Liquid Biopsy in Breast Cancer: Molecular Characterization and Support to Clinical Management

Thesis

The Open University
Milton Keynes, United Kingdom

Affiliated Research Centre
Fondazione IRCCS Istituto Nazionale dei Tumori
Milano, Italia

Liquid biopsy in breast cancer: molecular characterization and support to clinical management

Thesis presented for the Degree of Doctor of Philosophy The Open University, Milton Keynes (UK) School of Life, Health and Chemical Sciences

Serena Di Cosimo
Medical Doctor
Personal Identifier: G4607089

March 2021

ABSTRACT

The classical paradigm of breast cancer (BC) management has been challenged by recent advances in the understanding of molecular biology, which is increasingly encouraging attempts to move away from protocol-based in favor of a personalized approach. This project aimed to test circulating tumor DNA (ctDNA) and circulating tumor cells (CTCs) to characterize the disease, and monitor patients with BC including both non-invasive and invasive cases.

For this purpose, 41 patients with newly diagnosed ductal carcinoma in situ (DCIS), and 61 triple negative BC (TNBC) patients treated with neoadjuvant chemotherapy were analyzed. Primary tumor mutations identified by targeted-gene sequencing were validated, and tracked in 182 plasma samples longitudinally collected at multiple time-points by droplet-digital Polymerase Chain Reaction. In recurrent TNBC patients, plasma DNA underwent direct targeted-gene assay, and CTCs were analyzed for Copy Number Alterations (CNAs).

In DCIS patients, ctDNA from baseline samples was evaluable in 71% of cases, with Variant Allele Frequency (VAF) values ranging between 0.001% and 5%, independently of clinico-pathological features, and likewise at initial diagnosis and at relapse. In TNBC patients, the number of primary tumor mutations showed a downward trend after treatment; proficient genes involved in the immune pathways were associated to response; and surgical samples of non responsive cases presented alterations in druggable pathways. ctDNA after treatment was associated with increased risk of relapse, and its prognostic value remained worthy even after adjusting for age, residual disease, systemic inflammatory indices, and Ki-67 (HR 1.91; 95%CI 0.51-7.08). During follow-up,

ctDNA was undetectable in non-recurrent cases. Conversely, ctDNA was detected in 83% of recurrent cases, and antedated the clinical diagnosis up to 13 months. Notably, recurrent cases without ctDNA developed loco-regional, contralateral, and bone-only disease. At clinical progression, CTCs presented chromosome 10 and 21q CNAs whose network analysis showed connected modules including HER/PI3K/Ras/JAK signaling, and immune response.

These findings extend the value of liquid biopsy to further clinical scenarios by reporting - unlike previous studies - the presence of ctDNA in noninvasive BC; and substantiate the development of ctDNA in invasive BC as an exploitable tool, either alone or with CTCs, for personalized TNBC management.

DEDICATION

To my son Giovanni

ACKNOWLEDGEMENTS

I want to thank my supervisor, Dr. Maria Grazia Daidone, for introducing me to biomarkers study, for her precious feedbacks on my work, and for inspiring me along the way. She helped me grow, taught me to be persistent, and not to give up; for that I will always be thankful.

I am very grateful to Dr. Valentina Appierto, Dr. Vera Cappelletti, and Dr. Luca Roz: their feedbacks and questions helped me to gain a perspective in progressing this research to its outcome.

I am truly in debt to my family, with their love and patience encouraged all throughout my life to become a better version of myself. I owe everything to them, and their support. My dear Francesco for his kindness, positivity, being there for me, and for inspiring me to dream (and for tending to our child during the long afternoons and week ends I spent studying and writing this essay).

This study was conducted under the supervision of Drs. Valentina Appierto and Maria Grazia Daidone at the Biomarkers Unit, headed by Dr. Maria Grazia Daidone, at the Department of Applied Research and Technological Development, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy, between October 1, 2017 and October 30, 2020.

TABLE OF CONTENTS

ABBREVIATIONS

ADO	Allelic Drop Out
AI	Allelic Imbalance
BC	Breast Cancer
CCP	Comprehensive Cancer Panel
CD	Cluster of Differentiation
cfDNA	cell free DNA
CI	Confidence Interval
CISH	Chromogenic In Situ Hybridization
CK	CytoKeratin
CNA	Copy Number Alteration
CTC	Circulating Tumor Cell
ctDNA	Circulating tumor DNA
DCIS	Ductal Carcinoma In Situ
DEP	DielEctroPhoresis
DIN	DNA Integrity Number
DNA	DeoxyriboNucleic Acid
gDNA	genomic DNA
GE	Genomic Equivalent
GII	Genome Integrity Index
ddPCR	droplet digital Polymerase Chain Reaction
eCTC	epithelial-Circulating Tumor Cell
EFS	Event-Free Survival
EpCAM	Epithelial Cell Adhesion Molecule
ER	Estrogen Receptor
FDA	United States Food and Drug Administration
FFPE	Formalin-Fixed Paraffin Embedded
HER2	Human Epidermal Growth Factor Receptor 2
HR	Hazard Ratio
HS	HotSpot
IBC	Invasive Breast Cancer
IBTR	Ipsilateral Breast Tumor Recurrence

IHC	ImmunoHistoChemistry
INT	Istituto Nazionale Tumori
IQR	InterQuartile Range
LMR	Lymphocyte-to-Monocyte Ratio
lpWGS	low-pass Whole Genome Sequencing
MDA	Multiple Displacement Amplification
NAC	Neoadjuvant chemotherapy
ncCTC	non-conventional Circulating Tumor Cell
NGS	Next Generation Sequencing
NLR	Neutrophil-to-Lymphocyte Ratio
PBMC	Peripheral Blood Mononuclear Cells
pCR	pathological Complete Response
PR	Progesterone Receptor
RT	Room Temperature
SNV	Single Nucleotide Variant
sTILs	Stromal tumor infiltrating lymphocytes
T	primary Tumor
TNBC	Triple Negative Breast Cancer
UMI	Unique Molecular Identifier
VAF	Variant Allele Frequency
WBC	White Blood Cells
WES	Whole Exome Sequencing
WGA	Whole Genome Amplification
WGS	Whole Genome Sequencing
wt	wild type

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1.

INTRODUCTION

1.1 Breast cancer

Epidemiology **-** Breast cancer is a significant public health problem, accounting for approximately 24% of new cancer cases diagnosed in women each year worldwide [Bray F et al., 2018]. According to the International Agency for Research on Cancer, nearly 2 million new cases were diagnosed in 2018, and incident cases are expected to increase by almost 50% in the next two decades, rendering it the most prevalent cancer [Global Cancer Observatory, 2020].

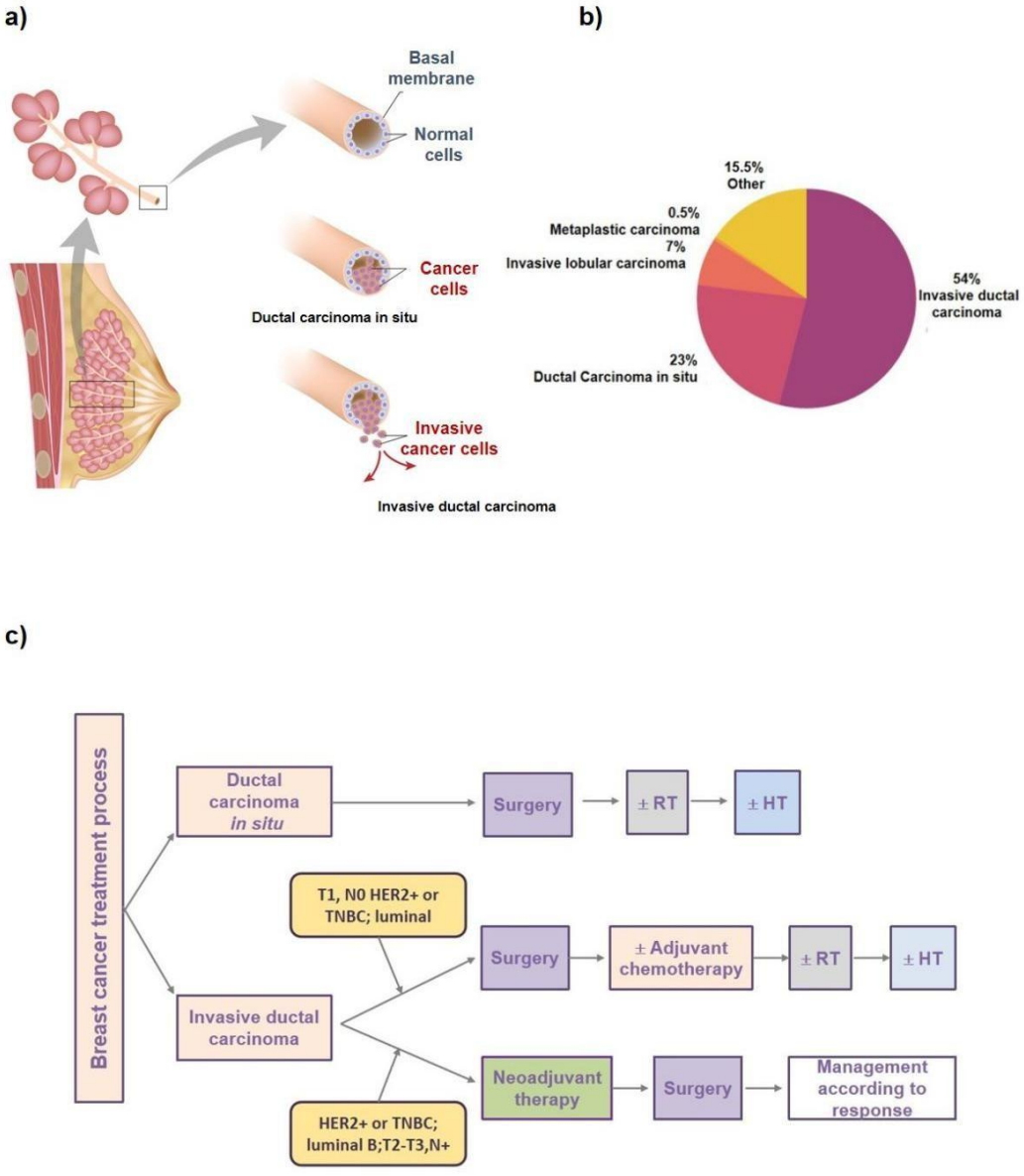


Figure 1. Breast cancer a) Stages of disease progression, modified from https://www.caperay.com; b) Types and relative incidence; c) Treatment algorithm.

The risk of developing breast cancer increases with age, with a probability of developing breast cancer of 2.3% up to the age of 49 (1 in 43 women), 5.4% in the 50-69 age group (1 in 18 women) and 4.5% in the 70-84 age group (1 in 22 women) [AIRTUM - AIOM, 2020].

This correlation with age could be linked to the endocrine proliferative stimulus that the mammary epithelium undergoes over the years and to the possible accumulation of genetic and epigenetic alterations that can cause an imbalance between oncogenes and suppressor genes. The incidence curve rises exponentially until the menopause years (around 50-55 years), then slows down with a plateau after the menopause, and then rises again after 60 years.

This specific trend is related both to the endocrinological history of the woman and to the presence of screening programs.

Thanks to mammography screening, breast cancer primary tumors are now routinely found at the sub-clinical level [Siegel RL et al., 2018] (Figure 1, a). With the discovery of breast cancer at earlier stages there is the risk that some indolent or slowly growing lesions, *i.e.,* ductal carcinoma in situ will be overdiagnosed or overtreated [Siu AL et al., 2016; Morrow M et al., 2015]. Reasonable estimates show the overdiagnosis rate with mammography screening to be 1-10% [reviewed by Pashayan N et al., 2020]. Until we are able to discriminate with certainty between cases of breast cancer that will not cause damage or relapse and those that will, this small overdiagnosis is bound to persist. It should not, however, be used as a reason to delay screening or to widen the screening interval, because neither strategy will decrease the small amount of overdiagnosis that exists. Improvements in personalized treatment will hopefully diminish the morbidity of treatment and, therefore, the significance of overdiagnosis. Though most women present with localized potentially curable tumors, incurable and lethal relapses remain frequent in clinical practice. Indeed it is estimated that 7% of all cancer

deaths are attributed to breast cancer, making it the second leading cause of cancer-related deaths, and the first among females [Siegel RL et al., 2018].

Histological classification **-** Most breast malignancies arise from epithelial elements and are categorized as carcinomas (Figure 1, b). Breast carcinomas are a diverse group of lesions that differ in microscopic appearance and biologic behavior, although these disorders are often discussed as a single disease. The in situ carcinomas of the breast are either ductal (also known as intraductal carcinoma) or lobular. This distinction is primarily based upon the growth pattern and cytologic features of the lesions, rather than their anatomic location within the mammary ductal-lobular system. The invasive breast carcinomas consist of several histologic subtypes; the estimated percentages are reported in Figure 1, b. Other subtypes, including metaplastic breast cancer and invasive micropapillary breast cancer, all account for less than five percent of cases [Dillon DA et al., 2009].

Molecular classification **-** Breast cancer is a heterogeneous disease and patients with tumors apparently similar in clinicopathological features may present a different clinical course (Figure 2). Following investigations of gene expression analysis by microarray methods that identified an intrinsic gene list of 496 genes [Sorlie T et al., 2001; Sorlie T et al., 2003], four subtypes of invasive carcinomas were identified [Van de Vijver MJ et al., 2002]:

- "Luminal A": neoplasms with marked expression of hormone receptors;
- "Luminal B": neoplasms that, although expressing hormone receptors, have a high risk of recurrence, due to the high proliferative index correlated with high expression of proliferation genes;
- "HER2": characterized by the presence of HER2 expression;

- "Basal like": tumors characterized by the absence of expression of hormone receptors and HER2 and increased expression of basal cytokeratins (e.g. CK5/6 and CK14).

These subgroups have also proved to be prognostically important, Luminal A carcinomas having a favourable prognosis, much better than Luminal B, and HER2 positive and Basal-like carcinomas having the worst prognosis of all subgroups.

Within these subtypes there is a high heterogeneity. In the light of new pathological and molecular knowledge there is a definition of further subtypes of breast cancer. For example, another subgroup of breast cancers has been identified with no expression of hormone receptors and HER2, but with stem cell markers, low expression of claudin (cell-cell junction proteins) and accompanying lymphocytic infiltrate, termed "claudin low" and characterised by poor prognosis [Prat A et al., 2010] . In addition, a gene expression analysis of 587 triple negative breast cancers identified six different subtypes with different molecular biology and clinical behaviour: Basal like 1 and 2 (BL1 and BL2), immunomodulatory (IM), mesenchymal (M), mesenchymal stem-like (MSL) and Luminal Androgen Receptor (LAR) [Lehmann BD et al., 2011]. In clinical practice, immunohistochemical evaluation of hormone receptor status, Ki67 and HER2, allows surrogate identification of the four phenotypic subgroups of breast cancer that present a relative correspondence with the four derived from gene expression profiles [Cheang MC et al., 2009].

Immunophenotypic groups of clinical relevance and with important therapeutic implications, including systemic therapy, are:

- Luminal A: hormone receptor positive, HER2 negative and low proliferative activity (which frequently includes some special histotypes such as tubular carcinoma, lobular carcinoma classic type). Luminal A breast cancer cases are

represented by tumors with positive estrogen receptors, with positive progesterone receptors (*i.e.*, ≥ 20%), with negative HER2, and Ki67<20%.

- Luminal B: hormone receptor positive, HER2 negative and high proliferative activity;

- HER2 positive luminals: hormone receptor positive, HER2 overexpressed (immunohistochemistry reaction score 3+) or amplified, any proliferative activity value;

- HER2 positive (non-luminal): HER2 overexpressed (3+ score of immunohistochemistry reactions) or amplified (FISH or other methods) and both hormone receptors negative

- Triple-negative: absence of hormone receptor expression and HER2 negativity. The correspondence between the immunohistochemically detected triple negative phenotype and the intrinsic Basal Like subgroup detected on a gene basis only exists in about 80% of cases, further demonstrating the extreme heterogeneity present within these subgroups. The triple negative subgroup includes some special histological types such as typical medullary and adenoid cystic, which have a low risk of recurrence. Retrospective analyses have associated the four subtypes with differences in disease-free survival, sites of recurrence and overall survival (Figure 2) [Kennecke H et al., 2010].

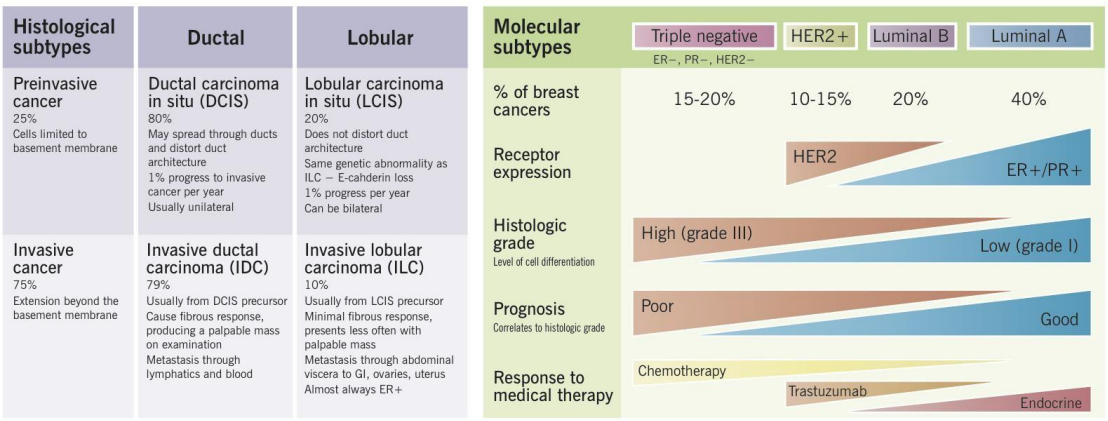


Figure 2. Breast cancer classification according to histological findings and molecular subtypes, adpated from http://www.pathophys.org.

1.1.1 Current management

Systemic treatment was initially employed along with surgery and radiotherapy for extensive loco-regional disease control, but became a main component of breast cancer management once breast cancer was recognized as a systemic disease at onset due to the potential for lymphatic and blood dissemination [Veronesi U et al., 1981].

Traditionally, systemic therapy has been administered to breast cancer patients after surgery (adjuvant therapy). Recently, however, neoadjuvant systemic therapy has been regarded as an equally effective option when compared to adjuvant therapy [Early Breast Cancer Trialists' Collaborative Group, 2018]. Whereas neoadjuvant anti-hormonal therapy is mainly recommended for postmenopausal patients with hormone receptor positive breast cancer [Yeo B and Dowsett M, 2015], neoadjuvant chemotherapy is increasingly utilized for all breast cancer subtypes (Figure 1, c).

The use of systemic treatment in the neoadjuvant setting is associated with equivalent overall survival rates compared to the use of the same systemic treatment in the post-operative setting. Specifically, a single-patient meta-analysis compared the outcomes of neoadjuvant as compared to adjuvant systemic therapy, based on data from 4,756 women enrolled in ten studies initiated between 1983 and 2002 [Early Breast Cancer Trialists' Collaborative Group, 2018]. No significant difference was observed in terms of distant recurrence rates (15-year risk: 38.2% vs 38%, Relative Risk [RR] 1.02, CI 95% 0.92-1.14, p=0.66), breast cancer-related mortality (15-year risk: 34. 4% vs 33.7%, RR 1.06, CI 95% 0.95-1.18, p=0.31) or overall mortality (15-year risk: 40.9% vs 41.2%, RR 1.04, CI 95% 0.94-1.15, p=0.45), between neoadjuvant and adjuvant chemotherapy. Systemic neoadjuvant treatment was found to be associated with an increased risk of local recurrence (local recurrence at 15 years: 21.4% vs. 15.9%, HR 1.37, CI 95% 1.17-1.61,

p=0.001), attributable at least in part to the increased use of conservative surgery, with the increased risk essentially confined to women in trials where surgery had been omitted.

Thus, in the face of comparable distant recurrence rates and overall survival between systemic treatment administered in the adjuvant as compared to neoadjuvant setting, the advantages of neoadjuvant systemic treatment are listed below:

- Primary tumor down-staging to allow conservative surgical treatment in patients who would otherwise be candidates for demolitive (*i.e.*, mastectomy) surgery;
- Treatment of locally advanced breast cancer (stage IIB-IIIC), since, irrespective of the subtype, conservative surgery is not amenable in the majority of cases;
- Treatment of early-stage breast cancer (including stages I-IIA) when conservative surgery is not feasible (e.g. because of a high tumour-to-mammary ratio or when the expected cosmetic outcome is suboptimal because of a particular tumour location).
- Possibility of allowing prognostic assessment based on pathological response after neoadjuvant systemic treatment. The positive prognostic role - at individual patient level - of achieving complete pathological response after neoadjuvant systemic treatment is well established. In particular, in a large meta-analysis sponsored by FDA (CtNeoBC), which included 11,955 patients enrolled in 12 clinical trials, a highly significant association - at individual patient level - was reported between achievement of pathological complete response (defined as absence of residual invasive disease in the breast and lymph nodes - ypT0/is, N0) and survival, in terms of event-free survival (EFS) and overall survival (OS) [Cortazar P et al., 2014]. This association has been reported in the overall population and in all breast cancer subtypes. However, the strength of this

association was found to be more robust in subtypes characterized by greater biological aggressiveness, such as HER2 positive and triple negative.

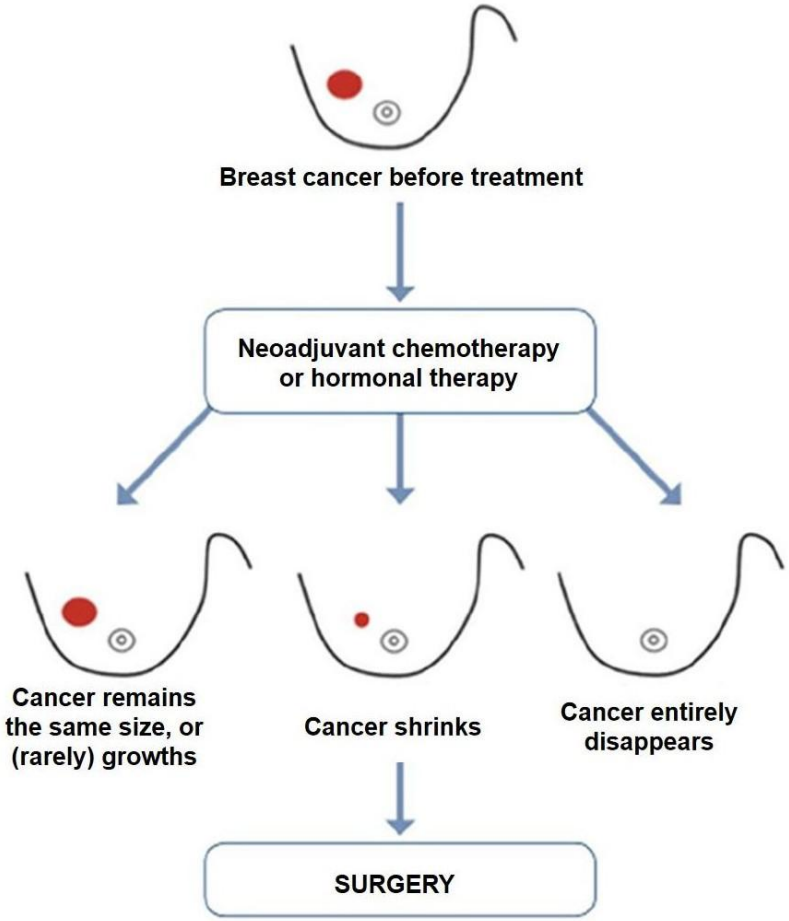


Figure 3. Neoadjuvant therapy overview

- Possibility of calibrating post-operative treatment on the basis of the presence of residual disease after neoadjuvant systemic treatment. Indeed, the presence of residual disease after neoadjuvant systemic treatment is a recognized negative prognostic factor. In this context, depending on the subtype of breast cancer, several treatment options are now available for patients who have not achieved pathological complete response after standard neoadjuvant systemic treatment (therefore considered "high risk"), including “rescue” adjuvant chemotheray and/or treatment including targeted therapy, immunotherapy and conjugated antibodies.

Hence, neoadjuvant chemotherapy increases the rate of breast conserving surgery [Kaufmann M et al., 2012], and allows monitoring of treatment and tumor biological response, providing unique opportunities for development of both individualized treatment strategies and drug development [Bardia A and Baselga J, 2013] (Figure 3).

1.1.2 Future perspectives

During the last decades, enormous advances have been made in the study and dissection of tumor molecular characteristics, thanks to the advent of the -omics era. These advances have led to the origin of precision oncology as a branch of precision medicine [Meric-Bernstam F and Mills GB, 2012], shifting from a *"one size fits all"* therapeutic approach towards the identification of *"the right treatment, for the right patient, at the right time"*. The applications of tumor molecular profiling encompass all the stages of cancer management [Cronin M and Ross JS, 2011] including screening, prognosis, patient stratification for predicting response to therapy, and for the selection of personalized therapies, monitoring of treatment effectiveness, and follow-up for early detection of relapse and metastasis occurrence [http://precision.fda.gov].

The heterogeneity of breast cancer can be addressed by mutational profile, and recent studies have demonstrated that individual cases are composed of a mosaic of different clones with different sets of genetic aberrations [Gagan J and Van Allen EM, 2015; Zarvadas D et al., 2015]. Massively parallel sequencing analyses revealed that primary breast cancers have a complex repertoire of somatic genetic alterations, which vary according to hormone receptor status [Cancer Genome Atlas Network, 2012; Nik-Zainal S et al., 2016].

Recurrent mutated genes are few, namely *TP53, PIK3CA, GATA3, MAP3K1* and *CDH1*. Notably, among the subset of rarely mutated genes,

druggable genetic alterations have been reported, affecting *HER2* [Bose R et al., 2013], *ESR1* [Toy W et al., 2013], and *AKT1* [Bleeker FE et al., 2008]. More recently, studies have demonstrated that a large subset of breast cancer present intra-tumor genetic heterogeneity, whose level and type differ between estrogen receptor-positive and -negative cases [Ellsworth RE et al., 2017; Ng CK et al., 2015].

