

UNIVERSIDADE FEDERAL DO PARANÁ

TAMYRES MINGORANCE CARVALHO



**PCR DIGITAL PARA IDENTIFICAR MUTAÇÕES SOMÁTICAS EM PACIENTES
COM CÂNCER COLORRETAL E CÂNCER DE PULMÃO**

CURITIBA

2018

TAMYRES MINGORANCE CARVALHO

**PCR DIGITAL PARA IDENTIFICAR MUTAÇÕES SOMÁTICAS EM PACIENTES
COM CÂNCER COLORRETAL E CÂNCER DE PULMÃO**

Dissertação apresentada ao Curso de Pós-Graduação em Genética, Setor de Ciências Biológicas, Universidade Federal do Paraná, como parte das exigências para obtenção do título de Mestre em Genética.

Orientador(a): Profª Drª Angelica B.W. Boldt

Co-orientador(a): Drª Sueli M. Nakatani

Curitiba

2018

Universidade Federal do Paraná. Sistema de Bibliotecas.
Biblioteca de Ciências Biológicas.
(Telma Terezinha Stresser de Assis –CRB/9-944)

Carvalho, Tamyres Mingorance
PCR digital para identificar mutações somáticas em pacientes com câncer colorretal e câncer de pulmão. / Tamyres Mingorance Carvalho. – Curitiba, 2018.

111 p.: il. ; 30cm.

Orientadora: Angelica Beate Winter Boldt

Co-orientador: Sueli M. Nakatani

Dissertação (Mestrado) - Universidade Federal do Paraná, Setor de Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Fator de Crescimento Epidérmico. 2. Pulmões - Câncer. 3. Câncer Colorretal. 4. Reação em cadeia de polimerase. I. Título. II. Boldt, Angelica Beate Winter. III. Nakatani, Sueli M. III. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Genética.

CDD (20. ed.) 574.87328



MINISTÉRIO DA EDUCAÇÃO
SETOR CIÊNCIAS BIOLÓGICAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO GENÉTICA

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GENÉTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **TAMYRES MINGORANCE CARVALHO** intitulada: **PCR DIGITAL PARA IDENTIFICAR MUTAÇÕES SOMÁTICAS EM PACIENTES COM CÂNCER COLORRETAL E CÂNCER DE PULMÃO**, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua aprovação no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 28 de Março de 2018.

ANGÉLICA BEATE WINTER BOLDT
Presidente da Banca Examinadora (UFPR)

MARCIA HOLSBACK BELTRAME
Avaliador Externo (UFPR)

VANESSA SANTOS SOTOMAIOR
Avaliador Externo (FIOCRUZ)

GABRIEL ADELMAN CIPOLLA
Avaliador Externo (UFPR)

AGRADECIMENTOS

A palavra “gratidão” deriva de “graça”, que por sua vez se traduz como algo que nos remete a bem-estar, alegria e, por que não, felicidade. Dessa forma, sou grata a todos que direta ou indiretamente foram responsáveis pelo meu bem-estar nestes últimos dois anos.

À Universidade e ao Departamento de Genética pela oportunidade de crescimento.

À minha orientadora Professora Dr^a Angelica B.W. Boldt, pela orientação e suporte durante os anos que nos conhecemos. Agradeço imensamente a confiança no meu trabalho e por ter aceitado esse projeto inovador. Definitivamente, este foi um crescimento para nós.

À minha co-orientadora, Dr^a Sueli M. Nakatani, pela paciência em compartilhar seu vasto conhecimento em Biologia Molecular Diagnóstica, orientar e ensinar, sem medir esforços. À equipe do Laboratório Genoprimer, essencial para o desenvolvimento deste projeto. De estagiários a funcionários, todo o projeto foi possível graças ao esforço de cada um. Em especial agradeço imensamente a Renata Dourado, amiga e colega de profissão, pela coragem de abraçar esse projeto e fazer parte de tudo com muita dedicação.

Ao grupo de pesquisa do LGMH da UFPR, em especial aos amigos Luana C. Oliveira, Veronica C. Solis, Sara Lobo, Fernanda Brandão e Leonardo Amorim, por me apoiar nesse projeto e me acolher nestes últimos anos, embora eu me ausentasse por conta dos experimentos. Às amigas geneticistas Gabrielle e Nina, pelo auxílio nos conhecimentos básicos, e essenciais para o meu bem-estar emocional.

À minha família, agradeço e dedico esse trabalho, visto que são os principais responsáveis pelo meu sucesso profissional e pessoal. Aos meus pais que foram meus heróis e companheiros ao longo de toda vida, me amando e me fortalecendo em todos os momentos. Ao meu cunhado pelas tardes de estudo e conversas. E por fim ao meu companheiro Nathan pelo apoio e carinho, e pela nova rotina que estamos constituindo; estar todos os dias ao seu lado foi essencial para o meu sucesso.

“Sejamos agradecidos com as pessoas que nos fazem felizes, elas são os jardineiros encantadores que fazem a nossa própria alma florescer” (M.Proust)

“Com o tempo você aprende que as circunstâncias e os ambientes têm influência sobre nós, mas nós somos responsáveis por nós mesmos. Começa a aprender que não se deve comparar com os outros, mas com o melhor que pode ser. Descobre que se leva muito tempo para se tornar a pessoa que quer ser, e que o tempo é curto. E você aprende que realmente pode suportar; que realmente é forte, e que pode ir muito mais longe, depois de pensar que não se pode mais!”

W. Shakespeare

RESUMO

O rastreio de mutações no gene que codifica o receptor de fator de crescimento epidérmico (*EGFR*) é essencial para direcionamento terapêutico no carcinoma de pulmão de células não pequenas (NSCLC, do inglês *non-small-cell lung cancer*), enquanto que mutações no gene *KRAS* (do inglês, *Kirsten Rat Sarcoma viral oncogene homolog*) devem ser investigadas no câncer colorretal (CCR). A PCR digital (dPCR) é proposta como substituição ao sequenciamento de Sanger ou à PCR em tempo real, para detecção e quantificação de mutações raras. Com o intuito de validar a técnica de dPCR para este fim, realizamos a extração de DNA de 57 biópsias de NSCLC, 41 biópsias de CCR e 10 biópsias livres de neoplasia de pacientes brasileiros, todas fixadas em formalina e emblocadas em parafina (FFPE). Investigamos as mutações DEL19, p.L858R, p.T790M, p.L861Q, p.S768I e p.G719X nos exons 18-21 no gene *EGFR* e mutações no códon 12 e p.G13D no exon 2 no gene *KRAS* utilizando dPCR, sequenciamento de Sanger e PCR competitiva alelo-específica por TaqMan® (castPCR), e determinamos a sensibilidade e os limites de detecção (LoD) e quantificação (LoQ) da técnica de dPCR. Foram encontradas 27 amostras de câncer de pulmão com mutações somáticas no gene *EGFR* por dPCR, comparado a 11 por castPCR ($p=0.014$) e 4 por sequenciamento ($p=0.0002$). Além disso, foram encontradas 28 amostras portadoras de mutações somáticas no gene *KRAS* por dPCR, 8 destas com mutações no códon 13, em comparação a 13 amostras com mutações somáticas em *KRAS*, por sequenciamento de Sanger ($p=0.04$). Não houve diferença entre as técnicas de dPCR e castPCR para a quantificação de mutações no gene *KRAS* ($p>0.05$). LoD foi de 100 moléculas de DNA/mL, enquanto que a LoQ, 1%. Grande parte das amostras (86,96%) identificadas como portadoras do alelo mutante por dPCR, e que não tiveram tal alelo detectado por castPCR, apresentaram menos de 10% de moléculas mutadas (média 4,57%). Não foram observadas mutações nas amostras livres de neoplasia. A acurácia da técnica de dPCR foi de 94,44% para detecção de *EGFR* e 85,71% para detecção de *KRAS*, de acordo com o Colégio Americano de Patologia (CAP). Estes resultados indicam maior sensibilidade da técnica de dPCR para a investigação de mutações nos genes *EGFR* e *KRAS* em biópsias de NSCLC e CCR, comparado ao sequenciamento de Sanger e à castPCR.

Palavras-chave: *EGFR*; câncer de pulmão; *KRAS*; câncer colorretal; PCR digital.

ABSTRACT

Screening of endothelial growth factor receptor (*EGFR*) mutations is mandatory in non-small-cell lung cancer (NSCLC), whereas *KRAS* (*Kirsten Rat Sarcoma viral oncogene homolog*) mutations shall be investigated in colorectal cancer (CRC). Digital PCR (dPCR) is proposed to replace Sanger sequencing or real-time PCR for detection and quantification of rare mutations. In order to validate dPCR to detect *EGFR* somatic mutations in NSCLC and *KRAS* somatic mutations in CRC, we extracted DNA from formalin-fixed paraffin-embedded (FFPE) samples of 57 NSCLC, 41 CRC and 10 free-cancer Brazilian patients. We are looking for *EGFR* mutations in exons 18-21 and exon 2 *KRAS* mutations using dPCR, Sanger Sequencing and Competitive Allele-Specific PCR (castPCR), and determine dPCR sensitivity, limits of detection (LoD) and quantification (LoQ). We found 27 samples with *EGFR* mutations by dPCR, versus 11 by castPCR ($p=0.014$) and 4 by Sanger Sequencing ($p=0.0002$). We also found 28 samples with *KRAS* somatic mutations by dPCR, 8 of which had mutations in codon 13, comparing to 13 by Sanger Sequencing ($p=0.04$). We found no difference between quantification of *KRAS* mutations by dPCR and castPCR ($p>0.05$). LoD was 100 molecules of DNA/mL and LoQ, 1%. Most of the samples (86.96%) identified by castPCR as wild-type and by dPCR as mutated, presented less than 10% mutated DNA molecules (mean 4.57%). We found no mutations in free-cancer samples. Accuracy was 94.44% for *EGFR* detection and 85.71% for *KRAS* detection, as measured with the assay recommended by the College of American Pathologist. These results indicated higher sensitivity and specificity of dPCR for screening *EGFR* and *KRAS* mutations in NSCLC and CRC biopsies, compared with Sanger Sequencing and castPCR.

Keywords: *EGFR*; lung cancer; *KRAS*; colorectal cancer; digital PCR.

LISTA DE FIGURAS

FIGURA 1. VIAS DE SINALIZAÇÃO DE EGFR.....	18
FIGURA 2. SEQUENCIAMENTO DE DNA PELO MÉTODO DE SANGER.....	28
FIGURA 3. SISTEMA DE DETECÇÃO TAQMAN®.....	32
FIGURA 4. DETECÇÃO DE MUTAÇÃO NOS GENES <i>EGFR</i> E <i>KRAS</i> POR CASTPCR.....	33
FIGURA 5. DETECÇÃO DE VARIANTES PELA TÉCNICA DE PCR DIGITAL.....	35
FIGURA 6. COMPARAÇÃO DOS PROCEDIMENTOS DE PCR EM TEMPO REAL CONVENCIONAL E PCR DIGITAL.....	38
FIGURA 7. SISTEMA DE DETECÇÃO PARA AS MUTAÇÕES NO <i>rs28929495</i> NO EXON 18 DO GENE <i>EGFR</i>	45
FIGURA 8. SISTEMA DE DETECÇÃO DA DELEÇÃO DO EXON 19 DO GENE <i>EGFR</i>	46
FIGURA 9. ETAPAS DO PROTOCOLO DE (A) REMOÇÃO DE PARAFINA SEGUIDO DE (B) EXTRAÇÃO DE DNA DAS AMOSTRAS DE CÂNCER DE PULMÃO E CÂNCER COLORRETAL.....	50
FIGURA 10. EFICIÊNCIA DAS SONDAS PARA DETECÇÃO DAS VARIANTES....	57
FIGURA 11. DISTRIBUIÇÃO DAS CONCENTRAÇÕES DETECTADAS PELA DPCR DE CONTROLES SINTÉTICOS CONTENDO AS VARIANTES COMUNS DE <i>EGFR</i> E <i>KRAS</i> DE ACORDO COM OS SEUS NÍVEIS DE DILUIÇÃO.....	58
FIGURA 12. DISTRIBUIÇÃO DAS CONCENTRAÇÕES DETECTADAS PELA DPCR DE CONTROLES SINTÉTICOS CONTENDO AS VARIANTES COMUNS DE <i>EGFR</i> E <i>KRAS</i> DE ACORDO COM OS SISTEMAS DE DETECÇÃO.....	59
FIGURA 13. REGRESSÃO LINEAR PARA ANÁLISE DE LIMITE DE QUANTIFICAÇÃO (LOQ) DO MÉTODO DE DPCR PARA OS SISTEMAS DE DETECÇÃO DAS VARIANTES p.T790M e p.L861Q DE <i>EGFR</i>	60
FIGURA 14. IMAGEM REPRESENTATIVA DE AMOSTRA DE CÂNCER DE PULMÃO AVALIADA PELO SOFTWARE <i>QUANTSTUDIO™ 3D ANALYSIS SUITE</i> PORTADORA DE MUTAÇÃO DEL19.....	63
FIGURA 15. NÚMERO DE MUTAÇÕES SOMÁTICAS DETECTADAS NO GENE <i>KRAS</i> NAS AMOSTRAS DE CÂNCER COLORRETAL POR DPCR E CASTPCR....	65
FIGURA 16. CORRELAÇÃO DAS FREQUÊNCIAS DE MUTAÇÕES SOMÁTICAS DETECTADAS NOS GENES <i>EGFR</i> E <i>KRAS</i> POR DPCR E CASTPCR.....	67

FIGURA 17. FREQUÊNCIA DE MUTAÇÕES SOMÁTICAS NOS GENES EGFR E KRAS DAS AMOSTRAS DE CÂNCER DE PULMÃO E CÂNCER COLORRETAL, CUJA MUTAÇÃO FOI DETECTADA APENAS POR DPCR, EM COMPARAÇÃO ÀQUELAS DETECTADAS POR CASTPCR.....	69
FIGURA 18. ANÁLISE DE LIMITE DE QUANTIFICAÇÃO (LOQ) DAS AMOSTRAS CONTENDO AS VARIANTES RARAS AVALIADAS PARA EGFR E KRAS POR DPCR.....	70
FIGURA 19. RESULTADO REPRESENTATIVO DE UMA AMOSTRA DE TECIDO LIVRE DE NEOPLASIA POR DPCR.....	71
FIGURA 20. FLUXOGRAMA DO FLUXO DE TRABALHO NA OCASIÃO DO RECEBIMENTO DE AMOSTRAS DE CÂNCER DE PULMÃO OU CÂNCER COLORRETAL EM FFPE.....	72

LISTA DE TABELAS

TABELA 1. FREQUÊNCIA DE MUTAÇÕES CLÁSSICAS, NÃO CLÁSSICAS E COMPLEXAS NO GENE EGFR, ASSOCIADA AO ADENOCARCINOMA DE PULMÃO E À RESISTÊNCIA AO TRATAMENTO COM INIBIDORES DE EGFR.....	21
TABELA 2. FREQUÊNCIA DE MUTAÇÕES SOMÁTICAS NO GENE KRAS ASSOCIADAS AO ADENOCARCINOMA COLORRETAL E DE RESISTÊNCIA AO TRATAMENTO.....	24
TABELA 3. TÉCNICAS COMUMENTE UTILIZADAS NA IDENTIFICAÇÃO DE MUTAÇÕES SOMÁTICAS NA ROTINA DIAGNÓSTICA.....	26
TABELA 4. PRINCIPAIS PLATAFORMAS DE PCR DIGITAL.....	37
TABELA 5. MISTURA DE REAÇÃO PARA REALIZAÇÃO DOS EXPERIMENTOS DE PCR DIGITAL.....	43
TABELA 6. PROTOCOLO DE CICLAGEM DOS EXPERIMENTOS DE PCR DIGITAL.....	43
TABELA 7. INFORMAÇÕES RELATIVAS A SEXO E IDADE DOS PACIENTES, ASSIM COMO GRAU DE DIFERENCIAÇÃO TUMORAL, PARA CÂNCER DE PULMÃO E CÂNCER COLORRETAL.....	48
TABELA 8. MISTURA DE REAÇÃO PARA A REALIZAÇÃO DOS EXPERIMENTOS DE AMPLIFICAÇÃO DOS FRAGMENTOS PELA DNA POLIMERASE.....	52
TABELA 9. PROTOCOLO DE CICLAGEM DOS EXPERIMENTOS DE AMPLIFICAÇÃO DOS FRAGMENTOS PELA DNA POLIMERASE.....	52
TABELA 10. PROTOCOLO DE CICLAGEM DAS REAÇÕES DE SEQUENCIAMENTO DE SANGER.....	53
TABELA 11. MISTURA DE REAÇÃO PARA A REALIZAÇÃO DOS EXPERIMENTOS DE PCR COMPETITIVA ALELO ESPECÍFICA.....	54
TABELA 12. PROTOCOLO DE CICLAGEM DOS EXPERIMENTOS DE PCR COMPETITIVA ALELO ESPECÍFICA.....	54
TABELA 13. VALORES DE CICLO DE AMPLIFICAÇÃO (CT) PARA DETECÇÃO DE MUTAÇÕES EM EGFR POR CASTPCR.....	61
TABELA 14. NÚMERO DE AMOSTRAS DE CÂNCER DE PULMÃO COM MUTAÇÕES SOMÁTICAS NO GENE EGFR DETECTADAS PELAS TÉCNICAS DE: DPCR, CASTPCR E SEQUENCIAMENTO DE SANGER.....	62

TABELA 15. MUTAÇÕES SOMÁTICAS DETECTADAS NO GENE EGFR NAS AMOSTRAS DE CÂNCER DE PULMÃO PELAS DIFERENTES TÉCNICAS UTILIZADAS.....	64
TABELA 16. VALORES DE CICLO DE AMPLIFICAÇÃO (CT) OBTIDOS ATRAVÉS DA DETECÇÃO DE MUTAÇÕES NO GENE KRAS PELA TÉCNICA DE CASTPCR.....	66
TABELA 17. NÚMERO DE AMOSTRAS DE CÂNCER COLORRETAL COM MUTAÇÕES SOMÁTICAS NO GENE KRAS DETECTADAS PELAS TÉCNICAS DE DPCR, CASTPCR, E SEQUENCIAMENTO DE SANGER.....	67
TABELA 18. MUTAÇÕES SOMÁTICAS DETECTADAS NO GENE KRAS NAS AMOSTAS DE CÂNCER COLORRETAL PELAS DIFERENTES TÉCNICAS UTILIZADAS.....	68
TABELA 19. RESULTADOS DA TIPAGEM DOS CONTROLES PARA MUTAÇÕES NO GENE EGFR, PARA ACREDITAÇÃO INTERNACIONAL DE ACORDO COM O COLÉGIO AMERICANO DE PATOLOGIA (CAP).....	73
TABELA 20. RESULTADOS DA TIPAGEM DOS CONTROLES PARA MUTAÇÕES NO GENE KRAS, PARA ACREDITAÇÃO INTERNACIONAL DE ACORDO COM O COLÉGIO AMERICANO DE PATOLOGIA (CAP).....	73
TABELA 21. CUSTO CALCULADO NA UTILIZAÇÃO DAS TÉCNICAS DE DPCR, CASTPCR E SEQUENCIAMENTO DE SANGER PARA DETECÇÃO DE MUTAÇÕES SOMÁTICAS NOS GENES EGFR E KRAS.....	74

LISTA DE ABREVIACOES

BCL2 - clula B de linfoma 2.

BEAM - do ingls, *Beads, Emulsion, Amplification and Magnetics*.

BIN - protena de membrana inibida por BLC2.

BRAF – oncogene B1 de sarcoma viral murino (do ingls, *V-Raf murine sarcoma viral oncogene B1*).

castPCR - PCR competitiva em tempo real por TaqMan®.

CCR - cncer colorretal.

CpG - citosina-fosfato-guanina.

CTC - clulas tumorais circulantes.

ctDNA - DNA tumoral circulante (do ingls, *circulating tumor DNA*).

ddNTPs – didesoxinucleotdeos.

dNTPs – desoxinucleotdeos.

dPCR - PCR digital.

EGF - fator de crescimento epidrmico.

EGFR - receptor do fator de crescimento epidrmico (do ingls, *epidermal growth factor receptor*).

EGFR-TKI - inibidor de tirosina quinase especfico para o receptor do fator de crescimento epidrmico.

EMA - *European Medicines Agency*.

EMT - protena epitlio mesnquima.

ERK - cinase regulada por sinal extracelular.

FDA - *Food and Drug Administration*.

FFPE - amostras fixadas em formalina e emblocadas em parafina.

GAS6 - protena *growth arrest-specific 6*.

HER - receptor de fator de crescimento epidrmico humano.

IGFR - fator de crescimento do tipo insulina.

IMS - instabilidade de microssatlites.

INC - instabilidade cromossmica.

JAK - protena janus cinase.

KRAS – oncogene homlogo ao sarcoma viral de rato (do ingls, *Kirsten rat sarcoma viral oncogene homolog*).

LoD - limite de deteco.

LoQ - limite de quantificao.

MAPK - cinase ativada por mitógeno.

MEK - cinase ativadora de MAP cinase.

MGB - (do inglês, *minor groove binder*)

mTOR - proteína alvo da rapamicina em mamíferos.

NF-kB - fator nuclear kappa beta.

NGS - sequenciamento de nova geração (do inglês, *next generation sequencing*)

NSCLC - câncer de pulmão são do tipo não pequenas células (do inglês, *non small cell lung cancer*).

p53 - proteína 53.

PCR - reação em cadeia da polimerase.

PI3K - fosfatidilinositol 3 cinase.

PIK3CA - fosfatidilinositol 3 cinase subunidade catalítica alfa.

PLC - fosfolipase C.

PPi - pirofosfato.

PTEN - homólogo de fosfatase tensina.

qPCR - PCR em tempo real.

RAF - cinase fibrosarcoma rapidamente acelerado.

RAS - oncogene homólogo de sarcoma viral de rato.

RNA-seq - sequenciamento direto ou de transcriptoma.

SOS - proteína derivada do gene filho de *sevenless*.

SRC - cinase SRC, abreviação de sarcoma.

STAT - transdutor de sinal e ativador de transcrição.

TKI - inibidor de tirosina-quinase (do inglês, *tyrosine kinase inhibitor*).

VEGFR - receptor de fator de crescimento endotelial vascular.

WES - sequenciamento completo de exoma (do inglês, *whole-exome sequencing*).

WGS - sequenciamento inteiro de genoma (do inglês *whole-genome sequencing*).

SUMÁRIO

1. INTRODUÇÃO	16
2. REVISÃO DE LITERATURA	17
2.1. ONCOGÊNESE E AS VIAS DE SINALIZAÇÃO EGFR E KRAS	17
2.1.1. Câncer de pulmão e EGFR.	19
2.1.2. Câncer colorretal e KRAS	22
2.2. TÉCNICAS PARA IDENTIFICAÇÃO DE MUTAÇÕES SOMÁTICAS NA ROTINA DIAGNÓSTICA	24
2.2.1. Identificação de mutações somáticas por sequenciamento	27
2.2.2. Identificação de mutações somáticas por PCR em tempo real (qPCR)	31
2.2.3. Identificação de mutações somáticas por PCR digital (dPCR)	34
3. HIPÓTESE E JUSTIFICATIVA	40
4. OBJETIVOS	41
4.1. OBJETIVO GERAL	41
4.2. OBJETIVOS ESPECÍFICOS	41
5. MATERIAIS E MÉTODOS	41
5.1. COMITÊ DE ÉTICA	41
5.2. VALIDAÇÃO ANALÍTICA	41
5.2.1. Produção do DNA sintético como controle positivo para mutações nos genes <i>EGFR</i> e <i>KRAS</i>	42
5.3. PCR DIGITAL	42
5.3.1. Detecção de mutações de ponto nos genes EGFR e KRAS	44
5.3.2. Detecção da deleção do éxon 19 do gene EGFR	45
5.3.3. Limite de detecção, eficiência e linearidade	46
5.4. VALIDAÇÃO DIAGNÓSTICA	47
5.4.1. Amostras	47
5.4.2. Extração de DNA	48
5.4.3. Sequenciamento de Sanger	51
5.4.3.1. Amplificação dos fragmentos pela DNA polimerase	51
5.4.3.2. Purificação dos produtos amplificados	52
5.4.3.3. Sequenciamento de DNA	52
5.4.4. PCR competitiva alelo específica (CASTPCR)	53
5.5. ANÁLISE ESTATÍSTICA	55

6. RESULTADOS	56
6.1. VALIDAÇÃO ANALÍTICA	56
6.2. VALIDAÇÃO DIAGNÓSTICA	60
6.2.1. Análises qualitativas	61
6.2.1.1. Câncer de pulmão e EGFR	61
6.2.1.2. Câncer colorretal e KRAS	65
6.2.2. Análises quantitativas	69
6.3. CONTROLE DE QUALIDADE E ACREDITAÇÃO INTERNACIONAL	72
6.4. CUSTO EFETIVO DA REALIZAÇÃO DAS TÉCNICAS	74
7. DISCUSSÃO	75
REFERÊNCIAS	84
ANEXOS	92
ANEXO 1. OLIGONUCLEOTÍDEOS SINTÉTICOS UTILIZADOS COMO CONTROLES POSITIVOS	92
ANEXO 2. OLIGONUCLEOTÍDEOS INICIADORES E SONDAS DE HIDRÓLISE DOS SISTEMAS DE DETECÇÃO DE MUTAÇÕES DO GENE <i>EGFR</i> .	94
ANEXO 3. OLIGONUCLEOTÍDEOS INICIADORES E SONDAS DE HIDRÓLISE DOS SISTEMAS DE DETECÇÃO DE MUTAÇÕES DO GENE <i>KRAS</i> .	95
ANEXO 4. MUTAÇÕES SOMÁTICAS DETECTADAS NO GENE <i>EGFR</i> NAS AMOSTRAS DE CÂNCER DE PULMÃO ANALISADAS POR PCR DIGITAL, <i>CASTPCR</i> E SEQUENCIAMENTO.	96
ANEXO 5. PRESENÇA DE MUTAÇÃO NO GENE <i>KRAS</i> NAS AMOSTRAS DE CÂNCER COLORRETAL ANALISADAS POR PCR DIGITAL, <i>CASTPCR</i> E SEQUENCIAMENTO DE SANGER.	98
APÊNDICES	100
APÊNDICE 1. LAUDO REPRESENTATIVO DE PACIENTE COM AUSÊNCIA DE MUTAÇÃO NO GENE <i>KRAS</i> .	100
APÊNDICE 2. LAUDO REPRESENTATIVO DE PACIENTE COM AUSÊNCIA DE MUTAÇÃO NO GENE <i>EGFR</i> .	101
APÊNDICE 3. ARTIGO EM DESENVOLVIMENTO.	102

1. INTRODUÇÃO

O câncer de pulmão e o câncer colorretal são neoplasias que podem apresentar mutações em diferentes componentes da via de ativação de tirosina-quinase, incluindo os genes *EGFR* (do inglês, *Epidermal Growth Factor Receptor*) e *KRAS* (do inglês, *Kirsten Rat Sarcoma viral oncogene homolog*). Mutações no gene *EGFR* resultam em ativação constitutiva da atividade tirosina-quinase do receptor, contribuindo para o processo oncogênico (LICHTENBERGER et al., 2010), e a detecção precoce de mutações é essencial para a escolha do tratamento ideal e melhor prognóstico da doença (FERLAY et al., 2010). Logo, na presença de mutações somáticas em *EGFR*, diretrizes clínicas atuais recomendam que pacientes com câncer de pulmão usem como primeira escolha, a terapia com inibidor de tirosina-quinase (TKI, do inglês *tyrosine kinase inhibitor*). Em contrapartida, pacientes portadores de mutações no gene *KRAS* não são passíveis de receber tratamento com EGFR-TKIs, visto que a proteína *KRAS* é uma molécula intermediária envolvida na cascata de sinalização intracelular mediada pelo receptor *EGFR*. Por essa razão, o bloqueio de *EGFR* não inibe as etapas seguintes da via (ROOCK et al., 2011).

Atualmente, o método padrão-ouro para detecção de mutações em neoplasias é o sequenciamento de DNA (DIDELLOT et al., 2012), extraído de biópsias teciduais. No entanto, tal técnica requer que pelo menos 10% das células contenham a variante de interesse, dentre a população celular estudada (FRANCIS; STEIN, 2015). Em situações como esta, métodos que empreguem amplificação privilegiada das variantes raras e/ou baseados em amplificação de moléculas de DNA isoladas por reação em cadeia em cadeia da polimerase (PCR), têm se mostrado promissores para detecção de mutações somáticas, além de mais sensíveis que o sequenciamento clássico (BOLTON et al., 2015).

Uma tecnologia que apresenta vantagem em relação às demais para detecção de mutações somáticas é a PCR de terceira geração - a PCR digital. Com essa técnica, é possível detectar e quantificar as mutações na reação, utilizando a combinação de diluição da amostra, PCR alelo-específica e estatística de Poisson (SYKES et al., 1992). Dessa forma, a PCR digital é capaz de detectar frequências na ordem de 0,01% de cópias do gene com a variante de interesse (PEKIN et al., 2010), permitindo a identificação de mutações com maior sensibilidade e acurácia (HUGGETT; WHALE, 2013).

