U.PORTO

MESTRADO

MEDICINA E ONCOLOGIA MOLECULAR

Gene rearrangements in lung cancer: towards the detection in cell free nucleic acids

Joana Catarina Pereira Reis

M

2018

Mestrado em Medicina e Oncologia Molecular

Joana Catarina Pereira Reis

GENE REARRANGEMENTS IN LUNG CANCER: TOWARDS THE DETECTION IN CELL FREE NUCLEIC ACIDS

Dissertação de candidatura ao grau de Mestre em

Medicina e Oncologia Molecular, submetida

à Faculdade de Medicina da Universidade do Porto.

Orientador: Doutor José Luís Costa

Afiliação: IPATIMUP/i3S

“Remember to look up at the stars and not down at your feet. Try to make sense of what you see and wonder about what makes the Universe exist.”

- *Stephen Hawking*

1. Acknowledgments

Gostaria de, em primeiro lugar, agradecer ao meu orientador, Professor Doutor José Luís Costa. Pela oportunidade de realizar este projeto, pela motivação, disponibilidade, e paciência. O meu maior obrigado por ter sido um verdadeiro mentor no meu percurso científico e por sempre me encorajar a ir mais longe.

Ao Professor Doutor José Carlos Machado, por ter possibilitado o desenvolvimento deste trabalho no grupo *Genetic Dynamic of Cancer Cells* e por todo o que conhecimento científico transmitido.

Aos meus colegas GeDy. Por terem tornado esta jornada mais fácil com a boa disposição que já os caracteriza. Em especial à Carolina por ser um exemplo a seguir em ciência. À Susana, pelas conversas na hora certa que ajudaram em tantas situações. À Joana Marques e à Helena, por todos os momentos partilhados.

Às "meninas dos diagnósticos". Por conseguirem sempre melhorar um dia menos bom e tornarem cada momento banal em algo para recordar. Um obrigada muito especial para a minha companheira de luta, Márcia, que sempre esteve disponível para ajudar. À Rafaela, por estar sempre a uma mensagem de distância, obrigada.

À Cláudia, à Sara, à Liliana e à Raquel. Por serem o meu porto de abrigo para qualquer situação e me ajudarem sempre a ver o lado positivo de cada situação, entre cafés matinais e jantares difíceis de marcar. Um enorme obrigada!

À Joana e ao Fábio por estarem sempre presentes e disponíveis. Obrigada! (E que nunca nos faltem as jantaradas!)

Ao Tiago, meu companheiro em tudo. Por todo o apoio e compreensão. Por não só aceitar os meus objetivos mas por me ajudar a alcançá-los. Um simples obrigada não é suficiente!

A toda a minha família, por serem a melhor família que podia pedir. Obrigada.
E claro, aos meus pais e à minha irmã. Por em momento algum me deixarem desistir, por sempre me encorajarem a ir mais além e por todo o apoio e motivação que me dão, diariamente. Devo-vos tudo.

2. Relevant abreviations

A	
AF	Allelic frequency
ALK	Anaplastic lymphoma kinase
B	
BRAF	B-Raf proto-oncogene
C	
cfDNA	Cell free DNA
cfRNA	Cell free RNA
cfTNA	Cell free total nucleic acids
CTCs	Circulating tumour cells
ctDNA	Circulating tumour DNA
CLTC	Clathrin heavy chain 1
CNV	Copy number variation
CCDC66	Coiled-Coil Domain Containing 66
CUX1	Cut Like Homeobox 1
D	
DNA	Desoxirribonuleic acid
dPCR	Digital PCR
dsDNA	Double stranded DNA
E	
EML4	Echinoderm microtubule associated protein like 4
EGFR	Epidermal growth factor receptor
EMA	European Medicines Agency
Evs	Extracellular vesicles
EZR	Ezrin
F	
FGFR 1/2/3	Fibroblast Growth Factor Receptor
FISH	Fluorescence in situ hibridization
FDA	Food and Drug Administration
FFPE	formalin fixed paraffin embedded
H	
HBMS	Hydroxymethylbilane Synthase
I	
IBP	Small heat shock protein IBP

ITGB7	Integrin Subunit Beta 7
K	
KRAS	Kirsten Rat Sarcoma
L	
LMNA	Lamin A/C
LRIG3	Leucine Rich Repeats And Immunoglobulin Like Domains 3
M	
mRNA	Messenger RNA
MET	MET Proto-Oncogene
miRNA	Micro RNA
MYC	MYC proto-oncogene
N	
NRG1	Neuregulin 1
NTRK1	Neurotrophic Receptor Tyrosine Kinase 1
NGS	Next generation sequencing
NSCLC	Non small cell lung cancer
NCOA	Nuclear Receptor Coactivator 1
NTA	Nanoparticle tracking analysis
P	
PBS	Phosphate buffered saline
PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
Q	
qPCR	Quantitative PCR
R	
ROS1	Proto-oncogene tyrosine-protein kinase ROS
RON	Recepteur d'Origine Nantais
RET	Ret proto-oncogene
RNA	Ribonucleic acid
S	
SCLC	Small cell lung cancer
SLC24A2	Solute Carrier Family 24 Member 2
SDC4	Syndecan 4
T	
TRIM33	Tripartite Motif Containing 33
TPM3	Tropomyosin 3
TP53	Tumour protein 53

TKI	Tyrosine kinase inhibitor
TFG	TRK-Fused Gene

3. Summary

The survival of lung cancer patients has greatly improved with target therapies. After a promising initial response, resistance is developed and patients undergo disease progression. Only a small part of the patients are eligible for target therapies. The assortment is performed through genetic analysis of the tumor and, based on the alterations, the patient is selected for the most suitable therapy. This analysis is usually performed on a tumor biopsy, or more recently, through liquid biopsy. For the patients with genetic mutations, a non-invasive disease monitoring is already possible, based on the continuous analysis of the driver alteration in cfDNA. However, for the patients whose tumours harbor gene rearrangements, the lack of liquid biopsy strategies is related to the difficulty of handling and evaluating cell free RNA. This work aimed to make liquid biopsy accessible to lung cancer patients with gene rearrangements. The aim of this work was to develop an approach that allowed the simultaneous evaluation of DNA and RNA in cell free total nucleic acids allowing, not only to detect and monitor the driver alteration, but also the rise of resistance mutations. Moreover, we intended to apply a similar strategy but to FFPE material, to allow the analysis of tumour tissue.

For that, we optimized and validated the cfTNA extraction method in conditioned media from cell lines with different genetic alterations, prior to its application in plasma samples from NSCLC patients, which were sequenced with a panel that allowed the detection of point mutations, deletions and fusion transcripts in cfTNA, later confirmed by dPCR.

We observed that cfRNA is resistant to nuclease degradation in circulation even without shielding in extracellular vesicles, probably due to association with protein complexes. Moreover, we determined that the EV cfRNA only contributes in a small part to the overall amount of cfRNA. Moreover, we determined that the cfRNA and cfDNA yield are severely compromised with freezing. Additionally, by applying the methodology to clinical samples, aside from being able to detect the driver genetic rearrangement, we were also able to detect different mutations with possible clinical impact in response to therapy.

In summary, we developed a feasible methodology that allows not only the extraction of cfTNA but also the simultaneous detection of genetic mutations and gene translocations, allowing to expand the clinical application of liquid biopsy monitoring to patients that otherwise would not be enrolled in this type of non-invasive monitoring.

4. Resumo

A sobrevivência dos pacientes de cancro de pulmão sofreu desenvolvimentos significativos com a implementação de terapias dirigidas. No entanto, após uma resposta inicial promissora, ocorre o desenvolvimento de resistência que culmina em doença progressiva. Apenas alguns doentes são elegíveis para este tipo de terapia, o que leva à necessidade de testar o perfil molecular do tumor para decidir qual o melhor plano terapêutico para cada doente. Este análise é tradicionalmente realizada em biópsias tumorais, ou, mais recentemente, através de biópsia líquida. Para os pacientes com mutações conhecidas, é já possível realizar uma monitorização não invasiva através da da mutação fundadora em ADN circulante tumoral. No entanto, para os pacientes cujos tumores contêm rearranjos genéticos, a falta de opções de biópsia líquida relaciona-se principalmente com a dificuldade em analisar ARN circulante. O principal objetivo deste projeto foi desenvolver uma estratégia que tornasse possível a avaliação simultânea de ADN e ARN em ácidos nucleicos totais circulantes, permitindo a deteção da mutação transformante mas também o aparecimento de mutações de resistência. Para além disso, pretendeu-se também desenvolver uma estratégia semelhante, aplicável a material tumoral fixado.

Para isso, otimizámos e validamos um método de extração de ácidos nucleicos totais circulantes em meio condicionado de linhas celulares com diferentes alterações genéticas, precedendo a aplicação em amostras de plasma de pacientes cancro do pulmão. As amostras foram sequenciadas com um painel que permite a deteção de mutações pontuais, deleções e transcritos de fusão, posteriormente confirmadas por PCR digital.

Observámos que o ARN circulante é resistente à degradação por nucleases em circulação, o que se deverá à associação com complexos proteicos, e não derivado da combinação com vesículas extracelulares, cujo conteúdo em ARN circulante apenas contribui numa pequena percentagem para a quantidade de ARN encontrada em circulação. Determinamos que a quantidade de ADN e ARN circulante é severamente comprometida com o processo de congelação. Adicionalmente, através da aplicação do método a amostras clínicas, foi possível a deteção do rearranjo genético transformante, bem como diferentes mutações, com possível impacto na resposta à terapia.

Sumariamente, desenvolvemos uma metodologia que permite não só a extração de ácidos nucleicos totais mas também a deteção simultânea de mutações e translocações, permitindo a expansão da aplicabilidade da biópsia líquida a pacientes que, de outra forma, jamais seriam incluídos neste tipo de monitorização não invasiva.

Table of contents

5. Figure index

6. Table index

7. Introduction

7.1 Molecular basis of cancer

Cancer is a disease primarily driven by the accumulation of genomic and epigenomic alterations that, altogether culminate in uncontrolled proliferation and impaired cell death (1). It is now one of the main causes of deaths throughout the world and, with the continuous increase in the incidence rates worldwide, it is considered an emergent public health issue. According to the latest statistics, in the United States of America alone, more than 2 million people were diagnosed with cancer in 2017 and more than 600000 cancer related deaths were reported (2).

After several decades of research to understand the basic aspects of tumour biology, cancer is now considered an heterogeneous disease, developed from the interaction between environmental and genetic factors (3, 4). The carcinogenic process can be divided into three main stages: initiation, promotion and progression (5). The initiation phase starts through the occurrence of an irreversible genetic modification, such as point mutations, translocations, deletions or insertions in the desoxirribonuleic acid (DNA) sequence. Following initiation, the promotion phase takes place. During this stage, the promoting agents support clonal expansion of the previously initiated cells towards malignant conversion. Tumour progression follows, and transformed cells advance to higher degrees of malignancy where characteristics such as anaplasia, growth rate and invasive capacity become more pronounced and eventually culminate in metastatic growth (5). It is now accepted that the development of a tumour follows a pathways similar to Darwinian evolution, where a succession of genetic alterations occur and each one confers some type of growth advantage, ultimately leading to the conversion into cancer cells (Figure 1) (6).

It is described that a single cancer cell can harbour thousands of different genetic and epigenetic alterations, as a result of the rounds of cell division (7, 8). However, not all alterations have the same impact on cancer development. Only a few mutations, described as “driver” confer growth advantage. The number of driver mutations in a single cancer cell is directly proportional to the degree of deregulation of the biological processes (7). All the other alterations that cancer cells harbour are defined as “passengers”. They are accumulated with every cell division and may or may not contribute to the malignant phenotype, but reflect the number of mitotic cell divisions that have occurred (7-9). So, the set of somatic mutations of a cancer cell represents the genomic changes that were accumulated over the lifetime of the affected individual,

including both mutations that indeed confer a malignant phenotype and also the marks of the mutational process that led to the development of the disease (3, 8).

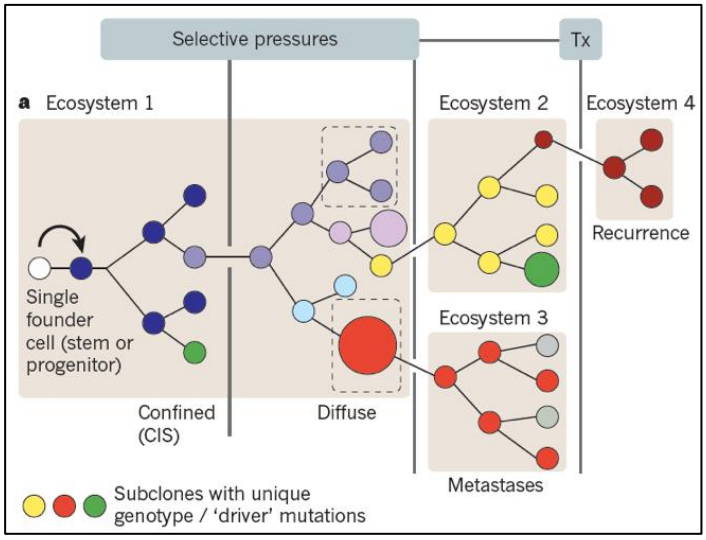

Figure 1: The branching architecture of cancer evolution representing the Darwinian evolution of a tumour. In this case, the selective pressure is exerted by the treatment that lead to the extinction or dominance of certain clones. Greaves et al, 2012.

Despite all the improvements achieved in the oncology field, cancer incidence is expected to continue to increase. Stomach, colorectal and lung cancer are the most common cancers in both genders, apart from breast and prostate in women and men, respectively (2). Certain lifestyle habits were already associated with the increase in cancer incidence, such as smoking, sedentary lifestyle and non-healthy diets (2, 10).

7.2 Lung cancer

In the last 100 years, lung cancer shifted from being considered an extremely rare disease to become the most frequently diagnosed and the main cause of cancer related deaths. In 2017, lung cancer was responsible for the death of more than 150000 people in the United States of America alone (2, 10, 11). The reason behind the quick rise in prevalence of lung cancer is tobacco smoking. However, chronic respiratory pathologies, pollution and exposure to radon gas or asbestos are also important agents (12).

Lung cancer can be divided into 2 major groups according to histopathological features displayed by tumour cells: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). The last category accounts for over 87% of reported lung cancer cases with only a 15% 5-year survival rate (13, 14). The NSCLC category includes squamous cell

carcinoma, large cell carcinoma and, representing about 50% of all lung cancers, adenocarcinoma (15, 16) (Figure 2).

The treatment for lung cancer patients is dependent on the stage the disease is diagnosed. When the tumour is detected in an early stage (I and II) surgery is used as a treatment option with curative intent, being associated with very good long term survival. However, most of the patients are diagnosed at later stages (III or IV), when the tumour is no longer resectable. The available treatment plans for these patients are most frequently platinum-based chemotherapy, targeted therapy or, more recently, immunotherapy (17).

7.3 Molecular classification of cancer

A major part of the improvement in the management of lung cancer patients observed in the last decades was the effort made into building a detailed profile of the tumour, namely in the detections of alterations that explain the oncogenic properties of tumour cells. Now we have a much more comprehensive understanding of its molecular basis and, consequently, its pathogenesis (18, 19). In the last decade, the molecular-based classification gained relevance alongside with histology and it is now a fundamental part of the diagnosis (Figure 2) (20, 21).

Multiple genetic alterations have already been identified as therapeutic targets for NSCLC (22). Point mutations affecting the Epidermal growth factor receptor (*EGFR)* pathway (Kirsten Rat Sarcoma (*KRAS),* B-Raf proto-oncogene *(BRAF)*, etc) are the most common driver alteration in NSCLC (Figure 2) (23), leading to uncontrolled cell proliferation and evasion of apoptosis (24). However, tumours with gene rearrangements as driver alterations represent about 10% of all cases, even though this value is highly variable, according with the cohort evaluated by each study. The incidence of echinoderm microtubule associated protein like 4 (*EML4*) - anaplastic lymphoma kinase (*ALK*) (EML4-ALK) rearrangement in NSCLC is the most commonly detected translocation (25), and its incidence can be as high as 17%, if a non-smoker young population is studied (26-28).

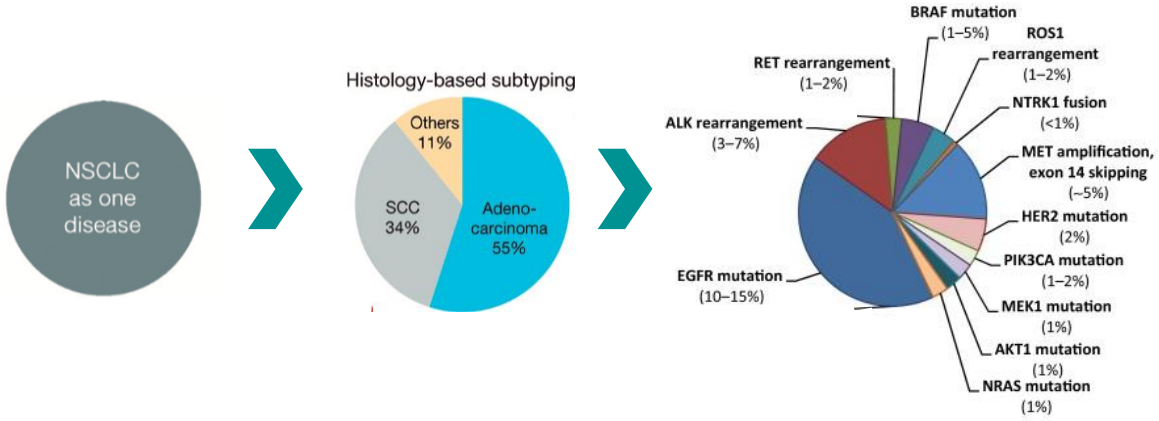

Figure 2: Evolution of the classification of lung tumours. Adapted from Lin et al, 2016.

7.4 EML4-ALK rearrangement in NSCLC

The EML4-ALK rearrangement is derived from the fusion of the 5' end of *EML4* to the intracellular tyrosine kinase of *ALK*, leading to a constitutive expression of an oncogenic fusion protein and, consequently, aberrant expression of the ALK fusion protein, inducing a constitutive activation of downstream signalling pathways, ultimately culminating in uncontrolled cellular proliferation and survival (28-31). This rearrangement is caused by the inversion of the short arm of chromosome 2 of *ALK* (2p23) and *EML4* (2p21). The most common variants (v1:E13; A20, v2: E20;A20 and v3:E6;A20) (Figure 3) account for more than 90% of EML4-ALK associated NSCLC (28, 32). Besides these variants, there are currently over than 15 variants of the EML4-ALK translocation already identified, whereas some of these variants are expressed as multiple isoforms (33).

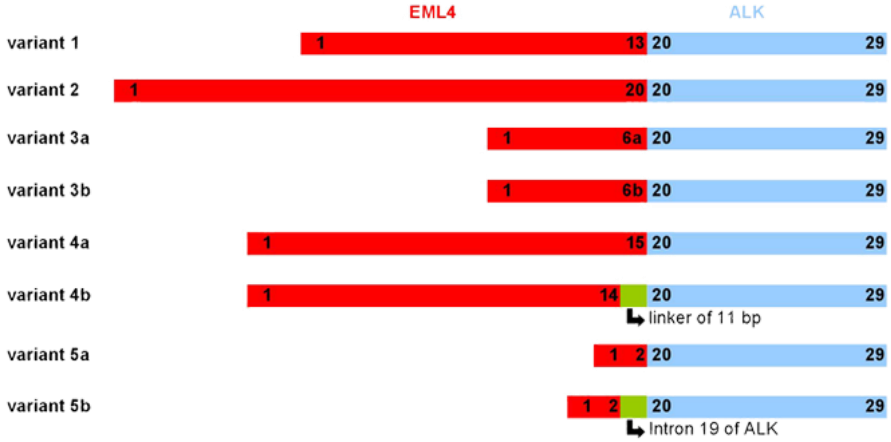

Figure 3: Described variants of the EML4-ALK translocation. Variant 1: exon 1-13 (EML4) + exon 20-29. Variant 2: exon 1-20 (EML4) + exon 20-29. Variant 3a: exon 1-6a (EML4) + exon 20-29. Variant 3b: exon 1-6b (EML4) + exon 20-29. Variant 4a : exon 15 (EML4) + exon 20-29. Variant 4b : exon 14 (EML4) + linker of 11bp + exon 20-29. Variant 5a : exon 2 (EML4) + exon 20-29 Variant 5b : exon 2 (EML4) + intron 19 + exon 20-29. Adapted from Atlas of Genetics and Cytogenetics in Oncology and Haematology.

Nevertheless, *EML4* is not an exclusive fusion partner for *ALK*. More recently, other fusion partners were described in NSCLC, such as the TRK-Fused Gene (*TFG)*, however, their incidence is considerably lower. Regardless of the fusion partner, the breakpoint on the *ALK* gene occurs mostly in intron 19, only in rare cases the exon 20 is affected, which results in the inclusion of *ALK* kinase domain in the fused gene (28, 34).

More recently, the list of genes in which translocations occur in NSCLC was expanded, and oncogenes such as RET proto-oncogene (*RET*), Proto-oncogene tyrosine-protein kinase ROS (*ROS1*), Neurotrophic Receptor Tyrosine Kinase 1 (*NTRK1*), Neuregulin 1 (*NRG1*) and Fibroblast Growth Factor Receptor 1/2/3 (*FGFR1/2/3*) were added to the list, even though affecting a much smaller fraction of lung cancer patients. The occurrence of fusions affecting the *RET* gene was discovered in 2012, and, since then, seven different partner genes were identified, such as the Kinesin-1 heavy chain (*KIF5B)* more often, but also Coiled-Coil Domain Containing 6 (*CCDC6*), Nuclear Receptor Coactivator 1 (*NCOA*), Tripartite Motif Containing 33 (*TRIM33*) and Cut Like Homeobox 1 (*CUX1*). It is estimated that *RET* fusions are present in about 1-2% of all lung adenocarcinoma cases, in a mutually exclusive fashion with other genetic alterations (35-42). Regarding *ROS1* fusions, *CD74* is the most common partner gene, but rearrangements affecting Ezrin (*EZR*), Solute Carrier Family 24 Member 2 (*SLC24A2*), Tropomyosin 3 (*TPM3*), Syndecan 4 (*SDC4*), Leucine Rich Repeats And Immunoglobulin Like Domains 3 (*LRIG3*), *CCDC6* or Clathrin heavy chain 1 (*CLTC*) gene have also been described (35, 42-51). *ROS 1* translocations are more commonly detected in young females with no or light history of smoking habits. They are described has specific for lung adenocarcinoma and they also occur in a mutually exclusive manner with other oncogenic alterations (35, 42-51).

7.5 Targeted therapy in NSCLC

The identification of the most adequate targeted therapy plan according with the molecular profile of the tumour resulted in significant improvements when comparing with the standard of care, first-line platinum doublet chemotherapy.

Whenever an *ALK* or *ROS 1* rearrangement is detected, the therapy regimen is directed towards tyrosine kinase inhibitors (TKI) for these targets. The currently approved treatment pipeline relies on the use of crizotinib as a first-line drug. Crizotinib is an *ALK* oral TKI, also inhibiting hepatocyte growth factor receptor (c-MET), ROS1 and Recepteur d'Origine Nantais (RON). Resistance to crizotinib is usually observed around 9 months after the start of the treatment, and it can be related to different causes, mainly related

with mutations or copy number variations affecting *ALK*. Also, mutations in the *EGFR* or *KRAS* gene can also play a role in resistance to crizotinib. (Figure 4) (52, 53).

Upon the establishment of resistance to crizotinib, a second line therapy regimen is applied, such as ceritinib. Ceritinib is an oral *ALK* TKI, also inhibiting insulin-like growth factor-1 receptors, insulin receptors, and *ROS1*, but not *MET*. In vitro studies show that ceritinib is 20x more potent in inhibiting *ALK* than crizotinib (54, 55). Similarly to ceritinib, alectinib is a second generation oral *ALK* TKI with ability to also inhibit RET, but not *MET* or *ROS1*. Its capacity to inhibit *ALK* is about 5x higher than crizotinib, also having clinical impact when *ALK*-resistance mutations are present (56-58).