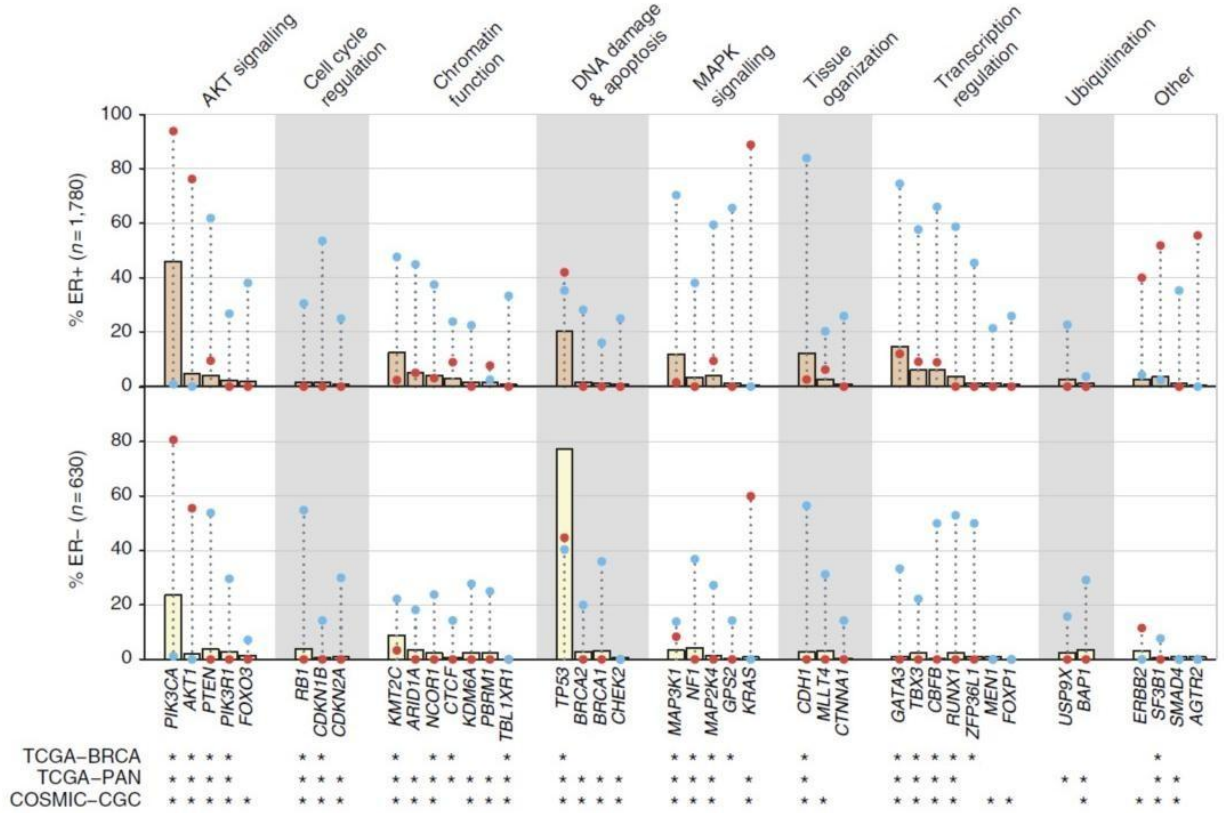


Figure 4. Primary tumor somatic mutations in breast cancer. Bars depict proportions of ER positive and ER negative samples harbouring mutations in mutation-driver genes. Red and blue points indicate for each gene, the proportions of recurrent and inactivating mutations, respectively. Asterisks indicate genes highlighted in: COSMIC, Cancer gene census from the Catalogue of Somatic Mutations in Cancer; TCGA-BRCA, TCGA breast cancer study; TCGA-PAN, TCGA pan-cancer analysis. Estrogen Receptor (ER) status available for 2,410 tumors. MAPK, mitogen-activated protein kinase. The genes are grouped by pathway or function. Modified from Pereira B et al., 2016.

Given this genetic heterogeneity, it is conceivable that in the transition from primary to metastatic disease, the metastatic process itself as well as the therapeutic interventions administered in the adjuvant setting may result in clonal selection. In this context, the evaluation of primary versus metastatic breast cancer would lead to the identification of genes whose genetic alterations are specific or enriched in metastatic lesions compared to the respective primary tumors. Hence, therapeutic strategies targeted at predominant aberrations could be ineffective, especially in the case of recurrent disease were temporal evolution is likely to occur [Nik-Zainal S et al., 2012; Stephens PJ et al, 2012; Siu LL, 2013]. Academic enterprises are therefore encouraged to build a longitudinal map of breast cancer evolution that interrogates both primary and recurrent disease sites.

To achieve this, traditionally, multiple biopsies would be necessary. Instead of repeated biopsies, a more convenient strategy could be blood-born biomarkers, such as circulating tumor DNA (ctDNA) and circulating tumor cells [Alix-Panabieres C, Pantel K., 2013; Dawson SJ et al., 2013; Speicher MR, Pantel K, 2014].

1.2 Circulating tumor DNA

DNA circulates freely in the blood of both healthy and sick individuals [Mandel P and Metais P, 1948]. A crucial discovery was made when Leon and colleagues demonstrated that the levels of cell-free DNA (cfDNA) were higher in patients than in healthy subjects [Leon SA et al., 1977]. This was followed by the work conducted by Stroun and colleagues which traced the neoplastic origin of cfDNA present in the plasma of cancer patients [Stroun M et al., 1989]. The mechanism by which the tumor-derived fraction of cfDNA, commonly named circulating tumor

DNA (ctDNA), is released in circulation is not entirely known. Apoptosis and necrosis of cancer cells are thought to be the major contributors. Other possibilities include active secretion of DNA by tumor cells, as well as potential if not minor contributions from circulating tumor cells [reviewed by Kustanovich A et al. 2019].

ctDNA has received enormous attention during the last decade owing to its huge potential as a minimal invasive tumor biomarker to aid in the diagnosis, treatment, and prognosis of tumor diseases, showing greater sensitivity compared to the previously used blood biomarkers [reviewed by Bronkhorst AJ et al., 2019]. However, the detection of ctDNA at an early stage has shown greater challenge due to its presence in relatively low quantities [Heitzer E et al. 2019].

Although the amount of cfDNA in patients with cancer varies widely, most studies report that 1 ml plasma from a patient with cancer contains approximately 1,500 diploid genome equivalents (GE) (~10 ng DNA) (Figure 5, a), with considerably higher amounts often observed in patients with metastatic cancer. A typical 10 ml blood draw yields on average 4 ml plasma containing 6,000 GE (12×10^3 molecules per region or gene), which implies a theoretical sensitivity limit of ~0.01% (that is, the ability to detect 1 somatic variant in 12,000 copies). If the VAF of ctDNA corresponds to 0.1%, there are on average just six molecules per tube carrying the respective mutation, which may be affected by stochastic sampling (Figure 5, b) as reported by Heitzer E et al. 2019.

Furthermore, when one circulating primary tumor mutation is admixed with 20,000 or more normal copies, such as in the case of patients with primary tumors ≤1 cm, the use of 4 ml of plasma will likely not contain a single cancer genome for sequencing, thus rendering the diagnosis of cancer impossible due to sampling error [Lievens A et al. 2016]. Consequently, the reliability, accuracy, and

reproducibility of ctDNA detection represents a challenge, and requires specialized technologies.

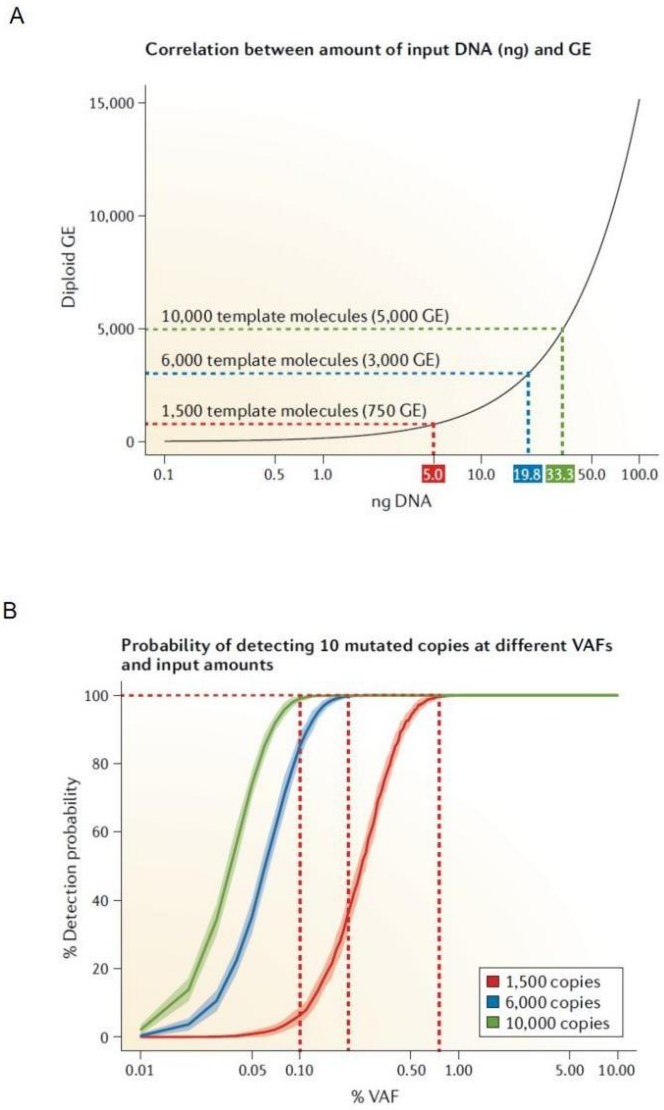


Figure 5. Cell-free DNA input and somatic variants detection

A) The graph shows the relationship between the amount of input DNA (in nanograms) and the number of template molecules in terms of diploid genome equivalents (GE). B) The graph shows the relationship between the number of input molecules (copies) and the probability that 10 mutated copies will be detected in a total of 1,500, 6,000 or 10,000 copies at different variant allele frequencies (VAFs). For a reliable detection of a mutation with a VAF of 0.1% (10 mutated copies), at least 10,000 template molecules (5,000 GE or 33ng of input DNA) are necessary. Modified from Heitzer E et al. 2019.

1.2.1 Technological approaches to ctDNA analysis

Current methods to identify sequence alterations in cfDNA can be divided in polymerase chain reaction (PCR)- and Next-Generation Sequencing (NGS)-based techniques [Elazezy M and Joosse SA, 2018] (Figure 6).

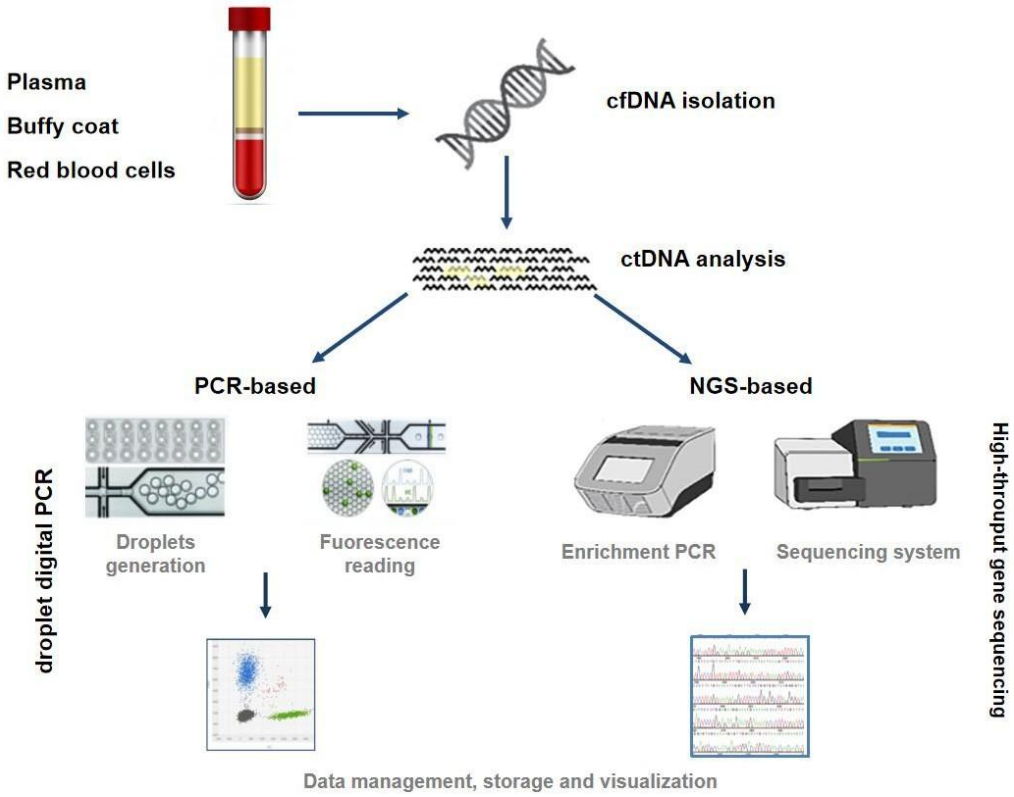


Figure 6. PCR- and NGS-based approaches for ctDNA analysis

PCR-based techniques are used for single locus assays, and were historically the first to be implemented to identify mutant *KRAS* in pancreatic carcinoma [Sorenson GD et al., 1994]. Although useful, allele specific PCR has very limited analytical sensitivity [Wan JCM et al., 2017], and it is currently being substituted by more sensitive approaches, such as droplet digital PCR (ddPCR) [Heitzer E et al., 2019]. While in traditional PCR a single sample offers only a single measurement, in ddPCR the sample is partitioned into 20,000 nl-sized droplets. This partitioning enables the measurement of thousands of independent amplification events within

a single sample. ddPCR-based approaches have high sensitivity (ranging between 0.001% and 0.01%), allow ctDNA absolute quantification, and are currently being used to detect specific known mutations, such as the main driver mutations of the primary tumor, or variants associated with response to drugs in individual tumor types.

NGS-based techniques have been developed to interrogate multiple mutations in parallel and to identify *de novo* mutations, and range from the analysis of several tens of mutations, to a genome-wide analysis of cfDNA by whole-exome sequencing, or whole-genome sequencing [Newman AM et al., 2014; Zehir A et al., 2017; Gray ES et al., 2019].

In general, the larger the analysis, the more expensive and difficult is to obtain high sensitivity for mutation calling. Targeted sequencing (also known as NGS panels) involves the sequencing of specific loci after selective amplification by PCR (amplicon-based) or hybrid capture-based enrichment [Forshew T et al. 2012; Gale D et al., 2018; Chicard M et al., 2018]. NGS panels, usually designed to include cancer specific genes of interest with a diagnostic or prognostic value, are being widely used to assess tumor heterogeneity, to follow clonal evolution under and after treatment, and to identify potential resistance mutations. Non-targeted approaches are instead reserved to research in the field of immuno-therapy, where mutation load *(i.e.*, the number of non synonymous mutations found in a tumor) has emerged as a putative biomarker of the response to treatment.

1.3 Circulating tumor cells

Circulating tumor cells (CTCs) were discovered by Thomas Ashworth in the 1860s during an autopsy of a patient with metastatic cancer. These cells have the ability to detach from a primary or metastatic tumor site, penetrate the vessel wall and enter the circulatory system [Zhang et al., 2017; Domínguez-Vigil et al., 2018].

CTCs can be isolated as single cells or clusters from the blood of patients with cancer.

1.3.1 Detection and enumeration

The differences between CTCs and normal blood cells in physical and biological properties had led to the development and commercialization of several CTC detection platforms, which can be categorized as label-dependent or -independent (Habli Z et al., 2020).

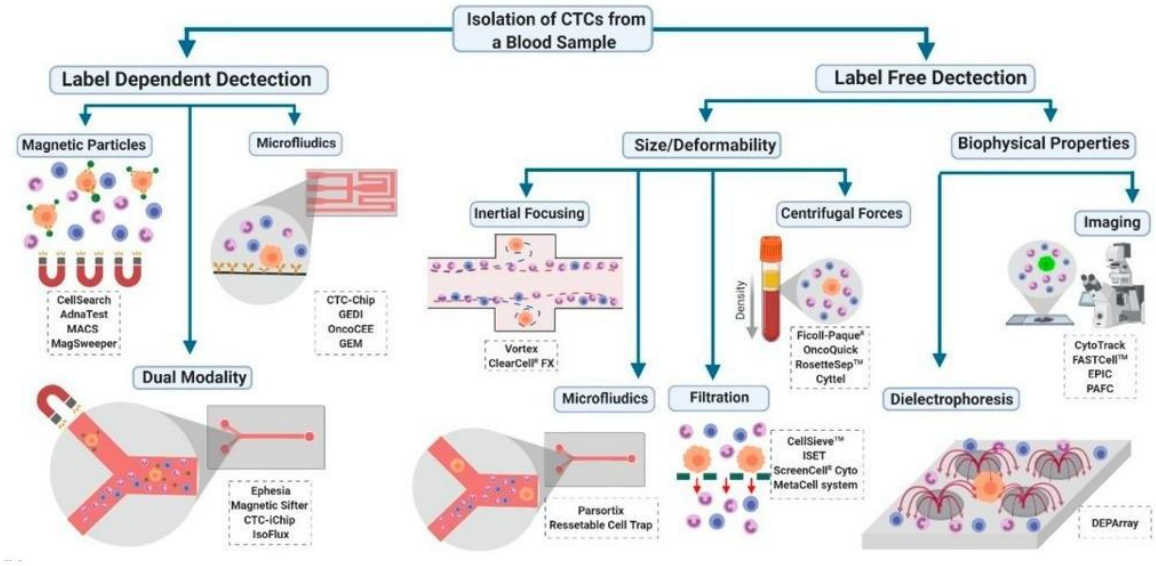


Figure 7. Label-dependent and -independent CTCs detection. Commercialized technologies for CTC detection, enumeration and count. CTCs can be detected based on their molecular or physical properties. Examples of commercial boxes. Adapted from Habli et al., 2020.

Among the leading label-dependent platforms, the FDA-approved Cell Search® system (Menarini Silicon Biosystems) captures CTCs using antibodies directed against EpCAM, and defines CTCs as nucleated CK8-positive, CK18-positive or CK19-positive, and CD45-negative cells. Important correlations between CTC count and prognosis have already been reported using Cell Search® in patients with different type of cancer, including breast cancer [Cristofanilli M et al., 2004; Cristofanilli M et al., 2005; Riethdorf S et al., 2007]. Specifically, in the metastatic setting, patients with ≥5 CTCs per 7.5 ml blood showed reduced progression-free

survival, and those with <5 CTCs per 7.5 ml improved overall survival [Pantel K et al., 2008; Bidard FC et al., 2014]. Furthermore, an increasing baseline CTC count was associated with a progressively worse prognosis in a recent metanalysis of clinical trials with neoadjuvant therapy for patients with early-stage breast cancer [Bidard FC et al., 2018]. However, Cell Search® CTC enumeration has some limitations. First, EpCAM expression in tumor cells decreases during the epithelial-to-mesenchymal transition that accompanies tumor progression [Rao CG et al., 2005]; and different studies showed that CTCs may not show EpCAM expression on their surface [Sieuwerts AM et al., 2009]. Therefore, an exquisitely EpCAM-based technology is able to detect just a limited number of CTCs. Secondly, cell isolation through the Cell Search® system is followed by cell fixation, which prevents further viable CTC characterization. Finally, sensitivity of Cell Search® system is 1 cell per ml, and detected CTCs are usually contaminated, the purity of captured cells ranging between 60–70% [Habli Z et al., 2020]. Label-independent approaches allow the isolation of CTCs with a low epithelial phenotype because these platforms discriminate CTCs from other cells based on physical characteristics such as size, density, deformability, and electrical properties (Figure 7).

The Parsortix® system (Angle plc, Guildford, UK) is an example of such an approach because of its scale architecture, which gradually decreases in width to 4.5 µm, favors the embedding of CTCs, which are larger than other circulating cells while preserving their morphology [Obermayr E et al., 2015]. This microfluidic system is used as an initial enrichment step for the subsequent application of dielectrophoresis-based techniques (DEP) capable of separating CTCs, and leukocytes based on differences in their conductivities through the use of a non-uniform alternating electric field [De Luca F et al., 2016]. The advantage of the

DEP strategy is that it allows enumeration and also isolation of single CTCs for downstream molecular analyses.

1.3.2 Molecular characterization

To achieve accurate genomic analysis of CTCs at the single-cell level, whole genome amplification (WGA) of DNA from a single cell must be performed with sufficient breadth and precision (Figure 8) [reviewed by Wang Y and Navin NE, 2015].

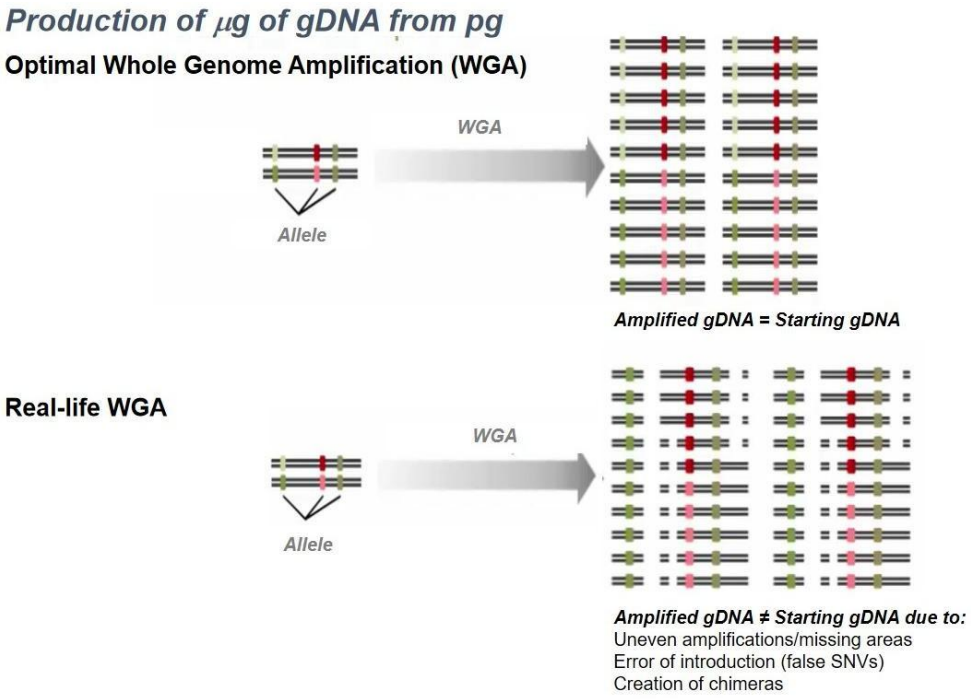


Figure 8. Genomic DNA (gDNA) amplification from single cells

Currently WGA methods include PCR, and multiple displacement amplification (MDA) [Blainey PC, 2013]. PCR-based approaches typically conduct PCR amplification using degenerate oligos as primers (*i.e.*, DOP-PCR [PicoPlex]), or linker adaptors with universal sequences ligated to the DNA fragments (*i.e.*, LM-PCR [GenomePlex and Ampli1]). On the other hand, the isothermal MDA-based methods (*i.e.*, Repli-g) use high-fidelity Phi29 DNA polymerase for amplification of larger DNA fragments across the genome with a 1000 fold higher fidelity than PCR-based approaches. WGA methods that hybridize the principles of PCR and

MDA-based approaches have also been reported (*i.e.*, multiple annealing and loop-based amplification cycling [MALBAB]). The optimal WGA method *per se* does not exist as all these approaches present technical issues (Table 1). Common amplification errors include preferential amplification with secondary allelic imbalance (AI), PCR reaction failure in an allele, (*i.e.,* allele drop out [ADO]), and false positive detection of SNVs.

Table 1. Strength and weakness of WGA systems

	MDA-based	PCR-based
Amplified Fragment Lengths	10 – 100 kb	~1-2 kb
Nucleotide Error Rate	10^{-8}	10^{-4} - 10^{-5}
Chimera Formation	Higher (?)	Lower (?)
Completeness of Genome Coverage	High	10 – 70%
Variability of Amplification	High	Low
CNV Detection	Poor	Good
Duplicate Formation	Lower	Higher
Allelic Dropout (ADO) (AB → AA or BB)	5 -50% (?)	?
SNV Detection	OK	+/-
Protocol	Simple	Often multi-step

Thanks to random primers, which bind the DNA at multiple points to ensure extensive coverage, PCR-based approaches are able to amplify about 96% of the genome at least 1000 times. However, PCR-based methods can produce artifacts and generate DNA fragments less than 1 kb in length, which are unusable for many subsequent applications. In contrast, MDA methods ensure intact products of significant length (>10 kb) and a reduction in amplification errors up to four orders of magnitude. However, the caveat of MDA is that it generates non-uniform coverage and can therefore result in important distortions of copy-number states. In general, PCR-based methods are thought to generate increased uniformity, but cause more errors at the single nucleotide level than MDA-based methods. MDA-based methods suffer from ADO and AI drawbacks, which may prevent applications in CNA analysis. Hybrid approaches finally have broader genomic coverage while maintaining sufficient uniformity for CNA analysis, can still result

in >30% base dropout, a potentially significant sacrifice in the sensitivity to detect single nucleotide mutations. In addition to considering the characteristics of the individual methods, there is also the need to use the one with fewer steps and reduced reaction volume to minimize or completely avoid possible contamination of the DNA sample by the surrounding environment. After WGA the amplified DNA is used to construct libraries for NGS. While high-coverage WGS is the most comprehensive and data-rich approach, alternative methods can deliver genetic information for a lower cost. Low-pass WGS (typically defined as <1X coverage of the genome) is twenty-fold less expensive than complete genome sequencing, yet provides data for useful comparative genomic analyses especially for structural variation such as CNAs.

1.3.3 Copy Number Alteration (CNA) analysis

Studies exploring the landscape of CNAs have considerably advanced our knowledge of breast cancer biology, with translational efforts leading to advances in the clinic. This was the case for the amplification of the *ERBB2* (*HER2*) oncogene, which defines a biological subtype of breast cancer, and whose targeting led to the development of therapeutic agents that increased the survival of HER2 amplified breast cancer patients [Baselga J and Albanell, 2001; Arteaga and Engelman, 2014]. Additionally, numerous studies have demonstrated the utility of copy number information in the prognostic stratification of breast cancer patients [Russnes HG et al., 2010; Shah SP et al., 2012], and studies of the genes targeted by this class of somatic mutations have paved the way for new therapeutic targets [GatzaML et al., 2014; Cai Y and Sablina AA, 2016].

However few studies have explored the role of CNAs in the CTCs of breast cancer patients [Riebensahm C et al., 2019; Kanwar N et al., 2015]. Specifically, CNA patterns showed a high clonality among different CTCs of the same patient, and aberrations thought to be unique to CTCs were detected at subclonal frequencies

within primary tumors, indicating pre-existing genetic signatures for metastasis. These subclonal alterations were enriched in CTCs and metastases, pointing towards the selection of a more “fit” component of tumor cells with survival advantages. Lastly, these CTCs may also be the chemo-resistant portion that escapes systemic treatment, or acquire resistance during disease progression, as unique genomic alterations may impinge on pathways of tumor progression, such as NOTCH, or treatment resistance, such as PI3K [Kim C et al., 2018; Li et al., 2020; Litviakov NV et al., 2020]. Although these studies have limitations, represented by the low number of cases analysed, and the different disease settings investigated, their data suggest that CTCs are a useful tool for the study of intratumoural heterogeneity to identify rare aberrations within primary tumors that make them more able to spread; moreover, they may be useful in monitoring the development of treatment-resistant subclones as breast cancer progresses.

1.4 Scope of the thesis

To gain a perspective on the scope of the problem, we should consider that each year approximately 2 million new cases of breast cancer will be diagnosed worldwide, and about one-fourth of those patients will experience recurrence [Bray F et al., 2018]. The classic paradigm of breast cancer management, *i.e.,* decision making based on clinical and pathological parameters, has not significantly changed over decades [Curigliano G et al., 2017], and most clinical decisions are guided by balancing treatment benefits (reduced risk of local recurrence and increased survival), and drawbacks (toxicity and costs). Randomized clinical trials have shown that adjuvant treatment of resected invasive and non invasive, *i.e.,* ductal carcinoma in situ (DCIS), breast cancer are able to reduce the risk of relapse as compared to observation. However, up to half of patients receiving these therapies would in fact not have failed, and thus had been unnecessarily treated. Therefore, it is necessary to develop predictive and prognostic biomarkers

to minimize overtreatment, and to promptly identify patients at risk of relapse for case strategies [Appierto V et al., 2017].

HYPOTHESIS

Liquid biopsy may identify women with non invasive breast cancer, provide an early readout of the effect of systemic therapy, help to explore breast cancer heterogeneity, and guide treatment of recurrent disease.

SPECIFIC AIMS

- To assess ctDNA in newly diagnosed and recurrent cases of DCIS;
- To use ctDNA as a marker of response to neoadjuvant chemotherapy;
- To unravel differences in tissue and ctDNA baseline/post-treatment pairs;
- To identify mechanisms of tumor progression and resistance to therapy.

EXPERIMENTAL OUTLINE

ctDNA in women at high risk of invasive breast cancer

Women diagnosed with DCIS may relapse with invasive breast (IBC).

