**UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF**

Institut für Tumorbiologie

unter der Direktion von Prof. Dr. Klaus Pantel

# Deciphering Molecular Subtypes in Advanced Prostate Cancer by Transcriptional Profiling of Circulating Tumor Cells

Dissertation

zur Erlangung des Grades eines Doktors (Dr. rer. biol. hum.)
an der Medizinischen Fakultät der Universität Hamburg.

vorgelegt von:
Lina Bergmann (geb. Merkens)
aus Buxtehude

Hamburg 2024

Angenommen von der
Medizinischen Fakultät der Universität Hamburg am: 06.05.2024

Veröffentlicht mit Genehmigung der
Medizinischen Fakultät der Universität Hamburg.

Prüfungsausschuss, der/die Vorsitzende: Prof. Dr. Gunhild von Amsberg

Prüfungsausschuss, zweite/r Gutachter/in: Prof. Dr. Klaus Pantel

# Contents

# List of Figures

# List of Tables

# 1 Introduction

## 1.1 Prostate Cancer

### 1.1.1 Epidemiology and Etiology

Worldwide, approximately 1.4 million men were diagnosed with prostate cancer (PCa) in 2020, making it the second most common non-cutaneous cancer in men after lung cancer. In terms of lethality, PCa was ranked fifth with about 370,000 deaths in 2020 [1]. The major risk factors for PCa are age, familiar predisposition, hereditary syndromes and ethnicity, whereas life style factors such as obesity and smoking show only weak association with PCa mortality [2]. Pathogenic germline variants in homologous recombination repair genes, including *BRCA2*, are among the most common hereditary alterations that predispose for PCa and are also associated with a worse prognosis [3], [4].

Serum levels of the prostate-specific antigen (PSA) and digital rectal examination are standard methods for prostate cancer screening, while tissue biopsies are mandatory to confirm a diagnosis. The majority of patients is diagnosed with early stage, localized disease and has a good prognosis compared to patients with metastatic disease at initial diagnosis [5]. Regular assessment of PSA in the serum is key to monitoring PCa patients and an increase of PSA levels in the course of therapy is referred to as biochemical recurrence.

The prostate is part of the male reproductive system and serves as an accessory gland for the seminal fluid. Its epithelium contains three major cell types: luminal, basal and neuroendocrine cells, with the exocrine luminal cells being the largest cell population [6]. There is strong evidence that prostate adenocarcinoma (PRAD) arises from the luminal cells, but also the basal cells are discussed as cell of origin [7]. At the molecular level, androgen receptor (AR) signaling is the main driver of tumor growth. PRAD has a lower tumor mutational burden compared

to other tumor entities [8]. Common somatic mutations are found in genes such as *TP53*, *SPOP*, *FOXA1* and *MYC* [9]. However, structural aberrations predominate the aberrant genomic landscape in PRAD [10]. Especially, fusions with the ETS transcription factor family, such as the most common *TMPRSS2:ERG* fusion, are present in about half of the tumors [11]. In addition, aberrations such as amplification of the *AR* locus or *MYC* as well as loss of *NKX3-1* and *PTEN* are frequently found [9]. Due to the low tumor mutational burden and the infiltration with immunosuppressive myeloid cells, PRAD is an immunologically cold tumor [12]. Overall, PRAD shows a high degree of inter- and intra-patient heterogeneity. This includes intraprostatic spatial heterogeneity between individual tumor foci even in a localized disease stage [13]. In 74 % of patients, primary tumor foci show significant differences in terms of point mutations and copy number aberrations [14]. In the metastatic setting, monoclonal and polyclonal metastasis-to-metastasis seeding is common and multiple sites of metastasis exhibit convergent evolution of resistance mechanisms [15].

### 1.1.2 Treatment Resistance

As the deregulation of AR signaling is the major driver of PCa growth, targeted systemic therapy options include androgen deprivation therapy (ADT) and novel hormonal agents (NHA), such as enzalutamide or abiraterone in addition to chemotherapy. In recent years, several clinical studies showed a benefit of intensified first line treatment in metastatic hormone-sensitive PCa (mHSPC) compared to ADT alone [16]–[18]. The combined use of ADT and NHA and/or taxane-based chemotherapy significantly prolonged the overall survival (OS) and progression-free survival (PFS) of the included patients [16], [18], [19]. However, progression to metastatic castration-resistant PCa (mCRPC) occurs when cancer cells acquire mutations that allow activation of the AR despite the deprived concentration of its physiological ligands. Several molecular mechanisms have been described that enable AR activation, such as increased expression of the AR mediated by gene amplification, enhanced translation or attenuated protein degradation of AR [20]–[22]. Mutations in the AR ligand binding domain can cause constitutive activation of the receptor [23]. Depending on the prior therapy line, NHA are also applied in the therapy of mCRPC, due to the underlying deregulation of AR signaling [24], [25]. However, the increased use of NHA has also led to the emergence of more complex, additional resistance mechanisms by which the tumor cells evade the increased therapeutic pressure. For instance, additional muta-

tions in the *AR* gene can render enzalutamide from an inhibitor to an agonist of AR signaling [26]. Furthermore, the increased expression of long non-coding RNAs can facilitate the expression of AR target genes [27]. Alternative splicing of the *AR* omitting the C-terminal binding domain, for example the AR-V7 splice variant, can induce NHA resistance [28]. Additionally, the oncogene c-Myc which is highly expressed in mCRPC promotes AR transcription and enhances protein stability of AR and splice variants, thereby supporting enzalutamide resistance [29]. Another mechanism of escape of AR blockade is the induction of the glucocorticoid receptor. As AR and glucocorticoid receptor have partially overlapping target genes, the up-regulation is sufficient to maintain the resistance phenotype in the absence of AR signaling [30]. An additional resistance mechanism is lineage plasticity and the emergence of AR-indifferent tumors that do not longer depend on the classical AR signaling pathways for survival and proliferation [31].

### 1.1.3 Aggressive Variant Prostate Cancer

While the majority of CRPC tumors retains aberrant AR signaling as major driver of cell proliferation also during NHA resistance, a subset of patients acquires resistance to AR-targeted therapy by losing its dependency on the AR pathway. This AR-independent disease stage is known as aggressive variant prostate cancer (AVPC). The frequency of AR-negative PCa increased from 11.6 % in the pre-NHA era (1998-2011) to 36.7 % in the modern era (2012-2016) [31]. Importantly, the standard biomarker in PRAD, serum PSA, is AR-dependent and, thus, no longer reliable in AVPC [32]. Clinically, AVPC can be characterized by at least one of the following criteria defined by Epstein *et al.* and Aparicio *et al.* [33], [34]:

- Histologic evidence of small-cell prostate carcinoma (pure or mixed);
- Exclusively visceral metastases;
- Radiographically predominant lytic bone metastases by plain x-ray or CT scan;
- Bulky ($\geq 5\,\mathrm{cm}$) lymphadenopathy or bulky ($\geq 5\,\mathrm{cm}$) high-grade (Gleason $\geq$ 8) tumor mass in prostate/pelvis;
- Low PSA ($\leq 10\,\mathrm{ng/mL}$) at initial presentation (before ADT or at symptomatic progression in the castrate setting) plus high volume ($\geq$ 20) bone metastases;
- Presence of neuroendocrine markers on histology (chromogranin A or syn-

aptophysin) or in serum (chromogranin A or GRP) at initial diagnosis or at progression. Plus any of the following in the absence of other causes: A. elevated serum LDH; B. malignant hypercalcemia; C. elevated serum CEA;

- Short interval ($\leq$ 6 month) to androgen-independent progression following the initiation of hormonal therapy with or without the presence of neuroendocrine markers.

An overview of the stages of PCa progression from localized PCa over mHSPC and mCRPC to AVPC is shown in figure 1.1. Both molecularly and histologically, AVPC is a heterogeneous disease and lineage plasticity is the main driver of resistance. Neuroendocrine transdifferentiation (NET) is a mechanism that has been observed in patients who lost AR-dependency and progressed to neuroendocrine prostate cancer (NEPC) [35]. NEPC can refer to different forms of neuroendocrine tumors of the prostate comprising small and large cell neuroendocrine carcinoma, Peneth-cell like differentiation and carcinoid tumors [36]. However, upon NET pure or admixed small cell carcinomas are observed and, therefore, the term NEPC is used in this thesis to refer to these subtypes. Similar to the prostate, NET is observed in lung cancer, where adenocarcinomas have been shown to progress to small cell lung cancer as a mechanism of treatment resistance [37]. However, loss of AR and AR signaling do not necessitate the emergence of NEPC [38]. Tumors that are negative for AR as well as neuroendocrine markers are referred to as double-negative prostate cancer (DNPC). While the molecular drivers of DNPC have long been unknown, current research is starting to identify major signaling pathways and transcription factors (TF) that govern cell proliferation in this subtype. For example, Bluemn *et al.* have identified FGF signaling and subsequent MAP kinase pathway activation as drivers of DNPC [31]. Further, Tang *et al.* have found two AR-independent subtypes of CRPC next to the neuroendocrine tumors. One subtype has acquired a stem-cell like phenotype with the AP1 family of TFs and YAP/TAZ driving the proliferation [39]. The acquisition of cancer stemness features mediated by YAP1 was demonstrated in an enzalutamide-resistant cell model before [40]. The other subtype is driven by Wnt signaling and the downstream TFs TCF/LEF [39]. Additional studies on mCRPC tissue and case reports have also identified a squamous cell differentiation in patients with DNPC [41], [42], while a series of DNPC patient-derived xenografts (PDX) has shown a mix of basal and luminal expression signatures [38].

Treatment options for AVPC are still very limited and mostly rely on chemother-

apy. Platinum-based chemotherapy has been shown to have a higher activity in patients with AVPC as compared to unselected mCRPC patients [34]. Typically, a combination of chemotherapeutic agents is applied, such as Carboplatin-Cabazitaxel or Platinum-Etoposide for pure NEPC [43], [44]. Alternatively, salvage chemotherapy with Cisplatin, Ifosfamide, and Paclitaxel has been shown to improve survival of heavily pretreated patients [45]. For a subpopulation of patients with aberrations in DNA repair genes, PARP inhibitors have been approved [46]. To date, immune checkpoint inhibition has not led to significant improvement of therapy, as seen in other entities. For instance, PD-L1 blockade has shown poor overall response in a small cohort of AVPC patients, not selected based on microsatellite instability [47]. However, checkpoint inhibitors are now being tested in combination with other small molecule inhibitors or chemotherapy, which might increase their benefit in the future [12].

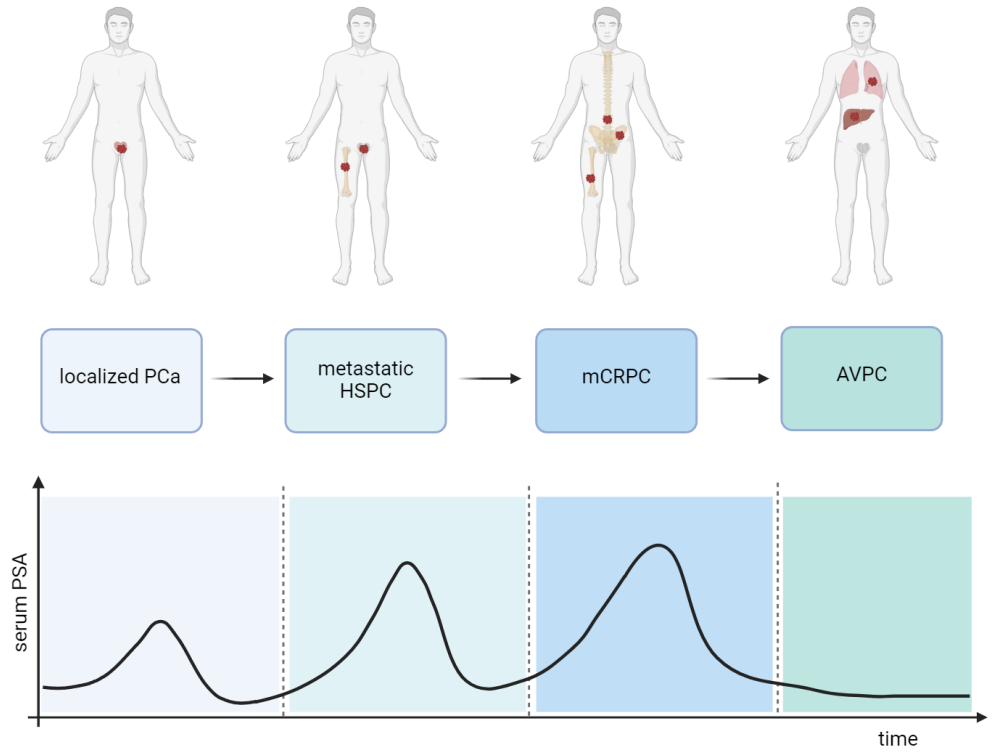



Figure 1.1: **Stages of PCa progression.** PCa is typically detected in a localized or metastatic hormone-sensitive stage that is accompanied by rising serum PSA; patients are treated with local therapy or systemic ADT; progression to CRPC is detected by rising serum PSA; following treatment with NHA patients can progress with AR-independent AVPC which is not detected with PSA.

### 1.1.4 Neuroendocrine Transdifferentiation

While DNPC is still poorly understood, NET is the better studied resistance mechanism of AR-independence in AVPC. As a result of NET, patients progress with treatment-emergent NEPC. NEPC is a highly proliferative and aggressive disease with dismal prognosis [48]. In histology, NEPC often displays the typical morphology of a small cell carcinoma with scant or hardly any cytoplasm and lack of nucleoli. Chromogranin A, synaptophysin, neuron-specific enolase 2 (NSE) and neural cell adhesion molecule 1 (CD56) are commonly used markers for NEPC in immunohistochemistry. However, the histological findings can be heterogeneous and apart from pure small cell carcinomas, many patients have tumors with mixed histology consisting of a high-grade adenocarcinoma and and an admixed small cell neuroendocrine component [36]. In addition, some patients have amphicrine tumors that are double-positive for AR and neuroendocrine markers on a cellular level [49]. Therefore, the categorization and nomenclature of these tumors remain a challenge.

*De novo* small cell PCa is a rare disease that is found in less than 1 % of patients at initial diagnosis [50]. In contrast, approximately 15 % - 20 % of patients progress with NEPC after enduring treatment in the setting of CRPC [51], [52]. A study by Abida *et al.* has found that the frequency of NEPC increased from 2.3 % to 10.5 % in mCRPC tissue biopsies from NHA-treated patients compared to NHA-naïve patients [51]. The role of AR-targeted therapy in the emergence of NEPC is further supported by an increase in NEPC incidence after the approval of enzalutamide and abiraterone from 6.3 % (1998-2011) to 13.3 % (2012-2016) [31]. Of note, a rising age-adjusted incidence rate of NEPC was also observed in the years 2004 to 2011 suggesting other factors might as well promote this phenomenon [53]. These factors might include an increased life-expectancy, advanced testing or therapeutic pressure of taxanes or radiation therapy [54], [55].

In principle, there are two contradictory hypotheses to explain the primary cellular origin of NEPC. According to the first hypothesis, NEPC emerges after clonal evolution of basal or neuroendocrine cell populations in the prostate [56], [57]. Before ADT and NHA treatment, AR-positive adenocarcinoma cells are thought to have a growth advantage compared to the transformed basal or neuroendocrine cells. However, upon AR inhibition, their AR-independence is a major growth advantage resulting in the development of NEPC [58]. More precisely, lineage tracing experiments in a transgenic mouse model of PCa by Lee *et al.* have

suggested basal cells as the origin of NEPC [56].

The second hypothesis assumes a mechanism of transdifferentiation from luminal adenocarcinoma cells to neuroendocrine tumor cells. Lotan *et al.* showed that the frequency of ERG gene rearrangements is similar in patients with small cell carcinoma compared to PRAD and in patients with mixed adenocarcinoma and small cell carcinoma of the prostate the ERG status was highly congruent between the two components [59], [60]. Since ERG fusion products such as TMPRSS2:ERG are activated in an AR-dependent manner, this suggests that the NEPC cells had previously been relying on AR signaling, which is typical only for luminal cells. In addition, amplifications of the *AR* gene are similarly distributed in NEPC and PRAD, again supporting an AR-driven origin [52].

In fact, the loss of AR in NEPC rarely occurs due to genomic aberrations, but rather by epigenetic or post-transcriptional mechanisms [61]. In general, many genomic alterations, such as point mutations or copy number aberrations, are concurring in PRAD and NEPC, whereas considerable differences in transcriptional and epigenetic regulation have been described [62]. On the transcriptional level, single-cell RNA sequencing and intra-tumoral RNA velocity analysis suggests that NE cells are directly originating from luminal-like adenocarcinoma cells [63]. Additionally, re-exposure to androgens has been shown to reverse neuroendocrine transdifferentiation in *in vitro* models of androgen deprivation [64]. Taken together these results strongly support the notion of a transdifferentiation mechanism driving the emergence of treatment-induced NEPC in the majority of patients.

However, the exact mechanism of transdifferentiation has not been fully elucidated, yet. First, the presence of amphicrine cells might indicate that the activation of neuroendocrine pathways and markers precedes the loss of AR signaling. In this case, the amphicrine cells represent an intermediate stage in the transdifferentiation process, after which the loss of AR signaling is required to progress to small cell carcinoma. Labrecque *et al.* found that transcriptional programs in amphicrine tumors differ from those in small cell tumors and that amphicrine tumors express a smaller subset of neuroendocrine TFs and markers, suggesting that the loss of AR is necessary to activate the full-blown neuroendocrine expression signature [41]. However, the same study also found a subset of AR-low tumors without the expression of neuroendocrine markers. Dedifferentiation of luminal cells to a stem-like state might be the first step to the development of DNPC, but might as well be an intermediate step to the development of NEPC, in

which the adenocarcinoma cells first have to lose AR signaling and gain plasticity before they differentiate to NEPC [65].

Importantly, NE differentiation alone is not sufficient for the development of this rapidly growing cancer: the neuroendocrine cells in healthy prostate tissue do not exhibit enhanced proliferation capabilities and, in cell culture experiments, the induction of a neuroendocrine phenotype reduced cell growth and proliferation [66], [67]. Consequently, deregulation of additional pathways involved in lineage plasticity, such as stemness signaling or epithelial-to-mesenchymal transition (EMT) is required for the fast and aggressive proliferation found in NEPC. The combination of neuroendocrine differentiation and a plastic cell state is illustrated, for example, by over-expression of the basal marker TROP2 and the EMT-inducer SNAIL in aggressive PRAD cell models. Depending on the genetic background, both have been shown to be sufficient to induce a neuroendocrine phenotype individually, indicating that NET, basal-like gene expression and EMT are interwoven in NEPC [68], [69]. Meanwhile, several genes have been identified that are involved in the process of NET and these include transcriptional regulators, chromatin modulators, histone modifiers and enhancers of proliferation [70]. An overview of the central pathways and main players of NET is given in figure 1.2.

As mentioned before, there are few genomic aberrations that are indicative of NEPC. Combined loss-of-function alterations in the tumor suppressor genes *TP53*, *RB1* and *PTEN* are increased in patients with AVPC as compared to mCRPC [71]. *In vivo* knock-out studies have shown that the combinatorial, but not the single knock-out of those tumor suppressors significantly enhances tumor aggressiveness as well as lineage plasticity and induces a neuroendocrine-like tumor growth in mice [72], [73]. However, patient-derived xenograft models have also shown that the combined loss of the three tumor suppressors does not inevitably cause the emergence of NEPC [38].

In contrast, many TFs have been identified as drivers of the NET process. ASCL1 and FOXA2 are pioneering TFs that can bind to closed chromatin regions to induce genes associated with neuronal differentiation and cancer stem cells [74], [75]. Additionally, the TFs POU3F2, SOX11, NKX2-1 and LMO3 have been recognized to promote the neuroendocrine phenotype [41], [76]–[78]. This is accompanied by an increased expression of chromatin remodelers and histone modifiers such as ACTL6B, EZH2 and LSD1 [79]–[81]. Moreover, RNA processing factors such as SRRM4 and LIN28B are deregulated to facilitate the epigenetic

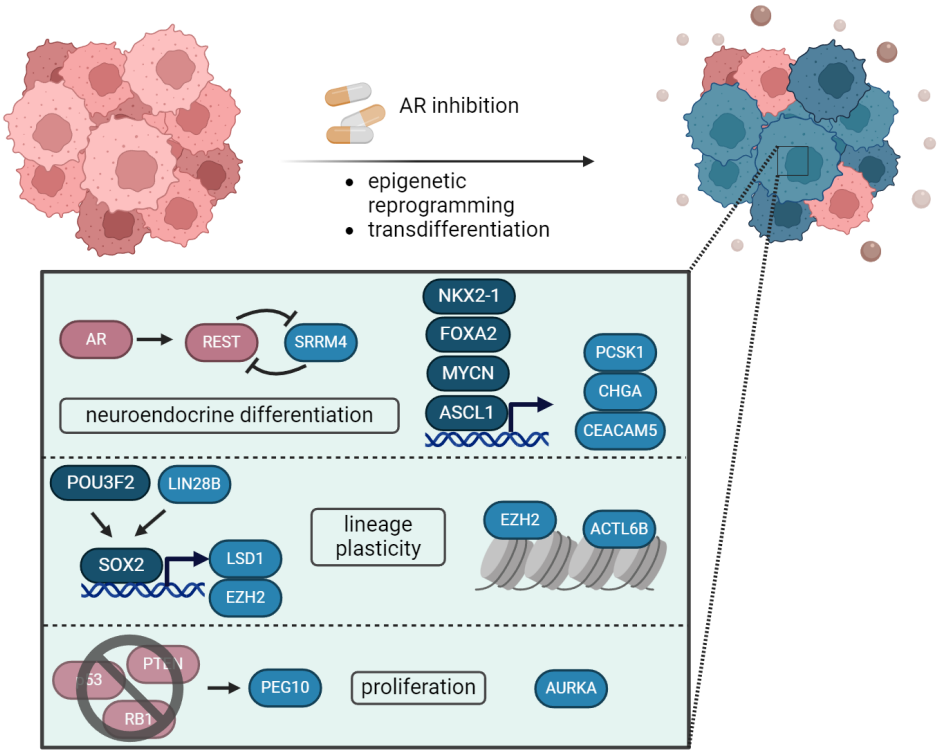



Figure 1.2: **Signaling pathways in NET.** Emergence of NEPC is caused by AR inhibition and requires transdifferentiation and epigenetic reprogramming; this involves induction of multiple TFs and epigenetic regulators that facilitate proliferation, lineage plasticity and neuroendocrine differentiation; red - down-regulated, blue - up-regulated.

reprogramming during NET [82], [83]. Up-regulation of PEG10 and AURKA is required to maintain the high proliferation capacity of the transdifferentiated tumor cells [84], [85]. Next to the drivers of the NET process mentioned before, downstream genes are used as NEPC markers, although there is no evidence for a functional role in the NET process. One example is the most commonly used marker for NEPC, CHGA, which is activated by ASCL1 and thereby indicates a successful reprogramming of the cell fate [74]. PCSK1 encodes a pro-hormone convertase involved in the maturation of proneuropeptides and prohormones in neuroendocrine cells [86]. PCSK1 was identified as significantly upregulated in NEPC compared to PRAD in different patient cohorts [41], [87]. CEACAM5 is another marker of neuroendocrine differentiation in PCa [88]. As a cell surface protein, it is of special interest for the development of new substances for the detection or targeted therapy of NEPC. CEACAM5 expression in NEPC tissue samples showed little overlap with other well-known cell surface markers such as PSMA and TROP2 [89].

## 1.2 Liquid Biopsy

Tissue biopsies are still the basis for diagnosis and therapy decisions in modern oncology. However, the collection of tissue biopsies is associated with high risk and pain for the patients and, therefore, also not suitable for longitudinal sampling. As tissue biopsies are often taken as a punch biopsy, they have the risk not to cover the heterogeneity within the tumor or between different sites of metastasis. The term liquid biopsy has been introduced for the analysis of tumor cells or tumor cell-derived biomolecules from body fluids [90]. Body fluids that are harnessed for liquid biopsy include blood, urine, saliva or cerebrospinal fluid [91]. Liquid biopsies are considered to be less risky and they can be taken repeatedly. This makes liquid biopsy a valuable tool at different disease stages, including screening of people at risk, detection of minimal residual disease and monitoring of treatment response, resistance and tumor progression [92]. Due to the multitude of information that can be extracted from a liquid biopsy sample - be it mutations, gene silencing or protein expression - liquid biopsy is a key to precision oncology and personalized therapy. A schematic overview of liquid biopsy and potential analytes is shown in figure 1.3. Circulating nucleic acids (cfDNA, cfRNA), extracellular vesicles (EV), proteins and other tumor metabolites can be isolated from the blood plasma. cfDNA fragments can be analyzed to obtain information about tumor mutations, copy number aberrations, DNA methylation profiles and chromatin accessibility [93]. Due to their role in cell communication, EVs are considered important mediators in the formation of the metastatic niche. Therefore, EVs may carry biomarkers that are relevant for disease progression [94]. Among the cellular fraction, tumor-educated platelets and cancer-assiciated fibroblasts have been identified as biomarkers, while circulating tumor cells (CTC) remain the most studied cell type in liquid biopsy [95], [96].

### 1.2.1 Circulating Tumor Cells

CTCs are malignant cells that have left the tumor tissue and entered the blood stream. CTCs can either be passively shed from the tumor or actively leave the tumor after acquiring mesenchymal traits [97]. However, loss of cell-cell-contacts and the shear stress in the circulation are harsh conditions and, thus, the half-life time of CTCs is limited to about $1\,\mathrm{h}$ to $3\,\mathrm{h}$ [98]. Some CTCs are able to survive the stress, extravasate and migrate to distant organs. These cells become the mediators of metastasis via the blood stream [99]. Whole exome sequencing of CTCs

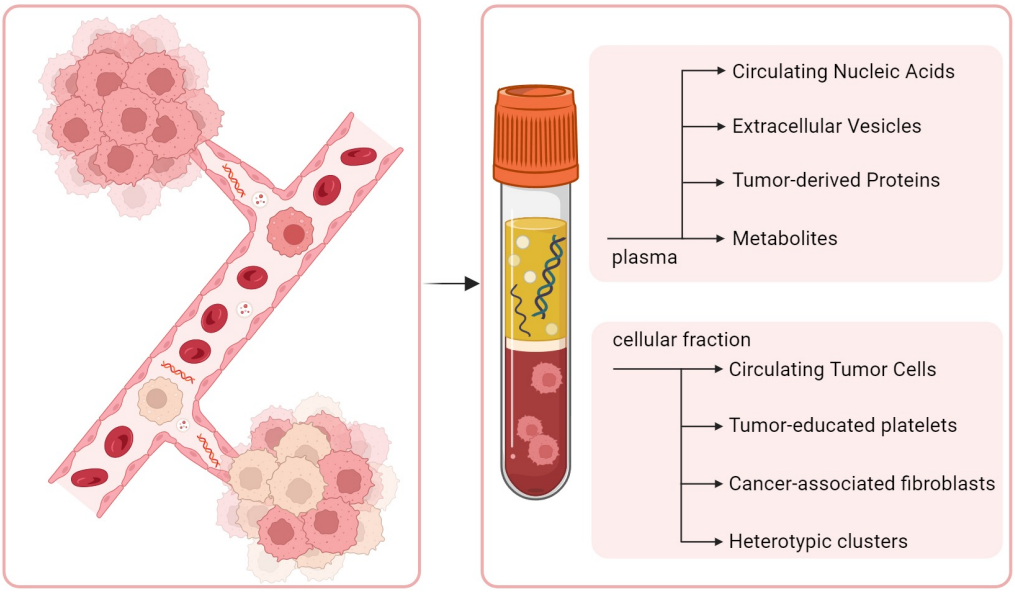



Figure 1.3: **Overview of analytes in liquid biopsy.** Primary tumors and metastases shed tumor cells and additional biomolecules into the blood circulation, these can be isolated from peripheral blood; circulating nucleic acids, extracellular vesicles and small metabolites can be analyzed in the blood plasma while the mononuclear cell fraction contains CTCs and additional cancer-associated cells.

as well as primary and metastatic tissue has revealed that not only the primary tumor but also metastases contribute to the pool of CTCs found in the circulation [100].

In the circulation, a few CTCs are diluted in millions of leukocytes and erythrocytes. Therefore, enrichment of CTCs is the first step followed by detection and analysis [91]. Enrichment of CTCs can generally be performed based on the expression of epithelial or tumor-specific markers (positive selection) and leukocyte markers (negative selection), or based on physical properties such as deformability and size. For positive selection of CTCs, EPCAM is the most common epithelial cell surface marker used for immunomagnetic enrichment. The CellSearch system is the first FDA-cleared device for automated EPCAM-based enrichment and enumeration of CTCs. It has been successfully used for CTC detection in PCa, breast cancer and colorectal cancer, among others [101]–[103]. Importantly, CTCs can serve as prognostic and predictive biomarkers as well as surrogate markers for survival in clinical trials [104]. Additional markers such as EGFR, HER2 or organ-specific markers are also used by other enrichment technologies, such as the AdnaTest [105]. In contrast, label-independent enrichment methods are not

constrained by cell differentiation and have been shown to improve the detection of CTCs undergoing EMT. The Parsortix device relies on cell size for CTC enrichment. Its key part is a microfluidic cassette, in which cells are moving through a stepped structure narrowing down to $6.5\,\mu m$ to leave leukocytes through but trap larger and more rigid tumor cells [106]. Other technologies, for instance microfluidic spirals or filters, also harness other physical properties of CTCs such as density, deformability or electrical charges [92].

Immunocytochemistry staining of cytokeratins (CK) is commonly applied for the detection of CTCs, but other methods such as the EPISPOT assay or PCR-based assays are also available. CTCs can further be isolated for genomic, transcriptomic or other analysis. These analyses can either be performed on the single cell level or on the bulk CTC fraction. Copy number aberrations and mutations can be analyzed by next generation sequencing (NGS), but this needs to be coupled with whole genome amplification for single cell analysis [107]. In-situ hybridization or amplification assays are used to study structural variants, but also mRNA and microRNA expression in single CTCs [108]–[110]. If enough cells are available, CTC transcriptomes can be analyzed by next-generation single cell RNA-sequencing [111].

In summary, CTCs can be enriched from peripheral blood and enumeration as well as transcriptional profiling of these cells can inform about tumor burden and specific aberrations. Consequently, CTCs serve as valuable biomarkers in clinical trials.

## 1.3 Use of Liquid Biopsy in Prostate Cancer

In PCa, the detection of tumor-derived proteins in serum in the form of PSA monitoring is the earliest application of liquid biopsy. PSA is used as a standard biomarker for PCa screening, disease monitoring and detection of biochemical recurence. However, PSA is not a tumor-, but a prostate-specific marker. It indicates AR signaling, but does not yield information about driver mutations or resistance mechanisms, especially in the case of AR-independent disease. Therefore, more advanced liquid biopsy assays for PCa are developed, focusing on complex analytes such as CTCs or cfDNA. In comparison to other metastatic carcinomas, detection of CTCs appears to be rather frequent in mCRPC [112]. This facilitates the enrichment and, consequently, also the analysis of CTCs. With regard to cfDNA, patients seem to be less frequently positive for tumor signatures

compared to other advanced carcinomas. However, if positive, patients have a comparably high number of mutant fragments [113].

For a reliable use in the clinic, the genomic aberrations found in CTCs or ctDNA need to be congruent with tissue samples. Several studies confirmed a general congruence of cancer tissue and liquid biopsy results. For instance, a high concordance of selected somatic mutations of matched tissue biopsy and ctDNA samples has been found in a cohort of metastatic PCa [114].

### 1.3.1 Prognostic and predictive markers in the setting of prostate adenocarcinoma

CTC count has been one of the first prognostic liquid biopsy-based biomarkers in PCa. Evaluation of CTC count in a CRPC cohort before and after therapy has led to the definition of a threshold of 5 CTCs per $7.5\,\mathrm{mL}$ blood to assess disease prognosis [101]. Meanwhile, the CTC count is included as an intermediate endpoint of survival in clinical trials, thus being also used as a response biomarker [115]. Apart from the count of CTCs, also the presence of CTC clusters has been associated with a worse prognosis [116]. Similarly, the baseline cfDNA concentration can serve as an independent prognostic biomarker for PFS in clinical trials of first- and second-line taxane chemotherapy [117]. In addition, early changes in ctDNA fraction during NHA treatment have been shown to be associated with duration of response to therapy and survival in a prospective trial in CRPC patients [118].

Since alternative splicing is a well-defined resistance mechanism against NHA, the expression of the AR-V7 splice variant has been analyzed in CTCs. Indeed, clinical studies have shown that patients with AR-V7 positive CTCs have a worse outcome after NHA compared to patients with AR-V7 negative CTCs and the integration of AR-V7 detection and CellSearch-based enumeration in CTCs might have an additional benefit as a predictive biomarker set [119], [120]. Moreover, AR-V7 has been identified in exosomal RNA of mCRPC patients and also in this setting, AR-V7 has been identified as a marker for resistance to hormonal therapy [121]. Likewise, other known mutations from PCa tissue have been detected in CTCs and ctDNA. For instance, amplification or mutation of *AR* in ctDNA predicted the outcome of CRPC patients on NHA [122], [123]. Further, specific tumor-derived, non-coding RNA expression levels in blood have been described to be associated with prognosis and response to treatment, but applications in clinical trials are not as common as for CTCs or ctDNA [124], [125].

A step beyond the analysis of single markers is the assessment of a marker panel. For instance, a signature of expression levels of cancer-related genes in CTCs from mCRPC patients has been described as a classifier to predict resistance to NHA and PFS with a higher accuracy than AR-V7 alone [126]. Similarly, evaluation of CTC count and gene expression in CRPC and mHSPC patients predicted outcome after docetaxel therapy [127]. In general, the use of NGS, such as targeted-NGS for the evaluation of aberrations in large panel of genes or with low-pass whole genome sequencing drastically increases the amount of data that can be extracted from the low quantities of analytes in the blood. Similarly, analysis of mutations and DNA methylation in CRPC patients can inform about prognosis and therapy options [123], [128], [129].

In conclusion, CTC count and detection of tumor-specific transcripts such as AR-V7 are well-established biomarkers for survival and treatment resistance in advanced PCa.

### 1.3.2 Liquid Biopsy in the detection of neuroendocrine transdifferentiation

In recent years, scientists have started to apply liquid biopsy approaches to detect progression from PRAD to NEPC.

The easiest approach tested is the quantification of neuromediators released into the plasma by neuroendocrine cells. In a small pilot study, the neuromediators CHGA, NSE and pro-GRP have been measured in plasma from mCRPC patients before and after cabazitaxel treatment. However, no significant correlation with disease progression or survival has been found [130]. Heck *et al.* have analyzed serum NSE and CHGA in mCRPC patients treated with abiraterone. NSE and CGA levels were not correlated and did not predict PSA response, but the combination of both was a significant predictor of PFS and OS [131].

The expression of neuroendocrine and stemness markers has also been analyzed on the level of cell-free RNA in plasma. Overall, the patients in advanced stage PCa exhibited a high degree of heterogeneity in their expression profiles. At least one of four neuroendocrine marker transcripts was identified in 15 % of CRPC patients and this was associated with worse prognosis. Neuroendocrine markers have always been detected in combination with other transcripts, such as luminal or stemness markers. However, no tissue samples were analyzed to determine sensitivity or specificity of that assay [132].

At CTC level, staining of cell surface markers is an easy method for phenotype analysis. Synaptophysin is one of the most commonly used markers for neuroendocrine cell differentiation and it has been tested for detection of neuroendocrine CTCs in CellSearch. In mCRPC, detection of SYP positive cells is correlated with short time to progression on NHA, but neuroendocrine differentiation was not confirmed by histology [133]. Similarly, expression of the Notch ligand DLL3 has been found to be significantly up-regulated in NEPC compared to adenocarcinoma and detection of DLL3 on CTCs was highly congruent with matched tissue biopsies. About 60 % of patients had DLL3-positive CTCs with variable fractions of the DLL3-positive subpopulation [134].

CTC morphology has also been described as a marker for NEPC. Beltran and colleagues have used the non-selection-based *EPIC Sciences* platform to detect CTCs in mCRPC patients with known histology based on morphology and immunocytochemistry staining. CTCs from NEPC patients have lower AR expression and a higher nuclear to cytoplasmic ratio. About 60 % of NEPC patients with CTCs were positive for the neuroendocrine marker CD56 on at least one cell, while it was absent in PRAD samples. NEPC CTCs were further characterized by smaller cell area, cell circularity and an increased number of CTC clusters. Identification of NEPC-like CTCs in CRPC patients was associated with aggressive disease [135]. Using the same platform, CTCs have been analyzed in a case report of a patient with primary high grade adenocarcinoma and a liver metastasis with neuroendocrine transdifferentiation. Based on morphology, a fraction of the cells showed a small cell phenotype and almost all cells expressed CD56 while none expressed AR. Single cell whole genome sequencing of CTCs has revealed a high concordance of copy number aberrations with the liver metastasis but not the primary tumor [136]. In CTCs, combined loss of the three tumor suppressors *TP53*, *RB1* and *PTEN* was correlated with the expression of neuroendocrine markers and lower AR expression [136]. In depth analysis of *TP53*, *RB1* and *PTEN* loss in cfDNA with targeted NGS has shown concordance in all paired plasma-tissue samples with a tumor fraction larger than 10 %. Loss of the tumor suppressors was correlated with worse clinical outcome [137].

Transcript analysis of CTCs enriched from PCa patients using the EPCAM-based VERSA platform revealed a high degree of heterogeneity between patients. In addition to prostate markers, CHGA and SYP have been included as neuroendocrine markers, but the panel allowed only for poor clustering of NEPC patients, mainly because only about half of CTCs from the NEPC patients had a positive

CHGA signal. However, in one case, positivity for neuroendocrine markers preceded clinical symptoms about 3 months, thus emphasizing the potential use of liquid biopsy for NEPC detection [138]. The same platform for CTC analysis has been used in a study evaluating enzalutamide in mCRPC patients and identified the expression of neuroendocrine markers as a resistance mechanism [139].

Taken together, various methods have been applied to identify patients undergoing NET with liquid biopsy. On the protein level, neuroendocrine markers have been detected in plasma and on CTCs, but the sensitivity and specificity of these markers is not well studied. Gene expression profiling of CTCs has shown promising results for early detection of NET in single case studies, however, higher sensitivity is required for reliable application in clinical testing.

## 1.4 Aim of the study

As a consequence of enduring therapeutic pressure on the AR signaling pathway, PRAD may lose its dependence on AR for cell proliferation. AVPC is AR-indifferent, grows more aggressively and treatment options are limited. NET is a resistance mechanism by which tumor cells acquire an AR-independent phenotype. Hence, PSA no longer is a reliable biomarker in the setting of AVPC and new markers are urgently needed to detect AR-indifferent disease and monitor treatment response in these patients. However, metastatic biopsies are often difficult to obtain and associated with risks and pain for patients. Additionally, a single biopsy does not capture the heterogeneity that is observed in AVPC patients. Within a single patient, metastases with luminal or neuroendocrine differentiation can coexist and even a single tumor can consist of different subclones. Therefore, this study aimed to identify new biomarkers for AVPC and NEPC based on CTC analysis. CTCs depict the heterogeneity within a patient better than a tissue biopsy and the analysis is easy to repeat. Classical serum biomarkers for neuroendocrine disease such as SYP, CHGA and CD56 are not expressed universally in NEPC and their specificity is limited as focal expression of these markers is also observed in PRAD tumors [140]. As NET is mainly driven by alterations in the epigenome, the decision was made to identify potential markers on the transcriptome, but not the genome level. For this purpose, CTCs were chosen as liquid biopsy analytes.

