Monitoring genotypic and phenotypic progression of systemic melanoma by cell lineage tree analysis and for molecular disease staging

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

vorgelegt von
Julia Greindl-Junghans,
geb. Junghans

aus
Regensburg

im Jahr 2020

Der Promotionsgesuch wurde eingereicht am:
30.10.2020

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

Unterschrift:

Table of contents

Table of abbreviations

ASB-PCR	Allele-specific PCR with a blocking reagent
ATP	Adenosine triphosphate
bp	Base pair
BSA	Bovine serum albumin
CNV	Copy number variation
DCC	Disseminated cancer cell
DCCD	DCC-density; number of DCCs per million mononuclear cells
EDTA	Ethylenediaminetetraacetic acid
FBS	Fetal bovine serum
FFPE	Formalin-fixed paraffin-embedded
gDNA	Genomic DNA
GII	genomic integrity index
h	hour
H2O	water
HUVECs	Human umbilical vein endothelial cells
LAD	Lymphadenectomy
LN	Lymph node
LT	Lineage tree
MCSP	Melanoma-associated chondroitin sulfate proteoglycan
Mets	Metastases
min	Minute(s)
MIPs	Molecular inversion probes
NBT/BCIP	nitro-blue tetrazolium chloride/5-bromo-4-chloro-3’-indolylphosphate p-tuluidine salt
OPA	One phor all buffer
PBLs	Peripheral blood lymphocytes
PBMCs	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PT	Primary tumour
QC	Quality control
RT	Room temperature
sec	Second(s)
SLN	Sentinel LN
STR	Short tandem repeat
TBE	Tris/Borate/EDTA buffer
T-Vec	Talimogen laherparepvec
UMI	Unique molecular identifier
WGA	Whole genome amplification

Summary

A recent study on metastatic seeding in melanomas showed that lymphatic dissemination occurs very early. Disseminated cancer cells (DCCs) of melanoma patients can leave the primary tumor (PT) in a genomically immature state, then evolve within the lymph nodes (LNs) and adapt to the ectopic site until they start to proliferate and form metastasis. Since the PT and the metastasis are often genetically disparate, the focus for treating metastases should be on the DCCs. The molecular characterization of the DCCs that left the PT at an early stage could reveal new therapeutic targets against metastasis. After routine LN removal in melanoma patients, staining of LNs against the tumour marker MCSP identified two different phenotypes: small MCSP-positive and large MCSP-positive DCCs. While the small phenotype appears mostly in LNs with a low DCCD (DCC-density; number of DCCs per million mononuclear cells), the large phenotype could be found in LNs with a higher DCCD. Furthermore, we also observed LNs with both small and large DCCs, that had a medium DCCD.

Based on these findings we hypothesized that small MCSP-DCCs are precursors of large MCSP-DCCs and represent very early DCCs. In addition, we wanted to have a closer look at the two most common *BRAF* mutations in malignant melanoma and its association with the DCCD of the LNs. We hypothesized that acquisition of *BRAF* mutations marks the transition from pre-colonizing DCCs to colonizing DCCs and hence a significant progression step in systemic cancer development.

The hypothesis if small MCSP-positive DCCs are the precursors of large MCSP-positive DCCs should be investigated with the help of a cell lineage tree reconstruction based on short tandem repeats (STRs). To study the incidence of the *BRAF* mutations we established an allele-specific PCR with a blocking reagent (ASB-PCR) for DCCs.

The cell lineage tree reconstruction of patient MM15-127 resulted in three distinct clusters of DCCs. Two of the clusters were found in close proximity to the PT, while one DCC cluster was closer to the metastatic tumour cells than the PT. Both small and large MCSP-positive DCCs were found in the two clusters close to the PT. The cluster closer to the metastatic tumour cells only contained large MCSP-positive DCCs.

Retrospective testing of 80 DCCs with the established ASB-PCR resulted in the correct identification of wild type and mutant DCCs in 98% and 96% of the samples, respectively. From patient MM16-423, DCCs were isolated from the sentinel lymph node (SLN) and the non-SLNs and tested for *BRAF* mutations by the ASB-PCR. While the PT and the DCCs isolated from the SLN at primary diagnosis were wild type, the DCCs isolated from non-SLNs after LN relapse harboured a *BRAF* mutation. Testing a cohort of 150 malignant melanoma patients for *BRAF* mutations in DCCs, showed that 19.8% patients with a pathologically negative LN and 59.4% with a pathologically positive LN harboured a mutation. However, studying the incidence of the *BRAF* mutation depending on the DCCD, we found out that there is a large increase of the *BRAF* mutation from 14.9% in LNs with a DCCD>1≤10 to 62.5% in LNs with a DCCD>10≤30.

Based on the result of the cell lineage tree reconstruction of patient MM15-127 our hypothesis that small MCSP-positive DCCs are the precursors of large MCSP-positive DCCs could neither be confirmed nor rejected. The resolution of the cell lineage tree is no yet good enough

to provide such accurate insights. However, three distinct clusters of DCCs were identified which could be an indication that DCCs disseminated at different time points. The ASB-PCR of DCCs from patient MM16-423 showed that *BRAF* mutations were acquired outside of the PT at a later time point of disease progression, when metastases were detected in the non-SLN. However, 62.5% of patients with a DCCD>10≤30 harboured a *BRAF* mutation, indicating that the *BRAF* mutation could be acquired early before colonisation of the DCCs.

1 Introduction

1.1 Malignant melanoma

1.1.1 Epidemiology and etiology

The prevalence of malignant melanoma is rising worldwide, especially in light-skinned populations with excessive sun exposure. In Europe, the incidence rate is <10-25 new melanoma cases per 100000 inhabitants, in the United States 20-30 per 100000 inhabitants, and in Australia 50-60 per 100000 inhabitants (Garbe et al. 2020). In many parts of Europe, the incidence is increasing at all ages and is predicted to continue to rise (Whiteman, Green, and Olsen 2016).

Cutaneous melanoma is classified as invasive melanoma when atypical melanocytes progressively invade into the dermis. Four different subtypes of invasive melanoma can be distinguished by clinical and histopathological features: superficial spreading melanoma is present in 41% of cases, nodular melanoma in 16%, 2.7-14% of cases represent lentigo maligna melanoma and 1-5% acral lentiginous melanoma. Nevertheless, histopathological subtypes are not included as prognostic factors in the current American Joint Committee on Cancer (AJCC) staging system for malignant melanoma (Garbe et al. 2020).

Characteristic alterations in signal transduction pathways have long been the focus of research. Mutations in the respective genes often lead to constitutive activation of the pathways and play a central role in the development and progression of malignant melanoma. The majority of melanomas have mutations in the mitogen-activated protein kinase (MAPK) pathway, involved in cell growth, proliferation, and survival (L. Cheng et al. 2018). Mutation leading to constitutive activation of the BRAF kinase are most common and occur in about 40-60% of melanoma cases. The second most common mutation is a NRAS mutation, occurring in about 15-30% of melanomas (Colombino et al. 2012; Curtin et al. 2005).

The major external factor is exposure to ultraviolet (UV) irradiation, particularly high sun exposure. While genetics and pathophysiological factors contribute to early-onset melanomas, cumulative sun exposure contributes to late-onset melanomas. Melanomas originating from chronically sun damaged (CSD) skin are often found on anatomic sites with higher exposure to UV light, like the head, the neck, and the dorsal surfaces of distal extremities of older individuals (>55 years). They have a high mutational burden and are associated with *NF1*, *NRAS*, *BRAF* (non-V600E), and *KIT* mutations. In contrast, non-CSD melanomas are found on intermediate sun-exposed areas such as the trunk or proximal extremities of younger individuals (<55 years) and are characterized by distinct mutations such as the $BRAF^{V600E}$ (Shain and Bastian 2016).

1.1.2 Diagnosis

The clinical diagnosis of malignant melanoma by a dermatologist is based on three analyses: (1) Examination with the naked eye according to the ABCD rule. Suspicious melanocytic lesions often fulfil one or more of the following criteria: Asymmetry (A), irregular borders (B), inhomogeneous colour (C) and a diameter ≥ 5mm (D). (2) Intra-individual comparative analysis, which means looking for a lesion that differs from the others (ugly duckling sign). (3) Chronic analysis of rapid and recent changes in a pigmented lesion, when it can be attested by the patient or compared to archived pictures (Grob and Bonerandi 1998).

To clarify the differential diagnosis, a dermatoscopic examination should be performed. Characteristic features considered by the dermatologist include atypical pigment network, irregular brown-black dots/globules, streaks, and pigmentation with multiple colours asymmetrically distributed (Garbe et al. 2020).

If a suspicious skin lesion is removed, a histological examination is required. Staging of the disease is important for the prognosis and the treatment recommendations for the individual patient that is based on the AJCC Melanoma Staging System. The recommendations of the up-to-date 8th edition AJCC Melanoma Staging System are supported by the analysis and monitoring of 46000 stage I-III melanoma patients since 1998. The classification is composed of the staging of the primary tumour (PT; T category), the regional lymph nodes (N category) and distant metastasis (M category), the so called TMN-status (Scolyer et al. 2020).

The T category is based on the tumour thickness and ranges from T0, meaning no evidence of PT (unknown primary or completely regressed primary melanoma), to T4 with a thickness >4.0 mm. The presence of ulceration is another important prognostic factor (explained in detail in **1.1.3**). Therefore, the T category is subdivided into PTs without ulceration (*a*) und with ulceration (*b*) (Scolyer et al. 2020).

The N category describes the sentinel lymph node (SLN) status after biopsy and is required for pathological staging of all patients with a PT >1.0 mm. The N category ranges from N0 to N3 and depends on the number of affected LNs. Stages N1 to N3 are further subdivided in *a* for micro-metastasis and *b* for macro-metastasis (explained in detail in **1.1.3**) (Scolyer et al. 2020).

The third category is the M category, describing whether distant metastases are present. M0 indicates no evidence of distant metastasis, while M1 means that distant metastases are present. Stage M1 is further subdivided depending on the anatomic site of the distant metastasis (Scolyer et al. 2020).

The T, N and M categories are summarized in the pathological stage group (**Table 1**). Malignant melanomas without regional LN metastasis or distant metastases are summarized in stages 0 to IIC. Affection of one or more reginal LNs with either micro- or macro-metastases are divided into groups IIIA to IIID. As soon as a patient is diagnosed with distant metastasis, the pathological classification is IV, independent of the PT and the SLN staging (Garbe et al. 2020).

Table 1: AJCC pathological prognostic stage groups

Tis: melanoma *in situ*, adapted from Garbe *et al.* 2020

T stage	N stage	M stage	pathological stage group
Tis	N0	M0	0
T1a or T1b	N0	M0	IA
T2a	N0	M0	IB
T2b or T3a	N0	M0	IIA
T3b or T4a	N0	M0	IIB
T4b	N0	M0	IIC
T0	N1b or N1c	M0	IIIB
T0	N2b, N2c, N3b or N3c	M0	IIIC
T1a/b or T2a	N1a or N2a	M0	IIIA
T1a/b or T2a	N1b/c or N2b	M0	IIIB
T2b or T3a	N1a - N2b	M0	IIIB
T1a – T3a	N2c or N3a/b/c	M0	IIIC
T3b or T4a	Any N ≥ N1	M0	IIIC
T4b	N1a – N2c	M0	IIIC
T4b	N3a – N3c	M0	IIID
Any T	Any N	M1	IV

1.1.3 Prognosis

If the melanoma is detected at an early stage, the chances of survival are very good. About 90% of melanomas are diagnosed without the presence of metastasis (stage 0 to IIC). The 5-year-survival rate for those patients is 75-95% (Garbe et al. 2020). The most important histopathological prognostic factors are the Breslow's depth, measured from the granular layer of the epidermis to the deepest point of invasion (Balch et al. 2000) and the presence of histological ulceration, defined as combination of full-thickness epidermal defect, evidence of host response and thinning, effacement, or reactive hyperplasia of the surrounding epidermis (Spatz et al. 2003).

The prognosis is poorer when the melanoma already metastasised to regional LNs (stage IIIA to IIIC). Regional metastasis can appear as satellite metastasis, defined as located up to 2 cm from the PT or in-transit metastasis, located between 2 cm from the PT and the first draining LN. Micro-metastasis, which are clinically not recognizable and can only be identified via sentinel LN biopsy can also appear in the regional LN. In contrast to micro-metastasis, macro-metastasis are clinically or radiologically recognizable in the regional LN (NCI Dictionary of Cancer Terms 2020). The 10-year survival rate of patients with satellite or in-transit metastasis is about 50%, 45-60% for patients with micro-metastasis and 25-35% for those macro-metastasis in the regional LN (Balch et al. 2009).

The prognosis for patients with distant metastasis is worse (stage IV). Depending on the localisation of the metastasis, the 5-year survival rate is 10 to 30%. Patients with distant metastasis in the skin, subcutaneous or in distant LNs have a better prognosis compared to patients with metastasis in the lung or non-pulmonary visceral metastasis for example in bone, brain or liver.

1.1.4 Treatment

The first step of malignant melanoma treatment is the surgical excision of the PT. All patients with invasive malignant melanoma are at risk of lymphogenic metastasis. Therefore, for the initial diagnosis and the treatment, it is important to examine the LN stations of these patients. If the tumour thickness exceeds 0.8 mm, or is thinner but ulcerated (stage T1b or higher), SLN biopsy is performed (Davis, Shalin, and Tackett 2019). SLN biopsy increases the 10-year survival rate of patients with intermediate-thickness melanomas (1.20-3.50 mm) and nodal metastasis. The 10-year survival rate is 42% without biopsy and increases to 62% when SLN biopsy is performed. Additionally, SLN biopsy significantly improved the distant disease-free survival of patients with nodal metastasis and intermediate-thickness melanomas (Morton et al. 2014). Besides the therapeutic benefit, SLN biopsy is a diagnostic measure to determine the stage, prognosis and adjuvant therapy.

For patients with cytological or histological proven LN metastasis, therapeutic lymphadenectomy (LAD) is recommended (S3-Leitlinie zur Diagnostik, Therapie und Nachsorge des Melanoms 2019). A recent study including 1755 patients showed that the rate of regional disease control and therefore the disease-free survival of patients who had a LAD significantly increased at 3 years. However, a significant increase in the melanoma-specific survival could not be measured (Faries et al. 2017).

The primary goal of adjuvant drug therapy is to extend the relapse-free survival or the overall survival of the patients. For patients with high risk malignant melanoma (stage IIB and above), adjuvant treatment with interferon alpha was investigated in several studies. Although studies included in meta-analyses were very heterogeneous in terms of patient population and dosing regimens, a significant, albeit small, benefit for overall survival and a significant benefit for progression-free survival was found (Ives et al. 2017). Patients with stage III or IV malignant melanoma can also be treated with checkpoint-inhibitors, such as anit-PD-1-inhibitors. Studies investigating the treatment with Pembrolizumab and Nivolumab showed a significant benefit for progression-free survival (Eggermont et al. 2018; Weber et al. 2017). BRAF and MEK inhibitors are another option for adjuvant therapy. In stage III patients with a BRAF V600E or V600K mutation, combinational treatment with Dabrafenib and Trametinib led to a benefit for progress-free and overall survival (Long, Hauschild, et al. 2017).

To date, therapeutic approaches for metastatic melanoma have not achieved satisfactory results for patients. For this reason, research and development of new treatment methods is currently focusing increasingly on personalized and targeted therapy at the molecular level. Fundamental to this is the understanding of the underlying mechanisms of the systemic dissemination of tumour cells.

1.2 Metastasis

Metastases account for the great majority of cancer-related deaths, yet the complex process of metastasis is not yet fully understood. About 90% of cancer-associated deaths are caused by metastatic disease rather than PTs because there are hardly any therapeutic treatment options up to now (Lambert, Pattabiraman, and Weinberg 2017). The dissemination of cancer cells from the PT and the formation of metastasis is a multi-step process (**Figure 1**). After tumour formation, vascularization of the PT is required for continued growth. Synthesis and secretion of

angiogenesis factors plays a crucial role in establishing a vascular network within the surrounding host tissue. Tumour cell invasion can occur by thin-walled venules or lymphatic channels and interaction of circulating tumour cells with hematopoietic cells can result in the formation of multicell aggregates. Surviving tumour cells can arrest in a capillary bed where they extravasate into organ parenchyma forming micrometastatic colonies. Subsequent proliferation of micrometastasis, vascularization and defense against the host immune response results in the formation of clinically detectable metastasis. Reinitiating the whole process leads to metastases from metastases (Talmadge and Fidler 2010).

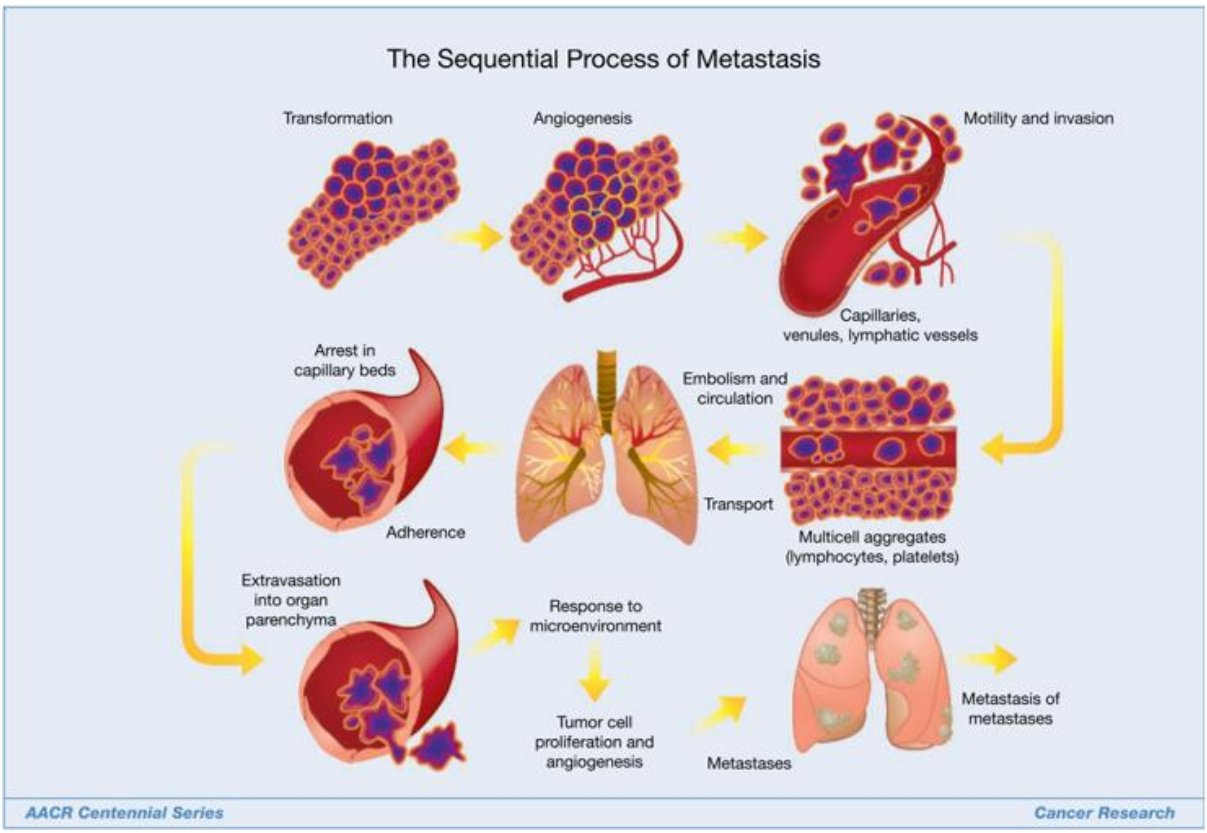


Figure 1: The process of cancer metastasis from the PT formation to clinically detectable metastasis. Schematic illustration adapted form Talmadge and Fidler, 2010

The initial step of the invasion-metastasis cascade is the invasion of tumour cells into blood or lymphatic vessels. To do this, the cells must acquire certain abilities to leave the primary site and travel to distant tissues. One important process enabling these steps is the epithelial-mesenchymal transition (EMT). Epithelial cells undergoing the EMT program gain increased motility, invasiveness and the ability to degrade components of the extracellular matrix. The EMT program is coordinated by a series of master EMT-inducing transcription factors like Snail, Slug, Twist and Zeb1 (Nieto et al. 2016; Pearlman et al. 2017). Although EMT is widely accepted to be important for the carcinoma cell dissemination, its exact role remains unclear. The migration of PT cells usually involves large cell aggregates rather than individual carcinoma cells. Cells residing within these cell aggregates still express epithelial markers such as E-cadherin, which help to maintain the cohesion between the epithelial cells within these aggregates (Cheung et al. 2016). This observation conflicts with the EMT program in which cells lost cohesive cell-cell

interactions. Therefore, this raises the question whether collective migration represents an alternative to EMT and if these cell-biological are mutually exclusive for the dissemination of carcinoma cells (Friedl et al. 2012).

After invasion of single cells or multi-cellular clumps into the vasculature, the so-called circulating tumour cells (CTCs) can travel to distant sites, where they may seed metastatic colonies. Once the tumour cells have entered the bloodstream, the CTCs are exposed to natural killer cells or fragmentation due to physical stresses encountered in transit through the circulation. On the other hand, CTCs gain physical and immune protection through interaction with platelets, which coat CTCs after intravasation. Furthermore, neutrophils can protect CTCs from natural killer cells and contribute to extravasation (Kang and Pantel 2013). Nevertheless, some experiments have shown very low efficiency (0.01%) of metastasis after intravenous injection of tumour cells (Chambers, Groom, and MacDonald 2002).

After extravasation, the disseminated cancer cells (DCCs, every tumour cell that left the primary lesion) find themselves in a new microenvironment to which they are poorly adapted. Therefore, most of the carcinoma cells are either eliminated from the tissue parenchyma or enter a state of dormancy, either as single DCCs or as small micrometastatic clusters, for weeks, months or even years (Luzzi et al. 1998). Metastatic colonization depends, as presently understood, on two preconditions of the DCCs: they must possess a tumour-initiating ability and they must adapt to the microenvironment of the distant tissue. The seed and soil hypothesis, put forward by Stephen Paget in the 19th century, states that the organ distribution of metastases does not happen by chance, but rather that tumour cells (the “seed”) grow preferentially in the microenvironment of selected organs (the “soil”) (Paget 1889). Paget’s hypothesis is supported by several experimental studies (Auerbach et al. 1987; Greene and Harvey 1964; Kinsey 1960). The opposing hypothesis is that the primary factor determining the pattern of tumour metastasis is the anatomy of vascular and lymphatic vessels draining from the site of the PT. According to this hypothesis, the tumour cells follow the circulatory route from the PT and arrest non-specifically in the first organ encountered (Ewing 1928). The prevailing opinion is that none of the two mechanisms is mutually exclusive, but that the extent to which either or both mechanisms are active depends on the tumour type (Langley and Fidler 2011).

The autopsy of 216 patients with advanced malignant melanoma showed that 95% of the patients had multiple organ metastases. Most metastases were found in the LNs (73.6%) and the lungs (71.3%). 58.3% were found in the liver and 49.1% - 43.5% in the brain, the bone, the heart, the adrenal glands, and the gastrointestinal tract (Patel et al. 1978). The fact that most metastases were found in the LNs is consistent with the observation that cancer cells often metastasize regionally through the lymphatic system before metastasizing systematically through the blood (Leong et al. 2011). In the blood stream, cancer cells are exposed to oxidative stress, which kills most of the DCCs. A recent study from Ubellacker *et al.* in mice showed that human melanoma cells in the bloodstream are killed by ferroptosis, leading to cell death by lipid oxidation of the polyunsaturated cell membrane. Analysing the blood and lymph from the mice, they found that the main lipids in the lymph were triglycerides containing oleate groups, derived from a monounsaturated fatty acid. These monounsaturated fatty acids were more abundant in lymph that in blood, as was iron, which is also essential for ferroptosis. In addition, Ubellacker *et al.* isolated mouse melanoma cells from subcutaneous tumours and from tumours in LNs and injected them into the bloodstream of mice. Melanoma cells from LNs were more likely to form metastasis

than cells from subcutaneous tumours, supporting their hypothesis that metastasis to the lymph nodes occurs before that to other locations because the lymph induces a metabolic advantage in the melanoma cells (Ubellacker et al. 2020). To what extent these findings apply to humans remains to be determined.

1.3 Concepts of metastasis

Once a PT is resected, metastases arising at a later time point must originate from tumour cells disseminated to ectopic sites before the PT was resected. However, it is not known which DCCs will grow into overt metastasis and which molecular features they require. Therefore, models of cancer progression are the key to predict which DCCs are important targets of systemic therapies. There are two fundamental models of metastasis, the linear progression model and the parallel progression model (**Figure 2**). The linear progression model states that cancer cells of the PT first accumulate genetic and epigenetic alterations before they start proliferation. Once the tumour reaches a certain size, individual cancer cells leave the PT and seed clonal metastases as a late event in the progression of the disease (Foulds 1954; Klein 2009). On the other hand, the parallel progression model concludes that metastasis may be initiated long before the PT is diagnosed. This implies that the DCCs already acquired the ability to leave the PT at an early stage, settle down in an ectopic microenvironment and evolve independently of the PT. Therefore, the parallel progression model predicts greater disparity between the metastatic founder cells and the PT than the linear progression model due to an independent and site specific accumulation of genetic and epigenetic alterations (Klein 2009).

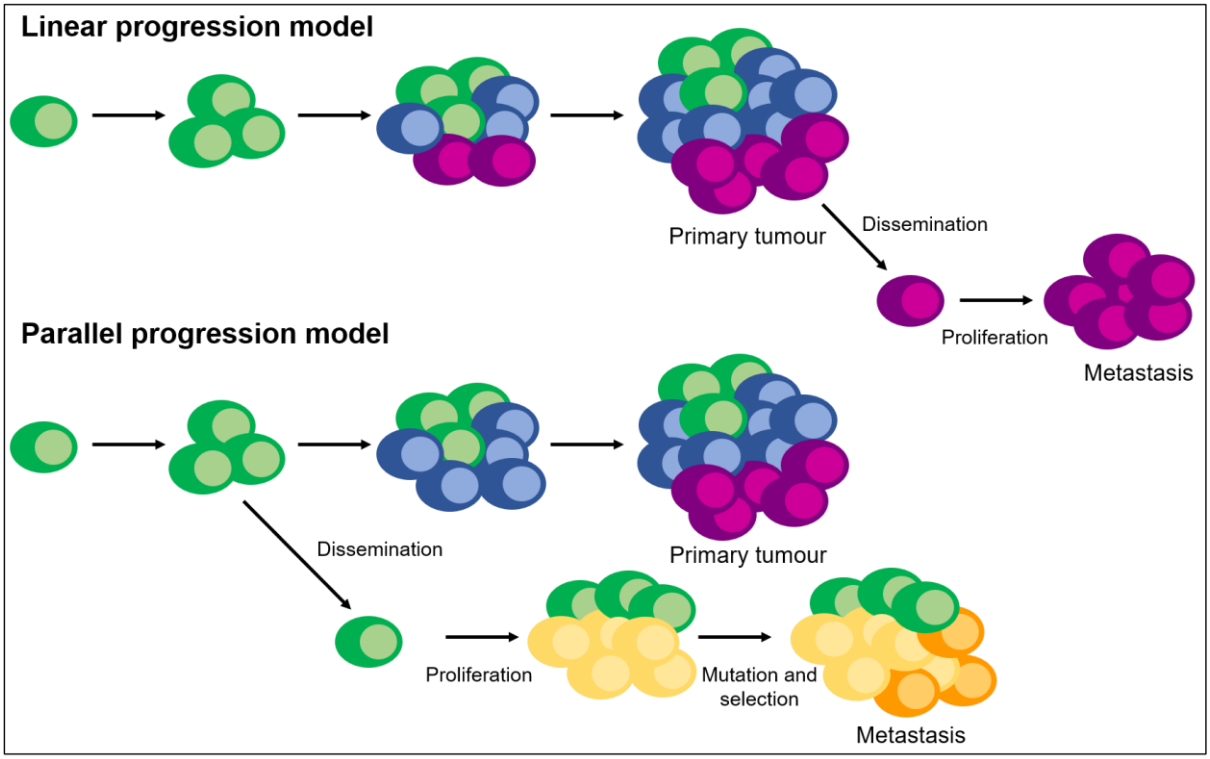


Figure 2: The linear progression model and the parallel progression model of metastasis

According to the linear progression model, tumour cells go through several rounds of mutation and selection before they start proliferation. Once the tumour reached a certain size, DCCs leave the PT forming clonal metastases at ectopic sites. In contrast, the parallel progression model states that dissemination is an early event and that DCCs mutate and adapt to the ectopic microenvironment in parallel to the PT.

A look at the current data, based on the course of disease of patients and the genetics of DCCs, favors the parallel progression model over the linear progression model (Klein 2009). According to the model of linear dissemination, the number of DCCs should increase with the tumour growth. However, DCCs were also detected in patients with small tumours, consistent with the parallel progression model predicting that dissemination is an early event. Furthermore, no increase of dissemination with tumour size was observed (Engel et al. 2003; Hölzel, Eckel, and Engel 2009; Hüsemann et al. 2008). Systemic treatment of cancer patients is increasingly moving in the direction of personalized medicine, targeting essential pathways on which the DCCs are dependent through genetic alterations. But genetic analyses of paired samples from PTs and DCCs have uncovered a striking genetic disparity, indicating that the PT is not suitable as a surrogate for the genetics of systemic cancer (Stoecklein and Klein 2009).

A study on the metastatic seeding of malignant melanoma showed that the lymphatic dissemination of DCCs occurs early. Clinical data from 1027 patients with clinically node-negative lymph nodes from previous studies was examined, where DCCs were isolated with the tumour marker gp100. Although there was a weak positive correlation of between the PT thickness and the disseminated cancer cell density (DCCD; number of DCCs per million mononuclear cells), the percentage of patients with DCCs in the SLN increased only marginally from T1 stage (PT ≤ 1 mm) with 45.8% to T4 stage (PT > 4 mm) with 59.4%. This observation suggests that dissemination occurs preferentially early (Werner-Klein et al. 2018). In addition, computational analyses revealed that lymphatic dissemination has occurred in 50% of the DCC-positive patients at a tumour thickness of <0.5 mm. The comparison of genomic profiles showed a striking disparity

between the PTs and their matching DCCs from SLNs. The PTs displayed significantly more deletions than DCCs, while the corresponding difference in gains was nonsignificant. Mutations in the *BRAF* and the *NRAS* oncogenes occur in about 40% and 21% of malignant melanoma patients, respectively. A comparison of the mutation status of matched patient samples showed that *BRAF* was mutated more frequently in PTs (34%) than in DCCs (15%), while no significant difference was observed for *NRAS* mutations. Mutations in the two oncogenes were shared in 16% of matching samples, 47% of samples shared a wild type status and 37% of samples differed in the mutation status. Matching DCCs were usually not sharing mutations with the PT, indicating that the DCCs left the primary site before the mutation was acquired. The copy number alterations and the targeted mutation analysis showed that the PT and the paired DCCs of melanoma patients are mostly genetically disparate, hints at an early dissemination of tumour cells (Werner-Klein et al. 2018).

1.4 Detection and analysis of DCCs from patients

Based on the previously discussed findings, DCCs disseminate at an early stage of melanoma progression. The identification and molecular characterization of the early DCCs is therefore the key to understanding the metastatic formation and for the identification of potential therapy targets. Compared to standard histopathology, immunocytochemistry staining against the melanocytic marker gp100 (antibody clone HMB45) resulted in an improved detection of DCCs in lymph nodes of melanoma patients (Ulmer et al. 2005). A subsequent study has proven the prognostic value of DCC detection in the SLN. Analysis of 1834 lymph nodes from 1027 malignant melanoma patients showed that an increased DCCD is associated with an increased risk to die of melanoma. Even patients with a very low number of DCCs detected in the SLN (DCCD ≤ 3) had an increased risk of dying compared to patients with a DCCD = 0 (Ulmer et al. 2014).

Another marker used for the detection of DCCs in LN is MCSP (Melanoma-associated-chondroitin-sulfate-proteoglycan; antibody clone 9.2.27). The antibody was already used for immunomagnetic cell sorting and isolation of CTCs in the blood of malignant melanoma cells. The detection of two or more CTCs correlated significantly with a reduced survival of the patients (Ulmer et al. 2004). The MCSP-staining was successfully established for the isolation of DCCs from LN of melanoma patients by Sebastian Scheitler in 2013 (Scheitler 2013). Using MCSP-staining, two different phenotypes of MCSP-positive DCCs were observed: small MCSP-positive DCCs with a diameter < 10 µm and intensively fluorescent large MCSP-positive DCCs with a diameter > 10 µm. After 233 LN samples were stained against MCSP, positive cells were detected in 145 samples. Large MCSP-positive cells were found in 26 samples (11.2%), small MCSP-positive in 103 samples (44.2%) and both phenotypes were found in 16 samples (6.9%). Although the small phenotype was the most abundant one, the median DCCD was only 4. When only large MCSP-positive cells were detected, the number of positive cells was significantly higher ($p<0.0001$ Mann-Whitney) with a median DCCD of 60. In mixed samples with both phenotypes the median DCCD was 15 (Scheitler 2013).

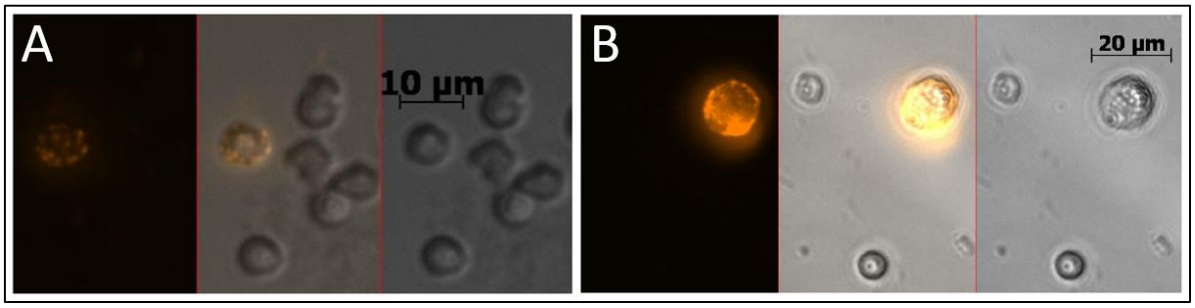


Figure 3: Morphology and phenotype of MCSP-positive DCCs in LNs

(**A**) Fluorescent, merged and bright field picture (from left to right) of the same small MCSP-positive cells with a diameter <10 µm, approximately the same size as the surrounding negative lymphocytes. (**B**) Fluorescent, merged and bright field picture (from left to right) of the same large MCSP-positive cells with a diameter >10 µm, an intensive staining. The cells are considerably larger than surrounding negative lymphocytes. The figure is adapted from Scheitler, 2013.

For these reasons, DCCs were isolated from LN that had to be resected for histological examination due to their stage. One half of the LN was histologically examined in routine diagnostics, the other half was used by our laboratory for immunocytochemistry. After mincing into small pieces and the mechanical disaggregation of the LN, the cell suspension was either stained for MCSP or transferred to adhesion slides for subsequent gp100 staining. Detected DCCs were isolated by micromanipulation and the single-cell genome was amplified by whole genome amplification (WGA), developed by our research group (Klein et al. 1999, 2002; Stoecklein et al. 2002). After the WGA, the genomic material was analysed by two different downstream analyses, the cell lineage tree reconstruction and the allele-specific PCR with a blocking reagent, that were described in more detail in the next chapter.

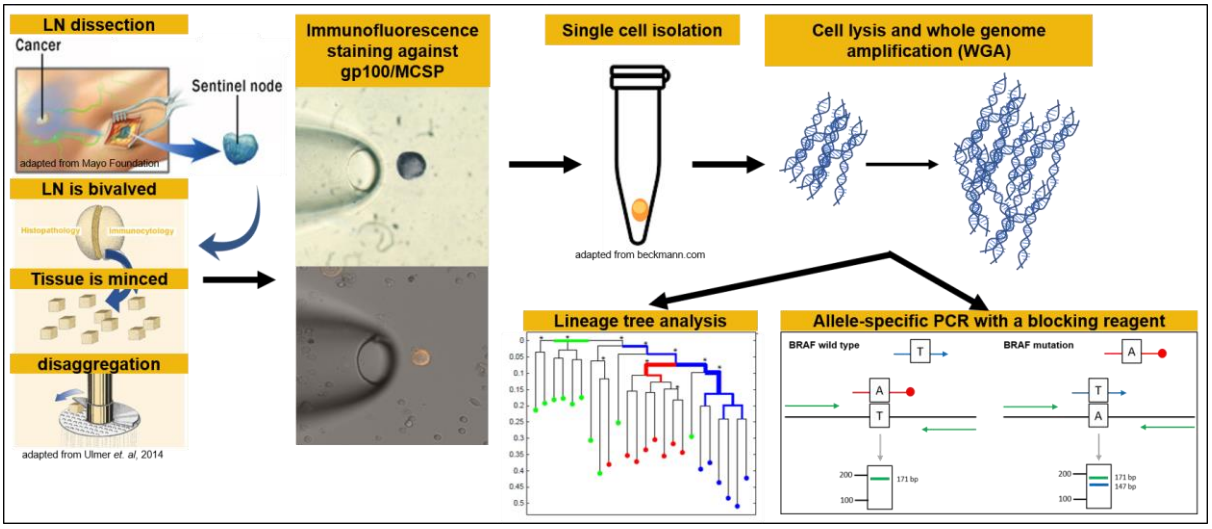


Figure 4: Workflow for the isolation of DCCs from LNs

The schematic illustrates how DCCs were isolated from LNs of malignant melanoma patients, followed by immunofluorescent staining against gp100 and MCSP and single cell isolation with a micromanipulator. Single cells were lysed, WGA was performed and the amplified DNA was used for downstream analysis. To investigate the sequence of mutational events happening during cancer progression lineage tree analysis and ASB-PCR was performed. The illustration for the LN dissection was adapted from the Mayo Foundation for Medical Education and Research, the LN processing from the publication of Ulmer *et al*., 2014 and the reaction tube from beckmann.com.

1.5 Downstream analysis

1.5.1 Cell lineage tree analysis

Uncovering the lineage relationships between cells in an organism is not only a fundamental interest of developmental biology, but can also advance the understanding of pathological states, including cancer. The first approach, in which the descendants of individual cells were characterized, was done by observation of *Caenorhabditis elegans* with time-lapse microscopy (Sulston et al. 1983). The revolution of next-generation sequencing led to the development of genetic lineage tracing methods using nucleotide sequences that serve as lineage barcodes (Shapiro, Biezuner, and Linnarsson 2013). For prospective genetic lineage tracing cells have to be genetically labeled with a barcode that is passed to the next generation of cells (Baron and van Oudenaarden 2019). This approach is suitable for organisms that could be genetically modified, such as the zebrafish (Junker et al. 2016). However, genetic modifications are impossible in the context of human development and disease. Therefore, naturally occurring barcodes in the nuclear DNA are used for the retrospective lineage tracing in humans, such as copy number variants (CNVs), single-nucleotide variants (SNV), long interspersed nuclear element 1 (LINE-1) retroelements and short tandem repeats (STRs) (Baron and van Oudenaarden 2019). STR, also known as microsatellites, are highly abundant regions with repetitive sequences of 1-6 bases. They are very prone to *de novo* mutations due to slippage events during DNA replication (Willems et al. 2014; Woodworth, Girskis, and Walsh 2017). Therefore, STRs are a promising mutational source to unravel the cell lineage of selected cells. Cell lineage analysis using STRs was already successfully applied to a mouse model to investigate the tumour development in lymphoma (Frumkin et al. 2008) and to human patients with acute leukemia to identify the relapse initiating tumour cells (Shlush et al. 2012). Our aim is the molecular characterisation of DCCs and the identification potential therapeutic targets in malignant melanoma. We hypothesize that early DCCs share the mutations with most of the DCCs found in the LN. The aim of this thesis is the identification of those early DCCs. Based on the findings of Sebastian Scheitler (Scheitler 2013) that were discussed above, we hypothesize that small MCSP-positive DCCs are precursors of large MCSP-positive DCCs and therefore the early DCCs. With the cell lineage tree reconstruction we aimed to test this hypothesis (see **1.4**).

1.5.2 Allele-specific PCR with a blocking reagent

The majority of melanomas have a mutation associated with the mitogen-activated protein kinase pathway that is important for cell growth, proliferation and survival (Burotto et al. 2014). The most common mutation occurs in *BRAF*, leading to a constitutive activation of the BRAF kinase. The *BRAF* c.1799G>A mutation is the most common one and occurs in up to 90% of cases, resulting in a substitution of valine for glutamic acid at position 600 (V600E). *BRAF* c.1798_1799GT>AA is the second most frequent one, leading to a substitution of valine for lysine (V600K) (L. Cheng et al. 2018). The observation that many melanomas harbour *BRAF* mutations led to the development of selective inhibitors of the BRAF V600-mutated kinase or the downstream MEK kinase (Long, Flaherty, et al. 2017; Ugurel et al. 2017). Therefore, several methods for the fast and easy detection of *BRAF* mutations in melanoma patients have been developed in recent years (Barbano et al. 2015; C. Cheng et al. 2013; Vallée et al. 2019). One of these methods is the allele-specific PCR with a blocking reagent (ASB-PCR), in which the mutated template can be amplified

with a background of abundant wild-type DNA (Morlan, Baker, and Sinicropi 2009). This method has so far only been performed with paraffin-embedded formalin fixed (FFPE) tissue but has not been established for single cells. With the establishment of the ASB-PCR for DCCs we aimed for a fast and easy assay for the detection of the two most common *BRAF* mutations in melanoma. First, a patient with several LNs resected at different time points of disease progression should be tested for the *BRAF* mutations. The aim was to investigate the DCCs isolated from different LNs for the incidence of the *BRAF* mutation. In the next step, we planned to study a larger cohort of early (LN pathologically negative) and progressed stage (LN pathologically positive) melanoma patients.

1.6 Aims of the thesis

For the identification of new therapeutic targets against metastatic cells, tumour cells that left the PT at an early stage have to be identified. Molecular characterization of those early DCCs could reveal genetic targets that are shared by most metastatic cells. The first aim of this thesis was the identification of early DCC in malignant melanoma patients. In malignant melanoma, metastasis usually begins in the regional lymph nodes before the tumour cells spread to distant organs (Leong et al. 2011). Using the tumour marker MCSP small and large phenotypes of DCCs were observed. While the small MCSP-positive DCCs were mainly found in LNs with a small DCCD (number of DCCs per million mononuclear cells; median DCCD of 4), the large MCSP-positive DCCs were mainly detected in LNs with a higher DCCD (median DCCD of 60). LNs containing both small and large MCSP-positive DCCs had a median DCCD of 15. Based on these findings we hypothesized that small MCSP-positive DCCs are progenitors of large MCSP-positive DCCs and represent very early DCCs. To test this hypothesis, we adapted the cell lineage tree reconstruction to single DCCs of malignant melanoma patients (together with the Shapiro Laboratory, Weizmann Institute of Science, Israel).

The second aim of the thesis was to study the incidence of the two most common *BRAF* mutations in malignant melanoma patients and its association with the DCCD in LNs. We hypothesized that the acquisition of *BRAF* mutations marks the transition from pre-colonizing DCCs to colonizing DCCs and hence a significant progression stop in systemic cancer development. Therefore, we wanted to establish a fast and easy detection method, the so-called ASB-PCR, to detect the two most common mutations in *BRAF* (c.1799T>A and c.1798_1799GT>AA) in DCCs. The investigation of the incidence of *BRAF* mutations depending on the lymph node status and the DCCD should provide us with the information when the mutations first occur.

In summary, the two central questions of this thesis are the following:

- Are small MCSP-positive DCCs the precursors of large MCSP-positive DCCs and therefore represent very early DCCs?
- Do the *BRAF* mutations mark the transition from pre-colonizing DCCs to colonizing DCCs?

2 Materials

2.1 Reagents

2.1.1 Antibodies and staining solutions

Name	Manufacturer	Catalogue number
AP-conjugate substrate Kit	Bio Rad	1706432
AP-polymer anti-rabbit solution	Zytomed Systems	ZUC031-006
Cy™3 AffiniPure goat anti-mouse IgG	Jackson Immunores.	115-166-071
Goat anti-mouse IgG Alexa Fluor 488	Invitrogen	A-11029
Goat anti-rabbit IgG Alexa Fluor 555	Invitrogen	A-21428
Melanosome (gp100), clone HMB-45	Agilent Dako	M0634
Mouse anti-CD68, clone KP1	Agilent Dako	M081401-2
Mouse anti-human CD31, clone JC70A	Dako	M082329-2
Mouse IgG1k, clone MOPC21	Sigma Aldrich	M29269
Purified mouse anti-chondroitin sulfate (MCSP), clone 9.2.27	Becton Dickinson	554275
Rabbit anti-CD3	Sigma Aldrich	C7930
Rabbit IgG	Southern Biotech	0111-01

2.1.2 Oligonucleotides

Oligonucleotides used for the ASB-PCR were obtained from Metabion, all other oligonucleotides were obtained from Eurofins Germany in HPSF grade.