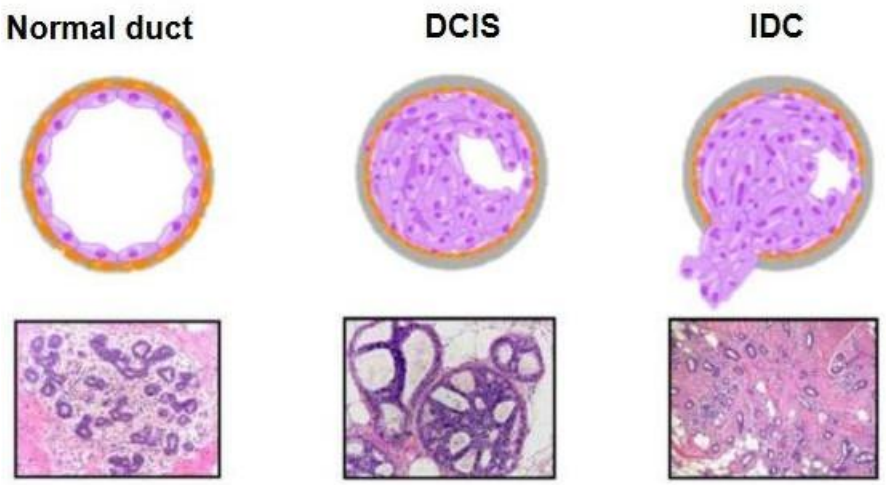


Figure 9. Evolution of DCIS from a normal duct to invasive cancer - Schematic cross sections of the duct (top row), and corresponding microscopic images (bottom row), modified from Casasent et al., 2017.

Free surgical margins and postoperative radiotherapy are the most important and significant prognostic factors for local disease control, but cannot reliably predict the risk of invasive recurrence after DCIS [reviewed by Salvatorelli L et al., 2020]. It has been reported that DCIS displays genomic heterogeneity similarly to that found in IBC. Likewise, the majority of synchronous DCIS and IBC exhibit similar genomic profiles [Shee K et al., 2019]. Nevertheless, recent analysis of mutational profiling of DCIS and matched adjacent IBC revealed intratumor genetic heterogeneity, and clonal selection [Martelotto LG et al., 2017].

Based on these considerations, we seek to evaluate ctDNA mutations in DCIS and matched primary tumor specimens in women who underwent treatment with curative intent and developed or not loco-regional relapse, with the final aim of paving the way for using ctDNA as a tool to stratify DCIS patients according to their risk of recurrence.

ctDNA to identify patients less likely to respond to systemic therapy

Neoadjuvant therapy is increasingly being used to treat breast patients with the expectation that treatment can reduce the size or extent of cancer before surgery, thus making the surgical procedures easier, and more likely to be curative.

Numerous studies have demonstrated that the burden of pathologically detected residual disease after neoadjuvant therapy is associated with long term prognosis, as the absence of invasive cancer in surgical specimens – *i.e.*, pathological complete response (pCR) – defines patients with favorable outcome [Cortazar P et al., 2014]. Unfortunately, only a minority of patients treated with neoadjuvant therapy achieve a pCR. The underlying reasons remain a major challenge.

This project aims to analyze the mutational profile of primary tumor, residual disease, and matched plasma samples to develop minimally invasive biomarkers to anticipate tumor response, to address the determinants of poor response at the molecular level, and to seek alternative strategies for unresponsive patients.

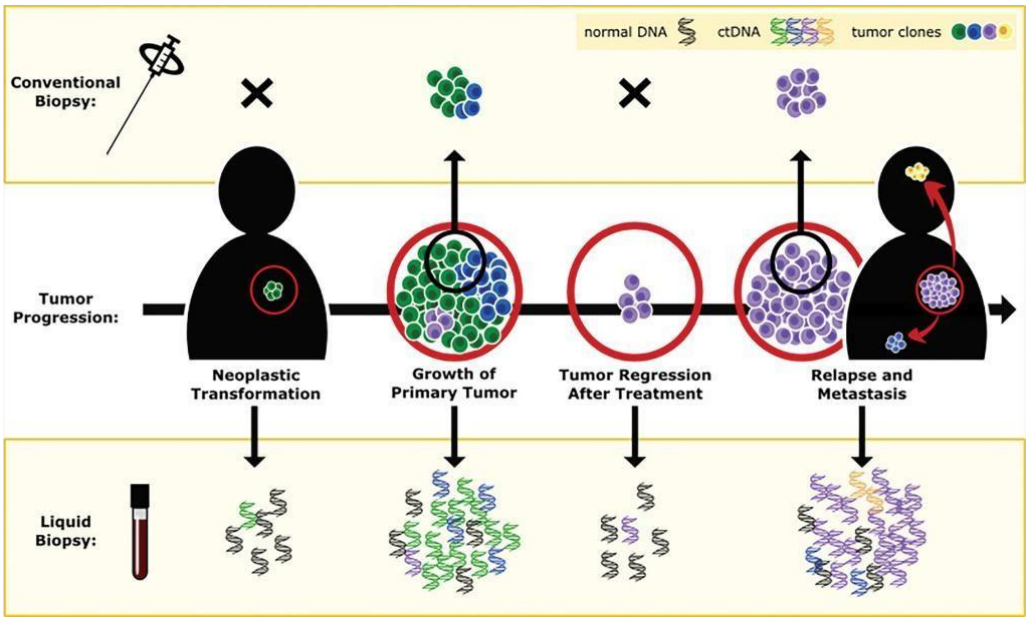


Figure 10. Clinical application of liquid biopsy in the neoadjuvant setting.
Adapted from https://www.genengnews.com

ctDNA to fully capture tumor heterogeneity

High-depth sequencing of plasma DNA can be used to interrogate the mutational profile of recurrent disease before overt relapse and, in cancer cases with intratumor heterogeneity, assess whether ctDNA reflected that of the original primary breast cancer or rather the subsequent recurrence. Therefore, this work aims at using direct ctDNA sequencing to evaluate tumor heterogeneity in recurrent BC cases.

CTCs to inform on the mechanisms of tumor progression

Finally we aimed to assess if the characterization of single CTCs could provide information on tumor molecular features and on response/resistance to treatment in patients with breast cancer. Therefore, CTCs were collected, enriched and analyzed in recurrent cases after neoadjuvant therapy, and single CTC molecular characterization was used to study the evolution of breast cancer disease in response to treatment.

CHAPTER 2.

MATERIALS AND METHODS

2.1 Patient information and biological samples collection

The work of this dissertation was conducted on different patient populations, namely patients diagnosed with ductal carcinoma in situ that subsequently relapsed or not (DCIS study); and patients with invasive triple negative breast cancer treated with preoperative chemotherapy, surgery with curative intent, and followed up over time (Neoadjuvant study).

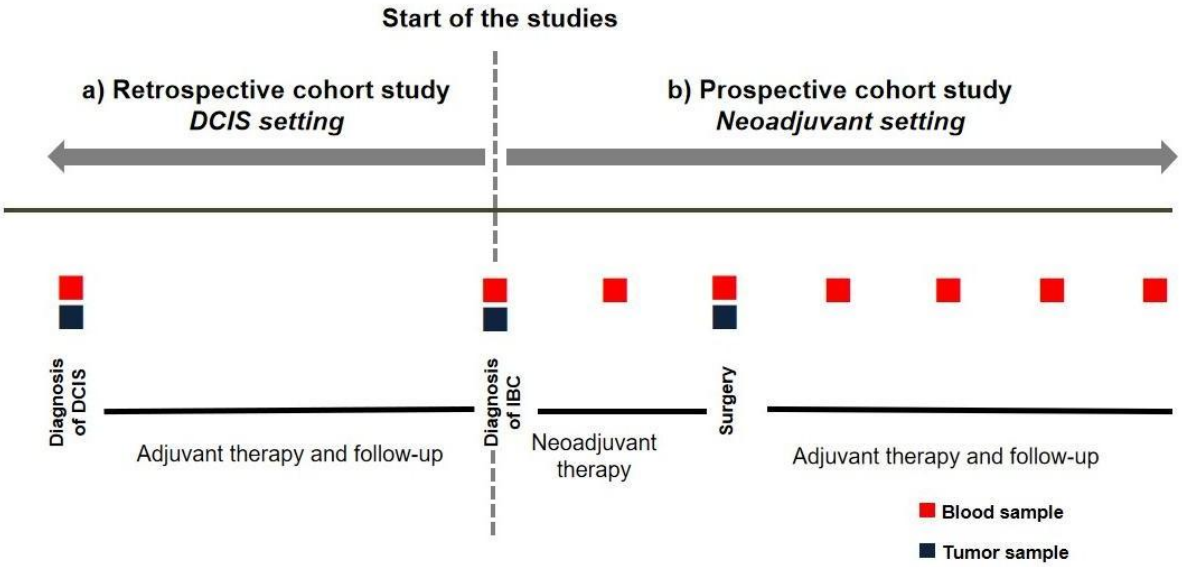


Figure 11. Schematic representation of DCIS and neoadjuvant study design

DCIS study - This was a retrospective study on newly diagnosed DCIS patients referred to Fondazione IRCCS Istituto Nazionale dei Tumori - Milano between January 2012 and June 2018, and treated with breast conserving surgery or mastectomy fulfilling the following criteria: availability of a formalin-fixed paraffin-embedded (FFPE) sample from the surgical specimen of the primary DCIS and matched invasive breast cancer in recurrent cases; no distinct invasive neoplastic components other than DCIS by revision of two dedicated pathologists; minimal tumor cellularity of 50%; negative margins (no tumor on ink); availability of plasma samples (minimum volume required 2.5 ml) at initial diagnosis and at recurrence; PBMC as a source of germline DNA. Patients with a previous or synchronous

contralateral breast cancer (DCIS or invasive) were excluded, as were patients with DCIS associated with lobular carcinoma in situ. Cases with and without recurrence were matched according to known prognostic variables, *i.e.*, age at diagnosis, tumor size and grade, hormone receptor and HER2 expression, type of surgery, and adjuvant treatment, *i.e.,* endocrine and/or radiotherapy.

Neoadjuvant study - This was composed of two portions: the first was a retrospective analysis of mutational profile of pre- and post treatment primary tumor samples; the second was a prospective study to test the clinical value of ctDNA and CTCs in the context of clinical practice. Patients included were all treated at Fondazione IRCCS Istituto Nazionale dei Tumori - Milano, for invasive breast cancer, fulfilling the following criteria: ER and PR negative status as defined by immunohistochemical staining of fewer than 1% of tumor cell nuclei; HER2 negative status as defined by IHC score 0-1 or IHC: 2+ with a negative chromogenic in situ hybridization (CISH) result; anthracycline/taxane-based neoadjuvant chemotherapy and breast and axillary surgery; availability of a FFPE sample from initial tumor biopsy and/or surgical specimen; minimal tumor cellularity of 50%; availability of plasma samples (minimum volume required 2.5 ml) at different time points, prior, during, after neoadjuvant therapy, and/or follow-up (minimum two drawings); PBMC as a source of germline DNA. Treatment efficacy was assessed by pCR as defined by the absence of invasive breast cancer in the breast and axillary nodes, and event free survival (EFS). EFS events were defined as post-surgery breast cancer relapse, second primary malignancy, or death for any cause. EFS time was measured from the date of NAC starting to the date of event, whichever occurred first; time was censored at the date of last follow-up for patients alive and without events. Stromal Tumor-Infiltrating Lymphocytes (sTILs) were assessed on a single hematoxylin-eosin stained slide, and scored according to pre-defined criteria [Salgado R et al., 2015]. All cases were independently

evaluated by two experienced breast cancer pathologists blinded to treatment data and follow-up. Breast surgery was performed after a maximum of six weeks from the conclusion of neoadjuvant chemotherapy.

Collection, use, and analysis of specimens in DCIS and neoadjuvant studies were approved by the institutional review board, and all patients signed an informed consent form to allow the research on tumor samples after the completion of diagnostic procedures, and on blood drawing. Due to the explorative nature of the study, no statistical hypothesis was postulated.

2.2 DNA extraction from FFPE tissues and buffy coat

Tumor DNA was isolated from four sections of primary tumor FFPE tissues (10 µm thick slides with tumor cellularity ≥50%) using the GeneRead DNA FFPE Kit (Qiagen, Valencia, CA) according to manufacturer's instruction. Buffy coat DNA (for germ line DNA) was extracted using the QIAamp DNA Mini Kit (Qiagen) according to manufacturer's instruction. The quality and the quantity of isolated DNA were assayed using the Agilent BioAnalyzer DNA High Sensitivity kit (Agilent Technologies, Santa Clara, CA) and the Qubit dsDNA HS kit (Life Technologies), respectively.

2.3 Tissue targeted Next Generation Sequencing

Sequencing of hotspot regions of 50 genes contained in the Ion AmpliSeq Cancer Hotspot Panel v2 (Thermo Fisher), *i.e.,* HS panel, was performed by employing Ion S5 XL System (Thermo Fisher).

Patients with enough DNA samples were also analyzed by sequencing of 409 all-exon genes contained in the Ion AmpliSeq Comprehensive Cancer Panel, *i.e.,* CCP, with germ line DNA extracted from PBMC belonging to the same patients using the same Thermo Fisher platform.

Ten ng or 40 ng of genomic DNA were used to amplify 50 gene hotspot regions using the Ion AmpliSeq Library Kit2.0 (Thermo Fisher) or the 409 all-exon genes

contained in the Comprehensive Cancer Panel (Thermo Fisher) according to the manufacturer's manual (MAN0006735 rev 5.0). Amplicons were ligated to P1 and Barcode adapters using DNA Ligase. Barcoded libraries were purified using AMPure Beads XP (Beckman Coulter, Pasadena, CA), and PCR-amplified for a total of five cycles. After a second round of purification with AMPure Beads, the amplified libraries were sized, their quality was assessed using the Agilent BioAnalyzer DNA High Sensitivity kit (Agilent Technologies, Santa Clara, CA), and DNA was quantified using the Qubit dsDNA HS kit (Thermo Fisher). In particular, due to the FFPE origin of the DNA, dsDNA fragments with a length range of 50-150 bp were selected for both panels. Emulsion PCR and sample enrichment were performed using the Ion One touch 2 instrument according to the manufacturer's instruction. Briefly, an input concentration of DNA library obtained with the first amplification step was added to the emulsion PCR master mix, and the Ion sphere particles (ISPs) and a double phase (oil/water) PCR was performed. Template preparation was performed using the Ion Chef instrument according to the manufacturer's instruction. 530 and 540 chips were used to sequence samples on Ion S5 XL next-generation sequencing system using Ion S5 XL sequencing kits and following the manufacturer's instructions.

2.3.1 Sequencing data analysis

Sequencing data were mapped to the human reference genome (hg19) using the TMAP algorithm implemented in the Torrent Suite software version 4.4.3 with default parameters. Tissue samples processed with Ion AmpliSeq Cancer Hotspot Panel v2 and Ion AmpliSeq Comprehensive Cancer Panel were analyzed with AmpliSeq CHPv2 single sample workflow and AmpliSeq CCP w1.1 – Tumor-Normal pair workflow of Ion Reporter version 5.10, respectively. To reduce the number of false positive calls, and obtain a list of confident somatic mutations we discarded variants passing at least one of the following filters: (i)

phred quality score< 30; (ii) strand bias p-value< 0.01; (iii) number of variant-supporting reads< 7; (iv) variant allelic frequency (VAF)< 5% and depth< 500 X; (v) allele frequency≥ 1% in the 1000 Genomes Project European population, or ExAC non-Finnish European population; (vi) variant allelic frequency (VAF) in normal sample≥ 10% (only for CCP); (vii) depth in normal sample< 10X. The confirmation of tumor somatic mutations was assessed using *ad hoc* ddPCR assays. When ddPCR assays wet-validated by the manufacturer were not available, custom mutation-specific ddPCR assays were designed using the Thermo Fisher Scientific custom SNP genotyping assay tool or Bio-Rad Mutation Detection Assay online design tool. The specific assays are detailed in the table below.

Table 2 Mutation-specific ddPCR assays designed for each case

PIK3CA p.His1047Arg	GCCTTAGATAAAACTGAGCAAGAGGCTT TGGAGTATTTCATGAAACAAATGAATGAT GCAC**[A/G]**TCATGGTGGCTGGACAACAA AAATGGATTGGATCTTCCACACAATTAAA CAGCATGCATTG	Bio-Rad Laboratories
PIK3CA p.Glu542Ala	GAATTAAGGGAAAATGACAAAGAACAGC TCAAAGCAATTTCTACACGAGATCCTCTC TCTG**[A/C]**AATCACTGAGCAGGAGAAAGA TTTTCTATGGAGTCACAGGTAAGTGCTAA AATGGAGATTC	Bio-Rad Laboratories
TP53 p.Leu252fs	CAGGCCAGTGTGCAGGGTGGCAAGTGG CTCCTGACCTGGAGTCTTCCAGTGTGAT GATGGT**[GAG/*]**GATGGGCCTCCGGTTC ATGCCGCCCATGCAGGAACTGTTACACA TGTAGTTGTAGTGGATG	Bio-Rad Laboratories
TNK2 p.Arg1020Trp	ATCTTGTCTGCTGGCCGGCCCGCCTCTG GCCCATCGCCAGGGCAGCCCCTCTGTG GCAGCC**[G/A]**GCAGTGGCCCTCGGGGG TGGTGGCCGGGCCCCTGGGTTGCTGTT GTTGGTGGAGAAGTTG	Bio-Rad Laboratories
TP53 p.Gln136STOP	CGGGTGCCGGGCGGGGGTGTGGAATCA ACCCACAGCTGCACAGGGCAGGTCTTG GCCAGTTGG**[C/T]**AAAACATCTTGTTGAG GGCAGGGGAGTACTGTAGGAAGAGGAA GGAGACAGAGTTGAAA	Bio-Rad Laboratories
TP53 p.Cys238Phe	TTCCAGTGTGATGATGGTGAGGATGGGC CTCCGGTTCATGCCGCCCATGCAGGAAC TGTTA**[C/A]**ACATGTAGTTGTAGTGGATG GTGGTACAGTCAGAGCCAACCTAGGAGA TAACACAGGCCCA	Bio-Rad Laboratories
PIK3CA p.His1047Arg	GCCTTAGATAAAACTGAGCAAGAGGCTT TGGAGTATTTCATGAAACAAATGAATGAT GCAC**[A/G]**TCATGGTGGCTGGACAACAA AAATGGATTGGATCTTCCACACAATTAAA CAGCATGCATTG	Bio-Rad Laboratories

TP53 p.Cys176Ser	CAGCCCCAGCTGCTCACCATCGCTATCT GAGCAGCGCTCATGGTGGGGGC**[A/T]**GC GCCTCACAACCTCCGTCATGTGCTGTGA CTGCTTGTAGATGGCCATGG	Thermo Fisher Scientific
PIK3CA p.His1047Arg	GCCTTAGATAAAACTGAGCAAGAGGCTT TGGAGTATTTCATGAAACAAATGAATGAT GCAC**[A/G]**TCATGGTGGCTGGACAACAA AAATGGATTGGATCTTCCACACAATTAAA CAGCATGCATTG	Bio-Rad Laboratories
SF3B1 p.His1210Arg	AACTGCCTGAATTACATGAGGAGATGTC TCAAATACATTGGGCCATACATAGTTCAA CAAG**[T/C]**GATTCAGCGAATCTTCACAAC CAAATCCATAAACCCCAAGTGACATGTG CTGTACCACTGC	Bio-Rad Laboratories
ATM p.Ile1441Met	CCATATGTGAGCAAGCAGCTGAAACAAA TAATGTTTATAAGAAGCACAGAATTCTTA AAAT**[A/G]**TATCACCTGTTTGTTAGTTTAT TACTGAAAGATATAAAAAGTGGCTTAGGA GGAGCTTGGG	Bio-Rad Laboratories
TP53 p.Arg213Gln	CCCAGAGACCCCAGTTGCAAACCAGACC TCAGGCGGCTCATAGGGCACCACCACA CTATGT**[C/T]**GAAAAGTGTTTCTGTCATC CAAATACTCCACACGCAAATTTCCTTCCA CTCGGATAAGATG	Bio-Rad Laboratories
TP53 p.His179Gln	GCAACCAGCCCTGTCGTCTCTCCAGCCC CAGCTGCTCACCATCGCTATCTGAGCAG CGCTCA**[T/A]**GGTGGGGGCAGCGCCTCA CAACCTCCGTCATGTGCTGTGACTGCTT GTAGATGGCCATGG	Bio-Rad Laboratories
TP53 p.Gly266Glu	CAGCTGCACAGGGCAGGTCTTGGCCAG TTGGCAAAACATCTTGTTGAGGGCAGGG GAGTAC**[T/G]**GTAGGAAGAGGAAGGAGA CAGAGTTGAAAGTCAGGGCACAAGTGAA CAGATAAAGCAACTG	Bio-Rad Laboratories
TP53 p.Lys132Asn	GCGGGGGTGTGGAATCAACCCACAGCT GCACAGGGCAGGTCTTGGCCAGTTGGC AAAACAT**[C/G]**TTGTTGAGGGCAGGGGA GTACTGTAGGAAGAGGAAGGAGACAGA GTTGAAAGTCAGGGCAC	Bio-Rad Laboratories
FGFR3 p.Phe384Leu	GGTGGAGGCTGACGAGGCGGGCAGTGT GTATGCAGGCATCCTCAGCTACGGGGTG GGCTTC**[T/C]**TCCTGTTCATCCTGGTGGT GGCGGCTGTGACGCTCTGCCGCCTGCG CAGCCCCCCCAAGAA	Bio-Rad Laboratories
TP53 p.Arg273His	CCCTTTCTTGCGGAGATTCTCTTCCTCTG TGCGCCGGTCTCTCCCAGGACAGGCAC AAACA**[C/T]**GCACCTCAAAGCTGTTCCGT CCCAGTAGATTACCACTACTCAGGATAG GAAAAGAGAAGCA	Bio-Rad Laboratories
TP53 p.Arg213Ter	CTGAAAATGTTTCCTGACTCAGAGGGGG CTCGACGCTAGGATCTGACTGC**[G/A]**GC TCCTCCATGGCAGTGACCCGGAAGGCA GTCTGGCTGCTGCAAGAGGAA	Thermo Fisher Scientific
TP53 p.Arg248Trp	GGGTGGCAAGTGGCTCCTGACCTGGAG TCTTCCAGTGTGATGATGGTGAGGATGG GCCTC**[C/T]**GGTTCATGCCGCCCATGCA GGAACTGTTACACATGTAGTTGTAGTGG ATGGTGGTACAGTCA	Bio-Rad Laboratories
TP53 p.Pro177_Cys182del	GGACCCTGGGCAACCAGCCCTGTCGTC TCTCCAGCCCCAGCTGCTCACCATCGCT ATCTGA**[GCAGCGCTCATGGTGGGGG/*]** CAGCGCCTCACAACCTCCGTCATGTGCT GTGACTGCTTGTAGATGGCCATGGCGCG GACGC	Bio-Rad Laboratories
TP53 p.His168Arg	TATCTGAGCAGCGCTCATGGTGGGGGC AGCGCCTCACAACCTCCGTCATG**[T/C]**G CTGTGACTGCTTGTAGATGGCCATGGCG CGGACGCGGGTGCCGGGCGGG	Thermo Fisher Scientific

TP53 p.His193Pro	TCGAAAAGTGTTTCTGTCATCCAAATACT CCACACGCAAATTTCCTTCCACTCGGAT AAG**[A/C]**TGCTGAGGAGGGGCCAGACCT AAGAGCAATCAGTGAGGAATCAGAGGCC TGGGGACCCTGGG	Bio-Rad Laboratories
TP53 p.Ala159Val	ATGGTGGGGGCAGCGCCTCACAACCTC CGTCATGTGCTGTGACTGCTTGTAGATG GCCATG**[G/A]**CGCGGACGCGGGTGCCG GGCGGGGGTGTGGAATCAACCCACAGC TGCACAGGGCAGGTCTT	Bio-Rad Laboratories
TP53 p.Tyr107*	CAAGTCACAGACTTGGCTGTCCCAGAAT GCAAGAAGCCCAGACGGAAACC**[G/T]**TA GCTGCCCTGGTAGGTTTTCTGGGAAGGG ACAGAAGATGACAGGGGCCA	Thermo Fisher Scientific
EGFR p.Asp587Asn	AACATTTTTCTCCACCTTGGTGCAGGGA CCAGACAACTGTATCCAGTGTGCCCACT ACATT**[G/A]**ACGGCCCCCACTGCGTCAA GACCTGCCCGGCAGGAGTCATGGGAGA AAACAACACCCTGGT	Bio-Rad Laboratories
AKT1 p.Gln17Lys	GCTTTCCAACCTAGGAAGGCAGGGGAGT AGGGCCAGGAAGGGGCTGAGGTCACTC ACCTGG**[A/G]**GTGAGCCCTGCTCCCCCC TGGCTCCTTCCCAGCCTGGGCATCCTTG AGTTCCAAGGCCTCA	Bio-Rad Laboratories
TP53 p.Glu294fs	TCGTCTCTCCAGCCCCAGCTGCTCACCA TCGCTATCTGAGCAGCGCTCATGGTGGG GGCAG**[C/*]**GCCTCACAACCTCCGTCATG TGCTGTGACTGCTTGTAGATGGCCATGG CGCGGACGCGGGT	Bio-Rad Laboratories
TP53 p.Cys176Ser	CAGCCCCAGCTGCTCACCATCGCTATCT GAGCAGCGCTCATGGTGGGGGC**[A/T]**GC GCCTCACAACCTCCGTCATGTGCTGTGA CTGCTTGTAGATGGCCATGG	Thermo Fisher Scientific

In the table are listed the mutation selected for each patient and the sequences used to design ddPCR assays on the online tools according to manufacturer instructions. Nucleotide changes are marked in red and enclosed in square brackets.

2.4 DNA extraction from plasma samples

Blood was separated into plasma aliquots within 2 hour from blood drawn by centrifugation at 4°C and stored at −80°C until assayed; no thawing accident occurred during storage. A second full speed centrifugation was performed after plasma thawing before cfDNA extraction using the QIAamp Circuling Nucleic Acid Kit (Qiagen) according to manufacturer's instruction. DNA was eluted in 70 μl of sterile distilled water.

2.5 ddPCR analysis of tissue and plasma DNA

The confirmation of tumor somatic mutations in tissue and the presence of ctDNA in plasma were assessed using *ad hoc* ddPCR assays. One single mutation was selected per patient. When multiple mutations were identified in a single tumor

sample, the mutation with the highest allelic fraction or the one with an already manufactured assay available in our laboratory was selected. Custom mutation-specific ddPCR assays were designed using the Thermo Fisher custom SNP genotyping assay tool or alternatively Bio-Rad Mutation Detection Assay online design tool.

ddPCR was performed on the QX200 ddPCR™ system (Bio-Rad) in a total reaction volume of 21 µl including 1x ddPCR supermix for probes without dUTP (Bio-Rad), 9 µM each primer, 5 µM probe, and 7 µl of pre-amplified plasma DNA. Reactions were partitioned into a median of ~14,000 droplets per well in a QX200 droplet generator according to the manufacturer's instructions. Droplets from each sample were transferred into a 96-well PCR plate for amplification on a C1000 Touch™ Thermal Cycler (Bio-Rad), incubating the plates at 95°C for 10 min followed by 40 cycles of 94°C for 30 s and 55°C for 60 s, then 10 min incubation at 98°C. Plates were read on a Bio-Rad QX200 droplet reader (Bio-Rad) using QuantaSoft program (Bio-Rad). All the experiments were performed in triplicate including in every run three negative controls: i) no DNA template (NTC); ii) wild type (wt) genome (Promega); iii) cfDNA extracted from a pool of healthy donors. A sample was considered positive only when three or more FAM-positive droplets were detected in each replicate. Samples displaying only double-positive droplets containing both wt and mutant DNA templates were considered as false positives. Variant allele fraction (VAF) was calculated as follows: VAF (%) = [mutant copies / (wild-type + mutant copies)] x 100.