Based on whole genome or whole exome sequencing experiments in tissue biopsies, several signatures have previously been identified that differentiate quite

accurately between AR-driven adenocarcinoma and NEPC [41]. They are based on the expression levels of different stem cell, NE- and AR- related transcripts. Subsequently, the knowledge on these NEPC signatures has to be combined with the recent advances in liquid biopsy approaches in order to identify patients progressing to NEPC based on blood samples. The aim of this study was to enrich CTCs from patients with AVPC and NEPC as well as from patients with hormone-sensitive disease as control and analyze the transcript profiles in these cells. As the test should be suitable for future routine use in the clinic, multiplex qPCR analysis of bulk RNA from enriched CTCs was chosen. This allowed a rapid turn around time and was less expensive than NGS analysis, for instance. For this purpose, a transcript panel was identified and validated based on recent publications, tissue data sets and PCa cell lines. In addition, the use of label-dependent CTC enrichment methods was compared to label-independent CTC enrichment. The marker panel was afterwards measured by qPCR on cDNA generated from AdnaTest-enriched CTCs. In parallel, CTC counts were determined using the CellSearch system. The acquired data was afterwards compared to clinical parameters of the patients.

# 2 Materials and methods

## 2.1 Materials and Reagents

### 2.1.1 Cell lines

All cell lines used in this study were provided by the Institute of Tumor Biology or the lab for Experimental Oncology at the second medical clinic and had been purchased from ATCC. LNCaP cells (RRID: CVCL_0395) originate from a lymph node of a male patient with hormone-sensitive PRAD [141]. VCaP cells (RRID: CVCL_22359) were derived from a bone metastasis of a patient with castration-resistant PCa still expressing PSA [142]. NCI-H660 (RRID: CVCL_1576) originate from a lymph node metastasis of a patient with treatment-naive small cell PCa [143]. DU145 cells (RRID: CVCL_0105) are castration-resistant PCa cells without PSA expression derived from a brain metastasis [144]. PC3 cells (RRID: CVCL_0035) originate from a bone metastasis of a patient with high grade castration-resistant PRAD [145]. LASCPC-01 cells (RRID: CVCL_UE17) are a model of neuroendocrine PCa generated from a PDX model of benign prostate basal cells with overexpression of MYCN and constitutively active AKT1 [146]. NCI-H209 cells (RRID: CVCL_1525) originate from bone marrow of a male patient suffering from treatment-naive small cell lung cancer [147]. All cells were cultured in their respective medium shown in table 2.1 at $37\,^{\circ}\mathrm{C}$. Cells grown in DMEM were kept at 10 % $CO_2$, while the other cells cultured in RPMI-based medium were kept at 5 % $CO_2$.

Table 2.1: Cell culture media.

| **Medium** | **Supplements** | **Cell lines** |
|---|---|---|
| DMEM | $2\,\mathrm{mM}$ L-glutamine, 10 % FCS | DU145, VCaP, PC3 |

| Medium | Supplements | Cell lines |
|---|---|---|
| HITES | 2 mM L-glutamine, 5 % FCS, 1 % Insulin-Transferrin-Selenium A, 10 nM β-estradiol, 10 nM hydrocortisone in RPMI | LASCPC-01, NCI-H660 |
| RPMI | 2 mM L-glutamine, 10 % FCS | LNCaP, NCI-H209 |

### 2.1.2 Chemicals and Reagents

All chemicals and reagents used in this study are listed in table 2.2.

Table 2.2: Chemicals and Reagents

| Reagent | Company |
|---|---|
| AB-Serum | Bio-Rad Medical Diagnostics (Dreieich, DE) |
| β-estradiol | Sigma-Aldrich (St. Louis, US) |
| β-mercaptoethanol | Merck (Darmstadt, DE) |
| DMEM High Glucose | PAN-Biotech (Aidenbach, DE) |
| DRAQ5™staining solution | Miltenyi Biotec (Bergisch Gladbach, DE) |
| Dulbecco's phosphate buffered saline | Gibco (Eggenstein, DE) |
| Erythrocyte Lyse Buffer (10X) | R&D Systems (Minneapolis, US) |
| Ethanol absolut | ChemSolute/Th. Geyer (Renningen, DE) |
| Fetal Bovine Serum, Advanced | Capricorn Scientific (Ebsdorfergrund, DE) |
| Ficoll-Paque Plus™ | Amersham Bioscience (Buckinghamshire, UK) |
| Fluoromount-G™ | Thermo Fisher Scientific (Carlsbad, US) |
| Hydrocortisone | Sigma-Aldrich (St. Louis, US) |
| Insulin-Transferrin-Selenium A, 100X | Gibco (Eggenstein, DE) |
| Isopropanol | ChemSolute/Th. Geyer (Renningen, DE) |
| L-glutamine 200 mM | Gibco (Eggenstein, DE) |
| Nuclease-free water | Qiagen (Hilden, DE) |
| Paraformaldehyde | ChemSolute/Th. Geyer (Renningen, DE) |
| RNasin®Plus Ribonuclease Inhibitor | Promega (Madison, US) |
| RPMI 1640 | PAN-Biotech (Aidenbach, DE) |

| Reagent | Company |
|---|---|
| Triton X-100 | Sigma-Aldrich (St. Louis, US) |
| Trypan blue staining solution | Sigma-Aldrich (St. Louis, US) |
| 0.25 % Trypsin-EDTA | Gibco (Eggenstein, DE) |
| Water, deionised | B. Braun (Melsungen, DE) |
| Water, nuclease-free | Qiagen (Hilden, DE) |

### 2.1.3 Antibodies

All antibodies including conjugates are given in table 2.3.

Table 2.3: Antibodies

| Target | RRID | Clone | Conjugate | Manufacturer |
|---|---|---|---|---|
| CD45 | AB_893337 | HI30 | PerCp | Biolegend (San Diego, US) |
| pan-cytokeratin | AB_1834350 | AE1/AE3 | AF488 | Invitrogen (Carlsbad, US) |
| pan-cytokeratin | AB_2616664 | C-11 | AF488 | Biolegend (San Diego, US) |
| NCAM1 | AB_2905076 | REA196 | BV423 | Miltenyi Biotec (Bergisch Gladbach, DE) |

### 2.1.4 Commercial Kits

Table 2.4 lists all commercial kits used in this study.

Table 2.4: Commercial Kits

| Kit | Manufacturer |
|---|---|
| AdnaTest ProstateCancerDetect | Qiagen (Hilden, DE) |
| AdnaTest ProstateCancerSelect | Qiagen (Hilden, DE) |
| CellSearch® Circulating Tumor Cell Kit | Menarini Silicon Biosystems (Castel Maggiore, IT) |
| NucleoSpin RNA Mini Kit | Macherey & Nagel (Düren, DE) |
| RevertAid First Strand cDNA Synthesis Kit | Thermo Fisher Scientific (Waltham, US) |
| Sensiscript RT Kit | Qiagen (Hilden, DE) |
| TATAA Interplate Calibrator FAM | TATAA Biocenter (Göteborg, SW) |
| TATAA Probe GrandMaster® Mix | TATAA Biocenter (Göteborg, SW) |

| Kit | Manufacturer |
|---|---|
| TATAA SYBR® GrandMaster® Mix | TATAA Biocenter (Göteborg, SW) |
| ValidPrime® Human Probe assay | TATAA Biocenter (Göteborg, SW) |

### 2.1.5 Oligonucleotides

Commercial qPCR assay were bought from TATAA Biocenter (Göteborg, SW) (table 2.5). Primers that were designed in the institute or chosen based on recent literature were purchased from MWG Eurofins (Ebersberg, Germany). Primers designed and provided by TATAA Biocenter (Göteborg, SW) were produced by Integrated DNA Technologies (München, Germany). All gBlocks were purchased from Integrated DNA Technologies (München, Germany). Primer sequences are listed in table 2.6 and gBlock sequences are given in table 2.7.

Table 2.5: Commercial probe-based qPCR assays.

| Transcript | Reference number | Design |
|---|---|---|
| ACTB | qA-01-0104P | TATAA Biocenter |
| AR | qA-01-0364P | TATAA Biocenter |
| AR-V7 | qA-01-0368P | TATAA Biocenter |
| CCND1 | qA-01-0203P | TATAA Biocenter |
| CD45/PTPRC | qA-01-0240P | TATAA Biocenter |
| EGFR | qA-01-0211P | TATAA Biocenter |
| EPCAM | qA-01-0212P | TATAA Biocenter |
| FOLH1 | qA-01-0357P | TATAA Biocenter |
| GAPDH | qA-01-0101P | TATAA Biocenter |
| HOXB13 | qA-01-0889P | TATAA Biocenter |
| HPRT1 | qA-01-0112P | TATAA Biocenter |
| KLK3 | qA-01-0356P | TATAA Biocenter |
| KRT19 | qA-01-0225P | TATAA Biocenter |
| PROM1 | qA-01-0371P | TATAA Biocenter |
| RAI2 | qA-01-0890P | TATAA Biocenter |
| TACSTD2 | qA-01-0372P | TATAA Biocenter |

Table 2.6: Single-stranded DNA oligonucleotides

| Transcript | Sequence | Design |
|---|---|---|
| ACTL6B-F | GGC CTG TTT GAT CCC TCG AA | [148] |
| ACTL6B-R | GGT GAC AAT GAC ACT CCC GT | [148] |

| Name | Sequence | Design |
|---|---|---|
| ASCL1-F | TCG CCG GTC TCA TCC TAC TC | TATAA Biocenter |
| ASCL1-R | GTT GTG CGA TCA CCC TGC TT | TATAA Biocenter |
| ASCL1-P | AGG AGC TTC TCG ACT TCA CCA ACT GG – FAM/ZEN | TATAA Biocenter |
| CEACAM5-F | GGG ACC TAT GCC TGT TTT GTC | TATAA Biocenter |
| CEACAM5-R | AGA GAC CAG GAG AAG TTC CAG A | TATAA Biocenter |
| CEACAM5-P | TTC CAT AGT CAA GAG CAT CAC AGT CTC TGC – FAM/ZEN | TATAA Biocenter |
| CHGA-F | TCC CTG TGA ACA GCC CTA TG | TATAA Biocenter |
| CHGA-R | AAG GAT CCG TTC ATC TCC TCG G | TATAA Biocenter |
| CHGA-P | CTC CGA CAC ACT TTC CAA GCC CAG CCC – FAM/ZEN | TATAA Biocenter |
| FOXA2-F | CGG TGA AGA TGG AAG GGC A | TATAA Biocenter |
| FOXA2-R | CAT GTT GCT CAC GGA GGA GT | TATAA Biocenter |
| FOXA2-P | AGC CGT CCG ACT GGA GCA GC – FAM/ZEN | TATAA Biocenter |
| KRT6A-F | TAG TGC CCT CAC TTC TTC TCT CTC | TATAA Biocenter |
| KRT6A-R | GCT CAG CCT CAG AGA TAG AAC AC | TATAA Biocenter |
| KRT6A-P | TGT AAT CAC CAC TGG AGC TTC ACT GTT – FAM/ZEN | TATAA Biocenter |
| LIN28B-F | TGT AGT CTA CCT CCT CAG CCA A | [83] |
| LIN28B-R | ATT CTG CTT CCT GTC TTC CCT G | [83] |
| LMO3-F | TCT GAG GCT CTT TGG TGT AAC G | [149] |
| LMO3-R | CCA GGT GGT AAA CAT TGT CCT TG | [149] |
| NKX2-1-F | CGT ACC AGG ACA CCA TGA G | TATAA Biocenter |
| NKX2-1-R | ATG CCG CTC ATG TTC ATG C | TATAA Biocenter |
| NKX2-1-P | CCA TCT CCC GCT TCA TGG GC – FAM/ZEN | TATAA Biocenter |
| NKX3-1-F | CCC ACA CTC AGG TGA TCG AG | ITB in-house |
| NKX3-1-R | GAG CTG CTT TCG CTT AGT CTT | ITB in-house |
| PCSK1-F | CCG ACC AGA GAA TCA CGA GC | TATAA Biocenter |
| PCSK1-R | ACC AGG TGC TGC ATA TCT CG | TATAA Biocenter |
| PCSK1-P | CCA GAG CGA AGA TGC CAG CAG CC – FAM/ZEN | TATAA Biocenter |

| Name | Sequence | Design |
|---|---|---|
| PEG10-F | GCC TCC ATC CCC ACA GAA GTG AAG C | TATAA Biocenter |
| PEG10-R | CAC TCT TAT GGC CGG TGT GCT TGG A | TATAA Biocenter |
| PEG10-P | CCC AAC CGT CAC CCT GGG TCC CGA CTG CCC – FAM/ZEN | TATAA Biocenter |
| POU3F2-F | GTA ACT GTC AAA TGC GCG GC | TATAA Biocenter |
| POU3F2-R | GAG GTG AGC AGG CTG TAG TG | TATAA Biocenter |
| POU3F2-P | CGG TCG CCA TGA CTC TCG GAG CC – FAM/ZEN | TATAA Biocenter |
| SOX11-F | CCA GGA CAG AAC CAC CTG AT | [150] |
| SOX11-R | CCC CAC AAA CCA CTC AGA CT | [150] |
| SRRM4-F | GCC CAT CGC CTG TCA AGA AA | TATAA Biocenter |
| SRRM4-R | TTT GGG CTA GAG GAG CTG TG | TATAA Biocenter |
| SRRM4-P | AAA GTT CCA AGA AAC ACA AGC GAC G – FAM/ZEN | TATAA Biocenter |

Table 2.7: Double-stranded DNA oligomers (gBlocks)

| Name | Sequence |
|---|---|
| ASCL1 | TCG CCG GTC TCA TCC TAC TCG TCG GAC GAG GGC TCT TAC GAC CCG CTC AGC CCC GAG GAG CAG GAG CTT CTC GAC TTC ACC AAC TGG TTC TGA GGG GCT CGG CCT GGT CAG GCC CTG GTG CGA ATG GAC TTT GGA AGC AGG GTG ATC GCA CAA C |
| CEACAM5 | TCG CCA AAA TCA CGC CAA ATA ATA ACG GGA CCT ATG CCT GTT TTG TCT CTA ACT TGG CTA CTG GCC GCA ATA ATT CCA TAG TCA AGA GCA TCA CAG TCT CTG CAT CTG GAA CTT CTC CTG GTC TCT CAG CTG GGG CCA CT |
| CHGA | TCC CTG TGA ACA GCC CTA TGA ATA AAG GGG ATA CCG AGG TGA TGA AAT GCA TCG TTG AGG TCA TCT CCG ACA CAC TTT CCA AGC CCA GCC CCA TGC CTG TCA GCC AGG AAT GTT TTG AGA CAC TCC GAG GAG ATG AAC GGA TCC TT |

| Name | Sequence |
|---|---|
| FOXA2 | TGC CAT GCA CTC GGC TTC CAG TAT GCT GGG AGC GGT GAA GAT GGA AGG GCA CGA GCC GTC CGA CTG GAG CAG CTA CTA TGC AGA GCC CGA GGG CTA CTC CTC CGT GAG CAA CAT GAA CGC CGG CCT GGG GAT GAA CGG CAT GAA CAC GTA |
| KRT6A | TAG TGC CCT CAC TTC TTC TCT CTC TCT CTA TAC CAT CTG AGC ACC CAT TGC TCA CCA TCA GAT CAA CCT CTG ATT TTA CAT CAT GAT GTA ATC ACC ACT GGA GCT TCA CTG TTA CTA AAT TAT TAA TTT CTT GCC TCC AGT GTT CTA TCT CTG AGG CTG AGC |
| NKX2-1 | TAC TGC AAC GGC AAC CTG GGC AAC ATG AGC GAG CTG CCG CCG TAC CAG GAC ACC ATG AGG AAC AGC GCC TCT GGC CCC GGA TGG TAC GGC GCC AAC CCA GAC CCG CGC TTC CCC GCC ATC TCC CGC TTC ATG GGC CCG GCG AGC GGC ATG AAC ATG AGC GGC AT |
| PCSK1 | CCG ACC AGA GAA TCA CGA GCG CTG ACC TGC ACA ATG ACT GCA CGG AGA CGC ACA CAG GCA CCT CGG CCT CTG CAC CTC TGG CTG CTG GCA TCT TCG CTC TGG CCC TGG AAG CAA ACC CAA ATC TCA CCT GGC GAG ATA TGC AGC ACC TGG T |
| PEG10 | GCC TCC ATC CCC ACA GAA GTG AAG CTA CAG CTG GGA GGT CTC CTC CCA CCC CAA CCG TCA CCC TGG GTC CCG ACT GCC CAC CTC CTC CTC CTC CCC CTC CCC CCA ACA TCA GCA ACA TCA ACA ACT CCA AGC ACA CCG GCC ATA AGA GTG |
| POU3F2 | GAG AGG GAG CCC GAG GCG AAA AAG TAA CTG TCA AAT GCG CGG CTC CTT TAA CCG GAG CGC TCA GTC CGG CTC CGA GAG TCA TGG CGA CCG CAG CGT CTA ACC ACT ACA GCC TGC TCA CCT CCA GCG CCT CCA TCG TGC AC |
| SRRM4 | GAA GGA GGT CCT CAT CCT ATA GCC CAT CGC CTG TCA AGA AAA AGA AGA AGA AAA GTT CCA AGA AAC ACA AGC GAC GCA GGT CAT TCT CCA AGA AGA GAA GGC ACA GCT CCT CTA GCC CAA AAA GCA AAA GAA GAG ATG AG |

### 2.1.6 Consumables

All consumables used in this study are listed in table 2.8

Table 2.8: Consumables

| **Product** | **Manufacturer** |
|---|---|
| BD Vacutainer K2E | Becton, Dickinson and Company (Plymouth, UK) |
| Cell culture flask T75, suspension cells | Sarstedt (Nümbrecht, DE) |
| Cell culture flask T25, adherent cells | Sarstedt (Nümbrecht, DE) |
| CellSave tube | Menarini Silicon Biosystems (Castel Maggiore, IT) |
| Conical tubes (15 mL, 50 mL) | Greiner Bio-One (Frickenhausen, DE) |
| Cover slides (22 x 22 mm) | Marienfeld (Lauda-Königshofen, DE) |
| ImmEdge® hydrophobic barrier pen | Vector Laboratories (Burlingame, US) |
| Microseal® B plate seal | Bio-Rad Laboratories (Hercules, US) |
| Monovette, EDTA | Sarstedt Sarstedt (Nümbrecht, DE) |
| Parsortix casette EU GEN 6.5 | ANGLE plc (Guildford, UK) |
| Pipette tips (10 µL, 100 µL, 200 µL, 1,000 µL ) | Sarstedt (Nümbrecht, DE) |
| Pipette tips with filter (10 µL, 100 µL, 200 µL, 1,000 µL) | Sarstedt (Nümbrecht, DE) |
| PCR plate, 96 wells | Biorad Laboratories (Hercules, US) |
| Reaction tubes (0.5 mL, 1.5 mL, 2.0 mL) | Sarstedt (Nümbrecht, DE) |
| Serological pipettes (2 mL, 5 mL, 10 mL) | Sarstedt (Nümbrecht, DE) |
| SuperFrost® Plus microscope slides | R. Langenbrinck (Emmendingen, DE) |
| Ultra-Low attachment plate, 24 well | Corning (Kennebunk, US) |

### 2.1.7 Laboratory Devices

A list of all laboratory devices used in this study is found in table 2.9.

Table 2.9: Laboratory Devices

| **Device** | **Manufacturer** |
|---|---|
| AdnaMag-L | Qiagen (Hilden, DE) |
| AdnaMag-S | Qiagen (Hilden, DE) |
| Axio Oberver Fluorescence microscope | Carl Zeiss (Jena, DE) |
| Celltracks Analyzer II® | Menarini Silicon Biosystems (Castel Maggiore, IT) |
| Celltracks® Autoprep® System | Menarini Silicon Biosystems (Castel Maggiore, IT) |
| Centrifuge 5415 R | Eppendorf (Hamburg, DE) |
| Cytocentrifuge Rotofix 32 | Hettich (Tuttlingen, DE) |
| Sterile hood Herasafe KS12 | Heraeus Kendro (Langenselbold, DE) |
| Incubator HERAcell 150i | Thermo Scientific (Waltham, US) |
| Nanodrop® ND-1000 spectrophotometer | PEQLAB Biotechnology (Erlangen, DE) |
| Micropipette (2.5 µL, 10 µL, 100 µL, 200 µL, 1,000 µL) | Eppendorf (Hamburg, DE) |
| Mini Plate Spinner MPS 1000 | Labnet (Corning, US) |
| Mini Star Centrifuge | VWR (Darmstadt, DE) |
| Multifuge 3 S-R | Heraeus (Hanau, DE) |
| Neubauer counting chamber | Hecht Assistent (Sondheim, DE) |
| Parsortix™PR1 | ANGLE plc (Guildford, UK) |
| PCR Workstation Pro | PEQLAB Biotechnology (Erlangen, DE) |
| Pipette boy | Hirschmann (Eberstadt, DE) |
| Roll mixer RM5 | Hecht Assistent (Sondheim, DE) |
| Rotator 2-1175 | neoLab (Heidelberg, DE) |
| Thermal Cycler PTC-200 | MJ Research (Waltham, US) |
| Thermomixer compact | Eppendorf (Hamburg, DE) |
| Thermoycler C1000 Touch CFX96™ | Bio-Rad Laboratories (Hercules, US) |
| Thermocycler peqSTAR 96 Universal Gradient | PEQLAB Biotechnology (Erlangen, DE) |
| Vortex-Genie 2 | Scientific Industries (Bohemia, US) |
| Water bath | Lauda-GFL (Lauda-Königshofen, DE) |

### 2.1.8 Software and Online Resources

Published RNA sequencing data sets for PCa tissue samples were downloaded from *cBioPortal* (cbioportal.org). For the Neuroendocrine Prostate Cancer dataset (Multi-Institute, Nat Med 2016), mRNA expression data were downloaded as log(FPKM) values [135]. For the Metastatic Prostate Adenocarcinoma dataset (SU2C/PCF Dream Team, PNAS 2019), sequencing data were downloaded at z-scores relative to all samples.

qPCR data were inspected and $C_q$ values were calculated using the CFX Manager 3.1 software (Biorad, Germany). Plots and statistical analyses were done with Graphpad Prism version 10 and R Studio version 2023.06.2 Build 561 with R version 4.2.2. *NormFinder* version 5 for R was used to check the stability of housekeeping genes and select housekeepers for subsequent analysis [151]. The R package *ComplexHeatmap* (V 2.14.0) was used to draw all heatmaps. For supervised analysis, the R packages *ranger* (V 0.16.0) and *caret* (V 6.0-94) were used. Survival data were plotted and evaluated using the R package *survminer* (V 0.4.9). Microscopy images were acquired and processed with *ZEN 3.1 blue edition*, (Carl Zeiss, Jena, DE). The illustrations in the introduction were created with *BioRender.com*.

## 2.2 Methods

### 2.2.1 Cell culture

#### 2.2.1.1 Passaging

All cells were cultivated in their respective media as listed in 2.1. To passage adherent cells, the medium was discarded and cells were washed with PBS. Cells were detached by incubation with $1\,\mathrm{mL}$ trypsin-EDTA at $37\,^{\circ}\mathrm{C}$. $4\,\mathrm{mL}$ medium were added and cells were resuspended by pipetting. Cells were collected by centrifugation at $1{,}200\,\mathrm{rpm}$ for $3\,\mathrm{min}$ and the supernatant was discarded. The cell pellet was resuspended in $10\,\mathrm{mL}$ medium and a fraction of cells was transferred to a new T25 cell culture flask and fresh medium was added to a final volume of $5\,\mathrm{mL}$. NCI-H660, NCI-H209 and LASCPC-01 cells, which all grow in suspension, were spun down at $1{,}200\,\mathrm{rpm}$ for $3\,\mathrm{min}$ and resuspended in fresh medium. A fraction of the cells was transferred to a new T75 cell culture flask and fresh medium was added to a final volume of $15\,\mathrm{mL}$. In case of NCI-H660, the supernatant was

not discarded, but $5\,\mathrm{mL}$ were add to the new flask to foster the cell growth by paracrine stimuli.

#### 2.2.1.2 Spike-in experiments

For spike-in experiments, cells were harvested as described above. For small cell numbers, the resuspended cells were diluted in PBS and $2\,\mathrm{mL}$ cell suspension were transferred to an ultra-low attachment 24-well plate and placed on the microscope. Cells were allowed to settle down for $5\,\mathrm{min}$ before single cells were soaked up with a $10\,\mu\mathrm{L}$ pipette under the microscope and transferred to a new tube or a blood sample, respectively. Spike-in samples were afterwards treated similar to the patient samples.

### 2.2.2 Patient samples

Blood samples from PCa patients were collected in $4.9\,\mathrm{mL}$ or $7.5\,\mathrm{mL}$ EDTA monovettes at the University Medical Center Hamburg-Eppendorf between November 2020 and June 2023. Blood collection was approved by the ethical commission of the city of Hamburg (Hamburger Ärztekammer) and all patients provided informed consent. Patients were selected by an experienced clinician based on clinical parameters and serum markers according to modified criteria defined by Epstein *et al.* and Aparicio *et al.* [33], [34]. Samples were pseudonymized in the clinic before transport to the research lab. For negative controls, blood was collected in EDTA tubes from male donors at the blood donation facility at the Department for Transfusion medicine, University Medical Center Hamburg-Eppendorf. Blood samples from healthy donors were anonymized, but information on the age of the donors was available. All blood samples were processed within $3\,\mathrm{h}$ after blood collection.

### 2.2.3 Processing of Blood Samples

#### 2.2.3.1 Isolation of Peripheral Blood Mononuclear Cells

Peripheral blood mononuclear cells (PBMCs) were isolated by gradient centrifugation. First, a whole blood sample or the cell fraction following the first step of plasma isolation were filled up with PBS to final volume of $30\,\mathrm{mL}$. The diluted blood sample was gently layered onto $20\,\mathrm{mL}$ ficoll and centrifuged at $400\,\mathrm{xg}$ for $30\,\mathrm{min}$ with low acceleration and break off. The aqueous upper phase and the

silk-like PBMC phase on top of the ficoll phase were transferred to a new tube and PBS was added to a final volume of $50\,\mathrm{mL}$. PBMCs were collected by centrifugation at $400\,\mathrm{xg}$ for $10\,\mathrm{min}$ with break. In case the pellet was of a reddish color, it was resuspended in $1\,\mathrm{mL}$ erythrocyte lysis buffer and incubated for $2\,\mathrm{min}$ at room temperature. $20\,\mathrm{mL}$ PBS were added and the PBMCs were collected by centrifugation at $400\,\mathrm{xg}$ for $10\,\mathrm{min}$ with break. PBMCs were resuspended in $10\,\mathrm{mL}$ PBS and counted with an automated cell counter.

#### 2.2.3.2 CTC enrichment using CellSearch

For CellSearch analysis, $7.5\,\mathrm{mL}$ EDTA blood were transferred to a CellSave tube and mixed by inversion. Within $72\,\mathrm{h}$, CTCs were enriched using the automated CellSearch device with the CellSearch Circulating Tumor Cell kit. In brief, CTCs were automatically enriched based on EPCAM expression and stained with fluorescence-labeled antibodies against pan-cytokeratin and CD45 as well as DAPI. Enriched and stained cells were scanned in a cassette using the Celltracks Analyzer II. Events with positive DAPI and pan-cytokeratin staining and absent CD45 staining were automatically filtered. All suspicious events were manually checked for cell morphology and CTC counts were documented. CTCs with a peri-nuclear dot-like cytokeratin staining were indicative of a small cell-like phenotype and the presence of these small cell-like CTCs was documented.

#### 2.2.3.3 CTC enrichment using Parsortix

CTC enrichment using Parsortix is based on microfluidics and harnesses the increased cell size and decreased deformability of tumor cells for enrichment. $7.5\,\mathrm{mL}$ EDTA blood were used for CTC enrichment. If plasma was isolated from the same tube, the respective volume was replaced with PBS after the first step of plasma isolation. According to manufacturer's instructions, the Parsortix device was primed with the *PX2_PF* program and a GEN 6.5 cassette was inserted. The CTCs were enriched using the S50F protocol or with the S99F protocol in case more pressure was needed to keep the blood moving through the cassette. CTCs were harvested using the *PX2_H* program with the *Further Flush* option. Cells were either harvested directly into a cytospin funnel or into a $1.5\,\mathrm{mL}$ reaction tube. From the latter, $0.5\,\mathrm{mL}$ were transferred into a cytospin funnel and the remaining cells were spun down at $800\,\mathrm{xg}$ for $10\,\mathrm{min}$. The supernatant was discarded and $200\,\mu\mathrm{L}$ AdnaTest lysis buffer were added. Cytospins were centrifuged at $1{,}200\,\mathrm{rpm}$ for $7\,\mathrm{min}$ and air-dried over night. Lysates and cytospin slides were

stored at $-80\,^{\circ}\mathrm{C}$.

#### 2.2.3.4 CTC enrichment using AdnaTest

For subsequent transcript analysis, CTCs were enriched using the label-dependent AdnaTest Prostate Cancer Select Kit. CTC enrichment relies on the expression of the cell surface markers EPCAM, HER2 and EGFR. CTC lysates were prepared according to the manufacturer's instructions within $3\,\mathrm{h}$ after blood collection. First, antibody-coupled magnetic beads were washed three times in $1\,\mathrm{mL}$ PBS for $1\,\mathrm{min}$. Next, $5\,\mathrm{mL}$ blood were incubated with $100\,\mu\mathrm{L}$ equilibrated beads for $30\,\mathrm{min}$ in slow rotation at room temperature. The beads were captured with a magnet and the blood was discarded. The beads were washed three times with $5\,\mathrm{mL}$ PBS and transferred to a $1.5\,\mathrm{mL}$ reaction tube in $1\,\mathrm{mL}$ PBS. Last, the beads were incubated in AdnaTest lysis buffer to release the cell contents. The supernatant containing the RNA was transferred to a new collection tube and stored at $-80\,^{\circ}\mathrm{C}$.

### 2.2.4 Gene expression analysis

#### 2.2.4.1 RNA isolation from cultured cells

Total RNA was isolated from cultured cells using the NucleoSpin RNA Mini kit according to the manufacturer's instructions. Briefly, up to five million cells were lysed in $350\,\mu\mathrm{L}$ lysis buffer with $3.5\,\mu\mathrm{L}$ $\beta$-mercaptoethanol. The lysate was cleared by centrifugation through a filter column. $350\,\mu\mathrm{L}$ 70 % ethanol were added and the lysate was loaded onto an RNA binding column. The column was desalted once with $350\,\mu\mathrm{L}$ MDB buffer and DNA was digested on the column with recombinant DNase for $15\,\mathrm{min}$. DNase was inactivated by washing buffer RAW2.

After two additional washing steps with RA3 buffer, the column was dried by centrifugation and the RNA was eluted in $40\,\mu\mathrm{L}$ nuclease-free water. RNA concentration was measured using the Nanodrop photometer and the RNA was stored at $-80\,^{\circ}\mathrm{C}$.

#### 2.2.4.2 cDNA synthesis from RNA of cultured cells

cDNA synthesis was set up using the RevertAid FirstStrand cDNA Synthesis Kit. Reaction mixtures were prepared according to the table 2.10. First, RNA, water and oligo(dT) primer were mixed and incubated at $65\,^{\circ}\mathrm{C}$ for $5\,\mathrm{min}$ and on ice for $2\,\mathrm{min}$. Next, the remaining reagents were added and reverse transcription was

Table 2.10: Composition of RevertAid cDNA synthesis reaction.

| reagent | volume [µL] |
|---|---|
| 5X Reaction buffer | 4 |
| dNTP mix, 10 mM | 2 |
| Oligo(dT)$_{18}$ primer, 100 µM | 4 |
| RiboLock RNase inhibitor | 1 |
| RevertAid M-MuLV RT | 2 |
| purified RNA | 1 µg |
| RNase-free water | ad 20 µL |

Table 2.11: Thermocycler protocol for RevertAid cDNA synthesis.

| temperature | time |
|---|---|
| 37 °C | 60 min |
| 70 °C | 5 min |

performed in the thermocycler at 37 °C and the reverse transcriptase was heat-inactivated (table 2.11). cDNA was stored at −20 °C until further use.

#### 2.2.4.3 mRNA isolation and cDNA synthesis from CTC lysates

mRNA isolation and cDNA synthesis were performed on AdnaTest CTC lysates using the AdnaTest Detect Kit according to the manufacturer's instructions. First, oligo(dT) magnetic beads were equilibrated in lysis buffer and 20 µL beads were added to 200 µL CTC lysate. Following incubation for 10 min in rotation at room temperature, beads were captured with a magnet and washed twice with wash buffer A. Next, the beads were transferred to a new tube and washed twice in buffer B and once in ice-cold Tris. For cDNA synthesis, the beads were resuspended in 14.75 µL nuclease-free water and incubated at 65 °C for 5 min. The complete volume of beads in water was mixed with 5.25 µL cDNA synthesis master mix as shown in table 2.12 and incubated in a thermal cycler according to table 2.13. Lysates derived from Parsortix-enriched cells were processed similar to AdnaTest lysates.

Table 2.12: Composition of AdnaTest cDNA synthesis master mix.

| reagent | volume [µL] |
|---|---|
| 10X reaction buffer | 2 µL |
| dNTP mix | 2 µL |
| RNasin Ribonuclease inhibitor | 0.25 µL |
| reverse transcriptase | 1 µL |

Table 2.13: Thermal cycler protocol for AdnaTest cDNA synthesis.

| temperature | time |
|---|---|
| 37 °C | 60 min |
| 93 °C | 5 min |

Table 2.14: Composition of cDNA pre-amplification reaction.

| reagent | volume [µL] |
|---|---|
| 2X TATAA SYBR GrandMaster Mix | 25 |
| assay mix (500 nM per assay) | 5 |
| cDNA on beads | 5 |
| nuclease-free water | 15 |

#### 2.2.4.4 Pre-amplification of AdnaTest cDNA

In order to analyse a higher number of target transcripts in the small cDNA volume from the enriched CTCs, transcripts were pre-amplified in a multiplex PCR before every single transcript was quantified by real-time qPCR. Due to the pre-amplification step, quantification of the transcripts was only possible in a semi-quantitative manner. Each pre-amplification reaction was set up in 50 µL with TATAA SYBR GrandMaster Mix, 5 µL cDNA on beads and a mix of the single assays with a final concentration of 50 nM per assay as shown in table 2.14. Reactions were carefully mixed by pipetting and not centrifuged to keep the beads in suspension. cDNA was preamplified in 18 cycles according to the protocol in table 2.15. Each cDNA sample was pre-amplified with three different assays mixes. Assay mix 1 contained the single assays for *AR*, *AR-V7*, *KLK3*, *FOLH1*, *EPCAM*, *KRT19*, *TACSTD2*, *PROM1*, *CCND1*, *RAI2*, *CD45*, *ACTB*, *GAPDH*, *HPRT1*, *ValidPrime*. In assay mix 2 the single assays for *ASCL1*, *FOXA2*, *PCSK1*, *POU3F2*, *PEG10*, *SRRM4*, *KRT6A*, *CHGA*, *CEACAM5*, *NKX2-1* were combined. *LMO3*, *SOX11*, *LIN28B*, *ACTL6B*, *NKX3-1* were mixed for assay mix 3.

#### 2.2.4.5 Quantitative PCR

cDNA was detected and quantified by real-time qPCR using the TATAA GrandMaster Mixes. For cell line-derived cDNA, 10 µL reaction mixes were prepared ac-

Table 2.15: Thermal cycler protocol for AdnaTest cDNA synthesis.

| | | |
|---|---|---|
| 95 °C | 3 min | |
| 95 °C | 15 s | |
| 60 °C | 60 s | repeat 18x |
| 72 °C | 20 s | |

Table 2.16: Composition of realtime qPCR mixtures for cell-line derived cDNA samples.

| reagent | volume [µL] |
|---|---|
| 2X TATAA GrandMaster Mix | 5 |
| assay ($10\,\mu M$ per primer) | 0.4 |
| cDNA (diluted 1:10) | 1 |
| nuclease-free water | 3.6 |

Table 2.17: Composition of realtime qPCR mixtures for CTC-derived cDNA samples.

| reagent | volume [µL] |
|---|---|
| 2X TATAA GrandMaster Mix | 12.5 |
| assay ($10\,\mu M$ per primer) | 1 |
| cDNA (diluted 1:8) | 5 |
| nuclease-free water | 6.5 |

cording to table 2.16. AdnaTest pre-amplified cDNA was diluted 1:8 in nuclease-free water and $25\,\mu L$ reaction mixes were prepared to allow for a larger input volume of the template (table 2.17). All probe-based assays were used in combination with the TATAA Probe GrandMaster Mix and the remaining assays were analyzed using the TATAA SYBR GrandMaster Mix. The ValidPrime qPCR required a minimally different reaction mix as shown in table 2.18. Set up of the interplate calibrator reaction is shown in table 2.19. The thermocycler protocol was the same for all reactions and is shown in table 2.20.

Thresholds were calculated automatically by the *CFX Manager* software and melt curves were checked for the SYBR-based assays. Cell line samples were measured in triplicates while patients samples, due to restriction of available material, were measured with one reaction per assay. For quantification of gene expression in cell lines, $\Delta C_q$ values were calculated by subtracting the mean $C_q$ value of the reference genes *ACTB* and *HPRT1*. ΔCq values were visualized as reversed values ($19 - \Delta C_q$) and missing values were set to 0.

