UNIVERSIDADE DE LISBOA

FACULDADE DE MEDICINA

GENOTYPING TUMOR DNA FOR PRECISION ONCOLOGY

Catarina Sofia Urbano Silveira de Sá

Orientador: Professora Doutora Maria do Carmo Salazar Velez Roque da Fonseca, Instituto de Medicina Molecular João Lobo Antunes e Faculdade de Medicina da Universidade de Lisboa

Co-orientador: Dra. Teresa Cristina de Almeida Porta Nova Soares Lopes, GenoMed - Diagnósticos de Medicina Molecular, S.A.

Tese especialmente elaborada para a obtenção do grau de Doutor em Ciências Biomédicas, Especialidade de Genética

2022

UNIVERSIDADE DE LISBOA

FACULDADE DE MEDICINA

GENOTYPING TUMOR DNA FOR PRECISION ONCOLOGY

Catarina Sofia Urbano Silveira de Sá

Orientador: Professora Doutora Maria do Carmo Salazar Velez Roque da Fonseca, Instituto de Medicina Molecular João Lobo Antunes e Faculdade de Medicina da Universidade de Lisboa

Co-orientador: Dra. Teresa Cristina de Almeida Porta Nova Soares Lopes, GenoMed - Diagnósticos de Medicina Molecular, S.A.

Tese especialmente elaborada para a obtenção do grau de Doutor em

Ciências Biomédicas, Especialidade de Genética

Júri:

Presidente: Doutor Mário Nuno Ramos de Almeida Ramirez, Professor Associado com Agregação e Vice-Presidente do Conselho Científico da Faculdade de Medicina da Universidade de Lisboa

Vogais:

- Doutor Fernando Carlos de Landér Schmitt, Professor Catedrático da Faculdade de Medicina da Universidade do Porto
- Doutor Manuel António Rodrigues Teixeira, Professor Catedrático Convidado do Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.
- Doutora Joana Cancela de Amorim Falcão Paredes, Investigadora Principal do Instituto de Investigação e Inovação em Saúde da Universidade do Porto
- Doutora Maria do Carmo Salazar Velez Roque da Fonseca, Professora Catedrática da Faculdade de Medicina da Universidade de Lisboa (Orientadora)
- Doutor João António Augusto Ferreira, Professor Associado da Faculdade de Medicina da Universidade de Lisboa
- Doutor João Pedro Taborda Barata, Professor Associado da Faculdade de Medicina da Universidade de Lisboa

Instituições Financiadoras:

- Fundação para a Ciência e a Tecnologia – FCT (SFRH/BDE/110544/2015 e LISBOA-01-0145-FEDER-016394, projeto co-financiado por FEDER, POR Lisboa 2020: Programa Operacional Regional de Lisboa PORTUGAL 2020)
- GenoMed - Diagnósticos de Medicina Molecular, S.A., Lisboa, Portugal
- Memorial Sloan Kettering Cancer Center, Nova Iorque, EUA
- European Union's Horizon 2020-MSCA-RISE, grant agreement No 734825 (LysoMod)
- FEDER/POR Lisboa 2020-Programa Operacional Regional de Lisboa, PORTUGAL 2020 (Infogene, 045300)

2022

A impressão desta tese foi aprovada pelo Conselho Científico da Faculdade de Medicina da Universidade de Lisboa em reunião de 21 de Dezembro de 2021.

As opiniões expressas nesta publicação são da exclusiva responsabilidade do seu autor.

Preface

The present thesis, entitled ***Genotyping tumor DNA for Precision Oncology***, contains results of a 5 years' work, developed at GenoMed - *Diagnósticos de Medicina Molecular*, S.A., *Lisboa*, *Portugal*. This project was developed under the supervision of Professora Doutora Maria Carmo-Fonseca and Dra. Teresa Porta Nova.

This dissertation is organized in four main chapters, preceded by a list of publications, a list of abbreviations, Portuguese and English summaries, and the corresponding keywords.

The first chapter introduces the concepts of precision oncology and precision medicine and presents an overview of the different types of methodologies that are available for this new *era* of standard care in Oncology.

The second chapter states the objectives of this work.

The results obtained are presented in the third chapter. This chapter is subdivided into three subchapters, which represent the three main achievements of my research. The first subchapter, ***Detection of rare and novel EGFR mutations in NSCLC patients: Implications for treatment-decision***, highlights the limitations of targeted-methods. In this study, we found that approximately 3% of NSCLC patients have uncommon genetic alterations in the *EGFR* gene, which could be missed if targeted-methods were used. Moreover, six new single mutations, previously unidentified in NSCLC patients, were reported, as well as seven novel combinations of doublet mutations. The second subchapter, ***Detection and quantification of EGFR T790M mutation in liquid biopsies by droplet digital PCR***, describes the optimization of an ultra-sensitive method that allows mutation detection in a cfDNA allele frequency as low as 0.5%. The results suggest that repeated ddPCR tests in cfDNA may obviate tissue re-biopsy in patients unable to provide a tumor tissue sample. The third subchapter, ***Comprehensive Genomic Profiling of Cell-Free Circulating Tumor DNA Detects Response to Ribociclib Plus Letrozole in a Patient with Metastatic Breast Cancer***, highlights the utility of cfDNA analysis by massively parallel sequencing for therapy monitoring in metastatic breast cancer patients.

The fourth chapter includes a general discussion, addressing pros and cons of the main methodologies currently available, as well as trends for future directions.

Table of contents

Acknowledgments

This project was supported by *Fundação para a Ciência e a Tecnologia* (SFRH/BDE/110544/2015 and LISBOA-01-0145-FEDER-016394, project co-funded by FEDER, POR Lisboa 2020: *Programa Operacional Regional de Lisboa* PORTUGAL 2020), *GenoMed - Diagnósticos de Medicina Molecular, SA*, Memorial Sloan Kettering Cancer Center - Department of Pathology, and European Union's Horizon 2020-MSCA-RISE (grant agreement No 734825, LysoMod). I am also thankful to iMM-Laço for support in the early stages of this project.

I would like to thank all the people who helped me throughout this project:

My first word of gratitude goes to my outstanding supervisor, Professor Maria Carmo-Fonseca. She is the principal responsible for this work, thanks for the time, patience, and precious help. Thank you for believing in me, more than I do.

To my co-supervisor, Dr. ª Teresa Porta Nova, who always gave me support to carry on, even when I doubted.

A big thanks to Jorge S. Reis-Filho, Britta Weigelt and, what I like to call, their *dream team*, in particular Arnaud, Fresia, Ana Paula, and Edaise. Thank you for letting me be part of your team. Thank you for trusting in me. The time that I spent with you was one of the most important periods of my journey.

I would like to state and recognize the importance of ALL my colleagues in the lab, during the last 5 years. Their support was fundamental for this work, mainly: Ana Carla Sousa, André Janeiro, Sara Malveiro, Ana Coutinho, Ana Rita Dario, Ana Flores, and Fernanda Barbosa.

To my PhD advisory committee: Professor Luís Costa, Professor João Barata and Professor João Ferreira. Thank you for your support and advice.

Big thanks to these VIP, which stood by my side during a big part of this journey: Célia Carvalho, Dra. Ana Rita Sousa, Dra. Inês Vendrell, Dra. Catarina Pulido, Patrícia Corredeira, Marta Martins.

To all the patients who were tested, allowing me to achieve these results. To all the medical doctors who trusted our diagnostic testing team.

I would also like to thank my family (especially my parents and my parents-in-law), and friends, for all your support during these troubled years. Thank you for believing in me.
Last but not least, a BIG thank you to my husband. Thank you for all the help in these last years. For your patience and support. For being there for me in every step of the way, and never letting me fall. You are my better half, and I am so lucky to have you.
Thank you all for being by my side.

♥

I want to dedicate this work to all my stars that are lighting my way: tia Miquelina, tio Francisco, Mónica, Aida, Avô Zé, Cátia, Avó Delmira.

♥

Para ti, Benjamin ♥

Publications list

First author

Silveira C, Sousa AC, Corredeira P, Martins M, Sousa AR, Da Cruz Paula A, Selenica P, Brown DN, Golkaram M, Kaplan S, Zhang S, Liu L, Weigelt B, Reis-Filho JS, Costa L, Carmo-Fonseca M. *Comprehensive Genomic Profiling of Cell-Free Circulating Tumor DNA Detects Response to Ribociclib Plus Letrozole in a Patient with Metastatic Breast Cancer.* Biomolecules 2022, 12,1818. PMID: 36551247. DOI: 10.3390/biom12121818.

First co-author

Silveira C[*], Sousa AC[*], Janeiro A[*], Malveiro S, Teixeira E, Brysch E, Pantarotto M, Felizardo M, Madureira R, Nogueira F, Guimarães C, Matos C, Canário D, Bruges-Armas J, Carmo-Fonseca M. *Detection and quantification of EGFR T790M mutation in liquid biopsies by droplet digital PCR*. Transl Lung Cancer Res. 2021 Mar;10(3):1200-1208. PMID: 33889502. DOI: 10.21037/tlcr-20-1010.

Sousa AC[*], **Silveira C**[*], Janeiro A, Malveiro S, Oliveira AR, Felizardo M, Nogueira F, Teixeira E, Martins J, Carmo-Fonseca M. *Detection of rare and novel EGFR mutations in NSCLC patients: Implications for treatment-decision.* Lung Cancer. 2020 Jan;139:35-40. PMID: 31715539. DOI: 10.1016/j.lungcan.2019.10.030.

Co-author

Guo K, Shao C, Han L, Liu H, Ma Z, Yang Y, Feng Y, Pan M, Santarpia M, Carmo-Fonseca M, **Silveira C**, Lee KY, Han J, Li X, Yan X. *Detection of epidermal growth factor receptor (EGFR) mutations from preoperative circulating tumor DNA (ctDNA) as a prognostic predictor for stage I-III non-small cell lung cancer (NSCLC) patients with baseline tissue EGFR mutations*. Transl Lung Cancer Res. 2021 Jul;10(7):3213-3225. PMID: 34430359. DOI: 10.21037/tlcr-21-530.

da Silva EM, Beca F, Sebastiao APM, Murray M, **Silveira C**, Da Cruz Paula A, Pareja F, Wen HY, D'Alfonso T, Edelweiss M, Weigelt B, Brogi E, Reis-Filho JS, Zhang H. *MED12 Exon 2 Mutations in Complex Fibroadenomas of the Breast*. Journal of Clinical Pathology. J Clin Pathol. 2020 Dec29;jclinpath-2020-207062. PMID: 33376197. DOI: 10.1136/jclinpath-2020-207062.

Kim SH, Da Cruz Paula A, Basili T, Dopeso H, Bi R, Pareja F, da Silva EM, Gularte-Mérida R, Sun Z, Fujisawa S, Smith CG, Ferrando L, Martins Sebastião AP, Bykov Y, Li A, **Silveira C**, Ashley CW, Stylianou A, Selenica P, Samore WR, Jungbluth AA, Zamarin D, Abu-Rustum NR, Helin K, Soslow RA, Reis-Filho JS, Oliva E, Weigelt B. *Identification of recurrent FHL2-GLI2 oncogenic fusion in sclerosing stromal tumors of the ovary*. Nat Commun. 2020 Jan2;11(1):44. PMID: 31896750. DOI: 10.1038/s41467-019-13806-x.

Marchiò C, Da Cruz Paula A, Gularte-Mérida R, Basili T, Brandes A, da Silva EM, **Silveira C**, Ferrando L, Metovic J, Maletta F, Annaratone L, Pareja F, Rubin BP, Hoschar AP, De Rosa G, La Rosa S, Bongiovanni M, Purgina B, Piana S, Volante M, Weigelt B, Reis-Filho JS & Papotti M. *PAX8–GLIS3 gene fusion is a pathognomonic genetic alteration of hyalinizing trabecular tumors of the thyroid*. Mod Pathol. 2019 Dec;32(12):1734-1743. PMID: 31273314. DOI: 10.1038/s41379-019-0313-x.

Sebastiao APM, Pareja F, Kumar R, Brown DN, **Silveira C**, da Silva EM, Lee JY, Katabi N, Chiosea S, Weigelt B, Reis-Filho JS, Seethala RR. *Genomic analysis of recurrences and high-grade forms of polymorphous adenocarcinomas*. Histopathology. 2019 Aug;75(2):193-201. PMID: 30843621. DOI: 10.1111/his.13854.

Pareja F, Lee JY, Brown DN, Piscuoglio S, Gularte-Mérida R, Selenica P, Da Cruz Paula A, Arunachalam S, Kumar R, Geyer FC, **Silveira C**, da Silva EM, Li A, Marchiò C, Ng CKY, Mariani O, Fuhrmann L, Wen HY, Norton L, Vincent-Salomon A, Brogi E, Reis-Filho JS, Weigelt B. *The genomic landscape of mucinous breast cancer.* J Natl Cancer Inst. 2019 Jan 11. PMID: 30649385. DOI: 10.1093/jnci/djy216.

Abbreviations

AMP	Association for Molecular Pathology
BC	Breast cancer
bp	base pair
BRAF	B-Raf proto-oncogene, serine/threonine kinase
BRCA	Breast cancer genes
BRCA1	BRCA1 DNA repair associated
BRCA2	BRCA2 DNA repair associated
CAP	College of American Pathologists
CCF	cancer cell fraction
CCND1	cyclin D1
cfDNA	circulating cell-free DNA
cfNAs	cell-free nucleic acids
CNA	copy number alteration
CNR	contrast-to-noise ratio
CNS	central nervous system
CNV	copy number variation
COSMIC	Catalogue of Somatic Mutations in Cancer
CR	complete response
CT	computerized tomography
CTCs	circulating tumor cells
ctDNA	circulating tumor DNA
CTX	chemotherapy
ddPCR	droplet digital PCR
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
dPCR	digital PCR
dsDNA	double-strand DNA
EDTA	ethylenediamine tetra-acetic acid

EGFR	epidermal growth factor receptor
EGFR-TKIs	EGFR-tyrosine kinase inhibitors
ER	estrogen receptor
ERBB2	erb-b2 receptor tyrosine kinase 2
ESMO	European Society for Medical Oncology
FAM	fluorescein amidite
FFPE	formalin-fixed paraffin-embedded
FGFR1	fibroblast growth factor receptor 1
FISH	fluorescent in-situ hybridization
GATA3	GATA binding protein 3
gDNA	germline DNA
H&E	hematoxylin and eosin
HER2	ERBB2 alias
HEX	hexachloro-6-carboxy-fluorescine
HS	High Sensitivity
IASLC	International Association for the Study of Lung Cancer
indels	small insertions and deletions
LOH	loss of heterozygosity
$MgCl_2$	magnesium chloride
mL	millilitre
MPS	Massively Parallel Sequencing
MSK-IMPACT	Integrated Mutation Profiling of Actionable Cancer Targets
MTB	Molecular Tumor Board
MYC	MYC proto-oncogene, bHLH transcription factor
NA	not available or not assessed
NCCN	National Comprehensive Cancer Network
NGS	Next-Generation Sequencing
NSCLC	Non-small cell lung cancer
NTC	No template control
ORR	overall response rates

OS	overall survival
PCR	polymerase chain reaction
PD	progression of disease
PFS	progression-free survival
PgR	progesterone receptor
PIK3CA	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha
PR	partial response
PS	performance status
PTEN	phosphatase and tensin homolog
QA	Quality Assessment
qPCR	quantitative PCR
RNA	ribonucleic acid
RT	radiotherapy
SBS	sequencing by synthesis
SD	stable disease
SNVs	single-nucleotide variants
Ta	annealing temperature
TKIs	tyrosine kinase inhibitors
TNBC	triple-negative breast cancer
TP53	tumor protein p53
VAF	variant allele frequency
WES	whole-exome sequencing
WGS	whole-genome sequencing
WT	wild-type

Resumo

O início do século XXI foi marcado por uma revolução genómica nas áreas da Biologia e da Medicina. Com a conclusão do Projeto do Genoma Humano, em 2003, inúmeras inovações tecnológicas surgiram, levando à descodificação de toda a informação genética, na saúde e na doença, e iniciando o caminho na era da Medicina de Precisão. Em particular, esta revolução genómica modificou a investigação na área do cancro e transformou o nosso conhecimento e entendimento sobre a origem do cancro e a sua biologia. Nos últimos anos, temos assistido à reclassificação de vários tumores, com base no seu perfil genómico, e inúmeros fármacos têm vindo a ser desenvolvidos, com o objetivo de atuar num alvo molecular específico de um respetivo tumor – mutações acionáveis.

Para o sucesso da Oncologia de Precisão, é crucial que disponhamos de ensaios que sejam sensíveis e precisos na obtenção da assinatura genómica dos tumores. Além disso, e para além da genotipagem do tecido tumoral, tanto do tumor primário como das suas metástases, metodologias que permitam realizar genotipagem tumoral em ácidos nucleicos extraídos a partir de sangue periférico e de outros fluidos corporais, têm atraído muitas atenções.

Este projeto foi iniciado com o objetivo de implementar e validar metodologias inovadoras, na genotipagem de ADN tumoral, para o uso na prática clínica.

Numa primeira abordagem, recorremos à sequenciação de Sanger para determinar a frequência com que mutações raras ou novas, no gene que codifica o recetor do fator de crescimento epidérmico (*EGFR*), não seriam detetadas por metodologias baseadas na técnica de PCR, como seja o caso do *kit Idylla™ EGFR Mutation Assay*. Em doentes com carcinoma do pulmão de não pequenas células (CPNPC), a pesquisa de mutações no gene *EGFR* é realizada na prática clínica, pois estas mutações são consideradas biomarcadores preditivos de resposta às terapias dirigidas a este gene. Estas terapias são conhecidas como inibidores da tirosina cinase (EGFR-TKI). Assim, a genotipagem do domínio tirosina cinase do gene *EGFR* fornece informação crucial na escolha e decisão terapêutica para estes doentes. Rotineiramente são usadas duas principais metodologias na genotipagem deste gene: a sequenciação de Sanger e o PCR em Tempo Real. No entanto, as metodologias

baseadas na técnica de PCR, como o PCR em Tempo Real, estão desenhadas para detetar apenas mutações frequentes, não detetando mutações raras ou novas. Neste estudo, o nosso objetivo passou por determinar a prevalência de mutações novas ou raras no domínio tirosina cinase do gene *EGFR* (exões 18 a 21), mutações essas que não constam da lista de possíveis mutações a detetar se recorrermos aos *kits* comerciais mais comummente usados, como o cobas® *EGFR Mutation Test* e o Idylla™ *EGFR Mutation Assay*. Com recurso à sequenciação completa dos exões do domínio tirosina cinase do gene *EGFR*, foram estudados um conjunto de 1228 doentes com CPNPC. Observámos que 252 doentes (~20%) apresentavam pelo menos uma mutação nas regiões estudadas do gene *EGFR* e que 38 doentes (~3%) apresentavam variantes genéticas raras que poderiam não ser identificadas pelos *kits* comerciais cobas® ou Idylla™. Adicionalmente foram identificadas seis novas mutações e sete mutações compostas (conjuntos de duas mutações) que não tinham sido reportadas previamente na literatura. A informação clínica e o resultado dos tratamentos aplicados a estes doentes foi apresentada. Em conclusão, este estudo demonstra a importância das metodologias de sequenciação na identificação de mutações raras ou novas. Adicionalmente, importa referir que, em muitas situações como seja, por exemplo, a presença de inserções ou deleções, apenas recorrendo às técnicas de sequenciação conseguimos apurar que mutação está realmente presente, sendo que essa informação poderá ser determinante na decisão terapêutica a aplicar àquele doente. Os nossos resultados acrescentam novas mutações ao reportório de mutações descritas no gene *EGFR*, contribuindo assim para a melhoria do tratamento de precisão em doentes com cancro do pulmão.

Após a deteção de uma mutação sensível aos EGFR-TKI, ao diagnóstico inicial, e após tratamento com inibidores de 1ª ou 2ª geração, cerca de metade dos tumores desenvolvem uma mutação de resistência. Sendo imprescindível identificar essa mutação para que o tratamento possa ser reajustado, é necessário que a amostra e o método sejam cuidadosamente escolhidos. Assim, o próximo passo foi a implementação e avaliação de desempenho de um ensaio de dr*oplet digital* PCR (ddPCR) para a deteção da conhecida mutação de resistência, no gene *EGFR*, T790M, em biópsias líquidas. Com recurso à biópsia líquida, é possível identificar mutações acionáveis em cancro, de forma minimamente invasiva. Em doentes com CPNPC, a técnica de ddPCR é amplamente usada para a

genotipagem do gene *EGFR* no ADN livre circulante (cfDNA). No entanto, a sensibilidade desta abordagem ainda não está totalmente esclarecida. Foi otimizado o ensaio para a pesquisa da mutação T790M no gene *EGFR* por ddPCR, e testámos 77 amostras de plasma de doentes com CPNPC em progressão de doença. Este ensaio permitiu-nos detetar e quantificar a mutação T790M no gene *EGFR* em cfDNA com frequência alélica até 0.5%. Esta mutação foi detetada em 40 amostras de plasma, correspondendo a uma taxa de positividade de 52%. O número de moléculas mutadas por mililitro de plasma variou entre 1 e 6.000. Em 12 doentes que apresentaram um resultado negativo no plasma, testámos uma rebiópsia de tecido, tendo sido identificada a mutação em dois doentes. Em 6 doentes com resultado negativo no plasma, foi repetida a biópsia líquida e a mutação foi identificada em 3 casos. Em conclusão, este estudo demonstra a capacidade da técnica de ddPCR em detetar e quantificar a mutação T790M no gene *EGFR* em biópsias líquidas, na prática clínica. Os nossos resultados demonstram ainda que a repetição do estudo numa segunda amostra de plasma pode obviar uma rebiópsia tecidular, nomeadamente em doentes em que não seja possível aceder a uma amostra de tecido adequada à análise molecular.

Apesar das vantagens acima descritas, o ddPCR tem como grande desvantagem a dificuldade em testar vários alvos em simultâneo. Como tal, foi explorada a aplicabilidade da tecnologia de sequenciação massiva em paralelo na identificação de mutações em cfDNA extraído a partir de sangue periférico. Foi selecionada uma doente com cancro da mama metastático, que tinha sido elegida para participar num ensaio clínico de fase IIIb que consistia na administração de tratamento combinado com ribociclib e letrozol. Foram testadas duas amostras de cfDNA isolado a partir de sangue periférico colhidas antes (Pre_cfDNA) e depois (Post_cfDNA) do tratamento combinado com ribociclib e letrozol. A análise da amostra Pre_cfDN, feita com recurso ao painel *Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets* (MSK-IMPACT), revelou duas alterações: uma mutação *missense* no gene *PIK3CA* (3:178952085, A>G, H1047R) e um aumento do número de cópias (amplificação) do gene *CCND1*. A presença destas alterações foi posteriormente confirmada numa amostra de tecido do tumor primário desta doente. Após o início do tratamento, a doente apresentou melhorias clínicas significativas e ambas as alterações deixaram de ser detetadas em amostra de cfDNA (Post_cfDNA). Este trabalho

mostra que a natureza minimamente invasiva da biópsia líquida permite colher amostras seriadas e monitorizar a progressão da doença oncológica, representando uma alternativa prática à biópsia tecidular. O recurso à sequenciação massiva em paralelo permite interrogar vários genes simultaneamente, pesquisar vários tipos de mutação e identificar variantes com baixa frequência alélica.

Com base nos resultados obtidos no ponto anterior, decidimos prosseguir com a implementação de um ensaio de ddPCR dirigido à pesquisa de mutações no gene *PIK3CA* em cfDNA. Este ensaio foi utilizado para testar cfDNA em amostras de urina da doente com a mutação *PIK3CA* detetada previamente no plasma, mas não foi possível obter nenhum resultado positivo, ou seja, concordante com o resultado obtido no plasma.

Em conclusão, na GenoMed foram implementados os principais métodos que são hoje em dia recomendados no diagnóstico molecular aplicado à Oncologia de Precisão, desde a sequenciação massiva em paralelo ao ddPCR. Foram apresentadas as principais vantagens e desvantagens de cada uma das tecnologias, aplicadas a uma variedade de amostras biológicas, desde tecido tumoral a cfDNA obtido a partir de sangue e outros fluidos. Em conjunto, os resultados deste trabalho representam uma importante contribuição no sentido de caminhar para uma Medicina de Precisão mais acessível e menos dispendiosa para que possa ser aplicada na prática clínica.

Palavras-chave

Diagnóstico molecular; biópsia líquida; ddPCR; sequenciação massiva em paralelo; cancro do pulmão; cancro da mama.

Summary

The genomic revolution marked the beginning of the 21st century in biology and medicine. Completion of the Human Genome Project in 2003 has brought a flood of technological innovations that enabled decoding the entire genetic information in health and disease, leading to the foundation of the precision medicine movement. Genomics has revolutionized cancer research and transformed our understanding of how cancer arises. In recent years cancers have been re-classified based on their mutations, and a multitude of new drugs were developed that target specific molecular features of the tumor.

Accurate and sensitive assays for cancer genotyping are crucial to enable precision oncology. Moreover, in addition to genotyping tissue samples from the primary tumor and metastasis, methodologies for cancer genotyping in circulating blood and other biological fluids are attracting much attention.

This project was initiated with the goal of implementing and validating cutting-edge methodological approaches for profiling tumor DNA in the clinic.

First, we used Sanger sequencing to determine the frequency of rare mutations in the gene that encodes epidermal growth factor receptor (*EGFR*) that are not detected by the widely used PCR-based Idylla™ EGFR Mutation Assay. Mutations in the *EGFR* gene are biomarkers that predict how non-small cell lung cancer (NSCLC) patients respond to EGFR-targeted therapies collectively known as tyrosine kinase inhibitors (TKIs). Thus, *EGFR* genotyping provides crucial information for treatment decision. Both Sanger sequencing and real-time PCR methodologies are used for *EGFR* genotyping. However, methods based on real-time PCR have limitations, as they may not detect rare or novel mutations. We sought to determine the prevalence of rare mutations in the tyrosine kinase domain (exons 18 to 21) of the *EGFR* gene not targeted by the most frequently used real-time PCR approaches, i.e., the cobas® EGFR Mutation Test, and the Idylla™ EGFR Mutation Assay. A total of 1228 NSCLC patients were screened for mutations in exons 18 to 21 of the *EGFR* gene using Sanger sequencing. We observed that 252 patients (~20%) had at least one mutation in the *EGFR* gene, and 38 (~3%) carried uncommon genetic alterations that could not be identified by the cobas® or the Idylla™ tests. We further found six new single mutations and seven

previously unreported compound mutations. Clinical information and patient outcome are presented for these cases. In conclusion, this study highlights the value of sequencing-based approaches to identify rare mutations. Our results add to the inventory of known *EGFR* mutations, thus contributing to improved lung cancer precision treatment.

After the detection of an EGFR-TKI-sensitive mutation, at initial diagnosis, and after treatment with 1st or 2nd generation inhibitors, about half of tumors develop a resistance mutation. To identify this mutation, so that the treatment can be readjusted, it is mandatory to carefully choose the sample and the method to be used. Therefore, we implemented and assessed the performance of a droplet digital PCR (ddPCR) assay for detecting the *EGFR* T790M mutation in liquid biopsies. Liquid biopsy allows the identification of targetable cancer mutations in a minimally invasive manner. In patients with NSCLC, ddPCR is increasingly used to genotype the *EGFR* gene in circulating cell-free DNA (cfDNA). However, the sensitivity of this method is still under debate. We optimized a ddPCR assay and used it to detect the *EGFR* T790M mutation in plasma samples from 77 patients with NSCLC in progression. Our ddPCR assay enabled the detection and quantification of the *EGFR* T790M mutation at cfDNA allele frequency as low as 0.5%. The mutation was detected in 40 plasma samples, corresponding to a positivity rate of 52%. The number of mutant molecules per mL of plasma ranged from 1 to 6,000. A re-biopsy was analyzed for 12 patients that had a negative plasma test and the mutation was detected in 2 cases. A second liquid biopsy was performed for 6 patients and the mutation was detected in 3 cases. In conclusion, this study highlights the value of ddPCR to detect and quantify the *EGFR* T790M mutation in liquid biopsies in a real-world clinical setting. Our results suggest that repeated ddPCR tests in cfDNA may obviate tissue re-biopsy in patients unable to provide a tumor tissue sample suitable for molecular analysis.

Despite the advantages above described, the major disadvantage of ddPCR is the difficulty in testing multiple targets simultaneously. As such, we explored massively parallel sequencing methodologies for mutation discovery in cfDNA extracted from circulating blood. We analysed a patient with metastatic breast cancer who was selected for enrolment in an open-label clinical phase IIIb trial with ribociclib combined with letrozole. We genotyped cfDNA isolated from blood samples collected before (Pre_cfDNA) and after (Post_cfDNA) treatment with ribociclib and letrozole. Analysis of Pre_cfDNA using the

Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT) revealed two alterations: a missense hotspot mutation in the *PIK3CA* gene (3:178952085, A>G, H1047R) and a copy number amplification of the *CCND1* gene. The presence of both molecular alterations was confirmed in the primary tumor. After treatment, the patient presented a significant clinical improvement, and the two alterations were no longer detected in cfDNA (Post_cfDNA). This work underlines that the minimally invasive nature of liquid biopsy allows monitoring of disease progression, demonstrating a practical alternative to tissue biopsy. The use of massively parallel sequencing enables to interrogate of several genes simultaneously, the search for different types of mutation, and identify variants with low allele frequency.

Based on the previously mentioned results, we decided to implement a targeted ddPCR assay for detecting *PIK3CA* mutations in cfDNA. We used this assay to screen urine samples from the patient with a *PIK3CA* mutation identified in the plasma, but we failed to detect any positive result in the urine.

In conclusion, we implemented at GenoMed the main methods that are currently recommended for molecular diagnosis in Precision Oncology, from massively parallel sequencing to ddPCR. We characterized the main advantages and limitations of each technology applied to a variety of biological samples including tumor tissue, blood, and other fluids. Taken together, the results of our work represent an important contribution towards a more achievable and cost-effective Precision Medicine in the context of real-world clinical practice.

Keywords

Molecular diagnosis; liquid biopsy; ddPCR; massively parallel sequencing; lung cancer; breast cancer.

1. Introduction

Cancers are among the leading causes of death worldwide, and their incidence is increasing as the population ages. Molecular lesions that drive oncogenesis have been identified and therapeutic strategies were designed to block the effect of specific cancer drivers. Currently, many successful cancer therapies are targeted to a specific molecular lesion, and most recently developed therapeutic approaches are only effective if the cancer has a particular molecular signature. As treatments become more precisely tailored to specific molecular targets, the concept of "Precision Medicine" has emerged (Carmo-Fonseca, 2021). At present, several studies are looking for genetic abnormalities in patients' tumors. Based on the genetic results, patients with advanced tumors that have stopped responding to standard therapy are enrolled into a clinical trial of an investigational drug that might target the specific molecular pathology driving their cancer (Carmo-Fonseca, 2021).

Accurate and sensitive assays for cancer genotyping are crucial to enable precision oncology. Moreover, in addition to genotyping tissue samples from the primary tumor and metastasis, methodologies for cancer genotyping in circulating blood and other biological fluids are attracting much attention.

1.1 Precision Medicine

Precision or Personalized Medicine is based upon the idea of adapting personalized healthcare choices to each patient, according to some patient's characteristics (Bidard, Weigelt and Reis-Filho, 2013; Carmo-Fonseca, 2021).

In recent years, the Oncology field has been the area where this personalization has gained more power. Precision Oncology is the use of tumor molecular profiles to determine diagnostic, prognostic, and therapeutic implications for specific cancer under study (Schwartzberg *et al.*, 2017). This concept relies on the hypothesis that tumor biomarkers can predict disease phenotype, clinical outcomes, and therapeutic responses. Imatinib for chronic myeloid leukemia was the first cancer precision medicine under this model (Prasad, Fojo and Brada, 2016; Li and Warner, 2020). Later, Dagher R. *et al* showed that imatinib

was effective in gastrointestinal stromal tumors, a previously untreatable cancer (Dagher *et al.*, 2002).

The identification of actionable alterations marked a new era in oncology, being seen as a way to personalize treatment (Hyman, Taylor and Baselga, 2017; Gambardella *et al.*, 2020; Murciano-Goroff *et al.*, 2020). As result, patients have greater efficacy and less toxicity and avoids pointless treatments, increasing the chances of survival (Hey *et al.*, 2020; Murciano-Goroff *et al.*, 2020). The main goal is to perform a treatment aimed at a given molecular change – actionable mutation, usually in a gene, which is a tumor driver. This allows for a more specific treatment choice and possibly with fewer adverse effects, compared to treatments that do not target specific molecular changes, as will be the case with chemotherapy (CTX).

Another example is the use of EGFR-tyrosine kinase inhibitors (EGFR-TKIs), namely gefitinib, erlotinib, afatinib and osimertinib, which showed efficacy on patients with L858R mutations or exon 19 deletions on the *EGFR* gene. These mutations increase receptor dimerization and decrease ATP binding, increasing inhibitor affinity compared to wild-type *EGFR*. On the other hand, EGFR-TKIs are less effective against exon 20 insertions than against wild-type *EGFR*, limiting their tolerability (Murciano-Goroff *et al.*, 2020).

Even with the increasing availability of new drugs that are developed to target specific molecular features of the tumor, targeted therapies still benefit only a small percentage of cancer patients (Carmo-Fonseca, 2021). Therefore, it will be necessary to develop diagnostic tests that allow for the identification of the actionable mutations in clinical practice and to have evidence in clinical trials on the effectiveness of the treatments. Genomic sequencing can identify genetic variants that could lead to new treatments for cancer. A panel of specialists who will decide if the treatment is relevant and eligible must interpret the results. Patients with rare tumors or those who have already exhausted all treatment possibilities may perform genomic sequencing to look for these actionable mutations.

Patients would ideally be enrolled in one of the many precision medicine trials available, or they may receive off-label targeted medication for a specific mutation if adequate proof existed. A molecular tumor board can help the clinician by making therapy recommendations based on the existing literature and interdisciplinary knowledge.

Because laboratory reporting, read depth, tissue sources, assay type, and many other criteria may vary, the molecular tumor boards are increasingly useful to understand these highly specialized molecular testing data (Johnson, 2017).

There are some challenges that must be handled. The fact that the unit of testing in precision medicine is a complicated intervention ensemble that includes a therapeutic agent, a marker, and a diagnostic assay for identifying that marker presents one challenge. Each component of this complex intervention ensemble - therapy, marker, and assay - must be tailored for a specific condition in order to pass rigorous testing (Hey *et al.*, 2020). A second challenge is that diagnostic assays require their own multistep development and validation process, which includes assessing the assay's pre-analytical validity (proper specimen handling and processing), analytical validity (test accuracy, reliability, and reproducibility), and clinical validity (strong association between the test result and a clinical outcome of interest). Failure to complete validation process can lead to low diagnostic technology adoption, patient misclassification, poor patient outcomes, and wasteful healthcare spending (Hey *et al.*, 2020).

1.2 Liquid Biopsies

One of the main challenges of precision oncology is tumor sampling and to avoid sampling bias. In fact, in some patients, it may be difficult to provide a tumor tissue sample suitable for molecular analysis, either because of the patient's status or because of the tumor's location (inaccessible or unknown). Furthermore, tumor biopsies do not represent the overall disease (tumor heterogeneity or metastasis). For these reasons, and knowing that tumor cells shed genetic material into the circulation, emerged the opportunity of using the so-called "liquid biopsy" approach to access tumor genetics and its evolution (Bidard, Weigelt and Reis-Filho, 2013; Siravegna and Bardelli, 2014; Rowlands *et al.*, 2019; Siravegna *et al.*, 2019). Pantel and Alix-Panabières, in 2010, coined the term "liquid biopsy" (Pantel and Alix-Panabières, 2010).

Mandel and Métais were the pioneers in referring to the presence of cell-free nucleic acids (cfNAs) circulating in human blood, in 1948. Unlike circulating tumor cells (CTCs), which are tumor-derived by definition, the release of cell-free DNA into the blood by dying cells is not limited to cancer patients; cell-free DNA (cfDNA) can be perceived in the plasma of healthy subjects, and fetal DNA can be found in the plasma of pregnant women. The fraction of circulating tumor DNA (ctDNA) detectable within the total of plasma cfDNA in cancer patients is related to tumor burden or disease extension, and in 1977 it was shown that the cancer patients had higher levels of cfDNA than healthy subjects did. These levels of cfDNA were later related, in 1989, with tumor cells by the identification of biophysical characteristics (Bidard, Weigelt and Reis-Filho, 2013; Siravegna and Bardelli, 2014; Siravegna *et al.*, 2019). However, it is possible that ctDNA yields are linked to factors that cause cancer cells to die, such as hypoxia, proliferation, therapy response, and tumor management (surgery or biopsy). These factors must be taken into consideration since they may affect the prognostic and predictive usefulness of ctDNA-derived biomarkers (Bidard, Weigelt and Reis-Filho, 2013).

Liquid biopsy is a minimally invasive test that has the ability to detect tumor material (tumor cells, or circulating cfDNA) that is shed into body fluids (e.g.: blood, cerebrospinal fluid (CSF), urine, and saliva) due to necrosis or apoptosis (**Figure 1**) (Siravegna and Bardelli,

2014; Izquierdo *et al.*, 2021). Regarding blood samples for liquid biopsy, cfDNA is present in both plasma and serum. However, plasma has been proved the best cfDNA source, even though there is more quantity of cfDNA in serum. This is because of the clotting on serum samples, which leads to the lysis of the white blood cells, leading to a dilution of the cfDNA with genomic DNA (Siravegna and Bardelli, 2014; Rolfo *et al.*, 2020; Cheng *et al.*, 2021). Plasma cfDNA has been seen as a promising tool to evaluate the dynamic of solid tumors (Gambardella *et al.*, 2020). By using liquid biopsy analysis, besides the ability to overcome the challenge of intratumor heterogeneity when compared to a single-tumor biopsy, it is possible to evaluate the presence of driver mutations, the genomic landscape, tumor evolution in response to therapy, the level of genetic instability and seek potential resistance mechanisms that may be informative regarding future therapeutic approaches. All of this can be done at multiple time points (longitudinal studies), avoiding the jeopardies of a biopsy (Bidard, Weigelt and Reis-Filho, 2013; Siravegna and Bardelli, 2014; Merker *et al.*, 2018; Rowlands *et al.*, 2019; Izquierdo *et al.*, 2021; O'Rourke *et al.*, 2021). Nevertheless, the cfDNA level can be scant and the ctDNA fraction bearing a mutation even lower (Rowlands *et al.*, 2019). Therefore, mutation detection in ctDNA is a challenge and needs highly sensitive assays (Rowlands *et al.*, 2019).

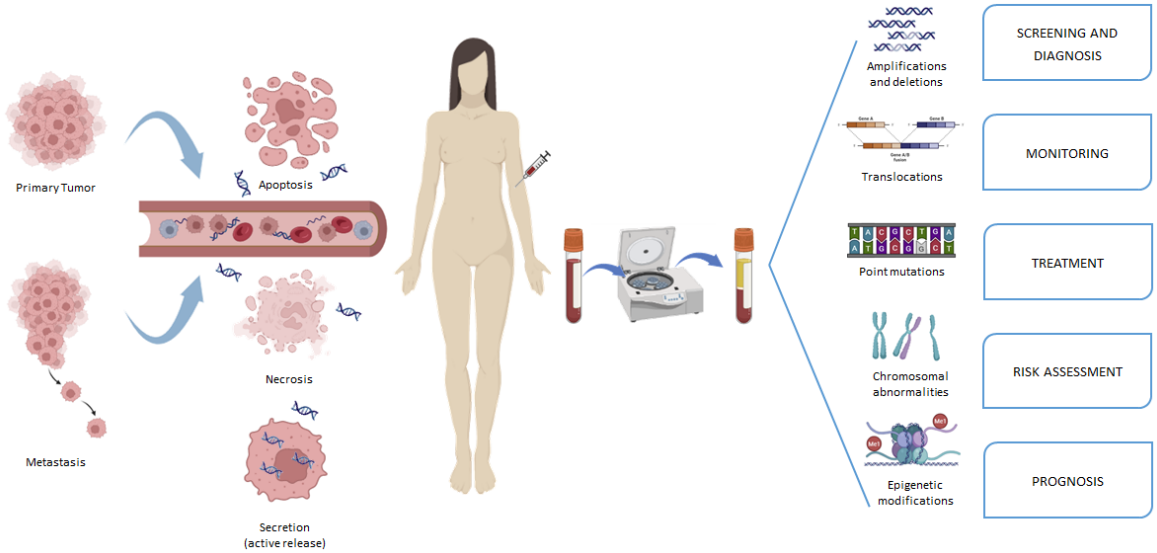


Figure 1 *- Schematic presentation of liquid biopsy.*

The size of the cfDNA fragments varies between 70 to 200 base pairs but longer fragments have been detected (Siravegna and Bardelli, 2014). The amount of cfDNA is very variable due to physiologic and pathological events, but also due to the tumor type. For this reason, it is not informative to look only for cfDNA concentrations (Siravegna and Bardelli, 2014). Because circulating tumor DNA is generally at a low concentration, short, fragmented, and diluted by the presence of genomic DNA, detecting tumor-derived mutations in ctDNA is difficult (Rowlands *et al.*, 2019). Knowing that mutant allelic frequencies in plasma are proportional to the ratio of ctDNA to normal DNA, which is linked to disease extension, it is possible that somatic mutations can only be reliably detected in patients with a high tumor burden (Bidard, Weigelt and Reis-Filho, 2013). As a result, very sensitive and specific assays are required for analysis. To find and monitor mutations in ctDNA, a variety of techniques can be used, each with its own set of benefits and drawbacks (Rowlands *et al.*, 2019).

1.3 Sequencing technologies – the era of Massively Parallel Sequencing

Frederick Sanger and colleagues first described Sanger dideoxynucleotide sequencing method, in 1977 (**Figure 2**) (Sanger, Nicklen and Coulson, 1977). This methodology is an enzymatic reaction that simulates the DNA replication mechanism. It involves the DNA polymerase with its template, primed with a synthetic oligonucleotide primer to provide a free 3'OH for the polymerase-catalyzed addition of native nucleotides and dideoxynucleotide analogs. Sanger sequencing produces a pool of molecules in each reaction mix that includes some molecules that are terminated at each residue within the growing chain. This creates a chain in which the dideoxynucleotide in that specific reaction is incorporated. Each reaction mix contains one of four enzymatic primer extension reactions. Then, an electrophoretic separation occurs in the sequencing machine, leading to the sequencing detection stage where a fluorescence detection provides four-color plots to reveal the DNA sequence (Metzker, 2010; McCombie, McPherson and Mardis, 2019). Sanger sequencing scalability and broad use have seen considerable technological breakthroughs, allowing large-scale projects to be completed. However, the use of Sanger techniques was still time-consuming and expensive (Margulies *et al.*, 2005; McCombie, McPherson and Mardis, 2019). One of the main challenges associated with Sanger sequencing technologies is the difficulty to scale for high throughput applications and the fact that this methodology may not be sensitive enough to identify low-level somatic alterations (Yan *et al.*, 2021). With the release of the first massively parallel DNA sequencing technology in 2005, this scenario began to shift (Margulies *et al.*, 2005).

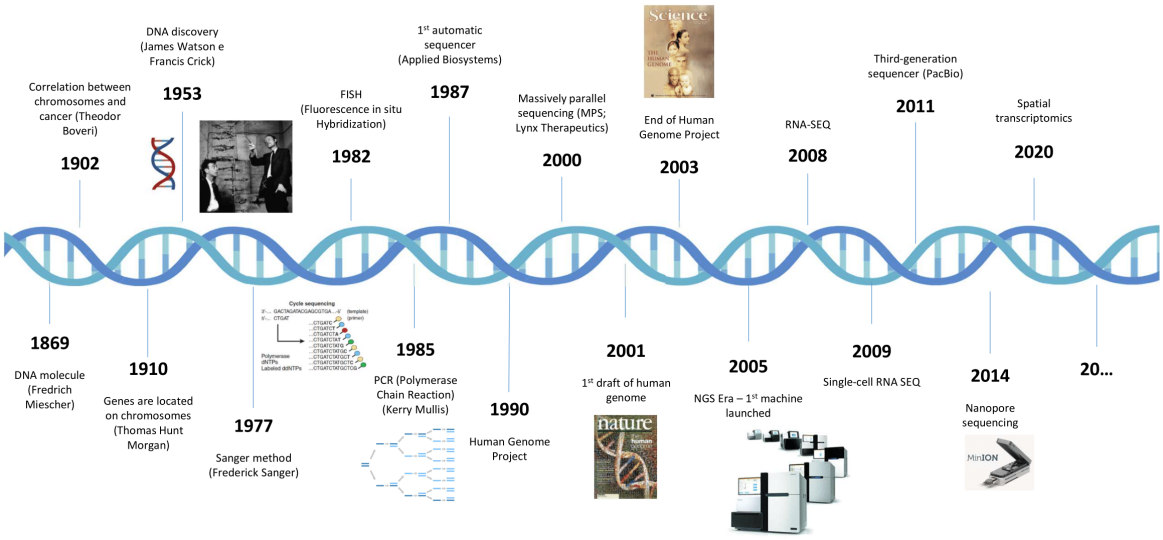


Figure 2 *- Sequencing technology timeline*

Next-generation sequencing (NGS) or massively parallel sequencing (MPS), known as high-throughput molecular technologies, have considerably improved the possibilities of genome analysis (Bidard, Weigelt and Reis-Filho, 2013; Sikkema-Raddatz *et al.*, 2013). This technology allows DNA or RNA sequencing for variant detection and allow studying a high number of genes in parallel, or the whole exome/genome, at a low cost, and in a short period of time (Qin, 2019; Mosele *et al.*, 2020).

Nowadays, disease diagnosis, prognosis, therapeutic decision, and patients follow-up have been facilitated by the use of MPS and new opportunities have arisen for personalized precision oncology (Qin, 2019).

An MPS instrument manages and progressively performs both enzymology and data capture, allowing sequence data to be created from tens of thousands to billions of templates at the same time. As a result, the term "massively parallel" alludes to this

increased data-generation capability, which has resulted in substantial improvements in DNA sequencing (McCombie, McPherson and Mardis, 2019).

The general procedure has several steps. Starts with the physical or enzymatic shearing of the DNA/RNA sample, which is then ligated to adaptors (library preparation), sequenced in several parallel reactions, bioinformatics analysis, and variant annotation and interpretation (**Figure 3**). This sequencing step can occur either in a solid matter (flow cells or sequencing chips) or in microreactors, depending on the platform used (Mardis, 2008; Metzker, 2010; Bidard, Weigelt and Reis-Filho, 2013; McCombie, McPherson and Mardis, 2019; Qin, 2019). The sequencing depth and the errors rates will determine the sensitivity of the procedure. After bioinformatics analysis, it is possible to extract qualitative and quantitative information (Bidard, Weigelt and Reis-Filho, 2013). It is important to be said that whole-genome sequencing (WGS) or whole-exome sequencing (WES) is less employed for cancer somatic mutations since it has a restricted depth of coverage. Variant allelic frequency (VAF) in a tumor might vary, as can the percentages of tumor cells. In such situations, detecting distinct alterations with different VAF generally necessitates deep sequencing, which is challenging using the WGS or WES method. For this reason, targeted panels, which interrogate dozens or hundreds of genes, are the most often utilized test for cancer patients. Typically, focused MPS assays are created for a certain cancer type or a group of diseases. As a result, it enables much more in-depth sequencing, which is required to cover a wide range of alterations with varying VAF (Qin, 2019).

