UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

PERFIL MOLECULAR DO CÂNCER DE PULMÃO DE NÃO PEQUENAS CÉLULAS EM UM HOSPITAL TERCIÁRIO NO SUL DO BRASIL

Tiago Finger Andreis

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para a obtenção do grau de **Doutor em Genética e Biologia Molecular**

Orientadora: Prof.ª Dr.ª Patricia Ashton-Prolla
Coorientador: Dr. Gabriel de Souza Macedo

Porto Alegre, julho de 2022.

INSTITUIÇÕES FINANCIADORAS: Este trabalho foi desenvolvido no Laboratório de Medicina Genômica (LMG), com apoio do Laboratório de Pesquisa em Resistência Bacteriana (LABRESIS) e da Unidade de Pesquisa Laboratorial, todos localizados no Centro de Pesquisa Experimental do Hospital de Clínicas de Porto Alegre (HCPA). Foi financiado pelo Fundo de Incentivo à Pesquisa e Eventos do HCPA (projetos 2018-0099 e 2018-0121), pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) com a concessão de uma bolsa de doutorado (processo n 140361/2018-9) e com recursos da Chamada Universal Nº 28/2018 – Faixa A (processo n 436348/2018-7) e pelas empresas AstraZeneca, Astellas Pharma, MSD Oncology, Novartis e Genentech.

Dedicatória

Ao meu eterno amigo, Felipe Momberger Machado, que encarou a vida, na saúde e na doença, com muita garra, força e determinação.

Ao meu tio, José Inácio Andreis, que nos aguarda para mais um almoço de domingo.

Aos pacientes, sem os quais este trabalho não seria possível.

Agradecimentos

Primeiramente agradeço à minha orientadora Dra. Patricia Ashton-Prolla, que me acolheu e acreditou no meu trabalho, e ao meu coorientador e amigo Dr. Gabriel de Souza Macedo, fonte de inspiração e que sempre me motivou.

A todo grupo do Laboratório de Medicina Genômica, principalmente Clévia, Bárbara e Igor que muito me ajudaram nos primeiros passos da pesquisa oncológica.

Ao Prof Afonso Luis Barth e equipe do LABRESIS do HCPA, por permitirem o uso do sequenciador Illumina MiSeq, fundamental para desenvolver este trabalho.

Aos agora ex-ICs do laboratório, em especial Julia Tsao Schein, Guilherme Danielski Viola, Bruno Correa e Débora Leite, sem os quais esse trabalho jamais seria realizado.

Aos coautores dos diversos trabalhos que desenvolvemos ao longo desses anos, os quais enriqueceram muito essa jornada, em especial ao ao Dr. Vinicius Lorandi, que ajudou na idealização deste projeto.

A equipe da Engenharia Biomédica do HCPA, que ajudou a viabilizar a extração do cfDNA.

A todo corpo clínico do serviço de oncologia do HCPA, em especial os Drs Marco Azeredo, Gustavo De Agostin, Nícolas Peruzzo e Marcus Teixeira, que participaram ativamente da coleta de dados e rastreamento dos pacientes.

Ao programa de Medicina Personalizada do HCPA e as empresas AstraZeneca, Astellas Pharma, MSD Oncology, Novartis e Genentech que fomentaram a testagem dos biomarcadores desta tese.

A equipe da Igenomix, em especial Gabriela Felix e Rodrigo Borges, que estão auxiliando a validação da técnica utilizada no projeto de biópsia líquida.

A minha namorada e companheira Clévia, que tanto me ajudou e apoiou (e aturou) nessa reta final de doutorado e aos meus pais, Madalena e Tarcísio, que são a base de tudo.

A todo corpo docente do PPGBM/UFRGS e também ao Elmo e

Ao CNPq e HCPA, que fomentaram o desenvolvimento deste doutorado.

Sumário

Lista de abreviaturas

AACR - Associação Americana para a Pesquisa do Câncer (*American Association for Cancer Research*)

ACMG - Colégio Americano de Genética Médica (*American College of Medical Genetics*)

ALK - Gene ALK (*Anaplastic Lymphoma Kinase*)

ALK - Proteína ALK (*Anaplastic Lymphoma Kinase*)

AMP - Associação Americana de Patologia Molecular (*Association for Molecular Pathology*)

ANVISA - Agência Nacional de Vigilância Sanitária

AR - Gene *AR (androgen receptor*)

AREG - Proteína Anfiregulina (*Amphiregulin*)

ARMS - Amplificação por Mutação Refratária (*Amplification-Refractory Mutation System*)

ASCO - Sociedade Americana de Oncologia Clínica (*American Society of Clinical Oncology*)

ATM - Gene *ATM* (*Ataxia Telangiectasia Mutated; ATM Serine/Threonine Kinase*)

B-Raf - Proteína B-Raf (*B-Raf Proto-Oncogene, Serine/Threonine Kinase*)

MET - Gene *MET* (*MET Proto-Oncogene, Receptor Tyrosine Kinase*)

BRAF - Gene *BRAF* (*B-Raf Proto-Oncogene, Serine/Threonine Kinase*)

BRCA1- Gene *BRCA1 (Breast Cancer Type 1 Susceptibility Protein)*

BRCA2 - Gene *BRCA2 (Breast Cancer Type 2 Susceptibility Protein)*

BRIP1 - Gene *BRIP1 (BRCA1-Associated C-Terminal Helicase 1)*

CAAE - Certificado de Apresentação para Apreciação Ética

CAP - Colégio Americano de Patologia (*College of American Pathology*)

CCNE1- gene *CCNE1 (Cyclin E1)*

CDK6 - gene *CDK6 (Cyclin Dependent Kinase 6)*

cfDNA - DNA livre circulante (*circulanting cell-free DNA*)

CHIP - Hematopoiese clonal de potencial indeterminado (*Clonal hematopoiesis of indeterminate potential*)

CLEA - Cromatografia Líquida Desnaturante de Alta Performance *(Liquid Chromatography Clean-up Method*)

CNVs - *Variação no Número de Cópias* (*Copy Number Variation*)

CPCP - Câncer de pulmão de células pequenas
CPNPC - Câncer de pulmão de não-pequenas células
CRKL - Gene *CRKL* (*CRK Like Proto-Oncogene, Adaptor Protein*)
CTC - Células tumorais circulantes
ctDNA - DNA tumoral circulante (*circulating tumor DNA*)
CTNNB1- Gene *CTNNB1 (Catenin Beta 1)*
ddPCR - PCR Digital em Gotas (*Digital Droplet PCR*)
DNA - Ácido desoxirribonucleico (*Deoxyribonucleic acid*)
DRM - Doença residual mínima
EDTA - Ácido etilenodiamino tetra-acético (*Ethylenediamine tetraacetic acid*)
EGF - Fator de crescimento epidermal (*Epidermal growth factor*)
EGFR- Gene *EGFR (Epidermal Growth Factor Receptor)*
EGFR - Proteína EGFR (*Epidermal Growth Factor Receptor*)
EML4 - Gene *EML4 (EMAP Like 4)*
EML4-ALK- Fusão gênica entre os genes *EML4* e *ALK*
EMQN - Rede europeia de qualidade para genética molecular (*European Molecular Genetics Quality Network*)
ERBB2 - Gene *ERBB2 (Erb-B2 Receptor Tyrosine Kinase 2)*
ERBB3 - Gene *ERBB3 (Erb-B2 Receptor Tyrosine Kinase 3)*
FDA - *Food and Drug Administration*
FGFR1- Gene *FGFR1 (Fibroblast Growth Factor Receptor 1)*
FGFR3 - Gene *FGFR3 (Fibroblast Growth Factor Receptor 3)*
gDNA - DNA genômico
GENIE - Projeto *The AACR Project Genomics Evidence Neoplasia Information Exchange*
GLTP - Proteína de transferência de glicolipídios (*Glycolipid transfer protein*)
HCPA - Hospital de Clínicas de Porto Alegre
HER2 - receptor HER2 (*Erb-B2 Receptor Tyrosine Kinase 2*)
HER4 - receptor HER4 (*Erb-B2 Receptor Tyrosine Kinase 4*)
INCA - Instituto Nacional de Câncer José Alencar Gomes da Silva
Indels Pequenas Inserções e Deleções
ITQ - Inibidor de tirosina quinase

JAK/STAT - *Janus Kinase/Signal Transducer and Activator of Transcription Protein Family*

JAK2 - Gene *JAK2 (Janus quinase 2)*

KEAP1 - Gene *KEAP1 (Kelch Like ECH Associated Protein 1)*

KIT - Gene *KIT (KIT Proto-Oncogene, Receptor Tyrosine Kinase)*

KRAS - Gene *KRAS (Kirsten Rat Sarcoma Viral Oncogene Homolog)*

KRAS - Proteínas KRAS (*Kirsten Rat Sarcoma Viral Oncogene Homolog*)

KRT7 - Queratina 7 (*Keratin 7*)

MEK - Proteína quinase ativada por mitógeno (*mitogen-activated protein kinase kinase*)

MEK1 - Gene *MEK1 (mitogen-activated protein kinase kinase 1)*

MYC - Gene *MYC (MYC Proto-Oncogene, BHLH Transcription Factor)*

MYC - Gene *MYC (MYC Proto-Oncogene, BHLH Transcription Factor)*

NCCN - *National Comprehensive Cancer Network*

NGS - Sequenciamento de Nova Geração (*Next Generation Sequencing*)

NRAS - Gene *NRAS* (*NRAS Proto-Oncogene, GTPase*)

NTRK1/2/3 - Genes *NTRK1, NTRK2 e NTRK3 (Neurotrophic Receptor Tyrosine Kinase 1/2/3)*

p63 - Proteína tumoral p63 (*Tumor protein p63*)

PD-L1- Ligante de morte celular programada 1 (*Programmed death-ligand 1*)

Pi3K/AKT/mTOR - *Phosphoinositide 3-kinase/Protein Kinase B/Mechanistic Target of Rapamycin*

P*IK3CA* - gene *PIK3CA (Phosphoinositide 3-kinase)*

RAS - vírus do sarcoma de rato (*Rat Sarcoma virus*)

RAS/MAPK - *Rat Sarcoma/Mitogen-activated Protein Kinases*

RET - Gene *RET (Ret Proto-Oncogene)*

ROS1- Gene *ROS1 (ROS Proto-Oncogene 1, Receptor Tyrosine Kinase)*

RT-PCR - Reação em Cadeia de Polimerase em Tempo Real (*Real Time Polymerase Chain Reaction*)

RTQs - Receptores de Tirosina Quinase

SEER - Programa de Vigilância, Epidemiologia e Resultados Finais

(The Surveillance, Epidemiology, and End Results)
SFTPA1 - Gene *SFTPA1 Surfactant Protein A1*
SFTPA2 - Gene *SFTPA2 Surfactant Protein A2*
SNVs - Variantes de Base Única (*Single Nucleotide Variant*)
SOX2 - fator de transcrição Sox2 (*SRY-Box Transcription Factor 2*)
STK11 - Gene *STK11 (Serine/Threonine Kinase 11)*
SUS - Sistema Único de Saúde
TBP - Câncer de traqueia, brônquio e pulmão
TCGA - Projeto Atlas Genômico do Câncer (*The Cancer Genome Atlas Project*)
TCLE - Termo de consentimento livre e esclarecido
TNM - Classificação de tumores malignos TNM (*TNM Classification of Malignant Tumors*)
TRK - receptores de tropomiosina relacionados à quinase (*Tropomyosin receptor kinase*)
TTF1 - Fator de transcrição de tireoide 1 (*Thyroid transcription factor 1*)
UMI - Identificadores moleculares únicos (Unique molecular identifiers)
VPN - Valor preditivo negativo
VPP - Valor preditivo positivo
WHO - Organização Mundial da Saúde (World Health Organization)

Lista de figuras

Lista de quadros

Resumo

O Câncer de Pulmão é a neoplasia que mais causa óbitos no mundo. O número de mortes relacionadas à doença é três vezes maior em homens se comparado ao câncer de próstata e quase duas vezes maior em mulheres quando comparado ao câncer de mama. Novas opções terapêuticas têm surgido nos últimos anos, incluindo drogas de alvo molecular indicadas quando há alterações acionáveis nos genes *EGFR*, *KRAS*, *BRAF*, *ALK*, *ROS1, RET, ERBB2* e *MET*. Tumores que apresentam tais alterações nesses genes podem responder a inibidores de tirosina quinase (ITQs). Estes apresentam uma menor toxicidade e resultam em uma sobrevida livre de progressão e global superior à quimio e radioterapia convencionais, desde que os pacientes sejam acuradamente selecionados por meio do diagnóstico molecular. Diversos estudos ao redor do mundo já demonstraram que as frequências dos biomarcadores de resposta a ITQs variam grandemente entre as diferentes populações, porém poucos trabalhos exploraram o tema no Brasil.

Este trabalho foi o primeiro a descrever a frequência de importantes biomarcadores de resposta a tratamento em uma série de pacientes com câncer de pulmão de não pequenas células (CPNPC) da região Sul do Brasil. Os dados revelaram que o gene *EGFR* está alterado em 19,1% dos casos, a menor frequência já relatada no país. Mutações em *KRAS*, por outro lado, estavam presentes em 30,2% dos casos, a maior até então descrita no Brasil. Também foi observado que aproximadamente 6% dos tumores investigados possuíam alterações raras de significado clínico incerto, cuja resposta a ITQs é ainda desconhecida. Uma vez que estudos investigando a resposta terapêutica de pacientes com tumores apresentando tal perfil molecular são escassos, o presente trabalho também busca avaliar a escolha dos profissionais médicos no tratamento deste grupo de pacientes na ausência de diretrizes específicas para conduta terapêutica e, para aqueles que optaram pelo uso de ITQs, o perfil de resposta observado após uso do medicamento.

O acesso ao diagnóstico molecular e diferenças no perfil molecular de pacientes atendidos no sistema público e particular/suplementar de saúde também é um tema importante. Estima-se que menos de 40% dos pacientes brasileiros com

CPNPC tenham acesso ao teste do gene *EGFR*, situação ainda mais precária para os demais biomarcadores e pacientes atendidos pelo sistema público. Nesse sentido, o estudo está avaliando os registros de dois hospitais terciários na cidade de Porto Alegre, a fim de coletar informações que possam fomentar futuras decisões quanto a políticas públicas de diagnóstico e tratamento do CPNPC.

Por fim, uma vez que o CPNPC é diagnosticado em estado avançado em mais de 70% dos casos, torna-se importante a implementação de técnicas alternativas para o diagnóstico molecular da doença. Em tais casos, é relativamente comum que o tecido tumoral não esteja disponível devido à invasividade da biópsia convencional ou inacessibilidade do tumor. Sendo assim, a biópsia líquida pode permitir a investigação do DNA tumoral circulante (ctDNA) e é uma ferramenta diagnóstica de grande valor, auxiliando na exploração da heterogeneidade tumoral e na busca de biomarcadores preditivos e/ou prognósticos. Pelo nosso conhecimento, este foi o primeiro estudo a analisar alterações moleculares no DNA livre circulante (cfDNA) isolado de uma série de pacientes diagnosticados com CPNPC atendidos pelo sistema público de saúde. Resultados preliminares demonstraram que foi possível acompanhar a variação alélica de duas mutações somáticas em um dos pacientes, corroborando com a evolução clínica do mesmo. Espera-se que este trabalho, assim que concluído, torne a medicina de precisão mais próxima dos pacientes diagnosticados com CPNPC no sistema público do país, auxiliando na tomada de decisões em políticas públicas aplicadas ao diagnóstico molecular.

Abstract

Lung cancer is the leading cause of cancer death in the world. The disease-related deaths are three times higher than prostate cancer in men and almost twice as high as breast cancer in women. New therapeutic options have emerged in recent years, including molecular target drugs which are indicated when actionable alterations in the genes *EGFR, KRAS, BRAF, ALK, ROS1, RET, ERBB2*, and *MET* are present. Lung tumors carrying mutations in such genes could respond to tyrosine kinase inhibitors (TKIs). These drugs show a lower toxicity and result in a superior progression-free and overall survival comparing to conventional chemotherapy and radiotherapy approaches when patients are properly selected through molecular diagnosis. Several studies around the world have already demonstrated that the frequency of biomarkers that indicate response to TKIs varies greatly among different populations, but few works have explored this topic in Brazil.

This is the first work describing the frequency of important treatment response biomarkers in a series of patients with non-small cell lung cancer (NSCLC) from Southern Brazil. The results revealed that the *EGFR* gene is altered in 19.1% of cases, the lowest frequency reported in the country. On the other hand, *KRAS* mutations were present in 30.2% of the cases which the highest described in the country. It was also observed that approximately 6% of the investigated patients carried rare alterations of uncertain clinical significance, which confer an unknown response to TKIs. Studies investigating the therapeutic response in patients with tumors presenting this molecular profile are also scarce. Therefore, in the absence of specific guidelines for therapeutic management, the present work is also evaluating the treatment choices recommended by medical professionals for a group of patients with alterations of uncertain significance.

The access to molecular diagnosis and differences in the molecular profile of patients treated in the public and private/supplementary health system is another important issue. It is estimated that less than 40% of Brazilian patients with NSCLC have access to basic results like *EGFR* mutational status. The access to other biomarker testing and for patients treated by the public system is even limited. In this

sense, we are collecting data from two tertiary hospitals in the city of Porto Alegre, Rio Grande do Sul, Brazil, in order to gather data that could guide future decisions regarding public policies for NSCLC diagnosis and treatment.

Finally, since NSCLC in Brazil is diagnosed at an advanced stage in more than 70% of the cases, it is important to implement alternative techniques for molecular diagnosis. In the advanced cases tumor tissue is frequently unavailable due to the invasiveness of conventional biopsy and/or tumor inaccessibility. The liquid biopsy allows the investigation of the circulating tumor DNA (ctDNA). It is an alternative tool of great value which can also be used to explore tumor heterogeneity and aid in the searching for predictive and/or prognostic biomarkers. To our knowledge, this was the first study to access molecular alternations in the cell-free DNA (cfDNA) isolated from a series of patients diagnosed with NSCLC treated in the public health system. Our preliminary results showed that it was possible to evaluate the allelic variation of two somatic mutations identified in the cfDNA from one patient, which corroborates with his clinical evolution. Once completed, we hope that our work could bring precision medicine closer to patients diagnosed with NSCLC in Brazil by aiding public policies in decision making applied to molecular diagnosis.

CAPÍTULO I – INTRODUÇÃO GERAL

1.1. Câncer de pulmão: epidemiologia e fatores de risco

O câncer de pulmão apresenta a maior letalidade dentre todos os tumores, sendo a sexta causa de óbitos no mundo, resultando em quase 2 milhões de mortes pela doença a cada ano (Sung *et al.*, 2021). É a neoplasia que mais afeta homens, com quase 1,5 milhão dos diagnósticos, e a terceira em mulheres, com 695 mil a cada ano. Estima-se um crescimento de 37% no número de diagnósticos de câncer de pulmão em um período de apenas 10 anos (2007– 2017), tendência que deve se manter nos próximos anos (Fitzmaurice *et al.*, 2019). Mesmo com avanços na descoberta de alvos terapêuticos e novos tratamentos direcionados, nos últimos anos a taxa de sobrevida em 5 anos foi de apenas 22,9% para pacientes diagnosticados com neoplasias de brônquios e pulmão (SEER, 2019). Atualmente, espera-se que um a cada 17 homens e uma a cada 43 mulheres irão desenvolver a doença antes dos 79 anos. Por fim, atualmente o câncer de pulmão é a principal causa de morte em 110 países (Fitzmaurice *et al.*, 2019).

No Brasil, em 2020 foram estimados 30.200 novos casos de câncer de traqueia, brônquio e pulmão (TBP), sendo 17.760 em homens (7,9% do total de casos de câncer) e 12.440 em mulheres (5,6% do total de casos de câncer). Em relação a mortalidade, ocupa o primeiro lugar em homens e o segundo em mulheres, ficando atrás somente do câncer de mama. No Rio Grande do Sul, em 2020 foram estimados 2.300 e 1.440 novos casos da doença em homens e mulheres, respectivamente, o que torna o câncer de pulmão o segundo tumor mais frequente em homens e o terceiro em mulheres no Estado (Figura 1; (Brasil, 2019).

A

Localização Primária	Casos	%
Próstata	65.840	29,2%
Cólon e reto	20.520	9,1%
Traqueia, brônquio e pulmão	17.760	7,9%
Estômago	13.360	5,9%
Cavidade oral	11.180	5,0%
Esôfago	8.690	3,9%
Bexiga	7.590	3,4%
Linfoma não Hodgkin	6.580	2,9%
Laringe	6.470	2,9%
Leucemias	5.920	2,6%

Homens

Mulheres

Localização Primária	Casos	%
Mama feminina	66.280	29,7%
Cólon e reto	20.470	9,2%
Colo do útero	16.590	7,4%
Traqueia,brônquio e pulmão	12.440	5,6%
Glândula tireoide	11.950	5,4%
Estômago	7.870	3,5%
Ovário	6.650	3,0%
Corpo do útero	6.540	2,9%
Linfoma não Hodgkin	5.450	2,4%
Sistema nervoso central	5.220	2,3%

B

- Em homens, Brasil, 2020

Localização Primária	Óbitos	%
Traqueia, Brônquios e Pulmões	16.009	13,6
Próstata	15.841	13,5
Cólon e Reto	9.889	8,4
Estômago	8.772	7,5
Esôfago	6.465	5,5
Fígado e Vias biliares intrahepáticas	6.093	5,2
Pâncreas	5.882	5,0
Sistema Nervoso Central	4.787	4,1
Cavidade oral	4.767	4,1
Laringe	3.896	3,3
Todas as neoplasias	**117.512**	**100,0**

- Em mulheres, Brasil, 2020

Localização Primária	Óbitos	%
Mama	17.825	16,5
Traqueia, Brônquios e Pulmões	12.609	11,6
Cólon e Reto	10.356	9,6
Colo do útero	6.627	6,1
Pâncreas	6.011	5,5
Estômago	5.078	4,7
Fígado e Vias biliares intrahepáticas	4.670	4,3
Sistema Nervoso Central	4.567	4,2
Ovário	3.921	3,6
Leucemias	3.035	2,8
Todas neoplasias	**108.318**	**100,0**

Figura 1. Incidência (A) e óbitos (B) por câncer no Brasil. Apesar das neoplasias de traqueia, brônquios e pulmões não liderarem em número de casos, são a principal causa de morte por câncer no país. Figura adaptada do Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA), estimativa de 2020 (https://www.inca.gov.br/).

Entre os fatores de risco ligados ao câncer de pulmão, o principal e mais bem estabelecido é o hábito tabagista, associado com aproximadamente 90% dos casos (Ozlü and Bülbül, 2005). Acredita-se que tabagistas apresentam de 20 a 30 vezes mais risco de desenvolver a neoplasia em comparação com indivíduos que nunca fumaram (Ridge, McErlean and Ginsberg, 2013). Diversas substâncias presentes na fumaça do cigarro, incluindo hidrocarbonetos policíclicos aromáticos (HPAs) e aldeídos, produzem adutos no DNA, induzindo mutações. Quando tais alterações acometem genes supressores de tumor, a exemplo do *TP53*, ou oncogenes, como *KRAS*, pode iniciar-se então o processo de carcinogênese, culminando na formação de um tumor

(Hecht, 2012). O advento dos cigarros eletrônicos traz novos desafios para a epidemiologia do câncer de pulmão. Apesar destes também apresentarem substâncias potencialmente carcinogênicas, como nitrosaminas, e induzirem danos no DNA em modelos *in*-vivo, seus efeitos a longo prazo ainda são desconhecidos e podem ditar novos rumos para a prevenção e tratamento da doença (Bracken-Clarke *et al.*, 2021; Lee *et al.*, 2018).

Outros fatores de risco incluem exposição ao cromo, arsênico, cádmio, sílica, níquel, poluentes atmosféricos e até mesmo fatores ligados à dieta (Sun, Schiller and Gazdar, 2007). Alguns estudos também demonstram uma maior incidência da doença em pacientes com repetidas infecções pulmonares, histórico de tuberculose, além de deficiência e/ou excesso de vitamina A (Dela Cruz, Tanoue and Matthay, 2011). Fatores hereditários que aumentam o risco no desenvolvimento da doença também devem ser levados em conta, incluindo indivíduos acometidos pela síndrome de Li-Fraumeni e aqueles com variantes germinativas no gene *EGFR.* Outros genes candidatos ainda carecem de maiores investigações, incluindo pacientes com alterações germinativas em *ATM*, *BRCA2, SFTPA1* e *SFTPA2* (Benusiglio *et al.*, 2021).

O câncer de pulmão tem um grande custo para os cofres públicos do Brasil, onde 80% da população depende do Sistema Único de Saúde (SUS), com gastos anuais que chegam próximos dos R$30 milhões quando levados em conta os custos diretos no tratamento da doença. Se levado em conta o sistema de saúde suplementar, as cifras ultrapassam os R$140 milhões (Ferreira *et al.*, 2016). Estimativas atribuem mais de 70% desses gastos ao atendimento ambulatorial, principalmente aos regimes de quimio e radioterapia (Knust *et al.*, 2017). Tais gastos refletem o estágio tardio em que a neoplasia é diagnosticada no Brasil, uma vez que 70,6% dos casos se apresentam já como doença avançada, quando intervenções cirúrgicas são pouco eficazes (Ismael *et al.*, 2011). Fármacos de uso oral, como os inibidores de tirosina quinase (ITQs) que serão abordados adiante, poderiam reduzir esses gastos em até 13% (Ferreira *et al.*, 2016). Estudos mais atuais são necessários, uma vez que o cenário de atenção oncológica mudou drasticamente nos últimos anos, seja com

situações que aumentam a proporção de diagnósticos com estádio avançado (pandemia da COVID-19), seja com situações que poderiam modificar custos do tratamento sistêmico (p.ex. o advento de novas drogas para o tratamento das neoplasias de pulmão). O impacto econômico da pandemia na pesquisa e desenvolvimento para o tratamento da doença foi significativo, atrasando ou até mesmo impedindo a realização de novos estudos, situação mais agravante em economias emergentes como o Brasil (Mukherji *et al.*, 2021).

1.2. Classificação histopatológica do câncer de pulmão

Atualmente o câncer de pulmão é classificado em dois grandes grupos: câncer de pulmão de células pequenas (CPCP), representando em torno de 15% dos diagnósticos; e câncer de pulmão de não-pequenas células (CPNPC), presente em aproximadamente 85% dos casos (Lortet-Tieulent *et al.*, 2014; Pirker and Filipits, 2016; Gridelli *et al.*, 2015) (Figura 2). Nas últimas décadas, a fim de auxiliar na decisão terapêutica, uma subclassificação foi incorporada para CPNPC, incluindo os subtipos: adenocarcinoma, câncer de pulmão de células escamosas e câncer de pulmão de células grandes (Johnson *et al.*, 2004; Scagliotti *et al.*, 2008).

Os adenocarcinomas de pulmão representam pelo menos 70% dos tumores não-escamosos e em torno de 50% de todos os diagnósticos de câncer de pulmão (Lortet-Tieulent *et al.*, 2014; Gridelli *et al.*, 2015). Estes tumores têm origem em células glandulares, são mais frequentemente diagnosticados em mulheres (~45% versus ~35% em homens) e tendem a apresentar um comportamento mais indolente quando comparado a outros subtipos. Expressam frequentemente biomarcadores consistentes com sua origem pulmonar, incluindo o fator de transcrição de tireoide 1 (TTF1, do inglês *Thyroid Transcription Factor 1*) e queratina 7 (KRT7, do inglês *Keratin 7*). Além disso, este é o tipo de câncer de pulmão mais comum em não tabagistas (Varughese *et al.*, 2012; Davidson, Gazdar and Clarke, 2013).

Já o carcinoma de células escamosas, também conhecido como carcinoma epidermóide, corresponde a aproximadamente 30% dos diagnósticos de CPNPC(Lortet-Tieulent *et al.*, 2014) (Figura 2). São mais frequentemente observados

em fumantes do que qualquer outro subtipo de CPNPC (Heist, Sequist and Engelman, 2012). Outros fatores de risco importantes para este tipo de neoplasia incluem idade, história familiar e exposição a diferentes carcinógenos. Os carcinomas de células escamosas distinguem-se dos adenocarcinomas pela expressão de citoqueratina 5/6, SOX2 e p63 (Lu *et al.*, 2010).

Por fim, representando em torno de 10% dos casos de câncer de pulmão não-escamosos, está o câncer de pulmão de células grandes (Lortet-Tieulent *et al.*, 2014) (Figura 2). A frequência deste subtipo tem caído significativamente, desde que muitos casos estão sendo reclassificados como adenocarcinomas ou carcinomas de células escamosas (Davidson, Gazdar and Clarke, 2013).

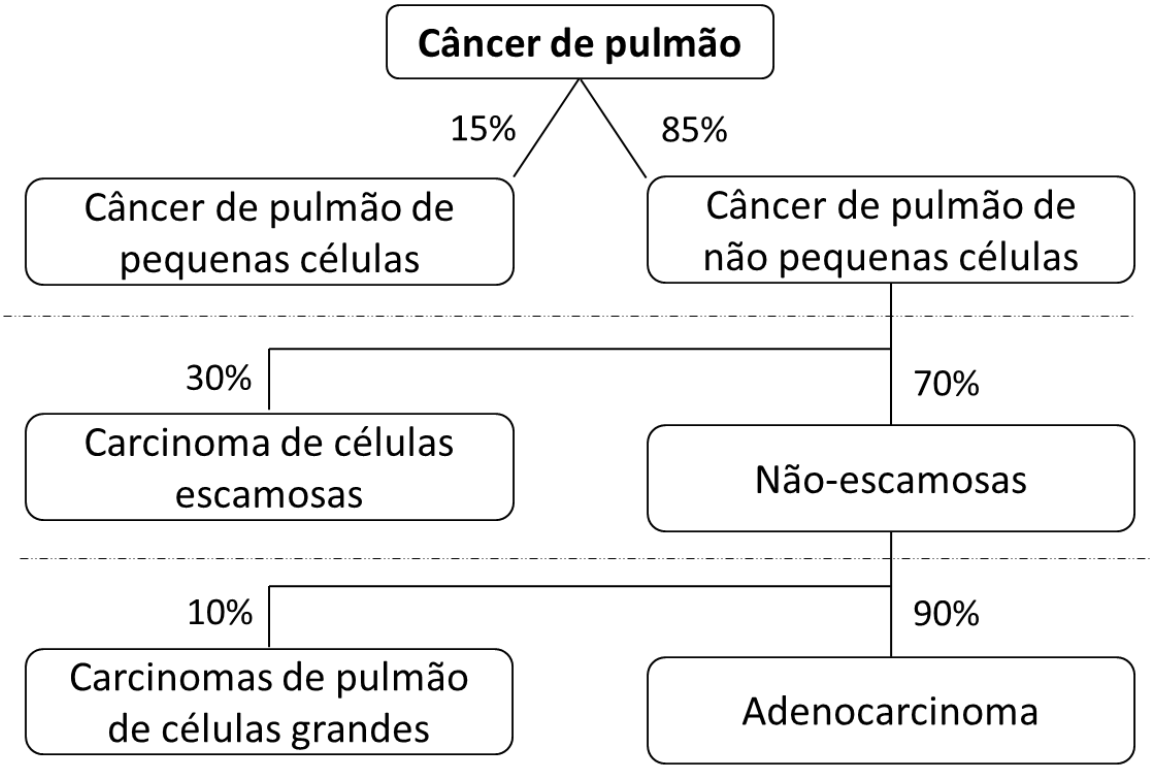


Figura 2. Classificação histológica do câncer de pulmão e suas respectivas frequências. Figura adaptada de Gridelli *et al.* (2015).

No Brasil, o diagnóstico do câncer de pulmão normalmente se dá por exames de imagem simples, como a radiografia de tórax. Esta, no entanto, possui sensibilidade

de apenas 70% e especificidade de 89% a 99%. Técnicas alternativas, como tomografía axial computadorizada de tórax e broncoscopia possuem melhor desempenho no diagnóstico, possibilitando também uma melhor definição do estadiamento do tumor (INCA, 2021 - https://www.inca.gov.br/tipos-de-cancer/cancer-de-pulmao/profissional-de-saude).

Quanto ao estadiamento, que basicamente indica o grau de invasividade do tumor, normalmente é definido pelo sistema TNM da *American Joint Committee on Cancer*. A letra "T" diz respeito às características morfológicas do tumor, como seu tamanho e a sua dispersão nos tecidos adjacentes. Já a letra "N" diz respeito à presença ou não de metástases nos linfonodos e, por fim, a "M" trata da distância das metástases em relação ao sítio primário do tumor. Com base nesses critérios, os tumores são classificados de I a IV. Tumores em estadiamento I são localizados e normalmente passíveis de cirurgia curativa. Já tumores em estadiamento IV são avançados por apresentarem metástase a distância em relação ao sítio primário de diagnóstico. Nesses casos, abordagens cirúrgicas são pouco eficazes, e tratamentos sistêmicos com radio e quimioterapia, além de terapias com drogas de alvo-molecular, são frequentemente indicados (Goldstraw *et al.*, 2016; Brierley, Gospodarowicz and Wittekind, 2017) (Quadro 1)

Quadro 1. Resumo da classificação de tumores malignos utilizada para definição do estadiamento de neoplasias de pulmão, com base nos critérios da *American Joint Committee on Cancer*, 8ª edição.

	N0 [a]	**N1** [b]	**N2** [c]	**N3** [d]
T1a – tumor ≤ 1 cm	IA1	IIB	IIIA	IIIB
T1b – tumor <1 cm, mas ≤ 2 cm	IA2	IIA	IIIA	IIIB
T1c – tumor > 2 cm, ≤ 3cm	IA3	IIA	IIIA	IIIB
T2a – tumor > 3 cm, mas ≤ 4 cm	IB	IIB	IIIA	IIIB
T2b – tumor > 4, mas ≤ 5 cm	IIA	IIB	IIIA	IIIB
T3 – tumor > 5 cm, mas ≤ 7 cm ou que invade determinada estrutura [e]	IIB	IIIA	IIIB	IIIC
T3 – tumor > 7 cm	IIIA	IIIA	IIIB	IIIC
T3 – invade o diafragma	IIIA	IIIA	IIIB	IIIC
T3 – tumor endobrônquico com atelectasia, > 3 cm, mas ≤ 4 cm	IB	IIB	IIIB	IIIB
T3 – tumor endobrônquico com atelectasia, > 4 cm, mas ≤ 5	IIA	IIB	IIIA	IIIB
T4 – tumor > 7 cm ou associado a nódulo(s) tumoral(is) separado(s) no mesmo lobo ipsilateral do tumor primário, ou que invade determinada estrutura [f]	IIIA	IIIA	IIIB	IIIC
M1a [h]	IVA	IVA	IVA	IVA
M1b [i] **com uma lesão**	IVA	IVA	IVA	IVA
M1c [i] **com múltiplas lesões**	IVB	IVB	IVB	IVB

a = Sem metástase em linfonodo regional
b = Metástase em linfonodo peribrônquico ipsilateral e/ou hilar ipsilateral e intrapulmonares, incluindo envolvimento por extensão direta.
c = Metástase em linfonodo mediastinal ipsilateral e/ou subcarinal
d = Metástase em linfonodo mediastinal contralateral; hilar contralateral; escalênico ipsilateral ou contralateral; ou supraclavicular.
e = parede do peito, incluindo pleura parietal e tumores de sulco; nervo frênico; pericárdio parietal.
f = parede torácica, diafragma, nervo frênico, pleural mediastinal, pericárdio
h = Nódulo (s) tumoral (ais) em lobo contralateral; tumor com nódulo pleural ou derrame pleural ou pericárdico maligno.
i = metástase a distância.

1.3. Diagnóstico molecular

Nas últimas décadas, avanços significativos foram alcançados no tratamento das neoplasias de pulmão. A maioria deles se deve à maior compreensão do perfil genômico de mutações somáticas do tumor, ou seja, daquelas alterações adquiridas ao longo do processo de oncogênese. Estas alterações, por sua vez, são classificadas em mutações direcionadoras/condutoras (*drivers*), as quais possuem papel fundamental no processo de carcinogênese, e mutações passageiras (*passengers*), as quais não possuem papel relevante na manutenção e progressão do tumor (Haber and Settleman, 2007).

Os avanços mais recentes para compreensão da genômica tumoral foram impulsionados pelo advento do sequenciamento massivo paralelo, também conhecido como Sequenciamento de Nova Geração (NGS, do inglês *Next Generation Sequencing*). A redução de custos e consequente maior acessibilidade à técnica permitiu o surgimento de grandes consórcios internacionais, como, por exemplo, o *The Cancer Genome Atlas* (TCGA) e o *The AACR Project Genomics Evidence Neoplasia Information Exchange* (GENIE), os quais mapearam alterações no código genético das neoplasias e permitiram que a classificação dos tumores fosse além de sua histologia, agrupando-os de maneira ainda mais precisa conforme suas assinaturas moleculares (Swami *et al.*, 2021; Alexandrov and Stratton, 2014; Network, 2014; Consortium, 2017; Weinstein *et al.*, 2013). Essas assinaturas atuam como marcadores moleculares, denominados biomarcadores, os quais podem auxiliar em um diagnóstico, prognóstico e tratamento da doença mais personalizado e individualizado, conhecido como oncologia de precisão (Ciardiello *et al.*, 2014; Yang *et al.*, 2020b).

1.4. Tratamento do câncer de pulmão com drogas de alvo molecular e inibidores de *checkpoint* imunológico

Como consequência do maior conhecimento do perfil genômico dos tumores, diversas drogas de alvo molecular vêm sendo desenvolvidas em ritmo acelerado nos últimos anos. Estas atuam de maneira mais específica sobre determinados receptores, ou ainda proteínas intracelulares, que estão alterados em decorrência das mutações

encontradas no tumor (Malone *et al.*, 2020). Tais avanços não somente levaram a um aumento da sobrevida dos pacientes, mas também em ganhos significativos na qualidade de vida durante o tratamento, graças a diminuição da toxicidade e eventos adversos associados, comuns durante regimes terapêuticos convencionais à base de quimioterápicos (Ai *et al.*, 2018).

O sucesso no uso de drogas de alvo molecular depende diretamente da correta seleção dos pacientes que potencialmente podem se beneficiar das novas opções terapêuticas. Diversos aspectos devem ser observados, incluindo a heterogeneidade molecular dos tumores de pulmão, mesmo entre aqueles de mesmo padrão histológico ((CLCGP) and (NGM), 2013; Kandoth *et al.*, 2013). Em outras palavras, tumores que são fenotipicamente semelhantes podem apresentar mutações muito distintas. Portanto, no contexto atual do manejo do CPNPC, o uso de técnicas de biologia molecular para correta e precisa identificação de biomarcadores de predição terapêutica é imprescindível na tomada de decisões, de maneira que estas sejam baseadas no perfil molecular do tumor, prevendo a provável resposta ao tratamento (Politi and Herbst, 2015).

Uma vez identificadas, as mutações, formalmente descritas como variantes somáticas, devem ser corretamente classificadas. Da mesma forma que o *American College of Medical Genetics* (ACMG) fornece diretrizes para a classificação das variantes de origem germinativa em cinco categorias (Patogênica, Provavelmente Patogênica, Variante de Significado Indeterminado, Benigna e Provavelmente Benigna), o padrão ouro para classificação de variantes somáticas segue as recomendações da *Association for Molecular Pathology, American Society of Clinical Oncology* e *College of American Pathologists* (AMP/ASCO/CAP). Atualmente seis níveis (em inglês, *tiers*) de classificação são utilizados, indicando o grau de evidências que suportam se determinada mutação tem significado terapêutico (sensibilidade ou resistência à determinada droga), prognóstico e/ou diagnóstico, conforme Quadro 2 abaixo (Li *et al.*, 2017).

Quadro 2. Níveis (*tiers*) de evidência para classificação de mutações somáticas, segundo recomendações da AMP/ASCO/CAP.

Nível de evidência	Descrição
Tier IA	Mutações com significado clínico comprovado e/ou com protocolos clínicos e diretrizes terapêuticas aprovados por agências de saúde.
Tier IB	Mutações com nível de evidências robusto, incluindo consenso de especialistas na área.
Tier IIC	Significado clínico comprovado para outros subtipos tumorais, ou quando a mutação é citada em terapias investigacionais, ou quando diversos estudos de menor impacto em consenso indicando o impacto da mutação.
Tier IID	Estudos pré-clínicos e/ou relatos de caso indicando o significado clínico da mutação.
Tier III	Mutações ausentes ou ocorrendo em baixa frequência em bancos de dados populacionais, porém sem evidências consistentes na literatura quanto ao seu impacto funcional.
Tier IV	Variantes benignas ou provavelmente benignas, observadas em frequência relativamente alta em bancos de dados populacionais (>1%) e sem evidências de associação com o processo de oncogênese.

Exemplos já bem difundidos na rotina clínica são o rastreamento de mutações no oncogene *EGFR* (do inglês *Epidermal Growth Factor Receptor*) e rearranjos em *ALK* (do inglês *Anaplastic Lymphoma Kinase*), utilizados como biomarcadores de resposta desde os anos de 2004 e 2007, respectivamente (Lynch *et al.*, 2004; Soda *et al.*, 2007). Somam-se também a esses marcadores as mutações em *KRAS* (do inglês *Kirsten Rat Sarcoma Viral Oncogene Homolog*) e *BRAF* (do inglês *B-Raf Proto-Oncogene, Serine/Threonine Kinase*), alterações em *MET* (do inglês *MET Proto-Oncogene, Receptor Tyrosine Kinase*), que incluem amplificações, exclusão do éxon 14 e mutações no códon D1010, além de fusões gênicas envolvendo *ROS1* (do inglês *ROS Proto-Oncogene 1, Receptor Tyrosine Kinase*), *RET* (do inglês *Ret Proto-Oncogene*) e *NTRK1/2/3* (do inglês *Neurotrophic Receptor Tyrosine Kinase 1/2/3*) (Imyanitov, Iyevleva and Levchenko, 2021). Um resumo dos fármacos utilizados nas terapias de alvo molecular e seus respectivos marcadores são apresentados no Quadro 3. Atualmente apenas dois estão aprovados para seu uso no SUS no

tratamento do CPNPC. São eles: o gefitinibe, aprovado pela Comissão Nacional de Incorporação de Tecnologias no Sistema Único de Saúde em 2013; e o erlotinibe, aprovado em 2015, quase 10 anos após sua aprovação pela Agência Nacional de Vigilância Sanitária (CONITEC, 2013; CONITEC, 2015).

Quadro 3. Linha do tempo com as aprovações de drogas de alvo molecular para o tratamento de CPNPC no mundo e no Brasil com seus respectivos biomarcadores de resposta terapêutica (alvos).

Alvo(s)	Fármaco(s)	Ano de aprovação	
		Mundo	Brasil
EGFR del19, L858R	Gefitinibe	2003	2011
EGFR del19, L858R	Erlotinibe	2004	2006
Fusões *ALK*, *MET* ex14, D1010X	Crizotinibe	2011	2016
EGFR del19, L858R, G719, S768I, L861Q	Afatinibe	2013	2016
Fusões *ALK*	Alectinibe Ceritinibe	2014	2019 NA*
EGFR del19, L858R, T790M	Osimertinibe	2015	2017
Fusões *ROS1*	Crizotinibe	2016	2018
*BRA*F V600E	Dabrafenibe + Trametinibe	2017	2018
Fusões *ALK*	Brigatinibe	2017	2019
Fusões *ALK*, *ALK* G1202R	Lorlatinibe	2018	2020
EGFR del19, L858R	Dacomitinibe	2018	NA
Fusões *NTRK*	Larotrectinibe	2018	2019
Fusões *NTRK* e *ROS1*	Entrectinibe	2019	NA
MET ex14, D1010X	Capmatinibe	2020	2021
	Tepotinibe		2021
Fusões *RET*	Pralsetinibe	2020	NA
	Selpercatinibe		NA
EGFR ins20	Amivantamabe	2021	2021
	Mobocertinibe		NA
KRAS G12C	Sotorasibe	2021	2022

* NA = Não aprovado.

Todos os fármacos citados acima envolvem direta ou indiretamente os Receptores de Tirosina Quinase (RTQs), que compõem um dos principais elementos reguladores da sobrevivência e proliferação celular. A estrutura básica dos RTQs consiste em um sítio de ligação extracelular, uma α-hélice transmembrana e um domínio tirosina quinase no citoplasma. Em humanos, atualmente são classificados em 20 famílias. A ativação destes envolve ligantes que mediam a conexão entre os receptores monoméricos ou oligoméricos, formando dímeros/oligômeros, os quais ativam, mediante fosforilação no meio intracelular, os fatores de transcrição (Lemmon and Schlessinger, 2010; Pottier *et al.*, 2020). Uma vez que um respectivo ligante adere a um RTQ, este sofre autofosforilação, desencadeando uma cascata de sinalização que envolve diversas vias intermediárias a jusante, incluindo: Pi3K/AKT/mTOR (do inglês *Phosphoinositide 3-kinase/Protein Kinase B/Mechanistic Target of Rapamycin*), a qual controla o crescimento celular, metabolismo, sobrevivência celular e manutenção da pluripotência; RAS/MAPK (do inglês, *Rat Sarcoma/Mitogen-activated Protein Kinases*), via reguladora do metabolismo, ciclo celular, proliferação, diferenciação e migração celular; e JAK/STAT (do inglês, *Janus Kinase/Signal Transducer and Activator of Transcription Protein Family*), um fator de transcrição citoplasmático que media citocinas e sinalização de fatores de crescimento (Liang and Yang, 2019; Santoni-Rugiu *et al.*, 2019; Pottier *et al.*, 2020) (Figura 2).

Alterações moleculares em genes que codificam RTQs e outras proteínas de nas vias associadas são um dos principais mecanismos de oncogênese em diversos tipos de tumores, incluindo CPNPC (Porter and Vaillancourt, 1998; Du and Lovly, 2018) (Figura 2). É na sequência que codifica o domínio tirosina quinase, presente no citoplasma, que muitas das mutações oncogênicas são encontradas. Neste domínio atuam a maioria das terapias de alvo molecular, utilizando fármacos conhecidos como ITQs (Mok *et al.*, 2017; Arora and Scholar, 2005).

Em condições normais, a fosforilação dos RTQs é estritamente regulada por fatores antagonistas e tirosina fosfatases. Alterações no domínio tirosina quinase propiciam uma ativação constitutiva das vias associadas, mesmo na ausência de ligantes, levando a uma proliferação celular descontrolada (Margiotta, 2021). Dentre

os principais mecanismos que alteram os genes que codificam RTQs estão: as mutações de ganho de função e as amplificações no domínio tirosina quinase; e os rearranjos gênicos que levam a codificação de oncoproteínas nas quais a porção auto-inibitória do domínio tirosina quinase está ausente, ou cujo o parceiro da fusão proporcione o recrutamento de proteínas que contribuem para a sinalização da via, estabilidade da proteína, localização celular e oligomerização (Lemmon and Schlessinger, 2010; Medves and Demoulin, 2012; Kubo *et al.*, 2009) (Figura 2).

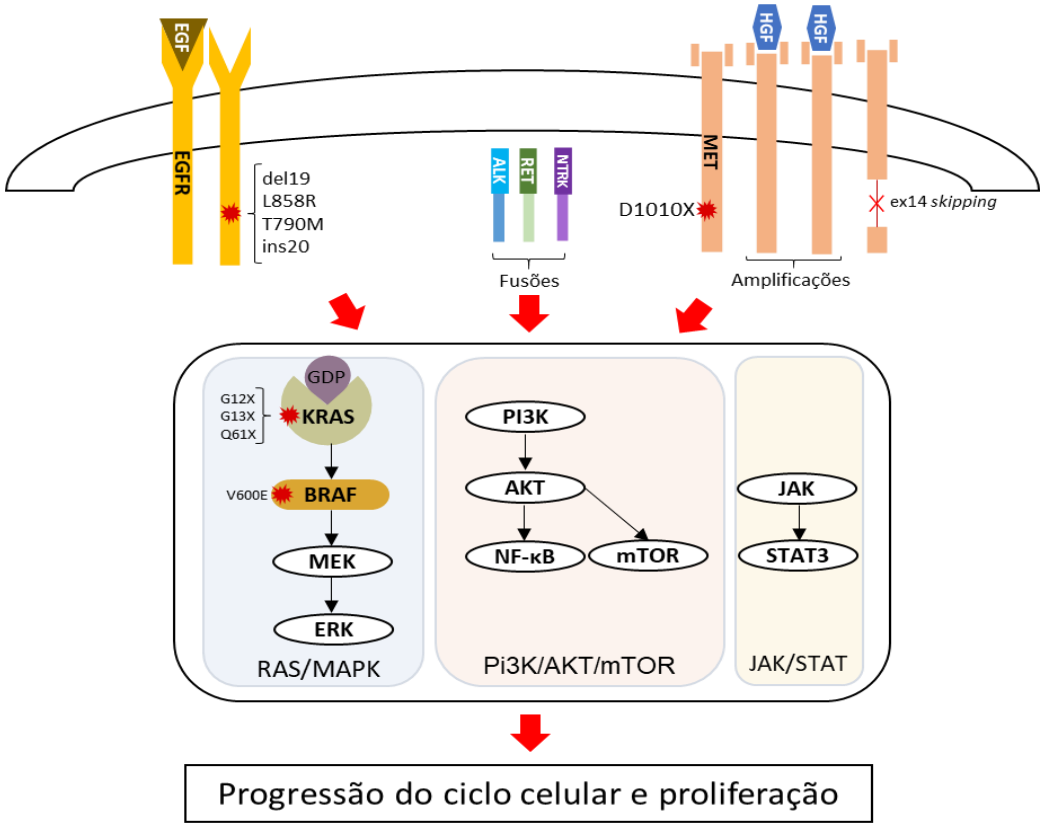


Figura 3. Representação das principais alterações moleculares envolvendo RTQs e vias de sinalização celular associadas a jusante. Em condições normais, ligantes levam a autofosforilação dos RTQs, desencadeando as vias a jusante que ativam fatores de transcrição e sinalizam a progressão do ciclo celular. Esse processo é controlado em condições de homeostasia, sendo ativado em situações específicas, como no reparo de tecidos. Quando ocorrem mutações em genes que codificam proteínas associadas a esta cascata de sinalização, as vias tornam-se constitutivamente ativas, com consequente descontrole da proliferação celular. Imagem original.

Quanto às alterações no gene *EGFR*, as deleções do éxon 19 e a substituição de uma arginina por uma leucina no códon 858 (L858R) do éxon 21 representam aproximadamente 90% das alterações somáticas do gene e são responsáveis por levar a uma atividade tirosina quinase constitutiva da proteína (Greulich *et al.*, 2005; Okabe *et al.*, 2007). Mutações complexas envolvendo mais de uma mutação em *EGFR* também já foram reportadas em uma minoria de casos, de 1% a 8% (revisado por (Santoni-Rugiu *et al.*, 2019). Os primeiros ITQs desenvolvidos para atuar sobre a proteína mutante do EGFR, chamados de primeira geração, ligam-se de maneira reversível ao domínio tirosina quinase, consequentemente bloqueando a cadeia de sinalização subsequente. Já as drogas de segunda geração ligam-se não somente irreversivelmente ao EGFR, como também bloqueiam outras proteínas associadas à família do fator de crescimento epidermal (EGF), incluindo HER2 (ou ErbB2, do inglês *Erb-B2 Receptor Tyrosine Kinase 2*) e HER4 (ou ErbB4, do inglês *Erb-B2 Receptor Tyrosine Kinase 4*) (Solca *et al.*, 2012; Engelman *et al.*, 2007). Já os ITQs de terceira geração foram desenvolvidos em resposta à mutação de resistência T790M do EGFR, que será abordada adiante. Nessa classe, o primeiro a ser aprovado é o fármaco Tagrisso (Osimertinibe), o qual também se liga de maneira irreversível a uma variedade de resíduos mutados da proteína, além de apresentar maior seletividade em comparação aos medicamentos das gerações anteriores, com menor afinidade pela proteína selvagem e, consequentemente, maior eficiência e menor toxicidade (Tan *et al.*, 2018).

As fusões gênicas, apesar de menos frequentes, possuem importante papel na oncogênese do CPNPC, sendo que as mais relevantes envolvem os genes *ALK*, *RET* e *ROS1*. Quando estes genes são fusionados com seus respectivos parceiros, traduzem uma proteína de fusão com o domínio tirosina quinase ativado constitutivamente, sobre o qual atuam também diferentes tipos de ITQs (Takeuchi *et al.*, 2012). A proteína de fusão *EML4-ALK* está presente de 2% a 5% dos pacientes com adenocarcinoma de pulmão (Kwak *et al.*, 2010). Já as fusões que incluem *ROS1* e *RET* são identificadas em 1 a 2% dos casos (Takeuchi *et al.*, 2012). Atualmente, somam-se as fusões acima aquelas envolvendo os genes *NTRK1, NTRK2* e *NTRK3*,

que codificam os receptores de tirosina quinase neurotrófica 1, 2 e 3, respectivamente. Em CPNPC, a frequência destas fusões é rara, sendo presentes em menos do que 0,5% dos casos (Liu *et al.*, 2022b). O ITQ Vitrakvi (Larotrectinibe), desenvolvido para inibir a ativação dos receptores de tropomiosina relacionados à quinase (TRKs, do inglês) pelas neurotrofinas (NTs), foi a primeira droga desta classe de medicamentos tumor-agnóstica, ou seja, aprovada para ser utilizada em qualquer neoplasia sólida, desde que na presença de fusões *NTRK1/2/3* (Mullard, 2018).

Ao contrário dos exemplos acima, mutações nos genes *KRAS* e *NRAS* indicam baixa probabilidade de resposta a ITQs. Mutações em *KRAS* são mais comuns em indivíduos com histórico de tabagismo e indicam um pior prognóstico, porém parecem estar associadas a maior expressão de PD-L1 e resposta a inibidores de *checkpoint* imunológicos (Pao *et al.*, 2005b; Dogan *et al.*, 2012; Slebos *et al.*, 1990; Herbst *et al.*, 2019). Estudos nos EUA, Europa e Ásia demonstraram que substituições de aminoácidos nos códons 12, 13 e 61 foram detectadas em 11%-29%, 12%-37% e 2%-9% dos pacientes, respectivamente (Lee *et al.*, 2010; Gao *et al.*, 2010; Fassina *et al.*, 2009; Chiosea *et al.*, 2010). Já as mutações em *NRAS* são raras, ocorrendo em aproximadamente 1% dos casos (Ohashi *et al.*, 2013).

Dada a alta frequência e pior prognóstico na presença de mutações em *KRAS*, não é exagero afirmar que um dos maiores avanços recentes no tratamento do CPNCP foi a aprovação do uso do inibidor de GTPase da família RAS Sotorasibe, indicado especificamente para tumores com a mutação G12C no gene, presente em mais de 10% da totalidade dos casos (Blair, 2021). Ademais, recentemente foram aprovados os fármacos Amivantamabe e Mobocertinibe, os quais atuam em tumores com inserções no éxon 20 de *EGFR*, presente em aproximadamente 5% dos pacientes e, até recentemente, marcadores de resistência aos ITQs de primeira e segunda geração (Yun *et al.*, 2020; Gonzalvez *et al.*, 2021). Com base no banco de dados do projeto GENIE, consórcio internacional que reúne dados genômicos de mais de 19 mil amostras de CPNPC, até 2021 estimava-se que aproximadamente 25% dos pacientes com CPNPC potencialmente se beneficiariam com o uso de terapias-alvo. Atualmente, com o advento de novos medicamentos direcionados para as inserções do éxon 20 no

EGFR e a mutação G12C de *KRAS*, mais de 40% dos pacientes passam a se beneficiar com os testes genômicos para a detecção de mutações acionáveis (Consortium, 2017).

Mutações em *BRAF* são comuns em melanoma e câncer de tireóide, mas também estão presentes de 3,5% a 5% dos CPNPC, sendo a mais frequente a V600E, presente em 1,5% - 2% dos casos (Marchetti *et al.*, 2011; Consortium, 2017). Este gene codifica a proteína B-Raf, a qual atua na ativação da via de sinalização MAPK/ERK, uma via de proliferação celular a jusante da de EGFR (Adjei, 2005) (Figura 2). Tumores com mutações ativadoras no gene também indicam baixa probabilidade de resposta no uso de ITQs direcionados a EGFR, porém os pacientes podem se beneficiar com o uso combinado de inibidores de B-Raf e MEK, Dabrafenibe e Trametinibe, respectivamente (Weart, Miller and Simone, 2018).

As amplificações gênicas têm importante papel na oncogênese de diversos tipos tumorais. Em CPNPC, a principal delas afeta o gene *MET*, sendo esta observada em aproximadamente 5% dos pacientes e responsável pela ativação constitutiva da via de sinalização do PIK3/AKT com consequente resistência aos inibidores do EGFR (Liang and Wang, 2020; Sequist *et al.*, 2011). Por outro lado, tumores com amplificaão de MET podem se beneficiar com o uso do ITQ Crizotinibe, utilizado também no tratamento de tumores com fusões em *ALK* e *ROS1*. Outras alterações em *MET* incluem mutações que afetam o *splicing* do pré-mRNA, resultando na exclusão do éxon 14, além de mutações do códon D1010X. Para tais alterações, já estão disponíveis os inibidores de c-Met Capmatinibe e Tepotinibe (Brazel, Zhang and Nagasaka, 2022).

Outra amplificação mais rara em CPNPC, presente em aproximadamente 1% dos casos, porém não menos importante, envolve o gene *ERBB2*, o qual codifica a proteína transmembrana e também receptor tirosina quinase HER2 (Consortium, 2017; Greulich *et al.*, 2012). Alterações neste gene são preditivas de resposta ao Trastuzumabe emtansina e Trastuzumabe deruxtecan, originalmente indicados para câncer de mama (Huang *et al.*, 2020).

Apesar das alterações citadas acima serem essencialmente de origem esporádica, suas prevalências diferem substancialmente entre diferentes populações. Por exemplo, a ocorrência de mutações em *EGFR* varia de 7% a 76% em pacientes com adenocarcinoma de pulmão (Midha, Dearden and McCormack, 2015). Tais alterações são mais frequentemente observadas em pacientes asiáticos e em mulheres não tabagistas (Lynch *et al.*, 2004). Em países com dimensões continentais e grande diversidade étnica como o Brasil, espera-se que diferenças expressivas entre diferentes regiões também sejam observadas. Um estudo publicado pelo nosso grupo, o qual consiste no segundo capítulo da presente tese, demonstrou que a prevalência de mutações em *EGFR* é menor no sul do Brasil (19,1%) em comparação com as demais regiões. Já a prevalência de alterações envolvendo o gene *KRAS* é a maior já registrada no país, presente em 30,2% dos casos analisados (Andreis *et al.*, 2019).

No entanto, mesmo com tantos avanços, o acesso a novas terapias ainda é um limitante, principalmente em países em desenvolvimento, como o Brasil. Estima-se que menos de 40% dos pacientes elegíveis para testes moleculares em CPNPC recebam o diagnóstico básico para o tratamento da neoplasia no país (Palacio *et al.*, 2019). Ademais, conforme demonstrado no Quadro 1, não é incomum que o processo de aprovação para o uso de novos fármacos leve mais de dois anos no país. Portanto, é fundamental o entendimento da epidemiologia molecular do CPNPC em países como o Brasil, possibilitando a melhor alocação de recursos para políticas de prevenção, diagnóstico e tratamento da neoplasia.

A frequência de marcadores moleculares em CPNPC também depende do subtipo histológico. Tumores escamosos possuem uma menor prevalência de mutações acionáveis em comparação com adenocarcinomas. Dados do projeto GENIE, que reúne dados de 1.826 casos de tumores escamosos de pulmão, indicam que menos de 2% apresentam a mutação G12C em *KRAS* ou mutações de sensibilidade a ITQs em *EGFR* (Network, 2014; Consortium, 2017). Entretanto, frente a falta de alternativas terapêuticas, a *National Comprehensive Cancer Network* (NCCN) ainda recomenda o teste molecular de *EGFR, ALK, KRAS, ROS1, BRAF,*

MET, RET e *NTRK1/2/3* em doença avançada ou metastática para este subtipo histológico (NCCN, 2022).

Algumas associações entre as alterações moleculares em CPNPC e aspectos clínico-patológicos também já foram observadas. Por exemplo, tumores com fusões gênicas em *ALK, ROS1* e *RET* são mais comumente encontradas em adenocarcinomas predominantemente sólidos e tendem a apresentar histologia mucinosa, padrão cribriforme, células de anel de sinete e citologia hepatóide (Pan *et al.*, 2014). Já mutações em *EGFR* estão frequentemente associadas com adenocarcinomas lipídico e micropapilar, sendo raramente encontradas em adenocarcinomas predominantemente sólidos ou invasivos mucinosos e mais comuns em pacientes asiáticas do sexo feminino não-tabagistas(Chapman *et al.*, 2016; Zhang *et al.*, 2016; Shim *et al.*, 2011). Mutações em *KRAS*, por outro lado, são frequentemente encontradas em tumores provenientes de indivíduos tabagistas, sendo em sua maioria adenocarcinomas invasivos mucinosos e raramente apresentam padrão predominante lipídico ou acinar (Jiang *et al.*, 2019; Suda, Tomizawa and Mitsudomi, 2010).

Apesar das associações descritas acima, nenhuma delas é utilizada como critério clínico para seleção de pacientes elegíveis para teste molecular e/ou para o uso de terapias de alvo molecular. Diferentes sociedades médicas recomendam teste molecular do tumor para todo CPNPC avançado ou metastático. No entanto, o NCCN já recomenda a testagem de mutações em *EGFR* no caso de tumores em estadiamento IB a IIIA. Naqueles positivos para mutação acionável, é indicado o uso de Osimertinibe para os já receberam quimioterapia adjuvante ou que sejam inelegíveis para receberem quimioterapia baseada em platina (NCCN, 2022).

Por fim, vale citar o tratamento por meio do bloqueio de *checkpoints* imunológicos (do inglês, *immune checkpoint inhibitors - ICIs*), como os sinais das proteínas PD-1 (do inglês, *programmed cell death-1*), PD-L1 (do inglês, *programmed cell death ligand-1*) e CTLA4 (do inglês, cytotoxic T-lymphocyte antigen 4), os quais possuem grande potencial para o tratamento de CPNPC para determinados grupos de pacientes (Rizvi *et al.*, 2015). PD-L1 é superexpresso em 23 - 27% dos casos de

câncer de pulmão de células grandes e adenocarcionoma, e de 19 a 56% dos carcinomas de células escamosas. Estudos já demonstraram que CPNPC que expressam altos níveis de PD-L1 respondem bem a imunoterapias que utilizem fármacos anti-PD-1/PD-L1 (Herbst *et al.*, 2014). No entanto, uma resposta satisfatória com o uso de um único agente anti-PD-1/PD-L1 tem variado de 10% - 45% em pacientes positivos para este biomarcador, demonstrando seu limitado valor preditivo de resposta (Garon *et al.*, 2015; Jiang *et al.*, 2017). Tais diferenças se devem em parte ao perfil molecular de cada tumor, incluindo a chamada carga mutacional, a qual indica o número de mutações por megabase do genoma tumoral. Conforme já demonstrado por diversos estudos, quanto maior o número de mutações no tumor, maior o potencial deste apresentar neoantígenos, criando um microambiente mais imunogênico e propício à resposta a ICIs (Blons *et al.*, 2019).

1.5. Mecanismos de resistência a drogas de alvo molecular

Apesar do impacto clínico positivo das novas terapias de alvo molecular, a resistência ao tratamento tem sido reportada como um dos principais desafios clínicos dessa abordagem terapêutica (Morgillo *et al.*, 2016). Em geral, pacientes tratados com ITQs começam a apresentar perda de resposta ao tratamento entre 10 e 14 meses após iniciarem o uso do medicamento (Wu and Shih, 2018). Tumores são compostos em sua maioria por um grupo dominante de células neoplásicas (clones) que compartilharam as mesmas alterações genéticas. No entanto, também podem apresentar subpopulações clones menos frequentes e que são geneticamente diferentes. Da mesma forma, o clone dominante original pode se diversificar em subclones que carregam mutações *de novo*. Tal diversidade, denominada heterogeneidade intratumoral, permite à neoplasia evadir estratégicas terapêuticas e apresentar resistência ao tratamento (Dagogo-Jack and Shaw, 2018; Kemper *et al.*, 2015; Zhang *et al.*, 2014). Terapias direcionadas, como ITQs, desencadeiam uma pressão seletiva sobre os clones presentes no tumor, selecionando negativamente aqueles que carregam as mutações que conferem sensibilidade ao medicamente e positivamente aqueles não apresentam resposta (Lim and Ma, 2019).

A resistência intrínseca (primária) ocorre quando um clone que possui uma determinada mutação de resistência já está presente desde antes do início do tratamento e, frequentemente, em uma baixa frequência. Este pode se expandir à medida que o clone dominante regride em virtude do tratamento. Por exemplo, estima-se que de 20% a 30% dos pacientes com mutações do *EGFR* que conferem sensibilidade ao uso de ITQs tenham ausência de resposta ou resposta de curta duração (menor que 3 meses) devido à presença de mecanismos de resistência intrínseca (Santoni-Rugiu *et al.*, 2019). O segundo tipo de mecanismo de resistência é denominado adquirido (secundário) e desenvolve-se a partir de mutações *de novo* no decorrer da terapia. Estes subclones tendem a substituir os clones originais ao longo do tratamento, levando a progressão da doença (Lim *et al.*, 2018).

Um exemplo bem difundido de resistência ao uso de ITQs direcionados ao EGFR ocorre quando se identifica no tumor a mutação T790M do éxon 20. Inicialmente foi descrito como um mecanismo de resistência adquirida, porém estudos demonstraram que alguns casos de CPNPC virgens de tratamento podem apresentar a alteração em uma baixa frequência alélica (Rosell *et al.*, 2011; Ye *et al.*, 2020). Esta foi uma das razões pelas quais o ITQ de terceira geração Osimertinibe, desenvolvido inicialmente para reverter a resistência em decorrência da T790M, passou a ser utilizado na primeira linha de tratamento do CPNPC (Ramalingam *et al.*, 2018).

