Liquid biopsy-based analysis of CTCs and ctDNA in tumor biology, biomarkers screening, and monitoring tumor progression in metastatic cancer patients.

Doctoral Thesis

With the aim of obtaining a doctoral degree (Dr. rer. nat.)

at the Department of Biology

Faculty of Sciences

University of Hamburg

Submitted by:

Maha Elazezy(M.Sc)

From Cairo, Egypt

Hamburg, 2022

The present study was carried out between April 2015 and January 2022 at the University Medical Center Hamburg-Eppendorf in the Institute of Tumour Biology under the supervision of Prof. Dr. med. Klaus Pantel and prof. Tobias Lenz.

Evaluation by

Prof. Dr. med. Klaus Pantel.

Director of Tumor Biology Institute

Center of Experimental Medicine

University Medical Center Hamburg-Eppendorf

Campus Forschung N27, 4. Etage, R 04.004

Martinistraße 52

20246 Hamburg

Prof. Dr. Tim Gilberger

Universität Hamburg

CSSB Centre for Structural Systems Biology c/o Deutsches Elektronen-Synchrotron DESY

Notkestraße 85, Building 15

22607 Hamburg

Date of oral defense: June 10th 2022

I. List of publications

Publication #1:

Elazezy, M., S. Schwentesius, L. Stegat, H. Wikman, S. Werner, W. Y. Mansour, A. V. Failla, S. Peine, V. Muller, J. P. Thiery, M. Ebrahimi Warkiani, K. Pantel, and S. A. Joosse. "Emerging Insights into Keratin 16 Expression During Metastatic Progression of Breast Cancer." Cancers (Basel) 13, no. 15 (Jul 31, 2021). https://dx.doi.org/10.3390/cancers13153869.

Publication #2:

Elazezy, M., K. Prieske, L. Kluwe, L. Oliveira-Ferrer, S. Peine, V. Muller, L. Woelber, B. Schmalfeldt, K. Pantel, and S. A. Joosse. "*BRCA1* Promoter Hypermethylation on Circulating Tumor DNA Correlates with Improved Survival of Patients with Ovarian Cancer." Mol Oncol (Oct 3, 2021). https://dx.doi.org/10.1002/1878-0261.13108.

Publication #3:

Retno Ningsi[†], **Maha Elazezy**[†], Luisa Stegat, Elena Laakmann, Sven Peine, Sabine Riethdorf, Volkmar Müller, Klaus Pantel, Simon A. Joosse. "Detection and characterization of estrogen receptor α expression of circulating tumor cells as a prognostic marker." Pathology and Oncology 2021.

Publication #4:

Maha Elazezy, Retno Ningsi, Alexander Sartori, Sabine Riethdorf, Harriet Wikman, Volkmar Müller, Klaus Pantel and Simon A. Joosse. "Detection of

endocrine therapy resistance mutations in metastatic breast cancer patients using a MassARRAY-Based liquid biopsy Assay." (**In processing**)

Publication #5:

Elazezy, M., and S. A. Joosse. "Techniques of Using Circulating Tumor DNA as a Liquid Biopsy Component in Cancer Management." Comput Struct Biotechnol J 16 (2018): 370-78. https://dx.doi.org/10.1016/j.csbj.2018.10.002.

II. Contents

III. List of abbreviations

ABD	Adapter binding domain
AF1	Activation functional domain1
AF1	Activation functional domain2
AF488	Alexa Fluor 488
AI	Aromatase inhibitors
AKT1	AKT Serine/Threonine Kinase 1
bp	Base pairs
BRCA1	BReast CAncer gene 1
BSA	Bovine Serum Albumin
°C	Degree Celsius
C2	Calcium-dependent phospholipid-binding domain
CDH1	E-Cadherin
CDK4	Cyclin Dependent Kinase 4
CDK6	Cyclin Dependent Kinase 6
cfDNA	Circulating free DNA
CNAs	Copy number aberrations
CO$_2$	Carbon Dioxide
Cq	Quantitation cycle
CTCs	Circulating tumor cells
ctDNA	Circulating tumor DNA
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA binding domain

DKK-1	Dickkopf WNT Signaling Pathway Inhibitor 1
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleoside triphosphates
DPBS	Dulbecco's PBS - Phosphate Buffered Saline
DTCs	Disseminated tumor cells
ΔΔCT	Relative gene expression
E2	17β-estradiol
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptors
EMT	Epithelial to mesenchymal transition
ER	Oestrogen receptor
ERBB2	Erb-B2 Receptor Tyrosine Kinase 2
ERK	Extracellular signal-regulated kinase
ERα	Estrogen receptor alpha
Erβ	Estrogen receptor beta
ESR1	Estrogen Receptor 1
FCS	Fetal calf serum
FGFR	Fibroblast growth factor receptor
FOXA1	Forkhead Box A1
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GATA3	GATA Binding Protein 3

gDNA	Genomic desoxyribonucleic acid
h	Hour/s
H20	Water
HER2	Erb-B2 Receptor Tyrosine Kinase 2
HGFR	Hepatocyte growth factor receptor
HR	Homologous recombination
HRD	Homologous recombination deficiency
HSC70	Heat shock cognate 71 kDa protein
IDC	Invasive ductal carcinoma
IF	Intermediate filament
IF	Immunofluorescence
ILC	Invasive lobular carcinoma
ITB	Institute of Tumour Biology
K16	Keratin 16
kDa	Kilodaltons
Ki-67	Ki-67 protein
LBD	Ligand binding domain
MAPK	Mitogen-Activated Protein Kinase
MET	Mesenchymal to epithelial transition
miRNA	MicroRNA
mM	Millimolar
mRNA	Messenger ribonucleic acid
MS	Methylation specific

mTOR	Mechanistic Target Of Rapamycin Kinase
ng	Nanogram
NGS	Next Generation Sequencing
NHEJ	Non-homologous end-joining
nM	Nanomolar
NOTCH1	Notch Receptor 1
PARP	Poly-ADP-Ribose-Polymerase
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PI3K/AKT	Phosphatidylinositol-3-Kinase and Protein Kinase B
	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic
PI3KCA	Subunit Alpha
PR	Progesterone receptor
PVDF	Polyvinylidene fluoride
qRT-PCR	Quantitative Real-time PCR
RBD	Ras binding domain
RIPA	Radioimmunoprecipitation assay buffer
RPMI	Rosewell Park Memorial Institute
RTKs	Tyrosine kinase receptors
RT-qPCR	Real-time quantitative polymerase chain reaction
SDS	Dimethyl sulfoxide
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SERD	Selective oestrogen receptor down-regulators
SERM	Selective oestrogen receptor modulators
SFRP-1	Secreted Frizzled Related Protein 1
siRNA	Small interfering RNA
SNAI1	Snail Family Transcriptional Repressor 1
SNAI2	Snail Family Transcriptional Repressor 2
TA	Transactivation domains
TF-AP2A	Transcription factor AP-2 alpha
TGF-β	Transforming Growth Factor Beta 1
TNBC	Triple-negative breast cancer
TP53	Tumor Protein P53
TWIST1	Twist Family BHLH Transcription Factor 1
UKE	University Medical Center Hamburg-Eppendorf
μm	Micrometre
μM	Micromolar
μM	Micromolar
VAF	Variant allele frequency
VEGFR	Vascular endothelial growth factor receptor
VIM	Vimentin
WNT-5	Wnt Family Member 5
ZEB1	Zinc Finger E-Box Binding Homeobox 1
ZEB2	Zinc Finger E-Box Binding Homeobox 2
Zn	Zinc finger

1. Abstract

Cancer patients' long-term survival largely depends on when the primary tumor and/or metastases are diagnosed. Liquid biopsy, including circulating tumor cells (CTCs) and circulating tumor DNA (ctDNA) taken from the blood of cancer patients can reveal critical details about the tumor status and progression. CTCs are considered the seeds of metastases, and their presence in the blood of patients predicts relapse-free and overall survival. Therefore, the phenotyping of cancer patients' CTCs may help translate the mechanisms that lead to tumor metastasis. In metastatic breast cancer patients, Keratin 16 (K16) expression was identified in 64.5% of detected CTCs, and it was associated with shorter relapse-free survival. K16 was found to be a metastasis-associated protein for breast cancer, promoting EMT and enhancing cell motility. Thus, assessing K16-CTCs status may provide predictive information that helps identify patients whose cancers are most likely to metastasize.

ctDNA, on the other hand, can be used to study the entire tumor genome as well as monitor drug response and resistance. *BRCA1* promoter methylation is a common epigenetic gene expression regulator in ovarian cancer. The conversion of methylation status is assumed to be the cause of disease recurrence. Liquid biopsy showed high potential and feasibility of monitoring the *BRCA1* methylation status in ovarian cancer patients. By developing an MS-PCR-based liquid biopsy assay, we could identify down to 0.03% of methylated DNA in a high background of normal DNA with 100% specificity. *BRCA1* promoter hypermethylation was detected in 60% of ovarian cancer patients, and

Abstract

we found that 24% of them lost their hypermethylation patterns during treatment. Multivariate survival analyses showed that the relapses are independent events, and the hypermethylation and methylation conversion are independently correlated to more prolonged relapse-free survival. Indeed, longitudinal monitoring of *BRCA1* methylation status in cfDNA may be a predictive marker.

In estrogen receptor (ER)-positive metastatic breast cancer, ER is considered a direct target of endocrine therapy. The investigation of ER-CTCs status may have a predictive value for endocrine therapy response. We found that the status of CTC was positively associated with progression-free survival. A higher number of CTCs during therapy was linked to disease progression, while a lower or stable number of CTCs was associated with a better outcome. However, only a third of metastatic breast cancer patients with ER-positive initially diagnosed had detectable ER-positive CTCs. The detection and monitoring of ER-CTCs status seem to be essential in the management of breast cancer patients, which could be tested as a potential source of endocrine therapy resistance.

ctDNA and CTCs are becoming crucial in clinical analysis for screening cancer by providing information about the genetic make-up of the total tumor burden present in the patient. In metastatic breast cancer, increasing *ESR1* and *PIK3CA* mutations result in resistance to endocrine therapy. The clinical significance of these genes with *FOXA1* and *GATA3* is unknown. Using MassARRAY-UltraSEEK® technology could detect and monitor the hotspot

Abstract

mutations of *ESR1*, *PI3KCA*, *FOXA1*, *GATA3*, *AKT1*, *ERBB2*, and *TP53* genes in cfDNA and CTCs. Longitudinal analysis of cfDNA and CTCs revealed a signature subclone with significant prognostic information in breast tumor guidelines. Patients with subclones of *FOXA1* (pE24K) and *GATA3* (pD336fs17) mutations, their tumors were more likely to progress. Both mutations were considerably raised upon chemotherapy alone or combined with endocrine therapy agents during tumor progression and at a progressive phase of a tumor. The combination of *GATA3* with *FOXA1*, *PIK3CA,* and/or *ESR1* was a strong predictor of tumor resistance and progression in patients with ER-positive metastatic breast cancer.

Overall, Liquid biopsy-based biomarkers like ctDNA and CTCs together could provide unique opportunities for real-time monitoring disease progression and give more detailed information on genetic variations, as well as predictive therapeutic information for clinical management and patient outcomes.

2. Zusammenfassung

Das langfristige Überleben von Krebspatienten hängt stark davon ab, wann der Primärtumor und/oder die Metastasen diagnostiziert werden. Flüssigbiopsien, zirkulierende Tumorzellen (CTCs) und zirkulierende Tumor-DNA (ctDNA) einschließend, die aus dem Blut von Krebspatienten entnommen werden, können kritische Details über den Tumorstatus und den Fortschritt aufdecken. CTCs gelten als die Saat von Metastasen, und ihre Anwesenheit im Blut von Patienten gibt Aussagen über einen möglichen Rückfall und das Gesamtüberleben. Daher kann die Phänotypisierung der CTCs von Krebspatienten helfen die Mechanismen zu übersetzen, die zu Tumormetastasen führen. Bei Patientinnen mit metastasiertem Brustkrebs wurde die Expression von Keratin 16 (K16) in 64,5% der nachgewiesenen CTCs identifiziert und war mit einem kürzeren rückfallfreien Überleben verbunden. Es wurde festgestellt, dass K16 ein Metastasen-assoziiertes Protein für Brustkrebs ist, das EMT fördert und die Zellmotilität erhöht. Daher kann die Beurteilung des K16-CTCs-Status prädiktive Informationen liefern, die dabei helfen, Patienten zu identifizieren, deren Krebs höchstwahrscheinlich metastasiert.

ctDNA hingegen kann verwendet werden, um das gesamte Tumorgenom zu untersuchen sowie das Ansprechen und die Resistenz von Medikamenten zu überwachen. Die Methylierung des *BRCA1*-Promotors ist ein häufiger epigenetischer Genexpressionsregulator bei Eierstockkrebs. Als Ursache für das Wiederauftreten der Erkrankung wird die Konversion des

Zusammenfassung

Methylierungsstatus angenommen. Die Flüssigbiopsie zeigte das hohe Potenzial und die mögliche Durchführung der Überwachung des *BRCA1*-Methylierungsstatus bei Patientinnen mit Eierstockkrebs. Durch die Entwicklung eines MS-PCR-basierten Flüssigbiopsie-Assays konnten wir bis zu 0,03 % methylierter DNA in einem hohen Hintergrund normaler DNA mit 100 % Spezifität identifizieren. Bei 60 % der Patientinnen mit Eierstockkrebs wurde eine BRCA1-Promotor-Hypermethylierung festgestellt, und wir fanden heraus, dass 24 % von ihnen ihre Hypermethylierungsmuster während der Behandlung verloren. Multivariate Überlebensanalysen zeigten, dass die Rückfälle unabhängige Ereignisse sind und die Hypermethylierung und Methylierungskonversion unabhängig mit einem verlängerten rückfallfreien Überleben korreliert sind. Tatsächlich kann eine Längsüberwachung des BRCA1-Methylierungsstatus in cfDNA ein prädiktiver Marker sein.

Bei Östrogenrezeptor (ER)-positivem metastasierendem Brustkrebs gilt ER als direktes Ziel der endokrinen Therapie. Die Untersuchung des ER-CTCs-Status kann einen prädiktiven Wert für das Ansprechen der endokrinen Therapie haben. Wir fanden, dass der CTC-positiver Status mit dem progressionsfreien Überleben assoziiert war. Eine höhere Anzahl von CTCs während der Therapie war mit dem Fortschritt der Krankheit verbunden, während eine niedrigere oder stabilere Anzahl von CTCs mit einem besseren Ergebnis für die Patientinnen assoziiert war. Allerdings hatte nur ein Drittel der Patientinnen mit metastasiertem Brustkrebs, welche zu Beginn als ER-positiv diagnostiziert wurden, nachweisbare ER-positive CTCs. Der Nachweis und die

Zusammenfassung

Überwachung des ER-CTC-Status scheinen bei der Behandlung von Brustkrebspatientinnen von wesentlicher Bedeutung zu sein, was als potenzielle Quelle für endokrine Therapieresistenz in ferner Zukunft werden könnte.

ctDNA und CTCs werden in der klinischen Analytik zum Screening von Krebs immer wichtiger, da sie Informationen über die genetische Ausstattung der gesamten im Patienten vorhandenen Tumorlast liefern. Bei metastasiertem Brustkrebs führen zunehmende *ESR1*- und *PIK3CA*-Mutationen zu einer Resistenz gegenüber einer endokrinen Therapie. Die klinische Bedeutung dieser Gene bei *FOXA1* und *GATA3* ist unbekannt. Mithilfe der MassARRAY-UltraSEEK®-Technologie konnten die Hotspot-Mutationen der *ESR1*, *PI3KCA*, *FOXA1*, *GATA3*, *AKT1*, *ERBB2* und *TP53* Gene in cfDNA und in CTCs nachgewiesen und überwacht werden. Die Längsanalyse von cfDNA und CTCs ergab einen charakteristischen Subklon mit signifikanten prognostischen Informationen in den Leitlinien für Brusttumore. Bei Patienten mit Subklonen von *FOXA1* (pE24K) und *GATA3* (pD336fs17) Mutationen war die Wahrscheinlichkeit einer Tumorprogression höher. Beide Mutationen waren bei alleiniger Chemotherapie oder in Kombination mit endokrinen Therapiemitteln während der Tumorprogression und in einer progressiven Phase eines Tumors erheblich erhöht. Die Kombination von *GATA3* mit *FOXA1*, *PIK3CA* und/oder *ESR1* war ein starker Prädiktor für die Tumorresistenz und den Fortschritt bei Patientinnen mit ER-positivem und metastasiertem Brustkrebs.

Zusammenfassung

Insgesamt könnten Flüssigbiopsie-basierte Biomarker wie ctDNA und CTCs zusammen einzigartige Möglichkeiten für die Echtzeitüberwachung des Krankheitsverlaufs bieten und detailliertere Informationen zu genetischen Variationen sowie prädiktive therapeutische Informationen für das klinische Management und die Patientenergebnisse liefern.

3. Introduction

3.1. Cancer

Cancer is a disease of genes triggered by a build-up of errors and mistakes in a cell's DNA of tumor suppressors and proto-oncogenes genes. Proto-oncogenes promote biological process and proliferation, whereas tumor suppressors induce caspase-mediated cell death and are negative regulators of cell proliferation[1-4]. Genetic and epigenetic aberrations aim to activate proto-oncogenes and inactivate tumor suppressor genes [5]. Genetic alterations encompass chromosomal aberrations such as translocations, insertions, deletions, and copy number aberrations (CNAs), as well as single nucleotide point mutations [6]. Epigenetic changes include aberrant methylation and histone modification [4, 7-9]. Both genetic and epigenetic alterations play a key role in gene activity, cell differentiation, tumorigenesis, and other cellular regulatory processes.

The major cause of cancer-related death is metastatic relapse, resulting from tumor cell colonization from the primary tumor into distant organs accompanied by organ failure [10, 11]. The dissemination route takes place principally through the circulation of the blood, where only a few circulating tumor cells (CTCs) can survive from several natural obstacles such as the immune system, anoikis, shear forces, and oxidative stress [10]. Tumor cell extravasation is commonly assumed to occur in distant organs, such as the brain, bone marrow, lungs, or liver, where disseminated tumor cells (DTCs) may remain dormant for several years before ultimately growing into an overt

metastasis [12]. Detection of DTCs in bone marrow was strongly associated with disease recurrence [13].

Epithelial to mesenchymal transition (EMT) and it is reverse process mesenchymal to epithelial transition (MET) are considered to be essential for metastatic process and formation of secondary tumor [14]. EMT seems critical for the initial escape by enabling individual tumor cells to migrate and invade [15]. Through participating in dynamic cellular and molecular changes in tumor cells such as loss of epithelial cell polarity, downregulation of junctional complexes (e.g., E-Cadherin (CDH1)) [16], upregulation of mesenchymal markers (e.g., Vimentin (VIM)) [16, 17], and reorganization of the actin cytoskeleton to produce the migratory phenotype required for cellular migration. EMT can be triggered by a wide range of intrinsic signaling molecules, such as tyrosine kinase receptors (RTKs) and TGF-β receptors, which are contributed to the activation of intracellular signaling pathways that induces EMT by upregulation of selected zinc finger (e.g., SNAI1, SNAI2, ZEB1, ZEB2) or TWIST1 and NOTCH1 transcription factors [18, 19]. MET contributes in the formation of secondary tumors after extravasation of tumor cells at a distant organ [15]. The tumor cell starts to undergo MET to sustain with epithelial-like-phenotype and conserve its polarity. The epithelial like-phenotype allows tumor cells able to proliferate and form macrometastases of a secondary tumor. Indeed the epithelial cell plasticity enables the tumor cells to undergo a dynamic and reversible transition between the epithelial and mesenchymal like-phenotype [20]. It has been shown that tumor cells undergoing a partial

transformation present a greater risk of metastasis than those undergoing either a mesenchymal or epithelial phenotype.

Tumor heterogeneity is one of the key issues limiting the effectiveness of cancer target medicines and reducing treatment outcomes [21]. Within a primary tumor and its metastases, tumor heterogeneity refers to subpopulations of cells with diverse genotypes and phenotypes that may have different biological behaviors (e.g., Intra-tumour heterogeneity) [21, 22]. Since tumor heterogeneity promotes resistance, a precise assessment of tumor heterogeneity is critical for developing successful medicines. Multiregion sequencing [23], single-cell sequencing [21], and longitudinal analysis of liquid biopsy samples [24] have shown considerable ability to analyze complicated clonal structures of cancer.

3.2. Liquid biopsy

To monitor the molecular characterization of the tumor in real-time and identify possible therapeutic targets, material taken directly from the tumor should be screened [24]. Liquid biopsy is an emerging field dealing with detecting tumor states and progression from body fluids of cancer patients [25]. Liquid biopsy based on minimally invasive blood tests through studying blood-borne biomarkers. These biomarkers include circulating tumor cells (CTCs) present in the mononuclear cell fraction and cell-free DNA (cfDNA), microRNA (miRNA), exosomes, and platelets derived from the plasma fraction [26]. This

method is used to detect biomarkers in blood or other body liquids for prognostic and predictive purposes and has several advances than using tissue only [27].

CTCs detections were found to be predictive and pronounced in various early-stage of cancer entities [28]. CTCs have a limited half-life (between 1 and 2.5 hours) in blood circulation [29]. However, the low number of tumor cells in the incredibly high background of normal cells requires highly sensitive techniques [25]. Different approaches for obtaining CTCs are based on either specific cellular markers expressed on the cell surface or depending on the cells' physical characteristics [30]. Antigens expressed by tumor cells enable positive enrichment, whereas white blood cell depletion will achieve negative enrichment [31]. Many commercially available instruments and test systems are approved as clinical diagnostic devices, which have allowed CTC to be identified, enumerated, and analyzed [32-35]. The most famous system for the enumeration and isolation of CTCs is the gold standard, the FDA-cleared CellSearch® system [33, 36]. This system detects CTCs based on binding to anti-EpCAM, cytokeratin (CK), and CD45 expression. CTCs count carries independent prognostic information in metastatic breast cancer patients. Consequently, the phenotyping of the CTCs can provide crucial information on the evolving characteristics of the tumor during progression and treatment resistance.

Circulating tumor DNA (ctDNA) are cell-free DNA fragments released into the blood by tumor cells through cell death either by apoptosis or necrosis. The rate of ctDNA release into the blood circulation depends on the location, size,

and vascularity of the tumor, causing a variation in ctDNA levels among patients [37]. In blood circulation, a half-life of cfDNA is between 16 minutes and 2.5 hours [38]. The cfDNA concentration in healthy individuals is an average of 30 ng/ml of plasma ranging from 0 to 100 ng/ml, whereas it can go up to 1,000 ng/ml for cancer patients [39, 40]. ctDNA accounts for 0.01% of the total cfDNA. This is extremely low concentrations making the downstream analysis is a challenge, particularly in the early stages of tumor development. ctDNA provides direct information about the genetic and epigenetic variations in the tumor, drug response, and resistance to therapy [37, 41].

In recent years, there has been a remarkable development in ctDNA detection and analysis technologies, such as NGS-based methods, that have made considerable progress in overcoming many of the challenges to reduce the error rate and increase the sensitivity of ctDNA detection [37]. Nonetheless, NGS-based methods are also relatively costly and time-consuming [42]. On the other hand, mass-spectrometry methods are promising tools for ctDNA screening due to their low cost, time, and DNA input requirements, as well as their high sensitivity and specificity [43, 44]. Furthermore, for a limited number of biomarkers, analysis using Real-Time PCR-based techniques is cost-effective, fast, and practical in routine clinical practice [37, 45, 46]. Eventually, further standardization of these methods would make ctDNA a valuable substrate in cancer diagnostics.

3.3. Breast cancer

Breast cancer is the most common cancer among women worldwide. It accounts for 11.7% of all new cancer cases in 2020 and about 25% of all cancer cases among women [47]. Incident cases of breast cancer are expected to increase by more than 46% by 2040, according to the GLOBOCAN Cancer Tomorrow prediction tool [47].

Breast cancer usually starts off in the inner lining of milk ducts, the lobules, or the tissue in between. Breast cancer is a heterogeneous disease comprising multiple entities associated with distinctive histological and molecular subtypes identified based on their hormone status and/or gene expression patterns [48]. The most common histopathological type of breast cancer is invasive ductal carcinoma (IDC), with a prevalence of approximately 80% of all breast cancers [49]. It is followed by invasive lobular carcinoma (ILC) with 10% of all invasive breast cancers [50].

On the molecular level, breast cancer is classified into five different clinical intrinsic subtypes [51]. The most prominent subtype is the Luminal A-like subtype presenting up to 60% of all breast cancer cases [52]. Luminal A-like subtype tumors are characterized with estrogen receptor (ER), progesterone receptor (PR) positive, negative for the human epidermal growth receptor 2 (ERBB2), and has a low expression level of the Ki-67 protein [53]. These tumors are dependent on hormones for growth and proliferation. Luminal A tumors characterize less aggressive and better prognosis. Luminal B-like subtype represents 30% of all breast tumors and is characterized by ER and /or

PR positive, and either ERBB2 positive or negative with high levels of Ki-67. Luminal B tumors generally grow slightly faster than luminal A tumors, and their prognosis is slightly worse [54]. The ERBB2-enriched subtype is hormone receptor-negative (ER and PR negative) and ERBB2 positive. ERBB2 -enriched cancers tend to grow faster than luminal cancers and can have a worse prognosis, but they are often successfully treated with targeted therapies for the ERBB2 protein [55]. Triple-negative/basal-like breast cancer subtype (TNBC) is hormone receptor-negative (ER and PR negative) and ERBB2 negative. This type of cancer is more common in women with *BRCA1* and *TP53* genes mutations. Patients with basal-like tumors have a worse prognosis than patients with luminal tumors [55]. Finally, the Normal-like subtype is similar to luminal A disease, hormone receptor-positive (ER and/or PR positive), ERBB2 negative, and a low expression level of Ki-67 protein. Its prognosis is slightly worse than luminal A tumors [55].

3.3.1. The functional role of estrogen receptor in breast cancer

The estrogen receptor (ER) signaling plays an important role in the growth of both normal and neoplastic breast tissue [56]. The ER is a nuclear transcription factor and member of the steroid-thyroid-retinoid receptor superfamily (nuclear receptor superfamily), located at chromosome 6 [57]. The ER is comprised of two subdivisions ERα and Erβ [57]. ER is involved in regulating several physiological functions, including cell cycle progression and proliferation [58]. It is activated by binding to its ligands (17β-estradiol (E2)),

leading to induction and regulation of the development of secondary female sex characteristics after puberty and during pregnancy, as well as regulation of the menstrual cycle, and forming breast tissue and its further development [59, 60].

Since ER is a major driver of breast cancer, a large number of therapeutic strategies were established to inhibit hormone synthesis through selective ER modulators (SERM), selective ER down-regulators (SERD), or aromatase inhibitors (AI) in combination with either mTOR inhibitor or CDK4/CDK6 inhibitors to disrupt of the ER signaling pathway in cancer cells [61-63]. Endocrine therapy is commonly used in women with ER-positive breast cancer as adjuvant therapy. However, endocrine therapy failure is noted in 15-20% of women whose tumors are intrinsically resistant to treatment, and 30-40% acquire resistance to treatment over many years [64].

Various mechanisms may cause resistance to endocrine therapy, resulting in either a deficiency of ER protein expression or ER pathway dysfunction. The resistance to the therapy causes tumor progression and metastasis, which is the cause of cancer-related deaths [63]. One of the resistance mechanisms to endocrine therapy is the overexpression of ERBB2 in hormone receptor-positive cells, leading to downregulation of ER expression [63]. Also, upregulation of Receptor Tyrosine Kinases such as epidermal growth factor receptors (EGFR), fibroblast growth factor receptor (FGFR), hepatocyte growth factor receptor (HGFR), and vascular endothelial growth factor receptor (VEGFR) led to the activation of MAPK/ERK or PI3K/AKT signaling pathway that involved in endocrine resistance development [65].

Introduction

Mutations of the ER gene play an essential role in the effectiveness of anti-breast cancer drugs [48]. The majority of mutations are in the ligand-binding domain (LBD) region, leading to constitutive activation of the ER [66]. Multiple studies of next-generation sequencing and liquid biopsy in clinical trial cohorts have shown interest in the high prevalence of ERα mutations in ER-positive metastatic breast cancer patients who received prior AI treatment [67-69]. The most prevalent mutations presented in breast cancer patients are the D538G, Y537S, Y537N, Y537C, and E380Q [67, 70] (Figure 1). The ERα mutations were found to be uncommon in primary tumors but appear to be relatively common in endocrine resistance progression [66]. Thus, these mutations might be used as a predictive marker for endocrine treatment resistance in advanced breast cancer patients [63].
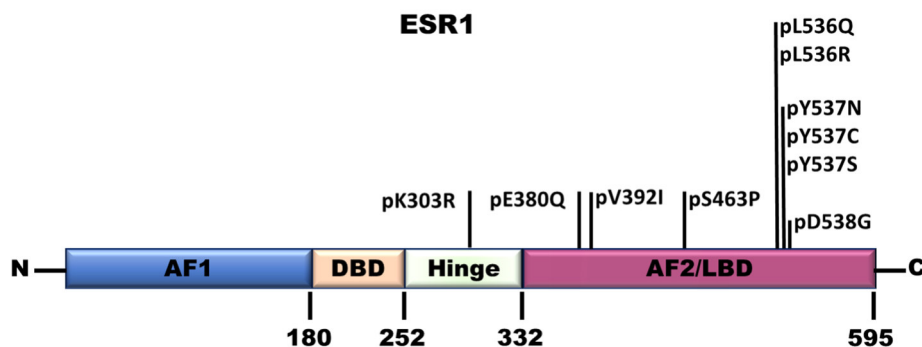

Figure 1. Protein domain structure of ESR1 with hotspot mutations. The ESR1 is divided into four functionally separate domains: an amino-terminal

domain, harboring the N-terminal ligand-independent activation functional domain (AF1), a DNA binding domain (DBD), and a flexible hinge region, connecting the DBD domain with the carboxy-terminal ligand-binding domain (LBD), including the second transcriptional activation domain AF2.

In addition to *ESR1* mutations, *PIK3CA* mutations frequently occur in 30% of breast cancer patients [71]. PI3K (PIK3CA) is commonly activated in breast cancer [72]. The activation of the PI3K/AKT pathway contributes to chemoresistance in breast cancer. Hot spot mutations H1047R, E542K, E545K, N345K, and H1047L are account for 73% of all *PIK3CA* mutations (Figure 2) [73]. Recent clinical studies have proposed the importance of *PIK3CA* mutations as a predictive marker for responses to PI3K inhibitors [74].
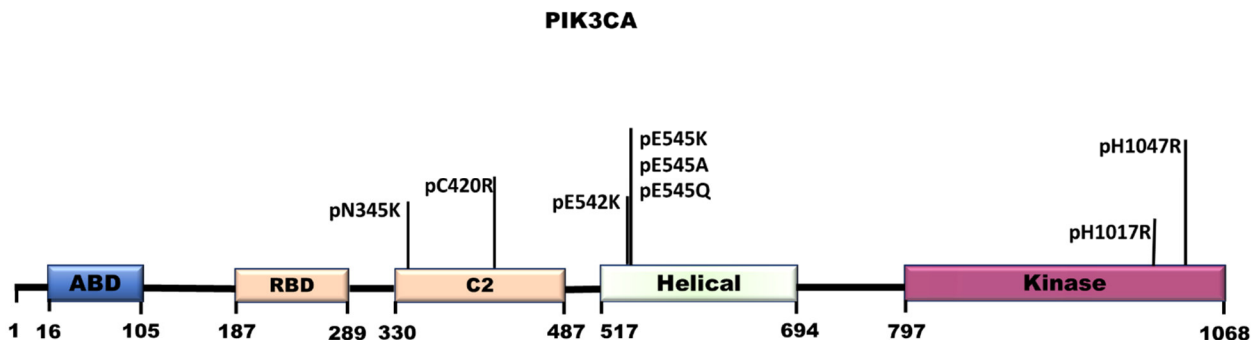

Figure 2. Protein domain structure of PIK3CA with hotspot mutations. PIK3CA encompasses regulatory subunit binding domain; ABD (adapter binding domain); RBD (Ras-binding domain); C2 (calcium-dependent

phospholipid-binding domain); Helical (PI3K helical domain); Kinase (PI3/4-kinase domain).

Restoring the function of ER by reprogramming the ER-dependent transcriptome is one of the promoting endocrine-resistant cell growth. *FOXA1* and *GATA3* are transcription factors required for ER binding and growth [75]. *FOXA1* binds to achromatized DNA and opens the chromatin, enhancing ERα binding to its target genes [76]. *GATA3* is involved in the differentiation of luminal epithelial cells and the subsequent development of differentiated epithelial cells' ductal tree [77]. *FOXA1* and *GATA3* are associated mainly with the luminal transcriptional program. *FOXA1* expression was observed in 42% of invasive carcinomas, while *GATA3* expression was found in 48% [78]. *ESR1* binding was mediated by both *GATA3* and *FOXA1* to the cis-regulatory elements that drive the transcription of *ESR1* target genes [75]. In breast cancer, *FOXA1* is found to be mutated in 4.18% and *GATA3* in 15% of breast cancer patients [79, 80] (Figure 3).
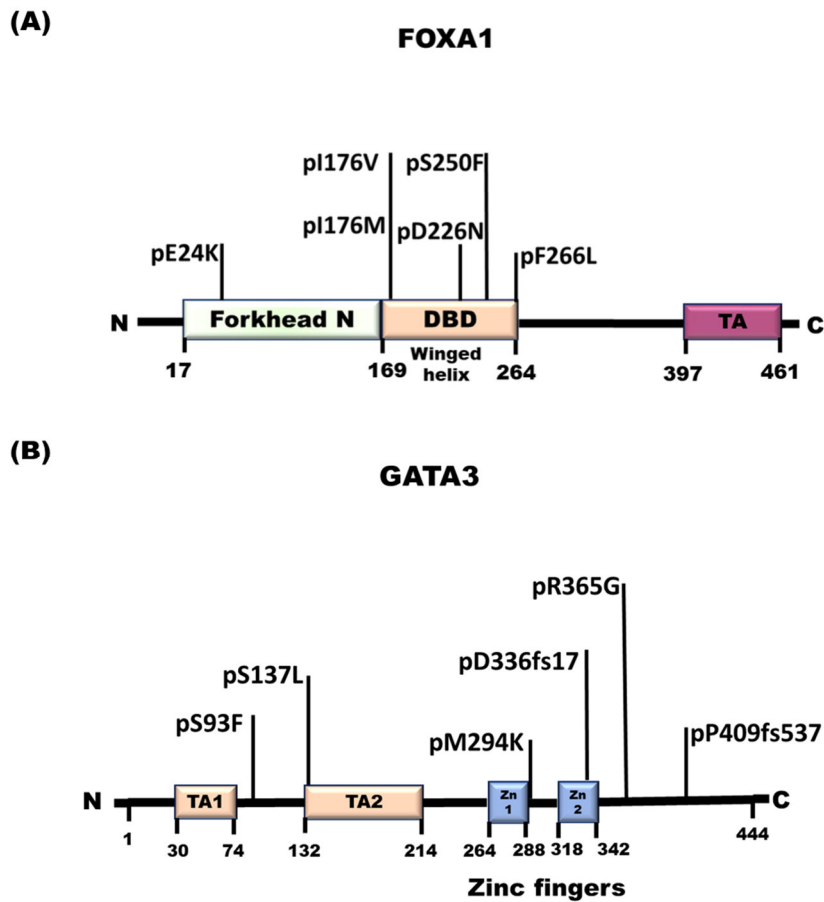
Figure 3. Protein domain structure of FOXA1 and GATA3 with hotspot mutations. (**A**) FOXA1 functional domains: Forkhead domain in N-terminal region; Winged helix DBD (winged helix–turn–helix DNA-binding domain); TA (transactivation domains). (**B**) GATA3 functional domains: TA1 and TA2 (two transactivation domains); Zn1 and Zn2 (two zinc fingers).

3.4. Keratins as diagnostic and prognostic markers in breast cancer

Keratins are intermediate filament (IF) proteins that have been used in cancer diagnosis and prognosis [81]. These proteins are found mainly in epithelial cells, where they disperse across the cytoplasm to maintain cell structure and rigidity [82] as well as regulate intracellular signaling pathways [83]. Keratins are classified into type I (acidic) and type II (basic). Currently, there are 40 different keratins described in human epithelial cells, including 20 (type I) and 20 (type II) keratins [84]. Type I epithelial keratins comprise K9–K28, while type II encompasses K1–K8 and K71–K80 [81, 84].

Keratin's expression patterns are specifically related to the epithelial type and stage of the cell differentiation of certain tissues [84]. Therefore, keratins are commonly used as an indicator of immunochemistry in diagnostic tumor pathology to identify tumor cells according to the original tissues, such as increased regulation of K7, K8, K18, and K19 in most breast adenomas in the breast, while K5/ 6, K14, and K17 are expressed in a basal subtype [85].

Keratin 16 (K16) is one of the IFs located at chr.17q21.2, which encodes for the type I cytoskeletal 16 protein [83]. Previous studies showed that K16 expression influences the keratinocyte organization, contributing to the changes in the morphology of epithelial cells and directly impacting the adhesion, differentiation, and migration of cells during wound healing [86-88].

Although differences in keratin expression patterns between metastatic and non-metastatic tumors have been published several times, little is known

about the potential role of keratins in metastatic development. *In silico* analyses have previously shown a correlation between the gene expression level of K16 and the period of metastasis-free survival of metastatic breast cancer patients. The analysis revealed that patients with high K16 expression levels in their primary tumor have shorter relapse-free survival compared to patients with a tumor expressing less K16 [85]. Therefore, it may be speculated that K16 is upregulated in tumor cells with high metastatic potential and that K16 might be associated with metastatic progression leading to a more aggressive course of breast cancer and shorter relapse-free survival. Thus, K16 could be a promising new prognostic marker for the metastatic capacity of poor-prognosis breast tumors. Further investigations of the cellular and differential background that promotes K16 regulation may help to better understand tumor malignancy.

3.5. Ovarian cancer

Ovarian cancer is one of the most prevalent female genital cancers in Germany, with one in approximately 71 women developing ovarian cancer over the course of her life [89]. Furthermore, it accounts for 3.2% of all malignant neoplasms in women and 5.3% of all female cancer-related deaths [89]. Ovarian cancer is typically diagnosed at an advanced stage (stage III-IV) with disseminated intra-abdominal metastasis that is associated with a long-term survival rate of only 20% [90]. The primary treatment of newly diagnosed ovarian cancer is cytoreductive surgery to achieve complete macroscopic

resection [91]. In addition, the standard management consists of (neo)adjuvant systemic treatment with carboplatin in combination with paclitaxel [92, 93].

Although most ovarian cancer patients respond well to the initial chemotherapy, metastatic recurrence of the disease occurs in more than half of the cases within approximately two years and 70% within five years [91, 94]. Patients with defects in homologous recombination-directed DNA repair mechanisms are expected to be a benefit for most of the treatment with platinum-based therapy or PARP (Poly-ADP-Ribose-Polymerase) inhibitors in first-line therapy or in the case of a platinum-sensitive recurrence [95, 96]. However, the development of therapy resistance after disease recurrence is a major clinical challenge.

Double-strand DNA breaks are repaired via homologous recombination (HR) and non-homologous end-joining (NHEJ) [97]. Unrepaired DNA damage can result in an accumulation of mutations and unregulated cell division. Therefore, homologous recombination deficiency (HRD) is related to cancer susceptibility and progression [98]. Germline or somatic mutations in HR genes have been identified in approximately one-third of ovarian carcinomas, and their presence is highly predictive of primary platinum and PARP-inhibitors sensitivity and favorable progression-free and overall survival [95, 99-101].

BRCA1 (BReast CAncer gene 1) is the most frequently implicated gene in ovarian cancer. *BRCA1* is essential for DNA repair, cell cycle checkpoint modulation, mitosis, chromatin remodeling, and transcriptional regulation [98, 102]. Epigenetics, such as promoter hypermethylation is a dynamic mechanism

that plays an essential role in tumor evolution and in developing therapy resistance. In breast cancer, the hypermethylation of the *BRCA1* promoter leads to downregulation of *BRCA1* mRNA expression, resulting in defective homologous recombination characterized by the typical chromosomal aberrations seen in *BRCA1* mutation carriers [103, 104].

In ovarian cancer tissue, The *BRCA1* promoter was often found to be hypermethylated. However, the methylation status was unstable and lost in recurrent disease, indicating a potential resistance mechanism through the development of cancer-induced by the treatment [105]. Another possibility is that tumors comprise many (epi)genetic clones with hypermethylated and unmethylated *BRCA1* promoters [106, 107]. Due to platinum-based therapy and PARP inhibitors, tumor cells with dysfunctional *BRCA1* will be killed, whereas slow-growing tumor cells with functional *BRCA1* will eventually overcome these therapies [108]. Whether the change in methylation status occurs through selection or evolution, detecting and monitoring *BRCA1* promoter hypermethylation using liquid biopsy may significantly influence the clinical management of ovarian cancer patients who lack *BRCA1* mutation [109].

4. Aim of the work

The main goal of the current study is to use liquid biopsy-derived materials (e.g., CTCs, cfDNA) to investigate biomarkers that may help in the early detection of micro-metastasis and tumor screening by providing information about the genetic and epigenetic variations in cancer patients.

From the current work, five publications were conducted to address the above-indicated aim(s).

Publication #1:

This study set out to investigate the biological role of K16 in metastatic breast cancer cell lines and evaluate the clinical relevance of K16 in metastatic breast cancer patients by analyzing the K16 expression in CTC, i.e., the metastatic seeds.

Publication #2:

The purpose of this study was to develop a liquid biopsy assay that could determine the methylation status of the *BRCA1* promoter to monitor hypermethylation of the *BRCA1* promoter and investigate its clinical significance as a predictive biomarker in ovarian cancer patients.

Publication #3:

The main aim of this study was to evaluate the ER-CTC status employing ERα monoclonal murine ER-119.3 antibody used by Paoletti *et al*. in patients with

Introduction

ER-positive metastatic breast cancer using the CellSearch System for the quantification of CTCs.

Publication #4:

This study aimed to screen the major hotspot mutations in *ESR1*, *PI3KCA*, *AKT1*, *ERBB2*, *TP53*, *FOXA1*, and *GATA3* occurrence using UltraSEEK® Breast panel in one hundred one patients with ER-positive metastatic breast cancer during the course of treatment by MassARRAY® System and assessing the clinical value of identified mutations in respect of tumor progression and overall survival.

<u>Publication #5:</u>

This research aimed to provide a comprehensive overview of several approaches used to extract and characterize ctDNA. In addition, it highlighted the challenges that still need to be overcome to implement ctDNA-based liquid biopsy for precision medicine.

5. Publications

5.1. Emerging insights into keratin 16 expression during metastatic progression of breast cancer

The mechanisms leading to tumor metastasis remain poorly understood, and therefore phenotyping of CTCs from cancer patients may contribute to translating these mechanisms. We have previously shown in silico analysis that K16 mRNA expression upregulation might be associated with higher tumor aggressiveness. In the presented study, we found that K16 is a metastasis-associated protein that promotes EMT and acts as a positive regulator of cellular motility by reorganizing the actin cytoskeleton, which is the driving force behind disrupting intercellular adhesion and directional migration. In metastatic breast cancer patients, 64.5% of the detected CTCs expressed K16, which was associated with shorter relapse-free survival (P=0.0024). This study, to our knowledge, is the first report indicating that K16 might be a metastasis-promoting gene in breast cancer.

cancers

MDPI

Article

Emerging Insights into Keratin 16 Expression during Metastatic Progression of Breast Cancer

Maha Elazezy [1], Sandra Schwentesius [1], Luisa Stegat [1], Harriet Wikman [1], Stefan Werner [1], Wael Y. Mansour [2], Antonio Virgilio Failla [3], Sven Peine [4], Volkmar Müller [5], Jean Paul Thiery [6], Majid Ebrahimi Warkiani [7], Klaus Pantel [1] and Simon A. Joosse [1,*]

1   Department of Tumor Biology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; m.elazezy@uke.de (M.E.); s.schwentesius@uke.de (S.S.); luisa@leifheit-stegat.de (L.S.); h.wikman@uke.de (H.W.); st.werner@uke.de (S.W.); pantel@uke.de (K.P.)
2   Department of Radiotherapy and Radiation Oncology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; wmansour@uke.de
3   UKE Microscopy Imaging Facility (UMIF), University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; a.failla@uke.de
4   Department of Transfusion Medicine, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; s.peine@uke.de
5   Department of Gynecology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany; v.mueller@uke.de
6   Bioland Laboratory, Guangzhou Regenerative Medicine and Health Guangdong Laboratory, Guangzhou 510320, China; tjp@nus.edu.sg
7   School of Biomedical Engineering, University of Technology Sydney, Sydney 2007, Australia; majid.warkiani@uts.edu.au
*   Correspondence: s.joosse@uke.de; Tel.: +49-(0)-40-7410-51970

Simple Summary: The mechanisms leading to tumor metastasis remain poorly understood, and therefore, phenotyping of circulating tumor cells from cancer patients may contribute to translating these mechanisms. In in silico analysis, high expression of keratin 16 was associated with higher tumor aggressiveness. According to our results, keratin 16 is a metastasis-associated protein that promotes EMT and acts as a positive regulator of cellular motility by reorganizing the actin cytoskeleton, which is the driving force behind disrupting intercellular adhesion and directional migration. In metastatic breast cancer patients, circulating tumor cells expressing keratin 16 were associated with shorter relapse-free survival. This is an important issue for future research to determine the exact function of keratin 16 in tumor dissemination and metastasis development by analyzing keratin 16 status in disseminating tumor cells. Furthermore, gaining a better knowledge of keratin 16's biology would give crucial mechanistic insights that might lead to a unique treatment option.

Abstract: Keratins are the main identification markers of circulating tumor cells (CTCs); however, whether their deregulation is associated with the metastatic process is largely unknown. Previously we have shown by in silico analysis that keratin 16 (*KRT16*) mRNA upregulation might be associated with more aggressive cancer. Therefore, in this study, we investigated the biological role and the clinical relevance of K16 in metastatic breast cancer. By performing RT-qPCR, western blot, and immunocytochemistry, we investigated the expression patterns of K16 in metastatic breast cancer cell lines and evaluated the clinical relevance of K16 expression in CTCs of 20 metastatic breast cancer patients. High K16 protein expression was associated with an intermediate mesenchymal phenotype. Functional studies showed that K16 has a regulatory effect on EMT and overexpression of K16 significantly enhanced cell motility ($p < 0.001$). In metastatic breast cancer patients, 64.7% of the detected CTCs expressed K16, which was associated with shorter relapse-free survival ($p = 0.0042$). Our findings imply that K16 is a metastasis-associated protein that promotes EMT and acts as a positive regulator of cellular motility. Furthermore, determining K16 status in CTCs provides prognostic information that helps to identify patients whose tumors are more prone to metastasize.

1. Introduction

Breast cancer is a heterogeneous disease encompassing different molecular subtypes that are identified based on their hormone status and/or gene expression patterns [1]. Long-term survival of breast cancer patients largely depends on when the primary tumor and especially the metastases are detected [2]. Epithelial to mesenchymal transition (EMT) is thought to play an essential role in initiating cancer dissemination and metastasis. During this process, intercellular adhesive complexes, such as E-cadherin-based adherens junctions, are downregulated, leading to a mesenchymal-like phenotype [3,4]. The reverse process, i.e., mesenchymal to epithelial transition (MET), plays a critical role in metastatic tumor formation [5]. High plasticity of carcinoma cells enables them to undergo a dynamic and reversible transition between the epithelial and mesenchymal-like phenotype [6].

An important aspect of EMT is the reorganization of the cytoskeleton, including changes in intermediate filaments, which may contribute to the induction of cell motility. Keratins are intermediate filament proteins routinely used for cancer diagnostics [3,7]. Keratins are mainly present in epithelial cells anchored to desmosomes, hemidesmosomes, and the nuclear membrane. Keratins contribute to the control of cell shape and rigidity [8], as well as regulating intracellular signaling pathways [9]. The keratin 16 (*KRT16*) gene is located at chromosome 17q21.2, encoding the type I cytoskeletal 16 protein K16 [9]. Previous studies have shown that K16 expression influences keratinocyte organization, which contributes to the changes in the morphology of epithelial cells and directly impacts adhesion, differentiation, and migration of cells during wound-healing [10–12].

Little is known about the deregulation of K16 in cancer and metastasis. Through in silico analysis, a positive correlation between *KRT16* gene expression and shorter relapse-free survival was shown in two large breast cancer patients' data sets [3]. These data indicate that *KRT16* expression is associated with higher tumor aggressiveness and shorter relapse-free survival. To further elucidate the role of K16 in cancer progression and metastasis, this study set out to investigate the biological role of K16 in metastatic breast cancer cell lines and evaluate the clinical relevance of K16 in metastatic breast cancer patients by analyzing the K16 expression in CTC, i.e., the seeds of metastasis.

2. Materials and Methods

2.1. In Silico Analysis

KRT1-20, *CDH1*, and *VIM* gene expression data of 51 breast cancer cell lines were obtained from GEO, accession number GSE69017 [13]. A hierarchical cluster analysis was performed on the gene expression data to compare expression levels of *KRT1-20* in the different breast cancer molecular subtypes and to evaluate the epithelial- and mesenchymal-like phenotype based on *CDH1* and *VIM*. The dataset was normalized to the mean value of each probe set.

2.2. Cell Culture

Ten human breast cancer cell lines (MDA-MB-468, MDA-MB-231, BT549, HS-578T, MCF7, T47D, MDA-MB-361, BT474, SKBR3, GI-101A), one normal-like breast epithelial cell line (MCF-10A), and one skin squamous carcinoma cell line (A431) were brought into a culture. All cell lines were obtained from ATCC. The cells were cultured in either DMEM media (catalog no. P04-03600, Aidenbach, Germany) or RPMI 1640 media (catalog no. P04-17500, Aidenbach, German) and incubated at 37 °C and 5% CO_2 or 10% CO_2, according to ATCC's instructions. Both media were supplemented with 10% fetal bovine serum (FBS) (Gibco—Life Technologies), 1% L-glutamine (catalog no. 25030-024, Gibco—Life Technologies), and 1% penicillin/streptomycin (catalog no. 15140-122, Gibco—Life Technologies). All cells were grown in a 25 cm^2 flask until confluence was reached. Cells were washed with DPBS (catalog

no. 14190-094, Gibco, Life Technologies) and harvested using trypsin/EDTA (catalog no. 25200-072; Thermo Fisher Scientific). A test for mycoplasma was regularly performed in all cultures to detect and prevent any potential mycoplasma contamination [14].

2.3. EMT Induction Assay

EMT was induced using StemXVivo EMT Inducing Media Supplement (catalog no. CCM017; R & D Systems, Wiesbaden-Nordenstadt, Germany). This media includes a cocktail of E-cadherin, SFRP-1, and DKK-1 blocking antibodies and WNT-5 and TGF-β1 recombinant proteins. MCF7 cells were seeded in standard culture media containing $1\times$ StemXVivo EMT Inducing Media Supplement according to the manufacturer's instructions [15,16]. Different culture conditions were tested through EMT treatment, such as hypoxia (1% O_2) and starvation (0.5% FCS) compared to standard conditions (21% O_2, 10% FCS) at different time-points of 24 h, 72 h, and 120 h. Control cells were seeded and put through the same conditions as EMT-treated cells. The experiment was processed in duplicate. Characterization of EMT-induced cells was performed by western blot and RT-qPCR.

2.4. KRT16 Overexpression

Keratin 16 plasmid DNA (catalog no. OHu24939D; GenScript) and pcDNA3.1+/c-(K)-DYK vector were used for transfection into MCF7 cells to overexpress *KRT16* and as a vector control, respectively. ORFs cloned in the pcDNA3.1+/C-(K) DYK vector were expressed in MCF7 cells as a tagged protein with a C-terminal DYKDDDDK tag. First, cells were seeded in 6-well plates (9×10^5 cells/well) in a standard medium containing 10% FBS, the day prior to transfection. Then cells were transfected with plasmid constructs (final concentration 2.5 μg) using Lipofectamine 3000 (catalog no. L3000008; Invitrogen) and Opti-MEM medium (Thermo Fisher Scientific), following the manufacturer's protocol [17]. The experiment was processed in triplicate. Keratin 16 overexpression was assessed by western blot using a DYKDDDDK Tag monoclonal anti-mouse clone [5A8E5] (catalog no. A00187; GenScript) and RT-qPCR. Transfection efficiency was assessed by immunofluorescence staining. After 24 h of transfection, cells were fixed in 4% paraformaldehyde (PFA) for 15 min on a culture slide and washed with PBS two times. Next, the cells were incubated with 0.2% Triton X-100 for 10 min and washed with PBS three times. Then, the cells were incubated with 1% BSA at room temperature for 30 min, followed by incubation with the primary antibody against DYKDDDDK Tag overnight at 4 ℃. On the second day, cells were washed with PBS and incubated with the secondary antibody goat anti-mouse IgG labeled with Alexa Fluor 488 (catalog no. A28175; Life technologies, 1:200) at room temperature for 90 min. After being washed with PBS, the cells were incubated for 1 min with DAPI (Janssen Diagnostics, 1:5000). The cells were examined by fluorescence microscopy (ZEISS Axio Observer). Pictures were taken for four fields and cells were counted by the Cell counter program [18]. The experiment was performed in duplicate. The results are expressed as the average percentage number of positive cells within a transfected cell population relative to the total number of cells.

2.5. KRT16 Knockdown

Transfection using a smart pool of *KRT16* interfering RNA (siRNA) duplexes, *KRT16*siRNA1 (GGAGAUGCUUGCUCUGAGA); *KRT16*siRNA2 (GGCCAGAGCUCCUA-GAACU); *KRT16*siRNA3 (GGAACAAGAUCAUUGCGGC); *KRT16*siRNA4 (GCGGAGAU-GUGAACGUGGA) (catalog no. L-017550-02-0005; Dharmacon), and Lipofectamine™ RNAiMAX Transfection Reagent (catalog no. 13778075; Thermofisher) was performed on MDA-MB-468 cells, according to the manufacturer's protocol [19]. ON-TARGETplus Non-targeting Pool (catalog no. D-001810-10-05; Dharmacon) was used as a control. The experiment was processed in duplicate, with a final oligonucleotide concentration of 20 nM. *KRT16* knockdown efficiency was assessed by RT-qPCR, and western blot and cells were used for further experiments 48 h after transfection.

2.6. Protein Level Assessment

Protein levels of the cells were measured after treatment with StemXVivo EMT Inducing Media, *KRT16* knockdown, and *KRT16* overexpression as follows: the cells were scraped in PBS and centrifuged at 1500×*g* for 2 min, the supernatant was removed, and the cell pellet was resuspended in RIPA lysis buffer containing the Protease Inhibitor Cocktail (Thermo Fisher Scientific). After incubation on ice for 30 min, cells were homogenized by ultrasonic treatment for 5 s and centrifuged at 16100×*g* for 15 min at 4 ˚C. The protein concentration was determined by a BSA protein assay kit (catalog no. 23227; Thermo Scientific), according to the manufacturer's instructions [20].

Protein extracts were loaded in 1x SDS buffer and denatured at 95 ˚C for 5 min. The denatured proteins were separated by SDS-PAGE using 10% polyacrylamide gels and blotted onto PVDF membrane. Detection of proteins was performed by incubation with the following specific antibodies: keratin 16 monoclonal anti-mouse clone [Ag11240] (catalog no. 66802-1-Ig; Proteintech), E-cadherin monoclonal anti-rabbit clone [EP700Y] (catalog no. ab40772; Abcam), vimentin monoclonal anti-mouse clone [RV202] (catalog no. 550513; BD Pharmingen™), SNAI2 monoclonal anti-mouse clone [A7] (catalog no. sc-166476; Santa Cruz Biotechnology), and HSC70 monoclonal anti-mouse clone [B-6] (catalog no. sc- 7298; Santa Cruz Biotechnology). Protein bands were determined using SignalFire™PlusECL reagent (Cell Signaling Technology, Danvers, USA), X-ray films (CEA, Hamburg, Germany), and LI-COR C-DiGit Chemiluminescence Western Blot Scanner, according to the manufacturers' instructions [21].

2.7. Quantitative Transcript Analysis

Relative RNA expression levels were measured by RT-qPCR using an equivalent of 15 ng of total RNA, isolated using the NucleoSpin RNA kit (catalog no. 740955.50, Macherey Nagel), according to manufacturer's instructions [22]. The RNA was reverse transcribed, performing the First Strand cDNA Synthesis Kit (Thermo Scientific) according to the manufacturer's instructions. Primers and product length for each gene are described in Table 1. AccuPower® 2X GreenStar™ master mix solution (catalog no. K-6253, Bioneer) RT-qPCR was used in 10 µL reaction volumes containing 10 pmol of each primer, 15 ng cDNA, and 1 x master mix (Tris-HCl, 60 mM KCl, 1.5 mM MgCl$_2$, SYBR Green I Dye, Hotstart DNA polymerase (1U), dNTP mixture (each 250 µM)). The PCR conditions were as follows: initial denaturation at 95 ˚C for 5 min, followed by 40 amplification cycles of 95 ˚C for 15 s and 61.4 ˚C for 30 s, and a melting curve of 65.0 ˚C to 95.0 ˚C, with increments of 0.5 ˚C every 5 s. The RT-qPCR reactions were run in duplicate using a CFX96 Touch™ Real-Time PCR Detection System. Data were analyzed by applying the ΔΔCT calculations [23], using the GAPDH expression for normalization to calculate mRNA expressed as fold changes $2^{-(\Delta\Delta Ct)}$.

Table 1. Specific primers used in RT-qPCR.

Gene	Forward Primer (5′-3′)	Reverse Primer (5′-3′)	Product Length
E-cadherin (*CDH1*)	CGAGAGCTACACGTTCACGG	GGGTGTCGAGGGAAAAATAGG	119 bp
N-cadherin (*CDH2*)	TGCGGTACAGTGTAACTGGG	GAAACCGGGCTATCTGCTCG	123 bp
Vimentin (*VIM*)	GACGCCATCAACACCGAGTT	CTTTGTCGTTGGTTAGCTGGT	238 bp
SNAI2 (SLUG)	TGTGACAAGGAATATGTGAGCC	TGAGCCCTCAGATTTGACCTG	203 bp
SNAI1	ACTGCAACAAGGAATACCTCAG	GCACTGGTACTTCTTGACATCTG	242 bp
ZEB1	GATGATGAATGCGAGTCAGATGC	ACAGCAGTGTCTTGTTGTTGT	86 bp
ZEB2	GGAGACGAGTCCAGCTAGTGT	CCACTCCACCCTCCCTTATTTC	107 bp
TWIST1	AAGGCATCACTATGGACTTTCTCT	GCCAGTTTGATCCCAGTATTTT	96 bp
WNT5A	ATTCTTGGTGGTCGCTAGGTA	CGCCTTCTCCGATGTACTGC	159 bp
NOTCH1	GAGGCGTGGCAGACTATGC	CTTGTACTCCGTCAGCGTGA	140 bp
KRT8	CAGAAGTCCTACAAGGTGTCCA	CTCTGGTTGACCGTAACTGCG	194 bp
KRT18	GCTCAGATCTTCGCAAATACTGT	CTTCCTCTTCGTGGTTCTTCTTC	250 bp
KRT19	ACCAAGTTTGAGACGGAACAG	CCCTCAGCGTACTGATTTCCT	181 bp
KRT16	GACCGGCGGAGATGTGAAC	CTGCTCGTACTGGTCACGC	91 bp
CD24	CTCCTACCCACGCAGATTTATTC	AGAGTGAGACCACGAAGAGAC	166 bp
CD44	CTGCCGCTTTGCAGGTGTA	CATTGTGGGCAAGGTGCTATT	109 bp
GAPDH	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGTCATACTTCTCATGG	197 bp

2.8. Migration Assay

Two migration assays were performed in this study, the Wound-healing and transwell (Boyden chamber) assays, to ensure the viability of modulated cells throughout the migration process. For the wound-healing assay, 1.2×10^6 MDA-MB-468 cells were plated in serum-free DMEM media in a 6-well plate and incubated at 37 ˚C. After 24 h, a scratch was made, and pictures were taken under the microscope at different time-points. The experiment was processed in duplicate. The images were analyzed by ImageJ software (MRI_Wound_Healing_Tool.ijm) [24]. For the transwell migration assay, 1×10^5 MCF7 cells were cultured in serum-free DMEM media in the upper chambers, with 8.0 μm poresof BD Cell Culture (BD Falcon). DMEM containing 10% FCS was used as a chemoattractant in the lower chamber. Plates were incubated under standard conditions, and migration could proceed for 24 h. Non-migrated cells in the upper chambers were removed with cotton swabs, and the remaining cells were fixed in 4% PFA and stained with crystal violet. Using a converted microscope (ZEISS Axio Observer), pictures were taken of four fields and counted by the Cell counter program [18]. The experiment was performed in triplicate, and the results are expressed as the average number of cells with standard deviation.