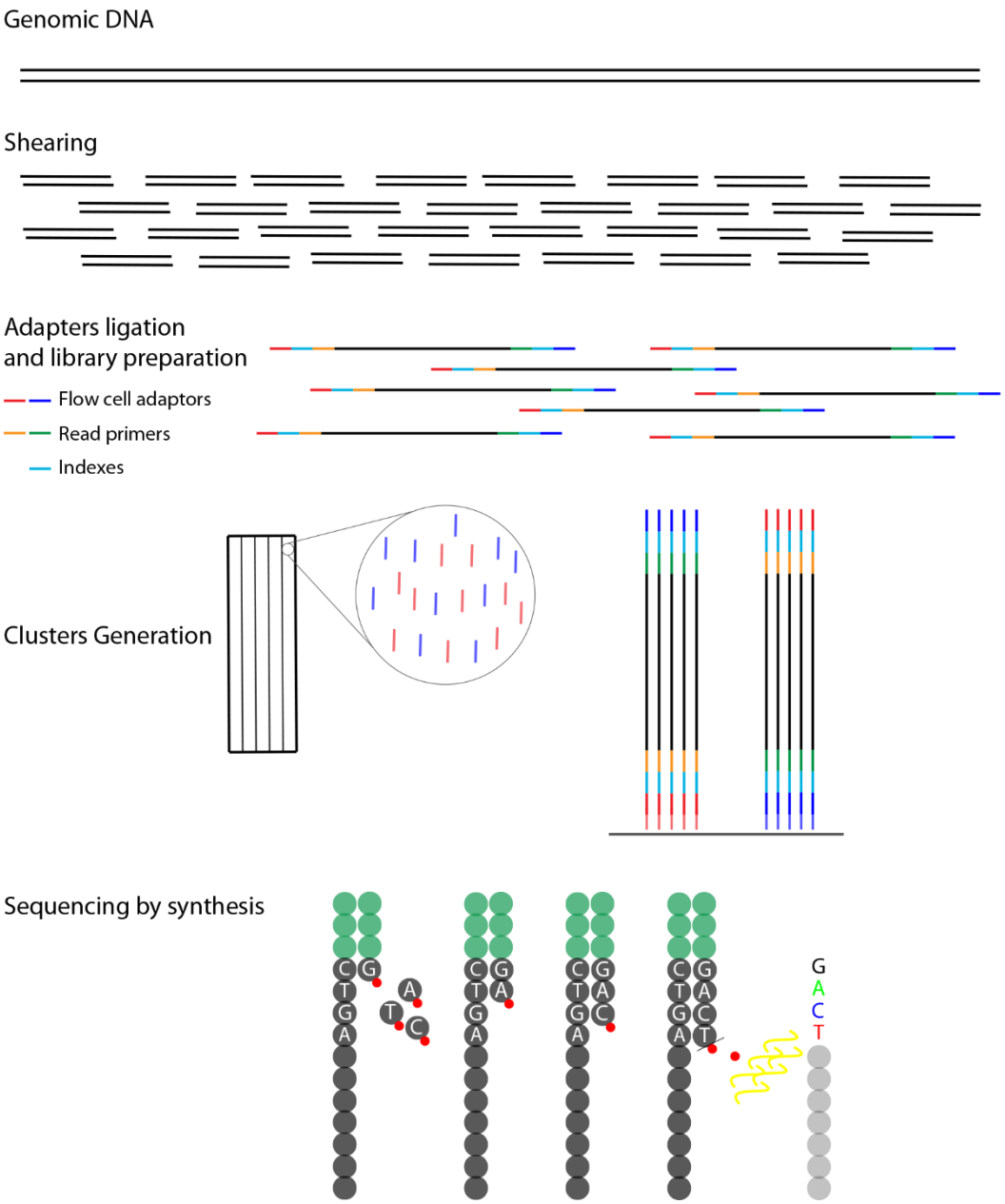


Figure 3 *- Schematic presentation of an MPS technology, using sequencing by synthesis chemistry.*

MPS has made its way into clinics, with the goal of sequencing multiple genes per tumor sample, to find driver and/or actionable changes. This technology has been found to have strong analytical validity in detecting clonally dominant changes in pioneering research (Mosele *et al.*, 2020). However, it is important to understand the recommendations regarding the use of this technology across different types of cancer. At this time, some cancers have an indication for multigene testing (e.g. NSCLC, cholangiocarcinoma), while others either do not have an indication for multigene (e.g. breast cancer, pancreatic cancer,

gastric cancer) or have indication only if represents lower costs (e.g. prostate cancer, colon cancer) (Mosele *et al.*, 2020).

In cancer patient treatment, for example, a single tumor may present a large repertoire of mutations – tumor mutation heterogeneity. If standard molecular assays are employed (e.g. Sanger sequencing, FISH, MLPA) in such clinical situations, numerous assays may be required. For these many assays, a larger volume of tissue may be required. These targets can be investigated using MPS technology in a single test contributing to cancer treatment by merging clinicopathological characteristics with a tumor's genomic profile to define diagnostic, prognostic, and therapeutic implications to a specific patient (Bidard, Weigelt and Reis-Filho, 2013; Qin, 2019). For immunotherapy purpose, tumor mutation burden has been an important parameter to examine. For this, many mutations in a tumor sample must be investigated. Again, traditional molecular test procedures are ineffective in such situations. As a result, MPS technology is required for such evaluation in patient care (Qin, 2019).

1.4 PCR-based technologies – the advent of Droplet Digital PCR

Conventional Polymerase Chain Reaction (PCR) is a technique used to amplify DNA, generating several copies of a particular fragment of DNA, by exponential amplification after each PCR cycle (Quan, Sauzade and Brouzes, 2018).

In 1988 Saiki et al. first suggested the concept of digital PCR, showing a Poisson distribution of individual PCR target sequences. Later, in 1999, Volgelstein et al., coined the term "Digital PCR". dPCR has shown to have superior sensitivity, precision, and reproducibility than the popular quantitative PCR (qPCR), which quantifies PCR templates by comparing the rate of PCR amplification of an unknown sample to the rate of a set of known qPCR standards (Dueck *et al.*, 2019).

Droplet digital PCR (ddPCR™) is an emerging and attractive option in the clinic for precision oncology and patient follow-up. This is a quantitative PCR method based on water-oil emulsion droplet technology (Hindson *et al.*, 2011; Taylor, Laperriere and Germain, 2017; Rolfo *et al.*, 2020; O'Rourke *et al.*, 2021). ddPCR™ is a third generation PCR solution for a wide variety of applications, for instance to search for cancer driver mutations, guide treatment decisions, monitor residual disease, study tumor evolution and treatment follow-up (Hindson *et al.*, 2011, 2013; Rowlands *et al.*, 2019).

Hindson *et al*., briefly describe the technology, as well as the possible applications (e.g. detection of CNV, detection of rare alleles and the absolute quantitation of cfDNA in plasma). The ddPCR workflow includes several steps as described below. The first step is to prepare an 8-well droplet generator cartridge with samples and droplet generation oil. This cartridge is then loaded into the droplet generator machine where samples and oil are mixed and, using vacuum, droplets are generated at a rate of ~1000/second, being the sample fractionated into 20,000 droplets (Hindson *et al.*, 2011). Each droplet contain either one, a few or no target sequences (Quan, Sauzade and Brouzes, 2018). Then, the suspension containing the droplets is transferred to a 96-well plate and follows a PCR amplification of both the mutated and wild type molecules within each droplet, in a conventional thermal cycler machine. After the PCR, the plate is loaded into the droplet reader machine and a two-color detector assigns the droplets as positive or negative,

according to their fluorescence amplitude (Hindson *et al.*, 2011). The amplified molecules are quantified, according to the Poison distribution, without the need for standard curves like those used in qPCR (**Figure 4**). Furthermore, it is not necessary to use external calibrators or endogenous controls. These characteristics help to avoid the drawbacks associated with variations in the efficiency of reactions (Hindson *et al.*, 2011, 2013; Quan, Sauzade and Brouzes, 2018; Dueck *et al.*, 2019).

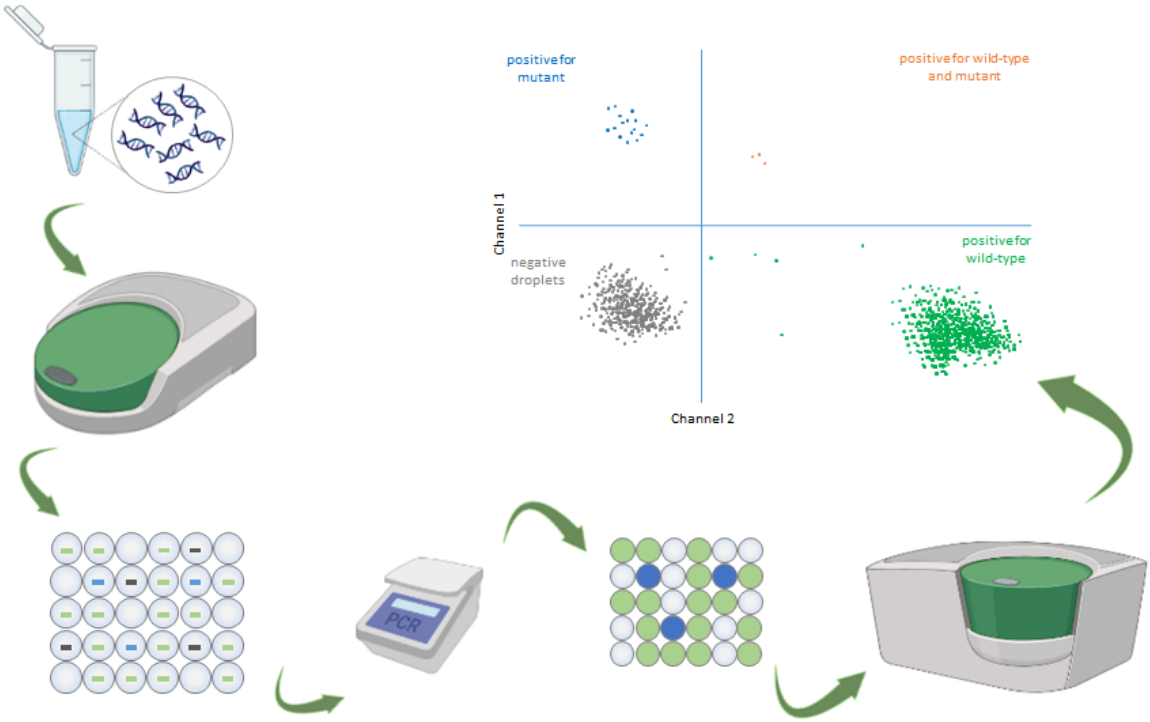


Figure 4 *- ddPCR technology.*

Regarding the applications of this technology, some examples are explored below. For example, for detection and quantification of somatic copy number alterations (amplifications and deletions), without the high-throughput technology, fluorescent in-situ hybridization (FISH) is used. However, FISH and related techniques are expensive, laborious and subject to large performance losses due to other analytical factors. This ddPCR technology provides the ability to scale the number of partitions by combining replicate wells to resolve fine copy number differences in heterogeneous mixtures and can predictably form the basis of more efficient diagnostic testing (Hindson *et al.*, 2011).

Concerning the rare mutant detection, it is known that qPCR can detect down to 1% of allele fraction. With ddPCR, due to the fact of the reaction efficacy is independent of the

amount of the initial sample, provides the possibility of accurately quantify low-abundance targets, being able to detect as low as 0,001% of allele fraction (Hindson *et al.*, 2011; O'Rourke *et al.*, 2021).

A third possible application is the quantitation of plasma cfDNA. Hindson *et al.*, demonstrate that ddPCR is able of absolute quantitation of cfDNA, even though being a highly fragmented sample and is present at low levels, being diluted by genomic DNA (Hindson *et al.*, 2011).

When compared with qPCR, this technology brings several advantages, including an absolute quantitative result, improved precision, reproducibility, accuracy, sensitivity and greater tolerance to PCR inhibitors (Hindson *et al.*, 2011, 2013; Taylor, Laperriere and Germain, 2017; O'Rourke *et al.*, 2021).

Albeit this technology represents a lower cost when compared to MPS, there is a main disadvantage connected to the difficulty of evaluating multiple targets simultaneously, being limited to single mutations or sets of related mutations at the same *locus* (Hindson *et al.*, 2011; Rowlands *et al.*, 2019; Rolfo *et al.*, 2020).

2. Objectives

This project was initiated with the goal of implementing and validating cutting-edge methodological approaches for profiling tumor DNA in the clinic.

The work plan had three specific objectives.

The first objective was to assess the limitations of targeted qPCR approaches, by characterizing uncommon genetic alterations in the *EGFR* gene using Sanger sequencing.

Considering the obtained results and knowing that after the detection of an EGFR-TKI-sensitive mutation, at initial diagnosis, and after treatment with 1st or 2nd generation inhibitors, about half of tumors develop a resistance mutation, our second objective was to implement and assess the performance of a droplet digital PCR (ddPCR) assay for detecting the *EGFR* T790M mutation in liquid biopsies.

Recognizing that the major disadvantage of ddPCR is the difficulty in testing multiple targets simultaneously, our third objective was to explore massively parallel sequencing methodologies for mutation discovery in blood samples (cfDNA).

3. Results

3.1 Detection of rare and novel *EGFR* mutations in NSCLC patients: Implications for treatment-decision.

3.1.1 Highlights

- Approximately 3% of NSCLC patients have uncommon genetic alterations in the *EGFR* gene.
- We report six new single mutations previously unidentified in NSCLC patients.
- We report seven novel combinations of doublet mutations.

3.1.2 Authors

Ana Carla Sousa[1,*], **Catarina Silveira**[1,*], André Janeiro[1], Sara Malveiro[1], Ana Rita Oliveira[1], Margarida Felizardo[2], Fernando Nogueira[3], Encarnação Teixeira[4], Joana Martins[4], Maria Carmo-Fonseca[5,#]

Authors Affiliations:

[1] GenoMed, Diagnósticos de Medicina Molecular, SA, Lisboa, Portugal, [2] Hospital Beatriz Ângelo, Serviço de Pneumologia, Lisboa, Portugal, [3] Hospital Egas Moniz, Serviço de Pneumologia, Lisboa, Portugal, [4] Hospital Pulido Valente, Hospital de Dia de Pneumologia Oncológica, Centro Hospitalar Lisboa Norte, Centro Académico de Medicina de Lisboa, Lisboa, Portugal, [5] Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal. * These authors contributed equally to this study. [#] Corresponding author.

3.1.3 Abstract

Objectives: Mutations in the gene that encodes epidermal growth factor receptor (*EGFR*) are biomarkers that predict how non-small cell lung cancer (NSCLC) patients respond to EGFR-targeted therapies collectively known as tyrosine kinase inhibitors

(TKIs). Thus, *EGFR* genotyping provides crucial information for treatment decision. Both Sanger sequencing and real-time PCR methodologies are used for *EGFR* genotyping. However, methods based on real-time PCR have limitations, as they may not detect rare or novel mutations. The aim of this study was to determine the prevalence of rare mutations in the tyrosine kinase domain (exons 18 to 21) of the *EGFR* gene not targeted by the most frequently used real-time PCR approaches, i.e., the cobas® EGFR Mutation Test, and the Idylla™ EGFR Mutation Assay. Methods**:** A total of 1228 NSCLC patients were screened for mutations in exons 18 to 21 of the *EGFR* gene using Sanger sequencing. Results**:** We observed that 252 patients (~20%) had at least one mutation in the *EGFR* gene, and 38 (~3%) carried uncommon genetic alterations that would not be identified by the cobas® or the Idylla™ tests. We further found six new single mutations and seven previously unreported compound mutations. Clinical information and patient outcome are presented for these cases. Conclusions**:** This study highlights the value of sequencing-based approaches to identify rare mutations. Our results add to the inventory of known *EGFR* mutations, thus contributing to improved lung cancer precision treatment.

Key words: *EGFR*, NSCLC, tyrosine-kinase domain mutations, molecular diagnosis.

3.1.4 Introduction

Every year, almost 2 million people receive a diagnosis of lung cancer (Ferlay, Parkin and Steliarova-Foucher, 2010). Of these, 85% have non–small cell lung cancer (NSCLC) (Dela Cruz, Tanoue and Matthay, 2011). Tobacco smoking is associated with 80% of all NSCLC cases but incidence in non-smokers has increased in the last years (Shaw *et al.*, 2009). Several genetic alterations define NSCLC subtypes (Pikor *et al.*, 2013). In particular, mutations in exons 18 to 21 of the *EGFR* gene are often detected in tumor samples of NSCLC patients. Mutations in these exons, which encode the tyrosine kinase domain of the Epidermal Growth Factor Receptor (EGFR), result in a gain in function, leaving the intracellular signaling pathway of EGFR constitutively active. About 10 to 30% of NSCLC samples harbor somatic mutations in exons 18, 19, 20 and / or 21 of the *EGFR* gene kinase domain (Sharma *et al.*, 2007).

In the last years, novel drugs known as EGFR-targeted therapies, or EGFR-tyrosine kinase inhibitors (TKIs), namely gefitinib, erlotinib, afatinib and osimertinib, have been successful in delaying disease progression in a subgroup of NSCLC patients (AF Gazdar, 2009).

Mutations in the *EGFR* gene are considered the most robust predictive biomarkers for response to EGFR-TKIs. Additionally, it has been suggested that *EGFR* mutations can also be useful for diagnosis purposes and therefore should be tested along with standard clinical examination, pathology and imaging studies (Park *et al.*, 2018). The guidelines set by the National Comprehensive Cancer Network (NCCN) and the European Society for Medical Oncology (ESMO) for the diagnosis, treatment and follow-up of NSCLC patients, state that *EGFR* genotyping is fundamental for decision making, as it may indicate which treatment is most likely to be effective in each patient (Kerr *et al.*, 2014; Masters *et al.*, 2015; Ettinger *et al.*, 2018). Therefore, great efforts have been made to standardize diagnostic tests. Panels of international experts from reference institutions such as the College of American Pathologists (CAP), International Association for the Study of Lung Cancer (IASLC), and the Association for Molecular Pathology (AMP) (Lindeman *et al.*, 2018) recommend the screening for mutations in the *EGFR* gene to be performed by one of the two complementary methodologies: PCR followed by direct

sequencing or real-time PCR. The most commonly used real-time PCR-based techniques are cobas® EGFR Mutation Test and Idylla™ EGFR Mutation Assay. These tests are highly sensitive and can detect mutations in samples with tumor infiltration as low as 10-20%. Both tests use a targeted approach designed to detect the most common *EGFR* mutations (Ellison *et al.*, 2013). Indeed, approximately 90% of all *EGFR* mutations consist of either deletion in exon 19 or a single specific point mutation, namely c.2753C>T (p.L858R) in exon 21 (Leary *et al.*, 2012).

In this study, we analyzed the complete sequence of *EGFR* exons 18 to 21 in 1228 NSCLC patients by Sanger sequencing. We identified novel mutations and we found that approximately 3% of the patients carried genetic alterations that would not be identified by the cobas® and the Idylla™ tests.

3.1.5 Materials and Methods

Study design and patients

We analyzed data from patients who had received a diagnosis of NSCLC and were tested at GenoMed, in Lisbon, Portugal, for the presence of mutations in the *EGFR* gene between 2010 and 2017. Tumor samples from 1228 patients, including 539 females and 689 males, were studied. Patients' ages at the time of testing ranged between 21 and 94 years old. This is a retrospective study for which patient consent and ethical committee approval were not necessary.

Sample collection

For this study, we analyzed data obtained from either formalin-fixed paraffin-embedded (FFPE) blocks or FFPE sections, from the following hospitals in Portugal: *Hospital Pulido Valente* and *Hospital Santa Maria* (*Centro Hospitalar Lisboa Norte*); *Hospital Professor Doutor Fernando Fonseca*; *Hospital da Luz*; *Hospital Beatriz Ângelo*; *Hospital Egas Moniz* and *Hospital São Francisco Xavier* (*Centro Hospitalar Lisboa Ocidental*); *CUF*; *Hospital das Forças Armadas*. Additional hospitals located elsewhere in Portugal are *Hospital Vila Franca Xira*; *Hospital de Santo André* (*Centro Hospitalar de Leiria)*; *Hospital de Faro*.

Sample preparation

Only samples with tumor content higher than 20% were analyzed because for samples containing less than 20 to 30% of tumor infiltration and/or less than 200 neoplastic cells the Sanger sequencing procedure may not be sensitive enough for mutation detection (Ladanyi and Pao, 2008; Leary *et al.*, 2012). DNA was extracted from tissue samples using a QIAmp Blood mini kit (catalog #51106), according to manufacturer's instructions. The DNA solution was quantified using a full spectrum (220-750nm) spectrophotometer, NanoDrop® ND-1000 (NanoDrop® Technologies).

Mutation screening

We screened the tyrosine kinase domain (exons 18 to 21) of the *EGFR* gene (7p11.2, OMIM#131550, NM_005228.3, LRG_304(t1)). Exonic regions of interest were amplified

by polymerase-chain-reaction (PCR) with flanking intronic primers. For exons 18 to 21, the reaction mixtures contained 1µL of template DNA (150ng), 2.5µL of 10×Buffer (Bioline), 2.5µL of 2mM dNTPs mixture (Bioline), a range from 0.6 to 0.75µL of 50mM $MgCl_2$ (Bioline), a range from 0.6 to 0.75µL of each primer (10µM), 0.3µL of BioTaq DNA polymerase (5U/µL) (Bioline) and ddH_2O, in a final volume of 25µL. Thermal cycling conditions were as follows: initial denaturation at 96°C for 5 minutes; followed by 40 cycles of 94°C for 30 seconds, annealing temperature between 63°C and 67°C for 30 seconds and extension at 72°C for 30 seconds. The final extension was for 10 minutes at 72°C. **Supplementary Table 1.1** describes the primers used, fragment sizes (bp) and annealing temperature of each exon. PCR products were purified using the vacuum purification system Montage™ MultiScreen™ PCR96 Cleanup Kit (LSKM PCR50, Millipore) or Exo/SAP Go - PCR Purification Kit (GRISP), according to the manufacturer's instructions, and sequenced using the Big Dye v3.1 Cycle Sequencing kit (Applied Biosystems) on an automatic sequencer (ABI Prism 3100-Avant Capillary Array, 36cm, Applied Biosystems). Data was analyzed using the DNA Sequencing Analysis Software 6™ Version 6.0 (Applied Biosystems). The Sequencher™ software was used to align samples against a reference on Ensembl. The nomenclature used to describe the genetic variants follows the guidelines of the Human Genome Variation Society.

3.1.6 Results

Sanger sequencing of tumor samples from 1228 individuals revealed at least one mutation in the *EGFR* gene in 252 cases, representing ~20% of the total population. This prevalence of mutations is higher than previously reported in unselected European patients (Sharma *et al.*, 2007), and the discrepancy is probably caused by the clinical selection bias of patients referred for *EGFR* genotyping.

Among the 252 samples with mutated *EGFR*, we identified 38 cases with mutations not explicitly listed in the specifications of the cobas® or the Idylla™ EGFR Mutation Tests (**Tables 1.1 and 1.2**). Among these 38 patients, 24 carried single mutations (**Table 1.1**) and 14 had compound mutations (**Table 1.2**). Clinical data for these patients is presented in **Supplementary Table 1.2**. The genetic alterations identified include point mutations, insertions and delins. Because many of these insertions and delins are located in the same region where other more common mutations that are targeted by the kits occur, it is possible that they can be detected by the PCR assays (these cases are indicated as "Not clear" in **Tables 1.1 and 1.2**). However, even if a PCR assay detects a hit in these regions, the precise genetic alteration would remain unknown.

As indicated in **Tables 1.1** and **1.2**, many of the 38 mutations identified by Sanger sequencing were previously described as either oncogenic or likely oncogenic. However, we found four single variants (L747_A755delinsNRQG, A763_Y764insLQEA, N771delinsHH and G857Wfs*40) that have not been previously reported (patients 4, 15, 29, and 30; **Table 1.1**). Noteworthy, these new variants are similar to previously described mutations (**Table 1.3**). We also detected a K823E mutation (patient 32; **Table 1.1**), that although not described in the COSMIC database or in the OncoKB is mentioned in a protein modelling study (Gajiwala, 2013). According to the aforementioned study, K823E greatly reduces the phosphorylation of *EGFR*. This is a surprising finding as most mutations localized at the protein dimer interface have very little or no impact on the phosphorylation pattern (Gajiwala, 2013). We were not able to find any study reporting an NSCLC patient carrying this or a similar mutation. We additionally detected a W817X (c.2450G>A) mutation (patient 6; **Table 1.1**), that was previously reported only in breast

cancer (Weber *et al.*, 2005); a closely related mutation (W817X, c.2450_2451insAT) was however described in brain metastasis from a primary lung cancer (**Table 1.3**).

We further found tumor samples from seven patients with combinations of two mutations that have not been previously reported co-occurring in the same patient (patients 5, 8, 13, 14, 18, 27, and 34; **Table 1.2**). Doublet mutations were detected in 19 cases (~1.5%), of which 14 (~74%) include at least one variant not catalogued in the cobas® and the Idylla™ tests (**Table 1.2**). Amino acids E709, G719, L854 were frequently altered, in agreement with a previous study (Chen *et al.*, 2008).

Finally, data on treatment response of patients with novel and rare mutations reveals very heterogeneous clinical outcomes (**Table 1.4**).

3.1.7 Discussion

Targeted-methods for *EGFR* genotyping are widely used but can only identify the most common genetic alterations. Our study highlights the value of sequencing-based approaches to identify rare mutations. Our results show that approximately 3% of NSCLC patients have rare mutations not catalogued in the cobas® and the Idylla™ tests.

To date, it remains unclear how tumors carrying uncommon mutations respond to first-line TKIs (Watanabe *et al.*, 2014). Previous studies have shown that exon 19 deletions are associated with EGFR-TKI sensitivity while exon 20 insertions are associated with treatment resistance (Lynch *et al.*, 2004; Paez, 2004; Pao *et al.*, 2004). However, one exon 20 insertion (p.A763_Y764insFQEA) was found associated with good response to EGFR-TKI therapy (Yasuda *et al.*, 2013). We observed that patient 30, who carried the novel but closely related mutation A763_Y764insLQEA and who was treated with Erlotinib, had disease progression in two months and an overall survival of 9 months.

The clinical significance of compound mutations is also unknown. Previous studies have reported mutations L858R and G719A co-occurring with other less common mutations (Kim *et al.*, 2016). The compound mutation L858R+A871G detected in patient 3 was previously reported being associated with resistance to erlotinib (Kobayashi *et al.*, 2013). The doublet L833V+H835L detected in patients 2 and 33 was previously described in patients that responded well to EGFR-TKI (Yang *et al.*, 2011; Frega *et al.*, 2016).

In this study, patients with novel or rare mutations that were treated with EGFR-TKIs had an overall survival ranging between 3 and 32 months. Further investigation of larger patient cohorts with sequencing-based approaches is needed in order to understand how uncommon genetic signatures influence treatment response. By expanding the inventory of *EGFR* mutations associated with NSCLC this study may contribute to improved lung cancer precision treatment.

3.1.8 Funding and Author Contributions

This work was supported by *Fundação para a Ciência e Tecnologia* (FCT) / *Ministério da Ciência, Tecnologia e Ensino Superior* through *Fundos do Orçamento de Estado* (UID/BIM/50005/2019), and FCT/FEDER/POR Lisboa 2020, *Programa Operacional Regional de Lisboa*, PORTUGAL 2020 (LISBOA-01-0145-FEDER-016394). C.S. was a recipient of a FCT fellowship (SFRH/BDE/110544/2015).

Ana Carla Sousa, **Catarina Silveira**, and Maria Carmo-Fonseca conceived the study. Margarida Felizardo, Fernando Nogueira, Encarnação Teixeira, and Joana Martins provided most of the analyzed samples. They also provided clinical information and their analysis and interpretation of the data. Laboratory tasks were performed by Ana Carla Sousa, **Catarina Silveira**, André Janeiro, Sara Malveiro, and Ana Rita Oliveira. Ana Carla Sousa and **Catarina Silveira** compiled and interpreted the results and wrote the first draft. All authors read, edited and approved the final manuscript.

Tables

Table 1.1. Single mutations identified by Sanger sequencing that are not catalogued in the cobas® and the Idylla™ kit specifications. The mutations are listed according to the amino acid position.

Patient ID	Protein level	Nucleotide Level	Exon	COSMIC ID	Detectable by cobas/Idylla	Ref
1 **11** **22**	**E709_T710delinsD** p.(Glu709_Thr710delinsAsp)	c.2127_2129delAAC	18	COSM51525	No	(Wu *et al.*, 2011; Kobayashi *et al.*, 2013)
28	**K745_E746insIPVAIK** p.(Lys745_Glu746insIleProValAlaIleLys)	c.2234_2235ins18 (or c.2217_2234dup)	19	COSM51504	Not clear	(He *et al.*, 2012; Arcila *et al.*, 2013)
21	**K745_E746insVPVAIK** p.(Lys745_Glu746insValProValAlaIleLys)	c.2236_2237ins18 (or c.2219_2236dup)	19	COSM26444	Not clear	(He *et al.*, 2012; Arcila *et al.*, 2013)
24	**E746_A750delinsQP** p.(Glu746_Ala750delinsGlnPro)	c.2236_2248delinsCAAC	19	COSM13557	Not clear	(Leduc *et al.*, 2017)
4	**L747_A755delinsNRQG** p.(Leu747_Ala755delinsAsnArgGlnGly)	c.2239_2265delinsAACCGACAAGGA	19	-	Not clear	**Not reported before**

37	**A750_I759delinsPT** p.(Ala750_Ile759delinsProThr)	c.2248_2276delinsCCAAC	19	COSM5023004	Not clear	(Shahi *et al.*, 2015; De Grève *et al.*, 2016)
25	**T751_I759delinsN** p.(Thr751_Ile759delinsAsn)	c.2252_2276delinsA	19	COSM96856	Not clear	(Penzel *et al.*, 2011)
30	**A763_Y764insLQEA** p.(Ala763_Tyr764insLeuGlnGluAla)	c.2289_2290insCTCCAGGAAGCC	20	-	No	**Not reported before**
26	**S768_V769delinsIL** **(or S768I + V769L)** p.(Ser768_Val769delinsIleLeu) (or p.(Ser768Ile) + p.(Val769Leu))	c.2303_2305GCG>TTT (or c.2303_2305delinsTTT)	20	COSM1651576[§]	Not clear	(Asahina *et al.*, 2006; Kobayashi *et al.*, 2013)
23	**V769_D770insCV** p.(Val769_Asp770insCysVal)	c.2307_2308insTGCGTG	20	COSM12379	Not clear	(Shigematsu *et al.*, 2005; Yasuda, Kobayashi and Costa, 2012)
31	**D770_N771insGF** (p.Asp770_Asn771insGlyPhe)	c.2310_2311insGGGTTT	20	COSM655155	Not clear	(Arcila *et al.*, 2013)
10	**D770_N771insGV** p.(Asp770_Asn771insGlyVal)	c.2310_2311insGGGCTT	20	-	Not clear	(Naidoo *et al.*, 2015)
29	**N771delinsHH** p.(Asn771delinsHisHis)	c.2311delinsCACC	20	-	Not clear	**Not reported before**

12	**H773_V774insPH** p.(His773_Val774insProHis)	c.2319_2320insCCACAC	20	COSM28944[§]	Not clear	(Yasuda, Kobayashi and Costa, 2012; Arcila *et al.*, 2013)
16 **36**	**H773_V774insNPH** p.(His773_Val774insAsnProHis)	c.2319_2320insAACCCCCAC	20	COSM12381	Not clear	(Shigematsu *et al.*, 2005; Yasuda, Kobayashi and Costa, 2012)
20	**V774L** p.(Val774Leu)	c.2320G>T	20	COSM25090	Not clear	(Pugh *et al.*, 2007)
7	**V774_C775insHV** **(or H773_V774dup)** p.(Val774_Cys775insHisVal) (or p.(His773_Val774dup))	c.2321_2322insCCACGT (or c.2316_2321dup)	20	COSM18432	Not clear	(Yasuda, Kobayashi and Costa, 2012; Arcila *et al.*, 2013)
17	**G779F** p.(Gly779Phe)	c.2335_2336GG>TT (or c.2335_2336delinsTT)	20	COSM13007	No	(Yang *et al.*, 2005)
6	**W817X (or W817*)** p.(Trp817X) (or p.(Trp817*))	c.2450G>A	20	-	No	(Weber *et al.*, 2005)
32	**K823E** p.(Lys823Glu)	c.2467A>G	20	-	No	(Gajiwala, 2013)
15	**G857Wfs*40** p.(Gly857Trpfs*40)	c.2568_2569insT (or c.2568dup)	21	-	No	**Not reported before**

[§] COSMIC ID information for protein only.

Table 1.2. Compound mutations in which at least one is not catalogued in the cobas® and the Idylla™ kit specifications.

Patient ID	Protein level	Nucleotide Level	Exon	COSMIC ID	Detectable by cobas/Idylla	Ref
35	**E709A** p.(Glu709Ala)	c.2126A>C	18	COSM13427	No	(Chen *et al.*, 2008; Kobayashi *et al.*, 2013)
	G719S p.(Gly719Ser)	c.2155G>A		COSM6252	Yes	
14	**E709V** p.(Glu709Val)	c.2126A>T	18	COSM12371	No	**Doublet not reported before**
	G719A p.(Gly719Ala)	c.2156G>C		COSM6239	Yes	
8	**G719A** p.(Gly719Ala)	c.2156G>C	18	COSM6239	Yes	**Doublet not reported before**
	D761Y p.(Asp761Tyr)	c.2281G>T	19	COSM21984	No	
9	**E709K** p.(Glu709Lys)	c.2125G>A	18	COSM12988	No	(Chen *et al.*, 2008; Wu *et al.*, 2011)
	L858R p.(Leu858Arg)	c.2573T>G	21	COSM6224	Yes	

19	**E709G** p.(Glu709Gly)	c.2126A>G	18	COSM13009	No	(Chen *et al.*, 2008; Wu *et al.*, 2011)
	L858R p.(Leu858Arg)	c.2573T>G	21	COSM6224	Yes	
13	**G719A** p.(Gly719Ala)	c.2156G>C	18	COSM6239	Yes	**Doublet not reported before**
	L833V p.(Leu833Val)	c.2497T>G	21	COSM13424	No	
18	**L747_T751delLREAT** p.(Leu747_Thr751del)	c.2240_2254del15	19	COSM12369	Yes	**Doublet not reported before**
	K754E p.(Lys754Glu)	c.2260A>G		COSM85993	No	
27	**V769M** p.(Val769Met)	c.2305G>A	20	COSM13425	No	**Doublet not reported before**
	N771_P772dup or P772_H773insNP p.(Asn771_Pro772dup) or p.(Pro772_His773insAsnPro)	c.2311_2316dup (or c.2316_2317insAACCCC)		COSM1738101[§]	No	
34	**H773R** p.(His773Arg)	c.2318A>G	20	COSM13433	No	**Doublet not reported before**

	V774M p.(Val774Met)	c.2320G>A		COSM13006	No	
2 **33**	**L833V** p.(Leu833Val)	c.2497T>G	21	COSM13424	No	(Chen *et al.*, 2008)
	H835L p.(His835Leu)	c.2504A>T		COSM6227	No	
38	**L833V** p.(Leu833Val)	c.2497T>G	21	COSM13424	No	(Chen *et al.*, 2008; Wu *et al.*, 2011)
	L858R p.(Leu858Arg)	c.2573T>G		COSM6224	Yes	
3	**L858R** p.(Leu858Arg)	c.2573T>G	21	COSM6224	Yes	(Chen *et al.*, 2008)
	A871G p.(Ala871Gly)	c.2612C>G		COSM13008	No	
5	**A864T** p.(Ala864Thr)	c.2590G>A	21	COSM13197	No	**Doublet not reported before**
	H870Y p.(His870Tyr)	c.2608C>T		COSM53292	No	

§ COSMIC ID information for protein only.

Table 1.3. Similarities with previously described mutations

Patient ID	New mutations found in this study	Previously described mutations
4	L747_A755delinsNRQG	L747_A755delinsSMS [§] L747_A755delinsSKS [§] L747_A755delinsAT [§]
6	W817X (G>A)	W817X (insAT) (Preusser *et al.*, 2015)[,§]
15	G857Wfs*40	G857E/V/R [§], G857R (Hsieh *et al.*, 2006) G857E (Liu *et al.*, 2013)
29	N771delinsHH	N771delinsSTH [§] N771delinsGY (Harada *et al.*, 2011) N771delinsTH (Sequist *et al.*, 2007) N771delinsFH (Robichaux *et al.*, 2018) N771delinsGP (Yasuda, Kobayashi and Costa, 2012)
30	A763_Y764insLQEA	A763_Y764insFQEA (Yasuda *et al.*, 2013)[,§]

[§] Mutations described in the following databases: cBioPortal, COSMIC and/or OncoKB.

Table 1.4. Treatment and outcome of patients with mutations that are not catalogued in the cobas® and the Idylla™ kit specifications.

Patient ID	Mutation	Initial treatment	Available information	OS
1	E709_T710delinsD	Gefitinib	PS at diagnosis: PS1 Best response: PD PFS: 3 months Stage at progression: IVB PS at progression: PS1 2nd Treatment: CTX Best response: SD PFS: 14 months Stage at progression: IVB	24 months

2	L833V + H835L	Surgery + RT for bone metastases.	2nd Treatment: CTX + Erlotinib Best response: SD PFS: 9 months	44 months
5	A864T + H870Y	Surgery + adjuvant CTX	PS at diagnosis: PS1 Best response: CR PFS: 11 months Stage at progression: IVA PS at progression: PS1 2nd Treatment: Gefitinib Best response: PD PFS: 3,5 months Stage at progression: IVA	46 months
6	W817X (or W817*)	CTX + Erlotinib	Best response: PD PFS: 4 months	19 months
7	V774_C775insHV (or H773_V774dup)	Erlotinib	PS at diagnosis: PS1 Best response: PD PFS: 4 months Stage at progression: IVB PS at progression: PS2 2nd Treatment: Palliative	9 months
8	G719A + D761Y	Erlotinib	PS at diagnosis: PS2 Best response: SD Stage at progression: IVB	9 months
10	D770_N771insGV	Surgery + adjuvant CTX and RT	Best response: CR PFS: 79 months	103 months
11	E709_T710delinsD	Surgery + adjuvant CTX	PS at diagnosis: PS1 Best response: PD PFS: 2,5 months Stage at progression: IVA PS at progression: PS2 2nd Treatment: Erlotinib Best response: PD	26 months

			PFS: 4 months PS at progression: PS2	
13	G719A + L833V	Gefitinib	PS at diagnosis: PS2 Best response: PD PFS: 2 months Stage at progression: IVB PS at progression: PS2 2nd Treatment: CTX Best response: SD PFS: 5 months Stage at progression: IVB	17 months
14	E709V + G719A	Gefitinib	PS at diagnosis: PS1 Best response: PD PFS: 3 months Stage at progression: IVB PS at progression: PS3 2nd Treatment: Palliative	7 months
15	G857Wfs*40	Gefitinib + RT (lung and CNS)	Best response: SD PFS: 7 months	9 months
16	H773_V774insNPH	CTX	PS at diagnosis: PS1 Best response: PR PFS: 8 months Stage at progression: IVA PS at progression: PS1 2nd Treatment: CTX Best response: PR PFS: 8 months Stage at progression: IVA PS at progression: PS2 3rd Treatment: Immunotherapy Best response: SD PFS: 3 months	22 months

17	G779F	Gefitinib	PS at diagnosis: PS2 Best response: PD Stage at progression: IVB	3 months
18	L747_T751delLREAT + K754E	Surgery + adjuvant CTX	PS at diagnosis: PS0 Best response: CR PFS: 4 years and 9 months	Alive
19	E709G + L858R	CTX	PS at diagnosis: PS1 Best response: SD PFS: 5 months Stage at progression: IVA PS at progression: PS1 2nd Treatment: Erlotinib Best response: PR PFS: 18 months Stage at progression: IVA	Alive
20	V774L	None	PS at diagnosis: PS3	4 months
21	K745_E746insVPVAIK	Gefitinib	PS at diagnosis: PS1 Best response: PR PFS: 6 months Stage at progression: IVB PS at progression: PS2 2nd Treatment: Erlotinib Best response: SD PFS: 10 months PS at progression: PS3	17 months
22	E709_T710delinsD	CTX	PS at diagnosis: PS1 Best response: SD PFS: 7 months Stage at progression: IVB PS at progression: PS2 2nd Treatment: Erlotinib Best response: PD	18 months

			PFS: 3 months PS at progression: PS3	
23	V769_D770insCV	CNS RT	PS at diagnosis: PS2 Best response: PD PFS: 2 months Stage at progression: IVB	2 months
24	E746_A750delinsQP	Erlotinib	PS at diagnosis: PS1 Best response: PD PFS: 3 months Stage at progression: IVA PS at progression: PS1 2nd Treatment: CTX Best response: PR PFS: 24 months Stage at progression: IVB	32 months
25	T751_I759delinsN	Erlotinib	PS at diagnosis: PS3 Best response: PD PFS: 2 months PS at progression: PS4 2nd Treatment: Palliative	3,5 months
26	S768_V769delinsIL (or S768I + V769L)	Erlotinib	PS at diagnosis: PS2 Best response: PD PFS: 3 months Stage at progression: IVB PS at progression: PS3 2nd Treatment: Palliative + RT for bone metastases	7 months
27	V769M + P772_H773insNP	CTX	PS at diagnosis: PS1 Best response: SD PFS: 14 months Stage at progression: IVB PS at progression: PS2 2nd Treatment: CTX	31 months

			Best response: SD PFS: 5 months Stage at progression: IVB	
28	K745_E746insIPVAIK	Surgery	2nd Treatment: Gefitinib Best response: SD PFS: 7 months	69 months
29	N771delinsHH	Surgery	PS at diagnosis: PS1 Best response: CR	NA
30	A763_Y764insLQEA	CTX + Erlotinib	Best response: PD PFS: 2 months	9 months
32	K823E	Surgery	PS at diagnosis: PS1 Best response: CR PFS: 9 months Stage at progression: IIIB PS at progression: PS1 2nd Treatment: CTX Best response: PD PFS: 3 months Stage at progression: IIIB	31 months
33	L833V + H835L	Erlotinib + Bevacizumab	Best response: PD	NA
35	E709A + G719S	Surgery + adjuvant CTX	PS at diagnosis: PS1 Best response: CR PFS: 11 months Stage at progression: IVA 2nd Treatment: Erlotinib Best response: SD PFS: Free of progession for 7 months	Alive
36	H773_V774insNPH	CTX	PS at diagnosis: PS0 Best response: CR PFS: 8 months Stage at progression: IVB	19 months

			PS at progression: PS3 2nd Treatment: CNR RT Best response: SD PFS: 5 months	
37	A750_I759delinsPT	Surgery + adjuvant CTX	PS at diagnosis: PS1 Best response: CR PFS: 4 months Stage at progression: IVA PS at progression: PS1 2nd Treatment: Erlotinib Best response: PR PFS: 15 months Stage at progression: IVB	Alive

Supplementary Table 1.1. Primers used, fragment sizes (bp) and annealing temperature per exon.

Exon	Forward Primer	Reverse Primer	Fragment size (bp)	Annealing temperature (Ta)
18	CTGGCACCCAAGCCCATG	CCCACCAGACCATGAGAGG	302	67^{o}C
19	CCAGTGTCCCTCACCTTC	AGCAGGGTCTAGAGCAGAGCAGCTGCC	308	63^{o}C
20	GATCGCATTCATGCGTCTTCACC	TATCCCAGGAGCGCAGACC	358	65^{o}C
21	GGCATGAACATGACCCTGAAT	AGCTGCTGCGAGCTCACC	371	66^{o}C

Primers 19R e 20F were designed by Shigematsu and collaborators(Shigematsu *et al.*, 2005).

Supplementary Table 1.2. Patient information

Patient ID	Sex	Age at molecular test	Age at diagnosis	Ethnicity	NSCLC Histology	Smoking status	Stage at diagnosis
1	F	66	66	Caucasian	Adenocarcinoma	Current heavy (>15 pack year)	IVB
2	F	74	71	Asian	Adenocarcinoma	Never	IB
3	M	84	NA	NA	NA	NA	NA
4	M	67	NA	NA	Adenocarcinoma	NA	NA

5	M	49	48	Caucasian	Adenocarcinoma	Never	IIB
6	M	81	80	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IV
7	M	46	46	Caucasian	Adenocarcinoma	Never	IVB
8	M	68	68	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IVB
9	F	65	NA	NA	NA	NA	NA
10	M	59	51	Caucasian	Adenocarcinoma	Former light (≤15 pack year)	IIIA
11	F	46	46	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IIA
12	F	59	59	Caucasian	Adenocarcinoma	Current heavy (>15 pack year)	IV
13	F	59	59	Caucasian	Adenocarcinoma	Never	IVB
14	M	76	76	Caucasian	Adenocarcinoma	Never	IVB
15	M	75	75	Caucasian	Adenocarcinoma	Current heavy (>15 pack year)	IV
16	F	72	72	Caucasian	Adenocarcinoma	Never	IVA
17	M	72	72	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IVB
18	F	60	60	Caucasian	Adenocarcinoma	Never	IB
19	F	55	55	African	Adenocarcinoma	Former heavy (>15 pack year)	IVA
20	M	75	75	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IVB
21	F	67	66	Caucasian	Adenocarcinoma	Never	IVB
22	F	57	57	Caucasian	Adenocarcinoma	Never	IVB
23	F	52	52	African	Adenocarcinoma	Never	IVB
24	M	63	62	Caucasian	Adenocarcinoma	Former light (≤15 pack year)	IVA

25	F	81	81	Caucasian	Adenocarcinoma	Never	IVA
26	M	72	72	Caucasian	Adenocarcinoma	Never	IVB
27	F	69	69	African	Adenocarcinoma	Never	IVB
28	F	79	74	Caucasian	Adenocarcinoma	Never	IA
29	M	70	70	Caucasian	Adenocarcinoma	Current	IB
30	M	73	72	Caucasian	Adenocarcinoma	Former light (≤15 pack year)	IV
31	M	53	NA	Caucasian	Adenocarcinoma	NA	IIA
32	F	72	71	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IA
33	M	64	NA	NA	Adenocarcinoma	NA	IB
34	F	83	83	Caucasian	Adenocarcinoma	NA	NA
35	M	64	64	Caucasian	Adenocarcinoma	Never	IIA
36	F	74	73	Caucasian	Adenocarcinoma	Never	IIIA
37	F	57	57	Caucasian	Adenocarcinoma	Never	IIIA
38	F	52	52	Caucasian	Adenocarcinoma	Current	IIIA

3.2 Detection and quantification of *EGFR* T790M mutation in liquid biopsies by droplet digital PCR

3.2.1 Highlights

- We optimized a ddPCR strategy that detected the *EGFR* T790M mutation in plasma samples from 77 patients with NSCLC in progression with a positivity rate of 52%.
- The mutation was detected at cfDNA allele frequency as low as 0.5%.
- Our results suggest that repeated ddPCR tests in cfDNA may obviate tissue re-biopsy in patients unable to provide a tumor tissue sample.