* These primers amplify a polymorphic DNA section on chromosome 5. Since this is a length polymorphism, the amplicon size may vary for each individual and between two alleles.
** in combination with the 3'-allele-independent primer

Name	Base sequence (5' → 3')	Amplicon size (bp)	Application
ddMSE11	TAA CTG ACAG ddC	--	WGA
Lib1	AGT GGG ATT CCT GCT GTC AGT	--	WGA
hD5S2117 for hD5S2117 re	CCA GGT GAG AAC CTA GTC AG ACT GTG TCC TCC AAC CAT GG	140*	WGA-QC
hKRT19 for hKRT19 rev	GAA GAT CCG CGA CTG GTA C TTC ATG CTC AGC TGT GAC TG	621	WGA-QC
hKRAS for hKRAS rev	ATA AGG CCT GCT GAA AAT GAC CTG AAT TAG CTG TAT CGT CAA GG	91	WGA-QC
hTP53 Exon 2/3 for hTP53 Exon 2/3 rev	GAA GCG TCT CAT GCT GGA TC CAG CCC AAC CCT TGT CCT TA	301	WGA-QC
KRAS_91bp_5' KRAS_91bp_3'	ATAAGGCCTGCTGAAAATGAC CTGAATTAGCTGTATCGTCAAGG	91	FFPE WGA-QC Mix 1
PIK3CA_Exon20 (HS2)-5' PIK3CA_Exon20 (HS2)-3'	TCTAGCTATTCGACAGCATGC TTGTGTGGAAGATCCAATCCAT	221	FFPE WGA-QC Mix 1
EGFR_Exon 19_5' EGFR_Exon 19_3'	TCCTCGATGTGAGTTTCTGC ATGCCTCCATTTCTTCATCC	350	FFPE WGA-QC Mix 1
BRAF_Exon 15_5' BRAF_Exon 15_3'	CTCTTCATAATGCTTGCTCTG TCCAGACAACTGTTCAAACTG	171	FFPE WGA-QC Mix 2
TP53_Exon 7_5'	GAGGCTGAGGAAGGAGAATG	400	

TP53_Exon 7_3‘	AGTATGGAAGAAATCGGTAAGAGG		FFPE WGA-QC Mix 2
NRAS_Exon 2_5‘ NRAS_Exon 2_3‘	ACACCCCCAGGATTCTTACA TCCGCAAATGACTTGCTATT	174	FFPE WGA-QC Mix 3
TP53_Exon 8_5‘ TP53_Exon 8_3‘	AGGTAGGACCTGATTTCCTTACTG AGGCATAACTGCACCCTTG	245	FFPE WGA-QC Mix 3
EGFR_Exon 21_5‘ EGFR_Exon 21_3‘	CAGCGGGTTACATCTTCTTTC AAACAATACAGCTAGTGGGAAGG	418	FFPE WGA-QC Mix 3
EGFR_Exon 18_5‘ EGFR_Exon 18_3‘	TTGTCCTTCCAAATGAGCTG TGCCTTTGGTCTGTGAATTG	496	FFPE WGA-QC Mix 3
TP53_Exon 5/6_5‘ TP53_Exon 5/6_3‘	ACGCATGTTTGTTTCTTTGC ACCCCTCCTCCCAGAGAC	1034	FFPE WGA-QC Mix 4
EGFR_Exon 20_5‘ EGFR_Exon 20_3‘	AAACGTCCCTGTGCTAGGTC CATGGCAAACTCTTGCTATCC	1324	FFPE WGA-QC Mix 4
CKND2A_Exon3_5‘ CKND2A_Exon3_3‘	TGGCTCTGACCATTCTGTTC TGGAAGCTCTCAGGGTACAA	1367	FFPE WGA-QC Mix 4
5‘-mutant-specific primer	GGACCCACTCCATCGAGATTTCT	147**	ASB-PCR
blocking oligonucleotide	CGAGATTTCACTGTAGCTAG-*PO_4*	--	ASB-PCR
5‘-allele-independent primer 3‘-allele-independent primer	TCCAGACAACTGTTCAAACTG CTCTTCATAATGCTTGCTCTG	171 bp	ASB-PCR

2.1.3 Chemicals and commercial solutions

Name	Manufacturer	Catalogue number
1 kb Plus DNA Ladder + Dye	New England Biolabs	N3200L
AB serum, human	Bio Rad	805135
Adenosine triphosphate (ATP) 100 mM	Roche Diagnostics	11140965001
Agarose LE	Anprotec	AC-GN-00009
AMPure XP purification beads	Beckman Coulter	A63882
Bacto™ Peptone	Becton Dickinson	211677
Boric acid	Sigma Aldrich	31146-500G
Bovine serum albumin (20 mg/ml) for PCR	Roche Diagnostics	10711454001
Bovine serum albumin for picking	Sigma Aldrich	B8667-5ml
Collagenase A	Sigma Aldrich	C0130
DMEM/F12 with 15 mM HEPES	PAN Biotech	P04-41250
DNaseI	Roche Diagnostics	471672801
dNTP Set; 100 mM each; A, C, G, T; 4x 24 µM	GE Healthcare	28-4065-51
Ethanol absolut ≥ 99.8%	VWR Chemicals	20821.330
Ethidium Bromide Solution (10 mg/ml)	Sigma Aldrich	E1510-1ML
Ethylendiamintetraacetic acid (EDTA)	J.T. Baker	B1073.1000
Expand Long Template Buffer 1	Roche Diagnostics	11759060001
FastStart DNA Polymerase	Roche Diagnostics	4738420001
FastStart dNTP mix	Roche Diagnostics	4738420001
FastStart PCR buffer with $MgCl_2$	Roche Diagnostics	4738420001
Fetal bovine serum sera Plus	PAN Biotech	P30-3702
FFPE repair mix	New England Biolabs	M6630S
Formaldehyde, 37%	Merck	104003
Gel loading dye 6x, purple, no SDS	New England Biolabs	B7025S
Hank’s solution	Biochrom AG	L2045

Hyaluronidase	Sigma Aldrich	H4272
Igepal CA-630 sixcous liquid	Sigma Aldrich	I3021-5ml
Levamisol	Sigma Aldrich	L9756
Low molecular weight ladder	New England Biolabs	N3233L
Mayer's hematoxylin	Sigma Aldrich	MHS16
Monopotassium phosphate for analysis	VWR International	1048731000
Orange G	Sigma Aldrich	O7252-25G
Penicillin (10000 U/ml) / Streptomycin (10 mg/ml)	PAN Biotech	P06-07100
Percoll™	GE Healthcare	17089101
Phosphate buffered saline for picking	Gibco	10-010-023
Sodium chloride solution 5 M	Sigma Aldrich	71386-1L
Tris ultrapure for biochemistry	AppliChem	A1086,1000
Trypan blue	Sigma Aldrich	T8154-20ml
Trypsin 0.05%/EDTA 0.2%	PAN Biotech	P10-024100
Tween® 20, for molecular biology	Sigma Aldrich	P9416-50ml
Water for chromatography, LiChrosolv®, LC-MS grade (PCR-water)	Merck	1.15333.1000
Water UltraPure, DEPC-treated (DEPC-water)	Invitrogen	750023
Water, aqua ad iniectabilia (NGS-water)	Braun	2351744

2.1.4 Commercial kits

Name	**Manufacturer**	**Catalogue number**	**Application**
Ampli1™ LowPass Kit (Set A+B)	Agilent Technologies	5067-4626	LowPass Library preparation
AP conjugate substrate Kit (BCIP/NBT)	Bio Rad	170-6432	Staining
Bioanalyzer DNA High Sensitivity Kit	Agilent Technologies	5067-4626	LowPass sequencing
DNeasy blood & tissue Kit	Qiagen	69504	DNA isolation from LN
Expand Long Template PCR System	Roche Diagnostics	11759060001	PCR Reaction
FastStart™ Taq DNA Polymerase	Roche Diagnostics	4738420001	WGA
KAPA Library Quantification Kit	Roche Diagnostics	07960298001	LowPass library quantification
MiSeq Reagent Kit v3 (150 cycles)	Illumin	MS-102-3001	LowPass library quantification
QIAamp DNA FFPE Tissue Kit	Qiagen	56404	gDNA isolation from PT
Qubit® dsDNA BR Assay Kit	Thermo Fisher Scientific	Q32853	DNA quantification
Qubit® dsDNA HS Assay Kit	Thermo Fisher Scientific	Q32854	DNA quantification

2.1.5 Custom buffers and solutions

Name	Components	Application
AB serum 10%/ peptone 2% solution	5 ml AB serum, human 5 ml 20% peptone solution 40 ml 1x PBS (pH 7.4)	MCSP staining
Igepal 10%	2 ml Igepal CA-630 18 ml DEPC-water	WGA
Loading dye	0.3% Orange G 25% Ficoll TBE	Agarose gel electrophoresis
NBT/BCIP-solution	100 µl AP development buffer* 25 µl AP reagent 1* 25 µl AP reagent 2* 25 µl Levamisol Distilled water ad 2.5 ml * from AP-conjugate substrate kit, Bio Rad	Staining
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 1L Sterile filtrated	WGA
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	Staining
PBS/10% AB-Serum	1 ml AB serum, human 9 ml 1x PBS, pH 7.4	Staining
Peptone solution 20%	100 g Bacto™Peptone 500 ml 1x PBS (pH 7.4) Sterile filtration with 0.45 µM filter	MCSP staining
Percoll 100%	100 ml Percoll stock 9 ml Hank's solution Sterile filtrate	Preparation of Percoll 60%
Percoll 60%	30 ml Percoll 100 ml 20 ml NaCl	Processing of LNs
PFA 0.5%	1.35 g formaldehyde solution 100 ml Mili-Q water	Staining
Tris/Borate/EDTA (TBE) buffer 10x	539 g Tris 275 g Boric acid 37 g EDTA 5 l demineralized water	Agarose gel electrophoresis
Tween 10%	2 ml 100% Tween® 20 18 ml DEPC-Water	WGA

2.1.6 Consumables

Name	Manufacturer	Catalogue number
CellSave Preservative Tubes	Menarini Silicon Biosystems	7900005
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
Diagnostic adhesion slides	Thermo Fisher Scientific	ER-203B-CE24
Eppendorf twin.tec®-PCR-Platte 96 LoBind, skrited	Eppendorf	0030129512
Eppendorf. protein LoBind tube 0.5ml	Eppendorf	0030 108.094
Eppendorf. protein LoBind tube 1,5ml	Eppendorf	0030108.116
Eppendorf. protein LoBind tube 2ml	Eppendorf	0030 108.132
Eppendorf. protein LoBind tube 5ml	Eppendorf	0030 108.302
Falcon® 40 µm cell strainer	Becton Dickinson	352340
LabTek Chamber Slides, glass, 8 fields	Nunc	11367764
MAXYMum Recovery™ 0.2 ml	Axygen Scientific	11370145
Medicons	Becton Dickinson	340591
Microcentrifuge tube 1.5 ml	Greiner Bio-One	616201
Microcentrifuge tube 2 ml	Greiner Bio-One	623201
Micro-hematocrit capillary	Brand	749321
Multichannel Reagent reservoirs	Integra	4331
PCR SingleCap 8er-SoftStrips 0.2 ml	Biozym	710970
PCR tube 0.2 ml	4titude Deutschland	4ti-0795
SafeSeal Surphob filertips 10 µl	Biozym	770010
SafeSeal Surphob filertips 1250µl	Biozym	VT0270
SafeSeal Surphob filertips 20 µl	Biozym	VT0220
SafeSeal Surphob filertips 200 µl	Biozym	VT0240
Superfrost® Plus slides	Thermo Fisher Scientific	10149870
Transparent 96-well PCR plate	Biozym	710884

2.1.7 Devices

Name	Manufacturer	Application
Bioanalyzer 2100	Agilent Technologies	LowPass sequencing
CellTram Pump	Eppendorf	DCC isolation
Centrifuge 5424	Eppendorf	WGA reamp purification
Centrifuge Plate Fuge	Benchmark Scientific	PCR plate centrifugation
DMZ Universal Puller	Zeitz	DCC isolation
DNA Engine Peltier Thermal Cycler	Bio Rad	WGA
DNA Engine Tetrad2 Peltier Thermal Cycler	Bio Rad	PCRs
Electrophoresis chamber 40-1214	Peqlab	Agarose gel electrophoresis
Genetouch thermal cycler	Bioer	PCRs
HiSeq4000	Illumina	RNA-Seq
Incubator Heraeus BB15	Thermo Fisher Scientific	Cell culture

Laminar flow bench, Her Safe KS18	Thermo Fisher Scientific	LN processing, WGA
Magnetic Rack, FastGene MagnaStand 2.0	Nippon Genetics Europe	DNA purification
Manual pipettes (2µl, 10µl, 20µl, 200 µl, 1000 µl)	Gilson	Molecular biology
Medimachine	Dako	LN preparation
Microscope Axiovert 200M	Zeiss	DCC isolation
Microscope CX23	Olympus	DCC isolation
Microscope IX81, inverted	Olympus	DCC isolation
Microwave	Micromaxx	Agarose gel electrophoresis
MiSeq	Illumina	LowPass sequencing
Multipette Stream	Eppendorf	PCRs
Nanodrop 2000c	Thermo Fisher Scientific	DNA quantity measurement
Neubauer hemocytometer	Schubert & Weiss	Cell counting
Patchman NP2 micromanipulator	Eppendorf	DCC isolation
pH-meter PB-11	Sartorius	pH adjustment of buffers
PCR bench UCT-S-AR	Thermo Fisher Scientific	PCR, LowPass
Power Supply MP-250N	Kisker Biotech	Agarose gel electrophoresis
Qubit3 fluorometer	Thermo Fisher Scientific	NGS

3 Methods

3.1 Patient samples and collaborations

3.1.1 Clinical cooperation

Sentinel LNs (SLN) of malignant melanoma patients were obtained from patients in cooperation with Dr. med. Sebastian Haferkamp and Professor Dr. med. Mark Berneburg from the Department of Dermatology, University Hospital Regensburg. The non-sentinel LNs and the metastasis from malignant melanoma patients were received in cooperation with Dr. med. Philipp Renner from the Department of Surgery, University Hospital Regensburg.

3.1.2 Cooperation within the working group

The sample collection and preparation for the lineage tree analysis as well as the CNV analysis were performed in cooperation with Sandra Huber. The patients MM15-127 and MM16-423 were of interest to both of us, but with different objectives. While Sandra Huber was interested in the order of disseminated cancer cell (DCC) seeding from sentinel LNs to non-sentinel LNs and the formation of metastasis, I was interested in the progression of disease based on the cell morphology and mutation status. Due to an enormous size of the sample collective of around 1000 single cells we decided to share the workload.

The establishing work for the lineage tree analysis (see **4.5.2** and **4.5.4**) was performed in cooperation with Sandra Huber and Manjusha Ghosh. The samples used for the establishment were from three different projects, including three melanoma patients from the project of Sandra Huber, one breast cancer patient from Manjusha Ghosh and two melanoma patients from this thesis, overlapping with the patients of Sandra Huber. The reason for that was to have a big sample collective for the best possible establishment.

3.1.3 Inclusion criteria

Inclusion criteria for the study of the metastatic progression on the morphological level and for the lineage tree analysis were i) phenotypically small and large MCPS-positive DCCs, ii) a DCC-derived cell line of the patient and iii) a DCCD below 100. For studying the metastatic progression on the anatomical level the inclusion criteria were: i) Patient from whom we received a sentinel LN and non-sentinel LNs, ii) from which we were able to isolate DCCs and iii) patients who developed metastasis.

3.1.4 Ethics

Written informed consent was obtained from all patients involved in all the experiments. The study was approved by the local ethics committee of the University of Regensburg (ethics vote numbers 07-079, 18-948-101).

3.1.5 Bioinformatic collaborations

Cell lineage tree analysis was performed in collaboration with the group of Dr. Ehud Shapiro, Department of Computer Science and Applied Mathematics, Weizmann Institute of Science Rehovot, Israel.

3.2 Sample preparation of different tissue types

3.2.1 Preparation of LN tissue

With the patient's consent, the LN is divided into two halves. One half remains in the pathology department for routine histological examination while the second half of the sample is further processed as quickly as possible for the planned single cell detection. The sample is stored in RPMI 1640 medium at 8 °C until mechanical disaggregation to produce the cell suspension. For the mechanical disaggregation of the tissue, it is first divided manually with scalpels into small pieces of approximately 2 x 2 mm. The pieces are disaggregated one by one in Medicons (BD-Bioscience), inserts of the Medimaschine (Dako). This disaggregation is performed in cell culture medium (RPMI 1640). The cell suspension is first separated from coarse impurities in a 50 ml Falcon with Hank's saline solution for 10 min at 200 g and 4 °C, the supernatant is removed to about 7 ml and discarded. The cell pellet is resuspended, transferred to a 15 ml Falcon for density gradient centrifugation and carefully layered on Percoll 60%. When adding the suspension, ensure that the two phases do not mix and that a clear dividing line is formed. Centrifugation is carried out for 20 min at 1000 g and 4 °C. The mononuclear cells separate from the erythrocytes and collect in the interphase. They are transferred into a new 50 ml Falcon and washed with PBS for another 10 min at 500 g and 4 °C to remove any remaining Percoll residues. After removal of the supernatant, the cell pellet is resuspended in 0.5 to 5 ml of PBS, depending on its size, and the cell count is determined in a Neubauer counting chamber. The cell suspension is then either used for direct staining (see 3.6) or for the preparation of adhesion slides (see 3.3).

3.2.2 Preparation of metastasis

The metastatic tissue was put in Basal medium consisting of DMEM/F12 Medium with 15mM Hepes, 1% penicillin-streptomycin and 1% BSA. It was cut into smaller pieces using a scalpel. To obtain single cell suspension, tissue pieces were enzymatically digested in a final concentration of 0.33 mg/ml collagenase, 100 µg/ml hyaluronidase and 100 µg/ml DNaseI for 20-40 minutes at 37 °C. At the end of the digestion step, samples were mixed by pipetting and transferred into a 50 ml tube. Following a washing step with PBS at 300g for 10 min at RT, the cell pellet was resuspended in pre-warmed (37 °C) Trypsin 0.05%/EDTA 0.2% in PBS and treated for five minutes at RT. Digestion was stopped by adding 5 ml of Basal medium supplemented with 10% FCS. Next, cells were filtered using a 40 µm cell strainer and washed with 14 ml PBS at 300g for 10 min at RT. The resulting cell pellet was resuspended in PBS and transferred onto adhesion slides (see 3.3).

3.2.3 Preparation of peripheral blood

Blood samples for the isolation of circulating tumour cells (CTCs) were collected in CellSave tubes containing a fixative. PBMCs were washed with Hanks solution at 200g for 10 min at 4°C. The cell pellet was diluted in 7 ml Hanks solution followed by centrifugation at 1000g for 20 min at 4 °C on a 60% Percoll density gradient. The interphase containing PBMCs was collected and washed with PBS at 200 g for 10 min at 4 °C. PBS was discarded, and the cell pellet was resuspended in fresh PBS and counted in a Neubauer counting chamber. PBMCs were transferred onto adhesion slides at a density of 10^6 cells/side (0,33 x 10^6cells/spot). Following sedimentation of cells on the slide for 1 h at RT, residual PBS was discarded, slides were air-dried overnight at RT and stored at -20 °C.

3.3 Preparation of adhesion slides

For the preparation of adhesion slides, the cell density of the single cell suspension after the tissue preparation (see 3.2) was set to 10^6 cells/ml in PBS. 333 µl of the cell suspension were transferred onto each of the three spots of an adhesion slide, resulting in 0.33 x 10^6 cells per spot. After sedimentation for 1 h at RT, the residual PBS was discarded, and the slides were air-dried overnight at RT. The slides were stored at -20 °C until later use.

3.4 Isolation of gDNA from primary tumour FFPE blocks

First, sections of the FFPE blocks were prepared and stained with haematoxylin and eosin (HE staining). Pathologist Dr. Florian Weber marked the tumour containing areas on the HE sections, which were then superimposed on the block to locate the tumour area. The punches were taken with a disposable biopsy punch (1.5 mm) and placed on an UV sterilized glass slide. The excess paraffin was cut out using a scalpel. An additional UV sterilized glass slide was used to grind the tissue, followed by cutting the tissue into pieces with a scalpel. The pulverized tissue was carefully transferred into a 1.5 ml tube that was shortly centrifuged to collect the material at the bottom. Next, gDNA was isolated using the QIAamp FFPE Tissue Kit according to the manufacturer's instructions. The isolated gDNA was quantified using the NanoDrop ND-1000.

3.5 Staining of non-tumour control cells on adhesion slides

3.5.1 Single staining against CD3

For the establishment of the double staining, staining with the rabbit anit-CD3 antibody had to be tested first. After thawing of adhesion slides with LN cell suspensions for 30 min at RT, unspecific binding was blocked using PBS/10% AB-serum for 30 min at RT. To test different concentrations, the adhesion slide was stained with 0,2 mg/ml and 0,1 mg/ml rabbit anti-CD3 in PBS/10% AB-serum for 60 min at RT. Each spot of the adhesion slide was covered with 150 µl antibody dilution. Rabbit IgG was used as isotype control in the higher concentration on a separate adhesion slide. After washing the slides with PBS 3 times for 3 min, the cells were stained with 3 drops of the AP-polymer anti-rabbit solution for 30 min at RT. After washing again 3 times for 3 min with PBS, 150 µl of the NBT/BCIP-solution was added for 10 min at RT. The solution was prepared with the AP-

conjugate substrate kit. After washing 3 times for 3 min with PBS, the cells were fixated with 1% PFA in PBS for 5 min. After washing 3 times for 3 min with PBS, the slides were stored in PBS at 4 °C.

After successful establishment of a primary antibody concentration of 0,1 mg/ml, secondary antibody dilutions were tested to replace the AP-polymer anti-rabbit solution. A dilution of 8 µg/ml and 4 µg/ml goat anti-rabbit IgG-Alexa Fluor 555 in PBS was incubated for 30 min at RT. The incubation of the secondary antibody and all subsequent steps were performed in the dark. After washing again 3 times for 3 min with PBS, the cells were fixated with 1% PFA in PBS for 5 min. After washing 3 times for 3 min with PBS, the slides were stored in PBS at 4 °C.

3.5.2 Double staining against CD3 and CD68

Non-tumour control cells were isolated from adhesion slides of LN samples. After thawing of adhesion slides for 30 min at RT, unspecific binding was blocked using PBS/10% AB-serum for 30 min at RT. To isolated T-cells and macrophages from the same adhesion slide, it was stained with 0,1 mg/ml rabbit anti-CD3 and a dilution of 1:500 mouse anti-CD68 antibody in PBS/10% AB-serum for 60 min at RT. Each spot of the adhesion slide was covered with 150 µl antibody dilution. Rabbit IgG and mouse IgG1k were used as isotype control in the same concentration and dilution for CD3 and CD68 on a separate adhesion slide for each patient. After washing the slides with PBS 3 times for 3 min, the cells were stained with 4 µg/ml goat anti-rabbit IgG-Alexa Fluor 555 and goat anti-mouse IgG-Alexa Fluor 488 in PBS for 30 min at RT. The incubation of the secondary antibodies and all subsequent steps were performed in the dark. After washing again 3 times for 3 min with PBS, the cells were fixated with 1% PFA in PBS for 5 min. After washing 3 times for 3 min with PBS, the slides were stored in PBS at 4 °C.

3.5.3 Single staining against CD31

For the isolation of endothelial cells as non-tumour control cells, staining with the mouse anti-human CD31 antibody had to be established. After thawing of adhesion slides with LN cell suspensions for 30 min at RT, unspecific binding was blocked using PBS/10% AB-serum for 30 min at RT. To test different dilutions, the adhesion slide was stained with a 1:250 and a 1:500 dilution of the mouse anti-human CD31 antibody in PBS/10% AB-serum for 60 min at RT. Each spot of the adhesion slide was covered with 150 µl antibody dilution. Mouse IgG1k was used as isotype control in the higher concentration on a separate adhesion slide. After washing the slides with PBS 3 times for 3 min, the cells were stained with 3 drops of the AP-polymer anti-mouse solution for 30 min at RT. After washing again 3 times for 3 min with PBS, 150 µl of the NBT/BCIP-solution was added for 10 min at RT. The solution was prepared with the AP-conjugate substrate kit. After washing 3 times for 3 min with PBS, the cells were fixated with 1% PFA in PBS for 5 min. After washing 3 times for 3 min with PBS, the slides were stored in PBS at 4 °C.

3.6 Staining of tumour cells for single cell isolation

We use two different melanoma markers for the isolation of DCCs from LNs: i) Adhesion slides are stained against gp100 for our routine diagnosis of malignant melanoma patients. ii) For

research purposes, fresh LN cell suspension is stained against MCSP. CTCs were identified and isolated using gp100 staining.

3.6.1 gp100 staining of adhesion slides

Adhesion slides of LNs or peripheral blood were thawed for 30 min at RT. All following incubation steps were also performed at RT. Blocking was done by adding 300 µl 10% AB-serum in PBS per spot for 30 min in a slide chamber. After discarding the blocking solution by tapping the adhesion slides on a paper towel, 150 µl antibody solution (1:50 dilution of anti-gp100 antibody in PBS with 10% AB-serum) was incubated for 60 min. Mouse IgG1k was used as isotype control on a separate adhesion slide. After washing three times for 3 min with PBS in a cuvette, 150 µl of an anti-mouse AP-polymer solution with 10% AB-serum were incubated for 30 min. After washing again three times for 3 min with PBS, the slides were incubated with 150 µl NBT/BCIP-solution was added for 10 min at RT. The solution was prepared with the AP-conjugate substrate kit. After washing three times for 3 min with PBS, the cells were fixated with 1% PFA in PBS for 5 min. After washing three times for 3 min with PBS, the slides were stored in PBS at 4 °C.

3.6.2 MCSP staining of cell suspensions

If LN cells were left over after the preparation of adhesion slides, up to 3 million cells were transferred into a 1.5 ml tube. All centrifugation steps were performed for 5 min at 500 g and RT. After centrifugation, the supernatant was discarded. The cell pellet was resuspended in 95 µl 10% AB-Serum and 2% peptone in PBS and incubated for 10 min at 4 °C on a shaker to block unspecific binding sites. Subsequently, 5 µl of the monoclonal murine MCSP antibody (25 µg/ml) (see 2.1.1) were added and again incubated for 10 min at 4°C on a shaker. After centrifugation, the supernatant was discarded and the cell pellet was washed in 500 µl PBS with 2% peptone. After discarding the supernatant, the cell pellet was resuspended in 98 µl 10% AB-Serum and 2% peptone in PBS and 2 µl of the secondary Cy3 conjugated goat anti-mouse antibody (75 µg/ml; see 2.1.1) were added. The duration of incubation was 10 min at 4 °C on a shaker. After centrifugation and discarding the supernatant, the cell pellet was again washed with 500 µl PBS with 2% peptone. At the end, the pellet was resuspended in 200 µl PBS.

3.7 Single-cell isolation by micromanipulation

Single cells were isolated from cell suspensions or adhesion slides using a micromanipulator (PatchMan NP2, Eppendorf) in combination with an inverted fluorescence microscope (Olympus IX81 or Zeiss Axiovert 200M). The slides were systematically screened for positive cells. Single cells were isolated from the cell suspension or carefully scratched from the adhesion slide without destroying their integrity. Cells were aspirated using a glass capillary covered with FCS and a diameter of 30 µm. The single cells were transferred into a separate field of an 8 chambered slide that was precoated with BSA and 200 µl PBS. The single cells were isolated manually with a micropipette, by aspirating each cell in 1 µl PBS and transferring it into a 0.2 ml PCR tube (Axygen, MAXYMum Recovery) containing 2 µl cell lysis buffer for WGA.

3.8 Isolation of genomic DNA from the SLN

For the isolation of genomic DNA from the SLN cell suspension of patient MM15-127 the DNeasy Blood and Tissue Kit (QIAGEN) was used according to manufacturer's instructions. The isolated gDNA was eluted in 50 µl of H_2O.

3.9 Whole genome amplification

3.9.1 Principle

The WGA protocol is used to amplify minute amounts of single cell DNA (Klein et al. 1999; Stoecklein et al. 2002) now commercially available as *Ampli1*™ WGA Kit. After the cells are lysed in the picking buffer, template DNA is digested using the MseI restriction enzyme, to obtain fragments that are around 150-1500 bp long. In parallel, asymmetric double stranded adaptors are ligated in a step called pre-annealing. In the next step, the DNA fragments are then ligated to asymmetric double stranded adaptors. One of the adaptors' strand is lacking a phosphate to prevent its ligation to the template DNA. The non-ligated adaptor strand is removed by heat denaturation creating an overhang of the so-called Lib1 oligonucleotide. The primary PCR reaction fills up the previously generated overhangs and uses the adapted Lib1 molecule as a primer for amplification. The WGA products are then stored at -20 °C or -80 °C for long-time storage.

3.9.2 Experimental procedure

Compositions of all master mixes are shown in **Table 2**. After isolation, single cells, cell pools or gDNA isolated from FFPE blocks were collected in lysis buffer containing proteinase K. The samples were incubated for 10 h at 42 °C (for single cells and cell pools) or 15 h at 42 °C (for gDNA isolated from FFPE blocks). Proteinase K was inactivated for 10 min at 80 °C and then cooled down to 4 °C. In the next step, the DNA template was digested with MseI for 3 h at 37 °C. In parallel, double stranded adaptors were prepared by pre-annealing at 65 °C and gradually decreasing temperature by 1 °C per minute until the temperature reached 15 °C. After 3 h of digestion, the MseI reaction was inactivated for 5 min at 65 °C and afterwards cooled down to 4 °C. Next, ATP and T4 liagse were added to the pre-annealing mixture, which was subsequently added to the MseI-digested samples. Ligation was performed over night at 15 °C. The next day, the primary PCR master mix was added to the samples and amplified according to the program in **Table 3**.

Table 2: Composition of master mixes for WGA
* amount for single cells and cell pools
** amount for gDNA isolated from FFPE blocks

Master Mix	Reagent	*Volume/sample (µl)	**Volume/sample (µl)
Lysis buffer	OPA 10x	0.2	0.5
	Tween 10%	0.13	0.13
	Igepla 10%	0.13	0.13
	Proteinase K (10 mg/ml)	0.26	0.26
	PCR-water	1.28	3.48
MseI digestion	OPA 10x	0.2	--
	MseI 50000 U/µl	0.2	0.25
	PCR-water	1.6	0.25
Pre-annealing	OPA 10x	0.5	0.5
	Lib1 100 µM	0.5	0.5
	ddMSE11 100 µM	0.5	0.5
	PCR-water	1.5	1.5
Primary PCR	Buffer 1 (Roche)	3	3
	dNTPs 10 mM	2	2
	DNA Pol Mix	1	1
	PCR-water	34	34

Table 3: Cycler program for WGA

Step	Temperature (°C)	Duration (h:min:sec)	Cycles
1	68	00:03:00	1
2	94	00:00:40	15
3	57	00:00:30	
4	68	00:01:30 + 1 sec/cycle	
5	94	00:00:40	9
6	57 + 1/cycle	00:00:30	
7	68	00:01:45 + 1 sec/cycle	
8	94	00:00:40	23
9	65	00:00:30	
10	68	00:01:53 + 1 sec/cycle	
11	68	00:03:40	1
12	4	forever	1

3.9.3 WGA of gDNA isolated from primary tumour FFPE blocks

The WGA of gDNA isolated from FFPE blocks was either performed after the standard protocol (see 3.9.2) or with an additional repair step. To test if the performance of the samples improves in the lineage tree analysis, we used a DNA repair enzyme mix from New England Biolabs including various DNA repair enzymes (Taq DNA Ligase, Endonuclease IV, Bst DNA Polymerase, Fpg, Uracil-DNA Glycosylase, T4 Endonuclease V and Endonuclease VIII). The DNA repair was integrated prior to the MseI digestion step in the WGA protocol (see 3.9.2). 1 µl of a gDNA concentration around 50 ng/µl was added to the lysis buffer, followed by 15 h incubation at 42 °C. Next, the DNA was treated with the FFPE repair mix by incubating 20 min at 37 °C, followed by enzyme inactivation at 80 °C for 20 min. After cooling down to 4 °C, the MseI digestions was

performed with half the amount of the standard WGA. The master mix composition is provided in **Table 4**. The rest of the WGA was identical to chapter 3.9.2.

Table 4: Composition of master *mix for FFPE repair*

Master Mix	Reagent	Volume/sample (µl)
FFPE Repair	FFPE repair buffer	0.4
	FFPE repair mix	0.1
MseI digestion	OPA 10x	0.1
	MseI 50000 U/µl	0.1
	PCR-water	0.8

3.9.4 Quality control PCR for WGA samples of single cells

Successful amplification of the genome was confirmed by endpoint PCR. For this purpose, four genes *KRAS*, *KRT19* and *TP53* Exon 2/3, as well as a polymorphic DNA area on chromosome 5 using the D5S2117 primer located on different MseI fragments were amplified. All primers are listed in **2.1.2**. 9 µl of the reaction master mix, shown in **Table 5**, were mixed with 1 µl of the primary WGA product. To control for the functionality of the assay and the purity of reagents, a positive and a negative control were included. The PCR reaction was performed after the following protocol: the DNA was denatured at 95 °C for 4 min, followed by 32 cycles of 95 °C for 30 sec, 58 °C for 30 sec and 72 °C for 90 sec. At the end, a final elongation step was performed at 72 °C for 7 min followed by cooling down to 4 °C. The amplified DNA was loaded on an agarose gels for analysis (see 3.10). The number of bands displayed on the gel is translated into the genome integrity index ranging from 0 to 4 (GII; Polzer et al. 2014). With a GII 4 all three largest amplicons (*KRT19*, *TP52* Exon 2/3 and D5S2117) were amplified, with a GII 3, two of the three largest amplicons were amplified and with GII 2 only one of the three largest amplicons was amplified. GII 2 to GII 4 are independent of the *KRAS* fragment, the smallest amplicon. GII 1 means that only the *KRAS* fragment was amplified. Absence of all bands results in a GII 0 and means that the starting marterial might not have contained any cell. Samples with a GII>2 were considered to be of good quality.

Table 5: Master mix for WGA quality control

Reagent	Amount per reaction [µl]
10x FastStart PCR Buffer (with 20 mM $MgCl_2$)	1
Primer mix (8 µM/primer)	1
dNTPs (included in FastStart Kit)	0.2
BSA (20 mg/ml)	0.2
FastStart Taq Polymerase (5 U/µl)	0.1
PCR-water	6.5

3.9.5 Quality control PCR for WGA samples of FFPE blocks

The quality of FFPE DNA is usually worse compared to DNA isolated from unfixed tissue. The DNA is often damaged and heavily fragmented resulting in lower GII values. To judge the suitability of the FFPE DNA for downstream analysis, we performed an endpoint PCR with different sets of primers targeting shorter and medium length MseI fragments. The FFPE QC PCR

consists of three different multiplex PCRs with a total of 12 primer sets. Primer sequences and amplicon lengths are provided in 2.1.2. The master mix was prepared according to **Table 6**. 9 µl master mix was used with 1 µl of template DNA. To control for the functionality of the assay and the purity of reagents, a positive and a negative control were included. The PCR cycler program is described in 3.9.4. The products of the PCR reaction were visualized by agarose gel electrophoresis (3.10).

Table 6: FFPE WGA multiplex master mixes

Multiplex	Reagents	Volume per sample [µl]
1	10x FastStart PCR Buffer (with 20mM MgCl2)	1
	Primer mix 1 (KRAS_Ex1, PIK3CA_Ex20, EGFR_Ex19, 8 µM/ primer)	1
	Primer mix 2 (BRAF_Ex15, TP53_Ex7, 8 µM/primer)	1
	dNTPs (from FastStart kit)	0.2
	BSA (20 mg/ml)	0.2
	FastStart Taq Polymerase (5 U/µl)	0.1
	PCR-water	5.5
2	10x FastStart PCR Buffer (with 20mM MgCl2)	1
	Primer mix 3 (EGFR_Ex18, EGFR_Ex21, TP53_Ex8, NRAS_Ex2, 8 µM/primer)	1
	dNTPs (from FastStart kit)	0.2
	BSA (20 mg/ml)	0.2
	FastStart Taq Polymerase (5 U/µl)	0.1
	PCR-water	6.5
3	10x FastStart PCR Buffer (with 20mM MgCl2)	1
	Primer mix 4 (CKND2A_Ex3, EGFR_Ex20, TP53_Ex 5/6, 8 µM per primer)	1
	dNTPs (from FastStart kit)	0.2
	BSA (20 mg/ml)	0.2
	FastStart Taq Polymerase (5 U/µl)	0.1
	PCR-water	6.5

3.10 Agarose gel electrophoresis

Gels were prepared by dissolving 1.5 g agarose in 100 ml 1x TBE buffer in an Erlenmeyer falsk and heating it up in the microwave until the solution was clear. After cooling down to skin temperature, 4 µl of 10 mg/ml ethidium bromide solution were added and mixed with the liquid gel by gentle swinging. The liquid gel was transferred from the Erlenmeyer flask into the gel tray equipped with two combs for 20 pockets each. The gel was left at RT for polymerization for at least 20 min. 3 µl of gel loading dye was mixed with the samples by pipetting. After polymerization of the gel, the samples were carefully loaded into the pockets. For the determination of the size, 8 µl of 1 kb DNA ladder was loaded next to the samples. At the end, the DNA was separated at 160 V and 400 mA for 45 min and imaged using UV light.

3.11 Reamplification of *Ampli1*™ WGA samples

To conserve valuable primary WGA material, samples can be reamplified for downstream analysis. For this 49 µl master mix shown in **Table 8** were added to 1 µl of primary WGA product to reach a final volume of 50 µl. The reaction was performed in a PCR cycler using the program described in **Table 7**. The two variations of the cycler program, the elongation temperature of 65

°C or 68 °C, result in different distributions of the fragment length of the PCR products. For the lineage tree analysis we combined two both variants by splitting the reaction into two (25 µl each), running both variants of the cycler program and mixing both PCR reactions afterwards 1:1. Therefore we aim to get a broad fragment length distribution. For the *Ampli1*™ LowPass CNA analysis we only used the reamplification reaction with 68 °C. Afterwards, WGA-QC PCR (see 3.9.4) was performed to check if the reaction was successful.

Table 7: Master mix for reamplification of Ampli1™ WGA products

Reagent	Amount per reaction [µl]
Expand Long Template Buffer 1	5
Lib1 (10 µM)	5
dNTPs (10 mM)	1.75
BSA (20 mg/ml)	1.25
DNA Pol Mix	0.5
PCR-water	35.5

Table 8: Cycler program for reamplification

Step	Temperature (°C)	Duration (h:min:sec)	Cycles
1	94	00:01:00	
2	60	00:00:30	
3	65 **or** 68	00:02:00	
4	94	00:00:30	
5	60	00:00:30	10
6	65 **or** 68	00:02:00 + 20 sec/cycle	
7	4	forever	

3.12 Double-strand synthesis of *Ampli1*™ WGA samples

Double-strand synthesis was performed to fill up the single strand overhangs that arise in the WGA. During the optimization of the sample preparation for the lineage tree analysis (see 4.5.2) it didn't show any improvement of the library performance and was therefore not done for the samples finally undergoing lineage tree analysis. Components of the master mix were are provided in **Table 9**. For the double-strand synthesis, 10 µl of the primary WGA product or the reamplified WGA were mixed with 2 µl of the master mix. The samples were incubated in a PCR cycler at 68 °C for 2 h, followed by a cool down to 12 °C.

Table 9: Master mix composition for double-strand DNA synthesis

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

3.13 Purification of reamplified WGA samples with AMPure XP beads

After reamplification all samples for the lineage tree analysis were purified with AMPure XP beads according to manufacturer's instructions to remove residual PCR reagents. Briefly, 15 µl of reamplified samples (1:1 mixture of reamplification with 65 and 68°C elongation temperature) were mixed with 1.8 times the amount of AMPure XP beads. After 5 min of incubation time to allow the beads to bind the DNA, the samples were put on a magnetic rack for another 5 min. Due to the magnets, the beads were attached to the wall of the tubes which allows the supernatant to be discarded. Still on the magnetic rack, the beads were washed two times with 80% ethanol. After the ethanol was removed the second time, the beads were air-dried for about 5-10 min. The samples were removed from the magnetic rack and DNA was eluted in 13 µl of nuclease free water, vortexing and incubation for 5 min. The samples were again put on the magnetic rack for 5 min, until the beads separated, and the solution was clear. The eluted DNA was collected in a fresh Eppendorf Lobind 96-well plate and stored at -20 °C until it was shipped to the Shapiro laboratory on dry ice.

3.14 Lineage tree analysis and bioinformatics

The target enrichment of STRs by the MIP pipeline, sequencing and the data analysis were performed by the Laboratory of Prof. Dr. Ehud Shapiro at the Department of Computer Science and Biological Chemistry at the Weizmann Institute of Science, Rehovot, Israel. The pipeline was published by Tao *et al*., 2018. Following targeted enrichment with duplex MIPs, library preparation and sequencing, the resulting reads were aligned against a custom reference genome of all possible STRs variations in the panel. Reference sequences for an STR locus are shown as an example (**Figure 5**). Sequenced data with a minimal coverage of 10X reads was genotyped and a confidence threshold of 0.05 (correlation above 0.95) between the measured histogram and the reported model was set. The reconstruction based on these resulting genotypes was performed using the FastTree2 algorithm with the mutation count distance matrix.

Figure 5: Example of reference sequences for STR loci

Reference sequences for this locus are created for all possible STR lengths (3-21 repeats in this case)

3.15 ASB-PCR assay

Gp100-positive DCCs from LNs of malignant melanoma patients tested for the *BRAF* mutation were stained and isolated according to our established protocols described in sections **3.6.1** and **3.7**. Subsequently, whole genome amplification was performed as described in **3.9**.

3.15.1 Isolation of DNA from cell lines and PBMCs

Bulk DNA from SKMEL28 (SK-MEL-28, HTB-72™, ATCC®), harbouring a homozygous *BRAF* c.1799T>A mutation was first used for the establishment of the ASB-PCR. Bulk DNA was isolated using the DNeasy Blood and Tissue Kit (QIAGEN) and quantified using the Nanodrop ND-1000 spectrophotometer.

The MelHo (MEL-HO, ACC-62, DMSZ) cell line carrying a heterozygous *BRAF* c.1799T>A mutation was used for the establishment of the ASB-PCR on single-cell WGA products. Single cells were isolated with a micromanipulator (PatchMan NP2, Eppendorf) and subjected to whole genome amplification (see **3.9**).

Peripheral blood mononuclear cells (PBMCs) from a male healthy donor were isolated according to bone marrow preparation protocol as previously described (REF Braun, Pantel et al. 2000). Mononuclear cells were separated using density gradient centrifugation in 65% Percoll (GE Healthcare). The cell pellet was resuspended in 3 ml PBS. Genomic DNA was isolated using the DNeasy Blood and Tissue Kit (QIAGEN) according to manufacturer's instructions and eluted in 50 µl of H_2O. DNA was fragmented using Covaris S220X focused-ultrasonicator according to manufacturer's instructions to achieve an average DNA fragment size of approximately 175 bp. The success of the DNA fragmentation was visualized utilizing agarose gel electrophoresis and the Agilent Bioanalyzer and quantified using the Nanodrop ND-1000 spectrophotometer.

3.15.2 Oligonucleotide design

NCBI sequence NG_007873.3 was used as reference sequence for *BRAF* and Human genome assembly GRCh38 was considered as reference throughout the study. The 5'-mutant-specific primer was designed with a complementary base to the hotspot *BRAF* mutation c.1799T>A at its 3'-end. The blocking oligonucleotide was fully complementary to the wild type sequence of the *BRAF* gene with the base complementary to the mutation site located in the middle of the oligonucleotide. In addition, the blocking oligonucleotide contained a modified base at its 3'-end, containing a 3'-phosphate group to prevent elongation during PCR. All oligonucleotide sequences can be found in **2.1.2** and used in combination in different PCR assays as outlined in **Table 10**.

Table 10: Oligonucleotide combinations used in the ASB-PCR assay

Assay	Primer combination	Fragment length
Singleplex ASB-PCR	5'-mutation-specific + 3'-allele-independent primer blocking oligonucleotide	147 bp
Control PCR	5'allele-independent + 3'allele-independent primer	171 bp
Multiplex ASB-PCR	5'-mutation-specific + 3'-allele-independent primer 5'-allele-independent + 3'-allele-independent primer blocking oligonucleotide	147 bp 171 bp

3.15.3 Singleplex ASB-PCR

The establishment of the ASB-PCR on fragmented DNA and on single-cell WGA products was first performed as singleplex ASB-PCR. The reaction was performed by adding 1 µl of gDNA or

WGA product to 9 µl master mix shown in **Table 11**. In order to increase specificity and selectivity in mutational analyses, a touch-down PCR was performed. The initial denaturation was 4 min at 95°C, followed by 10 cycles of 30 sec at 95°C, 30 sec at 68°C decreasing by 1°C per cycle down to 58°C and 90 sec at 72°C. Reaching the annealing temperature of 58°C, another 28 cycles followed before the final elongation of 7 min at 72°C were added.

Table 11: Master mix of the singleplex ASB-PCR reaction

Reagent	Amount per reaction [µl]
10x FastStart PCR Buffer (with 20 mM $MgCl_2$)	1
dNTPs (included in FastStart Kit)	0.2
BSA (20 mg/ml)	0.2
5'-mutant specific primer (8 µM)	0.5
3'-allele-independent primer (8 µM)	0.5
blocking oligonucleotide (32 µM)	0.5
FastStart Taq Polymerase (5 U/µl)	0.1
PCR-water	6.0

3.15.4 Control PCR

To confirm the presence of the *BRAF* gene sequence of interest, the control PCR was performed. The reaction was performed by adding 1 µl SKMEL28 bulk DNA or MelHo single-cell WGA products to 9 µl master mix shown in **Table 12**. The samples were amplified using the touch-down PCR program as described in **3.15.3**.

Table 12: Master mix of the control PCR of the ASB-PCR reaction

Reagent	Amount per reaction [µl]
10x FastStart PCR Buffer (with 20 mM $MgCl_2$)	1
dNTPs (included in FastStart Kit)	0.2
BSA (20 mg/ml)	0.2
5'-allele-independent primer (8 µM)	0.5
3'-allele-independent primer (8 µM)	0.5
blocking oligonucleotide (32 µM)	0.5
FastStart Taq Polymerase (5 U/µl)	0.1
PCR-water	6.0

3.15.5 Multiplex ASB-PCR

The multiplex ASB-PCR was first established using 1 µl SKMEL28 fragmented DNA and later single-cell WGA products as template in a total volume of 10 µl, the master mix is described in **Table 13**. In contrast to the singleplex reactions, the volume of the 3'-allele-independent primer was doubled for the multiplex reaction since the primer amplifies the template with both the 5'-mutant specific and the 5'-allele-independent primer. The samples were amplified using the touch-down PCR program as described in **3.15.3**.

Table 13: Master mix of the multiplex ASB-PCR reaction

Reagent	Amount per reaction [µl]
10x FastStart PCR Buffer (with 20 mM $MgCl_2$)	1
dNTPs (included in FastStart Kit)	0.2
BSA (20 mg/ml)	0.2
5'-mutant specific primer (8µM)	0.5
5'-allele-independent primer (8 µM)	0.5
3'-allele-independent primer (8 µM)	1.0
blocking oligonucleotide (32 µM)	0.5
FastStart Taq Polymerase (5 U/µl)	0.1
PCR-water	5.0

3.15.6 Sanger sequencing

For the retrospective testing of isolated DCCs by ASB-PCR, the *BRAF* mutational status of the DCCs was evaluated by Sanger Sequencing before (Sequiserve, Vaterstetten, Germany). Sanger Sequencing was performed after gene-specific amplification from WGA samples. For the *BRAF* exon 15 amplification we used the 5'-allele-independent and 3'-allele-independent primers designed for the ASB-PCR, encompassing the hotspot mutations of codon 600 (c.1799T>A, c.1798_1799GT>AG). Cycling temperatures were set to 94°C for 2 min, 60°C for 30 sec and 72°C for 2 min for one clycle, followed by 94°C for 15 sec, 60°C for 30 sec and 72° for 20 sec for 14 cycles. Then, 94°C for 15 sec, 60°C for 30 sec and 72° for 30 sec for 24 cycles and an additional final extension step at 72°C for 2 min.

4 Results

The first aim of the thesis was to test our hypothesis, if the phenotypical small MCSP-positive DCCs are precursors of large MCSP-positive DCCs with the help of a cell lineage tree reconstruction. In the second part of this thesis, I took a closer look at a common mutation in the *BRAF* gene to evaluate the incidence and its association with the DCCD in LNs.

4.1 Selection of patients and single cells for the study of the genotype and phenotype of early DCCs

In the following sections, an overview of the patient selection criteria and their respective cell collective is described.

4.1.1 Selection of patients

For the first part, the lineage tree analysis, we selected a patient with phenotypically different DCCs. Between 2008 and 2016, we received sentinel and non-SLN from 521 malignant melanoma patients. The lymph nodes (LNs) from these patients were disaggregated into a single-cell suspension and stained against the melanoma marker MCSP. We identified two distinct phenotypes of MCSP-positive DCCs, small and large DCCs. The first selection criteria for our lineage tree collective was having patients that showed both phenotypically small and large MCSP-positive DCCs in one LN. We found 51 patients with both small and large MCSP-positive DCCs. Another criterion was an existing DCC-derived cell line of the patient to perform drug-screening assays in the future, which exceeded the scope of this work. This criterion reduced the collective to 6 patients. Since the overall aim of this thesis was to study early dissemination in malignant melanoma, we decided to choose patients with a DCC-density (DCCD; defined as the number of DCCs per million cells) below 100. This threshold was defined by a recent publication in our laboratory that showed that colony formation in the LN starts in most patients with a DCCD of around 100 (Werner-Klein et al. 2018). Including all of these criteria, two patients were left, MM15-127 and MM16-394 (**Figure 6 A**). In addition, we received metastatic tissue from patient MM15-127 and therefore selected this patient for the lineage tree analysis.

For the second application, the *BRAF* mutation analysis, we aimed to have a closer look at a specific patient with DCCs isolated from different anatomic sites. Therefore, we compared DCCs from the sentinel LN and non-sentinel LNs during different times of disease progression with CTCs from blood and tumour cells from metastases. From 2015 to 2018, we received non-SLNs from 23 patients that underwent a LAD. From 15 of those patients, the SLN was previously extirpated. We found DCCs both in the SLN and the non-SLNs of seven patients. From two of these patients, we also received metastatic tissue after disease progression (MM-412 and MM-423). In these patients, we aimed to study the mutational progression by having a closer look at the melanoma-specific *BRAF* hotspot mutations (c.1799_T>A, c.1798_1799GT>AA). Patient MM16-423 was tested positive for the *BRAF* mutation and was therefore included in further testing (**Figure 6 B**).