Table 2.18: Composition of realtime qPCR mixtures for ValidPrime detection of genomic DNA.

| reagent | volume [µL] |
|---|---|
| 2X TATAA Probe GrandMaster Mix | 12.5 |
| ValidPrime primer mix ($10\,\mu M$ per primer) | 1 |
| ValidPrime probe ($10\,\mu M$) | 0.5 |
| cDNA (diluted 1:8) | 5 |
| nuclease-free water | 6.5 |

Table 2.19: Composition of realtime qPCR mixtures for the interplate calibrator.

| reagent | volume [µL] |
|---|---|
| 2X TATAA Probe GrandMaster Mix | 12.5 |
| IPC primer mix (µM per primer) | 5 |
| IPC probe (10 µM) | 2.5 |
| IPC template | 2 |
| nuclease-free water | 3 |

Table 2.20: Thermal cycler protocol for realtime qPCR and melting curve.

| | | |
|---|---|---|
| 95 °C | 3 min | |
| 95 °C | 15 s | repeat 40x |
| 60 °C | 20 s | |
| 72 °C | 20 s | |
| 60 °C to 95 °C | 10 min | |

PCR efficiency was calculated from the linear regression parameters of a four step dilution series of gBlock or cDNA according to the following formula:

$$\% \, PCR \, efficiency = ( \, 10^{\frac{-1}{slope}} \, - 1 \, ) \, * \, 100$$

For data analysis, single plates were first merged into one large data set using the interplate calibrator (IPC). All raw $C_q$ values were corrected by first subtracting the the average $C_q$ of the IPC triplicate on the respective plate and after that adding the average IPC $C_q$ value of all plates analyzed, as shown below.

$$C_q \, corrected = C_q \, raw - C_q \, IPC \, intraplate \, average + C_q \, IPC \, interplate \, average$$

Next, all $C_q$ values were corrected for potential contamination with genomic DNA (gDNA) using the ValidPrime Kit. First, the $C_q$ values of the reverse transcriptase negative control were calculated using the $C_q$ value of the gene of interest (GOI) on gDNA, and the $C_q$ value of the ValidPrime primer on the sample. The assays for CD45, ACTB, EPCAM, TACSTD2, POU3F2, PEG10, KRT6A, HOXB13, LMO3, SOX11, LIN28B and NKX3-1 required correction. For assays without a signal on gDNA and samples without detection of gDNA by the ValidPrime assay this step was omitted.

$$C_q RT^{-} control = C_q GOI \, gDNA + (C_q ValidPrime sample - C_q ValidPrime gDNA)$$

The calculated RT⁻ control was used to correct the $C_q$ values of the respective samples and GOIs for the background signal from gDNA.

$$C_q\, corrected = -log_2(2^{-C_q\, GOI\, RT^+} - 2^{-C_q\, GOI\, RT^-})$$

According to the AdnaTest manual a cut-off of 35 was chosen and all corrected $C_q$ values $\geq 35$ were defined negative. For data normalization, $C_q$ values for the housekeeping genes were first checked for stability of expression using *NormFinder*. *NormFinder* analysis revealed that the combination of GAPDH and ACTB had the best stability in gene expression. $\Delta C_q$ values were calculated by subtracting the $C_q$ value of the housekeeping genes from the $C_q$ value of the GOI. Based on the results from the blood samples of ten healthy individuals, a threshold for marker positivity was defined as the mean $\Delta C_q$ value minus one standard deviation to corrected for potential signals from contaminating leukocytes. Normalized and healthy donor corrected $C_q$ values were reversed by subtraction from 26 (rounding up the highest $\Delta C_q$ value) and transcripts without a signal were set to 0.

### 2.2.5 Immunofluorescence staining

Immunofluorescence staining was performed on cytospins derived from Parsortix-enriched CTC samples. As controls, cell line suspensions were admixed with PBMCs and cytospins were similarly prepared. All cytospins were stored at $-80\,°\mathrm{C}$. For staining, cells were encircled with a hydrophobic barrier pen and incubated with $4\,\%$ para-formaldehyde in PBS for $10\,\mathrm{min}$. Slides were washed three times for $3\,\mathrm{min}$ with PBS. The cell membrane was permeabilized in $0.2\,\%$ triton-X 100 in PBS for $20\,\mathrm{min}$ and cells were washed three times as before. Unspecific antibody binding was blocked with $10\,\%$ AB-serum in PBS for $30\,\mathrm{min}$. Cells were stained with a cocktail of the directly conjugated antibodies (1:200, only anti-CD45 1:100) for and the DNA dye DRAQ5 (1:1000) at room temperature for $60\,\mathrm{min}$. The cells were washed and mounted with one drop Fluoromount-G and covered with cover slides. Slides were allowed to air-dry in the dark for $30\,\mathrm{min}$ and stored at $4\,°\mathrm{C}$ until imaging. For all incubation steps longer than $20\,\mathrm{min}$, slides were kept in a humid chamber.

All slides were scanned at a fluorescence microscope at 20X magnification. Sus-

picious cells with positive CK signal were captured at 40X magnification. Cells were considered to be CTCs, when they had an intact nucleus, positive CK staining and negative CD45 staining. The CK staining was required to surround the nuclear staining.

### 2.2.6 Statistical analyses

Differential gene expression between the groups was analyzed using Kruskal-Wallis test and the Wilcoxon rank sum test together with Dunn's correction for multiple testing for pair-wise comparison. Chi-squared test and Fisher's exact test were used to assess differences among categorical variables, such as positivity of a certain marker per group. Hierarchical clustering on published data sets was performed with the *complete* method and *pearson* as distance measure, while clustering of patient samples based on gene expression profiles was performed with the *ward.D* method and *euclidean* as distance measure. For the random forest analysis, *mtry* was set to 2, *importance* was set to *impurity* and *splitrule* was set to *extratrees*, while default values were kept for all other parameters. Leave-on-out cross-validation was included to calculate the out-of-bag error rate. P-values are indicated as $*$ for $P < 0.05$, $**$ for $P < 0.01$, $***$ for $P < 0.001$ and $****$ for $P < 0.0001$.

# 3 Results

## 3.1 Selection and validation of the marker panel

The leading hypothesis of this study was that the gene expression patterns of CTCs isolated from advanced PCa patients can be used as a biomarker to identify disease subtypes with impact on treatment decisions. To identify patients with NEPC based on their CTC transcriptome profile, suitable markers first had to be identified and PCR assays had to be selected and validated. Potential marker transcripts were selected following literature review. The marker panel was afterwards validated on PCa cell lines representing different disease stages and on published RNA sequencing data sets of mCRPC patients with known histology. The assays were also technically validated and tested for potential interference in multiplex PCR and their sensitivity was tested in controls with spike-in tumor cells to validate the complete workflow of CTC enrichment, RNA isolation and PCR-based detection.

### 3.1.1 Selection of a transcript panel

The first step in assay development was to select a suitable marker panel for multiplex analysis. The epithelial markers *EPCAM* and *KRT19* were included for the purpose of classifying samples as CTC positive or negative [152]. Next, well-known prostate-specific transcripts were chosen to serve as markers for CTC positivity but also to monitor the activity of the AR pathway. These included *AR*, *AR-V7* and their downstream targets *KLK3* and *NKX3-1*. *HOXB13* and *FOLH1* are further markers of luminal cell differentiation with the latter also being an important marker in theranostics [153], [154]. *RAI2* was included to study its role in the loss of AR signaling dependency [153].

To identify transcripts that indicate NE differentiation or that are involved in

NET, recent literature was searched in PubMed (National Center for Biotechnology Information) using the terms *neuroendocrine prostate cancer* and *small cell prostate cancer*. Genes were chosen for which a functional role in the NET process was proven or which were identified as differentially expressed genes in data sets of patients with PRAD versus NEPC [70]. Next, expression of those genes in blood and immune cells was examined in *The Human Protein Atlas* (www.proteinatlas.org) [155]. In the section *IMMUNE CELL TYPE EXPRESSION (RNA)*, markers were only selected if they were classified as *Not detected in immune cells*. This was required as CTC-enriched samples still contain a considerable number of leukocytes that would interfere with the detection in the CTCs. As a consequence, several well-known markers and drivers of NET had to be excluded, such as *SYP*, *NCAM1*, *EZH2* or *REST*, because they are regularly expressed in the contaminating fraction of leukocytes.

Table 3.1: List of transcripts selected for gene expression analysis in CTCs.

| **Gene symbol** | **expr. in NEPC** | **function** | **expr. in PBMCs** | **reference** |
|---|---|---|---|---|
| ACTL6B | up | chromatin remodelling | absent | [79] |
| ASCL1 | up | lineage-pioneering TF | absent | [74] |
| CCND1 | down | cell cycle progression | detected in single | [156] |
| CEACAM5 | up | cell adhesion molecule | absent | [88] |
| CHGA | up | secretory protein | absent | [74] |
| EGFR | up | receptor tyrosine kinase | absent | [157] |
| FOXA2 | up | lineage-pioneering TF | absent | [75] |
| KRT6A | down | cytoskeletal protein | absent | [41] |
| LIN28B | up | mRNA maturation | absent | [158] |
| LMO3 | up | TF | absent | [41] |
| NKX2-1 | up | TF | absent | [159] |
| PCSK1 | up | pro-hormone convertase | absent | [41] |
| PEG10 | up | cell cycle progression | absent | [84] |
| POU3F2 | up | TF | absent | [76] |
| PROM1 | up | membrane protein | absent | [160] |
| SOX11 | up | TF | absent | [161] |
| SRRM4 | up | splice factor | absent | [162] |
| TACSTD2 | up | membrane receptor | detected in some | [68] |

Genes with no expression in leukocytes selected for the NET panel are shown in table 3.1 and include the TFs *ASCL1*, *FOXA2*, *POU3F2*, *NKX2-1*, *LMO3* and *SOX11* as well as the epigenetic regulators *PEG10*, *SRRM4*, *LIN28B* and *ACTL6B*

[41], [74]–[79], [82]–[84]. *CEACAM5, PCSK1* and *CHGA* were included as markers for NE differentiation [41], [74], [88]. *CCND1* was included as a negative marker, that is down-regulated upon NET [156]. In addition, *TACSTD2* and *PROM1* were chosen as stemness markers [68], [160]. *KRT6A* was chosen as a marker for squamous cell differentiation that had been observed in DNPC and *EGFR* is a known resistance marker for chemotherapy [41], [157].

### 3.1.2 Validation of marker transcripts in PCa cell lines

The panel of 27 markers, designed based on literature research, was first tested on PCa cell lines to validate the expression in different PCa phenotypes, such as NEPC. The LNCaP cell line represents hormone-sensitive PCa, while VCaP cells represent CRPC [141], [142]. PC3 and DU145 cells resemble double-negative, AR-independent PCa [144], [145]. NCI-H660 and LASCPC-01 served as examples for NEPC [143], [146]. Because of the limited number of NEPC cell lines, the small cell lung cancer cell line NCI-H209 served as an additional reference for the NE phenotype [147]. Relative gene expression of the markers in all seven cell lines is shown in a heatmap (fig. 3.1).

The overall expression of the tested genes showed a remarkable similarity between the two NE cell lines NCI-H209 and NCI-H660. At the same time, their expression profiles were most different from the other PCa cell lines, as both highly expressed the NE markers but had low to absent expression of *AR* or its related genes. The expression of epithelial markers and especially *EPCAM*, which is used for CTC enrichment, was not reduced in the NE cell lines. Similarly, the genetically modified LASCPC-01 cell line showed low expression of *AR* and its related genes, while several NE markers such as *LIN28B* and *SOX11* were induced. In contrast, the only hormone-sensitive cell line, LNCaP, could clearly be distinguished by its high expression of AR signaling pathway genes and the absence of NE markers. VCaP cells have a CRPC phenotype and maintained the expression of *AR* accompanied by focal weak expression of NE markers such as *CEACAM5*. PCa cell lines representing DNPC, such as DU145 and PC3, showed a decrease in the AR pathway in comparison to LNCaP and a focal increase in NE marker expression such as *FOXA2* or *PEG10*. Interestingly, DNPC cells maintained a high expression of *CCND1* and *NKX3-1*, which indicated their PRAD origin. Overall, analysis of the marker panel in PCa cell lines showed that the selected genes were sufficient to differentiate between the represented disease stages of PCa. Especially the expression of NE markers was clearly enriched in

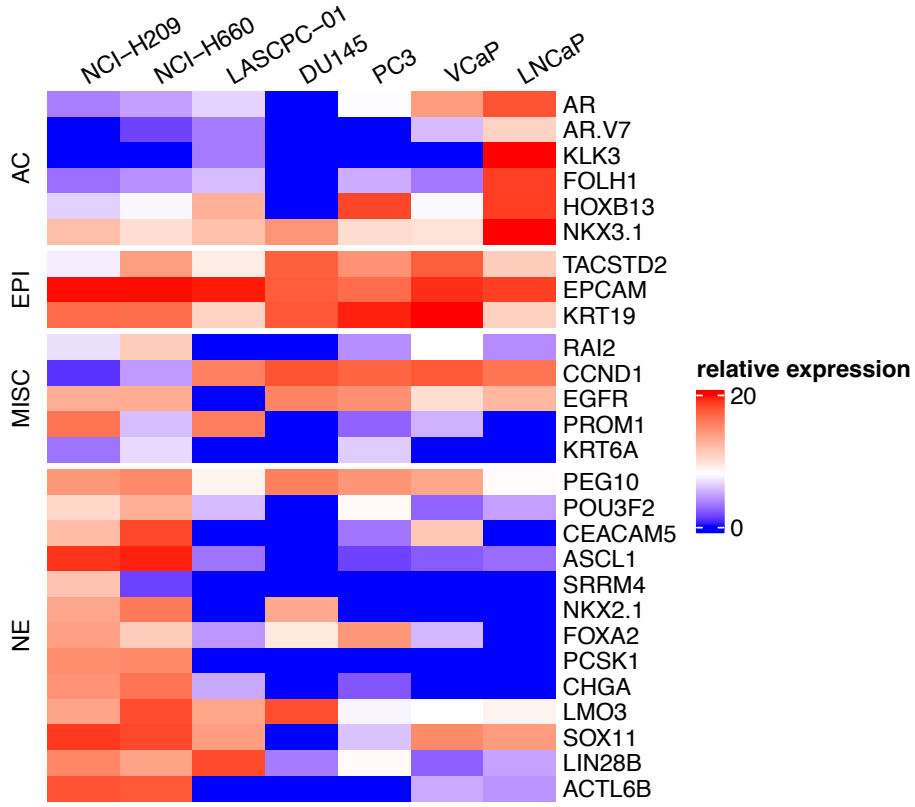



Figure 3.1: **Expression of the marker panel in PCa cell lines.** Gene expression in cell lines was measured by qPCR and the relative, normalized gene expression is shown as reversed $\Delta C_q$; AC – adenocarcinoma, EPI – epithelial, MISC – miscellaneous, NE – neuroendocrine.

NCI-H660 and NCI-H209, whereas PRAD marker expression was highly abundant in hormone-sensitive LNCaP cells. Thus, the analysis of the marker panel in PCa cell lines illustrated that different stages of PCa show unique expression patterns that allow to differentiate between the cell lines.

### 3.1.3 Validation of marker transcripts in published PCa patient tissue data sets

To test the marker panel on human PCa tissue samples, two published RNA sequencing data sets were downloaded from *cBioPortal*. The Metastatic Prostate Adenocarcinoma data set contained 208 mCRPC samples with available gene expression data [51]. NEPC scores were assigned to single samples and with a

classification threshold of 0.4, 10 samples were classified as NEPC [135]. The Neuroendocrine Prostate Cancer data set contained 49 samples of which 15 were classified as NEPC [135].

Hierarchical clustering based on the selected marker panel showed a precise grouping of the NEPC samples in the Metastatic Prostate Adenocarcinoma data set with all cases grouped into one cluster (fig. 3.2 B). NEPC samples showed a high expression of the selected NE markers and reduced expression of the prostate-specific markers. The majority of NEPC samples was positive for epithelial markers, but negative for *CCND1* and the stemness marker *TACSTD2* while *PROM1* was up-regulated in some samples.

In the Neuroendocrine Prostate Cancer data set, NEPC samples mostly clustered together, while four samples clustered with PRAD (fig. 3.2 A). In these samples, expression of NE markers was mostly absent, but three out of the four NEPC samples were positive for at least one of the NE markers. Similar to the Metastatic Prostate Adenocarcinoma data set, NEPC samples were predominantly positive for epithelial markers, negative for *CCND1* and *TACSTD2*, but positive for *PROM1*.

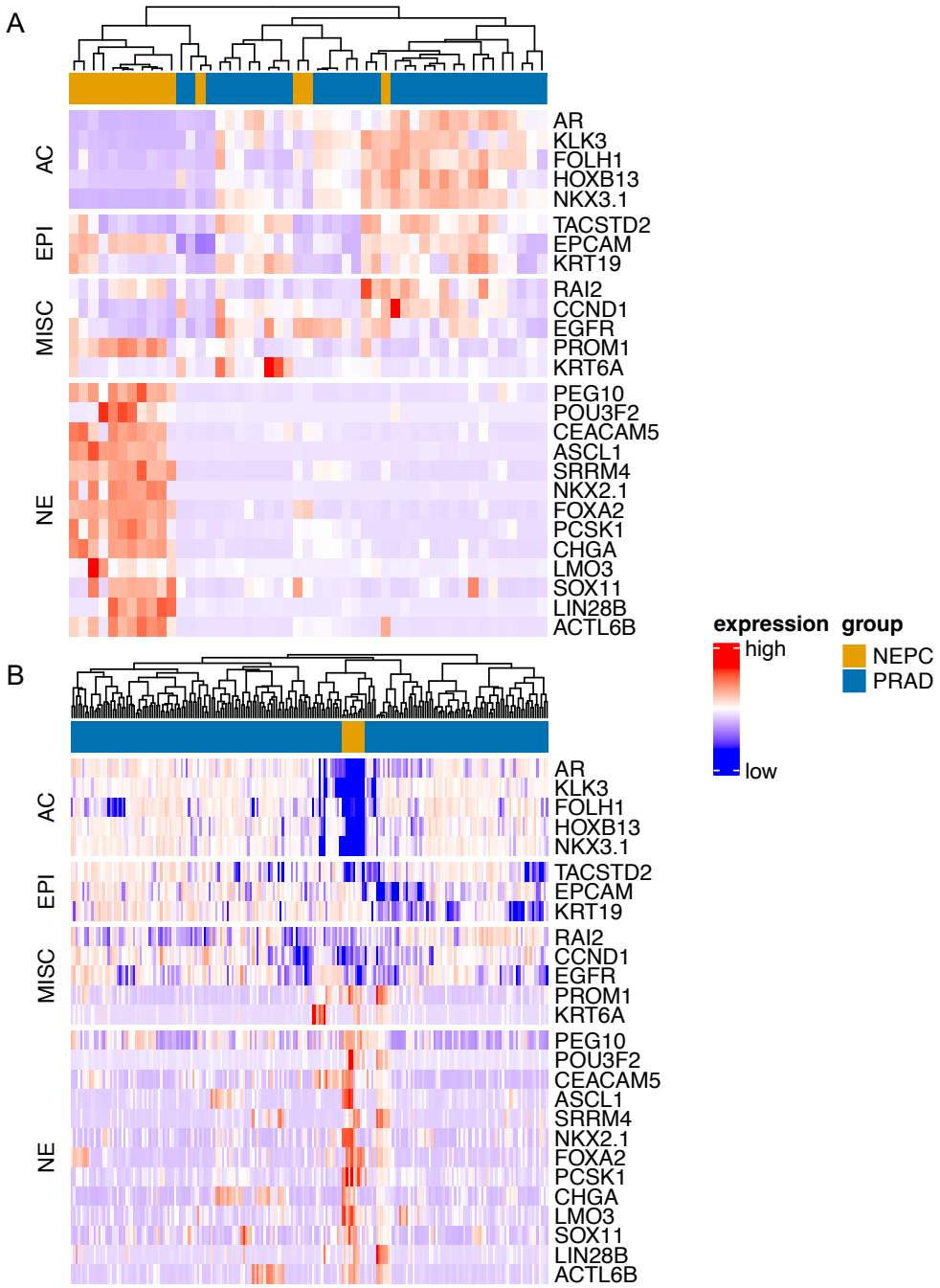



Figure 3.2: **Expression of the marker panel in published PCa datasets.** Gene expression data were downloaded from cBioPortal and annotated as NEPC or PRAD; hierarchical clustering was performed based on the selected marker panel; A: Neuroendocrine Prostate Cancer (Multi-Institute, Nat Med 2016); B: Metastatic Prostate Adenocarcinoma (SU2C/PCF Dream Team, PNAS 2019); AC – adenocarcinoma, EPI – epithelial, MISC – miscellaneous, NE – neuroendocrine.

When the the marker panel was used to predict the sample group in a supervised learning approach, a random forest model allowed for good classification of sam-

ples in the Neuroendocrine Prostate Cancer data set with an out-of-bag error rate of 8.16 % and an AUC of 96.47 %. In the Metastatic Prostate Adenocarcinoma data set the prediction was even more accurate with an out-of-bag estimated error rate of 0.48 % and an AUC of 99.95 % (fig. 3.3).

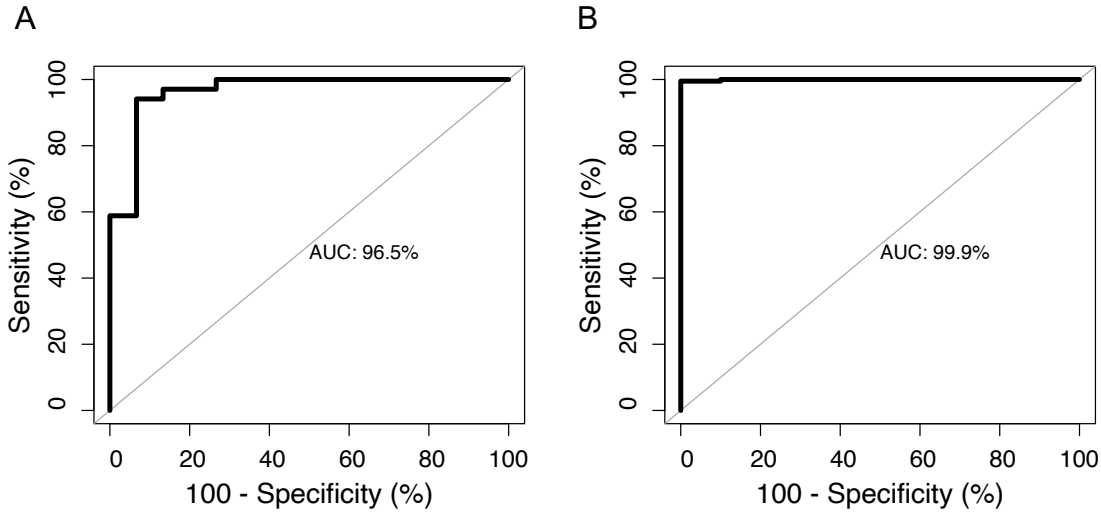



Figure 3.3: **Performance of the marker panel for classification of published PCa datasets.** Relative expression of the genes in the marker panel was used to train a random forest classifier with bootstrapping to predict the group status; performance of the classifier is shown in a receiver operator curve; A: Neuroendocrine Prostate Cancer (Multi-Institute, Nat Med 2016); B: Metastatic Prostate Adenocarcinoma (SU2C/PCF Dream Team, PNAS 2019); AUC – area under the curve.

Table 3.2: PCR efficiency of single qPCR assays.

| **assay** | **PCR efficiency** [%] |
|---|---|
| ACTL6B | 86 |
| ASCL1 | 92 |
| CEACAM5 | 86 |
| CHGA | 94 |
| FOXA2 | 96 |
| KRT6A | 93 |
| LIN28B | 102 |
| LMO3 | 97 |
| NKX2-1 | 93 |
| NKX3-1 | 96 |
| PCSK1 | 87 |
| PEG10 | 85 |
| POU3F2 | 91 |
| SOX11 | 97 |
| SRRM4 | 95 |

### 3.1.4 Technical validation of the gene expression analysis in CTCs

After validating the marker panel in appropriate luminal and NE PCa cell lines and published data sets, individual assays as well as the entire analysis pipeline had to be tested and optimized to ensure reliable detection of the respective markers also from small sample quantities which are common in CTC analysis. Assays for *AR*, *AR-V7*, *KLK3 EPCAM*, *KRT19*, *HOXB13*, *FOLH1*, *RAI2*, *EGFR*, *CCND1*, *PROM1* and *TACSTD2* are commercial assays validated by the manufacturer. These assays had successfully been used before in AdnaTest enriched samples and were therefore not subjected to PCR efficiency calculation [153]. For all other assays, PCR efficiency was determined from a four-step dilution series and the results are shown in table 3.2.

PCR efficiencies range from 85 % to 102 % with a mean PCR efficiency of 93 % which was close to an optimal PCR efficiency of 100 %.

Pre-amplification was required to measure all selected assays in the small cell numbers that are common after CTC enrichment. Therefore, the assays were tested for their compatibility in multiplex PCR that is required for pre-amplification. Each assay was subjected to pre-amplification either as single assay or in the respective assay mix (see Method section) and afterwards quantified by qPCR. The differences in $C_q$ values between single and multiplexed assays are shown in figure 3.4. The absolute differences in $C_q$ values ranged from 0.09 to 3.93 with a

mean difference of 1.03. Although five assays had absolute $\Delta C_q$ values higher than 2, the $C_q$ were smaller in the multiplexed reaction, meaning that no significant inhibition of single assays in the multiplex reaction was found and validation was continued with all markers.

| assay | $\Delta C_q$ | assay | $\Delta C_q$ |
|---|---|---|---|
| ACTB | 0.23 | KRT6A | 0.81 |
| ACTL6B | 0.17 | LIN28B | 0.88 |
| ASCL1 | 3.36 | LMO3 | 0.37 |
| CEACAM5 | 0.63 | NKX2-1 | 3.16 |
| CHGA | 0.15 | NKX3-1 | 0.34 |
| GAPDH | 0.33 | PCSK1 | 0.78 |
| EGFR | 0.42 | PEG10 | 2.07 |
| FOXA2 | 2.24 | POU3F2 | 0.16 |
| HOXB13 | 0.09 | SOX11 | 0.11 |
| HPRT1 | 0.47 | SRRM 4 | 3.93 |

Figure 3.4: **Comparison of the assay performance in single versus multiplex pre-amplification.** cDNA was pre-amplified with a single assay or a mix of assays; pre-amplified DNA was quantified by qPCR and the difference between singleplex and multiplex was calculated; $\Delta C_q$ – absolute difference between raw $C_q$ values after singleplex and multiplex pre-amplification.

As assays performed well in multiplex pre-amplification reactions, the workflow was tested in combination with mRNA isolation and cDNA synthesis from small sample quantities. For this, 25 single cells of different PCa cell lines were spiked into AdnaTest lysis buffer. Detection of the transcripts in these spike-in controls was compared to the expression level in the respective pure cell line determined before in 3.1.2. Figure 3.5 A shows the detection of the single markers for each cell line in comparison for the pure cell line and the lysis buffer spike-in control.

Gene expression showed a good concordance between pure cell lines and spike-in controls. Markers with sensitive and specific detection in the respective positive control cell line were classified as confident, while markers with unspecific results in more than one cell line were excluded from subsequent analysis of patient

samples. As multiple NE markers were measured, insufficient specificity, could cause false-positive results. The markers *KRT6A*, *POU3F2* and *SRRM4* were excluded from analysis in patient samples, because their expression was detected in spike-in controls while they were absent in the respective cell lines. Interestingly, the assays for *KRT6A* and *POU3F2* together with *PEG10* were the only assays that showed activity on genomic DNA in validation experiments conducted by *TATAA Biocenter*, suggesting that the presence of genomic DNA in the samples might have been the cause for the distorted expression between pure cell lines and spike-in samples.

After the assays had successfully been tested on lysis buffer spike-in controls, CTC enrichment was added to the analysis pipeline. During the CTC enrichment step target cells can get lost and the presence of leukocytes can introduce background signals [91]. Therefore, 25 single cells were spiked into healthy donor blood to validate the detection of the markers following CTC enrichment. The results were compared to the gene expression measured in the respective lysis buffer spike-in controls. Figure 3.5 B illustrates the comparison of the marker detection between lysis buffer and blood spike-in controls. In addition to the three markers excluded after the first spike-in experiment, *LIN28B* and *PEG10* were excluded from further analysis. Both samples showed unspecific detection in the samples enriched from healthy donor blood.

After completing the technical validation, 22 out of 27 markers were confidently detected in enriched CTC fractions. 17 out of these markers showed highly confident overlap in marker detection. The marker panel was sufficient to differentiate between PCa cell lines representing variable disease stages and and between tissue samples from PRAD and NEPC patients. An analysis pipeline was developed to robustly detect the marker expression in as few as 25 cells enriched from peripheral blood.

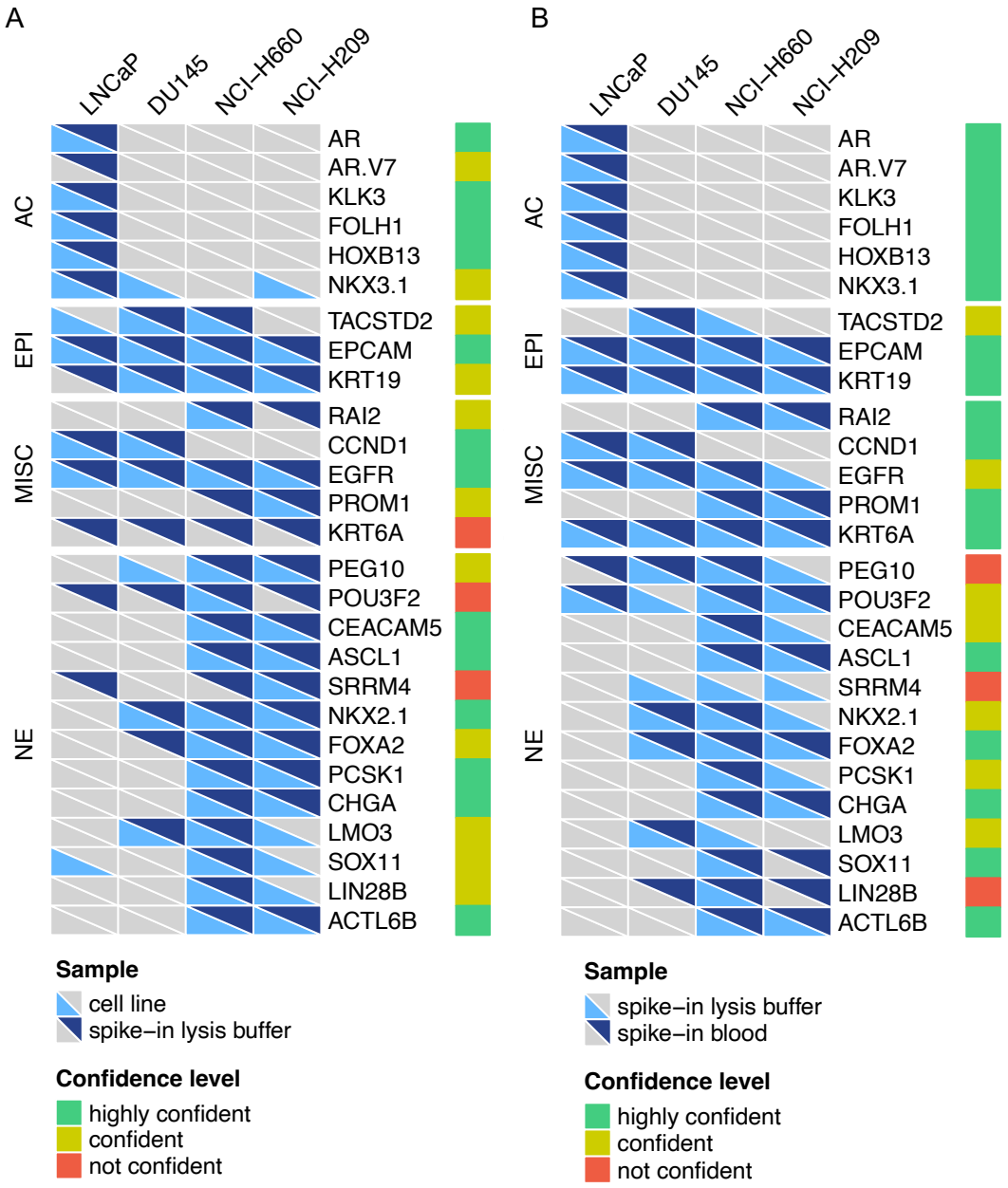



Figure 3.5: **Comparison of marker detection between pure cell lines and spike-in controls.** 25 single cells per cell line were spiked into AdnaTest lysis buffer or healthy donor blood and subjected to RNA purification, cDNA synthesis, pre-amplification and qPCR; the individual markers were classified as positive or negative based on a threshold set by means of healthy donor samples; A: comparison of the marker detection in pure cell lines and 25 cells spiked into lysis buffer, light blue – positive in pure cell line, dark blue – positive in lysis buffer spike-in; B: comparison of marker detection between 25 cells spiked in lysis buffer and 25 cells spiked in healthy donor blood and enriched by AdnaTest, light blue – positive in lysis buffer spike in, dark blue – positive in 25 cells enriched from healthy donor blood; the confidence level is based on the number of congruent results for each marker: green: all samples congruent, yellow – all but one sample congruent, red – two or more samples not congruent.

### 3.1.5 Comparison of label-dependent and label-independent CTC enrichment

Gene expression profiling strongly depends on the CTC enrichment method, as size and cell surface marker expression may vary along the different steps of the metastatic cascade and tumor progression. Consequently, one enrichment method might be more suitable for the purpose of enriching NEPC-derived CTCs. For this purpose, label-dependent CTC detection in form of the AdnaTest was compared to the size-based Parsortix analysis. Multiple blood samples were drawn from a single patient at one time point and analyzed in parallel. According to the manual, AdnaTest samples were declared positive, if one or more epithelial or prostate-specific markers were positive. For Parsortix, stained cells as well as gene expression data were available for evaluation. When gene expression data were available, criteria for CTC positive samples were similar to the AdnaTest. Positive CK staining and absent CD45 signal were required to identify CTCs in immunocytochemical analysis. Representative images of Parsortix-enriched CTCs are shown in figure 3.6. Figure 3.7 displays the comparison of both enrichment methods.

Comparing AdnaTest and Parsortix, CTCs were significantly more frequently detected with the AdnaTest with a positivity rate of 100 % in contrast to 62 % in Parsortix ($p = 0.0007$). This indicates that an EPCAM-based enrichment strategy might be more suitable than size-based enrichment in the selected disease stage. In contrast, morphological information and CTC count are lost during AdnaTest-based CTC enrichment. The representative images of CTCs shown in figure 3.6 illustrate the heterogeneity regarding the cell shape, expression of NE markers and cluster formation that is observed in Parsortix-enriched cells.

13 matched samples were available for the comparison of gene expression between AdnaTest and Parsortix. For both enrichment methods, the samples were subjected to an identical workflow of RNA isolation, cDNA synthesis and PCR-based detection. Based on ten healthy donor samples, thresholds were calculated to classify samples as marker-positive or marker-negative depending on their normalized expression data. The comparison of expression data from both techniques for 13 patient samples is shown in figure 3.7.

Regarding the detection of epithelial markers, AdnaTest and Parsortix showed mixed concordance with 56.4% of single marker-patient combinations showing a congruent results. 38.5 % of single measurements were only positive in AdnaT-

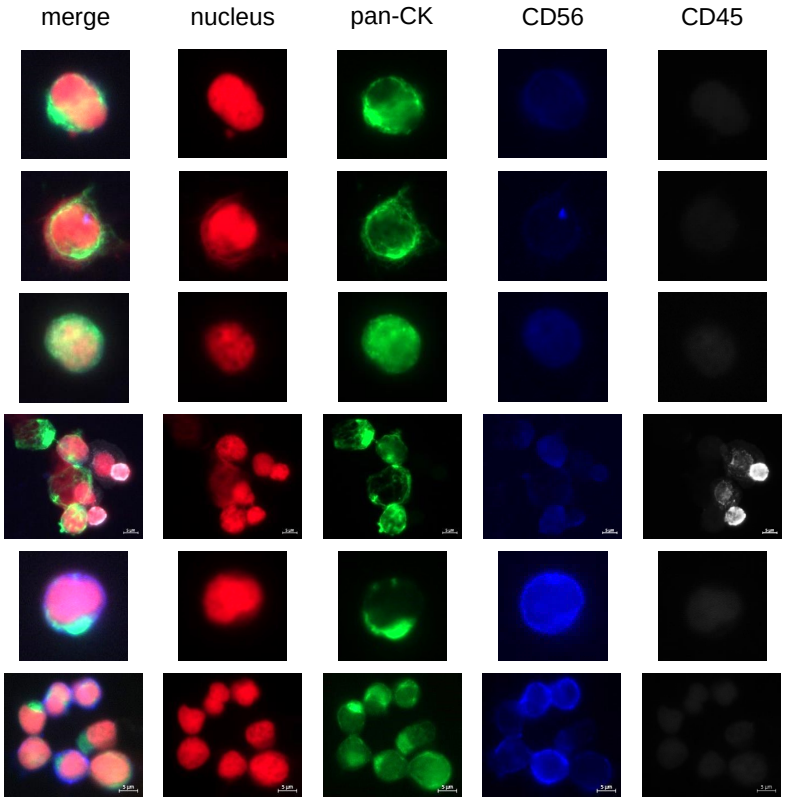



Figure 3.6: **Representative images of Parsortix-enriched CTCs.** Cytospins were prepared from Parsortix-enriched CTCs and immunofluorescence staining was used to detect pan-CK (AF488) and CD56 (BV421), CD45 (PerCP) served as negative marker and nuclei were counterstained with DRAQ5.

est, while 5.1 % were only positive in Parsortix-enriched cells. However, just one sample was only positive for CTCs in AdnaTest when combining all single markers, while most samples were positive for at least one marker in both tests. This indicates, that the AdnaTest might be more sensitive to enrich epithelial cells from blood. In contrast, CTCs might be lost in Parsortix due to small size or Parsortix enriches another cell population, that might have down-regulated epithelial cell programs. With regard to the prostate-specific markers, detection of single markers was mostly concordant with 69.2 % of congruence in the individual marker-patient combinations. With 21.8 % AdnaTest-only positive and 8 % Parsortix-only positive comparisons, more markers were detected in the AdnaTest enriched cell lysates. Comparing the overall marker expression between samples, the majority was positive for at least one prostate-specific marker with both method. Two out of thirteen samples were only prostate marker positive in AdnaTest and one samples was only positive in Parsortix.

As the detection of NE markers was among the central aims of this project, these markers were closely compared between AdnaTest- and Parsortix-enriched CTCs. In total, ten out of thirteen samples were positive for at least one NE marker. Of those, five were positive with both tests, three were only positive in AdnaTest and two samples were only positive after Parsortix enrichment. In four of the five samples that were positive with both methods, more NE markers were detected in AdnaTest. In all AdnaTest-only positive samples, more than one NE marker was detected. In contrast, both Parsortix-only positive cases were positive for a single marker. Precisely, 56.8 % of detected markers were overlapping with both methods, however, 37.8 % were only found in Adna-Test samples and 5.4 % only in Parsortix enriched samples. Thus, NE markers were significantly more frequently detected after AdnaTest enriched ($p = 0.0013$).

In summary, both CTC enrichment methods - AdnaTest and Parsortix - showed intermediate concordance with regard to CTC positivity. However, samples with discordant results were more often only positive using the label-dependent AdnaTest. Thus the respective epitopes EPCAM, EGFR and HER2 covered by the AdnaTest might increase the sensitivity of this analysis method compard to size-based Parsortix enrichment. With regard to gene expression, the AdnaTest performed better for the detection of an epithelial cell population and CTCs with NE marker expression were more frequently detected. Thus, AdnaTest was chosen for subsequent gene expression analysis in CTCs.

In conclusion, a panel of 22 markers was identified by literature research and validated *in vitro* and *in silico*. An analysis pipeline was developed to robustly detect the marker expression in as few as 25 cell enriched from peripheral blood by the AdnaTest.