3.2.2 Authors

Catarina Silveira[1,*], Ana Carla Sousa [1,*], André Janeiro[1,*], Sara Malveiro[1], Encarnação Teixeira[2], Eva Brysch[2], Marcos Pantarotto[3], Margarida Felizardo[4], Rosa Madureira[5], Fernando Nogueira[6], Cátia Guimarães[6], Cristina Matos[6], Dolores Canário[7], Jácome Bruges-Armas[8], Maria Carmo-Fonseca[9,#]

Authors Affiliations:

[1] GenoMed, Diagnósticos de Medicina Molecular, SA, Lisboa, Portugal, [2] Hospital Pulido Valente, Hospital de Dia de Pneumologia Oncológica, [3] Hospital da Luz, Serviço de Oncologia Médica, [4] Hospital Beatriz Ângelo, Serviço de Pneumologia, [5] Hospital Beatriz Ângelo, Serviço de Anatomia Patológica, [6] Hospital Egas Moniz, Serviço de Pneumologia, [7] Hospital Garcia De Orta, Serviço de Pneumologia, [8] Hospital de Santo Espírito da Ilha Terceira, Serviço Especializado de Epidemiologia e Biologia Molecular, [9] Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Portugal. * These authors have contributed equally to the work. [#] Corresponding author.

3.2.3 Abstract

Introduction: Liquid biopsy allows the identification of targetable cancer mutations in a minimally invasive manner. In patients with advanced non-small cell lung cancer (NSCLC), droplet digital PCR (ddPCR) is increasingly used to genotype the epidermal growth factor receptor (*EGFR*) gene in circulating cell-free DNA (cfDNA). However, the sensitivity of this method is still under debate. The aim of this study was to implement and assess the performance of a ddPCR assay for detecting the *EGFR* T790M mutation in liquid biopsies. Methods: A ddPCR assay was optimized to detect the *EGFR* T790M mutation in plasma samples from 77 patients with NSCLC in progression. Results: Our ddPCR assay enabled the detection and quantification of the *EGFR* T790M mutation at cfDNA allele frequency as low as 0.5%. The mutation was detected in 40 plasma samples, corresponding to a positivity rate of 52%. The number of mutant molecules per mL of plasma ranged from 1 to 6,000. A re-biopsy was analyzed for 12 patients that had a negative plasma test and the mutation was detected in 2 cases. A second liquid biopsy was performed for 6 patients and the mutation was detected in 3 cases. Conclusions: This study highlights the value of ddPCR to detect and quantify the *EGFR* T790M mutation in liquid biopsies in a real-world clinical setting. Our results suggest that repeated ddPCR tests in cfDNA may obviate tissue re-biopsy in patients unable to provide a tumor tissue sample suitable for molecular analysis.

Keywords: lung cancer; *EGFR* T790M mutation; liquid biopsy; ddPCR

3.2.4 Introduction

Lung cancer is the most commonly diagnosed cancer and remains the leading cause of cancer death (Bray *et al.*, 2018). A significant improvement of progression-free survival has been achieved with receptor-tyrosine kinase inhibitors (TKIs) that target the epidermal growth factor receptor (EGFR) in patients with non-small cell lung cancer (NSCLC) harboring activating EGFR mutations (Mok *et al.*, 2009; Rosell *et al.*, 2012; Yang *et al.*, 2015; Wu *et al.*, 2017; Soria *et al.*, 2018).

Binding of the EGFR extracellular domain to its ligands triggers autophosphorylation at key tyrosine residues and activates several downstream signaling pathways. Certain mutations and/or amplification of the *EGFR* gene lead to constitutive activation of EGFR signaling and play an important role as oncogenic drivers in NSCLC. The prevalence of *EGFR*-activating mutations in a Caucasian population with lung adenocarcinoma is approximately 10-20%, and the most common (>90%) are small in-frame deletions in exon 19 and an amino acid substitution in exon 21 (L858R) (Paez *et al.*, 2004; Stewart *et al.*, 2015; Planchard *et al.*, 2018). These alterations confer sensitivity to EGFR-TKI therapy, resulting in response rates up to 70% and median survival up to 24-30 months (Morgillo *et al.*, 2016).

Despite initial responses, most patients with *EGFR*-mutant NSCLC and treated with EGFR-TKIs (such as gefitinib, erlonitib, and afatinib) will have disease progression within 9-14 months after starting the treatment (Stewart *et al.*, 2015; Thress *et al.*, 2015). The major mechanism of acquired resistance to EGFR-TKIs is the occurrence of a secondary *EGFR* kinase domain mutation in exon 20, the T790M substitution, which accounts for about half of the cases (AF Gazdar, 2009; Sequist *et al.*, 2011; Stewart *et al.*, 2015). This mutation leads to an enhanced affinity for ATP, thus reducing the ability of ATP-competitive reversible EGFR tyrosine kinase inhibitors, including gefitinib and erlotinib, to bind to the tyrosine kinase domain of *EGFR* (Yun *et al.*, 2008).

Recently, a third generation of EGFR-TKIs was developed that irreversibly block T790M mutant *EGFR* with maintained activity against the original exon 19del and L858R mutations (Murtuza *et al.*, 2019). Thus, testing for the *EGFR* T790M mutation has become routine clinical practice in patients with NSCLC that become resistant to first- and second-generation EGFR-TKIs. Ideally, detection of this new mutation should be done in tumor

tissue obtained by re-biopsy (Lindeman *et al.*, 2018; Planchard *et al.*, 2018). However, many patients on progression develop lesions in inaccessible locations. Moreover, the poor performance status of the patients also makes re-biopsy difficult. It is estimated that up to 40% of relapsed NSCLC patients may be unable to provide a tumor tissue sample suitable for molecular analysis (Jenkins *et al.*, 2017). For these patients it is acceptable to perform a liquid biopsy, which allows genotyping cell-free tumor DNA (cfDNA) present in the plasma and other body fluids (Rolfo et al., 2020).

Early comparisons between tumor tissue samples and liquid biopsy for determining *EGFR* mutation status concluded that analysis of cfDNA detected fewer EGFR mutation positive patients (Goto *et al.*, 2012; Mok *et al.*, 2015). However, subsequent studies using more sensitive assays such as the Inivata InVision™ (eTAm-Seq™) assay or the cobas EGFR Mutation Test, reported detection of the T790M mutation in plasma samples from 50% and 61% of the patients with NSCLC at disease progression after previous EGFR-TKI therapy (Jenkins *et al.*, 2017; Remon *et al.*, 2017).

Droplet digital PCR (ddPCR) is emerging as a very attractive option in the clinic to genotype cfDNA in liquid biopsies (Rolfo et al., 2020). This is a PCR method based on water-oil emulsion droplet technology. A cfDNA sample is fractionated into 20,000 droplets, PCR amplification of both the mutated and wild-type DNA molecules occurs in each individual droplet, and fluorescent specific probes are used to quantify the amplified molecules. Whether this approach has the required rigor to be used in the clinical setting remains debatable. A prospective validation study showed that plasma ddPCR detected *EGFR* T790M mutation with a sensitivity of 77%, supporting the use of this assay to direct clinical care (Sacher *et al.*, 2016). However, in a real-world setting, the practical sensitivity of the ddPCR assay may vary. Indeed, recent studies that analyzed plasma cfDNA by ddPCR reported values for the prevalence of the T790M mutation in patients with acquired resistance to EGFR-TKIs ranging between 30.4% (Li *et al.*, 2018) and 42.7% (Guo *et al.*, 2019).

Here we present an optimized ddPCR strategy that was used to test for the presence of the resistance *EGFR* T790M mutation in plasma samples from 77 patients with NSCLC in progression, resulting in a positivity rate of 52%.

3.2.5 Materials and Methods

Study population

This is a retrospective study including a total of 111 patients with NSCLC in progression after treatment with EGFR-tyrosine kinase inhibitors (EGFR-TKIs), who were tested for the presence of the resistance mutation T790M in exon 20 of the *EGFR* gene. Most patients were diagnosed with lung adenocarcinoma in TNM stage IV. The average age of the patients at progression was 67, ranging from 31 to 91 (**Table 2.1**). The EGFR-TKIs used as first- or second-line treatment are indicated in **Table 2.1**. All tests were performed at GenoMed, in Lisbon, Portugal, during the period from July 2015 to December 2019. The samples were obtained from the following hospitals in Portugal: *Hospital Pulido Valente*; *Hospital de Santa Maria*; *Hospital Egas Moniz*; *Hospital dos Capuchos*; *Hospital Beatriz Ângelo*; *Hospital de Santo Espírito da Ilha Terceira*; *Hospital Professor Doutor Fernando Fonseca*; *Hospital da Luz*; *Hospital Garcia de Orta*; *CUF Descobertas*; *Unidade Local Saúde Norte Alentejano* and *Hospital de Cascais*. The Lisbon Academic Medical Center Ethics Committee approved the study (311/20). Because the study involves the secondary analysis of existing data collected during standard of care procedures, a waiver of informed consent was requested and approved. All the data were analyzed anonymously, thus protecting the privacy and personal identity information of participating individuals. Samples were collected for standard clinical care and were not used for any other purpose. The sample collection procedures used have minimal risks and no adverse events were reported. The performers of the DNA tests were informed that patients had a diagnosis of advanced NSCLC in progression after treatment with EGFR-TKIs. Additional patient-related information was provided by clinicians after DNA data analysis was completed.

DNA analysis from FFPE samples

A trained pathologist made estimations of tumor cell content in formalin-fixed paraffin-embedded (FFPE) tissue samples resected from NSCLC patients. For macrodissection, the pathologist marked tumor areas on hematoxylin and eosin (H&E) stained tissue slides. Two to five serial FFPE sections of 10µm thickness were then sliced from the selected areas. DNA was extracted using cobas® DNA Sample Preparation Kit, quantified using the

NanoDrop® ND-1000 spectrophotometer (NanoDrop® Technologies), and analyzed by real-time PCR using cobas® EGFR Mutation Test. For Idylla™ EGFR Mutation Assay, sample preparation was done according to the manufacturer's instructions.

Isolation of cfDNA from plasma

To extract cfDNA from plasma, ~10mL blood samples were collected in EDTA or Cell-Free DNA BCT® (Streck) tubes. After centrifugation (900g for EDTA tubes and 1,600g for Streck tubes) for 10 minutes at room temperature, the supernatant was transferred into 2mL microtubes, wasting about 5mm of plasma in order to avoid buffy-coat disturbance. The samples were then centrifuged at 16,000g for 10 additional minutes at room temperature to remove cell debris, thus reducing contamination with cellular DNA. The supernatants were collected into microtubes (2 or 5mL), frozen, and stored at -80^{o}C. The cfDNA was isolated from 2mL thawed plasma samples using QIAamp® MinElute® ccfDNA (Qiagen), according to the manufacturer's instructions. The cfDNA was eluted in 50µL of ultra-clean water and quantified using the Qubit® 3.0 Fluorometer (Invitrogen, Life Technologies) with Qubit® dsDNA HS Assay kit (Invitrogen, Life Technologies). The purified cfDNA was stored at 4^{o}C for up to 16 hours or at -80^{o}C for longer periods.

T790M screening by Droplet Digital PCR (ddPCR)

To screen for the T790M mutation in plasma cfDNA, we used the QX200™ Droplet Digital PCR System (Bio-Rad, Hercules, CA, USA), and Bio-Rad assay primer/probe mixtures (dHsaCP2000019 and dHsaCP2000020). As reference for wild type and mutant *EGFR*, we used Horizon DNA standards HD709 and HD258 (Horizon Discovery Ltd., Cambridge, UK). The optimal annealing temperature was established at 59^{o}C. Altogether, the thermocycling conditions included an initial incubation at 95^{o}C for 10 minutes, 47 cycles of 94^{o}C for 30 seconds and 59^{o}C for 1 minute, followed by an inactivation step at 98^{o}C for 10 minutes. At least two replicates were analyzed per sample, and the total amount of DNA loaded (considering the two replicates), was approximately 15ng. No template control (NTC) was used to exclude PCR contamination. Amplification results were analyzed using QX200™ Droplet Reader and QuantaSoft™ software. Samples with three or more positive mutant droplets were considered positive, as recommended by the best practice guidelines for rare

mutation detection (Biorad, 2018). If one or two droplets were observed, the result was considered inconclusive and whenever possible a second sample was collected and tested.

Statistical analysis

The concentrations of target alleles were calculated using QuantaSoft^{TM} software (Bio-Rad) based on Poisson distribution. Mutant cfDNA molecules were reported as number of copies per millilitre (mL) of plasma. The mutant allelic frequency was determined as the ratio of mutant droplets relative to the sum of mutant and wild type droplets.

3.2.6 Results

Quality Assessment

DNA reference standards for wild type and mutant *EGFR* confirmed the specificity of the ddPCR primers and probes (**Figure 2.1**). The wild type DNA exhibited high fluorescence signal in the hexachloro-6-carboxy-fluorescine (HEX) channel (channel 2, green) and no fluorescence signal in the fluorescein amidite (FAM) channel (channel 1, blue) (**Figure 2.1A**). In contrast, when the wild type and the mutant DNA were mixed at 1% allele frequency, fluorescence signal in the FAM channel (channel 1, blue) was detected (**Figure 2.1B**).

Next, we tested a series of 13 non-clinical liquid biopsy reference samples spanning different mutations and allele frequencies. These samples were provided by AstraZeneca and none of the laboratory technicians were informed about the genotype or mutant allelic frequencies at the time of execution of the tests. The results show that our ddPCR assay accurately detected the T790M allele in frequencies ranging from 5 to 0.5% (**Table 2.2**).

Finally, we applied the ddPCR assay to analyze DNA extracted from FFPE tissue biopsies from NSCLC patients. These DNA samples were previously genotyped in our laboratory by real-time PCR using cobas® EGFR Mutation Test. The ddPCR assay detected the mutation in all samples that previously tested positive for the T790M allele.

EGFR mutation testing in clinical plasma samples

The ddPCR assay was used to screen plasma samples from 77 patients for the T790M mutation in *EGFR*. A total of 40 plasma samples were identified as positive for the T790M mutation, corresponding to 52%. The number of mutant molecules per mL of plasma ranged from 1 to 6,000 as depicted in **Figure 2.2**. A second sample was collected for 18 patients that initially tested negative (**Table 2.3**). As recommended by current guidelines (Lindeman *et al.*, 2018; Planchard *et al.*, 2018), when the liquid biopsy was negative, re-biopsy was performed for analysis of tumor tissue. However, this was only possible in 12 cases, either because the patient status did not allow re-biopsy or the tumor was not accessible. We detected the T790M mutation in 2 of the 12 re-biopsies (17%) (**Table 2.3**). In the remaining cases, we analyzed either a second plasma sample (5 patients) or a sample

of bronchoalveolar fluid (1 patient), and detected the T790M mutation in 3 of these 6 liquid biopsies (50%) (**Table 2.3**).

Finally, we analyzed the association of T790M detected in progression with *EGFR* activating mutations detected at diagnosis (**Table 2.4**). The majority (61%) of tumors positive for T790M had an *EGFR* del19 mutation at diagnosis, whereas the majority (65%) of tumors negative for T790M had either L858R or other *EGFR* mutations (**Table 2.4**).

3.2.7 Discussion

We implemented a liquid biopsy ddPCR assay that enabled the detection and quantification of the *EGFR* T790M mutation at cfDNA allele frequency as low as 0.5%. With a turnaround time of 3-5 business days, the assay was routinely used in clinical practice since it played an important role in deciding the next line of treatment in NSCLC patients that acquired resistance to first- or second-generation EGFR-TKIs.

We analyzed plasma samples from 77 patients and detected the *EGFR* T790M mutation in 52%. Although this was a retrospective study potentially influenced by bias associated with patient selection, our result is in good agreement with previous evidence indicating that about half of lung cancers that become resistant to EGFR TKIs acquire the *EGFR* T790M mutation (Sequist *et al.*, 2011). Most important, our results revealed that performing the ddPCR assay in a second liquid biopsy collected 1-8 months after a first negative plasma test increased by 50% the number of positive cases. Thus, the use of repeated ddPCR-based cfDNA genotyping may obviate tissue re-biopsy in cases that the tumor is not accessible, or the patient has a poor functional status.

In our study, among tumors harboring the T790M mutation detected in either liquid biopsy or tumor tissue at disease progression, the majority (61%) had an *EGFR* del19 mutation at diagnosis. This is consistent with previous reports indicating that the T790M mutation is more frequent in patients with an EGFR exon 19 deletion mutation. In a cohort of 314 Japanese patients studied by re-biopsy, the T790M mutation was detected in 55.6% of cases with del19 mutation and in 43.0% of cases with the L858R mutation (Nosaki *et al.*, 2016). In another study, the T790M mutation was present in 63% (26/41) of patients with *EGFR* exon 19 deletion and in 38% (12/32) of patients with L858R mutation (Matsuo *et al.*, 2016). A more recent literature review confirmed that detection of the T790M mutation was more frequent in del19 mutated patients (53%) than in L858R mutated patients (36%) with acquired resistance to EGFR-TKIs (Liang *et al.*, 2018). One possibility is that patients with the del19 mutation are more sensitive to TKIs, and therefore cells with the T790M mutation are more likely to be selected and enriched (Liang *et al.*, 2018).

An important advantage of ddPCR-based assays is the ability to provide absolute quantification of DNA molecules. In our study cohort, the number of T790M mutant

molecules per mL of plasma ranged from 1 to 6,000. Quantification of mutant molecules in plasma may be useful to monitor response to treatment. Indeed, a recent study reported that after 6 weeks of treatment with osimertinib, the T790M mutation load assessed by ddPCR decreased to very low level (Li *et al.*, 2018). However, whether the number of mutant molecules affects disease progression and response to therapy remains unknown. Clearly, future studies using quantitative diagnostic assays such as ddPCR are needed to maximize the benefits of precision therapy for advanced NSCLC.

3.2.8 Acknowledgements, Funding and Author Contributions

We thank AstraZeneca for providing non-clinical control samples.

This work was supported by Fundação para a Ciência e Tecnologia (FCT) / Ministério da Ciência, Tecnologia e Ensino Superior - Fundos do Orçamento de Estado (UID/BIM/50005/2019), and FCT/FEDER/POR Lisboa 2020, Programa Operacional Regional de Lisboa, PORTUGAL 2020 (LISBOA-01-0145-FEDER-016394). C.S. was a recipient of a FCT fellowship (SFRH/BDE/110544/2015).

Catarina Silveira, Ana Carla Sousa, André Janeiro, and Maria Carmo-Fonseca conceived the study. Encarnação Teixeira, Eva Brysch, Marcos Pantarotto, Margarida Felizardo, Rosa Madureira, Fernando Nogueira, Cátia Guimarães, Cristina Matos, Dolores Canário, and Jácome Bruges-Armas provided most of the analyzed samples, together with the clinical information. **Catarina Silveira**, Ana Carla Sousa, André Janeiro, and Sara Malveiro performed laboratory tasks. **Catarina Silveira**, Ana Carla Sousa, and André Janeiro compiled and interpreted the results and wrote the first draft. Maria Carmo-Fonseca wrote the final manuscript. All authors read, edited and approved the final manuscript.

Tables

Table 2.1. Clinical data of NSCLC patients enrolled in the study

Age [1]	**N (%)**
≤65	51 (46%)
>65	60 (54%)
Gender	N (%)
Female	77 (69%)
Male	34 (31%)
EGFR –TKI [2]	N (%)
Erlotinib	50 (56%)
Gefitinib	27 (30%)
Afatinib	13 (14%)

[1] Age at progression

[2] EGFR-tyrosine kinase inhibitors used in first or second-line treatment. No information was available for 21 of the 111 patients.

Table 2.2. Assessing the ddPCR assay with liquid biopsy reference samples

Sample ID	**Pre-defined *EGFR* genotype**	**Pre-defined mutant AF**	**Amount of cfDNA obtained per sample (ng/µL)**	**Result of ddPCR assay**	**Mutant AF estimated by ddPCR**
QA-01	T790M	0.05%	7,82	Not detected	-
QA-02	T790M	0.5%	6,59	Detected	0,6%
QA-03	T790M	5%	6,08	Detected	4,6%
QA-04	L858R	0.05%	7,42	Not detected	-
QA-05	L858R	0.5%	6,75	Not detected	-
QA-06	L858R	5%	6,60	Not detected	-
QA-07	ΔE746-A750	0.05%	6,86	Not detected	-

QA-08	ΔE746-A750	0.5%	6,74	Not detected	-
QA-09	ΔE746-A750	5%	5,82	Not detected	-
QA-10	Wild type	0%	4,81	Not detected	-
QA-11	T790M/L858R	5%	6,44	Detected	5,7%
QA-12	T790M/L858R	0.5%	6,35	Detected	0,3%
QA-13	T790M/L858R	0.05%	7,42	Not detected	-

AF – allelic frequency

Table 2.3. Genotyping results for samples collected after an initial negative plasma test.

1st test (plasma samples)		2nd test		
Patient #	**Date**	**Date**	**Sample type**	**Result**
P#1	29/11/2016	03/01/2017	FFPE	Not detected
P#2	05/12/2016	27/12/2016	FFPE	Not detected
P#3	09/03/2017	06/04/2017	FFPE	Not detected
P#4	18/07/2017	07/08/2017	FFPE	Detected
P#5	18/08/2017	07/11/2017	FFPE	Not detected
P#6	15/09/2017	13/10/2017	Bronchoalveolar fluid	Detected
P#7	25/09/2017	02/11/2017	FFPE	Not detected
P#8	14/12/2017	22/01/2018	FFPE	Not detected
P#9	07/02/2018	15/03/2018	FFPE	Detected
P#10	07/03/2018	19/04/2018	FFPE	Not detected
P#11	15/03/2018	20/11/2018	Plasma	Detected
P#12	22/05/2018	22/08/2018	FFPE	Not detected
P#13	28/06/2018	20/11/2018	Plasma	Detected
P#14	26/11/2018	05/12/2018	FFPE	Not detected
P#15	23/01/2019	13/03/2019	FFPE	Not detected
P#16	24/04/2019	21/05/2019	Plasma	Not detected
P#17	18/07/2019	12/09/2019	Plasma	Not detected
P#18	12/09/2019	16/10/2019	Plasma	Not detected

Table 2.4. Association of T790M detected in progression with *EGFR* activating mutations detected at diagnosis.

***EGFR* mutation at diagnosis**	**T790M detected**	**T790M not detected**
del19	N = 17 (61%)	N = 8 (35%)
L858R	N = 9 (32%)	N = 10 (43%)
Other	N = 2 (7%)	N = 5 (22%)

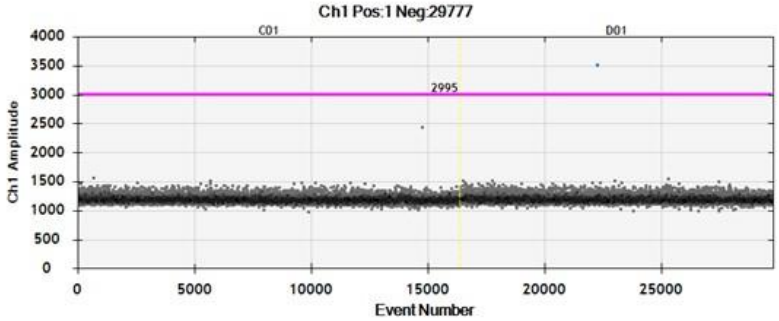
del19 - any deletion or delins in exon 19

L858R - substitution of amino acid leucine to arginine at codon 858 in exon 21

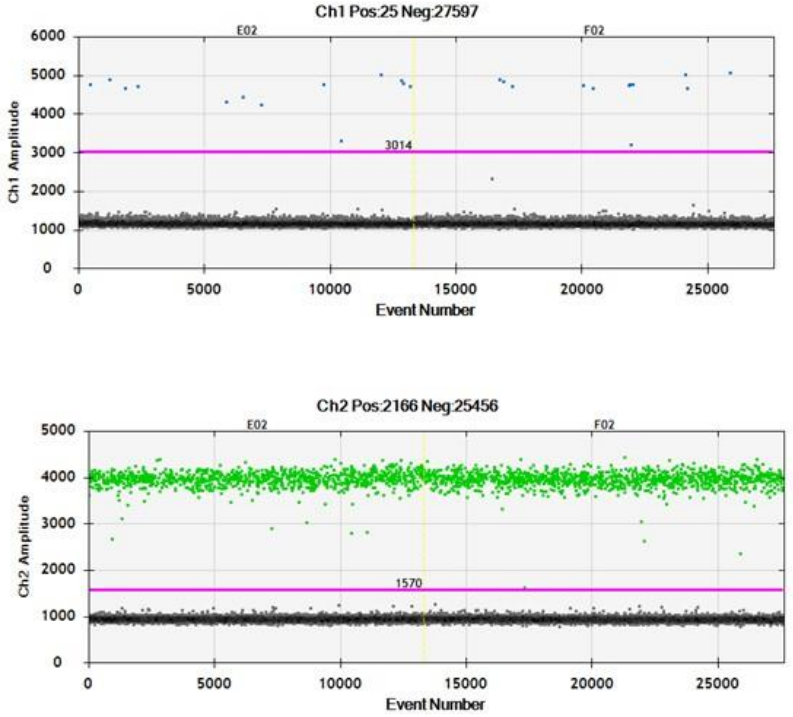
Figures

A

WT EGFR



EGFR T790M (1% allele frequency)



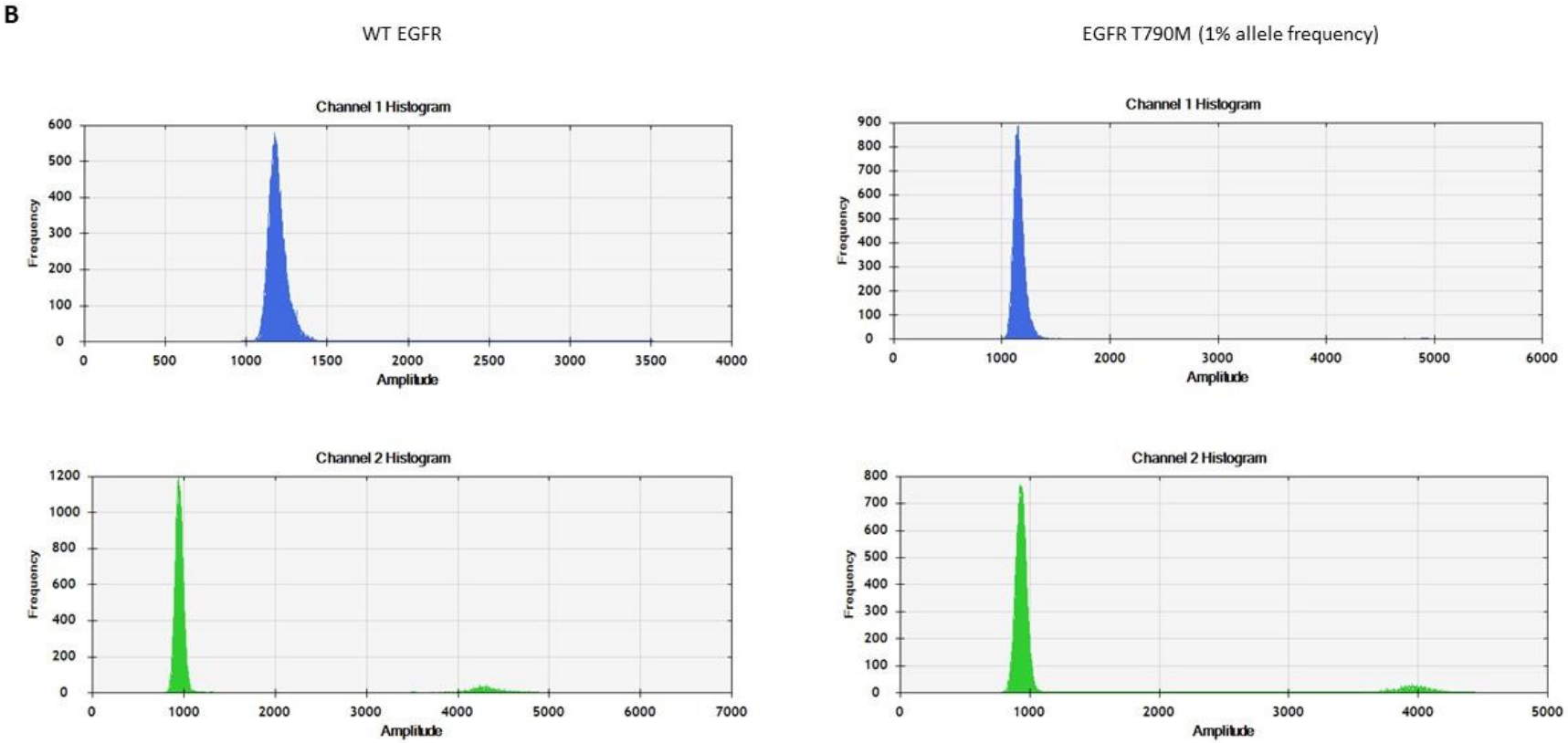


Figure 2.1 *- Visualization of two-channel ddPCR data. Either wild type (WT) reference DNA only (left) or a mix of WT and mutant reference DNA (right) were analyzed. Scatter plots* ***(A)*** *and histograms* ***(B)*** *are shown. Channel 2 (green) depicts the signal corresponding to WT EGFR. Channel 1 (blue) depicts the signal corresponding to mutant EGFR. In scatter plots, positive droplets are depicted in blue and green, negative droplets are depicted in grey, and the threshold line is depicted in pink.*

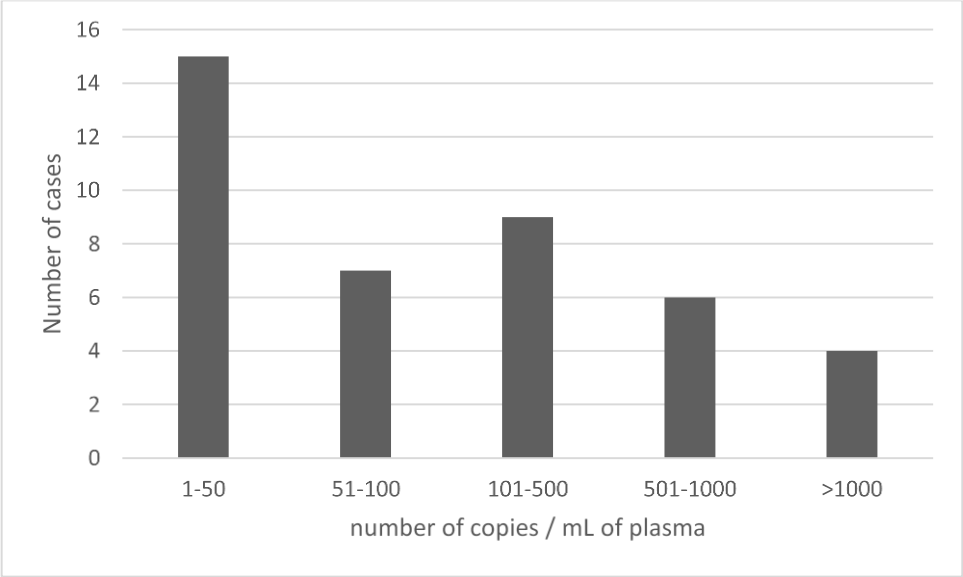


Figure 2.2 *- Distribution of the number of T790M mutant molecules in 1 mL of plasma from 77 patients tested by ddPCR.*

3.3 Comprehensive Genomic Profiling of Cell-Free Circulating Tumor DNA Detects Response to Ribociclib Plus Letrozole in a Patient with Metastatic Breast Cancer

3.3.1 Highlights

- Using massively parallel sequencing, we found a *PIK3CA* mutation and amplification of the *CCND1* gene in the plasma of a patient with metastatic breast cancer.
- The presence of both molecular alterations was confirmed in the primary tumor.
- After enrolment of the patient in a clinical trial, the two molecular alterations were no longer detected, reflecting the clinical response.
- This study demonstrates the clinical utility of combining liquid biopsy and comprehensive genomic profiling to monitor treatment response.

3.3.2 Authors

Catarina Silveira[1], Ana Carla Sousa[1], Patrícia Corredeira[2], Marta Martins[2], Ana Rita Sousa[3], Arnaud Da Cruz Paula[4], Pier Selenica[4], David N. Brown[4], Mahdi Golkaram[5], Shannon Kaplan[5], Shile Zhang[5], Li Liu[5], Britta Weigelt[4], Jorge S. Reis-Filho[4], Luís Costa[2,3], Maria Carmo-Fonseca[2,#]

Authors Affiliations:

[1] GenoMed - Diagnósticos de Medicina Molecular, S.A., Avenida Professor Egas Moniz, 1649-028 Lisboa, Portugal, [2] Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Avenida Professor Egas Moniz, 1649-028 Lisboa, Portugal, [3] Serviço de Oncologia Médica, Hospital de Santa Maria, Centro Hospitalar Lisboa Norte, Avenida Professor Egas Moniz, 1649-035 Lisboa, Portugal, [4] Department of Pathology and Laboratory Medicine, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. [5] Illumina Inc., 5200 Ilumina Way, San Diego, CA 92122, USA. # Corresponding author

3.3.3 Abstract

Analysis of cell-free circulating tumor DNA obtained by liquid biopsy is a non-invasive approach that may provide clinically actionable information when conventional tissue biopsy is inaccessible or infeasible. Here, we followed a patient with hormone receptor-positive and human epidermal growth factor receptor (HER) 2-negative breast cancer who developed bone metastases seven years after mastectomy. We analyzed circulating cell-free DNA (cfDNA) extracted from plasma using high-depth massively parallel sequencing targeting 468 cancer-associated genes, and we identified a clonal hotspot missense mutation in the *PIK3CA* gene (3:178952085, A>G, H1047R) and amplification of the *CCND1* gene. Whole-exome sequencing revealed that both alterations were present in the primary tumor. After treatment with ribociclib plus letrozole, the genetic abnormalities were no longer detected in cfDNA. These results underscore the clinical utility of combining liquid biopsy and comprehensive genomic profiling to monitor treatment response in patients with metastasized breast cancer.

Keywords: liquid biopsy; circulating cell-free DNA; metastatic breast cancer; ribociclib plus letrozole

3.3.4 Introduction

The development of metastases is a major cause of death in cancer patients (Chaffer and Weinberg, 2011). While many tumors can be cured when detected early, once metastasis forms, most cancers be-come incurable (Seyfried and Huysentruyt, 2013). In the case of breast cancer, metastasis can be found in the lungs, liver and brain, in addition to lymph nodes, but the bone is the most-affected site (Jin and Mu, 2015). Survival outcomes of breast cancer patients differ depending on metastatic sites, with bone metastasis associated with the best prognosis, and brain metastasis associated with the worst survival (Wang *et al.*, 2019).

Although metastatic disease can be present at diagnosis, most often, metastases are detected months or years following initial diagnosis and treatment. In breast cancer, metastatic recurrence has been reported ranging from months to decades after surgery (Karrison, Ferguson and Meier, 1999). Such long-term relapse of disease in a patient who was clinically asymptomatic can be attributed to cancer dormancy (Uhr and Pantel, 2011). This phenomenon is likely caused by cancer cells that escaped from the primary tumor and disseminated throughout the body (Sosa, Bragado and Aguirre-Ghiso, 2014). For unknown reasons, these micrometastases remain 'dormant' in secondary sites and evade anti-cancer therapies (Riggio, Varley and Welm, 2021). How the body controls the proliferation of microclusters of disseminated tumor cells and why they occasionally transform into growing metastases is unclear (Riggio, Varley and Welm, 2021).

The emergence of targeted therapies to treat metastatic breast cancer is extending life expectancy (Caswell-Jin *et al.*, 2018; Rinnerthaler, Gampenrieder and Greil, 2018). In parallel, the development of non-invasive biomarker assays based on liquid biopsy promises to enable the early detection of breast cancer relapse. Recent studies showed that genomic alterations identified in the primary tumor of breast cancer patients could be detected in circulating cell-free DNA (cfDNA) analyzed from plasma samples collected approximately 10 months prior to clinical or radiological relapse (Coombes *et al.*, 2019; Garcia-Murillas *et al.*, 2019). Thus, liquid biopsy may contribute to initiating treatment of metastatic disease at an earlier stage. In this regard, a meta-analysis study concluded that

the earlier detection of all breast cancer recurrences would result in an absolute reduction in mortality of 17–28% (Lu *et al.*, 2009).

Liquid biopsy tests might not only identify recurrence early, but also inform the selection of optimal treatment strategies (Riggio, Varley and Welm, 2021). Indeed, the clonal heterogeneity of tumor cells limits efficacy and duration of response to targeted treatments in metastatic cancer (Aparicio and Caldas, 2013). The analysis of cfDNA in plasma may be sufficient to identify somatic alterations contributed by distinct metastases, potentially circumventing the problem of lacking access to multiple metastatic tumor tissue samples due to associated risks and costs (Murtaza *et al.*, 2013).

The detection of circulating cell-free tumor DNA (ctDNA) requires very sensitive molecular assays. Although PCR-based technologies, including droplet digital PCR, are powerful methods for the accurate quantification of a scarce amount of circulating nucleic acids in plasma, they can only test a few mutations per assay (Dang and Park, 2022). In contrast, recent developments in massively parallel sequencing (also known as next generation sequencing) technologies allow for the comprehensive genomic profiling of entire exonic regions of hundreds of cancer-relevant genes, identifying base substitutions, insertions or deletions, copy number alterations, and gene rearrangements (Woodhouse *et al.*, 2020; Caputo *et al.*, 2022; Poh *et al.*, 2022).

Here, we used an established tumor-normal massively parallel sequencing assay (Cheng *et al.*, 2015; Zehir *et al.*, 2017) to characterize the genetic alterations present in the cfDNA extracted from the plasma of a breast cancer patient who developed bone metastases. We identified two DNA abnormalities that were already present in the primary tumor genomic DNA. After treatment with ribociclib plus letrozole, the patient showed a significant clinical improvement, and the two genetic alterations were no longer detected in the plasma.

3.3.5 Methods

cfDNA Extraction from Blood Samples

Within 1 to 2h after blood collection in EDTA tubes, whole blood was centrifuged at 1600g for 10min at room temperature. Then, the supernatant was transferred into falcon tubes, wasting about 5mm of plasma to avoid buffy-coat disturbance. Next, the plasma samples were centrifuged at 3000g for 10min at room temperature. This high g-force centrifugation step removes cellular debris and thereby reduces the amount of cellular or genomic DNA and RNA in the sample. After this step, the supernatants were collected into microtubes (2 or 5mL), without disturbing the pellet containing cell debris. Plasma samples were frozen and stored at -80^{o}C.

For cfDNA extraction, 4mL of plasma were thawed. The cfDNA was purified using QiAamp® MinElute® ccfDNA kit from Qiagen, according to the manufacturer's instructions. The cfDNA was eluted in 50µL of ultra-clean water and quantified using the Qubit® 3.0 Fluorometer (Invitrogen, Life Technologies) with Qubit® dsDNA HS Assay kit (Invitrogen, Life Technologies), according to the manufacturer's protocol. Sample quality was assessed using High Sensitivity D1000 ScreenTape (TapeStation, Agilent Technologies), according to the manufacturer's instructions. The purified cfDNA was stored at -80^{o}C.

To assess quality of extracted cfDNA, samples were analyzed using a fragment analyzer (TapeStation 4200, Agilent Technologies). This assay uses a fluorescently stained double-stranded DNA and separates nucleic acids by means of electrophoresis. The TapeStation Analysis software automatically determines size, quantity, and purity of each sample. The size determination is based on a known ladder with specific sizing standards. The known concentration of the upper marker is used to determine concentration values.

Genomic DNA Extraction from Blood Samples

Genomic DNA (gDNA) was extracted from the buffy-coat of the blood sample collected before enrollment in the clinical trial. After removal of plasma, a red blood cell lysis buffer (in-house solution) was added and incubated for 10min at 4^{o}C. Then, the sample was centrifuged at 250g for 10min. The supernatant was discarded and the pellet containing white blood cells was washed with PBS 1x (Sigma). After centrifugation at 250g for 10min,

the supernatant was discarded, and the pellet was resuspended in 1mL of PBS 1x (Sigma) and again centrifuged at 250g for 10min. The supernatant was discarded, and the dry pellet was stored at -80^{o}C.

DNA was extracted using a QIAmp® Blood mini kit from Qiagen, according to the manufacturer's instructions. The DNA was quantified using Qubit® 3.0 Fluorometer (Invitrogen, Life Technologies) with the Qubit® dsDNA HS Assay kit (Invitrogen, Life Technologies), according to the manufacturer's protocol.

Sequencing and Analysis of cfDNA and gDNA Extracted from Blood Samples

Both the cfDNA and matched normal gDNA were subjected to massively parallel sequencing using an established tumor-normal assay (Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets; MSK-IMPACT) that targets 468 cancer-related genes (Cheng *et al.*, 2015; Zehir *et al.*, 2017). Sequencing data were processed and analyzed as previously reported (De Mattos-Arruda *et al.*, 2014; Weigelt *et al.*, 2018; Da Cruz Paula *et al.*, 2020). Briefly, reads were aligned to the reference human genome GRCh37 using the Burrows–Wheeler Aligner (v0.7.15) (Li and Durbin, 2009). Local realignment, duplicate removal, and base quality recalibration were performed using the Genome Analysis Toolkit (v3.7) (DePristo *et al.*, 2011). Somatic single-nucleotide variants (SNVs) were detected by MuTect (v1.0) (Cibulskis *et al.*, 2013), and small insertions and deletions (indels) were detected using a combination of Strelka (v2.0.15) (Saunders *et al.*, 2012), VarScan2 (v2.3.7) (Koboldt *et al.*, 2012), Lancet (v1.0.0) (Narzisi *et al.*, 2018), Scalpel (v0.5.3) (Narzisi *et al.*, 2014), and Platypus (Rimmer *et al.*, 2014). Pathogenic mutations were defined as variants that were deleterious and/or mutational hotspots. Allele-specific copy number alterations (CNAs) and loss of heterozygosity (LOH) were defined using FACETS (Shen and Seshan, 2016), as previously described (Pareja *et al.*, 2019; Da Cruz Paula *et al.*, 2020). The fraction of the genome altered was computed from the CNAs obtained from FACETS. The cancer cell fraction of each mutation was determined using ABSOLUTE (v1.0.6) (Carter *et al.*, 2012), as previously described (Weigelt *et al.*, 2018; Pareja *et al.*, 2019; Da Cruz Paula *et al.*, 2020).

Sequencing and Analysis of Primary Tumor Genomic DNA

DNA was extracted from the FFPE primary tumor sample, and Illumina DNA Prep with Enrichment was used for generating whole-exome sequencing libraries, with 40ng input DNA, as previously described (Golkaram *et al.*, 2021). In brief, following quantification with Qubit® dsDNA High Sensitivity assay, four libraries were pooled for enrichment (4-plex) such that 500ng of each library was used. Target enrichment was performed using IDT xGen Exome Research Panel. A single hybridization was done overnight at 58^{o}C, with 12 cycles of post-enrichment PCR. Libraries were quantified by Qubit® dsDNA High Sensitivity assay, normalized, and pooled. Samples were sequenced with 151bp paired-end reads on the NovaSeq 6000 S4 flow cell using the XP workflow for individual lane loading.

Whole-exome sequencing data were processed as previously described (Golkaram *et al.*, 2021). An un-paired normal sample was used to perform variant calling. All germline variants observed in a database curated in-house, which includes the most common germline variants present in dbSNP (Sherry *et al.*, 2001), were removed. Copy number changes were estimated as previously described (Golkaram *et al.*, 2021). Tumor purity and ploidy were estimated using Sequenza 2.1, and sciClone 1.1 was used for clonality estimation (Golkaram *et al.*, 2021).

3.3.6 Results

Clinical Case

The patient is a woman who was first admitted to hospital in June 2009, at the age of 34 years. She presented with a palpable mass (4 × 4 cm) in the upper outer quadrant of the right breast, with no skin alterations, and an axillary lymphadenopathy on the right side (0.5 × 1.0 cm). A diagnostic mammogram and breast ultrasound showed a hypoechoic area in the upper outer quadrant, with irregular borders and 18 mm of diameter. A microbiopsy was performed that revealed an invasive ductal carcinoma of not otherwise specified (NOS) that was estrogen receptor-positive (ER+), progesterone receptor-positive (PR+), human epidermal growth factor receptor 2-negative (HER2-). p53 was normal as detected by immunohistochemistry.

The patient started neoadjuvant chemotherapy (CTX) with doxorubicin together with cyclophosphamide. After the 5th cycle of treatment, a computed tomography (CT) scan of the abdomen and pelvis revealed a tumor in the right ovary (5 cm). In December 2009, the patient was subjected to a breast conservative surgery with axillary lymph node dissection, and a right salpingo-oophorectomy. The histological exam revealed a residual invasive ductal carcinoma NOS in multiple areas with positive margins, an axillary lymph node metastasis of the same type, and a mature cystic teratoma of the ovary (6 cm).

In January 2010, the patient underwent a mastectomy. Post-surgical treatment was adjuvant CTX with docetaxel and, subsequently, hormonal therapy with goserelin and tamoxifen. The patient also underwent adjuvant radiotherapy. Analysis of genomic DNA extracted from a blood sample revealed no pathogenic germline mutations in the *BRCA1* and *BRCA2* genes.

In September 2017, the patient presented with knee pain. A chest-abdomen-pelvis CT scan showed multiple lytic bone lesions, with soft-tissue involvement in the right iliac (**Table 3.1**). Lytic bone lesions were also detected in lumbar vertebral bodies (**Table 3.1**). The patient enrolled in an open-label clinical phase 3b trial with ribociclib combined with letrozole (CompLEEment-1, NCT02941926). A considerable clinical improvement was observed after treatment, including a decrease in pain score and partial remission of the target bone lesion in the pelvis at the CT scan (**Table 3.1**). In November 2018, the patient

was treated with denosumab (120 mg at 4-week intervals). A partial response of the target and non-target lesions was observed until the 24th cycle (**Table 3.1**).

Genomic Profiling of Plasma cfDNA

Blood samples were collected before (Pre-cfDNA) and after (Post-cfDNA) the patient enrolled in the clinical trial, and cfDNA was extracted as indicated in **Table 3.2**.

A gDNA sample was additionally extracted from the buffy-coat obtained from the blood collected in September 2017. In both cfDNA samples, we detected cfDNA fragments with sizes ranging between 70 to 200 base pairs (bp), with a peak at approximately 150bp (**Figure 3.1**).

Pre-cfDNA was subjected to massively parallel sequencing using the MSK-IMPACT assay that targets 468 cancer-related genes, detecting all protein-coding mutations, copy number alterations, and selected promoter mutations and structural rearrangements (Cheng *et al.*, 2015; Zehir *et al.*, 2017). The sequencing panel includes oncogenes, tumor suppressor genes, and members of pathways deemed actionable by targeted therapies, and are recurrently altered in cancer. This sequencing assay has been employed for the study of >25,000 tumors (Nguyen *et al.*, 2022) as well as cfDNA (De Mattos-Arruda *et al.*, 2014). Two genomic alterations were detected: a missense mutation in the *PIK3CA* gene (3q26.32, **Figure 3.2**) and an amplification of the *CCND1* gene (11q13.3, **Figure 3.3**). The *PIK3CA* hotspot mutation (3:178952085, A>G, H1047R) was present at a variant allele frequency (VAF) of 0.14 (28 out of 200 reads). This variant was present in an estimated cancer cell fraction (CCF) of 0.97, indicating that the variant is likely clonal. The matched normal gDNA sample had a coverage of 199 reads at this position (3:178952085), and no altered reads were detected.

The presence of both molecular alterations was confirmed in gDNA from the primary tumor. In the primary tumor tissue, the hotspot mutation in *PIK3CA* (H1047R) was present with an estimated CCF of 1. This mutation was detected with an estimated purity of 0.34 and a ploidy of 1.8, with a normal allelic depth of 392 and a tumor allelic depth of 93 (VAF = 0.2). Additionally, amplification of the *CCND1* gene was observed, with an estimated fold change of 1.25. No other molecular changes were identified.

