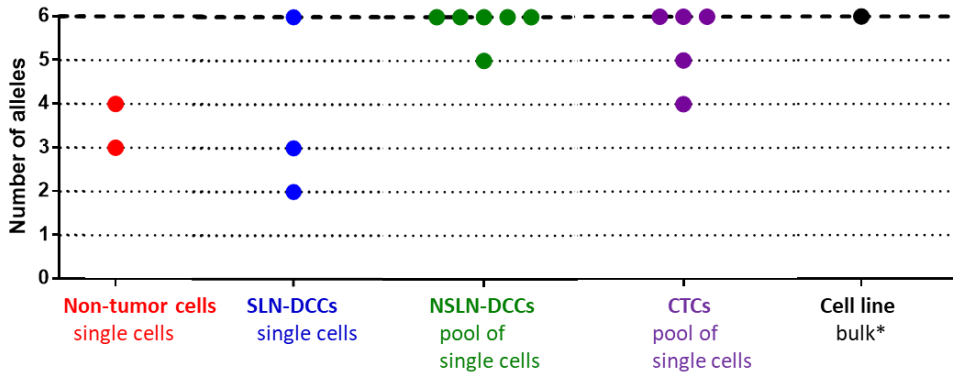


Figure 50: Number of detected HLA-I alleles of samples from MM16-423 using OptiType
SLN: sentinel lymph node; DCC: disseminated cancer cell; NSLN: non-sentinel lymph node; CTC: circulating tumor cell; (*) confirmed by classical HLA typing.

For patient MM16-412, in bulk samples of the primary tumor and of metastatic tissues sequenced at a coverage of 500x all six alleles were detected, in SLN-DCCs and pools of single cells from NSLNs or metastatic tissues sequenced at 200x less than six alleles were found (Figure 51, Table 45). For this patient several aspects suggested real LOHs for HLA-I alleles of some tumor samples. (I) In bulk DNA from the cell line, which was derived from NSLN-DCCs, already classical HLA typing revealed a LOH. (II) Additionally, that bulk DNA from metastatic tissues showed six alleles in contrast to pools of single cells thereof might be explained by the presence of non-tumor cells in bulk samples representing the tumor microenvironment (TME). Regarding the primary tumor sample, also sequenced as bulk, a contamination with non-tumor cells could also be possible explaining the detection of six alleles. However, a loss of chromosome 6, which includes regions encoding for the HLA-I alleles (chromosome 6p), was found for the primary tumor sample indicating LOH (Figure 52 D).

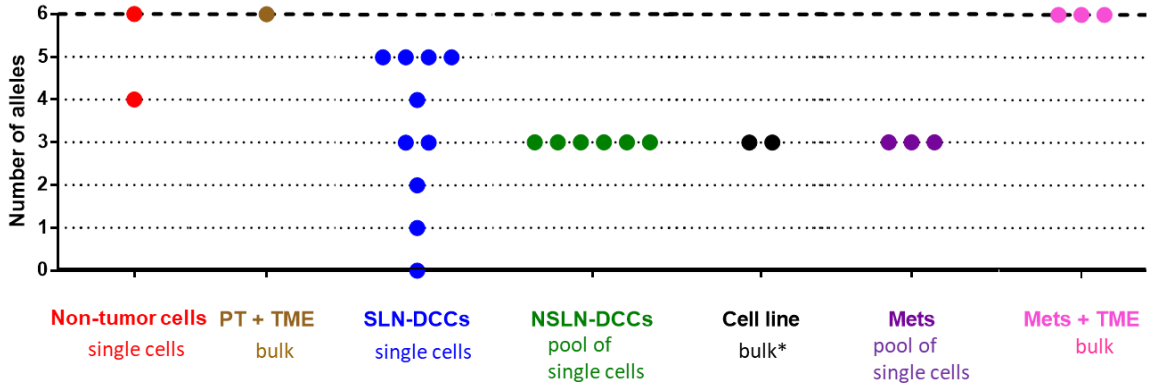


Figure 51: Number of detected HLA-I alleles of samples from MM16-412 using OptiType
PT: primary tumor; TME: tumor microenvironment; SLN: sentinel lymph node; DCC: disseminated cancer cell; NSLN: non-sentinel lymph node; Met: metastasis; (*) confirmed by classical HLA typing.

Moreover, based on the CNA analysis single tumor cells with a potential LOH also showed deletions on chromosome 6p (Figure 52). In SLN-DCCs a maximum of five HLA-I alleles were detected indicating that the HLA-LOH already started in the SLN. However, SLN-DCCs were found to be a heterogenous population regarding the presence of HLA-I alleles ranging from zero to five detected alleles. CNA analysis further showed less deletions on chromosome 6p of SLN-DCCs with five HLA-I alleles compared to SLN-DCCs with three alleles (Figure 52 B, C). In summary, these findigs indicated that for this patient HLA-LOH started at an early time point of disease and that already SLN-DCCs were exposed to immune pressure.

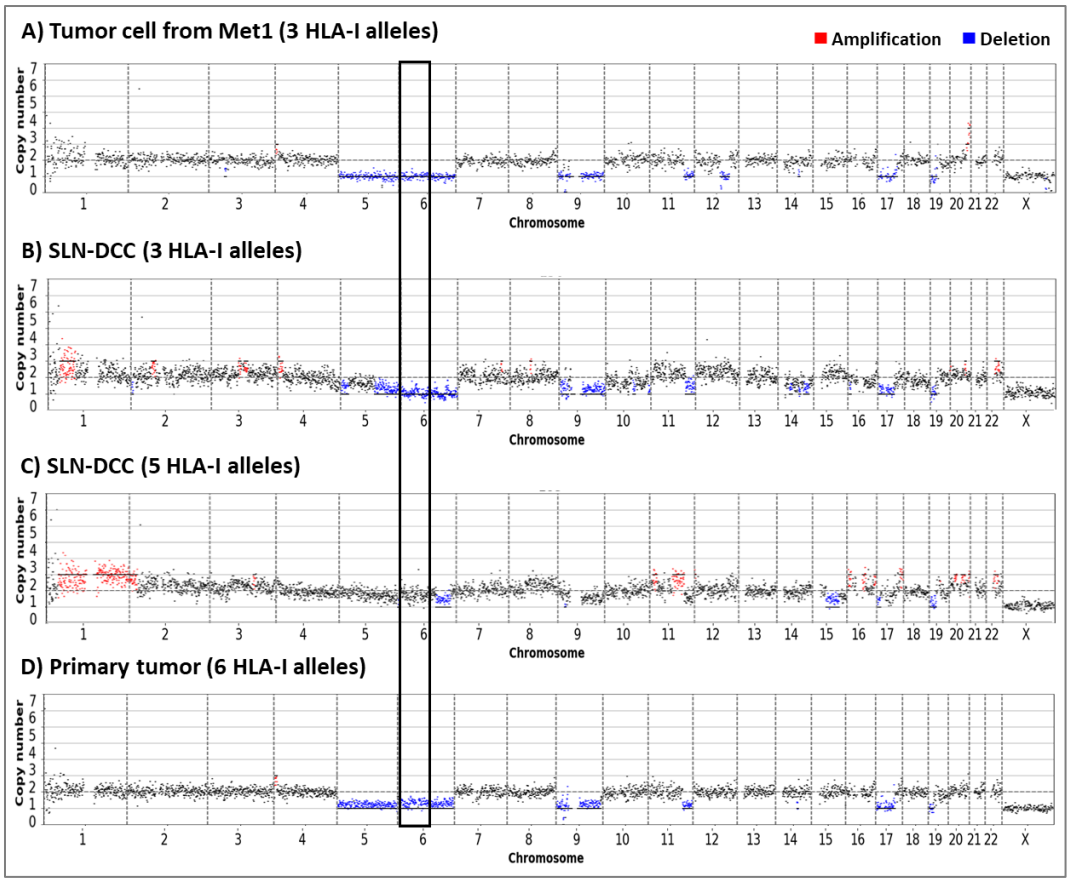


Figure 52: Representative CNA profiles of tumor samples from MM16-412
CNA profile of a metastatic tumor cell **(A)**, a SLN-DCC **(B)** with both having three HLA-I alleles, CNA profile of a SLN-DCC **(C)** having five HLA-I alleles and of the primary tumor **(D)** with six HLA-I alleles are shown. Amplifications are marked in red, deletions in blue. Deletions on chromosome 6p (marked with frame) indicated loss of heterozygosity of HLA alleles.

4.3.3 Mutation detection

Using the data from WES, the aim was to assess the neoantigen and thereof predicted neopeptide load of lymph nodes (SLN, NSLNs), metastases (Mets, CTCs) and those that were shared among these samples. However, mutational signature analysis of all samples was performed first, in order to see whether specific signatures associated for example with UV light exposure could be detected in the tumor samples of the melanoma patients. For mutational calling of the signature analysis two variant calling tools were applied: Varscan and Platypus. In general less mutations were detected when applying Platypus, as it is a stricter algorithm with more filtering parameters. In order to increase the chance to identify shared exonic mutations only Varscan was used for the analysis of shared mutations as it detected in total more mutations than Platypus. To eliminate false-positive mutations, mutations were called appearing in at least two non-pool tumor samples (i.e. bulk samples and single cells). In the mutational signature analysis all coding mutations were used, whereas for the identification of shared mutations only exonic mutations changing the protein sequence were considered, i.e. non-synonymous mutations, frameshifts or missense mutations. Regarding peptide selections of potential neoantigens for later functional validation only exonic non-synonymous mutations were analysed.

4.3.3.1 Mutational signatures

For MM15-127 bulk DNA of the SLN-derived cell line, 15 single SLN-DCCs and two pools of single cells from two cutaneous metastases were analysed. The mutational signature analysis of MM15-127 using the Varscan algorithm showed that UV signature 7 was found in all of the tumor samples, apart from one SLN-DCC (Figure 53 A, marked with *). Interestingly, this SLN-DCC showed less mutations compared to the other tumor cells of this patient, which might be the reason that no UV signature was

found for this cell (Figure 53 A, log10dsmut). However, although no UV signature could be detected for this SLN-DCC that was located within the control cell cluster of the MS-based lineage tree (SLN-DCC 4 in Figure 59), this cell showed clear copy number alterations also detected in different tumor cells of this patient (see appendix Figure 81). The signal of the UV signature was strong in metastastic samples, the cell line and in five of 15 SLN-DCCs (Figure 53 A). In the remaining SLN-DCCs the UV signature was still detectable, but weaker. However, the expression of the UV signature did not correlate with the number of detected mutations as weak UV signatures were also found in samples with a high number of mutations (Figure 53 A, log10dsmut). When the Platypus algorithm was applied the UV signature 7 was stronger, but the total number of mutations decreased (see appendix Figure 80 A, log10dsmut). The presence of a UV signature in the tumor samples from MM15-127 indicated that corresponding parts of the body where the tumor tissues were removed (Figure 53 B) were exposed to UV light creating characteristic UV-associated mutations. The second most prominent signature mainly found in UV signature weak samples was signature 5, which is of unknown aetiology.

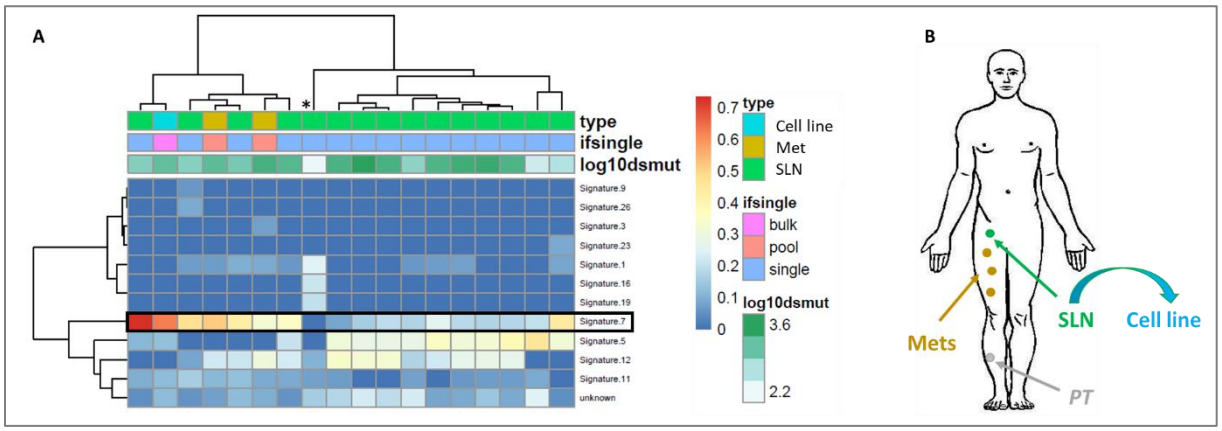


Figure 53: Mutational signature analysis of MM15-127 using Varscan
A: Heatmap of mutational signature analysis of MM15-127 using the Varscan algorithm. The UV signature 7 is marked with a black frame. Expression levels ranged from blue marked 0 to red marked 0.7. The number of detected mutations is represented by log10dsmut ranging from 2.2 to 3.6. The SLN-DCC4, for which no UV signature was detected, is marked with *. **B:** Body map of corresponding tumor tissues. DNA from the PT was not used for WES due to quality reasons. PT: primary tumor; SLN: sentinel lymph node; Mets: metastases; WES: Whole exome sequencing.

Next, the mutational signatures of samples from patient MM16-412 were analysed using bulk DNA (primary tumor, cell line and metastasis 1, 2 and 4), single cells (10 SLN-DCCs and two non-tumor cells) and pools of single cells (NSLN 1-6 and metastasis 1 and 4). Using Varscan, UV signatures were detected in almost all tumor samples, but not in the two single non-tumor cells serving as controls (Figure 54 A). Only for one SLN-DCC no UV signature could be found (Figure 54 A, marked with *). In addition, this cell only showed small deletions on chromosome 5, 10 and 15 and did not harbor characteristic CNAs that were present in other tumor cells of this patient (see appendix Figure 82). Moreover, for this cell a very low number of mutations were detected compared to the remaining tumor cells (Figure 54 A, log10dsmut). Thus, it was not clear whether this cell represented the earliest tumor cell of this patient or whether it was a normal cell. Strinkingly, this cell dropped out from MS-QC and was therefore not found in the MS-based lineage tree. When the Platypus algorithm was applied the UV signature 7 was stronger, but in total less mutations were detected (see appendix Figure 80 B, log10dsmut). The detection of UV signatures in almost all tumor samples from MM16-412 showed that corresponding parts of the body were probably exposed to UV light leading to UV-related mutations (Figure 54 B). However, there was no correlation between the UV signature expression and the number of detected mutations as strong UV signatures were also detected in samples with an intermediate number of mutations (Figure 54 B, log10dsmut). The strongest signature detected in control cells was signature 5, which is of unkown origin.

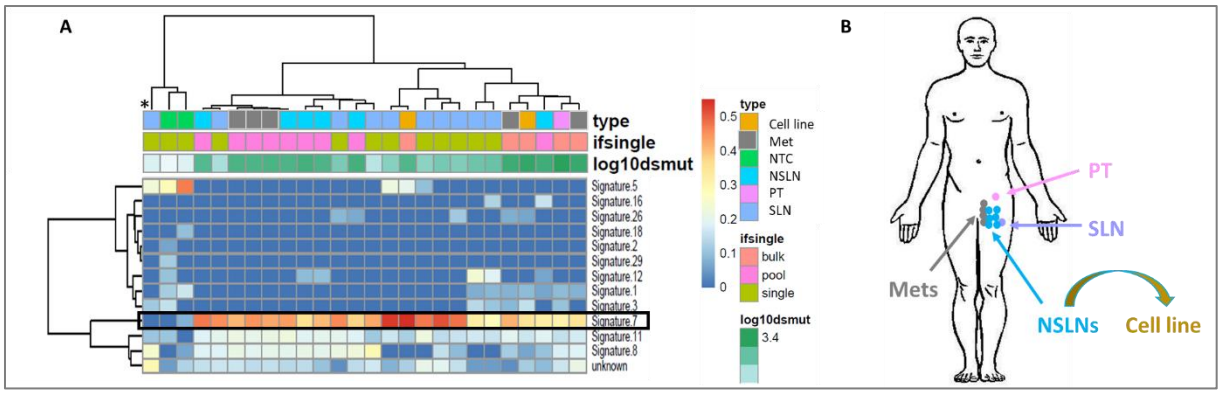


Figure 54: Mutational signature analysis of MM16-412 using Varscan
A: Heatmap of mutational signature analysis of MM16-412 using the Varscan algorithm. The UV signature 7 is marked with a black frame. Expression levels ranged from blue marked 0 to red marked 0.5. The number of detected mutations is represented by log10dsmut ranging up to 3.4. The tumor sample (SLN-DCC10), for which no UV signature was found, is marked with *. **B:** Body map of corresponding tumor tissues for WES. PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; Mets: metastases; NTC: non-tumor cell; WES: Whole exome sequencing.

At last, the mutational analysis was performed with samples from patient MM16-423. Here, bulk DNA from the cell line, single cells (three SLN-DCCs, five CTCs and two non-tumor cells) and pools of single cells from NSLN 1-6 were used. For patient MM16-423 no UV signature 7 was found in the tumor samples using Varscan although showing high numbers of mutations (Figure 55 A, log10dsmut). Only one SLN-DCC showed a weak expression of signature 7 (Figure 55 A, marked with *). Even with the Platypus algorithm no signature 7 was found for the tumor samples (see appendix Figure 80 C). That no UV signatures were found for MM16-423 might be explained by the fact that the primary tumor was located on the heel of the patient and thus most likely protected from UV light (Figure 55 B). The two strongest signatures detected in most of the samples were signature 5 and 12, which are both of unkown aetiologies. However, signature 12 was reported to be associated with liver cancers (Letouzé et al. 2017) consistent with the patients disease progression developing liver metastases.

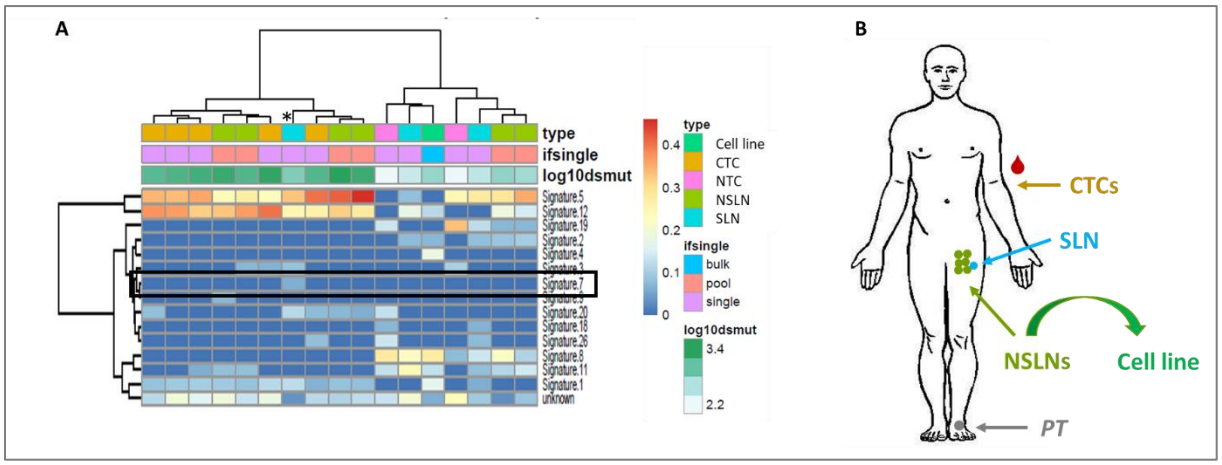


Figure 55: Mutational signature analysis of MM16-423 using Varscan
A: Heatmap of mutational signature analysis of MM16-423 using the Varscan algorithm. The UV signature 7 is marked with a black frame. Expression levels ranged from blue marked 0 to red marked 0.4. The number of detected mutations is represented by log10dsmut ranging from 2.2 to 3.4. The tumor sample with a weak UV signature is marked with *. **B:** Body map of corresponding tumor tissues. DNA from the PT was not used for WES due to quality reasons. PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; CTCs: circulating tumor cells isolated from blood; NTC: non-tumor cell; WES: Whole exome sequencing.

4.3.3.2 Shared exonic mutations

Next, exonic mutations shared between the different tumor samples were analysed for the three patients. For this, only mutations were considered that did not appear in non-tumor controls.

MM15-127

For patient MM15-127, all mutations of the cell line or metastases were also present in cells from the SLN (Figure 56 A). With SLN samples harboring 1395 mutations and the cell line or metastases 261 and 440 mutations, respectively, a higher mutational load of the SLN samples indicated a higher clonal heterogeneity compared to the cell line or metastatic samples. Moreover, 229 from 440 mutations, which were found in the metastases (Met1 and Met3), were shared between both metastatic samples (Figure 56 B). Metastases, cell line and SLN samples shared in total 216 mutations reflecting 15% of all mutations (Figure 56 A). As SLN samples comprised 15 single DCCs, the number of mutations shared between the different tumor tissues was analysed depending on the number of SLN samples in which they appeared. The number of mutations shared between all tumor tissues (SLN, MET and cell line) was similar to the number of mutations only found in SLN samples, when they appeared in more than four SLN-DCCs (Figure 56 C, similar height of orange and pink bar). This means that if mutations were shared by more than four SLN samples the same number of mutations was found in all tumor tissues.

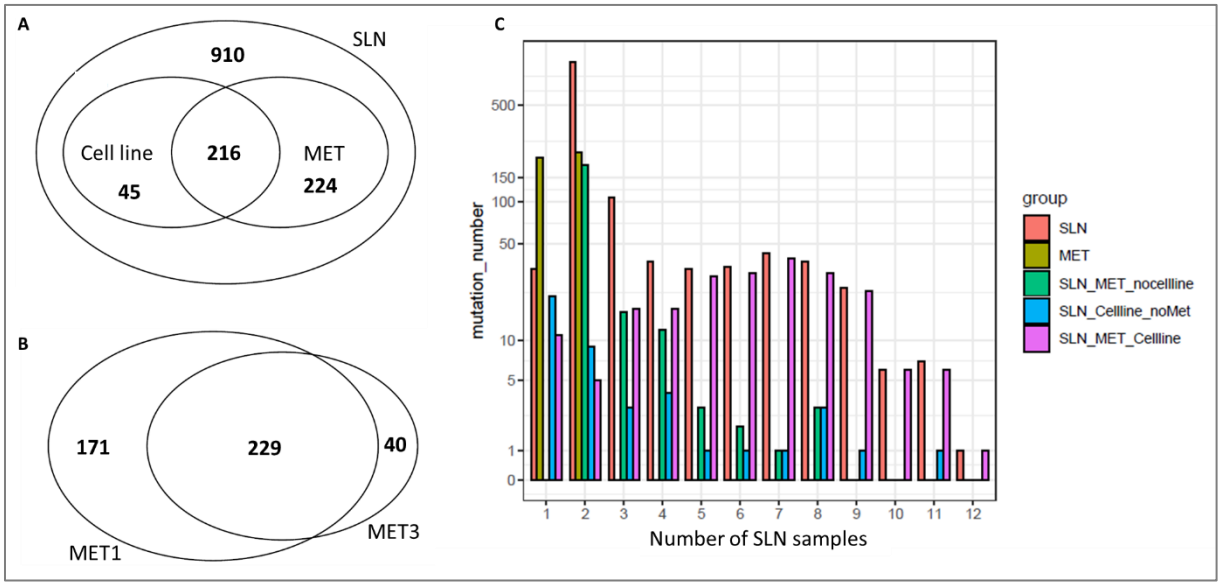


Figure 56: Overlap of exonic mutations for MM15-127
Overlap of exonic mutations of SLN, cell line and metastases **(A)** and between two metastatic samples **(B)**. **C:** SLN mutation sample frequency. Y-axis shows the number of mutations for different tissue groups, whereas x-axis describes in how many SLN samples the mutations were found. If the mutations were shared by more the four SLN samples the amount of mutations shared between all tumor tissues was similar to the amount of mutations only found in SLN samples. SLN: sentinel lymph node; METs: metastases.

MM16-412

Then, shared exonic mutations were analysed for the different samples of MM16-412. In both non-tumor control cells, 15 exonic mutations were found with two overlapping and six or seven private mutations (Figure 57 A). For the different tumor tissues, a similar mutational load was detected by identifying 268 exonic, tumor-specific mutations in samples from the SLN, 274 in samples from NSLNs, 311 in samples from cutaneous metastases and 314 exonic mutations within the primary tumor. Of all mutations, 43% [171/394] were shared between all tumor samples (Figure 57 B). Next, the overlap of exonic mutations between each tumor tissue and metastases were compared. Between the SLN and metastases 219 mutations were shared, 235 between NSLNs and metastases and 274 between the primary tumor and metastases, indicating a similar pattern of mutational overlap between different

tumor tissues and the metastatic samples. Interestingly, more than half of mutations from SLN-DCCs [171/268] were also found in other tumor samples from the different tissues, suggesting that SLN-DCCs might cover the mutational spectrum of later-arising and metastases. From this patient several metastatic samples were analysed enabling the evaluation of not only the mutational overlap between different metastases, but also between bulk samples and pools of single cells from the same metastatic tissue. Exonic mutations were shared to a high degree between bulk samples and pools of single cells from the same metastatic tissue (Figure 57 D and E). In addition, bulk samples showed more mutations in total than pools of single cells from the same metastasis. However, with bulk and pools sharing around 60 to 70% of mutations (66% for Met1 and 58% for Met4), pools of single cells sequenced at 200x did not completely reflect the entirety of the clonal repertoire of the bulk samples. Sequencing at 500x for bulk samples also did not detect all mutations, but this effect was weaker than sequencing of pools. This suggested that it would be better to sequence bulk samples instead of pools of single cells, if enough material is available, as more mutations were detected. Next, the mutational overlaps between different metastatic tissues were compared. As simultaneously removed metastasis 1 and 2 were mainly found within one big cluster of the MS-based tree (Figure 38) and melanoma marker-negative metastasis 4 clustered separately, the question was whether metastasis 1 and 2 also shared more mutations in comparison to metastasis 4. In this case, pools of single cells were considered as from metastasis 2 no bulk sample was available for WES. Of all mutations from the metastatic samples 65% [169/233] were shared between the three metastases showing a high mutational overlap (Figure 57 C). However, there was no difference in the degree of shared mutations between metastases 1 and 2 compared to between either metastases 1 or 2 and later removed metastasis 4 (Figure 57 C). Thus, the mutational analysis did not add to understand their different locations within the MS-based tree. In summary, a high number of mutations were shared between all tumor tissues and between different cutaneous metastases indicating a large overlap of the mutational spectrum of this patient.

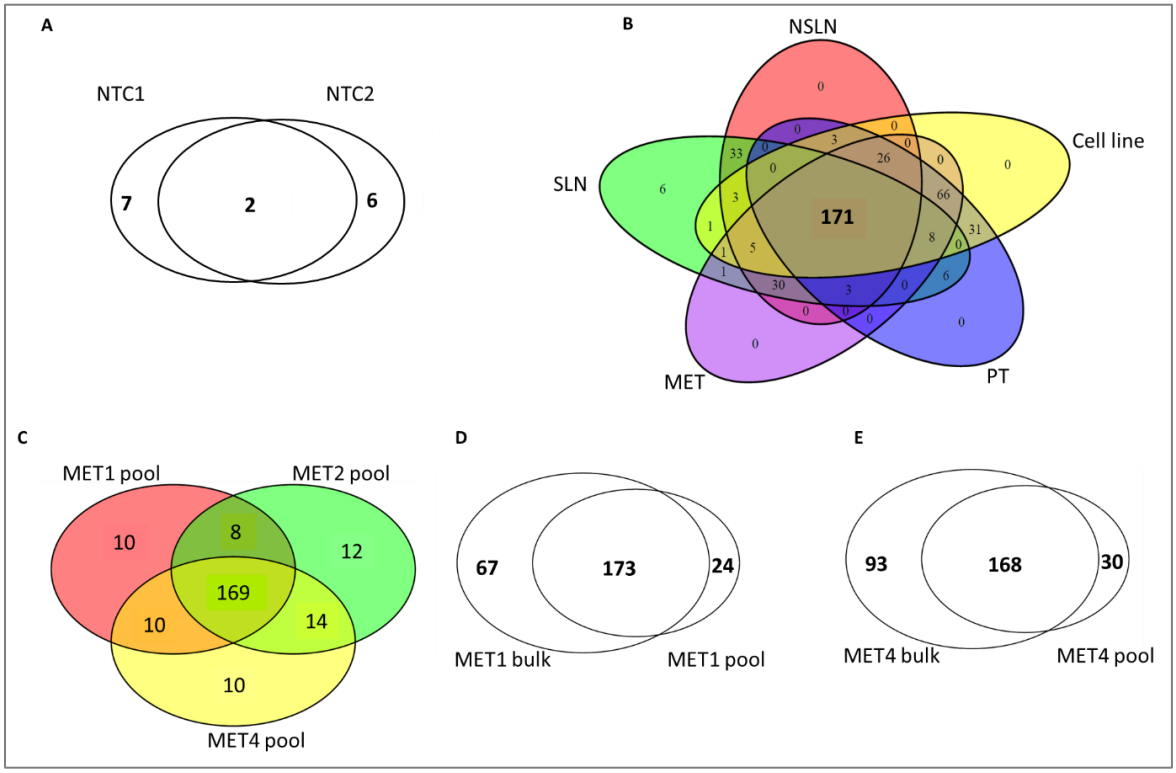


Figure 57: Overlap of exonic mutations for MM16-412
A: Overlap of exonic mutations of single non-tumor controls (NTCs). **B:** Overlap of exonic mutations of PT, SLN, NSLN, Mets and cell line. **C:** Overlap of exonic mutations of pools of single cells from three metastatic samples. **D:** Overlap of exonic mutations of bulk and pool of single cells from metastasis 1. **E:** Overlap of exonic mutations of bulk and pool of single cells from metastasis 4. PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; METs: metastases.

MM16-423

Next, the mutational spectrum of different samples from patient MM16-423 was evaluated. Both non-tumor controls showed in total 16 exonic mutations sharing seven (Figure 58 A). In the tumor samples, the highest number of exonic, tumors-specific mutations was found in CTCs and the lowest mutational load was detected in bulk DNA from the cell line (Figure 58 B). Exonic mutations were shared to a high degree between CTCs and NSLN-DCCs, whereas only 2% [11/544] of all mutations were shared between all tumor samples. However, 11 of 98 mutations detected in SLN-DCCs (11%) were also found in all other tumor samples including the cell line and 79 mutations from all 98 SLN mutations (81%) were shared with NSLNs and CTCs. These findings indicated that some mutations of CTCs and NSLN-DCCs, removed at a later time point in disease progression, were already present in early SLN-DCCs.