Visto isso, o objetivo geral desse projeto é desenvolver um método de tipagem por PCR digital para a detecção de mutações somáticas nos genes *EGFR* e *KRAS*, em DNA extraído de biópsias teciduais e avaliar o seu desempenho, de forma a validá-lo para a aplicação na rotina clínica. Para tanto, comparar-se-á o seu resultado com o obtido por *castPCR* (PCR competitiva em tempo real por TaqMan®) e por sequenciamento de Sanger. Com a implementação dessa ferramenta para uso clínico, espera-se que o presente trabalho contribua grandemente na precisão do diagnóstico e tomada de decisão terapêutica para o tratamento dos pacientes.

2. REVISÃO DE LITERATURA

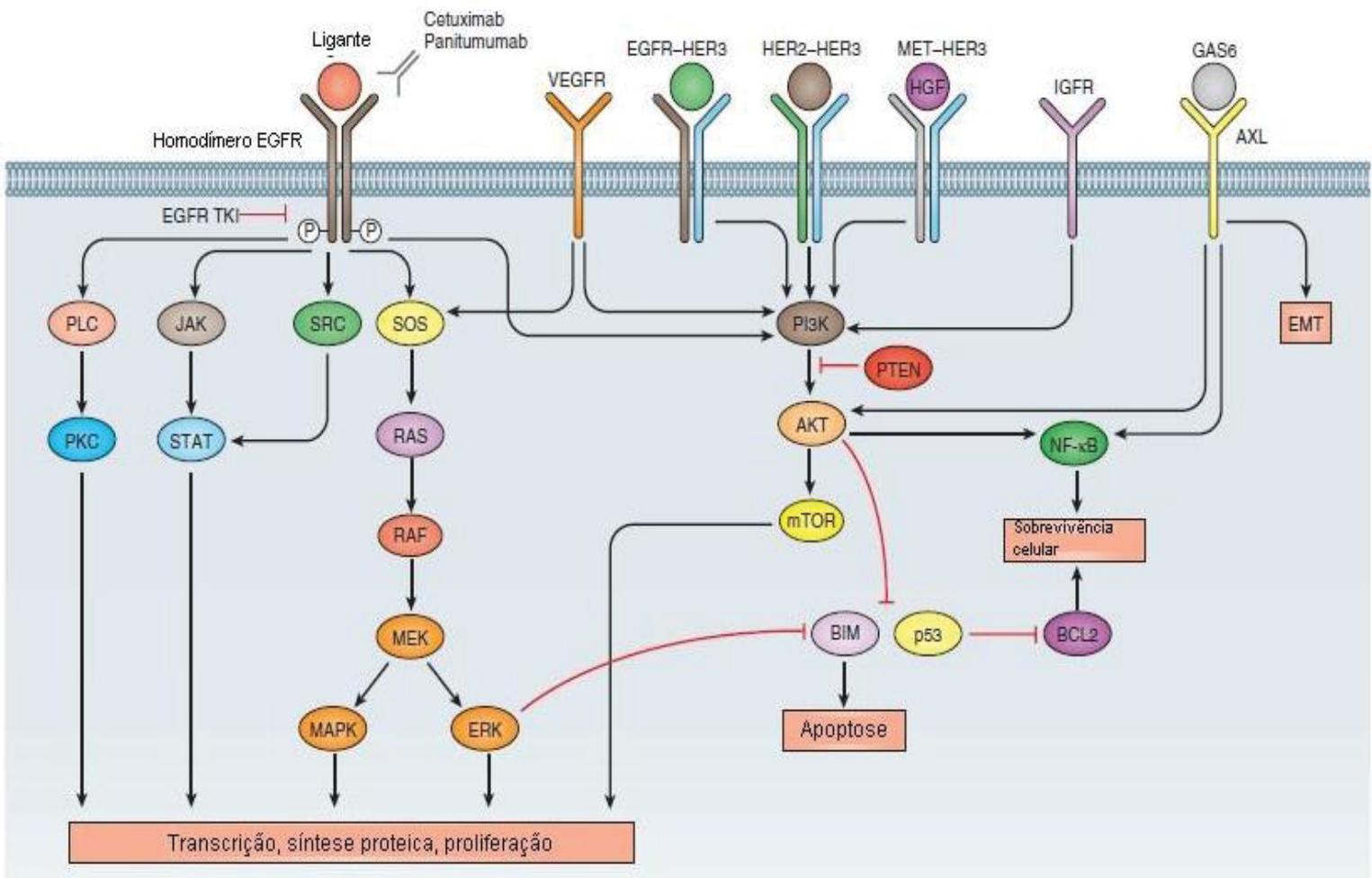
2.1. ONCOGÊNESE E AS VIAS DE SINALIZAÇÃO EGFR E KRAS

A carcinogênese é um processo de múltiplas etapas, tanto em nível fenotípico, quanto genético (LOEB; LOEB; ANDERSON, 2003). O dano genético não letal pode desregular a atividade normal de genes envolvidos na promoção do crescimento, supressão tumoral, apoptose e reparo, contribuindo para a origem ou progressão de tumores e levando à transformação de células normais, em células tumorais (KUMAR, 2011). Hanahan e Weinberg, em 2011, propuseram um modelo de características comuns das neoplasias, no qual a maioria das células cancerosas adquire essas propriedades durante o seu desenvolvimento, tipicamente por meio de mutações nos genes relevantes. São estas: (i) autossuficiência aos sinais de crescimento e (ii) insensibilidade aos sinais de anti crescimento, (iii) invasão tecidual e metástase, (iv) potencial replicativo ilimitado, (v) angiogênese sustentada e (vi) evasão à apoptose (HANAHAN; WEINBERG, 2011).

Todas as células normais requerem estimulação por fatores de crescimento, para se submeterem à proliferação (KUMAR, 2011). O receptor de fator de crescimento epidérmico (EGFR) é uma proteína de superfície celular com atividade de tirosina-quinase, e a ligação do ligante EGF ao seu receptor induz dimerização e autofosforilação do receptor, levando a uma cascata de proliferação celular (FIGURA 1). Visto que o câncer é também uma doença de proliferação celular inapropriada, o gene *EGFR* possui ação no desenvolvimento e progressão de neoplasias, como no câncer de pulmão de células não pequenas e câncer colorretal (ONN et al., 2004). Através de vias de sinalização que influenciam a angiogênese, ativação, regulação e proliferação celular, sua ativação impede a apoptose celular (LICHTENBERGER et al., 2010). Logo, há um consenso clínico que as terapias anti-EGFR se mostram

eficazes, na maioria dos casos, para a redução da progressão em tumores que apresentam alguma alteração genética em *EGFR*. Da mesma forma, o uso de anticorpos inibidores de EGFR como cetuximabe e panitumumabe também se mostra eficaz no tratamento desses pacientes, por inibir a cascata de ativação de RAS/MAPK, inibindo a proliferação celular e induzindo apoptose (ROOCK et al., 2011) (FIGURA 1).

Figura 1. Vias de sinalização de EGFR.



Legenda: EGFR: receptor de fator de crescimento epidérmico; EGFR-TKI: inibidores de tirosina-quinase; VEGFR: receptor de fator de crescimento endotelial vascular; HER: receptor de fator de crescimento epidérmico humano; PLC: fosfolipase C; IGFR: fator de crescimento tipo insulina; GAS6: proteína *growth arrest-specific 6*; JAK: proteína Janus cinase; STAT: transdutor de sinal e ativador de transcrição; SRC: (cinase SRC, abreviação de sarcoma); SOS: proteína derivada do gene filho de *seven less*; RAS: oncogene homólogo de sarcoma viral de rato; RAF: cinase fibrosarcoma rapidamente acelerado; MEK: cinase ativadora de MAP cinase; MAPK: cinase ativada por mitógeno; ERK: cinase regulada por sinal extracelular; PI3K: fosfatidilenositol 3 cinase; PTEN: homólogo de fosfatase tensina; mTOR: proteína alvo da rapamicina em mamíferos; NF-κB: fator nuclear kappa beta; BIM: proteína de

membrana inibida por BCL2; p53: proteína 53; BCL2: célula B de linfoma 2; EMT: proteína epitélio-mesênquima.

Nota: os receptores de tirosina-quinase podem estar na forma de homodímeros EGFR, ou heterodímeros EGFR-HER3, HER2-HER3 ou MET-HER3. A ativação de EGFR por seu ligante leva a uma cascata de ativação que resulta em proliferação tumoral e bloqueio de apoptose. Inibidores de tirosina-quinase são utilizados por serem capazes de impedir essa cascata.

Fonte: (CHONG; JÄNNE, 2013)

A sinalização nas cascatas posteriores aos receptores de fatores de crescimento é mediada por KRAS, uma proteína efetora da superfamília das GTPases de extrema importância na patogênese das neoplasias. O gene *KRAS* pertence à família de genes *RAS* (*HRAS*, *NRAS* e *KRAS*) e codifica essa GTPase emissora de sinal (KUMAR, 2011). Mutações em *KRAS* resultam em uma proteína incapaz de exercer sua função GTPásica normal, e o estado ativado é mantido, resultando em proliferação celular aumentada (ROOCK et al., 2011). Aproximadamente 20% de todas as neoplasias humanas contêm versões mutadas das proteínas RAS e cerca de um terço dos adenocarcinomas de cólon e reto, possuem mutações em *KRAS* (SAMOWITZ et al., 2000).

2.1.1. Câncer de pulmão e EGFR.

Sem considerar os tumores de pele não-melanoma, o câncer de pulmão é o mais comum de todas as neoplasias malignas e a principal causa de morte por neoplasia no mundo, sendo responsável por aproximadamente uma em cada cinco mortes por neoplasia (1,6 milhões em 2012) (FERLAY et al., 2015). Estima-se que, em 2018, ocorram mais 18.740 novos casos em homens e 12.530 novos casos em mulheres no Brasil (Estimate/2018 – Cancer Incidence in Brazil). O principal fator de risco para o desenvolvimento do câncer de pulmão, bem como tumores em traqueia e brônquios, é o tabagismo (Estimate/2018 – Cancer Incidence in Brazil). No final do século XX, o câncer de pulmão foi considerado uma das principais neoplasias de causa de morte evitável, visto que fumantes têm cerca de 10 vezes mais chances de desenvolver neoplasias associadas às vias aéreas superiores, quando comparados aos indivíduos que nunca fumaram (KUMAR, 2011).

De acordo com a classificação histológica, mais de 80% de todos os casos de câncer de pulmão são do tipo não pequenas células (NSCLC, do inglês *non small cell lung cancer*) (FANG; WANG, 2014). O adenocarcinoma de pulmão é o principal tipo

de NSCLC, e o gene com a maior frequência de mutações associadas a esta doença, é o *EGFR* (ARRIETA et al., 2015). Mais de 100 diferentes mutações já foram identificadas no gene *EGFR* associadas a adenocarcinoma de pulmão, mas cerca de dez se repetem com mais frequência (YEH et al., 2013). Com exceção de alguns casos raros (OXNARD et al., 2012), grande parte dessas mutações são somáticas.

As mutações somáticas mais comuns identificadas no gene *EGFR* associadas a adenocarcinoma de pulmão são *rs121913424* (DEL19 ou ENSP00000275493.2:p.Glu746_Ser752delinsVal) e *rs121434568* (p.L858R) (ARRIETA et al., 2015), e alelos selvagens para essas mutações são mais frequentemente encontrados em mulheres jovens com câncer de pulmão não fumantes (PAO et al., 2004). Através da desestabilização da sua conformação, DEL19 e p.L858R são capazes de induzir ativação constitutiva do receptor EGFR, e por essa razão são ditas mutações ativadoras (ROSELL et al., 2009). Alguns pesquisadores sugerem que a presença dessas mutações somáticas é um marcador de prognóstico favorável (FANG; WANG, 2014), pois conferem sensibilidade ao tratamento com inibidores reversíveis de tirosina quinase (TKIs, do inglês *tyrosine kinase inhibitors*) como gefitinibe, erlotinibe e afatinibe em NSCLC avançado (BAEK et al., 2015) (TABELA 1). Por meio de estudos de análise multivariada, observou-se que a presença dessas mutações no gene *EGFR* influencia positivamente o tempo de sobrevida geral (D'ANGELO et al., 2012) e que estão associadas ao menor risco de recorrência, e maior probabilidade de alcançar linfonodos negativos para células tumorais e ausência de metástases a distância, dependente de tratamento (KOSAKA et al., 2009).

A maioria dos autores utiliza a classificação para as mutações no gene *EGFR* nos pacientes com NSCLC de acordo com sua frequência e resposta a primeira linha de tratamento. Dessa forma, as mutações DEL19 e p.L858R são ditas clássicas, visto suas altas frequências e maior resposta ao tratamento, enquanto as outras mutações são não clássicas, ou raras (BAEK et al., 2015). Indivíduos que possuem mais de uma mutação associada, são classificados como portadores de mutações complexas (HATA et al., 2010) (TABELA 1).

Tabela 1: Frequência de mutações clássicas, não clássicas e complexas no gene *EGFR*, associadas ao adenocarcinoma de pulmão e à resistência ao tratamento com inibidores de EGFR.

MUTAÇÃO	CÂNCER DE PULMÃO (%)	RESPOSTA AO TRATAMENTO (%)	EFEITO GERAL
<i>Clássicas</i>			
DEL19	45	72,9	Sensibilidade
p.L858R	40	60,1	Sensibilidade
<i>Não clássicas</i>			
20indel	2 – 9	*	Resistência primária
p.G719X	3	50	Resposta favorável
p.L861Q	2	60	Resposta favorável
p.T790M	0,5 - 3	0	Resistência adquirida
<i>Complexas</i>	5 – 9		**

* a resposta para 20indel (*rs146024686*) varia progressivamente de acordo com o tamanho do fragmento inserido

** Mutações cujo efeito não é totalmente elucidado.

Nota: Os valores de DEL19 (*rs121913424*) e p.G719X (*rs28929495*) estão representados em porcentagem da média de todas as variações conhecidas associadas. Para as mutações complexas, são consideradas as combinações já vistas em estudos anteriores. A porcentagem média de resposta ao tratamento é com base em estudos com inibidores reversíveis de EGFR (Gefitinib de primeira, e Erlotinib de segunda geração).

Fontes: (ARRIETA et al., 2015; CHONG; JÄNNE, 2013)

Os pacientes com NSCLC que inicialmente respondem aos EGFR-TKIs devem ser monitorados para a possível ocorrência e seleção de mutações somáticas de resistência (THRESS et al., 2015). A mais comum é a *rs121434569*, uma substituição de uma treonina por metionina na posição 790 (p.T790M) no éxon 20, que ocorre em aproximadamente 50% dos casos de resistência adquirida (TABELA 1), estando presente também, em 6% das mutações complexas (HATA et al., 2010). Tal mutação parece resultar em redução da afinidade da ligação entre os TKIs e o domínio cinase de EGFR, o que poderia justificar seu mecanismo de resistência (MA et al., 2017). Para estes pacientes com resistência adquirida devido a mutação p.T790M, inibidores de EGFR de terceira geração como Osimertinibe e Rociletinibe parecem ser efetivos (HIDAKA et al., 2017). No entanto, Hidaka e colaboradores observaram a presença de uma substituição de cisteína por serina na posição 797 (p.C797S) em pacientes que vivenciaram uma falha nas terapias de primeira e terceira geração com inibidores de EGFR. Esta mutação pode estar em *cis* ou em *trans* com p.T790M (HIDAKA et al., 2017).

O desenvolvimento de inibidores EGFR irreversíveis como Neratinibe e Canertinibe foi descontinuado pela falta de eficácia e uma dose limitante com efeitos colaterais (KWAK, 2011). Uma outra linha de tratamento que se mostrou efetiva para pacientes com NSCLC é a quimioterapia a base de platina. No entanto, apenas indivíduos com mutações raras em *EGFR* possuem a resposta adequada, quando recebem esta quimioterapia como primeira escolha de tratamento (THRESS et al., 2015).

2.1.2. Câncer colorretal e KRAS

O câncer colorretal (CCR) é o terceiro tipo de câncer mais frequente no mundo e assim como o câncer de pulmão, é considerado uma “doença de estilo de vida”, visto que sua incidência é influenciada por fatores comportamentais (BAE; KIM; KANG, 2016), como hábitos alimentares ricos em gorduras, carnes vermelhas e processadas, pouca ingestão de fibras e frutas, alta prevalência de obesidade e sobrepeso e pouca prática de exercícios físicos (Estimate/2018 – Cancer Incidence in Brazil). Estima-se que até o final do ano de 2018, o Brasil tenha 17.380 novos casos de câncer colorretal em homens e 18.980 em mulheres, o que torna essa uma das neoplasias mais frequentes na região sul (22,17 / 100 mil habitantes). Em 2012, foram estimados mais de 600 mil casos de morte por câncer colorretal no mundo, o que reflete 8,5% do total de óbitos, sendo a maioria em países com IDH baixo (Estimate/2018 – Cancer Incidence in Brazil).

Mais de dois terços dos casos de CCR são do tipo esporádico, enquanto que 20 – 30% possuem um componente hereditário incomum, como por exemplo a síndrome de Lynch (3% dos casos) e a polipose adenomatosa familiar (1% dos casos) (WHIFFIN et al., 2014). A instabilidade genética (predisposição a mutações no número de repetições de microssatélites ou número/alterações cromossomos) aumenta à medida que as células normais se transformam em células neoplásicas. O epitélio normal possui duas vias possíveis de mudança de padrão para o desenvolvimento em CCR esporádico (KUIPERS et al., 2015): (i) a via clássica, que resulta na progressão de adenomas tubulares para adenocarcinoma (FEARON; VOGELSTEIN, 1990) e (ii) a via alternativa, que compreende pólipos e sua progressão para CCR serrátil. De acordo com as características moleculares, o CCR esporádico pode ser classificado em três principais subgrupos: (i) aqueles cuja característica principal é a instabilidade cromossômica (INC), com alterações no número e estrutura dos cromossomos, (ii)

aqueles que apresentam instabilidade de microssatélites (IMS) (MÜLLER et al., 2016) e (iii) aqueles com instabilidade epigenética, ou taxa elevada de metilação das ilhas CpG (regiões com elevada frequência de citosina-guanina, no sentido 5'-3', frequentemente presentes na região reguladora de genes constitutivos e tecido-específicos) (BAE; KIM; KANG, 2016). Tumores com instabilidade cromossômica são mais comuns e geralmente resultam da combinação da ativação de um oncogene (como *KRAS*) e a inativação de um gene supressor tumoral (MÜLLER et al., 2016).

Cerca de um terço dos adenocarcinomas de cólon e reto possuem mutações em *KRAS*, a maioria destas mutações ocorrem nos códons 12 (*rs121913530*) ou 13 (*rs121913529*) e apenas 5% no códon 61 (*rs121913238*) (SAMOWITZ et al., 2000) (TABELA 2). Estudos retrospectivos (PEETERS et al., 2010) e prospectivos (AMADO et al., 2008) confirmam a falta de resposta de pacientes que apresentam tumores com mutação em *KRAS*, submetidos a monoterapia ou terapia combinada. Ajustes no tratamento, com base em mutações, já demonstrou aumentar o tempo de sobrevida global dos pacientes em 40 meses, após o diagnóstico de câncer de cólon e reto (GRAHAM et al., 2016). Tanto o FDA (*Food and Drug Administration*), quanto o EMA (*European Medicines Agency*), recomendam que, antes do tratamento com inibidores de EGFR, os pacientes com câncer colorretal sejam previamente testados para as mutações mais comuns em *KRAS*, visto que a presença de mutações permite a transmissão do sinal de crescimento e proliferação tumoral, independente da ativação de EGFR (KRISTENSEN et al., 2012). Em 2010, De Roock e colaboradores encontraram que a mutação *rs112445441* (p.G13D) pode ser benéfica para pacientes com carcinoma colorretal em tratamento com cetuximabe, estando associada a maior tempo livre de doença e maior sobrevida geral (DE ROOCK et al., 2010). Os autores sugerem que tumores que apresentam essa substituição são mais sensíveis ao cetuximabe, o que beneficia o uso desse tratamento em indivíduos com essa mutação. No entanto, mesmo o tratamento com cetuximabe ou panitumumabe sendo indicado apenas para os pacientes com genótipo do tipo selvagem ou p.G13D para *KRAS*, a taxa de resposta varia entre 10 a 40%, haja visto a frequente presença de desregulação de outros oncogenes da via, como *BRAF* (do inglês, *V-Raf murine sarcoma viral oncogene B1*), *PTEN* (do inglês, *phosphatase and tensin homolog*) ou *PIK3CA* (fosfatidilinositol 3 cinase subunidade catalítica alfa), por exemplo (ROOCK et al., 2011).

Tabela 2. Frequência de mutações somáticas no gene *KRAS* associadas ao adenocarcinoma colorretal e de resistência ao tratamento.

MUTAÇÃO	CÂNCER COLORRETAL (%)	RESPOSTA AO TRATAMENTO (%)	
		MUTADO	SELVAGEM
Códon 12	80,2	0 - 6	10 - 40
Códon 13	14,8	6,5	26
Códon 61	3 - 5	*	*

* efeito não é totalmente elucidado.

Fontes: (DE ROOCK et al., 2010; SAMOWITZ et al., 2000).

2.2. TÉCNICAS PARA IDENTIFICAÇÃO DE MUTAÇÕES SOMÁTICAS NA ROTINA DIAGNÓSTICA

A identificação de mutações somáticas associadas ao câncer promoveu uma mudança nas estratégias terapêuticas, partindo da quimioterapia generalizada para formas de tratamento personalizadas, cuja eficácia clínica é muito maior, em comparação com a terapia padrão (MAEMONDO et al., 2010). Diversas ferramentas clínicas são atualmente utilizadas para detecção de mutações somáticas (TABELA 3), e a biópsia de tecido é considerada o substrato padrão-ouro para identificar o perfil genético tumoral. Contudo, biópsias de câncer submetidas à fixação na formalina e parafinização, muitas vezes fornecem material de baixa quantidade e qualidade (YU; O'TOOLE; TRENT, 2015). A amplificação deste material é possível pela reação em cadeia da polimerase (PCR), idealizada por Kary Mullis na década de 80 (MULLIS et al., 1986). Milhões de cópias de fragmentos podem ser geradas através da amplificação exponencial de uma única molécula de DNA, com apenas alguns ciclos de alternância térmica. Um ciclo típico consiste em três etapas distintas: (1) desnaturação, (2) acoplamento dos iniciadores e (3) extensão da cadeia (AHRBERG; MANZ; CHUNG, 2016). Durante a fase de desnaturação, altas temperaturas (cerca de 95°C) são responsáveis por separar a dupla fita de DNA em fitas simples. Em seguida, a temperatura é reduzida para o acoplamento de cerca de 50% dos iniciadores à fita simples de DNA, a uma temperatura que depende do seu conteúdo de G-C (em geral, entre 50-68°C). Os iniciadores são oligonucleotídeos de 17-26 nucleotídeos, complementares às sequências que flanqueiam a região a ser amplificada. Ao se hibridizarem à sequência simples de DNA em uma região alvo, permitem que a enzima Taq DNA polimerase, originalmente extraída da *Archaea Thermus aquaticus* e resistente a altas temperaturas, inicie a polimerização da fita de

DNA. A atividade da polimerase é ótima a 72°C, na etapa de extensão da cadeia. Nesta fase, o segmento de DNA complementar é sintetizado a partir dos nucleotídeos livres na solução, resultando em um aumento exponencial de fragmentos dupla fita de DNA, obtidos pela repetição das três etapas. O sucesso da reação é observável após corrida eletroforética e incorporação de fluoróforos entre as bases nitrogenadas do produto amplificado, que permite a visualização da banda após excitação sob comprimento de onda apropriado (SAIKI et al., 1988).

Tabela 3. Técnicas comumente utilizadas na identificação de mutações somáticas na rotina diagnóstica.

TÉCNICA	BASE DA TÉCNICA	VANTAGENS	DESVANTAGENS
Sequenciamento de Sanger	PCR + dideoxinucleotídeos fluorescentes + eletroforese capilar	Padrão-ouro; analisa base por base; diversas mutações analisadas.	Pouca sensibilidade; impossibilidade de ensaios multiplex.
Pirosequenciamento	A base incorporada durante a PCR libera um pirofosfato que pode ser detectado como luz visível através de um sistema enzimático	Boa sensibilidade; alta resolução; diversas mutações analisadas; bom para uso com amostras FFPE.	Poucas leituras; impossibilidade de ensaios multiplex; alto custo.
Espectrometria de massa (MALDI-TOF)	PCR + alongamentos dos oligonucleotídeos com terminadores	Boa sensibilidade; alta resolução; bom para uso com amostras FFPE; possibilidade de ensaios multiplex.	Necessidade de desenho experimental acurado; apenas análise de mutações previamente selecionadas.
PCR em tempo real alelo-específica	PCR com oligonucleotídeos e sondas para a região de interesse + análise de fluorescência	Alta sensibilidade; sem necessidade de testes pós PCR.	Necessidade de desenho experimental acurado e otimização cuidadosa; dificuldade em diferenciar nucleotídeos em baixa frequência.
<i>High resolution melting</i>	Produtos de PCR possuem diferentes temperaturas de <i>melting</i>	Boa sensibilidade; sem necessidade de testes pós PCR.	Necessidade de desenho experimental acurado e otimização cuidadosa.
Sequenciamento de nova geração (NGS)	Preparo de biblioteca + PCR em emulsão + sequenciamento ligado a adaptadores + fluorescência de incorporação	Altíssima sensibilidade quando em alta cobertura; analisa base por base; diversas mutações analisadas.	Alto custo; dificuldade na interpretação dos dados.
PCR digital (dPCR)	Diluição amostral + PCR + leitura	Altíssima sensibilidade; bom para uso com amostras FFPE; possibilidade de ensaios multiplex;	Necessidade de desenho experimental acurado e otimização cuidadosa; apenas análise de mutações previamente selecionadas.

Legenda: PCR - reação em cadeia da polimerase; MALDI-TOF - *matrix-assisted laser desorption/ionization—time of flight*; NGS - *next generating sequencing*;

FFPE: amostras fixadas em formalina e emblocadas em parafina; multiplex: combinação de diferentes genes detectados no mesmo ensaio.

Nota: em geral, a PCR digital não é comercialmente utilizada na identificação de mutações somáticas. No entanto, a dPCR foi adicionada à tabela a fim de comparação, visto que a proposta da inserção desta técnica na rotina diagnóstica faz parte do objetivo principal deste trabalho.

Fonte: adaptado de YU, 2015.

2.2.1. Identificação de mutações somáticas por sequenciamento

Após a amplificação do material genético, o sequenciamento genético de Sanger permite a identificação de variantes, em um processo similar à PCR, envolvendo amplificação, purificação e eletroforese com radioisótopos ou fluoróforos marcando as bases nitrogenadas (SIMMONS, SNUSTAD, 2010). O sequenciamento baseia-se na incorporação de desoxinucleotídeos (dNTPs) e de didesoxinucleotídeos (ddNTPs, quimicamente modificados de forma a não apresentarem a hidroxila 3'OH) a uma cadeia de DNA em crescimento. A polimerização é interrompida no momento em que a DNA polimerase, por acaso, insere um ddNTP, visto que a ausência do grupo hidroxila 3' OH livre impede a ligação do próximo dNTP (SIMMONS, SNUSTAD, 2010). Nas formas mais antigas de sequenciamento, os dNTPs eram radioativamente marcados (SANGER; NICKLEN; COULSON, 1977). Desde os anos 90, a visualização dos fragmentos é possível pelo uso de ddNTPs marcados com fluoróforos diferentes (um por base). Esta estratégia de amplificação gera fragmentos marcados de tamanhos distintos. Após purificação, os fragmentos marcados por fluoróforos são ordenados por tamanho e cor, utilizando um software que reflete a migração eletroforética e marcação dos fragmentos, na forma de eletroferogramas (FIGURA 2).