The analysis of cfDNA after treatment (Post-cfDNA) did not detect either of the two alterations. In the *PIK3CA* gene, we identified 266 reads covering the position of interest (3:178952085), and none presented this variant (**Figure 3.2**). Moreover, amplification of the *CCND1* gene was no longer observed (**Figure 3.3**). Thus, the results in cfDNA mirror the clinical response.

3.3.7 Discussion

This study highlights the utility of cfDNA analysis for therapy monitoring in metastatic breast cancer patients. Our results are consistent with previous reports indicating that circulating tumor DNA can be used as surrogate marker of treatment outcome (Dawson *et al.*, 2013; Murtaza *et al.*, 2013). Recently, genotyping cfDNA in plasma samples from patients in the randomized phase III PALOMA-3 study of CDK4/6 inhibitor palbociclib and fulvestrant for women with advanced ER+ breast cancer showed that a reduction in the levels of mutant *PIK3CA* DNA detected in circulation correlated with improved progression-free survival (PFS) after treatment (O'Leary *et al.*, 2018). Similarly, patients with ER+ advanced metastatic breast cancer enrolled in the phase I/II randomized BEECH trial (paclitaxel plus placebo versus paclitaxel plus AKT inhibitor capivasertib) with decreased levels of mutant cfDNA detected in plasma after 4 weeks of treatment had substantially improved PFS (Hrebien *et al.*, 2019).

Using massively parallel sequencing to analyze cfDNA in the patient plasma before treatment, we detected the *PIK3CA* hotspot mutation H1047R. *PIK3CA* is one of the two most frequently mutated genes in breast cancers, occurring in 30-40% of cases, and H1047R is the most common mutation in this gene (Pereira *et al.*, 2016; Martínez-Sáez *et al.*, 2020). The *PIK3CA* gene encodes the catalytic subunit of phosphatidylinositol 3-kinase (PI3K), and the H1047R mutation induces gain of enzymatic function, allowing PI3K to signal without regulation and triggering oncogenic properties (Isakoff *et al.*, 2005; Bader, Kang and Vogt, 2006). When present, *PIK3CA* mutations are typically found in both the primary tumor and in the relapsed/metastatic tissue (Yates *et al.*, 2017). Consistent with the finding that *PIK3CA* mutations are predominantly truncal events in breast cancer, we identified the H1047R mutation to be clonal and likely early occurrence in tumor evolution. Notably, a previous study showed that truncal mutations in *PIK3CA* detected by liquid biopsy predicted sensitivity to palbociclib, whereas sub clonal mutations were weak predictors of outcome (O'Leary *et al.*, 2018). More recently, the sequencing of circulating tumor DNA in patients enrolled in the phase III MONALEESA-7 trial revealed a treatment response to endocrine therapy plus ribociclib independent of the *PIK3CA* mutational status (Bardia *et al.*, 2021).

In addition to the *PIK3CA* hotspot mutation, our cfDNA analysis identified the amplification of *CCND1*, an oncogene that encodes the protein cyclin D1. The cyclin dependent kinases 4 and 6 (CDK4/6) form complexes with D-type cyclins that act on the retinoblastoma protein Rb and drive cell cycle progression (Topacio *et al.*, 2019). *CCND1* amplification leads to increased cyclin D1 expression and inappropriate cyclin D-CDK4/6 activity (Elsheikh *et al.*, 2008; Jeffreys *et al.*, 2022), thus promoting sustained cell proliferation, which is one of the hallmarks of cancer (Hanahan and Weinberg, 2011). *CCND1* amplification occurs in 10-35% of breast cancers and is typically associated with positive ER status (Elsheikh *et al.*, 2008; Jeffreys *et al.*, 2022). Breast cancer patients with *CCND1* amplification tend to show a poor response to endocrine therapy (Jeffreys *et al.*, 2022), which may be related to the ability of cyclin D1 to stimulate the growth of estrogen responsive tissues through a CDK-independent mechanism by activating the transcription of ER-regulated genes in the absence of estrogen (Zwijsen *et al.*, 1997). However, the clinical benefit after ribociclib and endocrine therapy was observed in advanced ER+/HER2- breast cancer patients with altered *CCND1* (Bardia *et al.*, 2021).

The patient reported in this study was treated with the CDK4/6 inhibitor ribociclib combined with the aromatase inhibitor letrozole. After treatment, the patient had a significant clinical improvement, and no molecular abnormalities were detected by massively parallel sequencing of cfDNA. A drawback of sequencing cfDNA is the problem of false negatives. Indeed, not all cancer cells release their DNA into circulating blood, and the concentration of cell-free tumor DNA in the plasma may be below the sensitivity of available technologies. However, in this case, the cfDNA results mirrored the clinical response. A limitation of our study is that we did not monitor the patient cfDNA prospectively to determine whether detectable genetic alterations could be detected prior to clinical relapse.

Based on the results of recent trials, ribociclib plus letrozole is currently considered the frontline treatment option in postmenopausal patients with advanced ER+/PR+/HER2- breast cancer (De Laurentiis *et al.*, 2022; Hortobagyi *et al.*, 2022). Although these trials showed a consistent overall survival benefit, future studies are needed to stratify drug response according to sub-groups defined by patient and disease characteristics. In this regard, comprehensive profiling of cfDNA isolated from plasma samples may contribute a

real-time assessment of driver and actionable mutations and their clonal evolution in response to treatment.

3.3.8 Author Contributions

Conception and design: C.S. and M.C.-F. Provision of study materials and clinical data from patients: A.R.S. and L.C. Technical support: A.C.S., P.C. and M.M.; Collection and assembly of sequencing data and bioinformatics analysis: A.D.C.P., P.S., D.N.B., M.G., S.K., S.Z., L.L., B.W., and J.S.R.-F.; Data analysis and interpretation: C.S. and A.D.C.P.; Manuscript writing: C.S. and M.C.-F. All authors have read and agreed to the published version of the manuscript.

3.3.9 Funding

This work was supported by *Fundação para a Ciência e Tecnologia (FCT)/Ministério da Ciência, Tecnologia e Ensino Superior*, Portugal, through *Fundos do Orçamento de Estado* to Instituto de Medicina Molecular João Lobo Antunes (LA/P/0082/2020), and FCT/FEDER/POR Lisboa 2020, *Programa Operacional Regional de Lisboa*, PORTUGAL 2020 (LISBOA-01-0145-FEDER-016394), and FEDER/POR Lisboa 2020-*Programa Operacional Regional de Lisboa*, PORTUGAL 2020 (Infogene, 045300). C.S. was a recipient of a FCT fellowship (SFRH/BDE/110544/2015). This work was funded in part by the National Institutes of Health (NIH)/National Cancer Institute (NCI) Cancer Center Support Grant (P30 CA008748; MSK). J.S.R-F. and B.W. are funded in part by the NIH/NCI P50 CA247749 01 grant and a Breast Cancer Research Foundation grant. J.S.R.-F. is also funded in part by a Susan G Komen leadership grant, and B.W. by a Cycle for Survival grant.

3.3.10 Institutional Review Board Statement

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Lisbon Academic Medical Center (Oncodynamics, ref. n. º 343/16, 12 October 2016).

3.3.11 Informed Consent Statement

Informed consent was obtained from the patient involved in the study.

3.3.12 Acknowledgments

Authors are grateful to the patient who consented to participate in this study.

3.3.13 Conflicts of Interest

J.S.R.-F. reports receiving personal/consultancy fees from Goldman Sachs, REPARE Therapeutics, Paige.AI, Personalis and Bain Capital, membership of the scientific advisory boards of VolitionRx, REPARE Therapeutics, Paige.AI and Personalis, membership of the Board of Directors of Grupo Oncoclinicas, and ad hoc membership of the scientific advisory boards of Roche Tissue Diagnostics, Ventana Medical Systems, Novartis, Genentech, MSD, Daiichi Sankyo and InVicro, outside the scope of this study. B.W. reports research support by REPARE Therapeutics, outside of the scope of this study. M.C.-F. is a founder of "GenoMed-Diagnósticos de Medicina Molecular, SA" and a member of its scientific advisory board; she reports research support by AbbVie and Gilead Sciences, outside of the scope of this study. The remaining authors have nothing to disclose.

Tables

Table 3.1. Lesions follow-up according to the RECIST 1.1 criteria

Target lesion	**Lytic bone lesions, right iliac with soft-tissue involvement**							
Follow-up date	19-Dec-2017	13-Mar-2018	5-Jun-2018	27-Sep-2018	4-Dec-2018	11-Mar-2019	4-Jun-2019	29-Aug-2019
Size	81 mm	58 mm	57 mm	54 mm	48 mm	47 mm	46 mm	42 mm
Non-target lesion	**Lytic bone lesions, lumbar vertebral bodies**							
Follow-up date	19-Dec-2017	13-Mar-2018	5-Jun-2018	27-Sep-2018	4-Dec-2018	11-Mar-2019	4-Jun-2019	29-Aug-2019
Number	Multiple	Multiple	Stable	Stable	Stable	Stable	Stable	Stable

Table 3.2. Blood and plasma sample details.

Sample ID	**Collection date**	**Sample type**	**Sample Concentration (ng/uL)**	**Sample Volume (µL)**
Pre gDNA	18/Sep/2017	Buffy-coat	70,5	40,0
Pre cfDNA	18/Sep/2017	Plasma (4mL)	0,7	45,0
Post cfDNA	25/Jun/2018	Plasma (4mL)	0,3	45,0

Figures

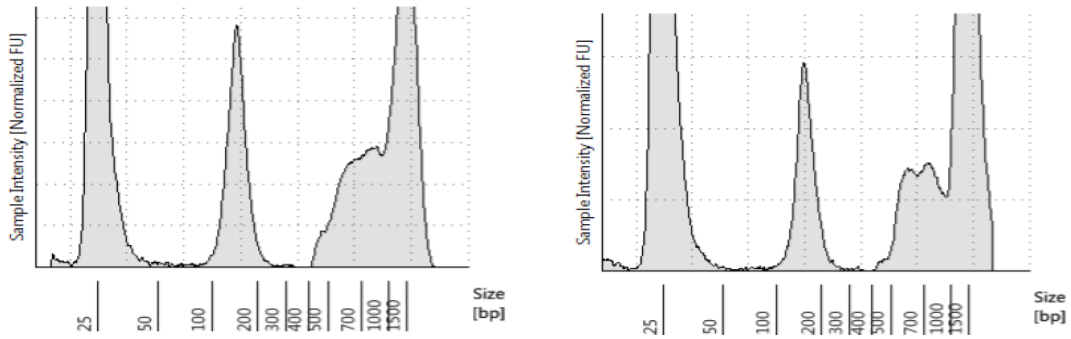


Figure 3.1 *- TapeStation analysis of DNA extracted from plasma collected before (left) and after (right) treatment with ribociclib plus letrozole. Arrows indicate cfDNA fragments (with a peak at ~150 bp).*

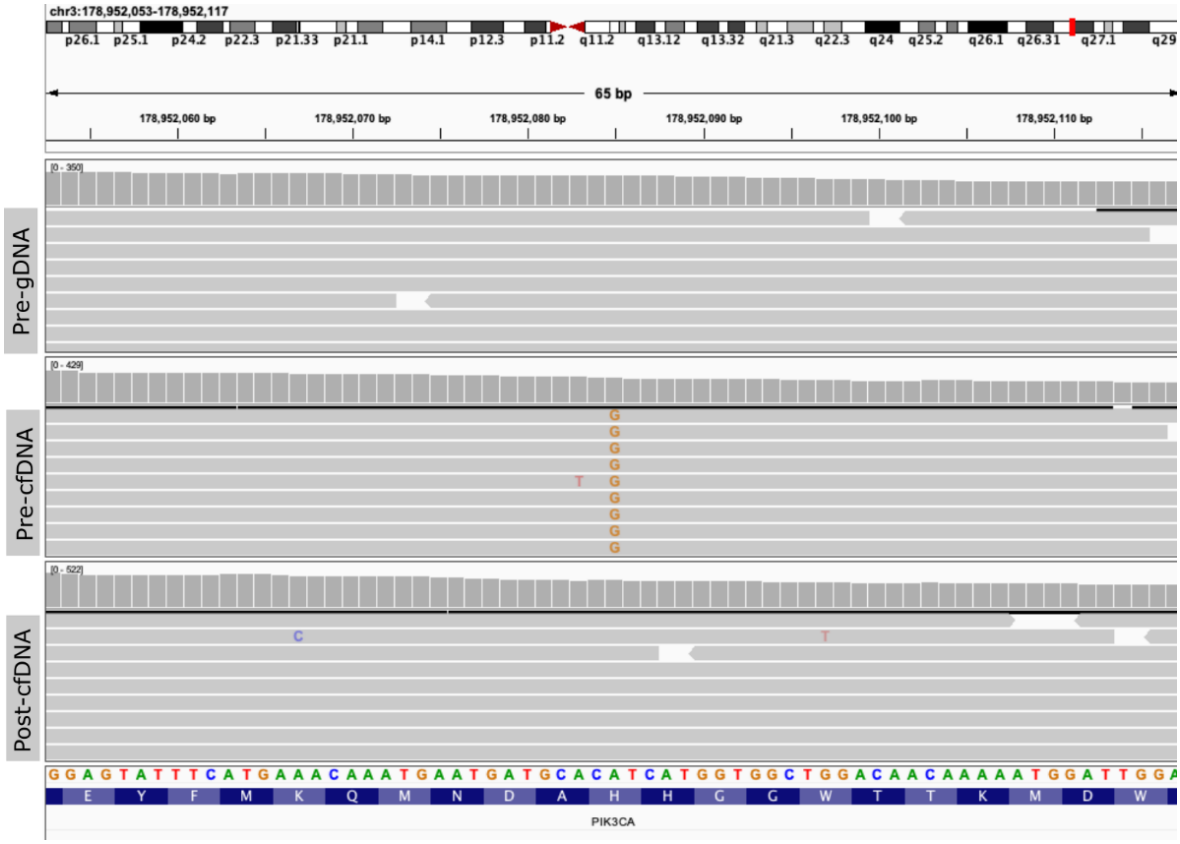


Figure 3.2 *- Integrative Genomics Viewer (IGV) screenshot of the PIK3CA variant detected in cell-free DNA using targeted massively parallel sequencing. The top panel depicts the sequencing reads of the genomic DNA from buffy-coat collected before treatment (Pre-gDNA). The middle and bottom panels depict the sequencing reads of the cell-free (cf)DNA collected*

before (Pre-cfDNA) and after (Post-cfDNA) treatment with ribociclib plus letrozole. PIK3CA hotspot mutation (3:178952085, A>G, H1047R) was detected only in cfDNA before treatment (middle, alternate alleles are shown in orange).

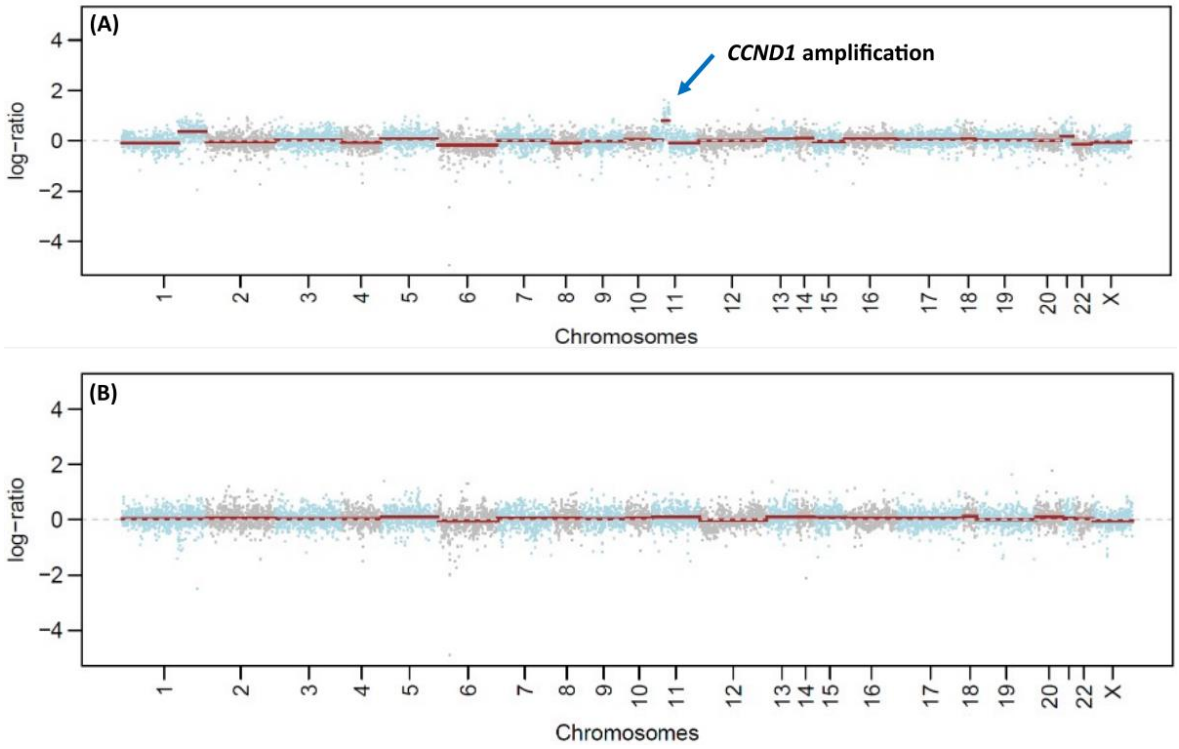


Figure 3.3 - *Copy number analysis of cell-free DNA (cfDNA) extracted from plasma collected before* ***(A)*** *and after* ***(B)*** *treatment with ribociclib plus letrozole. Amplification of the CCND1 gene in chromosome 11 (arrow) was detected at baseline before treatment, but not after therapy.*

3.3b Explore liquid biopsies in urine

3.3b.1 Motivation

Liquid biopsies in urine have the advantage of being less invasive than blood analysis. To determine whether cancer-associated mutations can be identified in urine, we implemented a targeted ddPCR assay for the most common *PIK3CA* mutations, and we screened urine samples from the patient with a *PIK3CA* mutation detected in the plasma.

3.3b.2 Methods

Isolation of cfDNA from urine

Urine samples were collected into two 50mL tubes containing Urine Preservative Single Dose Solution (Norgen Biotek), previously supplied to each patient. According to the manufacturer's instructions, this solution prevents microbial growth and preserves nucleic acids for up to two years, at room temperature. The first urine in the morning was collected. Extraction was performed from the first urine flow (U1), which is expected to be more enriched in cell-free DNA. Midstream urine (U2) was additionally used in order to understand if there were differences between samples.

Urine samples were centrifuged at 200g for 10 minutes and the supernatant was transferred into falcon tubes. Then, samples were centrifuged at 16000g for 10minutes in a high-speed centrifuge at 4^{o}C. After this, the supernatant was transfer to a new tube. At this point, the samples were ready for cfDNA extraction or could be kept at 4^{o}C until further use (**Figure 3.4**).

For cfDNA isolation, two commercial kits were used. MagMAX™ Cell-Free DNA Isolation kit (Applied Biosystems™) and QIAamp® Circulating Nucleic Acid (Qiagen). Using the MagMAX™ Cell-Free DNA Isolation kit (Applied Biosystems™) 10mL of urine sample were used, while using QIAamp® Circulating Nucleic Acid kit (Qiagen), 4mL of urine sample were used. The extractions were conducted according to the manufacturer's instructions. The cfDNA was eluted in 20µL (MagMAX™ Cell-Free DNA Isolation kit) or 30-50µL (QIAamp®

Circulating Nucleic Acid kit) of ultra-clean water. cfDNA was quantified using the Qubit® 3.0 Fluorometer (Invitrogen, Life Technologies) with Qubit® dsDNA HS Assay kit (Invitrogen, Life Technologies). Sample quality was accessed using High Sensitivity D1000 ScreenTape (TapeStation 4200, Agilent Technologies), according to the manufacturer's instructions (**Figures 3.5 to 3.9**). The purified cfDNA was stored at 4^{o}C for up to 16 hours or at -80^{o}C for longer periods.

Detection of *PIK3CA* E545K and H1047R mutations by Droplet Digital PCR (ddPCR)

We used the QX200™ Droplet Digital PCR System (Bio-Rad, Hercules, CA, USA), and Bio-Rad unique assays (dHsaMDV2010075 and dHsaMDV2010077). For the *PIK3CA* E545K mutation the reference standard used was the Horizon DNA HD689 (Horizon Discovery Ltd., Cambridge, UK). For the H1047R we used MFM-223 cell line as reference standard. Each of the standards worked as negative control for the other mutation.

After several tests regarding the annealing temperature and the number of cycles, the optimal annealing temperature was established at 54^{o}C. Altogether, the thermocycling conditions included an initial incubation at 95^{o}C for 10 minutes, 45 cycles of 94^{o}C for 30 seconds and 54^{o}C for 1 minute, followed by an inactivation step at 98^{o}C for 10 minutes. Two replicates were analyzed per sample, and the total amount of DNA loaded (considering the two replicates), range from 1 to 30ng (**Table 3.3**). No template control (NTC) was used to exclude PCR contamination.

Amplification results were analyzed using QX200™ Droplet Reader and QuantaSoft™ software. Samples with three or more positive mutant droplets were considered positive, as recommended by the best practice guidelines for rare mutation detection (Biorad, 2018).

Statistical analysis

The concentrations of target alleles were calculated using QuantaSoft^{TM} software (Bio-Rad) based on Poisson distribution. Mutant cfDNA molecules were reported as number of copies per millilitre (mL) of urine. The mutant allelic frequency was determined as the ratio of mutant droplets relative to the sum of mutant and wild-type droplets.

3.3b.2 Results

Urine samples analysis using ddPCR

As demonstrated in **Figures 3.10** to **3.13** the ddPCR assay was successfully implemented. The 2D plots presented are divided in four quadrants: positive droplets for the mutations (blue, up left), empty droplets (grey, bottom left), double positive droplets (orange, up right), and wild-type droplets (green, down right).

In **Figures 3.10** and **3.12** we are able to see the 2D plot of the positive controls (VAF=1%), with the representation of the positive droplets (blue, up left quadrant) for each mutation (**Figures 3.10.** E545K and **Figure 3.12.** H1047R). In **Figures 3.11** and **3.13** we are able to see the 2D plot of the negative controls with the representation of the wild-type droplets (green, bottom right quadrant) together with the absence of positive droplets, which demonstrate the specificity of the assay.

The results for the hotspot mutation of interest (H1047R) obtained in urine samples collected before the patient enrolled in the clinical trial are represented in **Figures 3.14** to **3.18**. Overall, 2 samples were tested (urine first flow and midstream urine) and 5 attempts were made. Three independent extractions were done for first flow urine (one with MagMax kit and two with QIAamp kit), and two independent extractions for midstream urine (both with QIAamp kit). All samples were validated since all presented >10.000 droplets per well (**Table 3.3**) (Biorad, 2018), however, the scant concentration present in all urine first flow attempts in patent in the figures since the vast majority of droplets found were empty droplets (without DNA template). The midstream urine, showed better results, with a good quantity of wild-type droplets, but no positive droplet was detected.

3.3b.3 Discussion

This study sought to monitor response to therapy by liquid biopsy, in a case with metastatic breast cancer, using cfDNA extracted from plasma and urine. These would be two less invasive approaches to monitor disease evolution and both have been exploited in previous studies. Recently, Guan et al. and Zhang et al. showed the feasibility of the use of plasma and midstream urine cfDNA in real-time monitoring of BC patients. They studied a group of patients with BC and *PIK3CA* mutation and/or a *TP53* mutation that they seek to follow. Concordance between tissue and cfDNA was access. Moreover, they have demonstrated the specificity of their studies by comparing the amount of urinary cfDNA with healthy volunteers and reported an increased amount of urinary cfDNA in the BC patients. Additionally, cfDNA measurements in the healthy donors are stable when compared with cancer patients (Guan *et al.*, 2020; Zhang, Zhang and Shen, 2020; Zuo *et al.*, 2020). Guan et al. also reported that BC patients with either *PIK3CA* or *TP53* mutation showed higher levels of urinary cfDNA when compared with wild-type BC patients. Another observation of interest was that after surgery, there was a significant drop in the amount of cfDNA in the BC patients, while healthy donors maintained the baseline levels of cfDNA. For this reason, there was several patients in which they could not detect mutations by ddPCR in the post-surgery sample (Guan *et al.*, 2020).

The first limitation present in this work is poor sampling. A single patient to test, with a known MPS result (*PIK3CA* H1047R), allowing us to seek the same variant in urine by a different methodology. The second limitation was the absence of enough plasma sample to try to reproduce the MPS result with the ddPCR technology. The third limitation was the uncertainty regarding the stability of cfDNA in urine. The use of preservative solution and storage at room temperature for more than a year until de cfDNA extraction might have influenced the performance of our assays. Additionally, there was an unexpected result regarding the cfDNA extraction from urine samples. It was expected to obtain a higher quantity of cfDNA in the first flow urine, compared with midstream urine. It was also expected that the first flow urine might present a considerable amount of genomic DNA and that was not noticed in the TapeStation results in any of the three extractions. This might reflect the poor stability of the nucleic acids in this sample type, even using the

preservative solution. Nevertheless, it is now well established that the use of the midstream urine must prevail and that the samples should be processed almost immediately (Guan *et al.*, 2020; Zhang, Zhang and Shen, 2020; Zuo *et al.*, 2020).

Albeit the lack of success using ddPCR and urine samples to follow-up the treatment response, these results demonstrate that profiling of cancer-associated genes by MPS in plasma samples is an effective strategy for monitoring treatment response in a BC patient with metastatic disease.

Tables

Table 3.3. Detailed information about extraction kit used in urine samples, ddPCR DNA input and obtained results

		E545K		H1047R	
Sample ID	**Input (ng)**	**Accepted Droplets**	**Positive droplets**	**Accepted Droplets**	**Positive droplets**
ddPCR #1 assay					
NTC	0,0	32636	0	38432	0
CTRL NEG	15,0	28556	0	35533	0
CTRL POS 1%	15,0	30675	14	35981	19
Pre cfDNA #1.U1	1,0	35435	0	38603	0
ddPCR #2 assay					
NTC	0,0	34702	0	32317	0
CTRL NEG	15,0	33088	0	31373	0
CTRL POS 1%	15,0	33676	14	36581	15
Pre cfDNA #2.U1	5,0	33399	0	33539	0
Pre cfDNA #3.U2	30,2	33111	0	30056	0
Pre cfDNA #4.U1	15,5	30793	0	27475	0
Pre cfDNA #5.U2	30,1	34693	0	29331	0

U1 – urine first flow

U2 – midstream urine

Figures

Figure 3.4 *- Schematic representation of the urine sample testing with ddPCR.*

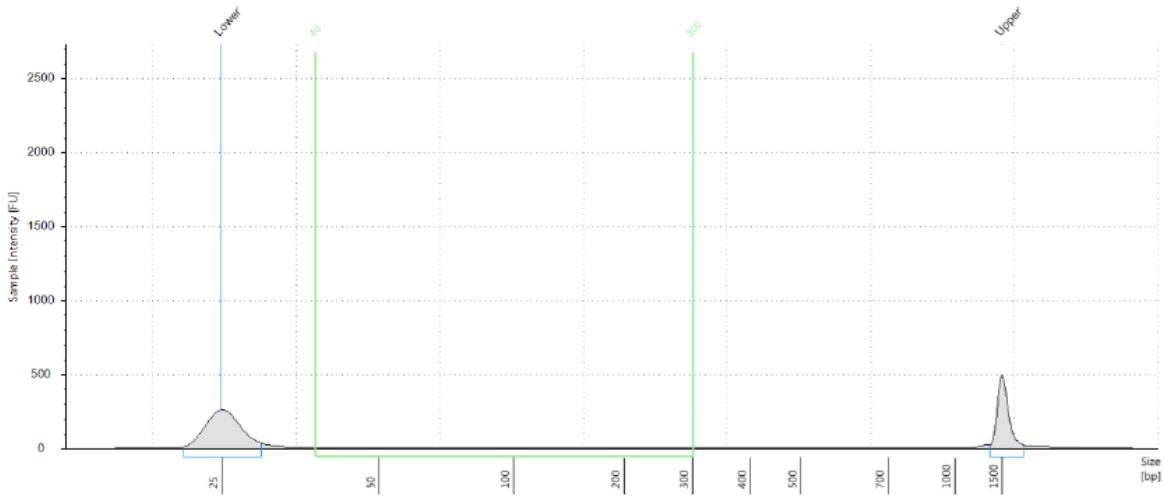


Figure 3.5 *- Pre cfDNA #1.U1. cfDNA sample extracted from first flow urine, using MagMAX™ Cell-Free DNA Isolation kit. This image shows a too low concentrated sample. A region was selected, between 40 and 300bp to calculate the possible cfDNA present and a 18,4pg/µL concentration was obtained (0,0184ng/µL) This sample presented a 0,13ng/µL concentration when quantified with Qubit® 3.0.*

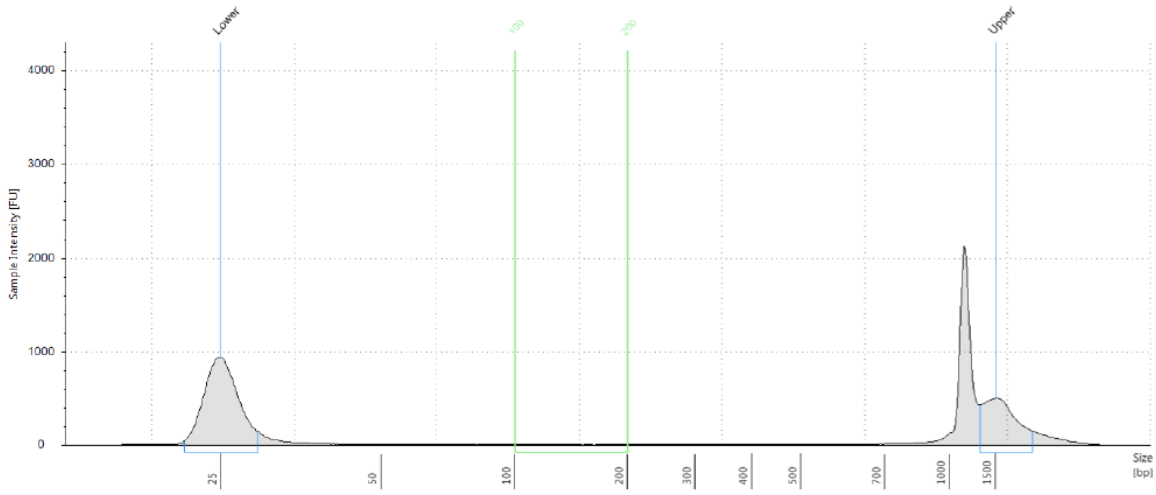


Figure 3.6 *- Pre cfDNA #2.U1. cfDNA sample extracted from first flow urine, using QIAamp® Circulating Nucleic Acid kit. This image shows a too low concentrated sample. A region was selected, between 100 and 200bp to calculate the possible cfDNA present and a 0.597pg/µL concentration was obtained (0,000597ng/µL) This sample presented a 0,315ng/µL concentration when quantified with Qubit® 3.0.*

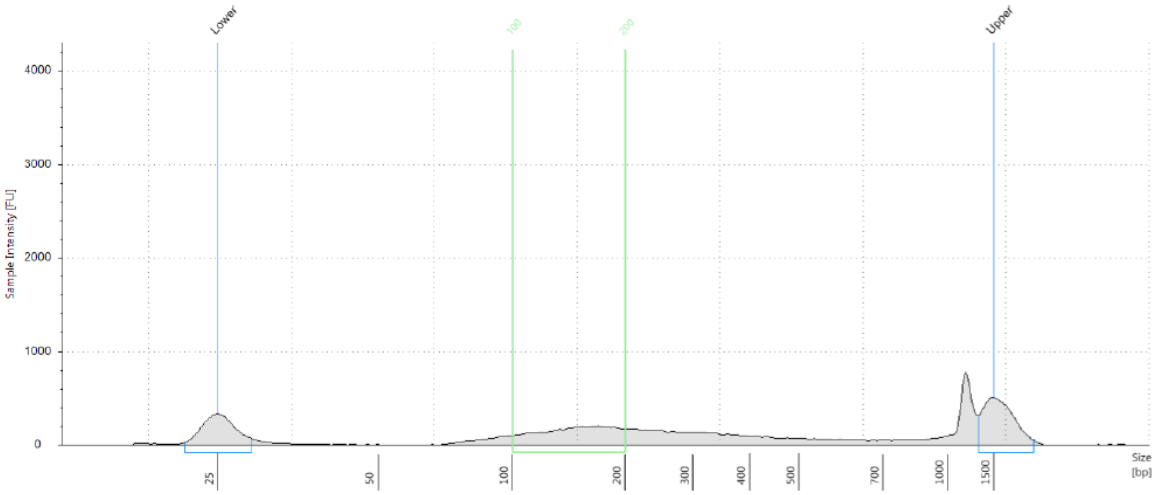


Figure 3.7 *- Pre cfDNA #3.U2. cfDNA sample extracted from midstream urine, using QIAamp® Circulating Nucleic Acid kit. A region was selected, between 100 and 200bp to calculate the possible cfDNA present and a 242pg/µL concentration was obtained (0,242ng/µL) This sample presented a 3,97ng/µL concentration when quantified with Qubit® 3.0.*

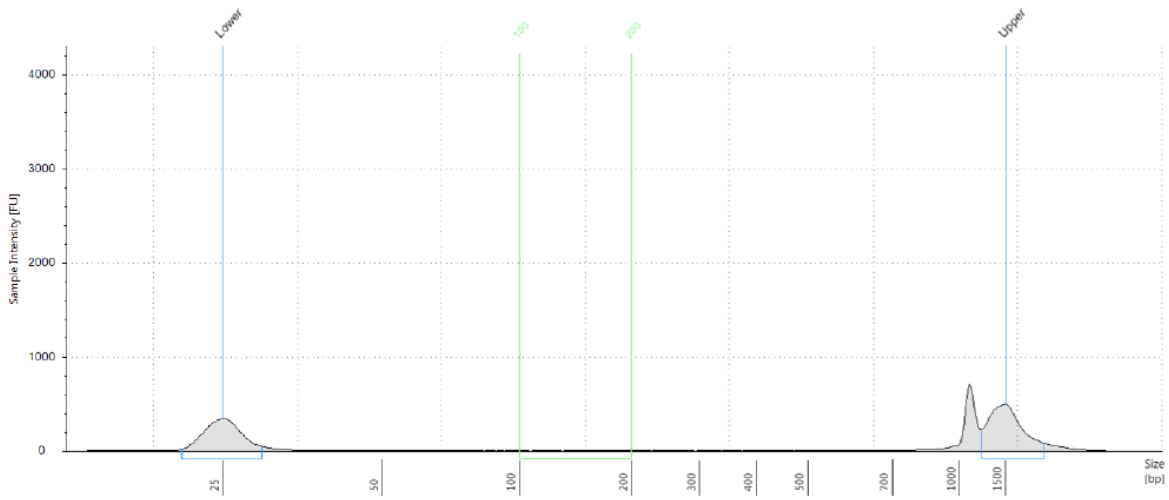


Figure 3.8 *- Pre cfDNA #4.U1. cfDNA sample extracted from first flow urine, using QIAamp® Circulating Nucleic Acid kit. This image shows a too low concentrated sample. A region was selected, between 100 and 200bp to calculate the possible cfDNA present and a 0.528pg/µL concentration was obtained (0,000528ng/µL) This sample presented a 0,980ng/µL concentration when quantified with Qubit® 3.0.*

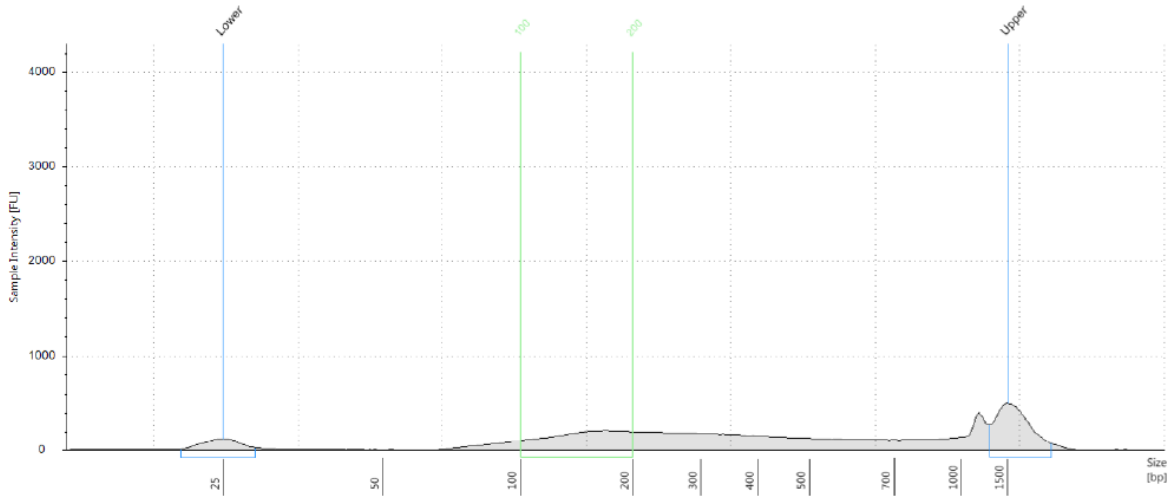


Figure 3.9 *- Pre cfDNA #5.U2. cfDNA sample extracted from midstream urine, using QIAamp® Circulating Nucleic Acid kit. A region was selected, between 100 and 200bp to calculate the possible cfDNA present and a 245pg/µL concentration was obtained (0,245ng/µL) This sample presented a 8,13ng/µL concentration when quantified with Qubit® 3.0.*

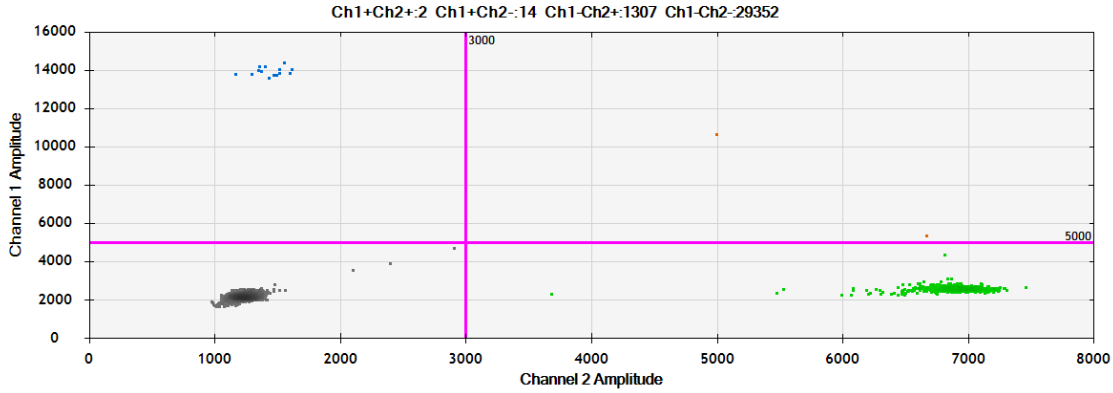


Figure 3.10 *- 2D plot results for PIK3CA E545K positive control (VAF=1%). In the up left quadrant, we are able to see the positive droplets (blue) for E545K mutation.*

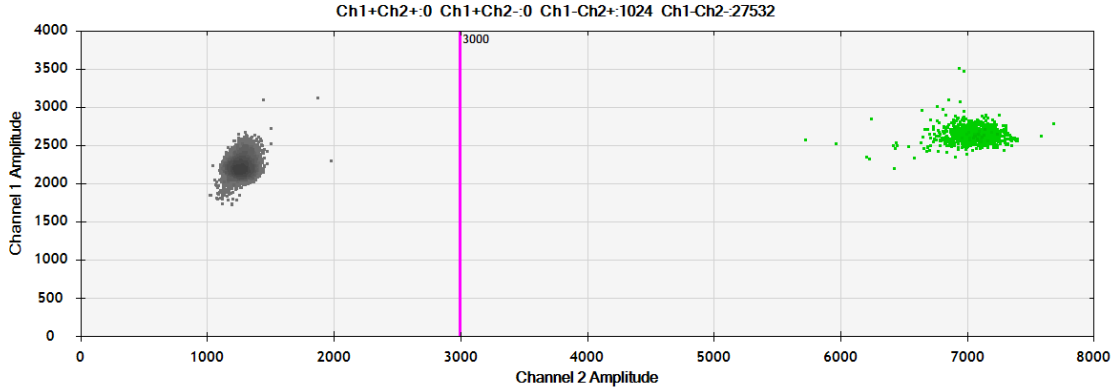


Figure 3.11 *- 2D plot results for PIK3CA E545K WT, represented by the green droplets present in the bottom right quadrant.*

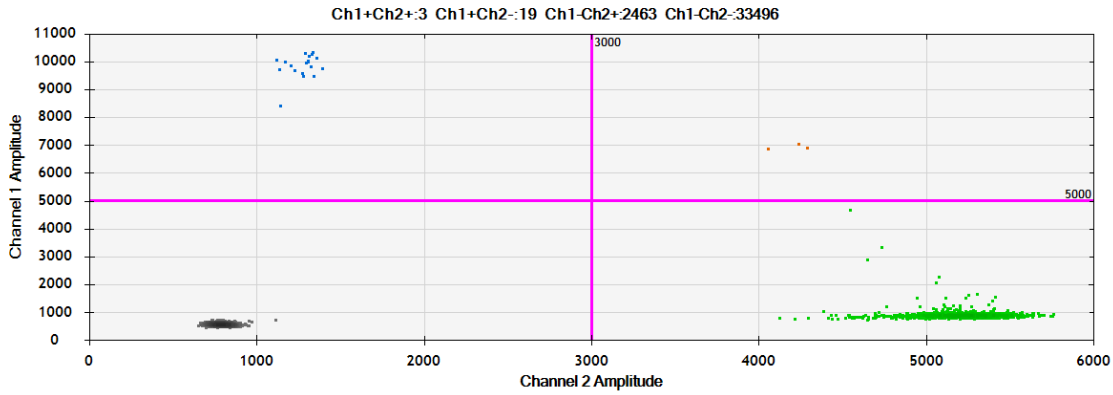


Figure 3.12 *- 2D plot results for PIK3CA H1047R positive control (VAF=1%). In the up left quadrant, we are able to see the positive droplets (blue) for H1047R mutation.*

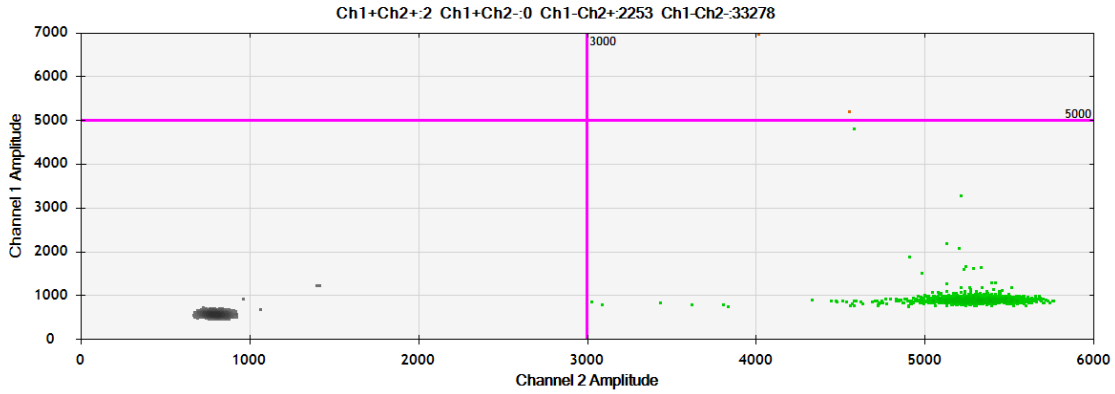


Figure 3.13 *- 2D plot results for PIK3CA H1047R WT represented by the green droplets present in the bottom right quadrant.*

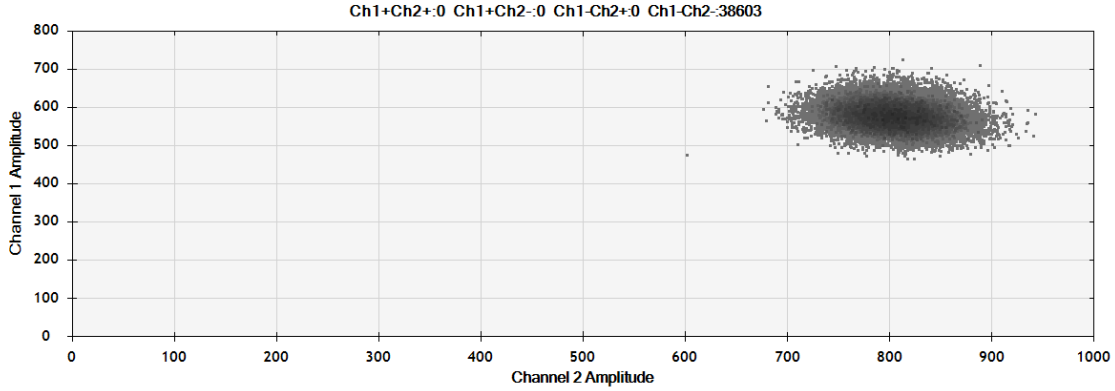


Figure 3.14 *- 2D plot results for urine sample Pre cfDNA #1.U1 – PIK3CA H1047R. The absence of WT or positive droplets indicates the unsuccess of this experiment, due to the fact of having scarce cfDNA input.*

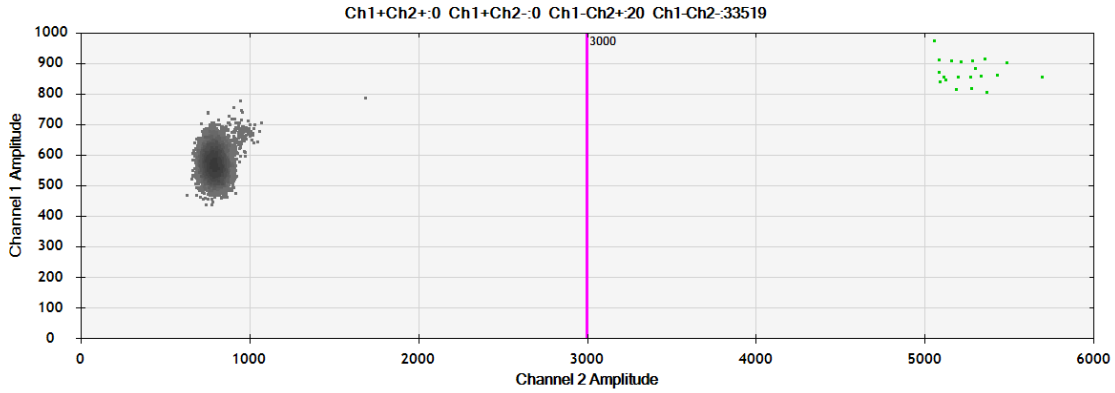


Figure 3.15 *- 2D plot results for urine sample Pre cfDNA #2.U1 – PIK3CA H1047R. In the bottom right, the presence of scarce WT droplets demonstrates the presence of a low concentrated sample.*

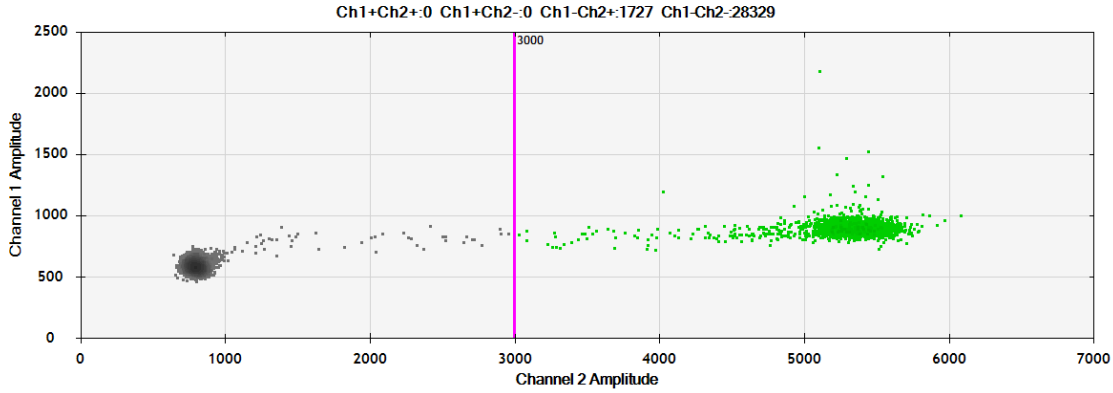


Figure 3.16 *- 2D plot results for urine sample Pre cfDNA #3.U2 – PIK3CA H1047R. WT droplets present in a good number, but no positive droplets to show.*

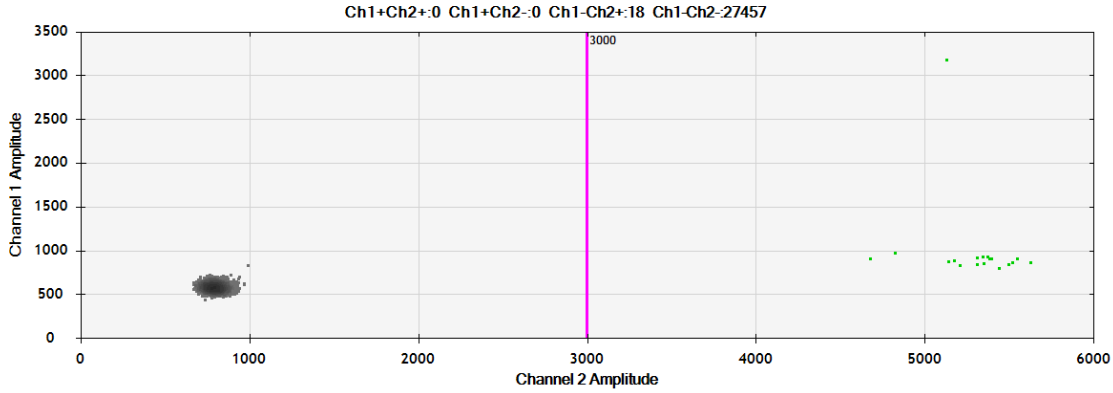


Figure 3.17 *- 2D plot results for urine sample Pre cfDNA #4.U1 – PIK3CA H1047R. In the bottom right, the presence of scarce WT droplets demonstrates the presence of a low concentrated sample.*

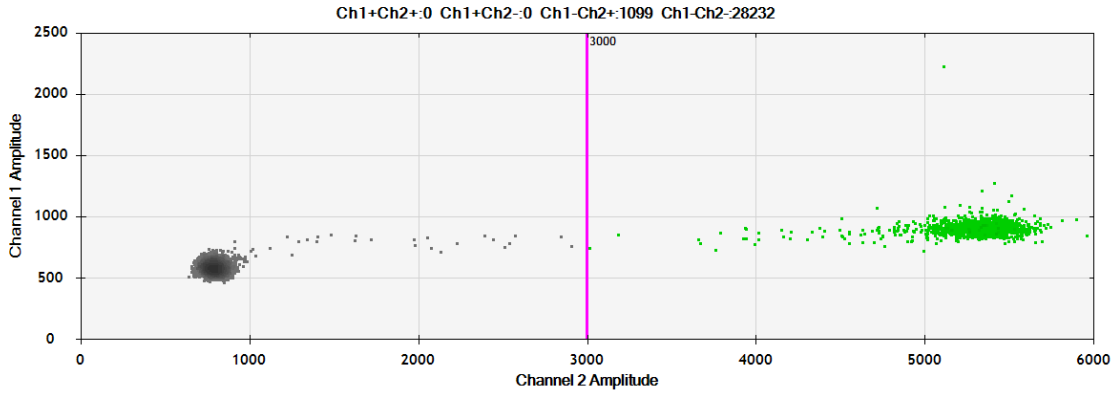


Figure 3.18 *- 2D plot results for urine sample Pre cfDNA #5.U2 – PIK3CA H1047R. WT droplets present in a good number, but no positive droplets to show.*

4. Discussion

Cancers have a diverse repertoire of somatic mutations and intratumor genetic heterogeneity has been shown, as well as discrepancies in the mutational status of original tumors and metastasis. This intratumor genetic heterogeneity, in both space and time, might represent a challenge for precision medicine and remains the doubt whether the best way to assess cancer's genetic repertoire is to look at primary tumor tissue, metastatic sites, circulating biomarkers, or all three (Bidard, Weigelt and Reis-Filho, 2013).