2.6 Pre-amplification of plasma DNA

Prior to ddPCR analysis, plasma DNA was pre-enriched by amplification using TaqMan® PreAmp Master Mix Kit (Thermo Fisher). Sample elution volume was reduced by Eppendorf Concentrator 5301 (Epperdorf) to 14 µl. Pre-amplification reaction was performed in a volume of 10 µl containing 4 µl of DNA template, 5 µl

of pre-amplification mastermix, and 1 µl of the same specific primers and probes designed for dPCR (at a final dilution of 0.05x). The amplification reaction was initiated by incubation of samples at 95°C for 10 minutes followed by 12 cycles of 95°C for 15 seconds, 60°C for 4 minutes. The pre-amplified PCR products were then diluted 1:100-1:500 and 7 µl of dilutions were used to perform ddPCR.

2.7 Plasma direct sequencing

Circulating-free DNA (cfDNA) sequencing was performed using Oncomine Pan-Cancer Cell-Free Assay (Thermo Fisher) to identify genomic alterations among 272 amplicons from 52 cancer-related genes. The assay minimizes false positives by removing randomly incorporated errors through unique molecular tags attached to the gene-specific primers. Libraries were prepared following manufacturer's instructions.

Library purifications were carried out with AMPure XP magnetic beads (Beckman Coulter, Inc., Brea, CA, USA), and library quantification was performed using the Ion Library TaqMan Quantitation kit. For sequencing, libraries were diluted to a final concentration of 100 pM and after pooling, the concentration was adjusted to 50 pM. Template preparation and chip loading were carried out on Ion Chef System (Thermo Fisher, Palo Alto, CA, USA). Eight samples were loaded on Ion 550 chips and sequenced in an Ion Torrent S5XL Sequencer (Thermo Fisher, Palo Alto, CA, USA).

2.7.1 Sequencing data analysis

Raw sequencing data were analyzed using Torrent Suite Software (v5.12). Sequencing coverage was analyzed using the Coverage Analysis (v.5.10.0.3) plug-in. Raw reads were aligned to the human reference genome hg19.

Plasma samples processed with Oncomine™ Pan-Cancer Cell-Free Assay were analyzed with Oncomine TagSeq Pan-Cancer Liquid Biopsy w2.2 Single Sample workflow of Ion Reporter version 5.12 with default setting of germline low-

stringency parameters (minimal variant frequency of 0.1%, minimum variant quality of 10, maximum strand bias of 0.98, and minimum variant score of 10).

2.8 Circulating Tumor Cells

2.8.1 Enrichment and identification

Blood samples (10 ml) collected in K2EDTA tubes were subjected to CTC enrichment with Parsortix® (Angle plc, Guildford, UK) within 1 h from blood draw. Enriched cells were harvested according to manufacturer's instructions and fixed for 20 min at room temperature (RT) with 2% paraformaldehyde. Fixed samples were stained immediately or within 24 h from enrichments. Mean Parsortix's recovery rates (previously investigated using 5 different cell lines spiked in healthy donor blood at final concentrations ranging between 25–50 cells/10 mL) were 81% (range 75–90%, depending on the cell line).

Fixed samples were fluorescently stained with phycoerythrin (PE)-labeled antibodies against epithelial markers EpCAM (clone HEA-125, Miltenyi Biotec, Bergisch Gladbach, Germany, working dilution 1:11 for 10 min at 4° C), cytokeratins (pan cytokeratin clone C11, Abcam, San Francisco, CA, USA, and pan cytokeratin clone AE1/AE3, NSJ Bioreagents, San Diego, CA, USA, working dilution 1:10 for 10 min at RT), and EGFR (clone 423103, SantaCruz Biotechnology, Dallas, TX, USA, working dilution 1:11 for 10 min at 4° C), and with allophycocyanin (APC)-labeled antibodies recognizing leukocytes and monocytes: CD45 (clone 5B1, Miltenyi Biotec, working dilution 1:11 for 10 min at 4° C),

CD14 (clone M5E2, BD Biosciences Pharmigen, San Diego, CA, USA, working dilution 1:20 for 10 min at 4° C), and CD16 (clone 3G8, BD Biosciences Pharmigen, working dilution 1:20 for 10 min at 4° C). Nuclei were stained with 1 g/mL Hoechst 33342 (Sigma-Aldrich, Saint Louis, MI, USA) for 5 min at RT. The labeling procedure did not lead to significant loss of cells. Labeled cells were analyzed using the DEPArray™ (Menarini Silicon Biosystems, Bologna, Italy) within 2 days

from staining to visualize and recover single cells manually selected based on fluorescence labeling and morphology.

Selected single epithelial or double-negative (PE-ve/APC-ve) cells were recovered for downstream molecular analyses. CTC enrichments by Parsortix® lasted about 3 h, fixation lasted 20 min, and the cell selection and recovery process with the DEAPrray lasted about 3–4 h depending on the number of cells recovered from the patient 4.5.

2.8.2 Molecular characterization

Recovered single cells and pools of white blood cells (WBC) were subjected to whole genome amplification employing the Ampli1 WGA kit (Menarini Silicon Biosystems). Amplified DNA quality was checked with the Ampli1— QC kit (Silicon Biosystems), and a low-pass whole genome sequencing (lpWGS) to detect copy number aberrations was performed using the Ampli1-Low Pass kit (Menarini Silicon Biosystems) for barcoded libraries preparation, followed by sequencing with the IonTorrent Ion S5-system (Thermo Fisher, Waltham, MA, USA), using the Ion530 chip according to manufacturer's instructions.

2.8.3 Sequencing data analysis

WGS sequences were aligned to the human reference genome (hg19) using tmap aligner tool on Torrent_Suite 5.4.0. CNAs were predicted by using QDNAseq 11.0 with the following settings: minMapq=37, window=500kb. "Gain" and "loss" calls were filtered out by residual (> 4 SD) and black list regions reported in ENCODE database [Scheinin et al., 2014]. Segmented copy number data of each sample were extracted starting from logRatio value and used for downstream analysis performed by R software [https://www.R-project.org/.]. Functional enrichment analysis of gene lists extracted from most altered regions between CTCs for using Gene Ontology (GO) biological process terms and KEGG pathways was performed using the ClusterProfiler Bioconductor package [Yu G et al., 2012].

Enrichment with *p*-value< 0.05 was considered for KEGG pathways or GO terms, respectively.

2.9 Development of a protocol for ctDNA and CTC analysis

The standard paradigm for ctDNA analysis involves the identification of somatic genomic alterations in tumor tissue, the design of PCR-based assays to detect and confirm these alterations, and the application of these assays to quantify accurately ctDNA in plasma (Figure 12).

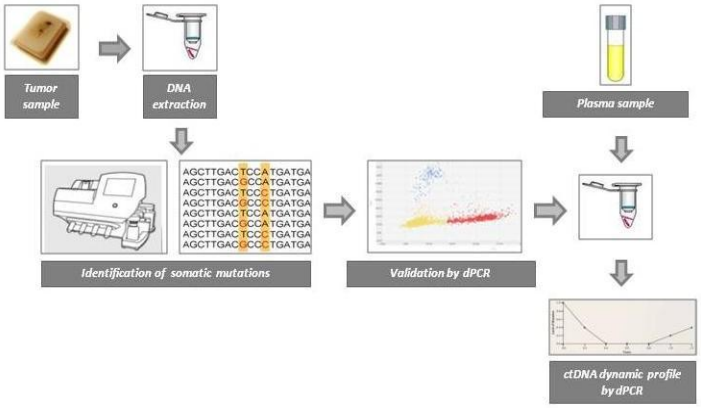


Figure 12. Workflow for tissue and plasma mutations analysis

As mentioned, the main challenge for the use of liquid biopsy is the detection of a minor allelic fraction of mutant DNA on fragmented cfDNA. This requires the implementation of standardized operating procedures to ensure reliable results. To this end, the following steps were fine-tuned for cfDNA extraction, and ctDNA analysis.

2.9.1 Optimization of pre-analytical aspects

1. Source - ctDNA was obtained from plasma rather than from serum, according to the joint American Society of Clinical Oncology (ASCO), and College of American Pathologists (CAP) [Merker JD et al., 2018]. Indeed, leukocyte-derived normal DNA, which could dilute the ctDNA fraction, is lower in plasma fraction.

2. Tubes - Blood samples were collected in EDTA, and processed within a maximum of two hours from blood drawing to avoid the risk of leukocyte lysis, which could dilute the ctDNA fraction.

3. Plasma extraction - A first centrifugation at 1700 × g for 10 min was performed to remove cells. After supernatant harvesting, plasma was collected and centrifuged again at 14,000 × g (high-speed centrifugation) for 10 min to remove all cellular-derived contaminants.

4. Storage - Plasma was aliquoted into single-use fractions and stored at -80 °C.

5. Isolation - Studies directly comparing different ctDNA extraction methods are lacking. However, a recent experience has found comparable ctDNA yields when using either magnetic beads or silica-based membranes [Carpi FM et al., 2011].
The spin-column based method - currently used at our labs - takes advantage of the fact that DNA binds to silica. The sample containing DNA is added to a column containing a silica membrane and chaotropic salts. The chaotropic salts disrupt the hydrogen bonding between strands and facilitate binding of the DNA to silica by causing the nucleic acids to become hydrophobic. This exposes the phosphate residues so they are available for adsorption. The DNA binds to the silica, while the rest of the solution is washed out using ethanol to remove chaotropic salts and other unnecessary constituents (Figure 13). The DNA can then be re-hydrated with aqueous low salt solutions allowing for elution of the DNA from the beads. This method yields high-quality which can be used for subsequent

analysis.

6. Quantification - Nucleic acid quantification was performed in a cuvette spectrophotometer, where the monochromator optical system provided light at 260 nm, the absorbance peak for DNA. In addition, nucleic acid samples were also measured at 280 nm, which is the absorbance peak for protein. The ratio of the 60

nm and 280 nm measurements provided a determination of the purity of the nucleic acid, with a ratio near 2 indicating a highly pure nucleic acid sample.

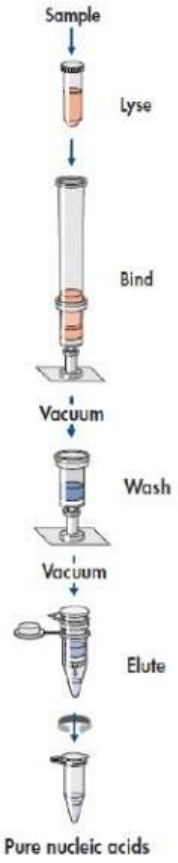


Figure 13. Spin-column based method for DNA extraction

2.9.2 Optimization of analytical aspects

In theory, all the molecular techniques allowing to detect a mutation can be used. However, dealing with minimal fraction of ctDNA requires highly sensitive techniques. ddPCR approach allows absolute quantification with a sensitivity as high as 0.001%. ddPCR consists of the emulsification of extracted ctDNA, master mix, and Taqman reagents in droplets in order to isolate a single DNA molecule in each droplet. Two colors can be distinguished, the target mutant codon of interest and the conserved sequence, defining negative or positive droplets, respectively. As ddPCR system is prone to variability in the number of compartments that are

generated and/or accepted into the analysis, this essentially corresponds to a form of sub-sampling which may in turn add variation to the quantification.

This is especially true for reactions with targets at very low abundance. It has been reported that a minimum compartmentalisation of 30% (7,000 droplets) is required to quantify down to 1%, and this percentage reaches 50% up to 12,000 to quantify down to 0.5%. In addition, limited yields of cfDNA may also decrease the specificity of mutation detection as a very small amount of PCR-generated errors can cause considerable noise, even in samples from healthy individuals [Lievens A et al., 2016; Ono Y et al., 2017].

To circumvent the limitations posed by the limited volume of plasma available and maximize ctDNA detection, we decided to perform a targeted enrichment step of the cfDNA prior to ddPCR analysis (Figure 14).

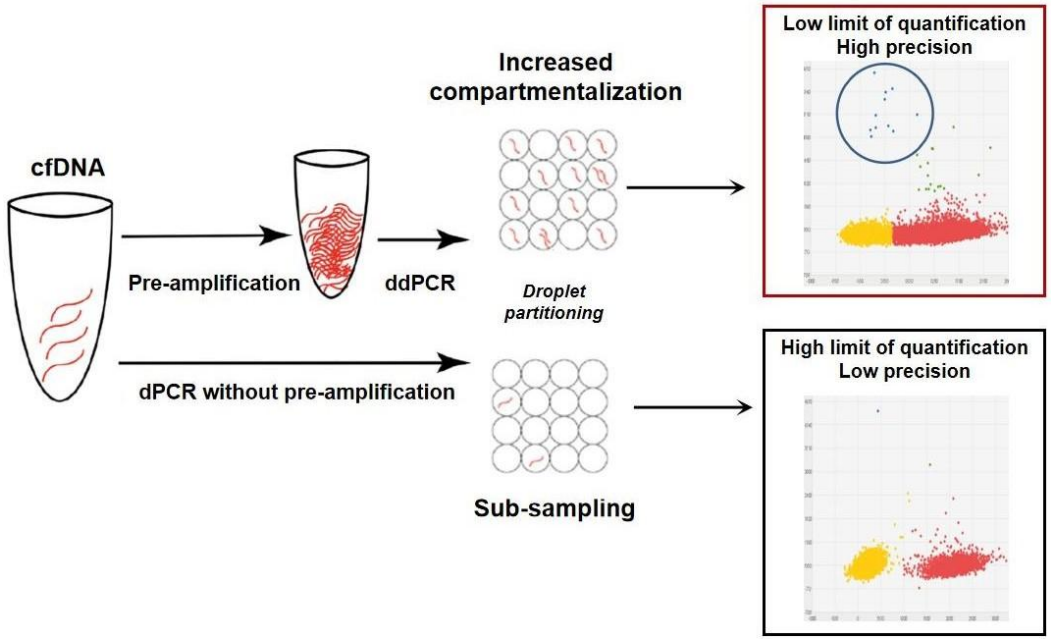


Figure 14. Output of pre-amplification to overcome sub-sampling Blue circle: ctDNA identified in pre-amplified samples.

Specifically, pre-amplification (12 cycles) of the regions of interest was performed by conventional PCR with the same mutation detection assays used for the

subsequent ddPCR analysis. The pre-amplified samples were diluted five fold in PCR grade water prior to further manipulation.The following controls were set up for each pre-amplification reaction: (i) a wild type genome sample as a negative control, to verify the specificity of the assay; (ii) a sample without template DNA to test for contamination; and (iii) the DNA sample extracted from the tumor to check that the mutation frequency was not altered by pre-amplification (Figure 15). Pre-amplified products were stored at -20°C, and analysed by ddPCR.

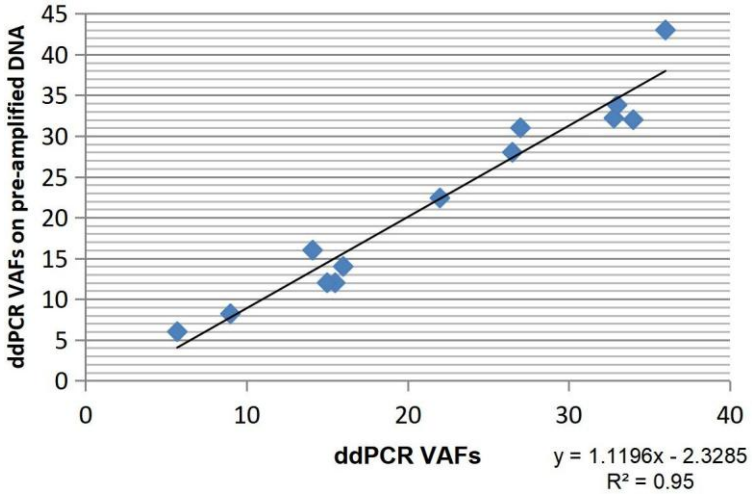


Figure 15. VAFs estimated by NGS, ddPCR with or without pre-amplification. Somatic primary tumor mutations tested in ddPCR with (Y axis) or without (X axis) pre-amplification showed a high level of agreement in estimated VAFs (R^2=0.95).

2.9.3 Workflow of targeted sequencing of plasma cell-free DNA

Genomic profiling of plasma samples by direct targeted NGS was performed by Oncomine Pan-Cancer Cell-Free Assay, covering DNA target regions containing hotspots variants of 52 relevant genes in pathogenetic, and resistance molecular mechanisms in breast cancer.

1. ***Enrichment process*** - This NGS assay applied unique molecular identifiers (UMIs) to improve the sensitivity by decreasing the amount of sequencing artifacts. With the recommended input of 20 ng cfDNA, the use of UMIs enabled a limit of detection (LoD) as low as 0.1%.

2. ***Library preparation*** - Library quantities were determined by qPCR and then diluted to a final concentration of 50 pM. Next, sample barcoded libraries were pooled together for template preparation, loaded onto a chip and sequenced on a Ion Torrent S5XL Sequencer (ThermoFisher Scientifific).

3. ***Annotation*** - Analysis of variants were locally carried out using preconfigured parameter settings for liquid biopsy application. Raw sequencing data were analyzed using Torrent Suite Software (v5.12). Sequencing coverage was analyzed using the Coverage Analysis (v.5.10.0.3) plug-in. The manufacturer recommends a Median Read Coverage (MedReadCov) >25,000, and Median Molecular Coverage (MedMolCov) >2500 to detect a variant with a VAF of 0.1%. Raw reads were aligned to the human reference genome hg19. Plasma samples were analyzed with Oncomine TagSeq Pan-Cancer Liquid Biopsy w2.2 Single Sample workflow of Ion Reporter version 5.12 with default setting of germline low-stringency parameters.

4. ***Reproducibility*** - ddPCR was tested on cfDNA samples to report concordance.

2.9.4 CTC analysis in house pipeline

This study aimed at assessing the role of molecular characterization of single CTC as a strategy to selectively distinguish clones responsible for treatment resistance/ metastatic progression. To such a purpose, the recently published protocol developed in our lab [Reduzzi C et al., 2017] (Figure 16) was applied according to the following steps:

1. Enrichment by the Parsortix® Cell Separation System which is an epitope independent process, and consequently agnostic to cellular genotype or (immuno)phenotype.

2. Labeling with antibody cocktails against epithelial and leukocyte markers.

3. Visualization through DEPArray™ (Menarini Silicon Biosystems, Bologna, Italy).

4. Recovery of epithelial and double-negative cells.

5. Downstream molecular analyses.

CTC enrichment by Parsortix® lasted about 3 hours, fixation lasted 20 min, and the cell selection and recovery process with the DEPArray™ lasted about 3-4 hours depending on the number of collected cells.

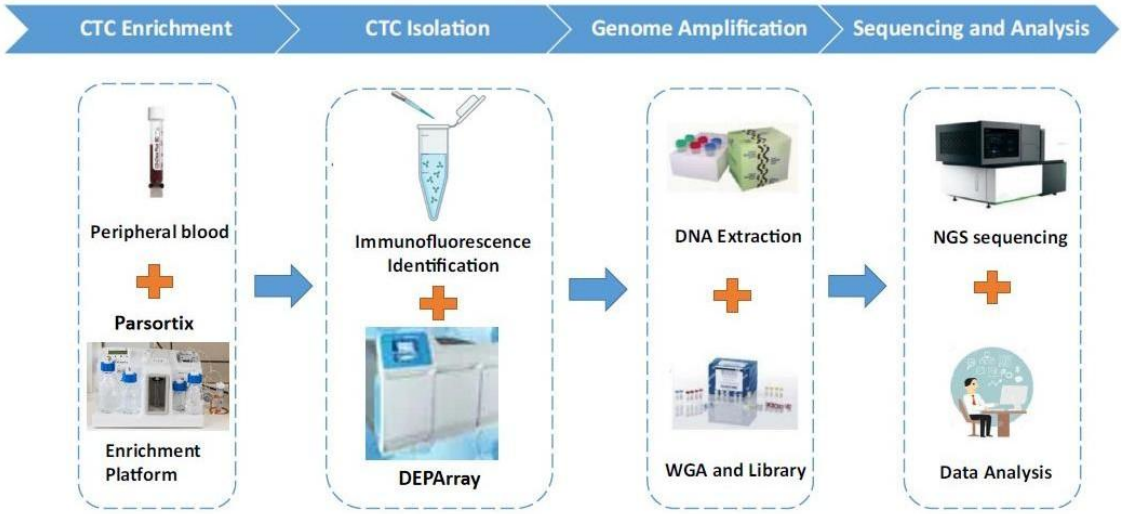


Figure 16. The workflow of Circulating Tumor Cell analysis

At the end of the DEPArray™ processing, single CTCs were collected in 0.2 ml PCR-tubes in a drop (around 20 µl) of SB115 buffer. Before starting the WGA it was necessary to perform a volume reduction procedure to leave the cell in 1 µl of PBS. After volume reduction, cells were lysed to allow WGA, which was followed by a quality assessment check through a PCR-based approach. This way, we assigned a genome integrity index (GII) to each amplified DNA. The GII value could range from 0 to 4: 0 indicated a failure in the WGA, probably due to the loss of the cell during volume reduction; 1-2 corresponded to a low-quality sample; 3-4 was indicative of high-quality amplified DNA. After quality check, DNA was sequenced by low pass WGS to confirm or exclude the malignant nature of each cell based on CNA profile.

CHAPTER 3.

RESULTS

3.1 DCIS study

3.1.1 Patient cohort

The search for DCIS cases with available tumor tissue and perioperative blood samples stored at the Institutional bio-bank between October 2012 and July 2018 retrieved a total of 160 cases. Among these, 10 (6.25%) secondary events occurred, *i.e.*, 7 invasive ipsilateral breast tumor recurrences, and 3 contralateral breast cancer. This figure is consistent with literature data, considering both the median follow-up of this case series, and a yearly rate≤ 1% of breast cancer recurrence [EBCTCG, 2010].

We initially matched 1:2 the 7 cases of invasive ipsilateral breast tumor recurrence (IBTR) with 14 controls according to age, primary tumor (T) size, grading, presence of comedotype, hormone receptor status. Recurrent patients had a median age of 43 years (range 28-69), and a median T size of 7 mm (range 2-15). Grade III and comedotype occurred in 3 and 2 cases, respectively, all had a positive hormone receptor status. Unfortunately, tumor tissue from 2 of these recurrent cases and 1 control were not received from the pathological anatomy service, which was to be expected as this was a study carried out on the leftover of the diagnostic material and not collected *ad hoc*. We obtained a DIN value >2 by the TapeStation 4200 (Agilent) in 12 of 18 evaluable cases, for an overall adequate DNA rate of 67%. This result is little in line with laboratory experience with nucleic acid studies performed on paraffin-embedded tumor tissue samples typically yielding optimal DIN values in about 20% more cases examined. We next proceeded to NGS of primary tumors. Low coverage hampered the analysis of one additional control. No formal statistical design was planned due to the explorative nature of the study, nevertheless we next extended our analysis to an intended cohort of 20 additional cases. This second cohort was enriched of cases with T size superior to 2 cm. Nevertheless, again 2 cases were not provided by the

service of pathology due to unavailable tissue. After discarding 2 cases with low coverage, mutational analysis was carried out on the remaining 16 samples.

3.1.2 Primary tumor mutations

Considering the DCIS primary tumor samples with adequate DNA and coverage for mutational analysis from both the case-control (N= 11) and the extended study (N=16) as a unique study cohort, 13 cases resulted wild type by the 50 HS panel, and by the CCP in 4 cases with still available tissue.

Fourteen cases had at least one mutation, for an overall mutation rate of 52%, with 2 mutations in 7 cases (Table 3), all validated by ddPCR. *PIK3CA* was the most commonly mutated gene, harboring 8 mutations in 7 cases. All were missense: 4 in the helical (E542K mutations), 3 in the kinase (H1047R), and 1 in the C2 (N345K) domains. One case presented co-existing E542K and H1047R mutations with variant allele frequencies (VAF) of 33% and 32%, respectively. One additional case had mutated *AKT*. All these data are consistent with the early involvement of PI3K/Akt dysregulation during BC development. The next most frequently mutated gene was *TP53* which was altered in 3 cases.

Table 3. DCIS patient features, and mutation analysis of matched primary tumor and plasma samples

Patient number	Age (years)	Detection	Surgery	Histology	Grade	Size (mm)	HR[a] status	HER2	Ki-67 (%)	Adjuvant therapy	Mutational profile			
											Tissue	VAF (%)	Plasma	VAF (%)
DCIS patients with evaluable tissue and plasma mutational profile														
1	28	Clinical	Mastectomy	Comedo	III	12	+	1+	NE	**None**	*PIK3CA N345K*	8	*PIK3CA N345K*	0.04
2	43	Clinical	Mastectomy	Comedo	III	30	+	2+	25	**None**	*PIK3CA E542Q*	33	*PIK3CA E542Q*	0.12
											PIK3CA H1047R	32		
											KIT V532I	49		
3	64	Clinical	Mastectomy	Non comedo	II	30	NE	NE	NE	**None**	*PIK3CA H1047R*	9	*PIK3CA H1047R*	0.11
4	60	Screening	BCS	Non comedo	I	6	+	0	3	**None**	*PIK3CA E542K*	34	*PIK3CA E542K*	0.14
5	47	Screening	Mastectomy	Comedo	II	35	+	3+	35	**None**	*TP53 T163C*	11	*TP53 T163C*	0.9
											ATM V410A	62		
6	44	Screening	BCS	Comedo	III	7	+	2+	30	TAM	*TP53 R282W*	23	*TP53 R282W*	5
											IDH2 R149Q	19		
7	36	Clinical	Mastectomy	Comedo	II	25	+	0	20	None	*TP53 R282Q*	27	*TP53 R282Q*	0.1
											PIK3CA H1047R	28		
8	52	Screening	Mastectomy	Non comedo	III	30	+	1+	<5	None	*BRAF*^{V600E}	34	*BRAF*^{V600E}	0.1
											PIK3CA E542Q	29		
9	36	Screening	Mastectomy	Non comedo	II	11	+	0	25	None	*NTRK3 R793Ter*	15	*NTRK3 R793Ter*	0.34
10	52	Clinical	BCS	Non comedo	II	10	–	0	NE	RT	*KRAS G12V*	5	*KRAS G12V*	0.013
11	73	Screening	BCS	Non comedo	III	25	–	3+	15	RT	*PIK3CA E545K*	34	*PIK3CA E545K*	0
											GNAS R201C	32		
12	72	Screening	Mastectomy	Comedo	III	22	–	1+	NE	None	*ERBB2 V777L*	18	*ERBB2 V777L*	0
											KRAS G12D	18		
13	50	Screening	BCS	Comedo	II	12	+	1+	25	None	*AKT1 E17K*	6	*AKT1 E17K*	0
14	36	Clinical	Mastectomy	Non comedo	II	22	+	1+	12	None	*GATA3 P409fs*	22	*GATA3 P409fs*	0
15	48	Screening	BCS	Non comedo	II	22	+	2+	10	TAM	Wild type			
16	46	Screening	BCS	Non comedo	III	6	+	1+	NE	TAM+RT	Wild type			
17	51	Screening	BCS	Non comedo	II & III	7+2	+	1+	NE	TAM	Wild type			
18	54	Screening	BCS	Non comedo	II	30	+	1+	15	TAM	Wild type			
19	50	Screening	Mastectomy	Non comedo	II	50	+	2+	5	None	Wild type			
20	79	Clinical	Mastectomy	Non comedo	I & II	23	–	1+	<5	None	Wild type			
21	53	Screening	BCS	Non comedo	I	2	+	0	22	None	Wild type			
22	58	Clinical	Mastectomy	Comedo	II & III	35	+	0	15	None	Wild type			
23	78	Screening	Mastectomy	Non comedo	II	34	+	1+	14	None	Wild type			
24	43	Screening	BCS	Non comedo	I	5	NE	NE	NE	None	Wild type			
25	42	Screening	Mastectomy	Comedo	III	6	+	2+	20	None	Wild type			
26	73	Clinical	BCS	Non comedo	II	28	+	1+	14	TAM+RT	Wild type			
27	65	Screening	BCS	Comedo	III	21	+	1+	20	TAM+RT	Wild type			
DCIS patients not evaluable for tissue and plasma mutational profile														
28	71	Screening	BCS	Comedo	II & III	6	+	1+	35	TAM	Low coverage			
29	74	Screening	BCS	Non comedo	I	27	+	2+	4	RT	Low coverage			
30	47	Screening	Mastectomy	Non comedo	II	45	+	3+	15	None	Low coverage			
31	45	Screening	BCS	Non comedo	III	7	+	3+	NE	TAM+RT	Low DIN			
32	55	Screening	BCS	Non comedo	II & III	15	+	2+	10	TAM+RT	Low DIN			
33	58	Screening	BCS	Comedo	III	40	+	3+	NE	TAM+RT	Low DIN			
34	46	Screening	BCS	Non comedo	I	15	+	2+	<5	None	Low DIN			
35	69	Screening	BCS	Non comedo	II	2	+	1+	NE	TAM	Low DIN			
36	43	Screening	BCS	Non comedo	III	15	+	3+	NE	TAM	Low DIN			

Mutations assayed in plasma were those with the highest VAF (patients 6 and 8) or for which a sensitive digital droplet PCR assay was already available (patients 2, 5, and 7). Primary tumor tissue mutations were not tested in plasma when plasma DNA was not available (patients 11 and 12).