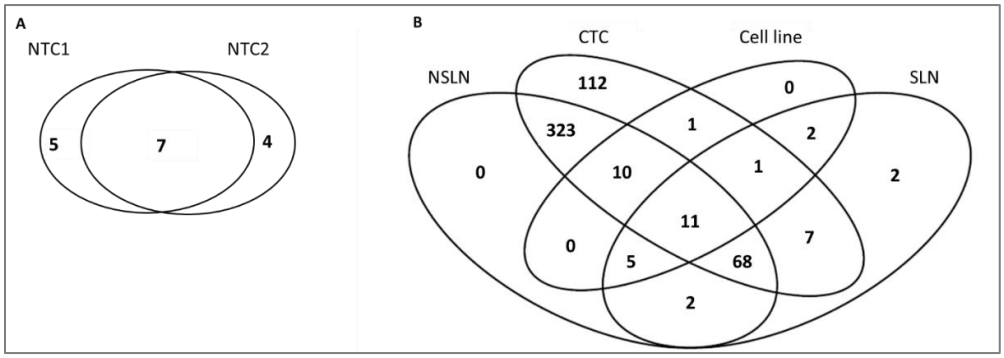


Figure 58: Overlap of exonic mutations for MM16-423
A: Overlap of exonic mutations of single non-tumor contols (NTCs). **B:** Overlap of exonic mutations of SLN, cell line, CTCs and NSLNs. SLN: sentinel lymph node; CTCs: circulating tumor cells; METs: metastases.

4.3.3.3 Predicted class I peptide load

In the last step, HLA-I binding affinities of peptides created by identified exonic mutations of the tumor samples was predicted using NetMHCpan taking the information on class I HLA alleles from the patients into account. Only 8 to 11 amino acids long peptides were considered as this peptide length is presented in context of HLA-I. The predicted potential neopeptides were categorized into strong binders (SB) or weak binders (WB) depending on their %Rank, which describes the rank of predicted affinity in comparison to a set of 400000 random natural peptides. Strong binders were defined by having a %Rank smaller than 0.5 and weak binders by a %Rank between 0.5 and 2.

MM15-127

In total, 5190 different peptides were predicted from 1386 exonic mutations. Less than 1000 predicted peptides were found in seven SLN-DCCs, between 1000 and 2000 in five SLN-DCCs and three SLN-DCCs showed more than 2000 predicted peptides (Table 46). Of those 15 single SLN-DCCs 11 cells were also considered in phylogenetic analyses and therefore present in MM15-127's lineage tree using microsatellite markers (Figure 59 A). SLN-DCC4, for which no UV signature was detected, showed the lowest number of predicted peptides (Table 49). In addition, SLN-DCC4 had an aberrant CNA profile (Figure 53 A (*), appendix Figure 81) and was indeed located early in the MS-based tree (Figure 59 A), within the control cell cluster, indicating that this SLN-DCC might represent the earliest cancer cell. However, the location within the tree of SLN-DCCs having more than 100 predicted peptides did not correlate with the number of predicted peptides as for example SLN-DCC3 allowed prediction of a very high number of peptides, but was located at the beginning of the tumor cell cluster. In pools of single cells from metastatic samples 1167 and 1758 peptides were predicted, whereas bulk DNA from the cell line allowed detection of 1111 peptides (Table 46). In average 30% of predicted peptides were strong binders with high HLA-I binding afffinities (Table 46). With regard to the binding affinities of predicted peptides to the six different HLA-I alleles, the proportion of strong binders ranged from 15

to 31% (Figure 59 B). The highest number of predicted peptides showed binding affinities to HLA-A*68:01 and HLA-B*07:02 with both alleles also revealing the highest percentages of strong binders.

Table 46: Number of class I predicted peptides of MM15-127

Sample	Total	WB	SB	SB/Total [%]
SLN-DCC 4	10	5	5	0.50
SLN-DCC 14 *	171	122	49	0.29
SLN-DCC 6	241	169	72	0.30
SLN-DCC 11	556	386	170	0.31
SLN-DCC 15 *	557	403	154	0.28
SLN-DCC 7	694	477	217	0.31
SLN-DCC 10	713	494	219	0.31
Cell line (bulk)	**1111**	**772**	**339**	**0.31**
Met3 (pool)	1167	823	344	0.29
SLN-DCC 5	1426	996	430	0.30
SLN-DCC 8	1531	1093	438	0.29
SLN-DCC 13	1677	1194	483	0.29
SLN-DCC 17 *	1721	1216	505	0.29
Met1 (pool)	1758	1256	502	0.29
SLN-DCC 16 *	1791	1280	511	0.29
SLN-DCC 3	2057	1452	605	0.29
SLN-DCC 9	2077	1493	584	0.28
SLN-DCC 12	3186	2290	896	0.28

Note: SLN-DCCs marked with * dropped out from MS-based lineage tree analysis. WB: weak binder, SB: strong binder (%Rank < 0.5). Samples are ranked according to their number of total predicted peptides.

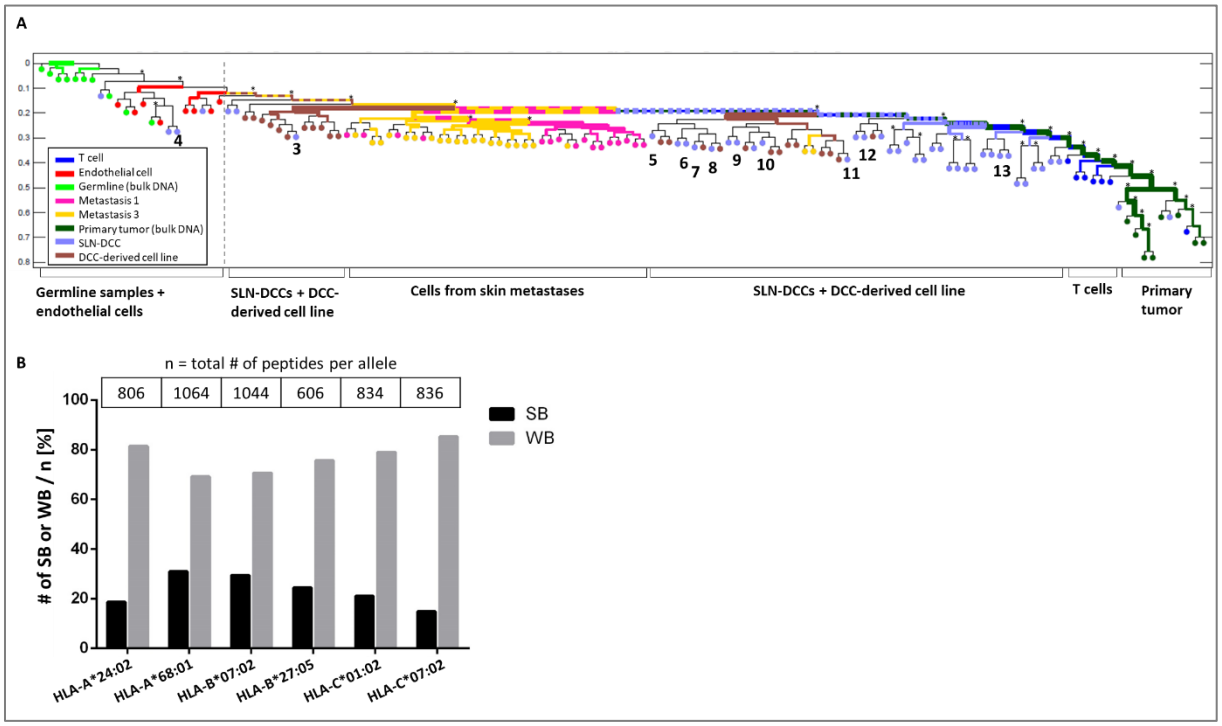


Figure 59: MM15-127's lineage tree with WES samples and predicted peptide analysis
A: MS-based lineage tree of MM15-127. SLN-DCCs that were analysed by whole exome sequencing (WES) are labeled in phylogenetic tree with 3-13. **B:** Peptide analysis predicted for the different HLA-I alleles. #: number; SB: strong binder; WB: weak binder.

MM16-412

Next, potential neopeptides were predicted for each tumor sample of MM16-412 identifying in total 1316 peptides from 383 exonic mutations. Most predicted peptides were found in bulk DNA of the primary tumor (Table 47). In bulk DNA from the two patient-derived cell lines 922 and 1179 peptides were predicted. In pools of single cells and bulk samples of metastatic tissues between 864 and 1105 peptides were predicted, whereas pools of NSLN-DCCs revealed 584 to 1102 predicted peptides (Table 47). All ten single SLN-DCCs showed less than 700 predicted peptides ranging from 62 to 672. The fact that in individual SLN-DCCs less peptides were predicted than in the primary tumor was due to having single cells compared to bulk sample. Both tissues, the SLN and the primary tumor, showed a similar mutational load (Figure 57 B). The SLN-DCC with the lowest number of predicted peptides was SLN-DCC10, for which no UV signature was found, probably due to the low mutational load (Figure 53 A, log10dsmut). However, one could not clearly say whether SLN-DCC10 that dropped out from the MS-based lineage analysis was an early tumor cell with few exonic mutations and few CNAs (see appendix Figure 82). Next, it was analysed whether the number of predicted peptides correlated with the tumor burden of the different lymph nodes reflected by the disseminated cancer cell density (DCCD). A significant correlation between the number of predicted peptides and the tumor burden of the lymph nodes was found (Figure 60 B; $p < 0.0001$). Cells from lymph nodes with a higher tumor burden also allowed prediction of more peptides indicating a higher mutational load. In addition, six of the ten single SLN-DCCs with more than 400 predicted peptides were included in the reconstruction of the cell lineage of the patient using microsatellite markers (Figure 60 A). From the remaining four SLN-DCCs (SLN-DCC10/11/12/13) that dropped out from MS-based lineage reconstruction CNA analysis was performed (see appendix Figure 82 A to D). Interestingly, SLN-DCC10, 11 and 12 showed fewer CNAs than SLN-DCC13 that had characteristic CNAs also detected in other tumor cells of this patient (see appendix Figure 82 E). Regarding SLN-DCCs included in the MS-based lineage tree, their location did not correlate with the number of predicted peptides as for example SLN-DCC1 showed more predicted peptides than SLN-DCC3/4/5/6 but was located at the beginning of the mixed DCC branch (Figure 60 A). HLA-I typing revealed a loss of heterozygosity (LOH) for some HLA-I alleles in several samples of this patient. The number of predicted peptides significantly increased with decreasing number of HLA-I alleles (Figure 60 C; $p = 0.0006$). Consequently, a high mutational load accompanied by a high number of predicted peptides indicated more immune pressure and HLA-LOH. Of all predicted peptides, 29% showed high binding affinities to HLA-I determined as strong binders. The percentage of strong and weak binding peptides from peptides binding the different HLA-I alleles ranged from 10 to 33% (Figure 60 D). Notably, within the different HLA-I alleles more peptides were predicted to bind to those HLA-I alleles that were not found in all tumor samples most likely due to LOH (Figure 60 D, red alleles). Moreover, high percentages of strong binders (24% and 33%) were also detected within peptides predicted to bind to those HLA-I alleles that were lost in some tumor samples pointing towards immune pressure as potential reason for LOH.

Table 47: Number of class I predicted peptides of MM16-412

Sample	DCCD	HLA alleles	Total	WB	SB	SB/Total [%]
SLN-DCC 10 *	9.5	5	62	42	20	0.32
SLN-DCC 11 *	9.5	5	137	98	39	0.28
SLN-DCC 12 *	9.5	5	240	162	78	0.33
SLN-DCC 13 *	9.5	5	409	283	126	0.31
SLN-DCC 3	9.5	2	439	321	118	0.27
SLN-DCC 4	9.5	3	451	322	129	0.29
SLN-DCC 6	9.5	1	465	331	134	0.29
SLN-DCC 5	9.5	3	515	368	147	0.29
SLN-DCC 1	9.5	4	531	378	153	0.29
NSLN6 (pool)	50 000	3	584	413	171	0.29
SLN-DCC 7	9.5	0	672	476	196	0.29
NSLN5 (pool)	258	3	719	512	207	0.29
Met1 (pool)	--	3	864	600	264	0.31
Met4 (pool)	--	3	919	649	270	0.29
Cell line a (bulk)	--	3	922	662	260	0.28
NSLN4 (pool)	750 000	3	961	681	280	0.29
Met2 (pool)	--	3	968	688	280	0.29
NSLN3 (pool)	750 000	3	1028	729	299	0.29
Met1 (bulk)	--	(6)**	1059	746	313	0.30
NSLN1 (pool)	750 000	3	1070	754	316	0.30
NSLN2 (pool)	500 000	3	1102	782	320	0.29
Met4 (bulk)	--	(6)**	1105	795	310	0.28
Cell line b (bulk)	--	3	1179	842	337	0.29
PT (bulk)	--	(6)**	1342	973	369	0.27

Note: SLN-DCCs marked with * dropped out from MS-based lineage tree analysis. ** 6 alleles in bulk samples of Met1/4 were detected due to TME contamination. Also for PT, a TME contamination could not be excluded. TME: tumor microenvironment including non-tumor cells. DCCD: disseminated cancer cell density; WB: weak binder, SB: strong binder (%Rank < 0.5). Samples are ranked according to their number of total predicted peptides.

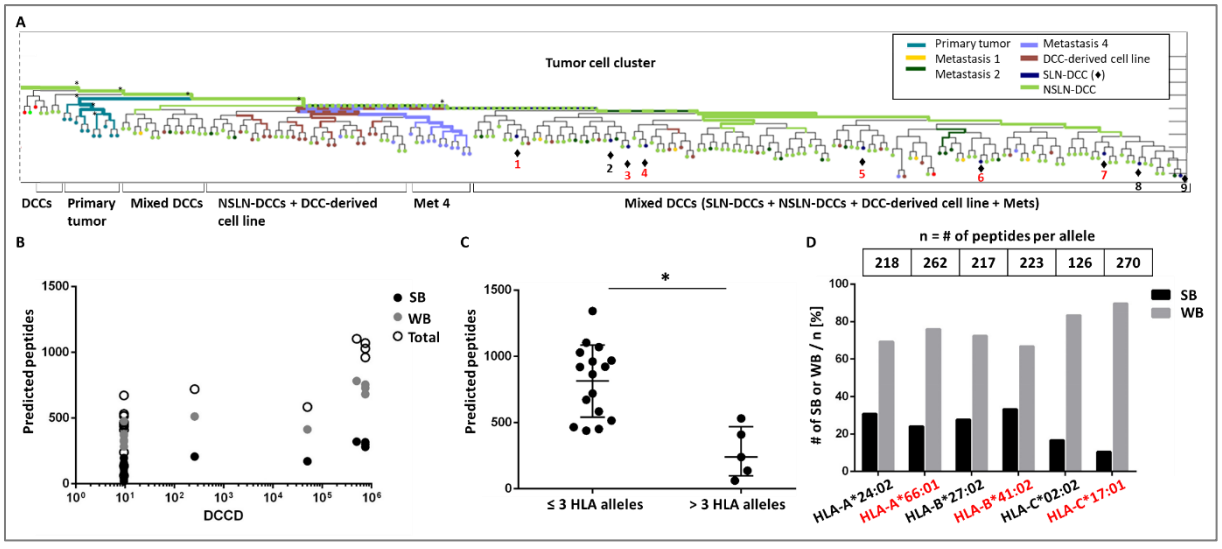


Figure 60: MM16-412's lineage tree with WES samples and predicted peptide analysis
A: Tumor cell cluster of MM16-412's cell lineage reconstruction. All nine SLN-DCCs are labelled with ♦. SLN-DCCs that were analysed by whole exome sequencing (WES) are labeled in red color. **B:** Correlation of DCCD and predicted peptides of lymph node samples from MM16-412 ($p < 0.0001$ using Pearson correlation). **C:** Number of predicted peptides and number of HLA-I alleles, without bulk samples of PT and Met1+4, due to TME contamination. TME: tumor microenvironment including non-tumor cells. Mean and standard deviation are shown ($p = 0.0006$ using unpaired t-test). **D:** Predicted peptide analysis for the different HLA-I alleles. Alleles not detected in some tumor samples by HLA typing are marked in red. #: number; SB: strong binder; WB: weak binder.

MM16-423

Potential class I neopeptides were predicted for tumor samples of MM16-423. In total 1643 peptides were predicted from 484 exonic mutations. Most peptides were predicted in pools of single cells from NSLNs ranging from 138 to 1522, followed by single CTCs with 615 to 1441 predicted peptides (Table 48). For the SLN-DCCs, 83 to 378 peptides were predicted. All three SLN-DCCs and two of five CTCs were included in the cell lineage reconstruction using microsatellite markers (Figure 61 A). The fewest predicted peptides were detected in SLN-DCC1 that was located within the control cell cluster of the phylogenetic tree (Figure 61 A). This cell also showed CNAs (Figure 42) supporting the finding that it represented the earliest tumor cell. Moreover, analogous to MM16-412 a significant correlation between the number of predicted peptides and the lymph node tumor burden reflected by the disseminated cancer cell density was detected (Figure 61 B; $p = 0.0442$). For bulk DNA of the cell line a relatively low number of peptides (Table 48) was predicted in line with the low number of detected exonic mutations (Figure 58), indicating a high clonality and low heterogeneity of the cell line. For all samples, in average 31% of predicted peptides were strong binders with high affinities to HLA-I alleles (Table 48). Regarding the binding affinities to the different HLA-I alleles the percentage of strong binders ranged from 17 to 31% showing a similar pattern of distribution (Figure 61 C).

Table 48: Number of class I predicted peptides of MM16-423

Sample	DCCD	Total	WB	SB	SB/Total [%]
SLN-DCC 1	2	83	56	27	0.33
SLN-DCC 3	2	86	54	32	0.37
Cell line (bulk)	**--**	**113**	**82**	**31**	**0.27**
NSLN3 (pool)	10	138	93	45	0.33
NSLN5 (pool)	42	178	118	60	0.34
SLN-DCC 2	2	378	261	117	0.31
CTC4 *	--	615	433	182	0.30
CTC3 *	--	691	507	184	0.27
CTC7 *	--	793	548	245	0.31
NSLN4 (pool)	58	886	593	293	0.33
CTC5	--	1074	784	290	0.27
NSLN2 (pool)	50 000	1080	755	325	0.30
NSLN6 (pool)	138	1159	824	335	0.29
CTC6	--	1441	1026	415	0.29
NSLN1 (pool)	90 0000	1522	1093	429	0.28

Note: CTCs marked with * dropped out from MS-based lineage tree analysis. DCCD: disseminated cancer cell density; WB: weak binder, SB: strong binder (%Rank < 0.5). Samples are ranked according to their number of total predicted peptides.

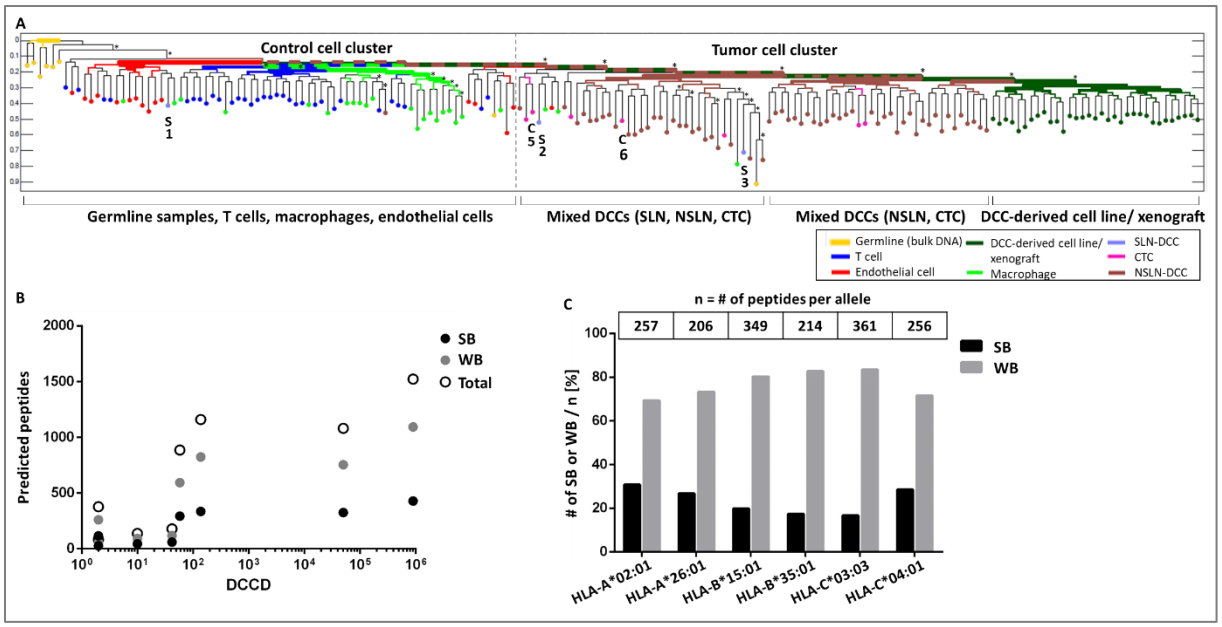


Figure 61: MM16-423's lineage tree with WES samples and predicted peptide analysis
A: MS-based lineage tree of MM16-423. SLN-DCCs and CTCs that were analysed by whole exome sequencing (WES) are labeled the phylogenetic tree with S1, S2, S3 and C5, C6. **B:** Correlation of DCCD and predicted peptides of lymph node samples from MM16-423 ($p = 0.0442$ using Pearson correlation). **C:** Predicted peptide analysis for the different HLA-I alleles. #: number; SB: strong binder; WB: weak binder.

Peptide selection for functional evaluation

In order to select candidate peptides for functional evaluation of identified potential neoantigens, predicted peptides were ranked according to their HLA-I binding affinity. Only strong binding peptides were considered that were shared between all tumor tissues. For all patients predicted strong binding peptides deriving from five different exonic non-synonymous mutations are shown as examples.

For patient MM15-127, predicted peptides shared between the SLN, cutaneous metastases and the cell line were detected in four to eight from in total 15 single SLN-DCCs (Table 49). For one exonic mutation two different peptides were predicted with high binding affinities. Moreover, the majority of those predicted peptides showed high binding affinities to more than one HLA-I allele.

Table 49: List of shared strong binding class I peptides from MM15-127

#	Peptide	Chr	Allele position	Ref base	Alt base	Gene name	HLA	%Rank	# Samples			
									SLN	Met	CL	Sum
1	SPRTALSAF	22	31861414	C	T	DEPDC5	HLA-B*07:02	0.0071	4/15	2/3	1/1	9/19
	FVRSPRTAL	22	31861414	C	T	DEPDC5	HLA-C*01:02	0.0312	4/15	2/3	1/1	9/19
							HLA-C*07:02	0.4613				
2	FVFGGPPSL	16	84976452	G	A	ZDHHC7	HLA-C*01:02	0.0099	4/15	1/3	1/1	8/19
							HLA-C*07:02	0.0831				
3	ESAFVSYEK	7	82949664	G	A	PCLO	HLA-A*68:01	0.0085	7/15	2/3	1/1	10/19
4	YSPWMPSSL	6	34133427	C	T	GRM4	HLA-C*01:02	0.0092	8/15	2/3	1/1	11/19
							HLA-C*07:02	0.2719				
5	YSRPHSHLL	11	18286595	G	A	HPS5	HLA-C*01:02	0.0095	7/15	2/3	1/1	10/19
							HLA-C*07:02	0.0503				
							HLA-B*07:02	0.4265				

Chr: chromosome; Ref: reference; Alt: alternative; %Rank < 0.5 = strong binder; # Samples: number of tumor samples in which the peptide appeared; SLN: sentinel lymph node; Met: metastasis; CL: cell line. Samples are ranked according to their %Rank.

For patient MM16-412, predicted strong binders were shared between 13 and 18 tumor samples of the different tissues, i.e. the primary tumor, the SLN, NSLNs, cutaneous metastases and the cell lines (Table 50). Shared predicted peptides were found in at least one and in up to four single SLN-DCCs and in five or six NSLN samples. For one exonic mutation, two different peptides were predicted with high binding affinities to the same HLA-I allele. Moreover, two strong binders showed high binding affinities

to an HLA-I allele that was not found in some tumor samples (Figure 48; Table 50, marked in bold), suggesting that some tumor samples might have lost those HLA-I alleles as a result of immune evasion.

Table 50: List of shared strong binding class I peptides from MM16-412

#	Peptide	Chr	Allele position	Ref base	Alt base	Gene name	HLA	%Rank	# Samples					
									PT	SLN	NSLN	Met	CL	Sum
1	KYVKVFYKF	19	22392745	G	A	ZNF98	HLA-A*24:02	0.0064	1/1	3/10	6/6	5/5	2/2	17/24
2	SYFSFIFPNF	3	75730493	C	A	ZNF717	HLA-A*24:02	0.0081	1/1	2/10	5/6	5/5	1/2	14/24
	HFWPSYFSF	3	75730493	C	A	ZNF717	HLA-A*24:02	0.0188	1/1	2/10	6/6	5/5	1/2	15/24
3	KELLLTASL	19	43269382	C	T	PSG9	**HLA-B*41:02**	0.0082	1/1	1/10	6/6	3/5	2/2	13/24
4	SRLQWVAS LW	16	77293086	G	A	ADAMTS 18	HLA-B*27:02	0.0097	1/1	2/10	5/6	5/5	2/2	15/24
5	KELHYASVV	17	74484013	G	A	CD300A	**HLA-B*41:02**	0.0099	1/1	4/10	6/6	5/5	2/2	18/24

Chr: chromosome; Ref: reference; Alt: alternative; %Rank < 0.5 = strong binder; # Samples: number of tumor samples in which the peptide appeared; PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentiel lymph node; Met: metastasis; CL: cell line. Samples are ranked according to their %Rank. HLA alleles lost in tumor samples of MM16-412 are marked in bold.

For MM16-423, peptides predicted from five different mutations were shared between five and 11 tumor samples and were predicted in one to three SLN-DCCs, up to four pools of NSLN-DCCs and in up to four CTCs (Table 51). Notably, two mutations led to the prediction of three and four different strong binding peptides. Four predicted peptides showed high binding affinities to more than one HLA-I allele.

Table 51: List of shared strong binding class I peptides from MM16-423

#	Peptide	Chr	Allele position	Ref base	Alt base	Gene name	HLA	%Rank	# Samples				
									SLN	NSLN	CTC	CL	Sum
1	GLLSFFFAV	7	22493414	C	G	STEAP1B	HLA-A*02:01	0.0149	1/3	2/6	3/5	1/1	7/15
1b	KQFGLLSFF	7	22493414	C	G	STEAP1B	HLA-B*15:01	0.0170	1/3	2/6	3/5	1/1	7/15
1c	RKQFGLLSFF	7	22493414	C	G	STEAP1B	HLA-B*15:01	0.0361	1/3	2/6	3/5	1/1	7/15
1d	KQFGLLSFFF	7	22493414	C	G	STEAP1B	HLA-B*15:01	0.0697	1/3	2/6	3/5	1/1	7/15
2a	YAMMRSYRY	7	22493367	C	A	STEAP1B	HLA-B*35:01	0.0195	2/3	4/6	4/5	1/1	11/15
							HLA-A*26:01	0.2397					
							HLA-B*15:01	0.3855					
2b	LSYAMMRSY	7	22493367	C	A	STEAP1B	HLA-B*15:01	0.0931	2/3	4/6	4/5	1/1	11/15
							HLA-B*35:01	0.4272					
							HLA-A*26:01	0.4424					
2c	TLSYAMMRSY	7	22493367	C	A	STEAP1B	HLA-B*15:01	0.2881	2/3	4/6	4/5	1/1	11/15
							HLA-A*26:01	0.4555					
3	YVNGGWSSL	10	71287619	G	T	UNC5B	HLA-C*03:03	0.0393	1/3	1/6	3/5	1/1	6/15
							HLA-A*26:01	0.2268					
4	FPGNCGTAL	2	174572129	G	A	WIPF1	HLA-B*35:01	0.0738	1/3	1/6	2/5	1/1	5/15
5	NVYVNTAHL	11	88312415	A	G	CTSC	HLA-C*03:03	0.4186	3/3	4/6	3/5	1/1	11/15

Chr: chromosome; Ref: reference; Alt: alternative; %Rank < 0.5 = strong binder; # Samples: number of tumor samples in which the peptide appeared; SLN: sentinel lymph node; NSLN: non-sentiel lymph node; CTC: circulating tumor cell; CL: cell line. Samples are ranked according to their %Rank.

In summary, analysing a small set of shared predicted peptides, strong binders of SLN-DCCs were also predicted in later-arising metastatic samples for all patients. Whether predicted peptides were indeed neopeptides leading to an immune response and thus would have been suitable for adjuvant immunotherapy, potential neoantigens needed to be evaluated functionally.

4.4 Test system for functional validation of shared tumor-specific mutations

To assess whether shared non-synonymous mutations identified by WES are indeed neoantigens that can elicit an immune response, they have to be tested for their immunogenicity. Only if neopeptides are presented on the surface of the tumor cells -a prerequisite for recognition by T cells and anti-tumor immunity-, identified neoantigens are therapeutically relevant. In order to address this, a test system was established to detect neoantigen-specific T cells among tumor-reactive T cells in the blood of the patient using biobanked cell line and immune cells (Figure 62). Due to the low frequency of antigen-specific T cells, tumor-reactive T cells were exposed to autologous tumor cells for *in vitro* expansion. The pool of expanded T cells could contain both, T cells specific for tumor-associated antigens and T cells specific for tumor-specific antigens, i.e. neoantigens. To identify neoantigen-specific T cells, expanded T cells should be exposed to autologous antigen-presenting cells electroporated with *in vitro*

transcribed RNA coding for up to 25 candidate neoantigens ('minigene') and test for secretion of Interferon-γ (IFN-γ). This analysis will inform whether candidate neoantigens could have been of potential therapeutic relevance for adjuvant therapy of early systemic cancer.

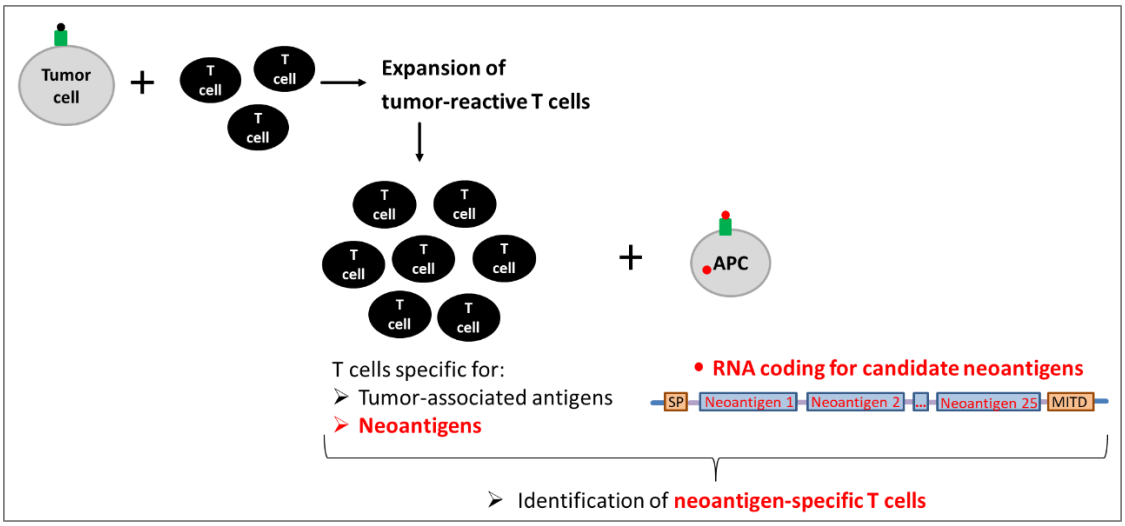


Figure 62: Assay for detection of neoantigen-specific T cell responses
Tumor-reactive T cells are expanded by stimulating them with tumor cells, autologous cell line cells. Within expanded tumor-reactive T cells, neoantigen-specific T cells should be identified. Thus, tumor-reactive T cells are co-cultured together with antigen-presenting cells (APCs) electroporated with *in vitro* transcribed RNA coding for up to 25 identified neoantigens. Candidate neoantigens within the minigene are separated by linker sequences and flanked by a MHC class I signal peptide fragment (SP) and a MHC class I trafficking signal (MITD) for optimized peptide presentation and improved translational efficiency. In the end, neoantigen-specific T cells are identified by Interferon-γ production.