Improvements in molecular profiling have resulted in a rapid increase in the number of predicted molecular biomarkers and related targeted therapies, requiring large-scale, prospective tumor profiling studies for all cancer types. Tumor tissue is the primary specimen of choice for biomarker detection in the majority of comprehensive MPS-based profiling approaches. Despite their widespread use, acquiring a sufficient tissue sample remains one of the main challenges in some situations due to the requirement for invasive biopsies (Rose Brannon *et al.*, 2021). Moreover, there are several potential limitations regarding tissue sampling, which lead to assay failures, such as low tumor cellularity, low DNA yield, or quality (due to DNA fragmentation in FFPE samples, for instance) (Rolfo *et al.*, 2020; Rose Brannon *et al.*, 2021). Besides, a single tumor biopsy may not be representative of genetic heterogeneity present in the tumor, and, for this reason, the genetic repertoire will not be fully accessed. Taken together, this leads us towards the advantage of use cfDNA sample to perform tumor profiling (Thorat, 2019; Rakha and Pareja, 2021; Rose Brannon *et al.*, 2021). However, there are still some limitations regarding the use of ctDNA as a biomarker of interest. The limit of detection for detecting ctDNA is still not low enough, implying that the assay's sensitivity is insufficient to detect mutations when there is a small tumor burden, when there is low tumor shed into the blood, or when low mutant allele frequencies are present. Even in metastatic disease with relatively higher ctDNA quantities, the current limit of detection leads to poor repeatability between assays, leading to discordance between ctDNA results (Qin, 2019; Thorat, 2019; Li *et al.*, 2020; Turner *et al.*, 2020; Rose Brannon *et al.*, 2021).

More sensitive assays, like ddPCR could be applied to overcome this challenge. However, they are quite limited by the fact of interrogating a single *locus* at a time, given an insufficient response to the necessities (Rose Brannon *et al.*, 2021). Another alternative is the use of more sensitive assays using MPS technology. The used of unique molecular identifiers (barcodes), together with deep sequencing (~20.000x raw coverage), helps to reduce the error rates and gives a more sensitive and precise result (Qin, 2019; Rose Brannon *et al.*, 2021). However, considering the coverage needed to obtain these trustworthy results a high-throughput sequencing machine is needed, and the costs are substantial. Moreover, to improve analysis, and to accurate report only somatic mutations, there is a need of sequencing gDNA together with the cfDNA sample, duplicating the cost with the test (Rose Brannon *et al.*, 2021). The use of a normal sample is crucial namely to filter germline variants as well as somatic variants related to clonal hematopoiesis, thus avoiding false-positive results (Rolfo *et al.*, 2020; Cheng *et al.*, 2021; Rose Brannon *et al.*, 2021).

cfDNA analysis for therapy selection, treatment response monitoring, and resistance mechanisms detection has been widely used, namely in NSCLC. PCR-based approaches have been validated and approved for this purpose, however, it is known that MPS panels would be able to give more complete information (Rolfo *et al.*, 2020; Cheng *et al.*, 2021). In BC, the clinical value of ctDNA analysis is also undeniable and has emerged as a diagnostic tool for patient selection and disease monitoring (Rolfo *et al.*, 2020). According to NCCN guidelines, luminal BC patients should be tested for *PIK3CA* mutations in cfDNA, but in the presence of a negative result, tumor testing is recommend (Cheng *et al.*, 2021).

Even though there are no doubts about the utility of the use of ctDNA MPS profiling in cancer, the use of liquid biopsies in clinical practice is not well established but it will be in a near future (Rolfo *et al.*, 2020; Cheng *et al.*, 2021).

Regarding the four main objectives of the present work, we were able to study three main methodologies (MPS, Sanger sequencing, and PCR-based solutions), and we managed to work with three types of samples (tumor-tissue, plasma cfDNA, and urinary cfDNA).

In the first study, a cohort of patients was screened to seek uncommon genetic alterations in the *EGFR* gene, which were not detectable by commercial PCR-based methods targeted-methods. We conclude that PCR-based targeted-methods for *EGFR* genotyping can only

identify the most common genetic alterations, missing about 3% of rare mutations that are only detected by sequencing methods and that might give therapeutic options to the patients.

Regarding the second objective of this work, a ddPCR strategy was optimized to offer *EGFR* T790M mutation detection in plasma samples NSCLC patients in progression and with acquired resistance to first- or second-generation EGFR-TKIs. It was possible to detect *EGFR* T790M mutation in cfDNA with an allele frequency as low as 0.5%. Moreover, it was showed that repeated ddPCR tests in cfDNA may obviate tissue re-biopsy in patients unable to provide a tumor tissue sample.

Moving into the third challenge, MPS technology was tested to look for mutations in cfDNA that would allow a close follow-up of a patient with metastatic breast cancer. We were managed to find two alterations in two common cancer genes (*PIK3CA* and *CCND1*), and in the follow-up sample, these two alterations were no longer detected in the cfDNA, highlighting the utility of liquid biopsy to monitor drug response. In the continuation of this study, a ddPCR assay was implemented to screen the *PIK3CA* mutation in the urinary cfDNA of the same patient. Although the implementation was well succeeded, we failed to detect the mutation in the urine.

4.1 Conclusions and future steps

In this work, we analyzed the most frequently used methods in Precision Oncology and sought to address the main advantages and disadvantages of each technology. Moreover, the importance of the sample chosen for each scenario was also addressed and has been demonstrated that this is a crucial step for the success of the decision.

Although the use of cfDNA it is not well established in the clinical routine, there are no doubts that liquid biopsy is a very promising tool in cancer detection and diagnosis, and big research efforts are underway to tackle the limitations so that it can reach its full potential.

At GenoMed we are committed to providing the most cutting-edge solutions for our doctors and patients, and we will continue to work towards a more achievable and cost-effective Precision Medicine.

With the increasing number of variants detected in the MPS panels used in the clinical practice, we strongly recommend that the results are evaluated in a context of a multidisciplinary molecular tumor board (MTB), and we are available to participate in these meetings and provide our expertise so that the patients can benefit from their results.

References

AF Gazdar (2009) 'Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors', *Oncogene*, 28(Suppl 1), pp. S24–S31. doi: 10.1038/onc.2009.198.

Aparicio, S. and Caldas, C. (2013) 'The Implications of Clonal Genome Evolution for Cancer Medicine', *New England Journal of Medicine*, 368(9), pp. 842–851. doi: 10.1056/nejmra1204892.

Arcila, M. E. *et al.* (2013) 'EGFR Exon 20 Insertion Mutations in Lung Adenocarcinomas: Prevalence, Molecular Heterogeneity, and Clinicopathologic Characteristics', *Molecular Cancer Therapeutics*, 12(2), pp. 220–229. doi: 10.1158/1535-7163.MCT-12-0620.

Asahina, H. *et al.* (2006) 'Non-responsiveness to gefitinib in a patient with lung adenocarcinoma having rare EGFR mutations S768I and V769L', *Lung Cancer*, 54(3), pp. 419–422. doi: 10.1016/j.lungcan.2006.09.005.

Bader, A. G., Kang, S. and Vogt, P. K. (2006) 'Cancer-specific mutations in PIK3CA are oncogenic in vivo.', *Proceedings of the National Academy of Sciences of the United States of America*, 103(5), pp. 1475–9. doi: 10.1073/pnas.0510857103.

Bardia, A. *et al.* (2021) 'Genomic Profiling of Premenopausal HR+ and HER2- Metastatic Breast Cancer by Circulating Tumor DNA and Association of Genetic Alterations With Therapeutic Response to Endocrine Therapy and Ribociclib.', *JCO precision oncology*, 5. doi: 10.1200/PO.20.00445.

Bidard, F. C., Weigelt, B. and Reis-Filho, J. S. (2013) 'Going with the flow: From circulating tumor cells to DNA', *Science Translational Medicine*, 5(207). doi: 10.1126/scitranslmed.3006305.

Biorad (2018) 'Droplet Digital ™ PCR Droplet Digital ™ PCR Applications Guide', *Biorad*, p.

145. Available at: http://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6407.pdf.

Bray, F. *et al.* (2018) 'Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries', *CA: A Cancer Journal for Clinicians*, 68(6), pp. 394–424. doi: 10.3322/caac.21492.

Caputo, V. *et al.* (2022) 'Comprehensive genome profiling by next generation sequencing of circulating tumor DNA in solid tumors: a single academic institution experience.', *Therapeutic advances in medical oncology*, 14(X), p. 17588359221096878. doi: 10.1177/17588359221096878.

Carmo-Fonseca, M. (2021) 'Precision Medicine: Enabling Healthcare Progress in the Twenty-First Century', in Fonseca, A. C. and Ferro, J. M. (eds) *Precision Medicine in Stroke*. Cham: Springer International Publishing, pp. 9–19. doi: 10.1007/978-3-030-70761-3_2.

Carter, S. L. *et al.* (2012) 'Absolute quantification of somatic DNA alterations in human cancer', *Nature Biotechnology*. Nature Publishing Group, 30(5), pp. 413–421. doi: 10.1038/nbt.2203.

Caswell-Jin, J. L. *et al.* (2018) 'Change in Survival in Metastatic Breast Cancer with Treatment Advances: Meta-Analysis and Systematic Review', *JNCI Cancer Spectrum*, 2(4), pp. 1–10. doi: 10.1093/JNCICS/PKY062.

Chaffer, C. L. and Weinberg, R. A. (2011) 'A perspective on cancer cell metastasis', *Science*, 331(6024), pp. 1559–1564. doi: 10.1126/science.1203543.

Chen, Z. *et al.* (2008) 'EGFR somatic doublets in lung cancer are frequent and generally arise from a pair of driver mutations uncommonly seen as singlet mutations: One-third of doublets occur at five pairs of amino acids', *Oncogene*, 27(31), pp. 4336–4343. doi: 10.1038/onc.2008.71.

Cheng, D. T. *et al.* (2015) 'Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT)', *The Journal of Molecular Diagnostics*. American

Society for Investigative Pathology and the Association for Molecular Pathology, 17(3), pp. 251–264. doi: 10.1016/j.jmoldx.2014.12.006.

Cheng, M. L. *et al.* (2021) 'Circulating tumor DNA in advanced solid tumors: Clinical relevance and future directions', *CA: A Cancer Journal for Clinicians*, 71(2), pp. 176–190. doi: 10.3322/caac.21650.

Cibulskis, K. *et al.* (2013) 'Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples', *Nature Biotechnology*. Nature Publishing Group, 31(3), pp. 213–219. doi: 10.1038/nbt.2514.

Coombes, R. C. *et al.* (2019) 'Personalized detection of circulating tumor DNA antedates breast cancer metastatic recurrence', *Clinical Cancer Research*, 25(14), pp. 4255–4263. doi: 10.1158/1078-0432.CCR-18-3663.

Dela Cruz, C. S., Tanoue, L. T. and Matthay, R. a. (2011) 'Lung Cancer: epidemiology, etiology and prevention', *Clin Chest Med*, 32(4), pp. 1–61. doi: 10.1016/j.ccm.2011.09.001.Lung.

Da Cruz Paula, A. *et al.* (2020) 'Genomic profiling of primary and recurrent adult granulosa cell tumors of the ovary', *Modern Pathology*. Springer US, 33(8), pp. 1606–1617. doi: 10.1038/s41379-020-0514-3.

Dagher, R. *et al.* (2002) 'Approval summary: Imatinib mesylate in the treatment of metastatic and/or unresectable malignant gastrointestinal stromal tumors', *Clinical Cancer Research*, 8(10), pp. 3034–3038. doi: 10.1634/theoncologist.2008-0255.

Dang, D. K. and Park, B. H. (2022) 'Circulating tumor DNA: current challenges for clinical utility.', *The Journal of clinical investigation*, 132(12). doi: 10.1172/JCI154941.

Dawson, S.-J. *et al.* (2013) 'Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer', *New England Journal of Medicine*, 368(13), pp. 1199–1209. doi: 10.1056/nejmoa1213261.

DePristo, M. A. *et al.* (2011) 'A framework for variation discovery and genotyping using

next-generation DNA sequencing data', *Nature Genetics*, 43(5), pp. 491–498. doi: 10.1038/ng.806.

Dueck, M. E. *et al.* (2019) 'Precision cancer monitoring using a novel, fully integrated, microfluidic array partitioning digital PCR platform.', *Scientific reports*, 9(1), p. 19606. doi: 10.1038/s41598-019-55872-7.

Ellison, G. *et al.* (2013) 'EGFR mutation testing in lung cancer: a review of available methods and their use for analysis of tumour tissue and cytology samples.', *Journal of clinical pathology*. BMJ Publishing Group, 66(2), pp. 79–89. doi: 10.1136/jclinpath-2012-201194.

Elsheikh, S. *et al.* (2008) 'CCND1 amplification and cyclin D1 expression in breast cancer and their relation with proteomic subgroups and patient outcome.', *Breast cancer research and treatment*, 109(2), pp. 325–35. doi: 10.1007/s10549-007-9659-8.

Ettinger, D. S. *et al.* (2018) 'NCCN Guidelines Insights: Non–Small Cell Lung Cancer, Version 5.2018', *Journal of the National Comprehensive Cancer Network*, 16(7), pp. 807–821. doi: 10.6004/jnccn.2018.0062.

Ferlay, J., Parkin, D. M. and Steliarova-Foucher, E. (2010) 'Estimates of cancer incidence and mortality in Europe in 2008', *European Journal of Cancer*, 46(4), pp. 765–781. doi: 10.1016/j.ejca.2009.12.014.

Frega, S. *et al.* (2016) 'A Triple Rare E709K and L833V/H835L EGFR Mutation Responsive to an Irreversible Pan-HER Inhibitor: A Case Report of Lung Adenocarcinoma Treated with Afatinib', *Journal of Thoracic Oncology*, 11(5), pp. e63–e64. doi: 10.1016/j.jtho.2016.01.023.

Gajiwala, K. S. (2013) 'EGFR: Tale of the C-terminal tail', *Protein Science*, 22(7), pp. 995–999. doi: 10.1002/pro.2283.

Gambardella, V. *et al.* (2020) 'Personalized Medicine: Recent Progress in Cancer Therapy.', *Cancers*, 12(4), pp. 1–3. doi: 10.3390/cancers12041009.

Garcia-Murillas, I. *et al.* (2019) 'Assessment of Molecular Relapse Detection in Early-Stage Breast Cancer', *JAMA Oncology*, 5(10), pp. 1473–1478. doi: 10.1001/jamaoncol.2019.1838.

Golkaram, M. *et al.* (2021) 'HERVs establish a distinct molecular subtype in stage II/III colorectal cancer with poor outcome', *npj Genomic Medicine*. Springer US, 6(1), p. 13. doi: 10.1038/s41525-021-00177-w.

Goto, K. *et al.* (2012) 'Epidermal growth factor receptor mutation status in circulating free DNA in serum: From IPASS, a phase III study of gefitinib or carboplatin/paclitaxel in non-small cell lung cancer', *Journal of Thoracic Oncology*. International Association for the Study of Lung Cancer, 7(1), pp. 115–121. doi: 10.1097/JTO.0b013e3182307f98.

De Grève, J. *et al.* (2016) 'Prospective Evaluation of First-Line Erlotinib in Advanced Non-Small Cell Lung Cancer (NSCLC) Carrying an Activating EGFR Mutation: A Multicenter Academic Phase II Study in Caucasian Patients (FIELT)', *PLOS ONE*. Edited by R. Rosell, 11(3), p. e0147599. doi: 10.1371/journal.pone.0147599.

Guan, G. *et al.* (2020) 'Utility of urinary ctDNA to monitoring minimal residual disease in early breast cancer patients.', *Cancer biomarkers : section A of Disease markers*, 28(1), pp. 111–119. doi: 10.3233/CBM-190523.

Guo, Q. mei *et al.* (2019) 'Detection of plasma EGFR mutations in NSCLC patients with a validated ddPCR lung cfDNA assay', *Journal of Cancer*, 10(18), pp. 4341–4349. doi: 10.7150/jca.31326.

Hanahan, D. and Weinberg, R. A. (2011) 'Hallmarks of cancer: the next generation.', *Cell*. Elsevier Inc., 144(5), pp. 646–74. doi: 10.1016/j.cell.2011.02.013.

Harada, T. *et al.* (2011) 'Characterization of epidermal growth factor receptor mutations in non-small-cell lung cancer patients of African-American ancestry', *Oncogene*, 30(15), pp. 1744–1752. doi: 10.1038/onc.2010.545.

He, M. *et al.* (2012) 'EGFR Exon 19 Insertions: A New Family of Sensitizing EGFR Mutations in Lung Adenocarcinoma', *Clinical Cancer Research*, 18(6), pp. 1790–1797. doi:

10.1158/1078-0432.CCR-11-2361.

Hey, S. P. *et al.* (2020) 'The evidence landscape in precision medicine', *Science Translational Medicine*, 12(540), pp. 1–5. doi: 10.1126/scitranslmed.aaw7745.

Hindson, B. J. *et al.* (2011) 'High-throughput droplet digital PCR system for absolute quantitation of DNA copy number', *Analytical Chemistry*, 83(22), pp. 8604–8610. doi: 10.1021/ac202028g.

Hindson, C. M. *et al.* (2013) 'Absolute quantification by droplet digital PCR versus analog real-time PCR', *Nature Methods*. Nature Publishing Group, 10(10), pp. 1003–1005. doi: 10.1038/nmeth.2633.

Hortobagyi, G. N. *et al.* (2022) 'Overall Survival with Ribociclib plus Letrozole in Advanced Breast Cancer', *New England Journal of Medicine*, 386(10), pp. 942–950. doi: 10.1056/nejmoa2114663.

Hrebien, S. *et al.* (2019) 'Early ctDNA dynamics as a surrogate for progression-free survival in advanced breast cancer in the BEECH trial.', *Annals of oncology : official journal of the European Society for Medical Oncology*, 30(6), pp. 945–952. doi: 10.1093/annonc/mdz085.

Hsieh, M. H. *et al.* (2006) 'Complex mutation patterns of epidermal growth factor receptor gene associated with variable responses to gefitinib treatment in patients with non-small cell lung cancer', *Lung Cancer*, 53(3), pp. 311–322. doi: 10.1016/j.lungcan.2006.06.005.

Hyman, D. M., Taylor, B. S. and Baselga, J. (2017) 'Implementing Genome-Driven Oncology.', *Cell*, 168(4), pp. 584–599. doi: 10.1016/j.cell.2016.12.015.

Isakoff, S. J. *et al.* (2005) 'Breast cancer-associated PIK3CA mutations are oncogenic in mammary epithelial cells.', *Cancer research*, 65(23), pp. 10992–1000. doi: 10.1158/0008-5472.CAN-05-2612.

Izquierdo, E. *et al.* (2021) 'Droplet digital PCR-based detection of circulating tumor DNA from pediatric high grade and diffuse midline glioma patients.', *Neuro-oncology advances*,

3(1), p. vdab013. doi: 10.1093/noajnl/vdab013.

Jeffreys, S. A. *et al.* (2022) 'Prognostic and Predictive Value of CCND1/Cyclin D1 Amplification in Breast Cancer With a Focus on Postmenopausal Patients: A Systematic Review and Meta-Analysis.', *Frontiers in endocrinology*, 13(June), p. 895729. doi: 10.3389/fendo.2022.895729.

Jenkins, S. *et al.* (2017) 'Plasma ctDNA Analysis for Detection of the EGFR T790M Mutation in Patients with Advanced Non–Small Cell Lung Cancer', *Journal of Thoracic Oncology*. Elsevier Inc, 12(7), pp. 1061–1070. doi: 10.1016/j.jtho.2017.04.003.

Jin, X. and Mu, P. (2015) 'Targeting Breast Cancer Metastasis.', *Breast cancer : basic and clinical research*, 9(Suppl 1), pp. 23–34. doi: 10.4137/BCBCR.S25460.

Johnson, T. M. (2017) 'Perspective on Precision Medicine in Oncology.', *Pharmacotherapy*, 37(9), pp. 988–989. doi: 10.1002/phar.1975.

Karrison, T. G., Ferguson, D. J. and Meier, P. (1999) 'Dormancy of mammary carcinoma after mastectomy', *Journal of the National Cancer Institute*, 91(1), pp. 80–85. doi: 10.1093/jnci/91.1.80.

Kerr, K. M. *et al.* (2014) 'Second ESMO consensus conference on lung cancer: Pathology and molecular biomarkers for non-small-cell lung cancer', *Annals of Oncology*, 25(9), pp. 1681–1690. doi: 10.1093/annonc/mdu145.

Kim, E. Y. *et al.* (2016) 'Compound EGFR mutation is frequently detected with co-mutations of actionable genes and associated with poor clinical outcome in lung adenocarcinoma', *Cancer Biology & Therapy*, 17(3), pp. 237–245. doi: 10.1080/15384047.2016.1139235.

Kobayashi, S. *et al.* (2013) 'Compound EGFR Mutations and Response to EGFR Tyrosine Kinase Inhibitors', *Journal of Thoracic Oncology*, 8(1), pp. 118–122. doi: 10.1097/JTO.0b013e3182781e35.

Koboldt, D. C. *et al.* (2012) 'VarScan 2: Somatic mutation and copy number alteration

discovery in cancer by exome sequencing', *Genome Research*, 22(3), pp. 568–576. doi: 10.1101/gr.129684.111.

Ladanyi, M. and Pao, W. (2008) 'Lung adenocarcinoma: guiding EGFR-targeted therapy and beyond.', *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc*, 21 Suppl 2(S2), pp. S16-22. doi: 10.1038/modpathol.3801018.

De Laurentiis, M. *et al.* (2022) 'Safety and Efficacy of Ribociclib in Combination with Letrozole in Patients with HR+, HER2− Advanced Breast Cancer: Results from the Italian Subpopulation of Phase 3b CompLEEment-1 Study', *Targeted Oncology*. Springer International Publishing, (0123456789). doi: 10.1007/s11523-022-00913-x.

Leary, A. F. *et al.* (2012) 'Establishing an EGFR mutation screening service for non-small cell lung cancer - sample quality criteria and candidate histological predictors.', *European journal of cancer (Oxford, England : 1990)*, 48(1), pp. 61–7. doi: 10.1016/j.ejca.2011.09.022.

Leduc, C. *et al.* (2017) 'Clinical and molecular characteristics of non-small-cell lung cancer (NSCLC) harboring EGFR mutation: results of the nationwide French Cooperative Thoracic Intergroup (IFCT) program', *Annals of Oncology*, 28(11), pp. 2715–2724. doi: 10.1093/annonc/mdx404.

Li, C. *et al.* (2018) 'EGFR T790M detection and osimertinib treatment response evaluation by liquid biopsy in lung adenocarcinoma patients with acquired resistance to first generation EGFR tyrosine kinase inhibitors', *Diagnostic Pathology*. Diagnostic Pathology, 13(1), pp. 1–6. doi: 10.1186/s13000-018-0728-6.

Li, H. and Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows-Wheeler transform', *Bioinformatics*, 25(14), pp. 1754–1760. doi: 10.1093/bioinformatics/btp324.

Li, J. *et al.* (2020) 'Non-Invasive Biomarkers for Early Detection of Breast Cancer.', *Cancers*, 12(10), pp. 1–28. doi: 10.3390/cancers12102767.

Li, X. and Warner, J. L. (2020) 'A Review of Precision Oncology Knowledgebases for

Determining the Clinical Actionability of Genetic Variants.', *Frontiers in cell and developmental biology*, 8(February), p. 48. doi: 10.3389/fcell.2020.00048.

Liang, H. *et al.* (2018) 'The alteration of T790M between 19 del and L858R in NSCLC in the course of EGFR-TKIs therapy: a literature-based pooled analysis', *Journal of Thoracic Disease*, 10(4), pp. 2311–2320. doi: 10.21037/jtd.2018.03.150.

Lindeman, N. I. *et al.* (2018) 'Updated Molecular Testing Guideline for the Selection of Lung Cancer Patients for Treatment With Targeted Tyrosine Kinase Inhibitors: Guideline From the College of American Pathologists, the International Association for the Study of Lung Cancer, and the ', *Archives of pathology & laboratory medicine*, 142(3), pp. 321–346. doi: 10.5858/arpa.2017-0388-CP.

Liu, Y. *et al.* (2013) 'Screening for EGFR and KRAS mutations in non-small cell lung carcinomas using DNA extraction by hydrothermal pressure coupled with PCR-based direct sequencing.', *International journal of clinical and experimental pathology*, 6(9), pp. 1880–9.

Lu, W. L. *et al.* (2009) 'Impact on survival of early detection of isolated breast recurrences after the primary treatment for breast cancer: A meta-analysis', *Breast Cancer Research and Treatment*, 114(3), pp. 403–412. doi: 10.1007/s10549-008-0023-4.

Lynch, T. J. *et al.* (2004) 'Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non-Small-Cell Lung Cancer to Gefitinib', *New England Journal of Medicine*, 350(21), pp. 2129–2139. doi: 10.1056/NEJMoa040938.

Mardis, E. R. (2008) 'Next-generation DNA sequencing methods', *Annual Review of Genomics and Human Genetics*, 9, pp. 387–402. doi: 10.1146/annurev.genom.9.081307.164359.

Margulies, M. *et al.* (2005) 'Genome sequencing in microfabricated high-density picolitre reactors', *Nature*, 437(7057), pp. 376–380. doi: 10.1038/nature03959.

Martínez-Sáez, O. *et al.* (2020) 'Frequency and spectrum of PIK3CA somatic mutations in

breast cancer.', *Breast cancer research : BCR*. Breast Cancer Research, 22(1), p. 45. doi: 10.1186/s13058-020-01284-9.

Masters, G. A. *et al.* (2015) 'Systemic Therapy for Stage IV Non-Small-Cell Lung Cancer: American Society of Clinical Oncology Clinical Practice Guideline Update', *Journal of Clinical Oncology*, 33(30), p. JCO.2015.62.1342-. doi: 10.1200/JCO.2015.62.1342.

Matsuo, N. *et al.* (2016) 'Association of EGFR Exon 19 Deletion and EGFR-TKI Treatment Duration with Frequency of T790M Mutation in EGFR-Mutant Lung Cancer Patients', *Scientific Reports*. Nature Publishing Group, 6(October), pp. 1–6. doi: 10.1038/srep36458.

De Mattos-Arruda, L. *et al.* (2014) 'Capturing intra-tumor genetic heterogeneity by de novo mutation profiling of circulating cell-free tumor DNA: a proof-of-principle.', *Annals of oncology : official journal of the European Society for Medical Oncology*, 25(9), pp. 1729–1735. doi: 10.1093/annonc/mdu239.

McCombie, W. R., McPherson, J. D. and Mardis, E. R. (2019) 'Next-Generation Sequencing Technologies.', *Cold Spring Harbor perspectives in medicine*, 9(11). doi: 10.1101/cshperspect.a036798.

Merker, J. D. *et al.* (2018) 'Circulating tumor DNA analysis in patients with cancer: American society of clinical oncology and college of American pathologists joint review', *Journal of Clinical Oncology*, 36(16), pp. 1631–1641. doi: 10.1200/JCO.2017.76.8671.

Metzker, M. L. (2010) 'Sequencing technologies the next generation', *Nature Reviews Genetics*. Nature Publishing Group, 11(1), pp. 31–46. doi: 10.1038/nrg2626.

Mok, T. *et al.* (2015) 'Detection and dynamic changes of EGFR mutations from circulating tumor DNA as a predictor of survival outcomes in NSCLC Patients treated with first-line intercalated erlotinib and chemotherapy', *Clinical Cancer Research*, 21(14), pp. 3196–3203. doi: 10.1158/1078-0432.CCR-14-2594.

Mok, T. S. *et al.* (2009) 'Gefitinib or Carboplatin–Paclitaxel in Pulmonary Adenocarcinoma', *New England Journal of Medicine*, 361(10), pp. 947–957. doi: 10.1056/NEJMoa0810699.

Morgillo, F. *et al.* (2016) 'Mechanisms of resistance to EGFR-targeted drugs: Lung cancer', *ESMO Open*, 1(3), pp. 1–9. doi: 10.1136/esmoopen-2016-000060.

Mosele, F. *et al.* (2020) 'Recommendations for the use of next-generation sequencing (NGS) for patients with metastatic cancers: a report from the ESMO Precision Medicine Working Group', *Annals of Oncology*. Elsevier Ltd., 31(11), pp. 1491–1505. doi: 10.1016/j.annonc.2020.07.014.

Murciano-Goroff, Y. R. *et al.* (2020) 'Toward a More Precise Future for Oncology', *Cancer Cell*. Elsevier Inc., 37(4), pp. 431–442. doi: 10.1016/j.ccell.2020.03.014.

Murtaza, M. *et al.* (2013) 'Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA', *Nature*. Nature Publishing Group, 497(7447), pp. 108–112. doi: 10.1038/nature12065.

Murtuza, A. *et al.* (2019) 'Novel third-generation EGFR tyrosine kinase inhibitors and strategies to overcome therapeutic resistance in lung cancer', *Cancer Research*, 79(4), pp. 689–698. doi: 10.1158/0008-5472.CAN-18-1281.

Naidoo, J. *et al.* (2015) 'Epidermal growth factor receptor exon 20 insertions in advanced lung adenocarcinomas: Clinical outcomes and response to erlotinib', *Cancer*. NIH Public Access, 121(18), pp. 3212–3220. doi: 10.1002/cncr.29493.

Narzisi, G. *et al.* (2014) 'Accurate de novo and transmitted indel detection in exome-capture data using microassembly', *Nature Methods*, 11(10), pp. 1033–1036. doi: 10.1038/nmeth.3069.

Narzisi, G. *et al.* (2018) 'Genome-wide somatic variant calling using localized colored de Bruijn graphs', *Communications Biology*. Springer US, 1(1), p. 20. doi: 10.1038/s42003-018-0023-9.

Nguyen, B. *et al.* (2022) 'Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients.', *Cell*. Elsevier Inc., 185(3), pp. 563-575.e11. doi: 10.1016/j.cell.2022.01.003.

Nosaki, K. *et al.* (2016) 'Re-biopsy status among non-small cell lung cancer patients in Japan: A retrospective study', *Lung Cancer*. Elsevier Ireland Ltd, 101, pp. 1–8. doi: 10.1016/j.lungcan.2016.07.007.

O'Leary, B. *et al.* (2018) 'Early circulating tumor DNA dynamics and clonal selection with palbociclib and fulvestrant for breast cancer', *Nature Communications*. Springer US, 9(1), pp. 1–10. doi: 10.1038/s41467-018-03215-x.

O'Rourke, D. *et al.* (2021) 'Fit-for-purpose quantitative liquid biopsy based droplet digital PCR assay development for detection of programmed cell death ligand-1 (PD-L1) RNA expression in PAXgene blood samples', *PLoS ONE*, 16(5 May 2021), pp. 1–14. doi: 10.1371/journal.pone.0250849.

Paez, J. G. *et al.* (2004) 'EGFR mutations in lung, cancer: Correlation with clinical response to gefitinib therapy', *Science*, 304(5676), pp. 1497–1500. doi: 10.1126/science.1099314.

Pantel, K. and Alix-Panabières, C. (2010) 'Circulating tumour cells in cancer patients: Challenges and perspectives', *Trends in Molecular Medicine*, 16(9), pp. 398–406. doi: 10.1016/j.molmed.2010.07.001.

Pao, W. *et al.* (2004) 'EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib', *Proceedings of the National Academy of Sciences*, 101(36), pp. 13306–13311. doi: 10.1073/pnas.0405220101.

Pareja, F. *et al.* (2019) 'The Genomic Landscape of Mucinous Breast Cancer', *JNCI: Journal of the National Cancer Institute*, 111(7), pp. 737–741. doi: 10.1093/jnci/djy216.

Park, I. K. *et al.* (2018) 'The prognostic effect of the epidermal growth factor receptor gene mutation on recurrence dynamics of lung adenocarcinoma', *European Journal of Cardio-Thoracic Surgery*, 0(June), pp. 1–6. doi: 10.1093/ejcts/ezy220.

Penzel, R. *et al.* (2011) 'EGFR mutation detection in NSCLC-assessment of diagnostic application and recommendations of the German Panel for Mutation Testing in NSCLC',

Virchows Archiv, 458(1), pp. 95–98. doi: 10.1007/s00428-010-1000-y.

Pereira, B. *et al.* (2016) 'The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes', *Nature Communications*, 7(May). doi: 10.1038/ncomms11479.

Pikor, L. A. *et al.* (2013) 'Genetic alterations defining NSCLC subtypes and their therapeutic implications', *Lung Cancer*, 82(2), pp. 179–189. doi: 10.1016/j.lungcan.2013.07.025.

Planchard, D. *et al.* (2018) 'Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up', *Annals of Oncology*, 29(Supplement 4), pp. iv192–iv237. doi: 10.1093/annonc/mdy275.

Poh, J. *et al.* (2022) 'Analytical and clinical validation of an amplicon-based next generation sequencing assay for ultrasensitive detection of circulating tumor DNA', *PLoS ONE*, 17(4 April), pp. 1–25. doi: 10.1371/journal.pone.0267389.

Prasad, V., Fojo, T. and Brada, M. (2016) 'Precision oncology: Origins, optimism, and potential', *The Lancet Oncology*. Elsevier Ltd, 17(2), pp. e81–e86. doi: 10.1016/S1470-2045(15)00620-8.

Preusser, M. *et al.* (2015) 'Spectrum of gene mutations detected by next generation exome sequencing in brain metastases of lung adenocarcinoma', *European Journal of Cancer*. Elsevier Ltd, 51(13), pp. 1803–1811. doi: 10.1016/j.ejca.2015.06.107.

Pugh, T. J. *et al.* (2007) 'Correlations of EGFR mutations and increases in EGFR and HER2 copy number to gefitinib response in a retrospective analysis of lung cancer patients.', *BMC cancer*. BioMed Central, 7, p. 128. doi: 10.1186/1471-2407-7-128.

Qin, D. (2019) 'Next-generation sequencing and its clinical application', *Cancer Biology and Medicine*, 16(1), pp. 4–10. doi: 10.20892/j.issn.2095-3941.2018.0055.

Quan, P. L., Sauzade, M. and Brouzes, E. (2018) 'DPCR: A technology review', *Sensors (Switzerland)*, 18(4). doi: 10.3390/s18041271.

Rakha, E. A. and Pareja, F. G. (2021) 'New Advances in Molecular Breast Cancer Pathology', *Seminars in Cancer Biology*. Elsevier, 72(March), pp. 102–113. doi: 10.1016/j.semcancer.2020.03.014.

Remon, J. *et al.* (2017) 'Osimertinib benefit in EGFR-mutant NSCLC patients with T790M-mutation detected by circulating tumour DNA', *Annals of Oncology*. Elsevier Masson SAS, 28(4), pp. 784–790. doi: 10.1093/annonc/mdx017.

Riggio, A. I., Varley, K. E. and Welm, A. L. (2021) 'The lingering mysteries of metastatic recurrence in breast cancer', *British Journal of Cancer*. Springer US, 124(1), pp. 13–26. doi: 10.1038/s41416-020-01161-4.

Rimmer, A. *et al.* (2014) 'Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications', *Nature Genetics*. Nature Publishing Group, 46(8), pp. 912–918. doi: 10.1038/ng.3036.

Rinnerthaler, G., Gampenrieder, S. P. and Greil, R. (2018) 'ASCO 2018 highlights: metastatic breast cancer.', *Memo*, 11(4), pp. 276–279. doi: 10.1007/s12254-018-0450-9.

Robichaux, J. P. *et al.* (2018) 'Mechanisms and clinical activity of an EGFR and HER2 exon 20-selective kinase inhibitor in non-small cell lung cancer', *Nature Medicine*, 24(5), pp. 638–646. doi: 10.1038/s41591-018-0007-9.

Rolfo, C. *et al.* (2020) 'Challenges and opportunities of cfDNA analysis implementation in clinical practice: Perspective of the International Society of Liquid Biopsy (ISLB)', *Critical Reviews in Oncology/Hematology*. Elsevier, 151(April), p. 102978. doi: 10.1016/j.critrevonc.2020.102978.

Rose Brannon, A. *et al.* (2021) 'Enhanced specificity of clinical high-sensitivity tumor mutation profiling in cell-free DNA via paired normal sequencing using MSK-ACCESS', *Nature Communications*, 12(1), pp. 1–12. doi: 10.1038/s41467-021-24109-5.

Rosell, R. *et al.* (2012) 'Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer

(EURTAC): A multicentre, open-label, randomised phase 3 trial', *The Lancet Oncology*. Elsevier Ltd, 13(3), pp. 239–246. doi: 10.1016/S1470-2045(11)70393-X.

Rowlands, V. *et al.* (2019) 'Optimisation of robust singleplex and multiplex droplet digital PCR assays for high confidence mutation detection in circulating tumour DNA', *Scientific Reports*. Springer US, 9(1), pp. 1–13. doi: 10.1038/s41598-019-49043-x.

Sacher, A. G. *et al.* (2016) 'Prospective validation of rapid plasma genotyping for the detection of EGFR and kras mutations in advanced lung cancer', *JAMA Oncology*, 2(8), pp. 1014–1022. doi: 10.1001/jamaoncol.2016.0173.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) 'DNA sequencing with chain-terminating inhibitors.', *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), pp. 5463–7. doi: 10.1073/pnas.74.12.5463.

Saunders, C. T. *et al.* (2012) 'Strelka: accurate somatic small-variant calling from sequenced tumor–normal sample pairs', *Bioinformatics*, 28(14), pp. 1811–1817. doi: 10.1093/bioinformatics/bts271.

Schwartzberg, L. *et al.* (2017) 'Precision Oncology: Who, How, What, When, and When Not?', *American Society of Clinical Oncology educational book. American Society of Clinical Oncology. Annual Meeting*, 37, pp. 160–169. doi: 10.1200/EDBK_174176.

Sequist, L. V. *et al.* (2007) 'Response to Treatment and Survival of Patients with Non-Small Cell Lung Cancer Undergoing Somatic EGFR Mutation Testing', *The Oncologist*, 12(1), pp. 90–98. doi: 10.1634/theoncologist.12-1-90.

Sequist, L. V. *et al.* (2011) 'Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors', *Science Translational Medicine*, 3(75). doi: 10.1126/scitranslmed.3002003.

Seyfried, T. N. and Huysentruyt, L. C. (2013) 'On the origin of cancer metastasis.', *Critical reviews in oncogenesis*, 18(1–2), pp. 43–73. doi: 10.1615/critrevoncog.v18.i1-2.40.

Shahi, R. B. *et al.* (2015) 'Detection of EGFR-TK domain-activating mutations in NSCLC with generic PCR-based methods', *Applied Immunohistochemistry and Molecular Morphology*, 23(3), pp. 163–171. doi: 10.1097/PDM.0000000000000035.

Sharma, S. V. *et al.* (2007) 'Epidermal growth factor receptor mutations in lung cancer', *Nature Reviews Cancer*, 7(3), pp. 169–181. doi: 10.1038/nrc2088.

Shaw, A. T. *et al.* (2009) 'Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK', *Journal of Clinical Oncology*, 27(26), pp. 4247–4253. doi: 10.1200/JCO.2009.22.6993.

Shen, R. and Seshan, V. E. (2016) 'FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing', *Nucleic Acids Research*, 44(16), pp. e131–e131. doi: 10.1093/nar/gkw520.

Sherry, S. T. *et al.* (2001) 'dbSNP: the NCBI database of genetic variation.', *Nucleic acids research*, 29(1), pp. 308–11. doi: 10.1093/nar/29.1.308.

Shigematsu, H. *et al.* (2005) 'Clinical and Biological Features Associated With Epidermal Growth Factor Receptor Gene Mutations in Lung Cancers', *JNCI Journal of the National Cancer Institute*. Oxford University Press, 97(5), pp. 339–346. doi: 10.1093/jnci/dji055.

Sikkema-Raddatz, B. *et al.* (2013) 'Targeted Next-Generation Sequencing can Replace Sanger Sequencing in Clinical Diagnostics', *Human Mutation*, 34(7), pp. 1035–1042. doi: 10.1002/humu.22332.

Siravegna, G. *et al.* (2019) 'How liquid biopsies can change clinical practice in oncology', *Annals of Oncology*. Elsevier Masson SAS, 30(10), pp. 1580–1590. doi: 10.1093/annonc/mdz227.

Siravegna, G. and Bardelli, A. (2014) 'Genotyping cell-free tumor DNA in the blood to detect residual disease and drug resistance', *Genome Biology*, 15(8), pp. 4–9. doi: 10.1186/s13059-014-0449-4.

Soria, J. C. *et al.* (2018) 'Osimertinib in untreated EGFR-Mutated advanced non-small-cell lung cancer', *New England Journal of Medicine*, 378(2), pp. 113–125. doi: 10.1056/NEJMoa1713137.

Sosa, M. S., Bragado, P. and Aguirre-Ghiso, J. A. (2014) 'Mechanisms of disseminated cancer cell dormancy: An awakening field', *Nature Reviews Cancer*. Nature Publishing Group, 14(9), pp. 611–622. doi: 10.1038/nrc3793.

Stewart, E. L. *et al.* (2015) 'Known and putative mechanisms of resistance to EGFR targeted therapies in NSCLC patients with EGFR mutations-a review', *Translational Lung Cancer Research*, 4(1), pp. 67–81. doi: 10.3978/j.issn.2218-6751.2014.11.06.

Taylor, S. C., Laperriere, G. and Germain, H. (2017) 'Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: From variable nonsense to publication quality data', *Scientific Reports*. Springer US, 7(1), pp. 1–8. doi: 10.1038/s41598-017-02217-x.

Thorat, M. A. (2019) 'Liquid biopsy for cancer diagnosis and screening - The promise and challenges.', *Annals of clinical biochemistry*, 56(4), pp. 420–423. doi: 10.1177/0004563219837592.

Thress, K. S. *et al.* (2015) 'Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M.', *Nature medicine*, 21(6), pp. 560–2. doi: 10.1038/nm.3854.

Topacio, B. R. *et al.* (2019) 'Cyclin D-Cdk4,6 Drives Cell-Cycle Progression via the Retinoblastoma Protein's C-Terminal Helix.', *Molecular cell*, 74(4), pp. 758-770.e4. doi: 10.1016/j.molcel.2019.03.020.

Turner, N. C. *et al.* (2020) 'Circulating tumour DNA analysis to direct therapy in advanced breast cancer (plasmaMATCH): a multicentre, multicohort, phase 2a, platform trial', *The Lancet Oncology*. The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license, 21(10), pp. 1296–1308. doi: 10.1016/S1470-2045(20)30444-7.

Uhr, J. W. and Pantel, K. (2011) 'Controversies in clinical cancer dormancy', *Proceedings of the National Academy of Sciences of the United States of America*, 108(30), pp. 12396–12400. doi: 10.1073/pnas.1106613108.

Wang, R. *et al.* (2019) 'The Clinicopathological features and survival outcomes of patients with different metastatic sites in stage IV breast cancer', *BMC Cancer*. BMC Cancer, 19(1), pp. 1–12. doi: 10.1186/s12885-019-6311-z.

