Genetic Diversity and Treatment Resistance in Prostate Cancer Cell Lines

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt

dem Bereich Mathematik und Naturwissenschaften der Technischen Universität Dresden

von

M.Sc. Lukas Donix

Eingereicht am 04.11.2022

Verteidigt am 26.04.2023

Begutachtet durch

Prof. Dr. Christian Dahmann

Prof. Dr. Helge Taubert

Die Dissertation wurde in der Zeit von April 2018 bis Oktober 2022
am Nationalen Centrum für Tumorerkrankungen Dresden (NCT/UCC)
angefertigt.

Table of Contents

List of Figures

List of Tables

Abbreviations

7-AAD	7-Aminoactinomycin D
ABCB1 / C1 / G2	ATP-binding cassette transporter B1 / C1 / G2
ABCT	ATP-binding cassette transporter
AD	Adenocarcinoma
ADT	Androgen deprivation therapy
Akt	Protein kinase B
AR	Androgen receptor
ARSI	Androgen receptor signaling inhibitor
ATP7A & ATP7B	P-type ATPase 7A & B
Bcl	B-cell lymphoma (protein family, e.g., Bcl-2)
BCR	Biochemical relapse
BicaR	Bicalutamide-resistant
BRCA1 & BRCA2	Breast cancer early onset 1 & 2
CAR	Carboplatin
CCLE	Cancer Cell Line Encyclopedia
CD	Cluster of differentiation proteins (e.g., CD56)
CDDP	Cisplatin
CDDPR	Cisplatin-resistant
CDH1	E-cadherin (epithelial-cadherin)
CDH2	N-cadherin (neuronal cadherin)
CDK1	Cyclin dependent kinase 1
CHGA	Chromogranin A
CNV	Copy number variant
CRPC	Castration-resistant prostate cancer
CSPC	Castration-sensitive prostate cancer
CSS	Cancer-specific survival
CT	Cycle threshold
CTR1	Copper transporter-1
DepMap	Cancer Dependency Map Project
DHT	Dihydrotestosterone
DRE	Digital-rectal examination
DSB	(DNA) double strand break
DTX	Docetaxel
DTXR	Docetaxel-resistant
EBRT	External beam radiotherapy
EMAST	Elevated microsatellite alterations at selected tetranucleotide repeats
EMT	Epithelial to mesenchymal transition
EnzaR	Enzalutamide-resistant
FBS	Fetal bovine serum
FA	Fanconi anemia
FANC	Fanconi anemia complementation group proteins (e.g., FANCA)
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GDSC	Genomics of Drug Discovery in Cancer Project
GO: MF/BP/CC	Gene Ontology: molecular function / biological process / cellular component
GoF	Gain of function
GS	Gleason score
GSH	Reduced glutathione
HGVS	Human Genome Variation Society
HKG	House keeping gene
HLA	Human leukocyte antigen
HRR	Homologous recombination repair
IC50	Half maximal inhibitory concentration
ICL	(DNA) Interstrand cross-link
ISUP	International Society of Urological Pathology
kbp	kilo base pairs

kDa	kilo Daltons
KRT	(Cyto)keratin proteins (e.g., KRT8)
LLT	Length (of the) longest transcript
LoF	Loss of function
MAF	Mutation annotation formatted
mCRPC	metastatic CRPC
mCSPC	metastatic CSPC
MDR	Multidrug resistance
MLH	MutL homolog proteins (e.g., MLH1)
MMP	Matrix metalloprotease
MMR	Mismatch repair
MRI	Magnetic resonance imaging
MSH	MutS homolog proteins (e.g., MSH2)
MSI	Microsatellite instability
MT	Microtubule
NE	Neuroendocrine
NEPC	Neuroendocrine prostate cancer
NER	Nucleotide excision repair
NGS	Next generation sequencing
NHEJ	Nonhomologous end joining
NSE / ENO2	Neuron specific enolase / Enolase 2
OS	Overall survival
PARP1	Poly (ADP-ribose) polymerase 1
PBS	Phosphate buffered saline
PCa	Prostate cancer
PCR	Polymerase chain reaction
PE	Phycoerythrin
PI3K	Phosphatidylinositol 3-kinase
PNGase F	Peptide-N-glycosidase F
PSA / KLK3	Prostate-specific antigen / Kallikrein 3
PSMA	Prostate-specific membrane antigen
Pt	Platinum
PTEN	Phosphatase and tensin homologue deleted on chromosome 10
PTX	Paclitaxel
qPCR	Quantitative PCR
RP	Radical prostatectomy
RR	Radio-resistant
RT	Room temperature
SBS	Single base substitution
SD	Standard deviation
SLCO	Solute carrier of organic anions proteins (e.g., SLCO4A1)
SNP / SNV	Single nucleotide polymorphism / variant
STR	Satellite tandem repeat
SV	Structural (genetic) variant
SYP	Synaptophysin
TF	Transcription factor
TMB	Tumor mutational burden
TMD	Transmembrane domain
TNM	Tumor-node-metastases
TNR	Tetranucleotide repeat
TP63	Tumor protein 63
TUBB3	ßIII-tubulin
UTR	Untranslated region
VCF	Variant calling formatted
VIM	Vimentin
ZEB1 & ZEB2	Zinc finger E-box binding homeobox 1 & 2
ZNF	Zinc finger proteins

1. Introduction

1.1 Cancer

There is a compelling correlation between the human development index and the prevalence of cancer. In highly developed geographical regions, cancer is the leading cause of premature mortality, surpassing stroke and coronary heart disease[1]. Although there are many exceptions, the majority of cancers are more common in the old than in the young. Cancer is uncontrolled and malignant cell proliferation and is ubiquitously found in higher animals[2]. In healthy tissues, cell proliferation and programmed cell death (apoptosis) are tightly controlled and balanced. In turn, carcinogenesis may be explained to a considerable extend by the *multistage hypothesis*[3], which states that cancerous cells arise from normal cells through a series of several sequential and heritable genetic aberrations that together overrule cell cycle and apoptotic control checkpoints. Despite the sophisticated DNA repair mechanisms that have evolved and despite the removal of aberrant cells by apoptosis, somatic mutations accumulate throughout life due to external and internal factors. Eventually, a fatal combination of genetic events may trigger the onset of cancerous growth. Almost all tissues of the human body can give rise to cancer. This dissertation is dealing with cancer of the prostate gland (ICD-10 code 61), which is an example for a cancer that primarily affects the elderly.

1.2 Prostate Cancer Etiology and Diagnosis

1.2.1 The Prostate

The prostate is a glandular organ located beneath the bladder that encapsulates the urethra and the ejaculatory ducts that connect the urethra to the seminal vesicles (**Figure 1 a** and **b**). The tissue of the prostate is interspersed with tubular *acini* that merge with the urethra. Together with the seminal vesicles, the prostate produces an alkaline carrier fluid that nourishes and protects the spermatozoa after ejaculation. In young men, the prostate is walnut-sized. Around the age of 40, the gland begins to grow, reaching an average total volume of 38 cm^3 in men aged 75 or older[4]. In addition to normal physiological prostate enlargement, 50% of men aged 51 – 60 and 80% of men aged 81 or older exhibit benign prostatic hyperplasia[5], a condition characterized by increased but benign prostate growth.
The prostate can be divided into zones based on histological features (**Figure 1 b** and **c**). The peripheral zone constitutes the outer, dorsolateral part of the prostate and hosts for 70% of the glandular tissue. The peripheral zone surrounds the central zone, which accounts for 25% of the glandular tissue and encapsulates the ejaculatory ducts. The transitional zone (5% of the glandular tissue) is located ventrally to the central zone and surrounds the urethra. Finally, the fibromuscular zone (also referred to as anterior region) builds the outer ventrolateral part of the prostate and consists of fibrotic and muscular connective tissue without acini.

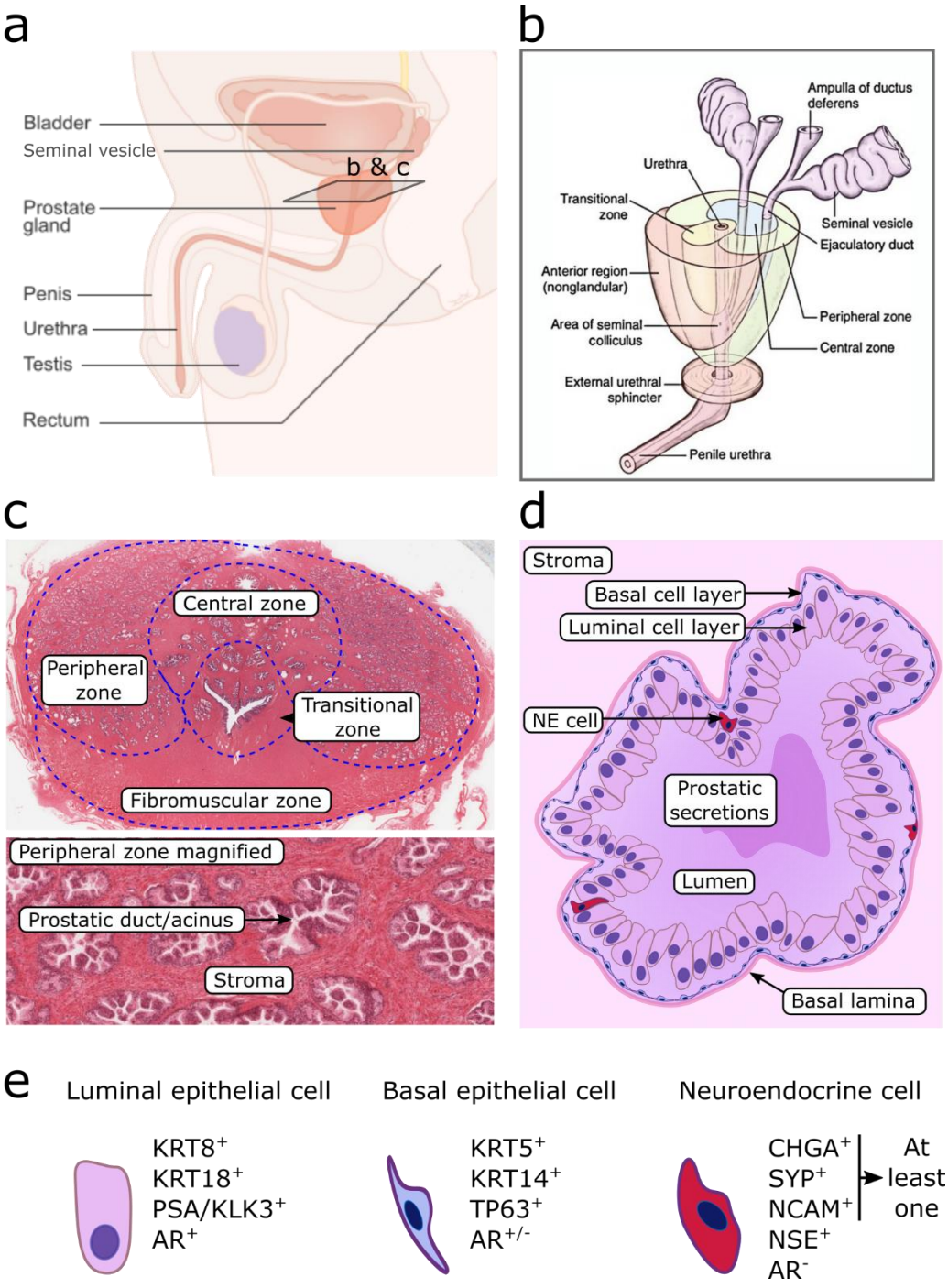

Figure 1: Macroscopic and microscopic anatomy of the healthy prostate. (a) The location of the prostate relative to the bladder and other anatomic structures is shown. The warped square (caption b & c) represents a cross-sectional plane. (b) Volumetric model of the prostate and the seminal vesicles. The prostate features a cross-sectional cut and the anatomic zones of the prostate are highlighted in color. (c) Cross-section of a prostate. The upper image shows a complete cross-section. The four major anatomic zones are demarcated by dashed lines. The lower image depicts a magnification of the peripheral zone. Individual prostatic acini are visible. (d) Schematic of a prostatic acinus. Luminal and basal epithelial cells interspersed by neuroendocrine (NE) cells together form a stratified epithelium and surround the inner lumen where the prostatic secretions flow towards the urethra. (e) Marker profiles of cell types found in the acinar epithelium. AR: androgen receptor, CHGA: chromogranin A, KRT: (cyto)keratin, NCAM: neural cell adhesion molecule 1, NSE: neuron-specific enolase, PSA: prostate-specific antigen, TP63: tumor protein 63, SYP: synaptophysin. For image references see **section 9**.

The acini consist of a two-layered, stratified and convoluted epithelium forming a central lumen (**Figure 1 d**). The luminal epithelial cells produce and secrete the prostatic fluid and form the inner layer. Luminal epithelial cells are positive for the cytokeratins KRT8 and KRT18, for prostate-specific antigen (PSA/KLK3) and for the androgen receptor (AR)[6] (**Figure 1 e**). The basal epithelial cells are connected with the basal lamina and form the outer layer (**Figure 1 d**). Basal epithelial cells are positive for KRT5, KRT14 and tumor protein 63 (TP63) and characterized by low levels or absence of the AR[6] (**Figure 1 e**).
Neuroendocrine (NE) cells represent ~1% of the epithelial cell population[7], are filled with eosinophilic granules and are observed within both epithelial layers. Their exact role is incompletely understood, however, there is evidence that they provide paracrine signaling cues to surrounding epithelial cells by secretion of neuropeptides and cytokines[7-9]. On a molecular level, NE cells are defined by immunohistochemical positivity of at least chromogranin A (CHGA) or synaptophysin (SYP) or neural cell adhesion molecule 1 (NCAM)[10]. Expression of neuron-specific enolase (NSE) is very common in NE cells but not sufficient on its own. In addition, NE cells lack AR expression[10] (**Figure 1 e**).

1.2.2 Epidemiology and Risk Factors

In men, prostate cancer (PCa) has the second highest incidence among all cancers in the developed world and was responsible for 6.8% of cancer-related mortality worldwide in 2020[1]. The risk factors for PCa are manifold. The single, most important risk factor is age. Approximately 85% of PCa cases are diagnosed in men older than 65 years[11]. Ethnicity is another important factor. The annual incidence of PCa cases per 100.000 individuals is 1.9 in Asia (China, India and Japan), 161 in the US, and 272 among African Americans living in the US[11]. Men of Japanese origin that migrated to the US have a higher risk of developing PCa than Japanese men living in Japan, suggesting an influence of external factors like diet and lifestyle[11]. Familial history of PCa is a strong risk factor too. The individual risk of developing PCa increases 2-fold when either the father or a brother has previously been diagnosed with PCa[11]. The increased risk associated with ethnicity and familial history is rooted in genetics. The list of PCa susceptibility genes and single nucleotide polymorphism (SNPs) is long and is a prolific field of research. Especially early-onset PCa (age ≤ 55 years) is often associated with specific genetic mutations. For example, detrimental mutations in breast cancer early onset 2 (BRCA2), checkpoint kinase 2 (CHEK2) and 8-oxoguanine DNA glycosylase (OGG1) increase PCa risk. In addition, the fusion oncogenes TMPRSS2-ERG and TMPRSS2-ETV1 increase PCa risk and are associated with more aggressive disease[11]. Finally, studies link infection, inflammation and serum androgen levels to increased PCa risk[11].

1.2.3 Subtypes and their Origins

The prostatic adenocarcinoma (AD) of epithelial origin represents over 95% of all PCa cases[12]. Approximately 90 – 95% of prostatic ADs are of acinar and 5 – 10% of ductal origin, respectively[12,13]. In contrast to urothelial AD, which originates from the epithelium lining the urethra, ductal prostatic AD originates from primary acini (ducts) of the *periurethral* region. Acinar AD arises from branches of the acini that are more distal to the urethra[13]. About 70% of prostatic ADs arise from the peripheral zone, 20% from the transitional zone, and 10% from the central zone, respectively[12] (**Figure 1 b** and **c**). It was believed that prostatic ADs exclusively arise from luminal cells (**Figure 1 d**) until eventually, it was shown that basal cells have the ability to initiate PCa as well[14]. There is an ongoing discourse regarding the predominant cellular origin of prostatic AD[15,16].

Next to prostatic AD, there are less common types of PCa. Among these, neuroendocrine PCa (NEPC) is clinically most significant. NEPC is characterized by malignant cells that exhibit NE features and express NE markers such as CHGA and SYP. It was proposed that NEPCs arise from the small population of native NE cells that resides in the healthy prostate[17]. However, newer evidence detailed that NEPCs primarily arise from prostatic ADs of luminal or basal origin through trans-differentiation and clonal selection[18]. For example, treatment with drugs targeting the AR pathway is one factor driving typical AD cells toward the NE phenotype[19,20]. NEPCs can be further subdivided into several subtypes based on histology[10]. Finally, there are other types of carcinomas that can emerge from the prostate. Squamous cell cancer of the prostate is a rare (≤ 1% of PCa cases) malignant entity that arises from squamous epithelial cells that line the surface of the prostate[21]. In addition, sarcomas of prostatic origin (≤ 1% of PCa cases) have been described[22].

1.2.4 Screening and Early Detection

Early detection of PCa primarily relies on digital-rectal examination (DRE) of the prostate and the PSA blood test[23]. Due to its proximity to the rectal wall, trained physicians are able to notice suspicious enlargements of the prostate by palpation, especially if the tumor is located in the peripheral zone (**Figure 1 b** and **c**). The serum PSA level may be determined after a suspect DRE or on a stand-alone basis. The PSA level has to be interpreted in the context of patient age[23]. Serum PSA levels ≥ 4 or ≥ 10 ng/mL are associated with a ~25% or ~50% chance of PCa, respectively. DRE and PSA cannot definitively discriminate between malignant neoplasia and other conditions such as benign prostatic hyperplasia or inflammation and thus, biopsy is the only way of definitive PCa diagnosis. Due to biopsy-associated risks, the European Association of Urology recommends early DRE and PSA testing (before 50 years of age) only for well-informed men at elevated risk of having PCa[23]. There are numerous other tests and scoring modalities. Some of them complement DRE and PSA results, while others inform treatment decisions after initial diagnosis. Olleik et al. review the dynamic landscape of novel tests that guide PCa diagnosis and early treatment[24].

1.2.5 Clinical Diagnosis, Staging and Stratification

A definitive PCa diagnosis requires a trans-perineal or trans-rectal prostate needle biopsy with subsequent histologic evaluation by a pathologist. Typically, this invasive procedure is guided by magnetic resonance imaging (MRI), computer tomography or ultrasound and around 10-12 prostate tissue samples (cores) are extracted[25]. The standard framework for assessing the aggressiveness of localized PCa is the Gleason grading system devised by Donald F. Gleason in the 1960's[26]. This system was modified in 2005[27] and complemented by the International Society for Urological Pathology (ISUP) grading system in 2014[28]. The Gleason grading system defines groups of histological patterns numbered 1 to 5, that correspond to the aggressiveness of the tumor (**Figure 2 a** and **b**). First, the most aggressive and the most abundant histological patterns in the biopsy sample are determined. Then, the values of the two patterns are added to derive the Gleason score (GS). The most aggressive pattern is reported first and the most abundant pattern is reported second. Thus, a pattern combination of 4 + 3 (GS 7b) indicates a more aggressive tumor than a pattern combination of 3 + 4 (GS 7a). The GS can also be determined from prostatectomy samples. In this case, the two most abundant histological patterns are considered irrespective of their aggressiveness. The ISUP grades that were established in 2014 are based on the Gleason grading system[28] (**Table 1**). In addition to the GS and ISUP grading, the tumor-node-metastases (TNM) classification system for malignant tumors is applied. The TNM code describes the size and location of the primary tumor (T), the presence of localized nodal infiltration (N) and the presence distant metastases (M)[29] (**Figure 3 a**). Additional biopsies or contrasted MRI are necessary to determine the N and the M stages. Finally, the D'Amico classification system[30] stratifies PCa patients into low-, intermediate- and high-risk groups based on PSA, GS and TNM status (**Table 2**).

Table 1: Relationship between Gleason patterns, Gleason scores and ISUP Grades.

Gleason Patterns	Gleason Score	ISUP Grade	Description of ISUP Grade Group
3 + 3 and lower	≤ 6	1	Only individual discrete well-formed glands
3 + 4	7a	2	Predominantly well-formed glands with lesser component of poorly-formed/fused/cribriform glands
4 + 3	7b	3	Predominantly poorly-formed/fused/cribriform glands with lesser component of well-formed glands
4 + 4 **or** 3 + 5 **or** 5 + 3	8	4	Only poorly-formed/fused/cribriform glands **or** predominantly well-formed glands and lesser component lacking glands **or** predominantly lacking glands and lesser component of well-formed glands
4 + 5 **or** 5 + 4 **or** 5 + 5	9-10	5	Lacks gland formation **or** with necrosis: with or without poorly formed/fused/cribriform glands

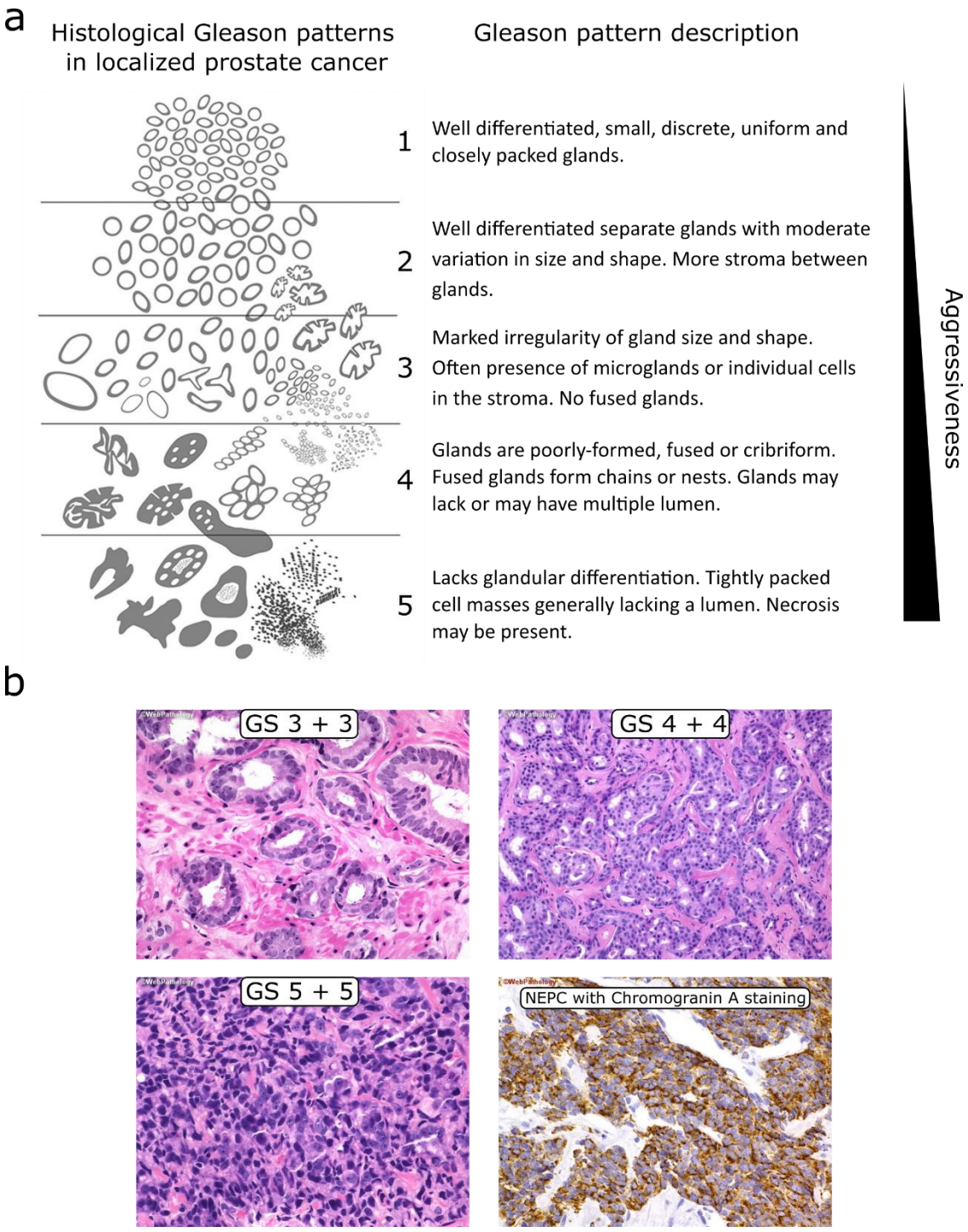

Figure 2: Gleason patterns. (a) Left: revised version of the original drawing made by D. F. Gleason MD, established after the 2014 ISUP conference. Right: description of the respective Gleason patterns. (b) The histological stainings show representative Gleason stages of prostatic AD. Upper left (GS 3+3): the aberrant acini have small round or oval lumina without epithelial convolutions. The basal cell layer is missing. Upper right (GS 4+4): the epithelial neoplasia is progressing. Individual glands are hard to define. Lower left (GS 5+5): tightly packed tumor cells. No discernible gland formation. Lower left (NEPC sample): cells are positive for CHGA in this sample. For image references see **section 9**.

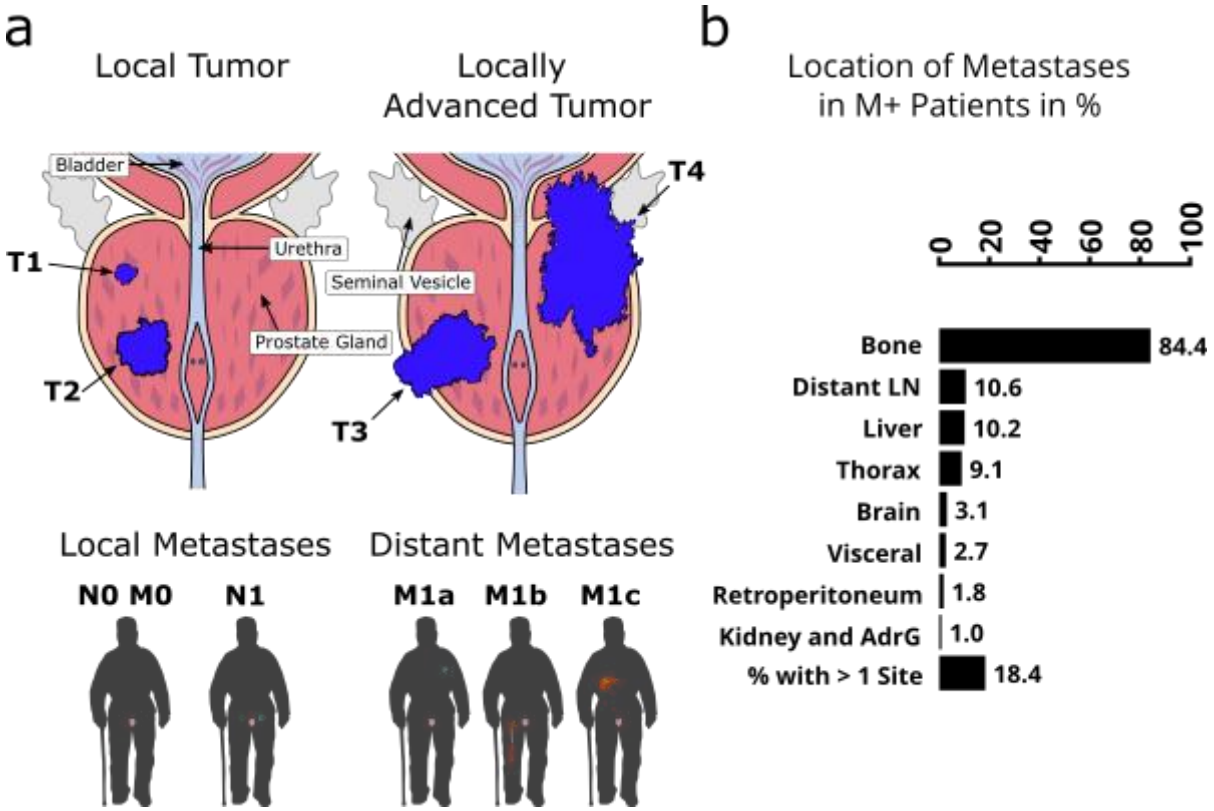

Figure 3: TNM stages and location of PCa metastases. (a) Schematic illustrating the T, N, and M stages of PCa. (b) Bar-plot showing the distribution of metastatic sites found in PCa. The data are adapted from[31]. LN: lymph node, AdrG: adrenal gland.

Table 2: Risk groups for localized PCa.[23] Except for minor modifications these risk groups are equivalent to the risk groups defined by D'Amico in 1998[30]. * refer to [29] for details on TNM stages.

Low-risk	Intermediate-risk	High-risk	
PSA < 10 ng/mL	PSA 10-20 ng/mL	PSA > 20 ng/mL	**any** PSA
and GS < 7 (ISUP grade 1)	**or** GS 7 (ISUP grade 2/3)	**or** GS > 7 (ISUP grade 4/5)	**any** GS (**any** ISUP Grade)
and cT1-2a*	**or** cT2b*	**or** cT2c*	cT3-4 or cN+*
Localized			Locally advanced

1.2.6 The Androgen Dependency of Prostate Cancer

PCa is initially castration-sensitive, meaning that tumor cells depend on external androgen stimuli and AR signaling to proliferate and survive. Testosterone is the primary circulating androgen and is produced in the testicular Leydig cells and – to a small extend – in the adrenal cortex. After diffusion into the cytoplasm, testosterone is converted to dihydrotestosterone (DHT) by the enzyme 5α-reductase. DHT has a 5-fold higher potential to activate the AR than testosterone[32]. Binding of testosterone, DHT or other androgens leads to a conformational change and to homodimerization of the AR. The AR homodimer is then translocated to the nucleus where it binds to androgen-response elements in the genome and promotes the expression of AR-controlled genes[32]. The transcription of AR-controlled genes like TMPRSS2 and others is essential for proliferation and cell survival in castration-sensitive PCa cells.
Androgen deprivation therapy (ADT), which will be mentioned several times in the following sections on PCa treatment, refers to the lowering of the serum androgens to castrate levels. ADT can be achieved by orchiectomy or medical castration through administration of luteinizing hormone-releasing hormone analogues or antagonists[33]. Castration-sensitive PCa responds to ADT for typically 2 – 3 years[32]. Eventually, the tumor acquires the ability to grow despite castrate serum androgen levels. The mechanisms that abrogate the natural androgen-dependency of PCa and lead to castration resistance are discussed in section **1.3.5**.

1.3 Prostate Cancer Treatment

1.3.1 Localized Disease

Locally confined PCa (**Figure 3 a** and **Table 2**) has a favorable outlook and is curable by either radical prostatectomy (RP) or external beam radiotherapy (EBRT)[23]. Active surveillance by regular DRE and PSA tests is a third approach that can be considered for low-risk patients (**Table 2**). If the surveillance indicates unacceptable disease progression, RP or EBRT will be initiated[23]. Furthermore, palliative watchful waiting without regular surveillance may be considered in frail patients with <10 years of life expectancy[23]. RP is the surgical removal of the prostate and the attached seminal vesicles. EBRT is the umbrella term for radiation therapies with an external radiation source. X-ray is the typical, but not the only type of radiation used in EBRT. Modern EBRT employs imaging techniques to target the radiation with high precision, reducing the exposure of surrounding healthy tissue and limiting side effects[23]. Brachytherapy (radioactive implants) is an alternative to EBRT in selected patients[23]. For low- and intermediate-risk PCa patients, RP and EBRT yield a comparable cancer-specific survival (CSS). A prospective randomized trial found that the 10-year CSS was at 99% regardless of whether the patients received RP or EBRT. The cohort consisted of roughly three quarters low- and one quarter intermediate-risk patients[34] (**Table 2**). Lower values for a 10-year CSS of ~80 – 90% were reported in cohorts with higher risk profiles[35] and in older cohorts[36].

1.3.2 Locally Advanced Disease

Locally advanced disease (**Figure 3 a,** T3-4) is characterized by neoplastic infiltration of tissue surrounding the prostate. Watchful waiting may be considered as described. The treatment is usually multimodal and may include RP, EBRT and ADT as well as sequential applications or combinations of these modalities[23]. The 10-year CSS is considerably lower than for locally confined PCa and patients have a higher risk of biochemical relapse (section **1.3.3**) after initial treatment. In high-risk patients, the risk of infiltration of local lymph nodes is 10–40%[23]. Only selected T3–4 patients can be considered for RP and adjuvant EBRT after RP is recommended[37]. If RP is not a viable option, EBRT combined with ADT is the standard of care[23].

1.3.3 Relapse after Curative Treatments

After RP, the serum PSA level generally falls below 0.2 ng/mL, whereas after EBRT, it drops and remains stable at a low level (the nadir). A biochemical relapse (BCR) is defined as two consecutive rising PSA measurements above 0.2 ng/mL after RP, or as a PSA level that is more than 2 ng/mL above the nadir after EBRT[33]. A BCR indicates that the initial curative treatment has failed. BCR after initial presentation with a T1–2 tumor has a good prognosis. If the BCR is detected early after RP, patients still have a 60% chance to be cured by salvage EBRT with or without ADT[33]. Similarly, salvage RP can be considered for fit patients that present with BCR after EBRT and after confirming local recurrence by biopsy or imaging. For most patients that have a BCR after initial treatment of a T3 or T4 tumor, curative treatment is not feasible anymore. These patients have a high chance of metastatic disease.

1.3.4 Metastatic Castration-Sensitive Disease

PCa metastasizes to various tissues, however, bone- and lymphatic metastases are most common[31] (**Figure 3 b**). Patients that present with metastasized castration-sensitive PCa (mCSPC) have a median overall survival (OS) of 30 – 45 months and require systemic treatment[38]. The current standard of care is the systemic ADT combined with either docetaxel (DTX)[33] or with AR signaling inhibitors (ARSIs). Abiraterone acetate[38], enzalutamide[39] and apalutamide[40] are approved in combination with ADT.

1.3.5 Castration-Resistant Disease

Castration resistance is a major challenge in PCa treatment and refers to the acquired ability of tumor cells to grow despite castrate serum androgen levels[32]. Castration-resistant PCa (CRPC) is defined by a BCR or a radiological progression despite a serum testosterone level < 50 ng/dl[41]. Numerous mechanisms promote castration resistance[32]. PCa cells can develop androgen hypersensitivity, either through AR amplification or AR mutations. The AR can acquire mutations that allow activation through alternative substrates. Other mutations can lead to constitutively active AR variants. Mutations that alter AR signaling can also occur in transcriptional co-activators or co-repressors. Finally, intra-tumoral production of androgens can occur by utilizing androgen precursors of adrenal origin[32]. ARSIs were developed to counteract these mechanisms. Abiraterone acetate[42] inhibits CYP17, an enzyme involved in androgen production. Enzalutamide[43], apalutamide[44] and darolutamide[45] act more directly on the AR. They are AR antagonists that inhibit substrate binding, nuclear translocation and DNA binding of the AR[43-45].

1.3.5.1 Non-metastatic CRPC

Patients that progressed to castration resistance on initial ADT but exhibit no evidence of metastases have non-metastatic CRPC. In these patients, ADT is continued and usually combined with either enzalutamide, apalutamide or darolutamide[41] to overcome castration resistance.

1.3.5.2 Metastatic CRPC

The treatment of metastatic CRPC (mCRPC) is complex and should be carried out in an experienced center within a multidisciplinary team[41]. The first-line treatment options include ADT in combination with DTX, abiraterone or enzalutamide. All patients with mCRPC eventually progress on their first-line treatment[41]. Second- and third-line treatments depend on the prior-line treatment and include ADT in combination with enzalutamide, abiraterone, DTX, cabazitaxel (after progression on DTX), the poly (ADP-ribose) polymerase 1 (PARP1) inhibitor olaparib (in patients with BRCA1/2 mutations, section **1.3.7**) and radium-223 (third-line, exclusively for bone metastases)[33,41]. Furthermore, lutetium-177-PSMA therapy can be considered for selected patients after progression on at least one ARSI and one taxane[46]. Most PCa cells express the surface marker prostate-specific membrane antigen (PSMA). This therapy exploits PSMA expression to perform a targeted delivery of radioactive lutetium-177 isotopes to the cancer cells[46].

1.3.6 Neuroendocrine Prostate Cancer

NEPC is very rare at initial diagnosis (<1% of cases), however, it is common after tumor progression on treatment with AR signaling-ablative therapies such as abiraterone acetate, enzalutamide or darolutamide. Prospective biopsies of metastatic tissue revealed NE features in 11.2% of the samples[47] and autopsies revealed their presence in 10–20% of men dying with mCRPC[48-50]. NEPC is characterized by a distinct set of markers (**Figure 1 e**, **Figure 2 b** and section **1.2.3**) and by reduced sensitivity to androgen ablation.
Differentiated treatment approaches are taken to mitigate NEPC. However, NEPC has a poor prognosis with a median OS of usually less than a year[51]. NEPC shows a frequent albeit short-lived response to platinum-based chemotherapy. Therapies consist of either cisplatin or carboplatin in combination with taxanes or etoposide[51].

1.3.7 Prostate Cancer with Defects in Homologous Recombination Repair Genes

Genomic profiling across disease stages revealed gradually increasing numbers of deleterious DNA repair gene mutations from locally confined PCa to mCRPC[47]. 10% of mCRPC-derived tumor specimen had BRCA2 mutations. Patients with defects in homologous recombination (HR) repair genes – most notably in *BRCA1* & *BRCA2* – are an important subset of mCRPC patients. Inhibition of the DNA repair enzyme PARP1 with olaparib proved to be effective in these patients and is the treatment of choice after progress on ARSIs[41,52]. Several case and retrospective studies suggest that platinum-based chemotherapy exhibits increased effectiveness in HRR-deficient PCa[53-56].

1.4 Cellular Mechanisms of Chemoresistance in Prostate Cancer

1.4.1 Resistance against Taxanes

1.4.1.1 Taxanes: Mechanism of Action

Taxanes bind the ß-tubulin subunit of α/ß-tubulin dimers, which constitute the cellular microtubules (MTs). Taxanes stabilize MTs and prevent their physiological depolymerization. The complex processes that occur during cell division heavily depend on MT dynamics. Taxanes generally cause a cell cycle arrest in the G_2/M-phase, ultimately culminating in apoptotic cell death[57]. Paclitaxel (PTX) is a natural taxane isolated from the bark of the Pacific yew tree *Taxus brevifolia* in the 1970s and was the first taxane used in cancer therapy[58]. The semi-synthetic DTX was discovered in 1981[58]. Finally, cabazitaxel is a semi-synthetic taxane that has cytotoxic activity in DTX-resistant cancers and was approved as a second-line therapy for CRPC after progression on DTX in 2010[59,60]. Tumor cells employ various strategies to achieve taxane resistance[61] (**Figure 4**). These are discussed in the following sections.

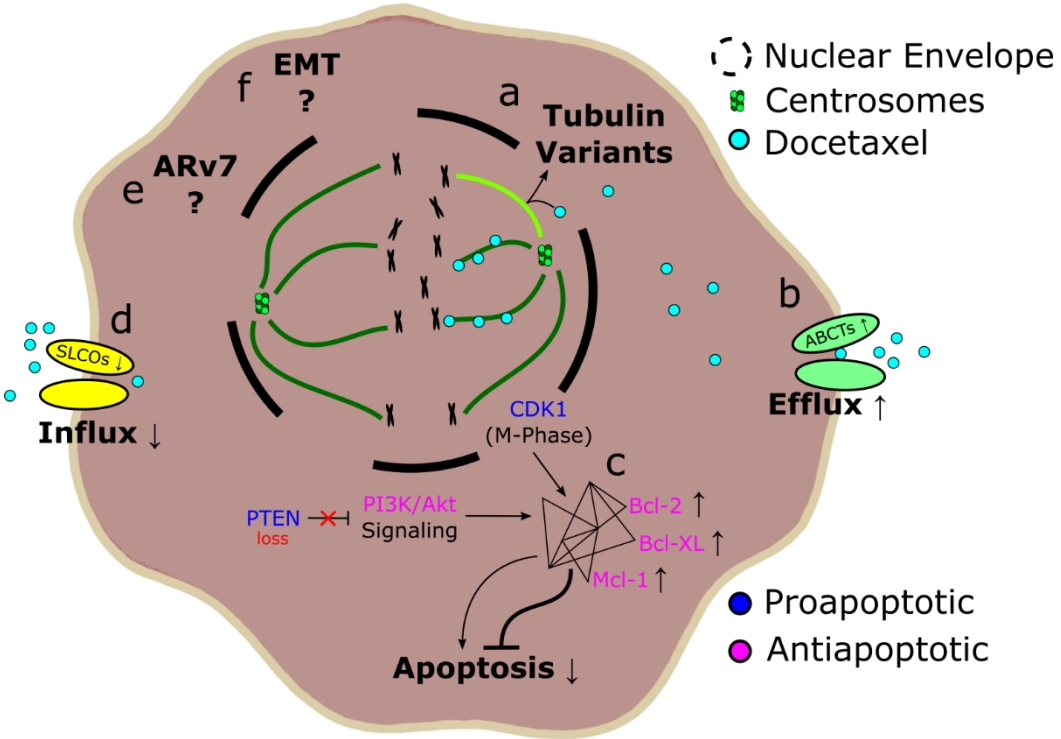

Figure 4: Resistance mechanisms against taxanes. The schematic depicts a dividing cell in M-phase as seen by the disintegrated nuclear envelope. The various mechanisms of DTX resistance discussed in section **1.4.1** are summarized. (a) Increased expression of tubulin variants such as *ß*III-tubulin counteracts the effectiveness of DTX (section **1.4.1.2**). (b) Increased efflux of DTX (section **1.4.1.3**). (c) Changes to the apoptotic signaling network result in reduced propensity towards programmed cell death (section **1.4.1.4**). Pro- and antiapoptotic elements are colored in blue and magenta, respectively. The chronic expression of CDK1 in M-phase-arrested cells provides a link between taxane treatment and the induction of apoptosis. (d) Decreased influx of DTX (section **1.4.1.5**). (e) A role for the AR variant ARv7 is discussed in the literature but lacks comprehensive evidence (section **1.4.1.7**). (f) The EMT program drives cellular plasticity and is activated upon taxane treatment but concrete links between EMT and effectors of taxane resistance have not been unraveled so far (section **1.4.1.8**).

1.4.1.2 ß-Tubulin Isotypes and Mutations

In humans, *ß*-tubulin exists as eight distinct isotypes that exhibit differential expression in different tissues. Certain tubulin isotypes contribute to taxane resistance (**Figure 4 a**). *ß*Ia-tubulin is expressed in almost all tissues and is the most widespread isoform. Mutations in the taxane-binding site of *ß*Ia-tubulin have been implicated with increased resistance of tumor cells to PTX and DTX[62]. Increased expression of *ß*III-tubulin (TUBB3) has been linked to taxane resistance in various cancer entities including PCa[62]. The expression of *ß*III-tubulin can be triggered by taxane treatment. For example, Sobue et al. reported increased expression of *ß*III-tubulin in PTX-resistant compared to PTX-naïve PC-3 PCa cells[63]. Moreover, Terry et al. demonstrated that in PCa, the expression of *ß*III-tubulin is elevated after androgen ablation *in vivo* and in castration-resistant cell culture models[64].

1.4.1.3 Increased Taxane Efflux

Transmembrane proteins of the ATP-binding cassette (ABC) family facilitate the transport of numerous structurally unrelated compounds across cell membranes[65]. The ABC family of membrane transporters comprises of 49 genes and is divided into seven subfamilies (A-G)[65]. ABC transporters (ABCTs) are primarily expressed in epithelial tissues with barrier or secretory function and typically facilitate an efflux of molecules from the cytosol to the extracellular matrix or lumen. ABCTs protect sensitive tissues and cells from detrimental molecules but also facilitate the directed flow of physiologic metabolites. Szakacs et al. give an excellent overview over the physiological roles of ABCTs[66]. ABCTs are N-glycosylated transmembrane proteins. The prototypical ABCT has two transmembrane domains (TMDs), two intracellular ATP-binding sites and a central ligand-binding site. However, some ABCTs have three TMDs and about one third of ABCTs are transcribed as half transporters with only one TMD. The half transporters need to homo- or heterodimerize with other half transporters in order to become functional[67]. Some ABCTs have very broad substrate specificities and are associated with a phenomenon coined multidrug resistance (MDR). MDR refers to a state where a cell has become resistant to an array of several cytotoxic compounds. The first ABCT implicated in MDR was the ABCB1 protein[68], also known as P-glycoprotein (P-gp) or multi-drug resistance protein 1 (MDR1). ABCB1, for example, effluxes at least 20 substances including PTX, DTX, 5-fluorouracil, doxorubicin and mitoxantrone[67]. Other important ABCTs that are associated with MDR include ABCC1 and ABCG2[65]. Many ABCTs including ABCA3, -B1, -B4, -C2 and -C10 efflux taxanes[67] (**Figure 4 b**). However, *in vitro* studies indicate that ABCB1 is the primary ABCT mediating taxane efflux in taxane-resistant PCa cell culture models[69-75]. Inhibition of ABCB1 completely resensitized taxane-resistant cell lines in several of these studies. Potent and specific small molecule inhibitors (e.g., elacridar[76] and tariquidar[77]) were designed to target ABCTs, especially ABCB1 and ABCG2. These molecules proved very effective *in vitro* and were subsequently evaluated in combination with classical chemotherapies in many clinical trials and various cancer entities[78]. However, auxiliary systemic inhibition of ABCB1 either lacked effectiveness or was associated with significant treatment-related toxicity and thus, none of these combinatorial treatments have been approved by the regulatory authorities so far[78].

1.4.1.4 Apoptotic Escape

Apoptotic escape is a general resistance mechanism that applies to many drugs, including taxanes (**Figure 4 c**). Apoptosis is a physiologically ubiquitous suicide program eliminating unnecessary or dysfunctional cells[79]. The apoptotic program is executed by cysteine proteases, which are controlled by a complex gene regulatory network revolving around the B cell lymphoma 2 (Bcl-2) protein family[79]. Cancer cells are under constant stress caused – for example – by the hypoxic tumor microenvironment or by genomic instability. These stimuli would usually trigger the apoptotic program, but cancer cells frequently acquire the capacity to escape programmed cell death. They achieve this by loss, downregulation or loss of function (LoF) mutation of proapoptotic regulatory genes or by amplification, upregulation or gain of function (GoF) mutation of antiapoptotic regulatory genes[79].
In the context of taxane treatment, the link between mitotic cell cycle arrest and the induction of apoptosis is of specific interest. The cyclin dependent kinase 1 (CDK1) is transiently expressed during the mitotic M-phase under physiological conditions, but chronically in cells that are caught in a mitotic arrest. CDK1 phosphorylates antiapoptotic proteins such as Bcl-2, Bcl-XL[80] and Mcl-1[81], which antagonizes their antiapoptotic activity. Consequentially, chronic expression of CDK1 is a potential link between M-phase arrest and the induction of apoptosis[80,81]. Targeting these antiapoptotic proteins with siRNA or with small molecule inhibitors led to decreased taxane tolerance in taxane-resistant cell culture models of PCa[61]. Together, these studies demonstrate that antiapoptotic proteins contribute to taxane resistance in PCa[61].

1.4.1.5 Decreased Taxane Influx

In addition to increased taxane efflux mediated by ABCTs, the decreased influx of taxanes due to downregulation or LoF mutation of solute carrier of organic anions (SLCO) genes is another mechanism of controlling the cytosolic taxane concentration that has been described in PCa cells and patients[61] (**Figure 4 d**).

1.4.1.6 The PI3K/Akt Pathway

The PI3K/Akt pathway is a pro-survival signaling pathway and its increased activity has been linked to taxane resistance in PCa[61]. Specifically, loss or inactivation of phosphatase and tensin homologue deleted on chromosome 10 (PTEN), which is a negative regulator of the PI3K/Akt pathway and occurs frequently in progressed PCa, has been implicated with increased taxane resistance[61].

1.4.1.7 AR Signaling

An interplay between AR signaling and taxane resistance is debated in the literature. After ligand-binding and homodimerization in the cytosol, the AR is translocated to the nucleus. This translocation depends on MTs and is disturbed by treatment with taxanes[61,82]. It has been shown that splice variants of the AR differentially associate with MTs. In one study, one of these variants – ARv7 – was associated with increased taxane resistance after expression in a PCa cell culture model[83]. However, another study failed to reproduce these results[84]. At the present time, the existence of a link between AR signaling and taxane resistance lacks strong evidence and needs further investigation[61] (**Figure 4 e**).

1.4.1.8 Epithelial to Mesenchymal Transition

Various groups have shown that treatment of PCa cells with taxanes leads to reduced expression of epithelial markers such as E-cadherin (CDH1) and increased expression of mesenchymal markers such as vimentin (VIM)[85,86]. The shift from an epithelial marker profile toward a more mesenchymal marker profile is termed epithelial to mesenchymal transition (EMT) and is associated with increased cellular capacity for migration, invasion and survival[87]. It is well established that taxane-treated PCa cells undergo EMT[85,86], and that EMT has wide-ranging implications for cellular behavior and gene expression[87]. For example, the knockdown of the EMT-inducing transcription factor (TF) zinc finger E-box binding homeobox 1 (ZEB1) caused a moderate resensitization of DTX-resistant PCa cells[86]. However, how exactly EMT is connected with actual effectors of taxane resistance is incompletely understood (**Figure 4 f**).

1.4.2 Resistance against Platinum Compounds

1.4.2.1 Platinum Compounds: Mechanism of Action

Platinum (Pt) compounds are among the oldest chemotherapeutic drugs. Cisplatin (CDDP) was the first clinically used Pt compound and was accidentally discovered in the 1960s[88]. Carboplatin (CAR) was discovered in the early 1980s[88]. Due to their high antitumor activity, CDDP and CAR are essential for the treatment of – for example – testicular and ovarian cancer. Newer compounds include oxaliplatin and satraplatin. Pt compounds react with DNA and generate intra- or interstrand Pt-DNA adducts[88]. Typically, these DNA lesions impede genome replication during S-phase, rendering cells unable to progress through this phase of the cell cycle. Ultimately, the stalled cell cycle and the damaged DNA trigger apoptotic cell death[88]. As for DTX, there are numerous mechanisms potentially contributing to Pt resistance (**Figure 5**).

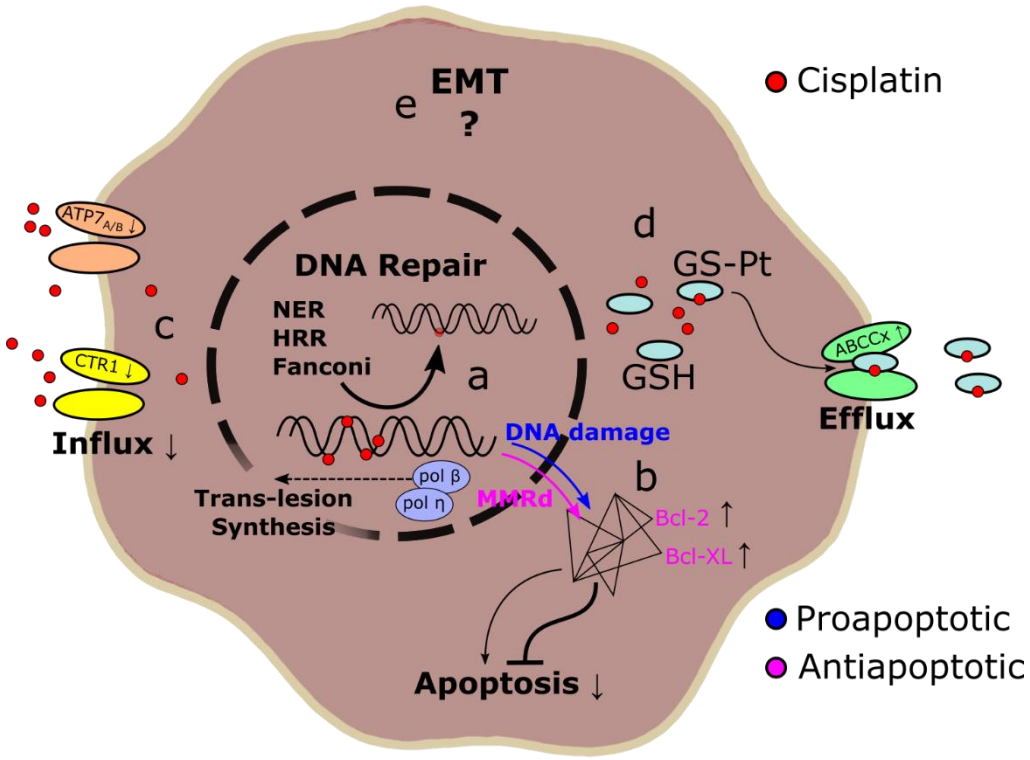

Figure 5: Resistance against platinum compounds. The schematic summarizes the resistance mechanisms against Pt compounds discussed in section **1.4.2**. (a) DNA repair pathways (sections **1.4.2.2 - 1.4.2.5**) are at the core of Pt resistance. Nucleotide excision repair (NER), homologous recombination repair (HRR) and the Fanconi anaemia pathway repair different kinds of Pt-induced DNA lesions and reduce the overall load of lesions. Trans-lesion synthesis (section **1.4.2.6**) bypasses DNA lesions during S-phase and promotes progression through the cell cycle. (b) Changes in the regulation of apoptosis (section **1.4.2.7**) with regard to Pt resistance are shown. Pro- and antiapoptotic elements are colored in blue and magenta, respectively. DNA lesions and stalled DNA mismatch repair (MMR) complexes provide proapoptotic signals. Increased expression of antiapoptotic genes and MMR deficiency (MMRd) can lead to escape from apoptosis. (c) Reduced expression of Pt-influxing membrane transporters reduces the intracellular concentration of Pt compounds (section **1.4.2.8**). (d) Reduced glutathione (GSH) binds and thereby inactivates cytosolic Pt compounds. GS-Pt complexes are then effluxed by ABCTs of the C-family (section **1.4.2.8**). (e) Pt treatment is associated with EMT but how exactly EMT translates into altered Pt tolerance is incompletely understood (section **1.4.2.9**).