4.4.1 Analysis of patient-derived cell lines

For functional validation of neoantigens, optimal (neo)peptide presentation on the HLA complexes of the tumor cells was required. Therefore, autologous DCC-derived cell line cells were analysed for their surface marker expression in context of HLA. To induce upregulation of (neo)peptide-presenting HLA complexes, the cell line cells were stimulated with IFN-γ shown to induce and enhance the expression of HLA class I (HLA-ABC) and class II (HLA-DR,-DP,-DQ) antigens. However, this could also result in an increased expression of immune inhibitory checkpoint ligands possibly interfering with the *in vitro* expansion of T cells (Table 52). For this, it was necessary to analyse whether any inhibitory immune checkpoint ligands were upregulated by the cell lines in response to IFN-γ stimulation.

Table 52: Inhibitory checkpoint molecules used for tumor cell line analysis

Location	Inhibitory checkpoint molecule				
Ligand on tumor cell	HLA-II	PD-L1	PVR	CD80/86	Galectin-9
Receptor on T cell	LAG3	PD-1	TIGIT	CTLA-4	TIM-3

4.4.1.1 Establishment of IFN-γ stimulation parameters

At first, the optimal IFN-γ concentration and treatment time was determined as IFN-γ is a cytotoxic cytokine. For this, the MelDTC_11 cell line from patient MM16-423 was stimulated with IFN-γ (0.01 - 100 U/ml) for either 24 or 48 hours and analysed by flow cytometry for HLA-ABC (class I) and HLA-DR,-DP,-DQ (class II) expression. Without stimulation, all MelDTC_11 cells were positive for HLA-ABC and negative for HLA-DR,-DP,-DQ (Figure 63 A). An increase in HLA-II expression was already detected after 24 hours when 100 U/ml of IFN-γ were used (Figure 63 B). However, the effect was much higher after 48 hours (Figure 63 C) without affecting the cell viability (see appendix Figure 83). The effect, i.e. upregulation of HLA-II, was present even 72 hours after IFN-γ (100 U/ml) removal (Figure 63 D), but negatively affected the cell viability (Figure 63 E and F). At last, the expression of HLA-I, which is

recognized by CD8+ cytotoxic T cells, was not affected when cells were stimulated 24 or 48 hours with 100 U/ml IFN-γ (Figure 63 G). In summary, as there was a positive effect on HLA-II when using 100 U/ml of IFN-γ and no negative effect on the cell viability and HLA-I expression after a treatment time of 48 hours, these parameters were used for testing the cell lines of the other patients.

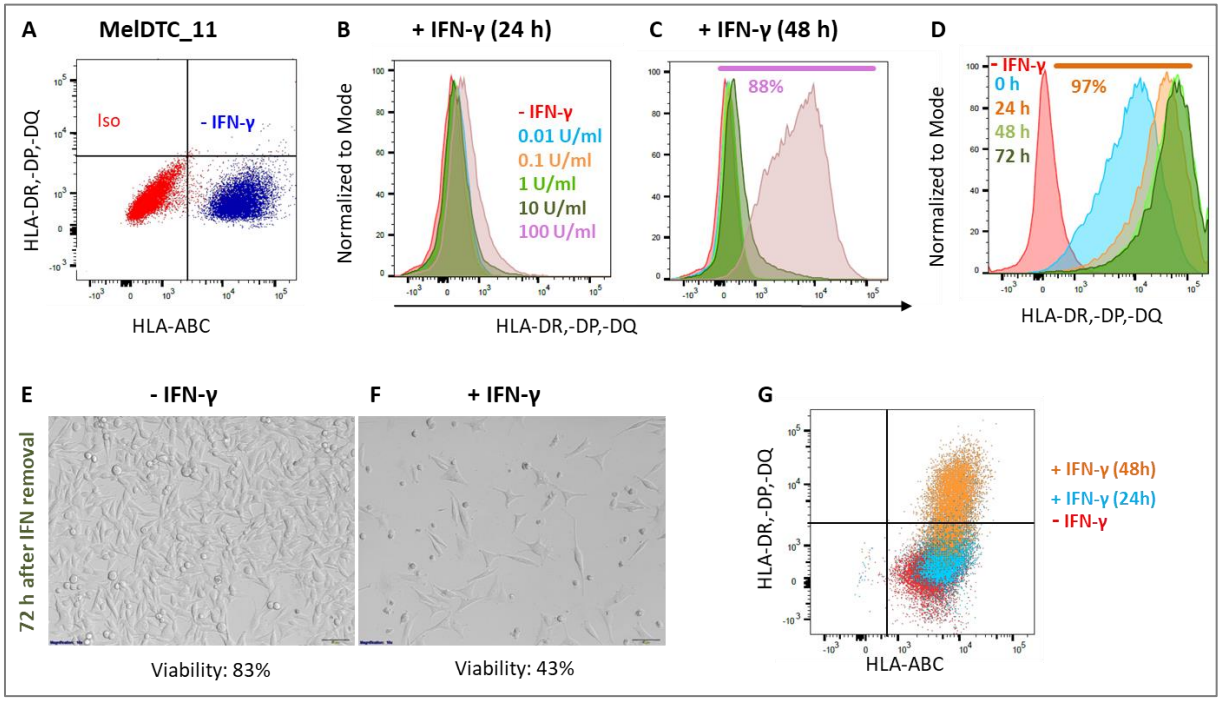


Figure 63: Effects of IFN-γ stimulation on HLA expressions and cell viability of MelDTC_11
A: Flow cytometric analysis of HLA-ABC (class I) and HLA-DR,-DP,-DQ (class II) without Interferon-γ stimulation (- IFN-γ) gated on viable cells of MelDTC_11. Iso: isotype control. **B** and **C:** Flow cytometric analysis of HLA-DR,-DP,-DQ with (+) IFN-γ stimulation (0.01-100 U/ml) for 24 **(B)** and 48 hours **(C)** gated on viable cells of MelDTC_11. **D:** HLA-DR,-DP,-DQ flow cytometric analysis 0 to 72 hours after IFN-γ (100 U/ml; 48 h) removal gated on viable cells of MelDTC_11. Time indication reflected the time after IFN-γ removal. **E** and **F:** Pictures of MelDTC_11 without IFN-γ stimulation **(E)** and 72 hours after IFN-γ (100 U/ml) removal **(F)**. The cell viability was measured by flow cytometry. Magnification:10x. **G:** Flow cytometric analysis of HLA-ABC and HLA-DR,-DP,-DQ with IFN-γ stimulation (100 U/ml) gated on viable cells of MelDTC_11.

4.4.1.2 Surface marker expression of autologous cell lines

All four autologous cell lines of the three melanoma patients were analysed for their surface marker expressions of HLA class I and II and inhibitory checkpoint molecules in response to IFN-γ stimulation (100 U/ml; 48 h). All cell lines were positive for HLA-I independent on IFN-γ stimulation (Figure 63 A and G; Figure 64 A to C and appendix Figure 84 A and B). However, IFN-γ stimulation led to an increased expression of HLA-II in three of four cell lines (Figure 63 G, Figure 64 A and C, appendix Figure 84 A and B, Table 53). Regarding inhibitory checkpoint molecules, the expression of the following ligands were evaluated upon IFN-γ stimulation as their expression could negatively affect T cell responses in combination with corresponding receptors of the T cells: PD-L1, CD80/CD86, PVR and Galectin-9. All four cell lines were positive fo PVR and negative for CD80/86 not upregulated upon stimulation (Table 53; Figure 64 D and appendix Figure 84 C and D). In addition, IFN-γ stimulation led to an increase of PD-L1 expression in all cell lines (Figure 64 E to M) showing a stronger upregulation for MelDTC_9 and MelDTC_11 (Figure 64 F and H) than for MelDTC_10a/b (Figure 64 K and M).

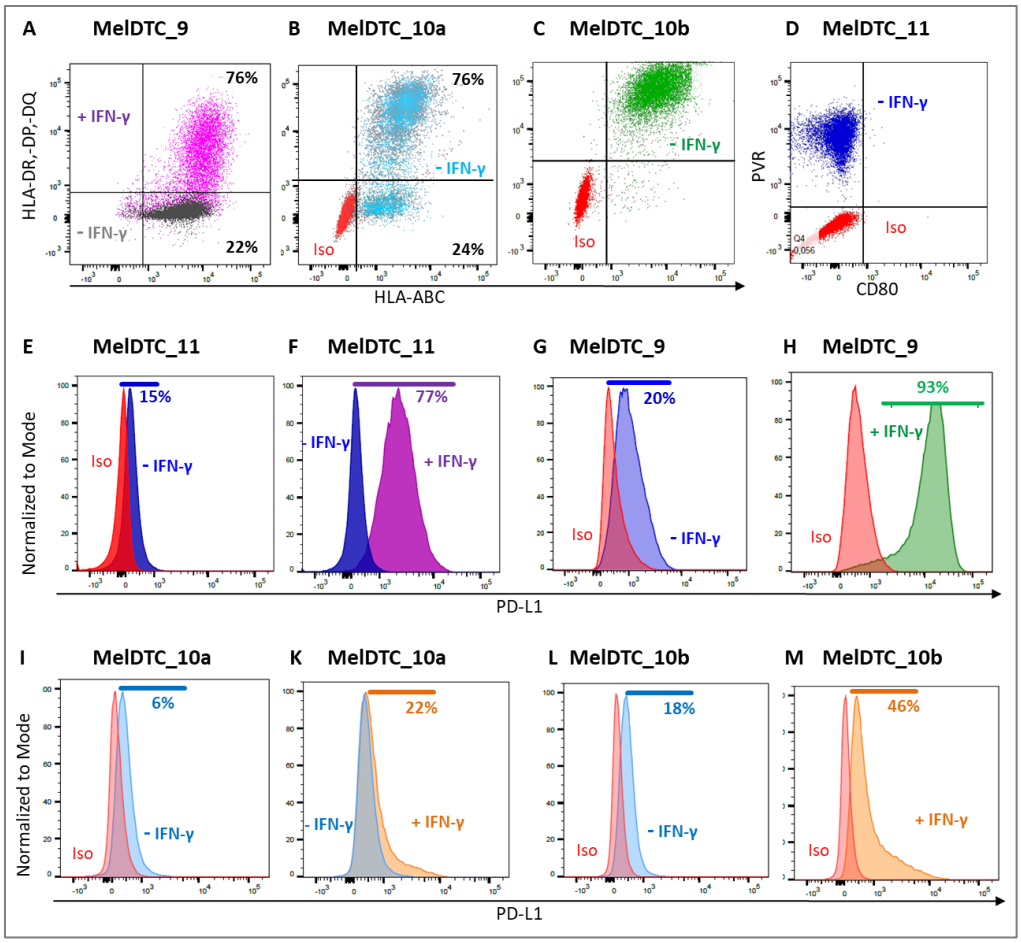


Figure 64: Flow cytometric analysis of surface marker expression of melanoma cell lines
A to **C:** Flow cytometric analysis of HLA-ABC (class I) and HLA-DR,-DP,-DQ (class II) with (+) or (-) without IFN-γ stimulation of MelDTC_9 **(A)**, MelDTC_10a **(B)** and MelDTC_10b **(C)**. **D:** Flow cytometric analysis of CD80 and PVR expression of MelDTC_11 without IFN-γ stimulation (- IFN-γ). **E** and **F:** Flow cytometric analysis of PD-L1 expression of MelDTC_11 with **(F)** or without **(E)** IFN-γ stimulation. **F** and **H:** Flow cytometric analysis of PD-L1 expression of MelDTC_9 with **(H)** or without **(F)** IFN-γ stimulation. **I** and **K:** Flow cytometric analysis of PD-L1 expression of MelDTC_10a with **(I)** or without **(K)** IFN-γ stimulation. **L** and **M:** Flow cytometric analysis of PD-L1 expression of MelDTC_10b with **(L)** or without **(M)** IFN-γ stimulation. It was always gated on viable cells and percentages were calculated based on isotype controls (iso); IFN-γ treatment: 100 U/ml, 48h.

All findings of HLA and checkpoint molecule expressions of the four cell lines were additionally summarized in Table 53. Stimulating cell lines with IFN-γ increased the HLA-II expression, which is important for CD4+ T cells, but did not negatively affect the HLA-I expression necessary for CD8+ T cells. Hence, after IFN-γ stimulation more than 76% of cell line cells were positive for both HLA molecules, thus leading to optimal requirements for peptide presentation and recognition by CD4+ and CD8+ T cells. However, IFN-γ stimulation further upregulated the expression of the inhibitory checkpoint molecule PD-L1 in all cell lines. This could stop T cell responses and interfere with the expansion of reactive T cells. To circumvent this, blocking antibodies were considered.

Table 53: Summary of HLA/ checkpoint molecule expression of autologous melanoma cell lines

Patient	**MM15-127**	**MM16-412**		**MM16-423**
Cell line (+ IFN-γ)	**MelDTC_9**	**MelDTC_10a**	**MelDTC_10b**	**MelDTC_11**
HLA-I	**+ / +**	**+ / +**	**+ / +**	**+ / +**
HLA-II	**- / ↑**	high / ↑	**+ / +**	**- / ↑**
PD-L1	low / ↑↑	low / ↑	low / ↑	low / ↑↑
PVR	**+ / +**	**+ / +**	**+ / +**	**+ / +**
CD80/86	**- / -**	**- / -**	**- / -**	**- / -**
Galectin-9	**- / -**	**- / -**	**- / -**	**- / -**

Surface marker expression of autologous cell lines (MelDTC_9, MelDTC_10a/b, MelDTC_11) without or with Interferon-γ (IFN-γ) stimulation (100 U/ml; 48 h), marked in blue. (-): negative; low: 5-49%; high: 50-90%; (+): positive; (↑): upregulation.

4.4.2 Detection of tumor-reactive T cells

In order to identify rare tumor-reactive T cells among the plethora of non-tumor-reactive T cells within the blood of the patients, *in vitro* expansion was conducted with biobanked immune cells. Due to the limited sample availability, the *in vitro* expansion was established for patient MM16-423 of whom a leukapheresis had been performed for CTC isolation. The aim was to find the best conditions to expand the most tumor-reactive T cells and also to design the assay conditions in a way that T cells could develop their maximum cytotoxic activity. In the end, this workflow will enable identifying neoantigens by detecting neoantigen-specific T cells with the highest possible sensitivity.

Expansion of tumor-reactive T cells from blood of patient MM16-423

To ensure the expansion of mainly tumor-reactive T cell clones, blood cells were exposed to autologous DCC-derived cell line cells for four weeks. IFN-γ stimulation of cell line cells led to an upregulation of HLA-II molecules recognized by CD4+ T cells. As CD4+ T cell responses are also important for CD8+ T cells, DCC-derived cell lines were pre-stimulated with IFN-γ. However, as this resulted in an enhanced expression of the immune inhibitory checkpoint ligand PD-L1 (Table 53) possibly interfering with *in vitro* expansion of T cells, the necessity of adding blocking antibody anti-PD-1 to the co-culture was tested. Every two weeks, cells were analysed by flow cytometry for the production of effector molecules (IFN-γ, CD107a) or the expression of PD-1 or CD39/CD103 to check for tumor-reactive T cells. The absolute cell number increased exponentially over time resulting in a 1.6-fold higher cell number when anti-PD-1 was added to the co-culture medium (Figure 65 A).

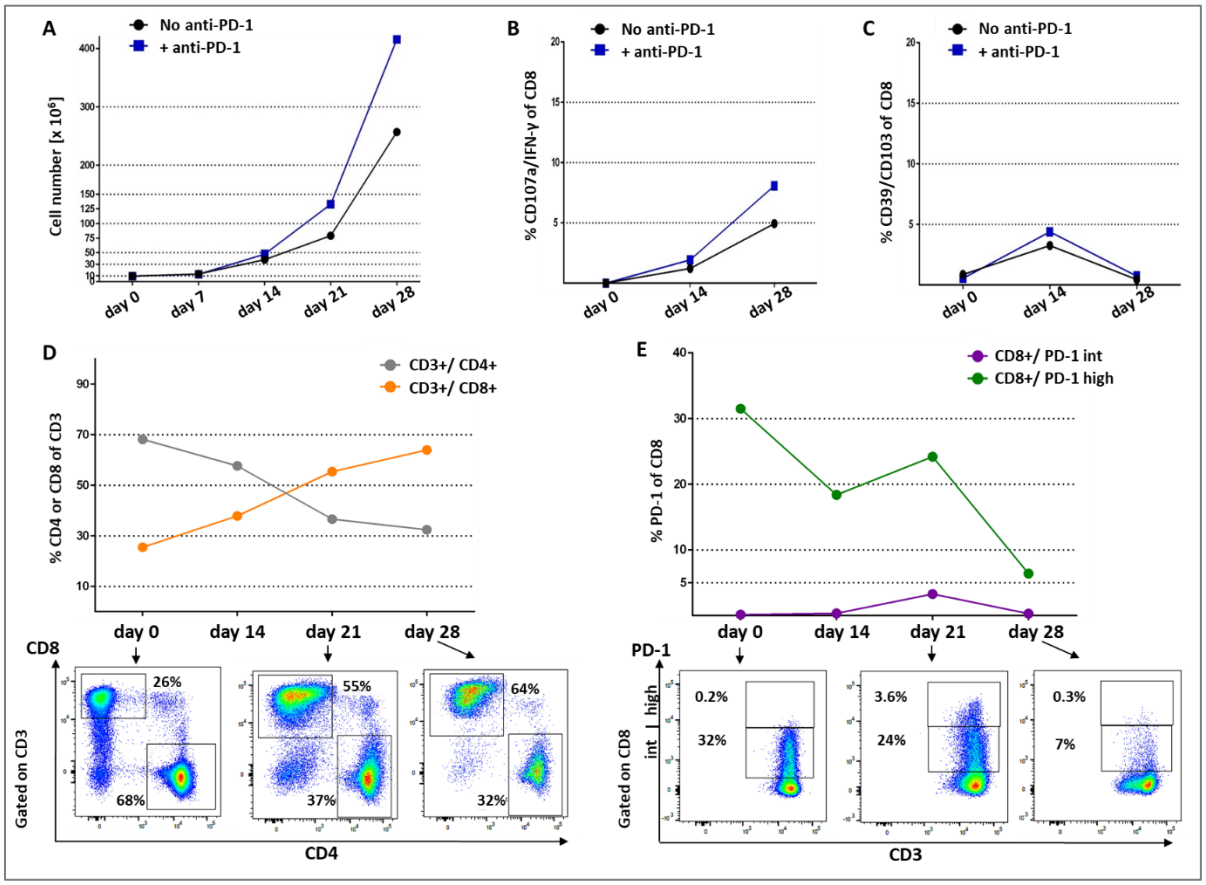


Figure 65: Expansion of tumor-reactive T cells from MM16-423's blood
A: Cell number during expansion with (blue) or without (black) anti-PD-1; **B** and **C:** Quantification of flow cytometric analysis for tumor-reactive T cell markers (CD107a+/IFN-γ+ in **B** and CD39+/CD103+ in **C**) gated on CD3+/CD8+ T cells expanded with (blue) or without (black) anti-PD-1. **D:** CD4+ and CD8+ expression levels of expanded CD3+ T cells without anti-PD-1 and corresponding flow cytometry data. **E:** PD-1 expression levels of expanded CD3+/CD8+ T cells without anti-PD-1 and corresponding flow cytometry data. PD-1 expression was separated in intermediate (PD-1 int) and high (PD-1 high) levels.

During the co-culture mainly CD3+/ CD8+ T cells were expanded (Figure 65 D). With later evaluating the killing activity of expanded T cells, expanded CD8+ cytotoxic T cells were analysed for tumor-reactivity marker. The total PD-1 expression decreased over time indicating T cell exhaustion after four weeks (Figure 65 E). The PD-1 expression was further separated into intermediate (PD-1^{int}) and high (PD-1^{high}) PD-1 expression levels. Interestingly, only CD8+ T cells from day 21 showed high expression levels of PD-1 (Figure 65 E). In addition, the production of effector molecules CD107a and IFN-γ increased over time (Figure 65 B, and appendix Figure 85 A and B). In contrast, the expression of CD39 and CD103 showed a different dynamic with an upregulation until day 14 and a downregulation from day 14 to day 28 (Figure 65 D). This indicated that an expansion over four weeks might be too long, thus leading to exhaustion of T cells or expansion of T cells that were not tumor-specific.

In order to test at which time point of the expansion T cells were able to kill best, the tumor-specific killing activity of expanded T cells was evaluated over time. Naiive T cells did not show any killing acitivity, whereas all expanded T cells from day 14, 21, and 28 showed tumor-specific killing of autologous tumor cells (Figure 66). T cells expanded over three weeks killed best (Figure 66 A), in line with the dynamic of the tumor-reactivity marker expressions. Showing the highest percentage of PD-1^{high} cells after 21 days indicated an expansion time of three weeks as optimal.

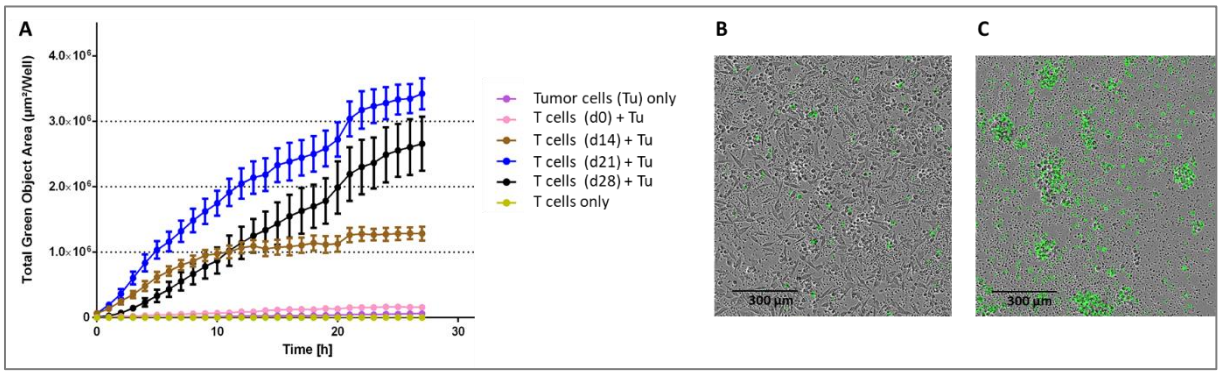


Figure 66: Immune cell-mediated killing using expanded T cells from MM16-423
A: Immune cells-mediated killing using expanded T cells (without anti-PD-1) from different time points. Killing was measured by size of Total Green Objective Area. Mean Total Green Objective Area is shown with standard errors. **B** and **C:** Example pictures taken 24 hours after starting the killing assay using naiive T cells **(C)** and T cells expanded for 21 days **(D)**; d: day.

As the addition of anti-PD-1 to the co-culture led to an enhanced number of expanded cells and a slightly higher expression of tumor-reactivity markers, the effect of anti-PD-1 during the expansion on the killing activity of T cells expanded for 21 days was analysed. Indeed, those T cells that were expanded in the presence of anti-PD-1 showed higher a killing activity than T cells that were expanded without anti-PD-1 (Figure 67).

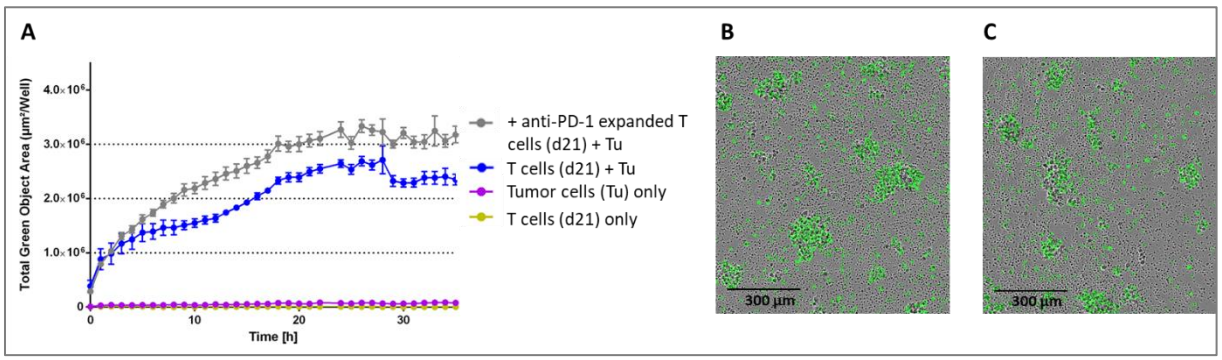


Figure 67: Immune cell-mediated killing using T cells from MM16-423 expanded with anti-PD-1
A: Immune cells-mediated killing using T cells expanded for 21 days in prescence of anti-PD-1 (5 µg/ml). Killing was measured by size of Total Green Objective Area. Mean Total Green Objective Area is shown with standard errors. **B** and **C:** Example pictures taken 24 hours after starting the assay using T cells expanded with **(B)** and without **(C)** anti-PD-1 for 21 days; d: day.

In the end, as adding anti-PD-1 during the expansion improved the killing activity of expanded T cells, it was tested whether the killing activity could further be enhanced by blocking PD-1 during the killing assay. Regarding T cells expanded for 14 and 28 days, there was no improvement of the killing activity (Figure 68 A and C). However, T cells expanded for 21 days indeed showed a stronger killing activity when anti-PD-1 was present during the assay (Figure 68 B, D and E).

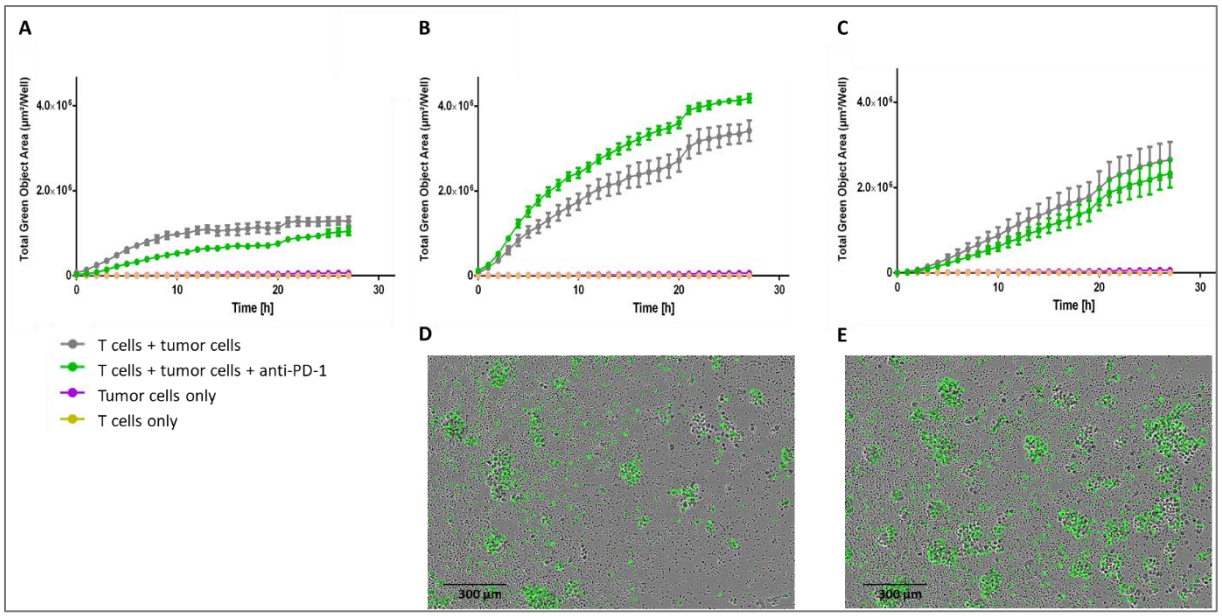


Figure 68: Immune cell-mediated killing using expanded T cells from MM16-423 with anti-PD-1
Immune cells-mediated killing using T cells expanded for 14 **(A)**, 21 **(B)** and 28 days **(C)** with or without the prescence of anti-PD-1 during the assay. Killing was measured by size of Total Green Objective Area. **D** and **E:** Example pictures taken 24 hours after starting the assay using T cells expanded for 21 days with **(D)** and without **(E)** the presence of anti-PD-1 during the assay; d: day. Mean Total Green Objective Area is shown with standard errors.

In summary, all expanded T cells from patient MM16-423 were able to kill autologous tumor cells. The strongest killing activity was seen for T cells that were expanded for 21 days showing the highest proportion of CD8+/PD-1^{high} cells. The killing activity of those cells was even stronger when anti-PD-1 was added during the killing assay. Moreover, expanding T cells in the presence of anti-PD-1 also led to an enhanced killing effect. Combining the expansions with immune cell-mediated killing assays, the best way to enrich for tumor-reactive T cells would be expanding T cells for 21 days in the presence of anti-PD-1 leading to a high number of expanded T cells with a strong killing activity of putative tumor-reactive CD8+/PD-1^{high} T cells.

4.4.3 Identification of antigen-specific T cells

In order to identify neoantigen-specific T cells among expanded tumor-reactive T cells, antigen-presenting cells (APCs) are transfected with RNA coding for candidate neoantigens. For this, the assay was established by using RNA coding for test-antigens (MART-1$^{26-35,modified}$; Survivin$^{95-104}$; Influenza A$^{MP58-66}$) and human T cell clones specifically recognizing these antigens.

At first, two protocols for generating antigen-presenting dendritic cells were compared. Dendritic cells (DCs) were generated based on plastic adherence of monocytes with a Fast-DC or a one-week DC protocol. Fast-DCs were generated within two days using differentiation cytokines (GM-CSF, IL-4) and maturation cytokines (IL-1β, TNF-α, IL-6, PGE2), whereas for the generation of one-week DCs only differentiation cytokines were used for 7 days. Viable DCs were analysed for the expression of CD209, a marker for *in vitro* generated DCs. Of all viable cells, 36% expressed CD209 when the Fast-DC protocol was used, compared to 1% CD209-positive viable cells using the one-week protocol (Figure 69). However, the one-week protocol resulted in a higher viability with 84% viable cells compared to 60%

using the Fast-DC protocol (Figure 70). As only mature DCs are able to really present antigens on their surface leading to T cell activation, the expression of HLA-II and CD80/86 were analysed. CD209+ DCs generated with the Fast-DC protocol showed high expression levels of HLA-II and CD80/86 (Figure 69 C and D, red). One-week DCs were also positive for HLA-II but only showed intermediate levels of CD80/86 (Figure 69 C and D, blue). In summary, the Fast-DC protocol outperformed the one-week protocol as more mature DCs were generated in shorter time using the Fast-DC protocol (Figure 69 E).