Antes do uso de Osimetinib na primeira linha de tratamento, a mutação T790M no gene *EGFR* explicava 50-60% dos casos de resistência secundária em pacientes submetidos a ITQs de primeira geração, chegando a mais de 70% naqueles tratados com Afatinibe, um inibidor de segunda geração (Pao *et al.*, 2005a; Hochmair *et al.*, 2019). Outros mecanismos de resistência adquirida a ITQs de primeira e segunda geração direcionadas à EGFR estão associadas a mutações nos códons L747X, D761X e T854X de *EGFR* (menos que 5% dos casos), mutações em *PIK3CA* (5%) e *BRAF* (1%), amplificações do *EGFR* (8%), *MET* (5-22%), *ERBB2* (12%) e *CRKL* (9%), além de alterações fenotípicas incluindo transformação para histologia de câncer de pulmão de pequenas células (3-14%) (Gainor and Shaw, 2013). A transição Epitélio-

Mesenquimal também parece ser um mecanismo importante de resistência não dependente de *EGFR* (Bronte *et al.*, 2018).

Atualmente o mecanismo mais frequente de resistência adquirida a ITQ de terceira geração é a mutação C797S, também no gene *EGFR*, presente entre 22 a 40% dos casos tratados com Osimertinibe (Wang *et al.*, 2016; Tan *et al.*, 2018). Outros mecanismos não relacionados a T797S incluem aqueles semelhantes aos observados após uso de fármacos de primeira e segunda geração, incluindo amplificações de outros genes associados à via (*EGFR*, *ERBB2, MET, FGFR1* e *KRAS*), fusões de gênicas (*RET, ALK, FGFR3* e *NTRK1*) e transformação para histologia de câncer de pulmão de pequenas células. Também são observadas fusões de *BRAF* e mutações em *KRAS*. Normalmente esses mecanismos estão associados ao desaparecimento dos clones T790M e o reaparecimento da mutação ativadora de *EGFR* original, demonstrando a importância de um monitoramento molecular constante e seriada da evolução da neoplasia (Nakatani *et al.*, 2019; Ramalingam *et al.*, 2018; Minari, Bordi and Tiseo, 2016; Lim *et al.*, 2018; Oxnard *et al.*, 2018; Yang *et al.*, 2018; Piotrowska *et al.*, 2018; Schrock *et al.*, 2018; Yu, Planchard and Lovly, 2018; Leonetti *et al.*, 2021).

O uso de Osimertinibe na primeira linha de tratamento para pacientes *EGFR* mutados ainda é relativamente recente, porém alguns estudos já exploraram os mecanismos de resistência associados. Ao contrário do que ocorre quando utilizado na segunda linha de tratamento, Fuchs e colaboradores demonstraram que a mutação C797S não foi encontrada em pacientes que receberam o medicamento na primeira linha. Nestes casos, amplificações de *MET* foi o evento mais comum (Fuchs *et al.*, 2021). O número de pacientes avaliados, no entanto, foi relativamente baixo, apenas 15 para cada grupo. O estudo AURA, uma das principais evidências favoráveis ao uso de Osimertinibe na primeira linha de tratamento, identificou, dentre os 38 pacientes que apresentaram progressão, amplificações de *MET, EGFR* e *KRAS*; mutações de *MEK1, KRAS, PIK3CA* e *JAK2*; inserção no éxon 20 de *ERBB2*; e dois pacientes com C797S (Ramalingam *et al.*, 2018).

Mecanismos emergentes de resistência a ITQs direcionados à *EGFR* incluem micro-RNAs. Um exemplo é o miR-196a, para o qual níveis elevados de expressão

foram associados a resistência a Gefitinibe mediante inibição da expressão de proteína de transferência de glicolipídios (GLTP), que induz apoptose em CPNPC (Liu *et al.*, 2022a).

No que diz respeito a resistência secundária em pacientes com fusões *ALK* tratados com Crizotinibe, o mecanismo mais comum é a ativação de EGFR não direcionada por mutações, mas sim por ligantes do EGF e anfiregulina (AREG), presente em mais de 40% dos casos. Outros escapes comuns envolvem mutações de ponto no próprio gene *ALK*, sendo a L1196M a principal delas, encontrada em até 36% dos casos. Além disso, amplificações de *ALK* e *KIT*, presentes em até 18% e 15% dos casos, respectivamente, já foram descritas (Gainor and Shaw, 2013). Amplificações do gene *MYC*, encontradas 4% dos CPNPC, já foi reportada como mecanismo de resistência primária a Crizotinibe.

Em relação à resistência adquirida a inibidores de MET, diversas alterações moleculares já foram descritas, sendo aquelas envolvendo os códons H1094X, G1163X, L1195X, D1228X e Y1230X e amplificação focal do éxon 14 as mais comuns. Outros mecanismos que não envolvem o gene, presente em aproximadamente 45% dos casos, incluem amplificações de *ERBB3, EGFR, KRAS* e *BRAF*, bem como mutações em *KRAS*. Estima-se que 25% dos tumores tratados com inibidor de MET ainda apresentem mecanismos de resistência adquirida indeterminada (Recondo *et al.*, 2020).

1.6. DNA livre circulante: a mais nova fronteira do diagnóstico molecular?

Como discutido acima, a resistência ao tratamento ainda é um dos principais desafios na terapêutica do CPNPC, seja ela intrínseca ou adquirida. A análise molecular constante e contínua, com um acompanhamento em tempo real da evolução dos clones que carregam alterações em genes associados à oncogênese (*drivers* oncogênicos), permite, em teoria, que o tratamento do paciente seja ainda mais personalizado, antecipando-se à evolução clínica da doença. Um exemplo do potencial desta abordagem já foi demonstrado por Fujita e colaboradores, os quais chegaram a encontrar 79% de casos virgens de tratamento positivos para mutação T790M

utilizando a técnica de hibridização de colônia, atingindo uma sensibilidade de 0.01% (Fujita *et al.*, 2012). No entanto, técnicas dessa natureza ainda demandam tempo e custos relativamente altos, além de exigirem instalações altamente especializadas.

Já para detecção de resistência adquirida, um caminho óbvio parece ser a realização de biópsias seriadas do tecido tumoral. No entanto, é comum encontrar em tumores avançados uma alta taxa de heterogeneidade intratumoral (Dagogo-Jack and Shaw, 2018; Kemper *et al.*, 2015). Portanto, para que haja uma maior representatividade do perfil molecular do tumor, seriam necessárias diversas biópsias de variadas frações do tumor, além de espécimes provenientes dos diversos sítios metastáticos, quando presentes. Novamente, a logística e custo para tal abordagem a tornam pouco viável para prática clínica. Finalmente, um outro obstáculo para diagnóstico molecular de biópsias teciduais é a alta invasividade da técnica, sendo que por vezes o tecido a ser biopsiado é inacessível, p.ex. quando há metástases no sistema nervoso central, principalmente em pacientes que apresentam complicações clínicas importantes para realização de procedimentos cirúrgicos (Heitzer *et al.*, 2013).

Um dos grandes avanços na investigação de biomarcadores tumorais foi o advento da análise em material tumoral circulante, conhecida como biópsia líquida. Esta técnica se baseia no rastreamento de alterações da neoplasia que estejam presentes no sangue ou outros fluidos corpóreos do paciente (por exemplo líquido pleural e cefalorraquidiano, urina, etc), tanto a nível celular (CTC, do inglês *circulating tumor cells*), quando a nível molecular, utilizando o DNA tumoral circulante (ctDNA, do inglês *circulating tumor DNA*). Esta última se baseia no rastreamento dos fragmentos do material genético tumoral naturalmente liberados no sistema circulatório após as células neoplásicas sofrerem apoptose ou necrose, fenômeno observado já há muitas décadas (Leon *et al.*, 1977; Stroun *et al.*, 1987; Sorenson *et al.*, 1994). Dentre suas principais vantagens estão a baixa invasividade e a possibilidade de realizar múltiplas coletas ao longo do tratamento do paciente, bastando a coleta de aproximadamente 10 mL de sangue ou outros fluidos, conforme necessidade. Há perspectiva no uso de ctDNA no monitoramento de doença residual mínima (DRM) após ressecção cirúrgica do tumor ou mesmo no rastreamento de tumores na população clinicamente saudável

(Abbosh, Birkbak and Swanton, 2018; Aravanis, Lee and Klausner, 2017). No caso da DRM, as alterações identificadas no tecido neoplásico ressecado podem ser monitoradas no DNA livre circulante (cfDNA, do inglês *circulanting cell-free DNA*), e caso sejam encontradas, mesmo que em baixa frequência, podem indicar um risco de recidiva da doença (Chin *et al.*, 2019; Yang *et al.*, 2020a). Já o monitoramento de populações fenotipicamente saudáveis para detecção de neoplasias ainda é mais desafiante, uma vez que tumores em estágios iniciais tendem a liberar baixíssimas quantidades de ctDNA, abaixo do limite de detecção da maioria das técnicas (Bettegowda *et al.*, 2014). Portanto, atualmente seu maior potencial reside no monitoramento de doença avançada, auxiliando inclusive na readequação do tratamento conforme evolução clonal do tumor (Yanagita *et al.*, 2016; Remon *et al.*, 2017; Wan *et al.*, 2017) (Figura 3).

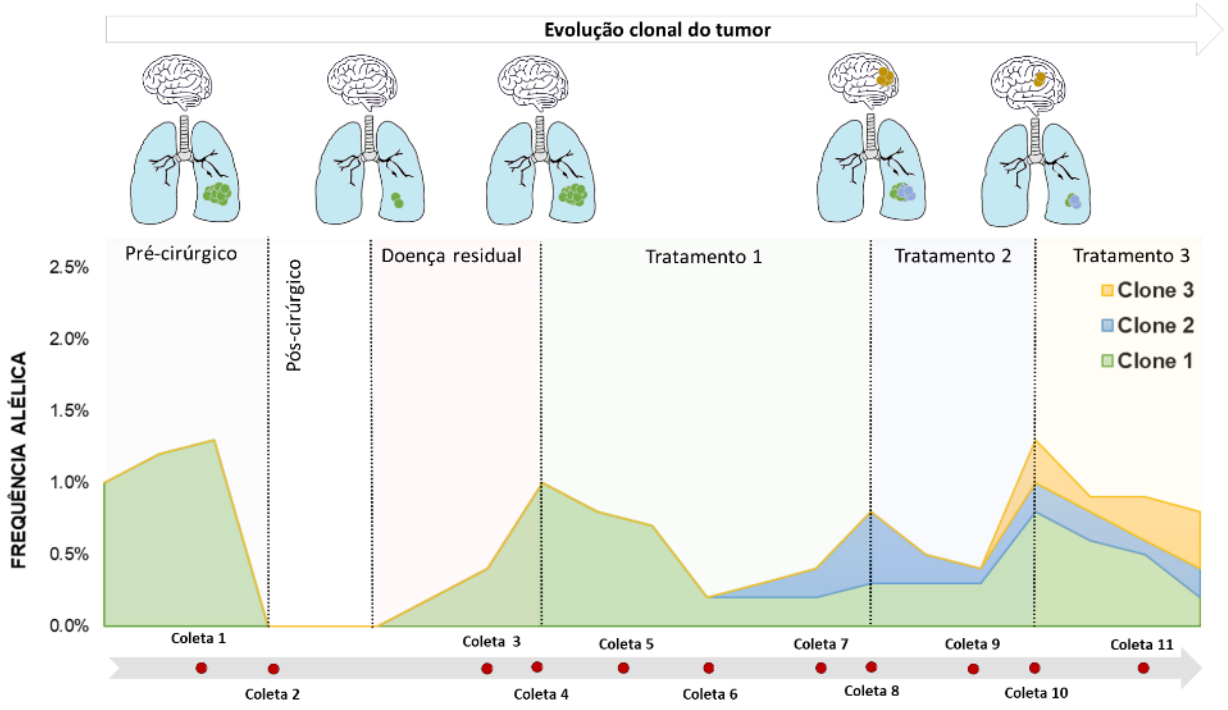


Figura 4. Ilustração das potenciais aplicações da biópsia líquida para investigação do DNA tumoral circulante (ctDNA). A identificação de mutações acionáveis pode auxiliar no monitoramento de doença residual mínima (DRM) e na decisão terapêutica, inclusive indicando uma readequação do tratamento conforme a evolução clonal do tumor. Figura original.

O cfDNA ocorre naturalmente em baixas concentrações no plasma, sendo derivado em sua grande maioria de células geneticamente normais (genótipo selvagem). Apenas uma pequena fração desse material corresponde aos alelos derivados do tumor (ctDNA), normalmente 0,01% a <1%, o que exige técnicas de alta sensibilidade para sua detecção (Volckmar *et al.*, 2018). A quantidade de cfDNA em pacientes oncológicos é maior do que em indivíduos saudáveis, principalmente devido ao aumento da atividade apoptótica e necrótica sobre as células neoplásicas (El Messaoudi *et al.*, 2013; Schwarzenbach, Hoon and Pantel, 2011). Portanto, além da busca de mutações no ctDNA, a quantificação total de cfDNA também pode ser, em alguns cencários, utilizada como marcador preditivo e/ou de prognóstico. Por exemplo, altos níveis de DNA livre circulante no plasma de pacientes com câncer de reto e hepatocelular foram considerados um marcador independente de pior prognóstico após tratamento (Schou *et al.*, 2018; Nakatsuka *et al.*, 2021).

Diversas técnicas para avaliação molecular do ctDNA já foram desenvolvidas, porém ainda carecem de especificidade e sensibilidade em níveis aceitáveis para o uso clínico (Quadro 3). A evolução das novas plataformas de sequenciamento de nova geração (NGS, do inglês *Next Generation Sequencing*), foram fundamentais para a evolução do campo da biópsia líquida. O NGS permite a detecção de mutações com frequências inferiores a 1%, e tem sido uma das principais ferramentas para a investigação de ctDNA, apresentando diversas vantagens frente a outras tecnologias (Siravegna *et al.*, 2017) (Quadro 3). Artefatos naturalmente gerados no processo de preparo das bibliotecas genômicas utilizadas pelo NGS também têm sido uma dificuldade importante para detecção de alterações em baixa frequência alélica. Novas tecnologias que incluem identificadores moleculares únicos (UMI, do inglês *unique molecular identifiers*), reduziram significativamente o índice de falsos-positivos, aumentando a sensibilidade e especificidade do teste, ao custo da necessidade de maiores coberturas de sequenciamento (Deveson *et al.*, 2021; Weber *et al.*, 2020).

Quadro 4. Resumo das principais técnicas atualmente disponíveis para a análise de DNA tumoral circulante (ctDNA) a partir de biópsias líquidas, incluindo suas principais vantagens, desvantagens, limite de detecção e tipos de alterações detectadas.

Técnica/Teste	Alterações detectadas	Limite de detecção	Sensibilidade	Especificidade	Vantagens	Desvantagens
Espectrometria de massa	SNVs pré-estabelecidos	1-10%	38,9%	84,6%	Agilidade na obtenção dos resultados.	Baixa sensibilidade. Restrito a alterações já descritas. Limite de detecção alto para muitas aplicações.
CLEA	SNVs pré-estabelecidos	1-5%	81,8%	89,5%	Agilidade na obtenção dos resultados.	Restrito a alterações já descritas. Limite de detecção alto para muitas aplicações.
Cobas (RT-PCR)	SNVs pré-estabelecidos	1-3%	61,4%	78,6%	Baixo custo. Agilidade na obtenção dos resultados.	Baixa sensibilidade e especificidade. Restrito a alterações já descritas. Limite de detecção alto para muitas aplicações.
Scorpion/ARMS	SNVs pré-estabelecidos	1-3%	61,8-85,7%	94,3-100%	Alta especificidade. Agilidade na obtenção dos resultados.	Sensibilidade pode ser baixa. Restrito a alterações já descritas. Limite de detecção alto para muitas aplicações.
HRMA	SNVs e indels pré-estabelecidos	0,1-10%	91,67%	100%	Alta especificidade. Agilidade na obtenção dos resultados. Limite de detecção baixo em algumas situações.	Restrito a alterações já descritas. Limite de detecção pode ser alto em algumas aplicações.
ddPCR	SNVs pré-estabelecidos	>0,1%	77%	63%	Baixo limite de detecção. Agilidade na obtenção dos resultados.	Baixa sensibilidade e especificidade. Restrito a alterações já descritas.
BEAMing	SNVs, CNVs e fusões pré-estabelecidos	>0,1-0,01%	70%	69%	Baixo limite de detecção, detecta alterações complexas	Baixa sensibilidade e especificidade. Restrito a alterações já descritas.

Técnica/Teste	Alterações detectadas	Limite de detecção	Sensibilidade	Especificidade	Vantagens	Desvantagens
NGS	SNVs, indels, CNVs e fusões novos e pré-estabelecidos	0,01-5%	93%	94%	Baixo limite de detecção. Alta sensibilidade/especificidade. Detecta novas alterações, incluindo aquelas de alta complexidade	Alto custo. Dependendo da plataforma, limite de detecção pode se tornar alto para algumas aplicações.
PNA-PCR	SNVs e indels pré-estabelecidos	0,01%	78%	100%	Baixo limite de detecção e alta especificidade	Baixa sensibilidade. Restrito a alterações já descritas.

Adaptado de Wu & Shih, 2018. CLEA = Cromatografia Líquida Desnaturante de Alta Performance (do inglês, *Liquid Chromatography Clean-up Method*); RT-PCR = Reação em Cadeia de Polimerase (PCR) em Tempo Real (do inglês, *Real Time Polymerase Chain Reaction*); ARMS = Amplificação por Mutação Refratária (do inglês, *Amplification-Refractory Mutation System*); HRMA = Análise de Fusão de Alta Resolução de Fragmentos de DNA (do inglês, *High-Resolution Melting Analysis*); ddPCR = PCR Digital em Gotas (do inglês, *Digital Droplet PCR*); BEAMing = Esferas, Emulsão, Amplificação e Magnetismo (do inglês, *Beads, Emulsion, Amplification, Magnetics*); NGS = Sequenciamento de Nova Geração (Next Generation Sequencing); PNA-PCR = PCR Mediada por Ácidos Nucleicos Peptídicos (do inglês, *Peptide Nucleic Acid-Mediated PCR*); SNVs = Variantes de Base Única (do inglês, *Single Nucleotide Variant*); indels = Pequenas Inserções e Deleções (<1 kb); CNVs = Variação no Número de Cópias (do inglês, *Copy Number Variation*).

Uma importante limitação no rastreamento de alterações no ctDNA é a ocorrência Hematopoiese Clonal de Potencial Indeterminado (CHIP). A CHIP consiste no acúmulo de mutações somáticas em células tronco hematopoiéticas, levando a expansão destas no processo natural de envelhecimento do indivíduo ou após regimes terapêuticos intensivos, como quimio e radioterapia (Steensma, 2018). Estas células mutadas podem se expandir de forma clonal, causando ruído na análise do cfDNA, levando a classificação equivocada de variantes que não contribuem para o processo de tumorigênese (Jaiswal *et al.*, 2014; Razavi *et al.*, 2019; Hu *et al.*, 2018). Estima-se que a presença de CHIP seja responsável por diminuir a especificidade da técnica, com índices de falsos positivos que podem chegar a 50% (Razavi *et al.*, 2019; Chan *et al.*, 2020). Uma forma de contornar o problema é a análise em paralelo do DNA germinativo do paciente, normalmente extraído da fração leucoplaquetária da amostra, de maneira a filtrar aquelas variantes que são de fato derivadas do tumor (Leal *et al.*, 2020). Estudos em CPNPC revelaram a presença de CHIP em mais de 50% dos casos durante a avaliação do cfDNA. Até 20% destes poderiam ter sido erroneamente reportados caso não houvesse uma análise pareada do DNA germinativo (Chabon *et al.*, 2020; Zhang *et al.*, 2021; Razavi *et al.*, 2019). Por outro lado, a grande maioria das alterações encontradas na CHIP não tem implicações clínicas para o manejo e tratamento de muitos tumores sólidos, questionando a necessidade de sua inclusão na prática clínica (Hu *et al.*, 2018). Ademais, a inclusão da análise em pareada com DNA germinativo agrega custo ao teste, o que o tornaria ainda menos acessível em países como o Brasil (Liu *et al.*, 2019).

Apesar das dificuldades expostas acima, um dos grandes potenciais da biópsia líquida é sua possibilidade de vislumbrar a heterogeneidade intratumoral (Chabon *et al.*, 2016). Utilizando técnicas de biópsia tecidual convencionais, mesmo com a seleção de pacientes elegíveis para tratamentos com terapia-alvo, estudos já demonstraram que a resposta objetiva varia de 50% a 83%, sendo raros os casos onde há resposta completa. Parte dessa variação pode ser explicada pela coocorrência de outros *drivers* oncogênicos (heterogeneidade *intradriver*), cujos clones podem expandir-se frente a seleção negativa ocasionada pelo tratamento-alvo sobre aqueles que possuem alterações acionáveis, de maneira semelhante ao que ocorre na resistência intrínseca

(Skoulidis and Heymach, 2019). Muitas dessas alterações podem ser sub-notificadas na biópsia tecidual, estando presentes em frações não avaliadas do tumor e/ou em metástases distantes não-biopsiadas (Lim and Ma, 2019; Wu *et al.*, 2021; Zhang *et al.*, 2019b). Nesse sentido, a avaliação do ctDNA poderia ser uma ferramenta valiosa para validação de novos protocolos de probabilidade de resposta a terapias-alvo (Guibert *et al.*, 2020). Uma das aplicações seria o uso de tratamentos combinados. Alguns estudos já relataram tentativas desta natureza com algum grau de sucesso, porém a resposta foi variada e a toxicidade parece ser um limitador importante nessa abordagem (Fuchs *et al.*, 2021).

Levando em consideração a baixa quantidade e estabilidade do cfDNA presente no plasma, diversos fatores devem ser considerados na fase pré-analítica para um diagnóstico preciso. No caso de sangue total, este deve ser coletado em tubos com anticoagulante EDTA (do inglês *ethylenediamine tetraacetic acid*), mantidos resfriados (2 C° a 8 C°) e processados até duas horas para evitar degradação enzimática. Atualmente já existem outras alternativas, com preservativos que garantem estabilidade da amostra de até três dias em temperatura ambiente. Uma vez coletado o sangue, o plasma deve ser imediatamente separado por centrifugação em baixa rotação. Este então pode ser armazenado em -20 C ou -80 C para preservação por maiores períodos de tempo ou imediatamente submetido à extração de ácidos nucléicos. Antes da extração, o plasma passa por uma segunda rodada de centrifugação em alta rotação, de maneira a separar detritos celulares e aumentar a seletividade dos fragmentos de cfDNA. Também é importante a escolha de kits de extração de ácidos nucléicos seletivos para fragmentos pequenos (<166 pb), de maneira a amenizar a contaminação por gDNA. Por fim, uma vez extraído, o material deve ser armazenado em tubos com baixa retenção de DNA, uma vez que o rendimento por muitas vezes pode ser baixo (Meddeb, Pisareva and Thierry, 2019; Greytak *et al.*, 2020; Godsey *et al.*, 2020; Heitzer *et al.*, 2022).

Já na etapa analítica, ferramentas para avaliação da sensibilidade e especificidade da técnica são imprescindíveis. A sensibilidade é definida pelo limite superior de detecção de determinadas alterações pela técnica, enquanto a especificidade é definida pela capacidade do teste reportar de maneira precisa determinadas alterações, indicando a acurácia do teste associada a sua sensibilidade.

Tais indicadores utilizam valores preditivos positivo (VPP) e negativo (VPN), onde o VPP indica a porcentagem de pacientes/amostras que de fato apresentam a alteração/mutação detectada pelo teste, enquanto o VPN indica a porcentagem de pacientes/amostras negativos no teste que de fato não carregam a alteração/mutação. Um VPP elevado se traduz em baixas taxas de falso negativos, enquanto um VPN alto se traduz em baixas taxas de falsos positivos (Poole *et al.*, 2019; Trevethan, 2017). Atualmente, controles comerciais com frequências alélicas conhecidas de determinadas alterações são opções acessíveis para determinar a especificidade e sensibilidade das técnicas disponíveis para avaliação de ctDNA. Atualmente o *College of American Pathology* (CAP) e o *European Molecular Genetics Quality Network* (EMQN) já oferecem programas de controle de qualidade externos para ctDNA e tem auxiliando na difusão da técnica na prática clínica, porém ainda não existe um consenso de quais alvos devem ser incluídos e em qual contexto (Heitzer *et al.*, 2022).

Apesar de seu alto custo e falta de diretrizes clínicas bem estabelecidas para o seu uso, no Brasil diversos laboratórios particulares já oferecem soluções baseadas na investigação do ctDNA. No entanto, essa ainda é uma realidade distante no SUS e atualmente o teste não é nem mesmo oferecido pelos planos de saúde do país. Nesse sentido, é de extrema necessidade a inclusão de mais estudos que avaliem sua aplicabilidade na realidade clínica brasileira, auxiliando, em um futuro próximo, a tomada de decisões considerando seu custo-efetividade (IJzerman *et al.*, 2021). Apesar do alto investimento para implementação da técnica, seus potenciais benefícios devem ser levados em conta, especialmente considerando a potencial redução da necessidade de intervenções clínicas de alta complexidade e, acima de tudo, nos ganhos na qualidade de vida do paciente e eventualmente aumento da sobrevida (Cheng, Akalestos and Scudder, 2020).

1.7. Biópsia líquida aplicada ao monitoramento do câncer de pulmão

Em CPNPC, até 80% dos pacientes com doença avançada terão acesso apenas a biópsias com pequenas quantidades de tecido, o que limita sua utilização no diagnóstico molecular. Outros levantamentos demonstram também que até 31% dos pacientes não possuem tecido acessível para biópsia e daqueles biopsiados, até 20%

poderão ser inadequados para testagem devido à baixa quantidade de material (Chouaid *et al.*, 2014; Liam, Mallawathantri and Fong, 2020). Portanto, a avaliação do ctDNA em CPNPC apresenta grande potencial de aplicação efetiva. Já foi demonstrado que a técnica é uma ferramenta de utilidade clínicas na seleção de pacientes para o tratamento com terapias de alvo molecular, incluindo a detecção de mutações acionáveis em um terço dos casos onde o tecido não estava disponível (Thompson *et al.*, 2016; Remon *et al.*, 2020).

Assim como ocorre em outros subtipos tumorais, o uso da biópsia líquida tem maior potencial em CPNPC avançados, uma vez que a abundância do ctDNA está correlacionada ao estadiamento da doença (Bettegowda *et al.*, 2014). Outros fatores que influenciam a detecção de mutações driver em ctDNA na doença incluem uma histologia não-adenocarcinoma, necrose tumoral, índices elevados de proliferação (marcação para Ki67) e invasão linfovascular (Abbosh *et al.*, 2017). O sítio metastático também deve ser levado em conta para correta escolha de amostragem. Por exemplo, metástases no sistema nervoso central apresentam um desafio mesmo para detecção de ctDNA, uma vez que a barreira hematoencefálica impede que o material seja liberado na circulação. Nesses casos, recomenda-se a coleta de fluido cerebroespinhal (Ma *et al.*, 2020).

A biópsia líquida já demonstrou ser uma alternativa viável para monitorar a resposta de pacientes durante o uso de ITQs (Yanagita *et al.*, 2016; Remon *et al.*, 2017). A técnica já tem sido utilizada na prática clínica nos casos onde há progressão da doença em decorrência de resistência adquirida via mutação T790M no gene *EGFR* (Wan *et al.*, 2017). No entanto, resultados negativos apara esta mutação devem ser interpretados com cautela e sempre que possível confirmados na avaliação tecidual. Estudos utilizando esta abordagem também já encontraram outros *drivers* coocorrendo com a T790M, incluindo *CDK6, CCNE1, CTNNB1, AR, MYC* e *BRCA1*, o que poderia auxiliar na predição de resposta ao tratamento (Blakely *et al.*, 2017). Em estudo conduzido por Weber e colaboradores foi observado que mudanças na dinâmica do ctDNA durante o tratamento com inibidores de *checkpoint* imunológico apresentam valor preditivo, incluindo a presença de mutações nos genes *STK11* e *KEAP1*, que indicaram um pior prognóstico quando presentes após o início do tratamento, além de encontrarem uma associação entre o decaimento de ctDNA e uma maior sobrevida (Weber *et al.*, 2021).

Apesar dos avanços e potenciais aplicações, conforme ocorre em outros tumores, inúmeras limitações como a baixa carga tumoral, estadiamento inicial, tratamentos utilizados previamente, e acurácia das metodologias de detecção têm sido descritos como fatores que influenciam diretamente na sensibilidade identificação de alterações no ctDNA. Atualmente ainda não existem diretrizes clínicas bem estabelecidas para investigação de ctDNA em pacientes com CPNPC e a re-biópsia do tecido tumoral continua sendo a alternativa mais desejável, uma vez que o índice de falsos-negativos varia de 30% a 40% (Lindeman *et al.*, 2018). Exceções para as quais há indicação no uso da metodologia se aplicam a pacientes sem condições clínicas para realização de novas biópsias ou para aqueles com resultado negativo para *drivers* oncogênicos cujo tecido é insuficiente para re-análise (NCCN, 2022).

Mesmo com as limitações citadas acima, já existe algum grau de consenso do que deve ser testado no ctDNA e quando este deve ser analisado. No caso do monitoramento da DRM, painéis focados são mais recomendados para acompanhar a evolução de mutações previamente detectadas no tumor, uma vez que tendem a apresentar maior sensibilidade e menor índice de falso-positivos. Já no monitoramento de mecanismos emergentes de resistência, como por exemplo durante o tratamento com ITQs, painéis mais abrangentes são recomendados para identificação de alterações *de novo* (Heitzer et al., 2022). Em suma, a Figura 4 abaixo ilustra alguns dos possíveis cenários clínicos onde a biópsia líquida para avaliação do ctDNA em CPNPC poderia ser utilizada.

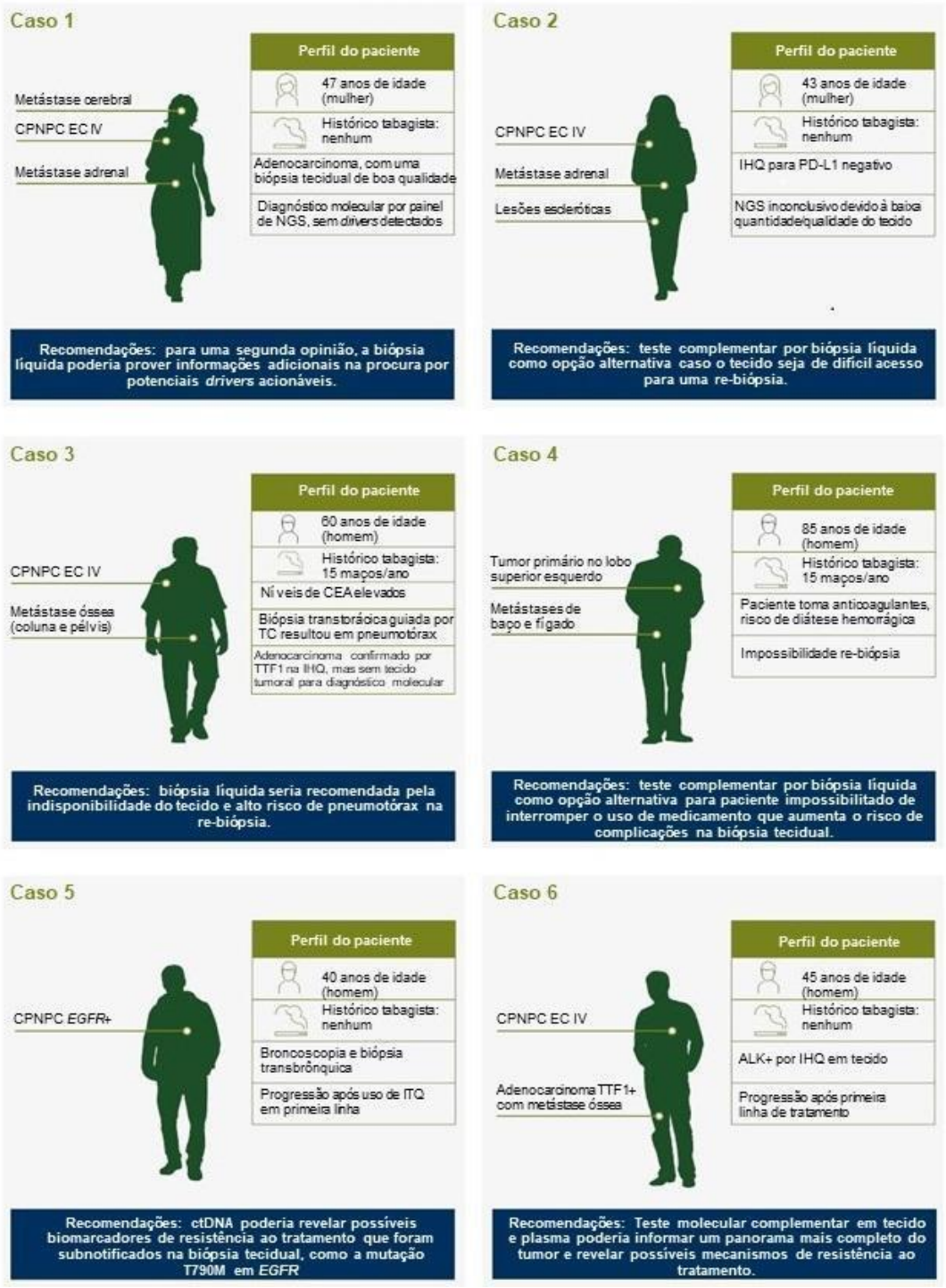


Figura 5. Possíveis cenários onde a avaliação do ctDNA poderia apresentar vantagens ao paciente com CPNPC. Adaptado de Heitzer et al. (2021).

CAPÍTULO II – JUSTIFICATIVA

O câncer de pulmão é a principal causa de morte por câncer no mundo. A sobrevida média cumulativa total em cinco anos varia entre 13 e 22% em países desenvolvidos e entre 7 e 10% nos países em desenvolvimento. No Brasil, é a terceira e quarta neoplasia mais comum em homens e mulheres, respectivamente, e a que mais leva a óbitos se somados ambos os sexos.

Avanços significativos foram feitos na identificação e caracterização de mutações *drivers* e acionáveis no CPNPC. Em paralelo a estes avanços, novas drogas de alvo molecular foram desenvolvidas para subtipos específicos da neoplasia. Em países da Ásia, Europa e América do Norte, as frequências mutacionais em genes clinicamente relevantes utilizados como preditores para o tratamento da doença é relativamente bem conhecida, mas esta informação ainda é escassa no Brasil. A avaliação do perfil molecular de CPNPC na população brasileira, a qual é altamente diversa e miscigenada, aliada a uma correlação destes dados com os parâmetros clínicos dos pacientes, poderia auxiliar na tomada de decisões e guiar novas políticas públicas para o tratamento da neoplasia.

A biópsia líquida tem representado uma nova alternativa para análise molecular das neoplasias através do rastreamento de mutações presentes no ctDNA. Este é um método minimamente invasivo e permite coletas seriadas para o acompanhamento da evolução da doença. No câncer de pulmão, em situações específicas, seu uso clínico já é aprovado pelo FDA (do inglês, *Food and Drug Administration*) e estudos recentes têm demonstrado seu potencial como uma ferramenta acessória à análise em tecido. No entanto, a técnica ainda foi pouco explorada no Brasil, principalmente em pacientes atendidos pelo SUS. Além de auxiliar no diagnóstico, a abordagem oportuniza uma melhor compreensão da heterogeneidade molecular dos tumores de pulmão que acometem a população brasileira.

CAPÍTULO III – OBJETIVOS

3.1. Objetivo geral

Avaliar o perfil molecular do câncer de pulmão de não pequenas células em pacientes do sul do Brasil e implementar a técnica de biópsia líquida para investigação da heterogeneidade tumoral pelo sequenciamento do DNA tumoral circulante (ctDNA).

3.2. Objetivos específicos

- Investigar a frequência das alterações somáticas nos *genes EGFR, KRAS, NRAS* e *BRAF*, além da expressão dos biomarcadores ALK e PD-L1, em uma série de pacientes do sul do Brasil diagnosticados com câncer de pulmão não pequenas células (CPNPC);
- Correlacionar o perfil molecular do CPNPC com os parâmetros clínicos dos pacientes;
- Investigar a escolha e resposta terapêutica em pacientes com CPNPC que tenham mutações compostas ou de significado clínico indeterminado em *EGFR*;
- Utilizar a metodologia de biópsia líquida para investigação da heterogeneidade tumoral através do ctDNA em pacientes diagnosticados com CPNPC em estadiamento IIIC não ressecáveis atendidos no Hospital de Clínicas de Porto Alegre (HCPA).
- Monitorar a frequência clonal das alterações genéticas de pacientes com diagnóstico de CPNPC no HCPA;
- Produzir conhecimento que possa auxiliar na melhoria de políticas públicas e que em última análise garantam um melhor acesso ao tratamento do CPNPC.

CAPÍTULO IV – MANUSCRITO I

"***Analysis of Predictive Biomarkers in Patients with Lung Adenocarcinoma from Southern Brazil Reveals a Distinct Profile from Other Regions of the Country***"

Artigo publicado no periódico *Journal of Global Oncology* (ISSN 2687-8941) - 2019 Sep;5:1-9. doi: 10.1200/JGO.19.00174.

original report

Analysis of Predictive Biomarkers in Patients With Lung Adenocarcinoma From Southern Brazil Reveals a Distinct Profile From Other Regions of the Country

Tiago F. Andreis, MSC[1,2]; Bruno S. Correa[2]; Fernanda S. Vianna, PhD[1,2]; Fernanda De-Paris, PhD[2]; Marina Siebert, PhD[2]; Sandra Leistner-Segal, PhD[2]; Eriza C. Hahn, MSc[1]; Jane M. Ulbrich, PhD, MD[2]; Luis F.R. Rivero, PhD, MD[2]; Francine H. De Oliveira, MD[2]; Vinícius Lorandi, MD[3]; Patricia Ashton-Prolla, PhD, MD[1,2]; and Gabriel S. Macedo, PhD[2]

abstract

PURPOSE Adenocarcinoma is the most common histologic subtype of non–small-cell lung cancer, representing 40% of all diagnoses. Several biomarkers are currently used to determine patient eligibility for targeted treatments, including analysis of molecular alterations in *EGFR* and *ALK*, as well as programmed death-ligand 1 (PD-L1) protein expression. Epidemiologic data reporting the frequency of these biomarkers in Brazilian patients with lung adenocarcinoma (LUAD) are limited, and existing studies predominantly included patients from the southeast region of the country.

MATERIALS AND METHODS The goal of this study was to investigate the frequency of somatic mutations in the *EGFR, KRAS, NRAS,* and *BRAF* genes, ALK, and PD-L1 expression in a series of Brazilian patients diagnosed with LUAD predominantly recruited from centers in southern Brazil. Molecular analysis of the *EGFR, KRAS, NRAS,* and *BRAF* genes was performed by next-generation sequencing using DNA extracted from tumor tissue. Immunohistochemistry was used to detect ALK and PD-L1 expression.

RESULTS Analysis of 619 tumors identified *KRAS* mutations in 189 (30.2%), *EGFR* mutations in 120 (19.16%), and *BRAF* mutations in 19 (3%). Immunohistochemistry demonstrated ALK and PD-L1 expression in 4% and 35.1% of patients, respectively.

CONCLUSION To our knowledge, this is the first study investigating the molecular epidemiology of patients with LUAD from southern Brazil and the largest assessing the frequency of multiple predictive biomarkers for this tumor in the country. The study also reveals a distinct mutation profile compared with data originating from other regions of Brazil.

INTRODUCTION

Lung cancer (LC) is the leading cause of cancer mortality worldwide and responsible for 1.7 million deaths every year.[1] In Brazil, the National Cancer Institute estimated there would be 31,270 new patients with LC from 2018 to 2019, accounting for the second most common tumor type in the country. It is the leading cause of deaths among men, ahead of prostate cancer, and the second leading cause among women, only behind breast cancer. In southern Brazil in 2018, 5,350 and 3,110 new cases were estimated in men and women, respectively, which makes LC the third most frequent cancer in the region.[2]

Non–small-cell LC (NSCLC) accounts for approximately 85% of pulmonary neoplasm diagnoses.[3,4] Effective treatments remain scarce, considering that the 5-year survival rate does not reach 20%, even in countries such as the United States.[5] In Brazil, this number is even lower, estimated at 16%.[6]

The use of predictive biomarkers allows therapeutic decisions to be based on tumor molecular profile.[7] For instance, certain somatic changes in the *EGFR, ALK, ERBB2,* and *BRAF* genes are substantial targets for tyrosine kinase inhibitors (TKIs).[8] In addition, new treatments for NSCLC using immune checkpoint inhibitors have recently been approved.[9] Its prescription depends on the expression of certain biomarkers on the tumor cell surface, such as the programmed death-ligand 1 (PD-L1) protein, a molecule in which the binding to its programmed death-1 receptor on T cells allows immune escape and tumor cell proliferation. The use of anti–programmed death-1/PD-L1

ASSOCIATED CONTENT

Data Supplement

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on July 30, 2019 and published at ascopubs.org/journal/jgo on September 18, 2019: DOI https://doi.org/10.1200/JGO.19.00174

drugs blocks such binding and reactivates the patient's immune response.[10]

Although the molecular profile of predictive biomarkers in LUAD is already well documented in Europe, the United States, and some regions of Asia, there are few studies exploring these data in Latin America. In Brazil, only a few reports have been published since 2012, and all were essentially restricted to the southeast region. Thus, these data may not be representative of all regions in Brazil, given the differences in ancestry according to regions.[11,12]

On the basis of this information, the main goal of this study was to investigate the frequency of somatic alterations in *EGFR, KRAS, NRAS,* and *BRAF* genes by next-generation sequencing (NGS), as well as ALK and PD-L1 expression in a series of Brazilian patients diagnosed with LUAD. To our knowledge, this is the first study to include a large number of patients who were tested by a biomarker panel in southern Brazil. These results might be important for new public policies in the treatment of LUAD.

MATERIAL AND METHODS

Study Population

This was a retrospective study conducted by the Precision Medicine Program of the Hospital de Clínicas de Porto Alegre (HCPA) in Brazil, which enrolled a case series of patients with LUAD who underwent molecular testing from September 2016 to January 2019.

Samples from 619 individuals were obtained from different hospitals and clinics distributed in 22 centers located in the three states of the southern region of Brazil: Rio Grande do Sul (N = 516), Santa Catarina (N = 24), and Paraná (N = 74). The five remaining patients were obtained from Rio de Janeiro. All included patients had confirmed adenocarcinoma histology. The diagnostic slides and formalin-fixed, paraffin-embedded tissue blocks were retrieved and reviewed by pathologists with expertise in LC. This project was approved by the HCPA Research Ethics Committee (No. 18-0099) and registered under the Certificate of Presentation for Ethical Appreciation (No. 83557418.5.0000.5327).

Tumor Selection and DNA Extraction

For all patients, 10-μm thick sections representative of the tumor tissue were cut, and regions with a higher percentage of tumor cells were selected for DNA extraction. DNA from the tissue samples was extracted using the ReliaPrep FFPE gDNA Miniprep System (Promega, Madison, WI) according to the manufacturer's recommendations. After extraction, the DNA samples were quantified using the fluorescence method (Qubit 2.0 Fluorometer, Invitrogen, Carlsbad, CA), which provides an accurate DNA quantification.

Of the 799 NSCLC samples that were received by the Precision Medicine Program of the HCPA, 619 (77.4%) were considered suitable for NGS analysis on the basis of tumor cell content, DNA amount, and purity.

Molecular Analysis by NGS

Molecular analysis of the *EGFR, KRAS, NRAS,* and *BRAF* genes was performed with an NGS platform (Ion Torrent PGM, server version 5.0; ThermoFisher Scientific, Waltham, MA). We used an AmpliSeqTM customized panel (ThermoFisher Scientific) to identify mutations in the *EGFR* gene (exons 18 to 21), *KRAS* (exons 2 and 3), *NRAS* (exons 2 and 3), and *BRAF* (exons 11 and 15). Data were analyzed on the bioinformatics platform Ion Torrent Suite and Ion Reporter version 5.0 with a minimum depth of 500x. Sequences NM_005228.3 (*EGFR*), NM_0033360.3 (*KRAS*), MM_002524.3 (*NRAS*), and NM_004333.4 (*BRAF*) were used as references. The test was performed using research-use-only reagents with internal validation.

PD-L1 and ALK Analysis

For PD-L1 analysis, which included 202 patients, the immunohistochemistry (IHC) expression scores were based on the Interpretation Guide for Window PD-L1 (SP263) Assay Staining of NSCLC 2017 (Roche, Basel, Switzerland). The analysis considers the overall percentage of positive neoplastic cells in the sample of any intensity above the eventual observed background staining in the negative control slide showing a cytoplasmic membrane pattern (circumferential, discontinuous, or basolateral). A negative control slide with an antibody provided in the PD-L1 kit was used for all patients. A tissue sample with PD-L1 positivity was also used as a positive control. Positivity in immune cells was not considered for determining scores in the Window PD-L1 test.

Similarly, for the ALK assay, paraffin sections of the lesions from 350 patients were submitted to the IHC technique. Tissue sections were stained with the ALK antibody clone D5F3 (Roche), which is able to recognize ALK fusion proteins and EML4-ALK variant expression. The slides were analyzed using the Ventana BenchMark XT Automated System (Roche). All reactions had negative and positive controls on the blade itself.

Statistical Analysis

To verify any possible association of the predictive biomarkers with gender, age, and whether there was any statistically significant difference between our study and others performed in the Brazilian population, we performed χ^2 tests. The results were considered statistically significant when the P was $< .05$.

RESULTS

NGS Analysis of EGFR, KRAS, NRAS, and BRAF Genes

NGS results of 619 tumors revealed 189 patients (30.2%) with mutations in *KRAS,* 120 (19.16%) with mutations in *EGFR*, and 19 (3%) with mutations in *BRAF*. No *NRAS* mutations were identified, and seven patients harbored double mutations. In 298 patients (47.6%), we did not detect any alteration using our NGS panel (Fig 1).

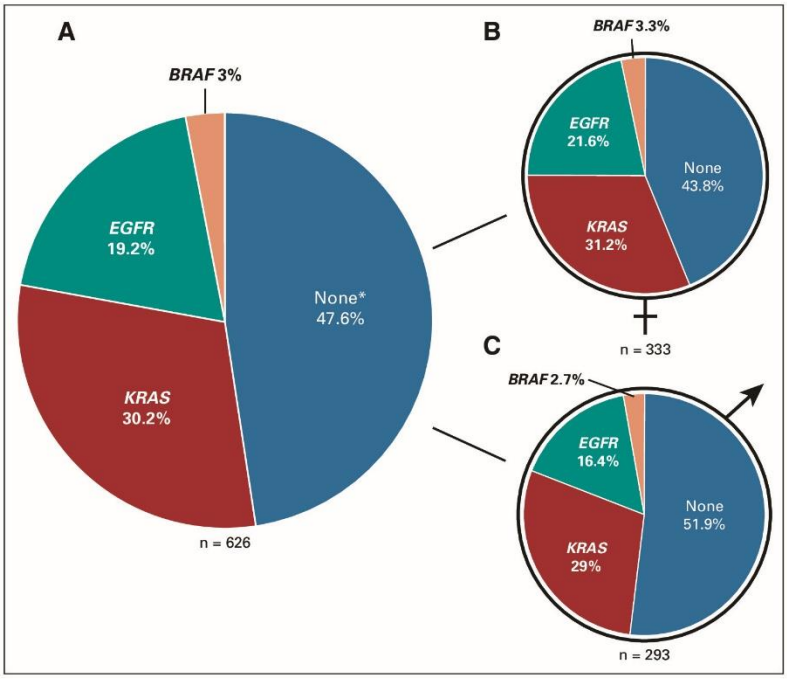


FIG 1. Frequency of somatic mutations in *EGFR*, *KRAS*, and *BRAF* in lung adenocarcinoma tumors from patients in southern Brazil. (A) Female and male patients, (B) only female patients, and (C) only male patients. (*) Not detected using our next-generation sequencing panel. It does not exclude the presence of alterations in other driver genes.

As expected, the frequency of *EGFR* mutations was 5.3% higher in women compared with men; however, this difference was not statistically significant (P = .096). In contrast, the frequency of patients with no mutations detected by our panel was higher in men, with a statistically significant difference (P = .044; Fig 1B and 1C; Data Supplement).

Among patients with *KRAS* mutations, p.(Gly12Cys) was the most frequent, accounting for 37% of the mutations found in this gene. This alteration was followed by p.(Gly12Val) and p.(Gly12Asp), found in 43 (22.75%) and 30 (15.9%) patients, respectively. Another 29 mutations were registered in exon 2 (15.3%). *KRAS* mutations in codon 61 of exon 3 were less frequently found in 17 patients (8.9%; Data Supplement).

Regarding the effect of *EGFR* mutations on TKI response, the majority was classified as sensitive (n = 98; 81.67%). Most of the mutations were located in exon 19 (n = 57; 47.50%), including the common deletion p.(Glu746_Ala750del; n = 36; 30.00%), followed by the point mutation p.(Leu858Arg; n = 35; 29.17%) in exon 21. Resistance alterations were restricted to exon 20 (n = 15; 12.50%). We also reported seven (5.83%) missense mutations of unknown clinical significance, including four in exon 19 and two in exon 21, and one complex mutation p.(Glu709_Thr710delinsAsp) in exon 18. In exon 20, there were 12 (10%) in-frame insertions and two (1.7%) duplications. Interestingly, one patient with no progression of disease was diagnosed with the p.(Thr790Met) mutation at a low frequency (3.2% of 883 sequence reads), which might be a case of primary resistance to first- and second-generation TKIs (Table 1; Fig 2).

Finally, among the 19 *BRAF* mutations, 13 (68.42%) were the well-known p.(Val600Glu) in exon 15. Another four mutations were located in exon 15, and two were located in exon 11, which also belongs to the kinase domain of the BRAF protein.

Although positive results of all predictive biomarkers were more commonly found in elderly patients (≥ 60 years of age), only *KRAS* had a statistically significant difference (P = .044; Data Supplement). The mean age at diagnosis of patients with *KRAS*, *EGFR*, and *BRAF* alterations was 67.5, 66.6, and 61.0 years, respectively.

PD-L1 and ALK IHC Analysis

PD-L1 staining by IHC was positive in 71 of the 202 analyzed tumors (35.1%). Approximately one fifth of tumors

TABLE 1. *EGFR* Mutations Identified in This Series of Patients and Grouped According to Their Sensitivity to Tyrosine Kinase Inhibitors

Sensitivity and Exon	Mutation	No. of Patients	Frequency. %
Sensitive			
Exon 18		5	4.17
	p.(Gly719Ala)	4	3.33
	p.(Gly719Ser)	1	0.83
Exon 19		57	47.50
	p.(Glu746_Ala750del)	36	30.00
	p.(Glu746_Arg748del)	3	2.50
	p.(Glu746_Thr751del)	1	0.83
	p.(Glu746_Thr751delinsAla)	1	0.83
	p.(Glu746Valfs*11)	2	1.67
	p.(Leu747_Pro753delinsSer)	2	1.67
	p.(Leu747_Ser752del)	1	0.83
	p.(Leu747_Thr751del)	6	5.00
	p.(Leu747fs)	1	0.83
	p.(Leu747Glnfs*16)	1	0.83
	p.(Thr751Lysfs*9)	1	0.83
	p.(Ser752_Ile759del)	1	0.83
	p.(Lys754fs)	1	0.83
Exon 20		1	0.83
	p.(Ser768Ile)[a]	1	0.83
Exon21		35	29.17
	p.(Leu858Arg)	35	29.17
Total		98	81.67
Resistant			
Exon 20		15	12.50
	p.(Met766_Ala767insAlaSerVal)	3	2.50
	p.(Ala767_Ser768insSerValAsp)	1	0.83
	p.(Ser768_Asp770dup)	2	1.67
	p.(Ser768_Val769insValAspAsn)	1	0.83
	p.(Val769_Asp770insCysVal)	1	0.83
	p.(Asp770_Asn771insGly)	1	0.83
	p.(Asp770_Asn771insGlyPhe)	1	0.83
	p.(Asp770_Asn771insPro)	1	0.83
	p.(Asn771_Pro772insProHis)	1	0.83
	p.(Asn771_Pro772insThr)	1	0.83
	p.(His773_Val774insAsnProTyr)	1	0.83
	p.(Thr790Met)	1	0.83
Total		15	12.50
Unknown			
Exon 18		1	0.83
	p.(Glu709_Thr710delinsAsp)	1	0.83

(Continued on following page)

TABLE 1. *EGFR* Mutations Identified in This Series of Patients and Grouped According to Their Sensitivity to Tyrosine Kinase Inhibitors (Continued)

Sensitivity and Exon	Mutation	No. of Patients	Frequency. %
Exon 19		4	3.33
	p.(Ala750Pro)	1	0.83
	p.(Ser752Phe)	3	2.50
Exon 21		2	1.67
	p.(Leu861Arg)	1	0.83
	p.(Leu861Gln)	1	0.83
Total		7	5.83

[a]Limited clinical data demonstrating the efficacy of TKIs.

(21.3%) had moderate expression, with positive staining noted between 1% and fewer than 50% of the cells. High expression of PD-L1, which is indicated by positive staining in 50% or more of the cells, was observed in 28 tumors (13.8%). ALK staining by IHC was positive in 14 of the 350 patients analyzed (4%), most of them identified in women (78.6%).

DISCUSSION

Although the prevalence of predictive biomarkers for molecular-targeted treatment in NSCLC has been established in several continents, including Asia, Europe, and North America, only a few studies have been conducted in Latin America. Furthermore, most of the studies characterizing samples from patients with LC in Brazil have recruited patients from the southeast, which limits the true understanding of the scenario in the entire country. In accordance with this assumption, in our study, we identified a particular frequency pattern of predictive biomarkers compared with that observed in other Brazilian regions.

To our knowledge, the frequency of *EGFR* mutations (19.16%) was the lowest ever reported when considering

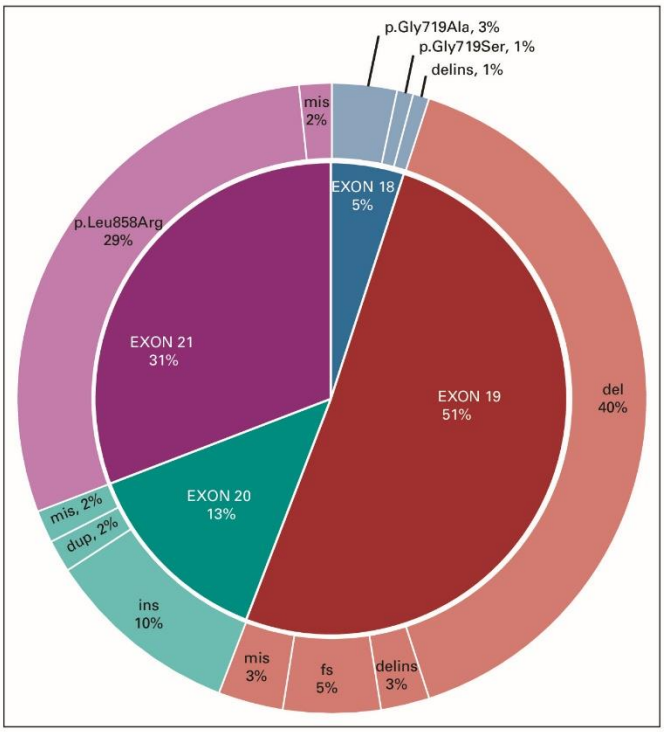


FIG 2. *EGFR* mutation frequencies according to their location in the gene. Del, deletions; delins, insertions/deletions; dup, duplications; fs, frameshifts; ins, insertions; mis, missense mutations.

TABLE 2. Comparison of Our Data With Four Previous Studies Conducted in Brazil

Alteration	Andreis (this study)	First Author, No (%); *P*			
		Bacchi[13]	de Melo[14]	Palacio[15]*	Leal[16]
EGFR variants (overall)	120 (19.1)	**64 (30.9); 6.9 × 10^{-4}**	27 (20.6); .704	**857 (25.5%); 7.5 × 10^{-4}**	108 (23.9); .058
EGFR – Exon 18	5 (4.1)	5 (7.9); .298	1 (3.7); .912	NA	2 (1.8); .311
EGFR – Exon 19 del/ins	57 (47.5)	38 (59.3); .125	8 (29.6); .091	NA	52 (48.1); .922
EGFR - L858R	35 (29.1)	17 (26.5); .708	3 (11.1); .052	NA	32 (29.6); .938
EGFR – Exon 20 del/ins plus T790M	15 (12.5)	3 (4.6); .089	1 (3.7); .184	NA	11 (10.2); .582
EGFR - unknown significance	8 (6.6)	1 (1.5); .089	**14 (51.8); 2.7 × 10^{-9}**	NA	11 (10.2); .337
KRAS variants (overall)	189 (30.2)	**30 (14.5); 1 × 10^{-5}**	33 (26.4); .396	NA	**90 (20.2); 2.7 × 10^{-4}**

NOTE. Bold font indicates values with $P < .05$ using the χ^2 test.
Abbreviations: del/ins, insertions and deletions; T790M, the p.(Thr790Met) point mutationL858R, p.(Leu858Arg) point mutation; NA, not applicable.
*Data on specific *EGFR* variants and *KRAS* were not available in this study.

those observed in four previous studies performed in Brazil (Table 2; Data Supplement).[13-16] In agreement with this observation, it is expected that the proportion of *EGFR*-mutated LUAD would decrease as the proportion of European ancestry increases in a given population. Several studies conducted in Europe showed a frequency of *EGFR* mutation of approximately 10%.[17-19]

In line with these observations, the prevalence of European ancestry in southern Brazil has been estimated at 80% to 90%, the highest among all Brazilian regions. African and Amerindian populations have a lower but no less important contribution.[11,12] Studies comparing the human leukocyte antigen allelic diversity in the region also confirmed a high similarity to Europeans and a significant difference from Asians or even Brazilian indigenous populations.[20-23] Together, this information supports the reason why the frequency of *EGFR* mutations in southern Brazil is lower than other regions of the country, but still higher than in Europe.

Considering the work of Bacchi et al,[13] which showed *EGFR* mutation in 30.9% of patients, it is important to note that the majority of the patients were from southeast and northeast Brazil. A recent and larger study performed in southeast Brazil also found a higher frequency of mutations in the gene (23.9%), but this difference was not statistically significant compared with our data ($P = .058$). Although European ancestry was predominant in this series of patients, a relatively high proportion of them had Asian ancestry (7.3%),[13] a population in which the frequency of *EGFR* mutations may reach 60%.[16,24] Although de Melo et al[14] found a percentage of patients with *EGFR*-mutated disease similar to ours, more than half (51.8%) were variants of unknown therapeutic impact compared with only 6.6% in our study ($P = 2.7 \times 10^{-9}$). Finally, the largest published study including data from *EGFR* molecular testing in Brazil (N = 3,364) also identified a higher and statistically significant mutation frequency (25.5%; $P = 7.5 \times 10^{-4}$; Table 2; Data Supplement).[15] Interestingly, a study conducted in Uruguay, a country geographically close to the southern region of Brazil, showed similar results regarding *EGFR* mutations (18.3%). The proportion of actionable alterations in exon 19 was similar (48.7% *v* 47.5%), but the missense substitution p.(Leu858Arg) in exon 21 was lower (22% *v* 29.17%; χ^2,1.025; $P = .311$). Our proportion of resistance mutations in exon 20 was also slightly higher (8.5% *v* 12.5%; χ^2; 0.646; $P = .422$).[25]

Previous reports showed a statistically significant association between female patients and a higher prevalence of *EGFR* mutations,[14,16,17,26] which was not observed in our participants (Data Supplement). In addition, we reported six rare *EGFR* mutations of unknown therapeutic impact: p.(Glu709_Thr710delinsAsp), p.(Ala750Pro), p.(Ser752Phe), p.(Leu861Arg), and p.(Leu861Gln), and p.(Ser768Ile). The efficacy of TKIs on tumors harboring these mutations is not well established, but some case reports suggest a benefit from treatment with first-generation TKIs.[27-32]

We also found *EGFR* alterations co-occurring in the same tumor in four patients. Two had the p.(Gly719Ala) coexisting with the p.(Ser768Ile) and the p.(Leu861Gln). Two other patients were carriers of two sensitive mutations in exon 19—p.(Glu746Valfs*11) plus p.(Lys754fs) and p.(Leu747Glnfs*16) plus p.(Thr751Lysfs*9). Another three patients with coexisting resistance mutations were also reported. One had the *EGFR* alteration in exon 20 p.(Asp770_Asn771insGly) plus *BRAF* p.(Val600Glu), another had a double mutation in *KRAS* (p.Gly13Cys plus p.Gln61His), and one had the *KRAS* p.(Gly12Cys) plus the *BRAF* p.(Asp594Asn). Only p.(Gly719Ala) plus p.(Ser768Ile) was already reported in a Brazilian patient.[14] Our data are in accordance with other studies conducted in Brazil, which revealed that all *EGFR* and *KRAS* alterations were mutually exclusive.[13,14,16,33]

Regarding *KRAS* alterations, some studies reported mutations in tumors diagnosed at more advanced ages,[26,34]

which were confirmed in our participants (Data Supplement). The general frequency of such alterations identified in our study is notable (30.2%) and significantly higher than that observed in two previous Brazilian cohorts (Table 2; Data Supplement).[13,16] However, it is important to note that these studies used Sanger sequencing for the molecular analysis, a method that requires the presence of the mutant allele in at least 15% to 20% of tumor DNA. For comparison purposes, if we had excluded the positive patients with a mutant allele frequency below 15% and 20%, our *KRAS* mutation percentage would decrease from 30.2% to 18.21% and 15.49%, respectively. For *EGFR* mutations, the frequencies in these two scenarios also decreased, ranging from 14.85% to 13.57% when excluding samples with mutant allele frequencies less than 15% and less than 20%, respectively. These numbers are closer to the previous studies cited earlier and could indicate that our NGS panel is more sensitive in detecting these driver mutations in LUAD. In addition, actionable mutations would have been missed in our study if we used a cutoff of allele frequency at 15%, including seven sensitive deletions in exon 19, four p.(Leu858Arg) substitutions in exon 21 and five resistance mutations, four insertions in exon 20, and one p.(Thr790Met) point mutation.

We next focused on biomarkers with a lower mutation prevalence. To our knowledge, our study is the first to report *BRAF* mutations in Brazilian patients with LUAD, found in 3% of the participants. This prevalence is in agreement with results in non-Latin populations.[33,35,36] The European Medicines Agency and US Food and Drug Administration recently approved the combined use of dabrafenib and trametinib for the treatment of patients with NSCLC harboring the *BRAF* p.(Val600Glu) mutation.[37]

In Brazil, the only study to assess ALK expression was performed in the northeast region.[38] The authors found a much higher prevalence of ALK-positive tumors (13.3% compared with 4% in our study).[38] A recent report that included samples from nine countries in Latin America found that the prevalence of *ALK* rearrangements ranged from 4.1% (Colombia) to 10.8% (Peru). Argentina and Uruguay, which are more geographically close to the southern region of Brazil, had a frequency of 5.4% and 4.4%,[39] respectively, which is similar to our findings. Co-occurrence of *ALK* rearrangements with *EGFR* or *KRAS* mutations is rare, with only a few reports in the literature.[26,40-46] We found two individuals with positive IHC for ALK protein who also harbored the *KRAS* variants p.(Gly12Val) and p.(Gly12Asp). Ulivi et al[47] reported that patients with *EML4-ALK* translocations co-occurring with *KRAS* point mutations had decreased responsiveness to crizotinib.

Alves da Silva et al[38] evaluated the PD-L1 expression in LC patients from the northeast region of Brazil and observed a similar proportion of positivity (40.5% v 35.1% in our study). The prevalence of tumors scoring from 1% to 50% or less and 50% or more was also similar between both studies (24% *v* 21.2% and 16.5% *v* 13.8%, respectively). These data reveal a much lower proportion of tumors with a staining score of 50% or more, due to a larger case series that had found high PD-L1 expression in 30.2% of the patients.[48]

Our study has some limitations. Because of its retrospective approach and the use of data from different private and public medical centers, we had to anonymize patients' information, including clinical data on treatment response, survival, and smoking status. However, we believe that this fact does not compromise the relevance of our results since the association between biomarkers and these clinical aspects has been widely studied.

In summary, to our knowledge, this is the first and largest study assessing the frequency of multiple predictive biomarkers for LUAD in Brazil. In addition, it reveals a unique pattern of mutation frequencies in different genes compared with data originating from other regions of the country. The frequency of *EGFR* mutations is the lowest found in Brazilian patients, possibly reflecting a higher proportion of individuals with European ancestry.[11,12] Our results also underscore the need to expand LUAD molecular testing in the Brazilian public health system, given that approximately 15% of patients with LUAD from the southern region would benefit from the use of TKIs. Araujo et al[49] estimated that fewer than half of the patients have their tumors submitted for molecular testing in Brazil. This number can be even lower in public health care institutions.[49] We expect that these data, together with results from other studies, will help to change this scenario and accelerate the implementation of new public policies for the treatment of LC in the country, on the basis of cost-effective analysis of our population.

AFFILIATIONS

[1]Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
[2]Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil
[3]Hospital Mãe de Deus, Porto Alegre, Brazil

CORRESPONDING AUTHOR

Tiago F. Andreis, MSc, Hospital de Clínicas de Porto Alegre, Rua Ramiro Barcelos, 2350, Porto Alegre/RS, Brazil 90035-903; e-mail: tiago.f.andreis@gmail.com.

PRIOR PRESENTATION

Presented in part at Next Frontiers to Cure Cancer, São Paulo, Brazil, May 10 to 12, 2018; and International Congress of Genetics, Foz do Iguaçu, Brazil, September 10 to 14, 2018.

SUPPORT

Supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico and Research Incentive Fund of Hospital de Clínicas de Porto Alegre (Grant no. 2018-0099). PhD scholarship awarded to T.F.A.