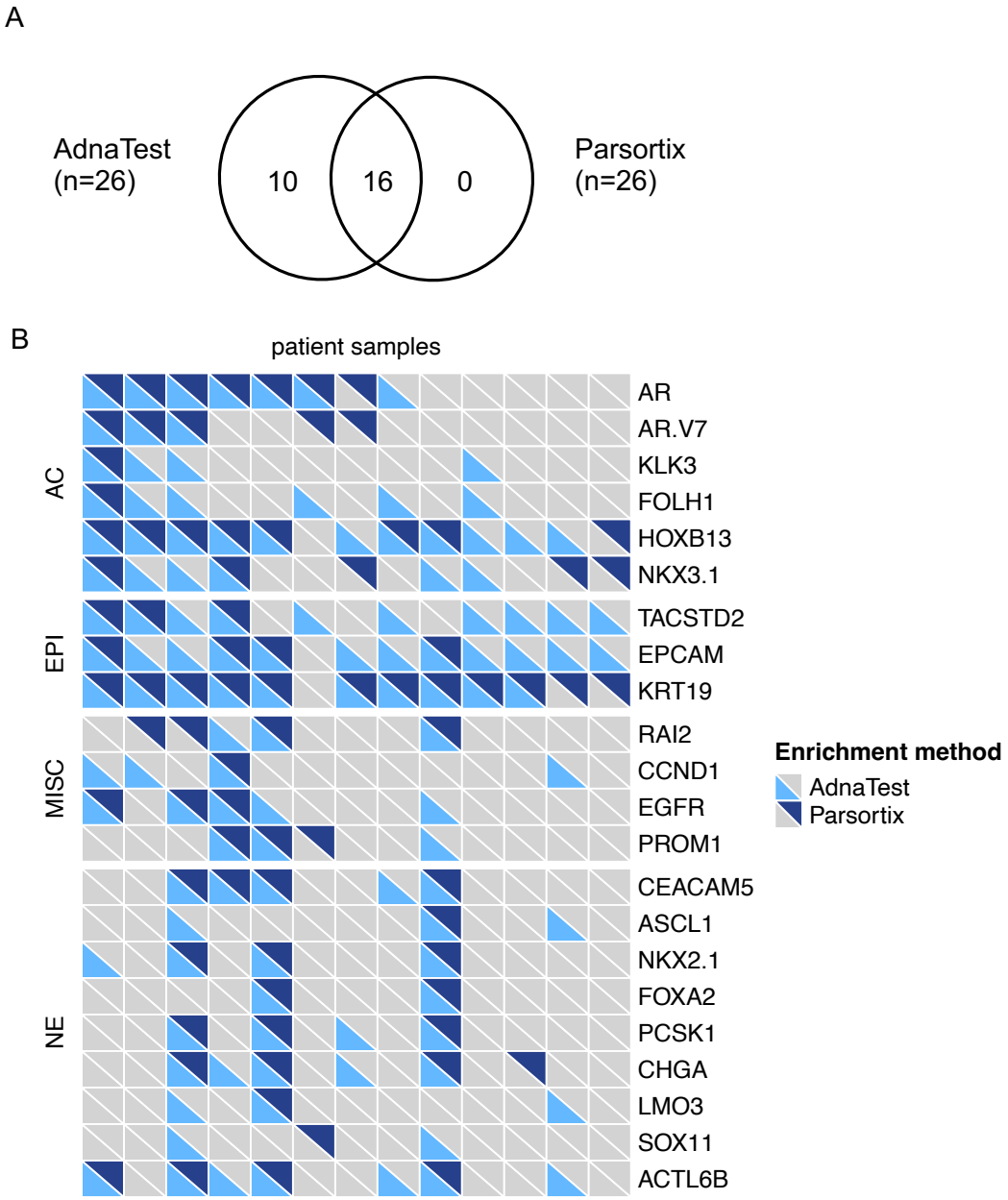
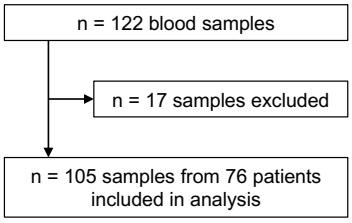



Figure 3.7: **Comparison of AdnaTest and Parsortix-enriched CTCs.** Two blood samples per patient were drawn to enrich CTCs with AdnaTest and Parsortix in parallel; A: number of CTC positive samples in AdnaTest and Parsortix (number of samples in total in brackets), CTC positivity was determined either with immunocytochemical staining of enriched cells or gene expression analysis; B: comparison of gene expression in bulk CTCs enriched by the respective enrichment method, gene expression was measured by semi-quantitative PCR after an identical workflow of RNA isolation, cDNA synthesis and pre-amplification; based on ten healthy donor samples, thresholds were calculated to classify samples as marker-positive or marker-negative depending on their normalized expression data; light-blue: marker positive in AdnaTest, dark-blue: marker positive in Parsortix.

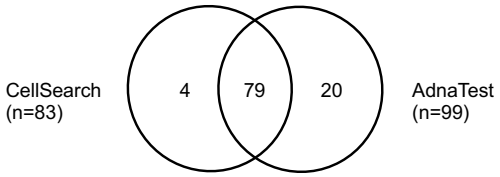
## 3.2 CTC analysis in patient samples

In total, 122 samples were collected from PCa patients and ten samples were collected from male healthy donors. Sex and age of the healthy donors was chosen to match the patient cohort. Average patient age was 65.5 years (46-88 years) and the mean age of the healthy donors was 63.6 years (59-69 years). 17 samples had to be excluded from further analysis due to missing clinical data or because no conclusive group assignment was possible. This left 105 samples from 76 patients for a comprehensive CTC analysis. An overview of the distribution of the samples between the groups and the applied CTC enrichment methods is given in figure 3.8. 99 samples were analyzed with the AdnaTest and CellSearch counts were available from 83 samples. According to their clinical manifestation, patients were assigned to one of four groups. Patients with aggressive disease according to modified Aparicio criteria, except for neuroendocrine histology, were classified as AVPC [34]. This group exhibited highly elevated median serum PSA and a slightly increased median serum NSE with a broad range of serum NSE detection between patients. Patients with histological evidence of small cell, neuroendocrine prostate cancer, pure or admixed, were defined as NEPC group. This was based on the pathology report with either positivity for SYP and/or CHGA or a small cell morphology, which are widely accepted NEPC markers [36]. This group was characterized by low serum PSA but elevated NSE and CEA. As control group, patients with metastatic hormone-sensitive PCa (HSPC) who were naive to hormonal treatment were chosen. In the HSPC group, serum PSA was elevated but NSE and CEA were not. The grouping of patients relied on a combination of clinical parameters. At the time point of blood collection, histology analysis was not available for all patients. Also for patients with multiple blood draws, histology findings were not available for all time points. As many of the AVPC classified patients had multiple metastases with expected spatial heterogeneity, the AVPC group is rather heterogeneous and not as precisely defined as the NEPC group. Overall, patients were at an advanced disease stage with several lines of treatment. The majority of patients was treated with ADT, NHA and taxane-based chemotherapy. Some patients with NEPC were treated with a combination of cabazitaxel and carboplatin. At the time point of blood collection most patients had progressive disease.

A




B




C

| | NEPC n=39 | AVPC n=52 | DNPC n=3 | HSPC n=11 |
|---|---|---|---|---|
| median age (years) | 65 (51-81) | 68.5 (46-88) | 70.5 (66-75) | 62.5 (53-80) |
| median PSA (ng/ml) | 0.36 (0.01-206) | 243.465 (0-5647) | 54.82 (37.82-71.8) | 47.8 (10.86-342) |
| median NSE (ng/ml) | 49.7 (9.5-962) | 19.6 (10.1-114.6) | 48.35 (42-57.7) | 14.7 (10.8-18.6) |
| median CEA (ng/ml) | 3 (1.2-6179) | 2.3 (0-285.8) | 1.3 | 1.3 (0-8.2) |
| median LDH (U/l) | 293 (154-2923) | 283 (152-1748) | 796 (417-1175) | 214 (167-253) |
| liver metastasis | 18 (62.0 %) | 8 (17.0 %) | 0 | 1 (12.5 %) |

Figure 3.8: **Overview of sample numbers and patient characteristics.** A: Numbers of collected and excluded samples; B: Venn diagram showing the number of samples analyzed per enrichment method; C: patient characteristics for the individual groups.

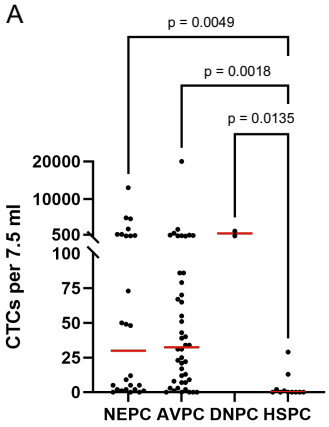
### 3.2.1 CTC enumeration with CellSearch

CellSearch is the first FDA-cleared technology for CTC enumeration and still represents the gold standard for CTC analysis. Enumeration relies on EPCAM expression for CTC enrichment and epithelial CK expression for CTC detection [163]. In total, 83 samples were analyzed by CellSearch. Potential CTCs were identified based on positive pan-CK and absent CD45 expression.

In the NEPC group, 88.5 % of samples were positive for CTCs with a mean count of 30 CTCs per $7.5\,\mathrm{mL}$ and a range of 0-13,000 CTCs. This was significantly higher than in the HSPC group with 36.4 % positive samples and a median CTC count of 0 CTCs per $7.5\,\mathrm{mL}$ ($p = 0.0049$). Similarly, the AVPC group showed a significantly increased CTC positivity of 88.6 % with a median count of 32.5 CTCs per $7.5\,\mathrm{mL}$ and a range of 0-20,000 CTCs ($p = 0.0018$). As only two samples were available for analysis in the DNPC group, the informative value of this comparison was limited, but still both samples showed exceeding CTC counts of 187 and 1480 CTCs per $7.5\,\mathrm{mL}$. In AVPC, NEPC and DNPC the median CTC count was above

the threshold of 5 CTC per $7.5\,\mathrm{mL}$ for a dismal prognosis in PCa as defined by De Bono *et al.* [101]. This is in line with the clinically observed aggressive course of disease in these patients.

A




B

| | % CTC positive | median count | range |
|---|---|---|---|
| NEPC (n = 26) | 88.5 | 30 | 0 – 13,000 |
| AVPC (n = 44) | 88.6 | 32.5 | 0 – 20,000 |
| DNPC (n = 2) | 100.0 | 833.5 | 187 – 1,480 |
| HSPC (n = 11) | 36.4 | 0 | 0 – 29 |

Figure 3.9: **CellSearch CTC counts in patient samples.** CTCs were enriched from blood samples using the automated CellSearch system; CTCs were detected and enumerated based on positive CK staining, absent CD45 staining and an intact nucleus; A: CTC count of single samples per group, red line indicates median; B: median CTC count and range per group.

As illustrated in figure 3.10, CTCs detected by CellSearch analysis showed variable morphology. Some cells were round with bright CK staining surrounding the entire nucleus. Other cells had less intense CK staining, while some were characterized by a perinuclear dot-like CK staining. In these cells, a bright spot of CK staining was observed next to the nucleus, but only a very faint staining surrounded the complete nucleus. This CK staining pattern is indicative of a small cell morphology that is associated with pure NEPC [36]. In addition, CTC clusters were detected in five samples in total. Some clusters consisted only of CTCs while in other clusters the CTCs were accompanied by immune cells. Interestingly, all of the samples with clusters had extraordinary high CTC counts of more than 100 CTCs per $7.5\,\mathrm{mL}$ blood and none belonged to the NEPC group. However, none of these features showed an association with the patient groups.

In conclusion, high CTC counts were measured in all three aggressive disease groups compared to HSPC. This illustrates the dismal prognosis of these patients, but facilitates the gene expression analysis due to the overall high abundance of relevant tumor cell populations in the blood samples.

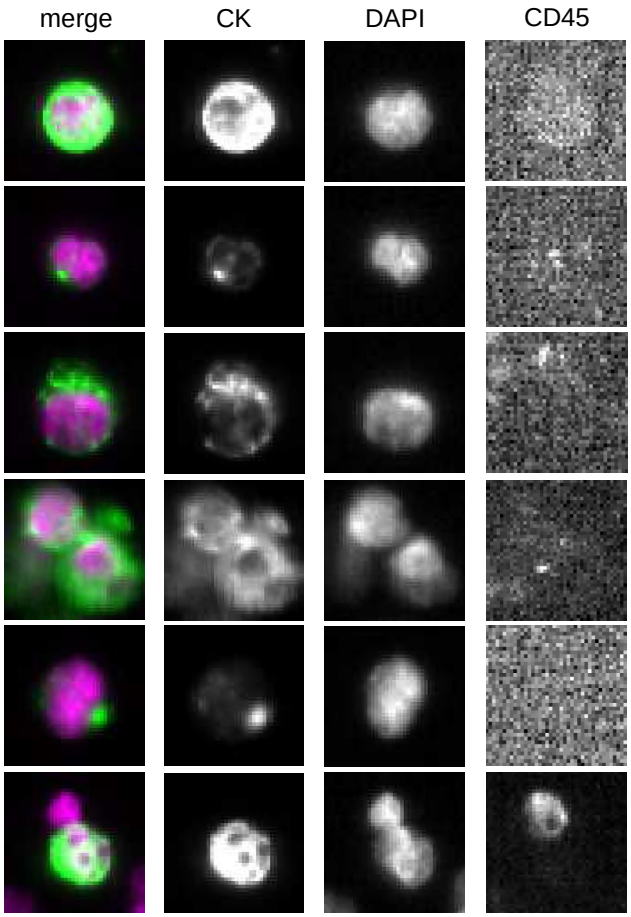



Figure 3.10: **Representative morphology of CellSearch-enriched CTCs.** CTCs were enriched from blood samples using the automated CellSearch system; CTCs were detected and enumerated based on positive cytokeratin staining, absent CD45 staining and an intact nucleus; for each event the merged image as well as the individual signals from CK, DAPI and CD45 channel are depicted.

### 3.2.2 CTC detection and gene expression profiling with the AdnaTest

In total, 99 samples were analyzed using the AdnaTest and complete expression profiles were available for 94 samples and ten healthy donors. The frequency of CTC positive samples was compared between AdnaTest and CellSearch. To get a deeper insight into gene expression patterns between the groups and the overall profiles of the single samples, the data were analyzed on the level of single genes as well as in a supervised and unsupervised learning approach.

#### 3.2.2.1 Comparison of CTC detection by AdnaTest and CellSearch

79 samples were available for the comparison of AdnaTest and CellSearch. Overall, both methods showed high positivity rates for CTCs (figure 3.11). With the AdnaTest, 95 % of samples were positive for CTCs compared to 81 % in CellSearch ($p = 0.014$). 13 samples were not concordant between both methods: One sample was negative in AdnaTest but not CellSearch, while twelve samples were CTC negative in CellSearch and positive in AdnaTest. Differences in CTC positivity rates can occur due to statistical variability in true CTC counts between two blood samples especially at low CTC numbers. However, as most non-concordant samples were positive in AdnaTest and not CellSearch, this might not be due to chance but due to the two additional antigens that are enriched with AdnaTest but not in CellSearch. Comparison of the four patient groups revealed a higher concordance of AdnaTest and CellSearch in the NEPC and AVPC groups that both share high CTC counts compared to the HSPC group ($p = 0.0021$, $p = 0.0003$). While more than 85 % of CTC positive samples in the NEPC and AVPC group were CTC positive with both enrichment methods, only 36 % of CTC positive samples in the HSPC group were positive with both methods. The other 64 % of samples were only CTC positive with the AdnaTest. This indicates that both enrichment methods robustly detect CTCs in samples with high CTC counts, while the AdnaTest more frequently detects CTC at smaller cell counts.

#### 3.2.2.2 Gene expression analysis of single markers

After the comparison to CellSearch, the AdnaTest was applied to examine the expression of the 22 marker genes that had been selected and validated before. Raw $C_q$ values were normalized by the housekeeping genes *ACTB* and *GAPDH* and for every marker, thresholds were set based on the expression in healthy

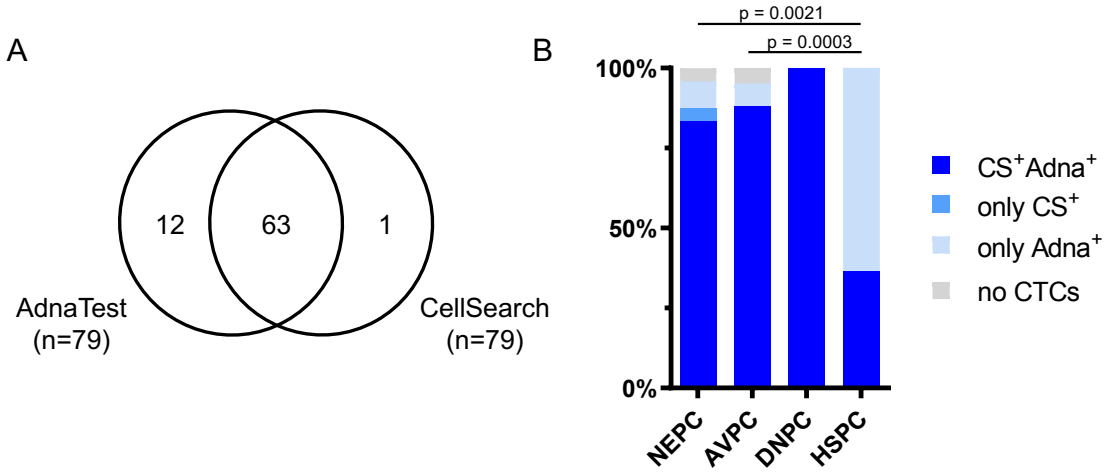



Figure 3.11: **Comparison of AdnaTest and CellSearch based CTC enrichment.** Multiple blood samples were collected from a single patient and analyzed in parallel with AdnaTest and CellSearch; A: CTC positivity was compared between both methods, the Venn diagram illustrates the numbers of CTC positive samples comparing both enrichment methods; B: comparison of the percentage of CTC positive samples between the four patient groups and the two enrichment methods; CS – CellSearch.

donors to avoid false-positive results. The normalized expression for every single marker was compared between the four patient groups (figure 3.12).

The most prominent prostate-specific marker, *AR*, was detected in most patient samples. *AR* expression in the AVPC group was significantly higher than in the NEPC and the HSPC group ($p = 0.0012$; $p = 0.0265$). Although overall expressed in fewer samples, the *AR-V7* splice variant showed a similar pattern with increased expression in AVPC compared to NEPC and HSPC samples ($p = 0.0017$; $p = 0.0015$). Gene expression of *KLK3*, *FOLH1* and *NKX3-1*, which represent targets of AR-mediated gene regulation, was detected in all groups, while the expression was significantly reduced in the NEPC patients compared to the AVPC patients ($p = 0.002$; $p = 0.0004$; $p = 0.0002$). The prostate-specific TF *HOXB13* was up-regulated in AVPC compared to HSPC ($p = 0.0455$). Interestingly, no significant differences were found between AVPC and NEPC ($p = 0.1489$), indicating that it might be a suitable marker to prove the prostatic origin of a neuroendocrine carcinoma. Taken together, these results highlight the loss of AR-dependency in the NEPC group, while the detection of AR and associated markers was maintained in the AVPC group.

The epithelial marker EPCAM was used as an epitope in CTC enrichment and,

thus, expected to be detected in all samples. Indeed, *EPCAM* was highly expressed in the majority of samples and expression showed no significant differences between the groups, which confirms the utility of the corresponding gene product for CTC enrichment. Similarly, *KRT19* and *TACSTD2* expression was not significantly altered between the groups. This indicates that the epithelial differentiation was maintained through all disease stages analyzed.

*EGFR*, which is also used for CTC enrichment in AdnaTest, was only detected in a minority of samples and no differences were identified between the groups. The cell cycle regulator *CCND1* was significantly reduced in NEPC samples compared to AVPC ($p = 0.0084$). *RAI2* was expressed in a minority of samples and showed no significant differences between the groups, similar to the stemness marker *PROM1*. While *CCND1* was identified as an additional marker for PRAD, the other three markers had no added value for the characterisation of the four groups.

In addition, significant differences between the groups were found for six out of nine NE markers. *CHGA* was one of the most abundantly expressed NE markers and its expression was significantly increased in NEPC samples compared the the HSPC and the AVPC samples ($p = 0.0065$; $p < 0.0001$). Similarly, *CEACAM5* was significantly higher expressed in NEPC samples compared to HSPC and AVPC ($p = 0.0081$; $p = 0.0003$). *ACTL6B* was another NE marker with significantly enhanced expression in NEPC samples compared to HSPC and AVPC ($p = 0.04$; $p = 0.0019$). Similarly, the transcription factor *ASCL1* was significantly increased in NEPC compared to AVPC, albeit being expressed in only the minority of samples ($p = 0.0072$). In addition, *PCSK1* and *FOXA2* were induced in NEPC compared to AVPC ($p = 0.0003$; $p < 0.0001$), although only expressed in a small number of samples. The NE markers *LMO3, SOX11* and *NKX2-1* did not show significant differences in gene expression between the groups and were overall rarely expressed.

In summary, the analysis of differential gene expression of the selected markers highlights the significant alterations in the NEPC group especially in comparison to the AVPC group. The NEPC group was mainly characterized by a downregulation of PRAD markers and expression of multiple NE markers.

Next, the individual markers were analyzed for the correlation of their expression to identify clusters of markers with similar or opposite expression patterns. The results of the correlation analysis are shown in figure 3.13.

A strong positive correlation was observed for prostate-specific markers *FOLH1*, *AR*, *AR-V7*, *KLK3*, *NKX3-1* and *HOXB13* as well as the epithelial markers *KRT19*, *EPCAM* and *TACSTD2*. This observation is supported by the fact, that the prostate-specific markers depend on each other for their gene expression. It indicates that all prostate-specific markers are equally well suited to identify the adenocarcinoma differentiation. The cell cycle regulator *CCND1* was also positively correlated with that cluster, suggesting that its expression is characteristic of a prostate luminal cell differentiation. The NE markers also formed a defined cluster, although less clearly delimited and with an overall lower degree of correlation. Within the NE marker cluster, *CHGA* and *NKX2-1* showed the strongest correlation to the other NE markers, and even weak correlation to some of the PRAD markers. Interestingly, only a fraction of the NE markers including *FOXA2*, *PCKS1* and *ACTL6B* showed significant anti-correlation to a number of PRAD markers including *AR*, *KLK3* and *FOLH1*. This might suggest, that different profiles of NE marker expression can be observed comparing PRAD-high and PRAD-low NEPC samples. Especially *PCSK1* and *FOXA2*, seem to be associated with an AR-low phenotype, while the markers *CEACAM5*, *CHGA* and *NKX2-1* might as well be correlated with double-positivity in this study cohort. Additionally, the stemness-marker *PROM1* was correlated with the NEPC exclusive markers and anti-correlated to some PRAD markers. Thus, PROM1 might be involved in the lineage plasticity that is required for complete transdifferentiation. In contrast, *EGFR* showed rather strong correlation with both NE and PRAD markers. It might therefore be associated with the double-positive phenotype or with highly heterogeneous tumors in general. Similarly, *RAI2* was positively correlated with *CHGA* and *NKX2-1* among other NE markers but not with PRAD markers. Thus, *RAI2* expression might as well be associated with a heterogeneous double-positive phenotype in PCa. Overall, the NE markers seemed to be more variable in their expression and did not form a well-defined cluster as seen for the PRAD markers. This indicates that tumor cells undergoing NET do not simply switch from one differentiation state to another, but are rather acquiring a heterogeneous state that is dominated by plasticity.

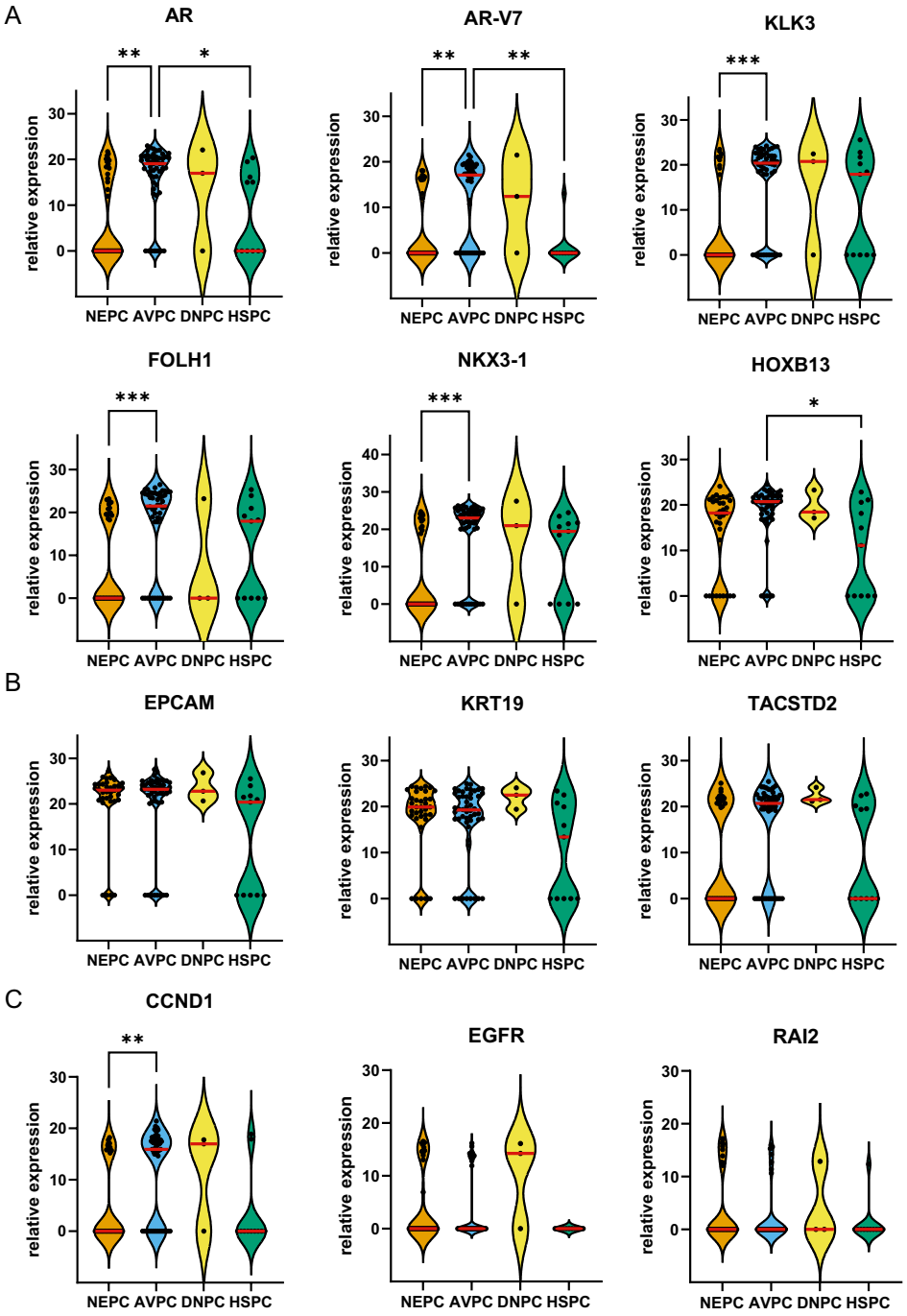



Figure 3.12: **Normalized gene expression of single markers in enriched CTCs.**

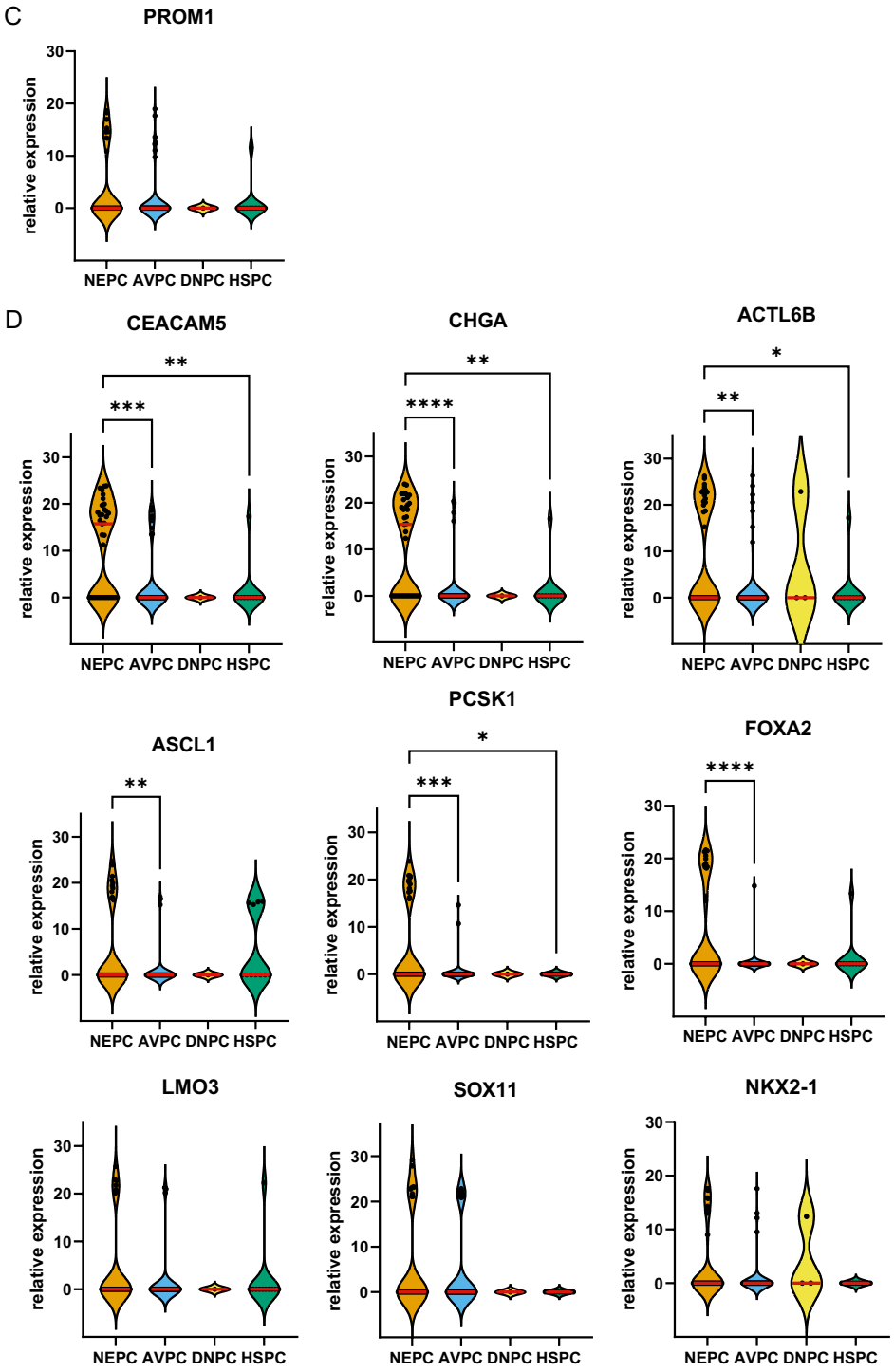



Figure 3.12: **Normalized gene expression of single markers in enriched CTCs, continued.** CTCs were enriched from blood samples and gene expression was measured by semi-quantitative PCR following RNA isolation, cDNA synthesis and pre-amplification; gene expression is shown as reversed normalized $C_q$ with healthy donor based thresholds applied; violin plots show all single samples as dots with the median expression per group indicated as red line; A: adenocarcinoma markers, B: epithelial markers, C: miscellaneous markers, D: neuroendocrine markers.

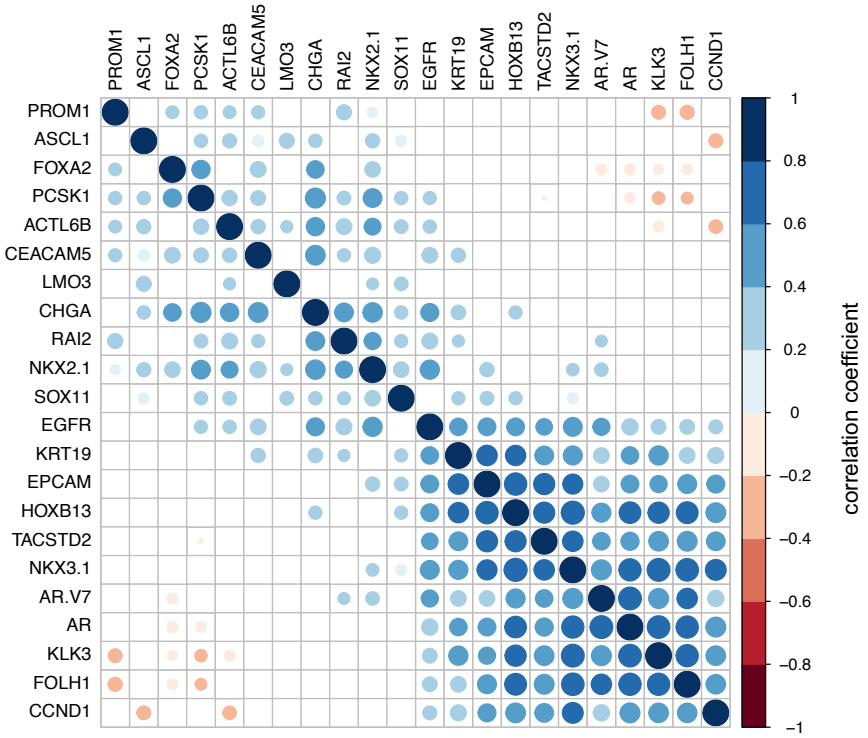



Figure 3.13: **Correlation of the gene expression of single markers.** Normalized gene expression of all marker transcripts irrespective of the patient group was subjected to correlation analysis by Spearman correlation coefficient; significant correlations are shown as dots with the colour resembling the correlation coefficient: blue – positive correlation, red – negative correlation.

#### 3.2.2.3 Unsupervised analysis of gene expression profiles

Expression analysis of single genes revealed significant differences between the clinical groups and correlation analysis indicated distinct profiles of NE and PRAD markers. Thus, the next step was to examine the gene expression profiles of all samples in an unsupervised analysis. For this, 94 samples with a complete expression profile were subjected to hierarchical clustering. The resulting heatmap and clusters are depicted in figure 3.14.

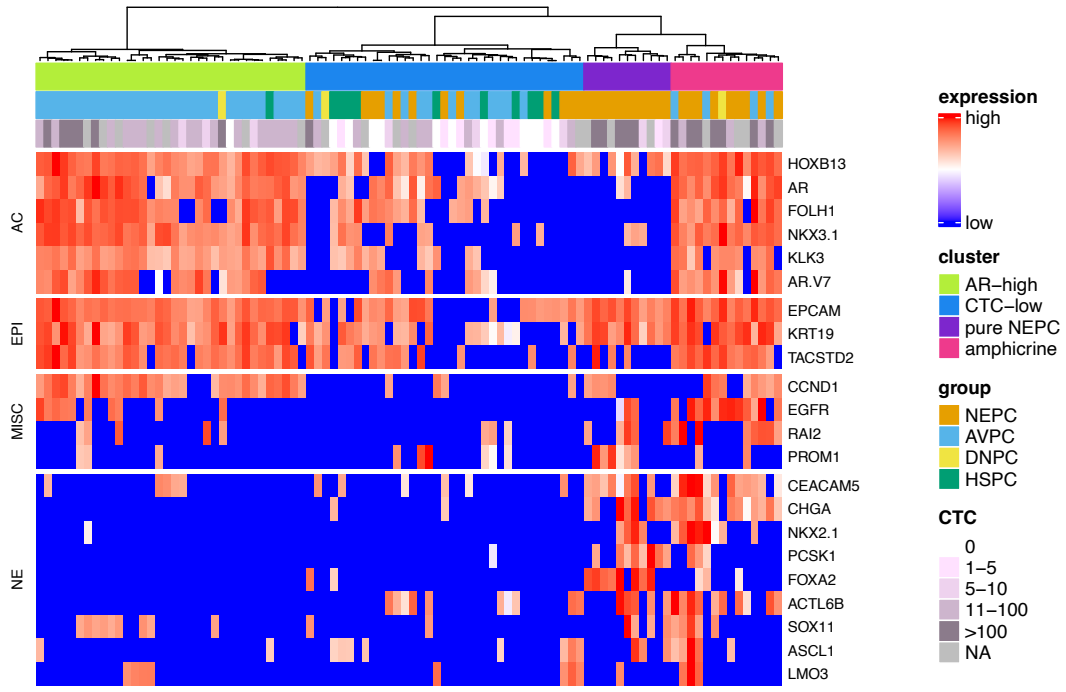



Figure 3.14: **Hierarchical clustering of patient samples based on gene expression profiles.** Normalized gene expression data were scaled and subjected to hierarchical clustering; based on the dendrogram, samples were split into four clusters; sample cluster, group and CTC count per 7.5 mL are displayed above the heatmap.

The hierarchical clustering defined four clusters. The first cluster comprised 34 samples and the expression profile was dominated by a high expression of the PRAD markers and a high expression of epithelial markers. Although some samples showed expression of one to two NE markers, including *CEACAM5, SOX11* and *LMO3*, NE marker expression was almost absent in samples of this cluster. Thus, this cluster was termed the $AR^{high}$ cluster. Samples in this cluster predominantly showed clinical characteristics of AVPC patients and none of the samples

belonged to the NEPC group. Samples in the AR$^{high}$ cluster had high CTC counts with a median of 40 CTCs per $7.5\,\mathrm{mL}$ of blood. In addition, this cluster was characterized by high expression of *CCND1* and some samples also expressed *EGFR*. The second cluster included 35 samples with mixed expression of PRAD markers and reduced to absent expression of epithelial markers. In combination with the low median CTC count of 1.5 CTCs per $7.5\,\mathrm{mL}$ of blood, this cluster was designated the CTC$^{low}$ cluster. Consistent with this, the NE markers were as well negative in this group, except for a few samples with expression of especially *CEACAM5* and *ACTL6B*. Regarding the clinical group, the CTC$^{low}$ cluster showed a mixed composition with samples of all groups. This underscores that this cluster does not reflect a clinical subtype but rather an overall CTC$^{low}$ phenotype that is independent of the clinical phenotype. With eleven samples, the third cluster was the smallest of the four clusters. Its expression profile was characterized by an almost absent expression of PRAD markers except for *HOXB13*. The epithelial markers *KRT19* and *EPCAM* were still highly expressed, although *TACSTD2* was down-regulated in comparison to the AR$^{high}$ and the amphicrine cluster ($p = 0.006$, $p = 0.0044$). NE markers were highly expressed in this group with all samples being positive for multiple markers. In particular, the markers *FOXA2* and *PCSK1* were specific for this cluster showing significant increase in comparison to the other three clusters ($p < 0.0001, p < 0.0001$). Therefore, this cluster was considered the pure NEPC cluster. In addition, the stemness-associated marker *PROM1* was significantly enriched in this cluster compared to the AR$^{high}$ and the amphicrine cluster ($p = 0.0028$, $p = 0.0401$). In line with the expression profiles, cluster three only contained samples with histological evidence of NEPC. Patients with AVPC and HSPC were absent from the NEPC cluster, indicating the high specificity of this gene expression signature. For the cases assessed, CTC counts were very high in the NEPC cluster with a median count of 381 CTCs per $7.5\,\mathrm{mL}$ blood. Lastly, the fourth cluster with 14 samples exhibited a double-positive gene expression profile. Next to high expression of epithelial markers, all samples showed intense expression of PRAD markers and positivity for multiple NE markers, predominantly *CHAG*, *CEACAM5* and *ACTL6B*. Thus, this cluster was designated the amphicrine cluster, as it showed a combination of luminal and NE marker expression. Further, this cluster was characterized by a significantly increased expression of *EGFR* compared to all other clusters ($p < 0.0001$). Similar to the AR$^{high}$ cluster, the samples in the amphicrine cluster were mostly positive for *CCND1*. Compared to the AR$^{high}$ and the CTC$^{low}$ clusters, *RAI2* was significantly up-regulated ($p = 0.0013$, $p = 0.0002$). Samples in

the amphicrine cluster also showed high CTC counts with a median of 743 CTCs per $7.5\,\mathrm{mL}$ blood. Patients in this cluster had either been classified as NEPC or AVPC based on their clinical parameters, indicating the difficulties in identifying the amphicrine patients in the clinic, especially with lack of complete histopathological analysis.