Watanabe, S. *et al.* (2014) 'Effectiveness of Gefitinib against Non–Small-Cell Lung Cancer with the Uncommon EGFR Mutations G719X and L861Q', *Journal of Thoracic Oncology*, 9(2), pp. 189–194. doi: 10.1097/JTO.0000000000000048.

Weber, F. *et al.* (2005) 'Variability in organ-specific EGFR mutational spectra in tumour epithelium and stroma may be the biological basis for differential responses to tyrosine kinase inhibitors', *British Journal of Cancer*, 92(10), pp. 1922–1926. doi: 10.1038/sj.bjc.6602557.

Weigelt, B. *et al.* (2018) 'The Landscape of Somatic Genetic Alterations in Breast Cancers From ATM Germline Mutation Carriers', *JNCI: Journal of the National Cancer Institute*, 110(9), pp. 1030–1034. doi: 10.1093/jnci/djy028.

Woodhouse, R. *et al.* (2020) 'Clinical and analytical validation of foundation one liquid CDx, a novel 324-Gene cfDNA-based comprehensive genomic profiling assay for cancers of solid tumor origin', *PLoS ONE*, 15(9 September), pp. 1–18. doi: 10.1371/journal.pone.0237802.

Wu, J.-Y. *et al.* (2011) 'Effectiveness of Tyrosine Kinase Inhibitors on "Uncommon" Epidermal Growth Factor Receptor Mutations of Unknown Clinical Significance in Non-Small Cell Lung Cancer', *Clinical Cancer Research*, 17(11), pp. 3812–3821. doi: 10.1158/1078-0432.CCR-10-3408.

Wu, Y. L. *et al.* (2017) 'Dacomitinib versus gefitinib as first-line treatment for patients with EGFR-mutation-positive non-small-cell lung cancer (ARCHER 1050): a randomised, open-label, phase 3 trial', *The Lancet Oncology*. Elsevier Ltd, 18(11), pp. 1454–1466. doi:

10.1016/S1470-2045(17)30608-3.

Yan, Y. H. *et al.* (2021) 'Confirming putative variants at ≤ 5% allele frequency using allele enrichment and Sanger sequencing.', *Scientific reports*. Nature Publishing Group UK, 11(1), p. 11640. doi: 10.1038/s41598-021-91142-1.

Yang, J. C. H. *et al.* (2015) 'Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon EGFR mutations: A combined post-hoc analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6', *The Lancet Oncology*, 16(7), pp. 830–838. doi: 10.1016/S1470-2045(15)00026-1.

Yang, S. H. *et al.* (2005) 'Mutations in the Tyrosine Kinase Domain of the Epidermal Growth Factor Receptor in Non-Small Cell Lung Cancer', *Clinical Cancer Research*, 11(6), pp. 2106–2110. doi: 10.1158/1078-0432.CCR-04-1853.

Yang, T. Y. *et al.* (2011) 'Good response to gefitinib in a lung adenocarcinoma harboring a heterozygous complex mutation of L833V and H835L in epidermal growth factor receptor gene', *Journal of Clinical Oncology*, 29(16), pp. 468–469. doi: 10.1200/JCO.2010.33.5802.

Yasuda, H. *et al.* (2013) 'Structural, Biochemical, and Clinical Characterization of Epidermal Growth Factor Receptor (EGFR) Exon 20 Insertion Mutations in Lung Cancer', *Science Translational Medicine*, 5(216), pp. 216ra177-216ra177. doi: 10.1126/scitranslmed.3007205.

Yasuda, H., Kobayashi, S. and Costa, D. B. (2012) 'EGFR exon 20 insertion mutations in non-small-cell lung cancer: Preclinical data and clinical implications', *The Lancet Oncology*. Elsevier Ltd, 13(1), pp. e23–e31. doi: 10.1016/S1470-2045(11)70129-2.

Yates, L. R. *et al.* (2017) 'Genomic Evolution of Breast Cancer Metastasis and Relapse.', *Cancer cell*, 32(2), pp. 169-184.e7. doi: 10.1016/j.ccell.2017.07.005.

Yun, C. H. *et al.* (2008) 'The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP', *Proceedings of the National Academy of Sciences of the United States of America*, 105(6), pp. 2070–2075. doi: 10.1073/pnas.0709662105.

Zehir, A. *et al.* (2017) 'Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients.', *Nature medicine*, 23(6), pp. 703–713. doi: 10.1038/nm.4333.

Zhang, J., Zhang, X. and Shen, S. (2020) 'Treatment and relapse in breast cancer show significant correlations to noninvasive testing using urinary and plasma DNA.', *Future oncology (London, England)*, 16(13), pp. 849–858. doi: 10.2217/fon-2020-0074.

Zuo, Z. *et al.* (2020) 'Probing of breast cancer using a combination of plasma and urinary circulating cell-free DNA.', *Bioscience reports*, 40(11), pp. 1–10. doi: 10.1042/BSR20194306.

Zwijsen, R. M. L. *et al.* (1997) 'CDK-independent activation of estrogen receptor by cyclin D1', *Cell*, 88(3), pp. 405–415. doi: 10.1016/S0092-8674(00)81879-6.

Annexes

Annex 1

Silveira C, Sousa AC, Corredeira P, Martins M, Sousa AR, Da Cruz Paula A, Selenica P, Brown DN, Golkaram M, Kaplan S, Zhang S, Liu L, Weigelt B, Reis-Filho JS, Costa L, Carmo-Fonseca M. *Comprehensive Genomic Profiling of Cell-Free Circulating Tumor DNA Detects Response to Ribociclib Plus Letrozole in a Patient with Metastatic Breast Cancer*. Biomolecules 2022, 12,1818. PMID: 36551247. DOI: 10.3390/biom12121818.

Annex 2

Silveira C[*], Sousa AC[*], Janeiro A[*], Malveiro S, Teixeira E, Brysch E, Pantarotto M, Felizardo M, Madureira R, Nogueira F, Guimarães C, Matos C, Canário D, Bruges-Armas J, Carmo-Fonseca M. *Detection and quantification of EGFR T790M mutation in liquid biopsies by droplet digital PCR*. Transl Lung Cancer Res. 2021 Mar;10(3):1200-1208. PMID: 33889502. DOI: 10.21037/tlcr-20-1010.

Annex 3

Sousa AC[*], **Silveira C**[*], Janeiro A, Malveiro S, Oliveira AR, Felizardo M, Nogueira F, Teixeira E, Martins J, Carmo-Fonseca M. Detection of rare and novel EGFR mutations in NSCLC patients: Implications for treatment-decision. Lung Cancer. 2020 Jan;139:35-40. PMID: 31715539. DOI: 10.1016/j.lungcan.2019.10.030.

Annex 4

cobas® EGFR Mutation Test – list of mutations

Annex 5

Idylla™ EGFR Mutation Test – list of mutations

biomolecules

MDPI

Case Report

Comprehensive Genomic Profiling of Cell-Free Circulating Tumor DNA Detects Response to Ribociclib Plus Letrozole in a Patient with Metastatic Breast Cancer

Catarina Silveira [1], Ana Carla Sousa [1], Patrícia Corredeira [2], Marta Martins [2], Ana Rita Sousa [3], Arnaud Da Cruz Paula [4], Pier Selenica [4], David N. Brown [4], Mahdi Golkaram [5], Shannon Kaplan [5], Shile Zhang [5], Li Liu [5], Britta Weigelt [4], Jorge S. Reis-Filho [4], Luís Costa [2,3] and Maria Carmo-Fonseca [2,*]

1 GenoMed—Diagnósticos de Medicina Molecular, S.A., Avenida Professor Egas Moniz, 1649-028 Lisboa, Portugal
2 Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina da Universidade de Lisboa, Avenida Professor Egas Moniz, 1649-028 Lisboa, Portugal
3 Serviço de Oncologia Médica, Hospital de Santa Maria, Centro Hospitalar Lisboa Norte, Centro Académico de Medicina de Lisboa, Avenida Professor Egas Moniz, 1649-035 Lisboa, Portugal
4 Department of Pathology and Laboratory Medicine, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA
5 Illumina Inc., 5200 Illumina Way, San Diego, CA 92122, USA
* Correspondence: carmo.fonseca@medicina.ulisboa.pt

check for updates

Citation: Silveira, C.; Sousa, A.C.; Corredeira, P.; Martins, M.; Sousa, A.R.; Da Cruz Paula, A.; Selenica, P.; Brown, D.N.; Golkaram, M.; Kaplan, S.; et al. Comprehensive Genomic Profiling of Cell-Free Circulating Tumor DNA Detects Response to Ribociclib Plus Letrozole in a Patient with Metastatic Breast Cancer. *Biomolecules* **2022**, *12*, 1818. https://doi.org/10.3390/biom12121818

Academic Editors: Sridurga Mithraprabhu and Andrew Spencer

Received: 4 November 2022
Accepted: 1 December 2022
Published: 6 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: Analysis of cell-free circulating tumor DNA obtained by liquid biopsy is a non-invasive approach that may provide clinically actionable information when conventional tissue biopsy is inaccessible or infeasible. Here, we followed a patient with hormone receptor-positive and human epidermal growth factor receptor (HER) 2-negative breast cancer who developed bone metastases seven years after mastectomy. We analyzed circulating cell-free DNA (cfDNA) extracted from plasma using high-depth massively parallel sequencing targeting 468 cancer-associated genes, and we identified a clonal hotspot missense mutation in the *PIK3CA* gene (3:178952085, A > G, H1047R) and amplification of the *CCND1* gene. Whole-exome sequencing revealed that both alterations were present in the primary tumor. After treatment with ribociclib plus letrozole, the genetic abnormalities were no longer detected in cfDNA. These results underscore the clinical utility of combining liquid biopsy and comprehensive genomic profiling to monitor treatment response in patients with metastasized breast cancer.

Keywords: liquid biopsy; circulating cell-free DNA; metastatic breast cancer; ribociclib plus letrozole

1. Introduction

The development of metastases is a major cause of death in cancer patients [1]. While many tumors can be cured when detected early, once metastasis forms, most cancers become incurable [2]. In the case of breast cancer, metastasis can be found in the lungs, liver and brain, in addition to lymph nodes, but the bone is the most-affected site [3]. Survival outcomes of breast cancer patients differ depending on metastatic sites, with bone metastasis associated with the best prognosis, and brain metastasis associated with the worst survival [4].

Although metastatic disease can be present at diagnosis, most often, metastases are detected months or years following initial diagnosis and treatment. In breast cancer, metastatic recurrence has been reported ranging from months to decades after surgery [5]. Such long-term relapse of disease in a patient who was clinically asymptomatic can be attributed to cancer dormancy [6]. This phenomenon is likely caused by cancer cells that escaped from the primary tumor and disseminated throughout the body [7]. For unknown reasons, these micrometastases remain 'dormant' in secondary sites and evade anti-cancer

therapies [8]. How the body controls the proliferation of microclusters of disseminated tumor cells and why they occasionally transform into growing metastases is unclear [8].

The emergence of targeted therapies to treat metastatic breast cancer is extending life expectancy [9,10]. In parallel, the development of non-invasive biomarker assays based on liquid biopsy promises to enable the early detection of breast cancer relapse. Recent studies showed that genomic alterations identified in the primary tumor of breast cancer patients could be detected in circulating cell-free DNA (cfDNA) analyzed from plasma samples collected approximately 10 months prior to clinical or radiological relapse [11,12]. Thus, liquid biopsy may contribute to initiating treatment of metastatic disease at an earlier stage. In this regard, a meta-analysis study concluded that the earlier detection of all breast cancer recurrences would result in an absolute reduction in mortality of 17–28% [13].

Liquid biopsy tests might not only identify recurrence early, but also inform the selection of optimal treatment strategies [8]. Indeed, the clonal heterogeneity of tumor cells limits efficacy and duration of response to targeted treatments in metastatic cancer [14]. The analysis of cfDNA in plasma may be sufficient to identify somatic alterations contributed by distinct metastases, potentially circumventing the problem of lacking access to multiple metastatic tumor tissue samples due to associated risks and costs [15].

The detection of circulating cell-free tumor DNA (ctDNA) requires very sensitive molecular assays. Although PCR-based technologies, including droplet digital PCR, are powerful methods for the accurate quantification of a scarce amount of circulating nucleic acids in plasma, they can only test a few mutations per assay [16]. In contrast, recent developments in massively parallel sequencing (also known as next generation sequencing) technologies allow for the comprehensive genomic profiling of entire exonic regions of hundreds of cancer-relevant genes, identifying base substitutions, insertions or deletions, copy number alterations, and gene rearrangements [17–19].

Here, we used an established tumor-normal massively parallel sequencing assay [20,21] to characterize the genetic alterations present in the cfDNA extracted from the plasma of a breast cancer patient who developed bone metastases. We identified two DNA abnormalities that were already present in the primary tumor genomic DNA. After treatment with ribociclib plus letrozole, the patient showed a significant clinical improvement, and the two genetic alterations were no longer detected in the plasma.

2. Materials and Methods

2.1. cfDNA Extraction from Blood Samples

Within 1 to 2 h after blood collection in EDTA tubes, whole blood was centrifuged at 1600*g* for 10 min at room temperature. Then, the supernatant was transferred into falcon tubes, wasting about 5 mm of plasma to avoid buffy-coat disturbance. Next, the plasma samples were centrifuged at 3000*g* for 10 min at room temperature. This high g-force centrifugation step removes cellular debris and thereby reduces the amount of cellular or genomic DNA and RNA in the sample. After this step, the supernatants were collected into microtubes (2 or 5 mL), without disturbing the pellet containing cell debris. Plasma samples were frozen and stored at −80 °C.

For cfDNA extraction, 4 mL of plasma were thawed. The cfDNA was purified using QiAamp® MinElute® ccfDNA kit from Qiagen, according to the manufacturer's instructions. The cfDNA was eluted in 50 µL of ultra-clean water and quantified using the Qubit® 3.0 Fluorometer (Invitrogen, Life Technologies, Carlsbad, CA, USA) with Qubit® dsDNA HS Assay kit (Invitrogen, Life Technologies), according to the manufacturer's protocol. Sample quality was assessed using High Sensitivity D1000 ScreenTape (TapeStation, Agilent Technologies, Santa Clara, CA, USA), according to the manufacturer's instructions. The purified cfDNA was stored at −80 °C.

To assess quality of extracted cfDNA, samples were analyzed using a fragment analyzer (TapeStation 4200, Agilent Technologies). This assay uses a fluorescently stained double-stranded DNA and separates nucleic acids by means of electrophoresis. The TapeStation Analysis software automatically determines size, quantity, and purity of each sample.

The size determination is based on a known ladder with specific sizing standards. The known concentration of the upper marker is used to determine concentration values.

2.2. Genomic DNA Extraction from Blood Samples

Genomic DNA (gDNA) was extracted from the buffy-coat of the blood sample collected before enrollment in the clinical trial. After removal of plasma, a red blood cell lysis buffer (in-house solution) was added and incubated for 10 min at 4 °C. Then, the sample was centrifuged at 250*g* for 10 min. The supernatant was discarded and the pellet containing white blood cells was washed with PBS 1x (Sigma, St. Louis, MI, USA). After centrifugation at 250*g* for 10 min, the supernatant was discarded, and the pellet was resuspended in 1 mL of PBS 1x (Sigma) and again centrifuged at 250*g* for 10 min. The supernatant was discarded, and the dry pellet was stored at −80 °C.

DNA was extracted using a QIAmp® Blood mini kit from Qiagen, according to the manufacturer's instructions. The DNA was quantified using Qubit® 3.0 Fluorometer (Invitrogen, Life Technologies) with the Qubit® dsDNA HS Assay kit (Invitrogen, Life Technologies), according to the manufacturer's protocol.

2.3. Sequencing and Analysis of cfDNA and gDNA Extracted from Blood Samples

Both the cfDNA and matched normal gDNA were subjected to massively parallel sequencing using an established tumor-normal assay (Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets; MSK-IMPACT) that targets 468 cancer-related genes [20,21]. Sequencing data were processed and analyzed as previously reported [22–24]. Briefly, reads were aligned to the reference human genome GRCh37 using the Burrows–Wheeler Aligner (v0.7.15) [25]. Local realignment, duplicate removal, and base quality recalibration were performed using the Genome Analysis Toolkit (v3.7) [26]. Somatic single-nucleotide variants (SNVs) were detected by MuTect (v1.0) [27], and small insertions and deletions (indels) were detected using a combination of Strelka (v2.0.15) [28], VarScan2 (v2.3.7) [29], Lancet (v1.0.0) [30], Scalpel (v0.5.3) [31], and Platypus [32]. Pathogenic mutations were defined as variants that were deleterious and/or mutational hotspots. Allele-specific copy number alterations (CNAs) and loss of heterozygosity (LOH) were defined using FACETS [33], as previously described [23,34]. The fraction of the genome altered was computed from the CNAs obtained from FACETS. The cancer cell fraction of each mutation was determined using ABSOLUTE (v1.0.6) [35], as previously described [22,23,34].

2.4. Sequencing and Analysis of Primary Tumor Genomic DNA

DNA was extracted from the FFPE primary tumor sample, and Illumina DNA Prep with Enrichment was used for generating whole-exome sequencing libraries, with 40 ng input DNA, as previously described [36]. In brief, following quantification with Qubit® dsDNA High Sensitivity assay, four libraries were pooled for enrichment (4-plex) such that 500 ng of each library was used. Target enrichment was performed using IDT xGen Exome Research Panel. A single hybridization was done overnight at 58 °C, with 12 cycles of post-enrichment PCR. Libraries were quantified by Qubit® dsDNA High Sensitivity assay, normalized, and pooled. Samples were sequenced with 151 bp paired-end reads on the NovaSeq 6000 S4 flow cell using the XP workflow for individual lane loading.

Whole-exome sequencing data were processed as previously described [36]. An unpaired normal sample was used to perform variant calling. All germline variants observed in a database curated in-house, which includes the most common germline variants present in dbSNP [37], were removed. Copy number changes were estimated as previously described [36]. Tumor purity and ploidy were estimated using Sequenza 2.1, and sciClone 1.1 was used for clonality estimation [36].

3. Results

3.1. Clinical Case

The patient is a woman who was first admitted to hospital in June 2009, at the age of 34 years. She presented with a palpable mass (4 × 4 cm) in the upper outer quadrant of the right breast, with no skin alterations, and an axillary lymphadenopathy on the right side (0.5 × 1.0 cm). A diagnostic mammogram and breast ultrasound showed a hypoechoic area in the upper outer quadrant, with irregular borders and 18 mm of diameter. A microbiopsy was performed that revealed an invasive ductal carcinoma of not otherwise specified (NOS) that was estrogen receptor-positive (ER+), progesterone receptor-positive (PR+), human epidermal growth factor receptor 2-negative (HER2-). p53 was normal as detected by immunohistochemistry.

The patient started neoadjuvant chemotherapy (CTX) with doxorubicin together with cyclophosphamide. After the 5th cycle of treatment, a computed tomography (CT) scan of the abdomen and pelvis revealed a tumor in the right ovary (5 cm). In December 2009, the patient was subjected to a breast conservative surgery with axillary lymph node dissection, and a right salpingo-oophorectomy. The histological exam revealed a residual invasive ductal carcinoma NOS in multiple areas with positive margins, an axillary lymph node metastasis of the same type, and a mature cystic teratoma of the ovary (6 cm).

In January 2010, the patient underwent a mastectomy. Post-surgical treatment was adjuvant CTX with docetaxel and, subsequently, hormonal therapy with goserelin and tamoxifen. The patient also underwent adjuvant radiotherapy. Analysis of genomic DNA extracted from a blood sample revealed no pathogenic germline mutations in the *BRCA1* and *BRCA2* genes.

In September 2017, the patient presented with knee pain. A chest-abdomen-pelvis CT scan showed multiple lytic bone lesions, with soft-tissue involvement in the right iliac (Table 1). Lytic bone lesions were also detected in lumbar vertebral bodies (Table 1). The patient enrolled in an open-label clinical phase 3b trial with ribociclib combined with letrozole (CompLEEment-1, NCT02941926). A considerable clinical improvement was observed after treatment, including a decrease in pain score and partial remission of the target bone lesion in the pelvis at the CT scan (Table 1). In November 2018, the patient was treated with denosumab (120 mg at 4-week intervals). A partial response of the target and non-target lesions was observed until the 24th cycle (Table 1).

Table 1. Lesions follow-up according to the RECIST 1.1 criteria.

Target Lesion	**Lytic Bone Lesions, Right Iliac with Soft-Tissue Involvement**							
Follow-up date	19 December 2017	13 March 2018	5 June 2018	27 September 2018	4 December 2018	11 March 2019	4 June 2019	29 August 2019
Size	81 mm	58 mm	57 mm	54 mm	48 mm	47 mm	46 mm	42 mm
Non-target lesion	**Lytic bone lesions, lumbar vertebral bodies**							
Follow-up date	19 December 2017	13 March 2018	5 June 2018	27 September 2018	4 December 2018	11 March 2019	4 June 2019	29 August 2019
Number	Multiple	Multiple	Stable	Stable	Stable	Stable	Stable	Stable

3.2. Genomic Profiling of Plasma cfDNA

Blood samples were collected before (Pre-cfDNA) and after (Post-cfDNA) the patient enrolled in the clinical trial, and cfDNA was extracted as indicated in Table 2.

Table 2. Blood and plasma sample details.

Sample ID	Collection Date	Sample Type	Sample Concentration(ng/uL)	Sample Volume(μL)
Pre gDNA	18 September 2017	Buffy-coat	70.5	40.0
Pre cfDNA	18 September 2017	Plasma (4 mL)	0.7	45.0
Post cfDNA	25 June 2018	Plasma (4 mL)	0.3	45.0

A gDNA sample was additionally extracted from the buffy-coat obtained from the blood collected in September 2017. In both cfDNA samples, we detected cfDNA fragments with sizes ranging between 70 to 200 base pairs (bp), with a peak at approximately 150 bp (Figure 1).

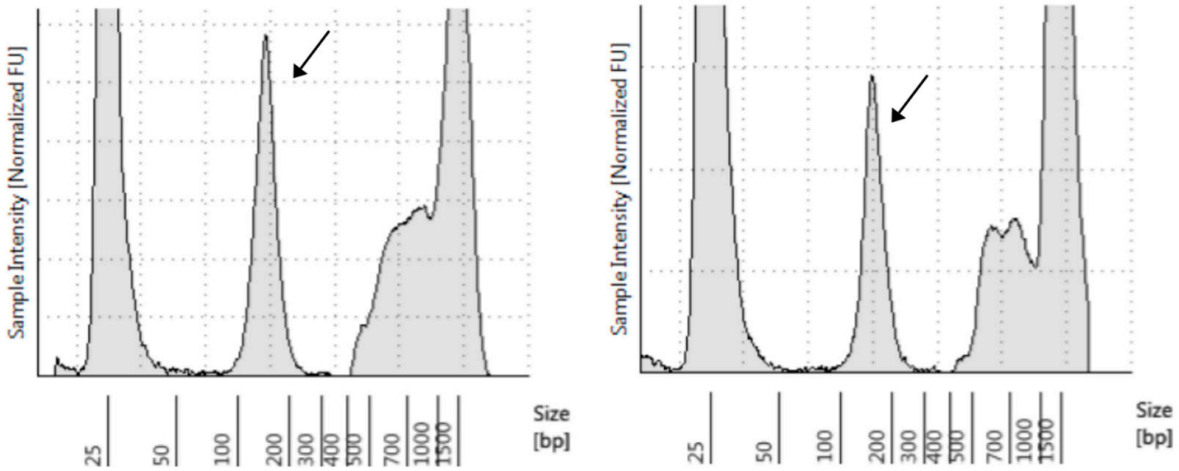


Figure 1. TapeStation analysis of DNA extracted from plasma collected before (**left**) and after (**right**) treatment with ribociclib plus letrozole. Arrows indicate cfDNA fragments (with a peak at ~150 bp).

Pre-cfDNA was subjected to massively parallel sequencing using the MSK-IMPACT assay that targets 468 cancer-related genes, detecting all protein-coding mutations, copy number alterations, and selected promoter mutations and structural rearrangements [20,21]. The sequencing panel includes oncogenes, tumor suppressor genes, and members of pathways deemed actionable by targeted therapies, and are recurrently altered in cancer. This sequencing assay has been employed for the study of >25,000 tumors [38] as well as cfDNA [24]. Two genomic alterations were detected: a missense mutation in the *PIK3CA* gene (3q26.32, Figure 2) and an amplification of the *CCND1* gene (11q13.3, Figure 3). The *PIK3CA* hotspot mutation (3:178952085, A > G, H1047R) was present at a variant allele frequency (VAF) of 0.14 (28 out of 200 reads). This variant was present in an estimated cancer cell fraction (CCF) of 0.97, indicating that the variant is likely clonal. The matched normal gDNA sample had a coverage of 199 reads at this position (3:178952085), and no altered reads were detected.

The presence of both molecular alterations was confirmed in gDNA from the primary tumor. In the primary tumor tissue, the hotspot mutation in *PIK3CA* (H1047R) was present with an estimated CCF of 1. This mutation was detected with an estimated purity of 0.34 and a ploidy of 1.8, with a normal allelic depth of 392 and a tumor allelic depth of 93 (VAF = 0.2). Additionally, amplification of the *CCND1* gene was observed, with an estimated fold change of 1.25. No other molecular changes were identified.

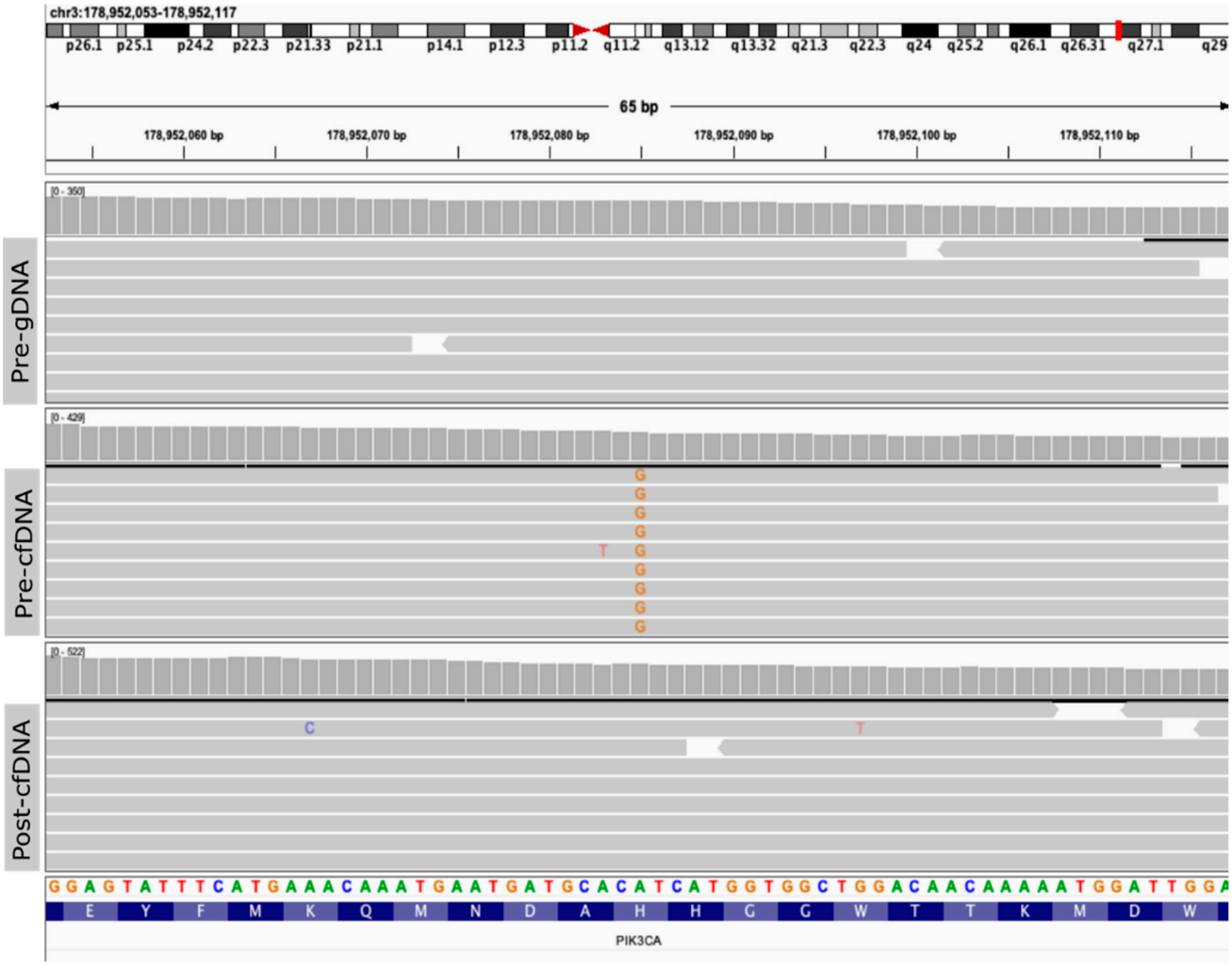


Figure 2. Integrative Genomics Viewer (IGV) screenshot of the *PIK3CA* variant detected in cell-free DNA using targeted massively parallel sequencing. The top panel depicts the sequencing reads of the genomic DNA from buffy-coat collected before treatment (Pre-gDNA). The middle and bottom panels depict the sequencing reads of the cell-free (cf)DNA collected before (Pre-cfDNA) and after (Post-cfDNA) treatment with ribociclib plus letrozole. *PIK3CA* hotspot mutation (3:178952085, A > G, H1047R) was detected only in cfDNA before treatment (middle, alternate alleles are shown in orange).

(**A**)

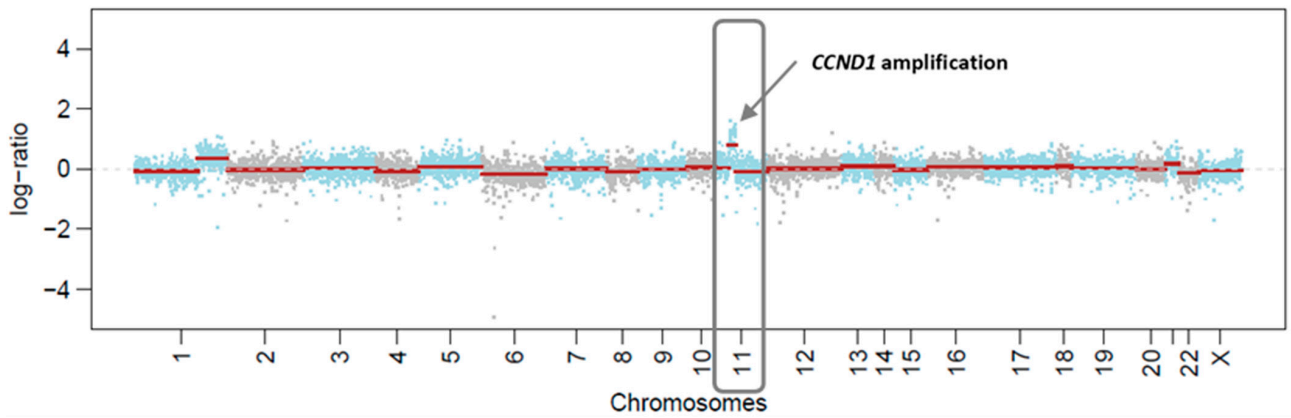


Figure 3. *Cont.*

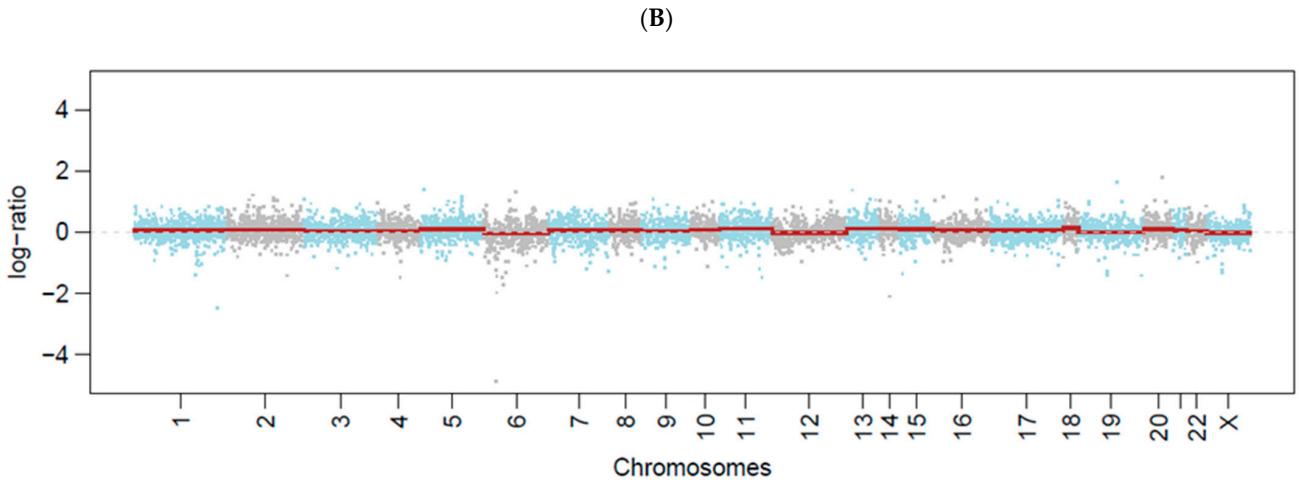


Figure 3. Copy number analysis of cell-free DNA (cfDNA) extracted from plasma collected before (**A**) and after (**B**) treatment with ribociclib plus letrozole. Amplification of the *CCND1* gene in chromosome 11 (arrow) was detected at baseline before treatment, but not after therapy.

The analysis of cfDNA after treatment (Post-cfDNA) did not detect either of the two alterations. In the *PIK3CA* gene, we identified 266 reads covering the position of interest (3:178952085), and none presented this variant (Figure 2). Moreover, amplification of the *CCND1* gene was no longer observed (Figure 3). Thus, the results in cfDNA mirror the clinical response.

4. Discussion

This study highlights the utility of cfDNA analysis for therapy monitoring in metastatic breast cancer patients. Our results are consistent with previous reports indicating that circulating tumor DNA can be used as surrogate marker of treatment outcome [15,39]. Recently, genotyping cfDNA in plasma samples from patients in the randomized phase III PALOMA-3 study of CDK4/6 inhibitor palbociclib and fulvestrant for women with advanced ER+ breast cancer showed that a reduction in the levels of mutant *PIK3CA* DNA detected in circulation correlated with improved progression-free survival (PFS) after treatment [40]. Similarly, patients with ER+ advanced metastatic breast cancer enrolled in the phase I/II randomized BEECH trial (paclitaxel plus placebo versus paclitaxel plus AKT inhibitor capivasertib) with decreased levels of mutant cfDNA detected in plasma after 4 weeks of treatment had substantially improved PFS [41].

Using massively parallel sequencing to analyze cfDNA in the patient plasma before treatment, we detected the *PIK3CA* hotspot mutation H1047R. *PIK3CA* is one of the two most frequently mutated genes in breast cancers, occurring in 30–40% of cases, and H1047R is the most common mutation in this gene [42,43]. The *PIK3CA* gene encodes the catalytic subunit of phosphatidylinositol 3-kinase (PI3K), and the H1047R mutation induces gain of enzymatic function, allowing PI3K to signal without regulation and triggering oncogenic properties [44,45]. When present, *PIK3CA* mutations are typically found in both the primary tumor and in the relapsed/metastatic tissue [46]. Consistent with the finding that *PIK3CA* mutations are predominantly truncal events in breast cancer, we identified the H1047R mutation to be clonal and likely early occurrence in tumor evolution. Notably, a previous study showed that truncal mutations in *PIK3CA* detected by liquid biopsy predicted sensitivity to palbociclib, whereas sub clonal mutations were weak predictors of outcome [40]. More recently, the sequencing of circulating tumor DNA in patients enrolled in the phase III MONALEESA-7 trial revealed a treatment response to endocrine therapy plus ribociclib independent of the *PIK3CA* mutational status [47].

In addition to the *PIK3CA* hotspot mutation, our cfDNA analysis identified the amplification of *CCND1*, an oncogene that encodes the protein cyclin D1. The cyclin dependent kinases 4 and 6 (CDK4/6) form complexes with D-type cyclins that act on the retinoblas-

toma protein Rb and drive cell cycle progression [48]. *CCND1* amplification leads to increased cyclin D1 expression and inappropriate cyclin D–CDK4/6 activity [49,50], thus promoting sustained cell proliferation, which is one of the hallmarks of cancer [51]. *CCND1* amplification occurs in 10–35% of breast cancers and is typically associated with positive ER status [49,50]. Breast cancer patients with *CCND1* amplification tend to show a poor response to endocrine therapy [50], which may be related to the ability of cyclin D1 to stimulate the growth of estrogen responsive tissues through a CDK-independent mechanism by activating the transcription of ER-regulated genes in the absence of estrogen [52]. However, the clinical benefit after ribociclib and endocrine therapy was observed in advanced ER+/HER2- breast cancer patients with altered *CCND1* [47].

The patient reported in this study was treated with the CDK4/6 inhibitor ribociclib combined with the aromatase inhibitor letrozole. After treatment, the patient had a significant clinical improvement, and no molecular abnormalities were detected by massively parallel sequencing of cfDNA. A drawback of sequencing cfDNA is the problem of false negatives. Indeed, not all cancer cells release their DNA into circulating blood, and the concentration of cell-free tumor DNA in the plasma may be below the sensitivity of available technologies. However, in this case, the cfDNA results mirrored the clinical response. A limitation of our study is that we did not monitor the patient cfDNA prospectively to determine whether detectable genetic alterations could be detected prior to clinical relapse.

Based on the results of recent trials, ribociclib plus letrozole is currently considered the frontline treatment option in postmenopausal patients with advanced ER+/PR+/HER2- breast cancer [53,54]. Although these trials showed a consistent overall survival benefit, future studies are needed to stratify drug response according to subgroups defined by patient and disease characteristics. In this regard, comprehensive profiling of cfDNA isolated from plasma samples may contribute a real-time assessment of driver and actionable mutations and their clonal evolution in response to treatment.

Author Contributions: Conception and design: C.S. and M.C.-F. Provision of study materials and clinical data from patients: A.R.S. and L.C. Technical support: A.C.S., P.C. and M.M.; Collection and assembly of sequencing data and bioinformatics analysis: A.D.C.P., P.S., D.N.B., M.G., S.K., S.Z., L.L., B.W. and J.S.R.-F.; Data analysis and interpretation: C.S. and A.D.C.P.; Manuscript writing: C.S. and M.C.-F. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Fundação para a Ciência e Tecnologia (FCT)/Ministério da Ciência, Tecnologia e Ensino Superior, Portugal, through Fundos do Orçamento de Estado to Instituto de Medicina Molecular João Lobo Antunes (LA/P/0082/2020), and FCT/FEDER/POR Lisboa 2020, Programa Operacional Regional de Lisboa, PORTUGAL 2020 (LISBOA-01-0145-FEDER-016394), and FEDER/POR Lisboa 2020-Programa Operacional Regional de Lisboa, PORTUGAL 2020 (Infogene, 045300). C.S. was a recipient of a FCT fellowship (SFRH/BDE/110544/2015). This work was funded in part by the National Institutes of Health (NIH)/National Cancer Institute (NCI) Cancer Center Support Grant (P30 CA008748; MSK). J.S.R-F. and B.W. are funded in part by the NIH/NCI P50 CA247749 01 grant and a Breast Cancer Research Foundation grant. J.S.R.-F. is also funded in part by a Susan G Komen leadership grant, and B.W. by a Cycle for Survival grant.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Lisbon Academic Medical Center (Oncodynamics, ref. n.º 343/16, 12 October 2016).

Informed Consent Statement: Informed consent was obtained from the patient involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: Authors are grateful to the patient who consented to participate in this study.

Conflicts of Interest: J.S.R.-F. reports receiving personal/consultancy fees from Goldman Sachs, REPARE Therapeutics, Paige.AI, Personalis and Bain Capital, membership of the scientific advisory boards of VolitionRx, REPARE Therapeutics, Paige.AI and Personalis, membership of the Board of Directors of Grupo Oncoclinicas, and ad hoc membership of the scientific advisory boards of Roche Tissue Diagnostics, Ventana Medical Systems, Novartis, Genentech, MSD, Daiichi Sankyo and InVicro,

outside the scope of this study. B.W. reports research support by REPARE Therapeutics, outside of the scope of this study. M.C.-F. is a founder of "GenoMed-Diagnósticos de Medicina Molecular, SA" and a member of its scientific advisory board; she reports research support by AbbVie and Gilead Sciences, outside of the scope of this study. The remaining authors have nothing to disclose.