1.4.2.2 DNA Repair: Nucleotide Excision Repair

DNA repair is a complex network of specialized sub-pathways that have partial overlaps. Most DNA repair pathways are involved in the repair of Pt-DNA lesions to some extent (**Figure 5 a**). The pathways that are of specific importance to Pt tolerance are discussed hereafter. The nucleotide excision repair (NER) pathway involves about two dozen proteins, most notably the excision repair cross-complementing family (e.g., ERCC1) and the xeroderma pigmentosum complementation groups (e.g., XPA)[89]. NER recognizes and repairs various types of helix-distorting DNA damage and is very effective at repairing intrastrand Pt-DNA adducts. However, NER is unable to repair interstrand crosslinks (ICLs)[89], which are especially problematic for DNA replication and gene transcription. Duan et al. give a comprehensive overview over the many studies that link the NER pathway with Pt resistance[89].

1.4.2.3 DNA Repair: The Fanconi Anemia Pathway

Fanconi anemia (FA) is a rare, genetically heterogeneous inherited syndrome associated with extreme sensitivity to DNA-crosslinking agents[90]. The FA complementation (FANC) group members (e.g., FANCA) are a set of 14 genes that together constitute the core of the FA pathway. The FA pathway is specialized in the repair of DNA ICLs[90] (**Figure 5 a**).

1.4.2.4 DNA Repair: Processing of Platinum-Induced Double Strand Breaks

As mentioned, ICLs are specifically problematic for DNA replication. When the replicative machinery encounters an ICL, DNA double strand breaks (DSBs) can occur[91]. DSBs are a secondary consequence of Pt treatment that occur during DNA replication[91]. DSBs can be repaired either through variants of homologous recombination repair (HRR) (**Figure 5 a**) or through variants of nonhomologous end joining (NHEJ)[92]. Defective HRR is associated with a better response to Pt treatment in several entities including breast[93], ovarian[94] and pancreatic[95] cancer. LoF of either BRCA1 or BRCA2 is the most common cause of HRR deficiency in PCa.

1.4.2.5 DNA Repair: Mismatch Repair – A Counterintuitive Mechanism

The DNA mismatch repair (MMR) pathway recognizes and repairs mismatched bases[96]. The MutS homologs 2,3 & 6 (MSH2, MSH3 & MSH6) and the MutL homolog 1 (MLH1) take central roles in this pathway[96]. Most mismatched bases are generated during genome replication due natural error rates inherent to DNA polymerases[96]. The MMR system also recognizes certain lesions caused by CDDP and CAR[88], however, the MMR system is unable to repair these[97,98].
It is hypothesized that the stalled MMR complexes provide signaling cues that translate into the initiation of apoptotic cell death[88,97-99]. Thus, MMR deficiency (MMRd) is associated with increased CDDP and CAR resistance and vice versa.

1.4.2.6 Trans-lesion Synthesis

Trans-lesion synthesis allows the replication of the genomic DNA by bypassing Pt-induced lesions[88,100]. The DNA polymerases α, δ and ε, which are the dominant polymerases involved in the replication of the DNA during S-phase, are not able to pass Pt-induced DNA lesions[100]. However, the DNA polymerases ß and η (pol ß and pol η), which have physiological functions in DNA repair pathways, are able to bypass Pt-induced DNA lesions[100] (**Figure 5 a**). This process introduces mutations at the original site of the Pt-DNA lesion[100]. Despite that, it has been shown that trans-lesion synthesis via pol η confers resistance to Pt compounds *in vitro*[101].

1.4.2.7 Apoptotic Escape

The importance of apoptotic escape for cancer chemoresistance was outlined in section **1.4.1.4**. In cells that were treated with Pt compounds, the major origins of apoptotic signals are DNA lesions (especially DSBs)[102] and stalled DNA repair proteins (e.g., MMR complexes)[99]. Pt-resistant cells acquire features that interfere with the induction of apoptosis, for example, the upregulation of antiapoptotic proteins[103] (**Figure 5 b**).

1.4.2.8 Control of Platinum Concentration through Membrane Transporters

Pt compounds poorly diffuse across the cell membrane passively[88]. Membrane transporters play an important role in the import and export of Pt compounds (**Figure 5 c** and **d**). The copper transporter-1 (CTR1) imports Pt compounds[104]. Resistance against CDDP, CAR and oxaliplatin was increased in a $CTR1^{-/-}$ cell culture model[105]. Exposure to CDDP promotes the internalization and proteasomal degradation of CTR1[106], which represents a negative feedback mechanism promoting resistance. The P-type ATPases ATP7A and ATP7B physiologically function as copper exporters. ATP7A and ATP7B efflux Pt compounds and contribute to Pt resistance[107]. Finally, ABCTs play a role in the export of Pt compounds. However, before efflux via ABCTs occurs, Pt compounds must first react with and be captured by reduced glutathione (GSH)[88]. GSH is a tripeptide that aids in cellular detoxification by binding and inactivating small molecules. After reaction with GSH, GS-X molecules (for example GS-Pt) are effluxed by members of the C-family of ABCTs including ABCC1 and ABCC2[88].

1.4.2.9 EMT

As discussed in section **1.4.1.8**, EMT is a pathway that drives plasticity in gene transcription and cellular behavior. Similar to taxane treatment, there is evidence that a treatment with Pt compounds activate the EMT program[108]. However, if and how exactly EMT influences the effectors of Pt tolerance is unclear as of now (**Figure 5 e**).

1.5 Prostate Cancer Cell Lines

In vitro cell culture models of PCa are an important resource for PCa research. There are three particularly widespread cell lines in PCa research[109]: DU145 (ATCC: HTB-81, Cellosaurus: CVCL_0105)[110], PC-3 (ATCC: CRL-1435, Cellosaurus: CVCL_0035)[111] and LNCaP (ATCC: CRL-1740, Cellosaurus: CVCL_1379)[112]. Less widely used are LAPC-4 cells[113] as well as the LNCaP derivatives C4-2[114,115] and MR49F[116]. All mentioned cell lines were isolated from PCa metastases, are immortal and grow as adherent monolayers. Karyotyping revealed aneuploidy with variable chromosome counts in all cell lines. While DU145 and PC-3 are castration-resistant, wild-type LNCaP and LAPC-4 cells are sensitive to castration. Numerous sublines have been established from DU145, PC-3 and LNCaP by various techniques including chemical treatment, intermediate culture in animals and viral transformation[109].

1.5.1 DU145

DU145 cells were isolated in 1978 from PCa metastases of the central nervous system of a 69-year-old Caucasian man. DU145 cells are castration-resistant and have lost PSA and AR expression. The number of chromosomes in DU145 cells ranges from 46 to 143 with a modal number of 64 and the population doubling time is approximately 34 h[109]. DU145 cells express the luminal cytokeratins KRT8 and KRT18 as well as the basal cytokeratin KRT5. The cadherin expression pattern $CDH1^{+}/CDH2^{-}$ suggests an epithelial phenotype of DU145 cells. On the other hand, DU145 cells also express the mesenchymal marker VIM. Furthermore, DU145 cells do not express any NE markers. Noteworthy is a homozygous genetic alteration (chr3:g.37038108A>T) affecting a splice site within the *MLH1* gene. This alteration leads to an early truncation and complete LoF of the protein causing MMRd and microsatellite instability (MSI) in DU145 cells[117]. Furthermore, DU145 cells feature a homozygous variant (chr13:g.49037903A>T) resulting in a premature stop codon in the *RB1* gene[117].

1.5.2 PC-3

PC-3 cells were isolated from lumbar vertebral metastases of a 62-year-old Caucasian man with PCa in 1979[111]. Like DU145 cells, PC-3 cells do not express the AR and PSA. PC-3 cells have a modal number of 58 chromosomes and an approximate population doubling time of 33 h[109]. The cytokeratin expression pattern $KRT8^{+}$, $KRT18^{+}$ and $KRT5^{+}$ suggests a mixed luminal/basal phenotype. According to Sobel *et al*.[109], PC-3 cells are negative for the mesenchymal marker CDH2. Furthermore, PC-3 cells are NSE^{+} [118] and according to some studies $CHGA^{+}$ [119], which are two classical NE markers. Therefore, PC-3 cells can be described as a cell line with NE features. However, PC-3 cells are negative for the NE marker SYN. Notably, PC-3 cells harbor a homozygous mutation (chr17:g.7578516del) in the *tumor protein 53* gene. This mutation causes a frameshift producing an early stop codon and complete LoF of the *tumor protein 53* protein[120]. Furthermore, PC-3 cells also feature a homozygous deletion spanning several exons in the *PTEN* gene[121]. Finally, a 500kb genomic region spanning the *STAT3* gene is homozygously deleted in PC-3 cells[122].

1.5.3 LNCaP

LNCaP cells were isolated from a lymph node needle aspiration biopsy of a 50-year-old Caucasian man with PCa in 1980. They are highly aneuploid with a modal number of 76 chromosomes. The population doubling time of LNCaP cells is with approximately 60 h markedly slower than for PC-3 and DU145 cells[109]. LNCaP cells are sensitive to castration and express the AR and PSA. LNCaP cells are KRT8/18^{+} and KRT5/14^{-}, which indicates a stricter luminal phenotype as compared to DU145 and PC-3 cells, in which the basal KRT5 is expressed. LNCaP cells are CDH1^{+}/CDH2^{-}/VIM^{-}, which is characteristic of well-defined epithelial cells. LNCaP cells do not express any NE markers. The *MSH2* gene features a long homozygous deletion spanning from exon 9 to 16[123], causing MMRd in LNCaP cells. More than 60 derivative cell lines of LNCaP have been established, including several castration-resistant sublines[109].

1.5.3.1 The LNCaP Derivative C4-2

C4-2 cells are a derivative of LNCaP cells established by repeated transient (murine) *in vivo* culture including the culture in castrated animals[115]. C4-2 are able to generate tumors in castrated animals and thus, represent a castration-resistant derivative of LNCaP cells[115]. Karyotyping revealed a modal number of 83 chromosomes in C4-2 cells. With a doubling time of approximately 48 h, C4-2 cells are faster than their parental cell line LNCaP[115]. Like its parental cell line, C4-2 cells are PSA^{+} and AR^{+}.

1.5.3.2 The LNCaP Derivative MR49F

MR49F cells were derived from LNCaP cells using a similar protocol as for C4-2 cells. LNCaP cells were injected into male athymic nude mice. After castration and subsequent tumor recurrence, the animals were treated with the ARSI enzalutamide[116]. The surviving tumors were harvested, minced and reinjected into castrated and enzalutamide-dosed host animals for three additional cycles, finally resulting in the castration- and enzalutamide-resistant cell line MR49F[116]. Although the initial steps in their establishment follow the same protocol, the only relation between C4-2 and MR49F cells is that both of them originate from LNCaP cells. The establishment of C4-2 and MR49F cells took place in separate laboratories at different times.

1.5.4 LAPC-4

LAPC-4 cells were obtained from a lymph node needle aspiration biopsy of a patient with metastatic PCa. The biopsy material was minced and subcutaneously implanted into immunodeficient mice. Biopsy material from the resulting tumors was extracted, minced and reimplanted for several (>12) *in vivo* passages[113]. Although *in vivo* passaging worked for LAPC-1, 3, 4, 7, 8 and 9, only LAPC-4 cells were suitable for *in vitro* culture and resulted in a stable PCa cell line. LAPC-4 cells survive implantation into castrated host animals without any additional materials or factors. After subcutaneous implantation, the tumor volume regresses. However, after six to eight weeks of dormancy, androgen-independent growth occurs[113]. LAPC-4 cells express PSA and a wild-type AR. Karyotype analysis revealed a modal number of 89 chromosomes.

1.6 Bulk Whole Exome Sequencing of Cell Cultures

Within this work, bulk whole exome sequencing (WES) of 24 PCa cell lines was performed and the analysis of the resulting dataset constitutes a significant part of the presented results. This section provides theoretical background on WES and on the techniques employed in the analysis of the WES data.

1.6.1 Variant Calling and the Reference Genome

WES yields the sequence information of the portions of the genetic material that are translated to protein, whereas whole genome sequencing (WGS) aims to cover the entire genetic material. Depending on the specific assay used, WES may or may not cover exon-flanking regions as well as 3'- and 5'-untranslated regions (UTRs). The individual reads as well as quality scores for each individual base are stored in fastq files. Variant calling starts with fastq files and is the process of identification of genetic variants in relation to a reference sequence. Currently, GRCh38[124] (hg38) is the most recent and most widely used reference sequence for the human genome. The variant calling workflow consists of pre-processing, alignment of the reads to the reference sequence, variant calling and post-processing/review[125]. Numerous bioinformatic software tools exist to carry out these tasks[125]. Which tools and algorithms are most suitable depends on the employed next-generation sequencing (NGS) technique, the experimental design, the samples and the pursued research interest[125]. Called variants are stored in variant calling formatted (VCF) files. Mutation annotation formatted (MAF) files can be generated from VCF files using custom scripts or conversion tools.

1.6.2 Mutation Annotation Formatted Files

In contrast to VCF files, MAF files require that each variant must be mapped to only one of all possible gene transcripts/isoforms that it might affect. In addition, MAF files only allow a single effect per variant. For example, a single nucleotide variant (SNV) located close to a splice site can be labelled either as 'Missense_Mutation' or as 'Splice_Site' but not as both, while in VCF files several labels are possible. Like VCF files, MAF files are standardized tab-delimited text files designed to store aggregated information about genetic variants. There is an extensive list of possible columns MAF files can contain[126]. Additional custom columns may be added to project-specific files, however, the available software packages that aid in analyzing MAF files depend on a set of about 10 columns providing basic information on the variant.

1.6.3 The HGVS Variant Nomenclature

The Human Genome Variation Society (HGVS) has developed a standardized nomenclature for genetic variants. HGVS IDs[127] consist of a reference sequence and a description separated by a colon. For example, 'chr20:g.43264927C>T' is a valid ID for the human reference genome GRCh37[128]. 'chr20' is the reference sequence and the lower case 'g' indicates that a genomic reference sequence is used. The number 43264927 refers to the position of the variant on chr20 and C>T indicates that a cytosine to thymine transition has occurred. HGVS IDs were used as variant identifiers, such that variants occurring in several samples shared the same ID. The IDs were used to identify shared, partially shared and unique variants in sample groups.

1.6.4 Variant Types and Classifications

Variants types and variant classifications are summarized in **Table 3**. In MAF files, point mutations are abbreviated 'SNPs' (single nucleotide polymorphisms), suggesting that the variant exists with a specific frequency within a population. This is not necessarily true for genetic variants that are a consequence of treatment with mutagenic substances. The term 'SNV' (single nucleotide variant) is a better fit in this case and is used instead of SNP in this work. Furthermore, the term 'InDel' is used to simultaneously refer to insertions and deletions. 'Non-coding variants' are variants that are not located within exons, corresponding to the variant classifications 'IGR', 'Intron', '3'UTR' and '5'UTR' (**Table 3**). 'Synonymous variants' refers to the variant classification 'Silent' (**Table 3**). Finally, 'coding and non-synonymous variants' refers to all variant classifications except 'IGR', 'Intron', '3'UTR', '5'UTR' and 'Silent'.

1.6.5 Limitations to Variant Detection by Whole Exome Sequencing

The central advantages of WES over WGS are its significantly lower costs and the lower requirements in terms of compute and data storage. However, WES comes with drawbacks that have to be considered.

In contrast to WGS, WES generally does not succeed in covering all known exons[129]. This is due to the incomplete capturing of exons by the commercially available assays. This problem cannot be solved by increasing the read depth[129]. Furthermore, certain structural variants (SVs) cannot be detected by WES. Balanced SVs do not change the absolute amount of genetic material (e.g., inversions or translocations of stretches of DNA). The detection of these SVs is generally not possible by WES[130], in particular if long stretches of DNA are affected. Unbalanced SVs change the absolute amount of genetic material and are also referred to as copy number variants (CNVs). CNVs can either be deletions or duplications/amplifications. Small CNVs contained by single reads can be detected by WES. The detection of larger CNVs requires an analysis of the coverage across larger genomic regions. The coverage is the number of reads that got mapped to a certain interrogated gene or genomic region. However, due to capturing and enrichment, the technical uniformity of the coverage is considerably lower in WES than in WGS. The non-uniformity of the coverage makes it necessary to apply non-trivial normalization procedures, making WGS the preferred technique to detect large CNVs[129,130].

Table 3: Variant types and classifications.

Variant Types	Description	Example
SNP	Single nucleotide polymorphism. A base is exchanged for another base. SNV (variant) is an equivalent term.	C > T
DNP	Di-nucleotide polymorphism. Like SNP but two bases.	CC > TT
TNP	Tri-nucleotide polymorphism. Like SNP but three bases.	CCC > TTT
ONP	Oligo-nucleotide polymorphism. Like SNP but affecting more than three consecutive bases.	CCCC > TTTT
MNV	Multi-nucleotide variant. Umbrella term for DNP, TNP & ONP.	See above
DEL	A deleted base or sequence of bases without replacement.	… delTT …
INS	An inserted base or sequence of bases.	… insTT …
Variant Classifications	Description	Example
IGR	Mutation of any variant type located in intergenic regions.	Any type
Intron	Mutation of any variant type located in intronic regions.	Any type
3'UTR	Mutation of any variant type located in 3'UTRs.	Any type
5'UTR	Mutation of any variant type located in 5'UTRs.	Any type
Silent	Mutation of type SNP or MNV located in an exon without changing the amino acid sequence of the resulting protein.	GGG > GGA Glycine > Glycine
Missense_Mutation	Mutation of type SNP or MNV resulting in the change of one or several consecutive amino acids in the resulting protein.	GGA > AGA Glycine > Arginine
In_Frame_Insertion	Insertions of base pair fragments of length n × 3. Not resulting in a frame shift.	… insGGG … Glycine inserted
In_Frame_Deletion	Deletions of base pair fragments of length n × 3. Not resulting in a frame shift.	… delGGG … Glycine deleted
Frame_Shift_Insertion	Insertions of base pair fragments of length n × m, where m modulo 3 ǂ 0. This results in a frame shift.	… insGG …
Frame_Shift_Deletion	Deletions of base pair fragments of length n × m, where m modulo 3 ǂ 0. This results in a frame shift.	… delGG …
Splice_Site	Mutation of any type located near a splice site	Any type
Nonstop_Mutation	Mutation of any type causing the loss of a wild-type stop codon.	TAA > TAC Stop > Tyrosine
Nonsense_Mutation	Mutation of any type introducing novel stop codon.	TAC > TAA Tyrosine > Stop
Translation_Start_Site	Mutation of any type causing the loss of a wild-type start codon.	ATG > ACG Start > Threonine

1.6.6 The Variant Allele Frequency

The variant allele frequency (VAF) is calculated by dividing the number of reads in which the variant was identified (the count) by the total number of reads that covered the particular region (the coverage). A coverage of 25 and a count of 12 would equate to a VAF of 48%. In homogeneous samples with known ploidy, the zygosity can be inferred from the VAF.
For clonally heterogeneous samples with unknown ploidy, any VAF lower than 100% must be cautiously interpreted. The PCa cell lines studied herein are highly aneuploid. PC-3, DU145, LNCaP and LAPC-4 cells have modal chromosomal numbers of 58, 64, 76 and 89, implying that the average gene has 2.5, 2.8, 3.3 and 3.8 copies in these cell lines, respectively (section **1.5**). Some chromosomes may have two, while others may have three or more copies. VAFs around 33% and 66% could indicate that one or two out of three genetic copies carry a certain variant. However, robust conclusions cannot be drawn from the VAF in the context of bulk sequencing of aneuploid cancer cell lines. It is best to think about the VAF as an estimate of the prevalence of the variant in the genetic material of the cell population.

1.6.7 Mutational Signatures

Mutational signatures were developed by Alexandrov et al.[131,132] in 2013. There are internal (e.g., defective DNA repair) and external (e.g., smoking) mutagenic processes. The most common types of DNA damage induced by such processes can be determined. For example, C:G>A:T transversions are particularly common in smoking-associated lung cancer[132]. The idea of mutational signatures is to infer underlying mutational process(es) from the prevalence of certain types of mutations. Albeit the concept has been extended[133], mutational signatures are based on SNVs in their original version[131,132]. For that matter, SNVs are classified into six subtypes: C:G>A:T, C:G>G:C, C:G>T:A, T:A>A:T, T:A>C:G, and T:A>G:C. Additionally, the 3'- and 5'-adjacent bases are taken into account, resulting in 96 possible single base substitution (SBS) events (6 types of substitution × 4 types of 5' bases × 4 types of 3' bases). The fractions of these 96 SBS events plotted as a histogram make the mutational signature (**Figure 6**). Several mutational processes can be simultaneously operative. A single mutational signature will be a mixture of several sub-signatures where every sub-signature corresponds to one particular mutational process[132]. The primary signature can be unmixed, however, the mathematical method requires a minimum of ~20 samples with at least a few hundred SNVs per sample[132].

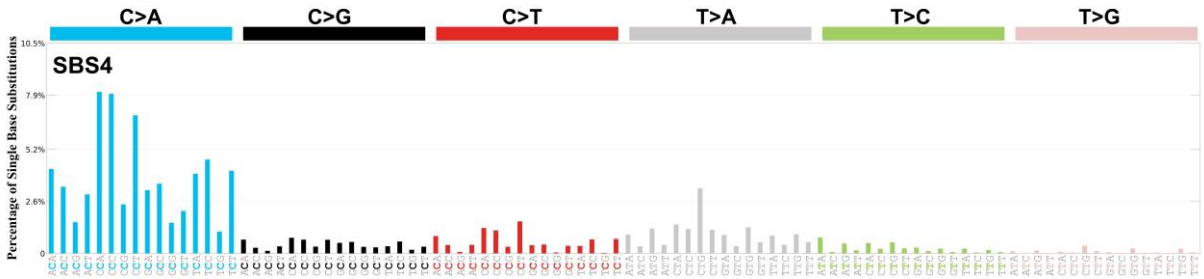

Figure 6: The mutational signature SBS4. The proposed etiology of this signature is 'Tobacco Smoking'[134]. Note the abundance of C>A (C:G>A:T) transversions relative to SBS events of other subtypes. Within each subtype, there are 16 different possible SBS events depending on the 3'- and 5'-adjacent bases.

2 Aims and Objectives

2.1 The Study of Genetic Diversity among PCa Cell Lines

Within this work, WES of 24 PCa cell lines was performed. The 24 cell lines consisted of five groups, namely eight PC-3, six DU145, five LNCaP, three LAPC-4 and two LNCaP derivative cell lines. The cell lines of one particular group (e.g., the eight PC-3 cell lines) originated from different laboratories and encompassed treatment-naïve and age-matched treatment-resistant derivative cell cultures. A faithful assumption that is necessary to compare studies done with cancer cell lines is that the cell lines used in different laboratories are largely isogenic. One goal of this work was to challenge this assumption and to study putative genetic diversity within the above-mentioned cell line groups. Furthermore, genetic comparisons of parental treatment-naïve cell lines and their direct and age-matched treatment-resistant sublines were done.

2.2 Implications of the Previous Treatment History for CDDP Tolerance

This work studied whether previous treatment with either DTX or ionizing radiation influences the response to CDDP treatment. As elaborated in section **1.3.6**, Pt compounds such as CDDP are an effective albeit short-lived treatment option for patients with NEPC. However, NEPC is a late-stage phenomenon and when diagnosed, patients usually have already been subjected to multiple prior treatment modalities including DTX and/or ionizing radiation. The treatment history may influence the efficacy of CDDP. The diversity of prior treatment modalities paired with inhomogeneous patient genetics and the small pool of eligible patients makes it hard to conduct clinical studies to address this question. This work utilized DTX-resistant and radio-resistant DU145 and PC-3 cell lines as well as their age-matched parental control cells. The cells were treated with CDDP *in vitro*. Their CDDP sensitivity was assessed and compared by measurements of growth rates, viability, apoptosis, metabolic activity and colony formation ability.

2.3 The Role of ABCB1 in Docetaxel-resistant DU145 and PC-3 Cells

As described in section **1.4.1.3**, the transmembrane transporter ABCB1 plays a pivotal role for DTX resistance in many cell culture models. This work studied the relevance of ABCTs, particularly of ABCB1 in DTX-resistant DU145 and PC-3 cells. To this end, the expression of ABCTs was analyzed. ABCB1 was ablated by several approaches and the effect of ABCB1 ablation on DTX resistance was evaluated. Beyond that, a second objective was to study if and to what extend the glycosylation of ABCB1 was important to maintain the DTX resistance in DTX-resistant DU145 and PC-3 cells. The ABCB1 project was initiated and conceptualized by Mr. Lukas Donix (the author of this dissertation). Consecutive experiments were at large conducted by Ms. Dinah Linke within her master thesis under the practical supervision of Mr. Lukas Donix (section **11** on page **142**).

3 Materials

Material and equipment manufacturers and suppliers are abbreviated in this section. For their full names and descriptions please refer to **Appendix 1** on page **140**.

3.1 Cell Lines

Table 4: List of Cell Lines. The table lists the cell lines used in this work. The origins of the respective cell lines are letter-coded A-D as follows. Cell lines labelled "A" (except LAPC-4^{A}) originate from the research laboratory of Prof. Dr. Susanne Füssel at the Dept. of Urology of the University Hospital Carl Gustav Carus in Dresden, Germany. This site was the primary site where all experimental work was performed. The resistant sublines were established by Dr. Doreen Hübner and are unpublished material as of the time of writing. LAPC-4^{A} originate from the laboratory of (retired) Prof. Dr. Günter Vollmer, which belonged to the Section of Natural Sciences (Dept. of Biology) of the Technische Universität Dresden in Dresden, Germany. Cell lines labelled with "B" originate from the laboratory of Prof. Zoran Culig at the Dept. of Experimental Urology of the Innsbruck Medical University in Innsbruck, Austria. The respective cells were transferred to the primary site by Dr. Holger H. H. Erb in 2018. Cell lines labelled "C" originate from the laboratory of Prof. Dr. Anna Dubrovska at the Center for Radiation Research in Oncology in Dresden, Germany. Cell lines labelled "D" were kindly provided by Dr. Martin Michaelis and Prof. Dr. Jindrich Cinatl and are part of the Resistant Cancer Cell Line (RCCL) collection[135], which is based in the Industrial Biotechnology Center at the University of Kent in Kent, UK. C4-2[114] and MR49F EnzaR[116] cells are not letter-coded and originate from the respective laboratories in which they were derived from LNCaP cells (see publications). These two cell lines were transferred to the primary site by Dr. Holger H. H. Erb in 2018. MTA: material transfer agreement, DE: Dose escalation, CDDPR: CDDP-resistant, CR: castration-resistant, DTXR: DTX-resistant, BicaR: bicalutamide-resistant, EnzaR: enzalutamide-resistant, RR: radio-resistant. Standard culture conditions are described in section **4.1**.

Internal name	Parent	Resistance	Culture conditions	Purchased from / MTA
DU145^{A}	-	-	standard	ATCC in 1996
DU145^{A} CDDPR	DU145^{A}	CDDPR. Established ~2009 by DE.	+0.5 µg/mL CDDP	-
PC-3^{A}	-	-	standard	ATCC in 1996
PC-3^{A} CDDPR	PC-3^{A}	CDDPR. Established ~2009 by DE.	+0.5 µg/mL CDDP	-
LNCaP^{A}	-	-	standard	LGC in 2004
LAPC-4^{A}	-	-	standard	MTA, Vollmer, 2014
DU145^{B}	-	-	standard	MTA, Erb, 2018
DU145^{B} DTXR	DU145^{B}	DTXR. Established ~2012 by DE[85].	+10 nM DTX	MTA, Erb, 2018
PC-3^{B}	-	-	standard	MTA, Erb, 2018
PC-3^{B} DTXR	PC-3^{B}	DTXR. Established ~2012 by DE[85].	+10 nM DTX	MTA, Erb, 2018
LNCaP^{B}	-	-	standard	MTA, Erb, 2018
LNCaP^{B} BicaR	LNCaP^{B}	BicaR. Established ~2006 by prolonged exposure to bicalutamide[136].	charcoal stripped FBS +0.01 nM R1881 +10 nM bicalutamide	MTA, Erb, 2018

(Continued on next page.)

Continuation of Table 4: List of Cell Lines.

Internal name	Parent	Resistance	Culture conditions	Purchased from / MTA
LAPC-4[B]	-	-	DMEM	MTA, Erb, 2018
LAPC-4[B] EnzaR	LAPC-4[B]	EnzaR. Established ~2016 by DE[137].	DMEM +10 µM enzalutamide	MTA, Erb, 2018
C4-2	LNCaP	CR. Established ~2000[114].	DMEM, low glucose	MTA, Erb, 2018
MR49F EnzaR	LNCaP	CR and EnzaR. Established ~2013[116].	+10 µM enzalutamide	MTA, Erb, 2018
DU145[C]	-	-	standard	MTA, Dubrovska, 2018
DU145[C] RR	DU145[C]	RR. Established ~2015 X-ray treatment[138].	standard, max. six weeks	MTA, Dubrovska, 2018
PC-3[C]	-	-	standard	MTA, Dubrovska, 2018
PC-3[C] RR	PC-3[C]	RR. Established ~2015 X-ray treatment[138].	standard, max. six weeks	MTA, Dubrovska, 2018
LNCaP[C]	-	-	standard	MTA, Dubrovska, 2018
LNCaP[C] RR	LNCaP[C]	RR. Established ~2015 by X-ray treatment[138].	standard, max. six weeks	MTA, Dubrovska, 2018
PC-3[D]	-	-	standard	MTA, Michaelis/Cinatl, 2019
PC-3[D] CDDPR	PC-3[D]	CDDPR. Established by DE.	+ 0.5 – 2 µg/mL CDDP	MTA, Michaelis/Cinatl, 2019

3.2 Chemicals, Reagents and Kits

Table 5: List of chemicals, reagents and kits.

Product	Manufacturer
7-AAD (7-Aminoactinomycin D)	Biolegend
Accutase	Sigma
Antioxidant (NuPAGE)	Thermo Fisher
Aqua dest.	Fresenius Kabi
BCA Protein Assay Kit (Pierce™)	Thermo Fisher
Crystal Violet aqueous solution (1%)	Sigma
dNTPs (10 mM)	Promega
Docetaxel (DTX)	Selleck
Direct-zol™ RNA Miniprep RNA isolation kit	Zymo Research
Dithiothreitol (DTT, 0.1 M)	Thermo Fisher
Dulbecco's Modified Eagle Medium (DMEM) (1X)	Thermo Fisher
Elacridar	APExBIO
Ethanol (96%, methylated)	Berkel AHK
Ethanol (absolute)	VWR
Fetal Bovine Serum (FBS)	Thermo Fisher
First-Strand Buffer (5X) for cDNA synthesis	Thermo Fisher
GoTaq Probe qPCR Mastermix (2X)	Promega
Hoechst 33342 (ab228551)	Abcam
iBlot® Gel Transfer Stacks Nitrocellulose, Regular	Thermo Fisher
Lipofectamin 2000	Thermo Fisher
Non-Fat Dry Milk (NFDM) powder	Heirler
β-Mercaptoethanol	Sigma
Methanol absolute	AppliChem
MES SDS Running Buffer (20X)	Thermo Fisher
MOPS SDS Running Buffer (20X)	Thermo Fisher
MUSE Count & Viability-Kit	Luminex
MUSE Multi Caspase kit	Luminex
Mycoalert™ Mycoplasma Detection kit	Lonza
NP-40 (TERGITOL™)	Sigma
NucLight Rapid Red	Sartorius
NuPAGE LDS sample buffer (4X) for Western Blotting	Thermo Fisher
OptiMEM	Thermo Fisher
PBS (Dulbecco's phosphate buffered saline, 1X)	Sigma
pd(N)6 Random Primers (0.2 µg/µL)	Roche

Continuation of Table 6: List of chemicals, reagents and kits.

Product	Manufacturer
Phosphatase inhibitor cocktail	Sigma
PNGase F	Promega
Ponceau S	Sigma
Polyacrylamid gels (NuPAGE Novex Bis-Tris, 4-12%)	Thermo Fisher
PowerPlex 18D System kit	Promega
Protease inhibitor cocktail	Santa Cruz
Protein marker Novex Magic Mark XP	Thermo Fisher
Protein marker Spectra Multicolor Broad Range	Thermo Fisher
ReliaPrep™ genomic DNA Tissue Miniprep System kit	Promega
Rhodamin 123	J&K Scientific
Roswell Park Memorial Institute (RPMI) Medium 1640 (1X)	Thermo Fisher
RT^2 SYBR Green master mix	Thermo Fisher
Sodium chloride (NaCl)	Carl Roth
Sodium citrate	Riedel-de Haën
Sodium deoxycholate	Sigma
Sodium dodecyl sulfate (SDS)	Carl Roth
SureSelectXT HS Human All Exon V5 + UTRs for enrichment-based WES	Agilent
SuperScript II Reverse Transcriptase	Thermo Fisher
Tariquidar	APExBIO
TriFast peqGOLD	PeqLab
Tris(hydroxymethyl)aminomethan (TRIS)	Carl Roth
TRIS-HCl	Carl Roth
Triton X-100	Sigma
Trypsin EDTA (10X)	Sigma
Tunicamycin	Merck Millipore
Tween 20	Sigma
Verapamil	Sigma
WesternBright Sirius chemiluminescence kit	Advansta
WST-1 cell proliferation reagent	Roche

3.3 Laboratory Consumables

Table 6: List of laboratory consumables.

Product	Manufacturer
0.2 mL reaction tubes for PCR	Sarstedt
0.5, 1 and 2 mL reaction tubes	Eppendorf, Sarstedt
10 cm culture dishes	Greiner
15 and 50 mL tubes	Greiner
6-well plates	Corning, TPP
96-well plates	Corning, Greiner, TPP
96-well plates, µCLEAR®, black (for fluorescence)	Greiner
96-well plates, thin, V, white (for LightCycler 480 instrument)	Biozym
Cell culture flasks 75 and 175 cm^2, CELLSTAR®	Greiner
Cell scraper	Corning
Combitips Advanced (0.5, 2.5 and 5 mL)	Eppendorf
Filtropur S 0.45 µm syringe filters	Sarstedt
Gloves latex micro touch	Ansell
Gloves purple nitrile extra	Halyard
Pipette tips (20 µL, 5000 µL)	Sarstedt
Pipette tips (100 µL, 1000 µL)	Greiner
Pipette tips with filter, low binding (1000 µL)	Greiner
Pipette tips with filter, XL (1000 µL)	Sarstedt
Pipette tips with filter, Maxymum Recovery (10 µL, 100 µL)	Corning
Pipette tips with filter, SurPhob (10 µL)	Biozym
Reaction tubes with screw top (1.5 mL)	Sarstedt
Serological pipettes for cell culture (5 mL, 10 mL, 25 mL, 50 mL)	Corning
Sterile pads	Fink & Walter

3.4 Laboratory Devices

Table 7: List of laboratory devices.

Product	Manufacturer
ABI 3500 Genetic Analyzer	Thermo Fisher
Autoclave Systec VX-75 and VE-55	Systec
BD Celesta flow cytometer	BD
Chemiluminescence scanner, MicroChemi 4.2	Biostep
Clean bench LaminAir Model 1.2	Heto-Holten
CO_2 Incubator, Heracell 240i	Thermo Fisher
Electronic multi-channel pipette, Research Pro 1200 µL	Eppendorf
Electrophoresis chamber, XCell SureLock Mini-Cell	Thermo Fisher
EVOS XL Core Microscope	Thermo Fisher
iBlot and iBlot2 Dry Blotting-System	Thermo Fisher
iBlot and iBlot2 Gel transfer stacks Nitrocellulose	Thermo Fisher
Live Cell Imaging Microscope IncuCyte S3	Sartorius
LightCycler 480 Instrument II	Roche
Multiple dispenser Handy Step	Brand
Multi-channel pipette 50-100 µL	VWR
Microliter pipette Research Plus (0.5-10, 10-100 and 100-1000 µL)	Eppendorf
MUSE Cell Analyzer	Luminex
NanoDrop 2000c Spectrophotometer	Thermo Fisher
NextSeq 500 Sequencer	Illumina
PCR Workstation Pro	PeqLab
Pipetboy acu	Integra
Plate reader Mithras LB 940	Berthold
Refrigerated centrifuge 5415 R	Eppendorf
Scale Kern 510	Kern & Sohn
Table centrifuge MiniSpin	Eppendorf
Thermocycler PTC-100	MJ Research
ThermoMixer C	Eppendorf

3.5 Buffers and Solutions

Table 8: List of buffers and solutions.

0.1% Crystal violet staining solution	
50 mL	Crystal violet solution (1%)
450 mL	Aqua dest.
1X MES/MOPS SDS running buffer	
50 mL	20X MES/MOPS SDS running buffer, respectively
950 mL	Aqua dest.
10X TBS	
585.4 g	NaCl
8.48 g	Tris
52.0 g	Tris-HCl
dissolve in 2 L	Aqua dest.
1X TBS	
100 mL	10X TBS
900 mL	Aqua dest.
1X TBS-T	
1000 mL	1X TBS
1 mL	Tween 20
1% MMP/TBS-T	
0.5 g	MMP
dissolve in 50 mL	TBS-T
Filtered through Filtropur S 0.45 µm filters before use.	
10% MMP/TBS-T	
5 g	MMP
dissolve in 50 mL	TBS-T
0.1 M Sodium citrate solution	
29 g	Sodium citrate
dissolve in 1000 mL	50% Ethanol
EDTA stock solution (100 mM)	
3.72 g	EDTA
100 mL	PBS
Flow cytometry staining buffer (FCS-buffer)	
1 mL	EDTA stock solution (100 mM)
1 mL	HEPES (1 M)
5 mL	FBS
93 mL	PBS
+ 0.5 µg/mL	7-AAD (optional depending on staining)
+ 0.1 µg/mL	Rhodamin 123 (optional depending on staining)

Continuation of Table 9: List of buffers and solutions.

RIPA buffer (final concentrations in Aqua dest.)	
25 mM	TRIS-HCl, pH 7.5
1% (v/v)	NP-40 (TERGITOL™)
0.5% (v/w)	Sodium deoxycholate
0.1% (v/w)	SDS
150 mM	NaCl
2 mM	EDTA
1% (v/v)	Triton X-100
1%	Protease inhibitor cocktail (added immediately before use)
1%	Phosphatase inhibitor cocktail (added immediately before use)
RPMI 1640 for cell culture	
450 mL	RPMI 1640 (1X)
50 mL	FBS
Trypsin/EDTA solution	
100 mL	10X Trypsin/EDTA
900 mL	Aqua dest.

3.6 TaqMan Probes

Table 9: List of TaqMan probes used in qPCR measurements. All were 20X concentrated and manufactured by Thermo Fisher. All assays are also found under catalog number 4331182.

Gene	Assay ID	Gene	Assay ID	Gene	Assay ID
ABCA3	Hs00184543_m1	BRCA1	Hs00173237_m1	PTEN	Hs02621230_s1
ABCA8	Hs00992362_m1	BRCA2	Hs00609073_m1	RRM1	Hs01040698_m1
ABCB1	Hs00184500_m1	CDH1	Hs01023895_m1	RRM2	Hs00357247_g1
ABCB11	Hs00994811_m1	CDH2	Hs00983056_m1	SNAI1	Hs00195591_m1
ABCB4	Hs00983957_m1	CTDSP1	Hs01105502_m1	SNAI2	Hs00161904_m1
ABCC1	Hs01561483_m1	CTNNB1	Hs00355049_m1	TIMP3	Hs00165949_m1
ABCC10	Hs01056200_m1	EGFR	Hs01076090_m1	TUBB3	Hs00964962_g1
ABCC2	Hs00960489_m1	ERCC1	Hs01012157_m1	TWIST1	Hs00361186_m1
ABCG2	Hs01053790_m1	ERCC2	Hs00361161_m1	VEGFA	Hs00900055_m1
AMACR	Hs01091292_m1	ERCC5	Hs01557031_m1	VIM	Hs00185584_m1
AR	Hs00171172_m1	EZH2	Hs00544830_m1	XIAP	Hs00236913_m1
ATM	Hs00175892_m1	FANCC	Hs00984545_m1	ZEB1	Hs01566408_m1
ATP7A	Hs00163707_m1	GSK3B	Hs01047719_m1	ZEB2	Hs00207691_m1
ATP7B	Hs01075310_m1	HIF-1a	Hs00153153_m1	**Reference**	**Assay ID**
ATR	Hs00992123_m1	MLH1	Hs00179866_m1	HPRT1	Hs02800695_m1
BCL2	Hs01048932_g1	MTF1	Hs00232306_m1	RPLP0	Hs00420895_gH
BIRC5	Hs00977611_g1	PSA	Hs02576345_m1		
BMI1	Hs00180411_m1	PSMA	Hs00379515_m1		

3.7 RT^2 Profiler PCR Arrays

Table 10: List of RT^2 profiler PCR arrays used in qPCR measurements.

Targeted Pathway(s)	Cat. #	Assay ID	Manufacturer
Epithelial to Mesenchymal Transition	330231	PAHS-090ZA	Qiagen
DNA Repair	330231	PAHS-042ZA	Qiagen

3.8 Antibodies

3.8.1 For Western Blotting

Table 11: List of antibodies used in western blotting. Ms: mouse, Rb: rabbit.

Primary Antibodies					
Target Protein	**Type**	**Manufacturer**	**Host Animal**	**Clone**	**Dilution**
ABCB1/MDR1	Monoclonal, IgG	CST	Rb	D3H1Q	1 : 1,000
ABCG2/BCRP	Monoclonal, IgG	CST	Rb	D5V2K	1 : 1,000
ABCC1/MRP1	Monoclonal, IgG1	Abcam	Ms	IU2H10	1 : 1,000
AR	Monoclonal, IgG	CST	Rb	D6F11	1 : 2,000
CDH1	Monoclonal, IgG	CST	Rb	24E10	1 : 1,000
CDH2	Monoclonal, IgG	Abcam	Rb	EPR1791-4	1 : 1,000
MLH1	Monoclonal, IgG1	CST	Ms	4C9C7	1 : 1,000
PSA	Monoclonal, IgG	CST	Rb	D6B1	1 : 1,000
PTEN	Monoclonal, IgG	CST	Rb	138G6	1 : 1,000
Vimentin	Monoclonal, IgG	CST	Rb	D21H3	1 : 1,000
Primary Antibodies, Reference Proteins					
Alpha-Tubulin	Monoclonal, IgG1	Sigma	Ms	DM1A	1 : 5,000
GAPDH	Monoclonal, IgG1	Acris	Ms	6C5	1 : 5,000
Secondary Antibodies, Horseradish Peroxidase Conjugated					
Rb	Polyclonal	Dako	Pig	P0217	1 : 2,000 - 10,000
Ms	Polyclonal	Dako	Rb	P0260	1 : 2,000 - 10,000

3.8.2 For Flow Cytometry

Table 12: List of antibodies used in flow cytometry. One test is equivalent to approximately 500,000 cells. PE: Phycoerythrin (fluorophor), Ms: Mouse.

Primary Antibodies					
Target Protein	**Conjugate / Type**	**Manufacturer**	**Host Animal**	**Clone / Cat. #**	**Dilution**
ABCB1	PE / IgG2a κ	Biolegend	Ms	UIC2 / 348606	5 µL per Test
ABCG2	PE / IgG2b κ	Biolegend	Ms	5D3 / 313620	5 µL per Test
Isotype Control Antibodies					
-	PE / IgG2a κ	Biolegend	Ms	MPC-11 / 400313	5 µL per Test
-	PE / IgG2b κ	Biolegend	Ms	MOPC-173 / 400213	2.5 µL per Test

3.9 siRNAs

Table 13: List of siRNAs.

Internal name	Description	Concentration of stock	Manufacturer	Cat. #
siR-CON	Stealth™ siRNA Negative Control, Low GC	20 µM	Thermo Fisher	12935200
siR-MLH1	Stealth™ siRNA against MLH1	20 µM	Thermo Fisher	HSS106567
siR-ABCB1	Stealth™ siRNA against ABCB1	20 µM	Thermo Fisher	HSS182278

3.10 Control DNA for Cell Line Authentication

Table 14: Control DNA for cell line authentication.

Cell Line	Product	Cat. #	Manufacturer
DU145	DU145 gDNA, 500 ng	300168GD05	CLS
PC-3	LNCaP gDNA, 500 ng	300265GD05	CLS
LNCaP	PC-3 gDNA, 500 ng	300312GD05	CLS

3.11 Software

3.11.1 Programming Languages

R (https://www.r-project.org) was used in version 4.1.2 and Python (https://www.python.org) in version 3.9, respectively. Writing and editing of R code was done in Rstudio 2021.09.1 (https://www.rstudio.com) and writing and editing of Python code was done in PyCharm Community Edition v2021.3.1 (https://www.jetbrains.com/pycharm). R, Python, Rstudio and PyCharm Community Edition are open-source software entities under GNU General Public Licenses (R, and Rstudio), Apache 2 License (PyCharm) and OSI – Python License (Python), respectively.

3.11.2 For Generation of Graphs and Compilation of Figures

Graphs and individual panels were generated using R and/or Python and/or GraphPad Prism 9 (GraphPad Software, San Diego, CA, USA). Panels and figures were arranged using either Inkscape 1.2 (https://inkscape.org) or Microsoft Powerpoint 2016 (Microsoft, Redmond, WA, USA). Inkscape is an open-source software entity under GNU General Public License. GraphPad Prism 9 and Microsoft Powerpoint 2016 are proprietary software entities developed and distributed by GraphPad Software and Microsoft, respectively.

3.11.3 For Tabular Data Analysis and Statistics

For tabular data analysis tasks for which the use of R or Python was not necessary, Microsoft Excel 2016 and/or Graphpad Prism 9 were used. Microsoft Excel 2016 (Microsoft) is a proprietary software developed and distributed by Microsoft. All statistical tests were done using GraphPad Prism 9.

3.11.4 Analysis of Flow Cytometry Data

Flow cytometry data were analyzed and graphed using FlowJo cell analysis software, a proprietary software developed by Becton Dickinson Inc. (Becton Dickinson, Franklin Lakes, New Jersey, USA).

3.11.5 Densitometric analysis of western blot images

Densitometric analysis of western blot images was done using Fiji in version 2.1.0, which is an open-source software entity developed by Wayne Rasband and contributors. It is available under the GNU General Public License.

3.11.6 For Word Processing and Citation Management

The writing and formatting of this work took place using Microsoft Word 2016 (Microsoft) as a word processor and EndNote X9 (Clarivate, Boston, MA, USA) for managing citations. Microsoft Word 2016 and EndNote X9 are proprietary software entities developed and distributed by Microsoft and Clarivate, respectively.

4 Methods

4.1 Culture of Cell Lines

For a list of cell lines and a description of their properties please refer to **Table 4**. If not otherwise mentioned in **Table 4**, all cells were cultured under standard conditions (37°C, humidified atmosphere with 5% CO_2) and received RPMI 1640 medium supplemented with 10% FBS. DU145 and PC-3 cell lines were split three times a week. For splitting, the medium was removed, the monolayer was washed twice with PBS and an aqueous solution containing 0.05% trypsin/EDTA was used to detach adherent cells. Cell counting was done using the MUSE Count & Viability kit on a MUSE Cell Analyzer according to the manufacturer's instructions.

4.2 Cell Line Authentication and Contamination Control

Prior to extraction of genomic DNA for WES and other experimentation, all cell lines were authenticated by satellite tandem repeat (STR) analysis (see **Table 15** in the Results section). STR profiling was performed at the Institute of Legal Medicine (TU Dresden). For STR analysis, cell line DNA was amplified using the PowerPlex 18D System kit according to manufacturer's instructions. Capillary electrophoresis was then performed with an ABI 3500 Genetic Analyzer and analyzed by GeneMapper ID-X software version 1.4 (Thermo Fisher). The STR profiles were compared to expected profiles as documented in the cellosaurus database[139]. Additionally, the profiles of DU145, PC-3 and LNCaP cells were compared to commercially available control DNA (**Table 14**). For LAPC-4, C4-2 and MR49F cells, control DNA was not available. Furthermore, the cellosaurus database listed only an incomplete expected profile for C4-2 cells. For MR49F cells, no data on STR markers was found. However, both C4-2 and MR49F cells are derivatives of LNCaP and hence, a profile closely resembling LNCaP cells can be expected.
Furthermore, all cells were regularly tested for contamination with mycoplasma using the Mycoalert™ Mycoplasma Detection kit according to manufacturer's instruction.

4.3 Microsatellite Instability Testing

The treatment-naïve parental cell lines DU145^{A}, DU145^{B}, DU145^{C}, PC-3^{A}, PC-3^{B}, PC-3^{C}, and PC-3^{D} were subjected to MSI-testing (**Table 20**) at the Institute of Pathology (TU Dresden). The Bethesda panel recommended by the National Cancer Institute was used for MSI-testing. This panel consists of five mononucleotide repeat markers (BAT-25, BAT-26, NR-21, NR-24 and MONO-27) for the identification of samples with MSI as well as two pentanucleotide repeat markers (Penta C and Penta D) to prevent sample mix-ups and/or contamination[140].

4.4 Whole Exome Sequencing of Cell Lines

The total collection of 24 cell lines (**Table 4**) was subjected to WES and subsequent *in silico* data analysis. For extraction of genomic DNA, the ReliaPrep™ genomic DNA Tissue Miniprep System kit was used according to manufacturer's instructions. Library preparation and target enrichment was performed using the SureSelectXT HS Human All Exon V5 + UTRs kit, according the manufacturer's instruction. Sequencing was performed on a NextSeq 500 sequencer with an average coverage of 150 reads per target. Resulting fastq files were processed using the CLC Genomics Workbench v12. Software. Reads were aligned to the human reference genome GRCh37 (hg19)[128] and variants were called without assumption of known sample ploidy (low frequency variant detection algorithm), a required minimum variant read count of three, a minimum coverage of 10 reads and a minimum VAF of 10%. Variants in homopolymer regions longer than three nucleotides were removed. Furthermore, variants with a minor allele frequency of at least 1% in the dbSNP database 'dbSNP common v151'[141] were considered reasonably common in the general population (i.e., likely not tumor-associated) and thus excluded from further analysis. Called variants were annotated using the CLC Genomics Workbench v12 software as well as the myVariant.info API[142] through custom Python scripts. The resulting data was stored as a MAF file for further analysis.

