Establishment of PCR based methods for detection of ctDNA in blood

Bente Risberg

2013

HØGSKOLEN I OSLO OG AKERSHUS

Oslo University Hospital
Norwegian Radium Hospital

Establishment of PCR based methods for detection of ctDNA in blood

By

Bente Risberg

Master in Biomedicine
Department of Health Sciences
Thesis submitted for the Master`s degree in Biomedicine, 60 ECTS,
Oslo University Hospital, Institute for Cancer Research, Department of Genetics and
Oslo and Akershus University College of Applied Sciences

Supervisor Hege G. Russnes
Co-supervisors Anne Tierens and Sarah-Jane Dawson

21.05.2013

Acknowledgements

The work presented in this thesis was carried out at Cambridge Research Institute in UK from October 2012 to May 2013 for the Master degree in Biomedicine at Oslo and Akershus University College of Applied Sciences.

First I would like to thank the Department of Pathology and Laboratory for Molecular Pathology for giving me the opportunity to take this master, as a part time job since autumn 2009. Next I want to thank my three excellent supervisors, Hege Russnes (Dep. of Path/Dep. of Genetics), Anne Tierens (Det. of Path) and Sarah-Jane Dawson (Cambridge Research Institute, UK) for wonderful support. You are clever and inspiring! I also would like to thank the Radiumhospital Foundation and Astrid and Birger Torstedts Legat for financing my stay in Cambridge.

My thesis has been carried out in collaboration between several Departments and Institutes and I would like to thank Professor Anne-Lise Børresen-Dale at Department of Genetics and Professor Carlos Caldas at Cambridge Research Institute for giving me the possibility to visit Prof. Caldas group and stay there for half a year while working on my thesis. A special thanks to Sarah-Jane who was there for me one hundred percent, I never felt lost. I would also like to thank Caldas and the whole group for the warm welcome and especially Bernard Pereira who has aligned my sequences, Suet-Feung Chin who showed me IgV and Wendy Greenwood for taking care of my orders. I also would like to thank Nitzan Rosenfelds group for help and support, and especially Dana Tsui who has been my closest contact since Sarah-Jane went on maternity leave in December. Dana and also Davina Gale have taught me about digital PCR, and Tim Forshew has willingly shared both his experience and his new reagents for testing sequencing. Thanks to both Dana and Daniel Nebdal at Dep. of Genetics for helping me with figures and thanks to Christine Parkinson and James Brenton for providing me some of the patient material.

At last I want to thank my family for their never ending support. I also would like to thank friends and family for visiting me in Cambridge. A special thanks to Margunn Koksæter who looked after my apartment while I was gone, Helen Vålerhaugen and Paal Ruud for reading through the manuscript, giving me valuable feedback. And thanks to the whole group at

Laboratory for Molecular Pathology for their positive attitude despite busy days and lab leader Sarah L. Ariansen for being helpful and flexible all the way along

I have had an inspiring and interesting year! Thanks to everybody!

Bente Risberg

21-May-2013

Abstract

Cell-free DNA appears in the circulation when cells undergo apoptosis and necrosis. This is a natural process and small amounts of cell-free DNA can be found in the blood of healthy individuals. When tumors are present, a variable part of the cell-free DNA will derive from tumor. Monitoring ctDNA is attractive in solid cancers because a blood sample is less invasive than a biopsy and it might represent all lesions including metastasis. "Liquid biopsies" also give the opportunity for serial monitoring during therapy and follow up when tumor biopsies are difficult to achieve. Little is still known about the stability of ctDNA *in vitro*, so one of the subprojects in this thesis was to elucidate this. Cell-free DNA is extracted from plasma and ctDNA is often a minor fraction of the total plasma DNA. Our results showed that blood can be left for 24 hours in room temperature without significantly affecting the total plasma-DNA concentration. It was however not clear if ctDNA showed a slight decrease due to delayed processing or whether it was due to assay variation. More data needs to be collected. ERBB2 is amplified in 15-20 % of breast cancers and targeted therapy is available against its protein (HER2). Studies have shown that HER2-status might change during disease progression, so detection of the amplification in plasma would be desirable. Other methods like quantitative PCR (qPCR) and digital PCR (dPCR) have been used to detect the amplification, but these rely on the use of an unamplified reference gene. Breast cancers are heterogeneous and show a variety of genetic alterations. A common unamplified gene can therefore be hard to find. One of our subprojects detected ERBB2 amplification by targeted massively parallel sequencing (MPS). ERBB2 is a monoallelic event, and SNP ratios could be used to detect amplification. No reference gene was needed as the unamplified allele was the reference. The method achieved the same limit of detection as dPCR, 1.2 fold. Breast cancer also has few recurrent mutations. Massively parallel sequencing (MPS) has given the opportunity to detect the broad range of mutations that can be seen in breast tumors and to use these mutations as tumor markers for detection of ctDNA. The last subproject screened 17 genes frequently mutated in breast cancer with targeted MPS and detected somatic mutations in 94 % of the patients (49/52). 48 % of the patients had two mutations that could be tracked in serial blood samples taken during therapy. Both of the cases presented in this thesis showed evidence of subclonality. In one case, the KRAS mutation emerged during therapy. This might be due to acquisition of the mutation during treatment or that it belonged to minor subclone that was not detected at

presentation. This emphasizes the importance of detecting all subclones at disease presentation and not only tracking the dominant clone in the primary tumor. Monitoring ctDNA is an important means to detect therapy response and early relapse and might be a step towards more personalized medicine.

Sammendrag

Ved apoptose og nekrose frigis cellefritt DNA til sirkulasjonen. Dette er en naturlig prosess og små mengder fritt sirkulerende DNA finnes i blodet hos friske. Ved tilstedeværelse av tumor vil en del av det frie sirkulerende DNA stamme fra tumor og sirkulerende tumor DNA (ctDNA) vil gjerne være representativ for både primærtumor og eventuelle metastaser. Å ta en blodprøve er mindre invasivt enn å ta en biopsi, derfor er måling av ctDNA en attraktiv metode for å monitorere kreftsvulster. Hvor stabilt ctDNA er *in vitro* vet man ennå lite om, så ett av delprosjektene i denne masteroppgaven var å undersøke dette. Cellefritt DNA ekstraheres fra plasma og det er viktig å unngå leukocytt-lysering og frigjøring av villtype-DNA. Resultatene viste at blodprøvene kunne oppbevares i romtemperatur inntil 24 timer uten at den totale konsentrasjonen av cellefritt DNA ble signifikant endret. Det kunne se ut til at andelen ctDNA var svakt synkende, men materialet var for lite til å trekke en sikker konklusjon.

Et annet delprosjekt var deteksjon av *ERBB2*-amplifikasjon i plasma. *ERBB2* er amplifisert i 15-20 % av brystkrefttilfellene og det finnes målrettet behandling mot proteinet (HER2). Studier har vist at HER2-status kan endres under sykdomsforløpet. Å måle *ERBB2*-amplifikasjon i plasma er derfor ønskelig. Når *ERBB2* amplifiseres skjer dette kun med det ene allelet, det andre forblir normalt. Vårt prosjekt benyttet derfor målrettet dypsekvensering til å detektere SNP-ubalanse i *ERBB2* som en indikasjon på amplifikasjon. Metoden kunne detektere 1.2 ganger amplifikasjon, dette er like sensitivt som dPCR som per i dag er den mest sensitive metoden til å detektere kopitallsendringer.

Det siste delprosjektet benyttet målrettet dypsekvensering til å screene pasienter med metastatisk kreft for somatiske mutasjoner i 17 selekterte gener, som er av betydning innenfor brystkreft. Både tumor, normalmateriale og gjentatte blodprøver tatt under behandling ble screenet. Resultatene viste at én mutasjon kunne påvises hos 94 % av pasientene (49/52) mens 48 % hadde to mutasjoner som vil kunne benyttes som tumormarkører for oppfølging i plasma. Et par av resultatene ble tatt med i denne masteroppgaven og begge tydet på underliggende subklonalitet, da de to mutasjonene responderte ulikt på behandling. Den ene pasienten hadde en KRAS-mutasjon som ikke var detektert på diagnosetidspunktet, dette kan skyldes at mutasjonen ble ervervet under

behandling eller at den var minimalt representert i biopsien. Dette viser viktigheten av å screene for flere markører både ved diagnose og oppfølging.

Måling av ctDNA i blod er et viktig verktøy for å kunne følge med på terapirespons og tidlig detektere tilbakefall av sykdom. Kvantitering av ctDNA kan bli et skritt videre mot individuelt tilpasset kreftbehandling.

Abbreviations

BC	Buffy coat
bp	Base pair
cfDNA	Circulating free DNA/ cell free DNA
CDH1	Cadherin 1
CNA	Copy number abbreviation
CS1/CS2	Common sequence 1/Common sequence 2
CT	Computed tomography
CTC	Circulating tumor cell
ctDNA	Circulating tumor DNA
DETECT	Identification and classification of circulating tumor cells and circulating nucleic acids in patients with metastatic breast cancer
DNA	Deoxyribonucleic acid
DTC	Disseminated tumor cell
dPCR	digital PCR
EMT	Epithelial- mesenchymal transition
ER	Estrogen receptor
ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2
FISH	Fluorescence in situ hybridization
FFPE	Formalin fixed paraffin embedded
gDNA	Genomic DNA
HER2	Human epidermal growth factor receptor 2
IHC	Immunohistochemistry
IFC	Integrated fluidic circuit
Ki67	Antigen KI-67
KRAS	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
LNA	Locked nucleic acid
LOD	Limit of detection
MAF	Minor allele frequency
MBC	Metastatic breast cancer
MBG	Minor groove binder
MPS	Massively parallel sequencing
NEC	No extraction control

nt	Nucleotide
PEG	Polyethylene glycol
PIK3CA	Phosphatidylinositol-4, 5-bisphosphate 3-kinase, catalytic subunit alpha
PCR	Polymerase chain reaction
PPC	Pooled normal plasma control
PR	Progesterone receptor
PT	Primary tumor
qPCR	Quantitative polymerase chain reaction
Rn	Normalized reporter
RPP30	RNase P subunit 30
SBS	Sequencing by synthesis
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SPRI	Solid phase reversible immobilization
TNBC	Triple negative breast cancer
TP53	Tumor protein 53
WT	Wild type
XenT	Xenopus Tropicalis

Contents

Acknowledgements ... 3
Abstract ... 5
Sammendrag ... 7
Abbreviations ... 9
Introduction ... 13
Breast cancer ... 13
Tumor Heterogeneity ... 14
Genomic tumor markers ... 17
Circulating tumor cells ... 18
Circulating tumor DNA ... 19
Aim ... 21
1. Evaluating the effect of delayed sample processing of plasma on ctDNA measurements ... 22
Background ... 22
Material ... 22
Methods ... 23
Plasma separation ... 23
Plasma DNA extraction ... 23
Cell-free DNA quantification ... 24
Quantification of total plasma DNA ... 24
Quantification of circulating tumor DNA ... 25
Running a digital array ... 25
Analyzing a dPCR run ... 27
Allelic discrimination assays ... 28
Results ... 33
Extraction control ... 33
Total plasma DNA ... 33
Circulating tumor DNA ... 37
PIK3CA p.E545K ... 40
TP53 p.N239D ... 43
TP53 p.Y220C ... 45
2. Analysis of circulating tumor DNA to monitor metastatic breast cancer ... 48
Material ... 48
Methods ... 49
Illumina sequencing ... 50
Primers ... 52
Pre-amplification ... 52

Target specific amplification 53
Harvest 53
Barcoding 54
Library Cleanup 54
Library Quantification 55
Bioinformatics 56
Preliminary results 57
3. Detection of tumor specific amplification in blood by targeted MPS 59
Background 59
Material 59
Methods 60
Primer design and validation 60
Target specific PCR and MPS 61
Results 63
SNP frequency 63
Lower limit of detection (LOD) 63
Discussion 66
Evaluating the effect of delayed sample processing of plasma on ctDNA measurements 66
Allelic discrimination assays 69
Analysis of ctDNA by MPS to monitor metastatic breast cancer 71
Detection of tumor specific amplification in blood by targeted MPS 72
Comparisons of the two methods 73
Conclusion 74
References 74
Appendix A 74
Primer- and probe sequences: total DNA assay
Primer- and probe sequences: allelic discrimination assays
Appendix B 80
Phenol-chloroform extraction protocol
Primer plate for targeted MPS
Purification of harvested PCR products
Appendix C 83
Primer sequences ERBB2
Dilution curve ERBB2, each SNP separately

Introduction

Breast cancer

Breast cancer is the most common cancer in women worldwide, comprising 60 % of all female cancers. The incidence rates vary greatly, with the highest incidence in North America and Western Europe, and the lowest incidence in East Africa. The incidence is overall increasing, both as a result of population growth and aging, and environmental and life style factors (3)

In Norway, 2839 women and 13 men were diagnosed with breast cancer in 2010. For the first time since the beginning of breast cancer registration, a decrease in incidence (5 %) was observed between 2006 and 2010, compared with the previous 5-year period. In the same period the relative survival increased from 86 to 89 %. This positive trend is linked to an increased focus on the disease, screening programs and better treatment alternatives (figure 1) (1)

More than 90 % of women with early stage diagnosis survive their disease for at least 5 years. For women with more advanced stage disease, survival is only ~ 15 % (4)

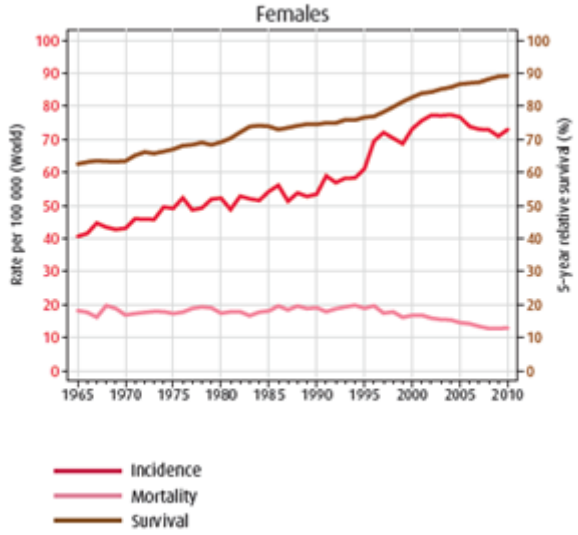

Figure 1. Breast cancer incidence in Norway 2010 (1)

Standard treatment of breast cancer includes surgery with or without local radiation. The surgery can either be breast conserving (lumpectomy) or include complete removal of the breast (mastectomy). The sentinel node is localized, removed and examined during surgery and guides the surgeon in the removal of other lymph nodes. With no signs of metastasis fewer lymph nodes have to be removed thereby avoiding side effects, such as lymph edema, for the patient.

Systemic therapy might be given after surgery to increase the chance for long-term survival. Treatment decisions are made on the basis of tumor size, histological grade, lymph node involvement, Ki67 expression (proliferative marker) and age in addition to a handful of predictive molecular markers (5). Targeted therapy is indicated by the expression of estrogen receptor (ER), progesterone receptor (PR) and/or overexpression of human epidermal growth receptor 2 (HER2) which can be identified by tissue immunohistochemistry (IHC). Fluorescence in situ (FISH) can be used instead of IHC to detect HER2 amplification by targeting the gene (v-erb-b2 erythroblastic leukemia viral oncogene homolog 2; *ERBB2*) instead of the protein, HER2 (Figure 2). Triple negative tumors; ER-/PR-/HER2-(TNBC) have no targeted therapy option at present. Cytostatic adjuvant treatment might be given when predictive markers are missing. Neoadjuvant treatment (adjuvant treatment before surgery) is increasingly being used in breast cancer and is administrated before surgery to decrease the tumor.

Tumor Heterogeneity

Breast cancer can be divided into four subgroups with predictive and prognostic differences based on the protein markers ER, PR and HER2. A more detailed stratification can be done by looking at gene expression levels and recently 10 integrated clusters with different outcomes could be recognized based on copy number aberrations (CNA) and gene expression of the nearby genes (6, 7).

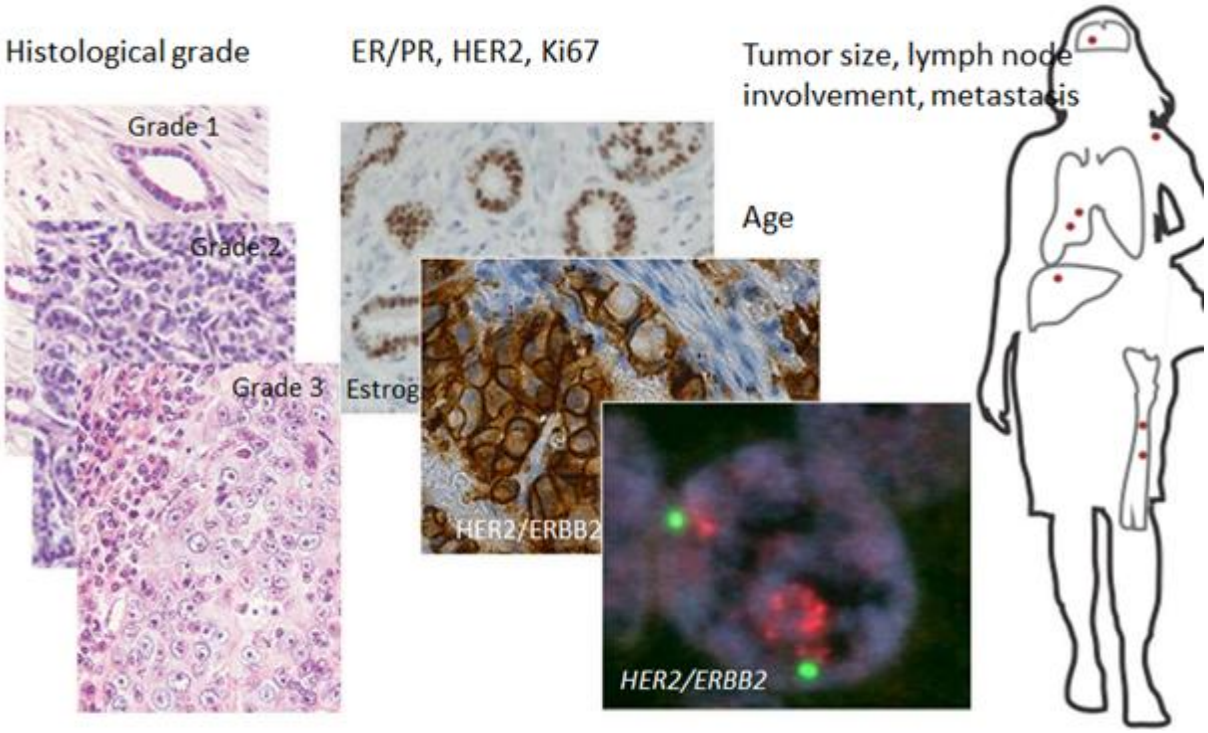

Figure 2. Several prognostic and predictive markers are used to identify those patients that might benefit from adjuvant therapy. These markers include histological grade; less differentiated tumors are more aggressive than tumors that resemble normal tissue. IHC staining of estrogen receptor (ER), progesterone receptor (PR) and HER2 identifies patients that are suited for targeted therapeutics. In some instances FISH is used instead of IHC to detect HER2 overexpression (*ERBB2* amplification). Other parameters that affect treatment decisions are age, tumor size, lymph node involvement and metastasis.

In addition to the vast heterogeneity observed across breast tumors, extensive genomic heterogeneity is also seen within tumors. Intratumor heterogeneity is defined as co-existence of multiple clonal subpopulations within a single neoplasm (8). Tumors arise as a consequence of acquisition of mutations that confer growth advantages. As subclones evolve from the original primary clone, a selection process occurs and subclones with a mutation resulting in a selective advantage, will eventually outgrow subclones with less potential (9, 10)

At least two forms of clonal evolution seem to exist. The linear model where a single progenitor gradually accumulate more somatic mutations giving rise to more aggressive subclones within the original neoplastic clone. And the branched model where multiple subclones are developed and retained from the earliest days of tumor evolution (11) (figure 3). Shah et al saw both patterns in TNBC's. By comprehensive molecular analysis including

ultra-deep targeted sequencing (20 000 x depth) of 104 TNBC's, it was shown that some only had a few subclones, whereas others had a huge degree of heterogeneity (12).

Another study, of 21 breast cancers of different subtypes, supported the branched model but found them all to harbor a dominant subclone contributing to more than 50 % of the tumor cells (13). The major clone may however not be the one that eventually acquires metastatic potential. Studies comparing primary and metastatic tumors have shown that the metastasis may derive from a low-frequency subclone within the primary (14)

By the time of diagnosis, breast cancers are composed of heterogeneous populations of tumor cells and it is important to consider the possibility of co-existing subclones that may be regionally separated within the tumor (15). Regional heterogeneity might introduce sampling bias which impairs the interpretation of genomic data from single biopsies (16). Seol et al found HER2 regional heterogeneity present in subset of HER2 amplified breast cancers and especially in patients with low grade HER2 amplification and equivocal HER2 expression. This group had shorter disease-free survival, indicating the need of examining a larger biopsy (17).

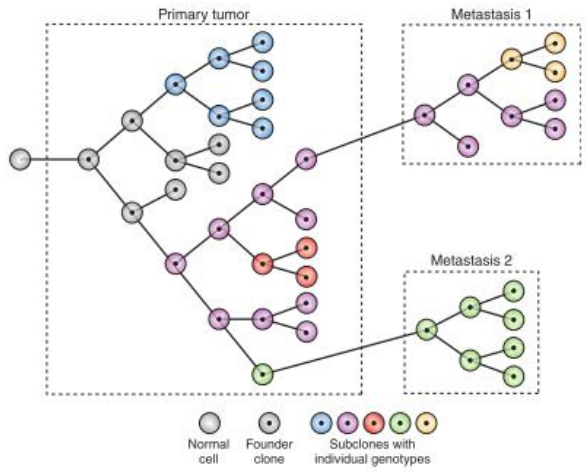

Figure 3. Clonal evolution. A normal cell acquire a mutation that confers growth advantage and accumulation of further mutations expands the clone and creates diversity. Only a limited number of clones have the potential to metastasize. Further acquisition of mutations can create subclones in the metastasis (18)

Genomic tumor markers

A wide range of genomic alterations can lead to the development of cancer. Advances in sequencing technologies have enabled rapid identification of these genomic alterations in individual tumors. Massive parallel sequencing (MPS) of breast tumors show numerous alterations such as point mutations, small insertions and deletions (indels) and larger gene rearrangements arising from deletions, inversions, duplications and translocations (figure 4). A subset of the mutations can be classified as drivers, but the vast majority of mutations will be passengers. A driver mutation has been selected for in the tumor because of growth advantages while passenger mutations have not been selected for and are not involved in the carcinogenesis (19, 20). Any chromosome alterations might be used as a tumor marker for follow-up, whether it is a driver or a passenger.

Breast cancer has few recurrent structural variations, so most of these will be private rearrangements in the individual tumors. Some genes like tumor protein 53 (*TP53*) and phosphatidylinositol-4, 5-bisphosphate 3-kinase, catalytic subunit alpha (*PIK3CA*) are frequently mutated. *PIK3CA* has two hotspot regions, while *TP53* mutations are spread throughout the gene. Gene amplification is a frequent alteration in breast cancer that affects multiple genomic regions (21). The genetic events producing gene amplifications are complex and may include telomere shortening, breakage-fusion-bridge cycles and incorporation of sequences of many chromosomes (22). The most common amplified gene is *ERBB2* which encodes a transmembrane tyrosine kinase receptor, human epidermal growth factor receptor-2 (HER2) and the receptor is overexpressed and/or amplified in 15-20 % of breast cancers. To date this is the only genomic tumor marker recommended for diagnostic use.