(Grant no. 140361/2018-9). All results presented here were performed by the Precision Medicine Program of Hospital de Clínicas de Porto Alegre, a diagnostic program that provides clinical services to AstraZeneca Brazil, Bristol-Myers Squibb, and Pfizer. Although the diagnostic tests were paid by these companies, none of them participated in or interfered with the design of the study, result analyses, and conclusions.

AUTHOR CONTRIBUTIONS

Conception and design: Tiago F. Andreis, Marina Siebert, Sandra Leistner-Segal, Vinícius Lorandi, Patricia Ashton-Prolla, Gabriel S. Macedo

Administrative support: Jane M. Ulbrich

Provision of study materials or patients: Sandra Leistner-Segal, Jane M. Ulbrich, Luis F.R. Rivero, Patricia Ashton-Prolla

Collection and assembly of data: Tiago F. Andreis, Bruno S. Correa, Fernanda De-Paris, Marina Siebert, Jane M. Ulbrich, Luis F.R. Rivero, Francine H. De Oliveira, Patricia Ashton-Prolla, Gabriel S. Macedo

Data analysis and interpretation: Tiago F. Andreis, Fernanda S. Vianna, Sandra Leistner-Segal, Jane M. Ulbrich, Luis F.R. Rivero, Vinícius Lorandi, Patricia Ashton-Prolla, Gabriel S. Macedo

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jgo/site/misc/authors.html.

Vinícius Lorandi

Travel, Accommodations, Expenses: Astellas Pharma, MSD Oncology, Novartis, AstraZeneca, Genentech

Patricia Ashton-Prolla

Research Funding: AstraZeneca Brazil (Inst)

No other potential conflicts of interest were reported.

ACKNOWLEDGMENT

We thank the undergraduate students Julia Tsao Schein, from the Medicine School of Universidade Federal do Rio Grande do Sul, and Guilherme Danielski Viola, from the Institute of Biosciences of Universidade Federal do Rio Grande do Sul, for their collaboration.

REFERENCES

1. Fitzmaurice C, Allen C, Barber RM, et al: Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: A systematic analysis for the global burden of disease study. JAMA Oncol 3:524-548, 2017
2. Brasil: Estimativa 2018: Incidência de Câncer no Brasil. Rio de Janeiro - RJ, Brazil, Ministério da Saúde, Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA), Coordenação de Prevenção e Vigilância, 2017 pp. 128
3. Lortet-Tieulent J, Soerjomataram I, Ferlay J, et al: International trends in lung cancer incidence by histological subtype: Adenocarcinoma stabilizing in men but still increasing in women. Lung Cancer 84:13-22, 2014. https://www.inca.gov.br/sites/ufu.sti.inca.local/files//media/document//estimativa-incidencia-de-cancer-no-brasil-2018.pdf
4. Pirker R, Filipits M: Personalized treatment of advanced non-small-cell lung cancer in routine clinical practice. Cancer Metastasis Rev 35:141-150, 2016
5. SEER: Cancer Stat Facts: Lung and Bronchus Cancer 5-year survival rates (2009-2015), in Surveillance E, and End Results (SEER) Program, National Cancer Institute, DCCPS, Surveillance Research Program (ed), 2019. https://seer.cancer.gov/statfacts/html/lungb.html
6. Polato CPB, Bonfante MDS, Andrae IG, et al: Análise de sobrevida em pacientes com câncer de pulmão tratados no Sistema Único de Saúde no Brasil entre 2002 e 2003 [in Portuguese]. Cad Saude Colet 21:173-181, 2013
7. Politi K, Herbst RS: Lung cancer in the era of precision medicine. Clin Cancer Res 21:2213-2220, 2015
8. Pao W, Girard N: New driver mutations in non-small-cell lung cancer. Lancet Oncol 12:175-180, 2011
9. Rizvi NA, Mazières J, Planchard D, et al: Activity and safety of nivolumab, an anti-PD-1 immune checkpoint inhibitor, for patients with advanced, refractory squamous non-small-cell lung cancer (CheckMate 063): A phase 2, single-arm trial. Lancet Oncol 16:257-265, 2015
10. Herbst RS, Soria JC, Kowanetz M, et al: Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. Nature 515:563-567, 2014
11. Ruiz-Linares A, Adhikari K, Acuña-Alonzo V, et al: Admixture in Latin America: Geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. PLoS Genet 10:e1004572, 2014
12. Chacón-Duque JC, Adhikari K, Fuentes-Guajardo M, et al: Latin Americans show wide-spread Converso ancestry and imprint of local Native ancestry on physical appearance. Nat Commun 9:5388, 2018
13. Bacchi CE, Ciol H, Queiroga EM, et al: Epidermal growth factor receptor and KRAS mutations in Brazilian lung cancer patients. Clinics (São Paulo) 67:419-424, 2012
14. de Melo AC, Karen de Sá V, Sternberg C, et al: Mutational profile and new IASLC/ATS/ERS classification provide additional prognostic information about lung adenocarcinoma: A study of 125 patients from Brazil. Oncology 89:175-186, 2015
15. Palacio S, Pontes L, Prado E, et al: *EGFR* mutation testing: Changing patterns of molecular testing in Brazil. Oncologist 24:e137-e141, 2019
16. Leal LF, de Paula FE, De Marchi P, et al: Mutational profile of Brazilian lung adenocarcinoma unveils association of EGFR mutations with high Asian ancestry and independent prognostic role of KRAS mutations. Sci Rep 9:3209, 2019
17. Gahr S, Stoehr R, Geissinger E, et al: EGFR mutational status in a large series of Caucasian European NSCLC patients: Data from daily practice. Br J Cancer 109:1821-1828, 2013
18. Lemée F, Bergoglio V, Fernandez-Vidal A, et al: DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. Proc Natl Acad Sci USA 107:13390-13395, 2010
19. Smits AJJ, Kummer JA, Hinrichs JWJ, et al: EGFR and KRAS mutations in lung carcinomas in the Dutch population: Increased EGFR mutation frequency in malignant pleural effusion of lung adenocarcinoma. Cell Oncol (Dordr) 35:189-196, 2012
20. Castro MS, Issler HC, Gelmini GF, et al: High-resolution characterization of 12 classical and non-classical HLA loci in southern Brazilians. HLA 93:80-88, 2019
21. Reis PG, Ambrosio-Albuquerque EP, Fabreti-Oliveira RA, et al: HLA-A, -B, -DRB1, -DQA1, and -DQB1 profile in a population from southern Brazil. HLA 92:298-303, 2018

22. Boquett J, Schüler-Faccini L, Jobim LF, et al: Self-assessment of color categories and its relationship with HLA profiling in Brazilian bone marrow donors. Biol Blood Marrow Transplant 21:1140-1144, 2015
23. Santos NP, Ribeiro-Rodrigues EM, Ribeiro-Dos-Santos AK, et al: Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. Hum Mutat 31:184-190, 2010
24. Shi Y, Au JS, Thongprasert S, et al: A prospective, molecular epidemiology study of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (PIONEER). J Thorac Oncol 9:154-162, 2014
25. Berois N, Touya D, Ubillos L, et al: Prevalence of EGFR mutations in lung cancer in Uruguayan population. J Cancer Epidemiol 2017:6170290, 2017
26. Lee B, Lee T, Lee SH, et al: Clinicopathologic characteristics of EGFR, KRAS, and ALK alterations in 6,595 lung cancers. Oncotarget 7:23874-23884, 2016
27. Wu JY, Yu CJ, Chang YC, et al: 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 17:3812-3821, 2011
28. Wu SG, Chang YL, Lin JW, et al: Including total EGFR staining in scoring improves EGFR mutations detection by mutation-specific antibodies and EGFR TKIs response prediction. PLoS One 6:e23303, 2011
29. Ackerman A, Goldstein MA, Kobayashi S, et al: EGFR delE709_T710insD: A rare but potentially EGFR inhibitor responsive mutation in non-small-cell lung cancer. J Thorac Oncol 7:e19-e20, 2012
30. Kobayashi Y, Togashi Y, Yatabe Y, et al: EGFR exon 18 mutations in lung cancer: Molecular predictors of augmented sensitivity to afatinib or neratinib as compared with first- or third-generation TKIs. Clin Cancer Res 21:5305-5313, 2015
31. Chiu CH, Yang CT, Shih JY, et al: Epidermal growth factor receptor tyrosine kinase inhibitor treatment response in advanced lung adenocarcinomas with G719X/L861Q/S768I mutations. J Thorac Oncol 10:793-799, 2015
32. Leventakos K, Kipp BR, Rumilla KM, et al: S768I mutation in EGFR in patients with lung cancer. J Thorac Oncol 11:1798-1801, 2016
33. Carneiro JG, Couto PG, Bastos-Rodrigues L, et al: Spectrum of somatic EGFR, KRAS, BRAF, PTEN mutations and TTF-1 expression in Brazilian lung cancer patients. Genet Res 96:e002, 2014
34. Dacic S, Shuai Y, Yousem S, et al: Clinicopathological predictors of EGFR/KRAS mutational status in primary lung adenocarcinomas. Mod Pathol 23:159-168, 2010
35. Brustugun OT, Khattak AM, Trømborg AK, et al: BRAF-mutations in non-small cell lung cancer. Lung Cancer 84:36-38, 2014
36. Davies H, Bignell GR, Cox C, et al: Mutations of the BRAF gene in human cancer. Nature 417:949-954, 2002
37. Khunger A, Khunger M, Velcheti V: Dabrafenib in combination with trametinib in the treatment of patients with BRAF V600-positive advanced or metastatic non-small cell lung cancer: Clinical evidence and experience. Ther Adv Respir Dis 12:1753466618767611, 2018
38. Alves da Silva AV, Martins Neto F: The frequency of high PD-L1 expression is low in lung adenocarcinoma patients from northeast Brazil. Surg Exp Pathol 2:4, 2019
39. Arrieta O, Cardona AF, Bramuglia G, et al: Molecular epidemiology of ALK rearrangements in advanced lung adenocarcinoma in Latin America. Oncology 96:207-216, 2019
40. Soda M, Choi YL, Enomoto M, et al: Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 448:561-566, 2007
41. Zhang X, Zhang S, Yang X, et al: Fusion of EML4 and ALK is associated with development of lung adenocarcinomas lacking EGFR and KRAS mutations and is correlated with ALK expression. Mol Cancer 9:188, 2010
42. Kuo YW, Wu SG, Ho CC, et al: Good response to gefitinib in lung adenocarcinoma harboring coexisting EML4-ALK fusion gene and EGFR mutation. J Thorac Oncol 5:2039-2040, 2010
43. Rossing HH, Grauslund M, Urbanska EM, et al: Concomitant occurrence of EGFR (epidermal growth factor receptor) and KRAS (V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog) mutations in an ALK (anaplastic lymphoma kinase)-positive lung adenocarcinoma patient with acquired resistance to crizotinib: A case report. BMC Res Notes 6:489, 2013
44. Jürgens J, Engel-Riedel W, Prickartz A, et al: Combined point mutation in KRAS or EGFR genes and EML4-ALK translocation in lung cancer patients. Future Oncol 10:529-532, 2014
45. Yang JJ, Zhang XC, Su J, et al: Lung cancers with concomitant EGFR mutations and ALK rearrangements: Diverse responses to EGFR-TKI and crizotinib in relation to diverse receptors phosphorylation. Clin Cancer Res 20:1383-1392, 2014
46. Lee JK, Kim TM, Koh Y, et al: Differential sensitivities to tyrosine kinase inhibitors in NSCLC harboring EGFR mutation and ALK translocation. Lung Cancer 77:460-463, 2012
47. Ulivi P, Chiadini E, Dazzi C, et al: Nonsquamous, non-small-cell lung cancer patients who carry a double mutation of EGFR, EML4-ALK or KRAS: Frequency, clinical-pathological characteristics, and response to therapy. Clin Lung Cancer 17:384-390, 2016
48. Reck M, Rodríguez-Abreu D, Robinson AG, et al: Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. N Engl J Med 375:1823-1833, 2016
49. Araujo LH, Baldotto C, Castro G Jr, et al: Lung cancer in Brazil. J Bras Pneumol 44:55-64, 2018

CAPÍTULO V – RESULTADOS PARCIAIS PARA O MANUSCRITO II

Short communication em preparação: “***Mutações de significando clínico incerto em pacientes com câncer de pulmão de não pequenas células: decisão terapêutica e resposta a inibidores de tirosina-quinase.***”

INTRODUÇÃO

Aproximadamente 6% dos casos positivos para *drivers* oncogênicos em câncer de pulmão de não-pequenas células possuem alterações moleculares raras de significado clínico incerto envolvendo os genes analisados, incluindo: mutações em *EGFR* cuja a resposta a inibidores de tirosina quinases (ITQs) é desconhecida, ocorrência de mais de uma mutação em *EFGR* (duplo-mutados) e coocorrência de mutações em *EGFR* e *KRAS* (Andreis *et al.*, 2019). Tais alterações são um desafio para o tratamento do câncer de pulmão de não pequenas células (CPNPC), uma vez que dados clínicos são escassos devido a sua frequência relativamente baixa (Kohsaka et al., 2017).

Apesar de estudos clínicos demonstrarem que tumores com mutações raras em *EGFR* serem menos sensíveis ao tratamento ITQs de primeira geração, existe certo nível de consenso entre os clínicos quanto ao uso do medicamento diante desses casos (Harrison, Vyse and Huang, 2020). Um exemplo de avanços neste campo é a inclusão das mutações G719X, S768I e L861Q, presentes em menos de 5% dos casos de CPNPC, na lista de alterações que podem indicar benefício no uso de Afatinibe, um ITQ de segunda geração para o qual originalmente não havia tal indicação (Zhang et al., 2019).

Mutações em *KRAS* normalmente indicam baixa probabilidade de resposta a ITQs e podem sugerir um pior prognóstico (NCCN, 2022). Porém, casos de coocorrência de mutações no gene com *EGFR*, apesar de raros, já foram reportados e pouco se conhece quanto a sua resposta terapêutica a ITQs (Swami, 2021; Skoulidis e Heymach, 2019; Blakely et al., 2017).

Frente a ausência de diretrizes clínicas bem estabelecidas para condução de casos com alterações moleculares raras de significado clínico incerto em CPNPC, este trabalho objetivou verificar qual foi a escolha dos profissionais médicos no tratamento deste grupo de pacientes e, para aqueles que optaram pelo uso de ITQs, qual foi resposta após uso do medicamento.

MATERIAL E MÉTODOS

Este foi um estudo retrospectivo que analisou os dados de 990 tumores de pacientes diagnosticados com CPNPC entre setembro de 2016 e maio de 2020, provenientes do Programa de Medicina Personalizada do Hospital de Clínicas de Porto Alegre. As amostras foram submetidas ao sequenciamento de nova geração (NGS) nos equipamentos Ion Torrent PGM e Ion S5 (ThermoFisher Scientific) após enriquecimento e seleção de regiões-alvo com painel AmpliSeq (ThermoFisher Scientific) customizado, incluindo os genes *EGFR* (exons 18 a 21), *KRAS* (exons 2 e 3), *NRAS* (exons 2 e 3) e *BRAF* (exons 11 e 15). Os dados gerados foram analisados com a plataforma Ion Torrent Suite e Ion Reporter versão 5.0 (ThermoFisher Scientific). Os testes foram fomentados pelas empresas AstraZeneca, Astellas Pharma, MSD Oncology, Novartis e Genentech.

Foram selecionadas amostras com mutações raras de significado clínico incerto em *EGFR*, com mutações compostas (mais de uma mutação em *EGFR*) e com coocorrência de mutações em *EGFR* e *KRAS*. Em seguida, um questionário foi aplicado aos médicos responsáveis pelos pacientes, incluindo dados dos esquemas de tratamento utilizados e outras informações clínicas (Anexo I). Os dados clínicos de cada paciente foram então compilados a fim de avaliar a escolha terapêutica e, para aqueles cuja a escolha foi o uso de ITQs, qual foi a resposta ao tratamento. O projeto encontra-se aprovado no CAAE sob nº 83557418.5.0000.5327 e HCPA 2018-0099..

RESULTADOS PARCIAIS E DISCUSSÃO

Dados referentes ao tratamento de nove dos 25 pacientes selecionados foram obtidos a partir de do questionário aplicado e busca ativo junto aos oncologistas. As mutações e dados clínicos básicos dos participantes estão resumidos na Tabela 1. Destes, cinco possuem mutações em *EGFR* cuja resposta a ITQs é descrita como desconhecida, dois apresentam mutação composta em *EGFR* e dois apresentaram coocorrência de mutação em *EGFR* com *KRAS*. Curiosamente, o paciente 3 diagnosticado com a mutação p.(Thr790Met) em *EGFR* não havia sido tratado previamente, sugerindo que este seja um caso de resistência intrínseca (primária).

Tabela 1. Dados clínicos e moleculares dos pacientes incluídos no estudo.

ID	Idade ao diagnóstico	Sexo	Tabagismo	Histologia	Variantes			Frequência no TCGA (%)[2]	Significado clínico[3]
					Gene	Alteração na proteína (HGVS)	VAF[1] (%)		
1	63	M	S	Adenocarcinoma	*EGFR*	p.(Gly719Ala)	20.7	0.3	Tier IA
						p.(Ser768Ile)	SD[4]	SD[4]	Tier IA
2	64	M	N	Adenocarcinoma	*EGFR*	p.(Glu709_T710insAsn)	32.9	0	Tier IB
3	56	F	N	Carcinoma escamoso	*EGFR*	p.(Leu747_Thr753delinsSer)	60.5	0	Tier IA
						p.(Thr790Met)	29.6	0.3	Tier IA
4	67	F	N	Adenocarcinoma	*EGFR*	p.(Glu746_Ala750del)	12.9	2.4	Tier IA
					KRAS	p.(Gly12Asp)	16.1	1.7	Tier III
5	54	F	S	Adenocarcinoma	*EGFR*	p.(Ser752Phe)	9.7	0	Tier III
6	79	M	N	Adenocarcinoma	*EGFR*	p.(Met766Leu)	6.9	0	Tier III
7	52	F	S	Adenocarcinoma	*EGFR*	p.(Ala750Pro)	6.4	0	Tier III
8	60	M	N	Adenocarcinoma	*EGFR*	p.(Lys806Arg)	8.2	0	Tier III
					KRAS	p.(Gly12Val)	21.6	3.7	Tier III
9	58	M	S	Adenocarcinoma	*EGFR*	p.(Lys806Arg)	4.6	0	Tier III

1. *Variant allele frequency* (frequência do alelo variante).
2. Pan-Lung Cancer; Nature Genetics, 2016.
3. AMP/ASCO/CAP Consensus Recommendations, specific for Lung Cancer.
4. SD = Sem Dados.

Dentre os que optaram por ITQs, quatro utilizaram o fármaco de primeira geração (gefitinibe) e um de terceira geração (osimertinibe). Os quatro casos tratados com quimioterapia utilizaram carboplatina e pemetrexede. Na segunda linha, a escolha de quimioterapia foi a mais frequente (N = 5), as quais utilizaram esquemas diversos. O imunoterápico nivolumabe foi utilizado em um caso. Já ITQs não foram utilizados em segunda linha em nenhum dos pacientes (Tabela 2). A média de sobrevida global dos pacientes que receberam ITQs foi de 27,8 meses (11 meses a 42 meses), a maioria deles com resposta parcial. Já no grupo que recebeu quimioterapia, a média de sobrevida global foi de 10,7 meses (2 meses a 20 meses).

Tabela 2. Esquemas de tratamento e resposta terapêutica para os nove pacientes com dados disponíveis.

ID	Variante(s)	Sequenciamento do tumor	1ª linha de tratamento para doença metastática (resposta em meses)	Resposta	2ª linha de tratamento	Sobrevida global (dias)
1	p.(Gly719Ala) p.(Ser768Ile)	Após progressão para doença metastática[1]	Afatinibe (12.9)	RP > RP	Carboplatina + Paclitaxel + Bevacizumabe	1.153
2	p.(Glu709_T710insAsn)	No diagnóstico da doença	Carboplatina + Pemetrexede (0.8)	ND	Não	155
3	p.(Leu747_Thr753delinsSer) p.(Thr790Met)	No diagnóstico da doença	Osimertinibe (6.9)	PD	Carboplatina + Paclitaxel	332
4	p.(Glu746_Ala750del) p.(Gly12Asp)	No diagnóstico da doença	Gefitinibe (32.2)	RP > PD > DE	Gemcitabina	1.201
5	p.(Ser752Phe)	No diagnóstico da doença	Gefitinibe (45,1)	RP > DE	Não	1.353
6	p.(Met766Leu)	No diagnóstico da doença	Carboplatina + Pemetrexede (6.0)	PD	Nivolumabe	489
7	p.(Ala750Pro)	No diagnóstico da doença	Gefitinibe (3.7)	PD > RP	Carboplatina Paclitaxel	320
8	p.(Lys806Arg) p.(Gly12Val)	No diagnóstico da doença	Carboplatina + Pemetrexede (7.7)	RP > DE	Docetaxel	609
9	p.(Lys806Arg)	No diagnóstico da doença	Carboplatina + Pemetrexede (1.5)	PD	Não	47

RP = resposta parcial; DE = doença estável; PD = progressão da doença; ND = não disponível.

Dos nove pacientes, cinco receberam tratamento com ITQ e quatro quimioterapia na primeira linha, demonstrando uma heterogeneidade na escolha do tratamento frente a falta de diretrizes clínicas bem estabelecidas. Apesar dos dados indicarem um aparente benefício no uso de ITQs na maioria dos casos, nenhum dos dois com a mutação p.(Lys806Arg) recebeu esta categoria de medicamento. Curiosamente, o paciente 3 que carrega duas mutações que conferem sensibilidade ao ITQ osimertinibe, apresentou uma sobrevida global modesta (332 dias). Por outro lado, o paciente 1 que também apresentou duas mutações de sensibilidade teve uma resposta duradoura, com uma sobrevida global de 1.153 dias.

Chama atenção nesta análise preliminar que mesmo na ausência de diretrizes clínicas e evidências mais robustas, a maioria dos clínicos (5 dos 9 incluídos) optou pelo uso de ITQs para o tratamento dos pacientes com alterações moleculares de significado indeterminado. Espera-se que, quando concluído, este trabalho incentive a discussão e conscientize a comunidade quanto ao manejo de casos de CPNPC com mutações raras, compostas ou com mais de um *driver* alterado.

REFERÊNCIAS BIBLIOGRÁFICAS

Blakely, C. M., Watkins, T. B. K., Wu, W., Gini, B., Chabon, J. J., McCoach, C. E., McGranahan, N., Wilson, G. A., Birkbak, N. J., Olivas, V. R., Rotow, J., Maynard, A., Wang, V., Gubens, M. A., Banks, K. C., Lanman, R. B., Caulin, A. F., St John, J., Cordero, A. R., Giannikopoulos, P., Simmons, A. D., Mack, P. C., Gandara, D. R., Husain, H., Doebele, R. C., Riess, J. W., Diehn, M., Swanton, C. and Bivona, T. G. (2017) 'Evolution and clinical impact of co-occurring genetic alterations in advanced-stage EGFR-mutant lung cancers', Nat Genet, 49(12), pp. 1693-1704.

Harrison, P. T., Vyse, S. and Huang, P. H. (2020) 'Rare epidermal growth factor receptor (EGFR) mutations in non-small cell lung cancer', Semin Cancer Biol, 61, pp. 167-179.

Kohsaka, S., Nagano, M., Ueno, T., Suehara, Y., Hayashi, T., Shimada, N., Takahashi, K., Suzuki, K., Takamochi, K., Takahashi, F. and Mano, H. (2017) 'A method of high-throughput functional evaluation of', Sci Transl Med, 9(416).

NCCN, N. C. C. N. 2022. Lung cancer (version 3.2022).

Skoulidis, F. and Heymach, J. V. (2019) 'Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy', Nat Rev Cancer, 19(9), pp. 495-509.

Swami, N., Hwang, W. L., Guo, J. A., Hoffman, H., Abramowitz, M. C., Elbakouny, Z., Beltran, H., Chipidza, F., Choueiri, T., Pra, A. D., Huang, F., Kaochar, S., Kantoff, P., Kim, D. W., Kishan, A. U., Kobetz, E., Marinac, C., Mucci, L. A., Muralidhar, V., Pollack, A., Sanford, N. N., Schaeffer, E. M., Spratt, D. E., Zhao, S. G., Rebbeck, T. R., Nguyen, P. L., Feng, F. Y., Mahal, B. A. and Alshalalfa, M. (2021) 'Novel genomic signature predictive of response to immune checkpoint blockade: A pan-cancer analysis from project Genomics Evidence Neo-plasia Information Exchange (GENIE)', Cancer Genet, 258-259, pp. 61-68.

Zhang, T., Wan, B., Zhao, Y., Li, C., Liu, H., Lv, T., Zhan, P. and Song, Y. (2019a) 'Treatment of uncommon', *Transl Lung Cancer Res,* 8(3), pp. 302-316.

ANEXO I – Questionário aplicado no estudo

1. Nome do paciente:
2. Esquema de 1ª linha para tratamento, se aplicável:
3. Data de início do esquema 1ª, se aplicável(dd/mm/aa):
4. Data da progressão após 1ª linha, se aplicável (dd/mm/aa):
5. Esquema de 2ª linha para tratamento anti-neoplásico (NA caso não-aplicável):
6. Data de início do esquema 2ª, se aplicável (dd/mm/aa):
7. Data da progressão após 2ª linha, se aplicável (dd/mm/aa):
8. O paciente é tabagista (incluindo passivos):
9. Histologia:
10. Melhor performance status durante o tratamento (ECOG):
11. Performance status atual (ECOG):
12. Doença em estadiamento IV *upfront*:
13. Já recebeu tratamento abaltivo em sítio metastático:
14. RECIST 1.1:
15. Data de óbito, se aplicável (dd/mm/aa):
16. Outras informações relevantes:

CAPÍTULO VI – RESULTADOS PARCIAIS PARA O MANUSCRITO III

Manuscrito em preparação, em fase de análise dos dados e complementações: "***Sequenciamento de nova geração para avaliação do DNA tumoral circulante em câncer de pulmão não pequenas células em estadiamento IIIC não ressecáveis*.”**

INTRODUÇÃO

O câncer de traqueia, brônquios e pulmão (TBP) é a neoplasia que mais causa mortes no mundo, sendo responsável por quase dois milhões de óbitos a cada ano (Fitzmourice *et al*., 2019). No Brasil, a taxa de sobrevida estimada em 5 anos é de apenas 18% (INCA, 2022). Frente a alta letalidade desta doença, nos últimos 20 anos diversas drogas de alvo molecular foram desenvolvidas para o tratamento do câncer de pulmão de não pequenas células (CPNPC), o qual representa aproximadamente 80% dos casos diagnosticados de TBP (Kumarakulasinghe, 2015; Salgia *et al*., 2021; Singh et al., 2021).

Atualmente o órgão norte-americano *National Comprehensive Cancer Network* (NCCN), recomenda a análise de pelo menos 11 biomarcadores no CPNPC, incluindo mutações de pontos e/ou pequenas deleções/inserções nos genes *EGFR* e *KRAS*, fusões/rearranjos de *ALK, RET, ROS1* e *NTRK1/2/3*, exclusão do éxon 14 de *MET*, e a mutação V600E em *BRAF.* Marcadores emergentes incluem amplificações de *MET* e mutações em *ERBB2* (NCCN, 2022). As alterações citadas acima levam à ativação constitutiva das vias de Raf-MEK-ERK a jusante dos receptores de tirosina quinase (RTQs) (Arora & Scholar, 2005; Mok *et al*., 2017). Por esta razão, a maioria das drogas de alvo de molecular atualmente disponíveis para CPNPC são da classe de inibidores de tirosina quinase (ITQs).

Vários estudos demonstraram a superioridade das terapias de alvo molecular em relação aos tratamentos convencionais com quimio e radioterapia, com aumentos significativos na sobrevida livre de progressão e global, além de uma diminuição nos índices de efeitos adversos devido à alta toxicidade (Ai *et al*., 2018). Mesmo com tais avanços, a perda de resposta e consequentemente progressão da doença é observada de 10 a 14 meses após início do tratamento com ITQs (Wu & Shih, 2018). Apesar de serem compostos, em sua maioria, de uma população clonal dominante com um ou poucos *drivers* oncogênicos, tumores são doenças heterogêneas (Dagogo-Jack & Shaw, 2018; Kemper *et al*., 2015). A resistência à terapia emerge a partir da expansão de clones intrínsecos, pré-existentes em baixa frequência, que possuem alterações não-acionáveis (resistência primária) ou a partir de sub-clones com mutações adquiridas ao longo do tratamento (resistência secundária) (Lim & Ma, 2019).

O monitoramento da evolução clonal da neoplasia a partir de biópsias seriadas é a estratégia ideal para prevenir e minimizar o efeito da resistência ao tratamento. A biópsia líquida, que avalia o DNA tumoral circulante (ctDNA), surge como uma alternativa promissora, sendo uma técnica menos invasiva e muitas vezes a única possível em casos onde o tecido não está acessível (Guibert *et al*., 2019). Em países como o Brasil, onde com frequência o diagnóstico do CPNPC é feito já em estágio avançado, tal abordagem torna-se ainda mais desejável (Malta *et al*., 2007). O método permite uma melhor representação da heterogeneidade intratumoral, levando a uma melhor compreensão do perfil molecular da doença. Até mesmo em tumores com baixa frequência de mutações acionáveis, como nos carcinomas escamosos, a técnica poderia ser aplicada, principalmente em casos com baixa acessibilidade ao tecido e cuja a biópsia tecidual possa não ser representativa para detecção de porções não-escamosas (tumores adenoescamosos) (Nicholson, *et al*., 2021).

No Brasil, onde menos da metade dos pacientes elegíveis para o teste não recebem sequer diagnóstico molecular no tecido para mutações em *EGFR,* a biópsia líquida ainda dá seus primeiros passos (Palacio *et al*., 2018). A técnica está acessível essencialmente para pacientes que podem custear o teste, sendo a grande maioria restrita a metodologias baseadas em PCR em tempo real (RT-PCR) e suas derivações, como PCR digital (ddPCR). Esta abordagem possui como principal limitação a possibilidade de detecção de poucos alvos já pré-estabelecidos (Wu & Shih, 2018).

No âmbito do SUS, onde existem poucas opções para o uso de terapia-alvo, o uso de ctDNA pode compor um importante biomarcador para o monitoramento de doença residual mínima, sendo capaz de indicar a resposta ao tratamento, mesmo para aqueles a base de radio e quiometerapia. Tal informação poderia orientar a tomada de decisão terapêutica, auxiliando na estratificação de risco e priorizando o atendimento daqueles pacientes com maior chance de progressão.

Este é um dos primeiros trabalhos no Brasil a avaliar, com a técnica de NGS, o ctDNA de pacientes com CPNPC e, pelo nosso conhecimento, o primeiro a utilizar tal abordagem para avaliar pacientes atendidos pelo sistema único de saúde (SUS). Além de explorar potenciais biomarcadores de predição terapêutica e prognósticos, o estudo pode estimular a utilização da técnica para futuros projetos designados a elucidar a

heterogeneidade tumoral em uma população ainda pouco conhecida e cujo o acesso ao diagnóstico de precisão é escasso.

MATERIAL E MÉTODOS

Seleção dos pacientes

Durante o período de maio/2020 a dezembro/2021, foram rastreados pacientes com câncer de pulmão atendidos pelo Serviço de Oncologia do Hospital de Clínicas de Porto Alegre (HCPA), Rio Grande do Sul/Brasil. Para o estudo, foram selecionados pacientes com diagnostico de CPNPC com estádio IIIC e irressecável, visto que este grupo apresenta menor acessibilidade a biópsias teciduais e que poderiam, em teoria, beneficiar-se do monitoramento de doença residual mínima após o tratamento e da descoberta de mutações que indiquem sensibilidade ao tratamento com drogas de alvo-molecular após a progressão para o estadiamento IV.

Inicialmente, apenas o subtipo adenocarcinoma seria incluído, porém devido ao baixo número de pacientes identificados amostral, optou-se por incluir também aqueles com histologia escamosa. As coletas foram realizadas em três momentos: logo após o diagnóstico (pré-tratamento), até 14 dias após tratamento padrão com quimio e radioterapia (pós-tratamento) e após progressão da doença (Figura 1). O termo de consentimento livre e esclarecido (TCLE) foi aplicado e assinado por cada um dos participantes antes da primeira coleta. O projeto foi aprovado pelo comitê de ética do HCPA e está inscrito sob o CAAE 70045017.7.0000.5327 e HCPA 2018-0121.

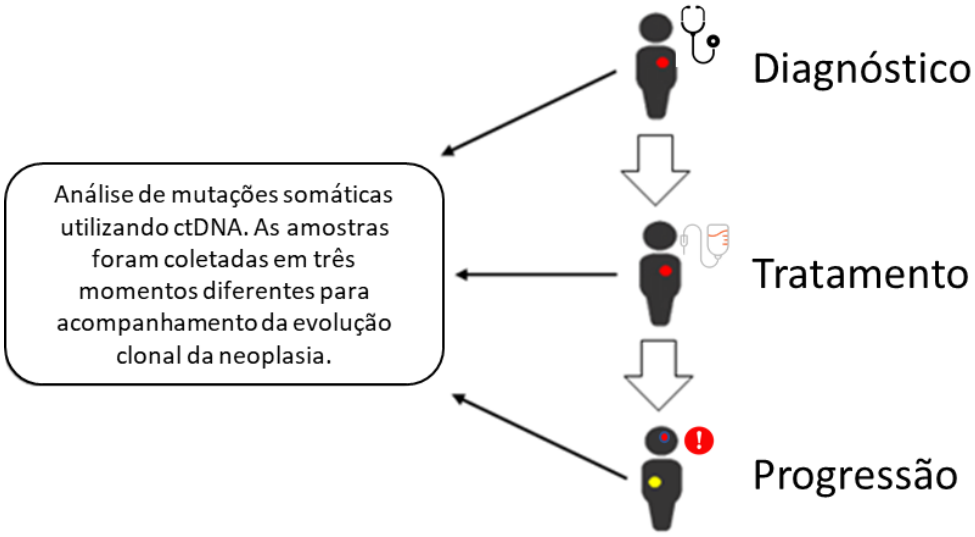


Figura 1. Ilustração do esquema de coleta elaborado para o estudo. No total nove pacientes diagnosticados com CPNPC IIIC foram acompanhados e tiveram suas amostras coletadas ao longo do estudo.

Coleta, processamento e extração do cfDNA

Foram realizadas 24 coletas de aproximadamente 8 mL de sangue periférico em tubo de ácido etilenodiamino tetra-acético (EDTA), os quais foram imediatamente acondicionados em gelo. O material foi pré-processado em um intervalo inferir a 2 h com centrifugação de 800 g a 4°C durante 10 min para separação do plasma e evitar a degradação enzimática do DNA livre circulante (cfDNA) e a contaminação com gDNA proveniente da lise de leucócitos. Após esta etapa, as amostras foram armazenadas a -20°C ou submetidas a uma segunda centrifugação de 7.900g a 4°C por 20 min para remoção de detritos remanescentes e subsequente extração.

A extração e isolamento do cfDNA foi realizada a partir de 4 mL de plasma utilizando o kit QIAamp MinElute ccfDNA (© Qiagen), conforme orientações do fabricante. A concentração do material extraído foi determinada no fluorômetro Qubit 3 (© Thermo Fisher Scientific) com kit Qubit dsDNA HS Assay (© Thermo Fisher Scientific) e imediatamente armazenado em -20°C até processamento.

Preparo das bibliotecas, enriquecimento dos alvos e sequenciamento

Dois protocolos foram utilizados para análise do ctDNA. O primeiro consiste em um painel customizado SureSelect XT-HS (© Agilent), otimizado para o processamento de amostras com baixa quantificação de DNA e que inclui *barcodes* moleculares para detecção de variantes em baixa frequência. As etapas foram seguidas conforme protocolo do fabricante, utilizando 17 uL de cfDNA como input inicial. Foram selecionados 70 genes, dentre aqueles associados a vias de reparo do DNA (supressores de tumor) e oncogenes, os quais estão frequentemente alterados em CPNPC, conforme consulta no *The Cancer Genomas Atlas* (Quadro 1).

Quadro 1. Lista de genes selecionados no painel utilizado neste estudo. Foram incluídas todas regiões codificantes (éxons) e 10 pb dos íntrons adjacentes a 3'e 5'.

ABCB1	*ATR*	*CUL3*	*MDC1*	*NOTCH4*	*RAD50*	*SETD2*
ABCB4	*BRAF*	*DMC1*	*MDM2*	*NRAS*	*RAD51*	*SMARCA4*
ABCC1	*BRCA1*	*EGFR*	*MET*	*NTRK1*	*RAD51B*	*SPEN*
ABCG2	*BRCA2*	*ERBB2*	*MLH1*	*NTRK2*	*RAD51C*	*STK11*
AKT1	*CCND1*	*ERBB4*	*MRE11A*	*NTRK3*	*RAD51D*	*TP53*
ALK	*CCNE1*	*ERCC1*	*MSH2*	*PDGFRA*	*RB1*	*TSC1*
ARID1A	*CDK4*	*KDR*	*MVP*	*PIK3CA*	*RBM10*	*TSC2*
ARID1B	*CDKN2A*	*KEAP1*	*NBN*	*PIK3R1*	*RET*	*U2AF1*
ARID2	*CHEK1*	*KRAS*	*NF1*	*POLQ*	*RIT1*	*XRCC2*
ATM	*CHEK2*	*MAP2K1*	*NOTCH1*	*PTEN*	*ROS1*	*XRCC3*

A qualidade da biblioteca resultante foi verificada com kit D1000 ScreenTape (© Agilent) no equipamento TapeStation 4200 (© Agilent). Uma vez qualificadas, seguiram para enriquecimento dos alvos incluídos do painel e nova verificação com kit D1000 HightSensitivity ScreenTape (© Agilent). Os produtos resultantes foram então quantificados no fluorômetro Qubit 3 (© Thermo Fisher Scientific) com kit Qubit dsDNA HS Assay (© Thermo Fisher Scientific). Finalmente, as amostras foram diluídas e submetidas a sequenciamento utilizando o kit MiSeq Reagente V3 (© Illumina) configurado em duas leituras de 151 pb no sistema Illumina MiSeq (© Illumina), conforme orientações do fabricante.

A segunda abordagem para análise do ctDNA utilizou o sistema Ion Torrent Genexus (© Thermo Fisher Scientific), com o painel Oncomine Precision Assay (©

Thermo Fisher Scientific). Este inclui 78 variantes em 50 genes, frequentemente identificadas em tumores sólidos, incluindo CPNPC (Quadro 2). As amostras de cfDNA já extraídas foram carregadas na plataforma automatizada de preparo de bibliotecas, sequenciamento e análise dos dados, conforme especificações do fabricante. Esta etapa foi executada em parceria com a Igenomix Brasil Laboratório de Medicina Genética LTDA.

Análise dos resultados

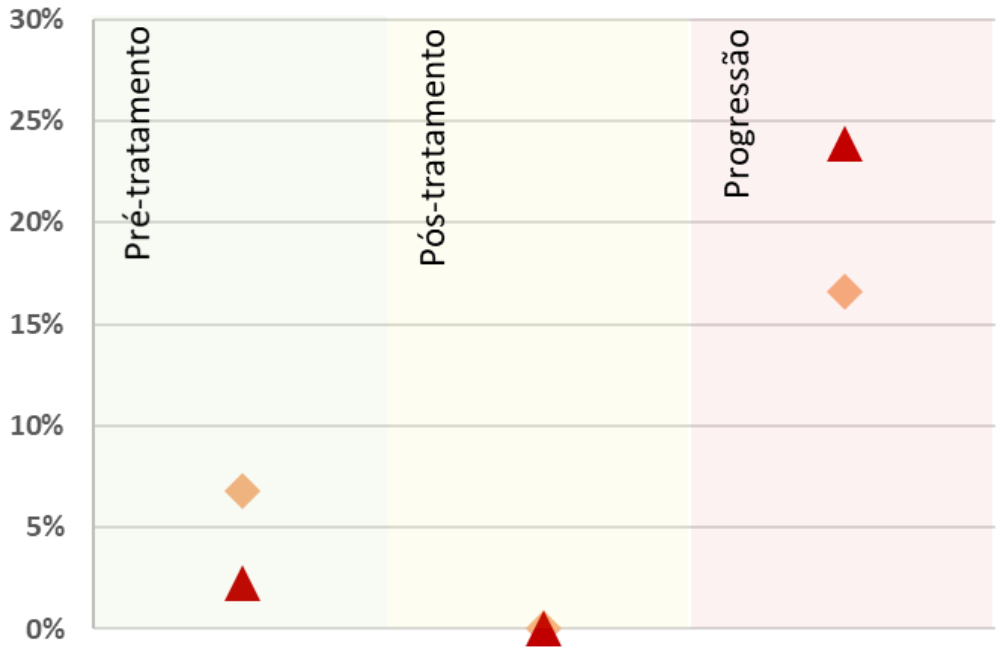
Uma vez gerados, os arquivos FASTq resultantes do sequenciamento foram analisados no software Agilent SureCall 4.2.2.3 (© Agilent). Dois métodos de análise foram utilizados, um contendo *barcodes* moleculares (MBC) e outro não. A introdução de MBCs permite a redução erros aleatoriamente gerados durante o sequenciamento de amostras com baixas concentrações de DNA, como no caso do cfDNA, reduzindo as taxas de falso-positivos. Ambos arquivos (VCFs) gerados foram então analisados manualmente para seleção de variantes. Foram mantidas alterações que não apresentassem viés de fita, baixa frequência ou ausência em bancos de dados populacionais e que estivessem ausentes ou descritas como de significado clínico incerto ou como patogênicas no banco de dados *ClinVar*. Variantes fora das regiões dos alvos selecionados ou que não contemplem sítios de processamento de RNA canônicos foram desconsideradas.

RESULTADOS PARCIAIS

Dos nove pacientes que estavam sendo acompanhados, cinco apresentaram progressão da doença antes da conclusão deste estudo. Um dos pacientes (paciente 6) recebeu dois ciclos adicionais de quimioterapia e também foi incluído nas análises, porém evoluiu à óbito antes de realizar nova coleta. A quantificação média de cfDNA extraído foi de 0,9 ng/µL (0,19 ng/µL – 2,81 ng/µL). Até o presente momento, 16 bibliotecas foram geradas a partir das amostras coletadas. O tamanho médio após o enriquecimento/captura das bibliotecas foi de 304 pb (256 pb – 356 pb), dentro dos parâmetros esperados para o kit.

Os dados de sequenciamento foram analisados de maneira preliminar. Apesar das estimativas do fabricante para uma profundidade de cobertura média de 2.000x por amostra para o painel desenvolvido, sugere-se a utilização de painéis mais focados, incluindo menos alvos/regiões para genes de interesse, principalmente para sequenciadores de média capacidade como o Illumina MiSeq. Mesmo apresentando um Q30 de 93,7%, a profundidade média de sequenciamento por paciente foi de 422x (254x – 570x) e 213x (162x – 270x) para a análise sem e com o uso de MBCs, respectivamente. A profundidade obtida em ambas as análises foi abaixo da ideal para detecção de alterações em baixa frequência presentes no ctDNA. Dentre as razões para a baixa profundidade obtida está a alta taxa de *soft-clipping* após a análise, provavelmente devido ao tamanho reduzido dos fragmentos de cfDNA que acabam sendo descartados, uma vez que não são corretamente alinhados à sequência de referência.

Mesmo com a baixa profundidade, foi possível ilustrar o potencial de aplicação da técnica para um dos pacientes analisados. O paciente 2 apresentou duas mutações possivelmente provenientes do tumor, uma de significado clínico incerto p.(Ser62Trp) em *ALK* e outra provavelmente patogênica p.(Arg337Leu) em *TP53*, localizada no domínio de oligomerização da proteína p53. Ambas foram detectadas em baixa frequência antes do início do tratamento, não foram reportadas logo após o tratamento e voltaram a ser detectadas no momento da progressão da doença com frequências relativamente altas.



Mutação	Frequência alélica		
	Pré-tratamento	Pós-tratamento	Progressão
***ALK* S62T** ◆	6.79%	0%	16.60%
***TP53* R337L** ▲	2.27%	0%	23.90%

Figura 1. Ilustração da variação da frequência alélica de duas mutações de origem somática ao longo da evolução clínica do Paciente 2.

Como perspectivas para viabilizar a publicação do artigo, espera-se sequenciar o as bibliotecas já construídas em um sequenciador de maior capacidade (e.g. Illumina NextSeq 550/1000/2000), a fim de aumentar o número de fragmentos sequenciados e, consequentemente, a profundidade de leitura. Tal abordagem irá permitir diminuir o limite de detecção do teste e aumentar sua sensibilidade.

Paralelamente, as amostras também estão sendo analisadas com o painel *Oncomine Focus Assay* (© Thermo Fisher Scientific) para validação dos resultados utilizando outra metodologia. Até então, duas amostras do estudo foram processadas nesta abordagem, referentes às coletas realizadas após a progressão dos pacientes 4 e 5. Foram detectadas as mutações em *TP53* R248Q (paciente 4) e S241F (paciente 5), com frequências de 0,58% e 0,86%, respectivamente. A cobertura média nas regiões alvo deste painel foi de 46.177x para a amostra do Paciente 4 e 32.375x para a amostra

do Paciente 5. Tais alterações não foram identificadas no painel originalmente desenhado para este estudo, reforçando a necessidade de se atingir maior cobertura de sequenciamento.

Ainda para validação da técnica, amostras de tecido tumoral fixadas em parafina, até então disponíveis para quatro pacientes, serão também sequenciadas, a fim de excluir variantes provenientes da CHIP, uma vez que gDNA de origem germinativa não foi isolado para o presente estudo. Por fim, os dados clínicos serão reunidos e confrontados com os dados moleculares, a fim de avaliar os possíveis benefícios da aplicação da técnica no contexto de cada paciente.

REFERÊNCIAS BIBLIOGRÁFICAS

Ai, X., Guo, X., Wang, J., Stancu, A. L., Joslin, P. M. N., Zhang, D. and Zhu, S. (2018) 'Targeted therapies for advanced non-small cell lung cancer', Oncotarget, 9(101), pp. 37589-37607.

Arora, A. and Scholar, E. M. (2005) 'Role of tyrosine kinase inhibitors in cancer therapy', J Pharmacol Exp Ther, 315(3), pp. 971-9.

Brasil (2019) Estimativa 2020: Incidência de Câncer no Brasil. Rio de Janeiro - RJ, Brazil: Ministério da Saúde, Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA), Coordenação de Prevenção e Vigilância, p. 122.

Dagogo-Jack, I. and Shaw, A. T. (2018) 'Tumour heterogeneity and resistance to cancer therapies', Nat Rev Clin Oncol, 15(2), pp. 81-94.

Fitzmaurice, C. and Abate, D. and Abbasi, N. and Abbastabar, et al. (2019) 'Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study', JAMA Oncol, 5(12), pp. 1749-1768.

Guibert, N., Pradines, A., Favre, G. and Mazieres, J. (2020) 'Current and future applications of liquid biopsy in nonsmall cell lung cancer from early to advanced stages', Eur Respir Rev, 29(155).

Kemper, K., Krijgsman, O., Cornelissen-Steijger, P., Shahrabi, A., Weeber, F., Song, J. Y., Kuilman, T., Vis, D. J., Wessels, L. F., Voest, E. E., Schumacher, T. N., Blank, C. U., Adams, D. J., Haanen, J. B. and Peeper, D. S. (2015) 'Intra- and inter-tumor heterogeneity in a vemurafenib-resistant melanoma patient and derived xenografts', EMBO Mol Med, 7(9), pp. 1104-18.

Kumarakulasinghe, N. B., van Zanwijk, N. and Soo, R. A. (2015) 'Molecular targeted therapy in the treatment of advanced stage non-small cell lung cancer (NSCLC)', Respirology, 20(3), pp. 370-8.

Lim, Z. F. and Ma, P. C. (2019) 'Emerging insights of tumor heterogeneity and drug resistance mechanisms in lung cancer targeted therapy', J Hematol Oncol, 12(1), pp. 134.

Malta, D. C., Moura, L., Souza, M. e. F., Curado, M. P., Alencar, A. P. and Alencar, G. P. (2007) 'Lung cancer, cancer of the trachea, and bronchial cancer: mortality trends in Brazil, 1980-2003', J Bras Pneumol, 33(5), pp. 536-43.

Mok, T. S., Wu, Y. L., Ahn, M. J., Garassino, M. C., Kim, H. R., Ramalingam, S. S., Shepherd, F. A., He, Y., Akamatsu, H., Theelen, W. S., Lee, C. K., Sebastian, M., Templeton, A., Mann, H., Marotti, M., Ghiorghiu, S., Papadimitrakopoulou, V. A. and Investigators, A. (2017) 'Osimertinib or Platinum-Pemetrexed in EGFR T790M-Positive Lung Cancer', N Engl J Med, 376(7), pp. 629-640.

NCCN, N. C. C. N. 2022. Lung cancer (version 3.2022).

Nicholson, A. G., Tsao, M. S., Beasley, M. B., Borczuk, A. C., Brambilla, E., Cooper, W. A., Dacic, S., Jain, D., Kerr, K. M., Lantuejoul, S., Noguchi, M., Papotti, M., Rekhtman, N., Scagliotti, G., van Schil, P., Sholl, L., Yatabe, Y., Yoshida, A. and Travis, W. D. (2022) 'The 2021 WHO Classification of Lung Tumors: Impact of Advances Since 2015', J Thorac Oncol, 17(3), pp. 362-387.

Palacio, S., Pontes, L., Prado, E., Arshad, J., Ali, R., Piha, T., Bacchi, C. E., Mudad, R. and Lopes, G. (2019) 'Mutation Testing: Changing Patterns of Molecular Testing in Brazil', Oncologist, 24(4), pp. e137-e141.

Salgia, R., Pharaon, R., Mambetsariev, I., Nam, A. and Sattler, M. (2021) 'The improbable targeted therapy: KRAS as an emerging target in non-small cell lung cancer (NSCLC)', Cell Rep Med, 2(1), pp. 100186.

Singh, T., Fatehi Hassanabad, M. and Fatehi Hassanabad, A. (2021) 'Non-small cell lung cancer: Emerging molecular targeted and immunotherapeutic agents', Biochim Biophys Acta Rev Cancer, 1876(2), pp. 188636.

Wu, S. G. and Shih, J. Y. (2018) 'Management of acquired resistance to EGFR TKI-targeted therapy in advanced non-small cell lung cancer', Mol Cancer, 17(1), pp. 38.

CAPÍTULO VII – DISCUSSÃO

Apesar do tema já ter sido relativamente bem explorado nos Estados Unidos, na Europa e em alguns países asiáticos, pouco se conhece sobre o perfil molecular do CPNPC na América Latina e no Brasil, áreas com altas taxas de miscigenação. A exemplo das mutações em *EGFR*, cuja prevalência em pacientes diagnosticados com adenocarcinoma de pulmão varia de 7% a 76%, conforme estudos realizados em diversas partes do mundo, é de se esperar que haja diferenças regionais importantes na composição genética do CPNPC em um país de dimensões continentais e com uma população etnicamente diversa como o Brasil (Midha, Dearden and McCormack, 2015). Considerando que os poucos estudos realizados no Brasil estão concentrados no Sudeste, principalmente no estado de São Paulo, e que aproximadamente 40% das alterações moleculares no CPNPC podem indicar o uso de terapias-alvo, evoluir no maior conhecimento destes aspectos é de grande relevância, principalmente em um país onde a grande maioria da população depende do sistema público de saúde. Por exemplo, estimativas já demonstraram que fármacos de uso oral, como os inibidores de tirosina quinase (ITQs), poderiam reduzir os gastos com a doença no sistema de saúde em até 13% (Ferreira *et al.*, 2016). Portanto, conhecer o perfil molecular do CPNPC é a base para guiar políticas públicas atuais no tratamento da neoplasia, de maneira a distribuir e direcionar estrategicamente recursos, conforme as particularidades de cada região do país.

Nesse sentido, no primeiro manuscrito desta tese, demonstramos em uma série de amostras provenientes do Sul do Brasil que a frequência de pacientes com mutações no gene *EGFR*, um importante biomarcador para o CPNPC, foi a menor até então descrita no país (19%). Tal informação tem relevância terapêutica, uma vez que muitas das mutações acionáveis que conferem sensibilidade a diversos ITQs estão localizadas neste gene. Já a frequência de mutações em *KRAS* foi a maior até então descrita no país (30%), 40% delas representando a mutação G12C (Bacchi *et al.*, 2012; de Melo *et al.*, 2015; Palacio *et al.*, 2019; Leal *et al.*, 2019). Tal perfil é único e, apesar de assemelhar-se ao que foi descrito em populações europeias, há uma diferença importante na frequência de alterações em *EGFR*, a qual é de aproximadamente 10% (Gahr *et al.*, 2013; Kerr *et al.*, 2021).

Apesar de não aprovado ainda no momento da publicação do estudo, o fármaco Sotorasibe ganha especial importância na região Sul do Brasil, uma vez que quase 40% das alterações encontradas no gene correspondem a mutação G12C, sob a qual este inibidor de GTPase da família RAS atua (Blair, 2021). O presente estudo também foi o primeiro a reportar, pelo nosso conhecimento, a frequência de mutações em *BRAF* no país, estimada em 3%, corroborando com achados de outros locais (Marchetti *et al.*, 2011; Consortium, 2017; Alvarez and Otterson, 2019). A expressão de ALK, outro importante biomarcador que infere a presença de rearranjos envolvendo o gene, foi reportada em 4% dos casos, inferior ao que já havia sido reportado na região nordeste, mas semelhante ao que já foi observado no Sudeste(Alves da Silva *et al.*, 2019; Arrieta *et al.*, 2019; Gomes *et al.*, 2015; Lopes and Bacchi, 2012) . Por fim, nosso estudo também avaliou a expressão de PD-L1, um biomarcador de resposta aos inibidores de *checkpoint* imunológico. Verificamos que as amostras que expressam pontuação proporcional do tumor (PPT) ≥50% representaram 14% do total, indicando que talvez uma porção menor da população avaliada poderia se beneficiar da imunoterapia quando comparado aos 23% - 28% estimados no KEYNOTE-024, estudo que avaliou a eficácia do imunoterápico Pembrolizumabe no CPNPC (Reck *et al.*, 2016).

Outro dado revelado pelo estudo foi a presença de mutações raras e de significado clínico indeterminado em *EGFR* ou a presença de mais de uma alteração envolvendo o mesmo *driver* ou outro gene (mutações compostas). Neste trabalho, até então, 25 pacientes com este perfil foram identificados e os dados clínicos de sete deles já foram reunidos e pré-analisados. Tais estudos são de suma importância, uma vez que poucos dados estão disponíveis na literatura ou mesmo em grandes estudos clínicos. Exemplos recentes da inclusão de mutações raras em *EGFR* na lista daquelas associadas com sensibilidade aos ITQs estão a G719X, a S768I e a L861Q, presentes em menos de 5% dos casos de CPNPC e para as quais a agência reguladora norte-americana FDA (do inglês, *Food and Drug Administration*) indicou, em 2018, o uso de Afatinibe (Zhang *et al.*, 2019a).

No Brasil, é inevitável que fatores socioeconômicos também tenham um impacto importante no diagnóstico molecular do CPNPC. Apesar dos avanços tanto nas técnicas utilizadas para detecção das mutações, quanto no tratamento, o custo de ambos ainda

é elevado e, consequentemente, restringem o acesso a maioria da população. Por exemplo, o custo do NGS em países em desenvolvimento pode ser de quatro a cinco vezes maior do que em países desenvolvidos (Helmy, Awad and Mosa, 2016). Como consequência, estimativas demonstram que menos de 40% dos pacientes elegíveis para testes moleculares em CPNPC recebam o diagnóstico básico para o tratamento da neoplasia no país. A situação é ainda mais precária para pacientes atendidos pelo sistema público de saúde, onde menos de 10% dos pacientes diagnosticados com adenocarcinoma de pulmão realizam sequer o teste molecular de *EGFR* (Palacio *et al.*, 2019; Mathias *et al.*, 2020). No entanto, estas estimativas foram levantadas com dados retrospectivos de 2011 a 2016 e pode não refletir o cenário atual do país, uma vez que nos últimos anos programas da indústria farmacêutica passaram financiar diversos testes moleculares aplicados ao CPNPC. Ademais, as particularidades de cada região devem ser consideradas, reforçando a importância de estudos mais recentes e específicos para retratar o cenário atual do diagnóstico molecular no Brasil.

A ancestralidade e exposição a fatores ambientais podem levar a diferenças importantes no perfil molecular do CPNPC de pacientes com acesso ao tratamento privado e/ou a saúde suplementar, em comparação com aqueles cuja única opção é o sistema público. Nos EUA e no México, foram observadas diferenças no perfil molecular e resposta ao tratamento com drogas de alvo molecular entre pacientes de origem hispânica e não-hispânica, inclusive entre hispânicos nascidos nos EUA e imigrantes (Arrieta *et al.*, 2015; Patel *et al.*, 2013).

Outro ponto fundamental que deve ser levado em conta é a relativa demora para aprovação de novas drogas de alvo molecular no Brasil. Um exemplo foi a aprovação do Crizotinibe, droga-alvo indicada para CPNPC com fusões de *ALK* e mais recentemente exclusão do éxon 14 de *MET*. Entre o início de seu uso nos EUA em agosto de 2011 até aprovação no Brasil apenas em fevereiro de 2016, estima-se que 804 anos de vida foram perdidos devido a essa demora (Ferreira, 2019). Por outro lado, o Sotorasibe teve sua primeira aprovação em agosto de 2021 para tratamento de CPNPC localmente avançado ou metastático com a alteração *KRAS* G12C e em menos de 7 meses foi aprovado pela Agência Nacional de Vigilância Sanitária (ANVISA) para o uso no Brasil (Brasil, 2022). Outro grande agravante no cenário nacional é a acessibilidade a tais medicamentos, pois

seu custo é elevado e atualmente somente dois ITQs estão incorporados pelo SUS, o Gefitinibe e o Erlotinibe (Brasil, 2013b; Brasil, 2013a). Portanto, é fundamental o conhecimento do perfil molecular dos tumores de pulmão no Brasil e como este pode ser influenciado por diferentes cenários socioeconômicos. Dados complementares ao manuscrito I desta tese, incluindo também informações clínicas dos pacientes, poderão servir como base para guiar novos estudos que visem a gerência de políticas públicas mais atuais no tratamento do CPNPC e acelerar o acesso de tais medicamentos para a população de menor poder aquisitivo, principalmente em uma era onde ocorrem rápidos avanços em virtude da expansão do conhecimento da genômica tumoral.

Ainda no contexto de acesso ao diagnóstico molecular, já foi demonstrado que 70,6% dos casos de CPNPC do Brasil se apresentam já como doença avançada, quando intervenções cirúrgicas são pouco eficazes e quando frequentemente o tecido está pouco acessível (Ismael *et al.*, 2011). Por exemplo, em nosso estudo 16,5% (123/742) das amostras de tecido foram inadequadas para o teste ou tiveram resultados inconclusivos devido à ausência/baixa quantidade de células tumorais, ou à baixa quantidade e/ou qualidade do gDNA extraído, respectivamente, semelhante ao que já foi observado em outros trabalhos (Chouaid *et al.*, 2014; Liam, Mallawathantri and Fong, 2020)(Chouaid et al, 2014; Ilé & Hofman, 2016; Liam et al., 2020). Estudos já demonstraram que foi possível a detecção de mutações acionáveis em um terço dos casos onde o tecido não estava disponível (Thompson *et al.*, 2016; Remon *et al.*, 2020). Diante deste número expressivo de resultados inadequados/inconclusivos, este trabalho também buscou validar e implementar a técnica de biópsia líquida para avaliação do ctDNA como ferramenta diagnóstica acessória em pacientes com CPNPC estadiamento IIIC não ressecáveis atendidos em um hospital terciário no Sul do país.

A avaliação do ctDNA em CPNPC pode ir além do monitoramento de mecanismos de resistência a ITQs, auxiliando no mapeamento da heterogeneidade intratumoral e revelando possíveis novos marcadores prognósticos (Chabon *et al.*, 2016; Rohanizadegan, 2018; von Felden *et al.*, 2021). Ademais, o acompanhamento contínuo e dinâmico do paciente permite acessar em tempo real a evolução da doença, permitindo inclusive a readequação ou mesmo antecipação do tratamento conforme evolução clonal do tumor (Yanagita *et al.*, 2016; Remon *et al.*, 2017; Wan *et al.*, 2017). Os resultados

preliminares obtidos no manuscrito IV são promissores, uma vez que foi possível acompanhar em um dos pacientes as variações na frequência alélica de duas alterações de origem somática conforme evolução clínica do caso.

Algumas limitações importantes permanecem em aberto neste último estudo. Tendo em mente que apenas uma pequena fração do material genético circulante corresponde ao ctDNA (0,01% a <1%), a baixa cobertura atingida na primeira rodada de sequenciamento foi um fator limitante importante para uma melhor análise e apresentação dos resultados (Volckmar *et al.*, 2018). Por outro lado, as etapas que envolveram coleta, isolamento, amplificação e captura do cfDNA foram bem sucedidas, demonstrando que, mesmo com recursos limitados, seria viável a implementação da técnica. Portanto, para atingir uma maior cobertura de sequenciamento e um limite de detecção satisfatório, painéis amplos como este que foi utilizado no estudo talvez seriam inviáveis para a maioria das intuições, uma vez que demandariam equipamentos e insumos de alto custo. Outro limitador importante é a presença da hematopoiese clonal de potencial indeterminado (CHIP), o que pode levar a classificação equivocada de variantes que não contribuem para o processo de tumorigênese (Jaiswal *et al.*, 2014; Razavi *et al.*, 2019; Hu *et al.*, 2018; Chan *et al.*, 2020). A necessidade de inclusão de uma análise pareada, de DNA somático e de DNA de origem germinativa, agregaria mais custos ao teste, tornando a técnica ainda menos acessível. Porém, a grande maioria das alterações encontradas na CHIP não tem implicações clínicas para o manejo e tratamento de tumores de pulmão, questionando a necessidade e a real relevância de sua inclusão na prática clínica (Hu *et al.*, 2018; Liu *et al.*, 2019).

Em suma, este trabalho revelou um panorama do perfil molecular dos principais biomarcadores no CPNPC em uma série de pacientes da região Sul do Brasil, algo ainda não realizado até então. Também demonstrou a importância da medicina personalizada em um hospital terciário que atende majoritariamente uma população de baixa renda e com pouco acesso a tratamentos de alvo molecular. Não o bastante, foi proposta a implementação de uma tecnologia de alta complexidade, a biópsia líquida, para avaliação do perfil molecular de pacientes cujo tecido está pouco acessível ou é mesmo inexistente para uma avaliação molecular adequada. Apesar das dificuldades impostas pela pandemia, as quais atrasaram de maneira significativa o andamento do trabalho, os

resultados preliminares do manuscrito III demonstram que é possível a implementação da técnica, porém o custo de tal abordagem continua sendo uma barreira importante para difusão da técnica, a qual dificilmente ficará disponível no SUS em um futuro próximo sem a contrapartida do setor privado, a exemplo do que já ocorre com a avaliação molecular em tecido. Por outro laudo, técnicas que apesar de mais limitadas quanto ao número de alvos incluídos, mas que possuem custo expressivamente menor, como o PCR digital, poderiam viabilizar a incorporação do monitoramento do ctDNA, principalmente no contexto de estratificação de risco após adjuvância ou no diagnóstico de pacientes cujo biópsia tecidual esteja indisponível, com foco em *drivers* mais comumente mutados na neoplasia. Espera-se que este trabalho, o qual ainda está sendo concluído, auxilie na tomada de decisões para o direcionamento estratégico de recursos, de maneira a trazer a medicina de precisão mais próxima daqueles que mais necessitam dela.

CAPÍTULO VIII – PERSPECTIVAS

Este foi o primeiro trabalho a explorar o perfil molecular do câncer de pulmão de não pequenas células na região Sul do Brasil e também um dos poucos que explorou o tema no país. Além da conclusão das etapas que ainda estão em aberto, este trabalho cria novas perspectivas:

- Comparar o perfil molecular do CPNPC de pacientes atendidos em dois hospitais terciários no sul do Brasil, sendo um grupo atendido pelo Sistema Único de Saúde (SUS) e outro pelo sistema particular e/ou de saúde suplementar. Esta análise poderá ser desenvolvida por meio de um estudo retrospectivo;
- Avaliar a acessibilidade ao tratamento com terapias de alvo molecular em dois hospitais terciários no sul do Brasil, sendo um grupo atendido pelo SUS e outro pelo sistema particular e/ou de saúde suplementar. Esta análise poderá ser desenvolvida por meio de um estudo retrospectivo;
- Estimular a criação de uma rede nacional de diagnóstico molecular para aumentar o acesso do diagnóstico de precisão aos pacientes com condições financeira limitadas;
- Estimular a incorporação de um número maior de drogas de alvo molecular no SUS;
- Motivar a incorporação da biópsia líquida para avaliação do ctDNA como ferramenta diagnóstico complementar em hospitais públicos, principalmente para pacientes cujo tecido esteja inadequado ou inacessível para o diagnóstico molecular. Técnicas mais custo-efetivas, como o PCR digital, seriam uma alterativa mais viável para o SUS, apesar de suas limitações quanto ao número de alvos incluídos por ensaio.

PRODUÇÃO CIENTÍFICA ADICIONAL DURANTE O PERÍODO

"*Prevalence of the Brazilian TP53 Founder c.1010G>A (p.Arg337His) in Lung Adenocarcinoma: Is Genotyping Warranted in All Brazilian Patients?"*

Autores: Igor Araujo Vieira, Tiago Finger Andreis, Bruna Vieira Fernandes, Maria Isabel Achatz, Gabriel de Souza Macedo, Daniel Schramek, Patricia Ashton-Prolla

Artigo publicado no periódico *Frontiers in Genetics* (ISSN 1664-8021) - 2021 Feb 2;12

doi: 10.3389/fgene.2021.606537.