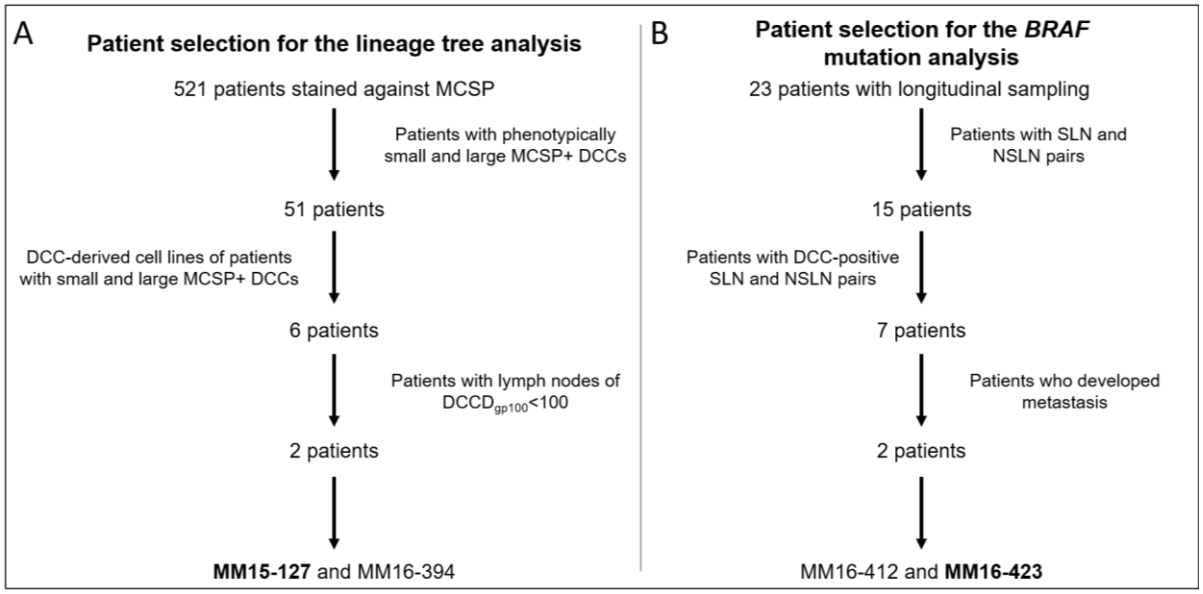


Figure 6: Patient selection for the study of progression based on mutational analysis

(A) Morphological progression was studied in patients with small and large DCCs in the SLN. Therefore, LNs of 521 patients were stained against MCSP of which 51 patients showed both phenotypically small and large MCSP-positive DCCs. From 6 of those patients, a DCC-derived cell line was generated and 2 of those patients had a DCCD<100. Patient MM-127 was chosen since this patient´s disease further progressed and metastatic tissue was available. **(B)** Local progression was studied in patients where we found DCCs in both the SLN and non-SLNs. We collected non-SLNs from 23 patients of which 15 patients also had their SLN extirpated previously. In 7 patients we found DCCs in both the sentinel and the non-SLNs. We received metastatic tissue from 2 patients with disease progression. In the end, we chose patient MM-423 as this patient tested positive for the *BRAF* mutation.

4.1.2 Baseline patient characteristics and sample collection of patients MM15-127 and MM16-423

Patient MM15-127, selected for the cell lineage tree analysis, was female and 79 years old at the time point of malignant melanoma diagnosis in 2015. The primary tumour (PT) was located at the lower leg with a tumour thickness of 6 mm and ulceration, corresponding to classification 4b. The tumour was removed completely (R0). After PT removal, one half of the LN was analysed by routine histopathology and microscopic metastases (N1a) were found. The patient had no distant metastases at the time of diagnosis in 2015. Shortly after the SLN surgery, a LAD was carried out. By the end of 2015, the patient had two times LN recurrence. Consequently, the patient received chemotherapy with Bleomycin and immunotherapy with Pembrolizumab and T-Vec. By the end of 2016, the patient developed skin metastases. In 2017 the patient died (**Figure *7* A**).

We received several tissue samples from patient MM15-127. After examination through the pathologist, we had access to the PT in form of a formalin-fixed paraffin-embedded (FFPE) tissue block to isolate genomic DNA. Additionally, we prepared sections from the FFPE block for immunohistochemistry. With the consent of the patient, we received one half of the SLN, which was disaggregated into a cell suspension by our laboratory. After immunocytochemistry, we successfully isolated DCCs for further analysis and generated a DCC-derived cell line. Furthermore, residual disaggregated LN tissue was frozen for future use. Unfortunately, we

did not receive any LN tissue from the LAD. The skin metastases the patient developed in 2016 after treatment were disaggregated and single tumour cells were isolated **(Figure *7* B)**.

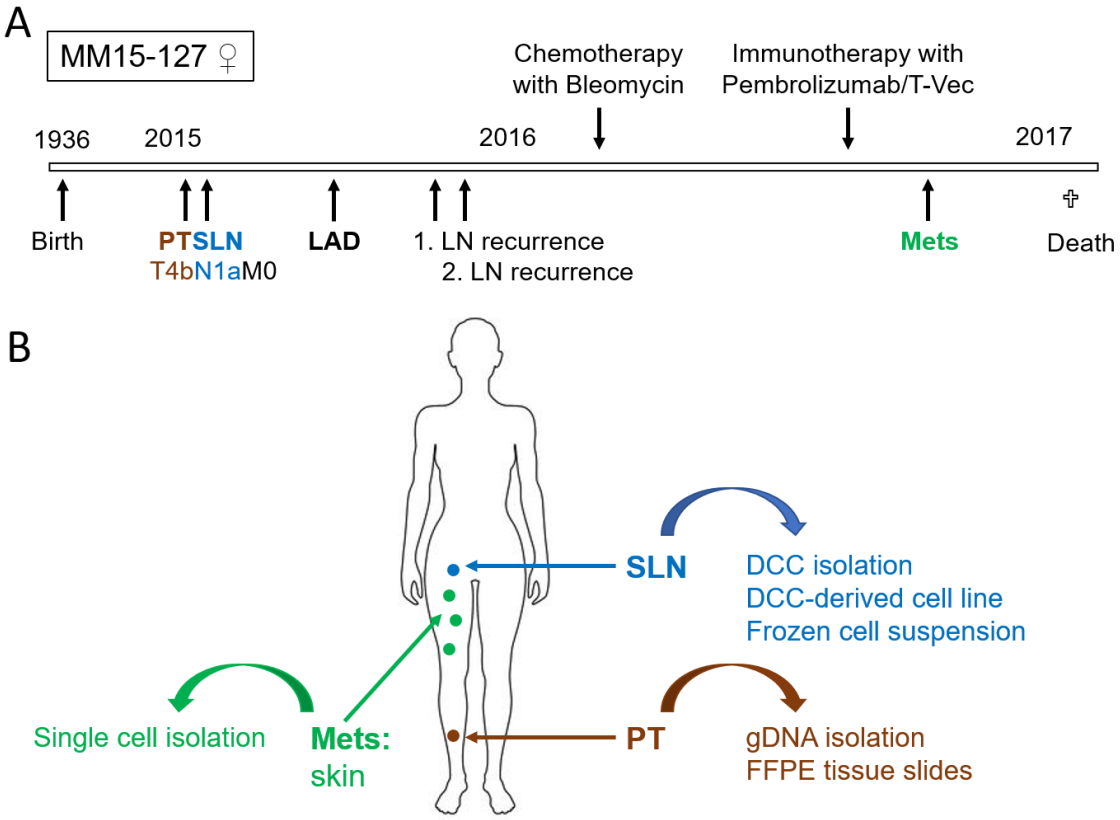


Figure 7: Course of disease and samples collection of patient MM15-127

(A) The course of disease of patient MM15-127 starts 2015 with the diagnosis of malignant melanoma. The PT and the SLN were removed followed by a lymphadenectomy (LAD). Due to LN recurrence, the patient received different treatments. After developing skin metastasis in 2016, the patient died 2017. **(B)** We received the FFPE tissue block from the patient for gDNA isolation and tissue sections. The SLN was disaggregated for single-cell isolation and the generation of a DCC-derived cell line. The residual suspension was frozen for further experiments. From the metastases, we isolated single tumour cells.

Patient MM16-423, who was selected for the *BRAF* mutation analysis, was a male patient of age 61 at the time of malignant melanoma diagnosis in 2013. He was diagnosed with malignant melanoma at the heel. The PT of stage 4b had a thickness of 4.4 mm, displayed ulceration and was completely removed (R0). The pathological examination showed a micro-metastases in the sentinel lymph node (SLN; N1a), but no distant metastases (M0) at the time point of diagnosis. In 2013, the patient received adjuvant immunotherapy with interferon. In 2016, LN recurrence was diagnosed and a LAD was performed. In 2017 we received a blood donation from the patient after leukapheresis for the isolation of CTCs. Afterwards, the patient developed multiple metastases and died (**Figure 8 A**).

After the examination through the pathologist, we received the PT tissue in form of an FFPE tissue block for genomic DNA isolation. With the consent of the patient, the SLN was divided between the pathologist and our laboratory. After disaggregation, we isolated DCCs from the SLN. In 2016, the patient had a LN recurrence and underwent a LAD. We received the halves of six non-SLN that were disaggregated for DCC isolation. In addition, we generated a DCC-

derived cell line from DCCs isolated from one of the non-SLN. In 2017, the patient donated leukapheresis blood to our laboratory that we used for the isolation of CTCs **(Figure 8 B)**.

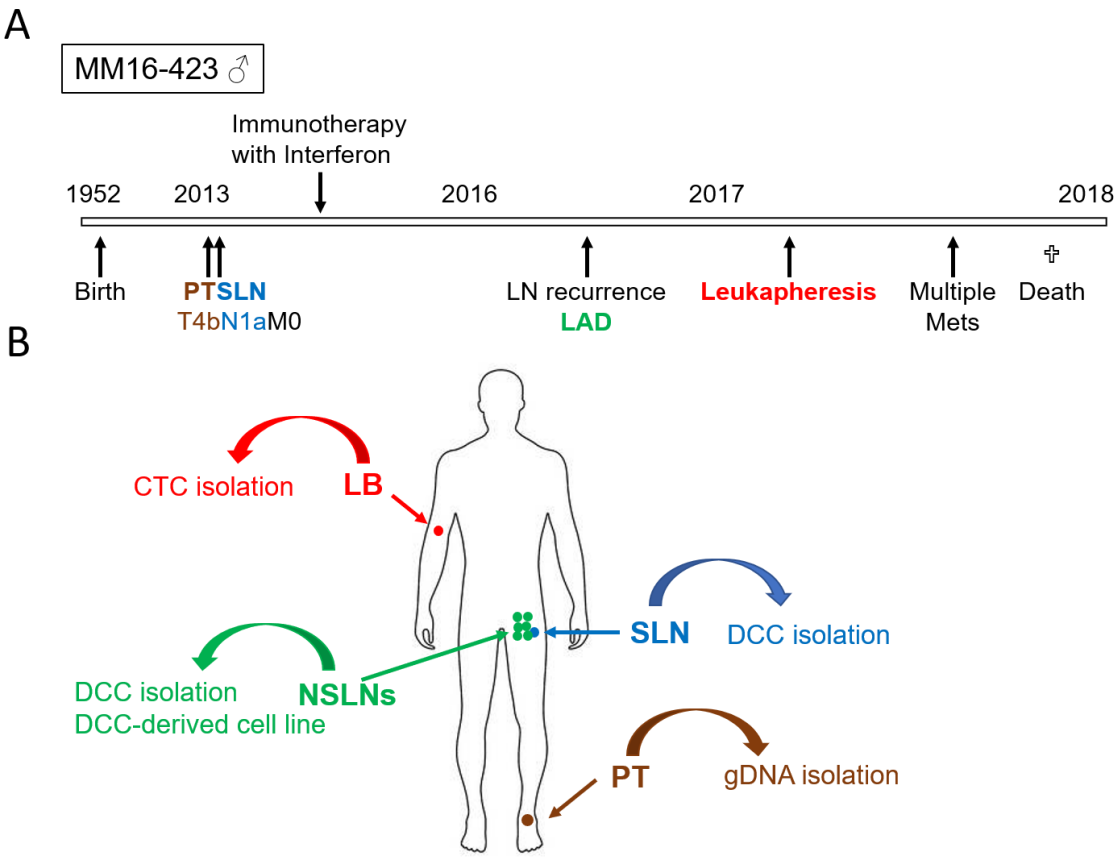


Figure 8: Course of disease and sample collection of patient MM16-423

(A) The course of disease of patient MM16-423 started in 2013 with the diagnosis of malignant melano. After resection of the PT and the SLN, the patient received an immunotherapy. 2016, the patient had a LN recurrence leading to a lymphadenectomy (LAD). For the isolation of CTCs, the patient underwent leukapheresis in 2017. After developing multiple metastases (mets), the patient died. **(B)** We received the PT in form of a FFPE tissue block for genomic DNA isolation. Half of the SLN was disaggregated for DCC isolation. We also received the non-SLNs after LAD for DCC isolation and the generation of a DCC-derived cell line. The leukapheresis blood was used for CTC isolation.

4.2 Detection and isolation of tumour cells for lineage tree analysis of patient MM15-127

For the study of the phenotypical progression in malignant melanoma we aimed to investigate our hypothesis that small MCSP-positive DCCs are the precursors of large MCSP-positive DCCs and that large MCSP-positive DCCs might then have the potential to form metastases. After disaggregation of the LN into a single cell suspension, the cells were counted and divided. Three million cells were stained right after the LN preparation (**Table** 14, “MCSP+ fresh”), the residual cells were frozen for later use (**Table** 14, “MCSP+ frozen”). Thereby, we isolated two phenotypes of MCSP-positive cells: On the one hand small MCSP-positive cells with a diameter below 10 µm (**Figure 9 A**) and on the other hand large MCSP-positive cells with a diameter above 15 µm (**Figure 9 B**). Besides MCSP, we also used another tumour marker for DCC isolation from malignant melanoma patients, namely HMB45/gp100. The

gp100 staining was performed on adhesion slides prepared from the LNs single cell suspension of patient MM15-127 (**Figure 9 C**). By using the marker gp100, we did not observe phenotypically different tumour cells, but only large cells with a diameter of ~20µm. We included those cells in the lineage tree analysis to study their role in the metastatic progression of the patient. The metastatic tissue of the patient was digested into a single cell suspension, transferred onto adhesion slides and stained against HMB45/gp100 (**Figure 9 D**). Single cells from the DCC-derived cell line from patient MM15-127 were isolated without any prior staining. Their malignant origin was later confirmed by copy number variation (CNV) analysis (see **4.5.10**).

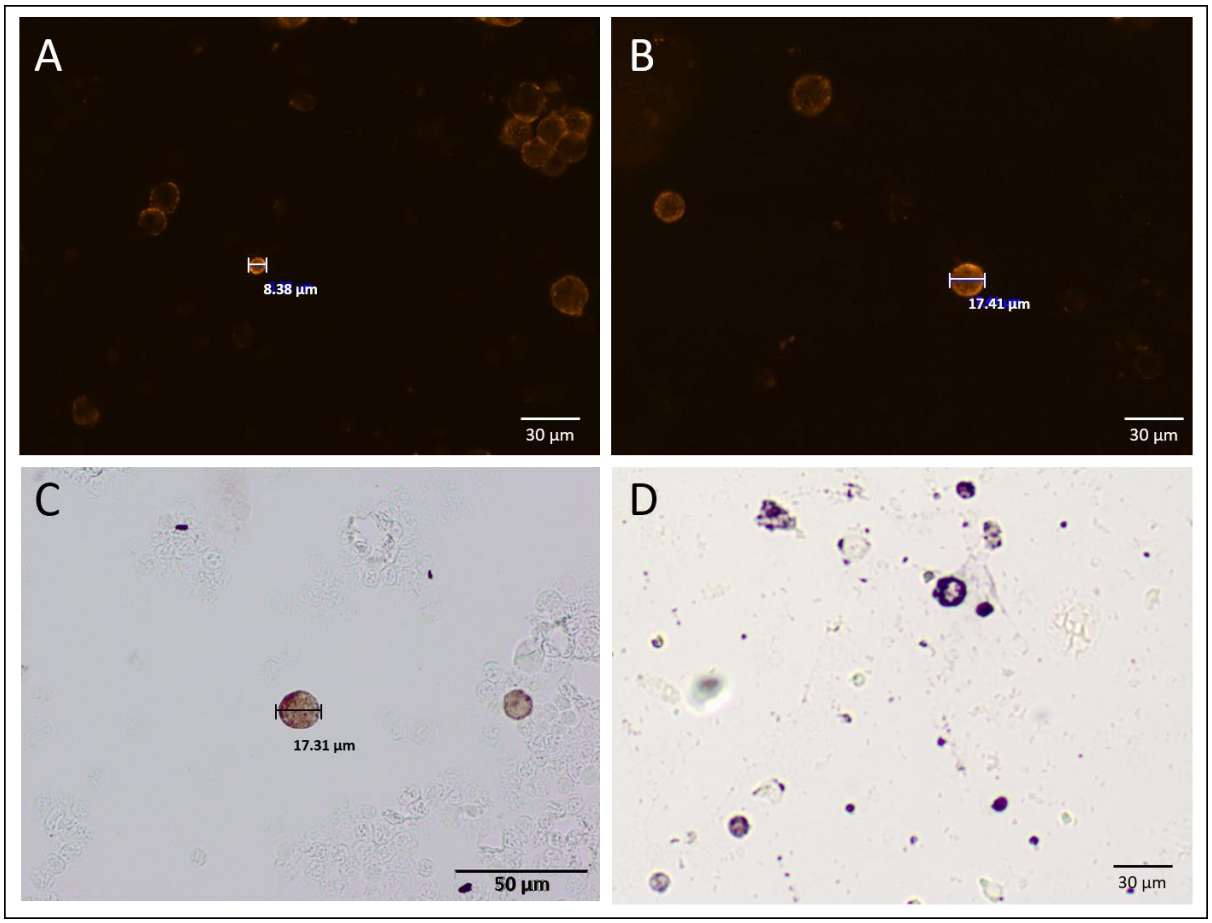


Figure 9: Pictures of tumour cells isolated for lineage tree analysis from patient MM15-127

Staining of DCCs against MCSP was performed with single cell suspensions. Phenotypically **(A)** small and **(B)** large MCSP-positive cells were isolated. **(C)** DCCs were in addition stained against gp100 on adhesion slides. **(D)** The metastatic tissue was digested, transferred onto adhesion slides and stained against gp100.

From the fresh LN suspension, we isolated two small MCSP-positive cells by micromanipulation. Both cells had a good DNA quality (GII>2) after WGA and were selected for lineage tree analysis. From the same cell suspension, five large MCSP-positive cells were isolated. All five cells showed a good DNA quality after WGA and were selected for lineage tree analysis. To enlarge the cell collection of MCSP-positive cells we thawed a frozen LN suspension from patient MM15-127, stained it against MCSP and isolated additional MCSP-positive cells. From the 32 small MCSP-positive cells that were isolated, 28 showed good DNA

quality and were selected for lineage tree analysis. 24 large MCSP-positive cells were isolated of which 20 were of good DNA quality. With the second melanoma marker, gp100, 20 cells were isolated. Of these, 18 cells had a good DNA quality and were selected for lineage tree analysis. From the DCC-derived cell line, 24 of 26 isolated cells showed a good DNA quality. In addition, we received three metastases. From metastasis 1 and metastasis 3, 42 and 44 cells were isolated with the melanoma marker gp100, whereof 18 and 24 cells had a good DNA quality and were selected for lineage tree analysis (**Table 14**).

Table 14: Selection of tumour cells for lineage tree analysis

Number of cells	Small MCSP+ cells (fresh/frozen)	Large MCSP+ cells (fresh/frozen)	gp100+ cells	Single-cells from DCC-derived cell line	Tumour cells from metastasis 1+3
Isolation by micromanipulation	2 / 32	5 / 24	20	26	42 + 44
with GII>2	2 / 28	5 / 20	18	24	18 + 24
Selected for LT analysis	2 / 28	5 / 20	18	24	18 + 24

4.3 Isolation of gDNA from FFPE tissue of the PT of patient MM15-127

In addition to non-tumour control cells and single tumour cells for the lineage tree analysis of patient MM15-127, we aimed to include cells from the PT. Unfortunately, the PT was only available as FFPE tissue. The quality of the FFPE samples is usually not optimal and therefore the quality of isolated single cells is often poor (Ghosh 2020). That is why we decided to isolate bulk DNA from the PT using the FFPE DNA isolation kit from Qiagen (see 3.4). We took three punches from the tumour area of the FFPE block, that was previously marked by Dr. Florian Weber our pathologist. After DNA isolation, we measured the DNA concentration and prepared eight replicates with a concentration of around 50 ng/µl of each punch. Four replicates of each punch were amplified by the standard WGA protocol, using 1 µl of isolated gDNA. The other four replicates were amplified by WGA including a treatment with the FFPE repair mix (see 3.9.3). The aim of the different approaches was to compare the performance of the samples in the lineage tree analysis. At the end, we included 4 replicates with and 4 replicates without DNA repair of the three punches, resulting in a total of 24 samples.

4.3.1 Detection and isolation of tumour cells for the *BRAF* mutation analysis of patient MM16-423

To investigate the sequence of mutational events happening in the *BRAF* gene of a specific patient, tumour cells isolated from different tissue entities of patient MM16-423 were tested with the previously established ASB-PCR. We isolated DCCs from the SLN (**Figure 10 A**) and from 6 non-SLNs after LAD (**Figure 10 B**) during the patient's progress of disease. In addition, we isolated CTCs from a leukapheresis (**Figure 10 C**) and single cells from the DCC-derived cell line (**Figure 10 D**) of the patient.

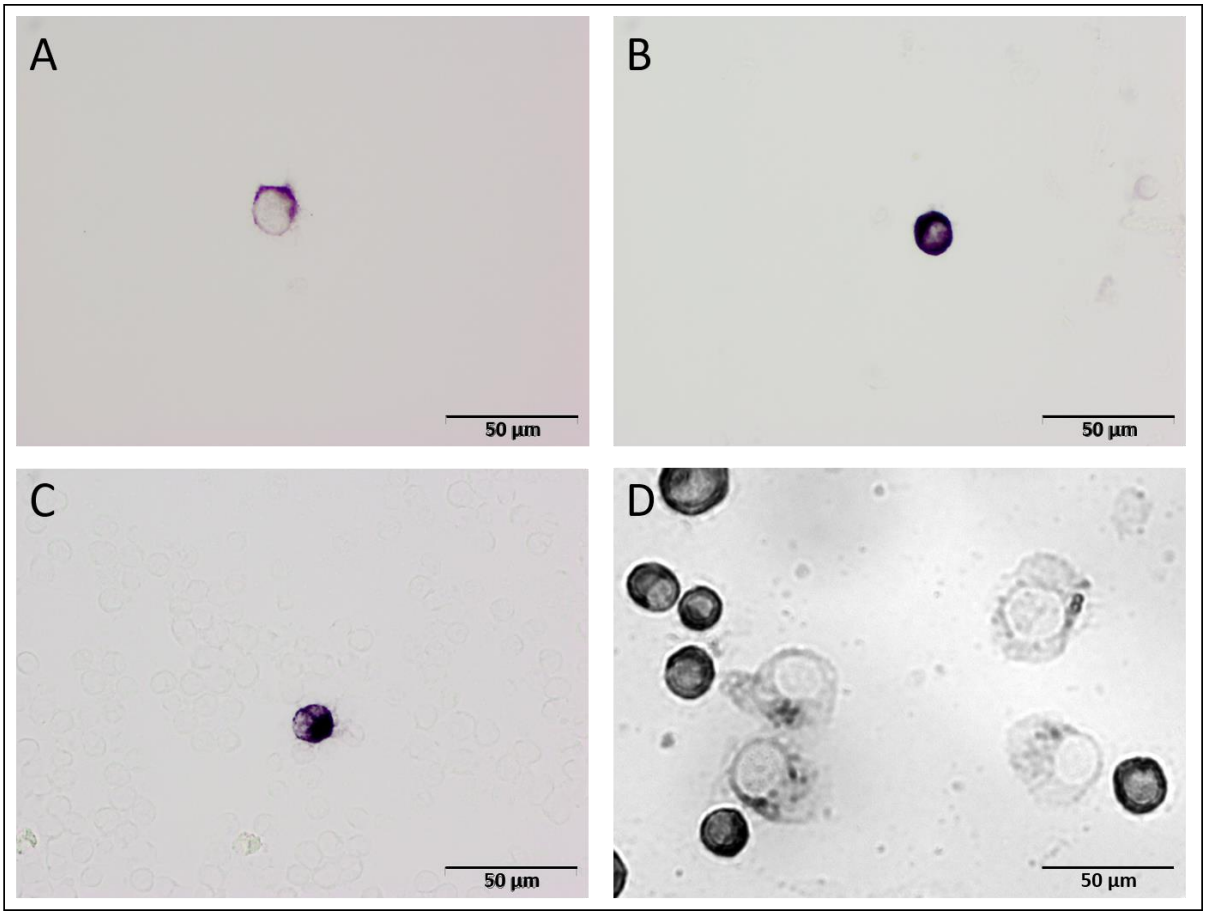


Figure 10: Tumour cells isolated from patient MM16-423 for *BRAF* mutation analysis

The tumour cells were stained against gp100 on adhesion slides. DCCs were isolated from **(A)** the sentinel lymph node and **(B)** non-sentinel lymph nodes. **(C)** Circulating tumour cells were isolated from blood. **(D)** The DCC-derived cell line was also stained against gp100 and positive cells were isolated.

The tumour cells were all stained on adhesion slides and against the tumour marker gp100+. Gp100-positive tumour cells were isolated and tested for their genomic quality reflected in a genome integrity index (GII) defined by multiplex PCR (see **3.9.4**). Cells with a GII>2 were assumed to be of good quality and selected for the analysis. Furthermore, the cells were subjected to low pass sequencing to check for copy number variations (CNVs). A reliable tumour origin can only be assumed of cells with an aberrant CNV profile. For that reason, only tumour cells with an aberrant CNV profile were included in this analysis. The CNV analysis of the tumour cells of patient MM16-423 was performed by Sandra Huber as part of her doctorate and will therefore not be discussed in detail here. From the SLN, four cells were isolated of which 3 had a good DNA quality and an aberrant CNV profile. From the six non-SLN, a total of 112 cells were isolated. 59 cells were of good DNA quality and showed an aberrant CNV profile. All DCCs from the SLN and the six non-SLN that passed the quality control were analysed by the ASB-PCR for *BRAF* mutation. 15 of the 19 CTCs that were isolated from blood had a good DNA quality and an aberrant CNV profile. All of those CTCs were tested by the ASB-PCR. From the DCC-derived cell line, 35 cells were isolated and 33 cells had a good quality and an aberrant CNV profile. A subset of 18 cells was selected for ASB-PCR **(Table 15)**.

Table 15: Selection of tumour cells for *BRAF* mutation analysis

Number of cells	Isolation by micromanipulation	with GII>2 and aberrant CNV profile	Selected for *BRAF* analysis
cells isolated from the SLN	4	3	3
cells isolated from the NSLN 1	20	14	14
cells isolated from the NSLN 2	20	11	11
cells isolated from the NSLN 3	13	1	1
cells isolated from the NSLN 4	20	17	17
cells isolated from the NSLN 5	19	3	3
cells isolated from the NSLN 6	20	13	13
CTCs	19	15	15
cell line cells	35	33	18

4.4 Detection and isolation of non-tumour control cells for lineage tree analysis of patient MM15-127

Besides the DCCs, non-tumour control cells had to be included in the lineage tree analysis. The non-tumour control cells were needed for the calculation of the root that anchors the beginning of the tree. Besides, we aimed to test if the non-tumour control cells can be clearly separated from the malignant disseminated cancer cells. We decided to use macrophages and T-cells from the haematopoietic cell lineage and endothelial cells from the endothelium to have control cells from a second cell lineage.

4.4.1 Double-staining against CD3 and CD68 for T-cells and macrophages

The staining of T-cells and macrophages was established as a double-staining to save valuable patient material. First, both stainings were performed individually and combined later. The double-staining was established on lymphatic tissue. As LN tissue from malignant melanoma patients is used for DCC isolation, archived slides could be used for staining establishment. Lineage tree analysis was performed also with non-small cell lung cancer patients for another project and therefore the staining was established for those two tissue types.

4.4.1.1 Establishment of the CD3-staining against T-cells

For the isolation of T-cells from LN suspension on tissue slides the rabbit anti-CD3 antibody was used. CD3 (cluster of differentiation 3) is a cell-surface protein expressed on T-cells and required for T-cell activation. The polyclonal antibody was tested in two different

concentrations of 0,2 mg/ml and 0,1 mg/ml of total protein on a LN slide of a DCC-negative patient. Secondary staining was performed with an AP-polymer anti-rabbit solution. Cells of the LN showed a violet staining of the cell surface that was equal for both concentrations. The concentration of 0,1 mg/ml of the rabbit anti-CD3 antibody was sufficient for a clear staining of T-cells. As isotype control rabbit IgG was used in the higher concentration of 0,2 mg/ml. The isotype control was negative (**Figure 11 A+B**).

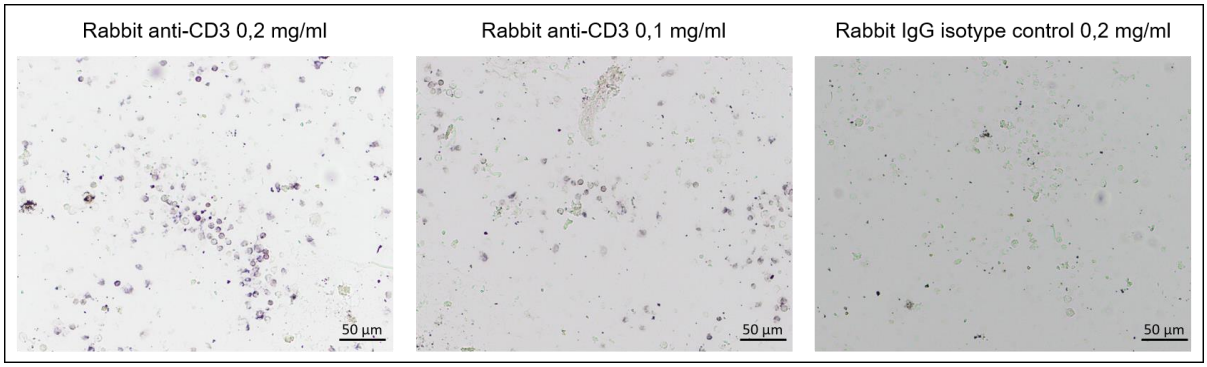


Figure 11: Establishment of CD3-staining against T-cells on LN slides

Anti-CD3 staining was established on **(A)** a LN slide of a DCC-negative patient. Two concentrations, 0,2 mg/ml and 0,1 mg/ml of the primary antibody were tested. In both tissues, cells showed a violet staining of the cells surface for both concentrations. The isotype control was negative.

For a double-staining of T-cells and macrophages on the same slide, secondary antibodies conjugated to two different fluorochromes had to be used to distinguish the different cell types. Therefore, we tested immunofluorescent staining for the CD3 antibody with a goat anti-rabbit Alexa Fluor 555-conjugated secondary antibody on a LN slide of a DCC-negative patient. The primary antibody was used in a concentration of 0,1 mg/ml, the secondary antibody in 8 µg/ml (Figure 12 A) and 4 µg/ml (Figure 12 B) on LN slides. Both concentrations showed a specific staining of cells in the LN. Although 8 µg/ml showed a slightly stronger staining, a concentration of 4 µg/ml was sufficient for a clear staining to identify T-cells on adhesion slides and used throughout all further stainings.

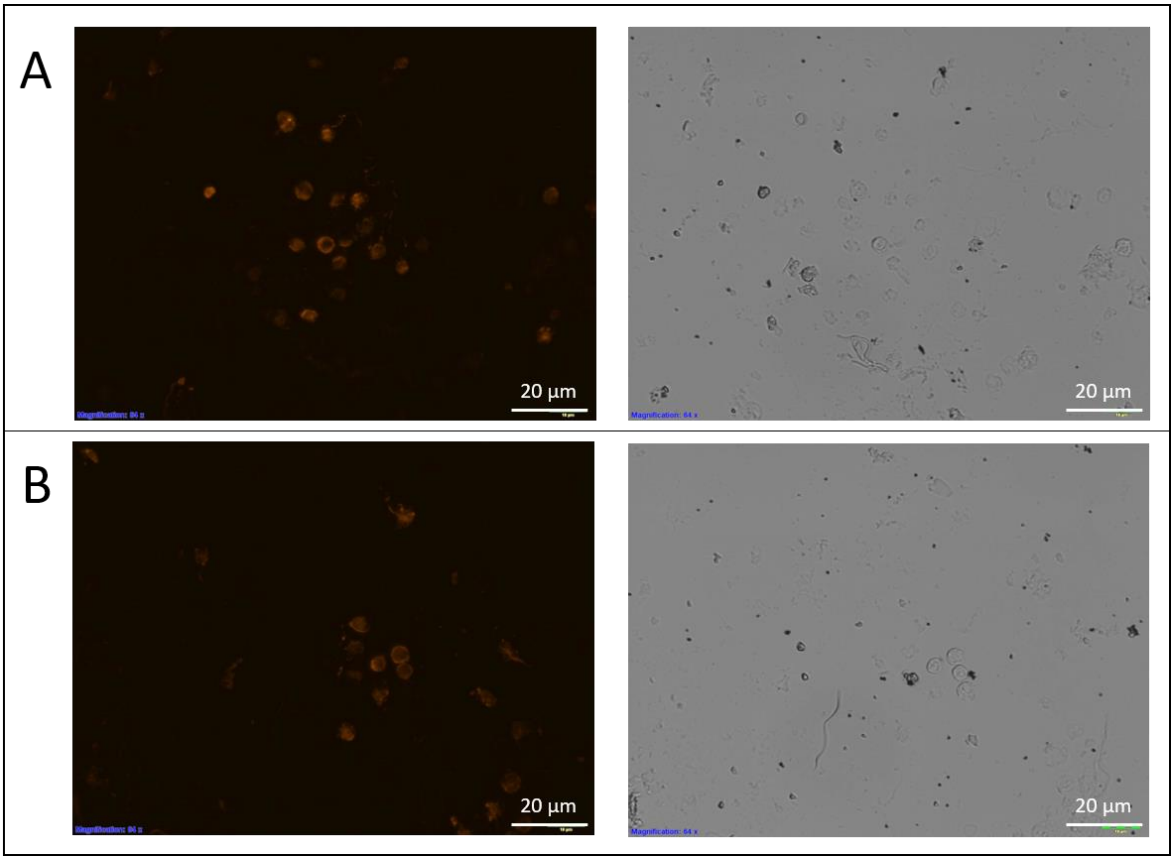


Figure 12: Combination of the rabbit anti-CD3 antibody and an Alexa Fluor 555-conjugated anti-rabbit secondary antibody

The rabbit anti-CD3 primary antibody was used in a concentration of 0,1 mg/ml and was tested with two different concentrations of the Alexa Fluor 555-conjugated anti-rabbit secondary antibody. Both concentrations of **(A)** 8 µg/ml and **(B)** 4 µg/ml showed a strong and specific immunofluorescent staining of the cells (left side), although the 8 µg/ml showed a slightly stronger staining. Bright-field images of slides are displayed on the right.

4.4.1.2 Cross-reaction test of rabbit anti-CD3 and mouse anti-CD68 antibodies

The CD68-staining against macrophages had already been established by Manjusha S. Ghosh (Ghosh 2020). CD68 (cluster of differentiation 68) is a transmembrane glycoprotein expressed on cells of the macrophage lineage and required for recruitment and activation of macrophages and the promotion of phagocytosis. To distinguish the macrophage staining from the T-cell staining, a secondary antibody conjugated to a fluorochrome of another emission spectrum had to be used. Therefore, a 1:500 dilution of the mouse anti-CD68 primary antibody was combined with 4 µg/ml goat anti-mouse Alexa Fluor 488-conjugated antibody. As isotype control the mouse IgG1κ MOPC21 clone was used in a 1:500 dilution, which was negative on all tested slides (see results section of thesis Ghosh 2020).

To exclude the possibility of a cross-reaction between the antibodies to be combined, the rabbit anti-CD3 antibody was tested with the Alexa Fluor 488 anti-mouse antibody while the mouse anti-CD68 was tested with the goat anti-rabbit Alexa Fluor 555 antibody. The concentrations and dilutions of the different antibodies were used as previously established. Neither the combination of the rabbit anti-CD3 antibody with the Alexa Fluor 488 anti-mouse antibody, nor the combination of the mouse anti-CD68 antibody with the goat anti-rabbit Alexa

Fluor 555 antibody resulted in a positive staining (**Figure 13**). This result demonstrated the specificity of both secondary antibodies and showed that both antibodies can be combined.

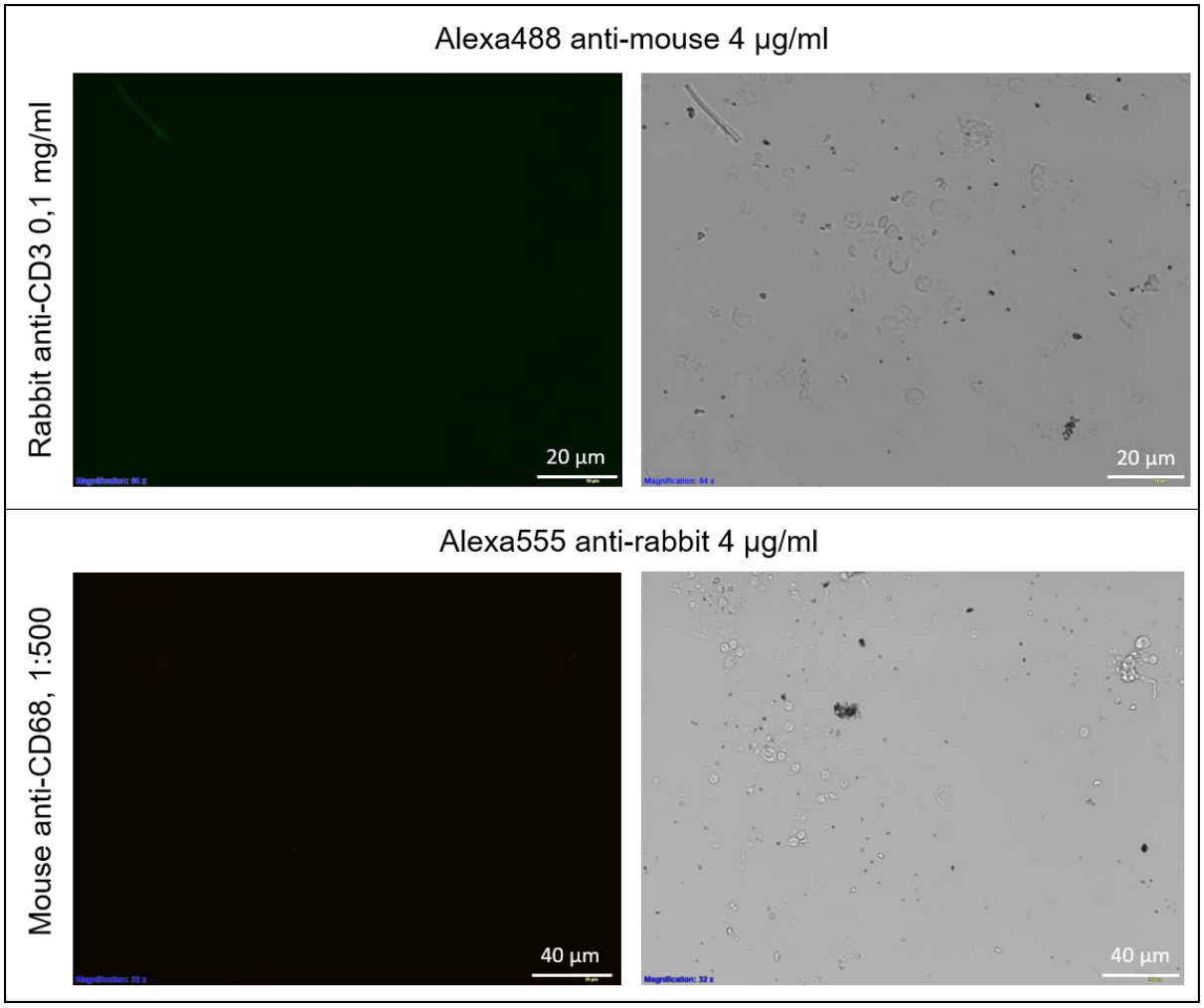


Figure 13: Cross-reaction test of rabbit anti-CD3 and mouse anti-CD68 antibodies with Alexa Fuor 488 anti-mouse and Alexa 555 anti-rabbit antibodies

By combining the rabbit anti-CD3 antibody with a final concentration of 0,1 mg/ml with 1:500 diluted Alexa Fluor 488 anti-mouse antibody no staining was detected (left side). The combination of the mouse anti-CD68 in a dilution of 1:500 and the Alexa 555 Fluor anti-rabbit in a 1:500 dilution also resulted in no positive staining (left side). Bright-field images of slides are displayed on the right.

4.4.1.3 Testing of the double-staining against T-cells and macrophages on LN slides

Finally, the double staining was tested on a LN slide from a DCC-negative patient. The primary antibodies rabbit anti-CD3 and mouse anti-CD68 were used in a final concentration of 0,1 mg/ml and a dilution of 1:500, respectively. For the detection of the cells secondary antibodies conjugated to different fluorochromes were used in a final concentration of 4 µg/ml each. The goat anti-rabbit Alexa Fluor 555 antibody detecting the T-cells emits an orange light visible in the Cy3 channel (**Figure 14, B**). The antibody against CD3 stains the cell surface of the T-cells that could now be isolated and used as non-tumour control cells for the lineage tree analysis. (**Figure 14**, arrow). For the detection of macrophages the secondary antibody Alexa

Fluor 488 anti-mouse emitting a green light visible in the FITC channel was used (**Figure 14**, **C**). The antibody stains the cell surface of macrophages (**Figure 14**, arrowhead) enabling a separate isolation of T-cells and macrophages from one slide followed by further amplification with the Ampli1 WGA protocol and inclusion as non-tumour control cells for the lineage tree analysis.

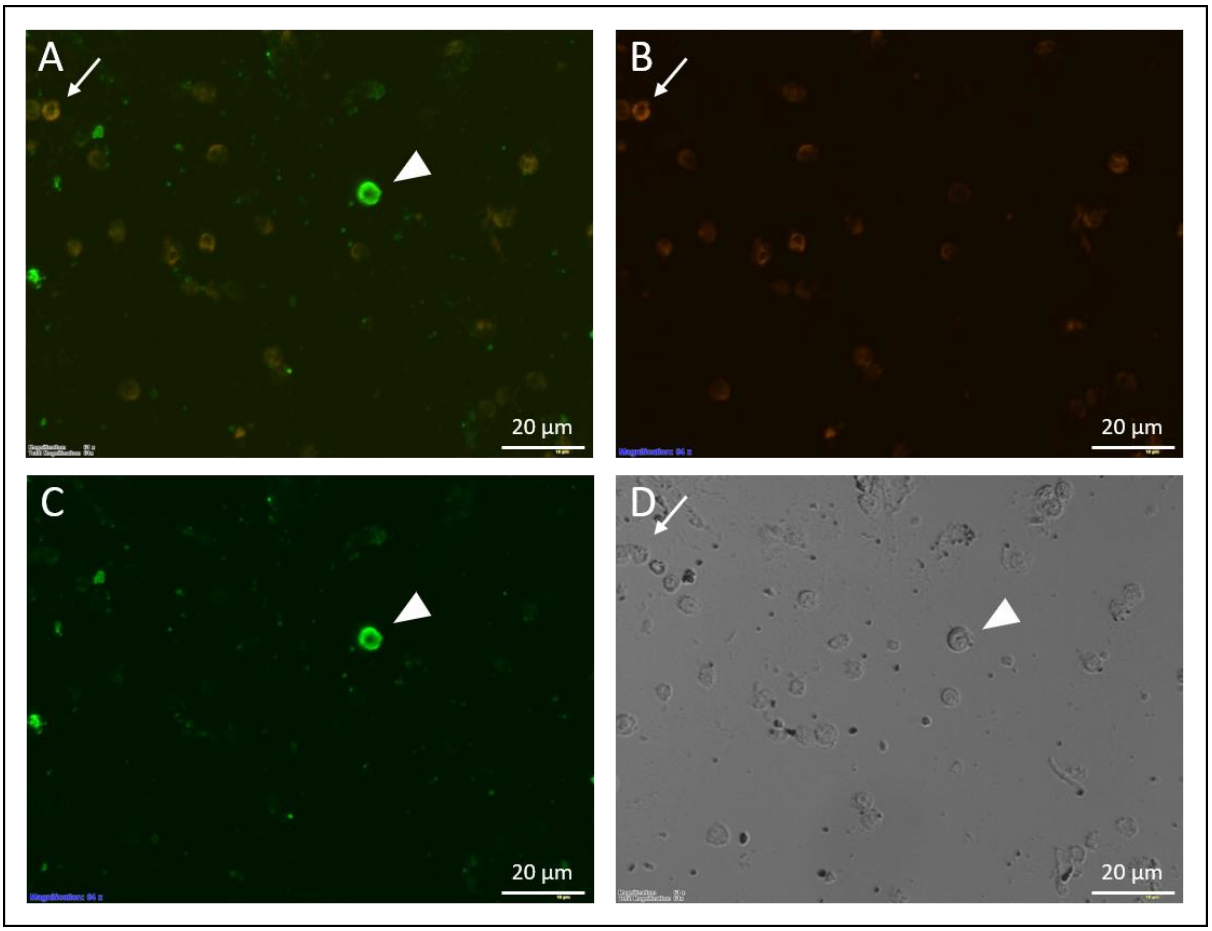


Figure 14: Double-staining against T-cells and macrophages on a LN adhesion slide

The cell surface of the T-cells was stained with a rabbit anti-CD3 antibody and detected with the Alexa Fluor 555 anti-rabbit secondary antibody emitting an orange colour (arrow). The macrophages were detected with a mouse anti-CD68 antibody and the Alexa Fluor 488 anti-mouse antibody emitting a green colour (arrowhead). **(A)** Cy3 and FITC channel overlap, **(B)** Cy3 channel, **(C)** FITC channel, **(D)** bright field.

4.4.2 CD31-staining against endothelial cells

In addition to macrophages and T-cells, CD31-positive endothelial cells were added as non-tumour control cells. CD31 is a transmembrane protein that is expressed on all continuous endothelial (i.e. blood vessels) and plays a role in adhesive interactions between adjacent endothelial cells as well as between leucocytes and endothelial cells. The staining of the polyclonal mouse anti-human CD31 antibody was established using commercially available human umbilical vein endothelial cells (HUVECs). As no concentration was specified by the manufacturer of the polyclonal CD31 antibody, the amount used was indicated as dilution. A dilution of 1:250 mouse anti-human CD31 was tested on adhesion slides containing 10%

HUVECs mixed with peripheral blood leucocytes (PBLs). Mouse IgG1κ was used as an isotype control. After incubation with the CD31 antibody, and secondary staining with an AP-polymer anti-mouse solution, the HUVECs showed a strong violet staining of the cell surface, while the isotype control was negative (**Figure** 15).

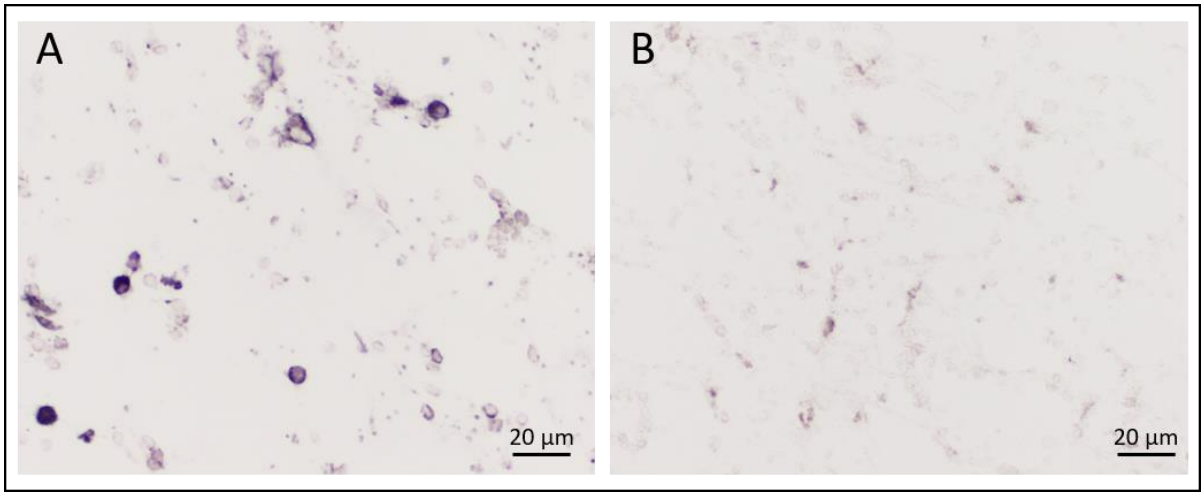


Figure 15: Single staining of HUVECs with a mouse anti-human CD31 antibody

10% HUVECs were mixed with PBLs and transferred to adhesion slides. (A) The slide was stained with a 1:250 mouse anti-human dilution of the antibody directed against CD31 to detect cells of endothelial origin. (B) Mouse IgG1k was used as isotype control in the same concentration. No stained cells could be found on this slide.