2.9. Proliferation Assay

Cell proliferation was assessed after *KRT16* knockdown and *KRT16* overexpression. The cells were seeded and incubated at 37 ˚C with 5% CO_2. Monitoring the cell density and viability over time was measured in triplicate with the Vi-CELL™ XR 2.04 (Beckman Coulter), which depended on performing the Trypan Blue Dye Exclusion method [25]. The experiment was performed in triplicate, and the results are represented as the average number of cells.

2.10. Immunocytochemical Staining

To visualize the actin filaments and a cell–cell adhesion protein (E-cadherin) in K16-expressing cells, staining was performed using phalloidin Alexa Fluor 555 (catalog no. A34055; Life Technologies, 1:00) and E-cadherin monoclonal anti-rabbit clone [EP700Y] (catalog no. ab40772; Abcam, 1:500) as follows: 30,000 cells were seeded onto a chamber slide and incubated at the standard culture condition for 48 h. The cells were fixed with 4% PFA for 10 min and incubated with 0.1% Triton X-100 for 10 min, 10% AB-serum at room temperature for 20 min, primary antibody against E-cadherin overnight at 4 ˚C, and secondary antibody goat anti-rabbit IgG labeled with Alexa Fluor 488 (catalog no. A1008; Life technologies, 1:200) at room temperature for 60 min. Cells were washed with PBS, incubated with phalloidin and DAPI (1:1000) for 60 min, and examined by confocal microscopy (Leica TCS SP8). Using the Bitlplane Imaris software, the cell thickness was estimated by counting the number of image planes required to cover the entire cell volume. Actin filament length was determined by calculating the distance between the extreme points of the detected filament structures. Movies were produced after 3D reconstruction by the mean of image interpolation of confocal images stacks, voxel $(80 \times 80 \times 300)$ nm.

1.4. Blood Collection and Processing

Twenty metastatic breast cancer patients were included into the study after giving written informed consent (local ethical committee approval number: PV5392). Patients were treated at the Department of Gynecologic Oncology, University Medical Center Hamburg-Eppendorf and received therapy according to international guidelines. The selection criteria were women with metastatic breast cancer, independent of the tumor's hormone status. Blood was taken upon a progression of the disease. Peripheral blood samples were collected into EDTA-containing tubes, kept at room temperature, and processed within 1 h. The density gradient Ficoll (catalog no. 17-1440-03; GE Healthcare) was used for mononuclear cell enrichment as before [26].

2.12. Enrichment of CTCs using a Spiral Microfluidic Chip

A spiral microfluidic chip was designed to isolate CTCs according to size and density [27], in which the larger and denser particles (i.e., CTCs) focused and aligned near the inner wall, while the small particles occupied the lateral positions near the outer wall. One inlet and two outlet tubes were connected to a spiral chip. The inlet tubing was connected to a syringe pump (catalog no. 78-8110; Kd. Scientific), and the outlet tubing was connected to two sterile 15 mL collection BD Falcon tubes. An initial washing was performed before sample processing using 5 mL of 5% NaClO, 5 mL H_2O, and finally 5 mL sterile 1×PBS at a flow rate of 1 mL/min. The sample was transferred into a syringe and pumped through the spiral chip at a flow rate of 1.7 mL/min. The enriched CTC fraction was collected on a microscope slide in a cytospin funnel and spun down at 190× g for 5 min.

2.13. CTC Immunophenotyping

K16 staining was established using 7.5 mL blood samples from anonymous healthy donors spiked with tumor cell line cells. The blood was processed as mentioned above, and the mononucleated cell layer was collected and spiked with MCF7 cells (K16−/C11+) as the K16 negative control and A431 cells (K16+/C11−) or MDA-MB-468 (K16+/C11+) as K16 positive control. The spiked samples were spun on a microscope slide. Immunofluorescence staining was used to identify the enriched CTCs and spiked tumor cell line cells. Briefly, cells were fixed in 4% PFA for 10 min on a microscope slide and washed with PBS three times. Next, the cells were incubated with 10% AB-serum at room temperature for 1 h, followed by incubation with the primary antibody against keratin 16 (catalog no. 66802-1-Ig; Proteintech, 1:500) overnight at 4 °C. On the second day, cells were washed with PBS and were incubated with the secondary antibody rabbit anti-mouse IgG labeled with Alexa Fluor 546 (catalog no. A11060; Life technologies, 1:200) at room temperature for 1 h. After being washed with PBS, the cells were incubated for 1 h with DAPI (1:1000), CD45 Alexa Fluor 647 (catalog no. 130-110-633; MACS, 1:200), and C11 Alexa Fluor 488 (catalog no. ab187773; Abcam, 1:300) in order to detect keratins (K4/5/6/8/10/13/18). After staining, the cells were washed with PBS, covered with a coverslip, and examined by fluorescence microscopy (ZEISS Axio Observer). CD45−/K16+/C11+ cells with an intact nucleus were interpreted as CTCs.

2.14. Statistical Analyses

Statistical analyses were performed using R (R Foundation for Statistical Computing, version 4.0.1) and In-Silico Online, version 2.3.0 [28], and graphs were generated using GraphPad Prism (GraphPad Software, San Diego, CA, USA). Mean values were given with standard deviations. Statistical significance was defined as $p < 0.05$. Linear regression was used to determine the rate of migrated cells over time. Relapse-free survival (RFS) was determined using a Kaplan–Meier curve and log-rank test. The primary endpoint of RFS was defined as the time in months from primary diagnosis until the first progression. The clinical variables (estrogen receptor (ER) status, progesterone receptor (PR) status, ERBB2 status, T-stage, and age) were compared to the K16-positive CTCs of metastatic breast cancer patients. A power analysis for sample size was performed by PASS, version 20.0.2.

3. Results

3.1. Keratin 16 Is Overexpressed in the Basal-Like Breast Cancer Subtype

In in silico analysis, we investigated the expression of 20 keratins in 54 breast cancer cell lines (GSE69017, Figure S1) and could show that *KRT16* was upregulated in 24% (13/54) of breast cancer cell lines of predominantly the basal-like (53.8%), ERBB2 enriched (23%), claudin-low (15%), and normal-like (7.7%) subtypes that mainly overexpress *CDH1* and *VIM*, whereas *KRT16* was downregulated in cell-lines of the luminal A and luminal B molecular subtypes. Furthermore, we investigated the correlation between mRNA expression levels of *KRT16* and the EMT-associated genes *CDH1* and *VIM*. A significant positive correlation ($R^2 = 0.643$, $p = 0.018$, Pearson's r) was observed between *CDH1* and

KRT16. In order to extend these data at the protein level, we assessed the protein expression of K16 in different breast cancer cell lines presenting various breast cancer subtypes using western blot analysis. In addition to K16, the expressions of E-cadherin (CDH1) and Vimentin (VIM) were determined to assess the degree of epithelial- and mesenchymal- like phenotype, respectively. Western blot analysis revealed that K16 protein was more abundant in carcinoma cells of the basal-like A and normal-like subtypes that also express CDH1 and VIM, while K16 expression was completely absent in cell lines of the luminal A and B subtypes that express CDH1 but not VIM and the ERBB2-overexpressing subtype that does not express CDH1 and VIM (Figure 1A,B). A significant correlation between mRNA and protein levels in the expression of CDH1 (R^2 = 0.851, p = 0.0001), VIM (R^2 = 0.653, p = 0.0047), and K16 (R^2 = 0.773, p = 0.0008) were observed (Figure 1C). Taken together, these data reveal a correlation between K16 expression and mesenchymal-like phenotype.
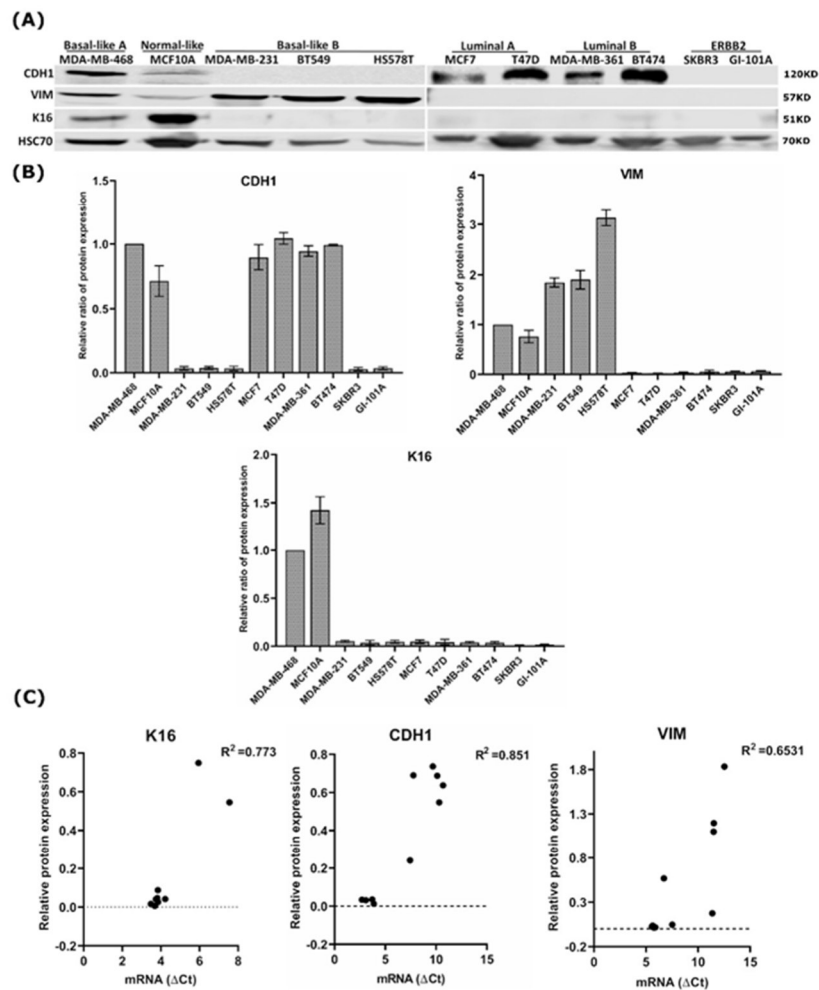

Figure 1. Analysis of the K16 expression in breast cancer cell lines. (**A**) Expression of K16, adherent junction protein CDH1, and mesenchymal intermediate filament VIM in a panel of human breast cancer cell lines using western blot. (**B**) The relative ratio of protein expression of K16, VIM, and CDH1; all quantified values are normalized to the expression level of HSC70 protein and then to MDA-MB-468 as a control (**C**) Correlation between mRNA and protein expression levels of K16, CDH1, and VIM in breast cancer cell lines.

3.2. EMT Induction Leads to Overexpression of K16

Next, we sought to address whether the induction of the EMT process leads to K16 upregulation. Therefore, we investigated changes in K16 expression during EMT induction in the MCF7 cell line, which normally does not express K16 (Figure 1A). Microscope imaging of EMT-induced MCF7 cells showed substantial changes in their morphology after treatment with the EMT-inducing media supplement. The EMT-induced cells showed a mesenchymal-like, spindle-shaped morphology, losing all intercellular contacts, whereas the untreated cells showed a typical epithelial-like morphology with extended cell–cell contacts (Figure 2A). In order to assess EMT induction efficiency, the mRNA expression levels of the MET/EMT related genes *CDH1*, *KRT8*, *KRT18*, *KRT19*, *CD24*, *VIM*, *CDH2*, *SNAI2*, *CD44*, *ZEB1*, *ZEB2*, *SNAI1*, *TWIST1*, *WNT5A*, and *NOTCH1* were examined at various EMT culture conditions and different time-points (Figure 2B). Mesenchymal markers *VIM*, *CDH2*, *SNAI2*, and *CD44* were significantly ($p < 0.0001$) upregulated in the EMT-induced cells as compared to untreated controls, in particular, the cells that were under standard conditions (10% FCS, 21% O_2) and starvation conditions (0.5% FCS, 21% O_2) (Figure 2B). These cells showed slight changes in the expression of epithelial markers, such as *CDH1* and *KRT8*. We further observed that EMT-induced cells could continue to transition after five days (Figure 2B).
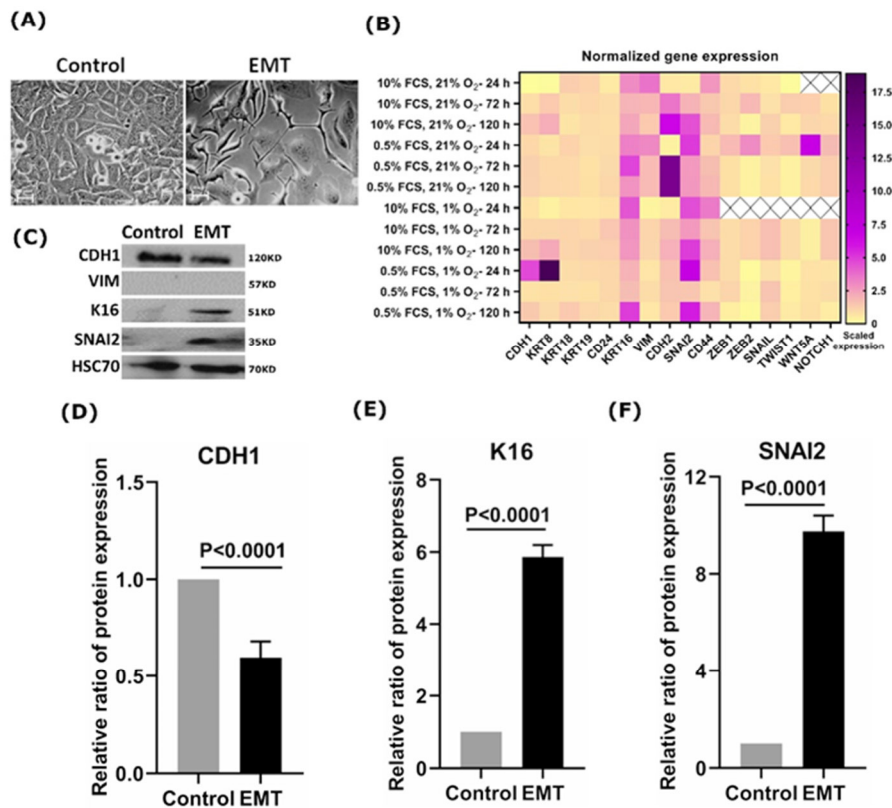

Figure 2. Analysis of the K16 expression in EMT-induced cells. (**A**) Microscopic images showed morphological changes in MCF7 cells during EMT induction; scale bar represents 100 μm. (**B**) Heat-map of EMT/MET specific genes expression in EMT-induced cells at different culture conditions and time points; relative mRNA expression values were normalized to *GAPDH* and subsequently displayed relative to gene expression as a fold change $2^{-(\Delta\Delta CT)}$. (**C**) Protein expression profile of K16, CDH1, VIM, and SNAI2 during EMT induction on MCF7 cells. HSC70 was used as a reference control. The relative ratio of protein expression of (**D**) CDH1; (**E**) K16; (**F**) SNAI2.

Interestingly, K16 was upregulated in EMT-induced cells under all conditions in which cells were exposed at all time points (Figure 2B). Immunoblot analysis showed overexpression of K16 and SNAI2 (Figure 2C,E,F) with decreased expression of CDH1 compared to untreated cells (Figure 2C,D), which indicates the disruption in adhesion junction formation and loss of epithelial properties, whereas no change in VIM expression was detected (Figure 2C). Overall, EMT promotes the expression of K16 which highlights a crucial role of K16 in EMT execution.

3.3. Overexpression of KRT16 Leads to EMT Induction

To further investigate the association between K16 and the EMT, *KRT16* was overexpressed in the K16-negative MCF7 cells (Figure 1A). MCF7 cells were transfected with a *KRT16* coding sequence containing plasmid or non-target plasmid as a control, resulting in transfection efficiency of 81.4% of the total transfected cell population (Figure S2 and Table S1). Transfection of *KRT16* increased the mean fold-change of mRNA gene expression levels by 30,000-fold in the treated MCF7 cells, compared to the non-target treated control cells (Figure 3A). Furthermore, K16 protein expression was strongly induced in the treated MCF7 cells, whereas no expression of K16 could be detected in the non-target treated control cells (Figure 3B,C). The treated MCF7 showed a spindle-shaped morphology with long and thin stress fibers, characteristic of epithelial-to-mesenchymal transition, whereas non-target treated and untreated control cells were regular in morphology (Figure 3E and Figure S3). In line with this, mRNA analysis revealed overexpression of the mesenchymal specific genes *VIM*, *CDH2*, *SNAI2*, *ZEB1*, *ZEB2*, *TWIST1*, and *WNT5A* in *KRT16* induced cells (p = 0.0001, Two-sample Wilcoxon rank test; Figure 3F), but no changes in the expression levels of *CDH1*, *KRT8*, *KRT18*, *KRT19* and *CD24*. Furthermore, no changes were detected in VIM and CDH1 protein expression levels (p = 0.430; Figure 3B,D) upon *KRT16* induction. Immunocytochemical staining showed abundant expression of E-cadherin localized in the cytoplasm and weakening cell–cell adhesion in K16-induced MCF7 cells, whereas untreated MCF7 cells demonstrated stable adherens junctions at the cell edges (Figure 3G). Moreover, overexpression of K16 resulted in a reorganization of actin microfilaments to form long, parallel, thin stress fibers compared to untreated MCF7 cells (Figure 3G and Figure S4 and Video S1). A highly significant increase (p = 0.0027, Welch's two sample t-test) in the length of actin microfilaments was detected with a reduction in the cell thickness (p = 0.0162, Welch's two sample t-test) of MCF7 cells that induced K16 compared to control cells (Figure 3H,I).

The phenotypic changes induced by *KRT16* overexpression were further investigated by using the Boyden chamber assay for migration analysis (Figure 3J). Compared to the non-target treated cells, MCF7 cells overexpressing *KRT16* had a significantly higher mean migration rate of 1.88 times (95% CI: 1.47–2.42; p < 0.001, negative-binomial generalized linear mixed-effects model; Figure 3K). The equal mean proliferation rates of 79.4 and 78.3 of the target and empty-vector-control treated cells, respectively, indicate that the proliferation rate was not modified by *KRT16* overexpression and that the increased migration wasnot the consequence of a higher proliferation rate (p = 0.8514, Welch's two sample t-test; Figure 3L). Hence, we could show that overexpression of *KRT16* contributes to EMT and toa more aggressive phenotype.
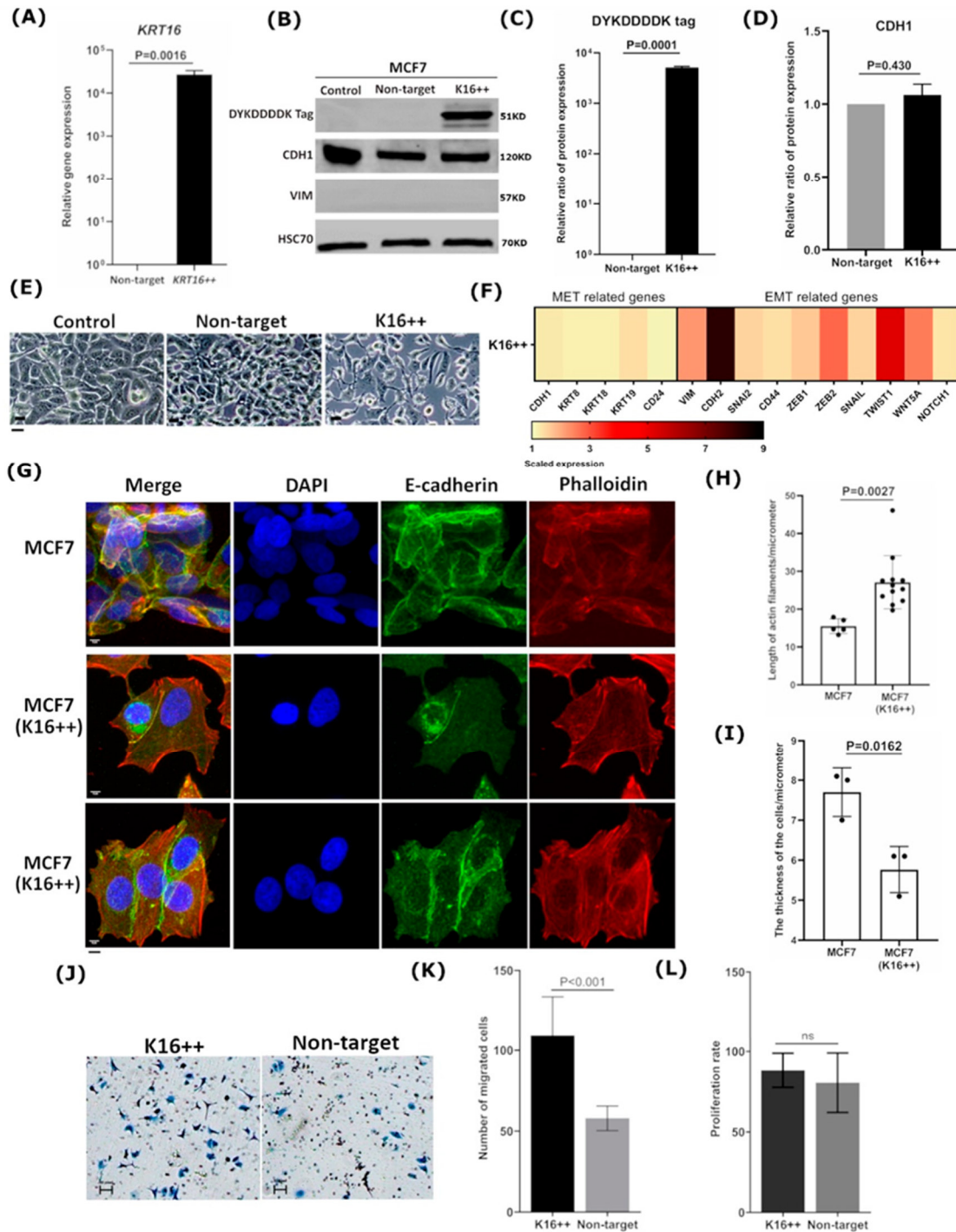
Figure 3. Overexpressing of K16 in MCF7 cells promotes a mesenchymal phenotype. (**A**) Relative mRNA expression was verified after inducing KRT16. (**B**) Immunoblot: DYKDDDDK tag was used to detect the expression of transfected K16 in MCF7 cells (K16++) compared to transfected empty vector (non-target) and untreated cells (control); CDH1 and VIM expression were

investigated after 48h of treating MCF7 cells, loading control: HSC70. The relative ratio of protein expression of (**C**) DYKD-DDDK tag and (**D**) CDH1. (**E**) Cellular morphological changes in MCF7 cells after inducing the expression of K16 protein (K16++) were observed under a normal microscope compared to transfected non-target and untreated cells (control); scale bars 20 μm. (**F**) Heat-map of EMT/MET-specific genes was investigated after KRT16 enhancement; mRNA expression values were normalized to GAPDH and subsequently displayed relative to gene expression as a fold change $2^{-(\Delta\Delta CT)}$. (**G**) Immunocytochemistry staining of actin filaments by phalloidin (red), intercellular adhesion by E-cadherin (green), and nucleus by DAPI (blue) to visualize the morphological changes in MCF7 cells that induced K16; scale bar represents 5 μm. Differences between MCF7 cells and MCF7 cells with induced K16 in (**H**) the length of actin filaments and (**I**) the thickness of the cells. (**J**) Microscopy images of migrated cells by Boyden Chamber to analyze the motility of MCF7 cells overexpressing K16 relative to the non-target control; the cells were seeded on transwell chambers and incubated for 24 h; scale bar represents 100 μm. (**K**) The migrated cells were counted after staining cells with crystal violet, the data were generated as the mean ± SD., n = 3. (**L**) Proliferation assay: cell proliferation rates of K16++ and non-target control cells were determined using Trypan Blue Dye, and the cells were counted by Vi-CELL Counter device; the data are expressed as the mean ± SD.; n = 3. The p-values were calculated with Welch's two sample t-test ($p < 0.05$).

3.4. K16 Knockdown Changes the Mesenchymal Phenotype

Next, we sought to investigate whether K16 expression contributes to EMT-associated cell (de)differentiation. To that end, *KRT16* was knocked down in MDA-MB-468 cells using siRNA directed against *KRT16* mRNA. MDA-MB-468 is a basal-like cell line expressing K16, but it also contains cells having an epithelial-like phenotype (Figure 1A). *KRT16* was successfully depleted, leading to a mean downregulation of *KRT16* mRNA of five times compared to control cells ($p < 0.0001$, Welch's t-test; Figure 4A). Because of gene silencing by siRNA, protein K16 levels were downregulated to 0.18 as compared to the control cells (Figure 4B,C). In addition, CDH1 and VIM expression were lowered by 0.56 and 0.35 in K16-depleted cells, respectively, compared to untreated cells (Figure 4B,D,E), indicating a less mesenchymal phenotype upon K16-knockdown. Furthermore, the mRNA expression of selected mesenchymal-associated genes (*VIM, CDH2, SNAI2, CD44, TWIST1, NOTCH1, ZEB1, ZEB2, SNAI1*, and *WNT5A*) were significantly downregulated (p = 0.0135, Two- sample Wilcoxon rank test), whereas the expression of selected epithelial-associated genes (*CDH1, KRT8, KRT18, KRT19*, and *CD24*) stayed unaltered (p = 0.127, Two-sample Wilcoxon rank test; Figure 4F), indicating that MDA-MB-468 expressed mesenchymal markers to a lesser extent after *KRT16* knockdown.

Because one of the characteristics of breast cancer cells with a mesenchymal-like phenotype is increased migration properties, we assessed whether knockdown of *KRT16* impacts cell migration. Using the wound-healing assay, it could be observed that *KRT16* knockdown led to a mean 70% slower migration (Figure 4G,H; p = 0.0032, ANCOVA test), indicating that the depletion of K16 protein expression in the MDA-MB-468 cell line impaired cell migration. Equal means of proliferation rates in the siRNA treated and control cell lines confirmed that the migration was not the consequence of a higher proliferation rate (p = 0.9714; Wilcoxon rank-sum test, Figure 4I). Taken together, these results indicate that loss of K16 directly influences the capacity of cells to migrate, which might be the consequence of loss of mesenchymal properties.
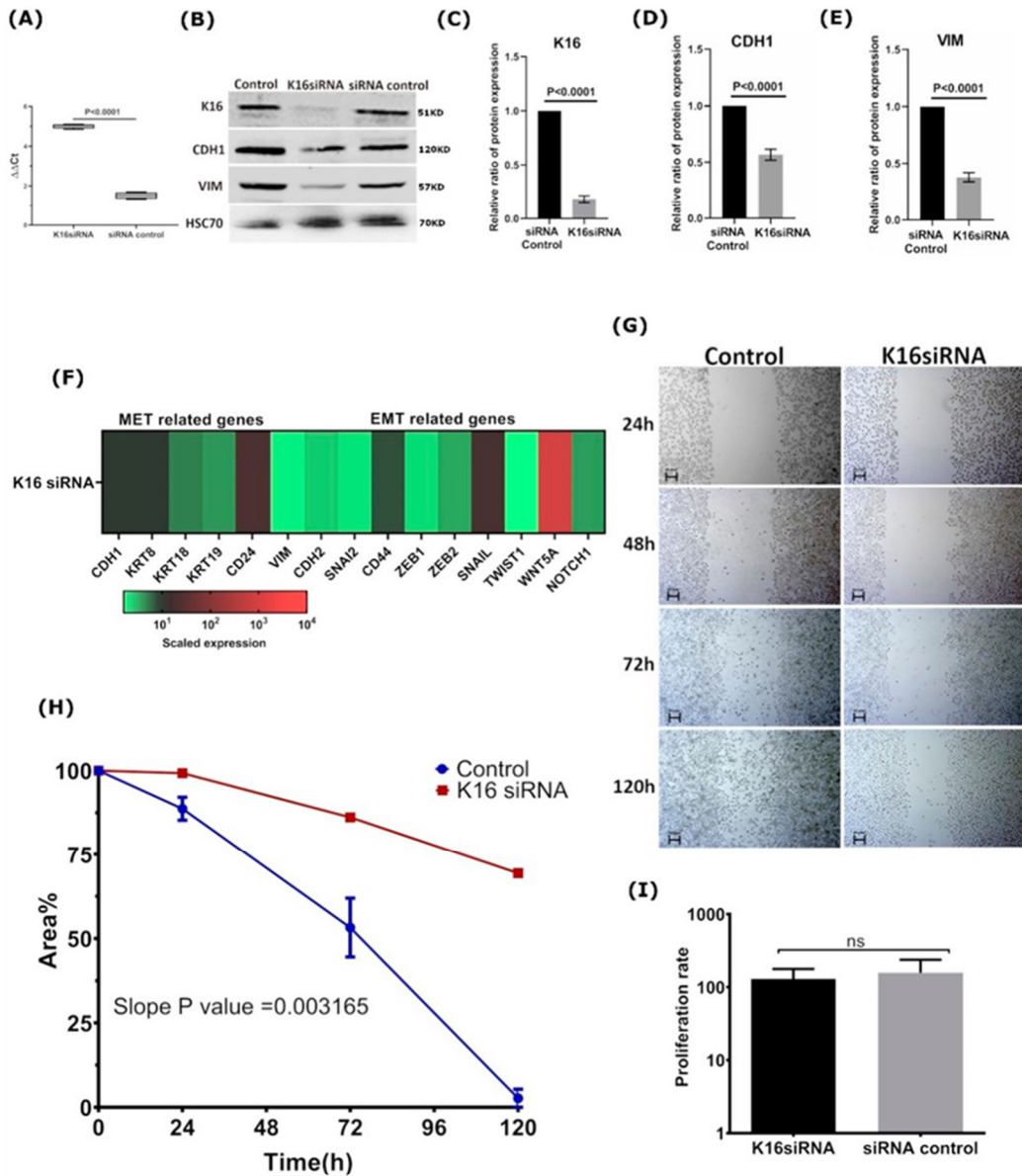
Figure 4. *KRT16* knockdown in MDA-MB-468 impaired cell migration (**A**) mRNA expression level was verified after *KRT16* knockdown compared to the non-target (siRNA control). (**B**) Immunoblot to verify the expression levels of CDH1 and VIM in *KRT16* depleted cells. The relative ratio of protein expression of (**C**) K16; (**D**) CDH1; and (**E**) VIM. (**F**) Relative mRNA expression of EMT/MET-specific genes was investigated after *KRT16* depletion; mRNA expression values were normalized to *GAPDH* and subsequently displayed relative to gene expression as a fold change $2^{-(\Delta\Delta CT)}$. (**G**) Microscopy images of migrated cells by wound-healing assay were performed on *KRT16* depleted cells (*KRT16* siRNA) compared to control cells (siRNA control) at different time points of 0 h, 24 h, 72 h, and 120 h; $n = 2$, scale bar represents 100 µm. (**H**) In migration, linear regression analysis was performed for the velocity slope of *KRT16*-siRNA transfected cells and untreated cells (control). (**I**) Cell proliferation rates of *KRT16*-siRNA transfected and siRNA control cells were determined using Trypan Blue Dye, and the cells were counted by a Vi-CELL Counter device, the data are expressed as the mean ± SD.; $n = 3$. The *p*-values were calculated by the Wilcoxon rank-sum test ($p < 0.05$).

3.5. K16 Expression in CTCs Correlates with a Worse Survival

To evaluate the clinical relevance of the expression of K16 in CTCs, an immunofluorescence staining protocol was established by spiking blood with K16 positive and negative cancer cell lines (Figure 5A). Using the optimal staining conditions, specific detection of tumor cell line cells MDA-MB-468 and A431 was achieved with no background in the K16-negative cell line MCF7.
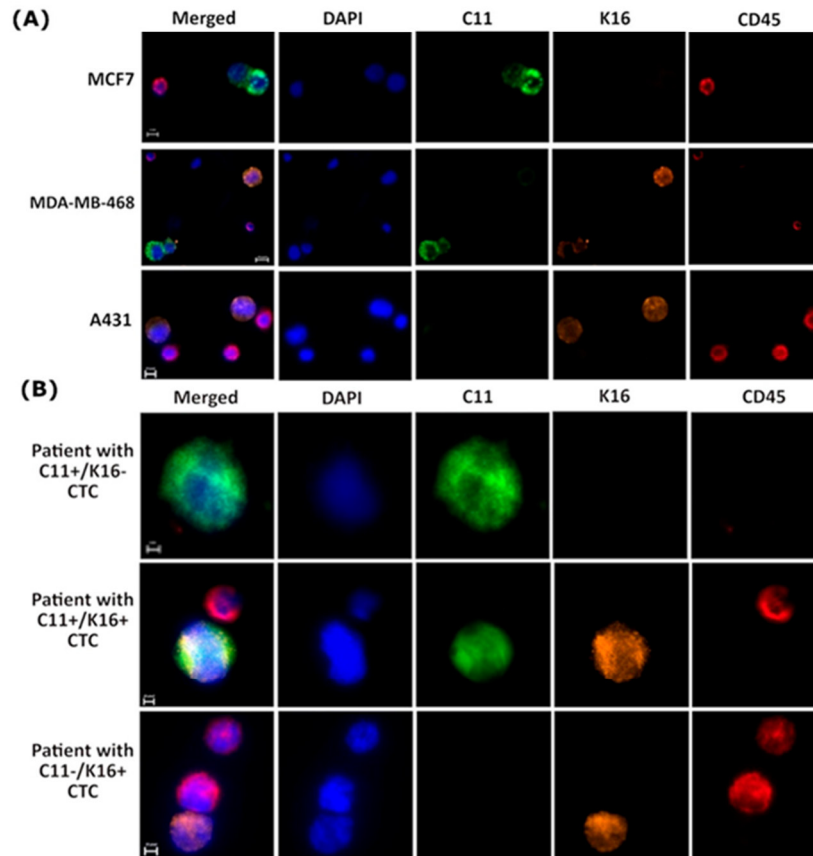

Figure 5. K16 immunocytochemistry staining (**A**) Establishment K16 immunocytochemistry staining; cells were stained by DAPI (blue), CD45 (AF647; red), C11 (AF488; green), and K16 (AF546; orange); scale bar represents 10 µm. (**B**) Detection of CTCs in metastatic breast cancer patients (*n* = 20); CTCs cells (DAPI+, C11+, K16+, and CD45−) were differentiable from the WBCs (DAPI+ and CD45−); scale bar represents 10 µm.

Next, blood was acquired from 20 metastatic breast cancer patients that were experiencing disease progression. CTCs (keratin+/DAPI+/CD45−) were detected in 19/20 patients, follow-up was not available for one patient (Table S2). The blood of five patients counted ≤5 CTCs, and 14 patients counted ≥5 CTCs per 7.5 mL blood. A total of 64.7% of detected CTCs were K16+/C11−, and 16.6% were C11+/K16−, whereas 18.7% of detected CTCs were K16+/C11+ (Figure 5B and Table S2).

The median follow-up time of breast cancer patients was 32 months. Relapse-free survival (RFS) analysis was performed to test the difference in survival between metastatic breast cancer patients diagnosed with K16 expressing CTCs only (i.e., K16+/C11−and/or K16+/C11+) and patients that were diagnosed with at least one K16-negative CTC(i.e., C11+/K16−), patients without detectable CTCs were excluded from this analysis. Pa-

tients with K16+ CTCs had shorter relapse-free survival (median: 12.9 months) compared to patients who had CTCs with negative expression of K16 (median: 75 months; $p = 0.0042$; Figure 6).
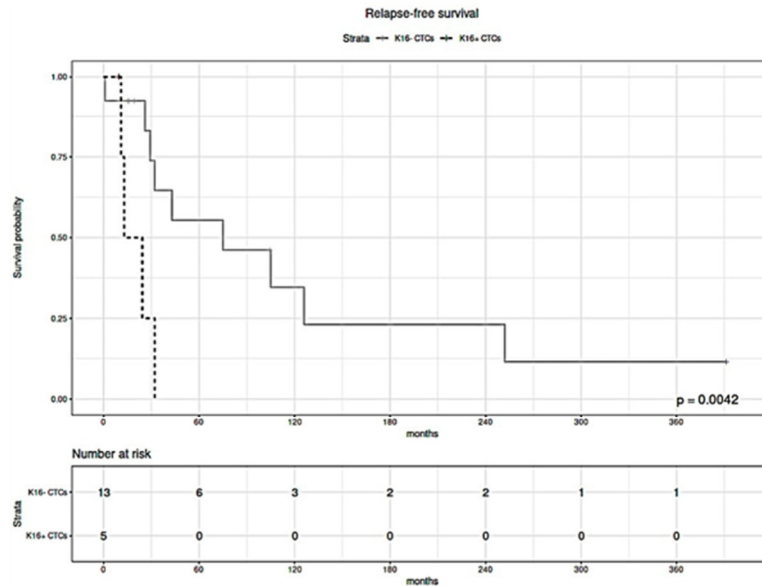

Figure 6. Kaplan–Meier estimates of the relapse-free survival of metastatic breast cancer patients with K16+ CTCs (black dashed line) and K16– CTCs (solid gray line). Statistical significance was determined by a log-rank test. Shorter relapse-free survival correlates with the presence of K16+ CTCs in the blood ($p = 0.0042$).

Both in uni-and multivariable analysis, the presence of K16-positive CTCs was correlated with shorter survival (Table 2). Although the expression of estrogen receptor (ER) was correlated with improved RFS in multivariable analysis, ER-status was not significantly associated with survival in univariable analysis, possibly due to the small number of cases. These results indicate that K16 expression is independently associated with high aggressiveness.

Table 2. Cox proportional hazard ratios. Estimated coefficients of relapse-free survival on breast cancer subjects. Calculated are the corresponding hazard ratio (HR), 95% confidence interval (CI) of the hazard ratio, and *p*-value in uni- and multivariable Cox proportional hazard analysis for the presence of K16-positive CTCs, estrogen receptor (ER) status, progesterone receptor (PR) status, ERBB2 status, and T-stage (reference T1-2).

	Univariable Analysis				Multivariable Analysis			
Covariate	Coefficient (bi)	HR [exp(bi)]	HR 95% CI	*p*-Value	Coefficient (bi)	HR [exp(bi)]	HR 95% CI	*p*-Value
K16+ CTCs	2.061	7.856	(1.677, 36.81)	0.0089	4.864	129.6	(1.6, 1.120)	0.0287
Age	−0.072	0.930	(0.867, 0.998)	0.0432	−0.217	0.804	(0.627, 1.032)	0.0869
ER+	−0.638	0.528	(0.166, 1.680)	0.280	−5.909	0.003	(0.0, 0.951)	0.0481
PR+	−0.150	0.861	(0.275, 2.693)	0.797	3.857	47.3	(0.108, 2.670)	0.2136
ERBB2+	0.078	1.081	(0.283, 4.123)	0.910	−0.435	0.647	(0.030, 14.1)	0.7820
T3-4	0.272	1.312	(0.286, 6.420)	0.737	0.117	1.352	(0.080, 15.9)	0.9308

CTCs = circulating tumor cells, ER = estrogen receptor, PR = progesterone receptor, ERBB2 = Erb-B2 Receptor Tyrosine Kinase 2, HR = hazard ratio.

4. Discussion

Tumor metastasis is the major cause of cancer-related deaths. Therefore, identifying a potential biomarker that can detect micrometastasis early could be helpful in clinical

tumor management. We have previously shown by in silico analysis that K16 mRNA expression upregulation might be associated with a more aggressive course of cancer [3]. Accordingly, we investigated the biological role of K16 in metastatic breast cancer. In this study, K16 was positively correlated to an intermediate mesenchymal phenotype, and this K16 expression was mainly observed in cell lines that had a hybrid phenotype of epithelial and mesenchymal cell features. These results provide further support for the hypothesis that K16 appears to contribute to the hyperplastic epithelial phenotype.

EMT has been recognized as an important mechanism driving cell migration and invasion during tumor growth; however, the process of partial and full reversion of EMT in the stages of metastasis is not well understood yet [29,30]. Although our results indicate a clear association between *KRT16* expression and EMT regulation, further experiments have to be performed to determine whether *KRT16* is upstream or downstream of EMT or whether it can function in a transcription-translation feedback loop. The relative gene and protein expression of epithelial markers such as CDH1 were not changed upon transfection of K16 in MCF7 cells; E-cadherin was relocalized from the cell membrane to the cytoplasm and thereby compromised cell–cell adhesion, as well as acting possibly as transcriptional gene regulator [4,31]. Our findings are similar to the study reported by Huang and colleagues who showed the phenotypic EMT characterization of 43 ovarian cancer cell lines and found that CDH1 expression in the cytoplasm was correlated with an intermediate mesenchymal phenotype [32]. Furthermore, the knockdown of *KRT16* also led to the downregulation of mesenchymal-associated genes but not epithelial-related genes. While the expression level of CDH1 protein was reduced in K16-depleted cells, our results suggest that expression of K16 is correlated with the regulation of the mesenchymal-regulatory genes of EMT, as well as the expression of CDH1 as a crucial epithelial marker. Overall, the functional role of K16 in the EMT appears to be essential to allow epithelial carcinoma cells to undergo multiple morphological and biochemical changes to have higher plasticity and thus be able to migrate. Carcinoma cells could acquire a mesenchymal-like phenotype through EMT before or during intravasation. A reverse mechanism, i.e., MET, may occur at the secondary site following extravasation (Figure 7). A recent study reported evidence that the transcription factor TF-AP2A in EMT-related pathways induced K16 expression in lung adenocarcinoma [33]. This study is consistent with our finding that K16 has a regulatory role in EMT, and thus, more in-depth studies are needed to provide additional information on the role of K16 during EMT and MET.

K16 is one of the intermediate filament members that forms a heterodimer interaction with K6 but is also able to form homodimers [9]. Homomeric K16 formation further dis-plays unique properties that may contribute to morphological changes of epithelial cells during migration [12,34]. In the present study, modifications of the actin microfilament morphology has been observed in MCF7 cells under conditions inducing K16 expression. These changes demonstrate the mechanical effect of K16 in the reorganization of intermediate filaments, which is required for changes in cell shape and locomotion as the first step to induce migration [2]. In this study, K16 was shown to significantly enhance the migration in MCF7 cells by reorganizing the actin cytoskeleton, which is the driving force behind disrupting intercellular adhesion and directional migration. In contrast, K16-depleted cells exhibited impaired cell migration. In accordance with the present results, previous studies on wound-healing have demonstrated a crucial role of K16, along with other partners such as K6 and K17 that have contributed to the enhancement of migration after skin injury [12,35,36]. Taken together, K16 appears to be an important protein in cell migration, which is considered a major event in the metastatic cascade.
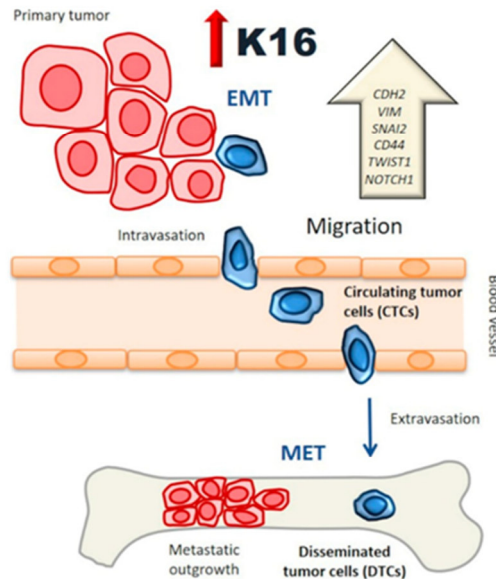
Figure 7. K16 functional model: K16 is involved in a functional transition of polarized epithelial cells to motile by promoting EMT regulator genes while preserving the epithelial phenotype, enhancing the migration of tumor cells to access the bloodstream and survive as a circulating tumor cell (CTCs). These tumor cells have a semi-mesenchymal phenotype with a high capacity to generate metastasis.

CTCs in breast cancer are known to be an indicator for poor outcome [37]. However, the potential of individual CTCs to initiate metastasis is heterogeneous [38]. Here, we identified three CTC subsets that exhibited different expression patterns of keratins (e.g., K16+/C11−, K16 −/C11+, K16+/C11+) during tumor progression in patients. Previous studies have shown that phenotypic and genetic heterogeneity among CTCs are associated with a higher risk for metastasis [39–41]. In our preliminary analysis, K16 expression was associated with higher tumor aggressiveness, and it was mainly associated with epithelial–mesenchymal plasticity. Indeed, a significant relationship has been reported in several studies between the CTCs with high epithelial–mesenchymal plasticity and poor clinical outcomes in different cancer entities [5,42–46]. These CTCs have dynamic plasticity to adapt to the selective microenvironment during their dissemination in distant organs [5,43]. Although our pilot study has only analyzed a small collection of blood samples, a significantly shorter relapse-free survival was observed among patients diagnosed with K16-positive CTCs, which is consistent with our previous in silico analyses [3]. Although additional studies with larger cohorts and the analyses of more confounding clinical factors are required to elucidate the true correlation between K16 expression among CTCs and relapse-free survival, this study can be considered hypothesis generating.

To our knowledge, this has been the first study to assess the role of K16 in metastatic breast cancer. Our results suggest that K16 may represent a novel metastasis-associated protein that acts as a positive regulator of cell motility and promoter of EMT regulator genes in breast cancer. This is an important issue for future research to analyze the exact function of K16 in tumor dissemination and metastasis development by assessing K16 status in disseminating tumor cells (DTCs) and CTCs. Additionally, understanding the biology of K16 would provide valuable mechanistic insights translating into a novel therapeutic opportunity.

5. Conclusions

Our results provide new insights into K16, representing a novel metastasis-associated protein for breast cancer, enhancing cell motility and promoting EMT. K16 expression in CTCs contributes to shorter relapse-free survival in metastatic breast cancer patients. This has been the first report indicating that K16 might be a metastasis-promoting gene in breast cancer.

References

1. Dai, X.; Cheng, H.; Bai, Z.; Li, J. Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. *J. Cancer* **2017**, *8*, 3131–3141. [CrossRef] [PubMed]
2. Pantel, K.; Hayes, D.F. Disseminated breast tumour cells: Biological and clinical meaning. *Nat. Rev. Clin. Oncol.* **2017**, *15*, 129–131. [CrossRef]
3. Joosse, S.A.; Hannemann, J.; Spötter, J.; Bauche, A.; Andreas, A.; Müller, V.; Pantel, K. Changes in Keratin Expression during Metastatic Progression of Breast Cancer: Impact on the Detection of Circulating Tumor Cells. *Clin. Cancer Res.* **2012**, *18*, 993–1003. [CrossRef]
4. Thiery, J.P.; Acloque, H.; Huang, R.Y.-J.; Nieto, M.A. Epithelial-Mesenchymal Transitions in Development and Disease. *Cell* **2009**, *139*, 871–890. [CrossRef]
5. Alix-Panabières, C.; Mader, S.; Pantel, K. Epithelial-mesenchymal plasticity in circulating tumor cells. *J. Mol. Med.* **2016**, *95*, 133–142. [CrossRef]
6. Williams, E.D.; Gao, D.; Redfern, A.; Thompson, E.W. Controversies around epithelial–mesenchymal plasticity in cancer metastasis. *Nat. Rev. Cancer* **2019**, *19*, 716–732. [CrossRef]

7. Werner, S.; Keller, L.; Pantel, K. Epithelial keratins: Biology and implications as diagnostic markers for liquid biopsies. *Mol. Asp. Med.* **2020**, *72*, 100817. [CrossRef]

8. Moll, R.; Divo, M.; Langbein, L. The human keratins: Biology and pathology. *Histochem. Cell Biol.* **2008**, *129*, 705–733. [CrossRef]

9. Trost, A.; Costa, I.; Jakab, M.; Ritter, M.; Haim, M.; Hintner, H.; Bauer, J.W.; Önder, K. K16 is a further new candidate for homotypic intermediate filament protein interactions. *Exp. Dermatol.* **2009**, *19*, e241–e250. [CrossRef]

10. Wawersik, M.J.; Mazzalupo, S.; Nguyen, D.; Coulombe, P.A. Increased Levels of Keratin 16 Alter Epithelialization Potential of Mouse Skin Keratinocytes In Vivo and Ex Vivo. *Mol. Biol. Cell* **2001**, *12*, 3439–3450. [CrossRef]

11. Rotty, J.D.; Coulombe, P.A. A wound-induced keratin inhibits Src activity during keratinocyte migration and tissue repair. *J. Cell Biol.* **2012**, *197*, 381–389. [CrossRef]

12. Wawersik, M.; Coulombe, P.A. Forced Expression of Keratin 16 Alters the Adhesion, Differentiation, and Migration of Mouse Skin Keratinocytes. *Mol. Biol. Cell* **2000**, *11*, 3315–3327. [CrossRef] [PubMed]

13. Neve, R.M.; Chin, K.; Fridlyand, J.; Yeh, J.; Baehner, F.L.; Fevr, T.; Clark, L.; Bayani, N.; Coppe, J.-P.; Tong, F.; et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **2006**, *10*, 515–527. [CrossRef]

14. Young, L.; Sung, J.; Stacey, G.; Masters, J.R. Detection of Mycoplasma in cell cultures. *Nat. Protoc.* **2010**, *5*, 929–934. [CrossRef] [PubMed]

15. System, R.D. a biotechne. Brand-manual.pdf. StemXVivo®EMT Inducing Media Supplement (100X). p. Catalog Number: CCM017. Available online: https://resources.rndsystems.com/pdfs/datasheets/ccm017.pdf?v=20210506&_ga=2.157864314.1993146906.1620297005-771850332.1620297004. (accessed on 30 July 2011).

16. Tang, Y.; Herr, G.; Johnson, W.; Resnik, E.; Aho, J. Induction and Analysis of Epithelial to Mesenchymal Transition. *J. Vis. Exp.* **2013**, *2013*, e50478. [CrossRef]

17. GenScript. User Manual_GenEZ ORF Clone Products.pdf. Available online: https://www.genscript.com/gsfiles/GenEZ%20ORF%20Clone%20Technical%20Manual.pdf. (accessed on 31 July 2018).

18. Li, X.; Yang, H.; Huang, H.; Zhu, T. CELLCOUNTER: Novel Open-Source Software for Counting Cell Migration and Invasion In Vitro. *BioMed Res. Int.* **2014**, *2014*, 863564. [CrossRef] [PubMed]

19. Life technologies, Lipofectamine_RNAiMAX_Reag_protocol.pdf. Available online: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/Lipofectamine_RNAiMAX_Reag_protocol.pdf. (accessed on 30 July 2013).

20. ThermoScientific, MAN0011430_Pierce_BCA_Protein_Asy_UG.pdf. Available online: https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0011430_Pierce_BCA_Protein_Asy_UG.pdf&title=VXNlciBHdWlkZTogUGllcmNlIEJDQSBQcm90ZWluIEFzc2F5IEtpdpdA== (accessed on 30 June 2015).

21. C-DiGit Blot Scanner–manual.pdf. Available online: https://www.licor.com/documents/uhcazdtkjwc3xax5yqtm6xcaiqtfdqva. (accessed on 25 May 2013). NucleoSpin RNA XS Total RNA Isolation User Manual_Rev_09.pdf. Available online: https://www.takarabio.com/documents/User%20Manual/NucleoSpin%20RNA%20XS%20Total%20RNA%20Isolation%20User%20Manual/NucleoSpin%20RNA%20XS%20Total%20RNA%20Isolation%20User%20Manual_Rev_09.pdf. (accessed on 24 July 2015).

22. Livak, K.J.; Schmittgen, T.D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta C_T}$ Method. *Methods* **2001**, *25*, 402–408. [CrossRef]

23. Suarez-Arnedo, A.; Figueroa, F.T.; Clavijo, C.; Arbeláez, P.; Cruz, J.C.; Muñoz-Camargo, C. An image J plugin for the high throughput image analysis of in vitro scratch wound healing assays. *PLoS ONE* **2020**, *15*, e0232565. [CrossRef] [PubMed]

24. LF_biochemlab_BeckmanCoulter_Vi-cell_datasheet_EN.pdf. Available online: http://www.vtpup.cz/common/manual/LF_biochemlab_BeckmanCoulter_Vi-cell_datasheet_EN.pdf. (accessed on 30 May 2010).

25. Riebensahm, C.; Joosse, S.A.; Mohme, M.; Hanssen, A.; Matschke, J.; Goy, Y.; Witzel, I.; Lamszus, K.; Kropidlowski, J.; Petersen, C.; et al. Clonality of circulating tumor cells in breast cancer brain metastasis patients. *Breast Cancer Res.* **2019**, *21*, 1–11. [CrossRef]

26. Warkiani, M.E.; Guan, G.; Luan, K.B.; Lee, W.C.; Bhagat, A.A.; Chaudhuri, P.K.; Tan, D.S.-W.; Lim, W.T.; Lee, S.C.; Chen, P.C.Y.; et al. Slanted spiral microfluidics for the ultra-fast, label-free isolation of circulating tumor cells. *Lab Chip* **2014**, *14*, 128–137. [CrossRef]

27. Joosse, S.A. In-Silico Online (Version 2.3.0). 2020. Available online: http://in-silico.online (accessed on 16 June 2020).

28. Thiery, J.P. Epithelial–mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* **2002**, *2*, 442–454. [CrossRef]

29. Vasaikar, S.V.; Deshmukh, A.P.; Hollander, P.D.; Addanki, S.; Kuburich, N.A.; Kudaravalli, S.; Joseph, R.; Chang, J.T.; Soundararajan, R.; Mani, S.A. EMTome: A resource for pan-cancer analysis of epithelial-mesenchymal transition genes and signatures. *Br. J. Cancer* **2021**, *124*, 259–269. [CrossRef]

30. Du, W.; Liu, X.; Fan, G.; Zhao, X.; Sun, Y.; Wang, T.; Zhao, R.; Wang, G.; Zhao, C.; Zhu, Y.; et al. From cell membrane to the nucleus:
31. An emerging role of E-cadherin in gene transcriptional regulation. *J. Cell. Mol. Med.* **2014**, *18*, 1712–1719. [CrossRef]

32. Huang, R.Y.-J.; Wong, M.K.; Tan, T.Z.; Kuay, K.T.; Ng, A.H.C.; Chung, V.Y.; Chu, Y.-S.; Matsumura, N.; Lai, H.-C.; Lee, Y.F.; et al. An EMT spectrum defines an anoikis-resistant and spheroidogenic intermediate mesenchymal state that is sensitive to e-cadherin restoration by a src-kinase inhibitor, saracatinib (AZD0530). *Cell Death Dis.* **2013**, *4*, e915. [CrossRef]

33. Yuanhua, L.; Pudong, Q.; Wei, Z.; Yuan, W.; Delin, L.; Yan, Z.; Geyu, L.; Bo, S. TFAP2A Induced KRT16 as an Oncogene in Lung Adenocarcinoma via EMT. *Int. J. Biol. Sci.* **2019**, *15*, 1419–1428. [CrossRef]

34. Mazzalupo, S.; Wong, P.; Martin, P.; Coulombe, P.A. Role for keratins 6 and 17 during wound closure in embryonic mouse skin. *Dev. Dyn.* **2003**, *226*, 356–365. [CrossRef]

35. Zhang, X.; Yin, M.; Zhang, L.-J. Keratin 6, 16 and 17—Critical Barrier Alarmin Molecules in Skin Wounds and Psoriasis. *Cells* **2019**, *8*, 807. [CrossRef]

36. Tomikawa, K.; Yamamoto, T.; Shiomi, N.; Shimoe, M.; Hongo, S.; Yamashiro, K.; Yamaguchi, T.; Maeda, H.; Takashiba, S. Smad2 Decelerates Re-epithelialization during Gingival Wound Healing. *J. Dent. Res.* **2012**, *91*, 764–770. [CrossRef]

37. Pantel, K.; Alix-Panabières, C. Liquid biopsy and minimal residual disease—Latest advances and implications for cure. *Nat. Rev. Clin. Oncol.* **2019**, *16*, 409–424. [CrossRef]

38. Keller, L.; Pantel, K. Unravelling tumour heterogeneity by single-cell profiling of circulating tumour cells. *Nat. Rev. Cancer* **2019**, *19*, 553–567. [CrossRef]

39. Massagué, J.; Obenauf, A. Metastatic colonization by circulating tumour cells. *Nat. Cell Biol.* **2016**, *529*, 298–306. [CrossRef]

40. Pantel, K.; Brakenhoff, R.H. Dissecting the metastatic cascade. *Nat. Rev. Cancer* **2004**, *4*, 448–456. [CrossRef]

41. Agnoletto, C.; Corrà, F.; Minotti, L.; Baldassari, F.; Crudele, F.; Cook, W.J.J.; Di Leva, G.; Pio d'Adamo, A.; Gasparini, P. Heterogeneity in Circulating Tumor Cells: The Relevance of the Stem-Cell Subset. *Cancers* **2019**, *11*, 483. [CrossRef]

42. Lecharpentier, A.; Vielh, P.; Perez-Moreno, P.; Planchard, D.; Soria, J.C.; Farace, F. Detection of circulating tumour cells with a hybrid (epithelial/mesenchymal) phenotype in patients with metastatic non-small cell lung cancer. *Br. J. Cancer* **2011**, *105*, 1338–1341. [CrossRef]

43. Bednarz-Knoll, N.; Alix-Panabières, C.; Pantel, K. Plasticity of disseminating cancer cells in patients with epithelial malignancies. *Cancer Metastasis Rev.* **2012**, *31*, 673–687. [CrossRef]

44. Nel, I.; Baba, H.A.; Ertle, J.; Weber, F.; Sitek, B.; Eisenacher, M.; Meyer, H.E.; Schlaak, J.F.; Hoffmann, A.-C. Individual Profiling of Circulating Tumor Cell Composition and Therapeutic Outcome in Patients with Hepatocellular Carcinoma. *Transl. Oncol.* **2013**, *6*, 420–428. [CrossRef]

45. Satelli, A.; Mitra, A.; Brownlee, Z.; Xia, X.; Bellister, S.; Overman, M.J.; Kopetz, S.; Ellis, L.M.; Meng, Q.H.; Li, S. Epithelial–Mesenchymal Transitioned Circulating Tumor Cells Capture for Detecting Tumor Progression. *Clin. Cancer Res.* **2015**, *21*, 899–906. [CrossRef]

46. Armstrong, A.J.; Marengo, M.S.; Oltean, S.; Kemeny, G.; Bitting, R.; Turnbull, J.D.; Herold, C.I.; Marcom, P.K.; George, D.J.; Garcia-Blanco, M.A. Circulating Tumor Cells from Patients with Advanced Prostate and Breast Cancer Display Both Epithelial and Mesenchymal Markers. *Mol. Cancer Res.* **2011**, *9*, 997–1007. [CrossRef]

5.2. *BRCA1* promoter hypermethylation on circulating tumor DNA correlates with improved survival of patients with ovarian cancer

The development of highly sensitive and specific assays has made minimally invasive liquid biopsy in oncology possible. Especially in the clinical management of ovarian cancer patients, new assays for real-time monitoring of therapy response are direly required. In the study presented here, we developed and employed a liquid biopsy-based assay to monitor *BRCA1* promoter hypermethylation during the treatment of ovarian cancer patients. Our results suggest that hypermethylation (HR, 0.5614; 95%CI, 0.3774-0.8352; p=0.0044) and methylation conversion but in a lesser extent (HR, 0.6004; 95% CI, 0.3738 - 0.9644; p=0.0349) are independently correlated to longer progression-free survival. We present a highly sensitive and specific liquid biopsy assay that can assess *BRCA1* gene promoter hypermethylation on circulating cell-free DNA from blood plasma. Our assay provides predictive information in ovarian cancer and might be used to enrich high-risk patients in clinical trials.