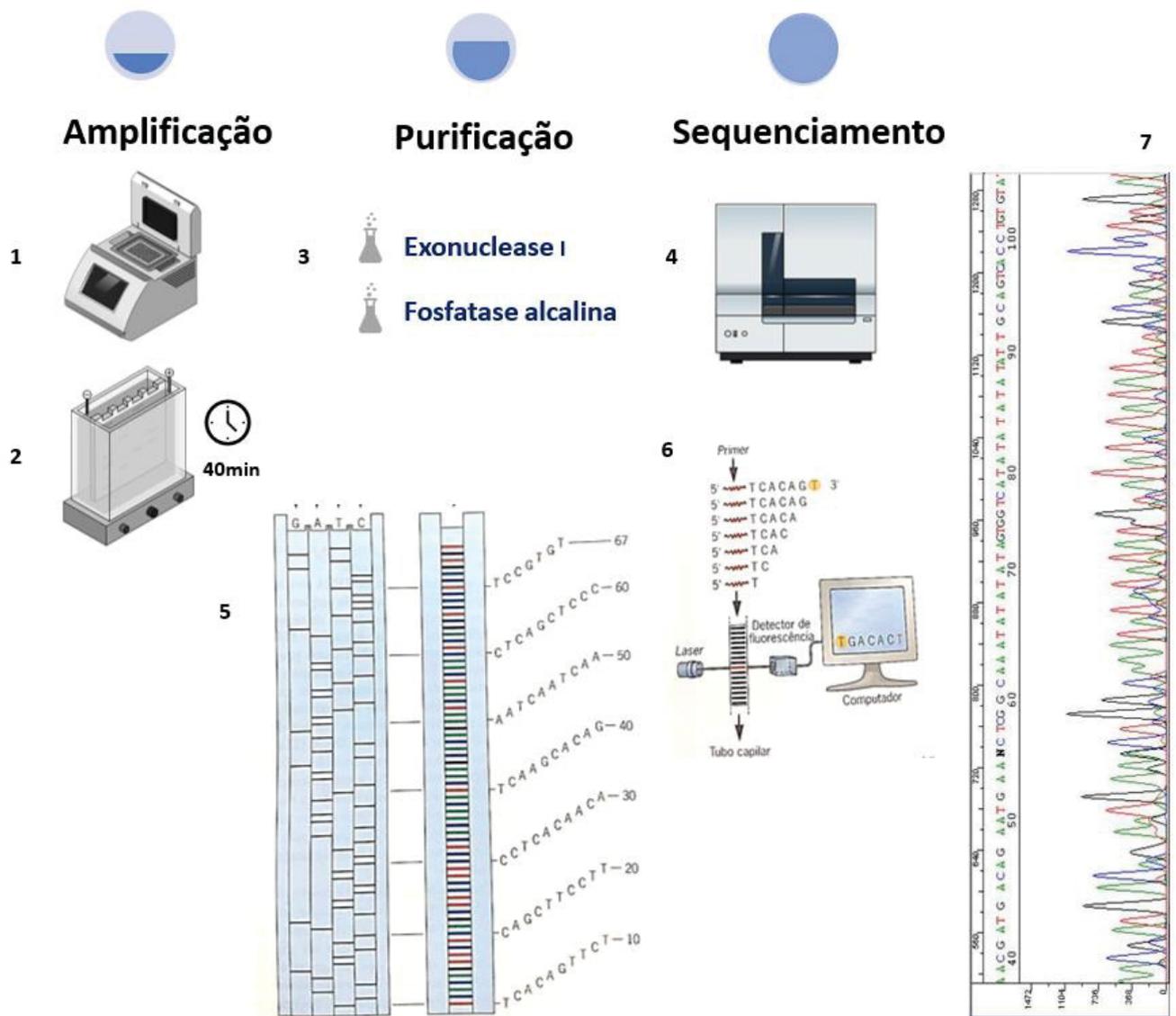


Figura 2. Sequenciamento de DNA pelo método de Sanger.

Legenda: O protocolo consiste em etapas de amplificação em termociclador (1) e confirmação dos fragmentos em gel de agarose (2), purificação do produto de PCR com as enzimas exonuclease I e fosfatase alcalina (3) e sequenciamento seguido de precipitação (4). Os produtos das reações de sequenciamento são separados com base no tamanho por eletroforese de capilaridade (5), e a fluorescência é detectada à medida que cadeias nascentes de DNA se movem através de um feixe de laser (6). A análise dos eletroferogramas (7) é realizada utilizando um software de análise.

Fonte: adaptado de SNUSTAD, 2010.

O sequenciamento genético de Sanger é considerado o método padrão-ouro para análise de sequências de DNA, permitindo a detecção de mutações. A grande vantagem dessa técnica é a possibilidade de detecção de diferentes tipos de mutação – inserções, deleções e substituições de base (VAN KRIEKEN et al., 2016). No entanto, é necessário que pelo menos 10 - 25% das células na biópsia contenham a

mutação, para que estas sejam detectadas com precisão (CAROTENUTO et al., 2010). Além disso, esta técnica apresenta um custo alto para doenças com heterogeneidade genética e alélica como o câncer, nas quais é necessário sequenciar vários genes e diferentes sequências, dentro dos mesmos (PANDEY et al., 2016).

Em contraste, o sequenciamento multiparalelo ou de nova geração (NGS, do inglês *next generation sequencing*) fornece uma quantidade muito maior de resultados que o método de Sanger (tecnologia de primeira geração) (BEHJATI; TARPEY, 2013). O NGS tem sido utilizado como um método complementar ao sequenciamento de Sanger, garantindo um teste genético eficiente e acessível para identificação e confirmação de mutações causadoras de doenças na pesquisa clínica (PANDEY et al., 2016). A maior parte das plataformas de NGS comercialmente disponíveis realizam o sequenciamento de bilhões de pequenos fragmentos de DNA em paralelo (BEHJATI; TARPEY, 2013; VOELKERDING; DAMES; DURTSCHI, 2009). O sequenciamento inteiro de genoma (WGS, do inglês *whole-genome sequencing*) fornece o mais completo conjunto de informações genômicas, o que requer organização e recursos para a interpretação do grande volume de dados gerado (DEL VECCHIO et al., 2017). A fim de evitar essas dificuldades e, ao mesmo tempo, obter alto rendimento de dados, foram elaborados métodos como sequenciamento completo de exoma (WES, do inglês *whole-exome sequencing*), sequenciamento direto ou de transcriptoma (RNA-Seq) (DEL VECCHIO et al., 2017).

A Roche™ foi pioneira na tecnologia de NGS com o *454 Genome Sequencer* (www.roche.com) (ROTHBERG; LEAMON, 2008). A plataforma utiliza uma PCR em emulsão, na qual as moléculas de DNA fragmentadas são amplificadas individualmente. Os nucleotídeos incorporados durante a amplificação são identificados pela detecção de diferenças na energia emitida, quando há a liberação dos pirofosfatos (PPi). Esta técnica é denominada de pirosequenciamento e independe de fluoróforos (DEL VECCHIO et al., 2017). Já a plataforma da *Illumina® (Illumina Genome Analyzer sequencing)* é a mais amplamente conhecida e utiliza nucleotídeos marcados com fluoróforos (METZKER, 2009). O fragmento fita simples de DNA é previamente ligado a um “adaptador”, que apresenta sequência complementar a um oligonucleotídeo fixado numa célula de fluxo, em uma lâmina. Quando adicionado a esta célula, tal fragmento é ancorado por hibridação a este oligonucleotídeo. A amplificação inicia quando o fragmento forma uma “ponte”, hibridando com um iniciador também preso à lâmina. Diversos ciclos de amplificação

geram agrupamentos de clones amplificados. Após a amplificação, estes clones são desnaturados e, portanto, linearizados. A reação de sequenciamento é iniciada com adição dos iniciadores, enzima polimerase e nucleotídeos com terminadores reversíveis, marcados com fluoróforos. Se houver complementaridade com a sequência do clone, os nucleotídeos terminadores acoplam-se a ele, gerando um sinal fluorescente e bloqueando temporariamente a continuação da síntese. O excesso de fluoróforo é removido antes do próximo ciclo de síntese, que se inicia após a remoção do terminador e do fluoróforo (VOELKERDING; DAMES; DURTSCHI, 2009).

Além dessas, a plataforma SOLiD® é disponibilizada pela *Thermo Fisher Scientific*. A preparação da amostra compartilha similaridades com a tecnologia do *454 Genome Sequencer* (Roche), na qual o DNA fragmentado é ligado a oligonucleotídeos adaptadores, incorporado em esferas e amplificado por uma reação de PCR em emulsão. Cada esfera com um clone amplificado é imobilizada em uma superfície de vidro, e sequenciada a partir do acoplamento dos oligonucleotídeos iniciadores, complementares ao DNA-adaptador. Diferentes sondas competem pela hibridação às sequências adjacentes aos oligonucleotídeos iniciadores, e a fluorescência de pós-incorporação é registrada (DEL VECCHIO et al., 2017; VOELKERDING; DAMES; DURTSCHI, 2009). A plataforma *Ion torrent™* (*Thermo Fisher Scientific*) também é amplamente utilizada, e se baseia na detecção da alteração de pH, provocada pela liberação de íons hidrogênio (H⁺) durante a polimerização do DNA, o que permite transformar este sinal químico em sinal digital (MERRIMAN; ROTHBERG, 2012).

Inúmeros estudos são realizados com a técnica de sequenciamento individual do genoma, incluindo o Projeto 1000 Genomas (<http://www.internationalgenome.org/>) e o Atlas do Genoma do Câncer (<http://cancergenome.nih.gov/>), que utilizam a plataforma da *Illumina®* para sequenciar todo o genoma (DEMICCO et al., 2017; RADOVICH et al., 2018; SUDMANT et al., 2015). Independentemente da plataforma, o NGS tem se mostrado um método sensível e específico para determinação de mutações presentes em baixas frequências (DEL VECCHIO et al., 2017). Dependendo da cobertura (número de fragmentos sequenciados que correspondem à sequência de interesse), o mesmo permite a detecção de mutações somáticas com prevalências inferiores a 1. Também possibilita a identificação de outras variações nos genes de interesse e subsequente avaliação de seu significado, através de ferramentas de bioinformática (FRANCIS; STEIN, 2015). No entanto, para identificar

as possíveis mutações com grande sensibilidade e especificidade, é essencial garantir dados de sequenciamento alta qualidade (PANDEY et al., 2016).

2.2.2. Identificação de mutações somáticas por PCR em tempo real (qPCR)

Métodos baseados em PCR quantitativa em tempo real (qPCR) parecem ser promissores para detecção de mutações somáticas e mais sensíveis que o sequenciamento, mas sua aplicação na rotina ainda é incerta (BOLTON et al., 2015). A maioria dos ensaios de qPCR possui os ciclos típicos de uma reação de PCR que consistem em desnaturação ($\approx 95^\circ\text{C}$), acoplamento e extensão ($\approx 60^\circ\text{C}$), associados à adição de fluoróforos livres ou ligados a sondas. A validação apropriada dos oligonucleotídeos iniciadores é crucial para uma reação de PCR específica, sensível e robusta (BUSTIN; HUGGETT, 2017). A qPCR permite o monitoramento contínuo do sinal fluorescente ao longo dos ciclos de amplificação, e a análise do tempo necessário para que este alcance o ponto de saturação do sistema. Este período é tanto menor, quanto maior for a quantidade de substrato inicial para amplificação, razão pela qual a qPCR permite quantificação relativa a um controle ou padrão de diluição, de uma sequência a ser pesquisada. Ela pode ser associada a sondas de hidrólise marcadas com fluoróforos e complementares à sequência que contém ou não, a mutação pesquisada. No sistema Taqman®, a sonda está marcada com um fluoróforo repórter na extremidade 5', que, ao ser estimulado por uma fonte de luz, transfere a energia para um extintor (*quencher*) na extremidade 3', o qual suprime a emissão de fluorescência pelo repórter (APPLIED BIOSYSTEMS, 2005) (FIGURA 3). Caso haja complementaridade entre a sequência a ser pesquisada e a sonda, ocorre emissão de fluorescência na etapa de extensão da cadeia, no momento em que a sonda é destruída pela atividade exonucleásica 5'-3' da polimerase, separando o repórter do extintor. A intensidade de fluorescência é proporcional ao número de amplicons com a sequência específica, possibilitando a conversão do sinal fluorescente em valor numérico para cada amostra (SIMMONS, SNUSTAD, 2010)(ROMA et al., 2013).

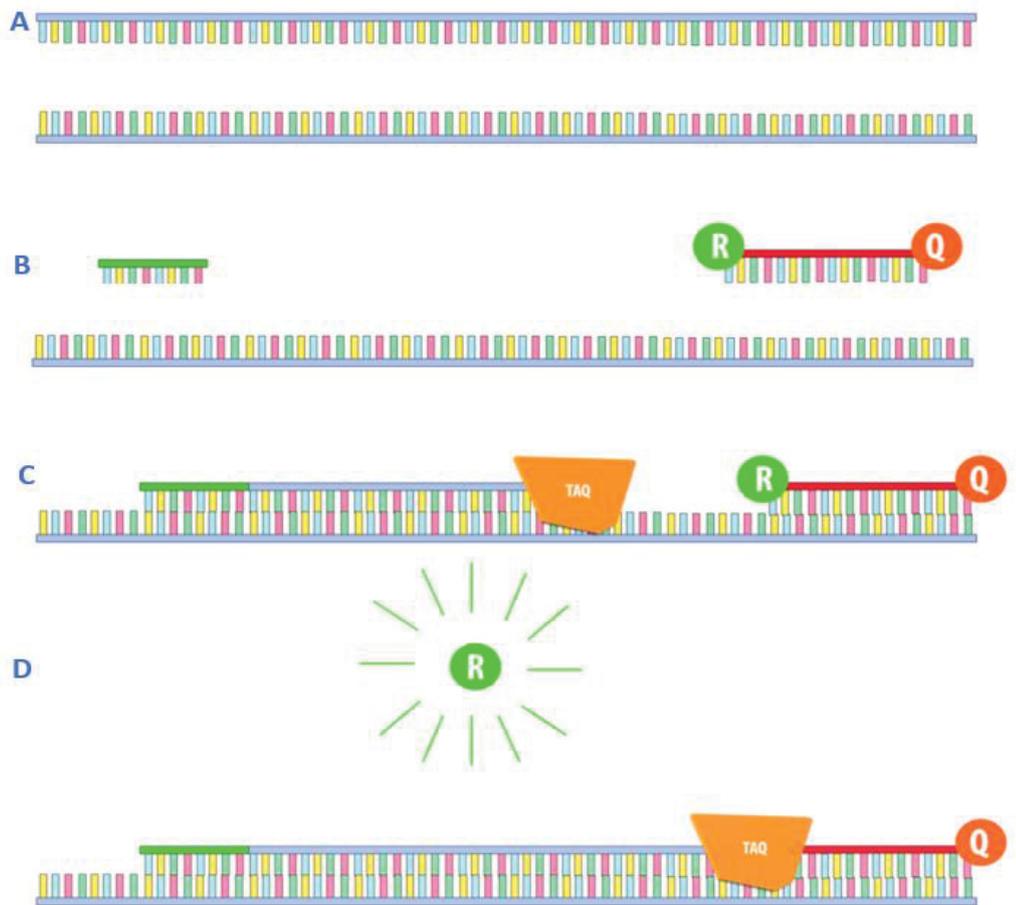


Figura 3. Sistema de detecção TaqMan®.

Legenda: Ciclos de uma reação de PCR em tempo real que consistem em (A) desnaturação, (B) acoplamento dos oligonucleotídeos iniciadores e da sonda fluorescente (C) e extensão pela Taq polimerase (em laranja) (D). A sonda TaqMan® é específica para a sequência de interesse e marcada com um fluoróforo repórter e um extintor. Em caso de complementaridade entre a sequência pesquisada e a sonda (C), ocorre emissão de fluorescência na etapa de extensão da cadeia (D).

Fonte: www.youtube.com/MrSimpleScience. Acesso em 16/03/2018

A PCR competitiva alelo-específica por TaqMan® (*castPCR*) é uma qPCR modificada, com sensibilidade confirmada para detectar mutações raras em DNA oriundo de amostras parafinadas (DIDELOT et al., 2012). Nesta técnica, um dos iniciadores de um par é alelo-específico, sendo o outro, complementar à sequência flanqueadora do fragmento de interesse. Iniciadores alelo-específicos também são utilizados em PCRs sequência-específicas e apresentam o último nucleotídeo a 3', complementar à variante rara a ser pesquisada (ROMA et al., 2013). Caso não haja complementaridade nesta posição, a hidroxila 3'OH do iniciador não estará disposta de forma a permitir a polimerização pela Taq polimerase, o que impede a extensão da

cadeia. Na *castPCR*, a variante comum é ocultada da Taq polimerase, por ser complementar a um nucleotídeo no meio da sequência de um oligonucleotídeo, específico para a sequência flanqueadora da mutação. Este oligonucleotídeo também apresenta um ligante de sulco secundário de DNA (MGB, do inglês *minor groove binder*) na extremidade 3', que garante alta estabilidade da fita dupla, resultante da hibridação do oligonucleotídeo com a fita simples de DNA, e impede a polimerização da cadeia. Com essa estratégia, é possível inibir a amplificação das fitas com as variantes comuns, sem interferir na amplificação dos alelos raros. A mesma fita complementar ao iniciador específico ou ao oligonucleotídeo também é reconhecida por sondas de hidrólise, similares às da qPCR (TaqMan®). Contudo, somente a fita contendo a variante rara, será amplificada, emitindo um sinal fluorescente (BARBANO et al., 2015) (FIGURA 4). Para determinar se a presença ou a ausência de amplificação é específica, os ensaios de detecção de mutação amplificam uma região conservada do gene de interesse, chamada reação de referência (ROMA et al., 2013).

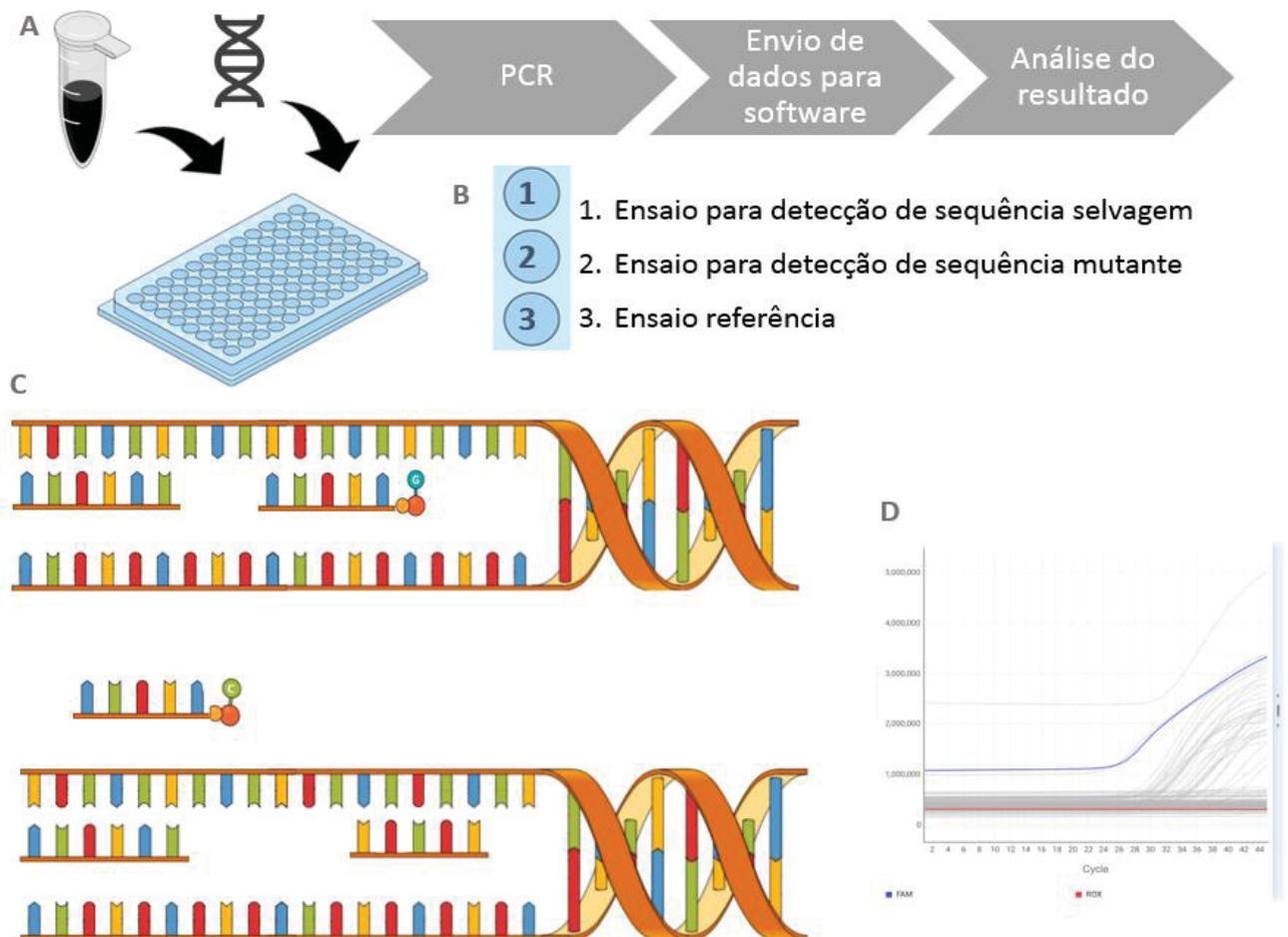


Figura 4. Detecção de mutações nos genes *EGFR* e *KRAS* por *castPCR*.

Legenda: (A) protocolo de *castPCR*, a partir da adição da mistura de reação contendo iniciadores, sondas de hidrólise e oligonucleotídeo com MGB, e DNA, seguida de reação de PCR até a análise do resultado; (B) modelo de distribuição das amostras na placa de PCR, na qual cada amostra foi submetida a um (1) ensaio para detecção da sequência selvagem, (2) um ensaio para detecção da sequência mutante e (3) um ensaio referência; (C) ensaio *castPCR* para detecção da sequência mutante. Na figura está representado apenas o oligonucleotídeo iniciador direto. A figura acima representa a sequência de DNA contendo a mutação de interesse, enquanto a figura abaixo representa sequência de DNA que não a contém. Sondas alelo específicas foram elaboradas para cada uma das sequências. Junto à sequência de DNA selvagem foi adicionada uma sequência complementar a região de interesse, não permitindo a hibridização da sonda específica e, por consequência, inibindo a emissão de fluorescência da sequência selvagem; (D) exemplo de resultado de curva de amplificação *castPCR* resultante da análise de dados pelo software *QuantStudio™ Design and Analysis*.

Fonte: (A autora, 2018)

2.2.3. Identificação de mutações somáticas por PCR digital (dPCR)

A PCR digital (dPCR) tem se tornado uma estratégia promissora na Biologia Molecular, por combinar quantificação absoluta e alta sensibilidade na detecção de sequências mutadas em ácidos nucleicos (HUGGETT et al., 2013; ZONTA et al., 2016). A dPCR é baseada no conceito de diluição limitante, onde as moléculas alvo são randomicamente divididas em múltiplas partições (HINDSON et al., 2013; RUTSAERT et al., 2018). Dessa forma, a técnica é capaz de detectar até uma célula com uma mutação somática específica, em aproximadamente 200.000 células que não apresentam a mesma mutação (PEKIN et al., 2010).

Primeiramente descrita em 1992 por Sykes e colaboradores como BEAM (*Beads, Emulsion, Amplification and Magnetics*), a dPCR está baseada na combinação de diluição limitante, PCR alelo-específica e estatística de Poisson para mensurar as concentrações de ácidos nucleicos na reação (SYKES et al., 1992). O termo “PCR digital” foi proposto em 1999 por Vogelstein e Kinzler (VOGELSTEIN; KINZLER, 1999), num estudo com amostras previamente diluídas de pacientes com câncer colorretal, para detecção de mutação no gene *RAS*. Neste trabalho, os autores já sugeriram potenciais aplicações para dPCR, como detecção de mutações pontuais, translocações cromossômicas, alterações de expressão gênica e discriminação alélica. Atualmente, seu uso como ferramenta de quantificação está consolidado na Biologia Molecular (RUTSAERT et al., 2018).

A técnica é baseada na distribuição das reações de amplificação em milhares de compartimentos individuais, de modo que uma única molécula de DNA sirva como substrato para a reação, por compartimento (VOGELSTEIN; KINZLER, 1999). Antes da amplificação, a amostra é diluída a determinada concentração e dispersa em micro reações unitárias, o que resulta em uma ou nenhuma molécula/sequência alvo de DNA em cada unidade (HUGGETT et al., 2013; TONG et al., 2017). Após a amplificação com qPCR, um valor de limiar (*threshold*) é aplicado aos valores de fluorescência, dividindo as partições em populações de reações positivas e negativas (RUTSAERT et al., 2018). Cada unidade contendo as cópias da molécula/sequência alvo de DNA exibe um sinal positivo (definido como “1”), enquanto apenas o fundo (*background*) de fluorescência é observado nas unidades sem nenhuma molécula de DNA (definido como “0”) (RUTSAERT et al., 2018; TONG et al., 2017; VOGELSTEIN; KINZLER, 1999). A estatística da distribuição de Poisson é aplicada para quantificar a probabilidade média de detecção do analito e a fração numérica de unidades positivas, de forma a reduzir os erros gerados para presença de mais de uma molécula alvo, na mesma unidade (HINDSON et al., 2013; RUTSAERT et al., 2018). Dessa forma, o número de moléculas inicial e a concentração de moléculas de DNA pode ser obtida, sem a necessidade de uma curva padrão ou de controles endógenos (HUGGETT et al., 2013) (FIGURA 5).

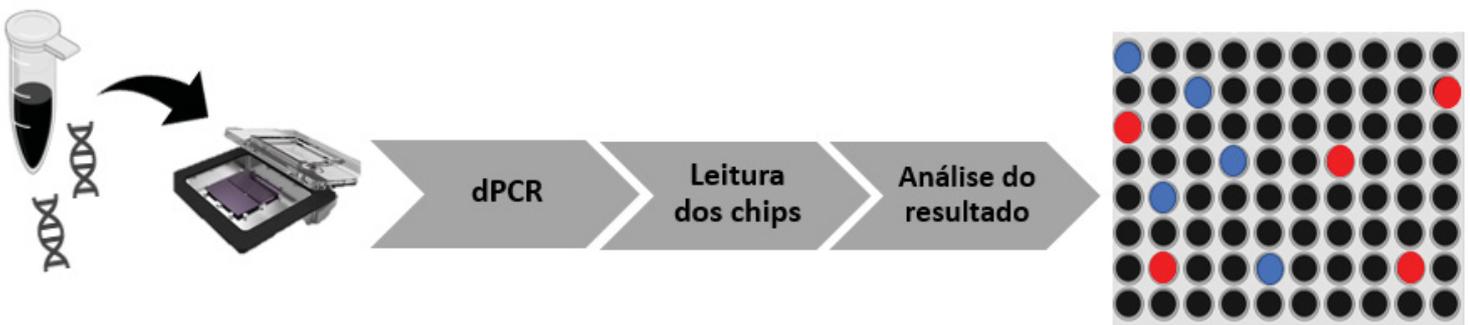


Figura 5. Detecção de variantes pela técnica PCR digital.

Legenda: À mistura de reação com as sondas específicas para detecção dos diferentes alelos é adicionado DNA da amostra de interesse e realizada reação de PCR, com posterior leitura dos chips e análise dos resultados.

Fonte: (A autora, 2018).

Originalmente, as partições do protocolo de dPCR eram criadas distribuindo a amostra manualmente através dos poços de PCR (VOGELSTEIN; KINZLER, 1999). A metodologia proposta por Vogelstein e colaboradores foi adaptada, de forma a ser

comercialmente viável, e atualmente é possível a realização da amplificação por diferentes tecnologias (TABELA 4). O sistema de detecção por dPCR disponível pela *Applied Biosystems* (QuantStudio 3D) consiste em reações individuais de PCR em um chip, no qual cada amostra é dispensada em um ensaio (LAIG et al., 2016). A leitura dos chips é realizada em um equipamento detector de fluorescência, que analisa cada partição individual como positiva ou negativa para a presença de amplificação do DNA alvo. Já os ensaios disponibilizados pela Bio-Rad (QX200™ Droplet Digital™ PCR) e *RainDance™ technologies* (*RainDrop plus™ Digital PCR system*) usam o sistema de PCR em emulsão para gerar as partições (*droplets*), nas quais cada uma contém uma ou nenhuma cópia do alvo de interesse, de forma a possibilitar amplificações individuais (RUTSAERT et al., 2018). Um detector de fluorescência é utilizado para analisar os *droplets*, e o modelo de Poisson é aplicado para mensurar o número absoluto de cópias das sequências alvo (ZHANG et al., 2015). As plataformas *BioMark™ HD System (Fluidigm)* e *CONSTELLATION® Digital PCR System (Formulatrix)* utilizam microplacas que permitem a distribuição da amostra, de forma que ocorram reações de PCR individuais em cada partição, e a plataforma *Naica (Stilla)* utiliza uma combinação de ensaios alelo específicos e PCR em emulsão (RUTSAERT et al., 2018).

Tabela 4. Principais plataformas de PCR digital.