Afterwards, adhesion slides with a LN sample of a patient who was negative for DCCs were stained with a dilution of 1:250 of the CD31 antibody. As an isotype control, mouse IgG1k was used in the same concentration. The staining with the 1:250 dilution of the mouse anti-human CD31 antibody was successful. CD31 positive cells with a violet staining of the cell surface were found in LN (**Figure 16**) samples. The isotype control was negative.

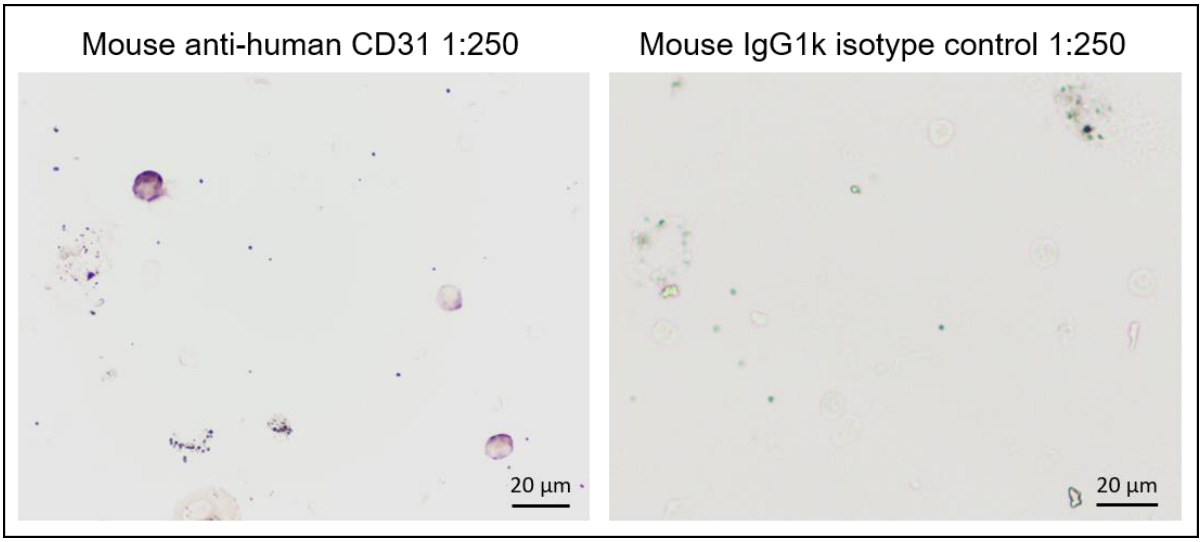


Figure 16: Anti-CD31 staining on LN adhesion slides

Adhesion slides of LN sample was stained with a 1:250 dilution of the mouse anti-human CD31 antibody. The CD31-positive cells showed a violet staining of the cell surface. The isotype control stained with the same concentration of the mouse IgG1k showed no cell-specific staining.

4.4.3 Isolation of non-tumour control cells for the lineage tree analysis of patient MM15-127

After successfully establishing the staining on test slides, LN slides from patient MM15-127 were used to stain and isolate non-tumour control cells. Therefore, the LN was disaggregated into a single cell suspension, put onto adhesion slides and frozen away until further usage. T-cells were stained against CD3 in a fluorescent staining on an adhesion slide (**Figure 17 A**). Endothelial cells were stained with the anti-CD31 antibody that was detected with an alkaline phosphatase system on an additional slide. (**Figure 17 B**).

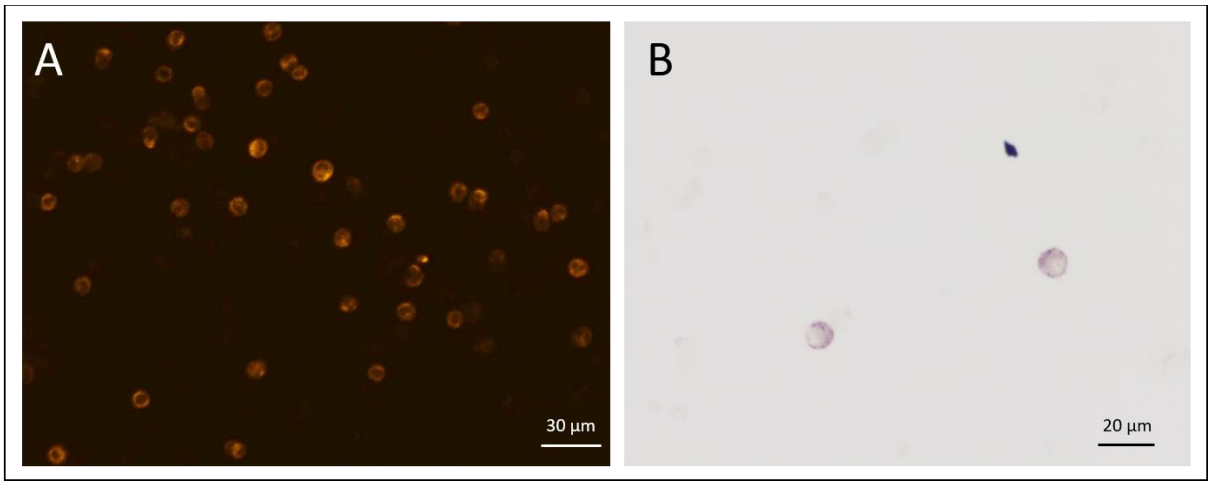


Figure 17: Staining of non-tumour control cells from patient MM15-127

(A) Staining of T-cells against CD3 and **(B)** staining with an anti-CD31 antibody for the isolation of endothelial cells. All stainings were performed on adhesion slides prepared from single cell suspensions of LN tissue.

After staining of the LN slides (detailed procedure see **3.5.2** and **3.5.3**), single cells were isolated manually by micromanipulation (see **3.7**). Unfortunately, the macrophages were isolated from another patient and were therefore not available for the cell lineage reconstruction of patient MM15-127. To compensate, 40 T-cells and 16 endothelial cells were isolated, which were subsequently processed by *Ampli1* WGA (see **3.9**) 34 T-cells and 15 endothelial cells showed acceptable quality with a genome integrity index (GII) >2 defined by multiplex PCR (see **3.9.4**). The GII is defined by the number of PCR bands after the multiplex PCR and reflects the quality of the WGA sample. The GII ranges from 0 to 4, with 4 being the highest quality. We selected 25 T-cells with a GII of 3 or 4 and all 15 endothelial cells with a GII of 3 or 4 for cell lineage tree analysis. In addition, we isolated bulk gDNA from the SLN using a DNeasy Blood & Tissue Kit (see **3.8**). We used five replicates of 1 ng gDNA and five replicates of 10 ng gDNA for WGA. All ten samples had a good DNA quality (GII>2) and were selected for lineage tree analysis (**Table 16**).

Table 16: Number of non-tumour control cells selected for lineage tree (LT) analysis of patient MM15-127

Cell / tissue types	# Macrophages	# T-cells	# Endothelial cells	Bulk gDNA from LN
Isolated by micromanipulation	Due to technical issues, no macrophages were isolated	40	16	5 x 1 ng amplified 5 x 10 ng amplified
With GII>2	0	34	15	10 replicates
Selected for LT analysis	0	25	15	10 replicates

4.5 Lineage tree analysis of patient MM15-127

The aim of the lineage tree analysis of patient MM15-127 was to test our hypothesis if phenotypically small MCSP-positive DCCs are precursors of large MCSP-positive DCCs and therefore represent very early DCCs. Our collection of samples and cells for the lineage tree cell comprised 50 non-tumour samples (**Table 17**, 25 T-cells, 15 endothelial cells, 10 bulk gDNA replicates from LN) for the stabilization of the tree and to test if non-tumour cells could be separated from tumour cells. In addition, the cohort comprises 163 tumour samples (**Table 17**, 30 small and 25 large MCSP+ cells, 18 gp100+ cells, 24 cells from the DCC-derived cell line, 18 and 24 cells from metastases 1 and 3, 24 replicates from 3 PT punches).

Table 17: Summary of cell collection for the lineage tree analysis of patient MM15-127

Sample type	Number of samples
Non-tumour control samples	**50**
- T-cells	25
- Endothelial cells	15
- Bulk gDNA from LN	10 replicates
Tumour samples	**163**
- Small MCSP+ cells	30
- Large MCSP+ cells	25
- gp100+ cells	18
- DCC-derived cell line cells	24
- Tumour cells from metastasis 1	18
- Tumour cells from metastasis 3	24
- Bulk gDNA from PT	24 samples (8 replicates from 3 punches)

4.5.1 The cell lineage tree reconstruction workflow

The lineage tree analysis aims to uncover the developmental history of a collection of cells based on the somatic mutations that occur naturally during cell division. For the reconstruction, the laboratory of Ehud Shapiro at the Weizmann Institute of Science, Rehovot, Israel, uses short tandem repeats (STRs). STR, also known as microsatellites, are highly abundant regions with repetitive sequences of 1-6 bases. They are prone to *de novo* mutations due to slippage events during DNA replication (Willems et al. 2014; Woodworth, Girskis, and Walsh 2017). Therefore, STRs are a promising mutational source to unravel the cell lineage of selected cells. To amplify the approximately 12000 selected STRs the Shapiro lab uses MIPs (molecular inversion probes), single strand DNA molecules composed of two targeting arms and a linker in between. Compared to PCR multiplexing, they have a much higher targeting throughput and better specificity (Tao et al. 2018). First, MIPs were designed for targeting highly mutable STRs selected from a database (**Figure 18 a**). MIP precursors were synthesized on a microarray. They consist of a pair of universal adaptors, two-3bp unique molecular identifiers (UMIs), two target specific arms and one Illumina sequencing compatible spacer with a total size of 150 bp. The universal adaptors (orange and green) serve as a primer binding site for PCR amplification and are removed after amplification by enzymatic digestion. After purification and quality control, the duplex MIPs are ready to use (**Figure 18 b**). The duplex MIPs are mixed with the template WGA products of single cells where the target arms (dark blue and yellow) anneal to the target region (light blue and red) on the WGA DNA. The MIPs are circularized by gap filling with a DNA polymerase and a ligase. Linear template DNA and excess MIPs are digested by exonucleases. The barcoding (black and green adaptors) for Illumina sequencing is performed in an additional step by PCR reaction. The libraries are pooled and sequenced by the Illumina NGS platform to receive the mutation

status of each amplified STR for lineage tree analysis (**Figure 18 c**). This whole workflow was performed by Ehud Shapiros laboratory at the Weizmann Institute of Science, Rehovot, Israel.

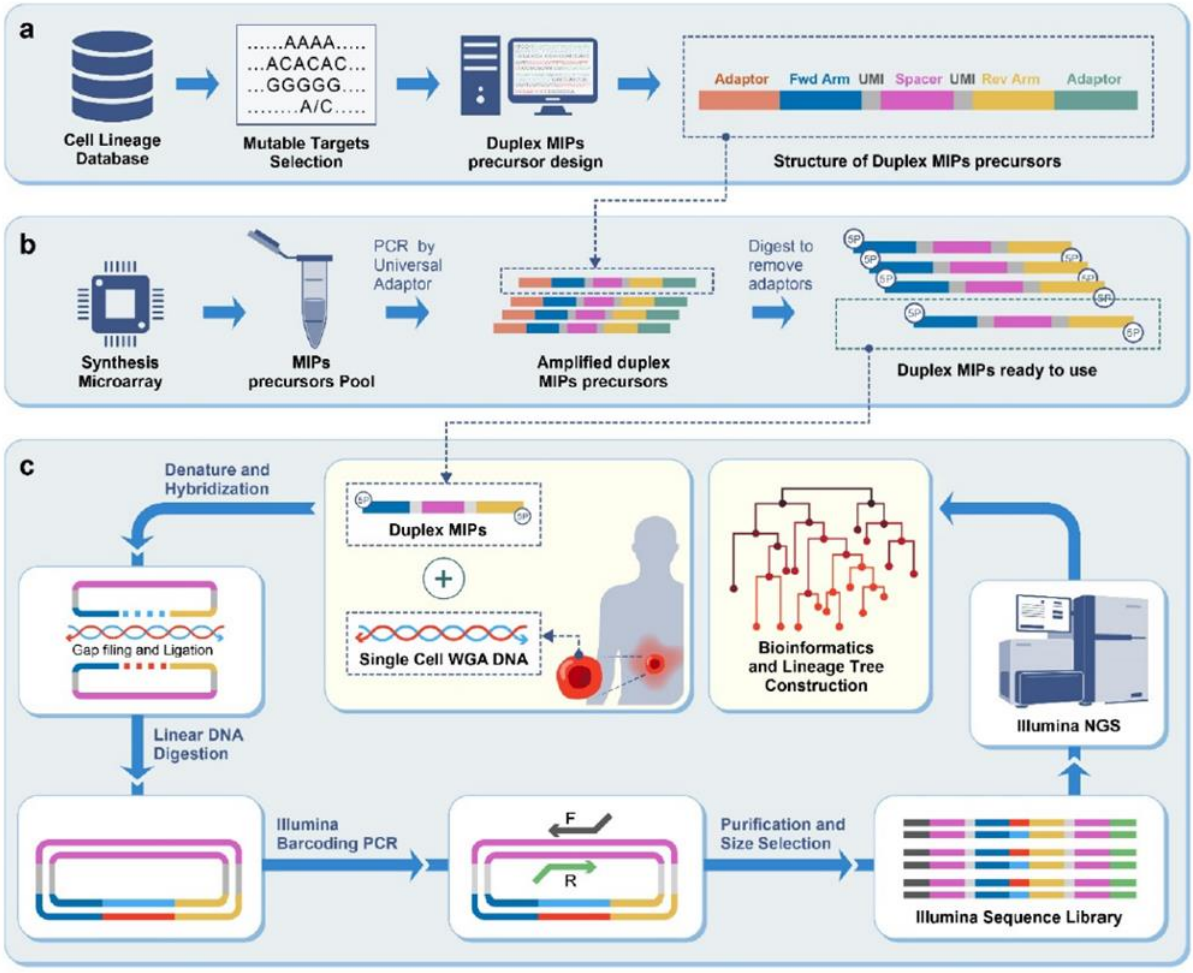


Figure 18: Lineage tree workflow

(a) Highly mutable STR targets were selected from the cell lineage database for MIP design. MIP precursors are composed of a spacer, two unique molecular identifiers (UMIs), two target specific arms and two universal adaptors for PCR amplification. (b) MIP precursors are synthesized on a microarray, amplified by PCR, digested to remove the universal adaptors and purified. (c) Mixing of duplex MIPs and template WGA DNA is followed by hybridization of template DNA and MIPs, gap filling, ligation and digestion of excess linear DNA and MIPs. Barcoding for Illumina sequencing is done by PCR reaction. After purification and pooling, Illumina sequencing is performed. Lineage trees are reconstructed based on mutational status of sequenced STRs. Figure adapted from Tao *et al.* 2018.

4.5.2 Optimization of sample preparation for the cell lineage workflow

Lineage tree analysis of single cells isolated from malignant melanoma patients has not been done before and there was no established protocol for the optimal sample preparation. There were different options for the preparation of our WGA samples for the lineage tree analysis. On the one hand, there was the option to provide some of our original WGA samples. Since these samples are very valuable, there was also the possibility to perform a reamplification of the samples (see 3.11) and use these instead. Next, the samples, original or reamplified, were be subjected to double-strand synthesis to fill up the overhanging single strands created by WGA. Finally, the samples were purified. Due to the various options of sample preparation,

we tested various combinations of the different preparation methods on 20 non-tumour single cells with good quality DNA. The single cells originated from different cancer patients included in different projects. Besides the melanoma patient MM15-127, a breast cancer patient from the project of Manjusha Ghosh and a melanoma patient from Sandra Huber were included. WGA was performed with all single cells. Group C7 (**Figure 19**) was not further processed but tested as original WGA samples. After WGA, double-strand synthesis was carried out with another group (**Figure 19, C1 and C2/5**). The samples were either in the old (**C1**) or the new MIP run of the pipeline (**C2/5**) of the Shapiro laboratory. Alternatively, the samples were subjected to a reamplification and also prepared by the Shapiro laboratory's old (**Figure 19, C3**) or new run of the pipeline (**Figure 19**, **C4**). After reamplification, the samples were either purified immediately with AMPure XP beads (**Figure 19, C9**) or a double-strand synthesis was performed (**Figure 19, C6**) followed by purification (**Figure 19, C8**). If not mentioned otherwise, the samples were prepared with the new MIP pipeline (C6-C9). Furthermore, samples from groups C1-C4 were delivered by us as primary WGA products and prepared completely by the Shapiro laboratory.

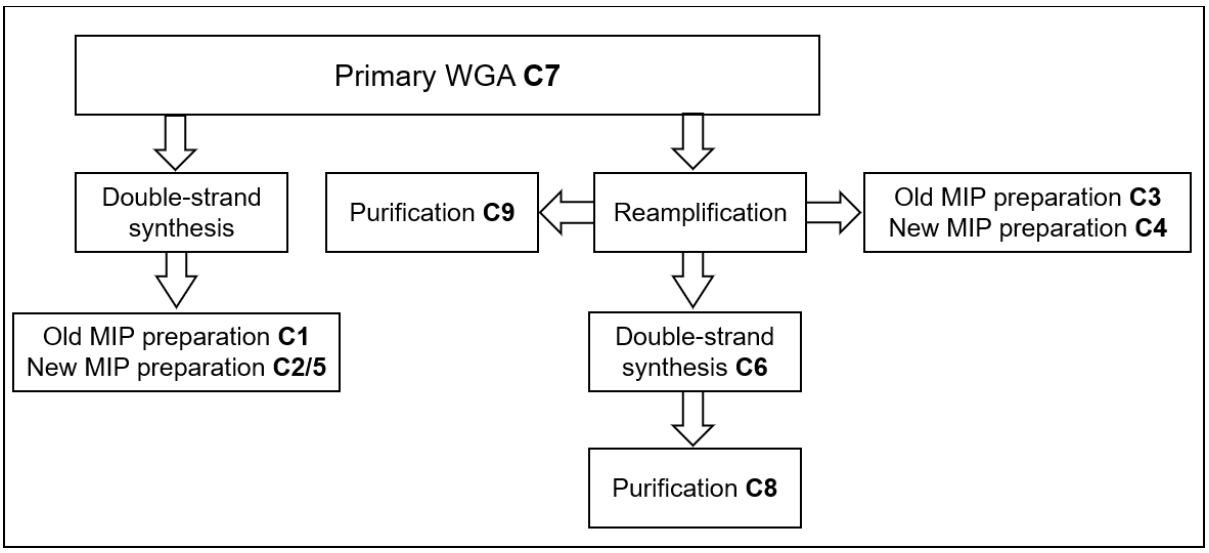


Figure 19: Different methods of sample preparation for lineage tree analysis

WGA was performed for all samples in this experiment (**C7**). The samples were then either subjected to double-strand synthesis and with the old (**C1**) or the new MIP run (**C2/5**) or they were prepared by reamplification with the old (**C3**) or the new MIP run (**C4**). Additionally, samples were purified after reamplification (**C9**). Alternatively, samples were prepared by double-strand synthesis after reamplification (**C6**) followed by purification (**C8**). If not mentioned otherwise, the samples were prepared with the new MIP protocol. Primary WGA samples of groups C1-C4 were prepared by the Shapiro laboratory.

For the comparison of the different groups, the Shapiro laboratory performed library preparation with the old or the new MIP pipeline, followed by sequencing of the samples and evaluation of the mapping rate, which is the percentage of successfully mapped reads to the reference genome. Groups C1 and C2 were prepared identically except for the MIP pipeline. The old MIP preparation provided a mapping rate of almost 90%, while the new preparation decreased the mapping rate to 40-70%. The samples of groups C3 and C4 that were reamplified performed equally poorly with a mapping rate of about 20-40%. Group C5 was

processed in the same way as group C2 and showed a mapping rate of 30-40%. Group C6 performed worst with a mapping rate of around 10%. These samples were prepared by reamplification followed by double-strand synthesis. These results show that the combination of reamplification and double-strand synthesis works worse than double-strand synthesis alone (C1, C2 and C5) or reamplification alone (C3 and C4). The original WGA samples of group C7 had a mapping rate of about 50%. The highest mapping rate of over 90% was achieved with the purification of the reamplified samples. There was no difference between the samples undergoing reamplification, double-strand synthesis and then purification (C8) and the samples being only reamplified and then purified with the AMPure XP beads (C9). These results demonstrate that the key to high quality samples for library preparation is the purification of the reamplified samples. At the end, we decided to reamplify and purify our samples before sending them to the Shapiro laboratory.

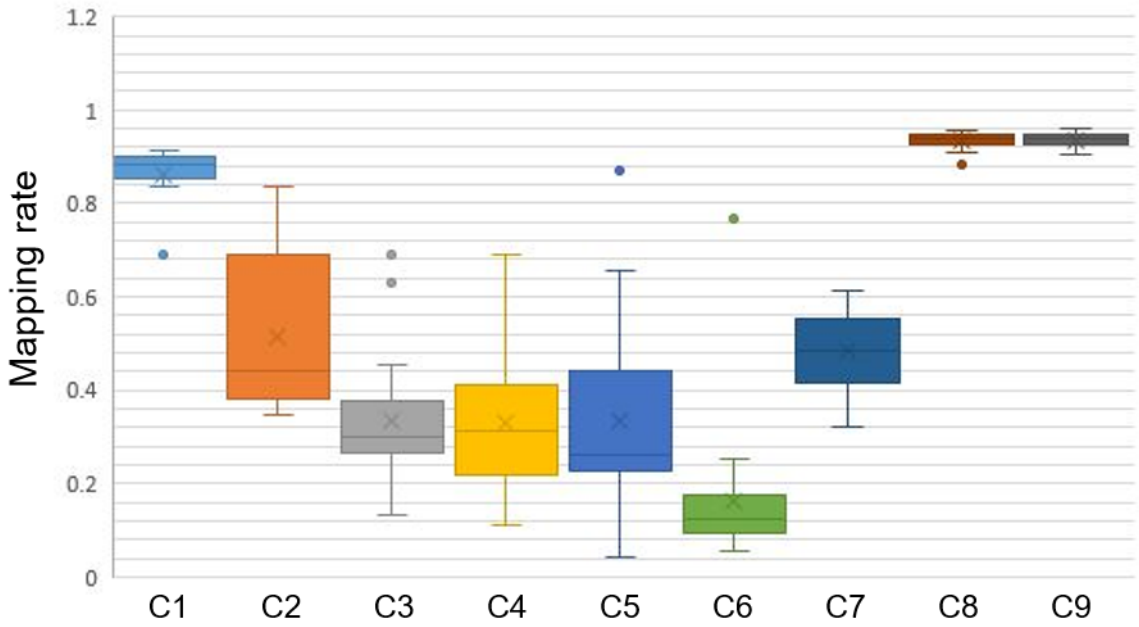


Figure 20: Mapping rates of different sample preparation methods for lineage tree analysis

Detailed description of the different groups C1-C9 can be found in **Figure 19**. **C1**: Primary WGA samples were processed by double-strand synthesis followed by the old MIP preparation. **C2**: identical procedure as C1 but with new MIP preparation. **C3**: Primary WGA samples were reamplified and sequencing library was prepared with the old MIP procedure. **C4**: Identical with C3, but with new MIP preparation. **C5**: Identical to the procedure of C2. **C6**: Primary WGA samples were reamplified, followed by double strand synthesis. **C7**: Primary WGA samples. **C8**: Primary WGA samples were reamplified, double-strand synthesis was performed, followed by purification. **C9**: Reamplification of primary WGA samples subsequently followed by purification. The **figure** was created by the Shapiro laboratory using data from the melanoma patient MM15-127, the melanoma patients from the project of Sandra Huber and the breast cancer patient of Manjusha Ghosh.

4.5.3 Evaluation of the optimal sequencing depth

After successful preparation of the samples from patient MM15-127, library preparation and sequencing were performed by the Shapiro laboratory. The samples were sequenced several times and the sequencing quality was measured as the number of loci covered by more than

30 reads. The more loci per cell are covered, the more can be used for lineage tree analysis and the more s**table** the analysis becomes. First, a NextSeq run was performed, which delivered around half a million reads per cell and between 2000 and 4000 covered loci per cell (**Figure 21, red dots**). To achieve more reads, a second NextSeq run was carried out, providing about a million reads (**Figure 21, blue dots**). The number of covered loci increased to 4000-5000. Most recently, a NovaSeq run was conducted, which generated up to eight million reads. The number of loci covered by more than 30 reads reached a plateau of 8000 at four to five million total reads (**Figure 21, green dots**). From this we conclude that about four to five million reads are sufficient to cover the largest possible number of loci.

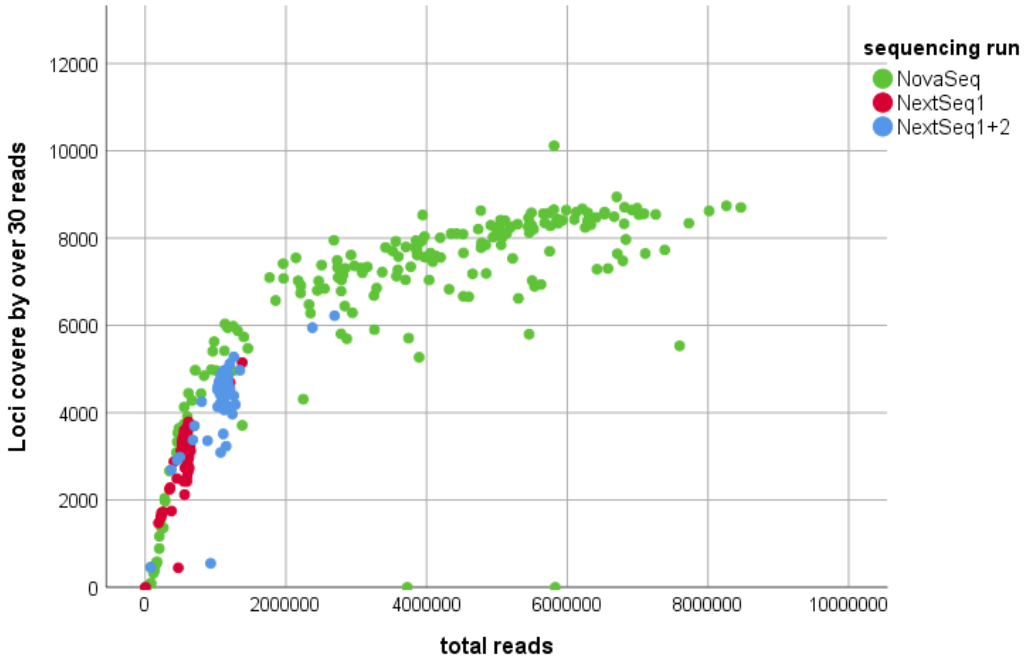


Figure 21: Loci covered by over 30 reads as a function of total reads

The number of loci covered by more than 30 reads of the patient MM15-127 increases as the number of total reads increases. The NextSeq1 run (red) delivered half a million reads and 2000 to 4000 loci covered by more than 30 reads. An additional NextSeq run (blue) increased the number of reads to about one million and the loci covered by over 30 reads to 5000. The NovaSeq run (green) provided up to eight million reads and showed that around four million reads a plateau is reached at 8000 loci covered by more than 30 reads.

4.5.4 Separating cell lineages of different patients

The first step of the lineage tree analysis was to show a successful separation of different patients. For this, we included the T-cells, that were used as non-tumour control cells, from four different cancer patients in a lineage tree analysis. The other three patients, the breast cancer patient included in the project of Manjusha Ghosh and two malignant melanoma patients from Sandra Huber's project, were used to show that patients can be reliably separated by lineage tree analysis. For the reconstruction of the tree, the Shapiro laboratory used the Triplet MaxCut (TMC) algorithm (Sevillya, Frenkel, and Snir 2016). For the analysis, only STR loci are included that present at least two different variants of the STR locus. Furthermore, both variants of one locus had to occur in at least three cells. The algorithm puts

all cells - in the analysis called “leaves” - at equal depth. Therefore, the depth had to be calculated in an additional step by comparing the number of mutations in the STRs from the patients with an *ex-vivo* tree of the DU-145 cancer cell line (Biezuner et al. 2016). The depth of the tree is proportional to the number of cell divisions, although it is not an absolute number. For the reconstruction of the tree with the T-cells of different patients, the cells were sequenced with about one million reads per cell. Overall, a clear separation the patients’ T-cells isolated from a breast cancer patient and three melanoma patients can be seen (**Figure 22**).

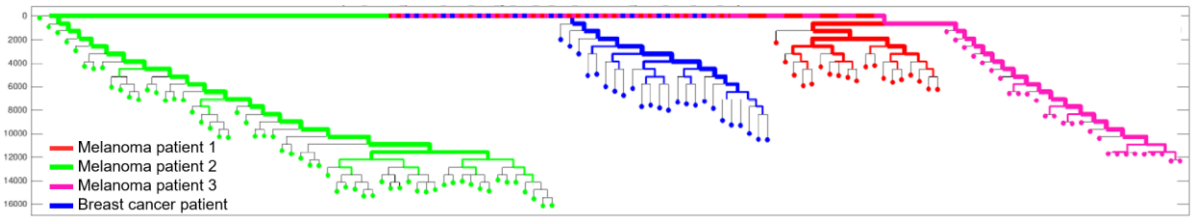


Figure 22: Lineage tree analysis of T-cells isolated from different patients

T-cells were isolated from four different cancer patients and sequenced with about one million reads per cell. The depth of the cells (y-axis) is proportional to the number of cells divisions the cells have undergone. The T-cells from the breast cancer patient (blue), the melanoma patient 1 (red), melanoma patient 2 (green) and melanoma patient 3 (pink) were clearly separated. The three additional patients included in this analysis were part of another project. The vertical axis on the left represents the estimated relative cell depth.

4.5.5 Separating non-tumour control cells by lineage tree analysis

In the next step of the analysis we tested if non-tumour control cells within one patient can be separated. For that we included T-cells, endothelial cells and the bulk gDNA isolated from the LN of patient MM15-127 (n=35) that were sequenced with the desired depth of around five million reads per cell. Simultaneously, improvements in tree reconstruction algorithms continued in the Shapiro laboratory. Namely, the introduction of the FastTree2 algorithm for the tree reconstruction, with a uniform transition table (all mutations are of equal probability). We could observe a separation of the bulk gDNA isolated from the LN (**Figure 23**, green dots), the endothelial cells (**Figure 23**, red dots) and the T-cells **(Figure 23**, blue dots). One endothelial cell ended up in a branch with the bulk DNA, while two bulk DNA samples ended up in the branch with the endothelial cells (**Figure 23**, black arrows). The endothelial cells were isolated from the same LN from which the bulk DNA was isolated.

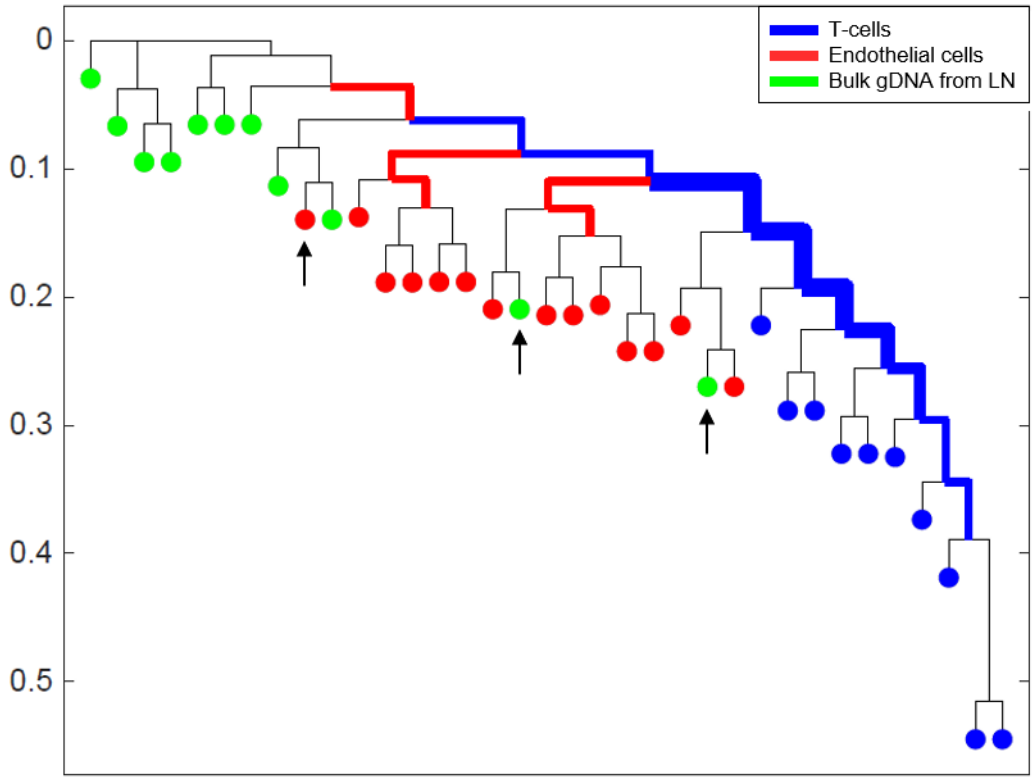


Figure 23: Lineage tree of healthy non-tumour control cells

The lineage tree was reconstructed with the non-tumour control cells of patient MM15-127. The gDNA from LNs (green dots), the endothelial cells (red dots) and the T-cells (blue dots) form three distinct clusters. One endothelial cell and two replicates of the gDNA from LNs are mislocated in another branch (black arrows). The final edges connecting the cells to their immediate ancestors have been trimmed to a fix length to improve the visibility of the higher edges. The vertical axis on the left represents the estimated relative cell depth.

4.5.6 Separating non-tumour control cells and DCCs by lineage tree analysis

The final step of the lineage tree analysis was the inclusion of the DCCs, the bulk DNA from the PT and tumour cells, isolated from the metastases. Together with the non-tumour control cells, their STR profiles were genotyped using Raz *et al.,* 2019 and reconstructed using FastTree2 (Price, Dehal, and Arkin 2010), resulting in the lineage tree shown below (**Figure 24**, n=190). In total, 8 different clusters could be observed. At the top of the tree (left side), the bulk gDNA from the LN was located (**Figure 24**, 1). The next branch was formed by the endothelial cells (**Figure 24**, 2). Then a cluster of both small and large MCSP-positive DCCs was seen (**Figure 24**, 3). Next we had two branches with T-cells (**Figure 24**, 4) and a cluster with the bulk gDNA from the PT (**Figure 24**, 5). On the opposite branch a cluster of the gp100-positive DCCs (**Figure 24**, 6), the single cells from the DCC-derived cell line (**Figure 24**, 7) and the metastatic tumour cells from metastases 1 and 3 was formed (**Figure 24**, 8). Although we could observe a clear separation of the different cell types, there were 3 small MCSP-positive DCCs mixed up with the endothelial cells (**Figure 24**, arrows). Furthermore, 3 T-cells and 2 gDNA samples from the LN were located in the DCC cluster with the small and large MCSP-positive DCCs (**Figure 24**, arrowheads).

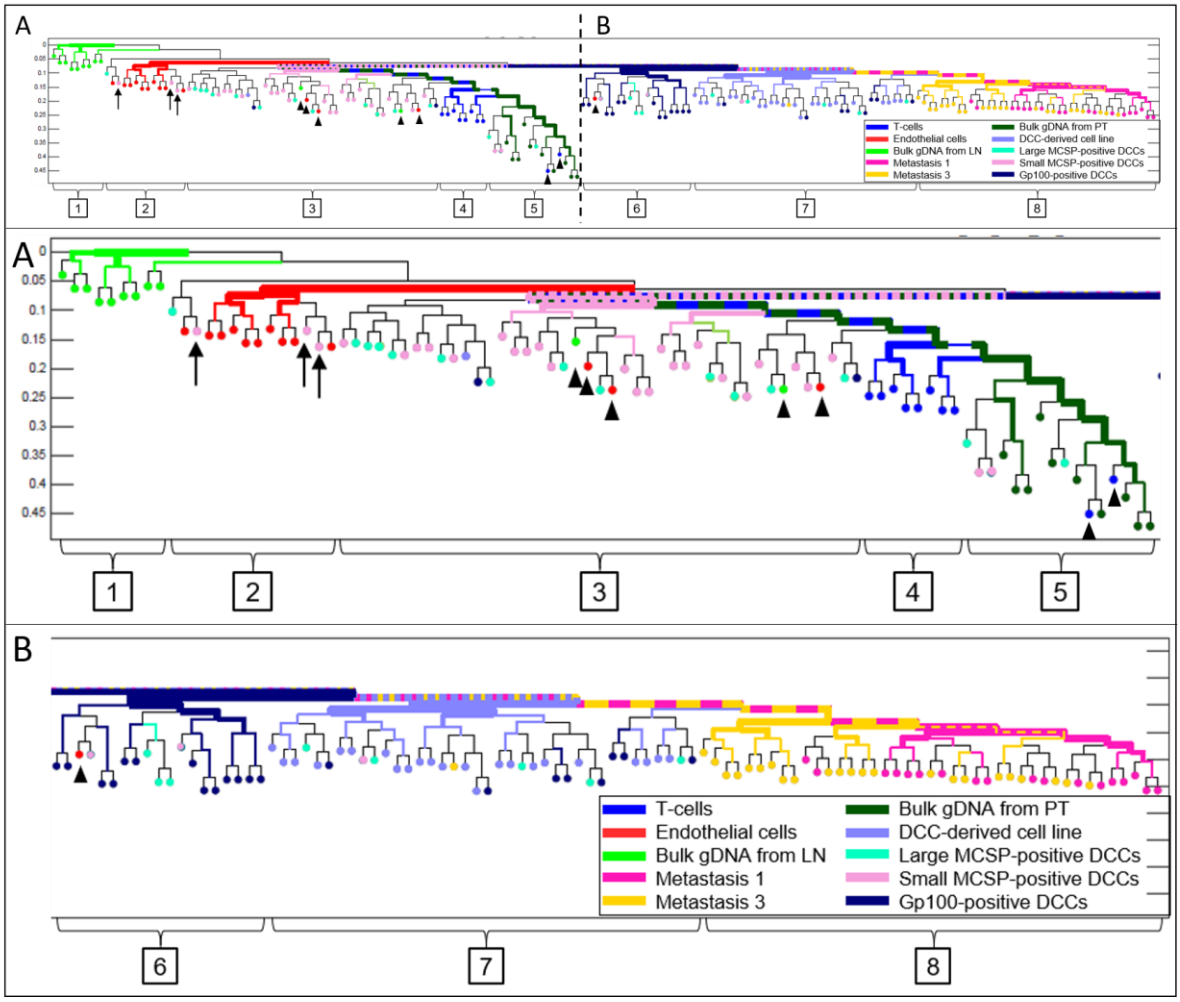


Figure 24: Lineage tree of non-tumour control cells and all tumour cells

The lineage tree of all tumour cells and non-tumour control cells (n=190) is first shown as an overview and then divided into two parts and enlarged (**A+B**). The lineage tree analysis showed 8 distinct clusters.(**A**) (**1**) bulk gDNA from LN (green), (**2**) endothelial cells (red), (**3**) a cluster with small (rose) and large MCSP-positive DCCs (turquoise), (**4**) T-cells (blue) and (**5**) bulk gDNA from PT (dark green). (**B**) On the other branch (**6**) gp100-positive DCCs (dark blue), (**7**) single cells from the DCC-derived cell line (purple) and (**8**) single cells from the metastases 1 (pink) and 3 (yellow) were located. Mislocated small MCSP-positive DCCs in the endothelial cluster were marked with an arrow, endothelial cells and bulk DNA from LNs mislocated in the DCC-cluster were marked with an arrowhead. The vertical axis on the left represents the estimated relative cell depth.

4.5.7 CNV analysis of tumour and non-tumour cells to determine the true origin of the cells

The CNV analysis was performed on all tumour cells and non-tumour control cells included in the lineage tree analysis of patient MM15-127. The aim of the CNV analysis was to improve the lineage tree analysis by only including those cells of which the origin is clear. Although we stained the tumour cells against the tumour markers MCSP and gp100, it cannot be excluded that non-tumour cells expressed one of the markers and was therefore wrongly isolated and determined a tumour cell. MCSP expression has also been described in a variety of normal tissues, like endothelial cells, chondrocytes and certain basal keratinocytes within the epidermis (Campoli et al. 2004). Furthermore, unspecific staining could also not be excluded.

On the other hand, tumour cells could undergo morphological mimicry to adapt to the ectopic environment of the LN. Vascular mimicry was observed in aggressive primary and metastatic melanoma, where tumour cells reconstituted the vascular channels in human tumour tissue and expressed endothelial cell markers (Maniotis et al. 1999). Therefore, we decided to evaluate the CNV profiles of all cells and include only DCCs and tumour cells with an aberrant genome and non-tumour cells with a balanced genome in the lineage tree analysis. Non-tumour cells with an aberrant genome could either indicate poor DNA quality or that they are not of healthy origin. Therefore, aberrant non-tumour control cells (**Figure 25, A**) were excluded from the analysis. All non-tumour control cells included in the lineage tree analysis after the CNV analysis had a balanced genome (**Figure 25, B**). The DCCs and tumour cells that finally remained in the lineage tree all showed an aberrant CNV profile (**Figure 25, C**). Tumour cells with a balanced genome (**Figure 25, D**) were excluded from the lineage tree analysis since we cannot be sure about their malignant origin.

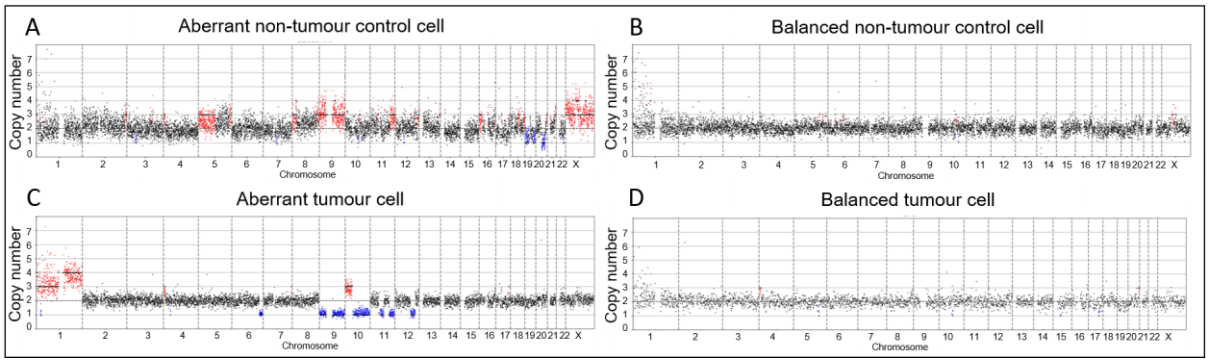


Figure 25: Examples for CNV profiles of aberrant and balanced non-tumour control cells and tumour cells

(**A**) After the CNV analysis of all cells included in the lineage tree analysis, aberrant non-tumour control cells were excluded from the lineage tree. (**B**) Non-tumour control cells with a balanced genome and (**C**) DCCs with an aberrant genome remained in the lineage tree. (**D**) Tumour cells with a balanced genome were removed from the lineage tree analysis. Genomic losses are depicted in blue, genomic gains in red.

The CNV analysis of all cells previously included in the lineage tree analysis has provided the following results: 3 of 39 non-tumour control cells had an aberrant genome (1 of 25 T-cells and 2 of 14 endothelial cells) and were excluded from the lineage tree analysis. Furthermore, 6 T-cells and 2 endothelial cells were also excluded because of poor CNV profile quality (**Table 18**). We had concluded that a poor CNV profile quality might result in a poor sequencing quality for lineage tree analysis. Moreover, cells with a poor CNV profile could not be clearly classified as balanced or aberrant. In total, we ended up with 18 T-cells and 10 endothelial cells for the lineage tree analysis (**Table 18**).

Of the 152 tumour samples, 119 had an aberrant genome, while 28 showed a balanced profile. Surprisingly, only 5 of 32 small MCSP-positive cells (15.6%) had an aberrant genome, while 26 cells were excluded due to a balanced CNV profile and one due to poor profile quality (**Table 18**). In contrast, 21 of 24 large MCSP-positive cells (87.5%) had an aberrant genome, while only 2 cells were excluded because of a balanced profile and 1 due to poor profile quality. There was a significant difference in the number of aberrant cells between small and large

MCSP-positive cells (chi-square, n=54, p=0.000). In total, we ended up with 5 small and 21 large MCSP-positive cells. The gp100-positive DCCs, cells from the DCC-derived cell line, the metastatic tumour cells as well as the bulk gDNA from the PT all showed aberrant CNV profiles. Solely 3 gp100-positive tumour cells had to be excluded due to poor profile quality. There was no significant difference in the number of aberrant cells between large MCSP-positive cells, gp100-positive cells, the cell line cells and the metastases (chi-square, n=110, p=0.103). In total, we ended up with 119 tumour samples having an aberrant genome (**Table 18**).

Table 18: Results of the CNV analysis of the cells included in the lineage tree analysis.

LT: lineage tree; met: metastasis; PT: primary tumour

Sample type	Number of samples for CNV analysis	Samples with aberrant profile	Samples with balanced profile	Aberrant control cells and balanced tumour cells	Samples with poor CNV profile quality
Non-tumour control cells	**39**	**3**	**28**	**3**	**8**
- T-cells	25	1	18	1	6
- Endothelial cells	14	2	10	2	2
Tumour samples	**152**	**119**	**28**	**28**	**5**
- Small MCSP+ cells	32	5	26	26	1
- Large MCSP+ cells	24	21	2	2	1
- gp100+ cells	24	21	0	0	3
- DCC-derived cell line cells	24	24	0	0	0
- Tumour cells from met 1	18	18	0	0	0
- Tumour cells from met 3	24	24	0	0	0
- Bulk gDNA from PT	6 replicates	6 replicates	0 replicates	0	0

4.5.8 Overview of the cell loss during the different steps before the cell lineage tree analysis

During the samples preparation for the cell lineage tree analysis, all cells underwent several quality controls. First, the DNA quality was determined by the genomic integrity index (GII; see **3.9.4**). Cells with a GII>2 were considered to be of good quality and were included in the CNV analysis. In this first quality control step, the number of cells was reduced by 25% (**Table 19**). Non-tumour control cells with an aberrant CNV profile were excluded from the lineage tree analysis. The aberrant profile could either indicate a poor DNA quality or a non-healthy origin of the control cells. Tumour cells with a balanced CNV profile were also excluded from further lineage tree analysis since their malignant origin could not be reliably determined. By this quality control, the number of cells was further reduced by 25% (**Table 19**). For the lineage tree analysis, around 12000 short tandem repeats (STRs) were sequenced of each single cell. Cells that were not sufficiently covered by the sequencing and did not deliver enough STR regions to be placed reliably in the tree were also excluded from the final tree reconstruction. This final quality control step reduced the number of cells by an additional 14% (**Table 19**).

Table 19: Overview of the cell loss during the different quality control steps of patient MM15-127

	Total number of cells	Small MCSP+ cells (fresh / frozen)	Large MCSP+ cells (fresh / frozen)	gp100+ cells	DCC-derived cell line cells	Tumour cells from metastasis 1+3	T-cells	Endothelial cells
Isolation by micromanipulation	251	2 / 32	5 / 24	26	26	42 + 44	40	16
Cells with GII>2	188 of 251	2 / 28	5 / 20	24	24	18 + 24	34	15
Cells with confirmed origin after CNV analysis	141 of 188	0 / 5	4 / 17	21	24	18 + 24	18	10
Sufficient coverage of STRs for LT analysis	121 of 141	0 / 4	4 / 16	19	24	16 + 22	8	8

4.5.9 Separating non-tumour control cells by lineage tree analysis after inclusion of CNV analysis

After the CNV analysis we ended up with 18 T-cells and 10 endothelial cells with a balanced CNV profile (**Table 18**). The bulk gDNA samples isolated from the LN were also included in the lineage tree analysis. After the analysis, only 8 of 18 T-cells and 8 of 10 endothelial cells were chosen for the lineage tree analysis (**Table 19**). Following a repeated reconstruction with the remaining cells, a clear separation between the bulk gDNA from the LN, T-cells and endothelial cells was visible (**Figure 26**, n=24). Nevertheless, there were still one mislocated bulk DNA sample and one endothelial cell (arrows in **Figure 26**).