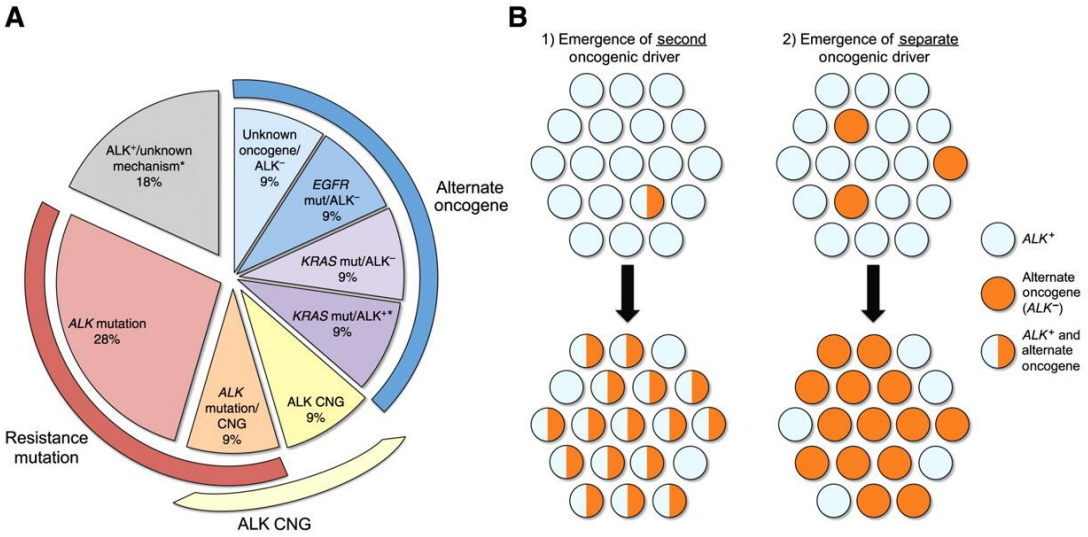

Figure 4: Molecular mechanisms associated with crizotinib resistance. Doebele et al, 2014.

7.6 Molecular pathology of lung tumours

The gold standard material to analyse the histology and the genetic landscape of the tumour is tissue biopsy, where a small portion of the tumour is removed with a needle. The complexity of the technique is mostly associated with the location or size of the tumour, and often the procedure cannot be performed due to the risk of complications (59, 60). Furthermore, considering the small portion of tumour removed, single-site biopsies might not be representative of tumour heterogeneity. In fact, up to 23% of tumour biopsies obtained during disease progression provide limited amount of biological material, with poor quality for genetic analysis (60), resulting in a possible failure in identifying the predominant resistance mechanism for each patient (61). Moreover, the pre-analytical handling of the sample is a critical factor, and when performed incorrectly, all the downstream analysis might be compromised. It is described

that a bad fixation procedure (either derived from a delay in fixation or extensive/insufficient fixation time) may impair the detection of *ALK* protein (62).

Moreover, the quality of nucleic acids is also significantly compromised in formalin fixed paraffin embeddeed (FFPE) samples, which occurs during tissue processing, due to the cross-linking action of some fixatives (formaldehyde, for example). The detrimental effects of fixation are directly proportional with the duration of the fixation protocol, affecting both the yield as well as the integrity of the nucleic acids (63, 64).

In order to analyse for the presence of genetic translocations, different methods can be used. Fluorescent *in situ* hybridization (FISH) was the first method to be developed to detect *ALK* rearrangement and is, up to today, widely considered as the gold standard approach. The currently Food and Drug Administration approved technique is based on hybridization with break apart FISH probes and evaluation by fluorescent microscopy (54). In order to report an *ALK* translocation, the following criteria have to be present:

- The *ALK* probe shows separated red (linked to the 3' *ALK* kinase encoding region) and green (bound to the 5' region to *ALK*) fluorophores, or
- Loss of the green signal in more than 15% of examined cells.

Despite its wide use, FISH presents several disadvantages, namely the high cost associated, the need for the samples to be analyzed by trained and experienced staff and the extended turnaround time. More importantly, it does not allow the specification of the translocation variant or partner (65).

Next generation sequencing (66) has more recently been applied to the field of fusion transcripts. The ability of this technique to analyze more than one gene at any given time revolutionized the molecular pathology field, and allowed translocations in *ROS1* or *NTRK1* to be detected without subsequent analysis, that would only increase the cost and the turnaround time of the test. Moreover, it also allows to detect the specific variants and fusion partners, which provides a more comprehensive insight into the tumour biology. Nonetheless, even though NGS usually requires small amounts of RNA (about 10ng), the application of this method to FFPE, whose genetic material is of very poor quality, can result in false negative results (28).

7.7 Liquid biopsy

In order to overcome the limitations associated with tissue biopsy, a lot of effort has been applied to the development of the liquid biopsy field. The whole concept of liquid biopsy

relies on the detection of tumour-associated biomarkers (point mutations, copy number aberrations, gene rearrangements) in body fluids, such as urine, cerebrospinal fluid, sputum, or, more commonly, peripheral blood. This is only possible because all cells, either malignant or normal, have the ability to shed their nucleic acids into circulation.

7.7.1 Cell free DNA

The first observation that cell free DNA (cfDNA) was present in the blood of healthy individuals was made in 1948 by Mandel and Metais (67). However, only in 1966 cfDNA was reported by Tan et al to be present also in non-physiological conditions, such as systemic lupus erythematous (68). The first association of cfDNA with the oncology field was made in 1977 by Leon et al, that observed a higher concentrations of cfDNA in cancer patients when comparing to healthy individuals (69). From that point on, and mainly after the discovery that it was possibly to detect tumour derived somatic variants in cfDNA (70, 71), the interest towards liquid biopsy continued to increase, providing a significant amount of information not only about the clinical use of cfDNA but also about its' biology. Regardless of all the developments achieved in the field, a lot it still unknown about cfDNA, namely the mechanisms through which it is released, that still represents a matter of discussion (Figure 5).

Apoptosis is accepted as the main source of release, as the size distribution pattern of cfDNA, assessed by electrophoresis, is composed mainly by fragments of around 170 base pairs (bp). This corresponds to mononucleossome units, which is in accordance with the pattern obtained when observing apoptotic cleaved DNA, a result of caspase activated DNAse action. The nucleosome is composed by a double stranded DNA (dsDNA) rolled around a histone octamer, corresponding in a size of about 147 bp. The linker DNA (around 20-90bp) is responsible to bind one nucleosome to another, thus guaranteeing the structural integrity of DNA, protecting it from the enzymatic degradation that occurs in the circulatory system (72-76). Necrosis, a very common occurrence in the tumour microenvironment, is another described mechanism for cfDNA release. Nonetheless, cfDNA of necrotic origin is expected to have longer DNA fragments (77). For both of these mechanisms, the genetic material is released into circulation in a free form, therefore more susceptible to nuclease degradation. In order to assess how cell free nucleic acids endure in circulation, some studies show that cancer cells may actively secrete their genetic material into exosomes, thus protecting from degradation (78-80). Exosomes are microvesicules secreted by most cells and have both functional and biological properties, having the ability to transfer biological material between cells. Exosomes are composed mostly by lipids, but also contain proteins, messenger RNA

(mRNA) and micro RNA (miRNA). The presence of tumour derived double stranded DNA (dsDNA) in exosomes has also been recently described in a report by Thakur et al in 2014, who showed that dsDNA can be present both on the membrane (usually dsDNA of larger size) but also inside the exosomes (normally dsDNA of smaller size) (81).

Circulating tumour cells (CTCs) also contribute for the presence of cfDNA in circulation, even though their impact is probably less significant. Considering that on average, there are less than 10 CTCs per 7,5mL of blood, which represents a very small percentage of the 17ng of DNA per mL of blood described to be present in advanced stage cancer patients (82, 83).

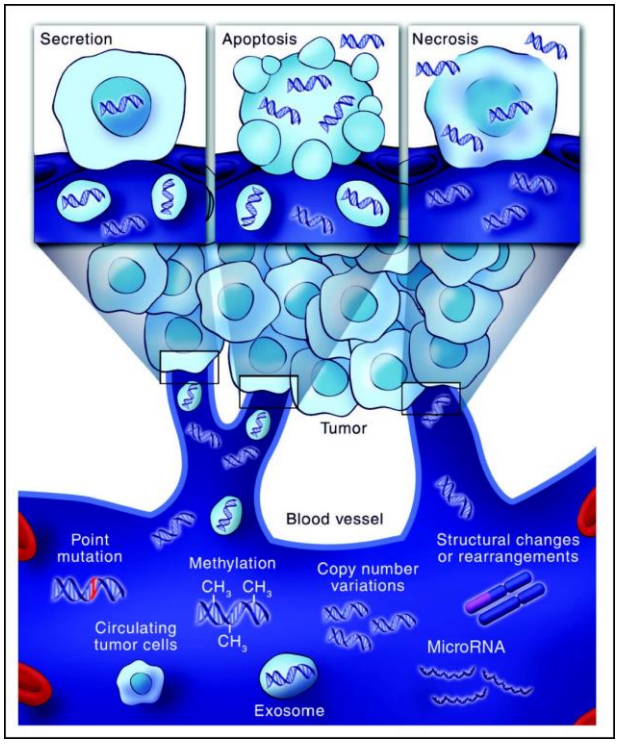

Figure 5: Mechanims of cfDNA release into the bloodstream. Tumour derived genetic material can be released as DNA/RNA fragments or shielded inside exosomes, circulating tumour cells (CTC's). Diaz et al, 2014.

Most of the studies recently or currently conducted focus on the possible role of liquid biopsy as a monitoring tool to detect resistance to therapy or minimal residual disease. (84). Several authors, and more recently Veldore et al (85) suggests the potential application of liquid biopsy as a valid substitute when tissue biopsy cannot be performed or when the low amount of biological material is limited to perform all the necessary genetic testing (86).

The overall quantification of overall cfDNA alone can be used as a prognostic factor, being directly proportional to the stage of disease (87, 88) considering that patients with more advanced tumours present higher levels of cfDNA in circulation (89). However, several factors can interfere with cfDNA amount, such as autoimmune diseases or inflammatory conditions and so, evaluating cfDNA levels alone might lead to an overestimation of tumour burden. Taking into consideration that tumour cells harbour specific somatic mutations, they represent the ideal biomarker to accurately quantify tumour derived cell free DNA fraction, more commonly referred to as circulating tumour DNA (ctDNA). The amount of ctDNA should be proportional to disease burden at any given time, which gave rise to a lot of interest towards its prognostic value. It is now described that by evaluating serial liquid biopsy samples from the same patient during treatment, it is possible to reflect the course of disease by evaluating the dynamic changes of the tumour derived mutations in circulation. More specifically, several reports show that upon response to treatment, there is a significant decrease of ctDNA levels, while during stable disease it is not expected that ctDNA levels vary significantly, increasing only in a scenario of disease progression. The increase of ctDNA can be predicted by an increase in the allelic frequency of tumour derived somatic mutations, an indication of an augmented tumour burden.

NSCLC has been a prime model of study for liquid biopsy, mainly due to its very well established mutational profile. For this reason, many studies have been published showing the potential of this approach in the management of late stage NSCLC patients, reporting that with this approach, the occurrence of disease progression can be detected in circulation up to 3 months earlier than any radiological signs are detected. However, in most studies, the biomarker used to identify and quantify the ctDNA fraction are tumour somatic mutations, which restrains the application of the method to patients with known actionable mutations. In order to assess the variants present in cfDNA, many methodologies are available, with variable cost and limit of detection. The most commonly used techniques are summarized in Table 1.

The break-throughs achieved with liquid biopsy have led to the application of tests in clinical practice. However, their implementation was preceded by approval of regulatory agencies, such as the Food and Drug Administration (54) (54) or European Medicines Agency (EMA), whereas the cobas® EGFR Mutation Test for non-small cell lung cancer and the QIAGEN's therascreen® EGFR RGQ PCR Kit represent, until now, the only validated liquid biopsy tests for cfDNA evaluation.

7.7.2 Cell free RNA

The existence of RNA in circulation was acknowledged several years ago, however, the potential application in clinical practice of RNA based-biomarker is not as successful as DNA-based biomarkers (90).

Table 1: Currently available technologies to evaluate liquid biopsy samples

Approach	Method	Detection limit
Specific mutation detection (mutational status previously determined)	qPCR	<10% (91, 92)
	dPCR	<0.001% (93)
	Scorpion ARMS (Amplification-Refractory Mutation System)	<0.1% (89)
	BEAMing (Beads, Emulsion, Amplification, Magnetics)	<0.1% (94)
	QIAGEN's therascreen® EGFR RGQ PCR Kit	>0.81% (95)
	cobas® EGFR Mutation Test	3% (96)
Deep sequencing (screening of unkown alterations)	Ampliseq	>2% (97)
	Tam-Seq	>2% (98)
	CAPP-Seq	0.01% (99)
	Safe-Seq	0.1% (100)
	Tag sequencing	0.1% (101, 102)
	Whole exome sequencing	>1% (103)

The explanation for that fact could rely of the biology of cfRNA. RNA, within cells, has a tightly regulated life cycle that results in its transient nature (104). Moreover, RNA molecules are very unstable and susceptible to different factors where alkaline pH, heavy metal ions, RNA hydrolyzing enzyme are only a few examples. Biological fluids are particularly rich in RNA hydrolyzing enzymes, as is the case of RNAse A superfamily members (105-107). The action of this enzyme in circulation can justify the very short half-time life of naked RNA of just 15 seconds (108). The association with proteins, lipoproteins and shielding within extracellular vesicles (EV's) can significantly increase the mean half time life of cfRNA up to a few hours, allowing to overcome the stability issue (109-111). Secretion is, therefore, considered as the main source of cfRNA in circulation, and not cell death, as it happen with cfDNA (90).

The promising potential of cfRNA goes beyond expression analysis, allowing the detection of tissue specific transcripts, evaluation of RNA molecules that underwent alternative splicing or post-transcriptional modifications and, most importantly for this work: the detection of fusion events (112, 113).

The detection of tumor derived somatic mutations, as well as gene rearrangements, is possible by analyzing cell free mRNA (114). Tumour-derived mRNAs are already

described to be present in relative abundance in peripheral blood, urine and sputum (115-117). Despite different potential mRNA biomarkers have been identified, such as the hTERT mRNA, not much progress was achieved in the area lately, mainly due to the difficulty of retrieving cfRNA from biological fluids, given its fragmented state and low abundance (118). Despite the limitations, it is described that cancer patients have increased levels of cfRNA in circulation, thus confirming the potential of cfRNA as a biomarker for diagnosis and monitoring (119, 120).

8. Rational and aims

The liquid biopsy field is in clear expansion, especially for NSCLC, however, this methodology is still not viable to be applied to all patients. For those whose tumours harbour transforming mutations, several liquid biopsy strategies are already available, based on the continuous monitoring of the driver mutation present on cfDNA. However, for the patients with gene rearrangements this is not yet possible.

The study of rearrangements in circulation implies the evaluation of cfRNA, which comprehends several challenges, including the low abundance in circulation and the rapid nuclease degradation (121). Moreover, the nonexistence of efficient extraction methods to retrieve both DNA and RNA from the same biological sample to allow the detection of the driver rearrangement but also the rise of mutations that confer resistance to targeted therapy has limited the application of liquid biopsy to this cohort of NSCLC patients.

Taking all of this into account, the major goal of this study is to develop a liquid biopsy workflow focused on NSCLC patients, also including the patients with gene translocations. Specifically, we aim:

- To optimize a total cell free nucleic acids (cfTNA) extraction method from liquid biopsy samples that allows simultaneous evaluation of DNA and RNA. This will allow us to enhance the output from the biological sample and to diminish the complexity of the process.
- To detect genetic rearrangements and mutations in cfTNA, using NGS and/ or digital PCR (dPCR). With this we expect to end up with an approach feasible to be applied at diagnosis but also during disease course, to allow continuous monitoring of tumour burden.
- To apply the developed strategy to FFPE specimens, to allow the pipeline to be used also in cases where liquid biopsy is not conclusive.

This project might have an impact in the future management of NSCLC patients towards a faster genetic diagnosis and disease monitoring due to a simplified process and, subsequently, a quicker administration of the most suitable therapy plan, for any stage of disease course.

9. Material and methods

9.1 Biological material

9.1.1 Cell lines

Within the scope of this work, different biological materials were used. For the *in vitro* experiments, NCI-H2228 (ATCC – American Type Culture Collection, Lockville, MD, USA), HCC78 (DSMZ, Braunschweig, Germany) and H1975 (ATCC) cell lines were used (Figure 6) (Table 2).

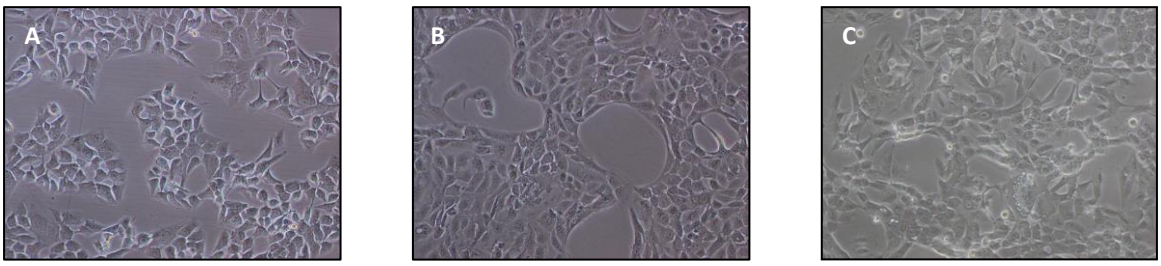

Figure 6: Inverted microscopy images of H2228 (A), HCC78 (B) and H1975 (C) cell lines in culture.

Table 2: Relevant genetic alterations of the lung cancer cell lines used in the study

Cell line	Genetic alteration
H2228	EML4-ALK fusion
HCC78	SLC4A2-ROS1 fusion
H1975	EGFR p.T790M and p.L858Rl

One vial for each cell line was thawed and maintained in culture with RPMI 1640 (GIBCO®, Life Technologies, Carlsbad, CA, USA) supplemented with 10% heat inactivated Fetal Bovine Serum (FBS) (Biowest, Nuaillé, France) and 1% Penicilin-Streptomyin (P/S) solution (Gibco). Cells were maintained in an humidified incubator at 37ºC with 5% Co_2, where they grew in an adherent monolayer. To subculture the cells, a Trypsin-EDTA (Sigma-Aldrich®, St. Louis, MO, USA) solution was used to harvest the cells. Cells were subcultured in a 1:3 ratio whenever 80% of confluency was observed.

The cultured cells were regularly tested for Mycoplasma spp. contamination by Polymerase Chain Reaction (PCR), combining a par of primers specific for the 16S RNA of Mollicutes (MGSO: 5`- TGC ACC ATG TGT CAC TCT GTT AAC CTC – 3` and GPO1:

5`- ACT CCT ACG GGA GGC AGC AGT A – 3) (IDT – Integrated DNA Technologies, Coralville, IA, USA). The cells were also genotyped to confirm their authenticity, through STR sequencing. Briefly, DNA from cells was amplified through PCR with primers for the following locus: Penta E, D18S51, D21S11, THO1, D3S1358, FGA, TPOX, D81179, vWA, Amelogenin, Penta D, CSF1PO, D16S539, D7S820, D14s317 and D5S818 with the Powerplex 16 HS system, that allows the co-amplification and simultaneous detection of the 16 described loci. The amplified fragments were then detected with capillary electrophoresis using the 3500 Genetic Analyser sequencer (Applied Biosystems). Afterwards, the genotypes were assigned with the GeneMapper v5.0 (Applied Biosystems).

9.1.2 Cell blocks

For HCC78 and H2228 cell lines, cell blocks were prepared. To prepare the cell blocks, cells in about 70-80% confluency were detached using Trypsin and pelleted by centrifugation. The pellet was resuspended in complete culture media and cells were counted using a TC20 Automated Cell Counter (Bio-Rad Laboratories, Hercules, CA, USA). Approximately 3 millions cells were fixed in 10% formalin (Bio-Optica, Milan, Italy) for 24 hours. Then, 2mL of Histogel (Thermo Scientific, Waltham, MA, USA) were liquefied in a microwave using a 500W potency. After the cells were pelleted again and the supernatant discarded, the cells were resuspended in the Histogel fluid and incubated at 4ºC until completely solidified. The Histogel-cell pellet was then removed and processed as an histological sample. Briefly, the cassettes were inserted in a tissue processor, where tissues were immersed in solutions of: 70% ethanol, 80% ethanol, 90% ethanol, 96% ethanol, 2x in 100% ethanol, 2x in Clear Rite for 1h each and in 2x Paraffin for 1h and 20 min each. Then, the processed cell blocks were embedded in paraffin wax for analysis.

In order to confirm the presence of cells prior to proceed to additional studies, tissue sections of 3 µm were obtained for the cell blocks of both cell lines in non-coated slides, using the MICROM HM 335E microtome followed by an incubation at 65ºC for 1 hour to adhere the sections to the slides. The sections were then stained using the hematoxylin-eosin staining. Samples were dewaxed with xylene, followed by submersion in solutions with decreasing concentrations of alcohol (100%, 100%, 70%) and then rinsed with running water to hydrate. After staining with Modified Gill II Hematoxylin (Merck Millipore, Burlington, MA, EUA) and differentiation with running water, the sections were dehydrated using solutions with increasing concentrations of alcohol (70%, 100%, 100%) and then stained with an alcoholic eosin solution (Thermo Scientific), quickly rinsed in

100% ethanol and then diaphanized twice with xylene. Samples were then mounted using DPX mounting medium (Sigma-Aldrich®, St. Louis, MO, USA) a glass coverslip (Normax, Marinha Grande, Portugal). To visualize the staining, an optical microscope was used.

9.1.3 Plasma

Human plasma samples were also used, from a previously established collaboration with Hospital de São João, we had access to a retrospective cohort of blood samples collected from stage IIIB/IV NSCLC patients, whose tumours had already been tested for the presence of fusion transcripts as driver alterations. The hospital's ethics committee approved the study and for every patient enrolled, an informed consent was obtained.

By analysing the clinical information for each patient, the timepoints corresponding to progressive disease were identified as being the most informative regarding tumour burden and development of resistance mechanisms. Therefore, the samples collected during that period were used for cfTNA extraction and analysis. We performed a comprehensive follow up for one patient, a 77 year old man with a history of smoking habits, diagnosed in August 2014 with a T4N2M1a non-small cell lung cancer.

The BD Vacutainer® PPT™ tubes (Becton Dickinson, Franklin Lakes, NJ, USA) were used for sample collecting, allowing a physical separation of plasma after a centrifugation of 12000g for 10minutes. The isolated plasma was stored at -20ºC until further analysis.

9.1.4 Synthetic controls

The synthetic cfTNA controls (HorizonDx, Cambridge, England) were used to mimic the possible outcomes to be encountered in clinical samples, namely variable MET amplification together with different genetic translocations (Lung low and high copy control), single nucleotide variations affecting different genes (Lung SNV control), and a wild type control to assure specificity (negative control). The variants present in each one of the controls are described in Table 3.