BRCA1 promoter hypermethylation on circulating tumor DNA correlates with improved survival of patients with ovarian cancer

Maha Elazezy[1] (iD), Katharina Prieske[2,3], Lan Kluwe[4], Leticia Oliveira-Ferrer[2] (iD), Sven Peine[5], Volkmar Müller[2], Linn Woelber[2], Barbara Schmalfeldt[2], Klaus Pantel[1] and Simon A. Joosse[1,3] (iD)

1 Department of Tumor Biology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
2 Department of Gynecology and Gynecologic Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
3 Mildred Scheel Cancer Career Center HaTriCS4, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
4 Department of Oral and Maxillofacial Surgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
5 Department of Transfusion Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Methylation of the *BRCA1* promoter is an epigenetic gene expression regula tor and is frequently observed in ovarian cancer; however, conversion o methylation status is thought to drive disease recurrence. Therefore, longitu dinal monitoring of methylation status by liquid biopsy in cell-free DN/ may be a predictive marker. In total, 135 plasma samples were collected fron 69 ovarian cancer patients before and during systemic treatment. Our liqui biopsy assay could detect down to a single molecule of methylated DNA in high background of normal DNA (0.03%) with perfect specificity in contro samples. We found that 60% of the cancer patients exhibited *BRCA1* pro moter hypermethylation at one point, although 24% lost hypermethylatio during treatment. Multivariate survival analyses indicate that relapses ar independent events and that hypermethylation and methylation conversio are independently correlated to longer relapse-free survival. We present highly sensitive and specific methylation-specific quantitative PCR-based liq uid biopsy assay. *BRCA1* promoter hypermethylation is frequently found ii ovarian cancer and is often reversed upon recurrence, indicating the selectio of therapy-resistant clones and unfavorable clinical outcome.

1. Introduction

Germline and somatic mutations in *BRCA1* have been identified in approximately one-third of ovarian carcinomas, and their presence is highly predictive of primary platinum and PARP (Poly-ADP-Ribose-Polymerase) inhibitors sensitivity and favorable progression-free and overall survival [1–4]. Hyperme thylation of the *BRCA1* promoter leads to downregula tion of *BRCA1* mRNA expression [5,6], resulting in defective homologous recombination characterized by typical chromosomal aberrations seen in *BRCA1*

mutation carriers [6,7]. Although hypermethylation o the *BRCA1* promoter has been shown in xenografts to predict response to PARP inhibitors as well [8], data o patient survival are still unclear [9]. Recently, we showe that the *BRCA1* promoter is frequently hypermethy lated in ovarian cancer tissue, but the methylation statu is often lost in recurrent disease, suggesting a potentia resistance mechanism either through therapy-induce cancer evolution or by clonal selection [10–12]. Thereby the detection and monitoring of *BRCA1* promote hypermethylation may have an important impact on th clinical management of ovarian cancer patients withou

BRCA1 mutation [8]. Precision medicine in oncology may be achieved through the diagnostic method 'liquid biopsy' as has been shown in several cancer entities [13–16]. This method utilizes the detection of biomarkers in blood or other body liquids for prognostic and predictive purposes and has several advances over using tissue alone [15]. Circulating tumor DNA (ctDNA) are cell-free DNA (cfDNA) fragments released into the circulation by tumor cells and can provide direct information about the methylomic make-up of the tumor currently present in the patient [17,18]. The purpose of this study was to develop a liquid biopsy assay that could determine the methylation status of the *BRCA1* promoter to be able to monitor hypermethylation of the *BRCA1* promoter and investigate its clinical significance as a predictive biomarker in ovarian cancer patients. We considered two models of cancer progression leading to therapy resistance: cancer evolution driven by therapy pressure and, secondly, selection and survival of tumor clones originating from the primary tumor (illustrated in Fig. S1).

2. Materials and methods

2.1. Sample collection

In this prospective study, 69 ovarian cancer patients treated during 2015–2020 were included. Selection criteria were primary or recurrent, high-grade serous ovarian cancer with platinum-based first-line therapy (carboplatin, $n = 68$; cisplatin, $n = 1$) (Table S1). Blood sampling was performed before treatment, at relapse before therapy change, and/or during the course of therapy. Additionally, 69 healthy, age-matched women were included as controls. This study was approved by the local ethical board (ethical approval number: PV5392) in accordance with the declaration of Helsinki, all participants enrolled into this study gave written informed consent. The patients' demographic statistics are described in Table 1.

2.2. Cell-free DNA isolation and bisulfite conversion

Peripheral blood of all women was collected in EDTA containing tubes and processed within 1 h. Blood was centrifuged at 360 *g* for 20 min, and plasma was centrifuged again at 5087 *g* for 10 min. cfDNA was isolated and treated with bisulfite by the full automated InviGenius® Plus instrument with the InviMag® Free Circulating DNA Kit/IG (cat.no. 2439320400; Invitek Molecular, Berlin, Germany) and InviMag® Bisulfite

Table 1. Demographic statistics. *P*-values were calculated using the *G*-test with Williams' correction for count data and ANOVA for continuous data. The study cohort was divided by the methylation status of the *BRCA1* promoter of the patients: hypermethylated detected in all blood samples (ME1), no methylated detected in any of the blood samples (M0), and methylation status changed during the course of treatment from positive to negative (MEc).

	ME1	ME0	MEc	Total	*P*-value
Mean age (years)	56.7	60.8	58.0	58.3	0.60
FIGO stage					
I–IIIB	2	2	3	7	0.39
IIIC	19	16	4	37	
IV	9	5	3	17	
Grade					
G2	4	2	1	7	0.92
G3	30	22	9	32	
T-stage					
T1	1	2	1	4	0.41
T2	1	3	0	4	
T3	28	15	7	50	
N-stage					
N0	6	4	1	11	0.79
N1	15	10	4	29	
Nx	4	3	0	7	
Residual tumor					
No (macroscopic complete resection)	18	12	4	34	0.93
Yes	15	8	3	26	
Lymphatic invasion					
L0	6	7	1	14	0.29
L1	22	11	7	40	
Venous invasion					
V0	24	15	7	37	0.99
V1	3	2	1	6	
PARP inhibitor treatment					
Yes	6	5	4	15	0.40
No	25	22	6	53	
Germline *BRCA1*					
Mutated	5	6	2	13	0.64
Wild-type/Unknown	29	18	8	55	

Conversion Kit/IG (cat.no. 3030200100; Invitek Molecular). cfDNA quantification and fragment size distribution were assessed using the Agilent 2200 TapeStation High Sensitivity D5000 and Qubit® 2.0 Fluorometer dsDNA HS Assay Kits (Thermo Fisher Scientific, Eugene, OR, USA) according to the manufacturers' instructions.

2.3. Methylation-specific, quantitative real-time PCR

BRCA1 promoter methylation status in cfDNA was assessed after bisulfite conversion using methylation-specific primers as before [10] with slight adjustments to the PCR protocol. The sequences of the primers for amplifying the 1543- to 1617-bp region (fragment

length, 75 bp) of the *BRCA1* (GenBank U37574.1) promoter in case of methylation were 5′-TCGTGGT AACGGAAAAGCGC-3′ (sense) and 5′-AAATCTCA ACGAACTCACGCCG-3′ (antisense). The primers for amplifying the 1536–1621 bp region (fragment size), 86 bp of the wild-type *BRCA1* promoter were 5′-T TGGTTTTTGTGGTAATGGAAAAGTGT-3′ (sense) and 5′-CAAAAAATCTCAACAAACTCACACCA-3′ (antisense). Methylation-specific, quantitative real-time PCR (MS-qPCR) was performed in 15 μL reaction volume containing $1\times$ PCR buffer (1.5 mM $MgCl_2$, 10 mM Tris/HCl of pH 8.3, 50 mM KCl), 100 μM of each dNTP, 0.2 μM of each primer, $0.1\times$ SYBR green, 0.25 μg BSA, and 0.5 U JumpStart™ Taq DNA Polymerase (cat no. D9307; Merck, Darmstadt, Germany). The MS-qPCR reaction was applied in a CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad Laboratories GmbH, Feldkirchen, Germany). The PCR conditions were as follow: initial denaturation at 94 °C for 1 min, followed by 40 amplification cycles of 94 °C for 30 s, 63 °C for 30 s, and 72 °C for 30 s, a final extension at 72 °C for 1 min, and a melting curve of 65.0–95.0 °C with increments of 0.5 °C every 5 s.

2.4. Liquid biopsy assay establishment

The sensitivity and specificity of our MS-qPCR-based liquid biopsy assay were assessed by testing serial dilutions of methylated reference DNA 'Human HCT116 DKO Methylated DNA' from Zymo Research (Freiburg, Germany) and unmethylated reference DNA isolated from a pool of healthy donors (cat. No. G1521; Promega, Walldorf, Germany). The DNA dilution mix ranged from 100 to 1 genome copies in a background of wild-type unmethylated DNA equivalent to 3000 genome copies. Methylated and unmethylated reference DNA was fragmented using Bioruptor® Plus sonication system for 20 cycles (20 s on, 30 s off) to a length similar to that of patients' cfDNA, followed by bisulfite treatment. Quantitative and qualitative detection of the methylation status was determined via melting curve analysis of the MS-qPCR amplified products, as well as from agarose gel electrophorese for confirmation.

2.5. cfDNA sequencing

The MS-qPCR products were separated by electrophoresis on a 3% agarose gel stained with ethidium bromide. The corresponding bands were cut out of the gel and purified using NucleoSpin Gel and PCR Clean-up for QIAcube (cat. No. 15116456; Macherey-Nagel, Dueren, Germany). Next, Sanger sequencing was performed to confirm the methylation status of the *BRCA1* promoter

as before [10]. Ideally, 15 ng of the purified PCR product was used for sequencing with the same reverse primer used for the MS-qPCRs. PCR sequencing was performed using Big Dye™ Terminator v1.1 Cycle Sequencing Kit (cat. No. 4337451; Applied Biosystems, Foster City, CA, USA) and analyzed by 3130 Genetic Analyzer (Thermo Fisher). CpG islands of bisulfite sequences were aligned and analyzed using QUMA [19].

2.6. Statistical analysis

Statistical analyses were performed using R (R Foundation for Statistical Computing, version 3.6.3, Vienna, Austria) and In-Silico Online, version 2.1.2 [20]. Because ovarian cancer patients can relapse multiple times during the course of their disease, analysis based only on the first event time cannot be used to examine the effect of the risk factors on the number of recurrences over time [21]. Therefore, progression-free survival analyses were performed using Kaplan–Meier curves for recurrent events using two multivariate models. The first model tests the correlation between methylation status and the gap times between successive progression events. This model assumes that the occurrence of an event is related to the previous event, however, this assumption only holds true if a relapse originates from the preceding tumor (illustrated in Fig. S1). The second model tests the correlation between methylation status and the time from initial diagnosis to each event of progression independently for each patient. This model assumes that the occurrence of subsequent events are not correlated, and this assumption only holds true if all relapses originate from the primary tumor (illustrated in Fig. S1). The endpoints were progression and cancer-related death according to REMARK [22]. The clinical variables (residual tumor, FIGO stage, grade, T-stage, N-stage, or treatment with PARP inhibitors) were compared to the methylation status of *BRCA1* promoter of ovarian cancer patients using ANOVA for continuous data and *P*-values were calculated using the *G*-test with Williams' correction for count data. The nonsignificant clinical variables (FIGO stage, T-stage, N-stage, and grade) were excluded from the multivariate analysis. The power of sample size was analyzed by PASS, version 20.0.2 (NCSS LLC, East Kaysville, UT, USA).

3. Results

3.1. Liquid biopsy assay establishment

Sensitivity of detection of *BRCA1* gene promoter hypermethylation was tested in a dilution series of

hypermethylated DNA mixed with wild-type, non-methylated DNA with a total amount of DNA equivalent to 3000 copies of gDNA (the average amount of DNA per mL plasma found in healthy individuals [23]). The dilution series of hypermethylated DNA consisted of 100% (3000 gDNA copies), 3.33% (100 gDNA copies), 1.67% (50 gDNA copies), 0.33% (10 gDNA copies), 0.17% (5 gDNA copies), and 0.03% (1 gDNA copy), and was measured in triplicate twice. Melting curve and gel electrophoresis analyses of the amplified products showed highly sensitive detection of all diluted hypermethylated DNA samples (Fig. 1A,B). The methylation status of MS-qPCR products was confirmed using Sanger sequencing (Fig. 1C,D). These results demonstrated the ability to detect hypermethylated *BRCA1* gene promoter, down to a single molecule, in a background of 99.97% normal DNA.

3.2. Study cohort and sample material

In total, 135 plasma samples from 69 advanced-stage ovarian cancer patients were obtained; 111 multiple longitudinal blood samples were collected from 41/69 patients during the course of disease. The patients' mean age was 58.3 years (range: 31–89); the healthy donors' mean age was 56.2 years (range: 30–73). The median follow-up was 39.3 months [95% confidence interval (CI): 25.0–49.6 months], starting from the time point of first diagnosis. The cfDNA fragment size in ovarian cancer patients was on average 166 bp ($s = 17.2$) and 338 bp ($s = 48.3$; Fig. 2A). The median concentration of total cfDNA obtained from all patients' blood samples at all time points was 306 ng·mL^{-1} plasma (range: 28.4–6750), whereas the median of cfDNA from a healthy donor was 192 ng·mL^{-1} plasma (range: 100–742; Fig. 2B). cfDNA concentration of ovarian cancer patients was significantly higher as compared to healthy controls ($P = 0.0002$, Wilcoxon rank sum test with continuity correction). No significant differences were observed between the median cfDNA concentrations of 282, 348, and 337 ng·mL^{-1} plasma of before, during, and after systemic therapy, respectively ($P = 0.535$, ANOVA; Fig. 2C). Out of 33 patients who were tested for germline mutations in the *BRCA1* gene, 12 tested positive and were therefore also analyzed separately.

3.3. *BRCA1* promoter hypermethylation in cfDNA

BRCA1 promoter hypermethylation was detected during the course of the disease until the end of follow-up in 46% (31/68) of patients, no methylation could be detected in 40% (27/68), and 15% (10/68) of patients

converted from having hypermethylation to no methylation until the end of follow-up; data of one patient could not be obtained (1/69). Two patients started with a negative methylation status, which was positive in subsequent cfDNA samples, and these patients were included into the methylation positive group (2/31). The methylation status of patients of whom only one plasma sample was obtained was assumed to remain stable, and the patients were grouped into either the methylation positive or negative group. The median of cfDNA concentrations in the methylation positive and negative ovarian cancer patients was 316 and 344 ng·mL^{-1} plasma (Fig. 2D), respectively, indicating that the lack of detection of hypermethylation was not correlated to cfDNA concentrations ($P = 0.612$, Wilcoxon rank sum test with continuity correction). Methylation status of the ovarian cancer patients (positive, negative, or converted) was not correlated to any of the recorded clinical variables: residual tumor, FIGO, grade, T-stage, N-stage, or treatment with PARP inhibitors (Table 1). Surprisingly, hypermethylation of the *BRCA1* gene promoter was also detected in 5/12 (41.7%) patients with germline *BRCA1* mutations. A nonsignificant difference ($P = 0.376$, Welch's two sample *t*-test) was detected in the median of cfDNA levels between patients who were carriers of germline *BRCA1* mutations (267 ng·mL^{-1}) and patients who were negative for germline *BRCA1* mutations (213 ng·mL^{-1}). In order to verify our results, all ($n = 62$) cfDNA samples showing signs of hypermethylation were processed by Sanger sequencing. In the MS-qPCR amplified DNA fragment, five CpG islands are present of which methylation was detected in 97%, 100%, 95%, 89%, and 95% of the sequenced samples, respectively. In each of the samples, methylation of 4 or 5 CpG islands could be confirmed, and one sample showed methylation of three CpG islands only. In all 69 healthy individuals, *BRCA1* promoter hypermethylation was not detected (0/69).

For illustrative purpose, Fig. 3 depicts the course of disease of two patients, including the concentration of the tumor marker CA-125 measured for routine diagnostics, the systemic treatments given, the disease status, and the *BRCA1* gene promoter methylation status. Both patients initially exhibited ovarian cancer with a hypermethylated *BRCA1* gene promoter, however, the status converted to presumably functional *BRCA1* during the course of therapy.

3.4. Survival analyses

Because the treatment regiments of all ovarian cancer patients were relatively heterogeneous and because
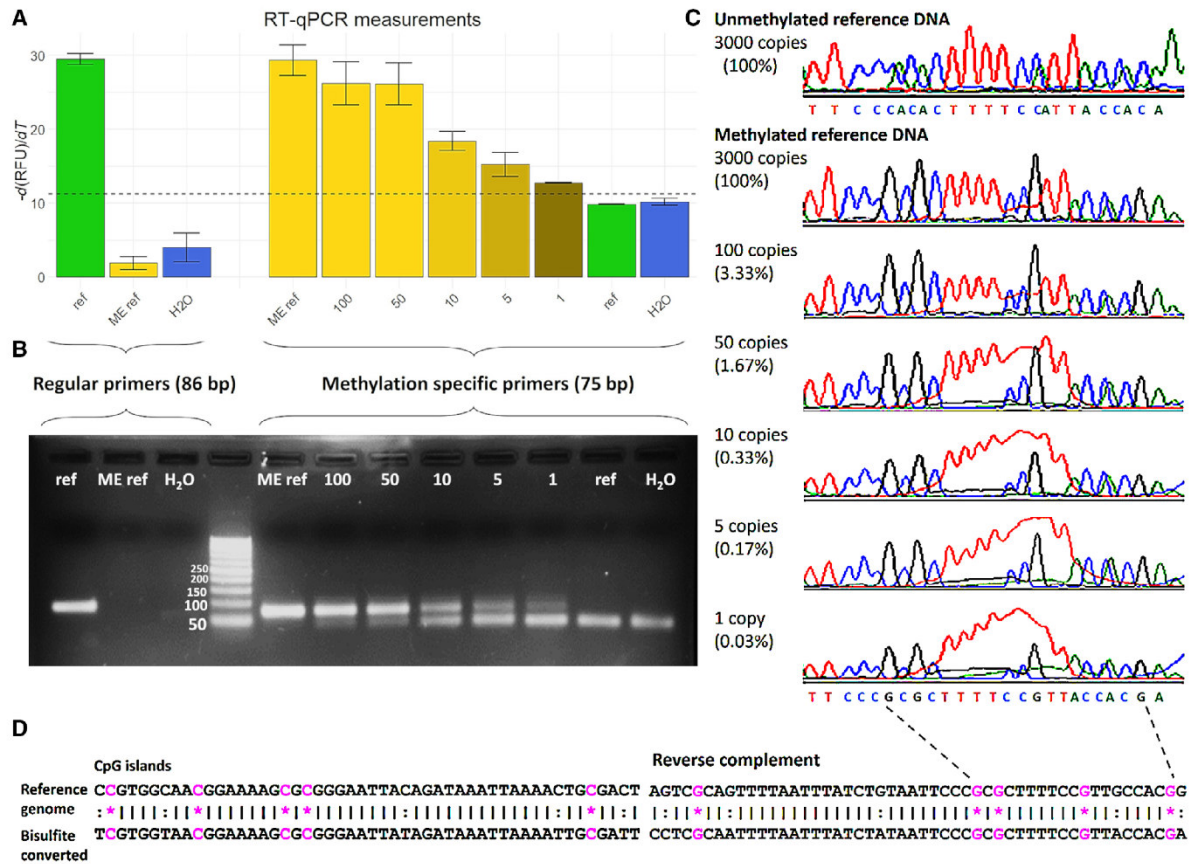
Fig. 1. Measurement of methylated *BRCA1* promoter by MS-qPCR. Two sets of primers were used to detect unmethylated *BRCA1* promoter DNA (regular primers, fragment length 86 bp) and methylated *BRCA1* promoter DNA (methylation-specific primers, fragment length 75 bp). Sensitivity and specificity of both primer sets and PCR protocol were assessed using unmethylated human reference DNA (ref), methylated human reference DNA (ME ref), water (H_2O), and a dilution series containing 100, 50, 10, 5, and 1 copy of methylated human reference DNA in a background of unmethylated human reference DNA with a total amount of 3000 copies of gDNA. (A) Average $-d(RFU)/dT$ values measured in RT-qPCR, error bars represent standard deviations ($n = 3$). (B) Gel electrophoreses photo of PCR amplified products. (C) Sanger sequencing results of the PCR amplified products. (D) Genome sequence of investigated CpG islands and sequence after bisulfite conversion and PCR amplification in case of methylation. Sequences outputted by Sanger sequencing are reversed complemented. *, methylated CpG site (pink); :, non-CpG converted cytosine to thymine; |, matching base.

ovarian cancer patients can suffer from multiple relapses (see Fig. 3 as example), the survival analyses were performed using two multivariate models testing two hypotheses of how a tumor develops therapy resistance: through therapy-induced evolution or by selection (illustrated in Fig. S1). In total, 239 events of progression were recorded during follow-up with a median of three events per patients (range 1–16), with no difference in number of events between the patients with or without mutated *BRCA1* ($P = 0.879$, Wilcoxon rank sum test) and no difference between patients with methylation, without methylation, or a change in methylation status of the *BRCA1* promoter ($P = 0.309$, Kruskal–Wallis rank sum test), suggesting

no bias in the survival analyses due to the number of events in single cases.

3.4.1. Model 1: Dependent survival model—Therapy-induced evolution

The first multivariate survival model was applied to test the correlation between the gap time between successive events and *BRCA1* mutation/promoter methylation status, assuming the dependency of subsequent relapses (illustrated in Fig. S1). The median times between events for ovarian cancer patients with *BRCA1* mutations, hypermethylation of the *BRCA1* promoter, and a negative methylation status of the

51
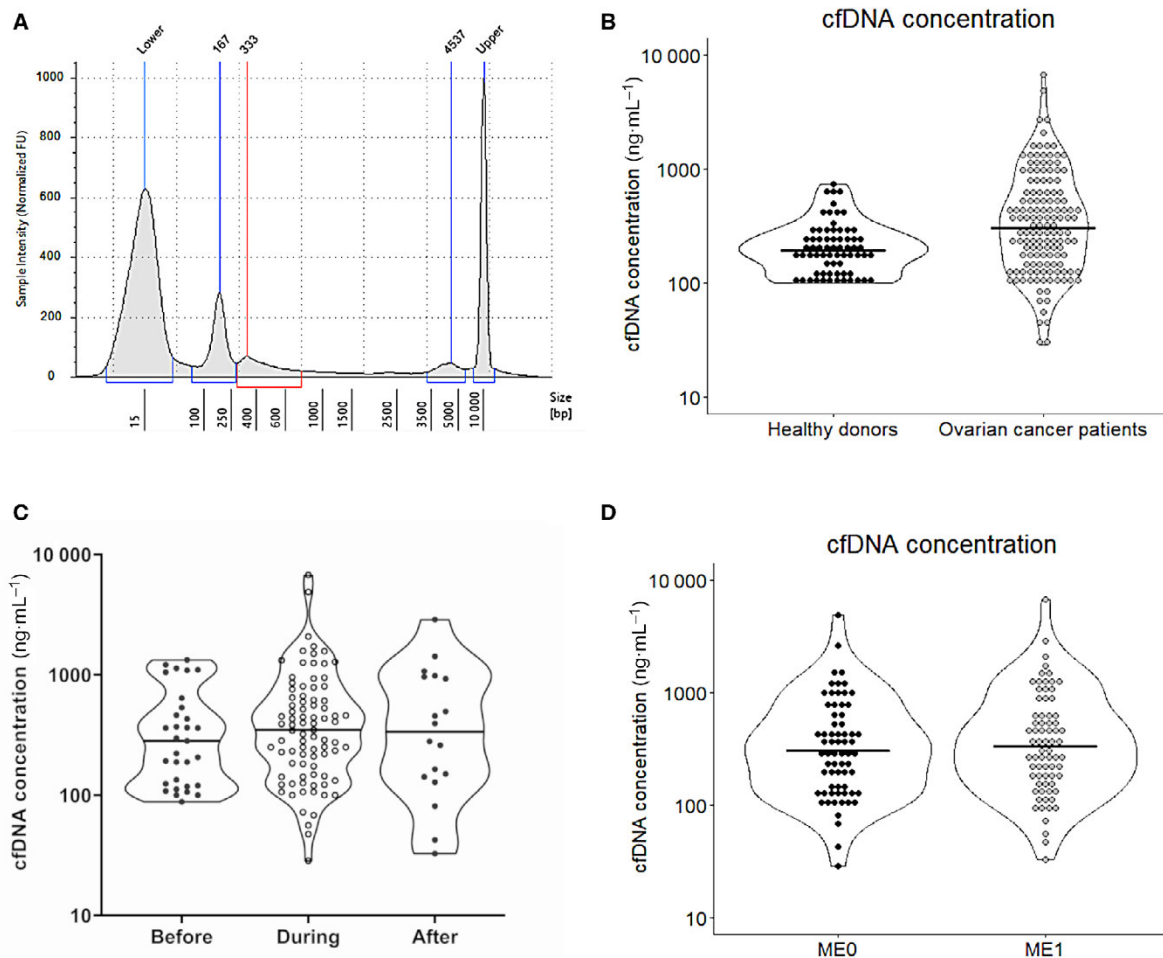
Fig. 2. Isolated cfDNA. (A) Fragment size distribution of cfDNA obtained from a single ovarian cancer case. Peaks marked *Lower* and *Upper* signify the size standards used by the TapeStation to estimate the other three peaks. The two peaks labeled 167 and 333 bp represent DNA originating from apoptotic and necrotic cells [16], whereas the 4537 bp peak represents the gDNA from lysed leukocytes after blood sampling. (B) Violin plot showing the distribution of the cfDNA concentrations in ng·mL^{-1} plasma from blood samples obtained from healthy donors ($n = 69$) and ovarian cancer patients ($n = 135$). (C) Violin plot showing the cfDNA concentrations distribution in blood taken before ($n = 31$), during ($n = 85$), and after ($n = 18$) the systemic therapy. (D) Violin plot depicting the distribution of the cfDNA concentrations of ovarian cancer patients with (ME1; $n = 61$) or without (ME0 $n = 71$) hypermethylated BRCA1 promoter.

BRCA1 promoter were 10.9 (95% CI: 7.6–15.9), 10.6 (95% CI: 8.0–12.5), and 12.0 (95% CI: 8.2–18.4) months, respectively ($P = 0.84$, log rank test; Fig. 4A). Excluding the cases with *BRCA1* mutations and separating the methylation positive group into cases with stable positive methylation status and those showing conversion, the median times between events were 10.0 (95% CI: 7.4–11.9) and 13.8 (95% CI: 9.2–19) months, respectively. There was no significant difference between the median gap times between successive events of patients with stable negative, stable positive, or conversion of methylation status ($P = 0.84$, log rank test; Fig. 4B). Progression-free survival to the

first progression only and overall survival were not correlated to methylation status due to the relatively low number of events that require to achieves 80% power at a significant level $P = 0.15$, HR = 0.8.

3.4.2. Model 2: Independent survival model—Therapy-induced selection

The second multivariate survival model was applied to test the correlation between time from initial diagnosis to each subsequent relapse and *BRCA1* mutation/promoter methylation status, assuming the independency of subsequent relapses and all originating from the

Fig. 3. Longitudinal assessment of *BRCA1* promoter methylation status in two cases. Methylation status of the *BRCA1* promoter was assessed in patients d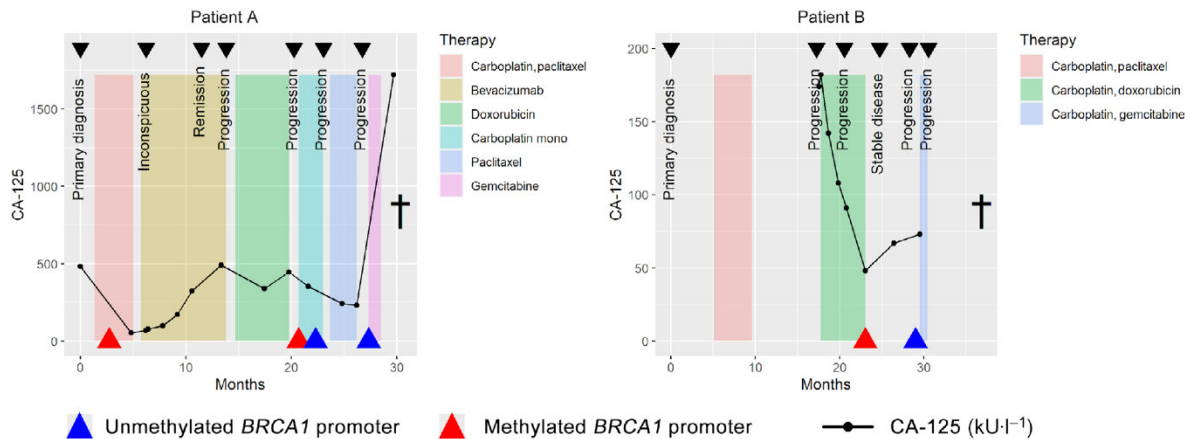uring the course of disease until death (†). Clinically recorded data were the systemic therapy (here, doxorubicin is 'PEG-liposomal doxorubicin'), CA-125 (kU·L^{-1}), and results from computed tomography (CT) scan-based staging. Two examples of a methylation conversion (left and right panel).

primary tumor (illustrated in Fig. S1). Ovarian cancer patients with methylated *BRCA1* promoter detected in cfDNA (median: 57 months; 95% CI: 44.1–72.3) had a comparable survival to ovarian cancer patients with germline *BRCA1* mutations (median: 62 months; 95% CI: 51.0–86.2), but a significantly longer survival than patients with unmethylated *BRCA1* promoter (median 37.5 months; 95% CI: 28.0–52.3; $P = 0.0019$, log rank test; Fig. 4C). The difference between survival of the patients with and without hypermethylated *BRCA1* promoter became more noticeable after excluding the cases with *BRCA1* mutations from the analysis ($P = 0.0014$, log rank test). Interestingly, patients from the methylation positive group who eventually converted to a negative methylation status had a shorter median survival (median: 50.3 months; 95% CI: 30.4–70.0) than the patients who had a stable positive methylation status throughout the whole course of disease (median: 63.8 months; 95% CI: 44.1–81.4), but a better survival than patients with unmethylated *BRCA1* promoter ($P = 0.0011$, log rank test; Fig. 4D). After removal of the nonsignificant clinical variables (FIGO stage, T-stage, N-stage, and grade), multivariable analyses showed that methylation status of the *BRCA1* promoter was an independent predictor of survival and that ovarian cancer patients with methylated *BRCA1* promoter had a significant lower risk for disease-related progression (HR: 0.5614; 95% CI: 0.3774–0.8352; $P = 0.0044$, Cox proportional hazard ratio) as well as patients that showed conversion of the methylation status (HR: 0.6004; 95% CI: 0.3738–0.9644; $P = 0.0349$) as compared to patients with unmethylated *BRCA1* promoter (Table 2). Residual

tumor after surgery was as well correlated with a worse survival as compared to who had no residual tumor.

4. Discussion

BRCA1 promoter hypermethylation was first shown in sporadic breast and ovarian tumors more than 20 years ago [5,24]. Since then, the development of highly sensitive and specific assays have made minimally invasive, liquid biopsy in oncology possible. Especially in the clinical management of ovarian cancer patients, new assays for real-time monitoring of therapy response are direly required.

With our liquid biopsy assay, we could show a high sensitivity by detecting down to a single molecule of DNA, minimizing the possibility of failing to detect the tumor's true methylation status. Five CpG sites were investigated to confirm the enrichment of methylated DNA sites, which previously have been documented to be strongly correlated with very low *BRCA1* expression in breast cancer cell lines [5]. Nevertheless, extremely low ctDNA concentrations and the complete absence of tumor DNA in the obtained blood sample may result in a false negative result. Although the latter cannot be excluded and is most likely the case for the two patients in which hypermethylation was detected after a negative plasma sample, cfDNA concentrations were not correlated with methylation status overall, whereas low concentrations of cfDNA have been shown to be associated with better survival [25,26]. In addition, cfDNA levels have previously shown to be raised in patients with *BRCA1*
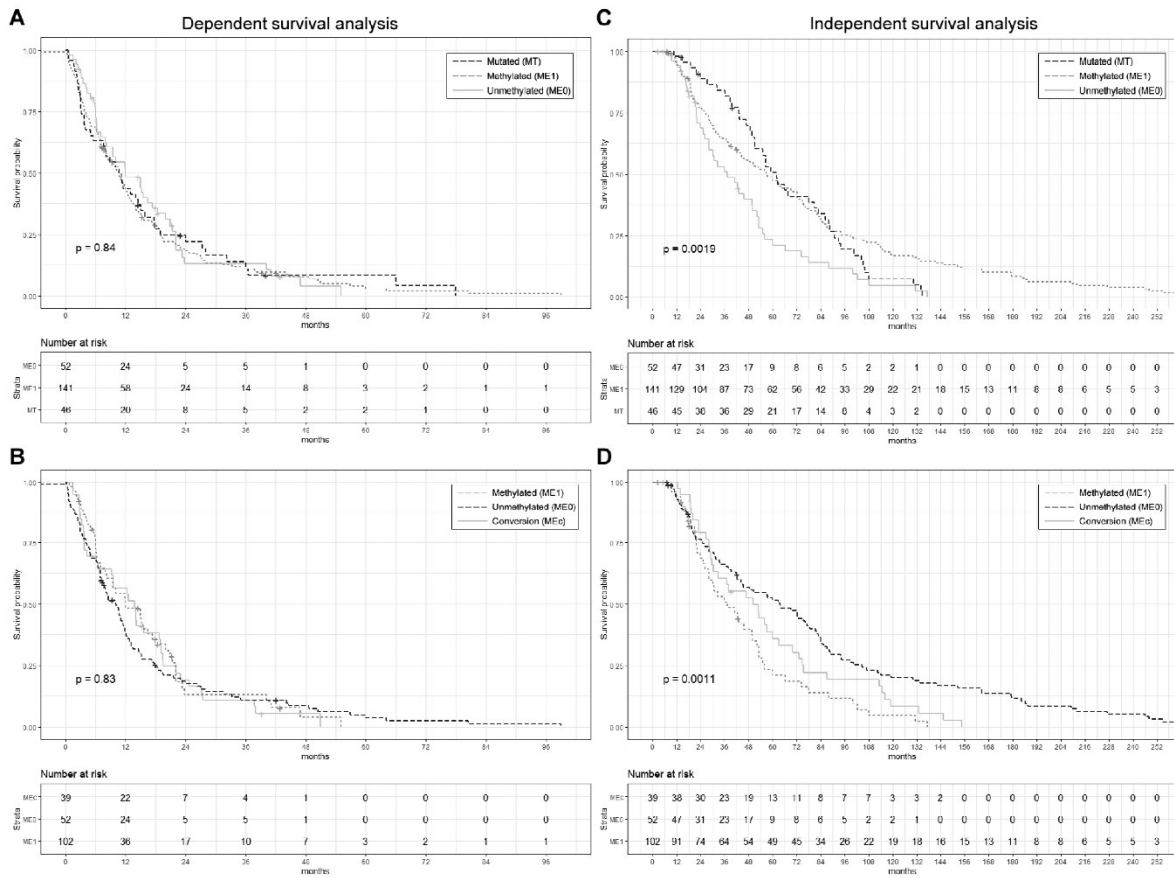
Fig. 4. Survival analyses. Kaplan–Meier curves using a multivariate-dependent survival model on the gap time between successive events of patients showing *BRCA1* gene mutations (A), hypermethylated *BRCA1* promoter (A, B), unmethylated (A, B), and conversion (B). Kaplan-Meier curves using a multivariate independent survival model on the event time between primary tumor diagnosis and every subsequent progression of disease of patients showing *BRCA1* gene mutations (C), hypermethylated *BRCA1* promoter (C, D), unmethylated (C, D), and conversion (D). *P*-values were calculated using the log rank test. +, censored.

Table 2. Cox proportional hazard ratios. Estimated coefficient of hazard ratios (HR) correlated to hypermethylated *BRCA1* promoter and methylation status conversion (reference: unmethylated) and residual tumor after surgery (reference: none), along with 95% CI and *P*-value. Cases with germline *BRCA1* mutations ($n = 13$) were excluded from this analysis.

Covariate	Coefficient (b_i)	HR [exp(b_i)]	HR 95% CI	*P*-value
Methylation positive	−0.5102	0.6004	0.3738–0.9644	0.0349
Methylation conversion	−0.5772	0.5614	0.3774–0.8352	0.0044
Tumor rest	1.0895	2.9727	2.0717–4.2656	<0.0001

mutation carriers irrespective of the disease [27]. Furthermore, our data show that a negative methylation status is correlated with a poor progression-free survival. Taken together, the majority of plasma samples contained enough tumor DNA for a reliable measurement; however, increased blood volumes could be considered. Due to the essential functions of *BRCA1*, hypermethylation of its promoter is not expected to occur in healthy tissue and has thus far not been reported. Because cfDNA is a mixture of DNA originating from practically all regenerating tissues in the body, our data on a group of elderly, cancer-free, healthy women strongly suggest that hypermethylation of the *BRCA1* promoter only occurs in (pre-) cancerous cells. Although our data are convincing, a limitation of this study is the lack of positive control for ctDNA presence and should be considered in future studies. Such marker could be *TP53*, the most common mutated gene in high-grade serous ovarian cancer [1], however, because there are no hotspot mutations

in *TP53* or any other known gene in ovarian cancer, a control marker for liquid biopsy may be challenging.

Previously, we showed that a conversion of hypermethylated to wild-type (unmethylated) *BRCA1* promoter frequently takes place in high-grade serous ovarian cancer upon recurrence [10]. In the current study, multiple blood samples were obtained from 60% of patients during the course of treatment and a conversion of methylation status could be detected in 24% of methylation positive patients. We hypothesize that methylation conversion is either an active mechanism of resistance as the tumor reactivated *BRCA1* in a response of therapy-induced DNA double-strand breaks (e.g., platinum-based therapy). Alternatively, we postulate that ovarian cancer is a methylomic heterogeneous entity consisting of subclones exhibiting both methylation statuses. Although deactivated DNA repair mechanisms may be advantageous for tumor growth at its early stages, systemic therapy will eventually select for (early-stage) subclones with active DNA repair mechanisms capable of overcoming chemotherapy. Further in-depth analyses on tumor methylomic evolution will confirm these hypotheses. Such studies may include ultra-deep sequencing of the primary tumor to discover minor subclones that may grow out in later stages of the disease as our laboratory has shown to happen in colorectal cancer [28]. Although more data with a bigger cohort are required to confirm our findings, our data suggest that methylation conversion shortens progression-free survival after initial response and thereby potentially decreasing the median overall survival times of all patients that are initially diagnosed with hypermethylation of the *BRCA1* promoter. Furthermore, in our survival analyses, we used a multivariate model, assuming that relapse is independent of previous relapses. These results may explain why hypermethylation of the *BRCA1* promoter has thus far not been correlated with overall survival and only with progression-free survival [9].

BRCA1 promoter hypermethylation in combination with gene mutation has so far been reported only once [29] and has been considered mutually exclusive. We, on the other hand, could show a relatively high frequency of double affected cases. Possible reasons may be temporal and spatial heterogeneity in which differences in *BRCA1* deactivation (LOH or promoter hypermethylation in combination with mutation) throughout the tumor take place during therapy selection over time, or a relatively low sensitivity of techniques in earlier reports compared to our assay. Subclonal investigation of both the genome and methylome is required to answer this question.

A wide range of 5–90% *BRCA1* promoter hypermethylation among ovarian cancer cases has been reported in the past [1,30–34]; the wide variability is possibly the consequence of different detection techniques, cohort selection criteria, and/or material sources (i.e., tissue vs. cfDNA). A recent meta-analysis showed that on average only 16.3% (430/2636) of all reported primary ovarian tumors are showing signs of *BRCA1* promoter hypermethylation, but that hypermethylation is strongly correlated with high-grade serous carcinomas [9]. Furthermore, in the analyzed literature, Kalachand *et al.* found three different methods for determining promoter methylation status: methylation-specific PCR, methylation sensitive restriction endonuclease digestion, and genome-wide methylation arrays; only methylation-specific PCR was correlated to progression-free survival (HR: 0.80; 95% CI: 0.66–0.97; $P = 0.02$), which is in line with our data. Taking previously published data together with our results, it can be concluded that the detection of *BRCA1* promoter hypermethylation can be achieved by methylation-specific PCR. Furthermore, methylation screening could help to identify patients who are likely not to respond to platinum re-challenge.

5. Conclusion

Here, we present the first prospective study in which liquid biopsy was used to assess and monitor the methylation status of the *BRCA1* promoter during platinum-based therapy in ovarian cancer patients. Our results suggest that hypermethylation of *BRCA1* promoter is correlated with a better survival and that conversion of methylation status is a consequence of therapy-induced selection rather than cancer evolution. This study opens the avenue for larger clinical studies in which liquid biopsy can be used to monitor the functional status of *BRCA1* by screening for its gene's promoter hypermethylation in real-time to predict the response to treatment.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

Author contributions

KaPr, KlPa, SAJ involved in conceptualization. ME, KaPr, LK, LW, and SAJ performed data curation. ME, LK, and SAJ performed formal analysis. SAJ, VM, and ME involved in funding acquisition. ME, LK, and SAJ investigated the study. ME and SAJ performed methodology and project administration. KaPr, LK, LO-F, SP, VM, LW, KlPa, and SAJ involved in resources. ME and SAJ performed software. KlPa and SAJ supervised the data. ME and SAJ validated the manuscript, involved in visualization, and wrote original draft. ME, KaPr, LK, LO-F, SP, VM, LW, KlPa, and SAJ reviewed and edited the manuscript.

Peer Review

The peer review history for this article is available at https://publons.com/publon/10.1002/1878-0261.13108.

Data accessibility

The data that support the findings of this study are available on request from the corresponding author (s.joosse@uke.de). The data are not publicly available due to privacy or ethical restrictions.

References

1 Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615.

2 Pennington KP, Walsh T, Harrell MI, Lee MK, Pennil CC, Rendi MH, Thornton A, Norquist BM, Casadei S, Nord AS *et al.* (2014) Germline and somatic mutations in homologous recombination genes predict platinum response and survival in ovarian, fallopian tube, and peritoneal carcinomas. *Clin Cancer Res* **20**, 764–775.

3 Hauke J, Hahnen E, Schneider S, Reuss A, Richters L, Kommoss S, Heimbach A, Marme F, Schmidt S, Prieske K *et al.* (2019) Deleterious somatic variants in 473 consecutive individuals with ovarian cancer: results of the observational AGO-TR1 study (NCT02222883). *J Med Genet* **56**, 574–580.

4 Harter P, Hauke J, Heitz F, Reuss A, Kommoss S, Marme F, Heimbach A, Prieske K, Richters L, Burges A *et al.* (2017) Prevalence of deleterious germline variants in risk genes including BRCA1/2 in consecutive ovarian cancer patients (AGO-TR-1). *PLoS One* **12**, e0186043.

5 Esteller M, Silva JM, Dominguez G, Bonilla F, Matias-Guiu X, Lerma E, Bussaglia E, Prat J, Harkes IC, Repasky EA *et al.* (2000) Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. *J Natl Cancer Inst* **92**, 564–569.

6 Joosse SA, Brandwijk KI, Mulder L, Wesseling J, Hannemann J & Nederlof PM (2011) Genomic signature of BRCA1 deficiency in sporadic basal-like breast tumors. *Genes Chromosom Cancer* **50**, 71–81.

7 Holstege H, van Beers E, Velds A, Liu X, Joosse SA, Klarenbeek S, Schut E, Kerkhoven R, Klijn CN, Wessels LF *et al.* (2010) Cross-species comparison of aCGH data from mouse and human BRCA1- and BRCA2-mutated breast cancers. *BMC Cancer* **10**, 455.

8 Kondrashova O, Topp M, Nesic K, Lieschke E, Ho GY, Harrell MI, Zapparoli GV, Hadley A, Holian R, Boehm E *et al.* (2018) Methylation of all BRCA1 copies predicts response to the PARP inhibitor rucaparib in ovarian carcinoma. *Nat Commun* **9**, 3970.

9 Kalachand RD, Stordal B, Madden S, Chandler B, Cunningham J, Goode EL, Ruscito I, Braicu EI, Sehouli J, Ignatov A *et al.* (2020) BRCA1 promoter methylation and clinical outcomes in ovarian cancer: an individual patient data meta-analysis. *J Natl Cancer Inst* **112**, 1190–1203.

10 Prieske K, Prieske S, Joosse SA, Trillsch F, Grimm D, Burandt E, Mahner S, Schmalfeldt B, Milde-Langosch K, Oliveira-Ferrer L *et al.* (2017) Loss of BRCA1 promotor hypermethylation in recurrent high-grade ovarian cancer. *Oncotarget* **8**, 83063–83074.

11 Joosse SA & Pantel K (2016) Genetic traits for hematogenous tumor cell dissemination in cancer patients. *Cancer Metastasis Rev* **35**, 41–48.

12 Guo M, Peng Y, Gao A, Du C & Herman JG (2019) Epigenetic heterogeneity in cancer. *Biomark Res* **7**, 23.

13 Joosse SA, Gorges TM & Pantel K (2015) Biology, detection, and clinical implications of circulating tumor cells. *EMBO Mol Med* **7**, 1–11.

14 Joosse SA & Pantel K (2013) Biologic challenges in the detection of circulating tumor cells. *Cancer Res* **73**, 8–11.

15 Joosse SA & Pantel K (2015) Tumor-educated platelets as liquid biopsy in cancer patients. *Cancer Cell* **28**, 552–554.

16 Elazezy M & Joosse SA (2018) Techniques of using circulating tumor DNA as a liquid biopsy component in cancer management. *Comput Struct Biotechnol J* **16**, 370–378.

17 Giannopoulou L, Mastoraki S, Buderath P, Strati A, Pavlakis K, Kasimir-Bauer S & Lianidou ES (2018) ESR1 methylation in primary tumors and paired circulating tumor DNA of patients with high-grade serous ovarian cancer. *Gynecol Oncol* **150**, 355–360.

18 Mastoraki S, Strati A, Tzanikou E, Chimonidou M, Politaki E, Voutsina A, Psyrri A, Georgoulias V & Lianidou E (2018) ESR1 methylation: a liquid biopsy-based epigenetic assay for the follow-up of patients with metastatic breast cancer receiving endocrine treatment. *Clin Cancer Res* **24**, 1500–1510.

19 Kumaki Y, Oda M & Okano M (2008) QUMA: quantification tool for methylation analysis. *Nucleic Acids Res* **36**, W170–W175.

20 Joosse SA (2020) In-Silico Online (version 2.1.2). Available at http://in-silico.online

21 Amorim LD & Cai J (2015) Modelling recurrent events: a tutorial for analysis in epidemiology. *Int J Epidemiol* **44**, 324–333.

22 Clark TG, Bradburn MJ, Love SB & Altman DG (2003) Survival analysis part IV: further concepts and methods in survival analysis. *Br J Cancer* **89**, 781–786.

23 Gormally E, Hainaut P, Caboux E, Airoldi L, Autrup H, Malaveille C, Dunning A, Garte S, Matullo G, Overvad K *et al.* (2004) Amount of DNA in plasma and cancer risk: a prospective study. *Int J Cancer* **111**, 746–749.

24 Dobrovic A & Simpfendorfer D (1997) Methylation of the BRCA1 gene in sporadic breast cancer. *Cancer Res* **57**, 3347–3350.

25 Capizzi E, Gabusi E, Grigioni AD, De Iaco P, Rosati M, Zamagni C & Fiorentino M (2008) Quantification of free plasma DNA before and after chemotherapy in patients with advanced epithelial ovarian cancer. *Diagn Mol Pathol* **17**, 34–38.

26 Kamat AA, Baldwin M, Urbauer D, Dang D, Han LY, Godwin A, Karlan BY, Simpson JL, Gershenson DM, Coleman RL *et al.* (2010) Plasma cell-free DNA in ovarian cancer: an independent prognostic biomarker. *Cancer* **116**, 1918–1925.

27 Douvdevani A, Bernstein-Molho R, Asraf K, Doolman R, Laitman Y & Friedman E (2020) Circulating cell-free DNA (cfDNA) levels in BRCA1 and BRCA2 mutation carriers: a preliminary study. *Cancer Biomark* **28**, 269–273.

28 Gasch C, Bauernhofer T, Pichler M, Langer-Freitag S, Reeh M, Seifert AM, Mauermann O, Izbicki JR, Pantel K & Riethdorf S (2013) Heterogeneity of epidermal growth factor receptor status and mutations of KRAS/PIK3CA in circulating tumor cells of patients with colorectal cancer. *Clin Chem* **59**, 252–260.

29 Rzepecka IK, Szafron L, Stys A, Bujko M, Plisiecka-Halasa J, Madry R, Osuch B, Markowska J, Bidzinski M & Kupryjanczyk J (2012) High frequency of allelic loss at the BRCA1 locus in ovarian cancers: clinicopathologic and molecular associations. *Cancer Genet* **205**, 94–100.

30 Abkevich V, Timms KM, Hennessy BT, Potter J, Carey MS, Meyer LA, Smith-McCune K, Broaddus R, Lu KH, Chen J *et al.* (2012) Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *Br J Cancer* **107**, 1776–1782.

31 Baldwin RL, Nemeth E, Tran H, Shvartsman H, Cass I, Narod S & Karlan BY (2000) BRCA1 promoter region hypermethylation in ovarian carcinoma: a population-based study. *Cancer Res* **60**, 5329–5333.

32 Geisler JP, Hatterman-Zogg MA, Rathe JA & Buller RE (2002) Frequency of BRCA1 dysfunction in ovarian cancer. *J Natl Cancer Inst* **94**, 61–67.

33 Catteau A, Harris WH, Xu CF & Solomon E (1999) Methylation of the BRCA1 promoter region in sporadic breast and ovarian cancer: correlation with disease characteristics. *Oncogene* **18**, 1957–1965.

34 Pradjatmo H, Dasuki D, Anwar M, Mubarika S & Harijadi (2014) Methylation status and immunohistochemistry of BRCA1 in epithelial ovarian cancer. *Asian Pac J Cancer Prev* **15**, 9479–9485.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Detailed summary of progression models. Two models explain the conversion of *BRCA1* promoter hypermethylation. *BRCA1* promoter hypermethylation is an early event in tumorigenesis. After detection of the primary tumor and multiple rounds of therapy after relapse, the tumor reactivates BRCA1 by evolving and reversing its methylation status and thereby developing therapy resistance (upper panel). Alternatively, multiple subclones may have already developed during tumorigenesis and through multiple rounds of therapy, the most therapy resistant clone eventually survives and thrives (lower panel). The arrows indicate the time from the development to detection of the tumor to be treated, illustrating the need for different statistical models for analysis.

Table S1. Detailed summary of the patients' characteristics.

57

5.3. Detection and characterization of estrogen receptor α expression of circulating tumor cells as a prognostic marker

Estrogen receptor is one of the most crucial biomarkers in breast cancer management. More than 60% of breast cancer patients express estrogen receptor (ER). Endocrine therapy is routinely recommended for metastatic breast cancer patients with estrogen receptor-positive. Nevertheless, 30-40% of patients treated with endocrine therapy developed resistance. Hence, detection and monitor estrogen receptor may serve as a predictor of endocrine therapy response. The presented study evaluated the ERα monoclonal murine ER-119.3 antibody used by Paoletti *et al*. in 109 blood samples from 60 metastatic ER-positive breast cancer patients using the CellSearch system. We found that a high number of CTCs during therapy was associated with disease progression, whereas a lower number of CTCs or a CTC-negative status was associated with stable disease. A favorable relationship between CTC status and progression-free survival was detected during the course of the disease (P = 0.0045). We detected ER-positive CTCs in 32% of the CTC-positive samples, although all patients were diagnosed with an ER-positive primary breast tumor. The most obvious finding is that ER expression in individual CTCs was heterogeneous within and between patients, which highlights the value of ER in endocrine therapy resistance, which may need to be addressed on a large scale in future studies.

Detection and characterization of estrogen receptor α expression of circulating tumor cells as a prognostic marker

Retno Ningsi[1†], Maha Elazezy[1†], Luisa Stegat[1], Elena Laakmann[2], Sven Peine[3], Sabine Riethdorf[1], Volkmar Müller[2], Klaus Pantel[1], Simon A. Joosse[1,4*]

[1]Department of Tumor Biology, [2]Department of Gynecology, [3]Department of Transfusion Medicine, [4]Mildred Scheel Cancer Career Center HaTriCS4, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany.

***Correspondence: s.joosse@uke.de

† Shared first author.

Abbreviations

Cyclin-dependent Kinase 4 and 6, CDK-4/6; confidence interval, CI; CTC, circulating tumor cell; CTC-Endocrine Therapy Index, CTC-ETI; ER, estrogen receptor; Food and Drugs Administration , FDA; Hazard Ratio, HR; s, standard deviation; K, keratin

Abstract

BACKGROUND. Detection of Circulating Tumor Cells (CTCs) in the blood is correlated to survival in metastatic breast cancer. Assessment of estrogen receptor (ER) expression in CTCs has been suggested as a potential marker to improve the clinical management of patients with hormonal positive breast cancer. Therefore, this study aimed to evaluate the expression of ER in CTCs using semi-automated quantification.

METHODS. From sixty metastatic ER-positive breast cancer patients, 109 longitudinal blood samples were prospectively collected and analyzed using the CellSearch System in combination with the ERα monoclonal murine ER-119.3 antibody.

RESULTS. Thirty-one cases were found to be CTC-positive and an increased number of CTCs during treatment was correlated with disease progression, whereas a decrease or stable amount of CTC number during treatment was correlated with a better clinical outcome. Survival analyses further indicate a positive association between CTC status and progression-free survival (HR, 66.17; 95%CI, 3.66-195.96; P = 0.0045). Incubation with permeabilization agent is required for ER staining in tumor cells. In metastatic breast cancer patient that were initially diagnosed with ER-positive primary breast cancer, only a third harbor detectable ER-positive CTCs.

CONCLUSION. As reported in other studies, CTC-positivity is correlated with shorter relapse-free survival and only a third of the CTC-positive patients harbor ER-positive CTCs. The expression of ER in individual CTCs was intra- and inter-patient heterogeneous, which might be further investigated as potential source of resistance to endocrine therapy in future studies.

1. Introduction

Breast cancer is the most common cancer among women worldwide with over 2 million new cases in 2018 and more than 600 000 deaths in the same year [1]. One of the most essential biomarkers in breast cancer management is the estrogen receptor (ER). More than 70% of breast cancer cases express ER, and the classification of breast cancer into ER-positive and ER-negative determines the type of therapy the patient will receive. Estrogen is a steroid hormone that upon binding to its receptor in the normal situation, affects the growth and differentiation of the mammary gland. In breast cancer, estrogen stimulates tumor cell proliveration. In general, patients with ER-positive breast cancer respond well to endocrine therapy due to the suppression of estrogen production or blocking of the receptor's binding site [2]. In addition, as ER-positive breast cancer has shown to exhibit hyperactivity of Cyclin-dependent Kinase 4 and 6 (CDK-4/6), a combination of endocrine therapy with CDK-4/6 inhibitors has been shown to further improve outcome [3]. Unfortunately, 10% of breast cancer cases is diagnosed with metastatic disease and many patients will develop systemic relapse over time. Although the hormone receptor status of relapses are usually maintained, due to crosstalk between hormone receptors and growth factors, but also due to genetic progression and point mutations [4], many patients develop therapy resistance after 24-36 months [5]. As a consequence, distant metastasis is still the leading cause of breast cancer-related death [6].

Cancer metastasis starts with single or clusters of tumor cells separating from the primary tumor and intravasating into the bloodstream [7]. The motility of these cells enabling tumor dissemination is made possible by the epithelial-to-mesenchymal transition [8]. Next, these so-called circulating tumor cells (CTCs) may extravasate into distant organs, adapt to a new environment, and grow out to become a new tumor mass. In breast cancer, as well as many other cancer entities, it has been shown that the number of detectable CTCs in the blood of early-stage and metastatic breast cancer patients is negatively associated with progression-free survival [9]. Because blood is easily acquired and can be obtained repeatedly, CTCs as a so called "liquid biopsy" have increasingly gained attention during the past decades [10]. Phenotyping CTCs can provide crucial information on the evolving characteristics of the tumor during progression and development of therapy resistance [11-13]. Although CTCs can be detected in more than 60% of metastatic breast cancer patients [14, 15], due to the low amount of CTCs (typically 1-10 CTCs/7.5 ml blood), reliable detection can still be challenging [7, 16, 17]. To overcome this challenge, many techniques for the quantification, characterization, and isolocation of CTCs have been developed based on different cell properties [10, 18-21]. Currently the only FDA-cleared method is the CellSearch® System, which is therefore considered the gold standard in CTC detection [11, 22]. This system enriches for EpCAM-positive cells and detects CTCs based on the expression of keratin (K), but negatively selects for CD45 expression. Besides the initial number of CTCs, the change in the status of CTCs during treatment with systemic therapy in metastatic breast cancer patients has recently been linked with prognosis in a meta-analysis by Yan et al. [23], showing the relevance of monitoring treatment efficacy.

Previously, we found that ER expression among CTCs is heterogeneous in metastatic breast cancer patients who were initially diagnosed with an ER-positive primary breast tumor [24]. The heterogeneous expression of ER may indicate estrogen independence of disseminated cells and thus a possible predictor for hormonal treatment failure. Not surprisingly, in a more recent study it was shown that a shift from ER-positvie breast cancer to ER-negative CTCs in metastatic disease was associated with a worse prognosis as compared to cases in which ER-positivity was stable [25]. Paoletti and colleagues developed a CTC-Endocrine Therapy Index (CTC-ETI) using the semi-automated CellSearch System for CTC quantification and characterization by assessing

the expression of ER, BCL-2, ErbB2, and Ki-67 [26-28]. The preliminary data from Paoletti *et al.* demonstrated the reproducibility of CTC-ETI as a predictive factor for resistance to endocrine therapy in ER-positive metastatic breast cancer patients [26]. Recently, they determined the CTC-ETI in a larger cohort of ER-positive metastatic breast cancer patients in their phase 2 trial [28]. Paoletti and colleagues found that patients with high CTC-ETI were more likely to have rapid progression after three months of treatment. Taken together, it may be concluded that ER expression on CTCs have a predictive value of endocrine therapy response.