4.4.1 Calculation of mutational burden

Only non-synonymous variants located within coding sequences (N = 110,059) were considered for the calculation of the mutational burden. The absolute number of such variants per cell line was divided by the sum of the length of all exons in the human exome and then multiplied by 1000 to calculate the implied variants / kbp (**Table 16**) for each cell line. The length of the human exome (1.887×10^7 bp) was calculated from data provided by the Consensus Coding Sequence Project (CCDS, 'CCDS_exons.current.txt' release #22, last updated 2022-08-24)[143] using a custom Python script. The CCDS data only includes exons that are consistently documented and annotated across major databases and therefore provides a conservative estimate of the length of the human exome, which in turn may lead to a certain degree of overestimating the mutational burden of the cell lines. Other sources estimate the length of the human exome to be up to 3×10^7 bp[144]. Furthermore, the fraction of genes with variants (**Table 16**) was calculated by dividing the number of genes with variants by 19,030, which is the total number of unique genes in the CCDS release #22.

4.4.2 Top Mutated Genes and Gene Set Enrichment Analyses

To estimate the variants per kbp for single genes, the total number of variants mapping to the respective gene was divided by the length of the genes longest transcript (LLT) and then multiplied by 1000. Only coding and non-synonymous variants were considered. The LLTs were derived from Ensemble data[145] downloaded from ensembl.org via BioMart. From the 'Ensemble Genes 107' database the dataset 'Human genes (GRCh38.p13)' was selected. For the purpose of estimating the variants per kbp, small differences in terms of exon lengths present between GRCh38 and GRCh37 may be tolerable. From the structures 'GENE' and 'EXON', the following data was selected: 'Ensembl Gene ID', 'Ensembl Transcript ID', 'Associated Gene Name', 'Ensembl Exon ID', 'Exon Chr Start (bp)', 'Exon Chr End (bp)', 'Exon Rank in Transcript'. Records without an associated HGNC ID (Hugo_Symbol)[146] were removed. A custom Python script was used to extract the LLT for each gene present in the downloaded file. Very small genes with LLTs less than 100 bp were removed from the resulting data because per kbp values calculated for such genes are heavily skewed by single mutations. The following formula was used to calculate the group average variants per kbp for each gene:

$$Group\ average\ variants\ per\ kbp = \frac{N_{total\ variants\ in\ group}}{N_{cell\ lines\ in\ group} \times LLT} \times 1000$$

To test for gene set enrichments, ordered queries with preset parameters of the top 15 mutated genes were done using the gProfiler webtool[147].

4.4.3 Unique, Partially Shared and Shared Variants

In order to identify unique, partially shared and shared variants within cell line groups, variants were annotated with HGVS IDs[127] using a custom Python script. Utilizing HVGS IDs as variant identifiers, a custom Python script was used to count and summarize unique, partially shared and shared variants across cell line pairs, parental cell lines and whole cell line groups, respectively. The numbers of such variants among cell line pairs and parental cell lines (**Figure 8** and **Figure 9**) were summarized using Venn diagrams, which were generated using the *pyvenn* package for Python available on GitHub (https://github.com/tctianchi/pyvenn)[148]. Venn diagrams exist for arbitrary numbers of sets[149]. However, because Venn diagrams for all DU145 cells (N=6) or all PC-3 cells (N=8) would be difficult to interpret visually, the UpSetR package[150] for R was used to summarize unique, partially shared and shared variants among complete cell line groups (**Figure 10**). Unless otherwise mentioned, the terms 'unique', 'partially shared' and 'shared' variants, refer only to such variants that were identified based on the analysis of complete cell line groups (**Figure 10**).

4.4.4 Phylogenetic Tree Inference

The Analyses of Phylogenetics and Evolution (ape)[151] package for R was used to infer phylogenetic trees from WES data. ape's bionj() function takes a distance matrix as an input. Distance matrices for cell line groups were prepared as follows. First, a binary genotype matrix (cell lines in columns, variants in rows) was generated. In this matrix, a one represents the presence and a zero represents the absence of a variant, respectively. HGVS IDs served as variant identifiers (rownames). Variants with a VAF ≤ 33% were discarded, however, partially shared and shared variants were preserved if the variant had a VAF ≥ 33% in at least one of the respective cell lines. Otherwise, all variants regardless of type or classification were considered. Next, the genotype matrix was sliced by cell line groups and R's dist() function (method = 'binary') was used to generate a distance matrix for each respective slice. Finally, the distance matrices were used as inputs for ape's bionj() function, which implements a neighbor joining algorithm developed for phylogenetic analyses[152]. The results were plotted using ape's plot.phylo() function (type = 'phylogram') (**Figure 11**).

4.4.5 Mutational Signatures

Mutational signatures[131,132] were extracted using the 'maftools' package[153] for R. All SNVs including synonymous and intronic SNVs are considered for signature calculation. Albeit several mutational processes may be active in one sample simultaneously, only one signature (the composite signature) per sample was extracted. This was done because of the limited number of cell lines within the respective cell line groups (N_{max}=8 for PC-3 cell lines), which is insufficient for effective unmixing of the composite signature[132]. Signatures were extracted for individual samples (e.g., PC-3^{B} DTXR) and for cell line groups (e.g., all DU145 cell lines). Furthermore, signatures were extracted considering only unique or only shared variants, respectively. Similarity between signatures was measured by calculating their cosine similarity[132], where a value of one represents perfect identity, and zero represents perfect disparity, respectively. The extracted signatures were compared to validated and annotated signatures listed by COSMIC[154,155] and the best matching signature (i.e., the signature with the highest cosine similarity) was identified.

4.5 Cisplatin Treatment of Cell Lines in Microwell Plates

For selected cell lines, the tolerance for CDDP was determined by treatment with serial dilutions followed by various readouts. 800 DU145 or 2,000 PC-3 cells per well were seeded in 96-well plates. The cells were treated with 0.01 – 1.0 µg/mL CDDP 24 h after seeding and then cultured in the presence of CDDP for six days before performing a WST-1 assay (section **4.6**) and a crystal violet assay (section **4.7**). During the entire experiment, the cell confluence was measured in 6 h-intervals (phase contrast, 10X objective) using the IncuCyte S3 Live-Cell Analysis System (IncuCyte S3). The half-maximal inhibitory concentrations for CDDP (IC_{50}, section **4.14.2**) were derived from WST-1 and crystal violet assay results as well as from the tracked cell confluence using the 144 h timepoint (five days after initiating the treatment).

For measurements of apoptosis (section **4.8**) and for colony formation assays (section **4.9**), cells were treated with CDDP in 6-well plates. The cells were seeded in 6-well plates (15,000 per well for all DU145 cell lines, 50,000 for PC-3^C RR and 30,000 for all other PC-3 cell lines), cultured to ~50% confluence and then treated with 0.01–1.0 µg/mL CDDP for 24 h. The treatment was removed and 24 h later, adherent cells were harvested from the 6-well plates by trypsinization.

4.6 WST-1 Assay

The WST-1 assay was performed according to the manufacturer's instructions. Tetrazolium salts like MTT, XTT and WST-1 are common chemicals to assess the *in vitro* cytotoxicity of chemicals. Tetrazolium salts are reduced to their corresponding formazans by enzymatic reactions that depend on the presence of NADH. In contrast to MTT and XTT, which are reduced intracellularly (by the mitochondrial succinate-tetrazolium-reductase system), WST-1 is primarily reduced at the outer plasma membrane by the enzyme NADH oxidase[156]. The readout of the WST-1 assay is positively proportional to the availability of NADH, which depends on glycolytic and citric acid cycle activity. Therefore, the WST-1 assay measures metabolic activity. The formazan absorption was measured at 450 nm with a reference measurement at 620 nm using a Berthold Mithras LB940 microplate reader.

4.7 Crystal Violet Assay

The crystal violet dye stains both living and dead cells. However, washing removes non-adherent cells during the assay protocol. The assay was performed directly after the WST-1 assay. The media was removed, cells were fixed with methanol for 10 min and then stained with an aqueous solution of 0.1% crystal violet for 10 min. After thorough washing with water, stained adherent cells were dissolved by adding 100 µL of 0.1 M sodium citrate in 50% ethanol per well and incubating for 30 min on a shaker at room temperature (RT). Crystal violet absorption was measured at 590 nm using the Berthold Mithras LB940 microplate reader.

4.8 Measurement of Apoptosis

Cells were seeded in 6-well plates and treated as described in section **4.5**. The medium was removed and the culture wells were washed with 1 mL PBS. Then, adherent cells were harvested with an aqueous solution containing 0.05% trypsin/EDTA (300 µL per well). Trypsinization was stopped by adding 1 mL RPMI cell culture medium per well. The cell suspensions were transferred to 1.5 mL reaction tubes, centrifuged (4°C, 500 g, 5 min) and the pellets were resuspended in 1 mL ice-cold PBS. For measurement of apoptosis based on pan-caspase activity, 4×10^4 viable cells were analyzed using the Muse® Multi Caspase kit on the Muse® Cell Analyzer according to the manufacturer's instructions.

4.9 Colony Formation Assay

CDDP treated cells were harvested from 6-well plates and counted as described in section **4.5**. For colony formation assays, technical triplicates of 100 (all DU145 cell lines) or 200 (all PC-3 cell lines) viable cells per well were seeded into 6-well plates. At day 8–12 of culture, phase contrast images (10X objective) of non-stained colonies were taken with the IncuCyte S3 Live-Cell Analysis System and the colonies were counted from these images. Only colonies with >20 cells were considered. Colonies were then fixed with methanol, stained with an aqueous solution of 0.1% crystal violet and recounted to validate the image-based counting method.

4.10 Transient Knockdown of *MLH1* by siRNA

The effect of a transient knockdown of MLH1 on CDDP tolerance was investigated in PC-3^B and PC-3^C cells. In preliminary experiments, 72 h was determined as the timepoint of strongest downregulation of MLH1 on the transcript and protein level. 1,000 (96-well) or 50,000 (6-well) cells per well were seeded and the transfection was performed 24 h after seeding. For the transfection, the RPMI medium was removed and the cells were treated with OptiMEM (60 or 600 µL in 96- or 6-well plates, respectively) containing 20 nM siRNA and 0.3% Lipofectamin 2000 for 4 h at 37°C. Then, fresh RPMI medium (140 or 1400 µL, respectively) was added and the cells were cultured for 72 h. Subsequently, the cells seeded in 96-well plates were treated with a serial dilution of CDDP (0.01 – 1 µg/mL) and cultured for an additional four days. Confluence was tracked in the IncuCyte S3 during the whole experiment and WST-1 as well as crystal violet assays (sections **4.6** and **4.7**, respectively) served as endpoints. The cells seeded in 6-well plates were harvested after transfection for RNA and protein isolation, which was followed by quantitative PCR (qPCR) and western blot analysis, respectively. The timepoint of RNA and protein isolation from 6-well plates corresponded with the timepoint at which CDDP treatment was initiated in the 96-well plates.

4.11 Measurement of Gene Expression

4.11.1 RNA Isolation

RNA was isolated from cells cultured in either 10 cm dishes or 6-well plates. The medium was removed on ice, the cells were washed twice with ice-cold PBS and then detached from the dish using a cell scraper in 1 mL PBS. Then, the cell suspensions were centrifuged (4°C, 500 g, 5 min), the pellets were thoroughly resuspended in 375 µL TrisFast peqGOLD and stored at -80°C until further processing. The spin-column based Direct-zol™ RNA Miniprep-Kit was used for RNA isolation according to the manufacturer's instruction. Finally, the concentration and purity of the eluted RNA was measured using a NanoDrop 2000c Spectrophotometer. Thereafter, samples were stored at -80°C until further use.

4.11.2 cDNA Synthesis

All steps were performed on ice. Sample-dependent volumes containing 500 ng eluted RNA were topped up to 11 µL with double-distilled (dd)H_2O in a 0.2 mL PCR tube. Next, 2 µL of a Master mix containing 10 mM dNTPs and 0.2 µg/mL pd(N)6 random primers were added followed by 5 min incubation at 65°C and 5 min on ice. Thereafter, 4 µL of 5X First Strand Buffer and 2 µL of 0.1 M DTT were added followed by 2 min incubation at RT. Finally, 1 µL (200 units) Superscript II Reverse Transciptase was added. The final volume per sample amounted to 20 µL. The synthesis of cDNA took place in a Thermocycler PTC-100 (25°C: 10 min, 42°C: 50 min and 70°C: 15 min). Undiluted cDNA was stored at -80°C until further use. cDNA intended for timely use was diluted 1:5 with ddH_2O and stored at 4°C.

4.11.3 Single-mRNA qPCR with TaqMan Assays

Sample preparation for qPCR was carried out on ice. One reaction consisted of 9 µL master mix and 1 µL 1:5 diluted cDNA (see section **4.11.2**), amounting to a reaction volume of 10 µL. The master mix consisted of 3.5 µL ddH_2O, 0.5 µL 20X TaqMan probe (see **Table 9**) and 5 µL 2X GoTaq probe qPCR mastermix. For each TaqMan probe, batches of the master mix were prepared in the volume required for the number of respective cDNA samples (multiples of 9 µL). Then, 9 µL master mix and 1 µL cDNA (or 1 µL ddH_2O as a negative control) per well were transferred into 96-well plates suitable for the LightCycler 480 instrument (see **Table 6**). The plates were sealed and centrifuged before placing them into the instrument. Initial denaturation was done for 10 min at 95°C (heating rate: 4.4°C/s). 45 cycles of amplification were performed. Each amplification cycle consisted of 15 s denaturation at 95°C (heating rate: 4.4°C/s) and 60 s of primer/probe annealing and extension at 60°C (cooling rate: 2.2°C/s). Finally, the samples were first cooled to 40°C (cooling rate: 1.5°C/s) for 1 min and then cooled down to 4°C. The cycle threshold (CT) values were determined by the LightCycler 480 software (automatic second derivative method). The geometric mean of the CT values of the house keeping genes (HKGs) RPLP0 and HPRT1 was used to calculate ΔCT values. Differential gene expression analyses were carried out using the ΔΔCT method. The CT values of two independent runs per sample were averaged and if the mean deviation was > 0.25, a third run was performed. CT values > 35 were considered negative.

4.11.4 RT^2 Profiler Arrays

Pathway-specific mRNA panels were quantified using qPCR by means of RT^2 Profiler Arrays (**Table 10**). 96-well formatted RT^2 Profiler Arrays measure 84 pathway-specific mRNAs and five HKGs including RPLP0 and HPRT1. The remaining seven wells consist of contamination controls, quality controls and negative controls. For one array, a master mix of 1350 µL of SYBR Green master mix, 102 µL of 1:5 diluted cDNA (section **4.11.2**) and 1248 µL ddH_2O was prepared. After gentle vortexing, 25 µL master mix was transferred into each well. The plate was sealed and shortly centrifuged. At last, qPCR was performed on the LightCycler 480 using a program described in section **4.11.3** with the following differences. During the amplification phase, the heating rate was at 1.5°C/s. After the last amplification cycle, samples were not cooled to 40°C, but instead heated to 95°C (heating rate: 0.03°C/s). During the final heating of the samples, 20 measurements per °C were taken in order to acquire melting curves. Likewise, differential gene expression analysis was done as described in section **4.11.3**.

4.12 Western Blot Analysis

4.12.1 Protein Isolation

For protein isolation, cells were seeded in 10 cm dishes and cultured to ~ 70% confluence. All following steps were carried out on ice. After removing the medium, the cells were washed two times with ice-cold PBS and then detached from the dish using a cell scraper in 1 mL PBS. The suspension was centrifuged (4°C, 500 g, 5 min) and the pellet was thoroughly resuspended in 50 – 300 µL RIPA cell lysis buffer depending on the pellet size. To open up and decompose the nucleic envelope – which is not sufficiently achieved by RIPA buffer alone – the suspensions were shock-frosted in liquid N_2 and then incubated at 4°C and 2000 rpm on a reaction tube shaker for at least 30 min. The resulting suspensions were centrifuged (4°C, 21,250 g, 15 min) and the supernatant was transferred to a fresh reaction tube in order to remove remaining supramolecular debris. 10 µL of this lysate was again transferred into a fresh reaction tube and diluted 1:10 with PBS for the measurement of protein concentration via BCA assay. The BCA assay was carried out using the Pierce™ BCA protein assay kit according the manufacturer's instructions on the Berthold Mithras LB940 microplate reader. Finally, one part of 4X loading buffer was added to three parts of protein lysate, β-mercaptoethanol was added to a final concentration of 1% and the lysates were denatured at 80°C (or 60°C for ABCB1 western blots) for 10 minutes. The final protein concentration was calculated taking into account the 1:10 dilution for the BCA assay and the dilution resulting from adding the loading buffer. Samples were stored at -20°C until they were used for western blotting.

4.12.2 Western Blotting

Western blots were performed with the NuPAGE system, consisting of the running chamber, premade gels as well as buffers and chemicals. The chamber was assembled with either one or two gels. The sample pockets were rinsed with running buffer using a Hamilton syringe prior to loading sample volumes equivalent to 5 – 20 µg of protein per lane, depending on the experiment. Every marker lane contained, 7 µL of Spectra Multicolor Broad Range protein marker (colored and visible by eyesight) as well as 1 µL Novex Magic Mark XP (visible in the chemiluminescence detection device). For large proteins (> 150 kDa) the MOPS running buffer was used. Otherwise, the running chamber was filled with MES running buffer. Before starting the electrophoresis, 500 µL antioxidant (NuPAGE) was added to the inner chamber. A current of 150 V was applied for 1.5 – 2 h for electrophoresis, after which the proteins were transferred onto nitrocellulose membranes (**Table 5**) using either the iBlot or the iBlot 2 semi dry blotting device (20 V for 13 min). To control the transfer, a Ponceau S staining was performed. After thorough destaining in TBS-T, the membranes were blocked in 10% MMP/TBS-T for at least 1 h at RT. If necessary, the membranes were cut and individual fragments or whole membranes were shrink-wrapped in plastic foil for primary antibody incubation. Primary antibodies (**Table 11**) were diluted in 1% MMP/TBS-T, which was filtered (0.45 µm pore size) before to remove larger undissolved NFDM particles. Primary antibody incubation was done overnight at 4°C. The next day, membranes were washed with TBS-T (three times 15 min) before they were again shrink-wrapped for secondary antibody incubation. Secondary antibodies (see **Table 11**) were diluted in filtered 1% MMP/TBS-T and incubation took place for 1.5 h at RT on a shaker. Thereafter, the membranes were washed two times for 10 min with TBS-T and once for 10 min with TBS before transferring them to Aqua dest. Finally, the WesternBright Sirius chemiluminescence kit was used according to the manufacturer's instructions and – depending on the protein and antibody – the chemiluminescent signal was collected for up to 2 h in a MicroChemi 4.2 chemiluminescence scanner. Densitometric analyses of western blot data were done using Fiji (section **3.11.5**). GAPDH or alpha-tubulin served as reference proteins.

4.13 Experiments related to ABCB1 and Docetaxel Resistance

The ABCB1 project was initiated and conceptualized by Mr. Lukas Donix (the author of this dissertation). Consecutive experiments were at large conducted by Ms. Dinah Linke within her master thesis under the practical supervision of Mr. Lukas Donix (section **11** on page **142**). A summary of the employed methods is provided in the following sections.

4.13.1 Docetaxel Treatment of Cell Lines

The tolerance for DTX was determined for all DU145 and PC-3 cell lines. 600 – 800 DU145 or 1,500 – 2,200 PC-3 cells per well were seeded in 96-well plates. Cells were treated at 30-50% confluence. In contrast to the CDDP treatment, which remained in the culture medium (section **4.5**), the serial dilution of DTX (0.01 – 10 µM) was removed after 24 h. The WST-1 assay (section **4.6**) and the crystal violet assay (section **4.7**) served as endpoints. IC_{50} values were determined from the resulting inhibition curves.

4.13.2 Hoechst-33342 Efflux

The efflux of the Hoechst-33342 dye, which is a substrate of ABCB1 as well as ABCG2, was studied in PC-3^{B}, PC-$3^{B\ DTXR}$, DU145^{B} and DU$145^{B\ DTXR}$ cells. 16,000 cells were seeded in black 96-well plates suitable for fluorescence measurements. 12 h later, the Hoechst-33342 dye was added in a final concentration of 5 µg/mL with or without verapamil in a final concentration of 50 µM. The cells were incubated at 37°C for 90 min. After removal of the treatment and a single wash with RPMI, 100 µL ice-cold PBS was added to the wells and the plates were stored on ice for a maximum of 30 min before measurement. Fluorescence was measured (excitation at 360 nm, emission at 450 nm) using a TECAN Infinite® 200 Pro microplate reader.

4.13.3 Flow Cytometry

In PC-3^{B}, PC-$3^{B\ DTXR}$, DU145^{B} and DU$145^{B\ DTXR}$ cells, the expression of ABCB1 and ABCG2 was measured by flow cytometry. Cells were seeded and harvested using accutase to reduce the digestion of cell surface proteins. Hereafter, cells and cell suspensions were kept at 4°C or on ice. Cell suspensions containing accutase were centrifuged (4°C, 500 g, 5 min) and the pellets were resuspended in 5 mL flow cytometry staining buffer (FCS buffer, **Table 8**). Next, the cells were counted and aliquots of 500,000 cells were prepared in reaction tubes. The suspensions were again centrifuged (4°C, 500 g, 5 min) and the resulting pellets were resuspended in 100 µL FCS buffer containing target-specific or isotype control antibodies, respectively (**Table 12**). The samples were stained at 4°C for 30 min in the dark. Thereafter, 1 mL of plain FCS buffer was added and the samples were centrifuged (4°C, 500 g, 5 min). The pellets were then resuspended in 300 µL FCS buffer containing 0.5 µg/mL 7-AAD. In addition, compensation control samples for the PE fluorophore (**Table 12**) and for 7-AAD were prepared, accordingly. The measurement was performed on a BD Celesta flow cytometer with excitation optics containing the blue (488 nm), yellow-green (561 nm) and violet (405 nm) laser. Flow cytometry data were analyzed and plotted using FlowJo cell analysis software.

4.13.4 PNGase F Digestion of Protein Lysates

Protein lysates were prepared as described in section **4.12.1** until the denaturation step. Denaturation was performed by adding 1% β-mercapthoethanol and incubating the samples with 10 mU peptide-N-glycosidase F (PNGase F) per 100 µg protein on a shaker at 37°C for 2 h with gentle shaking. The loading buffer was added after denaturation and the samples were stored at -20°C until western blot analysis.

4.13.5 Targeting of ABCB1 and Subsequent Dose Response to DTX Treatment

The experiments described herein were conducted by Ms. Dinah Linke M.Sc. Please refer to the data origination statement in section **9**. In summary, cells were seeded and treated with DTX in essence as described in section **4.13.1**. The cells were treated with either DTX alone or with DTX in combination with 50 nM of the ABCB1 inhibitors tariquidar or elacridar. During the treatment, the confluence was tracked in the IncuCyte S3. WST-1 as well as crystal violet assays were performed in the end. In addition, a siRNA-mediated knockdown of ABCB1 was performed, in essence as described in section **4.10**. This was followed by qPCR and western blot analysis. Furthermore, ABCB1 was knocked down in conjunction with DTX treatment, which was followed by colony formation assays as described in section **4.9**, WST-1 and crystal violet assays. Furthermore, a potential influence of an inhibition of N-linked glycosylation on DTX tolerance was studied. For this, cells were pretreated with either 100 ng/mL tunicamycin or 1 µM swainsonine before receiving a DTX treatment as described in **4.13.1.** The DTX tolerance was evaluated by WST-1 and crystal violet assays and the inhibition of ABCB1 glycosylation was confirmed by western blots.

4.14 Statistics

4.14.1 Statistical Tests

If not mentioned otherwise, a minimum of three independent biological replicates were performed for each experiment. Furthermore, all data are plotted as their mean with standard deviation (SD) if not otherwise indicated. Two-sided t-tests with Welch's correction (assumption of unequal variance) were used to test for statistically significant differences between the means of two groups. Due to the assumption of unequal variance, Welch's t-tests produce more conservative results than the normal t-tests.

4.14.2 Non-linear Regression and Calculation of the IC_{50} and $\beta_{1/2}$ Values

During the CDDP treatment of cells, their confluence was tracked in the IncuCyte S3 as described (section **4.5**). GraphPad Prism's non-linear 'log(agonist) vs. response' regression with variable slope was used to fit the untransformed confluence data (**Figure 22**). As cell confluence cannot be lower than 0% or higher than 100%, the lower plateau was constrained to be ≥ 0 and the upper plateau was constrained to be ≤ 100. Furthermore, the half-maximal inhibitory concentrations (IC_{50}) for CDDP and DTX treatments were calculated based on confluence data (**Figure 23**), WST-1 assays (**Figure 24**) and crystal violet assays (**Figure 25**). For the confluence data, the 144 h timepoint (five days after initiation of the treatment) was used. As mentioned, WST-1 assays and crystal violet assays were performed six days after the initiation of the treatment. The data were first normalized to the untreated control and then log-transformed. Then, GraphPad Prism's non-linear 'log(antagonist) vs. response' regression with variable slope was used for curve fitting. Due to data-normalization the upper plateaus of the inhibition curves were constrained to equal 1 and because negative values would not be reasonable in the experimental context (a zero represents 100% inhibition of the original measurement), the lower plateaus of the curves were constrained to be ≥ 0. The IC_{50} is one of GraphPad Prism's output parameters for the resulting inhibition curves.
Furthermore, non-linear regression was used to model the reduction of colonies in response to the CDDP treatment in the colony formation assays. First, the x-fold number of colonies was derived by normalizing the colony numbers in the treated groups to the untreated group. Then, GraphPad Prism's non-linear 'exponential decay' regression was used to fit curves onto the data. Due to previous data normalization, the curves were constrained to cross the y-axis at y = 1. The concentration $\beta_{1/2}$ at which the colony formation was reduced by 50% was one of the output parameters for the resulting inhibition curves.

4.14.3 Hierarchical Clustering

Hierarchical clustering based on log-transformed $2^{-\Delta CT}$ or $2^{-\Delta\Delta CT}$ values (**Figure 19**, **Figure 20** and **Figure 21**) and based on the relative abundance of SBS events in mutational signatures as well as the generation of dendrograms was done using R. In short, the default parameters for the calculation of the distance matrix (stats::dist(), method = 'euclidean') and for hierarchical clustering (stats::hclust(), method = 'complete') were used.

5 Results

5.1 Cell Line Authentication

All cell lines were subjected to authentication via STR analysis. In short, the identity of all cell lines was confirmed (**Table 15**). All detected deviations were minor changes in the expected repeat-length of individual markers and were not sufficient to suggest sample contamination. PC-3 cells had the least number of deviations from their expected profile. Most PC-3 cells had none. Radiation-treated PC-3^{C} RR and CDDP-treated PC-3^{D} CDDPR cells had deviations in three or two authentication markers, respectively (**Table 15**). DU145, LNCaP and LAPC-4 cells generally had deviations in several authentication markers (**Table 15**). The same was true for the LNCaP derivatives C4-2 and MR49F considering the profile of LNCaP cells as a reference (**Table 15**). DU145^{A} was (next to the control DNA) the only DU145 cell line without deviations from the expected profile. With 11 deviations DU145^{B} DTXR had most deviations from the expected profile among the DU145 cells (**Table 15**). The purchased LNCaP control DNA, LNCaP^{A} and LNCaP^{B} cells had 4 deviations from the expected profile. LNCaP^{B} BicaR, LNCaP^{C} and LNCaP^{C} RR on the other hand had 8, 9 and 12 deviations, respectively (**Table 15**). LAPC-4^{A}, LAPC-4^{B} and LAPC-4^{B} EnzaR had 6, 8 and 8 deviations, respectively. However, the profile among the three tested LAPC-4 cell lines was very similar (**Table 15**).

Table 15: Cell line authentication via STR analysis.

STR marker	Expected in PC-3	Purchased Control DNA	PC-3^{A}	PC-3^{A} CDDPR	PC-3^{B}	PC-3^{B} DTXR	PC-3^{C}	PC-3^{C} RR	PC-3^{D}	PC-3^{D} CDDPR
Amelogenin	X	X	X	X	X	X	X	X	X	X
CSF1PO	11	11	11	11	11	11	11	10, 11	11	11
D13S317	11	11	11	11	11	11	11	10, 11	11	11
D16S539	11	11	11	11	11	11	11	11	11	11
D18S51	14, 15	14, 15	14, 15	14, 15	14, 15	14, 15	14, 15	14, 15	14, 15	14, 15
D21S11	29, 31.2	29, 31.2	29, 31.2	29, 31.2	29, 31.2	29, 31.2	29, 31.2	29, 31.2	29, 31.2	29, 31.2
D3S1358	16	16	16	16	16	16	16	16	16	16
D5S818	13	13	13	13	13	13	13	13	13	13
D7S820	8, 11	8, 11	8, 11	8, 11	8, 11	8, 11	8, 11	8, 11	8, 11	8, 11
D8S1179	13	13	13	13	13	13	13	13	13	13
FGA	24	24	24	24	24	24	24	23, 24	24	24
Penta D	9	9	9	9	9	9	9	9	9	9
Penta E	10, 17	10, 17	10, 17	10, 17	10, 17	10, 17	10, 17	10, 17	10, 17	10, *
THO1	6, 7	6, 7	6, 7	6, 7	6, 7	6, 7	6, 7	6, 7	6, 7	6, *
TPOX	8, 9	8, 9	8, 9	8, 9	8, 9	8, 9	8, 9	8, 9	8, 9	8, 9
vWA	17	17	17	17	17	17	17	17	17	17
Deviations from expected profile:		0	0	0	0	0	0	3	0	2
Deviations from parent profile:		-	-	0	-	0	-	3	-	2

(Continued on next page)

Continuation of Table 15: Cell line authentication via STR analysis.

STR marker	Expected in DU145	Purchased control DNA	DU145^{A}	DU145^{A} CDDPR	DU145^{B}	DU145^{B} DTXR	DU145^{C}	DU145^{C} RR
Amelogenin	X, Y	X, Y	X, Y	X, Y	X, Y	X, Y	X, Y	X, Y
CSF1PO	10, 11	10, 11	10, 11	10, 11	10, 11	9, 10, 11, 12	9, 10, 11	10, 11
D13S317	12, 13, 14	12, 13, 14	12, 13, 14	12, 13, 14	12, 13,*	10, 11, 12, 13,*	12, 13, 14	12, 13, 14
D16S539	11, 13	11, 13	11, 13	11, 13	11, 12, 13	11, 12, 13	11, 12, 13	11, 13
D18S51	12	12	12	12	12	11, 12	12	11, 12
D21S11	30, 33	30, 33	30, 33	30, 33	30, 31, 32,*	29, 30, 32, 33	*, 33	30, 33
D3S1358	16	16	16	16	16	15, 16	16	16
D5S818	10, 13	10, 13	10, 13	10, 13, 14	10, 13	9, 10, 13	10, 13	10, 13
D7S820	7, 10, 11	7, 10, 11	7, 10, 11	7, 10, 11	7, *, 11	7, *, 11, 13	7, 10, 11	7, 10, 11
D8S1179	13, 14	13, 14	13, 14	13, 14	13, 14	13, 14, 15	13, 14	13, 14
FGA	21, 22	21, 22	21, 22	*, 22	*, 22	21, 22, 23	*, 22	*, 22
Penta D	9, 13	9, 13	9, 13	9, 13	9, 13	9, 13	9, 13	9, 13
Penta E	12, 14	12, 14	12, 14	12, 13, 14	12, 14	12, 14	12, *, 15	11, 12, 14
THO1	7	7	7	7	7	7	7	7
TPOX	11	11	11	11	11	11	11	10, 11
vWA	17, 18, 19	17, 18, 19	17, 18, 19	17, 18, *	*, 18, *	17, 18, *	17, 18, *	17, 18, *
Deviations from expected profile:		0	0	4	6	11	5	5
Deviations from parent profile:		-	-	4	-	10	-	6

(Continued on next page)

Continuation of Table 15. Cell line authentication via STR analysis.

STR marker	Expected in LNCaP	Purchased control DNA	LNCaP^{A}	LNCaP^{B}	LNCaP^{B} BicaR	LNCaP^{C}	LNCaP^{C} RR
Amelogenin	X, Y	X, Y	X, Y	X, Y	X, Y, OL	X, Y, OL	X, Y, OL
CSF1PO	10, 11	10, 11	10, 11	10, 11	10, 11	9, 10, 11	10, 11
D13S317	10, 12	10, 12	10, 12	10, 12	10, 12	10, *	10, 11, *
D16S539	11	11	11	11	11	11	11
D18S51	11, 12	11, 12	11, 12	11, 12	11, 12	11, 12	10, 11, 12
D21S11	29, 31.2, 32.2	29, 31.2, 32.2	29, *, 32.2	29, 31.2, *	28, 29, *, 32.2, 34	29, *, 32.2	29, *, 32.2
D3S1358	16	15, 16	16	16	16	15, 16, 17	15, 16
D5S818	11, 12	11, 12	11, 12	11, 12	11, 12, 13	11, 12	10, 11, 12
D7S820	9.1, 10.3	9.1, 10.3	9.1, 10.3	9.1, 10.3	9.1, 10.3	9.1, 10.3	9, 9.1, 10.3
D8S1179	12, 13, 14	12, *, 14	12, *, 14	12, *, 14	11, 12, *, 14, 15	11, 12, *, 14	11, 12, *, 14
FGA	19, 20, 21	19, 20, *	19, 20, *	19, 20, *	19, 20, 21	18, 19, 20, *	18, 19, *, *
Penta D	12, 12.4	12, *, OL	12, *, OL	12, *, OL	11, 12, *, OL	10, 12, *, 13	10, 12, *, 13
Penta E	12, 16	12, 16	12, 16	12, 16	12, 16, 17	12, 16	10, 12, 16
THO1	9	9	9	9	8, 9	9	9
TPOX	8, 9	8, 9	8, 9	8, 9	8, 9	8, 9	8, 9
vWA	16, 18	16, 18	16, 18	16, 18	16, 17, 18, 19	15, 16, 18	15, 16, 18, 19
Deviations from expected profile:		4	4	4	8	9	12
Deviations from parent profile:		-	-	-	9	-	9

(Continued on next page)

Continuation of Table 15. Cell line authentication via STR analysis.

STR marker	Expected in LAPC-4	LAPC-4^{A}	LAPC-4^{B}	LAPC-4^{B} EnzaR	Expected in C4-2	C4-2	Expected in MR49F	MR49F EnzaR
Amelogenin	X, Y	X, Y	X, Y	X, Y	X, Y	X, Y	no data	X, Y
CSF1PO	12, 13	12, *	12, *	12, *	9, 10, 11	9, 10, 11	no data	10, 11
D13S317	9, 10, 11, 12	*, *, 10, 12	*, *, 10, 12	*, *, 10, 12	10, 11	10, 11	no data	11
D16S539	9	9	9	9	10, 11	10, 11	no data	11
D18S51	14, 15, 22	13, 14, *, 21, 22	14, 15, 21, 22, 23	14, 15, 21, 22, 23	no data	10, 11, 12	no data	11, 12
D21S11	28, 20	28, 30	28, 30	28, 30	no data	29, 32.2	no data	29, 32.2
D3S1358	15, 17	15, 17	15, 16, 17	15, 16, 17	no data	15, 16	no data	16
D5S818	11, 12, 13	11, 12, 13	11, *, 13	11, *, 13	11, 12	11, 12	no data	11, 12
D7S820	10, 10.3, 11	*, 10.3, 11	*, 10.3, 11	*, 10.3, 11	9.1, 10. 3	9.1, 10.3	no data	9.1, 10.3
D8S1179	13	13	12, 13	12, 13	no data	12, 14, 15	no data	12, 14
FGA	21, 22	21, 22	21, 22	21, 22	no data	19, 20	no data	19, 20
Penta D	no data	10	10	10	no data	12, 13	no data	12, 13
Penta E	no data	12, 13	12, 13	12, 13	no data	12, 16, 17	no data	12, 15, 16
THO1	6, 9.3	6, 9.3	6, 9.3	6, 9.3	9	9	no data	9
TPOX	8	8, 9	8	8	8, 9	8, 9	no data	8, 9
vWA	15, 16	*, 16	*, 16	*, 16	16, 18	16, 18	no data	16, 18, 19
Deviations from expected profile:		6	8	8	-	0	-	-
Deviations from parent profile:		-	-	0	-	10 (LNcaP)	-	7 (LNCaP)

Legend: Cell line authentication of DU145 (Cellosaurus: CVCL_0105), PC-3 (CVCL_0035), LNCaP (CVCL_0395), LAPC-4 (CVCL_4744), C4-2 (CVCL_4782) and MR49F (CVCL_RW53) cells via STR analysis. Deviations from the expected STR profile are marked red. * denotes missing alleles. Every STR marker not matching the expected profile is counted as one deviation. The expected profiles were retrieved from cellosaurus[139] which lists multiple data sources. ATCC was the preferred data source and where ATCC data were not available, DMSZ data were used. Where neither was available, other data sources were considered. The profile information for C4-2 cells was directly retrieved from the ATCC website. Missing data for the expected STR profile markers are labelled with 'no data'. Control DNA for LAPC-4, C4-2 and MR49F cells was not commercially available. OL: off ladder allele (allele not covered by the reference ladder of the authentication kit)[157].

5.2 Whole Exome Sequencing

WES of all 24 cell lines listed within **Table 4** and **Table 15** was performed. The variants were called and filtered from the fastq files as described in section **4.4** and recorded into a MAF file for further analysis.

5.2.1 Dataset Characteristics

The average sequencing coverage per variant was at 150 reads, skewed by highly covered variants with coverages up to 3,227 reads (data not shown). The median coverage per variant was at 73 and the lower and upper quartiles were at 34 and 144 reads per variant, respectively (**Figure 7 a**). The coverage was slightly variable across the cell lines. With 91 and 58 reads, LNCaP^{C} RR cells had the highest and PC-3^{B} the lowest median coverage per variant in the dataset. Generally, LNCaP cell lines as well as C4-2 and MR49F EnzaR cells had elevated median coverages, while PC-3 cell lines had lower median coverages (**Figure 7 a**).
The median VAF of the dataset was 42.8%. With 63.7 and 34.4%, PC-3^{B} DTXR had the highest and MR49F EnzaR cells the lowest median VAF, respectively. The lower and upper VAF quartiles for the whole dataset were at 26.1 and 66.1%, respectively (**Figure 7 a**).
The dataset comprised of a total of 377,225 variants that met the inclusion criteria defined in section **4.4**. 317,386 (84.1%) of the variants were SNVs, 7.2% were deletions, 4.3% insertions and 4.4% MNVs, respectively (**Figure 7 b**). 267,166 (70.8%) variants were either synonymous or located in non-coding regions. For most of the consecutive analyses, these variants were excluded because they do not directly affect the amino acid sequence of their target proteins and are therefore significantly less likely to be of consequence.
110,059 (29.2%) of the variants were located in coding regions and were non-synonymous. Of these, 95,964 (87.2%) were missense mutations (predominantly SNVs) and 14,095 (12.8%) had other variant classifications. Among these 14,095 variants (other coding and non-synonymous variants), mutations located near splice sites constituted the largest subgroup (33.1%), followed by nonsense mutations (26.7%), frame shift InDels (23%) and in frame InDels (15.2%). With 158 (1.1%) and 123 (0.9%) variants, nonstop mutations and mutations affecting start codons were the rarest variant classes among other coding and non-synonymous variants, respectively (**Figure 7 b**).

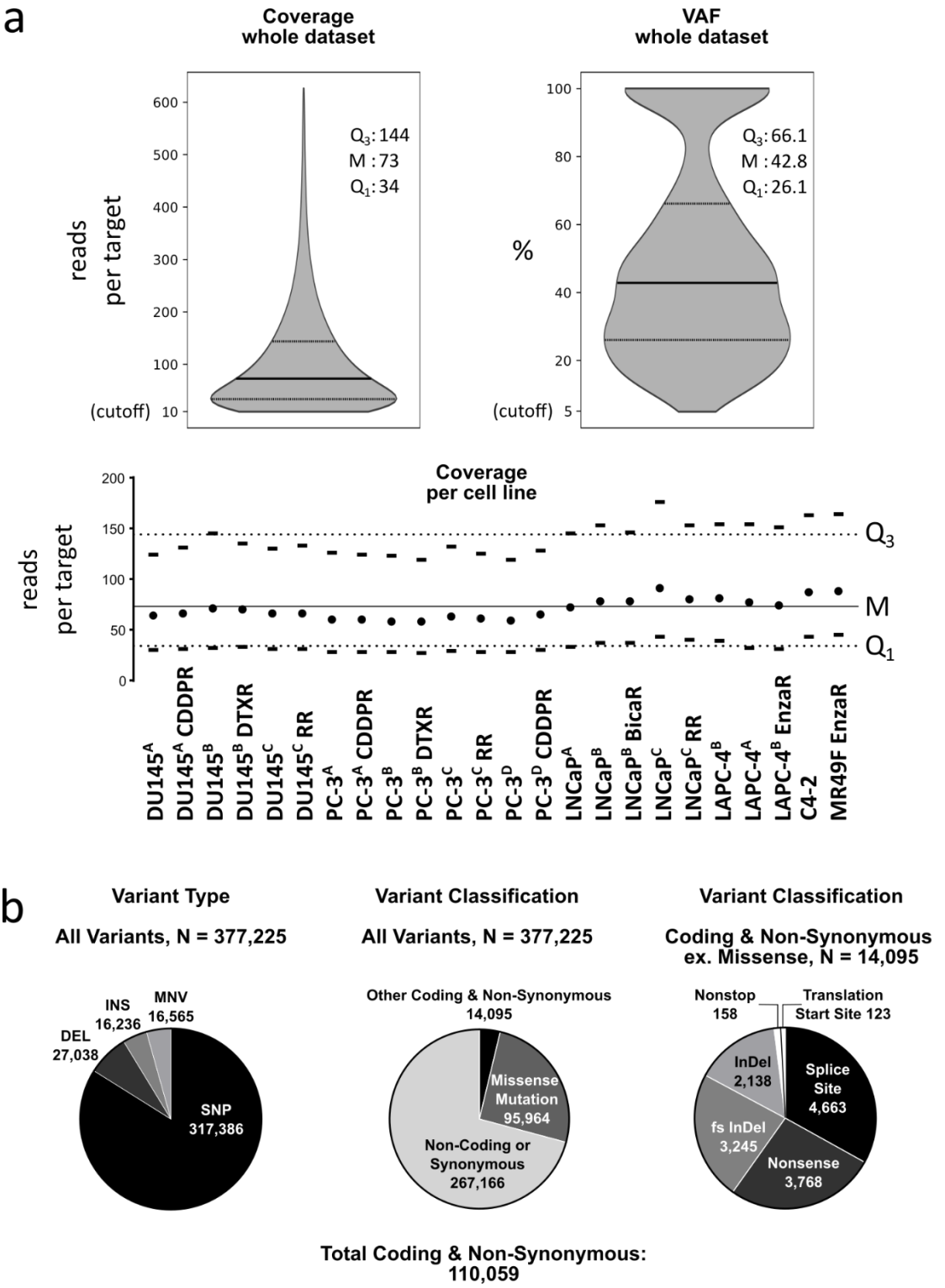

Figure 7: Summary of WES dataset characteristics. (a) On the left, the sequencing coverage per variant is shown. The top 1% of variants that were covered by 627 to 3,227 reads and were removed in order to better plot the data. However, they were considered in the calculation of the median and the quartiles displayed on the plot. On the right, the distribution of the VAF is shown. Below, the coverage per cell line is shown. The full line and the black circles represent the dataset median and cell line medians, respectively. The dotted lines and black bars represent the dataset and cell line lower and upper quartiles, respectively. (b) Pie charts give a summary of the distribution of variant types (left) and variant classifications (middle and right side). Variant types and classifications are described in more detail in **Table 3**. Q_1: lower quartile, M: median, Q_3: upper quartile.

5.2.2 Mutational Burden

The total mutational burden of the cell lines was studied. Only non-synonymous coding variants (N = 110,059) were considered. In absolute numbers, DU145 cells harbored an average of 4,935 ± 326, PC-3 cells of 2,065 ± 316, LNCaP cells of 6,106 ± 846 and LAPC-4 cells of 6,275 ± 274 variants, respectively (**Table 16**). The average of C4-2 and MR49F EnzaR cells was 7,284 variants. With 7,992 and 1,850, MR49F EnzaR harbored the highest and PC-3^{B} the lowest number of variants, respectively. The cell lines where the number of variants deviated more than one standard deviation (σ) from their respective group averages were $DU145^{A}$ (1.37 σ below), $DU145^{B}$ DTXR (1.52 σ above), PC-3^{D} CDDPR (2.4 σ above), LNCaPC RR (1.63 σ above) and LAPC-4^{B} (1.04 σ above).

From these numbers, the implied number of (protein-changing) variants per kilo base pairs (v / kbp) was estimated (**Table 16**). Since only variants located in exons were considered, the numbers were divided by the total length of the human exome, which was estimated as described in section **4.4.1**. DU145 cells harbored approximately one protein-altering mutation in every 4,000 bp of coding sequence. In the other cell line groups, one protein-altering mutations was present approximately in every 10,000 bp in PC-3 cells, every 3,000 bp in LNCaP cells and LAPC-4 cells and in every 2,500 bp in the group comprising of the cell lines C4-2 and MR49F EnzaR. For per cell line details please refer to **Table 16**.

Assuming a total number of unique genes of 19,030 (section **4.4.1**), 16.9 ± 1.1% of all genes carried at least one protein-altering mutation in DU145 cells (**Table 16**) and in the other cell line groups, 6.5 ± 1.3% (PC-3 cells), 19.2 ± 2.8% (LNCaP cells), 20.3 ± 0.4 (LAPC-4 cells) and 23.1% (C4-2 & MR49F EnzaR) of all genes carried at least one protein altering mutation, respectively. For per cell line details please refer to **Table 16**.

Table 16: Cell line mutational burden. Only non-synonymous variants located within coding sequences (N = 110,059) were considered. The implied variants per kbp and the % of genes with variants were approximated as described in section 4.4.1.

Cell Line	Variants	Implied variants / kbp	Genes with variants	% Genes with variants
DU145[A]	4,487	0.238	2,931	15.4
DU145[A] CDDPR	4,698	0.249	3,048	16.0
DU145[B]	4,906	0.260	3,176	16.7
DU145[B] DTXR	5,432	0.288	3,543	18.6
DU145[C]	5,073	0.269	3,302	17.3
DU145[C] RR	5,019	0.266	3,267	17.2
Group Average	4,935 ± 326	0.261 ± 0.017	3,211 ± 213	16.9 ± 1.1
Group Median	4,962	0.263	3,221	16.9
PC-3[A]	2,048	0.109	1,225	6.4
PC-3[A] CDDPR	2,031	0.108	1,202	6.3
PC-3[B]	1,850	0.098	1,085	5.7
PC-3[B] DTXR	1,859	0.098	1,084	5.7
PC-3[C]	1,935	0.103	1,140	6.0
PC-3[C] RR	2,004	0.106	1,199	6.3
PC-3[D]	1,965	0.104	1,182	6.2
PC-3[D] CDDPR	2,825	0.150	1,857	9.7
Group Average	2,065 ± 316	0.109 ± 0.017	1,246 ± 252	6.5 ± 1.3
Group Median	1,984	0.105	1,190	6.2
LNCaP[A]	5,510	0.292	3,180	16.7
LNCaP[B]	5,336	0.283	3,171	16.7
LNCaP[B] BicaR	6,061	0.321	3,690	19.4
LNCaP[C]	6,137	0.325	3,749	19.7
LNCaP[C] RR	7,488	0.397	4,475	23.5
Group Average	6,106 ± 846	0.324 ± 0.045	3,653 ± 534	19.2 ± 2.8
Group Median	6,061	0.321	3,690	19.4
LAPC-4[A]	6,015	0.319	3,777	19.8
LAPC-4[B]	6,561	0.348	3,939	20.7
LAPC-4[B] EnzaR	6,250	0.331	3,893	20.4
Group Average	6,275 ± 274	0.332 ± 0.015	3,869 ± 83	20.3 ± 0.4
Group Median	6,250	0.331	3,893	20.4
C4-2	6,577	0.348	4,054	21.3
MR49F EnzaR	7,992	0.423	4,735	24.9
Group Average:	7,284	0.386	4,394	23.1

5.2.3 Top Mutated Genes

Genes harboring particularly high numbers of variants were identified, ranked and gene set enrichment analyses were performed on the top mutated genes. To calculate the variants per kbp for specific genes, the number of variants associated with the gene was divided by the length of its longest transcript as detailed in section **4.4.2**.

The top mutated genes in individual cell lines (data not shown) were not differing substantially from the top mutated genes in their cell line groups (**Table 17**), respectively. Compared to the average mutational burden across the exome described in **Table 16**, the top mutated genes harbored significantly higher numbers of variants with up to 33.5 variants per kbp in *ZNF717* in LNCaP cells. The top mutated genes contained many members of gene families such as human leukocyte antigen (HLA) genes, T cell receptor beta variable (TRBV) genes, killer cell immunoglobulin-like receptor (KIR) genes, immunoglobulin heavy variable (IGHV) genes, immunoglobulin light variable (IGHL) and olfactory receptor (OR) genes (**Table 17**). In addition, there were some additional genes that did not belong to the above-mentioned gene families. Among these genes were zinc finger protein 717 (ZNF717), FSHD region gene 1 and 2c (FRG1 & FRG2C), serine protease 3 (PRSS3), cell division cycle protein 27 (CDC27), keratin associated protein 9-1 (KRTAP9-1), late cornified envelope 4A (LCE4A) and Tektin-4 (TEKT4).

Three Gene Ontology (GO) databases (MF: molecular function, BP: biological process and CC: cellular component) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database were queried to identify gene set enrichments among the top 15 mutated genes.

For the combined dataset, the sets GO:MF 0023026 (*MHC class II protein complex binding proteins*), GO:BP 0006955 (*immune response*), GO:CC 0042613 (*constituents of the MHC class II protein complex*) and KEGG 05332 (*graft-versus-host disease*) exhibited the highest enrichment among the top 15 mutated genes.

DU145 and PC-3 cells exhibited the same results with almost identical p-values (**Table 17**). For the GO:MF and GO:CC databases, the results for DU145 and PC-3 cells were the same as for the combined dataset. The queries in the GO:BP and KEGG databases gave 0002503 (*peptide antigen assembly with MHC class II protein complex*) and 04612 (*antigen processing and presentation*), respectively.

The cell line groups LNCaP, LAPC-4 and C4-2/MR49F EnzaR all gave the same top results with similar p-values (**Table 17**) across all four queried databases. The sets GO:MF 0003823 (*antigen binding*), GO:BP 0002250 (*adaptive immune response*), GO:CC 0042101 (*T cell receptor complex*) and KEGG 05310 (*asthma*) were most enriched among the top 15 mutated genes in these cell line groups (**Table 17**).

Table 17: Gene set enrichment analysis of top mutated genes. Only coding and non-synonymous variants were considered. Genes were ranked by their average variants per kbp within the specified groups and gene set enrichment analyses were performed as described in section **4.4.2**. The lowest p-value results from three GO databases and from the KEGG database are shown. MF: Molecular Function, BP: Biological Process, CC: Cellular Component.