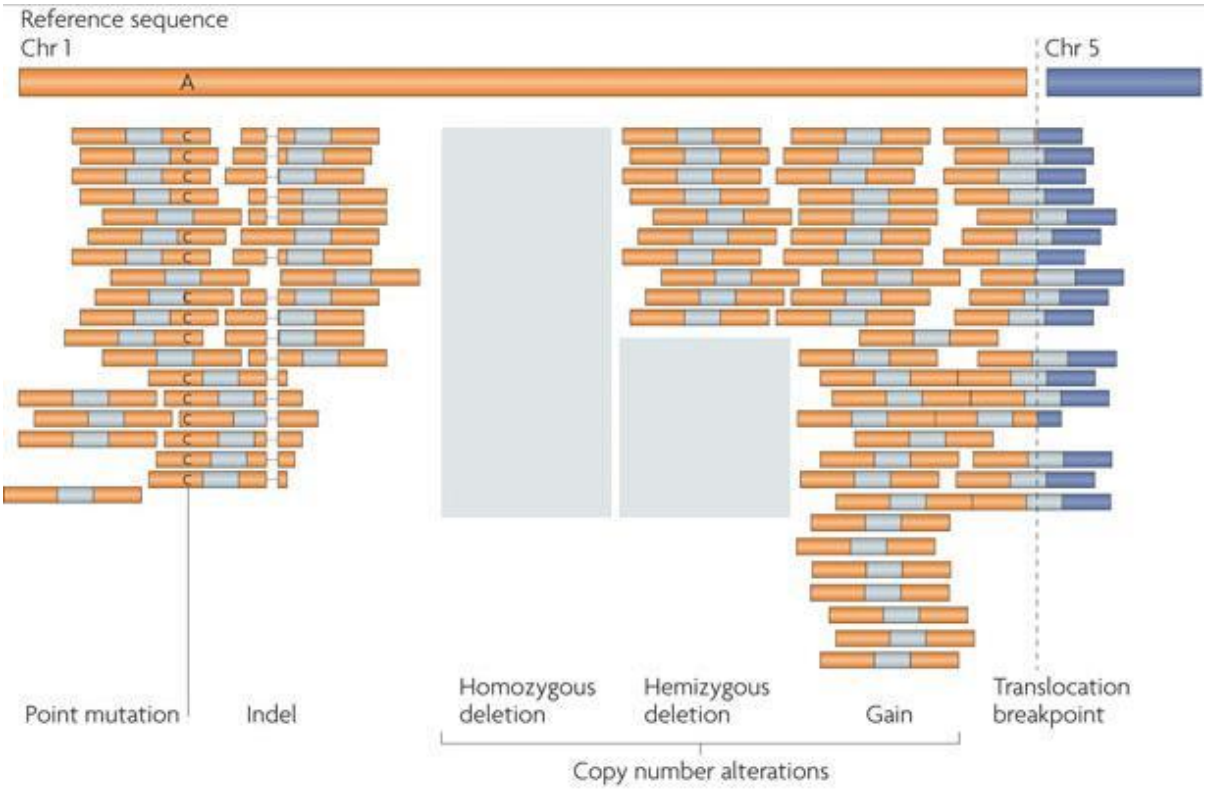

Figure 4. Types of alterations that can be detected by massively parallel sequencing (MPS). Sequenced fragments (reads) are aligned to the reference genome and the colored tips indicate that they have been sequenced from both ends (paired-end sequencing). The grey in the middle of the reads represent the unsequenced portion of the DNA fragment. Different types of DNA alterations can be detected; point mutations and indels are identified by not mapping to the reference genome. Copy number alterations are seen by a change in the sequencing depth compared to a normal reference (preferentially normal material from the same individual). Paired-end reads that map to different loci are evidence of gene rearrangements (Modified figure) (23)

Circulating tumor cells

Tumors shed whole cells into the circulation. Some of these circulating tumor cells, CTC's, have the capability to invade and spread to other sites, becoming disseminated tumor cells (DTC). Bone marrow is a frequent homing organ for DTC's in breast cancer, where they can rest for several years in a non-proliferative state (dormancy). Detection of DTC's is a strong risk factor for future metastases but a significant proportion of women with DTC's do not experience relapse.

CTC's go through a morphogenetic process called epithelial- mesenchymal transition (EMT) to acquire a migratory phenotype with invasive properties. This process downregulates the expression of epithelial antigens on the cell surface. CTC's entering the bone marrow can go

through the reverse process (mesenchymal-epithelial transition) and regain the epithelial characteristics (24).

Obtaining a blood sample is less invasive than a bone marrow aspirate. Several studies have therefore addressed the question whether CTC detection has prognostic value. Five or more CTC's per 7.5 ml blood (or even 1 CTC/7.5ml in some studies(24)) is associated with worse outcome, but there are some limiting factors in terms of sensitivity and specificity of current methods for CTC detection. Detection of CTC's is not currently used in routine diagnostics (25-27).

Circulating tumor DNA

Detection of circulating tumor DNA (ctDNA) provides a non-invasive method to analyze tumor specific markers from blood. The precise mechanism by which cell-free DNA is released into the bloodstream remains uncertain, but apoptosis and necrosis are proposed to be the main mechanisms (Figure 5) (28). Low amounts of circulating free DNA (cfDNA) can be detected in healthy individuals. Although the mean cfDNA levels are significantly higher in cancer patients, the concentration range is overlapping with both healthy individuals and non-malignant conditions like trauma and chronic disease (29).

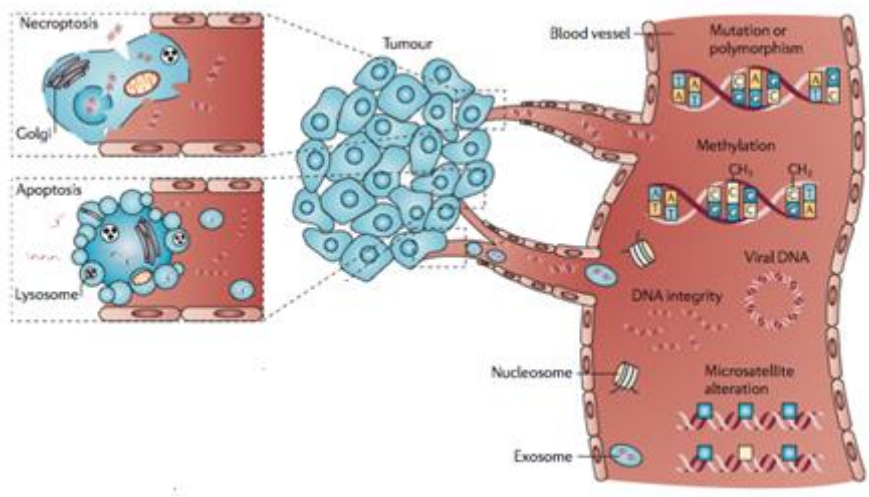

Figure 5. Release of ctDNA into the blood stream. Tumor cells as well as normal cells undergo necrosis and apoptosis and release DNA into the circulation. Whether cell-free DNA circulates as naked DNA or protein bound is unclear. ctDNA can be detected by tumor specific biomarkers (modified figure)(28)

Whether cfDNA circulates in blood as naked DNA, apoptotic bodies or bound to proteins may vary. The typical apoptotic ladder that is often seen in gel electrophoresis suggests that the main part circulates as mono- and oligonucleosomes. Nucleosomes consist of double stranded DNA wrapped twice around a histone core plus linker DNA, and are shed into the circulation during apoptosis when endonucleases cleave the DNA-helix. A mono-nucleosome has a DNA fragment around 150 bp. The size distribution of cfDNA varies between samples and high molecular weight DNA can also be seen on gels, supporting necrosis as a source of cfDNA as well (30-32)

In a cancer patient's blood, part of the cell free DNA derives from tumor. The proportion of ctDNA is variable and levels from less than 1 % to more than 90 % have been observed (33, 34). Studies have shown that ctDNA correlates with disease progression and can be used to monitor therapy response (35-37).

Aim

The aim of this thesis was to establish and evaluate different methods to detect ctDNA in blood. We identified two main challenges: First, ctDNA is frequently a minor fraction of total DNA and in order to detect ctDNA we need robust and very sensitive methods. Secondly it is important to analyze several tumor markers due to intratumor heterogeneity and clonal evolution. To elucidate this, three different projects were initiated:

1. **Evaluating the effect of delayed sample processing of plasma on ctDNA measurements**

2. **Analyzing the use of somatic point mutations in blood for disease monitoring in metastatic breast cancer patients**

3. **Detection of tumor specific amplification in blood by massively parallel sequencing**

The three projects will be presented separately in material, methods and results with a short introduction to each of them. The patient samples used in these projects have been approved for research purposes by the local research ethics committee in UK and all patients provided written informed consent (REC reference no.07/Q0106/63 and 08/H0306/61).

1. Evaluating the effect of delayed sample processing of plasma on ctDNA measurements

Background

Circulating tumor DNA is often a minor fraction of total cell free DNA in blood. It is therefore important to minimize the level of background wild type DNA. There has been a lack of standardization when it comes to processing blood for detection of ctDNA. Gradually, some guidelines have emerged like using plasma instead of serum because the clotting lyses leucocytes, producing an elevated background of wild type DNA (38). Some studies have compared different extraction kits and seen a difference in yield and fragment size-distribution (38, 39). Now a commercial kit tailored for extraction of circulating nucleic acids is on the market. There are still no guidelines on how fast a sample should be processed after collection; the general rule is to process it as soon as possible and preferentially within one hour. This may not always be feasible in a clinical setting, so the scope of the time course study was to see how delayed sample processing would affect total cell free DNA, and in particular the fraction of circulating tumor DNA.

Material

The study included 10 patients with metastatic breast cancer from the DETECT study (Identification and classification of circulating tumor cells and circulating nucleic acids in patients with metastatic breast cancer) and 14 patients with advanced ovarian cancer from the CTCR-OV04 study (Molecular Analysis of Response to Treatment in Ovarian Cancer). Five of the ovarian cancer patient samples were collected in EDTA- and BCT-tubes (Streck). BCT-tubes contain K_3EDTA and a proprietary preservative that stabilizes leukocytes to prevent lysis and thus an increased background of wild type DNA. BCT-blood can be stored at room temperature (RT) for up to 14 days, according to the manufacturer, before processing the plasma (40). The samples were split into two aliquots each, and processed immediately or after one week at RT. For the rest of the patients, one EDTA-tube for each time point was collected and processed as indicated in table 1.

Table 1. Number of blood samples processed at each time point. Blood was collected in EDTA- or BCT-tubes and stored at RT or at 4 °C at different time intervals before processing plasma.

	0 hr (RT)	6 hr (RT)	24 hr (RT)	24 hr (4 °C)	48 hr (RT)	48 hr (4 °C)	96 hr (RT)	96 hr (4 °C)	1 week (RT)
EDTA-blood	19	19	19	6	8	8	5	5	5
BCT-blood	5								5

hr= hour(s). RT=room temperature

Methods

Plasma separation

BCT- or EDTA-blood was centrifuged at 820g for 10 min. 1 ml plasma aliquots were transferred to 1.5 ml sterile tubes and centrifuged at 20,000g for 10 min to pellet debris. Plasma was then transferred to new 1.5 ml tubes without disturbing the pellets and stored at -80°C. All centrifugal steps were carried out at RT

Plasma DNA extraction

DNA was extracted from 0.46-2.6 ml plasma with QIAamp Circulating Nucleic Acid kit (Qiagen) according to the manufactures' 2009 protocol. Briefly, 5.6 µg carrier RNA per sample was added to lysis buffer ACL to facilitate binding of plasma DNA to the spin column. In addition to the protocol, buffer ACL was spiked with 6 µl of a single DNA amplicon from Xenopus Tropicalis (XenT) per sample to control the efficiency of the extraction. The XenT amplicon has no homology to human DNA. Plasma was incubated for 30 min at 60°C with Buffer ACL and proteinase K for protein digestion before processed through the spin columns with a vacuum manifold. Bound DNA was washed twice with buffers and eluted in 50 µl Buffer AVE. The first eluate was reapplied on the column to enhance the yield (41). A pooled plasma control (PPC) of five normal plasma samples was extracted with each extraction batch. A no extraction control (NEC) of 6 µl XenT and 44 µl Buffer AVE was prepared for each batch of extraction. Plasma DNA was stored at -20°C or -80°C.

Cell-free DNA quantification

Quantifications were done with microfluidic real-time digital PCR (dPCR) on BioMark HD System (Fluidigm). The concept of digital PCR (dPCR) was first described in 1992 (42)and the method depends on the ability of PCR to detect a single target molecule. In dPCR, samples are heavily diluted and partitioned into many parallel reactions, so that some aliquots receive one target molecule and some receive none (figure 6) (43). Plasma DNA concentrations are generally low, so dilutions are often unnecessary or limited. Hydrolysis probes were used to detect amplification. Hydrolysis probes are labeled with a reporter fluorochrome in the 5' end and have a 3'quencher that absorbs the fluorescence from the reporter as long as the probe is intact. The DNA polymerase has 5'-3' exonuclease activity and hydrolyzes the probe upon amplification. This separates the reporter from the quencher and the increased fluorescent signal is proportional to the amplified product (LifeTechnologies).

Figure 6. The principle of digital PCR is to dilute the sample and partition it into hundreds or thousands of nl or pl volumes so that some volumes get no target template whilst others get one. Absolute quantity is measured by counting positive partitions and the quantity is corrected by Poisson equation due to stochastic events leading to more than one target in some partitions (44)

Quantification of total plasma DNA

A 65 bp fragment of the single copy gene, RNase P subunit 30 (*RPP30*) was quantified in duplex dPCR with a 67 bp fragment of XenT. Primer- and probe sequences are listed in Appendix A. Sample- and reaction mix was made in a total volume of 6 µl. The final reaction consisted of 1x Universal TaqMan PCR Master Mix (Life Technologies), 1x GE Sample Loading

Reagent (Fluidigm), 900 nM primers and 250 nM probes. DNA was denatured at 96° C for 1 min and cooled on ice for minimum 1 min to create single stranded DNA fragments. Plasma DNA concentrations are usually low, so a standard volume of 1.8 µl template was added to 4.2 µl cooled reaction mix. Each sample was run in duplicate.

Quantification of circulating tumor DNA

Circulating tumor DNA (ctDNA) was quantified with allelic discrimination assays (allelic discrimination assays are described in a section below). Sample- and reaction mix was made in a total volume of 10 µl. The final concentration was 1x TaqMan Universal PCR Master Mix, 1xGE Loading Reagent, 900 nM primers and 250 nM probes. DNA was denatured at 96° C for 1 min and cooled on ice for minimum 1 min. Volume template was added based on the total amount of DNA quantified, but at a maximum of 4 µl. Each set up included the normal plasma control (PPC) extracted with the sample, as a negative template control for the mutant probe. Primers, probes and run parameters are listed in Appendix A.

Running a digital array

Fluidigm has two kinds of microfluidic digital arrays; qdPCR 37K IFC and 12.765 IFC (figure 7). Sample- and reagent mixes are pipetted into sample inlets and transferred to ~770 microwells per sample by air pressure. The differences between the two arrays are the number of samples that can be loaded per array and the volume of template that can be loaded and analyzed. Details are listed in table 2. We quantified total plasma DNA with qdPCR 37K IFC arrays that can take 48 samples per array while ctDNA was quantified with 12.765 IFC arrays that only take 12 samples but have the capacity to be loaded with more template.

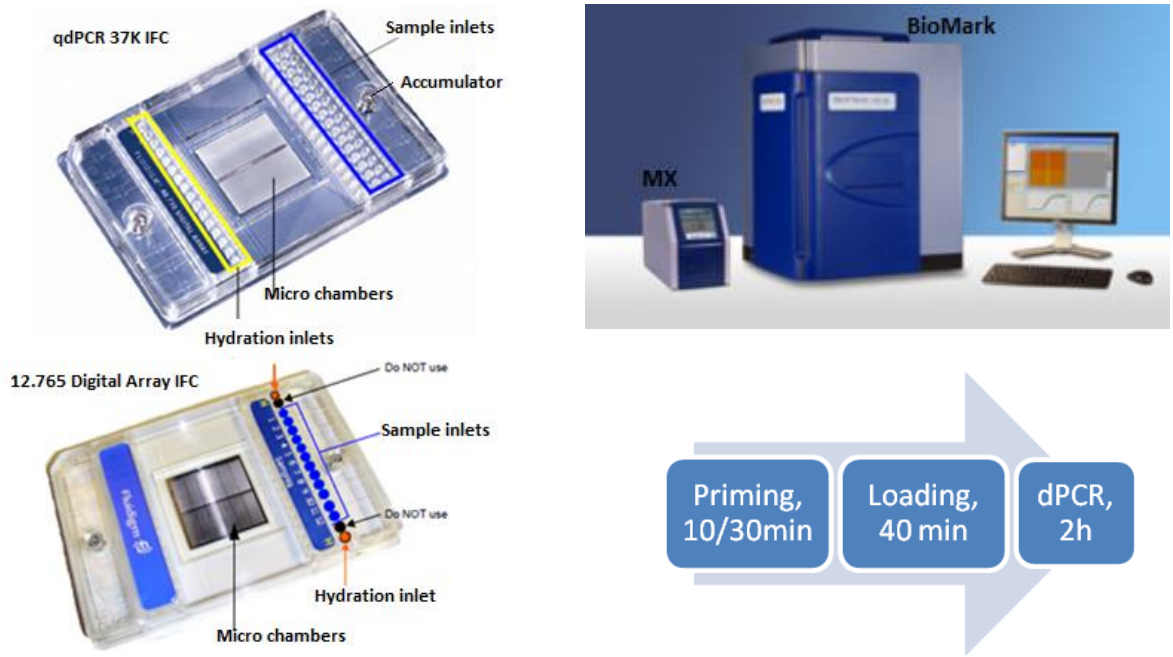

Figure 7. There are two types of microfluidic digital arrays, the qdPCR 37K IFC which can take for 48 samples and the 12.765 Digital Array IFC which can be loaded with 12 samples. The workflow includes priming and loading of the array in IFC Controller MX and running dPCR on the BioMark™ HD System (Modified figures) (45).

Running a digital array includes three steps; priming, loading and dPCR. Arrays have to be loaded within one hour of priming and dPCR has to be run within 4 hours of loading. **Priming:** A syringe with 300 µl Control Line Fluid was added to each of the two accumulators of the digital array before 40 min (37K) or 10 min (12.765) priming in IFC Controller MX (Fluidigm). **Loading qdPCR 37K IFC:** 10 µl 1xGE Loading Reagent was pipetted to the hydration inlets and 4 µl cooled sample- and reaction mix was added to the sample inlets of the digital array. The digital array was loaded with sample- and reaction mix for 40 min in IFC Controller MX and transferred to BioMark HD System for dPCR. The PCR program was an initial 50°C for 2 min and 95°C for 10 min followed by 55 cycles of 95°C for 15 seconds and 60°C for 1 min. **Loading 12.765 IFC:** 9 µl DEPC-treated H_2O was added to the hydration inlets and 9 µl cooled sample- and reaction mix was added to the sample inlets. The digital array was loaded with sample- and reaction mix for 40 min in IFC Controller MX and transferred to BioMark HD System for dPCR. Run conditions are specified in Appendix A, table 2.

Analyzing a dPCR run

Fluidigm's Digital PCR Analysis Software version 3.0 was used to analyze the results. The absolute quantity is calculated by dividing the number of positive aliquots by the total number of parallel reactions, so there is no need for standard curves. Since limited diluted target molecules will be randomly distributed, some wells will contain more than one target molecule. Simply counting positive wells, will underestimate the true number of target. The quantity is therefore corrected by Poisson equation, which estimates the average number of targets (λ) per well, $\lambda = -\ln\left(1 - \frac{H}{C}\right)$, where H is the number of positive wells and C is the number of total wells.

Estimated target counts were then calculated manually to copies per µl and copies per ml according to the formulas below:

Copies/ µl: $$\frac{estimated\ target\ per\ panel \times total\ reaction\ input}{panel\ volume \times template\ input} \times template\ dilution$$

Copies/ml: $$\frac{copies\ per\ \mu l \times elution\ volume}{ml\ plasma\ extracted}$$

Table 2. Digital Array Volumes

	qdPCR 37 K	12.765
Total reaction input *	4 µl	9 µl
Max template input**	1,2 µl	3,6 µl
Microwell volume	0,85 nl	6 nl
Panel volume	0,65 µl	4,59 µl
Samples/array	48	12
Number of microwells	770	765

***Total volume added to the sample inlets, inclusive template **Amount of template added to the sample inlets of the array. The template volume is slightly larger when preparing the sample and reaction mix due to dead volume**

Allelic discrimination assays

In this study all patients were known to have somatic point mutations in either PIK3CA or TP53 that had previously been identified by targeted sequencing. Allelic discrimination assays were designed in each case to detect these mutations in ctDNA using dPCR.

Allelic discrimination assays can be used for detection of single base mutations, either with allelic primers or allelic probes. There are some technical challenges though in the form of specificity. Discriminating between sequences differing by one base pair only, might give rise to unspecific binding and amplification. In this study, allelic hydrolysis probes were used to detect mutations in ctDNA using dPCR (Figure 8). Allelic discrimination probes should be as short as possible, because a mismatch will have bigger impact on the melting temperature (Tm) of a short sequence than a larger one. TaqMan minor groove binding (MGB) probes are often used because the 3' MGB increases Tm allowing design of shorter probes (46). Another useful probe for allelic discrimination is locked nucleic acid (LNA) probes. LNA's are nucleic acid analogs with a 2'-O -4 'C methylene bridge whose conformation enhances base stacking and phosphate backbone organization and results in improved affinity for complementary DNA (47-49).

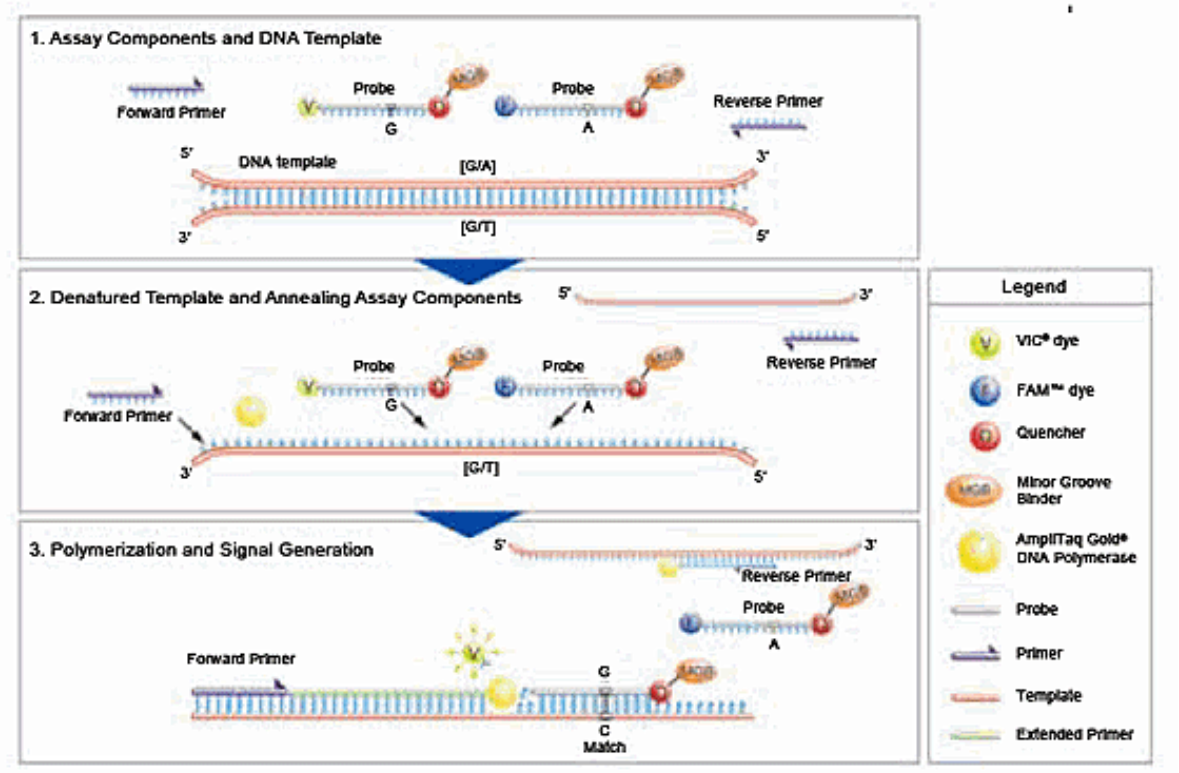

Figure 8. Allelic discrimination assay with two hydrolysis probes which are labeled with different fluorochromes. One probe binds to wild type sequences and the other to mutant sequences. Bound probes are fragmented by the DNA polymerase upon amplification with the subsequent increase of fluorescence when the reporter fluorochrome is separated from the 3' quencher (Life Technologies)

Primer and probe design

Three allelic discrimination assays were designed; PIK3CA p.E545K, TP53 p.D281G and TP53 p.y220C (Table 3). LNA-probes were made when optimal MGB-probes were difficult to obtain. Life Technologies guidelines for design of primers and MGB-probes were followed. The guidelines recommend the primers to have a Tm between 58-60°C and the amplicon size to be between 50-150 bp. To avoid cross-hybridization, probes must be able to discriminate between sequences that differ by one base only. To achieve this, the size of the probe should be as short as possible so that the single base mismatch will have greatest impact. Although less than 13 nt may reduce signal (probably due to non-specific binding). The GC-content should be between 30-80 % and Tm between 65-67°C so that probes bind to the template before primer elongation starts. The following base positions and motifs should be avoided in an MGB-probe; no G at the 5'; this will reduce signal, even after cleavage. No G in the second position 5' of 6FAM-labeled probes since this also might reduce the signal. The

following motifs should also be avoided: 4 consecutive G's, 6 consecutive A's, and C di-nucleotides or more in the middle of the probe. At the 3' end the following motifs should be avoided: …GGG-MGB and GGAG-MGB. The mismatch position should be central or towards the 3' end, but not at the very two last 3'bases.

Owczarcys paper was used as guideline for design of LNA-probes. To maximize mismatch discrimination, the general recommendation is to have a triplet of LNA-residues, with the mismatch site in the center of the triplet. In most cases this increases discrimination, but the destabilization is sequence dependent and not universally observed. If introduction of LNA-nucleotides cause a large duplex stabilization and the single mismatch destabilizes the duplex less, the mismatched LNA-DNA duplex will be more stable than a mismatched DNA duplex. This is often the case of +G·T mismatches and some +C·A mismatches (+ indicate an LNA base, the dot indicates base pairing) as shown in figure 9. The mismatch's flanking LNA-bases also have impact on the discrimination. +G·C and +C·G decrease the level of discrimination while +A·T or +T·A increase it. As with MGB-probes the position of the mismatch should be central and here they recommend no LNA-base in the last three positions 3'. In the case of +G·T mismatches, the paper suggests that it might give better discrimination by focusing on making a short probe to maximize the effect of mismatch and position the LNA's ≥2 bp from the mismatch (2).