For patients with blood samples from multiple time points, the clustering of the individual samples was often stable for samples that were collected over a short period of time. For example, NEPC patient 75 had two samples drawn within six weeks and both were clustering to the NEPC cluster. Similarly, two blood samples were collected from patient 10 within six weeks and both were assigned to the $AR^{high}$ cluster. In contrast, samples collected over a longer period of time displayed how patients could either switch between subtypes or stay stable during disease progression. For instance, patient 3 had two samples collected within an interval of four months. The first was taken before switching to carboplatin/etoposide based chemotherapy and the second at the time of progression. At both time points the patient fell into the $AR^{high}$ cluster. Opposed to that, patient 65 was first assigned to the amphicrine cluster, but after five month moved to the NEPC cluster due to loss of PRAD markers and increase of NE markers.

In conclusion, unsupervised analysis revealed four clusters with distinct gene expression profiles. These included a $CTC^{low}$ cluster, and three CTC positive clusters with $AR^{high}$, $NE^{high}$ or amphicrine subtype. Clustering of patients was stable for short time periods, but could display transdifferentiation processes over longer intervals.

#### 3.2.2.4 Supervised analysis of gene expression profiles to predict patient group

The aim of this project was to develop a liquid biopsy assay that allows to identify patients undergoing NET based on the CTC gene expression signature. For this purpose, methods for classification and prediction of the clinical group based on the gene expression profiles were explored. As the unsupervised analysis revealed an increased expression of NE markers in the NEPC samples but a mixed expression of PRAD markers, the first approach was to classify the samples based on the positivity for any NE marker. Every marker per sample was either deemed positive or negative based on an expression threshold determined with healthy donor samples. Based on this binary classification, 87.9 % of all AdnaTest positive NEPC samples were positive for at least one NE marker (see table 3.15 A). This was significantly higher than in the AVPC group with 60.4 % NE marker

positivity ($p = 0.0112$). In the DNPC group, only one out of three samples was positive for NE markers. However, the detection of NE markers was still too high in the non-NEPC groups to classify the samples based on this measure. In contrast, no significant differences were found between the samples with regard to the percentage of samples with detection of at least one PRAD marker. Similar to the AVPC and the HSPC group, PRAD markers were identified in 90.9 % of the NEPC samples. Only when *HOXB13*, which was identified to be the only PRAD marker in pure NEPC before, was excluded, the percentage of positive samples was significantly reduced from 91.7 % in the AVPC group to 66.7 % in the NEPC group ($p = 0.0044$).

As already in the non-NEPC groups, more than half of the samples were positive for at least one NE marker, this parameter was not sufficient for group prediction. For example, five blood samples were collected from patient 7 within four months. At all time points, the samples were grouped into the $AR^{high}$ cluster. However, the patient was positive for one or two NE markers at every time point. In line with the clustering, histopathological analysis taken in parallel to the last blood sample showed strong PSA and nuclear AR staining, while the neuroendocrine marker SYP was negative and the cell morphology was not suspicious of small cell carcinoma.

Thus, the number of positive NE markers per samples was counted and depicted in figure 3.15 A. The non-NEPC samples, including AVPC, HSPC and DNPC samples, expressed a significantly lower number of NE markers with a median of 1 positive marker in contrast to 3 markers in the NEPC group ($p < 0.0001$). The count of positive NE markers was used as a classifier in a ROC analysis to differentiate between NEPC and AVPC or HSPC, respectively. NEPC and HSPC could be separated with an AUC of 84.3 % (figure 3.15 B and C). Similarly, the positive marker count was sufficient to differentiate between NEPC and AVPC with an AUC of 82.8 %. At a threshold of two or more positive NE markers the classifier reached a sensitivity of 78.79 % and a specificity of 81.82 % for the differentiation of NEPC and HSPC samples. Similarly, a threshold of two ore more positive NE markers showed the most favorable combination of 78.79 % sensitivity and 79.17 % specificity for the comparison of NEPC and AVPC samples.

To test whether the sample group could be predicted with an even higher accuracy, a more sophisticated machine-learning based, supervised model was trained to include all variables that were acquired. A random forest classifier was chosen, as it can easily be trained without extensive optimization of hyperparameters and

A

| | no. CTC$^+$ samples | NE$^+$ | AC$^+$ |
|---|---|---|---|
| NEPC (n=35) | 33 | 87.9 % | 90.9 % |
| AVPC (n=50) | 48 | 60.4 % | 97.9 % |
| DNPC (n=3) | 3 | 33.3 % | 100 % |
| HSPC (n=11) | 11 | 54.5 % | 81.8 % |

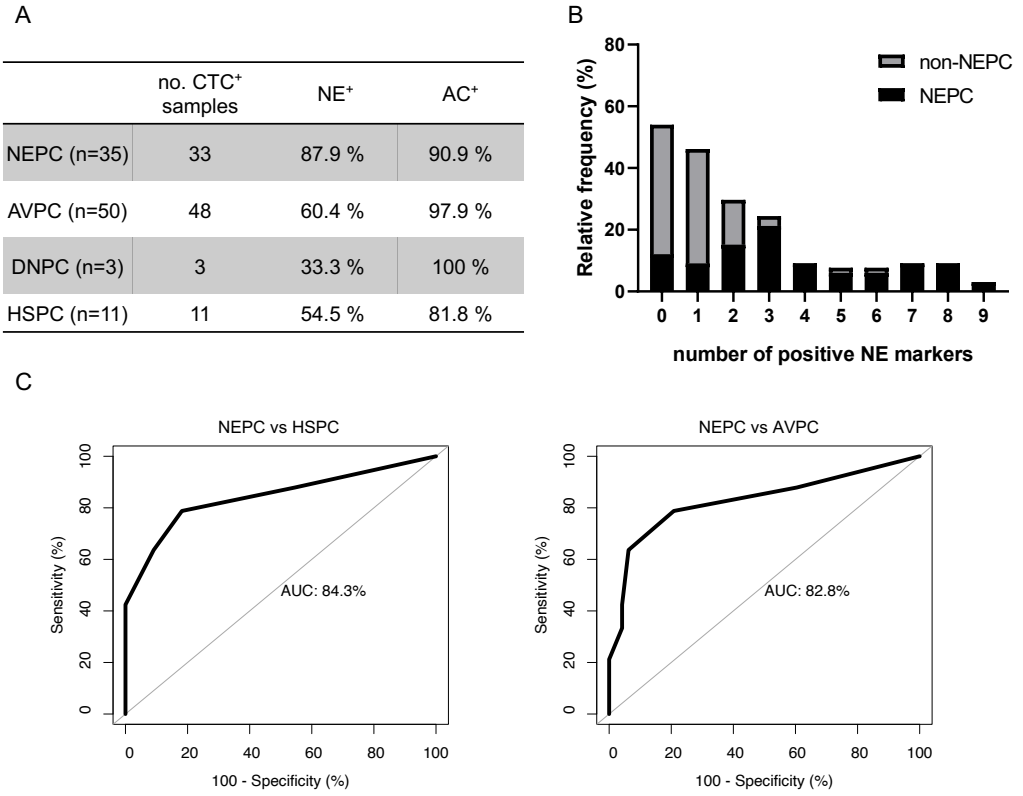



Figure 3.15: **Prediction of patient group based on binary analysis.** Normalized gene expression data were simplified to binary representation of marker positivity based on thresholds set by means of healthy donor samples, only CTC positive samples were considered for the analysis; A: percentages of samples positive for at least one neuroendocrine or adenocarcinoma (AC, HOXB13 excluded) marker, respectively; B: frequency distribution of the number of detected NE markers in NEPC and non-NEPC (AVPC, DNPC, HSPC) samples; C: ROC analysis comparing NEPC vs AVPC and HSPC, respectively, based on the number of positive NE markers per sample.

does not require careful scaling of the input data. In addition, the random forest algorithm includes bootstrapping of the samples which reduces the risk of over-fitting the model that arises from a small group size and a missing validation data set. In order to avoid over-fitting of the model, only one sample was included per patient in case multiple samples had been collected. For patients with multiple samples, only the first sample which was positive for CTCs was included in the analysis. Only samples that were CTC positive based on the AdnaTest results were included into the analysis to avoid training the classifier to identify CTC positive samples in general.

First, the model was trained to differentiate between NEPC and HSPC group

based on a subset of 22 NEPC and 11 HSPC samples. The results are shown in figure 3.16. The accuracy of sample classification could be improved in comparison to the simple count classifier, reaching an AUC of 95.5 % and a sensitivity of 90.9 % at 95.5 % specificity. In leave-one-out cross-validation, the out-of-bag error rate was 15.15 %. The predictions of the model are shown in 3.16 B. Samples that were mis-classified had borderline probabilities for both groups and were not clearly interpretable for the model. These samples were examined closer for the course of their disease. For instance, sample PC256 from patient 21 was predicted by the classifier to belong to the HSPC group, while the patient was categorized as NEPC based on clinical parameters. Indeed, histology analysis of metastatic tissue showed a double-positive tumor with AR expression and partial small cell morphology with dot-like lesions in CK staining. Overall, the patient also had low CTC counts, indicating that also the cell number might not have been sufficient to detect more NE markers. In contrast, the other four mis-classified patients originally belonged to the HSPC group. Two of the samples were positive for *ASCL1* and thus might have been attributed to the NEPC group. As NE markers were not observed in histology, this indicated that the *ASCL1* assay might be more error prone. Two samples as well showed low or partial PSA staining in histology and might therefore have been in the process of reduction of AR signaling and thus, the corresponding patients should be followed-up more closely. Another mis-classified HSPC sample was only positive for one epithelial marker and it therefore remains questionable whether this sample was positive for CTCs at all.

In addition, the importance of the single markers for the performance of the classifier was analyzed. As shown in figure 3.16 C, the epithelial markers *KRT19* and *EPCAM* were the most important features for group assignment. This indicates the maintained epithelial differentiation in the NEPC group and might as well be derived from the increased CTC counts observed in this group. The NE markers *CHGA*, *ACTL6B* and *CEACAM5* also highly contributed to the classification. These markers also showed significant changes in expression between the groups on a single marker level. The three NE markers were directly followed by the PRAD markers *HOXB13*, *AR-V7* and *KLK3*. All three clusters of genes thus appear to shape the expression profile indicative of the respective group. In contrast the markers *AR* and *SOX11* are of least importance for the model.

As the discrimination of aggressive but AR-expressing and aggressive neuroendocrine disease is relevant in the clinic, another model was trained to differen-

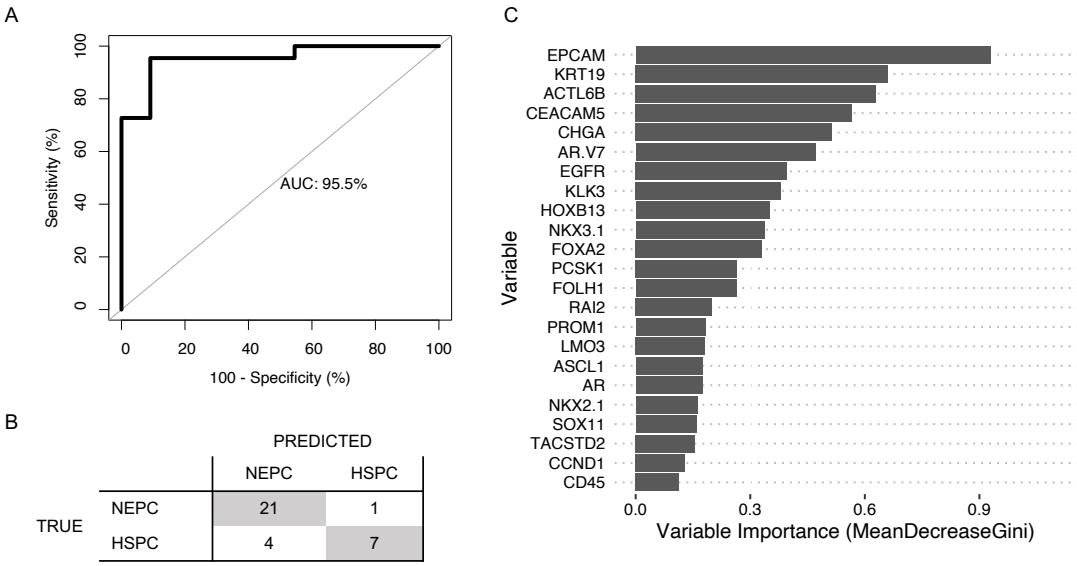



Figure 3.16: **Random forest classification of NEPC and HSPC patients.** Normalized gene expression of NEPC and HSPC samples from individual patients were used to train a random forest classifier with leave-on-out cross-validation; A: ROC analysis of the model trained, B: model prediction based on cross-validation, C: variable importance of the single markers in the classifier.

tiate between AVPC and NEPC patients. 22 NEPC and 32 AVPC patients were included in this analysis. The model was able to discriminate between the patient groups with an AUC of 88.2 % and an out-of-bag error rate of 23.53 % with a trend to incorrectly identify NEPC samples as AVPC and not vice versa (see figure 3.17 A, B). A closer examination of the mis-classified patients revealed that these cases were mostly assigned to the amphicrine cluster with double-positive marker expression. Other mis-classified patients from both the AVPC and the NEPC group belonged to the $CTC^{low}$ cluster. Thus, marker detection might have been hampered by the low CTC numbers in those samples. The most important variables in the model were *CHGA* and *KLK3* (see figure 3.17 C). This emphasizes the importance of both NE and PRAD markers for prediction and simultaneously explains the difficulties in the discrimination of double-positive samples. While the epithelial markers *EPCAM* and *KRT19* were intermediately important for the model, the NE markers *SOX11* and *LMO3* as well as the stemness markers *TACSTD2* and *PROM1* were only of limited value for the model.

In conclusion, the count of positive NE markers was sufficient to discriminate between NEPC and non-NEPC samples with an AUC of more than 80 %. A ran-

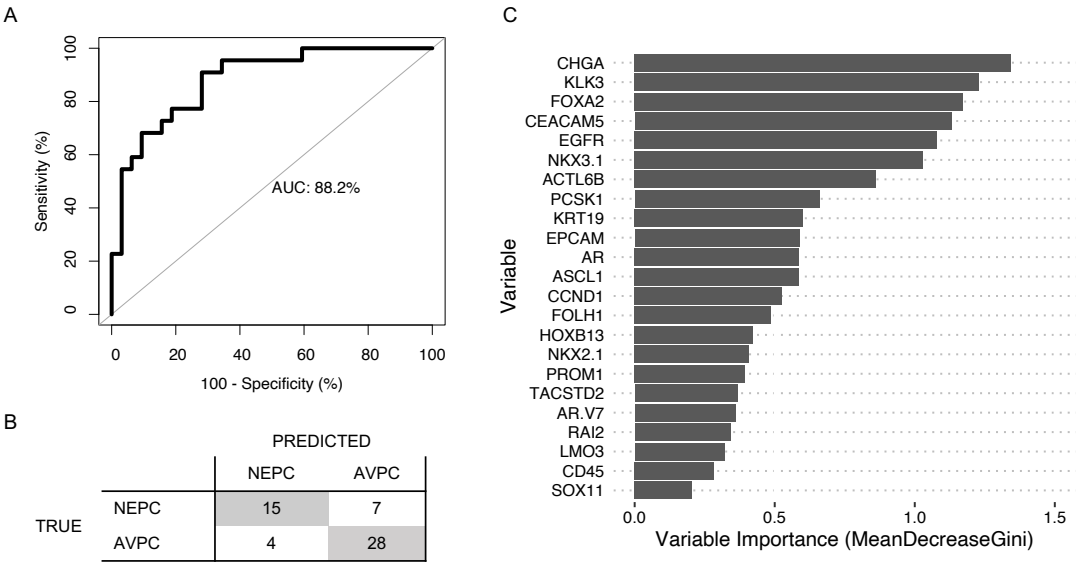



Figure 3.17: **Random forest classification of NEPC and AVPC patients.** Normalized gene expression of NEPC and AVPC samples from individual patients were used to train a random forest classifier with leave-on-out crossvalidation; A: ROC analysis of the model trained, B: model prediction based on cross-validation, C: variable importance of the single markers in the classifier.

dom forest classifier trained on the marker panel achieved an even higher accuracy and classified NEPC and HSPC samples with an AUC of 95.5 %. Epithelial and NE markers were most important for the correct classification of the samples. Classification of NEPC and AVPC samples was possible with an AUC of 88.2 %. NE and PRAD markers were necessary for the correct prediction of the group. Double-positive samples impeded a similarly accurate discrimination of NEPC and AVPC samples compared to HSPC samples.

### 3.2.3 Integration of CTC analysis and clinical data

#### 3.2.3.1 Correlation of liquid biopsy data and clinical parameters

Following the successful development of a liquid biopsy assay, the next step was to test whether patient clustering and subtype prediction based on the assay results correlated with clinical parameters and patient outcomes.

Visceral metastasis is rather uncommon in PCa, but frequently observed in aggressive disease including NEPC [34]. In order to find associations between the metastatic status and the liquid biopsy results, the incidence of liver metastasis was compared to the NE marker status and the expression profiles in AdnaTest (fig. 3.18). Information were available for 80 samples from 59 patients.

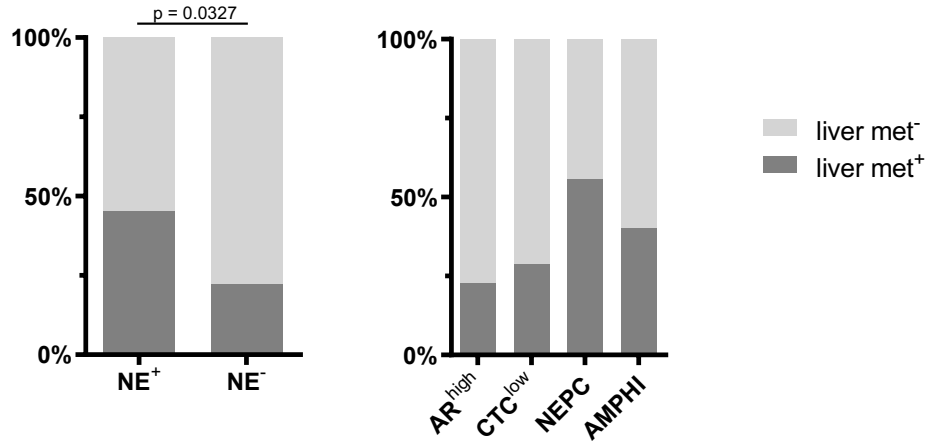



Figure 3.18: **Association of liver metastasis and liquid biopsy results.** Frequency of liver metastasis was compared between samples depending on the detection of NE markers and the overall expression profiles represented by the four clusters identified in unsupervised analysis; threshold for NE-positive samples > 1 NE marker positive; AMPHI: amphicrine.

First, liver status was evaluated depending on the detection of NE markers in AdnaTest. A threshold of two or more positive markers was chosen based on the results of the sample classification in 3.2.2.4. 45.2 % of $NE^+$ samples had liver metastasis and this was significantly more frequent than in the $NE^-$ group with 22.4 % ($p = 0.0327$). Next, the frequency of liver metastasis was compared for the four clusters identified in unsupervised analysis of CTC transcript profiles. Liver metastasis was most frequently detected in the NEPC cluster in which 55.6 % of samples were from patients with liver metastasis. In the amphicrine cluster, this frequency was slightly reduced to 40.0 %. The lowest frequency of liver metas-

tasis was found in the $AR^{high}$ cluster with 22.6 %. Similarly, 28.6 % of samples in $CTC^{low}$ cluster had liver metastasis. However, no significant differences in the frequency of liver metastasis between the clusters were found.

Serum markers such as PSA, NSE and LDH are used in the clinic to monitor response to treatment. Thus, these markers were correlated to the respective liquid biopsy assay read outs to analyze the concordance of these markers and the reliability of the liquid biopsy analyses (fig. 3.19).

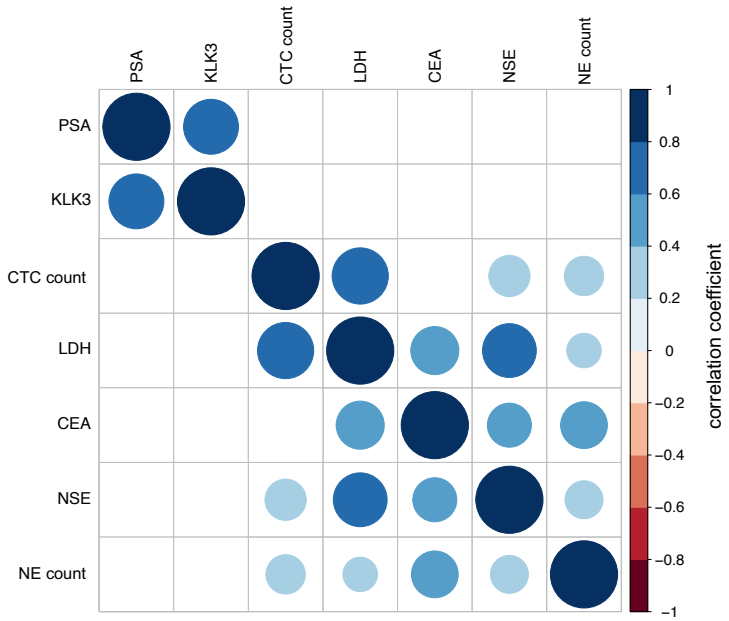



Figure 3.19: **Correlation of serum markers and liquid biopsy read outs.** The serum markers PSA, NSE, CEA and LDH were correlated to the associated liquid biopsy assay read outs *KLK3* expression, NE marker count and CTC count, respectively; the color of the dots indicates the Spearman correlation coefficient according to the color legend and only correlations with a p-values below 0.05 are shown.

*KLK3* expression in CTCs and serum PSA were positively correlated with a correlation coefficient of 0.67, as expected from the biological relation between the *KLK3* transcript and its gene product PSA ($p = 0.0008$). This emphasizes the accuracy of the developed pipeline for gene expression in CTCs. Serum LDH is an unspecific tumor marker that indicates cellular turn-over and, thus, is used as a surrogate for tumor burden [164]. In line with that, a positive correlation to the CTC count determined by CellSearch was observed ($r = 0.7$, $p = 0.0031$), indicating that also the CTC counts represented the tumor burden in the patients. This supports the hypothesis that EPCAM-based CTC enrichment captures the rele-

vant cell population in AVPC and NEPC patients. Serum LDH also correlated with serum NSE ($r = 0.64$, $p < 0.0001$) and this might reflect the high proliferation index usually found in NEPC [36]. Since NSE was not part of the transcript panel itself, serum NSE was compared to the count of positive NE markers to represent NET of the CTCs. Here, a moderate positive correlation with a coefficient of 0.32 was found ($p < 0.0001$). Similarly, serum CEA correlated with the NE marker count in CTCs ($r = 0.49$, $p < 0.0001$). In several samples with multiple NE markers in CTCs, only a slight increase of serum NSE was observed similar to the samples without detection of NE markers in CTCs. Thus, detection of NE markers in CTCs might be a more sensitive and specific marker for NET. The increased sensitivity of the CTC-based analysis was also illustrated by single cases. For instance, PC334 was identified as NEPC based on the detection of four NE transcripts while serum markers NSE and CEA were not elevated above normal in a parallel blood sample.

Survival of patients after first diagnosis was assessed for the groups as well as for the clusters and the CTC counts to explore prognostic values of the liquid biopsy analysis. The survival probability of the patients as well as the number of patients at risk is displayed in figure 3.20.

With regard to the clinical group no significant differences in OS were found (fig. 3.20 A). However, many patients were censored at early time points, thus leaving only few patients for long time follow-up. Expression of NE markers showed only a minimal effect on OS with a similar prognosis for patients who expressed more than one marker and those who were marker negative (fig. 3.20 B). Based on a threshold of 5 CTCs per $7.5\,\mathrm{mL}$ blood, no significant differences in OS were found between $CTC^{high}$ and $CTC^{low}$ samples [101] (fig. 3.20 C). Lastly, survival was compared for the four clusters identified in unsupervised analysis of the expression profiles (fig. 3.20 D). No significant differences in OS were found between the clusters.

In conclusion, the choice of the first diagnosis as starting point and the limited follow-up for some patients resulted in no significant differences in OS with regard to clinical group, CTC count and expression profile.

#### 3.2.3.2 Longitudinal analysis of individual patients

While longitudinal analysis was not the central aim of this study, multiple blood samples were collected and evaluated for twenty patients. Per patient, 2-5 sam-

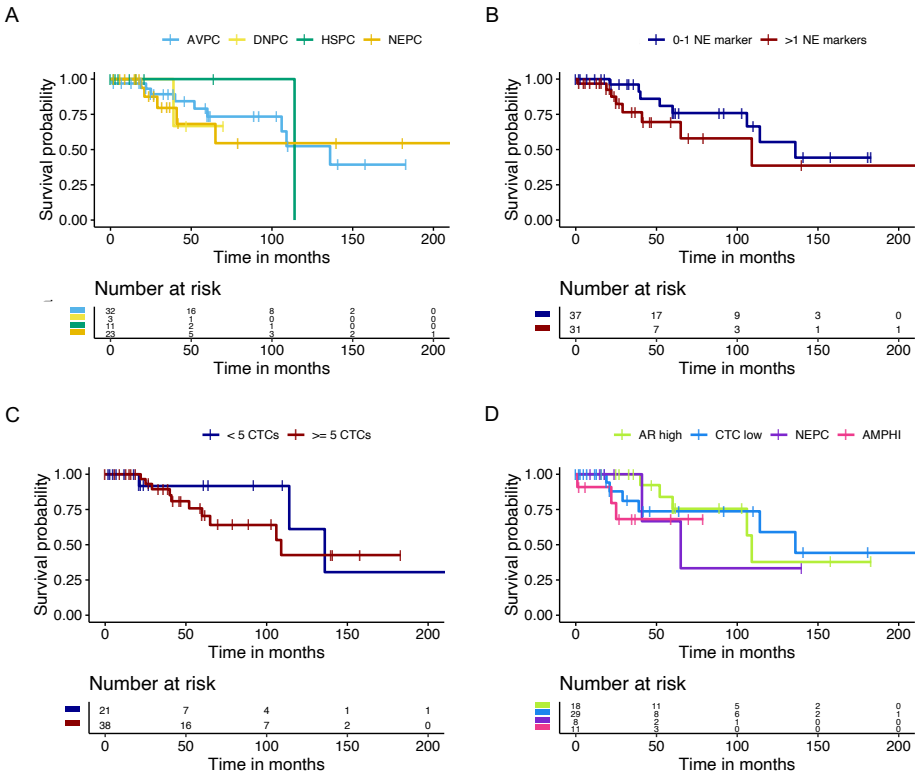



Figure 3.20: **Survival analysis based on liquid biopsy read-outs.** Overall survival of patients after first diagnosis was analyzed for all patients; A: Comparison of survival between clinical groups; B: comparison of survival for patients depending on the detection of neuroendocrine (NE) markers (threshold: > 1 marker); C: Comparison of patient depending on the CTC count based on CellSearch (threshold: >= 5 CTCs); D: Comparison of patient survival between the four cluster identified in the gene expression analysis, AMPHI: amphicrine.

ples were available. As the follow up samples were not drawn at pre-defined time points as required for systematic analysis, three case reports are presented in this section to highlight the value of longitudinal CTC analysis to recapitulate the clinical cause of the disease. First, patient 22 was initially diagnosed with Gleason 8 metastatic PCa and metastatic lesions in lymph nodes and liver. His initial PSA was $85.2\,\mathrm{ng/mL}$. Histopathology report from that time found an adenocarcinoma partially admixed with a neuroendocrine carcinoma. While PSA was positive in the adenocarcinoma proportion of the tumor, transdifferentiated areas were PSA negative and highly proliferative. In total, four samples were available for analysis and the combined results are displayed in figure 3.21.

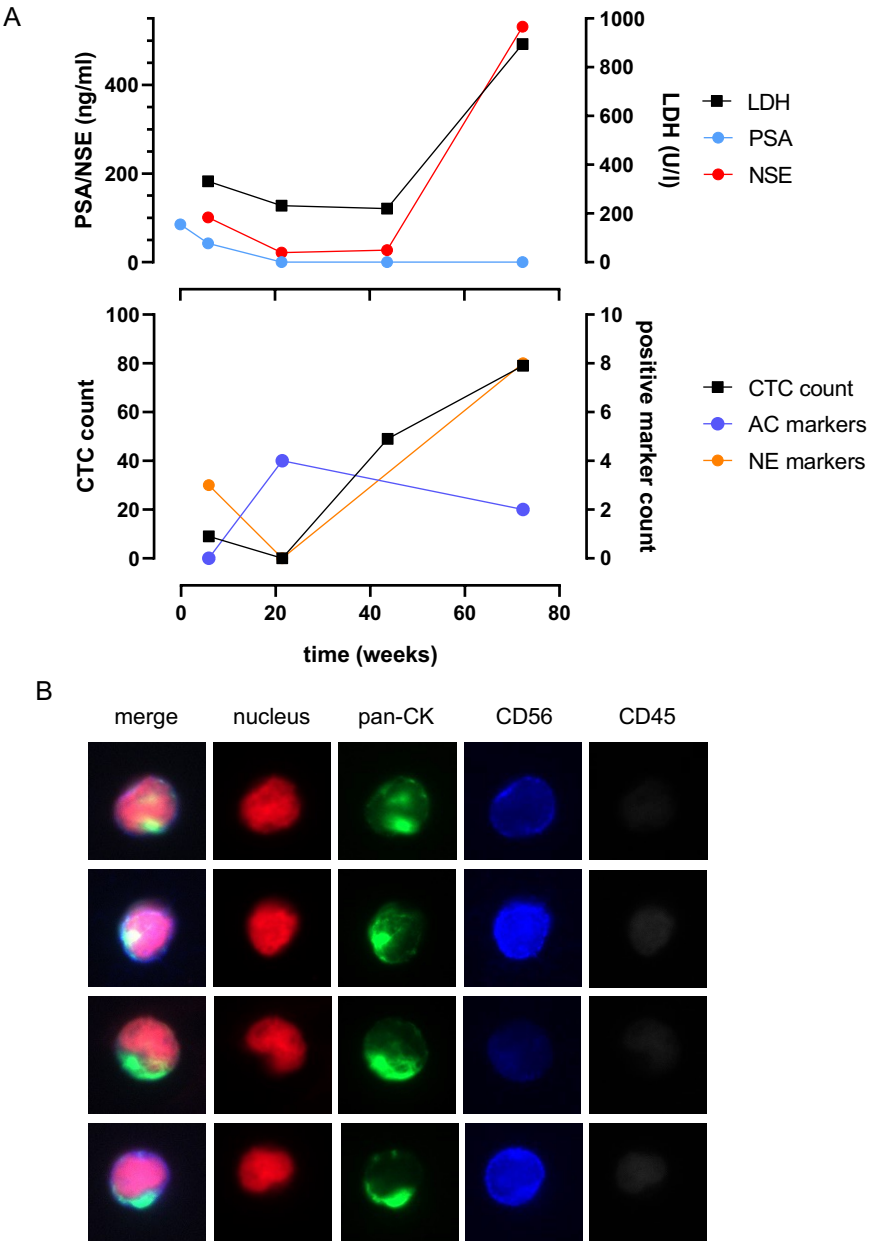



Figure 3.21: **Longitudinal analysis of patient 22.** A: Serum markers at the time point of blood collection are shown in the upper plot; CTC count determined by CellSearch and the number of positive markers in AdnaTest are depicted in the lower plot; B: representative images of CTCs enriched by Parsortix at the last time point of blood collection; immunofluorescence staining was used to detect pan-CK (AF488) and CD56 (BV421), CD45 (PerCP) served as negative marker and nuclei were counterstained with DRAQ5.

At the first time point of blood collection, the patient had a count of 9 CTCs/mL and the AdnaTest detected NE markers, which is in line with the histological analysis. PRAD markers were not detected, which might indicate that the highly

proliferative NE cells contributed more to the CTC pool in the blood. Serum PSA and NSE were both elevated. The patient was treated for his NE disease with carboplatin/etoposide chemotherapy and serum markers as well as CTC count and NE markers in AdnaTest were reduced at the second time point of blood collection, while the AdnaTest still detected PRAD markers. Five months later an increase in CTCs was detected that was not accompanied by an increase of any of the serum markers. Only at the last time point of blood collection after six additional months, serum LDH and NSE dramatically rose accompanied by an increase of the number of NE markers found in AdnaTest. For this time point, parallel Parsortix analysis was also available. CTCs had an NE phenotype with a partial small cell-like morphology as indicated by dot-like CK staining and intermediate to strong CD56 expression. In line with serum PSA, detection of adenocarcinoma markers with the AdnaTest was reduced, indicating an AR-independent progression. This case highlights the superiority of CTC count in detection of disease progression compared to the serum markers LDH and NSE. While both markers only showed an increase at the fourth time point, the CTC count indicated disease progression already at an earlier time point. AdnaTest results illustrated the response of the NEPC to the carboplatin/etoposide therapy during the second blood draw, while PRAD markers were still detectable. The NE markers *ASCL1, LMO3* and *CEACAM5* were the first transcript to be up-regulated in CTCs, while at the last time point all markers but *ASCL1* were detected. This supports the role of ASCL1 as a pioneering TF in the NET process.

Similarly, the temporal landscape of NE marker expression was illustrated in the disease course of patient 35. He was first diagnosed with hormone-sensitive PCa and treated with ADT. After multiple lines of treatment the patient progressed with NEPC. The patient experienced a fast progression of the disease with a continuous increase in CTC count and expression of NE markers as illustrated in figure 3.22.

The drastic tumor growth was represented by rising serum LDH and CTC count in CellSearch. At the same time, serum PSA increased less steeper and the number of detected PRAD markers in AdnaTest remained unchanged. In contrast, serum NSE and the number of positive NE markers in AdnaTest strongly increased. The histopathology report found a small-cell neuroendocrine carcinoma in the liver. The tumor cells had a dot-like pan-CK staining, were negative for CHGA and SYP, but focally positive for CD56 and highly proliferative. Comparison of the AdnaTest results for the three collection time points informed about

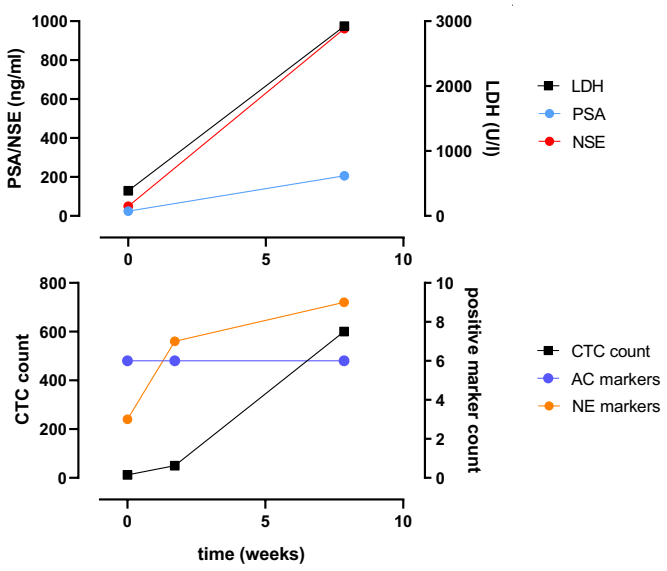



Figure 3.22: **Longitudinal analysis of patient 35.** Serum markers at the time point of blood collection are shown in the upper plot; CTC count determined by CellSearch and the number of positive markers in AdnaTest are depicted in the lower plot.

the time course of NE marker expression. *ASCL1*, *CEACAM5* and *CHGA* were the first markers to be detected two of which also were among the first markers detected in patient 22. While CHGA was not detected in histological examination, the AdnaTest gave a positive result for *CHGA*, suggesting that liquid biopsy better represented intra-patient tumor heterogeneity. At the second time point of blood collection about two weeks later, the markers *NKX2-1*, *LMO3* and *ACTL6B* were also found positive. Lastly, the markers *PCSK1* and *FOXA2* turned positive in the blood sample collected six weeks later. Interestingly, these two NE markers were significantly up-regulated in the pure NEPC cluster and, thus, might be indicative or even required to the progression to purely neuroendocrine disease. However, patient 35 was assigned to the amphicrine cluster at all time points due to the continuous expression of PRAD markers. This suggests that, other metastatic lesions might still be dominated by adenocarcinoma clones.

At diagnosis, patient 65 presented with neurological symptoms of primary osseous metastatic prostate carcinoma. Decompression surgery revealed an adenocarcinoma originating from the prostate. A high volume disease according to CHAARTED criteria was diagnosed in CT scan and bone scan. There was no sign of visceral metastasis. Additional immunohistochemical examinations of

the resected tumor tissue showed positive expression of AE1/AE3, PSAP, PSMA and partially weakly positive for PSA, as well as nuclear positivity for AR. At that time, the cells were negative for synaptophysin and there was no evidence for a neuroendocrine small cell component. Intensified hormonal therapy with abiraterone was initiated. Subsequently, a serological and morphological tumor response was observed. As shown in figure 3.23, neuroendocrine markers were detected in CTCs at the first time point of blood collection, when no macroscopic signs of NEPC were visible. As observed in the other two case reports, *CHGA* and *CEACAM5* were the first markers to be detected.

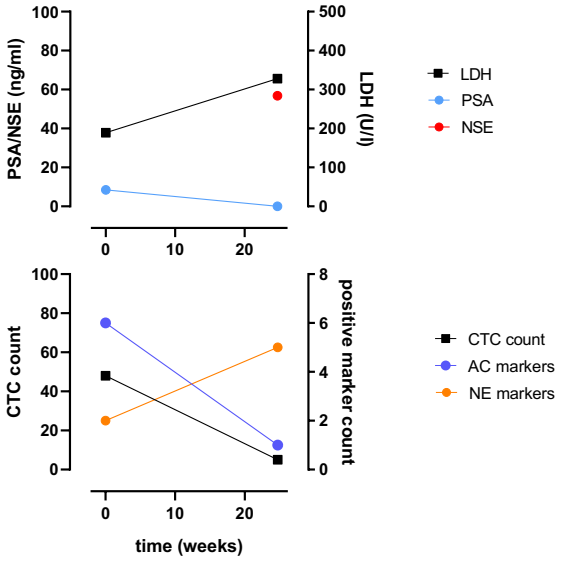



Figure 3.23: **Longitudinal analysis of patient 65.** Serum markers at the time point of blood collection are shown in the upper plot; CTC count determined by CellSearch and the number of positive markers in AdnaTest are depicted in the lower plot.

After six months, however, a local, PSA-negative progression of the primary tumor occurred. A biopsy showed extensive infiltrates of a neuroendocrine prostate carcinoma with positivity for synaptophysin and negativity for PSA, AR, PSMA and NKX3.1. The proliferation index Ki67 was >95 % of the tumor cells. Serum NSE was elevated. In line with that, the CTC count, the number of positive PRAD markers on CTCs and serum PSA were decreased. In contrast, the number of NE markers increased as *ASCL1, ACTL6B* and *PCSK1* turned positive. This resulted in a switch of the patient from the amphicrine to the pure NEPC cluster. This case report highlights the value of the developed CTC analysis pipeline for early de-

tection of NEPC. It also illustrates the switch of the patient from the amphicrine to the pure NEPC cluster in response to intense therapy and emphasizes the need for sensitive monitoring of treatment response.