BCS, breast conserving surgery; DCIS, ductal carcinoma *in situ*; DIN, DNA integrity number; ER, estrogen receptor; HR, hormone receptor; NE, not evaluated; PR, progesterone receptor; TAM, tamoxifen; VAF, variant allele frequency.

[a] HR status classified as positive when either estrogen receptor or progesterone receptor were ≥1%, or negative when both ER and PR were <1%.

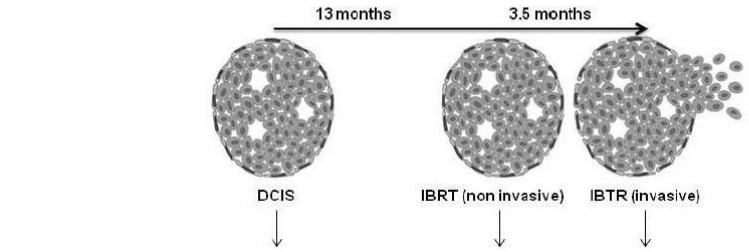
3.1.3 ctDNA in newly diagnosed and recurrent cases

Customized ddPCR assays detected mutations in pre-surgery blood from 10/14 (71%) patients harboring mutated DCIS, with variable levels among patients (median VAF= 0.14%; range 0.01%-5%), unrelated to clinico-pathological features (Figure 17).

Figure 17. Mutation profile according to patient features

Our series also included 3 recurrent patients with evaluable DNA from primary DCIS, resected IBTR, and matched plasma pairs. Patient #24 had no mutations in all surgical specimen. By contrast, patient #1 had mutated *PIK3CA* in each surgical specimen *(i.e.,* primary DCIS, non-invasive IBTR, and subsequent triple-negative BC) (Figure 18), and patient #6 had an IBTR with DCIS and invasive breast cancer which harbored the same *TP53* mutation of initial DCIS. The corresponding plasma samples revealed that primary DCIS mutations in *PIK3CA* and *TP53* were present at the time of recurrence.

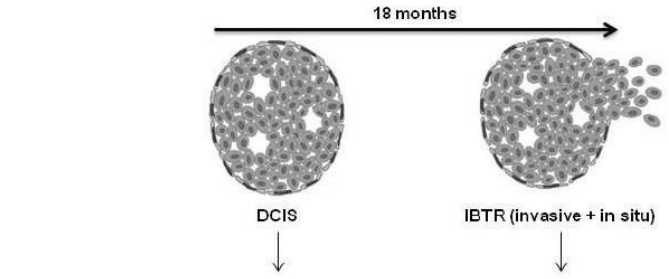
Patient #1 Female, 27 years old at initial diagnosis



Tissue *PI3CA* VAF	8%	15%	3.2%
Plasma *PI3KCA* VAF	0.04%	0.001%	0.03%

Initial Diagnosis	Recurrent disease manifestations		Follow up
02/27/12 Mastectomy of the right side + sentinel node biopsy: DCIS comedo type, G3, ∅ 12 mm, 0/2 node; ER 33-66%, PgR 1-10%, HER2= 1+. No adjuvant treatment.	***04/02/13*** Resection of IBTR: DCIS comedo type, G3, ∅ 20 mm; ER 66-100%, PgR 10-33%, HER2= 1+. Adjuvant tamoxifen.	***07/17/13*** Resection of IBTR + axillary node dissection: invasive ductal carcinoma, G2, ∅ 1.5 mm, G2, TNBC.	***07/31/18*** No evidence of disease.

Patient #6 Female, 44 years old at initial diagnosis



Tissue *TP53* VAF	23%	33% (invasive)	16% (in situ)
Plasma *TP53* VAF	5%	6.5%	

Initial Diagnosis	Recurrent disease manifestation	Follow up
10/19/12 Lumpectomy + SNB: DICS comedo type, G3, ∅ 7 mm, 0/3 nodes; ER 33-66%, PgR 10-33%, HER2= 2+, Ki-67= 30% Adjuvant tamoxifen.	***06/18/14*** Mastectomy of the left side + SNB for: IDC, G3, ∅ 7 mm;ER= 60%, PgR= 8%, HER2= 3+, Ki-67= 70%. Adjuvant trastuzumab-based therapy.	***06/07/19*** No evidence of disease.

Figure 18. Recurrent cases with primary tumor tissue and plasma samples mutation analysis at initial DCIS diagnosis and at the time of relapse

3.2 Targeted next generation sequencing in TNBC

3.2.1 Patient cohort

Table 4. Targeted NGS study cohort

Characteristics	N (%)
Patient age	
<50 years	12 (63)
≥50 years	7 (37)
Clinical T size	
2-5 cm	13 (68)
>5 cm	6 (32)
Clinical Nodal status	
N0	8 (42)
N1	11 (58)
Clinical stage	
IIA	5 (26)
IIB	10 (53)
IIIA	2 (11)
IIIB	1 (5)
IIIC	1 (5)
Tumor grade	
2	1 (5)
3	18 (95)
Type of NAC	
AT --> CMF	11 (58)
AT --> Eribulin	5 (26)
Other	3 (16)
Path findings	
ypT0N0	4 (21)
ypT1N0	6 (32)
ypT2-3N0	7 (37)
ypT2-4N1-3	2 (10)
Evaluable cases	
Pre-NAC	16 (84)
Post-NAC	15 (79)
Paired pre- and post-NAC	12 (63)
Events	
Distant Metastases	7 (37)

Pre-treatment and surgically resected tumor samples were from TNBC patients treated with anthracycline/taxane-based NAC and breast and axillary surgery between September 2010 and March 2016. Clinical and pathological data were retrieved from the prospectively maintained pathology-based institutional registry [Baili P et al., 2015]. Median patient age was 43 years (range 32-75). Fifteen patients presented with stage II disease, three with stage III, and one patient with suspicious bone metastases at initial diagnosis. All cases but one were grade 3. After a median number of six (range 4–8) cycles of NAC, the pCR rate was 21%. With a median follow-up of 70 months (95%CI 50–81), seven (37%) patients had relapsed. Pre- and post-NAC tumor sequencing was feasible in 16 (84%), and 15 (79%) of the 19 cases, respectively. Inadequate DNA content (pre-NAC cases = 3), and unavailability of post-NAC tumor tissue (in

four cases achieving pCR) prevented mutation analysis in missing cases. For 12 (63%) patients both pre- and post-NAC tumor specimens were analyzed.

3.2.2 TNBC somatic mutations

According to the 409 comprehensive cancer panel, somatic mutations were detected in all 31 sequenced tumor samples, with a median of four mutations per sample. The majority of detected mutations were missense, followed by truncating mutations and variants affecting splicing and the most frequent nucleotide variation was C > T/G > A (Figure 19, A). Among the 10 most frequently mutated genes we found *TP53*, altered in 87% of pre-NAC, and in 60% of post-NAC samples, followed by a long tail of less recurrently mutated gene (Figure 19, B). A decreased number of mutations was observed in patients ≤50 compared to those >50 years old, 4 (2-6) versus 7 (5-66), p= 0.002. This association is unlikely confounded by other factors since the proportion of stage, grade, and type of treatment was similar among different age categories in this case series.

3.2.3 Genes and pathways associated with pCR

Considering pre-NAC samples, we first assessed whether patients with pCR or residual disease (RD) had a different tumor mutational load. No significant difference was observed in the number of mutations between the two groups, both considering mutations with high/moderate impact, and overall detected somatic mutations. Furthermore, the frequency of mutated genes was similar in patients with and without pCR. Beside *TP53* that was mutated in 87% of pre-NAC tumors, most genes were mutated in few patients and no significant association between altered genes and NAC response was found. When we grouped mutated genes by pathway no significant association with pCR was observed. Nevertheless, 50% of patients without pCR (N= 6) were characterized by alterations in pathways related to adaptive immunity, and in particular in B cell and T cell signaling, and in ERBB signaling (Figure 20).

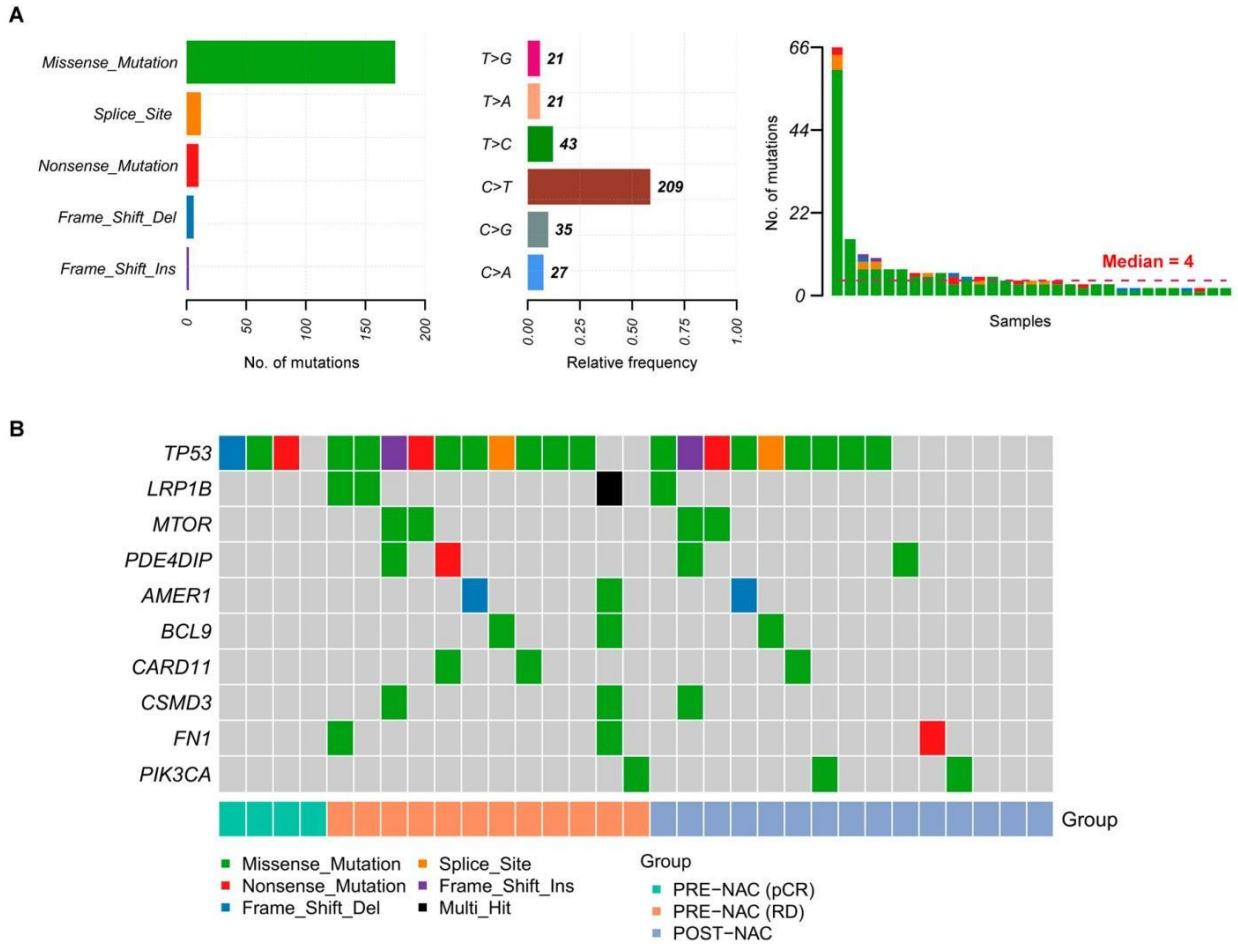


Figure 19. Mutational spectrum of triple negative breast cancer. (A) Bar plots show the number of somatic mutations by type, the relative frequency of nucleotide substitutions, and the number of mutations per sample by mutation type. (B) Oncoplot reporting the top-10 most recurrently mutated genes across the TNBC samples analyzed in this study. Samples are ordered according to response group. Genes are listed from top to bottom by decreasing frequency

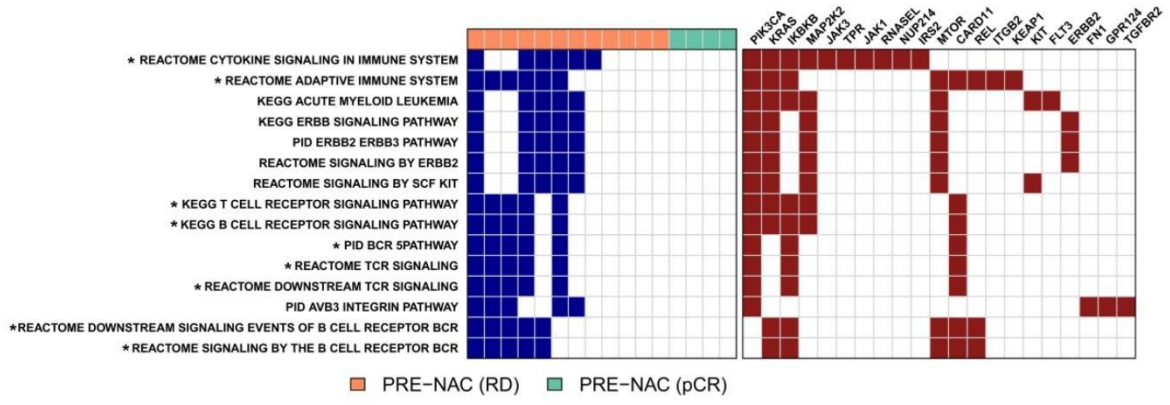


Figure 20. Signaling pathways associated with pCR in pre-treatment tumor specimens. Heatmap shows the 15 pathways (rows) preferentially mutated in patients with RD compared to those with pCR (columns). Asterisks indicate immune-related pathways. Blue squares indicate that the pathway is mutated in the sample. Mutated genes (red squares) included in each pathway are reported.

None of the four patients attaining pCR showed alterations in these pathways. Notably, primary tumors with stable and mutated mutational profile shared the same mean sTILs values, *i.e.*, 13% (range 5-25%, and range 1-35%), respectively. After treatment with NAC, sTILs remained the same, *i.e.*, 13% (range 3-30%) in stable cases whereas they reached an average value of 19% (range 5-50%) in cases with a modified mutational profile.

3.2.4 Comparison between pre- and post-treatment samples

Comparison of tumor mutational load between pre- (N = 16) and post-NAC (N = 15) samples showed no differences considering either high and moderate impact mutations, or all detected somatic mutations. Clustering of pre- and post-NAC tumor pairs according to VAF of somatic mutations showed that for some patients the mutational profiles of tumors were conserved during NAC, while for others pre- and post-NAC samples had different mutations (Figure 21, A). Examination of detected mutations showed that there was a portion of mutations private to (i*.e.*, observed only in) pre- and post-NAC tumors; the median number of mutations private to pre- and post-NAC tumors was 2 (range 0–65) and 0.5 (range 0–3), respectively; there were very few shared mutations, and the median number of common mutations was 2 (range 0–10) (Figure 21, B).

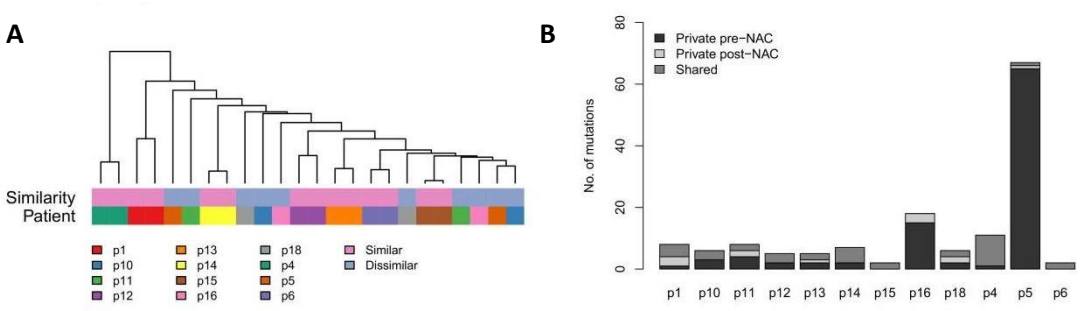


Figure 21. Evaluation of patients with matched pre- and post-NAC samples. (A) Hierarchical clustering analysis based on mutant allele frequency showing patients with similar (pink) or dissimilar (light blue) mutational profile between their pre and post-NAC samples. (B) Bar plots showing number of private pre-NAC (black), post-NAC (grey), and shared mutations in patients with matched samples.

To further assess the characteristics of genomic alteration in unresponsive TNBC patients, we examined individual cases evaluating the different mutant allele frequency distribution between pre- and post-NAC sample. In eight out of 12 (66%), the mutant alleles frequencies differ between pre- and post-NAC samples, identifying groups of residual mutations. The representative case of this behavior, patient #16, is characterized by loss of 15 genes in the post-NAC sample, and by the presence of residual mutations in cluster C2, represented by *SOX11*, *TAF1L*, and *TCF7L1* genes (Figure 22, A). In four out of 12 (33%) patients, the frequency of single mutations did not change between pre and post-NAC conditions. Considering patient #4 as a representative case, the highest VAF was detected for cluster C1 involving *LTK*, *SF3B1*, *TP53* genes in both pre- and post-NAC samples (Figure 22, B). Notably, all patients with stable clusters and mutation relapsed, as compared to 3/8 (37.5%) patients presenting different clusters in pre- and post-NAC samples. Globally, genes in post-NAC samples were involved in pathways associated with actionable targets in tumor treatment where, besides the canonical PI3K/Akt/mTOR, EGFR, and Ras signaling pathways the most represented terms were related to regulation of cell cycle processes (Table 5).

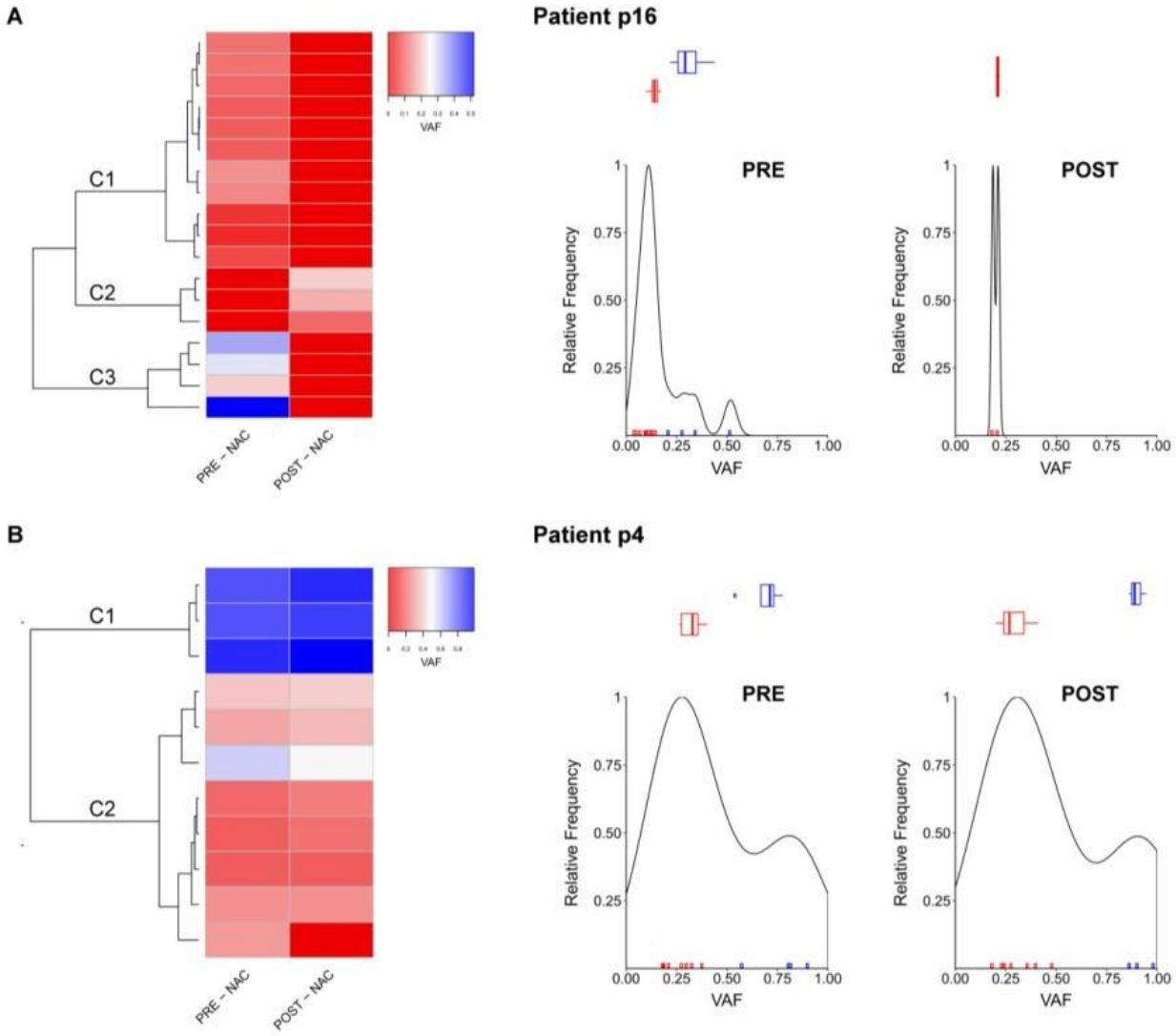


Figure 22. Mutant allele distribution before and after NAC. Non-synonymous mutations of representative patients with residual mutations in post-NAC samples (A) and without evidence of change from post-NAC samples (B) are grouped and represented in red-blue scale according to their frequency.

Table 5. Pathway enrichment analysis of mutations in cases with changed mutational profile before and after neoadjuvant treatment.

Patient	Cluster	No. of Genes in Cluster	No. of Genes in Cluster Mutated in POST-NAC Tumor	Enriched Pathways Including Genes Mutated in POST-NAC Tumor
p5	C1	42	1	No pathways found
	C2	10	1	No pathways found
	C3	15	0	NA
p10	C1	3	3	EGFR tyrosine kinase inhibitor resistance (KEGG - hsa01521); Ras signaling pathway (KEGG - hsa04014)
	C2	3	0	NA
p11	C1	3	2	Negative regulation of cell cycle process (GO:0010948); Androgen receptor signaling pathway (GO:0030521)
	C2	5	2	Ras signaling pathway (hsa04014); TOR signaling (GO:0031929); EGFR tyrosine kinase inhibitor resistance (KEGG - hsa01521)
p12	C1	3	3	mTOR signaling pathway (KEGG - hsa04150); PI3K-Akt signaling pathway (KEGG - hsa04151);
	C2	2	0	NA
p13	C1	1	1	PI3K-Akt signaling pathway (KEGG - hsa04151); Negative regulation of cell cycle process (GO:0010948)
	C2	4	2	TORC1 signaling (GO:0038202)
p16	C1	11	0	NA
	C2	3	3	Hippo signaling pathway (KEGG - hsa04390)
	C3	4	0	NA
p18	C1	2	2	Regulation of cell cycle arrest (GO:0071156); PI3K-Akt signaling pathway (KEGG - hsa04151)
	C2	2	2	Mismatch repair (KEGG - hsa03430); Platinum drug resistance (KEGG - hsa01524)
	C3	2	0	NA

3.3 Liquid biopsy to monitor TNBC progression

Based on our experience with DCIS, and the NGS analysis of pre- and post-neoadjuvant samples, we next conducted an observational prospective study to explore the value of ctDNA as a longitudinal test for early detection of relapse in TNBC patients, to challenge high-depth plasma sequencing and single cell CTC analysis from blood drawn as means to identify druggable molecular features in relapsed cases.

3.3.1 Study CONSORT

Starting from 42 patients, 7 and 4 patients were excluded due to the lack of serial blood sampling and primary tumor tissue, respectively (Figure 23).

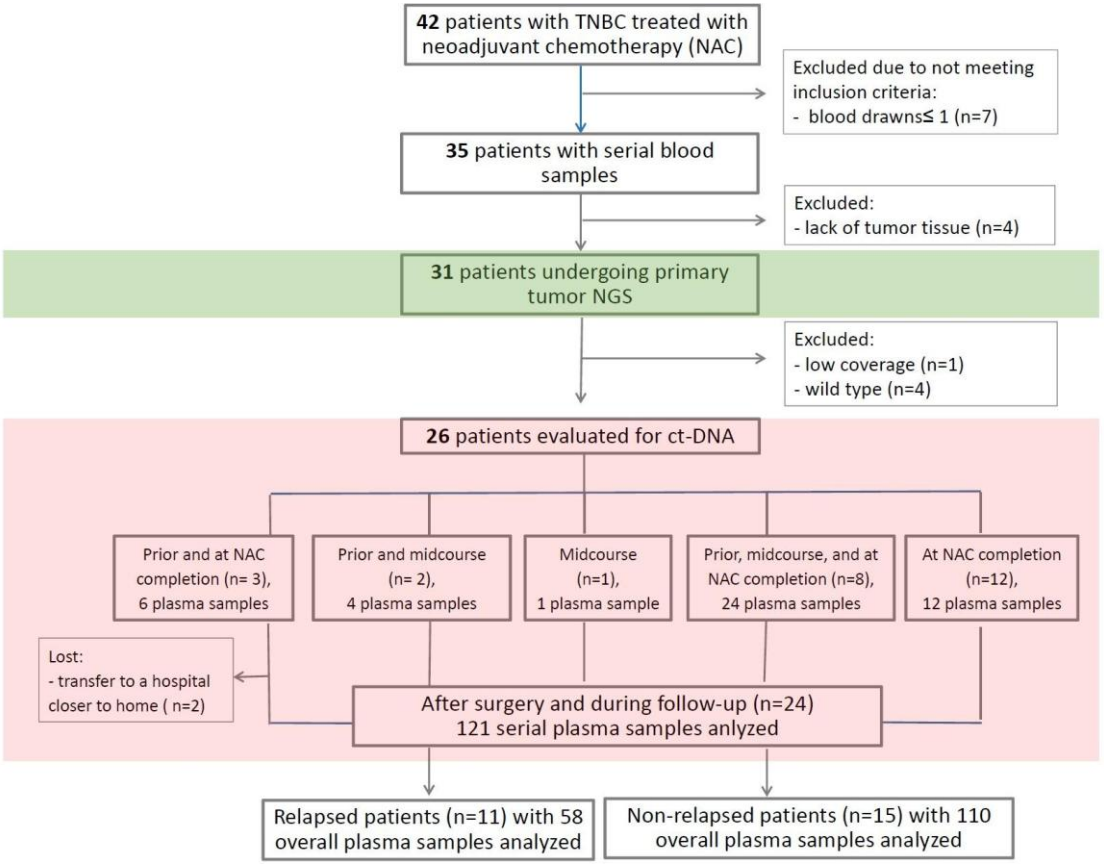


Figure 23. Liquid biopsy to monitor TNBC progression: study consort. The diagram show patients included in each analysis and reasons for their exclusion. Green shadow: patients with primary tumor bearing at least one mutation; Red shadow: patients with overall ctDNA assessment :and according to different treatment time points and follow-up.