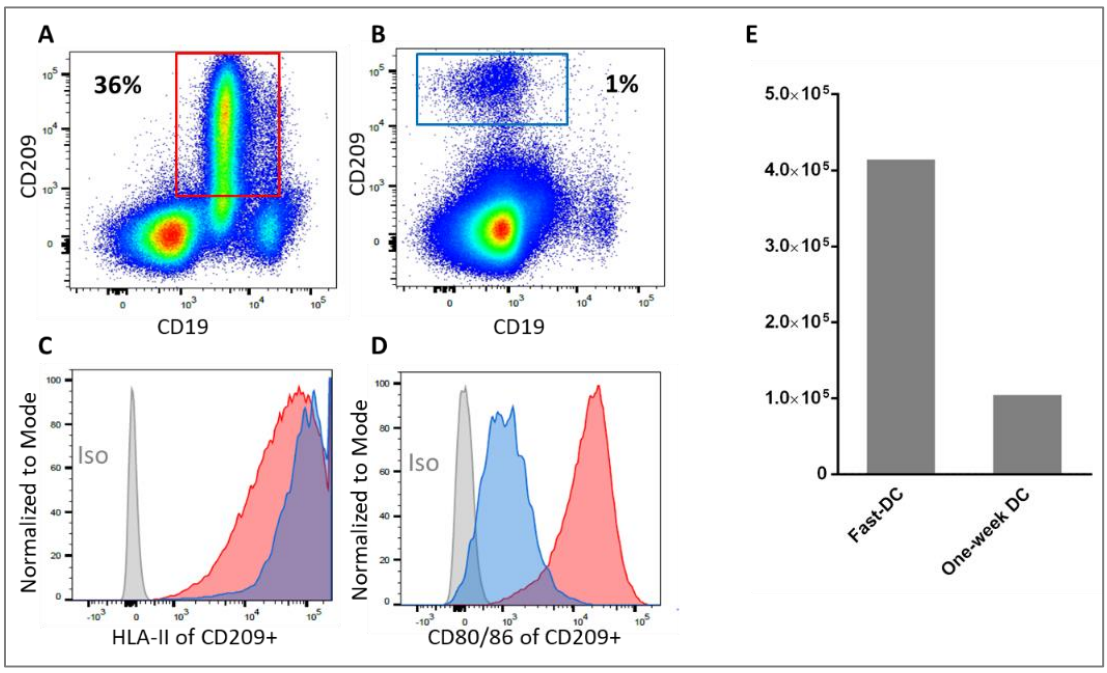


Figure 69: Generation of dendritic cells
A: Flow cytometric analysis of CD209+ dendritic cells generated with Fast-DC protocol. **B:** Flow cytometric analysis of CD209+ dendritic cells generated with one-week DC protocol. **C** and **D**: CD209+ Fast-DCs (red) and CD209+ one-week DCs (blue) were analysed for expression of maturation marker HLA-II and CD80/86. **E:** Total number of viable mature DCs generated with Fast-DC and one-week DC protocol.

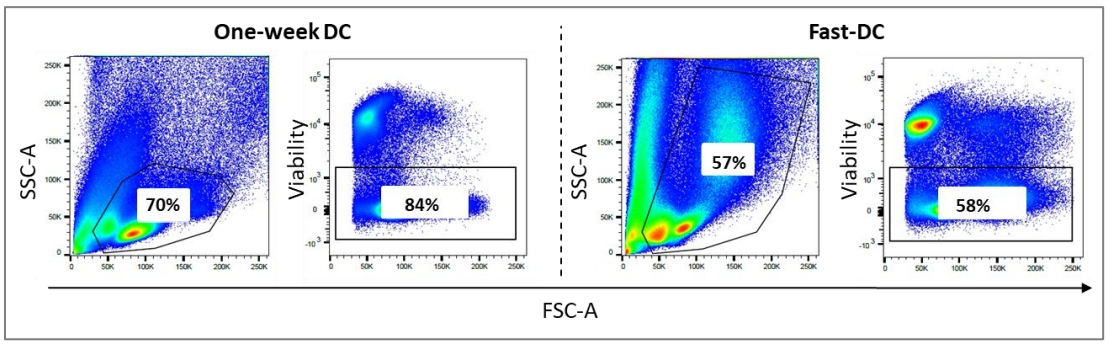


Figure 70: Flow cytometric analysis of dendritic cells
Gating strategy and cell viability of dendritic cells (DCs) generated with one-week DC (left) and Fast-DC protocol (right).

Next, a tandem minigene consisting of sequences for the test antigens MART-1, Survivin and Influenza A (FLU) was constructed and cloned into a pGEM4Z vector containing a T7 promotor for later RNA *in vitro* transcription. Afterwards, dendritic cells generated with the Fast-DC protocol were transfected via electroporation with RNA encoding for the three test antigens. The transfected DCs were co-cultured with test peptide-specific T cell clones and analysed for their release of the effector molecule IFN-γ. A significant increase of 6% or 7% of IFN-γ producing CD8+ T cells was found for all test peptide-specific T cells compared to untransfected controls (Figure 71 A and B). However, a high reactivity of Survivin-specific T cells was seen with untransfected DCs. To test if Survivin was already expressed on

mature DCs, RNA was isolated. The Survivin-specific PCR confirmed an endogenous Survivin-expression by mature dendritic cells (Figure 71 C) explaining the high background with Survivin-untransfected control samples. A specific control digestion with NcoI, cutting the PCR product in 207 and 69 bp long fragments verified the amplification of the correct fragment (Figure 71 C). To summarize, the assay was working for all three test peptides detecting test antigen-specific T cell responses. Hence, an assay was successfully established allowing to test candidate neoantigens for their immunogenicity. However, functional testing of all identified shared potential neoantigens exceeded the scope of the thesis and remains to be evaluated.

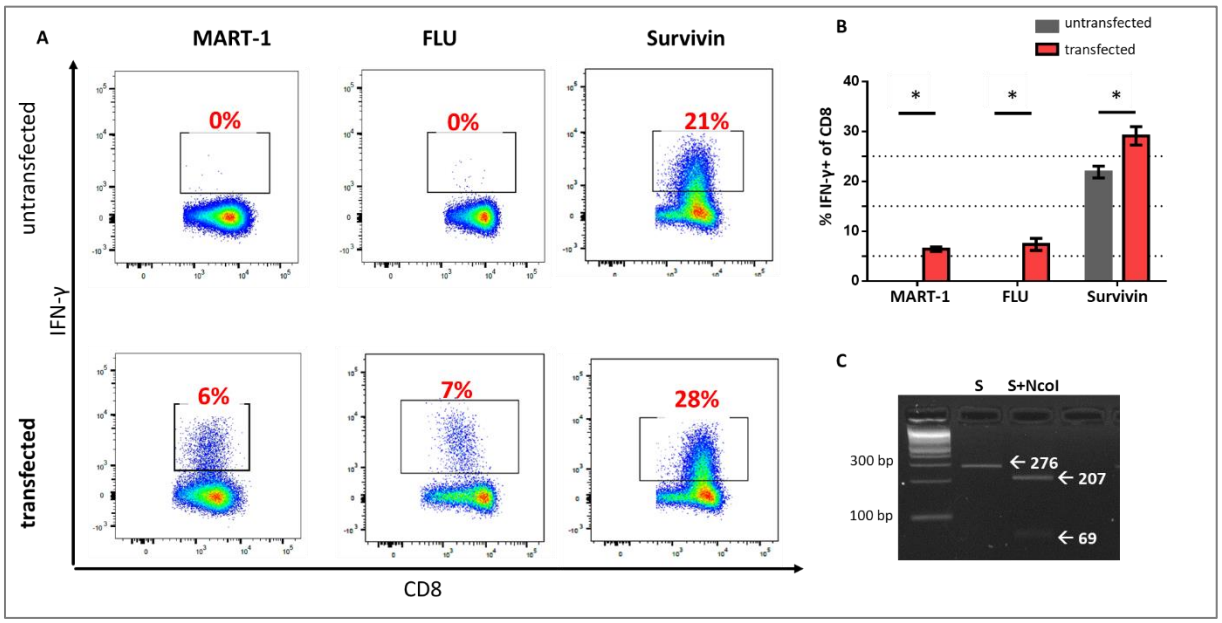


Figure 71: Results of assay for test antigens
A: Flow cytometric analysis for Interferon-γ (IFN-γ) of MART-1, Influenza A (FLU), and Surivin specific CD8+ T cells after co-culture with dendritic cells (DCs) transfected with *in vitro* transcribed RNA encoding for the test peptides. **B:** Quantification of flow cytometric analysis. Significance (indicated by *) was tested using 2way ANOVA multiple comparisons test. **C:** Agarose gel electrophoresis after PCR for Survivin (S) on cDNA from mature DCs. 276 bp long Survivin fragment was digested with enzyme NcoI leading to specific fragments of 207 and 69 bps.

5 Discussion

As metastatic melanoma is a very aggressive cancer type with a high mortality, it is aimed to treat the patient early on before lethal metastases have formed. As dissemination of melanoma cells starts early during primary tumor (PT) development, disseminated cancer cells (DCCs) that can be detected in sentinel lymph nodes (SLNs) are the target cells of adjuvant therapy. Moreover, neoantigen-reactive T cells directed towards somatically mutated tumor-specific antigens play an important role in controlling metastatic melanoma but is not described for the adjuvant immunotherapeutic setting. Therefore, the aim of the study was to investigate the potential of early DCCs isolated from SLNs to serve as surrogates for systemic DCCs and enable target identification for adjuvant therapy. To address this, the phylogenetic relationship and neoantigen conservation between local and systemically spread melanoma cells were explored.

5.1 Phylogenetic analysis of disseminated melanoma cells

The phylogeny of disseminated melanoma cells was analysed by microsatellite (MS) marker-based sequencing and by approaches considering copy number alterations (CNAs) and the *BRAF* mutation status of tumor cells. In contrast to MS mutations, which occur independent of selective pressures, CNAs and somatic mutations that affect crucial pathways are under selection pressure leading to survival and progression of tumor cells.

5.1.1 Lineage tree analysis based on microsatellite markers

Methodical aspects

In this study 12472 MS loci were sequenced (Tao et al. 2018). However, as MS loci are highly polymorphic they are predisposed to introduce noise during *in vitro* amplification, thereby hampering accurate genotyping. However, this was solved by designing a model for correct genotyping of amplification-biased cases (Raz et al. 2019). As MS markers have been used for the first time to uncover the phylogeny of disseminated melanoma cells of a single patient, efforts were made to improve several critical parameters. In terms of sample preparations it was shown that combining re-amplification and purification performed on the original WGA products of single cells led to the highest mapping rates within the target enrichment pipeline. For bulk DNA of primary tumors, the highest number of covered loci was found for WGA samples that were repaired and not re-amplified and purified. This indicated that different types of input material (single cells, bulk DNA from blood cells, bulk DNA from FFPE primary tumors) require different sample preparations for MS-based sequencing.

In addition, the number of accurately genotyped loci and the sequencing depth had influences on MS-based cell lineage reconstructions. In this study, with increasing sequencing depth a higher number of genotyped loci were covered with more than 10 reads reaching a plateau of around 8700 genotyped loci with four million total reads per sample. These findings were consistent with simulation studies of the cooperation partners. They showed that a minimum of three million reads per sample are required to retrieve an accurate tree and that the reconstruction accuracy increases with increasing number of genotyped MS loci (Spiro and Shapiro 2016).

For stabilizing the reconstructed phylogenetic tree, non-tumor control cells were needed in addition to tumor cells. CNA anaylsis was performed to discriminate tumor cells with aberrant genomes from control cells with balanced genomes (Shao et al. 2019). However, CNA analysis showed that some tumor cells had balanced CNA profiles or some cells isolated as control cells harbored aberrations. Of note, 18% of tumor cells from MM15-127 had balanced genomes and were small MCSP-positive cells with a diameter of 10 µm. This might be explained by the fact that MCSP is not 100% specific for tumor

cells but can be expressed on a variety of normal cells that are usually smaller than tumor cells (Campoli et al. 2004). In contrast, all 177 gp100-positive tumor cells from MM16-412 and 98.5% [128/130] of gp100-positive tumor cells from MM16-423 revealed aberrant profiles. Interestingly, 28% of control cells from MM16-412 that were isolated as CD3+ and CD68+ hematopoietic cells showed aberrations characteristic for tumor cells isolated from the same lymph node. Hence, these cells were not control cells, but most likely cancer cells undergoing molecular mimicry (Pisacane et al. 2007, Banerjee and Harris 2000). In the end, from three melanoma patients 84 of 734 single cells (11.4%) were excluded from phylogenetic analyses. Hence, only tumor cells having aberrant genomes and control cells with balanced genomes were included. For patient MM15-127, two small MCSP-positive cells with an amplification of chromosome 12 (named as 'balanced-12') were excluded in this study, as this aberration was not seen in any other tumor cell of this patient. However, MDM2, which is a negative regulator of the tumor suppressor p53, is located on chromosome 12 and is frequently overexpressed in several cancers (Hou et al. 2019). Therefore, those cells might be early tumor cells with a gain of chromosome 12 and lineage analyses will be repeated including those 'balanced-12' MCSP-positive cells. Nevertheless, it is important to implement a second approach in addition to single-marker staining to reliably predict whether a cell belonged to the tumor or control group and to eliminate low-quality cells from the analysis. Thus, integrating CNA analysis in the MS sequencing pipeline will improve the reliability of MS-based trees.

Cell lineage reconstructions

After protocol and algorithm optimizations, the compatibility of the established library preparation protocol with manually isolated single cells and tumor bulk DNA was shown. In order to evaluate the resolution of the method, MS-Seq based analysis was first applied to control cells from different individuals achieving a perfect separation. Next, control cells were analysed to test whether the method was able to separate non-tumor cells from different lineages. It was shown that the method was reliable as technical replicates of bulk samples indeed clustered together. In addition, clustering of hematopoietic versus either oral epithelial cells or versus endothelial cells was obtained for all patients (Tao et al. 2020) reflecting the early developmental separation of these lineages (Tavian and Péault 2005). However, a clear separation of cells from different lineages within the hematopoietic cluster (T cells, B cells and macrophages) was not possible for all patients. Notably, hematopoietic cells from different lineages originate from a common ancestor (Bujko et al. 2019). This showed that the resolution of MS-based analyses was limited with regard to separating different hematopoietic cells.

Taking all optimizations and findings regarding control cells into account, the phylogenetic relationship of melanoma cells was analysed for the three patients. It was further evaluated whether the MS-Seq based lineage analysis could reliably identify tumor cells. For all melanoma patients the majority of tumor cells were separated from non-tumor control cells. However, in all patients potentially mislocated cells were detected, i.e. tumor cells within the control cell cluster or control cells within the tumor cell cluster. However, whether the reason for these potential mislocations was biology or a technical problem remains unclear. Nevertheless, CNA analysis showed that control cells found within the tumor cells clusters had balanced genomes and tumor cells located in the control cell cluster harbored aberrant CNA profiles proving that these cells were really tumor or non-tumor cells.

Within the MS-based tree of MM15-127, there was no clear separation between cancer cells from different tissues. Therefore, the sentinel lymph node (SLN), the primary tumor (PT) and cutaneous metastatic tissues, which were removed around two years after the PT (Figure 72 A), were most likely seeded in parallel and not in cascades with the dissemination process starting from a precursor lesion (PL). The analysed clones within the primary tumor apparently evolved independently and in parallel (Figure 72 D). The majority of tumor cells from both cutaneous metastases were located in one big

cluster indicating a close relationship, simultaneous seeding of cutaneous tissues and resulting metastases that originated from a common ancestor. In addition, SLN-DCCs were spread across the tumor cluster and shared many CNAs with cells from cutaneous metastases and with the PT. Of note, four SLN-DCCs with aberrant genomes were located within the control cluster and two SLN-DCC at the beginning of the tumor cell cluster. This suggests that these SLN-DCCs (14% of all SLN-DCCs) represent not only the earliest DCCs, but also the earliest tumor cell clone.

A stepwise evolution was not detected in MM16-412. There was no clear separation of cells from SLN, NSLNs and cutaneous metastases within the MS-based tree. Only PT samples and single tumor cells from melanoma marker-negative metastasis 4 formed a separate branch within the tumor cell cluster. As PT samples were found at the beginning of the tumor cell cluster and lymph node-derived DCCs were mixed up with cells from cutaneous metastases, melanoma cells might have left the primary tumor and colonized lymph nodes and cutaneous tissues in parallel (Figure 72 E). CNA analysis rather pointed towards a late dissemination of tumor cells from the PT, as the majority of CNAs were shared between all tumor samples including the PT. Moreover, in the tumor cell cluster of this patient there was no specific order of the six different NSLNs also indicating a parallel colonization after acquisition of the most relevant genetic alterations. Early DCCs, isolated from the SLN shortly after the PT removal, were found together with the majority of NSLN-DCCs and tumor cells from two cutaneous metastases taken 1.5 years after the SLN removal (Figure 72 B) suggesting that those tissues were seeded simultaneously. Interestingly, SLN-DCCs clustered more with the earlier-occurring metastases (i.e. Met 1 and 2), than with the later-occurring metastasis 4 (removed one year after Met 1 and 2).

Comparable to MM16-412, for patient MM16-423 cells from the SLN, different NSLNs and metastases, represented by circulating tumor cells (CTCs), were not clearly separated within the MS-based tree indicating a parallel seeding (Figure 72 F, blue arrows). However, it remains unclear whether melanoma dissemination started from the primary tumor or from a precursor lesion, as the primary tumor was excluded from the analysis. CTCs, which were isolated from the blood years after the PT and SLN removal (Figure 72 C), were spread across the tumor cell cluster indicating heterogeneity among these samples. The early SLN-DCCs were also found separated from each other. As they were located within the control cell cluster and at the beginning of the tumor cell cluster between NSLN-DCCs and the majority of CTCs, they might also mirror the heterogeneity of later-arising CTCs of this patient.

In summary, the evaluation of MS-based phylogenetic reconstructions of three melanoma patients showed that lymph node DCCs were not separated, but phylogenetically closely related to systemically spread cancer cells. MS-based lineage analyses further pointed towards a parallel seeding of lymph nodes and cutaneous or blood-derived metastases. This is in line with clinical data that indirectly suggest that melanoma cells colonize distant metastases without passing and seeding lymphatic organs (Faries et al. 2017). It is further supported by studies of metastatic melanoma patients based on shared somatic single nucleotide variants showing that cells from the primary tumor simultaneously seed distant metastatic sites (Sanborn et al. 2015).

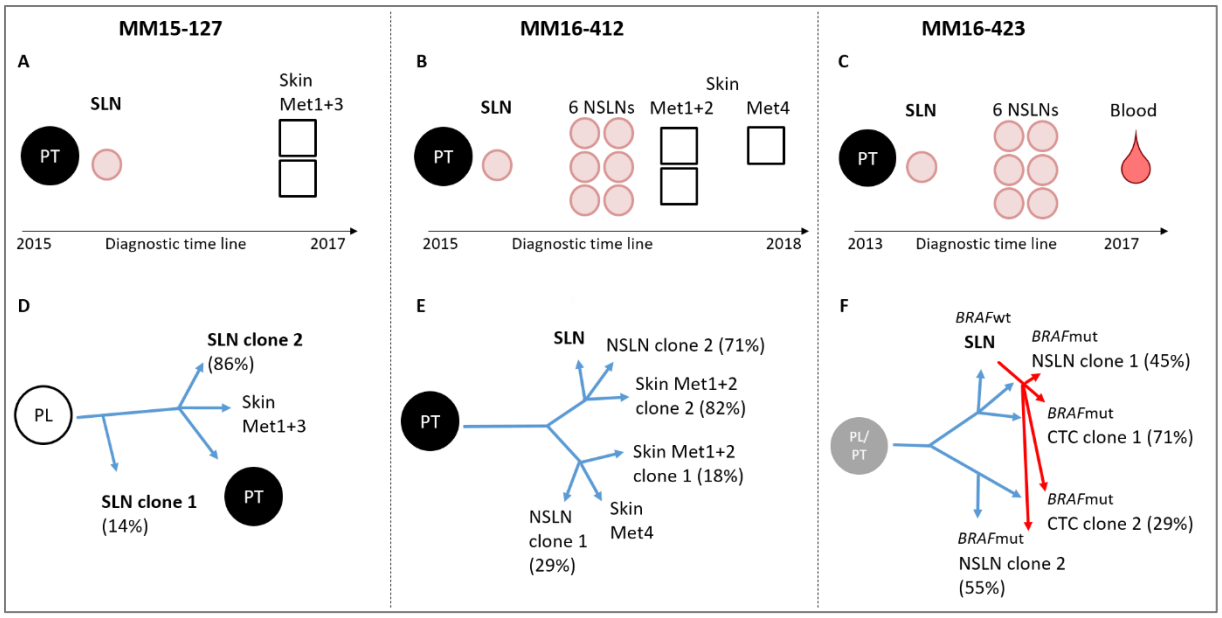


Figure 72: Phylogenetic relationship of disseminated melanoma cells
A, B, C: Diagnostic time lines when tissues were removed for tumor cell isolation from the patients (**A**: MM15-127, **B**: MM16-423, **C**: MM16-423). ○ lymph node; □ cutaneous metastasis. **D, E, F:** Phylogeny of tumor cells from the different tissues for patient MM15-127 **(D)**, MM16-412 **(E)** and MM16-423 **(F)**. Blue arrows indicate a parallel colonization suggested by microsatellite-based lineage analyses. For MM16-423 **(F)** a mixture of parallel and cascadic seeding (red arrows) was proposed by *BRAF* mutation analysis. The PT from MM16-423 was not considered in the analyses and therefore marked in grey. The length of the arrows has no meaning. The percentage shows the frequency of a clone from the corresponding sample type. PL: precursor lesion; PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; Met: metastasis; CTC: circulating tumor cell isolated from blood; wt: wild type; mut: mutated.

Limitations of MS-based analysis

However, MS-based analysis has its limitations in addressing the phylogeny of individual disseminated cancer cells. First of all, a very high number of cells are needed for reconstructing the cell lineage of a patient. Following successful isolation of high-quality DNA from single tumor and non-tumor control cells, CNA analysis was performed to exclude cells with low-quality CNA profiles and to confirm the correct group assignments. In addition, bioinformatic analysis after MS sequencing was performed to eliminate samples with insufficient coverage of MS loci. In this study, DNA was initially isolated from 989 single cells from three melanoma patients. However, considering the above mentioned restrictions and cell selection criteria, the sample size was reduced to 578 single cells resulting in a drop-out rate of 42%. This shows that a MS-based lineage reconstruction is reasonable only for patients from whom enough material is available to isolate a very large number of single cells.

Another technical limitation of the MS-based method already described was that separating all non-tumor cells from different hematopoietic lineages was not possible. Moreover, statements on the accurate position of individual cells within the lineage tree is currently restricted. Here, bootstrapping analysis was applied in order to validate the robustness of the different clusters (Lemoine et al. 2018). Branches supported with a transfer bootstrapping expectation (TBE) value larger than 70% are expected to be reliable and that corresponding clades are robust. In the lineage trees of all patients the majority of big clusters showed a TBE larger than 70%. However, many subbranches within big clusters had a TBE smaller than 70%. This indicates that one could suppose that the big clusters are true, but statements on the location of individual cells cannot be made with 100% reliability. This of course limits conclusions drawn for scattered single cells using microsatellite marker sequencing.

Furthermore, efforts must also be directed towards evaluating the impact of the different input material (FFPE versus single WGA samples) on the position of the samples within the tree. In the tree of MM15-127 primary tumor samples showed more depth than the remaining tumor samples. This

could indicate that PT samples had divided more often and thus separated from germline samples to a higher degree than other tumor cells. Alternatively, the deeper location resulted from formalin-fixation-induced DNA damage. On the other hand, the FFPE-derived PT samples from MM16-412 did not show more depth than other single tumor cells, indicating that at least some FFPE samples can be used for MS-based cell lineage tree reconstruction. However, more analyses are needed in this regard.

At last, macroscopic evaluation of single CNA profiles did not help to confirm the plausibility of MS-based trees of all patients by understanding individual locations of potentially mislocated cells and specific cancer cells within the tumor cell cluster. However, validating the conclusions drawn from MS-based trees might be possible by comparing MS-based trees to CNA-based trees.

5.1.2 Lineage tree analysis based on copy number alterations of MM16-423

The CNA data of patient MM16-423 were used to track the phylogeny of disseminated melanoma cells by analysing overlapping aberrations retrieved from *Ampli1*™ LowPass sequencing data (Ferrarini et al. 2018). The preliminary heatmap of 91 tumor samples showed that CNAs were shared between all tumor samples with prominent regions of chromosome 6, 11 and 15. Furthermore, these aberrations were also found in one of the two SLN-DCCs. As in the other SLN-DCC some, but not all shared alterations were found, this SLN-DCC might represent a very early DCC. These findings showed that SLN-DCCs may have the potential to serve as surrogates for systemically spread cancer cells as they shared CNAs with other tumor cells. Additionally, there were already some similarities between the preliminary CNA-tree and the corresponding MS-based tree: (i) cell line cells were clearly separated from remaining cancer cells; (ii) SLN-DCC2 and SLN-DCC3 were found within two different clusters; (iii) SLN-DCC2 clustered together with CTC-A; (iv) CTC-B was located separately from CTC-A. However, final phylogenetic trees from CNA data have not yet been generated due to the necessity of eliminating artefacts first. These artefacts, which were small putatively false-positive CNAs occurring in some, but not in all control cells, were most likely introduced as the used analysis tool was originally developed for Ion torrent sequencing technologies and not for the Illumina technology considered within this project. This might be overcome by using an alternative software developed by a group from the Department of Applied Research and Technical Development, Fondazione IRCCS Instituto Nazionale dei Tumori (Milano, Italy), allowing the normalization against a specific determined pool of control cells. In the end, lineage reconstructions based on CNA data without confounding artefacts will be performed for all three patients as previously shown for breast cancer and melanoma patients (Brown et al. 2017, Werner-Klein et al. 2018).

5.1.3 Phylogeny based on BRAF mutation rate of MM16-423

The phylogenetic relationship of disseminated melanoma cells based on their *BRAF* mutation status was performed exclusively for patient MM16-423 as the other patients did not have *BRAF* mutated tumor samples. The allele-specific PCR with a blocking reagent (ASB-PCR) for BRAF V600E hotspot mutations (c.1799T>A and c.1798_1799GT>AA) was developed together with Julia Greindl-Junghans. The sensitivity and specificity of the assay was shown in the context of her thesis. In this study, the ASB-PCR was applied to 85 tumor cells derived from different anatomic sites and time points of disease. Those included DCCs from SLN and NSLNs, CTCs isolated from the blood of the patient and DCC-derived cell line cells. In total, *BRAF* hotspot mutations were detected in 92.9% of analysed tumor cells. All SLN-DCCs were *BRAF* wild type, whereas all CTCs, DCCs from the cell line and from all NSLNs, except NSLN4, displayed *BRAF* mutations. NSLN4 harbored *BRAF* wild type and mutated cells. The analysis revealed that tumor cells isolated early during disease progression at the time of the PT removal were *BRAF* wild type and all NSLN-DCCs isolated from lymph nodes with a high tumor burden (DCCD > 100) were found to be *BRAF* mutated. This is in line with previous findings showing that a

mutation in the *BRAF* gene is acquired within lymph nodes during colony formation (Werner-Klein et al. 2018) and describes an early event in melanoma progression (Birkeland et al. 2018, Rabbie et al. 2020). As DCCs from NSLN4 with a relatively low tumor burden (DCCD = 58) showed a mixture of *BRAF* wild type and mutated cells, this lymph node might have been seeded first and only *BRAF* mutated cells with colony-forming potential left this lymph node and spread systemically. These findings rather point towards a mixture of parallel and cascadic seeding, suggesting that the SLN has been colonized first before seeding NSLNs and the blood stream simultaneously (Figure 72 F, red arrows). This is consistent with findings from *in vivo* data showing that melanoma cells use lymph node vessels as exit routes for systemic spread (Brown et al. 2018, Pereira and Kedrin 2018). It is further supported by studies revealing that lymph vessels protect cancer cells from cell death and so enables them to survive during dissemination (Ubellacker et al. 2020). However, without the *BRAF* information of the primary tumor it is not clear whether *BRAF* wild type cells have left the PT and seeded the SLN first, or whether *BRAF* wild type and mutated cells have left the PT and seeded lymph nodes and other sites in parallel. It might also be possible that melanoma cell dissemination started from a precursor lesion or that *BRAF* mutations are acquired independently by different clones.

5.2 Mutational analysis

Phylogenetic analyses of disseminated melanoma cells already indicated that SLN-DCCs may reflect at least parts of the mutational spectrum of systemically spread cancer cells. In order to analyse, whether early DCCs from SLNs are indeed surrogates for systemic DCCs it was evaluated whether SLN-DCCs could enable target identification within the context of shared neoantigens for adjuvant therapy. For this, whole exome sequencing (WES) was performed. WES data was first used for *in silico* HLA-I typing as the information on the HLA-I alleles was in the end intersected with the binding affinity of predicted peptides from identified mutations to these alleles.

In order to retrieve correct HLA-I alleles for single cells from WES data, a pipeline was established testing the OptiType and HLAminer tool. The alignment-based OptiType tool has a higher accuracy and success rate than the assembly-based HLAminer tool (Bauer et al. 2018). Indeed, OptiType outperformed HLAminer as HLAminer also detected HLA-I alleles that were not found by classical HLA typing. If the correct alleles were known from classical HLA typing and the allele was indeed present, correct HLA-I alleles could be retrieved from WES data using the OptiType tool. However, correct HLA-I alleles of single non-tumor cells could not be faithfully retrieved for all patients. Unbalanced amplification on HLA gene areas of single cells introduced during WGA may have led to amplification bias interfering with the *in silico* typing analysis of single cell WES data (Klasberg et al. 2019). Nevertheless, the established pipeline allowed assessment of the impact of immune evasion in tumor samples from patient MM16-412 by evaluating the time point of loss of heterozygosity (LOH) for HLA-I alleles. In DCCs from NSLNs and cutaneous metastases isolated more than one year after primary tumor and SLN removal three from possible six HLA-I alleles were lost. HLA-LOH already started at an early time point of the disease course within the SLN. In SLN-DCCs one to six HLA alleles were not detected, indicating heterogeneity among these samples. This showed that tumor cells of MM16-412 displayed a potential early escape mechanism to evade immune responses (Garrido et al. 2010). Moreover, the number of HLA-I alleles of SLN-DCCs correlated with the number of deletions on chromosome 6 where the HLA gene is located. As more alleles were detected in SLN-DCCs compared to systemically spread cancer cells isolated from NSLNs or cutaneous metastases, analysing SLN-DCCs can increase the chance to detect more candidate neoantigens. This is supported by lung cancer studies showing that HLA-LOH correlated with subclonal neoantigen burden (McGranahan et al. 2017).