References

1. Chaffer, C.L.; Weinberg, R.A. A perspective on cancer cell metastasis. *Science* **2011**, *331*, 1559–1564. [CrossRef] [PubMed]
2. Seyfried, T.N.; Huysentruyt, L.C. On the origin of cancer metastasis. *Crit. Rev. Oncog.* **2013**, *18*, 43–73. [CrossRef] [PubMed]
3. Jin, X.; Mu, P. Targeting Breast Cancer Metastasis. *Breast Cancer* **2015**, *9*, 23–34. [CrossRef] [PubMed]
4. Wang, R.; Zhu, Y.; Liu, X.; Liao, X.; He, J.; Niu, L. The Clinicopathological features and survival outcomes of patients with different metastatic sites in stage IV breast cancer. *BMC Cancer* **2019**, *19*, 1091. [CrossRef]
5. Karrison, T.G.; Ferguson, D.J.; Meier, P. Dormancy of mammary carcinoma after mastectomy. *J. Natl. Cancer Inst.* **1999**, *91*, 80–85. [CrossRef]
6. Uhr, J.W.; Pantel, K. Controversies in clinical cancer dormancy. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 12396–12400. [CrossRef]
7. Sosa, M.S.; Bragado, P.; Aguirre-Ghiso, J.A. Mechanisms of disseminated cancer cell dormancy: An awakening field. *Nat. Rev. Cancer* **2014**, *14*, 611–622. [CrossRef]
8. Riggio, A.I.; Varley, K.E.; Welm, A.L. The lingering mysteries of metastatic recurrence in breast cancer. *Br. J. Cancer* **2021**, *124*, 13–26. [CrossRef]
9. Rinnerthaler, G.; Gampenrieder, S.P.; Greil, R. ASCO 2018 highlights: Metastatic breast cancer. *Memo* **2018**, *11*, 276–279. [CrossRef]
10. Caswell-Jin, J.L.; Plevritis, S.K.; Tian, L.; Cadham, C.J.; Xu, C.; Stout, N.K.; Sledge, G.W.; Mandelblatt, J.S.; Kurian, A.W. Change in Survival in Metastatic Breast Cancer with Treatment Advances: Meta-Analysis and Systematic Review. *JNCI Cancer Spectr.* **2018**, *2*, pky062. [CrossRef]
11. Coombes, R.C.; Page, K.; Salari, R.; Hastings, R.K.; Armstrong, A.; Ahmed, S.; Ali, S.; Cleator, S.; Kenny, L.; Stebbing, J.; et al. Personalized detection of circulating tumor DNA antedates breast cancer metastatic recurrence. *Clin. Cancer Res.* **2019**, *25*, 4255–4263. [CrossRef] [PubMed]
12. Garcia-Murillas, I.; Chopra, N.; Comino-Méndez, I.; Beaney, M.; Tovey, H.; Cutts, R.J.; Swift, C.; Kriplani, D.; Afentakis, M.; Hrebien, S.; et al. Assessment of Molecular Relapse Detection in Early-Stage Breast Cancer. *JAMA Oncol.* **2019**, *5*, 1473–1478. [CrossRef] [PubMed]
13. Lu, W.L.; Jansen, L.; Post, W.J.; Bonnema, J.; van de Velde, J.C.; de Bock, G.H. Impact on survival of early detection of isolated breast recurrences after the primary treatment for breast cancer: A meta-analysis. *Breast Cancer Res. Treat.* **2009**, *114*, 403–412. [CrossRef] [PubMed]
14. Aparicio, S.; Caldas, C. The Implications of Clonal Genome Evolution for Cancer Medicine. *N. Engl. J. Med.* **2013**, *368*, 842–851. [CrossRef]
15. Murtaza, M.; Dawson, S.-J.; Tsui, D.W.Y.; Gale, D.; Forshew, T.; Piskorz, A.M.; Parkinson, C.; Chin, S.-F.; Kingsbury, Z.; Wong, A.S.C.; et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* **2013**, *497*, 108–112. [CrossRef]
16. Dang, D.K.; Park, B.H. Circulating tumor DNA: Current challenges for clinical utility. *J. Clin. Investig.* **2022**, *132*, e154941. [CrossRef]
17. Woodhouse, R.; Li, M.; Hughes, J.; Delfosse, D.; Skoletsky, J.; Ma, P.; Meng, W.; Dewal, N.; Milbury, C.; Clark, T.; et al. Clinical and analytical validation of foundation one liquid CDx, a novel 324-Gene cfDNA-based comprehensive genomic profiling assay for cancers of solid tumor origin. *PLoS ONE* **2020**, *15*, e0237802. [CrossRef]
18. Caputo, V.; De Falco, V.; Ventriglia, A.; Famiglietti, V.; Martinelli, E.; Morgillo, F.; Martini, G.; Corte, C.M.D.; Ciardiello, D.; Poliero, L.; et al. Comprehensive genome profiling by next generation sequencing of circulating tumor DNA in solid tumors: A single academic institution experience. *Ther. Adv. Med. Oncol.* **2022**, *14*, 17588359221096878. [CrossRef]
19. Poh, J.; Ngeow, K.C.; Pek, M.; Tan, K.-H.; Lim, J.S.; Chen, H.; Ong, C.K.; Lim, J.Q.; Lim, S.T.; Lim, C.M.; et al. Analytical and clinical validation of an amplicon-based next generation sequencing assay for ultrasensitive detection of circulating tumor DNA. *PLoS ONE* **2022**, *17*, e0267389. [CrossRef] [PubMed]
20. Cheng, D.T.; Mitchell, T.N.; Zehir, A.; Shah, R.H.; Benayed, R.; Syed, A.; Chandramohan, R.; Liu, Z.Y.; Won, H.H.; Scott, S.N.; et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT). *J. Mol. Diagn.* **2015**, *17*, 251–264. [CrossRef]
21. Zehir, A.; Benayed, R.; Shah, R.H.; Syed, A.; Middha, S.; Kim, H.R.; Srinivasan, P.; Gao, J.; Chakravarty, D.; Devlin, S.M.; et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med.* **2017**, *23*, 703–713. [CrossRef] [PubMed]
22. Weigelt, B.; Bi, R.; Kumar, R.; Blecua, P.; Mandelker, D.L.; Geyer, F.C.; Pareja, F.; James, P.A.; kConFab Investigators; Couch, F.J.; et al. The Landscape of Somatic Genetic Alterations in Breast Cancers From *ATM* Germline Mutation Carriers. *J. Natl. Cancer Inst.* **2018**, *110*, 1030–1034. [CrossRef] [PubMed]
23. Da Cruz Paula, A.; da Silva, E.M.; Segura, S.E.; Pareja, F.; Bi, R.; Selenica, P.; Kim, S.H.; Ferrando, L.; Vahdatinia, M.; Soslow, R.A.; et al. Genomic profiling of primary and recurrent adult granulosa cell tumors of the ovary. *Mod. Pathol.* **2020**, *33*, 1606–1617. [CrossRef] [PubMed]

24. De Mattos-Arruda, L.; Weigelt, B.; Cortes, J.; Won, H.H.; Ng, C.K.Y.; Nuciforo, P.; Bidard, F.-C.; Aura, C.; Saura, C.; Peg, V.; et al. Capturing intra-tumor genetic heterogeneity by de novo mutation profiling of circulating cell-free tumor DNA: A proof-of-principle. *Ann. Oncol.* **2014**, *25*, 1729–1735. [CrossRef] [PubMed]
25. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef]
26. DePristo, M.A.; Banks, E.; Poplin, R.; Garimella, K.V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; del Angel, G.; Rivas, M.A.; Hanna, M.; et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **2011**, *43*, 491–498. [CrossRef]
27. Cibulskis, K.; Lawrence, M.S.; Carter, S.L.; Sivachenko, A.; Jaffe, D.; Sougnez, C.; Gabriel, S.; Meyerson, M.; Lander, E.S.; Getz, G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **2013**, *31*, 213–219. [CrossRef]
28. Saunders, C.T.; Wong, W.S.W.; Swamy, S.; Becq, J.; Murray, L.J.; Cheetham, R.K. Strelka: Accurate somatic small-variant calling from sequenced tumor–normal sample pairs. *Bioinformatics* **2012**, *28*, 1811–1817. [CrossRef]
29. Koboldt, D.C.; Zhang, Q.; Larson, D.E.; Shen, D.; McLellan, M.D.; Lin, L.; Miller, C.A.; Mardis, E.R.; Ding, L.; Wilson, R.K. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **2012**, *22*, 568–576. [CrossRef]
30. Narzisi, G.; Corvelo, A.; Arora, K.; Bergmann, E.A.; Shah, M.; Musunuri, R.; Emde, A.-K.; Robine, N.; Vacic, V.; Zody, M.C. Genome-wide somatic variant calling using localized colored de Bruijn graphs. *Commun. Biol.* **2018**, *1*, 20. [CrossRef]
31. Narzisi, G.; O'Rawe, J.A.; Iossifov, I.; Fang, H.; Lee, Y.-h.; Wang, Z.; Wu, Y.; Lyon, G.J.; Wigler, M.; Schatz, M.C. Accurate de novo and transmitted indel detection in exome-capture data using microassembly. *Nat. Methods* **2014**, *11*, 1033–1036. [CrossRef] [PubMed]
32. Rimmer, A.; Phan, H.; Mathieson, I.; Iqbal, Z.; Twigg, S.R.F.; WGS500 Consortium; Wilkie, A.O.M.; McVean, G.; Lunter, G. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat. Genet.* **2014**, *46*, 912–918. [CrossRef] [PubMed]
33. Shen, R.; Seshan, V.E. FACETS: Allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res.* **2016**, *44*, e131. [CrossRef] [PubMed]
34. Pareja, F.; Lee, J.Y.; Brown, D.N.; Piscuoglio, S.; Gularte-Mérida, R.; Selenica, P.; Da Cruz Paula, A.; Arunachalam, S.; Kumar, R.; Geyer, F.C.; et al. The Genomic Landscape of Mucinous Breast Cancer. *J. Natl. Cancer Inst.* **2019**, *111*, 737–741. [CrossRef]
35. Carter, S.L.; Cibulskis, K.; Helman, E.; McKenna, A.; Shen, H.; Zack, T.; Laird, P.W.; Onofrio, R.C.; Winckler, W.; Weir, B.A.; et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat. Biotechnol.* **2012**, *30*, 413–421. [CrossRef]
36. Golkaram, M.; Salmans, M.L.; Kaplan, S.; Vijayaraghavan, R.; Martins, M.; Khan, N.; Garbutt, C.; Wise, A.; Yao, J.; Casimiro, S.; et al. HERVs establish a distinct molecular subtype in stage II/III colorectal cancer with poor outcome. *NPJ Genomic Med.* **2021**, *6*, 13. [CrossRef]
37. Sherry, S.T.; Ward, M.-H.; Kholodov, M.; Baker, J.; Phan, L.; Smigielski, E.M.; Sirotkin, K. dbSNP: The NCBI database of genetic variation. *Nucleic Acids Res.* **2001**, *29*, 308–311. [CrossRef]
38. Nguyen, B.; Fong, C.; Luthra, A.; Smith, S.A.; DiNatale, R.G.; Nandakumar, S.; Walch, H.; Chatila, W.K.; Madupuri, R.; Kundra, R.; et al. Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients. *Cell* **2022**, *185*, 563–575.e11. [CrossRef]
39. Dawson, S.-J.; Tsui, D.W.Y.; Murtaza, M.; Biggs, H.; Rueda, O.M.; Chin, S.-F.; Dunning, M.J.; Gale, D.; Forshew, T.; Mahler-Araujo, B.; et al. Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer. *N. Engl. J. Med.* **2013**, *368*, 1199–1209. [CrossRef]
40. O'Leary, B.; Hrebien, S.; Morden, J.P.; Beaney, M.; Fribbens, C.; Huang, X.; Liu, Y.; Bartlett, C.H.; Koehler, M.; Cristofanilli, M.; et al. Early circulating tumor DNA dynamics and clonal selection with palbociclib and fulvestrant for breast cancer. *Nat. Commun.* **2018**, *9*, 896. [CrossRef]
41. Hrebien, S.; Citi, V.; Garcia-Murillas, I.; Cutts, R.; Fenwick, K.; Kozarewa, I.; McEwen, R.; Ratnayake, J.; Maudsley, R.; Carr, T.H.; et al. Early ctDNA dynamics as a surrogate for progression-free survival in advanced breast cancer in the BEECH trial. *Ann. Oncol.* **2019**, *30*, 945–952. [CrossRef] [PubMed]
42. Pereira, B.; Chin, S.-F.; Rueda, O.M.; Vollan, H.-K.M.; Provenzano, E.; Bardwell, H.A.; Pugh, M.; Jones, L.; Russell, R.; Sammut, S.-J.; et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat. Commun.* **2016**, *7*, 11479. [CrossRef] [PubMed]
43. Martínez-Sáez, O.; Chic, N.; Pascual, T.; Adamo, B.; Vidal, M.; González-Farré, B.; Sanfeliu, E.; Schettini, F.; Conte, B.; Brasó-Maristany, F.; et al. Frequency and spectrum of *PIK3CA* somatic mutations in breast cancer. *Breast Cancer Res.* **2020**, *22*, 45. [CrossRef]
44. Isakoff, S.J.; Engelman, J.A.; Irie, H.Y.; Luo, J.; Brachmann, S.M.; Pearline, R.V.; Cantley, L.C.; Brugge, J.S. Breast cancer-associated *PIK3CA* mutations are oncogenic in mammary epithelial cells. *Cancer Res.* **2005**, *65*, 10992–11000. [CrossRef] [PubMed]
45. Bader, A.G.; Kang, S.; Vogt, P.K. Cancer-specific mutations in *PIK3CA* are oncogenic in vivo. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 1475–1479. [CrossRef] [PubMed]

46. Yates, L.R.; Knappskog, S.; Wedge, D.; Farmery, J.H.R.; Gonzalez, S.; Martincorena, I.; Alexandrov, L.B.; Van Loo, P.; Haugland, H.K.; Lilleng, P.K.; et al. Genomic Evolution of Breast Cancer Metastasis and Relapse. *Cancer Cell* **2017**, *32*, 169–184.e7. [CrossRef] [PubMed]
47. Bardia, A.; Su, F.; Solovieff, N.; Im, S.-A.; Sohn, J.; Lee, K.S.; Campos-Gomez, S.; Jung, K.H.; Colleoni, M.; Vázquez, R.V.; et al. Genomic Profiling of Premenopausal HR+ and HER2- Metastatic Breast Cancer by Circulating Tumor DNA and Association of Genetic Alterations with Therapeutic Response to Endocrine Therapy and Ribociclib. *JCO Precis. Oncol.* **2021**, *5*, 1408–1420. [CrossRef]
48. Topacio, B.R.; Zatulovskiy, E.; Cristea, S.; Xie, S.; Tambo, C.S.; Rubin, S.M.; Sage, J.; Kõivomägi, M.; Skotheim, J.M. Cyclin D-Cdk4,6 Drives Cell-Cycle Progression via the Retinoblastoma Protein's C-Terminal Helix. *Mol. Cell* **2019**, *74*, 758–770.e4. [CrossRef] [PubMed]
49. Elsheikh, S.; Green, A.R.; Aleskandarany, M.A.; Grainge, M.; Paish, C.E.; Lambros, M.B.K.; Reis-Filho, J.S.; Ellis, I.O. *CCND1* amplification and cyclin D1 expression in breast cancer and their relation with proteomic subgroups and patient outcome. *Breast Cancer Res. Treat.* **2008**, *109*, 325–335. [CrossRef]
50. Jeffreys, S.A.; Becker, T.M.; Khan, S.; Soon, P.; Neubauer, H.; de Souza, P.; Powter, B. Prognostic and Predictive Value of *CCND1*/Cyclin D1 Amplification in Breast Cancer With a Focus on Postmenopausal Patients: A Systematic Review and Meta-Analysis. *Front. Endocrinol.* **2022**, *13*, 895729. [CrossRef]
51. Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: The next generation. *Cell* **2011**, *144*, 646–674. [CrossRef] [PubMed]
52. Zwijsen, R.M.L.; Wientjens, E.; Klompmaker, R.; der Sman, J.; Bernards, R.; Michalides, R.J.A.M. CDK-independent activation of estrogen receptor by cyclin D1. *Cell* **1997**, *88*, 405–415. [CrossRef] [PubMed]
53. Hortobagyi, G.N.; Stemmer, S.M.; Burris, H.A.; Yap, Y.-S.; Sonke, G.S.; Hart, L.; Campone, M.; Petrakova, K.; Winer, E.P.; Janni, W.; et al. Overall Survival with Ribociclib plus Letrozole in Advanced Breast Cancer. *N. Engl. J. Med.* **2022**, *386*, 942–950. [CrossRef] [PubMed]
54. De Laurentiis, M.; Caputo, R.; Mazza, M.; Mansutti, M.; Masetti, R.; Ballatore, Z.; Torrisi, R.; Michelotti, A.; Zambelli, A.; Ferro, A.; et al. Safety and Efficacy of Ribociclib in Combination with Letrozole in Patients with HR+, HER2− Advanced Breast Cancer: Results from the Italian Subpopulation of Phase 3b CompLEEment-1 Study. *Target. Oncol.* **2022**, *17*, 615–625. [CrossRef] [PubMed]

Original Article

Detection and quantification of *EGFR* T790M mutation in liquid biopsies by droplet digital PCR

Catarina Silveira[1#], Ana Carla Sousa[1#], André Janeiro[1#], Sara Malveiro[1], Encarnação Teixeira[2], Eva Brysch[2], Marcos Pantarotto[3], Margarida Felizardo[4], Rosa Madureira[5], Fernando Nogueira[6], Cátia Guimarães[6], Cristina Matos[6], Dolores Canário[7], Jácome Bruges-Armas[8], Maria Carmo-Fonseca[9]

[1]GenoMed - Molecular Medicine Diagnostics, S.A., Lisbon, Portugal; [2]Hospital Pulido Valente, Department of Pulmonary Oncology, Lisbon, Portugal; [3]Hospital da Luz, Department of Medical Oncology, Lisbon, Portugal; [4]Hospital Beatriz Ângelo, Department of Pulmonology, Lisbon, Portugal; [5]Hospital Beatriz Ângelo, Department of Pathology, Lisbon, Portugal; [6]Hospital Egas Moniz, Department of Pulmonology, Lisbon, Portugal; [7]Hospital Garcia de Orta, Department of Pulmonology, Almada, Portugal; [8]Hospital de Santo Espírito, Department of Specialized Service of Epidemiology and Molecular Biology (SEEBMO), Angra do Heroísmo, Azores, Portugal; [9]Instituto de Medicina Molecular João Lobo Antunes, University of Lisbon Medical School, Lisbon, Portugal

Contributions: (I) Conception and design: C Silveira, AC Sousa, A Janeiro, M Carmo-Fonseca; (II) Administrative support: C Silveira, AC Sousa, A Janeiro; (III) Provision of study materials or patients: E Teixeira, E Brysch, M Pantarotto, M Felizardo, R Madureira, F Nogueira, C Guimarães, C Matos, D Canário, J Bruges-Armas; (IV) Collection and assembly of data: C Silveira, AC Sousa, A Janeiro; (V) Data analysis and interpretation: AC Sousa, C Silveira, A Janeiro, S Malveiro; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to the work.

Correspondence to: Maria Carmo-Fonseca. Instituto de Medicina Molecular João Lobo Antunes, Avenida Professor Egas Moniz, Edifício Egas Moniz, 1649-028 Lisboa, Portugal. Email: carmo.fonseca@medicina.ulisboa.pt.

Background: Liquid biopsy allows the identification of targetable cancer mutations in a minimally invasive manner. In patients with advanced non-small cell lung cancer (NSCLC), droplet digital PCR (ddPCR) is increasingly used to genotype the epidermal growth factor receptor (*EGFR*) gene in circulating cell-free DNA (cfDNA). However, the sensitivity of this method is still under debate. The aim of this study was to implement and assess the performance of a ddPCR assay for detecting the *EGFR* T790M mutation in liquid biopsies.

Methods: A ddPCR assay was optimized to detect the *EGFR* T790M mutation in plasma samples from 77 patients with NSCLC in progression.

Results: Our ddPCR assay enabled the detection and quantification of the *EGFR* T790M mutation at cfDNA allele frequency as low as 0.5%. The mutation was detected in 40 plasma samples, corresponding to a positivity rate of 52%. The number of mutant molecules per mL of plasma ranged from 1 to 6,000. A re-biopsy was analyzed for 12 patients that had a negative plasma test and the mutation was detected in 2 cases. A second liquid biopsy was performed for 6 patients and the mutation was detected in 3 cases.

Conclusions: This study highlights the value of ddPCR to detect and quantify the *EGFR* T790M mutation in liquid biopsies in a real-world clinical setting. Our results suggest that repeated ddPCR tests in cfDNA may obviate tissue re-biopsy in patients unable to provide a tumor tissue sample suitable for molecular analysis.

Keywords: Lung cancer; *EGFR* T790M mutation; liquid biopsy; droplet digital PCR (ddPCR)

Submitted Sep 03, 2020. Accepted for publication Dec 21, 2020.
doi: 10.21037/tlcr-20-1010
View this article at: http://dx.doi.org/10.21037/tlcr-20-1010

Introduction

Lung cancer is the most commonly diagnosed cancer and remains the leading cause of cancer death (1). A significant improvement of progression-free survival has been achieved with receptor-tyrosine kinase inhibitors (TKIs) that target the epidermal growth factor receptor (EGFR) in patients with non-small cell lung cancer (NSCLC) harboring activating *EGFR* mutations (2-6).

Binding of the EGFR extracellular domain to its ligands triggers autophosphorylation at key tyrosine residues and activates several downstream signaling pathways. Certain mutations and/or amplification of the *EGFR* gene lead to constitutive activation of EGFR signaling and play an important role as oncogenic drivers in NSCLC. The prevalence of EGFR-activating mutations in a Caucasian population with lung adenocarcinoma is approximately 10–20%, and the most common (>90%) are small in-frame deletions in exon 19 and an amino acid substitution in exon 21 (L858R) (7-9). These alterations confer sensitivity to EGFR-TKI therapy, resulting in response rates up to 70% and median survival up to 24–30 months (10).

Despite initial responses, most patients with *EGFR*-mutant NSCLC and treated with EGFR-TKIs (such as gefitinib, erlotinib, and afatinib) will have disease progression within 9–14 months after starting the treatment (8,11). The major mechanism of acquired resistance to EGFR-TKIs is the occurrence of a secondary *EGFR* kinase domain mutation in exon 20, the T790M substitution, which accounts for about half of the cases (8,12,13). This mutation leads to an enhanced affinity for ATP, thus reducing the ability of ATP-competitive reversible EGFR tyrosine kinase inhibitors, including gefitinib and erlotinib, to bind to the tyrosine kinase domain of EGFR (14).

Recently, a third generation of EGFR-TKIs was developed that irreversibly block T790M mutant *EGFR* with maintained activity against the original exon 19del and L858R mutations (15). Thus, testing for the *EGFR* T790M mutation has become routine clinical practice in patients with NSCLC that become resistant to first- and second-generation EGFR-TKIs. Ideally, detection of this new mutation should be done in tumor tissue obtained by re-biopsy (9,16). However, many patients on progression develop lesions in inaccessible locations. Moreover, the poor performance status of the patients also makes re-biopsy difficult. It is estimated that up to 40% of relapsed NSCLC patients may be unable to provide a tumor tissue sample suitable for molecular analysis (17). For these patients it is acceptable to perform a liquid biopsy, which allows genotyping cell-free tumor DNA (cfDNA) present in the plasma and other body fluids (18).

Early comparisons between tumor tissue samples and liquid biopsy for determining *EGFR* mutation status concluded that analysis of cfDNA detected fewer *EGFR* mutation positive patients (19,20). However, subsequent studies using more sensitive assays such as the Inivata InVision™ (eTAm-Seq™) assay or the cobas EGFR Mutation Test, reported detection of the T790M mutation in plasma samples from 50% and 61% of the patients with NSCLC at disease progression after previous EGFR-TKI therapy (17,21).

Droplet digital PCR (ddPCR) is emerging as a very attractive option in the clinic to genotype cfDNA in liquid biopsies (18). This is a PCR method based on water-oil emulsion droplet technology. A cfDNA sample is fractionated into 20,000 droplets, PCR amplification of both the mutated and wild-type DNA molecules occurs in each individual droplet, and fluorescent specific probes are used to quantify the amplified molecules. Whether this approach has the required rigor to be used in the clinical setting remains debatable. A prospective validation study showed that plasma ddPCR detected *EGFR* T790M mutation with a sensitivity of 77%, supporting the use of this assay to direct clinical care (22). However, in a real-world setting, the practical sensitivity of the ddPCR assay may vary. Indeed, recent studies that analyzed plasma cfDNA by ddPCR reported values for the prevalence of the T790M mutation in patients with acquired resistance to EGFR-TKIs ranging between 30.4% (23) and 42.7% (24).

Here we present an optimized ddPCR strategy that was used to test for the presence of the resistance *EGFR* T790M mutation in plasma samples from 77 patients with NSCLC in progression, resulting in a positivity rate of 52%.

We present the following article in accordance with the STARD reporting checklist (available at http://dx.doi.org/10.21037/tlcr-20-1010).

Methods

Study population

This is a retrospective study including a total of 111 patients with NSCLC in progression after treatment with EGFR-tyrosine kinase inhibitors (EGFR-TKIs), who were tested for the presence of the resistance mutation T790M in exon 20 of the *EGFR* gene. Most patients were diagnosed with lung adenocarcinoma in TNM stage IV. The average age of the patients at progression was 67, ranging from 31 to 91 (*Table 1*).

Table 1 Clinical data of NSCLC patients enrolled in the study

Clinical data	N (%)
Age[1]	
≤65	51 [46]
>65	60 [54]
Gender	
Female	77 [69]
Male	34 [31]
EGFR-TKI[2]	
Erlotinib	50 [56]
Gefitinib	27 [30]
Afatinib	13 [14]

NSCLC, non-small cell lung cancer; EGFR-TKI, EGFR-tyrosine kinase inhibitor. [1], Age at progression; [2], EGFR-TKI used in first or second-line treatment. No information was available for 21 of the 111 patients.

The EGFR-TKIs used as first- or second-line treatment are indicated in *Table 1*. All tests were performed at GenoMed, in Lisbon, Portugal, during the period from July 2015 to December 2019. The samples were obtained from the following hospitals in Portugal: Hospital Pulido Valente; Hospital de Santa Maria; Hospital Egas Moniz; Hospital dos Capuchos; Hospital Beatriz Ângelo; Hospital de Santo Espírito da Ilha Terceira; Hospital Professor Doutor Fernando Fonseca; Hospital da Luz; Hospital Garcia de Orta; CUF Descobertas; Unidade Local Saúde Norte Alentejano and Hospital de Cascais. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The Lisbon Academic Medical Center Ethics Committee approved the study (311/20). Because the study involves the secondary analysis of existing data collected during standard of care procedures, a waiver of informed consent was requested and approved. All the data were analyzed anonymously, thus protecting the privacy and personal identity information of participating individuals. Samples were collected for standard clinical care and were not used for any other purpose. The sample collection procedures used have minimal risks and no adverse events were reported. The performers of the DNA tests were informed that patients had a diagnosis of advanced NSCLC in progression after treatment with EGFR-TKIs. Additional patient-related information was provided by clinicians after DNA data analysis was completed.

DNA analysis from FFPE samples

A trained pathologist made estimations of tumor cell content in formalin-fixed paraffin-embedded (FFPE) tissue samples resected from NSCLC patients. For macrodissection, the pathologist marked tumor areas on hematoxylin and eosin (H&E) stained tissue slides. Two to five serial FFPE sections of 10 µm thickness were then sliced from the selected areas. DNA was extracted using cobas® DNA Sample Preparation Kit, quantified using the NanoDrop® ND-1000 spectrophotometer (NanoDrop® Technologies), and analyzed by real-time PCR using cobas® EGFR Mutation Test. For Idylla™ EGFR Mutation Assay, sample preparation was done according to the manufacturer's instructions.

Isolation of cfDNA from plasma

To extract cfDNA from plasma, ~10 mL blood samples were collected in EDTA or Cell-Free DNA BCT® (Streck) tubes. After centrifugation (900 g for EDTA tubes and 1,600 g for Streck tubes) for 10 minutes at room temperature, the supernatant was transferred into 2 mL microtubes, wasting about 5 mm of plasma in order to avoid buffy-coat disturbance. The samples were then centrifuged at 16,000 g for 10 additional minutes at room temperature to remove cell debris, thus reducing contamination with cellular DNA. The supernatants were collected into microtubes (2 or 5 mL), frozen, and stored at –80 ℃. The cfDNA was isolated from 2 mL thawed plasma samples using QIAamp® MinElute® ccfDNA (QIAGEN), according to the manufacturer's instructions. The cfDNA was eluted in 50 µL of ultra-clean water and quantified using the Qubit® 3.0 Fluorometer (Invitrogen, Life Technologies) with Qubit® dsDNA HS Assay kit (Invitrogen, Life Technologies). The purified cfDNA was stored at 4 ℃ for up to 16 hours or at –80 ℃ for longer periods.

T790M screening by droplet digital PCR (ddPCR)

To screen for the T790M mutation in plasma cfDNA, we used the QX200™ Droplet Digital PCR System (Bio-Rad, Hercules, CA, USA), and Bio-Rad assay primer/probe mixtures (dHsaCP2000019 and dHsaCP2000020). As reference for wild type and mutant *EGFR*, we used Horizon DNA standards HD709 and HD258 (Horizon Discovery Ltd., Cambridge, UK). The optimal annealing temperature was established at 59 ℃. Altogether, the thermocycling conditions included an initial incubation at 95℃ for 10 minutes, 47 cycles of 94 ℃ for 30 seconds and 59 ℃ for 1 minute, followed by an inactivation step at 98 ℃ for 10 minutes. At least two replicates were analyzed per

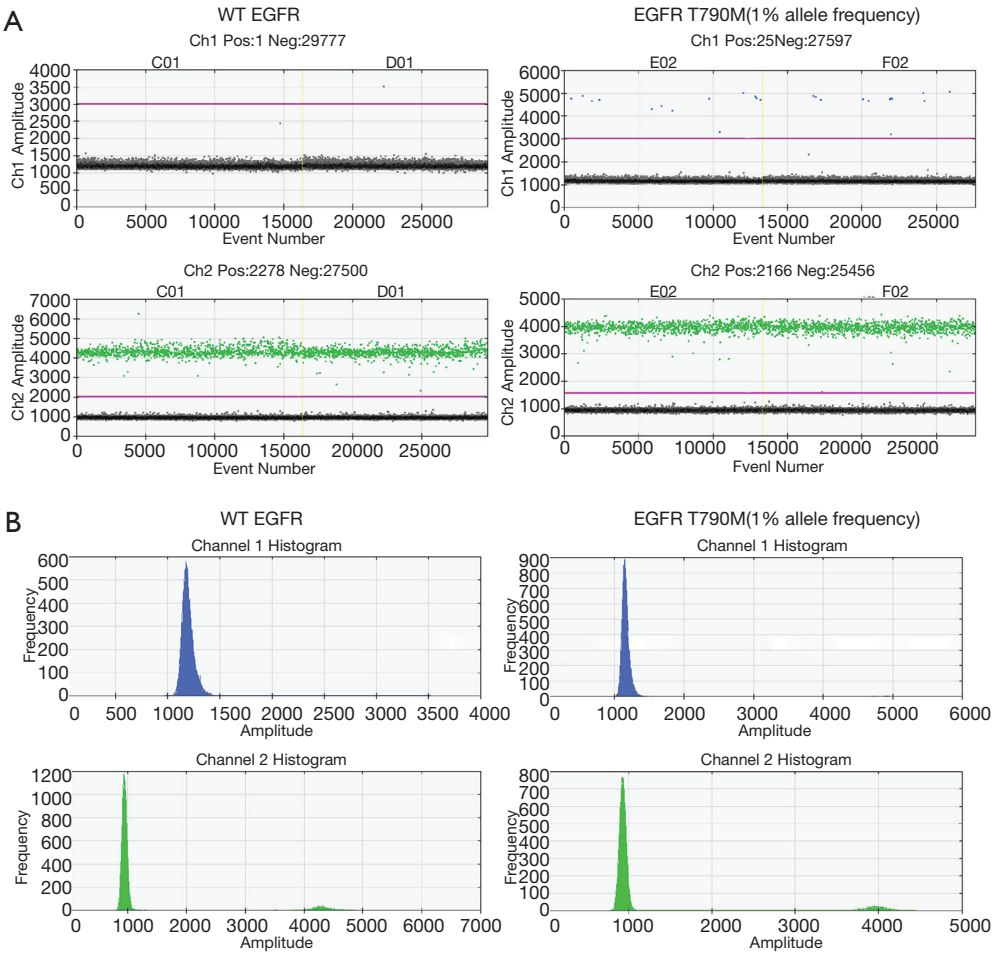


Figure 1 Visualization of two-channel ddPCR data. Either wild type (WT) reference DNA only (left) or a mix of WT and mutant reference DNA (right) were analyzed. Scatter plots (A) and histograms (B) are shown. Channel 2 (green) depicts the signal corresponding to WT *EGFR*. Channel 1 (blue) depicts the signal corresponding to mutant *EGFR*. In scatter plots, positive droplets are depicted in blue and green, negative droplets are depicted in grey, and the threshold line is depicted in pink.

sample, and the total amount of DNA loaded (considering the two replicates), was approximately 15 ng. No template control (NTC) was used to exclude PCR contamination. Amplification results were analyzed using QX200™ Droplet Reader and QuantaSoft™ software. Samples with three or more positive mutant droplets were considered positive, as recommended by the best practice guidelines for rare mutation detection (25). If one or two droplets were observed, the result was considered inconclusive and whenever possible a second sample was collected and tested.

Statistical analysis

The concentrations of target alleles were calculated using QuantaSoft™ software (Bio-Rad) based on Poisson distribution. Mutant cfDNA molecules were reported as number of copies per milliliter (mL) of plasma. The mutant allelic frequency was determined as the ratio of mutant droplets relative to the sum of mutant and wild type droplets.

Results

Quality assessment

DNA reference standards for wild type and mutant *EGFR* confirmed the specificity of the ddPCR primers and probes (*Figure 1*). The wild type DNA exhibited high fluorescence signal in the hexachloro-6-carboxy-fluorescine (HEX) channel (channel 2, green) and no fluorescence signal in the fluorescein

Table 2 Assessing the ddPCR assay with liquid biopsy reference samples

Sample ID	Pre-defined EGFR genotype	Pre-defined mutant allelic frequency	Amount of cfDNA obtained per sample (ng/μL)	Result of ddPCR assay	Mutant allelic frequency estimated by ddPCR
QA-01	T790M	0.05%	7.82	Not detected	–
QA-02	T790M	0.5%	6.59	Detected	0.6%
QA-03	T790M	5%	6.08	Detected	4.6%
QA-04	L858R	0.05%	7.42	Not detected	–
QA-05	L858R	0.5%	6.75	Not detected	–
QA-06	L858R	5%	6.60	Not detected	–
QA-07	ΔE746-A750	0.05%	6.86	Not detected	–
QA-08	ΔE746-A750	0.5%	6.74	Not detected	–
QA-09	ΔE746-A750	5%	5.82	Not detected	–
QA-10	Wild type	0%	4.81	Not detected	–
QA-11	T790M/L858R	5%	6.44	Detected	5.7%
QA-12	T790M/L858R	0.5%	6.35	Detected	0.3%
QA-13	T790M/L858R	0.05%	7.42	Not detected	–

cfDNA, circulating cell-free DNA; ddPCR, droplet digital PCR; QA, quality assessment.

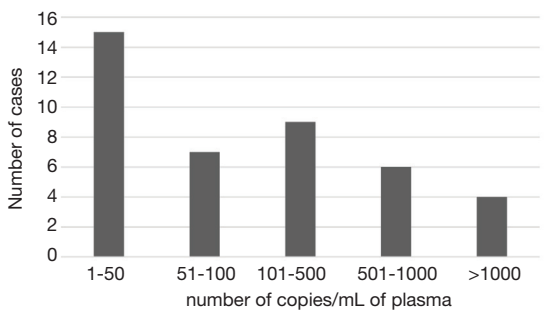


Figure 2 Distribution of the number of T790M mutant molecules in 1 mL of plasma from 77 patients tested by ddPCR.

amidite (FAM) channel (channel 1, blue) (*Figure 1A*). In contrast, when the wild type and the mutant DNA were mixed at 1% allele frequency, fluorescence signal in the FAM channel (channel 1, blue) was detected (*Figure 1B*).

Next, we tested a series of 13 non-clinical liquid biopsy reference samples spanning different mutations and allele frequencies. These samples were provided by AstraZeneca and none of the laboratory technicians were informed about the genotype or mutant allelic frequencies at the time of execution of the tests. The results show that our ddPCR assay accurately detected the T790M allele in frequencies ranging from 5 to 0.5% (*Table 2*).

Finally, we applied the ddPCR assay to analyze DNA extracted from FFPE tissue biopsies from NSCLC patients. These DNA samples were previously genotyped in our laboratory by real-time PCR using cobas® EGFR Mutation Test. The ddPCR assay detected the mutation in all samples that previously tested positive for the T790M allele.

EGFR mutation testing in clinical plasma samples

The ddPCR assay was used to screen plasma samples from 77 patients for the T790M mutation in *EGFR*. A total of 40 plasma samples were identified as positive for the T790M mutation, corresponding to 52%. The number of mutant molecules per mL of plasma ranged from 1 to 6,000 as depicted in *Figure 2*. A second sample was collected for 18 patients that initially tested negative (*Table 3*). As recommended by current guidelines (9,16), when the liquid biopsy was negative, re-biopsy was performed for analysis of tumor tissue. However, this was only possible in 12 cases, either because the patient status did not allow re-biopsy or the tumor was not accessible. We detected the T790M mutation in 2 of the 12 re-biopsies (17%) (*Table 3*). In the remaining cases, we analyzed either a second plasma sample (5 patients) or a sample of bronchoalveolar fluid (1 patient), and detected the T790M mutation in 3 of these 6 liquid biopsies (50%) (*Table 3*).

Table 3 Genotyping results for samples collected after an initial negative plasma test

1st test (plasma samples)		2nd test		
Patient #	Date	Date	Sample type	Result
P#1	29/11/2016	03/01/2017	FFPE	Not detected
P#2	05/12/2016	27/12/2016	FFPE	Not detected
P#3	09/03/2017	06/04/2017	FFPE	Not detected
P#4	18/07/2017	07/08/2017	FFPE	Detected
P#5	18/08/2017	07/11/2017	FFPE	Not detected
P#6	15/09/2017	13/10/2017	Bronchoalveolar fluid	Detected
P#7	25/09/2017	02/11/2017	FFPE	Not detected
P#8	14/12/2017	22/01/2018	FFPE	Not detected
P#9	07/02/2018	15/03/2018	FFPE	Detected
P#10	07/03/2018	19/04/2018	FFPE	Not detected
P#11	15/03/2018	20/11/2018	Plasma	Detected
P#12	22/05/2018	22/08/2018	FFPE	Not detected
P#13	28/06/2018	20/11/2018	Plasma	Detected
P#14	26/11/2018	05/12/2018	FFPE	Not detected
P#15	23/01/2019	13/03/2019	FFPE	Not detected
P#16	24/04/2019	21/05/2019	Plasma	Not detected
P#17	18/07/2019	12/09/2019	Plasma	Not detected
P#18	12/09/2019	16/10/2019	Plasma	Not detected

FFPE, formalin-fixed paraffin-embedded.

Table 4 Association of T790M detected in progression with *EGFR* activating mutations detected at diagnosis

EGFR mutation at diagnosis	T790M detected	T790M not detected
del19	N=17 (61%)	N=8 (35%)
L858R	N=9 (32%)	N=10 (43%)
Other	N=2 (7%)	N=5 (22%)

del19: any deletion or delins in exon 19; L858R: substitution of amino acid leucine to arginine at codon 858 in exon 21.

Finally, we analyzed the association of T790M detected in progression with *EGFR* activating mutations detected at diagnosis (*Table 4*). The majority (61%) of tumors positive for T790M had an *EGFR* del19 mutation at diagnosis, whereas the majority (65%) of tumors negative for T790M had either L858R or other *EGFR* mutations (*Table 4*).

Discussion

We implemented a liquid biopsy ddPCR assay that enabled the detection and quantification of the *EGFR* T790M mutation at cfDNA allele frequency as low as 0.5%. With a turnaround time of 3–5 business days, the assay was routinely used in clinical practice since it played an important role in deciding the next line of treatment in NSCLC patients that acquired resistance to first- or second-generation EGFR-TKIs.

We analyzed plasma samples from 77 patients and detected the *EGFR* T790M mutation in 52%. Although this was a retrospective study potentially influenced by

bias associated with patient selection, our result is in good agreement with previous evidence indicating that about half of lung cancers that become resistant to EGFR TKIs acquire the *EGFR* T790M mutation (13). Most important, our results revealed that performing the ddPCR assay in a second liquid biopsy collected 1–8 months after a first negative plasma test increased by 50% the number of positive cases. Thus, the use of repeated ddPCR-based cfDNA genotyping may obviate tissue re-biopsy in cases that the tumor is not accessible or the patient has a poor functional status.

In our study, among tumors harboring the T790M mutation detected in either liquid biopsy or tumor tissue at disease progression, the majority (61%) had an *EGFR* del19 mutation at diagnosis. This is consistent with previous reports indicating that the T790M mutation is more frequent in patients with an *EGFR* exon 19 deletion mutation. In a cohort of 314 Japanese patients studied by re-biopsy, the T790M mutation was detected in 55.6% of cases with del19 mutation and in 43.0% of cases with the L858R mutation (26). In another study, the T790M mutation was present in 63% (26/41) of patients with *EGFR* exon 19 deletion and in 38% (12/32) of patients with L858R mutation (27). A more recent literature review confirmed that detection of the T790M mutation was more frequent in del19 mutated patients (53%) than in L858R mutated patients (36%) with acquired resistance to EGFR-TKIs (28). One possibility is that patients with the del19 mutation are more sensitive to TKIs, and therefore cells with the T790M mutation are more likely to be selected and enriched (28).

An important advantage of ddPCR-based assays is the ability to provide absolute quantification of DNA molecules. In our study cohort, the number of T790M mutant molecules per mL of plasma ranged from 1 to 6,000. Quantification of mutant molecules in plasma may be useful to monitor response to treatment. Indeed, a recent study reported that after 6 weeks of treatment with osimertinib, the T790M mutation load assessed by ddPCR decreased to very low level (23). However, whether the number of mutant molecules affects disease progression and response to therapy remains unknown. Clearly, future studies using quantitative diagnostic assays such as ddPCR are needed to maximize the benefits of precision therapy for advanced NSCLC.

Acknowledgments

We thank AstraZeneca for providing non-clinical control samples.

Funding: This work was supported by Fundação para a Ciência e Tecnologia (FCT)/Ministério da Ciência, Tecnologia e Ensino Superior - Fundos do Orçamento de Estado (UID/BIM/50005/2019), and FCT/FEDER/POR Lisboa 2020, Programa Operacional Regional de Lisboa, PORTUGAL 2020 (LISBOA-01-0145-FEDER-016394). C.S. was a recipient of a FCT fellowship (SFRH/BDE/110544/2015).

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at http://dx.doi.org/10.21037/tlcr-20-1010

Data Sharing Statement: Available at http://dx.doi.org/10.21037/tlcr-20-1010

Peer Review File: Available at http://dx.doi.org/10.21037/tlcr-20-1010

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tlcr-20-1010). ACS reports receiving grants from AstraZeneca, as well as personal fees and non-financial support from AstraZeneca and Novartis. CS reports receiving personal fees from AstraZeneca. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics committee of Centro Hospitalar Universitário Lisboa Norte and Lisbon Academic Medical Center (n. 311/20). Because the study involves the secondary analysis of existing data collected during standard of care procedures, a waiver of informed consent was requested and approved.

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
2. Mok TS, Wu YL, Thongprasert S, et al. Gefitinib or Carboplatin-Paclitaxel in Pulmonary Adenocarcinoma. N Engl J Med 2009;361:947-57.
3. Rosell R, Carcereny E, Gervais R, et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): A multicentre, open-label, randomised phase 3 trial. Lancet Oncol 2012;13:239-46.
4. Yang JCH, Sequist LV, Geater SL, et al. Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon EGFR mutations: A combined post-hoc analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6. Lancet Oncol 2015;16:830-8.
5. Wu YL, Cheng Y, Zhou X, et al. Dacomitinib versus gefitinib as first-line treatment for patients with EGFR-mutation-positive non-small-cell lung cancer (ARCHER 1050): a randomised, open-label, phase 3 trial. Lancet Oncol 2017;18:1454-66.
6. Soria JC, Ohe Y, Vansteenkiste J, et al. Osimertinib in untreated EGFR-Mutated advanced non-small-cell lung cancer. N Engl J Med 2018;378:113-25.
7. Paez JG, Jänne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004;304:1497-500.
8. Stewart EL, Tan SZ, Liu G, et al. Known and putative mechanisms of resistance to EGFR targeted therapies in NSCLC patients with EGFR mutations-a review. Transl Lung Cancer Res 2015;4:67-81.
9. Planchard D, Popat S, Kerr K, et al. Metastatic non-small cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2018;29 Suppl 4:iv192-237.
10. Morgillo F, Della Corte CM, Fasano M, et al. Mechanisms of resistance to EGFR-targeted drugs: Lung cancer. ESMO Open 2016;1:e000060.
11. Thress KS, Paweletz CP, Felip E, et al. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. Nat Med 2015;21:560-2.
12. AF Gazdar. Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors. Oncogene 2009;28 Suppl 1:S24-31.
13. Sequist LV, Waltman BA, Dias-Santagata D, et al. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. Sci Transl Med 2011;3:75ra26.
14. Yun CH, Mengwasser KE, Toms AV, et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. Proc Natl Acad Sci U S A 2008;105:2070-5.
15. Murtuza A, Bulbul A, Shen JP, et al. Novel third-generation EGFR tyrosine kinase inhibitors and strategies to overcome therapeutic resistance in lung cancer. Cancer Res 2019;79:689-98.
16. Lindeman NI, Cagle PT, Aisner DL, et al. Updated Molecular Testing Guideline for the Selection of Lung Cancer Patients for Treatment With Targeted Tyrosine Kinase Inhibitors: Guideline From the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. Arch Pathol Lab Med 2018;142:321-46.
17. Jenkins S, Yang JCH, Ramalingam SS, et al. Plasma ctDNA Analysis for Detection of the EGFR T790M Mutation in Patients with Advanced Non–Small Cell Lung Cancer. J Thorac Oncol 2017;12:1061-70.
18. Rolfo C, Cardona AF, Cristofanilli M, et al. Challenges and opportunities of cfDNA analysis implementation in clinical practice: Perspective of the International Society of Liquid Biopsy (ISLB). Crit Rev Oncol Hematol 2020;151:102978.
19. Goto K, Ichinose Y, Ohe Y, et al. Epidermal growth factor receptor mutation status in circulating free DNA in serum: From IPASS, a phase III study of gefitinib or carboplatin/paclitaxel in non-small cell lung cancer. J Thorac Oncol 2012;7:115-21.
20. Mok T, Wu YL, Lee JS, et al. Detection and dynamic changes of EGFR mutations from circulating tumor DNA as a predictor of survival outcomes in NSCLC Patients treated with first-line intercalated erlotinib and chemotherapy. Clin Cancer Res 2015;21:3196-203.
21. Remon J, Caramella C, Jovelet C, et al. Osimertinib benefit in EGFR-mutant NSCLC patients with T790M-mutation detected by circulating tumour DNA. Ann Oncol

2017;28:784-90.
22. Sacher AG, Paweletz C, Dahlberg SE, et al. Prospective validation of rapid plasma genotyping for the detection of EGFR and kras mutations in advanced lung cancer. JAMA Oncol 2016;2:1014-22.
23. Li C, Jia R, Liu H, et al. EGFR T790M detection and osimertinib treatment response evaluation by liquid biopsy in lung adenocarcinoma patients with acquired resistance to first generation EGFR tyrosine kinase inhibitors. Diagn Pathol 2018;13:49.
24. Guo QM, Wang L, Yu WJ, et al. Detection of Plasma EGFR Mutations in NSCLC Patients with a Validated ddPCR Lung cfDNA Assay. J Cancer 2019;10:4341-9.
25. Rare Mutation Detection Best Practices Guidelines [Internet]. 2015 Feb [cited 2020 May]; Available online: https://www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6628.pdf
26. Nosaki K, Satouchi M, Kurata T, et al. Re-biopsy status among non-small cell lung cancer patients in Japan: A retrospective study. Lung Cancer 2016;101:1-8.
27. Matsuo N, Azuma K, Sakai K, et al. Association of EGFR Exon 19 Deletion and EGFR-TKI Treatment Duration with Frequency of T790M Mutation in EGFR-Mutant Lung Cancer Patients. Sci Rep 2016;6:36458.
28. Liang H, Pan Z, Wang W, et al. The alteration of T790M between 19 del and L858R in NSCLC in the course of EGFR-TKIs therapy: A literature-based pooled analysis. J Thorac Dis 2018;10:2311-20.

Cite this article as: Silveira C, Sousa AC, Janeiro A, Malveiro S, Teixeira E, Brysch E, Pantarotto M, Felizardo M, Madureira R, Nogueira F, Guimarães C, Matos C, Canário D, Bruges-Armas J, Carmo-Fonseca M. Detection and quantification of *EGFR* T790M mutation in liquid biopsies by droplet digital PCR. Transl Lung Cancer Res 2021;10(3):1200-1208. doi: 10.21037/tlcr-20-1010

Lung Cancer 139 (2020) 35–40

Contents lists available at ScienceDirect

Lung Cancer

journal homepage: www.elsevier.com/locate/lungcan

Detection of rare and novel *EGFR* mutations in NSCLC patients: Implications for treatment-decision

AC. Sousa[a,1], C. Silveira[a,1], A. Janeiro[a], S. Malveiro[a], AR. Oliveira[a], M. Felizardo[b], F. Nogueira[c], E. Teixeira[d], J. Martins[d], M. Carmo-Fonseca[e,*]

[a] GenoMed, Diagnósticos de Medicina Molecular, SA, Lisboa, Portugal
[b] Hospital Beatriz Ângelo, Serviço de Pneumologia, Lisboa, Portugal
[c] Hospital Egas Moniz, Serviço de Pneumologia, Lisboa, Portugal
[d] Hospital Pulido Valente, Hospital de Dia de Pneumologia Oncológica, Centro Hospitalar Lisboa Norte, Centro Académico de Medicina de Lisboa, Lisboa, Portugal
[e] Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal

ARTICLE INFO

Keywords:
EGFR
NSCLC
Tyrosine-kinase domain mutations
Molecular diagnosis

ABSTRACT

Objectives: Mutations in the gene that encodes epidermal growth factor receptor (*EGFR*) are biomarkers that predict how non-small cell lung cancer (NSCLC) patients respond to EGFR-targeted therapies collectively known as tyrosine kinase inhibitors (TKIs). Thus, *EGFR* genotyping provides crucial information for treatment decision. Both Sanger sequencing and real-time PCR methodologies are used for *EGFR* genotyping. However, methods based on real-time PCR have limitations, as they may not detect rare or novel mutations. The aim of this study was to determine the prevalence of rare mutations in the tyrosine kinase domain (exons 18–21) of the *EGFR* gene not targeted by the most frequently used real-time PCR approaches, i.e., the cobas® EGFR Mutation Test, and the Idylla™ EGFR Mutation Assay.

Methods: A total of 1228 NSCLC patients were screened for mutations in exons 18–21 of the *EGFR* gene using Sanger sequencing.

Results: We observed that 252 patients (~20%) had at least one mutation in the *EGFR* gene, and 38 (~3%) carried uncommon genetic alterations that would not be identified by the cobas® or the Idylla™ tests. We further found six new single mutations and seven previously unreported compound mutations. Clinical information and patient outcome are presented for these cases.

Conclusions: This study highlights the value of sequencing-based approaches to identify rare mutations. Our results add to the inventory of known *EGFR* mutations, thus contributing to improved lung cancer precision treatment.

1. Introduction

Every year, almost 2 million people receive a diagnosis of lung cancer [1]. Of these, 85% have non–small cell lung cancer (NSCLC) [2]. Tobacco smoking is associated with 80% of all NSCLC cases but incidence in non-smokers has increased in the last years [3]. Several genetic alterations define NSCLC subtypes [4]. In particular, mutations in exons 18–21 of the *EGFR* gene are often detected in tumor samples of NSCLC patients. Mutations in these exons, which encode the tyrosine kinase domain of the Epidermal Growth Factor Receptor (EGFR), result in a gain in function, leaving the intracellular signaling pathway of EGFR constitutively active. About 10–30% of NSCLC samples harbor somatic mutations in exons 18, 19, 20 and / or 21 of the *EGFR* gene kinase domain [5].

In the last years, novel drugs known as EGFR-targeted therapies, or EGFR-tyrosine kinase inhibitors (TKIs), namely gefitinib, erlotinib, afatinib and osimertinib, have been successful in delaying disease progression in a subgroup of NSCLC patients [6].

Mutations in the *EGFR* gene are considered the most robust predictive biomarkers for response to EGFR-TKIs. Additionally, it has been suggested that *EGFR* mutations can also be useful for diagnosis purposes and therefore should be tested along with standard clinical examination, pathology and imaging studies [7]. The guidelines set by the National Comprehensive Cancer Network (NCCN) and the European Society for Medical Oncology (ESMO) for the diagnosis, treatment and follow-up of NSCLC patients, state that *EGFR* genotyping is

* Corresponding author at: Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Avenida Professor Egas Moniz, 1649-028, Lisboa, Portugal.
E-mail address: carmo.fonseca@medicina.ulisboa.pt (M. Carmo-Fonseca).
[1] These authors contributed equally to this study.

https://doi.org/10.1016/j.lungcan.2019.10.030
Received 6 June 2019; Received in revised form 28 October 2019; Accepted 29 October 2019

Table 1
Single mutations identified by Sanger sequencing that are not catalogued in the cobas® and the Idylla™ kit specifications. The mutations are listed according to the amino acid position.