The clinico-pathological features of the 31 patients undergoing primary tumor NGS and with at least two serial blood sampling representing the study cohort are listed in Table 6.

Table 6. Prospective neoadjuvant study population

AGE	**N (%)**
<50 years	19 (61.3)
≥50 years	12 (38.7)
CLINICAL TUMOR SIZE	
2-5 cm	22 (71.0)
>5 cm	9 (29.0)
CLINICAL NODAL STATUS	
N0	11 (35.5)
N≥ 1	20 (64.5)
CLINICAL STAGE	
II	24 (77.4)
III	7 (22.6)
TUMOR GRADE	
G2	1 (3.2)
G3	26 (83.9)
Missing	4 (12.9)
Ki67	
<50%	6 (19.3)
≥50%	22 (71.0)
Missing	3 (9.7)
TYPE OF NAC	
Anthracycline/taxane	25 (80.6)
Anthracycline/taxane plus platins	4 (12.9)
Other	2 (6.5)
PATH FINDINGS	
ypT0N0 (pCR)	4 (12.9)
ypT1N0	13 (41.9)
ypT1Nx	2 (6.5)
ypT2-3N0	5 (16.1)
ypT1-3N1-3	7 (22.6)
BREAST CANCER EVENTS (N= 12)	
Distant metastases	9 (75.0)
Second primary	1 (8.3)
Loco-regional relapse	1 (8.3)
Death from any cause	1 (8.3)

All except two patients received anthracycline and taxane-based chemotherapy before surgery. All patients presented with primary tumors≥ 2 cm, and 20/31 (64.5%) with initial nodal involvement. The majority of patients were responsive to NAC except 4 that anticipated surgery due to progressive disease. Four patients attained the pCR. At the reporting census date (February 12, 2020), 12 of the 31 patients experienced an unfavorable event, including 1 contralateral BC, 1 loco-regional relapse, 9 metastatic dissemination, and 1 death for causes not related to BC and/or treatment. Overall a total of 45 specimens were processed, ie diagnostic biopsy (N= 5), surgical samples (N=12), and paired diagnostic

biopsy/surgical samples (N=14). At least 1 somatic mutation was found in 22/33 (66.6%) and in 15/21 (71.4%) samples by the 50-gene HS and CCP, respectively. Of note, CCP identified at least one mutation in 3/9 samples missed by the 50-gene HS, and in additional 12 samples processed upfront. No results were available in 6 samples with either panel. In total, 37 out of 45 (82.2%) cases were found with at least one mutation (range 1-9), and in particular in 19/37 (51%) multiple mutations were detected. The most frequently mutated gene was *TP53* (20/26, 76.9%), followed by *PI3KCA* (4/26, 15.3%), and *FGFR3* (2/26, 7.7%).

3.3.2 ctDNA in TNBC patients undergoing neoadjuvant chemotherapy

Among the 26 patients evaluated for ctDNA, blood drawings were available prior, during, after completion of NAC, and after surgery in 13, 11, 23, and 24 patients, respectively. Plasma DNA was extracted from 168 samples (median per patient, 6 samples; range 2-13 samples). The detection rate of ctDNA prior to commencing NAC was 10 out of 13 evaluable cases (77%) with a median VAF value of 1.36% and 17.55 copies/ml. At mid-course evaluation, half of these cases turned negative, 2 presented persistent though reduced levels of ctDNA (patients #17, and #4, with corresponding VAF values of 0.3% and 0.24%), and 3 were missed. Following NAC, ctDNA was still detectable in 10 out of 23 (11 with initial assessment prior to commencing NAC, and 12 with initial assessment at the end of NAC) evaluable cases (43%) with a median VAF value of 0.3%, and 28.29 copies/ml. Overall, these findings suggest that NAC reduced the levels of ctDNA soon after treatment start, and up to undetectable levels in most of the cases at the time of surgery (Figure 24).

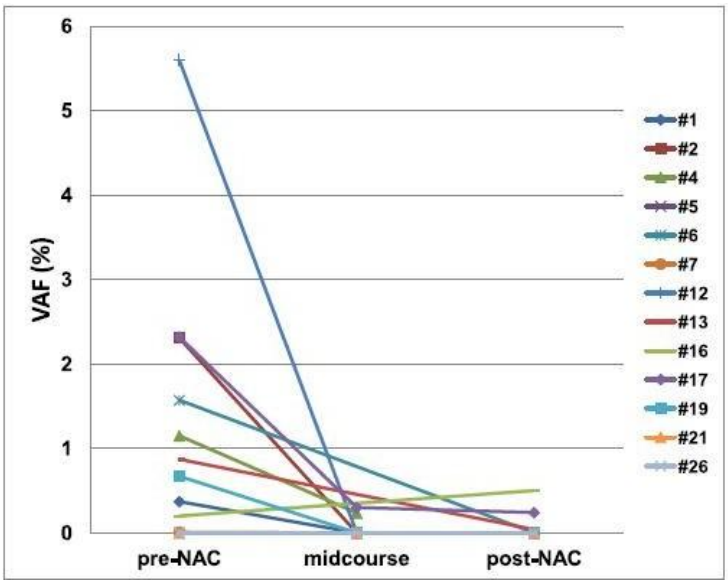


Figure 24. Midcourse and post-NAC ctDNA changes in non responsive cases. Three cases (#7, 21, 26) with pre-NAC undetectable ctDNA showed the same results also at midcourse and post-NAC.

3.2.3 ctDNA status, clinico-pathological features and prognosis

No clinico-pathological characteristics, including systemic inflammatory indices, were different between patients with detectable or undetectable ctDNA either at baseline or at the end of NAC, though ctDNA detection after NAC appeared more likely in cases with high LMR and low NLR ratios, and high Ki67 (Figure 25).

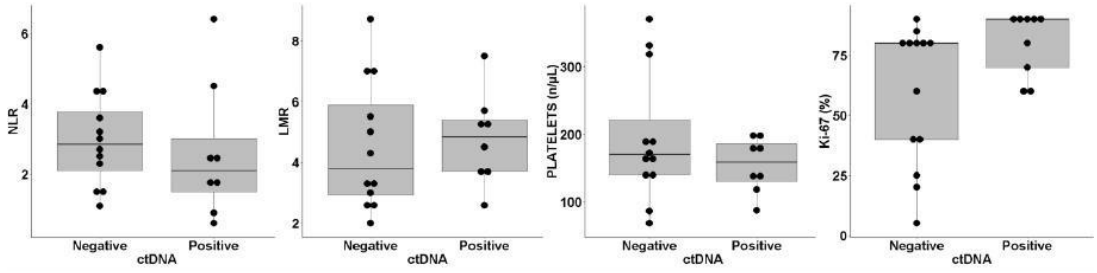


Figure 25. Boxplot showing the distribution of the post-NAC systemic inflammmatory indices (NLR, LMR, platelets), and Ki-67 according to ctDNA status

The summary of results obtained by the simultaneous consideration of all the variables (ctDNA, systemic inflammatory indices, Ki67 and patient age) and

presence/absence of unfavorable events during follow-up is detailed in the heatmap of Figure 26.

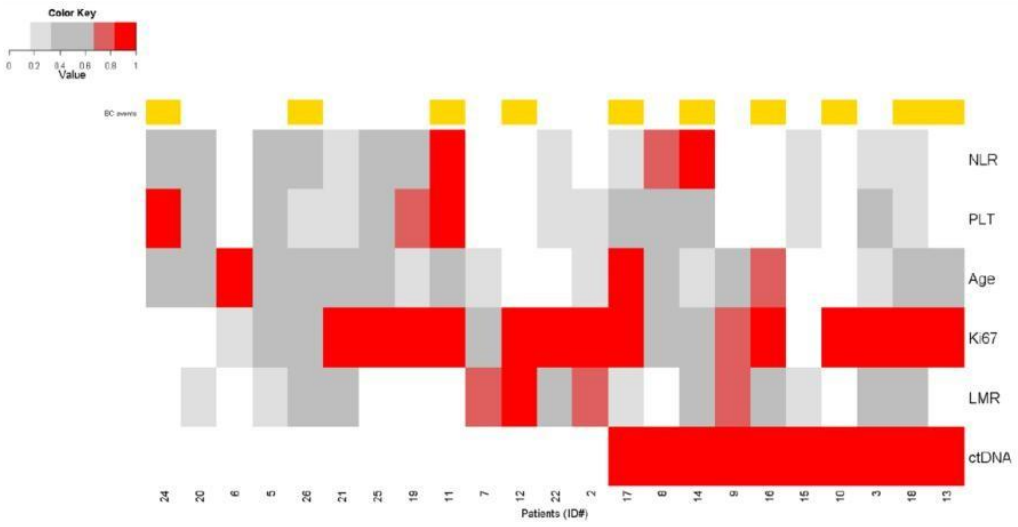


Figure 26. Patient age, post-NAC ctDNA, Ki67 and systemic inflammatory indices as a function of presence/absence of relapse. The numerical variables were rescaled to the 0-1 range

The conversion of ctDNA status from positive at baseline to negative during and after completing NAC was not associated with primary tumor response, as all 13 evaluable patients with post-NAC undetectable ctDNA levels had residual disease in surgical specimen. Nonetheless, all 3 evaluable patients (#3, #8, #10) who progressed during NAC had still detectable levels of ctDNA prior to surgery.

A total of 11 BC events occurred among the 26 patients with detectable ctDNA either pre- or post-NAC. ctDNA detection partially overlapped with higher Ki67 and LMR values and younger age, even though ctDNA seems to better explain the distribution of BC events in the study population. Specifically, negative ctDNA occurred in 5 out of 6 patients without recurrence despite high values of Ki67, whereas positive ctDNA occurred in 1 out of 3 relapsed patients despite low levels of Ki67.

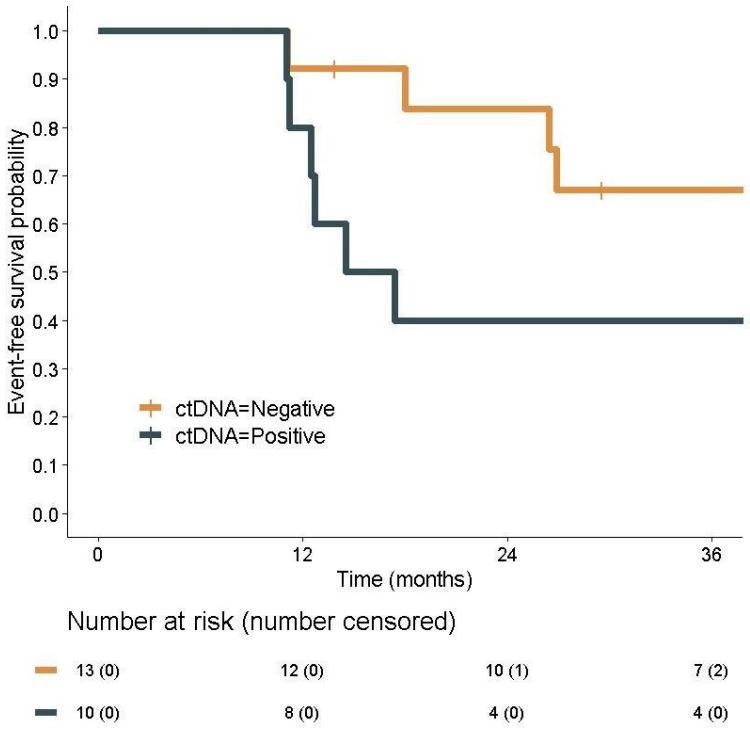


Figure 27. Kaplan-Meier event-free survival. The curves represents event-free survival according to post-NAC ctDNA status. Number of patients at risk and censored are shown at the figure bottom.

Survival curves showed no difference in clinical outcome between patients with detectable ctDNA at baseline (pre-NAC, N= 10) and those with no detectable ctDNA (N= 3), with 2-year EFS estimates of 57.1% (95% Confidence Interval [CI]: 32.6-100%) versus 66.7% (95% CI: 30.0-100%), respectively. More strikingly, patients with post-NAC ctDNA-positive status (N= 10) as compared to those with undetectable ctDNA (N= 13) had an increased risk of recurrence after surgery, with 2-year EFS estimates of 40.0% (95% CI: 18.7-85.5%) versus 83.9% (95% CI: 65.7-100%) (Figure 27, univariable Cox model HR 2.65; 95% CI: 0.74-9.44). Notably, the prognostic value of post-NAC ctDNA remained worth of consideration even after adjustment for age, Ki67, residual disease at surgery, and systemic inflammmatory indices (HR 1.91; 95% CI: 0.51-7.08) .

3.3.4 Serial ctDNA and patient outcome

All the patients with detectable levels of ctDNA at either pre-or post-NAC sampling were longitudinally monitored after surgery with the exception of two (#13 and #16) which had detectable ctDNA prior to surgery, recurred 7.2 and 9.3 months after surgery, and were not tested during follow-up. The dynamics of ctDNA during follow-up and the lead-time of molecular compared to clinical progression is detailed for individual patients in Figure 28. The first post-operative sample

showed undetectable levels of ctDNA in 20 of 24 (83.3%) evaluable cases. Among the patients with persistent levels of ctDNA, three (#14, #17 and #18) experienced a BC event after an average of 6.5 months from surgery, while patient #5, that showed a transient peak of ctDNA followed by consistently negative levels, is still disease-free up to 4 years from surgery. Three patients (#1, #11, and #12) turned ctDNA positive during follow-up after an average of 12.7 months from surgery, and in these patients ctDNA anticipated overt metastases by a mean of 8.9 months (range 6.5-13.1). Thus, the positive predictive value of detectable levels of ctDNA after surgery (either for persistence or reappearance) was evident in 6/7 cases (85.7%).

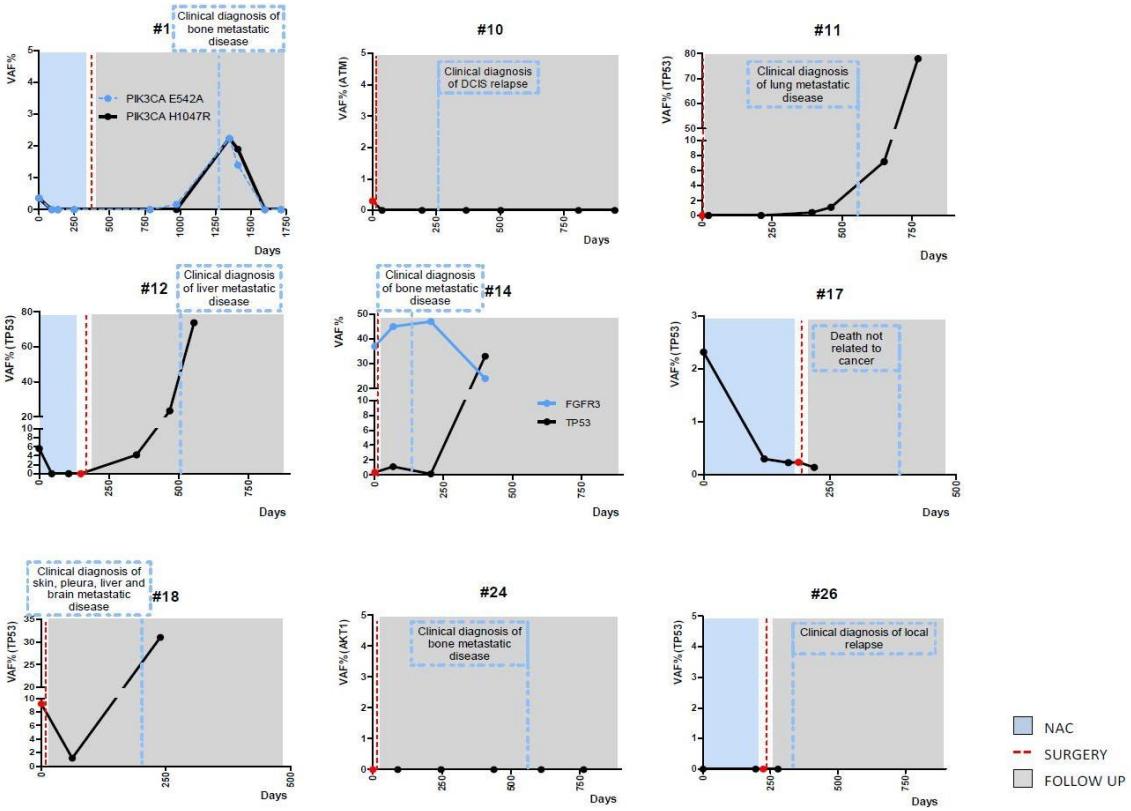


Figure 28. Serial ctDNA assessment using indiviual patient mutations in relapsed cases.

Among the 17 patients with persistently undetectable ctDNA after NAC, 14 remained disease-free at a median follow-up of 3 years (range, 0.5-6.5). Two patients (# 24 and #26) tested negative prior and after NAC, and one additional

patient (#10) with detectable post-NAC ctDNA, were diagnosed with bone metastases, loco-regional relapse, and contralateral BC in absence of detectable ctDNA. Thus, the negative predictive value of undetectable levels of ctDNA after surgery was demonstrated in 14/17 cases (82.3%).

3.3.5 Direct plasma sequencing in recurrent patients

Eight recurrent patients had plasma samples collected following NAC (#13), after surgery (#17), or at the time of clinical relapse (#1, 11, 12, 14, 18 and 24), allowing a further analysis using the NGS panel of 52 cancer-related genes. No mutations were found in 3 cases. Notably, 2 of these cases (#13 and 17) had instead detectable ctDNA by ddPCR, though with low VAF values of 0.04% and 0.14%, respectively. In the remaining 5 cases, a median of 2 mutations per sample was found, with median higher VAF values of 10.35% (0.7-80%) (Table 7). All patients retained at least one of the primary tumor mutations. However, patients #12 and #14 lost mutated *ATM* and *TP53* found in surgical samples, patient #11 gained an extra mutation in *TP53* (VAF 80%), and patient #18 a *de novo* mutation in *MAP2K2* (VAF 0.75%). Taken together, these data imply that ctDNA, while retaining the dominant features of the primary tumor, is informative of genetic alterations occurring during progression.

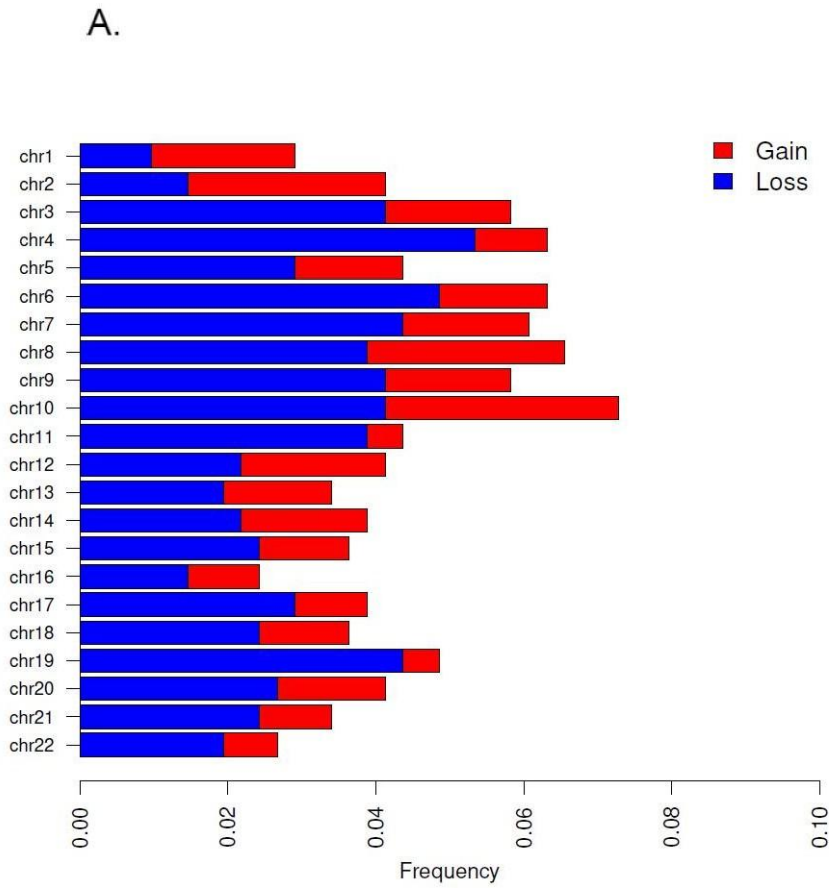
Table 7. Plasma sequencing at progression for relapsed patients

ID	PRIMARY TUMOR TISSUE				PLASMA		
	BIOPSY		RESIDUAL DISEASE		ddPCR VAF%	PAN CANCER (PC) MUTATIONS	PC VAF%
	MUTATION	VAF%	MUTATION	VAF %			
#1	**PIK3CA H1047R**	6	**PIK3CA H1047R**	52	2.24	**PIK3CA H1047R**	2
	PIK3CA E542A	13	**PIK3CA E542A**	48	2.2	**PIK3CA E542A**	7
#11	NOT AVAILABLE		**TP53 R213Q**	51	7.2	**TP53 R213Q**	7.7
			TP53 R158Q	49		TP53 R158Q	5.6
						TP53 P85sf	5.6
#12	**TP53 H179Q**	56	**TP53 H179Q**	58	73.6	**TP53 H179Q**	67
			ATRX Q2399H	10			
#13	**TP53 P72R**	42	UNDETECTABLE		0.04	0	
#14	**TP53 K132N**	60	**TP53 K132N**	18	33	**TP53 K132N**	25
	FGFR3 F384L	20	**FGFR3 F384L**	34	24	**FGFR3 F384L**	42.8
	TP53 R306Q	70	TP53 R306Q	16			
	FGFR3 P796L	19					
	FGFR2 A371V	31					
	IDH1 K115E	10					
#17	**TP53 R248W**	48	**TP53 R248W**	51	0.14	0	
#18	**TP53 p.P177_C182del**	70	**TP53 p.P177_C182del**	60	31	**TP53 p.P177_C182del**	13
						MAP2K2 F129L	0.7
#24	NOT AVAILABLE		**AKT1_E17K****	7.4	0	0	

3.3.6 Analysis of CTCs in recurrent cases

To further characterize liquid biopsies at time of disease progression, CTCs were detected and molecularly profiled in 6 relapsed patients. Among the 21 CTCs collected, 1 expressed epithelial markers (eCTCs) only, corresponding to the

classical CTC definition, and 20 CTCs lacked both epithelial and leukocyte markers or expressed more than one marker. Those latter CTCs that did not meet the classical CTC definition were characterized by aberrant genomes and were considered as non-conventional CTC (ncCTCs). The molecular characterization of recovered CTCs showed a prevalence of deletions on amplifications, and included chromosomes 5, 8, 17 (Figure 29, panel A), which have already been described in the literature as frequently loss in primary TNBCs [Li Z et al., 2020]. Moreover, the 10q and 21q were the most frequently altered chromosomal arms (Figure 29, panel B). A network analysis of physical interaction among BC-related genes from these altered regions identified a module of 28 nodes involved in therapeutically exploitable pathways including mismatch repair, PI3K/Akt, erbB, Raf, platinum-resistence signaling, and regulation of immune response.



B.

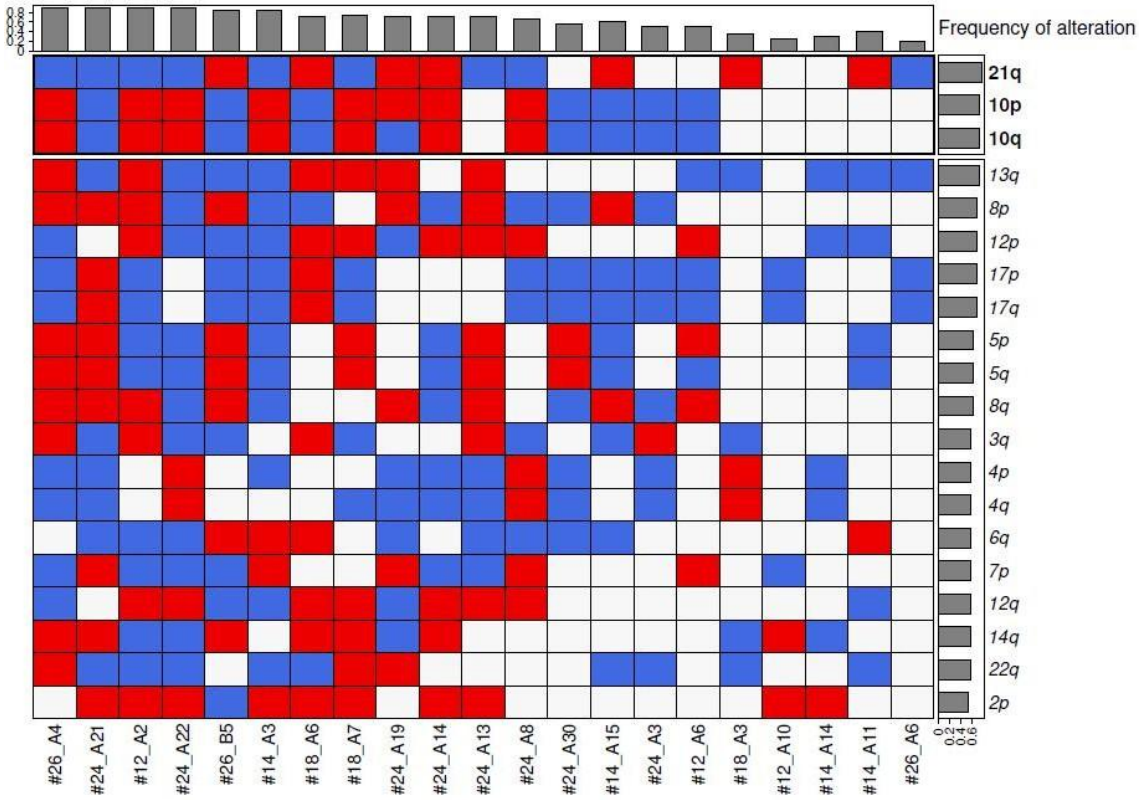


Figure 29. Copy number alterations (CNAs) detected in CTCs from relapsed cases.

A) The barplot shows the distribution of CNAs along chromosomes considering all the CTCs collected at time of relapse. Red and blue colors refer to amplifications and deletions, respectively. B) The heatmap reports patients on the column and the top 20 altered chromosomes arms on the rows. Red and blue colors refer to amplifications and deletions, respectively. 21q (43%), 10p (41%) and 10q (41%) were the most commonly altered arms.

CHAPTER 4.

DISCUSSION

This study provides an overview of the genomic alterations in matched pre-invasive and invasive primary breast cancer, and pre- and post-treatment triple negative breast cancer specimens. The characterization of molecular alterations represented in ductal carcinoma in situ, and unresponsive primary triple negative breast cancer cases, and their tracking in bloodstream might offer promising new strategies in the management of patients to prevent, and overcome tumor progression.