In addition to applying the guidelines above, Integrated DNA Technologies (IDT) OligoAnalyzer3.1 was used for Tm mismatch prediction (50). The tool predicts ΔTm for single mismatches but cannot take into account the effect of LNA-bases, so the calculation will only be approximate for LNA-probes. Exiqons LNA Oligo Tools were used for Tm prediction and to check for self-complementarity (51). All primers and MGB-probes were designed with Primer Express Software v3.0 (Life Technologies). Specificity was controlled with Primer-Blast (NCBI) and UCSC In-Silico PCR and the latter was also used to check for SNP's in the primer sites. Primers and LNA-probes were purchased from Sigma, MGB-probes from Life Technologies.

Assay Optimization

PIK3CA p.E545K assay

A previous assay had been designed with the automatic Primer Express Software program v.3.0 but this demonstrated high background levels. The WT-probe had Tm 69 °C and the mutant probe had Tm 64°C. Both had a calculated mismatch ΔTm of 5.6 °C. The assay annealing and elongation temperature was 56 °C. A new assay was therefore designed, on the same strand with LNA-probes containing a triplet of consecutive LNA-bases with the mutation site in center of the triplet. The mutant probe had Tm 65 °C and a calculated mismatch ΔTm of 3.4 °C. Three WT probes were designed. WT1 had Tm 67 °C and mismatch ΔTm 5.6 °C, WT2 had Tm 65 °C and ΔTm 7.4 °C and WT3 had Tm 65 °C and ΔTm 5.6 °C. All probes except WT3 had a triplet of LNA-bases with the mutation site in the center. WT3 probe had one LNA- base ≥2 bp from the mutation site on each side. This was suggested in Owczarcys paper because an LNA-base at a G·T mismatch might decrease ΔTm instead of increasing it (see figure 9). The effect of LNA-bases is not reflected in the calculated mismatch ΔTm.

TP53 p.Y220C assay

Primer Express Software v3.0's automatic program was used to get an assay proposal. Best mismatch discrimination would be achieved using the sense strand according to IDT's Tm mismatch tool (50). The automatic Primer Express program suggested a WT-probe with 67°C Tm. The probe had a calculated mismatch ΔTm of 6.1 °C. With a standard PCR annealing and elongation temperature of 60 °C, this WT-probe would be expected to be able to bind to mutant sequences, since the reduced Tm due to the mismatch is higher than the annealing/ elongation temperature. A shorter WT-probe was therefore designed, with a Tm of 64 °C and a mismatch ΔTm of 7.0 °C. The mutant probe had a Tm of 66 °C and ΔTm 8.2 °C and was not adjusted.

TP53 p.N239D assay

The TP53p.Y220C assay had to be designed on the anti-sense strand due to a triplet of C-bases close to the mutation site. A triplet of C's in the middle of a probe might decrease the signal. Primer Express v3.0's automatic program was used to get an assay proposal. Neither of the probes proposed could be labeled with 6FAM due to the second 5'G that might decrease signal in 6FAM-probes. The mutant probe was therefore designed as an LNA-probe and labeled with ROX. The proposed WT-probe was kept because it had the higher ΔTm-mismatch than the alternative LNA-probes that were designed. Primers were checked for SNP's with UCSC Genome Browser and an alternative reverse primer was designed due to the 3' SNP rs1800372 with minor allele frequency 0,008 (dbSNP, 1000Genomes).

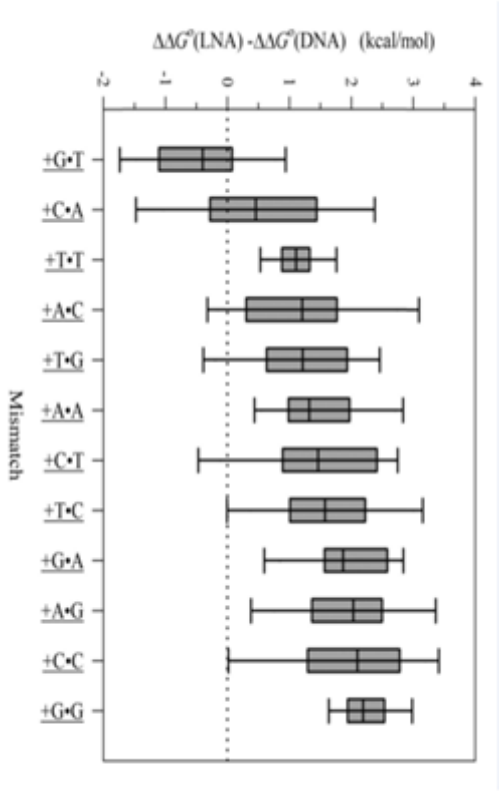

Figure 9. Differences in mismatch discrimination between LNA and DNA sequences. Free energies, ΔΔG°= ΔG° mismatch-ΔG° match, were predicted for triplets of LNA-bases with central mismatch and all possible neighbor base pair combinations. Positive values indicate that LNA increase mismatch discrimination relative to DNA (2)

Table 3. Developed assays. Mismatch site is indicated in red. [+base] indicates LNA-base. ΔTm mismatch was calculated with IDT's Tm mismatch tool in OligoAnalyzer 3.1 (50). The calculation is only approximate because the effect of LNA-bases is not taken into account. Three different WT-probes were tested. TP53 p.Y220C-assay, reverse primer 1 has a SNP indicated in yellow (rs1800372) with MAF 0.008 (dbSNP, 1000Genomes).

Assay	Primers	Sequence 5'-3'	Tm	Mismatch ΔTm (°C)	Mismatch	Neighbor bases
PIK3CA p.E545K (87 bp)	Forward	ACAGCTCAAAGCAATTTCTACACG	58.7			
	Reverse	AGCACTTACCTGTGACTCCATAGAAA	58,5			
	Mutant	6FAM-TCTGAAATCAC[+T][+A][+A]GCAGGAGA-BHQ1	65	7.4	A·C	TAA
	WT1-	HEX-CTGAAATCAC[+T][+G][+A]GCAGGA-BHQ1	67	5.6	G·T	TGA
	WT2-	HEX-AATCAC[+T][+G][+A]GCAGGAG-BHQ1	65	7.4	G·T	TGA
	WT3-	HEX-CTGAAAT[+C]ACTGAG[+C]AGGA-BHQ1	65	5.6	G·T	TGA
TP53 p.N239D (62 bp)	Forward1	TCTGACTGTACCACCATCCACTACA	59.4			
	Forward2	GCTCTGACTGTACCACCATCCA	58.7			
	Reverse	CATGCCGCCCATGCA	59.0			
	Mutant	6FAM-CATGTGTGACAGTTC-MGB	66.0	8.2	G·T	TGA
	WT	VIC-TACATGTGTAACAGTTC-MGB	64.0	7.0	A·C	TAA
TP53 p.Y220C (79 bp)	Forward	GAGACCCCAGTTGCAAACCA	59.5			
	Reverse1	TGGATGACAGAAACACTTTTCGAC	59.2			
	Reverse2	TTTGGATGACAGAAACACTTTTCG	59.2			
	Mutant	ROX-CGGCTCA[+C]AGGGCA-BHQ2	66	8.9	C·A	ACA
	WT	VIC-AGGCGGCTCATAG-MGB	66	4.3	T·G	ATA

Results

Extraction control

Sample copy numbers of XenT were divided by the copy number in the no extraction control (NEC) to monitor extraction efficiency. NEC copy numbers were divided by 1.1 to correct for dead volume. The average efficiency was 0.83 (median: 0.82, range: 0.19-2.12).

Total plasma DNA

Estimated copy numbers of *RPP30* were first adjusted to copies/ml plasma extracted. Then relative ratios were calculated within each patients sample collection (table 4). The sample processed within one hour was used as a baseline and all the other time points were divided by the baseline copy number. This was done because not all patients were represented at every time point and because the plasma DNA concentrations vary hugely between patients. For instance 19 patients had time point 24 hours at RT, but only 6 patients had time point 24

hours at 4 °C. The relative ratio of total plasma DNA levels at each time-point compared to baseline was analyzed for all samples. The ratios were compared with Wilcoxon signed-rank test and there were no significant difference in total plasma DNA yield from blood left up to 24 hours in room temperature (p=0.55, n=19) or stored at 4 °C for 24 hours (p=0.12, n=6). There was however a significant elevation in plasma DNA when samples were left at RT for 48 hours (p=0.008, n=8) or at 4 °C (p=0.02, n=8) prior to processing plasma. There was a strong trend towards elevation in the other time points; 96 hours at RT (p=0.06), 96 hours at 4 °C (p=0.06) or 1 week at RT (p=0.06), but these did not reach statistical significance, most likely because of the low number of samples (n=5) analyzed at these time-points (figure 10).

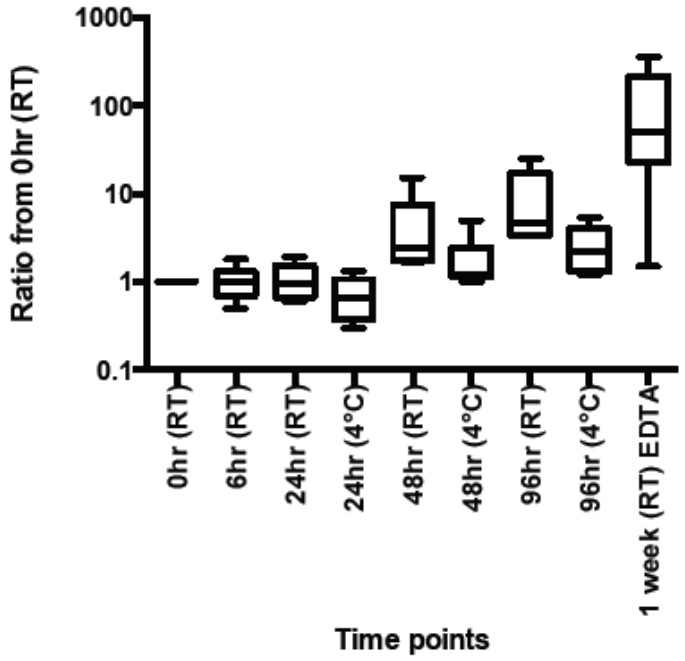

Figure 10. Box- and whisker plot for relative levels of total plasma DNA compared to samples processed within one hour. EDTA-blood was left at room temperature (RT) or at 4 °C for up to 1 week before processing plasma.

Samples stored in BCT-tubes (Streck) for one week at RT showed less elevation than EDTA-blood left at RT for the same time. BCT-blood showed a median 3.9 elevation (range 0.6-12) in total plasma DNA after one week compared to BCT-blood that was processed immediately. EDTA-blood had a median 49.7 increase (range 1.5 to 356) in the same period compared to EDTA-blood that was processed immediately. The median copy numbers of

total plasma DNA in EDTA-blood were 1954 copies/ml (0 hours) and 85613 copies/ml (1 week). BCT-blood had median 1757 copies/ml (0hr) and 10256 (1 week). There was a strong trend (p=0.06, n=5) towards less elevation of total plasma DNA in samples collected in BCT-tubes. There was no significant difference between cfDNA concentrations processed immediately or after one week in BCT tubes (p=0.31) but there was a trend seen in EDTA-blood (p=0.06) (figure11).

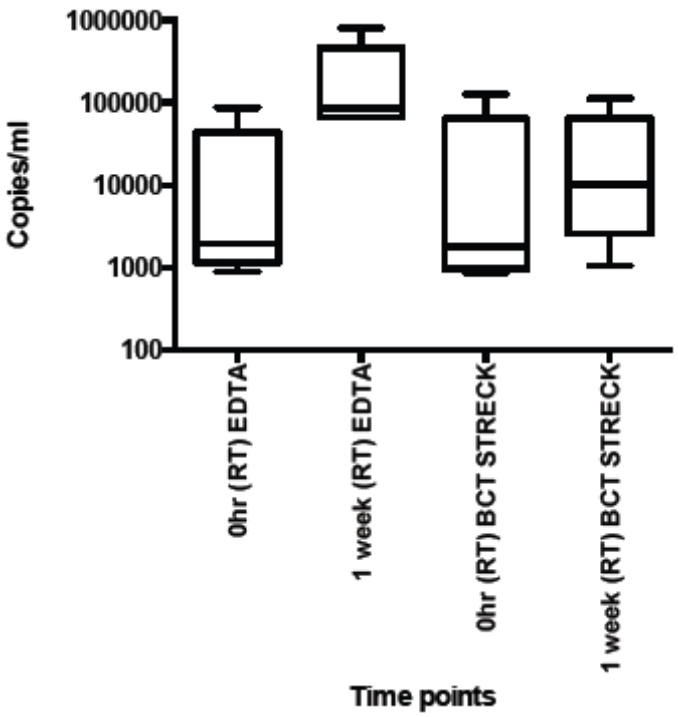

Figure 11. Box- and whisker plot for total plasma DNA (copies/ml) in blood from EDTA- and BCT-tubes. Samples processed immediately showed no difference between the tubes. After one week in room temperature (RT), EDTA-blood had a higher copy number elevation, although not statistically significant (p=0.06, n=5). There was a trend towards less elevation in BCT-tubes.

Table 4. Total DNA fragments per ml plasma extracted. Ratios of copy number change relative to 0 hours (0 hr) in parenthesis.

	EDTA									BCT	
Sample Name	**0hr RT**	**6hr RT**	**24hr RT**	**24hr 4°C**	**48hr RT**	**48hr 4°C**	**96hr, RT**	**96hr 4°C**	**1 week RT**	**0hr RT**	**1 week RT**
E26 V03	27502 (1.0)	14743 (0.5)	18335 (0.7)	-	-	-	92410 (3.4)	34123 (1.2)	-	-	-
E58 V02	7639 (1.0)	8234 (1.1)	4947 (0.6)	-	-	-	25742 (3.4)	10893 (1.4)	-	-	-
E56 V03	3972 (1.0)	3492 (0.9)	-	-	-	-	99999 (25.2)	8898 (2.2)	-	-	-
E65 V03	14423 (1.0)	13358 (0.9)	8328 (0.6)	-	-	-	129753 (9.0)	37903 (2.6)	-	-	-
OVO 129	6015 (1.0)	9082 (1.5)	7276 (1.2)	-	-	-	28026 (4.7)	32544 (5.4)	-	-	-
OVO 32	14490 (1.0)	13056 0.9)	9266 (0.6)	-	24050 (1.7)	31407 (2.2)	-	-	-	-	-
OVO 51	2467 (1.0)	2225 (1.1)	1512 (0.6)	-	4562 (1.8)	2769 (1.1)	-	-	-	-	-
OVO 114	22338 (1.0)	23767 (1.1)	25671 (1.1)	-	63081 (2.8)	21709 (1.0)	-	-	-	-	-
OVO 47	23973 (1.0)	21495 (0.9)	34802 (1.5)	-	58857 (2.5)	60738 (2.5)	-	-	-	-	-
OVO 90	2646 (1.0)	3498 (1.3)	4990 (1.9)	-	6446 (2.4)	3190 (1.2)	-	-	-	-	-
OVO 21	6715 (1.0)	4267 (0.6)	6240 (0.9)	-	61141 (9,1)	8092 (1.2)	-	-	-	-	-
OVO 116	2697 (1.0)	4750 (1.8)	5193 (1.9)	-	41385 (15.3))	13486 (5.0)	-	-	-	-	-
OVO 79	601272 (1.0)	894614 (1.4)	1199302 (1.9)	-	1017786 (1.7)	956200 (1.2)	-	-	-	-	-
E48 V02	4802 (1.0)	4584 (1.0)	4683 (1.0)	3274 (0.7)	-	-	-	-	-	-	-
E42 V05	28266 (1.0)	19291 (0.7)	22918 (0.8)	11421 (0.4)	-	-	-	-	-	-	-
E41 V03	6756 (1.0)	9604 (1.4)	10969 (1.6)	6425 (1.0)	-	-	-	-	-	-	-
E43 V04	62771 (1.0)	77399 (1.2)	84240 (1.3)	81219- (1.3)	-	-	-	-	-	-	-
E46 V03	32086 (1.0)	22248 (0.7)	24757 (0.8)	18957 (0.6)	-	-	-	-	-	-	-
E15 V05	12874 (1.0)	6239 (0.5)	10701 (0.8)	3820 (0.3)	-	-	-	-	-	-	-
OVO 65	2232 (1.0)	-	-	-	-	-	-	-	795482 (356,4))	2363 (1.0)	16117 (6.8)
OVO 52	902 (1.0)	-	-	-	-	-	-	-	66300 (73.5)	1045 (1.0)	4123 (3.9)
OVO 142	1954 (1.0)	-	-	-	-	-	-	-	85613 (43.8)	855 (1.0)	10256 (12)
OVO 57	85897 (1.0)	-	-	-	-	-	-	-	124679 (1.5)	124176 (1.0)	112927 (0.9)
OVO 36	1377 (1.0)	-	-	-	-	-	-	-	68483 (49.7)	1757 (1.0)	1045 (0.6)
n=	**24**	**19**	**19**	**6**	**8**	**8**	**5**	**5**	**5**	**5**	**5**

Circulating tumor DNA

Eight samples were quantified for ctDNA. Three samples had time points up to 24 hours (table 5), three had time points up to 48 hours (table 6) and two had time points up to 96 hours (table 7). The mutant fraction was calculated by dividing estimated mutant fragments by the sum of mutant and WT fragments. Each time point was corrected for differences in background relative to baseline sample (0 hr) to see if mutant fractions were the same. The Box-and whisker plots in figure 12 show mutant- and WT ratio at the different time points relative to baseline sample (0 hr). Median WT ratios have a clear increase after 48 hours, whether median mutant ratios are stable is hard to interpret from this dataset. Median ratio of 24 hour samples and 96 hour samples seem to decrease while 48 hour samples are more stable compared to baseline (0 hr). More data needs to be collected to be able to make a conclusion and to be able to perform statistical analysis

Table 5. EDTA blood delayed for up to 24 hr before processing plasma.

	Assay	Time Point	Estimated targets		Mutant fraction		Ratio	Mutant
			Mutant	WT	Replicate	Average	0hr	adjusted with ratio
48V02	PIK3CA p.E545K	0hr (RT)	10 15	330 347	2.6 % 3.7 %	3.2 %	1,0	3,2 %
		6hr (RT)	8 7	376 366	1.9 % 1.7 %	1.8 %	1,0	1,9 %
		24hr(RT)	5 9	400 457	1,1 % 1.7 %	1,4 %	1,0	1,6 %
		24hr (4 °C)	4 6	330 391	1,1 % 1.4 %	1.2 %	0,7	0,9 %
42V05	TP53 p.D281G	0hr (RT)	17 16	493 461	3.3 % 3.4 %	3.3 %	1,0	3,3 %
		6hr (RT)	24 17	461 406	4.9 % 4.0 %	4.5 %	0,7	3.1 %
		24hr (RT)	11 13	422 386	2,5 % 3,3 %	2,9 %	0,8	2,3 %
		24hr (4 °C)	16 28	586 514	2,7 % 5,2 %	3,9 %	0,4	1,6 %
15V05	PIK3CA p.H1047R	0hr (RT)	0 0 0	470 455 603	0 %	0 %	1,0	0 %
		6hr (RT)	1 3	319 366	0,3 % 0,8 %	0,6 %	0,5	0,3 %
		24hr (RT)	2 2	396 441	0,5 % 0,5 %	0,3 %	0,8	0,4 %
		24hr (RT)	0	378	0 %			
		24hr (4 °C)	1 3	423 476	0,2 % 0,6 %	0,4 %	0,3	0,1 %

Table 6. EDTA blood delayed for up to 48 hr before processing plasma

	Assay	Time Point	Estimated targets Mutant	WT	Mutant fraction	Ratio to Ohr	Mutant adjusted with ratio
OVO21 – on chemo	**TP53 p.C135R**	0hr (RT)	25	427	5,5 %	1,0	5,5 %
		6hr (RT)	22	244	8,3 %	0,6	5,3 %
		24hr (RT)	24	477	4,8 %	0,9	4,5 %
		48hr (RT)	6,5	1354	0,5 %	9,1	4,4 %
		48hr (4 °C)	15	422	3,3 %	1,2	4,0 %
OVO116	**TP53 p.C141Y**	0hr (RT)	6,5	139	4,5 %	1,0	4,5 %
		6hr (RT)	4	150	2,6 %	1,8	4,6 %
		24hr (RT)	2,5	274	0,9 %	1,9	1,7 %
		48hr (RT)	5	1892	0,3 %	15,3	4,0 %
		48hr (4 °C)	4	516	0,8 %	5,0	3,8 %
OVO79	**TP53 p.R213X**	0hr (RT)	16	371	4,1 %	1,0	4,1 %
		6hr (RT)	19	552	3,3 %	1,4	4,7 %
		24hr (RT)	30	740	3,9 %	1,9	7,4 %
		48hr (RT)	13	628	2,0 %	1,7	3,4 %
		48hr (4 °C)	19	590	3,1 %	1,2	3,9 %

Table 7. EDTA blood delayed for up to 96 hr before processing plasma.

	Assay	Time Point	Estimated targets		Mutant fraction		Ratio*	Mutant %
			Mutant	WT	Replicate	Average		Adjusted**
56V03	**TP53 p.N239D**	0hr (RT)	20 15	176 203	10,2 % 6,9 %	8,5 %	1,0	8,5 %
		6hr (RT)	41 17	200 194	17,0 % 8,1	12,5 %	0,9	11 %
		96hr (RT)[1]	2 0	502 360	0,4 % 0	0,2 %	21,4	4,9 %
		96hr (RT)	1 1	575 457	0,2 % 0,2 %	0,2 %	29,0	5,6 %
		96hr (4 °C)	23 41	347 373	6,2 % 9,9 %	8,1 %	2,2	18 %
65V03	**TP53 p.Y220C**	0hr (RT)	0 0 0 2	278 272 1109 824	0 0 0 0,2 %	0,1 %	1,0	0,1 %
		6hr (RT)	1 0 0 0	309 316 623 608	0,3 % 0 0 0	0,1 %	0,9	0,05 %
		24hr (RT)***	0 0 0 0	165 297 250 315	0 % 0 0 0	0 %	0,6	0 %
		96hr (RT)	1 2 0 0	301 307 815 599	0,3 % 0,6 % 0 0	0,2 %	9,0	1,33 %
		96hr (4 °C)	0 1 0 0	358 250 2307 2021	0 0,4 % 0 0	0,1 %	2,6	0,05 %

*Relative copy number ratio compared to 0 hr. **Mutant fraction (%) adjusted for elevated wild type background. ***Less template added because sign of inhibition in total DNA-assay.[1]No plasma was processed at 24 hr, extraction replicate for 96 hr.

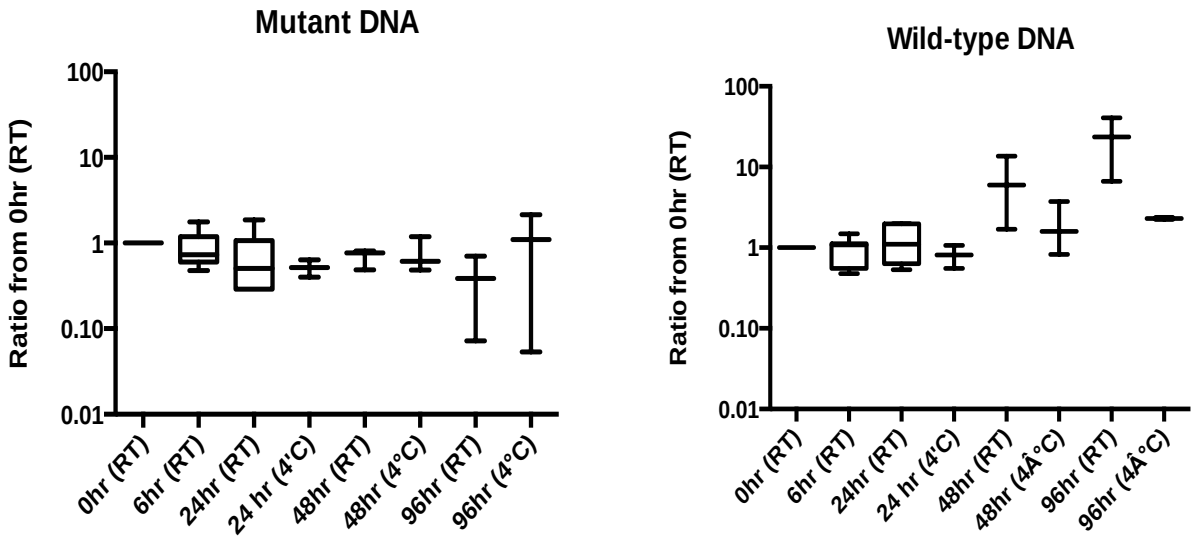

Figure 12. Box- and whisker plot for mutant- and WT fragments relative to baseline (0 hr). The number of samples was too low to do statistical analysis.

PIK3CA p.E545K

The former assay with MGB-probes, designed with Primer Express automatic program, had a high degree of background amplification from non-specific binding (figure 13).