Patient ID	Protein level	Nucleotide Level	Exon	COSMIC ID	Detectable by cobas/Idylla	Ref
1	**E709_T710delinsD** p.(Glu709_Thr710delinsAsp)	c.2127_2129delAAC	18	COSM51525	No	[24,27]
11						
22						
28	**K745_E746insIPVAIK** p.(Lys745_Glu746insIleProValAlaIleLys)	c.2234_2235ins18 (or c.2217_2234dup)	19	COSM51504	Not clear	[28,29]
21	**K745_E746insVPVAIK** p.(Lys745_Glu746insValProValAlaIleLys)	c.2236_2237ins18 (or c.2219_2236dup)	19	COSM26444	Not clear	[28,29]
24	**E746_A750delinsQP** p.(Glu746_Ala750delinsGlnPro)	c.2236_2248delinsCAAC	19	COSM13557	Not clear	[30]
4	**L747_A755delinsNRQG** p.(Leu747_Ala755delinsAsnArgGlnGly)	c.2239_2265delinsAACCGACAAGGA	19	–	Not clear	**Not reported before**
37	**A750_I759delinsPT** p.(Ala750_Ile759delinsProThr)	c.2248_2276delinsCCAAC	19	COSM5023004	Not clear	[31,32]
25	**T751_I759delinsN** p.(Thr751_Ile759delinsAsn)	c.2252_2276delinsA	19	COSM96856	Not clear	[33]
30	**A763_Y764insLQEA** p.(Ala763_Tyr764insLeuGlnGluAla)	c.2289_2290insCTCCAGGAAGCC	20	–	No	**Not reported before**
26	**S768_V769delinsIL (or S768I + V769 L)** p.(Ser768_Val769delinsIleLeu) (or p.(Ser768Ile) + p.(Val769Leu))	c.2303_2305 GCG > TTT (or c.2303_2305delinsTTT)	20	COSM1651576[§]	Not clear	[24,34]
23	**V769_D770insCV** p.(Val769_Asp770insCysVal)	c.2307_2308insTGCGTG	20	COSM12379	Not clear	[35,36]
31	**D770_N771insGF** (p.Asp770_Asn771insGlyPhe)	c.2310_2311insGGGTTT	20	COSM655155	Not clear	[28]
10	**D770_N771insGV** p.(Asp770_Asn771insGlyVal)	c.2310_2311insGGGCTT	20	–	Not clear	[37]
29	**N771delinsHH** p.(Asn771delinsHisHis)	c.2311delinsCACC	20	–	Not clear	**Not reported before**
12	**H773_V774insPH** p.(His773_Val774insProHis)	c.2319_2320insCCACAC	20	COSM28944[§]	Not clear	[28,36]
16	**H773_V774insNPH** p.(His773_Val774insAsnProHis)	c.2319_2320insAACCCCCAC	20	COSM12381	Not clear	[35,36]
36						
20	**V774L** p.(Val774Leu)	c.2320 G > T	20	COSM25090	Not clear	[38]
7	**V774_C775insHV (or H773_V774dup)** p.(Val774_Cys775insHisVal) (or p.(His773_Val774dup))	c.2321_2322insCCACGT (or c.2316_2321dup)	20	COSM18432	Not clear	[28,36]
17	**G779F** p.(Gly779Phe)	c.2335_2336GG > TT (or c.2335_2336delinsTT)	20	COSM13007	No	[39]
6	**W817X (or W817*)** p.(Trp817X) (or p.(Trp817*))	c.2450 G > A	20	–	No	[40]
32	**K823E** p.(Lys823Glu)	c.2467A > G	20	–	No	[15]
15	**G857Wfs*40** p.(Gly857Trpfs*40)	c.2568_2569insT (or c.2568dup)	21	–	No	**Not reported before**

[§] COSMIC ID information for protein only.

fundamental for decision making, as it may indicate which treatment is most likely to be effective in each patient [8–10]. Therefore, great efforts have been made to standardize diagnostic tests. Panels of international experts from reference institutions such as the College of American Pathologist, International Association for the Study of Lung Cancer and the Association for Molecular Pathology [11] recommend the screening for mutations in the *EGFR* gene to be performed by one of the two complementary methodologies: PCR followed by direct sequencing or real-time PCR. The most commonly used real-time PCR-based techniques are cobas® EGFR Mutation Test and Idylla™ EGFR Mutation Assay. These tests are highly sensitive and can detect mutations in samples with tumor infiltration as low as 10–20%. Both tests use a targeted approach designed to detect the most common *EGFR* mutations [12]. Indeed, approximately 90% of all *EGFR* mutations consist of either deletion in exon 19 or a single specific point mutation, namely c.2753C > T (p.L858R) in exon 21 [13].

In this study, we analyzed the complete sequence of *EGFR* exons 18–21 in 1228 NSCLC patients by Sanger sequencing. We identified novel mutations and we found that approximately 3% of the patients carried genetic alterations that would not be identified by the cobas® and the Idylla™ tests.

2. Materials and methods

2.1. Study design and patients

We analyzed data from patients who had received a diagnosis of NSCLC and were tested at GenoMed, in Lisbon, Portugal, for the presence of mutations in the *EGFR* gene between 2010 and 2017. Tumor samples from 1228 patients, including 539 females and 689 males, were studied. Patients' ages at the time of testing ranged between 21 and 94 years old. This is a retrospective study for which patient consent and ethical committee approval were not necessary.

2.2. Sample collection

For this study, we analyzed data obtained from either formalin-fixed paraffin-embedded (FFPE) blocks or FFPE sections, from the following hospitals in Portugal: *Hospital Pulido Valente* and *Hospital Santa Maria* (*Centro Hospitalar Lisboa Norte*); *Hospital Professor Doutor Fernando Fonseca*; *Hospital da Luz*; *Hospital Beatriz Ângelo*; *Hospital Egas Moniz* and *Hospital São Francisco Xavier* (*Centro Hospitalar Lisboa Ocidental*); *CUF*; *Hospital das Forças Armadas*. Additional hospitals located elsewhere in Portugal are *Hospital Vila Franca Xira*; *Hospital de Santo André* (*Centro Hospitalar de Leiria)*; *Hospital de Faro*.

2.3. Sample preparation

Only samples with tumor content higher than 20% were analyzed because for samples containing less than 20 to 30% of tumor infiltration and/or less than 200 neoplastic cells the Sanger sequencing procedure may not be sensitive enough for mutation detection [13,14]. DNA was extracted from tissue samples using a QIAmp Blood mini kit (catalog #51106), according to manufacturer's instructions. The DNA solution

Table 2
Compound mutations in which at least one is not catalogued in the cobas® and the Idylla™ kit specifications.

Patient ID	Protein level	Nucleotide Level	Exon	COSMIC ID	Detectable by cobas/Idylla	Ref
35	**E709A** p.(Glu709Ala)	c.2126A > C	18	COSM13427	No	[17,24]
	G719S p.(Gly719Ser)	c.2155 G > A		COSM6252	Yes	
14	**E709V** p.(Glu709Val)	c.2126A > T	18	COSM12371	No	**Doublet not**
	G719A p.(Gly719Ala)	c.2156 G > C		COSM6239	Yes	**reported before**
8	**G719A** p.(Gly719Ala)	c.2156 G > C	18	COSM6239	Yes	**Doublet not**
	D761Y p.(Asp761Tyr)	c.2281 G > T	19	COSM21984	No	**reported before**
9	**E709 K** p.(Glu709Lys)	c.2125 G > A	18	COSM12988	No	[17,27]
	L858R p.(Leu858Arg)	c.2573 T > G	21	COSM6224	Yes	
19	**E709G** p.(Glu709Gly)	c.2126A > G	18	COSM13009	No	[17,27]
	L858R p.(Leu858Arg)	c.2573 T > G	21	COSM6224	Yes	
13	**G719A** p.(Gly719Ala)	c.2156 G > C	18	COSM6239	Yes	**Doublet not**
	L833V p.(Leu833Val)	c.2497 T > G	21	COSM13424	No	**reported before**
18	**L747_I751delLREAT** p.(Leu747_Thr751del)	c.2240_2254del15	19	COSM12369	Yes	**Doublet not**
	K754E p.(Lys754Glu)	c.2260A > G		COSM85993	No	**reported before**
27	**V769M** p.(Val769Met)	c.2305 G > A	20	COSM13425	No	**Doublet not**
	N771_P772dup or P772_H773insNP p.(Asn771_Pro772dup) or p.(Pro772_His773insAsnPro)	c.2311_2316dup (or c.2316_2317insAACCCC)		COSM1738101[§]	No	**reported before**
34	**H773R** p.(His773Arg)	c.2318A > G	20	COSM13433	No	**Doublet not**
	V774 M p.(Val774Met)	c.2320 G > A		COSM13006	No	**reported before**
2 33	**L833V** p.(Leu833Val)	c.2497 T > G	21	COSM13424	No	[17]
	H835L p.(His835Leu)	c.2504A > T		COSM6227	No	
38	**L833V** p.(Leu833Val)	c.2497 T > G	21	COSM13424	No	[17,27]
	L858R p.(Leu858Arg)	c.2573 T > G		COSM6224	Yes	
3	**L858R** p.(Leu858Arg)	c.2573 T > G	21	COSM6224	Yes	[17]
	A871G p.(Ala871Gly)	c.2612C > G		COSM13008	No	
5	**A864T** p.(Ala864Thr)	c.2590 G > A	21	COSM13197	No	**Doublet not**
	H870Y p.(His870Tyr)	c.2608C > T		COSM53292	No	**reported before**

[§] COSMIC ID information for protein only.

Table 3
Similarities with previously described mutations.

Patient ID	New mutations found in this study	Previously described mutations
4	L747_A755delinsNRQG	L747_A755delinsSMS [§] L747_A755delinsSKS [§] L747_A755delinsAT [§]
6	W817X (G > A)	W817X (insAT)[[41]], [§]
15	G857Wfs*40	G857E/V/R [§], G857R [42] G857E [43]
29	N771delinsHH	N771delinsSTH[§] N771delinsGY [44] N771delinsTH [45] N771delinsFH [46] N771delinsGP [36]
30	A763_Y764insLQEA	A763_Y764insFQEA[22,§]

[§] Mutations described in the following databases: cBioPortal, COSMIC and/or OncoKB.

was quantified using a full spectrum (220–750 nm) spectrophotometer, NanoDrop® ND-1000 (NanoDrop® Technologies).

2.4. Mutation screening

We screened the tyrosine kinase domain (exons 18 to 21) of the *EGFR* gene (7p11.2, OMIM#131550, NM_005228.3, LRG_304(t1)). Exonic regions of interest were amplified by polymerase-chain-reaction (PCR) with flanking intronic primers. For exons 18 to 21, the reaction mixtures contained 1 µl of template DNA (150 ng), 2.5 µl of 10 × Buffer (Bioline), 2.5 µl of 2 mM dNTPs mixture (Bioline), a range from 0.6 to 0.75 µl of 50 mM $MgCl_2$ (Bioline), a range from 0.6 to 0.75 µl of each primer (10µM), 0.3 µl of BioTaq DNA polymerase (5U/µL) (Bioline) and ddH_2O, in a final volume of 25 µL. Thermal cycling conditions were as follows: initial denaturation at 96 °C for 5 min; followed by 40 cycles of 94 °C for 30 seg, annealing temperature between 63 °C and 67 °C for 30 seg and extension at 72 °C for 30 seg. The final extension was for 10 min at 72 °C. Supplementary Table 1 describes the primers used, fragment sizes (bp) and annealing temperature of each exon. PCR products were purified using the vacuum purification system Montage™ MultiScreen™ PCR96 Cleanup Kit (LSKM PCR50, Millipore) or Exo/SAP Go - PCR Purification Kit (GRISP), according to the manufacturer's instructions, and sequenced using the Big Dye v3.1 Cycle Sequencing kit (Applied Biosystems) on an automatic sequencer (ABI Prism 3100-Avant Capillary Array, 36 cm, Applied Biosystems). Data was analyzed using the DNA Sequencing Analysis Software 6™ Version 6.0 (Applied Biosystems). The Sequencher™ software was used to align samples against a reference on Ensembl. The nomenclature used to describe the genetic variants follows the guidelines of the Human Genome Variation Society.

3. Results

Sanger sequencing of tumor samples from 1228 individuals revealed at least one mutation in the *EGFR* gene in 252 cases, representing ~20% of the total population. This prevalence of mutations is higher than previously reported in unselected European patients [5], and the discrepancy is probably caused by the clinical selection bias of patients referred for *EGFR* genotyping.

Among the 252 samples with mutated *EGFR*, we identified 38 cases with mutations not explicitly listed in the specifications of the cobas® or the Idylla™ EGFR Mutation Tests (Tables 1 and 2). Among these 38 patients, 24 carried single mutations (Table 1) and 14 had compound mutations (Table 2). Clinical data for these patients is presented in Supplementary Table 2. The genetic alterations identified include point mutations, insertions and delins. Because many of these insertions and

Table 4
Treatment and outcome of patients with mutations that are not catalogued in the cobas® and the Idylla™ kit specifications.

Patient ID	Mutation	Initial treatment	Available information	OS
1	E709_T710delinsD	Gefitinib	PS at diagnosis: PS1 Best response: PD PFS: 3 months Stage at progression: IVB PS at progression: PS1 2nd Treatment: Chemotherapy Best response: SD PFS: 14 months Stage at progression: IVB	24 months
2	L833 V + H835L	Surgery + radiotherapy for bone metastases.	2nd Treatment: Chemotherapy + Erlotinib Best response: SD PFS: 9 months	44 months
5	A864 T + H870Y	Surgery + adjuvant chemotherapy	PS at diagnosis: PS1 Best response: CR PFS: 11 months Stage at progression: IVA PS at progression: PS1 2nd Treatment: Gefitinib Best response: PD PFS: 3,5 months Stage at progression: IVA	46 months
6	W817X (or W817*)	Chemotherapy + Erlotinib	Best response: PD PFS: 4 months	19 months
7	V774_C775insHV (or H773_V774dup)	Erlotinib	PS at diagnosis: PS1 Best response: PD PFS: 4 months Stage at progression: IVB PS at progression: PS2 2nd Treatment: Palliative	9 months
8	G719A + D761Y	Erlotinib	PS at diagnosis: PS2 Best response: SD Stage at progression: IVB	9 months
10	D770_N771insGV	Surgery + adjuvant chemotherapy and radiotherapy	Best response: CR PFS: 79 months	103 months
11	E709_T710delinsD	Surgery + adjuvant chemotherapy	PS at diagnosis: PS1 Best response: PD PFS: 2,5 months Stage at progression: IVA PS at progression: PS2 2nd Treatment: Erlotinib Best response: PD PFS: 4 months PS at progression: PS2	26 months
13	G719A + L833V	Gefitinib	PS at diagnosis: PS2 Best response: PD PFS: 2 months Stage at progression: IVB PS at progression: PS2 2nd Treatment: Chemotherapy Best response: SD PFS: 5 months Stage at progression: IVB	17 months
14	E709 V + G719A	Gefitinib	PS at diagnosis: PS1 Best response: PD PFS: 3 months Stage at progression: IVB PS at progression: PS3 2nd Treatment: Palliative	7 months
15	G857Wfs*40	Gefitinib + Radiotherapy (lung and CNS)	Best response: SD PFS: 7 months	9 months
16	H773_V774insNPH	Chemotherapy	PS at diagnosis: PS1 Best response: PR PFS: 8 months Stage at progression: IVA PS at progression: PS1 2nd Treatment: Chemotherapy Best response: PR PFS: 8 months Stage at progression: IVA PS at progression: PS2 3rd Treatment: Immunotherapy Best response: SD PFS: 3 months	22 months
17	G779F	Gefitinib	PS at diagnosis: PS2 Best response: PD Stage at progression: IVB	3 months
18	L747_T751delLREAT + K754E	Surgery + adjuvant chemotherapy	PS at diagnosis: PS0 Best response: CR PFS: 4 years and 9 months	Alive
19	E709 G + L858R	Chemotherapy	PS at diagnosis: PS1 Best response: SD PFS: 5 months Stage at progression: IVA PS at progression: PS1 2nd Treatment: Erlotinib Best response: PR PFS: 18 months Stage at progression: IVA	Alive
20	V774L	None	PS at diagnosis: PS3	4 months
21	K745_E746insVPVAIK	Gefitinib	PS at diagnosis: PS1 Best response: PR PFS: 6 months Stage at progression: IVB PS at progression: PS2 2nd Treatment: Erlotinib Best response: SD PFS: 10 months PS at progression: PS3	17 months
22	E709_T710delinsD	Chemotherapy	PS at diagnosis: PS1 Best response: SD PFS: 7 months Stage at progression: IVB PS at progression: PS2 2nd Treatment: Erlotinib Best response: PD PFS: 3 months PS at progression: PS3	18 months
23	V769_D770insCV	CNS radiotherapy	PS at diagnosis: PS2 Best response: PD PFS: 2 months Stage at progression: IVB	2 months
24	E746_A750delinsQP	Erlotinib	PS at diagnosis: PS1 Best response: PD PFS: 3 months Stage at progression: IVA PS at progression: PS1 2nd Treatment: Chemotherapy Best response: PR PFS: 24 months Stage at progression: IVB	32 months
25	T751_I759delinsN	Erlotinib	PS at diagnosis: PS3 Best response: PD PFS: 2 months PS at progression: PS4 2nd Treatment: Palliative	3,5 months
26	S768_V769delinsIL (or S768I + V769 L)	Erlotinib	PS at diagnosis: PS2 Best response: PD PFS: 3 months Stage at progression: IVB PS at progression: PS3 2nd Treatment: Palliative + radiotherapy for bone metastases	7 months
27	V769 M + P772_H773insNP	Chemotherapy	PS at diagnosis: PS1 Best response: SD PFS: 14 months Stage at progression: IVB PS at progression: PS2 2nd Treatment: Chemotherapy Best response: SD PFS: 5 months Stage at progression: IVB	31 months
28	K745_E746insIPVAIK	Surgery	2nd Treatment: Gefitinib Best response: SD PFS: 7 months	69 months
29	N771delinsHH	Surgery	PS at diagnosis: PS1 Best response: CR	NA
30	A763_Y764insLQEA	Chemotherapy + Erlotinib	Best response: PD PFS: 2 months	9 months
32	K823E	Surgery	PS at diagnosis: PS1 Best response: CR PFS: 9 months Stage at progression: IIIB PS at progression: PS1 2nd Treatment: Chemotherapy Best response: PD PFS: 3 months Stage at progression: IIIB	31 months
33	L833 V + H835L	Erlotinib + Bevacizumab	Best response: PD	NA
35	E709A + G719S	Surgery + adjuvant chemotherapy	PS at diagnosis: PS1 Best response: CR PFS: 11 months Stage at progression: IVA 2nd Treatment: Erlotinib Best response: SD PFS: Free of progession for 7 months	Alive
36	H773_V774insNPH	Chemotherapy	PS at diagnosis: PS0 Best response: CR PFS: 8 months Stage at progression: IVB PS at progression: PS3 2nd Treatment: CNR radiotherapy Best response: SD PFS: 5 months	19 months
37	A750_I759delinsPT	Surgery + adjuvant chemotherapy	PS at diagnosis: PS1 Best response: CR PFS: 4 months Stage at progression: IVA PS at progression: PS1 2nd Treatment: Erlotinib Best response: PR PFS: 15 months Stage at progression: IVB	Alive

CR = complete response, NA = not available, OS = overall survival, PD = progression of disease, PFS = progression-free survival, PR = partial response, PS = performance status, SD = stable disease, CNR = contrast-to-noise ratio, CNS = Central Nervous System.

delins are located in the same region where other more common mutations that are targeted by the kits occur, it is possible that they can be detected by the PCR assays (these cases are indicated as "Not clear" in Tables 1 and 2). However, even if a PCR assay detects a hit in these regions, the precise genetic alteration would remain unknown.

As indicated in Tables 1 and 2, many of the 38 mutations identified by Sanger sequencing were previously described as either oncogenic or likely oncogenic. However, we found four single variants (L747_A755delinsNRQG, A763_Y764insLQEA, N771delinsHH and G857Wfs*40) that have not been previously reported (patients 4, 15, 29, and 30; Table 1). Noteworthy, these new variants are similar to previously described mutations (Table 3). We also detected a K823E mutation (patient 32; Table 1), that although not described in the COSMIC database or in the OncoKB is mentioned in a protein modelling study [15]. According to the aforementioned study, K823E greatly reduces the phosphorylation of *EGFR*. This is a surprising finding as most mutations localized at the protein dimer interface have very little or no impact on the phosphorylation pattern [15]. We were not able to find any study reporting an NSCLC patient carrying this or a similar mutation. We additionally detected a W817X (c.2450 G > A) mutation (patient 6; Table 1), that was previously reported only in breast cancer [16]; a closely related mutation (W817X, c.2450_2451insAT) was however described in brain metastasis from a primary lung cancer (Table 3).

We further found tumor samples from seven patients with combinations of two mutations that have not been previously reported co-occurring in the same patient (patients 5, 8, 13, 14, 18, 27, and 34; Table 2). Doublet mutations were detected in 19 cases (~1.5%), of which 14 (~74%) include at least one variant not catalogued in the cobas® and the Idylla™ tests (Table 2). Amino acids E709, G719, L854 were frequently altered, in agreement with a previous study. [17]

Finally, data on treatment response of patients with novel and rare mutations reveals very heterogeneous clinical outcomes (Table 4).

4. Discussion

Targeted-methods for *EGFR* genotyping are widely used but can only identify the most common genetic alterations. Our study highlights the value of sequencing-based approaches to identify rare mutations. Our results show that approximately 3% of NSCLC patients have rare mutations not catalogued in the cobas® and the Idylla™ tests.

To date, it remains unclear how tumors carrying uncommon mutations respond to first-line TKIs [18]. Previous studies have shown that exon 19 deletions are associated with EGFR-TKI sensitivity while exon 20 insertions are associated with treatment resistance [19–21]. However, one exon 20 insertion (p.A763_Y764insFQEA) was found associated with good response to EGFR-TKI therapy [22]. We observed that patient 30, who carried the novel but closely related mutation A763_Y764insLQEA and who was treated with Erlotinib, had disease progression in two months and an overall survival of 9 months.

The clinical significance of compound mutations is also unknown. Previous studies have reported mutations L858R and G719A co-occurring with other less common mutations [23]. The compound mutation L858R + A871 G detected in patient 3 was previously reported being associated with resistance to erlotinib [24]. The doublet L833V +H835L detected in patients 2 and 33 was previously described in patients that responded well to EGFR-TKI [25,26].

In this study, patients with novel or rare mutations that were treated with EGFR-TKIs had an overall survival ranging between 3 and 32 months. Further investigation of larger patient cohorts with sequencing-based approaches is needed in order to understand how uncommon genetic signatures influence treatment response. By expanding the inventory of *EGFR* mutations associated with NSCLC this study may contribute to improved lung cancer precision treatment.

Funding

This work was supported by Fundação para a Ciência e Tecnologia (FCT) / Ministério da Ciência, Tecnologia e Ensino Superior through Fundos do Orçamento de Estado (UID/BIM/50005/2019), and FCT/FEDER/POR Lisboa 2020, Programa Operacional Regional de Lisboa, PORTUGAL 2020 (LISBOA-01-0145-FEDER-016394). C.S. was a recipient of a FCT fellowship (SFRH/BDE/110544/2015).

Declaration of Competing Interest

A.C. Sousa reports receiving grants from AstraZeneca, as well as personal fees and non-financial support from Novartis.

The remaining authors have nothing to disclose.

Acknowledgements

We would like to thank Marcia Triunfol for her assistance in producing this manuscript.

Author Contributions

ACS, CS and MCF conceived the study. MF, FN, ET and JM provided most of the analyzed samples. They also provided clinical information and their analysis and interpretation of the data. Laboratory tasks were performed by ACS, CS, AJ, SM, ARO. ACS and CS compiled and interpreted the results and wrote the first draft. All authors read, edited and approved the final manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.lungcan.2019.10.030.

References

[1] J. Ferlay, D.M. Parkin, E. Steliarova-Foucher, Estimates of cancer incidence and mortality in Europe in 2008, Eur. J. Cancer 46 (4) (2010) 765–781, https://doi.org/10.1016/j.ejca.2009.12.014.

[2] C.S. Dela Cruz, L.T. Tanoue, Ra Matthay, Lung cancer: epidemiology, etiology and prevention, Clin. Chest Med. 32 (4) (2011) 1–61, https://doi.org/10.1016/j.ccm.2011.09.001.Lung.

[3] A.T. Shaw, B.Y. Yeap, M. Mino-Kenudson, et al., Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK, J. Clin. Oncol. 27 (26) (2009) 4247–4253, https://doi.org/10.1200/JCO.2009.22.6993.

[4] L.A. Pikor, V.R. Ramnarine, S. Lam, W.L. Lam, Genetic alterations defining NSCLC subtypes and their therapeutic implications, Lung Cancer 82 (2) (2013) 179–189, https://doi.org/10.1016/j.lungcan.2013.07.025.

[5] S.V. Sharma, D.W. Bell, J. Settleman, D.A. Haber, Epidermal growth factor receptor mutations in lung cancer, Nat. Rev. Cancer 7 (3) (2007) 169–181, https://doi.org/10.1038/nrc2088.

[6] A. Gazdar, Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors, Oncogene 28 (Suppl 1) (2009) 1–14, https://doi.org/10.1038/onc.2009.198.Activating.

[7] I.K. Park, K. Hyun, E.R. Kim, S. Park, C.H. Kang, Y.T. Kim, The prognostic effect of the epidermal growth factor receptor gene mutation on recurrence dynamics of lung adenocarcinoma, Eur. J. Cardio-Thorac. Surg. 0 (June) (2018) 1–6, https://doi.org/10.1093/ejcts/ezy220.

[8] D.S. Ettinger, D.L. Aisner, D.E. Wood, et al., NCCN guidelines insights: non–small cell lung cancer, version 5.2018, J. Natl. Compr. Cancer Netw. 16 (7) (2018) 807–821, https://doi.org/10.6004/jnccn.2018.0062.

[9] G.A. Masters, S. Temin, C.G. Azzoli, et al., Systemic therapy for stage IV non-small-cell lung cancer: American Society of Clinical Oncology clinical practice guideline update, J. Clin. Oncol. 33 (30) (2015), https://doi.org/10.1200/JCO.2015.62.1342 JCO.2015.62.1342-..

[10] K.M. Kerr, L. Bubendorf, M.J. Edelman, et al., Second ESMO consensus conference on lung cancer: pathology and molecular biomarkers for non-small-cell lung cancer, Ann. Oncol. 25 (9) (2014) 1681–1690, https://doi.org/10.1093/annonc/mdu145.

[11] N.I. Lindeman, P.T. Cagle, D.L. Aisner, et al., Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: guideline from the College of American Pathologists, the International Association for the Study of Lung Cancer, Arch. Pathol. Lab. Med. (2018), https://doi.org/10.5858/arpa.2017-0388-CP arpa.2017-0388-CP.

[12] G. Ellison, G. Zhu, A. Moulis, S. Dearden, G. Speake, R. McCormack, EGFR mutation testing in lung cancer: a review of available methods and their use for analysis of

tumour tissue and cytology samples, J. Clin. Pathol. 66 (2) (2013) 79–89, https://doi.org/10.1136/jclinpath-2012-201194.
[13] A.F. Leary, Castro D.G. De, A.G. Nicholson, et al., Establishing an EGFR mutation screening service for non-small cell lung cancer - sample quality criteria and candidate histological predictors, Eur. J. Cancer 48 (1) (2012) 61–67, https://doi.org/10.1016/j.ejca.2011.09.022.
[14] M. Ladanyi, W. Pao, Lung adenocarcinoma: guiding EGFR-targeted therapy and beyond, Mod. Pathol. 21 (2008) S16–S22, https://doi.org/10.1038/modpathol.3801018.
[15] K.S. Gajiwala, EGFR: tale of the C-terminal tail, Protein Sci 22 (7) (2013) 995–999, https://doi.org/10.1002/pro.2283.
[16] F. Weber, K. Fukino, T. Sawada, et al., Variability in organ-specific EGFR mutational spectra in tumour epithelium and stroma may be the biological basis for differential responses to tyrosine kinase inhibitors, Br. J. Cancer 92 (10) (2005) 1922–1926, https://doi.org/10.1038/sj.bjc.6602557.
[17] Z. Chen, J. Feng, J.S. Saldivar, D. Gu, A. Bockholt, S.S. Sommer, EGFR somatic doublets in lung cancer are frequent and generally arise from a pair of driver mutations uncommonly seen as singlet mutations: one-third of doublets occur at five pairs of amino acids, Oncogene 27 (31) (2008) 4336–4343, https://doi.org/10.1038/onc.2008.71.
[18] S. Watanabe, Y. Minegishi, H. Yoshizawa, et al., Effectiveness of gefitinib against Non–Small-cell lung cancer with the uncommon EGFR mutations G719X and L861Q, J. Thorac. Oncol. 9 (2) (2014) 189–194, https://doi.org/10.1097/JTO.0000000000000048.
[19] T.J. Lynch, D.W. Bell, R. Sordella, et al., Activating mutations in the epidermal growth factor receptor underlying responsiveness of non–small-cell lung cancer to gefitinib, N. Engl. J. Med. 350 (21) (2004) 2129–2139, https://doi.org/10.1056/NEJMoa040938.
[20] J.G. Paez, EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy, Science (80) 304 (5676) (2004) 1497–1500, https://doi.org/10.1126/science.1099314.
[21] W. Pao, V. Miller, M. Zakowski, et al., EGF receptor gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib, Proc. Natl. Acad. Sci 101 (36) (2004) 13306–13311, https://doi.org/10.1073/pnas.0405220101.
[22] H. Yasuda, E. Park, C.-H. Yun, et al., Structural, biochemical, and clinical characterization of epidermal growth factor receptor (EGFR) exon 20 insertion mutations in lung cancer, Sci. Transl. Med. 5 (216) (2013), https://doi.org/10.1126/scitranslmed.3007205 216ra177-216ra177.
[23] E.Y. Kim, E.N. Cho, H.S. Park, et al., Compound EGFR mutation is frequently detected with co-mutations of actionable genes and associated with poor clinical outcome in lung adenocarcinoma, Cancer Biol. Ther. 17 (3) (2016) 237–245, https://doi.org/10.1080/15384047.2016.1139235.
[24] S. Kobayashi, H.M. Canepa, A.S. Bailey, et al., Compound EGFR mutations and response to EGFR tyrosine kinase inhibitors, J. Thorac. Oncol. 8 (1) (2013) 118–122, https://doi.org/10.1097/JTO.0b013e3182781e35.
[25] S. Frega, P. Conte, M. Fassan, V. Polo, G. Pasello, A triple rare E709K and L833V/H835L EGFR mutation responsive to an irreversible Pan-HER inhibitor: a case report of lung adenocarcinoma treated with afatinib, J. Thorac. Oncol. 11 (5) (2016) e63–e64, https://doi.org/10.1016/j.jtho.2016.01.023.
[26] T.Y. Yang, C.R. Tsai, K.C. Chen, K.A. Hsu, H.M. Lee, G.C. Chang, Good response to gefitinib in a lung adenocarcinoma harboring a heterozygous complex mutation of L833V and H835L in epidermal growth factor receptor gene, J. Clin. Oncol. 29 (16) (2011) 468–469, https://doi.org/10.1200/JCO.2010.33.5802.
[27] J.-Y. Wu, Chang Y.-C. Yu C-J, C.-H. Yang, J.-Y. Shih, P.-C. Yang, Effectiveness of tyrosine kinase inhibitors on “Uncommon” epidermal growth factor receptor mutations of unknown clinical significance in non-small cell lung cancer, Clin. Cancer Res. 17 (11) (2011) 3812–3821, https://doi.org/10.1158/1078-0432.CCR-10-3408.
[28] M.E. Arcila, K. Nafa, J.E. Chaft, et al., EGFR exon 20 insertion mutations in lung adenocarcinomas: prevalence, molecular heterogeneity, and clinicopathologic characteristics, Mol. Cancer Ther. 12 (2) (2013) 220–229, https://doi.org/10.1158/1535-7163.MCT-12-0620.
[29] M. He, M. Capelletti, K. Nafa, et al., EGFR exon 19 insertions: a new family of sensitizing EGFR mutations in lung adenocarcinoma, Clin. Cancer Res. 18 (6) (2012) 1790–1797, https://doi.org/10.1158/1078-0432.CCR-11-2361.
[30] C. Leduc, J.P. Merlio, B. Besse, et al., Clinical and molecular characteristics of non-small-cell lung cancer (NSCLC) harboring EGFR mutation: results of the nationwide French cooperative thoracic intergroup (IFCT) program, Ann. Oncol. 28 (11) (2017) 2715–2724, https://doi.org/10.1093/annonc/mdx404.
[31] R.B. Shahi, S. De Brakeleer, J. De Grève, C. Geers, P. In’t Veld, E. Teugels, Detection of EGFR-TK domain-activating mutations in NSCLC with generic PCR-based methods, Appl. Immunohistochem. Mol. Morphol. 23 (3) (2015) 163–171, https://doi.org/10.1097/PDM.0000000000000035.
[32] J. De Grève, J. Van Meerbeeck, J.F. Vansteenkiste, et al., Prospective evaluation of first-line erlotinib in advanced non-small cell lung cancer (NSCLC) carrying an activating EGFR mutation: a multicenter academic phase II study in caucasian patients (FIELT), Rosell R, ed. PLoS One 11 (3) (2016) e0147599, , https://doi.org/10.1371/journal.pone.0147599.
[33] R. Penzel, C. Sers, Y. Chen, et al., EGFR mutation detection in NSCLC-assessment of diagnostic application and recommendations of the German panel for mutation testing in NSCLC, Virch. Arch. 458 (1) (2011) 95–98, https://doi.org/10.1007/s00428-010-1000-y.
[34] H. Asahina, K. Yamazaki, I. Kinoshita, H. Yokouchi, H. Dosaka-Akita, M. Nishimura, Non-responsiveness to gefitinib in a patient with lung adenocarcinoma having rare EGFR mutations S768I and V769L, Lung Cancer 54 (3) (2006) 419–422, https://doi.org/10.1016/j.lungcan.2006.09.005.
[35] H. Shigematsu, L. Lin, T. Takahashi, et al., Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers, JNCI J. Natl. Cancer Inst. 97 (5) (2005) 339–346, https://doi.org/10.1093/jnci/dji055.
[36] H. Yasuda, S. Kobayashi, D.B. Costa, EGFR exon 20 insertion mutations in non-small-cell lung cancer: preclinical data and clinical implications, Lancet Oncol. 13 (1) (2012) e23–e31, https://doi.org/10.1016/S1470-2045(11)70129-2.
[37] J. Naidoo, C.S. Sima, K. Rodriguez, et al., Epidermal growth factor receptor exon 20 insertions in advanced lung adenocarcinomas: clinical outcomes and response to erlotinib, Cancer 121 (18) (2015) 3212–3220, https://doi.org/10.1002/cncr.29493.
[38] T.J. Pugh, G. Bebb, L. Barclay, et al., Correlations of EGFR mutations and increases in EGFR and HER2 copy number to gefitinib response in a retrospective analysis of lung cancer patients, BMC Cancer 7 (1) (2007) 128, https://doi.org/10.1186/1471-2407-7-128.
[39] S.H. Yang, L.E. Mechanic, P. Yang, et al., Mutations in the tyrosine kinase domain of the epidermal growth factor receptor in Non-small cell lung cancer, Clin. Cancer Res. 11 (6) (2005) 2106–2110, https://doi.org/10.1158/1078-0432.CCR-04-1853.
[40] F. Weber, K. Fukino, T. Sawada, et al., Variability in organ-specific EGFR mutational spectra in tumour epithelium and stroma may be the biological basis for differential responses to tyrosine kinase inhibitors, Br. J. Cancer 92 (10) (2005) 1922–1926, https://doi.org/10.1038/sj.bjc.6602557.
[41] M. Preusser, A.S. Berghoff, R. Koller, et al., Spectrum of gene mutations detected by next generation exome sequencing in brain metastases of lung adenocarcinoma, Eur. J. Cancer 51 (13) (2015) 1803–1811, https://doi.org/10.1016/j.ejca.2015.06.107.
[42] M.H. Hsieh, Y.F. Fang, W.C. Chang, et al., Complex mutation patterns of epidermal growth factor receptor gene associated with variable responses to gefitinib treatment in patients with non-small cell lung cancer, Lung Cancer 53 (3) (2006) 311–322, https://doi.org/10.1016/j.lungcan.2006.06.005.
[43] Y. Liu, B.-Q. Wu, H.-H. Zhong, P. Hui, W.-G. Fang, Screening for EGFR and KRAS mutations in non-small cell lung carcinomas using DNA extraction by hydrothermal pressure coupled with PCR-based direct sequencing, Int. J. Clin. Exp. Pathol. 6 (9) (2013) 1880–1889 http://www.ncbi.nlm.nih.gov/pubmed/24040454.
[44] T. Harada, A. Lopez-Chavez, L. Xi, M. Raffeld, Y. Wang, G. Giaccone, Characterization of epidermal growth factor receptor mutations in non-small-cell lung cancer patients of African-American ancestry, Oncogene 30 (15) (2011) 1744–1752, https://doi.org/10.1038/onc.2010.545.
[45] L.V. Sequist, V.A. Joshi, P.A. Janne, et al., Response to treatment and survival of patients with non-small cell lung cancer undergoing somatic EGFR mutation testing, Oncologist 12 (1) (2007) 90–98, https://doi.org/10.1634/theoncologist.12-1-90.
[46] J.P. Robichaux, Y.Y. Elamin, Z. Tan, et al., Mechanisms and clinical activity of an EGFR and HER2 exon 20-selective kinase inhibitor in non-small cell lung cancer, Nat. Med. 24 (5) (2018) 638–646, https://doi.org/10.1038/s41591-018-0007-9.

Supplementary Table 1. Primers used, fragment sizes (bp) and annealing temperature per exon.

Exon	Forward Primer	Reverse Primer	Fragment size (bp)	Annealing temperature (Ta)
18	CTGGCACCCAAGCCCATG	CCCACCAGACCATGAGAGG	302	67ºC
19	CCAGTGTCCCTCACCTTC	AGCAGGGTCTAGAGCAGAGCAGCTGCC	308	63ºC
20	GATCGCATTCATGCGTCTTCACC	TATCCCAGGAGCGCAGACC	358	65ºC
21	GGCATGAACATGACCCTGAAT	AGCTGCTGCGAGCTCACC	371	66ºC

Primers 19R e 20F were designed by Shigematsu and collaborators[35].

Supplementary Table 2. Patient information

Patient ID	Sex	Age at molecular test	Age at diagnosis	Ethnicity	NSCLC Histology	Smoking status	Stage at diagnosis
1	F	66	66	Caucasian	Adenocarcinoma	Current heavy (>15 pack year)	IVB
2	F	74	71	Asian	Adenocarcinoma	Never	IB
3	M	84	NA	NA	NA	NA	NA
4	M	67	NA	NA	Adenocarcinoma	NA	NA
5	M	49	48	Caucasian	Adenocarcinoma	Never	IIB
6	M	81	80	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IV
7	M	46	46	Caucasian	Adenocarcinoma	Never	IVB
8	M	68	68	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IVB
9	F	65	NA	NA	NA	NA	NA
10	M	59	51	Caucasian	Adenocarcinoma	Former light (≤15 pack year)	IIIA
11	F	46	46	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IIA
12	F	59	59	Caucasian	Adenocarcinoma	Current heavy (>15 pack year)	IV
13	F	59	59	Caucasian	Adenocarcinoma	Never	IVB
14	M	76	76	Caucasian	Adenocarcinoma	Never	IVB
15	M	75	75	Caucasian	Adenocarcinoma	Current heavy (>15 pack year)	IV
16	F	72	72	Caucasian	Adenocarcinoma	Never	IVA

17	M	72	72	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IVB
18	F	60	60	Caucasian	Adenocarcinoma	Never	IB
19	F	55	55	African	Adenocarcinoma	Former heavy (>15 pack year)	IVA
20	M	75	75	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IVB
21	F	67	66	Caucasian	Adenocarcinoma	Never	IVB
22	F	57	57	Caucasian	Adenocarcinoma	Never	IVB
23	F	52	52	African	Adenocarcinoma	Never	IVB
24	M	63	62	Caucasian	Adenocarcinoma	Former light (≤15 pack year)	IVA
25	F	81	81	Caucasian	Adenocarcinoma	Never	IVA
26	M	72	72	Caucasian	Adenocarcinoma	Never	IVB
27	F	69	69	African	Adenocarcinoma	Never	IVB
28	F	79	74	Caucasian	Adenocarcinoma	Never	IA
29	M	70	70	Caucasian	Adenocarcinoma	Current	IB
30	M	73	72	Caucasian	Adenocarcinoma	Former light (≤15 pack year)	IV
31	M	53	NA	Caucasian	Adenocarcinoma	NA	IIA
32	F	72	71	Caucasian	Adenocarcinoma	Former heavy (>15 pack year)	IA
33	M	64	NA	NA	Adenocarcinoma	NA	IB
34	F	83	83	Caucasian	Adenocarcinoma	NA	NA
35	M	64	64	Caucasian	Adenocarcinoma	Never	IIA
36	F	74	73	Caucasian	Adenocarcinoma	Never	IIIA
37	F	57	57	Caucasian	Adenocarcinoma	Never	IIIA
38	F	52	52	Caucasian	Adenocarcinoma	Current	IIIA

NA = not assessed.

Annex 4

cobas® EGFR Mutation Test – list of mutations

Exon 18		
Mutation type: G719X		
Nucleotide alteration	**Aminoacid change**	**ID COSMIC21**
c.2155G>A	p.G719S	6252
c.2155G>T	p.G719C	6253
c.2156G>C	p.G719A	6239

Exon 20		
Mutation type: T790M, S768I and exon 20 insertion		
Nucleotide alteration	**Aminoacid change**	**ID COSMIC21**
c.2303G>T	p.S768I	6241
c.2369C>T	p.T790M	6240
c.2319_2320insCAC	p.H773_V774insH	12377
c.2310_2311insGGT	p.D770_N771insG	12378
c.2307_2308ins9	p.V769_D770insASV	12376
c.2309_2310delins11	p.V769_D770insASV	13558
c.2311_2312ins9	p.D770_N771insSVD	13428

Exon 21		
Mutation type: L858R		
Nucleotide alteration	**Aminoacid change**	**ID COSMIC21**
c.2573T>G	L858R	6224
c.2573_2574TG>GT	L858R	12429

Exon 19		
Mutation type: exon 19 deletion		
Nucleotide alteration	**Aminoacid change**	**ID COSMIC21**
c.2235_2249del15	p.E746_A750del	6223
c.2236_2250del15	p.E746_A750del	6225
c.2240_2257del18	p.L747_P753delinsS	12370
c.2240_2254del15	p.L747_T751del	12369
c.2239_2256del18	p.L747_S752del	6255
c.2239_2251delinsC	p.L747_T751delinsP	12383
c.2237_2251del15	p.E746_T751delinsA	12678
c.2237_2255delinsT	p.E746_S752delinsV	12384
c.2239_2248delinsC	p.E747_A750delinsP	12382
c.2239_2253del15	p.L747_T751del	6254
c.2239_2247del9	p.L747_E749del	6254
c.2235_2252delinsAAT	p.E746_T751delinsI	13551
c.2236_2253del18	p.E746_T751del	12728
c.2237_2254del18	p.E746_S752delinsA	12367
c.2238_2255del18	p.E746_S752delinsD	6220
c.2238_2248_delinsGC	p.L747_A750delinsP	12422
c.2238_2252delinsGCA	p.L747_T751delinsQ	12419
c.2239_2258delinsCA	p.L747_P753delinsQ	12387
c.2240_2251del12	p.L747_T751delinsS	6210
c.2233_2247del15	p.K745_E749del	26038
c.2253_2276del24	p.S752_I759del	13556
c.2235_2248delinsAATTC	p.E746_A750delinsIP	13550
c.2237_2252delinsT	p.E746_T751delinsV	12386
c.2235_2251delinsAATTC	p.E746_T751delinsIP	13552
c.2235_2255delinsAAT	p.E746_S752delinsI	12385
c.2237_2253delinsTTGCT	p.E746_T751delinsVA	12416
c.2237_2257delinsTCT	p.E746_P753delinsVS	18427
c.2238_2252del15	p.L747_T751del	23571
c.2239_2256delinsCAA	p.L747_S752delinsQ	12403
2240T>C	L747S	26704

Annex 5

Idylla™ EGFR Mutation Test – list of mutations

Exon 18		
Mutation type	**Aminoacid change**	**Nucleotide alteration**
G719A	p.Gly719Ala	c.2156G>C
G719C	p.Gly719Cys	c.2155G>T
		c.2154_2155delinsTT
G719S	p.Gly719Ser	c.2155G>A

Exon 20		
Mutation type	**Aminoacid change**	**Nucleotide alteration**
T790M	p.Thr790Met	c.2369C>T
S768I	p.Ser768Ile	c.2303G>T
InsG	p.Asp770_Asn771insGly	c.2310_2311insGGT
InsASV(9)	p.Val769_Asp770insAlaSerVal	c.2307_2308insGCCAGCGTG
InsASV(11)	p.Val769_Asp770insAlaSerVal	c.2309_2310delinsCCAGCGTGGAT
InsSVD	p.Asp770_Asn771insSerValAsp	c.2311_2312insGCGTGGACA
InsH	p.His773_Val774insHis	c.2319_2320insCAC

Exon 21		
Mutation type	**Aminoacid change**	**Nucleotide alteration**
L858R	p.Leu858Arg	c.2573T>G L858R
		c.2573_2574delinsGT
		c.2573_2574delinsGA
L861Q	p.Leu861Gln	c.2582T>A

Exon 19		
Mutation type	**Aminoacid change**	**Nucleotide alteration**
deletion 9	p.Leu747_Ala750delinsPro	c.2238_2248delinsGC
		c.2239_2248delinsC
	p.Leu747_Ala750delinsSer	c.2240_2248del
	p.Leu747_Glu749del	c.2239_2247del
deletion 12	p.Leu747_Thr751delinsPro	c.2239_2251delinsC
	p.Leu747_Thr751delinsSer	c.2240_2251del
deletion 15	p.Glu746_Ala750del	c.2235_2249del
		c.2236_2250del
	p.Leu747_Thr751del	c.2239_2253del
		c.2240_2254del
		c.2238_2252del
	p.Glu746_Thr751delinsAla	c.2237_2251del
	p.Glu746_Thr751delinsIle	c.2235_2252delinsAAT
	p.Glu746_Thr751delinsVal	c.2237_2252delinsT
	p.Lys745_Ala750delinsThr	c.2234_2248del
	p.Glu746_Thr751delinsLeu	c.2236_2253delinsCTA
	p.Glu746_Thr751delinsVal	c.2237_2253delinsTA
	p.Glu746_Thr751delinsAla	c.2235_2251delinsAG
	p.Glu746_Thr751delinsGln	c.2236_2253delinsCAA
	p.Ile744_Ala750delinsValLys	c.2230_2249delinsGTCAA
deletion 18	p.Leu747_Pro753delinsSer	c.2240_2257del
	p.Glu746_Ser752delinsVal	c.2237_2255delinsT
	p.Leu747_Ser752del	c.2239_2256del
	p.Glu746_Thr751del	c.2236_2253del
	p.Leu747_Pro753delinsGln	c.2239_2258delinsCA
	p.Glu746_Ser752delinsAla	c.2237_2254del
	p.Glu746_Ser752delinsAsp	c.2238_2255del
	p.Glu746_Pro753delinsValSer	c.2237_2257delinsTCT
	p.Glu746_Ser752delinsIle	c.2236_2255delinsAT
		c.2236_2256delinsATC
	p.Glu746_Ser752delinsVal	c.2237_2256delinsTT
		c.2237_2256delinsTC
		c.2235_2255delinsGGT
deletion 21	p.Leu747_Pro753del	c.2238_2258del
	p.Glu746_Ser752del	c.2236_2256del
deletion 24	p.Ser752_Ile759del	c.2253_2276del