# 4 Discussion

Treatment-emergent neuroendocrine transdifferentiation and the lack of reliable biomarkers are major problems in late stage PCa. The aim of this study was to identify new biomarkers for NEPC based on liquid biopsies which depict tumor heterogeneity and allow repeated monitoring of the patients. For this purpose, a panel of 22 markers including specific epithelial, PRAD and NE transcripts was selected based on literature review and validated in PCa cell lines and patient tissue samples. A multiplexed, qPCR-based transcript analysis pipeline was developed as a reliable and cost-effective method to characterize enriched CTCs. The analysis of patient samples revealed heterogeneous CTC morphology and significantly increased CTC counts in AVPC and NEPC compared to HSPC. The gene expression analysis detected several markers with differential expression between the groups. Based on the transcript profiles, four major molecular subtypes were identified. The expression profiles allowed to robustly predict NEPC samples and longitudinal case studies highlighted the potential of the liquid biopsy assay for early detection of NEPC.

## 4.1 Marker panel selection and analysis pipeline validation

Liquid biopsy allows the analysis of patient samples at the level of variable biomolecules and, hence, at different points in the flow of information from genotype to phenotype [92]. Depending on the research question, a suitable type of biomarker needs to be chosen. In case of treatment resistance and transdifferentiation in PCa, the targeted analysis of gene expression profiles by PCR-based analysis of mRNA was chosen. Genomic aberrations at this disease stage have been shown to be largely consistent with castration-resistant but still AR-dependent adenocarcinoma [52]. Thus, the detection of mutations or structural aberrations

at either the level of CTCs or cfDNA would not enable the specific identification of patients undergoing NET. The combined loss of the three tumor suppressor genes *TP53, PTEN* and *RB1* has been shown to be characteristic of AVPC, but the increased combined defect from 4 % in CRPC to 31 % in AVPC as observed by Aparicio *et al.* is not sufficient for sensitive identification of patients [71]. In addition, the genomic loss of a gene is even more difficult to detect in liquid biopsy samples in which the background of healthy leukocyte DNA covers the deletion in the tumor DNA. In addition, the direct analysis of proteome profiles in very small sample quantities, as usually seen in liquid biopsy, is still technically challenging and expensive. Immunocytochemical staining allows the analysis of selected protein biomarkers in enriched CTCs, however the method is limited by the small number of proteins that can be measured simultaneously [92]. In contrast, the analysis of RNA species allows to identify differentiation stages and can rather easily and sensitively be accomplished with PCR or sequencing based methods [92]. In addition to mRNA, other regulatory RNA species such as microRNA can be analyzed. However, technical validation of microRNA measurement from liquid biopsy samples has shown large variability depending on several sampling parameters [165]. Consequently, mRNA was chosen as liquid biopsy analyte to capture cell differentiation. Due to the small sample volume, PCR-based detection in combination with a multiplexed pre-amplification reaction was chosen. This allows the semi-quantitative analysis of multiple markers from a small sample volume with high sensitivity. Compared to sequencing-based analysis, it is less expensive and has a shorter turnaround time. Thus, multiplexed PCR-based detection is a suitable method for repeated application in the clinical setting.

Phenotypic plasticity has recently been added to the hallmarks of cancer as it is an overarching principle of treatment resistance and not only observed in PCa [166]. When choosing suitable biomarkers for the detection of NET, the aim was to include not only markers of terminal NE differentiation, but also markers of lineage plasticity in general. However, many regulators of plasticity and epigenetic regulation, such as EZH2 and other histone modifiers, are universally expressed. In CTC analysis, this bears the problem of unspecific signals from the leukocyte background that is accompanying the enriched CTCs. For this reason, numerous well-established drivers of lineage plasticity and transdifferentiation such as *REST, NCAM1* and *DLL3* had to be excluded from the marker panel. Consequently, SOX11, TACSTD2 and PROM1 were included as less prominent

markers of plasticity and stemness phenotypes that have also been identified in the context of NET in recent literature [68], [160], [161].

In addition, several TFs that drive neuronal or NE differentiation as well as terminal NE markers were selected. Overall, TFs are lower expressed than for example structural or membrane proteins [167]. This is especially challenging for the sensitive detection of these markers in small cell numbers as commonly observed in liquid biopsy. Indeed, the most frequently detected neuroendocrine markers were CHGA, CEACAM5 and ACTL6B – all of which are not TFs but effector proteins of NE differentiation [74], [79], [88]. These markers were identified with high confidence in the spike-in experiments compared to the pure cell lines. In contrast, some of the TFs such as LMO3 or SOX11 were confidently identified in spike-in controls of the highest expressing NEPC cell line, but the signal was lost in spike-in samples of the neuroendocrine lung cancer cell line. This suggests that these TFs might show false-negative results in patients with overall low CTC numbers or with a minor NE subclone. Thus, this study has shown that it is important to include also effectors of NE differentiation and not only the drivers on the level of gene regulation.

The identification of DNPC patients was not the primary aim of this study, but a few patients with histological evidence of DNPC were included in our CTC analysis on an exploratory basis. A study by Bluemn *et al.* found a comparable frequency of DNPC and NEPC in a rapid autopsy study of CRPC patients. Due to the lack of suitable positive markers to identify this disease stage, DNPC is more difficult to recognize. The drivers of DNPC tumor growth are still poorly understood and only very recent publications point towards activation of Wnt and YAP/TAZ signaling as key signaling pathways [39]. In addition, Labrecque *et al.* found an up-regulation of a squamous cell like gene expression signature in DNPC patient tissue [41]. Based on this observation, *KRT6A* was included in the marker panel to explore these squamous cell traits in our patient cohort. Unfortunately, the *KRT6A* assays was unspecific in the spike-in controls and had to be excluded from further analysis. In future experiments, more markers should be tested including for instance *YAP1* and *TEAD1* which were identified by Tang *et al.* and are not expressed in leukocytes, unlike other drivers of the DNPC phenotype such as *FOSL1* and *TAZ*. In addition, published transcriptome data sets of CRPC patients can be used to search for DNPC samples and more potential marker transcripts. As explained before, the sensitivity of the assay panel might be improved by including not only TFs but also downstream effector proteins.

Such published data sets could be used to identify the most often and prominently up-regulated markers with the advantage of including more samples than a single study.

After careful validation of our marker panel in spike-in controls, five markers had to be excluded due to insufficient specificity and sensitivity. *SRRM4* was excluded due to insufficient sensitivity as it could not be detected in any of the positive control cell lines upon enrichment from blood. Although the *SRRM4* assay had a better PCR efficiency than other assays that performed well in spike-in controls, the overall expression of *SRRM4* as a splice factor might have been too low for robust detection in low cell numbers. *POU3F2, KRT6A, PEG10* and *LIN28B* were excluded due to unspecific signals in the spike-in controls compared to the cell line alone. Interestingly, the first three are exactly the three assays that have been annotated by the manufacturer to be active on genomic DNA. This suggests, that although mRNA enrichment by oligo(dT)-beads is part of the sample processing workflow, contamination with gDNA is still present in the purified RNA sample. In the pre-amplification step, even small amounts of gDNA can be amplified if mRNA-specific primers are not available, for example due to the gene structure. Thus, DNA and RNA signals are superimposed and quantification of RNA expression is not possible. The ValidPrime Kit was used with the intention to correct the detected expression for the gDNA signal [168]. It uses two control reactions to calculate the the pure RNA signal. First, the respective assay is tested on a gDNA control, second the samples are tested with a primer pair directed against an untranscribed DNA region. Approximately 20 % of the samples had a positive signal for the presence of gDNA in the cDNA sample. However, as seen in the spike-in controls, in this study the ValidPrime correction was not sufficient to disable unspecific results from assays with activity on gDNA. Thus, it is essential to carefully design new assays in order to avoid unspecific signals.

## 4.2 Influence of CTC enrichment methods on downstream analysis

CTCs are diluted in millions of leukocytes in the blood and need to be enriched for subsequent molecular analysis [92]. In this study, three different CTC enrichment methods were used. The CellSearch analysis is the only FDA-cleared technology for CTC enrichment and detection and was able to identify CTCs in the majority of patient samples analyzed in this project. The AdnaTest is as

well based on immunomagnetic enrichment and had concordant results with CellSearch. Only a small number of samples were positive by the AdnaTest but not by CellSearch. This was probably due to the additional epitopes for CTC enrichment compared to CellSearch. In contrast, size-based Parsortix enrichment performed worse with fewer CTCs counted after immunocytochemistry staining and fewer markers identified after gene expression analysis compared to AdnaTest. This illustrates the huge impact of the CTC enrichment method on the downstream molecular analysis of the CTCs. All enrichment methods come with their own advantages and disadvantages that need to be carefully considered when designing experiments with CTCs. First, label-dependent methods are well suited for specific enrichment of CTCs, but the choice of epitope might as well exclude relevant cell populations. EPCAM is the cell surface marker most frequently used for CTC enrichment, for instance in CellSearch and AdnaTest, as it is expressed in epithelial cells and can be used for the majority of solid tumors [169]. A key issue is the enrichment and detection of CTCs undergoing EMT, as these cells down-regulate EPCAM [169]. The use of the mesenchymal cell surface marker vimentin has been studied to identify CTCs that have undergone EMT. Transcript analysis of CTCs isolated after vimentin-based positive selection revealed an EPCAM-negative, mesenchymal phenotype. Additionally, vimentin-based CTC count has been found to have higher sensitivity and specificity for identifying patients with disease progression in a cohort of metastatic PCa patients [170]. However, as vimentin is also expressed by leukocytes, additional prostate cancer-specific markers are required for CTC detection and the higher background of leukocytes in the enriched fraction requires more sensitive detection methods for the tumor-specific markers. Nevertheless, EPCAM is widely used in PCa CTC detection and should not be neglected in enrichment of CRPC-derived CTCs. First, EPCAM has been associated with PCa stemness in 3D cell cultures [171]. In addition, EPCAM levels have been described to generally correlate with proliferation and dedifferentiation [169]. In CRPC patients, more EPCAM-high CTCs than EPCAM-low CTCs were found and only the number of EPCAM-high CTCs was associated with overall survival [172].

Transferring the techniques established in PRAD to NEPC, raises the questions whether NEPC CTCs are excluded from CTC enrichment by positive selection based on EPCAM. A study by Beltran *et al.* compared the CTC counts between the CellSearch system and the EPIC platform, which does not include an enrichment step, but proceeds with staining of all nucleated cells. They found that in some

patients with mixed or pure NEPC, the EPIC platform detected CTCs, that were not detected by CellSearch [135]. These cells had low to absent expression of AR and were negative for CK. In contrast, analysis of EPCAM expression in NE-differentiated tumor tissue from different primaries has revealed that the majority of these tumors have strong or at least variable EPCAM expression [173]. SCLC is an aggressive neuroendocrine tumor that shares many features with NEPC, also on the molecular level [174]. A study of nine CTC cell lines derived from SCLC showed high EPCAM expression in eight of the cell lines analyzed while being only partially positive for mesenchymal markers [175]. In a single-cell RNA sequencing-based study of CRPC samples, the neuroendocrine cell populations were the ones with the highest expression of EPCAM compared to luminal and basal cells [176]. Taken together, the evidence found in the literature indicates that EPCAM is an appropriate antigen for CTC enrichment of AVPC and NEPC derived CTCs.

Additional tumor-specific cell surface markers can be utilized to enrich CTCs with down-regulated or absent EPCAM expression. For example, the AdnaTest relies on three epitopes for enrichment, although anti-EPCAM antibodies are over-represented compared to the other two epitopes EGFR and HER2. Both proteins are growth factor receptors that are associated with CRPC. EGFR is up-regulated in tumor tissue during progression to CRPC to facilitate androgen-indpendent growth [177], [178]. Additionally, increased expression of EGFR has recently been associated with NET, where EGFR co-activates LIFR signaling and contributes to metabolic alterations that accompany NET [157]. Similarly, HER2 has been found to be contributing to CRPC progression and metastatic spread. Both receptors are as well associated with tumor initiating cells and CTCs [179]. The possibility of targeted therapy for aggressive disease is currently in preclinical testing, thus both markers are relevant for the purpose of this study [180].

To enrich NEPC-derived CTCs with a higher specificity, it is also possible to add NEPC-specific cell surface epitopes to the enrichment protocol. One candidate is DLL3, a membrane protein and inhibitory ligand of the Notch signaling pathway. DLL3 is currently evaluated *in vivo* as a specific target for diagnosis and treatment of NEPC. Radioactively labeled anti-DLL3 antibodies have been successfully evaluated for specific imaging of NEPC tumors in mouse xenografts [181]. Similarly, antibody-drug-conjugates targeting DLL3 resulted in durable responses in mouse xenografts and partial response in one NEPC patient in a basket trial [134]. The same study also analyzed DLL3 expression in CTCs and

found varying percentages of DLL3-positive cells in NEPC and castration-resistant PRAD [134]. Interestingly, a fraction of DLL3-positive CTCs were negative for CK, indicating that the use of DLL3 for CTC enrichment might capture cells which are overseen with the current EPCAM-based methods. Consequently, the integration of DLL3 or comparable NEPC-specific cell surface markers is a possibility to optimize CTC enrichment for NEPC patients in future studies. This can enhance sensitive detection of patients progressing with NEPC and might also point to therapeutic targets in a tumor subtype that still lacks effective therapy options. However, it is important to remember that adding more epitopes to an antibody cocktail for CTC enrichment simultaneously increases the leukocyte background due to more possibilities for unspecific binding. Thus the choice of antigens for label-dependent CTC enrichment has to be made carefully, including as many epitopes as necessary, but as few as possible.

As both AdnaTest and CellSearch mainly rely on EPCAM for CTC enrichment, it is not surprising that there was a broad overlap between the results of the two methods, with the AdnaTest detecting few more cases. In an analogical experiment, Gorges *et al.* compared CTC positivity rates between CellSearch and AdnaTest from paired colorectal cancer patient samples. The positivity rate for CTCs could be increase from approximately 30 % to 50 % when the results of both single assays were combined [103]. Similarly, in a direct comparison of AdnaTest and CellSearch results in a cohort of mCRPC patients, AdnaTest was more sensitive and detected CTCs in 92 % of CellSearch positive samples while only 68 % of AdnaTest positive samples had a positive CellSearch result [182].

However, direct comparison of CellSearch in AdnaTest remains difficult due to the different read outs obtained with both tests. According to the manufacturer and also in line with other studies, in this project samples were considered positive as soon as one of the prostate-specific or epithelial markers was detected [182]. Nevertheless, this method might also be prone for false-positive results. In this project ten age-matched healthy donor samples were analyzed with the AdnaTest to control for the background signal of the single assays. Indeed, PSA and EPCAM signals were detected in few healthy donor samples. This suggests, that samples with only one positive epithelial or prostate marker might have circulating epithelial cells but not necessarily circulating tumor cells. Expression of single prostate markers or epithelial markers in healthy donor samples has also been observed before, for example from total RNA, but also in PBMC from healthy donors [132], [183]. This emphasizes the need for careful background

correction using healthy donor samples to avoid false-positive results.

In contrast, label-independent methods use physical properties such as size and deformability to enrich CTCs. A bias due to epitope selection can therefore be avoided. However, no improvement in CTC detection nor in NE marker detection was observed in this study after Parsortix-based enrichment compared to the AdnaTest. This might be due to particular morphological features of the AVPC and NEPC derived CTCs. Chen *et al.* could show that in mCRPC CTCs nuclear and overall size was diminished compared to CTCs from non-metastatic patients. In addition, patients with visceral metastasis showed an even more pronounced reduction in nuclear size [184]. Very small CTCs detected especially in the most aggressive patients had a mean size of $6.8\,\mu\mathrm{m}$ which is almost identical to the size of the critical gap of the Parsortix casette with $6.5\,\mu\mathrm{m}$ [184]. As AVPC and NEPC patients progressed from mCRPC and visceral metastasis is one particular characteristic of AVPC, size-based CTC enrichment is not the most favorable CTC enrichment method at this late disease stage, as especially the smaller cells are derived from the more aggressive tumors.

To circumvent the issue of CTC enrichment, other studies have as well assessed luminal and NE markers in RNA isolated from whole blood of mCRPC patients. Detection of NE markers was associated with progression on NHA and overexpression of a higher number of markers was associated with poor survival [132]. However, the increased background of signals from leukocytes requires even more sensitive PCR assays and might therefore detect patients undergoing NET less early.

In both, AVPC and NEPC patients, extremely high CTC counts were measured. Indeed, high CTC counts have previously been associated with aggressiveness of disease. In mCRPC patients, elevated CTC counts were related to visceral metastasis, high alkaline phosphatase and high LDH [185]. This positive correlation between CTC count and serum LDH was reproduced in our patient cohort, emphasizing that EPCAM-based CTC enrichment captures the relevant cell populations that represent tumor burden. Several studies have investigated the CTC counts in CRPC applying different enrichment strategies. Using a label-independent approach, Ladurner *et al.* detected $\geq 5$ CTCs$/7.5\,\mathrm{mL}$ blood in 50% of mPCa patients compared to 69 % in our cohort which included more aggressive patients [186]. In line with our study, CellSearch analysis of samples from progressive CRPC patients had previously shown almost 80 % CTC positivity and a median CTC count of 45 CTCs$/7.5\,\mathrm{mL}$ blood [187]. Thus, the high CTC counts observed in this

project are consistent with recent literature. Similarly, Puca *et al.* found uniformly high CTC counts with a mean of 39 CTCs/mL and 37 CTCs/ml in CRPC-NE and CRPC-Adeno, respectively [134]. Clusters were found in both groups in contrast to our study where cluster were only seen in the AVPC group. However, as few clusters were found, no reliable conclusions can be drawn from this observation. In contrast, Conteduca *et al.* reported higher CTC counts in NEPC compared to CRPC-Adeno samples analyzed with the EPIC Science platform [136]. Small cell, neuroendocrine tumors typically have a very high proliferation rate as indicated by Ki67 staining and, thus, might indeed shed more cells into the blood stream. However, as neither the AVPC nor the HSPC group are congruent with the CRPC-Adeno cohort, it is difficult to extrapolate those findings to this study. Not only high CTC counts but also a small cell morphology have been observed in NEPC patients by Beltran *et al.* These cells had a higher nuclear to cytoplasmic ratio, an overall smaller cell area and higher cytoplasmic circularity [135]. Although CTC morphology was assessed in the CellSearch analysis, no significant differences between the groups were identified. Rather the detection of small cell CTCs was associated with increased CTC counts, suggesting that patients with a higher tumor burden and hence more aggressive disease had a higher heterogeneity in their CTC morphology. Small cell-like CTCs were identified based on manual inspection of the CellSearch galleries and the peri-nuclear dot-like CK staining. In order to systematically evaluate CTC morphology as a prognostic or predictive marker, it would be necessary to measure the size and shape of the nucleus and the CTC and calculate suitable ratios for reliable quantification of the morphological diversity.

In conclusion, EPCAM-based CTC enrichment is a suitable method to analyze NEPC derived CTCs. The high CTC counts observed in this cohort are consistent with results from previous studies on comparable patient populations.

## 4.3 Selection and classification of patient samples

Careful patient sample collection is a prerequisite for biomarker detection and precise allocation to groups is the basis of reliable inter-group comparison. However, this is particularly challenging in a heterogeneous disease setting such as mCRPC and AVPC. For the NEPC group, classification was rather straight-forward as patients were required to have histopathological report proving neuroendocrine disease. This included the detection of neuroendocrine markers such as CHGA,

SYP or CD56 in immunohistochemistry staining or a small cell morphology. Analysis of the pathology reports was limited by a non-standardized report format and missing stainings for NE markers in some tissue samples. In contrast, allocation to the AVPC group was based on a variety of clinical criteria including elevated serum markers and short progression on NHA among others. This resulted in a highly heterogeneous subgroup with very variable serum markers, CTC counts and survival. However, histopathological analysis was not performed for all AVPC patients and not every metastatic lesion was biopsied. Thus, it cannot be completely excluded that at least subclones of NEPC or DNPC were growing in the AVPC patients and contributed to the CTC pool in the blood. DNPC samples were defined based on the loss of expression of AR as well as neuroendocrine markers in histopathological analysis. However, analysis of DNPC samples was not a central aim of this study and the few samples that have been collected were the result of a selection bias for AVPC and NEPC features.

Most samples have been collected from patients at the time point of progression on their current therapy, but some patients also had stable disease at the time point of blood collection. However, the specific type of treatment and also the number of treatment lines differed between the patients. The majority of patients had previously received ADT and NHA or chemotherapy and typically had four or even more lines of treatment in total before blood collection. At the time point of blood collection most patients were progressing on taxane-based chemotherapy, but other treatments included also NHA, carboplatin-based chemotherapy or PARP inhibition and few also immune checkpoint blockade. This resulted in a heterogeneous patient cohort with tumors that might have developed different mechanisms of treatment resistance based on the respective sequence of therapies they had received. Additionally, this heterogeneity complicated the collection of survival data, because it was difficult to define a common day 0, since the patients had received different treatments in different orders. Thus, the initial diagnosis of PCa was chosen as starting point. Consequently, this might have led to limited significance of the survival analysis comparing the different groups. For instance, the association of OS and CTC count or AR-independent disease is well established in current literature, however this was not reproduced in this study [101], [52]. For future studies, the patient cohort and the definition of the day 0 should be defined in advance to increase the power of the analysis and reduce uncertainties in patient allocation. For instance, the date of progression to CRPC could be chosen for survival analysis, while the date of initial diagnosis could

be used only for patients with *de novo* NEPC. Still, in a project that runs two to three years, advances in the field may require a revision of the inclusion criteria and here it is of utmost importance that the clinicians involved in this process are blinded regarding the results of the liquid biopsy assays. For the longitudinal analysis, samples were not collected at predefined time points, but depending on when the patients visited the outpatient clinic and whether they underwent progression on their current therapy. Thus, longitudinal analyses allowed little conclusion about the lead time of liquid biopsy-based biomarkers, but rather showed stability or changes in the CTC profiles over time. Consequently, future longitudinal studies should start analyzing patient samples at the time point of progression to CRPC as those tumors are commonly still driven by AR signaling. After that, blood samples should be taken in predefined intervals and finally at the time point of clinical manifestation of AVPC or NEPC. This procedure would enable an estimation of the lead time of the CTC-based biomarkers compared to the clinical markers.

In conclusion, AVPC and NEPC are heterogeneous disease stages that require careful allocation of patient samples to the respective groups. As the transition between the groups is fluent, a correct clinical allocation can be challenging but might be improved by the developed molecular analysis. To identify significant differences between these heterogeneous groups, it is essential to precisely define inclusion criteria, blood sampling times and the intervals for survival analyses.

## 4.4 Gene expression profiles in enriched CTCs

Gene expression profiles were determined for 22 genes in 106 blood samples. The analysis of single markers showed several significant alterations in marker expression between groups. The AVPC group was characterized by increased expression of PRAD associated markers, whereas the NEPC group had reduced expression of PRAD markers with increased expression of NE markers. As only few DNPC cases were collected, this subgroup was not large enough to detect significant changes in the gene expression profiles.

Although a significant reduction in the expression of most PRAD markers was observed in the NEPC compared to the AVPC group, the NEPC samples showed a dichotomous distribution of PRAD marker expression. This is in line with the observations by Cancel *et al.* who found expression of AR and the androgen-regulated gene NKX3.1 in more than half of NEPC samples. These were more

frequently the samples with NEPC admixed to PRAD [159]. Subsequent hierarchical clustering indeed confirmed this observation and revealed two distinct clusters with NEPC patients. The role of CCND1 in NET has recently been discussed in the literature with divergent results about the regulation in NEPC [156], [188]. In this patient cohort, we found a clear down regulation of *CCND1* in the NEPC group and a positive correlation with several PRAD markers. This is consistent with evidence from cell models suggesting a role for CCND1 as a corepressor of neuronal genes in complex with histone deacetylases [189]. As pointed out before, markers of terminal NE differentiation such as *CHGA, CEACAM5* and *ACTL6B* were more frequently up-regulated than TFs such as *LMO3, SOX11* or *NKX2-1*. However, no single NE marker was detected in all NEPC samples. Consequently, the analysis of single markers is not sufficient to reliably discriminate between the disease stages

Correlation analysis of all individual markers revealed significant positive correlation of the PRAD specific markers. As for instance *KLK3* and *NKX3.1* are downstream targets of AR, this correlation was expected. Consequently, the number of PRAD markers can be reduced in future experiments, as there is only a limited gain of information from each individual marker. Thus, it may be sufficient to include *AR* as the master regulator, *KLK3* as one downstream target and *PSMA* due to its role in theranostics. For the NE markers, the significant positive correlation was less distinctive, most likely due to the overall lower detection of these markers. Only a subset of the NE markers showed significant anti-correlation with the PRAD markers. This indicates that different sets of NE markers might be associated with particular disease stages or steps in the NET process. While NE markers with no correlation to PRAD are those that are expressed early on in a double-positive state and stay activated in a pure NEPC state, the markers with anti-correlation to PRAD might be expressed at later stages in which tumor cells already switched off AR expression. A similar differentiation of NE markers in two distinct clusters has been made by Labrecque *et al.* [41]. NE markers were divided in REST-repressed genes and TFs regulating NE differentiation, with the letter being only up-regulated in AR-negative tumor tissue. While indeed *CHGA* and *PCSK1* showed corresponding results in our patient cohort and the Labrecque data, other markers such as *ASCL1* and *ACTL6B* were inconsistent. In contrast to this study, Labrecque *et al.* analyzed tissue samples derived from a rapid autopsy. Thus, differences in the co-expression of particular NE markers and PRAD markers might result from various metastatic lesions contributing to

the CTC pool. Alternatively, the NET process might include different intermediate steps in individual patients, hence resulting in heterogeneous patterns of NE marker expression.

Although the identified significant differences were in line with the expectations for each clinical group, overall quantification remains difficult, due to the pre-amplification step included in the transcript analysis workflow. In principle, quantification would be possible, as long as the pre-amplification reaction is performed with a reproducibly high efficiency. The pre-amplification includes a reduced number of cycles to ensure that the reaction is stopped in the exponential phase of amplification. Our PCR assays showed efficiencies between 85 % and 102 % in single reactions and amplification of the target was not inhibited in the multiplexed reaction. As all samples underwent the same analysis workflow, reduced performance of one assay compared to another would cause a systematic error that would affect all samples equally. However, normalization of gene expression adds another layer of complexity to the data analysis. The $C_q$ values of the individual markers were normalized using two housekeeping genes. Both housekeepers are also expressed in leukocytes, thus the number of leukocytes influences the normalization of the CTC private markers. As leukocytes dramatically outnumber the CTCs in all samples and the same concentration of beads is added per sample, a stable background of leukocytes in the enriched CTC fraction can be expected. The ranges observed in CellSearch counts were much higher than the range of the CTC background. However, different batches of antibody-conjugated beads or storage times of blood samples may introduce variation in the leukocyte background. For data analysis in this study, a threshold was calculated for each transcript with expression in healthy donor blood samples based on the mean $C_q$ value in the healthy donor controls. In patient samples, markers with lower or similar expression to healthy donors were considered negative. This did not only account for potential signals from leukocytes but also for rare circulating epithelial cells originating from tissue injury or other sources, thus further reducing potential false-positive results. Given that serum PSA correlates with PSA expression in the tumor and that CTCs recapitulate the properties of the tumor, positive correlation of serum PSA and *KLK3* expression in CTCs indicated, that the quantification of *KLK3* expression was reliable. Still, binary read outs were chosen for subsequent analysis to account for the semi-quantitative approach of transcript detection. This included the classification based on the number of positive markers and the random forest classifier based on multiple

decision trees with individual thresholds for marker expression.

In addition to the semi-quantitative multiplex PCR-based workflow, other methods are available to study mRNA expression in CTCs. Recently, the use of droplet digital PCR (ddPCR) has been validated for the detection of prostate biomarkers in enriched CTCs. The multiplex assay allowed for an absolute quantification of six transcripts and showed high specificity and increased sensitivity compared to classical qPCR [190]. Although the pre-amplification and, hence, the uncertainties in quantification can be neglected with ddPCR, a lower number of targets can be multiplexed. RNA sequencing offers the possibility to quantify expression of all transcripts and has been successfully applied for CTCs in the form of single cell RNA sequencing [191]. However, isolation of single cells and NGS require considerable resources and expertise and is therefore not the optimal method for clinical testing.

Four distinct clusters were identified by unsupervised analysis of patient samples based on the gene expression profiles. One cluster was characterized by a $CTC^{low}$ signature. The other three clusters were positive for CTCs and resembled $AR^{high}$, amphicrine and pure NEPC populations. The $AR^{high}$ cluster mainly comprised AVPC patients and was highly positive for PRAD markers, but negative for NE markers. This indicates that tumor growth in patients with this type of aggressive disease is still accompanied by AR expression. Although these patients have acquired resistance to several AR-targeting agents including NHA, the AR signaling may still be fostered by splice variants such as AR-V7 also observed in this cluster or by aberrant recruiting of other steroid receptors [30]. A similar $AR^{high}$ cluster was observed in CTCs of an advanced PCa patient cohort by Sperger *et al.* Patients in this cluster had a significantly worse OS and AR splice variants appeared as the most prominent resistance mechanism. The same study also investigated the expression of the NE markers CHGA and SYP, but these markers were not sufficient to identify an NE cluster [138]. In a rapid autopsy study, Labrecque *et al.* studied the transcriptome of 98 mCRPC tumors lesions by whole RNA sequencing. Five different phenotypes were identified based on hierarchical clustering that showed good concordance with the clusters identified in this study: $AR^{high}$, $AR^{low}$, amphicrine, neuroendocrine without AR expression and double-negative [41]. Similar to our clusters, amphicrine tumors had high expression of both PRAD and a limited number of NE markers, while pure NEPC patients showed a broader spectrum of NE marker expression. However, conclusions about amphicrine disease are overall limited due to the nature

of the bulk CTC analysis. As CTCs can be derived from different sites of metastasis within a patient, amphicrine expression profiles do not necessarily have to be the consequence of an amphicrine tumor with parallel expression of AR and NE markers in the same tumor or even the same cell, but might as well be the superposition of expression profiles, for instance from a bone metastasis with lumninal differentiation and a liver metastasis with pure small cell neuroendocrine phenotype. Single cell resolution is required to confirm a true amphicrine phenotype of the CTCs. A suitable method to identify a panel of transcripts in single cells without the need for expensive NGS analysis is the use of padlock probes for *in situ* staining of mRNA in enriched CTCs [109]. However, significant differences in NE marker profiles such as specific expression of *FOXA2* and *PCSK1* in the NEPC cluster, suggest biological variance between the two neuroendocrine clusters. In addition, histological analysis showed that patients in the amphicrine cluster rather had double positive tumors with expression of SYP but without morphological characteristics of small cell carcinoma. In contrast, most patients in the pure NEPC cluster were negative or only weakly positive for AR in histology and also showed morphological features of small cell carcinoma, such as dot-like CK staining. The frequency of liver metastasis was similar between both clusters, but survival analysis showed a trend towards worse prognosis for patients in the NEPC cluster. The collection of more patient samples is required in the future to explore the prognostic consequences of amphicrine compared to pure NEPC.

Interestingly, the $AR^{low}$ and the double-negative cluster identified by Labrecque *et al.* were most similar to our $CTC^{low}$ cluster, suggesting that the $CTC^{low}$ cluster might as well comprise patients with $AR^{low}$ or double-negative disease. As the DNPC cases were characterized by the absence of expression of AR and related markers as well as NE markers, these patients could hardly have been detected by our marker panel due to the lack of missing positive markers. The expression of epithelial markers is the only available indicator for CTCs in these samples, but epithelial markers alone are not sufficient to conclude the disease subtype. Similarly, $AR^{low}$ tumors with a high tumor burden would most likely be accompanied by CTC profiles that look similar to $AR^{high}$ tumors with a low CTC burden. The high CTC counts detected in a subset of patients in the $CTC^{low}$ cluster prove that an $AR^{low}$/double-negative subtype might indeed be hidden in the seemingly $CTC^{low}$ cluster. In addition, the frequency of liver metastasis and the survival of this cluster were also not significantly different from the $AR^{high}$ cluster and, thus,

suggest a similarly aggressive disease course for some patients in this cluster. As discussed before, more positive markers are required to identify DNPC tumors and confirm this hypothesis.

Significant differences were observed between the clusters that were not apparent in the clinical groups. One example is the epithelial and stemness marker *TACSTD2*. While *TACSTD2* positive samples were found in all clinical groups without significant differences in expression, significant up-regulation was identified in the $AR^{high}$ and amphicrine clusters compared to the NEPC cluster. Thus, TACSTD2 appears to be widely expressed in advanced PCa that maintains AR expression. Targeted therapies with antibody-drug conjugates directed against TACSTD2 have been developed and are already approved for triple-negative breast cancer [192]. Our data suggest the implementation of *TACSTD2* as a marker and therapeutic target also in personalized treatment of advanced PCa. Next, a significant up-regulation of RAI2 was observed in the amphicrine compared to the $AR^{high}$ cluster. RAI2 has recently been identified as a regulator of transcriptional repressors and histone methylation [193]. Thus, RAI2 is a promising biomarker for the progression to amphicrine PCa and a closer analysis of its function in transdifferentiation can improve our understanding of the underlying biological processes and facilitate the development of new therapies. *HOXB13* was the only PRAD associated marker that was not down-regulated in the NEPC cluster. This is in line with recent data from PCa tissue arrays confirming HOXB13 expression in AR-independent disease stages and its high specificity and sensitivity as a marker for prostate origin [194]. Thus, HOXB13 is a valuable marker to prove prostate origin, for instance in patients with a PSMA negative progress in the liver. Comparing the amphicrine and the pure NEPC clusters, *FOXA2* and *PCSK1* were the only NE markers with significant differential expression. PCSK1 has previously been shown to be specifically induced in pure small cell PCa, but not in amphicrine tumors [41]. These differences in NE marker expression profiles suggest, that the amphicrine cell state might be an intermediate step in the process of NET, whereas specific TFs such as FOXA2 that drive the full-blown NE differentiation require a loss of AR signaling to be induced [75]. This hypothesis was supported by observations from longitudinal CTC analysis. In three show cases with detection of NE markers at multiple time points, *FOXA2* and *PCSK1* were the last markers to become positive.

Additionally, PROM1 was specifically up-regulated in the NEPC cluster. It is known to be involved in the maintenance of the stem cell state and to be an in-

hibitor of differentiation [160]. Interestingly, in a cell culture model of androgen depletion, Sánchez *et al.* first observed the acquisition of NE features followed by the emergence of a cancer stem cell phenotype after prolonged androgen withdrawal [160]. This supports the idea of a multi-step transdifferentiation process, that also requires the activation of stemness pathways to ultimately switch off the AR pathway and develop a small cell neuroendocrine phenotype.

In conclusion, clusters identified based on gene expression profiles of CTCs are similar to the phenotypic clusters observed in primary tumors and metastatic lesions. Individual clusters might represent steps in the NET process and indicate specific targets for precision medicine. More positive markers are necessary to identify $AR^{low}$ and double-negative disease stages.

One of the major aims of the liquid biopsy test was to detect AR-negative, neuroendocrine relapse before the manifestation of clinical symptoms. A classifier was developed based on the gene expression profiles to predict the disease subtype. While the number of positive NE markers already enabled subtype prediction, a random forest classifier allowed robust identification of NEPC patients with an AUC of 95.5 % and a sensitivity of 90.9 % at 95.5 % specificity. In a similar study, Zhao and colleagues performed qPCR based detection of the NE markers CHGA and SYP in EPCAM-enriched CTCs. Longitudinal samples from mCRPC patients were collected for at least three time points. Their assay had a good specificity of 91 %, but a limited sensitivity of only 51 %. However, integrating the results of single samples from the same patient increased their per-patient prediction accuracy to 100 % [195]. In contrast, our assay showed that expression of only a single marker, even at multiple time points, was not predictive of NEPC. For example, patient 11 was positive for CEACAM5 at 3 out of 4 time points and showed aggressive but not neuroendocrine disease in histology. In contrast, patients with NEPC (e.g. patient 35) might have had varying numbers of NE markers expressed, but were positive at all time points, given they were CTC positive. Thus the inclusion of additional markers in our panel clearly improved the sensitivity for a single sample and improved the discrimination between aggressive disease and true NET.

*CHGA*, *ACTL6B*, *CECAM5*, *KLK3* and *KRT19* were among the most important markers for the model performance. This indicates that NE, epithelial as well as adenocarcinoma markers were required for accurate classification of advanced PCa patients. Other markers such as *LMO3*, *SOX11*, *PROM1* and *RAI2* had only a minimal influence on the classifier performance. Therefore, these markers could

be excluded from the panel to improve cost-effectiveness of the diagnostic test by avoiding irrelevant information. For instance, *SOX11* and *LMO3* showed no significant differences between the groups in the individual marker analysis. The adenocarcinoma markers were also highly correlated to each other. NKX3-1 was not among the highly ranked PRAD markers and was lost earlier than HOXB13. This could therefore be excluded from future analysis. A minimal marker panel that would be sufficient for clinical diagnosis could therefore include *EPCAM*, *KRT19*, *CHGA*, *ACTL6B*, *CEACAM5*, *FOXA2*, *PCSK1*, *AR*, *AR-V7*, *KLK3*, *HOXB13* and *EGFR*. The marker panel should include multiple NE markers, since it has been shown that the number of positive markers alone was a predictor for NEPC and because no individual marker was positive in all samples. Further markers such as *FOLH1* and *TACSTD2* should also be included, as these may provide information on potential targets for personalized treatment.

Cases of patients being incorrectly predicted by the classifier highlighted the weaknesses of the transcript panel. For instance, patients with double-positive expression profiles have repeatedly been mis-classified. This could be due to an imbalance of the data set with an over-representation of AVPC samples and the lack of specific markers for the AVPC and the amphicrine patients. This lack of specific markers is also reflected by the inability to discriminate the HSPC patients in a classifier for all three groups. Thus, more specific markers need to be included. Whole transcriptome analysis of amphicrine tumors compared to pure adenocarcinoma or pure NEPC might reveal more promising markers. In addition, the collection of more patient samples will result in a more balanced data set and will allow a sophisticated validation of the classifier based on an independent patient cohort.

## 4.5 Translation to clinical application

As new biomarkers are urgently needed to identify patients with PSA-negative relapse, a liquid biopsy test was developed to identify patients with NEPC. Although our assay showed robust performance in a cohort of more than one hundred PCa patients, additional steps are required to translate our pipeline into a diagnostic test for routine clinical application. Recently, establishment of clinical utility, standardization of platforms for analysis and reimbursement have been identified as the most challenging barriers for the application of liquid biopsy testing in clinical practice [196]. With respect to the liquid biopsy test validated

in this study, further data are required to prove the ability to overcome these barriers.