Before neoantigen identification, WES data was used to analyse tumor samples for their mutational signatures, particularly skin cancer-associated UV signature 7 (Alexandrov et al. 2019). UV signature 7

was detected in tumor samples from two melanoma patients (MM15-127 and MM16-412) most likely caused by enhanced UV exposure of corresponding body parts. This is in line with MM16-412's medical history as several sunburns were reported during his lifetime thus frequently exposed to UV radiation. No UV signatures were found for the tumor samples of patient MM16-423, even though the patient was diagnosed with multiple severe sunburns and enhanced UV exposure. Interestingly, the primary tumor of MM16-423 was diagnosed as nodular malignant melanoma and located on the heel, probably protected from UV radiation. This is consistent with data showing that UV signatures are more frequently found in melanomas detected in upper or peripheral parts of the body that are exposed to UV radiation (Craig et al. 2018, Hayward et al. 2017). In addition, the majority of tumor cells from MM16-423 showed BRAF V600E mutations, whose mutation pattern is not the classic UV-related C>T transition (Craig et al. 2018). This could further explain why no UV signatures were detected for MM16-423. Of note, not all melanomas show UV signatures, a prognostic factor for melanoma patients survival (Trucco et al. 2019). Another mutational signature detected in tumor samples from MM15-127 and MM16-423 and in some control samples of MM16-412 was signature 5, which is of unknown aetiology. However, the mutational signature 5 is found in several tumor types including skin melanoma (Alexandrov et al. 2019). Notably, there was no uniform pattern of mutational signatures between the three patients. This might be explained by the fact that differences in final compositions of melanomas including UV radiation-induced mutations are shaped by age and sex of the patients (Lotz et al. 2020).

Following this, WES data was used to determine the load of exonic mutations for the different samples of each patient. Eight to 12 exonic mutations were detected in the four single non-tumor control cells of MM16-412 and MM16-423 consistent with findings showing that somatic mutations also occur in healthy human tissues (García-Nieto et al. 2019). Regarding tumor-specific exonic mutations shared between the different tumor tissues and the cell line, UV signature-positive tumor samples of MM15-127 and MM16-412 revealed 216 and 171 overlapping mutations, respectively, whereas UV signature-negative tumor samples of MM16-423 shared only 11 exonic mutations. This is in line with studies using a cohort of immunotherapy-treated patients in which more immunogenic neoantigens are detected from UV-mutated genomes (Pham and Boichard 2020). However, there was no correlation between the signal of the UV signature and the total number of detected mutations for MM15-127 and MM16-412. Nevertheless, SLN-DCCs of both patients in which no UV signatures were found showed low numbers of total mutations possibly limiting the detection of UV signatures. It might also be possible that those SLN-DCCs disseminated before being exposed to excessive UV radiation.

Analysing a relatively high number of single SLN-DCCs from patient MM15-127 showed that the number of overlapping mutations between all tumor tissues was equal to the frequency of mutations only found in SLN-DCCs shared in at least five single cells. Thus, if mutations are detected within several SLN-DCCs they are most likely present in later-arising metastases. A high degree of shared mutations was found for cutaneous metastatic samples of MM15-127 and MM16-412. These findings regarding a high mutational overlap of samples from skin metastases are consistent with previous studies indicating a parallel metastatic seeding and branching evolution (Sanborn et al. 2015, Harbst et al. 2016). Moreover, different degrees of shared exonic mutations were found for the tumor samples of the melanoma patients (Figure 73 G, H, J). Regarding MM15-127, 32% [440/1395] of exonic mutations found in the SLN-DCCs were shared with tumor samples from metastatic tissues. For patient MM16-412, 78% [209/268] of mutations from the SLN were shared with tumor samples from NSLNs and cutaneous metastatic tissues. Of note, also 55% of mutations from the PT were found in lymph nodes and cutaneous metastatic tissues showing that a large amount of mutations of the primary tumor were already shared. This indicates that tumor cells that left the PT did not progress to a great extent and supports a late dissemination of melanoma cells for patient MM16-412. The highest percentage of

shared exonic mutations was found for MM16-423 as 81% [79/98] of mutations from the SLN were also detected in NSLNs and CTCs. Hence, for all three patients a high degree (32-81%) of shared exonic mutations between SLN-DCCs and DCCs from later-arising NSLNs or metastases, either cutaneous or in form of blood-derived CTCs, were found suggesting that SLN-DCCs were representative for systemic cancer cells sharing parts of the mutational spectrum.

Potential neopeptides arising from shared exonic mutations were predicted by assessing their HLA-I binding affinities. For patient MM15-127 there was no correlation between the position of analysed single SLN-DCCs within the MS-based phylogenetic tree and the number of predicted peptides from those samples. However, the number of predicted peptides significantly correlated with the tumor burden of the lymph nodes reflected by the disseminated cancer cell density in case of MM16-412 ($p < 0.0001$) and MM16-423 ($p = 0.0442$). In these patients, SLN tumor samples showed less mutations and less predicted peptides than tumor samples from later removed NSLNs and cutaneous metastases indicating tumor progression and non-clonal evolution (Harbst et al. 2016). However, these observations are in contrast to a case study of a metastatic melanoma patient showing reduced neoantigen loads in later tumors (Davidson et al. 2019). Different findings between patients might be explained by the patient-specific complex pattern of metastatic seeding (Sanborn et al. 2015). For patient MM16-412 having a loss of heterozygosity (LOH) for HLA-I alleles in some tumor samples, it was shown that the number of predicted peptides significantly increased with decreasing number of HLA-I alleles ($p = 0.0006$), contrary to findings from lung cancer studies (McGranahan et al. 2017). But, a high number of predicted peptides was also predicted to bind to HLA-I alleles of MM16-412 that were not detected by HLA typing in some tumor samples indicating immune pressure as reason for the LOH, a potential mechanism of immune escape. This is consistent with findings from different cancer types revealing an evolutionary selection of neoantigens as a consequence of immune evasion (Lakatos et al. 2020).

5.3 Suitability of early SLN-DCCs for adjuvant therapeutic target identification

In order to evaluate whether SLN-DCCs were suitable for adjuvant therapeutic target identification, phylogenetic and CNA analyses were combined with WES data. As previously described, important points support focusing on SLN-DCCs as source for target identification: (i) SLN-DCC dissemination is a quantitative risk to die from melanoma (Ulmer et al. 2014); (ii) SLN-DCCs provide prognostic information that allows prediction of the progression risk (Ulmer et al. 2018); (iii) SLNs are available as SLN biopsies are routinely taken from melanoma patients for diagnostic purposes.

Phylogenetic analyses based on sequencing of microsatellite mutations, which are not under selection pressure, indicated that SLNs are not seeded first but simultaneously with NSLNs and cutaneous tissues and the blood stream (Figure 73 D, E, F, blue arrows). In the light of lacking information on the primary tumor, the *BRAF* mutation rate of tumor samples from one patient (MM16-423) suggested two possible scenarios: (i) colonization of the SLN first, before simultaneous seeding of NSLNs and the blood stream (Figure 73 F, red arrows) or (ii) parallel dissemination of *BRAF* wild type cells to SLN and *BRAF* mutated cells to NSLNs and the blood stream (Figure 73 F, blue arrows). These findings showed that the order of metastatic seeding of melanoma cells is very complex and can vary from patient to patient consistent with findings uncovering diversification across lineages of melanoma metastases (Rabbie et al. 2020, Sanborn et al. 2015, Harbst et al. 2016). A mixture of cascadic and parallel colonization might also be possible. However, the MS-based lineage trees suggest that the entirety of the SLN-DCCs forms a good surrogate for systemically spread cells, as the SLN-DCCs did not form a separate cluster, but were mixed with NSLN-DCCs and also cells from cutaneous or blood-derived metastases. The suitability of early SLN-DCCs for adjuvant therapy was further supported by CNA analysis showing that the majority of DCCs from the SLN already harbored genomic aberrations that were found in NSLN-DCCs

and tumor cells from metastatic tissues. Thus, SLN-DCCs may already cover parts of the mutational spectrum of cancer cells isolated later in disease progression. To summarize, phylogenetic analyses combined with CNA analysis indicate so far that the information obtained from SLN-DCCs could be therapeutically relevant.

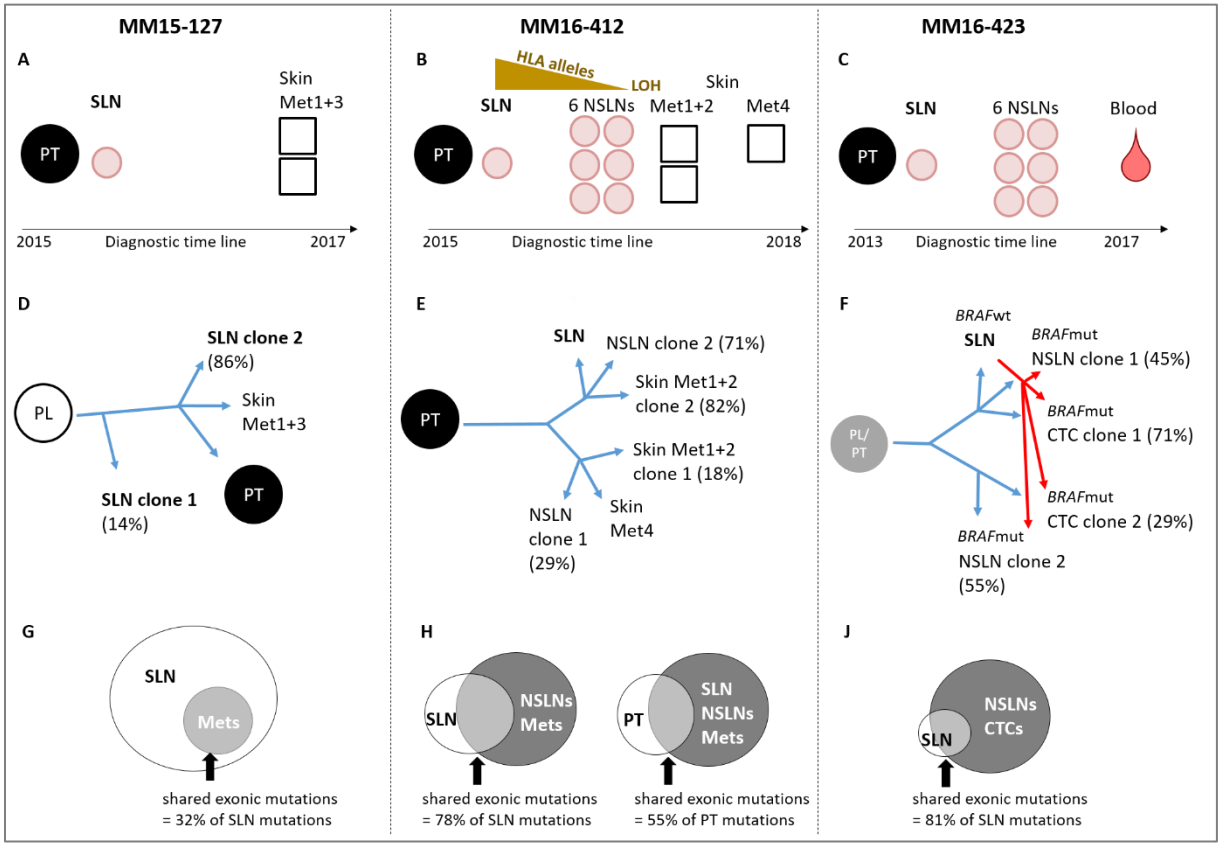


Figure 73: Summary of phylogenetic relationship and degree of shared exonic mutations
A, B, C: Diagnostic time lines when tissues were removed for tumor cell isolation from the patients (**A**: MM15-127, **B**: MM16-423, **C**: MM16-423). ○ lymph node; □ cutaneous metastasis. **D, E, F:** Phylogeny of tumor cells from the different tissues for patient MM15-127 **(D)**, MM16-412 **(E)** and MM16-423 **(F)**. Blue arrows indicate a parallel colonization for each patient suggested by microsatellite-based lineage analyses. For MM16-423 **(F)** a mixture of parallel and cascadic seeding (red arrows) was proposed by *BRAF* mutation analysis. The PT from MM16-423 was not considered in the analyses and therefore marked in grey. The length of the arrows has no meaning. The percentage shows the frequency of a clone from the corresponding sample type. **G, H, J:** Degree of shared tumor-specific exonic mutations from WES data of MM15-127 **(G)**, MM16-412 **(H)** and MM16-423 **(J)**. For MM16-412 a LOH in HLA-I alleles was found **(B)**. Cell lines are not shown. PL: precursor lesion; PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; Met: metastasis; CTC: circulating tumor cell isolated from blood; LOH: loss of heterozygosity; WES: whole exome sequencing; wt: wild type; mut: mutated.

Indeed, a high degree of exonic tumor-specific mutations shared between DCCs from SLNs and NSLNs and/or metastases (cutaneous or blood-derived) was detected for all patients using WES data (Figure 73 G, H, J). Thus, this findings indicate that SLN-DCCs are good surrogates for systemically spread cancer cells. Notably, primary tumor WES data was only available for patient MM16-412. The primary tumor of MM16-412 already allowed the detection of many shared exonic mutations that could be therapeutically relevant. However, a higher degree of exonic mutations were shared between DCCs from SLN and NSLNs and cutaneous metastases, suggesting that the entirety of SLN-DCCs sequenced as single cells better reflect the mutational spectrum of systemic cancer cells than bulk DNA from the primary tumor. For this patient, a loss of heterozygosity (LOH) for HLA-I alleles was indicated starting to occur in early SLN samples (Figure 73 B). As shared exonic mutations led to the prediction of peptides also binding to lost HLA-I alleles, HLA-LOH of this patient might be a potential mechanism of immune evasion.

Targeting tumor-specific neoantigens is of high importance in the field of cancer immunotherapy, as neoantigen-specific T cells are suggested to be the key mediators of effective immune therapies (Tran et al. 2017). The pivotal role of neoantigens as immune therapeutic targets has further been described in several studies analysing different cancer types (Efremova et al. 2017). Lung cancer studies showed that clonal neoantigens are immunogenic and sensitive to immune checkpoint blockade (McGranahan et al. 2016). Although two of the three analysed patients (MM15-127 and MM16-412) received anti-PD-1 therapies, they developed several metastases after immune checkpoint blockade indicating that metastases have evolved resistance or escaped the immune system. This is in line with another study showing neoantigen-directed mechanisms for immune evasion in non-small-cell lung cancers (Rosenthal et al. 2019). In melanoma patients a clinical benefit of adoptive T cell therapy is further predicted by their neoantigen load (Lauss et al. 2017). As Lauss et al. showed that having more than 600 somatic mutations is significantly associated with improved survival, patient MM15-127 most likely would have benefited from adoptive T cell therapy by showing more than 1000 mutations. Moreover, neoantigens were used as targets for vaccination-based therapies in metastatic melanoma patients eliciting T cell responses against multiple vaccine neoepitopes (Ott et al. 2017, Sahin et al. 2017, Hu et al. 2021, Sahin et al. 2020). In this line, Verdegaal et al. suggested not to focuss on single neoantigens but on inducing comprehensive neoantigen-specific T cell responses, as T cells mediated neoantigen immunoediting in metastatic melanoma patients (Verdegaal et al. 2016). As previous studies focused on neoantigens identified in advanced disease stage, their role in an early adjuvant therapy settings is currently unknown. Whether identified shared neoantigens detected in early SLN-DCCs of analysed patients were suitable for adjuvant immunotherapy, they had to be analysed for their immunogenicity.

5.4 Test system for functional validation of identified shared mutations

In this study, a test system for the functional evaluation of identified shared exonic tumor-specific mutations was successfully established. This test system enables the detection of neoantigen-specific T cells among expanded tumor-reactive T cells by specifically recognizing antigen-presenting cells transfected with neoantigen-encoding minigenes. As autologous DCC-derived cell line cells were used for the T cell expansion their HLA molecules and inhibitory checkpoint ligands were analysed. Upon IFN-γ stimulation the cell lines showed high expressions of HLA-I and HLA-II enabling optimal peptide presentation. However, as IFN-γ stimulation also led to the expression of the inhibitory checkpoint ligand PD-L1, a negative regulator of T cell activation, the necessity of adding anti-PD-1 during T cell expansion was tested. For patient MM16-423 tumor-reactive T cells were expanded from blood cells over four weeks resulting in a higher number of expanded T cells when PD-L1 was blocked. From expanded cytotoxic T cells around 5% were CD107a/IFN-γ-positive. These data are in line with studies using organoids of colorectal and non-small cell lung cancer patients to generate tumor-reactive T cells with T cell responses of 1-3% of cytotoxic T cells (Dijkstra et al. 2018). However, for one patient-derived organoid the tumor-reactivity was found in 50% of CD8+ T cells owed to the presence of a pre-existing tumor-reactive T cell population (Dijkstra et al. 2018). Whether the expanded tumor-reactive T cell clones of MM16-423 were already present within unexpanded T cells will be analysed by comparing corresponding RNA-Seq data. *In vitro* killing assays showed that expanded T cells were able to specifically kill melanoma cells. T cells expanded for three weeks revealed the strongest killing activity, most likely due to an enriched proportion of CD8+/PD-1^{high} cells. Expanding T cells in the presence of anti-PD-1 increased the number of expanded T cells and further led to enhanced killing activity with anti-PD-1 reducing the inhibition of T cells caused by PD-L1 expression of the cell lines. The tumor-reactivity of MM16-423 was suggested to be mediated by cytotoxic T cells expressing high levels of PD-1 (CD8+/PD-1^{high}). This is in line with studies from Gros et al. showing that PD-1 expression enables identification of patient-specific CD8+ tumor-reactive T cells and detection of neoantigen-specific T

cells in blood of melanoma patients (Gros et al. 2014, Gros et al. 2016). Consistent with this, a study of metastatic epithelial cancer revealed the detection of neoantigen-reactive T cells using T cells expressing PD-1 (Yossef et al. 2018). In this project, the addition of anti-PD-1 resulted in positive effects on the expansion of tumor-reactive T cells and on the killing activity of expanded T cells due to counteracting inhibitory effects on T cell activation of PD-L1 expressed by the cell line. Although the entire underlying mechanism of the PD-1 blockade is still unclear, pharmacodynamic changes in exhausted phenotypes of cytotoxic T cells detected in blood from advanced melanoma patients may play a role (Huang et al. 2017). In summary, tumor-reactive T cells were successfully expanded from peripheral blood cells of one patient and able to kill autologous melanoma cells. In order to increase the number of tumor-reactive T cells, lymph node cells might be a better source than peripheral blood cells, as within lymph nodes a higher density of tumor-reactive T cells is expected (Zhang et al. 2015).

5.5 Outlook

In conclusion, the aim of the study was to explore the phylogenetic relationship and neoantigen evolution and conservation between local and systemically spread melanoma cells. Due to the high mortality of metastatic melanoma the study addressed fundamental questions of adjuvant immunotherapy directed against early systemic cancer of melanoma, as represented by SLN-DCCs. Longitudinal samples from three melanoma patients were obtained, thus enabling isolation of high-quality DNA for downstream analyses. The study indicated that SLN-DCCs could indeed serve as surrogates for systemically spread cancer cells as (i) SLN-DCCs did not form a separate cluster in phylogenetic trees, thus indicating a parallel seeding of lymph nodes and cutaneous or blood-derived metastases, (ii) shared copy number alterations were found between SLN-DCCs and systemic melanoma cells, (iii) shared exonic tumor-specific mutations were identified allowing the prediction of shared tumor-specific peptides in an early adjuvant setting. Whether identified shared neoantigens were suitable for adjuvant immunotherapy remains open, but the immunogenicity of candidate neoantigens could be tested in the future using the minigene-based T cell assay that was established. Candidate neoantigens could be selected based on the following criteria: (i) shared between SLN-DCCs and tumor samples from NSLNs/ metastases and the patient-derived cell lines; (ii) having high binding affinities of predicted peptides to HLA alleles and (iii) expressed by the cell line to allow pre-clinical *in vitro* testing. Selected neoantigens could be used to generate minigenes for transfecting autologous antigen-presenting cells exposed to expanded T cells. If T cells specifically recognized the presented neopeptides, they would produce IFN-γ. In parallel HLA ligandome analysis could be performed to investigate whether shared predicted peptides were presented on the surface of the cell line cells (Rammensee and Singh-Jasuja 2013) allowing successful identification of neoantigens in melanoma patients suitable for immunotherapy (Kalaora et al. 2018). In the end, both approaches could provide information on whether one of the candidate neoantigens was of adjuvant therapeutic relevance in early systemic cancer. In addition, efforts will further be directed towards generating CNA-based trees and towards comparing phylogenetic analyses and WES data in more detail. By subjecting all DCCs to WES, the neoantigen load and conservation of individual single tumor cells located on different branches of phylogenetic trees could be assessed. Moreover, the degree of shared neoantigens could be analysed for single tumor cells found within the same branch of the lineage trees. This would allow the exploration of the neoantigen evolution in SLN-DCCs and CTCs on single cell level and enable identification of the earliest DCCs. Furthermore, for the patient with a potential HLA-LOH it could be evaluated whether predicted peptides showed high binding affinities to those HLA alleles that were lost at a later time point of disease (NSLNs and metastases) in all tumor samples including early SLN-DCCs in which those HLA alleles were still present. Then, those predicted peptides with high binding affinities to the lost HLA alleles could be tested to determine whether they were more immunogenic than predicted peptides with high binding affinities to HLA alleles that were present in all tumor

samples and not lost over time. This would give insight into the immunogenicity of neoantigens and allow the role of HLA-LOH for immune evasion in this patient to be studied. In summary, this thesis analysed the phylogeny of disseminated melanoma cells and pointed towards the pivotal role of early SLN-DCCs in neoantigen-based adjuvant settings. In the end, these findings drawn from melanoma studies will add to the importance of personalized immunotherapy already extended to different cancer types (Keskin et al. 2019, Wang et al. 2020, Ott et al. 2020, Ni et al. 2020).

6 Bibliography

Ahmadzadeh, M., Pasetto, A., Jia, L., Deniger, D. C., Stevanović, S., Robbins, P. F. and Rosenberg, S. A. (2019) 'Tumor-infiltrating human CD4+ regulatory T cells display a distinct TCR repertoire and exhibit tumor and neoantigen reactivity', *Sci Immunol,* 4(31), eaao4310.

Alexandrov, L. B., Kim, J., Haradhvala, N. J., Huang, M. N., Ng, A. W., Wu, Y., Boot, A., Covington, K. R., Gordenin, D. A., Bergstrom, E. N., Islam, S. M. A., Lopez-Bigas, N., Klimczak, L. J., McPherson, J. R., Morganella, S., Sabarinathan, R., Wheeler, D. A., Mustonen, V., Getz, G., Rozen, S. G. and Stratton, M. R. (2019) 'The Repertoire of Mutational Signatures in Human Cancer', *bioRxiv*, 322859.

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A., Behjati, S., Biankin, A. V., Bignell, G. R., Bolli, N., Borg, A., Borresen-Dale, A. L., Boyault, S., Burkhardt, B., Butler, A. P., Caldas, C., Davies, H. R., Desmedt, C., Eils, R., Eyfjord, J. E., Foekens, J. A., Greaves, M., Hosoda, F., Hutter, B., Ilicic, T., Imbeaud, S., Imielinski, M., Jager, N., Jones, D. T., Jones, D., Knappskog, S., Kool, M., Lakhani, S. R., Lopez-Otin, C., Martin, S., Munshi, N. C., Nakamura, H., Northcott, P. A., Pajic, M., Papaemmanuil, E., Paradiso, A., Pearson, J. V., Puente, X. S., Raine, K., Ramakrishna, M., Richardson, A. L., Richter, J., Rosenstiel, P., Schlesner, M., Schumacher, T. N., Span, P. N., Teague, J. W., Totoki, Y., Tutt, A. N., Valdes-Mas, R., van Buuren, M. M., van 't Veer, L., Vincent-Salomon, A., Waddell, N., Yates, L. R., Zucman-Rossi, J., Futreal, P. A., McDermott, U., Lichter, P., Meyerson, M., Grimmond, S. M., Siebert, R., Campo, E., Shibata, T., Pfister, S. M., Campbell, P. J. and Stratton, M. R. (2013) 'Signatures of mutational processes in human cancer', *Nature,* 500(7463), 415-21.

Alexandrov, L. B. and Stratton, M. R. (2014) 'Mutational signatures: the patterns of somatic mutations hidden in cancer genomes', *Curr Opin Genet Dev,* 24(100), 52-60.

Ali, Z., Yousaf, N. and Larkin, J. (2013) 'Melanoma epidemiology, biology and prognosis', *EJC Suppl,* 11(2), 81-91.

Alspach, E., Lussier, D. M., Miceli, A. P., Kizhvatov, I., DuPage, M., Luoma, A. M., Meng, W., Lichti, C. F., Esaulova, E., Vomund, A. N., Runci, D., Ward, J. P., Gubin, M. M., Medrano, R. F. V., Arthur, C. D., White, J. M., Sheehan, K. C. F., Chen, A., Wucherpfennig, K. W., Jacks, T., Unanue, E. R., Artyomov, M. N. and Schreiber, R. D. (2019) 'MHC-II neoantigens shape tumour immunity and response to immunotherapy', *Nature,* 574(7780), 696-701.

Ascierto, P. A., Kirkwood, J. M., Grob, J. J., Simeone, E., Grimaldi, A. M., Maio, M., Palmieri, G., Testori, A., Marincola, F. M. and Mozzillo, N. (2012) 'The role of BRAF V600 mutation in melanoma', *J Transl Med,* 10, 85.

Atkins, M. B., Kunkel, L., Sznol, M. and Rosenberg, S. A. (2000) 'High-dose recombinant interleukin-2 therapy in patients with metastatic melanoma: long-term survival update', *Cancer J Sci Am,* 6 Suppl 1, S11-4.

Banerjee, S. S. and Harris, M. (2000) 'Morphological and immunophenotypic variations in malignant melanoma', *Histopathology,* 36(5), 387-402.

Baron, C. S. and van Oudenaarden, A. (2019) 'Unravelling cellular relationships during development and regeneration using genetic lineage tracing', *Nature Reviews Molecular Cell Biology,* 20(12), 753-765.

Bauer, D. C., Zadoorian, A., Wilson, L. O. W. and Thorne, N. P. (2018) 'Evaluation of computational programs to predict HLA genotypes from genomic sequencing data', *Brief Bioinform,* 19(2), 179-187.

Bauer, J. and Bastian, B. C. (2006) 'Distinguishing melanocytic nevi from melanoma by DNA copy number changes: comparative genomic hybridization as a research and diagnostic tool', *Dermatol Ther,* 19(1), 40-9.

Beck-Engeser, G. B., Monach, P. A., Mumberg, D., Yang, F., Wanderling, S., Schreiber, K., Espinosa, R., 3rd, Le Beau, M. M., Meredith, S. C. and Schreiber, H. (2001) 'Point mutation in essential genes with loss or mutation of the second allele: relevance to the retention of tumor-specific antigens', *J Exp Med,* 194(3), 285-300.

Biezuner, T., Spiro, A., Raz, O., Amir, S., Milo, L., Adar, R., Chapal-Ilani, N., Berman, V., Fried, Y., Ainbinder, E., Cohen, G., Barr, H. M., Halaban, R. and Shapiro, E. (2016) 'A generic, cost-effective, and scalable cell lineage analysis platform', *Genome Research,* 26(11), 1588-1599.

Birkeland, E., Zhang, S., Poduval, D., Geisler, J., Nakken, S., Vodak, D., Meza-Zepeda, L. A., Hovig, E., Myklebost, O., Knappskog, S. and Lønning, P. E. (2018) 'Patterns of genomic evolution in advanced melanoma', *Nat Commun,* 9(1), 2665.

Blankenstein, T., Coulie, P. G., Gilboa, E. and Jaffee, E. M. (2012) 'The determinants of tumour immunogenicity', *Nat Rev Cancer,* 12(4), 307-13.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A. and Jemal, A. (2018) 'Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries', *CA Cancer J Clin,* 68(6), 394-424.

Broman, K. K., Dossett, L. A., Sun, J., Eroglu, Z. and Zager, J. S. (2019) 'Update on BRAF and MEK inhibition for treatment of melanoma in metastatic, unresectable, and adjuvant settings', *Expert Opin Drug Saf,* 18(5), 381-392.

Brown, D., Smeets, D., Székely, B., Larsimont, D., Szász, A. M., Adnet, P.-Y., Rothé, F., Rouas, G., Nagy, Z. I., Faragó, Z., Tőkés, A.-M., Dank, M., Szentmártoni, G., Udvarhelyi, N., Zoppoli, G., Pusztai, L., Piccart, M., Kulka, J., Lambrechts, D., Sotiriou, C. and Desmedt, C. (2017) 'Phylogenetic analysis of metastatic progression in breast cancer using somatic mutations and copy number aberrations', *Nat Commun,* 8(1), 14944.

Brown, M., Assen, F. P. and Leithner, A. (2018) 'Lymph node blood vessels provide exit routes for metastatic tumor cell dissemination in mice', 359(6382), 1408-1411.

Budden, T. and Bowden, N. A. (2013) 'The role of altered nucleotide excision repair and UVB-induced DNA damage in melanomagenesis', *Int J Mol Sci,* 14(1), 1132-51.

Bujko, K., Kucia, M., Ratajczak, J. and Ratajczak, M. Z. (2019) 'Hematopoietic Stem and Progenitor Cells (HSPCs)', *Adv Exp Med Biol,* 1201, 49-77.

Buson, G., Tononi, P., Forcato, C., Fontana, F., Medoro, G., Neves, R., Möhlendick, B., Stoecklein, N. and Manaresi, N. (2016) 'Abstract 2394: Scalable, rapid and affordable low-pass whole genome sequencing method for single-cell copy-number profiling on LM-PCR based WGA products', *Cancer Research,* 76, 2394-2394.

Cafri, G., Yossef, R., Pasetto, A., Deniger, D. C., Lu, Y.-C., Parkhurst, M., Gartner, J. J., Jia, L., Ray, S., Ngo, L. T., Jafferji, M., Sachs, A., Prickett, T., Robbins, P. F. and Rosenberg, S. A. (2019) 'Memory T cells targeting oncogenic mutations detected in peripheral blood of epithelial cancer patients', *Nat Commun,* 10(1), 449.

Campoli, M. R., Chang, C. C., Kageshita, T., Wang, X., McCarthy, J. B. and Ferrone, S. (2004) 'Human high molecular weight-melanoma-associated antigen (HMW-MAA): a melanoma cell surface chondroitin sulfate proteoglycan (MSCP) with biological and clinical significance', *Crit Rev Immunol,* 24(4), 267-96.

Campos, P. F. and Gilbert, T. M. (2012) 'DNA extraction from formalin-fixed material', *Methods Mol Biol,* 840, 81-5.

Carrel, S., Schmidt-Kessen, A. and Giuffrè, L. (1985) 'Recombinant interferon-gamma can induce the expression of HLA-DR and -DC on DR-negative melanoma cells and enhance the expression of HLA-ABC and tumor-associated antigens', *Eur J Immunol,* 15(2), 118-23.

Chapman, P. B., Hauschild, A., Robert, C., Haanen, J. B., Ascierto, P., Larkin, J., Dummer, R., Garbe, C., Testori, A., Maio, M., Hogg, D., Lorigan, P., Lebbe, C., Jouary, T., Schadendorf, D., Ribas, A., O'Day, S. J., Sosman, J. A., Kirkwood, J. M., Eggermont, A. M., Dreno, B., Nolop, K., Li, J., Nelson, B., Hou, J., Lee, R. J., Flaherty, K. T. and McArthur, G. A. (2011) 'Improved survival with vemurafenib in melanoma with BRAF V600E mutation', *N Engl J Med,* 364(26), 2507-16.

Choo, S. Y. (2007) 'The HLA system: genetics, immunology, clinical testing, and clinical implications', *Yonsei Med J,* 48(1), 11-23.

Cohen, C. J., Gartner, J. J., Horovitz-Fried, M., Shamalov, K., Trebska-McGowan, K., Bliskovsky, V. V., Parkhurst, M. R., Ankri, C., Prickett, T. D., Crystal, J. S., Li, Y. F., El-Gamil, M., Rosenberg, S. A. and Robbins, P. F. (2015) 'Isolation of neoantigen-specific T cells from tumor and peripheral lymphocytes', *J Clin Invest,* 125(10), 3981-91.