Starting with ductal carcinoma in situ, our results extended the value of ctDNA to pre-invasive forms and represent, to our knowledge, the first report showing that primary DCIS mutations can be found in plasma samples in a similar manner to invasive counterparts, possibly with lower VAF values. Despite the wealth of published data, evidence from the literature showed the absence of significant genetic differences between DCIS, and invasive cancer [Sheek K et al., 2019], and studies suggesting otherwise were limited by the small sample size analyzed, or lack of functional applicability [Pang JB et al., 2017; Abba MC et al., 2015]. With this in mind, we planned to analyze the DCIS population from current clinical practice, with a 6-year follow-up, and a number of secondary events consistent with literature data [EBCTCG, 2010]. Unfortunately, our initial intention to evaluate ctDNA in women with and without relapse was prevented by the fact that half of the cases were not evaluable. This was to be expected since the retrospective nature of the study, and in fact the analysis was carried out on leftover tissue rather than on *ad hoc* collected specimens. Besides, most DCIS specimens are usually small, or have low overall cellularity, posing significant challenges in the generation of high throughput sequencing libraries [Nachmanson D et al., 2020].

Despite this limitation, our data showed ctDNA in 2/3 (67%) of evaluable recurrent cases as compared with 4/8 (50%) matched evaluable controls. By enriching the study population with additional 20 non recurrent cases with primary tumor size

greater than 2 cm we had fewer technical problems. As a consequence, we increased our ability to successfully profile up to 16/18 (80%) cases finding ctDNA in all 8 cases bearing at least one primary tumor mutation. So that ctDNA marks at least half of newly diagnosed DCIS cases seems to us to be safe to say. Whether it can mark a DCIS more prone to recurrence remains to be demonstrated within planned clinical trials overcoming technical challenges Martínez-Pérez C et al., 2017], including more patients, and extending follow-up. Several biomarkers are being investigated to define patients who will develop invasive recurrence after DCIS. This would allow to reduce the prescription of adjuvant treatment to operated patients, or even to avoid surgery in extremely indolent cases [van Seijen M et al., 2019].

Our exploratory analysis suggests that ctDNA in operated DCIS could have the same application as in invasive cancer, *i.e.*, it could be used to monitor patients over time. Indeed, our data tell us on the one hand that women with DCIS may present with basal ctDNA, and on the other hand that the same ctDNA is present at the time of recurrence. Unfortunately, since we do not have intermediate samples, we cannot know if, as in the case of recurrence from operated invasive disease [Di Cosimo S et al., 2019], also that from operated DCIS can be anticipated. Finally, the DCIS findings are warning about the use of direct sequencing panels on healthy populations [Cohen JD et al., 2018], since ctDNA may be present also in the absence of a frankly invasive tumor.

Moving to NGS in TNBC analysis, we recognize that mutational profile of breast cancer has been widely explored in multiple contexts [Rossing M et al., 2019]. This study conducted a focused assessment of mutational changes in primary TNBC exposed to standard neoadjuvant chemotherapy, matched pairs in 63% of cases, over a limited time period, providing an opportunity to isolate anthracycline and taxane associated genomic changes in this clinical context.

Our primary results included the finding that TNBC is a very heterogeneous disease with few recurrent mutations. The unique exception was represented by *TP53,* which was found altered in 87% of pre-treatment samples, a finding in line with prior studies [Balko JM et al., 2014; Lips EH et al., 2015], including the TCGA dataset (http://portal.gdc.cancer.gov). In fact, we were not able to find specific tumor somatic mutations associated with response/resistance to neoadjuvant chemotherapy.

Nevertheless, albeit the sample size, by using external knowledge, *i.e.*, canonical pathways from MSigDB, we reported that none of the patients attaining pCR presented alterations in the immune related processes. As known, increasing data support the role of environmental factors in the escape of the tumor cells from the effect of cytotoxic agents. In particular, the influence of innate, and adaptive immune responses on the efficacy of chemotherapy has been intensely examined and recently reviewed [Binnewies M et al., 2018]. Tumor associated macrophages have been documented to contribute to doxorubicin and paclitaxel resistance by direct cathepsin protease Band S-mediated effects; in addition, their increased influx into tumors leads to a surge of regulatory T and B cells that is accompanied by impaired recruitment of cytotoxic T cells [Shree T et al., 2011]. On the other hand, several chemotherapeutic agents, including doxorubicin, are able to induce immunogenic cell death, antigen-presenting ability of dendritic cells, and a subsequent T cell response [Rapoport BL et al., 2019]. Our analysis revealed that TNBC tumors bearing proficient genes involved in the immune pathways are much more sensitive to chemotherapy.

Although our findings require confirmation in larger datasets, evaluation of additional variables including disease, diet, and lifestyle, preferably originating from prospective clinical studies, they provide the impetus for using a target gene profile to aid an effective strategy with near-term clinical impact on TNBC

management, by identifying those patients who might benefit from neoadjuvant chemotherapy, and directing the remaining patients to novel therapeutic strategies including immunotherapy, based on the profile of residual disease. As known, the FDA has recently approved atezolizumab as an immunotherapy for TNBC treatment (http://fda.gov), here we provide evidence of other pathways of the immune system that could also serve as targets for novel TNBC therapies. Notably, based on our analysis of deregulated immune response in unresponsive patients, we point out that deregulated NF-kappa-B signaling is known to be implicated in the pathogenesis of numerous human inflammatory disorders, and malignancies. Consequently, the NF-kappa-B pathway has attracted attention as a promising therapeutic target for drug discovery, and different compounds have been characterized, with many demonstrating promising efficacy in pre-clinical models of cancer and inflammatory disease [reviewed by Awasthee N et al., 2019]. Finally, treatments that increase RNase L activity directly or indirectly are predicted to increase the antitumor activities of well-known anticancer compounds, including 5-azacytidine [Banerjee S et al., 2019].

We further reported that the overall number of tumor somatic mutations was not increased after neoadjuvant chemotherapy. When considering previous studies [Balko JM et al., 2014; Lips EH et al., 2015], our findings reinforce the concept that mutation changes may be related to the different chemosensitivity levels of mutant, and wild-type cancer cells. For patients who had loss of mutations after neoadjuvant chemotherapy, it is likely that their cancer cells with certain mutations might be sensitive to chemotherapeutic agents. Therefore, cytotoxic therapy would primarily remove the mutant cells and consequently shift the evolutionary landscape in favor of the non-mutant subclones. Conversely, for patients without loss of mutations after neoadjuvant chemotherapy, cancer cells might be resistant to chemotherapeutic agents and in the residual tumors, the mutant alleles remain

or become enriched. We further observed tumor somatic mutations private to pre- and post-treatment samples, and found that they are primarily sub-clonal. Hence the overall effect of neoadjuvant chemotherapy is the reduction of the total number of mutations, and the appearance of new mutations not detected in pre-treatment samples. While this finding may be consistent with neoadjuvant chemotherapy-induced subclone reduction, and selection of pre-existing clones with low rate mutations, the confounding effects of tumor spatial heterogeneity, as observed in multi-regional biopsies, cannot be completely ruled out with this analysis alone [Martelotto LG et al., 2014].

By hierarchical clustering, we found that all patients with concordant pre- and post-treatment clusters or mutations (such as in the case of p15 which presented a unique *TP53* mutation before and after therapy) relapsed as compared to 37.5% of those with neoadjuvant chemotherapy-induced changes. Prior studies, reviewed by [Pribluda A et al., 2015], found that intratumoral heterogeneity (variably defined) was associated with survival and response to therapy. While this may reflect intrinsic tumor biology, our results suggest that stable post-treatment mutational profile (perhaps reflecting broad tolerance to treatment) may provide additional prognostic information; thereof, both pre- and post-treatment assessments of intratumoral heterogeneity may aid in risk stratification and should be explicitly assessed in future clinical trials. In addition, albeit the small sample size, basal and post treatment sTILs appear to be higher in tumors with changed mutational profile upon neoadjuvant chemotherapy, a result which is consistent with a possible anti-tumor effect of chemotherapy through modulation of the immune system [Dieci MV et al., 2014; Luen SJ et al., 2019]. Additional studies examining the tumor immunological landscape and response to chemotherapy through different “omics” in pre-and post-treatment settings may clarify this relation further. Recently, TNBC has been classified by gene expression into four different

subtypes by taking into consideration the contribution of transcripts from normal stromal and immune cells in the tumor environment [Lehmann BD et al., 2016]. These four TNBC subtypes differed in their response to standard neoadjuvant chemotherapy, with basal tumors having a better response than non-basal. Noteworthy, the luminal androgen receptor (LAR) subtype, which indeed presents increased AR signaling, had better survival despite a decreased response to neoadjuvant chemotherapy.

Although few patients shared similar tumor mutational profiles, common pathways in unresponsive cases included those of regulation of cell cycle progress, PI3K/Akt and mTOR signaling, EGFR tyrosine kinase inhibitor resistance. Network analysis revealed connectivity among somatic variants in unresponsive cases that defined highly connected modules including PI3K/Akt (KEGG-hsa04151), ERBB (KEGG-hsa04012), Ras (KEGG-hsa04014), Notch (KEGG-hsa04330), androgen receptor signaling pathways (GO:0030521), and mismatch repair (KEGG-hsa03430). Altogether, these results provide support to the concept that, despite the lack of common mutations in patients not attaining pCR, the mutations fell into several, and druggable shared functional categories. Our study has also several limitations. The power of analysis was compromised because of small capacity, and number of events. Moreover, it suffered the lack of sufficient pre-treatment tumor tissue for microdissection, and mutation analysis to address the issue of intratumor spatial heterogeneity. Furthermore, additional studies are warranted to explore the mechanisms resulting in neoadjuvant chemotherapy-induced mutational changes, the functional evaluation of emerging mutations involved in potential resistance, as well as proliferation advantage, and the value of immune response to conventional chemotherapy. Importantly, the costs of serial sequencing assays preclude its clinical implication, which might be overcome with technology improvement. Given that targeted gene sequencing dynamics might serve as a potential marker of

chemo-sensitivity, its clinical value should be further validated in larger series of patients with TNBC.

We next challenged the use of ctDNA as a tool to anticipate response to neoadjuvant chemotherapy, establish distinct prognostic groups after surgery, optimize follow-up by identifying recurrent patients in advance, and the use of CTCs to explore druggable targets at disease progression.

Few studies have evaluated whether ctDNA analysis can be informative of response to neoadjuvant chemotherapy. In a prospective cohort of 101 patients with different breast cancer subtypes (84 treated with neoadjuvant chemotherapy), ctDNA was detected before any treatment in 41 cases and significantly associated with poor prognosis. Notably, ctDNA was found to anticipate the diagnosis of overt metastases in 16 cases [Garcia-Murillas I et al., 2015]. An ancillary ctDNA analysis of the NeoALTTO trial showed that baseline detection of *PIK3CA,* or *TP53* mutations in plasma samples of 28 patients was associated with poor response to anti-HER2 targeted therapy [Rothé F et al., 2019]. More recently, ctDNA, singly and in association with CTCs, proved to be associated with disease outcome on more than 100 cases in a pre-planned correlative study from a phase 2 randomized clinical trial on TNBC patients treated with neoadjuvant chemotherapy [Radovich M et al., 2020]. In terms of monitoring treatment response, reduced levels of multiple plasma mutations in 22 patients have been recently correlated with increased chances of attaining pCR [McDonald BR et al., 2019].

In this study, by implementing a pragmatic approach with primary tumor targeted-gene sequencing and patient-specific point mutations ddPCR detection, we analyzed a homogeneous cohort of 31 TNBC patients to first assess the clinical value of ctDNA at baseline, during and after treatment with curative intent. Our basal ctDNA detection rate of 77% was comparable with the published literature [Riva F et al., 2017; McDonald BR et al., 2019; Li S et al., 2020; Garcia-Murillas I

et al., 2020], post-treatment detection rate of 43.4% was, however, higher and likely reflects the fact that we had several patients with large primary tumor size, and positive nodal status at initial presentation, which also justifies the low rate (13%) of pCR reported.

A reduction of ctDNA up to undetectable levels occurred at mid-treatment and continued until surgery in just over half of the patients. This finding, which is consistent with previous studies [Riva F et al., 2017; Kim JY et al., 2017], is only apparently counter-intuitive, as ctDNA is expected to become undetectable only for patients achieving pCR. One plausible mechanism is that chemotherapy kills the dividing cells most likely to contribute to ctDNA release and, because the half-life of ctDNA in the blood is short, leaves behind a tumor less prone to ctDNA release at our midcourse sampling time [Lo YM et al., 1999]. Consistent with this hypothesis, the findings reported in the metastatic setting show that patients with stable disease on-treatment presented decreased levels of ctDNA despite the lack of tumor response [Forshew T et al., 2012; Dawson SJ et al., 2013]. Hence, the same could happen in the neoadjuvant setting with the important difference that patients with non metastatic BC present lower ctDNA levels, which instead of reducing disappeared during treatment. Besides, ctDNA levels were detectable in patients who progressed during treatment, suggesting a role for ctDNA analysis in identifying tumor progression, and assaying response to neoadjuvant chemotherapy, as increasing ctDNA may be an indicator not only of tumor growth during treatment but of increased risk of recurrence.

Our results suggest that post-treatment ctDNA is a surrogate for the emergence of relapse with 6/10 patients with detectable ctDNA at the end of neoadjuvant chemotherapy developing a breast cancer event. Remarkably, despite the small sample size, detection of ctDNA after neoadjuvant chemotherapy retained its prognostic significance even after adjusting for other clinico-pathological variables,

including Ki-67, whose reduction has been associated with better prognosis in patients who do not obtain the pCR response [Burcombe R et al., 2006;
Jones RL et al. 2009], as 83% of non relapsed cases despite persistently high Ki-67 levels were negative for ctDNA, and 30% of relapsed patients despite decreased levels of Ki-67 were ctDNA positive. Hence, our data suggest that the evaluation of ctDNA after neoadjuvant chemotherapy could act as a clinically available tool that might allow clinicians to stratify patients into those who could benefit from “complementary” treatment, in agreement with recently published results [Radovich M et al., 2020]. Specifically, TNBC patients with persistent ctDNA levels at the end of neoadjuvant chemotherapy could benefit from treatment intensification, or alternative therapeutic strategies in an attempt to prevent the development of metastases.

Next, we showed that in 83% of cases ctDNA preceded clinical detection of distant metastases by 8.9 months (range, 6.5-13.1months), and with excellent specificity. Therefore, ctDNA is not to be considered another prognostic factor that "on average" associates with the prognosis, but is rather able to predict at the individual patient level whether the event will happen or not, and the “lead time” found, which is consistent with other reports, represents a unique window of opportunity for the introduction of non–cross-resistant therapies to prevent overt clinical relapse.

Our ctDNA analysis was restricted to the known mutation profile of the primary tumor, which could be considered a limitation as clonal evolution cannot be studied. However, focusing on the known mutation profile in the tumor may minimize the risk of false positives particularly in light of recent reports identifying plasma somatic mutations arising from clonal hematopoiesis [Razavi P et al., 2019]. On the other hand, the analysis with the Oncomine Pan-Cancer Cell-Free assay showed that primary tumors and metastases exhibit high genomic

concordance at the plasmatic level, though additional druggable mutations such as MAPK may arise.

In addition to ctDNA detection, using the marker-independent Parsortix® approach for CTC-enrichment coupled with positive and negative selection with the DEPArray™, we showed that CTCs are non conventional (*i.e.*, non epithelial) in most of recurrent cases and would not have been detected by any of the commercially available epithelial marker approach, including Cell Search® [Reduzzi C et al., 2017]. Notably, analysis for copy number alterations by NGS displayed a unique spectrum of genetic abnormalities, including gain/loss of chromosomes 10 and chromosome 21q. Although changes in gene copy number, large and small in scale, contributed to population diversity, our analysis revealed a network among genomic alterations in relapsed cases that defined highly connected modules including HER/PI3K/Ras/JAK signalling and immune response. Altogether, these results provide support to the concept that despite the lack of common CNAs in patients who progressed after NAC, they fell into several and more importantly druggable shared functional categories.

The fact that patients were prospectively recruited for the purpose of these analyses, which were conducted in the same laboratory with uniform methodology, represents the major strengths of our study. However, there are also several limitations to consider, some of which inherent to the observational design of the study, the small sample size and number of pre-NAC samples collected, and the variable (*i.e.*, non standardized) timing for post-surgical blood drawings, that could affect the evaluation of relapse/progression anticipation. Importantly, the costs of serial sequencing assays and single-CTC analyses preclude their clinical routinary application, which might be overcome by technology improvements. Although the data should be interpreted with caution, emerging findings - including our results -

may guide the direction of future studies since risk adapted treatment strategies continue to be a research priority.

In conclusion, our findings support blood-based genomic analyses as complementary tools to optimize monitoring and to guide therapy in TNBC patients treated with NAC and add and integrate previous studies demonstrating the clinical validity of ctDNA. Prospective trials are ongoing and will address the clinical utility of incorporating such strategies into routine clinical practice.

CHAPTER 5.

CONCLUSIONS AND FUTURE PERSPECTIVES

The recent approval of the first companion diagnostic assay to detect PIK3CA mutations both in tissue and blood samples has bridged the role of liquid biopsy in breast cancer management to the fore. However, its value in early-stage disease is yet to be clearly defined. Although our results are preliminary and need confirmation in larger settings, they suggest that circulating tumor DNA and tumor cells may be used to i) stratify patients at different risk of relapse, ii) monitor response to treatment, and iii) unravel treatment resistance mechanisms, thus discovering new putative treatment targets to be further validated, specifically:

- In DCIS of the breast, it is possible to detect ctDNA at the initial diagnosis in spite of technical problems, such as low cellularity and DNA quality. ctDNA does not seem to be associated with clinico-pathological features used in clinical practice for the definition of prognosis. Hence, this result suggests that ctDNA could carry an alternative, and complementary information to that of already known prognostic factors. Notably, patients with relapse showed in their blood the same mutations found at initial diagnosis, and this suggests that ctDNA could be extended to preinvasive form and used in the same way as in invasive cancer.These results are also worth of consideration in light of the development of ctDNA in screening of healthy people for invasive carcinomas .
- In triple-negative breast cancer, the implementation of targeted NGS in the context of neoadjuvant clinical setting revealed that although recurrent mutations are few, unresponsive cases are more likely to present with alterations in immune response. This finding is important in light of the current development of immunotherapy of solid tumors, including breast cancer. Neoadjuvant chemotherapy tends to select cases enriched with alterations in the druggable signalling pathways. Importantly, cases showing the same

mutational profile before and after treatment and thus unperturbed by treatment showed a dismal prognosis.

- Few studies have evaluated liquid biopsy in the context of neoadjuvant chemotherapy, the majority are retrospective, and analyzed patients who differed for disease stage, type of treatment and/or with a limited number of blood samples collected over time. Our prospective mono-institutional study enrolling 42 triple negative breast cancer patients, treated with anthracycline/taxane/platin-based, and sampled on average seven times prior, during and after treatment, provides the following: ctDNA is detected in 77% of cases prior to NAC; patients with ctDNA at the end of NAC are more than twice as likely to relapse as those with undetectable ctDNA; detection of ctDNA during follow-up antedates clinical overt metastases up to 13 months; at the time of clinical progression, molecular characterization of CTCs, and direct targeted sequencing of ctDNA revealed therapeutically exploitable genomic alterations.

This study also highlighted some limitations of the current workflow that need to be improved. In particular, i) mutational profiling is currently not feasible for all patients and needs to be improved by processing a larger volume of blood or by optimizing DNA analysis techniques, ii) optimal timing of blood sampling has still to be established, and iii) extensive use in clinical practice of direct sequencing of plasma DNA is currently prevented by its low sensitivity. Additional research is also needed to elucidate why ctDNA levels vary among cancer patients, such as determining whether this variability is due to tumor biology and/or other factors such as DNA clearance from plasma.

Our findings support the analysis of ctDNA and CTCs for monitoring TNBC patient response to NAC, assisting de-/escalating post-NAC strategies, anticipating the diagnosis of overt metastases during follow-up, and characterizing recurrent

disease with the ultimate goal to foster personalized management in the breast cancer subtype which is still defined by its lack of targets.

At the time of writing, a prospective study is currently underway at our Institution to evaluate ctDNA as a tool for early detection of disease recurrence and as a tracer for the introduction of additional new therapy. This and the other six currently ongoing prospective trials worldwide, registered at www.clinicaltrials.gov, will help to address clinically relevant questions, with a focus on using ctDNA to allow early adaptation of further treatment attempts, as well as on measuring the impact of ctDNA analysis on long-term outcomes.

CHAPTER 6.

REFERENCES

- Abba MC, Gong T, Lu Y, Lee J, Zhong Y, Lacunza E, et al. A molecular portrait of high-grade ductal carcinoma in situ. *Cancer Res.* 2015; 75: 3980-3990.
- AIRTUM-AIOM, I numeri del cancro in Italia 2020. Available at https://www.registri-tumori.it
- Alix-Panabieres C, Pantel K. Circulating tumor cells: liquid biopsy of cancer. *Clin Chem.* 2013;59:110-118. doi: 10.1373/clinchem.2012.194258.
- Appierto V, Di Cosimo S, Reduzzi C, Pala V, Cappelletti V, Daidone MG. How to study and overcome tumor heterogeneity with circulating biomarkers: The breast cancer case. *Semin Cancer Biol*. 2017;44:106-116. doi: 10.1016/j.semcancer.2017.04.007.
- Arteaga CL, Engelman JA. ERBB receptors: from oncogene discovery to basic science to mechanism-based cancer therapeutics. *Cancer Cell.* 2014; 25:282-303. doi: https://doi.org/10.1016/j.ccr.2014.02.025.
- Awasthee N, Rai V, Chava S, Nallasamy P, Kunnumakkara AB, Bishayee A, et al. Targeting IκappaB kinases for cancer therapy. *Semin Cancer Biol.* 2019; 56: 12–24.
- Baili P, Torresani M, Agresti R, Rosito G, Daidone MG, Veneroni S, et al. A Breast Cancer Clinical Registry in An Italian Comprehensive Cancer Center: An Instrument for Descriptive, Clinical, and Experimental Research. *Tumori.* 2015; 101:440-446.
- Balko JM, Giltnane JM, Wang K, Schwarz LJ, Young CD, Cook RS, et al. Molecular profiling of the residual disease of triple-negative breast cancers after neoadjuvant chemotherapy identifies actionable therapeutic targets. *Cancer Discov.* 2014;4: 232-245.

- Banerjee S, Gusho E, Gaughan C, Dong B, Gu, X, Holvey-Bates E, et al. OAS-RNase L innate immune pathway mediates the cytotoxicity of a DNA-demethylating drug. *Proc Natl Acad Sci USA* 2019;12: 5071-76.
- Bardia A, Baselga J. Neoadjuvant therapy as a platform for drug development and approval in breast cancer. *Clin Cancer Res*. 2013;19(23):6360-70. doi: 10.1158/1078-0432.CCR-13-0916.
- Baselga J, Albanell J. Mechanism of action of anti-HER2 monoclonal antibodies. *Ann Oncol*. 2001;12 Suppl 1:S35-41. doi: 10.1093/annonc/12.suppl_1.s35.
- Bidard FC, Michiels S, Riethdorf S, Mueller V, Esserman LJ, Lucci A, et al. Circulating tumor cells in breast cancer patients treated by neoadjuvant chemotherapy: a meta-analysis. *J Natl Cancer Inst*. 2018; 110, 560–567.
- Bidard FC, Peeters DJ, Fehm T, Nolé F, Gisbert-Criado R, Mavroudis D, et al. Clinical validity of circulating tumour cells in patients with metastatic breast cancer: a pooled analysis of individual patient data. *Lancet Oncol*. 2014;15: 406–414.
- Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med.* 2018; 24: 541–550.
- Blainey PC. The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol Rev*. 2013;37(3):407-27. doi: 10.1111/1574-6976.12015.
- Bleeker FE, Felicioni L, Buttitta F, Lamba S, Cardone L, Rodolfo M, et al. AKT1(E17K) in human solid tumours. *Oncogene* 2008;27:5648–50.

- Bose R, Kavuri SM, Searleman AC, Shen W, Shen D, Koboldt DC, et al. Activating HER2 mutations in HER2 gene amplification negative breast cancer. *Cancer Discov* 2013;3:224–37.
- Bray F, Ferlay J, Soerjomataram I, Siegel R, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018; 68: 394-424.
- Bronkhorst AJ, Ungerer V, Holdenrieder S.The emerging role of cell-free DNA as a molecular marker for cancer management. *Biomol Detect Quantif*. 2019;17:100087. doi: 10.1016/j.bdq.2019.100087.
- Burcombe R, Wilson GD, Dowsett M, Khan I, Richman PI, Daley F, Detre S, Makris A.Evaluation of Ki-67 proliferation and apoptotic index before, during and after neoadjuvant chemotherapy for primary breast cancer. *Breast Cancer Res.* 2006;8(3):R31. doi: 10.1186/bcr1508.
- Cai Y, Sablina AA. Cancer-associated chromosomal deletions: Size makes a difference. *Cell Cycle*. 2016 Nov;15(21):2850-2851. doi: 10.1080/15548627.2016.1204869.
- Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* 2012;490:61–70.
- Carpi FM, Di Pietro F, Vincenzetti S, Mignini F, Napolioni V. Human DNA extraction methods: patents and applications. *Recent Pat DNA Gene Seq*. 2011;5(1):1-7. doi: 10.2174/187221511794839264.
- Casasent AK, Edgerton M, Navin NE. Genome evolution in ductal carcinoma in situ: invasion of the clones. *J Pathol.* 2017;241(2):208-218.

- Cheang MC, Chia SK, Voduc D, et al. Ki67 Index, HER2 Status, and Prognosis of Patients With Luminal B Breast Cancer. *J Natl Cancer Inst*. 2009:101:736–50.
- Chicard M, Colmet-Daage L, Clement N, Danzon A, Bohec M, et al. Whole-Exome Sequencing of Cell-Free DNA Reveals Temporo-spatial Heterogeneity and Identifies Treatment-Resistant Clones in Neuroblastoma. *Clin Cancer Res*. 2018;24(4):939-949. doi: 10.1158/1078-0432.
- Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science*. 2018;359:926-930. doi: 10.1126/science.aar3247.
- Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet.* 2014;384(9938):164-72. doi: 10.1016/S0140-6736(13)62422-8.
- Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med*. 2004;351:781–791.
- Cristofanilli M, Hayes DF, Budd GT, Ellis MJ, Stopeck A, Reuben JM, Doyle GV, et al. Circulating tumor cells: a novel prognostic factor for newly diagnosed metastatic breast cancer. *J Clin Oncol*. 2005;23:1420–1430.
- Cronin M, Ross JS. Comprehensive next-generation cancer genome sequencing in the era of targeted therapy and personalized oncology. *Biomark Med*. 2011; 5:293–305.
- Curigliano G, Burstein HJ, Winer EP, Gnant M, Dubsky P, Loibl S, et al. De-escalating and escalating treatments for early-stage breast cancer: the St. Gallen International Expert Consensus Conference on the Primary Therapy of

Early Breast Cancer 2017. *Ann Oncol*. 2017;28:1700-1712. doi: 10.1093/annonc/mdx308.

- Dawson SJ, Rosenfeld N, Caldas C. Circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med*. 2013;369(1):93-4. doi: 10.1056/NEJMc1306040.

- De Luca F, Rotunno G, Salvianti F, Galardi F, Pestrin M, Gabellini S, et al. Mutational analysis of single circulating tumor cells by next generation sequencing in metastatic breast cancer. *Oncotarget* 2016; 7: 26107.doi: 10.1016/j.cell.2018.03.041.

- Di Cosimo S , Appierto V , Ortolan E , Dell'Angelo F, Silvestri M , Bianchi G, et al. Circulating tumor DNA and disease recurrence in early stage breast cancer: From a case-control study to a prospective longitudinal trial. *Ann Oncol.* 2019; 30 (Suppl 3): iii28-29.

- Dieci MV, Criscitiello C, Goubar A, Viale G, Conte P, Guarneri V, et al. Prognostic value of tumor-infiltrating lymphocytes on residual disease after primary chemotherapy for triple-negative breast cancer: A retrospective multicenter study. *Ann Oncol* 2014; 25: 611–618. doi: 10.1038/modpathol.2017.21.