First, prospective collection of patient samples is required to evaluate the findings in the initially retrospectively collected cohort and prove clinical utility. In order to improve patient care in the clinic, the test needs to detect the emergence of NEPC before of clinical relapse. The lead time of the liquid biopsy assay must therefore be examined by longitudinal sampling of CRPC patients and a close follow up to compare how long before clinical relapse CTC profiles indicate disease progression. The case of patient 22 illustrates that CTC count is able to identify progression of NEPC before an incline of serum markers. Sperger *et al.* reported a case study where CTC-based transcript analysis of CHGA and SYP allowed the identification of NEPC with a lead time of 3 month before clinical progression to NEPC. A larger cohort of prospectively collected samples is required to further validate these initial results. Repeated measurements at six to eight week intervals would allow a close coverage to detect early signs of transdifferentiation and inform about the most feasible monitoring interval for routine clinical application.

Next, technical details need to be refined to adapt the pipeline to the requirements of routine clinical testing. Several parameters impact the ease of clinical implementation. These include the turnaround and hands-on time of the analytical pipeline. The workflow of CTC enrichment, cDNA synthesis and PCR-based transcript detection currently takes less than 24 hours. This is considerably faster than NGS-based methods that also require sophisticated bioinformatic data evaluation. In contrast, our workflow could provide fast results for timely treatment adjustment. Simple data evaluation based on predefined thresholds and the prediction of a classifier allow easy interpretation of the data. Additionally, hands-on time could be further reduced by automating of the workflow. Recently, the AdnaTest was adapted to work on the automated *EZ Connect* platform (Qiagen, Hilden) [197]. This includes the CTC enrichment as well as the mRNA isolation steps. However, the system is currently developed for research use only and is not approved for diagnostic purposes. For broad use, including in smaller laboratories, it would be required to establish the CTC enrichment and mRNA isolation workflow on a standard platform for automated pipetting that is also used for other routine assays. However, as especially PCR assays are performed at high efficiency when multiple samples are run in parallel, cost-efficient establishment of the analysis pipeline is possibly subjected to larger laboratories.

The current analysis pipeline is based on EDTA tubes for blood collection. This preserves cell viability and RNA integrity, but requires fast sample processing within three hours to avoid degradation. Thus, the test could not be performed in a central laboratory as this needs longer storage and transportation times. To make the test also available for smaller outpatient clinics, the use of RNA stabilizing tubes is essential. As the type of blood collection tube might influence the downstream analysis, a careful choice of a suitable collection tube is required [165]. Recently, several companies have developed RNA stabilizing blood collection tubes, such as the PAXgene tube by Qiagen or TEMPUS tubes by Applied Biosystems [198]. A close comparison of the available stabilizing agents and the test of compatibility with the downstream transcript detection pipeline is necessary to bring the assay closer to clinical application.

From a health economic perspective, a liquid biopsy test must be cost-effective to pose a benefit not only for the patient but also for the health care system. Especially for a biomarker intended for disease monitoring and, hence, repeated sampling, low costs are essential for broad use in the clinic. We chose a PCR-based approach for our liquid biopsy assay because it is less expensive than NGS-based methods with regard to the reagents and instrumentation required. While the current workflow still requires manual work with a lot of hands-on time, automation will reduce this and, thus, the personal cost. Cost-effectiveness in terms of value for money will need to be demonstrated in order to achieve reimbursement from public health care providers. This is required to offer the test to all eligible patients. Although meta-analysis concerning the use of liquid biopsy in early stage prostate cancer show benefits in terms of cost and cost-effectiveness, additional studies are required to also confirm benefits in late stage disease monitoring [199].

Finally, biomarker improvement needs to be accompanied by the development of new personalized therapy options. Today, treatment options for NEPC are mostly limited to platinum-based chemotherapy and, thus, the earlier detection of progression to NEPC alone will not greatly improve patient outcomes. Several treatment options are currently in clinical testing in patients with NEPC and AVPC including novel small molecule inhibitors and immune therapies. For NEPC patients, ongoing clinical trials are evaluating the use of epigenetic therapies in the form of EZH2 and LSD1 inhibitors to counteract NET. An LSD1 inhibitor has recently shown promising results in advanced solid tumors and especially in neuroendocrine carcinomas [200]. Currently, this inhibitor is tested in a separate clin-

ical trial for CRPC patients with progression on enzalutamide (NCT04628988). Pre-clinical data points towards reversal of transdifferentiation following EZH2 inhibition in NEPC models [201]. Thus, EZH2 inhibition is currently being investigated for SCLC and CRPC with resistance to NHA (NCT03460977). Based on the high frequency of AURKA amplification observed in NEPC, inhibitors have been evaluated in clinical trials, however primary endpoints were not met despite some exceptional responders. This highlights the need for biomarkers and companion diagnostics [202]. In AVPC, the combination of cabazitaxel-carboplatin chemotherapy with olaparib maintenance therapy is currently being tested, but results are not available, yet (NCT03263650). This concept is based on the observation of a BRCAness phenotype in CRPC with synthetic lethality following PARP inhibition [203]. In addition, the combination of immune checkpoint blockade with either NHA or tyrosine kinase inhibitors is currently tested for patients with treatment-emergent NEPC (NCT04926181, NCT04848337). As upregulation of receptor tyrosine kinases such as RET, ALK or VEGFR is associated with NEPC, this concept has shown promising results in preclinical studies so far [204], [205]. In neuroendocrine carcinomas of various origins, an anti-DLL3 antibody drug-conjugate showed clinical benefit in the majority of NEPC patients although the objective response rate was limited [206]. Likewise, DLL3-targeting T cell engagers are currently being tested for DLL3 positive, neuroendocrine tumors including NEPC in phase I/II clinical trials and a previous study in SCLC has already suggested improved benefit from biomarker-assisted patient selection (NCT04471727, NCT04702737) [207], [208]. Similar to anti-DLL3 therapy, liquid biopsy enables a tumor-agnostic, personalized therapy selection, depending on the respective expression of mutational profiles. For example, anti-TROP2 antibody-drug conjugates are approved for triple-negative breast cancer and are now being investigated more broadly for advanced epithelial malignancies (NCT04152499) [209]. TROP2 was one of the markers analyzed in this study and results suggest a high expression in AVPC patients. In breast cancer, strong TROP2 expression in tissue biopsy was associated with improved overall survival, encouraging closer examination of the predictive value of TROP2 expression in CTCs for the selection of treatment, also in PCa [210]. Similarly, patients with mutations in *BRCA1/2* or other homologous recombination repair genes are eligible for PARP inhibitor treatment [211]. Consequently, the precise identification of actionable targets might be more important than the accurate classification of the tumor subtype in an anyhow highly heterogeneous disease. In conclusion, a sophisticated marker panel to detect AVPC and NEPC patients should include

prognostic biomarkers for the disease subtype and predictive biomarkers to support therapy decisions.

## 4.6 Conclusion and Outlook

The aim of this project was to identify biomarkers for neuroendocrine transdifferentiation in patients suffering from advanced prostate cancer. For this purpose, blood was collected from more than 100 PCa patients with known disease stage and CTCs were enriched. 22 markers identified by literature research and validated in cell lines and published data sets were analyzed in bulk CTC fractions. The transcript analysis revealed a high degree of inter-patient heterogeneity in CTC populations with distinct expression patterns between the disease subtypes. This allowed robust prediction of the NEPC samples and identified treatment relevant molecular targets. Longitudinal case studies emphasized the advantages of liquid biopsy-based disease monitoring for early detection of relapse.

The translation of the liquid biopsy assay for NEPC into clinical application will require systematic prospective, longitudinal analysis of patients to assess the lead time of the liquid biopsy assay compared to clinical parameters. Precise definition of inclusion criteria and careful allocation of the heterogeneous patients to the disease stages will be essential to obtain reliable results. Automation of the workflow and removal of redundant and informative markers can further increase the cost-effectiveness of the assay. Most importantly, therapy options need to be refined in AVPC and NEPC. The inter-patient heterogeneity observed in this cohort indicates the necessity for personalized medicine in advanced PCa that also considers molecular markers in addition to macroscopic clinical parameters. Our liquid biopsy assay enables the identification of therapeutic targets such as TROP2 as well as the monitoring of therapy response and, thus, paves the way for precision medicine in advanced PCa. Future research will need to deal with additional subtypes of AVPC such as $AR^{low}$ PCa and DNPC to develop more sophisticated assays that can detect all AR-independent relapses. CTCs as well as cfDNA offer the possibility to explore epigenetic features such as methylome profiles to further refine monitoring and diagnosis of AVPC patients beyond NET.

# Abstract

Standard therapy for prostate cancer is directed against the androgen receptor (AR) signaling pathway. As a result of increased therapeutic pressure from novel hormonal agents, patients are more frequently progressing to AR-independent, aggressive variant prostate cancer (AVPC). The acquisition of neuroendocrine features in a transdifferentiation process leads to the emergence of neuroendocrine prostate cancer (NEPC) and presents one resistance mechanism to AR-targeted therapy. As PSA no longer is a reliable biomarker in AVPC and NEPC, there is an urgent need to identify new biomarkers for AR-independent progression and transdifferentiation. The molecular analysis of circulating tumor cells (CTC) and tumor cell fragments shed into body fluids is known as liquid biopsy (LBx). LBx poses significantly reduced risk to patients, can be performed repeatedly and represents tumor heterogeneity more effectively than a tissue biopsy. Thus, this study aimed to establish a liquid biopsy-based biomarker for the detection and monitoring of patients with NEPC.

For this purpose, a transcript panel was selected based on literature review and validated in cell lines and published tumor tissue data sets. The combined workflow of CTC enrichment and transcript detection was validated in spike-in controls. Blood samples were collected from patients with AVPC, NEPC and hormone-sensitive prostate cancer at the University Medical Center Hamburg-Eppendorf. CTC counts were measured by CellSearch. After validation of the enrichment methods, the expression of a 22 gene panel was analyzed in AdnaTest-enriched CTCs. Gene expression profiles were evaluated using supervised and unsupervised approaches and correlated to clinical data.

Comparison of different enrichment methods showed that CTCs were more frequently detected with EPCAM-based CTC enrichment. The detection of neuroendocrine markers was not improved by Parsortix enrichment compared to AdnaTest. CellSearch analysis revealed significantly increased CTC counts in NEPC

and AVPC patients accompanied by a heterogeneous CTC morphology. Gene expression profiles revealed a high degree of heterogeneity between patients with significant deregulation of several individual markers. Unsupervised analysis identified four distinct clusters, which were termed $AR^{high}$, $CTC^{low}$, amphicrine and pure NEPC. Using a random forest model, HSPC and NEPC samples could be distinguished with an AUC of 95.5 % and an out-of-bag error rate of 15.5 % in cross-validation. In longitudinal samples of single patients, the detection of neuroendocrine markers in CTCs recapitulated the clinical course of therapy response and progression.

In conclusion, AVPC and NEPC patients had a high CTC burden, which facilitated subsequent molecular analyses. The gene expression profiles of the marker panel in CTCs reflected the histology of the tumors and were sufficient to distinguish intrinsic molecular subtypes of advanced prostate cancer. In the future, the convenient PCR-based analysis pipeline may allow monitoring of advanced prostate cancer patients for earlier adjustment of therapy for NEPC.

# Zusammenfassung

Die zielgerichtete Therapie des Prostatakarzinoms beruht üblicherweise auf der Hemmung des Androgenrezeptor (AR) Signalwegs. Neue hormonartige Wirkstoffe haben zuletzt gute Behandlungserfolge erzielt, jedoch kann der erhöhte Behandlungsdruck bei einigen Patienten zum Progress mit einer AR-unabhängigen, aggressiven Variante des Prostatakarzinoms (AVPC) führen. Der Erwerb neuroendokriner Eigenschaften im Rahmen einer Transdifferenzierung stellt einen möglichen Resistenzmechanismus dar und manifestiert sich klinisch als therapieinduziertes neuroendokrines Prostatakarzinom (NEPC). Da PSA in diesem Stadium der Erkrankung nicht länger als verlässlicher Biomarker verwendet werden kann, ist es dringend notwendig neue Marker für die AR-unabhängigen Stadien zu entwickeln. Die molekulare Analyse von zirkulierenden Tumorzellen (CTC) und anderen Tumorzell-assoziierten Biomolekülen in Körperflüssigkeiten wird als Flüssigbiopsie bezeichnet. Diese ist mit einem deutlich geringeren Risiko für die Patienten verbunden als eine herkömmliche Gewebebiopsie, kann einfacher wiederholt werden und bildet die Heterogenität der Erkankung häufig besser ab. Ziel dieser Arbeit war daher die Identifizierung und Validierung Flüssigbiopsiebasierter Biomarker für die Detektion von NEPC und das Monitoring von betroffenen Patienten.

Zu diesem Zweck wurde ein Panel von Transkripten auf Basis einer Literaturrecherche ausgewählt und in Zelllinien und publizierten Datensätzen Genomweiter Genexpression relevanter Tumorerkrankungen validiert. Der kombinierte Ablauf von CTC-Anreicherung und Transkript-Detektion wurde mit Zelllinienbasierten Kontrollen geprüft. Blutproben wurden von AVPC, NEPC und hormonsensitiven Kontrollpatienten am Universitätsklinikum Hamburg-Eppendorf gesammelt. Die CTC-Anzahl wurde mittels CellSearch bestimmt. Nach der Validierung verschiedener Anreicherungsmethoden wurde das 22-Gen-Panel in Adna-Test-angereicherten CTCs untersucht. Die Genexpressionsprofile wurden mit über-

wachten und unüberwachten Methoden analysiert und mit den klinischen Daten korreliert.

Im Vergleich der verschiedenen Anreicherungsmethoden wurden mit dem AdnaTest mehr CTC-positive Proben als mit CellSearch und Parsortix identifiziert. Die Detektion der neuroendokrinen Marker war in Parsortix-angereicherten CTCs im Vergleich mit dem AdnaTest nicht verbessert. Insgesamt zeigte die CellSearch Analyse eine heterogene Zellmorphologie und eine signifikant erhöhte Anzahl von CTCs in AVPC und NEPC Patienten. Die Genexpressionsprofile verdeutlichten die Heterogenität der Patientenkohorte, innerhalb derer zahlreiche neuroendokrine und Adenokarzinom-assoziierte Transkripte signifikant dereguliert waren. Basierend auf den Expressionsprofilen wurden vier Cluster identifiziert, die als $AR^{hoch}$, $CTC^{niedrig}$, rein neuroendokrin und amphikrin zusammengefasst werden konnten. Ein Random-Forest-Model mit einer AUC von 95.5 % konnte zudem zuverlässig zwischen der NEPC- und der Kontrollgruppe unterscheiden. In der longitudinalen Analyse einzelner Patienten konnte der klinische Verlauf anhand der CTC-Daten rekapituliert werden.

Zusammenfassend wurde in dieser Arbeit eine hohe CTC-Last in AVPC und NEPC Patienten nachgewiesen, was die nachfolgende molekulare Analyse erleichterte. Die Expressionsprofile des ausgewählten Markerpanels in den CTCs spiegelten die Histologie der Tumoren wider und erlaubten die Unterscheidung molekularer Subtypen innerhalb des fortgeschrittenen Prostatakarzinoms. Zukünftig könnte diese erfolgreich entwickelte, PCR-basierte Analyse genutzt werden, um die Entstehung von NEPC früher zu erkennen und die Therapie der entsprechenden Patienten gezielt anzupassen.

# List of Abbreviations

| Abbreviation | Meaning |
|---|---|
| AR | Androgen receptor |
| ADT | Androgen deprivation therapy |
| AMPC | Amphicrine prostate cancer |
| AUC | Area under the curve |
| AURKA | Aurora kinase A |
| AVPC | Aggressive variant prostate cancer |
| cDNA | Complementary DNA |
| CEA | Carcinoembryonic antigen |
| cfDNA | Cell-free DNA |
| CK | Cyto-keratin |
| ctDNA | Circulating tumor DNA |
| CRPC | Castration-resistant prostate cancer |
| CTC | Circulating tumor cell |
| ddPCR | Droplet digital PCR |
| DMEM | Dulbecco's modified Eagle medium |
| DNA | Deoxyribonucleic acid |
| DNPC | Double negative prostate cancer |
| EMT | Epithelial-to-mesenchymal transition |
| EV | Extracellular vesicle |
| FCS | Fetal calve serum |
| gDNA | Genomic DNA |
| GOI | Gene of interest |
| HSPC | Hormone-sensitive prostate cancer |
| IPC | Interplate calibrator |
| LDH | Lactate dehydrogenase |
| lncRNA | Long noncoding RNA |

| Abbreviation | Meaning |
|---|---|
| mRNA | Messenger RNA |
| NE | Neuroendocrine |
| NEPC | Neuroendocrine prostate cancer |
| NET | Neuroendocrine transdifferentiation |
| NGS | Next-generation sequencing |
| NHA | Novel hormonal agents |
| NSE | Neuron-specific enolase |
| OS | Overall survival |
| PCa | Prostate cancer |
| PBS | Phosphate-buffered saline |
| PBMC | Peripheral blood mononuclear cell |
| PCR | Polymerase chain reaction |
| PDX | Patient-derived xenograft |
| PFS | Progression-free survival |
| PRAD | Prostate adenocarcinoma |
| PSA | Prostate-specific antigen |
| qPCR | quantitative PCR |
| RNA | Ribonucleic acid |
| ROC | Receiver Operator Characteristic |
| SCLC | Small cell lung cancer |
| TF | Transcription factor |

# References


[1] H. Sung, J. Ferlay, R. L. Siegel, *et al.*, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA. Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, 2021.

[2] G. Gandaglia, R. Leni, F. Bray, *et al.*, "Epidemiology and Prevention of Prostate Cancer," *Eur. Urol. Oncol.*, vol. 4, no. 6, pp. 877–892, 2021.

[3] E. Castro, N. Romero-Laorden, A. del Pozo, *et al.*, "PROREPAIR-B: A Prospective Cohort Study of the Impact of Germline DNA Repair Mutations on the Outcomes of Patients With Metastatic Castration-Resistant Prostate Cancer," *J. Clin. Oncol.*, vol. 37, no. 6, pp. 490–503, 2019.

[4] T. Nyberg, D. Frost, D. Barrowdale, *et al.*, "Prostate Cancer Risks for Male BRCA1 and BRCA2 Mutation Carriers: A Prospective Cohort Study," *Eur. Urol.*, vol. 77, no. 1, pp. 24–35, 2020.

[5] R. J. Rebello, C. Oing, K. E. Knudsen, *et al.*, "Prostate cancer," *Nat. Rev. Dis. Prim.*, vol. 7, no. 1, p. 9, 2021.

[6] G. Wang, D. Zhao, D. J. Spring, and R. A. DePinho, "Genetics and biology of prostate cancer," *Genes Dev.*, vol. 32, no. 17-18, pp. 1105–1140, 2018.

[7] X. Ma, J. Guo, K. Liu, *et al.*, "Identification of a distinct luminal subgroup diagnosing and stratifying early stage prostate cancer by tissue-based single-cell RNA sequencing," *Mol. Cancer*, vol. 19, no. 1, p. 147, 2020.

[8] M. S. Lawrence, P. Stojanov, P. Polak, *et al.*, "Mutational heterogeneity in cancer and the search for new cancer-associated genes," *Nature*, vol. 499, no. 7457, pp. 214–218, 2013.

[9] A. Abeshouse, J. Ahn, R. Akbani, *et al.*, "The Molecular Taxonomy of Primary Prostate Cancer," *Cell*, vol. 163, no. 4, pp. 1011–1025, 2015.

[10] M. J. Ryan and R. Bose, "Genomic Alteration Burden in Advanced Prostate Cancer and Therapeutic Implications," *Front. Oncol.*, vol. 9, 2019.

[11] S. Perner, F. Demichelis, R. Beroukhim, *et al.*, "TMPRSS2:ERG Fusion-Associated Deletions Provide Insight into the Heterogeneity of Prostate Cancer," *Cancer Res.*, vol. 66, no. 17, pp. 8337–8341, 2006.

[12] G. von Amsberg, W. Alsdorf, P. Karagiannis, *et al.*, "Immunotherapy in Advanced Prostate Cancer-Light at the End of the Tunnel?" *Int. J. Mol. Sci.*, vol. 23, no. 5, 2022.

[13] P. C. Boutros, M. Fraser, N. J. Harding, *et al.*, "Spatial genomic heterogeneity within localized, multifocal prostate cancer," *Nat. Genet.*, vol. 47, no. 7, pp. 736–745, 2015.

[14] M. Løvf, S. Zhao, U. Axcrona, *et al.*, "Multifocal Primary Prostate Cancer Exhibits High Degree of Genomic Heterogeneity," *Eur. Urol.*, vol. 75, no. 3, pp. 498–505, 2019.

[15] G. Gundem, P. Van Loo, B. Kremeyer, *et al.*, "The evolutionary history of lethal metastatic prostate cancer," *Nature*, vol. 520, no. 7547, pp. 353–357, 2015.

[16] N. W. Clarke, A. Ali, F. C. Ingleby, *et al.*, "Addition of docetaxel to hormonal therapy in low- and high-burden metastatic hormone sensitive prostate cancer: long-term survival results from the STAMPEDE trial.," *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, vol. 30, no. 12, pp. 1992–2003, 2019.

[17] K. Fizazi, N. Tran, L. Fein, *et al.*, "Abiraterone acetate plus prednisone in patients with newly diagnosed high-risk metastatic castration-sensitive prostate cancer (LATITUDE): final overall survival analysis of a randomised, double-blind, phase 3 trial," *Lancet Oncol.*, vol. 20, no. 5, pp. 686–700, 2019.

[18] K. N. Chi, S. Chowdhury, A. Bjartell, *et al.*, "Apalutamide in Patients With Metastatic Castration-Sensitive Prostate Cancer: Final Survival Analysis of the Randomized, Double-Blind, Phase III TITAN Study," *J. Clin. Oncol.*, vol. 39, no. 20, pp. 2294–2303, 2021.

[19] K. Fizazi, J. Carles Galceran, S. Foulon, *et al.*, "LBA5 A phase III trial with a 2x2 factorial design in men with de novo metastatic castration-sensitive prostate cancer: Overall survival with abiraterone acetate plus prednisone in PEACE-1," *Ann. Oncol.*, vol. 32, S1299, 2021.

[20] K. Haapala, T. Kuukasjarvi, E. Hyytinen, I. RantalaA, H. Helin, and P. Koivisto, "Androgen receptor amplification is associated with increased cell proliferation in prostate cancer," *Hum. Pathol.*, vol. 38, no. 3, pp. 474–478, 2007.

[21] A. Köhler, Ü. Demir, E. Kickstein, *et al.*, "A hormone-dependent feedback-loop controls androgen receptor levels by limiting MID1, a novel translation enhancer and promoter of oncogenic signaling," *Mol. Cancer*, vol. 13, no. 1, p. 146, 2014.

[22] L. Gao, W. Zhang, J. Zhang, *et al.*, "KIF15-Mediated Stabilization of AR and AR-V7 Contributes to Enzalutamide Resistance in Prostate Cancer," *Cancer Res.*, vol. 81, no. 4, pp. 1026–1039, 2021.

[23] O. Snow, N. Lallous, K. Singh, N. Lack, P. Rennie, and A. Cherkasov, "Androgen receptor plasticity and its implications for prostate cancer therapy," *Cancer Treat. Rev.*, vol. 81, p. 101 871, 2019.

[24] K. Fizazi, H. I. Scher, A. Molina, *et al.*, "Abiraterone acetate for treatment of metastatic castration-resistant prostate cancer: final overall survival analysis of the COU-AA-301 randomised, double-blind, placebo-controlled phase 3 study," *Lancet Oncol.*, vol. 13, no. 10, pp. 983–992, 2012.

[25] T. M. Beer, A. J. Armstrong, D. Rathkopf, *et al.*, "Enzalutamide in Men with Chemotherapy-naïve Metastatic Castration-resistant Prostate Cancer: Extended Analysis of the Phase 3 PREVAIL Study," *Eur. Urol.*, vol. 71, no. 2, pp. 151–154, 2017.

[26] M. Korpal, J. M. Korn, X. Gao, *et al.*, "An F876L Mutation in Androgen Receptor Confers Genetic and Phenotypic Resistance to MDV3100 (Enzalutamide)," *Cancer Discov.*, vol. 3, no. 9, pp. 1030–1043, 2013.
[27] Y. Zhao, L. Wang, S. Ren, *et al.*, "Activation of P-TEFb by Androgen Receptor-Regulated Enhancer RNAs in Castration-Resistant Prostate Cancer," *Cell Rep.*, vol. 15, no. 3, pp. 599–610, 2016.
[28] D. Xu, Y. Zhan, Y. Qi, *et al.*, "Androgen Receptor Splice Variants Dimerize to Transactivate Target Genes," *Cancer Res.*, vol. 75, no. 17, pp. 3663–3671, 2015.
[29] S. Bai, S. Cao, L. Jin, *et al.*, "A positive role of c-Myc in regulating androgen receptor and its splice variants in prostate cancer," *Oncogene*, vol. 38, no. 25, pp. 4977–4989, 2019.
[30] V. K. Arora, E. Schenkein, R. Murali, *et al.*, "Glucocorticoid Receptor Confers Resistance to Antiandrogens by Bypassing Androgen Receptor Blockade," *Cell*, vol. 155, no. 6, pp. 1309–1322, 2013.
[31] E. G. Bluemn, I. M. Coleman, J. M. Lucas, *et al.*, "Androgen Receptor Pathway-Independent Prostate Cancer Is Sustained through FGF Signaling," *Cancer Cell*, vol. 32, no. 4, 474–489.e6, 2017.
[32] H. Beltran, S. Tomlins, A. Aparicio, *et al.*, "Aggressive variants of castration-resistant prostate cancer.," *Clin. Cancer Res.*, vol. 20, no. 11, pp. 2846–50, 2014.
[33] J. I. Epstein, M. B. Amin, H. Beltran, *et al.*, "Proposed Morphologic Classification of Prostate Cancer With Neuroendocrine Differentiation," *Am. J. Surg. Pathol.*, vol. 38, no. 6, pp. 756–767, 2014.
[34] A. M. Aparicio, A. L. Harzstark, P. G. Corn, *et al.*, "Platinum-Based Chemotherapy for Variant Castrate-Resistant Prostate Cancer," *Clin. Cancer Res.*, vol. 19, no. 13, pp. 3621–3630, 2013.
[35] R. R. Aggarwal, D. A. Quigley, J. Huang, *et al.*, "Whole-Genome and Transcriptional Analysis of Treatment-Emergent Small-Cell Neuroendocrine Prostate Cancer Demonstrates Intraclass Heterogeneity," *Mol. Cancer Res.*, vol. 17, no. 6, pp. 1235–1240, 2019.
[36] S. W. Fine, "Neuroendocrine tumors of the prostate," *Mod. Pathol.*, vol. 31, no. S1, pp. 122–132, 2018.
[37] W. Bai, C. Zhen, R. Zhang, W. Yu, and Z. Zhou, "Clinicopathological features of patients with transformation from EGFR mutant lung adenocarcinoma to small cell lung cancer," *Transl. Cancer Res.*, vol. 10, no. 8, pp. 3694–3704, 2021.
[38] W. N. Brennen, Y. Zhu, I. M. Coleman, *et al.*, "Resistance to androgen receptor signaling inhibition does not necessitate development of neuroendocrine prostate cancer," *JCI Insight*, 2021.
[39] F. Tang, D. Xu, S. Wang, *et al.*, "Chromatin profiles classify castration-resistant prostate cancers suggesting therapeutic targets," *Science (80-. ).*, vol. 376, no. 6596, 2022.
[40] H.-C. Lee, C.-H. Ou, Y.-C. Huang, *et al.*, "YAP1 overexpression contributes to the development of enzalutamide resistance by induction of cancer

stemness and lipid metabolism in prostate cancer," *Oncogene*, vol. 40, no. 13, pp. 2407–2421, 2021.

[41] M. P. Labrecque, I. M. Coleman, L. G. Brown, *et al.*, "Molecular profiling stratifies diverse phenotypes of treatment-refractory metastatic castration-resistant prostate cancer," *J. Clin. Invest.*, vol. 129, no. 10, pp. 4492–4505, 2019.

[42] H. Sawazaki, A. Asano, Y. Kitamura, J. Katsuta, and Y. Ito, "Androgen receptorneuroendocrine doublenegative tumor with squamous differentiation arising from treatmentrefractory metastatic castrationresistant prostate cancer," *IJU Case Reports*, vol. 4, no. 6, pp. 417–420, 2021.

[43] P. G. Corn, E. I. Heath, A. Zurita, *et al.*, "Cabazitaxel plus carboplatin for the treatment of men with metastatic castration-resistant prostate cancers: a randomised, open-label, phase 12 trial," *Lancet Oncol.*, vol. 20, no. 10, pp. 1432–1443, 2019.

[44] A. Fléchon, D. Pouessel, C. Ferlay, *et al.*, "Phase II study of carboplatin and etoposide in patients with anaplastic progressive metastatic castration-resistant prostate cancer (mCRPC) with or without neuroendocrine differentiation: results of the French Genito-Urinary Tumor Group (GETUG) P01 trial," *Ann. Oncol.*, vol. 22, no. 11, pp. 2476–2481, 2011.

[45] G. von Amsberg, M. Zilles, W. Mansour, *et al.*, "Salvage Chemotherapy with Cisplatin, Ifosfamide, and Paclitaxel in Aggressive Variant of Metastatic Castration-Resistant Prostate Cancer," *Int. J. Mol. Sci.*, vol. 23, no. 23, p. 14 948, 2022.

[46] J. de Bono, J. Mateo, K. Fizazi, *et al.*, "Olaparib for Metastatic Castration-Resistant Prostate Cancer," *N. Engl. J. Med.*, vol. 382, no. 22, pp. 2091–2102, 2020.

[47] L. C. Brown, S. Halabi, J. A. Somarelli, *et al.*, "A phase 2 trial of avelumab in men with aggressive-variant or neuroendocrine prostate cancer.," *Prostate Cancer Prostatic Dis.*, 2022.

[48] V. Conteduca, C. Oromendia, K. W. Eng, *et al.*, "Clinical features of neuroendocrine prostate cancer," *Eur. J. Cancer*, vol. 121, pp. 7–18, 2019.

[49] L. S. Graham, M. C. Haffner, E. Sayar, *et al.*, "Clinical, pathologic, and molecular features of amphicrine prostate cancer," *Prostate*, vol. 83, no. 7, pp. 641–648, 2023.

[50] H. Beltran, S. T. Tagawa, K. Park, *et al.*, "Challenges in Recognizing Treatment-Related Neuroendocrine Prostate Cancer," *J. Clin. Oncol.*, vol. 30, no. 36, e386–e389, 2012.

[51] W. Abida, J. Cyrta, G. Heller, *et al.*, "Genomic correlates of clinical outcome in advanced prostate cancer," *Proc. Natl. Acad. Sci.*, vol. 116, no. 23, pp. 11 428–11 436, 2019.

[52] R. Aggarwal, J. Huang, J. J. Alumkal, *et al.*, "Clinical and genomic characterization of treatment-emergent small-cell neuroendocrine prostate cancer: A multi-institutional prospective study," *J. Clin. Oncol.*, vol. 36, no. 24, pp. 2492–2503, 2018.

[53] S. Alanee, A. Moore, M. Nutt, *et al.*, "Contemporary Incidence and Mortality Rates of Neuroendocrine Prostate Cancer.," *Anticancer Res.*, vol. 35, no. 7, pp. 4145–50, 2015.

[54] G. Von Amsberg, S. Dyshlovoy, J. Hauschild, *et al.*, "Long-term taxane exposure and transdifferentiation of prostate cancer in vitro.," *J. Clin. Oncol.*, vol. 41, no. 6_suppl, pp. 254–254, 2023.

[55] X. Deng, B. D. Elzey, J. M. Poulson, *et al.*, "Ionizing radiation induces neuroendocrine differentiation of prostate cancer cells in vitro, in vivo and in prostate cancer patients.," *Am. J. Cancer Res.*, vol. 1, no. 7, pp. 834–44, 2011.

[56] D. K. Lee, Y. Liu, L. Liao, W. Li, D. Danielpour, and J. Xu, "Neuroendocrine prostate carcinoma cells originate from the p63-expressing basal cells but not the pre-existing adenocarcinoma cells in mice," *Cell Res.*, vol. 29, no. 5, pp. 420–422, 2019.

[57] H. Chen, Y. Sun, C. Wu, *et al.*, "Pathogenesis of prostatic small cell carcinoma involves the inactivation of the P53 pathway.," *Endocr. Relat. Cancer*, vol. 19, no. 3, pp. 321–31, 2012.

[58] Y.-H. Huang, Y.-Q. Zhang, and J.-T. Huang, "Neuroendocrine cells of prostate cancer: biologic functions and molecular mechanisms," *Asian J. Androl.*, vol. 21, no. 3, p. 291, 2019.

[59] T. L. Lotan, N. S. Gupta, W. Wang, *et al.*, "ERG gene rearrangements are common in prostatic small cell carcinomas," *Mod. Pathol.*, vol. 24, no. 6, pp. 820–828, 2011.

[60] V. J. Scheble, M. Braun, T. Wilbertz, *et al.*, "ERG rearrangement in small cell prostatic and lung cancer," *Histopathology*, vol. 56, no. 7, pp. 937–943, 2010.

[61] J. E. Vellky and W. A. Ricke, "Development and prevalence of castration-resistant prostate cancer subtypes," *Neoplasia (United States)*, vol. 22, no. 11, pp. 566–575, 2020.

[62] H. Beltran, D. Prandi, J. M. Mosquera, *et al.*, "Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer," *Nat. Med.*, vol. 22, no. 3, pp. 298–305, 2016.

[63] B. Dong, J. Miao, Y. Wang, *et al.*, "Single-cell analysis supports a luminal-neuroendocrine transdifferentiation in human prostate cancer," *Commun. Biol.*, vol. 3, no. 1, p. 778, 2020.

[64] J. A. Fraser, J. E. Sutton, S. Tazayoni, I. Bruce, and A. V. Poole, "hASH1 nuclear localization persists in neuroendocrine transdifferentiated prostate cancer cells, even upon reintroduction of androgen," *Sci. Rep.*, vol. 9, no. 1, p. 19 076, 2019.

[65] A. H. Davies, H. Beltran, and A. Zoubeidi, "Cellular plasticity and the neuroendocrine phenotype in prostate cancer," *Nat. Rev. Urol.*, vol. 15, no. 5, pp. 271–286, 2018.

[66] H. Bonkhoff, "Neuroendocrine cells in benign and malignant prostate tissue: morphogenesis, proliferation, and androgen receptor status.," *Prostate. Suppl.*, vol. 8, pp. 18–22, 1998.

[67] M. T. Spiotto and T. D. Chung, "STAT3 mediates IL-6-induced neuroendocrine differentiation in prostate cancer cells," *Prostate*, vol. 42, no. 3, pp. 186–195, 2000.

[68] E.-C. Hsu, M. A. Rice, A. Bermudez, *et al.*, "Trop2 is a driver of metastatic prostate cancer with neuroendocrine phenotype via PARP1," *Proc. Natl. Acad. Sci.*, vol. 117, no. 4, pp. 2032–2042, 2020.

[69] D. McKeithen, T. Graham, L. W. Chung, and V. Odero-Marah, "Snail transcription factor regulates neuroendocrine differentiation in LNCaP prostate cancer cells," *Prostate*, vol. 70, no. 9, pp. 982–992, 2010.

[70] L. Merkens, V. Sailer, D. Lessel, *et al.*, "Aggressive variants of prostate cancer: underlying mechanisms of neuroendocrine transdifferentiation," *J. Exp. Clin. Cancer Res.*, vol. 41, no. 1, p. 46, 2022.

[71] A. M. Aparicio, L. Shen, E. L. N. Tapia, *et al.*, "Combined Tumor Suppressor Defects Characterize Clinically Defined Aggressive Variant Prostate Cancers," *Clin. Cancer Res.*, vol. 22, no. 6, pp. 1520–1530, 2016.

[72] S. Y. Ku, S. Rosario, Y. Wang, *et al.*, "Rb1 and Trp53 cooperate to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance," *Sci. (80-. ).*, vol. 355, no. 6320, pp. 78–83, 2017.

[73] P. Mu, Z. Zhang, M. Benelli, *et al.*, "SOX2 promotes lineage plasticity and antiandrogen resistance in TP53-and RB1-deficient prostate cancer," *Science (80-. ).*, vol. 355, no. 6320, pp. 1–6, 2017.

[74] S. Nouruzi, D. Ganguli, N. Tabrizian, *et al.*, "ASCL1 activates neuronal stem cell-like lineage programming through remodeling of the chromatin landscape in prostate cancer," *Nat. Commun.*, vol. 13, no. 1, p. 2282, 2022.

[75] M. Han, F. Li, Y. Zhang, *et al.*, "FOXA2 drives lineage plasticity and KIT pathway activation in neuroendocrine prostate cancer," *Cancer Cell*, vol. 40, no. 11, 1306–1323.e8, 2022.

[76] J. L. Bishop, D. Thaper, S. Vahid, *et al.*, "The Master Neural Transcription Factor BRN2 Is an Androgen ReceptorSuppressed Driver of Neuroendocrine Differentiation in Prostate Cancer," *Cancer Discov.*, vol. 7, no. 1, pp. 54–71, 2017.

[77] S. Ji, Y. Shi, L. Yang, F. Zhang, Y. Li, and F. Xu, "miR-145-5p Inhibits Neuroendocrine Differentiation and Tumor Growth by Regulating the SOX11/-MYCN Axis in Prostate cancer," *Front. Genet.*, vol. 13, 2022.

[78] S. C. Baca, D. Y. Takeda, J.-H. Seo, *et al.*, "Reprogramming of the FOXA1 cistrome in treatment-emergent neuroendocrine prostate cancer," *Nat. Commun.*, vol. 12, no. 1, p. 1979, 2021.

[79] J. Cyrta, A. Augspach, M. R. De Filippo, *et al.*, "Role of specialized composition of SWI/SNF complexes in prostate cancer lineage plasticity," *Nat. Commun.*, vol. 11, no. 1, pp. 1–16, 2020.

[80] A. Davies, S. Nouruzi, D. Ganguli, *et al.*, "An androgen receptor switch underlies lineage infidelity in treatment-resistant prostate cancer," *Nat. Cell Biol.*, vol. 23, no. 9, pp. 1023–1034, 2021.

[81] A. Sehrawat, L. Gao, Y. Wang, *et al.*, "LSD1 activates a lethal prostate cancer gene network independently of its demethylase function," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 18, E4179–E4188, 2018.

[82] A. R. Lee, Y. Li, N. Xie, *et al.*, "Alternative RNA splicing of the MEAF6 gene facilitates neuroendocrine prostate cancer progression," *Oncotarget*, vol. 8, no. 17, pp. 27 966–27 975, 2017.

[83] J. Lovnicki, Y. Gan, T. Feng, *et al.*, "LIN28B promotes the development of neuroendocrine prostate cancer," *J. Clin. Invest.*, vol. 130, no. 10, pp. 5338–5348, 2020.

[84] S. Akamatsu, A. W. Wyatt, D. Lin, *et al.*, "The Placental Gene PEG10 Promotes Progression of Neuroendocrine Prostate Cancer," *Cell Rep.*, vol. 12, no. 6, pp. 922–936, 2015.

[85] H. Beltran, D. S. Rickman, K. Park, *et al.*, "Molecular Characterization of Neuroendocrine Prostate Cancer and Identification of New Drug Targets," *Cancer Discov.*, vol. 1, no. 6, pp. 487–495, 2011.