	Dataset, N=24		DU145, N=6		PC-3, N=8		LNCaP, N=5		LAPC-4, N=3		C4-2, N=2	
Rank	Symbol	v / kbp	Symbol	v / kbp	Symbol	v / kbp	Symbol	v / kbp	Symbol	v / kbp	Symbol	v / kbp
1	ZNF717	23.1	ZNF717	17.2	HLA-DRB5	17.8	ZNF717	33.5	ZNF717	31.1	ZNF717	29.9
2	HLA-DRB5	18.7	HLA-DRB5	17.1	ZNF717	16.3	OR4C3	22.8	OR4C3	24.2	OR4C3	24.2
3	OR4C3	10.3	KIR3DL1	15.5	PRSS3	7.5	HLA-DRB5	20.7	TRBV7-7	21.3	HLA-DRB5	19.3
4	KIR3DL1	8.8	KIR2DL1	12.7	FRG1	7.3	TRBV10-1	18.5	HLA-DRB5	20.1	TRBV6-5	17.2
5	HLA-DRB1	8.7	FRG1	8.3	FRG2C	6.6	TRBV6-5	18.1	TRBV3-1	19.7	IGLV5-45	16.5
6	PRSS3	8.6	PABPC3	8.1	KIR3DL1	6.5	TRBV7-7	18.0	IGHV5-51	19.5	TRBV7-7	15.9
7	FRG1	8.4	FRG2C	7.5	KIR2DL1	6.5	IGLV5-45	17.7	TRBV10-1	15.4	IGHV3-38	15.3
8	KIR2DL1	8.1	HLA-DRB1	6.6	HLA-DRB1	6.0	IGHV3-38	15.1	IGHV4-39	12.5	TRBV10-1	13.4
9	TRBV7-7	7.7	PRSS3	6.4	PABPC3	5.9	HLA-DRB1	12.6	TRBV6-5	12.3	HLA-DRB1	11.9
10	FRG2C	7.4	KRTAP9-1	6.2	KRTAP9-1	5.8	FRG1	11.3	HLA-DRB1	11.8	TRBV7-1	10.6
11	KRTAP9-1	7.0	CDC27	5.2	OR4A16	5.7	TRBV7-1	11.1	PRSS3	11.3	TRBV5-5	10.1
12	TRBV10-1	6.9	OR8U1	4.8	LCE4A	4.9	PRSS3	10.7	TRBV6-7	10.4	FRG1	9.9
13	TRBV6-5	6.7	KIR2DS4	4.5	KIR2DS4	4.0	TRBV6-7	9.9	IGLV5-45	10.1	PRSS3	9.8
14	IGLV5-45	6.3	KIR3DL2	4.5	TEKT4	3.7	IGHV4-31	9.4	IGHV1-3	9.3	IGHV4-31	9.4
15	CDC27	5.4	HNRNPCL1	4.4	KIR3DL2	3.7	TRBV4-2	8.8	KRTAP9-1	9.3	TRBV5-6	9.2
Database	Geneset ID	p_{adj}	Geneset ID	p_{adj}	Geneset ID	p_{adj}	Geneset ID	p_{adj}	Geneset ID	p_{adj}	Geneset ID	p_{adj}
GO:MF	0023026	$2.4E^{-3}$	0023026	$8.2E^{-3}$	0023026	$8.2E^{-3}$	0003823	$5.8E^{-4}$	0003823	$9.6E^{-6}$	0003823	$5.8E^{-4}$
GO:BP	0006955	$6.8E^{-4}$	0002503	$1.5E^{-2}$	0002503	$1.5E^{-2}$	0002250	$1.1E^{-10}$	0002250	$2.9E^{-11}$	0002250	$1.1E^{-10}$
GO:CC	0042613	$7.2E^{-4}$	0042613	$1.8E^{-3}$	0042613	$1.8E^{-3}$	0042101	$2.9E^{-8}$	0042101	$7.9E^{-7}$	0042101	$2.9E^{-8}$
KEGG	05332	$1.8E^{-6}$	04612	$6.6E^{-8}$	04612	$1.0E^{-7}$	05310	$2.2E^{-2}$	05310	$2.8E^{-2}$	05310	$2.2E^{-2}$

5.2.4 Similarity and Dissimilarity between Cell Lines

To study the similarity between pairs and groups of cell lines, unique, partially shared and shared variants were identified as described in section **4.4.3**. The numbers of such variants depend on which cell lines are included in the analysis. Three analyses were done. First, unique and shared variants were identified among pairs of parental and their derivative treatment-resistant cell lines (section **5.2.4.1**). Second, the parental cell lines (excluding their derivatives) were studied (section **5.2.4.2**), and third, the analysis was performed on complete cell line groups (treatment-naïve parental cell lines plus their derivatives, section **5.2.4.3**). Finally, phylogenetic trees (section **5.2.4.4**) of complete cell line groups were inferred from the WES data as described in section **4.4.4**.

5.2.4.1 Unique and Shared Variants among Cell Line Pairs

2-set Venn diagrams for each pair of parental and treatment-resistant derivative cell lines were prepared (**Figure 8**). With some exceptions, high proportions (up to 86.7%) of variants were shared between parental lines and their derivatives. The lowest proportion of shared variants (42.9%) was found between C4-2 and MR49F EnzaR cells. The 10,220 distinct variants comprised of 2,228 variants unique to C4-2 cells, 3,643 variants unique to MR49F EnzaR cells and 4,349 shared variants (**Figure 8**). While C4-2 and MR49F EnzaR cells share the same parent cell line (LNCaP), they are not actually derivatives of each other. In addition, these cell lines were subjected to several cycles of *in vivo* culture.

With 1,170 unique variants, PC-3^{D} CDDPR had more than three times as many unique variants than PC-3^{D}, which had 310. The high number of unique variants found in PC-3^{D} CDDPR cells may be explained by the DNA-damaging treatment with CDDP. However, neither CDDP-treated DU145^{A} CDDPR nor PC-3^{A} CDDPR exhibited increased numbers of unique variants compared to their parental counterparts, respectively (**Figure 8**).

Similarly, LNCaP^{C} RR cells harbored almost four times as many (1,931) unique variants as LNCaP^{C} (580), which might be attributable to the DNA-damaging treatment with ionizing radiation that these cells were subjected to. However, neither DU145^{C} RR nor PC-3^{C} RR exhibited a notable difference in the number of unique variants compared to their parental cell lines, respectively (**Figure 8**).

Two other cell line pairs exhibited an imbalance regarding the number of unique variants. DU145^{B} DTXR had about 1.5-fold more unique variants than DU145^{B} (1,459 vs 933) and LNCaP^{B} BicaR had about 1.9 times as many unique variants as LNCaP^{B} (1,539 vs 814), respectively.

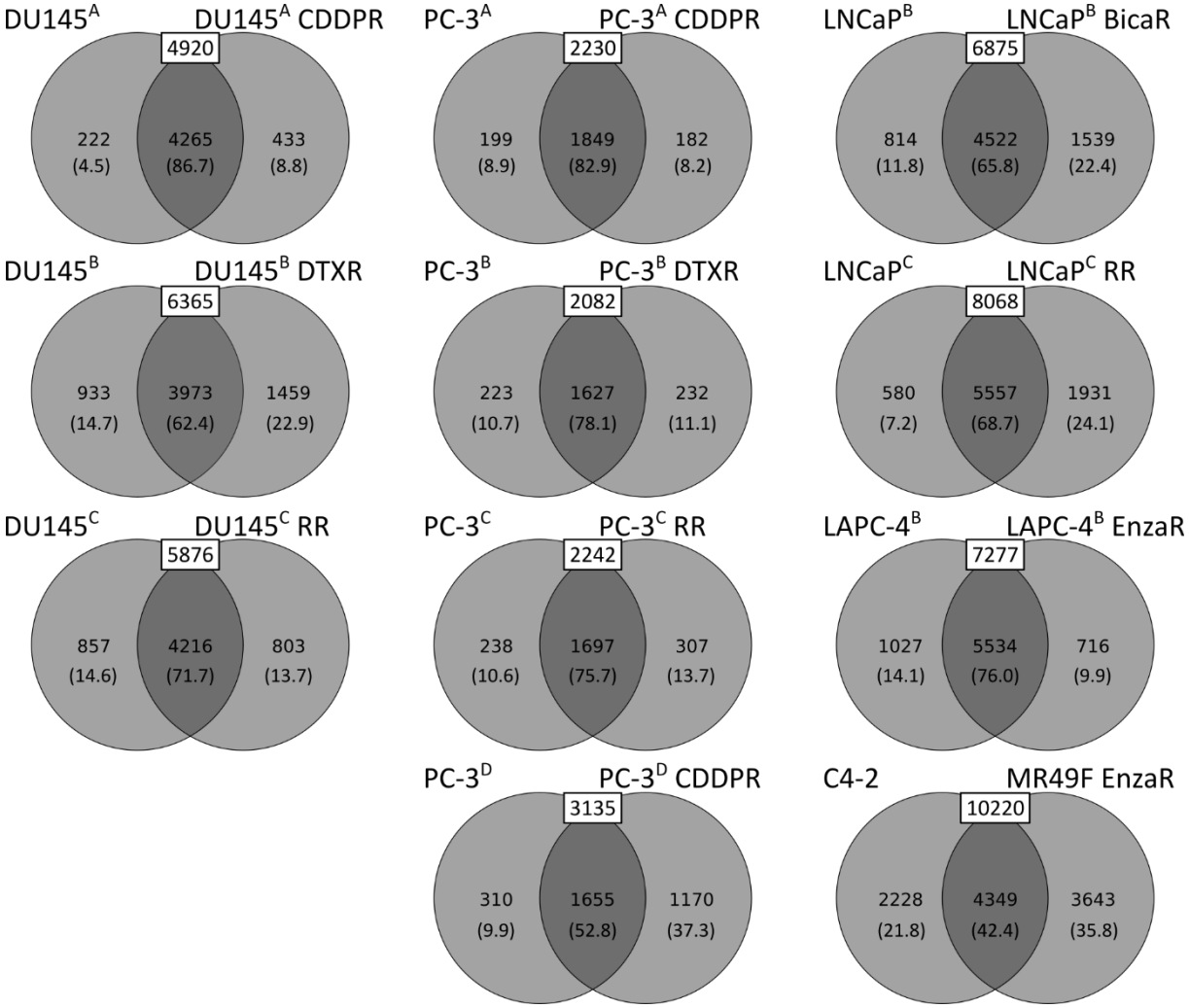

Figure 8: Venn diagrams of variants in cell line pairs. Only coding and non-synonymous variants were considered. The total number of distinct variants is displayed in white boxes on each diagram. Values in brackets indicate the % of total distinct variants.

5.2.4.2 Unique, Partially Shared and Shared Variants among Parental Cell Lines

Venn diagrams for parental DU145 (N=3), PC-3 (N=4), LNCaP (N=3) and LAPC-4 (N=2) cell lines were prepared (**Figure 9**). Among the three DU145 cell lines, a total of 6,197 distinct variants existed. Of these, 3,799 (61.3%) were shared among all lines, 671 (10.8%) were partially shared and 1,727 (27.9%) were unique to one line. DU145^{A} cells harbored substantially less unique variants (143) than DU145^{B} (761) and DU145^{C} (823) and shared twice as many variants with DU145^{C} (387) than with DU145^{B} (158). PC-3 cells generally harbored lower numbers of variants than other cell lines in the dataset (**Table 16**). The total number of distinct variants in parental PC-3 cell lines was 2,545. Of these, 1,501 (59%) were shared among all four, 203 (8%) among three and 172 (6.7%) among two cell lines, respectively. Furthermore, 669 (26.3%) variants in PC-3 cell lines were unique, comprising of 232, 100, 143 and 194 found in PC-3^{A}, PC-3^{B}, PC-3^{C} and PC-3^{D}, respectively (**Figure 9**). In LNCaP cell lines, a total of 7,643 distinct variants was found, comprising of 4,357 (57%) shared, 626 (8.2%) partially shared and 2,660 (34.8%) unique variants, respectively (**Figure 9**). LNCaP^{C} (1,554) harbored more unique variants than LNCaP^{A} (552) and LNCaP^{B} (554) combined. Furthermore, LNCaP^{A} and LNCaP^{B} shared 400 variants not present in LNCaP^{C}, while LNCaP^{C} shared only 201 variants with LNCaP^{A} and 25 variants with LNCaP^{B}, respectively. Based on these data, one could argue that LNCaP^{A} and LNCaP^{B} are genetically closer to each other than they are to LNCaP^{C}. The two treatment-naïve LAPC-4 cell lines harbored a total of 7,285 distinct variants of which 5,291 (72.6%) were shared and 1,994 (27.4%) were unique, respectively.

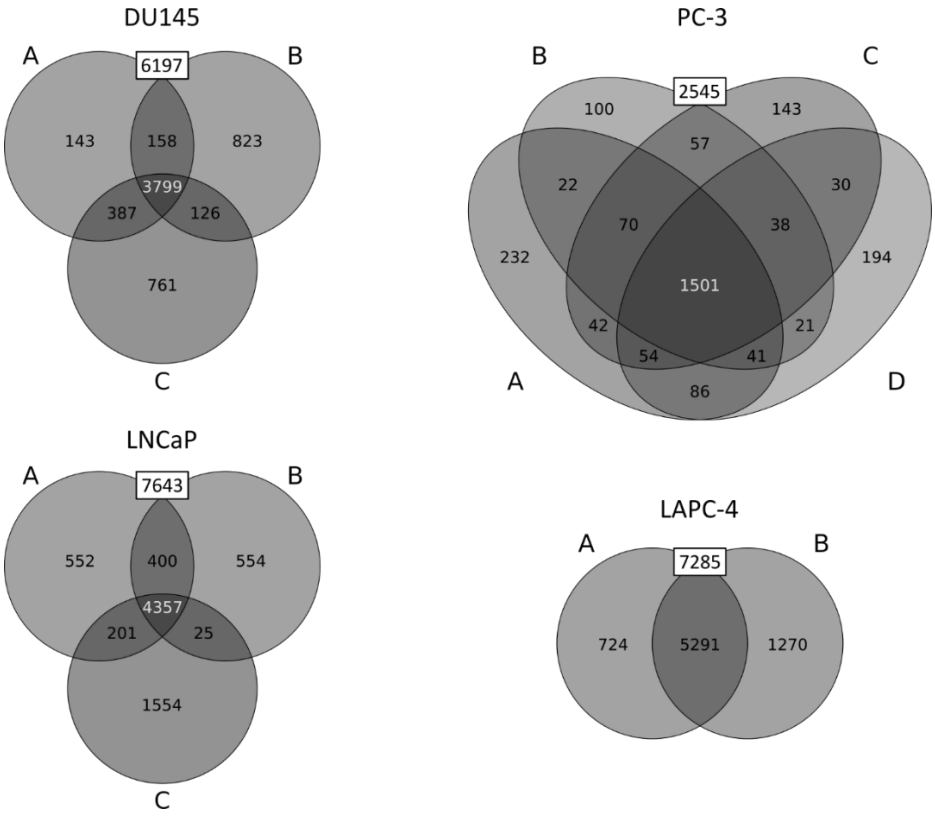

Figure 9: Venn diagrams of variants in parental cell lines. Only the treatment-naïve parental cell lines were included in the analysis and only coding and non-synonymous variant were considered. The total number of distinct variants is displayed in white boxes on each diagram.

5.2.4.3 Unique, Partially Shared and Shared Variants among Complete Cell Line Groups

Sets and their intersections in **Figure 8** and **Figure 9** were displayed as Venn diagrams. Since Venn diagrams for more than four sets would be difficult to interpret visually, the UpSetR package for R (section **4.4.3**) was used to summarize unique, partially shared and shared variants among whole cell line groups (**Figure 10**).

3,600, 1,294, 4,144 and 5,116 shared variants were identified among the DU145 (N=6), PC-3 (N=8), LNCaP (N=5) and LAPC-4 (N=3) cell line groups, respectively (**Figure 10**). On average, these numbers were 6.8% lower than the numbers of shared variants identified considering only parental cell lines (**Figure 9**). As expected, also the numbers of unique variants were reduced. For example, parental PC-3 cell lines harbored now 58 (PC-3^{A}), 39 (PC-3^{B}), 88 (PC-3^{C}) and 89 (PC-3^{D}) unique variants (**Figure 10**). This was less than half compared to the previously mentioned numbers (section **5.2.4.2** and **Figure 9**), where the treatment-resistant derivative cell lines were not considered.

Generally, unique and shared variants comprised the largest set-intersections (**Figure 10** blue and red colored bars). However, some set-intersections of partially shared variants had a considerable size as well. For example, 1,207 variants were shared between $LNCaP^{C}$ and $LNCaP^{C}$ RR but not present among the other LNCaP cell lines. PC-3^{A} and PC-3^{A} CDDPR cells harbored 127 shared variants not present in the other PC-3 cell lines. Finally, $DU145^{A}$, $DU145^{A}$ CDDPR, $DU145^{C}$ and $DU145^{C}$ RR cells shared 207 variants that were not present in $DU145^{B}$ and $DU145^{B}$ DTXR (**Figure 10**).

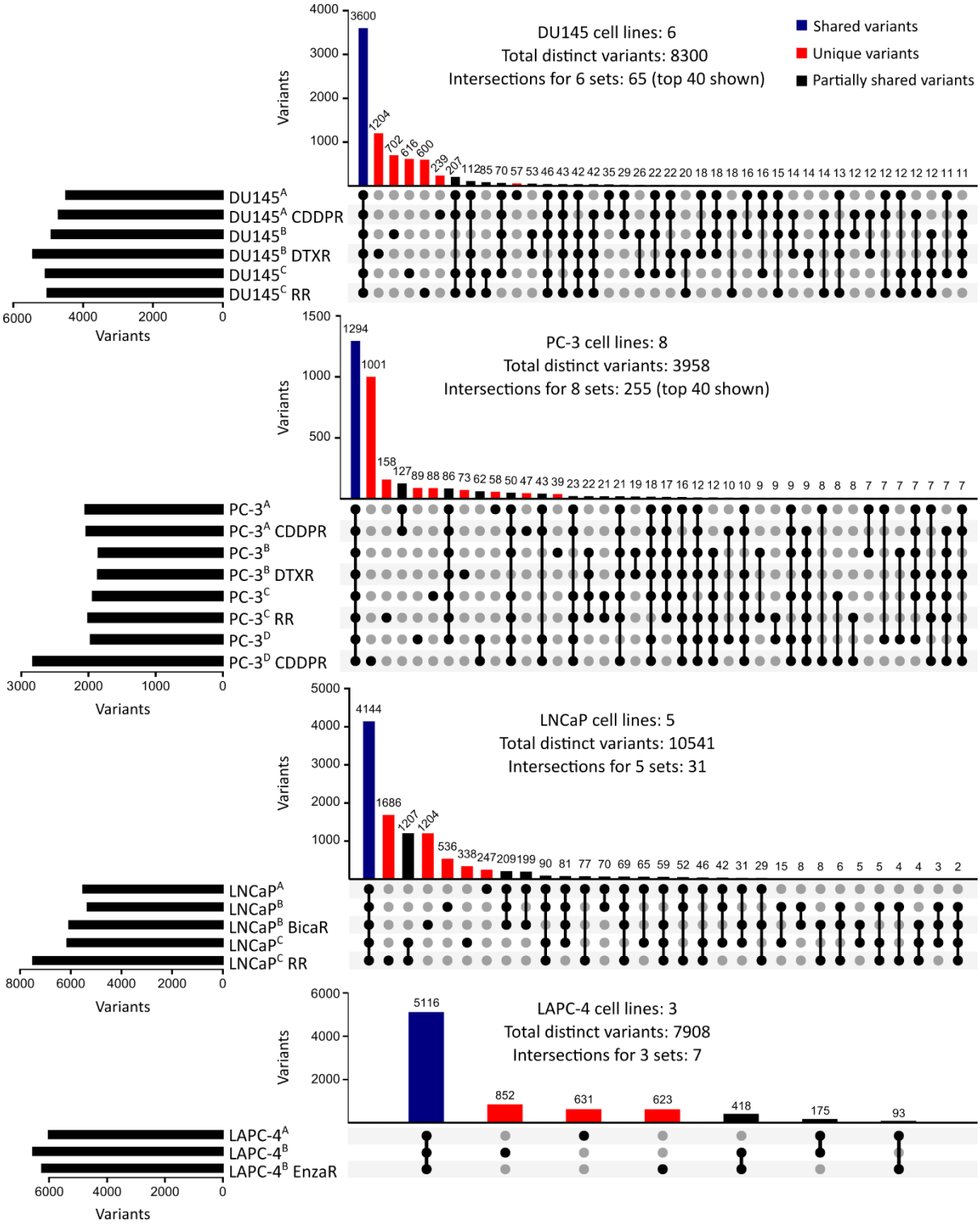

Figure 10: Unique, partially shared and shared variants among complete cell line groups. The plots were generated with the UpSetR package for R. Only coding and non-synonymous variants were considered. The total number of distinct variants is displayed on each graph. For DU145 and PC-3 cell lines, only the top 40 set-intersections are displayed. Intersections corresponding to shared, partially shared and unique variants are colored blue, black and red, respectively.

5.2.4.4 Phylogenetic Analysis Reveals Cell Line Ancestry

Based on WGS/WES data, phylogenetic relations among groups of samples can be inferred. Phylogenetic trees for cell line groups were generated as described (section **4.4.4**). The algorithm generated two-leafed branches for all pairs of parental and treatment-resistant cell lines except LAPC-4^{B} and LAPC-4^{B} EnzaR (**Figure 11**). The nodes of these branches represent the 'original' parental cell lines, before the cultures were split. The horizontal distance on the phylogram corresponds to the relative genetic distance between samples. Treatment-resistant cell lines as well as their age-matched counterparts were generally more distant to the 'original' parental cell line than the 'original' parental cell lines were to the common ancestors of the cell line groups. For LNCaP cells, LNCaP^{A} were most similar to the common ancestor of the cell line group, followed by LNCaP^{B}, LNCaP^{B} BicaR, LNCaP^{C} and LNCaP^{C} RR in that order. Furthermore, the 'original' parental LNCaP^{B} cell line was located significantly closer to the common ancestor of the cell line group than the 'original' LNCaP^{C} cell line. This can be well correlated to the number of variants presented on **Figure 10**. While LNCaP^{C} and LNCaP^{C} RR shared 1,207 variants that are not present in any other LNCaP cell line, LNCaP^{B} and LNCaP^{B} DTXR merely shared eight such variants (**Figure 10**). Furthermore, the algorithm inferred a common ancestor for the original PC-3^{A} and the original PC-3^{D} cell lines (**Figure 11**, green dot). This ancestor and its children were genetically more distant to the common ancestor of all PC-3 cell lines than the original PC-3^{B} and the original PC-3^{C} cell line. Interestingly, PC-3^{A}, PC-3^{A} CDDPR, PC-3^{D} and PC-3^{D} CDDPR shared 48 variants that were not present in the other four PC-3 cell lines (**Figure 10**).

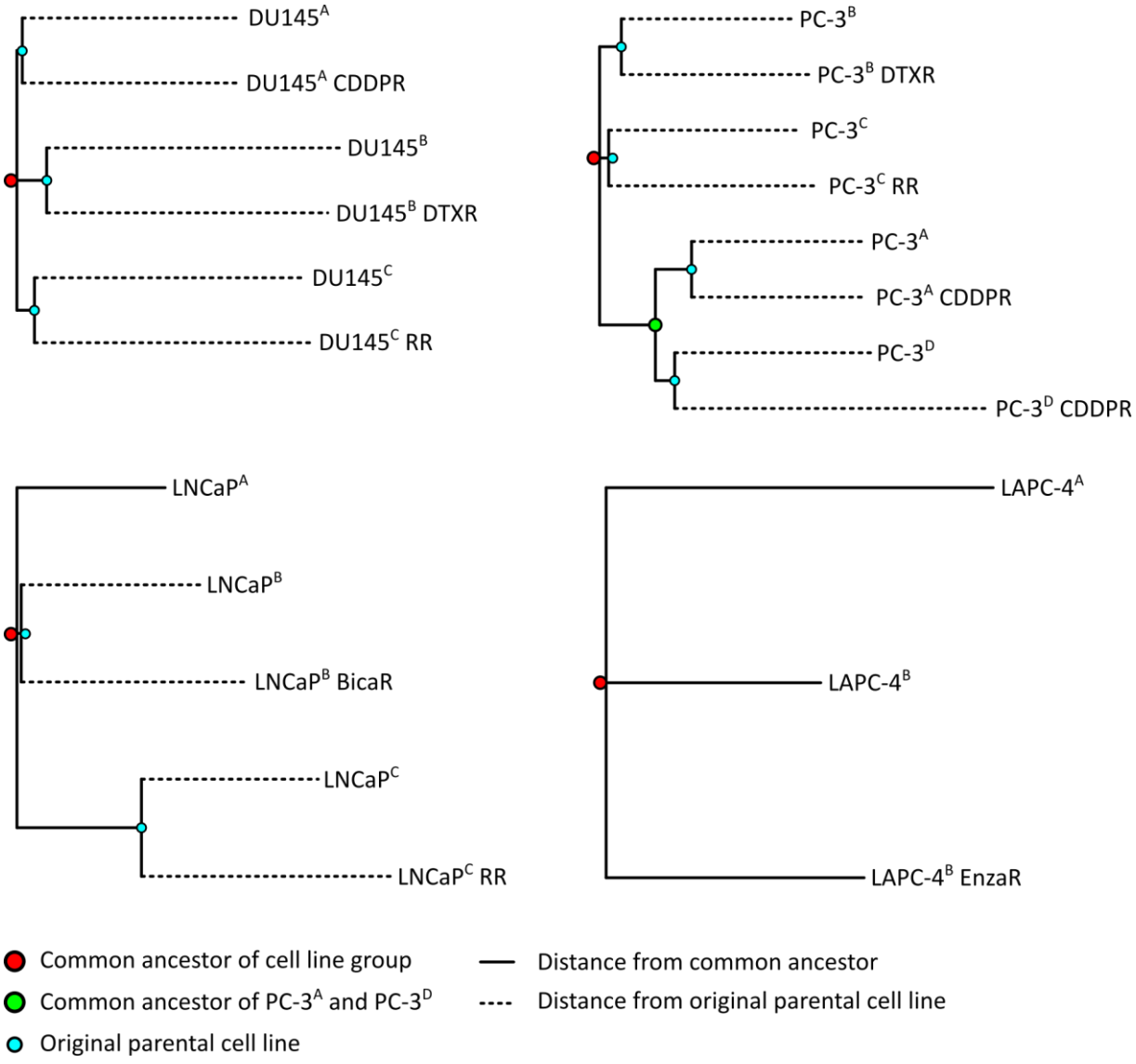

Figure 11: Phylogenetic Trees of Cell Line Groups. Phylogenetic trees were inferred as described in section **4.4.4**. All variant types and classes were considered. Red dots symbolize the hypothetical common ancestor of the cell line group, the green dot symbolizes the common ancestor of PC-3^{A} and PC-3^{D} and the turquoise dots symbolize the original parental cell line, respectively. Normal and dashed horizontal lines represent the genetic distance from the next common ancestor or from the original parental cell line, respectively.

5.2.5 The Variant Allele Frequency among Shared Variants

As described in section **1.6.6**, the VAF has to be cautiously interpreted in this dataset. The cell lines are highly aneuploid and their population may be heterogenic. It is best to think about the VAF as the average prevalence of the variant within the genetic material of the culture. The distribution of the VAF among shared and unique variants was studied. 3,600 variants were shared in all six DU145 cell lines and 892 (24.7%) of these were homozygous in the entire population (VAF = 100%, **Figure 12** upper panel). Three quarters of shared variants were somewhat evenly distributed between VAFs of 10% and 70%. Only a small number of variants had VAFs greater than 70% but lower than 100%. As mentioned, the remaining quarter had a VAF of 100%. In the eight PC-3 cell lines, a total of 1,294 shared variants existed (**Figure 12**). Of these, 584 (45.1%) were homozygous within the entire cell population. 4,144 shared variants existed among the five LNCaP cell lines and 353 (8.5%) of those had a VAF of 100%. A large part of the shared variants in LNCaP cells had a VAF above 40 but below 60% (**Figure 12**). The absolute numbers of shared variants as well as the distributions implied by the graphs and violin plots were strikingly similar between the LNCaP cell group and the group consisting of C4-2 and MR49F EnzaR cells. A manifest explanation for this is that both C4-2 and MR49F EnzaR cells have LNCaP as their parent. Thus, the shared variants shown on the two graphs are likely the same shared variants to a large extend (**Figure 12**).
VAF data were not available for LAPC-4^{B} and LAPC-4^{B} EnzaR. Thus, the VAF distribution among shared variants in LAPC-4^{A} is shown as a representation for the whole cell line group. LAPC-4 cells had a total of 5,116 shared variants. Similar as for DU145 cell lines, the majority of shared variants was evenly spread between VAFs of 10% and 70%. Only a small number of variants had VAFs greater than 70% but lower than 100% and 282 (5.5%) variants had a VAF of 100%, suggesting homozygosity (**Figure 12**).

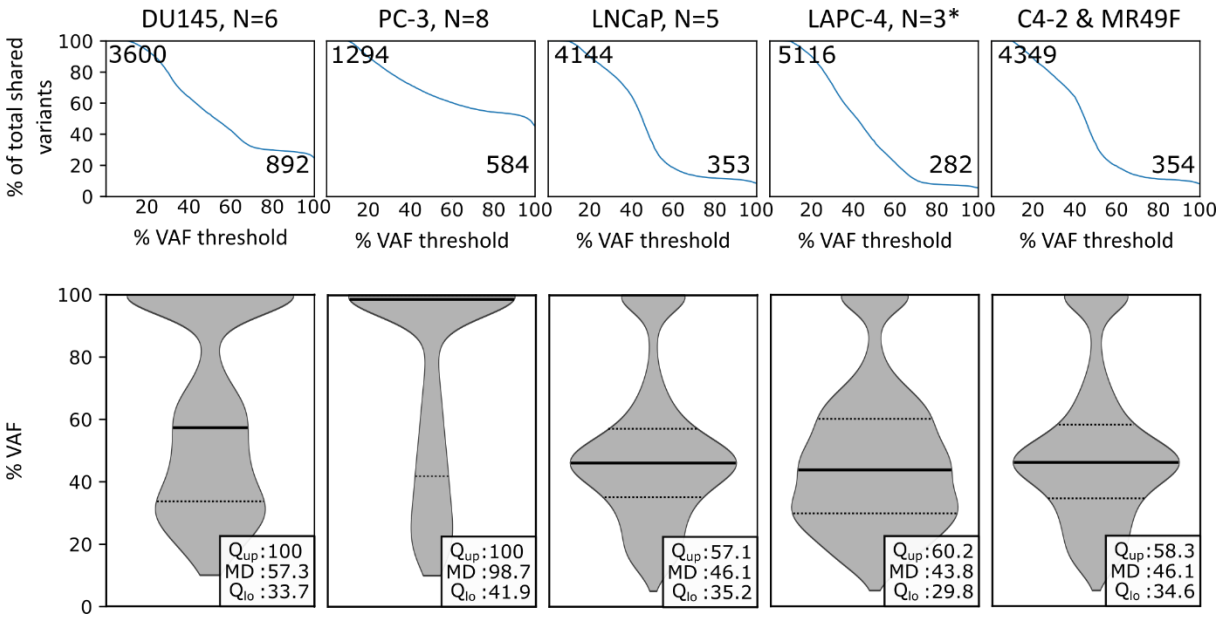

Figure 12: VAF distribution among shared variants. Only coding and non-synonymous variants were considered. The upper panel shows the number of shared variants as a function of a VAF threshold applied to the data. The data are normalized to a VAF threshold of 0% (no VAF threshold applied). The upper left corners of the graphs are labelled with the absolute number of shared variants (no VAF threshold applied) and the lower right corners of the graphs are labelled with the absolute number of shared variants with a VAF threshold of 100% applied. The lower panel shows the same distribution as violin plots. The median and the quartiles are shown in the lower right corners of the graphs. The asterisk indicates that although the absolute numbers of shared variants were based on all three LAPC-4 cells, VAF data was only available for LAPC-4^{A} but not for the other two cell lines. The graph and violin plot are based on these data.

5.2.6 The Variant Allele Frequency among Unique Variants

The distribution of the VAF among the unique variants is summarized on **Figure 13**. As mentioned, this analysis included the resistant cell lines and therefore, the absolute numbers of unique variants shown on **Figure 13** are lower than the numbers shown on the Venn diagrams in **Figure 9** and **Figure 8**.

The absolute number of unique variants in the cell lines was highly variable within the dataset as well as within cell line groups. For example, DU145^{A} had 57 while DU145^{B} DTXR had 1,204 unique variants, respectively. There were very few unique variants with VAFs above 50%. For example, 80% of 57 unique variants in DU145^{A} cells had a VAF between 10% (which was the cut-off) and 20% and only three variants were homozygous with a VAF of 100%. This pattern was essentially similar in all cell lines (**Figure 13**). Also, cell lines with more than 1,000 total unique variants (DU145^{B} DTXR, PC-3^{D} CDDPR, LNCaP^{B} BicaR and LNCaP^{C} RR) had only single digit numbers of homozygous unique variants with a VAF of 100% (**Figure 13**). C4-2 and MR49F EnzaR cells shall not be mentioned here because the number of unique variants in these cells is likely inflated as MR49F EnzaR were not derived from C4-2 cells and the group consists of only two cell lines.

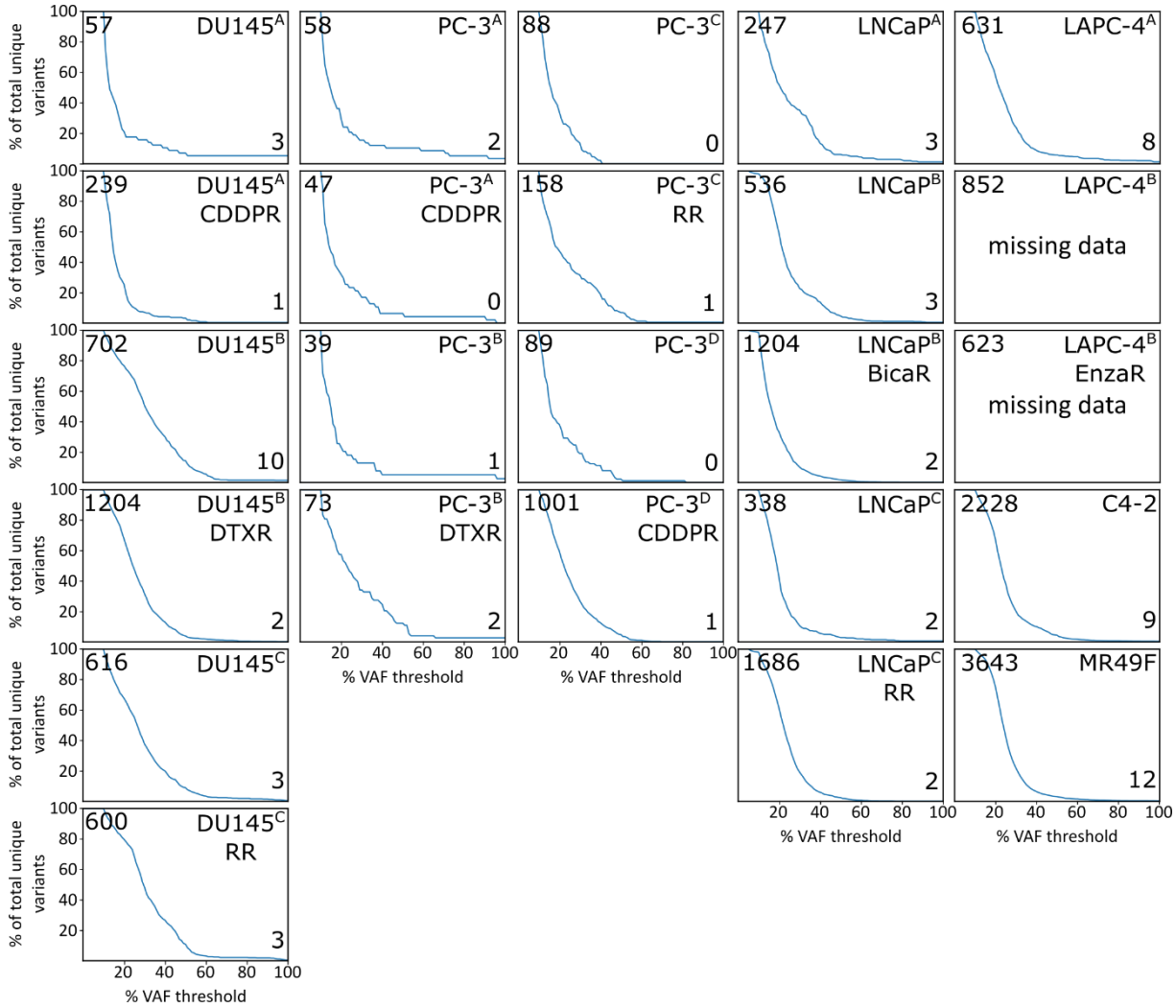

Figure 13: VAF distribution among unique variants. Only coding and non-synonymous variants were considered. Variants unique to exactly one cell line within a group are plotted as a function of a VAF threshold applied to the data. The data are normalized to a VAF threshold of 0% (no VAF threshold applied). The upper left corners of the graphs are labelled with the absolute number of unique variants (no VAF threshold applied) and the lower right corners of the graphs are labelled with the absolute number of unique variants with a VAF threshold of 100% applied. Note that there are very few strictly unique variants with a VAF of 100%. VAF data for LAPC-4^{B} and LAPC-4^{B} EnzaR cell lines was not available.

5.2.7 Mutational Signatures

Mutational signatures[131,132] were extracted from the data as described in section **4.4.5**. As discussed in the introduction (section **1.6.7**) mutational signatures look at the proportion of certain SBS events and take the adjacent 3' and 5' bases into account. The sum of all internal and external mutagenic processes acting on a sample produces the signature. Mutational signatures were extracted from all SNVs and from shared SNVs (**Figure 14**) as well as from unique SNVs (**Figure 15**).

First, signatures based on all SNVs (shared + partially shared + unique) were generated (**Figure 14 a**). Many of the resulting cell line signatures were most similar to SBS5. In fact, SBS5 was the most similar signature to all DU145 cell signatures, to the signatures of LNCaP^{A}, LNCaP^{B} and LNCaP^{C} as well as to the signatures of all PC-3 cell lines except PC-3^{D} CDDPR. The exact mutational process responsible for SBS5 is unknown, however, SBS5 correlates with patient/sample age[134]. Further, LNCaP^{B} BicaR, LNCaP^{C} RR, C4-2 and MR49F EnzaR cells exhibited signatures that were most similar to SBS44, which is associated with defective MMR[134]. The signatures of all three LAPC-4 cells were most similar to SBS15, which is also associated with defective MMR[134]. Finally, the signature of PC-3^{D} CDDPR cells was most similar to SBS31, which is associated with Pt treatment[134].

Next, mutational signatures from shared SNVs, excluding partially shared and unique SNVs, were generated (**Figure 14 b**). Except for the LAPC-4 cell line group, in which the signature was most similar to SBS15, the signatures of all other cell line groups were most similar to SBS5.

Finally, mutational signatures from unique SNVs were generated (**Figure 15 a**), yielding markedly different signatures as compared to the signatures shown on **Figure 14**. The signatures of DU145^{B}, DU145^{B} DTXR, DU145^{C} RR, all LNCaP cells and LAPC-4^{B} were now most similar to SBS44, indicating defective MMR. The signatures of DU145^{A} CDDPR, LAPC-4^{A} and LAPC-4^{B} EnzaR were now most similar to SBS6, which is another signature indicating defective MMR[134]. The signatures of PC-3^{B} DTXR and PC-3^{C} RR were now most similar to SBS7a (exposure to ultraviolet light) and SBS3 (defective HRR), respectively. The signature of DU145^{C} was still most similar to SBS5 and the result for C4-2 and MR49F EnzaR cells was still SBS44. The signature of PC-3^{D} CDDPR cells was still most similar to SBS31, however, the cosine similarity to SBS31 (0.958) was now significantly higher (compare **Figure 14 a** and **Figure 15 a**). Interestingly, the signatures of DU145^{A} was now most similar to SBS55 and the signatures of PC-3^{A}, PC-3^{A} CDDPR, PC-3^{B}, PC-3^{C} and PC-3^{D} were now most similar to SBS60, both of which are categorized as signatures possibly caused by sequencing artefacts[134] (**Figure 15 a**).

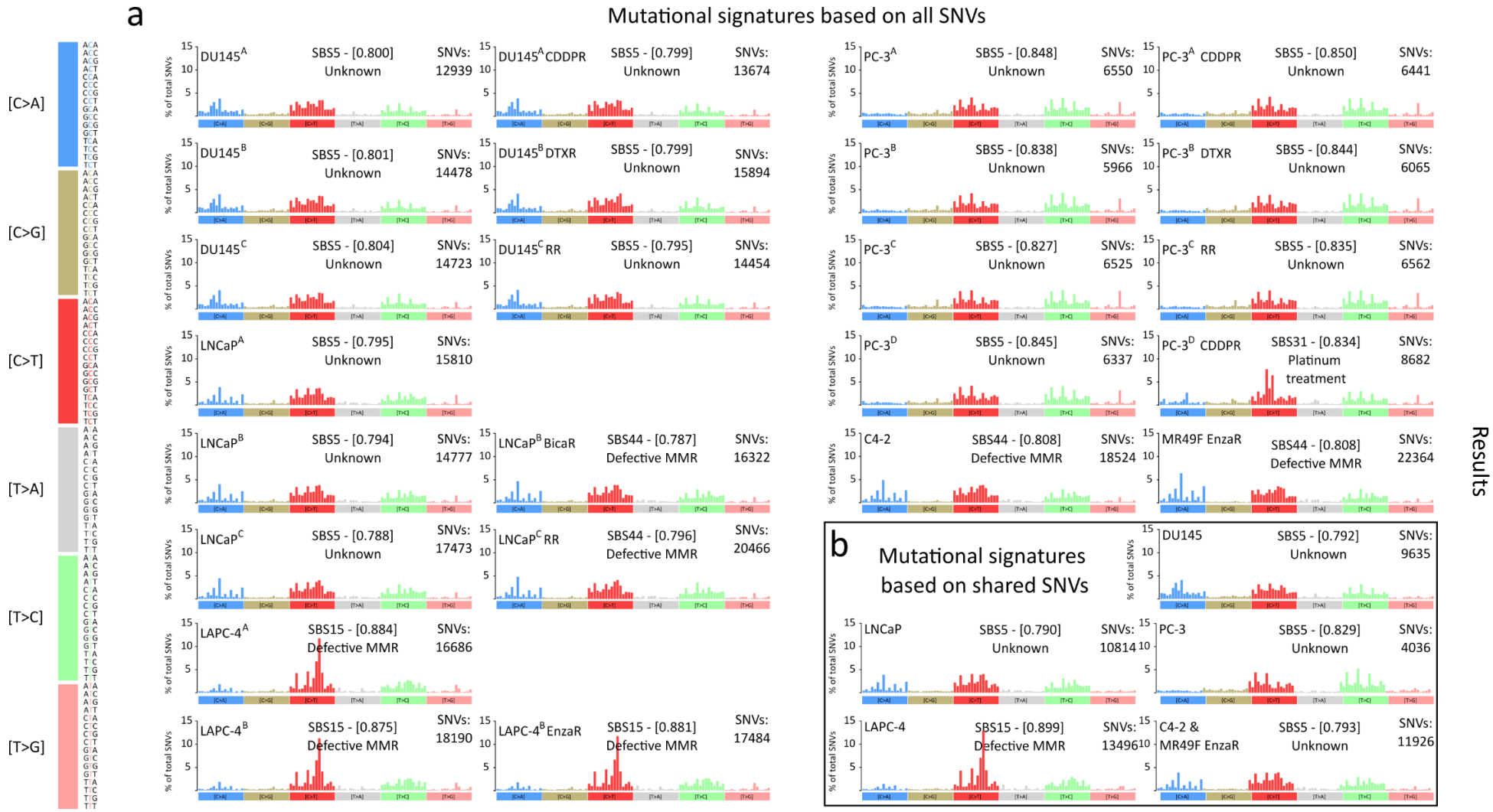

Figure 14: Mutational signatures based on all SNVs. On the left, a legend describes the color code and position of all 96 possible SBS events. (a) shows mutational signatures based on all SNVs and (b) shows the signatures of cell line groups based on shared SNVs. Each plot is labelled with the cell line (upper left), the best matching validated signature with the corresponding cosine similarity in brackets and its proposed etiology beneath (upper middle) and the number of SNVs from which the signature was generated (upper right).

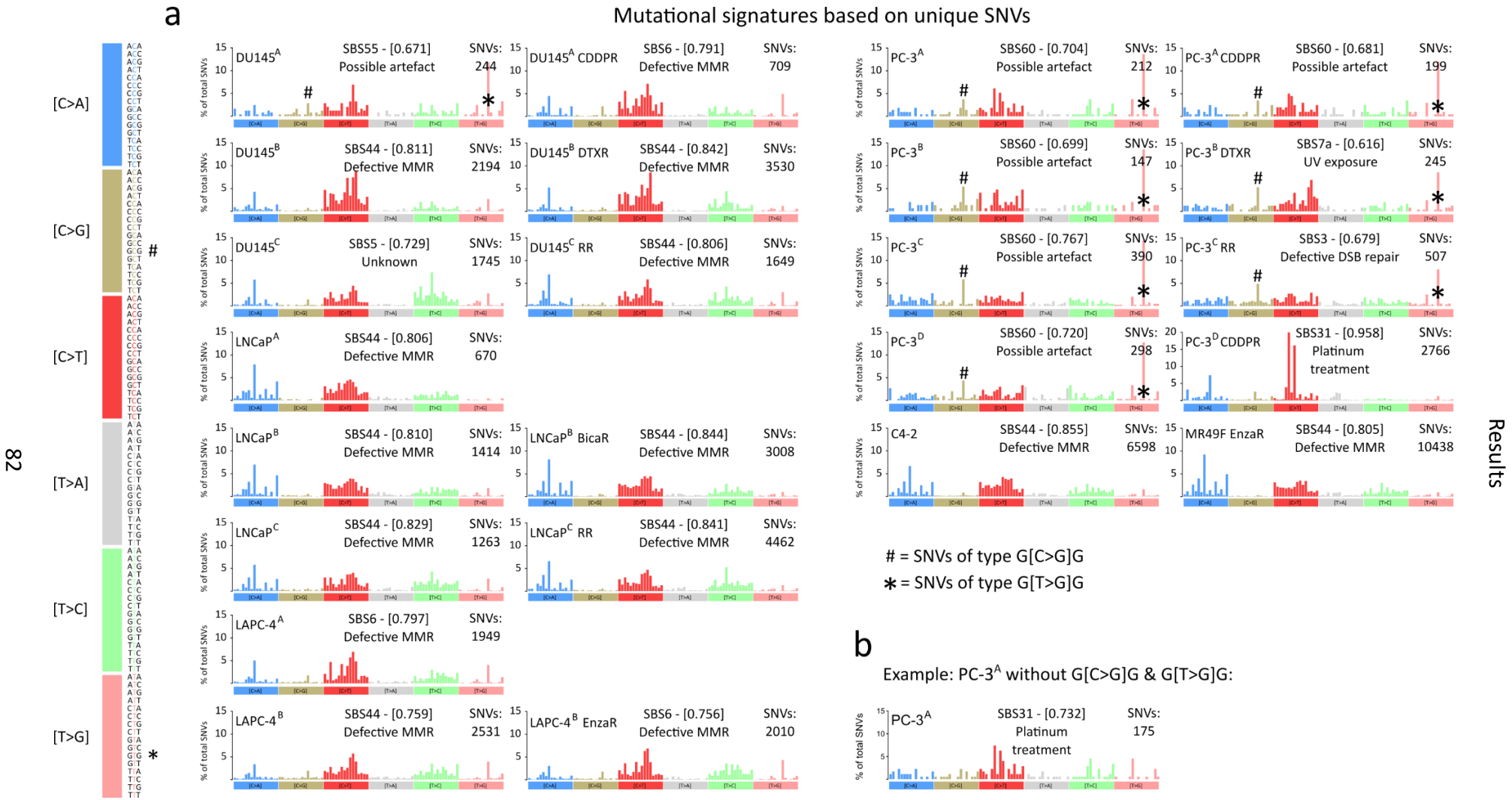

Figure 15: Mutational signatures based on unique SNVs. (a) Each plot is labelled with the cell line (upper left), the best matching validated signature with the corresponding cosine similarity in brackets and its proposed etiology beneath (upper middle) and the number of SNVs from which the signature was generated (upper right). G[C>G]G events are highlighted by # and G[T>G]G events are highlighted by *, respectively. (b) The signature of PC-3^{A} cells (G[C>G]G and G[T>G]G events excluded) is shown.

The signatures of the cell lines that were most similar to SBS55 or SBS60 were characterized by two particularly common SBS events: G[C>G]G and G[T>G]G (**Figure 15 a**, asterisk and hash symbols). **Table 18** lists the proportions of these events among all and among unique SNVs. The Spearman correlation coefficient (Spearman r) for all G[C>G]G + G[T>G]G events versus all SNVs was 0.58, while the average Spearman r for all possible SBS events was 0.72. Furthermore, the Spearman r for unique G[C>G]G + G[T>G]G events versus unique SNVs was 0.72 versus an average Spearman r of 0.81. Hence, G[C>G]G and G[T>G]G events were less correlated with the total number SNVs than the average SBS event. PC-3 cells, which generally had low variant numbers, had comparable numbers of G[C>G]G and G[T>G]G events as DU145 cells and other cell lines with high variant numbers. The proportions of G[C>G]G and G[T>G]G events was highest in samples with low variant numbers (**Table 18**). Additionally, the two events were further enriched in samples with less unique SNVs (**Table 18**, bold numbers).

Table 18: Prevalence of G[C>G]G and G[T>G]G events.

Cell line	All SNVs	All G[C>G]G + G[T>G]G	G[C>G]G + G[T>G]G % of all	Unique SNVs	Unique G[C>G]G + G[T>G]G	G[C>G]G + G[T>G]G % of unique
DU145^{A}	12939	311	2.4	244	36	**14.8**
DU145^{A} CDDPR	13674	319	2.3	709	51	**7.2**
DU145^{B}	14478	359	2.5	2194	75	3.4
DU145^{B} DTXR	15894	359	2.3	3530	73	2.1
DU145^{C}	14723	355	2.4	1745	74	4.2
DU145^{C} RR	14454	342	2.4	1649	64	3.9
PC-3^{A}	6550	291	4.4	212	37	**17.5**
PC-3^{A} CDDPR	6441	272	4.2	199	31	**15.6**
PC-3^{B}	5966	279	4.7	147	28	**19.0**
PC-3^{B} DTXR	6065	293	4.8	245	34	**13.9**
PC-3^{C}	6525	393	6.0	390	79	**20.3**
PC-3^{C} RR	6562	361	5.5	507	66	**13.0**
PC-3^{D}	6337	294	4.6	298	51	**17.1**
PC-3^{D} CDDPR	8682	274	3.2	2766	51	1.8
LNCaP^{A}	15810	525	3.3	670	4	0.6
LNCaP^{B}	14777	352	2.4	1414	34	2.4
LNCaP^{B} BicaR	16322	312	1.9	3008	46	1.5
LNCaP^{C}	17473	364	2.1	1263	42	3.3
LNCaP^{C} RR	20466	397	1.9	4462	86	1.9
LAPC-4^{A}	16686	412	2.5	1949	113	5.8
LAPC-4^{B}	18190	463	2.5	2531	149	5.9
LAPC-4^{B} EnzaR	17484	428	2.4	2010	123	6.1
C4-2	18524	328	1.8	6598	160	2.4
MR49F EnzaR	22364	300	1.3	10438	132	1.3

As elaborated, G[C>G]G and G[T>G]G events may partially represent sequencing artefacts. Therefore, they were excluded from the unique SNVs and mutational signatures were generated again to study the effect of their exclusion in the relevant cell lines (**Table 19**). After exclusion of G[C>G]G and G[T>G]G events, the signatures of PC-3^{A} CDDPR, PC-3^{B}, PC-3^{C}, PC-3^{C} RR and PC-3^{D} were most similar to SBS5, which is the same result obtained from the signatures based on all SNVs. The signature of PC-3^{B} DTXR cells was still most similar to SBS7a, which indicates genetic damage caused by UV irradiation[134]. The signature of DU145^{A} was now most similar to SBS44 (**Table 19**), and was now matching the results obtained for DU145^{A} CDDPR, DU145^{B}, DU145^{B} DTXR and DU145^{C} RR (see **Figure 15 a**).

Surprisingly, while the signature of PC-3^{A} CDDPR was now most similar to SBS5 (as mentioned), the signature of PC-3^{A} was now most similar to SBS31, indicating exposure to Pt compounds (see **Table 19** and compare **Figure 15 b** to the signature of PC-3^{D} CDDPR cells seen on **Figure 15 a**). The origin of the SBS31 signature in PC-3^{A} cells is not known. PC-3^{A} may have been exposed to CDDP as it is simultaneously cultured with PC-3^{A} CDDPR. Furthermore, a sample mix-up is conceivable as well.

Table 19: Effect of G[C>G]G and G[T>G]G exclusion on signature calculation. C. sim.<. cosine similarity, ex.: excluding.