"Molecular Profiling as Predictor of Outcomes in A Brazilian Cohort Of Stage IV Lung Cancer."

Autores: Nicolas Peruzzo, Juliano Ce Coelho, Gabriel de Souza Macedo, Tiago Finger Andreis, Gustavo Gössling, Pedro Grachinski Buiar, Débora Leite Rocha, Patricia Ashton-Prolla

Resumo com dados do Manuscrito II, apresentado na *American Society of Clinical Oncology* (ASCO) disponível no periódico *Journal of Clinical Oncology* (ISSN 1527-7755)

2019 May 26; 37:15_suppl

doi: 10.3389/fgene.2021.606537

"Clinical and molecular characterization of patients fulfilling Chompret criteria for Li-Fraumeni syndrome in Southern Brazil."

Autores: Camila Matzenbacher Bittar, Yasminne Marinho de Araújo Rocha, Igor Araujo Vieira, Clévia Rosset, Tiago Finger Andreis, Ivaine Tais Sauthier Sartor, Osvaldo Artigalás, Cristina B O Netto, Barbara Alemar, Gabriel de Souza Macedo, Patricia Ashton-Prolla

Artigo publicado no periódico *PLOS ONE* (ISSN 1932-6203) - 2021 Sep 16;16(9)

doi: 10.1371/journal.pone.0251639.

"TP53 variants of uncertain significance: increasing challenges in variant interpretation and genetic counseling"

Autores: Camila Matzenbacher Bittar, Igor Araujo Vieira, Cristina Silva Sabato, Tiago Finger Andreis, Bárbara Alemar, Osvaldo Artigalás, Henrique de Campos Reis Galvão, Gabriel de Souza Macedo , Edenir Inez Palmero, Patricia Ashton-Prolla

Artigo publicado no periódico *Familial Cancer* (ISSN 1573-7292) - Oct;18(4)

doi: 10.1007/s10689-019-00140-w

"Systems Biology Approaches Reveal Potential Phenotype-Modifier Genes in Neurofibromatosis Type 1"

Autores: Thayne Woycinck Kowalski, Larissa Brussa Reis, Tiago Finger Andreis, Patricia Ashton-Prolla, Clévia Rosset

Artigo publicado no periódico *Cancers* (ISSN 2072-6694) - 2020 Aug 26;12(9)
doi: 10.3390/cancers12092416.

"*MIR605 rs2043556 is associated with the occurrence of multiple primary tumors in TP53 p.(Arg337His) mutation carriers"*

Autores: Isabel Cristina Bandeira, Igor Araujo Vieira, Tiago Finger Andreis, Larissa Brussa Reis, Gabriel de Souza Macedo, Fernanda Sales Luiz Vianna, Patricia Santos-Silva, Edenir Inez Palmero, Henrique de Campos Reis Galvão, Cintia Regina Niederauer Ramos, Karina Miranda Santiago, Maria Isabel Achatz, Alexandre André Balieiro Anastácio da Costa, Patricia Ashton-Prolla

Artigo publicado no periódico *Cancer Genetics* (ISSN 2210-7762) - *2020 Jan;240*
doi: 10.1016/j.cancergen.2019.11.005.

frontiers in Genetics

BRIEF RESEARCH REPORT
published: 02 February 2021
doi: 10.3389/fgene.2021.606537

Prevalence of the Brazilian *TP53* Founder c.1010G>A (p.Arg337His) in Lung Adenocarcinoma: Is Genotyping Warranted in All Brazilian Patients?

Igor Araujo Vieira[1,2†], *Tiago Finger Andreis*[1,2†], *Bruna Vieira Fernandes*[2,3], *Maria Isabel Achatz*[4], *Gabriel S. Macedo*[2,5], *Daniel Schramek*[6,7] and *Patricia Ashton-Prolla*[1,2,5*]

[1] Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, [2] Laboratório de Medicina Genômica, Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil, [3] Curso de Graduação em Biomedicina, Universidade Federal de Ciências da Saúde de Porto Alegre, Porto Alegre, Brazil, [4] Departamento de Oncogenética, Hospital Sírio Libanês, São Paulo, Brazil, [5] Programa de Medicina Personalizada, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil, [6] Centre for Molecular and Systems Biology, Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON, Canada, [7] Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada

In Southern and Southeastern Brazil, there is a germline pathogenic variant with incomplete penetrance located in the oligomerization domain of *TP53*, c.1010G>A (p.Arg337His). Due to a founder effect, the variant is present in 0.3% of the general population of the region. Recently, this variant was identified in 4.4 and 8.9% of two apparently unselected, single center case series of Brazilian lung adenocarcinoma (LUAD) patients from the Southeastern and Central regions of the country, respectively. In the present study, our aim was to examine *TP53* c.1010G>A allele and genotype frequencies in LUAD samples obtained from patients diagnosed in Southern Brazil. A total of 586 LUAD samples (tumor DNA) recruited from multiple centers in the region were tested, and the mutant allele was identified using TaqMan® assays in seven cases (7/586, 1.2%) which were submitted to next generation sequencing analyses for confirmation. Somatic *EGFR* mutations were more frequent in *TP53* c.1010G>A carriers than in non-carriers (57.1 vs. 17.6%, respectively). Further studies are needed to confirm if *TP53* c.1010G>A is a driver in LUAD carcinogenesis and to verify if there is a combined effect of *EGFR* and germline *TP53* c.1010G>A. Although variant frequency was higher than observed in the general population, it is less than previously reported in LUAD patients from other Brazilian regions. Additional data, producing regional allele frequency information in larger series of patients and including cost-effectiveness analyses, are necessary to determine if *TP53* c.1010G>A screening in all Brazilian LUAD patients is justified.

Keywords: *TP53* gene, p53 protein, lung adenocarcinoma, founder variant, *TP53* (p.Arg337His), R337H, non-small cell lung cancer, Li-Fraumeni syndrome

OPEN ACCESS

Edited by:
Joan Brunet,
Catalan Institute of Oncology, Spain

Reviewed by:
Ernest Nadal,
Catalan Institute of Oncology, Spain
Judith Balmaña,
Vall d'Hebron University Hospital, Spain

***Correspondence:**
Patricia Ashton-Prolla
pprolla@hcpa.edu.br;
pprolla@gmail.com

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Genetics of Common and Rare Diseases, a section of the journal Frontiers in Genetics

Received: *15 September 2020*
Accepted: *11 January 2021*
Published: *02 February 2021*

Citation:
Vieira IA, Andreis TF, Fernandes BV, Achatz MI, Macedo GS, Schramek D and Ashton-Prolla P (2021) Prevalence of the Brazilian TP53 Founder c.1010G>A (p.Arg337His) in Lung Adenocarcinoma: Is Genotyping Warranted in All Brazilian Patients? Front. Genet. 12:606537. doi: 10.3389/fgene.2021.606537

INTRODUCTION

In Southern and Southeastern Brazil, a germline founder pathogenic variant with incomplete penetrance, c.1010G>A (rs121912664), also known as R337H or p.(Arg337His) has been detected in 0.3% of the general population (Achatz et al., 2009; Custódio et al., 2013). It is located in the oligomerization domain (exon 10) of *TP53*, and it is associated with Li-Fraumeni syndrome (LFS). Carriers are at a high risk for developing a wide spectrum of tumors and are prone to develop multiple primary cancers at different ages. The core tumors in LFS patients include early onset breast cancer, soft-tissue sarcomas, brain tumors and adrenocortical carcinomas (Achatz et al., 2007; Giacomazzi et al., 2013; Achatz and Zambetti, 2016). Lung cancer (LC), especially lung adenocarcinoma (LUAD) of the lepidic subtype, has been reported in LFS families and is included in the Chompret criteria for genetic testing of suspected patients, but there is limited evidence for a strong association of its occurrence with germline *TP53* variants (Tinat et al., 2009; Ricordel et al., 2015; Caron et al., 2017). Although *TP53* is considered one of the most commonly mutated genes in solid tumors, somatic occurrence of *TP53* c.1010G>A is extremely rare. In the IARC *TP53* database [International Agency for Research on Cancer TP53 database (Iarc), 2020], c.1010G>A has been described in only 4 of 28,869 solid tumors. In other public databases, its frequency in solid tumors is also very low. In COSMIC [Catalogue Of Somatic Mutations In Cancer (Cosmic), 2020], for instance, among over 20.000 samples, *TP53* c.1010G>A is not present, although G > T and G > C alterations in codon 1,010 do occur. In contrast, two lung tumors harboring germline c.1010G>A have been described in the IARC *TP53* database. Finally, Nogueira et al. (2015) reported a mixed acinar/bronchiolo-alveolar carcinoma in a known germline carrier of the *TP53* c.1010G>A variant (Nogueira et al., 2015).

LC is the leading cause of cancer related deaths worldwide, responsible for 1.7 million deaths every year. In Brazil, the National Cancer Institute [Instituto Nacional de Câncer José Alencar Gomes da Silva (Inca), 2020] estimated 30,200 new LC cases in 2020, rendering it the second most common solid tumor in the country. In Southern Brazil, LC is the third most common cancer diagnosed in adults, with non-small cell lung cancer (NSCLC) accounting for approximately 85% of all LCs cases. Recently, Couto et al. (2017) genotyped *TP53* c.1010G>A in 45 NSCLC patients from a single center in the central region of Brazil, where this variant has not been explored and its population prevalence has not been determined. They identified 4 (8.9%) heterozygotes, a surprisingly high variant frequency for a small, apparently unselected cohort (Couto et al., 2017). Importantly, a more recent single center study in Southeastern Brazil also observed a high prevalence of *TP53* c.1010G>A in an unselected series of 114 *EGFR*-positive LUAD patients: the variant was present in 4.4 and 12.5% of samples when considering diagnosis at any age or before the age of 50 years, respectively (Barbosa et al., 2020). Furthermore, the authors assessed LUAD tumors diagnosed in known *TP53* c.1010G>A carriers and observed that LUAD tumors from 8/9 (89%) *TP53* c.1010G>A carriers harbored an activating *EGFR* variant. To our knowledge, the *TP53* c.1010G>A variant has not been described in other studies assessing somatic *TP53* variants in sporadic lung adenocarcinoma (LUAD) patients (Greenman et al., 2007; Ding et al., 2008; Hammerman et al., 2012; Imielinski et al., 2012; Peifer et al., 2012; Rizvi et al., 2015; Jordan et al., 2017). Thus, in the present study, our goal was to examine the *TP53* c.1010G>A allele and genotype frequencies in a series of 586 LUAD samples obtained from patients diagnosed in multiple centers of the three states of Southern Brazil, a region with the highest population frequency of this particular variant observed to date, regardless of their clinical features.

METHODS

Study Subjects and Ethical Aspects

A total of 586 LC samples derived from a cohort described in a previous study from our group (Andreis et al., 2019) were analyzed. Patients were originally recruited for somatic mutation testing in *EGFR* (exons 18–21), *KRAS* (exons 2 and 3), *BRAF* (exons 11 and 15), and *NRAS* (exons 2 and 3) genes from different hospitals and clinics distributed in 22 healthcare centers located in the three states of the southern region of Brazil: Rio Grande do Sul (N = 496), Santa Catarina (N = 20), and Paraná (N = 70). Patients were not selected on previous cancer history or family history of cancer. Pathologic analyses confirmed typical adenocarcinoma histology in all cases. However, histological subtype data were available for only a small proportion (72/586) of patients. The mutation status of *EGFR*, *BRAF*, *KRAS*, and *NRAS* genes was evaluated using technical procedures as previously reported (Andreis et al., 2019). Genotyping was performed in a central laboratory, a diagnostic precision medicine program established in a tertiary care hospital in southern Brazil. Before initiation of this study, age at LUAD diagnosis, *EGFR*/*KRAS*/*BRAF* status, and histological subtype (when available) were annotated and samples were posteriorly de-identified. A consent waiver was approved by the Institutional Review Board specifically for the *TP53* c.1010G>A analysis, given that patient identification using only age at diagnosis, *EGFR*/*KRAS*/*BRAF*/*TP53* status and LUAD histological subtype would not allow patient re-identification.

DNA Extraction and *TP53* c.1010G>A Genotyping

Tumor DNA was extracted using the ReliaPrep™ FFPE gDNA Miniprep System (Promega), according to the manufacturer's recommendations. Next, TaqMan® allelic discrimination analyses of the pathogenic variant *TP53* c.1010G>A (rs121912664) were performed according to Applied Biosystems® standard protocols (Applied Biosystems, Carlsbad, CA United States), using fluorescent allele-specific probes as previously published (Fitarelli-Kiehl et al., 2016). An attempt to confirm rs121912664-positive results identified by TaqMan® by next generation sequencing (NGS) for determination of allele frequencies was made. Briefly, NGS of the *TP53* entire coding region (exons 2–11) and 70 pb exon-intron boundaries was done using a custom panel (Thermo Fisher Scientific, CA,

United States, reference number TP53.20140108.designed) on the Ion Torrent-Personal Genome Machine (PGM) platform. Amplicon library was prepared using the Ion AmpliSeq™ Library Kit 2.0 (Thermo Fisher Scientific, CA, United States). PCR products were then sequenced on the Ion GeneStudio S5 system (Ion Torrent Systems Inc., Gilford, NH, United States). NGS results were interpreted using the Ion Reporter software considering a minimum coverage of 100X by amplicon. Integrative Genomics Viewer (IGV) was used for visualization of the mapped reads. Human *TP53* cDNA sequence corresponding to the NM_000546.5 was used as a wild-type (WT) reference.

Statistical Analyses

Genotype and allele frequencies were estimated by simple counting. Clinical and molecular features of LUAD patients were assessed using descriptive statistics. Considering the low number of mutant alleles found in our study and limitations in clinical data availability, it was not possible to perform any meaningful statistical test in our comparisons between groups of carriers and non-carriers (see more in Results section). SPSS® version 18 (SPSS® Inc., Chicago, IL, United States) was used for data handling and for all descriptive analyses.

RESULTS

A total of 586 LUAD samples were included in this study. Clinical and molecular data are summarized in **Table 1**. Histopathological subclassifications were available only for 72 cases (12.3%). The most common subtypes were acinar (65.3%) and lepidic (34.7%). Moreover, the majority of LUAD samples (502/586, 85.7%) were tested for somatic mutations in *EGFR, KRAS, BRAF* and *NRAS* driver genes. Somatic mutations were mostly identified in *KRAS* (29.5%) and *EGFR* (21.1%).

TABLE 1 | Clinical and molecular features of 586 patients with lung adenocarcinoma (LUAD) included in this study.

Features	Median (IR)*	*N* (%)†
Age at tumor diagnosis, years	67 (16)	–
Gender		
Male		271 (46.2)
Female		315 (53.8)
Histological subtype		72 (12.3)
Acinar		47 (65.3)
Lepidic		25 (34.7)
Patients tested for somatic alterations‡		502 (85.7)‡
EGFR mutation		106 (21.1)
KRAS mutation		148 (29.5)
BRAF mutation		
None		233 (46.4)
Carriers of *TP53* c.1010G>A (p.Arg337His)§		7 (1.2)§

*IR, interquartile range.
† The percentage was calculated over the total number of genotyped samples (586) and over the number of cases for which the specified clinical data was available.
‡ Tested for somatic EGFR, KRAS, BRAF, and NRAS mutations. Specific regions evaluated in each gene are detailed in the "Materials and Methods."
§ Seven carriers were identified using TaqMan® assays, but heterozygous genotype was confirmed by NGS in only five positive cases due to limitations in sample availability for the remaining two patients.

Specific *TP53* c.1010G>A (p.Arg337His) genotyping by TaqMan® resulted in the identification of seven heterozygotes (GA genotype). Heterozygous genotype and mutant allele (A) frequencies were thus defined at 7/586 (1.2%) and 7/1,172 (0.6%), respectively. All patients in this subgroup were diagnosed with LUAD after age 50 years, and the median age at tumor onset was 60 years. In addition, most c.1010G>A tumors also had somatic *EGFR* (4/7, 57.1%) variants and none of them had a *KRAS* variant. Importantly, heterozygous genotype was confirmed by NGS in five (p.Arg337His)-positive cases (5/586, 0.85%). In the other two positive samples, further analyses were not possible due limitations in sample availability. Mutant allele frequencies determined by NGS were close to 50% in 3/5 cases, suggesting that at least in these cases, the variant may also be present in the germline. Details on each of the heterozygous samples are summarized in **Table 2**. Additional *TP53* alterations detected in the tumors by NGS are depicted in **Supplementary Table 1**.

Lastly, a comparison of the available clinical and molecular features between *TP53* c.1010G>A carriers identified by TaqMan® and non-carriers is presented in **Table 3** and **Supplementary Figure 1**. A difference between groups was observed for median age at cancer diagnosis and histological subtype but the number of mutant allele carriers (7/586) was too small to make meaningful comparisons. LUAD from three carriers were of the lepidic subtype (3/7, 42.8%), however, it is important to emphasize that three of the seven heterozygotes had no complete histological data, which hindered our ability to analyze a possible association between the c.1010G>A variant and occurrence of the lepidic subtype. Interestingly, the presence of somatic *EGFR* mutations was found in a much higher frequency in *TP53* c.1010G>A carriers than in non-carriers (57.1 vs. 17.6%, respectively).

DISCUSSION

The *TP53* founder variant c.1010G>A, widely referred as R337H, is likely responsible for a significant proportion of the cancer burden in the Southern and Southeastern regions of Brazil due to its high frequency in the population (Achatz et al., 2009; Garritano et al., 2010). In a landmark study conducted in 171,000 newborns from the Southern Brazilian State of Paraná the variant was identified in 461 individuals (~0.3%) (Custódio et al., 2013). Previous analyses performed by our research group (Palmero et al., 2008) identified a similar prevalence (2/750, ~0,3%) in a cohort of healthy, asymptomatic women participating in a community-based breast cancer screening program in the state of Rio Grande do Sul (Southern Brazil). Furthermore, it has been well-documented that the variant exhibits incomplete, variable penetrance resulting in significant inter- and intra-familiar heterogeneity in phenotypic presentation with some carriers surviving without any cancer diagnosis to older ages and others having the typical LFS phenotype, with one or more cancers diagnosed in childhood and/or adult life (reviewed in Achatz and Zambetti, 2016). Unlike most *TP53* somatic and/or germline

TABLE 2 | Characterization of LUAD tumors with an identifiable *TP53* c.1010G>A pathogenic variant.

Identifier	Gender	Age at LUAD diagnosis (years)	Histological subtype	Somatic *EGFR* mutation	*TP53* c.1010G>A zygosity	WT*/*TP53* c.1010G>A allele frequency (coverage)	Percentage of tumor cells in the sample
1	Male	57	NA†	p.(Leu858Arg)	Heterozygous	0.52/0.48 (2,400x)	40%
2	Male	65	Lepidic	p.(Ser768_Asp770dup)	Heterozygous	0.37/0.63 (2,259x)	30%
3	Female	55	NA	p.(Leu858Arg)	Heterozygous	0.24/0.76 (4,000x)	60%
4	Female	60	Lepidic	None	Heterozygous	0.42/0.58 (1,802x)	40%
5	Female	74	NA	p.(Leu858Arg)	Heterozygous	0.42/0.58 (1,835x)	70%
6	Female	54	Lepidic	Inconclusive§	Heterozygous	NP‡	5%
7	Female	62	Acinar	Inconclusive	Heterozygous	NP	40%

** WT, wild-type allele; † NA, not available; ‡ NP, not performed; § Inconclusive status due to technical limitations, such as availability of a low concentration of DNA extracted from the tumor tissue, and/or poor quality/purity of the tumor DNA, leading to inadequate results (low coverage) in the NGS analysis.*

deleterious variants involved in tumorigenesis (i.e., those located in the DNA-binding domain of p53 protein), c.1010G>A is located in the p53 oligomerization domain and has been associated with a unique intracellular pH-dependent effect on protein stability, through which the mutant protein retains some partial tumor suppressor activity (DiGiammarino et al., 2002; Zerdoumi et al., 2017). More recently, a knock-in animal model containing the homologous *TP53* R337H variant (mouse R334H) demonstrated that this alteration triggers reduced formation of p53 dimers and tetramers (deficient oligomerization capacity) compared to WT *p53* in mouse liver tissues after exposure to a specific carcinogen (Park et al., 2018). Interestingly, the homozygous *p53* R334H mutant mice exposed to this carcinogen showed an increased liver tumor development, while mice with the same genotype that were not under exposure to the liver carcinogen developed normally without any significant difference in terms of either cancer incidence or life span compared with WT mice (Park et al., 2018, 2019).

Given its relevance, prevalence of the mutant allele (germline and/or somatic DNA analysis) has been extensively investigated in different Brazilian cohorts of patients with several tumor types, regardless of cancer family history (FH). In three studies evaluating patients with adrenocortical carcinoma (ACC), c.1010G>A was present in 90–97%, independent of cancer FH (Ribeiro et al., 2001; Seidinger et al., 2011; Mastellaro et al., 2018). Moreover, the variant was found in 69, 8.4, and 7.3% of unselected subjects diagnosed with choroid plexus carcinoma, neuroblastoma and osteosarcoma, respectively (Seidinger et al., 2011, 2015). Prevalence of the variant in breast cancer (BC) patients from southern and southeastern Brazil was also studied. In three independent case series, prevalences of 3.4% (familial BC but without a clear LFS phenotype) and 2.5–8.6% (BC patients unselected for cancer FH diagnosed in different age groups) were observed (Giacomazzi et al., 2014; Hahn et al., 2018).

In the present study, we identified seven LUAD tumors harboring the founder *TP53* variant among 586 samples recruited from patients in southern Brazil, a cohort described in detail elsewhere (Andreis et al., 2019). Remarkably, all variant-positive tumors were diagnosed over age 50 years. The late onset of LC in *TP53* c.1010G>A carriers is in agreement with a recent finding showing a lifelong cancer risk pattern characterized by a bimodal age distribution: one peak in the first 10 years of life associated mainly with ACC and CNS tumors, and a second peak in the fifth decade in which different tumor types occur, including LC (Mastellaro et al., 2017). Although LC (especially LUAD) is not among the core, most frequent tumors originally described in LFS, it has been described in Brazilian LFS families with the *TP53* c.1010G>A variant (Barbosa et al., 2020), together with other atypical tumors such as papillary thyroid and renal cancers (Achatz and Zambetti, 2016).

Although *TP53* c.1010G>A prevalence in LUAD reported in the present study (1.2%) is substantially higher than the

TABLE 3 | Comparison of clinical and molecular features between LUAD samples of *TP53* c.1010G>A carriers and non-carriers.

Features	Carriers, *N* = 7	Non-carriers, *N* = 579
Median age at diagnosis, years (IR)*	60 (10)	67 (16)
Histological subtype, *N* (%)	4 (57.1)	68 (11.7)
Acinar	1 (14.3)	46 (7.9)
Lepidic	3 (42.8)	22 (3.8)
Patients with other somatic alterations, *N* (%)†	5 (71.4)	497 (85.8)
EGFR mutation	4 (57.1)	102 (17.6)
KRAS mutation	0	148 (25.6)
BRAF mutation	0	15 (2.6)
Median age at diagnosis restricted to *EGFR*-mutant subtype, years (IR)	61 (16)	67.5 (20)

**IR, interquartile range.*
†A total of 502/586 LUAD cases were tested for somatic EGFR, KRAS, BRAF, and NRAS mutations.

population frequency observed in large cohorts of healthy individuals from the same Brazilian region (0.3%) (Palmero et al., 2008; Custódio et al., 2013), it is much lower than the prevalences previously described by Couto et al. (2017), and Barbosa et al. (2020). The conflicting results between our study and the study performed by Couto et al. (2017) may be explained by two main reasons: (a) small sample size along with a putative selection bias in patient recruitment in the study conducted by Couto et al. indirectly evidenced by high proportion of carriers identified with a cancer FH; and (b) employment of different variant screening approaches (PCR-RFLP vs. TaqMan®), since use of restriction endonucleases may be associated with increased false-positive rates (Uemura et al., 2004; Kang et al., 2009). The study by Barbosa et al. (2020), although performed on a larger sample size, is also a single center study developed in an institution that has been a reference center for the diagnosis and follow-up of LFS families. In addition, one cannot exclude presence of specific, regional environmental factors (carcinogenic exposure, similar to what has been observed in mouse models) which might act triggering LUAD carcinogenesis in c.1010G>A carriers and result in a variable susceptibility to LUAD in carriers from different geographic regions.

Another interesting finding of our study was the co-occurrence of *TP53* c.1010G>A and activating *EGFR* mutations, which is in agreement with two previous reports (Barbosa et al., 2020; Mezquita et al., 2020). Studying LFS patients with a germline *TP53* pathogenic variant (either DNA-binding variants or the founder c.1010G>A) and LUAD, Barbosa et al. (2020) and Mezquita et al. (2020) identified somatic *EGFR* alterations in 89 and 85% of the tumors, respectively. In our study, *EGFR* variants occurred in 57.1% of the *TP53* c.1010G>A-positive tumors, a lower proportion than reported before, but our cohort was not of LFS patients. In addition, when looking at 114 LUAD patients with *EGFR* mutations from a non-LFS cohort,

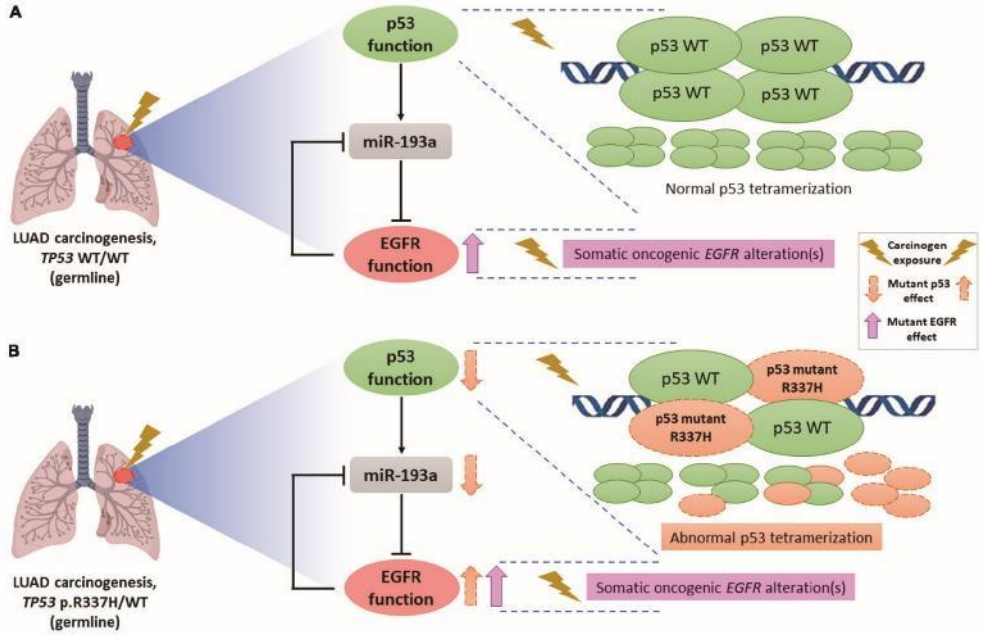


FIGURE 1 | Schematic representation of the hypothetical interaction between mutant p53 (due to hypomorphic *TP53* germline variants such as c.1010G>A, p.R337H) and mutant EGFR (due to somatic activating *EGFR* alterations) in NSCLC tumors. **(A)** Lung carcinogenesis in an individual without germline *TP53* pathogenic variants, i.e., a homozygous context of wild-type (WT) p53 (*TP53* WT/WT genotype) characterized by constitutive formation of normal p53 tetramers in response to genotoxic stress and/or carcinogenic exposure. Occurrence of an activating *EGFR* somatic event leads to an EGFR upregulation. Moreover, according to previous functional studies, there is a link between p53 and EGFR through miR-193a, forming a double-negative feedback loop that contributes to NSCLC tumorigenesis. **(B)** LUAD formation process in a carrier of the germline *TP53* p.R337H (*TP53* p.R337H/WT genotype). To explain the frequent co-occurrence of germline *TP53* and somatic *EGFR* variants recently described in different case series, we postulate that the deficient p53 oligomerization capacity conferred by the *TP53*-p.R337H founder allele, specifically in the context of exposure to carcinogenic environmental factors triggers reduced tumor-suppressive p53 functions in lung tissue, representing an initial factor (first hit) in NSCLC/LUAD carcinogenesis. Subsequent mutational events at the somatic level (possibly resulting also from environmental exposure factors), including oncogenic *EGFR* variants (second hit) could dramatically impair the p53/miR-193a/EGFR feedback loop regulation, acting as a mechanism driving the NSCLC/LUAD development.

Barbosa et al. (2020) observed a *TP53* codon 337 variant in 5.3%, which was further increased to 12.5% (4/32) when considering only patients diagnosed before age 50 years. In our study, *TP53* c.1010G>A was present in 4/106 (3.8%) LUAD samples with activating *EGFR* mutations, and all patients with both alterations were diagnosed after the age of 50 years. Importantly, ethnic ancestry has major impact on the incidence of *EGFR* mutation status in NSCLC patients, being lower (around 10%) in regions with higher European ancestry representation (Gahr et al., 2013). The prevalence of European ancestry in the population studied here, i.e., from Southern Brazil, has been estimated at 80–90%, the highest among all Brazilian regions (Ruiz-Linares et al., 2014). In agreement with this observation, several studies evaluating the human leukocyte antigen genotypic diversity in this region also confirmed a high European ancestry and a meaningful difference from Asians or even Brazilian indigenous populations (Castro et al., 2019).

Indeed, the frequent co-occurrence of *TP53* and *EGFR* sequence variants is striking and it might be explained by a p53/miR-193a/EGFR feedback loop mechanism previously reported as a driving force for NSCLC tumorigenesis (Wang et al., 2019). *In vitro* and *in vivo* studies demonstrated that WT p53 directly activates miR-193a transcription and, in turn, miR-193a directly targets *EGFR*, whereas *EGFR* functions as a transcriptional repressor to negatively control miR-193a expression, forming a feedback loop (**Figure 1A**). Considering the repeated reports of co-occurrence of *TP53* and *EGFR* variants in LUAD, including the present study, we hypothesize that occurrence of an activating *EGFR* somatic event on a background of a tissue already harboring a mutant (germline) *TP53* allele (such as c.1010G>A) may result in an impairment of this loop function, promoting NSCLC formation and tumor growth (**Figure 1B**). In parallel, it has been widely described that, upon conditions of genotoxic stress in the cellular context, presence of functional *TP53* germline variants, especially hypomorphic variants (c.1010G>A), leads to a loss of p53 tumor-suppressive functions (such as DNA damage repair and cell-cycle arrest), triggering a genomic instability that, in turn, may promote the accumulation of somatic mutations at different genes (Zerdoumi et al., 2017; Park et al., 2018), which could include the hotspot regions of somatic *EGFR* mutations. Additionally, a previous study suggested that radiation-induced cancers are more common among LFS patients (harboring germline *TP53* pathogenic variants) (Heymann et al., 2010). Although the underlying mechanism is not known, in recent reports of *EGFR*-mutated lung tumors in LFS patients, researchers hypothesized that the first genetic hit was a germline *TP53* mutation and suggested that chemotherapy or radiotherapy (i.e., genotoxic stress-inducing conditions) promoted a second hit, which might have consisted specifically of somatic *EGFR* mutations (Michalarea et al., 2014; Ricordel et al., 2015). Further functional studies are required to confirm this latter hypothesis, as well as to elucidate, mechanistically, the reason why patients with *TP53* germline mutations seem to harbor, more frequently, activating *EGFR* mutations in lung tissue.

Finally, the results of the current study must be interpreted with caution, considering its limitations. First, LUAD specimens were obtained retrospectively from a diagnostic cohort de-identified for this study, hindering the inquiry about clinical data, such as histological subtype, smoking habit, cancer FH, and ethnic ancestry. Second, for the same reason, germline screening for the variant was not done and, thus, the study design employed here did not allow contact with R337H-positive patients or their relatives. Lastly, the limited availability of samples did not allow additional analyses of the tumor samples such as LOH testing.

In conclusion, when compared to previous studies in Brazilian patients with LUAD, the prevalence of *TP53* c.1010G>A, although higher than expected for the general population, was much lower in our series from Southern Brazil, suggesting that there may be regional variations. The variability observed so far, in the absence of large prevalence studies in different regions of the country, and also without a more detailed cost-effectiveness analysis, do not allow, in our view, proposition of a general recommendation of testing all Brazilian LUAD patients for *TP53* c.1010G>A. Further studies assessing presence of *TP53* germline variants, or at least the founder c.1010G>A variant, in Brazilian LUAD patients, regardless of the age at tumor diagnosis and especially if they harbor activating *EGFR* mutations, should be undertaken in order to determine if universal screening for *TP53* c.1010G>A is justified. Ultimately, predictive testing in healthy family members of variant-positive LUAD probands might be useful to assess the clinical actionability toward the occurrence of other tumor types, and future cost-effectiveness analysis should include this clinical actionability after predictive testing.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board from Hospital de Clinicas de Porto Alegre (IRB-HCPA). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

IV, TA, GM, and PA-P conceived the work and conception design of the brief research report. IV, TA, and PA-P designed the draft of the manuscript and carried out the statistical analyses. TA, GM, and PA-P were involved in recruitment of tumor DNA samples and retrospective search of clinical data. TA, BF, IV, GM, MA, and DS performed the genotyping and NGS analyses, as well as the interpretation of results from these experiments. IV, TA, MA, GM, DS, and PA-P helped to draft the manuscript. PA-P supervised the work. All authors revised the manuscript critically,

contributed with interpretation of the findings and gave final approval of the version to be published.

FUNDING

This study was supported in part by funds from Fundo de Incentivo à Pesquisa do Hospital de Clínicas de Porto Alegre (FIPE-HCPA, grant number #18-0099 to GM) and from the Canada-Israel Health Research Initiative, jointly funded by the Canadian Institutes of Health Research, the Israel Science Foundation, the International Development Research Centre, Canada and the Azrieli Foundation (application #384428 2017-02-14/IDRC File #108587-001 to DS). PA-P. receives a researcher grant from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

ACKNOWLEDGMENTS

We would like to thank Bruno da Silveira Corrêa, Fernanda de Paris, Bárbara Alemar, and Kendi Nishino Miyamoto for their valuable contributions and technical support.

REFERENCES

Achatz, M. I., Hainaut, P., and Ashton-Prolla, P. (2009). Highly prevalent TP53 mutation predisposing to many cancers in the Brazilian population: a case for newborn screening? *Lancet Oncol.* 10, 920–925. doi: 10.1016/S1470-2045(09)70089-0

Achatz, M. I., Olivier, M., Le Calvez, F., Martel-Planche, G., Lopes, A., Rossi, B. M., et al. (2007). The TP53 mutation, R337H, is associated with Li-Fraumeni and Li-Fraumeni-like syndromes in Brazilian families. *Cancer Lett.* 245, 96–102. doi: 10.1016/j.canlet.2005.12.039

Achatz, M. I., and Zambetti, G. P. (2016). The Inherited p53 mutation in the Brazilian population. *Cold Spring Harb. Perspect. Med.* 6:a026195. doi: 10.1101/cshperspect.a026195

Andreis, T. F., Correa, B. S., Vianna, F. S., De-Paris, F., Siebert, M., Leistner-Segal, S., et al. (2019). Analysis of predictive biomarkers in patients with lung adenocarcinoma from southern brazil reveals a distinct profile from other regions of the country. *J. Glob. Oncol.* 5, 1–9. doi: 10.1200/JGO.19.00174

Barbosa, M. V. R., de Lima, V. C., Formiga, M. N., de Paula, C. A., Torrezan, G. T., and Carraro, D. M. (2020). High prevalence of egfr mutations in lung adenocarcinomas from Brazilian patients harboring the TP53 p.R337H variant. *Clin. Lung Cancer* 21, e37–e44. doi: 10.1016/j.cllc.2019.11.012

Caron, O., Frebourg, T., Benusiglio, P. R., Foulon, S., and Brugières, L. (2017). Lung adenocarcinoma as part of the Li-Fraumeni syndrome spectrum. *JAMA Oncol.* 3, 1736–1737. doi: 10.1001/jamaoncol.2017.1358

Castro, M. S., Issler, H. C., Gelmini, G. F., de Miranda, B. L. M., Calonga-Solís, V., Schmidt, A. H., et al. (2019). High-resolution characterization of 12 classical and non-classical HLA loci in Southern Brazilians. *HLA* 93, 80–88. doi: 10.1111/tan.13488

Catalogue Of Somatic Mutations In Cancer (Cosmic) (2020). *Graphical View of Mutations Across TP53 Gene.* https://cancer.sanger.ac.uk/cosmic/gene/analysis?ln=TP53 (Accessed September 15, 2020).

Couto, P. P., Bastos-Rodrigues, L., Schayek, H., Melo, F. M., Lisboa, R. G. C., Miranda, D. M., et al. (2017). Spectrum of germline mutations in smokers and non-smokers in Brazilian non-small-cell lung cancer (NSCLC) patients. *Carcinogenesis* 38, 1112–1118. doi: 10.1093/carcin/bgx089

Custódio, G., Parise, G. A., Kiesel Filho, N., Komechen, H., Sabbaga, C. C., Rosati, R., et al. (2013). Impact of neonatal screening and surveillance for the TP53 R337H mutation on early detection of childhood adrenocortical tumors. *J. Clin. Oncol.* 31, 2619–2626. doi: 10.1200/JCO.2012.46.3711

DiGiammarino, E. L., Lee, A. S., Cadwell, C., Zhang, W., Bothner, B., Ribeiro, R. C., et al. (2002). A novel mechanism of tumorigenesis involving pH-dependent destabilization of a mutant p53 tetramer. *Nat. Struct. Biol.* 9, 12–16. doi: 10.1038/nsb730

Ding, L., Getz, G., Wheeler, D. A., Mardis, E. R., McLellan, M. D., Cibulskis, K., et al. (2008). Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 455, 1069–1075. doi: 10.1038/nature07423

Fitarelli-Kiehl, M., Macedo, G. S., Schlatter, R. P., Koehler-Santos, P., Matte, U. S., Ashton-Prolla, P., et al. (2016). Comparison of multiple genotyping methods for the identification of the cancer predisposing founder mutation p.R337H in TP53. *Genet. Mol. Biol.* 39, 203–209. doi: 10.1590/1678-4685-GMB-2014-0351

Gahr, S., Stoehr, R., Geissinger, E., Ficker, J. H., Brueckl, W. M., Gschwendtner, A., et al. (2013). EGFR mutational status in a large series of Caucasian European NSCLC patients: data from daily practice. *Br. J. Cancer* 109, 1821–1828. doi: 10.1038/bjc.2013.511

Garritano, S., Gemignani, F., Palmero, E. I., Olivier, M., Martel-Planche, G., Le Calvez-Kelm, F., et al. (2010). Detailed haplotype analysis at the TP53 locus in p.R337H mutation carriers in the population of Southern Brazil: evidence for a founder effect. *Hum. Mutat.* 31, 143–150. doi: 10.1002/humu.21151

Giacomazzi, J., Graudenz, M. S., Osorio, C. A., Koehler-Santos, P., Palmero, E. I., Zagonel-Oliveira, M., et al. (2014). Prevalence of the TP53 p.R337H mutation in breast cancer patients in Brazil. *PLoS One* 9:e99893. doi: 10.1371/journal.pone.0099893

Giacomazzi, J., Selistre, S. G., Rossi, C., Alemar, B., Santos-Silva, P., Pereira, F. S., et al. (2013). Li-Fraumeni and Li-Fraumeni-like syndrome among children diagnosed with pediatric cancer in Southern Brazil. *Cancer* 119, 4341–4349. doi: 10.1002/cncr.28346

Greenman, C., Stephens, P., Smith, R., Dalgliesh, G. L., Hunter, C., Bignell, G., et al. (2007). Patterns of somatic mutation in human cancer genomes. *Nature* 446, 153–158. doi: 10.1038/nature05610

Hahn, E. C., Bittar, C. M., Vianna, F. S. L., Netto, C. B. O., Biazús, J. V., Cericatto, R., et al. (2018). TP53 p.Arg337His germline mutation prevalence in Southern Brazil: further evidence for mutation testing in young breast cancer patients. *PLoS One* 13:e0209934. doi: 10.1371/journal.pone.0209934

Hammerman, P. S., Lawrence, M. S., Voet, D., Jing, R., Cibulskis, K., Sivachenko, A., et al. (2012). Comprehensive genomic characterization of squamous cell lung cancers. *Nature* 489, 519–525. doi: 10.1038/nature11404

Heymann, S., Delaloge, S., Rahal, A., Caron, O., Frebourg, T., Barreau, L., et al. (2010). Radio-induced malignancies after breast cancer postoperative radiotherapy in patients with Li-Fraumeni syndrome. *Radiat. Oncol.* 5:104. doi: 10.1186/1748-717X-5-104

Imielinski, M., Berger, A. H., Hammerman, P. S., Hernandez, B., Pugh, T. J., Hodis, E., et al. (2012). Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* 150, 1107–1120. doi: 10.1016/j.cell.2012.08.029

Instituto Nacional de Câncer José Alencar Gomes da Silva (Inca) (2020). *Estimativa 2020: Incidência de Câncer no Brasil.* https://www.inca.gov.br (Accessed September 15, 2020)

International Agency for Research on Cancer TP53 database (Iarc) (2020). *Selected Statistics on Somatic Mutations.* http://p53.iarc.fr/TP53SomaticMutations.aspx (Accessed September 15, 2020)

Jordan, E. J., Kim, H. R., Arcila, M. E., Barron, D., Chakravarty, D., Gao, J., et al. (2017). Prospective comprehensive molecular characterization of lung adenocarcinomas for efficient patient matching to approved and emerging therapies. *Cancer Discov.* 7, 596–609. doi: 10.1158/2159-8290.CD-16-1337

Kang, S. H., Cho, S. I., Chae, J. H., Chung, K. N., Ra, E. K., Kim, S. Y., et al. (2009). False homozygous deletions of SMN1 Exon 7 Using Dra I PCR-RFLP caused by a novel mutation in spinal muscular atrophy. *Genet. Test. Mol. Biomark.* 13, 511–513. doi: 10.1089/gtmb.2008.0158

Mastellaro, M. J., Ribeiro, R. C., Oliveira-Filho, A. G., Seidinger, A. L., Cardinalli, I. A., Miranda, E. C. M., et al. (2018). Adrenocortical tumors associated with the TP53 p.R337H germline mutation can be identified during child-care consultations. *J. Pediatr.* 94, 432–439. doi: 10.1016/j.jped.2017.06.009

Mastellaro, M. J., Seidinger, A. L., Kang, G., Abrahão, R., Miranda, E. C. M., Pounds, S. B., et al. (2017). Contribution of the TP53 R337H mutation to the cancer burden in southern Brazil: insights from the study of 55 families of children with adrenocortical tumors. *Cancer* 123, 3150–3158. doi: 10.1002/cncr.30703

Mezquita, L., Jové, M., Nadal, E., Kfoury, M., Morán, T., Ricordel, C., et al. (2020). High prevalence of somatic oncogenic driver alterations in patients With NSCLC and Li-Fraumeni Syndrome. *J. Thorac. Oncol.* 15, 1232–1239. doi: 10.1016/j.jtho.2020.03.005

Michalarea, V., Calcasola, M., Cane, P., Tobal, K., Izatt, L., and Spicer, J. (2014). EGFR-mutated lung cancer in Li-Fraumeni syndrome. *Lung Cancer* 85, 485–487. doi: 10.1016/j.lungcan.2014.06.017

Nogueira, S. T. S., Lima, E. N. P., Nóbrega, A. F., Torres, I. C. G., Cavicchioli, M., Hainaut, P., et al. (2015). 18F-FDG PET-CT for surveillance of Brazilian Patients with Li-Fraumeni syndrome. *Front. Oncol.* 5:38. doi: 10.3389/fonc.2015.00038

Palmero, E. I., Schüler-Faccini, L., Caleffi, M., Achatz, M. I., Olivier, M., Martel-Planche, G., et al. (2008). Detection of R337H, a germline TP53 mutation predisposing to multiple cancers, in asymptomatic women participating in a breast cancer screening program in Southern Brazil. *Cancer Lett.* 261, 21–25. doi: 10.1016/j.canlet.2007.10.044

Park, J. H., Li, J., Starost, M. F., Liu, C., Zhuang, J., Chen, J., et al. (2018). Mouse homolog of the human TP53 R337H mutation reveals its role in tumorigenesis. *Cancer Res.* 78, 5375–5383. doi: 10.1158/0008-5472.CAN-18-0016

Park, J. H., Wang, P. Y., and Hwang, P. M. (2019). Modeling the prevalent germline *TP53* R337H mutation in mouse. *Oncotarget* 10, 631–632. doi: 10.18632/oncotarget.26603

Peifer, M., Fernández-Cuesta, L., Sos, M. L., George, J., Seidel, D., Kasper, L. H., et al. (2012). Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nat. Genet.* 44, 1104–1110. doi: 10.1038/ng.2396

Ribeiro, R. C., Sandrini, F., Figueiredo, B., Zambetti, G. P., Michalkiewicz, E., Lafferty, A. R., et al. (2001). An inherited p53 mutation that contributes in a tissue-specific manner to pediatric adrenal cortical carcinoma. *Proc. Natl. Acad. Sci.*U.S.A. 98, 9330–9335. doi: 10.1073/pnas.161479898

Ricordel, C., Labalette-Tiercin, M., Lespagnol, A., Kerjouan, M., Dugast, C., Mosser, J., et al. (2015). EFGR-mutant lung adenocarcinoma and Li-Fraumeni syndrome: report of two cases and review of the literature. *Lung Cancer* 87, 80–84. doi: 10.1016/j.lungcan.2014.11.005

Rizvi, N. A., Hellmann, M. D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J. J., et al. (2015). Mutational landscape determines sensitivity to PD-1 blockade in non–small cell lung cancer. *Science* 348, 124–128. doi: 10.1126/science.aaa1348

Ruiz-Linares, A., Adhikari, K., Acuña-Alonzo, V., Quinto-Sanchez, M., Jaramillo, C., Arias, W., et al. (2014). Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. *PLoS Genet.* 10:e1004572. doi: 10.1371/journal.pgen.1004572

Seidinger, A. L., Fortes, F. P., Mastellaro, M. J., Cardinalli, I. A., Zambaldi, L. G., Aguiar, S. S., et al. (2015). Occurrence of Neuroblastoma among TP53 p.R337H Carriers. *PLoS One* 10:e0140356. doi: 10.1371/journal.pone.0140356

Seidinger, A. L., Mastellaro, M. J., Paschoal Fortes, F., Godoy Assumpção, J., Aparecida Cardinalli, I., Aparecida Ganazza, M., et al. (2011). Association of the highly prevalent TP53 R337H mutation with pediatric choroid plexus carcinoma and osteosarcoma in Southeast Brazil. *Cancer* 117, 2228–2235. doi: 10.1002/cncr.25826

Tinat, J., Bougeard, G., Baert-Desurmont, S., Vasseur, S., Martin, C., Bouvignies, E., et al. (2009). 2009 Version of the chompret criteria for Li Fraumeni syndrome. *J. Clin. Oncol.* 27, e108–e109. doi: 10.1200/JCO.2009.22.7967

Uemura, T., Hibi, K., Kaneko, T., Takeda, S., Inoue, S., Okochi, O., et al. (2004). Detection of K-ras mutations in the plasma DNA of pancreatic cancer patients. *J. Gastroenterol.* 39, 56–60. doi: 10.1007/s00535-003-1245-1

Wang, W., Shen, X. B., Jia, W., Huang, D. B., Wang, Y., and Pan, Y. Y. (2019). The p53/miR-193a/EGFR feedback loop function as a driving force for non-small cell lung carcinoma tumorigenesis. *Ther. Adv. Med. Oncol.* 11:1758835919850665. doi: 10.1177/1758835919850665

Zerdoumi, Y., Lanos, R., Raad, S., Flaman, J. M., Bougeard, G., Frebourg, T., et al. (2017). Germline TP53 mutations result into a constitutive defect of p53 DNA binding and transcriptional response to DNA damage. *Hum. Mol. Genet.* 26, 2591–2602. doi: 10.1093/hmg/ddx106

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Journal of Clinical Oncology®
An American Society of Clinical Oncology Journal

Molecular profiling as predictor of outcomes in a Brazilian cohort of stage IV lung cancer.

Nicolas Peruzzo, Juliano Ce Coelho, Gabriel de Souza Macedo, Tiago Finger Andreis, Gustavo Gössling, Pedro Grachinski Buiar, Débora Leite Rocha, Patricia Ashton-Prolla

Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil; Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil; Hospital de Clinicas de Porto Alegre, Porto Alegre, RS, Brazil

Show Less

Abstract Disclosures

Abstract

e20668

Background: Lung cancer is the leading cause of cancer deaths globally. In stage IV non-small cell lung cancer (NSCLC), identification of a tumor driver mutation is critical, since it tailors treatment. We aimed to analyze clinical outcomes according to molecular profiling in patients (pts) with stage IV NSCLC. **Methods:** In this retrospective cohort study, we enrolled 57 pts with stage IV NSCLC treated at a public hospital in Southern Brazil between 2016 and 2018. **Results:** Median follow-up was 20.3 months, 53% were men, mean age was 65 years, 86% had smoked, 84% had *de novo* metastatic NSCLC and 96% had non-squamous carcinoma. Regarding molecular profiling, somatic mutations in KRAS, EGFR and BRAF were identified in 33%, 16% and 4% of pts, respectively. None of the pts had ALK mutations, and in 47% no identifiable driver mutation was found. In the EGFR-mutated subgroup, 67% had a deletion of exon 19, 22% had the exon 21 L858R mutation and 11% had exon 20 mutations. Palliative systemic therapy (PST) was delivered to 60% of the pts. Two or more lines of PST were delivered to 23%. The main reason for upfront best supportive care was ECOG PS 3-4 (poor). In the subgroup of pts with sensitizing EGFR mutations (8 pts), 75% received Gefitinib, the only anti-EGFR drug available in our public health system; the other pts died before recognition of the mutation. Median Progression Free-Survival was 6.3 months overall, 10.3 months for EGFR-mutated pts, 7.6 months for KRAS-mutated, 7.5 months for BRAF-mutated, and 5.2 months for pts with no mutation identified. Median Overall Survival (OS) was 7.7 months overall, 13.2 months for EGFR-mutated pts, 7.4 months for KRAS-mutated, 12.9 months for BRAF-mutated, and 5.3 months for pts with no mutation identified. In the Cox regression multivariate analysis, poor PS (HR 3.80, P < 0.0001) and second-line PST (HR 0.23, P = 0.002) were independent predictors of OS. No driver mutations were predictors of OS, although we did find a tendency towards better OS in pts with EGFR mutations and worse OS in pts with KRAS mutations or no identifiable mutation. **Conclusions:** In this cohort, genotyping results did not predict survival outcomes. This is probably due to the small number of pts studied, and data should be reanalyzed in larger cohorts.

OPTIONS & TOOLS

Export Citation
Track Citation
Add To Favorites
Rights & Permissions

COMPANION ARTICLES

No companion articles

ARTICLE CITATION

DOI: 10.1200/JCO.2019.37.15_suppl.e20668
Journal of Clinical Oncology 37, no. 15_suppl

Published online May 26, 2019.

WE RECOMMEND

Mutation spectrum of major cancer driver genes in Vietnamese NSCLC patients.
Uyen Vu Tran et al., J Glob Oncol, 2019

Phase II randomized study of biomarker-directed treatment for non-small cell lung cancer (NSCLC): The BATTLE (Biomarker-Integrated Approaches of Targeted Therapy for Lung Cancer Elimination) clinical trial program
E. S. Kim, J Clin Oncol, 2016

Outcomes in Lung Cancer: 9-Year Experience From a Tertiary Cancer Center in India
Aditya Navile Murali et al., J Glob Oncol, 2017

Molecular Characteristics of Bronchioloalveolar Carcinoma and Adenocarcinoma, Bronchioloalveolar Carcinoma Subtype, Predict Response to Erlotinib
Vincent A. Miller et al., J Clin Oncol, 2008

A phase II trial of salirasib in patients with stage IIIB/IV lung adenocarcinoma enriched for KRAS mutations
M. L. Johnson et al., J Clin Oncol, 2016

How Does Genetics Affect Outcomes in Nonsmokers With NSCLC?
Maurie Markman, MD, Medscape

The Distinct Molecular and Clinical Profile of Lung Cancer in Young Patients
H. Jack West, MD, Medscape

PGx Investigation of Bayers Nexavar in KRAS-Mutated NSCLC Patients Disappoints
GenomeWeb, 2012

PGx Investigation of Bayers Nexavar in KRAS-Mutated NSCLC Patients Disappoints
GenomeWeb, 2012

Prognostic model of survival outcomes in non-small cell lung cancer patients initiated on afatinib: pooled analysis of clinical trial data
Ashley M. Hopkins et al., Cancer Biology & Medicine, 2019

Powered by TREND MD

PLOS ONE

RESEARCH ARTICLE

Clinical and molecular characterization of patients fulfilling Chompret criteria for Li-Fraumeni syndrome in Southern Brazil

Camila Matzenbacher Bittar[1,2☯], Yasminne Marinho de Araújo Rocha[2☯], Igor Araujo Vieira[1,2☯], Clévia Rosset[1,2‡], Tiago Finger Andreis[1,2‡], Ivaine Tais Sauthier Sartor[3], Osvaldo Artigalás[3], Cristina B. O. Netto[4], Barbara Alemar[1,2], Gabriel S. Macedo[2], Patricia Ashton-Prolla[1,2,4]*

1 Programa de Pós-Graduação em Genética e Biologia Molecular (PPGBM), Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil, **2** Laboratório de Medicina Genômica, Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Rio Grande do Sul, Brazil, **3** Hospital Moinhos de Vento (HMV), Porto Alegre, Rio Grande do Sul, Brazil, **4** Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Rio Grande do Sul, Brazil

☯ These authors contributed equally to this work.
‡ These authors also contributed equally to this work.
* pprolla@gmail.com

OPEN ACCESS

Citation: Matzenbacher Bittar C, de Araújo Rocha YM, Vieira IA, Rosset C, Andreis TF, Sartor ITS, et al. (2021) Clinical and molecular characterization of patients fulfilling Chompret criteria for Li-Fraumeni syndrome in Southern Brazil. PLoS ONE 16(9): e0251639. https://doi.org/10.1371/journal.pone.0251639

Editor: Amanda Ewart Toland, The Ohio State University, UNITED STATES

Received: December 9, 2020

Accepted: April 29, 2021

Published: September 16, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pone.0251639

Abstract

Li-Fraumeni syndrome (LFS) is an autosomal dominant cancer predisposition syndrome caused by pathogenic germline variants in the *TP53* gene, characterized by a predisposition to the development of a broad spectrum of tumors at an early age. The core tumors related to LFS are bone and soft tissue sarcomas, premenopausal breast cancer, brain tumors, adrenocortical carcinomas (ACC), and leukemias. The revised Chompret criteria has been widely used to establish clinical suspicion and support *TP53* germline variant testing and LFS diagnosis. Information on *TP53* germline pathogenic variant (PV) prevalence when using Chompret criteria in South America and especially in Brazil is scarce. Therefore, the aim of this study was to characterize patients that fulfilled these specific criteria in southern Brazil, a region known for its high population frequency of a founder *TP53* variant c.1010G>A (p. Arg337His), as known as R337H. *TP53* germline testing of 191 cancer-affected and independent probands with LFS phenotype identified a heterozygous pathogenic/likely pathogenic variant in 26 (13.6%) probands, both in the DNA binding domain (group A) and in the oligomerization domain (group B) of the gene. Of the 26 carriers, 18 (69.23%) were R337H heterozygotes. Median age at diagnosis of the first tumor in groups A and B differed significantly in this cohort: 22 and 2 years, respectively ($P = 0.009$). The present study shows the clinical heterogeneity of LFS, highlights particularities of the R337H variant and underscores the need for larger collaborative studies to better define LFS prevalence, clinical spectrum and penetrance of different germline *TP53* pathogenic variants.

Introduction

Li-Fraumeni (LFS) syndrome is an autosomal dominant cancer predisposition disorder mainly caused by pathogenic and likely pathogenic germline variants (PV) in the *TP53* tumor

Data Availability Statement: All relevant data are within the manuscript and its S1 Fig and S1,S2 Tables files.

Funding: This study was funded by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Grant # 478430/2012-4), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) (Grant # 16/2551-0000486-2), and Fundo de Incentivo à Pesquisa do Hospital de Clínicas de Porto Alegre (FIPE-HCPA) to PA-P. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

suppressor gene encoding the p53 protein. Although any tumor can be identified in LFS carriers, "core" tumors of the syndrome have been reported and include premenopausal breast cancer, bone and soft-tissue sarcoma, brain cancer, leukemia and adrenocortical carcinoma (ACC). Carriers of germline *TP53* PV have a variable lifetime risk of developing cancer, and phenotype may vary from fully penetrant LFS to cancer-free over a lifetime. Nevertheless, about 50% of carriers develop at least one malignancy by age 30, especially those with *TP53* DNA-binding domain (DBD) variants, also called "classic" variants, which represent approximately 86% of the *TP53* pathogenic variants associated with the LFS phenotype in most countries [1–3].

Population prevalence studies have estimated that germline *TP53* PV occur at a frequency of 1 in 5,000 to 1 in 20,000 individuals [4]. In more recent studies, prevalence of *TP53* PV heterozygotes was proposed to reach 0.2% in Europeans [5, 6]. In addition, a germline *TP53* founder PV, c.1010G>A (p.Arg337His), widely referred as R337H, has been reported in Southern Brazil at a frequency of 1 in approximately 300 newborns [7–9], but tumor penetrance appears to be lower than that observed in carriers of DNA-binding domain (DBD) PV [10–13]. The arginine residue at codon 337 is involved in the protein oligomerization and functional data have shown that its replacement with histidine disrupts the tetramer form, making the domain unable to fully oligomerize in conditions of slightly elevated pH [14]. Although it was initially described as a "tissue-specific sequence variant" related only to ACC, today it is considered to be a PV related to the occurrence of multiple tumors, in a spectrum similar to that of LFS [15, 16]. Recent findings from a mouse model provided *in vivo* evidence that the R337H PV decreased p53 transactivation potential and renders mice susceptible to carcinogen-induced liver tumorigenesis [17].

Clinical criteria to define diagnosis of LFS were established based on the first study by Li and Fraumeni [18]. Approximately 70% to 80% of patients who fulfill classical criteria will have a germline PV in *TP53* [16, 19] When a broader LFS tumor spectrum was considered, a number of different sets of criteria started to be used to identify LFS patients, including the Chompret criteria and other criteria for Li-Fraumeni Like Syndrome (LFL) [19–21]. Importantly, diagnostic criteria defined by Chompret have increased the sensitivity of *TP53* germline PV detection by including patients with the core LFS tumors even without a family history. The revised Chompret criteria [21–23] had a PV detection rate of 18% and, when incorporated as part of *TP53* testing criteria along with classic LFS criteria, have been shown to improve the diagnostic sensitivity to 95% (Classic and Chompret criteria together) [2]. Therefore, the National Comprehensive Cancer Network (NCCN) and several other guidelines recommend using both the Classic LFS and the revised Chompret criteria to indicate germline *TP53* genetic testing [24].

So far, only a few studies showed the prevalence of germline *TP53* PV in individuals from Southern Brazil, in which the prevalence of 28,8% and 11,4% were found in a case series of 45 and 70 probands fulfilling any LFS criteria [25, 26]. In the present study, we aimed to characterize the clinical and molecular profile in a series of LFS patients fulfilling the 2015 revised Chompret criteria and recruited from cancer risk evaluation clinics in southern Brazil. These results can help to better define the LFS prevalence in Southern Brazil and also points out to differences in the clinical spectrum among carriers of distinct PV in *TP53*.

Materials and methods

Patients and ethical aspects

From July 2015 to January 2019, 211 independent cancer-affected patients from unique families with a suggestive clinical phenotype of LFS were identified at a public hospital and private cancer risk evaluation clinics in Southern Brazil. Of these, 191 were residents of the Southern region of Brazil, met the 2015 revised Chompret criteria and were included in the present

study. The majority of patients, 148 patients were from a reference public hospital (Hospital de Clínicas de Porto Alegre), seen at the institutional's outpatient cancer genetics clinic (108) and pediatric cancer ward (40). The additional 43 probands were identified in 4 private cancer genetics clinics in the same city. S1 Fig is a Consort Diagram that depicts the recruitment and testing process, while the S1 Table lists 2015 revised Chompret criteria. The study was approved by the Institutional Review Board. All participants underwent pre- and post-test genetic counseling, provided informed written or verbal consent for the study. When verbal consent was obtained, it was registered on participant clinical chart. Parents signed the consent for participants that were minors. Personal clinical history, self-reported family history and previous testing results were collected from patient interviews or medical records.

Molecular analysis

Of the 191 patients participating in this study, 43 had previously undergone multi-gene panel testing (MGPT) including *TP53* sequence variant and rearrangement testing using Next-Generation Sequencing (NGS, retrospectively tested), 99 patients had undergone previous analysis of the *TP53* coding region by Sanger sequencing and Multiplex Ligation-Dependent Probe Amplification (MLPA) (also retrospectively tested), and 49 patients were offered molecular testing in the institutional research laboratory at recruitment (prospectively tested). *TP53* genotyping in the latter was performed employing two methodologies: (1) NGS in peripheral blood samples using the Ion AmpliSeq ™ Panel *TP53* kit (Thermo Fisher Scientific) and Ion GeneStudio S5 system (Ion Torrent Systems Inc, Gilford, NH); and (2) MLPA using the SALSA MLPA P056 kit (MRC Holland, Amsterdam, Netherlands), followed by capillary gel electrophoresis with the Applied Biosystem 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) and analyses of the copy number variations conducted in the Coffalyser. Net software (MRC-Holland®) [27].

Statistical analysis

Tumor spectrum and clinical characteristics of carriers of DBD variants (group A) and R337H variant (group B) were compared. Data normality assumptions were verified on the age of group A and B and Mann-Whitney-Wilcoxon test was performed. To measure the association among the groups, gender, type of cancer and multiple tumors we used Pearson's Chi-Squared test or Fisher's exact test. Odds ratio with 95% confidence intervals were also calculated. To compare the pathogenic variant detection rate in this study and the rate found in Bougeard et al in 2015 [2], we used Pearson's Chi-Squared test with Yates continuity correction. We also divided our probands in three groups (hotspot DBD variant carriers; R337H carriers and DBD non hotspot variant carriers) and Kruskal-Wallis test followed by Benjamini-Hochberg correction for multiple comparisons was performed. All data analyses were performed in R 3.4.2 statistical software.

Results

Germline PV *TP53* were identified in twenty-six (13.6%) of the 191 probands included in the study. One of the carriers was homozygote and the other 25 carriers of germline PV were heterozygotes. 18 (69.2%) harboured the Brazilian founder R337H variant and 8 probands (30.8%) had a PV in the *TP53* DBD. MLPA analysis identified no *TP53* deletions and/or duplications in this series. Fig 1 shows the location of each pathogenic alteration detected in the gene and Table 1 summarizes the clinical and molecular results of all PV-positive probands (S2 Table exhibits the characterization of all probands analyzed). Fig 2 depicts the NGS results encompassing the entire *TP53* coding region from two probands.

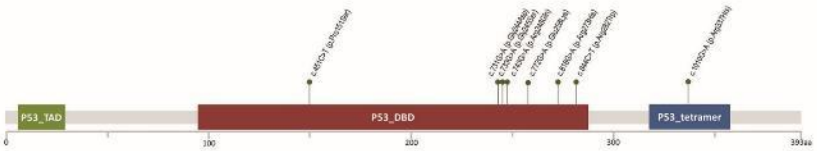


Fig 1. Location of the *TP53* pathogenic variants detected in the *p53* protein functional domains. Green dots represent the variants identified in the present study. P53_TAD, transactivation domain; P53_DBD, DNA binding domain; P53_oligomer, oligomerization domain.

https://doi.org/10.1371/journal.pone.0251639.g001

Important differences were observed when comparing the tumor spectrum and clinical characteristics of carriers of DBD variants (group A) and R337H variant (group B) (Table 2). The median age at first cancer diagnosis was 22 years in group A and 2 years in group B ($P = 0.009$; Mann-Whitney-Wilcoxon test). Fifteen patients (83.3%) in group B and only 3 (37.5%) in group A developed a tumor before age 18 years. Most of the tumors (13, 72.22%) observed in group B were ACC (all under 18 years), and only one ACC (12.5%) was observed in group A (diagnosed at age 44 years). Finally, multiple primary tumors were observed only among patients from group A, including 4 (50%) patients. Interestingly, one proband had been diagnosed with 4 primary tumors: osteosarcoma, bilateral breast cancer and soft tissue sarcoma; all before age 25 years. The tumor spectrum of the PV carriers is depicted in Fig 3 and shows evident differences between groups (DBD variant and R337H carriers), especially regarding ACC.

As observed in Table 3, a significant association was found in the comparative analyses including type of cancer and multiple tumors. A higher prevalence of ACC was observed in group B when compared to group A patients ($P = 0.043$; Chi-squared test) and the presence of multiple tumors was most frequent in group A ($P = 0.005$, Fisher exact test). Additionally, we classified the DBD variants in two groups, namely: group A non-hotspot PV, which comprised of p.(Pro151Ser), p.(Gly244Asp) and p.(Glu258Lis) variants; and group A hotspot PV (p.(Gly245Ser), p.(Arg248Gln), p.(Arg273His), p.(Arg282Trp)). When comparing the median age at first diagnosis of cancer in patients from group A non-hotspot PV, group A hotspot PV, and group B (R337H variants) we observed a significative difference ($P = 0.021$), being 31.8, 12.1 and 2.35 years respectively. The post-hoc analysis pointed out that age at first diagnosis was different between group B and A non-hotspot PV (data not shown).