In the presented study, we evaluated the ER-CTC status using the ERα monoclonal murine ER-119.3 antibody, which to our knowledge has been tested only once so far. We aimed to reproduce the previously published staining conditions *in vitro* followed by testing a cohort of metastatic breast cancer patients with ER-positive primary tumor using the CellSearch System for the quantification of CTCs in metastatic breast cancer. Based on the available literature, we expected an ER-positivity rate of approximately 30% among CTCs. Determination of the ER status of CTCs could help in therapy stratification, monitoring therapy efficacy, and predict the risk for metastatic relapse.

2. Materials and Methods

2.1. Cell lines

The two human breast cancer cell lines MCF-7 and SK-BR-3 (acquired from ATCC) were used in this study as positive and negative control for estrogen receptor (ER) expression to optimize the experimental conditions. Both cell lines were cultivated in DMEM (catalog no. E15-011, PAA Laboratories) at 37°C and 10% CO_2. The medium was supplemented with 10% fetal bovine serum, penicillin-streptomycin antibiotics and L-Glutamine (catalog no. E15-151, M11-004, P11-010, PAA Laboratories).

2.2. Patients

Sixty metastatic breast cancer patients with initially ER-positive primary tumors were included into the study. Patients were treated for metastatic breast cancer at the University Medical Center Hamburg-Eppendorf and received endocrine therapy as first-line therapy, followed by chemotherapy at the progression of the disease. The mean age of patients was 62 years (range: 34-86). All patients gave written informed consent to be included into the study. This study was approved by the local ethical board under number PV4367.

2.3. CTC detection

The CellSearch CXC kit (Menarini Silicon Biosystems, Bologna, Italy) was used to detect circulating tumor cells. This kit contains ferrofluid particles coated with anti-EpCAM antibodies, anti-keratin antibodies recognizing keratins (fluorescein-labeled) to specifically identify epithelial cells, an antibody against CD45 (labeled with allophycocyanin) as a negative selection marker for white blood cells, and a nuclear dye 4',6-diamidino-2-phenylindole (DAPI) [29]. We added ERα monoclonal murine ER-119.3 antibody (phycoerythrin-labeled , Ex_{max} 496 nm/Em_{max} 578 nm) to identify estrogen receptor expression in the nucleus.

The enrichment was done automatically by the CellTracks Autoprepsystem. The enriched cells were collected in the MagNest cartridge. Using CellSpotter Analyzer each cell in the cartridge was displayed based on the signal sent from the antibodies. Cells were enumerated as circulating tumor cells if the signal for keratin and

3

cell nucleus were positive and the signal for CD45 was negative. The estrogen receptor expression was shown by the positive nuclear signal of ERα antibody.

2.4. Statistics

The statistical analysis is performed with R (version 4.0.1) [30] and In-Silico Online, version 2.3.0 [31]. Because one-directional increase or decrease in the number of CTCs upon therapy failure or success is expected, one-sided tests were applied where appropriate. Survival analyses were performed using the logrank test and multivariable analysis by Cox proportional hazards function, with death by cancer as endpoint. An alpha level of 0.05 was applied to determine statistical significance.

3. Results

3.1. *in vitro* evaluation of the ER antibody

In order to evaluate previously published results on the ER antibody for semi-automated characterization of CTCs [10], an ER-positive (MCF-7) and ER-negative cell line (SK-BR-3) were cultured under identical conditions and used to establish the optimal experimental settings. Eight blood samples from healthy donors were spiked with 100 tumor cell line cells cultured at 37°C and 10% CO_2 with no incubation time after spiking. Four blood samples were spiked with MCF-7 cells and four with SK-BR-3 cells. Using the CellSearch System, the mean recovery of the cells was 94% ($s=7\%$) and 89% ($s=6.1\%$) of the MCF-7 and SKBR3 cells, respectively. However, only 15% ($s=2.1\%$) of the MCF-7 cells were found to be ER-positive on average. As expected, none of the detected SK-BR-3 cells were ER-positive.

Because more ER-positive MCF-7 cells were expected, the experiment was repeated, but spiked blood was incubated in CellSave tubes containing preservative buffer for approximately 24 hours to fix and permeabilize the cells. The mean percentage of ER-positive tumor cells among the total detected MCF-7 cells increased to 45% ($s=2.2$), and among SK-BR-3 cells stayed at 0% (Figure 1). In addition, we validated the results by determining ER protein expression using our previously published protocol as well [9]. The mean percentage of ER-positive MCF-7 cells as detected with our manual protocol was 40% ($s=4.3$), which was comparable with the results obtained with the CellSearch System (mean: 45%, $s=5.1$; p-value: 0.19, Welch's Two Sample t-test).

Based on these results, the final protocol was as follows: blood from metastatic breast cancer was drawn into a CellSave preservative tube and incubated at room temperature until processing the next day. The ER channel was analyzed using 0.2-second integration time; higher integration time resulted in too high a background.

3.2. Circulating tumor cells (CTCs) in metastatic breast cancer patients

Patients included in this study were selected prospectively based on the diagnosis of advanced disease with distant metastasis. Primary diagnosis of breast cancer among these patients was at a mean age of 52 years (range: 28-86), on average 7 years (range: 0-33) before study inclusion (Table 1). All patients were diagnosed with ER-positive primary breast cancer and twelve (20%) with primary metastasis. One hundred and nine blood samples from 60 patients were collected during this study (Figure 2). Nine samples were collected at the admission interview before the initiation of systemic therapy, 37 samples were collected during endocrine therapy, 51 samples

4

were collected at a progressive stage at which the patients received chemotherapy, and 12 samples failed the analysis. Out of all 97 blood samples, CTCs were found in 31; these blood samples were from 20 patients (Table 1 and Figure 3). In 15% (15/97) of the samples, 1-4 CTCs/7.5 ml were found, and in 16% (16/97) of samples, 5 or more CTCs could be detected. The presence of CTCs was not correlated to the TNM-stage at primary diagnosis or therapy at blood draw but was correlated to the stage of disease at the time of blood draw. Patients experiencing progression of disease were more frequently diagnosed with CTCs in their blood compared to patients with stable disease (p=0.0053, G-test).

3.3. Evaluation of ERα monoclonal murine ER-119.3 antibody on CTCs

ER-positive CTCs could be detected in 10 out of 31 CTC-positive blood samples (32%). In these ten cases, the mean percentage of ER-positive CTCs was 28% (range: 9-100%; Figure 3) and ranged from 1 to 207 CTCs. The total percentage of ER-positive CTCs among all detected CTCs was 18%. CTCs were detected in 4/9 (44%) of the samples that were collected before therapy; of these 2/4 (50%) cases were diagnosed with ER-positive CTCs. In 10/37 (33%) of the samples, CTCs were detected during hormone therapy, of which 1/10 (10%) exhibited ER-positive CTCs. In the blood samples from patients treated with chemotherapy, 17/52 (33%) CTCs were detected and 7/17 ER-positive CTCs. Although the fewest ER-positive CTCs could be detected among the patients treated with hormone therapy, no statistical significance was found (p=0.17, G-test). These results indicate a heterogeneous expression of ER among CTCs within individual patients.

3.4. Monitoring CTC count during therapy

Longitudinal blood samples were obtained from twenty-five patients. Three patients were initially diagnosed with ER-positive CTCs but changed to completely CTC-negative or ER-negative CTCs during the course of the study. All other patients were diagnosed with ER-negative CTCs only according to the CellSearch System results. If more than two blood draws were taken, only the first two before and after a change of clinical response were considered for further analyses.

Fourteen patients experienced progression of disease during this study. At the time of blood sampling before progression, two patients were found to be positive for CTCs. During the second blood sampling after the diagnosis of progression, one of the CTC-positive cases was diagnosed CTC-negative, and one stayed positive; five CTC-negative cases converted to CTC-positive, and seven remained CTC-negative. The increase in the median number of CTCs was significantly different (p-value: 0.0367, paired Wilcoxon signed rank test with continuity correction), as was the number of cases converting CTC-status from the first to second blood draw (p-value: 0.0352, Exact McNemar test).

Eleven cases were diagnosed with the stable disease throughout the period of the study or converted from progression to stable disease. At the first blood collection, seven patients were diagnosed CTC-positive and 4 CTC-negative. At the second blood draw, one case remained positive, whereas the other ten cases became or stayed CTC-negative. The decrease in the number of CTC from the second to the first blood draw was statistically significant (p-value: 0.0111, paired Wilcoxon signed rank test with continuity correction), as was the number of cases converting to a CTC-negative status (p-value: 0.0156, Exact McNemar test).

5

Because these data suggest that an increase or decrease of the median number of CTCs is associated with progression and stable disease, respectively, the median differences between the two blood draws of the two clinical response groups were compared (Figure 4). The median difference in CTC number between blood draws of the patients in stable disease was significantly less compared to the median difference in CTC number between blood draws of the patients in the progression of disease (p-value: 0.0102, Wilcoxon rank sum test with continuity correction). Among the patients with stable disease, a decrease in CTCs was seen in seven and no change in four patients. Among the patients experiencing a progression of disease, a decrease in CTC number was seen in 2 patients, an increase in 3 patients, and no change in 9 patients (p-value: 0.0206, G-test of independence with Williams' correction). In order to further study the clinical association with change in CTC status, survival analyses were performed next.

3.5. Survival analysis

The median follow-up after the first blood draw was 22.4 months (range: 1.8-61.6); during the time of the study, 28 patients died (47%). The patient data were divided on CTC status (positive or negative) at the time of the first blood draw, and overall survival was compared. More than 50% of the patients characterized as initially CTC-negative had an overall survival longer than the duration of the study, whereas the median survival of the patients diagnosed as CTC-positive was 8.6 months (p-value < 0.0001, Logrank test, Figure 5A). Also, in uni- and multivariable analyses, CTC-status was independently associated with survival (HR: 66.2, 95%CI: [3.7, 1196], p-value: 0.0045), whereas T-, N-, and M-stage at primary diagnosis were not (Table 2). These data indicate that a positive diagnosis for CTCs is highly associated with poor overall survival.

The 25 patients from which multiple blood samples were collected were separated into two groups: 1) cases with a stable CTC-status or conversion from positive to negative CTC-status and 2) cases with the conversion from negative to positive CTC-status. The individuals experiencing a CTC conversion (n=6) had a median survival of 13.5 months, whereas the individuals with a stable CTC status (n=19) had a median survival of 49.3 months, which were statistically different (p-value: 0.008, Logrank test, Figure 5B). In only three of the cases in which multiple samples were collected, ER+ CTCs were detected. Therefore, the effect of CTC ER-status on survival could not be assessed longitudinally. When considering only the first blood samples, the median overall survival of patients diagnosed with ER-positive CTCs was 7.3 months and 12.5 months of patients diagnosed with ER-negative CTCs only (p-value: 0.32, Logrank test).

4. Discussion

CTCs are considered the seeds for new metastases and their presence in the blood of patients with primary breast cancer may provide prognostic information in regards to relapse-free survival. Similarly, in metastatic setting the quantification of CTCs may offer prognostic information in regards to overall survival. Moreover, determining the expression of therapeutic markers to characterize the tumor as well as to monitor therapy efficiacy, could lead to personalized treatment regimens.

6

Earlier studies have shown that progression-free and overall survival of metastatic breast cancer is correlated with CTC presence (PFS: HR, 1.78; 95% CI, 1.52–2.09; OS: HR, 2.33; 95% CI, 2.09–2.60), irrespective of the CTC detection method and time point of blood withdrawal [32]. In line with these data, our study shows a clear correlation with overall survival and the detection of CTCs using the CellSearch System as well. Nevertheless, it should be noted that the sample cohort of our study is heterogeneous and the blood was collected at various time points in the patients' therapy regiments and confounding factors cannot be excluded. Therefore, an important feature of liquid biopsy that can be taken advantage of is repeated measurement and monitoring patient's CTC status during the course of therapy. In our study, fourteen patients experienced progression, whereas ten patients remained stable or even had a regression. As expected, patients with progression showed conversion of being CTC negative to positive or were diagnosed with an increased in the number of CTCs. Once the therapy was escalated from endocrine therapy to chemotherapy, we observed a decrease in CTC number. Taken together, our data are in line with earlier findings in which a significant correlation was found between CTCs and progression-free survival [23, 33].

The majority of clinical studies on breast cancer involving liquid biopsy are based on CTC quantification only. Although a positive diagnosis for CTCs is strongly associated with poor overall survival in metastatic breast cancer patients, CTC enumeration alone cannot elucidate the molecular mechanisms involved in therapy resistance or guide the choice of systemic therapy. Therefore, phenotyping CTCs can add another dimension to cancer management that may ultimately lead to personalized treatment. Especially the estrogen receptor (ER) expression status is an important therapeutic marker in the management of breast cancer. Forsare *et al.* showed that the precense of ER-positive CTCs at baseline and after initiation of systemic treatment is associated with a better prognosis as compared to ER-negative CTCs in patients with breast cancer [25]. The authors use a manual ER staining after CellSearch System based enrichment of CTCs and apply methanol for fixation and permeabilization, which is in contrast from the study of Paoletti *et al.* [28] and the study presented here where ER labeling is performed within the CellSearch System. Nevertheless, in all studies, approximately in only a third of the CTC-positive cases ER-positive CTCs can be detected detected: 36% (38/107), 25% (16/63), and 32% (10/31), respectively [25, 28]. In all studies, heterogeneity in ER expression status among CTCs within individual patients can be observed, consistent with our previous study in which we did not use the CellSearch System for CTC enrichment [24].

As a consequence of the limitations of the study, the relatively low number of patients with detectable ER-positive CTCs, the relatively short follow-up time, and heterogeneity in sample collection in our study, an association with survival could not be made with ER status of CTCs yet. Nevertheless, to our knowledge, this is one of few studies that evaluating ER status of CTCs using the CellSearch CXC kit. An important finding of this study is the requirement of permeabilization in order for ER antibodies to penetrate the cell membrane and reach the cell nucleus. Therefore, future studies conducted using the CellSearch System should make use of the CellSave preservation tubes for blood collection. Furthermore, (pre-)clinical studies may use the results of this study for sample-size calculations, taking the expected fraction of patients with detectable ER-positive CTCs of 32% into account. Overall, the identification and monitoring of ER status of CTCs is extremely important for the management of breast cancer patients. Further investigation towards a robust and reproducible assay is needed with larger cohorts.

7

5. Conclusion

Half of metastatic breast cancer patients that were diagnosed with an ER-positive primary tumor harbor detectable CTCs in their peripheral blood circulation. Most of these patients carry ER-negative CTCs only, whereas approximately a third show a mixture of ER-positive and -negative CTCs. The detection of CTCs remains a well-established marker for poor outcomes and shorter overall survival. Monitoring ER-CTC status could add a prognostic value to CTCs enumeration and may serve as a therapy resistance prediction. Therefore, further studies with large cohorts on long-term monitoring of ER-CTCs status are required to address the predictive value of ER-CTC for endocrine therapy efficacy in breast cancer patients.

8

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries**. *CA: a cancer journal for clinicians* 2018, **68**(6):394-424.
2. Waks AG, Winer EP: **Breast Cancer Treatment**. *JAMA* 2019, **321**(3):316.
3. Pernas S, Tolaney SM, Winer EP, Goel S: **CDK4/6 inhibition in breast cancer: current practice and future directions**. *Ther Adv Med Oncol* 2018, **10**:1758835918786451.
4. Joosse SA, Pantel K: **Genetic traits for hematogeneous tumor cell dissemination in cancer patients**. *Cancer Metastasis Rev* 2016, **35**(1):41-48.
5. Dixon JM: **Endocrine Resistance in Breast Cancer**. *New Journal of Science* 2014, **2014**:390618.
6. Hagemeister FB, Jr., Buzdar AU, Luna MA, Blumenschein GR: **Causes of death in breast cancer: a clinicopathologic study**. *Cancer* 1980, **46**(1):162-167.
7. Joosse SA, Gorges TM, Pantel K: **Biology, detection, and clinical implications of circulating tumor cells**. *EMBO Mol Med* 2015, **7**(1):1-11.
8. Joosse SA, Pantel K: **Biologic challenges in the detection of circulating tumor cells**. *Cancer research* 2013, **73**(1):8-11.
9. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, Reuben JM, Doyle GV, Allard WJ, Terstappen LW *et al*: **Circulating tumor cells, disease progression, and survival in metastatic breast cancer**. *N Engl J Med* 2004, **351**.
10. Joosse SA, Gorges TM, Pantel K: **Biology, detection, and clinical implications of circulating tumor cells**. *EMBO molecular medicine* 2014, **7**(1):1-11.
11. Guan X, Ma F, Liu S, Wu S, Xiao R, Yuan L, Sun X, Yi Z, Yang H, Xu B: **Analysis of the hormone receptor status of circulating tumor cell subpopulations based on epithelial-mesenchymal transition: a proof-of-principle study on the heterogeneity of circulating tumor cells**. *Oncotarget* 2016, **7**(40):65993-66002.
12. Saxena K, Subbalakshmi AR, Jolly MK: **Phenotypic heterogeneity in circulating tumor cells and its prognostic value in metastasis and overall survival**. *EBioMedicine* 2019, **46**:4-5.
13. Roßwag S, Cotarelo CL, Pantel K, Riethdorf S, Sleeman JP, Schmidt M, Thaler S: **Functional Characterization of Circulating Tumor Cells (CTCs) from Metastatic ER+/HER2- Breast Cancer Reveals Dependence on HER2 and FOXM1 for Endocrine Therapy Resistance and Tumor Cell Survival: Implications for Treatment of ER+/HER2- Breast Cancer**. *Cancers (Basel)* 2021, **13**(8).
14. Gwark S, Kim J, Kwon N-J, Kim K-Y, Kim Y, Lee CH, Kim YH, Kim MS, Hong SW, Choi MY *et al*: **Analysis of the serial circulating tumor cell count during neoadjuvant chemotherapy in breast cancer patients**. *Scientific Reports* 2020, **10**(1):17466.
15. Eroglu Z, Fielder O, Somlo G: **Analysis of circulating tumor cells in breast cancer**. *J Natl Compr Canc Netw* 2013, **11**(8):977-985.
16. Alix-Panabieres C, Pantel K: **Challenges in circulating tumour cell research**. *Nature reviews Cancer* 2014, **14**(9):623-631.
17. Alix-Panabières C, Pantel K: **Technologies for detection of circulating tumor cells: facts and vision**. *Lab Chip* 2014, **14**(1):57-62.
18. Liu HY, Koch C, Haller A, Joosse SA, Kumar R, Vellekoop MJ, Horst LJ, Keller L, Babayan A, Failla AV *et al*: **Evaluation of Microfluidic Ceiling Designs for the Capture of Circulating Tumor Cells on a Microarray Platform**. *Adv Biosyst* 2020, **4**(2):e1900162.
19. Koch C, Joosse SA, Schneegans S, Wilken OJW, Janning M, Loreth D, Muller V, Prieske K, Banys-Paluchowski M, Horst LJ *et al*: **Pre-Analytical and Analytical Variables of Label-Independent Enrichment and Automated Detection of Circulating Tumor Cells in Cancer Patients**. *Cancers (Basel)* 2020, **12**(2).

9

20. Gorges TM, Penkalla N, Schalk T, Joosse SA, Riethdorf S, Tucholski J, Lucke K, Wikman H, Jackson S, Brychta N *et al*: **Enumeration and Molecular Characterization of Tumor Cells in Lung Cancer Patients Using a Novel In Vivo Device for Capturing Circulating Tumor Cells**. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2016, **22**(9):2197-2206.

21. Gorges TM, Stein A, Quidde J, Hauch S, Rock K, Riethdorf S, Joosse SA, Pantel K: **Improved Detection of Circulating Tumor Cells in Metastatic Colorectal Cancer by the Combination of the CellSearch(R) System and the AdnaTest(R)**. *PloS one* 2016, **11**(5):e0155126.

22. Müller V, Riethdorf S, Rack B, Janni W, Fasching PA, Solomayer E, Aktas B, Kasimir-Bauer S, Pantel K, Fehm T: **Prognostic impact of circulating tumor cells assessed with the CellSearch System™ and AdnaTest Breast™ in metastatic breast cancer patients: the DETECT study**. *Breast Cancer Research* 2012, **14**(4):R118.

23. Yan W-T, Cui X, Chen Q, Li Y-F, Cui Y-H, Wang Y, Jiang J: **Circulating tumor cell status monitors the treatment responses in breast cancer patients: a meta-analysis**. *Scientific Reports* 2017, **7**(1):43464.

24. Babayan A, Hannemann J, Spotter J, Muller V, Pantel K, Joosse SA: **Heterogeneity of estrogen receptor expression in circulating tumor cells from metastatic breast cancer patients**. *PloS one* 2013, **8**(9):e75038.

25. Forsare C, Bendahl PO, Moberg E, Levin Tykjær Jørgensen C, Jansson S, Larsson AM, Aaltonen K, Rydén L: **Evolution of Estrogen Receptor Status from Primary Tumors to Metastasis and Serially Collected Circulating Tumor Cells**. *Int J Mol Sci* 2020, **21**(8).

26. Paoletti C, Muñiz MC, Thomas DG, Griffith KA, Kidwell KM, Tokudome N, Brown ME, Aung K, Miller MC, Blossom DL *et al*: **Development of circulating tumor cell-endocrine therapy index in patients with hormone receptor-positive breast cancer**. *Clin Cancer Res* 2015, **21**(11):2487-2498.

27. Paoletti C, Regan MM, Liu MC, Marcom PK, Hart LL, Smith JW, Tedesco KL, Amir E, Krop IE, DeMichele AM *et al*: **Abstract P1-01-01: Circulating tumor cell number and CTC-endocrine therapy index predict clinical outcomes in ER positive metastatic breast cancer patients: Results of the COMETI Phase 2 trial**. *Cancer Research* 2017, **77**(4 Supplement):P1-01-01.

28. Paoletti C, Regan MM, Niman SM, Dolce EM, Darga EP, Liu MC, Marcom PK, Hart LL, Smith JW, 2nd, Tedesco KL *et al*: **Circulating tumor cell number and endocrine therapy index in ER positive metastatic breast cancer patients**. *NPJ Breast Cancer* 2021, **7**(1):77.

29. Allard WJ, Matera J, Miller MC, Repollet M, Connelly MC, Rao C, Tibbe AG, Uhr JW, Terstappen LW: **Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases**. *Clin Cancer Res* 2004, **10**(20):6897-6904.

30. **R: a languange and enviroment for statistical computing**. *In Vienna, Austria: R Foundation for Statistical Computing* 2017.

31. **In-Silico Online (version 2.3.0) [http://in-silico.online]**.

32. Zhang L, Riethdorf S, Wu G, Wang T, Yang K, Peng G, Liu J, Pantel K: **Meta-analysis of the prognostic value of circulating tumor cells in breast cancer**. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2012, **18**(20):5701-5710.

33. Bidard FC, Hajage D, Bachelot T, Delaloge S, Brain E, Campone M, Cottu P, Beuzeboc P, Rolland E, Mathiot C *et al*: **Assessment of circulating tumor cells and serum markers for progression-free survival prediction in metastatic breast cancer: a prospective observational study**. *Breast Cancer Res* 2012, **14**(1):R29.

10

Table 1 – Demographic statistics. Number of CTC-positive (CTC+) and CTC-negative (CTC-) blood samples divided according to clinical variables of the patients at primary diagnosis. P values were calculated using Welch's two-sided t-test and Log likelihood ratio (G-test) test of independence with Williams' correction. CTC values were not available for 3 patients.

At primary diagnosis	N=60	CTC+ (n=20)	CTC- (n=37)	P value
Age (years)				
Mean	52	55	49	0.0848
Range	28-76	39-76	28-75	
ER				
Positive	60	20	37	-
PR				
Positive	54	18	34	0.8209
Negative	6	2	3	
ERBB2				
Positive	8	2	5	0.6896
Negative	43	15	26	
Grade				
G1-2	29	11	18	0.3889
G3	10	2	7	
T-stage				
1-2	34	12	21	0.333
3-4	12	3	8	
N-stage				
0	12	4	7	0.9077
1-3	33	11	21	
initial M-stage				
0	13	3	10	0.779
1	12	2	9	
At blood draw	N=97	CTC+ (n=31)	CTC- (n=68)	
Age (years)				
Mean	62	63	62	0.8006
Range	34-86	39-86	34-80	
Therapy				
Naïve	9	4	5	0.6062
Endocrine	37	10	27	
Chemo	53	17	35	
Stage				
Stable	37	6	31	0.0053
Progression	29	14	15	

Table 2 – Cox proportional hazard ratios. Estimated coefficients of overall survival on breast cancer subjects. Calculated are the corresponding hazard ratio (HR), 95% confidence interval (CI) of the hazard ratio, and p-value in uni- and multivariable Cox proportional hazard analysis for CTC-status with negative as reverence, and T-, N-, and M-stage at initial diagnosis.

Covariate	Univariable analysis					Multivariable analysis			
	Coefficient (b_i)	HR $[\exp(b_i)]$	HR 95% CI	p-value		Coefficient (b_i)	HR $[\exp(b_i)]$	HR 95% CI	p-value
CTC-positive	1.83	6.21	2.66, 14.47	0.0002		4.19	66.17	3.66, 1195.96	0.0045
T3-4 (rev: T1-2)	-0.17	0.84	0.31, 2.29	0.73		0.47	1.59	0.12, 20.54	0.72
N1 (rev: N0)	-0.30	0.74	0.27, 2.07	0.57		0.26	1.30	0.22, 7.73	0.77
M1 (rev: M0)	-0.71	0.49	0.15, 1.61	0.24		-1.92	0.15	0.01, 1.50	0.11
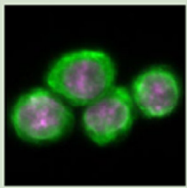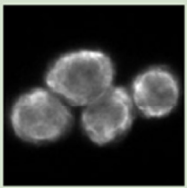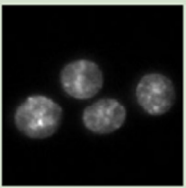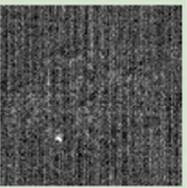
Figure 1 – Cell line cells. Image gallery of SK-BR-3 and MCF-7 cell line cells detected by the CellSearch System. K, keratin; ER, estrogen receptor.
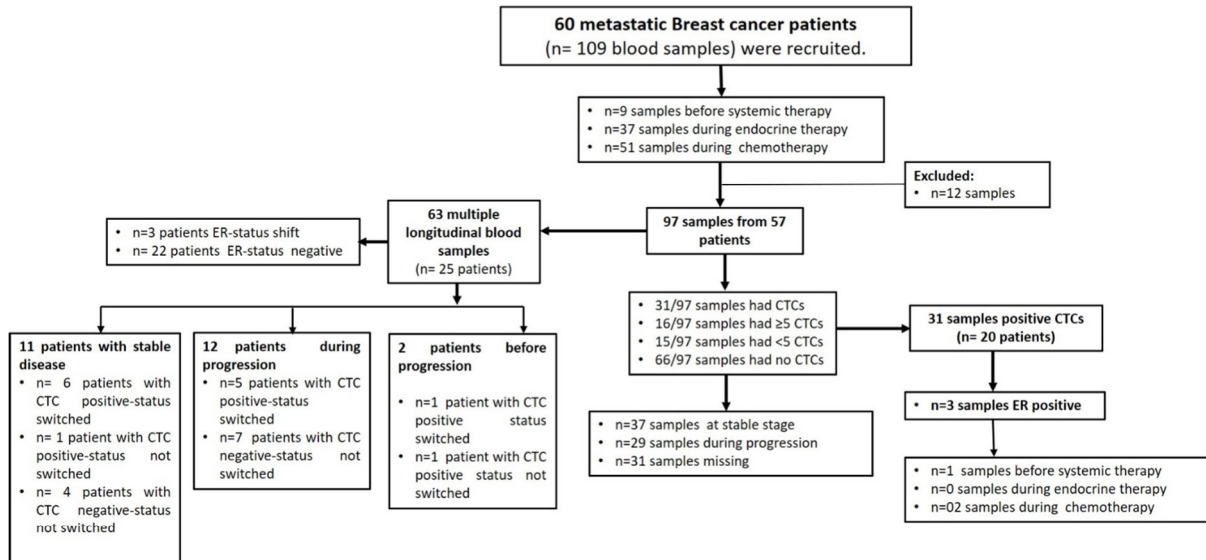
Figure 2 – Flow diagram of patient recruitment, exclusions and CTC status in metastatic breast cancer patients.
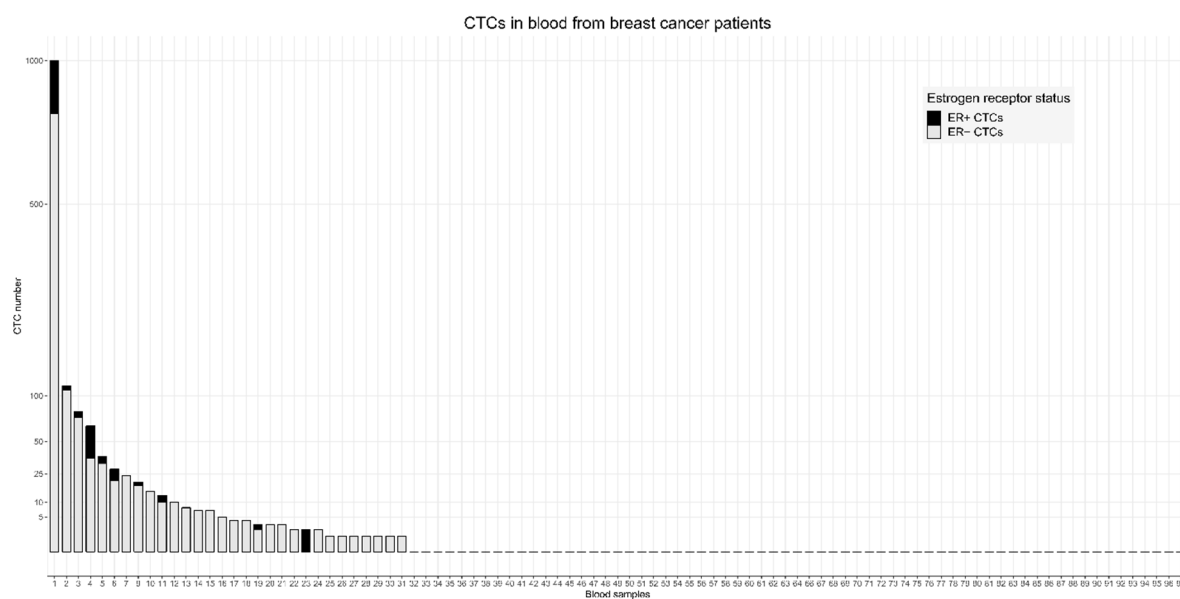
Figure 3 – Patient samples. Number of circulating tumor cells (CTCs) detected in blood samples from breast cancer patients. Black, number of ER (estrogen receptor) -positive CTCs; Gray, number of ER-negative CTCs.
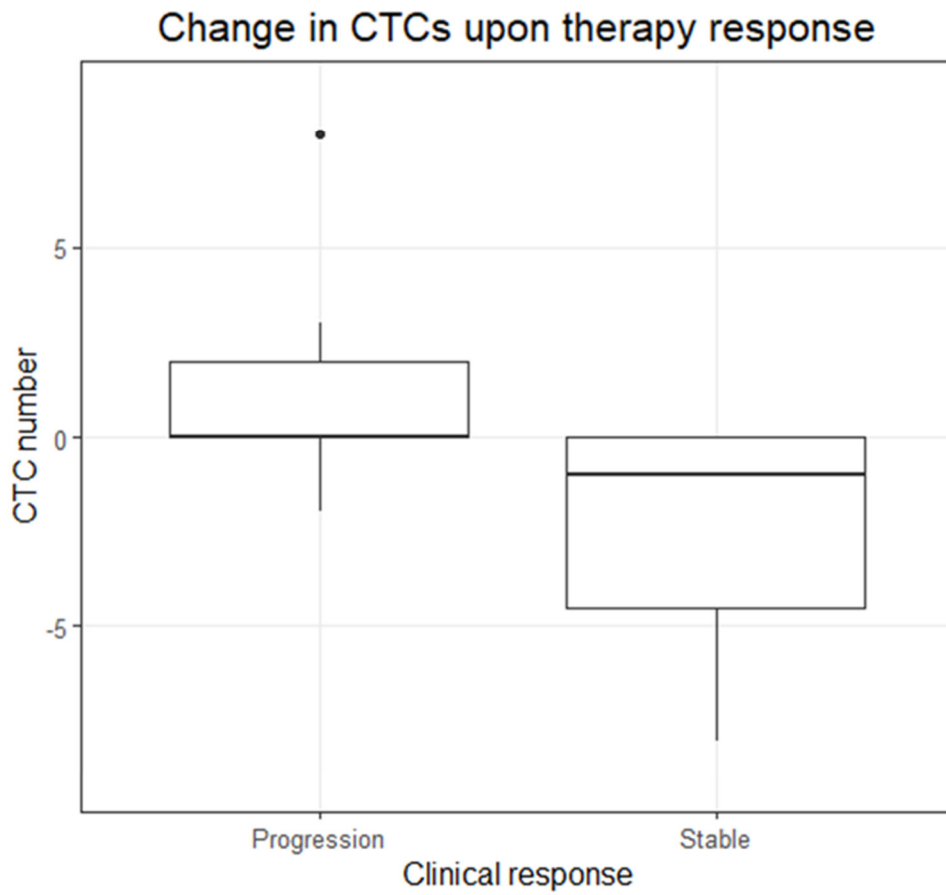
Figure 4 – CTC change. Boxplots showing the change in the number of circulating tumor cells (CTCs) at progression of disease and at stable disease.
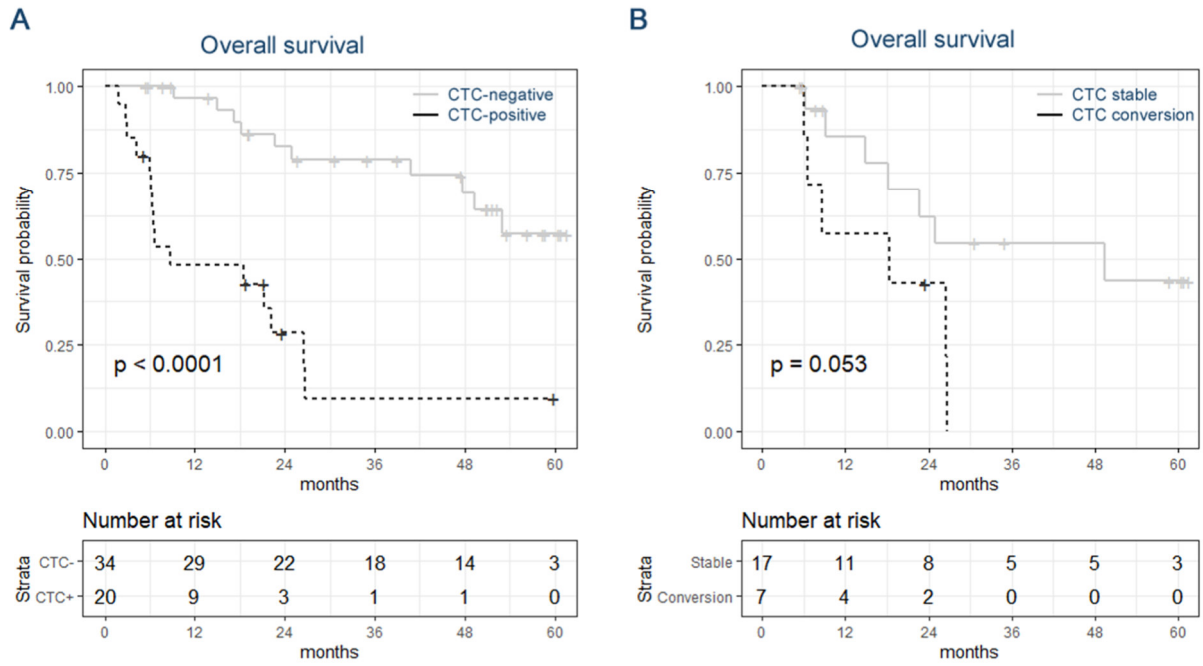
Figure 5 – Kaplan-Meier curves. Overall survival of breast cancer patients based on CTC-status (**A**) and based on stable CTC-status or CTC conversion (**B**).

5.4. Detection of *ESR1*, *PIK3CA*, *FOXA1,* and *GATA3* therapy resistance mutations in metastatic breast cancer patients using a MassARRAY-Based liquid biopsy Assay

Although many metastatic breast cancer patients initially respond to endocrine therapy, a large proportion of these patients develop resistance during systemic anti-estrogen therapies. Deep sequencing studies have highlighted the importance of acquired mutations of several genes, including the *ESR1*, *PI3KCA*, *FOXA1*, and *GATA3* in driving resistance to endocrine agents. A very promising marker is ctDNA in combination with CTCs, which can provide a comprehensive image of genetic make-up. Major hotspot mutations in *ESR1*, *PI3KCA*, *AKT1*, *ERBB2*, *TP53*, *FOXA1,* and *GATA3* were examined using MassARRAY-UltraSEEK® Breast panel in 272 blood samples from 101 metastatic breast cancer patients during the course of treatment. In the presented study, mutations were found in 82.8 % of metastatic breast cancer patients. The most frequent mutated genes were *ESR1*, *PIK3CA*, *FOXA1*, *GATA3*, and *TP53* in 53.5 %, 29.3 %, 31.3 %, 35.4 %, and 23.2 % of the patients, respectively. ctDNA and CTCs in longitudinal time-points showed a substantial level of intra- and inter-tumor heterogeneity involving polyclonal mutations with different clinical outcomes. During tumor progression, the *FOXA1* pE24K mutation was considerably raised upon chemotherapy and Fulvestrant treatment (p<0.0001), while *GATA3* pD336fs17 mutation significantly emerged in patients who received both endocrine agents and chemotherapy or chemotherapy alone (p<0.0001). The *GATA3* mutations

combined with *ESR1*, *FOXA1*, or *PIK3CA* mutations represent prognostic information for tumor progression and resistance to systemic therapy. Overall, liquid biopsy contributes in the identification of biomarkers that correlate to therapeutic targets and tumor progression that may improve breast cancer patients' management.

Publications

Detection of *ESR1*, *PIK3CA*, *FOXA1,* and *GATA3* therapy resistance mutations in metastatic breast cancer patients using a MassARRAY-Based liquid biopsy Assay

Maha Elazezy[1], Retno Ningsi[1], Alexander Sartori[2], Sabine Riethdorf[1], Harriet Wikman[1], Volkmar Müller[3], Klaus Pantel[1], and Simon A. Joosse[1]*

[1] Department of Tumor Biology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany.

[2] Agena Bioscience GmbH, 22761 Hamburg, Germany

[3] Department of Gynecology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany.

* Corresponding author (email: s.joosse@uke.de)

Abstract

Background: In metastatic breast cancer patients, the accumulation of *ESR1* and *PIK3CA* mutations resulted in resistance to endocrine treatment; however, the clinical value of these genes with *FOXA1* and *GATA3* is unknown. Therefore, we evaluated the clinical benefit of candidate mutations of these genes independently in circulating tumor DNA (ctDNA) and circulating tumor cells (CTCs).

Methods: In a prospective cohort of 101 patients with estrogen receptor-positive metastatic breast cancer, we assessed and monitored the candidate mutations in *ESR1*, *PIK3CA*, *FOXA1*, *GATA3*, *ERBB2*, *AKT1*, and *TP53* during the course of therapy using MassARRAY-UltraSEEK® technology and evaluated its clinical outcome.

Results: Mutations were detected in 82.8% of metastatic breast cancer patients. The *ESR1*, *PIK3CA*, *FOXA1*, *GATA3*, *ERBB2*, *AKT1*, and *TP53* genes were mutated in 53.5%, 29.3%, 31.3%, 35.4%, 3%, 5% and 23.2%, respectively of the patients. We identified clinically relevant somatic mutations during tumor progression, *FOXA1* (pE24K) significantly occurred upon chemotherapy and Fulvestrant treatment (p<0.0001), and *GATA3* (pD336fs17) was significantly raised upon both endocrine agents with chemotherapy or on chemotherapy alone (p<0.0001). A high degree of intra-and inter-tumor heterogeneity has been detected during longitudinal analysis. Emergence polyclonal mutations of *GATA3* in combining with *ESR1*, *FOXA1,* or *PIK3CA* were associated with tumor progression and worse overall survival (p<0.0001).

Conclusion: ctDNA and CTCs genotyping provide complementary prognostic information about disease progression and survival outcome. Mutations in the transcription factor *GATA3* combined with *FOXA1*, *ESR1* or *PIK3CA,* appear to be a hallmark of risk for estrogen receptor-positive metastatic breast cancer patients.

Keywords: Metastatic breast cancer; ctDNA; CTCs; mutations

1. Introduction

Breast cancer is one of the most prevalent cancers among women with an incidence of 13.3% and a mortality rate of 7.3% in Europe in 2020 [1]. Over 75% of breast cancer are estrogen receptor-positive (ER+) and thereby estrogen hormone-dependent for cell survival and proliferation [2, 3]. Hormone therapy, such as Tamoxifen, Fulvestrant, or Aromatase inhibitors (AIs), are the most preferred therapy to block estrogen receptor action or reduce estrogen concentration in primary and advanced disease settings [4-6]. Although, an initial response to hormone therapy can be observed in many metastatic breast cancer patients, a large proportion of these patients (30-40%) will develop therapy resistance [7].

Numerous mechanisms, including cell survival regulation and cell signaling pathways, have been implicated as the most resistance-acquired drivers. Additionally, mutations in *ESR1*, the gene encoding for ER, have been shown to negatively correlate with the effectiveness of anti-breast cancer therapy [8, 9]. The majority of *ESR1* mutations are in the ligand-binding domain (LBD) region containing p.D538G, p.Y537N, p.Y537C, p.E380Q, and p.L536R mutations that are more frequent in metastatic breast cancer patients and contribute to AI resistance [10]. Besides *ESR1* mutations, 30% of breast cancer patients often experience *PIK3CA* mutations [11]. The p.E542K and p.H1047R mutations are the most frequent driver mutations in *PIK3CA* [12]. These mutations contribute to the enhanced cell growth and serve as a predictor for response to treatment with PI3K/AKT/mTOR pathway inhibitors [13]. Increasing evidence correlates DNA binding affinity between *FOXA1*, *GATA3,* and *ESR1* [14]. *FOXA1* and *GATA3* are pioneer transcription factors involved in modulating chromatin condensation to permit ER recruitment in breast cancer cells. *FOXA1* and *GATA3* are frequently mutated in breast cancer, suggesting their contribution to endocrine-resistant breast cancer.[15, 16].

Liquid biopsy, such as circulating tumor cells (CTCs) and circulating tumor DNA (ctDNA) extracted from blood, is a novel minimally invasive approach to obtain longitudinally information about, e.g., the complete genetic make-up of a tumor [17-19]. MassARRAY liquid biopsy technology provides a high throughput mass spectrometry-based assay to screen a somatic mutation profile with a sensitivity of down to 0.1% [20-22]. In the study presented, we screened the major hotspot mutations in *ESR1, PI3KCA, FOXA1, GATA3, AKT1, ERBB2*, and *TP53* occurrence by using UltraSEEK® Breast panel in one hundred and one metastatic breast cancer patients during the course of treatment by MassARRAY® System and evaluating the clinical relevance of *FOXA1* and *GATA3* mutations in metastatic breast cancer patients.

2. Material and methods

2.1. Patients

In this prospective study, 101 metastatic breast cancer patients were recruited during 2015-2020. The patients were treated at the Department of Gynecologic Oncology, the University Medical Center Hamburg-Eppendorf, and received therapy according to international guidelines. All patients were with ER-positive primary tumor and selected based on the diagnosis of advanced disease with distant metastasis (Table 1). This study was approved by the local ethical board (ethical approval number: PV4367), and all patients enrolled into this study gave written informed consent.

2.2. Plasma Sample preparation and cfDNA isolation

One hundred seventy-six peripheral blood samples were collected into EDTA-containing tubes and processed within 1 hour at room temperature. After centrifugation at 360x *g* for 20 minutes, the plasma was transferred to a sterile 15 ml falcon tube and centrifuged again at 5087x *g* for 10 minutes at room temperature. The supernatant was transferred to a new sterile 15 ml falcon tube and stored at −20°C until cfDNA isolation. cfDNA was isolated by Qiagen QLAamp Circulating Nucleic Acid kit (cat. no. 55114, Qiagen, Hilden, Germany) according to the manufacturers' instructions [23]. cfDNA quantification was assessed using Qubit® 2.0 Fluorometer dsDNA HS Assay Kits (Thermo Fisher Scientific, Eugene, OR, USA) [24].

2.3. CTC enrichment and DNA isolation

Ninety-six blood samples were collected into CellSave Preservative tubes, CTCs were captured and enumerated from 7.5 ml blood by the CellSearch[TM] system using the CELLSEARCH® CXC Kit® (cat. No. 7900017) [25]. After CTCs enrichment, cells of all samples (e.g., CTC positive and CTC negative samples) were transferred from CellSearch-Cartridge into PCR tubes for DNA isolation [26]. DNA was isolated by using QLAamp DNA Micro Kit (cat. no. 56304, Qiagen, Hilden, Germany) according to manufacturers' instructions. DNA was quantified using Qubit® 2.0 Fluorometer dsDNA HS Assay Kits (Thermo Fisher Scientific, Eugene, OR, USA) [24].

2.4. Mutation Profiling with the UltraSEEK Breast Panel

cfDNA and CTCs were screened using the custom UltraSEEK® Breast and *GATA3*/*FOXA1* v1.0 panels (Agena Bioscience, San Diego, USA). The panels test for 50 mutations across 12 multiplex reactions in the genes *ESR1* (E380Q, S576L, V392I, L536Q, K303R, Y537S, Y537C, Y537N, D538G, S463P, L536R), *PIK3CA* (N345K, E542K, E545K, E545A, E545Q, C420R, H1017R, H1047R), *AKT1* (L52R, E17K), *ERBB2* (S310F, G309E, L755_T759del, L755R, L869R, G309A, L755S, D769H, V777L, D769Y), *TP53* (R248Q, R273C, R213X, R248W, R273H, Y220C, R175H), F*OXA1* (I176V, I176M, F266L, D226N, E42K, S250F), and *GATA3* (R365G, D336fs17, S93F, M294K, P409fs537, S137L). Initial amplification of the target regions that harbor the mutations of interest was done by two multiplex PCR reaction steps according to manufacturers' instructions. The PCR products were treated with shrimp alkaline phosphatase for dephosphorylation of dNTPs and primer digestion. Next, a single-base extension reaction with biotinylated chain terminator nucleotides specific to the mutant allele was performed in 12 multiplex reactions. The mutant-specific extension products were enriched by streptavidin-coated magnetic beads following the manufacturer's instructions, then transferred to the automated MassARRAY® System with Chip Prep Module 96 (CPM96). Data were acquired via matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry using the MassARRAY Analyzer.

2.5. Statistical analysis

Data analysis of MassARRAY was performed using Typer software version 4.0.26.74 (Agena Bioscience, San Diego, USA). Normalized intensity was calculated for the signal intensity of the mutant allele, which had been normalized against the capture control peaks found in the spectrum [27]. A value of one represents the peak intensity of the observed mutant allele equal to the peak intensity of the average of the five capture control peaks found in the spectrum [27, 28]. The capture control peaks are biotin-labeled, non-reactive oligos, which are added to the extension reaction and used as an internal control for the streptavidin-bead capture and elution of the mutant extension product steps. Graphs were generated performing GraphPad Prism (GraphPad Software, San Diego, CA, USA). Statistical significance was defined as $p < 0.05$. Survival analysis was determined using a Kaplan–Meier curve and logrank test. The primary endpoint of progression-free survival (PFS) and overall survival (OS) was defined as the time in months from a blood sample taken until the first progression and/or cancer-related death, according to REMARK [29].

3. Results

3.1. Study patient cohort and sample material

Two hundred seventy-two blood samples were obtained at 185 time-points from 101 metastatic breast cancer patients during this study. Two patients were excluded due to missing clinical data and pre-analytical errors (Figure 1). Five percent (10/183) of samples were collected before the initiation of systemic therapy, 16% (30/183) of samples were collected during endocrine therapy, 45% (82/183) of samples were taken during chemotherapy, 34% (63/183) of samples were obtained during both endocrine and chemotherapy, and 2.7% (5/183) of samples failed in the analysis. One hundred twenty-five multiple longitudinal blood samples were obtained from 41 patients during disease (Figure 1). The median follow-up was 20.3 months (range, 1-64.2), starting from the time point of blood analysis till the end of follow-up. At primary diagnosis, 96% of patients were diagnosed with ER-positive primary breast cancer; 11% of patients had ERBB2-positive, and 11% had primary metastasis. At blood draw, 75% of patients had distant metastasis (Table 1).

cfDNA was isolated from one hundred seventy-six blood samples. The median concentration of total cfDNA obtained from all patients' blood samples was 1085 ng/ml plasma at all time points (range, 104-76000 ng/ml; Supplementary Figure 1A). No significant differences were observed between the median cfDNA concentrations of cfDNA of patients who had no detectable mutations (wild-type (WT)) 888 ng/ml plasma and mutant cfDNA 1230 ng/ml plasma (p=0.4087, Welch's t-test; Supplementary Figure 1B). Simultaneously to plasma DNA isolation, 96 blood samples from 57 patients were investigated for the presence of circulating tumor cells (CTCs) using the CellSearch system. CD45-/ Epcam+/ Keratin+ cells with an intact nucleus were interpreted as CTCs (Supplementary Figure 1C). Three samples and three patients were excluded due to pre-analytical errors. In the total of 93 samples, CTCs were detected in 31% (29/93) of samples from 43.6% (24/55) of patients; 16 samples from 14 patients had ≥5 CTCs, 13 samples from 10 patients had <5 CTCs (Supplementary Figure 1D), and 68.8% (64/93) of samples were CTC negative from 55.5% (30/54) of patients. Forty-eight CTC and matched cfDNA samples were obtained from 54 patients.

3.2. Comprehensive detection of mutations in advanced breast cancer patients using UltraSEEK breast panel.

By performing MassARRAY technology, we screened 50 hotspot mutations in seven breast cancer-related genes (Supplementary Figure 2). The *ESR1* somatic mutations are in the ligand-binding domain (LBD) region (Figure 2A), *PIK3CA* mutations are in the helical and kinase domains (Figure 2B), while *FOXA1* mutations are in forkhead N and DNA binding domains (Figure 2C), *GATA3* candidate mutations are in zinc fingers domains (Figure 2D), in addition to *AKT1*, *ERBB2*, and *TP53* hot spot mutations.

The presence of somatic mutations in cfDNA was analyzed in 267 blood samples at 183 time-points from 99 metastatic breast cancer patients using the UltraSEEK breast panel. We could detect mutations in 82.8% (82/99) of metastatic breast cancer patients who harbored at least one of the candidate mutations (Figure 3). Among the 99 patients 53.5% (53/99) were diagnosed with mutations in *ESR1*, 29.3% (29/99) in *PIK3CA*, 31.3% (31/99) in *FOXA1*, 35.4% (35/99) in *GATA3*, 5% (5/99) in *AKT1*, 3% (3/99) in *ERBB2*, and 23.2% (23/99) in *TP53* (Figure 4A). The median variant allele frequency (VAF) of mutated genes was 0.525% (range 0.1-2.9%; Figure 4B). In circulating *ESR1* mutant DNA of metastatic breast cancer patients, pD538G and pY537S mutations were found in 44.4% (44/99) and 6% (6/99), respectively (Figure 4C). The median VAF of *ESR1* mutations was 1% (range 0.1-2.9%) (Figure 4B), pY537S and pD538G were with the highest VAF of 2.1% (range 0.3-2.9%) and 1.1% (range 0.4-2.7%) respectively (Figure 4D). Three *PIK3CA* hotspot mutations located in kinase and helical domains, pH1047R 11% (11/99), pE542K 11% (11/99), and pE545K 6% (6/9) were frequently detected in circulating DNA of metastatic breast cancer patients (Figure 5A). The median VAF of *PIK3CA* mutations was 0.8% (range 0.2-2.7%) (Figure 4E).

In *FOXA1*, the fork-head and DNA binding domains were the most commonly mutated region detected in circulating DNA, which harboring pE24K 11% (11/99), pI176M 11% (11/99), and pI176V 8% (8/99) mutations (Figure 4F). The median VAF of *FOXA1* mutations was 0.3% (range 0.1-2.1%) (Figure 4B). Frameshift pD336fs17 in *GATA3* was detected in 35% of patients (Figure 5G) at VAF of 0.4% (range 0.2-1.4%) in circulating DNA (Figure 4D). Somatic alterations were also detected in *AKT1* (pE17K (5%); Figure 4H), *ERBB2* (pS310F (1%) and pL755-T759del (1%); Figure 4I), *TP53* (pY220C (18%) and pR175H (2%); Figure 4J).

3.3. Mutation analysis of cfDNA and CTCs of patients with metastatic breast cancer.

One hundred seventy-four cfDNA samples from 97 patients were analyzed. The *ESR1*, *PIK3CA*, *FOXA1,* and *GATA3* genes were found to be mutated in cfDNA with a frequency of 33% (32/97), 28% (27/97), 27% (26/97), and 20% (19/97), respectively (Figure 4K). *ESR1* and *PIK3CA* mutations had a high VAF of 0.7% (range 0.1-2.9%) and 0.8% (range 0.2-2.7%), respectively (Figure 4L and Supplementary Figure 3A). In CTCs, 93 samples of 54 patients were analyzed. We detected mutation in 86% (25/29) of CTCs positive samples and in 62.5% (40/64) of CTCs negative samples. The *ESR1*, *GATA3*, and *TP53* mutations were the most frequently detected mutations in DNA derived from CTCs of patients with 67% (36/54), 41% (22/54), and 37% (20/54), respectively (Figure 4M). The VAF of *ESR1* and *GATA3* were was 1.2% (range 0.3-2.4%) and 0.45% (range 0.2-1.4%), respectively (Figure 4N). The pD538G in *ESR1* and pD336fs17 in *GATA3* were more frequent detected mutations in CTCs with VAF of 1.25% (range 0.4-2.4%) and 0.5% (range 0.2-1.4%), respectively (Supplementary Figure 3B). Comparing the detected mutations in cfDNA with matched DNA derived from CTCs of the same patients and its clinical value, we found an overlap in 36.5% (19/52) of patients. In total, the rate of detected mutations in CTCs enriched samples was 76.9% (40/52) and 65% (34/52) in cfDNA samples (Figure 4O). A high degree of genomic heterogeneity was observed in ctDNA compared to CTCs (Figure 4O); nevertheless, no big difference in clinical statistical significance was detected between ctDNA and CTCs (Supplementary Figur 4A and 4B). Therefore, we combined ctDNA and CTCs mutation profiles to provide additional genomic information of metastatic tumor and evolving genetic signatures during disease progression.

3.4. Association of frequently detected mutations with disease progression and systemic therapy in metastatic breast cancer patients.

Among 183 samples, 24.5% (45/183) were obtained from 34 patients diagnosed with the stable disease during the study, 12.6% (23/183) of samples were taken from 9 patients before they were diagnosed with progression. Whereas 53.5% (98/183) of samples were recruited from 44 patients after the diagnosis of progression, and 10.9% (20/183) of samples were collected from 8 patients at the progressive stage of the disease.

Of 82 patients with detectable mutations, 25.6% had a single mutation mainly detected at a stable disease stage. While 74.4% (61/82) of the patients showed a polyclonal mutation during disease progression or before. Out of 61 patients, 57% had two mutations detected either at the same gene or in two different genes, 29.5% (18/61) of cases showed three mutations detected during disease progression or before, and 15% (8/61) of patients had four mutations detected at a progressive stage of patients (Figure 5A and Supplementary Figure 5).

In metastatic breast cancer patients, the *ESR1* gene was mutated in 43.2% (79/183) of circulating DNA samples. Including pD538G, pY537S, and pY537N mutations, were detected in 81% (64/79), 11% (9/79), and 3.8% (3/79) of *ESR1* mutated samples, respectively. These mutations significantly emerged in patients who had received aromatase inhibitor and/or chemotherapy (p<0.0001; Supplementary Figure 6A) during disease progression and progressive stage of disease (p<0.0001; Supplementary Figure 6B). In addition, pE380Q mutation was identified in 5% (4/79) of *ESR1* mutated samples and was associated with patients treated with Fulvestrant therapy and chemotherapy (R^2=0.2415; p<0.0001).

PIK3CA mutations were found in 23% (42/183) of samples, including pH1047R, pE542K, and pE545K mutations were detected in 38% (16/42), 35.7% (15/42), and 21.4% (9/42), respectively of *PIK3CA* mutated samples. The pH1047R and pE542K mutations were significantly associated with patients who received chemotherapy alone or in combination with endocrine therapy (R^2=0.2053; p=0.0020), while pE545K mutations significantly emerged in patients under both endocrine agents and chemotherapy (p<0.0001; Supplementary Figure 7A). The pH1047R, pE545K, and pE542K mutations significantly occurred during disease progression (p<0.0001; Supplementary Figure 7B).

FOXA1 mutations were observed in 22.4% (41/183) of circulating DNA samples. The pE24K, and pI176V and, pI176M mutations were detected in 36.5% (15/41), 36.5% (15/41), and 26.8% (11/41) of *FOXA1* mutated samples, respectively. The pE24K and pI176V mutations were significantly occurred upon Fulvestrant and/or chemotherapy (p<0.0001; Figure 5B and 5C) during the disease progression and at a progressive phase (R^2=0.1646; p=0.0077; Figure 5D and 5E). While the pI176M mutation was significantly observed during the progression of disease and in patients who received chemotherapy alone or combined with aromatase inhibitor (p<0.0001; Figure 5F and 5G).

GATA3 mutations were detected in 27.9% (51/183) of circulating DNA samples, pD336fs17 and pS93F mutations were observed in 92% (47/51) and 9.8% (5/51) of *GATA3* mutated samples, respectively. The pD336fs17 mutation frequently emerged during the disease progression and at a progressive phase of the disease, particularly in patients treated with chemotherapy alone or combined with endocrine agents (p<0.0001; Figure 5H and 5I). Additionally, pS93F mutation was significantly raised in patients who received chemotherapy during disease progression and at a progressive phase of disease (p<0.0001; Figure 5 J and 5K).

Further analysis showed that *ESR1* and *PIK3CA* were mutated in 11% of patients (Figure 5L) with no significant difference in a mutation load observed (p=0.9966; Figure 5M). However, patients with a high mutation burden of *ESR1* exhibited a lower load of *PIK3CA* mutations and vice versa (Figure 5M). Interestingly, we found that 17% of patients had a mutation in both *ESR1* and *GATA3* genes, and 9% of patients exhibited a mutation in both *ESR1* and *FOXA1* genes, while 7% of patients had a mutated *ESR1, FOXA1,* and *GATA3* fusion genes (Figure 5N). A significant difference in mutation load was observed between patients with *ESR1* and *GATA3* or patients with *ESR1* and *FOXA1* mutations (P<0.0001 and p=0.0002, respectively; Figure 5O and 5P). Overall, *FOXA1* mutations appeared significantly during disease progression (R^2=0.295, p=0.0030), while *GATA3* mutations significantly occurred during disease progression (R^2=0.202, p=0.0446) and at the progressive stage of disease (R^2=0.272, p=0.0064; Supplementary Figure 5).