Table 3: Synthetic controls expected results

Control	SNV/Indel			CNV			Fusions (variants)
	Gene	AA change	MAF %	Gene	Gain/Loss	CNV Ratio	
Low Copy Control	NRAS	p.Q61K	NA	MET	↑	1,42	EML4-ALK.E6bA20.AB374362
	KRAS	P.G12A					SLC34A2-ROS1.S4R32.COSF1196
	TP53	p.R248L					SLC34A2-ROS1.S4R34.COSF1198
	TP53	p.G245C					CCDC6-RET.C1R12.COSF1271
							EML4-ALK.E6aA20.AB374361
							MET-MET.M13M15
High Copy Control	KRAS	P.G12A	NA	MET	↑	1,72	EML4-ALK.E6bA20.AB374362
	TP53	p.R248L					SLC34A2-ROS1.S4R32.COSF1196
	TP53	p.G245C					MET-MET.M13M15
							CCDC6-RET.C1R12.COSF1271
							SLC34A2-ROS1.S4R34.COSF1198
SNV Control	NRAS	p.G12D	0,922	ND	ND	ND	ND
	ALK	p.R1275Q	0,628				
	PIK3CA	p.H1047R	0,548				
	BRAF	p.V600E	0,584				
	KRAS	p.Q61H	0,887				
	KRAS	p.G12D	0,609				
	TP53	p.R273H	0,482				
	TP53	p.R248Q	0,471				
	TP53	p.Y234C	0,471				
	TP53	p.Y220C	0,471				
	ERBB2	p.Glu770_Ala771insAlaTyrValMet	0,471				
Negative Control	ND			ND			ND

ND: Not detected; NA: not applicable

9.2 Methodology

9.2.1 Nucleic acids extraction

9.2.1.1 DNA extraction

DNA was extracted from cell pellets containing 5 million cells were lysed using Protease K. Following precipitation with isopropanol and purification with ethanol, the DNA was eluted in a Tris-HCl/EDTA buffer and stored at -20ºC.

To extract DNA from the cell blocks, the MagMax FFPE DNA/RNA Ultra Kit (Applied Biosystems, Life technologies, Waltham, MA, USA) was used, a kit based on the use of magnetic beads to retrieve. Briefly, 40 µM from the cell block were used for protease K digestion. Following this digestion, a binding buffer and dynabeads were added to allow DNA binding to the beads. After the binding is complete, the beads were pelleted against a magnetic stand. A sequence of washing steps with buffers containing ethanol and isopropanol was executed to remove impurities and the DNA was eluted in 30µL of elution buffer.

9.2.1.2 RNA extraction

Around 6 millions cells were detached from the culture flask using trypsin and pelleted. The pellet was then resuspended in Trizol reagent (BioRad, Hercules, California, USA), and incubated at room temperature for 5 minutes. Then, chloroform was added, and, after centrifugation to allow the separation of phases, the aqueous phase containing the RNA was separated. To precipitate RNA, ice cold isopropanol (VWR, Radnor, Pennsylvania, USA) was added, followed by a 75% ethanol (VWR) wash. The RNA pellet was then eluted in nuclease free water and stored at -20ºC until further analysis.

RNA extraction from the cell blocks was performed using also the MagMax FFPE DNA/RNA Ultra Kit, using the supernatant of the DNA binding to the beads. Unlike DNA, RNA binding to the beads requires isopropanol in addition to the binding buffer. After the binding is complete, the beads were washed using buffers containing isopropanol and treated with DNAse to remove DNA contamination. This treatment was proceeded by another round of washing steps, and finally the beads were eluted in the same elution buffer used for DNA.

9.2.1.3 Total nucleic acids extraction

9.2.1.3.1 Cell culture media

The different cell lines were seeded in T75 flaks and were allowed to adhere. Approximately 24 hours after seeding, the medium was renewed and the cells were left in culture for 48h. The medium was then separated and centrifuged at 3180g to remove dead cells and cellular debris. Amicon Ultra-15 mL Centrifugal Filters (Merck, Darmstadt, Germany) columns were used to concentrate the culture medium. This columns are specific for concentration of over 25x biological samples with dsDNA, To concentrate the samples, 15mL of culture medium were placed in the Amicon and centrifuged at maximum speed (3180g) for 70 minutes.

To extract cfTNA, the MagMAX™ Cell-Free Total Nucleic Acid Isolation Kit (Applied Biosystems) was used. Shortly, the samples were digested using protease K. Then, binding/lysis buffer and dynabeads were added to each sample according to the starting volume, and incubated with vigorous shaking to allow cfTNA binding to the magnetic beads. After the binding step, beads were pelleted using a magnetic stand and the supernatant was discarded. The beads were then washed using 80% ethanol and an isopropanol based wash buffer to remove impurities. The cfTNA were first eluted in elution buffer and then bond again to dynabeads for another round of extraction to increase purification. The purified cfTNA pool was the eluted in 15 µL of elution buffer.

9.2.1.3.2 Patient plasma samples

The plasma samples were thawed on ice and centrifuged at 16000g during 10 minutes to remove cellular debris. The protocol used for cfTNA extraction from cell culture medium was used, without further modifications.

9.2.1.3.3 EVs

To precipitate the EVs fraction present in plasma, the Total Exosome Isolation Kit (from plasma) (Invitrogen, Life technologies, Waltham, MA, USA) was used. Briefly, plasma was mixed with phosphate buffered saline (PBS) and digested with protease K at 37ºC for 10 minutes. Then, Total Exosome Isolation Reagent was added and the mix were incubated at 4ºC for 30 minutes. The mix was centrifuged at 10000g for 1 hour at 4ºC, and the supernatant was discarded. The pellet was then resuspended in PBS. Nanoparticle tracking analysis (NTA) was used for EVs quantifications, according with

the manufacturer recommendations. The cfTNA extraction protocol was then applied, like previously described.

9.2.1.3.4 Cell blocks

In order to extract TNA from FFPE material, we used 40 μM cell block sections. To do so, after dewaxing and dehydrating the sections, we applied the protocol previously used to extract cfTNA from culture media and plasma, increasing the digestion time to 16 hours, without any further modifications.

9.2.2 Nucleic acids quantification and qualification

DNA and RNA concentrations were assessed using Qubit® 2.0 Fluorometer (Invitrogen), with the double stranded DNA high sensitivity assay or the RNA high sensitivity assay, respectively. This methodology relies on the specificity (either for DNA or for RNA) of a fluorescent dye that only emits fluorescence when bound to target molecules, allowing the extrapolation of the quantification based on the emitted fluorescence.

9.2.2.1 cfDNA quantification

The cfDNA fraction was quantified using the 2200 TapeStation (Agilent Technologies, Santa Clara, CA, USA). This device is an automated system for capillary electrophoresis that uses a DNA intercalating agent with fluorescence that allows the detection and quantification of each individual band in the resulting electropherogram. We used the High sensitivity DNA assay, taking into account that it is designed to analyse DNA molecules with a size from 35 to 1000bp, with a sensitivity down to 5pg/μL. Briefly, DNA sample buffer and the DNA sample were mixed in a 1:1 ratio, mixed and then placed in the Tape Station instrument. The results were analyzed using the TapeStation Analysis Software (Agilent Technologies).

9.2.2.2 cfRNA quantification

The quantification of the cfRNA fraction was conducted by quantitative real time PCR (qPCR), using aTaqman probe (Hs99999903_m1, Applied Biosystems) specific for the ACTB housekeeping gene. This type of m1 probe spans an exon junction, where the probe and one of the primers sit within one exon, and therefore, do not detect the DNA that is also present in the cfTNA pool.

The qPCR reaction was set up using TaqMan™ Fast Advanced Master Mix (Applied Biosystems), the Hs99999903_m1 probe, nuclease free water and finally, the cfTNA. The Human HL60 Total RNA was used for the standard curve creation, with dilutions close in proximity to the expected concentrations of cfRNA in the test samples (780 pg/μL, 200 pg/μL, 50 pg/μL, 20 pg/μL, 10 pg/μL and 5 pg/μL). The following PCR conditions were used: 50ºC for 20 minutes, 95ºC for 2 minutes and then 40 cycles at 95ºC for 2 seconds and 60ºC for 1 minute. A non-template control was used with every plate and each sample was run in triplicate reactions.

To assess cfRNA integrity, the Bioanalyser 2100 (Agilent) with the Eukaryote total RNA Pico assay, according to the manufacturer recommendations.

9.2.3 Digital PCR

Digital PCR (dPCR) (Applied Biosystems) is a recently developed technique that uses TaqMan probes to detect specific sequences, either wild type or mutated with an extremely high sensitivity (122). The probe used for the EGFR mutation contains two different sequences, differently labelled. The WT sequence is linked to VIC fluorophore, whereas the mutated sequence is linked to FAM fluorophore, allowing an absolute quantification of the mutation allelic fraction (AF). On the other hand, the probes for the translocated sequences only contain the FAM fluorochrome, allowing only the detection of molecules expressing the fusion transcript. This approach was used to detect specific alterations, with certain probes for each mutation, namely the Hs04396941_ft (for SLC34A2-ROS1 fusion, present in HCC78 cells), Hs04396571_ft (for variant 3 of EML4-ALK fusion, described in H2228 cells) and Hs000000029_rm (for EGFR p.T790M mutation, present in H1975 cells).

To perform the dPCR reaction 8.7 μL of master mix, 0.87 μL of TaqMan assay (primer/probe mix), 1.83 μL of H2O and 6 μL of DNA sample (or cDNA for the fusion probes) are combined and loaded into a QuantStudio™ 3D Digital PCR chip using the QuantStudio™ 3D Digital PCR chip loader. The loaded chips were placed in the GeneAmp 9700 Thermal Cycler with Flat Adapter and the following program was applied: 96°C for 10 min, 40 cycles at 60°C (or 62ºC in the case of Hs000000029_rm probe) for 2 min and 98°C for 30 sec. QuantStudio™ 3D digital PCR instrument was used to read the chips and the results were analyzed with QuantStudio™ 3D AnalysisSuite™ Software.

9.2.4 Next generation sequencing

9.2.4.1 DNA sequencing

The Ion AmpliSeq Colon and Lung Cancer Research Panel v2 (Ion Torrent, Waltham, MA, USA) was used to detect alterations in DNA. This panel targets the most commonly altered regions of the most commonly mutated genes in both colon and lung cancer, such as KRAS, EGFR, BRAF, PIK3CA, ALK, MET and TP53, among others. Briefly 10ng of DNA were mixed in the PCR reaction and amplified with the following program: 99ºC for 2 minutes; 22 cycles of 99ºC for 15 seconds and 60ºC for 4minutes. Prior to barcode ligation, samples were digested with FuPa reagent to obtain blunt ends. Agencourt AMPure XP (Beckman Coulter, Brea, CA, USA) beads were used to clean the barcoded samples. The final library was quantified by qPCR with the Ion Library TaqMan® Quantitation Kit (Ion Torrent), according with the manufacturer instructions. Following quantification, template preparation was performed using the Ion Chef (Ion Torrent) and the loaded chips were then sequenced in a S5XL sequencer. The sequencing quality was assessed through the coverage analysis plugin and the samples were analysed with Ion Reporter 5.6.

9.2.4.2 RNA sequencing

The RNA was also sequenced with the Ion AmpliSeq RNA Fusion Lung Cancer Research Panel (Ion Torrent) to detect translocations and in genes such as ALK, RET, ROS 1 and NTRK1, being able to detect over 70 fusion transcripts. Briefly, 10ng of RNA were used for cDNA synthesis with the SuperScript VILO cDNA Synthesis Kit (Invitrogen). The cDNA was then combined in a PCR reaction and amplified according with the following program: 99ºC for 2 minutes; 30 cycles of 99ºC for 15 seconds and 60ºC for 4 minutes. After the PCR, the same protocol previously used for DNA sequencing was followed.

9.2.4.3 TNA sequencing

9.2.4.3.1 Conditioned media and plasma

The Oncomine Lung Cell-Free Total Nucleic Acid Research Assay (Ion Torrent) was used to detect genetic alterations in the clinical samples. This panel is based on a tag sequencing strategy, allowing the detection of variants down to an allelic frequency of 0,1% with an input of just 20 ng. The tag sequencing methodology uses molecular tags to gene specific primers and allow, after amplification, the tagged molecules to be

grouped based on tags. The analysis algorithm will only consider to be truly positive a variant that is present in over 80% of the tag grouped, removing false positives from random errors generated through the library construction or sequencing process.

Briefly, a minimum of 10ng of cfTNA were used for cDNA synthesis, using SuperScript VILO Master Mix (Invitrogen). Then, the DNA/cDNA pool was amplified, purified with Agencourt beads and linked to tag barcodes, to allow sample multiplexing upon sequencing. Agencourt beads were once again used for sample purification and size selection of the amplified library. The Ion Library TaqMan® Quantitation Kit was used to quantify the samples, Ion Chef was used for template preparation and Ion S5XL sequencer was used to sequence the samples. Ion Reporter 5.6 was used to analyze the results. The variants detected by NGS were confirmed by dPCR.

9.2.4.3.2 Cell blocks

From the FFPE material, the feasibility of combining the DNA and RNA panels previsouly in only one reaction was evaluated. We tried 3 different mixing ratios, namely:

- **1:1 ratio** (1x RNA Fusion Lung Cancer Research Panel: 1x Ion Colon and Lung Cancer Research Panel)
- **2:1 ratio** (1x RNA Fusion Lung Cancer Research Panel: 2x Ion Colon and Lung Cancer Research Panel)
- **Standard ratio** (1x RNA Fusion Lung Cancer Research Panel: 2,5x Ion Colon and Lung Cancer Research Panel) (this last condition represents the volume described for each panel when used separately).

For all conditions, cDNA was synthetized like previously stated and amplified according to the protocol used for the Ion AmpliSeq RNA Fusion Lung Cancer Research Panel.

10. Results

10.1 Cell line characterization

We confirmed the legitimacy of the H2228 and HCC78 cell lines by analyzing STR profile and comparing with the data available on ATCC website (Table 4), thus validating the cells to be used in downstream applications. Moreover, we tested the presence of the alterations that were previously described for the cells. By dPCR we confirmed that H2228 harboured an EML4-ALK variant 3 translocation (Figure 7A) and that HCC78 cells held and SCLC4A32-ROS1 fusion (Figure 7B). For the H1975 cells, we confirmed that both the EGFR p. L858R and p.T790M mutations were present (Figure 7C and 7D, respectively).

Table 4: Autosomal STR DNA Profile from HCC78 and H2228 cell lines. Results analysed with POWERPLEX 16 HS kit (Promega). If allele imbalance is observed, the minor allele is represented with (-) when its proportion is less than 50% of the major allele; minor alleles with less than 20% signal (but over 10%) relative to the major allele are marked with "+".

		Cell lines					
		H2228		HCC78		H1975	
		Expected	Observed	Expected	Observed	Expected	Observed
STRs	D3S1358		16-18 (+17)	16-17	17-17	14-15	17 (+15, +16)
	TH01	7-8	7-8		7-9.3		6
	D21S11		28-30 (+29)	29-30	28-30	28	31(+30)
	D18S51		16	12-16	12-16	13	13
	Penta E		11-16	11-12	11-12	12-16	12-20
	D13S317	11	12	8-12	8-12	9	9
	D5S818		11	11-12	11-12	11-12	12
	D7S820	11	11	9	9	8-11	11-12
	D16S539	11-13	11-13	11-13	11-13 (+12)	9-12	12
	CSF1PO	12	12	11-12	11-12(-)	12	11
	Penta D		11-13	7-11	7 (-) -11	12-13	14
	Amel	XX	XX	XY	XY	XX	XX
	vWA	15-17	15-17	15-18	15-18 (+17)	18	18(+17)
	D8S1179		13	13-14	13-14	13-16	12
	TPOX	11	11	8-11	8-11	8-11	8(+11)
	FGA		21-22	21-22	21-22	21-24	22-23

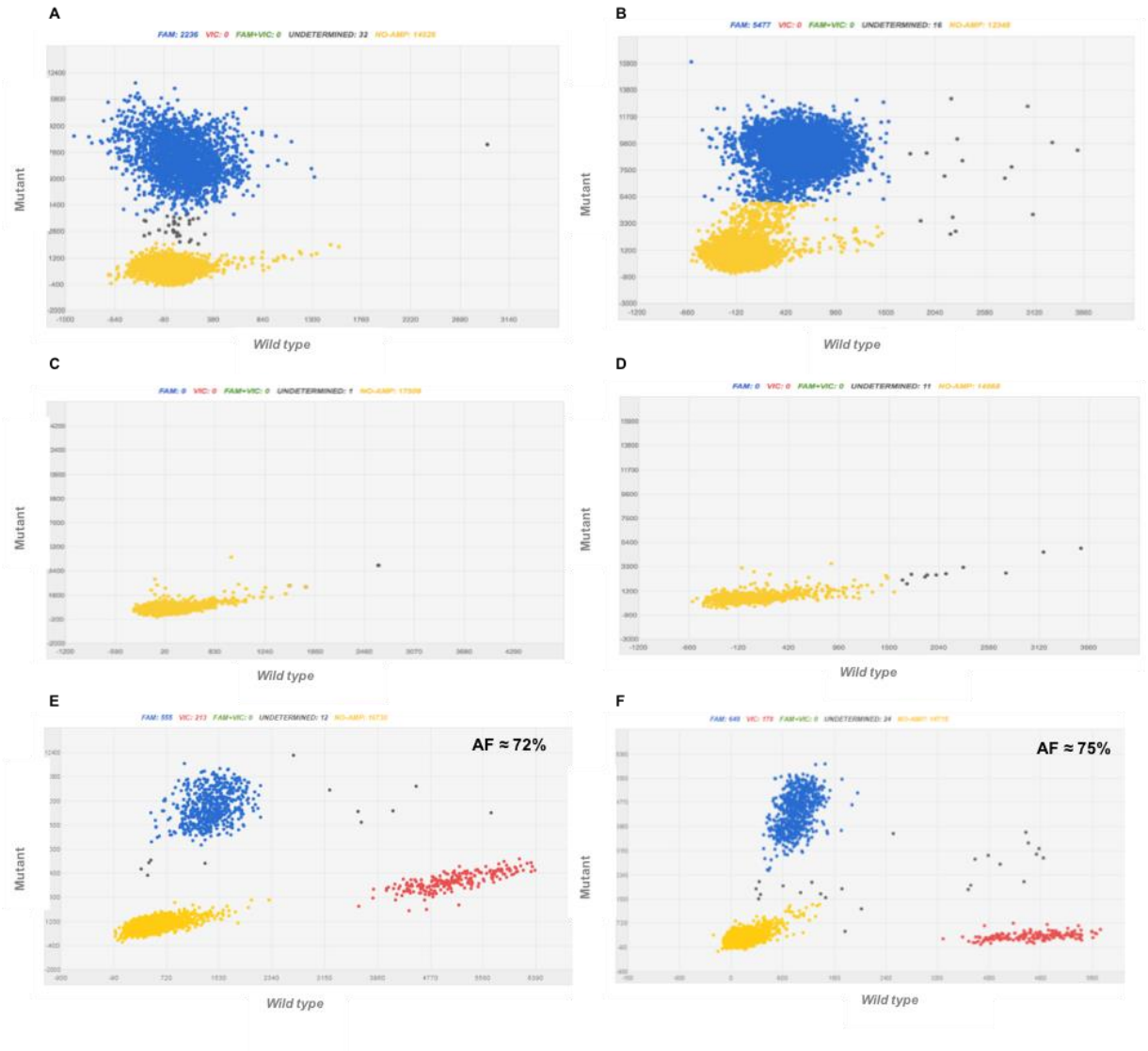

Figure 7: dPCR plots for EML4-ALK, SLC34A2-ROS1, EGFR p.T790M and p.L858R probes. The yellow dots represent the empty wells, the blue dots represent molecules that express the mutant variant, the red dots correspond to molecules expressing the wild-type sequence and the grey dots represent polyclonal wells and are not considered for analysis. A) H2228 RNA positive for the EML4-ALK fusions transcript. B) HCC78 RNA positive for the SLC34A2-ROS1 rearrangement. C and D) negative control for EML4-ALK and SLC34A2 with DNA from H228 and HCC78 cells respectively. E) H1975 cells DNA tested for the presence of the EGFR p.T790M mutation, present at 72% AF. F) H1975 cells DNA tested for the presence of the EGFR p.L858R mutation, present with 75% AF. Note: The AF is calculated dividing the number of blue dots by the red dots, then multiplied by 100.

10.2 cfTNA approach optimization

10.2.1 cfTNA extraction extraction

The optimization of the cfTNA extraction method was confirmed, assessing the yield of retrieved cfTNA from the cell culture media and also the presence of the somatic alterations of the cell lines. It was verified that the size distribution pattern of the cfDNA

fraction within the cfTNA pool was very similar to what was expected for cfDNA, with the majority of fragments corresponding to a peak of around 180bp (Figure 8A. As a control, we analyzed the size distribution pattern of the Horizon synthetic cfTNA samples was verified that the pattern is very similar between cfTNA from cell culture media and the control sample (Figure 8B).

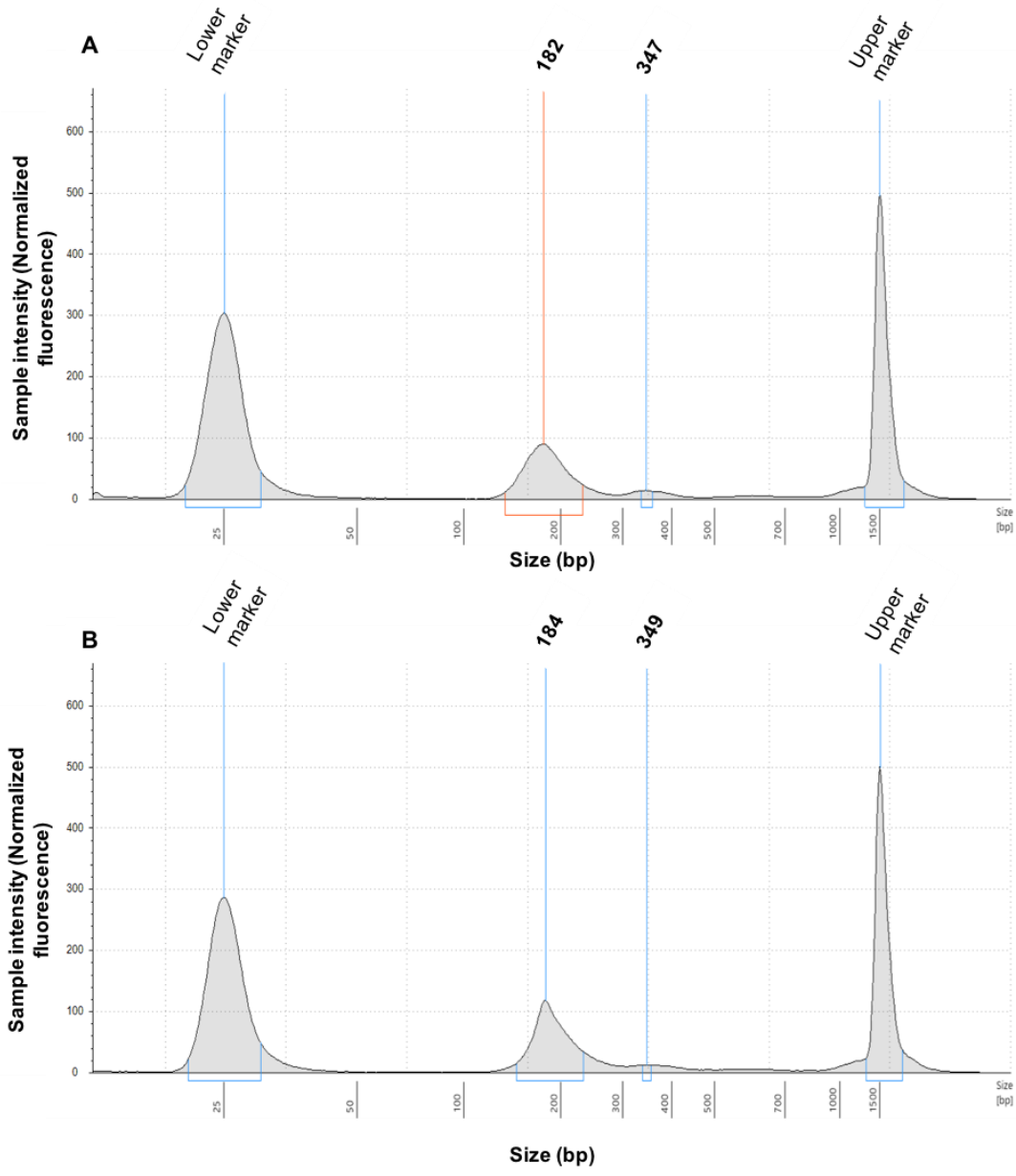

Figure 8: **cfDNA fraction size distribution pattern, assessed by capillary electrophoresis.** The x axis represents the size in bp, while the y axis is representative of the fluorescence intensity, proportional to the amount of DNA present. A) Representative image of cfDNA size distribution pattern from cell culture media, showing a major peak at around 180bp. B) Representative image of cfDNA from the synthetic controls displaying the same pattern of size distribution.