- Dillon DA, Guidi AJ, Schnitt SJ. Pathology of invasive breast cancer. In: Diseases of the Breast, 4th ed, Harris JR, Lippman ME, Morrow M, Osborne CK (Eds), Lippincott, Williams and Wilkins, Philadelphia 2009. p.386

- Domínguez-Vigil IG, Moreno-Martínez AK, Wang JY, Roehrl MHA, Barrera-Saldaña HA. The dawn of the liquid biopsy in the fight against cancer. *Oncotarget*. 2017 Dec 8;9(2):2912-2922. doi: 10.18632/oncotarget.23131.

- Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials. *Lancet Oncol*. 2018;19(1):27-39.
- Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Overview of the randomized trial of radiotherapy in duct carcinoma in situ of the breast. J Natl Cancer Inst Monog 2010; 41:162-177.
- Elazezy M, Joosse SA. Techniques of using circulating tumor DNA as a liquid biopsy component in cancer management. *Comput Struct Biotechnol J.* 2018;16,370–378.
- Ellsworth RE, Blackburn HL, Shriver CD, Soon-Shiong P, Ellsworth DL. Molecular heterogeneity in breast cancer: State of the science and implications for patient care. *Semin Cell Dev Biol* 2017;64:65–72
- Forshew T, Murtaza M, Parkinson C, Gale D, Tsui DW, Kaper F, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Sci Transl Med.* 2012;4:136ra168.
- Gagan J, Van Allen EM. Next-generation sequencing to guide cancer therapy. *Genome Med.* 2015;7(1):80. doi: 10.1186/s13073-015-0203-x.
- Gale D, Lawson ARJ, Howarth K, Madi M, Durham B, Smalley S, et al. Development of a highly sensitive liquid biopsy platform to detect clinically-relevant cancer mutations at low allele fractions in cellfree DNA, *PLoS One* 2018; 13:e0194630.
- Garcia-Murillas I, Chopra N, Comino-Méndez I, Beaney M, Tovey H, Cutts RJ, et al. Assessment of Molecular Relapse Detection in Early-Stage Breast Cancer. *JAMA Oncol.* 2019;5:1473-1478. doi: 10.1001/jamaoncol.2019.1838.

- Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci Transl Med.* 2015;7:302ra133.
- Gatza ML, Silva GO, Parker JS, Fan C, Perou CM. An integrated genomics approach identifies drivers of proliferation in luminal-subtype human breast Cancer. *Nature Genetics.* 2014;46:1051–1059. doi: https://doi.org/10.1038/ng.3073.
- Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. Available from: https://gco.iarc.fr/today (accessed March 19, 2021).
- Gray ES, Witkowski T, Pereira M, Calapre L, Herron K, Irwin D et al. Genomic analysis of circulating tumor DNA using a melanoma-specific UltraSEEK oncogene panel. *J Mol Diagn*. 2019; 21, 418–26.
- Habli Z, AlChamaa W, Saab R, Kadara H, Khraiche ML. Circulating Tumor Cell Detection Technologies and Clinical Utility: Challenges and Opportunities. *Cancers (Basel)*. 2020 Jul 17;12(7):1930. doi: 10.3390/cancers12071930.
- Heitzer E, Haque IS, Roberts CES, Speicher MR. Current and future perspectives of liquid biopsies in genomics-driven oncology. *Nat Rev Genet*. 2019;20(2):71-88. doi: 10.1038/s41576-018-0071-5.
- Jones RL, Salter J, A'Hern R, Nerurkar A, Parton M, Reis-Filho JS, Smith IE, Dowsett M. The prognostic significance of Ki67 before and after neoadjuvant chemotherapy in breast cancer. Breast Cancer Res Treat. 2009 Jul;116(1):53-68. doi: 10.1007/s10549-008-0081-7.

- Kanwar N, Hu P, Bedard P, Clemons M, McCready D, Done SJ. Identification of genomic signatures in circulating tumor cells from breast cancer. *Int J Cancer*. 2015;137:332-44. doi: 10.1002/ijc.29399.
- Kaufmann M, von Minckwitz G, Mamounas EP, et al. Recommendations from an international consensus conference on the current status and future of neoadjuvant systemic therapy in primary breast cancer. *Ann Surg Oncol*. 2012;19:1508-1516.
- Kim C, Gao R, Sei E, Barndt R, Hartman J, Hatschek T, et al. Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. *Cell*. 2018;173(4):879-893.e13.
- Kim JY, Park D, Son DS, Nam SJ, Kim SW, Jung HH, et al. Circulating tumor DNA shows variable clonal response of breast cancer during neoadjuvant chemotherapy. *Oncotarget.* 2017;8(49):86423-86434.
- Kustanovich A, Schwartz R, Peretz T, Grinshpun A. Life and death of circulating cell-free DNA. *Cancer Biol Ther.* 2019; 20,1057-1067.
- Lanman RB, Mortimer SA, Zill OA, Sebisanovic D, Lopez R, Blau S, 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:e0140712.
- Lehmann BD, Jovanovi´c B, Chen X, Estrada MV, Johnson KN, Shyr Y, et al. Refinement of triple-negative breast cancer molecular subtypes: implications for neoadjuvant chemotherapy selection. *PLoS ONE* 2016; 11: e0157368.
- Leon SA, Shapiro B, Sklaroff DM, Yaros MJ. Free DNA in the serum of cancer patients and the effect of therapy. *Cancer Res*.1977;37: 646–650.

- Li S, Lai H, Liu J, Liu Y, Jin L, Li Y, et al. Circulating tumor DNA predicts the response and prognosis in patients with early breast cancer receiving neoadjuvant chemotherapy. *JCO Precis Oncol*. 2020;4:244-257.
- Li Z, Zhang X, Hou C, Zhou Y, Chen J, Cai H, et al. Comprehensive identification and characterization of somatic copy number alterations in triple-negative breast cancer. *Int J Oncol*. 2020; 56(2):522-530.
- Lievens A, Jacchia S, Kagkli D, Savini C, Querci M. Measuring digital PCR quality: performance parameters and their optimization. *PLoS One*. 2016;11(5):e0153317. doi: 10.1371/journal.pone.0153317.
- Lips EH, Michaut M, Hoogstraat M, Mulder L, Besselink NJ, Koudijs MJ, et al. Next generation sequencing of triple negative breast cancer to find predictors for chemotherapy response. *Breast Cancer Res*. 2015;17:134.
- Litviakov NV, Ibragimova MK, Tsyganov MM, Doroshenko AV, Garbukov EY, Slonimskaya EM. Neoadjuvant Chemotherapy Induces the Appearance of New Copy Number Aberrations in Breast Tumor and is Associated with Metastasis. *Curr Cancer Drug Targets* 2020;20(9):681-688.doi: 10.2174/1568009620666200506104523.
- Lo YM, Zhang J, Leung TN, Lau TK, Chang AM, Hjelm NM. Rapid clearance of fetal DNA from maternal plasma. *Am J Hum Genet.* 1999;64:218-224.
- Luen SJ, Salagado M, Dieci V, Vingiani A, Curigliano G, Gould RE, et al. Prognostic implications of residual disease tumor-infiltrating lymphocytes and residual cancer burden in triple-negative breast cancer patients after neoadjuvant chemotherapy. *Ann Oncol.* 2019; 30: 236–242.

- Mandel P, Metais P. Les acides nucléiques du plasma sanguin chez l'homme. *Comptes rendus des seances de la Soc de biologie et de ses filiales*. 1948;142:241–243.
- Martelotto LG, Baslan T, Kendall J, Geyer FC, Burke KA, Spraggon L, et al. Whole-genome single-cell copy number profiling from formalin-fixed paraffin-embedded samples. *Nat Med.* 2017;23:376-385. doi: 10.1038/nm.4279.
- Martelotto LG, Ng CKY, Piscuoglio S, Weigelt B, Reis-Filho JS. Breast cancer intra-tumor heterogeneity. *Breast Cancer Res.* 2014; 16: 210.
- Martínez-Pérez C, Turnbull AK, Ekatah GE, Arthur LM, Sims AH, Thomas JS, Dixon JM. Current treatment trends and the need for better predictive tools in the management of ductal carcinoma in situ of the breast. *Cancer Treat Rev.* 2017;55:163-172. doi: 10.1016/j.ctrv.2017.03.009.
- McDonald BR, Contente-Cuomo T, Sammut SJ, Odenheimer-Bergman A, Ernst B, Perdigones N, et al. Personalized circulating tumor DNA analysis to detect residual disease after neoadjuvant therapy in breast cancer. *Sci Transl Med.* 2019;11(504):eaax7392.
- Meric-Bernstam F, Mills GB. Overcoming implementation challenges of personalized cancer therapy. *Nat Rev Clin Oncol*. 2012;9:542-8. doi: 10.1038/nrclinonc.2012.127.
- Merker JD, Oxnard GR, Compton C, Diehn M, Hurley P, Lazar AJ, et al. Circulating tumor DNA analysis in patients with cancer: American Society of Clinical Oncology and College of American Pathologists joint review
- Morrow M, Katz SJ. Addressing overtreatment in DCIS: what should physicians do now? *J Natl Cancer Inst*. 2015;107(12):djv290.

- Nachmanson D, Steward J, Yao H, Officer A, Jeong E, O'Keefe TJ, et al. Mutational profiling of micro-dissected pre-malignant lesions from archived specimens. *BMC Med Genomics.* 2020; 13: 173.

- Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med.* 2014 May;20(5):548-54. doi: 10.1038/nm.3519.

- Ng CK, Schultheis AM, Bidard FC, Weigelt B, Reis-Filho JS. Breast cancer genomics from microarrays to massively parallel sequencing: Paradigms and new insights. *J Natl Cancer Inst* 2015;107:10

- Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, et al. Breast Cancer Working Group of the International Cancer Genome Consortium: Mutational processes molding the genomes of 21 breast cancers. *Cell.* 2012;149:979-993.

- Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 2016;534:47–54.

- Obermayr E, Maritschnegg E, Speiser P, Singer C, Schuster E, Holzer B, et al. Circulating rare cells enable highly efficient cancer detection. In *Proceedings of the 106th Annual Meeting of the American Association for Cancer Research*, Philadelphia, PA, USA, 18–22 April 2015.

- Ono Y, Sugitani A, Karasaki H, Ogata M, Nozaki R, Sasajima J, et al. An improved digital polymerase chain reaction protocol to capture low-copy KRAS mutations in plasma cell-free DNA by resolving 'subsampling' issues. *Mol Oncol.* 2017;11(10):1448-1458. doi: 10.1002/1878-0261.12110.

- Pang JB, Savas P, Fellowes AP, Mir Arnau G, Kader T, Vedururu R, et al. Breast ductal carcinoma in situ carry mutational driver events representative of invasive breast cancer. *Mod Pathol.* 2017 Jul;30(7):952-963.

- Pantel K, Brakenhoff RH and Brandt B. Detection, clinical relevance and specific biological properties of disseminating tumour cells. *Nat Rev Cancer.* 2008; 8: 329–340.

- Pashayan N, Antoniou AC, Ivanus U, Esserman LJ, Easton DF, French D, et al. Personalized early detection and prevention of breast cancer: ENVISION consensus statement. *Nat Rev Clin Oncol*. 2020; 17(11): 687–705.

- Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, Pugh M, Jones L, Russell R, Sammut SJ, Tsui DW, Liu B, Dawson SJ, Abraham J, Northen H, Peden JF, Mukherjee A, Turashvili G, Green AR, McKinney S, Oloumi A, Shah S, Rosenfeld N, Murphy L, Bentley DR, Ellis IO, Purushotham A, Pinder SE, Børresen-Dale AL, Earl HM, Pharoah PD, Ross MT, Aparicio S, Caldas C. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun*. 2016 May 10;7:11479. doi: 10.1038/ncomms11479.

- Prat A, Parker JS, Karginova O. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res. 2010;12(5):R68.

- Pribluda A, de la Cruz CC, Jackson EL. Intratumoral heterogeneity: from diversity comes resistance. *Clin Cancer Res.* 2015; 21:2916-23.

- Radovich M, Jiang G, Hancock BA, Chitambar C, Nanda R, Falkson C, et al. Association of Circulating Tumor DNA and Circulating Tumor Cells After

Neoadjuvant Chemotherapy With Disease Recurrence in Patients With Triple-Negative Breast Cancer: Preplanned Secondary Analysis of the BRE12-158 Randomized Clinical Trial. *JAMA Oncol*. 2020;6(9):1410-1415.

- Rao GC, Chianese D, Doyle GV, Miller MC, Russell T, Sanders RA Jr, Terstappen LW. Expression of epithelial cell adhesion molecule in carcinoma cells present in blood and primary and metastatic tumors. *Int J Oncol.* 2005;27(1):49-57.

- Rapoport BL, Anderson R. Realizing the Clinical Potential of Immunogenic Cell Death in Cancer Chemotherapy and Radiotherapy. *Int J Mol Sci*. 2019; 20: 959.

- Razavi P, Li BT, Brown DN, Jung B, Hubbell E, Shen R, et al. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. *Nat Med*. 2019;25(12):1928-1937.

- Reduzzi C, Motta R, Bertolini G, Miodini P, Martinetti A, Sottotetti E, et al. Development of a Protocol for Single-Cell Analysis of Circulating Tumor Cells in Patients with Solid Tumors. *Adv Exp Med Biol*. 2017;994:83-103. doi: 10.1007/978-3-319-55947-6_4.

- Riebensahm C, Joosse SA, Mohme M, Hanssen A, Matschke J, Goy Y, et al.Clonality of circulating tumor cells in breast cancer brain metastasis patients. *Breast Cancer Res*. 2019;21:101. doi: 10.1186/s13058-019-1184-2.

- Riethdorf S, Fritsche H, Müller V, Rau T, Schindlbeck C, Rack B, et al. Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: a validation study of the CellSearch system. *Clin Cancer Res.* 2007; 13:920–928.

- Riva F, Bidard FC, Houy A, Saliou A, Madic J, Rampanou A, et al. Patient specific circulating tumor DNA detection during neoadjuvant chemotherapy in triple-negative breast cancer. *Clin Chem*. 2017;63:691-9.
- Rossing M, Sørensen C, Ejlertsen B, Ejlertsen B, Nielsen FC. Whole genome sequencing of breast cancer. *APMIS*. 2019 May;127(5):303-315. doi: 10.1111/apm.12920.
- Rothé F, Silva MJ, Venet D, Campbell C, Bradburry I, Rouas G, et al. Circulating tumor DNA in HER2-amplified breast cancer: a translational research substudy of the NeoALTTO phase III trial. *Clin Cancer Res*. 2019;25(12):3581-3588.
- Russnes HG, Vollan HKM, Lingjærde OC, Krasnitz A, Lundin P, Naume B, et al. Genomic architecture characterizes tumor progression paths and fate in breast cancer patients. *Sci Transl Med*. 2010;2:38ra47. doi: 10.1126/scitranslmed.3000611.
- Salgado R, Denkert C, Demaria S, Sirtaine, N, Klauschen, F, Pruneri, G, et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: Recommendations by an International TILs Working Group 2014. *Ann Oncol.* 2015, 26, 259–271.
- Salvatorelli L, Puzzo L, Vecchio GM, Caltabiano R, Virzì V, Magro G. Ductal carcinoma in situ of the breast: an update with emphasis on radiological and morphological features as predictive prognostic factors.
  Cancers (Basel). 2020;12(3):609. doi: 10.3390/cancers12030609.
- Scheinin I, Sie D, Bengtsson H, van de Wiel MA, Olshen AB, van Thuijl HF, et al. DNA copy number analysis of fresh and formalin-fixed specimens by

shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res.* 2014;24:2022-32. doi: 10.1101/gr.175141.114.

- Shah SP, Roth A, Goya R, Oloumi A, Ha G, Zhao Y, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature*. 2012;486:395-9. doi: 10.1038/nature10933.
- Shee K, Muller KE, Marotti J, Miller TW, Wells WA, Tsongalis GJ. Ductal carcinoma in situ biomarkers in a precision medicine era: current and future molecular-based testing. *Am J Pathol*. 2019;189:956-965. doi: 10.1016/j.ajpath.2018.08.020.
- Shree T, Olson OC, Elie BT, Kester JC, Garfall AL, Simpson K, et al. Macrophages and cathepsin proteases blunt chemotherapeutic response in breast cancer. *Genes Dev.* 2011; 25: 2465-79.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018;68:7e30.
- Sieuwerts AM, Kraan J, Bolt J, van der Spoel P, Elstrodt F, Schutte M, et al. Anti-epithelial cell adhesion molecule antibodies and the detection of circulating normal-like breast tumor cells. *J Natl Cancer Inst.* 2009;101(1):61-6. doi: 10.1093/jnci/djn419.
- Siu AL, U.S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med.* 2016; 164:279–296
- Siu LL. Accounting for tumor heterogeneity in the development of predictive biomarkers. *Clin Adv Hematol Oncol.* 2013;11:312–313.

- Sorenson GD, Pribish DM, Valone FH, Memoli VA, Bzik DJ, Yao SL. Soluble normal and mutated DNA sequences from single-copy genes in human blood. *Cancer Epidemiol Biomarkers Prev.* 1994;3:67-71.
- Sørlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*. 2001;98(19):10869-74.
- Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A*. 2003;100(14):8418-23.
- Speicher MR, Pantel K. Tumor signatures in the blood. *Nat Biotechnol*. 2014;32:441–443. doi: 10.1038/nbt.2897.
- Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. Oslo Breast Cancer Consortium: The landscape of cancer genes and mutational processes in breast cancer. *Nature.* 2012;486:400-404.
- Stroun M, Anker P, Maurice P, Lyautey J, Lederrey C, Beljanski M. Neoplastic characteristics of the DNA found in the plasma of cancer patients. *Oncology*. 1989;46:318–322.
- Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, et al. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat Genet* 2013;45:1439–45.
- van de Vijver MJ, He YD, van 't Veer LJ, et al. A Gene-expression signature as a predictor of survival in breast cancer. N Engl J Med. 2002;347:1999-2009.
- van Seijen M, Lips EH, Thompson AM, Nik-Zainal S, Futreal A, Hwang ES, et al; PRECISION team.Ductal carcinoma in situ: to treat or not to treat, that is the question. *Br J Cancer.* 2019;121(4):285-292. doi: 10.1038/s41416-019-0478-6.

- Veronesi U, Saccozzi R, Del Vecchio M, Banfi A, Clemente C, De Lena M, et al. Comparing radical mastectomy with quadrantectomy, axillary dissection, and radiotherapy in patients with small cancers of the breast. *N Engl J Med*. 1981;305:6–11.
- Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, Pacey S, Baird R, Rosenfeld N. Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nat Rev Cancer*. 2017;17(4):223-238. doi: 10.1038/nrc.2017.7.
- Wang Y, Navin NE**.** Advances and applications of single-cell sequencing technologies. *Mol Cell*. 2015;58:598-609. doi: 10.1016/j.molcel.2015.05.005.
- Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P, et al: Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat Med.* 2015;21:751-759.
- Yeo B, Dowsett M. Neoadjuvant endocrine therapy: Patient selection, treatment duration and surrogate endpoints. *Breast*. 2015;24 Suppl 2:S78-83. doi: 10.1016/j.breast.2015.07.019.
- Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16:284-287. doi:10.1089/omi.2011.0118
- Zardavas D, Irrthum A, Swanton C, Piccart M.Clinical management of breast cancer heterogeneity. *Nat Rev Clin Oncol*. 2015;12(7):381-94. doi: 10.1038/nrclinonc.2015.73.
- Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med*. 2017; 23:703–713.

❖ Zhang S, Wu T, Peng X, Liu J, Liu F, Wu S, et al. Mesenchymal phenotype of circulating tumor cells is associated with distant metastasis in breast cancer patients. *Cancer Manag Res.* 2017;9:691-700.

CHAPTER 7.

PUBLICATIONS

Publications related to the *"Liquid biopsy in breast cancer: molecular characterization and support to clinical management"* PhD project (Indexed in PubMed as of March 24, 2021)

- Ortolan E, Appierto V, Silvestri M, Miceli R, Veneroni S, Folli S, Pruneri G, Vingiani A, Belfiore A, Cappelletti V, Vismara M, Dell'Angelo F, De Cecco L, Bianchi GV, de Braud FG, Daidone MG, **Di Cosimo S**. Blood-based genomics of triple-negative breast cancer progression in patients treated with neoadjuvant chemotherapy. *ESMO Open*. 2021 Mar 17;6(2):100086. doi: 10.1016/j.esmoop.2021.100086.
- **Di Cosimo S**, Appierto V, Silvestri M, Ortolan E, De Cecco L, Veneroni S, Pruneri G, Vingiani A, Belfiore A, Scaperrotta G, Folli S, Daidone MG. Primary tumor somatic mutations in the blood of women with ductal carcinoma in situ of the breast. *Ann Oncol.* 2020;31(3):435-437.
- **Di Cosimo S**, Appierto V, Silvestri M, Pruneri G, Vingiani A, Perrone F, Busico A, Folli S, Scaperrotta G, de Braud FG, Bianchi GV, Cavalieri S, Daidone MG, Dugo M. Targeted-Gene Sequencing to Catch Triple Negative Breast Cancer Heterogeneity before and after Neoadjuvant Chemotherapy. *Cancers (Basel)*. 2019 Nov 8;11(11):1753. doi: 10.3390/cancers11111753.
- Reduzzi C, Vismara M, Gerratana L, Silvestri M, De Braud F, Raspagliesi F, Verzoni E, **Di Cosimo S**, Locati LD, Cristofanilli M, Daidone MG, Cappelletti V. The curious phenomenon of dual-positive circulating cells: Longtime overlooked tumor cells. *Semin Cancer Biol.* 2020; 60:344-350. doi: 10.1016/j.semcancer.2019.10.008.
- Gennaro M, Meneghini E, Baili P, Bravaccini S, Curcio A, De Santis MC, Lozza L, Listorti C, **Di Cosimo S**, Sant M, Folli S. High consistency between characteristics of primary intraductal breast cancer and subtype of subsequent

ipsilateral invasive cancer. *Tumori.* 2020 Feb;106(1):64-69. doi: 10.1177/0300891619867845.

- **Di Cosimo S**, Torri V, Porcu L. RE: Presence of Circulating Tumor Cells in High-Risk Early Breast Cancer During Follow-Up and Prognosis. *J Natl Cancer Inst*. 2019;111(11):1232-1233. doi: 10.1093/jnci/djz092.
- Rothé F, Silva MJ, Venet D, Campbell C, Bradburry I, Rouas G, de Azambuja E, Maetens M, Fumagalli D, Rodrik-Outmezguine V, **Di Cosimo S**, Rosa D, Chia S, Wardley A, Ueno T, Janni W, Huober J, Baselga J, Piccart M, Loi S, Sotiriou C, Dawson SJ, Ignatiadis M. Circulating Tumor DNA in HER2-Amplified Breast Cancer: A Translational Research Substudy of the NeoALTTO Phase III Trial. *Clin Cancer Res*. 2019 Jun 15;25(12):3581-3588. doi: 10.1158/1078-0432.CCR-18-2521. Epub 2019 Mar 12.PMID: 30862692
- Gennaro M, **Di Cosimo S**, Ardoino I, Veneroni S, Mariani L, Agresti R, Daidone MG, de Braud F, Apolone G, Biganzoli E, Demicheli R.Dynamics of the hazard for distant metastases after ipsilateral breast tumor recurrence according to estrogen receptor status: An analysis of 2851 patients. *Breast.* 2018 Aug;40:131-135. doi: 10.1016/j.breast.2018.05.004.
- Galli G, Bregni G, Cavalieri S, Porcu L, Baili P, Hade A, Di Salvo F, Sant M, Agresti R, Gennaro M, Folli S, De Santis MC, Paolini B, Carcangiu ML, de Braud F, **Di Cosimo S**. Neoadjuvant Chemotherapy Exerts Selection Pressure Towards Luminal Phenotype Breast Cancer. *Breast Care (Basel)*. 2017 Dec;12(6):391-394. doi: 10.1159/000479582.
- Appierto V, **Di Cosimo S**, Reduzzi C, Pala V, Cappelletti V, Daidone MG. How to study and overcome tumor heterogeneity with circulating biomarkers: The breast cancer case. *Semin Cancer Biol*. 2017;44:106-116. doi: 10.1016/j.semcancer.2017.04.007.

Presentations at International Meetings

- **Di Cosimo S** , Appierto V , Ortolan E , Dell'Angelo F, Silvestri M , Bianchi G, et al. Circulating tumor DNA and disease recurrence in early stage breast cancer: From a case-control study to a prospective longitudinal trial. Oral presentation, ESMO Breast Cancer, 2-4 May 2019, Berlin (Germany).

- **Di Cosimo S**, Silvestri M, Dugo M, Vismara M, Reduzzi C, Pruneri G, Folli S, Cappelletti V, Daidone MG. Primary tumor and circulating tumor cell (CTC) Copy Number Alterations (CNAs) in Triple Negative Breast Cancer (TNBC) patients (pts) treated with neoadjuvant chemotherapy (NAC). eposter 59P, ESMO Breast Cancer Virtual Meeting 2020.

- Daidone MG, **Di Cosimo S**, Veneroni S, Cascone F, De Cecco L, Dugo M, Folli S, Bianchi GV, Tamborini E, Busico A, Appierto V. Circulating tumor DNA detection anticipates disease recurrence in early stage breast cancer: A pilot study generating an observational confirmatory trial. Poster Session: P2-02-19. 2017 San Antonio Breast Cancer Symposium; December 5-9, 2017; San Antonio, Texas.

CHAPTER 8.

DECLARATION OF AUTHORSHIP

I hereby certify that the thesis I am submitting is entirely my own original work, except for the following parts:

- Recruitment of patients was also performed by Secondo Folli, Giulia Valeria Bianchi, Filippo Guglielmo de Braud, and Silvia Veneroni (Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy);
- Pathological review and DNA extraction from primary tumor tissue were performed by Giancarlo Pruneri, Andrea Vingiani, and Antonino Belfiore (Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy).
- Primary tumor NGS was performed by Loris de Cecco (Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy);
- DNA extraction from plasma and ctDNA analysis were also performed by Valentina Appierto, Elena Cavadini, and Elisa Ortolan;
- CTC characterization was performed by Vera Cappelletti, Patrizia Miodini, Rosita Mottadone, and Marta Vismara (Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy);
- Bioinformatic and statistical analyses were in charge of Marco Silvestri, and Matteo Dugo, and Rosalba Miceli (Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy).

APPENDIX

The pen moved fast, chasing the thoughts with difficulty. I still keep that yellow paper with blue lines, the kind you find at the reception of the convention center in Chicago. It is the treatment scheme for women with HER2 positive breast cancer applied today all over the world. At the bottom is the date 02/06/2007 and the signature of José Baselga.

A visionary of targeted therapy, Baselga has been instrumental in the development of drugs already in use to benefit millions of cancer patients such as trastuzumab, everolimus, lapatinib, pertuzumab, or in rapid clinical development for their great potential such as deruxetan.

Overcoming the division between preclinical and clinical research, the taboo of collaboration with the pharmaceutical industry, and academic self-reliance, Baselga has promoted translational research in the cooperative groups and scientific societies he has chaired.

A profound supporter of young people, he has shunned teaching ex cathedra in favor of training in the field: laboratory and animal encounters, medical staff in emergency rooms, hours and hours on the wards, interviews with the families of those who wouldn't make it, morning meetings with coffee in hand for multidisciplinary and molecular boards ante litteram. Believing in it up to the end, working hard, documenting, sharing, holding out, not settling and above all, hoping: Baselga was all of this.

He passed away on March 21, 2021 at the age of 61. He leaves us an extraordinary scientific legacy. It is up to us to make the best use of it, as he would have done. *Hasta siempre, Jefe!*

Printed in Milan on March 30, 2021