3.5. Monitoring hotspot mutations using serial cfDNA and CTCs samples from metastatic breast cancer patients.

One hundred eighty-seven multiple longitudinal blood samples at 125 time-points were analyzed from 41 patients who experienced disease progression. Among 125 time points, 63 samples (50.4%) were collected during the chemotherapy course, 8 samples (6.4%) were taken during endocrine therapy, 47 (37.6%) samples were obtained during the systemic course of endocrine agents in combination with chemotherapy, and five samples failed in analysis. In the chemotherapy cohort, 25% of blood samples were collected before patients were diagnosed with progression, and 48.4% of blood samples were obtained after and during a tumor is progressed. Before tumor progression, the *ESR1* (37.5%) and *GATA3* (43.8%) mutations were the most common. (Figure 6A). While after and during the patients were diagnosed with disease progression, the *ESR1* (38.7%), *PIK3CA* (25.8%), *FOXA1* (32.3%), *GATA3* (22.6%), and *TP53* (22.6%) mutations were the most detected (Figure 6B). In the endocrine therapy cohort, 73% of blood samples were received after patients were diagnosed with progression. The most mutated genes raised during the endocrine therapy course were ESR1(62.5%), FOXA1 (37.5%), GATA3 (25%), and TP53 (25%) mutations (Figure 6C). At advanced settings, *ESR1* and *GATA3* mutations were significantly correlated to a progressive phase of the disease (Figure 6D).

The median follow-up time of metastatic breast cancer patients was 20.2 months. Throughout follow-up visits (Figure 6E), *ESR1* (R^2=0.7048, p=0.0091), *FOXA1*(R^2=0.6228, p=0.0199), and *GATA3* (R^2=0.7040, p=0.0092) mutations were significantly raised along with disease progression. Figure 6F and 6G show the disease progression in two patients, including the concentration of the tumor markers CA15-3 and CEA assessed for regular diagnostics, the systemic therapy provided, the disease status, and the variant allele frequency of detected mutations. Both patients had distant metastasis in bone,

liver and /or lung. In the first patient, we tracked the profile of somatic mutations in seven serial blood samples throughout the course of treatment. *FOXA1* (pI176V) mutation was present at 0.2% VAF in the second follow-up sample and was absent in the following samples. *ESR1* (pD538G), *PIK3CA* (pE542K), and *GATA3* (pD336fs17) emerged during follow-up samples, and the mutation frequency increased during the treatment (Figure 6F). In the second patient, we monitored the somatic mutations in three serial blood samples during the progressive stage of disease at which the patients were treated with chemotherapy. *ESR1* (pD538G) and *GATA3* (pD336fs17) were detected during the treatment course at 0.5% and 0.3% VAF, respectively which increased to 1.8%(*ESR1*) and 0.6% (*GATA3*) VAF as the disease progressed (Figure 6G).

3.6. Survival analysis

PFS and OS analysis were performed to test the difference in survival between metastatic breast cancer patients who harbor *ESR1*, *PIK3CA*, *FOXA1*, and *GATA3* mutations compared to wild-type (WT) circulating DNA of patients who had no detectable mutations. We considered each time point as an event; In total, 183 events were recorded during the study.

Firstly, we analyzed the outcome of candidate mutations independently. Starting with patients whose circulating DNA harbor *ESR1* mutations, we found that patients with pY537N mutation showed significantly worse PFS (median: 5 months; p=0.0008) than WT circulating DNA (Figure 7A). We also observed that patients with pY537S (median: 15 months; p<0.0001) and pY537N (median: 18 months; p=0.0016) mutations had significantly shorter OS than patients with pD538G (median: 45 months; p=0.0154) mutations compared to WT cohort (Figure 7B). In patients with *PIK3CA* mutations, we found that patients with pH1047R (p=0.0017) mutation had worse PFS than the WT cohort (Figure 7C). At the same time, OS analysis showed that patients with pE545K (median: 18 months; p<0.0001) and pH1047R (median: 19 months; p<0.0001) mutations had significantly shorter OS than patients with pE542K (median: 24 months; p=0.0172) mutation compared to WT cohort (Figure 7D and Supplementary Table 1). Furthermore, in patients who had a mutant *FOXA1* fusion gene, we found that patients with pE24K mutation had significantly worse PFS (median: 5 months; p<0.0001) and OS (median: 8 months; p=0.0014) compared to WT circulating DNA (Figure 7E and 7F). Also, patients who had pI176V (median: 17 months; p<0.0001) mutation were associated with worse outcomes compared to the WT group (Figure 7F). Finally, patients with *GATA3* mutations who had frameshift pD336fs17 (p<0.0001) and pS93F (p=0.0256) mutations revealed significantly worse PFS than the WT cohort (Figure 7G). Also, those patients had shown a significantly shorter OS (median: 14 months;

p<0.0001 and median: 19 months; p=0.0276), respectively, compared to the WT cohort (Figure 7H and Supplementary Table 1).

Next, we analyzed the survival benefit of single and polyclonal mutant genes. Surprisingly, metastatic breast cancer patients whose circulating DNA harbor only *ESR1* mutations showed no significant difference in survival rate for PFS (p=0.948) and OS (p=0.819) compared to the WT circulating DNA cohort (Figure 8A, and 8B). Also, patients with *PIK3CA* mutations were not significantly for PFS (p=0.0503), but they showed shorter OS (median: 19.2 months; p=0.0002) than WT. Whereas metastatic breast cancer patients who had either *GATA3* (p=0.0048) or *FOXA1* (p=0.0051) mutations were more prone to tumor progression (Figure 8A) and had a shorter OS (median: 14.9 months; p<0.0001 and median: 18 months; p=0.0034) respectively, compared with patients who had *ESR1* mutations and/or WT cohort (Figure 8B, and Supplementary Table 2).

A significant reduction in PFS and OS survival median was observed in patients who had emerged more than one mutated gene compared to patients with a single mutated gene. In particular, patients with *GATA3* mutated subclones that appeared with other mutated genes had a shorter survival rate than other polyclonal mutant genes (Supplementary Figure 8). We found that patients who had either both *ESR1, FOXA1,* and *GATA3* or *ESR1*, *PIK3CA*, and *GATA3* mutant showed shorter PFS (median: 7 months; p=0.0002 and median: 12.5 months; p=0.017, respectively) and OS (median: 19 months; p=0.0047 and median: 17 months; p<0.0001, respectively) than patients who had *ESR1* and *FOXA1* or *ESR1* and *PIK3CA* only (Figure 8C and 8D).

Furthermore, patients with *FOXA1* and *GATA3* mutations were progressed significantly faster (median: 2.9 months; p<0.0001) than other polyclonal groups such as *ESR1*, *FOXA1,* and *GATA3* mutations (median: 7 months; p= 0.0022) or patients with *ESR1* and *GATA3* mutations (p=0.0016; Figure 8E). At the same time, patients with *FOXA1* and *GATA3* mutations had significantly emerged worse OS (median: 9 months; p<0.0001) than patients with combined *ESR1*, *FOXA1,* and *GATA3* (median: 19 months; p=0.0047) compared to WT cohort (Figure 8F and Supplementary Table 2).

4. Discussion

In this prospective study, we screened and monitored 50 different candidate mutations in *ESR1*, *PIK3CA*, *FOXA1*, *GATA3*, *AKT1*, *ERBB2*, and *TP53* genes using MassARRAY-UltraSEEK® technology and investigated the prognostic impact of detectable mutations in ctDNA and DNA-derived CTCs of estrogen receptor-positive metastatic breast cancer patients. The high sensitivity and specificity of the UltraSEEK® Breast Cancer Panel were recently verified [30]. The most crucial advantage of MassARRAY®-UltraSEEK® technology is that it can perform in ctDNA and CTCs, fast and cost-effectively.

During disease progression and prolonged exposure to hormonal and chemotherapy, subclones with different candidate point mutations are chosen and enriched in metastatic tumor settings. Development of the *ESR1* and *PIK3CA* mutations has recently been identified as an acquired mutational process that contributes to ER-positive breast cancer resistance and metastasis. In circulating DNA, *ESR1* and *PIK3CA* were mutated in 53% and 29% of breast cancer patients who experienced progression, respectively. The pY537S and pY537N mutations in *ESR1* were a good prognostic factor in predicting tumor progression upon Aromatase inhibitor and/or chemotherapy and were associated with shorter overall survival. Among the *PIK3CA* mutations, pH1047R, pE542K, and pE545K were the most frequently observed mutations during tumor progression, which are in accord with previously reported data [11, 12, 31-34].

Nevertheless, 26.8% of patients whose tumors progressed did not have *ESR1* and *PIK3CA* mutations indicating an alternative resistance mechanism involved in tumor progression upon systemic therapy. We found that these patients had a cumulation of *FOXA1* and/or *GATA3* mutations. Recently, restoration of ER function by recruitment transcription factors, particularly *FOXA1* and *GATA3*, has been highlighted as one of the acquired resistance mechanisms in ER-positive breast cancer [14, 35]. *FOXA1* and *GATA3* mutations were detected in 31% and 36%, respectively, of metastatic breast cancer patients. Wide variation in frequency distributions rate of *FOXA1* and *GATA3* mutations in breast cancer was observed, probably due to different detection techniques, cohort selection criteria, detectable mutation site, and material sources (i.e., tissue, cfDNA, and CTCs) [35-39]. The pI176V and pE24K mutations in *FOXA1* frequently occurred in metastatic breast cancer patients. The pI176V was a cluster in the winged-helix domain and was associated with worse outcomes. The functional role of winged domain mutations has been found to be a hallmark of risk for breast cancer by increasing the chromatin binding affinity at the estrogen receptor sites [35]. Whereas the pE24K mutation was in the forkhead-N domain, it was significantly correlated to a progressive stage of tumor and associated with poor PFS and OS outcome of patients. However, this mutation was reported in COSMIC from three comprehensive studies on breast cancer; its functional role in breast cancer is not yet clear [40-42]. The majority of *GATA3* mutations were located in the second zinc finger and C-terminus domains [43]. In the current study, frameshift pD336fs17 mutation in the second zinc finger domain correlated to disease progression and poor outcomes. A recent study showed that frameshift pD336fs17 mutation was contributed to a growth advantage of breast cancer [43]. We found in our study that *GATA3* mutations were significantly associated with chemotherapy regimens which were consistent with Tominaga *et al.* [44]. In contrast, Gustin *et al.* found that *GATA3* mutations did not impact the sensitivity to endocrine therapy or chemotherapeutic agents [43]. A strong relationship between *FOXA1*, *GATA3,* and *ESR1* has been reported in several studies [14, 45, 46]. One of the observed findings was a significant positive correlation between *ESR1* and *GATA3* mutations, which were frequently detected in CTCs than ctDNA.

In accordance with Davis and colleagues found that a high number of CTCs were strongly related to *ESR1* and *GATA3* mutations [47], which may contribute to the metastatic process of breast cancer.

A high degree of intratumor heterogeneity in the setting of disease progression or drug regimens was observed. Patients with a single mutant gene had a more favorable survival than those with two or more mutant genes. As shown by patients who had *ESR1* mutant gene did not vary from the WT cohort in terms of PFS and OS; during the study, most of these tumors were diagnosed as disease stable and had a single mutation of *ESR1*. We observed that 76% of *ESR1* mutant patients had pD538G mutation, which was associated with a better outcome. This finding possibly explained that not all mutations are equivalent and have the same impact on breast tumor progression and survival outcome. As well as, *PIK3CA* mutations were not a good predictor of PFS, but they showed significant prognostic information for OS. On the other hand, circulating tumor DNA of patients who emerged subclones of *FOXA1* or *GATA3* mutations, their tumors were more prone to progress. They also showed worse overall survival, indicating the risk factor of *FOXA1* and *GATA3* mutations in breast cancer.

Another interesting finding is that the accumulation of *GATA3* mutated subclones combined with *ESR1* and *FOXA1* mutations showed fast progression and worse outcomes than single mutated genes. Surprisingly, survival benefits were observed in patients who carry subclones mutations of *ESR1* and *FOXA1*. We found that 66.7% of this cohort carried pD538G and pI176M mutations that did not significantly impact on tumor progression and had a better survival rate than other subclonal mutations. Whereas patients whose tumor encompasses *FOXA1* and *GATA3* mutations progressed faster in 2.5 months and had a worse survival rate at nine months than other subclonal cohorts. In this cohort, 71% of patients encompassed subclones of pE24K and pD336fs17 mutations. Both mutations provided useful prognostic information about the tumor progression over chemotherapy as well as overall survival probability. Taken all together, the capacity to detect these mutations non-invasively might lead to a more personalized selection of effective treatment options in the future. ctDNA and CTCs serve as a source of tumor cells for genotyping and facilitate the longitudinal analysis of genomic heterogeneity. Our findings suggest the value of combining CTCs with ctDNA-derived biomarkers. Finally, mutations in *FOXA1* and *GATA3* transcription factors provide insight into the underlying therapy resistance that may directly influence treatment decision-making and cancer prognosis.

5. Conclusion

Breast tumors are comprised of sub-clonal populations with different mutations and clinical outcomes. ctDNA and CTCs in longitudinal time points allowed to identify hallmark subclones that have a prognostic value in tumor management. The *GATA3* with *PIK3CA*, *FOXA1,* and/or *ESR1* represent a high risk of metastatic breast cancer patients and a good predictor of tumor resistance and progression.

References

1. JRC and the European Network of Cancer Registries, t.w.t.I.A.f.R.o.C.I., *2020 Cancer incidence and mortality in EU-27 countries*. 2020, JRC
2. *American Cancer Society. Breast Cancer Facts & Figures 2019-2020. Atlanta, GA: American Cancer Society; 2019.*
3. Vella, V., et al., *Microenvironmental Determinants of Breast Cancer Metastasis: Focus on the Crucial Interplay Between Estrogen and Insulin/Insulin-Like Growth Factor Signaling.* Frontiers in Cell and Developmental Biology, 2020. **8**(1458).
4. Neven, P., et al., *Tamoxifen Metabolism and Efficacy in Breast Cancer: A Prospective Multicenter Trial.* Clin Cancer Res, 2018. **24**(10): p. 2312-2318.
5. Osborne, C.K., A. Wakeling, and R.I. Nicholson, *Fulvestrant: an oestrogen receptor antagonist with a novel mechanism of action.* Br J Cancer, 2004. **90 Suppl 1**(Suppl 1): p. S2-6.
6. Campos, S.M., *Aromatase inhibitors for breast cancer in postmenopausal women.* Oncologist, 2004. **9**(2): p. 126-36.
7. Anurag, M., M.J. Ellis, and S. Haricharan, *DNA damage repair defects as a new class of endocrine treatment resistance driver.* Oncotarget, 2018. **9**(91): p. 36252-36253.
8. Guttery, D.S., et al., *Noninvasive detection of activating estrogen receptor 1 (ESR1) mutations in estrogen receptor-positive metastatic breast cancer.* Clin Chem, 2015. **61**(7): p. 974-82.
9. Feng, Y., et al., *Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis.* Genes Dis, 2018. **5**(2): p. 77-106.
10. Rani, A., et al., *Endocrine Resistance in Hormone Receptor Positive Breast Cancer-From Mechanism to Therapy.* Frontiers in endocrinology, 2019. **10**: p. 245-245.

Publications

11.   Martinez-Saez, O., et al., *Frequency and spectrum of PIK3CA somatic mutations in breast cancer.* Breast Cancer Res, 2020. **22**(1): p. 45.

12.   Shimoi, T., et al., *PIK3CA mutation profiling in patients with breast cancer, using a highly sensitive detection system.* Cancer Sci, 2018. **109**(8): p. 2558-2566.

13.   Wang, M., et al., *The Predictive Role of PIK3CA Mutation Status on PI3K Inhibitors in HR+ Breast Cancer Therapy: A Systematic Review and Meta-Analysis.* BioMed research international, 2020. **2020**: p. 1598037-1598037.

14.   Theodorou, V., et al., *GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility.* Genome research, 2013. **23**(1): p. 12-22.

15.   Hurtado, A., et al., *FOXA1 is a key determinant of estrogen receptor function and endocrine response.* Nature genetics, 2011. **43**(1): p. 27-33.

16.   Albergaria, A., et al., *Expression of FOXA1 and GATA-3 in breast cancer: the prognostic significance in hormone receptor-negative tumours.* Breast Cancer Research, 2009. **11**(3): p. R40.

17.   Alix-Panabières, C. and K. Pantel, *Clinical Applications of Circulating Tumor Cells and Circulating Tumor DNA as Liquid Biopsy.* Cancer Discov, 2016. **6**(5): p. 479-91.

18.   Elazezy, M. and S.A. Joosse, *Techniques of using circulating tumor DNA as a liquid biopsy component in cancer management.* Computational and structural biotechnology journal, 2018. **16**: p. 370-378.

19.   Keller, L., et al., *Clinical relevance of blood-based ctDNA analysis: mutation detection and beyond.* Br J Cancer, 2021. **124**(2): p. 345-358.

20.   Lamy, P.J., et al., *Mass Spectrometry as a Highly Sensitive Method for Specific Circulating Tumor DNA Analysis in NSCLC: A Comparison Study.* Cancers (Basel), 2020. **12**(10).

21.   Gorges, K., et al., *Intra-Patient Heterogeneity of Circulating Tumor Cells and Circulating Tumor DNA in Blood of Melanoma Patients.* Cancers (Basel), 2019. **11**(11).

22.   Belloum, Y., et al., *Discovery of Targetable Genetic Alterations in NSCLC Patients with Different Metastatic Patterns Using a MassARRAY-Based Circulating Tumor DNA Assay.* Cells, 2020. **9**(11).

23.   *QIAamp® Circulating Nucleic Acid Handbook, HB-0202-006_HB_QA_ Circulating NucAcid_0819_WW.pdf, 2019.* https://www.qiagen.com/it/resources/resourcedetail?id=0c4b31ab-f4fb-425f-99bf-10ab9538c061&lang=en.

24.   *THE QUBIT® 2.0 FLUOROMETER.bioprobes-64-qubit, 2011.* https://www.thermofisher.com/content/dam/LifeTech/migration/en/filelibrary/support/bioprobes/bioprobes-64.par.71773.file.dat/bioprobes-64-qubit.pdf.

25.   Allard, W.J., et al., *Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases.* Clin Cancer Res, 2004. **10**(20): p. 6897-904.

26.   VYCAP, *Isolation-of-Single-CTC-from-CellSearch-cartridges-v1.0.pdf.* 2019.

27.   Fleitas, T., et al., *MassARRAY determination of somatic oncogenic mutations in solid tumors: Moving forward to personalized medicine.* Cancer Treatment Reviews, 2016. **49**: p. 57-64.

28.   Mosko, M.J., et al., *Ultrasensitive Detection of Multiplexed Somatic Mutations Using MALDI-TOF Mass Spectrometry.* J Mol Diagn, 2016. **18**(1): p. 23-31.

29.   Clark, T.G., et al., *Survival analysis part IV: further concepts and methods in survival analysis.* Br J Cancer, 2003. **89**(5): p. 781-6.

30. Giannoudis, A., et al., *Genomic profiling using the UltraSEEK panel identifies discordancy between paired primary and breast cancer brain metastases and an association with brain metastasis-free survival.* Breast Cancer Res Treat, 2021.

31. Allouchery, V., et al., *Circulating ESR1 mutations at the end of aromatase inhibitor adjuvant treatment and after relapse in breast cancer patients.* Breast Cancer Res, 2018. **20**(1): p. 40.

32. Chandarlapaty, S., et al., *Prevalence of ESR1 Mutations in Cell-Free DNA and Outcomes in Metastatic Breast Cancer: A Secondary Analysis of the BOLERO-2 Clinical Trial.* JAMA Oncol, 2016. **2**(10): p. 1310-1315.

33. Fribbens, C., et al., *Plasma ESR1 Mutations and the Treatment of Estrogen Receptor-Positive Advanced Breast Cancer.* J Clin Oncol, 2016. **34**(25): p. 2961-8.

34. Takeshita, T., et al., *Analysis of ESR1 and PIK3CA mutations in plasma cell-free DNA from ER-positive breast cancer patients.* Oncotarget, 2017. **8**(32): p. 52142-52155.

35. Arruabarrena-Aristorena, A., et al., *FOXA1 Mutations Reveal Distinct Chromatin Profiles and Influence Therapeutic Response in Breast Cancer.* Cancer Cell, 2020. **38**(4): p. 534-550 e9.

36. Smith, N.G., et al., *Targeted mutation detection in breast cancer using MammaSeq.* Breast Cancer Res, 2019. **21**(1): p. 22.

37. Afzaljavan, F., et al., *GATA3 somatic mutations are associated with clinicopathological features and expression profile in TCGA breast cancer patients.* Sci Rep, 2021. **11**(1): p. 1679.

38. Gaynor, K.U., et al., *GATA3 mutations found in breast cancers may be associated with aberrant nuclear localization, reduced transactivation and cell invasiveness.* Horm Cancer, 2013. **4**(3): p. 123-39.

39. Velimirovic, M., et al., *Landscape of GATA3 mutations identified from circulating tumor DNA clinical testing and their impact on disease outcomes in estrogen receptor-positive (ER+) metastatic breast cancers treated with endocrine therapies.* Journal of Clinical Oncology, 2021. **39**(15_suppl): p. 1065-1065.

40. Zehir, A., et al., *Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients.* Nat Med, 2017. **23**(6): p. 703-713.

41. Razavi, P., et al., *The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers.* Cancer Cell, 2018. **34**(3): p. 427-438 e6.

42. Li, Z., et al., *Loss of the FAT1 Tumor Suppressor Promotes Resistance to CDK4/6 Inhibitors via the Hippo Pathway.* Cancer Cell, 2018. **34**(6): p. 893-905 e8.

43. Gustin, J.P., et al., *GATA3 frameshift mutation promotes tumor growth in human luminal breast cancer cells and induces transcriptional changes seen in primary GATA3 mutant breast cancers.* Oncotarget, 2017. **8**(61): p. 103415-103427.

44. Tominaga, N., et al., *Clinicopathological analysis of GATA3-positive breast cancers with special reference to response to neoadjuvant chemotherapy.* Ann Oncol, 2012. **23**(12): p. 3051-3057.

45. Chaudhary, S., B.M. Krishna, and S.K. Mishra, *A novel FOXA1/ESR1 interacting pathway: A study of Oncomine breast cancer microarrays.* Oncol Lett, 2017. **14**(2): p. 1247-1264.

46. Hassani, B., et al., *Expression Analysis of Long Non-Coding RNAs Related With FOXM1, GATA3, FOXA1 and ESR1 in Breast Tissues.* Front Oncol, 2021. **11**: p. 671418.

47. Davis, A.A., et al., *Association of a novel circulating tumor DNA next-generating sequencing platform with circulating tumor cells (CTCs) and CTC clusters in metastatic breast cancer.* Breast Cancer Res, 2019. **21**(1): p. 137.

Table 1. Patient characteristics of metastatic breast cancer

Characteristic at primary diagnosis	Overall number (%)
Number of patients	101
ER status	
Positive	96 (96%)
Negative	4 (4%)
Missing	1
PR status	
Positive	88 (88%)
Negative	12 (12%)
Missing	1
ERBB2 status	
Positive	11 (11%)
Negative	82 (82%)
Missing	8
Grade	
G1-2	48 (48%)
G3	23 (23%)
Missing	30
T-stage	
T1	21 (21%)
T2	35 (35%)
T3	13 (14%)
T4	11 (10%)
Missing	21
N-stage	
N0	22 (22%)
N1	60 (60%)
Nx	2 (2%)
Missing	17

Distant metastasis

M0	30 (30%)
M1	18 (18%)
Mx	5 (5%)
Missing	48

Menopausal status

Pre/perimenopausal	12 (12%)
Postmenopausal	20 (20%)
Missing	69

At blood draw	N= 272 samples
Time-point	N=185 time-points

Distant metastasis

M1	75/101 (75%)
M0	18/101 (18%)
Missing	8

Adjuvant systemic therapy

Hormone therapy	30/185 (16%)
chemotherapy	82/185 (44%)
Both	63/185 (34%)
Non	3/185 (1.6%)
Missing	6

ER = estrogen receptor, PR = progesterone receptor, ERBB2 = Erb-B2 Receptor Tyrosine Kinase 2, HR = hazard ratio.
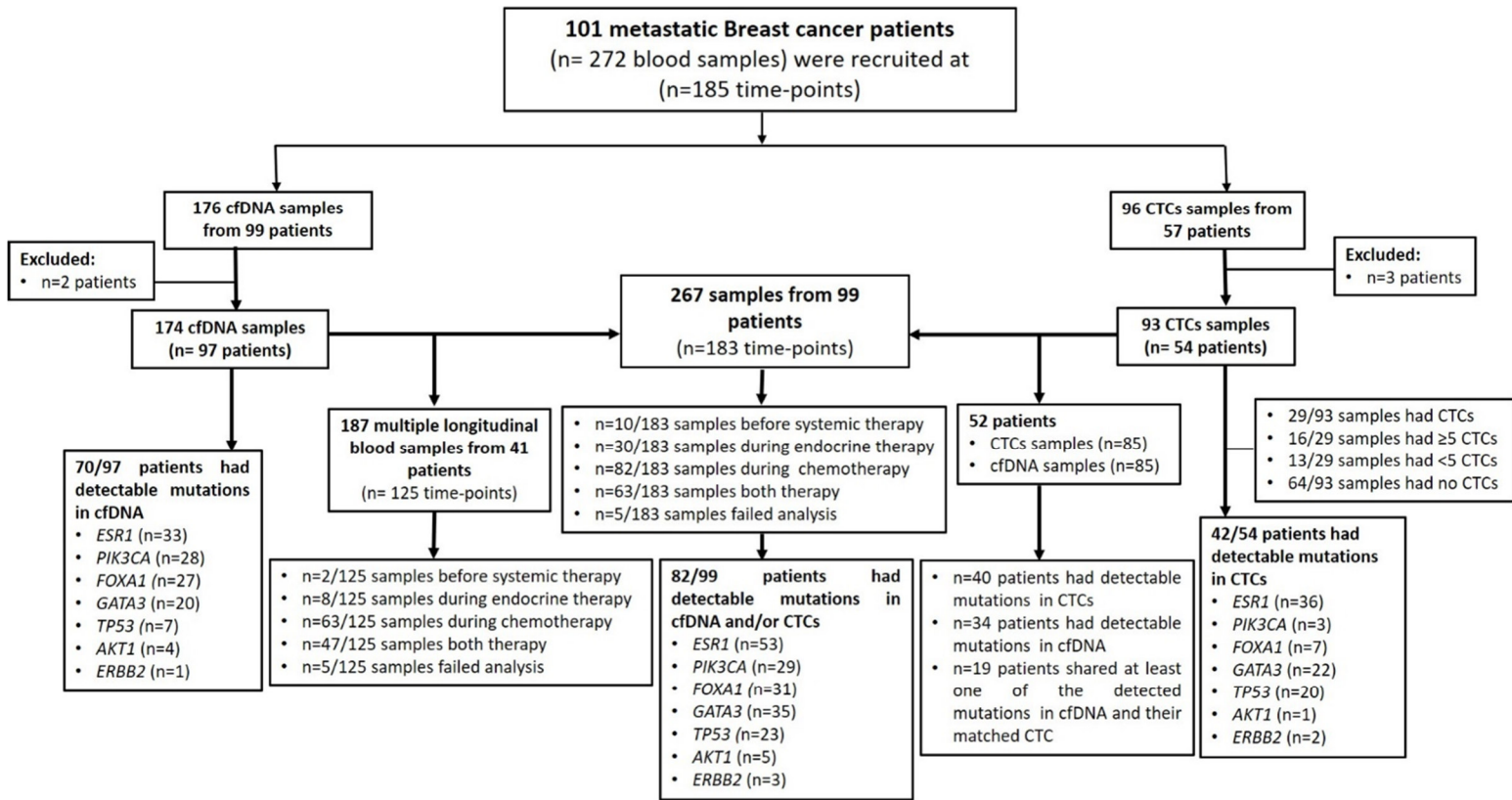
Figure 1. REMARK diagram for patient recruitment, exclusions, and distribution.
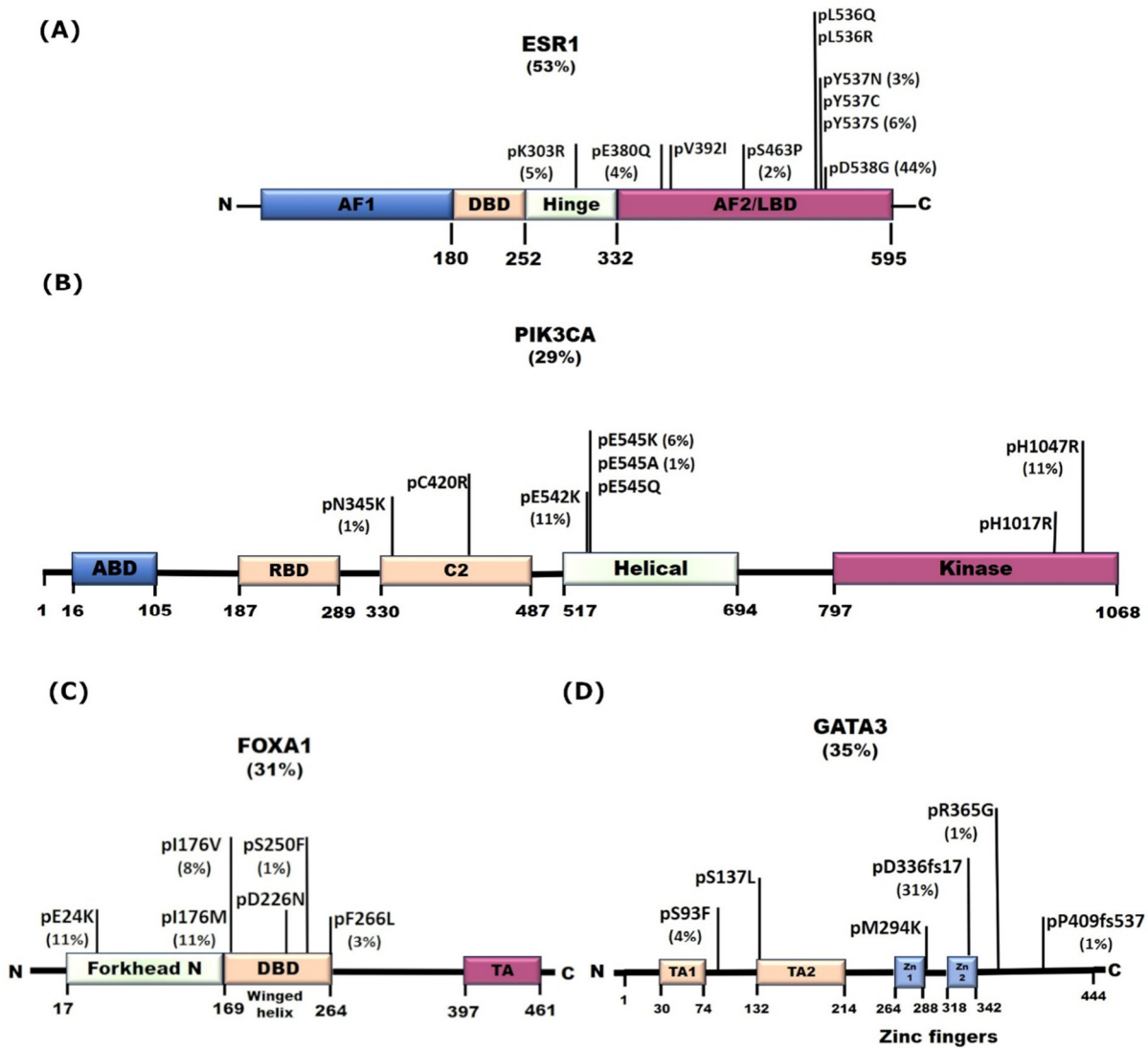
Figure 2. schematic of hotspot mutations identified in this study. (A) ESR1 structure domains; AF1(activation functional domain); DBD (DNA binding domain); hinge domain; LBD (ligand-binding domain) including AF2 (second transcriptional activation domain) encompasses hot spot mutations. (B) PIK3CA structure domains; ABD (adapter binding domain); RBD (Ras-binding domain); C2 (calcium-dependent phospholipid-binding domain); Helical (PI3K helical domain); Kinase (PI3/4-kinase domain) including the distribution of frequent alteration. (C) FOXA1 functional domains; Forkhead domain in N-terminal region; Winged helix DBD (winged helix–turn–helix DNA-binding domain) harbor hot spot alterations; TA (transactivation domains). (D) Distribution of mutations in GATA3 functional domains: TA1 and TA2 (two transactivation domains); Zn1 and Zn2 (two zinc fingers).
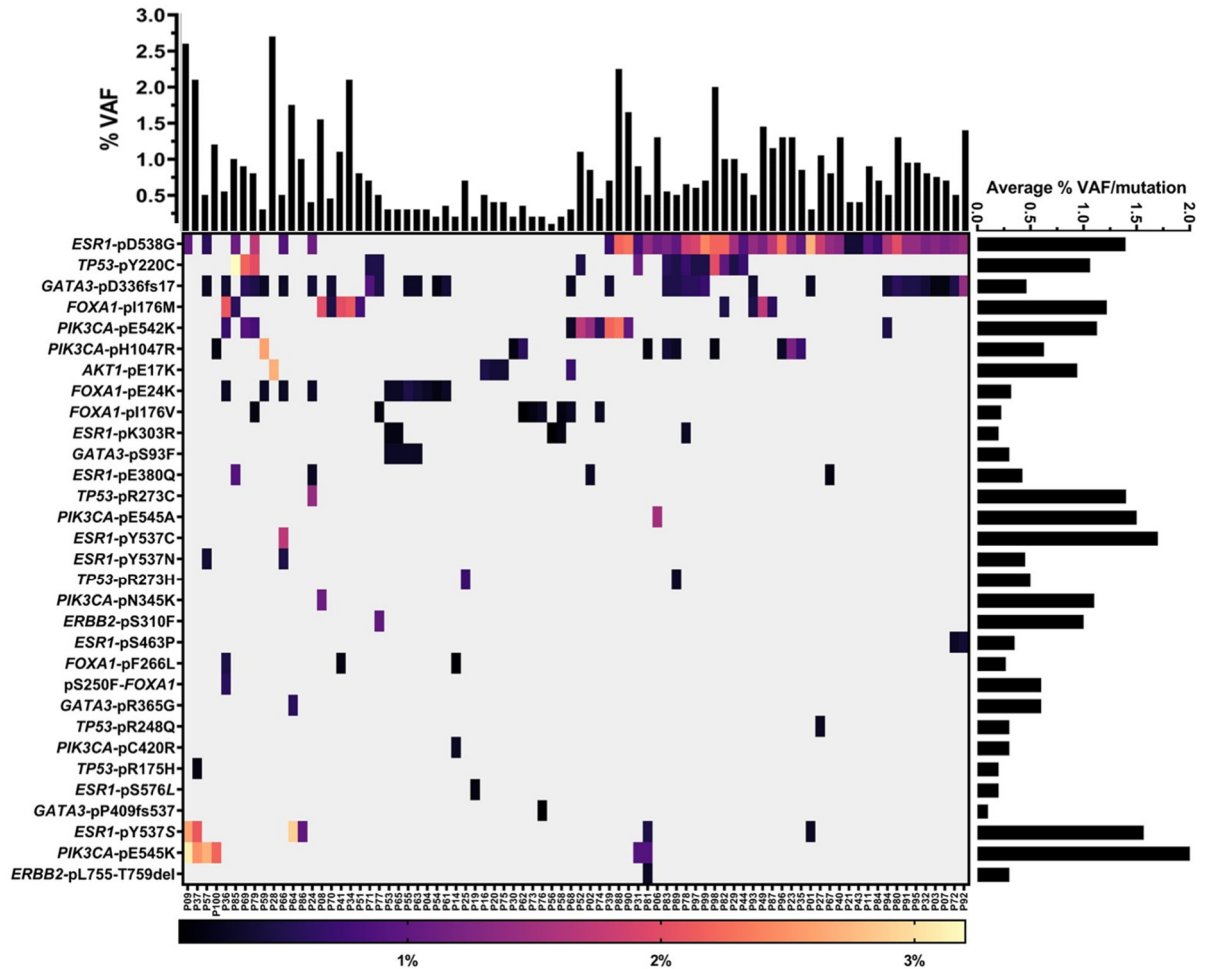
Figure 3. Heatmap depicting the distribution of the detected mutations in patients (n= 99); the upper panel shows the median of tumor load in each patient; the right panel shows the average of mutation load; scaled par represents the percentage of VAF.
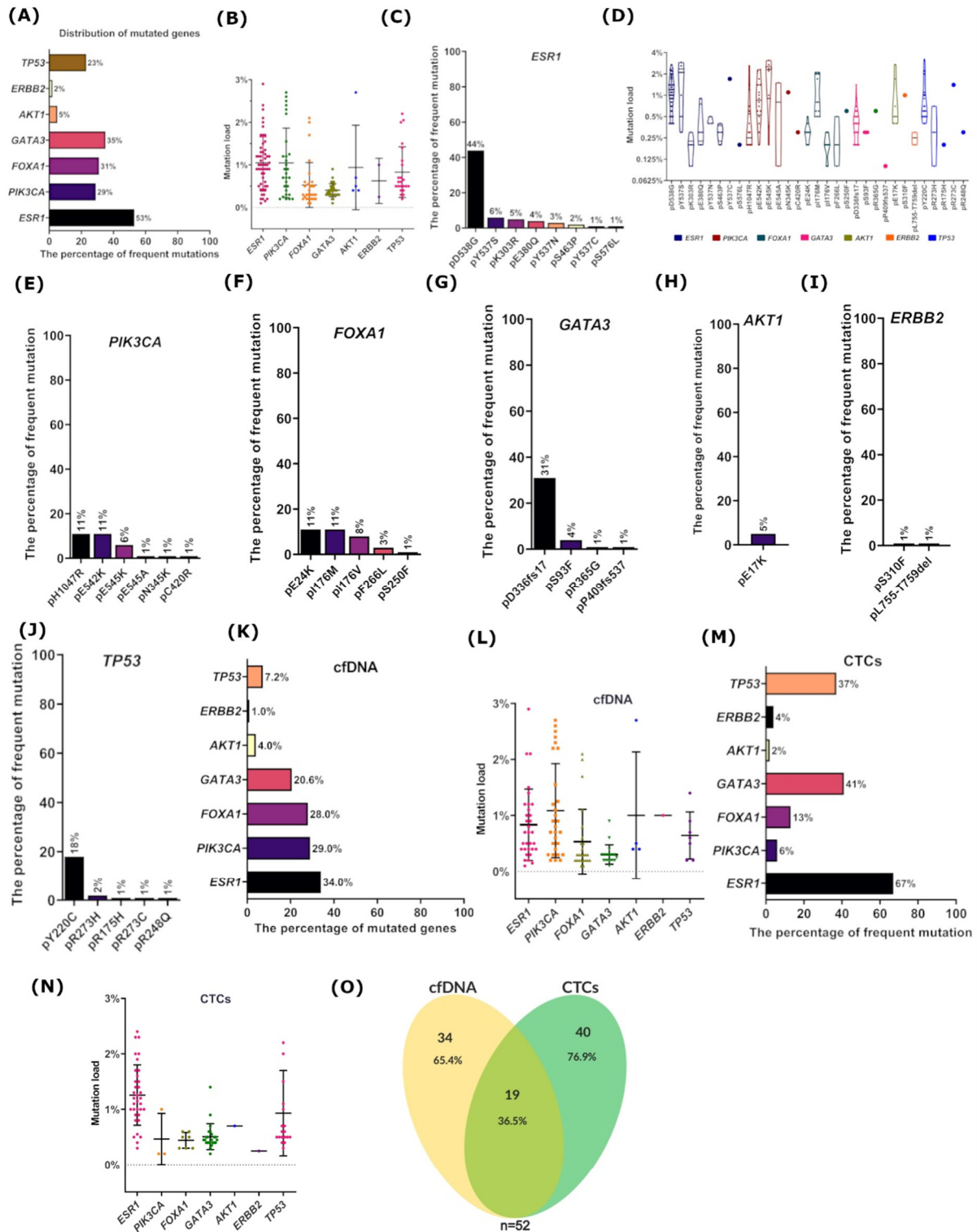
Figure 4. Overview of detected mutations in metastatic breast cancer patients and their variant allele frequency (VAF). (A) Distribution of mutated genes in patients (n= 99). (B) Scatter plot showing the distribution of mutation load in mutated genes. (C) The prevalence of *ESR1* detected mutations in patients (n= 99). (D) Violin plot showing the mutation load of detected mutations; the middle line

represents the median. The prevalence of detected mutations of (E) *PIK3CA*, (F) *FOXA1*, (G) *GATA3*, (H) *AKT1*, (I) *ERBB2*, and (J) *TP53* in patients (n=99). (K) The frequency of mutated genes in cfDNA of breast cancer patients (n=97). (L) Mutation load of mutated genes in cfDNA (n= 98). (M) The frequency of mutated genes in CTCs of patients (n= 54). (N) Mutation load of mutated genes in CTCs (n= 56). (O) Venn diagram depicting shared and distinct somatic mutations in cfDNA and CTCs.
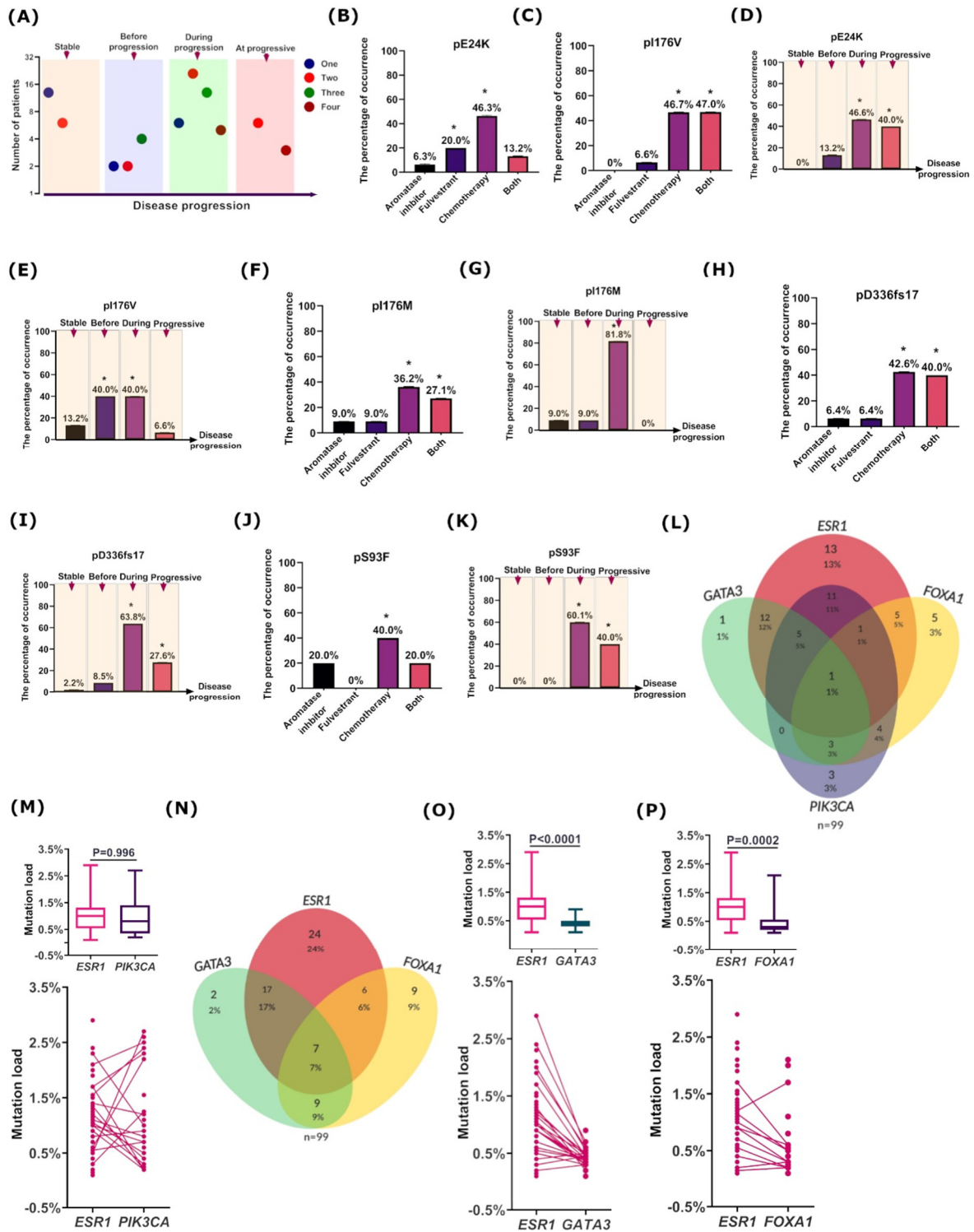
Figure 5. Mutations associated with therapy and disease progression. (A) Schematic diagram showing increased number of mutations along with disease progression. Occurrence of candidate mutations in patients received Aromatase inhibitor, Fulvestrant, Chemotherapy, or both endocrine therapy and chemotherapy (B) *FOXA1* (pE24K), (C) *FOXA1* (pI176V), (F) *FOXA1* (pI176M), (H) *GATA3* (pD336fs17), and (J) *GATA3* (pS93F). Occurrence of candidate mutations during disease

stability, before disease progression, during progression of disease, and in the progressive phase of disease (D) *FOXA1* (pE24K), (E) *FOXA1* (pI176V), (G) *FOXA1* (pI176M), (I) *GATA3* (pD336fs17), and (K) *GATA3* (pS93F). (L) Venn diagram showing shared somatic mutations of *ESR1*, *PIK3CA*, *FOXA1*, and *GATA3* in patients. (M) The difference between *ESR1* and *PIK3CA* in mutation load. (N) Venn diagram showing shared somatic mutations of *ESR1*, *FOXA1*, and *GATA3* in patients. The difference in mutation load between (O) *ESR1* and *GATA3* and (P) *ESR1* and *FOXA1*.
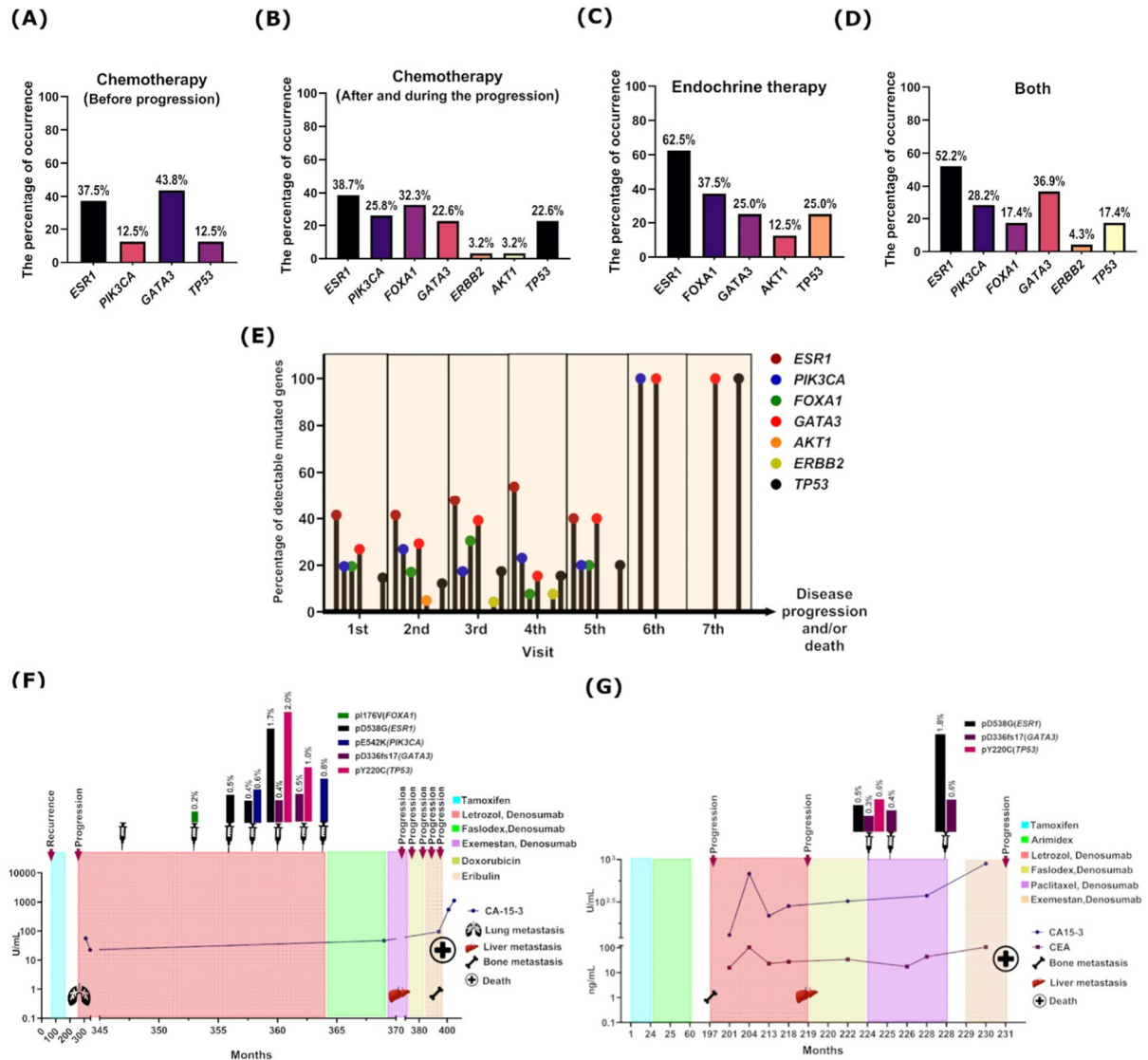
Figure 6. Longitudinal analysis of mutant genes in metastatic breast cancer patients. (A) Distribution of mutated genes in patients who received (A) chemotherapy before disease progression, (B) chemotherapy after and during disease progression, (C) endocrine therapy, (D) both endocrine and chemotherapy. (E) Schematic diagram showing the most mutated genes during follow-up visits of patients. Monitoring of candidate mutations in two patients (F) and (G) during the course of disease until death.
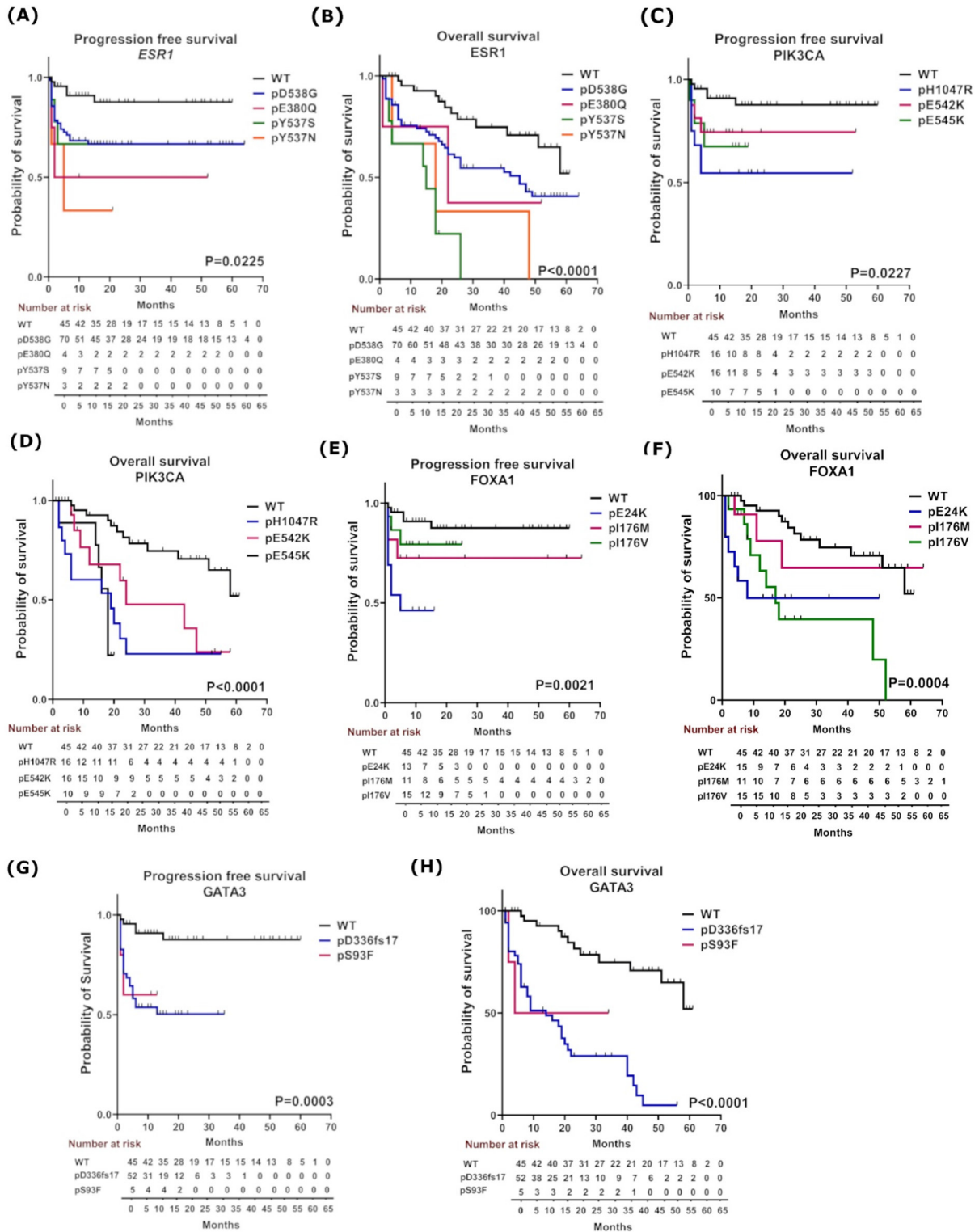
Figure 7. Survival analyses. Kaplan–Meier curves for progression-free survival (PFS) and overall survival (OS) according to the candidate mutations independently for ESR1 (A, B); PIK3CA (C, D); FOXA1 (E, F); GATA3 (G, H) compared to WT (wild-type cfDNA). P-values were calculated using the log rank test; ⊥ censored.
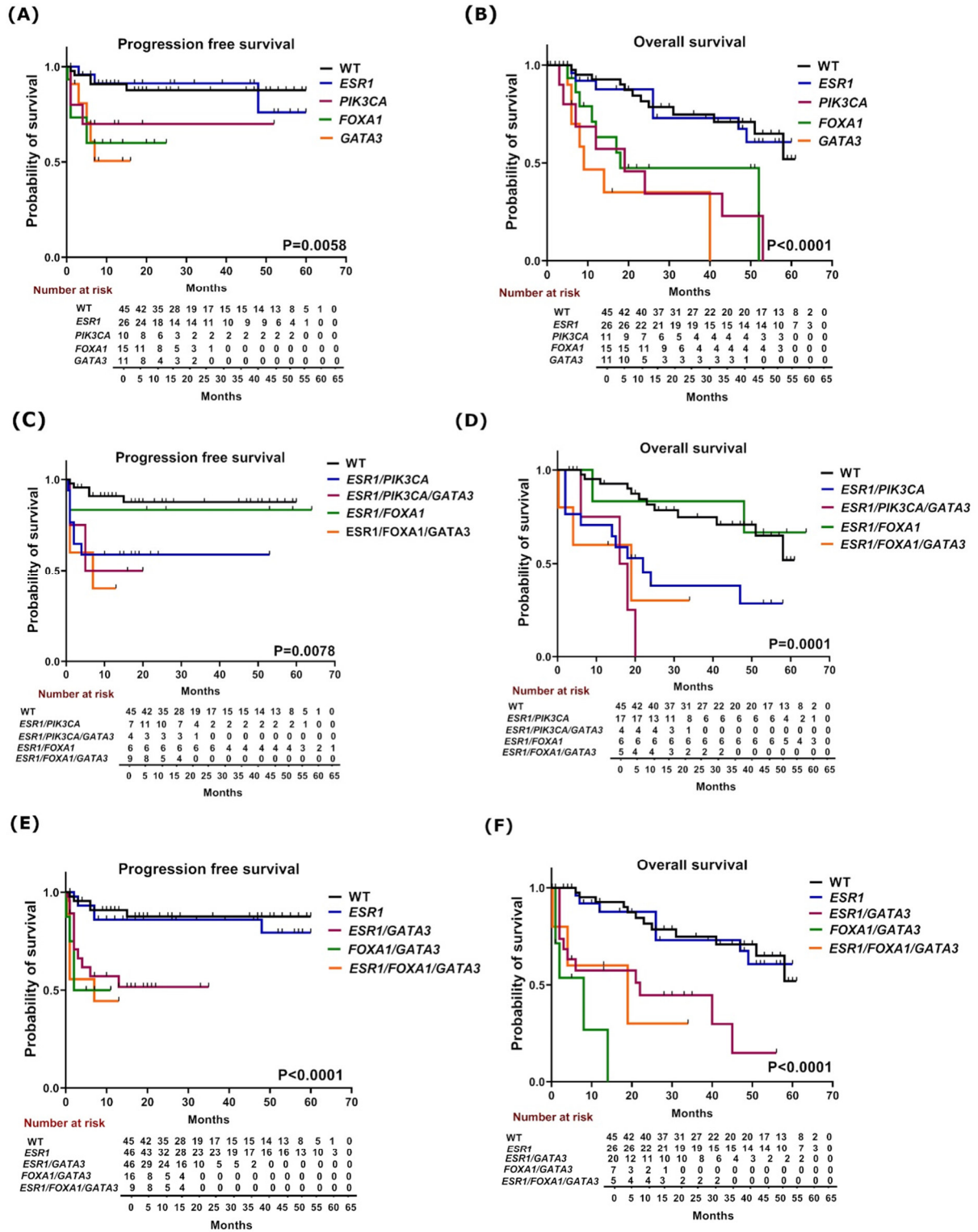
Figure 8. Survival analyses. Kaplan–Meier curves comparing patients with monoclonal (A, B) and polyclonal (C, D, E, and F) mutated genes compared to WT (wild-type cfDNA). P-values were calculated using the log rank test; ⌐ censored.

5.5. Techniques of using circulating tumor DNA as a liquid biopsy component in cancer management

This study focused on the use of circulating tumor DNA (ctDNA), which is derived from a tumor for the clinical treatment of cancer patients. The study provided a comprehensive overview of the many procedures used to extract and characterize ctDNA (e.g., Next-generation sequencing (NGS), Digital-PCR platforms, Real-time PCR-based methods, and Mass-spectrometry technology). Further, we addressed the challenges that still need to be overcome in order to accomplish ctDNA-based liquid biopsy for precision medicine.

Mini Review

Techniques of using circulating tumor DNA as a liquid biopsy component in cancer management

Maha Elazezy, Simon A. Joosse *

University Medical Center Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany

ABSTRACT

Precision medicine in the clinical management of cancer may be achieved through the diagnostic platform called "liquid biopsy". This method utilizes the detection of biomarkers in blood for prognostic and predictive purposes. One of the latest blood born markers under investigation in the field of liquid biopsy in cancer patients is circulating tumor DNA (ctDNA). ctDNA is released by tumor cells through different mechanisms and can therefore provide information about the genomic make-up of the tumor currently present in the patient. Through longitudinal ctDNA-based liquid biopsies, tumor dynamics may be monitored to predict and assess drug response and/or resistance. However, because ctDNA is highly fragmented and because its concentration can be extremely low in a high background of normal circulating DNA, screening for clinical relevant mutations is challenging. Although significant progress has been made in advancing the detection and analysis of ctDNA in the last few years, the current challenges include standardization and increasing current techniques to single molecule sensitivity in combination with perfect specificity. This review focuses on the potential role of ctDNA in the clinical management of cancer patients, the current technologies that are being employed, and the hurdles that still need to be taken to achieve ctDNA-based liquid biopsy towards precision medicine.