It was possible to detect the fusions transcripts, distinctive for H2228 and HCC78 cell lines (Figure 9A and 9B) in cfTNA. Moreover, we confirmed the presence of the *EGFR* p.T790M mutation in the DNA fraction of the cfTNA pool from H1975 cells, with an allelic frequency very similar to the one that is reported for the alteration in the cells themselves, around 80% (Figure 9C).

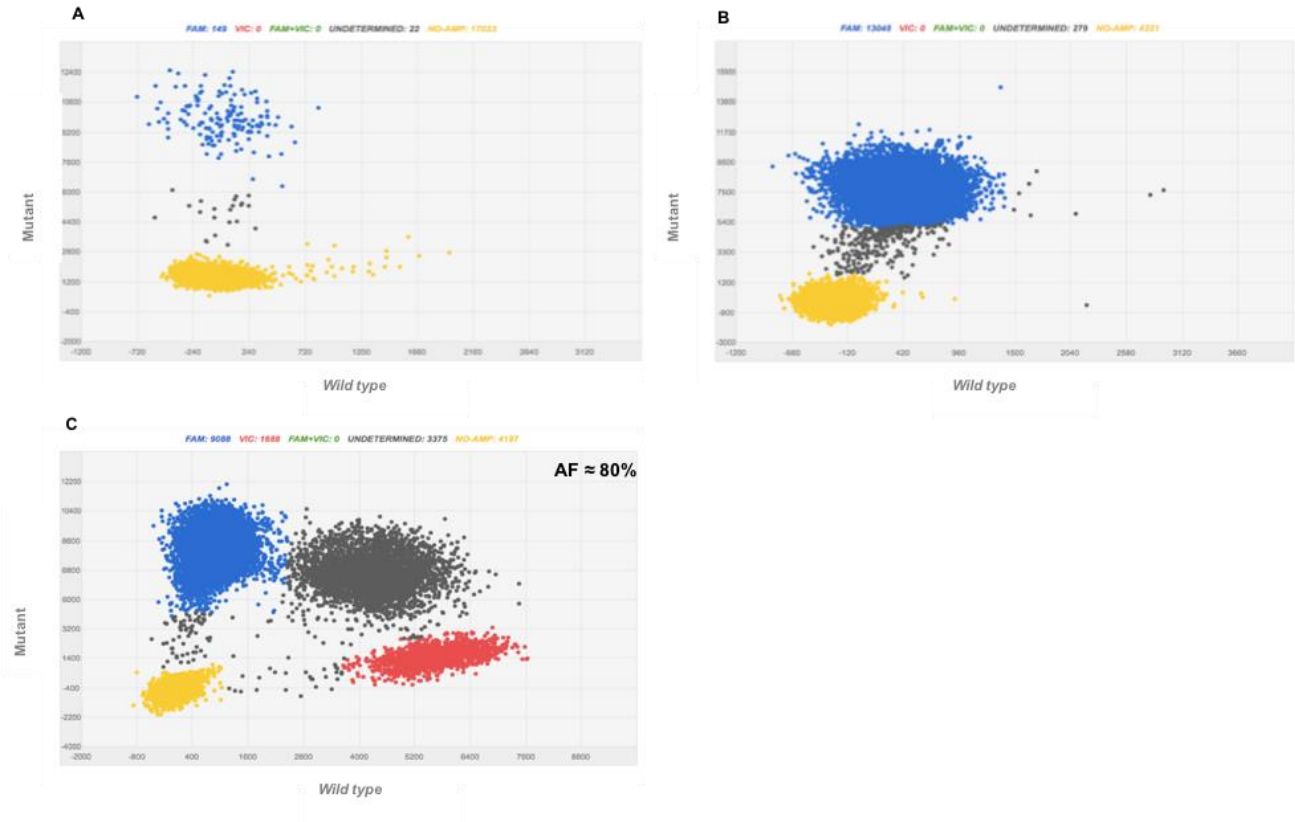

Figure 9: dPCR plots displaying the presence of somatic alterations in cfTNA from conditioned media. The presence of the EML4-ALK v3 is showed in 9A by the blue dots, as the occurrence of the SLC34A2-ROS1 is displayed in 9B. In 9C we detected an AF of 80%, which matched the AF present in the cell line.

10.2.2 cfTNA sequencing optimization

The HorizonDx synthetic controls were sequenced in order to assess the validity of the proposed approach. It was observed that the Oncomine Lung Cell-Free Total Nucleic Acid Research Assay allowed the detection of copy number variations (CNVs), single nucleotide variations (SNVs) and genetic rearrangements, as well as different genetic mutations (Table 4). The distinction between different *MET* copy number variations was also possible, allowing to separate the amplification level between the low copy control (1.32 CNV ratio) and the high copy control (1.74 CNV ratio). It was also possible to detect the different SNV present on each control: In the low copy control it was possible to detect mutation in the NRAS and KRAS, at 0.05% and 1.94% AF, respectively.

Additionally, two different mutations affecting *TP53* gene were detected, at 2.98% and 1.39% for the p.R248L and p.G245C, correspondingly.

Concerning the fusion transcript detection, we were able to detect translocations affecting *ALK* (two different fusion variants), *ROS1* (two different fusion variants as well), RET and MET. The results for the high copy control were comparable, except for the impossibility to detect the NRAS, differing also in the AF of the KRAS p.G12A mutation (5.8% in the high copy control) and of the TP53 mutations (6.07% for the p.R248L and 1.99% for the p.G245C). Regarding the SNV control- as expected- only point mutations and deletions were detected, with AF that ranged from 0.06% to 0.17%. No further alterations were found on this control sample, concerning whether copy number variations or fusions. Lastly, for the negative control, no alterations were detected at neither the RNA nor DNA level.

10.2.3 Patient plasma samples

Having the extraction panel optimized for in vitro conditions we moved to the analysis of cfTNA extracted from plasma samples. Additional optimizations steps had to be addressed, namely regarding plasma storage conditions. Therefore, in order to confirm if: a) whole plasma was the right biological material to extract cfTNA from and b) the storage of the plasma was affecting cfRNA retrieval, we used a plasma sample processed almost immediately after collection.

From part of that plasma sample, exosomes were isolated in order to evaluate if extracellular vesicles represented a better source for cfRNA. Also, during cfTNA extraction we increased the agitation during cfTNA binding to the dynabeads to the maximum. Plus, prior to cDNA synthesis we denatured the RNA at 80ºC for 10 minutes. The first observation was that the cfRNA yield from plasma as substantially higher than any other frozen sample previously tested (Figure 10).

Table 5: Synthetic control sequencing results

Control	SNV/Indel				CNV				Fusions (variants)	
	Gene	AA change	MAF %		Gene	Gain/Loss	CNV Ratio			
			Expected	Observed			Expected	Observed	Expected	Observed
Low Copy Control	NRAS	p.Q61K	NA	0,05	MET	↑	1,42	1.32	EML4-ALK.E6bA20.AB374362	EML4-ALK.E6bA20.AB374362
	KRAS	P.G12A		1,94					SLC34A2-ROS1.S4R32.COSF1196	SLC34A2-ROS1.S4R32.COSF1196
	TP53	p.R248L		2,98					SLC34A2-ROS1.S4R34.COSF1198	SLC34A2-ROS1.S4R34.COSF1198
	TP53	p.G245C		1,39					CCDC6-RET.C1R12.COSF1271	CCDC6-RET.C1R12.COSF1271
									EML4-ALK.E6aA20.AB374361	EML4-ALK.E6aA20.AB374361
									MET-MET.M13M15	MET-MET.M13M15
High Copy Control	KRAS	P.G12A	NA	5,8	MET	↑	1,72	1,74	EML4-ALK.E6bA20.AB374362	EML4-ALK.E6bA20.AB374362
	TP53	p.R248L		6,07					SLC34A2-ROS1.S4R32.COSF1196	SLC34A2-ROS1.S4R32.COSF1196
	TP53	p.G245C		1,99					MET-MET.M13M15	MET-MET.M13M15
									CCDC6-RET.C1R12.COSF1271	CCDC6-RET.C1R12.COSF1271
									SLC34A2-ROS1.S4R34.COSF1198	SLC34A2-ROS1.S4R34.COSF1198
SNV Control	NRAS	p.G12D	0,922	0,17	ND	ND	ND	ND	ND	ND
	ALK	p.R1275Q	0,628	0,06						
	PIK3CA	p.H1047R	0,548	0,14						
	BRAF	p.V600E	0,584	0,13						
	KRAS	p.Q61H	0,887	0,14						
	KRAS	p.G12D	0,609	0,1						
	TP53	p.R273H	0,482	0,16						
	TP53	p.R248Q	0,471	0,1						
	TP53	p.Y234C	0,471	0,1						
	TP53	p.Y220C	0,471	0,06						
	ERBB2	p.Glu770_Ala771insAlaTyrValMet	0,471	0,13						
Negative Control	ND				ND				ND	ND

ND: Not detected; NA: not applicable

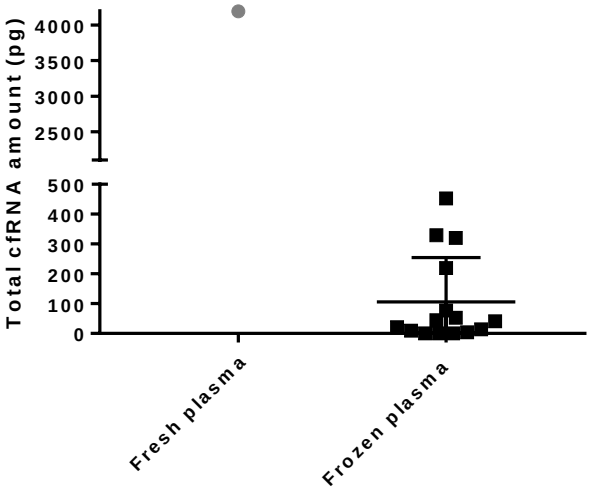

Figure 10: cfRNA yield comparison between a fresh and frozen plasma samples. Only one fresh plasma sample was used for comparison, while the results of the frozen samples are presented mean ± standard deviation.

The integrity of DNA and RNA was also evaluated, according with the size distribution pattern. TapeStation analysis as previously described assessed the cfDNA size distribution and no changes were observed between peak sizes regarding cfDNA from exosomes or from whole plasma (Figure 11 A and 11 B). A significant decrease in cfDNA from exosomes was detected. The cfRNA size distribution pattern was assessed by Bioanalyser analysis and showed that the RNA fragments are much smaller than the ones present in whole plasma (Figure 11 C and 11D). Additionally, the results demonstrated that both cfDNA and cfRNA amounts are significantly higher when using whole plasma as a source, comparing to exosomes (Figure 12A). Furthermore, upon sequencing of both samples, we noted that cfTNA from exosomes resulted in lower number of sequencing (Figure 12B), higher limit of detection (LOD) (Figure 12C) and lower number of reads detecting the mutant fusion transcript (therefore mentioned molecular coverage) towards the detection of the EML4-ALK (Figure 12D). Nevertheless- for both samples it was possible to detect the genetic rearrangement detected in the tumour biopsy.

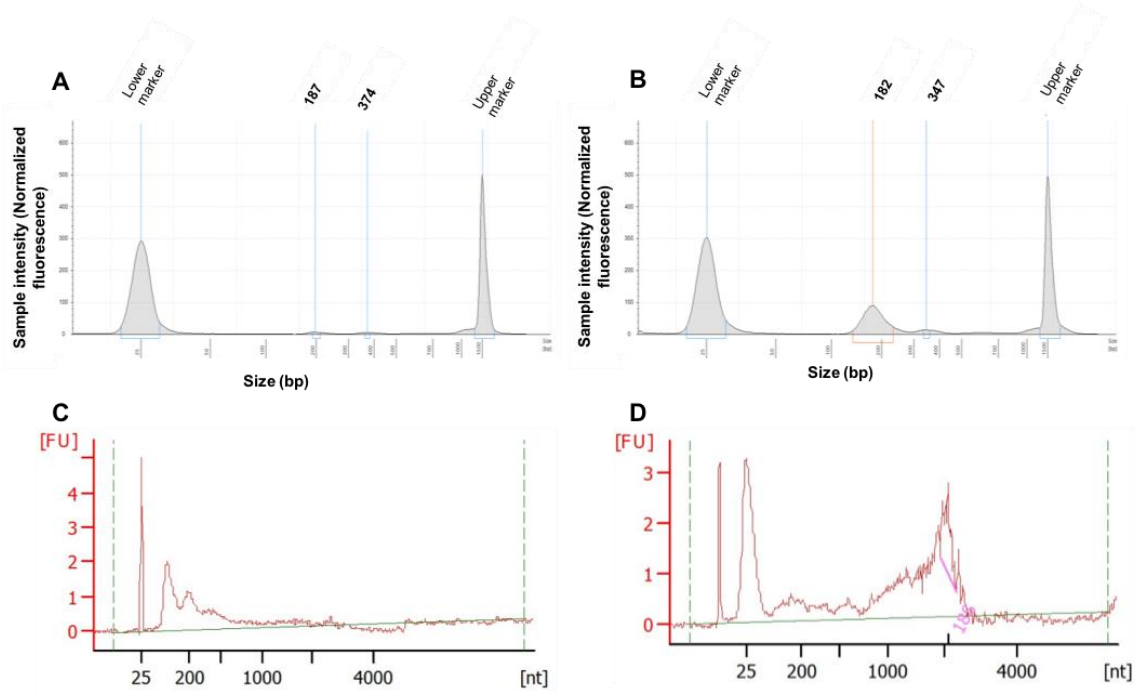

Figure 11: cfDNA and cfRNA size distribution pattern. A) cfDNA from exosomes, showing a major concentration around 180 bp. B) cfDNA from whole plasma. C) Bioanalyser profile of cfRNA from exosomes. D) Bioanalyser profile of cfRNA from whole plasma.

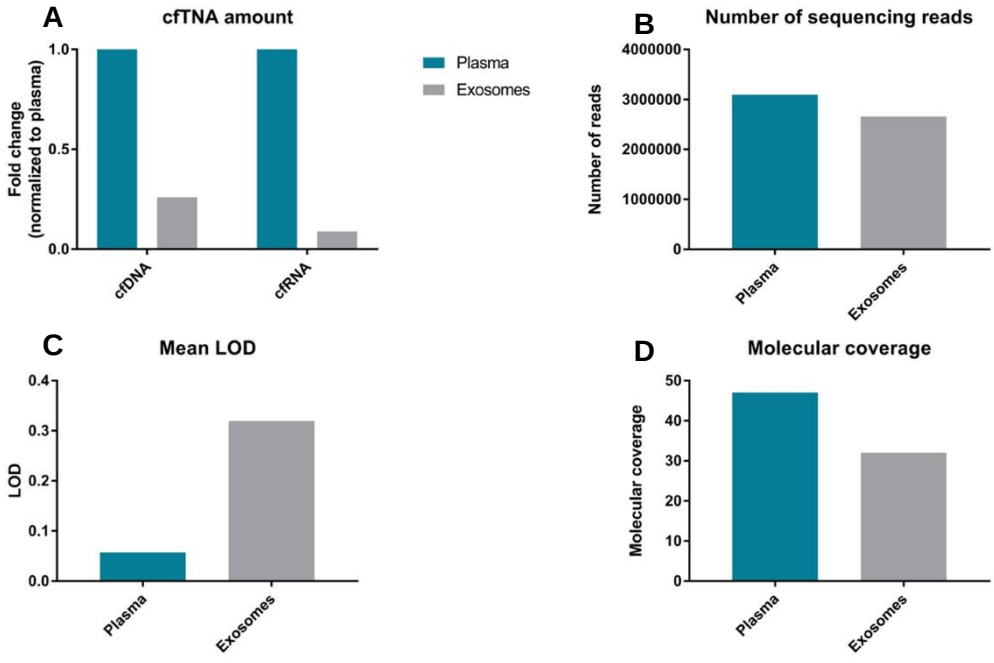

Figure 12: cfTNA in whole plasma *versus* exosomes. A) cfDNA and cfRNA amount in exosomes normalized to the amount retrieved from whole plasma, demonstrating a clear yield decrease on exosomes. B) Number of sequencing reads obtained with the Oncomine Lung Cell-Free Total Nucleic Acid Research Assay for plasma and exomes, whereas whole plasma resulted in a superior number of total sequencing reads. C) Mean LOD obtained for plasma (0,057%) and exosomes (0,32%). D) Molecular coverage for the mutant fusion transcript for both plasma (47 reads) and exosomes (32 reads).

10.3 Clinical application of the cfTNA approach

We moved to the comprehensive follow up of one patient throughout disease course applying the protocol modifications previously optimized, such as the vigorous shaking upon binding to the dynabeads and the denaturation step prior to cDNA synthesis. The presence of an EML4-ALK fusion was confirmed by FISH and a concomitant *KRAS* p.G12V mutation was detected through DNA sequencing. The treatment regimen started with first line pemetrexed for 10 cycles- followed by targeted therapy with TKI, namely crizotinib, ceritinib and alectinib. The complete treatment regimen is described in Figure 13.

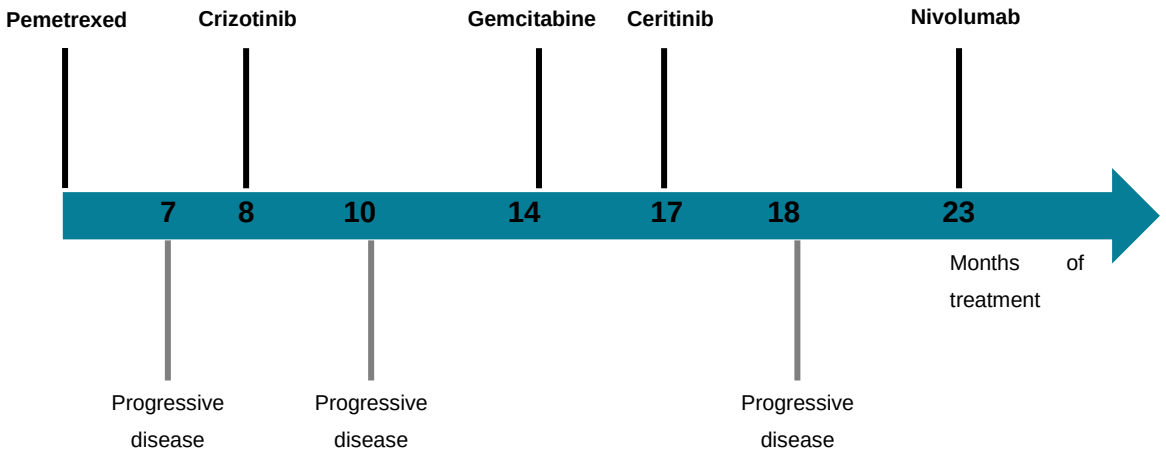

Figure 13: **Disease course of the patient selected for cfTNA based monitoring.**

The data obtained regarding the follow up of this patient was compiled and separated into 2 groups: one evaluated the variation of cfTNA amount in circulation throughout disease course and another one analysing the genetic alterations detected, either on DNA or RNA. We observed that the quantity of cfDNA and cfRNA in circulation changes significantly over time. The increase in cfDNA/cfRNA amount appears to be related with periods were clinical evidence of disease progression were identified (Figure 14). It was also possible to observe the similarity between cfDNA and cfRNA quantity, that despite the different scaling units ([cfDNA] is presented in ng/mL of plasma whereas [cfRNA] is presented in pg/mL of plasma), follow very a very analogous distribution pattern, with the exception of month 25 of treatment.

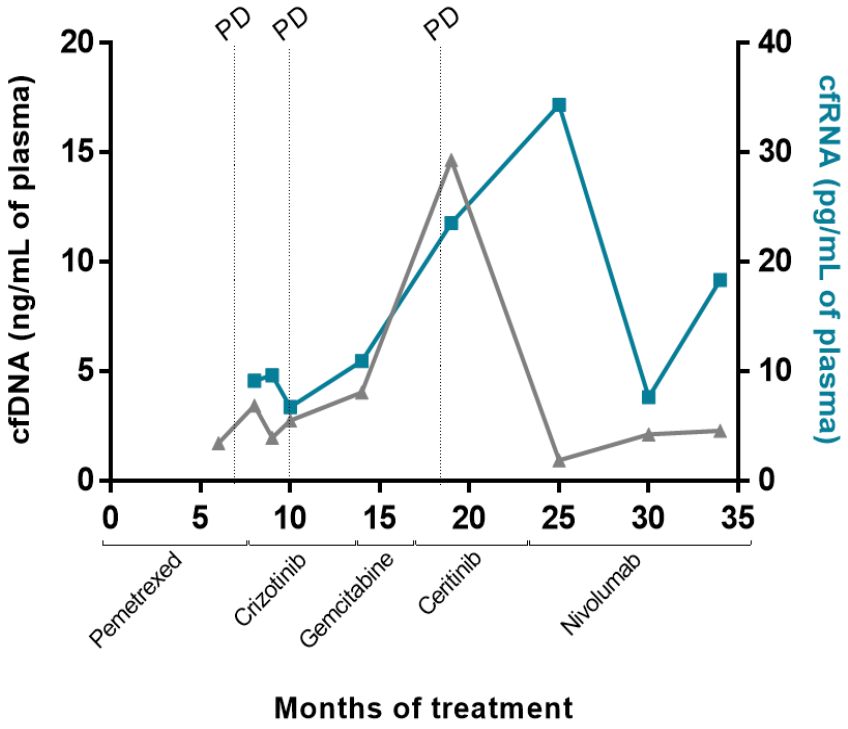

Figure 14: **Patient variation of overall cfTNA amount during the course of disease.**

Regarding the genetic alterations revealed with cfTNA sequencing, it is possible to replicate the atypical disease course presented by the patient, represented by the number of genetic alterations discovered but also their dynamic changes with time (Figure 15). At month 5 of treatment is was not possible to detect the EML4-ALK rearrangement, but instead, we detected a deletion in the *EGFR* gene, the p.E746_A750 del, with an AF of 2.5%, during treatment with pemetrexed-based chemotherapy. At month 8, through crizotinib treatment, the fusion transcript was still not detected but, at the DNA level, it was possible to detect the *KRAS* p.G12V mutation (previously detected on the tissue biopsy), with an AF of 0.36%. Moreover, two *EGFR* mutations were also detected, namely the formerly identified p.E746_A750 deletion and the p.T790M mutation, at 0.1% and 1.1% of AF, respectively. Around the nineth month of treatment, none of the *EGFR* mutations were detected. Only the *KRAS* point mutations was present at 0.5% AF. At month 10 we were only able to detect the *KRAS* p.G12V mutation and the *EGFR* p.T790M mutation, with an AF of 0.4% and 5.8%, correspondingly. Fourteen months after diagnosis, and during crizotinib treatment, only the *EGFR* p.T790M mutation (0.9% AF) and the *KRAS* p.G12V (0.9% AF) mutations were detected. We then evaluated the sample collected during month 19, where only the *KRAS* mutation was present, with an AF of 0.2%, which was also the case for the sample corresponding to

month 21, where the *KRAS* p.G12V mutation was detected with an AF of 0.4%. On month 25, during nivolumab treatment, it was possible to detect both *EGFR* mutations, namely the p.E746_A750 that was present at 0.5% AF and the p.T790M- detected with an AF of 0.3%. Similarly with the previous samples evaluated, we also detected the *KRAS* p.G12V mutation, with an AF of 1.4%. Likewise, we detected the EML4-ALK fusion transcript for the first time during this patient's course of disease. On month 30 after diagnosis, collected while the patient was still under nivolumab treatment, we could detect the EML4-ALK rearrangement, as well as the *KRAS* mutation with 2.1% AF. The last sample studied for this patient was collect during month 34 of treatment, and it was only possible to detect the *KRAS* mutation.