[86] P. Stijnen, B. Ramos-Molina, S. O'Rahilly, and J. W. M. Creemers, "PCSK1 Mutations and Human Endocrinopathies: From Obesity to Gastrointestinal Disorders," *Endocr. Rev.*, vol. 37, no. 4, pp. 347–371, 2016.

[87] C. Zhang, J. Qian, Y. Wu, *et al.*, "Identification of Novel Diagnosis Biomarkers for Therapy-Related Neuroendocrine Prostate Cancer," *Pathol. Oncol. Res.*, vol. 27, 2021.

[88] J. K. Lee, N. J. Bangayan, T. Chai, *et al.*, "Systemic surfaceome profiling identifies target antigens for immune-based therapy in subtypes of advanced prostate cancer.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 19, E4473–E4482, 2018.

[89] D. C. DeLucia, T. M. Cardillo, L. S. Ang, *et al.*, "Regulation of CEACAM5 and therapeutic efficacy of an anti-CEACAM5-SN38 antibody-drug conjugate in neuroendocrine prostate cancer," *Clin. Cancer Res.*, clincanres.3396. bibrangedash 2020, 2020.

[90] K. Pantel and C. Alix-Panabières, "Circulating tumour cells in cancer patients: challenges and perspectives," *Trends Mol. Med.*, vol. 16, no. 9, pp. 398–406, 2010.

[91] C. Alix-Panabières and K. Pantel, "Liquid biopsy: From discovery to clinical application," *Cancer Discov.*, vol. 11, no. 4, pp. 858–873, 2021.

[92] K. Pantel and C. Alix-Panabières, "Liquid biopsy and minimal residual disease latest advances and implications for cure," *Nat. Rev. Clin. Oncol.*, vol. 16, no. 7, pp. 409–424, 2019.

[93] M. Elazezy and S. A. Joosse, "Techniques of using circulating tumor DNA as a liquid biopsy component in cancer management," *Comput. Struct. Biotechnol. J.*, vol. 16, pp. 370–378, 2018.

[94] J. Park, N.-E. Kim, H. Yoon, *et al.*, "Fecal Microbiota and Gut Microbe-Derived Extracellular Vesicles in Colorectal Cancer," *Front. Oncol.*, vol. 11, 2021.

[95] J. Zhu, V. Giannakeas, S. A. Narod, and M. R. Akbari, "Emerging applications of tumour-educated platelets in the detection and prognostication of ovarian cancer," *Protein Cell*, 2023.

[96] Z. Ao, S. H. Shah, L. M. Machlin, *et al.*, "Identification of Cancer-Associated Fibroblasts in Circulating Blood from Patients with Metastatic Breast Cancer," *Cancer Res.*, vol. 75, no. 22, pp. 4681–4687, 2015.

[97] K. Pantel and M. R. Speicher, "The biology of circulating tumor cells," *Oncogene*, vol. 35, no. 10, pp. 1216–1224, 2016.

[98] S. Meng, D. Tripathy, E. P. Frenkel, *et al.*, "Circulating Tumor Cells in Patients with Breast Cancer Dormancy," *Clin. Cancer Res.*, vol. 10, no. 24, pp. 8152–8162, 2004.

[99] I. Baccelli, A. Schneeweiss, S. Riethdorf, *et al.*, "Identification of a population of blood circulating tumor cells from breast cancer patients that initiates metastasis in a xenograft assay," *Nat. Biotechnol.*, vol. 31, no. 6, pp. 539–544, 2013.

[100] J. G. Lohr, V. A. Adalsteinsson, K. Cibulskis, *et al.*, "Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer," *Nat. Biotechnol.*, vol. 32, no. 5, pp. 479–484, 2014.

[101] J. S. de Bono, H. I. Scher, R. B. Montgomery, *et al.*, "Circulating Tumor Cells Predict Survival Benefit from Treatment in Metastatic Castration-Resistant Prostate Cancer," *Clin. Cancer Res.*, vol. 14, no. 19, pp. 6302–6309, 2008.

[102] S. Riethdorf, H. Fritsche, V. Müller, *et al.*, "Detection of Circulating Tumor Cells in Peripheral Blood of Patients with Metastatic Breast Cancer: A Validation Study of the CellSearch System," *Clin. Cancer Res.*, vol. 13, no. 3, pp. 920–928, 2007.

[103] T. M. Gorges, A. Stein, J. Quidde, *et al.*, "Improved Detection of Circulating Tumor Cells in Metastatic Colorectal Cancer by the Combination of the CellSearchő System and the AdnaTestő," *PLoS One*, vol. 11, no. 5, pp. 1–13, 2016.

[104] H. I. Scher, G. Heller, A. Molina, *et al.*, "Circulating tumor cell biomarker panel as an individual-level surrogate for survival in metastatic castration-resistant prostate cancer," *J. Clin. Oncol.*, vol. 33, no. 12, pp. 1348–1355, 2015.

[105] M. Bredemeier, P. Edimiris, P. Mach, *et al.*, "Gene Expression Signatures in Circulating Tumor Cells Correlate with Response to Therapy in Metastatic Breast Cancer.," *Clin. Chem.*, vol. 63, no. 10, pp. 1585–1593, 2017.

[106] G. E. Hvichia, Z. Parveen, C. Wagner, *et al.*, "A novel microfluidic platform for size and deformability based separation and the subsequent molecular characterization of viable circulating tumor cells.," *Int. J. cancer*, vol. 138, no. 12, pp. 2894–904, 2016.

[107] A. Babayan, M. Alawi, M. Gormley, *et al.*, "Comparative study of whole genome amplification and next generation sequencing performance of single cancer cells.," *Oncotarget*, vol. 8, no. 34, pp. 56 066–56 080, 2017.

[108] A. Kulasinghe, Y. Lim, J. Kapeleris, M. Warkiani, K. O'Byrne, and C. Punyadeera, "The Use of Three-Dimensional DNA Fluorescent In Situ Hy-

bridization (3D DNA FISH) for the Detection of Anaplastic Lymphoma Kinase (ALK) in Non-Small Cell Lung Cancer (NSCLC) Circulating Tumor Cells," *Cells*, vol. 9, no. 6, p. 1465, 2020.

[109] A. El-Heliebi, C. Hille, N. Laxman, *et al.*, "In Situ Detection and Quantification of AR-V7, AR-FL, PSA, and KRAS Point Mutations in Circulating Tumor Cells.," *Clin. Chem.*, vol. 64, no. 3, pp. 536–546, 2018.

[110] C. Gasch, P. N. Plummer, L. Jovanovic, *et al.*, "Heterogeneity of miR-10b expression in circulating tumor cells," *Sci. Rep.*, vol. 5, no. 1, p. 15 980, 2015.

[111] Y.-F. Sun, L. Wu, S.-P. Liu, *et al.*, "Dissecting spatial heterogeneity and the immune-evasion mechanism of CTCs by single-cell RNA-seq in hepatocellular carcinoma," *Nat. Commun.*, vol. 12, no. 1, p. 4091, 2021.

[112] W. J. Allard, J. Matera, M. C. Miller, *et al.*, "Tumor Cells Circulate in the Peripheral Blood of All Major Carcinomas but not in Healthy Subjects or Patients With Nonmalignant Diseases," *Clin. Cancer Res.*, vol. 10, no. 20, pp. 6897–6904, 2004.

[113] C. Bettegowda, M. Sausen, R. J. Leary, *et al.*, "Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies," *Sci. Transl. Med.*, vol. 6, no. 224, 224ra24–224ra24, 2014.

[114] A. W. Wyatt, M. Annala, R. Aggarwal, *et al.*, "Concordance of Circulating Tumor DNA and Matched Metastatic Tissue Biopsy in Prostate Cancer," *JNCI J. Natl. Cancer Inst.*, vol. 109, no. 12, 2017.

[115] H. I. Scher, X. Jia, J. S. de Bono, *et al.*, "Circulating tumour cells as prognostic markers in progressive, castration-resistant prostate cancer: a reanalysis of IMMC38 trial data," *Lancet Oncol.*, vol. 10, no. 3, pp. 233–239, 2009.

[116] C. Wang, Z. Zhang, W. Chong, *et al.*, "Improved Prognostic Stratification Using Circulating Tumor Cell Clusters in Patients with Metastatic Castration-Resistant Prostate Cancer," *Cancers (Basel).*, vol. 13, no. 2, p. 268, 2021.

[117] N. Mehra, D. Dolling, S. Sumanasuriya, *et al.*, "Plasma Cell-free DNA Concentration and Outcomes from Taxane Therapy in Metastatic Castration-resistant Prostate Cancer from Two Phase III Trials (FIRSTANA and PROSELICA)," *Eur. Urol.*, vol. 74, no. 3, pp. 283–291, 2018.

[118] S. H. Tolmeijer, E. Boerrigter, T. Sumiyoshi, *et al.*, "Early on-treatment changes in circulating tumor DNA fraction and response to enzalutamide or abiraterone in metastatic castration-resistant prostate cancer.," *Clin. Cancer Res.*, 2023.

[119] E. S. Antonarakis, C. Lu, B. Luber, *et al.*, "Clinical Significance of Androgen Receptor Splice Variant-7 mRNA Detection in Circulating Tumor Cells of Men With Metastatic Castration-Resistant Prostate Cancer Treated With First- and Second-Line Abiraterone and Enzalutamide," *J. Clin. Oncol.*, vol. 35, no. 19, pp. 2149–2156, 2017.

[120] Hille, Gorges, Riethdorf, *et al.*, "Detection of Androgen Receptor Variant 7 (ARV7) mRNA Levels in EpCAM-Enriched CTC Fractions for Monitoring Response to Androgen Targeting Therapies in Prostate Cancer," *Cells*, vol. 8, no. 9, p. 1067, 2019.

[121] M. Del Re, E. Biasco, S. Crucitta, *et al.*, "The Detection of Androgen Receptor Splice Variant 7 in Plasma-derived Exosomal RNA Strongly Predicts Resistance to Hormonal Therapy in Metastatic Prostate Cancer Patients," *Eur. Urol.*, vol. 71, no. 4, pp. 680–687, 2017.

[122] V. Conteduca, D. Wetterskog, M. T. A. Sharabiani, *et al.*, "Androgen receptor gene status in plasma DNA associates with worse outcome on enzalutamide or abiraterone for castration-resistant prostate cancer: a multi-institution correlative biomarker study.," *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, vol. 28, no. 7, pp. 1508–1516, 2017.

[123] A. W. Wyatt, A. A. Azad, S. V. Volik, *et al.*, "Genomic Alterations in Cell-Free DNA and Enzalutamide Resistance in Castration-Resistant Prostate Cancer," *JAMA Oncol.*, vol. 2, no. 12, p. 1598, 2016.

[124] A. H. Zedan, P. J. S. Osther, J. Assenholt, J. S. Madsen, and T. F. Hansen, "Circulating miR-141 and miR-375 are associated with treatment outcome in metastatic castration resistant prostate cancer.," *Sci. Rep.*, vol. 10, no. 1, p. 227, 2020.

[125] Y.-H. Wang, J. Ji, B.-C. Wang, *et al.*, "Tumor-Derived Exosomal Long Noncoding RNAs as Promising Diagnostic Biomarkers for Prostate Cancer.," *Cell. Physiol. Biochem.*, vol. 46, no. 2, pp. 532–545, 2018.

[126] J. S. Chung, Y. Wang, J. Henderson, *et al.*, "Circulating Tumor CellBased Molecular Classifier for Predicting Resistance to Abiraterone and Enzalutamide in Metastatic Castration-Resistant Prostate Cancer," *Neoplasia (United States)*, vol. 21, no. 8, pp. 802–809, 2019.

[127] C. R. Davies, T. Guo, E. Burke, *et al.*, "The potential of using circulating tumour cells and their gene expression to predict docetaxel response in metastatic prostate cancer," *Front. Oncol.*, vol. 12, 2023.

[128] A. Jayaram, A. Wingate, D. Wetterskog, *et al.*, "Plasma tumor gene conversions after one cycle abiraterone acetate for metastatic castration-resistant prostate cancer: a biomarker analysis of a multicenter international trial," *Ann. Oncol.*, vol. 32, no. 6, pp. 726–735, 2021.

[129] A. Wu, P. Cremaschi, D. Wetterskog, *et al.*, "Genome-wide plasma DNA methylation features of metastatic prostate cancer," *J. Clin. Invest.*, vol. 130, no. 4, pp. 1991–2000, 2020.

[130] J. von Hardenberg, T. S. Worst, N. Westhoff, *et al.*, "Cell-Free DNA and Neuromediators in Detecting Aggressive Variant Prostate Cancer.," *Oncol. Res. Treat.*, vol. 41, no. 10, pp. 627–633, 2018.

[131] M. M. Heck, M. A. Thaler, S. C. Schmid, *et al.*, "Chromogranin A and neurone-specific enolase serum levels as predictors of treatment outcome in patients with metastatic castration-resistant prostate cancer undergoing abiraterone therapy," *BJU Int.*, vol. 119, no. 1, pp. 30–37, 2017.

[132] S. Derderian, Q. Vesval, M. D. Wissing, *et al.*, "Liquid biopsybased targeted gene screening highlights tumor cell subtypes in patients with advanced prostate cancer," *Clin. Transl. Sci.*, vol. 15, no. 11, pp. 2597–2612, 2022.

[133] S. K. Pal, M. He, L. Chen, *et al.*, "Synaptophysin expression on circulating tumor cells in patients with castration resistant prostate cancer under-

going treatment with abiraterone acetate or enzalutamide," *Urol. Oncol. Semin. Orig. Investig.*, vol. 36, no. 4, 162.e1–162.e6, 2018.

[134] L. Puca, K. Gavyert, V. Sailer, *et al.*, "Delta-like protein 3 expression and therapeutic targeting in neuroendocrine prostate cancer," *Sci. Transl. Med.*, vol. 11, no. 484, eaav0891, 2019.

[135] H. Beltran, A. Jendrisak, M. Landers, *et al.*, "The Initial Detection and Partial Characterization of Circulating Tumor Cells in Neuroendocrine Prostate Cancer," *Clin. Cancer Res.*, vol. 22, no. 6, pp. 1510–1519, 2016.

[136] V. Conteduca, S.-Y. Ku, L. Fernandez, *et al.*, "Circulating tumor cell heterogeneity in neuroendocrine prostate cancer by single cell copy number analysis," *npj Precis. Oncol.*, vol. 5, no. 1, p. 76, 2021.

[137] X. Dong, T. Zheng, M. Zhang, *et al.*, "Circulating Cell-Free DNA-Based Detection of Tumor Suppressor Gene Copy Number Loss and Its Clinical Implication in Metastatic Prostate Cancer," *Front. Oncol.*, vol. 11, 2021.

[138] J. M. Sperger, H. Emamekhoo, R. R. McKay, *et al.*, "Prospective Evaluation of Clinical Outcomes Using a Multiplex Liquid Biopsy Targeting Diverse Resistance Mechanisms in Metastatic Prostate Cancer," *J. Clin. Oncol.*, vol. 39, no. 26, pp. 2926–2937, 2021.

[139] R. R. McKay, L. Kwak, J. P. Crowdis, *et al.*, "Phase II Multicenter Study of Enzalutamide in Metastatic Castration-Resistant Prostate Cancer to Identify Mechanisms Driving Resistance.," *Clin. Cancer Res.*, vol. 27, no. 13, pp. 3610–3619, 2021.

[140] M. Kardoust Parizi, T. Iwata, S. Kimura, *et al.*, "Focal Neuroendocrine Differentiation of Conventional Prostate Adenocarcinoma as a Prognostic Factor after Radical Prostatectomy: A Systematic Review and Meta-Analysis," *Int. J. Mol. Sci.*, vol. 20, no. 6, p. 1374, 2019.

[141] J. S. Horoszewicz, S. S. Leong, E. Kawinski, *et al.*, "LNCaP model of human prostatic carcinoma.," *Cancer Res.*, vol. 43, no. 4, pp. 1809–18, 1983.

[142] S. Korenchuk, J. E. Lehr, L. MClean, *et al.*, "VCaP, a cell-based model system of human prostate cancer.," *In Vivo*, vol. 15, no. 2, pp. 163–8, 2001.

[143] S. L. Lai, H. Brauch, T. Knutsen, *et al.*, "Molecular genetic characterization of neuroendocrine lung cancer cell lines.," *Anticancer Res.*, vol. 15, no. 2, pp. 225–32, 1995.

[144] K. R. Stone, D. D. Mickey, H. Wunderli, G. H. Mickey, and D. F. Paulson, "Isolation of a human prostate carcinoma cell line (DU 145)," *Int. J. Cancer*, vol. 21, no. 3, pp. 274–281, 1978.

[145] M. E. Kaighn, K. S. Narayan, Y. Ohnuki, J. F. Lechner, and L. W. Jones, "Establishment and characterization of a human prostatic carcinoma cell line (PC-3).," *Invest. Urol.*, vol. 17, no. 1, pp. 16–23, 1979.

[146] J. K. Lee, J. W. Phillips, B. A. Smith, *et al.*, "N-Myc Drives Neuroendocrine Prostate Cancer Initiated from Human Prostate Epithelial Cells," *Cancer Cell*, vol. 29, no. 4, pp. 536–547, 2016.

[147] D. N. Carney, A. F. Gazdar, G. Bepler, *et al.*, "Establishment and identification of small cell lung cancer cell lines having classic and variant features.," *Cancer Res.*, vol. 45, no. 6, pp. 2913–23, 1985.

[148] L. Y. Ahn, G. C. Coatti, J. Liu, E. Gumus, A. E. Schaffer, and H. C. Miranda, "An epilepsy-associated ACTL6B variant captures neuronal hyperexcitability in a human induced pluripotent stem cell model.," *J. Neurosci. Res.*, vol. 99, no. 1, pp. 110–123, 2021.

[149] Z. Ling, X. Long, Y. Wu, J. Li, and M. Feng, "LMO3 promotes proliferation and metastasis of papillary thyroid carcinoma cells by regulating LIMK1-mediated cofilin and the $\beta$-catenin pathway.," *Open Med. (Warsaw, Poland)*, vol. 17, no. 1, pp. 453–462, 2022.

[150] J. Huang, E. H. Ji, X. Zhao, *et al.*, "Sox11 promotes head and neck cancer progression via the regulation of SDCCAG8.," *J. Exp. Clin. Cancer Res.*, vol. 38, no. 1, p. 138, 2019.

[151] C. L. Andersen, J. L. Jensen, and T. F. Ørntoft, "Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets.," *Cancer Res.*, vol. 64, no. 15, pp. 5245–50, 2004.

[152] S. A. Joosse and K. Pantel, "Biologic challenges in the detection of circulating tumor cells.," *Cancer Res.*, vol. 73, no. 1, pp. 8–11, 2013.

[153] K. Besler, A. Wglarz, L. Keller, *et al.*, "Expression Patterns and Corepressor Function of Retinoic Acid-induced 2 in Prostate Cancer.," *Clin. Chem.*, vol. 68, no. 7, pp. 973–983, 2022.

[154] T. S. C. Ng, X. Gao, K. Salari, D. V. Zlatev, P. Heidari, and S. C. Kamran, "Incorporating PSMA-Targeting Theranostics Into Personalized Prostate Cancer Treatment: a Multidisciplinary Perspective.," *Front. Oncol.*, vol. 11, p. 722 277, 2021.

[155] M. Uhlén, L. Fagerberg, B. M. Hallström, *et al.*, "Tissue-based map of the human proteome.," *Science*, vol. 347, no. 6220, p. 1 260 419, 2015.

[156] H. Tsai, C. L. Morais, M. Alshalalfa, *et al.*, "Cyclin D1 Loss Distinguishes Prostatic Small-Cell Carcinoma from Most Prostatic Adenocarcinomas," *Clin. Cancer Res.*, vol. 21, no. 24, pp. 5619–5629, 2015.

[157] S.-R. Lin, Y.-C. Wen, H.-L. Yeh, *et al.*, "EGFR-upregulated LIFR promotes SUCLG2-dependent castration resistance and neuroendocrine differentiation of prostate cancer.," *Oncogene*, vol. 39, no. 44, pp. 6757–6775, 2020.

[158] J. Balzeau, M. R. Menezes, S. Cao, and J. P. Hagan, "The LIN28/let-7 Pathway in Cancer," *Front. Genet.*, vol. 8, 2017.

[159] M. Cancel, C. Castellier, C. Debiais-Delpech, *et al.*, "Specificities of small cell neuroendocrine prostate cancer: Adverse prognostic value of TTF1 expression.," *Urol. Oncol.*, vol. 39, no. 1, 74.e17–74.e23, 2021.

[160] B. G. Sánchez, A. Bort, D. Vara-Ciruelos, and I. Díaz-Laviada, "Androgen Deprivation Induces Reprogramming of Prostate Cancer Cells to Stem-Like Cells," *Cells*, vol. 9, no. 6, pp. 1–19, 2020.

[161] A. Pugongchai, A. Bychkov, and P. Sampatanukul, "Promoter hypermethylation of SOX11 correlates with adverse clinicopathological features of human prostate cancer," *Int. J. Exp. Pathol.*, vol. 98, no. 6, pp. 341–346, 2017.

[162] Y. Li, Q. Zhang, J. Lovnicki, *et al.*, "SRRM4 gene expression correlates with neuroendocrine prostate cancer," *Prostate*, vol. 79, no. 1, pp. 96–104, 2019.
[163] S. Riethdorf, L. O'Flaherty, C. Hille, and K. Pantel, "Clinical applications of the CellSearch platform in cancer patients.," *Adv. Drug Deliv. Rev.*, vol. 125, pp. 102–121, 2018.
[164] G. Claps, S. Faouzi, V. Quidville, *et al.*, "The multiple roles of LDH in cancer," *Nat. Rev. Clin. Oncol.*, vol. 19, no. 12, pp. 749–762, 2022.
[165] S. Schneegans, L. Lück, K. Besler, *et al.*, "Preanalytical factors affecting the establishment of a single tube assay for multiparameter liquid biopsy detection in melanoma patients," *Mol. Oncol.*, vol. 14, no. 5, pp. 1001–1015, 2020.
[166] D. Hanahan, "Hallmarks of Cancer: New Dimensions," *Cancer Discov.*, vol. 12, no. 1, pp. 31–46, 2022.
[167] C. Vogel and E. M. Marcotte, "Insights into the regulation of protein abundance from proteomic and transcriptomic analyses.," *Nat. Rev. Genet.*, vol. 13, no. 4, pp. 227–32, 2012.
[168] H. Laurell, J. S. Iacovoni, A. Abot, *et al.*, "Correction of RT-qPCR data for genomic DNA-derived signals with ValidPrime.," *Nucleic Acids Res.*, vol. 40, no. 7, e51, 2012.
[169] L. Keller, S. Werner, and K. Pantel, "Biology and clinical relevance of EpCAM," *Cell Stress*, vol. 3, no. 6, pp. 165–180, 2019.
[170] A. Satelli, I. Batth, Z. Brownlee, *et al.*, "EMT circulating tumor cells detected by cell-surface vimentin are associated with prostate cancer progression," *Oncotarget*, vol. 8, no. 30, pp. 49 329–49 337, 2017.
[171] T. Eguchi, C. Sogawa, Y. Okusha, *et al.*, "Organoids with cancer stem cell-like properties secrete exosomes and HSP90 in a 3D nanoenvironment," *PLoS One*, vol. 13, no. 2, G. Papaccio, Ed., e0191109, 2018.
[172] S. De Wit, M. Manicone, E. Rossi, *et al.*, "EpCAMhigh and EpCAMlow circulating tumor cells in metastatic prostate and breast cancer patients," *Oncotarget*, vol. 9, no. 86, pp. 35 705–35 716, 2018.
[173] M. S. Khan, T. Tsigani, M. Rashid, *et al.*, "Circulating Tumor Cells and EpCAM Expression in Neuroendocrine Tumors," *Clin. Cancer Res.*, vol. 17, no. 2, pp. 337–345, 2011.
[174] L. Puca, P. J. Vlachostergios, and H. Beltran, "Neuroendocrine Differentiation in Prostate Cancer: Emerging Biology, Models, and Therapies.," *Cold Spring Harb. Perspect. Med.*, vol. 9, no. 2, 2019.
[175] S. Stickler, B. Rath, M. Hochmair, C. Lang, L. Weigl, and G. Hamilton, "Changes of protein expression during tumorosphere formation of small cell lung cancer circulating tumor cells.," *Oncol. Res.*, vol. 31, no. 1, pp. 13–22, 2023.
[176] J. Huang, D. Liu, J. Li, J. Xu, S. Dong, and H. Zhang, "A 12-gene panel in estimating hormone-treatment responses of castration-resistant prostate cancer patients generated using a combined analysis of bulk and single-cell sequencing data.," *Ann. Med.*, vol. 55, no. 2, p. 2 260 387, 2023.

[177] G. Di Lorenzo, G. Tortora, F. P. D'Armiento, *et al.*, "Expression of epidermal growth factor receptor correlates with disease relapse and progression to androgen-independence in human prostate cancer.," *Clin. Cancer Res.*, vol. 8, no. 11, pp. 3438–44, 2002.

[178] H. Ota, H. Sato, S. Mizumoto, *et al.*, "Switching mechanism from AR to EGFR signaling via 3-O-sulfated heparan sulfate in castration-resistant prostate cancer.," *Sci. Rep.*, vol. 13, no. 1, p. 11 618, 2023.

[179] K. C. Day, G. Lorenzatti Hiles, M. Kozminsky, *et al.*, "HER2 and EGFR Overexpression Support Metastatic Progression of Prostate Cancer to Bone.," *Cancer Res.*, vol. 77, no. 1, pp. 74–85, 2017.

[180] A. Rossini, M. Giussani, F. Ripamonti, *et al.*, "Combined targeting of EGFR and HER2 against prostate cancer stem cells.," *Cancer Biol. Ther.*, vol. 21, no. 5, pp. 463–475, 2020.

[181] J. A. Korsen, T. M. Kalidindi, S. Khitrov, *et al.*, "Molecular Imaging of Neuroendocrine Prostate Cancer by Targeting Delta-Like Ligand 3.," *J. Nucl. Med.*, vol. 63, no. 9, pp. 1401–1407, 2022.

[182] D. C. Danila, A. Samoila, C. Patel, *et al.*, "Clinical Validity of Detecting Circulating Tumor Cells by AdnaTest Assay Compared With Direct Detection of Tumor mRNA in Stabilized Whole Blood, as a Biomarker Predicting Overall Survival for Metastatic Castration-Resistant Prostate Cancer Patients," *Cancer J.*, vol. 22, no. 5, pp. 315–320, 2016.

[183] M. Marín-Aguilera, N. Jiménez, Ò. Reig, *et al.*, "Androgen Receptor and Its Splicing Variant 7 Expression in Peripheral Blood Mononuclear Cells and in Circulating Tumor Cells in Metastatic Castration-Resistant Prostate Cancer," *Cells*, vol. 9, no. 1, p. 203, 2020.

[184] J.-F. Chen, H. Ho, J. Lichterman, *et al.*, "Subclassification of prostate cancer circulating tumor cells by nuclear size reveals very small nuclear circulating tumor cells in patients with visceral metastases," *Cancer*, vol. 121, no. 18, pp. 3240–3251, 2015.

[185] T. Gu, J. Li, T. Chen, Q. Zhu, and J. Ding, "Circulating tumor cell quantification during abiraterone plus prednisone therapy may estimate survival in metastatic castration-resistant prostate cancer patients," *Int. Urol. Nephrol.*, vol. 55, no. 4, pp. 883–892, 2023.

[186] M. Ladurner, M. Wieser, A. Eigentler, *et al.*, "Validation of Cell-Free RNA and Circulating Tumor Cells for Molecular Marker Analysis in Metastatic Prostate Cancer," *Biomedicines*, vol. 9, no. 8, p. 1004, 2021.

[187] T. Pereira-Veiga, M. González-Conde, L. León-Mateos, *et al.*, "Longitudinal CTCs gene expression analysis on metastatic castration-resistant prostate cancer patients treated with docetaxel reveals new potential prognosis markers," *Clin. Exp. Metastasis*, 2021.

[188] J. T. Dankert, M. Wiesehöfer, E. D. Czyrnik, B. B. Singer, N. von Ostau, and G. Wennemuth, "The deregulation of miR-17/CCND1 axis during neuroendocrine transdifferentiation of LNCaP prostate cancer cells," *PLoS One*, vol. 13, no. 7, S. Papa, Ed., e0200472, 2018.

[189] W.-D. Liu, H.-W. Wang, M. Muguira, M. B. Breslin, and M. S. Lan, "INSM1 functions as a transcriptional repressor of the neuroD/beta2 gene through the recruitment of cyclin D1 and histone deacetylases.," *Biochem. J.*, vol. 397, no. 1, pp. 169–77, 2006.

[190] M. Zavridou, S. Smilkou, V. Tserpeli, *et al.*, "Development and Analytical Validation of a 6-Plex Reverse Transcription Droplet Digital PCR Assay for the Absolute Quantification of Prostate Cancer Biomarkers in Circulating Tumor Cells of Patients with Metastatic Castration-Resistant Prostate Cancer," *Clin. Chem.*, vol. 68, no. 10, pp. 1323–1335, 2022.

[191] H. Guo, J. A. Vuille, B. S. Wittner, *et al.*, "DNA hypomethylation silences anti-tumor immune genes in early prostate cancer and CTCs," *Cell*, vol. 186, no. 13, 2765–2782.e28, 2023.

[192] H. S. Rugo, A. Bardia, F. Marmé, *et al.*, "Sacituzumab Govitecan in Hormone ReceptorPositive/Human Epidermal Growth Factor Receptor 2Negative Metastatic Breast Cancer," *J. Clin. Oncol.*, vol. 40, no. 29, pp. 3365–3376, 2022.

[193] N. Goradia, S. Werner, E. Mullapudi, *et al.*, "Corepressor inactivation through tandem linear motif-induced polymerization," *bioRxiv*, p. 2023.06.16.545227, 2023.

[194] R. A. Patel, E. Sayar, I. Coleman, *et al.*, "Characterization of HOXB13 expression patterns in localized and metastatic castrationresistant prostate cancer," *J. Pathol.*, 2023.

[195] S. G. Zhao, J. M. Sperger, J. L. Schehr, *et al.*, "A clinical-grade liquid biomarker detects neuroendocrine differentiation in prostate cancer," *J. Clin. Invest.*, vol. 132, no. 21, 2022.

[196] M. J. IJzerman, J. de Boer, A. Azad, *et al.*, "Towards Routine Implementation of Liquid Biopsies in Cancer Management: It Is Always Too Early, until Suddenly It Is Too Late," *Diagnostics*, vol. 11, no. 1, p. 103, 2021.

[197] S. Hertlein, E. Haenssler, S. Hey, M. Otte, S. Hauch, and M. Schlumpberger, "Processing of large-volume liquid biopsies: cfDNA isolation and CTC enrichment with subsequent mRNA isolation," Qiagen, Hilden, Tech. Rep., 2022.

[198] A. H. Skogholt, E. Ryeng, S. E. Erlandsen, F. Skorpen, S. A. Schønberg, and P. Sætrom, "Gene expression differences between PAXgene and Tempus blood RNA tubes are highly reproducible between independent samples and biobanks.," *BMC Res. Notes*, vol. 10, no. 1, p. 136, 2017.

[199] K. Degeling, A. Pereira-Salgado, N. M. Corcoran, P. C. Boutros, P. Kuhn, and M. J. IJzerman, "Health Economic Evidence for Liquid- and Tissue-based Molecular Tests that Inform Decisions on Prostate Biopsies and Treatment of Localised Prostate Cancer: A Systematic Review.," *Eur. Urol. open Sci.*, vol. 27, pp. 77–87, 2021.

[200] A. Hollebecque, S. Salvagni, R. Plummer, *et al.*, "Clinical activity of CC-90011, an oral, potent, and reversible LSD1 inhibitor, in advanced malignancies.," *Cancer*, vol. 128, no. 17, pp. 3185–3195, 2022.

[201] E. Dardenne, H. Beltran, M. Benelli, *et al.*, "N-Myc Induces an EZH2-Mediated Transcriptional Program Driving Neuroendocrine Prostate Cancer," *Cancer Cell*, vol. 30, no. 4, pp. 563–577, 2016.

[202] H. Beltran, C. Oromendia, D. C. Danila, *et al.*, "A Phase II Trial of the Aurora Kinase A Inhibitor Alisertib for Patients with Castration-resistant and Neuroendocrine Prostate Cancer: Efficacy and Biomarkers.," *Clin. Cancer Res.*, vol. 25, no. 1, pp. 43–51, 2019.

[203] L. Li, S. Karanika, G. Yang, *et al.*, "Androgen receptor inhibitor-induced "BRCAness" and PARP inhibition are synthetically lethal for castration-resistant prostate cancer.," *Sci. Signal.*, vol. 10, no. 480, 2017.

[204] H. R. VanDeusen, J. R. Ramroop, K. L. Morel, *et al.*, "Targeting RET Kinase in Neuroendocrine Prostate Cancer," *Mol. Cancer Res.*, vol. 18, no. 8, pp. 1176–1188, 2020.

[205] K. Unno, Z. R. Chalmers, S. Pamarthy, *et al.*, "Activated ALK Cooperates with N-Myc via Wnt/$\beta$-Catenin Signaling to Induce Neuroendocrine Prostate Cancer.," *Cancer Res.*, vol. 81, no. 8, pp. 2157–2170, 2021.

[206] A. S. Mansfield, D. S. Hong, C. L. Hann, *et al.*, "A phase I/II study of rovalpituzumab tesirine in delta-like 3-expressing advanced solid tumors.," *NPJ Precis. Oncol.*, vol. 5, no. 1, p. 74, 2021.

[207] M. L. Johnson, G. K. Dy, H. Mamdani, *et al.*, "Interim results of an ongoing phase 1/2a study of HPN328, a tri-specific, half-life extended, DLL3-targeting, T-cell engager, in patients with small cell lung cancer and other neuroendocrine cancers.," *J. Clin. Oncol.*, vol. 40, no. 16_suppl, pp. 8566–8566, 2022.

[208] L. Paz-Ares, S. Champiat, W. V. Lai, *et al.*, "Tarlatamab, a First-in-Class DLL3-Targeted Bispecific T-Cell Engager, in Recurrent Small-Cell Lung Cancer: An Open-Label, Phase I Study.," *J. Clin. Oncol.*, vol. 41, no. 16, pp. 2893–2903, 2023.

[209] A. Bardia, S. A. Hurvitz, S. M. Tolaney, *et al.*, "Sacituzumab Govitecan in Metastatic Triple-Negative Breast Cancer," *N. Engl. J. Med.*, vol. 384, no. 16, pp. 1529–1541, 2021.

[210] A. Bardia, S. Tolaney, K. Punie, *et al.*, "Biomarker analyses in the phase III ASCENT study of sacituzumab govitecan versus chemotherapy in patients with metastatic triple-negative breast cancer," *Ann. Oncol.*, vol. 32, no. 9, pp. 1148–1156, 2021.

[211] N. Matsubara, J. de Bono, D. Olmos, *et al.*, "Olaparib Efficacy in Patients with Metastatic Castration-resistant Prostate Cancer and BRCA1, BRCA2, or ATM Alterations Identified by Testing Circulating Tumor DNA.," *Clin. Cancer Res.*, vol. 29, no. 1, pp. 92–99, 2023.

# Acknowledgements

I am profoundly grateful to all individuals whose support and contributions have been instrumental in the completion of this dissertation.

First, I would like to thank Prof. Dr. Gunhild von Amsberg for the opportunity to do my doctorate in such a fascinating research field, for her mentorship and for her inspiring motivation to advance patient care. I would also like to thank Prof. Dr. Klaus Pantel for granting me the opportunity to do my doctorate at the Institute of Tumor Biology in the intriguing field of liquid biopsy. Their expertise and insightful feedback have been pivotal in shaping this project.

Especially, I am grateful for the support of my supervisor and group leader Dr. Stefan Werner. I appreciate his guidance, encouragement and critical feedback throughout this research project and I would like to thank him for the opportunity to participate in an international conference.

I extend my thanks to the (former) members of my thesis committee Prof. Dr. Tobias Lange and Dr. Wael Mansour for their time and constructive criticism.

I would also like to thank the Graduate School "InTechCanDiTh" represented by Dr. Malte Kriegs and Dr. Natascha Kömm. I appreciate the diverse, educational course work that opened up my perspective on cancer research and patient care.

I would like to thank all members of the AG Werner including Bettina Steinbach, Sarah Greimeier, Stefanie Paigin, Simona Parretta and Dr. Marina Mutas for their helping hands and fruitful discussions.

Additionally, I am grateful for the help of Tina Rohlfing from my "second group", the AG Amsberg, for collection of clinical data. I also thank all other group members for their support, and especially Sergey for giving me the opportunity to participate in additional projects.

I would like to thank Prof. Dr. Sabine Riethdorf for her time and expertise spend

on the analysis of CellSearch data and to Antje, Conny and Malgorzata for sample preparation. Moreover, I am grateful for the support I got from all other ITB members including Haiju, Simon, Jolanthe, Sandra and Leonie for answering all my minor and major questions.

I am thankful for all the valuable little chats and countless lunch breaks with Annabelle, Jana, Vanessa, Kira, Arne, Jeannine and Kim-Lea who have been cheering me up whenever needed.

I would also like to appreciate the support from my friends and family who granted me the opportunity to always give my best.

Lastly, I thank all patients who selflessly decided to donate their blood to research.

# Publications

## First Authorship

**Merkens L**, Sailer V, Lessel D, Janzen E, Greimeier S, Kirfel J, Perner S, Pantel K, Werner S*, von Amsberg G*. Aggressive variants of prostate cancer: underlying mechanisms of neuroendocrine transdifferentiation. *J Exp Clin Cancer Res*. 2022 Feb 2;41(1):46. doi: 10.1186/s13046-022-02255-y

## Co-Authorship

von Amsberg G, Alsdorf W, Karagiannis P, Coym A, Kaune M, Werner S, Graefen M, Bokemeyer C, **Merkens L***, Dyshlovoy SA*. Immunotherapy in Advanced Prostate Cancer-Light at the End of the Tunnel? *Int J Mol Sci*. 2022 Feb 25;23(5):2569. doi: 10.3390/ijms23052569

Goradia N*, Werner S*, Mullapudi E, Greimeier S, **Merkens L**, Lang A, Mertens H, Wglarz A, Sander S, Chojnowski G, Wikman H, Ohlenschläger O, von Amsberg G, Pantel K*, Wilmanns M*. Corepressor inactivation through tandem linear motif-induced polymerization. *bioRxiv* 2023. 10.1101/2023.06.16.545227

## Conference Abstract

**Merkens L**, Greimeier S, Riethdorf S, Pantel K*, von Amsberg G*, Werner S*. Liquid biopsy approaches to determine tumor cell heterogeneity in advanced prostate cancer. American Association for Cancer Research Annual Meeting 2023, Orlando.

**Merkens L**, Greimeier S, Riethdorf S, Pantel K*, von Amsberg G*, Werner S*. Analysis of circulating tumor cells displays tumor heterogeneity in advanced prostate cancer. International Symposium on Minimal Residual Cancer 2023, Hamburg.

*these authors contributed equally to the publication

# Curriculum Vitae

entfällt aus datenschutzrechtlichen Gründen

# Eidesstattliche Erklärung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe. Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe. Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

---

Hamburg, Februar 2024 Lina Bergmann