Furthermore, for this tree the bootstrap method was applied, which is a statistical approach to study the robustness and variability of a dataset. Repeated resampling with replacement from the original dataset provides replication of the original estimate. With this approach, the variance and the distribution of this estimate can be calculated (Lemoine et al. 2018). This bootstrapping analysis was performed with 1000 iterations to prove that the proximity of the cells within the generated branches is substantially closer than within randomly sampled branches. From these iterations, an index called *transfer bootstrap expectation* (TBE) can be calculated. The TBE ranges from 0 to 1, with 0 meaning that the clustering of certain cells within one branch is random and 1 meaning that the clustering of certain cells appears in all bootstrap trees. The common threshold for a robust branch is typically 70% or higher (Lemoine et al. 2018). In the lineage tree with the non-tumour control cells, all branches marked with an asterisk appear in at least 70% of trees reconstructed on random subsets.

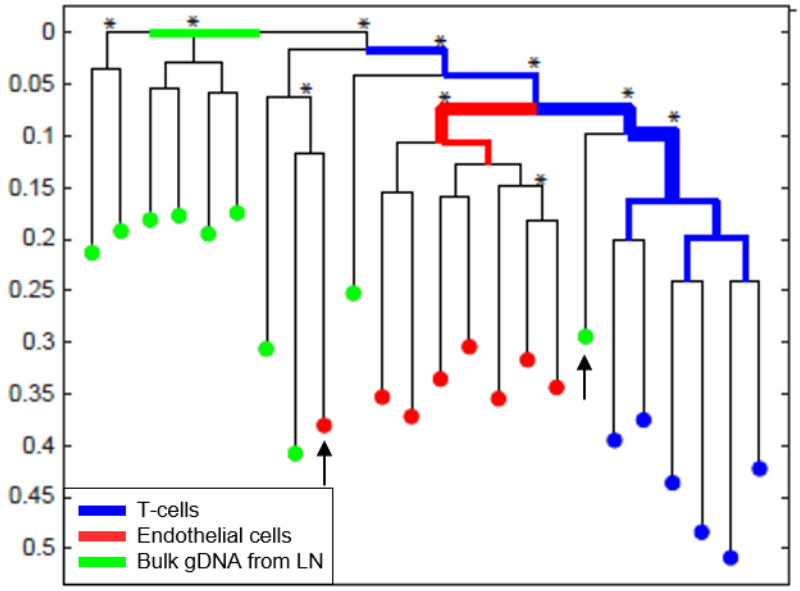


Figure 26: Lineage tree of non-tumour control cells after CNV analysis

The three different sample types, bulk gDNA for a LN (green dots), endothelial cells (red dots) and T-cells (blue dots) were clearly separated. Only one bulk gDNA sample and one endothelial cell were mislocated (arrow). The asterisks indicate a TBE of at least 70%, meaning that in at least 70% of randomly sampled trees, the cells in one particular branch cluster together. The vertical axis on the left represents the estimated relative cell depth. n=24

4.5.10 Separating non-tumour control cells and DCCs by lineage tree analysis after inclusion of CNV analysis

In the final tree reconstruction after the CNV analysis (n=121) we included all non-tumour control cells with balanced genomes, whereas DCCs, metastatic cells, DCC-derived cell line cells and PT gDNA had to display an aberrant CNV profile (all profiles shown in **7.1**). The tree could be separated into 9 different branches: Starting with the gDNA from the LN (**Figure 27**, 1) followed by a branch of endothelial cells and two large MCSP-positive DCCs (**Figure 27**, 2). Next a separate small branch with two large MCSP-positive DCCs (**Figure 27**, 3) followed. Then a big branch with two sub-branches is seen: One that contains mainly the DCC-derived cell line cells (**Figure 27**, 4) and one with metastatic cells (**Figure 27**, 5). The second big branch has 4 sub-branches: Two DCC-branches (**Figure 27**, 6 and 7), a T-cell branch (**Figure 27**, 8) and a PT-branch (**Figure 27**, 9). In total, we saw three different clades of DCCs. Clade number 4 is on the same big branch together with the metastatic cells (**Figure 27**, 4). This clade contains mainly single cells from the DCC-derived cell line plus some large MCSP-positive DCCs and gp100-positive DCCs. DCC clade number 6 is a sub-branch of the second big branch (**Figure 27**, 6). It contains DCC-derived cell line cells, small and large MCSP-positive DCCs, gp100-positive DCCs as well as two metastatic cells and four non-tumour control samples. DCC clade number 7 (**Figure 27**, 7) is located on the same sub-branch as the PT cluster 9 and is composed of small and large MCSP-positive DCCs as well as gp100-positive DCCs. Furthermore, the PT cluster 9 itself contains a small and a large MCSP-positive DCC. Each of the described clusters has a TBE > 0.7 (black asterisk on top of the branches, **Figure 27**), meaning that each of this clusters was generated in 70% of trees reconstructed

on random subsets. To get a more detailed insight into the relationship between the cells, we took a closer look at the CNV profiles of the respective clades.

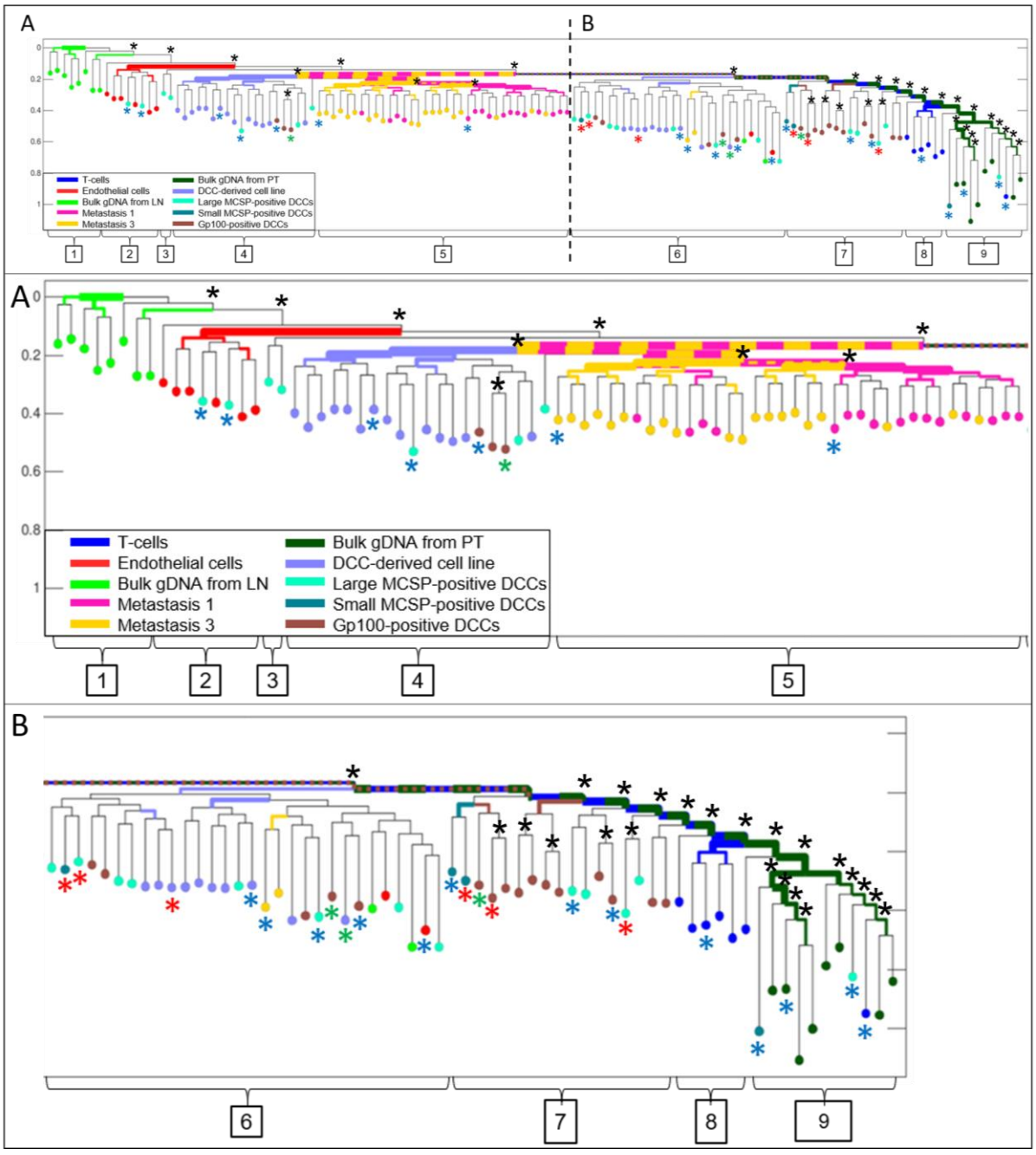


Figure 27: Lineage tree of non-tumour control cells and tumour cells after CNV analysis

The final lineage tree of all tumour and non-tumour control cells (n=121) is first shown as an overview and then divided into two parts and enlarged (**A+B**). The lineage tree showed 9 distinct clusters: (**A**) (1) a cluster of gDNA samples isolated from LN, (2) a cluster with endothelial cells and (3) a small branch with two large MCSP-positive DCCs. A big branch with two sub-branches containing a DCC clade (4) and a metastatic branch (5). (**B**) The second big branch can be subdivided into four sub-branches: (6) a second DCC clade, (7) a third DCC clade, (8) a T-cell branch and (9) a PT branch. Asterisks indicate the shown CNV profiles in **Figure 28**, **Figure 29**, **Figure 30** and **Figure 31**.The vertical axis on the left represents the estimated relative cell depth, the black asterisk on top of the branches the TBE > 0.7.

A comparison of the CNV profiles of all tumour cell included in the final lineage tree has shown that all DCCs (including the DCC-derived cell line) share genomic gains on chromosomes 1 and 10 and losses on chromosomes 6, 9, 10, 11 and 12. The DCCs and the metastatic cells share a gain on chromosome 1 and genomic losses on chromosomes 6, 9 and 10. The metastatic cells share genomic gains on chromosomes 1, 11, 20 and 22 and genomic losses on chromosomes 6, 9, 10 14 and 15. Those aberrations are also shared between metastatic cells and the PT.

The DCC cluster 4 and the metastatic cluster 5 are located on the same branch, that was opposite the branch with the DCC clusters 6 and 7 and the PT cluster. The DCCs share the gains and losses mentioned above. The large MCSP-positive DCC have an additional loss on chromosome 6 and the gp100-positive DCC a gain on chromosome 3 (**Figure 28**, **Figure 27 A**, blue and green asterisks).

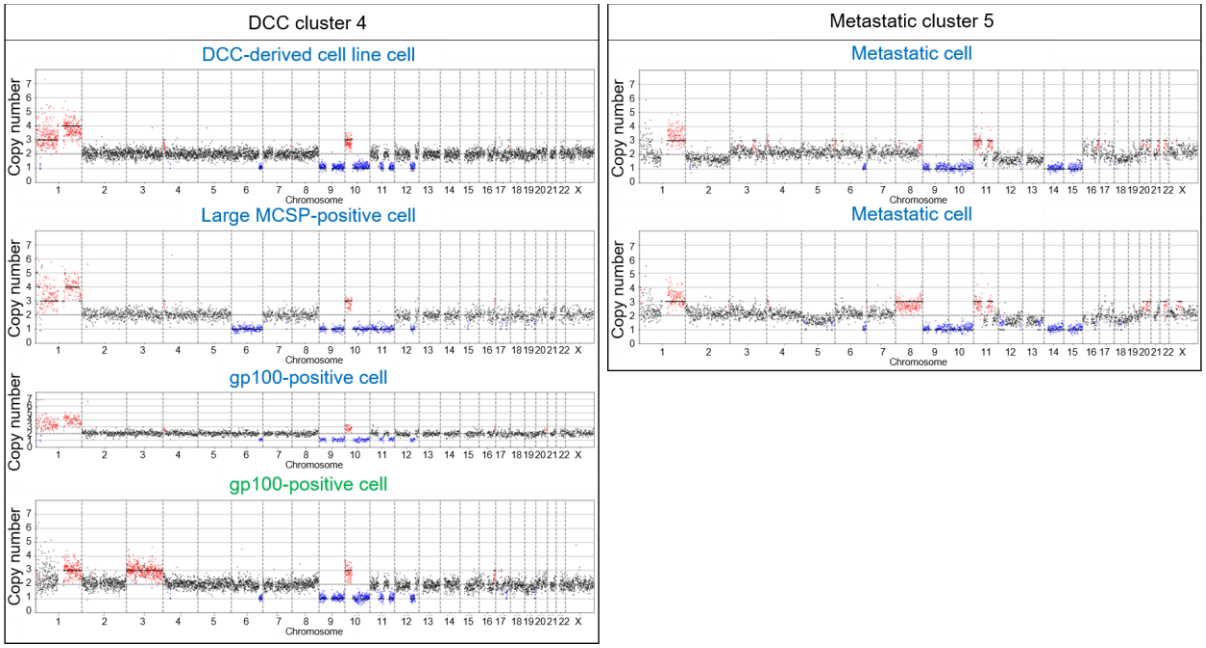


Figure 28: Comparison of DCC cluster 4 and metastatic cluster 5

While all DCCs and metastatic cells share a loss on chromosomes 9 and 10, the metastatic cells harbour an additional loss on chromosomes 14 and 15. In contrast, some large MCSP-positive DCCs and gp100-positive DCCs show a loss on chromosome 6 and a gain on chromosome 3, respectively. The colour of the headline indicates the asterisk in **Figure 27** in the respective clade or cluster. Genomic losses are depicted in blue, genomic gains in red.

DCC cluster 6 is located on the opposite branch of DCC cluster 4 and the metastatic cluster. Gains and losses shared between the DCCs including the DCC-derived cell line were described above. In addition, a small and a large MCSP-positive DCC share a loss of chromosome 5 (**Figure 29**, red headlines, **Figure 27 B**, red asterisks), while a cell of the DCC-derived cell line and a gp100-positive DCC share a gain of chromosome 3 (**Figure 29**, red and green headline, **Figure 27 B**, red and green asterisk).

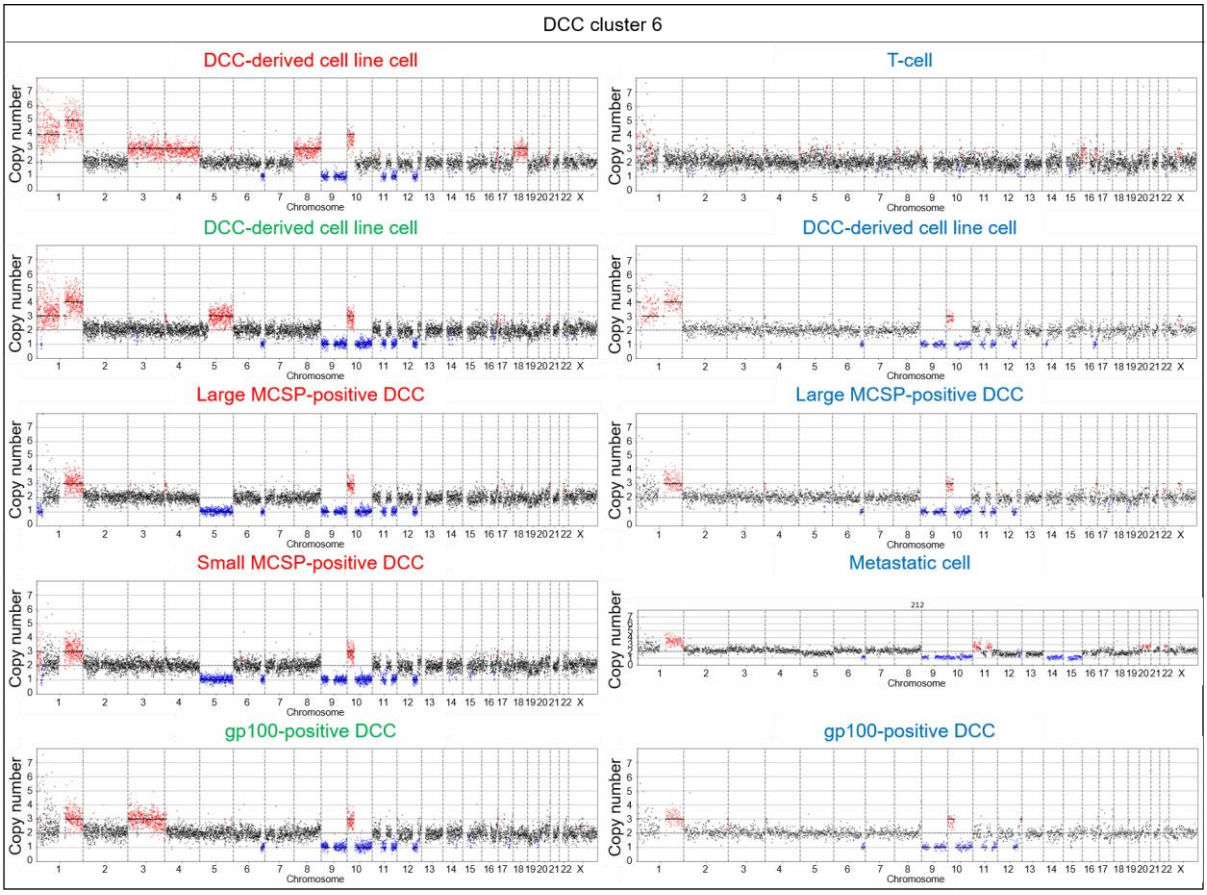


Figure 29: CNV profiles of DCC cluster 6

All tumour cells share the chromosomes 9 and 10 loss. The metastatic cells show an additional loss at chromosomes 14 and 15. The cell line cells show additional gains on chromosomes 3 and 4 or 5. The gain on chromosome 3 is shared by gp100-positive DCCs. The small MCSP-positive DCC and a large MCSP-positive DCC share a loss on chromosome 5. The colour of the headline indicates the asterisk in **Figure 27 B** in the respective clade or cluster. Genomic losses are depicted in blue, genomic gains in red.

The third DCC clade, DCC cluster 7, is located on the same branch as the PT and consisted of all three phenotypes of DCCs: small and large MCSP-positive DCCs as well as gp100-positive DCCs. All DCCs shared genomic gains and losses as described above. In addition, the small and the large MCSP-positive DCCs and the gp100-positive DCC share a genomic loss of chromosome 5 (**Figure 30**, red and green headlines; **Figure 27 B**, red and green asterisk).

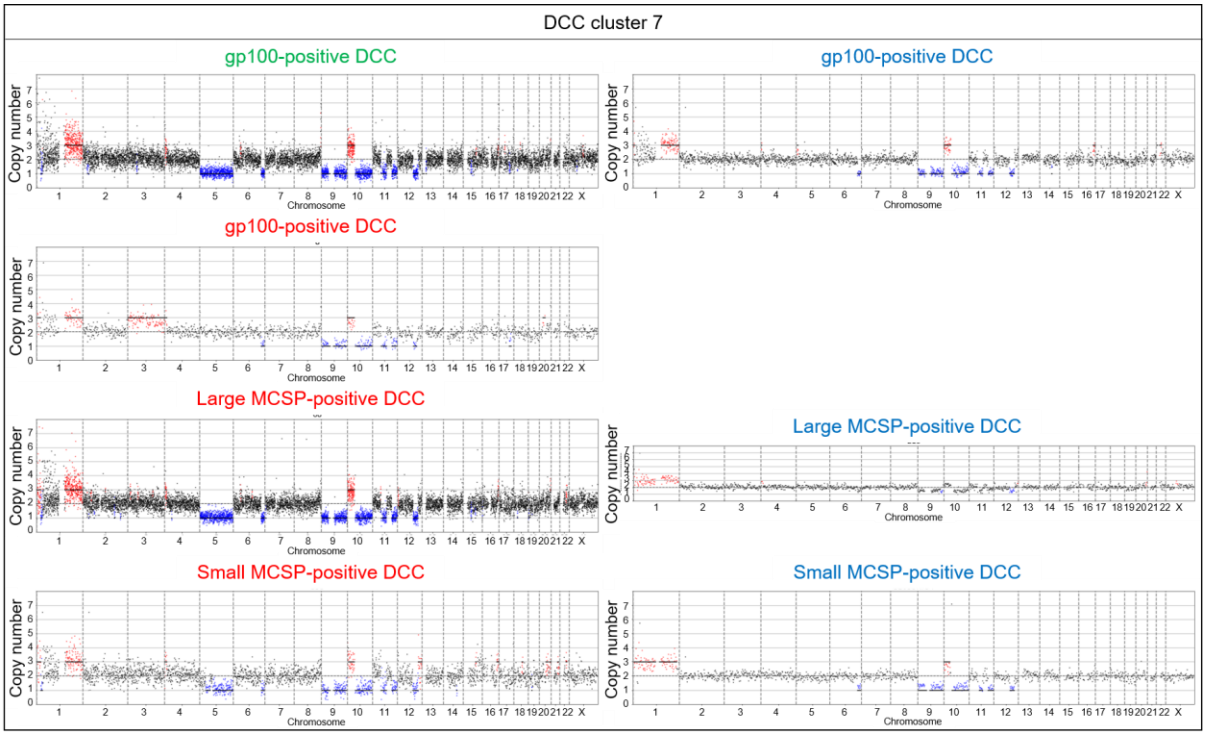


Figure 30: CNV profiles of DCC cluster 7

This cluster contains all phenotypes of DCCs (small and large MCSP-positive and gp100-positive DCCs). All DCCs share losses on chromosomes 9-12. One of the small MCSP-positive DCCs (blue headline) shares a gain on chromosome 10 with all gp100-positive DCCs and a large MCSP-positive DCC (red headline). An additional loss on chromosome 5 is shared by the other small MCSP-positive DCC (red headline), a large MCSP-positive DCC (red headline) and a gp100-positive DCC (green headline). The colour of the headline indicates the asterisk in Figure 27 in the respective clade or cluster. Genomic losses are depicted in blue, genomic gains in red.

The PT cluster 9 mainly consistes of bulk gDNA samples isolated from the PT of patient MM15-127. In addition, it contains a T-cell, a small MCSP-positive DCC and a gp100-positive DCC. The T-cell has a balanced CNV profile, while the small MCSP-positive and the gp100-positive DCC and the PT sample a gain on chromosome 1 and losses on chromosomes 6, 9 and 10. Besides, the PT showed additional gains on chromosomes 5, 7 and 11 and genomic losses on chromosomes 14 and 15 (**Figure 31**).

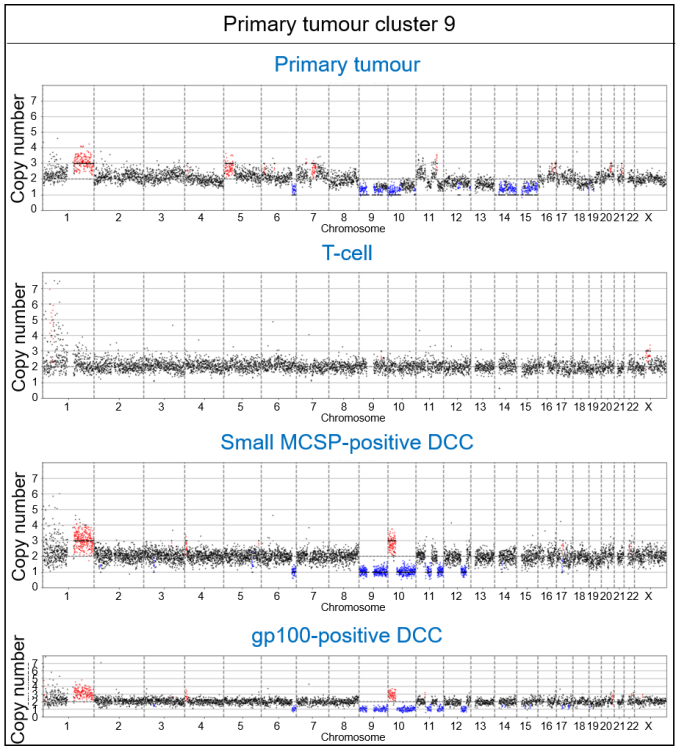


Figure 31: CNV profiles of cells in the PT cluster

The PT cluster consists mainly of bulk gDNA samples isolated from the PT but also contains a T-cell, a small MCSP-positive DCC and a gp100-positive DCC. While the T-cell has a balanced genome, the DCCs and the PT samples share losses on chromosomes 10 and 11. The two DCCs share an additional gain on chromosome 10 and losses on chromosomes 11 and 12. The PT showed an additional gain on chromosome 5 and losses on chromosomes 14 and 15. The colour of the headline indicates the asterisk in **Figure 27 B** in the respective clade or cluster. Genomic losses are depicted in blue, genomic gains in red.

In summary, with the help of the lineage tree analysis we identified three different DCC clades. Clade number 7, which is on the same branch as the PT consists of all three phenotypes of DCCs. The small and the large MCSP-positive DCCs as well as the gp100-positive DCC share an additional loss at chromosome 5 (**Figure 30**; **Figure 27 B**, 7). DCC clade number 6 is a sub-branch of the big branch containing the PT. It consists of all three phenotypes of DCCs, as well as DCC-derived cell line cells and metastatic cells. A small and a large MCSP-positive DCCs share a loss on chromosome 5. Furthermore, a gp100-positive DCC shares a gain with the DCC-derived cell line cell on chromosome 3 (**Figure 29**; **Figure 27 B**, 6). DCC clade number 4 is located on the same branch as the metastatic cluster (**Figure 27 A**, 5) and includes only large MCSP-positive DCCs and DCC-derived cell line cells. The DCCs share a genomic gain on chromosome 1 and losses on chromosomes 6, 9 and 10 with the metastatic cells (**Figure 28**).

4.6 *BRAF* mutation analysis by allele-specific PCR with a blocking reagent

The second aim of this thesis was to investigate the incidence of the two most common *BRAF* mutations in melanoma patients and its association with the DCCD in LNs. For this analysis we established a method called allele-specific PCR with a blocking reagent (ASB-PCR).

4.6.1 Description of the assay method

The purpose of this method is a rapid and cost-effective screening for the *BRAF* hotspot mutation c.1799T>A in single-cell WGA products, primarily in DCCs isolated from patients with malignant melanoma. The ASB-PCR (Morlan, Baker, and Sinicropi 2009) is a method developed to detect somatic mutations at a known base position such as hotspot mutations in oncogenes. Initially, the ASB-PCR assay was developed as a quantitative PCR (qPCR) assay to detect hotspot mutation in the *KRAS* gene in bulk DNA samples. We used the principle of the ASB-PCR and developed an endpoint PCR assay for detection of the *BRAF* c.1799T>A mutation in single-cell WGA products. In order to achieve selective amplification, the method utilizes a mutant-specific primer (**Figure 32**, blue arrow) in combination with a blocking oligonucleotide (**Figure 32**, red line) and an allele-independent reverse primer (**Figure 32**, green arrow). To suppress non-specific amplification of the wild type allele by the mutant-specific primer, a blocking reagent complementary to the wild type sequence is used. It is designed to have the variant base position approximately in the middle of the oligonucleotide and to partially overlap the sequence of the mutant-specific primer (Morlan, Baker, and Sinicropi 2009). Furthermore, the blocking reagent is phosphorylated (-PO_4) at its 3'-end to prevent elongation. In addition to the above-mentioned components our endpoint PCR assay contains an additional allele-independent forward primer, used to confirm that the *BRAF* gene sequence is present in the sample. Both allele-independent products generate a product of 171 bp in size in samples containing the DNA sequence of the *BRAF* gene, whereas combination of mutant-specific primer and the allele-independent reverse primer generate a second amplicon of 147 bp in samples harbouring the *BRAF* c.1799T>A mutation (**Figure 32**). The two reactions can either be performed as two singleplex reactions or combined in a multiplex reaction.

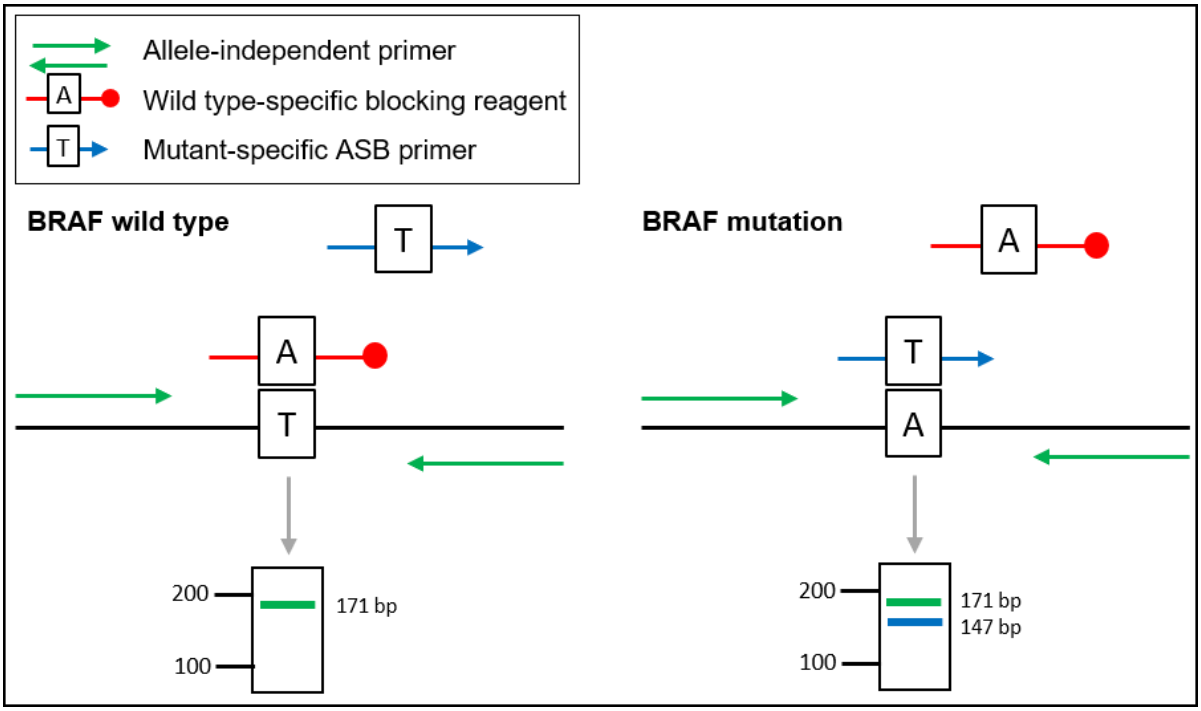


Figure 32: Schematic description of the ASB-PCR assay

For the ASB-PCR assay, two allele-independent primer, a wild type-specific blocking reagent and a mutant-specific ASB primer are used. When *BRAF* is in its wild type form, the wild type-specific blocking reagent binds to the wild type sequence, hindering the binding of the mutant-specific ASB primer. The allele-independent primers amplify a sequence of 171 bp. In case of a *BRAF* c.1799T>A mutation, the mutant-specific ASB primer can bind and amplify in combination with the reverse allele-independent primer a fragment of 147 bp. In addition, the allele-independent primer results in a 171 bp amplicon, verifying the presence of the *BRAF* gene.

4.6.2 Evaluation of the sensitivity, specificity and selectivity of the ASB-PCR assay on genomic DNA

For testing the performance of the ASB-PCR, varying amounts of template DNA were used. The experiment was performed by Franziska Mühlbauer as part of her bachelor's project (Mühlbauer 2017). The sensitivity was here defined as the lowest limit of *BRAF* mutation detection of the ASB-PCR assay, while the specificity was tested by using increasing amounts of wild type template DNA and checking for unspecific amplification of the ASB-PCR. To investigate the sensitivity of the assay the ASB-PCR was performed with varying amounts of DNA template ranging from 0.01 to 25 ng. The mutant DNA was obtained from the malignant melanoma cell line SKMEL28, harbouring the homozygous *BRAF* c.1799T>A mutation. The wild type DNA was isolated from peripheral blood mononuclear cells (PBMCs) from a healthy donor. The assay was carried out as a singleplex reaction using the mutant-specific primer, the blocking oligonucleotide and the allele-independent reverse primer. Using this approach, we were able to detect the mutated *BRAF* gene sequence in as little as 1 ng of bulk DNA. Moreover, we did not detect any unspecific byproducts of the ASB-PCR when using a wild type DNA template, showing a high level of specificity of our ASB-PCR assay (**Figure 33, A**).

Next, we tested the selectivity of the new ASB-PCR assay, here defined as the ability to detect the mutated allele in a background of wild type sequences. For this purpose we prepared DNA templates (with a total of 25 ng DNA per reaction) by mixing mutated (SKMEL28-derived) and

wild type (PBMC-derived) bulk DNA in various ratios with contents of mutated DNA ranging from 0 to 10 ng. Despite the abundant presence of wild type *BRAF* sequence, our ASB-PCR assay was positive even in samples where mutated DNA comprised only 4% (1 ng of mutated mixed with 24 ng of wild type DNA) of the DNA input (**Figure 33, B**).

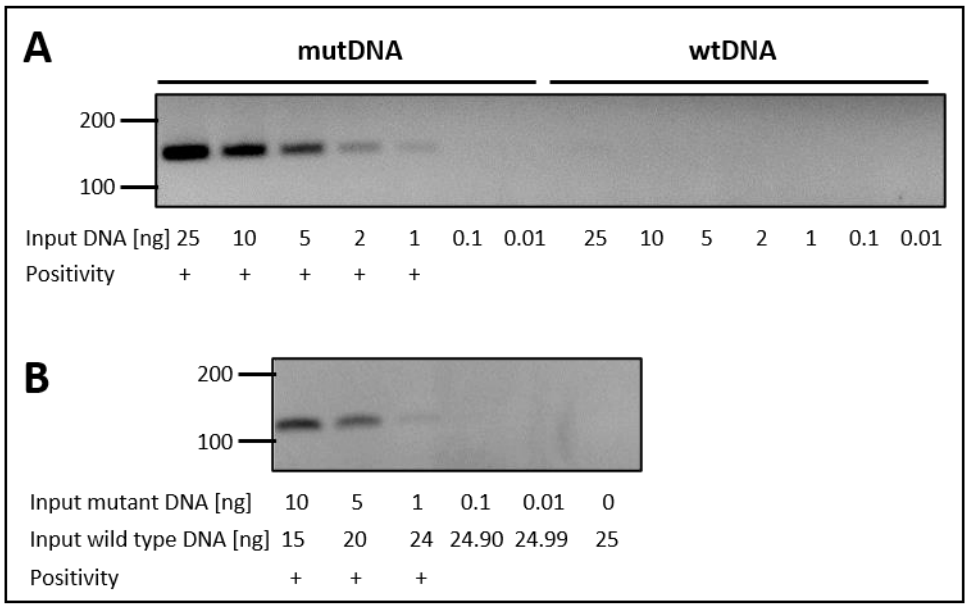


Figure 33: Sensitivity, specificity and selectivity of the ASB-PCR targeting BRAF c.1799T>A

(A) Sensitivity and specificity of the assay were analyzed using agarose gel electrophoresis. Varying amounts (25 to 0.01 ng) of mutant DNA of the SKMEL28 cell line (mutDNA) and wild type DNA (wtDNA) were tested. 1 ng of mutant DNA was amplified by the mutant-specific primer targeting *BRAF* c.1799T>A and the common reverse primer showing a specific band of 147 bp. The wild type DNA was not amplified, showing a high specificity of the assay. **(B)** Selectivity of the assay was analyzed using 10 to 0 ng of fragmented mutant DNA of the SKMEL28 cell line mixed with fragmented wild type DNA in a total amount of 25 ng. 1 ng of mutant DNA mixed with 24 ng of wild type DNA was still detectable. Figure adapted from (Mühlbauer 2017)

4.6.3 Multiplexing of the ASB-PCR with a corresponding control PCR

To check for the overall presence of the *BRAF* locus potentially bearing the c.1799T>A mutation a control PCR using the allele-independent primers was established. The primers used in the control assay were designed to produce DNA fragments representing the *BRAF* locus independently of the presence the *BRAF* c.1799T>A mutation. Consequently, the control PCR generates the same product from heterozygous and homozygous mutant as well as homozygous wild type samples. Amplification of the *BRAF* locus was achieved by combining the allele-independent reverse primer, already used for the ASB-PCR, with an allele-independent forward primer generating a 171 bp fragment. The following experiment was conducted by Franziska Mühlbauer as part of her bachelor's project (Mühlbauer 2017).

The next step of the establishment was to create a multiplex PCR combining the ASB-PCR with the control PCR in one step. Therefore, a touchdown cycler program comprising 33 cycles was tested on the mutated SKMEL28 DNA (**Figure 34, A**) and wild type DNA (**Figure 34, B**) with the input DNA ranging from 0.1 to 25 ng. This protocol showed a limit of detection at 1 ng of template DNA. The mutant specific PCR product was only detected in the *BRAF* mutant

template DNA, while the control fragment was detected in both, the mutant and the wild type template.

At this stage, two additional cycler programs with varying numbers of cycles were tested to find an optimal number of PCR cycles enabling amplification of minute amounts of mutated DNA while preventing detection of unspecific amplification products. As the blocking reagent is only hindering, but not completely abrogating amplification of the wild type sequence variants unwanted amplification of wild type alleles may become detectable in procedures with high number of PCR cycles. Two additional touchdown cycler programs, comprising 36 or 39 cycles were tested on the mutated SKMEL28 DNA (**Figure 34, A**) and the wild type DNA (**Figure 34, B**) with input DNA ranging from 0.1 to 25 ng. The touchdown PCR cycler program with 36 or 39 cycles amplified the respective template down to 0.1 ng in the multiplex approach. The results revealed a higher level of detection for the multiplex touchdown PCR with 36 or 39 cycles. The specificity of the multiplex approach was proven by the fact that the PCR product of the mutant allele was only detected in reactions containing the mutated SKMEL28 DNA (**Figure 34, A**), but not in the reactions containing only the wild type DNA (**Figure 34, B**).

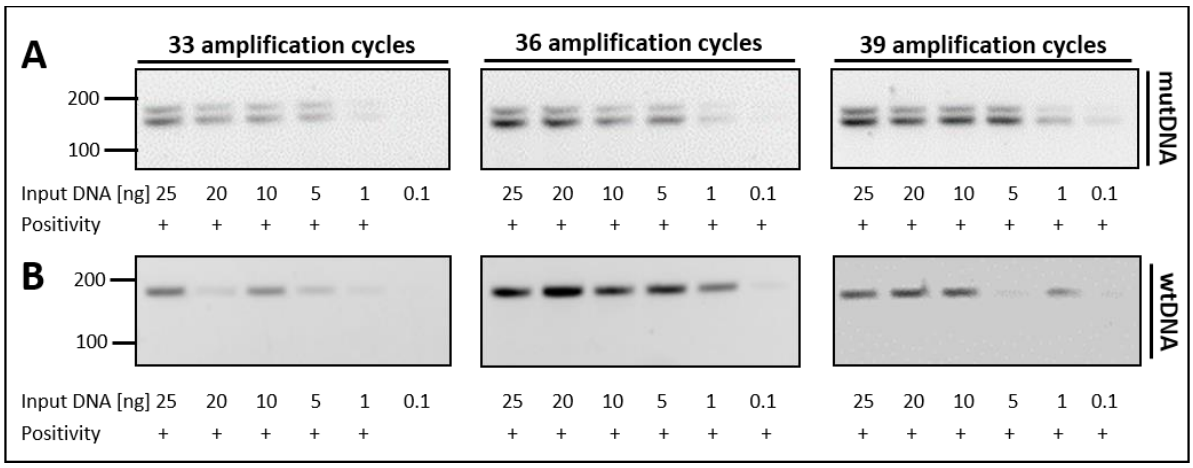


Figure 34: Comparison of three different touchdown PCR cylcer programs used to perform mutliplex PCR for detection of BRAF c.1799A>T

(A) Multiplex PCR was tested with mutant SKMel28 template gDNA in concentrations ranging from 25 to 0.1 ng/µl. The mutant-specific band produced by the mutant-specific and the common reverse primer has a length of 147 bp. The common forward and the common reverse primer used for control PCR produce allele unspecific DNA fragments of 171 bp. Agarose gel electrophoresis of mutant samples showed that 33 cycles enable visualization of mutant DNA down to 1 ng, while 36 and 39 amplification samples allow detection of mutant DNA even down to 0.1 ng. **(B)** Multiplex PCR was also tested with wild type template gDNA to prove the specificity of the ASB-PCR primers for mutant DNA. Agarose gel electrophoresis of wild type samples showed only the allele unspecific DNA fragment of 171 bp produced by the common forward and revers primer. Figure adapted from (Mühlbauer 2017)

4.6.4 Selectivity of the multiplex ASB-PCR on genomic DNA

Next, we evaluated selectivity of our assay, i.e. the extent to which our ASB-PCR method can be used to detect the *BRAF* V600 hotspot mutations without interference from the wild type DNA. To this end, we used mixtures of mutant unamplified bulk DNA isolated from the SKMEL28 cell line and wild type DNA from PBMCs of a healthy donor. The mutant DNA was ranging from 10 to 0 ng mixed with wild type DNA to a total amount of 25 ng, acting as template

for the ASB-PCR. The ASB-PCR was positive for the *BRAF* c.1799T>A mutation in samples comprising as few as 5 ng of mutated DNA mixed with 20 ng of wild type DNA (20% of the utilized template, **Figure 35 A**).

Given the possibility that allelic dropout (amplification bias inherent to all WGA methods; Stevens et al. 2017) may alter the ability of the ASB-PCR to detect the *BRAF* V600 mutation in DNA templates that had previously undergone WGA-based amplification, we tested the selectivity of our assay on WGA products generated from cell mixtures comprising defined percentages of mutated cells. For this purpose, we used cell mixtures comprising 5% - 80% of mutated SKMEL28 cell line mixed with wild type PBMCs. In addition, we prepared samples consisting of 100% SKMEL28 and 100% PBMCs and used them as positive and negative controls, respectively. Three independent WGA products were prepared from each cell mixture and were tested using our multiplex ASB-PCR assay. The *BRAF* c.1799T>A mutation could be detected consistently in all replicates in samples comprising as little as 5% of SKMEL28 cells (**Figure 35 B**). Therefore, we conclude that our multiplex ASB-PCR assay is suitable for both bulk DNA (**Figure 35 A**) and WGA samples generated from cell pools (**Figure 35 B**) allowing detection of *BRAF* c.1799T>A mutation in at least 20% or 5%, respectively, of cells included in the sample. This experiment was conducted by Franziska Mühlbauer during her bachelor's project.

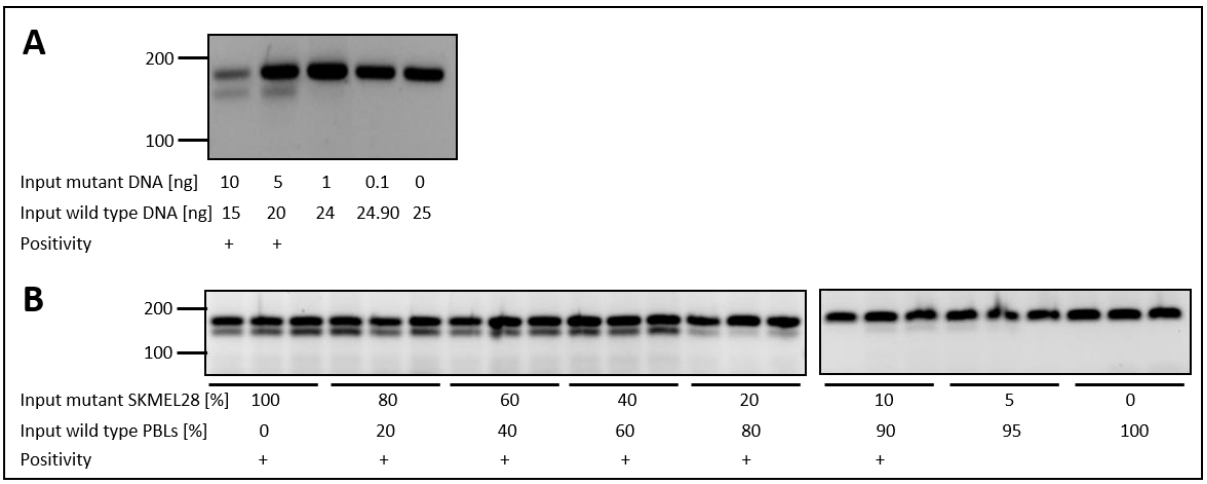


Figure 35: Selectivity of the multiplex assay targeting the BRAF c.1799T>A mutation

(A) The multiplex PCR for the detection of *BRAF* c.1799T>A mutation was first performed on mutant and wild type DNA of a total mixture of 25 ng. The mutation-specific primer in combination with the common reverse primer showed a mutation specific band of 147 bp down to 5 ng of mutant SKMEL28 DNA in a total amount of 25 ng of DNA mixture. The control PCR fragment of 171 bp produced by the common forward and reverse primers was visible in all the samples. **(B)** In the second step, multiplex PCR for the detection of *BRAF* c.1799T>A mutation was tested on WGA products from cell mixtures of mutant SKMEL28 cells with wild type PBLs tested in technical triplets. The mutant specific band of 147 bp is shown in agarose gel electrophoresis down to 5% of mutant SKMEL28 cells mixed with 95% wild type PBLs. The control band of 171 bp was detected in all samples. Figure adapted from (Mühlbauer 2017)

4.6.5 ASB-PCR on single-cell WGA products of malignant melanoma patients

After successful establishment of the multiplex ASB-PCR on WGA products generated from model samples, all of which exhibited excellent DNA quality of starting material reflected by high GII values of 3 or 4, we proceeded with the analysis of patient derived samples. Due to the procedures undertaken to store, fix and stain clinical specimens, the quality of single-cell DNA may be jeopardized, making downstream analysis more challenging. To control the performance of initial tests we selected samples that had been previously sequenced by Sanger sequencing to identify their *BRAF* mutation status. Subsequently, we analyzed five samples harbouring the *BRAF* c.1799T>A mutation and additional five *BRAF* wild type samples. Furthermore, we tested five additional samples with *BRAF* c.1798GT>AA mutation (causative for the BRAF V600E mutation on the protein level). This mutation has not been included in the initial tests of our ASB-PCR method due to the lack of suitable model samples (i.e. lack of cell lines harbouring the *BRAF* c.1798GT>AA). However, given the DNA sequence coding for this hotspot mutation and considering the design of our assay, we assumed that we should be able to detect the *BRAF* c.1798GT>AA equally well as the *BRAF* c.1799T>A mutation. As expected, the *BRAF* mutation was detected by ASB-PCR in all the five clinical samples with the *BRAF* c.1799T>A mutation (**Figure 36**, 1-5) and cells previously classified as wild type by Sanger sequencing were also graded as such by our method (**Figure 36**, 6-10). Importantly, however, our multiplex ASB-PCR could also detect the *BRAF* c.1798_1799GT>AA mutation in all five tested single tumour cells harbouring this mutation (**Figure 36**, 11-15).