Cell line	Unique SNVs	Best match signature – [C. sim.]	Description	Unique SNVs ex. G[C>G]G ex. G[T>G]G	Best match signature – [C. sim.]	Description
DU145^{A}	244	SBS55 – [0.671]	Possible artefact	208 (-15%)	SBS44 – [0.713]	Defective MMR
PC-3^{A}	212	SBS60 – [0.704]	Possible artefact	175 (-17%)	SBS31 – [0.732]	Pt treatment
PC-3^{A} CDDPR	199	SBS60 – [0.681]	Possible artefact	168 (-16%)	SBS5 – [0.694]	Unknown
PC-3^{B}	147	SBS60 – [0.699]	Possible artefact	119 (-19%)	SBS5 – [0.673]	Unknown
PC-3^{B} DTXR	245	SBS7a – [0.616]	UV exposure	211 (-14%)	SBS7a – [0.758]	UV exposure
PC-3^{C}	390	SBS60 – [0.767]	Possible artefact	311 (-20%)	SBS5 – [0.803]	Unknown
PC-3^{C} RR	507	SBS3 – [0.679]	Defective DSB repair	441 (-13%)	SBS5 – [0.825]	Unknown
PC-3^{D}	298	SBS60 – [0.720]	Possible artefact	247 (-17%)	SBS5 – [0.735]	Unknown

Next, cell lines were clustered by their mutational signatures (**Figure 16**). Based on the signatures generated from all SNVs (**Figure 16 a**), unsupervised hierarchical clustering separated the cell line groups. Only PC-3^{D} CDDPR was separated from the remaining PC-3 cell lines. The LNCaP derivatives C4-2 and MR49F EnzaR clustered together with the LNCaP cell lines. The structure of the dendrogram suggests that the pronounced differences were found between cell line groups and that the signatures within cell line groups were comparatively similar. Treatment-resistant cell lines and their parental counterparts clustered next to each other in most cases (**Figure 16 a**).

The clustering based on the signatures generated from unique SNVs (**Figure 16 b**) did not completely order the cell lines by their groups. Only PC-3 cell lines (except PC-3^{D} CDDPR) were contained within one clade. The other major clade on the right side of the dendrogram contained LAPC-4 cells, LNCaP cells, LNCaP derivatives and DU145 cells. The signatures of the cell lines within this clade were most similar to either SBS44 or SBS6 (expect for the signature of DU145^{C} which was most similar to SBS5). SBS6 and SBS44 are both associated with defective MMR, which may be the defining characteristic of this clade. In many but not all cases, parental cell lines neighbored their respective treatment-resistant sublines (**Figure 16 b**).

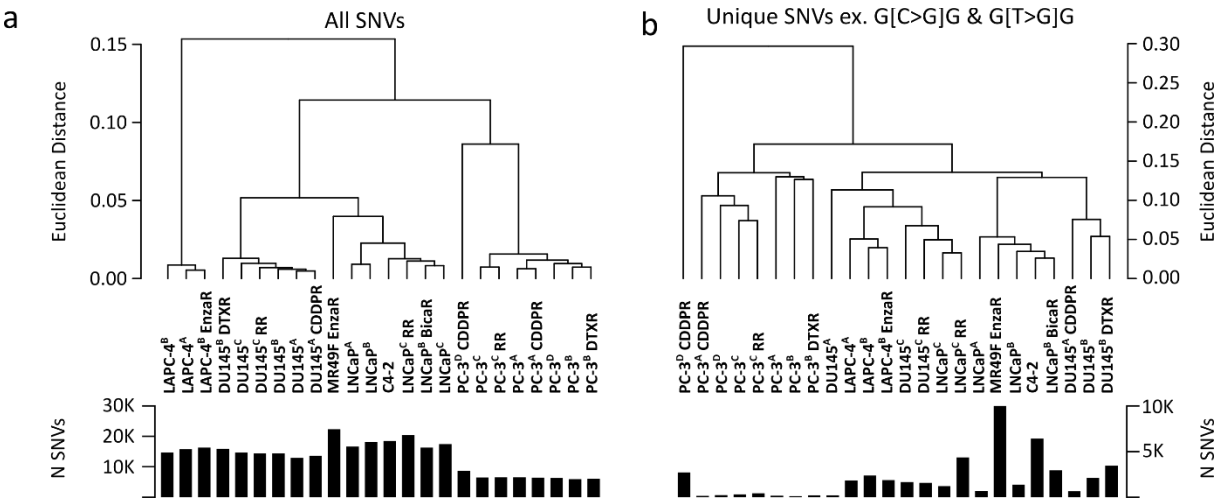

Figure 16: Clustering of cell lines based on their mutational signatures. The cells were clustered based on the signatures generated from (a) all SNVs or (b) unique SNVs, respectively. Assuming that they were primarily sequencing artefacts, G[C>G]G and G[T>G]G events were excluded from the unique SNVs.

5.3 Microsatellite Instability Analysis

Because the majority of experimental work conducted within this doctoral thesis focused on DU145 and PC-3 cells, only these were included in the MSI analysis. Furthermore, only the treatment-naïve parental cells were included. As MSI is an irreversible phenomenon, confirming MSI in the parental line indirectly confirms MSI in derivative cell lines as well. MSI was confirmed in all DU145 cells, while all PC-3 cell lines had stable microsatellites (**Table 20**). Of the five markers analyzed, DU145^{A} exhibited instability in two, DU145^{B} in three and DU145^{C} in three with an additional fourth marker that was classified as borderline stable (**Table 20**), respectively.

5.4 Cellular Morphology

The morphology of DU145 and PC-3 cell lines was studied (**Figure 17**). In their previous host laboratory, DU145^{C}, DU145^{C} RR, PC-3^{C} and PC-3^{C} RR received DMEM for culturing. To harmonize cell culture conditions, these cells were changed to RPMI 1640 medium. The switch from DMEM to RPMI in the mentioned cell lines did not produce any notable change in cell morphology or speed of proliferation (data not shown).

At high densities, all DU145 cell lines primarily exhibited a cuboidal epithelial-like growth pattern (**Figure 17 a**). The epithelial-like monolayer was interspersed with individual cells that exhibited a more spindle-shaped mesenchymal-like morphology (**Figure 17 a**). In contrast, the PC-3 cell lines – except PC-3^{B} and its subline PC-3^{B} DTXR – generally exhibited a spindle-shaped mesenchymal-like morphology that was clearly visible at both low and high densities (**Figure 17 b**). Compared to the other PC-3 cell lines, PC-3^{B} and PC-3^{B} DTXR exhibited a cuboidal more epithelial-like morphology at high density and both cell lines featured very large nuclei in proportion to their cytosol (**Figure 17 b**).

Phase contrast microscopy produces bright halos around objects that extend into the z-axis, while flat objects lack this particular brightness. The images suggest that DU145 cells grow as a flat and homogeneous monolayer, while many of the PC-3 cells had rounded and elevated profiles. Judging from the halo, PC-3^{B} cells seem flatter, while PC-3^{B} DTXR cells seem more spherical. Protruding structures resembling lamellipodia[158] were frequently seen in all PC-3 cells, however, they were most abundant and particularly prominent in PC-3^{C} cells (**Figure 17 b**, see red arrowheads). These structures were also – though less frequently – observed in DU145 cells (**Figure 17 a**, DU145^{C} RR). The morphology of PC-3^{C} was notably different than the morphology of its resistant subline PC-3^{C} RR. While PC-3^{C} featured the prominent multifocal protrusions just described, PC-3^{C} RR cells often featured a single prominent and far-reaching protrusion.

Table 20: MSI analysis of treatment-naïve parental DU145 and PC-3 cell lines.

Marker	Marker type	DU145[A]	DU145[B]	DU145[C]	PC-3[A]	PC-3[B]	PC-3[C]	PC-3[D]
Sample identification								
Penta C	pentanucleotide repeat	ok	ok	ok	ok	ok	ok	ok
Penta D	pentanucleotide repeat	ok	ok	ok	ok	ok	ok	ok
MSI markers								
BAT-25	mononucleotide repeat	stable	instable	instable	stable	stable	stable	stable
BAT-26	mononucleotide repeat	instable	instable	instable	stable	stable	stable	stable
MONO-27	mononucleotide repeat	instable	instable	instable	stable	stable	stable	stable
NR-21	mononucleotide repeat	stable	stable	stable*	stable	stable	stable	stable
NR-24	mononucleotide repeat	stable	stable	stable	stable	stable	stable	stable
Instable markers		2	3	3	0	0	0	0
MSI status:		MSI-high	MSI-high	MSI-high	MSS	MSS	MSS	MSS

Legend: The marker types as classified by Urquhart et al. 1994 are shown in the 'Marker type' column. A sample is microsatellite stable (MSS) when zero, 'MSI-low' when exactly one and 'MSI-high' when two or more markers among a panel of five recognized MSI markers exhibit instability[159,160]. * Denotes markers that were borderline stable.

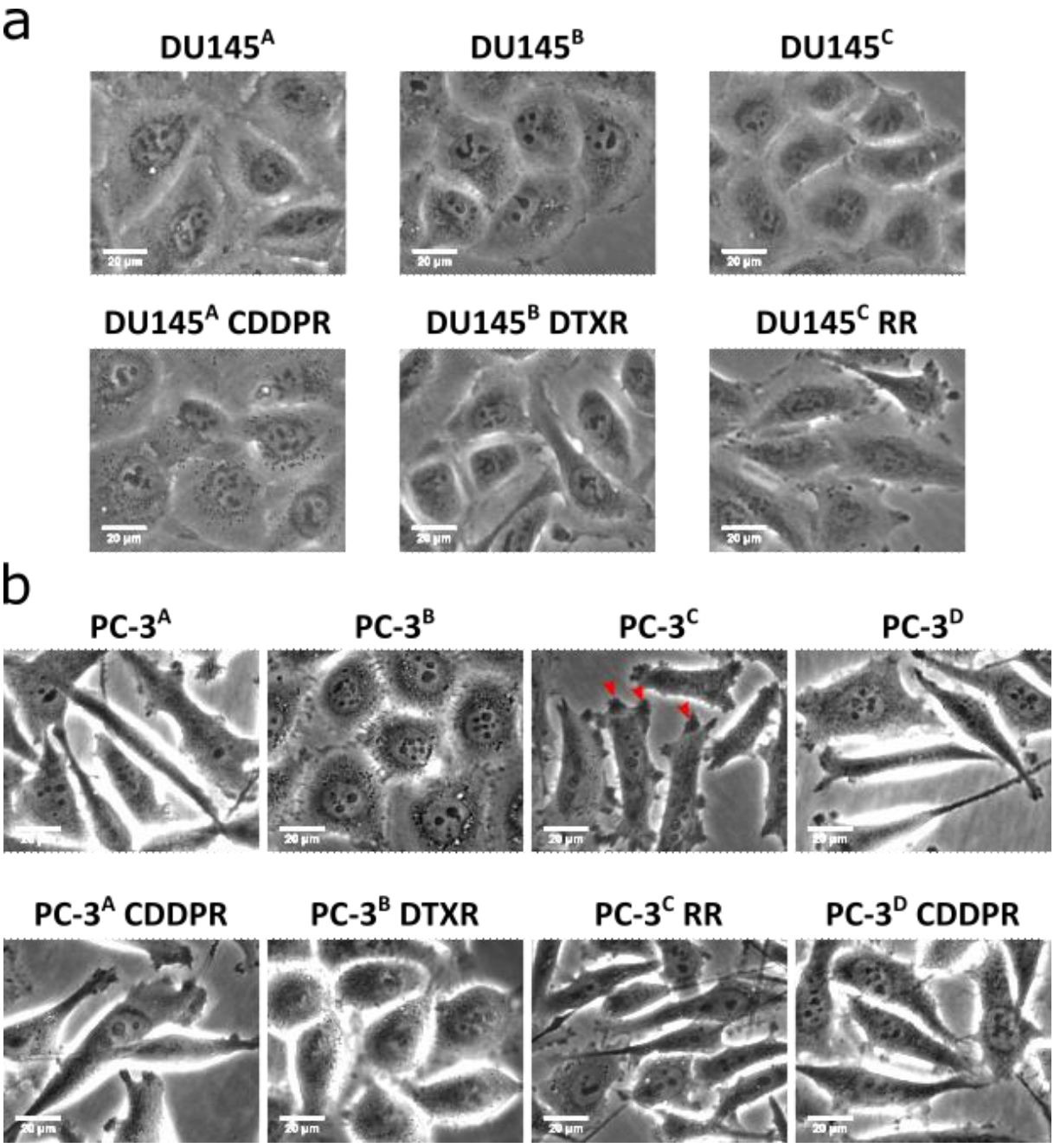

Figure 17: Morphology of DU145 and PC-3 cell cultures. DU145 (a) and PC-3 cells (b) were imaged with the EVOS XL Core Microscope. Phase contrast images were at 40X magnification. Desaturated greyscale versions of the original images are shown in order to better highlight the cellular morphology. Red arrowheads highlight structures that resemble lamellipodia. Scale bars (20 µm) are shown in all images.

5.5 Expression of Characteristic Proteins

The DU145 and PC-3 cells were studied by western blot (**Figure 18**). Authentication by STR analysis (section **5.1**) confirmed the cellular identity by providing a genetic fingerprint. However, confirming the presence and/or absence of specific markers on the mRNA and protein levels ensures that the cell lines also exhibit the right phenotypic properties as described in the literature.

DU145 and PC-3 cells are classically described as AR- and PSA-negative, while the castration-sensitive LNCaP cell line is AR- and PSA-positive[109,118]. In accordance, all DU145 cell lines used in this work lacked AR protein (**Figure 18**). However, while all other PC-3 cells were negative for AR protein, faint bands corresponding to the AR were detected in PC-3^{C} and PC-3^{C} RR. In agreement with the literature, all DU145 and PC-3 cell lines were negative for PSA (**Figure 18**). Furthermore, all PC-3 cells were negative for PTEN, which is a hallmark of these cells[121]. DU145 cell lines were negative for the MMR gene MLH1 on the protein level (**Figure 18**) but not on the mRNA level (**Figure 19**). This is in line with the fact that DU145 cells carry a *MLH1* point mutation in a splice site that yields an early premature stop codon in an otherwise wild-type mRNA transcript[117,161]. LNCaP^{B} cells were AR- and PSA-positive as expected.

Next, markers indicating the location on the spectrum between the mesenchymal and the epithelial cell states were studied. These markers are well studied and several works have described their expression in DU145, PC-3 and LNCaP cells previously[109,118]. The LNCaP^{B} cells served as an epithelial positive/mesenchymal negative control and fibroblasts (FB2010) as a mesenchymal positive/epithelial negative control vice versa. All DU145 cells and all PC-3 cells except PC-3^{C} RR were positive for the mesenchymal marker VIM (**Figure 18**). In PC-3^{B} and PC-3^{B} DTXR cells, VIM protein was detected as a single band at approximately 57 kDa, while in all other cell lines, the antibody reacted with two separate peptide entities at 50 and 57 kDa, respectively. With respect to the mesenchymal marker neuronal cadherin (CDH2), all DU145 cells were negative and all PC-3 cells were positive (**Figure 18**). The epithelial marker epithelial cadherin (CDH1) was present in all DU145 cells except DU145^{C} RR. In PC-3 cells, only PC-3^{B} and PC-3^{B} DTXR cells exhibited a clear CDH1 signal (**Figure 18**). However, at high exposure times, faint CDH1 bands were observed in other PC-3 cell lines as well (data not shown).

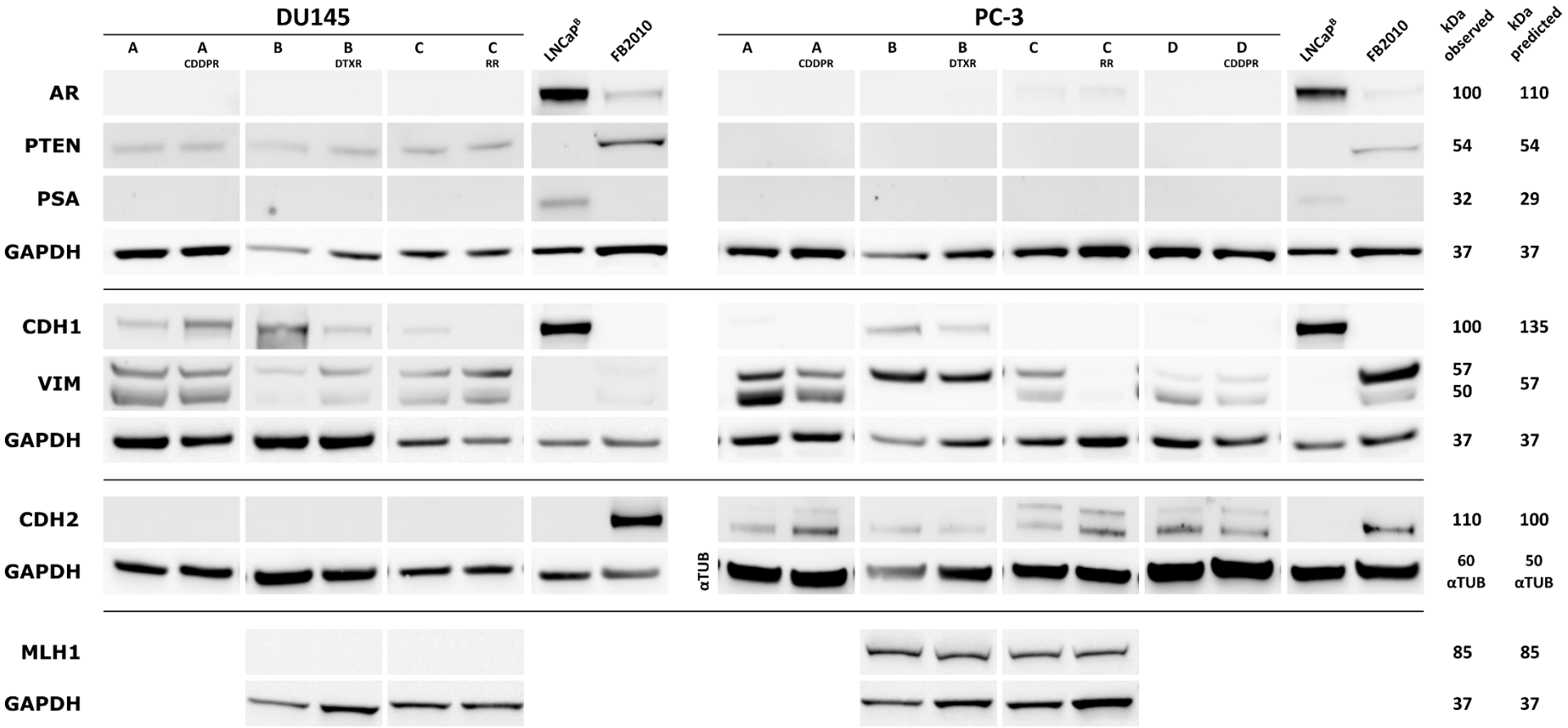

Figure 18: Expression of characteristic protein markers in DU145 and PC-3 cell lines. Western blots were performed as described in section **4.12**. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) or alpha-tubulin (αTUB) were used as references. LNCaP^{B} and FB2010 cells served as positive controls for the epithelial and mesenchymal phenotypes, respectively. LNCaP^{B} cells additionally served as an AR- and PSA-positive control. FB2010 cells (not listed in **Table 4**) are skin fibroblasts that were established from foreskin tissue at the Kinderklinik Dresden. The ladders are omitted. On the right, the observed and predicted molecular weights are shown in kilo Daltons (kDa). For information on the antibodies please refer to **Table 11**.

5.6 mRNA Expression

5.6.1 Selected genes

The mRNA expression levels of 49 selected genes were analyzed by qPCR (**Figure 19**). The selected genes included the *AR* and its target genes *PSA* and *PSMA*, epithelial markers (*CDH1*), mesenchymal markers (*CDH2*, *VIM*), TFs promoting the EMT process (*SNAI1*, *SNAI2*, *TWIST1*, *ZEB1*, *ZEB2*), DNA damage sensors (*ATM*, *ATR*), DNA repair genes (*BRCA1*, *BRCA2*, *ERCC1*, *ERCC2*, *FANCC*, *RRM1*, *RRM2*), ABCTs (*ABCA4*, *A8*, *B1*, *B4*, *B11*, *C1*, *C2*, *C10* and *G2*) and other handpicked genes associated with relevant resistance processes (e.g., *ATP7A*, *ATP7B*, *MTF1* and *TUBB3*). Unsupervised hierarchical clustering based on log-transformed $2^{-\Delta CT}$ values achieved meaningful separation of cell line groups (**Figure 19 a**). LNCaP cells, LNCaP derivatives and LAPC-4 cells were characterized by high relative expression of *AR*, *PSA*, *PSMA* and *CDH1* as well as low relative expression of *VIM*, *CDH2* and *ZEB1*. LNCaP, LNCaP derivatives and LAPC-4 cells together formed one of two major clades of the dendrogram (**Figure 19 a**). As expected, DU145 cells lacked *AR*, *PSA* and *PSMA* expression (**Figure 19 a**). PC-3 cell lines lacked *PSA* expression as well, however, some PC-3 cell lines had CT values lower than 35 for *AR* and *PSMA* (**Table 21**). With 28.8 and 28.4, PC-3^C and PC-3^C RR had the lowest CT values for the AR, respectively (**Table 21**). Accordingly, the AR antibody detected a faint band in both PC-3^C and PC-3^C RR cells (**Figure 18**). All PC-3 cells were negative for *PTEN* mRNA as expected. Compared with LNCaP and LAPC-4 cells, DU145 and PC-3 cells had higher relative expression of *VIM* and lower relative expression of *CDH1*, indicating more mesenchymal phenotypes. Furthermore, PC-3 cells were the only cell line expressing the mesenchymal marker *CDH2* (**Figure 19 a**). Together, DU145 and PC-3 cells formed the second major clade of the dendrogram.

In addition to the relative expression ($2^{-\Delta CT}$ values), x-fold expression ($2^{-\Delta\Delta CT}$ values) in resistant versus their parental cell lines was studied. At the root of the dendrogram, enzalutamide-resistant MR49F cells separated from the rest of the samples. This may be explained by the fact that C4-2 cells, which were taken as a reference, are not the age-matched parental cells of MR49F, as it is the case for the other cell line pairs. Thus, larger differences in gene expression could cause the prominent separation seen on the dendrogram.

The other major clade of the dendrogram contained cells of diverse types of resistance. Except for PC-3^A CDDPR and PC-3^D CDDPR, which exhibited similar gene expression, unsupervised hierarchical clustering did not generally group the resistance types (**Figure 19 b**). Many genes (e.g., *AMACR* in PC-3^D CDDPR or *ABCG2* in LAPC-4^B EnzaR) exhibited strong (> 10-fold) up- or downregulation in the resistant versus their parental cell lines (**Figure 19 b**). Notably, with CT values of 40.00 and 39.25 ABCB1 was negative in DU145^B and PC-3^B cells, respectively. However, ABCB1 had CT values of 25.48 in DU145^B DTXR and 25.50 in PC-3^B DTXR cells (**Figure 19 b**, boxes marked by white dots), respectively. $2^{-\Delta\Delta CT}$ values implied 10^4 to 10^5-fold upregulation of ABCB1 in the DTXR cell lines.

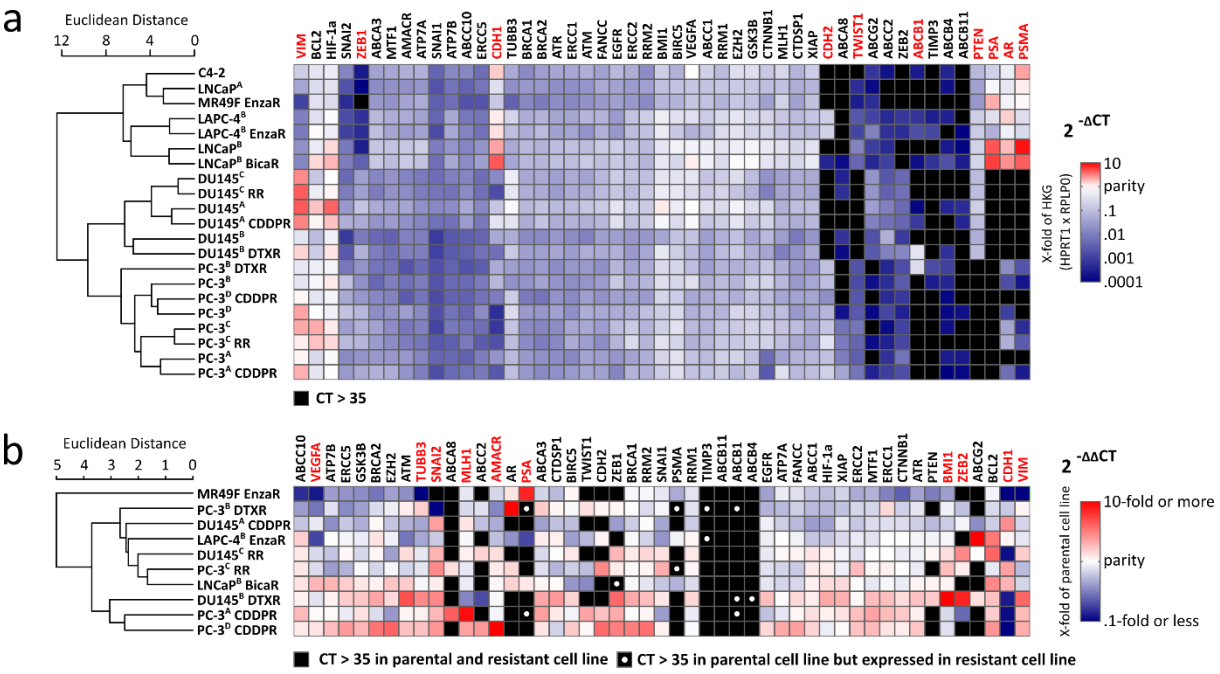

Figure 19: Gene expression of selected genes in various PCa cell lines. qPCR measurements were performed as described in section **4.11.3** and the TaqMan assays used are listed in **Table 9**. The upper panel (a) shows $2^{-\Delta CT}$ values, representing the gene expression relative to the house keeping gene (HKG) as a heatmap. Genes with CT values > 35 are colored black. The lower panel (b) shows $2^{-\Delta\Delta CT}$ values, representing the gene expression of the resistant cell lines relative to the parental cell line as a heatmap. Genes where the CT value of the parental and the treatment-resistant cell line was > 35 are colored black. Genes with a CT value > 35 in the parental line and < 35 in the resistant subline are additionally labelled with white dots if the difference of the CT values between the two cell lines was > 2. C4-2 cells were taken as a reference for MR49F EnzaR cells. Unsupervised hierarchical clustering was performed on both the cell lines and the measured genes. The dendrogram for the genes is omitted. The ten genes exhibiting the highest variance among the cell lines are colored red.

Table 21: Expression of the AR, PSA and PSMA in PC-3 cell lines. The obtained CT values for the AR and its target genes PSA and PSMA are shown. CT values > 35 (negative) are colored gray. The asterisks indicate the detection of the protein by western blot (**Figure 18**). AR and PSA but not PSMA were investigated by western blot analysis.

Cell Line	CT^{AR}	CT^{PSA}	CT^{PSMA}
PC-3^{A}	37.7	40.0	40.0
PC-3^{A} CDDPR	32.9	40.0	37.4
PC-3^{B}	33.1	38.4	36.1
PC-3^{B} DTXR	29.6	40.0	32.6
PC-3^{C}	28.8 *	40.0	35.1
PC-3^{C} RR	28.4 *	40.0	32.8
PC-3^{D}	38.5	40.0	39.0
PC-3^{D} CDDPR	33.8	40.0	39.7

5.6.2 RT^2 Profiler Arrays

In DU145 and PC-3 cells, gene expression was further interrogated using two RT^2 Profiler Arrays, covering genes connected with the EMT process (**Figure 20**) and genes involved in DNA repair pathways (**Figure 21**).

For the EMT array, unsupervised hierarchical clustering based on log-transformed $2^{-\Delta CT}$ values separated DU145 from PC-3 cells (**Figure 20**). In most cases, resistant cell lines clustered next to their parental counterparts. Exemplarily, *COL2A1*, *CDH2*, *KRT7* and *STAT3* were among the genes that most starkly discriminated PC-3 from DU145 cells. As opposed to the measurements done using TaqMan probes, *CDH2* was detected in DU145 cells on the EMT Profiler Array. In accordance with the literature[122], PC-3 cells were negative for *STAT3*.

Next, the x-fold gene expression ($2^{-\Delta\Delta CT}$) in parental DU145 and PC-3 cells relative to a hypothetical pooled sample (geometric mean of group) was studied (**Figure 20 b**). The expression of – for example – *IGFPB4*, *VIM* and *FGFBP1* was highly variable among DU145 parental cell lines. Parental PC-3 cell lines exhibited starkly disparate gene expression. The differences were most pronounced between $PC\text{-}3^B$ and $PC\text{-}3^C$, while $PC\text{-}3^A$ and $PC\text{-}3^D$ were in relative terms more similar. The genes with the highest variance included *SPARC*, *OCLN*, *SPP1*, *BMP2* and *BMP7*. Furthermore, x-fold gene expression in resistant versus parental cell lines was studied (**Figure 20 b**). On the EMT array, numerous genes with substantial up- or downregulation in the resistant cell lines were observed. These included – for example – *COL1A2*, *SNAI2* and *CDH1*.

For the DNA repair array, clustering based on log-transformed $2^{-\Delta CT}$ values correctly separated DU145 from PC-3 cell lines (**Figure 21 a**). Likewise, treatment-resistant cell lines were often located next to their parental counterparts. Across the cell lines, the expression of DNA repair genes was generally less variable than that of EMT genes.

X-fold gene expression relative to the group average indicated that many DNA repair genes had a lower expression in $DU145^B$ than in $DU145^A$ and $DU145^C$ (**Figure 21 b**). *TOP3A*, *TOP3B*, *DDB1* and *DDB2* were among the most variable genes among the parental DU145 cell lines. Among parental PC-3 cell lines, *ERCC1*, *ERCC2* and *ERCC6* were among the most variably expressed genes, indicating that the four cell lines may be discriminated by different levels of activity in the NER pathway. Finally, x-fold gene expression in resistant versus parental cell lines was studied (**Figure 21 b**). Compared to the EMT array, fewer genes on the DNA repair array exhibited substantial up- or downregulation in the resistant cell lines. For example, *MGMT* and MSH3 were downregulated 11.6- and 5.1-fold in $PC\text{-}3^C$ RR cells, respectively.

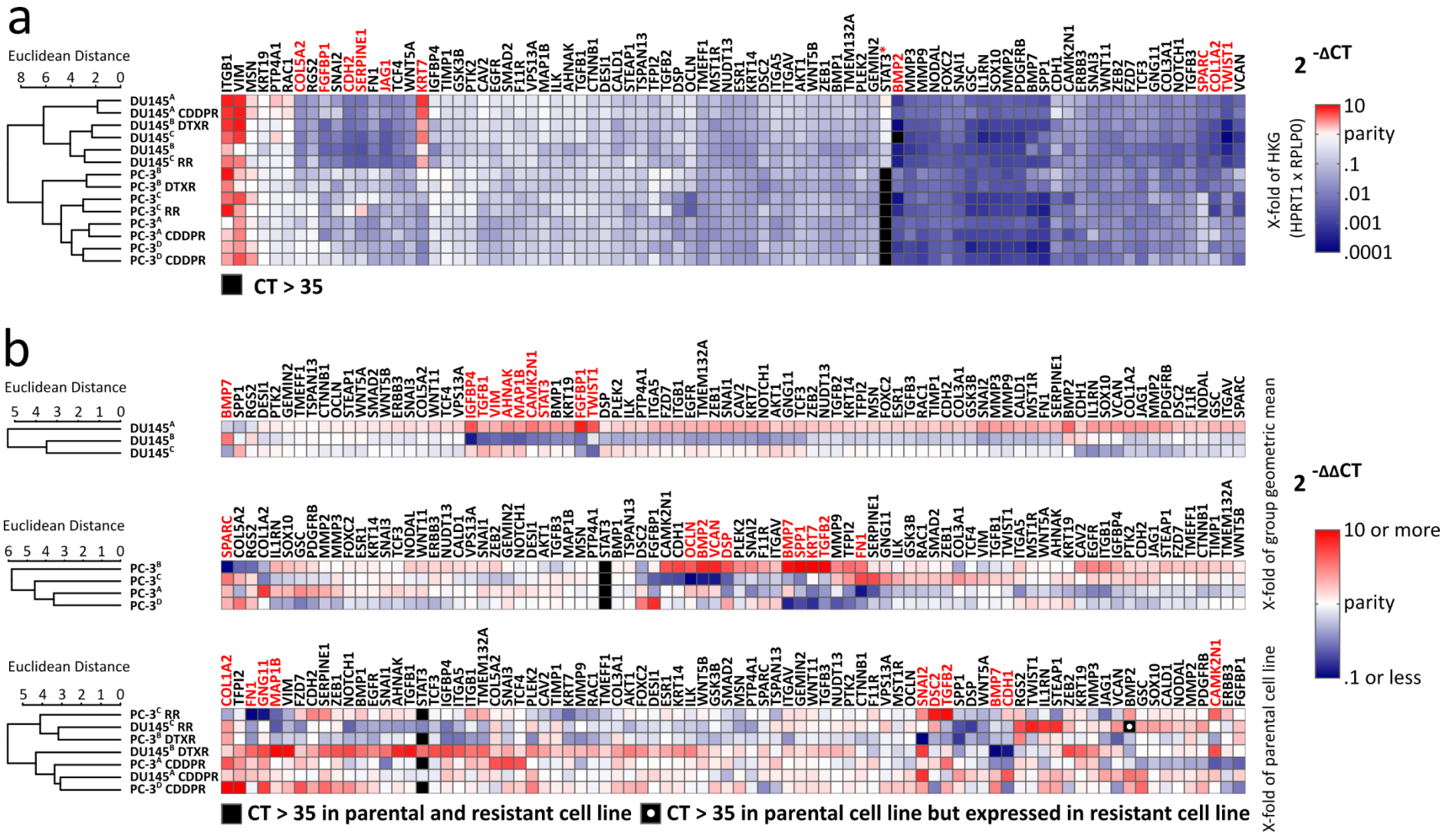

Figure 20: RT² Profiler Array – EMT. The array covers 84 genes connected (positively or negatively) with the EMT process. The upper panel (a) shows $2^{-\Delta CT}$ values (expression versus HKG) and the lower panel shows $2^{-\Delta\Delta CT}$ values (expression versus reference cell line), respectively. Unsupervised hierarchical clustering was performed on rows (cell lines) and columns (genes). The dendrograms for the genes are omitted. The ten genes exhibiting the highest variance among the cell lines are colored red.

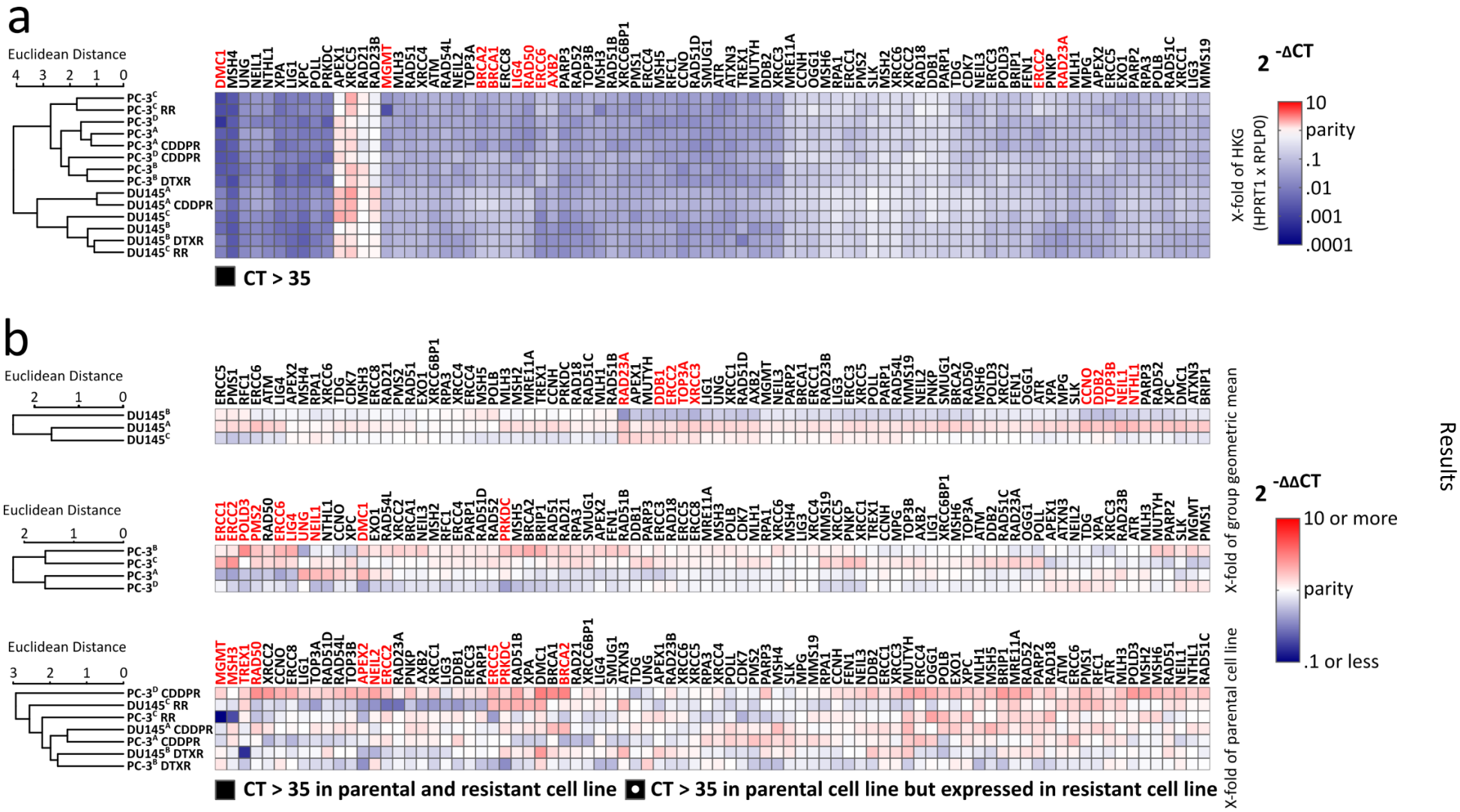

Figure 21: RT² Profiler Array – DNA repair. The array covers 84 genes involved in various DNA repair pathways. The upper panel (a) shows $2^{-\Delta CT}$ values (expression versus HKG) and the lower panel shows $2^{-\Delta\Delta CT}$ values (expression versus reference cell line), respectively. Unsupervised hierarchical clustering was performed on rows (cell lines) and columns (genes). The dendrograms for the genes are omitted. The ten genes exhibiting the highest variance among the cell lines are colored red.

5.7 Cisplatin Tolerance in Docetaxel- and Radio-Resistant Cell Lines

The CDDP tolerance of DU145^{B}, DU145^{B} DTXR, DU145^{C}, DU145^{C} RR, PC-3^{B}, PC-3^{B} DTXR, PC-3^{C} and PC-3^{C} RR was studied in order to assess whether DTX resistance and/or radio-resistance was associated with increased or decreased CDDP tolerance.

5.7.1 Cell Growth under Cisplatin Pressure

In the above-mentioned cell lines, the capacity to survive and proliferate under CDDP pressure was assessed. To this end, the cells were treated with a serial dilution of CDDP and their confluence was tracked for six days using the IncuCyte S3 (**Figure 22**). A first general observation was that DU145 cells tolerated higher CDDP concentrations than PC-3 cells. While the growth of DU145 cells was essentially unimpaired at CDDP concentrations of up to 0.1 µg/mL, the proliferation of PC-3 cells started to slow down at 0.025 µg/mL (**Figure 22**).
There was no apparent difference in the ability to grow under CDDP pressure between DU145^{B} and DU145^{B} DTXR cells. The highest concentration at which DU145^{B} and DU145^{B} DTXR cells were able to proliferate was 0.25 µg/mL CDDP (**Figure 22**). The same CDDP threshold for cell proliferation was observed in DU145^{C} cells. However, DU145^{C} RR cells were able to proliferate at up to 0.5 µg/mL CDDP, exhibiting an enhanced capacity to grow in the presence of CDDP compared to their parental cells (**Figure 22**).
As mentioned, PC-3 cells were generally more sensitive. There was no apparent difference between PC-3^{B} and PC-3^{B} DTXR cells as both cell lines proliferated at CDDP concentrations of up to 0.1 µg/mL (**Figure 22**). Contrary to what was described in the DU145 cells, radio-resistant PC-3^{C} RR cells were slightly more sensitive than their parental cell line PC-3^{C}, which exhibited a more robust growth at 0.1 and 0.25 µg/mL CDDP (**Figure 22**). Furthermore, low CDDP concentrations between 0.01 and 0.1 µg/mL affected cell proliferation in PC-3^{B} and PC-3^{B} DTXR more than in PC-3^{C} and PC-3^{C} RR, which tolerated these concentrations better.

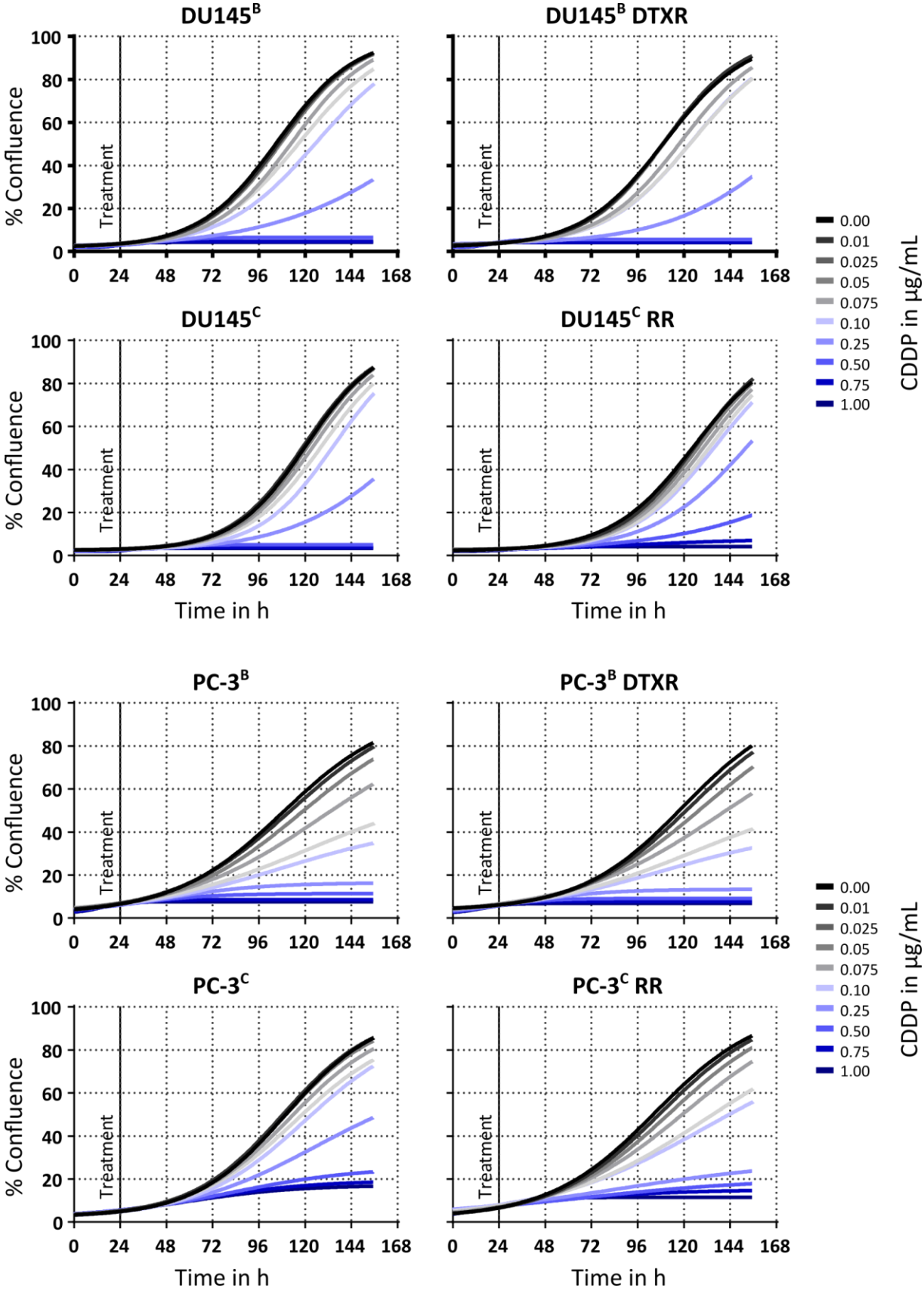

Figure 22: Cell growth of DU145 and PC-3 cells under cisplatin pressure. Cells were seeded in 96-well plates and 24 h later treated with a serial dilution of CDDP for six days. With the IncuCyte S3 Live-Cell Analysis System, the confluence was determined from acquired images every 6 h. Averaged datapoints from individual experiments (N≥4) are omitted to improve the visibility of fitted curves.

5.7.2 Confluence after Cisplatin Treatment

The cells were treated with serial dilutions of CDDP and dose response curves were generated from the confluence (x-fold of untreated cells) at 120 h after the initialization of the treatment (**Figure 23**). There was no difference between DU145^{B} vs DU145^{B} DTXR or between PC-3^{B} vs PC-3^{B} DTXR cells, respectively. The IC_{50} values (50% reduction of cell confluence relative to untreated cells) were 0.17 µg/mL CDDP for DU145^{B} and DU145^{B} DTXR cells and 0.07 µg/mL for PC-3^{B} and PC-3^{B} DTXR cells, again indicating that DU145 are able to tolerate CDDP better than PC-3 (**Figure 23**).

DU145^{C} RR (IC_{50} = 0.29 µg/mL) seemed to tolerate CDDP better than DU145^{C} (IC_{50} = 0.19 µg/mL). With P<0.01 based on the IC_{50} values of the individual experiments, this was statistically significant. In contrast, PC-3^{C} RR (IC_{50} = 0.10 µg/mL) demonstrated a higher sensitivity to CDDP than PC-3^{C}. Likewise, this was statistically significant (P<0.01, **Figure 23**).

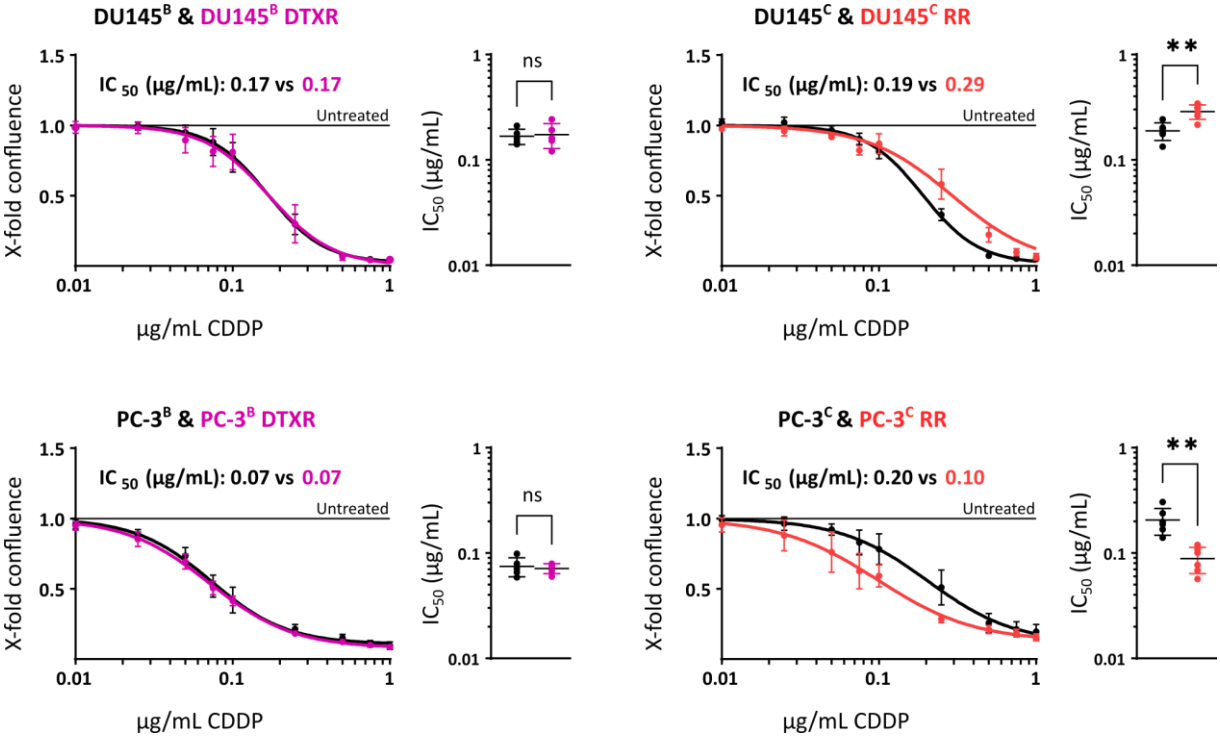

Figure 23: Confluence of DU145 and PC-3 cells after cisplatin treatment. DU145 and PC-3 cells were treated with CDDP for six days and their confluence was tracked in the IncuCyte S3. The confluence (x-fold of untreated cells) at 120 h after the initialization of the treatment (144 h after seeding) is plotted and IC_{50} values were calculated from the resulting inhibition curves. Averaged IC_{50} values are noted on the main graphs and IC_{50} values of individual experiments (N≥3) are plotted on the graphs adjacent to the right. All data are presented as mean ± SD. Non-paired t-tests with Welch's correction were performed to test for statistical significance of the differences between resistant cells and parental controls. ns: not significant, *: P<0.05, **: P<0.01.

5.7.3 Metabolic Activity after Cisplatin Treatment

To further assess the cells' ability to withstand CDDP toxicity, their metabolic activity was measured by a photometric WST-1 metabolization assay after six days of CDDP treatment (**Figure 24**). IC_{50} values representing the CDDP concentration eliciting the half of the maximal observed inhibition were calculated from the resulting inhibition curves. In accordance with other assays, DU145 cell lines (IC_{50}: 0.18 – 0.22 µg/mL) exhibited a greater tolerance for CDDP than PC-3 cell lines (IC_{50}: 0.06 – 0.12 µg/mL, **Figure 24**). Neither DTXR nor RR cell lines exhibited any significant difference in CDDP tolerance compared to their parental counterparts in this assay (**Figure 24**).

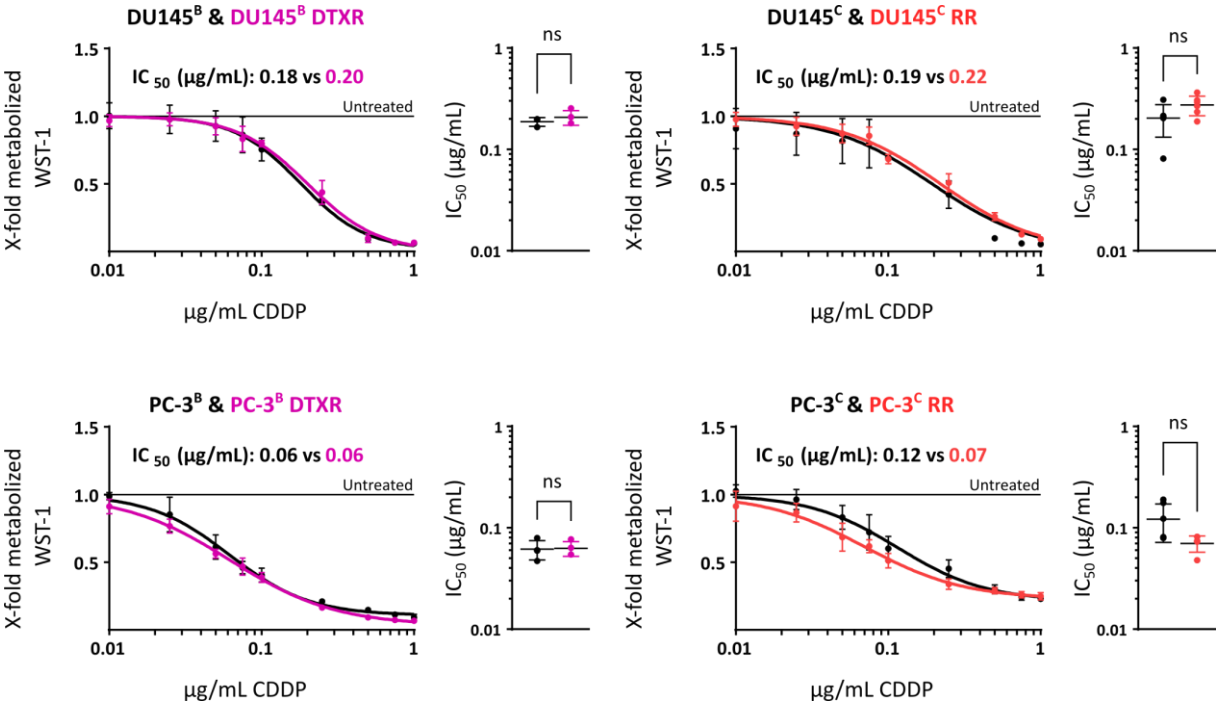

Figure 24: Metabolic activity of DU145 and PC-3 cells after cisplatin treatment. DU145 and PC-3 cells were treated with CDDP for six days and then incubated with WST-1. The absorption of the WST-1 metabolite formazan was measured. The x-fold metabolic activity of treated cells normalized to untreated cells is plotted. Averaged IC_{50} values are noted on the main graphs and IC_{50} values of individual experiments (N≥3) are plotted on the graphs adjacent to the right. All data are presented as mean ± SD. Non-paired t-tests with Welch's correction were performed to test for statistical significance of the differences between resistant cells and parental controls. ns: not significant, *: P<0.05, **: P<0.01.