© 2018 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Contents

1. Introduction

Cancer is the consequence of deregulation of tumor suppressors and proto-oncogenes caused by the accumulation of mutations in the genome of a normal cell [1,2]. Proto-oncogenes promote cell division and proliferation, whereas tumor suppressors can induce apoptosis and are negative regulators of cell proliferation [3]. The identification of the genetic and/or epigenetic modifications leading to pathogenesis can be exploited for anticancer therapy management, prediction, and prognosis [4]. Cancer-related mutations include chromosomal aberrations such as copy numbers alterations (CNAs), inversions,

* Corresponding author.
  E-mail address: s.joosse@uke.de (S.A. Joosse).

translocations, insertions, and deletions, as well as single nucleotide point mutations [3]. Epigenetics refers to the covalent modification of DNA resulting in changes to the function and/or regulation of the affected genes, without altering the primary sequences (a change in phenotype without a change in genotype). Epigenetic factors such as DNA methylation and histone modification, play a key role in gene activity, cell differentiation, tumorigenesis, X-chromosome inactivation, genomic imprinting, and other cellular regulatory processes [5].

Metastatic spread is the main cause of cancer-related death and is the result of colonization of tumor cells from the primary tumor into distant organs, which may finally be followed by organ failure. The route of dissemination takes place mainly through the blood circulation, in which only very few circulating tumor cells (CTCs) are able to survive [6]. Extravasation of the tumor cells is usually expected to occur in distant organs such as the brain, bone marrow, lungs, or liver in which the disseminated tumor cell (DTCs) can stay dormant for many years (Fig. 1) [7]. The observation of DTCs in bone marrow has been shown to be highly correlated with recurrence of disease [8].

In order to molecularly characterize the tumor and identify potential therapeutic targets, material directly taken from the tumor has to be investigated. The standard procedure to genotype a tumor is by obtaining a small piece of tissue using a tissue biopsy, which is a rather invasive procedure. Furthermore, neoadjuvant treatment may shrink the tumor to undetectable size, leaving no tissue for further investigation. Therefore, the procedure to obtain a tissue biopsy is severely hampered by spatial and temporal limitations; in addition, a single biopsy sample may not represent the full tumor load's heterogeneity [9,10]. As an alternative to characterize the tumor, blood can be used to obtain biomolecules or other markers originating from the tumor. One of these markers is circulating tumor cells (CTCs) that originate from the currently present tumor and thereby can function as a so-called "liquid biopsy" (Fig. 1) [11].

The identification of CTCs has been shown to have prognostic and predictive value in different entities of early-stage cancer [12]. However, highly sensitive techniques are required to identify the small number of cells in the extremely high background of normal cells. The different methods available for obtaining CTCs are either based on specific cellular makers expressed on the cell surface [13] or on the physical properties of the cells. Antigens expressed by the tumor cells enable positive enrichment whereas negative enrichment can be achieved by depletion of white blood cells [6]. Because the half-life time of CTCs is <2.5 hours [14] and the metastases are also able to shed tumor cells into the circulation, more CTCs can be expected in the advanced stages of the disease [15]. Other blood-borne biomarkers currently used as liquid biopsy include platelets, cell-free nucleotides, and extracellular vesicles such as exosomes (Fig. 1) [11]. Platelets may be altered through confrontation with tumor cells via transfer of tumor-associated biomolecules [16]. These so called tumor-educated platelets (TEPs) contain a variety of RNA transcripts and proteins that may influence the process of metastasis development by enhancing or blocking tumor cells, immune cells, and stromal cells, either by direct cell-to-cell contact or by releasing extracellular queues [17,18]. Exosomes are an effective way for cells to secrete mRNA and miRNA into the circulation that may lead to disease progression [19]. For example, exosome-mediated transfer of cancer-secreted miR-105 promotes metastasis in breast cancer [20]. Therefore, identification of such cell-free miRNAs can be used to serve as a biomarker for the early stage of metastasis [21]. Besides RNA, cell-free nucleotides also include cell-free DNA (cfDNA). As a consequence, liquid biopsy may also include the screening for fetal aneuploidy where the cfDNA originates either from the fetus or from apoptotic placental cells, circulating in a pregnant woman's plasma, is investigated [22]. This review will focus on the use of cfDNA originating from the tumor, i.e., circulating tumor DNA (ctDNA), for the clinical management of cancer patients and provide a comprehensive overview of the different techniques being applied to obtain and characterize ctDNA.

2. Circulating tumor DNA (ctDNA) properties

Two processes are involved in the release of ctDNA into the blood circulating [23]. The first is a passive release of DNA through cell death
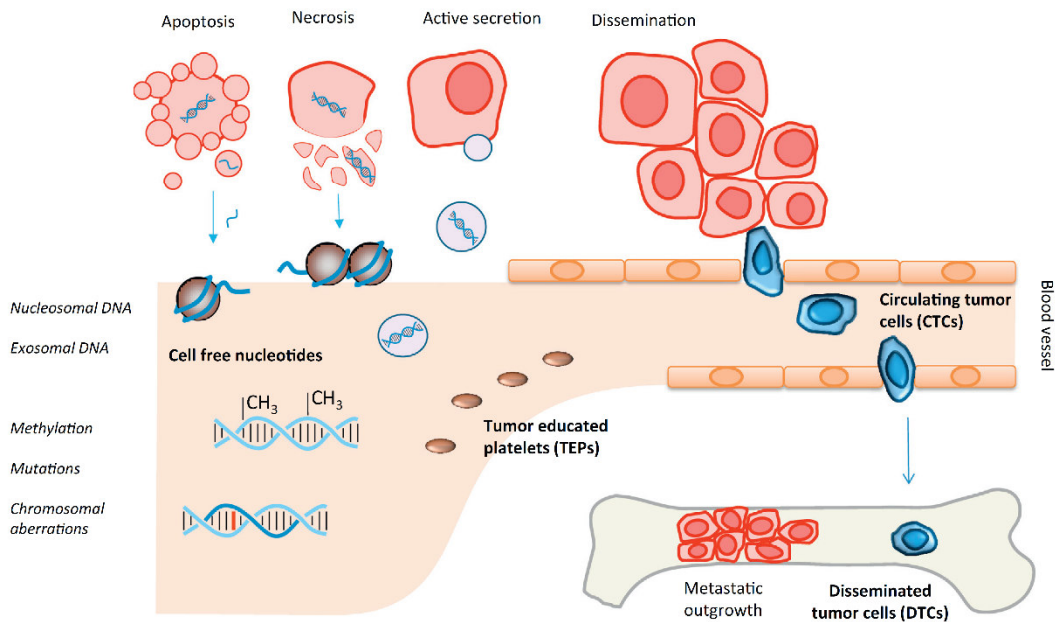
Fig. 1. Liquid biopsy markers. Biomarkers that are currently used as liquid biopsy include cell free nucleotides, circulating tumor cells (CTCs), tumor educated platelets (TEPs), and disseminated tumor cells (DTCs). Cell free nucleotides are released into the blood circulating by apoptotic or necrotic cells, or by active secretion of exosomes containing a cell's genetic material. Cell free DNA (cfDNA) is highly fragmented but is still wrapped around nucleosomes providing its typical length of 166 or 320 bp. cfDNA may be used to study a tumor's methylation patterns, chromosomal aberrations, or other mutations.

either by apoptosis or necrosis (Fig. 1). As a consequence of enzymatic cleavage of DNA during apoptosis, the resulting DNA fragments are still wrapped around single nucleosomes and the length plus linker is around 166 bp [24,25]. Larger fragments starting from 320 bp, the length of DNA wrapped around two nucleosomes, up to >1000 bp are released from phagocytosis of necrotic cells [23]. The second mechanism of ctDNA release is by active secretion [23]. Secretion of ctDNA takes place by the release of extracellular vesicles, such as exosomes and prostasomes, containing pieces of DNA around 150-250 bp [26]. Plasma DNA that originates specifically from tumors (ctDNA) typically represents 0.01–90% of the total cell-free DNA (cfDNA) found in blood [27,28]. It is hypothesized that ctDNA is secreted by tumor cells as a signaling molecule to drive tumor metastasis [29,30]. For example, two independent studies demonstrated that ctDNA may be involved in tumorigenesis and metastasis development. By incubating murine NIH-3 T3 cells with plasma from patients with *KRAS* mutated colorectal tumors followed by injection into mice, the development of tumors could subsequently be observed as well as the detection of human *KRAS* mutations in the mice' plasma [29,30]. Furthermore, it was observed that ctDNA could promote the proliferation of hormone receptor-positive breast cancer cells by activation of the TLR9-NF-κB-cyclin D1 pathway in vitro [31]. Finally, a small part of the ctDNA may originate from CTCs that die in the blood stream [32].

The rate of ctDNA shedding into the circulation depends on the location, size, and vascularity of the tumor, leading to a difference in ctDNA levels among patients [33,34]. The half-life time of ctDNA in the blood circulation ranges from 16 minutes to 2.5 hours [35]. The concentration of the total cfDNA in healthy individuals is on average 30 ng/ml plasma and ranges from 0 to 100 ng/ml, whereas in cancer patients this can be up to 1000 ng/ml [36,37]. In order to extract cfDNA from the blood, different methods have been developed. Magnetic enrichment of cfDNA can be achieved by positively charged magnetic beads that bind the negatively charged phosphate backbone of DNA [38–41], whereas silica column-based enrichment makes use of the binding affinity of DNA molecules [38–40,42–44]. Furthermore, cfDNA capturing can be performed by polymer mediated enrichment (PME) [39] or by a phenol-chloroform based extraction procedure in which DNA is not soluble [42]. Several studies have compared these extraction methods using DNA yield, fragment size distribution, and the quality of the obtained DNA in downstream analysis using for instance mutation detection as a read-out [38,39,42,43]. However, these studies have shown large variations in cfDNA yield and/or fragment size between the different extraction methods. For example, conventional extraction methods based on phenol-chloroform have shown higher yields than with DNA extraction kits, but DNA purity and thereby efficiency of downstream analyses was lower as compared to the magnetic-based method [40]. Some studies have favored the silica-based membrane method due to the high recovery of 82%–92% cfDNA from serum [45]. However, the silica-based membrane system has the disadvantages of a low yield and partial loss of DNA fragments smaller than 150 bp [46,47]. In contrast, a magnetic bead-based method seems to be more efficient in the recovery of short cfDNA fragments as compared to the silica-based membrane and conventional methods [48].

3. Clinical applications of ctDNA

The investigation of biomarkers that may help to detect cancer in its early stages before becoming clinically apparent could eventually lead to a decreased mortality [49]. The quantification of cfDNA concentration has been studied to discriminate between healthy individuals and malignant disease [50,51]. It was demonstrated that the levels of cfDNA in NSCLC cancer patients are significantly higher than in healthy individuals [50], in fact, a cutoff level of cfDNA >0.20 mg/ml is able to distinguish between lung cancer patients and control cases with a sensitivity of 69–79% and a specificity of 83–89% [50,51]. Furthermore, many studies have demonstrated that the cfDNA concentration is associated with

tumor volume leading to shorter overall survival (OS) of patients with breast [52], ovarian [53], lung [54,55], gastric [56], and colorectal cancer [35,57]. Interestingly, contradictory data have also been reported showing that the concentration of cfDNA did not seem to be associated with overall or progression-free survival [58]. Although, these data indicate that cfDNA levels can be used to monitor tumor progression, using cfDNA for diagnostic purposes is still of limited value.

Quantification of tumor-specific mutations in ctDNA appears to be more relevant for studying tumor progression. High levels of mutated *PIK3CA* in serum DNA of breast cancer patients are associated with short progression-free and overall survival as compared to patients with low or no detectable amounts of mutated ctDNA [59]. The analysis of single nucleotide variants in *KRAS, NRAS, PIK3CA, BRAF,* and *EGFR* using cfDNA has been shown to have >80% concordance when compared to tumor tissue of colorectal [60,61], lung [34,62], and breast [59,61] cancer patients. However, also the time-point at which liquid biopsy is performed in order to track minimal residual disease (MRD) seem to be important, as the ctDNA concentration may lay below the detection limit during certain stages of the treatment. For example, Murillas et al. demonstrated that the detection of ctDNA eight months after surgery is associated with a high risk of relapse in early-stage breast cancer patients, whereas this could not be discerned before the primary surgery based on the detected mutations [63].

ctDNA can also be used to monitor therapy efficiency by detecting mutation-driven resistance [61,64,65]. For example, early detection of *ESR1* mutations, which drive endocrine therapy resistance, may help to improve the outcome of patients by switching to other treatment before clinical progression of metastatic breast cancer patients [66]. Likewise, the detection of *KRAS* gene mutations in ctDNA of colorectal cancer patients may indicate resistance to epidermal growth factor receptor inhibitors [61]. Furthermore, decreasing sensitivity to tyrosine kinase inhibitors (TKIs) in patients with gastrointestinal stromal tumors could be demonstrated by tracking primary and secondary hotspot mutations in *KIT* (S821F) and *PDGFRA* (D842V) [67]. These data demonstrate the potential of ctDNA to detect and monitor the clonal evolution of cancer through serial genotyping, giving a more complete picture of the distinct genetic subclones that are related to drug resistance [68].

Methylation patterns found on ctDNA can be exploited as biomarkers to detect epigenetic deregulation of genes. Hypermethylation of the promoter of *RASSF1A, FHIT,* and *APC* found in plasma DNA was shown to be a useful diagnostic marker for early stage renal cancer with a sensitivity of 56.8% and specificity of 96.7% [69]. The detection of hypermethylation of the *MLH1* gene promoter in ctDNA could be employed as a predictive biomarker for acquired resistance in ovarian cancer and was associated with a poor overall and progression-free survival [70]. Similarly, the identification of methylation of *ESR1* promoter in ctDNA was found to be associated with a lack of response to everolimus/exemestane therapy in metastatic breast cancer patients [71]. Taken together, ctDNA has a high potential for monitoring clinically relevant cancer-related genetic and epigenetic modifications for discovering more detailed information on the tumor characterization [72].

4. ctDNA detection technologies

cfDNA is highly fragmented DNA and the total amount of ctDNA might make up as low as 0.01% of the total cfDNA. These extreme low concentrations make the detection challenging, particularly at the early stages of tumor development [27,73,74]. Two strategies have emerged to study the tumor's genomic material by liquid biopsy. First, targeted approaches in which a single or few tumor-specific mutations known from the primary tumor are used for monitoring residual disease in the peripheral blood. Such techniques include Q-PCR, BEAMing, Safe-SeqS, CAPP-Seq, and TAmSeq [57]. The disadvantage of this strategy is that it requires detailed information about the tumor genome. However, targeted monitoring can be extremely sensitive, as mutations can be detected at an allele frequency of down to 0.01% with high specificity and

at a fast and cost-effective rate [75–77]. The second strategy to investigate ctDNA involves untargeted screening and aims at a genome-wide analysis for copy number aberrations (CNAs) [78] or point mutations by whole-genome sequencing (WGS) or whole exome sequencing (WES) [79]. Advantages of untargeted strategies include (i) its ability to identify novel changes occurring during tumor treatment and (ii) prior information about the primary tumor's genome is not required. However, a disadvantage is that high concentrations of ctDNA are required for reliable reconstruction of tumor-specific genome-wide changes. Furthermore, untargeted approaches show an overall low sensitivity (5%–10%) [79]. Depending on which strategy is required to investigate the ctDNA or interest, different technologies are currently available (Table 1).

An additional strategy might be an alternative to "genotype-independent approaches" a non-invasive screening approach, which based on the fragmentation patterns of an individual's cfDNA that can include an evidence of the epigenetic profile of the origin cells. Such a footprint of nucleosome-bound cfDNA that can be used to determine the contributing cell types in the absence of genotypic differences [80].

4.1. Next-generation sequencing (NGS)

NGS has emerged in the past decade as an efficient technique for sequencing DNA and obtaining genetic information. NGS is based on the analysis of several millions of short DNA sequences in parallel followed by either sequence alignment to a reference genome or de novo sequence assembly. Despite its high sensitivity and specificity, NGS shows a random error rate between 0.1% and 1% depending on the applied platform [79] making the detection of ctDNA by rare mutations in the total cfDNA challenging. According to this observation, many protocols have been modified to improve and expand the detection of rare mutations [81] (Table 1).

Deep-sequencing is considered the first approach to detect mutations at an allele-frequency as low as <0.2% by sequencing the target regions with high coverage (>10,000×) [82–84]. As a result, the sensitivity of deep sequencing of finding mutations in cfDNA earlier discovered in tumor tissue can be up to 100%, although the specificity can be as low as 80% [83]. In early stage lung cancer patients (stages IA–IIIA), it was shown that deep sequencing for ctDNA resulted in a low sensitivity of 36.5% in detecting the *EGFR* (L858R) mutation present in the tumor tissue, whereas this increased to 72.7% in metastatic setting (stages IIIB–IV) [84]. The main advantage of deep sequencing is the ability to assess multiple biomarkers simultaneously while its disadvantage is the extreme high read depth that has to be performed in order to detect mutations at low allele frequency and thereby drastically increasing sequencing costs.

Bias-Corrected Targeted NGS is adapted to minimize PCR artifacts by using multifunctional adapters that facilitate read analysis and identify which probe captured the fragment. Bias-Corrected Targeted NGS was applied on cfDNA of NSCLC patients resulting in a detection of >0.4% mutant allele frequency with a specificity of 100% [81]. This technology

Table 1
Technologies for detecting circulating tumor DNA (ctDNA).

Technology	Platform	1-Sensitivity	Specificity	cfDNA input	Number of targets	Type of alteration	Limitations	References
NGS	Deep sequencing (>10,000×)	0.02%	80–90%	2 ng	Panel	Genome-wide copy number changes	Unable to detect rearrangements without assay customization	[82–84]
	TAm-Seq	0.02%	99.9997%	0.9-20 ng	Panel	Known point mutations	Detects only known mutations	[89]
	Safe-SeqS	0.1%	98.9%	3 ng	Panel	Known point mutations and copy number variations	Less comprehensive than WES	[90,91]
	FASTSeqS	>10%	80%	5-10 ng	Panel	Genome-wide copy number changes	Low sensitivity and specificity	[86,87]
	CAPP-Seq	0.004%	>99.99%	32 ng	Panel	Known point mutations, copy number variations, and rearrangements	High cfDNA input; detects only known mutations	[92–94]
	MCTA-Seq	0.25%	89%	7.5 pg	Panel	Known methylation sites		[130]
	Bias-Corrected Targeted NGS	>0.4%	100%		Panel	Known point mutations, copy number variations, and rearrangements		[81]
	Multiplex-PCR NGS	>0.1%	99.6%	2-50 ng	Panel	Known point mutations	Detects only known mutations	[85]
Digital-PCR	ddPCR	0.1%	100%	25 ng	1 to 3	Known point mutations	Detects specific genomic loci; limited in multiplexing	[78,109–111]
	BEAMing	0.01%	100%	1 ng	1 to 20	Known point mutations	Detects only known mutations	[112–115]
Real-Time PCR	AS-PCR	1%	98%	3–50 ng	1	Known point mutations	Low sensitivity; detects known mutations	[119–121]
	AS-NEPB-PCR	0.1%	100%	20 ng	1	Known point mutations	Detects only known point mutations	[76]
	(PNA-LNA) PCR clamp	0.1–1%	79%	30 ng	1	Known point mutations	Low specificity; detects only known point mutations	[122–124]
	(COLD-PCR)	0.1%	94.9%	1–10 ng	1–3	Known point mutations	Detect limited genomic loci; limited in multiplexing	[77]
	MS-PCR	0.62%	100%	20–100 ng	1	Known methylation sites	Detects only specific CpG islands	[71]
Mass-spectrometry technology	SERS	0.1%	100%	5 ng	3 to 10	Known point mutations	Detect limited genomic loci	[125]
	UltraSEEK	0.1%	100%	9 pg-4.2 ng	Up to 40	Known point mutations	Detect limited genomic loci	[126,127]

The performance of the different technologies for detecting ctDNA using different platforms. These technologies differ in sensitivity, specificity, the minimum input of cfDNA, the number of targets that can be analyzed in one reaction, and the type of alterations that can be detected. In addition, the limitations of each technology are indicated. Smallest allele frequencies = 1-sensitivity; TAm-Seq: Tagged-amplicon deep sequencing; Safe-SeqS: Safe-Sequencing System; WES: whole exome sequencing; CAPP-Seq: Cancer Personalized Profiling by deep sequencing; ddPCR: Droplet Digital polymerase chain reaction; BEAMing: Beads, Emulsion, Amplification and Magnetics; AS-PCR: Allele-specific amplification; AS-NEPB-PCR: Allele-Specific, Non-Extendable Primer Blocker PCR; (PNA-LNA) PCR clamp: Peptide Nuclei Acid-Locked Nucleic Acid; COLD-PCR: co-amplification at lower denaturation temperature; MS-PCR: methylation-specific PCR; SERS: surface-enhanced Raman spectroscopy.

showed a high specificity in the detection of genomic alterations without producing false positives.

Multiplex-PCR NGS is based on a designed PCR assay panel that facilitates amplification of specific target regions. Validation of the multiplex-PCR NGS platform on the early stage of lung cancer patients showed a highly sensitive detection of >99% of single-nucleotide variants (SNVs) at allele frequencies of >0.1% with a specificity of 99.6% with as less as 20 ng of cfDNA as input material [85].

FAST-SeqS is a simple and efficient method for the detection of aneuploidy by massive parallel sequencing [86,87]. FAST-SeqS can amplify approximately 38,000 amplicons with a single primer pair. During amplification, degenerate bases at the 5′-end of the primer are used as molecular barcodes to uniquely label each DNA template molecule. This ensures that each DNA template molecule is counted only once [88]. A modified version of FAST-SeqS (mFAST-SeqS) was established as a prescreening tool to estimate the ctDNA percentage by using a single primer pair to select and amplify distinct sections of the genome that occur on every chromosome and estimate a genome-wide z-score to evaluate the ctDNA percentage [75]. mFAST-SeqS has for example been used to monitor changing levels of ctDNA in prostate cancer patients before and after treatment, showing a decrease in the genome z-score in patients who responded to therapy [87]. The advantages of this approach include speed (<1 day) and it does not depend on prior knowledge of the genetic composition of tumor samples. Nevertheless, the lowest detection limit of 10% ctDNA is a clear disadvantage [87].

TAm-Seq (Tagged-amplicon deep sequencing) is based on a combination of efficient library preparation and statistically-based analysis algorithms. This technique is adapted to sequence, detect, and quantify tumor mutations across a gene panel including both tumor hotspots, as well as entire coding regions of selected genes [73]. The precision of this methods could be shown by the detection limit of 0.02% with 99.9997% specificity for point mutations in *EGFR* in circulating DNA [89]. The development of a bioinformatic method is a clear advantage that has helped to design more efficient gene panels, improve the detection sensitivity of mutant alleles, and reduce the detection of false positives.

Safe-SeqS was designed to further improve the sensitivity of NGS. Safe-SeqS includes two main steps, the first is to assign a unique identifier (UID) to each DNA template molecule and the second is to amplify each uniquely tagged template to create UID families and sequences [90]. The Safe-SeqS approach has for instance been applied to ctDNA of patients with metastatic colorectal and gastrointestinal stromal tumors (GIST) for tracking therapy response. Here, Safe-SeqS showed a highly sensitive detection of a mutant allele with a concentration of only 0.1% and with a specificity of 98.9% [91,92].

CAncer Personalized Profiling by deep Sequencing (CAPP-Seq) was developed to detect extremely low concentrations of ctDNA by the use of "selectors" consisting of biotinylated DNA oligonucleotides that are complementary to previously defined recurrent mutated areas. Hybridization of the "selectors" on the area-of-interest is followed by deep sequencing; thereby, multiple mutations can be detected by CAPP-Seq including single nucleotide variants, rearrangements, and copy number alterations [93]. Implementing CAPP-Seq on blood samples of patients with early and advanced stage NSCLC, showed a high efficiency for detecting an allele frequency of *EGFR* mutations of down to 0.02% with >96% specificity [93,94]. Further improving the sensitivity of the CAPP-Seq, Newman et al. employed an integrated digital error suppression (iDES), a computational tool that can correct sequencing or PCR system error, resulting in a theoretical detection rate of 0.00025% mutant allele frequency [95]. iDES-enhanced CAPP-Seq has shown to be highly sensitive in the detection of *EGFR* mutations with an allele frequency as low as 0.004% with >99.99% specificity using cfDNA of NSCLC patients; furthermore, the required amount for the library preparation was only 32 ng [95], making it a very practical test for investigating ctDNA.

Although many advances have been made, NGS is still a relatively expensive and time-consuming technique. Furthermore, skilled bioinformaticians are required for data analysis and interpretation.

Bioinformatics are an essential part for the analysis of NGS to enable the detection of single nucleotide polymorphisms (SNPs), copy number aberrations (CNAs), insertions and deletions (indels), epigenetic changes, or to assembling new genomes [96–98]. The lack of standardization thus far, has led to the development of different algorithms performing essentially similar tasks in analyzing sequencing data, but using different mathematics. For instance, Burrows-Wheeler Alignment tool (BWA) [99], Bowtie [100], STAR [101], TopHap, and Novoalign are all short reads alignment tools [102]. Furthermore, variant calling can be performed using, e.g., GATK [103], SAM tools [104], Atlas2 [105], and FreeBayes [106]. In order to come to a possible consensus, the performance of these different tools must be regularly compared under different conditions. To assess the accuracy in variant calls, Bao et al. evaluated the four variant-calling algorithms, GATK-UnifiedGenotyper, SAMtools mpileup, Atlas2, and FreeBayes after alignment to the human genome using BWA, Bowtie2, and NovoalignV3. The authors used the NIST-GIAB gold standard dataset to demonstrate the sensitivities of these methods. Variant calls by FreeBayes from Novoalign V3 mapped sequences showed the highest sensitivity and precision rate for SNV calling of 95.97% and 99.70% and for indel calling 83.39% and 99.57%, respectively [102]. However, using simulated data, conflicting results were demonstrated by Kockan et al., indicating a low sensitivity and accuracy by using FreeBayes compared to SiNVICT, MuTect, and VarScan2 [107]. In the same study, the authors evaluated the sensitivity and accuracy of SiNVICT in the detection of SNVs and short indels of cfDNA. By analyzing two different datasets obtained from cfDNA sequenced material of castrate-resistant prostate cancer with Ion Torrent (AmpliSeq) technology and from metastatic castration-resistant prostate cancer patients sequenced with Illumina MiSeq, the SiNVICT demonstrated highly sensitive detection of variant calls at a low variant allele frequency of 0.5% [107]. These studies show that further investigation has to be performed in order to determine the most accurate methods for analyzing ctDNA.

4.2. Digital-PCR platforms

Digital PCR is a robust method to detect point mutations in ctDNA at low allele fractions. This technique includes droplet-based systems, microfluidic platforms for parallel PCR such as droplet digital PCR (ddPCR), and BEAMing (beads, emulsions, amplification, and magnetics).

Droplet-digital PCR (ddPCR) was developed to provide high-precision, absolute quantification of copy number variation of target DNA, such as quantification of somatic mutations [108]. The ddPCR approach is based on water-oil emulsion droplet technology by the distribution of DNA sample into thousands to millions of droplets. A single droplet contains a single mutated or non-mutated DNA strand that can be distinguished by flow cytometry using fluorescent TaqMan-based probes. ddPCR has been applied in several notable publications on the detection and quantification of mutations in ctDNA [78,109,110]. ddPCR demonstrated accurate detection of *PIK3CA* mutations in early stage breast cancer patients using ctDNA compared to tumor tissue with 93.3% sensitivity and 100% specificity [78]. Furthermore, Picodroplet digital PCR facilitates simultaneous screening for multiple mutations in ctDNA from the plasma with a detection rate of >1% [111]. The advantages of ddPCR are the high sensitivity in detecting mutations and as well as it being an inexpensive technology for absolute quantification. The disadvantages of ddPCR are that only known variants can be screened and the limited number of variants that can be investigated within a single reaction.

BEAMing is a digital PCR method that is based on beads, emulsion, amplification, and magnetics. This technology uses water droplets in an oil emulsion as reaction vessels containing a mixture of template,

primers, PCR reagents, and magnetic beads. Fluorescently labeled dideoxynucleotide terminators are used to discriminate droplets containing sequences that diverge at positions of interest and analyzed by flow cytometry [112]. This technique can identify genetic variations present in the original DNA population and precisely quantify their number in comparison to the number of wild-type sequences [113]. BEAMing has shown a highly sensitive detection rate of 0.02% mutant allele frequency and a perfect specificity of 100%, with >90% concordance rate between tumor tissue and ctDNA from different patients with colorectal [35], breast [114], and lung [112,115,116] cancer. Although BEAMing is a highly sensitive and specific, its workflow is complicated and expensive to apply in routine clinical work.

4.3. Real-time PCR-based methods

Real-Time PCR represents a rapid and cheap method for amplification of nucleic acid. Its sensitivity to detect mutations in a background of wildtype DNA is 10–20% allele frequency, with almost no false positives [117,118]. To overcome the low sensitivity however, several PCR-based variations have been developed, such as Allele-Specific amplification (AS-PCR) [119–121], Allele-Specific Non-Extendable Primer Blocker PCR (AS-NEPB-PCR) [76], Peptide Nuclei Acid-Locked Nucleic Acid (PNA-LNA) PCR clamp [122–124], and co-amplification at lower denaturation temperature (COLD-PCR) [77]. Most of these assays are based on either using a blocking oligo at the 3′-end to block the amplification of the normal allele and allowing the amplification of the mutant allele or they make use of a modification step in the PCR protocol that enriches variant alleles from a mixture of wild-type and mutation-containing DNA. The AS-PCR is commonly used in clinical setting to detect single nucleotide variation (SNV) or small insertion/deletion in formalin-fixed, paraffin-embedded (FFPE) tumor tissues. However, as it exhibits 98% specificity and 92% sensitivity with a concordance of 96% of the mutant allele in ctDNA [119], it is not fully adequate for the detection of rare genetic events. The PNA-LNA PCR clamp method shows a high sensitivity with the detection of 0.1% mutant allele and a specificity of 79% [122–124]. COLD-PCR is a powerful method to detect single variants of approximately 0.1% and enables the enrichment of this amount of a mutant allele to improve the sensitivity of mutation detection by up to 100-fold [75,77]. Overall, PCR based assays are a promising tool for detecting mutations as a low-cost effective can be feasible in routine clinical practice.

4.4. Mass-spectrometry technology

The limited multiplexing ability of most PCR-based approaches represents a major limitation when dealing with clinical samples. Alternative technologies using mass-spectrometry have been developed to detect ctDNA mutations at low frequency, namely Surface-Enhanced Raman Spectroscopy (SERS) [125] and UltraSEEK [126,127].

The SERS-PCR detection method is based on using nanotags, which are nanoparticulate optical detection tags that function through surface-enhanced Raman Spectroscopy (SERS) for identification and tracking the binding target. Direct detection of multiple mutations at the same time using a Raman spectrometer is being enabled by laser excitation resulting in the emission of specific signals [128]. Multiplex PCR/SERS demonstrated high detection affinity of three hotspot mutations in melanoma showing a high sensitivity detection of <0.1% mutations with a low input amount of 5 ng DNA per reaction [125].

UltraSEEK is a high-throughput multiplex based method, using primers labeled with biotin that are specifically designed to anneal the mutant allele only [126]. The UltraSEEK assay panel covering the most frequent mutations in melanoma, showed a high sensitivity of detecting mutations at an allele frequency of <1% and a 100% specificity. Moreover, the minimum amount of cfDNA employed in the UltraSEEK analysis is between 9 pg/µl and 4.2 ng/µl [126]. Recently, the UltraSEEK's capacity has been further improved to a multiplexing of up to 40 targets

per reaction, with ultrasensitive detection of somatic mutations in ctDNA [127]. Taken together, the advantages of UltraSEEK are the high multiplex capability, fast turnaround time of less than a day, and the low input of DNA required for a single analysis.

4.5. Detection of hypermethylation in ctDNA

Methylation of DNA involves the addition of a methyl group to CpG dinucleotides at regions of the genome with a high density of CpG dinucleotides or so-called CpG islands [7]. The most common method for methylation detection of ctDNA relies on methylation-specific PCR (MS-PCR), which is based on treating DNA with bisulfite to chemically modify non-methylated cytosines into uracil [71]. Subsequently, the methylation profile of the converted DNA can be investigated using a downstream application such as PCR, NGS [129], or MCTA-Seq [130]. Methylation-specific PCR (MS-PCR) has shown to be highly sensitive in the detection of *ESR1* hypermethylation with a detection rate of 0.1% and a specificity of 100% [71]. Higher sensitivities may be reached by MCTA-Seq, which is able to detect thousands of hypermethylated CpG islands in parallel with a sensitivity of detecting methylated CpG alleles down to frequencies of <0.25%, but with a specificity of 89%. Nevertheless, the input amount of ctDNA of 7.5 pg is a clear advantage [130]. The costs, processing time, and the requirement of prior knowledge of the region of interest are disadvantages of MCTA-Seq. A genome-wide bisulfite sequencing for the identification of different methylated regions using >500 ng urinary cfDNA starting material, could show that the global methylation density in cancer is ranging from 61.1% to 73.5% [129]. However, the relatively large amount of 500 ng cfDNA that is required for the bisulfite conversion process increases the complexity of the methylation detection using ctDNA from plasma [129].

5. Outlook

As this review indicates, numerous studies have now shown the feasibility of using ctDNA in tracking and monitoring tumor dynamics, drug response, and therapy resistance. Although several technologies have shown an extremely high sensitivity with detection rates going down to single mutated DNA molecules, the use of ctDNA as a marker for liquid biopsy still lacks standardization in many aspects. The only tests thus far approved by the FDA in the USA and China include the DNA methylation-based test of *SEPT9* for the detection of colorectal cancer [131,132] and the qPCR-based test for mutated *EGFR* in NSCLC [133]. Further improvement in the standardization of liquid biopsy may include how the samples are obtained and how the analysis is performed.

Ideally, ctDNA should be investigated in combination with CTCs and/or exosomal miRNA, in order to extract as much biological information from the tumor as possible from a single blood sample. However, the type of collection tube and storage conditions may both have an effect on DNA stability as well as the stability of cells and thereby the amount of background and the quality of the material. Although fixatives may stabilize a tube's content required for transport of the material, not every fixative suitable for subsequent cellular or DNA analysis can be used in combination with RNA analysis. Also, too harsh fixation conditions can result in DNA interstrand crosslinking and thereby lowering the specificity of downstream analyses. It needs to be seen whether there will be one standard tube from which all analyses can be performed, although more likely will be that each biomarker will require its own dedicated collection tube.

An important aspect of mutation diagnostics, not limited to the analysis of ctDNA only, is the sheer amount of data that can be produced by current technologies such as NGS, which can be overwhelming from a clinical point of view. However, bioinformatic-based techniques are usually able to filter out the clinically most important information. Nevertheless, also standardization in regards of bioinformatic analysis needs to be achieved in order for such diagnostics to be reliably be applied in the clinic.

As discussed in this review, one of the hurdles of using ctDNA as liquid biopsy substrate is the usually low yield of material extracted from plasma. In order to obtain enough starting material for further downstream analyses such as deep sequencing, whole genome amplification (WGA) might be employed. However, further research has to be performed to study whether the currently available WGA methods are suitable for highly fragmented DNA, as well as whether the amplification is perfectly linear so that low frequency alleles are not lost.

Understanding the biological mechanisms of how ctDNA is released into the bloodstream may further improve the isolation of the tumor DNA as well as prognosis and prediction value. For instance, the specific enrichment of tumor-associated exosomes may provide undiluted information about potential metastatic sites and the resistance mechanisms of the still viable tumor cells under therapy. Equally important is to investigate the elimination rate of cfDNA from the bloodstream. Several mechanisms and organs appear to be responsible for cfDNA clearance from the bloodstream such as the kidney, liver, and spleen as well as nuclease degradation, and phagocytes [134–136]. Nevertheless, the kinetic dynamics of cfDNA still needs to be further investigated, as well as the best source of ctDNA, e.g., serum, plasma, urine, or other body liquids should be standardized.

ctDNA can play a vital complementary role along with other tumor-derived substrates as predictive biomarker. These other substrates include circulating tumor cells (CTCs) that provide essential information on tumor characteristics and metastatic development through investigation of DNA, RNA, or proteins, whereas cell-free nucleotides and exosomes can be an additional sources of information on tumorigenesis, possible therapeutic targets, and drug resistance mechanisms. Finally, platelets can carry information that may help to determine the tumor's origin. Overall, these tumors-substrates termed as liquid biopsy that can provide a more comprehensive picture together of the total clonal composition of tumor and therapy sensitivity and thereby, improve on clinical management and patient survival.

6. Conclusion

Liquid biopsy can provide valuable information about the biology and clinical characteristics of a tumor through different biomarkers released into the blood circulation. ctDNA can be employed to analyze the entire tumor genome and track drug response and/or therapy resistance. This can be achieved by either quantitative measurement of ctDNA in a blood sample or by the detection of mutations. A remarkable advancement in technologies for ctDNA detection and analysis has been observed in the last few years such as the significant progress made in NGS-based approaches in overcoming many of the challenges to reduce the error rate and improve sensitivity in ctDNA detection. Nevertheless, NGS-based approaches are still relatively expensive and consume much time. On the other hand, mass-spectrometry approaches provide a promising tool for ctDNA screening in terms of the cost, time, and low amounts of required input material, as well as their high sensitivity and specificity. Additionally, analysis by Real-Time PCR-based techniques is cost-effective, fast, and can be feasible in routine clinical practice for a limited number of biomarkers. Further development in the standardization of these techniques will make ctDNA a valuable substrate in the field of cancer diagnostics.

Acknowledgements

References

[1] Lee EY, Muller WJ. Oncogenes and tumor suppressor genes. Cold Spring Harb Perspect Biol 2010;2(10):a003236.
[2] Herceg Z, Hainaut P. Genetic and epigenetic alterations as biomarkers for cancer detection, diagnosis and prognosis. Mol Oncol 2007;1(1):26–41.
[3] Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100(1):57–70.
[4] Chakravarthi BV, Nepal S, Varambally S. Genomic and epigenomic alterations in cancer. Am J Pathol 2016;186(7):1724–35.
[5] Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nat Rev Genet 2016;17:487.
[6] Joosse SA, Gorges TM, Pantel K. Biology, detection, and clinical implications of circulating tumor cells. EMBO Mol Med 2015;7(1):1–11.
[7] Follain G, et al. Hemodynamic forces tune the arrest, adhesion, and extravasation of circulating tumor cells. Dev Cell 2018;45(1):33–52 e12.
[8] Pantel K, Brakenhoff RH. Dissecting the metastatic cascade. Nat Rev Cancer 2004;4: 448.
[9] McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. Cell 2017;168(4):613–28.
[10] Joosse SA, Pantel K. Genetic traits for hematogeneous tumor cell dissemination in cancer patients. Cancer Metastasis Rev 2016;35(1):41–8.
[11] Joosse SA, Pantel K. Tumor-educated platelets as liquid biopsy in cancer patients. Cancer Cell 2015;28(5):552–4.
[12] O'Flaherty L, Wikman H, Pantel K. Biology and clinical significance of circulating tumor cell subpopulations in lung cancer. Transl Lung Cancer Res 2017;6(4): 431–43.
[13] Joosse SA, et al. Changes in keratin expression during metastatic progression of breast cancer: impact on the detection of circulating tumor cells. Clin Cancer Res 2012;18(4):993–1003.
[14] Meng S, et al. Circulating tumor cells in patients with breast cancer dormancy. Clin Cancer Res 2004;10(24):8152–62.
[15] Tseng J-Y, et al. Dynamic changes in numbers and properties of circulating tumor cells and their potential applications. Cancer 2014;6(4):2369–86.
[16] Best, Myron G., et al., RNA-Seq of tumor-educated platelets enables blood-based pan-cancer, multiclass, and molecular pathway cancer diagnostics. Cancer Cell. 28(5): p. 666–676.
[17] McAllister SS, Weinberg RA. The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. Nat Cell Biol 2014;16:717.
[18] Sol N, Wurdinger T. Platelet RNA signatures for the detection of cancer. Cancer Metastasis Rev 2017;36(2):263–72.
[19] Heitzer E, et al. The potential of liquid biopsies for the early detection of cancer. NPJ Precis Oncol 2017;1(1):36.
[20] Zhou W, et al. Cancer-secreted miR-105 destroys vascular endothelial barriers to promote metastasis. Cancer Cell 2014;25(4):501–15.
[21] Bardelli A, Pantel K. Liquid biopsies, what we do not know (yet). Cancer Cell 2017; 31(2):172–9.
[22] Mackie FL, et al. Cell-free fetal DNA-based noninvasive prenatal testing of aneuploidy. Obstet Gynaecol 2017;19(3):211–8.
[23] Wan JCM, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. Nat Rev Cancer 2017;17(4):223–38.
[24] Heitzer E, et al. Establishment of tumor-specific copy number alterations from plasma DNA of patients with cancer. Int J Cancer 2013;133(2):346–56.
[25] Jiang P, et al. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. Proc Natl Acad Sci 2015;112(11):E1317–25.
[26] Minciacchi VR, et al. Extracellular vesicles for liquid biopsy in prostate cancer: where are we and where are we headed? Prostate Cancer Prostatic Dis 2017;20 (3):251–8.
[27] Schwarzenbach H, et al. Detection and monitoring of cell-free DNA in blood of patients with colorectal cancer. Ann N Y Acad Sci 2008;1137:190–6.
[28] Lanman RB, et al. Analytical and clinical validation of a digital sequencing panel for quantitative, highly accurate evaluation of cell-free circulating tumor DNA. PLOS One 2015;10(10):e0140712.
[29] Garcia-Olmo DC, et al. Cell-free nucleic acids circulating in the plasma of colorectal cancer patients induce the oncogenic transformation of susceptible cultured cells. Cancer Res 2010;70(2):560–7.
[30] Trejo-Becerril C, et al. Cancer progression mediated by horizontal gene transfer in an in vivo model. PLoS One 2012;7(12):e52754.
[31] Wang W, et al. Characterization of the release and biological significance of cell-free DNA from breast cancer cell lines. Oncotarget 2017;8(26):43180–91.
[32] Alix-Panabieres C, Schwarzenbach H, Pantel K. Circulating tumor cells and circulating tumor DNA. Annu Rev Med 2012;63:199–215.
[33] Kurdyukov S, Bullock M. DNA methylation analysis: choosing the right method. Biology (Basel) 2016;5(1).
[34] Fernandez-Cuesta L, et al. Identification of circulating tumor DNA for the early detection of small-cell lung cancer. EBioMedicine 2016;10:117–23.
[35] Diehl F, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med 2008;14 (9):985–90.
[36] Gedvilaitė V, Schveigert D, Cicėnas S. Cell-free DNA in non-small cell lung cancer. Acta Med Lituanica 2017;24(2):138–44.
[37] Esposito A, et al. The Emerging role of "liquid biopsies," circulating tumor cells, and circulating cell-free tumor DNA in lung cancer diagnosis and identification of resistance mutations. Curr Oncol Rep 2017;19(1):1.
[38] Perez-Barrios C, et al. Comparison of methods for circulating cell-free DNA isolation using blood from cancer patients: impact on biomarker testing. Transl Lung Cancer Res 2016;5(6):665–72.
[39] Sorber, L., et al., A comparison of cell-free DNA isolation kits. J Mol Diagnost. 19(1): p. 162–168.
[40] Jorgez CJ, et al. Quantity versus quality: optimal methods for cell-free DNA isolation from plasma of pregnant women. Genet Med 2006;8(10):615–9.
[41] Koo KM, et al. Poly(A) extensions of miRNAs for amplification-free electrochemical detection on screen-printed gold electrodes. Anal Chem 2016;88(4):2000–5.
[42] Fong SL, et al. Comparison of 7 methods for extracting cell-free DNA from serum samples of colorectal cancer patients. Clin Chem 2009;55(3):587–9.

[43] Kirsch C, et al. An improved method for the isolation of free-circulating plasma DNA and cell-free DNA from other body fluids. Ann N Y Acad Sci 2008;1137:135–9.

[44] Rather RA, Saha SC, Dhawan V. The most favourable procedure for the isolation of cell-free DNA from the plasma of iso-immunized RHD-negative pregnant women. J Circ Biomark 2015;4:12.

[45] Wu T-L, et al. Cell-free DNA: measurement in various carcinomas and establishment of normal reference range. Clin Chim Acta 2002;321(1):77–87.

[46] Schmidt B, et al. Improved method for isolating cell-free DNA. Clin Chem 2005;51 (8):1561–3.

[47] Muller I, et al. Identification of loss of heterozygosity on circulating free DNA in peripheral blood of prostate cancer patients: potential and technical improvements. Clin Chem 2008;54(4):688–96.

[48] Kloten V, et al. Liquid biopsy in colon cancer: comparison of different circulating DNA extraction systems following absolute quantification of KRAS mutations using Intplex allele-specific PCR. Oncotarget 2017;8(49):86253–63.

[49] Pantel K. Blood-based analysis of circulating cell-free DNA and tumor cells for early cancer detection. PLoS Med 2016;13(12):e1002205.

[50] Sozzi G, et al. Quantification of free circulating DNA as a diagnostic marker in lung cancer. J Clin Oncol 2003;21(21):3902–8.

[51] Catarino R, et al. Circulating DNA: diagnostic tool and predictive marker for overall survival of NSCLC patients. PLoS One 2012;7(6):e38559.

[52] Olsson E, et al. Serial monitoring of circulating tumor DNA in patients with primary breast cancer for detection of occult metastatic disease. EMBO Mol Med 2015;7(8): 1034–47.

[53] Shao X, et al. Quantitative analysis of cell-free DNA in ovarian cancer. Oncol Lett 2015;10(6):3478–82.

[54] Yoon KA, et al. Comparison of circulating plasma DNA levels between lung cancer patients and healthy controls. J Mol Diagn 2009;11(3):182–5.

[55] Sozzi G, et al. Analysis of circulating tumor DNA in plasma at diagnosis and during follow-up of lung cancer patients. Cancer Res 2001;61(12):4675–8.

[56] Kim K, et al. Circulating cell-free DNA as a promising biomarker in patients with gastric cancer: diagnostic validity and significant reduction of cfDNA after surgical resection. Ann Surg Treat Res 2014;86(3):136–42.

[57] Czeiger D, et al. Measurement of circulating cell-free DNA levels by a new simple fluorescent test in patients with primary colorectal cancer. Am J Clin Pathol 2011;135(2):264–70.

[58] Herrera LJ, et al. Quantitative analysis of circulating plasma DNA as a tumor marker in thoracic malignancies. Clin Chem 2005;51(1):113–8.

[59] Oshiro C, et al. PIK3CA mutations in serum DNA are predictive of recurrence in primary breast cancer patients. Breast Cancer Res Treat 2015;150(2):299–307.

[60] Kuo YB, et al. Comparison of KRAS mutation analysis of primary tumors and matched circulating cell-free DNA in plasmas of patients with colorectal cancer. Clin Chim Acta 2014;433:284–9.

[61] Bettegowda C, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 2014;6(224):224ra24.

[62] Yuan H, et al. A modified extraction method of circulating free DNA for epidermal growth factor receptor mutation analysis. Yonsei Med J 2012;53(1):132–7.

[63] Garcia-Murillas I, et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. Sci Transl Med 2015;7(302):302ra133-302ra133.

[64] Guttery DS, et al. Noninvasive detection of activating estrogen receptor 1 (ESR1) mutations in estrogen receptor-positive metastatic breast cancer. Clin Chem 2015;61(7):974–82.

[65] Thress KS, et al. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non–small cell lung cancer harboring EGFR T790M. Nat Med 2015;21:560.

[66] Fribbens C, et al. Plasma ESR1 mutations and the treatment of estrogen receptor–positive advanced breast cancer. J Clin Oncol 2016;34(25):2961–8.

[67] Kang G, et al. Detection of KIT and PDGFRA mutations in the plasma of patients with gastrointestinal stromal tumor. Target Oncol 2015;10(4):597–601.

[68] Amirouchene-Angelozzi N, Swanton C, Bardelli A. Tumor evolution as a therapeutic target. Cancer Discov 2017;7(8):805–17.

[69] Skrypkina I, et al. Concentration and methylation of cell-free DNA from blood plasma as diagnostic markers of renal cancer. Dis Markers 2016;2016:3693096.

[70] Gifford G, et al. The acquisition of Hmlh1methylation in plasma DNA after chemotherapy predicts poor survival for ovarian cancer patients. Clin Cancer Res 2004;10 (13):4420–6.

[71] Mastoraki S, et al. ESR1 methylation: a liquid biopsy-based epigenetic assay for the follow-up of patients with metastatic breast cancer receiving endocrine treatment. Clin Cancer Res 2018 Mar 15;24(6):1500–10.

[72] Lee H, et al. A novel strategy for highly efficient isolation and analysis of circulating tumor-specific cell-free DNA from lung cancer patients using a reusable conducting polymer nanostructure. Biomaterials 2016;101:251–7.

[73] Forshew T, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 2012;4(136):136ra68-136ra68.

[74] Kennedy SR, et al. Detecting ultralow-frequency mutations by duplex sequencing. Nat Protoc 2014;9(11):2586–606.

[75] Cristofanilli M, Braun S. Circulating tumor cells revisited. JAMA 2010;303.

[76] Wang H, et al. Allele-specific, non-extendable primer blocker PCR (AS-NEPB-PCR) for DNA mutation detection in cancer. J Mol Diagn 2013;15(1):62–9.

[77] Freidin MB, et al. Circulating tumor DNA outperforms circulating tumor cells for KRAS mutation detection in thoracic malignancies. Clin Chem 2015;61(10): 1299–304.

[78] Beaver JA, et al. Detection of cancer DNA in plasma of patients with early-stage breast cancer. Clin Cancer Res 2014;20(10):2643–50.

[79] Glenn TC. Field guide to next-generation DNA sequencers. Mol Ecol Resour 2011;11 (5):759–69.

[80] Snyder MW, et al. Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin. Cell 2016;164(0):57–68.

[81] Paweletz CP, et al. Bias-corrected targeted next-generation sequencing for rapid, multiplexed detection of actionable alterations in cell-free DNA from advanced lung cancer patients. Clin Cancer Res 2016;22(4):915–22.

[82] Narayan A, et al. Ultrasensitive measurement of hotspot mutations in tumor DNA in blood using error-suppressed multiplexed deep sequencing. Cancer Res 2012; 72(14):3492–8.

[83] Couraud S, et al. Noninvasive diagnosis of actionable mutations by deep sequencing of circulating free DNA in lung cancer from never-smokers: a proof-of-concept study from BioCAST/IFCT-1002. Clin Cancer Res 2014;20(17):4613–24.

[84] Uchida J, et al. Diagnostic accuracy of noninvasive genotyping of EGFR in lung cancer patients by deep sequencing of plasma cell-free DNA. Clin Chem 2015;61(9): 1191–6.

[85] Abbosh C, et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. Nature 2017;545(7655):446–51.

[86] Belic J, et al. mFast-SeqS as a monitoring and pre-screening tool for tumor-specific aneuploidy in plasma DNA. Adv Exp Med Biol 2016;924:147–55.

[87] Belic J, et al. Rapid identification of plasma DNA samples with increased ctDNA levels by a modified FAST-SeqS approach. Clin Chem 2015;61(6):838–49.

[88] Kinde I, et al. FAST-SeqS: a simple and efficient method for the detection of aneuploidy by massively parallel sequencing. PLoS One 2012;7(7):e41162.

[89] Gale D, et al. Development of a highly sensitive liquid biopsy platform to detect clinically-relevant cancer mutations at low allele fractions in cell-free DNA. PLoS One 2018;13(3):e0194630.

[90] Kinde I, et al. Detection and quantification of rare mutations with massively parallel sequencing. Proc Natl Acad Sci 2011;108(23):9530–5.

[91] Tie J, et al. Circulating tumor DNA as an early marker of therapeutic response in patients with metastatic colorectal cancer. Ann Oncol 2015;26(8):1715–22.

[92] Fredebohm J, et al. Detection and Quantification of KIT Mutations in ctDNA by Plasma Safe-SeqS. Cham: Springer International Publishing; 2016.

[93] Newman AM, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med 2014;20(5):548–54.

[94] Chabon JJ, et al. Circulating tumour DNA profiling reveals heterogeneity of EGFR inhibitor resistance mechanisms in lung cancer patients. Nat Commun 2016;7: 11815.

[95] Newman AM, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. Nat Biotechnol 2016;34(5):547–55.

[96] Roy S, et al. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines. J Mol Diagn 2018;20(1):4–27.

[97] Chen J, et al. Translational bioinformatics for diagnostic and prognostic prediction of prostate cancer in the next-generation sequencing era. Biomed Res Int 2013; 2013:901578.

[98] Wu D, Rice CM, Wang X. Cancer bioinformatics: a new approach to systems clinical medicine. BMC Bioinformatics 2012;13(1):71.

[99] Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 2009;25(14):1754–60.

[100] Langmead B, et al. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009;10(3):R25.

[101] Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29 (1):15–21.

[102] Bao R, et al. Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. Cancer Informatics 2014 (13s2):CIN.S13779.

[103] McKenna A, et al. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20(9): 1297 303.

[104] Chen K, et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nat Methods 2009;6:677.

[105] Challis D, et al. An integrative variant analysis suite for whole exome next-generation sequencing data. BMC Bioinformatics 2012;13(1):8.

[106] Garrison E, Marth G. Haplotype-Based Variant Detection From Short-Read Sequencing, vol. 1207; 2012.

[107] Kockan C, et al. SiNVICT: ultra-sensitive detection of single nucleotide variants and indels in circulating tumour DNA. Bioinformatics 2017;33(1):26–34.

[108] Kristensen LS, Hansen LL. PCR-based methods for detecting single-locus DNA methylation biomarkers in cancer diagnostics, prognostics, and response to treatment. Clin Chem 2009;55(8):1471–83.

[109] Russo M, et al. Tumor heterogeneity and lesion-specific response to targeted therapy in colorectal cancer. Cancer Discov 2016;6(2):147–53.

[110] Sefrioui, D., et al., Clinical value of chip-based digital-PCR platform for the detection of circulating DNA in metastatic colorectal cancer. Digest Liver Dis. 47(10): p. 884–890.

[111] Taly V, et al. Multiplex picodroplet digital PCR to detect KRAS mutations in circulating DNA from the plasma of colorectal cancer patients. Clin Chem 2013;59(12): 1722–31.

[112] Oxnard GR, et al. Association between plasma genotyping and outcomes of treatment with osimertinib (AZD9291) in advanced non-small-cell lung cancer. J Clin Oncol 2016;34(28):3375–82.

[113] Li M, et al. BEAMing up for detection and quantification of rare sequence variants. Nat Methods 2006;3(2):95–7.

[114] Higgins MJ, et al. Detection of tumor PIK3CA status in metastatic breast cancer using peripheral blood. Clin Cancer Res 2012;18(12):3462–9.

[115] Taniguchi K, et al. Quantitative detection of EGFR mutations in circulating tumor DNA derived from lung adenocarcinomas. Clin Cancer Res 2011;17 (24):7808–15.

[116] Thress KS, et al. EGFR mutation detection in ctDNA from NSCLC patient plasma: a cross-platform comparison of leading technologies to support the clinical development of AZD9291. Lung Cancer 2015;90(3):509–15.

[117] Michael P. High sensitivity detection of tumor gene mutations. BAOJ Cancer Res Therapy 2015;1(1):1–6.

[118] Lemmon GH, Gardner SN. Predicting the sensitivity and specificity of published real-time PCR assays. Ann Clin Microbiol Antimicrob 2008;7:18.

[119] Thierry AR, et al. Clinical validation of the detection of KRAS and BRAF mutations from circulating tumor DNA. Nat Med 2014;20(4):430–5.

[120] Little S. Amplification-refractory mutation system (ARMS) analysis of point mutations. Curr Protoc Hum Genet 2001 May Chapter 9:Unit 9.8.

[121] Veldore VH, et al. Validation of liquid biopsy: plasma cell-free DNA testing in clinical management of advanced non-small cell lung cancer. Lung Cancer (Auckl) 2018;9:1–11.

[122] Miyazawa H, et al. Peptide nucleic acid-locked nucleic acid polymerase chain reaction clamp-based detection test for gefitinib-refractory T790M epidermal growth factor receptor mutation. Cancer Sci 2008;99(3):595–600.

[123] Kim H-R, et al. Detection of EGFR mutations in circulating free DNA by PNA-mediated PCR clamping. J Exp Clin Cancer Res 2013;32(1):50.

[124] Watanabe K, et al. EGFR mutation analysis of circulating tumor DNA using an improved PNA-LNA PCR clamp method. Can Respir J 2016;2016:5297329.

[125] Wee EJ, et al. Simple, sensitive and accurate multiplex detection of clinically important melanoma DNA mutations in circulating tumour DNA with SERS nanotags. Theranostics 2016;6(10):1506–13.

[126] Mosko MJ, et al. Ultrasensitive detection of multiplexed somatic mutations using MALDI-TOF mass spectrometry. J Mol Diagn 2016;18(1):23–31.

[127] Wong SQ, et al. Assessing the clinical value of targeted massively parallel sequencing in a longitudinal, prospective population-based study of cancer patients. Br J Cancer 2015;112(8):1411–20.

[128] Harper MM, McKeating KS, Faulds K. Recent developments and future directions in SERS for bioanalysis. Phys Chem Chem Phys 2013;15(15):5312–28.

[129] Cheng THT, et al. Genomewide bisulfite sequencing reveals the origin and time-dependent fragmentation of urinary cfDNA. Clin Biochem 2017;50(9): 496–501.

[130] Wen L, et al. Genome-scale detection of hypermethylated CpG islands in circulating cell-free DNA of hepatocellular carcinoma patients. Cell Res 2015;25 (11):1250–64.

[131] Song L, et al. A systematic review of the performance of the SEPT9 gene methylation assay in colorectal cancer screening, monitoring, diagnosis and prognosis: SEPT9 assay in CRC diagnosis and therapy, Vol. 18; 2017; 1–8.

[132] Song L, et al. The performance of the SEPT9 gene methylation assay and a comparison with other CRC screening tests: A meta-analysis, Vol. 7; 2017.

[133] Brown P. The Cobas® EGFR Mutation Test v2 assay. Future Oncol 2016;12(4): 451–2.

[134] Yu SCY, et al. High-resolution profiling of fetal DNA clearance from maternal plasma by massively parallel sequencing. Clin Chem 2013;59(8):1228–37.

[135] Velders M, et al. Exercise is a potent stimulus for enhancing circulating DNase activity. Clin Biochem 2014;47(6):471–4.

[136] Choi J-J, Reich CF, Pisetsky DS. The role of macrophages in the in vitro generation of extracellular DNA from apoptotic and necrotic cells. Immunology 2005;115(1): 55–62.

6. Discussion

Tumor heterogeneity is a hallmark of cancer and one of the leading causes of cancer therapy resistance, tumor progression, and metastasis. There is a high degree of heterogeneity, which requires a comprehensive sampling of each metastatic lesion via several and repetitive tissue biopsies. In clinical practice, this is not possible because tissue biopsies are restricted to a few sampling time-points and available locations and because common sites of metastasis that are difficult to biopsies, such as bones, lungs and brain. Liquid biopsy is an alternative strategy for real-time monitoring of drug response and resistance, as well as assessing tumor heterogeneity and understanding the biology of metastatic development (see publication #5).