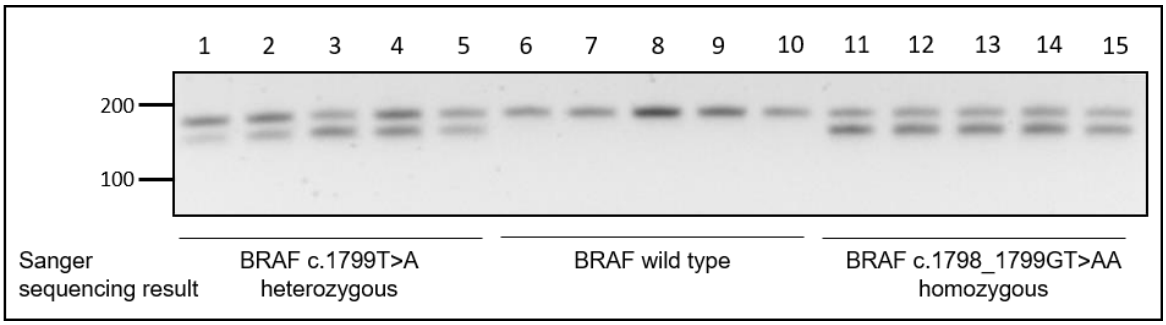


Figure 36: ASB-PCR on single-cell WGA products from malignant melanoma patients

Disseminated cancer cells isolated from LNs of malignant melanoma patients were Sanger sequenced to determine the *BRAF* status. Retrospective testing of the single-cell products harbouring the *BRAF* c.1799T>A mutation (1-5) showed the mutant-specific band (147 bp) as well as the allele-independent control band (171 bp). Single-cells with a wild-type *BRAF* (6-10) showed only the allele-independent control band. Remarkably, the cells harbouring the *BRAF* c.1798_1799GT>AA mutation (11-15) also showed the mutation-specific band. Figure adapted from (Mühlbauer 2017)

To further validate our multiplex ASB-PCR assay on an extended cohort of patient samples, we tested additional 80 single cell WGA products (all cells were derived from LNs of melanoma patients) that were previously tested using Sanger sequencing targeting the *BRAF* c.1799A>T and c.1798_1799GT>AA hotspot mutation. In the initial analysis by Sanger sequencing, 52 samples were tested as *BRAF* wild type, 14 samples harboured heterozygous *BRAF* c.1799T>A mutation, 6 samples were heterozygous for the *BRAF* c.1798_1799GT>AA

mutation and 8 samples were classified as homozygous for the *BRAF* c.1798_1799GT>AA mutation (**Table 20**).

The whole collective was next retrospectively analyzed by multiplex ASB-PCR (**Table 20**). As our aim was to establish a fast and easy to handle assay, we first tested the multiplex ASB-PCR assay with an undiluted template of all 80 samples that were first Sanger sequenced for the *BRAF* mutation. 51 of 52 samples that were classified wild type by Sanger sequencing were also wild type in the ASB-PCR assay (Negative Predictive Value = 98%). Only one of the samples could not be amplified by ASB-PCR. In 27 of 28 WGA products where we detected a *BRAF* mutation by Sanger sequencing, either the c.1799T>A or the c.1798_1799GT>AA mutation, the ASB-PCR confirmed a *BRAF* mutation (Positive Predictive Value = 96%). One sample with a c.1799T>A mutation could not be amplified by ASB-PCR. However, only two samples (one mutant and one wild type WGA product) successfully Sanger sequenced could not be amplified by ASB-PCR (Dropout Rate / Failure to Analysis = 2%).

To prevent amplification failure or false positive results due to excessive concentration of template DNA, we used 1:10 dilutions of the original WGA products as starting material for this experiment (**Table 20**). 51 of 52 cells that were tested as wild type by Sanger sequencing were also found to be wild type by ASB-PCR (Negative Predictive Value = 98%). The remaining wild type sample could not be amplified (technical dropout). Moreover, in 28 of 28 samples previously tested positive for the *BRAF* hotspot mutations (c.1799T>A or c.1798_1799GT>AA) by Sanger sequencing, were also positive in the ASB-PCR assay (Positive Predictive Value = 100%). Only one wild type sample could not be amplified by ASB-PCR due to technical dropouts related to this method (Dropout Rate / Failure to Analysis = 2%). Importantly, the ASB-PCR assay provided no false positive and false negative results (False Positive Rate = 0, False Negative Rate = 0) indicating high accuracy of this method.

Table 20: Sanger sequencing of single-cell WGA products and retrospective testing of the ASB-PCR

Mutation status of the cell	Cells tested by Sanger sequencing (n=80)	ASB-PCR result corresponding to Sanger Sequencing (undiluted template, n=80)	ASB-PCR result corresponding to Sanger Sequencing (1:10 template dilution, n=80)
Wild type	**52**	**51**	**51**
BRAF c.1799T>A	**14**	**13**	**14**
– heterozygous	14	With the ASB-PCR no distinction between homo- and heterozygous mutation is possible	
– homozygous	0		
BRAF c.1798_1799GT>AA	**14**	**14**	**14**
– heterozygous	6	With the ASB-PCR no distinction between homo- and heterozygous mutation is possible	
– homozygous	8		

4.6.6 Incidence of the *BRAF* mutation in patient MM16-423

The aim of testing the tumour cells isolated from different tissue entities of patient MM16-423 was to investigate the sequence of mutational events happening in the *BRAF* gene. We isolated DCCs from the SLN and from 6 LNs after LAD during the patient's progress of disease (**Figure 8**). In addition, we isolated CTCs from a leukapheresis and single cells from the DCC-derived cell line of the patient (**Table 21**). The DCCD of the LNs was ranging from 2 to 900000. We tested between 1 and 18 tumour cells (n) per tissue type by ASB-PCR, detecting variable numbers of *BRAF* mutations. Although we tested only cells with a good DNA quality (GII ≥2), we observed some drop outs, which means that the *BRAF* gene could not be amplified (**Table 21**).

Table 21: Overview of the samples from patient MM16-423 tested by ASB-PCR for the BRAF mutations c.1799T>A and c.1798_1799GT>AA.

SLN: sentinel lymph node; LAD: lymphadenectomy; CTCs: circulating tumour cells.

Tissue type	DCCD	n	Mutant	Wild type	Drop out
SLN	2	3	0	3	0
LAD LN1	900000	14	13	0	1
LAD LN2	50000	11	10	0	1
LAD LN3	10	1	1	0	0
LAD LN4	58	17	12	3	2
LAD LN5	42	3	1	0	2
LAD LN6	138	13	12	0	1
DCC-derived cell line		18	18	0	0
CTCs		15	12	0	3

After performing the ASB-PCR on the DCCs and tumour cells, collected at different time points of the patient's disease progress, we calculated the *BRAF* mutation rate, which is the percentage of cells detected with the *BRAF* c.1799T>A or c.1798_1799GT>AA mutation per successfully tested cells. The PT and DCCs isolated from the SLN removed right after diagnosis did not carry any of the tested *BRAF* hotspot mutations (**Figure 37**, brown and blue circles). Three years later, when the patient had a LN recurrence, multiple LNs were removed during a LAD. We received 6 of those LNs and isolated DCCs, which were then tested for the *BRAF* mutation by the ASB-PCR. The *BRAF* mutation rate was 80% for LN 4 (LN4) and 100% for LN 1 (LN), LN 2 (LN2), LN 3 (LN3), LN 5 (LN5) and LN 6 (LN6; **Figure 37**, green circles). The DCC-derived cell line, which was generated from LN1, also showed a mutation rate of 100% (**Figure 37**, black circle). CTCs isolated from the patient's blood after leukapheresis one year later also carried the *BRAF* hotspot mutation in 100% of investigated samples (**Figure 37**, red circle). The LNs with a DCCD above 100 (LN1, LN2 and LN6) showed a mutation rate

of 100%, while in LN4 with a DCCD of 58 not all DCCs harboured a *BRAF* mutation (mutation rate of 80%). LN3 and LN5 had a mutation rate of 100%, although their DCCD was below 100.

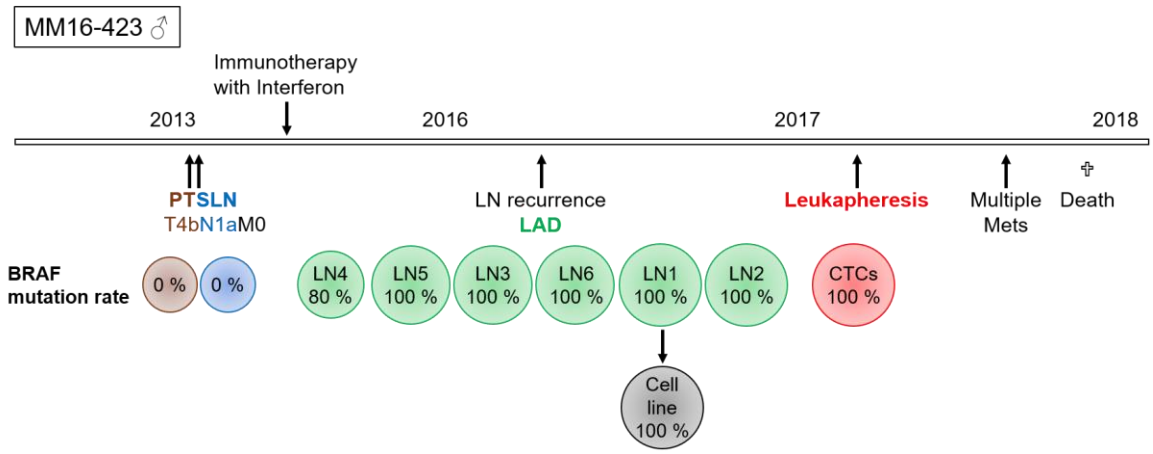


Figure 37: Progression of disease and *BRAF* mutation rate of tumour cells from patient MM16-423

While the primary tumour (brown circle) and the sentinel LN (blue circle), resected right after primary diagnosis did not have any *BRAF* mutation, the LNs (LN) removed during lymphadenectomy (green circles) had a *BRAF* mutation rate between 80% and 100%. The DCC-derived cell line (black circle) generated from LN1 also carried a *BRAF* mutation in 100% of investigated samples. The CTCs (red circle), isolated one year later from a leukapheresis showed a *BRAF* mutation rate of 100%.

4.6.7 Application of the ASB-PCR for the detection of *BRAF* mutations to a larger cohort of melanoma patients

The ASB-PCR assay was primarily established to facilitate fast, easy and cost-effective screening for most common hotspot mutations in the *BRAF* gene (*BRAF* c.1799T>A and c.1798_1799GT>AA) in DCCs isolated from melanoma patients. Therefore, I tested a larger cohort of melanoma patients for the *BRAF* mutation. In 2014, we developed a method for detection of DCCs in LN of melanoma patients and demonstrated its superior performance over conventional histopathology in a cohort of 1000 patients (Ulmer et al. 2014). Moreover, we showed previously that *BRAF* oncogenic mutation is part of the colonization signature of genetic alterations acquired by DCCs within the LN during the time of colony formation and associated with poorer clinical outcome of melanoma patients (Werner-Klein et al. 2018). We now combined our detection method of DCCs in LN of melanoma patients with our new ASB-PCR assay to assess the prevalence of *BRAF* mutations in DCCs in clinically LN positive compared to negative patients. Between 2008 and 2018, we examined LNs of 641 malignant melanoma patients from the University Hospitals in Regensburg and Tübingen. In 512 patients the pathologist did not find any involvement of the LNs (N0), while in 129 patients an invasion of the LN was detected (N+). Gp100-positive DCCs were detected in 136 and 95 patients with N0 and N+ LNs, respectively. The isotype control of 23 N0 LNs and 8 N+ patients was positive; therefore, the patients could not be evaluated. Subsequently, 719 DCCs from 86 patients with a pathologically negative LN status and 64 patients with a histologically positive LN status were evaluated either by ASB-PCR or by both ASB-PCR and Sanger sequencing in parallel (**Figure 38**).

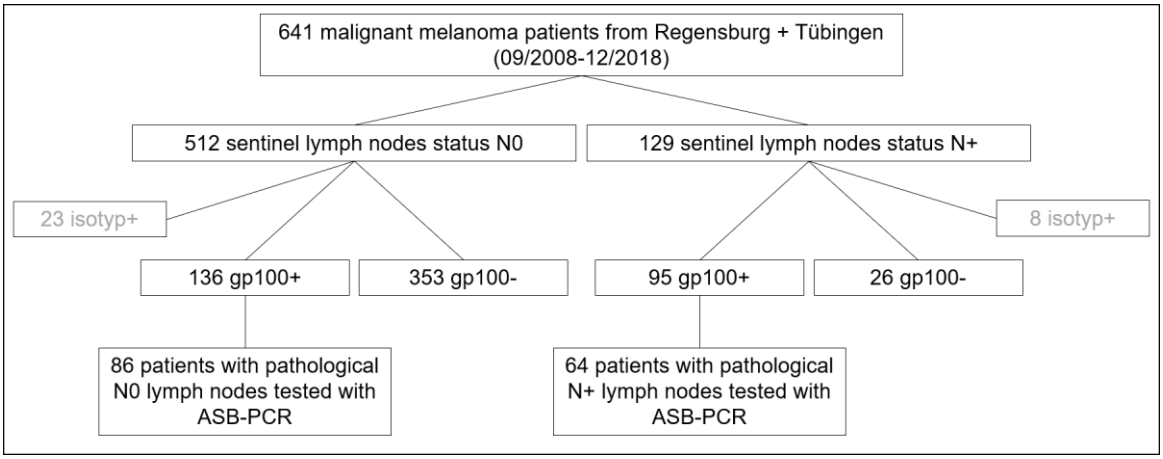


Figure 38: Overview of the patient cohort for the evaluation of the BRAF status of DCCs

Between 2008 and 2018, 641 LNs from malignant melanoma patients, recruited at the University Hospitals in Regensburg and Tübingen were examined. 512 LNs were negative (N0) according to conventional histopathology, while 129 LNs were affected (N+). Gp100-positive DCCs were found in 136 and 95 patients with a N0 stage LN and a N+ stage LN, respectively. Of these, 86 patients with N0 LNs and 64 patients with N+ LNs were tested for *BRAF* mutations.

298 cells were isolated from LN negative patients (pathologically staged as N0). In this group, *BRAF* c.1799T>A or c.1798_1799GT>AA mutations were found in 20.8% of the DCCs. In LN positive (N+) patients we detected and processed 421 DCCs. Within this group, *BRAF* mutations were found in 56.8% of the DCCs (**Figure 39, A**). Collectively, *BRAF* mutations were found in 19.8% of N0 and 59.4% of N+ patients (**Figure 39, B** and **0**). In summary, significantly more *BRAF* mutations were found in cells (chi-square, n=719, p=0.000) and patients (chi-square, n=150, p=0.000) from and with a pathologically positive LN (N+).

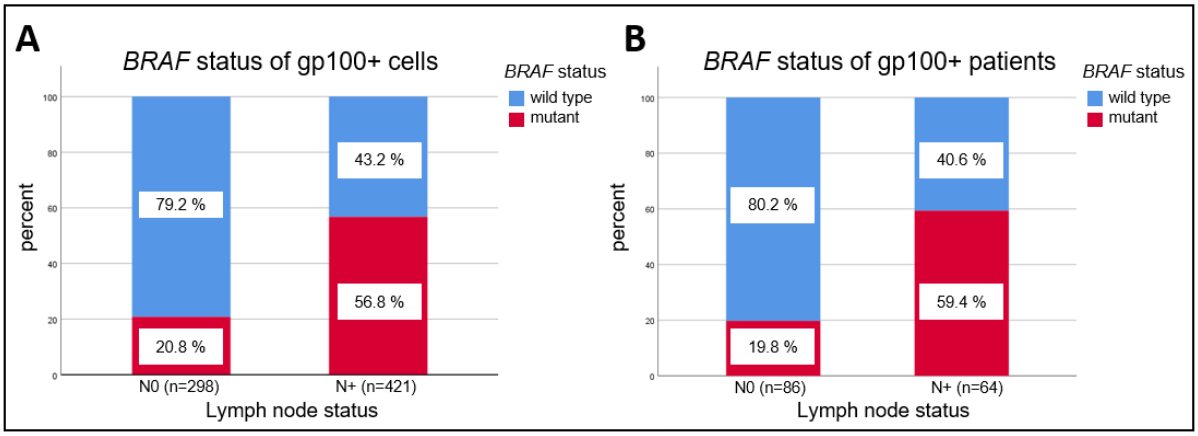


Figure 39: BRAF mutation status of DCCs isolated from malignant melanoma patients with N0 and N+ staged LNs

(A) *BRAF* mutation status of DCCs isolated from LNs stated as negative (N0) by a pathologist was compared to the mutation status of DCCs isolated from affected LNs (N+). 20.8% of DCCs isolated from N0 LNs harboured the *BRAF* c.1799T>A or the c.1798_1799 GT>AA mutation, while 56.8% of DCCs isolated from affected LNs were mutant by ASB-PCR, Sanger sequencing or both methods. **(B)** Of the patients with a pathologically negative LN 19.8% had the *BRAF* c.1799T>A or the c.1798_1799 GT>AA mutation, while 59.4% patients with a positive LN harboured a *BRAF* mutation.

Next, for each LN the DCC density (DCCD) was calculated which is defined as number of DCCs per one million of mononuclear cells in each LN. Previously we established that colonization, understood as formation of micrometastatic colonies by DCCs in LNs, was associated with a DCCD of around 100 in most patients, whereas DCCD values below 100 are indicative for early dissemination (Werner-Klein et al. 2018). Nevertheless, early colonization could also occur in LN with a DCCD below 100. Now we wanted to examine how DCCD values correlate with the occurrence of oncogenic mutations in the *BRAF* gene. Using a DCCD of 100 as a threshold, I identified 352 DCCs in LNs with a DCCD<100 and 367 DCC were detected in LNs with a DCCD≥100. Within these groups, oncogenic *BRAF* mutations were detected in 25.9% and 57.2% of DCCs, respectively (**Figure 40, A;** chi-square, n=719, p=0.000). Collectively, *BRAF* mutations were found in 22.7% of patients with a DCCD<100 and 62.3% of patients with a DCCD≥100 (**Figure 40, B** and **0**; chi-square, n=150, p=0.000). In summary, *BRAF* mutations were found in significantly less DCCs and patients with a DCCD<100.

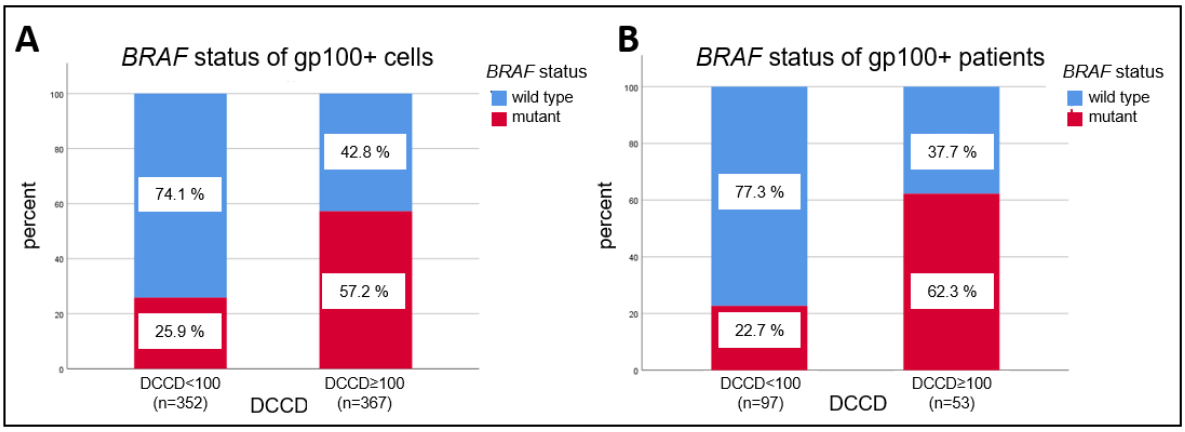


Figure 40: *BRAF* mutation status of DCCs isolated from malignant melanoma patients with DCCD<100 and DCCD≥100

(A) 25.9% of DCCs isolated from LNs with a DCCD<100 had a *BRAF* mutation. DCCs isolated from LNs with a DCCD≥100 were mutated in 57.2% of the cases. **(B)** 22.7% of patients with a DCCD<100 had at least one DCCs with a *BRAF* mutation, while 62.3% of patients with a DCCD≥100 harboured one of the tested *BRAF* mutations.

For a closer look at the proportion of *BRAF* mutations in relation to the DCCD, we have defined additional DCCD groups. In the groups with a DCCD>0≤1 and a DCCD>1≤10 15.1% and 14.9% of patients had a *BRAF* mutation, respectively (**Figure 41**). In the group of patients with a DCCD>10≤30 the proportion of patients with a *BRAF* mutation is significantly higher with 62.5%, compared to the group with a DCCD>1≤10 (chi-square, n=55, p=0.003). In the patient groups with a DCCD>30≤100 and a DCCD>100≤1000 the percentage is 55.5% and 60.9% (**Figure 41**). 83.3% of patients with a DCCD>1000≤10000 and 50.0% with a DCCD>10000≤100000 harbour a *BRAF* mutation (**Figure 41**).

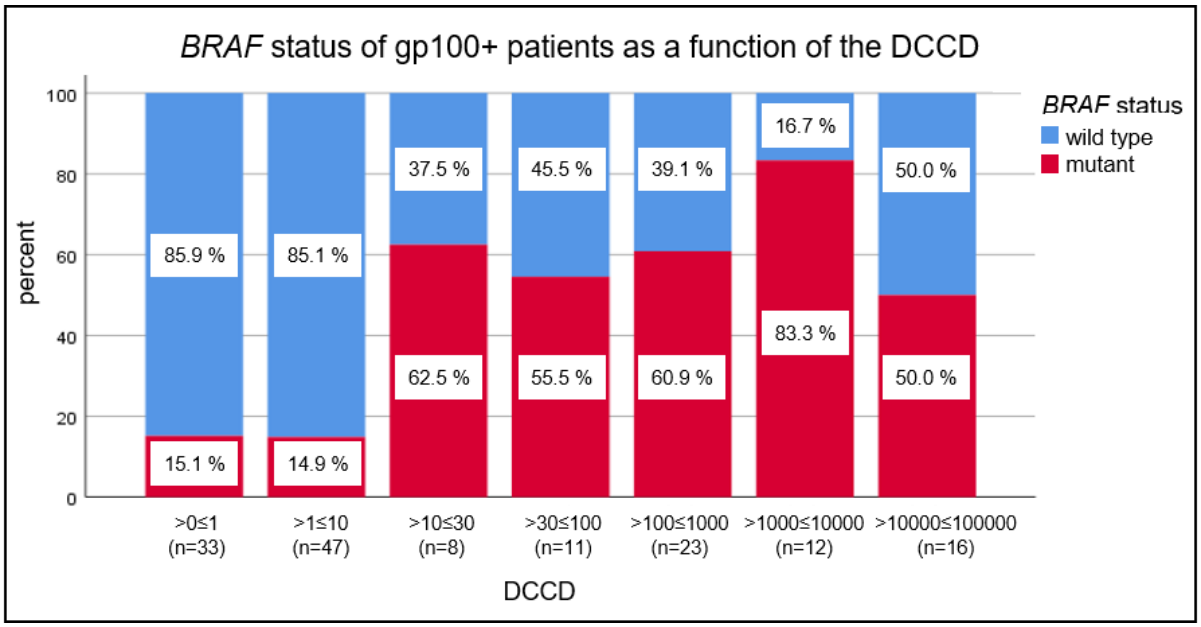


Figure 41: *BRAF* mutation status of malignant melanoma patients as a function of the DCCD
Patients with a DCCD>0≤1 and a DCCD>1≤10 had *BRAF* mutations in 15.1% and 14.9% of cases. Patients with a DCCD>10≤30 and a DCCD>30≤100 showed *BRAF* mutations in 62.5% and 55.5%. In 60.9%, 83.3% and 50.0% patients with a DCCD>100≤1000, a DCCD>1000 ≤10000 and a DCCD>10000≤100000 harboured a *BRAF* mutation.

5 Discussion

The parallel progression model predicts that disseminating cancer cells (DCCs) leave the primary tumour (PT) at an early stage. Mutation and adaption to the ectopic environment at the metastatic site happens in parallel to the formation of a PT. The model is supported by several patient studies showing that DCCs can disseminate in early stages of tumour progression, that dissemination does not increase with tumour size and that metastases and the PT display a striking genetic disparity (Brastianos et al. 2015; Stoecklein and Klein 2009; Werner-Klein et al. 2018). Therefore, we aimed to focus on the genetic alterations of early DCCs. Based on recent findings, colonisation defined as formation of a micrometastasis, was observed in most patients with a DCCD (DCC-density; number of DCCs per million mononuclear cells) around 100 while most LNs with a DCCD below 100 were negative by standard histopathology, representing early dissemination without colony formation but scattered single cells (Werner-Klein et al. 2018). Furthermore, Sebastian Scheitler observed in his thesis phenotypical small MCSP-positive DCCs in LNs with a low DCCD (median DCCD of 4), large MCSP-positive DCCs in LNs with a higher DCCD (median DCCD of 60) and LN with both phenotypes with a medium DCCD (median DCCD of 15) (Scheitler 2013). These observations led us to the hypothesis that small MCSP-positive DCCs are the precursors of large MCSP-positive DCCs and represent very early DCCs. This hypothesis should be tested by cell lineage tree reconstruction. Furthermore, we aimed to investigate the incidence of the two most common driver mutations in *BRAF* in melanoma patients and its association with the DCCD in LNs. We hypothesize that the acquisition of *BRAF* mutations marks the transition from pre-colonizing DCCs to colonizing DCCs and hence a significant progression step in systemic cancer development.

5.1 Patient cohort

To test our hypothesis if small MCSP-positive DCCs are the precursors of large MCSP-positive patients with both phenotypes in the LN were selected. In 51 of 521 patients both phenotypes were detected in the LN. This means that 9.8% of patients were detected with both small and large MCSP-positive DCCs, which is comparable with the observation of 6.9% of Sebastian Scheitler in 2013 (Scheitler 2013). The second criterion for the patient selection was an existing DCC-derived cell line from the patient. Though exceeding the scope of my project, the initial plan with the DCC-derived cell line was to use it for high-throughput drug screening assays. Sequencing of early disseminating DCCs could reveal unique genetic mutations that represent new targets for personalized therapies (Brastianos et al. 2015; Orgaz and Sanz-Moreno 2013). These targets could then be tested on high-throughput drug screening assays on the patient cell line in cooperation with the Fraunhofer Institute for Toxicology and Experimental Medicine. Since the overall aim of this thesis was to study early dissemination in malignant melanoma, the third criterion was to choose patients with a DCCD below 100. The median DCCD in mixed samples was 15 (Scheitler 2013) but for the generation of a DCC-derived cancer cell line a larger number of DCCs were generally needed. Therefore, the number of available DCC-derived cell lines was low with only 6 DCC-derived cell lines being generated from patients with both small and large MCSP-positive cells. Including all these criteria, two patients were left, MM15-127 and MM16-394. In addition, we received metastatic tissue from patient MM15-127. This opened up the possibility to compare the early DCCs with progressed tumour cells of the metastasis.

Using the allele-specific PCR with a blocking reagent (ASB-PCR), we aimed to have a closer look at the incidence of the two most common *BRAF* mutations in a specific patient with DCCs isolated from different anatomic sites. Therefore, we compared DCCs from the sentinel LN and non-sentinel LNs during different times of disease progression with CTCs from blood and tumour cells from metastases. All these criteria applied to two patients, whereas one patient didn't harbour any *BRAF* mutations and was consequently excluded.

5.2 Adaptation of the lineage tree analysis for single cells of melanoma patients

To the best of our knowledge, a STR-based lineage tree analysis has never been done before with DCCs of malignant melanoma patients. Therefore, protocols had to be established or adapted.

5.2.1 Establishment of the staining for the isolation of non-tumour control cells

Non-tumour control cells were needed for the calculation of the root that anchors the beginning of the cell lineage tree. Furthermore, non-tumour control cells were included in the lineage tree analysis as ground truth in order to see whether it is possible to separate the different cell types. We selected non-tumour outgroup cells that can be found in the LN tissue (Jalkanen and Salmi 2020). This had the advantage, that no additional sample collection from the patients was necessary, but already existing samples could be re-stained and analysed. As non-tumour control cells, T-cells and macrophages from the haematopoietic lineage and endothelial cells from the epithelial lineage were included. To save valuable patient material, the detection of T-cells and macrophages was established as a double-staining to isolate both cell types from the same slide. The successful establishment of the double-staining was shown on LN slides by the fluorescent staining of CD3-positive T-cells and CD68-positive macrophages and no double-positive cells (**Figure 14**). By the reverse combination of each primary antibody (anti-CD3 and anti-CD68) against T-cells and macrophages with the fluorescent secondary antibody, a cross-reaction was excluded (**Figure 13**). Unfortunately, the macrophages were accidentally isolated from another melanoma patients and had therefore to be excluded from the lineage tree analysis. To have a third non-tumour outgroup for the analysis of patient MM15-127, we included bulk gDNA that we isolated from lymph node tissue (**Table 16**). Next, staining of endothelial cells with the anti-CD31 antibody in combination with the AP-polymer anti-mouse solution was established on LN samples (**Figure 16**). No fluorescent double staining was used here, as only one cell type had to be detected.

The copy number variant (CNV) analysis of the non-tumour control cells revealed one of 25 aberrant T-cells and two of 14 aberrant endothelial cells. A staining artefact, resulting in the CD3- or CD31-staining of tumour cells could not be excluded. Besides, tumour cells could undergo morphological mimicry to adapt to the ectopic environment of the LN. Vascular mimicry was observed in aggressive primary and metastatic melanoma, where tumour cells reconstituted the vascular channels in human tumour tissue and expressed endothelial cell markers (Maniotis et al. 1999). However, the CNV profiles of the aberrant non-tumour control cells showed no common aberrations with other tumour cells (**Figure 25**). Usually, CNV are closely associated with

malignant transformation (Shaikh 2017). Therefore, it is currently unknown whether massive genomic rearrangements may also occur in T-cells or endothelial cells. Although all selected non-tumour control cells had a genomic integrity index (GII; for detailed explanation see **3.9.4**) > 2, I cannot exclude that these aberrant CNV profiles could result from poor DNA quality.

5.2.2 Optimization of the sample preparation and the sequencing depth

Before the lineage tree analysis was started, two steps had to be optimized in advance: First, the preparation of the samples in our laboratory and second, the sequencing depth of the samples performed in the Shapiro laboratory. There were several options for the preparation of the single cell samples (**Figure 19**). The library preparation and subsequent sequencing of the different samples by the Shapiro laboratory showed that either double-strand synthesis of the primary WGA samples or reamplification both with or without double-strand syntheses and purification resulted in mapping rates over 90%. To save valuable patient material, we decided to perform reamplification of the primary WGA samples. Additional double-strand synthesis of the reamplified samples before purification made no difference in the mapping rate of the samples and was therefore skipped in the preparation step (**Figure 20**).

The second step that had to be optimized was the sequencing depth. Computational simulations had established that the fewer cell divisions the cells have undergone, i.e. the closer they are related to each other, the more STRs are needed to successfully distinguish the cells from each other (Spiro and Shapiro 2016). Although we do not know the exact number of cell divisions of each cell, DCCs that left the primary tumour at an early stage as well as non-tumour control cells are expected to be closer to the germline than cancer cells isolated from a tumour mass resulting from massive proliferation. Therefore, we aimed at the maximum number of loci. Consecutive sequencing revealed that the number of loci covered by over 30 reads reached a plateau of 8000 at about four to five million total reads (**Figure 21**). From this we concluded to target a sequencing depth of four to five million total reads for all of our future experiments. However, it needs to be explored, why a plateau is reached at 8000 loci given a total number of sequenced loci of about 12000 (Tao et al. 2018). For future analyses it is important to know if the loci that were not covered by over 30 reads are the same for every cell or if they are randomly distributed. During the whole genome amplification (WGA) protocol, the DNA was digested into fragments of 100 – 1500 bp (Klein et al. 1999). We therefore ask ourselves, if some of the target loci have MseI sites, that prevent complete amplification of the loci. All those aspects are currently under investigation by the Shapiro laboratory, that performs the cell lineage tree reconstruction.

5.3 STR-based cell lineage tree reconstruction

The lineage tree analysis aims to uncover the developmental history of a collection of cells based on the somatic mutations that occur naturally during cell division. Somatic mutations used for cell lineage reconstructions should not confer a selective advantage or disadvantage, otherwise a bias could be introduced during the reconstruction. The somatic mutations should be associated with DNA replication and the dynamics should be well understood for reliable modelling (Shapiro, Biezuner, and Linnarsson 2013). STRs, also known as microsatellites, are highly abundant regions with repetitive sequences of 1-6 bases that are usually embedded in non-coding DNA and are generally assumed to evolve neutrally (Baron and van Oudenaarden 2019). Additionally, they

are prone to *de novo* mutations due to slippage events during DNA replication (Willems et al. 2014; Woodworth, Girskis, and Walsh 2017). Taking all those characteristics into account, STRs are a promising mutational source to unravel the cell lineage of selected cells.

5.3.1 Separation of samples from different patients

Before constructing the whole lineage tree of the complete sample collection of the patient MM15-127, we wanted to proof two ground truths of the analysis. The first ground truth is the separation of samples obtained from different patients. To show this, we used T-cells, also used as non-tumour control cells in the complete tree, from three malignant melanoma patients and a breast cancer patient. The breast cancer patient was included in the project of Manjusha Ghosh and the two other malignant melanoma patients in Sandra Huber's project. The lineage tree analysis of T-cells from different patients resulted in a clear separation of the samples obtained from different patients (**Figure 22**).

5.3.2 Separation of non-tumour control cells of patient MM15-127

The second ground truth we aimed to proof was the separation of the non-tumour control cells within one patient. This lineage tree analysis included T-cells from the hematopoietic lineage, endothelial cells from the epidermal lineage and bulk gDNA. A separation of the three non-tumour outgroups was clearly visible, although an endothelial cell and two replicates from the bulk gDNA were switched between the bulk and endothelial cell groups (**Figure 23**). The fewer cell divisions the cells undergo, the fewer mutations accumulate in the microsatellite regions (STRs) and the more STRs are needed for correct separation of two samples (Spiro and Shapiro 2016). It could be possible that the non-tumour control cells did not undergo enough cell divisions to accumulate enough distinctive loci. Sequencing of a larger STR panel could solve the problem.

5.3.3 Separation of non-tumour control cells and tumour cells of patient MM15-127

After showing the successful separation of outgroup cells from different patients and different outgroups within one patient we had sufficient confidence in the method. In addition to the non-tumour control cells the different phenotypes of DCCs (small and large MCSP-positive, gp100-positive), the bulk DNA from the PT and the tumour cells from metastases were added to the lineage tree analysis. The lineage tree showed eight distinct clusters consisting of the 3 outgroups, small and large MCSP-positive cells, gDNA from the PT, gp100-positive cells, cell line cells and the metastatic cells (**Figure 24**). Three small MCSP-positive DCCs were located in the endothelial cluster and eight non-tumour control cells in different tumour cell clusters. As the T-cells and endothelial cells are non-tumour cells, it is possible that they did not undergo enough cell divisions to accumulate enough distinctive mutations to be clearly separated from tumour cells. However, CNV analysis of the non-tumour control cells revealed that one of the T-cells clustering in the PT cluster and two of the endothelial cells in the DCCs cluster showed an aberrant CNV profile. Although the CNV profile had no similarity to the profiles of the tumour cells, its healthy origin could not be confirmed and it was therefore excluded from the final lineage tree reconstruction. The three DCCs located among the endothelial cells were stained against the tumour marker MCSP. MCSP expression has also been described in a variety of normal tissues, like endothelial

cells, chondrocytes and certain basal keratinocytes within the epidermis (Campoli et al. 2004). Therefore, it cannot be excluded that non-tumour control cells expressed this marker or that there has been a staining artefact. Actually, the CNV analysis showed that those three small MCSP-positive DCCs had a balanced profile. Since their malignant origin could not be confirmed, they were excluded from the final lineage tree reconstruction. From that we concluded that immunocytochemistry might not be sufficient to decide whether a cell is really a tumour cell or a non-tumour control cell. Therefore, CNV analysis was performed to confirm the origin of the included cells.

5.3.4 Separation of non-tumour control cells of patient MM15-127 after CNV analysis

After CNV profiling of all cells, lineage tree analysis was again performed. This time, aberrant non-tumour control cells, non-tumour control cells with poor CNV profile quality and cells that did not deliver enough STR regions for the lineage tree analysis to place them correctly in the tree were excluded (**Table 19**). Following a repeated reconstruction with the remaining cells, a clear separation between the bulk gDNA from the LN, T-cells and endothelial cells was visible (**Figure 26**). There was still one mislocated bulk DNA sample and one endothelial cell (arrows in **Figure 26**). As already explained above (**5.3.2**), they may not have undergone enough cell divisions to have acquired enough distinctive loci.

To study the robustness and variability of this tree, the bootstrap method was applied (explained in more detail in **4.5.9**). In the cell lineage tree reconstruction with the non-tumour control cells, all branches separating the different groups of non-tumour control cells as well as the two subbranches containing the mislocated endothelial cell and the bulk DNA samples had a TBE ≥ 0.7. Therefore, the separation of these samples based on the mutations detected in the STR loci can be assumed to be robust (Lemoine et al. 2018).

5.3.5 Separation of non-tumour control cells and tumour cells of patient MM15-127 after CNV analysis

In the final tree reconstruction after the CNV analysis we included all non-tumour control cells with balanced genomes, and all DCCs, metastatic cells, DCC-derived cell line cells and PT gDNA showing aberrant CNV profiles. Cells not included in the lineage tree did either not pass our DNA quality control, the CNV analysis, were not sufficiently covered during the sequencing or did not deliver enough STR regions to place them correctly in the tree (**Table 19**). The aim of the lineage tree analysis of patient MM15-127 was to test our hypothesis, if small MCSP-positive DCC are precursors of large MCSP-positive DCCs and therefore represent early DCCs.

Nine different clusters were observed in the tree. Three clusters consisted of non-tumour control cells (**Figure 27**, clusters 1, 2 and 8), the other six of tumour cells (**Figure 27**, clusters 3-7 and 9). The separation of each cluster had a TBE ≥ 0.7, meaning that they can be considered robust. If our hypothesis would be true, we would expect the small MCSP-positive DCCs to cluster far away from the PT, as they left at an early stage of disease acquired genomic mutations outside the PT. The large MCSP-positive DCCs would be expected to cluster close to the small-MCSP positive DCCs as they would be their descendants. The DCC-derived cell line cells, the metastatic cells

and the PT would be expected to cluster in separate clades as they would be in a more advanced metastatic stage. But in fact, two of the three observed DCC clusters contain both small and large MCSP-positive cells. Furthermore, all three clusters are mixed with DCC-derived cell line cells and gp100-positive DCCs. Although we could not confirm our hypothesis yet, we saw a clear separation of three DCC clusters. One could hypothesize that the DCC clade that clusters on the same branch as the metastatic clade and opposite of the PT cluster are early DCCs. If those DCC left the PT at an early stage, they would have accumulated distinctive mutations outside of the PT in the SLN. The other two DCC clades that cluster on the same branch as the PT could have disseminated at a later time point. If they have been in the PT for a longer period, they would have accumulated more mutations similar to the PT and therefore cluster closer to the PT cluster. Those hypotheses could be tested by the investigation of shared and unique mutations in the different clusters by whole exome sequencing.

5.4 CNV analysis to determine the origin of cells

CNV analysis was performed for all 39 non-tumour control cells and 152 tumour cells including the bulk gDNA from the PT. The aim of the CNV analysis was to improve the cell lineage analysis by only including the cells with a clear origin, meaning non-tumour control cells with a balanced profile and tumour cells with an aberrant profile (Shao et al. 2019)(**Figure 25**). Three of 39 non-tumour control cells showed aberrations in the CNV analysis (**Table 18**). These three non-tumour control cells were also mislocated in the first lineage tree reconstruction including the DCCs (**Figure 24**, black arrowhead). Two endothelial cells were mislocated in the DCCs cluster and the T-cell in the PT cluster. It could not be excluded that the two aberrant cells expressing the endothelial marker CD31 were tumour cells undergoing vasculogenic mimicry (Luo et al. 2020). Nevertheless, neither the aberrant T-cell nor the two aberrant CD31-positive cells had a similar CNV profile to the DCCs or the PT. Therefore, morphological mimicry or a staining artefact are unlikely since the the CD3-positive and the CD31-positive cells did not show any tumour specific aberrations. Other reasons for aberrations in control cells could be poor DNA quality. For that reason, we excluded all aberrant control cells from further analysis, as their non-malignant origin could not be confirmed.

From the 152 tumour cells included in the CNV analysis, 119 showed an aberrant CNV profile (**Table 18**). All gp100-positive cells isolated from the LN as well as the DCC-derived cell line cells, tumour cells from the metastasis and the bulk gDNA from the PT had an aberrant profile. Five of 32 small and 21 of 24 large MCSP-positive cells isolated from the LN showed an aberrant profile corresponding to 15.6% and 87.5%, respectively (**Table 18**). Significantly less small MCSP-positive cells had an aberrant genome compared to the large MCSP-positive cells (chi-square, n=54, p=0.000), while there was no significant difference between the number of aberrant large MCSP-positive cells and the gp100-positive DCCs, the DCC-derived cell line cells and the tumour cells from the metastasis (chi-square, n=110, p=0.103). Balanced cells expressing the tumour marker MCSP could either have disseminated in a genomically “immature” state in which they have not yet acquired typical mutations or copy number changes (Klein 2013; Werner-Klein et al. 2018) or they could be healthy cells of the melanocyte lineage, immune cells or endothelial cells being able to express MCSP (Campoli et al. 2004). However, as the malignant origin of the balanced MCSP-positive cells could not be proven, they were also excluded from further analysis.

The CNV analysis showed the importance of a second analysis in addition to the single-marker immune staining to reliably predict the origin of a cell.

5.5 Limitations of the lineage tree analysis

The first aspect that has to be considered when performing a lineage tree analysis is a collection of sufficient sample cells. Therefore, the WGA products of the isolated non-tumour and tumour cells have to pass the DNA quality control. Next, it is highly recommended to perform a CNV analysis to confirm the malignant or non-malignant origin of the cells. The cells which passed the first two control steps were then sequenced for the construction of the lineage trees. Again, the number of cells was reduced at this step due to insufficient coverage of the STR loci in some cells. These selection criteria reduced the initial cell number by 48% (see **4.5.8**). Therefore, to apply a lineage tree analysis is only reasonable if enough DCCs are available.

The second limitation of the whole workflow is the exclusion of balanced DCCs. Compared to the other tumour cell populations, significantly more small MCSP-positive DCCs were excluded due to a balanced genome (chi-square, n=54, p=0.000). However, it is not known if this is a population of non-tumour cells or if a part of it are early disseminating DCCs before chromosomal aberration occurred. For further characterisation of small MCSP-positive cells, the expression of tumour cell and lymphatic markers could be tested. Nevertheless, to analyse the marker expression, mRNA must be isolated from the cells which was not possible using our here applied methods.

The last limitation of the cell lineage tree analysis until now is the “resolution” of the single cells within the lineage tree. Branches with a TBE value of 0.7 and higher are usually expected to be reliable, as they appear in at least 70% of the tree reconstructed on random subsets. Although the described clusters of the lineage tree of patient MM15-127 showed a TBE > 0.7 (black asterisk, **Figure 27**), most subbranches of cells within the clusters have a TBE < 0.7. Therefore, we could assume that the different clusters we observed are true, but the conclusion we could draw for the single cells is still limited.

5.6 Conclusions drawn from the lineage tree analysis

After the adaption of the lineage tree analysis to single DCCs of malignant melanoma patients, the method was tested for two ground truths. The first one was the separation of single cells from different patients and the second one the separation of different cell lineages (haematopoietic and endothelial) within one patient. Both ground truths could be confirmed. With the final lineage tree reconstruction containing all tumour and non-tumour cells of patient MM15-127, our hypothesis that small MCSP-positive DCCs are precursors of large MCSP-positive DCCs could neither be confirmed nor rejected. Nevertheless, we observed three clades of DCCs clustering in different branches of the tree. If the cluster farthest away from the PT is representing early DCCs and the clusters closer to the PT is representing DCCs that disseminated later, must still be confirmed.

5.7 ASB-PCR assay for the detection of the two most common *BRAF* mutations in melanoma patients

The ASB-PCR was established as a fast and easy method to test for the two most common *BRAF* mutations (c.1799T>A and c.1798_1799GT>AA) in DCCs of malignant melanoma patients. First, we tested DCCs from a patient, of whom we isolated DCCs from different anatomic locations at different time points of the disease. DCCs were isolated from the SLN and non-SLNs, removed during lymphadenectomy. Additionally, circulating tumour cells (CTCs) and a DCC-derived cell line was tested. The aim of this assay was to identify the incidence of the *BRAF* mutation in the different cell types during the progression of disease. Next, a larger cohort of DCCs from melanoma patients was tested for the *BRAF* mutations.

5.7.1 Sensitivity and specificity of the ASB-PCR assay compared with other methods

By retrospective testing of previously Sanger sequenced single-cell WGA products, we identified the *BRAF* c.1799T>A or c.1798_1799GT>AA mutations in 96% of samples, while the wild type allele was correctly found in 98% of samples, resulting in a sensitivity of 96% and a specificity of 98%. Nevertheless, testing of 80 single-cell WGA products (of DCCs isolated from LN of melanoma patients) did not show a single false positive or false negative result, but only resulted in two dropouts. The mutated allele was detected in WGA samples generated from cell pools containing at least 5% of mutated cells, being the limit of detection (**Figure 35**). There are two commercially available companion diagnostic tests to detect the two most common *BRAF* mutations both based on real-time PCR (RT-PCR): the cobas 4800 BRAF V600 Mutation Test and the THxID-BRAF kit. The cobas 4800 test has a sensitivity of around 95% for the *BRAF* c.1799T>A mutation, meaning that 95% of mutations are correctly identified, while 5% of mutations are missed. The specificity of the cobas 4800 test is over 98%, indicating under 2% of false positive results. The limit of detection is 5% of c.1799T>A mutated DNA. Although detection of c.1798_1799GT>AA mutation is also possible with the cobas 4800 test, it is only approved for the c.1799T>A mutation. The THxID-BRAF kit detects both the c.1799T>A and c.1798_1799GT>AA mutation with a sensitivity of 96% and 92%, respectively. The specificity of the kit is 100% and the limit of detection is 5% for both mutations (L. Cheng et al. 2018). The limit of detection of the ASB-PCR assay, as well as the sensitivity and specificity of the ASB-PCR assay is comparable to the two companion diagnostic tests, the cobas 4800 test and the THxID-BRAF kit.

Further improvements on the ASB-PCR assay by other groups show even lower limits of detection of the mutated allele, providing an even higher selectivity. Using an allele-specific RT-PCR instead of an end-point PCR reaction, detection of the mutated allele down to 1%-0.1% was achieved (Lang et al. 2011; Mostert et al. 2013; Yang et al. 2017). The high selectivity of an allele-specific PCR is usually necessary, as the assay is often performed on tumour tissue or FFPE specimen with a low number of cancer cells (Szankasi et al., 2013; Suciu et al., 2016). For our application, a limit of detection of 5% of the end-point ASB-PCR is satisfying as the single cell DNA is amplified using our WGA protocol (see **3.9**). The cell is either homozygous wild type, homozygous mutant or heterozygous, meaning that a mutant cell carries at least one mutant allele that constitutes 50% of mutant DNA.

5.7.2 ASB-PCR analysis of patient MM16-423 for the identification of *BRAF* mutations during disease progression

The patient MM16-423 was tested to investigate the incidence of the *BRAF* mutation during different time points of disease progression. DCCs were isolated from the SLN that was removed after primary diagnosis. Additional DCCs were isolated from six non-SLN after lymphadenectomy due to a LN recurrence three years after primary diagnosis and CTCs after leukapheresis one year later. While the PT and the DCCs of the SLN displayed no *BRAF* mutation, the DCCs isolated from the non-SLNs as well as the DCC-derived cell line and the CTCs were 80-100% *BRAF* mutant (**Figure 37**). From this observation we conclude that the DCCs left the PT without a *BRAF* mutation and spread to the LNs. The SLN was removed at a time as the DCCs within it did not acquired a *BRAF* mutation. However, at the time the non-SLNs were removed, most DCCs acquired a *BRAF* mutation. Three of the six tested non-SLNs had a DCCD > 100 (138 – 900000), a value where most of the LNs had pathologically detected (micro)metastasis. This is consistent with the observation that the *BRAF* mutation is part of the colonisation signature of genetic alterations acquired by DCCs within the lymph node during the time of colony formation (Werner-Klein et al. 2018). The other three non-SLNs harbouring a *BRAF* mutation had a DCCD < 100 (10 – 58). Testing a larger cohort of DCCs isolated from LNs of melanoma patients we observed, that *BRAF* mutations were already acquired in significantly more frequent in LNs with a DCCD about 30 compared to LNs with a DCCD ≤ 10 (**Figure 41**). Based on this finding, it is possible that the *BRAF* mutation was acquired even before colony formation.