Discussion

LFS is considered a rare cancer predisposition disorder worldwide. In Southern Brazil, due to presence of a germline founder pathogenic variant in the *TP53* oligomerization domain (R337H), it is estimated that 0,3% of the general population carries this variant [12]. Despite significant heterozygote frequency at the population level, little information is available on the prevalence of germline *TP53* PV among individuals with a suggestive phenotype, i.e. fulfilling revised Chompret criteria. This information is important to guide health care policies for cancer prevention and treatment in the region. Identifying LFS patients is important to determine adequate clinical surveillance and follow up, not only in the proband but in his/her relatives, since detection of a carrier provides the opportunity for cascade testing and, if additional carriers are identified in the family, they can be referred to appropriate genetic counseling and specific high risk screening protocols [28]. Villani and colleagues (2016) demonstrated that carriers of pathogenic *TP53* variants benefit enormously from an enhanced surveillance protocol, including frequent physical examination plus targeted biochemical monitoring and

Table 1. Clinical and molecular characterization of all LFS probands harboring germline *TP53* pathogenic variants (PV) identified in this study.

Proband ID / Gender	Age at 1st cancer diagnosis (years)	Proband's type of cancer	Age at diagnosis, other tumors (years)	2015 Version Chompret Criterion(s)	Recruitment	Genetic Testing	chr17 position on Assembly GRCh37 (dbSNP rs ID)	*TP53* variant HGVS c.	*TP53* variant HGVS p.
166 / F	32	Breast	Breast (38)	Familial	PC	Sanger + MLPA	rs28934874	c. 451C>T	p. (Pro151Ser)
167 / F	30	Breast (bilateral)	Thyroid (37)	Familial, EOBC	PUB	Sanger + MLPA	rs1057517983	c.731G>A	p. (Gly244Asp)
168 / F	11	CNS	NA	Familial	PUB	Sanger + MLPA	rs28934575	c.733G>A	p. (Gly245Ser)
169 / F	12	OS	Breast (21), Breast (22), STS (24)	MT, EOBC	PC	MGPT	rs28934575	c.733G>A	p. (Gly245Ser)
170 / F	25	Breast	NA	EOBC	PC	MGPT	rs11540652	c.743G>A	p. (Arg248Gln)
171 / M	44	ACC	NA	RT	PUB	NGS + MLPA	rs121912652	c.772G>A	p. (Glu258Lys)
172 / F	19	OS	Breast (29), STS (38)	Familial, MT, EOBT	PUB	Sanger + MLPA	rs28934576	c.818G>A	p. (Arg273His)
173 / F	5	CNS (CPC)	NA	RT	PC	MGPT	rs28934574	c.844C>T	p. (Arg282Trp)
174 / F	0 (6 mo)	ACC	NA	RT	PED	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
175 / F	0 (4 mo)	ACC	NA	Familial, RT	PED	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
176 / F	0 (8 mo)	ACC	NA	RT	PED	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
177 / F	1	ACC	NA	Familial, RT	PUB	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
178 / M	1	ACC	NA	RT	PED	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
179 / M	2	ACC	NA	RT	PED	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
180 / M	2	ACC	NA	RT	PUB	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
181 / F	3	ACC	NA	RT	PUB	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
182 / F	3	ACC	NA	RT	PUB	NGS + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
183 / F	5	ACC	NA	RT	PED	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
184 / F	11	ACC	NA	RT	PED	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
185 / M	17	ACC	NA	RT	PED	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
186 / F	23	Breast	NA	Familial, EOBC	PUB	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
187 / F	57	Breast	NA	Familial	PUB	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
188 / F	49	Breast (bilateral)	NA	Familial	PUB	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
189 / M	1	CNS (CPC)	NA	RT	PED	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)
190 / M	1	CNS (CPC)	NA	RT	PED	Sanger + MLPA	rs121912664	c.1010G>A	p. (Arg337His)

(*Continued*)

Table 1. (Continued)

Proband ID / Gender	Age at 1st cancer diagnosis (years)	Proband's type of cancer	Age at diagnosis, other tumors (years)	2015 Version Chompret Criterion(s)	Recruitment	Genetic Testing	chr17 position on Assembly GRCh37 (dbSNP rs ID)	*TP53* variant HGVS c.	*TP53* variant HGVS p.
191 / F	1	ACC	NA	Familial, RT	PED	Sanger + MLPA	rs121912664	c.1010G>A	p.(Arg337His)*

ACC, Adrenocortical Carcinoma; CNS, Central Nervous System; CPC, Choroid Plexus Carcinoma; EOBC, Early Onset Breast Cancer; MGPT, Multigene Panel Testing; MT, Multiple Tumors; MO, months old; OS, Osteosarcoma; RT, Rare Tumors, STS, Soft tissue sarcoma; NA, not applicable; PUB, high-risk public clinic; PC, high-risk private clinic; PED, pediatric tumors database; NGS, Next-generation Sequencing; MLPA, Multiplex Ligation-Dependent Probe Amplification; WT, wild-type genotype

* homozygous for the R337H variant.

https://doi.org/10.1371/journal.pone.0251639.t001

periodic imaging screens (ultrasounds, brain magnetic resonance images, and rapid whole body MRI scans) [28]. Collectively, this approach has a significant impact in overall survival, compared to patients that do not undergo enhanced surveillance. In Brazil, although patients with health insurance have access to genetic testing if they fulfill the revised Chompret criteria, those that rely solely on the public health care system (about 70% of the population) must pay out of pocket to have this information, since genetic testing for cancer predisposition is not yet payed in the public setting.

In this cohort, tumoral spectrum in R337H carriers was similar to that already described in literature, especially when compared to previous studies performed in other Brazilian Centers. However, in the present study a strikingly higher prevalence of ACC was observed in R337H carriers when compared to carriers of DBD variants ($P = 0.043$; chi-squared test). From this observation we can conclude that in the series presented here, ACC was the most prevalent tumor observed in association with R337H whereas the previous Brazilian study described breast cancer as the most frequent tumor (30%) [25].

Regarding PV detection rate for the 2015 Chompret Criteria identified here (13,6%), this rate is similar to the 18% described by Bougeard *et al.* in 2015 in France ($P = 0.2482$; chi-squared test with Yates correction), but it is mainly due to the presence of the R337H variant [2]. Of note, in the previous study by Andrade *et al.* (2017) of Brazilian patients from the Southeastern region, PV detection rate in 17 probands with the 2015 Chompret Criteria was much higher, 35% [26]. These differences between the studies from Southern and Southeastern Brazil may reflect regional genetic modifiers of the phenotype (i.e. additional genetic risk factors), regional environmental factors or different recruitment strategies in each study.

Regarding genotype-phenotype correlations, it is well know that DBD hotspot variants with reported dominant negative effects, such as p.(Gly245Ser), p.(Arg248Gln), p.(Arg273His) and p.(Arg282Trp) are associated with earlier onset cancers and stronger family history of tumors within the LFS spectrum [29]. On the other hand, several previous studies from Brazilian cohorts have suggested that R337H is a PV with lower prevalence associated with cancer diagnoses at older ages, although a bimodal distribution of age at cancer diagnosis has also been suggested [30, 31]. Contrary to the expected phenotype, probands with the R337H variant in this study had earlier age at first tumor diagnosis when compared with carriers of DBD variants. To analyze this data in more detail, we divided our probands in three groups according to type of PV (non-hotspot DBD, hotspot DBD and R337H) and observed that median age at first tumor diagnosis among groups with the lowest mean age identified among R337H carriers.

The results of the present study are relevant for two main reasons. First, they underscore the importance of considering that significant regional differences may occur and that criteria

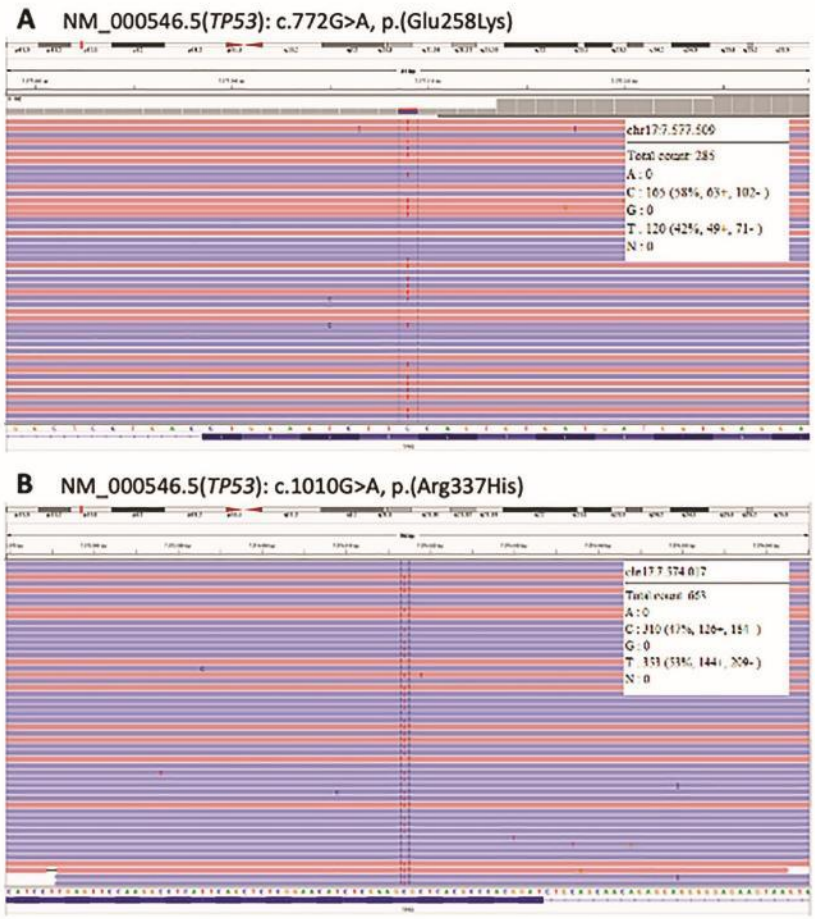


Fig 2. Representative next-generation sequencing results encompassing the *TP53* entire coding region (minimum coverage of 100X by amplicon) from two probands fulfilling the 2015 revised Chompret criteria for Li-Fraumeni syndrome. (A) Carrier of a germline pathogenic variant (PV) located in the p53 DNA binding domain (DBD); and (B) carrier of the Brazilian founder R337H PV located in the p53 oligomerization domain. Description of *TP53* sequence variants is provided according to updated Human Genome Variation Society (HGVS) recommendations. Human *TP53* sequence corresponding to the NM_000546.5 was used as a wild-type reference. Right panels show wild-type and variant allele counts, which were consistent with the expected germline occurrence of these genetic alterations (around 50% of reads for each allele). Note that both alleles were analyzed from antisense strand due to the *TP53* gene orientation. Chr17, position or genomic coordinate at chromosome 17 (GRCh37/hg19 human genome assembly).

https://doi.org/10.1371/journal.pone.0251639.g002

established for one population may not have the same performance in another population. Considering that the population of Southern Brazil is mostly of European ancestry, one would expect to see a prevalence of germline *TP53* PV variants similar to that observed in Europeans. A high frequency of R337H among probands with a phenotype suggesting LFS had been previously reported by Achatz et al. (2007) (46,1% of those with coding region *TP53* variants), but

Table 2. Distribution of tumor types in all LFS PV-positive patients (n = 26).

Tumor types diagnosed in PV carriers	Number of tumors per PV group (A/B)	OR (95% CI), p value	Number of patients per group (A/B)	% PV carriers per group (A/B)	Age at diagnosis (range when >1) in each group (A/B)
Adrenocortical Carcinoma	1 / 13	16.0 (1.5–875.8), **0.009**	1 / 13	12.5 / 72.2	44 / 0 to 17
Breast	8 / 4	0.18 (0.03–1.0), **0.03**	5 / 3	62.5 / 16.6	21 to 38 / 23 to 57
CNS	2 / 2	0.39 (0.02–6.53), 0.56	2/ 2	25 / 11.1	5 to 11 / 1
Osteosarcoma	3 / NA	-	2 / NA	25 / NI	12 to 38 / NA
Thyroid	1 / NA	-	1 / NA	12.5 / NI	37 / NA

DBD, pathogenic variants located in the DNA-binding domain; CNS, Central Nervous System tumors; NA, not applicable; NI, not identified.

https://doi.org/10.1371/journal.pone.0251639.t002

these authors did not restrict their recruitment to patients fulfilling Chompret criteria [25]. Second, results from the present analysis, in which overall *TP53* germline PV detection rate in Chompret criteria fulfilling probands was lower than expected from previous studies, may suggest that a different set of pathogenic variants, not yet mapped (i.e. located in intronic or regulatory regions of *TP53*) may be associated with the LFS phenotype in this particular region. It is also possible that PV in other, yet unidentified genes are associated with the LFS phenotype, accounting for the "missing heritability" of more than 85% observed here [32, 33]. An important limitation of the present study, that must be accounted for when analyzing the results is this study, is that a significant proportion of data on genetic testing were obtained retrospectively and with different variant detection strategies. Thus, further analyses on a prospectively recruited cohort of probands fulfilling Chompret criteria and then, clinical assessment of families carrying either DBD PV or R337H will be important to confirm these findings. Expanding

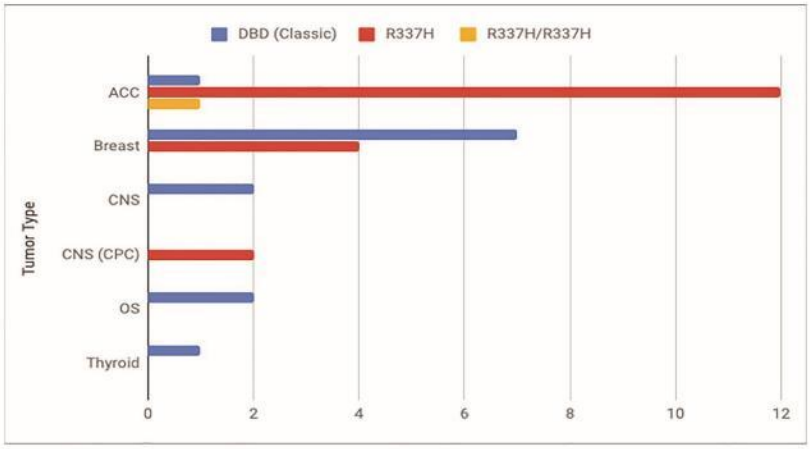


Fig 3. Graphic showing the differences between the tumor spectrum observed in carriers of the DBD variants, R337H variant and R337H homozygous proband. ACC, adrenocortical carcinoma; CNS, central nervous system; CPC, choroid plexus carcinoma; DBD, DNA binding domain; OS, osteosarcoma.

https://doi.org/10.1371/journal.pone.0251639.g003

Table 3. Association of gender, age at first tumor diagnosis, tumor type and development of multiple tumors among carriers of different groups of germline PV *TP53* (groups A and B).

	Group of pathogenic germline variants (PV)		
	A (n = 8)	B (n = 18)	*P* value
Gender			
Female	7	12	0.375*
Male	1	6	
Age at first cancer diagnosis, median (IQR)	22 (11.7–30.5)	2.0 (1.0–9.5)	**0.009****
Tumor types			
ACC	1	13	**0.043**†
Breast	2	2	
Breast bilateral	1	1	
CNS	1	0	
CNS (CPC)	1	2	
OS	2	0	
Multiple tumors			
Yes	4	0	**0.005***
No	4	18	

† Pearson Chi-squared test.

* Fisher exact test.

** Mann-Whitney-Wilcoxon test.

ACC, adrenocortical carcinoma; OS, osteosarcoma; CNS, central nervous system; CPC, choroid plexus carcinoma.

https://doi.org/10.1371/journal.pone.0251639.t003

this study in the region will be essential to instrument policy makers in establishing cancer screening protocols for these individuals.

Conclusions

The current study shows the impressive clinical heterogeneity of LFS, highlights particularities of the founder *TP53* pathogenic variant R337H and points to the need for larger and collaborative studies to better define LFS prevalence, clinical spectrum and penetrance of different types of PVs in the Brazilian population.

Supporting information

S1 Fig. Consort Diagram representing the patient recruitment and genetic testing process employed in the current study.
(DOCX)

S1 Table. 2015 revised Chompret criteria for LFS and *TP53* gene testing.
(DOCX)

S2 Table. Clinical and molecular characterization of all probands (n = 191) included in the study.
(DOCX)

Acknowledgments

We would like to thank Gustavo Stumpf da Silva and Patricia Santos-Silva for their valuable contributions and technical support.

Author Contributions

Conceptualization: Camila Matzenbacher Bittar, Igor Araujo Vieira, Cristina B. O. Netto, Barbara Alemar, Gabriel S. Macedo, Patricia Ashton-Prolla.

Data curation: Yasminne Marinho de Araújo Rocha, Clévia Rosset, Cristina B. O. Netto, Gabriel S. Macedo, Patricia Ashton-Prolla.

Formal analysis: Camila Matzenbacher Bittar, Igor Araujo Vieira, Clévia Rosset, Tiago Finger Andreis, Patricia Ashton-Prolla.

Funding acquisition: Patricia Ashton-Prolla.

Investigation: Camila Matzenbacher Bittar, Yasminne Marinho de Araújo Rocha, Igor Araujo Vieira, Clévia Rosset, Tiago Finger Andreis, Ivaine Tais Sauthier Sartor, Osvaldo Artigalás, Cristina B. O. Netto, Barbara Alemar, Gabriel S. Macedo, Patricia Ashton-Prolla.

Methodology: Camila Matzenbacher Bittar, Yasminne Marinho de Araújo Rocha, Igor Araujo Vieira, Clévia Rosset, Tiago Finger Andreis, Ivaine Tais Sauthier Sartor, Barbara Alemar, Gabriel S. Macedo, Patricia Ashton-Prolla.

Project administration: Camila Matzenbacher Bittar.

Resources: Osvaldo Artigalás, Cristina B. O. Netto.

Software: Yasminne Marinho de Araújo Rocha, Ivaine Tais Sauthier Sartor.

Supervision: Patricia Ashton-Prolla.

Validation: Yasminne Marinho de Araújo Rocha.

Writing – original draft: Camila Matzenbacher Bittar, Patricia Ashton-Prolla.

Writing – review & editing: Camila Matzenbacher Bittar, Igor Araujo Vieira, Clévia Rosset, Tiago Finger Andreis, Osvaldo Artigalás, Cristina B. O. Netto, Barbara Alemar, Gabriel S. Macedo, Patricia Ashton-Prolla.

References

1. Bougeard G, Sesbo È R, Baert-Desurmont S, Vasseur S, Martin C, Tinat J, et al. Molecular basis of the Li-Fraumeni syndrome: an update from the French LFS families. J Med Genet. 45. England 2008. p. 535–8. https://doi.org/10.1136/jmg.2008.057570 PMID: 18511570
2. Bougeard G, Renaux-Petel M, Flaman JM, Charbonnier C, Fermey P, Belotti M, et al. Revisiting Li-Fraumeni Syndrome From TP53 Mutation Carriers. J Clin Oncol. 2015; 33(21):2345–52. https://doi.org/10.1200/JCO.2014.59.5728 PMID: 26014290
3. Fischer NW, Prodeus A, Tran J, Malkin D, GariÈpy J. Association Between the Oligomeric Status of p53 and Clinical Outcomes in Li-Fraumeni Syndrome. J Natl Cancer Inst. 2018; 110(12):1418–21. https://doi.org/10.1093/jnci/djy114 PMID: 29955864
4. Lalloo F, Varley J, Ellis D, Moran A, O'Dair L, Pharoah P, et al. Prediction of pathogenic mutations in patients with early-onset breast cancer by family history. Lancet. 2003; 361(9363):1101–2. https://doi.org/10.1016/S0140-6736(03)12856-5 PMID: 12672316
5. Buys SS, Sandbach JF, Gammon A, Patel G, Kidd J, Brown KL, et al. A study of over 35,000 women with breast cancer tested with a 25-gene panel of hereditary cancer genes. Cancer. 2017; 123 (10):1721–30. https://doi.org/10.1002/cncr.30498 PMID: 28085182
6. Fortuno C, James PA, Spurdle AB. Current review of TP53 pathogenic germline variants in breast cancer patients outside Li-Fraumeni syndrome. Hum Mutat. 2018; 39(12):1764–73. https://doi.org/10.1002/humu.23656 PMID: 30240537
7. Achatz MI, Hainaut P, Ashton-Prolla P. Highly prevalent TP53 mutation predisposing to many cancers in the Brazilian population: a case for newborn screening? Lancet Oncol. 2009; 10(9):920–5. https://doi.org/10.1016/S1470-2045(09)70089-0 PMID: 19717094

8. Custodio G, Parise GA, Kiesel Filho N, Komechen H, Sabbaga CC, Rosati R, et al. Impact of neonatal screening and surveillance for the TP53 R337H mutation on early detection of childhood adrenocortical tumors. J Clin Oncol. 2013; 31(20):2619–26. https://doi.org/10.1200/JCO.2012.46.3711 PMID: 23733769

9. Seidinger AL, Caminha IP, Mastellaro MJ, Gabetta CS, Nowill AE, Pinheiro VRP, et al. TP53 p.Arg337His geographic distribution correlates with adrenocortical tumor occurrence. Mol Genet Genomic Med. 2020; 8(9):e1168. https://doi.org/10.1002/mgg3.1168 PMID: 32592449

10. Ribeiro RC, Sandrini F, Figueiredo B, Zambetti GP, Michalkiewicz E, Lafferty AR, et al. An inherited p53 mutation that contributes in a tissue-specific manner to pediatric adrenal cortical carcinoma. Proc Natl Acad Sci U S A. 2001; 98(16):9330–5. https://doi.org/10.1073/pnas.161479898 PMID: 11481490

11. Giacomazzi J, Selistre S, Duarte J, Ribeiro JP, Vieira PJ, de Souza Macedo G, et al. TP53 p.R337H is a conditional cancer-predisposing mutation: further evidence from a homozygous patient. BMC Cancer. 2013; 13:187. https://doi.org/10.1186/1471-2407-13-187 PMID: 23570263

12. Palmero EI, Schuler-Faccini L, Caleffi M, Achatz MI, Olivier M, Martel-Planche G, et al. Detection of R337H, a germline TP53 mutation predisposing to multiple cancers, in asymptomatic women participating in a breast cancer screening program in Southern Brazil. Cancer Lett. 2008; 261(1):21–5. https://doi.org/10.1016/j.canlet.2007.10.044 PMID: 18248785

13. Rana HQ, Clifford J, Hoang L, LaDuca H, Black MH, Li S, et al. Genotype-phenotype associations among panel-based TP53+ subjects. Genet Med. 2019; 21(11):2478–84. https://doi.org/10.1038/s41436-019-0541-y PMID: 31105275

14. DiGiammarino EL, Lee AS, Cadwell C, Zhang W, Bothner B, Ribeiro RC, et al. A novel mechanism of tumorigenesis involving pH-dependent destabilization of a mutant p53 tetramer. Nat Struct Biol. 2002; 9(1):12–6. https://doi.org/10.1038/nsb730 PMID: 11753428

15. Sandrini F, Villani DP, Tucci S, Moreira AC, de Castro M, Elias LL. Inheritance of R337H p53 gene mutation in children with sporadic adrenocortical tumor. Horm Metab Res. 2005; 37(4):231–5. https://doi.org/10.1055/s-2005-861373 PMID: 15952083

16. Varley JM, McGown G, Thorncroft M, James LA, Margison GP, Forster G, et al. Are there low-penetrance TP53 Alleles? evidence from childhood adrenocortical tumors. Am J Hum Genet. 1999; 65(4):995–1006. https://doi.org/10.1086/302575 PMID: 10486318

17. Park JH, Li J, Starost MF, Liu C, Zhuang J, Chen J, et al. Mouse Homolog of the Human TP53 R337H Mutation Reveals Its Role in Tumorigenesis. Cancer Res. 2018; 78(18):5375–83. https://doi.org/10.1158/0008-5472.CAN-18-0016 PMID: 30042151

18. Li FP, Fraumeni JF Jr. Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? Ann Intern Med. 1969; 71(4):747–52. https://doi.org/10.7326/0003-4819-71-4-747 PMID: 5360287

19. Olivier M, Goldgar DE, Sodha N, Ohgaki H, Kleihues P, Hainaut P, et al. Li-Fraumeni and related syndromes: correlation between tumor type, family structure, and TP53 genotype. Cancer Res. 2003; 63(20):6643–50. PMID: 14583457

20. Birch JM, Hartley AL, Tricker KJ, Prosser J, Condie A, Kelsey AM, et al. Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. Cancer Res. 1994; 54(5):1298–304. PMID: 8118819

21. Chompret A, Abel A, Stoppa-Lyonnet D, BrugiÈres L, PagÈs S, Feunteun J, et al. Sensitivity and predictive value of criteria for p53 germline mutation screening. J Med Genet. 2001; 38(1):43–7. https://doi.org/10.1136/jmg.38.1.43 PMID: 11332399

22. Gonzalez KD, Noltner KA, Buzin CH, Gu D, Wen-Fong CY, Nguyen VQ, et al. Beyond Li Fraumeni Syndrome: clinical characteristics of families with p53 germline mutations. J Clin Oncol. 2009; 27(8):1250–6. https://doi.org/10.1200/JCO.2008.16.6959 PMID: 19204208

23. Tinat J, Bougeard G, Baert-Desurmont S, Vasseur S, Martin C, Bouvignies E, et al. 2009 version of the Chompret criteria for Li Fraumeni syndrome. J Clin Oncol. 27. United States 2009. p. e108–9; author reply e10. https://doi.org/10.1200/JCO.2009.22.7967 PMID: 19652052

24. Daly MB, Pilarski R, Berry M, Buys SS, Farmer M, Friedman S, et al. NCCN Guidelines Insights: Genetic/Familial High-Risk Assessment: Breast and Ovarian, Version 3.2019.

25. Achatz MI, Olivier M, Le Calvez F, Martel-Planche G, Lopes A, Rossi BM, et al. The TP53 mutation, R337H, is associated with Li-Fraumeni and Li-Fraumeni-like syndromes in Brazilian families. Cancer Lett. 2007; 245(1–2):96–102. https://doi.org/10.1016/j.canlet.2005.12.039 PMID: 16494995

26. Andrade RC, Dos Santos AC, de Aguirre Neto JC, Nevado J, Lapunzina P, Vargas FR. TP53 and CDKN1A mutation analysis in families with Li-Fraumeni and Li-Fraumeni like syndromes. Fam Cancer. 2017; 16(2):243–8. https://doi.org/10.1007/s10689-016-9935-z PMID: 27714481

27. Leroy B, Ballinger ML, Baran-Marszak F, Bond GL, Braithwaite A, Concin N, et al. Recommended Guidelines for Validation, Quality Control, and Reporting of TP53 Variants in Clinical Practice. Cancer Res. 2017; 77(6):1250–60. https://doi.org/10.1158/0008-5472.CAN-16-2179 PMID: 28254861

28. Villani A, Shore A, Wasserman JD, Stephens D, Kim RH, Druker H, et al. Biochemical and imaging surveillance in germline TP53 mutation carriers with Li-Fraumeni syndrome: 11 year follow-up of a prospective observational study. Lancet Oncol. 2016; 17(9):1295–305. https://doi.org/10.1016/S1470-2045(16)30249-2 PMID: 27501770

29. Fortuno C, Pesaran T, Dolinsky J, Yussuf A, McGoldrick K, Kho PF, et al. p53 major hotspot variants are associated with poorer prognostic features in hereditary cancer patients. Cancer Genet. 2019; 235–236:21–7. https://doi.org/10.1016/j.cancergen.2019.05.002 PMID: 31296311

30. Zerdoumi Y, Lanos R, Raad S, Flaman JM, Bougeard G, Frebourg T, et al. Germline TP53 mutations result into a constitutive defect of p53 DNA binding and transcriptional response to DNA damage. Hum Mol Genet. 2017; 26(14):2591–602. https://doi.org/10.1093/hmg/ddx106 PMID: 28369373

31. Mastellaro MJ, Seidinger AL, Kang G, Abrah$_a$o R, Miranda ECM, Pounds SB, et al. Contribution of the TP53 R337H mutation to the cancer burden in southern Brazil: Insights from the study of 55 families of children with adrenocortical tumors. Cancer. 2017; 123(16):3150–8. https://doi.org/10.1002/cncr.30703 PMID: 28387921

32. Complexo, Southey MC, Park DJ, Nguyen-Dumont T, Campbell I, Thompson E, et al. COMPLEXO: identifying the missing heritability of breast cancer via next generation collaboration. Breast Cancer Res. 15 2013. p. 402. https://doi.org/10.1186/bcr3434 PMID: 23809231

33. Pinto EM, Figueiredo BC, Chen W, Galvao HCR, Formiga MN, Fragoso M, et al. XAF1 as a modifier of p53 function and cancer susceptibility. Sci Adv. 2020; 6(26):eaba3231. https://doi.org/10.1126/sciadv.aba3231 PMID: 32637605

Familial Cancer (2019) 18:451–456
https://doi.org/10.1007/s10689-019-00140-w

SHORT COMMUNICATION

TP53 variants of uncertain significance: increasing challenges in variant interpretation and genetic counseling

Camila Matzenbacher Bittar[1,2] · Igor Araujo Vieira[1,2] · Cristina Silva Sabato[3] · Tiago Finger Andreis[1,2] · Bárbara Alemar[1,2] · Osvaldo Artigalás[4] · Henrique de Campos Reis Galvão[3] · Gabriel S. Macedo[1,2] · Edenir Inez Palmero[3,5] · Patricia Ashton-Prolla[1,2,6]

Published online: 18 July 2019

Abstract
Li-Fraumeni syndrome (LFS) and Li-Fraumeni Like (LFL) are autosomal dominant cancer predisposition syndromes caused by pathogenic germline variants in the *TP53* gene. Recent studies have shown that the incorporation of next-generation sequencing by using multigene panels in clinical practice has resulted in the frequent identification of variants of uncertain significance (VUS). Given that there is no established medical management for VUS carriers, the identification of these variants may cause confusion and anxiety for both patients and practitioners. Herein, we aimed to verify VUS frequency and review VUS classification and interpretation in 1844 patients submitted for comprehensive germline *TP53* testing independent of clinical criteria. Variant characterization was done assessing clinical information whenever available, variant frequency in population databases, pathogenicity predictions using in silico tools and previous functional studies. All variants were classified based on the guidelines proposed by the American College of Medical Genetics and Genomics (2015) and by the Sherloc framework (2017). Of the twelve VUS (0.65%) identified in *TP53*, two were classified as likely pathogenic and two were classified as likely benign after re-evaluation, potentially resulting in significant management modification for the proband and relatives. This report cases highlights the challenges and impact of *TP53* variant interpretation especially when there is no clear LFS/LFL phenotype.

Keywords Li-Fraumeni syndrome · *TP53* gene · Variants of uncertain significance · Genetic counseling

Introduction

Li-Fraumeni (LFS) and Li-Fraumeni Like (LFL) syndromes are autosomal dominant cancer predisposition disorders caused by pathogenic germline variants in the *TP53* tumor suppressor gene. LFS core tumors are premenopausal breast cancer, bone and soft-tissue sarcomas, brain cancer, leukemia and adrenocortical carcinomas.

Camila Matzenbacher Bittar and Igor Araujo Vieira have contributed equally to this work and should be considered co-first authors.

Edenir Inez Palmero and Patricia Ashton-Prolla should be considered joint senior authors.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10689-019-00140-w) contains supplementary material, which is available to authorized users.

✉ Patricia Ashton-Prolla
pprolla@gmail.com

[1] Programa de Pós-Graduação em Genética e Biologia Molecular (PPGBM), Universidade Federal do Rio Grande do Sul (UFRGS), Av. Bento Gonçalves, 9500 - Prédio 43323 M, Porto Alegre, RS 91501-970, Brazil

[2] Laboratório de Medicina Genômica, Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre (HCPA), Rua Ramiro Barcelos, 2350, Porto Alegre, RS 90035-903, Brazil

[3] Molecular Oncology Research Center, Barretos Cancer Hospital, Rua Antenor Duarte Viléla, 1331 - Dr. Paulo Prata, Barretos, SP 14784-400, Brazil

[4] Hospital Moinhos de Vento (HMV), Rua Ramiro Barcelos, 910, Porto Alegre, RS 91790-560, Brazil

[5] Barretos School of Health Sciences, Dr. Paulo Prata – FACISB, Barretos, SP 14784-400, Brazil

[6] Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre (HCPA), Rua Ramiro Barcelos, 2350, Porto Alegre, RS 90035-903, Brazil

Carriers of pathogenic *TP53* germline variants have a variable lifetime risk of developing cancer, and phenotype may vary from fully penetrant LFS to cancer-free over a lifetime. Nevertheless, about 50% of carriers develop at least one malignancy by age 30–40 years, especially those with *TP53* DNA-binding domain variants [1, 2].

Population prevalence studies have estimated that germline *TP53* pathogenic variants occur 1 in 5000 to 1 in 20,000 individuals. However, a recent study by de Andrade et al. [3] evaluated the prevalence of likely pathogenic *TP53* variants in the gnomAD database totaling 138,632 unrelated individuals, and identified 64 distinct pathogenic or likely pathogenic *TP53* variants in 399/138,632 individuals, suggesting that prevalence may reach one carrier in every 3555–5476 individuals. In 2 other studies, prevalence of *TP53* pathogenic variant carriers was proposed to reach 0.2% in Europeans [4, 5]. In addition, a germline *TP53* founder mutation with reduced penetrance, c.1010G > A p.(Arg337His), has been reported in Southern Brazil at a frequency of 1 in approximately 300 newborns [6].

Recent studies have shown that the incorporation of next-generation sequencing (NGS) and multigene panels testing (MGPT) in clinical practice has resulted in the frequent identification of variants of uncertain clinical significance (VUS) [7]. O'Leary et al. [8] described the occurrence of VUS in 28% of patients submitted to MGPT and without an identifiable pathogenic variant. Interpreting these variants and using the information for management have become challenges for clinicians dealing with genetic counseling and cancer surveillance in families with suspected hereditary cancer predisposition syndromes, especially when the phenotype is not typical of LFS/LFL. Given that there is no established medical management for VUS carriers, the identification of these variants causes confusion and anxiety for both patients and practitioners. Thus, establishing the correct variant classification in the clinical setting is of utmost importance. The ACMG–AMP framework assigns a strength level to each evidence criterion and requires various combinations of strong, moderate, and supporting evidence for a confident classification [7]. Population prevalence of the variant, computational data, segregation analyses, cumulative knowledge of clinical outcomes and functional assays in experimental systems have become important information in the difficult task of variant classification [7].

In the present study, we aimed to verify VUS frequency and review VUS classification and interpretation in a large case series submitted to *TP53* germline analysis without a pre-test suspicion of LFS/LFL. We also provide a detailed characterization of 12 probands with germline *TP53* VUS and discuss the caveats of interpretation and genetic counseling in these patients and families. These cases highlight the challenges of germline genetic testing in the clinic and the impact of *TP53* variant interpretation especially when there is no clear LFS/LFL phenotype.

Materials and methods

Genetic germline testing results including analysis of the *TP53* gene of 1844 patients from cancer risk evaluation programs in Southern and Southeastern Brazil were analyzed. All participants underwent pre- and post-test genetic counseling, providing informed written or verbal consent for testing. Testing was performed by Sanger sequencing or NGS (MGPT) due to a diagnosis of cancer, independent of LFS/LFL testing criteria. This study was approved by the Institutional Review Boards from the participating institutions.

Aiming to gather further evidence about pathogenicity, in silico analyses were performed for all variants. For exonic missense variants, we used optimized Align-GVGD, REVEL and BayesDel [9] (further details in Supplementary Tables). An intronic variant was evaluated using Human Splicing Finder, which predicts the effect of the variant on splicing motifs (including exonic enhancers and silencers), as well as by the Berkeley Drosophila Genome Project (BDGP) splicing site predictor and the Alamut Visual software. Minor allele frequencies (MAF) for all variants were obtained from the 1000 Genomes Project, The Exome Aggregation Consortium (ExAC), Genome Aggregation Database (gnomAD, non-cancer dataset), Online Archive of Brazilian Mutations (ABraOM), and Fabulous Ladies Over Seventy (FLOSSIES) population databases. The evaluation of p53 transcriptional activity in a yeast-based functional assay by Kato et al. [10] (described in the IARC *TP53* database, version R19, August 2018) and two recent mammalian functional studies were considered as functional evidence (further details in Supplementary Tables). Segregation analysis, although attempted, was not possible in the families who responded to a specific inquiry about continued testing in the family. Ultimately, the variants were classified according to the guidelines proposed by the American College of Medical Genetics and Genomics (ACMG) [7] and by the Sherloc classification framework [11]. Clinical information and pedigrees containing maternal and paternal cancer histories in at least 3 generations were obtained during pre-test counseling and all probands were classified according to the clinical/genetic testing criteria previously proposed by Bougeard et al. [1, 2].

Results

Twelve unrelated probands carrying germline VUS in *TP53* were identified (seven by MGPT), resulting in a VUS prevalence of 0.65%. For the probands analyzed by NGS panels, the read percentage of the variant alleles was between 47 and

52%, consistent with what is expected for a heterozygous germline variant (data not shown). None of the probands with *TP53* VUS had pathogenic/likely pathogenic variants identified in other genes. Three families fulfilled criteria for hereditary breast and ovarian cancer (HBOC) syndrome and two of them fulfilled the Chompret criteria for LFS. All 3 probands with HBOC criteria were also tested for germline *BRCA1* and *BRCA2* pathogenic variants and were negative. Five probands had two primary tumors, one proband had 3 primary tumors and 6 were diagnosed with breast cancer, the most common tumor type reported in LFS families [1]. The clinical features and molecular characterization of each proband carrying *TP53* VUS are described in Table 1.

To support decision making regarding surveillance, management, risk reduction approaches and genetic counseling, we evaluated family history, as well as available results of functional testing, predicted impact on p53 function, and frequency in population databases for each of the VUS identified and the results are summarized in Table 2. Location of each VUS in the p53 functional domains is depicted in Supplemental Fig. S1. Seven of the variants (58.3%) were present in at least one population database. Of note, 2 variants (16.7%) were not described in ClinVar (Table 2, Supplemental Table S1). Moreover, eleven of the twelve VUS (91.7%) are missense, of which five (41.6%) were predicted to be "benign" according to Align-GVGD, REVEL and BayesDel, while 2 (16.7%) variants were predicted to be "pathogenic" by combined results from these computational tools (Supplemental Table S2). Five VUS (41.7%) were shown to affect p53 function in yeast-based functional studies (non-functional and partially functional *TP53* germline variants), while 3 (25%) were defined as "loss-of-function" alterations when analyzing with two recent mammalian functional assays (Supplemental Table S3). Human Splicing Finder, BDGP and ALAMUT analyses indicated that the intronic variant c.559 + 4A > G (proband number 7) is predicted to affect splicing by altering a wild-type donor site (Table 2, Supplemental Fig. S2). Although this variant was not reported in the population databases queried, there are no other criteria indicating its pathogenicity and thus there is not enough evidence to re-classify it. Taken together and considering that the mammalian functional studies performed were robust, two variants (c.417G > T and c.589G > A) would be classified as likely pathogenic according to the guidelines from ACMG and Sherloc. Pedigrees of the two probands harboring these now classified

Table 1 Clinical and molecular characterization of probands carrying VUS in the *TP53* gene

Proband ID	Proband's age at diagnosis (years), 1st cancer	Proband's type of cancer	Age at diagnosis, other tumors (years)	*TP53* mutation testing criteria	HBOC NCCN testing criteria? (yes/no)	Molecular analysis (method)	chr17 position on assembly GRCh37 (dbSNP rs ID)	*TP53* variant HGVS c.	*TP53* variant HGVS p.
1	44	Gastric	NA	None	No	Sanger	7579882(rs201382018)	c.31G > C	p.(Glu11Gln)
2	70	Bilateral breast	Thyroid (71)	None	No	Sanger	7579558(rs754332870)	c.129G > C	p.(Leu43Phe)
3	38	Breast	NA	None	Yes	NGS	7579367(rs587782447)	c.320A > G	p.(Tyr107Cys)
4	33	Breast	Thyroid (34)	Breast cancer < 36 years	Yes	NGS	7578513	c.417G > T	p.(Lys139Asn)
5	40	Renal	NA	Birch (family)	No	NGS	7578470(rs137852789)	c.460G > A	p.(Gly154Ser)
6	60	Sarcoma	NA	None	No	Sanger	7578408	c.522G > C	p.(Arg174Ser)
7	3	LLA	Breast (25)	Chompret	Yes	NGS	7578367	c.559 + 4A > G	(Intronic)
8	42	Renal	Lung (42)	Chompret	No	NGS	7578260(rs786204041)	c.589G > A	p.(Val197Met)
9	51	Lung	NA	None	No	Sanger	7578207(rs587781386)	c.642T > G	p.(His214Gln)
10	33	Melanoma	Thyroid (47) and Breast (49)	None	No	NGS	7577151(rs72661119)	c.787A > C	p.(Asn263His)
11	36	Rectal	Pancreas (46)	None	No	Sanger	7574030(rs769934890)	c.997C > T	p.(Arg333Cys)
12	46	Breast	NA	None	No	NGS	7572989(rs587781858)	c.1120G > C	p.(Gly374Arg)

LLA lymphoblastic acute leukemia, *NA* not applicable

Table 2 Assessment of pathogenicity of each variant of uncertain significance identified in our case series

Proband ID (HGVS name)	Frequency in population databases					Functional studies[b]	ClinVar[c]	Computational prediction as suggested by Fortuno et al. (2018)[d]	ACMG classification (code assignment)[f, g]	Sherloc classification[h]
	FLOSSIES	ExAC	AbraOM	1000 Genomes	gnomAD[a]					
1 (c.31G > C)	Not present	4/118912 (0.00003364)	1/1218 (0.000821)	Not present	2/268046 (0.0000075)	Benign (BS3)	Conflicting[1]	Benign (BP4)	Benign (BS1, BS3, BP4)	Likely benign
2 (c.129G > C)	Not present	1/121066 (0.00000826)	Not present	Not present	1/236844 (0.0000042)	Benign (BS3)	Not Present	Benign (BP4)	Likely benign (BS3, BP4)	VUS
3 (c.320A > G)	Not present	Not present	Not present	Not present	Not Present	Benign (BS3)	VUS[2]	Pathogenic (PP3)	VUS (PM2, PP3, BS3)	VUS
4 (c.417G > T)	Not present	Not present	Not present	Not present	Not Present	Pathogenic (PS3)	VUS[1]	No evidence	Likely pathogenic (PS3, PM2, PP4)	Likely pathogenic
5 (c.460G > A)	1/9884 (0.0001012)	1/121164 (0.00000825)	1/1218 (0.000821)	Not present	6/268114 (0.0000022)	Pathogenic (PS3)	VUS[2]	No evidence	VUS (PS3, PP4, BS1)	VUS
6 (c.522G > C)	Not present	Not present	Not present	Not present	Not Present	Benign (BS3)	Not present	No evidence	VUS (PM2, BS3)	VUS
7 (c.559 + 4A > G)	Not present	Not present	Not present	Not present	Not present	Not performed	VUS[1]	NA[e]	VUS (PM2, PP3, PP4)	VUS
8 (c.589G > A)	Not PRESENT	Not present	Not present	Not present	Not present	Pathogenic (PS3)	VUS[1]	No evidence	Likely pathogenic (PS3, PM2, PP4)	Likely pathogenic
9 (c.642T > G)	Not present	2/120946 (0.0001654)	Not present	Not present	1/236930 (0.0000042)	Benign (BS3)	Conflicting[1]	Benign (BP4)	Likely benign (BS3, BP4)	VUS
10 (c.787A > C)	Not present	10/92284 (0.0001083)	Not present	1/5008 (0.0002)	27/232496 (0.00012)	Benign (BS3)	VUS[1]	Benign (BP4)	Benign (BS1, BS3, BP4)	Likely benign
11 (c.997C > T)	Not present	1/115640 (0.000008648)	Not present	Not present	6/236154 (0.000025)	Benign (BS3)	VUS[2]	Pathogenic (PP3)	VUS (PP3, BS3)	VUS
12 (c.1120G > C)	1/9884 (0.0001012)	1/121396 (0.00000823)	Not present	Not present	1/31378 (0.000032)	Benign (BS3)	Conflicting[1]	Benign (BP4)	Likely benign (BS3, BP4)	VUS

VUS variants of uncertain significance, *NA* not applicable, *NP* not present

[a]gnomAD frequency was determined using non-cancer dataset

[b]Results from mammalian functional assays were prioritized in the decision tree for more strong functional evidence (further details showed in Supplemental Table S3)

[c]ClinVar review status, 1 = criteria provided by a single submitter or conflicting interpretations; 2 = criteria provided by multiple submitters, no conflicts; 3 = reviewed by expert panel; 4 = practice guideline. For more details, please refer to: https://www.ncbi.nlm.nih.gov/clinvar/docs/review_status/. Complete data reported by ClinVar for these variants are shown in Supplemental Table S1

[d]Fortuno and colleagues (2018) supported use of Align-GVGD and BayesDel prediction for different strength of evidence levels in ACMG-AMP rules (further details in Supplemental Table S2 and PMID: 29775997)

[e]For intronic variant c.559 + 4A > G, Human Splicing Finder, BDGP splicing site predictor and Alamut computational/software tools indicated that it potentially affects splicing by altering a wild-type donor site. Output analysis from Alamut software is showed in Supplemental Fig. S2

[f]These in silico results suggested PP3 code assignment for ACMG classification of this specific variant

[g]For the ACMG classification, there is no evidence of de novo alteration for those variants in the IARC database

[h]For variant classification using Sherloc framework, yeast functional assay was considered "weak evidence", while mammalian functional studies were considered "strong evidence"

likely pathogenic germline variants are shown in Supplemental Fig. S3. Of note, several variants showed discordant classifications when either ACMG or Sherloc were used. For instance, three variants (c.129G > C, c.1120G > C, and c.642T > G) were classified as likely benign by ACMG, while the same alterations were defined as VUS according to Sherloc (Table 2).

Discussion

Several studies have demonstrated evidence for significantly improved outcomes in LFS patients submitted to an enhanced surveillance protocol which includes frequent physical examination plus periodic biochemical and imaging monitoring [12]. In this context, identification of a variant classified as VUS in *TP53*, in addition to the potential anxiety caused in carriers and health care providers, must be reviewed by a more robust classification system to ensure that management and counseling offered to the patient are appropriate to molecular status. Misinterpretation of a likely pathogenic variant as VUS or as likely benign has the potential to leave a patient and his/her family unassisted, resulting in a missed opportunity of cancer risk reduction or early detection. On the other hand, misinterpretation of a likely benign variant or VUS as likely pathogenic, will have an impact on the patient and relatives, potentially resulting in frequent and unnecessary surveillance and risk-reducing procedures, not to mention psychological consequences.

In this report, we aimed at defining frequency of *TP53* VUS in a large cohort of probands investigating hereditary cancer predisposition and at reviewing VUS classification. Although overall VUS frequency was low, extensive variant reclassification efforts showed interesting results. Among 12 *TP53* VUS identified, two (c.417G > T and c.589G > A) were reclassified as likely pathogenic, supporting the decision to recommend family testing and inclusion of all carriers in specific, high risk surveillance protocols. Specially for the proband carrying c.417G < T, an isolated case of breast cancer diagnosed under age 36 years with no family history of cancer, misclassification would have resulted in a lack of surveillance for other tumors, family testing and inappropriate genetic counseling. In addition, two variants (c.787A > C and c.31G > C) were classified as likely benign by Sherloc, also contributing to optimize clinical management. Importantly, five variants were classified differently using either the ACMG or Sherloc classification systems. This is likely due to the fact that Sherloc, in addition to the strong framework of the 33 rules established by the ACMG, introduced 108 detailed refinements, being recently described as a more robust approach to variant classification [11]. Of the five variants with discordant classifications, only 3 would result in distinct management guidelines, in which by ACMG a likely benign classification would result in no further investigation, while the Sherloc VUS classification would result in a recommendation of periodic reassessment to verify if the VUS reclassifies to another category. These findings underscore the need of carefully considering which pathogenicity algorithm is used in variant classification.

Another important aspect of variant classification, is the risk of overreacting when a VUS is identified. For instance, proband carrier of *TP53* c.559 + 4A > G variant, a particularly challenging case where the phenotype strongly suggests the diagnosis (Chompret criteria for LFS), that could lead to a misinterpretation of such variant as likely pathogenic by a clinician not familiarized with in depth variant classification systems. In this case, after gathering all available evidence the variant still classifies as VUS and causality cannot be attributed.

The cases reported here illustrate some of the complexities involved in variant classification and clinical management of VUS carriers and underscore the importance of reviewing the use of established clinical criteria for germline *TP53* testing. All probands whose variants were analyzed had the opportunity to be informed about variant reanalysis and its results during a genetic counseling session with a medical geneticist. The review of variant classification is often beyond the training of most practicing clinicians, and referral to a trained genetics professional should be strongly encouraged. Lastly, it is important to highlight that the VUS status should be revised often, since the classification may change once additional evidence becomes available.

Identification of a VUS in *TP53* is a complex counseling issue and highlights the need for more research aiming to improve decision making in patient care. An important contribution to these efforts is the sharing of sequencing results in an open database. As a result of these collective attempts, the incidence of VUS is expected to decline in the next years. Despite the small overall frequency of VUS in *TP53*, it is essential that a periodic review of variant significance is done, and that any new definitions are communicated to the patient. Finally, we encourage further studies of penetrance modifiers and of individuals with *TP53* VUS to enhance our knowledge of variant curation for clinical purposes.

Acknowledgements This study was funded by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Grant # 478430/2012-4), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) (Grant # 16/2551-0000486-2), and Fundo de Incentivo à Pesquisa do Hospital de Clínicas de Porto Alegre (FIPE-HCPA) to P.A-P. Barretos Cancer Hospital was partially supported by FINEP-CT-INFRA (02/2010). P.A.-P and E.I.P receive a CNPq scholarship.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Bougeard G, Sesboüé R, Baert-Desurmont S et al (2008) Molecular basis of the Li-Fraumeni syndrome: an update from the French LFS families. J Med Genet 45:535–538
2. Bougeard G, Renaux-Petel M, Flaman JM et al (2015) Revisiting Li-fraumeni syndrome from TP53 mutation carriers. J Clin Oncol 33(21):23452352
3. de Andrade KC, Frone MN, Wegman-Ostrosky T et al (2019) Variable population prevalence estimates of germline *TP53* variants: a gnomAD-based analysis. Hum Mutat 40:97–105
4. Buys SS, Sandbach JF, Gammon A et al (2017) A study of over 35,000 women with breast cancer tested with a 25-gene panel of hereditary cancer genes. Cancer 123:1721–1730
5. Fortuno C, James PA, Spurdle AB (2018) Current review of *TP53* pathogenic germline variants in breast cancer patients outside Li-Fraumeni syndrome. Hum Mutat 39:1764–1773
6. Achatz MI, Zambetti GP (2016) The Inherited p53 Mutation in the Brazilian Population. Cold Spring Harb Perspect Med 6:a026195
7. Richards S, Aziz N, Bale S et al (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 17:405–424
8. O'Leary E, Iacoboni D, Holle J et al (2017) Expanded gene panel use for women with breast cancer: identification and intervention beyond breast cancer risk. Ann Surg Oncol 24:3060–3066
9. Fortuno C, James PA, Young EL et al (2018) Improved, ACMG-compliant, in silico prediction of pathogenicity for missense substitutions encoded by TP53 variants. Hum Mutat 39:1061–1069
10. Kato S, Han S-Y, Liu W, Otsuka K, Shibata H, Kanamaru R, Ishioka C (2003) Understanding the function–structure and function–mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. Proc Natl Acad Sci 100(14):8424–8429
11. Nykamp K, Anderson M, Powers M et al (2017) Sherloc: a comprehensive refinement of the ACMG-AMP variant classification criteria. Genet Med 19:1105–1117
12. Villani A, Shore A, Wasserman JD et al (2016) Biochemical and imaging surveillance in germline TP53 mutation carriers with Li-Fraumeni syndrome: 11 year follow-up of a prospective observational study. Lancet Oncol 17:1295–1305

cancers

MDPI

Article

Systems Biology Approaches Reveal Potential Phenotype-Modifier Genes in Neurofibromatosis Type 1

Thayne Woycinck Kowalski [1,2,3], Larissa Brussa Reis [1,2], Tiago Finger Andreis [1,2], Patricia Ashton-Prolla [1,2,4] and Clévia Rosset [1,5,*]

1 Laboratório de Medicina Genômica, Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre, Porto Alegre 90035-007, Rio Grande do Sul, Brazil; thaynewk@gmail.com (T.W.K.); brussareis@gmail.com (L.B.R.); tiago.f.andreis@gmail.com (T.F.A.); pprolla@hcpa.edu.br (P.A.-P.)
2 Programa de Pós-Graduação em Genética e Biologia Molecular, PPGBM, Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre 91501-970, Rio Grande do Sul, Brazil
3 CESUCA - Faculdade Inedi, Cachoeirinha 94935-630, Rio Grande do Sul, Brazil
4 Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre, Porto Alegre 90035-007, Rio Grande do Sul, Brazil
5 Unidade de Pesquisa Laboratorial, Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre, Porto Alegre 90035-007, Rio Grande do Sul, Brazil
* Correspondence: crosset@hcpa.edu.br; Tel.: +55-51-3359-7661

Received: 30 June 2020; Accepted: 20 July 2020; Published: 26 August 2020

Abstract: Neurofibromatosis type (NF1) is a syndrome characterized by varied symptoms, ranging from mild to more aggressive phenotypes. The variation is not explained only by genetic and epigenetic changes in the *NF1* gene and the concept of phenotype-modifier genes in extensively discussed in an attempt to explain this variability. Many datasets and tools are already available to explore the relationship between genetic variation and disease, including systems biology and expression data. To suggest potential NF1 modifier genes, we selected proteins related to NF1 phenotype and *NF1* gene ontologies. Protein–protein interaction (PPI) networks were assembled, and network statistics were obtained by using forward and reverse genetics strategies. We also evaluated the heterogeneous networks comprising the phenotype ontologies selected, gene expression data, and the PPI network. Finally, the hypothesized phenotype-modifier genes were verified by a random-walk mathematical model. The network statistics analyses combined with the forward and reverse genetics strategies, and the assembly of heterogeneous networks, resulted in ten potential phenotype-modifier genes: *AKT1, BRAF, EGFR, LIMK1, PAK1, PTEN, RAF1, SDC2, SMARCA4*, and *VCP*. Mathematical models using the random-walk approach suggested *SDC2* and *VCP* as the main candidate genes for phenotype-modifiers.

Keywords: neurofibromatosis type 1; phenotype-modifier genes; systems biology

1. Introduction

Neurofibromatosis type 1 (NF1) is a disease with a worldwide birth incidence of 1 in 2500 and a prevalence of at least 1 in 4000 [1]. The main clinical features are café-au-lait spots, axillary and inguinal freckling, cutaneous and subcutaneous neurofibromas, and Lisch nodules, occurring in almost every NF1 patient. Other less -ommon characteristics are scoliosis, macrocephaly, learning disabilities, plexiform neurofibromas, and multiple other benign and malignant tumors [2,3]. Inter-familial and intra-familial variability in NF1 is extensive: cutaneous neurofibromas may vary in number from dozens to thousands; about 30% to 50% of patients are affected by large plexiform neurofibromas; only about

10% of them develop malignant peripheral nerve sheath tumors (MPNSTs), an aggressive sarcoma and one of the most critical symptoms [2,4,5]. Other tumors outside the central nervous system occur in different frequencies between NF1 patients: low grade pilocytic astrocytomas, pheochromocytoma, gastrointestinal stromal tumor, thyroid tumors, ovary and lung tumors, breast cancer, juvenile myelomonocytic leukemia, myelodysplastic syndrome, osteosarcoma, and rhabdomyosarcoma [3–5].

NF1 is caused by dominant loss-of-function mutations in the tumor suppressor gene *NF1*, which encodes neurofibromin, an interactor of Ras GTPase proteins [6]. Although NF1 is a monogenic disorder of dominant character, only a few associations between a specific *NF1* variant and the disease phenotype have been reported to date. Four genotype–phenotype correlations are well described in the literature: NF1 patients harboring microdeletions have been reported to have an increased risk of malignant peripheral nerve sheath tumors, lower average intelligence, connective tissue dysplasia, skeletal malformations, and dysmorphic facial features [7–9]; the 3-bp in-frame deletion c.2970_2972delAA was previously associated with absence of neurofibromas [10]; the missense variant p.Arg1809Cys was associated with developmental delay and/or learning disabilities, pulmonic stenosis, and Noonan-like features, but no external plexiform neurofibromas [11]; and missense mutations affecting *NF1* codons 844–848 were associated with a more severe clinical presentation [12].

Apart from the aforementioned correlations, NF1patients with the same mutation may develop severe symptoms or a mild clinical expression [13–15]. Modifier genes, environmental factors, epigenetic factors, or a combination of them may be responsible for the remaining variability [16,17]. Modifier genes include any genes, protein-coding sequences, microRNA, and long noncoding RNA that influence one or various features of the NF1 phenotype. Primarily, modifier genes were found to be associated with phenotype variation in NF1 in large family studies, and posteriorly, NF1 animal models and knock-in and knockdown strategies have reinforced these assumptions [18]. Several strategies to discover and understand modifiers genes have been developed to help to explain the NF1 variability and were reviewed recently [19–25]. These strategies have identified important candidate modifier genes, and some hypotheses and associations have been established so far [16,20,22]. However, many NF1 characteristics and variability remain unexplained.

Systems biology is an integrative field that combines molecular biology experiments and computational analysis. Its aim is to understand the simplest interactions in the complexity of an organism by the evaluation of interaction networks [26,27]. By integrating genomics, proteomics, and phenotype information, it is possible to evaluate how each of these elements acts as disturber-mechanism in a specific network. This strategy consists of a very effective and economical approach to explore the disease, and might even be applied if there is little information obtained from differential gene expression studies. Hence, through systems biology tools it is possible to perform genomics research by introducing a forward or reverse strategy. The former is a strategy used by evaluating the candidate genes and how they could explain the phenotype, whilst the second strategy starts from the outcome (here NF1), and evaluates which the genes and mechanisms could be connected to it [28]. The use of a deep phenotype characterization is a good approach in conditions with heterogeneous phenotypes, when combined with next-generation sequencing data [29]. For a better comprehension of these molecular mechanisms, ontologies databases have been widely used for a correct assortment of the gene function and in the phenome characterization [30].

In this context, by using this approach, the present study searches for novel candidate NF1 modifier genes. Considering that the modifier genes could play a role in the NF1 signaling pathway or other related and unrelated pathways, in silico analyses were performed through systems biology tools involving the *NF1* gene, its protein–protein interaction network, and its related genes or phenotype ontologies. Network statistics suggested ten candidate genes and mathematical models highlighted the roles of two of them as NF1 phenotype modifiers.

2. Results

A scheme presenting the main steps of the present study is available in Figure 1. To better comprehend the parameters used in each analysis, please see the Methods section.

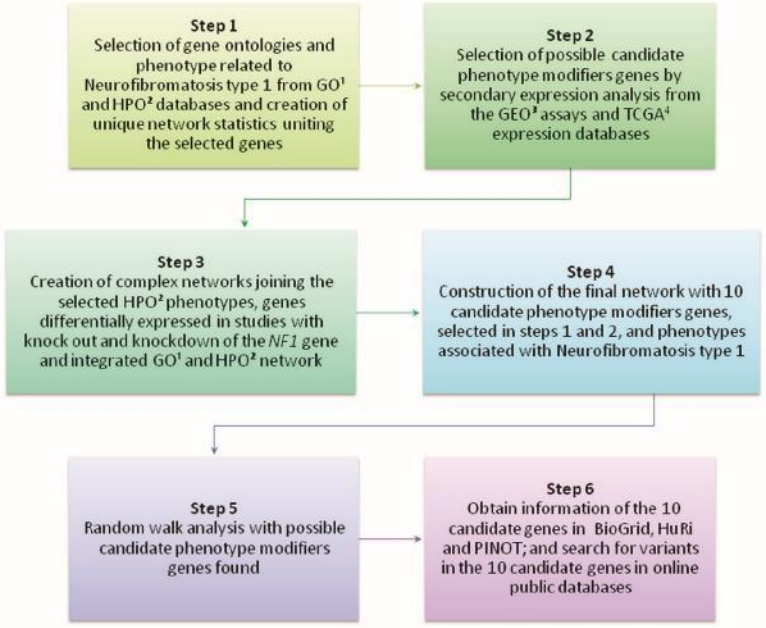


Figure 1. Main steps of the present study. The scheme shows the main steps in chronological order to identity potential NF1 phenotype-modifier genes. Gene Ontology (GO)[1]; Human Phenotype Ontology (HPO)[2]; Gene Expression Omnibus (GEO)[3]; The Cancer Genomic Atlas (TCGA)[4].

2.1. Gene and Phenotype Ontologies Analyses

Gene Ontology (GO) describes a biological domain considering three aspects: molecular function, cellular component, and biological process. The Human Phenotype Ontology (HPO) provides a standardized vocabulary of phenotypic abnormalities encountered in human diseases. NF1 GO biological processes and NF1 HPO were analyzed by two coworkers individually. The chosen ontologies are listed in Table S1.

HPO selection provided 1697 genes related to NF1 phenotype (OMIM 162200), whilst GO filter yielded 1449 genes included in the same ontologies previously selected for the *NF1* gene. When comparing both strategies, it was observed that HPO and GO analysis had 265 genes in common (Figure 2a). To assemble a network of the ontologies' selected genes, we used the STRING tool, observing protein–protein interactions (PPI) that were previously described by experimental assays. The separate networks generated for GO and HPO analyses are represented in Figure S1 and Figure S2, respectively. A combined network, comprising NF1 direct interactions (first neighbors) for both GO and HPO strategies is available in Figure 2b.

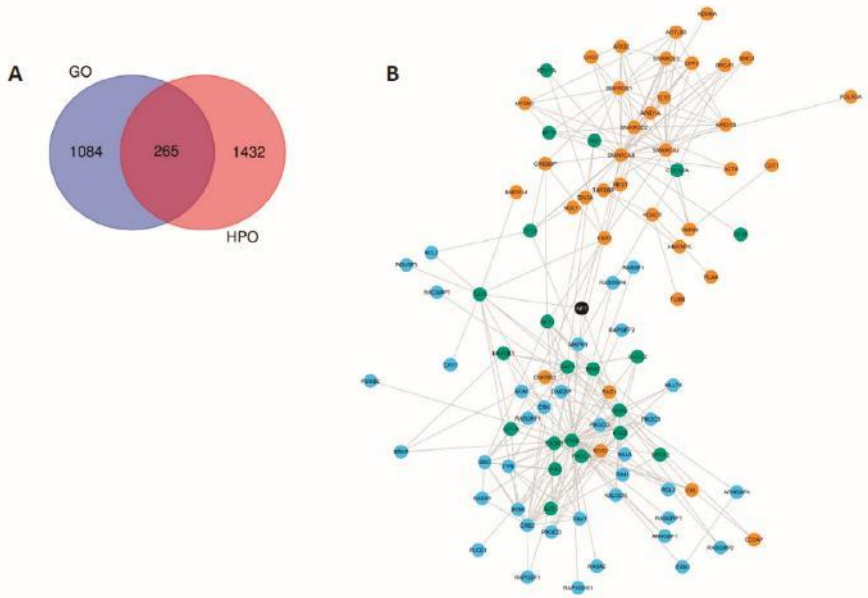


Figure 2. NF1 Gene Ontology (GO) and NF1 Human Phenotype Ontology (HPO) results. (**A**) Venn diagram showing in red genes related with neurofibromatosis type 1 exclusively found by the HPO project in the selected ontologies for NF1; genes exclusively found by the GO consortium using the ontologies selected for NF1 are shown in blue; and in purple genes shared by both GO and HPO analysis. (**B**) A combined network using the STRING tool using the 1697 genes selected exclusively in HPO (orange nodes); 1449 genes selected exclusively in GO (blue nodes); and 265 genes observed in both HPO and GO (green nodes). The network shows only direct protein–protein interactions with NF1 (first neighbors).

2.2. Network Statistics

To verify the nodes with relevant roles in the information flow from the network assembled in the previous section (Figure 2), systems biology network statistics were applied using Cytoscape v.3.7.2 software. Two main parameters were observed: (I) betweenness centrality, a measure based on the communication paths, meaning the nodes with high betweenness centrality could be important in the control of the information flow; and (II) the closeness centrality measure, which is based on the fastness of this information flow (from the central node to the others) [31]. The resulting network had 1561 nodes, making it difficult to visualize the main nodes. A simplified version, representing only the first neighbors, can be assessed in Figure 3.

According to this strategy, AKT1 presents the highest levels of betweenness and closeness centrality. However, despite NF1 being a highly connected protein in the network evaluated, it presented low levels of betweenness and closeness centrality, as can be observed by the node size (small) and color (light yellow, compared to the dark orange elements). Hence, we aimed to evaluate HPO and GO networks separately using the same approach to minimize the possibility of overlooking potential phenotype-modifier genes, as described in the following section.

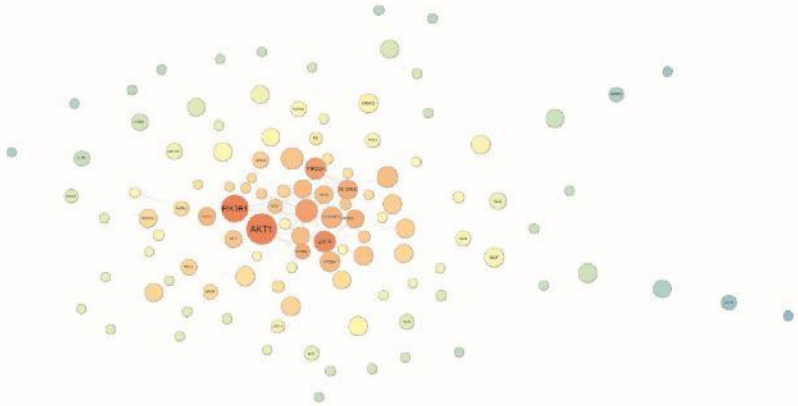


Figure 3. Betweenness centrality and closeness analysis of the STRING network previously generated in Figure 2b using GO and HPO. Large nodes have a more central role in communication among other nodes (hub/hub-like nodes), i.e., more connections. The darker orange the nodes are, the faster information flows towards the central node; i.e., they have the potential to impact the whole network even when having few connections (bottleneck nodes). Thus, nodes can be visualized in four categories: (1) large/dark-orange nodes = hub/bottleneck nodes; (2) large/blue nodes = hub/non-bottleneck nodes; (3) small/dark-orange nodes = non-hub/bottleneck nodes; and (4) small/blue nodes = non-hub/non-bottleneck hubs. NF1 is represented by the yellow node on the right side of AKT1.