Craig, S., Earnshaw, C. H. and Virós, A. (2018) 'Ultraviolet light and melanoma', 244(5), 578-585.

Dalal, K. M., Patel, A., Brady, M. S., Jaques, D. P. and Coit, D. G. (2007) 'Patterns of first-recurrence and post-recurrence survival in patients with primary cutaneous melanoma after sentinel lymph node biopsy', *Ann Surg Oncol,* 14(6), 1934-42.

Davidson, G., Coassolo, S., Kieny, A., Ennen, M., Pencreach, E., Malouf, G. G., Lipsker, D. and Davidson, I. (2019) 'Dynamic Evolution of Clonal Composition and Neoantigen Landscape in Recurrent Metastatic Melanoma with a Rare Combination of Driver Mutations', *Journal of Investigative Dermatology,* 139(8), 1769-1778.e2.

Dijkstra, K. K., Cattaneo, C. M., Weeber, F., Chalabi, M., van de Haar, J., Fanchi, L. F., Slagter, M., van der Velden, D. L., Kaing, S., Kelderman, S., van Rooij, N., van Leerdam, M. E., Depla, A., Smit, E. F., Hartemink, K. J., de Groot, R., Wolkers, M. C., Sachs, N., Snaebjornsson, P., Monkhorst, K., Haanen, J., Clevers, H., Schumacher, T. N. and Voest, E. E. (2018) 'Generation of Tumor-Reactive T Cells by Co-culture of Peripheral Blood Lymphocytes and Tumor Organoids', *Cell,* 174(6), 1586-1598.e12.

Duhen, T., Duhen, R., Montler, R., Moses, J., Moudgil, T., de Miranda, N. F., Goodall, C. P., Blair, T. C., Fox, B. A., McDermott, J. E., Chang, S.-C., Grunkemeier, G., Leidner, R., Bell, R. B. and Weinberg, A. D. (2018) 'Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors', *Nat Commun,* 9(1), 2724.

Eckert, K. A. and Hile, S. E. (2009) 'Every microsatellite is different: Intrinsic DNA features dictate mutagenesis of common microsatellites present in the human genome', *Mol Carcinog,* 48(4), 379-88.

Efremova, M., Finotello, F., Rieder, D. and Trajanoski, Z. (2017) 'Neoantigens Generated by Individual Mutations and Their Role in Cancer Immunity and Immunotherapy', *Front Immunol,* 8, 1679.

Eigentler, T. K., Caroli, U. M., Radny, P. and Garbe, C. (2003) 'Palliative therapy of disseminated malignant melanoma: a systematic review of 41 randomised clinical trials', *Lancet Oncol,* 4(12), 748-59.

Erdei, E. and Torres, S. M. (2010) 'A new understanding in the epidemiology of melanoma', *Expert Rev Anticancer Ther,* 10(11), 1811-23.

Ethun, C. G. and Delman, K. A. (2016) 'The importance of surgical margins in melanoma', *J Surg Oncol,* 113(3), 339-45.

Faries, M. B., Thompson, J. F., Cochran, A. J., Andtbacka, R. H., Mozzillo, N., Zager, J. S., Jahkola, T., Bowles, T. L., Testori, A., Beitsch, P. D., Hoekstra, H. J., Moncrieff, M., Ingvar, C., Wouters, M. W. J. M., Sabel, M. S., Levine, E. A., Agnese, D., Henderson, M., Dummer, R., Rossi, C. R., Neves, R. I., Trocha, S. D., Wright, F., Byrd, D. R., Matter, M., Hsueh, E., MacKenzie-Ross, A., Johnson, D. B., Terheyden, P., Berger, A. C., Huston, T. L., Wayne, J. D., Smithers, B. M., Neuman, H. B., Schneebaum, S., Gershenwald, J. E., Ariyan, C. E., Desai, D. C., Jacobs, L., McMasters, K. M., Gesierich, A., Hersey, P., Bines, S. D., Kane, J. M., Barth, R. J., McKinnon, G., Farma, J. M., Schultz, E., Vidal-Sicart, S., Hoefer, R. A., Lewis, J. M., Scheri, R., Kelley, M. C., Nieweg, O. E., Noyes, R. D., Hoon, D. S. B., Wang, H.-J., Elashoff, D. A. and Elashoff, R. M. (2017) 'Completion Dissection or Observation for Sentinel-Node Metastasis in Melanoma', *New England Journal of Medicine,* 376(23), 2211-2222.

Ferrarini, A., Buson, G., Bolognesi, C., Forcato, C., Tononi, P., del Monaco, V., Mangano, C., Fontana, F., Medoro, G. and Manaresi, N. (2017) 'Abstract 3945: Precise copy-number profiling of single cells isolated from FFPE tissues by low-pass whole-genome sequencing', *Cancer Research,* 77, 3945-3945.

Ferrarini, A., Forcato, C., Buson, G., Tononi, P., Del Monaco, V., Terracciano, M., Bolognesi, C., Fontana, F., Medoro, G., Neves, R., Mohlendick, B., Rihawi, K., Ardizzoni, A., Sumanasuriya, S., Flohr, P., Lambros, M., de Bono, J., Stoecklein, N. H. and Manaresi, N. (2018) 'A streamlined workflow for single-cells genome-wide copy-number profiling by low-pass sequencing of LM-PCR whole-genome amplification products', *PLoS ONE,* 13(3), e0193689.

Gandini, S., Sera, F., Cattaruzza, M. S., Pasquini, P., Picconi, O., Boyle, P. and Melchi, C. F. (2005) 'Meta-analysis of risk factors for cutaneous melanoma: II. Sun exposure', *Eur J Cancer,* 41(1), 45-60.

Garbe, C., Schadendorf, D., Stolz, W., Volkenandt, M., Reinhold, U., Kortmann, R.-D., Kettelhack, C., Frerich, B., Keilholz, U., Dummer, R., Sebastian, G., Tilgen, W., Schuler, G., Mackensen, A., Kaufmann, R. and Hauschild, A. (2008) 'Short German guidelines: Malignant melanoma', *JDDG: Journal der Deutschen Dermatologischen Gesellschaft,* 6(s1), S9-S14.

García-Nieto, P. E., Morrison, A. J. and Fraser, H. B. (2019) 'The somatic mutation landscape of the human body', *Genome Biology,* 20(1), 298.

Garrido, F., Cabrera, T. and Aptsiauri, N. (2010) '"Hard" and "soft" lesions underlying the HLA class I alterations in cancer cells: implications for immunotherapy', *Int J Cancer,* 127(2), 249-56.

Gershenwald, J. E. and Ross, M. I. (2011) 'Sentinel-lymph-node biopsy for cutaneous melanoma', *N Engl J Med,* 364(18), 1738-45.

Gershenwald, J. E., Scolyer, R. A., Hess, K. R., Sondak, V. K., Long, G. V., Ross, M. I., Lazar, A. J., Faries, M. B., Kirkwood, J. M., McArthur, G. A., Haydu, L. E., Eggermont, A. M. M., Flaherty, K. T., Balch, C. M. and Thompson, J. F. (2017) 'Melanoma staging: Evidence-based changes in the American Joint Committee on Cancer eighth edition cancer staging manual', *CA Cancer J Clin,* 67(6), 472-492.

Goldstein, A. M., Chan, M., Harland, M., Hayward, N. K., Demenais, F., Bishop, D. T., Azizi, E., Bergman, W., Bianchi-Scarra, G., Bruno, W., Calista, D., Albright, L. A., Chaudru, V., Chompret, A., Cuellar, F., Elder, D. E., Ghiorzo, P., Gillanders, E. M., Gruis, N. A., Hansson, J., Hogg, D., Holland, E. A., Kanetsky, P. A., Kefford, R. F., Landi, M. T., Lang, J., Leachman, S. A., MacKie, R. M., Magnusson, V., Mann, G. J., Bishop, J. N., Palmer, J. M., Puig, S., Puig-Butille, J. A., Stark, M., Tsao, H., Tucker, M. A., Whitaker, L. and Yakobson, E. (2007) 'Features associated with germline CDKN2A mutations: a GenoMEL study of melanoma-prone families from three continents', *J Med Genet,* 44(2), 99-106.

Gros, A., Parkhurst, M. R., Tran, E., Pasetto, A., Robbins, P. F., Ilyas, S., Prickett, T. D., Gartner, J. J., Crystal, J. S., Roberts, I. M., Trebska-McGowan, K., Wunderlich, J. R., Yang, J. C. and Rosenberg, S. A. (2016) 'Prospective identification of neoantigen-specific lymphocytes in the peripheral blood of melanoma patients', *Nat Med,* 22(4), 433-8.

Gros, A., Robbins, P. F., Yao, X., Li, Y. F., Turcotte, S., Tran, E., Wunderlich, J. R., Mixon, A., Farid, S., Dudley, M. E., Hanada, K., Almeida, J. R., Darko, S., Douek, D. C., Yang, J. C. and Rosenberg, S. A. (2014) 'PD-1 identifies the patient-specific $CD8^+$ tumor-reactive repertoire infiltrating human tumors', *J Clin Invest,* 124(5), 2246-59.

Hafner, C., Toll, A., Fernández-Casado, A., Earl, J., Marqués, M., Acquadro, F., Méndez-Pertuz, M., Urioste, M., Malats, N., Burns, J. E., Knowles, M. A., Cigudosa, J. C., Hartmann, A., Vogt, T., Landthaler, M., Pujol, R. M. and Real, F. X. (2010) 'Multiple oncogenic mutations and clonal relationship in spatially distinct benign human epidermal tumors', *Proc Natl Acad Sci U S A,* 107(48), 20780-5.

Hanahan, D. and Weinberg, Robert A. (2011) 'Hallmarks of Cancer: The Next Generation', *Cell,* 144(5), 646-674.

Harbst, K., Lauss, M., Cirenajwis, H., Isaksson, K., Rosengren, F., Törngren, T., Kvist, A., Johansson, M. C., Vallon-Christersson, J., Baldetorp, B., Borg, Å., Olsson, H., Ingvar, C., Carneiro, A. and Jönsson, G. (2016) 'Multiregion Whole-Exome Sequencing Uncovers the Genetic Evolution and Mutational Heterogeneity of Early-Stage Metastatic Melanoma', *Cancer Research,* 76(16), 4765-4774.

Hartmann, C. H. and Klein, C. A. (2006) 'Gene expression profiling of single cells on large-scale oligonucleotide arrays', *Nucleic Acids Res,* 34(21), e143.

Hayward, N. K., Wilmott, J. S., Waddell, N., Johansson, P. A., Field, M. A., Nones, K., Patch, A. M., Kakavand, H., Alexandrov, L. B., Burke, H., Jakrot, V., Kazakoff, S., Holmes, O., Leonard, C., Sabarinathan, R., Mularoni, L., Wood, S., Xu, Q., Waddell, N., Tembe, V., Pupo, G. M., De Paoli-Iseppi, R., Vilain, R. E., Shang, P., Lau, L. M. S., Dagg, R. A., Schramm, S. J., Pritchard, A., Dutton-Regester, K., Newell, F., Fitzgerald, A., Shang, C. A., Grimmond, S. M., Pickett, H. A., Yang, J. Y., Stretch, J. R., Behren, A., Kefford, R. F., Hersey, P., Long, G. V., Cebon, J., Shackleton, M., Spillane, A. J., Saw, R. P. M., Lopez-Bigas, N., Pearson, J. V., Thompson, J. F., Scolyer, R. A. and Mann, G. J. (2017) 'Whole-genome landscapes of major melanoma subtypes', *Nature,* 545(7653), 175-180.

Hillis, D. M. and Bull, J. J. (1993) 'An Empirical Test of Bootstrapping as a Method for Assessing Confidence in Phylogenetic Analysis', *Systematic Biology,* 42(2), 182-192.

Hokland, M., Basse, P., Justesen, J. and Hokland, P. (1988) 'IFN-induced modulation of histocompatibility antigens on human cells. Background, mechanisms and perspectives', *Cancer Metastasis Rev,* 7(3), 193-207.

Hou, H., Sun, D. and Zhang, X. (2019) 'The role of MDM2 amplification and overexpression in therapeutic resistance of malignant tumors', *Cancer Cell International,* 19(1), 216.

Howell, W. M., Carter, V. and Clark, B. (2010) 'The HLA system: immunobiology, HLA typing, antibody screening and crossmatching techniques', *Journal of Clinical Pathology,* 63(5), 387-390.

Hu, Z., Leet, D. E., Allesøe, R. L., Oliveira, G., Li, S., Luoma, A. M., Liu, J., Forman, J., Huang, T., Iorgulescu, J. B., Holden, R., Sarkizova, S., Gohil, S. H., Redd, R. A., Sun, J., Elagina, L., Giobbie-Hurder, A., Zhang, W., Peter, L., Ciantra, Z., Rodig, S., Olive, O., Shetty, K., Pyrdol, J., Uduman, M., Lee, P. C., Bachireddy, P., Buchbinder, E. I., Yoon, C. H., Neuberg, D., Pentelute, B. L., Hacohen, N., Livak, K. J., Shukla, S. A., Olsen, L. R., Barouch, D. H., Wucherpfennig, K. W., Fritsch, E. F., Keskin, D. B., Wu, C. J. and Ott, P. A. (2021) 'Personal neoantigen vaccines induce persistent memory T cell responses and epitope spreading in patients with melanoma', *Nature Medicine*.

Huang, A. C., Postow, M. A., Orlowski, R. J., Mick, R., Bengsch, B., Manne, S., Xu, W., Harmon, S., Giles, J. R., Wenz, B., Adamow, M., Kuk, D., Panageas, K. S., Carrera, C., Wong, P., Quagliarello, F., Wubbenhorst, B., D'Andrea, K., Pauken, K. E., Herati, R. S., Staupe, R. P., Schenkel, J. M., McGettigan, S., Kothari, S., George, S. M., Vonderheide, R. H., Amaravadi, R. K., Karakousis, G. C., Schuchter, L. M., Xu, X., Nathanson, K. L., Wolchok, J. D., Gangadhar, T. C. and Wherry, E. J. (2017) 'T-cell invigoration to tumour burden ratio associated with anti-PD-1 response', *Nature,* 545(7652), 60-65.

Hugo, W., Zaretsky, J. M., Sun, L., Song, C., Moreno, B. H., Hu-Lieskovan, S., Berent-Maoz, B., Pang, J., Chmielowski, B., Cherry, G., Seja, E., Lomeli, S., Kong, X., Kelley, M. C., Sosman, J. A., Johnson, D. B., Ribas, A. and Lo, R. S. (2017) 'Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma', *Cell,* 168(3), 542.

Jakubek, Y. A., Chang, K., Sivakumar, S., Yu, Y., Giordano, M. R., Fowler, J., Huff, C. D., Kadara, H., Vilar, E. and Scheet, P. (2020) 'Large-scale analysis of acquired chromosomal alterations in non-tumor samples from patients with cancer', *Nat Biotechnol,* 38(1), 90-96.

Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B. and Nielsen, M. (2017) 'NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data', *Journal of immunology (Baltimore, Md. : 1950),* 199(9), 3360.

Kalaora, S., Wolf, Y., Feferman, T., Barnea, E., Greenstein, E., Reshef, D., Tirosh, I. and Reuben, A. (2018) 'Combined Analysis of Antigen Presentation and T-cell Recognition Reveals Restricted Immune Responses in Melanoma', 8(11), 1366-1375.

Keskin, D. B., Anandappa, A. J., Sun, J., Tirosh, I., Mathewson, N. D., Li, S., Oliveira, G., Giobbie-Hurder, A., Felt, K., Gjini, E., Shukla, S. A., Hu, Z., Li, L., Le, P. M., Allesøe, R. L., Richman, A. R., Kowalczyk, M. S., Abdelrahman, S., Geduldig, J. E., Charbonneau, S., Pelton, K., Iorgulescu, J. B., Elagina, L., Zhang, W., Olive, O., McCluskey, C., Olsen, L. R., Stevens, J., Lane, W. J., Salazar, A. M., Daley, H., Wen, P. Y., Chiocca, E. A., Harden, M., Lennon, N. J., Gabriel, S., Getz, G., Lander, E. S., Regev, A., Ritz, J., Neuberg, D., Rodig, S. J., Ligon, K. L., Suvà, M. L., Wucherpfennig, K. W., Hacohen, N., Fritsch, E. F., Livak, K. J., Ott, P. A., Wu, C. J. and Reardon, D. A. (2019) 'Neoantigen vaccine generates intratumoral T cell responses in phase Ib glioblastoma trial', *Nature,* 565(7738), 234-239.

Klasberg, S., Surendranath, V., Lange, V. and Schöfl, G. (2019) 'Bioinformatics Strategies, Challenges, and Opportunities for Next Generation Sequencing-Based HLA Genotyping', *Transfus Med Hemother,* 46(5), 312-325.

Klein, C. A. (2013) 'Selection and adaptation during metastatic cancer progression', *Nature,* 501(7467), 365-72.

Klein, C. A. (2020) 'Cancer progression and the invisible phase of metastatic colonization', *Nature Reviews Cancer,* 20(11), 681-694.

Klein, C. A., Blankenstein, T. J., Schmidt-Kittler, O., Petronio, M., Polzer, B., Stoecklein, N. H. and Riethmuller, G. (2002a) 'Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer', *Lancet,* 360(9334), 683-9.

Klein, C. A., Schmidt-Kittler, O., Schardt, J. A., Pantel, K., Speicher, M. R. and Riethmuller, G. (1999) 'Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells', *Proc Natl Acad Sci U S A,* 96(8), 4494-9.

Klein, C. A., Seidl, S., Petat-Dutter, K., Offner, S., Geigl, J. B., Schmidt-Kittler, O., Wendler, N., Passlick, B., Huber, R. M., Schlimok, G., Baeuerle, P. A. and Riethmuller, G. (2002b) 'Combined transcriptome and genome analysis of single micrometastatic cells', *Nat Biotechnol,* 20(4), 387-92.

Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., Miller, C. A., Mardis, E. R., Ding, L. and Wilson, R. K. (2012) 'VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing', *Genome Res,* 22(3), 568-76.

Lakatos, E., Williams, M. J., Schenck, R. O., Cross, W. C. H., Househam, J., Zapata, L., Werner, B., Gatenbee, C., Robertson-Tessi, M., Barnes, C. P., Anderson, A. R. A., Sottoriva, A. and Graham, T. A. (2020) 'Evolutionary dynamics of neoantigens in growing tumors', *Nature Genetics,* 52(10), 1057-1066.

Lamers, M., Berlin, I. and Neefjes, J. (2018) 'Antigen Presentation: Visualizing the MHC Class I Peptide-Loading Bottleneck', *Current Biology,* 28(2), R83-R86.

Larkin, J., Chiarion-Sileni, V., Gonzalez, R., Grob, J.-J., Rutkowski, P., Lao, C. D., Cowey, C. L., Schadendorf, D., Wagstaff, J., Dummer, R., Ferrucci, P. F., Smylie, M., Hogg, D., Hill, A., Márquez-Rodas, I., Haanen, J., Guidoboni, M., Maio, M., Schöffski, P., Carlino, M. S., Lebbé, C., McArthur, G., Ascierto, P. A., Daniels, G. A., Long, G. V., Bastholt, L., Rizzo, J. I., Balogh, A., Moshyk, A., Hodi, F. S. and Wolchok, J. D. (2019) 'Five-Year Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma', *New England Journal of Medicine,* 381(16), 1535-1546.

Lauss, M., Donia, M., Harbst, K., Andersen, R., Mitra, S., Rosengren, F., Salim, M. and Vallon-Christersson, J. (2017) 'Mutational and putative neoantigen load predict clinical benefit of adoptive T cell therapy in melanoma', 8(1), 1738.

Lemoine, F., Domelevo Entfellner, J. B., Wilkinson, E., Correia, D., Dávila Felipe, M., De Oliveira, T. and Gascuel, O. (2018) 'Renewing Felsenstein's phylogenetic bootstrap in the era of big data', *Nature,* 556(7702), 452-456.

Letouzé, E., Shinde, J., Renault, V., Couchy, G., Blanc, J.-F., Tubacher, E., Bayard, Q., Bacq, D., Meyer, V., Semhoun, J., Bioulac-Sage, P., Prévôt, S., Azoulay, D., Paradis, V., Imbeaud, S., Deleuze, J.-F. and Zucman-Rossi, J. (2017) 'Mutational signatures reveal the dynamic interplay of risk factors and cellular processes during liver tumorigenesis', *Nat Commun,* 8(1), 1315.

Lotz, M., Budden, T., Furney, S. J. and Virós, A. (2020) 'Molecular subtype, biological sex and age shape melanoma tumour evolution'.

Lugowska, I., Teterycz, P. and Rutkowski, P. (2018) 'Immunotherapy of melanoma', *Contemp Oncol (Pozn),* 22(1a), 61-67.

McGranahan, N., Furness, A. J., Rosenthal, R., Ramskov, S., Lyngaa, R., Saini, S. K., Jamal-Hanjani, M., Wilson, G. A., Birkbak, N. J., Hiley, C. T., Watkins, T. B., Shafi, S., Murugaesu, N., Mitter, R., Akarca, A. U., Linares, J., Marafioti, T., Henry, J. Y., Van Allen, E. M., Miao, D., Schilling, B., Schadendorf, D., Garraway, L. A., Makarov, V., Rizvi, N. A., Snyder, A., Hellmann, M. D., Merghoub, T., Wolchok, J. D., Shukla, S. A., Wu, C. J., Peggs, K. S., Chan, T. A., Hadrup, S. R., Quezada, S. A. and Swanton, C. (2016) 'Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade', *Science,* 351(6280), 1463-9.

McGranahan, N., Rosenthal, R., Hiley, C. T., Rowan, A. J., Watkins, T. B. K., Wilson, G. A., Birkbak, N. J., Veeriah, S., Van Loo, P., Herrero, J. and Swanton, C. (2017) 'Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution', *Cell,* 171(6), 1259-1271.e11.

Merhavi-Shoham, E., Itzhaki, O., Markel, G., Schachter, J. and Besser, M. J. (2017) 'Adoptive Cell Therapy for Metastatic Melanoma', *Cancer J,* 23(1), 48-53.

Miller, K. D., Nogueira, L., Mariotto, A. B., Rowland, J. H., Yabroff, K. R. and Alfano, C. M. (2019) 'Cancer treatment and survivorship statistics, 2019', 69(5), 363-385.

Morlan, J., Baker, J. and Sinicropi, D. (2009) 'Mutation detection by real-time PCR: a simple, robust and highly selective method', *PLoS ONE,* 4(2), e4584-e4584.

Morris, L. G., Riaz, N., Desrichard, A., Senbabaoglu, Y., Hakimi, A. A., Makarov, V., Reis-Filho, J. S. and Chan, T. A. (2016) 'Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival', *Oncotarget,* 7(9), 10051-63.

Muñoz-Couselo, E., Adelantado, E. Z., Ortiz, C., García, J. S. and Perez-Garcia, J. (2017) 'NRAS-mutant melanoma: current challenges and future prospect', *Onco Targets Ther,* 10, 3941-3947.

Ni, Q., Zhang, F., Liu, Y., Wang, Z., Yu, G., Liang, B., Niu, G., Su, T., Zhu, G., Lu, G., Zhang, L. and Chen, X. (2020) 'A bi-adjuvant nanovaccine that potentiates immunogenicity of neoantigen for combination immunotherapy of colorectal cancer', *Sci Adv,* 6(12), eaaw6071.

Nielsen, M. and Andreatta, M. (2016) 'NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets', *Genome Med,* 8(1), 33.

Nielsen, M., Lundegaard, C., Blicher, T., Lamberth, K., Harndahl, M., Justesen, S., Røder, G., Peters, B., Sette, A., Lund, O. and Buus, S. (2007) 'NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence', *PLoS ONE,* 2(8), e796.

Ott, P. A., Hu-Lieskovan, S., Chmielowski, B., Govindan, R., Naing, A., Bhardwaj, N., Margolin, K., Awad, M. M., Hellmann, M. D., Lin, J. J., Friedlander, T., Bushway, M. E., Balogh, K. N., Sciuto, T. E., Kohler, V., Turnbull, S. J., Besada, R., Curran, R. R., Trapp, B., Scherer, J., Poran, A., Harjanto, D., Barthelme, D., Ting, Y. S., Dong, J. Z., Ware, Y., Huang, Y., Huang, Z., Wanamaker, A., Cleary, L. D., Moles, M. A., Manson, K., Greshock, J., Khondker, Z. S., Fritsch, E., Rooney, M. S., DeMario, M., Gaynor, R. B. and Srinivasan, L. (2020) 'A Phase Ib Trial of Personalized Neoantigen Therapy Plus Anti-PD-1 in Patients with Advanced Melanoma, Non-small Cell Lung Cancer, or Bladder Cancer', *Cell,* 183(2), 347-362.e24.

Ott, P. A., Hu, Z., Keskin, D. B., Shukla, S. A., Sun, J., Bozym, D. J., Zhang, W., Luoma, A., Giobbie-Hurder, A., Peter, L., Chen, C., Olive, O., Carter, T. A., Li, S., Lieb, D. J., Eisenhaure, T., Gjini, E., Stevens, J., Lane, W. J., Javeri, I., Nellaiappan, K., Salazar, A. M., Daley, H., Seaman, M., Buchbinder, E. I., Yoon, C. H., Harden, M., Lennon, N., Gabriel, S., Rodig, S. J., Barouch, D. H., Aster, J. C., Getz, G., Wucherpfennig, K., Neuberg, D., Ritz, J., Lander, E. S., Fritsch, E. F., Hacohen, N. and Wu, C. J. (2017) 'An immunogenic personal neoantigen vaccine for patients with melanoma', *Nature,* advance online publication.

Pereira, E. R. and Kedrin, D. (2018) 'Lymph node metastases can invade local blood vessels, exit the node, and colonize distant organs in mice', 359(6382), 1403-1407.

Pernick, N. L., DaSilva, M., Gangi, M. D., Crissman, J. and Adsay, V. (1999) '"Histiocytic markers" in melanoma', *Mod Pathol,* 12(11), 1072-7.

Pham, T. V. and Boichard, A. (2020) 'Role of ultraviolet mutational signature versus tumor mutation burden in predicting response to immunotherapy', 14(8), 1680-1694.

Pisacane, A. M., Picciotto, F. and Risio, M. (2007) 'CD31 and CD34 expression as immunohistochemical markers of endothelial transdifferentiation in human cutaneous melanoma', *Cell Oncol,* 29(1), 59-66.

Polzer, B., Medoro, G., Pasch, S., Fontana, F., Zorzino, L., Pestka, A., Andergassen, U., Meier-Stiegen, F., Czyz, Z. T., Alberter, B., Treitschke, S., Schamberger, T., Sergio, M., Bregola, G., Doffini, A., Gianni, S., Calanca, A., Signorini, G., Bolognesi, C., Hartmann, A., Fasching, P. A., Sandri, M. T., Rack, B., Fehm, T., Giorgini, G., Manaresi, N. and Klein, C. A. (2014) 'Molecular profiling of single circulating tumor cells with diagnostic intention', *EMBO Mol Med,* 6(11), 1371-86.

Poran, A., Scherer, J., Bushway, M. E., Besada, R., Balogh, K. N., Wanamaker, A., Williams, R. G., Prabhakara, J., Ott, P. A., Hu-Lieskovan, S., Khondker, Z. S., Gaynor, R. B., Rooney, M. S. and Srinivasan, L. (2020) 'Combined TCR Repertoire Profiles and Blood Cell Phenotypes Predict Melanoma Patient Response to Personalized Neoantigen Therapy plus Anti-PD-1', *Cell Rep Med,* 1(8), 100141.

Powderly, J. D., Koeppen, H., Hodi, F. S., Sosman, J. A., Gettinger, S. N., Desai, R., Tabernero, J., Soria, J.-C., Hamid, O., Fine, G. D., Xiao, Y., Mokatrin, A., Wu, J., Anderson, M., Irving, B. A., Chen, D. S. and Kowanetz, M. (2013) 'Biomarkers and associations with the clinical activity of PD-L1 blockade in a MPDL3280A study', *Journal of Clinical Oncology,* 31(15_suppl), 3001-3001.

Prehn, R. T. and Main, J. M. (1957) 'Immunity to methylcholanthrene-induced sarcomas', *J Natl Cancer Inst,* 18(6), 769-78.

Price, M. N., Dehal, P. S. and Arkin, A. P. (2010) 'FastTree 2--approximately maximum-likelihood trees for large alignments', *PLoS ONE,* 5(3).

Prickett, T. D., Crystal, J. S., Cohen, C. J., Pasetto, A., Parkhurst, M. R., Gartner, J. J., Yao, X., Wang, R., Gros, A., Li, Y. F., El-Gamil, M., Trebska-McGowan, K., Rosenberg, S. A. and Robbins, P. F. (2016) 'Durable Complete Response from Metastatic Melanoma after Transfer of Autologous T Cells Recognizing 10 Mutated Tumor Antigens', *Cancer Immunol Res,* 4(8), 669-78.

Puntervoll, H. E., Yang, X. R., Vetti, H. H., Bachmann, I. M., Avril, M. F., Benfodda, M., Catricala, C., Dalle, S., Duval-Modeste, A. B., Ghiorzo, P., Grammatico, P., Harland, M., Hayward, N. K., Hu, H. H., Jouary, T., Martin-Denavit, T., Ozola, A., Palmer, J. M., Pastorino, L., Pjanova, D., Soufir, N., Steine, S. J., Stratigos, A. J., Thomas, L., Tinat, J., Tsao, H., Veinalde, R., Tucker, M. A., Bressac-de Paillerets, B., Newton-Bishop, J. A., Goldstein, A. M., Akslen, L. A. and Molven, A. (2013) 'Melanoma prone families with CDK4 germline mutation: phenotypic profile and associations with MC1R variants', *J Med Genet,* 50(4), 264-70.

Rabbie, R., Ansari-Pour, N., Cast, O., Lau, D., Scott, F., Welsh, S. J., Parkinson, C., Khoja, L., Moore, L., Tullett, M., Wong, K., Ferreira, I., Gómez, J. M. M., Levesque, M., Gallagher, F. A., Jiménez-Sánchez, A., Riva, L., Miller, M. L., Allinson, K., Campbell, P. J., Corrie, P., Wedge, D. C. and Adams, D. J. (2020) 'Multi-site clonality analysis uncovers pervasive heterogeneity across melanoma metastases', *Nat Commun,* 11(1), 4306.

Raimondi, S., Suppa, M. and Gandini, S. (2020) 'Melanoma Epidemiology and Sun Exposure', *Acta Derm Venereol,* 100(11), adv00136.

Rajalingam, R., Cecka, M. and Reed, E. F. (2010) 'Chapter 30 - Molecular HLA Typing Methods Used in Clinical Laboratories' in Grody, W. W., Nakamura, R. M., Strom, C. M. and Kiechle, F. L., eds., *Molecular Diagnostics*, San Diego: Academic Press, 367-379.

Rammensee, H. G. and Singh-Jasuja, H. (2013) 'HLA ligandome tumor antigen discovery for personalized vaccine approach', *Expert Rev Vaccines,* 12(10), 1211-7.