5.7.4 Cell Survival after Cisplatin Treatment

Subsequent to the WST-1 assay (section **5.7.3**), a crystal violet assay was performed (**Figure 25)**. Crystal violet stains all cells and prior to measurement, non-adherent cells are removed by washing. Hence, the measured absorption is proportional to the adherent (viable) cell mass. IC_{50} values were calculated as described and DU145 cell lines were again generally more tolerant than the PC-3 cell lines (**Figure 25**, compare the IC_{50} values).

In this assay, PC-3^C RR (IC_{50} = 0.09 µg/mL) cells were significantly more sensitive to CDDP than their parental cell line PC-3^C (IC_{50} = 0.32 µg/mL, **Figure 25**). This observation matches the significant and non-significant results shown in this section (Cell growth under CDDP pressure: **Figure 22**, confluence: **Figure 23**, WST-1: **Figure 24** andcolony formation: **Figure 27**).

In DU145 cells, the IC_{50} values obtained through the crystal violet assay roughly matched the values obtained from the confluence and WST-1 assays. However, PC-3 cells exhibited higher IC_{50} values in the crystal violet assay than in the other assays (**Figures 23-25**). This suggests that in DU145 cells the decreased metabolic activity goes hand in hand with a loss of cell adherence and cell death. In contrast, in PC-3 cells the metabolic activity is diminished but the cells retain their adherence to the culture dish to a higher degree.

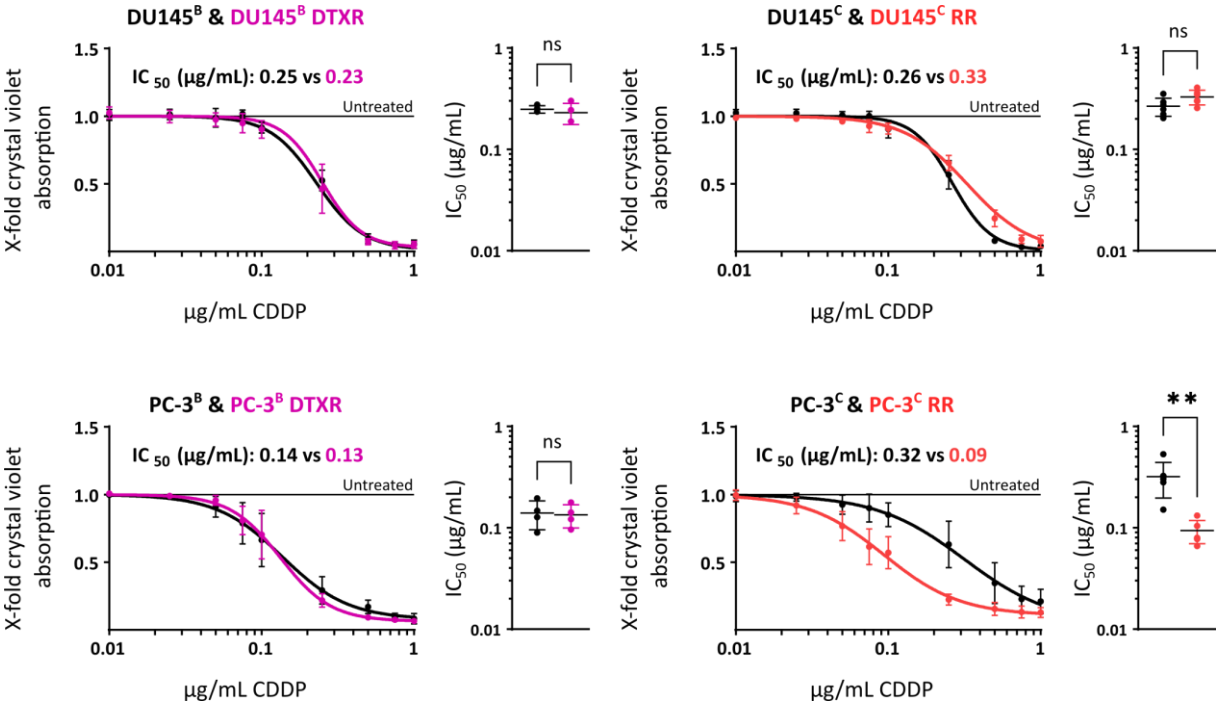

Figure 25: Cell survival of DU145 and PC-3 cells after cisplatin treatment. DU145 and PC-3 cells were treated with CDDP for six days and thereafter, the total adherent (viable) cell mass was stained using the crystal violet dye. The absorption of the crystal violet was measured. The x-fold crystal violet absorption of treated cells normalized to untreated cells is plotted. Averaged IC_{50} values are noted on the main graphs and IC_{50} values of individual experiments (N≥3) are plotted on the graphs adjacent to the right. All data are presented as mean ± SD. Non-paired t-tests with Welch's correction were performed to test for statistical significance of the differences between resistant cells and parental controls. ns: not significant, *: P<0.05, **: P<0.01.

5.7.5 Apoptotic Response to Treatment with Cisplatin

The impact of CDDP treatment on the induction of apoptosis was studied (**Figure 26**). In untreated DU145^{B} and DU145^{B} DTXR cells, the assay detected caspase activity in 14.1 ± 3.4% and 11.8 ± 2.5% of the cells, respectively. After treatment with 1 µg/mL CDDP, 24.8 ± 2.5% of DU145^{B} and 26.2 ± 5.2% of DU145^{B} DTXR cells exhibited caspase activity, respectively (**Figure 26**). The fraction of caspase^{+} cells was not different between DU145^{B} and DU145^{B} DTXR cells at any CDDP concentration. However, there was a significant induction of apoptosis mediated by the CDDP treatment in both cell lines (**Figure 26**).

With 6.6 ± 0.8% caspase^{+} cells in untreated cells versus 28.5 ± 6.3% in cells treated with 1 µg/mL CDDP, DU145^{C} exhibited the strongest induction of apoptosis among the tested cell lines (**Figure 26**). In contrast, DU145^{C} RR exhibited a high basal level of 17.1 ± 5.5% caspase^{+} cells, which increased just slightly to 19.2 ± 3.3% upon treatment with 1 µg/mL CDDP (**Figure 26**). PC-3^{B} and PC-3^{B} DTXR both exhibited a weak but significant induction of apoptosis mediated by the CDDP treatment. Between PC-3^{B} and PC-3^{B} DTXR, there was no significant difference in caspase^{+} cells at any CDDP concentration (**Figure 26**). Neither PC-3^{C} nor PC-3^{C} RR exhibited an induction of apoptosis mediated by the CDDP treatment. Similar to the DU145^{C} RR, PC-3^{C} RR also exhibited a high basal fraction of caspase^{+} cells (**Figure 26**).

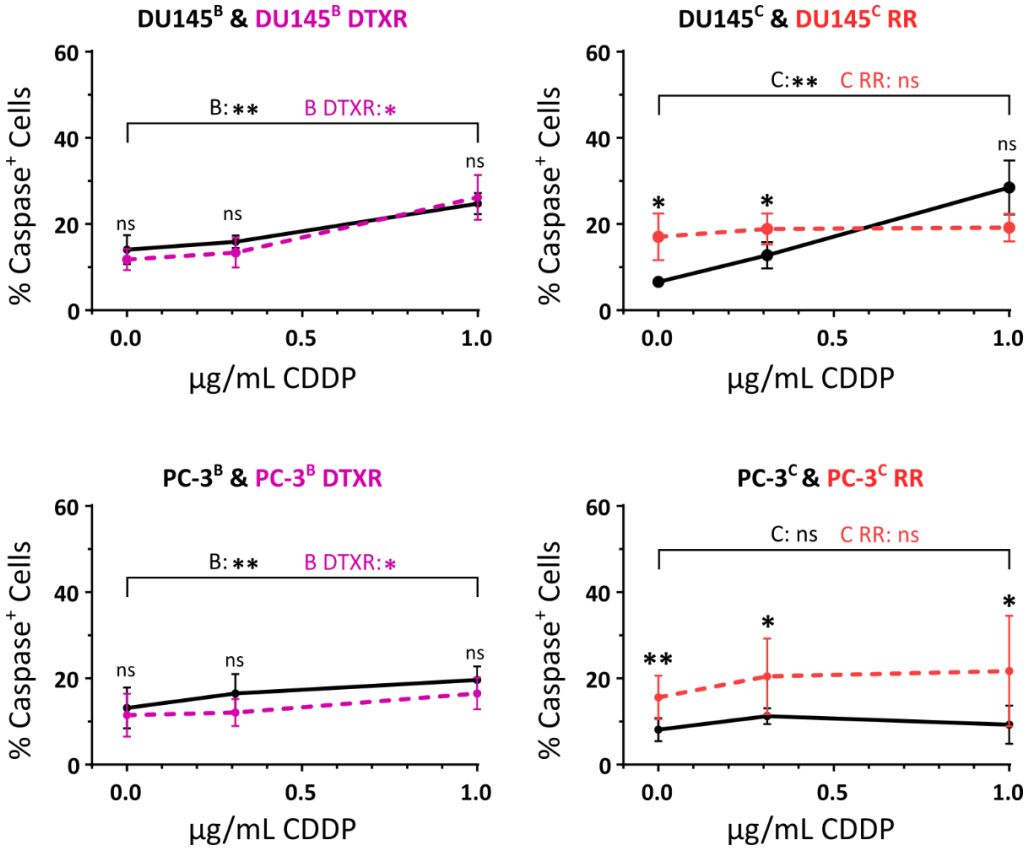

Figure 26: Apoptotic response of DU145 and PC-3 cells to cisplatin treatment. DU145 and PC-3 cells were treated with 0, 0.31 or 1 µg/mL CDDP and pan-caspase activity was measured. Caspase^{+} cells are plotted as % of total cells. Data points and error bars represent mean ± SD. All experiments were repeated N≥3 times. Non-paired t-tests with Welch's correction were performed to test for statistical significance of the differences between resistant cells and parental controls. ns: not significant, *: P<0.05, **: P<0.01.

5.7.6 Colony-Forming Ability after Cisplatin Treatment

The colony-forming ability of CDDP-pretreated cells was studied to evaluate the cells capacity to survive and recover from CDDP exposure in the long run (**Figure 27**). A stark difference between DU145 and PC-3 cell lines was observed. In DU145 cells, a CDDP treatment with 0.23 – 0.38 µg/mL was necessary to reduce the number of colonies by 50%, while in PC-3 cells the same reduction was achieved by CDDP concentrations between 0.03 and 0.06 µg/mL (**Figure 27**). DU145 cells were able to generate colonies after CDDP treatment with up to 0.5-µg/mL, while PC-3 cells tolerated a maximum of 0.2 – 0.3 µg/mL (**Figure 27**).
At concentrations up to 0.3 µg/mL, DU145^{B} DTXR cells tolerated CDDP better than DU145^{B} (**Figure 27**). Based on the $\beta_{1/2}$-values at which a 50% reduction of cell colonies was achieved, DU145^{B} DTXR tolerated more CDDP than DU145^{B} per trend (P = 0.056).
No differences regarding the CDDP tolerance were observed in DU145^{C} vs DU145^{C} RR and in PC-3^{B} vs PC-3^{B} DTXR cells, respectively. PC-3^{C} cells recovered from CDDP pretreatment slightly better than PC-3^{C} RR, however, this difference was not significant (**Figure 27**).

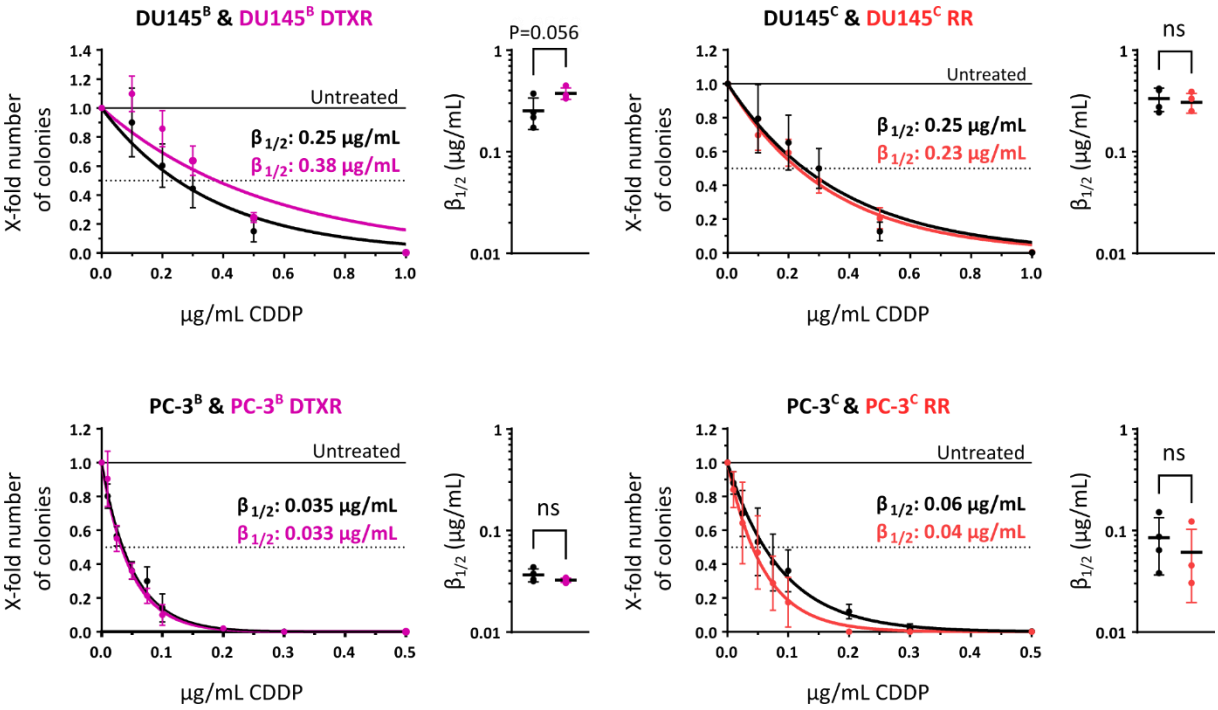

Figure 27: Colony-forming ability of DU145 and PC-3 cells after cisplatin treatment. After pretreatment with CDDP, DU145 and PC-3 cells were seeded for colony formation assays. The x-fold number of colonies of treated cells normalized to untreated cells is plotted. The concentrations $\beta_{1/2}$, at which colony formation was reduced by 50%, were inferred from the fitted curves and are noted in the graphs. $\beta_{1/2}$-values of individual experiments (N≥3) are plotted on the graphs adjacent to the right. Because of their greater sensitivity, PC-3 cells received additional CDDP concentrations between 0 and 0.1 µg/mL. Note that the x-axis scaling is different for DU145 and for PC-3 cells. All data are presented as mean ± SD and experiments were repeated for N≥3 times. Non-paired t-tests with Welch's correction (on $\beta_{1/2}$ values of individual experiments) were performed to test for statistical significance of the differences between resistant cells and parental controls. ns: not significant, *: P<0.05, **: P<0.01.

5.7.7 Knockdown of MLH1 in PC-3 Cells

A review of the literature suggested that the increased CDDP tolerance observed in DU145 compared to PC-3 cells may be caused by a genetic variant in the MMR gene *MLH1*. In DU145 cells, the *MLH1* variant chr3:g.37038108A>T is located near a splice site and causes an early truncation and LoF of the MLH1 protein (**section 1.5.1**).

Whole exome sequencing confirmed that the variant existed and was homozygous (VAF ~ 100%) in all six DU145 cell lines (data not shown/see deposited file). None of the other cell lines carried any variant in the *MLH1* gene. All DU145 and PC-3 cell lines contained MLH1 mRNA in detectable amounts (**Figure 19 a**, column #6). On the protein level, MLH1 was detected in PC-3 but not in DU145 cell lines (**Figure 28 a**). Preliminary experiments suggested that the most effective downregulation of the MLH1 protein (reduction by 80 – 95%) was achieved 72 h after siRNA transfection (data not shown). Knockdowns (N=3) of MLH1 were conducted in PC-3^B, PC-3^B DTXR, PC-3^C, PC-3^C RR cells (**Figure 28 b**). Experiments to determine whether the knockdown of MLH1 influenced the CDDP tolerance were done in PC-3^B and PC-3^C cells. In short, no differences regarding CDDP tolerance were found between siCON- and siMLH1-transfected cell lines. Exemplarily, the results obtained for PC-3^B cells are shown in **Figure 28 c** and **d**. The IC_{50} based on the WST-1 data were 0.29 µg/ml for siCON-transfected PC-3^B versus 0.28 µg/mL for siMLH1-transfected PC-3^B cells, respectively. These values cannot be compared to the values shown on **Figure 24** as the cells were treated later (96 h after seeding / 72 h after transfection versus 24 h after seeding) and the duration of the CDDP exposure was four days instead of six (compare **Figure 22** and **Figure 28 c**).

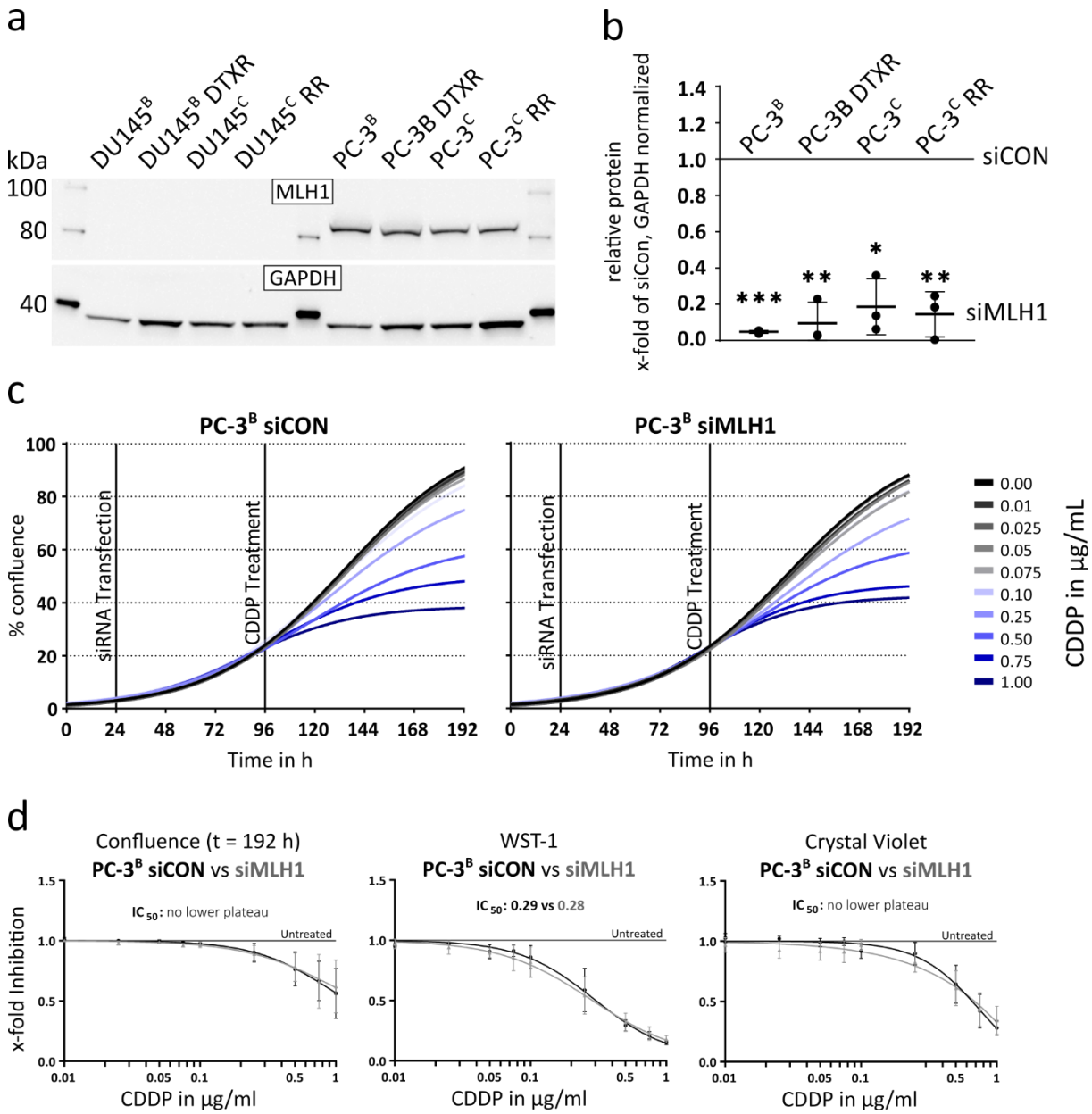

Figure 28: siRNA-mediated knockdown of MLH1 in PC-3^B cells. (a) shows a western blot confirming that DU145 but not PC-3 cells were negative for MLH1 protein. (b) shows a densitometric analysis of western blots (N=3) performed 72 h after transfection. The relative amount of protein (x-fold of siCON and normalized to GAPDH) is plotted. One sample t-tests (normalized data) were done to test for statistical significance of the MLH1 knockdown (*: P<0.05, **: P<0.01, ***: P<0.001). (c) With the IncuCyte S3, the confluence was determined from acquired images every 6 h. Averaged datapoints from individual experiments (N≥3) are omitted to improve the visibility of fitted curves. (d) Dose response curves based on the normalized confluence at 192 h (left, see also c), on the WST-1 assay (middle) and on the crystal violet assay (right) are shown. Because the lower plateaus of inhibition were not reached based on the confluence and crystal violet data, the calculation of IC_{50} values was only possible for the WST-1 data.

5.8 Targeting ABCB1 in Docetaxel-Resistant Cell Lines

5.8.1 Docetaxel Tolerance in DU145 and PC-3 Cell Lines

The DTX tolerance was measured in all DU145 and PC-3 cell lines (**Figure 29**). The DTX tolerance in DU145^{B} DTXR (IC_{50} = 400 ± 66 nM) and PC-3^{B} DTXR (IC_{50} = 649 ± 262 nM) was increased nearly 100-fold compared to their parental control cell lines (DU145^{B} = 5.3 ± 1.7 nM and PC-3^{B} = 9.4 ± 2.8 nM), respectively. In addition, PC-3^{D} CDDPR cells (4.0 ± 0.8 nM) were significantly more sensitive to DTX than their parental control PC-3^{D} (6.9 ± 1.4 nM). The other cell line pairs did not exhibit significant differences in DTX tolerance between parental control and treatment-resistant cells (**Figure 29**).

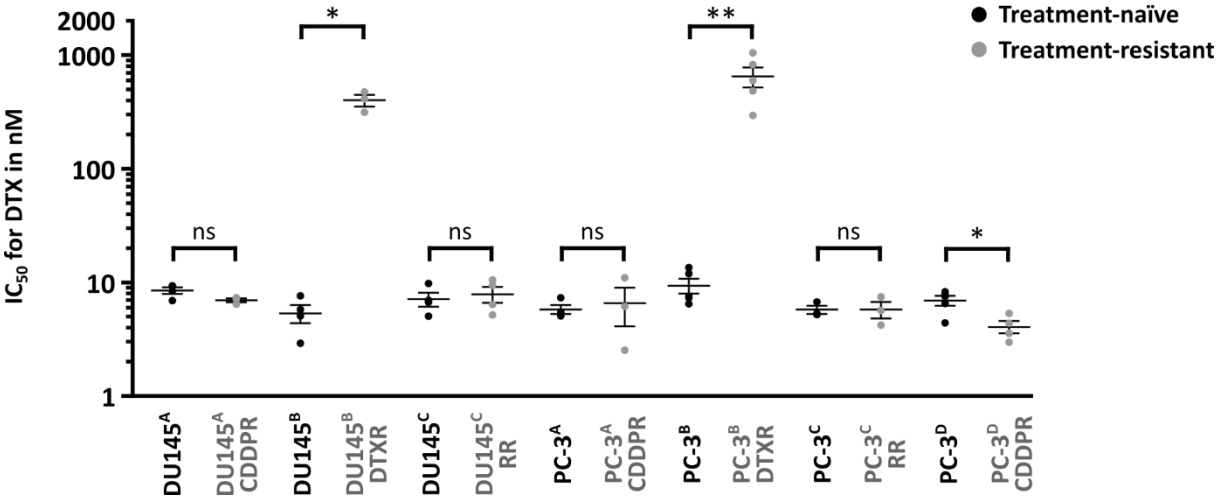

Figure 29: DTX tolerance in DU145 and PC-3 cell lines. The cells were treated with serial dilutions of DTX (section **4.13.1**) and WST-1 assays were performed. IC_{50} values of individual experiments (N≥3) are plotted. Treatment-naïve cell lines are colored black and their treatment-resistant sublines are colored grey. Data points and error bars represent mean ± SD. Non-paired t-tests with Welch's correction were performed to test for statistical significance of the differences between resistant cells and parental controls. ns: not significant, *: P<0.05, **: P<0.01.

5.8.2 DU145^{B} DTXR and PC-3^{B} DTXR Cells exclude Cell Staining Dyes

DU145^{B} and PC-3^{B} were stained with Nuclight Rapid Red – a nucleic dye that is used for cell counting in the IncuCyte S3. When DU145^{B} DTXR and PC-3^{B} DTXR were subjected to the same staining protocol, the dye failed to stain their nuclei (data not shown). After a review of the literature, it was hypothesized that ABCTs might be particularly abundant and/or active in the two cell lines, effluxing the nucleic dye and possibly facilitating DTX resistance by the same mechanism. Following the initial incidental finding of Nuclight Rapid Red exclusion, the cells were stained with Hoechst 33342 (Hoechst), which is a substrate of ABCG2[162]. Hoechst fluorescence was significantly lower in DU145^{B} DTXR and PC-3^{B} DTXR than in their respective parental cell lines. This was reversed by addition of the ABCT inhibitor verapamil (**Figure 30**). This experiment corroborated the assumption that ABCTs were particularly active/abundant in DU145^{B} DTXR and PC-3^{B} DTXR cells.

A consecutive experiment was performed by Ms. Dinah Linke (see **page 139**) using the fluorescent dye Rhodamin 123, which is a substrate of ABCB1[163]. The efflux of Rhodamin 123 was tracked in real time using the IncuCyte S3. Specific inhibitors of ABCB1 significantly slowed down the rate of Rhodamin 123 efflux in DU145^{B} DTXR and PC-3^{B} DTXR cells (data not shown). The effect of ABCB1 inhibition on Rhodamin 123 efflux considerably surpassed the effects seen in the Hoechst efflux assay.

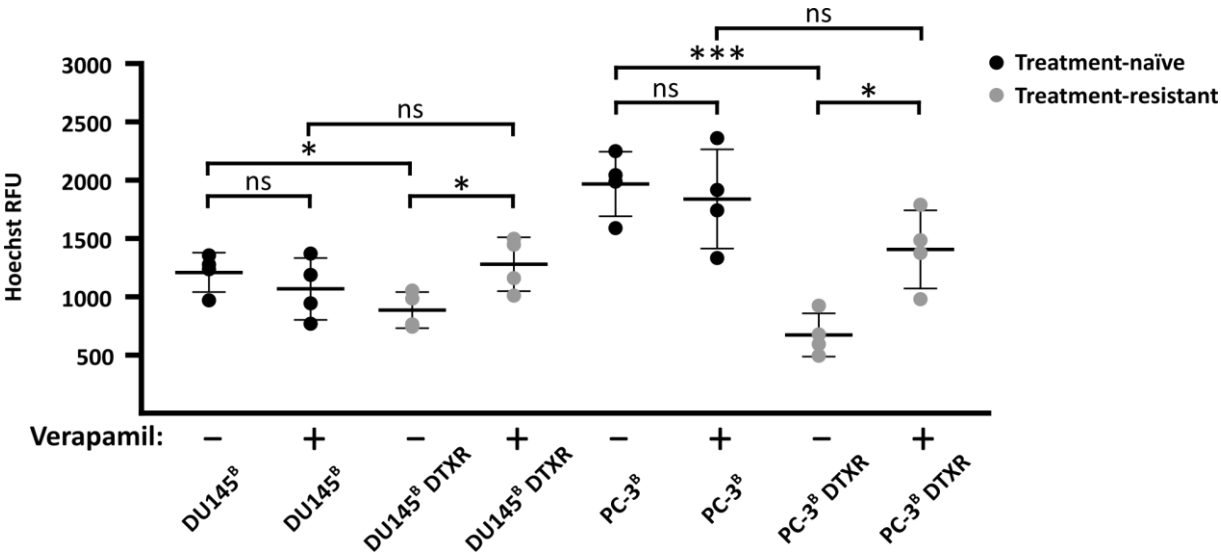

Figure 30: Verapamil enhances Hoechst retention in DTX-resistant cells. Cells were seeded and treated with Hoechst with or without the ABCT inhibitor verapamil as described (section **4.13.2**). After removal of the treatment and washing, the remaining Hoechst fluorescence was measured. Data points of individual experiments (N=4) and error bars represent mean ± SD. Non-paired t-tests with Welch's correction were performed to test for statistical significance of the differences between resistant cells and parental controls. ns: not significant, *: P<0.05, **: P<0.01, ***: P<0.001. RFU: Relative fluorescence units.

5.8.3 ABCB1 is Upregulated in DU145^{B} DTXR and PC-3^{B} DTXR Cells

To narrow down which of the 49 human ABCTs was active in DU145^{B} DTXR and PC-3^{B} DTXR cells, the expression of eight ABCTs (including ABCB1 and ABCG2) associated with taxane resistance[67,164] was determined by qPCR. Among those, only ABCB1 exhibited significant differential expression between treatment-naïve and DTXR cell lines. While ABCB1 mRNA was not detected in PC-3^{B} and DU145^{B} (CT values around 40), the CT values were around 25 in PC-3^{B} DTXR and DU145^{B} DTXR (**Figure 19 b**, see white dots). In addition to qPCR, flow cytometry was performed for ABCB1 and ABCG2 (**Figure 31**). Flow cytometry confirmed the upregulation of ABCB1 in DU145^{B} DTXR and PC-3^{B} DTXR cells. While nearly 100% of PC-3^{B} DTXR were ABCB1^{+}, DU145^{B} DTXR consisted of ABCB1^{+} (~85%) as well as ABCB1^{-} (~15%) cells. The ABCG2 antibody used for flow cytometry did not detect ABCG2 protein in DU145^{B} DTXR and PC-3^{B} DTXR cells (**Figure 31**).

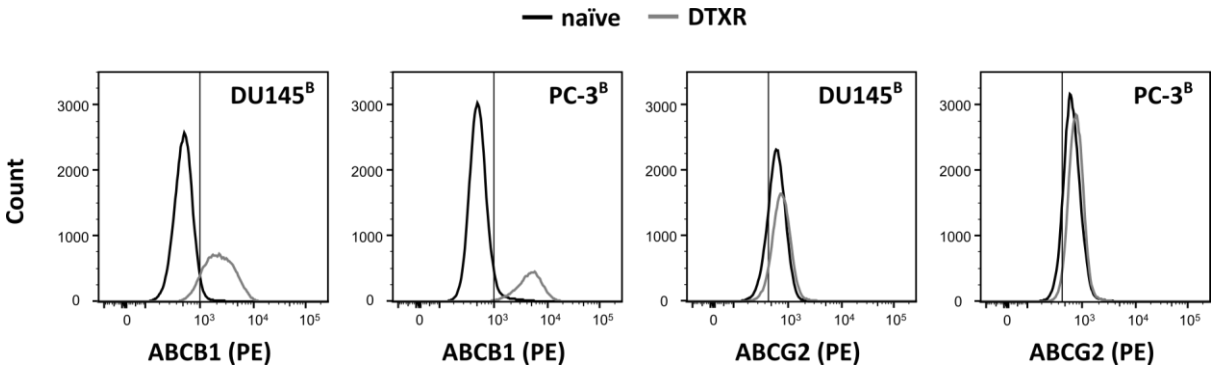

Figure 31: Flow cytometry for ABCB1 and ABCG2 in DU145 and PC-3 cell lines. Flow cytometry was performed as described in section **4.13.3**. Single channel stainings for ABCB1 and ABCG2 are shown as histograms, respectively. The horizontal line marks the detection threshold as determined by the isotype control antibodies.

5.8.4 ABCB1 is Differentially Glycosylated in DU145^{B} DTXR and PC-3^{B} DTXR Cells

Western blots for ABCB1, ABCG2 and ABCC1 were performed. DU145 and PC-3 cell lines were negative for ABCC1 and only minimal amounts of ABCG2 were detected. In contrast, ABCB1 was consistently detected in all four cell lines and was strongly upregulated in the DTXR cell lines (**Figure 32 a**). In addition to a 140 kDa band present in all four cell lines, the antibody detected an intense and wide additional band at 170 kDa in the two DTXR cell lines (**Figure 32 a**).

After reviewing the literature[165,166], it was reasoned that the band at 170 kDa in DTXR cells probably represents versions of ABCB1 modified with additional glycosyl residues. To test this, the harvested protein lysates of DTXR cells were digested with the bacterial N-glycosidase PNGase F[167] (section **4.13.4**). After digestion, the 170 kDa band was barely detected, the band at 140 kDa seemed diminished and a new band at 120 kDa emerged (**Figure 32 b**). Indeed, 120 kDa matches the theoretical molecular weight of the unmodified amino acid chain of the longest ABCB1 transcript (1350 amino acids, see https://www.ncbi.nlm.nih.gov/gene/5243). The identity of all bands was confirmed by a siRNA-mediated knockdown in all four cell lines (data not shown).

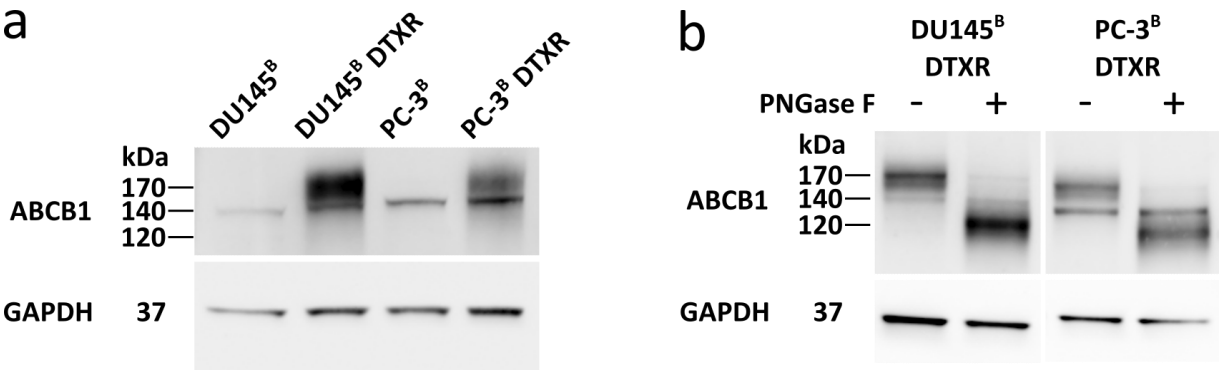

Figure 32: Differential glycosylation of ABCB1 in DU145^{B} DTXR and PC-3^{B} DTXR cells. (a) Western blot detecting the ABCB1 protein in DU145^{B}, DU145^{B} DTXR, PC-3^{B} and PC-3^{B} DTXR cells. The DTXR cell lines exhibit an additional band at 170 kDa. (b) Western blot detecting the ABCB1 protein in DU145^{B} DTXR and PC-3^{B} DTXR. Prior to performing the western blot, the protein lysates were digested with PNGase F. GAPDH was used as a reference.

5.8.5 ABCB1 is the Effector of Resistance in DU145^{B} DTXR and PC-3^{B} DTXR Cells

After finding that ABCB1 was likely the dominant ABCT expressed and active in DU145^{B} DTXR and PC-3^{B} DTXR cells, several approaches of targeting ABCB1 aiming to resensitize the two cell lines against DTX were conducted and evaluated. In short, two specific and non-competitive inhibitors of ABCB1 (elacridar, tariquidar), siRNA against ABCB1 and two inhibitors of glycosylation (tunicamycin, swainsonine) were tested. Measurements of confluence, metabolic activity (WST-1), total viable cell mass (crystal violet) and colony-forming ability served as readouts. The two small molecule inhibitors elacridar[76] and tariquidar[77] as well as siRNA-mediated knockdown of ABCB1 achieved a full resensitization of DU145^{B} DTXR and PC-3^{B} DTXR cells as seen by the WST-1 data (**Figure 33**) and other assays (data not shown). With ABCB1 inhibitor treatments, the IC_{50} values for DTX decreased by two orders of magnitude, arriving at the same level observed in the treatment-naïve parental cell lines. These data suggest that ABCB1 is not just one of many, but in fact the primary effector of DTX resistance in DU145^{B} DTXR and PC-3^{B} DTXR cells.

Treatment with tunicamycin[168], which inhibits the initial step of protein N-glycosylation[169], led to a notable but non-significant reduction of DTX tolerance in DU145^{B} DTXR cells. However, it was not effective in PC-3^{B} DTXR (**Figure 33**). Due to its toxicity for the cells, the use of tunicamycin had to be limited, both regarding to the treatment concentration and duration. Furthermore, swainsonine[170], which acts downstream of tunicamycin and inhibits the maturation of initial glycosyl-modifications in the Golgi apparatus[171], was not able to resensitize the DTX-resistant cells (data not shown). All experiments presented on **Figure 33** were conducted by Ms. Dinah Linke (see **page 139**).

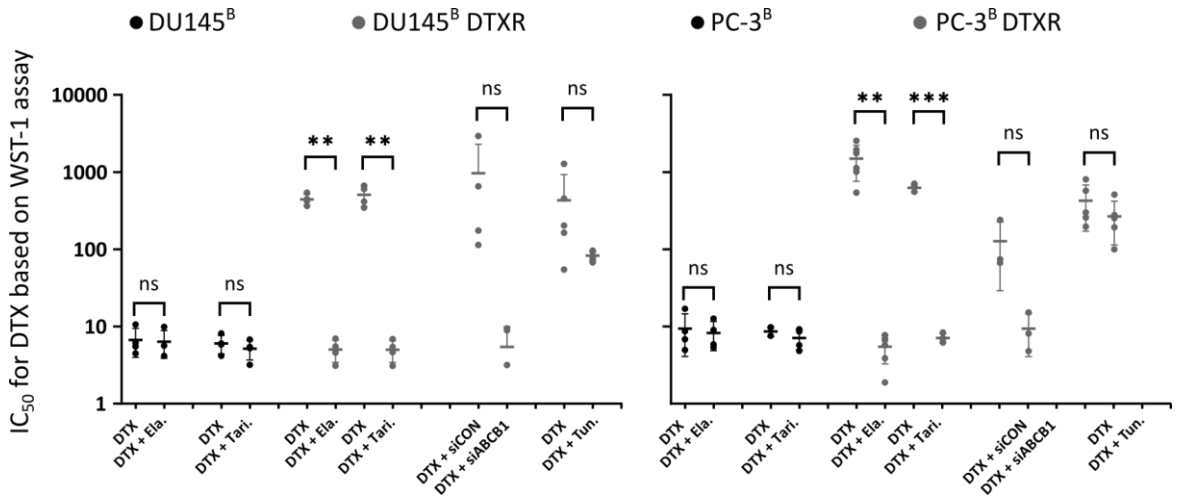

Figure 33: Influence of multiple approaches targeting ABCB1 on DTX tolerance. Cells were treated with small molecule inhibitors of ABCB1 (elacridar: Ela and tariquidar: Tari), siRNA against ABCB1 (siABCB1 and unspecific control construct siCON) and the glycosylation inhibitor tunicamycin (Tun). Cells were treated with either DTX alone, or DTX in combination with one of the above-mentioned compounds. IC_{50} values for DTX based on the WST-1 assay plotted. Data points of individual experiments (N≥3) and error bars represent mean ± SD. Non-paired t-tests with Welch's correction were performed to test for statistical significance of the differences between resistant cells and parental controls. ns: not significant, *: P<0.05, **: P<0.01, ***: P<0.001. All experiments were conducted by Ms. Dinah Linke.

6 Discussion

6.1 MMRd and MSI in Prostate Cancer Cell Lines

Microsatellites are repetitive DNA sequences with repeat units of one to six base pairs[172]. MSI refers to fluctuating numbers of repeat units in microsatellites. It was first described in colorectal cancer in 1993[173,174] and is defined by instable mono- and dinucleotide repeat sequences[140]. MSI is driven by MMRd[175,176], which introduces high numbers of point mutations and small InDels, specifically within repetitive DNA elements[96].

DU145, LNCaP and LAPC-4 cells are known to exhibit MMRd and MSI[117,161,177]. For PC-3 cells, conflicting reports regarding their MSI status exist[161,178]. However, Chen et al.[161] who reported MSI in PC-3 cells used an enzymatic assay that involved the introduction of a plasmid carrying a reporter gene. PCR- or sequencing-based approaches are more widely used for MSI detection. PCR-based MSI-testing using the NCI/Bethesda panel[140] of mononucleotide markers confirmed MSI in all three parental DU145 cells and showed that all four parental PC-3 cells had stable microsatellites (**Table 20**). LNCaP and LAPC-4 cells were not subjected to this test.

6.1.1 Diagnosing MMRd/MSI by Whole Exome Sequencing

In all six DU145 cell lines, WES rediscovered a homozygous SNV affecting the splicing of *MLH1* and which is known to be responsible for MMRd in these cells[117] (data not shown). LNCaP cells feature a homozygous deletion in the *MSH2* gene, which spans from exon 9 to 16[123]. This variant was not detected by WES, due to its size. Interestingly, literature regarding the exact molecular cause of MMRd and MSI in LAPC-4 cells was not found. Within the core mismatch repair genes (*MSH2-6*, *MLH1/3* and *PMS2*), the WES dataset contained three variants that were shared among LAPC-4^{A}, LAPC-4^{B} and LAPC-4^{B} EnzaR. One of the three variants (chr14:g.75498874C>T, rs116061019, affecting *MLH3*) was predicted to have a damaging effect on the protein while the other two were predicted to be benign/tolerated by the PolyPhen[179] and SIFT[180] algorithms, respectively. However, the respective variant in *MLH3* had an average VAF of just 43% across the three LAPC-4 cell lines. This suggests that at least one wild-type copy of the *MLH3* gene is likely still present in these cell lines. In summary, the variants found in the WES data are likely not able to explain MMRd in LAPC-4 cells. Possibly, LAPC-4 cells harbor MMR gene variants similar to LNCaP, which cannot be detected by WES. Likewise, there were no potentially deleterious variants among the core MMR genes in PC-3 cell lines.

While individual variants can potentially explain the cause(s) of MMRd, mutational signatures are a way to detect MMRd from sequencing data. Mutational signatures indicated MMRd in DU145, LNCaP and LAPC-4 but not in PC-3 cell lines (**Figure 14** and **Figure 15**).
In LNCaP^{B} BicaR, LNCaP^{C} RR and all three LAPC-4 cell lines, mutational signatures based on all variants were most similar to signatures associated with MMRd (**Figure 14**). However, in all DU145 and all thee parental LNCaP cell lines, signatures based on all variants were most similar to SBS5, which represents a mutational process associated with age. However, the exact molecular mechanism(s) resulting in the SBS events that together constitute SBS5 is/are unknown.
Compared to shared and partially shared variants, unique variants carry information about the more recent underlying mutational processes. By excluding shared and partially shared variants from the analysis, it was possible to reduce the influence of SBS5. Based on unique variants exclusively, all DU145, LNCaP and LAPC-4 cell lines except DU145^{A} exhibited mutational signatures associated with MMRd (**Figure 15**). However, the signatures of DU145^{A} and most PC-3 cell lines then indicated contamination by sequencing artefacts. Two SBS classes were identified as likely artefacts (**Table 18**). After removal of the putative artefacts, the signature of DU145^{A} was associated with MMRd and the signatures of most PC-3 cell lines were again most similar to SBS5. In summary, mutational signatures indicate MMRd in DU145, LNCaP and LAPC-4 but not in PC-3 cell lines.

6.1.2 Diagnosing MMRd/MSI by STR analysis

Tetranucleotide repeat (TNR) markers are microsatellites where the individual repeated units are four base pairs long. Generally, TNRs are stable in the sense that the number of repeat units will not change over the course of an individual's life. However, between individuals their lengths differ. Analyzing the number of repeat units in panels of TNRs and pentanucleotide markers forms the backbone of genetic fingerprinting, for example for forensic purposes but also for sample authentication.
In 2000, the instability of certain TNR markers, in particular of $AAAG_n$ and $ATAG_n$ sequences[172], was described[181] and coined elevated microsatellite alterations at selected TNRs (EMAST)[181]. EMAST has been described in numerous cancer entities[172] including PCa[182]. Initially, it was proposed that EMAST had distinct molecular root causes independent from MMRd. For example, an association between EMAST and *p53* mutations was proposed[181]. Burger et al. reported that in PCa, EMAST was neither related to MMRd/MSI nor to the *p53* status nor to any of the histopathological features that were studied[182]. However, the authors acknowledged that the size of the study cohort limited the conclusiveness of these findings[182]. Despite significant research efforts, the molecular root cause(s) of EMAST remained poorly understood[172]. In 2021, however, Kondelin et al. published a study in which WGS was performed on 248 colorectal cancer patients[183]. In this study, the authors conclusively showed that alterations in tri- and TNR markers were highly correlated with MSI and make a compelling case against an independent molecular root cause for EMAST. The authors suggest that observed alterations in TNRs are most likely caused by MMRd and normal MSI[183].

In 248 colorectal cancer patients, Kondelin et al. found a total of 49,055 small InDels located in TNRs, which were spread across 230 of the 256 (4^4) possible TNR motifs (see Kondelin et al. supplementary Table 3[183]). The 18 colorectal cancer samples with confirmed MSI accounted for 34,463 (70%) of the small InDels located in TNRs. Indeed, high numbers of small InDels affected TNR motifs that were previously associated with EMAST[172]. For example, the $AAAG_n$ motif ranked third with a total of 2,043 small InDels found in TNRs of this motif. Except for Penta D/E and Amelogenin, all STR markers listed in **Table 15** are TNRs. Strikingly, the TNR motifs present in the cell line authentication panel, i.e., $TATC_n$ (D13S317), $GATA_n$ (D16S539 and D7S820), $TCTA_n$ (D8S1179), $AGAT_n$ (CSF1PO and D5S818) and $AGAA_n$ (D18S51)[184] ranked 5, 6, 10, 13 and 19 among all 230 motifs, respectively. The markers D18S51, D3S1358, FGA and vWA consist of alternating TNR motifs[184] and thus cannot be ranked within this framework. The last two TNR markers in the STR panel – THO1 and TPOX – have the repeat motif $AATG_n$, which harbored less small InDels and ranked 56 of 230. In summary, Kondelin et al. showed that MSI affects TNRs of various repeat motifs – including and in particular the motifs present in the STR panel.

In DU145, LNCaP and LAPC-4 cells, TNR markers exhibited numerous deviations from the expected profiles. Small shifts of ± one repeat unit were particularly common (**Table 15**). Such small changes can be well explained by small InDels caused by MMRd. Considering the above-discussed data reported by Kondenlin et al., it is reasonable to assume that the deviations observed in TNR markers in DU145, LNCaP and LAPC-4 cells are likely a consequence of MMRd and MSI. It would be interesting to validate this hypothesis by a prospective study and to evaluate whether STR profiling using TNR markers can robustly identify MSI.

6.2 Genetic Diversity among Prostate Cancer Cell Lines

6.2.1 Shared and Unique Variants

It is unlikely that two or more independently cultured cell lines acquire the exact same genetic variant by chance. Thus, the shared variants among – for example – all six DU145 cell lines may be interpreted as the 'legacy' variants, which likely occurred either already *in vivo* or in early *in vitro* cultures of DU145. On the other hand, unique variants may be interpreted as either stochastic genetic events or – in the case of treatment-resistant cell lines – caused by mutagenic treatment (e.g., cisplatin or ionizing radiation). Unique variants are likely not as old as shared variants. For example, a unique variant present in $DU145^{A}$ (by definition not present in $DU145^{A}$ CDDPR) was most likely acquired after 2009, because that was the year when $DU145^{A}$ CDDPR were derived from $DU145^{A}$ (**Table 4**).

Furthermore, it is unlikely that two or several cells within one cell culture independently acquire the exact same genetic variant by chance. New variants emerge from single cells within the population and are propagated by cell division and clonal selection. Generally, new variants will initially exist only on one of two or several chromosomal copies. Assuming three chromosomal copies (the base case in DU145, LNCaP, LNCaP derivative and LAPC-4 cell lines), the entire culture population must be replaced by the progeny of the founder cell in order to reach a VAF of 33%. In short, it is reasonable to assume that most new (unique) variants will not establish themselves in the population in a stable manner unless they are associated with a survival or proliferation benefit for the progeny of the founder cell. Indeed, **Figure 13** shows that the overwhelming majority of unique variants among cell line groups had a VAF of less than 33%.

6.2.2 Pairwise Comparisons

Ben-David et al. conducted a study[185] on sets of cancer cell lines originating from different laboratories. The authors reanalysed two WES datasets (CCLE/DepMap headed by the Broad Institute and GDSC headed by the Sanger Institute) comprising of paired data for 106 different cancer cell lines. They reported that on average, 19% of the 'non-silent' (term used equivalent to 'coding and non-synonymous') variants found among the 106 cell line pairs were present in only one of the two samples[185] (i.e., unique). For pairwise comparisons of parental cell lines ($DU145^{A}$ vs $DU145^{B}$, $DU145^{A}$ vs $DU145^{C}$, etc.) in the herein presented dataset, an average of 28.1% of the variants were unique. These pairwise comparisons are not shown but can be inferred from **Figure 9**. This value is higher than what Ben-David and colleagues have reported. However, since the procedures used for variant calling were not uniform between this work and the discussed paper, a robust comparison is not possible. Furthermore, the herein presented dataset features cell lines from only one cancer entity while the datasets used by Ben-David et al. comprised of cell lines from various cancer entities. Finally, the proportion of cell lines exhibiting MMRd is very high in the herein presented dataset.

Pairwise comparisons between parental cell lines and their resistant counterparts (**Figure 8**) exhibited large variation in terms of the percentage of unique variants. For example, the PC-3^{A}/PC-3^{A} CDDPR pair featured 13.3% unique variants while the PC-3^{C}/PC-3^{C} RR pair featured 24.3%.

6.2.3 Factors Driving Genetic Variation

To identify and discuss potential factors that may drive genetic variation, it is useful to first introduce a simplified model aiming to describe how new variants accumulate. The formula $n = x \times t \times r \times p$ provides such a simplified model. In this formula, n is the number of stable new variants in the cell culture, x is the number of individual cells in culture, t is the total culture time, r is the rate of new mutations per time unit for a single cell and p is the probability of any new mutation to establish itself in a stable manner within the population. An example for an additional factor that is not incorporated in the model would be the number of freeze/thaw cycles as genetic damage can potentially occur during these procedures[186]. Unfortunately, most components of this formula are unknown and difficult to determine experimentally. However, the model helps to think about the potential sources of variation between laboratories, which in turn may help interpreting the numbers of variants shown in **Figure 8**, **9** and **10**.