5.7.3 Application of the ASB-PCR for the detection of *BRAF* mutations to a larger cohort of melanoma patients

Testing a larger cohort of melanoma patients for the incidence of the *BRAF* mutations revealed that they were found in 19.8% of patients with a pathologically negative LN (N0) and in 59.4% with a LN in which the pathologist detected (micro)metastasis (N+; **Figure 39**). The separation of LNs with a DCCD<100 and a DCCD≥100 was based on the observation that micrometastasis formation occurs in LNs with a DCCD around 100 (Werner-Klein et al. 2018). The incidence of the *BRAF* mutation in LN with a DCCD<100 was 22.7% and 62.3% for a DCCD≥100. The frequency of the *BRAF* mutation is comparable for both classifications, either by LN status or by DCCD. This indirectly confirms the observation, that in most LNs with a DCCD<100 no metastasis was detected by the pathologist, while LNs with a DCCD≥100 harboured in most cases detectable micrometastases. A more detailed look at the incidence of the *BRAF* mutation as a function of the DCCD showed that the proportion of *BRAF* mutated patients significantly rises already at a DCCD>10≤30 to 62.5% compared to the group with a DCCD>1≤10 and a *BRAF* incidence of 14.9%. These observations indicated that the *BRAF* mutation occurs before the formation of metastatic colonies. Mutations in *BRAF* are oncogenic driver mutations involved in cellular proliferation. Since early DCCs without *BRAF* mutations harbour often numerous CNAs and since *BRAF* mutation alone it is considered insufficient to induce proliferation but needs accompanying mutations (L. Cheng et al. 2018), we conclude that acquisition of *BRAF* mutations marks the onset of rapid proliferation leading to colony formation.

Tyrosine kinase inhibitors like Vemurafenib or Dabrafenib are selective inhibitors of the V600-mutated BRAF kinase and treatment of patients with *BRAF* c.1799T>A positive melanoma was shown to improve survival (Miller and Flaherty 2014). Combination treatments with BRAF and MEK inhibitors have shown even better response and a significant delay of acquired resistance (L. Cheng et al. 2018; Long, Hauschild, et al. 2017). Based on these findings, the National Comprehensive Cancer Network recommends *BRAF* mutation testing for patients with stage III at high risk for recurrence and stage IV malignant melanoma (National Comprehensive Cancer Network 2019). Due to the discrepancy of the *BRAF* status among primary and metastatic lesions in melanoma, it is recommended to test metastatic tissue for *BRAF* mutations (Bradish et al. 2015). Patients without nodal metastasis in stage I or II are usually not tested for *BRAF* mutations. However, DCCs were detected in 26.1% of those patients (**Figure 38**) and testing of DCCs with the ASB-PCR assay has shown that 19.8% of histologically N0 patients harboured a *BRAF* mutation (**Figure 39**). Twenty-nine per cent of N0 patients had a relapse within 15-49 months (median: 25 months). In the future, these patients could benefit from an adjuvant therapy with *BRAF* inhibitors if they were tested for the *BRAF* mutations.

5.7.4 Limitations of the ASB-PCR analysis

The ASB-PCR analysis as well as the other available companion diagnostic tests are limited to specific *BRAF* mutations. In contrast, genome wide sequencing is not limited to specific mutations. The Sanger sequencing method has a sensitivity of 92-98% to reliably detect mutations, while false positive results can be excluded with a specificity of 100%. Alternatively, the high-resolution melt assay is a combination of PCR and sequencing by DNA synthesis. Its sensitivity and specificity are 98-100% (L. Cheng et al. 2018). However, the disadvantage of both the Sanger sequencing and the high-resolution melt assay is the higher technical, financial and time effort compared to the ASB-PCR.

As the single-cell DNA undergoes WGA-based amplification there is a possibility of allelic dropout, which is a loss of one allele during PCR amplification of DNA (Stevens at al., 2017). The allelic drop-out (ADO) rate for *BRAF* was recently evaluated with single cells from cell lines, xenografts and primary tumours ranging between 0% and 8% for the mutant and the wild type allele, respectively (Werner-Klein et al. 2018). Furthermore, we tested the selectivity of our assay on WGA products prepared from cell mixtures comprising 5-80% of mutated SKMEL28 mixed with PBMCs. The mutated *BRAF* allele was detected in all samples containing as little as 10% of SKMEL28 cells. Six single-cell WGA products harbouring a heterozygous c.1798_1799GT>AA mutation and 15 with a heterozygous c.1799T>A mutation were previously subjected to Sanger sequencing. Our multiplex ASB-PCR confirmed the heterozygous mutation in all the 21 single-cell genomes amplified by WGA. Therefore, we conclude that the multiplex ASB-PCR assay is a suitable method for the detection of heterozygous *BRAF* mutations in WGA samples as no case of allelic dropout was observed.

5.7.5 Conclusions drawn from the ASB-PCR analysis

The aim of the ASB-PCR assay was to establish a fast and easy assay to investigate the incidence of the two most common *BRAF* mutations in melanoma patients on the single-cell level. The reliability and accuracy of the assay can keep up with commercially available companion

diagnostic tests. The limit of detection of the ASB-PCR assays established here is not as high as it is when using RT-PCR. However, when testing single cells, the frequency of the mutated allele is higher and a level of detection under 1% is not needed. A significant rise of the incidence of *BRAF* mutations was already observed in LNs with a DCCD>10≤30, indicating that *BRAF* mutations were acquired early before they started to proliferate. Early stage melanoma patients without LN metastasis could already benefit from adjuvant treatment when tested for *BRAF* mutations. The ASB-PCR could easily be integrated in the diagnostic workflow of a laboratory, where DCCs from malignant melanoma patients are routinely isolated and amplified by WGA.

6 Bibliography

Auerbach, Robert et al. 1987. “Specificity of Adhesion between Murine Tumor Cells and Capillary Endothelium: An in Vitro Correlate of Preferential Metastasis in Vivo.” *Cancer Research* 47(6).

Balch, Charles M. et al. 2000. “Long-Term Results of a Multi-Institutional Randomized Trial Comparing Prognostic Factors and Surgical Results for Intermediate Thickness Melanomas (1.0 to 4.0 Mm).” *Annals of Surgical Oncology* 7(2): 87–97. https://pubmed.ncbi.nlm.nih.gov/10761786/ (August 19, 2020).

———. 2009. “Final Version of 2009 AJCC Melanoma Staging and Classification.” *Journal of Clinical Oncology* 27(36): 6199–6206. /pmc/articles/PMC2793035/?report=abstract (August 20, 2020).

Barbano, Raffaela et al. 2015. “Competitive Allele-Specific TaqMan PCR (Cast-PCR) Is a Sensitive , Specific and Fast Method for BRAF V600 Mutation Detection in Melanoma Patients.” (November): 1–11.

Baron, Chloé S., and Alexander van Oudenaarden. 2019. “Unravelling Cellular Relationships during Development and Regeneration Using Genetic Lineage Tracing.” *Nature Reviews Molecular Cell Biology* 20(12): 753–65. www.nature.com/nrm (September 7, 2020).

Biezuner, Tamir et al. 2016. “A Generic , Cost-Effective , and Scalable Cell Lineage Analysis Platform.” *Genome biology* 26: 1588–99.

Bradish, Joshua R. et al. 2015. “Discordancy in BRAF Mutations among Primary and Metastatic Melanoma Lesions: Clinical Implications for Targeted Therapy.” *Modern Pathology* 28(4): 480–86.

Brastianos, Priscilla K. et al. 2015. “Genomic Characterization of Brain Metastases Reveals Branched Evolution and Potential Therapeutic Targets.” *Cancer Discovery* 5(11): 1164–77. https://pubmed.ncbi.nlm.nih.gov/26410082/ (September 10, 2020).

Burotto, Mauricio, Victoria L. Chiou, Jung Min Lee, and Elise C. Kohn. 2014. “The MAPK Pathway across Different Malignancies: A New Perspective.” *Cancer* 120(22): 3446–56. https://pubmed.ncbi.nlm.nih.gov/24948110/ (September 7, 2020).

Campoli, Michael R et al. 2004. “Human High Molecular Weight-Melanoma-Associated Antigen (HMW-MAA): A Melanoma Cell Surface Chondroitin Sulfate Proteoglycan (MSCP) with Biological and Clinical Significance.” *Critical reviews in immunology* 24(4): 267–96. http://www.ncbi.nlm.nih.gov/pubmed/15588226 (August 4, 2020).

Chambers, Ann F., Alan C. Groom, and Ian C. MacDonald. 2002. “Dissemination and Growth of Cancer Cells in Metastatic Sites.” *Nature Reviews Cancer* 2(8): 563–72. https://www.nature.com/articles/nrc865 (September 1, 2020).

Cheng, Changming et al. 2013. “Detection of Rare Point Mutation via Allele-Specific Amplification in Emulsion PCR.” *BMB Reports* 46(5): 270–75. /pmc/articles/PMC4133893/?report=abstract (September 7, 2020).

Cheng, Liang et al. 2018. “Molecular Testing for BRAF Mutations to Inform Melanoma Treatment Decisions: A Move toward Precision Medicine.” *Modern Pathology* 31(1): 24–38. http://dx.doi.org/10.1038/modpathol.2017.104.

Cheung, Kevin J. et al. 2016. “Polyclonal Breast Cancer Metastases Arise from Collective Dissemination of Keratin 14-Expressing Tumor Cell Clusters.” *Proceedings of the National Academy of Sciences of the United States of America* 113(7): E854–63. https://www.pnas.org/content/113/7/E854 (September 1, 2020).

Collins, V P, R K Loeffler, and H Tivey. 1956. “Observations on Growth Rates of Human Tumors.” *Am J Roentgenol Radium Ther Nucl Med* 76(5): 988–1000. https://pubmed.ncbi.nlm.nih.gov/13362715/ (September 2, 2020).

Colombino, Maria et al. 2012. “BRAF/NRAS Mutation Frequencies among Primary Tumors and Metastases in Patients with Melanoma.” *Journal of Clinical Oncology* 30(20): 2522–29.

Curtin, John A. et al. 2005. “Distinct Sets of Genetic Alterations in Melanoma.” *New England Journal of Medicine* 353(20): 2135–47. http://www.nejm.org/doi/abs/10.1056/NEJMoa050092 (August 21, 2020).

Davis, Lauren E., Sara C. Shalin, and Alan J. Tackett. 2019. “Current State of Melanoma Diagnosis and Treatment.” *Cancer Biology and Therapy* 20(11): 1366–79. https://www.tandfonline.com/doi/abs/10.1080/15384047.2019.1640032 (August 19, 2020).

Eggermont, Alexander M.M. et al. 2018. “Adjuvant Pembrolizumab versus Placebo in Resected Stage III Melanoma.” *New England Journal of Medicine* 378(19): 1789–1801. https://pubmed.ncbi.nlm.nih.gov/29658430/ (August 26, 2020).

Engel, J. et al. 2003. “The Process of Metastasisation for Breast Cancer.” *European Journal of Cancer* 39(12): 1794–1806. http://www.ejcancer.com/article/S0959804903004222/fulltext (September 3, 2020).

Ewing, James. 1928. “Neoplastic Diseases: A Treatise on Tumours. Third Edition.” *British Journal of Surgery* 16(61): 77–89. http://doi.wiley.com/10.1002/bjs.1800166126 (September 2, 2020).

Faries, Mark et al. 2017. “Completion Dissection or Observation for Sentinel-Node Metastasis in Melanoma.” *New England Journal of Medicine* 376(23): 2211–22.

Foulds, L. 1954. “The Experimental Study of Tumor Progression: A Review.” *Cancer Research* 14(5).

Friberg, Sten, and Stefan Mattson. 1997. 65 Journal of Surgical Oncology *On the Growth Rates of Human Malignant Tumors: Implications for Medical Decision Making*.

Friedl, Peter, Joseph Locker, Erik Sahai, and Jeffrey E. Segall. 2012. “Classifying Collective Cancer Cell Invasion.” *Nature Cell Biology* 14(8): 777–83. https://www.nature.com/articles/ncb2548 (September 1, 2020).

Frumkin, Dan et al. 2008. “Cell Lineage Analysis of a Mouse Tumor.” *Cancer Research* 68(14): 5924–31.

Garbe, Claus et al. 2011. “Systematic Review of Medical Treatment in Melanoma: Current Status and Future Prospects.” *The Oncologist* 16(1): 5–24. https://pubmed.ncbi.nlm.nih.gov/21212434/ (August 26, 2020).

———. 2020. “European Consensus-Based Interdisciplinary Guideline for Melanoma. Part 1: Diagnostics – Update 2019.” *European Journal of Cancer* 126: 141–58. www.sciencedirect.com (August 19, 2020).

Ghosh, Manjusha S. 2020. “Identification of Metastasis Founder Cells in Breast Cancer by Cell Lineage Tracing.” University of Regensburg.

Greene, Harry S. N., and Elizabeth K. Harvey. 1964. “The Relationship between the Dissemination of Tumor Cells and the Distribution of Metastases.” *Cancer Research* 24(5).

Grob, J.J., and J.J. Bonerandi. 1998. “The ‘Ugly Duckling’’ Sign: Identification of the Common Characteristics of Nevi in an Individual as a Basis for Melanoma Screening.’” *Arch Dermatol.* 143(1). https://jamanetwork.com/journals/jamadermatology/article-abstract/188553 (August 21, 2020).

Hendry, Shona et al. 2017. “Assessing Tumor-Infiltrating Lymphocytes in Solid Tumors: A Practical Review for Pathologists and Proposal for a Standardized Method from the International Immuno-Oncology Biomarkers Working Group: Part 2: TILs in Melanoma, Gastrointestinal Tract Carcinom.” *Advances in Anatomic Pathology* 24(6): 311–35. /pmc/articles/PMC5638696/?report=abstract (September 15, 2020).

Hölzel, D., R. Eckel, and J. Engel. 2009. “Colorectal Cancer Metastasis: Frequency, Prognosis, and Consequences.” *Chirurg* 80(4): 331–40. https://link.springer.com/article/10.1007/s00104-008-1603-x (September 3, 2020).

Hüsemann, Yves et al. 2008. “Systemic Spread Is an Early Step in Breast Cancer.” *Cancer Cell* 13(1): 58–68. https://pubmed.ncbi.nlm.nih.gov/18167340/ (September 3, 2020).

Ives, Natalie J. et al. 2017. “Adjuvant Interferon-α for the Treatment of High-Risk Melanoma: An Individual Patient Data Meta-Analysis.” *European Journal of Cancer* 82: 171–83. https://www.ejcancer.com/article/S0959-8049(17)31032-8/fulltext (August 26, 2020).

Jalkanen, Sirpa, and Marko Salmi. 2020. “Lymphatic Endothelial Cells of the Lymph Node.” *Nature Reviews Immunology* 20(9): 566–78. https://doi.org/10.1038/ (September 13, 2020).

Junker, Jan Philipp et al. 2016. “Massively Parallel Clonal Analysis Using CRISPR/Cas9 Induced Genetic Scars.” *bioRxiv*: 056499. https://doi.org/10.1101/056499 (September 7, 2020).

Kang, Yibin, and Klaus Pantel. 2013. “Tumor Cell Dissemination: Emerging Biological Insights from Animal Models and Cancer Patients.” *Cancer Cell* 23(5): 573–81.

Kinsey, David L. 1960. “An Experimental Study of Preferential Metastasis.” *Cancer* 13(4): 674–76.

Klein, Christoph A. et al. 1999. “Comparative Genomic Hybridization, Loss of Heterozygosity, and DNA Sequence Analysis of Single Cells.” *Proceedings of the National Academy of Sciences of the United States of America* 96(8): 4494–99.

———. 2002. “Genetic Heterogeneity of Single Disseminated Tumour Cells in Minimal Residual Cancer.” *Lancet* 360(9334): 683–89.

———. 2009. “Parallel Progression of Primary Tumours and Metastases.” *Nature reviews Cancer* 10(2): 156.

———. 2013. “Selection and Adaptation during Metastatic Cancer Progression.” *Nature* 501(7467): 365–72. http://www.ncbi.nlm.nih.gov/pubmed/24048069%5Cnhttp://www.nature.com/doifinder/10.1038/nature12628.

Lambert, Arthur W., Diwakar R. Pattabiraman, and Robert A. Weinberg. 2017. “Emerging Biological Principles of Metastasis.” *Cell* 168(4): 670–91. /pmc/articles/PMC5308465/?report=abstract (August 26, 2020).

Lang, Alois H. et al. 2011. “Optimized Allele-Specific Real-Time PCR Assays for the Detection of Common Mutations in KRAS and BRAF.” *Journal of Molecular Diagnostics* 13(1): 23–28. http://dx.doi.org/10.1016/j.jmoldx.2010.11.007.

Langley, Robert R., and Isaiah J. Fidler. 2011. “The Seed and Soil Hypothesis Revisited-The Role of Tumor-Stroma Interactions in Metastasis to Different Organs.” *International Journal of Cancer* 128(11): 2527–35. /pmc/articles/PMC3075088/?report=abstract (September 1, 2020).

Lemoine, F. et al. 2018. “Renewing Felsenstein’s Phylogenetic Bootstrap in the Era of Big Data.” *Nature* 556(7702): 452–56. http://www.nature.com/articles/s41586-018-0043-0 (August 5, 2020).

Leong, Stanley P.L. et al. 2011. “Cutaneous Melanoma: A Model to Study Cancer Metastasis.” *Journal of Surgical Oncology* 103(6): 538–49. http://doi.wiley.com/10.1002/jso.21816 (September 6, 2020).

Long, G. V., A. Hauschild, et al. 2017. “Adjuvant Dabrafenib plus Trametinib in Stage III BRAF-Mutated Melanoma.” *New England Journal of Medicine* 377(19): 1813–23. https://pubmed.ncbi.nlm.nih.gov/28891408/ (August 26, 2020).

Long, G. V., K. T. Flaherty, et al. 2017. “Dabrafenib plus Trametinib versus Dabrafenib Monotherapy in Patients with Metastatic BRAF V600E/ K-Mutant Melanoma: Long-Term Survival and Safety Analysis of a Phase 3 Study.” *Annals of Oncology* 28(7): 1631–39. http://www.annalsofoncology.org/article/S0923753419322707/fulltext (September 7, 2020).

Luo, Qingxi et al. 2020. “Vasculogenic Mimicry in Carcinogenesis and Clinical Applications.” *Journal of Hematology and Oncology* 13(1): 1–15.

Luzzi, Keith J. et al. 1998. “Multistep Nature of Metastatic Inefficiency: Dormancy of Solitary Cells after Successful Extravasation and Limited Survival of Early Micrometastases.” *American Journal of Pathology* 153(3): 865–73.

Maniotis, Andrew J. et al. 1999. “Vascular Channel Formation by Human Melanoma Cells in Vivo and in Vitro: Vasculogenic Mimicry.” *The American Journal of Pathology* 155(3): 739. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1866899/ (August 4, 2020).

Miller, David M., and Keith T. Flaherty. 2014. “Cyclin-Dependent Kinases as Therapeutic Targets in Melanoma.” *Pigment Cell and Melanoma Research* 27(3): 351–65.

Mocellin, Simone, Sandro Pasquali, Carlo R. Rossi, and Donato Nitti. 2010. “Interferon Alpha Adjuvant Therapy in Patients with High-Risk Melanoma: A Systematic Review and Meta-Analysis.” *Journal of the National Cancer Institute* 102(7): 493–501. https://pubmed.ncbi.nlm.nih.gov/20179267/ (August 26, 2020).

Morlan, John, Joffre Baker, and Dominick Sinicropi. 2009. “Mutation Detection by Real-Time PCR : A Simple , Robust and Highly Selective Method.” *PLoS One* 4(2).

Morton, Donald L. et al. 2014. “Final Trial Report of Sentinel-Node Biopsy versus Nodal Observation in Melanoma.” *New England Journal of Medicine* 370(7): 599–609. http://www.nejm.org/doi/10.1056/NEJMoa1310460 (August 25, 2020).

Mostert, Bianca et al. 2013. “KRAS and BRAF Mutation Status in Circulating Colorectal Tumor Cells and Their Correlation with Primary and Metastatic Tumor Tissue.” *International Journal of Cancer* 133(1): 130–41.

Mühlbauer, Franziska. 2017. “Development of Allele-Specific PCR Assays for Detection of Hot-Spot Mutations in Whole Genome Amplification Products.” University of Regensburg.

National Comprehensive Cancer Network. 2019. “NCCN Clinical Practice Guidelines in Oncology - Cutaneous Melanoma.” : 757–59.

“NCI Dictionary of Cancer Terms.” 2020. https://www.cancer.gov/publications/dictionaries/cancer-terms (August 20, 2020).

Nieto, M. Angela, Ruby Yun Y.J. Huang, Rebecca A A. Jackson, and Jean Paul P. Thiery. 2016. “EMT: 2016.” *Cell* 166(1): 21–45.

Orgaz, Jose L., and Victoria Sanz-Moreno. 2013. “Emerging Molecular Targets in Melanoma Invasion and Metastasis.” *Pigment Cell & Melanoma Research* 26(1): 39–57. http://doi.wiley.com/10.1111/pcmr.12041 (September 2, 2020).

Paget, Stephen. 1889. “THE DISTRIBUTION OF SECONDARY GROWTHS IN CANCER OF THE BREAST.” *The Lancet* 133(3421): 571–73.

Patel, J. K., M. S. Didolkar, J. W. Pickren, and R. H. Moore. 1978. “Metastatic Pattern of Malignant Melanoma. A Study of 216 Autopsy Cases.” *The American Journal of Surgery* 135(6): 807–10.

Pearlman, Ross, Mary Katherine Montes de Oca, Harish Chandra Pal, and Farrukh Afaq. 2017. “Potential Therapeutic Targets of Epithelial-Mesenchymal Transition in Melanoma.” *Cancer Letters* 176(5): 139–48.

Polzer, Bernhard et al. 2014. “Molecular Profiling of Single Circulating Tumor Cells with Diagnostic Intention.” 6(11): 1371–86.

Price, Morgan N., Paramvir S. Dehal, and Adam P. Arkin. 2010. “FastTree 2 - Approximately Maximum-Likelihood Trees for Large Alignments.” *PLoS ONE* 5(3): e9490. www.plosone.org (August 19, 2020).

Raz, Ofir et al. 2019. “Short Tandem Repeat Stutter Model Inferred from Direct Measurement of in Vitro Stutter Noise.” *Nucleic Acids Research* 47(5): 2436–45. https://academic.oup.com/nar/article/47/5/2436/5304323 (August 19, 2020).

S3-Leitlinie Zur Diagnostik, Therapie Und Nachsorge Des Melanoms. 2019. Version 3. https://www.awmf.org/uploads/tx_szleitlinien/032-024OLl_S3_Melanom-Diagnostik-Therapie-Nachsorge_2020-01.pdf%0Ahttps://www.awmf.org/uploads/tx_szleitlinien/032-024OL_l_S3_Melanom-Diagnostik-Therapie-Nachsorge_2018-07_01.pdf%0Ahttp://leitlinienprogramm-onk.

Scheitler, Sebastian. 2013. “Klonale Diversität Und Phänotypische Heterogenität Bei Der Progression Maligner Melanome.” Universität Regensburg.

Scolyer, Richard A. et al. 2020. “Melanoma Pathology Reporting and Staging.” *Modern Pathology* 33(1): 15–24. https://doi.org/10.1038/s41379-019-0402-x (August 21, 2020).

Sevillya, Gur, Zeev Frenkel, and Sagi Snir. 2016. “Triplet MaxCut: A New Toolkit for Rooted

Supertree" ed. Emmanuel Paradis. *Methods in Ecology and Evolution* 7(11): 1359–65. https://onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12606 (August 3, 2020).

Shaikh, Tamim H. 2017. "Copy Number Variation Disorders." *Current Genetic Medicine Reports* 5(4): 183–90.

Shain, a Hunter, and Boris C Bastian. 2016. "From Melanocytes to Melanomas." *Nature reviews. Cancer* 16(6): 345–58. http://www.ncbi.nlm.nih.gov/pubmed/27125352.

Shao, Xin et al. 2019. "Copy Number Variation Is Highly Correlated with Differential Gene Expression: A Pan-Cancer Study." *BMC Medical Genetics* 20(1): 175. https://bmcmedgenet.biomedcentral.com/articles/10.1186/s12881-019-0909-5 (September 28, 2020).

Shapiro, Ehud, Tamir Biezuner, and Sten Linnarsson. 2013. "Singel-Cell Sequencing-Based Technologies Will Revolutionize Whole-Organism Science." *Nature Reviews Genetics* 14(9): 618–30. http://dx.doi.org/10.1038/nrg3542.

Shlush, Liran I. et al. 2012. "Cell Lineage Analysis of Acute Leukemia Relapse Uncovers the Role of Replication-Rate Heterogeneity and Microsatellite Instability." *Blood* 120(3): 603–12. https://pubmed.ncbi.nlm.nih.gov/22645183/ (September 8, 2020).

Spatz, A. et al. 2003. "Interobserver Reproducibility of Ulceration Assessment in Primary Cutaneous Melanomas." *European Journal of Cancer* 39(13): 1861–65.

Spiro, Adam, and Ehud Shapiro. 2016. "Accuracy of Answers to Cell Lineage Questions Depends on Single-Cell Genomics Data Quality and Quantity." *PLOS Computational Biology* 12(6): 1–13. http://dx.plos.org/10.1371/journal.pcbi.1004983.

Stevens, Aaron J., Millie G. Taylor, Frederick Grant Pearce, and Martin A. Kennedy. 2017. "Allelic Dropout during Polymerase Chain Reaction Due to G-Quadruplex Structures and DNA Methylation Is Widespread at Imprinted Human Loci." *G3: Genes, Genomes, Genetics* 7(3): 1019–25.

Stoecklein, Nikolas H. et al. 2002. "SCOMP Is Superior to Degenerated Oligonucleotide Primed-Polymerase Chain Reaction for Global Amplification of Minute Amounts of DNA from Microdissected Archival Tissue Samples." *The American Journal of Pathology* 161(1): 43–51. https://www.sciencedirect.com/science/article/pii/S0002944010641557 (July 31, 2020).

Stoecklein, Nikolas H., and Christoph A. Klein. 2009. "Genetic Disparity between Primary Tumours, Disseminated Tumour Cells, and Manifest Metastasis." *International Journal of Cancer* 126(3): 589–98.

Sulston, J. E., E. Schierenberg, J. G. White, and J. N. Thomson. 1983. "The Embryonic Cell Lineage of the Nematode Caenorhabditis Elegans." *Developmental Biology* 100(1): 64–119. https://pubmed.ncbi.nlm.nih.gov/6684600/ (September 7, 2020).

Talmadge, James E., and Isaiah J. Fidler. 2010. "AACR Centennial Series: The Biology of Cancer Metastasis: Historical Perspective." *Cancer Research* 70(14): 5649–69. www.aacrjournals.org (August 30, 2020).

Tao, Liming et al. 2018. "A Duplex MIPs-Based Biological-Computational Cell Lineage Discovery Platform." *bioRxiv*: 191296. https://www.biorxiv.org/content/early/2017/10/23/191296%0Ahttps://www.biorxiv.org/content/early/2017/09/25/191296.

Ubellacker, Jessalyn M. et al. 2020. “Lymph Protects Metastasizing Melanoma Cells from Ferroptosis.” *Nature* 585(7823): 113. https://doi.org/10.1038/s41586-020-2623-z (September 6, 2020).

Ugurel, Selma et al. 2017. “Survival of Patients with Advanced Metastatic Melanoma: The Impact of Novel Therapies–Update 2017.” *European Journal of Cancer* 83: 247–57.

Ulmer, Anja et al. 2004. “Immunomagnetic Enrichment , Genomic Characterization , and Prognostic Impact of Circulating Melanoma Cells.” *Clinical Cancer Research* 10: 531–37.

———. 2005. “Detection of Melanoma Cells Displaying Multiple Genomic Changes in Histopathologically Negative Sentinel Lymph Nodes.” 11(15): 5425–33.

———. 2014. “Quantitative Measurement of Melanoma Spread in Sentinel Lymph Nodes and Survival.” *PLoS Medicine* 11(2).

Vallée, Audrey et al. 2019. “Prospective Evaluation of Two Screening Methods for Molecular Testing of Metastatic Melanoma: Diagnostic Performance of BRAF V600E Immunohistochemistry and of a NRAS-BRAF Fully Automated Real-Time PCR-Based Assay.” *PLoS ONE* 14(8). https://pubmed.ncbi.nlm.nih.gov/31415669/ (September 7, 2020).

Weber, Jeffrey et al. 2017. “Adjuvant Nivolumab versus Ipilimumab in Resected Stage III or IV Melanoma.” *New England Journal of Medicine* 377(19): 1824–35. https://pubmed.ncbi.nlm.nih.gov/28891423/ (August 26, 2020).

Werner-Klein, Melanie et al. 2018. “Genetic Alterations Driving Metastatic Colony Formation Are Acquired Outside of the Primary Tumour in Melanoma.” *Nature communications* 9(1): 595.

Whiteman, David C, Adele C Green, and Catherine M Olsen. 2016. “The Growing Burden of Invasive Melanoma: Projections of Incidence Rates and Numbers of New Cases in Six Susceptible Populations through 2031.” *Journal of Investigative Dermatology* 136: 1161–71. www.jidonline.org (August 19, 2020).

Willems, Thomas et al. 2014. “The Landscape of Human STR Variation.” *Genome research* 24(11): 1894–1904. http://www.ncbi.nlm.nih.gov/pubmed/25135957 (July 30, 2020).

Woodworth, Mollie B, Kelly M Girskis, and Christopher A Walsh. 2017. “Building a Lineage from Single Cells: Genetic Techniques for Cell Lineage Tracking.” *Nature reviews. Genetics* 18(4): 230–44. http://www.ncbi.nlm.nih.gov/pubmed/28111472 (July 30, 2020).

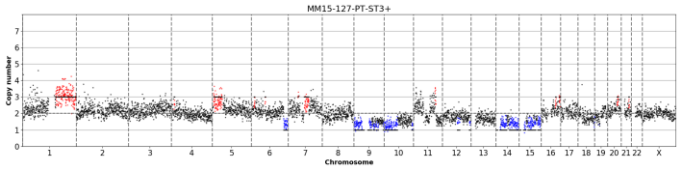
Yang, Zhao et al. 2017. “Improved Detection of BRAF V600E Using Allele-Specific PCR Coupled with External and Internal Controllers.” *Scientific Reports* 7(1): 1–12. http://dx.doi.org/10.1038/s41598-017-14140-2.

7 Appendix

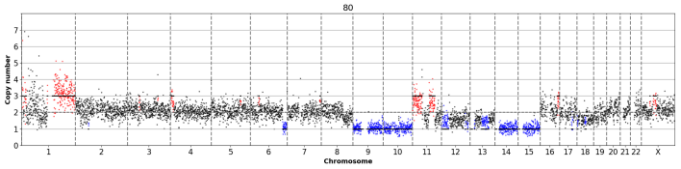
7.1 CNV profiles of samples included in the lineage tree analysis

CNV profiles from all tumour cells, non-tumour control cells and the bulk gDNA from the primary tumour included in the final lineage tree (**4.5.10**) are shown. The SRIDs are unique for each cell and can be assigned to each cell in the tree. The cell names are composed of the tissue origin, the marker used for isolation and the cell number.