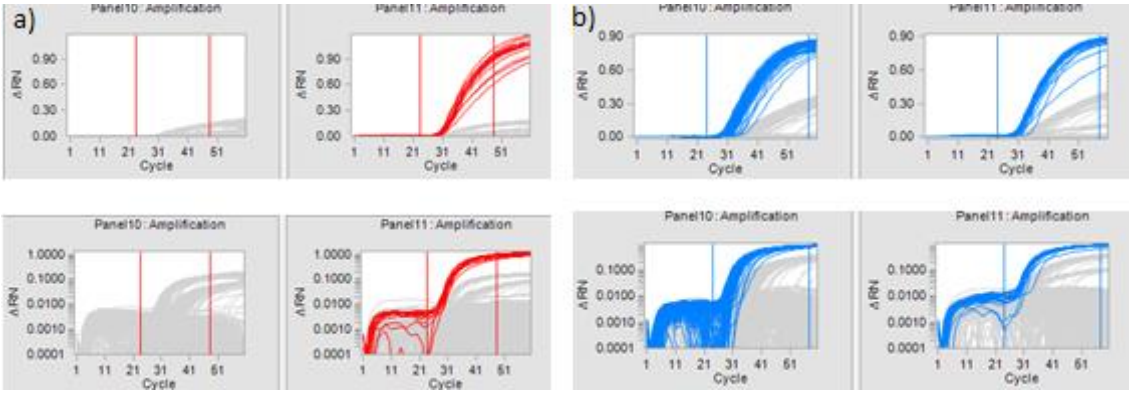

Figure13. Previous assay PIK3CA p.E545K, with wild type material (panel 10) and tumor DET43 (panel 11). a) Mutant probe in linear scale (upper row) and log scale (lower row). b) WT probe in linear- and log scale

A new assay was developed. One mutant probe and three WT probes were tested with tumor material (panel 3 and 4) and PPC (panel 5). The mutant probe discriminated well

(figure 14a). WT1 detected both wild type and mutant sequences (14b), WT2 (shorter WT1 probe) discriminated well but there was still background (14c). WT3 seem to be unable to

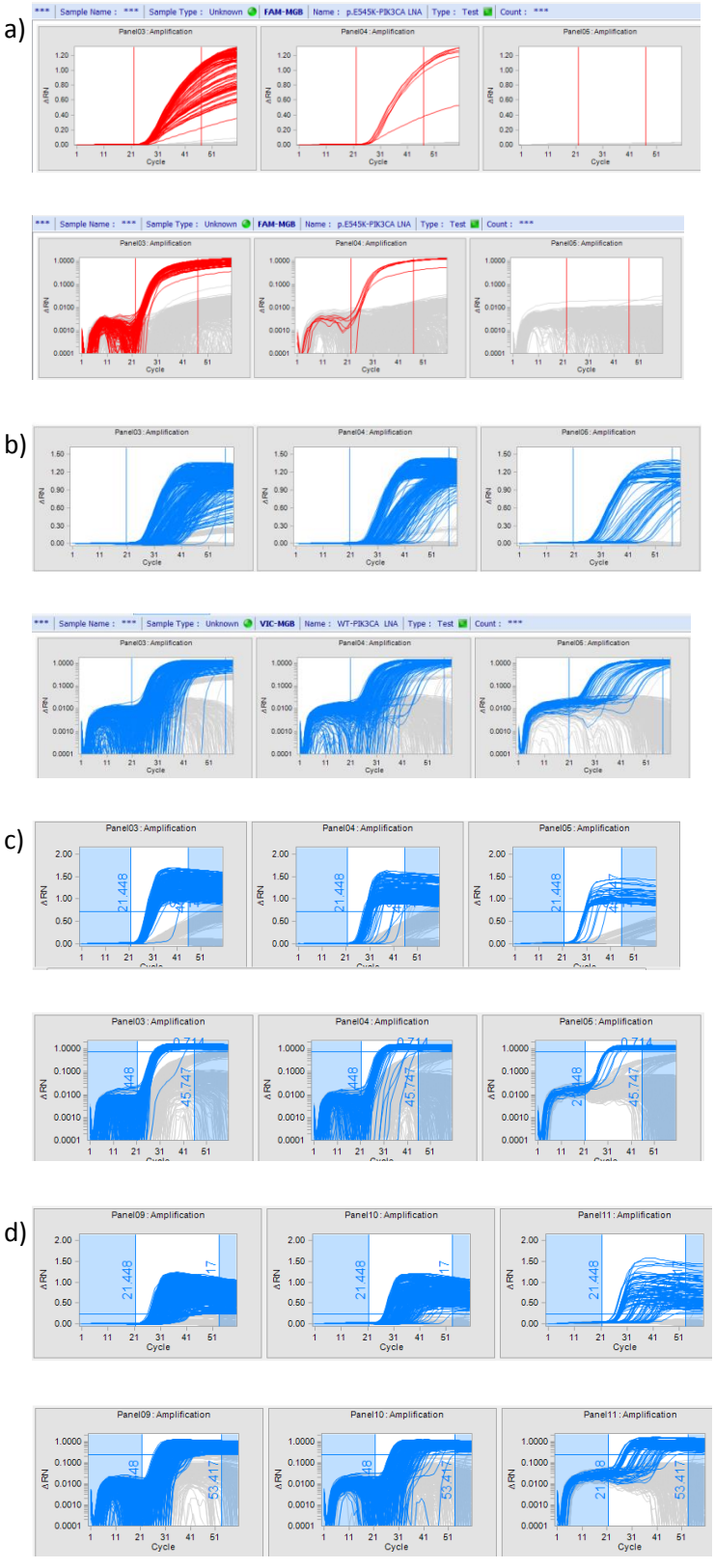

Figure 14. Primary tumor (PT) DET43 in panel 3, PT diluted in buffy coat (BC) in panel 4 and a pooled normal plasma control (PPC-3) in panel 5. a) The mutant LNA-probe (red curves) discriminated well between mutant and wild type alleles. The background level was very low. b) The WT1-probe (blue curves) was too strong and bound mutant sequences. c) WT2 was shorter then WT1 to increase mismatch-ΔTm and discriminated well. The background is probably due to non-specific amplification of chromosome 22 as discussed below. d) The WT3-probe had LNA –bases away from the mismatch site. This probe seems to be non-specific.

discriminate, the threshold between background and signal was hard to set (14d). The heatmap (figure 15) helps indicate if amplification is specific. This run was too loaded to draw conclusions.

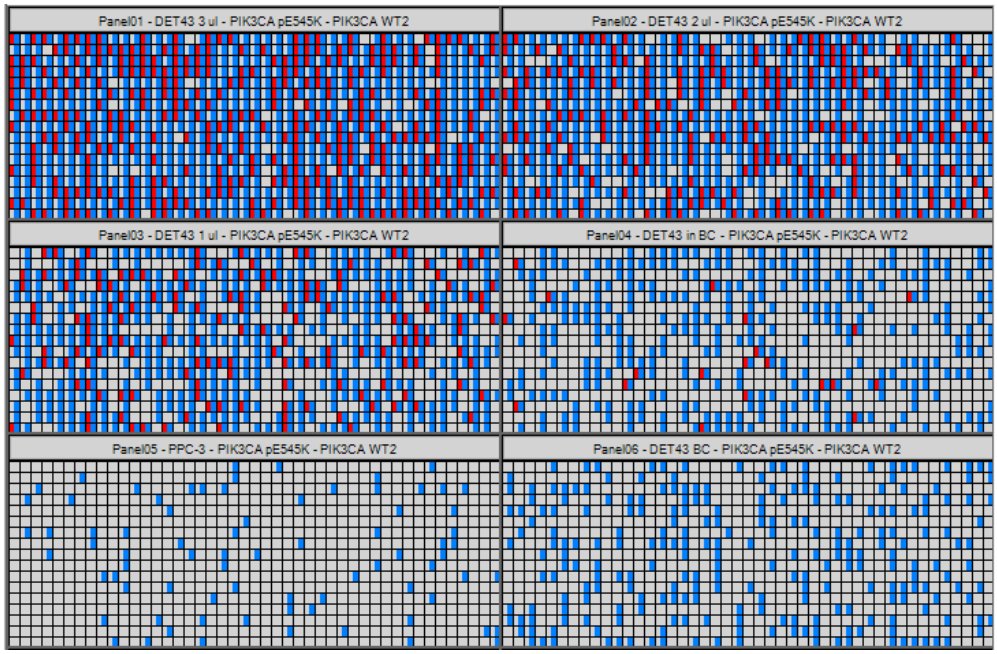

Figure 15. HeatMap of PIK3CA p.E545K with WT2 probe. The upper four panels are tumor material. The lower two are normal material. Red signals are mutant fragments and blue are WT

BLAST of the whole amplicon showed that PIK3CA is highly homolog to chromosome 22. The allelic probes have one mismatch to chromosome 22 and the reverse primer has 2 mismatches (figure 16).

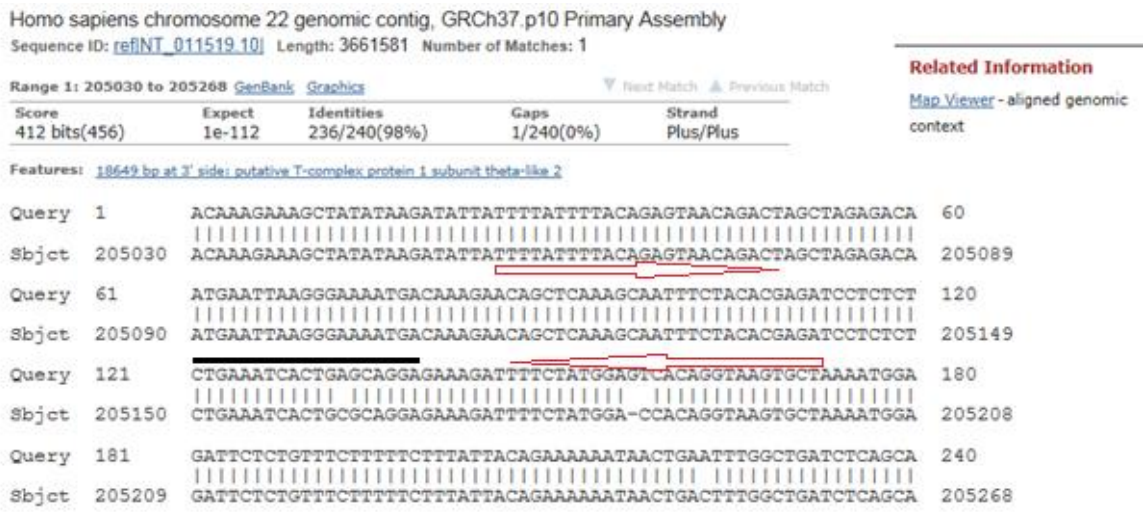

Figure16. *PIK3CA* homology with chromosome 22. PIK3CA p.E545K assay is indicated with arrows. Probe position is indicated with a black line. The reverse primer has two mismatches, wild type probes has one mismatch and mutant probe has two mismatches.

TP53 p.N239D

No cell line was available for testing this assay, so FFPE material from primary tumor (PT), DET56 was used. Forward primer 2 (panel 7 in figure 17) gave the highest yield of the two forward primers that were tested. The FFPE material was of suboptimal quality and had signal only in one of the PT replicates with the wild type probe. Targeted MPS had quantified this tumor to have 64 % mutant load.

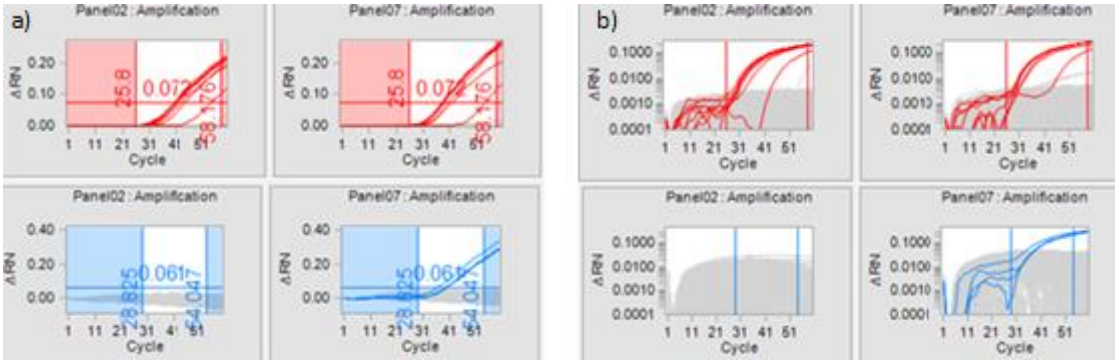

Figure 17. TP53 p.N239D assay tested on PT with forward primer 1 in panel 2 and forward primer 2 in panel 7. Mutant probe had red signals and WT probe had blue signals. a) Linear scale and b) log scale. PT was from FFPE material and had sub optimal quality.

The WT probe amplified as expected in fresh material, so it was not compromised (figure 18). Even though the test material is not optimal, one could interpret that both probes were specific. The background amplification that can be seen for instance in panel 11, upper row in figure 2, was due to bleed-through from spectral overlap and not unspecific binding. The signal-to-noise was not high, especially for the WT probe (VIC-MGB). The heat map shows all of the 765 wells per sample and which wells that has got signal (figure 19). The fragments were randomly distributed and both probes amplified on their own (panel 7), which strongly suggests that they are specific. If the probe with the lowest number of signals always amplifies in the same well as the other probe, this might be an indication of non-specificity.

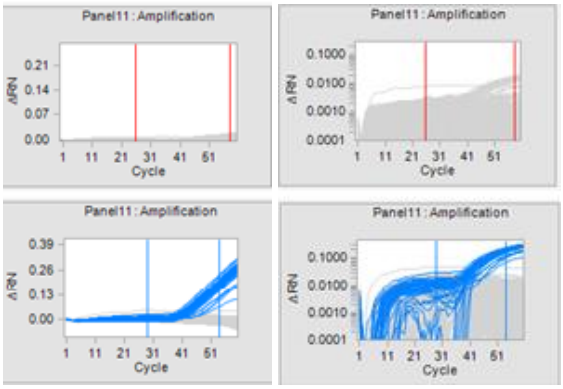

Figure 18. TP53 p.N239D assay. Normal plasma control with mutant probe in upper row and WT probe in lower row

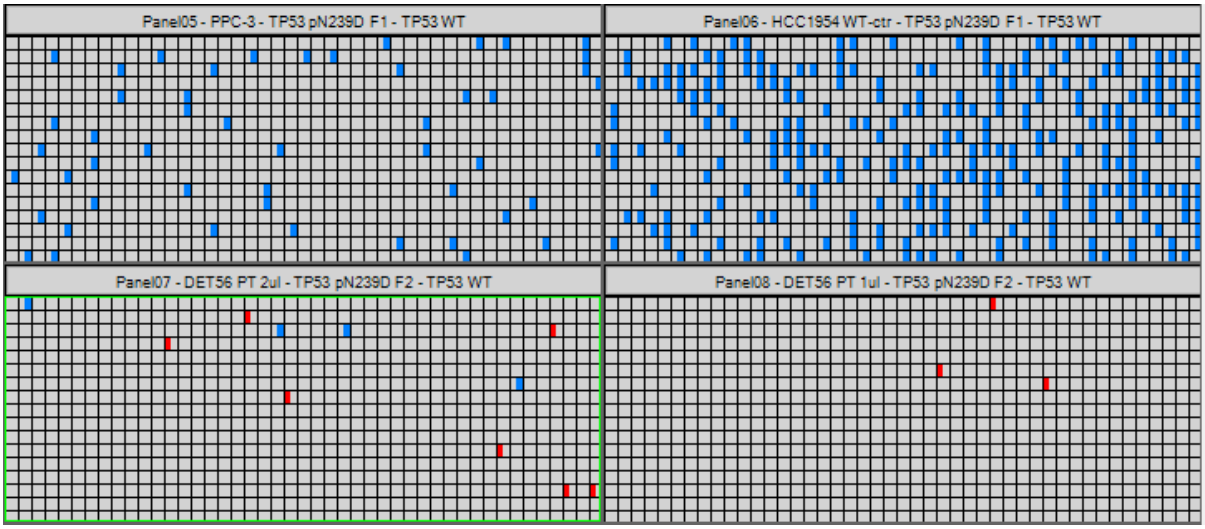

Figure 19. Heat Map for TP53 p.N239D Assay. Each panel represents an individual sample dispensed to 765 microwells. Blue signals indicate amplification of WT fragments, and red signal represents mutant fragments. The two upper panels had wild type material, and the two lower panels were tumor DNA from FFPE material.

TP53 p.Y220C

The assay was tested with the cell line HCC1419 which is supposed to be homozygous for the mutation according to the COSMIC database (52). The dPCR assay quantified the two replicates of HCC1419 to only have a mutation load of 32 % and 34 %. This is however in agreement with a capture based targeted deep sequencing done on the same material, which quantified the mutation load to be 31 %.

The assay was tested with two different reverse primers. Reverse primer 1 gave a higher yield than reverse primer 2 and was used in further assays with TP53 p.Y220C.

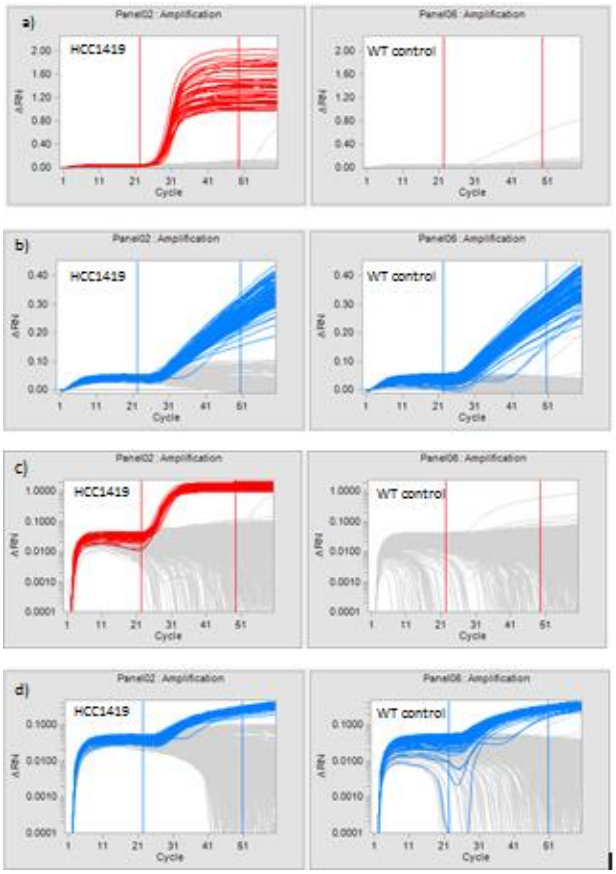

Figure 20. TP53 p.Y220C with reverse primer 1. The red curves represent the mutant LNA-probe labeled with ROX and the blue curves represent the WT MGB-probe labeled with VIC. The mutant cell line, HCC1419, is on the left side of the figure and the WT control is on the right side of the figure. a) Linear scale of mutant probe. WT control has a signal, but is not considered a true signal due to the shape of the curve. B) Linear scale of the WT probe. C) Mutant probe in log scale. D) Wild type probe in log scale. The signal-to-background ratio (ΔRn) is low with the wild type probe.

Figure 20 and 21 shows the results on HCC1419 and a WT control with reverse primer 1 and 2 respectively. Red signals represent the mutant LNA probe labeled with ROX and the blue signals represent the wild type MGB-probe labeled with VIC. The mutant probe has a high signal-to-noise ratio (ΔRn) and the signal in the WT control is not considered true due to the shape of the curve. The WT probe has a lower ΔRn but the threshold could be set. The heat map shows all of the 765 wells per sample and which wells that has got signal (figure 22). The fragments were randomly distributed and both probes amplified alone, which give strength to them being specific. Reverse primer 1 has higher signal-to-noise ratio than reverse primer 2.

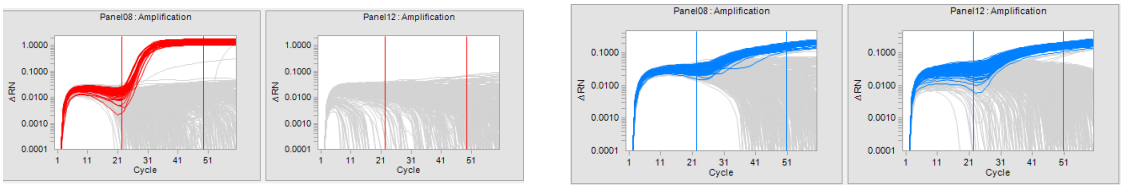

Figure 21. TP53 p.Y220C assay with reverse primer 2. Mutant probe in red and WT probe in blue. The assay gives a lower ΔRn than reverse primer 1.

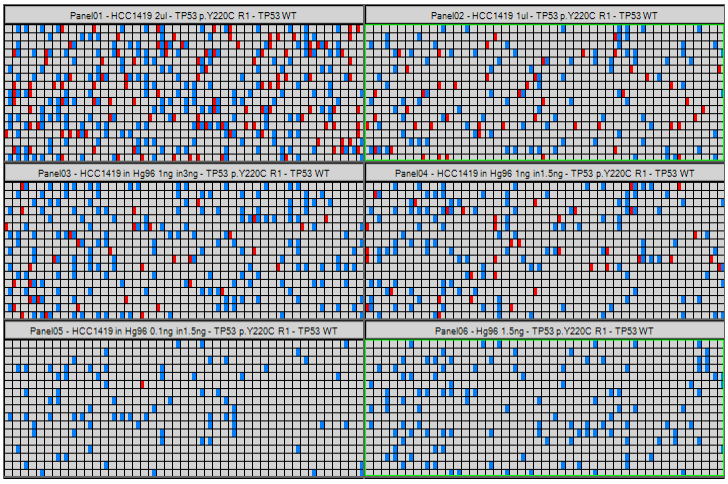

Figure 22. Heat Map of the TP53 p.Y220C assay showing 6 individual panels representing 6 different samples. Each panel has 765 microwells. The fragments showed random distribution and amplification with a single probe in many of the wells, supporting evidence that the probes were able to discriminate between wild type and mutant sequence. Panel 2 in the upper right is the HCC1419 in figure 1 and panel 6 in the lower right is WT control in figure 1.

There is no spectral calibration on BioMark™ HD System, so a certain degree of background is expected due to overlapping emission spectra. Figure 23a shows the TP53 p.Y220C assay run with mutant probe only (red signal). The grey curves in the right panel (figure 23b) are signals from mutant probe interpreted as wild type probe signal by the BioMark.

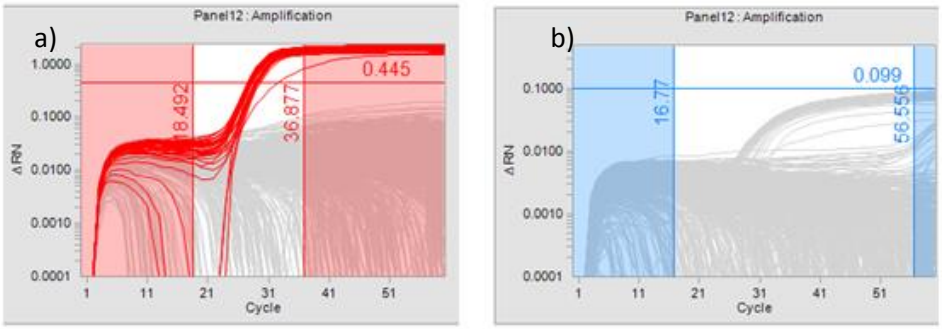

Figure23. Spectral overlap. TP53 p.Y220C assay run with mutant probe only (red). The signal in the panel to the right is due to spectral overlap between the probes.

2. Analysis of circulating tumor DNA to monitor metastatic breast cancer

The application of ctDNA analyses for disease monitoring requires the identification of somatic alterations in individual patients. In an initial study, already published on this cohort (35), ctDNA was detected through the identification and monitoring of *TP53* and *PIK3CA* mutations in plasma. *PIK3CA* and *TP53* mutations are found in up to 36% and 37% of breast cancers respectively and combined screening for these mutations identifies genomic alterations in ~50 % of the metastatic breast cancer (MBC) cases.

Recent large scale sequencing efforts have identified several novel cancer genes to be recurrently mutated in breast cancer (20, 53). The ability to detect these mutations in the plasma of women with MBC and their clinical application for ctDNA monitoring has not previously been studied. The initial study in MBC, established that tracking ctDNA through the quantification of *TP53* and *PIK3CA* mutations in plasma closely paralleled changes in tumor burden providing the rationale for using ctDNA in disease monitoring. In the remaining 50 % of patients without a *PIK3CA* or *TP53* mutation it is currently unclear what proportion will have at least one of the other recurrent mutations detected in plasma. It is also unclear whether changes in ctDNA levels, as assessed by following these low frequency mutations, will also correlate closely with clinical parameters.