2.3. Forward and Reverse Genetics Strategies

As mentioned before, GO and HPO databases are related, respectively, to the gene function and phenotype association. Hence, the observations of their independent networks, previously represented in Figures S1 and S2, were based on the forward and reverse genetics concepts.

When evaluating betweenness and closeness centrality by the forward genetics strategy (GO network), six genes were selected: *AKT1, RAF1, LIMK1, BRAF, EGFR*, and *PTEN*. In the other analysis, the reverse genetics approach (evaluating the HPO network) provided six genes as well: *PAK1, VCP, AKT1, SMARCA4, RAF1*, and *PTEN*. Together, the strategies provided nine candidate genes for neurofibromatosis phenotype modifiers. Besides, the network communities were evaluated, and *AKT1* was identified as the network main hub, whilst the *RAF1* gene had the highest score for authority. The authority score estimates the importance of the node itself, and the hub score measures this importance based on the other nodes which are linked to the main hub. Despite the network statistics having provided these candidates, we wanted to observe whether or not their expression was altered in the absence of *NF1*. Therefore, the next step was designed to conduct gene expression analyses to evaluate this hypothesis.

2.4. Differential Gene Expression Networks

To comprehend the differential gene expressions (DGEs) of the candidate phenotype-modifier genes, we performed secondary expression analysis on the data available in two public repositories: The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO).

Using the GEO database, *NF1* knockdown and knockout assays were selected: GSE14038 and GSE115406. The log fold-change (logFC) and false discovery rate (FDR) values for the ten potential modifier genes in each dataset are presented in Table S2.

In TCGA, we selected seven different types of tumors for which samples with nonsense mutations in *NF1* were available. We evaluated tumors that presented *NF1* nonsense mutations against tumors with wildtype *NF1*. The logFC and adjusted *p*-values, after the FDR correction, for the seven tumors evaluated and ten candidate genes are available in Table S3. Despite the somatic origins of these tumors,

we believe this information is valuable in order to check how NF1-loss could affect the global gene expression in a tissue/site-specific way, and check for signatures more related to a certain phenotype.

For both TCGA and GEO assays, few genes were evidenced to be significantly differentially expressed, demonstrating that the expression of all the candidates, and for the *NF1* gene, is strictly regulated. We did not identify a variable expression profile between the tumors evaluated, and knockout assays. Therefore, we performed other systems biology analysis with the PhenomeScape application v.1.0.4 from Cytoscape software, assembling a complex network (Figure 1, step 3). For that, we used as input (I) the ontologies selected from HPO database (Figure 1, step 1); (II) the network generated in STRING tool, as also mentioned in step 1 (Figure 1); and (III) expression data from studies selected from GEO database (Figure 1, step 2). The resulting upregulated genes (overexpressed) were presented in red and the downregulated (lower expression) in green. These networks are available in Figures S3–S7. *NF1* is downregulated (lower expression) in the knockdown and knockout studies, and upregulated (overexpressed) in the evaluation of malignant tumors when compared to benign neurofibromas. *NF1* is absent from the network when its expression is not significantly altered in the expression dataset.

Besides the genes previously selected in the forward and reverse genetics analysis, when evaluating the complex networks, we observed that the *SDC2* gene also had its expression altered when *NF1* was affected. Furthermore, *SDC2* is the first neighbor of *NF1*, which means both genes share a direct protein–protein interaction.

2.5. Systems Biology Approaches Reveal 10 NF1 Phenotype-Modifier Candidate Genes

Table 1 shows the final list of the 10 genes selected as potential phenotype-modifiers in this study and summarizes the strategies by which they were found. We then generated a complex network comprising all the candidate phenotype-modifier genes selected so far, and the NF1 phenotypes they are related to (Figure 4); the phenotypes were provided by the PhenomeScape tool, according to the data available in HPO database. Finally, we used a mathematical model to evaluate whether one of those genes could be a stronger candidate as a phenotype-modifier than the others, which is described in the following section.

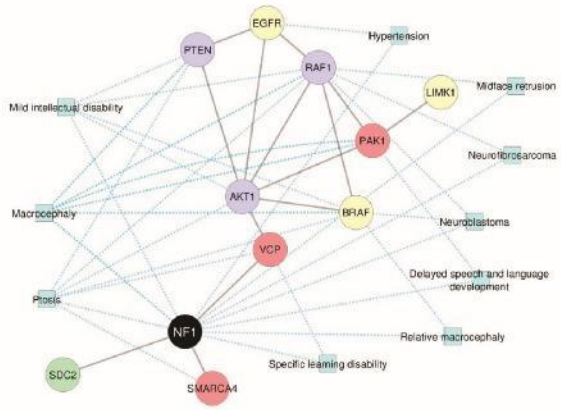


Figure 4. A complex network evaluation comprising the 10 candidate phenotype-modifier genes and the NF1 phenotypes they are related to. The ten candidate NF1 phenotype-modifier genes suggested by our analysis are represented with the NF1 phenotypes they are related. Yellow nodes: genes selected with the forward genetics strategy; red nodes: genes selected with the reverse genetics strategy; purple nodes: genes selected by forward and reverse genetics strategies; green node: gene selected by evaluating the differential gene expression in the complex network. Blue squares: phenotypes provided by the PhenomeScape tool, according to the associations presented in the HPO database.

Table 1. Potential phenotype-modifier genes selected in this study. Characterization of the 10 potential phenotype-modifier genes and the approaches used for their selection.

Gene (OMIM)	Aliases	Cytogenetic Location	Summary	PINOT [1]	STRING [2]	BioGrid	Human Interactome	HPO	GO	Phenome Scape	Direct Strategy	Forward Strategy
AKT1 (*164730)	*AKT, CWS6, PKB, PKB-ALPHA, PRKBA, RAC, RAC-ALPHA*	14q32.33	Serine/threonine kinase - development of the human nervous system; mediator of growth factor-induced neuronal survival; can inactivate components of apoptosis			X		X	X		X	X
BRAF (164757)	*NS7; B-raf; BRAF1; RAFB1; B-RAF1*	7q34	Serine/threonine kinase - role in regulating the MAP kinase/ERK signaling pathway						X		X	
EGFR (131550)	*ERBB, ERBB1, HER1, NISBD2, PIG61, mENA*	7p11.2	Cell surface protein - acts as a receptor for members of the epidermal growth factor family which induces cell proliferation	X		X			X		X	
LIMK1 (601329)	*LIMK; LIMK-1*	7q11.23	Serine/threonine kinase - regulates actin polymerization; it is ubiquitously expressed during development; associated with cytoskeletal structure						X		X	
PAK1 (602590)	*DKFZp686B20; p65-PAK; PAKalpha; alpha-PAK*	11q13.5-q14.1	Serine/threonine kinase - cytoskeleton reorganization and nuclear signaling, regulates cell motility and morphology; essential for the RAS-induced malignant transformation					X				X
PTEN (601728)	*BZS; DEC; CWS1; GLM2; MHAM; TEP1; MMAC1; PTEN1; PTENbeta*	10q23.31	Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase that functions as a tumor suppressor	X		X		X	X		X	X
RAF1 (164760)	*NS5; CRAF; Raf-1; c-Raf; CMD1NN*	3p25.2	MAP3 kinase - involved in the cell division cycle, apoptosis, cell differentiation and cell migration			X		X	X		X	X
SDC3 (142461)	*HSPG; CD362; HSPG3; SYND2 BRG1; CSS1; SNF2; SMI2; MRD16; RTPS2*	8q22.1	Syndecan proteoglycan protein - mediates cell binding, cell signaling, and cytoskeletal organization	X	X	X	X			X		
SMARCA4 (603254)	*BAT130; SNF2L4; SNF2LA; hSNF2b; BAF190A; SNF2-beta*	19p13.2	Part of the large ATP-dependent chromatin remodeling complex required for transcriptional activation of genes normally repressed by chromatin	X	X			X				X
VCP (601023)	*p97; TERA; CDC48*	9p13.3	Plays a role in protein degradation, intracellular membrane fusion, DNA repair and replication, regulation of the cell cycle, and activation of the NF-kappa B pathway	X	X	X	X	X				X

[1] PINOT = Protein Interaction Network Online Tool; [2] STRING = Search Tool for Recurring Instances of Neighboring Genes.

2.6. Random Walk Analysis

A random walk is a mathematical model known as a random process. It is based on the idea that a gene (node) is an imaginary particle that performs a succession of random steps (interactions) in a network [32]. Our aim was to evaluate whether these random steps could lead the gene to the phenotype, which was set as neurofibromatosis type 1 (OMIM 162200). For this goal, we performed the random walk analysis with the *RandomWalkRestartMH* package in R v.3.6.2.

According to this mathematical model, the *NF1* gene only had to take "one step" (one interaction) to reach the phenotype OMIM 162200. The genes *SDC2* and *VCP* had to take only two steps (two interactions) (Figure 5), whilst the other eight candidates needed more interactions to cause the syndrome (Figure S8). Genes *LIMK1* and *PAK1* also needed a higher number of potential interactions, and hence more steps, than the others.

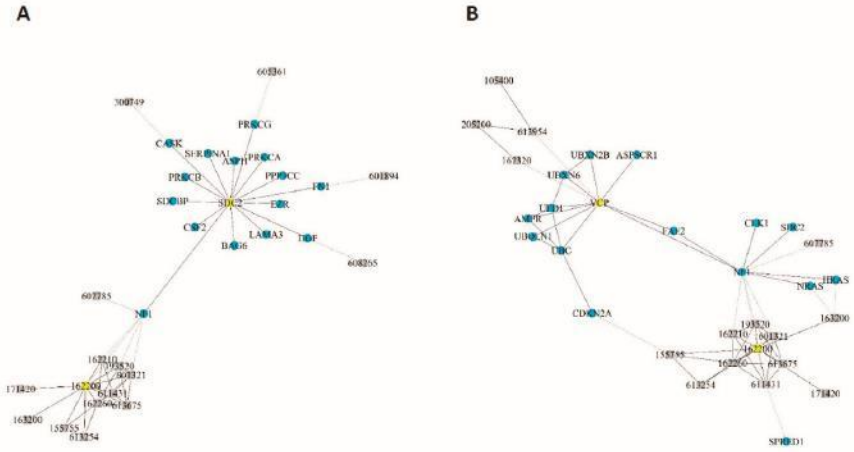


Figure 5. Random walk analysis. The minimum steps (interactions) between the selected genes (nodes) and the neurofibromatosis 1 phenotype were calculated. Two analyses are shown: (**A**) one for *SDC2*; and (**B**) one for *VCP*.

With this analysis, we confirmed all the candidates as potential phenotype-modifier genes. However, the results using this model pointed to *SDC2* and *VCP* as being more directly connected to the NF1 phenotype.

2.7. Literature Review and Genomic Databases Evaluation

To check for genetic variants already described in our 10 candidate genes, we looked at two databases: The Genome Aggregation Database (gnomAD v. 2.1.1), which spans 125,748 exomes and 15,708 genomes from unrelated individuals; and ClinVar, which aggregates information about genomic variation and its relationship to human health. For gnomAD, we found a total of 11,211 variants (Table 2). *SMARCA4* and *EGFR* have the highest numbers of variants (2575 and 2000, respectively); *SDC2* and *PTEN* the lower (335 and 456, respectively).

In ClinVar, germline variants were reported for all genes (N = 5216), with the exception of *SDC2*. Almost half (46.4%) of them were submitted as variants of uncertain significance (VUS). *EGFR* has the highest rate of VUS (67.3%), while all the 36 *LIMK1* variants are classified as benign/likely benign (B/LB). On the opposite way, *PTEN* and *BRAF* have the highest number of pathogenic/likely pathogenic (P/LP) variants, corresponding to 32.5% and 23.9% of all reported variants, respectively (Table 3).

Table 2. Variants reported in gnomAD for the 10 candidate genes. The total numbers and percentages of variants are presented according to their annotations.

GENE	All	Missense	Synonymous	Splice Site	Frameshift	Inframe Del/Ins	Intronic	Nonsense	Stop Lost	Start Lost	5′UTR	3′UTR
AKT1	932	166 (17.81%)	155 (16.63%)	80 (8.58%)	3 (0.32%)	4 (0.43%)	435 (46.67%)	1 (0.11%)	0	0	36 (3.86%)	52 (5.58%)
BRAF	1073	230 (21.44%)	170 (15.84%)	56 (5.22%)	1 (0.09%)	8 (0.75%)	561 (52.28%)	2 (0.19%)	1 (0.09%)	0	13 (1.21%)	31 (2.89%)
EGFR	2000	682 (34.10%)	387 (19.35%)	117 (5.85%)	15 (0.75%)	2 (0.10%)	867 (43.35%)	16 (0.80%)	1 (0.05%)	0	9 (0.45%)	104 (5.20%)
LIMK1	1133	322 (28.42%)	214 (18.89%)	58 (5.12%)	5 (0.44%)	1 (0.09%)	469 (41.39%)	3 (0.26%)	1 (0.09%)	0	44 (3.88%)	16 (1.41%)
PAK1	774	128 (16.54%)	121 (15.63%)	58 (7.49%)	4 (0.52%)	6 (0.78%)	401 (51.81%)	5 (0.65%)	1 (0.13%)	0	8 (1.03%)	42 (5.43%)
PTEN	456	83 (18.20%)	77 (16.89%)	17 (3.73%)	5 (1.10%)	0	223 (48.90%)	5 (1.10%)	0	0	28 (6.14%)	18 (3.95%)
RAF1	1005	264 (26.27%)	145 (14.43%)	65 (6.47%)	4 (0.40%)	2 (0.20%)	490 (48.76%)	7 (0.70%)	0	2 (0.20%)	8 (0.80%)	18 (1.79%)
SDC2	335	102 (30.45%)	52 (15.52%)	18 (5.37%)	5 (1.49%)	3 (0.90%)	114 (34.03%)	0	0	5 (1.49%)	22 (6.57%)	14 (4.18%)
SMARCA4	2575	473 (18.37%)	551 (21.40%)	164 (6.37%)	4 (0.16%)	22 (0.85%)	1318 (51.18%)	0	1 (0.04%)	0	9 (0.35%)	33 (1.28%)
VCP	928	135 (14.55%)	191 (20.58%)	72 (7.76%)	1 (0.11%)	1 (0.11%)	488 (52.59%)	0	0	0	20 (2.16%)	20 (2.16%)

Table 3. Germline variants submitted to ClinVar database for each candidate gene. The number of variants is presented according to its classification. Related syndromes and other relevant conditions for NF1 phenotype are also summarized.

GENE	Classification *					Related Syndromes ***	Other Relevant Reported Conditions
	All **	B/LB	P/LP	CI	VUS		
AKT1	182	94	8	4	76	Cowden, Proteus	E17K variant was associated with 22 conditions, including breast cancer
BRAF	334	125	80	12	117	Cardiofaciocutaneous, Dandy-Walker malformation, LEOPARD, PHACE, Noonan	Astrocytoma, glioma
EGFR	199	54	4	7	134	Cowden, not otherwise specified (NOS) hereditary cancer	Cerebral arteriovenous malformation, inflammatory skin and bowel disease
LIMK1	36	36	0	0	0	-	-
PAK1	6	2	3	-	1	-	intellectual developmental disorder with macrocephaly, seizures, and speech delay
PTEN	1567	367	510	22	668	Bannayan-Riley-Ruvalcaba, Cowden, Hereditary breast and ovarian cancer, NOS Hereditary cancer-predisposing, Proteus-like	Macrocephaly/autism, Phophatase and Tensin (PTEN) Homolog hamartoma tumor
RAF1	412	153	43	17	197	LEOPARD, Noonan	Chordoma, retinoblastoma, and
SDC2	-	-	-	-	-	-	Ieri pleonosteosis are reported in patients carrying the copy number gain of 8q22.1, which includes *SDC2*
SMARCA4	2310	980	95	81	1154	Coffin-Siris, NOS Hereditary cancer-predisposing, Rhabdoid tumor predisposition	Craniopharyngioma, intellectual deficiency, medulloblastoma, neurodevelopmental disorder, neuroblastoma
VCP	170	64	18	10	78	-	Amyotrophic lateral sclerosis, paget disease, Charcot-Marie-Thoth disease

* LB = likely benign; B = benign; LP = likely pathogenic; P = pathogenic; CI = conflicting interpretation; VUS = variant of uncertain significance. ** Unprovided interpretations and drug response variants were not considered. *** Syndromes/conditions reported for benign/likely benign variants, drug response variants, and not-interpreted variants were not considered.

Finally, we also explored TCGA and Genomics Evidence Neoplasia Information Exchange (GENIE) datasets to check for tumor samples harboring both genetic alterations in one of our candidate genes and *NF1*. Due to lack of samples or to the higher mutational and clinical heterogeneity, we managed to make reasonable assumptions only for *AKT1*, *VCP*, and *SDC2*. More details are presented in the discussion section.

3. Discussion

It is evident that genetic variants in *NF1* do not act alone to determinate disease phenotype. Many factors may contribute to disease variability, including environmental factors, the occurrence of epigenetic alterations, and somatic second hits in NF1-associated tumors. The accumulation of somatic NF1 mutations is much more difficult to evaluate since each tumor needs to be sequenced individually, but it may be responsible for some level of NF1 variability. Other symptoms, like delayed mental development, are less influenced by second hit mutations. Genetic modifiers in a single locus or the interaction between several genes may suppress or enhance disease severity, including genes involved in the pathways other than the NF1-Ras-mTOR pathway. There is evidence that genetic modifiers explain a major fraction of phenotypic variation in NF1 [16]. A few genes and their variants have already been described as phenotype modifiers in literature and were reviewed and summarized in Table 4, but they are still insufficient to explain all the variability found in NF1 patients.

Recently, a review pointed out the main methods with which to discover novel phenotype-modifier genes in Mendelian diseases and formulate hypotheses about other pathways than Ras-NF1 that could be phenotype modifiers [44]. The most used methods to select candidate modifier genes are whole-genome sequencing, genome-wide association studies, and experimental approaches using animal models. Other studies also select candidate modifier genes using differential gene expression analysis [20]. These strategies have proved their value in identifying a few phenotype-modifier genes to date; however, they have some disadvantages, such as the high costs involved, being time-consuming, the use and maintenance of animal models, and the confounding factors in studies with selected NF1 patients [18].

One of these limitations was observed in our expression analysis, for which a few candidate genes were actually differentially expressed. Furthermore, differential expression analysis using TCGA tumor samples (Table S2) generated distinct results when compared to GEO controlled knockdown/knockout experiments (Table S3). Expression may depend on the tumor heterogeneity, i.e., the number of cells that are actually not expressing functional NF1, and its location, since the expression profiles of *NF1* and related genes are tissue-dependent. Gene expression is by far the most common analysis among multi-omics studies. Despite that, in many studies it is not possible to obtain a clear scenario of the biological mechanisms disrupted in a disease by evaluating only the mRNA levels. Known disease genes are often not differentially expressed in affected individuals, once the mutations may only alter the protein function or post-translational mechanisms. As a consequence, much information contained in transcriptomics datasets are ignored, demanding an alternate strategy to evaluate these multi-omics assays [45]. On the other hand, the differential gene expression networks also allowed the selection of the *SDC2* gene as a new candidate, once its expression was altered when *NF1* was affected. This example highlighted the need to evaluate the multi-omics data in a more integrated and multidisciplinary perspective [46].

Table 4. List of genes and proteins previously described as NF1 phenotype modifiers in the literature.

Genes/Proteins	Consequence	Methodology Aspects	Reference
ADCY8	Genetic polymorphisms in *ADCY8* are correlated with glioma risk in NF1 in a sex-specific manner, elevating risk in females while reducing risk in males	- Genotyping of NF1 patients using Affymetrix whole-genome human SNP array - Primary astrocyte cultures from NF1-CKO mice and treatment with dideoxyadenosine to induce ADCY inhibition - cAMP regulator expression with qPCR and ELISA	Warrington et al. 2015 [33]
ANRIL allele T of SNP rs2151280	Higher number of plexiform neurofibromas; rs2151280 reduced *ANRIL* transcript levels	- High-resolution array comparative genomic hybridization (aCGH) of PNFs from NF1 patients	Pasmant et al. 2011 [22]
ATP6V0B SNP rs7161 *DPH2* SNP rs4660761 *MSH6* SNP rs1800934	*ATP6V0B* is associated with melanosome biology rs7161 and rs4660761 associated with café-au-lait macule (CALM) count; rs1800934 associated with development an NF1-like phenotype	- Lymphoblastoid cell lines with NF1-associated phenotypes - Gene expression (microarray and qPCR) - Sequencing of genes with increased expression in patients with high count CALM - Meta-analysis	Pemov et al. 2014 [19]
CRLF3, ADAP2, RNF135, UTP6, SUZ12, OMG, LRRC37B, EVI2A, EVI2B, RAB11FIP4, RAB11FIP3, TEFM, ATAD5, CORPS, *NF1* large 17q11 deletions encompassing the entire NF1 locus and neighboring genes	Dysmorphic features, learning disabilities, cardiovascular malformations, childhood overgrowth, a higher tumor burden and earlier onset of benign neurofibromas, and probably, a higher incidence of malignant peripheral nerve sheath tumors (MPSTs)	-MLPA, breakpoint-spanning PCR and FISH in NF1 deletion patients	Mautner et al. 2010 [13]
CXCR4 and its ligand, CXCL12	Highly expressed in mouse models of NF1-deficient MPNSTs, but not in normal precursor cells; Suppression of CXCR4 activity decreases MPNST cell growth in culture and inhibits tumorigenesis in allografts and in spontaneous genetic mouse models of MPNST; Demonstrated conservation of these activated molecular pathways in human MPNSTs	- NF1-deficient skin-derived precursor models (SKPs) and gene expression microarray (normal SKPs; pretumorigenic SKPs with either Nf1 deletion or Nf1 and p53 deletion) - qPCR, westerblot and IHC of CXCR4 and CXCL12 - CXCR4 shRNA for knockdown in SKP MPNST cells - Tissue microarray from plexiform neurofibromas in NF1 patients harboring MPNSTs and MPNSTs samples from NF1 and sporadic patients - CXCR4 cDNAs sequencing	Mo et al. 2013 [34]
GDNF R93W germline variant and maternally inherited *NF1* mutation	Congenital megacolon development	- Investigation of a family carrying variants in *GDNF* and *NRTN* genes with cutaneous manifestations of NF1 and megacolon - Haploinsufficient animal models for *Nf1* and *Trp53* that developed MPNSTs	Bahuau et al. 2001 [35]

Table 4. *Cont.*

Genes/Proteins	Consequence	Methodology Aspects	Reference
miR-34a	Down-regulation of miR-34a founded in most MPNSTs compared to neurofibromas; The p53 inactivation and subsequent loss of expression of miR-34a may contribute to MPNST development	- Microarray of MPNSTs, neurofibromas, Schwannomas, and synovial sarcomas - MPNST cell lines to check for miR-34a and other p53-dependent miRNAs by qRT-PCR after overexpressing wild-type p53	Subramanian et al. 2010 [36]
miR-21	Important in MPNST tumorigenesis and progression through its target, PDCD4	- Global miRNA expression profiling of MPNSTs and neurofibromas - qPCR of differentially expressed miRNAs in MPNSTs, 11 NFs, and 5 normal nerves and MPNST cell lines - Knockdown of miR-21 in MPNST cells	Ilani et al. 2012 [37]
miR-204	Down-regulation of miR-204 contributes to development and tumor progression of MPNSTs	- Global miRNA expression profiling of MPNSTs and benign NF1 neurofibroma tissues - qPCR of differentially expressed in tumor tissues and MPNST cell lines - Lentiviral system for miR-204 trasnfection in NF1 and non-NF1 MPNST cell lines - Non-NF1 MPNST cells Xenograft	Gong et al. 2012 [38]
MSH2, MSH6, MSH3, MLH1, PMS2	Phenotype overlapping between NF1 and Constitutional mismatch repair deficiency (CMMRD) Association with rare childhood malignancies	- Literature review about co-occurrence of symptoms and variants in genes associated with CMMRD and NF1	Wimmer, Rosenbaum and Messiaen 2017 [39]
SDHB	Cause gastrointestinal stromal tumor (GISTs)	- SDHB expression by immunohistochemically in NF1-associated GISTs	Wang, Lasota and Miettinen 2011 [40]
Serotonin receptor 5′*UTR* 5-HT_6 - HTR6 protein	Disrupting HTR6-neurofibromin interaction prevents agonist-independent HTR6-operated cAMP signaling in the prefrontal cortex, an effect that might underlie neuronal abnormalities in NF1 patients; 5-HT_6 receptor may be considered as a potentially therapeutic target to correct some NF1-related cognitive deficits	- Nf1+/− heterozygote mice - HEK-293T and NG108-15 cell lines - Immunoprecipitation followed by Western blot analysis	Deraredj Nadim et al. 2016 [41]
SPRED1 nonsense, frameshift and missense mutations	Complete *SPRED1* inactivation is needed to generate CALMs	- GWAS in unaffected and affected individuals. - *SPRED1* cDNA sequencing - Melanocyte cell culture from normal skin and CALM of NF1 patient - Mouse embryonic fibroblasts	Brems et al. 2007 [42]
TERT mRNA and telomerase activity	Telomere dysfunction may play a role in driving genomic instability and clonal progression in NF1-associated MPNST	- High-resolution Single Telomere Length Analysis (STELA) of cutaneous and diffused plexiforme neurofibromas, and MPNSTs	Jones et al. 2017 [43]

Network analysis makes it possible to combine different multi-omics studies, a strategy that has been applied in several personalized medicine studies evaluating genetic syndromes with phenotypic variability [47–49]. Recently, many projects and consortiums were created, gathering a huge amount of public results with germline and somatic mutation databases, transcriptomics, proteomics, and metabolomics data that can now be evaluated with a systems biology tools [44]. The analysis proposed here can be seen as an optimization in the search for candidate genes acting as phenotype modifiers in NF1, which can later be confirmed by the more robust molecular and functional assays. A brief description of the potential phenotype modifiers found in this study and their variants is provided below to show mechanistic insights and facilitate experimental studies. They are also summarized in Tables 1–3, respectively.

The ontologies selection was an important step in our analysis, especially because *NF1* is important in several molecular mechanisms. If not filtered, our analysis could be later compromised by evidencing genes that are not so deeply associated to neurofibromatosis type 1, but which are more frequently studied in cancer (i.e., *TP53* and developmental genes). This strategy has also been applied in studies that do not identify (or do not have access to) differentially expressed genes [50,51]. For Human Phenotype Ontology (HPO), our workflow was based on a deep phenotyping strategy (computational analysis of detailed, individual clinical abnormalities) [29], according to the heterogeneity of neurofibromatosis type 1 visualized in our patients. We also aimed to avoid ontologies related to congenital anomalies outside the NF1 spectrum of phenotypes, especially the ones that could have led to embryo lethality or severe impairments (major malformations) that would have been diagnosed before NF1. Embryo development is a critical, stepwise controlled process that can be disrupted by genetic or environmental factors, such as maternal infections or exposures [52]. Since the data on *NF1* expression in the embryonic period are scarce, and we do not have information about the maternal genome or environment, we focused on the functional anomalies (more related to the fetal period) that are better characterized in neurofibromatosis type 1.

In our forward strategy, we found the epidermal growth factor receptor (EGFR) that acts upstream of NF1 in the Ras signaling pathway. EGFR belongs to a family of receptor tyrosine kinases that are anchored to the cytoplasmic membrane. EGFR is frequently over-activated in cancer and studies have shown that it is not expressed by normal Schwann cells but it is overexpressed in subpopulations of NF1 mutant Schwann cells [53]. The great involvement of this gene in different cancers and its differential expression patterns in NF1-enriched tissues may indicate that the occurrence of minor variants in this gene could act as phenotype modifiers in NF1. There are 2000 *EGFR* variants registered in the gnomAD database, 34.1% of them missense mutations. In ClinVar, most of the 199 catalogued germline variants are VUS (N = 134). Among the four P/LP variants, one (C326F) was related to Cowden syndrome. Variants in promoter and UTR regions might also have a potential as phenotype modifiers, since animal models have already shown that high levels of *EGFR* expression modify the initiation of neurofibromas, increasing their numbers [54].

AKT1, *BRAF*, *LIMK1*, *PTEN*, and *RAF1* genes were also suggested by the forward strategy. They encode proteins that act downstream of NF1 in the Ras signaling pathway. Phosphoinositide 3-kinase PI3K/AKT is one of the most frequently activated pathways in cancer. This activation may occur through mutation of multiple genes, including *PTEN*, *PIK3R1*, and *mTORC1* [55]. *AKT1* presented the highest levels of betweenness and closeness centrality in systems biology network analysis, demonstrating itself to be the most relevant gene in the information flux among our selected genes (Figure 3).

AKT1 germline mutations are mainly associated with Cowden syndrome, characterized by the appearance of hamartomas, and an increased risk of developing multiple cancers, especially breast cancer [56]. One particular pathogenic variant, E17K, is reported to be linked to 22 different conditions in ClinVar. This variant was also found in gnomAD in a European individual. To explore how this alteration could modify the phenotype when co-occurring with *NF1* mutations, we looked at mutational and clinical data deposited in the GENIE database (v7.0). When excluding *AKT1*-mutated

patients and considering only *NF1* mutations with more deleterious effects (nonsense, frameshift, and splicing variants), the most frequent cancer types are non-small cell lung cancer (12%), glioma (10%), and melanoma (9%). However, when grouping samples with both *AKT1*-E17K and *NF1* mutations, breast cancer becomes the most prevalent, corresponding to 73% of all tumors. The link between breast cancer susceptibility and *NF1* alterations was already established [18]. However, the mechanism that leads to this specific phenotype remains to be elucidated and *AKT1* emerges as a strong candidate.

PI3K-Akt pathway activity is negatively regulated by phosphatase and tensin homolog protein (PTEN) [57]. PTEN is a tumor suppressor and its inactivation has a role in plexiform neurofibroma tumorigenesis and progression to high-grade peripheral nerve sheath tumors in the context of NF1 loss in Schwann cells, which is a very variable symptom in NF1, and may also participate in the mechanism of tumorigenesis of other tumors related to NF1 [58,59]. There are not many variants catalogued for *PTEN* in gnomAD (N = 456), but 1546 were already reported in ClinVar, with 35% still being classified as VUS. The phenotypes related to *PTEN* variants are, as expected, similar to *AKT1*, including Cowden syndrome and hereditary breast and ovarian cancer syndrome. Considering the importance of both *AKT1* and *PTEN* in tumorigenesis, variants in the corresponding genes, not necessarily pathogenic, could act as a modifier of NF1 disease and need to be further investigated.

In contrast, BRAF studies in NF1 patients were already conducted. *BRAF* gene encodes a protein belonging to the RAF family of serine/threonine protein kinases. This protein plays a role in regulating the MAPK/ERK signaling pathway, which affects cell division, differentiation, and secretion. Germline mutations in *BRAF* were previously associated with cardiofaciocutaneous, Noonan, and Costello syndromes [60]. In ClinVar, 334 germline variants were already submitted, 80 being classified as P/LP and 117 of the remainder as VUS. A recent study analyzed a cohort of 100 patients clinically suspected of NF1 and identified 73 *NF1* mutations and two *BRAF* novel variants. The clinical features of NF1 patients with co-occurrence of *NF1-BRAF* mutations were severe, and *BRAF* variants may have a synergistic role in determining NF1 phenotype [61].

Another member of the same family, the *RAF1* gene, was suggested by the forward strategy. The *RAF1* gene encodes a serine/threonine kinase protein that functions downstream of RAS and activates MEK1 and MEK2. In GENIE, *RAF1* mutations are observed in various cancers. LEOPARD and Noonan syndromes were already associated in ClinVar with 6 and 28 pathogenic *RAF1* variants, respectively. The other 197 variants were still classified as VUS and included conditions such as other rasopathies, chordoma, and retinoblastoma.

Finally, by the forward strategy, *LIMK1* was suggested as a phenotype modifier. The N-terminal domain of LIM kinase 1 (*LIMK1*) regulates actin dynamics, affects cell adhesion and migration by phosphorylating cofilin, and negatively regulates the Rac1/Pak1/LIMK1/cofilin pathway [62]. *NF1* is an upstream regulator of *LIMK1* by acting on cofilin phosphorylation. When *NF1* is mutated, this pathway is affected, possibly influencing neuronal development and cognitive deficits associated with the disease [62,63]. We found 1133 *LIMK1* variants in gnomAD, but only 36 reported as germline in ClinVar, all of them classified as benign/likely benign. That may indicate that mutated *LIMK1* alone is not pathogenic, but it does not exclude a combined effect of *NF1* variants acting as a phenotype modifier.

By the reverse strategy, *AKT1, PTEN,* and *RAF1* were also suggested, reinforcing the possible roles of these genes in NF1 phenotypes. Additionally, *PAK1, SMARCA4,* and *VCP* were found. The kinase PAK1 is a Rac/CDC42-dependent serine/threonine kinase that acts by activating several kinases such as RAF, ERK, and LIMK1, and other related pathways by activating TGFα and VEGF. PAK1 is required for the malignant growth of RAS transformants in NF1 neurofibrosarcoma cell lines [64,65]. There are not many PAK1 variants registered in gnomAD (N = 774) and only six in ClinVar. However, two LP variants (Y131C and Y429C) reported in ClinVar were associated with an intellectual developmental disorder with macrocephaly, seizures, and speech delay, phenotypes that are reported in NF1 patients.

SMARCA4 is a central component of the switch/sucrose-non-fermentable (SWI/SNF) chromatin remodeling complex. Inactivating mutations and loss of expression in several components of this complex have been implicated in carcinogenesis [66–68]. Thus, variants in one of these genes might

influence the NF1 phenotype. *SMARCA4* has the highest number of variants in both gnomAD and ClinVar among our candidate genes: 2575 and 2310, respectively. Peripheral nerve sheath tumors were already reported in patients with the syndrome carrying *SMARCB1* mutations, which belongs to the same family of *SMARCA4* [14]. Loss of *SMARCB1* was also related to Schwannoma, another phenotype found in NF1 despite being more frequent in NF2 [69].

The valosin-containing protein (*VCP*) appeared as a strong candidate for being an NF1 phenotype modifier by our random-walk analysis. *VCP* gene is associated with the multisystem degenerative autosomal dominant disorder of inclusion body myopathy with Paget disease of bone and frontotemporal dementia (IBMPFD) and mutations were related to 1–2% of amyotrophic lateral sclerosis cases [70]. However, missense mutations in *VCP*, and low-effect and low-penetrant mutations in this gene, have controversial roles in causing disease. Neurofibromin interacts with VCP through its Leucine- Rich Domain (LRD)-domain [71]. Patients with NF1 who have mutations in the LRD coding region were described to be more prone to developing cognitive deficits than those with mutations elsewhere in the *NF1* gene [72]. In the same study, it was observed that point mutations in the LRD coding region in the *NF1* gene abolished the ability of NF1 to interact with VCP, while *VCP* mutants were shown to have reduced affinity for NF1. Interestingly, non-disease-associated polymorphisms in the LRD region of the *NF1* gene may increase the risk of an IBMPFD patient developing dementia. In the same way, polymorphisms in the *VCP* gene that code for domains that interact with NF1 might influence the NF1 phenotype. These data obtained from literature research reinforce the accuracy of our systems biology analysis and random-walk mathematical model, which pointed *VCP* as a strong NF1 phenotype modifier candidate.

It would not be a surprise if *VCP* variants were found co-occurring with *NF1* alterations, especially the ones in ATPAse domains 1 and 2 (D1 and D2), responsible for interacting with neurofibromin's LRD-domain. For example, in ClinVar only two variants are reported as pathogenic D1/D2, one in each domain. On the other hand, 34 remain as VUS, 23 in D1, and 11 in D2. In gnomAD, more than half of the (N = 491) cataloged *VCP* variants are located in D1 and D2 domains. Looking at TCGA, there are few samples (N = 28) with *NF1* mutations co-occurring with variants in *VCP* D1/D2 domains, most of them (47.2%) from uterine corpus endometrial carcinomas.

The last gene that was suggested by our analysis is *SDC2*, which was found by differential gene expression networks and pointed to as a strong candidate by our mathematical model. The heparan sulfate part of SDC2 interacts with extracellular matrix proteins and growth factors to act as an adhesion molecule and as a coreceptor [73]. Variants in this gene might be associated with autism spectrum disorder [74]. Interestingly, some studies showed a higher frequency of autism spectrum disorder in NF1 children, and this is a variable condition in NF1 that might be influenced by variants in other genes [75]. Despite *SDC2* emerging as a strong candidate in our study, only 335 variants were found in the gnomAD database and none in ClinVar, suggesting a highly conserved gene. However, the lack of *SDC2* in ClinVar may merely reflect its absence from gene panels used for diagnostic purposes. On the other hand, TCGA somatic samples carrying both *NF1* and *SDC2* mutations are scarce (N = 14/10,437), most of them (57.1%) also related to uterine corpus endometrial carcinomas. This finding does not exclude the gene as a phenotype modifier, but variants co-occurring with *NF1* alterations might be a rare event.

The literature search for variants and functions of the candidate modifier genes identified by our strategy shows that this is an economical and accurate way to filter and select genes that would be further validated by experimental assays. As a perspective, variants in the ten genes selected by our strategy will be searched in NF1 patients with different symptoms. Many strategies could be used to subsequently evaluate and validate the selected genes. One of them is to identify variants in these genes or other nearby genes and genotype those variants in NF1 patients with different symptoms and control populations, followed by statistical methods to identify correlations with the phenotype. Moreover, in-vitro and in-vivo studies are also useful for validating previously selected genes, focusing on CRISPR/Cas9 assays to induce partially and complete loss of the proteins. Our candidate genes

could be also included in commercial gene panels with a low impact on their coast, which would help to feed public databases such as ClinVar.

One obvious limitation of the present study is the lack of proper validation for the candidate phenotype-modifier genes using benchwork. Hence, the results obtained must be evaluated with caution. Experimental validation is necessary and strongly recommended before clinical extrapolation. However, our purpose was to provide a new look in the strategies for evaluation of neurofibromatosis using the huge amount of data already available in shared public-curated datasets. For example, the protein–protein interactions identified by us were previously validated by several in vitro assays and here combined in a network. Together with our network analysis, we also performed random walk, a robust mathematical model that has been applied in the analysis of biological multiplex heterogeneous networks [76,77]. We hope this complex and robust systems biology approach will help to better understand the neurofibromatosis type 1 and its phenotypic variation.

4. Materials and Methods

4.1. Selection of NF1 Ontologies

The complete list of NF1 gene ontologies (GO) and phenotype ontologies were obtained in the AmiGO and Human Phenotype Ontology (HPO) databases, respectively. In modifier studies, the selection of which phenotypes to study is a key step, and in NF1, several phenotypic features are time-dependent [19]. Then, we selected from both lists (GO and HPO) the ontologies related to the less frequent and variable characteristics and not necessarily time-dependent, presented by NF1 patients, as reported in the literature and in our clinical experience in the Oncogenetics clinics of Hospital de Clínicas de Porto Alegre [16,78,79]. For example, cutaneous neurofibromas are common and may occur in up to 99% of NF1 patients in a cohort; this is a variable characteristic, mainly in the number of neurofibromas, but it is less variable when considering their presence in NF1 patients. Thus, we focused on ontologies related to characteristics that occur in a smaller number of patients to try to explain the variability of less common but more aggressive NF1 symptoms, such as breast cancer, delayed mental development, plexiform neurofibromas, and facial dysmorphism. The processes and phenotypes selected for the analysis involved NF1 and NF1-related signaling pathways, such as the MAPK cascade and regulation of the Ras pathway, considering the upper ontology in the hierarchy of each database. Hence, some ontologies were not selected because there was an upper term in the hierarchy that encompassed these ontologies. It is worth mentioning that we followed a guide for the correct selection of ontologies to try to limit the bias introduced by the choice of terms and keywords [80]. The processes were evaluated by two independent researchers and selected for subsequent analysis when both researchers considered it relevant.

4.2. Systems Biology Analysis

Networks were generated using STRING database v.11, comprising protein–protein interactions (PPI) for *Homo sapiens*. Only experimental interactions were selected, with a minimum required interaction score set in 0.400 (default). The assembled networks were transferred to Cytoscape v.3.7.2 software, with which the network statistics was obtained. Big nodes represented proteins with high betweenness centrality scores and warm colors comprised proteins with high closeness centrality measures.

Comparison between networks was performed with DyNet application for Cytoscape v.3.7.2. Complex networks comprising HPO selected phenotypes, gene expression, and PPI networks were assembled using the PhenomeScape app, also in Cytoscape v.3.7.2, using the default settings.

4.3. Gene Expression Evaluation

RNA-seq and microarray secondary analysis were performed using studies selected in Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases. For the GEO studies,

we looked for *NF1* knockdown or knockout assays and selected only the ones performed in human tumor cells. The data extraction was performed manually, and the robust multiarray averaging (RMA) normalization was applied using *oligo* or *affy* R packages (R v.3.6.2). The differentially expressed genes were obtained using the *limma* package (R v.3.6.2).

Firstly, for TCGA, we selected seven somatic tumors with nonsense mutations in *NF1* (*NF1*-ns): bladder urothelial carcinoma (BLCA), brain lower grade glioma (LGG), breast invasive carcinoma (BRCA), cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), colon adenocarcinoma (COAD), glioblastoma multiforme (GBM), and pheochromocytoma and paraganglioma (PCPG). Despite tumors having a somatic origin, this information can be useful to check how alterations in *NF1* could affect the global gene expression in a specific tissue/tumor (phenotype). We then compared these samples against wild type *NF1* tumors to check which genes were differentially expressed only in the *NF1*-ns group. The gene expression analysis for TCGA data was performed by extracting the data with *TCGAbiolinks* package and evaluating differential gene expression with *edgeR* package. All analyses were performed in R v.3.6.2.

4.4. Random Walk Analysis

The heterogeneous networks comprising both the genes and phenotypes selected were assembled in the *RandomWalkRestartMH* package, in R v.3.6.2, and the random walk analysis was performed with the same package.

4.5. Database Research

Other databases consulted to obtain data for the potential NF1 phenotype-modifier genes were: (1) BioGrid, for curated protein interactions; (2) PINOT tool, for literature data on curated protein interactions; (3) STRING database, for protein–protein interactions; and (4) The Human Reference Protein Interactome (HuRI), for the binary protein–protein interactions.

4.6. Variant Datasets

To explore variants in our candidate genes already reported in the general population or with clinical significance, we consulted The Genome Aggregation Database (gnomAD) v 2.1.1 and the ClinVar archive. For gnomAD, variants were classified according to its annotation. In ClinVar, only variants reported by at least one submitter as a germline were considered and classified according to their interpretations.

Finally, an additional analysis was performed consulting the 79,720 tumor samples made available by the AACR Project GENIE and the 10,967 samples from the TCGA PanCancer Atlas studies, using the cBioPortal for Cancer Genomics. Samples were filtered according to their mutational status: *NF1*-mutated patients including only nonsense, frameshift, and splicing; and patients with selected variants in our candidate genes, if available. Then, the clinical data were accessed and confronted with the mutational status to check which cancer types were predominant when *NF1* was exclusively altered and when *NF1* variant co-occurred with variants in our candidate genes.

5. Conclusions

We presented here a not yet explored systems biology strategy to investigate *NF1* phenotype modifiers. The public availability of multi-omics datasets makes possible the use of robust tools to generate complex networks including protein–protein interactions, differential expression data, and phenotypes, reinforced by mathematical models such as random-walk. Combining all these strategies, we found 10 candidate genes as potential NF1 phenotype modifiers. Resources and time may be scarce to carry out association studies and systems biology analyses makes possible to better explain the genetic heterogeneity of this complex syndrome. Our results must be interpreted cautiously in clinical application and may guide further in-vitro and in-vivo validation studies, saving time and

financial resources. The approach presented here may guide further in-vitro and in-vivo validation studies, saving time and financial resources.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-6694/12/9/2416/s1. Figure S1: Network for GO analysis. Figure S2: Network for HPO analysis. Figure S3: A complex network generated using GEO MPNST dataset, protein–protein interactions, and HPO ontologies as input data. Figure S4: A complex network generated using GEO neurofibroma dataset, protein–protein interactions, and HPO ontologies as input data. Figure S5: A complex network generated using GEO *NF1*-shRNA, protein–protein interactions, and HPO ontologies as input data. Figure S6: A network generated using GEO CRISPR-induced *knockout* of *NF1*, protein–protein interactions, and HPO ontologies as input data. Figure S7: A network generated using GEO neuroblastoma cell line, protein–protein interactions, and HPO ontologies as input data. Figure S8: Random walk analysis calculating the minimum steps (interactions) that a candidate gene (node) takes to reach the neurofibromatosis 1 phenotype. Table S1: Selected NF1 ontologies, Table S2: *NF1* knockdown and knockout assays in GEO database, Table S3: Tumors with *NF1* nonsense mutations from TCGA database, Table S4: List of genes and proteins previously described as NF1 phenotype modifiers in the literature.

Author Contributions: Conceptualization, T.W.K. and C.R.; formal analysis, T.W.K. and T.F.A.; methodology, T.W.K., L.B.R., T.F.A., and C.R.; project administration, C.R.; software, T.W.K. and T.F.A.; supervision, P.A.-P. and C.R.; validation, T.W.K. and T.F.A.; writing—original draft, T.W.K., L.B.R., T.F.A., and C.R.; writing—review and editing, L.B.R., T.F.A., P.A.-P., and C.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We would like to thank Fundo de Incentivo à Pesquisa (FIPE) of Hospital de Clínicas de Porto Alegre and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial and organizational support.

Conflicts of Interest: The authors stated that they had no interests which might be perceived as posing any conflict or bias.

References

1. Evans, D.; Howard, E.; Giblin, C.; Clancy, T.; Spencer, H.; Huson, S.; Lalloo, F.; Evans, D.G. Birth incidence and prevalence of tumor-prone syndromes: Estimates from a UK family genetic register service. *Am. J. Med. Genet. Part A* **2010**, *152*, 327–332. [CrossRef] [PubMed]
2. Jett, K.; Friedman, J.M. Clinical and genetic aspects of neurofibromatosis 1. *Genet. Med.* **2009**, *12*, 1–11. [CrossRef] [PubMed]
3. Ferner, R.E. Neurofibromatosis 1 and neurofibromatosis 2: A twenty first century perspective. *Lancet Neurol.* **2007**, *6*, 340–351. [CrossRef]
4. Evans, D.G.; Baser, M.E.; McGaughran, J.M.; Sharif, S.; Howard, E.; Moran, A. Malignant peripheral nerve sheath tumours in neurofibromatosis 1. *J. Med. Genet.* **2002**, *39*, 311–314. [CrossRef]
5. Seminog, O.; Goldacre, M.J. Risk of benign tumours of nervous system, and of malignant neoplasms, in people with neurofibromatosis: Population-based record-linkage study. *Br. J. Cancer* **2012**, *108*, 193–198. [CrossRef]
6. Carey, J.C.; Baty, B.J.; Johnson, J.P.; Morrison, T.; Skolnick, M.; Kivlin, J. The genetic aspects of neurofibromatosis. *Ann. N. Y. Acad. Sci.* **1986**, *486*, 45–56. [CrossRef]
7. Mensink, K.A.; Ketterling, R.P.; Flynn, H.C.; Knudson, R.A.; Lindor, N.M.; Heese, B.A.; Spinner, R.J.; Babovic-Vuksanovic, D. Connective tissue dysplasia in five new patients with NF1 microdeletions: Further expansion of phenotype and review of the literature. *J. Med. Genet.* **2005**, *43*, e8. [CrossRef]
8. De Raedt, T.; Brems, H.; Wolkenstein, P.; Vidaud, D.; Pilotti, S.; Perrone, F.; Mautner, V.; Frahm, S.; Sciot, R.; Legius, E. Elevated Risk for MPNST in NF1 Microdeletion Patients. *Am. J. Hum. Genet.* **2003**, *72*, 1288–1292. [CrossRef]
9. Descheemaeker, M.; Roelandts, K.; De Raedt, T.; Brems, H.; Fryns, J.; Legius, E. Intelligence in individuals with a neurofibromatosis type 1 microdeletion. *Am. J. Med. Genet.* **2004**, *131*, 325–326. [CrossRef]
10. Upadhyaya, M.; Huson, S.M.; Davies, M.; Thomas, N.; Chuzhanova, N.; Giovannini, S.; Evans, D.G.; Howard, E.; Kerr, B.; Griffiths, S.; et al. An absence of cutaneous neurofibromas associated with a 3-bp Inframe deletion in Exon 17 of the NF1 Gene (c.2970-2972 delAAT): Evidence of a clinically significant NF1 Genotype-Phenotype correlation. *Am. J. Hum. Genet.* **2007**, *80*, 140–151. [CrossRef]

11. Pinna, V.; Lanari, V.; Daniele, P.; Consoli, F.; Agolini, E.; Margiotti, K.; Bottillo, I.; Torrente, I.; Bruselles, A.; Fusilli, C.; et al. p.Arg1809Cys substitution in neurofibromin is associated with a distinctive NF1 phenotype without neurofibromas. *Eur. J. Hum. Genet.* **2014**, *23*, 1068–1071. [CrossRef] [PubMed]
12. Koczkowska, M.; Chen, Y.; Callens, T.; Gomes, A.; Sharp, A.; Johnson, S.; Hsiao, M.-C.; Chen, Z.; Balasubramanian, M.; Barnett, C.P.; et al. Genotype-Phenotype correlation in NF1: Evidence for a more severe Phenotype associated with Missense mutations affecting NF1 Codons 844–848. *Am. J. Hum. Genet.* **2018**, *102*, 69–87. [CrossRef]
13. Mautner, V.-F.; Kluwe, L.; Friedrich, R.; Roehl, A.C.; Bammert, S.; Högel, J.; Cooper, D.N.; Kehrer-Sawatzki, H.; Spöri, H. Clinical characterisation of 29 neurofibromatosis type-1 patients with molecularly ascertained 1.4 Mb type-1 NF1 deletions. *J. Med. Genet.* **2010**, *47*, 623–630. [CrossRef] [PubMed]
14. Upadhyaya, S.A.; McGee, R.B.; Wilky, B.A.; Broniscer, A. Malignant progression of a peripheral nerve sheath tumor in the setting of rhabdoid tumor predisposition syndrome. *Pediatr. Blood Cancer* **2018**, *65*, e27030. [CrossRef] [PubMed]
15. Grisart, B.; Rack, K.; Vidrequin, S.; Hilbert, P.; Deltenre, P.; Verellen-Dumoulin, C.; Destrée, A.; Destr, A. NF1 microduplication first clinical report: Association with mild mental retardation, early onset of baldness and dental enamel hypoplasia? *Eur. J. Hum. Genet.* **2008**, *16*, 305–311. [CrossRef] [PubMed]
16. Sabbagh, A.; Pasmant, E.; Laurendeau, I.; Parfait, B.; Barbarot, S.; Guillot, B.; Combemale, P.; Ferkal, S.; Vidaud, M.; Aubourg, P.; et al. Unravelling the genetic basis of variable clinical expression in neurofibromatosis 1. *Hum. Mol. Genet.* **2009**, *18*, 2768–2778. [CrossRef]
17. Easton, D.F.; Ponder, M.A.; Huson, S.M.; Ponder, B.A. An analysis of variation in expression of neurofibromatosis (NF) type 1 (NF1): Evidence for modifying genes. *Am. J. Hum. Genet.* **1993**, *53*, 305–313.
18. Pasmant, E.; Vidaud, M.; Vidaud, D.; Wolkenstein, P. Neurofibromatosis type 1: From genotype to phenotype. *J. Med. Genet.* **2012**, *49*, 483–489. [CrossRef]
19. Pemov, A.; Sung, H.; Hyland, P.L.; Sloan, J.L.; Ruppert, S.L.; Baldwin, A.M.; Boland, J.F.; Bass, S.E.; Lee, H.J.; Jones, K.M.; et al. Genetic modifiers of neurofibromatosis type 1-associated café-au-lait macule count identified using multi-platform analysis. *PLoS Genet.* **2014**, *10*, e1004575. [CrossRef]
20. Pasmant, E.; Sabbagh, A.; Spurlock, G.; Laurendeau, I.; Grillo, E.; Hamel, M.-J.; Martin, L.; Barbarot, S.; Leheup, B.; Rodriguez, D.; et al. NF1 microdeletions in neurofibromatosis type 1: From genotype to phenotype. *Hum. Mutat.* **2010**, *31*, E1506–E1518. [CrossRef]
21. Tritto, V.; Ferrari, L.; Esposito, S.; Zuccotti, P.; Bianchessi, D.; Natacci, F.; Saletti, V.; Eoli, M.; Riva, P. Non-coding RNA and tumor development in neurofibromatosis Type 1: ANRIL Rs2151280 is associated with optic glioma development and a mild phenotype in neurofibromatosis Type 1 Patients. *Genes* **2019**, *10*, 892. [CrossRef] [PubMed]
22. Pasmant, E.; Sabbagh, A.; Masliah-Planchon, J.; Ortonne, N.; Laurendeau, I.; Melin, L.; Ferkal, S.; Hernandez, L.; Leroy, K.; Valeyrie-Allanore, L.; et al. Role of Noncoding RNA ANRIL in Genesis of Plexiform Neurofibromas in Neurofibromatosis Type 1. *J. Natl. Cancer Inst.* **2011**, *103*, 1713–1722. [CrossRef] [PubMed]
23. Muram-Zborovski, T.M.; Stevenson, D.A.; Viskochil, D.H.; Dries, D.C.; Wilson, A.; Mao, R. SPRED 1 mutations in a neurofibromatosis clinic. *J. Child Neurol.* **2010**, *25*, 1203–1209. [CrossRef]
24. Yang, F.; Xu, S.; Liu, R.; Shi, T.; Li, X.; Li, X.; Chen, G.; Liu, H.; Zhou, Q.; Chen, J. The investigation for potential modifier genes in patients with neurofibromatosis type 1 based on next-generation sequencing. *OncoTargets Ther.* **2018**, *11*, 919–932. [CrossRef] [PubMed]
25. Sharafi, P.; Ayter, S. Possible modifier genes in the variation of neurofibromatosis type 1 clinical phenotypes. *J. Neurogenet.* **2018**, *32*, 65–77. [CrossRef]
26. Le Novère, N. Quantitative and logic modelling of molecular and gene networks. *Nat. Rev. Genet.* **2015**, *16*, 146–158. [CrossRef]
27. Kowalski, T.W.; Dupont, Á.D.V.; Rengel, B.D.; Sgarioni, E.; Gomes, J.D.A.; Fraga, L.R.; Schuler-Faccini, L.; Vianna, F.S.L. Assembling systems biology, embryo development and teratogenesis: What do we know so far and where to go next? *Reprod. Toxicol.* **2019**, *88*, 67–75. [CrossRef]
28. Verster, J.C.; Tzschentke, T.M.; O'Malley, K.; Colpaert, F.C.; Ellenbroek, B.; McAllister-Williams, R.H.; Liepert, J.; Hillard, C.J.; Preskorn, S.; Dahmen, M.M.; et al. Forward Genetics/Reverse Genetics. In *Encyclopedia of Psychopharmacology*; Springer Science and Business Media LLC: Berlin, Germany, 2010; p. 544.

29. Köhler, S.; Carmody, L.C.; Vasilevsky, N.; Jacobsen, J.O.B.; Danis, D.; Gourdine, J.-P.; Gargano, M.; Harris, N.L.; Matentzoglu, N.; McMurry, J.A.; et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res.* **2018**, *47*, D1018–D1027. [CrossRef]
30. Zhu, J.; Zhao, Q.; Katsevich, E.; Sabatti, C. Exploratory gene ontology analysis with interactive visualization. *Sci. Rep.* **2019**, *9*, 7793. [CrossRef]
31. Yu, H.; Kim, P.M.; Sprecher, E.; Trifonov, V.; Gerstein, M. The importance of bottlenecks in protein networks: Correlation with gene essentiality and expression dynamics. *PLoS Comput. Boil.* **2007**, *3*, e59. [CrossRef]
32. Valdeolivas, A.; Tichit, L.; Navarro, C.; Perrin, S.; Odelin, G.; Levy, N.; Cau, P.; Remy, E.; Baudot, A. Random walk with restart on multiplex and heterogeneous biological networks. *Bioinformatics* **2018**, *35*, 497–505. [CrossRef] [PubMed]
33. Warrington, N.M.; Sun, T.; Luo, J.; McKinstry, R.C.; Parkin, P.C.; Ganzhorn, S.; Spoljaric, D.; Albers, A.C.; Merkelson, A.; Stewart, U.R.; et al. The cyclic AMP pathway is a sex-specific modifier of glioma risk in type I neurofibromatosis patients. *Cancer Res.* **2014**, *75*, 16–21. [CrossRef]
34. Mo, W.; Chen, J.; Patel, A.; Zhang, L.; Chau, V.; Li, Y.; Cho, W.; Lim, K.; Xu, J.; Lazar, A.J.F.; et al. CXCR4/CXCL12 mediate autocrine cell- cycle progression in NF1-associated malignant peripheral nerve sheath tumors. *Cell* **2013**, *152*, 1077–1090. [CrossRef] [PubMed]
35. Bahuau, M.; Pelet, A.; Vidaud, M.; Lamireau, T.; Le Bail, B.; Munnich, A.; Vidaud, M.; Lyonnet, S.; Lacombe, D. GDNF as a candidate modifier in a type 1 neurofibromatosis (NF1) enteric phenotype. *J. Med. Genet.* **2001**, *38*, 638–643. [CrossRef] [PubMed]
36. Subramanian, S.; Thayanithy, V.; West, R.B.; Lee, C.-H.; Beck, A.H.; Zhu, S.; Downs-Kelly, E.; Montgomery, K.; Goldblum, J.R.; Hogendoorn, P.C.; et al. Genome-wide transcriptome analyses reveal p53 inactivation mediated loss of miR-34a expression in malignant peripheral nerve sheath tumours. *J. Pathol.* **2010**, *220*, 58–70. [CrossRef] [PubMed]
37. Itani, S.; Kunisada, T.; Morimoto, Y.; Yoshida, A.; Sasaki, T.; Ito, S.; Ouchida, M.; Sugihara, S.; Shimizu, K.; Ozaki, T. MicroRNA-21 correlates with tumorigenesis in malignant peripheral nerve sheath tumor (MPNST) via programmed cell death protein 4 (PDCD4). *J. Cancer Res. Clin. Oncol.* **2012**, *138*, 1501–1509. [CrossRef]
38. Gong, M.; Ma, J.; Li, M.; Zhou, M.; Hock, J.; Yu, X. MicroRNA-204 critically regulates carcinogenesis in malignant peripheral nerve sheath tumors. *Neuro-oncology* **2012**, *14*, 1007–1017. [CrossRef]
39. Wimmer, K.; Rosenbaum, T.; Messiaen, L. Connections between constitutional mismatch repair deficiency syndrome and neurofibromatosis type 1. *Clin. Genet.* **2017**, *91*, 507–519. [CrossRef]
40. Wang, J.H. Succinate Dehydrogenase subunit B (SDHB) is expressed in neurofibromatosis 1-associated Gastrointestinal Stromal Tumors (Gists): Implications for the SDHB expression based classification of Gists. *J. Cancer* **2011**, *2*, 90. [CrossRef]
41. Nadim, W.D.; Chaumont-Dubel, S.; Madouri, F.; Cobret, L.; De Tauzia, M.-L.; Zajdel, P.; Bénédetti, H.; Marin, P.; Morisset-Lopez, S. Physical interaction between neurofibromin and serotonin 5-HT6receptor promotes receptor constitutive activity. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12310–12315. [CrossRef]
42. Brems, H.; Chmara, M.; Sahbatou, M.; Denayer, E.; Taniguchi, K.; Kato, R.; Somers, R.; Messiaen, L.; De Schepper, S.; Fryns, J.-P.; et al. Germline loss-of-function mutations in SPRED1 cause a neurofibromatosis 1–like phenotype. *Nat. Genet.* **2007**, *39*, 1120–1126. [CrossRef] [PubMed]
43. Jones, R.E.; Grimstead, J.W.; Sedani, A.; Baird, D.M.; Upadhyaya, M. Telomere erosion in NF1 tumorigenesis. *Oncotarget* **2017**, *8*, 40132–40139. [CrossRef] [PubMed]
44. Rahit, K.T.H.; Tarailo-Graovac, M. Genetic modifiers and rare mendelian disease. *Genes* **2020**, *11*, 239. [CrossRef] [PubMed]
45. De La Fuente, A. From 'differential expression' to 'differential networking'—Identification of dysfunctional regulatory networks in diseases. *Trends Genet.* **2010**, *26*, 326–333. [CrossRef] [PubMed]
46. Moreno-Sánchez, R.; Saavedra, E.; Gallardo-Pérez, J.C.; Rumjanek, F.D.; Rodríguez-Enríquez, S. Understanding the cancer cell phenotype beyond the limitations of current omics analyses. *FEBS J.* **2015**, *283*, 54–73. [CrossRef] [PubMed]
47. Manczinger, M.; Kemény, L. Novel Factors in the Pathogenesis of Psoriasis and Potential Drug Candidates Are Found with Systems Biology Approach. *PLoS ONE* **2013**, *8*, e80751. [CrossRef]
48. Carapito, R.; Carapito, C.; Morlon, A.; Paul, N.; Jacome, A.S.V.; Alsaleh, G.; Rolli, V.; Tahar, O.; Aouadi, I.; Rompais, M.; et al. Multi-OMICS analyses unveil STAT1 as a potential modifier gene in mevalonate kinase deficiency. *Ann. Rheum. Dis.* **2018**, *77*, 1675–1687. [CrossRef]

49. Loviglio, M.N.; Beck, C.R.; White, J.; Leleu, M.; Harel, T.; Guex, N.; Niknejad, A.; Bi, W.; Chen, E.S.; Crespo, I.; et al. Identification of a RAI1-associated disease network through integration of exome sequencing, transcriptomics, and 3D genomics. *Genome Med.* **2016**, *8*, 105. [CrossRef]
50. Ghatge, M.; Nair, J.; Sharma, A.; Vangala, R.K. Integrative gene ontology and network analysis of coronary artery disease associated genes suggests potential role of ErbB pathway gene EGFR. *Mol. Med. Rep.* **2018**, *17*, 4253–4264. [CrossRef]
51. Zhang, P.; Wang, J.; Zhang, X.; Wang, X.; Jiang, L.; Gu, X. Identification of AIDS-associated Kaposi Sarcoma: A functional genomics approach. *Front. Genet.* **2020**, *10*, 1376. [CrossRef]
52. Worley, K.; Rico-Varela, J.; Ho, D.; Wan, L.Q. Teratogen screening with human pluripotent stem cells. *Integr. Boil.* **2018**, *10*, 491–501. [CrossRef] [PubMed]
53. Ling, B.C.; Wu, J.; Miller, S.J.; Monk, K.R.; Shamekh, R.; Rizvi, T.A.; Decourten-Myers, G.; Vogel, K.S.; DeClue, J.E.; Ratner, N. Role for the epidermal growth factor receptor in neurofibromatosis-related peripheral nerve tumorigenesis. *Cancer Cell* **2005**, *7*, 65–75. [CrossRef] [PubMed]
54. Wu, J.; Liu, W.; Williams, J.P.; Ratner, N. EGFR-Stat3 signalling in nerve glial cells modifies neurofibroma initiation. *Oncogene* **2016**, *36*, 1669–1677. [CrossRef] [PubMed]
55. Shaw, R.J.; Cantley, L.C. Ras, PI(3)K and mTOR signalling controls tumour cell growth. *Nature* **2006**, *441*, 424–430. [CrossRef] [PubMed]
56. Orloff, M.S.; He, X.; Peterson, C.; Chen, F.; Chen, J.-L.; Mester, J.L.; Eng, C. Germline PIK3CA and AKT1 mutations in Cowden and Cowden-like syndromes. *Am. J. Hum. Genet.* **2013**, *92*, 76–80. [CrossRef] [PubMed]
57. Madhunapantula, S.V.; Mosca, P.J.; Robertson, G.P. The Akt signaling pathway. *Cancer Boil. Ther.* **2011**, *12*, 1032–1049. [CrossRef]
58. Keng, V.W.; Rahrmann, E.P.; Watson, A.L.; Tschida, B.R.; Moertel, C.L.; Jessen, W.J.; Rizvi, T.A.; Collins, M.H.; Ratner, N.; Largaespada, D.A. PTEN and NF1 inactivation in Schwann cells produces a severe Phenotype in the peripheral nervous system that promotes the development and malignant progression of peripheral nerve sheath tumors. *Cancer Res.* **2012**, *72*, 3405–3413. [CrossRef]
59. Gregorian, C.; Nakashima, J.; Dry, S.M.; Nghiemphu, P.L.; Smith, K.B.; Ao, Y.; Dang, J.; Lawson, G.; Mellinghoff, I.K.; Mischel, P.S.; et al. PTEN dosage is essential for neurofibroma development and malignant transformation. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 19479–19484. [CrossRef]
60. Sarkozy, A.; Carta, C.; Moretti, S.; Zampino, G.; Digilio, M.C.; Pantaleoni, F.; Scioletti, A.P.; Esposito, G.; Cordeddu, V.; Lepri, F.R.; et al. Germline BRAF mutations in Noonan, LEOPARD, and cardiofaciocutaneous syndromes: Molecular diversity and associated phenotypic spectrum. *Hum. Mutat.* **2009**, *30*, 695–702. [CrossRef]
61. Wu-Chou, Y.-H.; Hung, T.-C.; Lin, Y.-T.; Cheng, H.-W.; Lin, J.-L.; Lin, C.-H.; Yu, C.-C.; Chen, P.; Yeh, T.-H.; Chen, Y.-R. Genetic diagnosis of neurofibromatosis type 1: Targeted next- generation sequencing with Multiple Ligation-Dependent Probe Amplification analysis. *J. Biomed. Sci.* **2018**, *25*, 72. [CrossRef]
62. Manetti, F. LIM kinases are attractive targets with many macromolecular partners and only a few small molecule regulators. *Med. Res. Rev.* **2011**, *32*, 968–998. [CrossRef] [PubMed]
63. Scott, R.W.; Olson, M.F. LIM kinases: Function, regulation and association with human disease. *J. Mol. Med.* **2007**, *85*, 555–568. [CrossRef] [PubMed]
64. Tang, Y.; Marwaha, S.; Rutkowski, J.L.; Tennekoon, G.I.; Phillips, P.C.; Field, J. A role for Pak protein kinases in Schwann cell transformation. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5139–5144. [CrossRef] [PubMed]
65. Tang, Y.; Chen, Z.; Ambrose, D.; Liu, J.; Gibbs, J.B.; Chernoff, J.; Field, J. Kinase-deficient Pak1 mutants inhibit Ras transformation of Rat-1 fibroblasts. *Mol. Cell. Boil.* **1997**, *17*, 4454–4464. [CrossRef] [PubMed]
66. Wilson, B.G.; Roberts, C.W. SWI/SNF nucleosome remodellers and cancer. *Nat. Rev. Cancer* **2011**, *11*, 481–492. [CrossRef] [PubMed]
67. Roberts, C.W.; Orkin, S.H. The SWI/SNF complex—Chromatin and cancer. *Nat. Rev. Cancer* **2004**, *4*, 133–142. [CrossRef]
68. Arnaud, O.; Le Loarer, F.; Tirode, F. BAFfling pathologies: Alterations of BAF complexes in cancer. *Cancer Lett.* **2018**, *419*, 266–279. [CrossRef]
69. Jo, V.Y.; Fletcher, C.D.M. SMARCB1/INI1 loss in Epithelioid Schwannoma. *Am. J. Surg. Pathol.* **2017**, *41*, 1013–1022. [CrossRef]

70. Johnson, J.O.; Mandrioli, J.; Benatar, M.; Abramzon, Y.; Van Deerlin, V.M.; Trojanowski, J.Q.; Gibbs, J.R.; Brunetti, M.; Gronka, S.; Wuu, J.; et al. Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* **2010**, *68*, 857–864. [CrossRef]
71. Weihl, C.C. Another VCP interactor: NF is enough. *J. Clin. Investig.* **2011**, *121*, 4627–4630. [CrossRef]
72. Wang, H.-F.; Shih, Y.-T.; Chen, C.-Y.; Chao, H.-W.; Lee, M.-J.; Hsueh, Y.-P. Valosin-containing protein and neurofibromin interact to regulate dendritic spine density. *J. Clin. Investig.* **2011**, *121*, 4820–4837. [CrossRef] [PubMed]
73. Hsueh, Y.-P.; Yang, F.-C.; Kharazia, V.; Naisbitt, S.; Cohen, A.R.; Weinberg, R.J.; Sheng, M. Direct interaction of CASK/LIN-2 and Syndecan Heparan sulfate Proteoglycan and their overlapping distribution in neuronal synapses. *J. Cell Boil.* **1998**, *142*, 139–151. [CrossRef] [PubMed]
74. Anney, R.J. Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Mol. Autism* **2017**, *8*, 21. [CrossRef]
75. Plasschaert, E.; Descheemaeker, M.-J.; Van Eylen, L.; Noens, I.; Steyaert, J.; Legius, E. Prevalence of Autism Spectrum Disorder symptoms in children with neurofibromatosis type 1. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2014**, *168*, 72–80. [CrossRef]
76. Zhang, W.; Lei, X.; Bian, C. Identifying cancer genes by combining two-rounds RWR based on multiple biological data. *BMC Bioinform.* **2019**, *20*, 1–12. [CrossRef]
77. Yoo, S.; Nam, H.; Lee, D.-J. Phenotype-oriented network analysis for discovering pharmacological effects of natural compounds. *Sci. Rep.* **2018**, *8*, 11667. [CrossRef]
78. Rieley, M.B.; Stevenson, D.A.; Viskochil, D.H.; Tinkle, B.T.; Martin, L.J.; Schorry, E. Variable expression of neurofibromatosis 1 in monozygotic twins. *Am. J. Med. Genet. Part A* **2011**, *155*, 478–485. [CrossRef]
79. Szudek, J.; Joe, H.; Friedman, J. Analysis of intrafamilial phenotypic variation in neurofibromatosis 1 (NF1). *Genet. Epidemiol.* **2002**, *23*, 150–164. [CrossRef]
80. Malone, J.R.; Stevens, R.; Jupp, S.; Hancocks, T.; Parkinson, H.; Brooksbank, C. Ten Simple Rules for Selecting a Bio-ontology. *PLoS Comput. Boil.* **2016**, *12*, e1004743. [CrossRef]

Cancer Genetics 240 (2020) 54–58

Contents lists available at ScienceDirect

Cancer Genetics

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

Short Communication

MIR605 rs2043556 is associated with the occurrence of multiple primary tumors in *TP53* p.(Arg337His) mutation carriers

Isabel Cristina Bandeira[a,b,1], Igor Araujo Vieira[a,b,1,*], Tiago Finger Andreis[a,b], Larissa Brussa Reis[a,b], Gabriel S. Macedo[b,c], Fernanda Sales Luiz Vianna[a,b], Patricia Santos-Silva[b], Edenir Inez Palmero[d,e], Henrique de Campos Reis Galvão[f], Cintia Regina Niederauer Ramos[d], Karina Miranda Santiago[g], Maria Isabel Achatz[g], Alexandre André Balieiro Anastácio da Costa[g], Patricia Ashton-Prolla[a,b,h,i]

[a] Post-Graduate Program in Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Rio Grande do Sul, Brazil
[b] Genomic Medicine Laboratory, Experimental Research Center, Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Rio Grande do Sul, Brazil
[c] Experimental Research Center, Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Rio Grande do Sul, Brazil
[d] Molecular Oncology Research Center, Barretos Cancer Hospital, Barretos, São Paulo, Brazil
[e] Barretos School of Health Sciences, Barretos, São Paulo, Brazil
[f] Oncogenetics Department, Barretos Cancer Hospital, Barretos, São Paulo, Brazil
[g] Oncogenetics Department, International Research Center, AC Camargo Cancer Center, National Institute of Oncogenomics, São Paulo city, São Paulo, Brazil
[h] Medical Genetics Service, HCPA, Porto Alegre, Rio Grande do Sul, Brazil
[i] Department of Genetics, UFRGS, Porto Alegre, Rio Grande do Sul, Brazil

ARTICLE INFO

Article history:
Received 9 August 2019
Revised 13 October 2019
Accepted 18 November 2019

Keywords:
Li-Fraumeni Syndrome
MIR605 gene
rs2043556
Multiple primary tumors

ABSTRACT

Li-Fraumeni and Li-Fraumeni-like (LFS/LFL) Syndrome are cancer predisposition syndromes caused by germline pathogenic variants in *TP53* and are associated with an increased risk of multiple early-onset cancers. In Southern and Southeastern Brazil, a germline founder variant with partial penetrance located in the oligomerization domain of *TP53*, c.1010G>A p.(Arg337His, commonly known as R337H), has been detected in 0.3% of the general population. Recently, the functional *MIR605* variant rs2043556 (*A*>*G*) has been identified as a novel LFS phenotype modifier in families with germline *TP53* DNA binding variants. In this study, our goal was to verify *MIR605* rs2043556 allele frequencies and further explore its possible effects on the phenotype of 238 Brazilian individuals carrying *TP53* p.(Arg337His). The *MIR605* rs2043556 G allele was detected in 136 (57.1%) individuals, including 25 homozygotes (10.5%), and although it had been previously associated with an earlier mean age of tumor onset, this effect was not observed in this study ($p = 0.8$). However, in p.(Arg337His) mutation carriers, the GG genotype was significantly associated with the occurrence of multiple primary tumors ($p = 0.005$). We provide further evidence of *MIR605* rs2043556 G allele's effect as a phenotype modulator in carriers of germline *TP53* pathogenic variants.