Plataforma de PCR digital (marca)	Tipo	Detecção por reação de PCR	Amostras por protocolo		Volume de reação por amostra	Reações por amostra	Software de análise	Detecção
			Carregamento	Amplificação				
Biomark (Fluidigm™)	Circuitos fluidicos integrados	Tempo real e ponto final	48 / 12 amostras por ensaio	48 / 12 amostras por ensaio	4 µL / 8 µL	770 / 765	Digital PCR Analysis	Até 3 fluoróforos diferentes por ensaio
QuantStudio™ 3D (Applied Biosystems Life Technologies)	Chip de dPCR	Ponto final	1 amostra por chip	24 amostras	14,5 µL	20.000	QuantStudio™ 3D AnalysisSuite™	FAM VIC
Nalca (Stilla Technologies)	Matriz de gotas de cristal	Ponto final	4 amostras por chip	3 chips ou 14 amostras	25 µL	25.000 – 30.000	Crystal Miner	FAM VIC Cy5
RainDrop plus™ (RainDance™ Technologies)	Gotas de tamanho pico	Ponto final	8 amostras	96 amostras	25 a 50 µL	5 – 10 milhões	RainDrop Analyst II™	FAM VIC
QX200™ Droplet Digital™ PCR (Bio-Rad)	Gotas de tamanho nano	Ponto final	96 amostras	96 amostras	20 µL	20.000	QuantaSoft™	FAM VIC
CONSTELLATION® DPCR (Formulatrix)	Câmaras microfluidicas	Ponto final	26 / 96 amostras	26 / 96 amostras	10 µL	36.000 / 8000	CONSTELLATION®	Até 5 fluoróforos diferentes por amostra

Legenda: Cy5= cianina 5, FAM = 6-carboxyfluoresceína, VIC=2'-cloro-7'fenil-1,4-dicloro-6-carbóxi-fluoresceína.

Nota: reação de ponto final (do inglês, *end-point*) permite estimar a quantidade de DNA molde do início da reação, com a utilização da quantidade de DNA presente no final da reação, visto que a concentração de DNA dobra a cada ciclo de PCR.

Fonte: adaptado de RUTSAERT et al., 2018

A técnica de dPCR tem sido extensivamente utilizada na detecção de DNA em diversas áreas (MORLEY, 2014), como monitoramento de biomarcadores neoplásicos (TATEMATSU et al., 2017), detecção de patógenos (RUTSAERT et al., 2018), análise de expressão gênica (PATEL et al., 2014) e monitoramento ambiental e alimentar (HUGGETT; COWEN; FOY, 2015; TONG et al., 2017). Ao comparar as técnicas de dPCR e qPCR na detecção de SNP em um modelo viral, Whale e colaboradores observaram que a dPCR foi capaz de identificar amostras com concentrações extremamente baixas de células portadoras da mutação de resistência observada (0,1%), superior ao encontrado pela qPCR (5%) (WHALE et al., 2016). Diferente da qPCR, na qual o produto de amplificação com a variante rara pode não ser mascarado pela análise global da sequência alvo em maior quantidade, a análise individual de moléculas realizada pela dPCR (FIGURA 6) permite que tal técnica tenha maior sensibilidade na detecção de variantes raras, podendo ser usada em estudos de variantes raras associadas resistência a medicamentos (PEREZ-TORALLA et al., 2015).

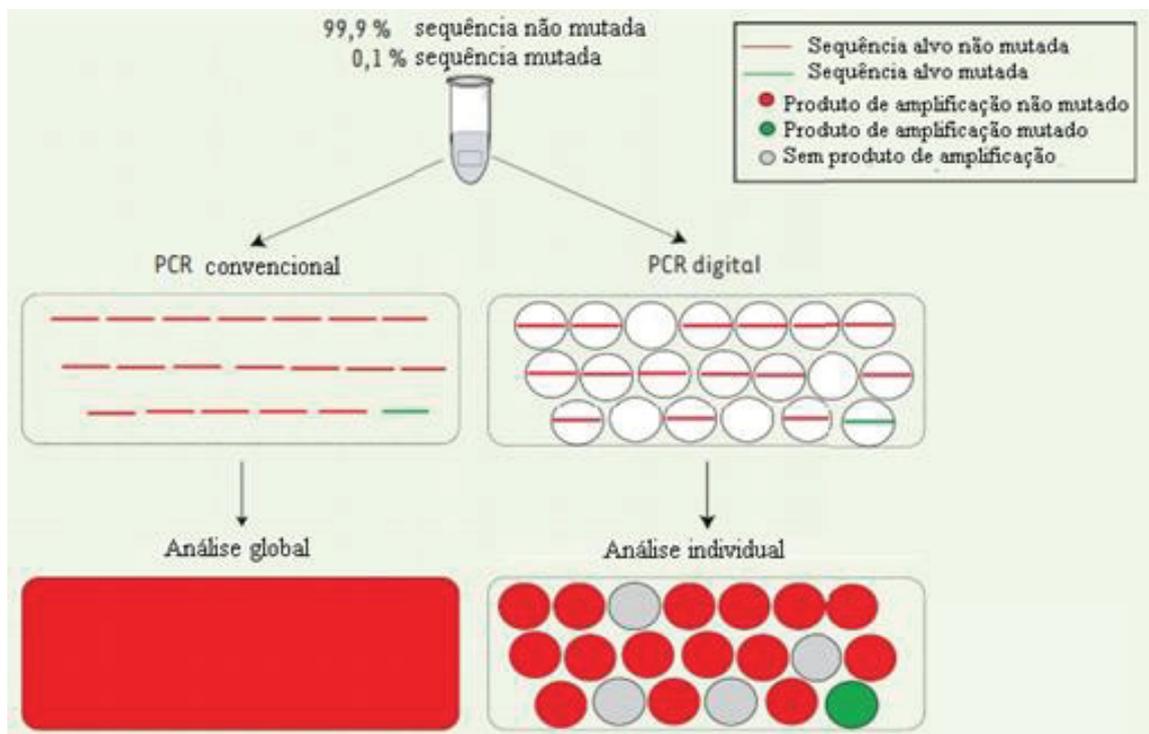


Figura 5. Comparação dos procedimentos de PCR em tempo real convencional e PCR digital.

Nota: exemplo de análise de uma amostra contendo 99,9% de alelos sem a mutação de interesse e 0,1% de alelos mutantes.

Fonte: Adaptado de PEREZ-TORALLA et al., 2015.

Alguns autores (RUTSAERT et al., 2018) indicam a dPCR como ferramenta crucial nos estudos das neoplasias (ZONTA et al., 2016), bem como no diagnóstico, tratamento personalizado e monitoramento (*follow-up*) dos pacientes (TONG et al., 2017), uma vez que marcadores genéticos são capazes de refletir a patogênese, progressão e prognóstico das neoplasias (MA et al., 2017; SEPULVEDA et al., 2017). Além disso, alguns marcadores são especificamente expressos em determinados tipos de câncer, considerados candidatos promissores no diagnóstico e pesquisa básica das neoplasias (TONG et al., 2017). Dessa forma, a detecção de mutações gênicas pode ser usada para monitorar metástases e/ou recidiva de câncer (SALVIANTI; PINZANI, 2017; TONG et al., 2017). Atualmente, a técnica de dPCR é utilizada na detecção de mutações em genes como: *EGFR* em câncer de pulmão de células não pequenas (WHALE et al., 2012; YUNG et al., 2009), *KRAS* em câncer colorretal e mutações e *ESR1* (do inglês, *estrogen receptor 1*) em câncer de mama (TONG et al., 2017).

A localização do tumor ou as condições do paciente podem fazer com que a biópsia tecidual se torne um procedimento inviável (FRANCIS; STEIN, 2015). Alguns pesquisadores sugerem que aproximadamente 20% de todas as biópsias de pacientes com câncer de pulmão sejam inadequadas para análise molecular, sugerindo o uso da biópsia líquida (provinda de urina, sangue ou plasma) para a detecção de mutações (OXNARD et al., 2012). Nessa abordagem, as análises genéticas são feitas a partir do plasma de pacientes com a neoplasia. Cânceres de todos os estágios são capazes de liberar células no sistema sanguíneo, e a detecção de mutações tumorais pode se dar através de DNA tumoral circulante (ctDNA, do inglês, *circulating tumor DNA* ou CTC, do inglês, *circulating tumor cells*) (DONALDSON; PARK, 2018). Contudo, a análise de ctDNA apresenta desafios para o diagnóstico clínico, haja visto a dificuldade na estabilidade do material após coleta e isolamento e rendimento do ctDNA, além da quantidade maior de DNA circulante não tumoral liberado pelas células normais, em comparação ao ctDNA liberado pelos tumores. Portanto, para sua detecção, é necessária tecnologia sensível e específica, capaz de detectar proporções tão pequenas quanto 0,1-1% do alvo (RACHIGLIO et al., 2016). Nesses casos, dPCR pode ser um método alternativo, não invasivo, de monitoramento da terapia com inibidores de tirosina-quinase através da detecção de mutações no gene *EGFR* em amostras de urina (TONG et al., 2017), plasma e sangue. Por meio desta técnica, a biópsia líquida permite um monitoramento mais eficaz dos

pacientes, além de ser um método menos invasivo, quando comparado a uma biópsia tecidual (SALVIANTI; PINZANI, 2017).

3. HIPÓTESE E JUSTIFICATIVA

O presente trabalho foi uma parceria entre a Universidade Federal do Paraná e o Laboratório Genoprimer – Diagnóstico Molecular e Medicina Personalizada. A empresa brasileira de inovação e pesquisa Finep realizou um processo de seleção por meio de chamada pública, visando criar condições financeiras favoráveis e apoiar a inovação. O chamado PROJETO TECNOVA teve como objetivo principal o crescimento rápido de empresas de micro e pequeno porte, com foco no apoio à inovação tecnológica, com apoio da Finep para realizar todas as atividades operacionais inerentes ao processo, incluindo fomento, análise e seleção das propostas, contratação, liberação dos recursos, acompanhamento físico e financeiro com a prestação de contas, assegurando o foco nos projetos de inovação e desenvolvimento tecnológico.

Para tal projeto, a nossa hipótese é que a técnica de PCR digital (dPCR) apresenta sensibilidade e especificidade superiores às técnicas de sequenciamento de Sanger e castPCR, atualmente utilizadas na rotina diagnóstica, para detecção de mutações somáticas associadas à resposta ao tratamento contra NSCLC e CRC, em amostras previamente fixadas em formalina e emblocadas em parafina (FFPE). Estas amostras frequentemente são as únicas a disposição, especialmente para estudos epidemiológicos retrospectivos. O método padrão atual para detecção de mutações somáticas em tais amostras é o sequenciamento de DNA. Entretanto, tal técnica requer que entre 10% a 25% das células contenham a variante de interesse, limitação que pode ter implicações na avaliação de possíveis marcadores que condicionam a eficiência terapêutica, como é o caso de mutações em *EGFR* e *KRAS*. Em situações como esta, métodos que empreguem amplificação privilegiada das variantes raras e/ou baseados em amplificação de moléculas de DNA isoladas por reação em cadeia em cadeia da polimerase (PCR), têm se mostrado promissores para detecção de mutações somáticas, além de mais sensíveis e acurados que o sequenciamento clássico (BOLTON et al., 2015).

4. OBJETIVOS

4.1. OBJETIVO GERAL

Desenvolver, padronizar e validar testes diagnósticos baseados em dPCR para detecção de mutações somáticas nos genes *EGFR* e *KRAS*, em DNA oriundo de biópsias FFPE.

4.2. OBJETIVOS ESPECÍFICOS

- Avaliar o desempenho analítico da dPCR para detecção de mutações somáticas no gene *EGFR* associadas ao câncer de pulmão do tipo células não pequenas, utilizando amostras de controle positivo sintético;
- Avaliar o desempenho analítico da dPCR para detecção de mutações somáticas no gene *KRAS* associadas ao câncer colorretal, utilizando amostras de controle positivo sintético.
- Avaliar o desempenho do uso clínico da dPCR para detecção de mutações somáticas em *EGFR* e *KRAS*, através da comparação de mutações detectadas por dPCR em amostras FFPE com sequenciamento de Sanger;
- Avaliar o desempenho do uso clínico da PCR digital para detecção e quantificação de mutações somáticas em *EGFR* e *KRAS*, através da comparação de mutações detectadas por dPCR em amostras FFPE com castPCR.

5. MATERIAIS E MÉTODOS

5.1. COMITÊ DE ÉTICA

O projeto foi aprovado pelo Comitê de Ética em Pesquisa da Liga Paranaense de Combate ao Câncer do Hospital Erasto Gaertner, segundo parecer 885.643/2014.

5.2. VALIDAÇÃO ANALÍTICA

A validação analítica foi realizada com o uso de controles positivos, nos testes de linearidade e eficiência, limites mínimos de detecção (LoD) e quantificação (LoQ), precisão e especificidade.

5.2.1. Produção do DNA sintético como controle positivo para mutações nos genes *EGFR* e *KRAS*

Oligonucleotídeos contendo as diferentes variantes para cada loco foram desenhados e sintetizados (IDT, Coralville, Iowa, USA), tanto para a variante mais comum, quanto para as mutações pontuais nos éxons 18, 20 e 21 do gene *EGFR* e 2 do gene *KRAS* (ANEXO 1). Através da ferramenta disponível online *OligoAnalyzer 3.1* os oligonucleotídeos foram analisados em relação a formação de dímeros e grampos. Após a síntese, cada um dos oligonucleotídeos foi ressuspensa em tampão TE e diluídos de forma seriada, partindo de 1.000.000.000 moléculas/ μ L até 1 molécula/ μ L, para posterior utilização na padronização dos experimentos.

5.3. PCR DIGITAL

Para a padronização do protocolo de PCR digital e publicação dos resultados, seguimos as recomendações do *Minimum Information for Publication of Quantitative Digital PCR Experiments* (HUGGETT et al., 2013).

Para a realização dos experimentos, os oligonucleotídeos iniciadores de reação foram desenhados para esta técnica, utilizando os seguintes parâmetros: tamanho entre 18 e 25 pares de bases, temperaturas de acoplamento próximas a 60°C e ausência de formação de dímeros e grampos, através da ferramenta disponível online *OligoAnalyzer 3.1*. Os mesmos foram sintetizados pela IDT (Coralville, Iowa, USA). Da mesma forma, as sondas foram desenhadas seguindo parâmetros de temperaturas de acoplamento e analisadas quanto ausência de formação de dímeros, através da ferramenta disponível online *OligoAnalyzer 3.1*. No entanto, a temperatura de acoplamento das sondas deve ser aproximadamente 10°C superior a dos oligonucleotídeos, visto que estas se hibridizam antes na reação de PCR em tempo real. Além disso, as sondas foram duplamente marcadas, contendo um fluoróforo *reporter* como FAM[™] (6-carboxylfluoresceína) ou VIC[®] (2'-cloro-7'fenil-1,4-dicloro-6-carbóxi-fluoresceína), covalentemente ligado a extremidade 5', pela *Life Technologies* (Carlsbad, Ca, USA). Preparamos uma mistura de reação com 900 nM de cada oligonucleotídeo iniciador, 200 nM de sonda de hidrólise TaqMan[®] para detecção do alelo mais frequente, 200 nM de sonda de hidrólise TaqMan[®] para detecção do alelo mutante, conforme as tabelas em anexo (ANEXO 2 e ANEXO 3), 1X *QuantStudio 3D Digital PCR Mastermix* (*Thermo Fisher Scientific*), 3 ng de DNA e água ultrapura (TABELA 5). Embora o recomendado para as reações de PCR digital

seja de 20 ng de DNA, utilizamos 3 ng devido à qualidade das amostras extraídas, e também com o objetivo de padronização das técnicas utilizadas.

Tabela 5. Mistura de reação para realização dos experimentos de PCR digital.

	Inicial	Final
<i>QuantStudio 3D Digital PCR Mastermix</i>	2X	1X
<i>Primer direto (nM)</i>	100000	900
<i>Primer reverso (nM)</i>	100000	900
Sonda para detecção de alelo selvagem (nM)	100000	200
Sonda para detecção de alelo mutante (nM)	100000	200
Água ultrapura		--
DNA	1ng/μL	3ng

A mistura de reação foi carregada em um chip de dPCR com 20 mil partições, utilizando o equipamento *QuantStudio™ 3D Digital PCR Chip Loader*. O volume médio individual por partição foi de 0,75 nL, e foi previsto um número de moléculas de DNA por partição inferior a 1, de acordo com a fórmula abaixo.

$$(1 \text{ genoma} \div 3.3\text{pg}) \times (1 \text{ ng} \div \mu\text{L}) \times (10^3\text{pg} \div 1 \text{ ng}) \times 3\mu\text{L} \div 20000$$

Dessa forma, algumas partições foram preenchidas com o alvo molecular de interesse, enquanto outras permaneceram apenas com a mistura de reação, situação essencial para o funcionamento adequado da técnica. Após o carregamento, o chip foi selado e levado ao termociclador, com a seguinte ciclagem: 96°C por 10 minutos, 39 ciclos de 60°C por 2 minutos e 98°C por 30 segundos, 60°C por 10 minutos e 10°C por 4 minutos (TABELA 6).

Tabela 6. Protocolo de ciclagem dos experimentos de PCR digital.

Temperatura	96°C	60°C	98°C	60°C	10°C
Tempo	00:10:00	00:02:00	00:00:30	00:10:00	00:04:00
Ciclos	1	39		1	1

Finalizada a corrida de PCR, o chip foi analisado no equipamento *QuantStudio™ 3D (Thermo Fisher Scientific)*, e o resultado da leitura exportado para o software *QuantStudio 3D AnalysisSuite Cloud*. A análise dos dados de dPCR foi feita pelo software na nuvem, considerando um intervalo de confiança de 95% e uma variação na precisão de até 10%, utilizando o modelo de estatística de Poisson.

5.3.1. Detecção de mutações de ponto nos genes *EGFR* e *KRAS*

Ensaio específicos foram elaborados para a detecção das variantes raras nos éxons 20 e 21 do gene *EGFR* e no éxon 2 do gene *KRAS*. Em cada mistura de reação, foram adicionados 900 nM de cada oligonucleotídeo iniciador, direto e reverso, e 200 nM de cada sonda de hidrólise marcadas com fluoróforo (repórter a 5' e *quencher* a 3' - TaqMan®) (ANEXO 2), uma para detecção da sequência com a mutação específica e outra para detecção da sequência sem a mutação. As sondas foram marcadas com fluoróforos excitáveis em diferentes comprimentos de onda (FAM e VIC), de forma que a detecção do alvo molecular de interesse seja específica. A essa mistura de reação, foi adicionada a *QuantStudio 3D Digital PCR Mastermix (Thermo Fisher Scientific)*, água ultrapura e o DNA das amostras de interesse. A reação de dPCR foi levada ao termociclador, conforme o protocolo padronizado no item 2.5, para posterior leitura dos chips no equipamento *QuantStudio™ 3D (Thermo Fisher Scientific)* e análise dos resultados.

Para as variantes do *rs28929495* no éxon 18 do gene *EGFR*, as sondas marcadas para as diferentes possibilidades de substituição nessa posição (p.G719A (*g.55174014G>C*), p.G719S (*g.55174014G>A*), p.G719C (*g.55174014G>T*)) foram unidas em uma única mistura de reação (FIGURA 7). O objetivo foi detectar a presença de mutação sem discriminação alélica, visto que a presença destas mutações é rara e a decisão terapêutica é independente da sua identidade. Dessa forma, para o éxon 18 do gene *EGFR*, foi possível apenas diferenciar o alelo com a mutação (FIGURA 7B) daquele selvagem (FIGURA 7A), detectando qualquer um dos alelos mutantes, seja este uma adenina, citosina ou timina (FIGURA 7B).

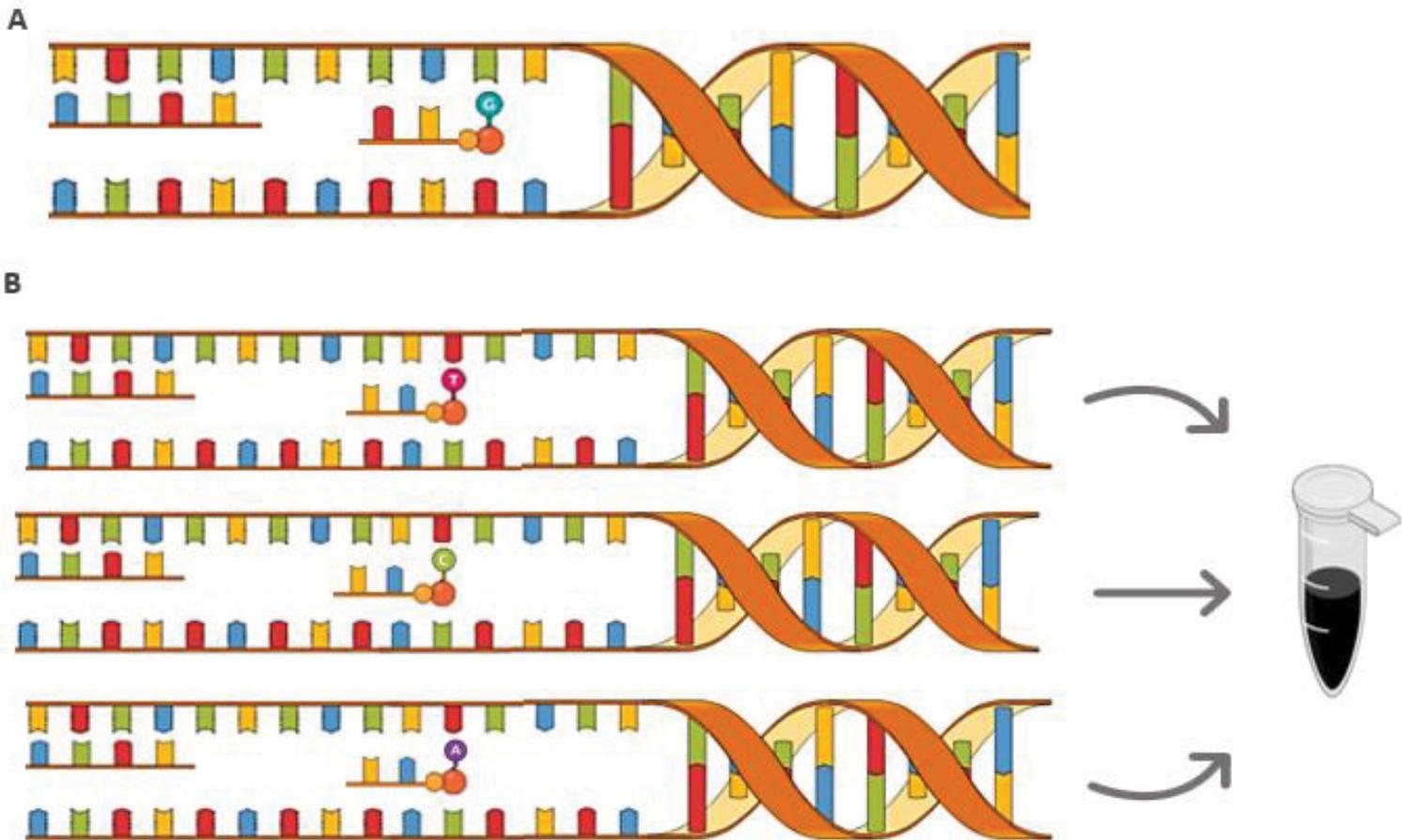


Figura 7. Sistema de detecção para a mutação *rs28929495* no éxon 18 do gene *EGFR*.

Legenda: (A) detecção de indivíduos sem mutação; (B) detecção de indivíduos com mutação pontual do éxon 18 por diferentes sondas específicas.

Nota: Quando em uma única mistura de reação, a detecção da mutação não possui discriminação alélica.

Fonte: (A autora, 2018).

5.3.2. Detecção da deleção do éxon 19 do gene *EGFR*

O ensaio elaborado para detecção da deleção do éxon 19 consiste em dois oligonucleotídeos iniciadores, um direto e um reverso, e duas sondas de hidrólise (TaqMan®) (ANEXO 2). Uma das sondas, marcada com o fluoróforo VIC®, foi desenhada para ser complementar a uma região conservada do gene *EGFR*. A outra sonda, marcada com o fluoróforo FAM™, foi desenhada para ser complementar a uma região onde ocorrem as deleções mais frequentes. Dessa forma, o indivíduo que não possui a deleção no éxon 19 possui essa região intacta, e com a hibridação da sonda e emissão de fluorescência, é possível identificar a amplificação. Ao passo que

quando essa região é deletada, a sonda não é capaz de se acoplar, e não há emissão de fluorescência da sonda marcada com FAM™. No entanto, a sonda marcada com VIC®, por estar em uma região independente da deleção, é detectada com a amplificação, o que garante o funcionamento da reação e a presença de amostra com a mutação de deleção do éxon 19 (FIGURA 8).

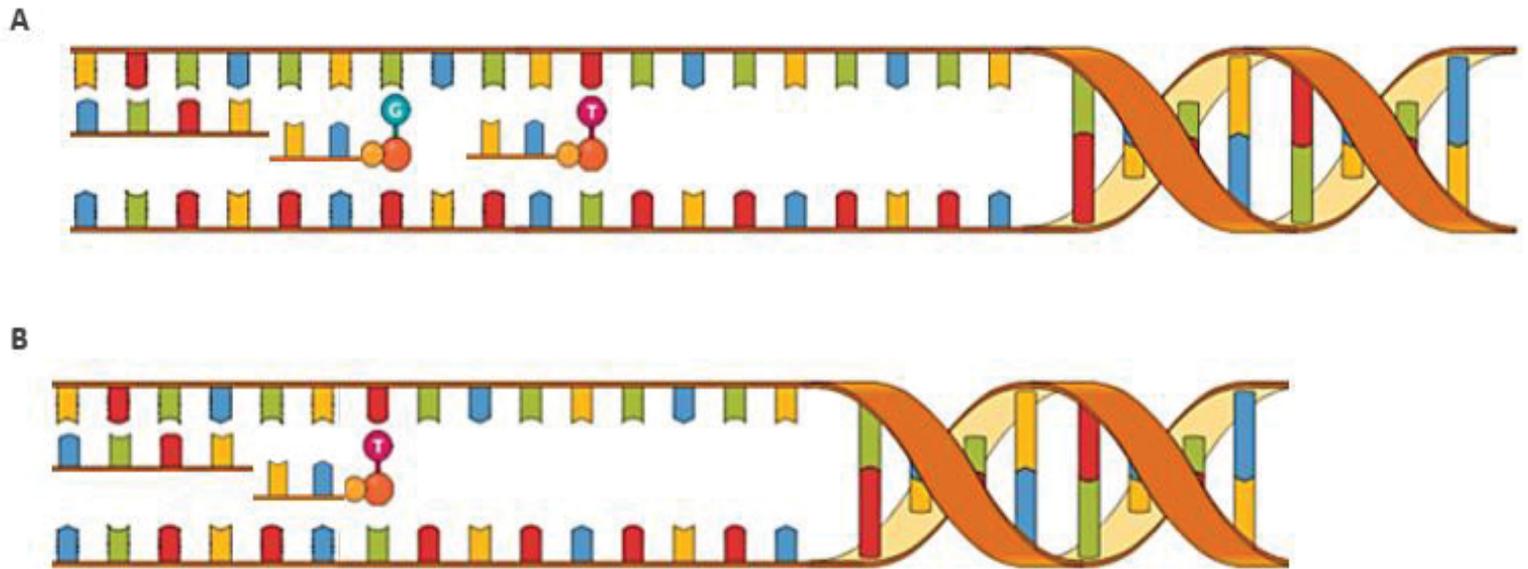


Figura 8. Sistema de detecção da deleção do éxon 19 no gene *EGFR*.

Legenda: (A) detecção de indivíduos sem mutação; (B) detecção de indivíduos com deleção do éxon 19. Na figura (A) é possível identificar duas sondas capazes de hibridizar ao DNA de um indivíduo sem a deleção do éxon 19, enquanto na figura (B) a região na qual uma das sondas poderia hibridizar foi deletada.

Nota: Os nucleotídeos representados são apenas modelos, não correspondendo, necessariamente, ao que acontece nas amostras avaliadas. Em (A) e (B) apenas o oligonucleotídeo iniciador direto está representado, no sentido 3'-5' da fita molde.

Fonte: (A autora, 2018).

5.3.3. Limite de Detecção, Eficiência e Linearidade

A fim de avaliar a qualidade de detecção do alvo molecular de interesse, as misturas de reação foram feitas com oligonucleotídeos iniciadores e sondas específicas marcadas para os alelos mutante e selvagem, e estas foram usadas durante toda a validação. A precisão e a especificidade também foram avaliadas nas amostras de blocos parafinados, durante a validação clínica.

Para o experimento de linearidade, foram realizadas diluições seriadas dos controles positivos, partindo de 1.000.000.000 moléculas/ μ L até 1 molécula/ μ L. Em cada tubo, foi adicionado 9X de tampão de eluição com 10% de carreador (*carrier*) (QIAGEN). A partir do controle de maior concentração foi aliquotado 1X, e a solução foi homogeneizada 5 vezes por pipetagens sucessivas, e passada no vórtex por 10 segundos. Essa etapa foi repetida até que todos os controles tenham sido diluídos. Após serem testados na plataforma de PCR em tempo real, os controles positivos sintéticos foram testados na plataforma de PCR digital (*QuantStudio® 3D Digital PCR System*), tanto para o alelo mais comum, quanto para alelos com as mutações. Os dados de linearidade foram utilizados para calcular a eficiência do sistema. A última diluição de DNA que continha a mutação encontrada no teste de linearidade, foi definida como o limite de detecção (LoD) do método, a menor quantidade de analito presente em uma amostra, que pode ser detectado.