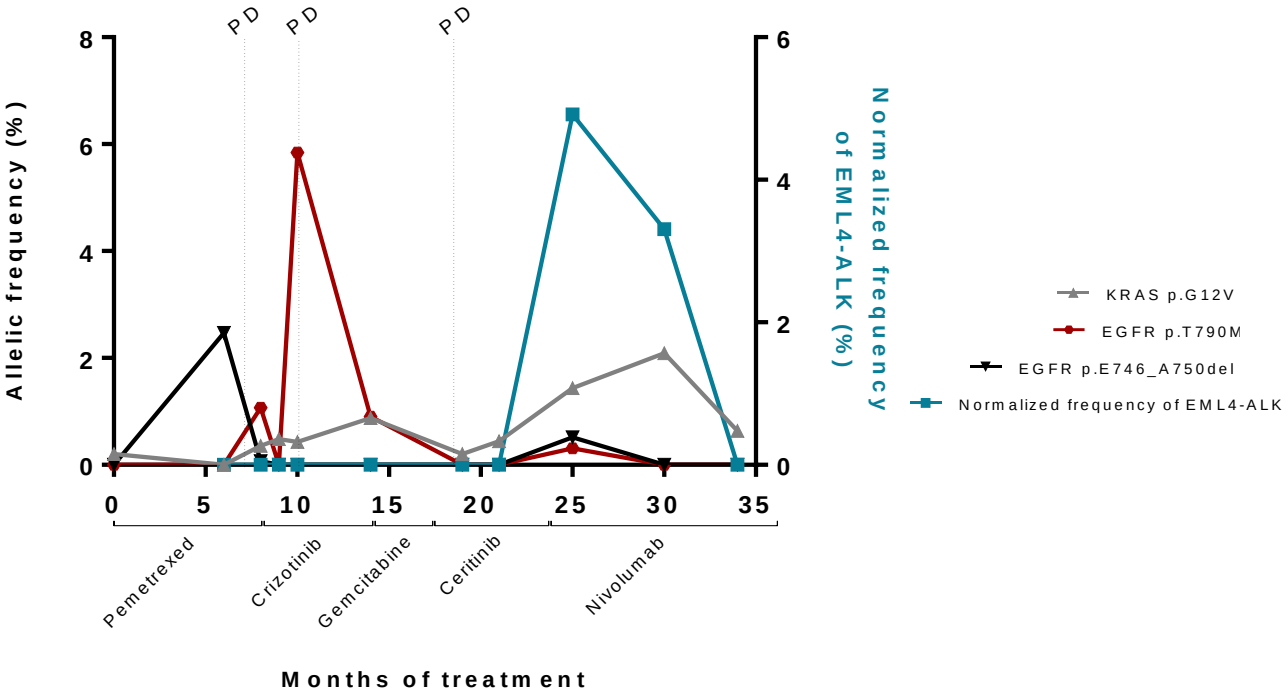

Figure 15: Genetic alterations detected with cfTNA sequencing. The AF of the mutations was extrapolated directly from the sequencing data. Normalized frequency of the EML4-ALK transcript was calculated dividing the mutant molecular coverage of the fusion transcript by the sum of mutant molecular reads of the internal controls.

10.4 TNA optimization in tissue blocks

Following TNA extraction from FFPE and sequencing, the results from the 3 conditions tested were analyzed and compared with the reference samples, that is, the results from DNA and RNA from H2228 cells sequenced separately (therefore mentioned as reference conditions), with the Colon and Lung and Fusion panel, the currently accepted gold standard methodology.

Regarding the RNA level, we took into account the expression of the EML4-ALK transcript throughout the different ratios (Figure 16). We observed that with the 3 ratios tested it was possible to detect the translocation, but the standard ratio resulted in the closest similarity of expression.

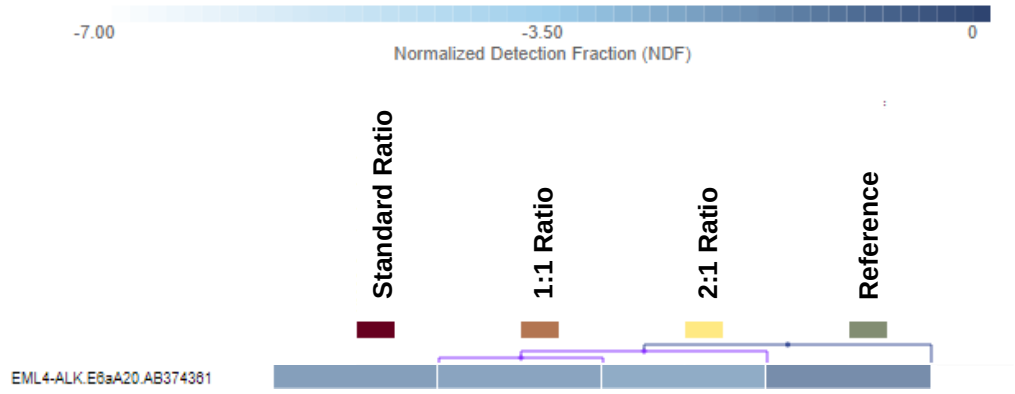

Figure 16: EML4-ALK fusion transcript expression for every condition tested. The standard ratio condition resulted in the most similar expression level, comparing to the reference conditions, followed by the 1:1 ratio and lastly by the 2:1 ratio. The detection of the fusion transcript was possible with all the conditions tested.

Nonetheless, in order to assess the sequencing quality, we also evaluated the coverage of internal controls, namely the coverage of housekeeping genes such as Hydroxymethylbilane Synthase (*HBMS*), Integrin Subunit Beta 7 (*ITGB7*), Lamin A/C (*LMNA*), MYC proto-oncogene (*MYC*) and Small heat shock protein IBP (*IBP*), that is directly proportional to sequencing quality. We observed that the 1:1 ratio was the one that yielded the most similar results, comparing with the reference conditions (Figure 17). The coverage of *HMBS* and *MYC* was very similar throughout all conditions, which was not the case for *ITGB7*, where the standard ratio result in no coverage of the housekeeping gene and in a very low coverage for the 2:1 ratio. The *LMNA* gene resulted in the biggest discrepancies, with the 1:1 ratio resulting in approximately half the coverage, followed by standard ratio and by the 2:1 ratio at last. Regarding *IBP*, the 1:1 ratio resulted in a slightly superior coverage than the one obtained for the reference conditions. The standard ratio and the 2:1 ratio resulted in similar values for this control, however, both considerably below the reference values.

In order to evaluate the impact of the merged panels on DNA sequencing we used amplicon coverage as a metric for sequencing quality. To evaluate this parameter we normalized the coverage of each condition to the total number of reads obtained. The normalized coverage for each condition was then compared individually with the reference conditions (Figure 18). We observed that every condition altered amplicon

coverage, increasing the dispersion of the relative coverage. For the standard ratio, the difference was statistically significant, indicating major changes in the overall sequencing coverage. The 2:1 ratio resulted in no statistically significant differences but in very dramatic coverage changes for some of the amplicons. The 1:1 ratio was the one that yield closest results to the reference conditions- which is why it was chosen as the ratio to be used on the validation test with human FFPE samples.

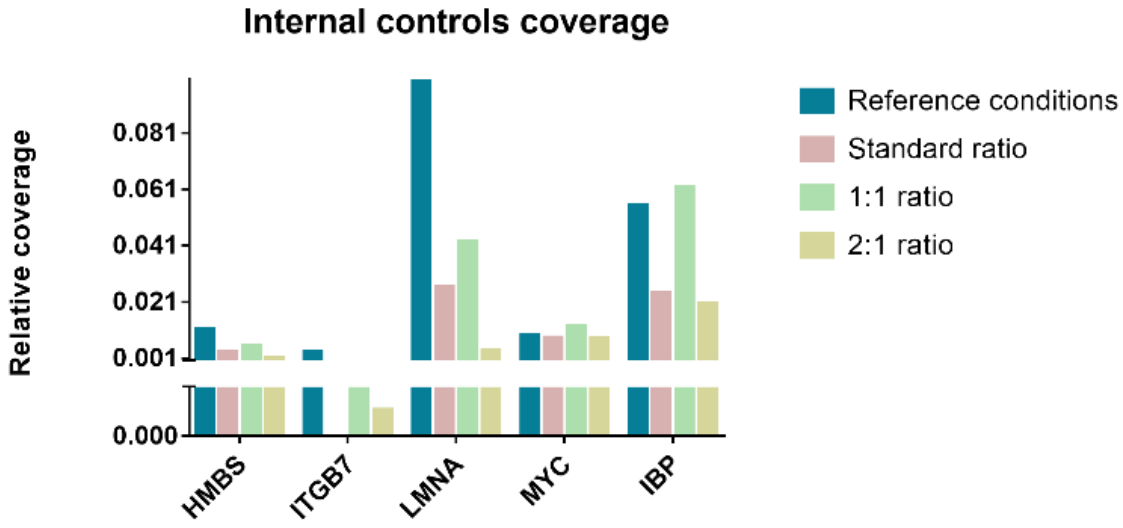

Figure 17: Internal controls normalized coverage comparison. Regarding RNA sequencing quality, the 1:1 ratio was the one that yield closest values comparing with the reference conditions. To normalize the results, the number of reads for each internal control was divided by the total number of reads obtained for each sample.

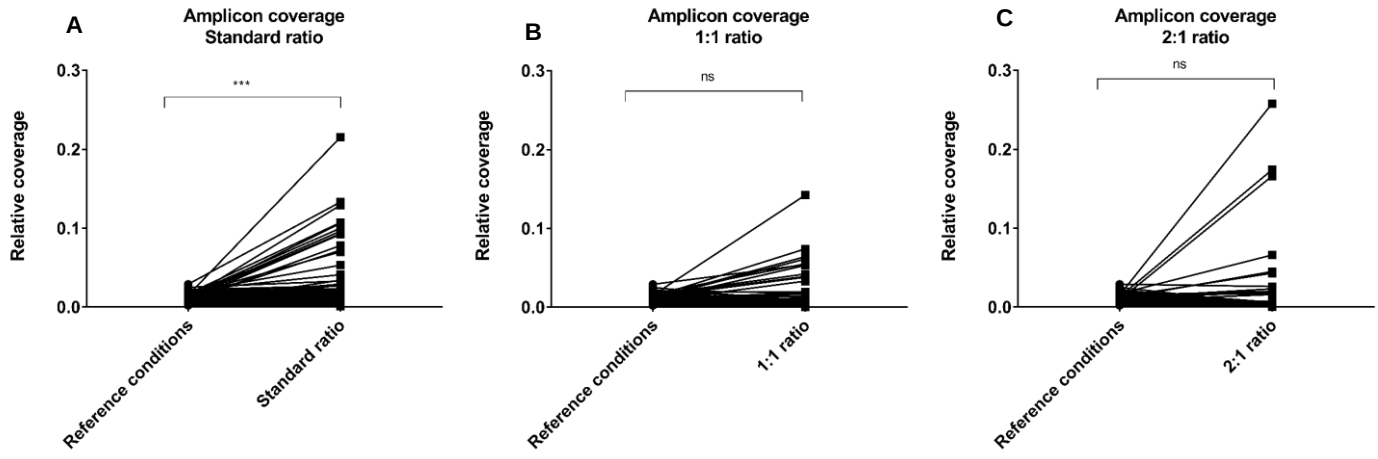

Figure 18: Normalized amplicon coverage compared with reference conditions. A) Amplicon coverage for standard ratio, which resulted in a statistically significant difference ($p<0.0001$) according with the Dunnett's multiple comparisons test. B) Amplicon coverage for the 1:1 ratio, where minor difference are observed comparing the individual coverage of each amplicon with the reference conditions. C) Amplicon coverage for the 2:1 ratio, which result in overall similarity with 3 very discrepant outliers, not statistically significant.

We were able to further validate the results from this experiment in one FFPE sample- from a lung cancer patient, where it was possible to detect a previously unknown *ALK* p.T1151M mutation- with an AF of 3.30%.

11. Discussion

The liquid biopsy approach has revolutionized the way NSCLC patients are monitored throughout the course of disease- sparing the use of invasive procedures such as tissue biopsy to acknowledge the molecular landscape of the tumour. The lack of methodologies that allow the evaluation of cfRNA has limited the use of this non-invasive approach to patients whose tumours are driven by genetic rearrangements, due to the requirement to evaluate this type of alteration at the RNA level. The possibility to study gene rearrangements and mutations from a single sample, in only one sequencing reaction has the potential to expand the clinical applications of liquid biopsy to other cohorts of patients, and not only the ones whose tumours harbour point mutations or deletions. Therefore, the need for the development of an approach feasible to be applied to this subset of patients is highly necessary. Taking that into account, the main objective of this project was to develop a liquid biopsy approach, based on the analysis of total nucleic acids, to allow the <monitoring of NSCLC patients with gene rearrangements in a non-invasive manner, enabling the detection of the driver translocation and the potential rise of resistance mutations.

Our first approach was based on *in vitro* studies, where three different cell lines were used, each one representing a different model of disease with potential interest for the future parts of the work. By using conditioned media from cell culture as a surrogate of plasma samples, we conducted in vitro liquid biopsy tests to assess the efficiency of the selected cfTNA extraction method. We analyzed the size distribution pattern of cfDNA, which showed fragments of mainly 180 bp, corresponding to the expected size of cfDNA and verified that cfRNA was also present by qPCR quantification. The extraction method was further validated by the confirmation that it was possible to detect the tumour derived somatic alterations (either fusion transcripts or point mutations) in cfTNA from cell culture medium. With this set of results we verified that the extraction method was efficient in the retrieval of both cfRNA and cfDNA enabling the detection of different types of genetic alterations.

The application of the approach in plasma samples from NSCLC patients was preceded by the analysis of synthetic cfTNA samples with the Oncomine Lung Cell-Free Total Nucleic Acid Research Assay. This validation played a very critical role in validating the panel as a tool efficient to distinguish between different types of alterations- some of which present at very low AF, close to the limit of detection of the technique. The results

of the four controls applied were comparable to what was expected. Moreover, the variety of alterations detected, from translocations to SNVs, assured to be possible to detect different scenarios upon application on clinical samples. Unfortunately, upon the first tests performed with cfTNA of plasma samples collected before the start of treatment, it was not possible to detect the driver rearrangement. Considering that these samples were naïve to any type of therapy and that they belong to stage IV NSCLC patients, we expected to be able to detect the fusion transcript. We thought that the problem could rely on the cfRNA amount that was significantly below to what was expected to obtain from a plasma sample. Moreover, most studies conducted on liquid biopsy use blood collection tubes with preservatives, such as the Streck tubes, which could allow the protection of the cfRNA upon freezing (123). The tubes used contained no fixative whatsoever, which could potentiate RNA degradation during storage. Also- the samples used for this study were frozen at -20ºC and not at -80ºC as it is described for RNA preservation. The impact of freezing was assessed by evaluating a fresh plasma sample. We verified that the retrieved amount of both cfDNA and cfRNA was significantly higher than every other sample we previously tested, indicating that the storage conditions, either the collection tubes or the freezing temperature were impairing the retrieval of cfTNA from the plasma samples. Moreover, because cfRNA is allegedly very susceptible to nuclease degradation and that shielding within extracellular vesicles, such as exosomes, could represent a protection mechanism, we analyzed cfTNA extracted from the fresh plasma sample (108- 111). We observed that the yield of both cfRNA and cfDNA is significantly lower in EVs than in whole plasma. This indicates that the cfRNA present in EVs does not represent a substantial portion of the overall cfRNA in circulation. Therefore, our results suggest that whole plasma represent the most suitable source to extract cfRNA, as it is already extensively described for cfDNA (124). Moreover, by analyzing the alterations to the protocol and their impact on cfRNA yield and ability to detect translocations, we were able to test that the affinity of the dynabeads towards cfRNA is enhanced with mechanical agitation, with the retrieved yield being proportional to the shaking speed. Additionally, the detection of fusions transcripts was only possible after we implemented a step of RNA denaturation. This fact made us assumed that cfRNA in circulation forms complex structures that impair the accessibility to the RNA molecules. Correlating this data with the one we previously obtained that the majority of cfRNA is not protected in EVs, we thought that the formation of these complex RNA conformations or even the association of RNA molecules with protein complexes could explain the survival of cfRNA in circulation.

With all the experimental conditions optimized, we analyzed the outcome of applying cfTNA monitoring throughout treatment. The patient used for cfTNA disease monitoring has a rather unusual disease course, with several changes in therapy due to lack of response presented as progressive disease. With cfTNA analysis, the singular disease progression observed in the clinical was also evident in the cfTNA analysis. Observing the cfDNA and cfRNA amount in circulation, a correlation with tumour burden was observed. Upon detection of clinical progression, as it is the case for month 10, it is possible to observe a preceding increase in the overall cfTNA quantity, which is also the case for month 17, where very significant increases in cfRNA and cfDNA were detected.

We also assessed the genetic alterations present in circulation in different time points- in order to obtain a clearer correlation with clinical data. However, the different types of alterations that were found made the interpretation of their biological significance complex. We suggest two different possibilities for the concomitant occurrence of the EML4-ALK translocation, *KRAS* p.G12V mutation and the *EGFR* p.E746_A750del and p.T790M mutations: the first one possibility is that the patients harbours more than one primary tumour, with different driver mutations. This is supported with the fact that both the EML4-ALK fusion and *KRAS* mutation were detected in the tissue biopsy, which was not the case for both EGFR mutations, even after the application of highly sensitive techniques such as dPCR. The *EGFR* mutations could be derived from another primary that was interpreted as a lung metastasis. The other possibility is that the primary tumour is highly heterogeneous. The different mutations could all be present at diagnosis but due to the lack of selective therapy exerted by therapy, no clonal expansion had yet occurred, making impossible to detect the alterations in the primary tumour.

We also intended to verify if the TNA approach was feasible to be applied to FFPE material, considering that tissue biopsies are still the gold standard material to perform genetic testing. By sequencing the RNA and DNA from H2228 cell block, we obtained the reference values that we aimed to mimic with the TNA approach, in which the panels for DNA and RNA were combined in different ratios. We verified that none of the conditions impaired the detection of the EML4-ALK fusion transcript, however, the standard ratio yielded the closest expression level comparing to the reference conditions. Considering that the coverage of the internal controls is directly proportional to sequencing quality, we considered the ratio in which the difference was less drastic throughout all controls and observed that the 1:1 ratio was the one whose coverage pattern was most similar to the reference conditions. Regarding DNA sequencing, we were not able to find any mutation likely to be used as a comparison metric between

experimental conditions. Alternatively, we evaluated amplicon coverage, looking for the ratio that resulted in less significant changes. We observed that all the conditions led to an overall increase of amplicon coverage. This could be related with the fact that besides genomic DNA, also cDNA was present in the sequencing reaction, increasing the number of available amplifiable molecules. The amplicon coverage analysis reinforced the previously obtained results regarding RNA that the 1:1 ratio was the more adequate experimental conditional to apply to analyze. By implementing this approach to a clinical sample, we were able to detect a previously unknown alteration in the *ALK* gene, which could have some impact in therapy response- further supporting the added value the application of the TNA approach could have in clinical practice.

12. Conclusions and future perspectives

To the best of our knowledge, this is the first report aiming to evaluate the impact of cfTNA analysis in clinical practice. This work provided evidence that by studying cfRNA and cfDNA simultaneously, it is possible to obtain a more comprehensive understanding of the tumour biology and evolution with treatment, where the detection of actionable mutations could help guide the therapy regimen in a scenario of disease progression. Some of that information would otherwise be lost if only cfDNA or cfRNA were to be analyzed.

We were able to develop a feasible methodology that allows not only the extraction of cfTNA but also the simultaneous detection of genetic mutations and gene translocations, allowing to expand the clinical application of liquid biopsy monitoring to patients that otherwise would not be enrolled in this type of non-invasive monitoring.

Moreover, the optimization tests provided a significant amount of information regarding cfRNA biology, a field still widely unexplored. We observed that in opposition to what is widely described, cfRNA seems to be resistant to nuclease degradation in circulation even without shielding in extracellular vesicles, probably due to association with protein complexes. Moreover, we determined that the EV cfRNA only contributes in a small part to the overall amount of cfRNA. Moreover, the cfRNA and cfDNA yield are severely compromised with freezing. In a scenario of a wide range application of this methodology, it would be recommended to use samples processed within the minimum period.

This set of results represent only the preliminary applications of the cfTNA approach in clinical samples. It would be of interest, in order to validate its utility, to perform comprehensive longitudinal studies with more patients. In addition, there is an ever-increasing need to standardize the biological significance of liquid biopsy results. In this matter, it should be evaluated the creation of threshold values that would allow the detection of progressive disease based on either on the amount of cfTNA or in the AF of the alterations detected, with a higher degree of confidence.

Finally, the expansion of the cfTNA approach to other cancer models and biological fluids, such as cerebral spinal fluid in brain tumours, would also be interesting.

13. Funding

This work resulted from "Institute for Research and Innovation in Health Sciences" (POCI-01-0145-FEDER-007274)- “GenomePT” (POCI-01-0145-FEDER-022184)- “Advancing cancer research: from knowledge to application” (NORTE-01-0145-FEDER-000029) supported by COMPETE 2020 - Operational Programme for Competitiveness and Internationalisation (POCI)- Norte Portugal Regional Programme (NORTE 2020)- Lisboa Portugal Regional Operational Programme (Lisboa2020)- Algarve Portugal Regional Operational Programme (CRESC Algarve2020)- under the PORTUGAL 2020 Partnership Agreement- through the European Regional Development Fund (ERDF)- by FCT - Fundação para a Ciência e a Tecnologia (PTDC/DTP-PIC/2500/2014) and by " Fundo de Investigação para a Saúde" (FIS-FIS-2015-01_ONC) projects.

14. References

1. Du W, Elemento O. Cancer systems biology: embracing complexity to develop better anticancer therapeutic strategies. Oncogene. 2015;34(25):3215,25.
2. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin. 2017;67(1):7,30.
3. Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012;481(7381):306,13.
4. Vineis P, Schatzkin A, Potter JD. Models of carcinogenesis: an overview. Carcinogenesis. 2010;31(10):1703,9.
5. Pitot HC. The molecular biology of carcinogenesis Cancer Biol Ther. 1993;72.
6. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57,70.
7. Stratton MR. Exploring the genomes of cancer cells: progress and promise. Science. 2011;331(6024):1553,8.
8. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., Kinzler KW. Cancer genome landscapes. Science. 2013;339(6127):1546,58.
9. Suva ML, Riggi N, Bernstein BE. Epigenetic reprogramming in cancer. Science. 2013;339(6127):1567,70.
10. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, et al. Cancer statistics, 2008. CA Cancer J Clin. 2008;58(2):71,96.
11. Didkowska J, Wojciechowska U, Manczuk M, Lobaszewski J. Lung cancer epidemiology: contemporary and future challenges worldwide. Ann Transl Med. 2016;4(8):150.
12. Lemjabbar,Alaoui H, Hassan OU, Yang YW, Buchanan P. Lung cancer: Biology and treatment options. Biochim Biophys Acta. 2015;1856(2):189,210.
13. Capelletto E, Novello S. Emerging new agents for the management of patients with non,small cell lung cancer. Drugs. 2012;72 Suppl 1:37,52.
14. Singhal S, Miller D, Ramalingam S, Sun SY. Gene expression profiling of non,small cell lung cancer. Lung Cancer. 2008;60(3):313,24.
15. Pao W, Girard N. New driver mutations in non,small,cell lung cancer. Lancet Oncol. 2011;12(2):175,80.
16. Shames DS, Wistuba, II. The evolving genomic classification of lung cancer. J Pathol. 2014;232(2):121,33.
17. Novello S, Barlesi F, Califano R, Cufer T, Ekman S, Levra MG, et al. Metastatic non,small,cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow,up. Ann Oncol. 2016;27(suppl 5):v1,v27.

18. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol. 2011;6(2):244,85.

19. Lindeman NI, Cagle PT, Beasley MB, Chitale DA, Dacic S, Giaccone G, et al. Molecular testing guideline for selection of lung cancer patients for EGFR and ALK tyrosine kinase inhibitors: guideline from the College of American Pathologists, International Association for the Study of Lung Cancer, and Association for Molecular Pathology. J Thorac Oncol. 2013;8(7):823,59.

20. Lindeman NI, Cagle PT, Beasley MB, Chitale DA, Dacic S, Giaccone G, et al. Molecular testing guideline for selection of lung cancer patients for EGFR and ALK tyrosine kinase inhibitors: guideline from the College of American Pathologists, International Association for the Study of Lung Cancer, and Association for Molecular Pathology. J Mol Diagn. 2013;15(4):415,53.

21. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger K, Yatabe Y, et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society: international multidisciplinary classification of lung adenocarcinoma: executive summary. Proc Am Thorac Soc. 2011;8(5):381,5.

22. Lin JJ, Shaw AT. Resisting Resistance: Targeted Therapies in Lung Cancer. Trends Cancer. 2016;2(7):350,64.

23. Miyamoto Y, Suyama K, Baba H. Recent Advances in Targeting the EGFR Signaling Pathway for the Treatment of Metastatic Colorectal Cancer. Int J Mol Sci. 2017;18(4).

24. Forcella M, Oldani M, Epistolio S, Freguia S, Monti E, Fusi P, et al. Non,small cell lung cancer (NSCLC), EGFR downstream pathway activation and TKI targeted therapies sensitivity: Effect of the plasma membrane,associated NEU3. PLoS One. 2017;12(10):e0187289.

25. Douillard JY, Ostoros G, Cobo M, Ciuleanu T, Cole R, McWalter G, et al. Gefitinib treatment in EGFR mutated caucasian NSCLC: circulating,free tumor DNA as a surrogate for determination of EGFR status. J Thorac Oncol. 2014;9(9):1345,53.

26. Shaw AT, Yeap BY, Mino,Kenudson M, Digumarthy SR, Costa DB, Heist RS, et al. Clinical features and outcome of patients with non,small,cell lung cancer who harbor EML4,ALK. J Clin Oncol. 2009;27(26):4247,53.

27. Barlesi F, Mazieres J, Merlio JP, Debieuvre D, Mosser J, Lena H, et al. Routine molecular profiling of patients with advanced non,small,cell lung cancer: results of a 1,year nationwide programme of the French Cooperative Thoracic Intergroup (IFCT). Lancet. 2016;387(10026):1415,26.

28. Hofman P. ALK in Non,Small Cell Lung Cancer (NSCLC) Pathobiology, Epidemiology, Detection from Tumor Tissue and Algorithm Diagnosis in a Daily Practice. Cancers (Basel). 2017;9(8).

29. Allouche M. ALK is a novel dependence receptor: potential implications in development and cancer. Cell Cycle. 2007;6(13):1533,8.

30. McDermott U, Iafrate AJ, Gray NS, Shioda T, Classon M, Maheswaran S, et al. Genomic alterations of anaplastic lymphoma kinase may sensitize tumors to anaplastic lymphoma kinase inhibitors. Cancer Res. 2008;68(9):3389,95.

31. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the transforming EML4,ALK fusion gene in non,small,cell lung cancer. Nature. 2007;448(7153):561,6.

32. Atlas of Genetics and Cytogenetics in Oncology and Haematology 2018 [updated 30,07,2017. Available from: http://atlasgeneticsoncology.org/Genes/EML4ID44353ch2p21.html.

33. Sabir SR, Yeoh S, Jackson G, Bayliss R. EML4,ALK Variants: Biological and Molecular Properties, and the Implications for Patients. Cancers (Basel). 2017;9(9).

34. Takeuchi K, Choi YL, Togashi Y, Soda M, Hatano S, Inamura K, et al. KIF5B,ALK, a novel fusion oncokinase identified by an immunohistochemistry,based diagnostic system for ALK,positive lung cancer. Clin Cancer Res. 2009;15(9):3143,9.

35. Kohno T, Ichikawa H, Totoki Y, Yasuda K, Hiramoto M, Nammo T, et al. KIF5B,RET fusions in lung adenocarcinoma. Nat Med. 2012;18(3):375,7.

36. Drilon A, Wang L, Hasanovic A, Suehara Y, Lipson D, Stephens P, et al. Response to Cabozantinib in patients with RET fusion,positive lung adenocarcinomas. Cancer Discov. 2013;3(6):630,5.

37. Lira ME, Choi YL, Lim SM, Deng S, Huang D, Ozeck M, et al. A single,tube multiplexed assay for detecting ALK, ROS1, and RET fusions in lung cancer. J Mol Diagn. 2014;16(2):229,43.

38. Tsuta K, Kohno T, Yoshida A, Shimada Y, Asamura H, Furuta K, et al. RET,rearranged non,small,cell lung carcinoma: a clinicopathological and molecular analysis. Br J Cancer. 2014;110(6):1571,8.

39. Wang R, Hu H, Pan Y, Li Y, Ye T, Li C, et al. RET fusions define a unique molecular and clinicopathologic subtype of non,small,cell lung cancer. J Clin Oncol. 2012;30(35):4352,9.

40. Ju YS, Lee WC, Shin JY, Lee S, Bleazard T, Won JK, et al. A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole,genome and transcriptome sequencing. Genome Res. 2012;22(3):436,45.

41. Mulligan LM. RET revisited: expanding the oncogenic portfolio. Nat Rev Cancer. 2014;14(3):173,86.
42. Kohno T, Nakaoku T, Tsuta K, Tsuchihara K, Matsumoto S, Yoh K, et al. Beyond ALK,RET, ROS1 and other oncogene fusions in lung cancer. Transl Lung Cancer Res. 2015;4(2):156,64.
43. Takeuchi K, Soda M, Togashi Y, Suzuki R, Sakata S, Hatano S, et al. RET, ROS1 and ALK fusions in lung cancer. Nat Med. 2012;18(3):378,81.
44. Lipson D, Capelletti M, Yelensky R, Otto G, Parker A, Jarosz M, et al. Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. Nat Med. 2012;18(3):382,4.
45. Pan Y, Zhang Y, Li Y, Hu H, Wang L, Li H, et al. ALK, ROS1 and RET fusions in 1139 lung adenocarcinomas: a comprehensive study of common and fusion pattern,specific clinicopathologic, histologic and cytologic features. Lung Cancer. 2014;84(2):121,6.
46. Yoshida A, Kohno T, Tsuta K, Wakai S, Arai Y, Shimada Y, et al. ROS1,rearranged lung cancer: a clinicopathologic and molecular study of 15 surgical cases. Am J Surg Pathol. 2013;37(4):554,62.
47. Chen YF, Hsieh MS, Wu SG, Chang YL, Shih JY, Liu YN, et al. Clinical and the prognostic characteristics of lung adenocarcinoma patients with ROS1 fusion in comparison with other driver mutations in East Asian populations. J Thorac Oncol. 2014;9(8):1171,9.
48. Bergethon K, Shaw AT, Ou SH, Katayama R, Lovly CM, McDonald NT, et al. ROS1 rearrangements define a unique molecular class of lung cancers. J Clin Oncol. 2012;30(8):863,70.
49. Davies KD, Doebele RC. Molecular pathways: ROS1 fusion proteins in cancer. Clin Cancer Res. 2013;19(15):4040,5.
50. Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014;511(7511):543,50.
51. Seo JS, Ju YS, Lee WC, Shin JY, Lee JK, Bleazard T, et al. The transcriptional landscape and mutational profile of lung adenocarcinoma. Genome Res. 2012;22(11):2109,19.
52. Xalkori (crizotinib) package insert [press release]. New York, Jan 2017 2017.
53. Doebele RC, Pilling AB, Aisner DL, Kutateladze TG, Le AT, Weickhardt AJ, et al. Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non,small cell lung cancer. Clin Cancer Res. 2012;18(5):1472,82.
54. FDA. FDA. Approved drugs: ceritinib. [Available from: www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm395386.htm.

55. Shaw AT, Kim DW, Mehra R, Tan DS, Felip E, Chow LQ, et al. Ceritinib in ALK,rearranged non,small,cell lung cancer. N Engl J Med. 2014;370(13):1189,97.
56. Sakamoto H, Tsukaguchi T, Hiroshima S, Kodama T, Kobayashi T, Fukami TA, et al. CH5424802, a selective ALK inhibitor capable of blocking the resistant gatekeeper mutant. Cancer Cell. 2011;19(5):679,90.
57. Kodama T, Tsukaguchi T, Yoshida M, Kondoh O, Sakamoto H. Selective ALK inhibitor alectinib with potent antitumor activity in models of crizotinib resistance. Cancer Lett. 2014;351(2):215,21.
58. Kim DW, Mehra R, Tan DSW, Felip E, Chow LQM, Camidge DR, et al. Activity and safety of ceritinib in patients with ALK,rearranged non,small,cell lung cancer (ASCEND,1): updated results from the multicentre, open,label, phase 1 trial. Lancet Oncol. 2016;17(4):452,63.
59. Bosc C, Ferretti GR, Cadranel J, Audigier,Valette C, Besse B, Barlesi F, et al. Rebiopsy during disease progression in patients treated by TKI for oncogene,addicted NSCLC. Target Oncol. 2015;10(2):247,53.
60. Sundaresan TK, Sequist LV, Heymach JV, Riely GJ, Janne PA, Koch WH, et al. Detection of T790M, the Acquired Resistance EGFR Mutation, by Tumor Biopsy versus Noninvasive Blood,Based Analyses. Clin Cancer Res. 2016;22(5):1103,10.
61. Hata A, Katakami N, Yoshioka H, Kaji R, Masago K, Fujita S, et al. Spatiotemporal T790M Heterogeneity in Individual Patients with EGFR,Mutant Non,Small,Cell Lung Cancer after Acquired Resistance to EGFR,TKI. J Thorac Oncol. 2015;10(11):1553,9.
62. Abe H, Kawahara A, Azuma K, Taira T, Takase Y, Fukumitsu C, et al. Heterogeneity of anaplastic lymphoma kinase gene rearrangement in non,small,cell lung carcinomas: a comparative study between small biopsy and excision samples. J Thorac Oncol. 2015;10(5):800,5.
63. von Ahlfen S, Missel A, Bendrat K, Schlumpberger M. Determinants of RNA quality from FFPE samples. PLoS One. 2007;2(12):e1261.
64. Dotti I, Bonin S, Nardon E, Basili G, Stanta G. EFFECTS OF FORMALIN, METHACARN AND FINEFIX FIXATION ON RNA RECOVERY AND PRESERVATION. [Poster]. In press.
65. Shackelford RE, Vora M, Mayhall K, Cotelingam J. ALK,rearrangements and testing methods in non,small cell lung cancer: a review. Genes Cancer. 2014;5(1,2):1,14.
66. Guttery DS, Page K, Hills A, Woodley L, Marchese SD, Rghebi B, et al. Noninvasive detection of activating estrogen receptor 1 (ESR1) mutations in estrogen receptor,positive metastatic breast cancer. Clin Chem. 2015;61(7):974,82.

67. Mandel P, Metais P. Les acides nucléiques du plasma sanguin chez l'homme. Comptes rendus des seances de la Societe de biologie et de ses filiales. 1948;142(3,4):241,3.
68. Tan EM, Schur PH, Carr RI, Kunkel HG. Deoxybonucleic acid (DNA) and antibodies to DNA in the serum of patients with systemic lupus erythematosus. J Clin Invest. 1966;45(11):1732,40.
69. Leon SA, Shapiro B, Sklaroff DM, Yaros MJ. Free DNA in the serum of cancer patients and the effect of therapy. Cancer Res. 1977;37(3):646,50.
70. Vasioukhin V, Anker P, Maurice P, Lyautey J, Lederrey C, Stroun M. Point mutations of the N,ras gene in the blood plasma DNA of patients with myelodysplastic syndrome or acute myelogenous leukaemia. Br J Haematol. 1994;86(4):774,9.
71. Sorenson GD, Pribish DM, Valone FH, Memoli VA, Bzik DJ, Yao SL. Soluble normal and mutated DNA sequences from single,copy genes in human blood. Cancer Epidemiol Biomarkers Prev. 1994;3(1):67,71.
72. 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.
73. Jiang P, Chan CW, Chan KC, Cheng SH, Wong J, Wong VW, et al. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. Proc Natl Acad Sci U S A. 2015;112(11):E1317,25.
74. Nagata S, Nagase H, Kawane K, Mukae N, Fukuyama H. Degradation of chromosomal DNA during apoptosis. Cell Death Differ. 2003;10(1):108,16.
75. Holdenrieder S, Stieber P, Bodenmuller H, Fertig G, Furst H, Schmeller N, et al. Nucleosomes in serum as a marker for cell death. Clin Chem Lab Med. 2001;39(7):596,605.
76. Holdenrieder S, Stieber P. Clinical use of circulating nucleosomes. Crit Rev Clin Lab Sci. 2009;46(1):1,24.
77. Wang W, Kong P, Ma G, Li L, Zhu J, Xia T, et al. Characterization of the release and biological significance of cell,free DNA from breast cancer cell lines. Oncotarget. 2017;8(26):43180,91.
78. Cherepanova AV, Tamkovich SN, Bryzgunova OE, Vlassov VV, Laktionov PP. Deoxyribonuclease activity and circulating DNA concentration in blood plasma of patients with prostate tumors. Ann N Y Acad Sci. 2008;1137:218,21.
79. Vlassov AV, Magdaleno S, Setterquist R, Conrad R. Exosomes: current knowledge of their composition, biological functions, and diagnostic and therapeutic potentials. Biochim Biophys Acta. 2012;1820(7):940,8.

80. Lowry MC, Gallagher WM, O'Driscoll L. The Role of Exosomes in Breast Cancer. Clin Chem. 2015;61(12):1457,65.

81. Thakur BK, Zhang H, Becker A, Matei I, Huang Y, Costa,Silva B, et al. Double,stranded DNA in exosomes: a novel biomarker in cancer detection. Cell Res. 2014;24(6):766,9.

82. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, et al. Detection of circulating tumor DNA in early, and late,stage human malignancies. Sci Transl Med. 2014;6(224):224ra24.

83. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: monitoring cancer,genetics in the blood. Nat Rev Clin Oncol. 2013;10(8):472,84.

84. Heitzer E, Ulz P, Geigl JB. Circulating tumor DNA as a liquid biopsy for cancer. Clin Chem. 2015;61(1):112,23.

85. Veldore VH, Choughule A, Routhu T, Mandloi N, Noronha V, Joshi A, et al. Validation of liquid biopsy: plasma cell,free DNA testing in clinical management of advanced non,small cell lung cancer. Lung Cancer (Auckl). 2018;9:1,11.

86. Kwapisz D. The first liquid biopsy test approved. Is it a new era of mutation testing for non,small cell lung cancer? Ann Transl Med. 2017;5(3):46.

87. Sozzi G, Conte D, Leon M, Ciricione R, Roz L, Ratcliffe C, et al. Quantification of free circulating DNA as a diagnostic marker in lung cancer. J Clin Oncol. 2003;21(21):3902,8.

88. Shao X, He Y, Ji M, Chen X, Qi J, Shi W, et al. Quantitative analysis of cell,free DNA in ovarian cancer. Oncol Lett. 2015;10(6):3478,82.

89. Hench IB, Hench J, Tolnay M. Liquid Biopsy in Clinical Management of Breast, Lung, and Colorectal Cancer. Front Med (Lausanne). 2018;5:9.

90. Zaporozhchenko IA, Ponomaryova AA, Rykova EY, Laktionov PP. The potential of circulating cell,free RNA as a cancer biomarker: challenges and opportunities. Expert Rev Mol Diagn. 2018;18(2):133,45.

91. Sorber L, Zwaenepoel K, Deschoolmeester V, Van Schil PE, Van Meerbeeck J, Lardon F, et al. Circulating cell,free nucleic acids and platelets as a liquid biopsy in the provision of personalized therapy for lung cancer patients. Lung Cancer. 2017;107:100,7.

92. Kohn L, Johansson M, Grankvist K, Nilsson J. Liquid biopsies in lung cancer,time to implement research technologies in routine care? Ann Transl Med. 2017;5(13):278.

93. Perez,Callejo D, Romero A, Provencio M, Torrente M. Liquid biopsy based biomarkers in non,small cell lung cancer for diagnosis and treatment monitoring. Transl Lung Cancer Res. 2016;5(5):455,65.

94. Cai X, Janku F, Zhan Q, Fan JB. Accessing Genetic Information with Liquid Biopsies. Trends Genet. 2015;31(10):564,75.
95. Qiagen. therascreen® EGFR RGQ PCR Kit Instructions for Use (Handbook). 2013.
96. Kim SS, Choi HJ, Kim JJ, Kim MS, Lee IS, Byun B, et al. Droplet digital PCR,based EGFR mutation detection with an internal quality control index to determine the quality of DNA. Sci Rep. 2018;8(1):543.
97. Rothe F, Laes JF, Lambrechts D, Smeets D, Vincent D, Maetens M, et al. Plasma circulating tumor DNA as an alternative to metastatic biopsies for mutational analysis in breast cancer. Ann Oncol. 2014;25(10):1959,65.
98. Forshew T, Murtaza M, Parkinson C, Gale D, Tsui DW, Kaper F, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med. 2012;4(136):136ra68.
99. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med. 2014;20(5):548,54.
100. Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B. Detection and quantification of rare mutations with massively parallel sequencing. Proc Natl Acad Sci U S A. 2011;108(23):9530,5.
101. Bennett CW, Berchem G, Kim YJ, El,Khoury V. Cell,free DNA and next,generation sequencing in the service of personalized medicine for lung cancer. Oncotarget. 2016;7(43):71013,35.
102. Christensen E, Nordentoft I, Vang S, Birkenkamp,Demtroder K, Jensen JB, Agerbaek M, et al. Optimized targeted sequencing of cell,free plasma DNA from bladder cancer patients. Sci Rep. 2018;8(1):1917.
103. Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating liquid biopsies into the management of cancer. Nat Rev Clin Oncol. 2017;14(9):531,48.
104. Wu X, Brewer G. The regulation of mRNA stability in mammalian cells: 2.0. Gene. 2012;500(1):10,21.
105. Koczera P, Martin L, Marx G, Schuerholz T. The Ribonuclease A Superfamily in Humans: Canonical RNases as the Buttress of Innate Immunity. Int J Mol Sci. 2016;17(8).
106. Hynie I, Meuffels M, Poznanski WJ. Determination of phosphodiesterase I activity in human blood serum. Clin Chem. 1975;21(10):1383,7.
107. Bryzgunova OE, Laktionov PP. Extracellular Nucleic Acids in Urine: Sources, Structure, Diagnostic Potential. Acta Naturae. 2015;7(3):48,54.

108. Tsui NB, Ng EK, Lo YM. Stability of endogenous and added RNA in blood specimens, serum, and plasma. Clin Chem. 2002;48(10):1647,53.
109. Fleischhacker M, Schmidt B. Circulating nucleic acids (CNAs) and cancer,,a survey. Biochim Biophys Acta. 2007;1775(1):181,232.
110. Hasselmann DO, Rappl G, Tilgen W, Reinhold U. Extracellular tyrosinase mRNA within apoptotic bodies is protected from degradation in human serum. Clin Chem. 2001;47(8):1488,9.
111. Crescitelli R, Lasser C, Szabo TG, Kittel A, Eldh M, Dianzani I, et al. Distinct RNA profiles in subpopulations of extracellular vesicles: apoptotic bodies, microvesicles and exosomes. J Extracell Vesicles. 2013;2.
112. Pellegrini KL, Sanda MG, Moreno CS. RNA biomarkers to facilitate the identification of aggressive prostate cancer. Mol Aspects Med. 2015;45:37,46.
113. Gallo A, Vukic D, Michalik D, O'Connell MA, Keegan LP. ADAR RNA editing in human disease; more to it than meets the I. Hum Genet. 2017;136(9):1265,78.
114. Feng H, Qin Z, Zhang X. Opportunities and methods for studying alternative splicing in cancer with RNA,Seq. Cancer Lett. 2013;340(2):179,91.
115. Garcia V, Garcia JM, Pena C, Silva J, Dominguez G, Lorenzo Y, et al. Free circulating mRNA in plasma from breast cancer patients and clinical outcome. Cancer Lett. 2008;263(2):312,20.
116. March,Villalba JA, Martinez,Jabaloyas JM, Herrero MJ, Santamaria J, Alino SF, Dasi F. Cell,free circulating plasma hTERT mRNA is a useful marker for prostate cancer diagnosis and is associated with poor prognosis tumor characteristics. PLoS One. 2012;7(8):e43470.
117. Souza MF, Kuasne H, Barros,Filho MC, Ciliao HL, Marchi FA, Fuganti PE, et al. Circulating mRNAs and miRNAs as candidate markers for the diagnosis and prognosis of prostate cancer. PLoS One. 2017;12(9):e0184094.
118. Yuan T, Huang X, Woodcock M, Du M, Dittmar R, Wang Y, et al. Plasma extracellular RNA profiles in healthy and cancer patients. Sci Rep. 2016;6:19413.
119. Ng EK, Tsui NB, Lam NY, Chiu RW, Yu SC, Wong SC, et al. Presence of filterable and nonfilterable mRNA in the plasma of cancer patients and healthy individuals. Clin Chem. 2002;48(8):1212,7.
120. Pos O, Biro O, Szemes T, Nagy B. Circulating cell,free nucleic acids: characteristics and applications. Eur J Hum Genet. 2018;26(7):937,45.
121. Tzimagiorgis G, Michailidou EZ, Kritis A, Markopoulos AK, Kouidou S. Recovering circulating extracellular or cell,free RNA from bodily fluids. Cancer Epidemiol. 2011;35(6):580,9.

122. Worrillow L, Baskaran P, Care MA, Varghese A, Munir T, Evans PA, et al. An ultra,deep sequencing strategy to detect sub,clonal TP53 mutations in presentation chronic lymphocytic leukaemia cases using multiple polymerases. Oncogene. 2016;35(40):5328,36.

123. Bartak BK, Kalmar A, Galamb O, Wichmann B, Nagy ZB, Tulassay Z, et al. Blood Collection and Cell,Free DNA Isolation Methods Influence the Sensitivity of Liquid Biopsy Analysis for Colorectal Cancer Detection. Pathology oncology research : POR. 2018.

124. Soo J, Kurtz DM, Jin M, Scherer F, Schroers,Martin JG, Esfahan MS, et al. Comparison of Circulating Tumor DNA Recovery from Plasma and Serum. Blood. 2017;vol. 130 no. Suppl 1 2756.

15. Appendix 1

Manuscript accepted for publication in *Acta Cytologica*

Circulating tumor DNA: a step into the future of cancer management

Short title: ctDNA: a step into the future of cancer management

Joana Marques MSc* [1-2], Joana Reis BSc* [1-2-3], Gabriela Fernandes MD [1-2-4], Venceslau Hespanhol MD PhD [1-2-4], José Carlos Machado PhD [1-2-3] , José Luís Costa- PhD [1-2-3]

*The authors contributed equally to this work

[1] i3S - Instituto de Investigação e Inovação em Saúde- Universidade do Porto- Porto- Portugal.

[2] IPATIMUP - Institute of Molecular Pathology and Immunology of the University of Porto- Porto- Portugal.

[3] Faculty of Medicine- University of Porto- Porto- Portugal.