6.1. Characterization of CTCs provide insights into the metastatic progression

CTCs are one of the liquid biopsy-derived materials. CTCs are derived from both primary and metastatic lesions; thus, CTCs represent an intermediate stage of metastasis [110]. Although not all tumor cell subpopulations are able to metastasize and most of CTCs die, only a few CTCs can extravasate into other tissues, survive in a dormant state, escape the immune system and systemic therapy, and ultimately grow and forming metastasis [11, 111]. Adaptation to a new microenvironment and growth of a single tumor cell in a distant location need the cell to possess or acquire specific characteristics to

Discussion

grow into an overt metastasis that can be detected by clinical techniques [12]. The epithelial to mesenchymal transition (EMT) is thought to play a crucial role in cancer propagation and metastasis. During this process, downregulation of intercellular adhesive complexes (e.g., E-cadherin-based adherens junctions), followed by loss of apicobasal polarity, gaining carcinoma cells' ability to migrate and invade [14]. The reverse process, known as mesenchymal to epithelial transition (MET), is essential in the formation of metastatic tumors. The high plasticity of carcinoma cells allows them to change from epithelial to mesenchymal-like phenotypes in a dynamic and reversible manner. A key aspect of the EMT is the reorganization of the cytoskeleton, which includes alterations in intermediate filaments, that may contribute to the induction of cell motility [85]. The heterogeneity of CTCs in phenotypic and genetic plasticity contribute to modulate therapy effectiveness and show different capacities to metastasize [21]. Therefore, CTCs phenotype is essential for identifying biomarkers that may help recognize tumor cells that can initiate metastatic colonization at distant sites. Using CTCs as a liquid biopsy, we identified in the current study three tumor cell subsets with different expression patterns of keratins (K16+/C11+, K16+/C11-, K16-/C11+) during the tumor progression of patients with metastatic breast cancer. A significantly shorter-free survival was observed for patients whose CTCs overexpressed K16 compared to patients who had CTCs with negative K16 expression. We found that K16 has a positive correlation to an intermediate mesenchymal phenotype and was mainly

116

observed in cells that have a hybrid phenotype of epithelial and mesenchymal cell features.

Another important finding is that K16 showed a regulatory effect on EMT, allowing epithelial carcinoma cells to undergo various morphological and biochemical changes that enable them to become more plastic and thus able to migrate (see publication #1). Yuanhua L. and colleagues showed that K16 has a regulatory role in EMT and that the transcription factor TF-AP2A in EMT-related pathways induced K16 expression in lung adenocarcinoma [112]. In consistence, we found that under the conditions of inducing K16 expression, a reorganization of actin microfilaments was observed in MCF7 cells forming long, parallel, thin stress fibers. These modifications in actin microfilaments seem to be a motivating force behind disrupting intercellular adhesion and directional migration. In line with this, on the one hand, induction of K16 expression has been shown to enhance migration in MCF7 cells, on the other hand, K16 depletion impaired cell migration. It is then tempting to hypothesize that K16 regulates the plasticity and reorganization of the actin microfilaments to facilitate cell migration, which is a critical step in the metastatic process. Based on our findings in the current study, K16 may represent a novel metastasis-associated protein in breast cancer via (i) regulating cell motility and (ii) inducing the EMT regulator genes. Therefore, assessing and monitoring the K16 status in CTCs may provide predictive information that helps identify patients whose tumors are more prone to metastasize. (see publication #1 and Appendix 8.1).

Discussion

6.2. Potential of ctDNA to monitor clinically relevant cancer-related epigenetic modifications

Another blood-based biomarker of liquid biopsy is ctDNA, which originates directly from the tumor or CTCs that shed from primary tumors and/or metastatic lesions. Here, we reported that the concentration of cfDNA was significantly higher in ovarian cancer patients compared to healthy controls, indicating the use of cfDNA as a potential biomarker for cancer diagnosis (see publications #2 and #5). The quantification of cfDNA concentration has been studied in order to discriminate between benign and malignant diseases (see publication #5).

In ovarian cancer, several studies have documented the distinct cfDNA concentrations between patients and healthy controls; they found that the cfDNA level has been increased in patients with advanced-stage compared to a benign tumor and healthy controls [42-44]. Furthermore, ctDNA has shown a high potential for tracking clinically significant cancer-related epigenetic changes and provides direct information about the methylomic make-up and genomic alteration of the tumor [37, 113, 114]. Previously, we have shown that *BRCA1* promoter hypermethylation was more frequent in primary tumors 73.7% than recurrent tumors 20.8% [105]. In the present study, *BRCA1* methylation frequency was significantly higher in ovarian cancer patients in which 46% of patients lost the function of *BRCA1* during their treatment; as a result of

Discussion

hypermethylation of the *BRCA1* promoter, which is contributing to impair the tumor cells to repair DNA cross-links introduced by chemotherapy agents (see publication #2). Therefore, we tested whether the identification of the methylation status of the *BRCA1* promoter using ctDNA might have a predictive role in ovarian cancer patients' outcomes.

A significant limitation in verifying the therapeutic importance of drug resistance is the challenge of collecting tumor biopsies following initial diagnosis, at a period when resistant subpopulations might be more evident. ctDNA has shown a high potential in tracking and monitoring tumor dynamics in several previous studies [115-118]. We developed a highly sensitive and specific liquid biopsy assay that could detect down to a single molecule (<0.03%) in a high background of normal DNA. Five CpG sites were analyzed to verify the enrichment of methylated DNA sites, which were previously associated with very low *BRCA1* expression in breast cancer cell lines [119] (see publication #2). Employing this sensitive method on multiple blood samples from ovarian cancer patients during the course of treatment, we detected a conversion of methylation status in 24% of methylation-positive patients, indicating a possible development of therapy resistance (see publication #2). Since all ovarian cancer patients' treatment regimens were heterogeneous, and because they can have several relapses, the survival analyses were conducted using two multivariate models. These models were applied to test two assumptions about how a tumor acquires medication resistance either due to therapy-induced evolution or as a result of selection (see publication #2 and

Discussion

Appendix 8.2). Multivariable analyses showed that the methylation status of the *BRCA1* promoter was an independent predictor of survival, assuming that relapse is independent of previous relapses. Further, our survival analysis showed that ovarian cancer patients with methylated *BRCA1* promoter had a significantly lower risk for disease-related progression. Interestingly, patients who had *BRCA1* hypermethylated and subsequently converted to unmethylated *BRCA1* had lower median survival than the patients who maintained a constant positive methylation status throughout the course of the disease but they had better survival than patients who had an unmethylated *BRCA1* promoter. Hence, the detection of methylation patterns in ctDNA can be used to monitor the methylation status of *BRCA1* by analyzing serial blood samples during the disease and predicting the survival of ovarian cancer patients (see publication #2).

6.3. CTC phenotyping as a surrogate marker for therapeutic selection and monitoring of tumor resistance

Another well-established marker for the response after hormone therapy in breast cancer is the status of ER. The expression of ER often facilitates the sensitivity of breast tumors to hormonal therapy with either selective estrogen receptor modulators, such as tamoxifen, or aromatase inhibitors. Unfortunately, 30-40% of patients develop therapy resistance after 24-36 months on average [120]. CTCs offer a non-invasive real-time screening of tumor progression and

Discussion

represent an alternative to serial tissue biopsies [121]. Phenotyping of the CTCs can provide crucial information on the evolving characteristics of the tumor during progression and treatment resistance [36, 122, 123]. CTCs are routinely found in more than 60% of metastatic breast cancer patients, with enumeration indicating prognostic significance in disease progression in patients undergoing chemotherapy or endocrine treatment [124, 125]. The CTC analysis provides an opportunity to study individual clones at a single cell level, originating in metastatic distant tumor sites [21, 126]. The isolation of CTCs is extremely challenging due to the very low amount of CTCs (typically 1 to 100 CTC/7.5 mL blood) [10, 110]. Many commercially available instruments and test systems are approved as clinical diagnostic devices, which have allowed CTC to be identified, enumerated, and analyzed [32, 34, 35, 127]. The most famous system for the enumeration and isolation of CTCs is the gold standard, the FDA-cleared CellSearch® system [36, 127]. This system detects CTCs based on binding to anti-EpCAM, cytokeratin (CK), and CD45 expression. A recent meta-analysis of Yan *et al*. showed that the status of CTC could be used to assess systemic therapy success for metastatic breast cancer patients since a change in CTC status between two-time points was prognostic [128]. Although CTCs count carries independent prognostic information, the phenotype of CTCs may serve as a guideline for therapeutic management of breast cancer patients, particularly estrogen receptor, which is considered one of the main candidate targets of endocrine therapy. We sought to evaluate the ER status in CTCs to test whether it could affect the response after hormonal therapy. Using the ERα

Discussion

monoclonal murine ER-119.3 antibody, ER-positive CTCs were monitored using the CellSearch System for CTCs quantification (see publication #3). CTCs were identified in 31.9% of analyzable blood samples, 15% showed 1-4 CTC, and 16% showed ≥5 CTC. Among all CTC-positive samples, the samples taken before the therapy initiation have the highest CTC detection rate (44%), and the samples taken during endocrine therapy have the lowest detection rate of CTC (27 %). A favorable relationship between CTC-positive status and progression-free survival was detected during the course of the disease. Furthermore, a higher number of CTCs during therapy was associated with disease progression ($p < 0.0001$), whereas a lower number of CTCs or a CTC-negative status was associated with stable disease. A high heterogeneous expression of ER in individual CTCs was observed in patients, the ER-positive CTCs were detected in 32% of the CTC-positive samples, comparable to that seen by Paoletti *et al*. [129]. Although a shift in ER status of CTCs has been noticed during tumor progression, no conclusion was formed on endocrine therapy response and/or resistance due to the low number of patients who had initially been diagnosed with ER-positive CTCs. Therefore, Further large-scale studies with long-term monitoring of ER-CTCs status are needed to determine the predictive relevance of ER-CTC for endocrine treatment efficacy (see publication #3).

6.4. ctDNA specific-PCR enrichment

The overall amount of ctDNA from the tumor accounts as low as 0.01% of the total cfDNA present in the blood [130]. Due to the extremely low levels of ctDNA make, its detection is a challenge, especially in the early phases of tumor formation [131, 132]. Two strategies have evolved to study the tumor's genomic material by liquid biopsy. First, targeted approaches are employed to detect residual disease in the peripheral blood using a single or few tumor-specific mutations identified from the primary tumor [133]. Such techniques include Safe-SeqS [134], TamSeq [135], CAPP-Seq [136], BEAMing [137], and q-PCR [138, 139]. The drawback of this strategy is that it needs extensive knowledge of the tumor genome. Targeted monitoring, on the other hand, can be highly sensitive, as mutations can be identified with high specificity at a fast and cost-effective rate at allele frequencies as low as 0.01% [140]. A second strategy to investigate ctDNA includes non-target screening, which aims for a genome-wide analysis for copy number aberrations (CNAs) or point mutations using whole-genome sequencing (WGS) or whole-exome sequencing (WES). Untargeted techniques have the advantages (i) Its capacity to detect unique alterations that occur during tumor therapy and (ii) not requiring prior knowledge of the initial tumor's genome. However, significant amounts of ctDNA are necessary for successful reconstruction of tumor-specific genome-wide alterations, which is a disadvantage. Furthermore, untargeted methods have a poor overall sensitivity (5%–10%) [141]. Depending on which strategy is

Discussion

imperative to investigate the ctDNA of interest, variou technologies are currently available.

PCR-based assays are promising tools for detecting hotspot mutations with low input materials (e.g. cfDNA, CTCs) and low cost (see publication #5). We developed ctDNA specific enrichment method to screen known and unknown mutations at high sensitivity. The enrichment of ctDNA starts with blocking all wild-type alleles that may be potentially mutated in breast cancer in the *ESR1* gene. For this, blocking oligos are used that are modified at the 3'-end, prohibiting further amplification. Through co-amplification at lower denaturation temperatures, mutated alleles can, however, be amplified using generic primers, which can only bind if no blocking oligo has hybridized. In order to measure the sensitivity of our method, we synthesized DNA plasmids that encompass possible mutations. Through dilution series, we could show that we can detect DNA point mutations at a concentration down to 0.16% in a background of normal DNA. Sanger sequencing confirmed the specific enrichment of mutated DNA (Appendix 8.3). The main challenge was to parallelly detect multiple driving mutations. We could successfully multiplex up to two sets to cover the five most frequent mutation sites in the *ESR1* gene (Appendix 8.3).

Nevertheless, several hotspot mutations still need to be investigated from different genes in one PCR reaction. MassARRAY®-UltraSEEK® technology is a promising technique allowing multiplexing more than 40 SNPs per single

reaction (see publication #5). This technology has the benefit that it can perform in ctDNA and CTCs, adding to this fast and cost-effective (see publication #4).

6.5. Potential of combining ctDNA and CTCs to monitor endocrine therapy resistance mutations.

Biomarkers based on liquid biopsies, such as ctDNA and CTCs, together provide a comprehensive picture of tumor heterogeneity and dynamic response to therapy. Using UltraSEEK® Breast panel, major hotspot mutations in *ESR1*, *PI3KCA*, *FOXA1, GATA3*, *AKT1*, *ERBB2,* and *TP53* were examined in cfDNA and CTCs of ER-positive metastatic breast cancer patients during the course of treatment (see publication #4). The *ESR1*, *PIK3CA*, *FOXA1*, *GATA3*, and *TP53* genes were frequently mutated in metastatic breast cancer patients. In circulating DNA, *ESR1* was mutated in 53% of breast cancer patients who experienced progression, similar to the 30–56.4% circulating *ESR1* mutations often seen in metastatic breast cancer patients who were progressed on endocrine therapy [142-144]. The most frequently detected *ESR1* mutations, pD538G, pY537S, and pY537N, these mutations were reported as a cause of hormone-independent ER activation, resulting in resistance to endocrine treatment and overexpression of metastasis-associated genes [145]. The pY537S and pY537N mutations were a good prognostic factor in predicting tumor development in response to Aromatase inhibitors and/or chemotherapy,

Discussion

as well as being related to shorter overall survival, which is consistent with previous research [143].

Besides the *ESR1* mutant gene, *PIK3CA* mutations were frequently activated in metastatic breast cancer. *PIK3CA* mutations were detected in 29% of ER-positive metastatic breast cancer patients. The most common *PIK3CA* mutations observed during tumor progression were pH1047R, pE542K, and pE545K, consistent with previous studies [71, 73]. However, the mutations in *ESR1* and *PIK3CA* have recently been found as part of an acquired mutational process that leads to ER-positive breast cancer resistance and metastasis, 26.8% of patients whose tumors progressed did not have *ESR1* and *PIK3CA* mutations, but they emerged accumulation of *FOXA1* and *GATA3* subclonal mutations, suggesting an additional resistance mechanism implicated in tumor development during systemic treatment.

Recently, restoring ER function by recruitment transcription factors *FOXA1* and *GATA3* has been emphasized as one of the acquired resistance mechanisms in ER-positive breast cancer [75, 146]. *FOXA1* and *GATA3* transcription factors are involved in modulating chromatin condensation to allow ER recruitment in breast cancer cells [78]. In breast cancer, the genes *FOXA1* and *GATA3* are often altered, suggesting that they have a role in endocrine-resistant breast cancer [76, 78]. *FOXA1* was mutated in 31% of ER-positive metastatic breast cancer patients. The pE24K and pI176V mutations in *FOXA1* were considerably increased upon Fulvestrant and/or chemotherapy during tumor progression and at an advanced stage of disease. Wherase *GATA3* gene

was mutated in 36% of patients, pD336fs17 and pS93F mutations were observed in 92% of *GATA3* mutated samples during disease progression and at a progressive phase of the disease.

The *FOXA1* (pE24K) and *GATA3* (pD336fs17) mutations represent a significant risk to a patient's survival. Both mutations were significantly raised upon chemotherapy alone or combined with endocrine therapy agents during tumor progression and at a progressive phase of a tumor. Furthermore, patients with two or more mutant genes showed shorter progression-free survival and worse overall survival than patients with single mutant subclones. As shown by patients with only *ESR1* (pD538G) mutation did not show significant information for tumor progression, but in combination with *FOXA1* (pE24K) and *GATA3* (pD336fs17) mutations subclones, it showed crucial prognostic information on tumor progression and survival outcome. Also, patients with *ESR1* and *PIK3CA* mutant subclones combined with *GATA3* mutations exhibited shorter progression-free survival and overall survival than patients with only *ESR1* and *PIK3CA* mutant subclones.

The most interesting finding was that patients with *FOXA1* and *GATA3* mutations progressed rapidly in 2.5 months and showed a poor survival rate of 9 months compared to other mutant cell populations. Both ctDNA and CTCs in longitudinal time points represent tumor dynamics, which allowed the identification of signature subclones that may be significant in tumor management. *GATA3* mutations with *PIK3CA*, *FOXA1*, and/or *ESR1* mutations

are considered a strong predictor of tumor resistance and progression in metastatic breast cancer patients (see publication #4 and Appendix 8.4).

In conclusion

Liquid biopsy can give valuable information about a tumor's biology and clinical features by releasing biomarkers into the bloodstream. CTCs are a valuable source of information on tumor features and metastatic progression, whereas ctDNA can be additional sources of knowledge on tumorigenesis, potential therapeutic targets, and medication resistance processes. In the last few years, a tremendous improvement has been observed in technologies for detection and analysis of CTCs and ctDNA, such as the substantial success achieved in NGS-based techniques in overcoming many of the hurdles in CTCs and ctDNA detection to lower the error rate and enhance sensitivity. NGS-based techniques, on the other hand, are still relatively expensive and time-consuming. Mass-spectrometry techniques are a potential tool for CTCs and ctDNA screening because of their minimal cost, time, and input material requirements, and also their high sensitivity and specificity. Furthermore, for a limited number of biomarkers, analysis using Real-Time PCR-based methods is cost-effective, quick, and feasible in routine clinical practice. Further development in the standardization, liquid biopsy will become a significant substrate in the field of cancer diagnosis.

Discussion

We provide here several evidence to rationalize the use of liquid biopsies like ctDNA and CTCs to identify biomarkers that can be used to provide more detailed information on the whole clonal make-up of the tumor and dynamic therapy response. Combining CTCs with ctDNA-derived biomarkers would indeed improve the clinical management of the disease and patient outcomes. Furthermore, we present promising, cost-effective, and fast PCR-based enrichment methods to screen gene mutations and epigenetic modification in liquid biopsies at high sensitivity ev,en in multiplexing conditions.

7. References

1. Fouad, Y.A. and C. Aanei, *Revisiting the hallmarks of cancer.* American journal of cancer research, 2017. **7**(5): p. 1016-1036.

2. Jiang, W.G., et al., *Tissue invasion and metastasis: Molecular, biological and clinical perspectives.* Seminars in Cancer Biology, 2015. **35**: p. S244-S275.

3. Delbridge, A.R.D., L.J. Valente, and A. Strasser, *The role of the apoptotic machinery in tumor suppression.* Cold Spring Harbor perspectives in biology, 2012. **4**(11): p. a008789.

4. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**.

5. Shortt, J. and R.W. Johnstone, *Oncogenes in cell survival and cell death.* Cold Spring Harbor perspectives in biology, 2012. **4**(12): p. a009829.

6. Hovhannisyan, G., et al., *DNA Copy Number Variations as Markers of Mutagenic Impact.* International journal of molecular sciences, 2019. **20**(19): p. 4723.

7. Ghavifekr Fakhr, M., et al., *DNA methylation pattern as important epigenetic criterion in cancer.* Genetics research international, 2013. **2013**: p. 317569-317569.

8. Chakravarthi, B.V., S. Nepal, and S. Varambally, *Genomic and Epigenomic Alterations in Cancer.* Am J Pathol, 2016. **186**(7): p. 1724-35.

References

9. Allis, C.D. and T. Jenuwein, *The molecular hallmarks of epigenetic control.* Nature Reviews Genetics, 2016. **17**(8): p. 487-500.

10. Joosse, S.A., T.M. Gorges, and K. Pantel, *Biology, detection, and clinical implications of circulating tumor cells.* EMBO Mol Med, 2015. **7**(1): p. 1-11.

11. Pantel, K. and R.H. Brakenhoff, *Dissecting the metastatic cascade.* Nature Reviews Cancer, 2004. **4**(6): p. 448-456.

12. Kang, Y. and K. Pantel, *Tumor cell dissemination: emerging biological insights from animal models and cancer patients.* Cancer cell, 2013. **23**(5): p. 573-581.

13. Pantel, K. and C. Alix-Panabières, *Bone marrow as a reservoir for disseminated tumor cells: a special source for liquid biopsy in cancer patients.* BoneKEy reports, 2014. **3**: p. 584-584.

14. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**(5): p. 871-90.

15. Armstrong, A.J., et al., *Circulating tumor cells from patients with advanced prostate and breast cancer display both epithelial and mesenchymal markers.* Mol Cancer Res, 2011. **9**(8): p. 997-1007.

16. Bruner, H.C. and P.W.B. Derksen, *Loss of E-Cadherin-Dependent Cell-Cell Adhesion and the Development and Progression of Cancer.* Cold Spring Harbor perspectives in biology, 2018. **10**(3): p. a029330.

References

17. Liu, C.-Y., et al., *Vimentin contributes to epithelial-mesenchymal transition cancer cell mechanics by mediating cytoskeletal organization and focal adhesion maturation.* Oncotarget, 2015. **6**(18): p. 15966-15983.

18. Stemmler, M., et al., *Non-redundant functions of EMT transcription factors.* Nature Cell Biology, 2019. **21**.

19. Moustakas, A. and C.-H. Heldin, *Mechanisms of TGFβ-Induced Epithelial–Mesenchymal Transition.* Journal of Clinical Medicine, 2016. **5**: p. 63.

20. Williams, E.D., et al., *Controversies around epithelial–mesenchymal plasticity in cancer metastasis.* Nature Reviews Cancer, 2019. **19**(12): p. 716-732.

21. Keller, L. and K. Pantel, *Unravelling tumour heterogeneity by single-cell profiling of circulating tumour cells.* Nat Rev Cancer, 2019. **19**(10): p. 553-567.

22. Gorges, K., et al., *Intra-Patient Heterogeneity of Circulating Tumor Cells and Circulating Tumor DNA in Blood of Melanoma Patients.* Cancers (Basel), 2019. **11**(11).

23. Caravagna, G., et al., *Detecting repeated cancer evolution from multi-region tumor sequencing data.* Nature Methods, 2018. **15**(9): p. 707-714.

24. Pantel, K. and C. Alix-Panabières, *Real-time Liquid Biopsy in Cancer Patients: Fact or Fiction?* Cancer Research, 2013. **73**(21): p. 6384-6388.

25. Pantel, K. and C. Alix-Panabières, *Liquid biopsy: Potential and challenges.* Molecular Oncology, 2016. **10**(3): p. 371-373.

References

26.   Joosse, S.A. and K. Pantel, *Tumor-Educated Platelets as Liquid Biopsy in Cancer Patients.* Cancer Cell, 2015. **28**(5): p. 552-554.

27.   Arechederra, M., M.A. Ávila, and C. Berasain, *Liquid biopsy for cancer management: a revolutionary but still limited new tool for precision medicine.* Advances in Laboratory Medicine / Avances en Medicina de Laboratorio, 2020. **1**(3).

28.   Hartkopf, A.D., et al., *Circulating Tumor Cells in Early-Stage Breast Cancer.* Geburtshilfe und Frauenheilkunde, 2011. **71**(12): p. 1067-1072.

29.   Tseng, J.Y., et al., *Dynamic changes in numbers and properties of circulating tumor cells and their potential applications.* Cancers (Basel), 2014. **6**(4): p. 2369-86.

30.   Chen, L., A.M. Bode, and Z. Dong, *Circulating Tumor Cells: Moving Biological Insights into Detection.* Theranostics, 2017. **7**(10): p. 2606-2619.

31.   Ferreira, M.M., V.C. Ramani, and S.S. Jeffrey, *Circulating tumor cell technologies.* Molecular oncology, 2016. **10**(3): p. 374-394.

32.   Habli, Z., et al., *Circulating Tumor Cell Detection Technologies and Clinical Utility: Challenges and Opportunities.* Cancers (Basel), 2020. **12**(7).

33.   Muller, V., et al., *Prognostic impact of circulating tumor cells assessed with the CellSearch System and AdnaTest Breast in metastatic breast cancer patients: the DETECT study.* Breast Cancer Res, 2012. **14**(4): p. R118.

References

34.   Giordano, A., et al., *Epithelial-mesenchymal transition and stem cell markers in patients with HER2-positive metastatic breast cancer.* Mol Cancer Ther, 2012. **11**(11): p. 2526-34.

35.   Scherag, F.D., et al., *Highly Selective Capture Surfaces on Medical Wires for Fishing Tumor Cells in Whole Blood.* Anal Chem, 2017. **89**(3): p. 1846-1854.

36.   Guan, X., et al., *Analysis of the hormone receptor status of circulating tumor cell subpopulations based on epithelial-mesenchymal transition: a proof-of-principle study on the heterogeneity of circulating tumor cells.* Oncotarget, 2016. **7**(40): p. 65993-66002.

37.   Elazezy, M. and S.A. Joosse, *Techniques of using circulating tumor DNA as a liquid biopsy component in cancer management.* Computational and structural biotechnology journal, 2018. **16**: p. 370-378.

38.   Diehl, F., et al., *Circulating mutant DNA to assess tumor dynamics.* Nat Med, 2008. **14**(9): p. 985-90.

39.   Herbreteau, G., et al., *Circulating free tumor DNA in non-small cell lung cancer (NSCLC): clinical application and future perspectives.* Journal of thoracic disease, 2019. **11**(Suppl 1): p. S113-S126.

40.   Esposito, A., et al., *The Emerging Role of "Liquid Biopsies," Circulating Tumor Cells, and Circulating Cell-Free Tumor DNA in Lung Cancer Diagnosis and Identification of Resistance Mutations.* Curr Oncol Rep, 2017. **19**(1): p. 1.

References

41.   Bronkhorst, A.J., V. Ungerer, and S. Holdenrieder, *The emerging role of cell-free DNA as a molecular marker for cancer management.* Biomolecular detection and quantification, 2019. **17**: p. 100087-100087.

42.   Paweletz, C.P., et al., *Bias-Corrected Targeted Next-Generation Sequencing for Rapid, Multiplexed Detection of Actionable Alterations in Cell-Free DNA from Advanced Lung Cancer Patients.* Clin Cancer Res, 2016. **22**(4): p. 915-22.

43.   Mosko, M.J., et al., *Ultrasensitive Detection of Multiplexed Somatic Mutations Using MALDI-TOF Mass Spectrometry.* J Mol Diagn, 2016. **18**(1): p. 23-31.

44.   Wong, S.Q., et al., *Assessing the clinical value of targeted massively parallel sequencing in a longitudinal, prospective population-based study of cancer patients.* Br J Cancer, 2015. **112**(8): p. 1411-20.

45.   Taly, V., et al., *Multiplex picodroplet digital PCR to detect KRAS mutations in circulating DNA from the plasma of colorectal cancer patients.* Clin Chem, 2013. **59**(12): p. 1722-31.

46.   Lemmon, G.H. and S.N. Gardner, *Predicting the sensitivity and specificity of published real-time PCR assays.* Ann Clin Microbiol Antimicrob, 2008. **7**: p. 18.

47.   Heer, E., et al., *Global burden and trends in premenopausal and postmenopausal breast cancer: a population-based study.* Lancet Glob Health, 2020. **8**(8): p. e1027-e1037.

References

48.   Feng, Y., et al., *Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis.* Genes Dis, 2018. **5**(2): p. 77-106.

49.   Badowska-Kozakiewicz, A.M., et al., *Retrospective evaluation of histopathological examinations in invasive ductal breast cancer of no special type: an analysis of 691 patients.* Archives of medical science : AMS, 2017. **13**(6): p. 1408-1415.

50.   McCart Reed, A.E., et al., *Invasive lobular carcinoma of the breast: the increasing importance of this special subtype.* Breast Cancer Research, 2021. **23**(1): p. 6.

51.   Dai, X., et al., *Breast cancer intrinsic subtype classification, clinical use and future trends.* American journal of cancer research, 2015. **5**(10): p. 2929-2943.

52.   Yersal, O. and S. Barutca, *Biological subtypes of breast cancer: Prognostic and therapeutic implications.* World journal of clinical oncology, 2014. **5**(3): p. 412-424.

53.   Laible, M., et al., *Impact of molecular subtypes on the prediction of distant recurrence in estrogen receptor (ER) positive, human epidermal growth factor receptor 2 (HER2) negative breast cancer upon five years of endocrine therapy.* BMC cancer, 2019. **19**(1): p. 694-694.

54.   Cancello, G., et al., *Progesterone receptor loss identifies Luminal B breast cancer subgroups at higher risk of relapse.* Ann Oncol, 2013. **24**(3): p. 661-8.

References

55. Dai, X., et al., *Cancer Hallmarks, Biomarkers and Breast Cancer Molecular Subtypes.* Journal of Cancer, 2016. **7**(10): p. 1281-1294.

56. Yaşar, P., et al., *Molecular mechanism of estrogen-estrogen receptor signaling.* Reproductive medicine and biology, 2016. **16**(1): p. 4-20.

57. Weikum, E.R., X. Liu, and E.A. Ortlund, *The nuclear receptor superfamily: A structural perspective.* Protein science : a publication of the Protein Society, 2018. **27**(11): p. 1876-1892.

58. Saha, S., S. Dey, and S. Nath, *Steroid Hormone Receptors: Links With Cell Cycle Machinery and Breast Cancer Progression.* Frontiers in Oncology, 2021. **11**(460).

59. Tang, Z.-R., et al., *Estrogen-Receptor Expression and Function in Female Reproductive Disease.* Cells, 2019. **8**(10): p. 1123.

60. Pescatori, S., et al., *A Tale of Ice and Fire: The Dual Role for 17β-Estradiol in Balancing DNA Damage and Genome Integrity.* Cancers, 2021. **13**(7): p. 1583.

61. Lewis-Wambi, J.S. and V.C. Jordan, *Treatment of Postmenopausal Breast Cancer with Selective Estrogen Receptor Modulators (SERMs).* Breast Dis, 2005. **24**: p. 93-105.

62. Robertson, J.F., *Selective oestrogen receptor modulators/new antioestrogens: a clinical perspective.* Cancer Treat Rev, 2004. **30**(8): p. 695-706.

References

63.   Rani, A., et al., *Endocrine Resistance in Hormone Receptor Positive Breast Cancer–From Mechanism to Therapy.* Frontiers in Endocrinology, 2019. **10**(245).

64.   Lei, J.T., et al., *Endocrine therapy resistance: new insights.* Breast (Edinburgh, Scotland), 2019. **48 Suppl 1**(Suppl 1): p. S26-S30.

65.   Butti, R., et al., *Receptor tyrosine kinases (RTKs) in breast cancer: signaling, therapeutic implications and challenges.* Molecular cancer, 2018. **17**(1): p. 34-34.

66.   Jeselsohn, R., et al., *ESR1 mutations—a mechanism for acquired endocrine resistance in breast cancer.* Nature reviews. Clinical oncology, 2015. **12**(10): p. 573-583.

67.   Reinert, T., et al., *Clinical Implications of ESR1 Mutations in Hormone Receptor-Positive Advanced Breast Cancer.* Frontiers in oncology, 2017. **7**: p. 26-26.

68.   Urso, L., et al., *ESR1 Gene Mutation in Hormone Receptor-Positive HER2-Negative Metastatic Breast Cancer Patients: Concordance Between Tumor Tissue and Circulating Tumor DNA Analysis.* Frontiers in Oncology, 2021. **11**(403).

69.   Spoerke, J.M., et al., *Heterogeneity and clinical significance of ESR1 mutations in ER-positive metastatic breast cancer patients receiving fulvestrant.* Nature Communications, 2016. **7**(1): p. 11579.

References

70. Takeshita, T., et al., *Prevalence of ESR1 E380Q mutation in tumor tissue and plasma from Japanese breast cancer patients.* BMC cancer, 2017. **17**(1): p. 786-786.

71. Martinez-Saez, O., et al., *Frequency and spectrum of PIK3CA somatic mutations in breast cancer.* Breast Cancer Res, 2020. **22**(1): p. 45.

72. Keraite, I., et al., *PIK3CA mutation enrichment and quantitation from blood and tissue.* Scientific Reports, 2020. **10**(1): p. 17082.

73. Shimoi, T., et al., *PIK3CA mutation profiling in patients with breast cancer, using a highly sensitive detection system.* Cancer Sci, 2018. **109**(8): p. 2558-2566.

74. Wang, M., et al., *The Predictive Role of PIK3CA Mutation Status on PI3K Inhibitors in HR+ Breast Cancer Therapy: A Systematic Review and Meta-Analysis.* BioMed research international, 2020. **2020**: p. 1598037-1598037.

75. Theodorou, V., et al., *GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility.* Genome research, 2013. **23**(1): p. 12-22.

76. Hurtado, A., et al., *FOXA1 is a key determinant of estrogen receptor function and endocrine response.* Nature genetics, 2011. **43**(1): p. 27-33.

77. Kouros-Mehr, H., et al., *GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland.* Cell, 2006. **127**(5): p. 1041-1055.

References

78.   Albergaria, A., et al., *Expression of FOXA1 and GATA-3 in breast cancer: the prognostic significance in hormone receptor-negative tumours.* Breast Cancer Research, 2009. **11**(3): p. R40.

79.   Arruabarrena-Aristorena, A., et al., *FOXA1 Mutations Reveal Distinct Chromatin Profiles and Influence Therapeutic Response in Breast Cancer.* Cancer Cell, 2020. **38**(4): p. 534-550.e9.

80.   Gustin, J.P., et al., *GATA3 frameshift mutation promotes tumor growth in human luminal breast cancer cells and induces transcriptional changes seen in primary GATA3 mutant breast cancers.* Oncotarget, 2017. **8**(61): p. 103415-103427.

81.   Werner, S., L. Keller, and K. Pantel, *Epithelial keratins: Biology and implications as diagnostic markers for liquid biopsies.* Mol Aspects Med, 2020. **72**: p. 100817.

82.   Moll, R., M. Divo, and L. Langbein, *The human keratins: biology and pathology.* Histochem Cell Biol, 2008. **129**(6): p. 705-33.

83.   Trost, A., et al., *K16 is a further new candidate for homotypic intermediate filament protein interactions.* Exp Dermatol, 2010. **19**(8): p. e241-50.

84.   Karantza, V., *Keratins in health and cancer: more than mere epithelial cell markers.* Oncogene, 2011. **30**(2): p. 127-138.

85.   Joosse, S.A., et al., *Changes in keratin expression during metastatic progression of breast cancer: impact on the detection of circulating tumor cells.* Clin Cancer Res, 2012. **18**(4): p. 993-1003.

References

86.   Wawersik, M.J., et al., *Increased levels of keratin 16 alter epithelialization potential of mouse skin keratinocytes in vivo and ex vivo.* Mol Biol Cell, 2001. **12**(11): p. 3439-50.

87.   Rotty, J.D. and P.A. Coulombe, *A wound-induced keratin inhibits Src activity during keratinocyte migration and tissue repair.* J Cell Biol, 2012. **197**(3): p. 381-9.

88.   Wawersik, M. and P.A. Coulombe, *Forced expression of keratin 16 alters the adhesion, differentiation, and migration of mouse skin keratinocytes.* Mol Biol Cell, 2000. **11**(10): p. 3315-27.

89.   Ferlay, J., et al., *Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012.* Eur J Cancer, 2013. **49**(6): p. 1374-403.

90.   Torre, L.A., et al., *Ovarian cancer statistics, 2018.* CA Cancer J Clin, 2018. **68**(4): p. 284-296.

91.   Berek, J.S., et al., *Cancer of the ovary, fallopian tube, and peritoneum.* Int J Gynaecol Obstet, 2018. **143 Suppl 2**: p. 59-78.

92.   Orr, B. and R.P. Edwards, *Diagnosis and Treatment of Ovarian Cancer.* Hematol Oncol Clin North Am, 2018. **32**(6): p. 943-964.

93.   Ledermann, J.A., et al., *Newly diagnosed and relapsed epithelial ovarian carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* Ann Oncol, 2013. **24 Suppl 6**: p. vi24-32.

References

94.   Harter, P., et al., *A Randomized Trial of Lymphadenectomy in Patients with Advanced Ovarian Neoplasms.* N Engl J Med, 2019. **380**(9): p. 822-832.

95.   Pennington, K.P., et al., *Germline and somatic mutations in homologous recombination genes predict platinum response and survival in ovarian, fallopian tube, and peritoneal carcinomas.* Clin Cancer Res, 2014. **20**(3): p. 764-75.

96.   Pujade-Lauraine, E., et al., *Olaparib tablets as maintenance therapy in patients with platinum-sensitive, relapsed ovarian cancer and a BRCA1/2 mutation (SOLO2/ENGOT-Ov21): a double-blind, randomised, placebo-controlled, phase 3 trial.* Lancet Oncol, 2017. **18**(9): p. 1274-1284.

97.   Li, X. and W.D. Heyer, *Homologous recombination in DNA repair and DNA damage tolerance.* Cell Res, 2008. **18**(1): p. 99-113.

98.   Joosse, S.A., *Prediction of "BRCAness" in breast cancer by array comparative genomic hybridization.* . 2012, Leiden University. p. 180.

99.   Cancer Genome Atlas Research, N., *Integrated genomic analyses of ovarian carcinoma.* Nature, 2011. **474**(7353): p. 609-15.

100.   Hauke, J., et al., *Deleterious somatic variants in 473 consecutive individuals with ovarian cancer: results of the observational AGO-TR1 study (NCT02222883).* J Med Genet, 2019. **56**(9): p. 574-580.

101.   Harter, P., et al., *Prevalence of deleterious germline variants in risk genes including BRCA1/2 in consecutive ovarian cancer patients (AGO-TR-1).* PLoS One, 2017. **12**(10): p. e0186043.

References

102. Rosen, E., *BRCA1 in the DNA damage response and at telomeres.* Frontiers in Genetics, 2013. **4**(85).

103. Joosse, S.A., et al., *Genomic signature of BRCA1 deficiency in sporadic basal-like breast tumors.* Genes Chromosomes Cancer, 2011. **50**(2): p. 71-81.

104. Holstege, H., et al., *Cross-species comparison of aCGH data from mouse and human BRCA1- and BRCA2-mutated breast cancers.* BMC Cancer, 2010. **10**: p. 455.

105. Prieske, K., et al., *Loss of BRCA1 promotor hypermethylation in recurrent high-grade ovarian cancer.* Oncotarget, 2017. **8**(47): p. 83063-83074.

106. Joosse, S.A. and K. Pantel, *Genetic traits for hematogeneous tumor cell dissemination in cancer patients.* Cancer Metastasis Rev, 2016. **35**(1): p. 41-8.

107. Guo, M., et al., *Epigenetic heterogeneity in cancer.* Biomark Res, 2019. **7**: p. 23.

108. Baloch, T., et al., *Sequential therapeutic targeting of ovarian Cancer harboring dysfunctional BRCA1.* BMC Cancer, 2019. **19**(1): p. 44.

109. Kondrashova, O., et al., *Methylation of all BRCA1 copies predicts response to the PARP inhibitor rucaparib in ovarian carcinoma.* Nat Commun, 2018. **9**(1): p. 3970.

110. Alix-Panabieres, C. and K. Pantel, *Challenges in circulating tumour cell research.* Nat Rev Cancer, 2014. **14**(9): p. 623-31.

References

111. Menyailo, M., M. Tretyakova, and E. Denisov, *Heterogeneity of Circulating Tumor Cells in Breast Cancer: Identifying Metastatic Seeds.* International Journal of Molecular Sciences, 2020. **21**: p. 1696.

112. Yuanhua, L., et al., *TFAP2A Induced KRT16 as an Oncogene in Lung Adenocarcinoma via EMT.* International journal of biological sciences, 2019. **15**(7): p. 1419-1428.

113. Mastoraki, S., et al., *ESR1 Methylation: A Liquid Biopsy-Based Epigenetic Assay for the Follow-up of Patients with Metastatic Breast Cancer Receiving Endocrine Treatment.* Clin Cancer Res, 2018. **24**(6): p. 1500-1510.

114. Guttery, D.S., et al., *Noninvasive detection of activating estrogen receptor 1 (ESR1) mutations in estrogen receptor-positive metastatic breast cancer.* Clin Chem, 2015. **61**(7): p. 974-82.

115. Harris, F.R., et al., *Quantification of Somatic Chromosomal Rearrangements in Circulating Cell-Free DNA from Ovarian Cancers.* Scientific Reports, 2016. **6**(1): p. 29831.

116. Martignetti, J.A., et al., *Personalized ovarian cancer disease surveillance and detection of candidate therapeutic drug target in circulating tumor DNA.* Neoplasia, 2014. **16**(1): p. 97-103.

117. Ratajska, M., et al., *Detection of BRCA1/2 mutations in circulating tumor DNA from patients with ovarian cancer.* Oncotarget, 2017. **8**(60): p. 101325-101332.

References

118. Kim, Y.M., et al., *Prospective study of the efficacy and utility of TP53 mutations in circulating tumor DNA as a non-invasive biomarker of treatment response monitoring in patients with high-grade serous ovarian carcinoma.* J Gynecol Oncol, 2019. **30**(3): p. e32.

119. Esteller, M., et al., *Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors.* J Natl Cancer Inst, 2000. **92**(7): p. 564-9.

120. Zheng, L.H., et al., *Endocrine resistance in breast cancer.* Climacteric, 2014. **17**(5): p. 522-8.

121. Alix-Panabières, C. and K. Pantel, *Clinical Applications of Circulating Tumor Cells and Circulating Tumor DNA as Liquid Biopsy.* Cancer Discov, 2016. **6**(5): p. 479-91.

122. Saxena, K., A.R. Subbalakshmi, and M.K. Jolly, *Phenotypic heterogeneity in circulating tumor cells and its prognostic value in metastasis and overall survival.* EBioMedicine, 2019. **46**: p. 4-5.

123. Roßwag, S., et al., *Functional Characterization of Circulating Tumor Cells (CTCs) from Metastatic ER+/HER2- Breast Cancer Reveals Dependence on HER2 and FOXM1 for Endocrine Therapy Resistance and Tumor Cell Survival: Implications for Treatment of ER+/HER2- Breast Cancer.* Cancers (Basel), 2021. **13**(8).

124. Gwark, S., et al., *Analysis of the serial circulating tumor cell count during neoadjuvant chemotherapy in breast cancer patients.* Scientific Reports, 2020. **10**(1): p. 17466.

References

125. Eroglu, Z., O. Fielder, and G. Somlo, *Analysis of circulating tumor cells in breast cancer.* J Natl Compr Canc Netw, 2013. **11**(8): p. 977-85.

126. Alix-Panabières, C. and K. Pantel, *Characterization of single circulating tumor cells.* FEBS Lett, 2017. **591**(15): p. 2241-2250.

127. Müller, V., et al., *Prognostic impact of circulating tumor cells assessed with the CellSearch System™ and AdnaTest Breast™ in metastatic breast cancer patients: the DETECT study.* Breast Cancer Research, 2012. **14**(4): p. R118.

128. Yan, W.T., et al., *Circulating tumor cell status monitors the treatment responses in breast cancer patients: a meta-analysis.* Sci Rep, 2017. **7**: p. 43464.

129. Paoletti, C., et al., *Circulating tumor cell number and endocrine therapy index in ER positive metastatic breast cancer patients.* NPJ Breast Cancer, 2021. **7**(1): p. 77.

130. Schwarzenbach, H., et al., *Detection and Monitoring of Cell-Free DNA in Blood of Patients with Colorectal Cancer.* Annals of the New York Academy of Sciences, 2008. **1137**(1): p. 190-196.

131. Forshew, T., et al., *Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA.* Sci Transl Med, 2012. **4**(136): p. 136ra68.

132. Kennedy, S.R., et al., *Detecting ultralow-frequency mutations by Duplex Sequencing.* Nat Protoc, 2014. **9**(11): p. 2586-606.

References

133. Cristofanilli, M. and S. Braun, *Circulating tumor cells revisited.* JAMA, 2010. **303**(11): p. 1092-3.

134. Kinde, I., et al., *Detection and quantification of rare mutations with massively parallel sequencing.* Proceedings of the National Academy of Sciences, 2011. **108**(23): p. 9530.

135. Gale, D., et al., *Development of a highly sensitive liquid biopsy platform to detect clinically-relevant cancer mutations at low allele fractions in cell-free DNA.* PLOS ONE, 2018. **13**(3): p. e0194630.

136. Chabon, J.J., et al., *Corrigendum: Circulating tumour DNA profiling reveals heterogeneity of EGFR inhibitor resistance mechanisms in lung cancer patients.* Nat Commun, 2016. **7**: p. 13513.

137. Higgins, M.J., et al., *Detection of tumor PIK3CA status in metastatic breast cancer using peripheral blood.* Clin Cancer Res, 2012. **18**(12): p. 3462-9.

138. Thierry, A.R., et al., *Clinical validation of the detection of KRAS and BRAF mutations from circulating tumor DNA.* Nat Med, 2014. **20**(4): p. 430-5.

139. Freidin, M.B., et al., *Circulating Tumor DNA Outperforms Circulating Tumor Cells for KRAS Mutation Detection in Thoracic Malignancies.* Clinical Chemistry, 2015. **61**(10): p. 1299-1304.

140. Wang, H., et al., *Allele-Specific, Non-Extendable Primer Blocker PCR (AS-NEPB-PCR) for DNA Mutation Detection in Cancer.* The Journal of Molecular Diagnostics, 2013. **15**(1): p. 62-69.

References

141. GLENN, T.C., *Field guide to next-generation DNA sequencers.* Molecular Ecology Resources, 2011. **11**(5): p. 759-769.

142. Allouchery, V., et al., *Circulating ESR1 mutations at the end of aromatase inhibitor adjuvant treatment and after relapse in breast cancer patients.* Breast Cancer Res, 2018. **20**(1): p. 40.

143. Chandarlapaty, S., et al., *Prevalence of ESR1 Mutations in Cell-Free DNA and Outcomes in Metastatic Breast Cancer: A Secondary Analysis of the BOLERO-2 Clinical Trial.* JAMA Oncol, 2016. **2**(10): p. 1310-1315.

144. Fribbens, C., et al., *Plasma ESR1 Mutations and the Treatment of Estrogen Receptor-Positive Advanced Breast Cancer.* J Clin Oncol, 2016. **34**(25): p. 2961-8.

145. Jeselsohn, R., et al., *Allele-Specific Chromatin Recruitment and Therapeutic Vulnerabilities of ESR1 Activating Mutations.* Cancer Cell, 2018. **33**(2): p. 173-186 e5.

146. Arruabarrena-Aristorena, A., et al., *FOXA1 Mutations Reveal Distinct Chromatin Profiles and Influence Therapeutic Response in Breast Cancer.* Cancer Cell, 2020. **38**(4): p. 534-550 e9.

8. Appendix

8.1. Supporting information for Publication #1.
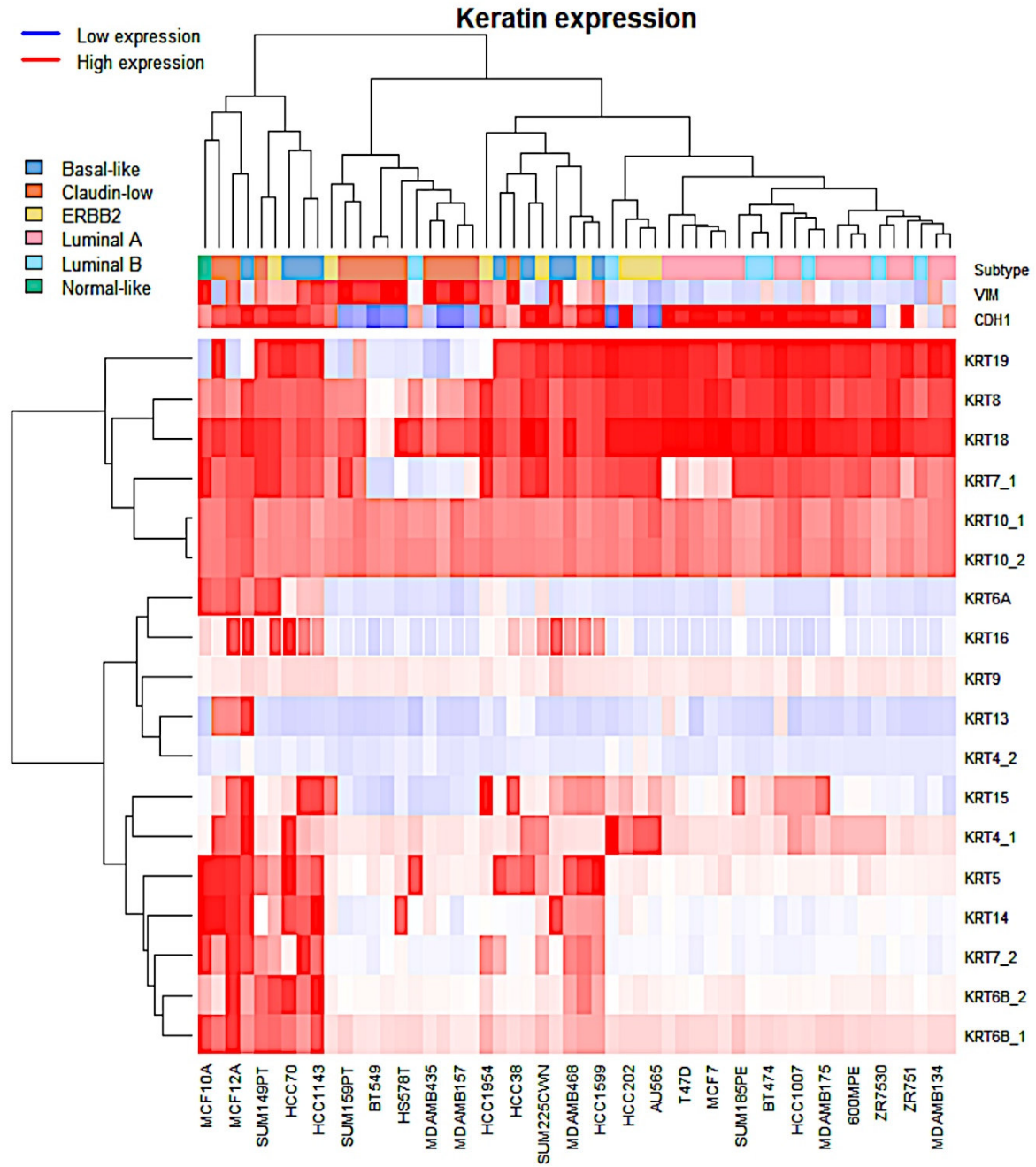
Supplementary Figure 1. *in silico* analysis, *KRT* expression in breast cancer cell lines, the red color represents high relative gene expression, and the blue color represents low relative gene expression
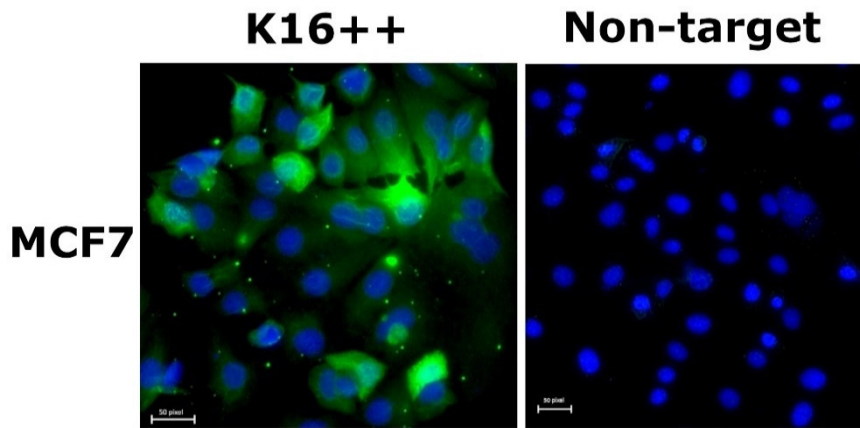
Supplementary Figure 2. Transfection efficiency of _KRT16_, immunocytochemistry staining of transfected cells compared to non-target vector, the cells were stained by DYKDDDDK Tag (AF488; green) and DAPI (blue), scale bar represents 20 µm.

Supplementary Table 1. Transfection efficiency of _KRT16_			
Replicate	Total cell population	Transfected cell population	Percentage of positive cells within a transfected cell population
1#	290	244	84.1
2#	293	260	88.7
3#	1013	672	66.3
4#	1042	770	73.9
5#	478	454	95.0
6#	351	314	89.5
7#	885	640	72.3
Average	621.7	479.1	81.4
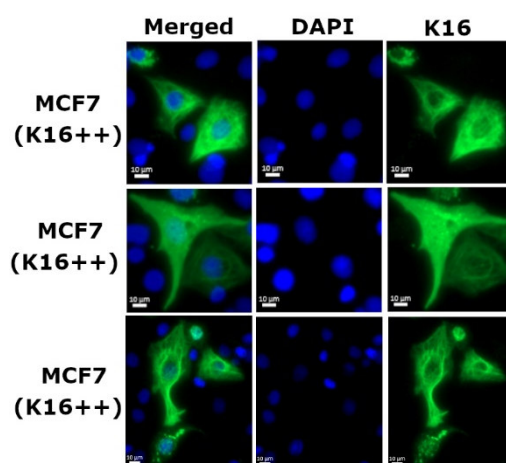
Supplementary Figure 3. Immunofluorescence staining of K16 (AF488; green) and DAPI (blue), scale bar represents 10 µm.
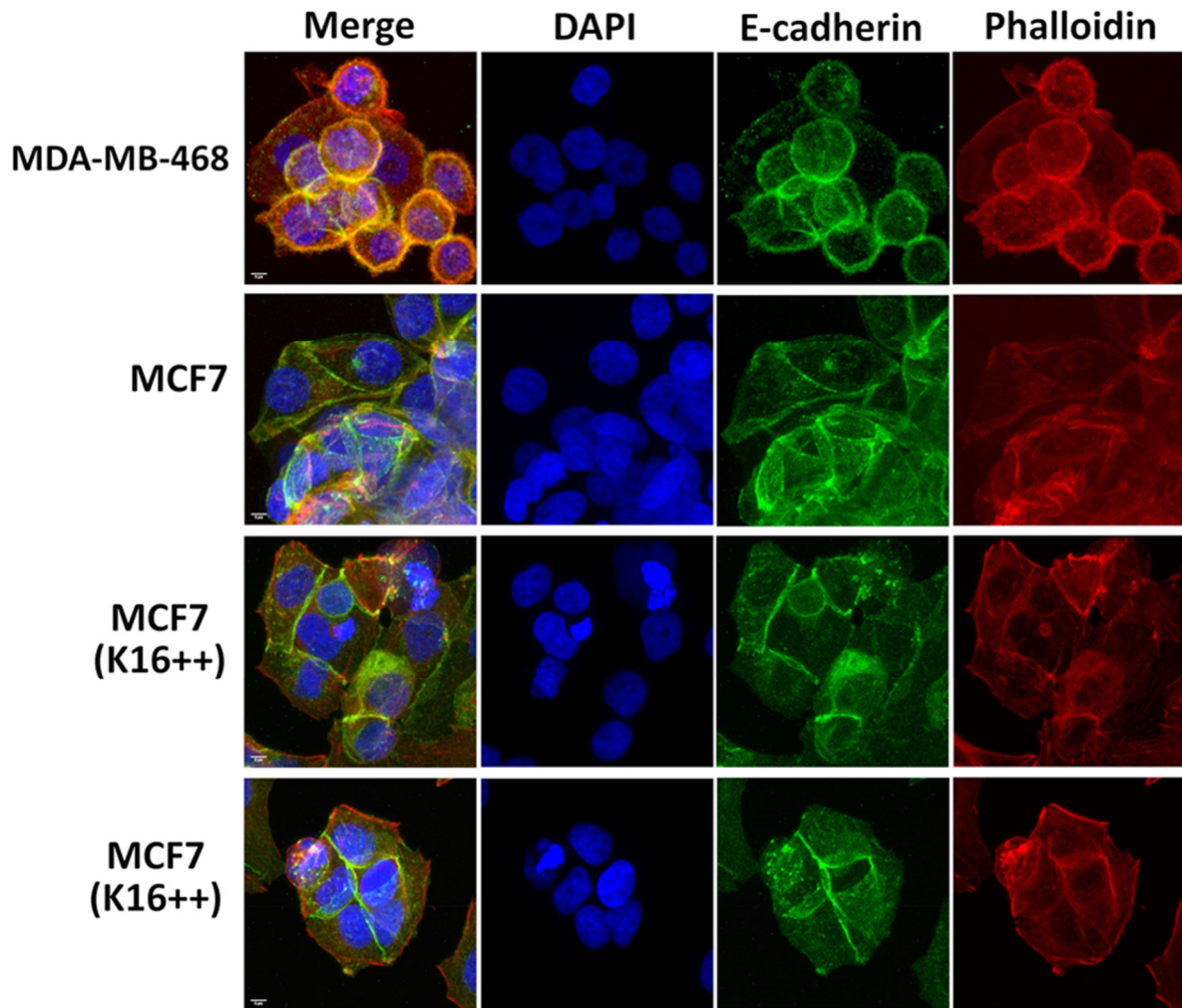
Supplementary Figure 4. Immunocytochemistry staining to visualize actin microfilaments by phalloidin (red), E-cadherin (green), and nucleus by DAPI (blue) in MDA-MB-468, MCF7, and MCF7 cells that induced K16. Scale bar represents 5 μm.

Table S2. Number of detected CTCs with patient's characteristics.

|	Total number of CTCs	K16+/C11- CTCs	K16-/C11+ CTCs	K16+/C11+CTCs	Biological subtype	Age	ER	PR	ERBB2	T-stage	Therapy	Progression	follow-up (months)	Survival analysis stratification
P01	10	90% (9/10)	0	10% (1/10)	TN	61	0	0	0	2	Eribulin	1	12.9	K16+
P02	1	100% (1/1)	0	0	LB	47	1	0	1	2	Emtansin	1	32.0	K16+
P03	5	20% (1/5)	60% (3/5)	20% (1/5)	LB	61	1	0	1	2	Denosumab	0	19.3	K16-
P04	3	100% (3/3)	0	0	LA	58	1	1	0	4	Paclitaxel	0	9.4	K16+
P05	37	62.1% (23/37)	16.2% (6/37)	21.6% (8/37)	LA	52	1	1	0	1	Denosumab, Capecitabine, Bevacizumab	1	32.1	K16-
P06	3	66.7% (2/3)	33.3% (1/3)	0	LA	48	1	1	0	3	Methotrexat	0	104.6	K16-
P07	3	66.7% (2/3)	0	33.3% (1/3)	LA	53	1	1	0		Bevacizumab	1	11.0	K16+
P08	0	0	0	0	LB	53	1	1	1		Doxorubicin	1	124.9	K16-
P09	8	85.7% (6/7)	14.3% (2/7)	0	LB	78	1	0	0			0	391.0	K16-
P10	6	0	66.7% (4/6)	33.3% (2/6)	TN	53	0	0	0	4	Doxorubicin	1	0.9	K16-
P11	25	72% (18/25)	28% (7/25)	0		59					Doxorubicin, Denosumab	1	252.0	K16-
P12	37	62% (23/37)	16.2% (6/37)	21.6% (8/37)	LA	52	1	1	0	1		1		
P13	12	8.3% (1/12)	66.7% (8/12)	25% (3/12)	ERBB2	50	0	0	1	2	Emtansin	1	29.2	K16-
P14	20	20% (4/20)	35% (7/20)	45% (9/20)	LA	57	1	1	0	2	Paclitaxel	1	43.0	K16-
P15	11	27.3% (3/11)	27.3% (3/11)	45.5% (5/11)	LA	63	1	1	0	2	Capecitabine, Bevacizumab	1	105.0	K16-
P16	8	12.5% (1/8)	50% (4/8)	37.5% (3/8)	TN	70	0	0	0	2	Capecitabine, Bevacizumab	0	15.7	K16-
P17	7	100% (7/7)	0	0	LA	63	1	1	0	4	Eribulin	1	24.3	K16+
P18	54	38.9% (21/54)	11.1% (6/54)	50% (27/54)	LA	42	1	1	0	1	Denosumab, Paclitaxel	1	26.0	K16-
P19	42	83.3% (35/42)	4.8% (2/42)	11.9% (5/42)	LA	75	1	1	0	1	Doxorubicin	1	126.0	K16-
P20	142	87.3% (124/142)	7.7% (11/142)	4.9% (7/142)	LA	54	1	1		2	Cyclophosphamid, Methotrexat, 5-Fluoruracil	1	75.0	K16-

TN = Triple-negative subtype, LA = Luminal A-like subtype, LB = Luminal B-like subtype, ER = Estrogen receptor, PR = Progesterone receptor, ERBB2 = Erb-B2 Receptor Tyrosine Kinase 2.