Os oligonucleotídeos sintetizados como controles que continham as sequências dos alelos mutados foram adicionados em meio às sequências dos alelos sem mutação, nas proporções de 0,5%, 1%, 5%, 10% e 50%. A menor quantidade do analito de interesse quantificada, foi definida como o limite de quantificação (LoQ) do método. A fim de avaliar se as quantificações se mantiveram consistentes, a cada corrida realizada, foi adicionada um controle positivo para uma das mutações analisadas. Dessa forma, foi possível avaliar todas as replicatas dos controles positivos, tanto na mesma corrida quanto em dias consecutivos, e o grau de concordância entre os valores das replicatas representa a precisão do método de dPCR. A precisão foi calculada pelo software de análise dos chips (*QuantStudio 3D Analysis Suite*), pois este avalia não só as replicatas dos chips, mas também cada reação de dPCR como uma reação individual, e considera como aceitável, um desvio de precisão de até 10%.

5.4. VALIDAÇÃO DIAGNÓSTICA

5.4.1. Amostras

Para a validação diagnóstica, dispomos de amostras previamente fixadas em formalina e emblocadas em parafina (FFPE) de biópsias realizadas entre os anos de 2003 a 2007 doados pelo Hospital Nossa Senhora das Graças, em colaboração com o Laboratório de Citogenética Humana e Oncogenética da Universidade Federal do Paraná. Por meio de uma parceria com o Hospital Erasto Gaertner, obtivemos biópsias

em tecido parafinado com presença de tumor pulmonar, realizadas entre os anos de 2007 a 2015. Todas as biópsias foram previamente avaliadas por médicos patologistas, e selecionadas e separadas de acordo com o diagnóstico: 41 amostras de câncer colorretal e 57 amostras de câncer de pulmão. Foram obtidas informações relativas a sexo, idade dos pacientes e grau de diferenciação tumoral (TABELA 7).

Tabela 7. Informações relativas a sexo e idade dos pacientes, assim como grau de diferenciação tumoral, para Câncer de Pulmão e Câncer Colorretal.

	Câncer de Pulmão (n = 57)	Câncer Colorretal (n = 41)
Média de idade	63 ± 11	63 ± 13
Masculino	46%	52%
Feminino	53%	48%
Sem informação	1%	0%
<i>Grau de diferenciação tumoral</i>		
Baixo	26,3%	2,4%
Moderado	35,1%	54,8%
Alto	15,8%	38,1%
Sem informação	22,8%	4,8%

Nota: as amostras de câncer colorretal foram obtidas do hospital Nossa Senhora das Graças, enquanto que as amostras de câncer de pulmão foram obtidas dos hospitais Nossa Senhora das Graças e Erasto Gaertner.

Como controle de especificidade, utilizamos 10 blocos parafinados de tecido pulmonar livre de neoplasia dos pacientes do Hospital Nossa Senhora das Graças, e 7 amostras de biópsia fresca de cólon de pacientes do mesmo hospital.

5.4.2. Extração de DNA

Parte selecionada das amostras FFPE foram dispostas em um tubo de 1,5 mL e submetidas a remoção da parafina com 1 mL de xileno, por 18 horas. Em seguida, as amostras foram centrifugadas por 5 minutos em rotação máxima (21382g), e o sobrenadante contendo o xileno foi cuidadosamente removido com pipeta Pasteur. Adicionou-se novamente 1 mL de xileno, e as amostras permaneceram em agitação em vórtex ou no equipamento *shaker* a 23°C por 30 minutos, para melhor eficiência na remoção da parafina. Repetiu-se a etapa de centrifugação por 5 minutos em rotação máxima (21382g), e o sobrenadante contendo o xileno foi cuidadosamente removido com pipeta Pasteur. Em situações nas quais o xileno demonstrasse opacidade, o procedimento de desparafinização com foi repetido com adição de 1 mL de xileno e incubação por 30 minutos em agitação, para completa remoção da parafina.

Para remoção do xileno (FIGURA 9A), adicionou-se 1mL de etanol absoluto, seguido por centrifugação por 1 minuto em rotação máxima (21382g) e remoção do sobrenadante contendo o etanol com pipeta Pasteur. Para garantir a completa remoção do xileno, prosseguiram-se sucessivas lavagens com etanol 70%, 50% e água miliQ, todos seguidos por centrifugação por 1 minuto em rotação máxima (21382g) e remoção do sobrenadante com pipeta Pasteur. Quando necessário, as amostras eram incubadas em temperatura ambiente, em repouso por 15 minutos com o tubo aberto em lugar protegido, ou em bloco de aquecimento a 37°C por 5 minutos, garantindo a secagem do material. Por fim, a amostra seca pode ser armazenada a -20°C, ou prosseguir às etapas de extração do DNA.

Após a remoção da parafina, o DNA foi extraído das amostras secas através do kit *QIAamp DNA Micro Kit* (QIAGEN) (FIGURA 9B). Inicialmente, foi adicionado à cada amostra 80 µL de tampão de lise (QIAGEN), 1 µL de RNA *Carrier* (concentração de 1µg/ µL) e 15 µL proteinase K. O protocolo original do fabricante foi adaptado para tempo de lise de 18 horas (*overnight*) a 56°C, a fim de aumentar o rendimento final da amostra. Em seguida, as amostras passaram por uma breve centrifugação. As amostras em lise foram incubadas em banho seco a 56°C em agitação de 700 rpm, com mais 15 µL proteinase K, por pelo menos uma hora, até que todo o tecido tenha sido lisado. Por fim, as amostras foram incubadas em banho seco a 98°C por 20 minutos, e após breve centrifugação, e prosseguiram-se sucessivas lavagens, com o auxílio de uma coluna de extração (*QIAamp Mini Spin Column*) e os tampões previamente cedidos pelo fabricante (QIAGEN). Após as lavagens, as amostras foram centrifugadas por 3 minutos em rotação máxima (21382g), a fim de remover de forma completa os tampões de lavagem, e em seguida cada amostra foi eluída em 50 µL de tampão de eluição (*Elution Buffer QIAGEN*) e incubada a temperatura ambiente por 5 minutos. Após o descarte da coluna, as amostras extraídas puderam ser armazenadas ou prosseguir com o protocolo de purificação sugerido pelo fabricante.

As concentrações de DNA foram mensuradas por fluorimetria no equipamento *Quantus Fluorimeter* (Promega), e a diluição de DNA final foi ajustada para $1\text{ng}\mu\text{L}$. Os dados de qualidade e pureza foram mensurados por espectrofotometria no equipamento NanoDrop (*ThermoFischer Scientific*).

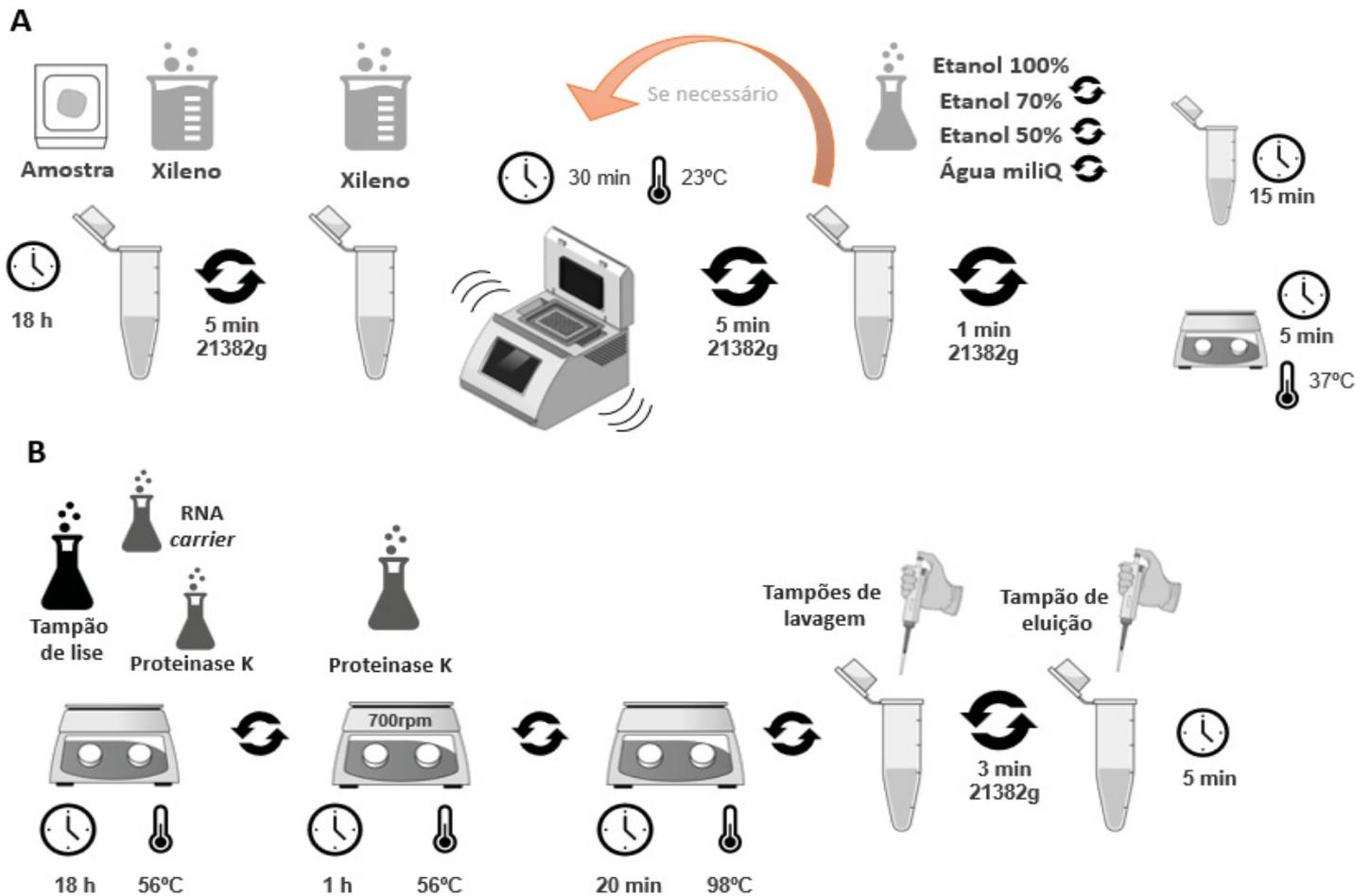


Figura 9. Etapas do protocolo de (A) remoção de parafina seguido de (B) extração de DNA das amostras de câncer de pulmão e câncer colorretal.

Nota: O protocolo consistiu em passos de lavagem com xileno, incubação com tampão de lise e proteinase K, seguido de diferentes etapas de lavagem, purificação e eluição. A concentração de DNA final foi ajustada para $1\text{ng}\mu\text{L}$.

Para a validação do uso na clínica da PCR digital, as amostras devem ser avaliadas pelo método padrão-ouro clínico ou laboratorial já estabelecido. Por essa razão, os resultados obtidos por PCR digital foram comparados com os de sequenciamento de Sanger. O sequenciamento de DNA foi realizado nos Laboratórios de Genética Molecular Humana e Laboratório de Imunogenética e

Histocompatibilidade da UFPR, e consiste em etapas de PCR convencional, purificação, sequenciamento e precipitação.

Além disso, utilizamos a PCR competitiva alelo específica como método comparativo adicional para avaliar a sensibilidade da dPCR.

5.4.3. Sequenciamento de Sanger

5.4.3.1. Amplificação dos fragmentos pela DNA polimerase

A reação de PCR foi realizada com os oligonucleotídeos iniciadores listados nas tabelas 4 e 5 para os éxons 18, 19, 20 e 21 do gene *EGFR* e 2 do gene *KRAS*. No total, foram necessárias 4 reações de amplificação. As amplificações foram realizadas para um volume final de 20 μ L, e cada reação continha: tampão 1X (Invitrogen); 0,3 mM de cada dNTP (*Life Technologies*); 2,5mM de MgCl₂ (Invitrogen); 0,4 μ M de cada primer específico; 0,8 U de *Taq polimerase Platinum*[®] (Invitrogen), 3 ng de DNA genômico e água ultrapura (TABELA 8). Embora o recomendado para as reações de amplificação seja de 20 ng de DNA genômico, utilizamos 3 ng visto a qualidade das amostras extraídas, e também com o objetivo de padronização das técnicas utilizadas.

Tabela 8. Mistura de reação para realização dos experimentos de amplificação dos fragmentos pela DNA polimerase.

	Inicial	Final
Tampão de reação	10X	1X
Primer direto (μ M)	100	0,4
Primer reverso (μ M)	100	0,4
dNTPs (nM)	2,5	0,3
<i>Taq</i> polimerase (U/ μ L)	5	0,8
Água ultrapura		--
DNA	1ng/ μ L	3ng

As reações foram executadas no termociclador *Veriti Thermal Cycler* (*Thermo Fisher Scientific*), utilizando as seguintes condições de ciclagem para cada amplificação: desnaturação inicial a 94°C por 5 minutos, 40 ciclos de 94°C por 20

segundos, 58°C por 30 segundos, 72°C por 30 segundos, seguido de um último passo de extensão final a 72°C por 5 minutos (TABELA 9).

Tabela 9. Protocolo de ciclagem dos experimentos de amplificação dos fragmentos pela DNA polimerase.

Temperatura	94°C	94°C	58°C	72°C	72°C
Tempo	00:05:00	00:00:20	00:00:30	00:00:30	00:05:00
Ciclos	1		40		1

Para confirmação da amplificação do fragmento, foram realizadas corridas eletroforéticas a 80 V por 40 minutos, aplicando-se 2 µL do produto de PCR junto a 2 µL de tampão de corrida, em gel de agarose 1,5% contendo 1 µL do corante de ácidos nucleicos *UniSafe Dye (Uniscience)*. Posteriormente, os géis de agarose foram visualizados e fotografados através de sistema de fotodocumentação computadorizado.

5.4.3.2. Purificação dos produtos amplificados

O total de 6µL de produto de PCR foi então homogeneizado e purificado através do uso de enzimas Exonuclease I (Fermentas), que degrada DNA de fita simples, e fosfatase alcalina (*Thermo Fisher Scientific*), que elimina grupos fosfato e inativa os nucleotídeos da reação anterior. Após centrifugação rápida, a reação foi levada ao termociclador em um ciclo de 60 minutos a 37°C, seguido de um ciclo de inativação enzimática a 80°C, por 20 minutos.

5.4.3.3. Sequenciamento de DNA

No total, foram realizadas 4 reações de sequenciamento, nas quais foi adicionado 1,9 µL de tampão de reação (*Big Dye® Terminator v.3.1 5X Sequencing Buffer*), 0,1 µL de cada oligonucleotídeo iniciador (direto e reverso) e 0,3 µL *BigDye® Terminator Cycle Sequencing Standard v3.1 (Life Technologies)*, a 2 µL do produto purificado. Para aquelas amostras cuja banda no gel de eletroforese era de baixa intensidade, utilizou-se 3 µL de produto purificado. Após centrifugação rápida, a mistura de reação foi levada ao termociclador para a reação de sequenciamento, que

consiste em um ciclo de 1 minuto a 95°C, 25 ciclos de 10 segundos a 95°C, um ciclo de 5 segundos a 50°C e, por fim, um ciclo de 4 minutos a 60°C (TABELA 10).

Tabela 10. Protocolo de ciclagem das reações de sequenciamento de Sanger.

Temperatura	95°C	95°C	50°C	60°C
Tempo	00:01:00	00:00:10	00:00:05	00:04:00
Ciclos	1	25	1	1

Após a reação de sequenciamento, foi realizada uma purificação do DNA antes da injeção das amostras no sequenciador, a fim de que os reagentes utilizados não interferissem na eletroforese capilar e leitura das bases. Logo, as amostras sequenciadas foram centrifugadas rapidamente, e foi adicionado 2,5µL de EDTA 125mM (Merck) e 30µL de etanol absoluto (Merck), ambos gelados, em cada amostra. Após homogeneização e centrifugação a 3220g por 45 minutos, o EDTA e o etanol absoluto foram removidos por inversão, e foi adicionado 30µL de etanol 70% em cada amostra, este em temperatura ambiente. Novamente houve homogeneização e centrifugação a 3220g e remoção do etanol por inversão. As amostras foram deixadas em temperatura ambiente e cobertas com papel absorvente, para completa remoção do etanol. Por fim, foi adicionado 10 µL de Hi-Di formamida (*Life Technologies*) por amostra, e a reação foi levada ao sequenciador.

A análise das sequências foi feita utilizando o software *Mutation Surveyor*® - *DNA Variant Analysis (Softgenetics)*, considerando as sequências de referência dos genes *EGFR* e *KRAS*. O sequenciamento do gene *EGFR* foi realizado duas vezes, utilizando os mesmos iniciadores e repetindo a PCR, para confirmação dos resultados.

5.4.4. PCR competitiva alelo específica (*castPCR*)

O experimento de *castPCR* foi realizado para detecção de mutações nos genes *EGFR* nas amostras de câncer de pulmão e *KRAS* nas amostras de câncer colorretal, com o objetivo de comparar tais resultados aos obtidos por dPCR. Cada amostra é submetida a três reações individuais de PCR, (i) uma para detecção do alelo mais frequente, (ii) uma para detecção do alelo contendo a mutação rara e, em todos os ensaios realizados, adicionamos (iii) uma reação de gene de referência. Esta reação detecta uma sequência conservada no gene de interesse, distante da mutação

estudada, a fim de determinar se a presença ou a ausência de amplificação é específica.

Para a realização dos experimentos de *castPCR*, preparamos uma mistura de reação com 1X de *TaqMan® Mutation Detection Assay*, que contém os oligonucleotídeos iniciadores e os bloqueadores, 2X *TaqMan Genotyping Mastermix*, 3 ng de DNA e água ultrapura (TABELA 11). Embora o recomendado para as reações de amplificação seja de 20 ng de DNA genômico, utilizamos 3 ng devido a qualidade das amostras extraídas, e também com o objetivo de padronização das técnicas utilizadas.

Tabela 11. Mistura de reação para realização dos experimentos de PCR competitiva alelo específica.

	Inicial	Final
<i>TaqMan® Mutation Detection Assay</i>	10X	1X
<i>TaqMan Genotyping Mastermix</i>	10X	2X
Água ultrapura		--
DNA	1ng/μL	3ng

Essa reação foi levada ao equipamento *QuantStudio™ 3 Real-Time PCR System* (*Thermo Fisher Scientific*) para realização de PCR competitiva em tempo real, que consiste em um ciclo de 10 minutos a 95°C, seguido de 5 ciclos de 15 segundos a 92°C e 1 minuto a 58°C, finalizado por 40 ciclos de 15 segundos a 92°C e 1 minuto a 60°C (TABELA 12).

Tabela 12. Protocolo de ciclagem dos experimentos de PCR competitiva alelo específica.

Temperatura	95°C	92°C	58°C	92°C	60°C
Tempo	00:10:00	00:00:15	00:01:00	00:00:15	00:01:00
Ciclos	1	5		40	

As reações de *castPCR* foram analisados no software *QuantStudio™ Design and Analysis*, com verificação de ruído automático e valores de limiar fixos em 0,2, tanto para as reações de detecção de mutação, quanto para as de gene de referência. Os resultados expressos em valores de ciclo de amplificação (Ct, do inglês, *cycle threshold*) foram exportados para o software *Mutation Detector™*. A quantificação das moléculas mutadas foi realizada através da diferença de Ct (ΔCt) entre as reações

para uma mesma amostra de detecção do alelo selvagem e detecção do alelo mutante, ou entre as reações de detecção do alelo de referência e detecção do alelo mutante. Para as análises, os valores de ciclo máximo aceitável (*cutoff*) foram ajustados a 40 ciclos de PCR, e o ΔCt entre a detecção do alelo raro e do alelo mais comum ou referência $< 9,96$, valor predeterminado pela *Life Technologies*. Dessa forma, amostras com mutação detectável são aquelas cujos valores de Ct de ambos os alelos analisados foram inferiores a 40 ciclos, e a diferença entre esses valores foi inferior a 9,96. Além disso, o software *Mutation Detector*TM considera amostras positivas apenas aquelas cujo valor de Ct do alelo de referência esteve detectável no intervalo de 18 a 28 ciclos de PCR. Visto a concentração reduzida de DNA adicionada às reações, optou-se por analisar cada caso de forma individual. Os resultados foram expressos em porcentagem de mutação (%) detectada. Para amostras cuja mutação não foi detectada, por ausência de amplificação do alelo mutante, no qual o software não é capaz de calcular o valor de ΔCt , ou $\Delta Ct > 9,96$, o valor expresso nos resultados é de 0%.

5.5. ANÁLISE ESTATÍSTICA

A normalidade na distribuição dos valores de Ct para detectar diferentes mutações somáticas, por PCR em tempo real na validação analítica, ou castPCR na validação clínica, foi avaliada pelo teste de Kolgomorov-Smirnov. Da mesma forma, a normalidade na distribuição das frequências de mutação encontrada nas três técnicas foi testada. As diferenças na distribuição dos resultados por técnica empregada foram avaliadas por testes paramétricos (t de Student ou Anova) ou não paramétricos (Mann-Whitney ou Kruskal-Wallis), a depender do resultado.

Para avaliar a eficiência dos sistemas de detecção elaborados foi utilizado teste de regressão linear com os valores de Ct obtidos pelo experimento de linearidade. Visto que os dados obtidos para as análises de correlações entre resultados obtidos por dPCR, castPCR e sequenciamento de Sanger foram de presença ou ausência de mutação, utilizamos o teste de correlação de Pearson. Todas estas análises foram realizadas no programa Graphpad Prism v 5.01.

A fim de comparar a eficiência na detecção qualitativa (presença/ausência de mutação), os números absolutos de mutações detectadas por cada técnica foram comparados entre si pelo teste exato de Fisher, utilizando o software SISA (<http://www.quantitativeskills.com/sisa/>). Para comparar a eficiência na detecção

quantitativa, as frequências cumulativas de mutações detectadas por dPCR e castPCR foram comparadas entre si através do teste de Mann-Whitney, utilizando o programa Graphpad Prism v 5.01.

6. RESULTADOS

6.1. VALIDAÇÃO ANALÍTICA

Para avaliar a eficiência de detecção, um experimento de linearidade foi realizado com os oligonucleotídeos sintéticos contendo as variantes comuns e raras. Diluições progressivas foram feitas de forma seriada, a partir de uma concentração aproximada de 100.000.000 ($\text{Log}_{10}=8$), até 0,01 ($\text{Log}_{10}=-2$) moléculas de DNA/mL. Os oligonucleotídeos diluídos foram submetidos à amplificação na plataforma digital de PCR em tempo real (*LightCycler480®Real Time PCR System*). As sondas complementares às variantes raras p.L861Q e p.T790M (ambas de de *EGFR*) foram capazes de detectar até uma única molécula de DNA em um mL de solução, após cerca de 40 ciclos de amplificação dos fragmentos. Já as sondas para *EGFR* p.S768I, e *KRAS* p.G12A, p.G12R e p.G13D detectaram até 10 moléculas contendo as possíveis variantes/mL, sendo que o mix de sondas para identificação das variantes no códon 719 de *EGFR* e a sonda para p.G12V de *KRAS* foram eficientes para detectar, no mínimo, 100 moléculas por mL, todas próximo ao ciclo 40. As demais sondas foram menos eficientes, detectando 1000 ou mais moléculas, perto do ciclo 40. Como esperado, o número de moléculas detectadas, com ou sem a mutação, foi inversamente proporcional ao número de ciclos de amplificação necessários para detectá-las, sendo as correlações muito altas, com r^2 maior que 0,92 (FIGURA 10).

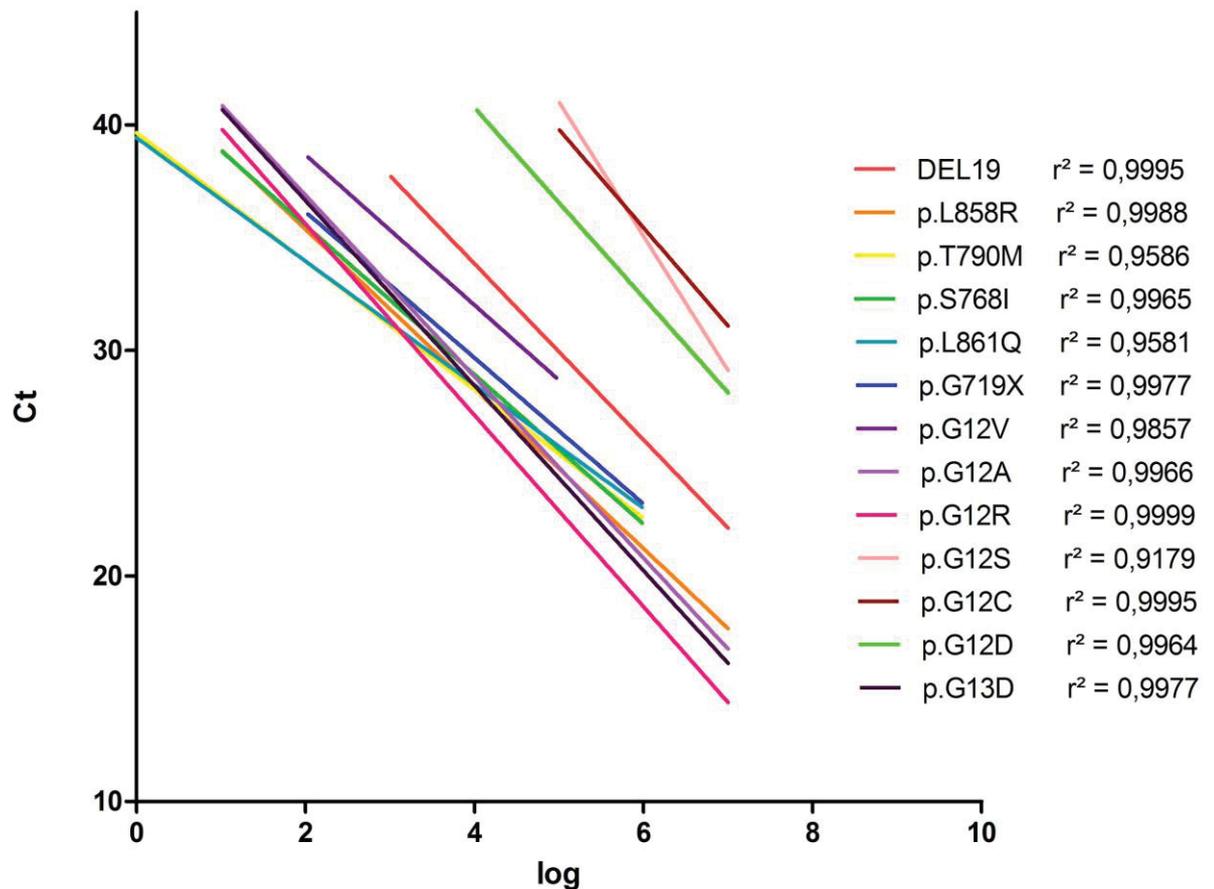


Figura 10. Eficiência das sondas para detecção das variantes.

Legenda: Ct - limiar de ciclo; log - LOG10 da concentração de DNA detectada.

Nota: Embora a diluição tenha sido feita até 0,01 moléculas de DNA/mL, não houve detecção abaixo de \log_{10}^0 , e por essa razão o eixo x não inclui estes valores. O objetivo foi detectar a eficiência dos sistemas. Cada ensaio de detecção é formado por um par de oligonucleotídeos e um par de sondas, um pra detecção do alelo selvagem e um mutado. Foram utilizados os controles positivos sintetizados para os alelos com as sequências selvagens e para as mutações de deleção no exon 19 e pontuais nos exons 18, 20 e 21 do gene *EGFR* e exon 2 do gene *KRAS*. Como o sistema é *duplex*, usamos dois controles positivos, em reações separadas, para avaliar o sistema. No entanto, unimos os valores como replicatas, visto que ambos os resultados avaliam o mesmo sistema de detecção.

Diluições progressivas dos controles sintéticos com as variantes comuns de *EGFR* e *KRAS* foram repetidamente testados para verificar a correspondência com o número de moléculas detectadas. O número de moléculas detectadas a partir da diluição de 100 moléculas de DNA/mL, para ambos os genes, foi significativamente superior ao controle negativo, ao contrário de concentrações entre 0,01 e 10 moléculas/mL, que também não diferiram entre si ($p > 0,05$). Em concentrações de DNA mais elevadas, houve menos dispersão na distribuição dos valores. Em outras

palavras, houve excelente correspondência entre o grau de diluição e o número de moléculas detectadas (log2 representa 100 moléculas por mL, o que de fato foi detectado com todas as sondas utilizadas) (FIGURA 11). Logo, o limite de detecção (LoD) do método de dPCR para todas as sondas, com exceção de DEL19 e p.S768I ficou definido como 100 moléculas de DNA/mL (FIGURA 12).