Material

52 patients from the metastatic breast cancer study DETECT (Identification and classification of circulating tumor cells and circulating nucleic acids in patients with metastatic breast cancer) were included. The cohort included patients receiving active therapy (either endocrine therapy, chemotherapy or biological therapy) for metastatic breast cancer. Serial blood samples were collected throughout the study (April 2010-April 2012) at intervals of ≥3 weeks for both circulating DNA and circulating tumor cell (CTC) analysis. EDTA –tubes (Sarstedt) were used to collect plasma DNA and CellSave® tubes (Veridex) to collect CTC's

Plasma was processed within one hour as described in at subproject 1 and frozen at -80°C until DNA extraction with QIAamp Circulating Nucleic Acid Kit (Qiagen) as described in subproject 1. After the collection of plasma, the remaining buffy coat was removed and

subjected to red cell lysis buffer (155 mM NH_4Cl, 10 mM $KHCO_3$ and 0.1 mM EDTA pH7.4). Buffy coats were stored at -80°C until DNA was extracted with a standard phenol-chloroform method (Appendix B). CellSave® tubes were processed within 96 hours using CellSearch® system (Veridex) according to the manufacturer's instructions. In brief, this is a semi-automated processing system which isolates CTC's using magnetic beads coated with antibodies against epithelial cell adhesion molecule (EpCam). Isolated cells are stained with the nuclear dye 4', 6-diamidino-2-phenylindole (DAPI) and antibodies against cytokeratin 8, -18, -19 and the common leukocyte antigen CD45. The criteria to define a cell as CTC include round to oval morphology, a visible nucleus, positive staining for cytokeratins and negative staining for CD45. Archival primary – or metastatic tumor tissue was obtained from all study participants. DNA from paraffin-embedded formalin-fixed (FFPE) tumor was extracted with Ex-Wax DNA Extraction Kit (Millipore) and fresh frozen tumor was extracted with DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's instructions (35).

Methods

Libraries for targeted MPS were made using 48.48 Access Array™ system (Fluidigm) and the sequencing was done on HiSeq2000 (Illumina). 48.48 Access Arrays™ are microfluidic chips with 48x48 micro wells for parallel amplifications. Samples and primer-mixes are transferred into the micro wells by air pressure, so that each of the 48 samples blends with each of the 48 primer mixes (Figure 24).

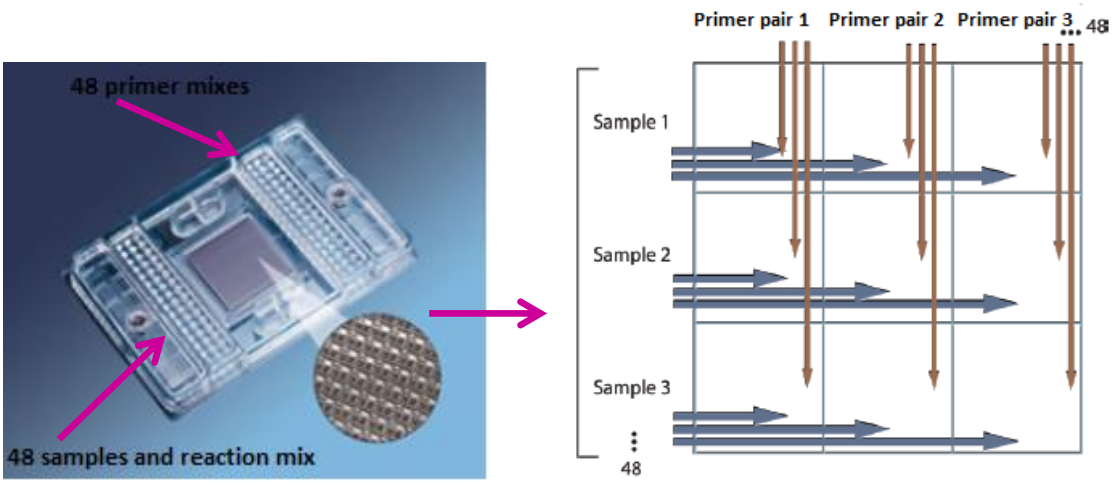

Figure24. The 48.48 Access Array mixes 48 samples with ≥48 primer pairs for parallel amplification in micro wells. The figures are modified (54, 55)

Illumina sequencing

12 libraries were prepared in total. Due to the 384 different barcodes, 8 and 4 libraries could be pooled together. Sequencing was done on HiSeq2000 and the libraries were injected to 3 of the 8 lanes in a flow cell. The flow cell has sequence adapters, PE1 and PE2, bound to the surface, to which single stranded library fragments bind after being denatured. The flow cell's sequence adapters are also used as primers for bridge amplification. The bound fragments will bend over and hybridize to a nearby complementary adapter and copied. Bridge amplification generates multiple copies around each fragment, called clusters. Each cluster represents one sequence read. After cluster generation, the reverse strand is removed and sequence is done on the remaining strand. The 4 dNTP's are differently labeled with fluorophores and 3'-blocked to prevent more than one dNTP to be incorporated per cycle. All 4 dNTP's are introduced at every cycle; this natural competition for binding will minimize bias. After laser excitation, fluorophores and 3' blocks are removed and the next nucleotide can bind. This form for sequencing is called sequencing-by-synthesis (SBS). A 100 bp of the sense strand libraries were sequenced, then 10 bp barcodes were read before the anti-sense strand was generated and sequenced for a 100 bp (Figure 25).

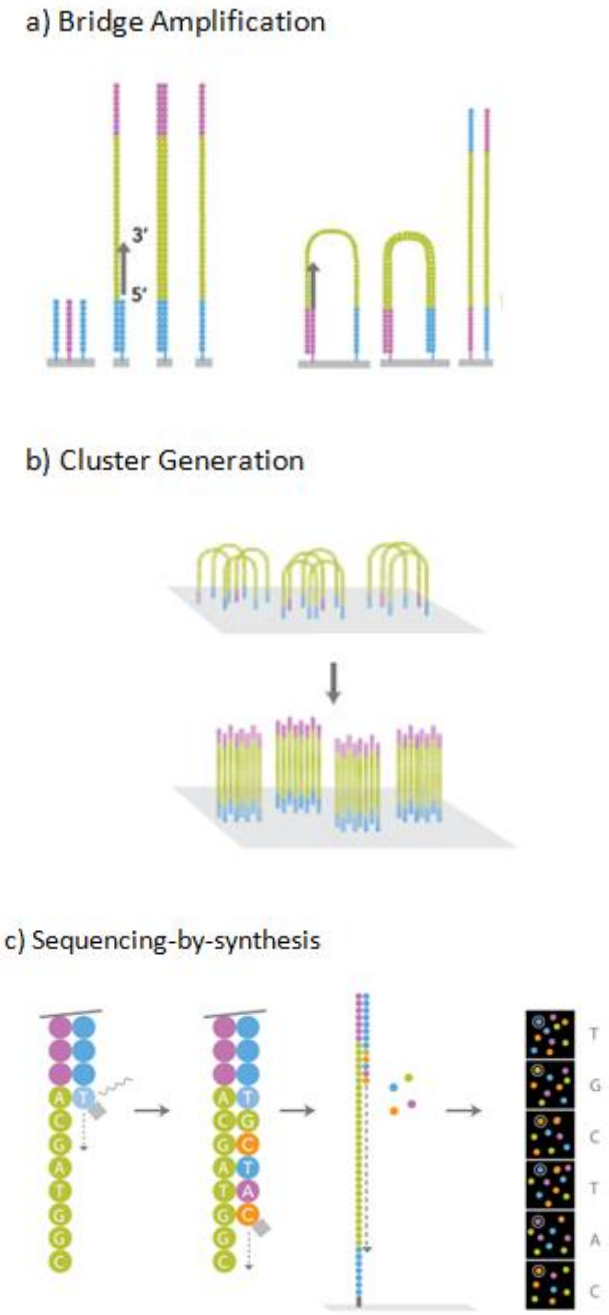

Figure 25. Sequencing-by-synthesis. a) Fragments are denatured and hybridized to sequence adapters on the flow-cell surface. b) Each fragment is multiplied by bridge amplification into clonal clusters. The reverse strand is cut off and clusters are sequenced in parallel. c) Sequence is read during synthesis. After incorporation of a nucleotide, elongation is terminated due to a 3'block on the nucleotide. Fluorescence from the labeled nucleotide is read before dye and 3'block are cleaved and a new nucleotide can be incorporated. For paired-end sequencing, the second strand is re-generated by bridge amplification when sequencing of the first strand is complete. The first strand is cut off and sequencing starts on the reverse strand.

Primers

Target specific primers were designed for somatic mutations in 17 genes (table 8) with Primer3 and tagged with 5' universal sequences (CS1 and CS2, Illumina). Primer sequences for *TP53* and *PIK3CA* are listed in Forshew's Supplementary Materials (56) . The rest of the primers will be published with the forthcoming paper on this cohort.

Pre-amplification

Because samples are split into many parallel reactions, some samples need pre-amplification due to low concentration or degraded material to ensure that each micro well receive template. 5 µl plasma DNA and 50 ng DNA from FFPE tumors were pre-amplified in a total volume of 10 µl. The pre-amplification was done in multiplex with a final concentration of 50 nM primers, 0.5 U FastStart High Fidelity Enzyme Blend (Roche), 1x High Fidelity Enzyme Buffer, 200 µM of each dNTP, 4.5mM $MgCl_2$ and 5 % DMSO. An initial 10 min at 95°C to activate the enzymes and denature DNA, was followed by 15 cycles of amplification at 95°C for 15 s and 60°C for 4 min. Pre-amplified templates were subjected to ExoSAP-IT to degrade unused primers and unincorporated dNTP's. ExoSAP-IT contains two enzymes, an exonuclease I that cuts single stranded molecules and Shrimp Alkaline Phosphatase that dephosphorylates dNTP's. The reaction took place at 37°C for 15 min , followed by 80°C for 15 min to inactivate the enzymes (57). Pre-amplified products were diluted 5-fold before target specific amplification in 48.48 Access Arrays.

Table 8. The 17 genes that were screened for SNV's and number of amplicons for each gene

Gene	Amplicon #	Gene	Amplicon #	Gene	Amplicon #
AKT1	2	GATA3	18	CDKN1B	5
AKT2	1	CDH1	25	PTEN	14
CASP8	11	EGFR	13	KRAS	1
AR	7	MAP3K1	44	TBX3	2
TP53	16	MAP2K4	12	BRAF	1
PIK3CA	2	SF3B1	1		

Target specific amplification

A 96-well plate containing 48 different primer mixes with 180 primer pairs in total was made according to Appendix B to a final concentration of 6 µM primers and 1x DNA Suspension Buffer (10 mM TRIS, pH8.0, 0.1 mM EDTA, TEKnova). A 4 µl pre-sample reaction mix was made of 0.05 U FastStart High Fidelity Enzyme Blend, 1x High Fidelity Enzyme Buffer, 200 µM of each dNTP, 4.5mM $MgCl_2$, 5 % DMSO and 1x Access Array Sample Loading Reagent (Fluidigm). 1 µl pre-amplified product or 50 ng DNA was added to the reaction mix, then 4 µl of the mix was loaded to the sample inlets of the 48.48 Access Array™ and 4 µl primer mix was added to the primer inlets. Wells named H1 to H4 were loaded with 500 µl 1x Access Array Harvest Reagent (0.05 % Tween-20). Samples and primers were loaded into the 48.48 Access Array™ using a Post-PCR IFC AX Controller (Fluidigm, figure 26) then the array was transferred to the thermal cycler FC1 (Fluidigm) for PCR. PCR-conditions were 50°C for 2 min, 70°C for 20 min, 95°C for 10 min. Then 10 cycles; 95°C for 15 sec, 60°C for 30 sec , 72°C for 1 min, 2 cycles; 95°C for 15 sec, 80°C for 30 sec, 60°C for 30 sec , 72°C for 1 min; 8 cycles of 95°C for 15 sec, 60°C for 30 sec , 72°C for 1 min; 2 cycles of 95°C for 15 sec, 80°C for 30 sec, 60°C for 30 sec , 72°C for 1 min; 8 cycles of 95°C for 15 sec, 60°C for 30 sec , 72°C for 1 min; 5 cycles of 95°C for 15 sec, 80°C for 30 sec, 60°C for 30 sec, 72°C for 1 min.

Harvest

After PCR, reagents in well H1 to H4 were replaced with 600 µl 1x Access Array Harvest Reagent and sample inlets were loaded with 2 µl 1x Access Array Harvest Reagent to harvest PCR-products in a Post-PCR IFC-AX Controller. 10 µl PCR-products were pipetted from the sample inlets on to a 96-well plate after harvest.

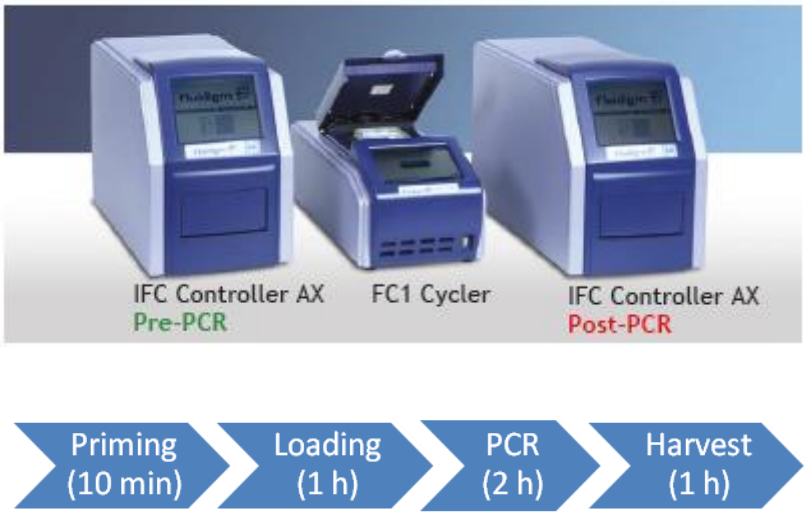

Figure 26. Workflow. Priming and loading of the 48.48 Access Array™ takes place in a Pre-PCR IFC Controller AX. The Access Array is transferred to FC1 Cycler for PCR, and PCR-products are harvested in a Post-PCR IFC Controller AX (Fluidigm).

Barcoding

Harvested PCR-products were diluted 1:100 in DEPC-free water. 1 µl of the diluted products were tagged with barcodes and Illumina sequence adapters, PE1 and PE2 in a new PCR. There are 384 unique barcodes which gives the possibility to pool 384 samples and sequence them simultaneously. The barcoding was done in a total volume of 10 µl, with a final concentration of 1xFastStart High Fidelity Reaction Buffer, 4.5 mM $MgCl_2$, 0.2 mM PCR Grade Nucleotide Mix, 5 % DMSO and 0.4 µM Access Array Barcode Library (Illumina). The thermal cycling was 95°C for 5 min, then 15 cycles of 98°C for 20 s, 60°C for 15 s and 72°C for 1 min, and a final hold at 72°C for 3 min. When each of the 48 pools of PCR-products had got its unique barcode, samples were pooled (Figure 27). This pool represented one library.

Library Cleanup

Cleanup was done with magnetic solid phase reversible immobilization (SPRI) beads to remove excess primers and primer-dimers. The beads bind DNA in the presence of polyethylen glycol (PEG) and salt. The size of the products binding to the beads depends on the ratio between PCR-products and beads. A 1.8 ratio was used to bind 300 bp amplicons. Bound PCR-products were washed twice with 70 % ethanol, dried at 37 °C until beads were slightly cracked and eluted in 20 µL DEPC-free water. See Appendix B for complete protocol.

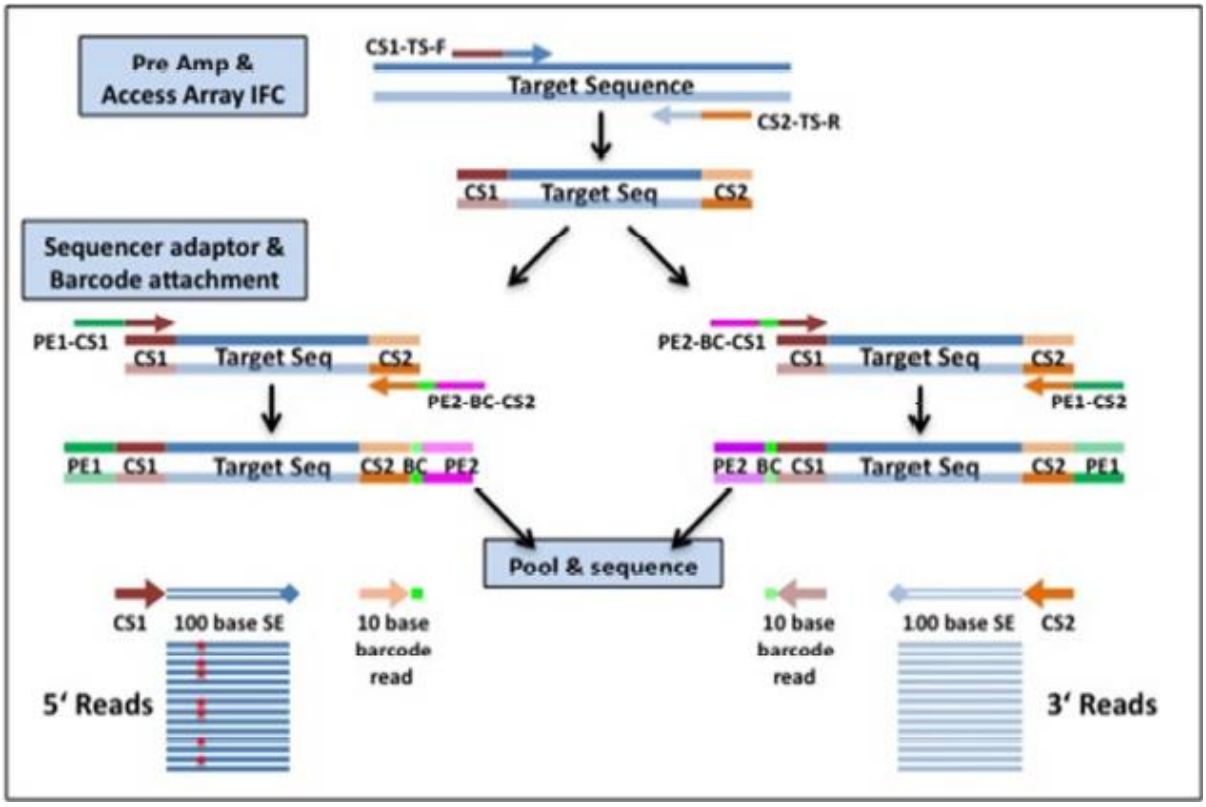

Figure 27. Target specific primers were tagged with universal adapter sequences, CS1 or CS2 at 5'end. The CS-adapters were used in the barcoding reaction to add Illumina sequence adapters (PE1 and PE2) and unique barcodes to the amplicons. Samples were then pooled and sequenced paired-end (56)

Library Quantification

Too much material applied on the flow cell will lead to overlapping clusters and noise while too little material will reduce the yield of reads. Libraries were therefore quantified with qPCR prior to sequencing. Libraries were also controlled for correct size with DNA1000 kit on Bioanalyzer (Agilent). Expected size was around 300 bp including tags and barcodes. The Bioanalyzer had been found to be inaccurate on quantification, so libraries were quantified with KAPA Library Quantification Kit (KAPA Biosystems) on ABI7900HT to ensure optimal sequencing. Libraries were diluted $1{:}10^3$, $1{:}10^4$, $1{:}10^5$ and $1{:}10^6$ and run in triplicate. The kit comprised six 452 bp standards and a KAPA SYBR FAST qPCR Master Mix that contained primers that annealed to PE1 and PE2. PCR was done in 10 µl volumes containing 6 µl Master Mix and 4 µl template on 384-well plates. The FAST PCR-program was 20s at 95 °C, then 35 cycles at 95°C for 1 s and 60°C for 20 s, then a dissociation curve (95°C for 15s, 60 °C for 15s

and 95°C-15 s, 2 % ramp rate) was run to check for primer dimers. Library concentrations (pM) were calculated with the following formula:

$$Average\ concentration\ x \frac{452}{average\ fragment\ lenght} x\ dilution$$

Bioinformatics

The sequence analysis pipeline used for this project has been described elsewhere (56). Briefly, sequences were first run through two initial quality control pipelines and de-multiplexed. The primary pipeline converted raw data to FastQ files. The secondary pipeline was a whole genome alignment to check that the sequences aligned to the correct species. Then sequences were de-multiplexed by grouping those with the same index (barcode) together and aligned to human genome 19 (hg19, UCSC Genome Browser). After alignment, data was broken into amplicons by using the primer sites. Only Q30 bases were kept for further analysis. The Q-score indicates the probability that a given base is called incorrectly, and with Q30 the probability is less than 1 in 1000. This gives a base call accuracy of 99.9 %. Potential single nucleotide variants (SNV's) were called and known single nucleotide polymorphisms (SNP's) from the 1000 Genome database were discarded leaving a list of candidate mutations. Already known SNV's was quantified directly. The list of somatic candidate mutations was further investigated by comparing primary tumor with its normal match, the buffy coat sample, to exclude germline mutations (Figure 28).

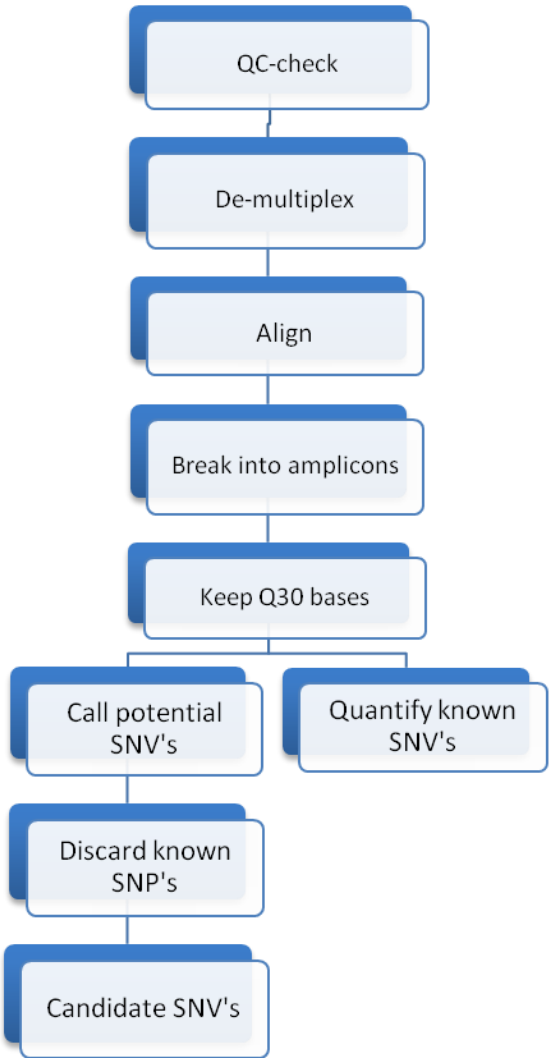

Figure 28. Sequencing analysis pipeline. After an initial quality check, samples were de-multiplexed and aligned to the human genome, hg19. The primer sites were used to break the data into amplicons and only Q30 bases were kept. Potential SNV's were called and known SNP's from the 1000 genome database were discarded, leaving a list of candidate SNV's. Known SNV's could be quantified directly.

Preliminary results

The detection rate in primary or metastatic tumor increased from 48 % to 94 % (49/52) by including more genes in the screening for somatic point mutations. More than one mutation was identified in 46 % (24/52) of the patients and 45 patients had point mutations detected in plasma (92 %). Some of the patients had mutations that showed different dynamics. Case 1 had somatic mutations detected both in *PIK3CA* and cadherin 1 gene (*CDH1*) at presentation. During treatment with capecitabine, *PIK3CA* mutation-load in plasma fell to non-detectable levels while the *CDH1* mutation persisted unchanged. *PIK3CA* reappeared in

plasma and both mutations increased following cessation of capecitabine chemotherapy. *CDH1* dropped to undetectable levels during subsequent treatment with exemestane while *PIK3CA* responded to some degree (figure 29). Case 18 had two mutations that showed more similar fluctuations. Only *TP53* was detected at diagnosis but *KRAS* appeared after treatment with exemestane. Both mutations stayed at stable high levels during further treatment with vinorelbine (figure 30). Disease status was measured with imaging (CT) and serum CA 15-3 and numbers of CTC's were measured simultaneously as ctDNA.

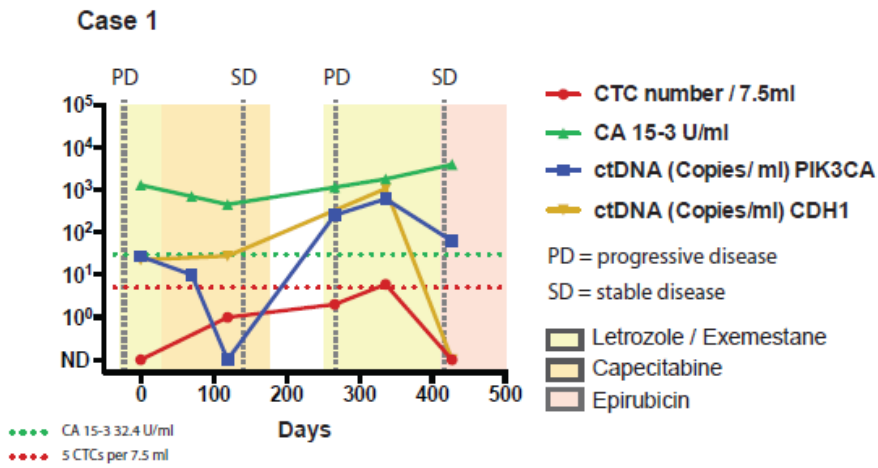

Figure 29. Case 1 had somatic point mutations detected in *PIK3CA* and cadherin 1 gene *(CDH1)* at presentation. Serial monitoring of the mutations in plasma DNA during treatment showed different dynamics. Disease status was ascertained by CT. ND denotes not detected.