Ravanat, J. L., Douki, T. and Cadet, J. (2001) 'Direct and indirect effects of UV radiation on DNA and its components', *J Photochem Photobiol B,* 63(1-3), 88-102.

Raz, O., Biezuner, T., Spiro, A., Amir, S., Milo, L., Titelman, A., Onn, A., Chapal-Ilani, N., Tao, L., Marx, T., Feige, U. and Shapiro, E. (2019) 'Short tandem repeat stutter model inferred from direct measurement of in vitro stutter noise', *Nucleic Acids Res,* 47(5), 2436-2445.

Read, J., Wadt, K. A. and Hayward, N. K. (2016) 'Melanoma genetics', *J Med Genet,* 53(1), 1-14.

Riaz, N., Havel, J. J., Makarov, V., Desrichard, A., Urba, W. J., Sims, J. S., Hodi, F. S., Martín-Algarra, S., Mandal, R., Sharfman, W. H., Bhatia, S., Hwu, W. J., Gajewski, T. F., Slingluff, C. L., Jr., Chowell, D., Kendall, S. M., Chang, H., Shah, R., Kuo, F., Morris, L. G. T., Sidhom, J. W., Schneck, J. P., Horak, C. E., Weinhold, N. and Chan, T. A. (2017) 'Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab', *Cell,* 171(4), 934-949.e16.

Richters, M. M., Xia, H., Campbell, K. M., Gillanders, W. E., Griffith, O. L. and Griffith, M. (2019) 'Best practices for bioinformatic characterization of neoantigens for clinical utility', 11(1), 56.

Rimmer, A., Phan, H., Mathieson, I., Iqbal, Z., Twigg, S. R. F., Wilkie, A. O. M., McVean, G. and Lunter, G. (2014) 'Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications', *Nat Genet,* 46(8), 912-918.

Robinson, J., Barker, D. J., Georgiou, X., Cooper, M. A., Flicek, P. and Marsh, S. G. E. (2020) 'IPD-IMGT/HLA Database', *Nucleic Acids Res,* 48(D1), D948-d955.

Rosenberg, S. A., Yang, J. C., Sherry, R. M., Kammula, U. S., Hughes, M. S., Phan, G. Q., Citrin, D. E., Restifo, N. P., Robbins, P. F., Wunderlich, J. R., Morton, K. E., Laurencot, C. M., Steinberg, S. M., White, D. E. and Dudley, M. E. (2011) 'Durable complete responses in heavily pretreated patients with metastatic melanoma using T-cell transfer immunotherapy', *Clin Cancer Res,* 17(13), 4550-7.

Rosenthal, R., Cadieux, E. L., Salgado, R., Bakir, M. A., Moore, D. A., Hiley, C. T., Lund, T., Tanić, M., Reading, J. L., Joshi, K., Henry, J. Y., Ghorani, E., Wilson, G. A., Birkbak, N. J., Jamal-Hanjani, M., Veeriah, S., Szallasi, Z., Loi, S., Hellmann, M. D., Feber, A., Chain, B., Herrero, J., Quezada, S. A., Demeulemeester, J., Van Loo, P., Beck, S., McGranahan, N. and Swanton, C. (2019) 'Neoantigen-directed immune escape in lung cancer evolution', *Nature,* 567(7749), 479-485.

Rotman, J., Koster, B. D., Jordanova, E. S., Heeren, A. M. and de Gruijl, T. D. (2019) 'Unlocking the therapeutic potential of primary tumor-draining lymph nodes', *Cancer Immunol Immunother,* 68(10), 1681-1688.

Sahin, U., Derhovanessian, E., Miller, M., Kloke, B.-P., Simon, P., Löwer, M., Bukur, V., Tadmor, A. D., Luxemburger, U., Schrörs, B., Omokoko, T., Vormehr, M., Albrecht, C., Paruzynski, A., Kuhn, A. N., Buck, J., Heesch, S., Schreeb, K. H., Müller, F., Ortseifer, I., Vogler, I., Godehardt, E., Attig, S., Rae, R., Breitkreuz, A., Tolliver, C., Suchan, M., Martic, G., Hohberger, A., Sorn, P., Diekmann, J., Ciesla, J., Waksmann, O., Brück, A.-K., Witt, M., Zillgen, M., Rothermel, A., Kasemann, B., Langer, D., Bolte, S., Diken, M., Kreiter, S., Nemecek, R., Gebhardt, C., Grabbe, S., Höller, C., Utikal, J., Huber, C., Loquai, C. and Türeci, Ö. (2017) 'Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer', *Nature,* advance online publication.

Sahin, U., Oehm, P., Derhovanessian, E., Jabulowsky, R. A., Vormehr, M., Gold, M., Maurus, D., Schwarck-Kokarakis, D., Kuhn, A. N., Omokoko, T., Kranz, L. M., Diken, M., Kreiter, S., Haas, H., Attig, S., Rae, R., Cuk, K., Kemmer-Brück, A., Breitkreuz, A., Tolliver, C., Caspar, J., Quinkhardt, J., Hebich, L., Stein, M., Hohberger, A., Vogler, I., Liebig, I., Renken, S., Sikorski, J., Leierer, M., Müller, V., Mitzel-Rink, H., Miederer, M., Huber, C., Grabbe, S., Utikal, J., Pinter, A., Kaufmann, R., Hassel, J. C., Loquai, C. and Türeci, Ö. (2020) 'An RNA vaccine drives immunity in checkpoint-inhibitor-treated melanoma', *Nature,* 585(7823), 107-112.

Sanborn, J. Z., Chung, J., Purdom, E., Wang, N. J., Kakavand, H., Wilmott, J. S., Butler, T., Thompson, J. F., Mann, G. J., Haydu, L. E., Saw, R. P., Busam, K. J., Lo, R. S., Collisson, E. A., Hur, J. S., Spellman, P. T., Cleaver, J. E., Gray, J. W., Huh, N., Murali, R., Scolyer, R. A., Bastian, B. C. and Cho, R. J. (2015) 'Phylogenetic analyses of melanoma reveal complex patterns of metastatic dissemination', *Proc Natl Acad Sci U S A,* 112(35), 10995-1000.

Sandmann, S., de Graaf, A. O., Karimi, M., van der Reijden, B. A., Hellström-Lindberg, E., Jansen, J. H. and Dugas, M. (2017) 'Evaluating Variant Calling Tools for Non-Matched Next-Generation Sequencing Data', *Sci Rep,* 7(1), 43169.

Sandru, A., Voinea, S., Panaitescu, E. and Blidaru, A. (2014) 'Survival rates of patients with metastatic malignant melanoma', *J Med Life,* 7(4), 572-6.

Sevillya, G., Frenkel, Z. and Snir, S. (2016) 'Triplet MaxCut: a new toolkit for rooted supertree', *Methods in Ecology and Evolution,* 7(11), 1359-1365.

Shain, A. H., Yeh, I., Kovalyshyn, I., Sriharan, A., Talevich, E., Gagnon, A., Dummer, R., North, J., Pincus, L., Ruben, B., Rickaby, W., D'Arrigo, C., Robson, A. and Bastian, B. C. (2015) 'The Genetic Evolution of Melanoma from Precursor Lesions', *New England Journal of Medicine,* 373(20), 1926-1936.

Shao, X., Lv, N., Liao, J., Long, J., Xue, R., Ai, N., Xu, D. and Fan, X. (2019) 'Copy number variation is highly correlated with differential gene expression: a pan-cancer study', 20(1), 175.

Snyder, A., Makarov, V., Merghoub, T., Yuan, J., Zaretsky, J. M., Desrichard, A., Walsh, L. A., Postow, M. A., Wong, P., Ho, T. S., Hollmann, T. J., Bruggeman, C., Kannan, K., Li, Y., Elipenahli, C., Liu, C., Harbison, C. T., Wang, L., Ribas, A., Wolchok, J. D. and Chan, T. A. (2014) 'Genetic basis for clinical response to CTLA-4 blockade in melanoma', *N Engl J Med,* 371(23), 2189-2199.

Spiro, A. and Shapiro, E. (2016) 'Accuracy of Answers to Cell Lineage Questions Depends on Single-Cell Genomics Data Quality and Quantity', *PLoS Comput Biol,* 12(6), e1004983-e1004983.

Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M. and Kohlbacher, O. (2014) 'OptiType: precision HLA typing from next-generation sequencing data', *Bioinformatics,* 30(23), 3310-6.

Tao, L., Raz, O., Marx, Z., Biezuner, T., Amir, S., Milo, L., Adar, R., Onn, A., Chapal-Ilani, N., Berman, V., Levy, R., Oron, B., Halaban, R. and Shapiro, E. (2018) 'A biological-computational human cell lineage discovery platform based on duplex molecular inversion probes', *bioRxiv*, 191296.

Tao, L., Raz, O., Marx, Z., Biezuner, T., Amir, S., Milo, L., Adar, R., Onn, A., Chapal-Ilani, N., Berman, V., Levy, R., Oron, B. and Shapiro, E. (2017) 'A duplex MIPs-based biological-computational cell lineage discovery platform', *bioRxiv*, 191296.

Tao, L., Raz, O., Marx, Z., Gosh, M., Huber, S., Greindl-Junghans, J., Biezuner, T., Amir, S., Milo, L., Adar, R., Levy, R., Onn, A., Chapal-Ilani, N., Berman, V., Arie, A. B., Rom, G., Oron, B., Halaban, R., Czyz, Z. T., Werner-Klein, M., Klein, C. A. and Shapiro, E. (2020) 'Retrospective cell lineage reconstruction in Humans using short tandem repeats', *bioRxiv*, 191296.

Tavian, M. and Péault, B. (2005) 'Embryonic development of the human hematopoietic system', *Int J Dev Biol,* 49(2-3), 243-50.

Tran, E., Robbins, P. F. and Rosenberg, S. A. (2017) ''Final common pathway' of human cancer immunotherapy: targeting random somatic mutations', *Nat Immunol,* 18(3), 255-262.

Trucco, L. D., Mundra, P. A., Hogan, K., Garcia-Martinez, P., Viros, A., Mandal, A. K., Macagno, N., Gaudy-Marqueste, C., Allan, D., Baenke, F., Cook, M., McManus, C., Sanchez-Laorden, B., Dhomen, N. and Marais, R. (2019) 'Ultraviolet radiation–induced DNA damage is prognostic for outcome in melanoma', *Nature Medicine,* 25(2), 221-224.

Ubellacker, J. M., Tasdogan, A., Ramesh, V., Shen, B., Mitchell, E. C., Martin-Sandoval, M. S., Gu, Z., McCormick, M. L., Durham, A. B., Spitz, D. R., Zhao, Z., Mathews, T. P. and Morrison, S. J. (2020) 'Lymph protects metastasizing melanoma cells from ferroptosis', *Nature,* 585(7823), 113-118.

Ulmer, A., Dietz, K., Hodak, I., Polzer, B., Scheitler, S., Yildiz, M., Czyz, Z., Lehnert, P., Fehm, T., Hafner, C., Schanz, S., Rocken, M., Garbe, C., Breuninger, H., Fierlbeck, G. and Klein, C. A. (2014) 'Quantitative measurement of melanoma spread in sentinel lymph nodes and survival', *PLoS Med,* 11(2), e1001604.

Ulmer, A., Dietz, K., Werner-Klein, M., Hafner, H. M., Schulz, C., Renner, P., Weber, F., Breuninger, H., Rocken, M., Garbe, C., Fierlbeck, G. and Klein, C. A. (2018) 'The sentinel lymph node spread determines quantitatively melanoma seeding to non-sentinel lymph nodes and survival', *Eur J Cancer,* 91, 1-10.

Ulmer, A., Fischer, J. R., Schanz, S., Sotlar, K., Breuninger, H., Dietz, K., Fierlbeck, G. and Klein, C. A. (2005) 'Detection of Melanoma Cells Displaying Multiple Genomic Changes in Histopathologically Negative Sentinel Lymph Nodes', *Clinical Cancer Research,* 11(15), 5425.

Ulmer, A., Schmidt-Kittler, O., Fischer, J., Ellwanger, U., Rassner, G., Riethmuller, G., Fierlbeck, G. and Klein, C. A. (2004) 'Immunomagnetic enrichment, genomic characterization, and prognostic impact of circulating melanoma cells', *Clin Cancer Res,* 10(2), 531-7.

Uong, A. and Zon, L. I. (2010) 'Melanocytes in development and cancer', *J Cell Physiol,* 222(1), 38-41.

Van Allen, E. M., Miao, D., Schilling, B., Shukla, S. A., Blank, C., Zimmer, L., Sucker, A., Hillen, U., Foppen, M. H. G., Goldinger, S. M., Utikal, J., Hassel, J. C., Weide, B., Kaehler, K. C., Loquai, C., Mohr, P., Gutzmer, R., Dummer, R., Gabriel, S., Wu, C. J., Schadendorf, D. and Garraway, L. A. (2015) 'Genomic correlates of response to CTLA-4 blockade in metastatic melanoma', *Science,* 350(6257), 207-211.

Verdegaal, E. M. E., de Miranda, N. F. C. C., Visser, M., Harryvan, T., van Buuren, M. M., Andersen, R. S., Hadrup, S. R., van der Minne, C. E., Schotte, R., Spits, H., Haanen, J. B. A. G., Kapiteijn, E. H. W., Schumacher, T. N. and van der Burg, S. H. (2016) 'Neoantigen landscape dynamics during human melanoma–T cell interactions', *Nature,* 536(7614), 91-95.

Wagner, J. D., Gordon, M. S., Chuang, T. Y., Coleman, J. J., 3rd, Hayes, J. T., Jung, S. H. and Love, C. (2000) 'Predicting sentinel and residual lymph node basin disease after sentinel lymph node biopsy for melanoma', *Cancer,* 89(2), 453-62.

Wang, G., Kang, X., Chen, K. S., Jehng, T., Jones, L., Chen, J., Huang, X. F. and Chen, S.-Y. (2020) 'An engineered oncolytic virus expressing PD-L1 inhibitors activates tumor neoantigen-specific T cell responses', *Nat Commun,* 11(1), 1395.

Wang, W., Green, M., Choi, J. E., Gijón, M., Kennedy, P. D., Johnson, J. K., Liao, P., Lang, X., Kryczek, I., Sell, A., Xia, H., Zhou, J., Li, G., Li, J., Li, W., Wei, S., Vatan, L., Zhang, H., Szeliga, W., Gu, W., Liu, R., Lawrence, T. S., Lamb, C., Tanno, Y., Cieslik, M., Stone, E., Georgiou, G., Chan, T. A., Chinnaiyan, A. and Zou, W. (2019) 'CD8(+) T cells regulate tumour ferroptosis during cancer immunotherapy', *Nature,* 569(7755).

Ward, W. H., Lambreton, F., Goel, N., Yu, J. Q. and Farma, J. M. (2017) 'Clinical Presentation and Staging of Melanoma',

Warren, R. L., Choe, G., Freeman, D. J., Castellarin, M., Munro, S., Moore, R. and Holt, R. A. (2012) 'Derivation of HLA types from shotgun sequence datasets', *Genome Med,* 4(12), 95.

Werner-Klein, M., Scheitler, S. and Hoffmann, M. (2018) 'Genetic alterations driving metastatic colony formation are acquired outside of the primary tumour in melanoma', 9(1), 595.

Willems, T., Gymrek, M., Highnam, G., Mittelman, D. and Erlich, Y. (2014) 'The landscape of human STR variation', *Genome Res,* 24(11), 1894-904.

Wolchok, J. D., Chiarion-Sileni, V., Gonzalez, R., Rutkowski, P., Grob, J.-J., Cowey, C. L., Lao, C. D., Wagstaff, J., Schadendorf, D., Ferrucci, P. F., Smylie, M., Dummer, R., Hill, A., Hogg, D., Haanen, J., Carlino, M. S., Bechter, O., Maio, M., Marquez-Rodas, I., Guidoboni, M., McArthur, G., Lebbé, C., Ascierto, P. A., Long, G. V., Cebon, J., Sosman, J., Postow, M. A., Callahan, M. K., Walker, D., Rollin, L., Bhore, R., Hodi, F. S. and Larkin, J. (2017) 'Overall Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma', *N Engl J Med,* 377(14), 1345-1356.

Yossef, R., Tran, E., Deniger, D. C., Gros, A., Pasetto, A., Parkhurst, M. R., Gartner, J. J., Prickett, T. D., Cafri, G., Robbins, P. F. and Rosenberg, S. A. (2018) 'Enhanced detection of neoantigen-reactive T cells targeting unique and shared oncogenes for personalized cancer immunotherapy', *JCI Insight,* 3(19).

Zhang, M., Graor, H., Visioni, A., Strohl, M., Yan, L., Caja, K. and Kim, J. A. (2015) 'T Cells Derived From Human Melanoma Draining Lymph Nodes Mediate Melanoma-specific Antitumor Responses In Vitro and In Vivo in Human Melanoma Xenograft Model', *J Immunother,* 38(6), 229-38.

7 Acknowledgements

Finally, I would like to thank everyone supporting to finish my thesis, for either contributing to the project, giving advice, helping with experiments or just keeping my motivation and mood up.

First, I want to thank my supervisor Prof. Dr. Christoph Klein for making it possible to work in his group. Thank you for the encouraging and supportive discussions. I would also like to thank him for giving me the chance to present my project on several conferences and for having the opportunity to be a part of AG Werner-Klein.

Next, I would like to thank Dr. Melanie Werner-Klein for the ability to work on this exciting project and for the continuous great support. Thank you, Melanie, for teaching me how to do science and the neverending motivation. I really enjoyed every second as part of AG MWK.

From AG MWK I would especially thank Severin Gütter for his any type of support at work, good mood and special nature he brought not only to the AG, but also to my life. Severin, thanks for everything!

Moreover, I would like to thank my RIGeL mentors Prof. Dr. Gunter Meister and Prof. Dr. Hans-Georg Rammensee for their supportive advice and helpful feedback during the project.

I also want to thank Prof. Dr. Christoph Klein, Prof. Dr. Gunter Meister and Prof. Dr. Stephan Schneuwly for examining my thesis and Prof. Dr. Richard Warth for being the substitute examiner.

For proofreading of my thesis, I want to thank Prof. Dr. Christoph Klein, Dr. Melanie Werner-Klein and Anthea Povall.

A special thanks goes to the patients who were willing to contribute to our research, making a decisive step towards fighting cancer.

I would also like to thank the 'Deutsche Forschungsgemeinschaft (DFG)' and the 'Studienstiftung des deutschen Volkes' for funding and supporting this project.

Another big thanks goes to the whole SCP Lab for processing the patient samples and for helping me all the time with stainings, screening and picking.

Furthermore, I want to thank all collaborators of this project for their support and help.

Thanks to Prof. Dr. Ehud Shapiro and his team, especially Ofir Raz, for the brilliant work regarding the cell lineage analyses and the friendly cooperation.

In line with this, I would also like to thank the whole 'lineage tree team': Dr. Melanie Werner-Klein, Dr. Zbigniew Czyz, Dr. Julia Greindl-Junghans and Dr. Manjusha S. Ghosh. Thanks for the awesome teamwork. A special thanks goes to Dr. Zbigniew Czyz, who alyways pushed the project and helped in any cases.

From the ITEM-R, I want to thank especially Dr. Stefan Kirsch and Stefanie Güldener supervising and managing the whole exome sequencing part. Thanks to Dr. Martin Hoffman and Dr. Huiqin Koerkel-Qu for their bioinformatic expertise. I would also like to thank Dr. Giancarlo Feliciello for supporting the first LowPass experiments and Dr. Steffi Treitschke for her support regarding the primary tumor samples and for performing the fingerprint analyses.

Further I want to thank the RCI, where it all started. Thank you, Dr. Markus Feuerer, for giving me the opportunity to start in AG MWK as part of the Department of Immunology. I would like to thank the whole AG Feuerer, particularly Dr. Sebastian Bittner and Dr. Michael Delacher, for their supportive feedback during my first two years. Thank you, Dorothea Weber-Steffens, Brigitte Ruhland and Luise Eder, for your helping hands whenever needed. A special thanks goes to Christina Fischer, Kathrin

Schambeck and Dania Riegel for not only spending time together in the kitchen, but also for enriching my life with you as lovely friends. And Christina, thank you for always being there!

From the RCI I would also like to thank the AG of Prof. Dr. Philipp Beckhove, in particular Dr. Slava Stamova and Dr. Ayse Nur Menevese, for supporting the immunological part of the project. Thanks to the AG of Prof. Dr. Simone Thomas, in particular Carina Mirbeth, for helping with several T cell assays.

Moreover, I would like to thank all PhD students, postdocs and TAs of the Chair of Experimental Medicine and Therapy Research (LEX) for their advice and helpful suggestions during the whole time. Special thanks go to Dr. Courtney König for supporting the melanoma team and to Christian Mulas, Regina Heindl and Lena Wasmaier, who always cared about a friendly and balanced atmosphere in office 1.

Another special thanks goes to Dr. Julia Greindl-Junghans and Tobias Mederer for the great time at and especially off from work.

Furthermore, I want to thank all members of LEX, ITEM-R and the Immunology Department for the nice atmosphere, the friendly cooperations and the unforgettable time.

Ich möchte auch all meinen Freunden und Verwandten danken, die mir auf meinem Weg zur Seite gestanden haben und immer für mich da sind!

I would like to thank my best friend in Canada, Maraea, and her daughter Amaeya, my personal cancer-fighting hero. Thanks for our special friendship and keeping my motivation up to do cancer research!

Natürlich danke ich auch meinen lieben Mitbewohnern der Residenz-WG, die meine Zeit in Regensburg unvergesslich gemacht haben: Jones mit H'i, Oz, Anna-Lena, Theresa, Flo, Ludwig und Andy!

Mein größter Dank geht jedoch an meine Eltern Robert und Sonja und an meinen Bruder Thomas, die mich in allen Lebenslagen unterstützen. Danke, dass ihr immer hinter mir steht! Ihr seid die Besten!

8 Appendix

Table 54: WGA-QC result of amplified bulk DNA from primary tumor of MM15-127

WGA-QC bulk		MM15-127 PT1								MM15-127 PT2								MM15-127 PT3									
Amplicon	**bp**	**- repair**				**+ repair**				**- repair**				**+ repair**				**- repair**				**+ repair**				**+**	**-**
TP53 Exon5/6	507																										
EGFR Exon20	442																										
CKND2A Exon3	380																										
NRAS Exon2	174																										
TP53 Exon8	245																										
EGFR Exon21	418																										
EGFR Exon18	496																										
KRAS	91																										
BRAF Exon15	171																										
PIK3CA Exon20	221																										
EGFR Exon19	350																										
TP53 Exon	400																										

WGA: whole genome amplification; QC: quality control; PT: primary tumor; +/- : positive/ negative control; grey color: amplicon was present. DNA input: 50 ng.

Table 55: WGA-QC result of amplified bulk DNA of MM16-412's and MM16-423's primary tumors

WGA-QC bulk		MM16-412 PT								MM16-423 PT					
Amplicon	**bp**	**- repair**			**+ repair**			**+**	**-**	**- repair**		**+ repair**		**+**	**-**
TP53 Exon5/6	507														
EGFR Exon20	442														
CKND2A Exon3	380														
NRAS Exon2	174														
TP53 Exon8	245														
EGFR Exon21	418														
EGFR Exon18	496														
KRAS	91														
BRAF Exon15	171														
PIK3CA Exon20	221														
EGFR Exon19	350														
TP53 Exon	400														

WGA: whole genome amplification; QC: quality control; PT: primary tumor; +/- : positive/ negative control; grey color: amplicon was present. DNA input: 10 ng for MM16-412 and 1 ng for MM16-423.

Table 56: Cell selection for BRAF ASB-PCR of single cells from MM16-423

MM16-423	DCCD	n = Number of analysed single cells	ASB + CNA drop-out	n – drop-outs
SLN	2	3	0	3
NSLN4	58	18	3	15
NSLN3	10	2	1	1
NSLN5	42	4	3	1
NSLN6	138	14	2	12
NSLN2	50 000	11	1	10
NSLN1	900 000	14	1	13
Cell line		18	0	18
CTC		15	3	12
		99	**14**	**85**

Table 57: WES samples of MM15-127

Sample ID	Internal ID	Cell type	Sequencing type	Pool #	Index	Sequence	Seq-ID
SLN-DCC 12	MM15-127 LK1 WGA TZ3	SLN_MCSP	Single	Pool 2	A02	AGCAGGAA	SK03_2_12
Met1	MM16-602 Met1 Pool (Q)	Met1	Pool of 16 SCs	Pool 3	A01	ATGCCTAA	SK03_3_1
Met3	MM16-602 Met3 Pool (R)	Met3	Pool of 20 SCs	Pool 3	B01	GAATCTGA	SK03_3_2
SLN-DCC 10	MM15-127 LK1 gp100+ 11	SLNgp100	Single	Pool 3	C01	AACGTGAT	SK03_3_3
SLN-DCC 5	MM15-127 LK1 gp100+ 12	SLNgp100	Single	Pool 3	D01	CACTTCGA	SK03_3_4
SLN-DCC 16	MM15-127 LK1 gp100+ 18	SLNgp100	Single	Pool 3	E01	GCCAAGAC	SK03_3_5
SLN-DCC 8	MM15-127 LK1 gp100+ 19	SLNgp100	Single	Pool 3	F01	GACTAGTA	SK03_3_6
SLN-DCC 17	MM15-127 LK1 gp100+ 20	SLNgp100	Single	Pool 3	G01	ATTGGCTC	SK03_3_7
SLN-DCC 13	MM15-0127 LK F2 TZ1	SLN_MCSP	Single	Pool 3	H01	GATGAATC	SK03_3_8
Pool of NTCs	MM15-127 LK CD31+ Pool (S)	NTCs	Pool of 15 SCs	Pool 3	B02	GAGCTGAA	SK03_3_9
Cell line	MM15-127 gDNA cell line	Cell line	Bulk	Pool 3	C02	AAACATCG	SK03_3_15
SLN-DCC 9	MM15-127 LK1 gp100+ 13	SLNgp100	Single	Pool 4	A01	ATGCCTAA	SK03_4_1
SLN-DCC 3	MM15-127 LK1 gp100+ 14	SLNgp100	Single	Pool 4	B01	GAATCTGA	SK03_4_2
SLN-DCC 18	MM15-127 LK1 gp100+ 16	SLNgp100	Single	Pool 4	C01	AACGTGAT	SK03_4_3
SLN-DCC 14	MM15-0127 LK F2 TZ10	SLN_MCSP	Single	Pool 4	D01	CACTTCGA	SK03_4_4
SLN-DCC 6	MM15-0127 LK F2 TZ12	SLN_MCSP	Single	Pool 4	E01	GCCAAGAC	SK03_4_5
SLN-DCC 15	MM15-0127 LK F2 TZ23	SLN_MCSP	Single	Pool 4	F01	GACTAGTA	SK03_4_6
SLN-DCC 7	MM15-127 LK1 WGA TZ1	SLN_MCSP	Single	Pool 4	G01	ATTGGCTC	SK03_4_7
SLN-DCC 4	MM15-127 LK1 WGA TZ2	SLN_MCSP	Single	Pool 4	H01	GATGAATC	SK03_4_8

SC: Single cell; NTC: non-tumor cell

Table 58: WES samples of MM16-412

Sample ID	Internal ID	Cell type	Sequencing type	Pool #	Index	Sequence	Seq-ID
SLN-DCC 12	MM15-59 TZ2	SLN	Single	Pool 1	A01	ATGCCTAA	SK03_1_1
SLN-DCC 3	MM15-59 TZ8	SLN	Single	Pool 1	B01	GAATCTGA	SK03_1_2
SLN-DCC 1	MM15-59 TZ10	SLN	Single	Pool 1	C01	AACGTGAT	SK03_1_3
SLN-DCC 6	MM15-59 TZ11	SLN	Single	Pool 1	D01	CACTTCGA	SK03_1_4
SLN-DCC 5	MM15-59 TZ12	SLN	Single	Pool 1	E01	GCCAAGAC	SK03_1_5
SLN-DCC 7	MM15-59 TZ13	SLN	Single	Pool 1	F01	GACTAGTA	SK03_1_6
SLN-DCC 13	MM15-59 TZ14	SLN	Single	Pool 1	G01	ATTGGCTC	SK03_1_7
SLN-DCC 11	MM15-59 TZ15	SLN	Single	Pool 1	H01	GATGAATC	SK03_1_8
Cell line a	Mel 10a gDNA bulk	Cell line	Bulk	Pool 1	B02	GAGCTGAA	SK03_1_9
SLN-DCC 4	MM15-59 TZ17	SLN	Single	Pool 1	A02	AGCAGGAA	SK03_1_12
PT	421835 Stanze I gDNA	FFPE PT	Bulk	Pool 1	C02	AAACATCG	SK03_1_13
Met1 (bulk)	MM16-412 Met1 gDNA	Met1	Bulk	Pool 1	D02	GAGTTAGC	SK03_1_14
SLN-DCC 10	MM15-59 TZ18	SLN	Single	Pool 2	B02	GAGCTGAA	SK03_2_9
NTC1	MM17-53 EDTA CD19 3	NTC	Single	Pool 2	F02	GATAGACA	SK03_2_11
Cell line b	Mel 10b gDNA	Cell line	Bulk	Pool 2	H02	GACAGTGC	SK03_2_14
Met4 (bulk)	MM17-48 met gDNA	Met4	Bulk	Pool 3	E02	CGAACTTA	SK03_3_10
NTC2	MM17-53 EDTA CD19 5	NTC	Single	Pool 3	G02	AAGGACAC	SK03_3_11
Germline	MM17-53 EDTA	Germline	Bulk	Pool 3	D02	GAGTTAGC	SK03_3_12
NSLN1	MM16-412 Pool LK1	NSLN	Pool of 11 SCs	Pool I	A1	ATGCCTAA	MM16-412 Pool LK1
NSLN2	MM16-412 Pool LK2	NSLN	Pool of 14 SCs	Pool I	B1	GAATCTGA	MM16-412 Pool LK2
NSLN3	MM16-412 Pool LK3	NSLN	Pool of 10 SCs	Pool I	C1	AACGTGAT	MM16-412 Pool LK3
NSLN4	MM16-412 Pool LK4	NSLN	Pool of 10 SCs	Pool I	D1	CACTTCGA	MM16-412 Pool LK4
NSLN5	MM16-412 Pool LK5	NSLN	Pool of 6 SCs	Pool I	E1	GCCAAGAC	MM16-412 Pool LK5
NSLN6	MM16-412 Pool LK6	NSLN	Pool of 21 SCs	Pool I	F1	GACTAGTA	MM16-412 Pool LK6
Met1 (pool)	MM16-412 Met Pool1	Met1	Pool of 9 SCs	Pool I	G1	ATTGGCTC	MM16-412 Met Pool1
Met2 (pool)	MM16-412 Met Pool2	Met2	Pool of 13 SCs	Pool I	H1	GATGAATC	MM16-412 Met Pool2
Met4 (pool)	MM16-412 Met Pool 4	Met4	Pool of 10 SCs	Pool I	A2	AGCAGGAA	MM16-412 Met Pool 4