[4] Department of Pulmonology- Hospital de São João- Porto- Portugal

Abstract

Liquid biopsy was introduced to the oncology field with the promise of revolutionizing the management of cancer patients, minimizing the exposure to invasive procedures such as tissue biopsy and providing more reliable information regarding therapy response and detection of disease relapse. Despite the significant increase in the number of published studies about circulating tumor DNA (ctDNA) in the past years, the emphasis of most studies relies on the development of new technologies or in the clinical significance of ctDNA, leaving a clear gap of knowledge concerning the biology of ctDNA, such as the fundamental mechanisms through which DNA from tumor cells is released into circulation. Moreover, considering that ctDNA analysis is now currently being applied in clinical practice, the need for rigorous quality control is arising and with it, the necessity to standardize the procedures, from sample collection to data analysis. This review discusses the main aspects of ctDNA and the approaches currently available to evaluate tumor genetics, as well as the core points that still require some improvement in order to make liquid biopsy a key player in precision medicine.

Key words: ctDNA, liquid biopsy, cancer, mutation detection, disease monitoring, early diagnosis

Biology of circulating tumor DNA

The discovery of circulating nucleic acids, namely circulating free DNA (cfDNA), in peripheral blood dates back to the late 40s when Mendel and Métais first identified DNA in circulation in peripheral blood of healthy individuals, pregnant women and clinical patients [1]. After this finding, only in 1977 new publications emerged on this topic. Leon *et al.*, applied the study of circulating nucleic acids to oncology being the first to describe an increased concentration of cfDNA in the serum of cancer patients compared to healthy individuals [2]. Over the years, the interest on detection of circulating tumor DNA (ctDNA) of cancer patients has expanded to study its biology, new methodologies and potential application in the clinic as a liquid biopsy approach.

The concept of ctDNA refers to fragmented DNA, single- or double-stranded, released by tumor cells (Figure 1). Tumor-derived DNA may be representative of the tumor genetics and heterogeneity. In fact, the “scanning” of plasma has allowed the identification of point mutations, copy number aberrations, chromosomal and epigenetic alterations, also present in the primary tumor [3,4]. These DNA fragments are continuously released by tumor cells and represent an instant picture of the tumor due to its short half-life in circulation (approximately 4 to 30 min) [5]. It is accepted that an increased concentration of DNA in plasma is observed in cancer patients when

compared to healthy individuals [6,7]. However, ctDNA concentration is highly variable between cancer patients [8]. Despite this, the amount of ctDNA often corresponds to a small fraction of total cfDNA (≈10%) due to the presence of circulating DNA derived from non-tumor cells [9,10].

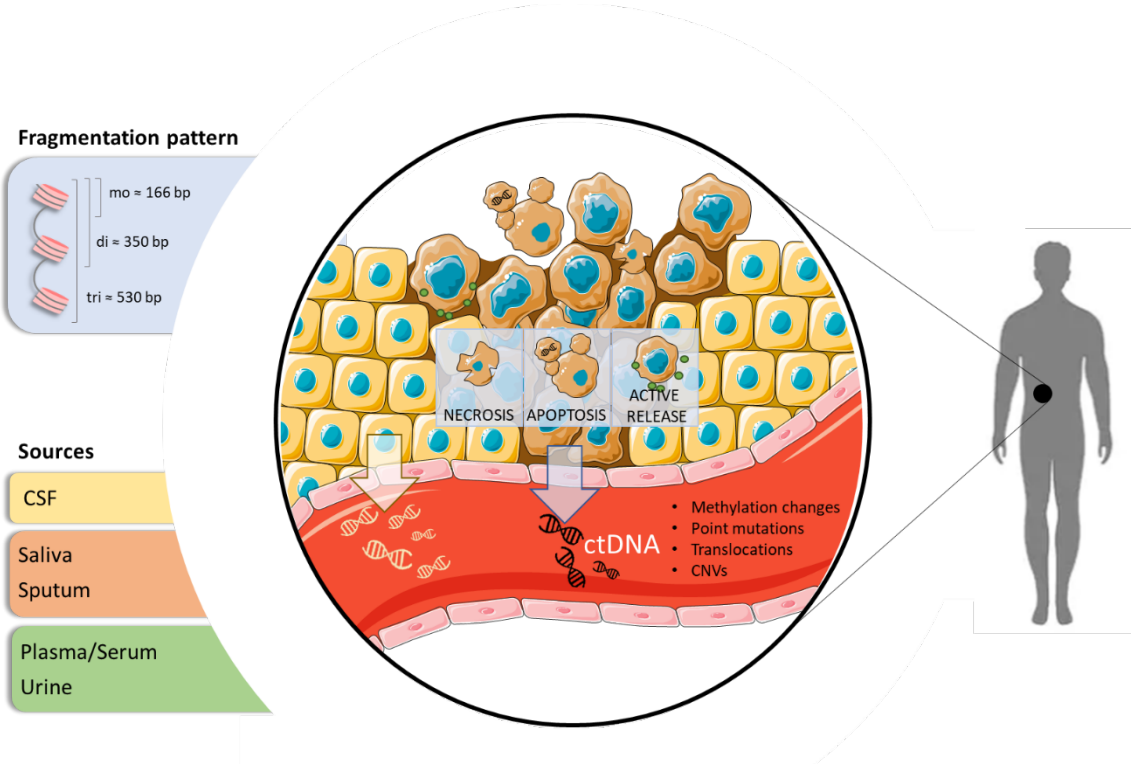

Figure 1. Biology and dynamics of cfDNA. DNA can be released into circulation during apoptotic, necrotic or active processes. Circulating tumor DNA correspond to small fragments with the size of a mononucleosome (mo) or multiple units of a nucleosome (dinucleosome (di) or trinucleosome (tri)). ctDNA can harbor different genetic alterations that mirror tumor genetics. Depending on the tumor anatomic location DNA can be enriched in specific body fluids (ex: plasma, urine, saliva or CSF).

The mechanism of ctDNA release is still a topic of discussion but evidence suggests apoptosis as one of the sources. Jiang *et al.* observed that cfDNA includes mostly fragments of around 166 bp, which corresponds to mononucleosome units, indicating that ctDNA is derived from apoptotic tumor cells [11,12]. This pattern of fragmentation has already been described in cancer patients [11,13,14]. Additionally, Mouliere *et al.* have demonstrated, for the first time, the presence of ctDNA with less than 100bp in plasma of cancer patients [13]. Altogether, these studies state that cfDNA from cancer patients is more fragmented than cfDNA present in healthy individuals [11,13]. The presence of bigger DNA fragments in circulation can indicate another mechanism of DNA release, such as necrosis. Secretion of DNA from living tumor cells

or circulating tumor cells (CTCs) might also account for the presence of cfDNA in circulation.

DNA can be virtually shed into all body fluids. Plasma is accepted a standard material for the extraction of ctDNA [15]. Several studies comparing ctDNA present in plasma and serum have reported increased levels of cfDNA in serum when compared to paired plasma samples [16-18]. Nevertheless, Soo *et al.* showed that plasma was enriched for ctDNA, resulting in increased sensitivity to detect tumor derived somatic mutations [15]. The overall higher amounts of cfDNA found in serum represent an enrichment for cfDNA with non-tumor origin, probably derived from lysis of white blood cells during the clothing process [18].

Non-blood body fluids such as urine, saliva, sputum and cerebrospinal fluid (CSF) can also contain ctDNA [19-21]. Specific body fluids might accumulate preferentially tumor-derived DNA, according to their anatomical location, increasing the sensitivity of detection compared to plasma DNA. For example, in the context of bladder or prostate cancer, urine ctDNA has been reported as a more representative source of genomic aberrations than plasma ctDNA [22,23]. Also, the detection of ctDNA in saliva in oral cancer, and sputum in neck and lung cancers may also represent a suitable source of ctDNA with relevant genetic alterations [24,25]. The difficulties associated with the characterization of genomic alterations in brain cancers and brain metastasis, due to the blood brain barrier, may be overcomed by the study of CSF since it has shown improved genetic representability than plasma ctDNA analysis [26]. Still major methodological aspects need to be optimized in order to make this a reality.

Methodological considerations

The growing interest in the liquid biopsy field and the fast development of technologies to analyze cfDNA has generated a lot of knowledge, mainly about tumor biology and disease progression [27]. Nonetheless, much of the technology currently used to analyze ctDNA was initially developed to detect fetal cell free DNA for noninvasive prenatal testing (NIPT). NIPT is considered the precursor of liquid biopsy, representing the first widely used clinical application of cfDNA [28].

Prior to a broad range application of liquid biopsy in clinical practice, there is an urgent need to standardize methodologies in order to obtain concordant results from different institutions. In this context, one of the aspects that is still missing for liquid biopsy is sample quality control. The first parameter that should be assessed upon the implementation of liquid biopsy is sample collection tubes. There are several commercially available tubes with numerous differences among them. The most frequently used, the Cell-Free DNA BCT® tubes, prevent the release of genomic DNA,

stabilizing cfDNA with the use of preservatives. These tubes allow the collected samples to be stored at room temperature up to 14 days, which represent a massive advantage regarding logistic handling. The K_3EDTA tubes are considerably cheaper, however, their use requires samples to be processed as soon as possible (preferably upon 4 hours after sample collection). Several reports evaluated the impact of each type of tubes on cfDNA yields, including Barták *et al.* that demonstrates that the levels of cfDNA are not affected by the use of neither one of the tubes, when samples are processed within standard conditions [29].

Another matter that should be addressed with major caution is the cfDNA extraction method. There are currently diverse commercially available solutions intended for cfDNA extraction from body fluids, either column or magnetic bead based. Regardless of the method of election, it is crucial to assess the specificity towards the recovery of low molecular weight DNA. In this matter, magnetic bead-based methodologies resulted in superior global cfDNA yield with minor contamination from high molecular weight (HMW) DNA [30]. The quantification of cfDNA also represents a critical step for quality control that will affect all the downstream processes. The most commonly used quantification methods, such as spectrophotometry or fluorometric based approaches account for the total double stranded DNA present in the sample. Spectrophotometry should not be used to quantify cfDNA samples, considering the low amount of DNA present and the high level of imprecision associated with this method [31], which could lead to an overestimation of the real amount present. The fluorometric approaches are more precise, however, provide no information about size distribution, making impossible to distinguish cfDNA from HMW DNA. Ideally, to accurately quantify the cfDNA fraction, the most informative way is by automated electrophoresis, considering the high resolution provided by these devices. This enables the assessment of the presence of HMW DNA that can impair the detection of alterations present in ctDNA [32].

The ultimate goal of cfDNA analysis in the oncology field is the detection of relevant genetic alterations. However, the abundance of ctDNA within cfDNA can be as low as 0.01% in an individual with early-stage cancer which stresses the importance of using highly sensitive methodologies to detect variants of interest [33]. Next Generation Sequencing (NGS) and Digital PCR (dPCR) strategies are amongst the most used techniques in liquid biopsy. Succinctly, NGS allows the detection of previously unknow variants, which is particularly important at the time of diagnosis when the screening of targetable mutations is fundamental to guide therapeutic decisions. The most commonly used NGS platforms can reach a limit of detection of about 1-5%. However, cfDNA samples often display alterations significantly below that limit. To overcome that

limitation, the use of unique molecular identifiers (UMI) [22,34], most commonly referred as molecular barcodes, is fundamental to reach a lower limit of detection, down to 0.01%. This is only possible because prior to any amplification step, each individual DNA molecule is identified with a UMI. Upon analysis of the sequencing results, it is possible to distinguish between the variants that were present in the original DNA template from the errors that were added during the process, thus reducing the signal to noise ratio, greatly improving the confidence in the obtained results [35,36]. Regarding dPCR, it requires the mutational status to be previously determined since it uses specific probes for each variant, which constrains the application of dPCR to the initial screening of the tumor mutational landscape in a comprehensive manner. Even with droplet dPCR, that allows the detection of multiple targets, it is still limited to a reduced number of target mutations at once [37]. Nevertheless, dPCR represents a more cost-effective manner to monitor tumor burden, more viable to be applied throughout the course of disease [38].

The amount of information generated should also be taken into consideration. NGS-based solutions usually result in a significant amount of bioinformatic data which could possibly increase the time necessary to properly analyze the results, and consequently, delay the conclusion of the test. Taking into account that one of the main advantages of liquid biopsy is the fast turnaround time (3 business days versus 12 business days required for a tissue biopsy) it is essential that the process is concluded as quickly as possible, as it happens with the dPCR based approaches [39].

All the currently available solutions have advantages and drawbacks and the choice of one method should always be in accordance with the biological question that is being addressed. When evaluating ctDNA, perhaps the most important features to take into account are the detection limit and the cost-effectiveness, so that implementation in clinical practice is feasible. The comparison between the most commonly used methods is summarized in Table 1.

In order to validate the experimental methodology and analytical pipeline, it is highly recommended to incorporate external controls, such as commercially available synthetic cfDNA samples [40]. These controls are derived from isogenic cell lines harboring different mutations that can be chosen according with the model of disease that is being studied. The genomic DNA within these samples is fragmented to mimic the fragmentation pattern of cfDNA by acoustic shearing, for example [34]. These controls are available with different allelic frequencies, down to 0.1%, representing an essential control to assess specificity and sensitivity that should be conducted prior to any other experiment.

The application of some liquid biopsy tests in clinical practice was preceded by approval of regulatory agencies, such as the Food and Drug Administration (FDA) or European Medicines Agency (EMA). The cobas® EGFR Mutation Test for non-small cell lung cancer and the QIAGEN's therascreen® EGFR RGQ PCR Kit represent, until now, the only validated liquid biopsy tests. According to a recent report by ASCO and CAP experts these were considered the only tests with clinical utility [41]. The methodologies that are not accredited can still be used in clinical practice, however, their application requires a strict analysis of analytical validity of the test, considering as well the pre-analytical factors that can affect the biological material [41].

Table 1*:* Comparison of different methodologies to analyze mutations present in cfDNA.

Approach	Method	Detection limit	Comments
Specific mutation detection (mutational status previously determined)	qPCR	<10% [67,68]	Low sensitivity
	dPCR	<0.001% [69]	High sensitivity
	Scorpion ARMS (Amplification-Refractory Mutation System)	<0.1% [47]	High cost
	BEAMing (Beads, Emulsion, Amplification,Magnetics)	<0.1% [70]	High complexity
	QIAGEN's therascreen® EGFR RGQ PCR Kit	>0.81% [71]	FDA approved
	cobas® EGFR Mutation Test	3% [72]	FDA approved
Deep sequencing (screening of unknown alterations)	Ampliseq	>2% [73]	High cost. Increased turnaround time
	Tam-Seq	>2% [74]	
	CAPP-Seq	0.01% [75]	
	Safe-Seq	0.1% [76]	
	Tag sequencing	0.1% [35,36]	
	Whole exome sequencing	>1% [77]	

Most of the studies recently or currently conducted focus on the possible role of liquid biopsy as a monitoring tool to detect resistance to therapy or minimal residual disease. [42]. Several authors, and more recently Veldore *et al.*, have compared the accuracy of tissue *versus* liquid biopsy and were able to obtain a sensitivity of 96.97% and 100% specificity by using allele specific real-time PCR and NGS [43]. This data suggests the potential application of liquid biopsy as a valid substitute when tissue biopsy cannot be performed or when the low amount of biological material is limited to perform all the necessary genetic testing [44].

The quantification of overall cfDNA can by itself be a prognostic factor, being directly proportional to the stage of disease [45,46] as it is reported that patients with more advanced tumors have higher levels of cfDNA in circulation [47]. However, systemic conditions such as autoimmune or inflammatory disorders may account for increased cfDNA levels [48]. Taking all this into account, one should always evaluate cfDNA levels prior to start treatment for each patient, to establish a standard value of reference for future analysis. Also, it is important to consider that cfDNA levels not only vary greatly among different tumor types, they also fluctuate between individuals with the same pathology, making very hard to establish the normal physiological levels [8]. Thus, evaluating cfDNA levels alone might overestimate tumor burden or misjudge therapy response, which is why a more accurate measure for the actual fraction of ctDNA should always be applied.

There is not a direct strategy to quantify the tumor-derived DNA. Bearing in mind that tumor cells harbor specific somatic mutations, they represent the ideal biomarker to accurately determine the tumor DNA fraction in circulation. In this context, diverse studies report that by evaluating serial liquid biopsies from the same patient under treatment, it is possible to reflect the course of disease by assessing the dynamic changes of mutations in circulation. Following response to treatment, there is a significant decrease of ctDNA and upon stable disease it is expected that ctDNA levels remain constant, only increasing when progressive disease is developed [49]. This increase in the ctDNA release can be foreseen by the increase in the allelic frequency of the tumor derived mutations, which is greatly associated with the development of therapy resistance mechanisms. Fortunately, these modifications can be observed in plasma prior to any clinical signs of progression, even when no resistance mechanism is detected, representing a critical role for ctDNA analysis [49]. Moreover, when progression is driven by a known resistance mechanism, such as the EGFR p.T790M point mutation in non-small cell lung cancer treated with tyrosine kinase inhibitors, it can be detected in cfDNA up to 3 months before signs of progressive disease are detected by radiological studies [50]. More recently, other cancer types have also been studied,

such as breast cancer, showing that ctDNA can be used to detect ESR1 mutations (conferring resistance to estrogen aromatase inhibitors) before disease progression is detected, allowing for an early change in the therapeutic regimen [51-53]. Colorectal cancer (CRC) has also been a prime model to study liquid biopsy application in clinic. Scholer *et al.* demonstrated that ctDNA changes are positively related with tumor burden in CRC patients submitted to surgery [54] . Moreover, Germano *et al.* reported that ctDNA analysis yields higher sensitivity in detecting clinically relevant alterations when compared to CTCs [55].

Even though the majority of research has been conducted with advanced stage patients, some recent reports also show the application of liquid biopsy in early stage cancer patients submitted to surgery with curative intent. In this cohort of patients, liquid biopsy can be used as a surrogate for tumor associated proteins such as carcinoembryonic antigen, prostate specific antigen, CA-125, MUC 1 or CA19-9 in detecting relapse, with increased sensitivity [41]. Moreover, it is now described that the evaluation of ctDNA levels after surgery can by itself predict relapse in the future [47,56]. As reviewed by Ellen Heitzer *et al.* there are two additionally scenarios for early detection including screening of patients at risk, due to known hereditary or lifestyle factors, and general population by searching genetic alterations indicative of tumorigenesis. The last scenario is the more controversial one, since false positive diagnosis can lead to unnecessary stress to the patient and increase the costs for cancer management [57].

The broad interest taken towards liquid biopsy in translational research is now being applied into clinical trials. The first prospective clinical trials that rely on the analysis of ctDNA alterations to guide therapeutic decisions have been approved, such as the APPLE Trial [58], in which liquid biopsy may play a major role in detecting tumor progression earlier than when applying exclusively RECIST criteria. Regarding breast cancer, there are currently trials aiming to recruit triple negative breast cancer patients treated with immune checkpoint inhibitors, for whom liquid biopsy will be used to detect minimal residual disease (c-TRAK-TN, ClinicalTrials.gov Identifier: NCT03145961CRC). Liquid biopsy in CRC is also being addressed in clinical trials, namely to detect micrometastatic disease in early or intermediate stage CRC patients (ClinicalTrials.gov Identifier: NCT03312374).

In summary, there is now sufficient knowledge and technology to apply ctDNA evaluation on different tumor types, different stages of disease and to patients submitted to different therapy regimens that can range from tyrosine kinase inhibitors to chemotherapy [59,60]. This has clear benefits for both patients, minimizing the exposure to invasive techniques and/or radiation and for the institutions considering that in some cases liquid biopsy can be more cost-effective than the traditional diagnostic techniques

[41]. However, there are still drawbacks to overcome since ASCO and CAP consider that the gathered data regarding the clinical application of liquid biopsy is not yet sufficient for the methodology to be applied outside of the clinical trial context [41].

Future Perspectives

Despite the recent breakthroughs in ctDNA characterization, the major challenges still delaying the wide application of liquid biopsy in clinic routine are the low abundance of ctDNA and the lack of standard and highly sensitive methodologies for detection and proper validation of ctDNA analysis.

A more comprehensive knowledge of ctDNA biology and dynamics may help to overcome the low abundance of ctDNA. Unpublished data from our lab demonstrates that it is possible to increase the yield of ctDNA in plasma by adjusting the timing of blood collection according to the patients' treatment schedule, leading to higher sensitivity of detection of genetic alterations in ctDNA. Moreover, it would be of interest to determine biomarkers to guide towards the tissue of origin of ctDNA allowing to distinguish ctDNA from the non-tumor cfDNA, which would have a dramatic impact in the diagnosis of cancers of unknown primary. To this end, studies in cancer epigenetics have described that cfDNA nucleosome fragmentation and methylation patterns in ctDNA can indicate the cell type of origin in cancer patients [14]. With the advances in the bioinformatics field, systematic prediction algorithms are being developed, based on methylation signatures and somatic mutations, to trace the tissue of origin of ctDNA [61,62]. In the future, more comprehensive studies to determine tissue of origin based on genetic alterations may help to guide clinical practice in advanced metastatic cancers.

One of the hot topics in cancer research is the development of new strategies for early-stage cancer detection to prevent cancer progression. Here, ctDNA analysis is a potential approach due to its genetic representativity of the tumor and non-invasive nature. A pioneering liquid biopsy strategy has been developed addressing the challenges of diagnostic sensitivity for detection of ctDNA in patients with early cancers and also their tissue of origin, through the application of a new multi-analyte blood test, called CancerSEEK [63]. This test combines assays for genetic alterations and protein biomarkers to discriminate eight cancer types. An important factor of these study was the inclusion of a big cohort of healthy controls which culminated in a specificity of >99% of the test. Moreover, the sensitivity for the CancerSEEK test was approximately 55% for all eight cancer types. This is still an emerging concept and new studies addressing different genetic alterations to define specific cancer types signatures are essential for an accurate diagnosis with reduced false negatives.

The detection and characterization of ctDNA has already proven to have a significant clinical value in cancer management. However, no standard guidelines have been established for ctDNA genetic analysis. It is essential to have a well-established standardization of procedures, as rapid and economic as possible to be applied to the majority of patients. Thus, there is an urgent need to set specific workflows so that the different strategies can become more equivalent with comparable results. In this context, our lab is now developing an integrated workflow for analysis of total nucleic acids, DNA and RNA, allowing a more complete, simple and economic analysis and consequently, a faster outcome to guide clinical follow up. Moreover, the development of National and International external quality assessment (EQA) programs to ensure better practices and standardize workflows are crucial to boost the confidence in ctDNA testing [64]. These policies will also help to expand the approval of platforms for clinical application of ctDNA tools by regulatory agencies, as already happened with EGFR mutation analysis in lung cancer patients.

Currently, the majority of studies in cohorts of patients have a retrospective design that may comprise limitations regarding long time storing of samples and poor assessment of clinical value [65]. Therefore, the development of prospective clinical trials including ctDNA analysis alone as a diagnostic and/or predicting marker for disease progression are essential to prove clinical utility of liquid biopsy and to consolidate technical standardization and EQA. Recently, few clinical trials are including an arm where ctDNA is used as a stand-alone marker to guide therapeutic decision when specific alterations related to therapy resistance are identified in plasma even before signs of clinical progression [58]. Additionally, these new studies should always include the correlation of plasma and tumor tissue genotyping results in order to confirm the power of liquid biopsy strategies and accelerate its application in the clinic.

The recent breakthroughs in oncology and the era of personalized medicine are having a great impact in extending the overall survival and improving quality of life, guiding cancer towards a chronic condition. Nevertheless, the development of therapy resistance constitutes a barrier to the successful management of cancer [66]. Therefore, it is crucial to develop comprehensive studies aiming to understand the mechanisms underlying drug resistance to support the identification of new genetic biomarkers and consequently establish liquid biopsy strategies to inform therapy response and resistance through ctDNA genetic profile.

Conclusion

Even though ctDNA is a fairly recent addition to the oncology field, its undeniable potential is already very well established. With all the studies conducted recently, it is

possible that liquid biopsy will soon be included as a management tool worldwide and to a variety of cancer types. The advantages in applying ctDNA analysis go beyond the non-invasive component, providing a clearer insight on tumor biology at any given time, which could result in improved clinical outcomes due to the real-time monitoring of disease burden that is possible with liquid biopsy.