Introduction

Li-Fraumeni Syndrome (LFS) (OMIM # 151623) and its variant, Li-Fraumeni-like Syndrome (LFL), are autosomal dominant cancer predisposition syndromes characterized by a high risk for development of multiple tumors at a young age [1,2]. Germline pathogenic variants in the *TP53* gene are identified in approximately 70% and 40% of families that meet clinical criteria for LFS and LFL, respectively [3,4]. The protein product of *TP53*, p53, is a transcription factor whose ability to mediate tumor suppression has been extensively studied. p53 exerts its multiple antiproliferative functions through the transcriptional control of several target genes and through protein-protein interactions [5].

In southern and southeastern Brazil, the *TP53* founder variant c.1010G>A or p.(Arg337His, commonly known as R337H), located in the oligomerization domain (OD) of the gene, is detected in 0.3% of the general population. Although penetrance is reduced in relation to DNA binding domain (DBD) variants, its frequency is higher than what is observed for any other germline variant possibly present in the gene [6]. Currently, c.1010G>A is associated with a broad spectrum of tumors, similar to the one observed in

* Corresponding author.
E-mail address: igvieira@hcpa.edu.br (I.A. Vieira).
[1] Isabel C. Bandeira and Igor A. Vieira have contributed equally to this work and should be considered co-first authors.

https://doi.org/10.1016/j.cancergen.2019.11.005
2210-7762/

families with LFS/LFL caused by DBD variants [7–9]. Recent studies have prioritized epidemiological aspects of this mutation, while the mechanism of cancer predisposition associated with this founder variant and its penetrance is still not completely understood [8,9]. Despite existing evidence linking the LFS/LFL phenotype with differences in mutant p53 activity in terms of cellular localization and functional effect, these genotype-phenotype correlations do not fully explain the global and intra-familial heterogeneity observed in carriers of this Brazilian founder mutation [10,11].

In this context, several recent studies have explored the role of coexisting or secondary genetic factors that might modify p53 function. Among these are microRNAs (miRNAs), a class of 18- to 25-nucleotide-long single-stranded non-coding RNAs involved in the post-transcriptional regulation of gene expression [12,13]. These small molecules play a role in several biological processes, including cell proliferation, differentiation, apoptosis, and development, acting as oncogenes or tumor suppressors [14–16]. In addition, it has been shown that several miRNAs contribute to a refined p53 expression control by interacting directly with target sites on the 3′UTR of the *TP53* mRNA and, thus, they could be considered clinically relevant oncogenes. A few examples of indirect p53 regulation through a miRNA network have been described, including the effect of miR-605, which directly modulates *MDM2* expression, the main p53 activity negative regulator [17,18]. MiR-605 also appears to be inducible by p53 in response to cell stress mechanisms [19]. Furthermore, the presence of a *MIR605* gene variant, rs2043556 (*A*>*G*), has been associated with an increased risk for developing different cancers [16]. More recently, this same single nucleotide polymorphism (SNP) was identified as a functional variant and a novel genetic modifier of the LFS phenotype, specifically associated with an earlier mean age of tumor onset in Canadian families with *TP53* pathogenic DBD variants [20]. Hence, the objective of this study was to determine the allelic and genotypic frequencies as well as the phenotypic effect of *MIR605* SNP rs2043556 (*A*>*G*) in Brazilian LFS/LFL individuals carrying the germline founder variant p.(Arg337His).

Materials and methods

Study subjects and ethical aspects

Patients selected for the study were recruited from three tertiary care hospitals in southern and southeastern Brazil. They all fulfilled the Chompret criteria for LFS/LFL and were carriers of a germline *TP53* pathogenic variant, either p.(Arg337His) (main study group) or a DBD variant (comparison group). A total of 238 p.(Arg337His) carriers were recruited, 65 from Hospital de Clínicas de Porto Alegre (HCPA, in the city of Porto Alegre, southern Brazil), 66 from A.C. Camargo Cancer Center (city of São Paulo, southeastern Brazil), and 107 from Hospital de Câncer de Barretos (city of Barretos, southeastern Brazil). Six LFS patients carrying germline DBD variants from HCPA were recruited and included in a comparison group. Genetic analyses were previously approved by research ethics committees of the involved collaborating centers (registered under the Certificate of Presentation for Ethical Appreciation – CAAE n° 52641616.0.0000.5327).

MIR605 *SNP genotyping*

TaqMan® allelic discrimination analyses of variant rs2043556 were performed according to Applied Biosystems® standard protocols (Applied Biosystems, Carlsbad, USA), using fluorescent allele-specific probes (reference number C_11737438_10). *TP53* p.(Arg337His) genotyping was performed following previously published protocols using custom TaqMan® assays [21]. Sanger sequencing confirmed all samples with a variant identified by TaqMan®. Sequencing of *TP53* exon 10 encompassing p.(Arg337His) was performed according to IARC standard protocols (primer sequences and PCR conditions available at http://p53.iarc.fr/download/tp53_directsequencing_iarc.pdf), while *MIR605* was analyzed according to the primers described by Id Said & Malkin [20].

Statistical analysis

Genotype and allele frequencies were estimated by simple counting. Differences between groups were compared using the Kruskal-Wallis' (median age at tumor onset), Pearson's chi-squared (multiple primary tumors and cancer personal history) or Fisher's exact tests (tumor type). All tests were two-tailed, significance was set at p less than 0.05 and statistical analyses were done using SPSS® version 18 (SPSS Inc., Chicago, USA).

Results

Clinical data on LFS/LFL patients included in the main study group are summarized in Supplementary Table 1. Among the 238 p.(Arg337His) carriers, the variant G-allele for *MIR605* rs2043556 was detected in 136 individuals (57.1%), including 25 homozygotes (GG genotype, 10.5%) (Table 1). Moreover, the G-allele had a high frequency in this series (0.34), similar to the frequency observed in population databases, especially those related to Latin cohorts (Supplementary Table 2). Presence of the *MIR605* rs204356 G allele was not associated with age of tumor onset ($p = 0.8$, Fig. 1, Supplementary Table 3) or tumor type ($p = 0.3$) (Table 2) or personal history of cancer ($p = 0.1$, Supplementary Table 4). Genotypic frequency distribution was also not related with the median age of breast and adrenocortical cancers (two common tumor types found in LFS/LFL families, Supplementary Table 5). However, the GG genotype was significantly associated with the occurrence of multiple primary tumors ($p = 0.005$) (Table 2). For instance, among these *TP53* p.(Arg337His) and rs204356[GG]-positive patients with tumor diagnosis in multiple primary sites, one female patient developed very early-onset breast and thyroid cancers (at 23 and 25 years, respectively; data not shown). Of note, 26 of 122 cancer-affected patients (10.3%) developed more than one primary tumor (Table 2). Among these 26 patients, 4 (15.4%) had synchronous tumors; for the remaining metachronous tumors, the average time between diagnoses was approximately 6 years (Supplementary Table 6).

In addition, *MIR605* rs2043556 was genotyped in a small group of LFS patients with DBD mutations (comparison group, $N = 6$). Clinical data from these patients is summarized in Supplementary Table 7. The variant G-allele was detected at a frequency of 0.33, very similar to the one observed in the p.(Arg337His) group. Interestingly, one patient in the comparison group heterozygous for the *TP53* p.(Arg273His) variant (identifier LFS-4) was GG homozygous for the *MIR605* SNP and was diagnosed with multiple tumors

Table 1
Genotypic and allelic frequencies of *MIR605* rs2043556 (*A*>*G*) in *TP53* p.(Arg337His) mutation carriers with the LFS/LFL phenotype.

Genotype frequencies	N (%) $n = 238$
AA	102 (42.9%)
AG	111 (46.6%)
GG	25 (10.5%)
AG+GG	136 (57.1%)
Allele frequency	
G*	0.34

* Risk/variant allele.

Genotype SNP rs2043556	Median age in years (IR)[a]	p
AA	35.5 (31.75)	
AG	31.5 (41.75)	0.808[b]
GG	40 (36)	

[a] IR, interquartile range.
[b] Kruskal-Wallis test.

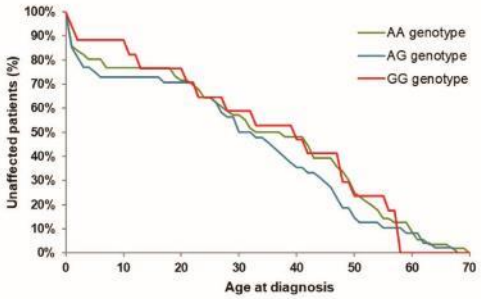


Fig. 1. Distribution of the median age of first cancer diagnosis according to *MIR605* rs2043556 genotype in *TP53* p.Arg337His carriers. The plot on the right shows the percentage of cancer-unaffected individuals harboring the founder variant for each *MIR605* genotype as a function of age of cancer onset.

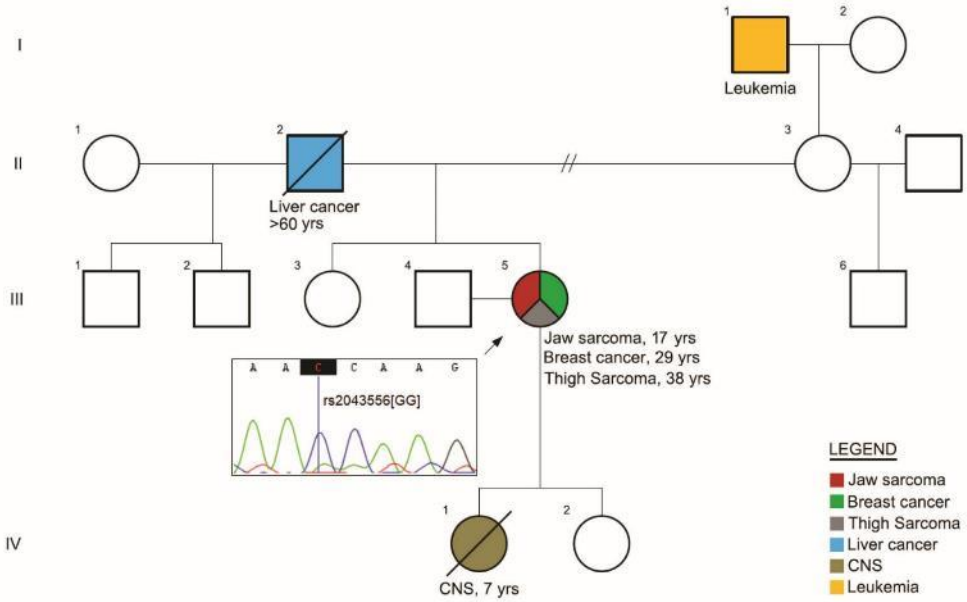


Fig. 2. Pedigree of an LFS proband (indicated by the black arrow; identifier LFS-4 in Supplementary Table 7) fulfilling classical clinical criteria and harboring the *TP53* DNA-binding domain germline variant p.(Arg273His), along with the *MIR605* variant G-allele in homozygosis (confirmed by Sanger sequencing as shown by the lower-left panel). Yrs, years; CNS, central nervous system tumor.

(sarcoma at age 17, breast cancer at age 29 and a new sarcoma at age 38 years). Fig. 2 shows the pedigree of this proband. To our knowledge, this is the first description of a patient with a DBD p53 pathogenic variant carrying the *MIR605* rs2043556 G-allele in homozygosis. Sanger sequencing confirmed the GG genotype in all cases identified by TaqMan® analyses (Supplementary Fig. 1).

Discussion

In a previous report, *MIR605* rs2043556 (*A*>*G*) was proposed to be a genetic modifier of the age of cancer onset in a Canadian cohort of LFS patients (classical LFS criteria) composed predominantly of *TP53* DBD mutation carriers [20]. Presence of the G-allele in heterozygosity was associated with a 10-year acceleration in the mean age of tumor onset ($p = 0.04$) in this LFS case series. Importantly, none of the participants in this previously mentioned study were carriers of the Brazilian founder mutation p.(Arg337His). We did not observe the same effect in our cohort, and although this could be due to the different genetic background between Canadian and Brazilian populations, the most important difference between both studies is the localization of the germline *TP53* mutations of their participants (DBD *vs.* OD). The modifying effect of *MIR605* rs2043556 G allele in heterozygosity on the age of cancer diagnosis previously observed in carriers of highly penetrant DBD mutations might be insufficient to cause a detectable phenotypic change in carriers of "milder", less penetrant *TP53* variants such as p.(Arg337His). In addition, it is worth emphasizing that a significant proportion of cancer diagnoses reported in p.(Arg337His) carriers occurs in adults, except for adrenocortical carcinomas and brain tumors, which occur more frequently in pediatric patients.

Table 2
Genotypic frequencies of *MIR605* rs2043556 in *TP53* p.Arg337His carriers according to tumor type and number of cancer diagnoses.

Genotype SNP rs2043556	Breast cancers, N (%)	Adrenocortical carcinomas, N (%)	CNS tumors[a], N (%)	Sarcomas and osteosarcomas[b], N (%)	Other tumor types[c], N (%)	*p*	Single tumor diagnosis, N (%)	Multiple primary tumors diagnosis, N (%)	*p*
AA	18 (40)	10 (38.5)	6 (66.7)	5 (33.3)	18 (64.3)		44 (45.8)	12 (46.1)	
AG	19 (42.2)	12 (46.1)	3 (33.3)	9 (60)	7 (25)	0.308[d]	44 (45.8)	6 (23.1)	**0.005**[e]
GG	8 (17.8)	4 (15.4)	0 (0)	1 (6.7)	3 (10.7)		8 (8.3)	8 (30.8)	

[a] CNS, central nervous system.
[b] Soft tissue sarcomas and osteosarcomas.
[c] Twenty-eight patients developed other tumor types, including prostate, thyroid, lung, kidney, uterine cancers, colorectal, gastric cancers, pheochromocytoma, melanoma, leukemia, lymphoma, and multiple myeloma.
[d] Fisher exact test.
[e] Pearson's chi-squared test.

A remarkable finding in the present study is a significant association between the homozygous GG genotype and the development of more than one primary tumor, suggesting that the G-allele might have a dose dependent effect associated with increased cancer predisposition, or might exert an influence on predisposition to a second primary tumor. This association between *MIR605* rs2043566 and multiple cancer diagnoses seems to be limited to *TP53* p.(Arg337His) carriers, since in the previous study ($N = 55$) [20] and in our small cohort ($N = 6$), among LFS patients with *TP53* DBD mutations, the frequency of multiple primary tumors did not differ between individuals harboring the variant G-allele and those with the AA genotype.

Furthermore, Id Said and Malkin (2015) [20] showed in their functional experiments that, in *MIR605* rs2043556 AG heterozygotes, the miR-605 processing efficiency from its precursor to its mature form was compromised, leading to reduced mature miR-605 levels. Therefore, it is reasonable to hypothesize that GG homozygotes could have dramatically impaired miRNA processing, triggering severe deregulation effects on MDM2 and/or p53 function. This effect would be profoundly deleterious in the presence of a germline *TP53* DBD mutation, but less so in the presence of a "milder" *TP53* OD mutation. In addition to MDM2 and p53, other validated individual target regulated by miR-605 includes Sec24D (miRTarBase database, http://mirtarbase.mbc.nctu.edu.tw/php/search.php#mirna), a cellular trafficking protein lacking a well-established role in carcinogenesis [22], as well as MAPK signaling pathway was shown to be induced by miR-605 overexpression, leading to tumor suppressive effects in *TP53* mutant cell lines according to the preliminary evidence reported by Malkin's group [23]. In agreement with this hypothesis, our study identified, for the first time, a considerable amount of *MIR605* rs2043556[GG] homozygotes meeting LFS/LFL clinical criteria and harboring the *TP53* OD mutation p.(Arg337His), differing from the results of the study by Id Said and Malkin (2015) in which no GG homozygotes were identified [20]. When considering the G allele frequencies described in population databases worldwide, the fact that none of the individuals in the Canadian study had the G-variant allele in homozygosis is striking. This difference in GG genotype frequency (0/55 *vs.* 25/258) between Canadian and Brazilian LFS/LFL cohorts exhibited a statistically significant difference, reinforcing that these findings are not random events ($P = 0.006$).

In conclusion, our results support previous studies in showing that the *MIR605* rs2043556 G allele is a potential phenotype modifier not only in LFS/LFL patients with DBD mutations but also in those with OD mutations, as demonstrated here in carriers of the exon 10 founder *TP53* variant p.(Arg337His). In these patients, we identified an association between the presence of the *MIR605* rs2043556 GG genotype and occurrence of multiple primary tumors. Larger studies including patients with different *TP53* germline mutations preferentially accompanied by functional analysis of the *MIR605* rs2043556 risk allele are required to confirm this hypothesis. Taken together, these findings emphasize the importance of analyzing miRNA genes that directly or indirectly regulate p53 expression as potential phenotype modifiers and as promising therapeutic targets in LFS/LFL.

Internet resources

IARC *TP53* Database. April 5, 2019. <http://p53.iarc.fr/Download/TP53_DirectSequencing_IARC.pdf>. miRTarBase Database. September 20, 2019. <http://mirtarbase.mbc.nctu.edu.tw/php/search.php#mirna>.

Declaration of Conflict Interest

H. C. R. G. has a disclosure with AstraZeneca Brazil (financial support for lectures) and Hermes Pardini Institute (Consultancy). Other authors do not have any conflict of interest to declare.

Acknowledgments

We would like to thank Gustavo Stumpf and Kendi Nishino Miyamoto for their valuable contributions and technical support. This study was funded by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Grant # 478430/2012-4), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS - Grant # 16/2551-0000486-2), and Fundo de Incentivo à Pesquisa do Hospital de Clínicas de Porto Alegre (FIPE-HCPA #16-0078) to Patricia Ashton-Prolla. This study was partially supported by FINEP-CT-INFRA (02/2010). EIP receives a National Council of Technological and Scientific Development (CNPq) scholarship.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cancergen.2019.11.005.

References

[1] Li FP, Fraumeni JF. Soft-tissue sarcomas, breast cancer, and other neoplasms. a familial syndrome? Ann Intern Med 1969;71:747–52. doi:10.7326/0003-4819-71-4-747.
[2] Malkin D, Li FP, Strong LC, Fraumeni JF, Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ, Tainsky MA. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 1990;250:1233–8. doi:10.1126/science.1978757.
[3] Varley JM, Evans DG, Birch JM. Li-Fraumeni syndrome-a molecular and clinical review. Br J Cancer 1997;76:1–14. doi:10.1038/bjc.1997.328.
[4] Malkin D. Li-fraumeni syndrome. Genes Cancer 2011;2:475–84. doi:10.1177/1947601911413466.
[5] Olivier M, Petitjean A, Marcel V, Pêtré A, Mounawar M, Plymoth A, de Fromentel CC, Hainaut P. Recent advances in p53 research: an interdisciplinary perspective. Cancer Gene Ther 2009;16:1–12. doi:10.1038/cgt.2008.69.
[6] Custódio G, Parise GA, Kiesel Filho N, Komechen H, Sabbaga CC, Rosati R, et al. Impact of neonatal screening and surveillance for the TP53 R337H mutation on early detection of childhood adrenocortical tumors. J Clin Oncol 2013;31(20):2619–26. doi:10.1200/JCO.2012.46.3711.
[7] Ribeiro RC, Sandrini F, Figueiredo B, Zambetti GP, Michalkiewicz E, Lafferty AR, DeLacerda L, Rabin M, Cadwell C, Sampaio G, Cat I, Stratakis CA, Sandrini R. An inherited p53 mutation that contributes in a tissue-specific manner to pediatric adrenal cortical carcinoma. Proc Natl Acad Sci USA 2001;98(16):9330–5. doi:10.1073/pnas.161479898.
[8] Achatz MI, Olivier M, Le Calvez F, Martel-Planche G, Lopes A, Rossi BM, Ashton-Prolla P, Giugliani R, Palmero EI, Vargas FR, Da Rocha JC, Vettore AL, Hainaut P. The TP53 mutation, R337H, is associated with li-fraumeni and li-fraumeni-like syndromes in brazilian families. Cancer Lett 2007;245(1-2):96–102. doi:10.1016/j.canlet.2005.12.039.
[9] Giacomazzi J, Selistre SG, Rossi C, Alemar B, Santos-Silva P, Pereira FS, Netto CB, Cossio SL, Roth DE, Brunetto AL, Zagonel-Oliveira M, Martel-Planche G, Goldim JR, Hainaut P, Camey SA, Ashton-Prolla P. Li-Fraumeni and li-fraumeni-like syndrome among children diagnosed with pediatric cancer in southern brazil. Cancer 2013;119(24):4341–9. doi:10.1002/cncr.28346.
[10] Olivier M, Goldgar DE, Sodha N, Ohgaki H, Kleihues P, Hainaut P, Eeles RA. Li-Fraumeni and related syndromes: correlation between tumor type, family structure, and TP53 genotype. Cancer Res 2003;63:6643–50 PMID: 14583457.
[11] Bougeard G, Sesboüé R, Baert-Desurmont S, Vasseur S, Martin C, Tinat J, Brugières L, Chompret A, de Paillerets BB, Stoppa-Lyonnet D, Bonaïti-Pellié C, Frébourg TFrench LFS working group. Molecular basis of the li-fraumeni syndrome: an update from the french lfs families. J Med Genet 2008;45:535–8. doi:10.1136/jmg.2008.057570.
[12] Ambros V. The functions of animal microRNAs. Nature 2004;431:350–5. doi:10.1038/nature02871.
[13] Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281–97. doi:10.1016/s0092-8674(04)00045-5.
[14] Ambros V. MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. Cell 2003;113:673–6. doi:10.1016/s0092-8674(03)00428-8.
[15] Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, Horvitz HR, Golub TR. MicroRNA expression profiles classify human cancers. Nature 2005;435(7043):834–8. doi:10.1038/nature03702.
[16] Hu Y, Yu CY, Wang JL, Guan J, Chen HY, Fang JY. MicroRNA sequence polymorphisms and the risk of different types of cancer. Sci Rep 2014;4:3648. doi:10.1038/srep03648.
[17] Hermeking H. MicroRNAs in the p53 network: micromanagement of tumour suppression. Nat Rev Cancer 2012;12:613–26. doi:10.1038/nrc3318.
[18] Li Y, Gordon MW, Xu-Monette ZY, Visco C, Tzankov A, Zou D, et al. Single nucleotide variation in the TP53 3′ untranslated region in diffuse large B-cell lymphoma treated with rituximab-CHOP: a report from the international dlbcl rituximab-chop consortium program. Blood 2013;121:4529–40. doi:10.1182/blood-2012-12-471722.
[19] Xiao J, Lin H, Luo X, et al. miR-605 joins p53 network to form a p53:miR-605:Mdm2 positive feedback loop in response to stress. EMBO J 2011;30:524–32. doi:10.1038/emboj.2010.347.
[20] Id Said B, Malkin D. A functional variant in miR-605 modifies the age of onset in li-fraumeni syndrome. Cancer Genet 2015;208(1-2):47–51. doi:10.1016/j.cancergen.2014.12.003.
[21] Fitarelli-Kiehl M, Macedo GS, Schlatter RP, et al. Comparison of multiple genotyping methods for the identification of the cancer predisposing founder mutation p.R337H in TP53. Genet Mol Biol 2016;39(2):203–9. doi:10.1590/1678-4685-GMB-2014-0351.
[22] Lee I, Ajay SS, Yook JI, et al. New class of microRNA targets containing simultaneous 5′-UTR and 3′-UTR interaction sites. Genome Res 2009;19(7):1175–83. doi:10.1101/gr.089367.108.
[23] Id Said B. The role of miR-605 and its variant in li-fraumeni syndrome. https://tspace.library.utoronto.ca/bitstream/1807/44027/1/Idsaid_Badr_201404_MSc_thesis.pdf. Thesis submitted for MSc. Degree; 2014 [accessed 20 September 2019].

Resumos publicados nas Semanas Científicas do HCPA

38ª SEMANA CIENTÍFICA DO HOSPITAL DE CLÍNICAS DE PORTO ALGREE

eP2735 **"*Analysis of predictive biomarkers in patients with lung adenocarcinoma from southern Brazil*"**

Tiago Finger Andreis; Bruno da Silveira Corrêa; Fernanda Salles Vianna; Fernanda de Paris; Marina Siebert; Sandra Leistner-Segal; Luis Fernando da Rosa Rivero; Francine Hehn de Oliveira; Patricia Ashton-Prolla; Gabriel de Souza Macedo

P1662 **"*Busca por novos biomarcadores preditivos de resposta à imunoterapia no câncer de pulmão: explorando dados do The Cancer Genome Atlas em tumores hipermutados*"**

Tiago Finger Andreis, Tiago Falcon Lopes, Ivaine Taís Sauthier Sartor, Patricia Ashton-Prolla, Gabriel de Souza Macedo

P1839 **"*Procura por novas assinaturas moleculares no adenocarcinoma de pulmão: análise da expressão diferencial em tabagistas e indivíduos não-tabagistas*"**

Bruno da Silveira Corrêa, Tiago Finger Andreis, Ivaine Taís Sauthier Sartor, Patricia Ashton-Prolla, Gabriel de Souza Macedo

P2041 **"*Epidemiologia molecular do adenocarcinoma de pulmão no sul do Brasil: análise dos primeiros 463 casos*"**

Julia Tsao Schein, Tiago Finger Andreis, Fernanda Sales Luiz Vianna, Patrícia Ashton-Prolla, Gabriel de Souza Macedo

39ª SEMANA CIENTÍFICA DO HOSPITAL DE CLÍNICAS DE PORTO ALEGRE

eP2264 **"*Prevalência da variante fundadora em TP53 C.1010G>A (P.ARG337HIS) em pacientes com adenocarcinoma de pulmão no sul do Brasil: é necessária uma nova política de rastreamento?*"**

Tiago Finger Andreis; Igor Araujo Vieira; Bruna Vieira Fernandes; Daniel Schramek; Gabriel de Souza Macedo; Patricia Ashton-Prolla

eP2284 **"*Prevalência de uma variante rara localizada no sinal de poliadenilação de TP53 em tecido tumoral de pacientes brasileiros diagnosticados com adenocarcinoma de pulmão*"**

Bruna Vieira Fernandes; Igor Araújo Vieira; Tiago Finger Andreis; Eduarda Heidrich Pezzi; Bruno da Silveira Corrêa; Guilherme Danielski Viola; Fernanda de Paris; Fernanda Sales Luiz Vianna; Gabriel de Souza Macedo; Patricia Ashton-Prolla

eP3107 **"*Caracterização do perfil demográfico, clínico e molecular de uma série de usuários do SUS com câncer de pulmão no sul do Brasil*"**

Débora Leite Rocha; Tiago Finger Andreis; Nícolas Peruzzo; Guilherme Geib; Gabriel de Souza Macedo; Patricia Ashton-Prolla

eP2497 **"*Desafios à eficácia do sequenciamento de nova geração em amostras tumorais*"**

Bruno da Silveira Corrêa; Fernanda de Paris; Tiago Finger Andreis; Fernanda Salles Vianna; Luis Fernando da Rosa Rivero; Francine Hehn de Oliveira; Patricia Ashton-Prolla; Gabriel de Souza Macedo

P1451 **"*Potential association between the functional variant RS2043556 in MIR605 gene and development of multiple primary tumors in TP53 P.ARG337HIS mutation carriers*"**

Igor Araujo Vieira, Isabel Cristina Bandeira da Silva, Larissa Brussa Reis, Tiago Finger Andreis, Edenir Inez Palmero, Karina Miranda Santiago, Maria Isabel Waddington Achatz, Patricia Ashton-Prolla

40ª SEMANA CIENTÍFICA DO HOSPITAL DE CLÍNICAS DE PORTO ALEGRE

2208 "**Variantes de regiões promotora e 5' UTR nos genes BRCA1/2: associação com metilação, importância no diagnóstico de HBOC e frequência na população brasileira"**

Kayana Isabel Weber De Souza; Tiago Finger Andreis; Patricia Ashton-Prolla; Clévia Rosset

2857 **"*Reclassificação de variantes de sentido trocado no gene BRCA1 associadas ao câncer de mama e ovário hereditários*"**

Bruno Da Silveira Corrêa; Igor Araujo Vieira; Tiago Finger Andreis; Guilherme Danielski Viola; Patricia Ashton-Prolla; Gabriel De Souza Macedo

2176 **"*Prevalência de uma variante funcional na região 3'UTR do gene TP53 em pacientes do sul do brasil diagnosticados com diferentes tumores esporádicos*"**

Eduarda Heidrich Pezzi; Igor Araujo Vieira; Bruna Vieira Fernandes; Tiago Finger Andreis; Luiza Mezzomo Donatti; Gabriela Dos Santos Sant'anna; Helena Von Eye Corleta; Ilma Simoni Brum; Gabriel De Souza Macedo; Patricia Ashton-Prolla

2505 **"*Biomarcadores De Resposta A Inibidores De Checkpoint Em Adenocarcinoma De Pulmão: Uma Análise Do The Cancer Genome Atlas*"**

Guilherme Danielski Viola; Tiago Finger Andreis; Patricia Ashton-Prolla; Gabriel De Souza Macedo

2420 **"*Systems biology approaches reveal potential phenotype modifier genes in neurofibromatosis type 1*"**

Larissa Brussa Reis; Thayne Woycinck Kowalski; Tiago Finger Andreis; Patricia Ashton-Prolla; Clévia Rosset

41ª SEMANA CIENTÍFICA DO HOSPITAL DE CLÍNICAS DE PORTO ALEGRE

2186 **"*Estratégias de bioinformática para selecionar variantes genéticas do hospedeiro humano que influenciam nos desfechos clínicos da COVID-19*"**

Nathan Araujo Cadore, Tiago Finger Andreis, Martiela Vaz de Freitas, Marilea Furtado Feira, Igor Araujo Vieira, Clevia Rosset, Kendi Nishino Miyamoto, Bruno Cesar Feltes, Thayne Woycinck Kowalski, Fernanda Sales Luiz Vianna

2475 **"*Reassessment of BRCA1/2 variants of unknown significance: an analysis using different classification guidelines*"**

Kayana Isabel Weber de Souza, Tiago Finger Andreis, Barbara Alemar Beserra, Osvaldo Artigalás, Camila Matzenbacher Bittar, Cristina Brinckmann Oliveira Netto, Patricia Ashton-prolla, Clevia Rosset

2258 **"*Searching for biomarkers to immune checkpoint inhibitors treatment response in lung adenocarcinoma: a The Cancer Genome Atlas Exploratory Analysis*"**

Guilherme Danielski Viola, Tiago Finger Andreis, Patricia Ashton-prolla, Gabriel de Souza

2387 **"*HLA frequencies and differential regional incidence of COVID-19 in Brazil: a population study based in a large bone marrow donors bank dataset*"**

Juliano André Boquett, Fernanda Sales Luiz Vianna, Nelson Jurandi Rosa Fagundes, Lucas Schroeder, Márcia Helena Barbian, Marcelo Zagonel-oliveira, Tiago Finger Andreis, Luis Cristóvão de Moraes Sobrino Pôrto, José Artur Bogo Chies, Lavínia Schüler-faccini, Patricia Ashton-prolla, Clevia Rosset

2669 **"*Alterações moleculares de significando clínico incerto em pacientes Com câncer de pulmão de não pequenas células: decisão terapêutica e Resposta a inibidores de tirosina-quinase*"**

Tiago Finger Andreis, Gabriel de Souza Macedo, Patricia Ashton-Prolla

Coorientação de Trabalho de Conclusão de Curso:

GUILHERME DANIELSKI VIOLA

BUSCA POR NOVOS BIOMARCADORES PREDITIVOS DE RESPOSTA A INIBIDORES DE CHECKPOINT NO CÂNCER DE PULMÃO: EXPLORANDO DADOS DO THE CANCER GENOME ATLAS EM TUMORES HIPERMUTADOS

Trabalho de conclusão de curso apresentado como requisito parcial para obtenção de título de bacharel em Ciências Biológicas pela Universidade Federal do Rio Grande do Sul

Orientadora: Prof. Patrícia Ashton-Prolla

Co-orientador: Tiago Finger Andreis

Porto Alegre

2018

REFERÊNCIAS BIBLIOGRÁFICAS

(CLCGP), C. L. C. G. P. and (NGM), N. G. M. (2013) 'A genomics-based classification of human lung tumors', *Sci Transl Med,* 5(209), pp. 209ra153.

Abbosh, C., Birkbak, N. J. and Swanton, C. (2018) 'Early stage NSCLC - challenges to implementing ctDNA-based screening and MRD detection', *Nat Rev Clin Oncol,* 15(9), pp. 577-586.

Abbosh, C. and Birkbak, N. J. and Wilson, G. A. and Jamal-Hanjani, M. and Constantin, T. and Salari, R. and Le Quesne, J. and Moore, D. A. and Veeriah, S. and Rosenthal, R. and Marafioti, T. and Kirkizlar, E. and Watkins, T. B. K. and McGranahan, N. and Ward, S. and Martinson, L. and Riley, J. and Fraioli, F. and Al Bakir, M. and Grönroos, E. and Zambrana, F. and Endozo, R. and Bi, W. L. and Fennessy, F. M. and Sponer, N. and Johnson, D. and Laycock, J. and Shafi, S. and Czyzewska-Khan, J. and Rowan, A. and Chambers, T. and Matthews, N. and Turajlic, S. and Hiley, C. and Lee, S. M. and Forster, M. D. and Ahmad, T. and Falzon, M. and Borg, E. and Lawrence, D. and Hayward, M. and Kolvekar, S. and Panagiotopoulos, N. and Janes, S. M. and Thakrar, R. and Ahmed, A. and Blackhall, F. and Summers, Y. and Hafez, D. and Naik, A. and Ganguly, A. and Kareht, S. and Shah, R. and Joseph, L. and Marie Quinn, A. and Crosbie, P. A. and Naidu, B. and Middleton, G. and Langman, G. and Trotter, S. and Nicolson, M. and Remmen, H. and Kerr, K. and Chetty, M. and Gomersall, L. and Fennell, D. A. and Nakas, A. and Rathinam, S. and Anand, G. and Khan, S. and Russell, P. and Ezhil, V. and Ismail, B. and Irvin-Sellers, M. and Prakash, V. and Lester, J. F. and Kornaszewska, M. and Attanoos, R. and Adams, H. and Davies, H. and Oukrif, D. and Akarca, A. U. and Hartley, J. A. and Lowe, H. L. and Lock, S. and Iles, N. and Bell, H. and Ngai, Y. and Elgar, G. and Szallasi, Z. and Schwarz, R. F. and Herrero, J. and Stewart, A. and Quezada, S. A. and Peggs, K. S. and Van Loo, P. and Dive, C. and Lin, C. J. and Rabinowitz, M. and Aerts, H. J. W. L. and Hackshaw, A. and Shaw, J. A. and Zimmermann, B. G. and Swanton, C. and consortium, T. and consortium, P. (2017) 'Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution', *Nature,* 545(7655), pp. 446-451.

Adjei, A. A. (2005) 'The role of mitogen-activated ERK-kinase inhibitors in lung cancer therapy', *Clin Lung Cancer,* 7(3), pp. 221-3.

Ai, X., Guo, X., Wang, J., Stancu, A. L., Joslin, P. M. N., Zhang, D. and Zhu, S. (2018) 'Targeted therapies for advanced non-small cell lung cancer', *Oncotarget,* 9(101), pp. 37589-37607.

Alexandrov, L. B. and Stratton, M. R. (2014) 'Mutational signatures: the patterns of somatic mutations hidden in cancer genomes', *Curr Opin Genet Dev,* 24, pp. 52-60.

Alvarez, J. G. B. and Otterson, G. A. (2019) 'Agents to treat', *Drugs Context,* 8, pp. 212566.

Alves da Silva, A. V., Martins Neto, F., de Oliveira, A. C. d. S. M., Carneiro, B. A., Alves, M., Nogueira, C. D. and Tavora, F. (2019) 'The frequency of high PD-L1 expression is low in lung adenocarcinoma patients from Northeast Brazil', *Surgical and Experimental Pathology,* 2(1), pp. 4.

Andreis, T. F., Correa, B. S., Vianna, F. S., De-Paris, F., Siebert, M., Leistner-Segal, S., Hahn, E. C., Ulbrich, J. M., Rivero, L. F. R., De Oliveira, F. H., Lorandi, V., Ashton-Prolla, P. and Macedo, G. S. (2019) 'Analysis of Predictive Biomarkers in Patients

With Lung Adenocarcinoma From Southern Brazil Reveals a Distinct Profile From Other Regions of the Country', *J Glob Oncol,* 5, pp. 1-9.

Aravanis, A. M., Lee, M. and Klausner, R. D. (2017) 'Next-Generation Sequencing of Circulating Tumor DNA for Early Cancer Detection', *Cell,* 168(4), pp. 571-574.

Arora, A. and Scholar, E. M. (2005) 'Role of tyrosine kinase inhibitors in cancer therapy', *J Pharmacol Exp Ther,* 315(3), pp. 971-9.

Arrieta, O., Cardona, A. F., Bramuglia, G., Cruz-Rico, G., Corrales, L., Martín, C., Imaz-Olguín, V., Castillo, O., Cuello, M., Rojas-Bilbao, É., Casas, G., Fernández, C., Arén Frontera, O., Denninghoff, V., Recondo, G., Avilés-Salas, A., Mas-Lopez, L. A., Oblitas, G., Rojas, L., Piottante, A., Jiménez-García, E., Sánchez-Sosa, S., Sáenz-Frias, J., Lupera, H., Ramírez-Tirado, L. A., Vargas, C., Carranza, H., Astudillo, H., Wills, L. B., Pichelbaur, E., Raez, L. E. and on behalf of the CLICaP (2019) 'Molecular Epidemiology of ALK Rearrangements in Advanced Lung Adenocarcinoma in Latin America', *Oncology,* 96(4), pp. 207-216.

Arrieta, O., Ramírez-Tirado, L. A., Báez-Saldaña, R., Peña-Curiel, O., Soca-Chafre, G. and Macedo-Perez, E. O. (2015) 'Different mutation profiles and clinical characteristics among Hispanic patients with non-small cell lung cancer could explain the "Hispanic paradox"', *Lung Cancer,* 90(2), pp. 161-6.

Bacchi, C. E., Ciol, H., Queiroga, E. M., Benine, L. C., Silva, L. H. and Ojopi, E. B. (2012) 'Epidermal growth factor receptor and KRAS mutations in Brazilian lung cancer patients', *Clinics (Sao Paulo),* 67(5), pp. 419-24.

Benusiglio, P. R., Fallet, V., Sanchis-Borja, M., Coulet, F. and Cadranel, J. (2021) 'Lung cancer is also a hereditary disease', *Eur Respir Rev,* 30(162).

Bettegowda, C., Sausen, M., Leary, R. J., Kinde, I., Wang, Y., Agrawal, N., Bartlett, B. R., Wang, H., Luber, B., Alani, R. M., Antonarakis, E. S., Azad, N. S., Bardelli, A., Brem, H., Cameron, J. L., Lee, C. C., Fecher, L. A., Gallia, G. L., Gibbs, P., Le, D., Giuntoli, R. L., Goggins, M., Hogarty, M. D., Holdhoff, M., Hong, S. M., Jiao, Y., Juhl, H. H., Kim, J. J., Siravegna, G., Laheru, D. A., Lauricella, C., Lim, M., Lipson, E. J., Marie, S. K., Netto, G. J., Oliner, K. S., Olivi, A., Olsson, L., Riggins, G. J., Sartore-Bianchi, A., Schmidt, K., Shih, I., Oba-Shinjo, S. M., Siena, S., Theodorescu, D., Tie, J., Harkins, T. T., Veronese, S., Wang, T. L., Weingart, J. D., Wolfgang, C. L., Wood, L. D., Xing, D., Hruban, R. H., Wu, J., Allen, P. J., Schmidt, C. M., Choti, M. A., Velculescu, V. E., Kinzler, K. W., Vogelstein, B., Papadopoulos, N. and Diaz, L. A. (2014) 'Detection of circulating tumor DNA in early- and late-stage human malignancies', *Sci Transl Med,* 6(224), pp. 224ra24.

Blair, H. A. (2021) 'Sotorasib: First Approval', *Drugs,* 81(13), pp. 1573-1579.

Blakely, C. M., Watkins, T. B. K., Wu, W., Gini, B., Chabon, J. J., McCoach, C. E., McGranahan, N., Wilson, G. A., Birkbak, N. J., Olivas, V. R., Rotow, J., Maynard, A., Wang, V., Gubens, M. A., Banks, K. C., Lanman, R. B., Caulin, A. F., St John, J., Cordero, A. R., Giannikopoulos, P., Simmons, A. D., Mack, P. C., Gandara, D. R., Husain, H., Doebele, R. C., Riess, J. W., Diehn, M., Swanton, C. and Bivona, T. G. (2017) 'Evolution and clinical impact of co-occurring genetic alterations in advanced-stage EGFR-mutant lung cancers', *Nat Genet,* 49(12), pp. 1693-1704.

Blons, H., Garinet, S., Laurent-Puig, P. and Oudart, J. B. (2019) 'Molecular markers and prediction of response to immunotherapy in non-small cell lung cancer, an update', *J Thorac Dis,* 11(Suppl 1), pp. S25-S36.

Bracken-Clarke, D., Kapoor, D., Baird, A. M., Buchanan, P. J., Gately, K., Cuffe, S. and Finn, S. P. (2021) 'Vaping and lung cancer - A review of current data and recommendations', *Lung Cancer,* 153, pp. 11-20.

Brasil (2019) *Estimativa 2020: Incidência de Câncer no Brasil.* Rio de Janeiro - RJ, Brazil: Ministério da Saúde, Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA), Coordenação de Prevenção e Vigilância, p. 122.

Brasil (2022) *Lumakras (sotorasibe): novo registro*. Available at: https://www.gov.br/anvisa/pt-br/assuntos/medicamentos/novos-medicamentos-e-indicacoes/lumakras-sotorasibe-novo-registro (Accessed: 6/11/2022 2022).

Brasil, M. d. S. 2013a. Erlotinibe para câncer de pulmão de células não pequenas. *In:* CONITEC (ed.) *Relatório de Recomendação da Comissão Nacional de Incorporação de Tecnologias no SUS – CONITEC – 63.*

Brasil, M. d. S. 2013b. Gefitinibe para câncer de pulmão de células não pequenas em primeira linha. *Relatório de Recomendação da Comissão Nacional de Incorporação de Tecnologias no SUS – CONITEC – 62.* CONITEC.

Brazel, D., Zhang, S. and Nagasaka, M. (2022) 'Spotlight on Tepotinib and Capmatinib for Non-Small Cell Lung Cancer with MET Exon 14 Skipping Mutation', *Lung Cancer (Auckl),* 13, pp. 33-45.

Brierley, J. D., Gospodarowicz, M. K. and Wittekind, C. (2017) *TNM Classification of Malignant Tumours, 8th Edition.* Oxford, UK ; Hoboken, NJ: John Wiley & Sons, Inc.

Bronte, G., Bravaccini, S., Bronte, E., Burgio, M. A., Rolfo, C., Delmonte, A. and Crinò, L. (2018) 'Epithelial-to-mesenchymal transition in the context of epidermal growth factor receptor inhibition in non-small-cell lung cancer', *Biol Rev Camb Philos Soc,* 93(4), pp. 1735-1746.

Chabon, J. J., Hamilton, E. G., Kurtz, D. M., Esfahani, M. S., Moding, E. J., Stehr, H., Schroers-Martin, J., Nabet, B. Y., Chen, B., Chaudhuri, A. A., Liu, C. L., Hui, A. B., Jin, M. C., Azad, T. D., Almanza, D., Jeon, Y. J., Nesselbush, M. C., Co Ting Keh, L., Bonilla, R. F., Yoo, C. H., Ko, R. B., Chen, E. L., Merriott, D. J., Massion, P. P., Mansfield, A. S., Jen, J., Ren, H. Z., Lin, S. H., Costantino, C. L., Burr, R., Tibshirani, R., Gambhir, S. S., Berry, G. J., Jensen, K. C., West, R. B., Neal, J. W., Wakelee, H. A., Loo, B. W., Kunder, C. A., Leung, A. N., Lui, N. S., Berry, M. F., Shrager, J. B., Nair, V. S., Haber, D. A., Sequist, L. V., Alizadeh, A. A. and Diehn, M. (2020) 'Integrating genomic features for non-invasive early lung cancer detection', *Nature,* 580(7802), pp. 245-251.

Chabon, J. J., Simmons, A. D., Lovejoy, A. F., Esfahani, M. S., Newman, A. M., Haringsma, H. J., Kurtz, D. M., Stehr, H., Scherer, F., Karlovich, C. A., Harding, T. C., Durkin, K. A., Otterson, G. A., Purcell, W. T., Camidge, D. R., Goldman, J. W., Sequist, L. V., Piotrowska, Z., Wakelee, H. A., Neal, J. W., Alizadeh, A. A. and Diehn, M. (2016) 'Circulating tumour DNA profiling reveals heterogeneity of EGFR inhibitor resistance mechanisms in lung cancer patients', *Nat Commun,* 7, pp. 11815.

Chan, H. T., Chin, Y. M., Nakamura, Y. and Low, S. K. (2020) 'Clonal Hematopoiesis in Liquid Biopsy: From Biological Noise to Valuable Clinical Implications', *Cancers (Basel),* 12(8).

Chapman, A. M., Sun, K. Y., Ruestow, P., Cowan, D. M. and Madl, A. K. (2016) 'Lung cancer mutation profile of EGFR, ALK, and KRAS: Meta-analysis and comparison of never and ever smokers', *Lung Cancer,* 102, pp. 122-134.

Cheng, M., Akalestos, A. and Scudder, S. (2020) 'Budget Impact Analysis of', *Diagnostics (Basel),* 10(6).

Chin, R. I., Chen, K., Usmani, A., Chua, C., Harris, P. K., Binkley, M. S., Azad, T. D., Dudley, J. C. and Chaudhuri, A. A. (2019) 'Detection of Solid Tumor Molecular Residual Disease (MRD) Using Circulating Tumor DNA (ctDNA)', *Mol Diagn Ther,* 23(3), pp. 311-331.

Chiosea, S., Shuai, Y., Cieply, K., Nikiforova, M. N. and Dacic, S. (2010) 'EGFR fluorescence in situ hybridization-positive lung adenocarcinoma: incidence of coexisting KRAS and BRAF mutations', *Hum Pathol,* 41(8), pp. 1053-60.

Chouaid, C., Dujon, C., Do, P., Monnet, I., Madroszyk, A., Le Caer, H., Auliac, J. B., Berard, H., Thomas, P., Lena, H., Robinet, G., Baize, N., Bizieux-Thaminy, A., Fraboulet, G., Locher, C., Le Treut, J., Hominal, S. and Vergnenegre, A. (2014) 'Feasibility and clinical impact of re-biopsy in advanced non small-cell lung cancer: a prospective multicenter study in a real-world setting (GFPC study 12-01)', *Lung Cancer,* 86(2), pp. 170-3.

Ciardiello, F., Arnold, D., Casali, P. G., Cervantes, A., Douillard, J. Y., Eggermont, A., Eniu, A., McGregor, K., Peters, S., Piccart, M., Popescu, R., Van Cutsem, E., Zielinski, C. and Stahel, R. (2014) 'Delivering precision medicine in oncology today and in future-the promise and challenges of personalised cancer medicine: a position paper by the European Society for Medical Oncology (ESMO)', *Ann Oncol,* 25(9), pp. 1673-1678.

CONITEC 2013. Gefitinibe para câncer de pulmão de células não pequenas em primeira linha.

CONITEC 2015. Erlotinibe para câncer de pulmão de células não pequenas.

Consortium, A. P. G. (2017) 'AACR Project GENIE: Powering Precision Medicine through an International Consortium', *Cancer Discov,* 7(8), pp. 818-831.

Dagogo-Jack, I. and Shaw, A. T. (2018) 'Tumour heterogeneity and resistance to cancer therapies', *Nat Rev Clin Oncol,* 15(2), pp. 81-94.

Davidson, M. R., Gazdar, A. F. and Clarke, B. E. (2013) 'The pivotal role of pathology in the management of lung cancer', *J Thorac Dis,* 5 Suppl 5, pp. S463-78.

de Melo, A. C., Karen de Sá, V., Sternberg, C., Olivieri, E. R., Werneck da Cunha, I., Fabro, A. T., Carraro, D. M., de Barros e Silva, M. J., Pimenta Inada, H. K., de Mello, E. S., Soares, F. A., Takagaki, T., Ferreira, C. G. and Capelozzi, V. L. (2015) 'Mutational Profile and New IASLC/ATS/ERS Classification Provide Additional Prognostic Information about Lung Adenocarcinoma: A Study of 125 Patients from Brazil', *Oncology,* 89(3), pp. 175-86.

Dela Cruz, C. S., Tanoue, L. T. and Matthay, R. A. (2011) 'Lung cancer: epidemiology, etiology, and prevention', *Clin Chest Med,* 32(4), pp. 605-44.

Deveson, I. W., Gong, B., Lai, K., LoCoco, J. S., Richmond, T. A., Schageman, J., Zhang, Z., Novoradovskaya, N., Willey, J. C., Jones, W., Kusko, R., Chen, G., Madala, B.

S., Blackburn, J., Stevanovski, I., Bhandari, A., Close, D., Conroy, J., Hubank, M., Marella, N., Mieczkowski, P. A., Qiu, F., Sebra, R., Stetson, D., Sun, L., Szankasi, P., Tan, H., Tang, L. Y., Arib, H., Best, H., Burgher, B., Bushel, P. R., Casey, F., Cawley, S., Chang, C. J., Choi, J., Dinis, J., Duncan, D., Eterovic, A. K., Feng, L., Ghosal, A., Giorda, K., Glenn, S., Happe, S., Haseley, N., Horvath, K., Hung, L. Y., Jarosz, M., Kushwaha, G., Li, D., Li, Q. Z., Li, Z., Liu, L. C., Liu, Z., Ma, C., Mason, C. E., Megherbi, D. B., Morrison, T., Pabón-Peña, C., Pirooznia, M., Proszek, P. Z., Raymond, A., Rindler, P., Ringler, R., Scherer, A., Shaknovich, R., Shi, T., Smith, M., Song, P., Strahl, M., Thodima, V. J., Tom, N., Verma, S., Wang, J., Wu, L., Xiao, W., Xu, C., Yang, M., Zhang, G., Zhang, S., Zhang, Y., Shi, L., Tong, W., Johann, D. J., Mercer, T. R., Xu, J. and Group, S. O. S. W. (2021) 'Evaluating the analytical validity of circulating tumor DNA sequencing assays for precision oncology', *Nat Biotechnol,* 39(9), pp. 1115-1128.

Dogan, S., Shen, R., Ang, D. C., Johnson, M. L., D'Angelo, S. P., Paik, P. K., Brzostowski, E. B., Riely, G. J., Kris, M. G., Zakowski, M. F. and Ladanyi, M. (2012) 'Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: higher susceptibility of women to smoking-related KRAS-mutant cancers', *Clin Cancer Res,* 18(22), pp. 6169-77.

Du, Z. and Lovly, C. M. (2018) 'Mechanisms of receptor tyrosine kinase activation in cancer', *Mol Cancer,* 17(1), pp. 58.

El Messaoudi, S., Rolet, F., Mouliere, F. and Thierry, A. R. (2013) 'Circulating cell free DNA: Preanalytical considerations', *Clin Chim Acta,* 424, pp. 222-30.

Engelman, J. A., Zejnullahu, K., Gale, C. M., Lifshits, E., Gonzales, A. J., Shimamura, T., Zhao, F., Vincent, P. W., Naumov, G. N., Bradner, J. E., Althaus, I. W., Gandhi, L., Shapiro, G. I., Nelson, J. M., Heymach, J. V., Meyerson, M., Wong, K. K. and Jänne, P. A. (2007) 'PF00299804, an irreversible pan-ERBB inhibitor, is effective in lung cancer models with EGFR and ERBB2 mutations that are resistant to gefitinib', *Cancer Res,* 67(24), pp. 11924-32.

Fassina, A., Gazziero, A., Zardo, D., Corradin, M., Aldighieri, E. and Rossi, G. P. (2009) 'Detection of EGFR and KRAS mutations on trans-thoracic needle aspiration of lung nodules by high resolution melting analysis', *J Clin Pathol,* 62(12), pp. 1096-102.

Ferreira, C. G. (2019) 'ES10.05 Brazilian Health Regulatory Agency - Ag□ia Nacional de Vigil↦ia Sanit□a', *Journal of Thoracic Oncology,* 14(10), pp. S39.

Ferreira, C. N., Blunk, V., Santana, C. F. S. and Squiassi, H. (2016) 'Impacto socioeconômico do câncer de

pulmão não-pequenas células no Brasil', *Jornal brasileiro de economia da saúde* 8(2), pp. 91-98, http://www.jbes.com.br/images/v8n2/91.pdf. DOI: 10.21115/JBES.v8.n2.p91-98.

Fitzmaurice, C. and Abate, D. and Abbasi, N. and Abbastabar, H. and Abd-Allah, F. and Abdel-Rahman, O. and Abdelalim, A. and Abdoli, A. and Abdollahpour, I. and Abdulle, A. S. M. and Abebe, N. D. and Abraha, H. N. and Abu-Raddad, L. J. and Abualhasan, A. and Adedeji, I. A. and Advani, S. M. and Afarideh, M. and Afshari, M. and Aghaali, M. and Agius, D. and Agrawal, S. and Ahmadi, A. and Ahmadian, E. and Ahmadpour, E. and Ahmed, M. B. and Akbari, M. E. and Akinyemiju, T. and Al-Aly, Z. and AlAbdulKader, A. M. and Alahdab, F. and Alam, T. and Alamene, G.