A larger number of cells in culture (x) proportionally increases the likelihood that one cell acquires a new variant that establishes itself in the population. Consequentially, in terms of introducing new genetic variation, it probably makes a substantial difference whether laboratories use T25, T75 or other formats as a standard for cell culture. T75 cell culture flasks host roughly three times as many cells than T25 at any given time.

The total culture time (t) is likely a major source of laboratory-associated variation. For example, PC-3^{A} CDDPR were derived from PC-3^{A} in 2009, while PC-3^{C} RR were derived from PC-3^{C} around 2014 (**Table 4**). However, the total culture-time of these cell lines between 2009/2014 and the present is not known in detail. Over such periods of time, keeping consistent records poses a significant challenge as numerous scientists work with the respective cell lines. Another point to consider are the practices related to cryoconservation, which may differ between laboratories.

Another large source of variation is the rate of new mutations per time unit for a single cell (r). Numerous factors potentially affect this rate. However, a significant part of r is likely explained by sample-inherent factors and is not dependent on the laboratory. For example, DNA repair defects like MMRd will increase r. Looking at **Figure 8** and also **Table 16**, it is reasonable to assume that DU145, LNCaP and LAPC-4 cells have a higher r than PC-3, regardless of the laboratory of origin. Some cell lines were deliberately treated with DNA-damaging agents in order to pursue various experimental questions, respectively. However, such treatments must be classified as sample-inherent as they are not systematic laboratory-associated factors. DNA-damaging treatments should theoretically increase r for their duration and one would generally expect higher numbers of unique variants in all CDDP- and radiation-treated cell lines versus their parental cell lines. However, this was observed

only in some (e.g., PC-3^{D} CDDPR and LNCaP^{C} RR), but not all (e.g., PC-3^{A} CDDPR and DU145^{C} RR) cases (**Figure 8**). In addition, there may be laboratory-specific factors that influence r. Such factors could be related to the handling of the cell lines or the culture conditions but are difficult to identify.

At last, there is the probability (p) of a new mutation to establish itself in a stable manner within the population. p likely has little dependence on the laboratory. Whether a new mutation is associated with a benefit depends on its location and nature, which determined its downstream effects. Possible downstream effects include GoF or LoF of affected proteins but potentially also altered transcription patterns in the case of mutations that affect regulatory element regions. The specific location of new mutations is likely stochastic to some extent. However, the relative prevalence of different types of mutations (e.g., SNVs vs InDels) may be determined by environmental (laboratory) factors and may to some degree influence p. In summary, there are numerous factors that affect how fast new genetic variants accumulate in a population of *in vitro*-cultured cells. Some factors are exclusively determined by laboratory-associated practices while others depend more on the specific cell line. As mentioned, it may be recommended to limit the number of cultured cells to the amount necessary and use smaller cell culture formats where those are sufficient. Limiting the culture-time and passage number should also be a goal in order to avoid excess genetic variation. Cryoconservation best-practises are key to this. Making a large number of cryoconserved batches initially, helps to revisit earlier timepoints more often. New cryoconserved batches should be made from the earliest possible timepoint if possible. As discussed, there are other factors that are less dependent on the laboratory. MMRd is an example for such a factor.

6.2.4 The Mutational Burden in Prostate Cancer Cell Lines

Compared to other cancer entities, primary as well as metastatic PCa is not *per se* known to exhibit a high tumor mutational burden (TMB)[187]. Only 2 – 4% of metastatic PCa patients exhibit somatic hypermutation[188,189]. Albeit other mechanisms may cause increased TMBs in other cancer entities, in PCa, a high TMB is generally caused by defective MMR genes[188,189]. According to TCGA data[187], primary PCa tumors host 35 protein-altering mutations in the median. Sequencing of metastatic PCa samples[47] suggested 57/90 (median/average) protein altering mutations per sample. In this study, DU145 and PC-3 cell lines harbored 4,962/1,984 protein altering mutations affecting a total of 3,221/1,190 proteins in the median, respectively (**Table 16**). Albeit a robust comparison would require harmonized wet-lab and bioinformatic procedures, it is evident that the numbers seen in PCa cell lines are in stark contrast to the numbers reported for patient material. The three classical PCa cell lines PC-3, DU145 and LNCaP were established in 1978, 1979 and 1980 (**section 1.5**), respectively. As discussed, this does not necessarily suggest that the actual time in culture for the PC-3, DU145 and LNCaP cell lines in this study is similar. However, assuming a similar time in culture, MMRd is a likely explanation for the higher numbers seen in DU145 and LNCaP cell lines compared to PC-3. In addition, LAPC-4 cells exhibit MMRd. LAPC-4 were established in 1997, however, their TMB is higher than the TMB of DU145 cells and comparable to LNCaP cells. C4-2 and MR49F cells

were derived from LNCaP and exhibited a very high TMB. LAPC-4, C4-2 and MR49F have in common that all three cell lines experienced several cycles of *in vivo* culture (section **1.5**).
PC-3, DU145 and LNCaP account for the vast majority of cell line-based publications on PCa. Two of these three cell lines exhibit MMRd, which is not representative of the prevalence of MMRd in patients. Even PC-3 cells – which do not exhibit MMRd – would be characterized as hypermutated, considering the number of accumulated protein-altering mutations in these cells. In summary, one must say that in terms of the TMB, the classical PCa cell lines are not representative of *in vivo* PCa material.

6.2.5 Evolutionarily Directed Mutagenesis *in vitro*?

In the short and intermediate term, cells and tissues respond to external cues and stresses by adapting their gene regulation. For example, a heat shock leads to changes in gene expression within seconds to minutes[190]. When external cues or stresses persist over generations, evolutionary mechanisms unfold. This is true for larger organisms as well as for bacterial or cell cultures. The field of directed evolution tries to leverage evolutionary mechanisms as opposed to directed mutagenesis, in order to find new enzymes with optimized poperties[191].
If one wanted to study the short-term adaptation of cultured cells to various treatments, RNA sequencing would be the technique of choice. As mentioned before, PC-3^{A} CDDPR were derived from PC-3^{A} in 2009 and the two cell lines have since then undergone a significant number of cell divisions (generations) in separation. During that time, PC-3^{A} CDDPR were constantly exposed to 0.5 µg/mL CDDP, while PC-3^{A} were not. All treatment-resistant cell lines in the dataset were established before 2016 (**Table 4**) and except for RR cell lines, all of them are cultured under continuous selective pressure. In this setting, it is not unreasonable to hypothesize that evolutionary processes may have been at work and that evolutionarily directed genomic adaptations to the respective treatments may have taken place. Such adaptations may be discovered by WES.
After normalizing for transcript length, the top mutated genes were identified and gene set enrichment analyses were performed for each cell line (data not shown) and for cell line groups (**Table 17**). The goal was to assess whether mutations preferentially accumulate in genes that belong to certain pathways. For example, if the set of genes GO:BP 0045786 (*negative regulation of cell cycle*) would have been enriched among the top mutated genes in CDDPR cell lines, it may have suggested that dysfunctional cell cycle control is beneficial to CDDP resistance. In any case, such analysis is explorative and demands subsequent manual examination of the respective variants and their putative effects in any case.
Unfortunately, no relevant pathways in any of the treatment-resistant cell lines were identified by this strategy. The gene families that constituted the top mutated genes were rather uniform within but also across cell line groups (**Table 17**). Gene families of known high sequence variability dominated the top mutated genes. For example, the gene *HLA-DRB5* belongs to the HLA gene family, which is known for its genetic diversity across the human population. *HLA-DRB5* ranked among the top 10 mutated genes in all cell lines (**Table 17**). In DU145 cell lines for example, the group had an average of 21.3 protein-altering variants in

HLA-DRB5 per cell line, resulting in 17.1 variants / kbp (**Table 17**), assuming a longest transcript of 1,288 nt for this gene. The mean VAF for the variants in *HLA-DRB5* in DU145 cell lines was 48 % versus 50.8 % for all protein-altering variants in DU145 cell lines. The individual distinct variants for *HLA-DRB5* were overwhelmingly shared across all six DU145 cell lines with similar VAF values suggesting a homogeneous zygosity of these variants across the cell lines (data not shown). Together, the data suggest that *HLA-DRB5* was not actually the target of increased mutagenesis. Rather, the data suggest that the variants found in *HLA-DRB5* were 'normal' variation of the *HLA-DRB5* gene and likely already present in the individual from which DU145 cells were established. Apparently, filtering against the 'dbSNP common' database (**section 4.4**) was insufficient to remove these variants from the dataset. The same phenomenon was observed for members of other gene families (TRBV, KIR, IGHV, IGHL) mentioned in **section 5.2.3**. The variants in these genes were overwhelmingly shared and had similar VAF values across the cell line groups, respectively, suggesting germline rather than somatic origin. In summary, regions of increased genetic variability pose a challenge to this kind to gene set enrichment analysis. More comprehensive filtering of variants from such regions would be a prerequisite for the success of such analyses.

Next to complications arising from high physiological variability of certain genetic regions, high degrees of sequence homology may distort the identification of preferentially mutated genes and pathways. When genes display a high degree of sequence homology, the algorithms used for variant calling might map certain variants incorrectly. This may have been a problem – for example – for olfactory receptor (OR) genes, which also frequently occurred among the top mutated genes (**Table 17**). The OR gene family putatively comprises of 339 full-length OR genes with sequence homologies of up to 99%[192].

In summary, there are significant technical challenges to the identification of meaningful preferentially mutated genomic regions.

6.2.6 ZNF717 is Highly Mutated in PCa Cell Lines

The zinc finger (ZNF) protein ZNF717 ranked first among the top mutated genes in all cell line groups except PC-3, where it ranked second. ZNF proteins constitute a large protein family of DNA-binding TFs with putatively 1,510 members[193]. Many ZNF proteins (e.g., SNAI1, ZEB1) have dedicated HUGO symbols and are well studied as they act as important gene-regulatory switches[193]. Other ZNFs – like ZNF717 – are numbered. Albeit the ZNF family is large and false positive variants due to sequence homology are possible, *ZNF717* exhibited particularly high numbers of variants (**Table 17**). The distinct variants were mostly shared across the individual cell lines of the groups (data not shown), respectively, suggesting that they were present in the common ancestor of the group. However, the average VAF of protein-altering variants in *ZNF717* for all cell lines was 26.4% versus 45.5% for all genes. The lower average VAF of *ZNF717* variants potentially indicates that an increased proportion of these variants may be of somatic origin. All cell lines harboured several fatal variants (frame shifts, nonsense or nonstop mutations) in *ZNF717*. For example, PC-3^{B} DTXR cells harboured nine distinct fatal mutations that mapped to *ZNF717* with VAFs ranging from 12 to 36%, respectively.

Individually, these variants likely affect only one chromosome each. However, the chance that each chromosomal copy of *ZNF717* carries at least one fatal variant is high considering that nine distinct fatal mutations were detected. The number of distinct potentially fatal mutations in *ZNF717* ranged from five (PC-3^{D}) to sixteen (LNCaP^{A}). In summary, the data suggest that the sequenced cell lines are likely not able to express functional ZNF717 protein.

ZNF717 is not well studied. A search on NCBI ("ZNF717" on 03.10.2022) found fourteen studies mentioning ZNF717, of which eight mentioned ZNF717 in the context of human cancer. Of these eight studies, five conducted WES, two WGS and one methylation sequencing. The covered cancer entities included hepatocellular carcinoma, colorectal carcinoma, uterine leiomyosarcoma, acute myeloid lymphoma, gastric cancer, calcifying fibrous cancer and cervical cancer. None of the studies were related to PCa and hepatocellular carcinoma was the only entity with two separate studies mentioning ZNF717[194,195].
Duan et al. included a functional analysis of ZNF717[195]. A knockdown of ZNF717 promoted cell growth, adhesion, migration and invasion, while inhibiting apoptosis in HepG2 and Huh7 hepatocellular carcinoma cell lines[195]. ZNF717 knockdown also accelerated subcutaneous tumor growth in xenografts. In subsequent experiments, Duan et al. found that ZNF717 binds the promoter region of STAT3 and that it acted as a repressor of STAT3 transcription. The authors summarized that ZNF717 acts as a tumor suppressor by inhibiting the STAT3 pathway and that loss of ZNF717 was to be characterized as a driver of tumor aggressiveness[195]. As mentioned, there are no studies focusing on ZNF717 in the context of PCa. A search in the public SU2C dataset[47] on cbioportal.org (metastatic PCa samples) indeed indicated a slight but significant negative correlation (spearman R: -0.246, p = $3.3e^{-4}$) between ZNF717 and STAT3 expression in metastatic PCa. In addition, a negative correlation (spearman R: -0.396, p = $3.1e^{-9}$) between ZNF717 and STAT5a was present. The data were based on mRNA sequencing of 208 metastatic PCa samples. However, among the total 444 samples included in SU2C, only one harboured a putatively damaging mutation in ZNF717[47]. This suggests that while ZNF717 may repress STAT3 and STAT5a expression in PCa, LoF of ZNF717 seems to be an infrequent event in patients. Another study conducted WES on three patient-derived samples of calcifying fibrous cancer[196]. The authors found that all three samples harboured large numbers of damaging mutations in ZNF717 as well as in fascioscapulohumeral muscular dystrophy-1 (FRG1) and cell division cycle 27 (CDC27)[196]. The authors did not engage in functional analyses regarding the genes. Strikingly, all three genes occur among the top mutated genes in **Table 17** (e.g., DU145). In the discussed study, the authors used the SureSelect Human All Exon V6 kit supplied by Agilent for enrichment-based WES, while this study used the same kit in version five.
In summary, it may be interesting to investigate ZNF717 and other highly mutated genes in PCa cell lines within future projects. As all PCa cell lines exhibited several fatal mutations in ZNF717, one could hypothesize that ZNF717 LoF could be a prerequisite for PCa-derived cells to become viable and immortal *in vitro* cell cultures. However, it is possible that the high number of mutations detected was influenced by technical procedures.

6.2.7 Sequencing Enables the Inference of Cell Line Phylogeny

As described in section **4.4.4**, the WES data was used to infer phylogenetic trees for each group of cell lines **Figure 11**. To improve the comparability of cell line-based research, laboratories should strive to work with cell lines that are genetically as close to their original version as possible. If several cell line strains are available, it is possible to use sequencing data to measure their genetic distance to a common ancestor. For example, among the LNCaP cell lines, the current LNCaP^{A} cell line was closer to the common ancestor than the current LNCaP^{B} and LNCaP^{C} cell lines (**Figure 11**). Furthermore, the phylogenetic trees for LNCaP, DU145 and PC-3 cell lines show that the 'original' parental cell lines (**Figure 11**, turquoise dots) are significantly closer to the hypothetical common ancestor of the group than the current cell lines. This data demonstrates that it is essential for laboratories to cryoconserve their cell lines before establishing pairs of treatment-resistant and age-matched parental cells. This process is generally associated with extensive passaging, which seems to result in increasing genetic distance to the common ancestor (**Figure 11**).

For the three parental DU145 cell lines, the original DU145^{A} was closest to the common ancestor, albeit the differences between the three original cell lines were not as stark as for LNCaP cells. For LAPC-4 cells, LAPC-4^{B} was closer to their ancestor than LAPC-4^{A}. For PC-3 cells, original PC-3^{B} and PC-3^{C} were significantly closer to the ancestor of the group than the original PC-3^{A} and PC-3^{D}. The analysis suggested that PC-3^{A} and PC-3^{D} had a common parent (**Figure 11**), which was likely defined by a total of 48 variants, which were shared among PC 3^{A}, PC-3^{A} CDDPR, PC-3^{D} and PC-3^{D} CDDPR but not present in the other four PC-3 cell lines.

Phylogenetic analyses based on sequencing technologies could help to make cell line research more comparable. However, such analyses are somewhat cost intensive and would require a high level of standardization. A generally accepted reference sample would be needed. If the 'original' sample is not available anymore, a reference would first need to be established, which would require sequencing as many cell line 'strains' as possible and then determining a reference sequence or a set of variants shared by their common ancestor. Even if the original sample was still available, sequencing at least a few different cell line 'strains' would still be necessary. This is because the degree of genetic variation that can be deemed 'acceptable' will vary and has to be determined for each cell line by analyzing the available material. Sequencing data for commonly used cancer cell lines are available online. As mentioned, 106 cancer cell lines were sequenced within the CCLE/DepMap dataset[197] as well as the GDSC dataset[198]. However, these datasets include only one 'strain' per cell line and though their experimental and bioinformatic procedures are surely sound, it is unclear how viable they are as a reference. To facilitate comparability between external data and a local sample, the bioinformatic workflows must be harmonized, requiring either re-analysis of the external data or the local establishment of foreign bioinformatic pipelines. In summary, it would be a major benefit to cell line-based research if harmonized datasets that include several strains plus a reference were available for cell lines. Such data could be used by laboratories to evaluate the genetic integrity of their local cell lines.

6.3 Phenotypic diversity among Prostate Cancer Cell Lines

As a complement to WES, DU145 and PC-3 cell lines originating from different laboratories were studied phenotypically. A focus of the phenotypic studies was the EMT status of the cells.

6.3.1 Not all PC-3 Cell Lines are AR-negative

As described, not all PC-3 cell lines were negative for the AR. Especially in PC-3^{C} and its subline PC-3^{C} RR, AR transcript was clearly detected. Concomitantly, low levels of PSMA expression were detected, while PSA remained negative (**Table 21**). The presence of PSMA transcript indicates the presence of AR protein in these cells. Strikingly, the AR antibody detected a band in PC-3^{C} and PC-3^{C} RR that was not present in the other PC-3 cell lines (**Figure 18**). To corroborate these findings, PSMA should be investigated on the protein level and the western blot for the AR should be confirmed by knockdown of the protein. Together, these results suggest that PC-3^{C} and its subline PC-3^{C} RR express the AR.

In the classical literature[109,118], PC-3 cells are described as AR-negative. However, the findings described herein are in line with sporadic reports describing AR expression in PC-3 cells[199,200]. Buchanan et al.[200] investigated the differences between PC-3^{AR-} and PC-3^{AR+} cells. The authors observed morphological differences between the two sublines and found that PC-3^{AR+} proliferated faster than PC-3^{AR-}. Furthermore, Buchanan et al. noticed that cultured PC-3^{AR+} remained unresponsive to androgens despite the presence of low levels of AR protein. The authors concluded that the amount of AR protein was insufficient to promote androgen responsiveness[200] in these cells. Finally, Buchanan et al. demonstrated that transfection with an AR plasmid conferred AR responsiveness more robustly in PC-3^{AR+} than in PC-3^{AR-}. Morphologically, PC-3^{AR+} cells resembled the cuboidal growth of PC-3^{B}, whereas PC-3^{AR-} cells best resembled the spindle-shaped growth of PC-3^{A} and PC-3^{D}. Buchanan et al. did not observe lamellipodia-like protrusions in PC-3^{AR+}, which were present in PC-3^{C} (**Figure 17**). The data presented by Buchanan et al. do not contain any incidental similarities between PC-3^{AR+} and PC-3^{C}. While PC-3^{AR+} indeed proliferated faster than PC-3^{AR-}, Buchanan et al. did not perform experiments to show that this effect was actually related to AR expression. Differences in the speed of proliferation were also observed among the four PC-3 cell lines studied herein. For example, the population doubling time of PC-3^{B} was approximately 25 h, while PC-3^{A} needed approximately 32 h (data not shown). However, neither PC-3^{A} nor PC-3^{B} expressed the AR.

In summary, the causes and the implications of sporadic AR expression in some PC-3 cell cultures remain not well understood. For research questions pertaining to the AR, AR-signaling and its downstream effects, it may be advised to check the cells' AR status before conducting experimental work with PC-3 cells.

6.3.2 Epithelial versus Mesenchymal PC-3 Cell Lines

Three different morphological types of PC-3 cells were observed (**Figure 17**). PC-3^{B} exhibited cuboidal growth while PC-3^{A} and PC-3^{D} exhibited spindle-shaped growth. PC-3^{C} were characterized by lamellipodia-like cellular protrusions (**Figure 17**). The differential expression of EMT-related genes in the parental PC-3 cell lines compared to the geometric mean of the group (**Figure 20 b**) suggested that PC-3^{A} and PC-3^{D} were (relatively) more similar, while PC-3^{B} and PC-3^{C} represented more separate entities. The ten genes that were most suitable to differentiate the parental PC-3 cell lines were *SPARC*, *OCLN*, *BMP2*, *VCAN*, *DSP*, *BMP7*, *SPP1*, *KRT7*, *TGFB2* and *FN1*.

OCLN and DSP are structural components of cell-cell junctions[201,202]. Their expression was highest in PC-3^{B}, lowest in PC-3^{C} and intermediate in PC-3^{A} and PC-3^{D} (**Figure 20 b**). KRT7 is an epithelial keratin that is expressed by various epithelia, including the urothelial epithelium[203]. Likewise, its expression was highest in PC-3^{B}, lowest in PC-3^{C} and intermediate in PC-3^{A} and PC-3^{D}. The same expression pattern was observed for the epithelial cell adhesion protein CDH1, which was not among the ten most variable genes (**Figure 20 b**). Furthermore, PC-3^{B} was the only PC-3 cell line in which a notable amount of CDH1 protein was detected (**Figure 18**). The mesenchymal adhesion protein CDH2 was detected in all PC-3 cell lines. The antibody detected two bands for CDH2. The detection of multiple bands for CDH1 or CDH2 is not uncommon. Both proteins undergo post-translational modifications, which can affect their molecular weight. In addition, proteolytic cleavage by matrix metalloproteases (MMPs) and other proteases can release the extracellular domains of CDH1 and CDH2, generating 80/90 kDa fragments, respectively[204,205]. In summary, gene expression analysis suggests that PC-3^{B} and PC-3^{C} may resemble epithelial- and mesenchymal-like phenotypes of PC-3, respectively. The spindle-shaped morphology of PC-3^{A} and PC-3^{D} suggests a mesenchymal cell state, however, gene expression data suggests an intermediate state compared to PC-3^{B} and PC-3^{C}.

Another interesting group of genes that exhibited high variability between PC-3 parental cell lines constituted of *SPARC*, *SPP1*, *BMP2*, *BMP7* and *TGFB2* (**Figure 20 b**). SPARC (osteonectin) and SPP1 (osteopontin) are involved in bone homeostasis[206,207] and promote bone formation and mineralization[208]. The bone morphogenetic proteins BMP2 and BMP7 are secreted cytokines and ligands of TGFB2[209] and stimulate the differentiation of mesenchymal cells to chondrocytes and osteoblasts[209]. Physiologically, osteoblasts are responsible for bone formation by secreting specific ECM constituents such as type I collagens, SPARC and SPP1[208,209]. With minor exceptions, the expression of *BMP2*, *BMP7*, *TGFB2*, and *SPP1* was high in PC-3^{B}, low in PC-3^{C} and intermediate in PC-3^{A} and PC-3^{D} (**Figure 20 b**). However, the expression of *SPARC* and *COL1A2* (a type I collagen) was low in PC-3^{B} compared to the other PC-3 cell lines. As PC-3 cells were isolated from bone metastatic tissue, one could hypothesize that the expression of BMP proteins and bone-specific ECM constituents represents a relic associated with the origin of these cells. This hypothesis would strengthen the notion that PC-3^{B} cells may represent a more original version of PC-3 compared to the other three cell lines. Lim et al. studied the effect of BMP7 treatment on PC-3 cells[210]. BMP7 treatment

disturbed the formation of 3D spheres, reduced CDH1 expression, increased MMP expression and increased the migratory capacity of PC-3 cell lines[210]. The authors concluded that BMP7 drives PC-3 cells toward a more mesenchymal phenotype. In summary, PC-3 cell lines are highly inhomogeneous with respect to their morphology and gene expression. Epithelial versus mesenchymal properties as well as the expression of genes related to the bone microenvironment seem best suitable to define different types of PC-3 cell lines.

6.4 Resistance to Radiation or Docetaxel does not confer Resistance to Cisplatin

Platinum chemotherapy can be considered to treat mCRPC with features of neuroendocrine differentiation and/or defects in HRR. However, platinum compounds are generally only applied after the failure of multiple prior-line treatment options such as DTX and radiotherapy. To find out whether the treatment history had an influence on the response to CDDP, DTXR and RR cell lines and their parental counterparts were treated with CDDP.

In general, DU145 cells tolerated higher CDDP concentrations than PC-3 cells regardless of acquired resistances (**Figure 22 - Figure 25**). This observation may be due to the fact that DU145 cells exhibit MMRd, which increases CDDP tolerance (section **1.4.2.5**).

As shown previously[85], the DTXR cells used in this study exhibited strongly increased DTX tolerance (**Figure 29**). Furthermore, DTXR cells exhibited numerous genes that were differentially expressed relative to their parental counterparts (**Figure 19**, **Figure 20** and **Figure 21**). The most differential gene expression was observed among hand-selected genes and EMT-related genes. Particularly striking was the differential expression of ABCB1, which was highly expressed in DTXR cell lines, while the parental cell lines were negative. In addition, some DNA repair genes (e.g., *TREX1* and DMC1 in DU145^{B} DTXR versus DU145^{B}, **Figure 21 b**) also exhibited differential expression. Despite the differences in DTX tolerance and in gene expression, DTXR cell lines exhibited the same degree of CDDP tolerance as their parental counterparts (**Figure 22 - Figure 25**).

The comparison of RR cells versus their parental counterparts yielded contradicting results for PC-3 and DU145 cells. DU145^{C} and DU145^{C} RR generally exhibited comparable CDDP tolerance. However, the cell growth under CDDP pressure (**Figure 22**) and cell confluence at 144 h (**Figure 23**) indicated a slightly higher CDDP tolerance in DU145^{C} RR than in DU145^{C} cells. In contrast, PC-3^{C} RR cells were clearly more sensitive to CDDP treatment than PC-3^{C} cells based on cell growth (**Figure 22**), colony formation assays (**Figure 27**), cell confluence at 144 h (**Figure 23**) and crystal violet assays (**Figure 25**).

Radiation-induced DNA lesions include single-strand breaks, DSBs, ICLs and single base modifications by reactive oxygen species[211]. DSBs are considered the main drivers of cytotoxicity and are repaired through either HRR or variants of NHEJ[211]. The base modification 8-oxo-guanine is a particularly common product of radiation treatment[212]. 8-oxo-guanine lesions are primarily repaired through base excision repair[212]. However, transcription-coupled repair[213], NER[214] and MMR[215] have been implicated as backup and auxiliary repair systems. Evidently, the repair of the various types of DNA damage induced by ionizing radiation involves

nearly all DNA repair pathways. Likewise, the repair of CDDP lesions depends on HRR, NHEJ, the FA pathway and NER (sections **1.4.2.2 - 1.4.2.4**). However, increased activity of the MMR pathway is detrimental to CDDP tolerance (section **1.4.2.5**). This offers a potential explanation for the disparate results observed in PC-3^C RR versus DU145^C RR. 8-oxo-guanine lesions caused by radiation treatment could potentially increase the activity of the MMR pathway, which would lower the cells' tolerance to CDDP. Since the MMR pathway is dysfunctional in DU145 cells due to their *MLH1* LoF mutation, this effect would only materialize in PC-3 cells.
To test this hypothesis, MLH1 was knocked down in PC-3^B cells (**Figure 28**). According to the hypothesis, a knockdown of MLH1 would potentially increase the CDDP tolerance. PC-3^B cells were chosen as they were generally more sensitive to CDDP than PC-3^C (**Figure 22 - Figure 25**). 72 h after the siRNA transfection, MLH1 protein was reduced by 90% and the cells were treated with CDDP. The knockdown of MLH1 did not increase the CDDP tolerance of PC-3^B cells (**Figure 28**). This result contradicts the hypothesis, however, there are technicalities to consider. Albeit MLH1 protein was reduced by 90% at the start of the treatment, the durability of the downregulation over the following four days of CDDP treatment was not tracked. Furthermore, even low amounts of MLH1 might be sufficient to maintain the functionality of the MMR pathway and may thereby mask desired effects. In summary, a stable gene knockout would be more suitable to study the influence of MLH1 on platinum tolerance.
Overall, the data suggest that DTX resistance does not influence CDDP. Radio-resistance was associated with sensitization to CDDP in PC-3, but not in DU145 cells. Furthermore, the differences in CDDP tolerance observed between the parental PC-3^B and PC-3^C cells (**Figure 22 - Figure 25**) corroborate the notion that these two cell lines represent functionally different types of PC-3 cells.

6.5 Targeting ABCB1 Resensitizes Docetaxel-resistant Prostate Cancer Cell Lines

Some ABCTs exhibit particularly broad substrate specificities and are linked to MDR – a cellular state characterized by the ability to resist multiple species of cytotoxic compounds[66]. Following the observation that the fluorescent nucleic dye NucLight Rapid Red was not able to stain DU145^B DTXR and PC-3^B DTXR cells, it was hypothesized that ABCTs might be particularly active and/or abundant in these cells. Several of the 49 ACBTs, including ABCA4, ABCB1, ABCB4, ABCB11, ABCC1, ABCC2, ABCC10 and ABCG2, have been reported to be able to mediate taxane resistance in various model systems[67,164]. Consequentially, qPCR of these eight ABCTs was performed. ABCB1 exhibited impressive differential expression between DTXR cell lines and their parental counterparts. However, also ABCA3 and (in the case of DU145^B DTXR) ABCG2 were slightly upregulated in the two DTXR cell lines (**Figure 19 b**). Next, fluorescent substrates of ABCG2 (Hoechst 33342) and ABCB1 (Rhodamin 123) were employed to further narrow down which of the ABCTs may be most active/abundant in the two DTXR cell lines (**Figure 30** and data not shown). Finally flow cytometry (**Figure 31**) and western blots (data not shown) were performed. In summary, all data pointed toward ABCB1 as the most abundant and active ABCT in DU145^B DTXR and PC-3^B DTXR cells. Other groups have previously

linked ABCB1 to taxane resistance in DU145 and PC-3 cell lines[69,74,216]. However, Puhr et al. did not perform experiments related to ABCTs after establishing DU145^{B} DTXR and PC-3^{B} DTXR cells[85]. Next to a significant differential expression of ABCB1, western blots suggested that the protein was also differentially glycosylated in the two DTXR cell lines (**Figure 32**). Finally, multiple approaches to resensitize DTXR cell lines by targeting ABCB1 were explored (**Figure 33**). The approaches included pharmacologic inhibition of ABCB1 by specific small molecule inhibitors (tariquidar and elacridar), siRNA-mediated knockdown of ABCB1 as well as global inhibition of N-linked glycosylation by tunicamycin and swainsonine. In summary, inhibition with small molecules as well as siRNA-mediated knockdown were effectively resensitizing the two DTXR cell lines (**Figure 33**). Tunicamycin exhibited limited effectiveness, presumably due to dose-limiting toxicity (**Figure 33**). Finally, swainsonine was non-toxic but also non-effective (data not shown).

Pharmacologic inhibition of ABCB1 in combination with chemotherapy has been explored in clinical studies in several cancer entities. However, studies in PCa have not been conducted as of now. The results of the available clinical studies utilizing elacridar and tariquidar[217-222] were mixed. Several studies reported the sporadic occurrence of severe toxicity stemming from the chemotherapeutic agent after the pharmacologic inhibition of ABCB1 was initiated[219,221,222]. Albeit studies generally reported some degree of response to the respective treatments, the positive effects did not offset the treatment-associated risks of adverse events. As of now, no combinations of chemotherapeutic agents and inhibitors of ABCB1 are approved in cancer therapy. ABCTs protect sensitive tissues and cells from detrimental molecules and at the same time facilitate the directed flow of physiologic metabolites[66]. Systemic inhibition of ABCB1 seems to be a futile strategy, especially in a cancer entity like PCa, where the average patient is less suitable for aggressive therapies due to the patient demographic. However, ABCB1 is clearly central to taxane resistance.

Alternative approaches targeting ABCB1 could potentially yield combinatorial treatments that are better tolerated than direct pharmacologic inhibition of the protein. In this context, the glycosylation of ABCB1 in DU145^{B} DTXR and PC-3^{B} DTXR cells is interesting (**Figure 32**). The glycosylation of ABCB1 is well studied[165,166,223,224]. The protein has ten Asn-X-Ser/Thr glycosylation motifs, of which three are N-glycosylated in the mature protein[68,165]. N-linked glycosylation is a multi-step process. First, a core glycosylation consisting of a glucose(3)-mannose-(9)-N-acetylglucosamine(2) (Glc3Man9GlcNAc2) residue is added co-translationally to Asn in the endoplasmic reticulum[225,226]. Additional saccharides of varying complexity are then added to the Golgi apparatus' core as the maturing glycoprotein is headed towards its final destination[226]. In the case of ABCB1, glycosylation accounts for a significant proportion of the protein's molecular weight. Non-glycosylated ABCB1 has a molecular weight of 130 kDa, core-glycosylated protein of 150 kDa and the fully glycosylated protein of 170 kDa[227]. These observations were confirmed in DU145^{B} DTXR and PC-3^{B} DTXR cells by digesting protein lysates with PNGase F (**Figure 32**). As mentioned, the specificity of both bands shown on **Figure 32 a** was confirmed by siRNA against ABCB1 (data not shown). In contrast, Seo et al. and Takeda et al. did not detect ABCB1 protein in their taxane-naïve DU145 and PC-3 cells and only observed a single band at around 170 kDa in their taxane-resistant cells[74,216].

Global inhibition of glycosylation exhibited limited effectiveness in DU145^{B} DTXR and PC-3^{B} DTXR cells. The specific relevance of the glycosylation for the functioning of the ABCB1 protein is debated in the literature. Several reports indicate that the glycosylations are potentially required for the correct trafficking of the protein to the cell membrane but are not essential for the effluxing function *per se*[165,223,228]. Furthermore, Gribar et al. could show that glycosylated ABCB1 was less susceptible to proteolytic degradation[223]. One contradictory study reported that inhibition of glycosylation did not affect the membrane localization of ABCB1 in L1210 (murine lymphocytic leukemia) cells[229]. Furthermore, Loo and Clarke studied various mutants of ABCB1 and provided evidence that the core-glycoslation of ABCB1 was probably necessary for correct folding of the protein in the endoplasmic reticulum. Furthermore, they showed that incorrectly folded protein was more prone to proteasomal digestion[166]. In summary, the glycosylation of ABCB1 seems to be important for correct folding, maturation and trafficking of the protein, while being dispensable for its effluxing activity. In that context, one could hypothesize that the 150 kDa band observed in treatment-naïve DU145 and PC-3 cells represents an immature core-glycosylated form of the protein and that the treatment-naïve cells lack some sort of stimulus allowing for the protein to mature to its 170 kDa form.

Apart from possible roles in ABCB1 folding and trafficking, the glycosylation may also play a role in stabilizing ABCB1 at the cell membrane by interacting with galectins, which preferentially bind to highly branched glycans and thereby oligomerize membrane-bound glycoproteins[230,231]. Membrane proteins that are cross-linked in this way can exhibit an increased lifespan at the cell membrane[230]. These mechanisms have been experimentally investigated for several crucial glycosylated cell surface molecules (e.g., the epidermal growth factor receptor and CDH1)[230]. For ABCB1, however, the relation between galectins and membrane stability has received little scientific attention so far. One study reported that a knockdown of galectin-3 resensitized multi-drug-resistant Caco-2 colon cancer cells to epirubicin by inhibiting ABC transporters[232]. However, the authors of this study did not attribute the effect to the property of galectin-3 to crosslink glycoproteins. Strikingly, an omics-scale mass spectrometry study reported a direct interaction between galectin-3 and ABCB1[233].

In summary, studying the relation between galectins and ABC transporters might present a viable direction for future investigations as new ways to target ABCB1 and other ABC transporters could emerge from such research.

7 Conclusion

This dissertation addressed questions related to treatment resistance in PCa cell lines and related to the comparability of cancer cell lines originating from different laboratories. WES as well as a morphological and molecular characterization of PCa cell lines revealed that the PCa cell lines originating from different laboratories can exhibit considerable genetic and phenotypic diversity. This was particularly apparent for different strains of the cell line PC-3. Different PC-3 cell lines exhibited several distinct morphological phenotypes and several distinct patterns of gene expression. In addition, age-matched pairs of treatment-naïve and treatment-resistant cell lines originating from the same laboratory exhibit considerable numbers of unique variants. These results underscore the importance of evaluating the genetic and phenotypic integrity of the biological model system before conducting the actual experiments. Simple cell line authentication by STR profiling is not able to capture genotypic and phenotypic variability. Preclinical cancer research finds itself in a reproducibility crisis[234]. Like other more comprehensive works[185], the herein presented results suggest that at least a part of this crisis may stem from the fact that many of the cancer cell lines used in current research are not homogeneous across laboratories. This problem might be particularly distressing for cell lines that were established a long time ago. Due to their ever-decreasing cost, NGS-based genomic, transcriptomic and phylogenetic analyses could replace classical cell line authentication in the future. These techniques have the potential to overcome the outlined problems, however, establishing a widely accepted reference is not trivial.

CDDP is a treatment option for selected mCRPC patients with either histological NE features and/or HRR defects following various prior-line treatments. This thesis also studied whether resistance to DTX or radiation influenced the CDDP tolerance in DU145 and PC-3 cells. It was found that DTXR cells exhibited the same tolerance for CDDP as their parental counterparts. The same was observed for DU145 RR cells, while PC-3 RR cells were less tolerant to CDDP than their parental cell line. However, similar differences regarding the CDDP tolerance were also observed among treatment-naïve PC-3 cells originating from different laboratories. Results from cell line experiments cannot be translated into clinical recommendations. However, these results would suggest that the previous treatment history (radiation and/or DTX) does not influence the effectiveness of CDDP.

Finally, this work investigated the role of ABCB1 in DTX-resistant DU145 and PC-3 cell lines. Albeit the importance of ABCB1 for DTX resistance in other strains of DU145 and PC-3 cells was reported before, the presented functional analyses had not been done in the herein used cell lines established by Puhr et al.[85]. Furthermore, the relevance of the glycosylation of ABCB1 had not been studied in PCa cell lines before. In summary, inhibition with small molecules as well as siRNA-mediated knockdown effectively resensitized the two DTXR cell lines, whereas a global inhibition of the glycosylation failed to do so. Direct pharmacological inhibition of ABCB1 has been viable in clinical settings so far. Alternative approaches targeting ABCB1 could overcome toxicological challenges. The glycosylation of ABCB1 should receive further scientific attention. In particular, the reported interaction between galectin-3 and ABCB1[233] should be studied in the future.

8 References

1 Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209-249, doi:10.3322/caac.21660 (2021).

2 Tollis, M., Boddy, A. M. & Maley, C. C. Peto's Paradox: how has evolution solved the problem of cancer prevention? *BMC Biol* **15**, 60, doi:10.1186/s12915-017-0401-7 (2017).

3 Peto, R. Epidemiology, multistage models, and short-term mutagenicity tests. *Int J Epidemiol* **45**, 621-637, doi:10.1093/ije/dyv199 (2016).

4 Berges, R. & Oelke, M. Age-stratified normal values for prostate volume, PSA, maximum urinary flow rate, IPSS, and other LUTS/BPH indicators in the German male community-dwelling population aged 50 years or older. *World J Urol* **29**, 171-178, doi:10.1007/s00345-010-0638-z (2011).

5 Berry, S. J., Coffey, D. S., Walsh, P. C. & Ewing, L. L. The development of human benign prostatic hyperplasia with age. *J Urol* **132**, 474-479, doi:10.1016/s0022-5347(17)49698-4 (1984).

6 Shen, M. M. & Abate-Shen, C. Molecular genetics of prostate cancer: new prospects for old challenges. *Genes Dev* **24**, 1967-2000, doi:10.1101/gad.1965810 (2010).

7 Butler, W. & Huang, J. Neuroendocrine cells of the prostate: Histology, biological functions, and molecular mechanisms. *Precis Clin Med* **4**, 25-34, doi:10.1093/pcmedi/pbab003 (2021).

8 Huang, J. *et al.* Function and molecular mechanisms of neuroendocrine cells in prostate cancer. *Anal Quant Cytol Histol* **29**, 128-138 (2007).

9 Huang, Y. H., Zhang, Y. Q. & Huang, J. T. Neuroendocrine cells of prostate cancer: biologic functions and molecular mechanisms. *Asian J Androl* **21**, 291-295, doi:10.4103/aja.aja_128_18 (2019).

10 Epstein, J. I. *et al.* Proposed morphologic classification of prostate cancer with neuroendocrine differentiation. *Am J Surg Pathol* **38**, 756-767, doi:10.1097/PAS.0000000000000208 (2014).

11 Patel, A. R. & Klein, E. A. Risk factors for prostate cancer. *Nat Clin Pract Urol* **6**, 87-95, doi:10.1038/ncpuro1290 (2009).

12 Anita, L., Roger, S. K. & Simon, C. Prostate cancer. *Medicine* **44**, 47-51, doi:doi.org/10.1016/j.mpmed.2015.10.001 (2016).

13 Baig, F. A., Hamid, A., Mirza, T. & Syed, S. Ductal and Acinar Adenocarcinoma of Prostate: Morphological and Immunohistochemical Characterization. *Oman Med J* **30**, 162-166, doi:10.5001/omj.2015.36 (2015).

14 Goldstein, A. S., Huang, J., Guo, C., Garraway, I. P. & Witte, O. N. Identification of a cell of origin for human prostate cancer. *Science* **329**, 568-571, doi:10.1126/science.1189992 (2010).

15 Wang, Z. A., Toivanen, R., Bergren, S. K., Chambon, P. & Shen, M. M. Luminal cells are favored as the cell of origin for prostate cancer. *Cell Rep* **8**, 1339-1346, doi:10.1016/j.celrep.2014.08.002 (2014).

16 Stoyanova, T. *et al.* Prostate cancer originating in basal cells progresses to adenocarcinoma propagated by luminal-like cells. *Proc Natl Acad Sci U S A* **110**, 20111-20116, doi:10.1073/pnas.1320565110 (2013).

17 Bonkhoff, H. Neuroendocrine cells in benign and malignant prostate tissue: morphogenesis, proliferation, and androgen receptor status. *Prostate Suppl* **8**, 18-22 (1998).

18 Dong, B. *et al.* Single-cell analysis supports a luminal-neuroendocrine transdifferentiation in human prostate cancer. *Commun Biol* **3**, 778, doi:10.1038/s42003-020-01476-1 (2020).

19 Berger, A. *et al.* N-Myc-mediated epigenetic reprogramming drives lineage plasticity in advanced prostate cancer. *J Clin Invest* **129**, 3924-3940, doi:10.1172/JCI127961 (2019).

20 Davies, A. H., Beltran, H. & Zoubeidi, A. Cellular plasticity and the neuroendocrine phenotype in prostate cancer. *Nat Rev Urol* **15**, 271-286, doi:10.1038/nrurol.2018.22 (2018).

21 Arva, N. C. & Das, K. Diagnostic dilemmas of squamous differentiation in prostate carcinoma case report and review of the literature. *Diagn Pathol* **6**, 46, doi:10.1186/1746-1596-6-46 (2011).

22 Musser, J. E., Assel, M., Mashni, J. W., Sjoberg, D. D. & Russo, P. Adult prostate sarcoma: the Memorial Sloan Kettering experience. *Urology* **84**, 624-628, doi:10.1016/j.urology.2014.05.036 (2014).

23 Mottet, N. *et al.* EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol* **71**, 618-629, doi:10.1016/j.eururo.2016.08.003 (2017).

24 Olleik, G. *et al.* Evaluation of New Tests and Interventions for Prostate Cancer Management: A Systematic Review. *J Natl Compr Canc Netw* **16**, 1340-1351, doi:10.6004/jnccn.2018.7055 (2018).

25 Tanaka, N. *et al.* The optimal number of initial prostate biopsy cores in daily practice: a prospective study using the Nara Urological Research and Treatment Group nomogram. *BMC Res Notes* **8**, 689, doi:10.1186/s13104-015-1668-9 (2015).

26 Gleason, D. F. Classification of prostatic carcinomas. *Cancer Chemother Rep* **50**, 125-128 (1966).

27 Epstein, J. I., Allsbrook, W. C., Jr., Amin, M. B., Egevad, L. L. & Committee, I. G. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol* **29**, 1228-1242, doi:10.1097/01.pas.0000173646.99337.b1 (2005).

28 Epstein, J. I. *et al.* The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. *Am J Surg Pathol* **40**, 244-252, doi:10.1097/PAS.0000000000000530 (2016).

29 Cheng, L., Montironi, R., Bostwick, D. G., Lopez-Beltran, A. & Berney, D. M. Staging of prostate cancer. *Histopathology* **60**, 87-117, doi:10.1111/j.1365-2559.2011.04025.x (2012).

30 D'Amico, A. V. *et al.* Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* **280**, 969-974, doi:10.1001/jama.280.11.969 (1998).

31 Gandaglia, G. *et al.* Distribution of Metastatic Sites in Patients With Prostate Cancer: A Population-Based Analysis. *Prostate* **74**, 210-216, doi:10.1002/pros.22742 (2014).

32 Chandrasekar, T., Yang, J. C., Gao, A. C. & Evans, C. P. Mechanisms of resistance in castration-resistant prostate cancer (CRPC). *Transl Androl Urol* **4**, 365-380, doi:10.3978/j.issn.2223-4683.2015.05.02 (2015).

33 Cornford, P. *et al.* EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part II: Treatment of Relapsing, Metastatic, and Castration-Resistant Prostate Cancer. *Eur Urol* **71**, 630-642, doi:10.1016/j.eururo.2016.08.002 (2017).

34 Hamdy, F. C. *et al.* 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. *N Engl J Med* **375**, 1415-1424, doi:10.1056/NEJMoa1606220 (2016).

35 Knipper, S. *et al.* Survival outcomes of radical prostatectomy vs. external beam radiation therapy in prostate cancer patients with Gleason Score 9-10 at biopsy: A population-based analysis. *Urol Oncol* **38**, 79 e79-79 e14, doi:10.1016/j.urolonc.2019.09.015 (2020).

36 Zincke, H. *et al.* Long-term (15 years) results after radical prostatectomy for clinically localized (stage T2c or lower) prostate cancer. *J Urol* **152**, 1850-1857, doi:10.1016/s0022-5347(17)32399-6 (1994).

37 Thompson, I. M. *et al.* Adjuvant radiotherapy for pathological T3N0M0 prostate cancer significantly reduces risk of metastases and improves survival: long-term followup of a randomized clinical trial. *J Urol* **181**, 956-962, doi:10.1016/j.juro.2008.11.032 (2009).

38 Mottet, N. *et al.* Updated Guidelines for Metastatic Hormone-sensitive Prostate Cancer: Abiraterone Acetate Combined with Castration Is Another Standard. *Eur Urol* **73**, 316-321, doi:10.1016/j.eururo.2017.09.029 (2018).

39 Armstrong, A. J. *et al.* ARCHES: A Randomized, Phase III Study of Androgen Deprivation Therapy With Enzalutamide or Placebo in Men With Metastatic Hormone-Sensitive Prostate Cancer. *J Clin Oncol* **37**, 2974-2986, doi:10.1200/JCO.19.00799 (2019).

40 Chi, K. N. *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* **39**, 2294-2303, doi:10.1200/JCO.20.03488 (2021).

41 Cornford, P. *et al.* EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer. Part II-2020 Update: Treatment of Relapsing and Metastatic Prostate Cancer. *Eur Urol* **79**, 263-282, doi:10.1016/j.eururo.2020.09.046 (2021).

42 Barrie, S. E. *et al.* Pharmacology of novel steroidal inhibitors of cytochrome P450(17) alpha (17 alpha-hydroxylase/C17-20 lyase). *J Steroid Biochem Mol Biol* **50**, 267-273, doi:10.1016/0960-0760(94)90131-7 (1994).

43 Tran, C. *et al.* Development of a second-generation antiandrogen for treatment of advanced prostate cancer. *Science* **324**, 787-790, doi:10.1126/science.1168175 (2009).

44 Clegg, N. J. *et al.* ARN-509: a novel antiandrogen for prostate cancer treatment. *Cancer Res* **72**, 1494-1503, doi:10.1158/0008-5472.CAN-11-3948 (2012).

45 Moilanen, A. M. *et al.* Discovery of ODM-201, a new-generation androgen receptor inhibitor targeting resistance mechanisms to androgen signaling-directed prostate cancer therapies. *Sci Rep* **5**, 12007, doi:10.1038/srep12007 (2015).

46 Sartor, O. *et al.* Lutetium-177-PSMA-617 for Metastatic Castration-Resistant Prostate Cancer. *N Engl J Med* **385**, 1091-1103, doi:10.1056/NEJMoa2107322 (2021).

47 Abida, W. *et al.* Genomic correlates of clinical outcome in advanced prostate cancer. *Proc Natl Acad Sci U S A* **116**, 11428-11436, doi:10.1073/pnas.1902651116 (2019).

48 Shah, R. B. *et al.* Androgen-independent prostate cancer is a heterogeneous group of diseases: lessons from a rapid autopsy program. *Cancer Res* **64**, 9209-9216, doi:10.1158/0008-5472.CAN-04-2442 (2004).

49 Turbat-Herrera, E. A. *et al.* Neuroendocrine differentiation in prostatic carcinomas. A retrospective autopsy study. *Arch Pathol Lab Med* **112**, 1100-1105 (1988).

50 Tanaka, M. *et al.* Progression of prostate cancer to neuroendocrine cell tumor. *Int J Urol* **8**, 431-436; discussion 437, doi:10.1046/j.1442-2042.2001.00347.x (2001).

51 Donix, L. *et al.* Acquired resistance to irradiation or docetaxel is not associated with cross-resistance to cisplatin in prostate cancer cell lines. *J Cancer Res Clin Oncol*, doi:10.1007/s00432-022-03914-5 (2022).

52 de Bono, J. *et al.* Olaparib for Metastatic Castration-Resistant Prostate Cancer. *N Engl J Med* **382**, 2091-2102, doi:10.1056/NEJMoa1911440 (2020).

53 Cheng, H. H., Pritchard, C. C., Boyd, T., Nelson, P. S. & Montgomery, B. Biallelic Inactivation of BRCA2 in Platinum-sensitive Metastatic Castration-resistant Prostate Cancer. *Eur Urol* **69**, 992-995, doi:10.1016/j.eururo.2015.11.022 (2016).

54 Pomerantz, M. M. *et al.* The association between germline BRCA2 variants and sensitivity to platinum-based chemotherapy among men with metastatic prostate cancer. *Cancer* **123**, 3532-3539, doi:10.1002/cncr.30808 (2017).

55 Zafeiriou, Z. *et al.* Genomic Analysis of Three Metastatic Prostate Cancer Patients with Exceptional Responses to Carboplatin Indicating Different Types of DNA Repair Deficiency. *Eur Urol* **75**, 184-192, doi:10.1016/j.eururo.2018.09.048 (2019).

56 Schmid, S. *et al.* Activity of Platinum-Based Chemotherapy in Patients With Advanced Prostate Cancer With and Without DNA Repair Gene Aberrations. *JAMA Netw Open* **3**, e2021692, doi:10.1001/jamanetworkopen.2020.21692 (2020).

57 Sorger, P. K., Dobles, M., Tournebize, R. & Hyman, A. A. Coupling cell division and cell death to microtubule dynamics. *Curr Opin Cell Biol* **9**, 807-814, doi:10.1016/s0955-0674(97)80081-6 (1997).

58 Clarke, S. J. & Rivory, L. P. Clinical pharmacokinetics of docetaxel. *Clin Pharmacokinet* **36**, 99-114, doi:10.2165/00003088-199936020-00002 (1999).

59 de Bono, J. S. *et al.* Prednisone plus cabazitaxel or mitoxantrone for metastatic castration-resistant prostate cancer progressing after docetaxel treatment: a randomised open-label trial. *Lancet* **376**, 1147-1154, doi:10.1016/S0140-6736(10)61389-X (2010).

60 Galsky, M. D., Dritselis, A., Kirkpatrick, P. & Oh, W. K. Cabazitaxel. *Nat Rev Drug Discov* **9**, 677-678, doi:10.1038/nrd3254 (2010).

61 Maloney, S. M., Hoover, C. A., Morejon-Lasso, L. V. & Prosperi, J. R. Mechanisms of Taxane Resistance. *Cancers (Basel)* **12**, doi:10.3390/cancers12113323 (2020).

62 Bumbaca, B. & Li, W. Taxane resistance in castration-resistant prostate cancer: mechanisms and therapeutic strategies. *Acta Pharm Sin B* **8**, 518-529, doi:10.1016/j.apsb.2018.04.007 (2018).