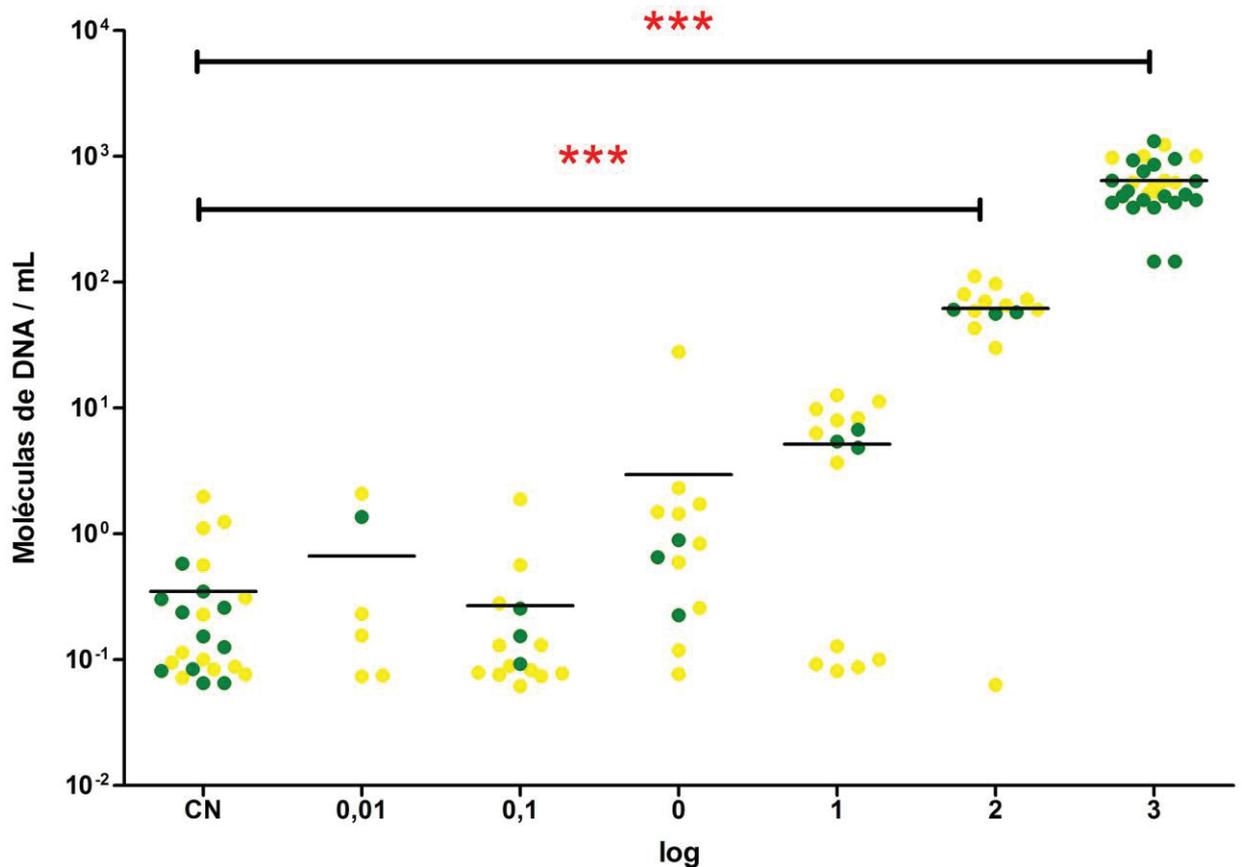
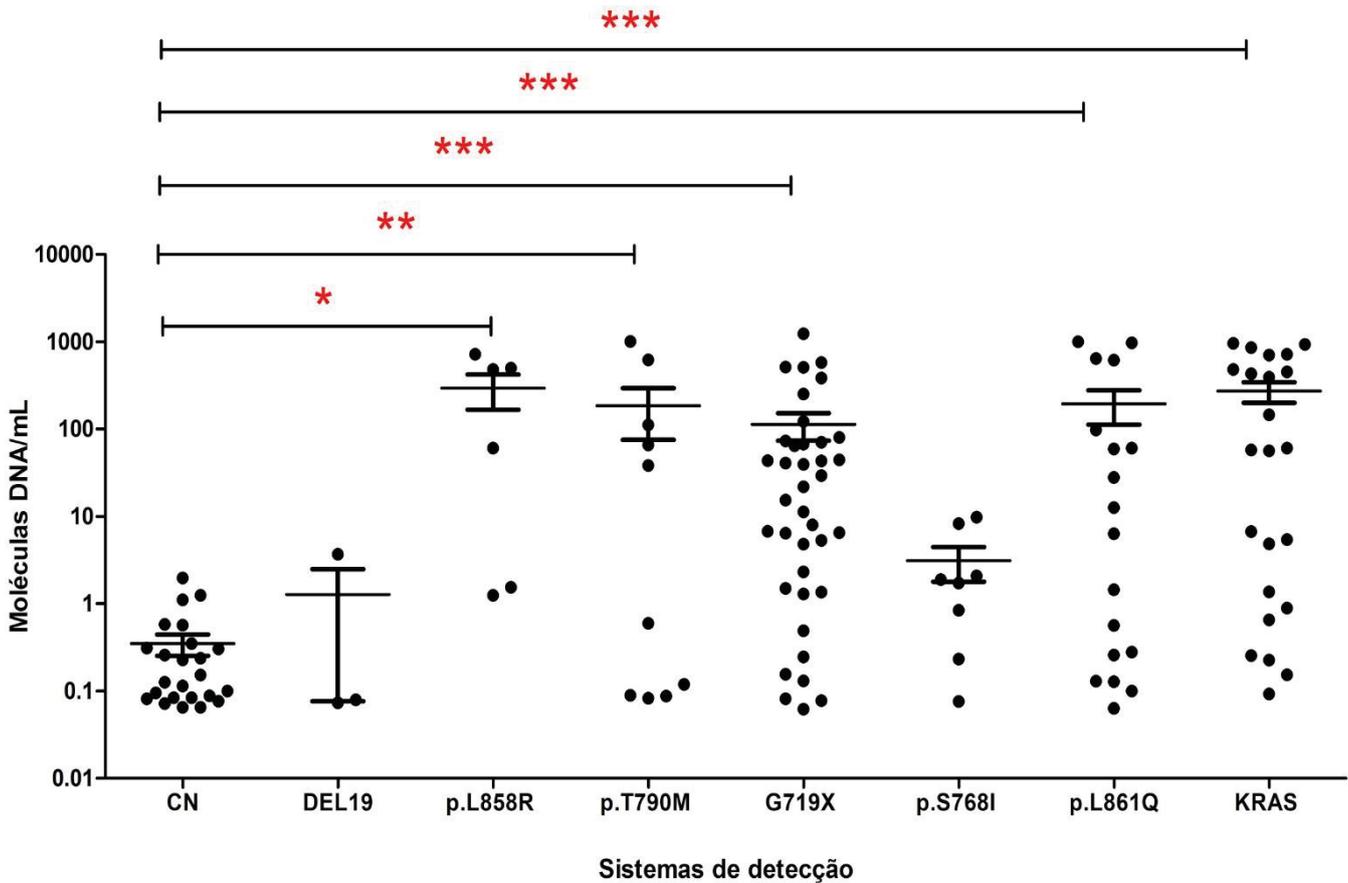


Figura 11. Distribuição das concentrações detectadas pela dPCR de controles sintéticos contendo as variantes comuns de *EGFR* e *KRAS*, de acordo com os seus níveis de diluição.

Legenda: pontos em amarelo representam variantes comuns de *EGFR* e pontos em verde representam variantes comuns de *KRAS*; (***) a partir de 100 moléculas de DNA/mL (log 2 e log3) houve diferença significativa em relação ao controle negativo (Kruskal Wallis $p < 0,0001$); CN= controle negativo (sem DNA).

Nota: visto que a técnica de PCR digital possui alta sensibilidade, é observada detecção de fluorescência residual no controle negativo, mesmo sem a adição de DNA.



Foram realizadas misturas de DNA que continham os alelos selvagem e mutado nas proporções de 1%, 5%, 10% e 50% de mutação, a fim de avaliar o limite de quantificação (LoQ) do método de dPCR. O LoQ das variantes raras foi testado apenas para as sondas específicas para as mutações p.L861Q e p.T790M, visto que estas obtiveram melhor desempenho no teste de linearidade. O número de moléculas detectadas a partir da diluição de 1% de mutação, para ambas as variantes, foi

superior ao controle negativo. A amplificação foi eficiente em todas as concentrações de DNA ($r^2 = 0,9636$ para p.T790M e $r^2 = 0,9981$ para p.L861Q), além de pouca dispersão na distribuição dos valores e precisão nos resultados, de acordo com o software *QuantStudio™ 3D Analysis Suite*. Logo, o limite de quantificação (LoQ) do método de dPCR para os sistemas de detecção das variantes p.L861Q e p.T790M ficou definido como 1 molécula de DNA contendo o alelo mutante em 100 moléculas de DNA contendo o alelo selvagem (Figura 13).

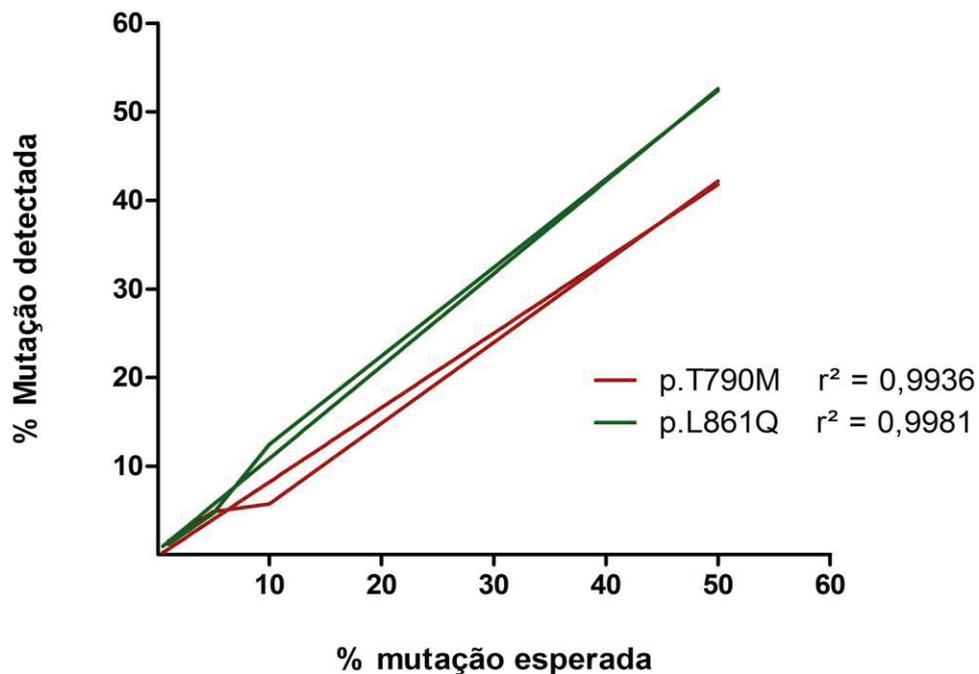


Figura 13. Regressão linear para análise de limite de quantificação (LoQ) do método de dPCR para os sistemas de detecção das variantes p.T790M e p.L861Q de EGFR.

Legenda: % mutação esperada: dada pelas misturas de DNA que foram realizadas de forma a conter os alelos selvagem e mutado nas proporções de 1%, 5%, 10% e 50% de mutação.

Nota: a amplificação foi eficiente em todas as concentrações de DNA; p.T790M ($r^2 = 0,9636$) e p.L861Q ($r^2 = 0,9981$); as demais sondas não foram testadas.

6.2. VALIDAÇÃO DIAGNÓSTICA

O protocolo de dPCR foi comparado aos de *castPCR* e Sequenciamento de Sanger para fins de validação analítica, em amostras de DNA extraído dos blocos de parafina de câncer de pulmão e câncer colorretal. A concentração final do DNA extraído das amostras foi ajustada para 1 ng/ μ L.

6.2.1. Análises qualitativas

6.2.1.1. Câncer de pulmão e *EGFR*

Para identificação de mutações do gene *EGFR* nas amostras de câncer de pulmão, foram utilizadas as técnicas de dPCR, *castPCR* e Sequenciamento de Sanger. Com exceção dos sistemas de detecção para as variantes p.L858R, p.T790M e p.S768I, os demais não foram capazes de identificar as variantes raras, apesar de terem detectado a variante comum em um Ct máximo de 40 (TABELA 13).

Tabela 13. Valores de ciclo de amplificação (Ct) para detecção de mutações em *EGFR*, por *castPCR*.

Sistema de detecção	Variante comum				Variante rara			
	Mínimo	Mediana	Máximo	Média ± DP	Mínimo	Mediana	Máximo	Média ± DP
DEL19	26,81	33,30	39,50	33,44 ± 4,50	--	--	--	--
p.L861Q	25,05	30,69	39,26	30,63 ± 3,76	--	--	--	--
p.L858R	20,87	25,25	33,03	25,71 ± 2,81	28,36	28,75	29,13	28,75 ± 0,54
p.G179A	25,19	31,62	38,38	31,77 ± 3,46	--	--	--	--
p.T790M	26,78	34,75	40,00	35,00 ± 3,37	37,31	39,59	40,00	39,19 ± 1,23
p.S768I	19,68	23,74	33,07	24,09 ± 2,72	38,16	38,16	38,16	38,16
p.G719S	28,47	35,04	40,00	35,20 ± 3,44	--	--	--	--
p.G719C	24,23	31,53	40,00	31,83 ± 3,63	--	--	--	--
REF	23,26	33,30	39,50	33,44 ± 4,5	--	--	--	--
NTC	--	--	--	--	--	--	--	--

Nota: Os resultados de *castPCR* foram expressos em valores de Ct, tanto para os ensaios para variante mais comum, quanto para os ensaios de mutação e referência. As análises qualitativas foram realizadas no software *QuantStudio™ Design and Analysis* pela presença ou ausência de curva de amplificação, com verificação de ruído automático e valores de *threshold* fixos em 0,2. Curvas de amplificação com valores de Ct acima de 40 foram desconsideradas.

Legenda: REF = ensaio referência; NTC = *no template control*, controle negativo. Curvas de amplificação com valores de Ct acima de 40, foram desconsideradas.

Das 57 amostras de câncer de pulmão avaliadas, 33 (57,9%) não apresentaram nenhuma mutação no gene *EGFR*, por dPCR. Das outras 24 amostras com mutação detectada por dPCR (42,1%), 18 apresentaram apenas uma mutação (18/24 = 75%) e 6 apresentaram mutações complexas (6/24 = 25%). Por castPCR, apenas 11 amostras apresentaram alguma das mutações avaliadas no projeto (19,3%), e por sequenciamento de Sanger, apenas 4 amostras apresentaram-se mutadas (7,0%) (TABELA 14). Portanto, a dPCR foi capaz de detectar um número maior de amostras mutadas, em comparação com a castPCR (Fisher $p=0,014$) e o sequenciamento de Sanger (Fisher $p=0,00002$) (TABELA 14).

Tabela 14. Número de amostras de câncer de pulmão com mutações somáticas no gene *EGFR*, detectadas pelas técnicas de dPCR, *castPCR* e sequenciamento de Sanger.

	Mutações totais	Mutações Simples	Mutações Complexas	Valor p*
dPCR	24	18	6	--
castPCR	11	8	3	0,014
Seq	4	4	0	0,00002

Nota: os números são valores absolutos de amostras que apresentaram alguma mutação no gene *EGFR* detectável. (*) o valor estatístico é em relação ao número total de mutações detectadas por cada uma das técnicas.

Legenda: Seq= Sequenciamento

Utilizando a técnica de dPCR, foram observadas 10 amostras com a mutação DEL19 (32,2% dos casos com mutação) (FIGURA 14), enquanto que por castPCR e sequenciamento, foram detectadas apenas 4 amostras com essa deleção (Fisher $p=0,049$). Destas, 6 amostras possuíam exclusivamente a mutação DEL19 (6/57 = 10,5% das amostras), enquanto as outras 4 eram portadoras de várias mutações (“mutação complexa”) (4/57 = 7,0% das amostras) (TABELA 15).

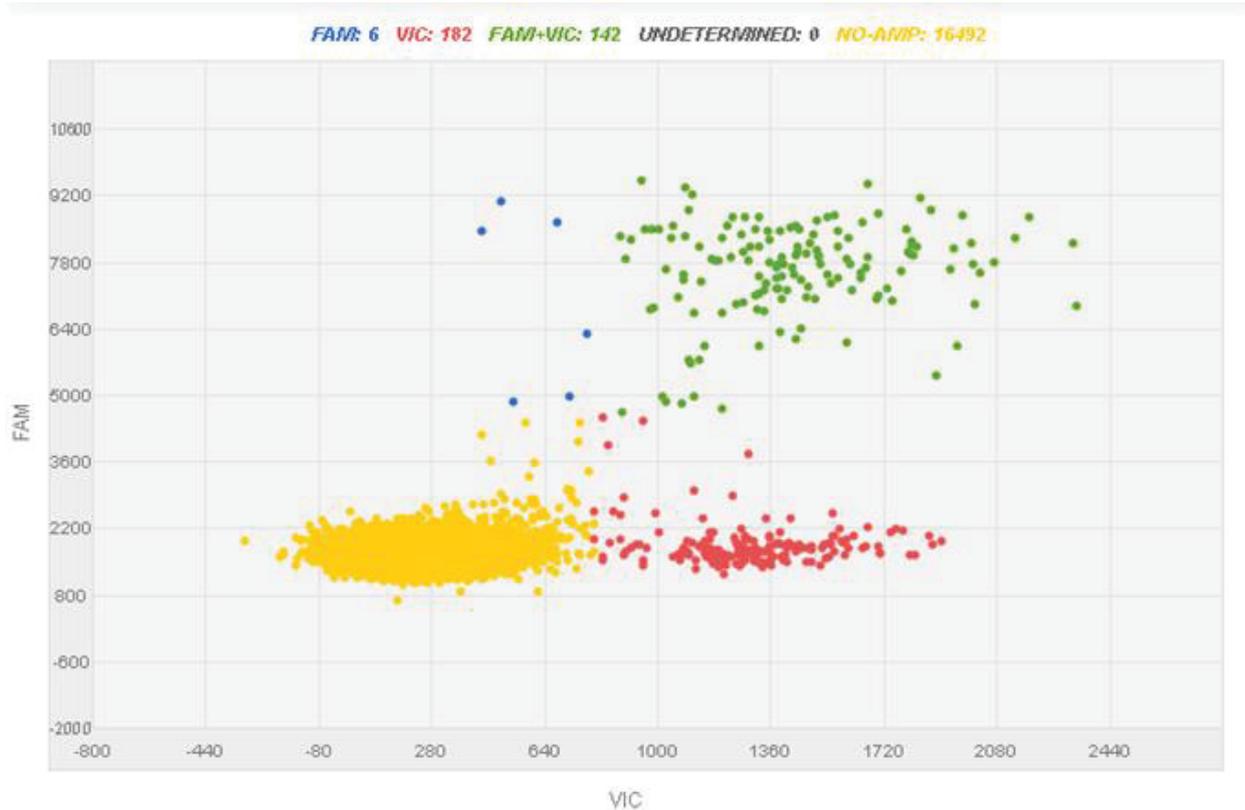


Figura 14. Imagem representativa de amostra de câncer de pulmão avaliada pelo software *QuantStudio™ 3D Analysis Suite* portadora de mutação DEL19.

Legenda: amostra de câncer de pulmão cuja mutação DEL19 foi detectada por dPCR. Cada ponto no gráfico representa uma partição no chip de dPCR. O eixo Y corresponde a fluorescência de VIC®, enquanto o eixo X corresponde a fluorescência de FAM™. Pontos amarelos são partições sem DNA amplificado (n=16492). Pontos vermelhos são partições cuja sonda detectada pelo software foi VIC® (n=182), característica de indivíduos portadores da deleção no exon 19. Pontos em azul são partições cuja sonda detectada pelo software foi FAM™ (n=6). Pontos em verde são partições nas quais ambas as sondas (FAM™ e VIC®) foram detectadas (n=142). Acima do gráfico estão representados os números de partições identificadas com cada fluoróforo.

Nota: a sonda marcada com o fluoróforo VIC® foi desenhada em uma região conservada do gene *EGFR*. A sonda marcada com o fluoróforo FAM™ foi desenhada em uma região onde ocorrem as deleções mais frequentes. A sonda FAM™ não é capaz de se acoplar nas moléculas de DNA do indivíduo que possui a deleção no exon 19. A emissão de fluorescência observada é exclusivamente da sonda marcada com VIC® nesses indivíduos, garantindo a eficiência da reação. Em indivíduos que não possuem alelos mutantes, ambas as sondas são capazes de se acoplar, de forma que há emissão de fluorescência de FAM™ e VIC® pela mesma molécula, portanto, na mesma partição. A emissão de fluorescência exclusiva de FAM™ pode indicar uma reação ineficiente nas partições específicas (nesse caso, 6), visto que a sonda VIC®, que garante a eficiência da reação, não foi capaz de se acoplar.

A mutação pontual p.L858R no éxon 21 foi a segunda mais frequente, nas amostras de câncer de pulmão (25,8% dos casos com mutação). Das 8 amostras com

a mutação p.L858R detectada por dPCR, 6 possuíam exclusivamente essa mutação (6/57 = 10,5% das amostras), enquanto 2 eram portadoras de mutação complexa (2/57 = 3,5% das amostras). Para a detecção desta mutação pontual, a dPCR foi mais eficaz na detecção quanto comparada ao sequenciamento de Sanger (Fisher p=0,04), e foi observada uma tendência ao melhor desempenho da técnica de dPCR quanto ao número de amostras detectadas com a mutação, em comparação à castPCR (Fisher p=0,09) (TABELA 15).

Para a mutação pontual p.L861Q no éxon 21, e para as mutações de ponto nos éxons 18 e 20, não foram observadas diferenças significativas entre a PCR digital e a castPCR. Além disso, o sequenciamento de Sanger não foi eficaz em detectar as mutações de ponto de *EGFR* no éxon 18, bem como as do éxon 20 - p.T790M e p.S768I, e éxon 21 - p.L858R e p.L861Q (TABELA 17).

Tabela 15. Mutações somáticas detectadas no gene *EGFR* nas amostras de câncer de pulmão pelas diferentes técnicas utilizadas.

Mutação	dPCR			castPCR			Valor p	Sequenciamento			Valor p
	A	B	Total	A	B	Total		A	B	Total	
DEL19	6	4	10	2	2	4	<u>0,04</u>	4	-	4	<u>0,04</u>
p.L858R	6	2	8	2	0	2	0,09	-	-	-	<u>0,04</u>
p.L861Q	-	1	1	-	-	-	ns	-	-	-	ns
p.G719X	-	4	4	-	2	2	ns	-	-	-	<u>0,04</u>
p.T790M	6	2	8	3	2	5	ns	-	-	-	<u>0,002</u>
p.S768I	-	1	1	1	-	1	ns	-	-	-	ns
Total			32			14		4		4	

Legenda: sublinhado - valor estatisticamente significativo, comparação entre os valores acumulados (totais) de mutações somáticas detectadas pelas técnicas de dPCR, castPCR e sequenciamento de Sanger. ns = não significativo. (A) presença de apenas uma mutação; (B) presença de múltiplas mutações em uma mesma amostra (mutação complexa).

Nota: os valores de p foram calculados com base no teste de Fisher, comparando os números absolutos de detecção pelas diferentes técnicas. Algumas das amostras apresentaram mais de uma mutação, e por essa razão, o valor total desta tabela é referente ao número de mutações detectadas, e não ao número de amostras mutadas. O sequenciamento de Sanger para *EGFR* foi realizado duas vezes para a confirmação dos dados.

Visto que nenhuma amostra obteve mais de uma mutação por sequenciamento, dPCR foi mais eficaz em detectar mutações complexas no gene *EGFR* em uma mesma amostra, quando comparada a essa técnica (Fisher p=0,004). No entanto, quando comparada a castPCR, seu desempenho foi equivalente para a detecção de duas ou mais mutações, em uma mesma amostra (Fisher p=0,17). Já para aquelas

amostras de câncer de pulmão com apenas uma mutação no gene *EGFR*, houve diferença significativa na detecção destas, quando comparada com *castPCR* (Fisher $p=0,028$) (FIGURA 15) e com sequenciamento de Sanger (Fisher $p=0,0003$).

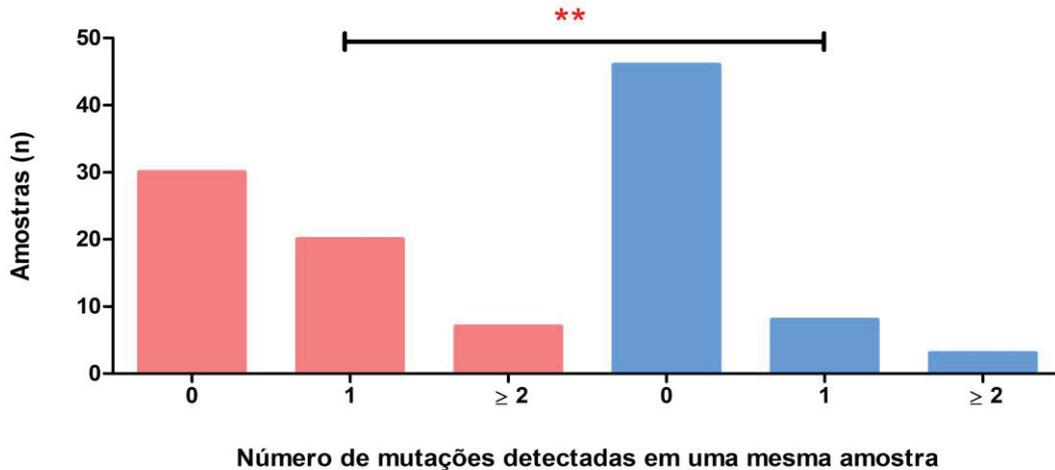


Figura 15. Número de mutações somáticas detectadas no gene *EGFR* nas amostras de câncer de pulmão por dPCR e por *castPCR*.

Legenda: em rosa, mutações detectadas por dPCR; em azul, mutações detectadas por *castPCR*; (**) foi observada diferença significativa na detecção de mutações em *EGFR* entre as técnicas de dPCR e *castPCR* naquelas amostras que possuem apenas uma mutação (mutação simples) (Fisher, $p = 0,004$). Para a detecção de duas ou mais mutações em uma mesma amostra o desempenho das técnicas foi equivalente ($p > 0,05$).

Através da avaliação individual das amostras de câncer de pulmão de acordo com (i) presença ou ausência de mutações no gene *EGFR*, e também quanto (ii) ao número de mutações detectadas em cada amostra, foi observado que as técnicas de dPCR e *castPCR* estão correlacionadas. No entanto, tal correlação é mediana (Pearson $r^2=0,53$, $p<0,0001$), visto que a dPCR detectou mais amostras mutadas (figura 13). Uma correlação ainda mais fraca foi observada em relação a dPCR e ao sequenciamento (Pearson $r^2=0,30$, $p=0,019$) (ANEXO 3).

6.2.1.2. Câncer colorretal e *KRAS*

Para identificação de mutações pontuais no gene *KRAS* nas amostras de câncer colorretal, foram utilizadas as técnicas de dPCR, *castPCR* e Sequenciamento de Sanger. A eficácia de detecção da mutação p.G13D foi comparada apenas entre a dPCR e o sequenciamento de Sanger (TABELA 17).

Tabela 16. Valores de ciclo de amplificação (Ct) obtidos através da detecção de mutações no gene *KRAS* pela técnica de *castPCR*.

Sistema de detecção	Variante comum				Variante rara			
	Mínimo	Mediana	Máximo	Média ± DP	Mínimo	Mediana	Máximo	Média ± DP
G12C	23,21	31,66	37,86	31,37 ± 3,26	--	--	--	--
G12S	23,47	34,16	39,50	33,24 ± 3,78	--	--	--	--
G12R	22,93	31,46	40,00	31,79 ± 3,79	--	--	--	--
G12V	24,01	33,34	39,59	33,04 ± 3,70	29,52	36,94	39,23	34,97 ± 4,14
G12D	23,67	30,97	39,41	31,22 ± 3,44	32,00	36,62	38,90	36,14 ± 2,44
G12A	23,47	30,19	39,52	30,66 ± 3,26	--	--	--	--
REF	22,12	29,95	39,60	29,54 ± 28,56				
NTC	--	--	--	--				

Nota: Os resultados de *castPCR* foram expressos em valores de Ct, tanto para os ensaios para variante mais comum, quanto para os ensaios de mutação e referência. As análises qualitativas foram realizadas no software *QuantStudio™ Design and Analysis* pela presença ou ausência de curva de amplificação, com verificação de ruído automático e valores de *threshold* fixos em 0,2. Curvas de amplificação com valores de Ct acima de 40 foram desconsideradas.