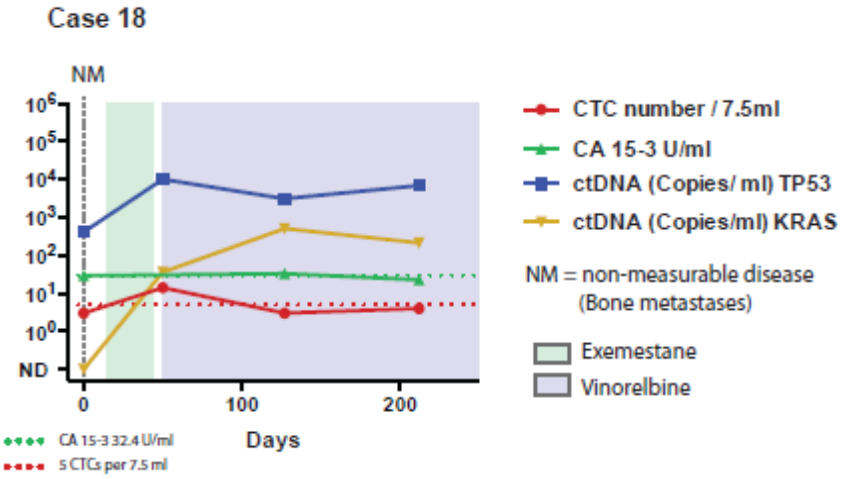

Figure 30. Only *TP53* mutation was detected in case 18 at presentation but *KRAS mutation* was subsequently detected in ctDNA following the commencement of therapy. The mutations show similar dynamics once *KRAS* is established. The patient had bone metastasis so disease status could not be ascertained by CT. ND denotes not detected.

3. Detection of tumor specific amplification in blood by targeted MPS

Background

Circulating tumor DNA is often a minor fraction of the total cell-free DNA in blood. Using an amplified gene as tumor marker might increase the possibility to detect ctDNA above the wild type background. Amplifications are in general a monoallelic event (58), i.e only one of the parental alleles amplify, the other remains normal. Single nucleotide polymorphisms (SNP's) can therefore be used to detect amplification based on allelic imbalance. *ERBB2* is the most frequently amplified gene in breast cancer and over-expression of its protein (HER2) can be treated with targeted therapy. HER2 over-expression is routinely diagnosed with IHC and FISH on primary tumor tissue. Studies have however shown that that HER2 status might change during disease progression (59, 60). This status change can be a consequence of intratumor heterogeneity and clonal evolution or induced by therapy. Seol et al examined 96 *ERBB2* amplified breast cancers and saw regional intratumor heterogeneity in a subset of the patients by examining several core biopsies from the same tumor in tissue microarrays. This pin points both the importance of examining several biopsies and to follow-up with repeated testing in case of equivocal results. Monitoring ctDNA in blood is non-invasive and has the potential to provide a powerful tool for the serial analysis of HER2 status during disease progression and treatment (35) but it requires a sensitive test. A 5-fold amplification of *ERBB2* will be seen as 1.2-fold in plasma if only 5 % of the total plasma DNA originate from tumor. Several papers have recently been published on detection of *ERBB2* amplification in plasma with qPCR and dPCR (61-63). Digital PCR is the more sensitive of the two methods, with a capacity to detect 1.2 fold changes. This projects intent was to elucidate if massively parallel sequencing using SNP ratios could increase sensitivity even further.

Material

The *ERBB2* amplified cell line HCC1954 (ATCC#CRL-2338D) was diluted into its normal match, the lymphoblastoid cell line HCC1954 BL (ATCC#CRL-2339D) in 50 ng/µl dilutions from 50 % tumor down to 0.1 %, as in table 9. In addition, five primary tumors from patients with metastatic breast cancers were analyzed. Two of them were formalin fixed paraffin

embedded (FFPE) primary tumors (DET28 and -56) and three were fresh frozen (DET42, -48 and -59). DNA had previously been extracted from these tumors with Ex-Wax™ DNA Extraction Kit (Millipore) and DNeasy Blood and Tissue Kit (Qiagen). DNA from each patient's buffy coat was used to screen for informative SNP sites. Buffy Coats had previously been extracted with a standard phenol-chloroform method. See Appendix B for protocol. DNA from three normal samples (B-lymphocytes), Hg00096, -97 and -98 (Coriell Institute for medical research) were also analyzed.

Table 9. HCC1954 was diluted into its normal match HCC1954 BL from

50 % tumor down to 0.1 % tumor

50 ng/µl Dilutions	Mutant, HCC1954 (µl)	Normal, HCC1954 BL(µl)	Total volume (µl)
50 %	5	5	10
30 %	3	7	10
20 %	2	8	10
10 %	1	9	10
5 %	µl 50 % HCC1954	µl HCC1954 BL	
	1	9	10
0.5 %	µl 5 % HCC1954	µl HCC1954 BL	
	1	9	10
1 %	µl 10 % HCC1954	µl HCC1954 BL	
	1	9	10
1 %	µl 20 % HCC1954	µl HCC1954 BL	
	0.5	9.5	10
0.1 %	µl 1 % HCC1954	µl HCC1954 BL	
	1	9	10
Total volume template	15.5	65.0	

Methods

Primer design and validation

We searched for the most frequent heterozygous SNP sites in *ERBB2* through NCBI's Variation Viewer and UCSC Genome Browser and found 24 SNP sites with minor allele frequency (MAF) over 0.16. 10 of them were situated in repetitive regions, identified with

Repeat Masker (UCSC Genome Browser) or close to repetitive motifs and could not be used. The gene sequence, *ERBB2*-001 ENST00000584601 was downloaded from Ensembl and primer 3 (http://frodo.wi.mit.edu/) was used to design 35 primer pairs covering the whole gene, including both frequent (figure 31) and less frequent SNP sites. Primers with melting temperature (Tm) of 60 °C and amplicons between 80-120 bp were designed if possible. Secondary structures and dimerization was checked for with Primer Express v3.0 (Life Technologies). Non-specific amplification was checked for with UCSC In-Silico PCR and Primer-Blast (NCBI). Forward primers were tagged with common sequence 1 (CS1) and reverse primers with common sequence 2 (CS2). These are universal sequences (Illumina) to be used for barcoding and adding of sequence tags after target specific PCR. Primers arrived diluted to 100 µM in RNase-free water (Sigma Aldrich) and were mixed to 50 µM primer pairs in a 96-well plate.

Primers were validated in 10 µl total volume with a final concentration of 1x KAPA HiFi Hotstart Ready Mix (KAPA BIOSYSTEMS), 0.45 µM primers, 5 % DMSO and 50 ng DNA. 1x KAPA HiFi Hotstart Ready Mix contained 2.5 mM $MgCl_2$ and 0.3 mM of each dNTP. The enzyme concentration was not stated by the manufacturer. KAPA HiFi DNA Polymerase is an antibody-based hotstart enzyme with proofreading activity. The fidelity is approximately a 100x higher than wild-type Taq and up to 10x higher than similar DNA polymerases.

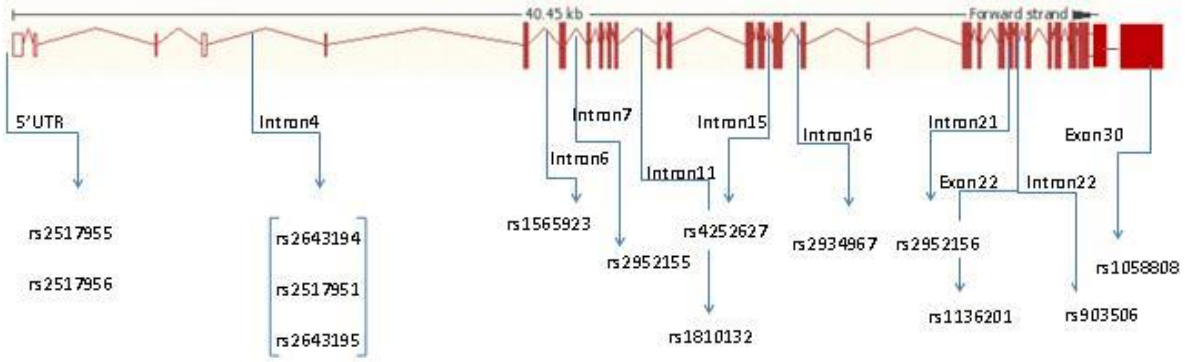

Figure 31. Frequent heterozygous SNP's in *ERBB2*. The three SNP's in intron 4 are on the same amplicon.

Target specific PCR and MPS

Singelplex target specific PCR was done with Fluidigm 48.48 Access Array IFC's (AA) which has 48 primer inlets and 48 sample mix inlets. A 20x primer plate was made by diluting 50

µM primer pairs with DNA Suspension Buffer (10 mM TRIS, pH8.0, 0.1 mM EDTA, TEKnova) and 20x Access Array Loading Reagent (Fluidigm) to a final concentration of 6 µM primer pairs and 1x Access Array Loading Reagent. 5 µl Sample Mix was made with a final concentration of 1x KAPA HiFi Hotstart Ready Mix (KAPA Biosystems), 5 % DMSO, 1x Access Array Loading Reagent and 50 ng DNA. **Priming:** 300 µl Control Line Fluid was added to each of the accumulators of the AA and 500 µl 1x Access Array Harvest Reagent (Fluidigm) was added to well H1 to H4. After priming in IFC Controller AX, 4 µl primers were added to the primer inlets and 4 µl Sample Mix was added to the sample inlets. Loading of AA was done, within one hour of priming, in IFC Controller AX, then the AA was transferred to FC1 (Fluidigm) for PCR. PCR-conditions were 50°C for 2 min, 70°C for 20 min, 95°C for 5 min. Then 10 cycles; 98°C for 20 sec, 61°C for 15 sec , 72°C for 15 sec, 2 cycles; 98°C for 20 sec, 80°C for 30 sec, 61°C for 15 sec , 72°C for 15 sec; 8 cycles of 98°C for 20 sec, 61°C for 15 sec , 72°C for 15 sec; 2 cycles of 98°C for 20 sec, 80°C for 30 sec, 61°C for 15 sec , 72°C for 15 sec; 8 cycles of 98°C for 20 sec, 61°C for 15 sec , 72°C for 15 sec; 5 cycles of 98°C for 20 sec, 80°C for 30 sec, 61°C for 15 sec, 72°C for 15 sec and a final cycle of 1 min at 72°C. **Harvest:** After PCR, reagents in well H1 to H4 were replaced with 600 µl 1x Access Array Harvest Reagent and sample inlets were loaded with 2 µl 1x Access Array Harvest Reagent. PCR-products were harvested in IFC Controller AX for 1 hour. Then 10 µl products were transferred from sample inlets to 6 columns of a 96 well-plate and diluted 1:100 in DEPC-treated water. **Barcoding:** Barcodes and sequence adapters were added to 1 µl diluted PCR-product in a new PCR. In a total volume of 10 µl, the final concentration was 1x KAPA HiFi Hotstart Ready Mix, 5 % DMSO and 0.4 µM Access Array Barcode Library (Illumina). PCR conditions were 95°C for 5 min, then 15 cycles of 98°C for 20 sec, 60°C for 15 sec and 72°C for 1 min, and a final hold at 72°C for 3 min. **Library Clean-Up**: 2 µl of each library was pooled in a 1.5 ml tube, 48 µl of the pool and 86.4 µl Solid Phase Reversible Immobilization (SPRI) beads were vortexed for 20 sec and left on magnet for 5 min at RT. The supernatant was discarded and beads were washed with 200 µl 70 % ethanol and left on magnet for 5 min. The ethanol was discarded and the washing was repeated once. Beads were dried at 37°C until the pellet was slightly cracked. 20µl DEPC treated water was applied and the beads were vortexed for 20 seconds and placed on magnet. After 5 min, the supernatant was transferred to a new 1.5 ml tube. **Library quantification:** The sequence library was checked on Agilent Bioanalyzer DNA 1000 Chip and quantified with Real-Time PCR on 7900HT ABI using KAPA Library

Quantification Kit (KAPA Biosystems). **Next generation sequencing:** Samples were sequenced on Illumina MiSeq. 100 bp paired-end (PE) reads were aligned to human genome hg19 with BWA. Integrated genome viewer (IgV) was used to determine allelic ratios.

Results

HCC1954 was found to have 49-fold amplification of *ERBB2*. This is in agreement with Kao et al who detected 45.01-fold with qPCR (64).

SNP frequency

Of the three normal blood samples, no informative SNP's were detected in Hg00096. Hg00097 had only one heterozygous SNP detected; rs4252596 in exon 5 (MAF=0.05) and Hg00098 had 11 informative SNP's. Most informative SNP's were found in HCC1954 which had 12, DET28 and DET59 which had 11 and DET56 which had 10. DET48 had only three heterozygous SNP's, all in exons; rs4252596 (exon5, MAF 0.05), rs4252612 (exon 10, MAF=0.04) and rs1136201 (exon22, MAF=0.16). One patient sample, DET42, had no informative SNP's detected.

Lower limit of detection (LOD)

SNP ratios in tumor samples were calculated by dividing % reads of the most frequent allele by % reads of the other allele. Normal sample allele-ratios were calculated the same way as matched tumor. Figure 32 shows calculated SNP ratios from 100 % tumor (HCC1954) down to zero tumor represented by the cell line's normal match, HCC1954 BL. One SNP, rs1565923, deviates from the rest. This SNP did not have paired-end reads like the rest of the SNP's. The 1 % dilution seems to be an outlier; this might be due to the low volume used when making the dilution curve, only 10 µl in total, which can introduce inaccuracy. Welch two sample t-test showed no significant difference between normal (HCC1954 BL) and 0.1 % dilution ($p=0.17$). The 0.1 % and 0.5 % dilution was significantly different ($p<0.0001$). Separate dilution curves for each SNP are in Appendix C.

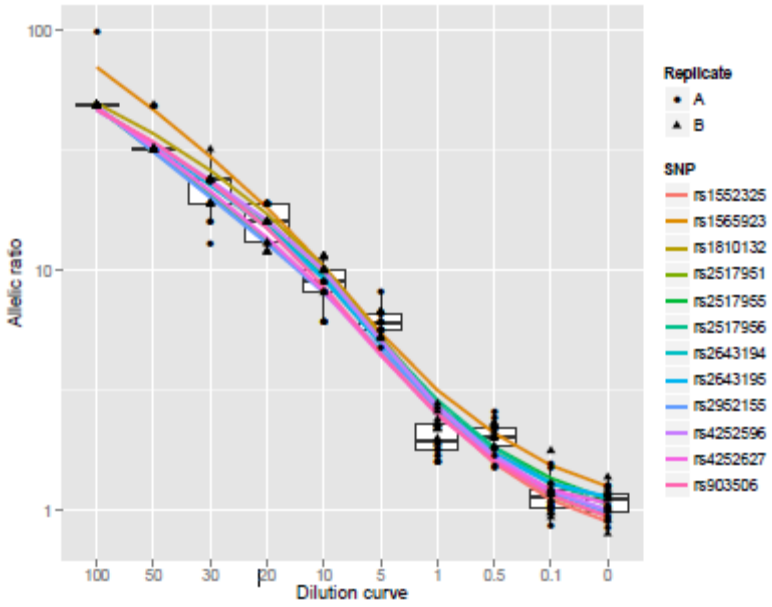

Figure 32. HCC1954 diluted in its normal match, HCC1954BL. 12 SNP rations were calculated per dilution. Undiluted HCC1954 had a 49-fold *ERBB2* amplification. Zero dilution represented the normal match HCC1954BL, which had an expected ratio of 1. All SNP's had paired-end sequencing except rs1565923.

Expected SNP ratios were calculated by summing up the alleles and calculating the representation of each of the two allels in each dilution (table 10).

Table 10. Expected allele ratios at the different dilutions for a 49 fold amplification of *ERBB2*

Expected ratios when ERBB2 is			49	fold amplified		
				A-allele	G-allele	
			Tumor cell	49	1	
Reads	10000		Normal cell	1	1	
					Allelic	Difference in
% tumor	A-allele		G-allele		Ratio	allele counts
100	9800	98 %	200	2 %	49	9600
50	9615	96 %	385	4 %	25	9231
30	9390	94 %	610	6 %	15	8780
20	9138	91 %	862	9 %	11	8276
10	8529	85 %	1471	15 %	5,8	7059
5	7727	77,3 %	2273	22,7 %	3,4	5455
1	5968	59,7 %	4032	40,3 %	1,48	1935
0,5	5536	55,4 %	4464	44,6 %	1,24	1071
0,1	5117	51,2 %	4883	48,8 %	1,05	234
0	5000		5000		1,00	0

Results indicate that the serial dilution curve actually was from 70 % tumor down to 0.14 %, not 50 % to 0.1 % when compared to theoretically calculated ratios (figure 33). This might be caused by inaccuracy when measuring DNA concentrations. HCC1954 was re-measured and gave the same result. There was not enough material to measure HCC1954 BL twice.

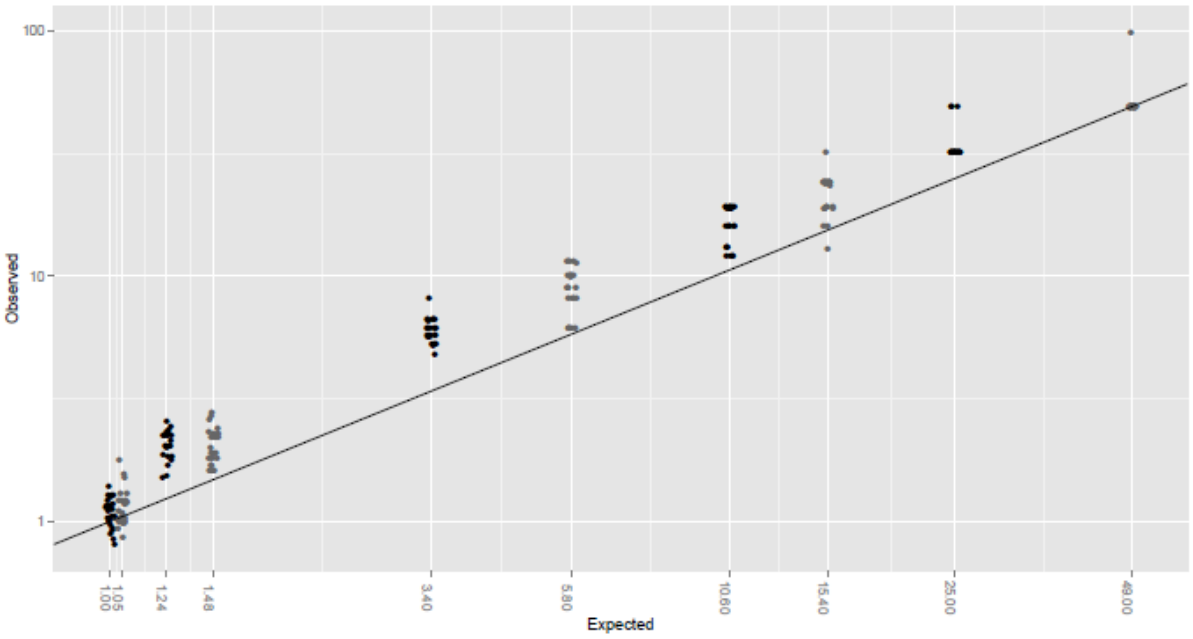

Figure 33. Correlation between expected and observed ratios in the dilution curve. Samples at each end are un-diluted, HCC1954 BL at the left side (1.00) and HCC1954 at the far right side (49 fold). Each sample has 12 points in total, from the 12 measured SNP ratios.

Discussion

Evaluating the effect of delayed sample processing of plasma on ctDNA measurements

Pre-analytical factors are an important source of variation and errors in laboratory measurements. Procedures therefore have to be standardized to prevent bias in the results. There is still no procedure for when to process blood for cell-free DNA analysis, but the general rule is to process blood as soon as possible. The time course study showed that EDTA-blood can be left for 24 hours at room temperature without significantly effecting cfDNA concentration. This is in agreement with Board et al and Kadam et al who investigated the effect of up to 24 hours delayed processing of EDTA-blood from healthy individuals and advanced cancer patients respectively (38, 65). Others have also reported no change in plasma DNA levels in a shorter time frame at room temperature (66, 67) whereas one study reports significant increase in cfDNA after 6 or 24 hours in room temperature. In that study aliquots were taken from the same sample at all the different time points and the samples were held at mild agitation in between aliquots (68).

Clearance of cell-free DNA *in vivo* seems to be rapid. Diehl et al monitored a patient with colorectal cancer after surgery and found ctDNA to have a half-life of 114 minutes (36), and a study which measured fetal DNA in 8 women's blood after giving birth found a mean half-life of 16.3 minutes. They also found that plasma nucleases seem to play a limited effect on degradation of cfDNA (69). Our results cannot draw a conclusion on whether ctDNA is stable *in vitro* yet. There was no dramatic change in ctDNA-fragment counts during delayed processing and even very low counts at baseline could still be detected in the delayed samples. But all samples showed a slight decrease in mutant load when corrected for background. More data will need to be collected to elucidate if this is assay variance or true decrease.

There was a significant increase in cfDNA beyond 24 hours, due to lysis of leukocytes. The ctDNA fraction is calculated relative to the total plasma DNA. Increased background of wild-type DNA will dilute the fraction of ctDNA, producing a false low result. OVO21 had 5.5 % ctDNA at baseline (0 hr), after 48 hours in RT ctDNA was calculated to 0.5 %, but this figure is 4.4 % when adjusting for the elevated background. This might raise the question if the

amount of leukocyte DNA should be measured in the samples (use rearranged immunoglobulin receptors or T-cell receptor genes as marker).

There are sample collection tubes (BCT tubes) on the market containing cell stabilizing substances. The time course study compared five blood samples taken with EDTA- and BCT tubes. Each sample was split in two parts and processed immediately or after one week at RT. There was no difference in yield between the tubes when processed immediately, but after a week in RT, EDTA blood had a median 49.7 fold increase (range 1.5-356) whereas BCT-tubes only had a median 3.9 fold (range 0.6-12). The sample number was too low to give statistically significant results, but there was a trend for less cell lysis in BCT compared to EDTA (p=0.06). A study by Hidestrand et al comparing EDTA- and BCT blood showed that BCT tubes shipped at RT and processed after three days had the same cfDNA concentrations as BCT- and EDTA blood processed within 2 hours, while EDTA blood had significant elevated cfDNA in the same period. They also showed that shipping BCT tubes at a temperature below RT significantly *increased* cfDNA concentration, showing that temperature is an important pre-analytical factor to consider (70). When it comes to EDTA blood, our study indicated less lysis at 4 °C. Another study comparing EDTA- and BCT tubes found significantly increased cfDNA in EDTA blood after 24 hours at RT, and no difference in BCT-tubes for up to 14 days at RT. Their study took aliquots from the same tube after 3 hours, 24 hours etc. Repeated mixing of the same sample before each aliquot might have introduced a bias in the results and influenced EDTA-blood more than BCT, because leukocytes in EDTA tubes are not stabilized and more prone to lysis (71).

Eight patients in the time course study had ctDNA quantified with dPCR. The estimated number of mutant targets was low, ranging from 0-41 fragments corresponding to a mutant fraction of 0-18 %. Fluidigm recommend between 200-700 positive wells per panel to achieve maximum accuracy (45) but less than 200 copies per panel have also been shown to give good precision and accuracy (72, 73). Our results show good agreement with targeted sequencing which supports dPCR in being accurate despite low counts. Case DET56V03 had 8.5 % mutant *TP53* detected in the baseline sample (processed within one hour), targeted MPS detected 11.9 % in a sample replicate from the same time point. DET65V03 and DET15V05 had low mutant burden detected by dPCR; between 0-1.33 % and 0-0.4 % when looking at all time-points. Targeted MPS gave 0.5 % and 0.8 % respectively for a sample

replicate at time point 0 hr. DET48V02 and DET42V05 had 2.6 % and 1.6 % mutant *TP53* detected by targeted MPS, dPCR detected 3.2% and 3.3 % in a sample replicate at the same time point. DET 42V05 is the only sample that gives some kind of different result. A second look at the assay could be worthwhile, but was not done in this thesis. Dawson et al compared dPCR and targeted MPS for seven mutations across 14 follow-up samples for two cases and showed that the methods correlated significantly (Spearman correlation, $r=0.878$, $p<0.0001$) (35)

Fluidigm also recommend running five replicate panels per sample for accurate quantification. It is especially important to rerun low counts to be more certain they are true positives. Most of the time course samples had two replicates, except DET65V03 that had four replicates. The baseline sample, 0 hr, had 0.08 % mutant calculated; of 2485 fragments in total, two fragments were detected as mutant. To increase confidence, results are always compared with the PPC, which should give signal only with the wild type assay. Raw data (heat maps) can also help giving confidence. DNA is denatured before dPCR to make single stranded fragments, these will be randomly distributed into the panel of wells and some wells will only contain one fragment. The appearance of single mutant signals (not sharing well with WT) strengthens the likelihood that it is a true positive.