M. and Alemnew, B. T. T. and Alene, K. A. and Alinia, C. and Alipour, V. and Aljunid, S. M. and Bakeshei, F. A. and Almadi, M. A. H. and Almasi-Hashiani, A. and Alsharif, U. and Alsowaidi, S. and Alvis-Guzman, N. and Amini, E. and Amini, S. and Amoako, Y. A. and Anbari, Z. and Anber, N. H. and Andrei, C. L. and Anjomshoa, M. and Ansari, F. and Ansariadi, A. and Appiah, S. C. Y. and Arab-Zozani, M. and Arabloo, J. and Arefi, Z. and Aremu, O. and Areri, H. A. and Artaman, A. and Asayesh, H. and Asfaw, E. T. and Ashagre, A. F. and Assadi, R. and Ataeinia, B. and Atalay, H. T. and Ataro, Z. and Atique, S. and Ausloos, M. and Avila-Burgos, L. and Avokpaho, E. F. G. A. and Awasthi, A. and Awoke, N. and Ayala Quintanilla, B. P. and Ayanore, M. A. and Ayele, H. T. and Babaee, E. and Bacha, U. and Badawi, A. and Bagherzadeh, M. and Bagli, E. and Balakrishnan, S. and Balouchi, A. and Bärnighausen, T. W. and Battista, R. J. and Behzadifar, M. and Bekele, B. B. and Belay, Y. B. and Belayneh, Y. M. and Berfield, K. K. S. and Berhane, A. and Bernabe, E. and Beuran, M. and Bhakta, N. and Bhattacharyya, K. and Biadgo, B. and Bijani, A. and Bin Sayeed, M. S. and Birungi, C. and Bisignano, C. and Bitew, H. and Bjørge, T. and Bleyer, A. and Bogale, K. A. and Bojia, H. A. and Borzì, A. M. and Bosetti, C. and Bou-Orm, I. R. and Brenner, H. and Brewer, J. D. and Briko, A. N. and Briko, N. I. and Bustamante-Teixeira, M. T. and Butt, Z. A. and Carreras, G. and Carrero, J. J. and Carvalho, F. and Castro, C. and Castro, F. and Catalá-López, F. and Cerin, E. and Chaiah, Y. and Chanie, W. F. and Chattu, V. K. and Chaturvedi, P. and Chauhan, N. S. and Chehrazi, M. and Chiang, P. P. and Chichiabellu, T. Y. and Chido-Amajuoyi, O. G. and Chimed-Ochir, O. and Choi, J. J. and Christopher, D. J. and Chu, D. T. and Constantin, M. M. and Costa, V. M. and Crocetti, E. and Crowe, C. S. and Curado, M. P. and Dahlawi, S. M. A. and Damiani, G. and Darwish, A. H. and Daryani, A. and das Neves, J. and Demeke, F. M. and Demis, A. B. and Demissie, B. W. and Demoz, G. T. and Denova-Gutiérrez, E. and Derakhshani, A. and Deribe, K. S. and Desai, R. and Desalegn, B. B. and Desta, M. and Dey, S. and Dharmaratne, S. D. and Dhimal, M. and Diaz, D. and Dinberu, M. T. T. and Djalalinia, S. and Doku, D. T. and Drake, T. M. and Dubey, M. and Dubljanin, E. and Duken, E. E. and Ebrahimi, H. and Effiong, A. and Eftekhari, A. and El Sayed, I. and Zaki, M. E. S. and El-Jaafary, S. I. and El-Khatib, Z. and Elemineh, D. A. and Elkout, H. and Ellenbogen, R. G. and Elsharkawy, A. and Emamian, M. H. and Endalew, D. A. and Endries, A. Y. and Eshrati, B. and Fadhil, I. and Fallah Omrani, V. and Faramarzi, M. and Farhangi, M. A. and Farioli, A. and Farzadfar, F. and Fentahun, N. and Fernandes, E. and Feyissa, G. T. and Filip, I. and Fischer, F. and Fisher, J. L. and Force, L. M. and Foroutan, M. and Freitas, M. and Fukumoto, T. and Futran, N. D. and Gallus, S. and Gankpe, F. G. and Gayesa, R. T. and Gebrehiwot, T. T. and Gebremeskel, G. G. and Gedefaw, G. A. and Gelaw, B. K. and Geta, B. and Getachew, S. and Gezae, K. E. and Ghafourifard, M. and Ghajar, A. and Ghashghaee, A. and Gholamian, A. and Gill, P. S. and Ginindza, T. T. G. and Girmay, A. and Gizaw, M. and Gomez, R. S. and Gopalani, S. V. and Gorini, G. and Goulart, B. N. G. and Grada, A. and Ribeiro Guerra, M. and Guimaraes, A. L. S. and Gupta, P. C. and Gupta, R. and Hadkhale, K. and Haj-Mirzaian, A. and Hamadeh, R. R. and Hamidi, S. and Hanfore, L. K. and Haro, J. M. and Hasankhani, M. and Hasanzadeh, A. and Hassen, H. Y. and Hay, R. J. and Hay, S. I. and Henok, A. and Henry, N. J.

and Herteliu, C. and Hidru, H. D. and Hoang, C. L. and Hole, M. K. and Hoogar, P. and Horita, N. and Hosgood, H. D. and Hosseini, M. and Hosseinzadeh, M. and Hostiuc, M. and Hostiuc, S. and Househ, M. and Hussen, M. M. and Ileanu, B. and Ilic, M. D. and Innos, K. and Irvani, S. S. N. and Iseh, K. R. and Islam, S. M. S. and Islami, F. and Jafari Balalami, N. and Jafarinia, M. and Jahangiry, L. and Jahani, M. A. and Jahanmehr, N. and Jakovljevic, M. and James, S. L. and Javanbakht, M. and Jayaraman, S. and Jee, S. H. and Jenabi, E. and Jha, R. P. and Jonas, J. B. and Jonnagaddala, J. and Joo, T. and Jungari, S. B. and Jürisson, M. and Kabir, A. and Kamangar, F. and Karch, A. and Karimi, N. and Karimian, A. and Kasaeian, A. and Kasahun, G. G. and Kassa, B. and Kassa, T. D. and Kassaw, M. W. and Kaul, A. and Keiyoro, P. N. and Kelbore, A. G. and Kerbo, A. A. and Khader, Y. S. and Khalilarjmandi, M. and Khan, E. A. and Khan, G. and Khang, Y. H. and Khatab, K. and Khater, A. and Khayamzadeh, M. and Khazaee-Pool, M. and Khazaei, S. and Khoja, A. T. and Khosravi, M. H. and Khubchandani, J. and Kianipour, N. and Kim, D. and Kim, Y. J. and Kisa, A. and Kisa, S. and Kissimova-Skarbek, K. and Komaki, H. and Koyanagi, A. and Krohn, K. J. and Bicer, B. K. and Kugbey, N. and Kumar, V. and Kuupiel, D. and La Vecchia, C. and Lad, D. P. and Lake, E. A. and Lakew, A. M. and Lal, D. K. and Lami, F. H. and Lan, Q. and Lasrado, S. and Lauriola, P. and Lazarus, J. V. and Leigh, J. and Leshargie, C. T. and Liao, Y. and Limenih, M. A. and Listl, S. and Lopez, A. D. and Lopukhov, P. D. and Lunevicius, R. and Madadin, M. and Magdeldin, S. and El Razek, H. M. A. and Majeed, A. and Maleki, A. and Malekzadeh, R. and Manafi, A. and Manafi, N. and Manamo, W. A. and Mansourian, M. and Mansournia, M. A. and Mantovani, L. G. and Maroufizadeh, S. and Martini, S. M. S. and Mashamba-Thompson, T. P. and Massenburg, B. B. and Maswabi, M. T. and Mathur, M. R. and McAlinden, C. and McKee, M. and Meheretu, H. A. A. and Mehrotra, R. and Mehta, V. and Meier, T. and Melaku, Y. A. and Meles, G. G. and Meles, H. G. and Melese, A. and Melku, M. and Memiah, P. T. N. and Mendoza, W. and Menezes, R. G. and Merat, S. and Meretoja, T. J. and Mestrovic, T. and Miazgowski, B. and Miazgowski, T. and Mihretie, K. M. M. and Miller, T. R. and Mills, E. J. and Mir, S. M. and Mirzaei, H. and Mirzaei, H. R. and Mishra, R. and Moazen, B. and Mohammad, D. K. and Mohammad, K. A. and Mohammad, Y. and Darwesh, A. M. and Mohammadbeigi, A. and Mohammadi, H. and Mohammadi, M. and Mohammadian, M. and Mohammadian-Hafshejani, A. and Mohammadoo-Khorasani, M. and Mohammadpourhodki, R. and Mohammed, A. S. and Mohammed, J. A. and Mohammed, S. and Mohebi, F. and Mokdad, A. H. and Monasta, L. and Moodley, Y. and Moosazadeh, M. and Moossavi, M. and Moradi, G. and Moradi-Joo, M. and Moradi-Lakeh, M. and Moradpour, F. and Morawska, L. and Morgado-da-Costa, J. and Morisaki, N. and Morrison, S. D. and Mosapour, A. and Mousavi, S. M. and Muche, A. A. and Muhammed, O. S. S. and Musa, J. and Nabhan, A. F. and Naderi, M. and Nagarajan, A. J. and Nagel, G. and Nahvijou, A. and Naik, G. and Najafi, F. and Naldi, L. and Nam, H. S. and Nasiri, N. and Nazari, J. and Negoi, I. and Neupane, S. and Newcomb, P. A. and Nggada, H. A. and Ngunjiri, J. W. and Nguyen, C. T. and Nikniaz, L. and Ningrum, D. N. A. and Nirayo, Y. L. and Nixon, M. R. and Nnaji, C. A. and Nojomi, M. and Nosratnejad, S. and Shiadeh, M. N. and Obsa, M. S. and Ofori-Asenso, R. and Ogbo, F. A. and Oh, I. H. and Olagunju, A.

T. and Olagunju, T. O. and Oluwasanu, M. M. and Omonisi, A. E. and Onwujekwe, O. E. and Oommen, A. M. and Oren, E. and Ortega-Altamirano, D. D. V. and Ota, E. and Otstavnov, S. S. and Owolabi, M. O. and P A, M. and Padubidri, J. R. and Pakhale, S. and Pakpour, A. H. and Pana, A. and Park, E. K. and Parsian, H. and Pashaei, T. and Patel, S. and Patil, S. T. and Pennini, A. and Pereira, D. M. and Piccinelli, C. and Pillay, J. D. and Pirestani, M. and Pishgar, F. and Postma, M. J. and Pourjafar, H. and Pourmalek, F. and Pourshams, A. and Prakash, S. and Prasad, N. and Qorbani, M. and Rabiee, M. and Rabiee, N. and Radfar, A. and Rafiei, A. and Rahim, F. and Rahimi, M. and Rahman, M. A. and Rajati, F. and Rana, S. M. and Raoofi, S. and Rath, G. K. and Rawaf, D. L. and Rawaf, S. and Reiner, R. C. and Renzaho, A. M. N. and Rezaei, N. and Rezapour, A. and Ribeiro, A. I. and Ribeiro, D. and Ronfani, L. and Roro, E. M. and Roshandel, G. and Rostami, A. and Saad, R. S. and Sabbagh, P. and Sabour, S. and Saddik, B. and Safiri, S. and Sahebkar, A. and Salahshoor, M. R. and Salehi, F. and Salem, H. and Salem, M. R. and Salimzadeh, H. and Salomon, J. A. and Samy, A. M. and Sanabria, J. and Santric Milicevic, M. M. and Sartorius, B. and Sarveazad, A. and Sathian, B. and Satpathy, M. and Savic, M. and Sawhney, M. and Sayyah, M. and Schneider, I. J. C. and Schöttker, B. and Sekerija, M. and Sepanlou, S. G. and Sepehrimanesh, M. and Seyedmousavi, S. and Shaahmadi, F. and Shabaninejad, H. and Shahbaz, M. and Shaikh, M. A. and Shamshirian, A. and Shamsizadeh, M. and Sharafi, H. and Sharafi, Z. and Sharif, M. and Sharifi, A. and Sharifi, H. and Sharma, R. and Sheikh, A. and Shirkoohi, R. and Shukla, S. R. and Si, S. and Siabani, S. and Silva, D. A. S. and Silveira, D. G. A. and Singh, A. and Singh, J. A. and Sisay, S. and Sitas, F. and Sobngwi, E. and Soofi, M. and Soriano, J. B. and Stathopoulou, V. and Sufiyan, M. B. and Tabarés-Seisdedos, R. and Tabuchi, T. and Takahashi, K. and Tamtaji, O. R. and Tarawneh, M. R. and Tassew, S. G. and Taymoori, P. and Tehrani-Banihashemi, A. and Temsah, M. H. and Temsah, O. and Tesfay, B. E. and Tesfay, F. H. and Teshale, M. Y. and Tessema, G. A. and Thapa, S. and Tlaye, K. G. and Topor-Madry, R. and Tovani-Palone, M. R. and Traini, E. and Tran, B. X. and Tran, K. B. and Tsadik, A. G. and Ullah, I. and Uthman, O. A. and Vacante, M. and Vaezi, M. and Varona Pérez, P. and Veisani, Y. and Vidale, S. and Violante, F. S. and Vlassov, V. and Vollset, S. E. and Vos, T. and Vosoughi, K. and Vu, G. T. and Vujcic, I. S. and Wabinga, H. and Wachamo, T. M. and Wagnew, F. S. and Waheed, Y. and Weldegebreal, F. and Weldesamuel, G. T. and Wijeratne, T. and Wondafrash, D. Z. and Wonde, T. E. and Wondmieneh, A. B. and Workie, H. M. and Yadav, R. and Yadegar, A. and Yadollahpour, A. and Yaseri, M. and Yazdi-Feyzabadi, V. and Yeshaneh, A. and Yimam, M. A. and Yimer, E. M. and Yisma, E. and Yonemoto, N. and Younis, M. Z. and Yousefi, B. and Yousefifard, M. and Yu, C. and Zabeh, E. and Zadnik, V. and Moghadam, T. Z. and Zaidi, Z. and Zamani, M. and Zandian, H. and Zangeneh, A. and Zaki, L. and Zendehdel, K. and Zenebe, Z. M. and Zewale, T. A. and Ziapour, A. and Zodpey, S. and Murray, C. J. L. and Collaboration, G. B. o. D. C. (2019) 'Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study', *JAMA Oncol,* 5(12), pp. 1749-1768.

Fuchs, V., Roisman, L., Kian, W., Daniel, L., Dudnik, J., Nechushtan, H., Goldstein, I., Dvir, A., Soussan-Gutman, L., Grinberg, R., Gillis, R. and Peled, N. (2021) 'The impact of osimertinib' line on clonal evolution in EGFRm NSCLC through NGS-based liquid biopsy and overcoming strategies for resistance', *Lung Cancer,* 153, pp. 126-133.

Fujita, Y., Suda, K., Kimura, H., Matsumoto, K., Arao, T., Nagai, T., Saijo, N., Yatabe, Y., Mitsudomi, T. and Nishio, K. (2012) 'Highly sensitive detection of EGFR T790M mutation using colony hybridization predicts favorable prognosis of patients with lung cancer harboring activating EGFR mutation', *J Thorac Oncol,* 7(11), pp. 1640-4.

Gahr, S., Stoehr, R., Geissinger, E., Ficker, J. H., Brueckl, W. M., Gschwendtner, A., Gattenloehner, S., Fuchs, F. S., Schulz, C., Rieker, R. J., Hartmann, A., Ruemmele, P. and Dietmaier, W. (2013) 'EGFR mutational status in a large series of Caucasian European NSCLC patients: data from daily practice', *Br J Cancer,* 109(7), pp. 1821-8.

Gainor, J. F. and Shaw, A. T. (2013) 'Emerging paradigms in the development of resistance to tyrosine kinase inhibitors in lung cancer', *J Clin Oncol,* 31(31), pp. 3987-96.

Gao, B., Sun, Y., Zhang, J., Ren, Y., Fang, R., Han, X., Shen, L., Liu, X. Y., Pao, W., Chen, H. and Ji, H. (2010) 'Spectrum of LKB1, EGFR, and KRAS mutations in chinese lung adenocarcinomas', *J Thorac Oncol,* 5(8), pp. 1130-5.

Garon, E. B., Rizvi, N. A., Hui, R., Leighl, N., Balmanoukian, A. S., Eder, J. P., Patnaik, A., Aggarwal, C., Gubens, M., Horn, L., Carcereny, E., Ahn, M. J., Felip, E., Lee, J. S., Hellmann, M. D., Hamid, O., Goldman, J. W., Soria, J. C., Dolled-Filhart, M., Rutledge, R. Z., Zhang, J., Lunceford, J. K., Rangwala, R., Lubiniecki, G. M., Roach, C., Emancipator, K., Gandhi, L. and Investigators, K.-. (2015) 'Pembrolizumab for the treatment of non-small-cell lung cancer', *N Engl J Med,* 372(21), pp. 2018-28.

Godsey, J. H., Silvestro, A., Barrett, J. C., Bramlett, K., Chudova, D., Deras, I., Dickey, J., Hicks, J., Johann, D. J., Leary, R., Lee, J. S. H., McMullen, J., McShane, L., Nakamura, K., Richardson, A. O., Ryder, M., Simmons, J., Tanzella, K., Yee, L. and Leiman, L. C. (2020) 'Generic Protocols for the Analytical Validation of Next-Generation Sequencing-Based ctDNA Assays: A Joint Consensus Recommendation of the BloodPAC's Analytical Variables Working Group', *Clin Chem,* 66(9), pp. 1156-1166.

Goldstraw, P., Chansky, K., Crowley, J., Rami-Porta, R., Asamura, H., Eberhardt, W. E., Nicholson, A. G., Groome, P., Mitchell, A., Bolejack, V., International Association for the Study of Lung Cancer Staging and Prognostic Factors Committee, A. v. B., and Participating Institutions and Institutions, I. A. f. t. S. o. L. C. S. a. P. F. C. A. B. a. P. (2016) 'The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer', *J Thorac Oncol,* 11(1), pp. 39-51.

Gomes, J. R., Amarante, M. P. F., D'Alpino, R. D. A., Moreira, R. B., Sousa, T. T. S. d., Lino, A. D. R., Cruz, M. R. S., Costa, M. A., Schutz, F., Kawamura Haddad, C., Schmerling, R., Maluf, F. C. and Buzaid, A. C. (2015) 'Mutation profile in non-small

cell lung cancer: Analysis of a Brazilian population', *Journal of Clinical Oncology,* 33(15_suppl), pp. e19115-e19115.

Gonzalvez, F., Vincent, S., Baker, T. E., Gould, A. E., Li, S., Wardwell, S. D., Nadworny, S., Ning, Y., Zhang, S., Huang, W. S., Hu, Y., Li, F., Greenfield, M. T., Zech, S. G., Das, B., Narasimhan, N. I., Clackson, T., Dalgarno, D., Shakespeare, W. C., Fitzgerald, M., Chouitar, J., Griffin, R. J., Liu, S., Wong, K. K., Zhu, X. and Rivera, V. M. (2021) 'Mobocertinib (TAK-788): A Targeted Inhibitor of', *Cancer Discov,* 11(7), pp. 1672-1687.

Greulich, H., Chen, T. H., Feng, W., Jänne, P. A., Alvarez, J. V., Zappaterra, M., Bulmer, S. E., Frank, D. A., Hahn, W. C., Sellers, W. R. and Meyerson, M. (2005) 'Oncogenic transformation by inhibitor-sensitive and -resistant EGFR mutants', *PLoS Med,* 2(11), pp. e313.

Greulich, H., Kaplan, B., Mertins, P., Chen, T. H., Tanaka, K. E., Yun, C. H., Zhang, X., Lee, S. H., Cho, J., Ambrogio, L., Liao, R., Imielinski, M., Banerji, S., Berger, A. H., Lawrence, M. S., Zhang, J., Pho, N. H., Walker, S. R., Winckler, W., Getz, G., Frank, D., Hahn, W. C., Eck, M. J., Mani, D. R., Jaffe, J. D., Carr, S. A., Wong, K. K. and Meyerson, M. (2012) 'Functional analysis of receptor tyrosine kinase mutations in lung cancer identifies oncogenic extracellular domain mutations of ERBB2', *Proc Natl Acad Sci U S A,* 109(36), pp. 14476-81.

Greytak, S. R., Engel, K. B., Parpart-Li, S., Murtaza, M., Bronkhorst, A. J., Pertile, M. D. and Moore, H. M. (2020) 'Harmonizing Cell-Free DNA Collection and Processing Practices through Evidence-Based Guidance', *Clin Cancer Res,* 26(13), pp. 3104-3109.

Gridelli, C., Rossi, A., Carbone, D. P., Guarize, J., Karachaliou, N., Mok, T., Petrella, F., Spaggiari, L. and Rosell, R. (2015) 'Non-small-cell lung cancer', *Nat Rev Dis Primers,* 1, pp. 15009.

Guibert, N., Pradines, A., Favre, G. and Mazieres, J. (2020) 'Current and future applications of liquid biopsy in nonsmall cell lung cancer from early to advanced stages', *Eur Respir Rev,* 29(155).

Haber, D. A. and Settleman, J. (2007) 'Cancer: drivers and passengers', *Nature,* 446(7132), pp. 145-6.

Harrison, P. T., Vyse, S. and Huang, P. H. (2020) 'Rare epidermal growth factor receptor (EGFR) mutations in non-small cell lung cancer', *Semin Cancer Biol,* 61, pp. 167-179.

Hecht, S. S. (2012) 'Lung carcinogenesis by tobacco smoke', *Int J Cancer,* 131(12), pp. 2724-32.

Heist, R. S., Sequist, L. V. and Engelman, J. A. (2012) 'Genetic changes in squamous cell lung cancer: a review', *J Thorac Oncol,* 7(5), pp. 924-33.

Heitzer, E., Auer, M., Ulz, P., Geigl, J. B. and Speicher, M. R. (2013) 'Circulating tumor cells and DNA as liquid biopsies', *Genome Med,* 5(8), pp. 73.

Heitzer, E., van den Broek, D., Denis, M. G., Hofman, P., Hubank, M., Mouliere, F., Paz-Ares, L., Schuuring, E., Sültmann, H., Vainer, G., Verstraaten, E., de Visser, L. and Cortinovis, D. (2022) 'Recommendations for a practical implementation of circulating tumor DNA mutation testing in metastatic non-small-cell lung cancer', *ESMO Open,* 7(2), pp. 100399.

Helmy, M., Awad, M. and Mosa, K. A. (2016) 'Limited resources of genome sequencing in developing countries: Challenges and solutions', *Appl Transl Genom,* 9, pp. 15-9.

Herbst, R. S., Lopes, G., Kowalski, D. M., Kasahara, K., Wu, Y. L., De Castro, G., Jr., Cho, B. C., Turna, H. Z., Cristescu, R., Aurora-Garg, D., Lunceford, J., Kobie, J., Ayers, M., Pietanza, M. C., Piperdi, B. and Mok, T. S. K. (2019) 'LBA4 Association of KRAS mutational status with response to pembrolizumab monotherapy given as first-line therapy for PD-L1-positive advanced non-squamous NSCLC in Keynote-042', *Annals of Oncology,* 30, pp. xi63-xi64.

Herbst, R. S., Soria, J. C., Kowanetz, M., Fine, G. D., Hamid, O., Gordon, M. S., Sosman, J. A., McDermott, D. F., Powderly, J. D., Gettinger, S. N., Kohrt, H. E., Horn, L., Lawrence, D. P., Rost, S., Leabman, M., Xiao, Y., Mokatrin, A., Koeppen, H., Hegde, P. S., Mellman, I., Chen, D. S. and Hodi, F. S. (2014) 'Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients', *Nature,* 515(7528), pp. 563-7.

Hochmair, M. J., Buder, A., Schwab, S., Burghuber, O. C., Prosch, H., Hilbe, W., Cseh, A., Fritz, R. and Filipits, M. (2019) 'Liquid-Biopsy-Based Identification of EGFR T790M Mutation-Mediated Resistance to Afatinib Treatment in Patients with Advanced EGFR Mutation-Positive NSCLC, and Subsequent Response to Osimertinib', *Target Oncol,* 14(1), pp. 75-83.

Hu, Y., Ulrich, B. C., Supplee, J., Kuang, Y., Lizotte, P. H., Feeney, N. B., Guibert, N. M., Awad, M. M., Wong, K. K., Jänne, P. A., Paweletz, C. P. and Oxnard, G. R. (2018) 'False-Positive Plasma Genotyping Due to Clonal Hematopoiesis', *Clin Cancer Res,* 24(18), pp. 4437-4443.

Huang, X., Jin, R., Lou, L., Zhao, J., Xia, L., Li, W., Xu, Z. and Xia, Y. (2020) 'The efficacy of ado-trastuzumab emtansine in patients with ERBB2-aberrant non-small cell lung cancer: a systematic review', *Transl Cancer Res,* 9(8), pp. 4507-4516.

IJzerman, M. J., de Boer, J., Azad, A., Degeling, K., Geoghegan, J., Hewitt, C., Hollande, F., Lee, B., To, Y. H., Tothill, R. W., Wright, G., Tie, J. and Dawson, S. J. (2021) 'Towards Routine Implementation of Liquid Biopsies in Cancer Management: It Is Always Too Early, until Suddenly It Is Too Late', *Diagnostics (Basel),* 11(1).

Imyanitov, E. N., Iyevleva, A. G. and Levchenko, E. V. (2021) 'Molecular testing and targeted therapy for non-small cell lung cancer: Current status and perspectives', *Crit Rev Oncol Hematol,* 157, pp. 103194.

Ismael, G., Coradazzi, A. L., Cantarelli, A. V., Abdalla, K. C., Mattos, E. R., Oliveira, J. d. S., Manzoni, C., Beato, C. A. d. M. and Segalla, J. G. M. (2011) 'Lung cancer histology in a Brazil: A retrospective study of 1,887 patients', *Journal of Clinical Oncology,* 29(15_suppl), pp. e18039-e18039.

Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P. V., Mar, B. G., Lindsley, R. C., Mermel, C. H., Burtt, N., Chavez, A., Higgins, J. M., Moltchanov, V., Kuo, F. C., Kluk, M. J., Henderson, B., Kinnunen, L., Koistinen, H. A., Ladenvall, C., Getz, G., Correa, A., Banahan, B. F., Gabriel, S., Kathiresan, S., Stringham, H. M., McCarthy, M. I., Boehnke, M., Tuomilehto, J., Haiman, C., Groop, L., Atzmon, G., Wilson, J. G., Neuberg, D., Altshuler, D. and Ebert, B. L. (2014) 'Age-related clonal hematopoiesis associated with adverse outcomes', *N Engl J Med,* 371(26), pp. 2488-98.

Jiang, L., Mino-Kenudson, M., Roden, A. C., Rosell, R., Molina, M., Flores, R. M., Pilz, L. R., Brunelli, A., Venuta, F., He, J. and Group, A. L. C. C. (2019) 'Association between the novel classification of lung adenocarcinoma subtypes and EGFR/KRAS mutation status: A systematic literature review and pooled-data analysis', *Eur J Surg Oncol,* 45(5), pp. 870-876.

Jiang, L., Su, X., Zhang, T., Yin, X., Zhang, M., Fu, H., Han, H., Sun, Y., Dong, L., Qian, J., Xu, Y., Fu, X., Gavine, P. R., Zhou, Y., Tian, K., Huang, J., Shen, D., Jiang, H., Yao, Y., Han, B. and Gu, Y. (2017) 'PD-L1 expression and its relationship with oncogenic drivers in non-small cell lung cancer (NSCLC)', *Oncotarget,* 8(16), pp. 26845-26857.

Johnson, D. H., Fehrenbacher, L., Novotny, W. F., Herbst, R. S., Nemunaitis, J. J., Jablons, D. M., Langer, C. J., DeVore, R. F., Gaudreault, J., Damico, L. A., Holmgren, E. and Kabbinavar, F. (2004) 'Randomized phase II trial comparing bevacizumab plus carboplatin and paclitaxel with carboplatin and paclitaxel alone in previously untreated locally advanced or metastatic non-small-cell lung cancer', *J Clin Oncol,* 22(11), pp. 2184-91.

Kandoth, C., McLellan, M. D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J. F., Wyczalkowski, M. A., Leiserson, M. D. M., Miller, C. A., Welch, J. S., Walter, M. J., Wendl, M. C., Ley, T. J., Wilson, R. K., Raphael, B. J. and Ding, L. (2013) 'Mutational landscape and significance across 12 major cancer types', *Nature,* 502(7471), pp. 333-339.

Kemper, K., Krijgsman, O., Cornelissen-Steijger, P., Shahrabi, A., Weeber, F., Song, J. Y., Kuilman, T., Vis, D. J., Wessels, L. F., Voest, E. E., Schumacher, T. N., Blank, C. U., Adams, D. J., Haanen, J. B. and Peeper, D. S. (2015) 'Intra- and inter-tumor heterogeneity in a vemurafenib-resistant melanoma patient and derived xenografts', *EMBO Mol Med,* 7(9), pp. 1104-18.

Kerr, K. M., Bibeau, F., Thunnissen, E., Botling, J., Ryška, A., Wolf, J., Öhrling, K., Burdon, P., Malapelle, U. and Büttner, R. (2021) 'The evolving landscape of biomarker testing for non-small cell lung cancer in Europe', *Lung Cancer,* 154, pp. 161-175.

Knust, R. E., Portela, M. C., Pereira, C. C. A. and Fortes, G. B. (2017) 'Estimated costs of advanced lung cancer care in a public reference hospital', *Rev Saude Publica,* 51, pp. 53.

Kubo, T., Yamamoto, H., Lockwood, W. W., Valencia, I., Soh, J., Peyton, M., Jida, M., Otani, H., Fujii, T., Ouchida, M., Takigawa, N., Kiura, K., Shimizu, K., Date, H., Minna, J. D., Varella-Garcia, M., Lam, W. L., Gazdar, A. F. and Toyooka, S. (2009) 'MET gene amplification or EGFR mutation activate MET in lung cancers untreated with EGFR tyrosine kinase inhibitors', *Int J Cancer,* 124(8), pp. 1778-84.

Kwak, E. L., Bang, Y. J., Camidge, D. R., Shaw, A. T., Solomon, B., Maki, R. G., Ou, S. H., Dezube, B. J., Jänne, P. A., Costa, D. B., Varella-Garcia, M., Kim, W. H., Lynch, T. J., Fidias, P., Stubbs, H., Engelman, J. A., Sequist, L. V., Tan, W., Gandhi, L., Mino-Kenudson, M., Wei, G. C., Shreeve, S. M., Ratain, M. J., Settleman, J., Christensen, J. G., Haber, D. A., Wilner, K., Salgia, R., Shapiro, G. I., Clark, J. W. and Iafrate, A. J. (2010) 'Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer', *N Engl J Med,* 363(18), pp. 1693-703.

Leal, A., van Grieken, N. C. T., Palsgrove, D. N., Phallen, J., Medina, J. E., Hruban, C., Broeckaert, M. A. M., Anagnostou, V., Adleff, V., Bruhm, D. C., Canzoniero, J. V., Fiksel, J., Nordsmark, M., Warmerdam, F. A. R. M., Verheul, H. M. W., van Spronsen, D. J., Beerepoot, L. V., Geenen, M. M., Portielje, J. E. A., Jansen, E. P. M., van Sandick, J., Meershoek-Klein Kranenbarg, E., van Laarhoven, H. W. M., van der Peet, D. L., van de Velde, C. J. H., Verheij, M., Fijneman, R., Scharpf, R. B., Meijer, G. A., Cats, A. and Velculescu, V. E. (2020) 'White blood cell and cell-free DNA analyses for detection of residual disease in gastric cancer', *Nat Commun,* 11(1), pp. 525.

Leal, L. F., de Paula, F. E., De Marchi, P., de Souza Viana, L., Pinto, G. D. J., Carlos, C. D., Berardinelli, G. N., Miziara, J. E., da Silva, C. M., Silva, E. C. A., Pereira, R., de Oliveira, M. A., Scapulatempo-Neto, C. and Reis, R. M. (2019) 'Mutational profile of Brazilian lung adenocarcinoma unveils association of EGFR mutations with high Asian ancestry and independent prognostic role of KRAS mutations', *Sci Rep,* 9(1), pp. 3209.

Lee, H. W., Park, S. H., Weng, M. W., Wang, H. T., Huang, W. C., Lepor, H., Wu, X. R., Chen, L. C. and Tang, M. S. (2018) 'E-cigarette smoke damages DNA and reduces repair activity in mouse lung, heart, and bladder as well as in human lung and bladder cells', *Proc Natl Acad Sci U S A,* 115(7), pp. E1560-E1569.

Lee, S. Y., Kim, M. J., Jin, G., Yoo, S. S., Park, J. Y., Choi, J. E., Jeon, H. S., Cho, S., Lee, E. B., Cha, S. I., Park, T. I., Kim, C. H. and Jung, T. H. (2010) 'Somatic mutations in epidermal growth factor receptor signaling pathway genes in non-small cell lung cancers', *J Thorac Oncol,* 5(11), pp. 1734-40.

Lemmon, M. A. and Schlessinger, J. (2010) 'Cell signaling by receptor tyrosine kinases', *Cell,* 141(7), pp. 1117-34.

Leon, S. A., Shapiro, B., Sklaroff, D. M. and Yaros, M. J. (1977) 'Free DNA in the serum of cancer patients and the effect of therapy', *Cancer Res,* 37(3), pp. 646-50.

Leonetti, A., Minari, R., Mazzaschi, G., Gnetti, L., La Monica, S., Alfieri, R., Campanini, N., Verzè, M., Olivani, A., Ventura, L. and Tiseo, M. (2021) 'Small Cell Lung Cancer Transformation as a Resistance Mechanism to Osimertinib in Epidermal Growth Factor Receptor-Mutated Lung Adenocarcinoma: Case Report and Literature Review', *Front Oncol,* 11, pp. 642190.

Li, M. M., Datto, M., Duncavage, E. J., Kulkarni, S., Lindeman, N. I., Roy, S., Tsimberidou, A. M., Vnencak-Jones, C. L., Wolff, D. J., Younes, A. and Nikiforova, M. N. (2017) 'Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists', *J Mol Diagn,* 19(1), pp. 4-23.

Liam, C. K., Mallawathantri, S. and Fong, K. M. (2020) 'Is tissue still the issue in detecting molecular alterations in lung cancer?', *Respirology,* 25(9), pp. 933-943.

Liang, H. and Wang, M. (2020) 'MET Oncogene in Non-Small Cell Lung Cancer: Mechanism of MET Dysregulation and Agents Targeting the HGF/c-Met Axis', *Onco Targets Ther,* 13, pp. 2491-2510.

Liang, Y. J. and Yang, W. X. (2019) 'Kinesins in MAPK cascade: How kinesin motors are involved in the MAPK pathway?', *Gene,* 684, pp. 1-9.

Lim, S. M., Syn, N. L., Cho, B. C. and Soo, R. A. (2018) 'Acquired resistance to EGFR targeted therapy in non-small cell lung cancer: Mechanisms and therapeutic strategies', *Cancer Treat Rev,* 65, pp. 1-10.

Lim, Z. F. and Ma, P. C. (2019) 'Emerging insights of tumor heterogeneity and drug resistance mechanisms in lung cancer targeted therapy', *J Hematol Oncol,* 12(1), pp. 134.

Lindeman, N. I., Cagle, P. T., Aisner, D. L., Arcila, M. E., Beasley, M. B., Bernicker, E. H., Colasacco, C., Dacic, S., Hirsch, F. R., Kerr, K., Kwiatkowski, D. J., Ladanyi, M., Nowak, J. A., Sholl, L., Temple-Smolkin, R., Solomon, B., Souter, L. H., Thunnissen, E., Tsao, M. S., Ventura, C. B., Wynes, M. W. and Yatabe, Y. (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 Association for Molecular Pathology', *Arch Pathol Lab Med,* 142(3), pp. 321-346.

Liu, B. J., Li, F. F., Xie, Y. X., Fan, C. Y., Liu, W. J., Qiu, J. G. and Jiang, B. H. (2022a) 'miR-196a Upregulation Contributes to Gefitinib Resistance through Inhibiting GLTP Expression', *Int J Mol Sci,* 23(3).

Liu, F., Wei, Y., Zhang, H., Jiang, J., Zhang, P. and Chu, Q. (2022b) 'NTRK Fusion in Non-Small Cell Lung Cancer: Diagnosis, Therapy, and TRK Inhibitor Resistance', *Front Oncol,* 12, pp. 864666.

Liu, J., Chen, X., Wang, J., Zhou, S., Wang, C. L., Ye, M. Z., Wang, X. Y., Song, Y., Wang, Y. Q., Zhang, L. T., Wu, R. H., Yang, H. M., Zhu, S. D., Zhou, M. Z., Zhang, X. C., Zhu, H. M. and Qian, Z. Y. (2019) 'Biological background of the genomic variations of cf-DNA in healthy individuals', *Ann Oncol,* 30(3), pp. 464-470.

Lopes, L. F. and Bacchi, C. E. (2012) 'Anaplastic lymphoma kinase gene rearrangement in non-small-cell lung cancer in a Brazilian population', *Clinics (Sao Paulo),* 67(7), pp. 845-7.

Lortet-Tieulent, J., Soerjomataram, I., Ferlay, J., Rutherford, M., Weiderpass, E. and Bray, F. (2014) 'International trends in lung cancer incidence by histological subtype: adenocarcinoma stabilizing in men but still increasing in women', *Lung Cancer,* 84(1), pp. 13-22.

Lu, Y., Futtner, C., Rock, J. R., Xu, X., Whitworth, W., Hogan, B. L. and Onaitis, M. W. (2010) 'Evidence that SOX2 overexpression is oncogenic in the lung', *PLoS One,* 5(6), pp. e11022.

Lynch, T. J., Bell, D. W., Sordella, R., Gurubhagavatula, S., Okimoto, R. A., Brannigan, B. W., Harris, P. L., Haserlat, S. M., Supko, J. G., Haluska, F. G., Louis, D. N., Christiani, D. C., Settleman, J. and Haber, D. A. (2004) 'Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib', *N Engl J Med,* 350(21), pp. 2129-39.

Ma, C., Yang, X., Xing, W., Yu, H., Si, T. and Guo, Z. (2020) 'Detection of circulating tumor DNA from non-small cell lung cancer brain metastasis in cerebrospinal fluid samples', *Thorac Cancer,* 11(3), pp. 588-593.

Malone, E. R., Oliva, M., Sabatini, P. J. B., Stockley, T. L. and Siu, L. L. (2020) 'Molecular profiling for precision cancer therapies', *Genome Med,* 12(1), pp. 8.

Marchetti, A., Felicioni, L., Malatesta, S., Grazia Sciarrotta, M., Guetti, L., Chella, A., Viola, P., Pullara, C., Mucilli, F. and Buttitta, F. (2011) 'Clinical features and outcome of patients with non-small-cell lung cancer harboring BRAF mutations', *J Clin Oncol,* 29(26), pp. 3574-9.

Margiotta, A. (2021) 'All Good Things Must End: Termination of Receptor Tyrosine Kinase Signal', *Int J Mol Sci,* 22(12).

Mathias, C., Prado, G. F., Mascarenhas, E., Ugalde, P. A., Zimmer Gelatti, A. C., Carvalho, E. S., Faroni, L. D., Oliveira, R., Cordeiro de Lima, V. C., de Castro, G. and Torácica, G. B. d. O. (2020) 'Lung Cancer in Brazil', *J Thorac Oncol,* 15(2), pp. 170-175.

Meddeb, R., Pisareva, E. and Thierry, A. R. (2019) 'Guidelines for the Preanalytical Conditions for Analyzing Circulating Cell-Free DNA', *Clin Chem,* 65(5), pp. 623-633.

Medves, S. and Demoulin, J. B. (2012) 'Tyrosine kinase gene fusions in cancer: translating mechanisms into targeted therapies', *J Cell Mol Med,* 16(2), pp. 237-48.

Midha, A., Dearden, S. and McCormack, R. (2015) 'EGFR mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII)', *Am J Cancer Res,* 5(9), pp. 2892-911.

Minari, R., Bordi, P. and Tiseo, M. (2016) 'Third-generation epidermal growth factor receptor-tyrosine kinase inhibitors in T790M-positive non-small cell lung cancer: review on emerged mechanisms of resistance', *Transl Lung Cancer Res,* 5(6), pp. 695-708.

Mok, T. S., Wu, Y. L., Ahn, M. J., Garassino, M. C., Kim, H. R., Ramalingam, S. S., Shepherd, F. A., He, Y., Akamatsu, H., Theelen, W. S., Lee, C. K., Sebastian, M., Templeton, A., Mann, H., Marotti, M., Ghiorghiu, S., Papadimitrakopoulou, V. A. and Investigators, A. (2017) 'Osimertinib or Platinum-Pemetrexed in EGFR T790M-Positive Lung Cancer', *N Engl J Med,* 376(7), pp. 629-640.

Morgillo, F., Della Corte, C. M., Fasano, M. and Ciardiello, F. (2016) 'Mechanisms of resistance to EGFR-targeted drugs: lung cancer', *ESMO Open,* 1(3), pp. e000060.

Mukherji, D., Murillo, R. H., Van Hemelrijck, M., Vanderpuye, V., Shamieh, O., Torode, J., Pramesh, C. S., Yusuf, A., Booth, C. M., Aggarwal, A., Sullivan, R. and Force, C.-a. C. T. (2021) 'Global cancer research in the post-pandemic world', *Lancet Oncol,* 22(12), pp. 1652-1654.

Mullard, A. (2018) 'FDA approves landmark tissue-agnostic cancer drug', *Nat Rev Drug Discov,* 18(1), pp. 7.

Nakatani, K., Yamaoka, T., Ohba, M., Fujita, K. I., Arata, S., Kusumoto, S., Taki-Takemoto, I., Kamei, D., Iwai, S., Tsurutani, J. and Ohmori, T. (2019) 'and', *Mol Cancer Ther,* 18(1), pp. 112-126.

Nakatsuka, T., Nakagawa, H., Hayata, Y., Wake, T., Yamada, T., Nishibatake Kinoshita, M., Nakagomi, R., Sato, M., Minami, T., Uchino, K., Enooku, K., Kudo, Y., Tanaka, Y., Kishikawa, T., Otsuka, M., Tateishi, R. and Koike, K. (2021) 'Post-treatment cell-free DNA as a predictive biomarker in molecular-targeted therapy of hepatocellular carcinoma', *J Gastroenterol,* 56(5), pp. 456-469.

NCCN, N. C. C. N. 2022. Lung cancer (version 3.2022).

Network, C. G. A. R. (2014) 'Comprehensive molecular profiling of lung adenocarcinoma', *Nature,* 511(7511), pp. 543-50.

Ohashi, K., Sequist, L. V., Arcila, M. E., Lovly, C. M., Chen, X., Rudin, C. M., Moran, T., Camidge, D. R., Vnencak-Jones, C. L., Berry, L., Pan, Y., Sasaki, H., Engelman, J. A., Garon, E. B., Dubinett, S. M., Franklin, W. A., Riely, G. J., Sos, M. L., Kris, M. G., Dias-Santagata, D., Ladanyi, M., Bunn, P. A. and Pao, W. (2013) 'Characteristics of lung cancers harboring NRAS mutations', *Clin Cancer Res,* 19(9), pp. 2584-91.

Okabe, T., Okamoto, I., Tamura, K., Terashima, M., Yoshida, T., Satoh, T., Takada, M., Fukuoka, M. and Nakagawa, K. (2007) 'Differential constitutive activation of the epidermal growth factor receptor in non-small cell lung cancer cells bearing EGFR gene mutation and amplification', *Cancer Res,* 67(5), pp. 2046-53.

Oxnard, G. R., Hu, Y., Mileham, K. F., Husain, H., Costa, D. B., Tracy, P., Feeney, N., Sholl, L. M., Dahlberg, S. E., Redig, A. J., Kwiatkowski, D. J., Rabin, M. S., Paweletz, C. P., Thress, K. S. and Jänne, P. A. (2018) 'Assessment of Resistance Mechanisms and Clinical Implications in Patients With EGFR T790M-Positive Lung Cancer and Acquired Resistance to Osimertinib', *JAMA Oncol,* 4(11), pp. 1527-1534.

Ozlü, T. and Bülbül, Y. (2005) 'Smoking and lung cancer', *Tuberk Toraks,* 53(2), pp. 200-9.

Palacio, S., Pontes, L., Prado, E., Arshad, J., Ali, R., Piha, T., Bacchi, C. E., Mudad, R. and Lopes, G. (2019) 'Mutation Testing: Changing Patterns of Molecular Testing in Brazil', *Oncologist,* 24(4), pp. e137-e141.

Pan, Y., Zhang, Y., Li, Y., Hu, H., Wang, L., Li, H., Wang, R., Ye, T., Luo, X., Li, B., Cai, D., Shen, L., Sun, Y. and Chen, H. (2014) 'ALK, ROS1 and RET fusions in 1139 lung adenocarcinomas: a comprehensive study of common and fusion pattern-specific clinicopathologic, histologic and cytologic features', *Lung Cancer,* 84(2), pp. 121-6.

Pao, W., Miller, V. A., Politi, K. A., Riely, G. J., Somwar, R., Zakowski, M. F., Kris, M. G. and Varmus, H. (2005a) 'Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain', *PLoS Med,* 2(3), pp. e73.

Pao, W., Wang, T. Y., Riely, G. J., Miller, V. A., Pan, Q., Ladanyi, M., Zakowski, M. F., Heelan, R. T., Kris, M. G. and Varmus, H. E. (2005b) 'KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib', *PLoS Med,* 2(1), pp. e17.

Patel, M. I., Schupp, C. W., Gomez, S. L., Chang, E. T. and Wakelee, H. A. (2013) 'How do social factors explain outcomes in non-small-cell lung cancer among Hispanics in California? Explaining the Hispanic paradox', *J Clin Oncol,* 31(28), pp. 3572-8.

Piotrowska, Z., Isozaki, H., Lennerz, J. K., Gainor, J. F., Lennes, I. T., Zhu, V. W., Marcoux, N., Banwait, M. K., Digumarthy, S. R., Su, W., Yoda, S., Riley, A. K., Nangia, V., Lin, J. J., Nagy, R. J., Lanman, R. B., Dias-Santagata, D., Mino-Kenudson, M., Iafrate, A. J., Heist, R. S., Shaw, A. T., Evans, E. K., Clifford, C., Ou, S. I., Wolf, B., Hata, A. N. and Sequist, L. V. (2018) 'Landscape of Acquired Resistance to Osimertinib in', *Cancer Discov,* 8(12), pp. 1529-1539.

Pirker, R. and Filipits, M. (2016) 'Personalized treatment of advanced non-small-cell lung cancer in routine clinical practice', *Cancer Metastasis Rev,* 35(1), pp. 141-50.

Politi, K. and Herbst, R. S. (2015) 'Lung cancer in the era of precision medicine', *Clin Cancer Res,* 21(10), pp. 2213-20.

Poole, J. C., Wu, S. F., Lu, T. T., Vibat, C. R. T., Pham, A., Samuelsz, E., Patel, M., Chen, J., Daher, T., Singh, V. M. and Arnold, L. J. (2019) 'Analytical validation of the Target Selector ctDNA platform featuring single copy detection sensitivity for clinically actionable EGFR, BRAF, and KRAS mutations', *PLoS One,* 14(10), pp. e0223112.

Porter, A. C. and Vaillancourt, R. R. (1998) 'Tyrosine kinase receptor-activated signal transduction pathways which lead to oncogenesis', *Oncogene,* 17(11 Reviews), pp. 1343-52.

Pottier, C., Fresnais, M., Gilon, M., Jérusalem, G., Longuespée, R. and Sounni, N. E. (2020) 'Tyrosine Kinase Inhibitors in Cancer: Breakthrough and Challenges of Targeted Therapy', *Cancers (Basel),* 12(3).

Ramalingam, S. S., Yang, J. C., Lee, C. K., Kurata, T., Kim, D. W., John, T., Nogami, N., Ohe, Y., Mann, H., Rukazenkov, Y., Ghiorghiu, S., Stetson, D., Markovets, A., Barrett, J. C., Thress, K. S. and Jänne, P. A. (2018) 'Osimertinib As First-Line Treatment of EGFR Mutation-Positive Advanced Non-Small-Cell Lung Cancer', *J Clin Oncol,* 36(9), pp. 841-849.

Razavi, P., Li, B. T., Brown, D. N., Jung, B., Hubbell, E., Shen, R., Abida, W., Juluru, K., De Bruijn, I., Hou, C., Venn, O., Lim, R., Anand, A., Maddala, T., Gnerre, S., Vijaya Satya, R., Liu, Q., Shen, L., Eattock, N., Yue, J., Blocker, A. W., Lee, M., Sehnert, A., Xu, H., Hall, M. P., Santiago-Zayas, A., Novotny, W. F., Isbell, J. M., Rusch, V. W., Plitas, G., Heerdt, A. S., Ladanyi, M., Hyman, D. M., Jones, D. R., Morrow, M., Riely, G. J., Scher, H. I., Rudin, C. M., Robson, M. E., Diaz, L. A., Solit, D. B., Aravanis, A. M. and Reis-Filho, J. S. (2019) 'High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants', *Nat Med,* 25(12), pp. 1928-1937.

Reck, M., Rodríguez-Abreu, D., Robinson, A. G., Hui, R., Csőszi, T., Fülöp, A., Gottfried, M., Peled, N., Tafreshi, A., Cuffe, S., O'Brien, M., Rao, S., Hotta, K., Leiby, M. A., Lubiniecki, G. M., Shentu, Y., Rangwala, R., Brahmer, J. R. and Investigators, K.-. (2016) 'Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer', *N Engl J Med,* 375(19), pp. 1823-1833.

Recondo, G., Bahcall, M., Spurr, L. F., Che, J., Ricciuti, B., Leonardi, G. C., Lo, Y. C., Li, Y. Y., Lamberti, G., Nguyen, T., Milan, M. S. D., Venkatraman, D., Umeton, R., Paweletz, C. P., Albayrak, A., Cherniack, A. D., Price, K. S., Fairclough, S. R., Nishino, M., Sholl, L. M., Oxnard, G. R., Jänne, P. A. and Awad, M. M. (2020) 'Molecular Mechanisms of Acquired Resistance to MET Tyrosine Kinase Inhibitors in Patients with MET Exon 14-Mutant NSCLC', *Clin Cancer Res,* 26(11), pp. 2615-2625.

Remon, J., Besse, B., Lacroix, L., Mezquita, L., Jovelet, C., Howarth, K., Plagnol, V., Morris, C. D., Nicotra, C., Pannet, C., Green, E. and Soria, J.-C. (2017) 'Evaluation of liquid biopsies for molecular profiling in patients with advanced non-small cell lung cancer (NSCLC) in the relapse treatment setting', *Journal of Clinical Oncology,* 35(15_suppl), pp. 11532-11532.

Remon, J., Swalduz, A., Planchard, D., Ortiz-Cuaran, S., Mezquita, L., Lacroix, L., Jovelet, C., Rouleau, E., Leonce, C., De Kievit, F., Morris, C., Jones, G., Mercier, K., Howarth, K., Green, E., Pérol, M., Saintigny, P. and Besse, B. (2020) 'Outcomes in oncogenic-addicted advanced NSCLC patients with actionable mutations identified by liquid biopsy genomic profiling using a tagged amplicon-based NGS assay', *PLoS One,* 15(6), pp. e0234302.

Ridge, C. A., McErlean, A. M. and Ginsberg, M. S. (2013) 'Epidemiology of lung cancer', *Semin Intervent Radiol,* 30(2), pp. 93-8.

Rizvi, N. A., Mazières, J., Planchard, D., Stinchcombe, T. E., Dy, G. K., Antonia, S. J., Horn, L., Lena, H., Minenza, E., Mennecier, B., Otterson, G. A., Campos, L. T., Gandara, D. R., Levy, B. P., Nair, S. G., Zalcman, G., Wolf, J., Souquet, P. J., Baldini, E., Cappuzzo, F., Chouaid, C., Dowlati, A., Sanborn, R., Lopez-Chavez, A., Grohe, C., Huber, R. M., Harbison, C. T., Baudelet, C., Lestini, B. J. and Ramalingam, S. S. (2015) 'Activity and safety of nivolumab, an anti-PD-1 immune checkpoint inhibitor, for patients with advanced, refractory squamous non-small-cell lung cancer (CheckMate 063): a phase 2, single-arm trial', *Lancet Oncol,* 16(3), pp. 257-65.

Rohanizadegan, M. (2018) 'Analysis of circulating tumor DNA in breast cancer as a diagnostic and prognostic biomarker', *Cancer Genet,* 228-229, pp. 159-168.

Rosell, R., Molina, M. A., Costa, C., Simonetti, S., Gimenez-Capitan, A., Bertran-Alamillo, J., Mayo, C., Moran, T., Mendez, P., Cardenal, F., Isla, D., Provencio, M., Cobo, M., Insa, A., Garcia-Campelo, R., Reguart, N., Majem, M., Viteri, S., Carcereny, E., Porta, R., Massuti, B., Queralt, C., de Aguirre, I., Sanchez, J. M., Sanchez-Ronco, M., Mate, J. L., Ariza, A., Benlloch, S., Sanchez, J. J., Bivona, T. G., Sawyers, C. L. and Taron, M. (2011) 'Pretreatment EGFR T790M mutation and BRCA1 mRNA expression in erlotinib-treated advanced non-small-cell lung cancer patients with EGFR mutations', *Clin Cancer Res,* 17(5), pp. 1160-8.

Santoni-Rugiu, E., Melchior, L. C., Urbanska, E. M., Jakobsen, J. N., Stricker, K., Grauslund, M. and Sørensen, J. B. (2019) 'Intrinsic resistance to EGFR-Tyrosine Kinase Inhibitors in', *Cancers (Basel),* 11(7).

Scagliotti, G. V., Parikh, P., von Pawel, J., Biesma, B., Vansteenkiste, J., Manegold, C., Serwatowski, P., Gatzemeier, U., Digumarti, R., Zukin, M., Lee, J. S., Mellemgaard, A., Park, K., Patil, S., Rolski, J., Goksel, T., de Marinis, F., Simms, L., Sugarman, K. P. and Gandara, D. (2008) 'Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer', *J Clin Oncol,* 26(21), pp. 3543-51.

Schou, J. V., Larsen, F. O., Sørensen, B. S., Abrantes, R., Boysen, A. K., Johansen, J. S., Jensen, B. V., Nielsen, D. L. and Spindler, K. L. (2018) 'Circulating cell-free DNA as predictor of treatment failure after neoadjuvant chemo-radiotherapy before surgery in patients with locally advanced rectal cancer', *Ann Oncol,* 29(3), pp. 610-615.

Schrock, A. B., Zhu, V. W., Hsieh, W. S., Madison, R., Creelan, B., Silberberg, J., Costin, D., Bharne, A., Bonta, I., Bosemani, T., Nikolinakos, P., Ross, J. S., Miller, V. A., Ali, S. M., Klempner, S. J. and Ou, S. I. (2018) 'Receptor Tyrosine Kinase Fusions

and BRAF Kinase Fusions are Rare but Actionable Resistance Mechanisms to EGFR Tyrosine Kinase Inhibitors', *J Thorac Oncol,* 13(9), pp. 1312-1323.
Schwarzenbach, H., Hoon, D. S. and Pantel, K. (2011) 'Cell-free nucleic acids as biomarkers in cancer patients', *Nat Rev Cancer,* 11(6), pp. 426-37.
SEER (2019) 'Cancer Stat Facts: Lung and Bronchus Cancer 5-year survival rates (2009-2015)'. Available at: https://seer.cancer.gov/statfacts/html/lungb.html (Accessed.
Sequist, L. V., Waltman, B. A., Dias-Santagata, D., Digumarthy, S., Turke, A. B., Fidias, P., Bergethon, K., Shaw, A. T., Gettinger, S., Cosper, A. K., Akhavanfard, S., Heist, R. S., Temel, J., Christensen, J. G., Wain, J. C., Lynch, T. J., Vernovsky, K., Mark, E. J., Lanuti, M., Iafrate, A. J., Mino-Kenudson, M. and Engelman, J. A. (2011) 'Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors', *Sci Transl Med,* 3(75), pp. 75ra26.
Shim, H. S., Lee, D. H., Park, E. J. and Kim, S. H. (2011) 'Histopathologic characteristics of lung adenocarcinomas with epidermal growth factor receptor mutations in the International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society lung adenocarcinoma classification', *Arch Pathol Lab Med,* 135(10), pp. 1329-34.
Siravegna, G., Marsoni, S., Siena, S. and Bardelli, A. (2017) 'Integrating liquid biopsies into the management of cancer', *Nat Rev Clin Oncol,* 14(9), pp. 531-548.
Skoulidis, F. and Heymach, J. V. (2019) 'Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy', *Nat Rev Cancer,* 19(9), pp. 495-509.
Slebos, R. J., Kibbelaar, R. E., Dalesio, O., Kooistra, A., Stam, J., Meijer, C. J., Wagenaar, S. S., Vanderschueren, R. G., van Zandwijk, N. and Mooi, W. J. (1990) 'K-ras oncogene activation as a prognostic marker in adenocarcinoma of the lung', *N Engl J Med,* 323(9), pp. 561-5.
Soda, M., Choi, Y. L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H., Bando, M., Ohno, S., Ishikawa, Y., Aburatani, H., Niki, T., Sohara, Y., Sugiyama, Y. and Mano, H. (2007) 'Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer', *Nature,* 448(7153), pp. 561-6.
Solca, F., Dahl, G., Zoephel, A., Bader, G., Sanderson, M., Klein, C., Kraemer, O., Himmelsbach, F., Haaksma, E. and Adolf, G. R. (2012) 'Target binding properties and cellular activity of afatinib (BIBW 2992), an irreversible ErbB family blocker', *J Pharmacol Exp Ther,* 343(2), pp. 342-50.
Sorenson, G. D., Pribish, D. M., Valone, F. H., Memoli, V. A., Bzik, D. J. and Yao, S. L. (1994) 'Soluble normal and mutated DNA sequences from single-copy genes in human blood', *Cancer Epidemiol Biomarkers Prev,* 3(1), pp. 67-71.
Steensma, D. P. (2018) 'Clinical consequences of clonal hematopoiesis of indeterminate potential', *Blood Adv,* 2(22), pp. 3404-3410.
Stroun, M., Anker, P., Lyautey, J., Lederrey, C. and Maurice, P. A. (1987) 'Isolation and characterization of DNA from the plasma of cancer patients', *Eur J Cancer Clin Oncol,* 23(6), pp. 707-12.
Suda, K., Tomizawa, K. and Mitsudomi, T. (2010) 'Biological and clinical significance of KRAS mutations in lung cancer: an oncogenic driver that contrasts with EGFR mutation', *Cancer Metastasis Rev,* 29(1), pp. 49-60.

Sun, S., Schiller, J. H. and Gazdar, A. F. (2007) 'Lung cancer in never smokers--a different disease', *Nat Rev Cancer,* 7(10), pp. 778-90.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A. and Bray, F. (2021) 'Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries', *CA Cancer J Clin,* 71(3), pp. 209-249.

Swami, N., Hwang, W. L., Guo, J. A., Hoffman, H., Abramowitz, M. C., Elbakouny, Z., Beltran, H., Chipidza, F., Choueiri, T., Pra, A. D., Huang, F., Kaochar, S., Kantoff, P., Kim, D. W., Kishan, A. U., Kobetz, E., Marinac, C., Mucci, L. A., Muralidhar, V., Pollack, A., Sanford, N. N., Schaeffer, E. M., Spratt, D. E., Zhao, S. G., Rebbeck, T. R., Nguyen, P. L., Feng, F. Y., Mahal, B. A. and Alshalalfa, M. (2021) 'Novel genomic signature predictive of response to immune checkpoint blockade: A pan-cancer analysis from project Genomics Evidence Neo-plasia Information Exchange (GENIE)', *Cancer Genet,* 258-259, pp. 61-68.

Takeuchi, K., Soda, M., Togashi, Y., Suzuki, R., Sakata, S., Hatano, S., Asaka, R., Hamanaka, W., Ninomiya, H., Uehara, H., Lim Choi, Y., Satoh, Y., Okumura, S., Nakagawa, K., Mano, H. and Ishikawa, Y. (2012) 'RET, ROS1 and ALK fusions in lung cancer', *Nat Med,* 18(3), pp. 378-81.

Tan, C. S., Kumarakulasinghe, N. B., Huang, Y. Q., Ang, Y. L. E., Choo, J. R., Goh, B. C. and Soo, R. A. (2018) 'Third generation EGFR TKIs: current data and future directions', *Mol Cancer,* 17(1), pp. 29.

Thompson, J. C., Yee, S. S., Troxel, A. B., Savitch, S. L., Fan, R., Balli, D., Lieberman, D. B., Morrissette, J. D., Evans, T. L., Bauml, J., Aggarwal, C., Kosteva, J. A., Alley, E., Ciunci, C., Cohen, R. B., Bagley, S., Stonehouse-Lee, S., Sherry, V. E., Gilbert, E., Langer, C., Vachani, A. and Carpenter, E. L. (2016) 'Detection of Therapeutically Targetable Driver and Resistance Mutations in Lung Cancer Patients by Next-Generation Sequencing of Cell-Free Circulating Tumor DNA', *Clin Cancer Res,* 22(23), pp. 5772-5782.

Trevethan, R. (2017) 'Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice', *Front Public Health,* 5, pp. 307.

Varughese, S., Jahangir, K. S., Simpson, C. E. and Boulmay, B. C. (2012) 'A paradigm shift in the treatment of advanced non-small cell lung cancer', *Am J Med Sci,* 344(2), pp. 147-50.

Volckmar, A. L., Sültmann, H., Riediger, A., Fioretos, T., Schirmacher, P., Endris, V., Stenzinger, A. and Dietz, S. (2018) 'A field guide for cancer diagnostics using cell-free DNA: From principles to practice and clinical applications', *Genes Chromosomes Cancer,* 57(3), pp. 123-139.

von Felden, J., Craig, A. J., Garcia-Lezana, T., Labgaa, I., Haber, P. K., D'Avola, D., Asgharpour, A., Dieterich, D., Bonaccorso, A., Torres-Martin, M., Sia, D., Sung, M. W., Tabrizian, P., Schwartz, M., Llovet, J. M. and Villanueva, A. (2021) 'Mutations in circulating tumor DNA predict primary resistance to systemic therapies in advanced hepatocellular carcinoma', *Oncogene,* 40(1), pp. 140-151.

Wan, J. C. M., Massie, C., Garcia-Corbacho, J., Mouliere, F., Brenton, J. D., Caldas, C., Pacey, S., Baird, R. and Rosenfeld, N. (2017) 'Liquid biopsies come of age: towards implementation of circulating tumour DNA', *Nat Rev Cancer,* 17(4), pp. 223-238.

Wang, S., Tsui, S. T., Liu, C., Song, Y. and Liu, D. (2016) 'EGFR C797S mutation mediates resistance to third-generation inhibitors in T790M-positive non-small cell lung cancer', *J Hematol Oncol,* 9(1), pp. 59.

Weart, T. C., Miller, K. D. and Simone, C. B. (2018) 'Spotlight on dabrafenib/trametinib in the treatment of non-small-cell lung cancer: place in therapy', *Cancer Manag Res,* 10, pp. 647-652.

Weber, S., Spiegl, B., Perakis, S. O., Ulz, C. M., Abuja, P. M., Kashofer, K., Leest, P. V., Azpurua, M. A., Tamminga, M., Brudzewsky, D., Rothwell, D. G., Mohan, S., Sartori, A., Lampignano, R., Konigshofer, Y., Sprenger-Haussels, M., Wikman, H., Bergheim, I. R., Kloten, V., Schuuring, E., Speicher, M. R. and Heitzer, E. (2020) 'Technical Evaluation of Commercial Mutation Analysis Platforms and Reference Materials for Liquid Biopsy Profiling', *Cancers (Basel),* 12(6).

Weber, S., van der Leest, P., Donker, H. C., Schlange, T., Timens, W., Tamminga, M., Hasenleithner, S. O., Graf, R., Moser, T., Spiegl, B., Yaspo, M. L., Terstappen, L. W. M. M., Sidorenkov, G., Hiltermann, T. J. N., Speicher, M. R., Schuuring, E., Heitzer, E. and Groen, H. J. M. (2021) 'Dynamic Changes of Circulating Tumor DNA Predict Clinical Outcome in Patients With Advanced Non-Small-Cell Lung Cancer Treated With Immune Checkpoint Inhibitors', *JCO Precis Oncol,* 5, pp. 1540-1553.

Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R., Ozenberger, B. A., Ellrott, K., Shmulevich, I., Sander, C., Stuart, J. M. and Network, C. G. A. R. (2013) 'The Cancer Genome Atlas Pan-Cancer analysis project', *Nat Genet,* 45(10), pp. 1113-20.

Wu, F., Fan, J., He, Y., Xiong, A., Yu, J., Li, Y., Zhang, Y., Zhao, W., Zhou, F., Li, W., Zhang, J., Zhang, X., Qiao, M., Gao, G., Chen, S., Chen, X., Li, X., Hou, L., Wu, C., Su, C., Ren, S., Odenthal, M., Buettner, R., Fang, N. and Zhou, C. (2021) 'Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer', *Nat Commun,* 12(1), pp. 2540.

Wu, S. G. and Shih, J. Y. (2018) 'Management of acquired resistance to EGFR TKI-targeted therapy in advanced non-small cell lung cancer', *Mol Cancer,* 17(1), pp. 38.

Yanagita, M., Redig, A. J., Paweletz, C. P., Dahlberg, S. E., O'Connell, A., Feeney, N., Taibi, M., Boucher, D., Oxnard, G. R., Johnson, B. E., Costa, D. B., Jackman, D. M. and Jänne, P. A. (2016) 'A Prospective Evaluation of Circulating Tumor Cells and Cell-Free DNA in EGFR-Mutant Non-Small Cell Lung Cancer Patients Treated with Erlotinib on a Phase II Trial', *Clin Cancer Res,* 22(24), pp. 6010-6020.

Yang, J., Gong, Y., Lam, V. K., Shi, Y., Guan, Y., Zhang, Y., Ji, L., Chen, Y., Zhao, Y., Qian, F., Chen, J., Li, P., Zhang, F., Wang, J., Zhang, X., Yang, L., Kopetz, S., Futreal, P. A., Zhang, J., Yi, X., Xia, X. and Yu, P. (2020a) 'Deep sequencing of circulating tumor DNA detects molecular residual disease and predicts recurrence in gastric cancer', *Cell Death Dis,* 11(5), pp. 346.

Yang, S. R., Schultheis, A. M., Yu, H., Mandelker, D., Ladanyi, M. and Büttner, R. (2020b) 'Precision medicine in non-small cell lung cancer: Current applications and future directions', *Semin Cancer Biol*.

Yang, Z., Yang, N., Ou, Q., Xiang, Y., Jiang, T., Wu, X., Bao, H., Tong, X., Wang, X., Shao, Y. W., Liu, Y., Wang, Y. and Zhou, C. (2018) 'Investigating Novel Resistance

Mechanisms to Third-Generation EGFR Tyrosine Kinase Inhibitor Osimertinib in Non-Small Cell Lung Cancer Patients', *Clin Cancer Res,* 24(13), pp. 3097-3107.

Ye, L., Mesbah Ardakani, N., Thomas, C., Spilsbury, K., Leslie, C., Amanuel, B. and Millward, M. (2020) 'Detection of Low-level EGFR c.2369 C > T (p.Thr790Met) Resistance Mutation in Pre-treatment Non-small Cell Lung Carcinomas Harboring Activating EGFR Mutations and Correlation with Clinical Outcomes', *Pathol Oncol Res,* 26(4), pp. 2371-2379.

Yu, H. A., Planchard, D. and Lovly, C. M. (2018) 'Sequencing Therapy for Genetically Defined Subgroups of Non-Small Cell Lung Cancer', *Am Soc Clin Oncol Educ Book,* 38, pp. 726-739.

Yun, J., Lee, S. H., Kim, S. Y., Jeong, S. Y., Kim, J. H., Pyo, K. H., Park, C. W., Heo, S. G., Yun, M. R., Lim, S., Lim, S. M., Hong, M. H., Kim, H. R., Thayu, M., Curtin, J. C., Knoblauch, R. E., Lorenzi, M. V., Roshak, A. and Cho, B. C. (2020) 'Antitumor Activity of Amivantamab (JNJ-61186372), an EGFR-MET Bispecific Antibody, in Diverse Models of', *Cancer Discov,* 10(8), pp. 1194-1209.

Zhang, J., Fujimoto, J., Wedge, D. C., Song, X., Seth, S., Chow, C. W., Cao, Y., Gumbs, C., Gold, K. A., Kalhor, N., Little, L., Mahadeshwar, H., Moran, C., Protopopov, A., Sun, H., Tang, J., Wu, X., Ye, Y., William, W. N., Lee, J. J., Heymach, J. V., Hong, W. K., Swisher, S., Wistuba, I. I. and Futreal, P. A. (2014) 'Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing', *Science,* 346(6206), pp. 256-9.

Zhang, T., Wan, B., Zhao, Y., Li, C., Liu, H., Lv, T., Zhan, P. and Song, Y. (2019a) 'Treatment of uncommon', *Transl Lung Cancer Res,* 8(3), pp. 302-316.

Zhang, Y., Chang, L., Yang, Y., Fang, W., Guan, Y., Wu, A., Hong, S., Zhou, H., Chen, G., Chen, X., Zhao, S., Zheng, Q., Pan, H., Zhang, L., Long, H., Yang, H., Wang, X., Wen, Z., Wang, J., Xia, X., Zhao, Y., Hou, X., Ma, Y., Zhou, T., Zhang, Z., Zhan, J., Huang, Y., Zhao, H., Zhou, N. and Yi, X. (2019b) 'Intratumor heterogeneity comparison among different subtypes of non-small-cell lung cancer through multi-region tissue and matched ctDNA sequencing', *Mol Cancer,* 18(1), pp. 7.

Zhang, Y., Yao, Y., Xu, Y., Li, L., Gong, Y., Zhang, K., Zhang, M., Guan, Y., Chang, L., Xia, X., Jia, S. and Zeng, Q. (2021) 'Pan-cancer circulating tumor DNA detection in over 10,000 Chinese patients', *Nat Commun,* 12(1), pp. 11.

Zhang, Y. L., Yuan, J. Q., Wang, K. F., Fu, X. H., Han, X. R., Threapleton, D., Yang, Z. Y., Mao, C. and Tang, J. L. (2016) 'The prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis', *Oncotarget,* 7(48), pp. 78985-78993.