Legenda: REF = ensaio referência; NTC = *no template control*, controle negativo.

Das 41 amostras de câncer colorretal avaliadas por dPCR, 31,7% não apresentaram nenhuma mutação de ponto no gene *KRAS*. Das amostras mutadas, 14 (34,1%) apresentaram a mutação de ponto p.G13D no éxon 2. Das outras 20 amostras mutadas por dPCR, 13 apresentaram apenas uma mutação (13/20 = 65%) e 7 apresentaram mutações complexas (7/20 = 35%). Por *castPCR*, 12 amostras apresentaram alguma das mutações avaliadas no projeto (29,2%), e por sequenciamento de Sanger, 13 amostras apresentaram-se mutadas (31,7%) (TABELA 19). Portanto, a dPCR foi capaz de detectar um número maior de amostras mutadas em comparação ao sequenciamento de Sanger ($p=0,002$), mas a diferença

não foi estatisticamente significativa, em comparação com a *castPCR* ($p=0,11$) (TABELA 17).

Tabela 17. Número de amostras de câncer colorretal com mutações somáticas no gene *KRAS*, detectadas pelas técnicas de dPCR, *castPCR* e sequenciamento de Sanger.

	Mutações totais	Mutações Simples	Mutações Complexas	Valor p*
dPCR	20	13	7	--
<i>castPCR</i>	12	11	1	ns
dPCR	28	17	11	--
Sequenciamento	13	13	0	0,04

Legenda: os números são valores absolutos de amostras que apresentaram alguma mutação no gene *KRAS* detectável; (*) o valor estatístico foi mensurado em comparação a técnica de dPCR; valores absolutos de amostras que apresentaram alguma mutação no gene *KRAS* detectada incluindo a mutação de ponto p.G13D foram utilizados para comparação com a técnica de sequenciamento, e por essa razão, estes valores são superiores aos usados para comparação com a técnica de *castPCR*; ns = não significativo.

Para amostras que possuíam apenas uma mutação no gene *KRAS*, houve diferença significativa na detecção destas, mas apenas por sequenciamento de Sanger ($p=0,0003$), e não por *castPCR* ($p=0,45$). No entanto, a dPCR foi mais eficaz em detectar mutações complexas no gene *KRAS* em uma mesma amostra quando comparada a *castPCR* ($p=0,023$) (FIGURA 16). Visto que nenhuma amostra obteve mais de uma mutação por sequenciamento, dPCR também foi mais eficaz em detectar mutações complexas no gene *KRAS* em uma mesma amostra quando comparada a essa técnica ($p=0,000003$).

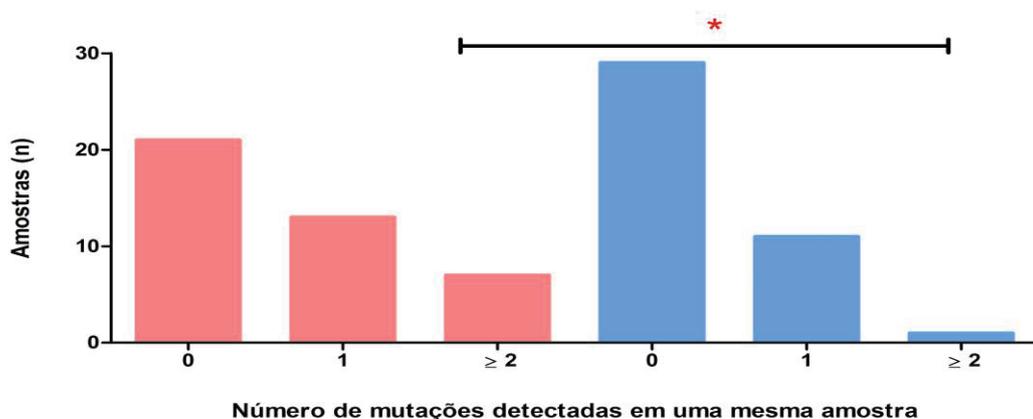


Figura 16. Número de mutações somáticas detectadas no gene *KRAS* nas amostras de câncer colorretal, por dPCR e por *castPCR*.

Legenda: em rosa, mutações detectadas por dPCR; em azul, mutações detectadas por castPCR; (*) houve diferença significativa na detecção de duas ou mais mutações por amostra, entre as técnicas de dPCR e *castPCR* (Fisher, $p = 0,023$).

A mutação p.G13D no gene *KRAS* foi detectada com maior frequência nas amostras de câncer colorretal pela técnica de dPCR (34,1%). Quando comparada ao sequenciamento de Sanger, a dPCR foi mais eficaz em detectar a mutação p.G13D (Fisher $p=0,037$), visto que essa mutação foi detectada em apenas 7,3% das amostras pelo sequenciamento. A detecção de p.G13D não foi avaliada por *castPCR*. A PCR digital também apresentou uma tendência a ser mais eficaz em relação a *castPCR* e sequenciamento em detectar as mutações p.G12R (Fisher $p=0,05$) e p.G12A (Fisher $p=0,05$) no gene *KRAS* nas amostras de câncer colorretal (TABELA 18).

As mutações p.G12V e p.G12D no gene *KRAS* foram detectadas em maior frequência nas amostras de câncer colorretal, por meio das três técnicas. A mutação p.G12V foi observada em 11 amostras utilizando dPCR (26,8%), 7 amostras utilizando *castPCR* (17,1%), e 5 amostras utilizando o sequenciamento de Sanger (12,2%). A mutação p.G12D foi observada em 11 amostras, utilizando dPCR (26,8%), 6 amostras utilizando *castPCR* (14,6%), e 7 amostras utilizando o sequenciamento de Sanger (9,7%) (TABELA 18). Não foram observadas diferenças entre as técnicas para a detecção das mutações p.G12V (Fisher $p=0,42$ para *castPCR* e Fisher $p=0,26$ para sequenciamento) e p.G12D (Fisher $p=0,57$ para *castPCR* e Fisher $p=0,55$ para sequenciamento) no gene *KRAS*.

Todas as outras mutações foram detectadas preferencialmente ou unicamente por dPCR, sendo que p.G13D não foi avaliada por *castPCR*, devido a restrições orçamentárias (TABELA 18).

Tabela 18. Mutações somáticas detectadas no gene *KRAS* nas amostras de câncer colorretal pelas diferentes técnicas utilizadas.

Mutação	dPCR	castPCR	Valor p	Sequenciamento	Valor p
p.G12S	1	-	ns	-	ns
p.G12C	1	-	ns	-	ns
p.G12R	4	-	<u>0,05</u>	-	<u>0,05</u>
p.G12D	11	6	ns	4	ns
p.G12V	11	7	ns	5	ns
p.G12A	4	-	<u>0,02</u>	1	<u>0,05</u>
p.G13D	14	NT	-	3	<u>0,03</u>
Total	46	13		13	

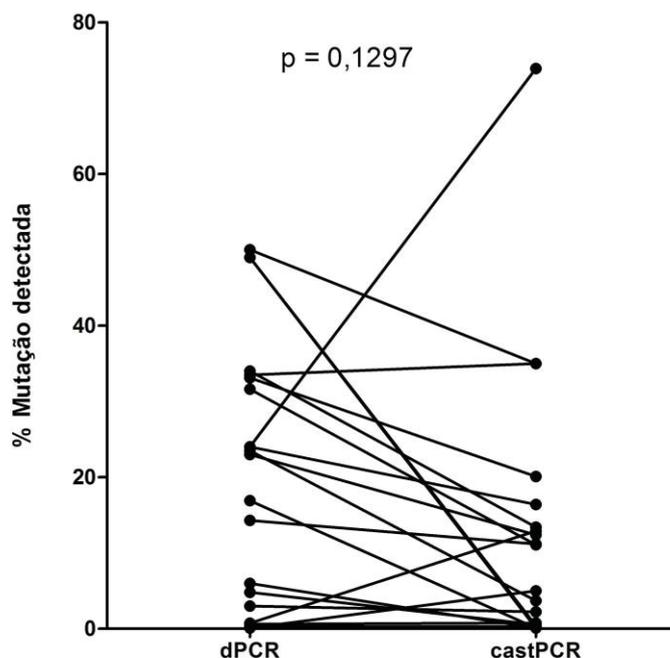
Nota: algumas das amostras (38%) apresentaram mais de uma mutação, e por essa razão, o valor total desta tabela é referente ao número de mutações detectadas, e não ao número de amostras mutadas.

Legenda: os valores de p foram calculados com base no teste de Fisher, comparando os números absolutos de detecção pelas diferentes técnicas. Os valores estatisticamente significativos estão sublinhados. ns = não significativo. NT = não tipado.

Através da avaliação individual das amostras de câncer de pulmão de acordo com (i) presença ou ausência de mutações no gene *KRAS*, e também quanto (ii) ao número de mutações detectadas em cada amostra, foi observado que as técnicas de dPCR e castPCR estão correlacionadas. No entanto, tal correlação é mediana (Pearson $r^2=0,55$, $p=0,0002$), visto que a dPCR detectou mais amostras mutadas. O mesmo foi observado em relação a dPCR e ao sequenciamento (Pearson $r^2=0,56$, $p=0,0001$) (ANEXO 4).

6.2.2. Análises Quantitativas

Ao observar exclusivamente as amostras de câncer de pulmão que apresentaram alguma mutação detectável no gene *EGFR* e de câncer colorretal no gene *KRAS*, não houve diferença na frequência de mutações detectadas, tanto por dPCR, quanto por castPCR (Mann-Whitney $p=0,1297$). Apenas duas amostras apresentaram grandes diferenças de detecção entre ambas as técnicas, o que possivelmente se deve a falsos positivos (FIGURA 17). O Sequenciamento de Sanger não foi avaliado por esse parâmetro, por ser uma técnica qualitativa.



Nota: no eixo y, a frequência de mutação observada por amostra, é dada pela frequência de mutação detectadas pelas técnicas (Pearson $r^2=0,3505$, $p=0,1297$). Três amostras apresentaram resultados muito diferentes, entre si.

Observando exclusivamente as amostras de câncer de pulmão e câncer colorretal com ausência de mutação por castPCR (selvagens), mas que possuíam mutações detectadas nos genes *EGFR* e *KRAS* através da dPCR, foi visto que grande parte dessas amostras (32 em 41, 78,05%) são aquelas cuja ocorrência de mutação é inferior a 20% (média 9,59 e mediana 4,16). Apenas 4 amostras (9,76%) não apresentaram mutação, de acordo com o método de detecção por dPCR (FIGURA 18). Considerando que este resultado é verdadeiro, apenas 4 pacientes ficariam com prognóstico inconclusivo, relativo à resposta à terapia convencional.

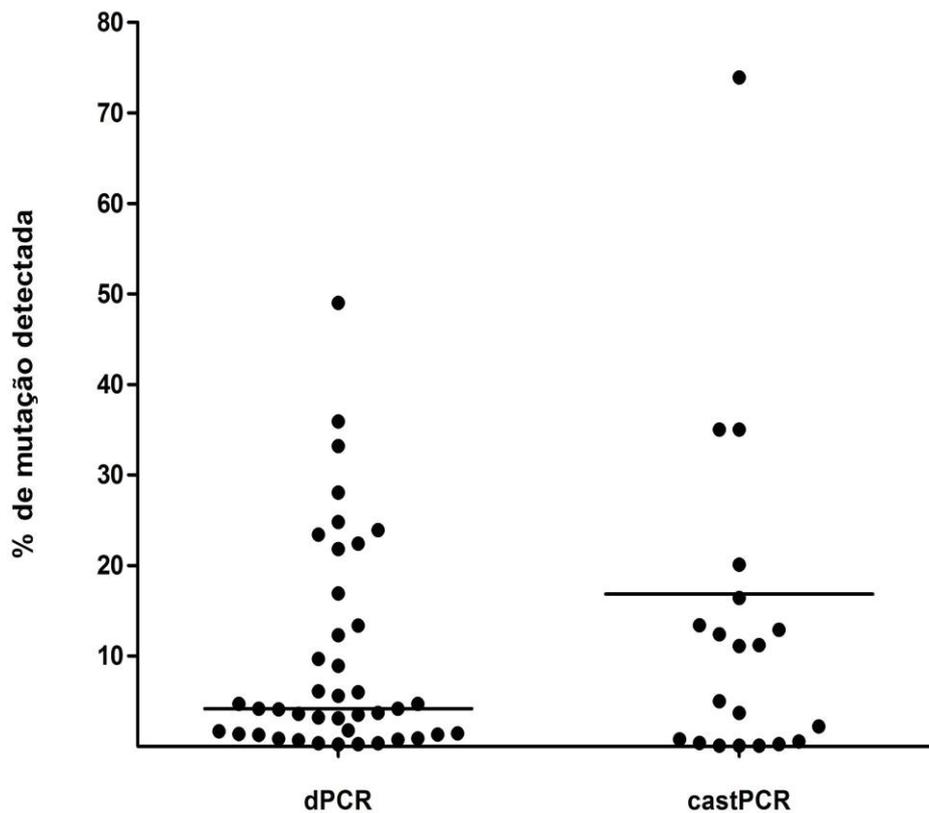


Figura 18. Quantificação de mutações somáticas nos genes *EGFR* e *KRAS* das amostras de câncer de pulmão e câncer colorretal, cuja mutação foi detectada apenas por dPCR, em comparação àquelas detectadas por castPCR.

Legenda: Cada ponto no gráfico representa uma amostra avaliada por dPCR ou castPCR. As amostras de câncer de pulmão e câncer colorretal representadas em dPCR são aquelas cuja técnica de castPCR não detectou mutações, enquanto a dPCR detectou; média 9,59 e mediana 4,16. As amostras de

câncer de pulmão e câncer colorretal representadas em castPCR são aquelas cuja técnica de castPCR detectou mutações, independente do resultado de dPCR; média 16,84 e mediana 11,10.

Nota: no eixo y, a frequência de mutação observada por amostra, é dada pela frequência de cada mutação, detectadas por dPCR. Amostras identificadas por ambas as técnicas não estão representadas para dPCR. Os casos não concordantes entre as técnicas de dPCR e sequenciamento não foram avaliados, em função dessa última não quantificar as moléculas mutadas detectadas.

As amostras contendo as variantes raras avaliadas estão distribuídas na figura para *EGFR* e *KRAS*. É possível observar que a partir da frequência de 1% de mutação, tanto para *EGFR* quanto para *KRAS*, houve diferença significativa com relação ao controle negativo, e que frequências inferiores a 1% não diferem do mesmo ($p > 0,05$), corroborando nosso resultado anterior de LoQ. Vimos também que, em concentrações mais elevadas de DNA com a sequência mutada em meio a DNA selvagem, a distribuição dos valores permaneceu mais homogênea e eficiente ($r^2 = 0,9906$) (FIGURA 19). Tais dados reforçam nossos resultados anteriores de limite de quantificação da técnica de PCR digital.

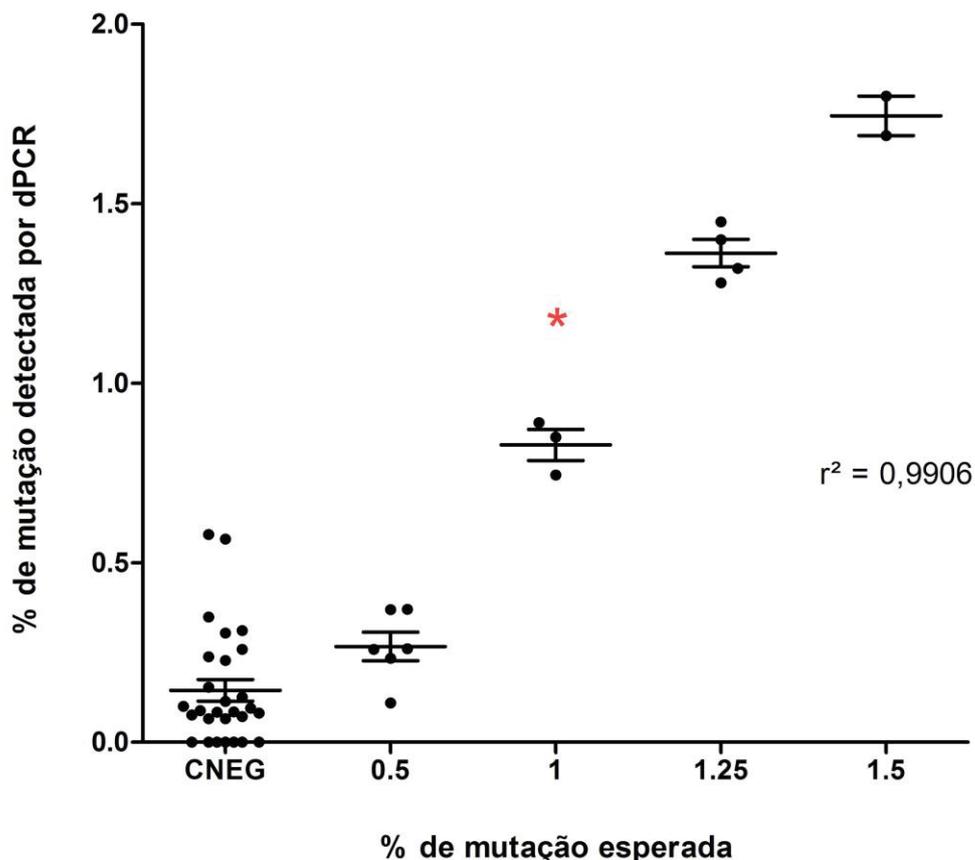


Figura 19. Análise de limite de quantificação (LoQ) das amostras contendo as variantes raras avaliadas para *EGFR* e *KRAS* por dPCR.

Legenda: (*) a partir da frequência de 1% de mutação detectada por dPCR, tanto em *EGFR* quanto em *KRAS*, houve diferença significativa do controle negativo (Kruskal-Wallis $p < 0.0001$). A dPCR foi capaz de detectar baixas frequências de mutação (a partir de 1%) de forma eficiente (regressão linear, $r^2 = 0,9906$). CNEG= controle negativo

6.3. CONTROLE DE QUALIDADE E ACREDITAÇÃO INTERNACIONAL

A fim de avaliar a especificidade da técnica de dPCR, foram utilizadas 10 amostras de DNA extraído de tecido pulmonar parafinado livre de neoplasia e 7 amostras de DNA extraído de biópsia de cólon livre de neoplasia. Não foram detectadas mutações em nenhuma dessas amostras, permitindo um controle de qualidade interno da técnica de dPCR (FIGURA 20).

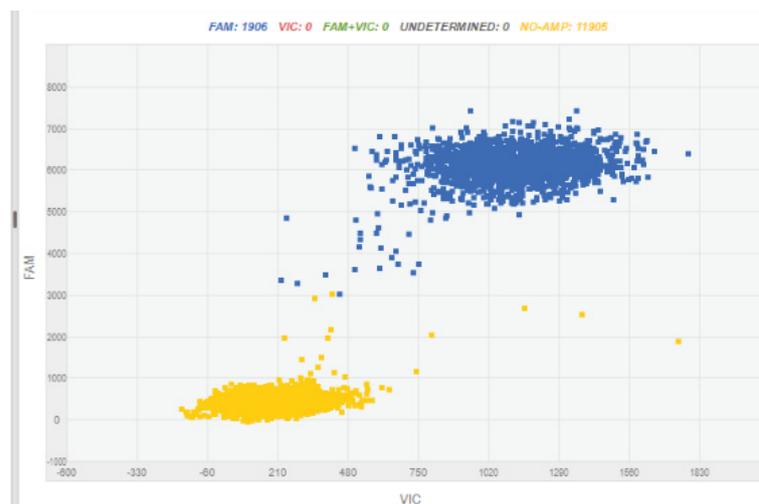


Figura 20. Resultado representativo de uma amostra de tecido livre de neoplasia por dPCR

Nota: Pontos em azul correspondem às moléculas de DNA marcadas com o fluoróforo FAM, indicando presença de DNA com a variante comum. Pontos em amarelo indicam partições do chip sem moléculas de DNA.

Além disso, foi possível avaliar a sensibilidade e acurácia através da realização do teste de acreditação do Colégio Americano de Patologia (CAP), no qual detectou-se de forma adequada 17 das 18 mutações no gene *EGFR* (94,44% de acurácia) (TABELA 19) e 18 das 21 mutações no gene *KRAS* (85,71% de acurácia) (TABELA 20). Os resultados foram enviados ao CAP para avaliação externa de qualidade e aguardam aprovação.

Tabela 19. Resultados da tipagem dos controles para mutações no gene *EGFR*, para acreditação internacional de acordo com o Colégio Americano de Patologia (CAP).

AMOSTRA	MUTAÇÃO	RESULTADO ENCONTRADO	RESULTADO ESPERADO
01	DEL19	-	-
	p.L858R	-	-
	p.T790M	+	-
	p.G719X	-	-
	p.S769I	-	-
02	p.L861Q	-	-
	DEL19	-	-
	p.L858R	-	-
	p.T790M	-	-
	p.G719X	+	+
03	p.S769I	-	-
	p.L861Q	-	-
	DEL19	+	+
	p.L858R	-	-
	p.T790M	-	-
	p.G719X	-	-
	p.S769I	-	-
	p.L861Q	-	-

Legenda: amostras 01, 02 e 03 analisadas para presença ou ausência das mutações no gene *EGFR* estudadas; (-) mutação não detectável; (+) mutação detectável; em **negrito**, resultado não concordante com o esperado pelo CAP.

Tabela 20. Resultados da tipagem dos controles para mutações no gene *KRAS*, para acreditação internacional de acordo com o Colégio Americano de Patologia (CAP).

AMOSTRA	MUTAÇÃO	RESULTADO ENCONTRADO	RESULTADO ESPERADO
01	p.G12S	+	-
	p.G12R	+	+
	p.G12C	+	-
	p.G12D	-	-
	p.G12A	-	-
	p.G12V	-	-
02	p.G13D	-	-
	p.G12S	-	-
	p.G12R	-	-
	p.G12C	+	+
	p.G12D	-	-
	p.G12A	-	-
03	p.G12V	-	-
	p.G13D	-	-
	p.G12S	-	-
	p.G12R	-	-
	p.G12C	-	-
	p.G12D	-	-
	p.G12A	-	-
	p.G12V	-	+
	p.G13D	-	-

Legenda: amostras 01, 02 e 03 analisadas para presença ou ausência das mutações no gene *KRAS* estudadas; (-) mutação não detectável; (+) mutação detectável; em **negrito**, resultado não concordante com o esperado pelo CAP.

6.4. CUSTO EFETIVO DA REALIZAÇÃO DAS TÉCNICAS

Visto que o objetivo inicial do trabalho é validar a técnica de dPCR para uso comercial, foi realizada uma estimativa do custo efetivo na utilização das técnicas de dPCR, castPCR e sequenciamento de Sanger, para realização de 100 reações. A técnica de dPCR mostrou-se mais econômica na realização em série para detecção de mutações (TABELA 21). Em relação ao tempo médio gasto na bancada para 10 reações, necessita-se de cinco a seis horas para realizar a dPCR, desde o preparo da reação até a leitura dos chips e análise de dados, visto que só a reação em termociclador demora cerca de três horas. Esse tempo é muito menor que o previsto para o sequenciamento de Sanger, que pode ser de 12 até 24 horas. Contudo, a dPCR exige maior tempo de bancada, que aquele empregado na castPCR (três a quatro horas).

Tabela 21. Custo calculado na utilização das técnicas de dPCR, castPCR e sequenciamento de Sanger para detecção de mutações somáticas nos genes *EGFR* e *KRAS*.

	dPCR /mutação	castPCR /ensaio (mutação)	Sequenciamento de Sanger /gene (mutação)
Ensaio de detecção	--	2120,89	--
<i>Mastermix</i>	64,35	88,88	58,50
<i>Primers</i>	10,00	--	10,00
Sondas	300,00	--	--
<i>Taq</i>	--	--	545,40
Chips para dPCR	350,00	--	--
Placas de PCR	--	399,00	399,00
Insumos	36,21	25,99	101,25
	760,56	2.634,76	1.114,21

Legenda: os valores são dados aproximados, calculados em reais, para realização de 10 reações para detecção de uma mutação somática validada, independente da técnica utilizada. Informações como coleta de material, processamento e extração de DNA não estão inseridos no valor final.

7. DISCUSSÃO

O método padrão-ouro para detecção de mutações em neoplasias é o sequenciamento de DNA (DIDELLOT et al., 2012). No entanto, além de ser uma técnica trabalhosa e de custo elevado, requer ao menos 10 a 20% das células contendo a variante de interesse, dentre a população celular estudada (FRANCIS; STEIN, 2015). O câncer de pulmão e o câncer colorretal, em geral, são neoplasias detectadas em estágios avançados (HON et al., 2017; KIM et al., 2013; KNIGHT et al., 2017). Logo, a detecção precoce de um pequeno número de células com mutações alvo é essencial para o melhor prognóstico da doença. Diante deste cenário, o objetivo deste projeto foi propor uma técnica alternativa, capaz de detectar variantes raras com sensibilidade, acurácia e precisão, a um custo reduzido, em amostras FFPE, que frequentemente estão disponíveis na rotina clínica. Para tanto, realizamos a validação da técnica de dPCR para detecção de mutações somáticas nos genes *EGFR* e *KRAS* em amostras FFPE de pacientes com NSCLC e CCR, respectivamente.

Os sistemas de detecção de mutações somáticas através da técnica de dPCR foram elaborados por meio da projeção e síntese de oligonucleotídeos iniciadores direto e reverso e sondas de hidrólise, testados quanto a sua eficiência. Este é um diferencial do nosso trabalho, visto que habitualmente, são escolhidos ensaios comerciais TaqMan® para os sistemas de detecção por dPCR (LAIG, 2016; WANG, 2010; FENG 2018). Além disso, os sistemas reportados na literatura, em geral não contemplam as mutações menos frequentes p.S768I, p.L861Q e p.G719X (TATEMATSU, 2017; DUFORT, 2011).

O limite de detecção (LoD) do método desenvolvido no presente estudo foi definido como 100 moléculas/mL. Foi possível detectar frequências de 1% de DNA contendo as mutações p.L861Q e p.T790M, em uma população de DNA selvagem, valor definido como o limite de quantificação (LoQ) do método de dPCR neste trabalho. Teoricamente, a dPCR é capaz de detectar cerca de 3 moléculas de DNA portadoras de determinada mutação, em 3333 moléculas com a variante comum (LoQ de 0,1%). Contudo, os protocolos de dPCR recomendam o uso de 20 ng de DNA, concentração cerca de 6,5 vezes superior à que usamos neste trabalho.

Para a validação clínica da técnica de dPCR, comparamos nossos resultados com aqueles obtidos por castPCR e sequenciamento de Sanger. A detecção de mutações somáticas foi feita em DNA extraído de amostras FFPE de pacientes com

câncer de pulmão e câncer colorretal. Embora os protocolos usuais recomendem cerca de 20 ng de DNA para amplificação na dPCR (MASAGO et al. 2018), utilizamos apenas 3 ng, devido ao baixo rendimento das extrações de DNA. Dessa forma, outro diferencial do nosso trabalho foi avaliar a performance clínica das técnicas em questão, na detecção precisa de mutações somáticas em situações com concentração limitada de DNA, sendo este, até onde temos conhecimento, o primeiro trabalho a avaliar esse cenário.

Devido à baixa qualidade do material, falsos negativos podem ter ocorrido pela presença de inibidores de reação, presentes nas amostras por serem oriundas de um material parafinado antigo. Em um trabalho recente, no qual mutações foram identificadas por NGS de genoma total (WGS, do inglês *whole genome sequencing*), em DNA extraído de amostras de tecido fresco congeladas, de FFPE e de sangue de 52 pacientes de câncer, a correlação entre o número de mutações identificadas no DNA oriundo de tecido fresco e no de FFPE foi relativamente baixa (0,44). Temperaturas mais baixas para reverter ligações cruzadas no processo de extração de DNA de amostras FFPE, contudo, aumentaram a acurácia na identificação de número de cópias, e o método foi considerado adequado para utilização na rotina clínica (REF: ROBB et al. 2018).

Métodos baseados em PCR em tempo real parecem ser promissores para detecção de mutações somáticas, além de mais sensíveis que o sequenciamento (BOLTON et al., 2015). Estudos mostram que a tecnologia de dPCR apresenta vantagens em relação ao sequenciamento genético, sendo capaz de detectar variantes raras com mais precisão (HUGGETT; WHALE, 2013; PEKIN et al., 2010), além de menos suscetível a inibição e com menor custo (HUGGETT et al., 2013). Nosso trabalho demonstra resultados similares, visto que a dPCR foi capaz de detectar um número maior de amostras com mutações somáticas nos genes *EGFR* e *KRAS*, em comparação com o sequenciamento de Sanger.