Digital PCR is useful in rare allele detection because the dilution and partitioning of samples reduce background signal and increase signal-to-noise ratio of low abundance targets compared to qPCR. Careful optimization is although needed, but dilutions of the template helps when threshold is difficult to set. Low amounts of template is needed, so dPCR is also valuable when sample concentration is low and sample volume is limited. The possibility to detect rare variants correlates with the number of replicates, and high numbers of aliquots will both increase the lower limit of detection and the accuracy of the analysis (74). Different factors can influence PCR efficiency such as inhibitors, suboptimal primers and secondary DNA structure. Bhat et al suggests that not only the DNA conformation, but also the fragment length might influence PCR in the initial cycles (75). Whale et al estimated the prevalence of molecular dropout, i.e. when targets fail to amplify, by analyzing targets on the same molecule in duplex PCR. They analyzed both genomic DNA (gDNA) and plasmids and found that molecular dropout was significantly higher in gDNA compared to linearized plasmid. DNA integrity (Bioanalyzer) was found to be good, but it was hypothesized that the

dropout might have been caused by PCR inhibitors, tertiary structure or nicks in the DNA (72). These limiting factors have less impact on dPCR than qPCR because there are no comparisons to standard curves.

Allelic discrimination assays

Allelic discrimination assays can have a certain degree of non-specific amplification because the allele-specific primer or probe has to discriminate between sequences that only differ by one base. Allelic probes should have a melting temperature 5-7 ° C above the annealing temperature to be certain that the probes have bound to the template before elongation starts. At the same time; a single-base mismatch has to change the melting temperature of the probe so that it falls below the annealing temperature for it to be specific. To achieve this, the probe must be short, so that the mismatch will have biggest impact. Both MGB-probes and LNA-probes can be made short due to their enhanced binding capacity to the template. To design the most optimal probe, ΔTm mismatch was calculated with OligoAnalyzer (IDT) and the best discriminating probe was chosen.

The previous assay on PIK3CA p.E545K had a high degree of background amplification. Both probes had a calculated ΔTm of 5.6 °C, and the mutant probe would have been expected to be specific at standard annealing temperature but the assay had a run temperature of 56 °C. The wild type probe's Tm was too high and therefore not expected to be specific. New primers and LNA-probes were designed. Mutant probe had Tm 65 °C and ΔTm 7.4 °C with mismatch, which means that the melting temperature would fall below 60 °C which is the annealing and elongation temperature during PCR (standard run temperature). The probe was specific, as anticipated. Three WT probes were designed. WT1 probe had a calculated mismatch Tm of 61.4 °C, so it could be expected to have a certain degree of non-specificity, but it turned out to bind stronger than expected, detecting all fragments. The OligoAnalyzer mismatch tool cannot take into account the effect of LNA-bases, so the ΔTm calculation is only approximate for LNA probes. The WT probe has a G·T mismatch which is the hardest single-base mismatch to discriminate and adding flanking LNA-bases could even decrease ΔTm further (2). These two factors might have given a probe which was more stable than OligoAnalyzer predicted, but the triplet of LNA bases around the G·T was still designed because the flanking bases was T and A. Adding LNA to flanking bases T and A might increase

discrimination according to Owczarcy's paper. WT2 probe was similar to WT1, only shorter. By lowering the Tm, the probe got a mismatch Tm below run temperature which helped discrimination. With WT3 probe the only focus was to make a short probe and avoid adding LNA at the mismatch site due to the G·T mismatch. But this probe did not work very well. The threshold between signal and background was difficult to set probably due to non-specific binding.

There was still background amplification with WT2 probe though. The primers had passed the specificity check with UCSC In-Silico PCR. A search with primer-BLAST (NCBI) with the new primers discovered that PIK3CA (on chromosome 3) has high homology with chromosome 22 (chr22) in this region. A search with the previous primer set did not give any indications of non-specificity with primer-BLAST either. This emphasized the need of checking the whole amplicon for homology in BLAST before setting up an assay. A closer look at chr22 showed that the mutant probe had two mismatches with chr.22, while WT probe had only one mismatch. An A·C mismatch is believed to discriminate well, but the probe has LNA bases aside of this that might stabilize more. The reverse primer had two mismatches positioned in the middle, slightly towards the 3' end. There was no possibility to make a forward primer specific for *PIK3CA* and the question was whether a reverse primer with 3'end exactly at the mismatch would prevent unspecific amplification or if this is due to linear amplification between forward primer and probe. A new set of reverse primers will be tested in the future. In case of linear amplification one could test if a slight increase in run temperature would prevent the WT probe from binding to chr22 or look at the possibility of designing it (both probes would then have to be moved) on the reverse strand so that elongation of the reverse primer will lead to fragmentation of the probe and not the forward primer.

Using OligoAnalyzer for ΔTm calculations and guide the choice of probe seems to work well. It seems from these experiments however, that LNA probes give a higher signal-to-noise ratio than MGB probes. A high signal-to-noise ratio is important to get sufficiently above the nonspecific signal due to spectral bleed-through. It also shows that the amplification curves differ between the two probes within an assay. For assay TP53 p.Y220C this might be due to the fact that the LNA-probe had a higher signal-to-noise ratio than the MGB-probe. Assay PIK3CA p.E545K had two LNA-probes with the same Tm, still wild type amplification had

better amplification curves than mutant probe. Both use the same primer pair, so it's not connected to amplification efficiency. And since both probes have the same Tm they should bind simultaneously to the template. Primers and probes were controlled for secondary structures, and even though no serious secondary binding was discovered, this might introduce differences. Non-specific binding without non-specific amplification, might also be a factor.

Analysis of ctDNA by MPS to monitor metastatic breast cancer

Metastatic breast cancer remains an incurable disease but it is treatable by means of endocrine, cytotoxic or biologic therapies. The monitoring of treatment response is important to avoid continuing ineffective therapies, to prevent unnecessary side effects and to determine the benefit of new therapeutics (35). Given the evidence of intratumor heterogeneity, more attention needs to be paid to the clonal evolution of tumors over the course of the disease and during and following drug therapy (16). Targeted MPS enables simultaneous screening of many potentially mutated genes with low input of DNA. The preliminary results from our study showed that by screening 17 selected genes for somatic point mutations, SNV's could be detected in 94 % of the patients (49/54). Half of the patients had SNV's in *TP53* and *PIK3CA* which are considered driver mutations, so tracking these would probably reflect the dominant clone. But cases have been reported where *PIK3CA* mutations have been lost during the metastatic progress (76). So due to the dynamics of clonal diversification and selection, several tumor markers should be used to monitor the disease. Our study identified two SNV's in 48 % of the patients. Both of the cases presented in this thesis indicated subclonality. Case 1 shows different dynamics of the *PIK3CA* mutation and *CDH1* mutation during therapy, and case 18 had only *TP53* detected at presentation, whereas *KRAS* was manifested at a later time point.

One important subclass of a driver is a mutation that confers resistance to therapy. These are typically found in recurrences of cancers that initially respond to treatment but later become resistant. Resistance mutations often confer limited growth advantages in the absence of therapy. Some seem to predate initiation of therapy, existing as passengers in minor subclones until the selective environment is changed by initiation of therapy. The passenger is then converted to a driver and preferentially expanded (19). Serial sampling of

the tumor genome during treatment might help identifying these drivers and eventually improve and lead to more personalized therapy. Murtaza et al identified somatic mutations associated with therapy resistance in a small cohort of advanced cancer patients by whole exome sequencing (77). Our study might also give some contribution to this field.

Detection of tumor specific amplification in blood by targeted MPS

HER2 over-expression can be detected by various methods like immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH). To detect more subtle fold changes, qPCR and dPCR have been used (62, 63, 72). We diluted the *ERBB2* amplified cell line HCC1954 into its normal match HCC1954BL from 50 % tumor down to 0.1 % to elucidate the sensitivity of MPS to detect the amplification. *ERBB2* amplification is a monoallelic event, which creates an allelic imbalance that can be measured with heterozygous SNP's. Our results indicate that MPS is as sensitive as dPCR but not superior, and can detect 1.2 fold changes in copy number.

Both dPCR and qPCR need the use of a non-amplified reference gene to measure *ERBB2* copy number change. Breast cancers have instable genomes and a high degree of heterogeneity, so choosing the right reference gene might be challenging (78). By using MPS and SNP frequencies, there is no need of a reference gene as the unamplified allele is the reference. As both alleles use the same primer set, there is neither any bias in PCR efficiency which might influence qPCR.

Two of nine samples in this study had no heterozygous SNP's identified among 12 frequent and seven less frequent heterozygous SNP sites. One sample had only one informative SNP and another sample had three heterozygous SNP's. The remaining five samples had 11 or 12 informative SNP sites. This implies the need to include more SNP's to detect the amplification. Of those with few heterozygous SNP's only low frequency SNP's were detected, so it could be worthwhile including more of the low frequency SNP sites in the assay. This can easily be done technically as 48.48 Access Arrays have the capacity to run 48 singelplex PCR's in parallel or up to 10x48 multiplex PCR's. This will however challenge the primer design. *ERBB2* is located on chromosome 17q12 and several 17q12-q21 genes are

variably co-amplified with *ERBB2*. Sircoulomb et al used microarrays to detect gene amplifications in 54 *ERBB2* amplified breast cancers and found *ERBB2*, *MIEN1* and *GRB7* to be the common core of the amplicons. The most common telomeric border of the amplicons was *IKZF3* and the most frequent centromeric border gene was *CRKRS* (21). Another study, comprising over 200 breast cancers identified parts of *PGAP3*, the whole of *ERBB2* and *MIEN1* as the common core (58). SNP ratios in these genes could be used in combination with SNP ratios in *ERBB2* to strengthen the confidence in the results.

A statistical analysis called sequential probability ratio test (SPRT) has been used to detect allelic imbalance with dPCR (79, 80). The method allows two hypothesis to be compared as data accumulates and this could give an indication on how many SNP ratios that will have to be calculated to detect that the ratio is different from normal samples. This was not done in this thesis, but could be interesting to do in the future for further evaluating of the assay.

Another question that emerges is if allelic imbalance could be detected with SNP's and dPCR for follow-up samples? Once the heterozygous SNP's have been identified in the primary tumor, selected SNP's could be run with dPCR. This would not necessarily imply the need of many assays. In our study we could calculate SNP ratios for 10-11 of the 12 most frequent SNP's in 56 % of the samples (5/9). Only two samples had no SNP marker identified.

Using targeted MPS to measure allelic imbalance does not tell if a gene is amplified or lost. Loss of heterozygocity (LOH) will create the same allelic imbalance as amplification. LOH in *ERBB2* is however a rare event, but it has been detected in the cell line MCF-7 (61)

Comparisons of the two methods

Both dPCR and targeted MPS are sensitive, quantitative methods. Limits of detection (LOD) down to 2 % and 0.14 % have been reported with targeted MPS (35, 56). Our study detecting *ERBB2* amplification had 0.5 % LOD. The sensitivity of dPCR can go beyond MPS, and it is connected to the number of aliquots analyzed. Increased number of partitions will increase LOD.

Digital PCR can only analyze a few different mutations at a time. Fluidigm can multiplex up to four assays. The method is also best for detecting known mutations. With MPS hundreds of

genes can be screened for mutations simultaneously with the potential of detecting evolved mutations in follow-up samples.

MPS needs the use of bioinformatics skills to map and align the sequences, while dPCR can be done without extra assistance. Digital PCR is also quicker to perform but it is not a high-throughput system like MPS can be.

Conclusion

Analysis of ctDNA is a valuable tool for monitoring treatment response and disease progression. A liquid biopsy is more accessible than tissue biopsies, and it is likely to contain a wider representation of the tumor genomes. Cell-free DNA from multiple lesions all mix together in the peripheral blood, whereas a single biopsy might miss a minor subclone or not be representative for all lesions due to clonal evolution (77). Care should be taken in the pre-analytical process because lysis of leukocytes will increase the background of wild type DNA in plasma and influence the quantitative measurements. This study showed that EDTA-blood can be stored for 24 hours before processing plasma without significantly affecting the total plasma DNA concentration. Whether ctDNA can be degraded *in vitro* is still unclear and more data needs to be collected. Targeted MPS can detect 1.2 fold amplifications of *ERBB2*, which is the same limit of detection that can be obtained by dPCR. In contrast to dPCR, MPS does not need a reference gene to detect copy number change, but uses the unamplified allele as the reference. Serial monitoring for a broad range of mutations by MPS gives a rapid identification of treatment response and relapse, and is a step towards personalized medicine.

References

1. Kreftregisteret. Institutt for populasjonsbasert kreftforskning [database on the Internet]. Giske Ursin. Available from: http://www.kreftregisteret.no.
2. Owczarzy R, You Y, Groth CL, Tataurov AV. Stability and mismatch discrimination of locked nucleic acid-DNA duplexes. Biochemistry (Mosc). 2011;50(43):9352-67. Epub 2011/09/21.
3. World Health Organization. Breast Cancer: Prevention and Control. World Health Organization; 2013 [cited 2013 16.05.13]; Available from: http://www.who.int/cancer/detection/breastcancer/en/index.html.
4. Cancer Research UK; 2013 [updated 13.02.13; cited 2013 15.02.13]; Available from: http://www.cancerresearchuk.org/cancer-info/cancerstats/types/breast/.
5. Norsk bryst cancer gruppe. [cited 2013 16.06.13]; Available from: http://www.nbcg.no/nbcg.blaaboka.html#Anchor-44867.
6. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486(7403):346-52. Epub 2012/04/24.
7. Dawson SJ, Rueda OM, Aparicio S, Caldas C. A new genome-driven integrated classification of breast cancer and its implications. EMBO J. 2013;32(5):617-28. Epub 2013/02/12.
8. Li J, Wang K, Jensen TD, Li S, Bolund L, Wiuf C. Tumor heterogeneity in neoplasms of breast, colon, and skin. BMC Res Notes. 2010;3:321.
9. Brosnan JA, Iacobuzio-Donahue CA. A new branch on the tree: next-generation sequencing in the study of cancer evolution. Semin Cell Dev Biol. 2012;23(2):237-42. Epub 2012/01/17.
10. Mardis ER. Genome sequencing and cancer. Curr Opin Genet Dev. 2012;22(3):245-50. Epub 2012/04/27.
11. Horswell S, Matthews N, Swanton C. Cancer heterogeneity and "The Struggle for Existence": Diagnostic and analytical challenges. Cancer Lett. 2012. Epub 2012/11/13.
12. 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(7403):395-9. Epub 2012/04/13.
13. Nik-Zainal S, Van Loo P, Wedge DC, Alexandrov LB, Greenman CD, Lau KW, et al. The life history of 21 breast cancers. Cell. 2012;149(5):994-1007. Epub 2012/05/23.
14. Swanton C. Intratumor heterogeneity: evolution through space and time. Cancer Res. 2012;72(19):4875-82. Epub 2012/09/25.
15. Navin N, Krasnitz A, Rodgers L, Cook K, Meth J, Kendall J, et al. Inferring tumor progression from genomic heterogeneity. Genome Res. 2010;20(1):68-80. Epub 2009/11/12.
16. Yap TA, Gerlinger M, Futreal PA, Pusztai L, Swanton C. Intratumor heterogeneity: seeing the wood for the trees. Science translational medicine. 2012;4(127):127ps10. Epub 2012/03/31.
17. Seol H, Lee HJ, Choi Y, Lee HE, Kim YJ, Kim JH, et al. Intratumoral heterogeneity of HER2 gene amplification in breast cancer: its clinicopathological significance. Mod Pathol. 2012;25(7):938-48. Epub 2012/03/06.
18. Caldas C. Cancer sequencing unravels clonal evolution. Nat Biotechnol. 2012;30(5):408-10. Epub 2012/05/09.
19. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458(7239):719-24. Epub 2009/04/11.
20. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012;486(7403):400-4. Epub 2012/06/23.
21. Sircoulomb F, Bekhouche I, Finetti P, Adelaide J, Ben Hamida A, Bonansea J, et al. Genome profiling of ERBB2-amplified breast cancers. BMC cancer. 2010;10:539. Epub 2010/10/12.
22. Rosenberg CL. Polysomy 17 and HER-2 amplification: true, true, and unrelated. J Clin Oncol. 2008;26(30):4856-8. Epub 2008/09/17.

23. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nature reviews Genetics. 2010;11(10):685-96. Epub 2010/09/18.
24. Pantel K, Alix-Panabieres C. Circulating tumour cells in cancer patients: challenges and perspectives. Trends in molecular medicine. 2010;16(9):398-406. Epub 2010/07/30.
25. Andreopoulou E, Cristofanilli M. Circulating tumor cells as prognostic marker in metastatic breast cancer. Expert review of anticancer therapy. 2010;10(2):171-7. Epub 2010/02/06.
26. Cabinakova M, Tesarova P. Disseminated and circulating tumour cells and their role in breast cancer. Folia Biol (Praha). 2012;58(3):87-97. Epub 2012/08/02.
27. Muller V, Alix-Panabieres C, Pantel K. Insights into minimal residual disease in cancer patients: implications for anti-cancer therapies. Eur J Cancer. 2010;46(7):1189-97. Epub 2010/03/30.
28. Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. Nature reviews Cancer. 2011;11(6):426-37. Epub 2011/05/13.
29. Zhong XY, Ladewig A, Schmid S, Wight E, Hahn S, Holzgreve W. Elevated level of cell-free plasma DNA is associated with breast cancer. Arch Gynecol Obstet. 2007;276(4):327-31. Epub 2007/04/14.
30. Holdenrieder S, Nagel D, Schalhorn A, Heinemann V, Wilkowski R, von Pawel J, et al. Clinical relevance of circulating nucleosomes in cancer. Ann N Y Acad Sci. 2008;1137:180-9. Epub 2008/10/08.
31. Thierry AR, Mouliere F, Gongora C, Ollier J, Robert B, Ychou M, et al. Origin and quantification of circulating DNA in mice with human colorectal cancer xenografts. Nucleic Acids Res. 2010;38(18):6159-75. Epub 2010/05/25.
32. Gormally E, Caboux E, Vineis P, Hainaut P. Circulating free DNA in plasma or serum as biomarker of carcinogenesis: practical aspects and biological significance. Mutat Res. 2007;635(2-3):105-17. Epub 2007/01/30.
33. Mouliere F, Thierry AR. The importance of examining the proportion of circulating DNA originating from tumor, microenvironment and normal cells in colorectal cancer patients. Expert opinion on biological therapy. 2012;12 Suppl 1:S209-15. Epub 2012/05/19.
34. Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer Res. 2001;61(4):1659-65. Epub 2001/03/14.
35. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, et al. Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer. N Engl J Med. 2013. Epub 2013/03/15.
36. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med. 2008;14(9):985-90. Epub 2008/08/02.
37. Leary RJ, Kinde I, Diehl F, Schmidt K, Clouser C, Duncan C, et al. Development of personalized tumor biomarkers using massively parallel sequencing. Science translational medicine. 2010;2(20):20ra14. Epub 2010/04/08.
38. Board RE, Williams VS, Knight L, Shaw J, Greystoke A, Ranson M, et al. Isolation and extraction of circulating tumor DNA from patients with small cell lung cancer. Ann N Y Acad Sci. 2008;1137:98-107. Epub 2008/10/08.
39. Fleischhacker M, Schmidt B, Weickmann S, Fersching DM, Leszinski GS, Siegele B, et al. Methods for isolation of cell-free plasma DNA strongly affect DNA yield. Clin Chim Acta. 2011;412(23-24):2085-8. Epub 2011/08/25.
40. Streck. Cell-free DNA BCT. 2011.
41. QIAGEN. QIAamp® Circulating Nucleic Acid Handbook. www.qiagen.comMay 2009.
42. Sykes PJ, Neoh SH, Brisco MJ, Hughes E, Condon J, Morley AA. Quantitation of targets for PCR by use of limiting dilution. Biotechniques. 1992;13(3):444-9. Epub 1992/09/01.
43. Baker M. Digital PCR hits its stride. Nat Meth. 2012;9(6):541-4.
44. Sedlak RH, Jerome KR.
45. Fluidigm. Digital PCR Analysis Software Version 3,0.
46. Biosystems A. Primer Express Version 1.5 and TaqMan MGB Probes for Allelic Discrimination. May 26 2000.

47. You Y, Moreira BG, Behlke MA, Owczarzy R. Design of LNA probes that improve mismatch discrimination. Nucleic Acids Res. 2006;34(8):e60. Epub 2006/05/04.
48. Eurogentec. Doube-Dye LNA Probes. 2013 [07.05.2013]; Available from: http://www.eurogentec.com/products/doube-dye-lna-probes.html.
49. Sigma-Aldrich. Locked Nucleic Acids® (LNA®) FAQ. 2011.
50. Integrated DNA Technologies I. OligoAnalyzer 3.1. Integrated DNA Technologies, Inc 2013; Available from: https://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/.
51. Exiqon. LNA™ Oligo Tools and Design Guidelines [cited 2013 17.05.13]; Available from: http://www.exiqon.com/oligo-tools.
52. Wellcome Trust Sanger Institute GRL. COSMIC, Catalogue of somatic mutations in cancer. [cited 2013 15.05.13]; Available from: http://cancer.sanger.ac.uk/cosmic/sample/overview?id=907045.
53. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61-70. Epub 2012/09/25.
54. Corporation RASaF. Targeted resequencing of the EGFR and MET genes using the Fluidigm Access Array System and the Roche GS Junior System BioTechniques2011.
55. Fluidigm. Access Array System for Illumina Sequencing Systems.
56. 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. Science translational medicine. 2012;4(136):136ra68. Epub 2012/06/01.
57. Affymetrics. ExoSAP-IT. PCR cleanup reagent. The gold standard. 2012.
58. Staaf J, Jonsson G, Ringner M, Baldetorp B, Borg A. Landscape of somatic allelic imbalances and copy number alterations in HER2-amplified breast cancer. Breast cancer research : BCR. 2011;13(6):R129. Epub 2011/12/16.
59. Krawczyk N, Banys M, Neubauer H, Solomayer EF, Gall C, Hahn M, et al. HER2 status on persistent disseminated tumor cells after adjuvant therapy may differ from initial HER2 status on primary tumor. Anticancer Res. 2009;29(10):4019-24. Epub 2009/10/23.
60. Wulfing P, Borchard J, Buerger H, Heidl S, Zanker KS, Kiesel L, et al. HER2-positive circulating tumor cells indicate poor clinical outcome in stage I to III breast cancer patients. Clin Cancer Res. 2006;12(6):1715-20. Epub 2006/03/23.
61. Whale AS, Huggett JF, Cowen S, Speirs V, Shaw J, Ellison S, et al. Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation. Nucleic Acids Res. 2012;40(11):e82. Epub 2012/03/01.
62. Page K, Hava N, Ward B, Brown J, Guttery DS, Ruangpratheep C, et al. Detection of HER2 amplification in circulating free DNA in patients with breast cancer. Br J Cancer. 2011;104(8):1342-8. Epub 2011/03/24.
63. Bechmann T, Andersen RF, Pallisgaard N, Madsen JS, Maae E, Jakobsen EH, et al. Plasma HER2 amplification in cell-free DNA during neoadjuvant chemotherapy in breast cancer. J Cancer Res Clin Oncol. 2013. Epub 2013/03/13.
64. Kao J, Salari K, Bocanegra M, Choi YL, Girard L, Gandhi J, et al. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. PloS one. 2009;4(7):e6146. Epub 2009/07/08.
65. Kadam SK, Farmen M, Brandt JT. Quantitative measurement of cell-free plasma DNA and applications for detecting tumor genetic variation and promoter methylation in a clinical setting. The Journal of molecular diagnostics : JMD. 2012;14(4):346-56. Epub 2012/05/15.
66. Jung M Fau - Klotzek S, Klotzek S Fau - Lewandowski M, Lewandowski M Fau - Fleischhacker M, Fleischhacker M Fau - Jung K, Jung K. Changes in concentration of DNA in serum and plasma during storage of blood samples. (0009-9147 (Print)).
67. Lui Yy Fau - Chik K-W, Chik Kw Fau - Chiu RWK, Chiu Rw Fau - Ho C-Y, Ho Cy Fau - Lam CWK, Lam Cw Fau - Lo YMD, Lo YM. Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. (0009-9147 (Print)).