Funding

The laboratory is supported by FEDER—Fundo Europeu de Desenvolvimento Regional funds through the COMPETE 2020—Operacional Programme for Competitiveness and Internationalization (POCI), Portugal 2020, and by FCT—Fundacao para a Ciencia e a Tecnologia/Ministerio da Ciencia, Tecnologia e Inovacao in the framework of the projects “Institute for Research and Innovation in Health Sciences” (POCI-01-0145-FEDER-007274) and (PTDC/DTP-PIC/2500/2014); and by Norte Portugal Regional Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF), in the framework of the project NORTE- 01-0145-FEDER-000029.

Conflict of interests

No potential conflicts of interest were disclosed by the other authors.

1 Mandel P, Metais P: Les acides nucléiques du plasma sanguin chez l'homme. Comptes rendus des seances de la Societe de biologie et de ses filiales 1948;142:241-243.
2 Leon SA, Shapiro B, Sklaroff DM, Yaros MJ: Free DNA in the serum of cancer patients and the effect of therapy. Cancer research 1977;37:646-650.
3 Leary RJ, Sausen M, Kinde I, Papadopoulos N, Carpten JD, Craig D, O'Shaughnessy J, Kinzler KW, Parmigiani G, Vogelstein B, Diaz LA, Velculescu VE: Detection of Chromosomal Alterations in the Circulation of Cancer Patients with Whole-Genome Sequencing. Science translational medicine 2012;4:162ra154-162ra154.
4 Imamura F, Uchida J, Kukita Y, Kumagai T, Nishino K, Inoue T, Kimura M, Oba S, Kato K: Monitoring of treatment responses and clonal evolution of tumor cells by circulating tumor DNA of heterogeneous mutant EGFR genes in lung cancer. Lung cancer (Amsterdam, Netherlands) 2016;94:68-73.
5 Lo YM, Zhang J, Leung TN, Lau TK, Chang AM, Hjelm NM: Rapid clearance of fetal DNA from maternal plasma. American journal of human genetics 1999;64:218-224.
6 Madhavan D, Wallwiener M, Bents K, Zucknick M, Nees J, Schott S, Cuk K, Riethdorf S, Trumpp A, Pantel K, Sohn C, Schneeweiss A, Surowy H, Burwinkel B: Plasma DNA integrity as a biomarker for primary and metastatic breast cancer and potential marker for early diagnosis. Breast Cancer Res Treat 2014;146:163-174.
7 Sozzi G, Conte D, Mariani L, Lo Vullo S, Roz L, Lombardo C, Pierotti MA, Tavecchio L: Analysis of circulating tumor DNA in plasma at diagnosis and during follow-up of lung cancer patients. Cancer research 2001;61:4675-4678.
8 Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM, Antonarakis ES, Azad NS, Bardelli A, Brem H, Cameron JL, Lee CC, Fecher LA, Gallia GL, Gibbs P, Le D, Giuntoli RL, Goggins M, Hogarty MD, Holdhoff M, Hong SM, Jiao Y, Juhl HH, Kim JJ, Siravegna G, Laheru DA, Lauricella C, Lim M, Lipson EJ, Marie SK, Netto GJ, Oliner KS, Olivi A, Olsson L, Riggins GJ, Sartore-Bianchi A, Schmidt K, Shih I M, Oba-Shinjo SM, Siena S, Theodorescu D, Tie J, Harkins TT, Veronese S, Wang TL, Weingart JD, Wolfgang CL, Wood LD, Xing D, Hruban RH, Wu J, Allen PJ, Schmidt CM, Choti MA, Velculescu VE, Kinzler KW, Vogelstein B, Papadopoulos N, Diaz LA, Jr.: Detection of circulating tumor DNA in early- and late-stage human malignancies. Science translational medicine 2014;6:224ra224.
9 Thierry AR: A Targeted Q-PCR-Based Method for Point Mutation Testing by Analyzing Circulating DNA for Cancer Management Care. Methods in molecular biology (Clifton, NJ) 2016;1392:1-16.
10 Diehl F, Li M, Dressman D, He Y, Shen D, Szabo S, Diaz LA, Jr., Goodman SN, David KA, Juhl H, Kinzler KW, Vogelstein B: Detection and quantification of mutations in the plasma of patients with colorectal tumors. Proceedings of the National Academy of Sciences of the United States of America 2005;102:16368-16373.
11 Jiang P, Chan CW, Chan KC, Cheng SH, Wong J, Wong VW, Wong GL, Chan SL, Mok TS, Chan HL, Lai PB, Chiu RW, Lo YM: Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. Proceedings of the National Academy of Sciences of the United States of America 2015;112:E1317-1325.
12 Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, Knippers R: DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer research 2001;61:1659-1665.
13 Mouliere F, Robert B, Arnau Peyrotte E, Del Rio M, Ychou M, Molina F, Gongora C, Thierry AR: High fragmentation characterizes tumour-derived circulating DNA. PloS one 2011;6:e23418.
14 Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J: Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. Cell 2016;164:57-68.
15 Soo J, Kurtz DM, Jin M, Scherer F, Schroers-Martin JG, Esfahani MS, Craig A, Stehr H, Neelapu SS, Nastoupil L, Roschewski M, Wilson WH, Diehn M, Alizadeh AA:

Comparison of Circulating Tumor DNA Recovery from Plasma and Serum. Blood 2017;130:2756-2756.
16 Park JL, Kim HJ, Choi BY, Lee HC, Jang HR, Song KS, Noh SM, Kim SY, Han DS, Kim YS: Quantitative analysis of cell-free DNA in the plasma of gastric cancer patients. Oncol Lett 2012;3:921-926.
17 Molina-Vila MA, de-Las-Casas CM, Bertran-Alamillo J, Jordana-Ariza N, Gonzalez-Cao M, Rosell R: cfDNA analysis from blood in melanoma. Ann Transl Med 2015;3:309.
18 van Ginkel JH, van den Broek DA, van Kuik J, Linders D, de Weger R, Willems SM, Huibers MMH: Preanalytical blood sample workup for cell-free DNA analysis using Droplet Digital PCR for future molecular cancer diagnostics. Cancer Med 2017;6:2297-2307.
19 Botezatu I, Serdyuk O, Potapova G, Shelepov V, Alechina R, Molyaka Y, Ananev V, Bazin I, Garin A, Narimanov M, Knysh V, Melkonyan H, Umansky S, Lichtenstein A: Genetic analysis of DNA excreted in urine: a new approach for detecting specific genomic DNA sequences from cells dying in an organism. Clin Chem 2000;46:1078-1084.
20 Su YH, Wang M, Brenner DE, Ng A, Melkonyan H, Umansky S, Syngal S, Block TM: Human urine contains small, 150 to 250 nucleotide-sized, soluble DNA derived from the circulation and may be useful in the detection of colorectal cancer. J Mol Diagn 2004;6:101-107.
21 Pan W, Gu W, Nagpal S, Gephart MH, Quake SR: Brain tumor mutations detected in cerebral spinal fluid. Clin Chem 2015;61:514-522.
22 Togneri FS, Ward DG, Foster JM, Devall AJ, Wojtowicz P, Alyas S, Vasques FR, Oumie A, James ND, Cheng KK, Zeegers MP, Deshmukh N, O'Sullivan B, Taniere P, Spink KG, McMullan DJ, Griffiths M, Bryan RT: Genomic complexity of urothelial bladder cancer revealed in urinary cfDNA. Eur J Hum Genet 2016;24:1167-1174.
23 Xia Y, Huang CC, Dittmar R, Du M, Wang Y, Liu H, Shenoy N, Wang L, Kohli M: Copy number variations in urine cell free DNA as biomarkers in advanced prostate cancer. Oncotarget 2016;7:35818-35831.
24 Perdomo S, Avogbe PH, Foll M, Abedi-Ardekani B, Facciolla VL, Anantharaman D, Chopard P, Calvez-Kelm FL, Vilensky M, Polesel J, Holcatova I, Simonato L, Canova C, Lagiou P, McKay JD, Brennan P: Circulating tumor DNA detection in head and neck cancer: evaluation of two different detection approaches. Oncotarget 2017;8:72621-72632.
25 Lianidou E, Alix-Panabières C: Early detection of lung cancer based on DNA methylation analysis in sputum and plasma. Translational Cancer Research 2017:S51-S53.
26 De Mattos-Arruda L, Mayor R, Ng CKY, Weigelt B, Martínez-Ricarte F, Torrejon D, Oliveira M, Arias A, Raventos C, Tang J, Guerini-Rocco E, Martínez-Sáez E, Lois S, Marín O, de la Cruz X, Piscuoglio S, Towers R, Vivancos A, Peg V, Cajal SRy, Carles J, Rodon J, González-Cao M, Tabernero J, Felip E, Sahuquillo J, Berger MF, Cortes J, Reis-Filho JS, Seoane J: Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma. Nature Communications 2015;6:8839.
27 Ramalingam N, Jeffrey SS: Future of Liquid Biopsies With Growing Technological and Bioinformatics Studies: Opportunities and Challenges in Discovering Tumor Heterogeneity With Single-Cell Level Analysis. Cancer J 2018;24:104-108.
28 Cohen PA, Flowers N, Tong S, Hannan N, Pertile MD, Hui L: Abnormal plasma DNA profiles in early ovarian cancer using a non-invasive prenatal testing platform: implications for cancer screening. BMC Med 2016;14:126.
29 Bartak BK, Kalmar A, Galamb O, Wichmann B, Nagy ZB, Tulassay Z, Dank M, Igaz P, Molnar B: Blood Collection and Cell-Free DNA Isolation Methods Influence the Sensitivity of Liquid Biopsy Analysis for Colorectal Cancer Detection. Pathology oncology research : POR 2018

30 Markus H, Contente-Cuomo T, Farooq M, Liang WS, Borad MJ, Sivakumar S, Gollins S, Tran NL, Dhruv HD, Berens ME, Bryce A, Sekulic A, Ribas A, Trent JM, LoRusso PM, Murtaza M: Evaluation of pre-analytical factors affecting plasma DNA analysis. Sci Rep 2018;8:7375.
31 Devonshire AS, Whale AS, Gutteridge A, Jones G, Cowen S, Foy CA, Huggett JF: Towards standardisation of cell-free DNA measurement in plasma: controls for extraction efficiency, fragment size bias and quantification. Anal Bioanal Chem 2014;406:6499-6512.
32 Zonta E, Nizard P, Taly V: Assessment of DNA Integrity, Applications for Cancer Research. Adv Clin Chem 2015;70:197-246.
33 Han X, Wang J, Sun Y: Circulating Tumor DNA as Biomarkers for Cancer Detection. Genomics Proteomics Bioinformatics 2017;15:59-72.
34 Gale D, Lawson ARJ, Howarth K, Madi M, Durham B, Smalley S, Calaway J, Blais S, Jones G, Clark J, Dimitrov P, Pugh M, Woodhouse S, Epstein M, Fernandez-Gonzalez A, Whale AS, Huggett JF, Foy CA, Jones GM, Raveh-Amit H, Schmitt K, Devonshire A, Green E, Forshew T, Plagnol V, Rosenfeld N: Development of a highly sensitive liquid biopsy platform to detect clinically-relevant cancer mutations at low allele fractions in cell-free DNA. PloS one 2018;13:e0194630.
35 Bennett CW, Berchem G, Kim YJ, El-Khoury V: Cell-free DNA and next-generation sequencing in the service of personalized medicine for lung cancer. Oncotarget 2016;7:71013-71035.
36 Christensen E, Nordentoft I, Vang S, Birkenkamp-Demtroder K, Jensen JB, Agerbaek M, Pedersen JS, Dyrskjot L: Optimized targeted sequencing of cell-free plasma DNA from bladder cancer patients. Sci Rep 2018;8:1917.
37 Dobnik D, Stebih D, Blejec A, Morisset D, Zel J: Multiplex quantification of four DNA targets in one reaction with Bio-Rad droplet digital PCR system for GMO detection. Sci Rep 2016;6:35451.
38 Perakis S, Speicher MR: Emerging concepts in liquid biopsies. BMC Med 2017;15:75.
39 Sacher AG, Paweletz C, Dahlberg SE, Alden RS, O'Connell A, Feeney N, Mach SL, Janne PA, Oxnard GR: Prospective Validation of Rapid Plasma Genotyping for the Detection of EGFR and KRAS Mutations in Advanced Lung Cancer. JAMA Oncol 2016;2:1014-1022.
40 Whale AS, Fernandez-Gonzalez A, Gutteridge A, Devonshire AS: Control Materials and Digital PCR Methods for Evaluation of Circulating Cell-Free DNA Extractions from Plasma. Methods in molecular biology (Clifton, NJ) 2018;1768:45-65.
41 Merker JD, Oxnard GR, Compton C, Diehn M, Hurley P, Lazar AJ, Lindeman N, Lockwood CM, Rai AJ, Schilsky RL, Tsimberidou AM, Vasalos P, Billman BL, Oliver TK, Bruinooge SS, Hayes DF, Turner NC: Circulating Tumor DNA Analysis in Patients With Cancer: American Society of Clinical Oncology and College of American Pathologists Joint Review. J Clin Oncol 2018;36:1631-1641.
42 Heitzer E, Ulz P, Geigl JB: Circulating tumor DNA as a liquid biopsy for cancer. Clin Chem 2015;61:112-123.
43 Veldore VH, Choughule A, Routhu T, Mandloi N, Noronha V, Joshi A, Dutt A, Gupta R, Vedam R, Prabhash K: Validation of liquid biopsy: plasma cell-free DNA testing in clinical management of advanced non-small cell lung cancer. Lung Cancer (Auckl) 2018;9:1-11.
44 Kwapisz D: The first liquid biopsy test approved. Is it a new era of mutation testing for non-small cell lung cancer? Ann Transl Med 2017;5:46.
45 Sozzi G, Conte D, Leon M, Ciricione R, Roz L, Ratcliffe C, Roz E, Cirenei N, Bellomi M, Pelosi G, Pierotti MA, Pastorino U: Quantification of free circulating DNA as a diagnostic marker in lung cancer. J Clin Oncol 2003;21:3902-3908.
46 Shao X, He Y, Ji M, Chen X, Qi J, Shi W, Hao T, Ju S: Quantitative analysis of cell-free DNA in ovarian cancer. Oncol Lett 2015;10:3478-3482.

47 Hench IB, Hench J, Tolnay M: Liquid Biopsy in Clinical Management of Breast, Lung, and Colorectal Cancer. Front Med (Lausanne) 2018;5:9.
48 Bartoloni E, Ludovini V, Alunno A, Pistola L, Bistoni O, Crino L, Gerli R: Increased levels of circulating DNA in patients with systemic autoimmune diseases: A possible marker of disease activity in Sjogren's syndrome. Lupus 2011;20:928-935.
49 Klein-Scory S, Maslova M, Pohl M, Eilert-Micus C, Schroers R, Schmiegel W, Baraniskin A: Significance of Liquid Biopsy for Monitoring and Therapy Decision of Colorectal Cancer. Transl Oncol 2018;11:213-220.
50 Provencio M, Torrente M, Calvo V, Perez-Callejo D, Gutierrez L, Franco F, Perez-Barrios C, Barquin M, Royuela A, Garcia-Garcia F, Bueno C, Garcia-Grande A, Camps C, Massuti B, Sotomayor E, Romero A: Prognostic value of quantitative ctDNA levels in non small cell lung cancer patients. Oncotarget 2018;9:488-494.
51 Chu D, Paoletti C, Gersch C, VanDenBerg DA, Zabransky DJ, Cochran RL, Wong HY, Toro PV, Cidado J, Croessmann S, Erlanger B, Cravero K, Kyker-Snowman K, Button B, Parsons HA, Dalton WB, Gillani R, Medford A, Aung K, Tokudome N, Chinnaiyan AM, Schott A, Robinson D, Jacks KS, Lauring J, Hurley PJ, Hayes DF, Rae JM, Park BH: Mutations in Circulating Plasma Tumor DNA from Metastatic Breast Cancer Patients. Clinical Cancer Research 2016;22:993.
52 Guttery DS, Page K, Hills A, Woodley L, Marchese SD, Rghebi B, Hastings RK, Luo J, Pringle JH, Stebbing J, Coombes RC, Ali S, Shaw JA: Noninvasive detection of activating estrogen receptor 1 (ESR1) mutations in estrogen receptor-positive metastatic breast cancer. Clin Chem 2015;61:974-982.
53 Schiavon G, Hrebien S, Garcia-Murillas I, Cutts RJ, Pearson A, Tarazona N, Fenwick K, Kozarewa I, Lopez-Knowles E, Ribas R, Nerurkar A, Osin P, Chandarlapaty S, Martin LA, Dowsett M, Smith IE, Turner NC: Analysis of ESR1 mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer. Sci Transl Med 2015;7:313ra182.
54 Scholer LV, Reinert T, Orntoft MW, Kassentoft CG, Arnadottir SS, Vang S, Nordentoft I, Knudsen M, Lamy P, Andreasen D, Mortensen FV, Knudsen AR, Stribolt K, Sivesgaard K, Mouritzen P, Nielsen HJ, Laurberg S, Orntoft TF, Andersen CL: Clinical Implications of Monitoring Circulating Tumor DNA in Patients with Colorectal Cancer. Clin Cancer Res 2017;23:5437-5445.
55 Germano G, Mauri G, Siravegna G, Dive C, Pierce J, Di Nicolantonio F, D'Incalci M, Bardelli A, Siena S, Sartore-Bianchi A: Parallel Evaluation of Circulating Tumor DNA and Circulating Tumor Cells in Metastatic Colorectal Cancer. Clinical colorectal cancer 2018;17:80-83.
56 Hu W, Yang Y, Zhang L, Yin J, Huang J, Huang L, Gu H, Jiang G, Fang J: Post surgery circulating free tumor DNA is a predictive biomarker for relapse of lung cancer. Cancer Med 2017;6:962-974.
57 Heitzer E, Perakis S, Geigl JB, Speicher MR: The potential of liquid biopsies for the early detection of cancer. npj Precision Oncology 2017;1:36.
58 Remon J, Menis J, Hasan B, Peric A, De Maio E, Novello S, Reck M, Berghmans T, Wasag B, Besse B, Dziadziuszko R: The APPLE Trial: Feasibility and Activity of AZD9291 (Osimertinib) Treatment on Positive PLasma T790M in EGFR-mutant NSCLC Patients. EORTC 1613. Clin Lung Cancer 2017;18:583-588.
59 Knebel FH, Bettoni F, Shimada AK, Cruz M, Alessi JV, Negrao MV, Reis LFL, Katz A, Camargo AA: Sequential liquid biopsies reveal dynamic alterations of EGFR driver mutations and indicate EGFR amplification as a new mechanism of resistance to osimertinib in NSCLC. Lung cancer (Amsterdam, Netherlands) 2017;108:238-241.
60 Sun Q, Liu Y, Liu B, Liu Y: Use of Liquid Biopsy in Monitoring Colorectal Cancer Progression Shows Strong Clinical Correlation. Am J Med Sci 2018;355:220-227.
61 Kang S, Li Q, Chen Q, Zhou Y, Park S, Lee G, Grimes B, Krysan K, Yu M, Wang W, Alber F, Sun F, Dubinett SM, Li W, Zhou XJ: CancerLocator: non-invasive cancer diagnosis and tissue-of-origin prediction using methylation profiles of cell-free DNA. Genome Biol 2017;18:53.

62 Marquard AM, Birkbak NJ, Thomas CE, Favero F, Krzystanek M, Lefebvre C, Ferte C, Jamal-Hanjani M, Wilson GA, Shafi S, Swanton C, Andre F, Szallasi Z, Eklund AC: TumorTracer: a method to identify the tissue of origin from the somatic mutations of a tumor specimen. BMC Med Genomics 2015;8:58.
63 Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, Douville C, Javed AA, Wong F, Mattox A, Hruban RH, Wolfgang CL, Goggins MG, Dal Molin M, Wang TL, Roden R, Klein AP, Ptak J, Dobbyn L, Schaefer J, Silliman N, Popoli M, Vogelstein JT, Browne JD, Schoen RE, Brand RE, Tie J, Gibbs P, Wong HL, Mansfield AS, Jen J, Hanash SM, Falconi M, Allen PJ, Zhou S, Bettegowda C, Diaz LA, Jr., Tomasetti C, Kinzler KW, Vogelstein B, Lennon AM, Papadopoulos N: Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science 2018;359:926-930.
64 van Krieken JH, Normanno N, Blackhall F, Boone E, Botti G, Carneiro F, Celik I, Ciardiello F, Cree IA, Deans ZC, Edsjo A, Groenen PJ, Kamarainen O, Kreipe HH, Ligtenberg MJ, Marchetti A, Murray S, Opdam FJ, Patterson SD, Patton S, Pinto C, Rouleau E, Schuuring E, Sterck S, Taron M, Tejpar S, Timens W, Thunnissen E, van de Ven PM, Siebers AG, Dequeker E: Guideline on the requirements of external quality assessment programs in molecular pathology. Virchows Arch 2013;462:27-37.
65 Shen W, Zhao GQ, Liang Z, Yellapantula V, Sun S, Tang LFP, Lin S, Ying K, Fields RC, Ding L: Retrospective analysis of circulating tumor DNA of archived plasma samples. Journal of Clinical Oncology 2016;34:e23175-e23175.
66 Chan BA, Hughes BG: Targeted therapy for non-small cell lung cancer: current standards and the promise of the future. Transl Lung Cancer Res 2015;4:36-54.
67 Sorber L, Zwaenepoel K, Deschoolmeester V, Van Schil PE, Van Meerbeeck J, Lardon F, Rolfo C, Pauwels P: Circulating cell-free nucleic acids and platelets as a liquid biopsy in the provision of personalized therapy for lung cancer patients. Lung Cancer 2017;107:100-107.
68 Kohn L, Johansson M, Grankvist K, Nilsson J: Liquid biopsies in lung cancer-time to implement research technologies in routine care? Ann Transl Med 2017;5:278.
69 Perez-Callejo D, Romero A, Provencio M, Torrente M: Liquid biopsy based biomarkers in non-small cell lung cancer for diagnosis and treatment monitoring. Transl Lung Cancer Res 2016;5:455-465.
70 Cai X, Janku F, Zhan Q, Fan JB: Accessing Genetic Information with Liquid Biopsies. Trends Genet 2015;31:564-575.
71 Qiagen: Therascreen® EGFR RGQ PCR Kit Instructions for Use (Handbook). 2013
72 Kim SS, Choi HJ, Kim JJ, Kim MS, Lee IS, Byun B, Jia L, Oh MR, Moon Y, Park S, Choi JS, Chae SW, Nam BH, Kim JS, Kim J, Min BS, Lee JS, Won JK, Cho SY, Choi YL, Shin YK: Droplet digital PCR-based EGFR mutation detection with an internal quality control index to determine the quality of DNA. Sci Rep 2018;8:543.
73 Rothe F, Laes JF, Lambrechts D, Smeets D, Vincent D, Maetens M, Fumagalli D, Michiels S, Drisis S, Moerman C, Detiffe JP, Larsimont D, Awada A, Piccart M, Sotiriou C, Ignatiadis M: Plasma circulating tumor DNA as an alternative to metastatic biopsies for mutational analysis in breast cancer. Ann Oncol 2014;25:1959-1965.
74 Forshew T, Murtaza M, Parkinson C, Gale D, Tsui DW, Kaper F, Dawson SJ, Piskorz AM, Jimenez-Linan M, Bentley D, Hadfield J, May AP, Caldas C, Brenton JD, Rosenfeld N: Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 2012;4:136ra168.
75 Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, Liu CL, Neal JW, Wakelee HA, Merritt RE, Shrager JB, Loo BW, Jr., Alizadeh AA, Diehn M: An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med 2014;20:548-554.
76 Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B: Detection and quantification of rare mutations with massively parallel sequencing. Proc Natl Acad Sci U S A 2011;108:9530-9535.

77 Siravegna G, Marsoni S, Siena S, Bardelli A: Integrating liquid biopsies into the management of cancer. Nat Rev Clin Oncol 2017;14:531-548.