63 Sobue, S. *et al.* Mechanism of paclitaxel resistance in a human prostate cancer cell line, PC3-PR, and its sensitization by cabazitaxel. *Biochem Biophys Res Commun* **479**, 808-813, doi:10.1016/j.bbrc.2016.09.128 (2016).

64 Terry, S. *et al.* Increased expression of class III beta-tubulin in castration-resistant human prostate cancer. *Br J Cancer* **101**, 951-956, doi:10.1038/sj.bjc.6605245 (2009).

65 Vasiliou, V., Vasiliou, K. & Nebert, D. W. Human ATP-binding cassette (ABC) transporter family. *Hum Genomics* **3**, 281-290, doi:10.1186/1479-7364-3-3-281 (2009).

66 Szakacs, G., Paterson, J. K., Ludwig, J. A., Booth-Genthe, C. & Gottesman, M. M. Targeting multidrug resistance in cancer. *Nat Rev Drug Discov* **5**, 219-234, doi:10.1038/nrd1984 (2006).

67 Ween, M. P., Armstrong, M. A., Oehler, M. K. & Ricciardelli, C. The role of ABC transporters in ovarian cancer progression and chemoresistance. *Crit Rev Oncol Hematol* **96**, 220-256, doi:10.1016/j.critrevonc.2015.05.012 (2015).

68 Chen, C. J. *et al.* Internal duplication and homology with bacterial transport proteins in the mdr1 (P-glycoprotein) gene from multidrug-resistant human cells. *Cell* **47**, 381-389, doi:10.1016/0092-8674(86)90595-7 (1986).

69 Zhu, Y. *et al.* Inhibition of ABCB1 expression overcomes acquired docetaxel resistance in prostate cancer. *Mol Cancer Ther* **12**, 1829-1836, doi:10.1158/1535-7163.MCT-13-0208 (2013).

70 Zhu, Y. *et al.* Antiandrogens Inhibit ABCB1 Efflux and ATPase Activity and Reverse Docetaxel Resistance in Advanced Prostate Cancer. *Clin Cancer Res* **21**, 4133-4142, doi:10.1158/1078-0432.CCR-15-0269 (2015).

71 Lombard, A. P. *et al.* ABCB1 Mediates Cabazitaxel-Docetaxel Cross-Resistance in Advanced Prostate Cancer. *Mol Cancer Ther* **16**, 2257-2266, doi:10.1158/1535-7163.MCT-17-0179 (2017).

72 Lombard, A. P. *et al.* Overexpressed ABCB1 Induces Olaparib-Taxane Cross-Resistance in Advanced Prostate Cancer. *Transl Oncol* **12**, 871-878, doi:10.1016/j.tranon.2019.04.007 (2019).

73 Lombard, A. P. *et al.* Activation of the ABCB1 Amplicon in Docetaxel- and Cabazitaxel-Resistant Prostate Cancer Cells. *Mol Cancer Ther* **20**, 2061-2070, doi:10.1158/1535-7163.MCT-20-0983 (2021).

74 Seo, H. K., Lee, S. J., Kwon, W. A. & Jeong, K. C. Docetaxel-resistant prostate cancer cells become sensitive to gemcitabine due to the upregulation of ABCB1. *Prostate* **80**, 453-462, doi:10.1002/pros.23946 (2020).

75 Ylitalo, E. B. *et al.* Marked response to cabazitaxel in prostate cancer xenografts expressing androgen receptor variant 7 and reversion of acquired resistance by anti-androgens. *Prostate* **80**, 214-224, doi:10.1002/pros.23935 (2020).

76 Hyafil, F., Vergely, C., Du Vignaud, P. & Grand-Perret, T. In vitro and in vivo reversal of multidrug resistance by GF120918, an acridonecarboxamide derivative. *Cancer Res* **53**, 4595-4602 (1993).

77 Mistry, P. *et al.* In vitro and in vivo reversal of P-glycoprotein-mediated multidrug resistance by a novel potent modulator, XR9576. *Cancer Res* **61**, 749-758 (2001).

78 Lai, J. I., Tseng, Y. J., Chen, M. H., Huang, C. F. & Chang, P. M. Clinical Perspective of FDA Approved Drugs With P-Glycoprotein Inhibition Activities for Potential Cancer Therapeutics. *Front Oncol* **10**, 561936, doi:10.3389/fonc.2020.561936 (2020).

79 Fernald, K. & Kurokawa, M. Evading apoptosis in cancer. *Trends Cell Biol* **23**, 620-633, doi:10.1016/j.tcb.2013.07.006 (2013).

80 Terrano, D. T., Upreti, M. & Chambers, T. C. Cyclin-dependent kinase 1-mediated Bcl-xL/Bcl-2 phosphorylation acts as a functional link coupling mitotic arrest and apoptosis. *Mol Cell Biol* **30**, 640-656, doi:10.1128/MCB.00882-09 (2010).

81 Harley, M. E., Allan, L. A., Sanderson, H. S. & Clarke, P. R. Phosphorylation of Mcl-1 by CDK1-cyclin B1 initiates its Cdc20-dependent destruction during mitotic arrest. *EMBO J* **29**, 2407-2420, doi:10.1038/emboj.2010.112 (2010).

82 Thadani-Mulero, M., Nanus, D. M. & Giannakakou, P. Androgen receptor on the move: boarding the microtubule expressway to the nucleus. *Cancer Res* **72**, 4611-4615, doi:10.1158/0008-5472.CAN-12-0783 (2012).

83 Thadani-Mulero, M. *et al.* Androgen receptor splice variants determine taxane sensitivity in prostate cancer. *Cancer Res* **74**, 2270-2282, doi:10.1158/0008-5472.CAN-13-2876 (2014).

84 Shan, X. *et al.* Replication study: androgen receptor splice variants determine taxane sensitivity in prostate cancer. *PeerJ* **6**, e4661, doi:10.7717/peerj.4661 (2018).

85 Puhr, M. *et al.* Epithelial-to-mesenchymal transition leads to docetaxel resistance in prostate cancer and is mediated by reduced expression of miR-200c and miR-205. *Am J Pathol* **181**, 2188-2201, doi:10.1016/j.ajpath.2012.08.011 (2012).

86 Marin-Aguilera, M. *et al.* Epithelial-to-mesenchymal transition mediates docetaxel resistance and high risk of relapse in prostate cancer. *Mol Cancer Ther* **13**, 1270-1284, doi:10.1158/1535-7163.MCT-13-0775 (2014).

87 Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol* **15**, 178-196, doi:10.1038/nrm3758 (2014).

88 Kelland, L. The resurgence of platinum-based cancer chemotherapy. *Nat Rev Cancer* **7**, 573-584, doi:10.1038/nrc2167 (2007).

89 Duan, M., Ulibarri, J., Liu, K. J. & Mao, P. Role of Nucleotide Excision Repair in Cisplatin Resistance. *Int J Mol Sci* **21**, doi:10.3390/ijms21239248 (2020).

90 Deans, A. J. & West, S. C. DNA interstrand crosslink repair and cancer. *Nat Rev Cancer* **11**, 467-480, doi:10.1038/nrc3088 (2011).

91 Frankenberg-Schwager, M. *et al.* Cisplatin-mediated DNA double-strand breaks in replicating but not in quiescent cells of the yeast Saccharomyces cerevisiae. *Toxicology* **212**, 175-184, doi:10.1016/j.tox.2005.04.015 (2005).

92 Scully, R., Panday, A., Elango, R. & Willis, N. A. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat Rev Mol Cell Biol* **20**, 698-714, doi:10.1038/s41580-019-0152-0 (2019).

93 Zhao, E. Y. *et al.* Homologous Recombination Deficiency and Platinum-Based Therapy Outcomes in Advanced Breast Cancer. *Clin Cancer Res* **23**, 7521-7530, doi:10.1158/1078-0432.CCR-17-1941 (2017).

94 Alsop, K. *et al.* BRCA mutation frequency and patterns of treatment response in BRCA mutation-positive women with ovarian cancer: a report from the Australian Ovarian Cancer Study Group. *J Clin Oncol* **30**, 2654-2663, doi:10.1200/JCO.2011.39.8545 (2012).

95 Pokataev, I. *et al.* Efficacy of platinum-based chemotherapy and prognosis of patients with pancreatic cancer with homologous recombination deficiency: comparative analysis of published clinical studies. *ESMO Open* **5**, e000578, doi:10.1136/esmoopen-2019-000578 (2020).

96 Kunkel, T. A. & Erie, D. A. DNA mismatch repair. *Annu Rev Biochem* **74**, 681-710, doi:10.1146/annurev.biochem.74.082803.133243 (2005).

97 Aebi, S. *et al.* Loss of DNA mismatch repair in acquired resistance to cisplatin. *Cancer Res* **56**, 3087-3090 (1996).

98 Fink, D. *et al.* The role of DNA mismatch repair in platinum drug resistance. *Cancer Res* **56**, 4881-4886 (1996).

99 Sawant, A., Kothandapani, A., Zhitkovich, A., Sobol, R. W. & Patrick, S. M. Role of mismatch repair proteins in the processing of cisplatin interstrand cross-links. *DNA Repair (Amst)* **35**, 126-136, doi:10.1016/j.dnarep.2015.10.003 (2015).

100 Bassett, E. *et al.* Frameshifts and deletions during in vitro translesion synthesis past Pt-DNA adducts by DNA polymerases beta and eta. *DNA Repair (Amst)* **1**, 1003-1016, doi:10.1016/s1568-7864(02)00150-7 (2002).

101 Albertella, M. R., Green, C. M., Lehmann, A. R. & O'Connor, M. J. A role for polymerase eta in the cellular tolerance to cisplatin-induced damage. *Cancer Res* **65**, 9799-9806, doi:10.1158/0008-5472.CAN-05-1095 (2005).

102 Roos, W. P. & Kaina, B. DNA damage-induced cell death by apoptosis. *Trends Mol Med* **12**, 440-450, doi:10.1016/j.molmed.2006.07.007 (2006).

103 Michels, J. *et al.* MCL-1 dependency of cisplatin-resistant cancer cells. *Biochem Pharmacol* **92**, 55-61, doi:10.1016/j.bcp.2014.07.029 (2014).

104 Ishida, S., Lee, J., Thiele, D. J. & Herskowitz, I. Uptake of the anticancer drug cisplatin mediated by the copper transporter Ctr1 in yeast and mammals. *Proc Natl Acad Sci U S A* **99**, 14298-14302, doi:10.1073/pnas.162491399 (2002).

105 Holzer, A. K., Manorek, G. H. & Howell, S. B. Contribution of the major copper influx transporter CTR1 to the cellular accumulation of cisplatin, carboplatin, and oxaliplatin. *Mol Pharmacol* **70**, 1390-1394, doi:10.1124/mol.106.022624 (2006).

106 Holzer, A. K. & Howell, S. B. The internalization and degradation of human copper transporter 1 following cisplatin exposure. *Cancer Res* **66**, 10944-10952, doi:10.1158/0008-5472.CAN-06-1710 (2006).

107 Safaei, R., Holzer, A. K., Katano, K., Samimi, G. & Howell, S. B. The role of copper transporters in the development of resistance to Pt drugs. *J Inorg Biochem* **98**, 1607-1613, doi:10.1016/j.jinorgbio.2004.05.006 (2004).

108 Brozovic, A. The relationship between platinum drug resistance and epithelial-mesenchymal transition. *Arch Toxicol* **91**, 605-619, doi:10.1007/s00204-016-1912-7 (2017).

109 Sobel, R. E. & Sadar, M. D. Cell lines used in prostate cancer research: a compendium of old and new lines--part 1. *J Urol* **173**, 342-359, doi:10.1097/01.ju.0000141580.30910.57 (2005).

110 Stone, K. R., Mickey, D. D., Wunderli, H., Mickey, G. H. & Paulson, D. F. Isolation of a human prostate carcinoma cell line (DU 145). *Int J Cancer* **21**, 274-281, doi:10.1002/ijc.2910210305 (1978).

111 Kaighn, M. E., Narayan, K. S., Ohnuki, Y., Lechner, J. F. & Jones, L. W. Establishment and characterization of a human prostatic carcinoma cell line (PC-3). *Invest Urol* **17**, 16-23 (1979).

112 Horoszewicz, J. S. *et al.* The LNCaP cell line--a new model for studies on human prostatic carcinoma. *Prog Clin Biol Res* **37**, 115-132 (1980).

113 Klein, K. A. *et al.* Progression of metastatic human prostate cancer to androgen independence in immunodeficient SCID mice. *Nat Med* **3**, 402-408, doi:10.1038/nm0497-402 (1997).

114 Thalmann, G. N. *et al.* LNCaP progression model of human prostate cancer: androgen-independence and osseous metastasis. *Prostate* **44**, 91-103 Jul 101;144(102), doi:10.1002/1097-0045(20000701)44:2<91::aid-pros1>3.0.co;2-l (2000).

115 Wu, H. C. *et al.* Derivation of androgen-independent human LNCaP prostatic cancer cell sublines: role of bone stromal cells. *Int J Cancer* **57**, 406-412, doi:10.1002/ijc.2910570319 (1994).

116 Kuruma, H. *et al.* A novel antiandrogen, Compound 30, suppresses castration-resistant and MDV3100-resistant prostate cancer growth in vitro and in vivo. *Mol Cancer Ther* **12**, 567-576, doi:10.1158/1535-7163.MCT-12-0798 (2013).

117 Boyer, J. C. *et al.* Microsatellite instability, mismatch repair deficiency, and genetic defects in human cancer cell lines. *Cancer Res* **55**, 6063-6070 (1995).

118 van Bokhoven, A. *et al.* Molecular characterization of human prostate carcinoma cell lines. *Prostate* **57**, 205-225, doi:10.1002/pros.10290 (2003).

119 Tai, S. *et al.* PC3 is a cell line characteristic of prostatic small cell carcinoma. *Prostate* **71**, 1668-1679, doi:10.1002/pros.21383 (2011).

120 Carroll, A. G., Voeller, H. J., Sugars, L. & Gelmann, E. P. p53 oncogene mutations in three human prostate cancer cell lines. *Prostate* **23**, 123-134, doi:10.1002/pros.2990230206 (1993).

121 Vlietstra, R. J., van Alewijk, D. C., Hermans, K. G., van Steenbrugge, G. J. & Trapman, J. Frequent inactivation of PTEN in prostate cancer cell lines and xenografts. *Cancer Res* **58**, 2720-2723 (1998).

122 Clark, J. *et al.* Genome-wide screening for complete genetic loss in prostate cancer by comparative hybridization onto cDNA microarrays. *Oncogene* **22**, 1247-1252, doi:10.1038/sj.onc.1206247 (2003).

123 Leach, F. S., Velasco, A., Hsieh, J. T., Sagalowsky, A. I. & McConnell, J. D. The mismatch repair gene hMSH2 is mutated in the prostate cancer cell line LNCaP. *J Urol* **164**, 1830-1833 (2000).

124 *GRCh38*, https://www.ncbi.nlm.nih.gov/data-hub/genome/GCA_000001405.15/ (2013).

125 Koboldt, D. C. Best practices for variant calling in clinical sequencing. *Genome Med* **12**, 91, doi:10.1186/s13073-020-00791-w (2020).

126 *GDC MAF Format v.1.0.0* https://docs.gdc.cancer.gov/Data/File_Formats/MAF_Format/

127 *HGVS Nomenclature*, https://varnomen.hgvs.org/

128 *GRCh37*, https://www.ncbi.nlm.nih.gov/data-hub/genome/GCA_000001405.1/ (2009).

129 Meienberg, J., Bruggmann, R., Oexle, K. & Matyas, G. Clinical sequencing: is WGS the better WES? *Hum Genet* **135**, 359-362, doi:10.1007/s00439-015-1631-9 (2016).

130 Escaramis, G., Docampo, E. & Rabionet, R. A decade of structural variants: description, history and methods to detect structural variation. *Brief Funct Genomics* **14**, 305-314, doi:10.1093/bfgp/elv014 (2015).

131 Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415-421, doi:10.1038/nature12477 (2013).

132 Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J. & Stratton, M. R. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep* **3**, 246-259, doi:10.1016/j.celrep.2012.12.008 (2013).

133 Alexandrov, L. B. *et al.* The repertoire of mutational signatures in human cancer. *Nature* **578**, 94-101, doi:10.1038/s41586-020-1943-3 (2020).

134 COSMIC. *Mutational Signatures (v3.3 - 2022)*, https://cancer.sanger.ac.uk/signatures/sbs/

135 *The Resistant Cancer Cell Line (RCCL) collection*, https://research.kent.ac.uk/industrial-biotechnology-centre/the-resistant-cancer-cell-line-rccl-collection/

136 Hobisch, A. *et al.* The androgen receptor pathway is by-passed in prostate cancer cells generated after prolonged treatment with bicalutamide. *Prostate* **66**, 413-420, doi:10.1002/pros.20365 (2006).

137 Hoefer, J. *et al.* Critical role of androgen receptor level in prostate cancer cell resistance to new generation antiandrogen enzalutamide. *Oncotarget* **7**, 59781-59794, doi:10.18632/oncotarget.10926 (2016).

138 Cojoc, M. *et al.* Aldehyde Dehydrogenase Is Regulated by beta-Catenin/TCF and Promotes Radioresistance in Prostate Cancer Progenitor Cells. *Cancer Res* **75**, 1482-1494, doi:10.1158/0008-5472.CAN-14-1924 (2015).

139 *Cellosaurus*, https://www.cellosaurus.org/

140 Bacher, J. W. *et al.* Development of a fluorescent multiplex assay for detection of MSI-High tumors. *Dis Markers* **20**, 237-250, doi:10.1155/2004/136734 (2004).

141 Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**, 308-311, doi:10.1093/nar/29.1.308 (2001).

142 Sean D. Mooney, C. W., Andrew Su and contributors. *myVariant.info: Variant Annotation as a Service*, https://myvariant.info/about

143 CCDS, *CCDS_exons.current.txt,* https://ftp.ncbi.nlm.nih.gov/pub/CCDS/current_human/

144 Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304-1351, doi:10.1126/science.1058040 (2001).

145 Cunningham, F. *et al.* Ensembl 2022. *Nucleic Acids Res* **50**, D988-D995, doi:10.1093/nar/gkab1049 (2022).

146 Tweedie, S. *et al.* Genenames.org: the HGNC and VGNC resources in 2021. *Nucleic Acids Res* **49**, D939-D946, doi:10.1093/nar/gkaa980 (2021).

147 Raudvere, U. *et al.* g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* **47**, W191-W198, doi:10.1093/nar/gkz369 (2019).

148 tctianchi. *pyvenn (A python package to create 2-6 set Venn diagrams)*, https://github.com/tctianchi/pyvenn

149 Griggs, J., Killian, C. E. & Savage, C. D. Venn Diagrams and Symmetric Chain Decompositions in the Boolean Lattice. *The Electronic Journal of Combinatorics* **11**, doi:10.37236/1755 (2004).

150 Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R. & Pfister, H. UpSet: Visualization of Intersecting Sets. *IEEE Trans Vis Comput Graph* **20**, 1983-1992, doi:10.1109/TVCG.2014.2346248 (2014).

151 Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526-528, doi:10.1093/bioinformatics/bty633 (2019).

152 Gascuel, O. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol Biol Evol* **14**, 685-695, doi:10.1093/oxfordjournals.molbev.a025808 (1997).

153 Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C. & Koeffler, H. P. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* **28**, 1747-1756, doi:10.1101/gr.239244.118 (2018).

154 Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res* **47**, D941-D947, doi:10.1093/nar/gky1015 (2019).

155 *SBS Signatures*, https://cancer.sanger.ac.uk/cosmic/signatures/SBS/

156 Berridge, M. V. & Tan, A. S. Trans-plasma membrane electron transport: a cellular assay for NADH- and NADPH-oxidase based on extracellular, superoxide-mediated reduction of the sulfonated tetrazolium salt WST-1. *Protoplasma* **205**, 74-82, doi:Doi 10.1007/Bf01279296 (1998).

157 Reid, Y., Storts, D., Riss, T. & Minor, L. in *Assay Guidance Manual* (2004).

158 Innocenti, M. New insights into the formation and the function of lamellipodia and ruffles in mesenchymal cell migration. *Cell Adh Migr* **12**, 401-416, doi:10.1080/19336918.2018.1448352 (2018).

159 Boland, C. R. *et al.* A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* **58**, 5248-5257 (1998).

160 Hegde, M. *et al.* ACMG technical standards and guidelines for genetic testing for inherited colorectal cancer (Lynch syndrome, familial adenomatous polyposis, and MYH-associated polyposis). *Genet Med* **16**, 101-116, doi:10.1038/gim.2013.166 (2014).

161 Chen, Y. *et al.* Defects of DNA mismatch repair in human prostate cancer. *Cancer Res* **61**, 4112-4121 (2001).

162 Seigel, G. M. & Campbell, L. M. High-throughput microtiter assay for Hoechst 33342 dye uptake. *Cytotechnology* **45**, 155-160, doi:10.1007/s10616-004-7256-9 (2004).

163 Chaudhary, P. M. & Roninson, I. B. Expression and activity of P-glycoprotein, a multidrug efflux pump, in human hematopoietic stem cells. *Cell* **66**, 85-94, doi:10.1016/0092-8674(91)90141-k (1991).

164 Mohammad, I. S., He, W. & Yin, L. Understanding of human ATP binding cassette superfamily and novel multidrug resistance modulators to overcome MDR. *Biomed Pharmacother* **100**, 335-348, doi:10.1016/j.biopha.2018.02.038 (2018).

165 Schinkel, A. H., Kemp, S., Dolle, M., Rudenko, G. & Wagenaar, E. N-glycosylation and deletion mutants of the human MDR1 P-glycoprotein. *J Biol Chem* **268**, 7474-7481 (1993).

166 Loo, T. W. & Clarke, D. M. Quality control by proteases in the endoplasmic reticulum. Removal of a protease-sensitive site enhances expression of human P-glycoprotein. *J Biol Chem* **273**, 32373-32376, doi:10.1074/jbc.273.49.32373 (1998).

167 Elder, J. H. & Alexander, S. endo-beta-N-acetylglucosaminidase F: endoglycosidase from Flavobacterium meningosepticum that cleaves both high-mannose and complex glycoproteins. *Proc Natl Acad Sci U S A* **79**, 4540-4544, doi:10.1073/pnas.79.15.4540 (1982).

168 Hasilik, A. & Tanner, W. Inhibition of the apparent rate of synthesis on the vacuolar glycoprotein carboxypeptidase Y and its protein antigen by tunicamycin in Saccharomyces cerevisiae. *Antimicrob Agents Chemother* **10**, 402-410, doi:10.1128/AAC.10.3.402 (1976).

169 Yoo, J. *et al.* GlcNAc-1-P-transferase-tunicamycin complex structure reveals basis for inhibition of N-glycosylation. *Nat Struct Mol Biol* **25**, 217-224, doi:10.1038/s41594-018-0031-y (2018).

170 Dorling, P. R., Huxtable, C. R. & Colegate, S. M. Inhibition of lysosomal alpha-mannosidase by swainsonine, an indolizidine alkaloid isolated from Swainsona canescens. *Biochem J* **191**, 649-651, doi:10.1042/bj1910649 (1980).

171 Kuntz, D. A. *et al.* Structural investigation of the binding of 5-substituted swainsonine analogues to Golgi alpha-mannosidase II. *Chembiochem* **11**, 673-680, doi:10.1002/cbic.200900750 (2010).

172 Watson, M. M., Berg, M. & Soreide, K. Prevalence and implications of elevated microsatellite alterations at selected tetranucleotides in cancer. *Br J Cancer* **111**, 823-827, doi:10.1038/bjc.2014.167 (2014).

173 Thibodeau, S. N., Bren, G. & Schaid, D. Microsatellite instability in cancer of the proximal colon. *Science* **260**, 816-819, doi:10.1126/science.8484122 (1993).

174 Ionov, Y., Peinado, M. A., Malkhosyan, S., Shibata, D. & Perucho, M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* **363**, 558-561, doi:10.1038/363558a0 (1993).

175 Bronner, C. E. *et al.* Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature* **368**, 258-261, doi:10.1038/368258a0 (1994).

176 Fishel, R. *et al.* The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* **75**, 1027-1038, doi:10.1016/0092-8674(93)90546-3 (1993).
177 Sun, X., Chen, C., Vessella, R. L. & Dong, J. T. Microsatellite instability and mismatch repair target gene mutations in cell lines and xenografts of prostate cancer. *Prostate* **66**, 660-666, doi:10.1002/pros.20390 (2006).
178 Lu, Y., Soong, T. D. & Elemento, O. A novel approach for characterizing microsatellite instability in cancer cells. *PLoS One* **8**, e63056, doi:10.1371/journal.pone.0063056 (2013).
179 Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* **7**, 248-249, doi:10.1038/nmeth0410-248 (2010).
180 Ng, P. C. & Henikoff, S. Predicting deleterious amino acid substitutions. *Genome Res* **11**, 863-874, doi:10.1101/gr.176601 (2001).
181 Ahrendt, S. A. *et al.* Microsatellite instability at selected tetranucleotide repeats is associated with p53 mutations in non-small cell lung cancer. *Cancer Res* **60**, 2488-2491 (2000).
182 Burger, M. *et al.* Elevated microsatellite alterations at selected tetranucleotides (EMAST) and mismatch repair gene expression in prostate cancer. *J Mol Med (Berl)* **84**, 833-841, doi:10.1007/s00109-006-0074-0 (2006).
183 Kondelin, J. *et al.* No evidence of EMAST in whole genome sequencing data from 248 colorectal cancers. *Genes Chromosomes Cancer* **60**, 463-473, doi:10.1002/gcc.22941 (2021).
184 Ruitberg, C. M., Reeder, D. J. & Butler, J. M. STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acids Res* **29**, 320-322, doi:10.1093/nar/29.1.320 (2001).
185 Ben-David, U. *et al.* Genetic and transcriptional evolution alters cancer cell line drug response. *Nature* **560**, 325-330, doi:10.1038/s41586-018-0409-3 (2018).
186 Shao, W., Khin, S. & Kopp, W. C. Characterization of effect of repeated freeze and thaw cycles on stability of genomic DNA using pulsed field gel electrophoresis. *Biopreserv Biobank* **10**, 4-11, doi:10.1089/bio.2011.0016 (2012).
187 Castle, J. C., Uduman, M., Pabla, S., Stein, R. B. & Buell, J. S. Mutation-Derived Neoantigens for Cancer Immunotherapy. *Front Immunol* **10**, 1856, doi:10.3389/fimmu.2019.01856 (2019).
188 Robinson, D. *et al.* Integrative Clinical Genomics of Advanced Prostate Cancer. *Cell* **162**, 454, doi:10.1016/j.cell.2015.06.053 (2015).
189 Ritch, E. *et al.* Identification of Hypermutation and Defective Mismatch Repair in ctDNA from Metastatic Prostate Cancer. *Clin Cancer Res* **26**, 1114-1125, doi:10.1158/1078-0432.CCR-19-1623 (2020).
190 Richter, K., Haslbeck, M. & Buchner, J. The heat shock response: life on the verge of death. *Mol Cell* **40**, 253-266, doi:10.1016/j.molcel.2010.10.006 (2010).
191 Molina, R. S. *et al.* In vivo hypermutation and continuous evolution. *Nature Reviews Methods Primers* **2**, 36, doi:10.1038/s43586-022-00119-5 (2022).
192 Malnic, B., Godfrey, P. A. & Buck, L. B. The human olfactory receptor gene family. *Proc Natl Acad Sci U S A* **101**, 2584-2589, doi:10.1073/pnas.0307882100 (2004).
193 Cassandri, M. *et al.* Zinc-finger proteins in health and disease. *Cell Death Discov* **3**, 17071, doi:10.1038/cddiscovery.2017.71 (2017).
194 Chen, Y., Wang, L., Xu, H., Liu, X. & Zhao, Y. Exome capture sequencing reveals new insights into hepatitis B virus-induced hepatocellular carcinoma at the early stage of tumorigenesis. *Oncol Rep* **30**, 1906-1912, doi:10.3892/or.2013.2652 (2013).
195 Duan, M. *et al.* Diverse modes of clonal evolution in HBV-related hepatocellular carcinoma revealed by single-cell genome sequencing. *Cell Res* **28**, 359-373, doi:10.1038/cr.2018.11 (2018).
196 Mehrad, M., LaFramboise, W. A., Lyons, M. A., Trejo Bittar, H. E. & Yousem, S. A. Whole-exome sequencing identifies unique mutations and copy number losses in calcifying

fibrous tumor of the pleura: report of 3 cases and review of the literature. *Hum Pathol* **78**, 36-43, doi:10.1016/j.humpath.2018.04.005 (2018).

197 Ghandi, M. *et al.* Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503-508, doi:10.1038/s41586-019-1186-3 (2019).

198 Yang, W. *et al.* Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* **41**, D955-961, doi:10.1093/nar/gks1111 (2013).

199 Alimirah, F., Chen, J., Basrawala, Z., Xin & H. DU-145 and PC-3 human prostate cancer cell lines express androgen receptor: implications for the androgen receptor functions and regulation. *FEBS Lett* **580**, 2294-2300, doi:10.1016/j.febslet.2006.03.041 (2006).

200 Buchanan, G. *et al.* PC-3 cells with enhanced androgen receptor signaling: a model for clonal selection in prostate cancer. *Prostate* **60**, 352-366, doi:10.1002/pros.20079 (2004).

201 Cummins, P. M. Occludin: one protein, many forms. *Mol Cell Biol* **32**, 242-250, doi:10.1128/MCB.06029-11 (2012).

202 Hatsell, S. & Cowin, P. Deconstructing desmoplakin. *Nat Cell Biol* **3**, E270-272, doi:10.1038/ncb1201-e270 (2001).

203 Moll, R., Divo, M. & Langbein, L. The human keratins: biology and pathology. *Histochem Cell Biol* **129**, 705-733, doi:10.1007/s00418-008-0435-6 (2008).

204 Derycke, L. *et al.* Soluble N-cadherin in human biological fluids. *Int J Cancer* **119**, 2895-2900, doi:10.1002/ijc.22219 (2006).

205 Grabowska, M. M. & Day, M. L. Soluble E-cadherin: more than a symptom of disease. *Front Biosci (Landmark Ed)* **17**, 1948-1964, doi:10.2741/4031 (2012).

206 Rosset, E. M. & Bradshaw, A. D. SPARC/osteonectin in mineralized tissue. *Matrix Biol* **52-54**, 78-87, doi:10.1016/j.matbio.2016.02.001 (2016).

207 Si, J., Wang, C., Zhang, D., Wang, B. & Zhou, Y. Osteopontin in Bone Metabolism and Bone Diseases. *Med Sci Monit* **26**, e919159, doi:10.12659/MSM.919159 (2020).

208 Lin, X., Patil, S., Gao, Y. G. & Qian, A. The Bone Extracellular Matrix in Bone Formation and Regeneration. *Front Pharmacol* **11**, 757, doi:10.3389/fphar.2020.00757 (2020).

209 Katagiri, T. & Watabe, T. Bone Morphogenetic Proteins. *Cold Spring Harb Perspect Biol* **8**, doi:10.1101/cshperspect.a021899 (2016).

210 Lim, M., Chuong, C. M. & Roy-Burman, P. PI3K, Erk signaling in BMP7-induced epithelial-mesenchymal transition (EMT) of PC-3 prostate cancer cells in 2- and 3-dimensional cultures. *Horm Cancer* **2**, 298-309, doi:10.1007/s12672-011-0084-4 (2011).

211 Mladenov, E., Magin, S., Soni, A. & Iliakis, G. DNA double-strand break repair as determinant of cellular radiosensitivity to killing and target in radiation therapy. *Front Oncol* **3**, 113, doi:10.3389/fonc.2013.00113 (2013).

212 David, S. S., O'Shea, V. L. & Kundu, S. Base-excision repair of oxidative DNA damage. *Nature* **447**, 941-950, doi:10.1038/nature05978 (2007).

213 Le Page, F., Klungland, A., Barnes, D. E., Sarasin, A. & Boiteux, S. Transcription coupled repair of 8-oxoguanine in murine cells: the ogg1 protein is required for repair in nontranscribed sequences but not in transcribed sequences. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 8397-8402, doi:10.1073/pnas.140137297 (2000).

214 Scott, A. D. *et al.* Spontaneous mutation, oxidative DNA damage, and the roles of base and nucleotide excision repair in the yeast Saccharomyces cerevisiae. *Yeast* **15**, 205-218, doi:10.1002/(SICI)1097-0061(199902)15:3<205::AID-YEA361>3.0.CO;2-1 (1999).

215 Brierley, D. J. & Martin, S. A. Oxidative stress and the DNA mismatch repair pathway. *Antioxid Redox Signal* **18**, 2420-2428, doi:10.1089/ars.2012.4994 (2013).

216 Takeda, M. *et al.* The establishment of two paclitaxel-resistant prostate cancer cell lines and the mechanisms of paclitaxel resistance with two cell lines. *Prostate* **67**, 955-967, doi:10.1002/pros.20581 (2007).

217 Fox, E. *et al.* Pharmacokinetic and pharmacodynamic study of tariquidar (XR9576), a P-glycoprotein inhibitor, in combination with doxorubicin, vinorelbine, or docetaxel in children and adolescents with refractory solid tumors. *Cancer Chemother Pharmacol* **76**, 1273-1283, doi:10.1007/s00280-015-2845-1 (2015).

218 Kelly, R. J. *et al.* A pharmacodynamic study of docetaxel in combination with the P-glycoprotein antagonist tariquidar (XR9576) in patients with lung, ovarian, and cervical cancer. *Clin Cancer Res* **17**, 569-580, doi:10.1158/1078-0432.CCR-10-1725 (2011).

219 Abraham, J. *et al.* A phase I study of the P-glycoprotein antagonist tariquidar in combination with vinorelbine. *Clin Cancer Res* **15**, 3574-3582, doi:10.1158/1078-0432.CCR-08-0938 (2009).

220 Malingre, M. M. *et al.* Co-administration of GF120918 significantly increases the systemic exposure to oral paclitaxel in cancer patients. *Br J Cancer* **84**, 42-47, doi:10.1054/bjoc.2000.1543 (2001).

221 Planting, A. S. *et al.* A phase I and pharmacologic study of the MDR converter GF120918 in combination with doxorubicin in patients with advanced solid tumors. *Cancer Chemother Pharmacol* **55**, 91-99, doi:10.1007/s00280-004-0854-6 (2005).

222 Pusztai, L. *et al.* Phase II study of tariquidar, a selective P-glycoprotein inhibitor, in patients with chemotherapy-resistant, advanced breast carcinoma. *Cancer* **104**, 682-691, doi:10.1002/cncr.21227 (2005).

223 Gribar, J. J., Ramachandra, M., Hrycyna, C. A., Dey, S. & Ambudkar, S. V. Functional characterization of glycosylation-deficient human P-glycoprotein using a vaccinia virus expression system. *J Membr Biol* **173**, 203-214, doi:10.1007/s002320001020 (2000).

224 Richert, N. D., Aldwin, L., Nitecki, D., Gottesman, M. M. & Pastan, I. Stability and covalent modification of P-glycoprotein in multidrug-resistant KB cells. *Biochemistry* **27**, 7607-7613, doi:10.1021/bi00420a006 (1988).

225 Beheshti Zavareh, R. *et al.* Suppression of cancer progression by MGAT1 shRNA knockdown. *PLoS One* **7**, e43721, doi:10.1371/journal.pone.0043721 (2012).

226 Stanley, P. Golgi glycosylation. *Cold Spring Harb Perspect Biol* **3**, doi:10.1101/cshperspect.a005199 (2011).

227 Loo, T. W. & Clarke, D. M. Prolonged association of temperature-sensitive mutants of human P-glycoprotein with calnexin during biogenesis. *J Biol Chem* **269**, 28683-28689 (1994).

228 Wojtowicz, K., Januchowski, R., Nowicki, M. & Zabel, M. Inhibition of protein glycosylation reverses the MDR phenotype of cancer cell lines. *Biomed Pharmacother* **74**, 49-56, doi:10.1016/j.biopha.2015.07.001 (2015).

229 Seres, M., Cholujova, D., Bubencikova, T., Breier, A. & Sulova, Z. Tunicamycin depresses P-glycoprotein glycosylation without an effect on its membrane localization and drug efflux activity in L1210 cells. *Int J Mol Sci* **12**, 7772-7784, doi:10.3390/ijms12117772 (2011).

230 Johannes, L. & Billet, A. Glycosylation and raft endocytosis in cancer. *Cancer Metastasis Rev* **39**, 375-396, doi:10.1007/s10555-020-09880-z (2020).

231 Nabi, I. R., Shankar, J. & Dennis, J. W. The galectin lattice at a glance. *J Cell Sci* **128**, 2213-2219, doi:10.1242/jcs.151159 (2015).

232 Lee, Y. K., Lin, T. H., Chang, C. F. & Lo, Y. L. Galectin-3 silencing inhibits epirubicin-induced ATP binding cassette transporters and activates the mitochondrial apoptosis pathway via beta-catenin/GSK-3beta modulation in colorectal carcinoma. *PLoS One* **8**, e82478, doi:10.1371/journal.pone.0082478 (2013).

233 Huttlin, E. L. *et al.* The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **162**, 425-440, doi:10.1016/j.cell.2015.06.043 (2015).

234 Errington, T. M., Denis, A., Perfito, N., Iorns, E. & Nosek, B. A. Challenges for assessing replicability in preclinical cancer biology. *Elife* **10**, doi:10.7554/eLife.67995 (2021).

9 Image References

Figure
1 a: https://commons.wikimedia.org/wiki/File:Diagram_showing_the_position_of_the_prostate_and_rectum_CRUK_358.svg
accessed 2022-01-10 12 AM
Figure 1 b: https://www.earthslab.com/wp-content/uploads/2017/03/031717_1922_Prostate1.jpg
accessed 2022-01-10 12 AM
Figure 1 c: https://histology.medicine.umich.edu/sites/default/files/images/slides/24male.jpg
accessed 2022-01-13 9 AM
Figure
2 a (left): https://media.springernature.com/full/springer-static/image/art%3A10.1038%2Fmodpathol.2017.133/MediaObjects/41379_2018_Article_BFmodpathol2017133_Fig1_HTML.jpg?as=jpg
accessed 2022-01-17 10AM
Figure 2 a (right): https://onlinelibrary.wiley.com/doi/full/10.1111/pin.12398
accessed 2022-01-17 11AM ©: Weinzerl | Visual Media, 2015 Indiana University
Figure 2 b (upper left): https://www.webpathology.com/image.asp?case=20&n=27
accessed 2022-01-21 3PM
Figure 2 b (upper right): https://www.webpathology.com/image.asp?case=20&n=43
accessed 2022-01-21 3PM
Figure 2 b (lower left): https://www.webpathology.com/image.asp?n=48&Case=20
accessed 2022-01-21 3PM
Figure 2 b (lower right): https://www.webpathology.com/image.asp?n=28&Case=23
accessed 2022-01-21 3PM
Figure 6: https://cancer.sanger.ac.uk/signatures/sbs/sbs4/
Accessed 2022-11-01 8PM

All images on **Figure 2 b** were used with kind allowance of Dharam Ramnani, M.D.

10 Appendix

Appendix 1: Full names of material, consumable and equipment manufacturers and suppliers.

Abbreviation	Full name (parent company), head quarter location
Abcam	Abcam plc., Cambridge, UK
Acris	Acris (OriGene Technologies Inc.), Rockville, Maryland, USA
Advansta	Advansta Inc., San Jose, California, USA
Agilent	Agilent Technologies Inc., Santa Clara, California, USA
Ansell	Ansell Ltd., Yarra City, Australia
APExBIO	APExBIO Technology LLC, Houston, Texas, USA
AppliChem	AppliChem GmbH, Darmstadt, Germany
ATCC	American Type Culture Collection, Manassas, Virginia, USA
BD	Becton Dickinson Inc., Franklin Lakes, New Jersey, USA
Berkel	Berkel AHK GmbH & Co. KG, Ludwigshafen, Germany
Berthold	Berthold Technologies GmbH & Co. KG, Bad Wildbad, Germany
Biolegend	BioLegend Inc., San Diego, California, USA
Biozym	Biozym Scientific GmbH, Hessisch Oldendorf, Germany
Brand	Brand GmbH & Co. KG, Wertheim, Germany
Carl Roth	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
CLS	CLS Cell Lines Service GmbH, Eppelheim, Germany
CST	Cell Signaling Technology Inc., Danvers, Massachusetts, USA
Corning	Corning Inc., Corning, New York, USA
Dako	Dako (Agilent Technologies Inc.)
Eppendorf	Eppendorf SE, Hamburg, Germany
Fink & Walter	Fink & Walter GmbH, Merchweiler, Germany
Fresenius Kabi	Fresenius Kabi Deutschland GmbH, Bad Homburg, Germany
Greiner	Greiner Bio-One GmbH, Frickenhausen, Germany
Halyard	Halyard (Owens & Minor Inc.), Mechanicsville, Virginia, USA
Heirler	Heirler, Hügli Nahrungsmittel GmbH, Radolfzell, Germany
Heto-Holten	Heto-Holten A/S, Allerød, Denmark
Illumina	Illumina Inc., San Diego, California, USA
Integra	Integra Holding AG, Wallisellen, Switzerland
J&K Scientific	J&K Scientific LLC, Beijing, China
Kern & Sohn	Kern & Sohn GmbH, Balingen-Frommern, Germany
LGC	LGC Standards Ltd. (formerly LGC Promochem Ltd.), Teddington, UK
Lonza	Lonza Group AG, Basel, Switzerland
Luminex	Luminex Corporation, Austin, Texas, USA
Merck Millipore	Merck Millipore (Merck KGaA), Darmstadt, Germany
MJ Research	MJ Research Inc., St. Bruno, Canada
PeqLab	PeqLab (Avantor), Radnor, Pennsylvania, USA
Promega	Promega Corporation, Madison, Wisconsin, USA
PromoCell	PromoCell GmbH, Heidelberg, Germany

Qiagen	Qiagen N.V., Venlo, Netherlands

Continuation of Appendix 2: Full names of material, consumable and equipment manufacturers and suppliers.

Abbreviation	Full name (parent company), head quarter location
Riedel-de Haën	Riedel-de Haën (Honeywell Inc.), Charlotte, North Carolina, USA
Roche	F. Hoffmann-La Roche Ltd., Basel, Switzerland
Santa Cruz	Santa Cruz Biotechnology Inc., Dallas, Texas, USA
Sarstedt	Sarstedt AG & Co. KG, Nümbrecht, Germany
Sartorius	Sartorius AG, Goettingen, Germany
Selleck	Selleck Chemicals LCC, Houston, Texas, USA
Sigma	Sigma-Aldrich (Merck KGaA), Darmstadt, Germany
Systec	Systec GmbH, Linden, Germany
PeqLab	PeqLab (Avantor), Radnor, Pennsylvania, USA
Thermo Fisher	Thermo Fisher Scientific Inc., Waltham, Massachusetts, USA
TPP	TPP Techno Plastic Products AG, Trasadingen, Switzerland
VWR	VWR (Avantor), Radnor, Pennsylvania, USA
Zymo Research	Zymo Research Corporation, Irvine, California, USA

11 Data Origination Statement

With the exception of the data shown on **Figure 32 b** and **Figure 33**, all data shown in this doctoral thesis were generated experimentally by the author – Mr. Lukas Donix M.Sc. The data shown in the above-mentioned figures were published within the master thesis of Ms. Dinah Linke M.Sc. The experiments associated with the above-mentioned figures were conducted by Ms. Linke. Mr Donix was, at the time, involved in the conceptualization and planning of these experiments as well as in the general supervision of Ms. Linke. Key results of the experiments conducted by Ms. Linke are briefly summarized within this doctoral thesis. Furthermore, the data shown on **Figure 31**, **33** and **34** are part of a manuscript in revision (manuscript ID: ijms-2027029). The manuscript was jointly written by Mr. Donix and Ms. Linke. A short summary of the methods used by Ms. Linke is provided in section **4.13**. The detailed description of the methods exists within Ms. Linke's master thesis, which can be made available upon request. Ms. Linke is aware that the mentioned data and results are summarized herein and has expressed her explicit consent in that matter.
Furthermore, the data shown on the Figures **Figure 22** - **Figure 27** have been published within the paper:

“Donix, L., Erb, H.H.H., Peitzsch, C. et al. Acquired resistance to irradiation or docetaxel is not associated with cross-resistance to cisplatin in prostate cancer cell lines. J Cancer Res Clin Oncol 148, 1313–1324 (2022). https://doi.org/10.1007/s00432-022-03914-5 ”,

which was conceptualized and written by Mr Donix. As the experimental work for this paper constituted a significant part of the doctoral work, the results discussed in the paper were presented and revised in this dissertation. Finally, the whole exome sequencing and cell line characterization data (**Figure 7** - **Figure 21**) are unpublished work as of now (2022-11-01). A manuscript related to these results is currently work-in-progress.

At the time of publishing this thesis (16.05.2023), the above-mentioned joint manuscript written by Mr. Donix and Ms. Linke is now published as:

“Linke, D.; Donix, L.; Peitzsch, C.; Erb, H.H.H.; Dubrovska, A.; Pfeifer, M.; Thomas, C.; Fuessel, S.; Erdmann, K. Comprehensive Evaluation of Multiple Approaches Targeting ABCB1 to Resensitize Docetaxel-Resistant Prostate Cancer Cell Lines. Int. J. Mol. Sci. 2023, 24, 666. https://doi.org/10.3390/ijms24010666 “.

12 Scientific Integrity Statement

I hereby confirm that the presented thesis

“Genetic Diversity and Treatment Resistance in Prostate Cancer Cell Lines”

is the result of my own independent scholarly and scientific work. In all cases, material and data from the work of others (in books, research articles, dissertations, and on the internet) is acknowledged and clearly indicated. No material other than that listed has been used.

Dresden, 16.05.2022

Lukas Donix

13 Short Summary

This thesis is dealing with cancer of the prostate gland. In men, prostate cancer (PCa) has the second highest incidence among all cancers in the developed world and was responsible for 6.8% of the worldwide cancer-related mortality in 2020. PCa primarily affects elderly men. Locally confined PCa has good chances to be cured by prostatectomy or radiotherapy, whereas advanced and metastatic PCa is a terminal disease that can only be managed without curative intent. The treatment landscape for advanced and metastatic PCa is multi-layered and complex. As it is common in the treatment of cancers, initial response to a treatment is followed by the development of treatment resistance and relapse of the disease. Especially in the final stages of metastatic PCa, multiple treatment options are available. Which one to employ at this stage depends on the prior treatments, on the patient's fitness and in some cases on genetic features of the patient.

Several research interests (1, 2 & 3) were investigated within this thesis. For the study of PCa in the laboratory, cell lines play an important role. There are three major cell lines that have been established from metastatic PCa and that account for the bulk of the *in vitro* research in this cancer entity: PC-3, DU145 and LNCaP.

(1) An assumption that is necessary to compare studies done with cancer cell lines is that the same cell lines used in different laboratories are largely isogenic. One goal of this work was to challenge this assumption for the above-mentioned cell lines. To this end, whole exome sequencing of several PC-3, DU145, LNCaP and other PCa cell lines was performed and the putative genetic diversity between different strains (originating from different laboratories) of the same cell line was studied. While the majority of the detected genetic variants were shared among all cell line strains within a respective group, considerable numbers of unique genetic variants were detected as well. These unique variants were present in one, but none of the other cell lines within the respective groups. The sequenced PC-3, DU145 and LNCaP cell lines were not entirely isogenic. Furthermore, age-matched pairs of treatment-resistant and treatment-naïve cells were to some degree non-isogenic as well. However, the majority of the unique variants had a low allelic frequency and likely affected only a single chromosomal copy of the respective genes. The sequencing data also allowed for the inference of a phylogenetic tree for each cell line group. Of the above-mentioned cell lines, individual PC-3 cell lines that originated from different laboratories exhibited the largest variability with respect to their phenotype and gene expression.

(2) Another question that was addressed within this thesis was whether the previous treatment history influences the efficacy of the chemotherapeutic drug cisplatin. Cisplatin is a treatment option for a subset of late-stage metastatic PCa patients with neuroendocrine features and/or defects in homologous recombination repair. In addition to other standard PCa treatments, a majority of these late-stage patients will have received radiotherapy and/or docetaxel, before cisplatin would be considered. Clinical studies to evaluate the influence of the treatment history on cisplatin efficacy are hard to conduct as the number of patients eligible for cisplatin treatment is small. Furthermore, the variation of the treatment histories and the variation associated with patient genetics add to the problem of the limited sample

size. In this thesis, the effect of radiation and docetaxel treatment on cisplatin tolerance was studied in PCa cell lines. Radio-resistant and docetaxel-resistant cell lines as well as age-matched treatment-naïve control cells were treated with cisplatin. Their cisplatin sensitivity was assessed and compared by measurements of growth rates, viability, apoptosis, metabolic activity and colony formation ability. It was found that docetaxel-resistant cells exhibited the same tolerance for cisplatin as their parental counterparts. In addition, radio-resistant DU145 cells exhibited the same tolerance for cisplatin as their treatment-naïve parental cells. Radio-resistant PC-3 cells, however, were more sensitive to cisplatin than their control cell line. However, similar differences regarding the cisplatin tolerance were also observed among treatment-naïve PC-3 cells originating from different laboratories. Results from cell line experiments cannot directly be translated into clinical recommendations. However, these results suggest that the previous treatment history (radiation and/or docetaxel) does not or only minimally influence the effectiveness of cisplatin in PCa.

(3) Finally, the mechanism of docetaxel resistance was investigated in PC-3 and DU145 cell lines. The ATP-binding cassette transporter B1 (ABCB1) was confirmed as the primary effector of docetaxel resistance in docetaxel-resistant DU145 and PC-3 cell lines. Direct inhibition of ABCB1 with small molecule inhibitors as well as a siRNA-mediated knockdown effectively resensitized the two docetaxel-resistant cell lines. While the role of ABCB1 for taxane resistance is well known, the functional relevance of its N-linked glycosylations is to some degree disputed in the literature and seems to depend on the model system in which it was studied. This thesis provides a first investigation of the role of ABCB1 glycosylations in docetaxel-resistant DU145 and PC-3 cell lines. The N-glycosylation pathway was targeted by small molecule inhibitors. The inhibition of both initial and secondary steps in the N-glycosylation pathway failed to resensitize the docetaxel-resistant cell lines. This confirmed the notion from other model systems that the glycosylations of ABCB1 is not *per se* necessary for its functioning.

14 Acknowledgements

I would like to thank Rajanya and Asha for their support as well as all other members of my family and extended family.
I would like to thank Kati Erdmann for having me as her student. Kati truseted me and gave me freedom as well as the necessary structure. My time as her doctoral student was a good one and will be remembered as such.
I would like to thank Susanne Füssel for keeping everything together. I wish that in the coming years and the urological research laboratory will be placed into the most modern and spaceous facilities and that Susanne will not have to deal with any more Bauaufträge.
I would like to thank all the other members and past members of the urological research laboratory. In particular I would like to thank the following poeple who trained me and who – at times – helped me with my experiments: Holger Erb, Jörg Hofmann, Andrea Lohse-Fischer, Ulrike Lotzkat, Katrin Kirsche and Carla Steinhauser.
I would like to thank Dinah Linke for being a great person, student and scientist. Supervising her as a bachelors student was a great experience. As the reader will know, Dinah took over the ABCB1-project. Without her, I would not have had the time to focus on the WES data.
I would also like to thank Alicia Mühlbrand. She too is a great person, student and scientist. I have no doubt that she will receive an excellent medical degree.
I would like to thank all other students that have worked in the urological research laboratory for making the place lively and for the nice talks.
I would like to thank all the people from other laboratories who have collaborated with Kati and me and who have helped with my administrative supervision. In particular I would like to thank Prof. Dr. Anna Dubrovska and Prof. Dr. Lorenz Hofbauer for being part of my TAC committee. I would like to thank Doreen William for performing the primary analysis of the WES data and for answering all my questions. I would like to thank Claudia Peitzsch for helping out with the flow cytometry and with the radiation treatments. I would like to thank Susanne Michen for helping out with consumable emergencies.
At last, I would like to thank Prof. Dr. Christian Dahmann and Prof. Dr. Helge Taubert for agreeing to evaluate this doctoral work and Prof. Dr. Christian Dahmann and Prof. Dr. Stefanie Schirmeier for taking roles as examiners in the viva voce.