SC: Single cell; NTC: non-tumor cell

Table 59: WES samples of MM16-423

Sample ID	Internal ID	Cell type	Sequencing type	Pool #	Index	Sequence	Seq-ID
CTC 7	MM17-224 LB TZ7	CTC	Single	Pool 1	G02	AAGGACAC	SK03_1_10
NSLN4	MM16-423 Pool LK4	NSLN	Pool of 16 SCs	Pool 1	F02	GATAGACA	SK03_1_11
SLN-DCC 2	H13-2396 TZ3	SLN	Single	Pool 2	A01	ATGCCTAA	SK03_2_1
SLN-DCC 3	H13-2396 TZ4	SLN	Single	Pool 2	B01	GAATCTGA	SK03_2_2
CTC 4	MM17-224 LB TZ4	CTC	Single	Pool 2	C01	AACGTGAT	SK03_2_3
CTC 5	MM17-224 LB TZ5	CTC	Single	Pool 2	D01	CACTTCGA	SK03_2_4
CTC 6	MM17-224 LB TZ6	CTC	Single	Pool 2	E01	GCCAAGAC	SK03_2_5
NTC1	H13-2396 NZ1	NTC	single	Pool 2	F01	GACTAGTA	SK03_2_6
NTC2	H13-2396 NZ2	NTC	Single	Pool 2	G01	ATTGGCTC	SK03_2_7
Cell line	Mel11 gDNA	Cell line	Bulk	Pool 2	H01	GATGAATC	SK03_2_8
NSLN3	MM16-423 Pool LK3	NSLN	Pool of 4 SCs	Pool 2	E02	CGAACTTA	SK03_2_10
NSLN5	MM16-423 Pool LK5	NSLN	Pool of 5 SCs	Pool 2	G02	AAGGACAC	SK03_2_13
SLN-DCC 1	H13-2396 TZ2	SLN	Single	Pool 3	H02	GACAGTGC	SK03_3_13
CTC 3	MM17-224 LB TZ3	CTC	Single	Pool 3	F02	GATAGACA	SK03_3_14
Germline	MM17-224 LB	Germline	Bulk	Pool 4	E02	CGAACTTA	SK03_4_9
NSLN2	MM16-423 Pool LK2	NSLN	Pool of 10 SCs	Pool I	D2	GAGTTAGC	MM16-423 Pool LK2
NSLN1	MM16-423 Pool LK1	NSLN	Pool of 11 SCs	Pool II	C2	AAACATCG	MM16-423 Pool LK1
NSLN6	MM16-423 Pool LK6	NSLN	Pool of 8 SCs	Pool II	H2	GACAGTGC	MM16-423 Pool LK6

CTC: Circulating tumor cell; SC: Single cell; NTC: non-tumor cell

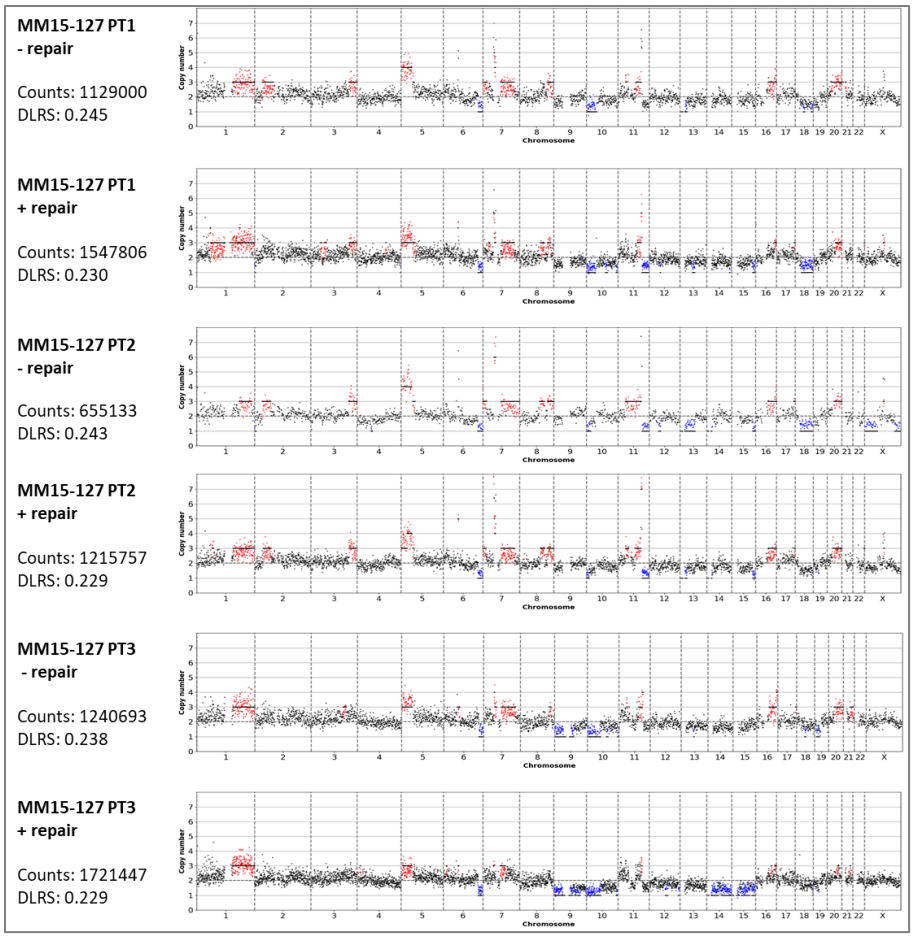


Figure 74: CNA profiles of DNA from primary tumor of MM15-127
CNA profiles of DNA isolated from 3 different punches (PT1-3) from primary tumor of MM15-127. WGA was performed without (-) or with (+) repair step. Amplifications are marked in red, deletions in blue color. DLRS: derivative log ratio spread.

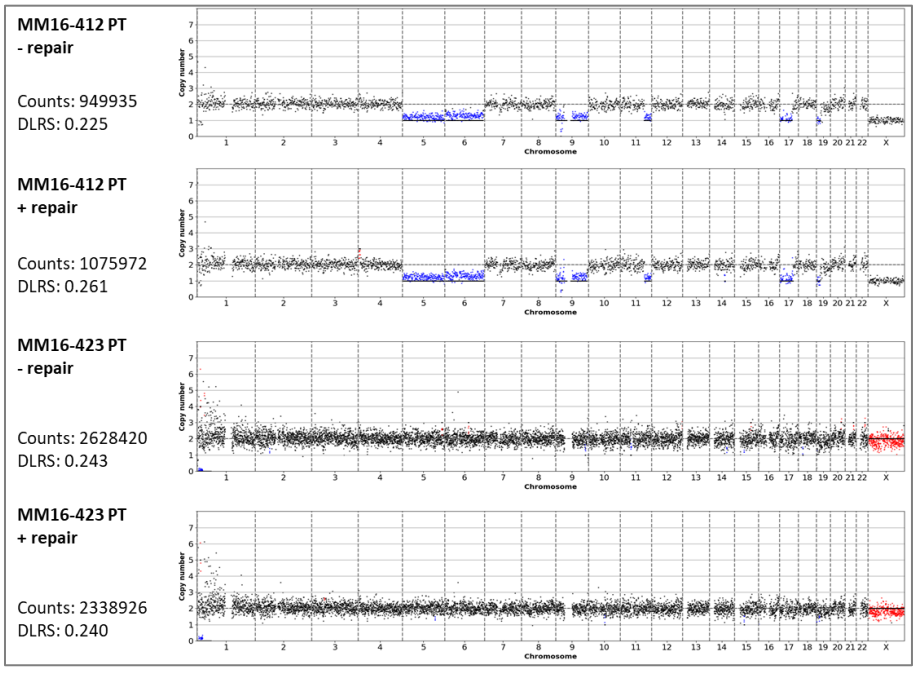


Figure 75: CNA profiles of DNA from primary tumors of MM16-412 and MM16-423
CNA profiles of DNA isolated from primary tumors of MM16-412 and MM16-423. WGA was performed without (-) or with (+) repair step. Amplifications are marked with red, deletions with blue color. DLRS: derivative log ratio spread.

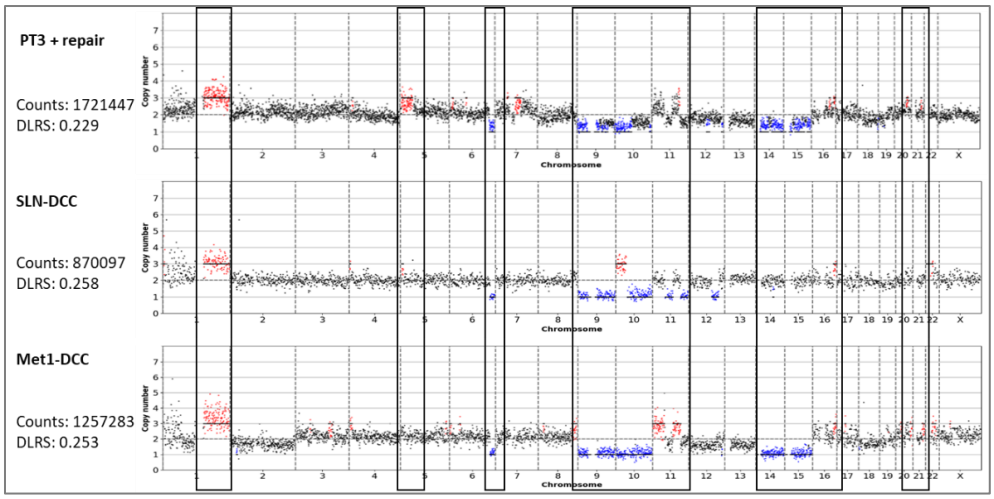


Figure 76: CNA profiles of PT, SLN-DCC and Met1-DCC from MM15-127
CNA profiles of primary tumor compared to CNA profiles of DCCs from SLN and Met1 of MM15-127. WGA of primary tumor sample was performed with (+) repair step. Amplifications are marked with red, deletions with blue color. PT: primary tumor; SLN: sentinel lymph node; Met: metastasis, DCC: disseminated cancer cell; DLRS: derivative log ratio spread.

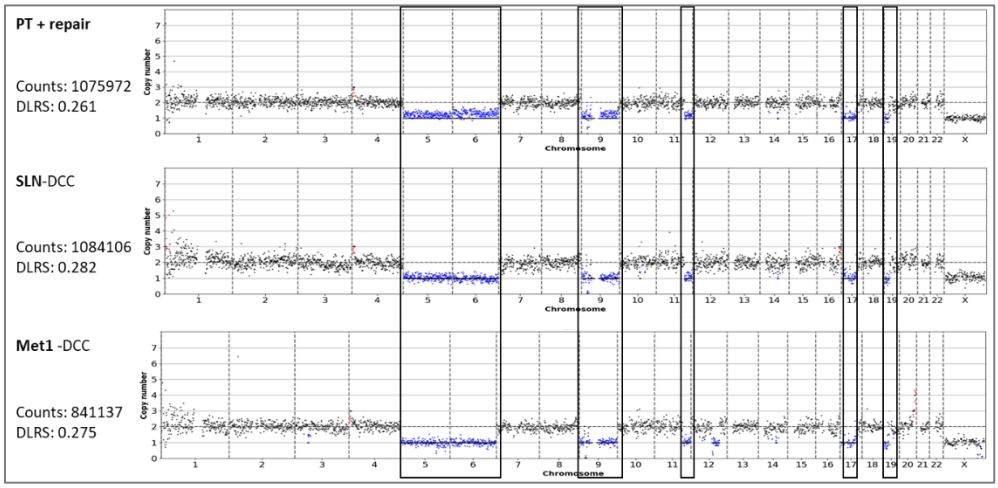


Figure 77: CNA profiles of PT, SLN-DCC and Met1-DCC from MM16-412
CNA profiles of primary tumor compared to CNA profiles of DCCs from SLN and Met1 of MM16-412. WGA of primary tumor sample was performed with (+) repair step. Amplifications are marked with red, deletions with blue color. PT: primary tumor; SLN: sentinel lymph node; Met: metastasis, DCC: disseminated cancer cell; DLRS: derivative log ratio spread.

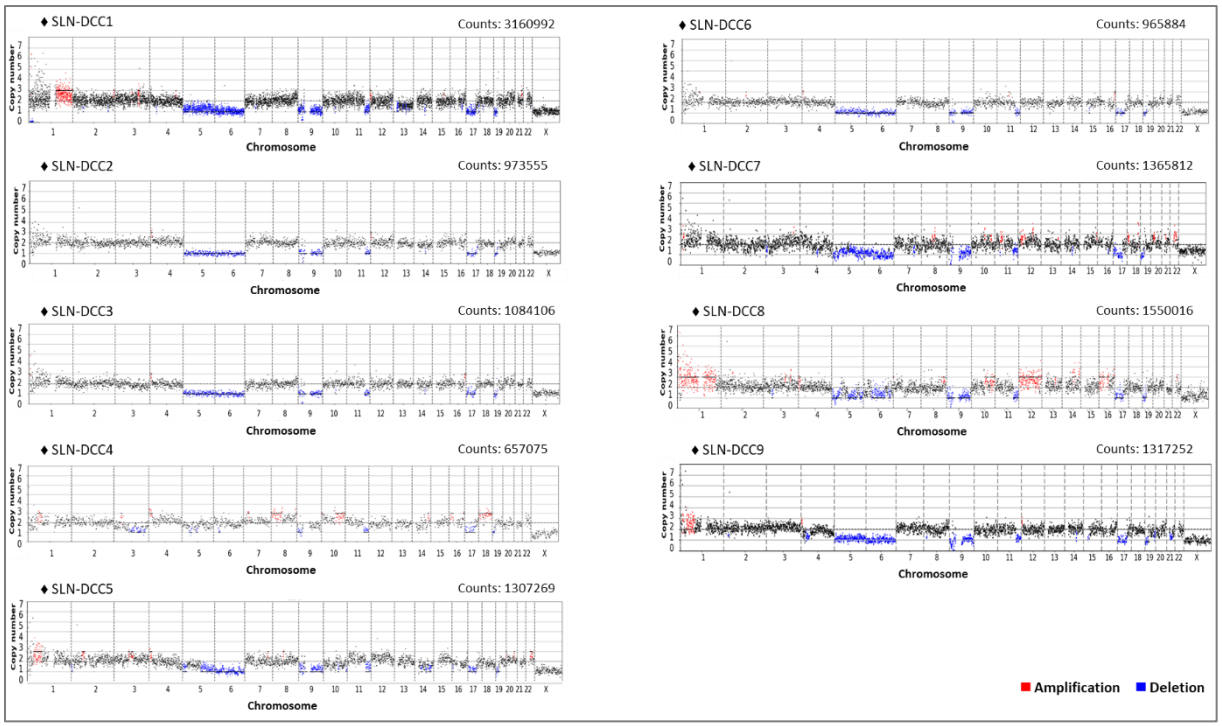


Figure 78: CNA profiles of SLN-DCCs from MM16-412
CNA profiles of SLN-DCCs from MM16-412's lineage tree. Amplifications are marked with red, deletions with blue color. All SLN-DCCs except SLN-DCC4 shared deletions on chromosome 5, 6, 9, 11q and 17.

Analysis	MM15-127	MM16-412	MM16-423	Total	Result
Screening for single tumor and control cells	249	454	286	989	Isolated single cells
WGA-QC	185	390	217	792	Cells with high quality DNA
Cell selection for WES	**127**	**163**	**93**	**383**	Cells with high quality DNA (no cell line cells, no control cells, but CD31+ for MM15-127, and 2 non-tumor cells for MM16-412/MM16-423)
QC for WES	106	116	64	286	Cells with sufficient quality and enough material for WES
Cell selection due to costs					
WES library preparation →	**66** **51** **15**	**116** **114** **2**	**64** **62** **2**	**246** **227** **19**	**Single cells** **Tumor cells** **Non-tumor cells** suitbale for WES

Figure 79: Single cell selection criteria for WES
Samples were selected by showing high-quality DNA after WGA, controls of WES (Qubit and Bioanalyzer). Single cell sequencing was not affordable for all high-quality SLN-DCCs. Libraries for WES were prepared with 246 single cells and additional bulk samples (not shown in the figure).

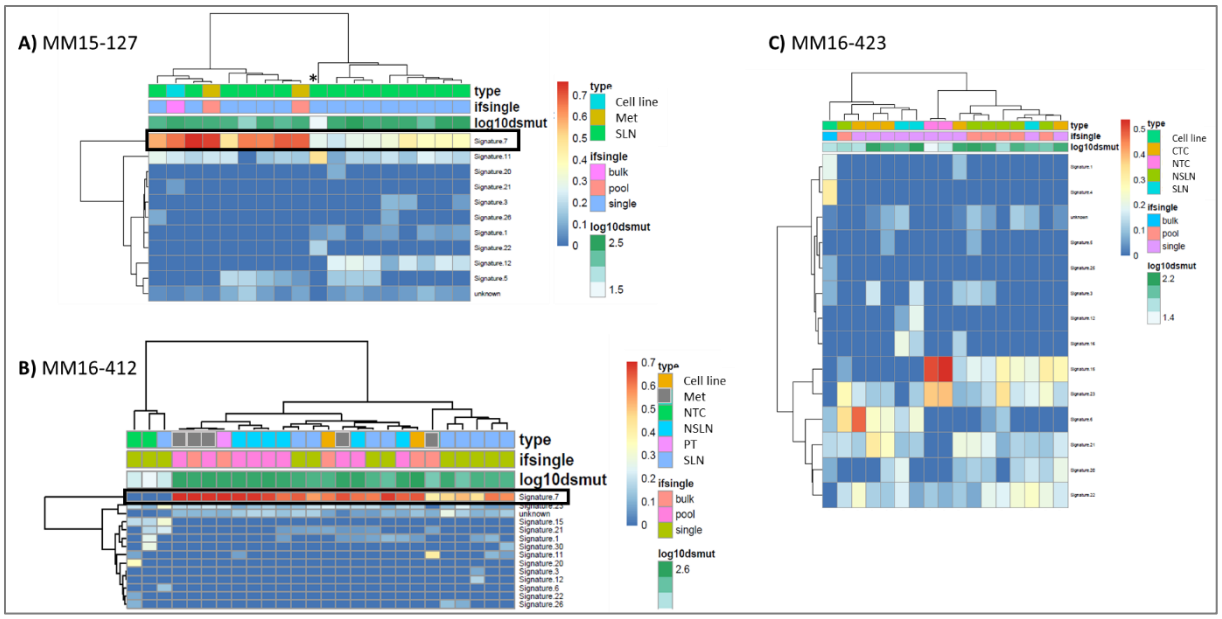


Figure 80: Mutational signature analysis using Platypus
Heatmap of mutational signature analysis of MM15-127 **(A)**, MM16-412 **(B)** and MM16-423 **(C)** using the Platypus algorithm. When detected, UV signature 7 was marked with black frame. Expression levels ranged from blue marked 0 to red marked 0.7 or 0.5. Number of detected mutations was represented by log10dsmut. Samples from MM15-127 **(A)** with no UV signature using varscan was marked with *. PT: primary tumor; SLN: sentinel lymph node; NSLN: non-sentinel lymph node; Mets: metastases; NTC: non-tumor cell.

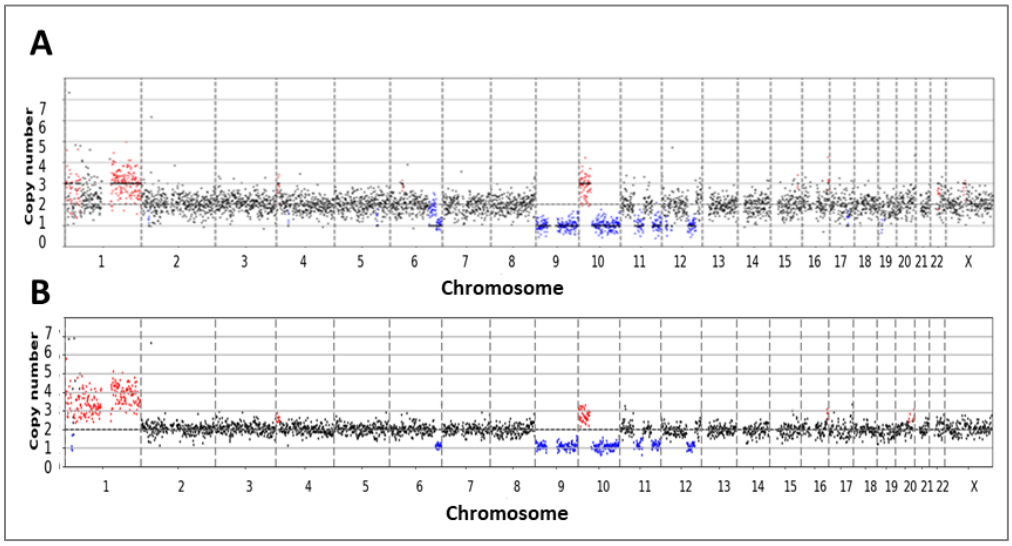


Figure 81: CNA profiles of SLN-DCCs from MM15-127
A: CNA profile of UV signature-negative SLN-DCC4. **B:** Representative CNA profile of UV signature-positive SLN-DCC. Amplifications are marked with red, deletions with blue color.

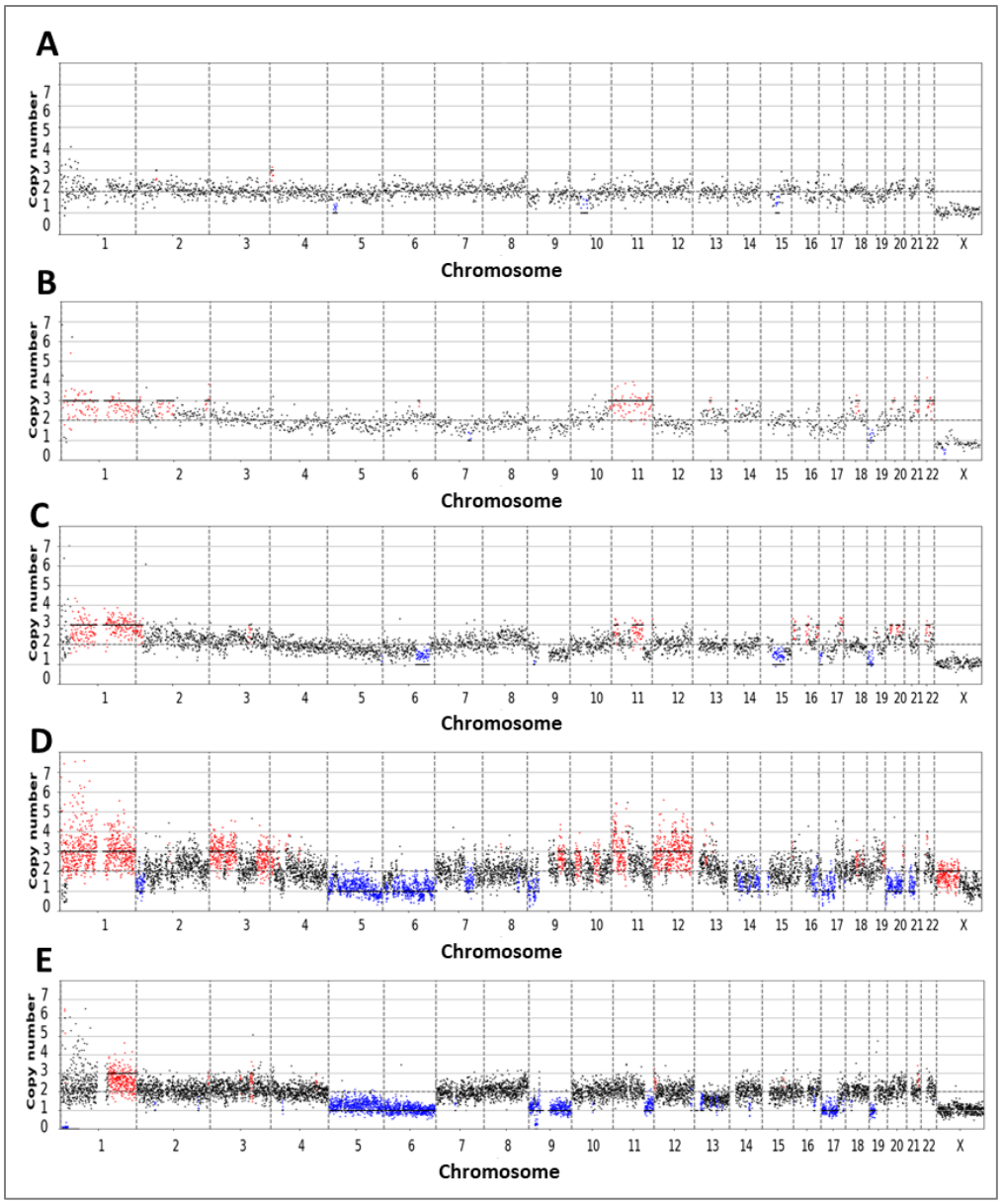


Figure 82: CNA profiles of SLN-DCCs from MM16-412
A to **D:** CNA profiles of SLN-DCCs not included in MS-based lineage tree. **A:** CNA profile of UV signature-negative SLN-DCC10. **B:** CNA profile of SLN-DCC11; **C:** CNA profile of SLN-DCC12; **D:** CNA profile of SLN-DCC13; **E:** Representative CNA profile of UV-signature positive SLN-DCC1. Amplifications were marked with red, deletions with blue color.

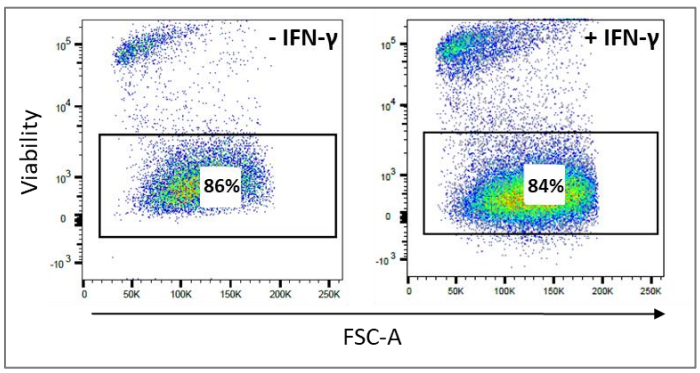


Figure 83: Viability analysis of MelDTC_11 by flow cytometry
Flow cytometric analysis of cell viability of MelDTC_11 with (right) or without (left) IFN-γ stimulation (100 U/ml) for 48 hours. (+): with; (-) without.

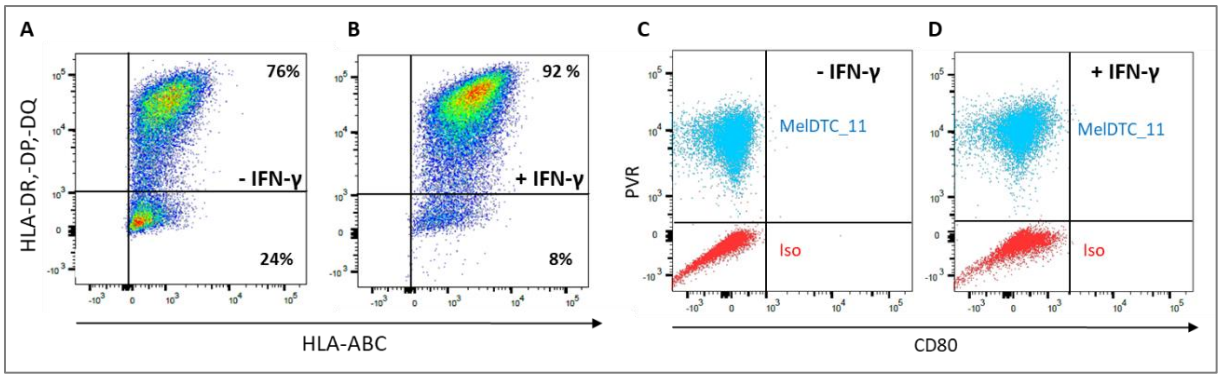


Figure 84: Flow cytometric analysis of HLA, CD80 and PVR expression
A and **B:** Flow cytometric analysis of HLA-ABC and HLA-DR,-DP,-DQ expression of MelDTC_10a with **(A)** or without **(B)** IFN-γ stimulation. **C** and **D:** Flow cytometric analysis of CD80 and PVR expression of MelDTC_11 with **(C)** or without **(D)** IFN-γ stimulation. IFN-γ treatment = 100 U/ml; 48 h; It was always gated on viable cells. (+): with; (-) without. Iso: isotype control.

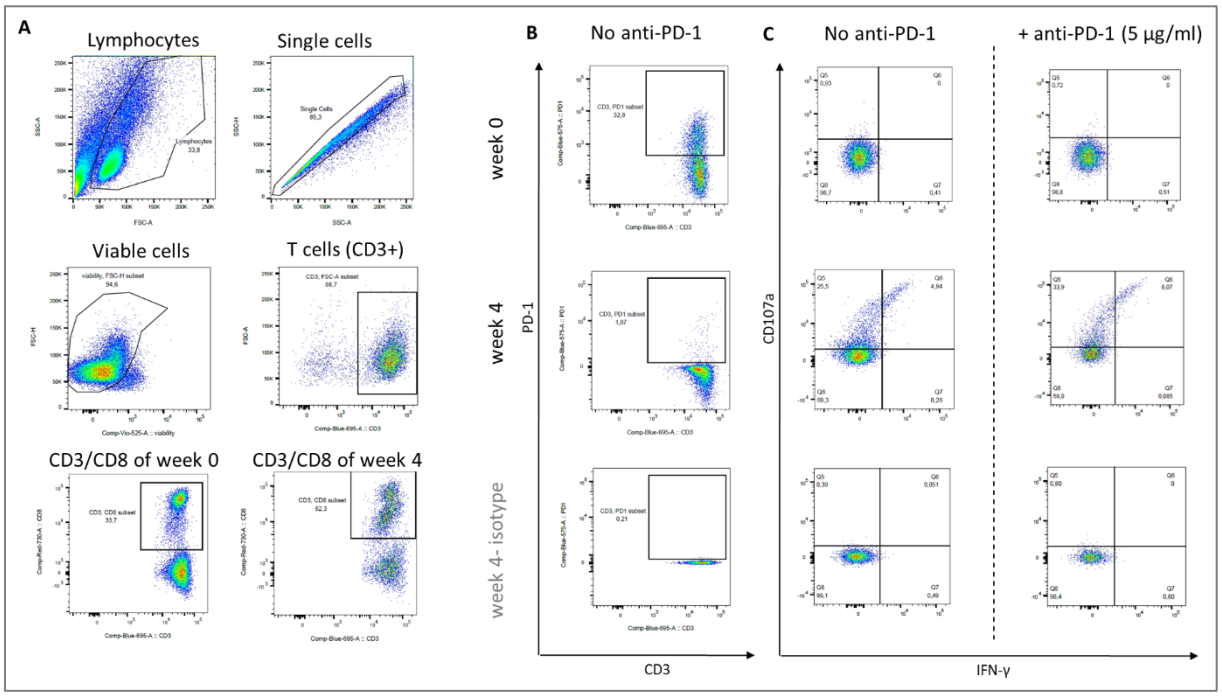


Figure 85: Gating strategy and marker expression of expanded T cells from MM16-423
A: Gating strategy for anti-PD1 expanded T cells. **B:** Flow cytometry analysis for PD-1 expression of CD8+ T cells from week 0 (upper picture) and week 4 (middle picture) compared to iso (lower picture). **C:** Flow cytometry analysis for CD107a/ IFN-γ expression of CD8+ T cells expanded without anti-PD-1 (left column) and with anti-PD-1 (right column). Marker expression was shown for T cells from week 0 (upper pictures) and week 4 (middle pictures) compared to iso (lower pictures).Iso: isotype.