8.2. Supporting information for Publication #2.
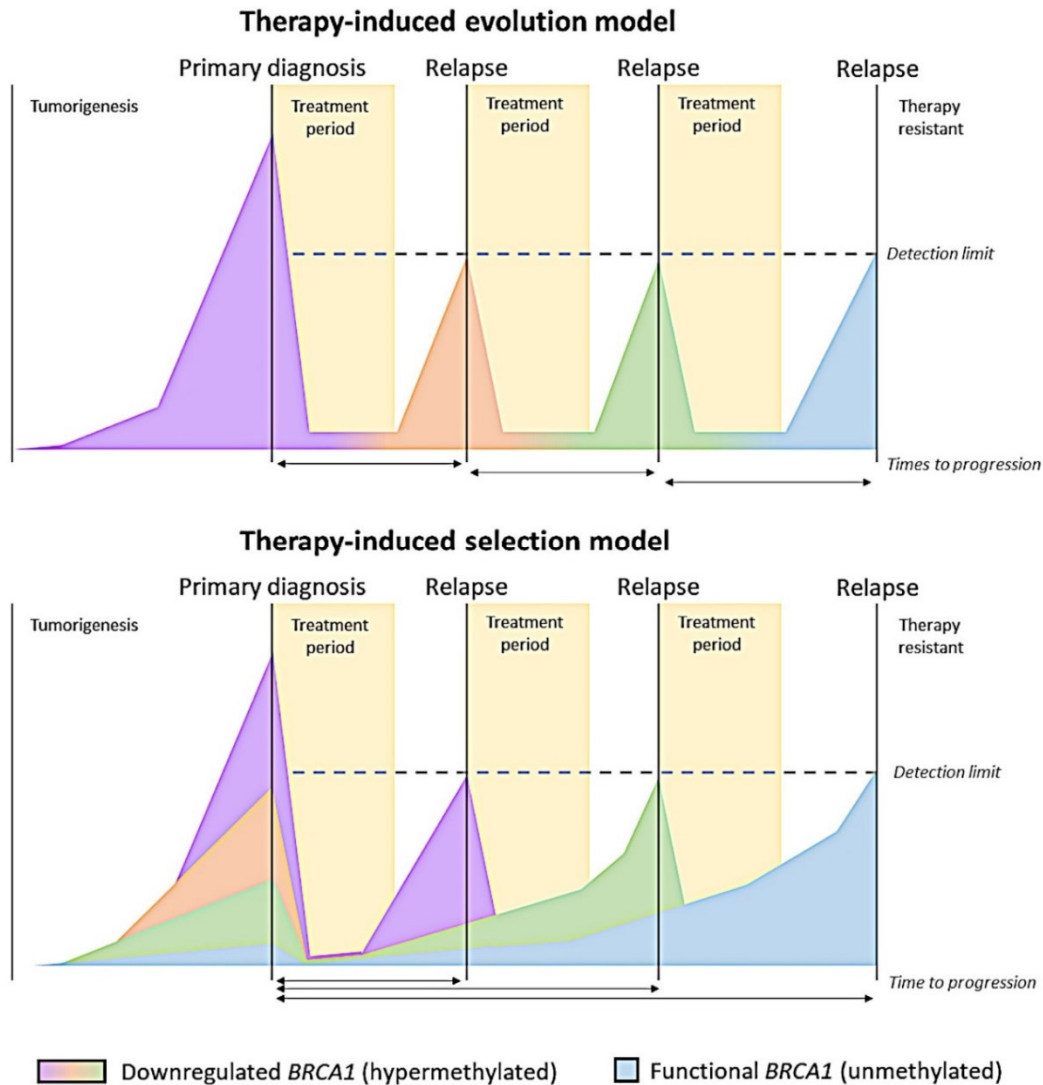
Supplementary Figure 1. Models of progression. Two models explaining the conversion of BRCA1 promoter hypermethylation. BRCA1 promoter hypermethylation is an early event in tumorigenesis. After detection of the primary tumor and multiple rounds of therapy after relapse, the tumor reactivates BRCA1 by evolving and reversing its methylation status and thereby developing therapy resistance (upper panel). Alternatively, multiple subclones may have already developed during tumorigenesis and through multiple rounds of therapy, the most therapy resistant clone eventually survives and thrives (lower panel). The arrows indicate the time from the development to detection of the tumor to be treated, illustrating the need for different statistical models for analysis.

Table S1. patient's characteristics.

Patinent ID – blood sample	cfDNA, conc.ng/ml	Methylation status	BRCA1/2 germline mutations testing	Blood obtained before/during/after therapy		Histology	Pathology
P02-1	184	Methylated	NA	During	Carboplatin,Caelyx	HGSOC	FIGO IV
P02-2	81	Unmethylated	NA	After	Carboplatin,Caelyx	HGSOC	FIGO IV
P03-1	134	Unmethylated	NA	Before	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P03-2	142	Unmethylated	NA	After	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P04-1	144	Methylated	Negative	During	Carboplatin,Caelyx	HGSOC	FIGO IV
P04-2	72	Methylated	Negative	During	Carboplatin,Caelyx	HGSOC	FIGO IV
P05	384	Methylated	Negative	During	Paclitaxel	HGSOC	FIGO IIIc
P06-1	398	Unmethylated	NA	During	Carboplatin,Caelyx	HGSOC	FIGO IV
P06-2	240	Unmethylated	NA	During	Carboplatin,Caelyx	HGSOC	FIGO IV
P07-1	370	Methylated	NA	Before	Carboplatin,Taxol	HGSOC	
P07-2	434	Methylated	NA	During	Carboplatin,Taxol	HGSOC	
P08	56	Methylated	NA	During	Carboplatin	HGSOC	FIGO Ic
P09	232	Methylated	NA	During	Gemcitabine	HGSOC	FIGO IV
P10-1	118	Methylated	Positive	Before	Carboplatin,Caelyx	HGSOC	
P10-2	148	Methylated	Positive	During	Carboplatin,Caelyx	HGSOC	
P10-3	452	Unmethylated	Positive	During	Avastin	HGSOC	
P11-1	108	Unmethylated	NA	Before	Carboplatin,Caelyx	HGSOC	
P11-2	122	Unmethylated	NA	During	Carboplatin,Caelyx	HGSOC	
P12	666	Methylated	NA	During	Caelyx	HGSOC	FIGO IV
P13-1	298	Unmethylated	NA	Before	Carboplatin	HGSOC	FIGO IIIc
P13-2	250	Unmethylated	NA	During	Carboplatin	HGSOC	FIGO IIIc
P14-1	282	Unmethylated	Negative	Before	Carboplatin,Paclitaxel	HGSOC	FIGO IV
P14-2	128	Unmethylated	Negative	After	Carboplatin,Paclitaxel	HGSOC	FIGO IV
P14-3	2080	Methylated	Negative	During	Carboplatin	HGSOC	FIGO IV
P15-1	222	Unmethylated	Positive	Before	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P15-2	250	Methylated	Positive	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P15-3	322	Unmethylated	Positive	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P15-4	1260	Methylated	Positive	During	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P15-5	42.4	Unmethylated	Positive	After	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P15-6	1720	Methylated	Positive	During	Olaparib	HGSOC	FIGO IIIc
P16-1	394	Methylated	NA	During	Cisplatin	HGSOC	FIGO IIIc
P16-2	174	Methylated	NA	During	Cisplatin	HGSOC	FIGO IIIc
P16-3	106	Unmethylated	NA	During	Caelyx	HGSOC	FIGO IIIc
P17-1	188	Methylated	Negative	Before	Carboplatin	HGSOC	FIGO IV
P17-2	47.3	Methylated	Negative	During	Carboplatin	HGSOC	FIGO IV
P17-3	126	Unmethylated	Negative	During	Carboplatin	HGSOC	FIGO IV
P17-4	68	Unmethylated	Negative	During	Carboplatin	HGSOC	FIGO IV
P18	494	Unmethylated	NA	During	Carboplatin	HGSOC	FIGO IIIc
P19	284	Methylated	Positive	During	Carboplatin	HGSOC	FIGO IIIc
P20-1	252	Methylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P20-2	260	Methylated	NA	After	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P21	642	Methylated	NA	Before	Carboplatin	HGSOC	
P22-1	88	Methylated	NA	Before	Carboplatin,Paclitaxel	HGSOC	FIGO IV
P22-2	132	Methylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IV
P23-1	188	Methylated	Negative	Before	Gemcitabine	HGSOC	FIGO IIIc
P23-2	460	Unmethylated	Negative	During	Gemcitabine	HGSOC	FIGO IIIc
P23-3	220	Methylated	Negative	During	Gemcitabine	HGSOC	FIGO IIIc
P24	228	Methylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P25	106	Unmethylated	Negative	During	Carboplatin,Paxlitaxel	HGSOC	FIGO Ia
P26-1	124	Methylated	Negative	Before	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P26-2	432	Unmethylated	Negative	During	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P26-3	364	Unmethylated	Negative	Before	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P27	150	Unmethylated	NA	After	Treaosulfan	HGSOC	FIGO IIIc
P28-1	206	Methylated	Negative	Before	Carboplatin	HGSOC	FIGO Ia
P28-2	394	Unmethylated	Negative	After	Carboplatin	HGSOC	FIGO Ia
P29	142	Unmethylated	Positive	During	Cisplatin	HGSOC	FIGO IIIc
P30	182	Methylated	Negative	During	Carboplatin,Caelyx	HGSOC	FIGO IIIc

155

P31-1	1210	Methylated	NA	Before	Carboplatin mono	HGSOC	FIGO IV
P31-2	142	Methylated	NA	During	Carboplatin mono	HGSOC	FIGO IV
P32-1	164	Unmethylated	Positive	After	Olaparib,Carboplatin mono	HGSOC	FIGO IIIc
P32-2	856	Methylated	Positive	During	Paclitaxel	HGSOC	FIGO IIIc
P33	106	Unmethylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P34-1	362	Unmethylated	Negative	Before	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P34-2	430	Unmethylated	Negative	Before	Avastin,Olaparib/Placebo	HGSOC	FIGO IIIc
P35	442	Methylated	Negative	During	Carboplatin,Gemcitabine	HGSOC	FIGO IIIc
P36	1330	Methylated	NA	Before	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P37-1	192	Methylated	Negative	Before	Cisplatin	HGSOC	FIGO IIIc
P37-2	120	Unmethylated	Negative	Before	Carboplatin mono	HGSOC	FIGO IIIc
P37-3	280	Unmethylated	Negative	After	Niraparib	HGSOC	FIGO IIIc
P38-1	218	Methylated	NA	During	Carboplatin,Paclitaxel,Avastin	HGSOC	FIGO IIIc
P38-2	534	Methylated	NA	Before	Carboplatin mono	HGSOC	FIGO IIIc
P38-3	100	Unmethylated	NA	During	Carboplatin mono	HGSOC	FIGO IIIc
P38-4	1090	Unmethylated	NA	Before	Gemcitabine	HGSOC	FIGO IIIc
P39-1	122	Unmethylated	Positive	During	Carboplatin mono	HGSOC	FIGO IV
P39-2	758	Unmethylated	Positive	During	Carboplatin mono	HGSOC	FIGO IV
P40	6750	Methylated	NA	During	Carboplatin	HGSOC	
P41-1	100	Unmethylated	Positive	During	Carboplatin	HGSOC	FIGO IIIc
P41-2	288	Unmethylated	Positive	During	Carboplatin	HGSOC	FIGO IIIc
P42	390	Methylated	Done	During	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P43-1	1240	Methylated	Done	During	Paclitaxel	HGSOC	FIGO IIIc
P43-2	518	Methylated	Done	During	Paclitaxel	HGSOC	FIGO IIIc
P44	456	Unmethylated	NA	After	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P45-1	112	Unmethylated	NA	Before	Carboplatin,Paclitaxel,Avastin	HGSOC	FIGO IIIc
P45-2	446	Unmethylated	NA	During	Carboplatin,Paclitaxel,Avastin	HGSOC	FIGO IIIc
P46	362	Unmethylated	NA	Before	Carboplatin,Taxol	HGSOC	FIGO IVa
P47-1	1050	Methylated	Done	Before	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P47-2	1420	Methylated	Done	After	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P48	456	Unmethylated	Done	During	Avastin	HGSOC	FIGO IIIc
P49	496	Methylated	NA	After	Carboplatin,Paclitaxel	HGSOC	FIGO IV
P50	618	Unmethylated	Done	During	Caelyx	HGSOC	FIGO IIIc
P51-1	1100	Methylated	NA	Before	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P51-2	610	Methylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P52-1	120	Unmethylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P52-2	786	Unmethylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P52-3	28.4	Unmethylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P52-4	1280	Unmethylated	NA	During	Avastin	HGSOC	FIGO IIIc
P53-2	124	Methylated	Done	During	Carboplatin,Paclitaxel	HGSOC	FIGO IV
P53-3	926	Methylated	Done	After	Carboplatin,Paclitaxel	HGSOC	FIGO IV
P54-2	348	Methylated	NA	During	Carboplatin mono	HGSOC	FIGO IIb
P54-3	960	Unmethylated	NA	After	Caelyx	HGSOC	FIGO IIb
P55-1	106	Methylated	Done	Before	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P55-2	596	Methylated	Done	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P55-3	528	Methylated	Done	During	Avastin	HGSOC	FIGO IIIc
P55-4	306	Unmethylated	Done	During	Avastin	HGSOC	FIGO IIIc
P56-1	100	Unmethylated	NA	Before	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P56-2	558	Unmethylated	NA	During	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P56-3	954	Unmethylated	NA	During	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P57-1	1580	Methylated	Positive	During	Cisplatin	HGSOC	FIGO IIIc
P57-2	2880	Methylated	Positive	After	Cisplatin	HGSOC	FIGO IIIc
P58-1	100	Unmethylated	NA	Before	Carboplatin mono	HGSOC	FIGO IV
P58-2	932	Unmethylated	NA	During	Carboplatin mono	HGSOC	FIGO IV
P58-3	274	Unmethylated	NA	During	Carboplatin mono	HGSOC	FIGO IV
P59	250	Methylated	Negative	During	Carboplatin,Doxorubicin	HGSOC	FIGO IIIc

156

P60-1	408	Unmethylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P60-2	1500	Unmethylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P60-3	1070	Unmethylated	NA	After	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P61	944	Unmethylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIb
P62-1	462	Methylated	Done	Before	Carboplatin,Paclitaxel	HGSOC	FIGO IIIb
P62-2	1320	Methylated	Done	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIb
P62-3	646	Unmethylated	Done	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIb
P63-1	1570	Unmethylated	Done	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P63-2	344	Unmethylated	Done	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P63-3	802	Unmethylated	Done	During	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P63-4	982	Unmethylated	Done	After	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc
P64-2	234	Unmethylated	NA	During	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P64-3	556	Unmethylated	NA	During	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P64-4	32.6	Unmethylated	NA	After	Carboplatin,Caelyx	HGSOC	FIGO IIIc
P65-1	210	Unmethylated	Positive	During	Carboplatin,Gemcitabine	HGSOC	FIGO IIIc
P65-2	1250	Unmethylated	Positive	During	Carboplatin,Gemcitabine	HGSOC	FIGO IIIc
P66	100	Methylated	NA	During	Caelyx	HGSOC	FIGO IVa
P67	4880	Unmethylated	Positive	During	Carboplatin,Caelyx	HGSOC	FIGO IV
P68-1	282	Methylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IVa
P68-2	802	Unmethylated	NA	During	Carboplatin,Paclitaxel	HGSOC	FIGO IVa
P69	1130	Unmethylated	NA	Before	Carboplatin,Paclitaxel	HGSOC	FIGO IIIc

High-Grade Serous Ovarian Carcinoma (HGSOC)

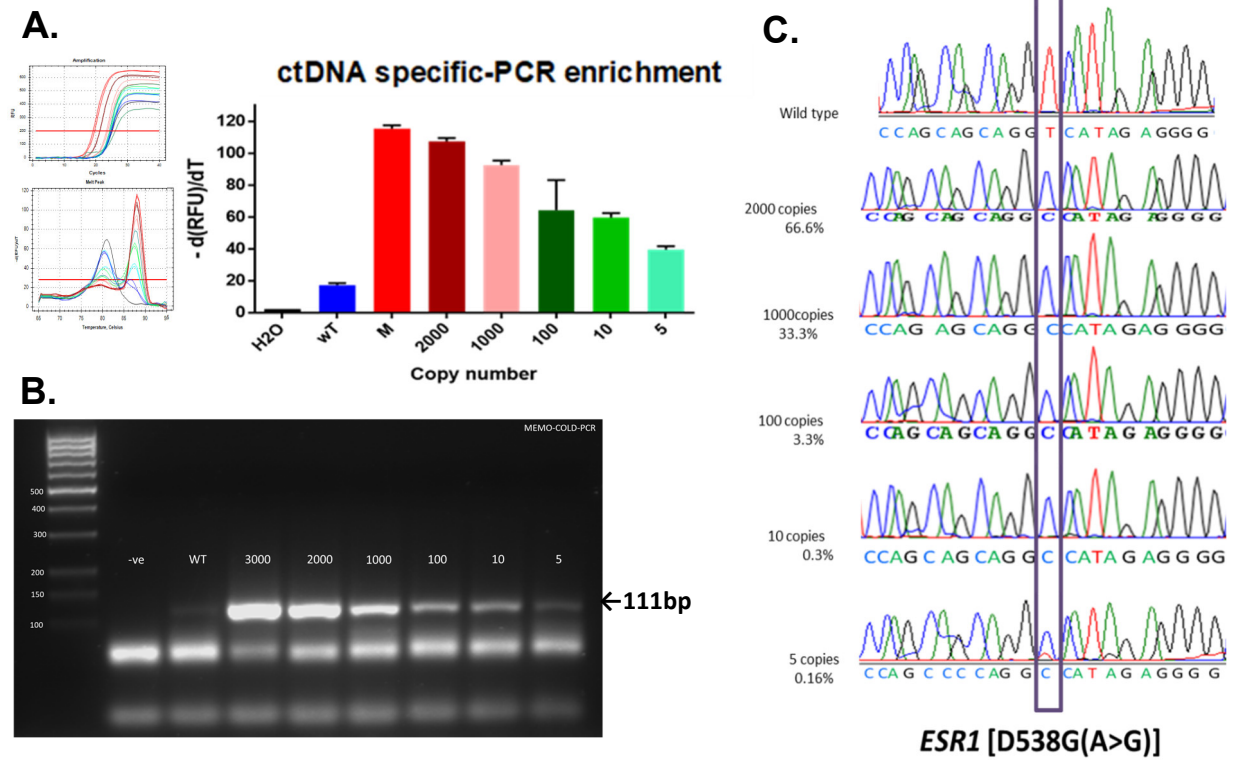## 8.3. Supporting information for ctDNA specific-PCR enrichment.

Figure S1. Evaluate the sensitivity and specificity of the ctDNA specific-PCR enrichment amplification protocol. (**A**) Melt analysis of ctDNA specific-PCR enrichment showing identification of mutation abundance down to 5copies. (**B**) Electrophoresis of PCR products on 1% agarose gel showing the difference in amplification between wild type (WT) and serial dilution of mutant. (**C**) Sanger sequencing analysis showing a high sensitivity and specificity of ctDNA specific-PCR enrichment to detect and enrichment the minimal abundance of mutant alleles at very low copy number down to 0.16% of mutant allele.
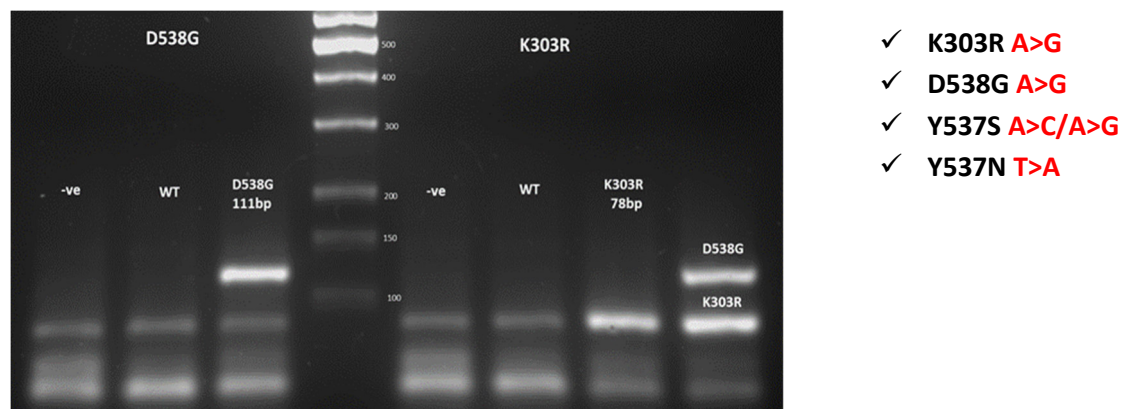
Figure S2. ctDNA specific-PCR multiplex enrichment. Electrophoresis of PCR products on 3% agarose gel showing the efficiency of ctDNA specific-PCR multiplex enrichment to multiplex two groups.

8.4. Supporting information for Publication #4.

Supplementary Table 1. Progression-free survival and overall survival analysis for candidate mutations independently.

| Progression-free survival						
Mutant gene	**Alteration**	**No.**	**Events**	**Median/month**	**HR (95%CI)**	**P Value**
Wild-type	WT	45	5	--	0.071-0.152	--
ESR1	pD538G	70	23	--	1.572-7.053	0.0055
	pY537S	9	3	--	0.4901-24.16	0.0458
	pY537N	3	2	5	0.2288-355.9	0.0008
	pE380Q	4	2	27	0.2765-143.2	0.0064
PIK3CA	pH1047R	16	7	--	1.273-19.68	0.0017
	pE542K	16	4	--	0.5798-13.37	0.1054
	pE545K	10	3	--	0.4869-23.34	0.0753
FOXA1	pE24K	13	7	8	1.521-31.16	<0.0001
	pI176M	11	3	--	0.4701-19.76	0.1035
	pI176V	15	3	17	0.3957-11.05	0.2968
GATA3	pD336fs17	52	24	--	2.636-11.34	<0.0001
	pS93F	5	2	--	0.2934-90.98	0.0256
Overall survival						
Mutant gene	**Alteration**	**No.**	**Events**	**Median/month**	**HR (95%CI)**	**P Value**
Wild-type	WT	45	12	--	0.224-0.285	--
ESR1	pD538G	70	37	45	1.230-3.837	0.0154
	pY537S	9	8	15	1.521-38.10	<0.0001
	pY537N	3	3	18	0.4544-77.54	0.0016
	pE380Q	4	2	22	0.2875-30.45	0.1344
PIK3CA	pH1047R	16	11	19	1.554-14.18	<0.0001
	pE542K	16	8	24	0.9234-8.562	0.0172
	pE545K	10	7	18	1.229-28.71	<0.0001
FOXA1	pE24K	15	7	8	1.027-14.96	0.0014
	pI176M	11	3	--	0.3088-4.382	0.8129
	pI176V	15	10	17	1.443-14.68	<0.0001
GATA3	pD336fs17	52	36	14	2.924-9.486	<0.0001
	pS93F	5	2	19	0.2663-78.32	0.0256
WT, Wild-type; NO, number; HR, hazard ratio						

Supplementary Table 2. Progression-free survival and overall survival analysis for single and polyclonal mutant genes.

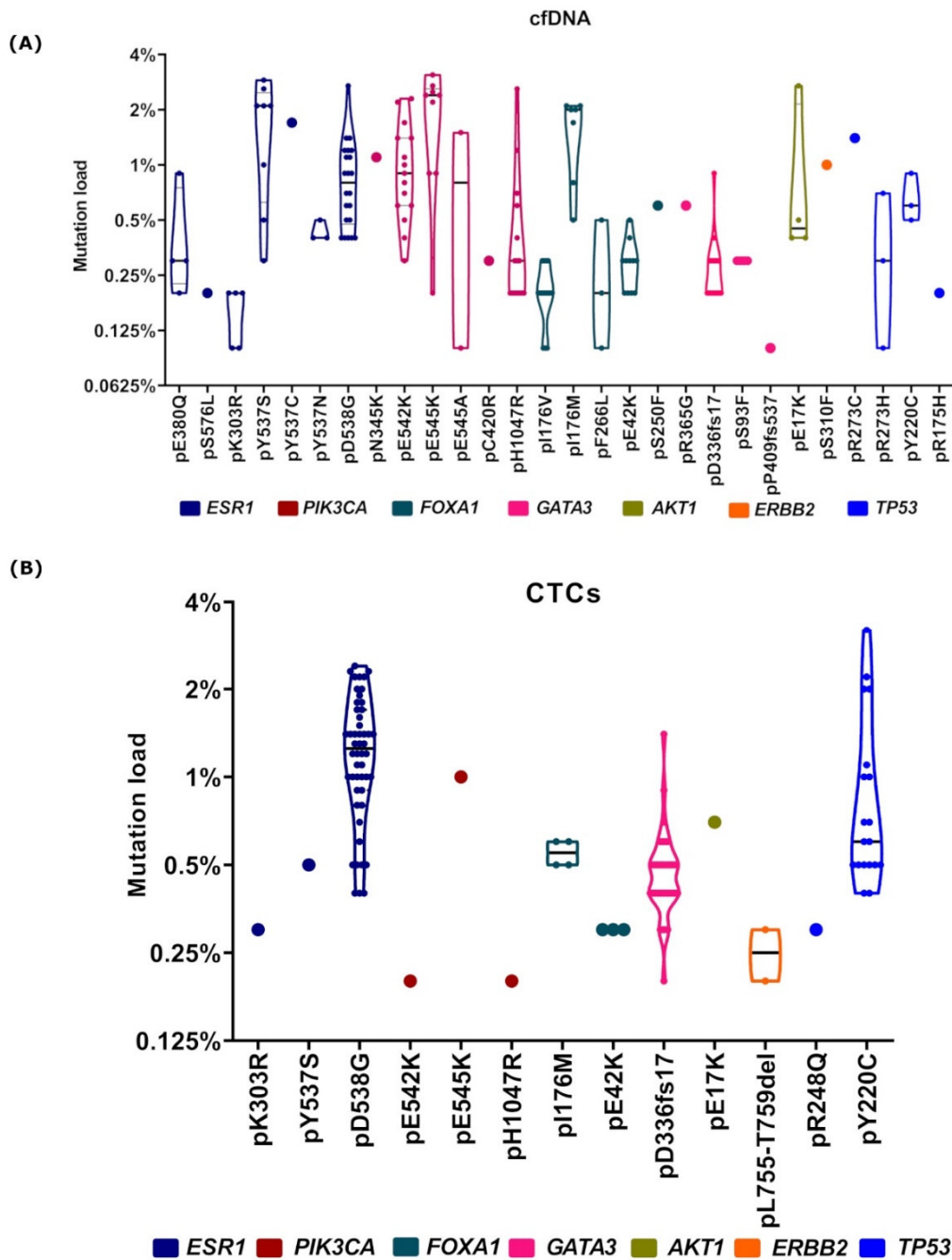| Progression-free survival					
Group	No.	Events	Median/month	HR (95%CI)	P Value
WT	45	5	--	0.227-3.988	--
ESR1	28	3	--	0.229-3.96	0.948
PIK3CA	10	3	--	0.503-28.15	0.0503
FOXA1	15	6	--	1.0744-19.99	0.0051
GATA3	11	5	--	0.9828-30.15	0.0048
ESR1/PIK3CA	17	7	--	1.205-17.45	0.0036
ESR1/FOXA1	6	1	--	0.1252-17.95	0.7075
ESR1/GATA3	19	8	--	1.409-17.84	0.0016
PIK3CA/FOXA1	5	2	--	0.2927-93.09	0.0254
FOXA1/GATA3	7	4	2.9	0.8277-103.3	<0.0001
ESR1/PIK3CA/GATA3	4	2	12.5	0.2850-116.2	0.017
ESR1/FOXA1/GATA3	9	8	7	1.088-49.11	0.0022
Overall survival					
Group	No.	Events	Median/month	HR (95%CI)	P Value
WT	45	12	58.9	0.224-0.285	--
ESR1	28	8	--	0.4180-2.513	0.819
PIK3CA	10	7	19.2	1.218-16.65	0.0002
FOXA1	15	8	18	1.016-10.51	0.0034
GATA3	11	7	14.9	1.627-32.30	<0.0001
ESR1/PIK3CA	17	11	22	1.238-9.054	0.0019
ESR1/FOXA1	6	2	--	0.1956-3.024	0.7256
ESR1/GATA3	19	11	22.7	2.271-18.48	0.0005
PIK3CA/FOXA1	5	2	--	0.2878-27.64	0.1505
FOXA1/GATA3	7	5	9	0.9870-71.22	<0.0001
ESR1/PIK3CA/GATA3	4	4	17	0.6394-131.2	<0.0001
ESR1/FOXA1/GATA3	5	3	19	0.4641-54.85	0.0051
WT, Wild-type; NO, number; HR, hazard ratio					
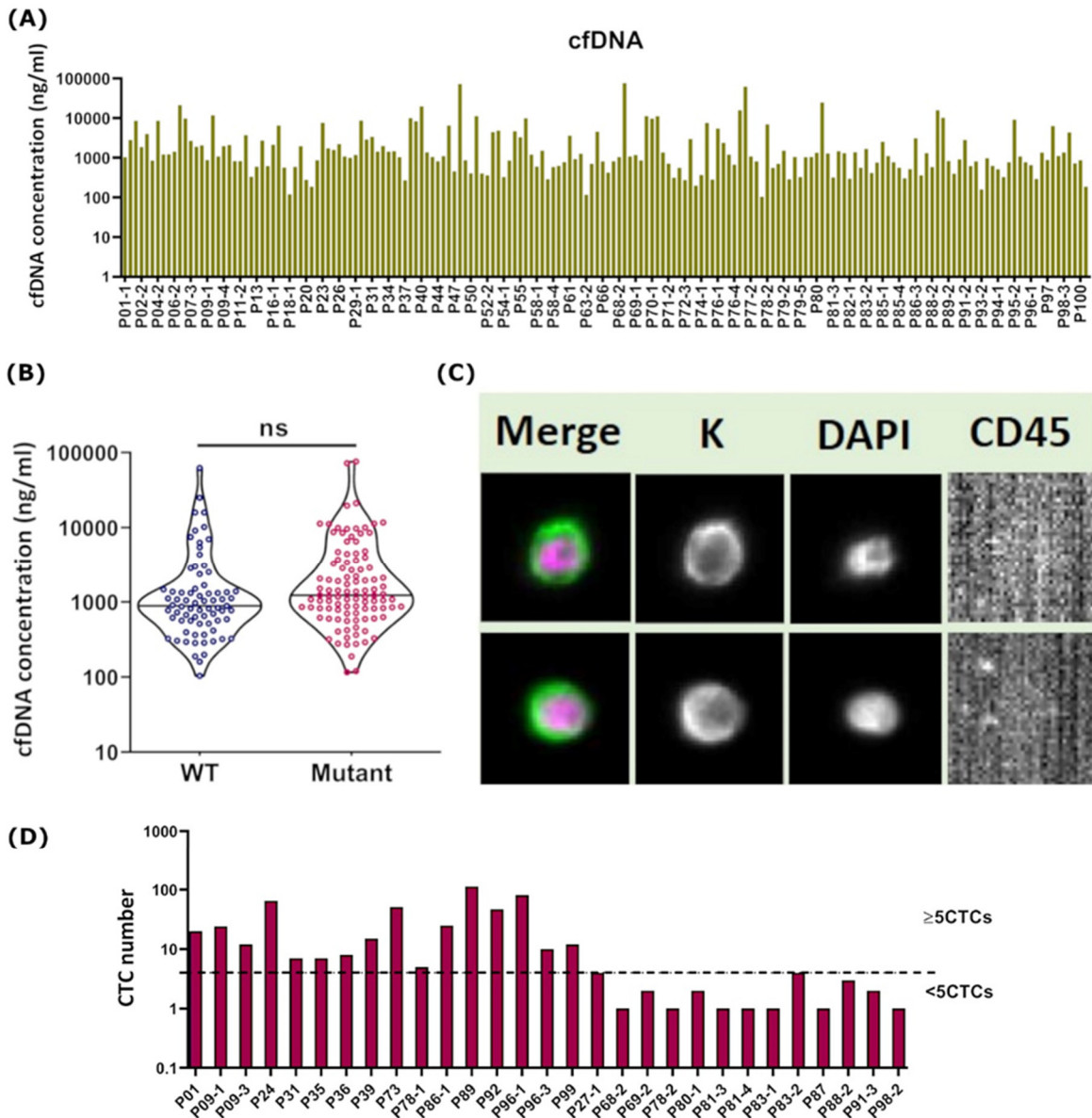
Supplementary Figure 1. Isolated cfDNA and CTCs. (A) Distribution of cfDNA in metastatic breast cancer patients (n= 176). (B) Violin plot depicting the distribution of the cfDNA concentrations of metastatic breast cancer patients with mutant ctDNA; (n= 103) and wildtype cfDNA (n= 73). (C) Image gallery of CTCs detected from one breast cancer patient by the Cell Search System; K (Keratin); DAPI; CD45. (D) Distribution of detected. CTCs in metastatic breast cancer patients (n= 29); dashed line represents the CTCs cut-off ≥ 5.
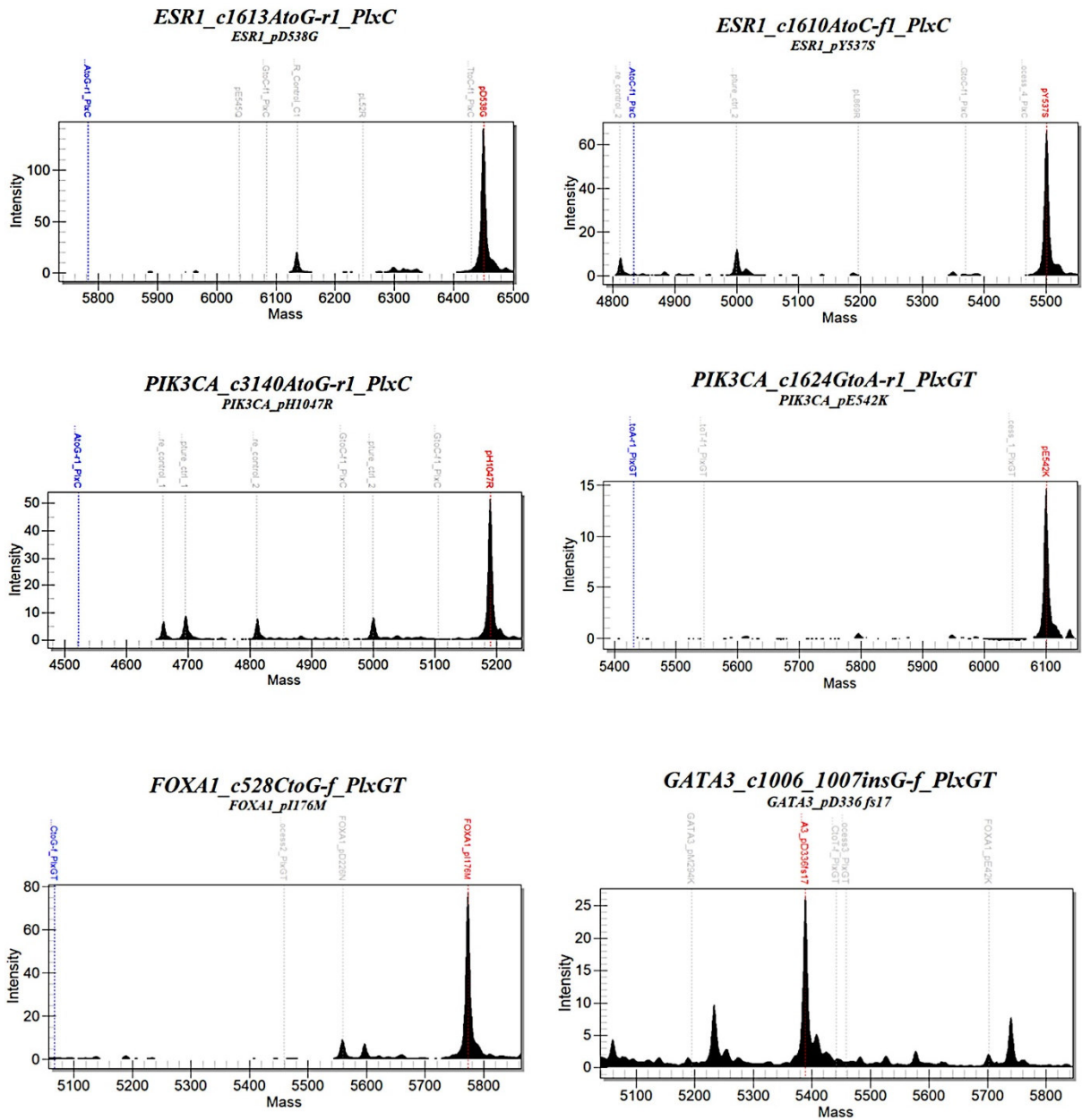
Supplementary Figure 2. MassARRAY plots of detected ESR1 (pD538G and pY537S), PIK3CA (pH1047R and pE542K), FOXA1 (pI176M) and GATA3 (pD336fs17) mutations in metastatic breast cancer patients.

Supplementary Figure 3. Violin plot displaying the mutation load of detected mutations in (A) cfDNA and (B) CTCs; the middle line represents the median.
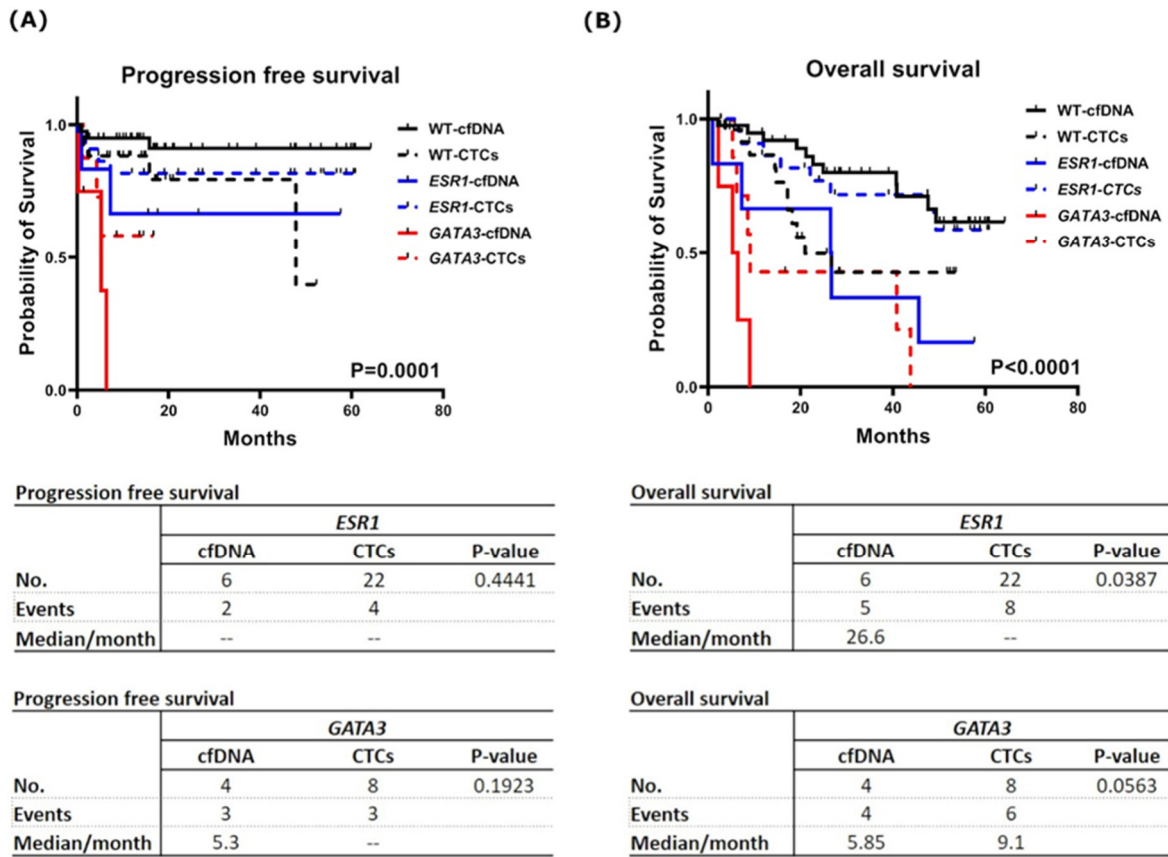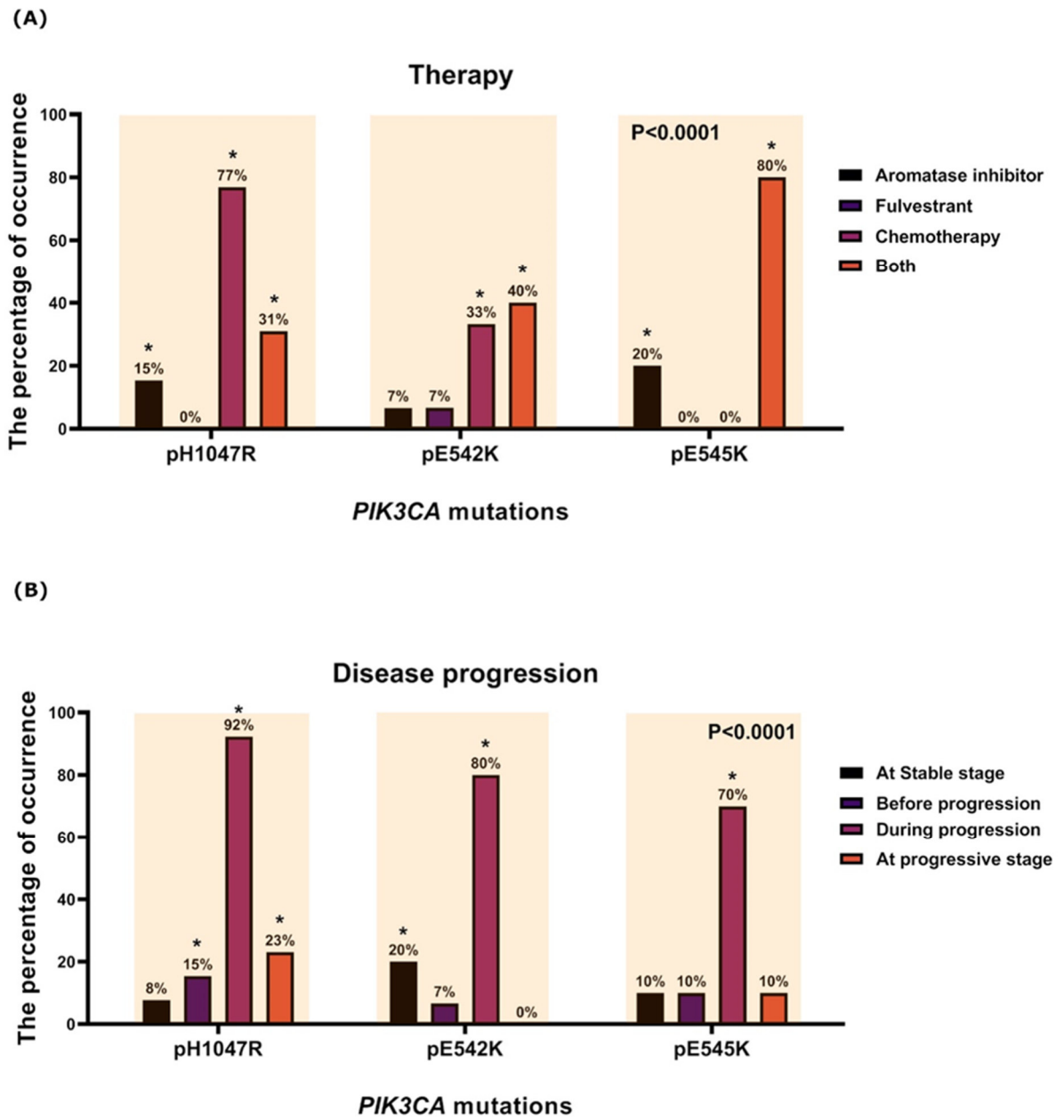
Supplementary Figure 4. Survival analyses. Kaplan–Meier curves for (A) progression-free survival (PFS) and (B) overall survival (OS) comparison the clinical value of cfDNA and CTCs from patients with ESR1and GATA3 mutations. P-values were calculated using the log rank test; ⊥ censored
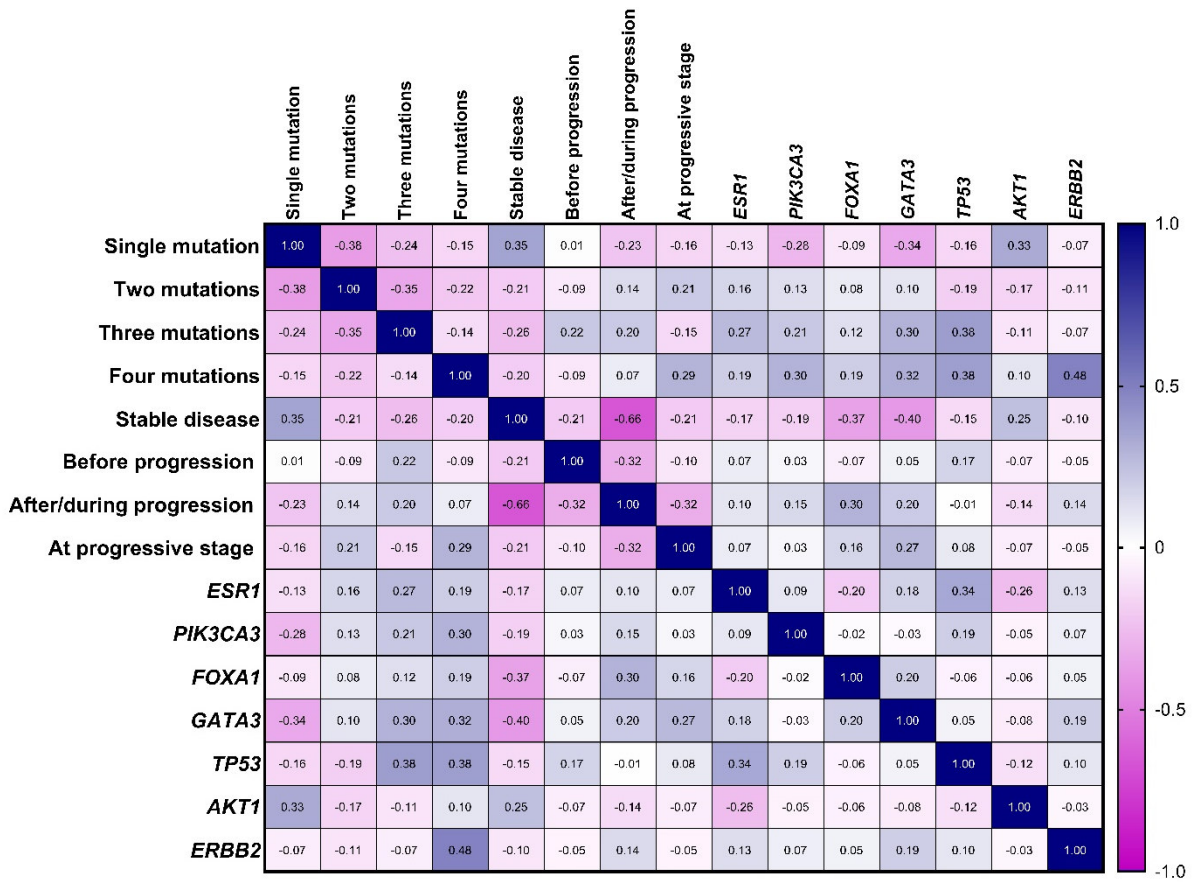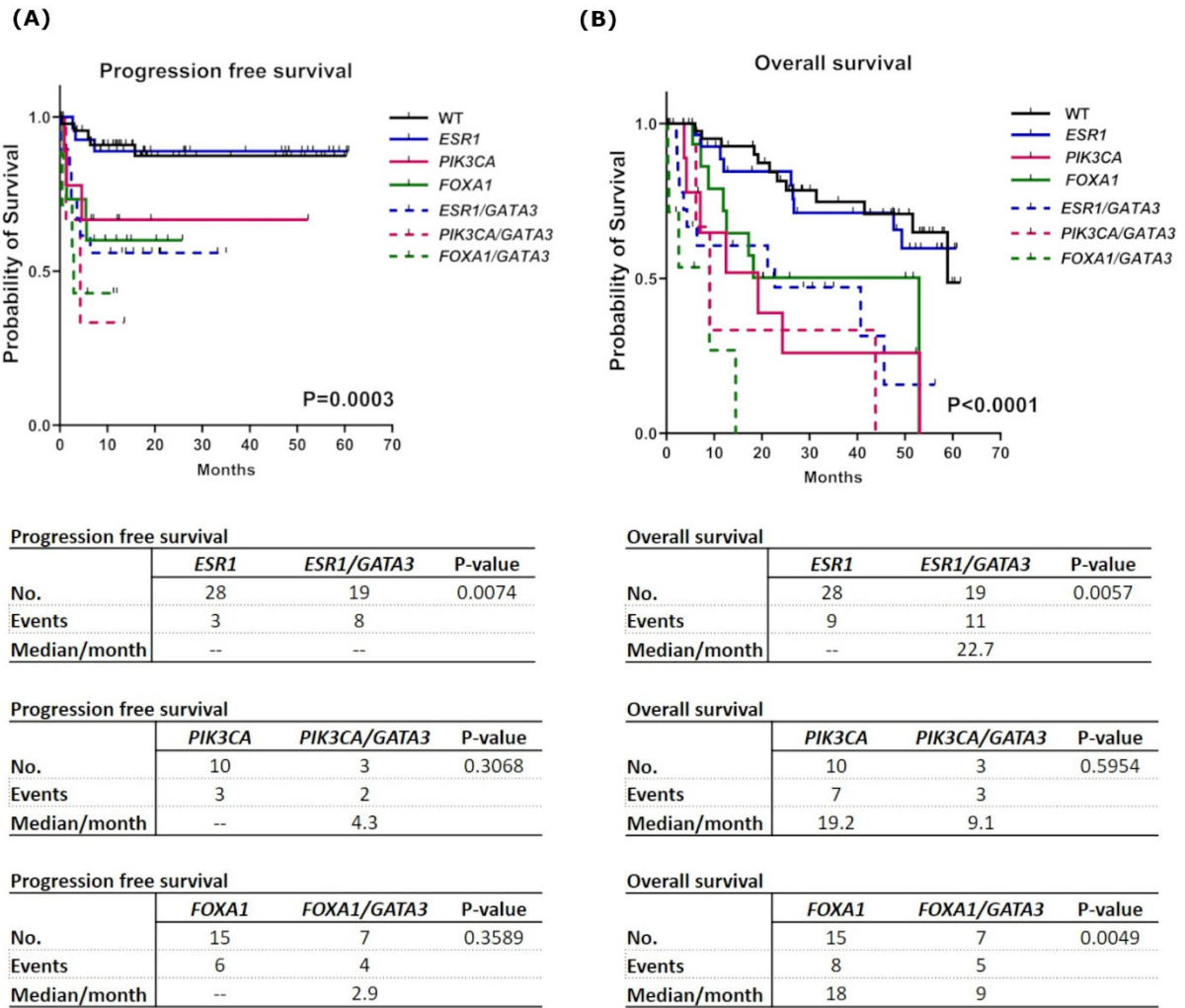
Supplementary Figure 5. Correlation analysis demonstrates the relationship between the number of detected mutations with the stages of disease progression and mutant genes.

(A)

(B)
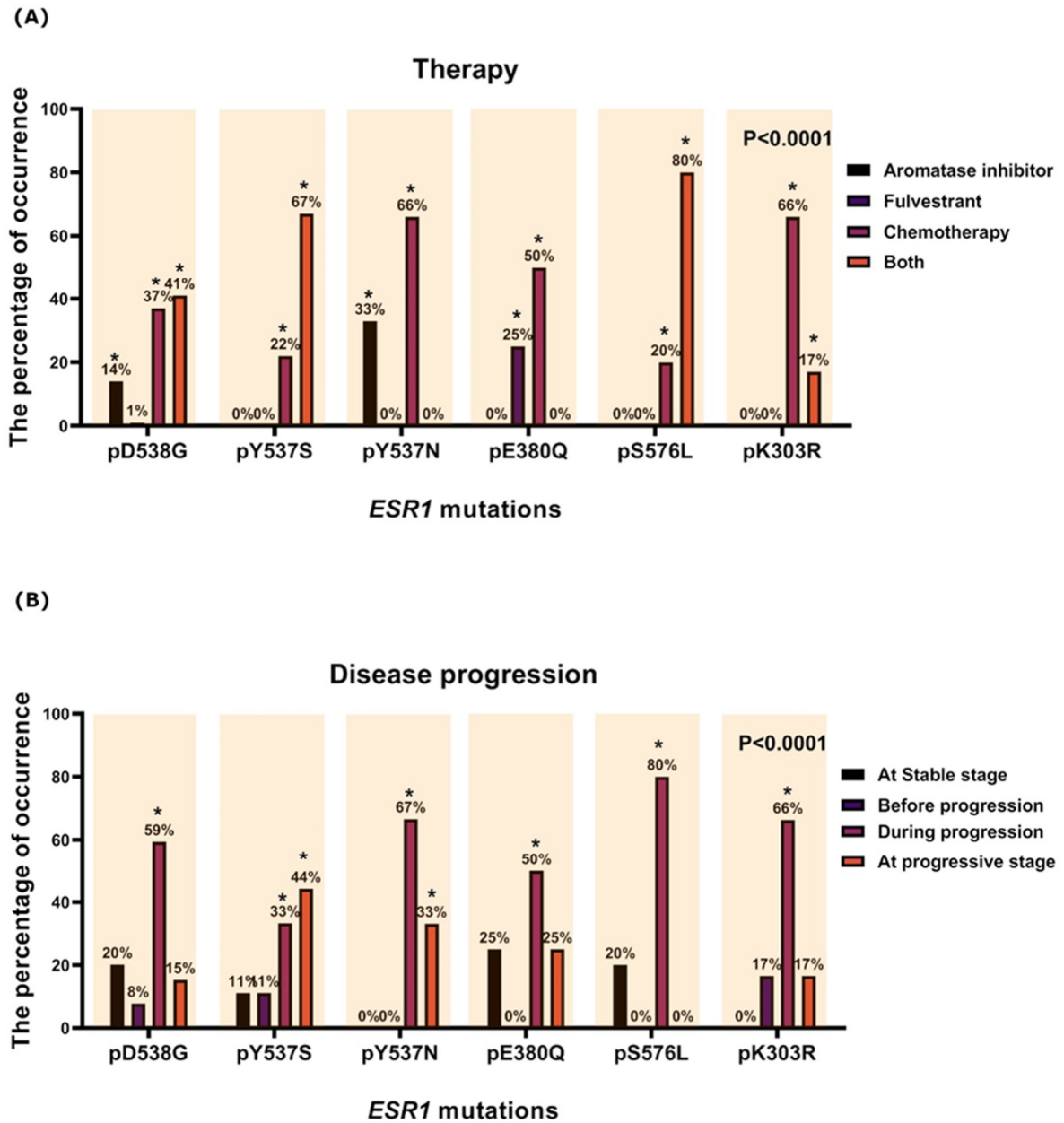
Supplementary Figure 6. Occurrence of ESR1 hot spot mutations in patients (A) received Aromatase inhibitor, Fulvestrant, Chemotherapy, or both endocrine therapy and chemotherapy; (B) during the phases of disease progression

Supplementary Figure 7. Occurrence of PIK3CA mutations in patients (A) received Aromatase inhibitor, Fulvestrant, Chemotherapy, or both endocrine therapy and chemotherapy; (B) during the phases of disease progression.

(A)

Progression free survival

(B)

Overall survival

Progression free survival

	ESR1	ESR1/GATA3	P-value
No.	28	19	0.0074
Events	3	8	
Median/month	--	--	

Progression free survival

	PIK3CA	PIK3CA/GATA3	P-value
No.	10	3	0.3068
Events	3	2	
Median/month	--	4.3	

Progression free survival

	FOXA1	FOXA1/GATA3	P-value
No.	15	7	0.3589
Events	6	4	
Median/month	--	2.9	

Overall survival

	ESR1	ESR1/GATA3	P-value
No.	28	19	0.0057
Events	9	11	
Median/month	--	22.7	

Overall survival

	PIK3CA	PIK3CA/GATA3	P-value
No.	10	3	0.5954
Events	7	3	
Median/month	19.2	9.1	

Overall survival

	FOXA1	FOXA1/GATA3	P-value
No.	15	7	0.0049
Events	8	5	
Median/month	18	9	

Supplementary Figure 8. Survival analyses. Kaplan–Meier curves for (A) progression-free survival (PFS) and (B) overall survival (OS) comparing patients with single mutant genes with patients who had subclones of the same mutated genes combined with GATA3 mutations. P-values were calculated using the log rank test; ⊥ censored

9. Acknowledgment

First and foremost, thanks to the God, the Almighty, for supporting and blessing me to complete this work successfully.

First of all, I would like to express my sincere appreciation and gratitude to my advisor Prof. Dr. Klaus Pantel, for giving me the opportunity to conduct my doctoral thesis under excellent conditions in such a fascinating field of research at the Tumor Biology Institute. Also, for help and support and for the many small conversations in the lobby or in the elevator about my research, which often give me a new idea. I would like to thank you for encouraging my research and allowing me to grow as a research scientist.

I would like to thank Prof. Dr. Tobias Lenz for his insightful supervision and comments on my Ph.D.

I would like to express my special appreciation and thanks to my mentor Dr. Simon Joosse for the continuous support of my Ph.D. study, motivation, and immense knowledge also provided valuable ideas for better versions of the publications. It was always exciting for me to research new topics with him, and given me various approaches to improve my dissertation. He showed me how to be a researcher and encouraged me to figure out the best way to arrange research data using proper solution techniques. I immensely enjoyed and learned from every meeting I had with him. My special thanks go to all members of the Joosse group, Reteno, Mais, Gresa, Nikhil, and Luisa.

10. Eidesstattliche Versicherung

Eidesstattliche Versicherung
Declaration on oath

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.
I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hamburg, den   **Unterschrift**

24-Jan-2022