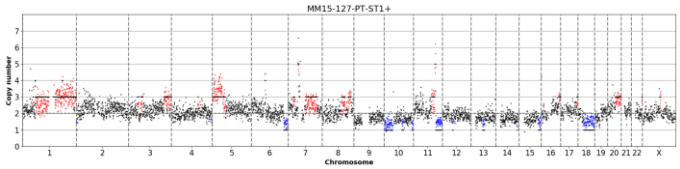
SRID40561, PTbulk3WGArepair4



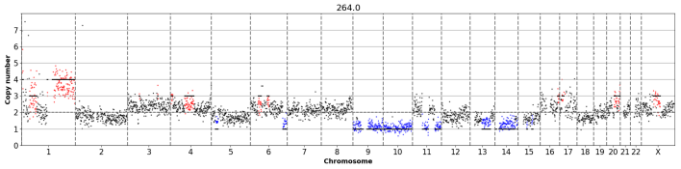
SRID40498, Met1gp100c30



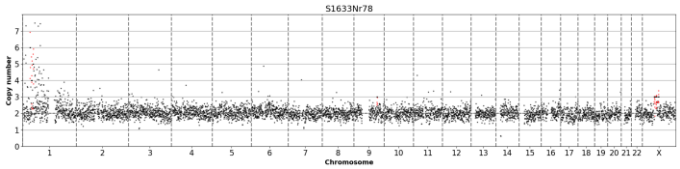
SRID40559, PTbulk1WGArepair4



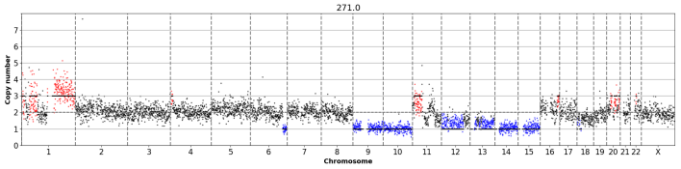
SRID40515, Met3gp100c24



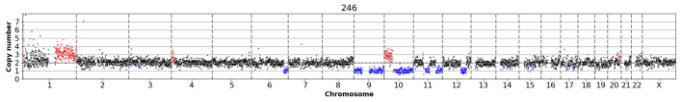
SRID40335, SLNCD3c25



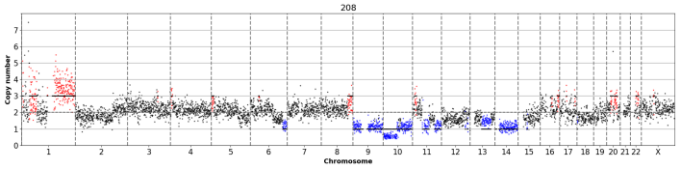
SRID40506, Met1gp100c32



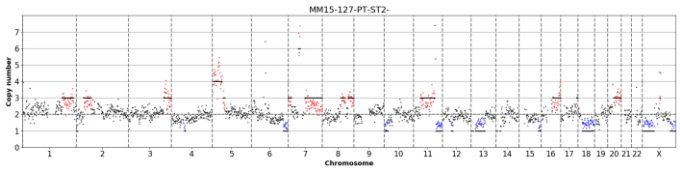
SRID40372, SLNDCClargeMCSPcryo1



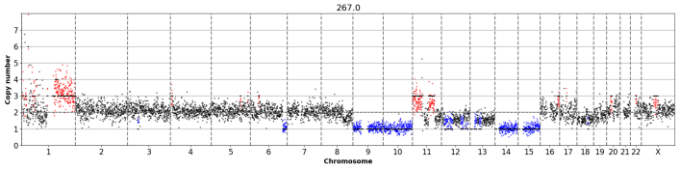
SRID40521, Met3gp100c4



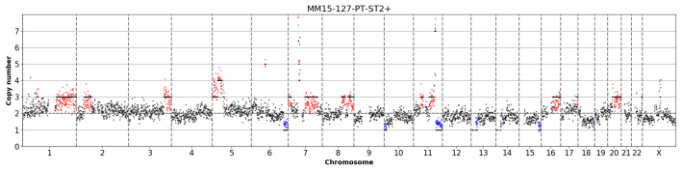
SRID40536, PTbulk2WGA1



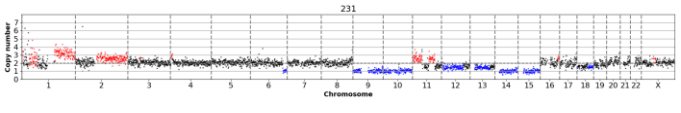
SRID40505, Met1gp100c15



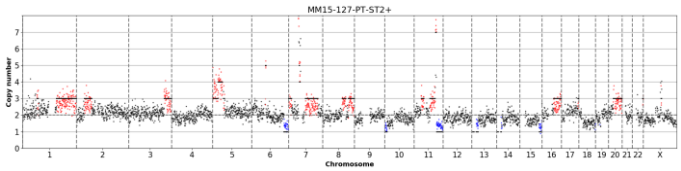
SRID40551, PTbulk2WGArepair1



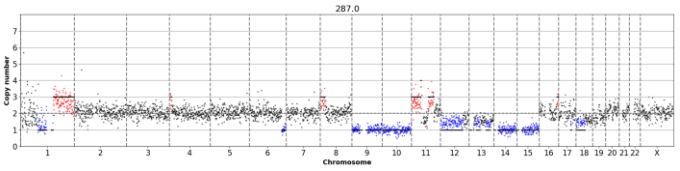
SRID40529, Met1gp100c27



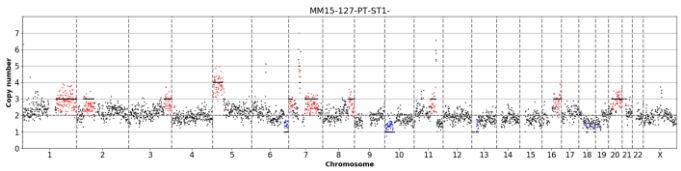
SRID40560, PTbulk2WGArepair4



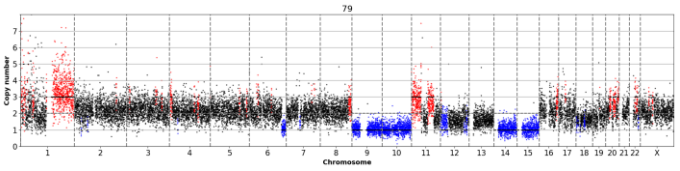
SRID40514, Met1gp100c42



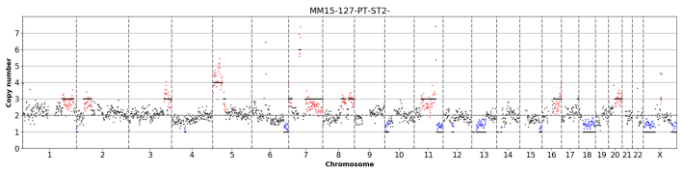
SRID40539, PTbulk1WGA2



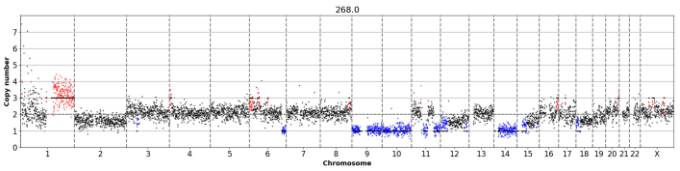
SRID40494, Met1gp100c29



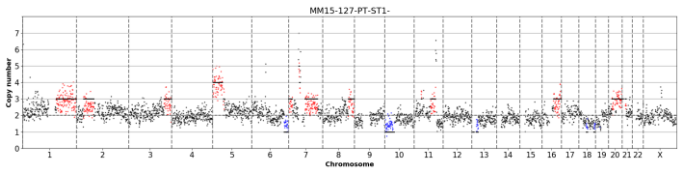
SRID40544, PTbulk2WGA3



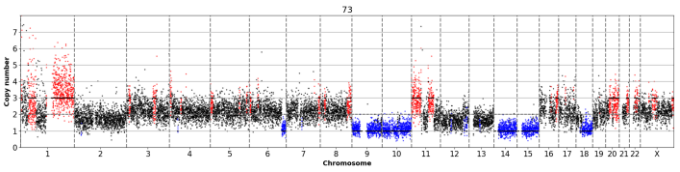
SRID40504, Met3gp100c41



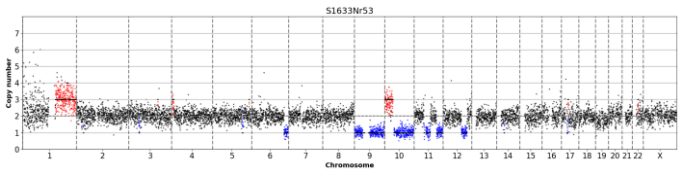
SRID40556, PTbulk1WGArepair3



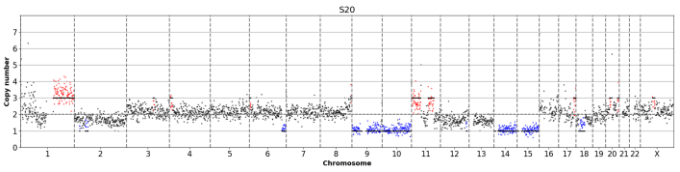
SRID40517, Met1gp100c21



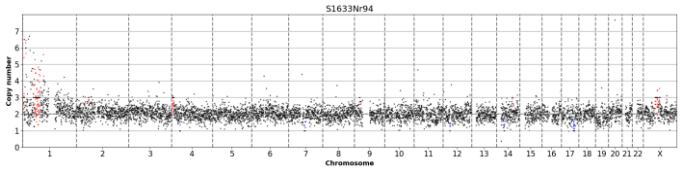
SRID40390, SLNDCCsmallMCSPcryo5



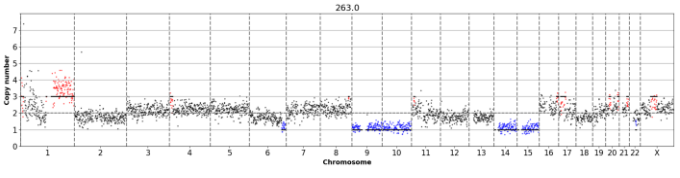
SRID40513, Met1gp100c20



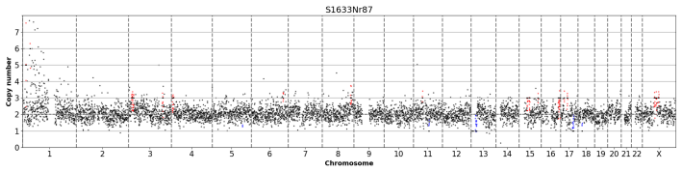
SRID40328, SLNCD3c2



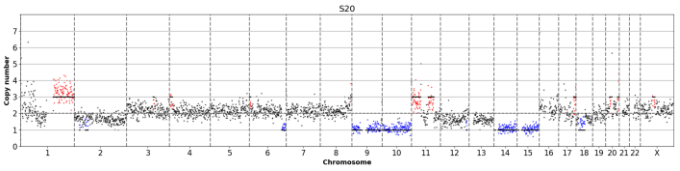
SRID40502, Met1gp100c31



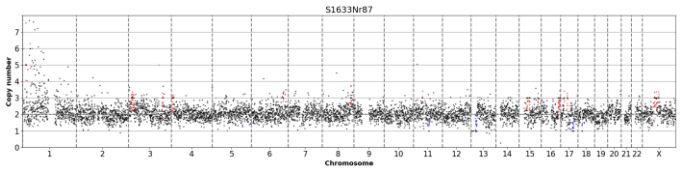
SRID40347, SLNCD3c27



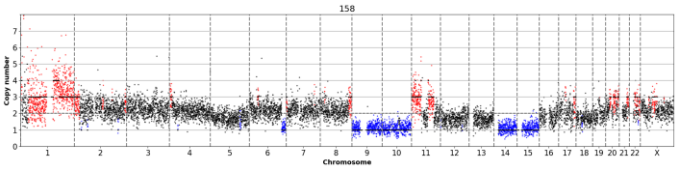
SRID40509, Met1gp100c19



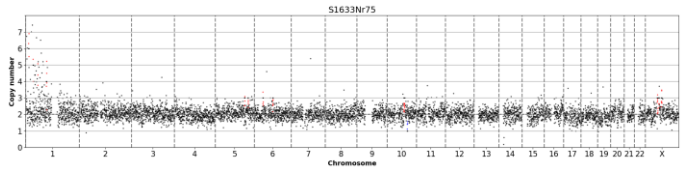
SRID40389, SLNCD3c38



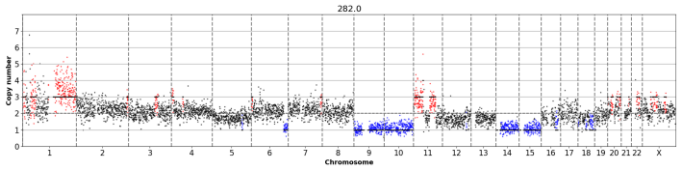
SRID40512, Met3gp100c43



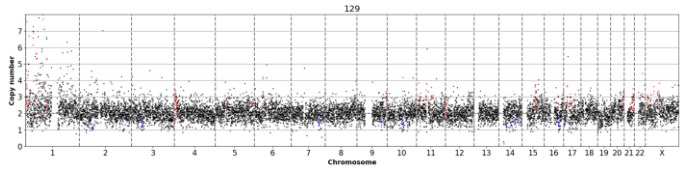
SRID40382, SLNCD3c18



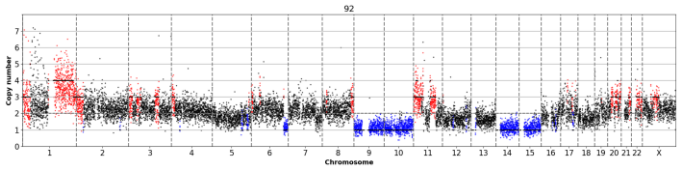
SRID40531, Met3gp100c36



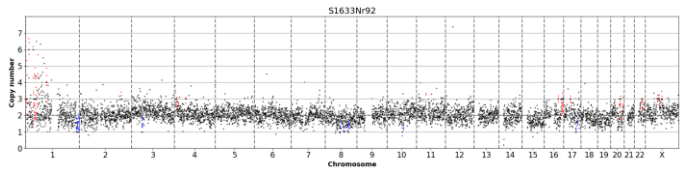
SRID40340, SLNCD3c4



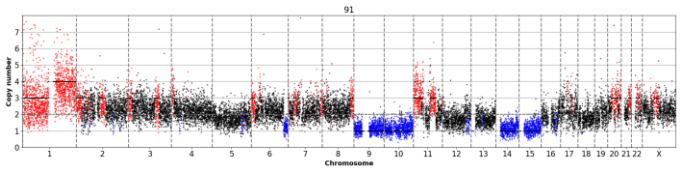
SRID40511, Met3gp100c23



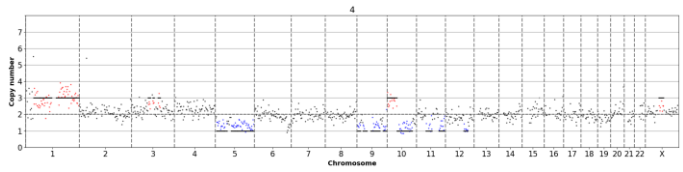
SRID40370, SLNCD3c9



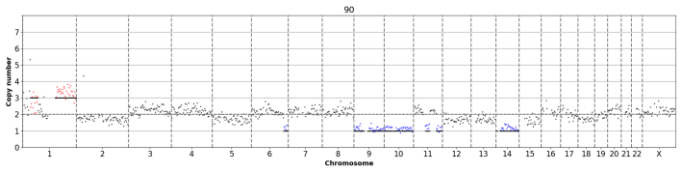
SRID40507, Met3gp100c21



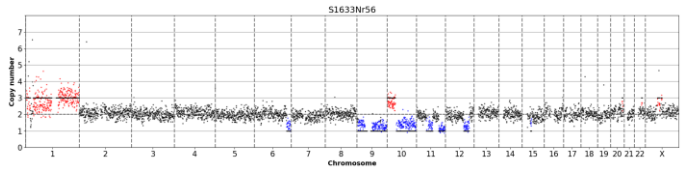
SRID40349, SLNDCCgp100c6



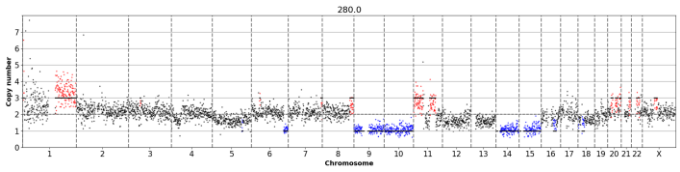
SRID40533, Met3gp100c16



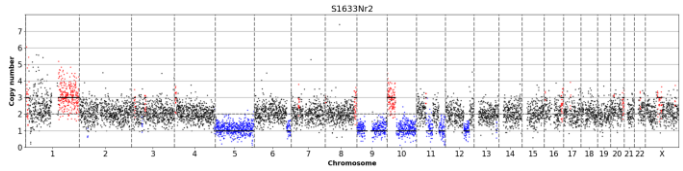
SRID40355, SLNDCCgp100c7



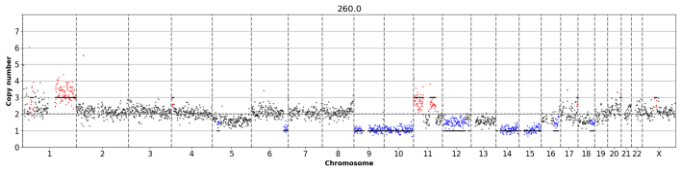
SRID40522, Met3gp100c32



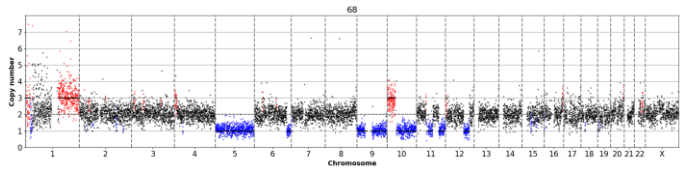
SRID40366, SLNDCClargeMCSPsn5



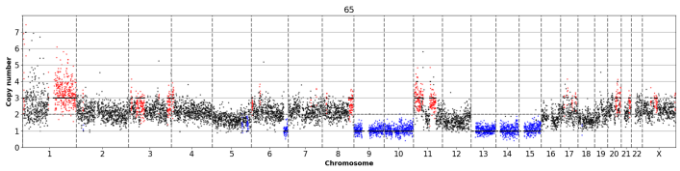
SRID40500, Met3gp100c40



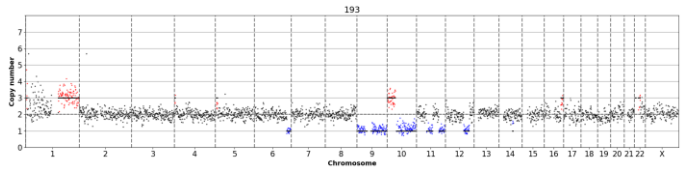
SRID40378, SLNDCClargeMCSPcryo2



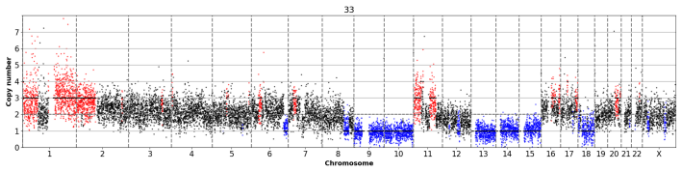
SRID40516, Met3gp100c44



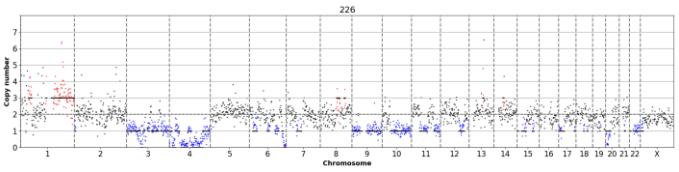
SRID40337, SLNDCCgp100c1



SRID40497, Met1gp100c7

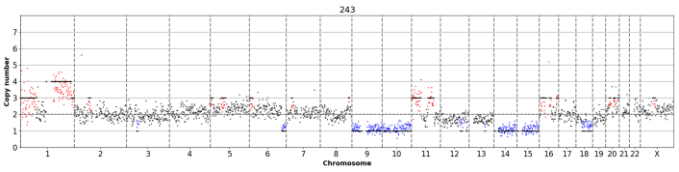


SRID40431, SLNDCCgp100c2

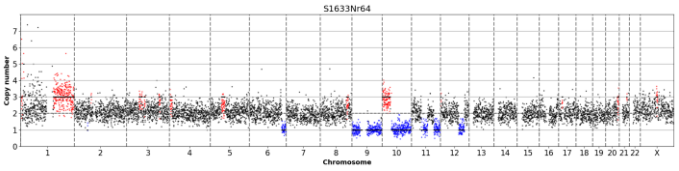


SRID40493, Met1gp100c1

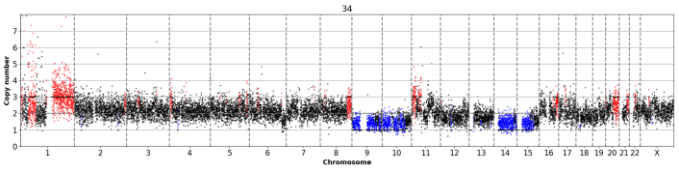
9



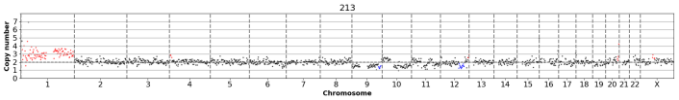
SRID40391, SLNDCClargeMCSPcryo4



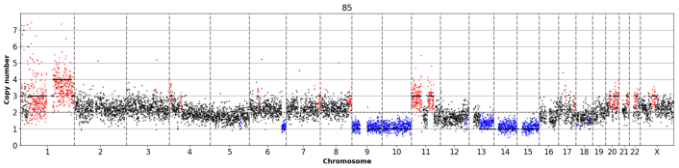
SRID40501, Met1gp100c11



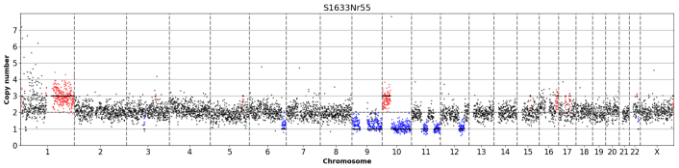
SRID40418, SLNDCClargeMCSPcryo40



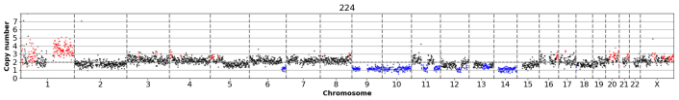
SRID40518, Met3gp100c2



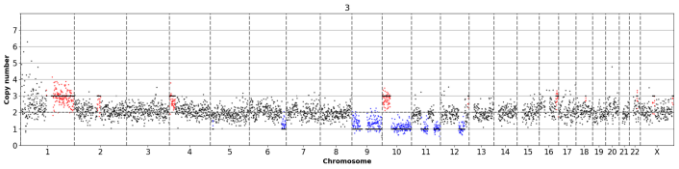
SRID40343, SLNDCCgp100c4



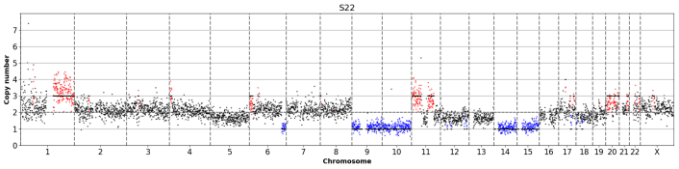
SRID40527, Met3gp100c12



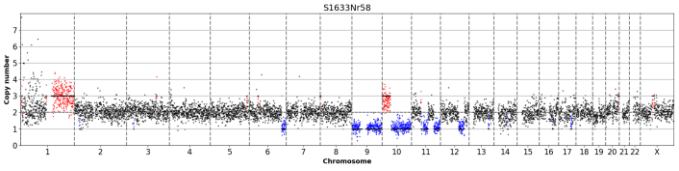
SRID40432, SLNDCCgp100c5



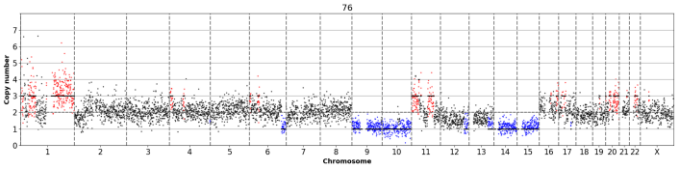
SRID40503, Met3gp100c20



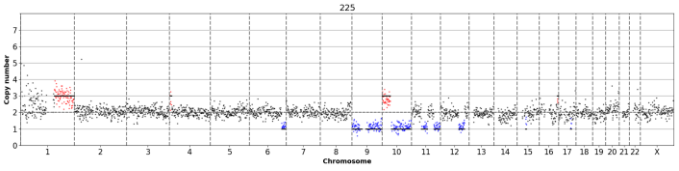
SRID40361, SLNDCCgp100c21



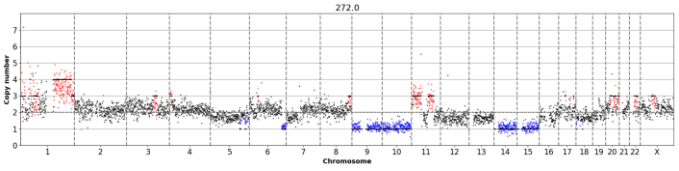
SRID40526, Met1gp100c25



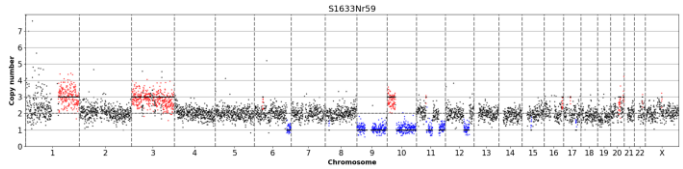
SRID40434, SLNDCCgp100c9



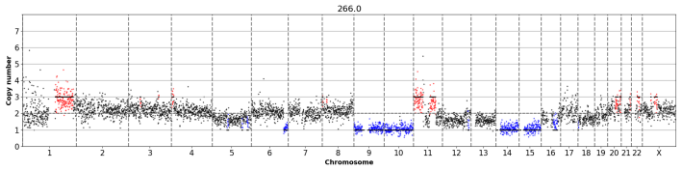
SRID40519, Met3gp100c31



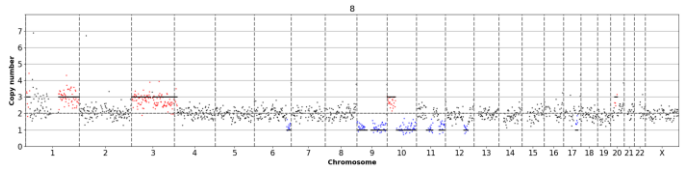
SRID40373, SLNDCCgp100c22



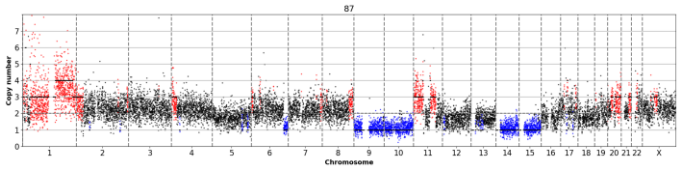
SRID40528, Met3gp100c34



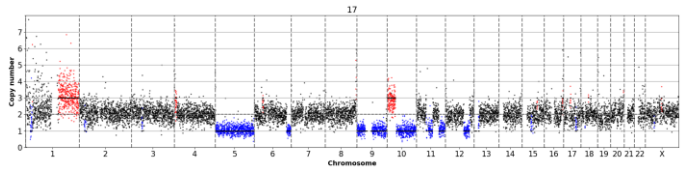
SRID40435, SLNDCCgp100c11



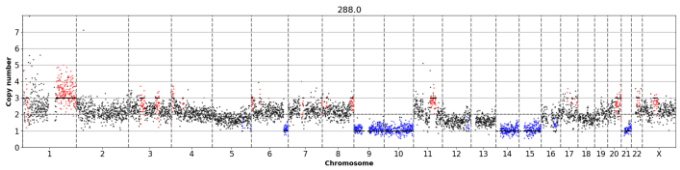
SRID40524, Met3gp100c10



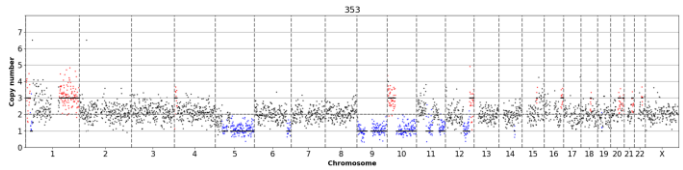
SRID40468, SLNDCCgp100c19



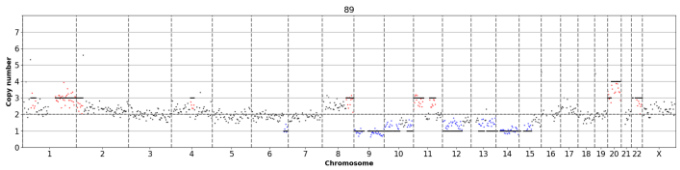
SRID40525, Met3gp100c33



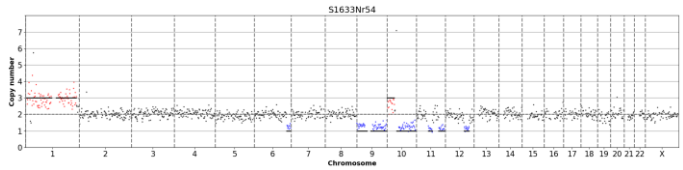
SRID40562, SLNDCCsmallMCSPcryosn6



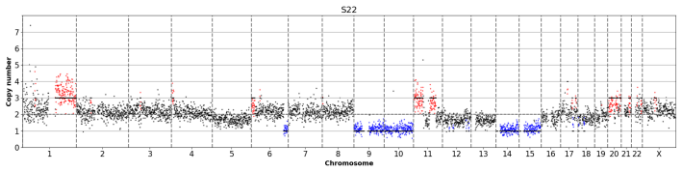
SRID40530, Met3gp100c15



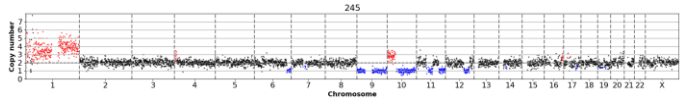
SRID40319, SLNDCCsmallMCSPcryosn4



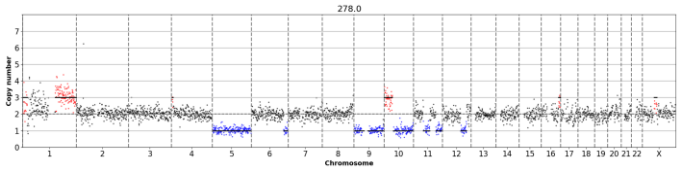
SRID40499, Met3gp100c19



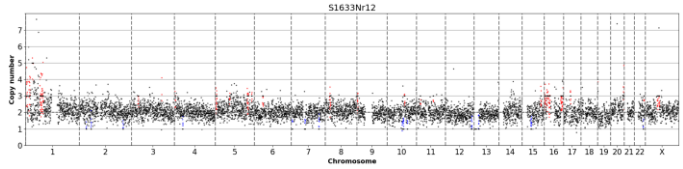
SRID40430, SLNDCClargeMCSPcryo44



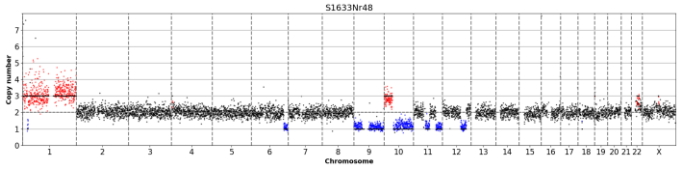
SRID40325, SLNDCClargeMCSPcryo5



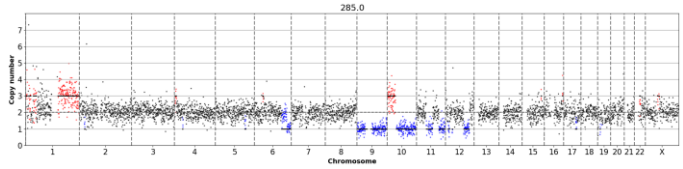
SRID40476, SLNCD31c4



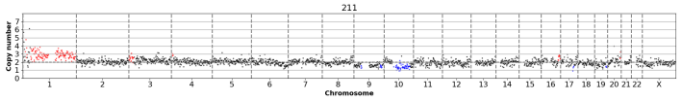
SRID40443, SLNcelllinec23



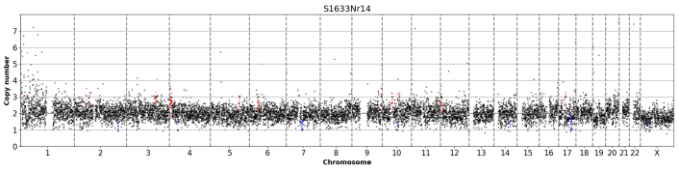
SRID40336, SLNDCClargeMCSP3



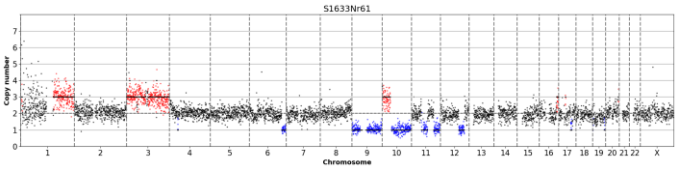
SRID40422, SLNDCClargeMCSPcryo12



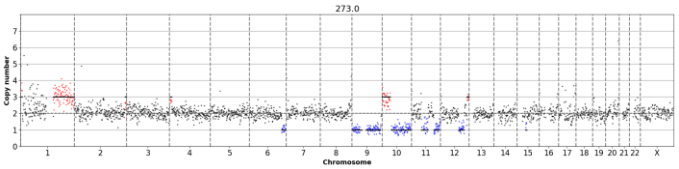
SRID40480, SLNCD31c7



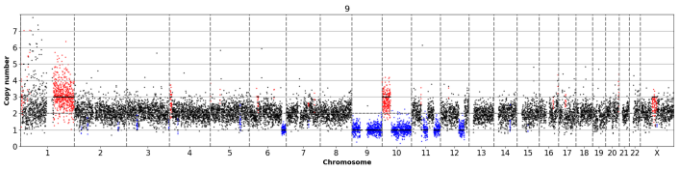
SRID40385, SLNDCCgp100c24



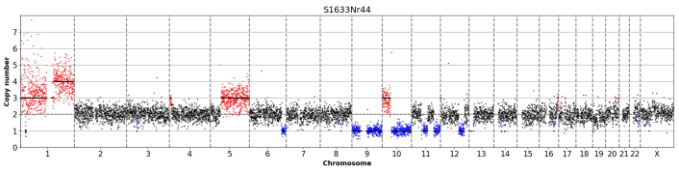
SRID40479, SLNDCCgp100c14



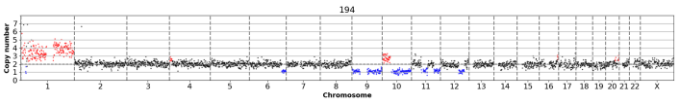
SRID40436, SLNDCCgp100c12



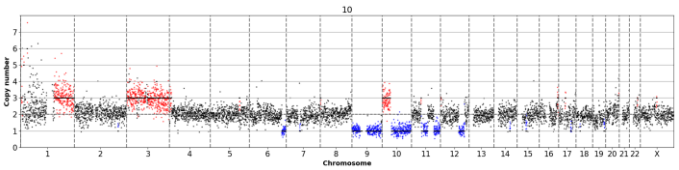
SRID40458, SLNcelllinec19



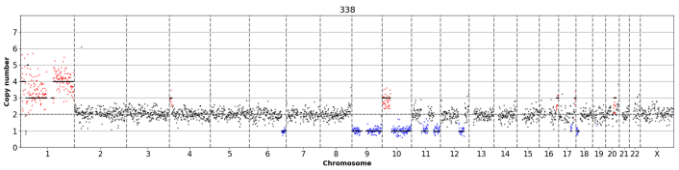
SRID40462, SLNDCCgp100c16



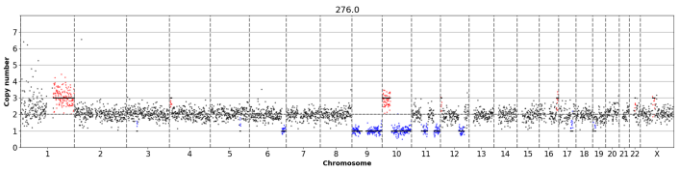
SRID40437, SLNDCCgp100c3



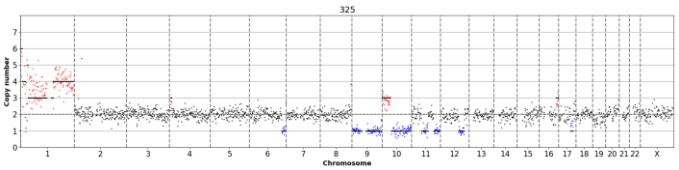
SRID40444, SLNcelllinec10



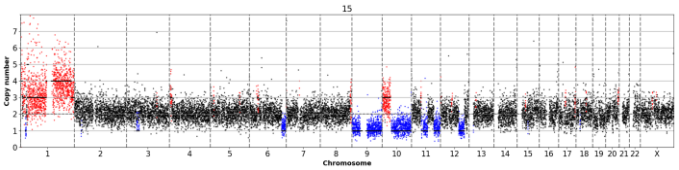
SRID40409, SLNDCClargeMCSPcryo36



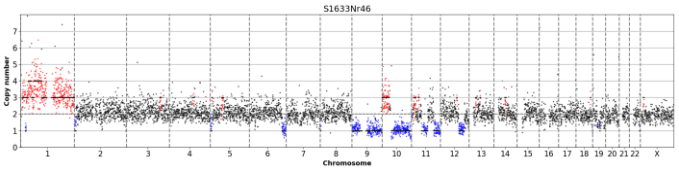
SRID40438, SLNcelllinec7



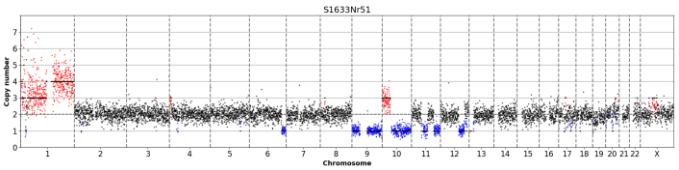
SRID40465, SLNDCCgp100c18



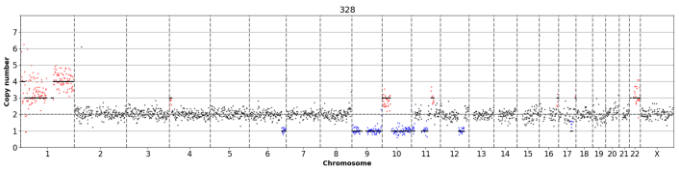
SRID40439, SLNcelllinec21



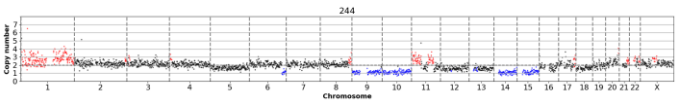
SRID40449, SLNcelllinec26



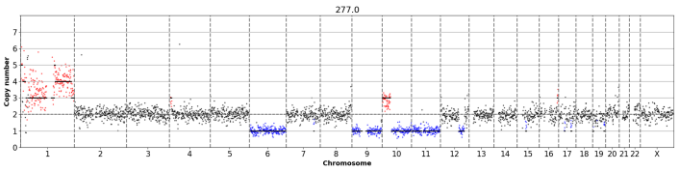
SRID40442, SLNcelllinec9



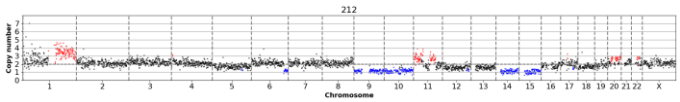
SRID40496, Met3gp100c39



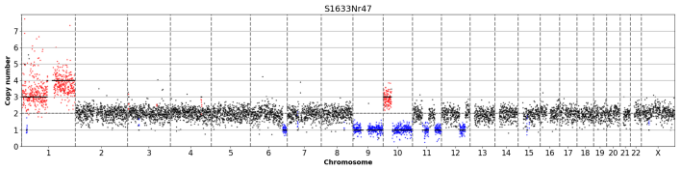
SRID40406, SLNDCClargeMCSPcryo48



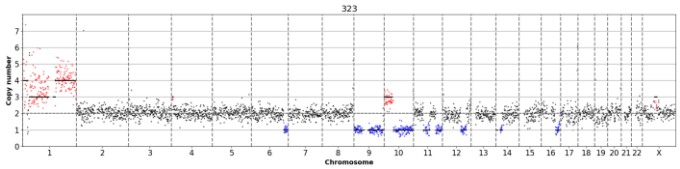
SRID40534, Met3gp100c38



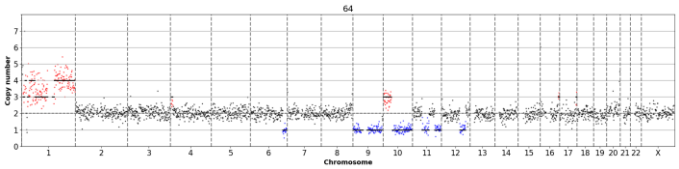
SRID40441, SLNcelllinec22



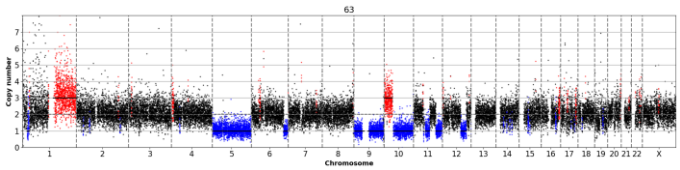
SRID40459, SLNcelllinec6



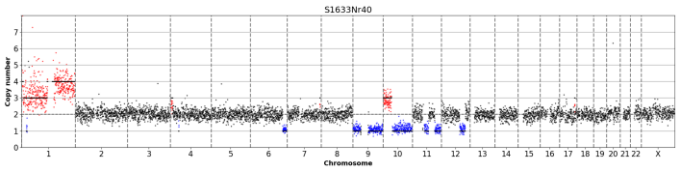
SRID40453, SLNcelllinec3



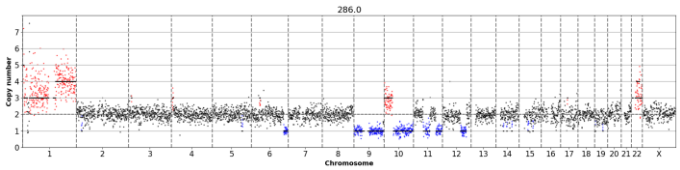
SRID40330, SLNDCClargeMCSP1



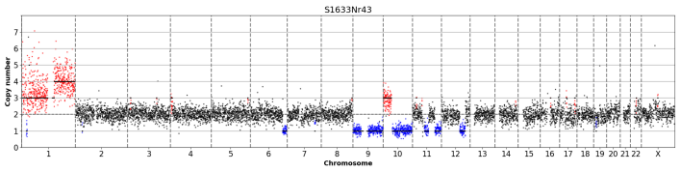
SRID40450, SLNcelllinec13



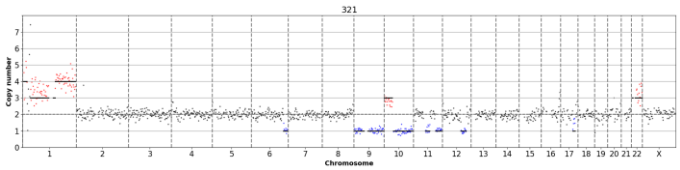
SRID40451, SLNcelllinec2



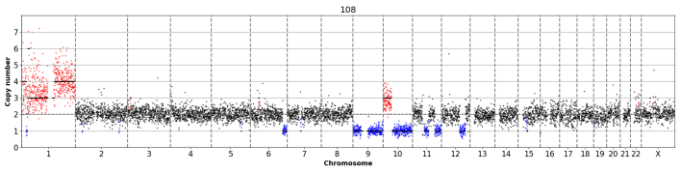
SRID40456, SLNcelllinec18



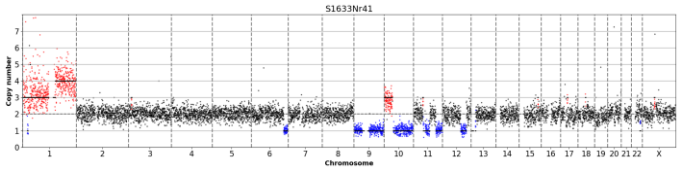
SRID40455, SLNcelllinec4



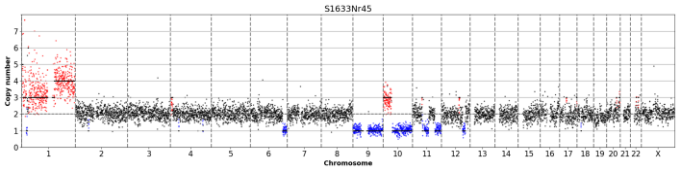
SRID42263, SLNcelllinec1



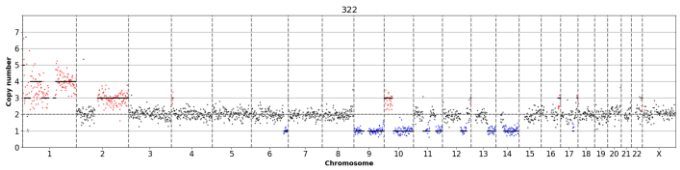
SRID40452, SLNcelllinec15



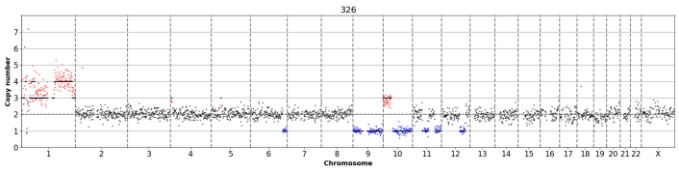
SRID40460, SLNcelllinec20



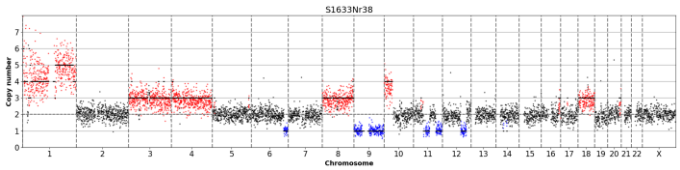
SRID40457, SLNcelllinec5



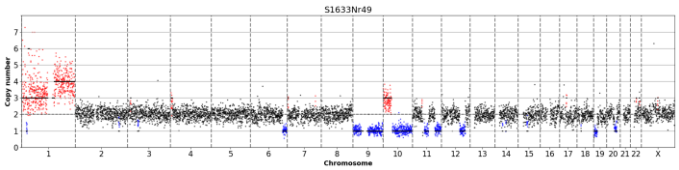
SRID40440, SLNcelllinec8



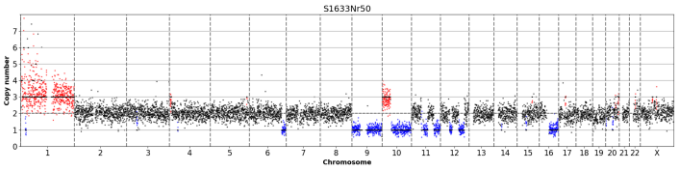
SRID40446, SLNcelllinec11



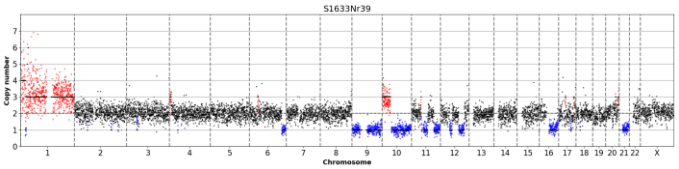
SRID40445, SLNcelllinec24



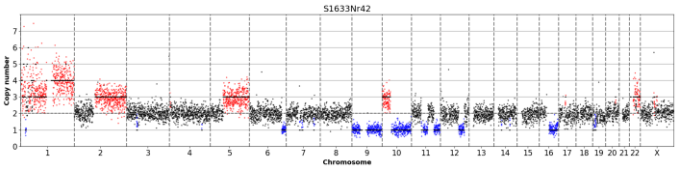
SRID40447, SLNcelllinec25



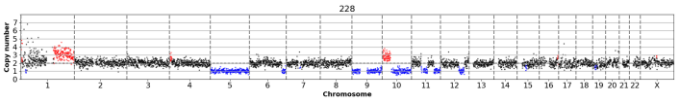
SRID40448, SLNcelllinec12



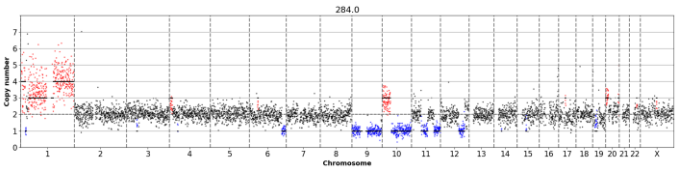
SRID40454, SLNcelllinec16



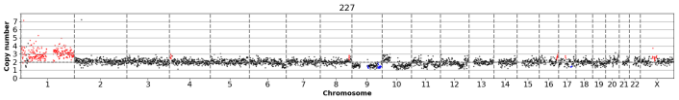
SRID40426, SLNDCClargeMCSPcryo27



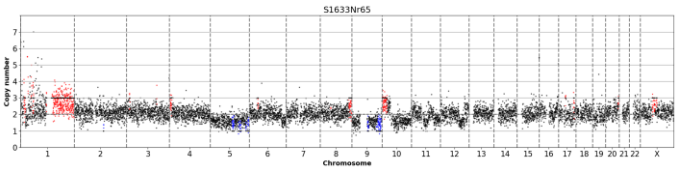
SRID40412, SLNDCClargeMCSPcryo37



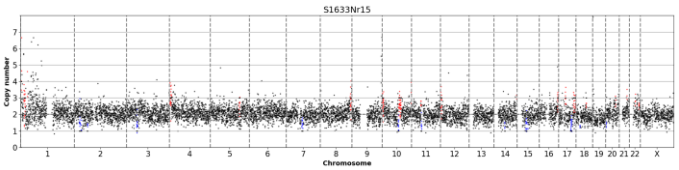
SRID40428, SLNDCClargeMCSPcryo14



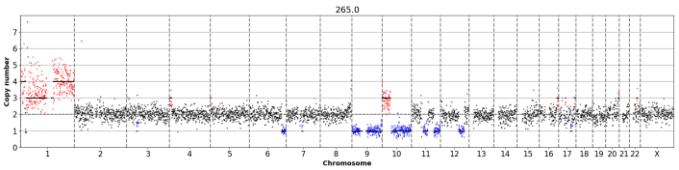
SRID40413, SLNDCClargeMCSPcryo9



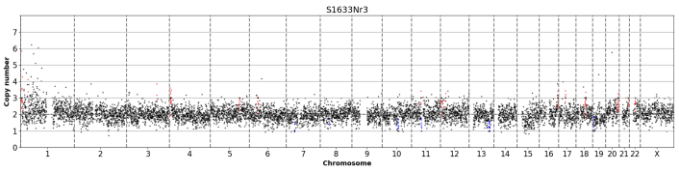
SRID40461, SLNCD31c8



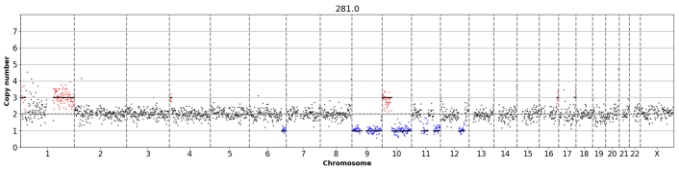
SRID40478, SLNDCCgp100c13



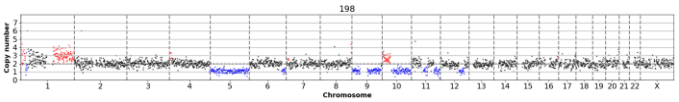
SRID40469, SLNCD31c1



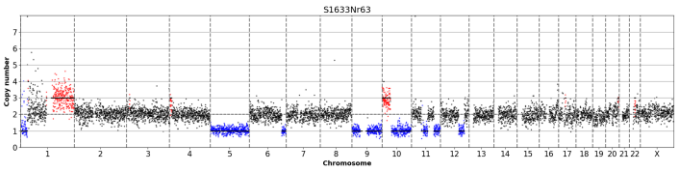
SRID40481, SLNDCCgp100c15



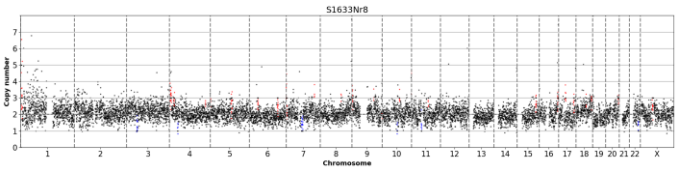
SRID40354, SLNDCClargeMCSP6



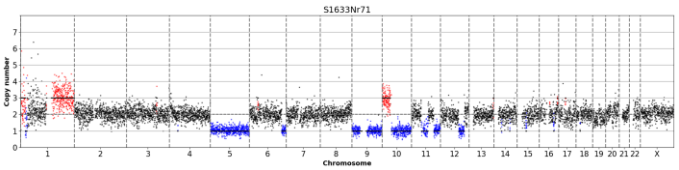
SRID40342, SLNDCClargeMCSP4



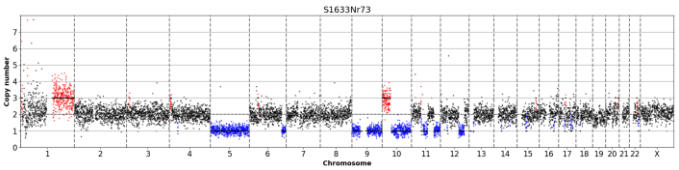
SRID40475, SLNCD31c15



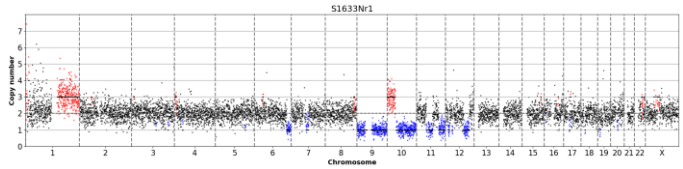
SRID40405, SLNDCCsmallMCSPcryo18



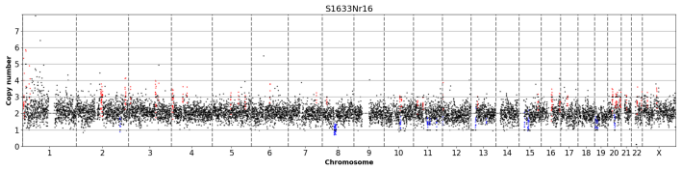
SRID40420, SLNDCClargeMCSPcryo7



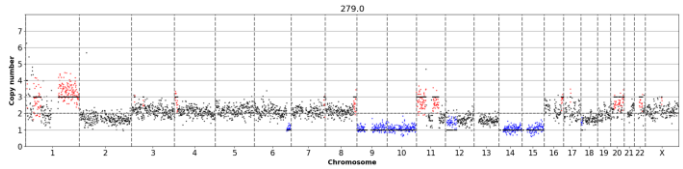
SRID40360, SLNDCClargeMCSPsn2



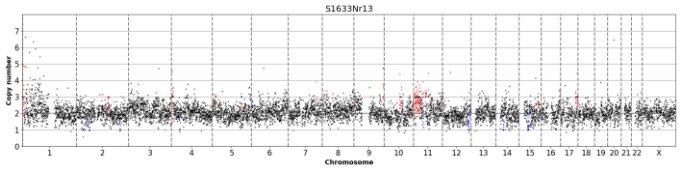
SRID40463, SLNCD31c9



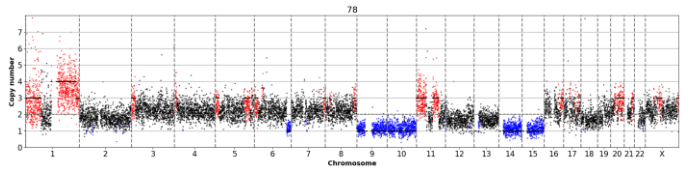
SRID40510, Met1gp100c40



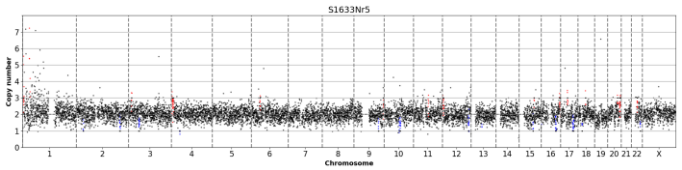
SRID42264, SLNCD31c6



SRID40532, Met1gp100c28



SRID40467, SLNCD31c12



7.2 Patient cohort tested by the ASB-PCR assay for the *BRAF* mutations c.1799T>A and c.1798_1799GT>AA

Patients from the Regensburg cohort were numbered by date of entry of the LN, patients from Tübingen were numbered by the University Hospital Tübingen and marked with a “T” to distinguish them from the Regensburg cohort.

N0: histopathologically negative LN; N+: histopathologically positive LN; DCCD: disseminated cancer cell density (number of DCCs per million cells); wt: wild type; mut: mutant

Patient number	cohort	lymph node status	DCCD (gp100)	DCCD group	BRAF status	gender
250	Regensburg	N0	2.5	<100	wt	male
288	Regensburg	N0	4	<100	wt	male
326	Regensburg	N+	147.5	≥100	mut	male
308	Regensburg	N+	13	<100	mut	male
312	Regensburg	N+	79	<100	wt	female
411	Regensburg	N+	50000	≥100	mut	male
204	Regensburg	N0	1	<100	mut	female
181	Regensburg	N0	1	<100	wt	male
184	Regensburg	N0	2	<100	wt	female
215	Regensburg	N+	2	<100	wt	male
247	Regensburg	N0	0.5	<100	wt	female
262	Regensburg	N0	0.5	<100	wt	female
265	Regensburg	N0	0.5	<100	wt	male
274	Regensburg	N0	10	<100	wt	male
276	Regensburg	N0	61	<100	mut	female
100	Regensburg	N+	180	≥100	wt	female
65	Regensburg	N+	106	≥100	mut	female
99	Regensburg	N+	300000	≥100	mut	male

10	Regensburg	N+	120	≥100	mut	female
104	Regensburg	N0	5	<100	wt	female
108	Regensburg	N+	500000	≥100	wt	male
124	Regensburg	N+	5000	≥100	mut	male
125	Regensburg	N+	706	≥100	wt	male
126	Regensburg	N0	1	<100	wt	male
134	Regensburg	N0	1	<100	wt	female
135	Regensburg	N+	174	≥100	mut	female
141	Regensburg	N+	105	≥100	mut	female
155	Regensburg	N+	10000	≥100	mut	female
165	Regensburg	N0	1	<100	mut	male
190	Regensburg	N+	173	≥100	mut	female
24	Regensburg	N0	3	<100	wt	female
25	Regensburg	N0	2	<100	wt	male
277	Regensburg	N+	20000	≥100	wt	male
30	Regensburg	N0	1	<100	wt	male
32	Regensburg	N0	1	<100	wt	male
34	Regensburg	N0	1	<100	wt	female
37	Regensburg	N+	224	≥100	wt	female
41	Regensburg	N+	400	≥100	wt	male
45	Regensburg	N0	2	<100	wt	male
48	Regensburg	N0	1	<100	wt	female
52	Regensburg	N0	9	<100	wt	male
53	Regensburg	N0	7	<100	wt	male
66	Regensburg	N0	3	<100	wt	male
71	Regensburg	N+	213	≥100	mut	female
81	Regensburg	N0	2	<100	wt	female
84	Regensburg	N0	2	<100	wt	female
91	Regensburg	N0	1	<100	wt	male
94	Regensburg	N0	2	<100	wt	male
293	Regensburg	N0	2.5	<100	wt	female
296	Regensburg	N0	9.5	<100	wt	male
323	Regensburg	N0	1	<100	wt	female
328	Regensburg	N0	0.5	<100	wt	female
404	Regensburg	N0	0.5	<100	mut	male
415	Regensburg	N0	0.5	<100	wt	female
430	Regensburg	N0	3	<100	wt	male
431	Regensburg	N0	3	<100	wt	female
433	Regensburg	N0	0.5	<100	wt	female
438	Regensburg	N0	0.5	<100	wt	male
447	Regensburg	N0	0.5	<100	wt	male
453	Regensburg	N0	1.5	<100	wt	male
454	Regensburg	N0	0.5	<100	wt	female
460	Regensburg	N0	0.5	<100	wt	male
476	Regensburg	N0	2	<100	wt	female
483	Regensburg	N0	1.5	<100	wt	female
488	Regensburg	N0	1.5	<100	wt	female

490	Regensburg	N0	0.5	<100	mut	female
495	Regensburg	N0	333.5	≥100	mut	male
515	Regensburg	N0	1.5	<100	wt	female
529	Regensburg	N0	2	<100	wt	female
534	Regensburg	N0	0.5	<100	wt	female
537	Regensburg	N0	1.5	<100	wt	female
546	Regensburg	N0	7	<100	wt	female
551	Regensburg	N0	1.5	<100	wt	male
555	Regensburg	N0	0.5	<100	wt	male
576	Regensburg	N0	2	≥100	wt	male
580	Regensburg	N0	19	<100	mut	male
582	Regensburg	N0	3	<100	mut	female
583	Regensburg	N0	35	<100	mut	male
T464	Tübingen	N+	412	≥100	mut	male
T519	Tübingen	N+	95	<100	mut	male
T600	Tübingen	N+	42500	≥100	wt	female
T574	Tübingen	N+	1587.5	≥100	mut	female
T603	Tübingen	N+	151.5	≥100	mut	male
T680	Tübingen	N0	168.5	≥100	wt	male
T588	Tübingen	N+	17200	≥100	mut	male
T529	Tübingen	N0	80.5	<100	wt	male
T996	Tübingen	N+	772	≥100	mut	female
T1612	Tübingen	N0	3	<100	wt	female
T1615	Tübingen	N0	24200	≥100	wt	female
T1620	Tübingen	N+	2500	≥100	wt	male
T1622	Tübingen	N+	160	≥100	mut	female
T1648	Tübingen	N0	3	<100	wt	male
T1669	Tübingen	N+	30.5	<100	mut	female
T1671	Tübingen	N+	348.5	≥100	wt	male
T1670	Tübingen	N+	3036.5	≥100	mut	male
T1694	Tübingen	N+	2100	≥100	mut	female
27	Regensburg	N+	180	≥100	mut	female
36	Regensburg	N+	2	<100	mut	male
72	Regensburg	N+	4	<100	wt	male
128	Regensburg	N+	34	<100	wt	male
168	Regensburg	N+	10000	≥100	mut	male
191	Regensburg	N+	?	≥100	mut	male
222	Regensburg	N0	1.5	<100	mut	male
245	Regensburg	N+	1.5	<100	mut	male
272	Regensburg	N+	97.5	<100	mut	male
273	Regensburg	N0	0.5	<100	wt	female
283	Regensburg	N0	1	<100	wt	male
284	Regensburg	N0	2	<100	wt	male
285	Regensburg	N+	2.5	<100	wt	female
290	Regensburg	N+	0.5	<100	wt	female
300	Regensburg	N+	12	<100	mut	female

338	Regensburg	N+	93	<100	wt	male
418	Regensburg	N0	5	<100	wt	male
420	Regensburg	N0	1	<100	wt	male
436	Regensburg	N+	73.5	<100	wt	male
439	Regensburg	N+	1	<100	wt	male
454	Regensburg	N0	0.5	<100	wt	female
478	Regensburg	N+	10000	≥100	mut	female
491	Regensburg	N+	146	≥100	mut	female
494	Regensburg	N+	1000	≥100	wt	male
510	Regensburg	N+	0.5	<100	wt	male
516	Regensburg	N+	50000	≥100	mut	female
521	Regensburg	N+	159.5	≥100	wt	male
T305	Tübingen	N0	25	<100	mut	male
T337	Tübingen	N0	15	<100	mut	male
T525	Tübingen	N0	36.5	<100	mut	male
T564	Tübingen	N0	15	<100	wt	male
T1623	Tübingen	N0	8	<100	wt	female
T1796	Tübingen	N0	5.5	<100	mut	male
T1846	Tübingen	N0	6	<100	wt	female
T1857	Tübingen	N0	11	<100	wt	female
T1912	Tübingen	N0	17.5	<100	wt	male
T2064	Tübingen	N0	4.5	<100	mut	female
T2167	Tübingen	N0	4.5	<100	mut	male
T2044	Tübingen	N0	3.5	<100	wt	female
T1908	Tübingen	N0	3	<100	wt	female
T1764	Tübingen	N0	3.5	<100	wt	male
T217	Tübingen	N0	4000	≥100	wt	male
T307	Tübingen	N0	3230	≥100	mut	female
T335	Tübingen	N+	3928	≥100	mut	male
T340	Tübingen	N+	18000	≥100	mut	female
T345	Tübingen	N+	500000	≥100	wt	female
T415	Tübingen	N+	57000	≥100	wt	male
T324	Tübingen	N+	412	≥100	wt	male
T459	Tübingen	N+	28900	≥100	mut	female
T518	Tübingen	N+	17100	≥100	wt	male
T560	Tübingen	N+	950000	≥100	wt	female
T570	Tübingen	N+	68400	≥100	mut	male
T1012	Tübingen	N+	1500	≥100	mut	male
T1327	Tübingen	N+	584700	≥100	mut	male