68. Angert RM, LeShane ES, Lo YM, Chan LY, Delli-Bovi LC, Bianchi DW. Fetal cell-free plasma DNA concentrations in maternal blood are stable 24 hours after collection: analysis of first- and third-trimester samples. Clin Chem. 2003;49(1):195-8. Epub 2003/01/01.
69. 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(1):218-24. Epub 1999/01/23.
70. Hidestrand M, Stokowski R, Song K, Oliphant A, Deavers J, Goetsch M, et al. Influence of temperature during transportation on cell-free DNA analysis. Fetal Diagn Ther. 2012;31(2):122-8. Epub 2012/01/21.
71. Fernando MR, Chen K, Norton S, Krzyzanowski G, Bourne D, Hunsley B, et al. A new methodology to preserve the original proportion and integrity of cell-free fetal DNA in maternal plasma during sample processing and storage. Prenat Diagn. 2010;30(5):418-24. Epub 2010/03/23.
72. Whale AS, Cowen S, Foy CA, Huggett JF. Methods for applying accurate digital PCR analysis on low copy DNA samples. PloS one. 2013;8(3):e58177. Epub 2013/03/09.
73. Sanders R, Huggett JF, Bushell CA, Cowen S, Scott DJ, Foy CA. Evaluation of digital PCR for absolute DNA quantification. Anal Chem. 2011;83(17):6474-84. Epub 2011/03/31.
74. Day E, Dear PH, McCaughan F. Digital PCR strategies in the development and analysis of molecular biomarkers for personalized medicine. Methods. 2012. Epub 2012/08/29.
75. Bhat S, Herrmann J, Armishaw P, Corbisier P, Emslie KR. Single molecule detection in nanofluidic digital array enables accurate measurement of DNA copy number. Analytical and bioanalytical chemistry. 2009;394(2):457-67. Epub 2009/03/17.
76. Dupont Jensen J, Laenkholm AV, Knoop A, Ewertz M, Bandaru R, Liu W, et al. PIK3CA mutations may be discordant between primary and corresponding metastatic disease in breast cancer. Clin Cancer Res. 2011;17(4):667-77. Epub 2010/10/14.
77. Murtaza M, Dawson SJ, Tsui DW, Gale D, Forshew T, Piskorz AM, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature. 2013. Epub 2013/04/09.
78. Gevensleben H, Garcia-Murillas I, Graeser MK, Schiavon G, Osin P, Parton M, et al. Non-invasive detection of HER2 amplification with plasma DNA Digital PCR. Clin Cancer Res. 2013. Epub 2013/05/03.
79. Lo YM, Lun FM, Chan KC, Tsui NB, Chong KC, Lau TK, et al. Digital PCR for the molecular detection of fetal chromosomal aneuploidy. Proc Natl Acad Sci U S A. 2007;104(32):13116-21. Epub 2007/08/01.
80. Zhou W, Galizia G, Goodman SN, Romans KE, Kinzler KW, Vogelstein B, et al. Counting alleles reveals a connection between chromosome 18q loss and vascular invasion. Nat Biotech. 2001;19(1):78-81.

Appendix A

Table1. Primers and probes for the total DNA assay

Target Gene	Primer/probe sequence 5'→3'	Amplicon bp	Annealing °C /PCR Cycles
RPP30	AGATTTGGACCTGCGAGCG GAGCGGCTGTCTCCACAAGT ROX-TTCTGACCTGAAGGCTCTGCGCG-BHQ2	65	60/55
XenT	GTGATCATGGGATTTGTAGCTGTT AAACCAACCTGAAAACCATGGA 6FAM-CCCATGGATTATCG-MGB	67	

Table 2. Allelic discrimination assays. Primer and probe sequences are in 5'-3' direction. All wild type probes are labeled with VIC and mutant probes are labeled with 6FAM or ROX. Mismatch sites are indicated in red. [+base] indicates the position of an LNA-base.

Case no.	Gene	Mutation	Primers and probes (Tm) 5'→3'	Amplicon bp	Annealing °C /PCR Cycles
15	PIK3CA	p.H1047R A>G	F-AAGAGGCTTTGGAGTATTTCATGAA (58.4) R-TGTTTAATTGTGTGGAAGATCCAATC (59.2) VIC-CAAATGAATGATGCACATC-MGB (67,0) 6FAM-TGATGCACGTCATGGT-MGB (67,0)	94	60/60
42	TP53	p.D281G T>C	F-GGTGAGGCTCCCCTTTCT (55,2) R-TTTGAGGTGCGTGTTTGTG (55,2) VIC- CGCCGGTCTCTC-MGB (69) 6FAM-CGCCGGCCTCTC-MGB (71)	82	60/60
48	PIK3CA	p.E545K G>A	F-GCAATTTCTACACGAGATCCTCTCT (58,2) R-CATTTTAGCACTTACCTGTGACTCCAT (59,8) VIC-TGAAATCACTGAGCAGGAG-MGB (69,0) 6FAM-TGAAATCACTAAGCAGGA-MGB (64,0)	83	56/60
56	TP53	p.N239D A>G	F-GCTCTGACTGTACCACCATCCA (58,7) R-CATGCCGCCCATGCA (59,0) VIC-TACATGTGTAACAGTTC-MGB (64) 6FAM-CATGTGTGACAGTTC-MGB (66)	62	60/60
65	TP53	p.Y220C T>C	F-GAGACCCCAGTTGCAAACCA (59,5) R-TGGATGACAGAAACACTTTTCGAC (59,2) VIC-AGGCGGCTCATAG-MGB (66) ROX-CGGCTCA[+C]AGGGCA-BHQ2 (66)	79	60/60

Appendix B

Phenol-chloroform extraction protocol

1. Re-suspend pellet in 1ml lysis buffer + 40µl 10 % SDS + 50µl proteinase K (20mg/ml)

DNA extraction lysis buffer: 10mM Tris pH 7.4, 10mM EDTA, 150mM $NaCl_2$

2. Leave on thermal shaker at 55'C overnight 300-500rpm

3. Add equal volume of phenol-chloroform (PCIA). Mix well. Aliquot into phase-lock tubes (light 2ml tubes, 5PRIME). Centrifuge at max speed for 15 min at 4 °C

4. Transfer aqueous layer into new tube. Add 1/10 volume (ie 100µl) of 3M NaAc pH5.2.

5. Add 2.5X volume (ie 2.5ml) of ice cold 100% ethanol. Mix. Remove pellet.

6. Add 1ml 70 % ethanol. Centrifuge at max speed for 15min (4 °C).

7. Remove ethanol. Centrifuge again and remove any remaining ethanol.

8. Dry pellet for 5-10 min and re-suspend in water (500µl).

9. Add 5µl RNase A (stock 10mg/ml) and 5µl of RNase one. Leave at 37 °C for 1 hour.

10. Add equal volume of PCIA. Mix well. Aliquot into phase lock tubes (light 2ml tubes). Centrifuge at max speed for 15 min at 4 °C.

11. Transfer aqueous layer into new tube. Add 1/10 volume of 3M NaAc pH5.2.

12. Add 2.5X volume of ice cold 100% ethanol. Mix. Remove pellet.

13. Add 1ml 70 % ethanol. Centrifuge at max speed for 15min (4 °C).

14. Remove ethanol. Centrifuge again and remove any remaining ethanol.

15. Dry pellet for 5-10 min and re-suspend in water (500µl).

Primer plate for targeted MPS (screening 17 genes):

A	1	AKT1_Ex_F1	TP53_E00001255919_1	GATA3_Ex4_F1	CDH1_Ex_F1	Multiplex-1
B	1	AKT1_Ex_F2	TP53_E00001255919_3	GATA3_Ex5_F1	CDH1_Ex_F2	Multiplex-2
C	1	AKT2_Ex_F1	TP53_E00001255919_6	GATA3_Ex6_F3	CDH1_Ex_F3	Multiplex-3
D	1	CASP8_Ex_F5	TP53_E00001612188_1	EGFR_E00001779947_1	CDH1_Ex_F8	Multiplex-4
E	1	AR_Ex4_F1	TP53_E00001665758_1	EGFR_E00001779947_2	CDH1_Ex_F9	Multiplex-5
F	1	AR_Ex4_F2	TP53_E00001757276_1	EGFR_E00001801208_1	CDH1_Ex_F10	Multiplex-6
G	1	AR_Ex5_F1	TP53_E00001789298_2	EGFR_E00001801208_2	CDH1_Ex_F21	Multiplex-7
H	1	AR_Ex6_F1	TP53_E00001789298_3	EGFR_E00001601336_2	CDH1_Ex_F22	Multiplex-8
A	2	CDH1_Ex_F12	CASP8_Ex_F7	MAP3k1_Ex_F43	MAP2K4_Ex_F1	Multiplex-9
B	2	CDH1_Ex_F13	CASP8_Ex_F8	MAP3k1_Ex_F44	MAP2K4_Ex_F2	Multiplex-10
C	2	CDH1_Ex_F14	CASP8_Ex_F9	MAP3k1_Ex_F45	MAP2K4_Ex_F5	Multiplex-11
D	2	CDH1_Ex_F15	CASP8_Ex_F10	MAP3K1_Ex_F15	MAP2K4_Ex_F6	Multiplex-12
E	2	CDH1_Ex_F16	CASP8_Ex_F11	MAP3K1_Ex_F21	MAP2K4_Ex_F7	Multiplex-13
F	2	CDH1_Ex_F17	CASP8_Ex_F2	MAP3K1_Ex_F23	MAP2K4_Ex_F10	Multiplex-14
G	2	CDH1_Ex_F18	CASP8_Ex_F3	MAP3k1_Ex_F29	MAP3K1_Ex_F1	Multiplex-15
H	2	CDH1_Ex_F19	CASP8_Ex_F4	MAP3k1_Ex_F30	MAP3K1_Ex_F2	Multiplex-16
A	3	GATA3_Ex2_F4	GATA3_Ex3_F4	CDKN1B_Ex_F3	TP53_E00001255919_5	Multiplex-17
B	3	GATA3_Ex2_F5	GATA3_Ex3_F5	CDH1_Ex_F24	TP53_E00001596491_1	Multiplex-18
C	3	GATA3_Ex3_F1	GATA3_Ex2_F6	CDKN1B_Ex_F1	TP53_E00001612188_2	Multiplex-19
D	3	GATA3_Ex3_F2	GATA3_Ex6_F2	EGFR_E00001601336_1	TP53_E00001728015_1	Multiplex-20
E	3	GATA3_Ex3_F6				Multiplex-21
F	3	MAP3k1_Ex_F39	PTEN_E00001456541_1	MAP3K1_Ex_F10		Multiplex-22
G	3	MAP3k1_Ex_F33	PTEN_E00001456541_2	MAP3k1_Ex_F46		Multiplex-23
H	3	EGFR_E00001730701_1	KRAS_AA_HS			Multiplex-24
A	4	MAP3K1_Ex_F16	CASP8_Ex_F1	MAP2K4_Ex_F3	PTEN_E00001156315_5	Multiplex-25
B	4	MAP3K1_Ex_F17	CASP8_Ex_F6	MAP2K4_Ex_F8	PTEN_E00001156315_7	Multiplex-26
C	4	MAP3K1_Ex_F18	GATA3_Ex6_F1	MAP2K4_Ex_F9	PTEN_E00001156321_4m	Multiplex-27
D	4	MAP3K1_Ex_F19	TBX3_Ex_F2	MAP2K4_Ex_F11	PTEN_E00001156327_1	Multiplex-28
E	4	MAP3K1_Ex_F20	TP53_00001404886_13	MAP3K1_Ex_F4	PTEN_E00001156327_4	Multiplex-29
F	4	MAP3K1_Ex_F22	BRAF_AA_HS	MAP3K1_Ex_F8	PTEN_E00001156330_1	Multiplex-30
G	4	MAP3k1_Ex_F28	PIK3CA_Ex10_gap	MAP3K1_Ex_F12	PTEN_E00001156330_3	Multiplex-31
H	4	MAP3k1_Ex_F32	PTEN_E00001156344_1	MAP3K1_Ex_F13	PTEN_E00001156337_4	Multiplex-32
A	5	CDH1_Ex_F4	CDKN1B_Ex_F5	MAP3K1_Ex_F5	MAP3k1_Ex_F40	Multiplex-33
B	5	CDH1_Ex_F5	AR_Ex3_F1	MAP3K1_Ex_F6	MAP3k1_Ex_F41	Multiplex-34
C	5	CDH1_Ex_F6	AR_Ex7_F1	MAP3K1_Ex_F9	MAP3k1_Ex_F42	Multiplex-35
D	5	CDH1_Ex_F7	AR_Ex8_F1	MAP3K1_Ex_F11	SF3B1_Ex_F1	Multiplex-36
E	5	CDH1_Ex_F11	TP53_E00001789298_1	MAP3k1_Ex_F31	TBX3_Ex_F1	Multiplex-37
F	5	CDH1_Ex_F20	EGFR_Exon19	MAP3k1_Ex_F35	PTEN_E00001456562_1	Multiplex-38
G	5	GATA3_Ex2_F1	PTEN_E00001156321_1	MAP3k1_Ex_F36	EGFR_E00001631695_1	Multiplex-39
H	5	GATA3_Ex2_F2	PTEN_E00001156321_2	MAP3k1_Ex_F47	EGFR_E00001773562_1	Multiplex-40
A	6	EGFR_E00001681524_2	CDKN1B_Ex_F4	EGFR_Ex18	EXP0116_TP53_E3	Multiplex-41
B	6	CDH1_Ex_F23	GATA3_Ex1_F1	MAP3k1_Ex_F37	EXP0116_TP53_E4	Multiplex-42
C	6	CDKN1B_Ex_F2	EGFR_E00001681524_1	MAP3k1_Ex_F38		Multiplex-43
D	6	CDH1_Ex_F25	GATA3_Ex2_F3	MAP2K4_Ex_F12	MAP3k1_Ex_F24	Multiplex-44
E	6	MAP3k1_Ex_F27	GATA3_Ex3_F3	MAP3k1_Ex_F26	MAP3k1_Ex_F25	Multiplex-45
F	6	MAP2K4_Ex_F4	PTEN_E00001156315_1m	MAP3k1_Ex_F34		Multiplex-46
G	6	MAP3K1_Ex_F7	MAP3K1_Ex_F14			Multiplex-47
H	6	MAP3K1_Ex_F3	PTEN_E00001156351_1	PIK3CA_Exn21_gap		Multiplex-48

Purification of Harvested PCR Products

1) Keep the AMPure XP magnetic beads in RT for 30 min before use
2) Pipette 2 µl from each of the harvested PCR product pools into a 1.5 ml tube to create a PCR product library. Vortex for 10 sec
3) Transfer 48 µl PCR product library to a new 1.5 ml tube. Vortex the AMPure XP magnetic beads and add 86.4 µl of the beads to the library. Vortex for 20 sec, spin and place on magnet for 5-10 min to separate the beads from the solution
4) Aspirate the clear solution and discard
5) Dispense 200 µl 70 % ethanol and incubate for 5 min4
6) Aspirate the ethanol and discard without disturbing the beads
7) Repeat step 5 and 6
8) Place the tube at 37 °C with open lid until beads are slightly cracked
9) Add 20 µl DEPC free water and vortex for 30 sec
10) Place the bead mix solution on magnet for 5-10 min to separate beads from solution
11) Carefully transfer the clear solution to a new 1.5 ml tube
12) Store purified library at -20 °C until sequencing

Appendix C

ERBB2 primers:

Primer Name	CS1 and CS2 Adaptors	Primer Sequence	Length	Tm	Amplicon bp*	GC content %	Tm	SNP's	MAF
HER2-5'UTR-1F	ACACTGACGACATGGTTCTACA	TTGCCCACTATGGTCCAAAT	20	60,2					
HER2-5'UTR-1R	TACGGTAGCAGAGACTTGGTCT	GTCAGTTCCGGCTTTGAAGA	20	60,4	104	50	79	rs2517955	0.483
HER2-5'UTR-2F	ACACTGACGACATGGTTCTACA	GGGACCCTGGTCTGATTTCT	20	60,3					
HER2-5'UTR-2R	TACGGTAGCAGAGACTTGGTCT	AGCCTCTCTTACTTTCCCAGAAG	23	59,6	118	41	76	rs2517956	0,425
HER2-Ex1-F	ACACTGACGACATGGTTCTACA	GTACGCAGTCGCGGTACAC	19	60,4					
HER2-Ex1-R	TACGGTAGCAGAGACTTGGTCT	GAGCCAGGGAGAAAGGATG	19	59,8	116	73	89		
HER2-Ex2-F	ACACTGACGACATGGTTCTACA	CTGGCCTGGTCTGAATCACT	20	60,3					
HER2-Ex2-R	TACGGTAGCAGAGACTTGGTCT	TGGTGAACAGGACAGCAAAG	20	59,9	116	51	80		
HER2-Ex4-1F	ACACTGACGACATGGTTCTACA	CAGGAGAAGGATGGCTCAGT	20	59,4					
HER2-Ex4-1R	TACGGTAGCAGAGACTTGGTCT	TGGACAGAATGGTTCCTGGT	20	60,4	110	50	79		
HER2-Ex4-2F	ACACTGACGACATGGTTCTACA	TTCAAGACCACATGCAAAGC	20	59,9					
HER2-Ex4-2R	TACGGTAGCAGAGACTTGGTCT	TTTGTCCTGGCCATTCTACTG	21	60,1	80	49	76		
HER2-Intron4-F	ACACTGACGACATGGTTCTACA	ATGGCGTCCACAGTAGCTTT	20	59,8				rs2643194 rs2517951	0,386 0,488
HER2-Intron4-R	TACGGTAGCAGAGACTTGGTCT	CCATCGGGATGTTAGGATCA	20	60,7	115	57	83	rs2643195	0,463
HER2-Ex5-F	ACACTGACGACATGGTTCTACA	TGTGACTGTCTCCTCCCAAA	20	59,2					
HER2-Ex5-R	TACGGTAGCAGAGACTTGGTCT	TAGTCAAGGAGGGGGATGTG	20	59,9	133	47	79	rs4252596	0,051
HER2-Intron6-F	ACACTGACGACATGGTTCTACA	CCTCCAGCACCCAGATTT	18	58,6					
HER2-Intron6-R	TACGGTAGCAGAGACTTGGTCT	AAACCCATACCTCTAAGAAAGGAT	24	57,8	126	56	83	rs1565923	
HER2-Ex7-F	ACACTGACGACATGGTTCTACA	CAGGGTAGTTGTGCCTGTCC	20	60,6					
HER2-Ex7-R	TACGGTAGCAGAGACTTGGTCT	ATGGGTGGCCAAACATAAGA	20	60,2	119	52	81		
HER2-Intron7-F	ACACTGACGACATGGTTCTACA	TTTCTCTTGGGAGTCTGTTGG	21	59,3					
HER2-Intron7-R	TACGGTAGCAGAGACTTGGTCT	AAGGGTCCATGTCCCTGAG	19	59,9	103	61	83	rs2952155	0,366
HER2-Ex9-F	ACACTGACGACATGGTTCTACA	GACCCATCTGCTCTCTCCTG	20	59,9					
HER2-Ex9-R	TACGGTAGCAGAGACTTGGTCT	AGCATGTCCAGGTGGGTCT	19	60,5	88	64	83		
HER2-Ex10-1F	ACACTGACGACATGGTTCTACA	GCGTCTGTGTTTCCGCTAA	19	60,0					
HER2-Ex10-1R	TACGGTAGCAGAGACTTGGTCT	CTCACCAGCCCTTTCTTCTG	20	60,0	114	54	81	rs4252612	0,04
HER2-Ex10-F	ACACTGACGACATGGTTCTACA	ACAACCAAGTGAGGCAGGTC	20	60,2					
HER2-Ex10-R	TACGGTAGCAGAGACTTGGTCT	GTATTGTTCAGCGGGTCTCC	20	59,6	115	57	83		
HER2-Intron10-F	ACACTGACGACATGGTTCTACA	GCATGCCCAGAGTAGAGGTG	20	60,8					
HER2-Intron10-R	TACGGTAGCAGAGACTTGGTCT	CCAGGGAGGAGTGAGTTGTC	20	59,7	89	57	81		

HER2-Ex11-F	ACACTGACGACATGGTTCTACA	GTCTTGATCCAGCGGAACC	19	60,6					
HER2-Ex11-R	TACGGTAGCAGAGACTTGGTCT	AGCCAGCTGGTTGTTCTTGT	20	59,9	84	54	79	rs35938081	0,024
HER2-Intron11-1F	ACACTGACGACATGGTTCTACA	GCTGCTGTTTGTGCCTCTCT	20	60,7					
HER2-Intron11-1R	TACGGTAGCAGAGACTTGGTCT	GAGCCCTTACACATCGGAGA	20	60,2	118	55	82	rs1810132	0,414
HER2-Intron11-2F	ACACTGACGACATGGTTCTACA	ATAGACACCAACCGCTCTCG	20	60,3					
HER2-Intron11-2R	TACGGTAGCAGAGACTTGGTCT	GCAGGCACTGGGTTGTAAGT	20	60,2	119	61	84		
HER2-Ex13-F	ACACTGACGACATGGTTCTACA	CACTGACTGCTGCCATGAG	19	59,1					
HER2-Ex13-R	TACGGTAGCAGAGACTTGGTCT	GCACAAAGCAGAGGCACATA	20	60,0	85	62	83		
HER2-Ex15-F	ACACTGACGACATGGTTCTACA	CCCGAGGTACCCACTCACT	19	59,0					
HER2-Ex15-R	TACGGTAGCAGAGACTTGGTCT	TAGGTCCCAAGAGGGTCTGA	20	59,7	113	65	86	rs4252625	0,02
HER2-Intron15-F	ACACTGACGACATGGTTCTACA	CGGAGAGCTTTGATGGGTAA	20	60,2					
HER2-Intron15-R	TACGGTAGCAGAGACTTGGTCT	ACTTCTGTCTCCTGCCATCC	20	59,3	115	56	82	rs4252627	0,482
HER2-Intron16-F	ACACTGACGACATGGTTCTACA	AAATTCCCTTTCACACTTCCTTT	23	59,4					
HER2-Intron16-R	TACGGTAGCAGAGACTTGGTCT	CGGTCTAGGTGAACACACACA	21	59,7	83	51	78	rs2934967	0,404
HER2-Ex17-F	ACACTGACGACATGGTTCTACA	GTGATGTCCACCCTGTTCCT	20	59,8					
HER2-Ex17-R	TACGGTAGCAGAGACTTGGTCT	AACACTTGGAGCTGCTCTGG	20	60,6	103	64	85		
HER2-Ex19-F	ACACTGACGACATGGTTCTACA	CTACTCGCTGACCCTGCAA	19	60,1				rs 4252633	0,02
HER2-Ex19-R	TACGGTAGCAGAGACTTGGTCT	AGCAGAGGTGGGTGTTATGG	20	60,0	107	64	85	rs2230698	0,04
HER2-Ex20-F	ACACTGACGACATGGTTCTACA	CGGTGTGAAACCTGACCTCT	20	60,2					
HER2-Ex20-R	TACGGTAGCAGAGACTTGGTCT	AAAAGACCGTTGGACTCACG	20	60,2	112	56	82		
HER2-Ex21-F	ACACTGACGACATGGTTCTACA	CTGTTTCTCCTGCAGCTGTG	20	59,8					
HER2-Ex21-R	TACGGTAGCAGAGACTTGGTCT	GTGAAGGGCAATGAAGGGTA	20	59,9	104	62	84		
HER2-Intron21-F	ACACTGACGACATGGTTCTACA	TCCGTGAAATGAGCATGTGT	20	60,1					
HER2-Intron21-R	TACGGTAGCAGAGACTTGGTCT	TTGTCACCCTCTCCTCTCCA	20	60,8	84	51	78	rs2952156	0,401
HER2-Ex22-F	ACACTGACGACATGGTTCTACA	CCCAAACTAGCCCTCAATCC	20	60,8					
HER2-Ex22-R	TACGGTAGCAGAGACTTGGTCT	AAGACCACGACCAGCAGAAT	20	59,7	97	61	83	rs1136201	0,16
HER2-Intron22-F	ACACTGACGACATGGTTCTACA	CTGCAGGAAACGGAGGTG	18	60,4					
HER2-Intron22-R	TACGGTAGCAGAGACTTGGTCT	CATCGCTCCGCTAGGTGT	18	60,0	128	71	89	rs903506	0,489
HER2-Ex24-F	ACACTGACGACATGGTTCTACA	AATTCCAGTGGCCATCAAAG	20	59,9					
HER2-Ex24-R	TACGGTAGCAGAGACTTGGTCT	GGGTCCTTCCTGTCCTCCTA	20	60,5	111	51	80		
HER2-Ex25-F	ACACTGACGACATGGTTCTACA	CATATGTCTCCCGCCTTCTG	20	60,6					
HER2-Ex25-R	TACGGTAGCAGAGACTTGGTCT	TCCCGGACATGGTCTAAGAG	20	60,1	97	58	82		

HER2-Ex26-F	ACACTGACGACATGGTTCTACA	CTCCTGAGCAGAACCTCTGG	20	60,1					
HER2-Ex26-R	TACGGTAGCAGAGACTTGGTCT	TCATCAGCTCCCACACAGTC	20	59,8	101	57	82		
HER2-Ex27-F	ACACTGACGACATGGTTCTACA	CCCAGGTTGGATGATTGACT	20	59,8					
HER2-Ex27-R	TACGGTAGCAGAGACTTGGTCT	CCCAGTACCTGGATGACCAC	20	60,2	111	58	82		
HER2-Ex30-F	ACACTGACGACATGGTTCTACA	CTGGTGCCACTCTGGAAAG	19	59,4					
HER2-Ex30-R	TACGGTAGCAGAGACTTGGTCT	CAAGTACTCGGGGTTCTCCA	20	60,1	104	58	82	rs1058808	0,49
HER2-Ex31-F	ACACTGACGACATGGTTCTACA	ACTGACCCAGCAGGTGTTCT	20	59,8					
HER2-Ex31-R	TACGGTAGCAGAGACTTGGTCT	CAGTGCTGTGAAGGAGCAGT	20	59,2	90	63	83		

*lenght without adaptors and barcode

Dilution curve *ERBB2*. Curves for each SNP ratio separately