Além disso, Bartels e colaboradores (2017), em sua análise molecular em pacientes com câncer de pulmão, observaram que a qPCR apresenta sensibilidade inferior à dPCR, e que a quantidade de DNA necessária para a realização do protocolo de qPCR deve ser superior à de dPCR (BARTELS et al., 2017). Nossos resultados fortalecem as conclusões de Bartels, visto que a dPCR também foi mais eficaz na detecção de mutações somáticas no gene *EGFR* em comparação com a castPCR, sendo que a concentração de DNA para a realização dos protocolos, foi a mesma

(1ng/ μ L). No entanto, acreditamos que a concentração do DNA tenha sido limitante no protocolo de castPCR. Ao observarmos os dados de castPCR, vimos que algumas amostras possuem valores de ciclo de amplificação (Ct) superiores a 30, inclusive para identificação de variantes comuns. Uma amostra é classificada como portadora de uma variante rara através da técnica de castPCR quando, dentre outros parâmetros, a diferença entre os valores de ciclo de amplificação entre a detecção das variantes rara e comum (Δ Ct) é inferior a 9,96. Dessa forma, algumas amostras ditas como portadoras de alelos exclusivamente selvagens, na realidade apenas não tiveram o alelo mutante identificado, sendo esta uma possível limitação do nosso trabalho.

Em pacientes com adenocarcinoma de pulmão, 93% das mutações somáticas no gene *EGFR* ocorrem entre os exons 18 - 21, região que codifica para o domínio tirosina cinase do receptor. Logo, tais mutações podem resultar na ativação independente de ligante, o que justifica a importância da sua detecção precoce (O'Kane, 2017). Nesse projeto, a dPCR permitiu identificar uma frequência aproximadamente duas vezes superior de amostras portando mutações no gene *EGFR*, comparado a castPCR, e seis vezes superior, comparada ao sequenciamento de Sanger. A deleção no exon 19 (DEL19) e a variante p.L858R no exon 21 são as mutações mais frequentes encontradas nos pacientes de câncer de pulmão (LADANYI; PAO, 2008; MITSUDOMI; YATABE, 2010). De fato, estas totalizaram mais de 50% das mutações encontradas no gene *EGFR* por dPCR, sendo dez amostras com DEL19 e oito amostras com p.L858R, dentre todas as 24 que continham mutações. Em contraste, pela castPCR e o sequenciamento de Sanger, foram detectadas apenas 4 amostras com a DEL19, evidenciando a maior sensibilidade da dPCR em detectar esta deleção. Para p.L858R, duas amostras foram detectadas por castPCR, indicando apenas uma tendência ao melhor desempenho da dPCR, enquanto que o sequenciamento de Sanger não foi capaz de detectar nenhuma amostra com tal variante.

O sequenciamento de Sanger não foi eficaz em detectar a variante p.T790M no gene *EGFR*, enquanto que por dPCR e por castPCR foram detectadas 8 e 5 amostras portadoras dessa variante, respectivamente. Além disso, também foi ineficaz para detectar as variantes raras menos frequentes no gene *EGFR* (p.L861Q no exon 21, p.S768I e mutações de ponto no exon 18). A vista disso, a concentração reduzida de DNA pode ter sido um fator limitante. Além disso, a utilização de ensaios altamente

sensíveis para a detecção da mutação p.T790M já foi reportada por apresentar resultados falsos positivos, principalmente em amostras com reduzida quantidade de DNA, dificultando a distinção em relação ao verdadeiro positivo (BORSU, 2017). Todavia, visto que estas variantes são menos frequentes nas amostras de pacientes com câncer de pulmão (ARRIETA, 2015), nossa hipótese é que isso se deva ao pequeno número de amostras. Somado a isso, não houve diferença entre a detecção dessas variantes por castPCR e PCR digital. Esses resultados reforçam nossa hipótese de que a técnica de dPCR é capaz de detectar um número maior de mutações no gene *EGFR*, em comparação com o sequenciamento de Sanger ($p=0,00002$), e também quando comparada à PCR competitiva de forma ampla ($p=0,014$). Tal situação já foi vista em outros estudos, que avaliaram a eficácia da técnica de dPCR em comparação com a técnica de sequenciamento (HENNEBIQUE et al., 2017) e com a PCR tempo real (ALIKIAN et al., 2017; CAMPOMENOSI et al., 2016).

O perfil molecular de *KRAS* possui papel clinicamente relevante, uma vez que mutações ativadoras no gene *KRAS* são associadas à resistência ao tratamento com inibidores de EGFR, enquanto que a presença de alelo selvagem prevê uma resposta positiva ao tratamento (Harlé, 2016). Das 41 amostras de câncer colorretal avaliadas por dPCR neste projeto, 68,3% apresentaram alguma mutação de ponto no gene *KRAS*, enquanto que por castPCR, 29,2% das amostras apresentaram mutações ($p=0,11$), e por sequenciamento de Sanger, apenas 31,7% das amostras apresentaram alguma variante rara ($p=0,002$).

Utilizando dPCR para detecção de mutações no gene *KRAS*, cerca de um terço dos casos apresentou a variante p.G13D no exon 2, o que está de acordo com o esperado (Seekhantod, 2016; MATIKAS, 2017). Por sequenciamento de Sanger, foram detectadas apenas três amostras portadoras dessa variante, evidenciando a maior sensibilidade da técnica de dPCR em detectar tal mutação no exon 2 do gene *KRAS* ($p=0,03$), o que também já foi reportado por outros pesquisadores (SHO et al., 2017). Tal como Seekhantod e colaboradores (2016) em seu estudo de validação técnica, as mutações p.G12V e p.G12D no gene *KRAS* foram detectadas em maior frequência nas amostras de câncer colorretal pelas três técnicas aqui avaliadas. A variante p.G12V foi observada em 26,8% das amostras utilizando a técnica de dPCR, 17,1% utilizando castPCR e 12,2%, utilizando sequenciamento de Sanger. A variante p.G12D foi observada em 26,8% das amostras utilizando a técnica de dPCR, 14,6%

utilizando castPCR e 9,7%, utilizando sequenciamento de Sanger. Não foram observadas diferenças entre as técnicas para a detecção das mutações p.G12V ($p=0,42$ para castPCR e $p=0,26$ para sequenciamento) e p.G12D ($p=0,57$ para castPCR e $p=0,55$ para sequenciamento) no gene *KRAS*. No entanto, para as mutações p.G12R e p.G12A, a PCR digital apresentou uma tendência a ser mais eficaz, em relação a castPCR e sequenciamento ($p=0,05$ para ambas) em detectar as mutações nas amostras de câncer colorretal. Da mesma forma que para *EGFR*, nossa hipótese é que isso se deva à baixa frequência dessas variantes, nas amostras de pacientes com câncer colorretal.

Visto que nenhuma amostra obteve mais de uma mutação por sequenciamento, podemos dizer que a dPCR foi mais eficaz em detectar mutações complexas tanto no gene *EGFR*, quanto no gene *KRAS* em uma mesma amostra, quando comparada a essa técnica. Tal dado poderia sugerir uma detecção equivocada da técnica de PCR digital. No entanto, quando comparada à PCR competitiva em tempo real, o desempenho da dPCR foi equivalente para a detecção de duas ou mais mutações somáticas no gene *EGFR* em uma mesma amostra. Além disso, com a utilização de amostras de tecido pulmonar parafinado livre de neoplasia, a técnica de PCR digital mostrou-se específica na detecção de variantes no *EGFR*, e como o esperado, não foram detectadas mutações em nenhuma dessas amostras. Somado a isso, o teste de acreditação internacional do Colégio de Patologia Americana (CAP) apresentou 94,44% de acurácia para *EGFR*. Acreditamos que esse cenário seja justificado, devido ao número reduzido de amostras com mutações somáticas complexas, esperado em menos de 15% dos indivíduos (O'KANE et al., 2017), raramente observadas no sequenciamento.

Os ensaios realizados para detecção de variantes em *KRAS*, de fato, demandaram uma validação mais trabalhosa, dado que as variantes analisadas são próximas (algumas inclusive no mesmo nucleotídeo), o que facilitaria ligações inespecíficas das sondas. Inclusive, detecção de duas ou mais mutações somáticas no gene *KRAS* em uma mesma amostra foi mais comum utilizando-se a dPCR, do que a castPCR, o que poderia confirmar uma detecção equivocada das variantes de *KRAS* pela técnica de PCR digital. O teste de acreditação internacional do Colégio de Patologia Americana (CAP) apresentou 85,7% de acurácia para *KRAS*, valor inferior ao estabelecido para detecção de mutações somáticas em *EGFR*. Embora não tenham sido detectadas mutações em nenhuma das biópsias de cólon livres de

neoplasia, evidenciando a especificidade da técnica de PCR digital na detecção de variantes no gene *KRAS* (uma vez que a castPCR é capaz de amplificar o alvo de interesse somado à inibição de outro alelo), os dados levam-nos a sugerir maior confiabilidade nos resultados de castPCR, para essa situação específica. Uma opção mais viável para os ensaios de detecção de mutações somáticas no exon 2 do gene *KRAS* por dPCR poderia ser uma mistura de reação, unindo todas as variantes. Dessa forma, a detecção de mutações somáticas não permitiria discriminação alélica entre as variantes raras, no entanto, permitiria diferenciação entre alelo selvagem e alelo mutante. Esta é uma alternativa econômica e suficiente para o manejo terapêutico dos pacientes, visto que seriam realizadas apenas duas reações por amostra.

A fim de avaliar se as técnicas de PCR digital, PCR competitiva em tempo real e Sequenciamento de Sanger apresentavam resultados concordantes na detecção de mutações, foi realizado o teste de correlação entre as técnicas. Avaliamos cada amostra individualmente, tanto para presença ou ausência de mutações somáticas nos genes *EGFR* e *KRAS*, quanto para número de mutações detectadas em uma mesma amostra. Foi observada uma correlação fraca entre as técnicas, visto que a PCR digital detectou um maior número de amostras portadoras de mutações somáticas. Ao observarmos apenas as amostras com mutação somática detectável, não foi observada diferença entre a porcentagem de mutação nos genes *EGFR* e *KRAS* das amostras detectadas por PCR digital ou PCR competitiva em tempo real. No entanto, notamos que mais de 70% dos casos não concordantes, ou seja, das amostras que são consideradas selvagens pelas técnicas de sequenciamento e castPCR e que apresentaram mutações pela PCR digital, apresentaram porcentagem de mutação por dPCR, inferior a 20%. Um contexto plausível, visto que a maioria das amostras com mutação detectável possui alta frequência de mutação, o que permite a equivalência entre as técnicas. No entanto, a correlação é fraca, dado que a técnica de PCR digital promete ser mais sensível que as outras técnicas na detecção de pequenas quantidades de moléculas mutadas (SHO et al., 2017; ZHU et al., 2017). Foram observadas apenas três amostras discordantes com mais de 30% de mutação, que podem ser falsos positivos detectados por dPCR.

Foram detectadas variantes raras com baixa frequência nas amostras avaliadas para *EGFR* e *KRAS*, sendo que a partir da frequência de 1% de mutação, tanto para *EGFR* quanto para *KRAS*, houve diferença significativa com relação ao controle negativo (frequências inferiores a 1% não diferem do mesmo). Esses dados

reforçam nossos resultados anteriores de limite de quantificação da técnica de PCR digital. Novamente, a baixa qualidade de DNA oriundo de amostras FFPE frequentemente é responsável pela amplificação de apenas 200 moléculas de DNA, aumentando consideravelmente o LoQ (MASAGO et al. 2018). Comparações entre quantificações obtidas de material experimental ou fresco, vs. obtido de FFPE, exemplificam este problema. Em um trabalho anterior de Milbury et al. (2014), foi possível detectar uma molécula de DNA portando a variante rara em 180.000 moléculas sem a mutação, em 3,3 ug de DNA de amostras experimentais, utilizando-se ensaios com oligonucleotídeos comercialmente disponíveis para *EGFR* p.T790M e p.L858R (MILBURY et al. 2014). Em outro trabalho com DNA obtido de 25 amostras FFPE, identificou-se até 5% de mutação p.T790M para *EGFR*, usando kit comercial (MASAGO et al. 2018).

Com a validação da técnica de PCR para a pesquisa de variantes raras, foi possível a implementação dessa nova ferramenta clínica no Laboratório Genoprimer, através da elaboração de um procedimento operacional padrão (apêndice 1). A figura 20 representa o fluxo de trabalho a ser realizado na ocasião do recebimento das amostras de câncer de pulmão ou câncer colorretal em FFPE. Dessa forma, pretende-se auxiliar no diagnóstico e tomada de decisão terapêutica para o tratamento dos pacientes, além de possibilitar a validação futura de testes para detecção de mutações comumente associadas a outros tipos de câncer, em tecido parafinado.

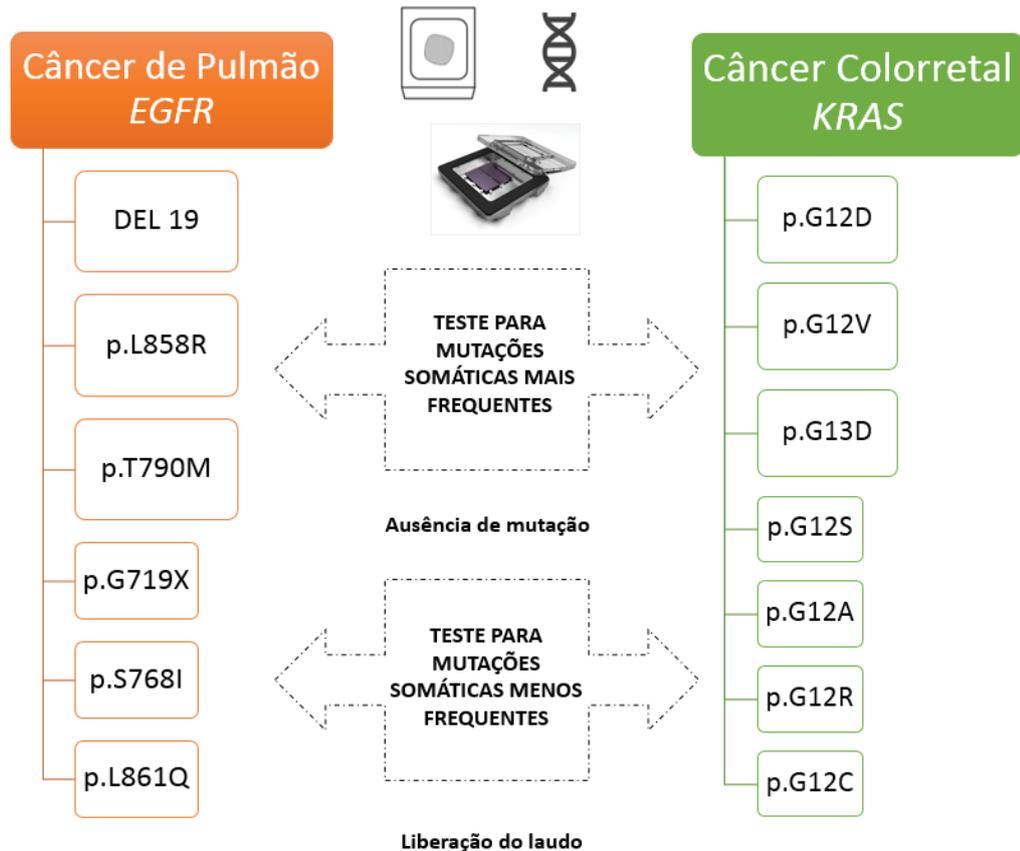


Figura 20. Fluxograma do fluxo de trabalho na ocasião do recebimento de amostras de câncer de pulmão ou câncer colorretal em FFPE.

Legenda: as amostras são recebidas em material parafinado e passam pelo procedimento de extração de DNA anteriormente ao protocolo de PCR digital. A primeira etapa de testes é realizada buscando a identificação de mutações somáticas de maior frequência ou de resistência ao tratamento, como é o caso de p.T790M no gene *EGFR*. Caso estas não sejam identificadas, prossegue-se para a busca de variantes raras.

Ainda que as técnicas baseadas em amplificação de DNA por PCR limitem-se a um painel de variantes genéticas pré-estabelecidas, o conhecimento acerca dos marcadores oncogênicos vem proporcionando avanço na abordagem terapêutica oncológica, que progressivamente deixa de ser baseada na localização anatômica do tumor, para ser fundamentada nas características genéticas das subpopulações tumorais. Isso permite que a terapia seja altamente personalizada e tenha as maiores chances possíveis de sucesso, acarretando para o paciente menor morbimortalidade e maior tempo de vida livre, sem progressão. E diante disso, a proposta do diagnóstico por PCR digital como um teste genômico padrão para busca de mutações somáticas em neoplasias não se limita apenas ao uso de amostras em bloco parafinados.

Embora o uso da biópsia tecidual configure o atual padrão para a pesquisa molecular em câncer (DONALDSON e PARK, 2018), uma das principais limitações na rotina diagnóstica se dá em casos nos quais é necessário a re-biópsia (MASAGO et al., 2018). Além disso, amostras teciduais são suficientes para um número limitado de testes moleculares (DONALDSON e PARK, 2018), e nem todos os pacientes podem ser submetidos a tal procedimento. Ainda que o tecido adequado possa ser obtido de aproximadamente 67% dos pacientes, eventos adversos como pneumotórax e hemorragia intrapulmonar já foram reportados em pacientes submetidos a uma segunda biópsia (MASAGO et al., 2018). Nesse cenário, a biópsia líquida é uma ferramenta valiosa no manejo de tumores de difícil acesso, sendo este um procedimento pouco invasivo. Por essa abordagem, é possível pesquisar células tumorais e DNA livre circulante do tumor e, conseqüentemente, detectar mutações, metástases e resistência ao tratamento, utilizando sangue periférico do paciente. Desde 2016, o uso da biópsia líquida é aprovado pela FDA (*Food and Drug Administration*) para identificação de mutações em *EGFR* em pacientes com câncer de pulmão de células não pequenas, em situações nas quais a biópsia tecidual não é possível de ser realizada (DONALDSON e PARK, 2018). O monitoramento seriado do perfil genético dos pacientes através da coleta de sangue periférico e com a utilização de uma técnica sensível como a PCR digital, poderia permitir um melhor gerenciamento do tratamento destes indivíduos, aumentando o tempo de sobrevivência. Para tanto, espera-se que o presente projeto permita embasar estudos futuros com biópsias líquidas e detecção de mutações em DNA circulante.

REFERÊNCIAS

- AHRBERG, C. D.; MANZ, A.; CHUNG, B. G. Polymerase Chain Reaction in Microfluidic Devices. **Lab on a Chip**, v. 11, n. 41, p. 1–21, 2016.
- AMADO, R. G. et al. Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. **Journal of Clinical Oncology**, v. 26, n. 10, p. 1626–1634, 2008.
- APPLIED BIOSYSTEMS. Essentials of Real Time PCR. **ReVision**, p. 1–8, 2005.
- ARRIETA, O. et al. The impact of common and rare EGFR mutations in response to EGFR tyrosine kinase inhibitors and platinum-based chemotherapy in patients with non-small cell lung cancer. **Lung Cancer**, v. 87, n. 2, p. 169–175, 2015.
- BAE, J. M.; KIM, J. H.; KANG, G. H. Molecular Subtypes of Colorectal Cancer and Their Clinicopathologic Features, With an Emphasis on the Serrated Neoplasia Pathway. **Archives of pathology & laboratory medicine**, v. 140, n. 5, p. 406–12, 2016.
- BAEK, J. H. et al. Efficacy of EGFR tyrosine kinase inhibitors in patients with EGFR-mutated non-small cell lung cancer except both exon 19 deletion and exon 21 L858R: A retrospective analysis in Korea. **Lung Cancer**, v. 87, n. 2, p. 148–154, 2015.
- BARBANO, R. et al. Competitive allele-specific TaqMan PCR (Cast-PCR) is a sensitive, specific and fast method for BRAF V600 mutation detection in Melanoma patients. **Scientific reports**, v. 5, n. November, p. 18592, 2015.
- BEHJATI, S.; TARPEY, P. S. What is next generation sequencing? **Archives of Disease in Childhood. Education and Practice Edition**, v. 98, n. 6, p. 236–238, 2013.
- BOLTON, L. et al. KRAS mutation analysis by PCR: A comparison of two methods. **PLoS ONE**, v. 10, n. 1, p. 1–13, 2015.
- BUSTIN, S.; HUGGETT, J. qPCR primer design revisited. **Biomolecular Detection and Quantification**, v. 14, n. April, p. 19–28, 2017.
- CAROTENUTO, P. et al. Detection of KRAS mutations in colorectal carcinoma patients with an integrated PCR/sequencing and real-time PCR approach. **Pharmacogenomics**, v. 11, n. 8, p. 1169–79, 2010.
- CHONG, C. R.; JÄNNE, P. A. The quest to overcome resistance to EGFR-targeted therapies in cancer. **Nature medicine**, v. 19, n. 11, p. 1389–400, 2013.
- D'ANGELO, S. P. et al. Distinct clinical course of EGFR-mutant resected lung cancers:

- results of testing of 1118 surgical specimens and effects of adjuvant gefitinib and erlotinib. **Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer**, v. 7, n. 12, p. 1815–22, 2012.
- DE ROOCK, W. et al. Association of KRAS p.G13D mutation with outcome in patients with chemotherapy-refractory metastatic colorectal cancer treated with cetuximab. **Jama**, v. 304, n. 16, p. 1812–1820, 2010.
- DEL VECCHIO, F. et al. Next-generation sequencing: Recent applications to the analysis of colorectal cancer. **Journal of Translational Medicine**, v. 15, n. 246, p. 1–19, 2017.
- DEMICCO, E. G. et al. Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. **Cell**, v. 171, n. 4, p. 950–965.e28, 2017.
- DENIS, J. A. et al. The Role of BEAMing and Digital PCR for Multiplexed Analysis in Molecular Oncology in the Era of Next-Generation Sequencing. **Molecular Diagnosis and Therapy**, p. 1–14, 2017.
- DIDELOT, A. et al. Competitive allele specific TaqMan PCR for KRAS , BRAF and EGFR mutation detection in clinical formalin fixed paraffin embedded samples. **Experimental and Molecular Pathology**, v. 92, n. 3, p. 275–280, 2012.
- DONALDSON, J.; PARK, B. H. Circulating Tumor DNA: Measurement and Clinical Utility. **Annual Review of Medicine**, v. 69, n. 1, p. 5–12, 2018.
- FANG, S.; WANG, Z. EGFR mutations as a prognostic and predictive marker in non-small-cell lung cancer. **Drug Design, Development and Therapy**, v. 8, p. 1595–1611, 2014.
- FEARON, E.; VOGELSTEIN, B. A genetic model for colorectal tumorigenesis. **Cell**, v. 61, n. 5, p. 759–767, 1990.
- FERLAY, J. et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. **International Journal of Cancer**, v. 127, n. 12, p. 2893–2917, 2010.
- FERLAY, J. et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. **International Journal of Cancer**, v. 136, n. 5, p. E359–E386, 2015.
- FRANCIS, G.; STEIN, S. Circulating cell-free tumour DNA in the management of cancer. **International Journal of Molecular Sciences**, v. 16, n. 6, p. 14122–14142, 2015.
- GRAHAM, D. M. et al. Molecular Subtypes and Personalized Therapy in Metastatic Colorectal Cancer. **Current Colorectal Cancer Reports**, v. 12, n. 3, p. 141–150, 2016.

- HANAHAN, D.; WEINBERG, R. A. The hallmarks of cancer. **Cancer Immunology, Immunotherapy**, v. 60, n. 3, p. 319–326, 2011.
- HATA, A. et al. Complex mutations in the epidermal growth factor receptor gene in non-small cell lung cancer. **Journal of thoracic oncology: official publication of the International Association for the Study of Lung Cancer**, v. 5, n. 10, p. 1524–8, 2010.
- HIDAKA, N. et al. Most T790M mutations are present on the same EGFR allele as activating mutations in patients with non–small cell lung cancer. **Lung Cancer**, v. 108, p. 75–82, 2017.
- HINDSON, C. M. et al. Absolute quantification by droplet digital PCR versus analog real-time PCR. **Nature Methods**, v. 10, n. 10, p. 1003–1005, 2013.
- HUGGETT, J. F. et al. The digital MIQE guidelines: Minimum information for publication of quantitative digital PCR experiments. **Clinical Chemistry**, v. 59, n. 6, p. 892–902, 2013.
- HUGGETT, J. F.; COWEN, S.; FOY, C. A. Considerations for digital PCR as an accurate molecular diagnostic tool. **Clinical Chemistry**, v. 61, n. 1, p. 79–88, 2015.
- HUGGETT, J. F.; WHALE, A. Digital PCR as a novel technology and its potential implications for molecular diagnostics. **Clinical Chemistry**, v. 59, n. 12, p. 1691–1693, 2013.
- KOSAKA, T. et al. Prognostic implication of EGFR, KRAS, and TP53 gene mutations in a large cohort of Japanese patients with surgically treated lung adenocarcinoma. **Journal of thoracic oncology: official publication of the International Association for the Study of Lung Cancer**, v. 4, n. 1, p. 22–9, 2009.
- KRISTENSEN, L. S. et al. Competitive amplification of differentially melting amplicons (CADMA) improves KRAS hotspot mutation testing in colorectal cancer. **BMC cancer**, v. 12, n. 1, p. 548, 2012.
- KUIPERS, E. J. et al. Colorectal cancer. **Nature Reviews Disease Primers**, 2015.
- KUMAR, V. **Robbins Patologia Básica**. [s.l.: s.n.].
- KWAK, E. The role of irreversible HER family inhibition in the treatment of patients with non-small cell lung cancer. **The oncologist**, v. 16, n. 11, p. 1498–507, 2011.
- LAIG, M. et al. Multiplex TaqMan Assays for Rare Mutation Analysis Using Digital PCR. p. 92008, 2016.
- LICHTENBERGER, B. M. et al. Autocrine VEGF Signaling Synergizes with EGFR in Tumor Cells to Promote Epithelial Cancer Development. **Cell**, v. 140, n. 2, p. 268–279,

2010.

LINDEMAN, N. I. et al. Molecular Testing Guideline for Selection of Lung Cancer Patients for EGFR and ALK Tyrosine Kinase Inhibitors: Guideline from the College of American Pathologists, International Association for the Study of Lung Cancer, and Association for Molecular Patho. **Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer**, v. 8, n. 7, p. 823–859, 2013.

LOEB, L. A.; LOEB, K. R.; ANDERSON, J. P. Multiple mutations and cancer. **Proceedings of the National Academy of Sciences**, v. 100, n. 3, p. 776–781, 2003.

MA, G. et al. The prognostic role of pretreatment epidermal growth factor receptor T790M mutation in advanced non-small cell lung cancer patients treated with EGFR tyrosine kinase inhibitors. **Oncotarget**, v. 8, n. 31, p. 50941–50948, 2017.

MAEMONDO, M. et al. Gefitinib or Chemotherapy for Non–Small-Cell Lung Cancer with Mutated EGFR. **New England Journal of Medicine**, v. 362, n. 25, p. 2380–2388, 2010.

MAJUMDAR, N. et al. Poisson Plus Quantification for Digital PCR Systems. **Scientific Reports**, n. December 2016, p. 1–10, 2017.

MASAGO, K. et al. Validation of the digital PCR system in tyrosine kinase inhibitor-resistant *EGFR* mutant non-small-cell lung cancer. **Pathology International**, p. 1–7, 2018.

MERRIMAN, B.; ROTHBERG, J. M. Progress in Ion Torrent semiconductor chip based sequencing. **Electrophoresis**, p. 3397–3417, 2012.

METZKER, M. L. Sequencing technologies — the next generation. **Nature Reviews Genetics**, v. 11, n. 1, p. 31–46, 2009.

MORLEY, A. A. Digital PCR: A brief history. **Biomolecular Detection and Quantification**, v. 1, n. 1, p. 1–2, 2014.

MÜLLER, M. F. et al. Molecular pathological classification of colorectal cancer. **Virchows Archiv**, p. 125–134, 2016.

MULLIS, K. et al. **Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction** Cold Spring Harbor Symposia on Quantitative Biology, 1986.

ONN, A. et al. Synchronous overexpression of epidermal growth factor receptor and HER2-neu protein is a predictor of poor outcome in patients with stage I non-small cell lung cancer. **Clinical cancer research: an official journal of the American**

Association for Cancer Research, v. 10, n. 1 Pt 1, p. 136–43, 2004.

OXNARD, G. R. et al. Brief report: Screening for germline EGFR T790M mutations through lung cancer genotyping. **Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer**, v. 7, n. 6, p. 1049–1052, 2012.

PANDEY, R. V. et al. ClinQC: A tool for quality control and cleaning of Sanger and NGS data in clinical research. **BMC Bioinformatics**, v. 17, n. 56, p. 1–9, 2016.

PAO, W. et al. EGF receptor gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, p. 13306–13311, 2004.

PATEL, S. et al. **Precise and Accurate Determination of MicroRNA Precursors by Digital PCR** .

PEETERS, M. et al. Randomized phase III study of panitumumab with fluorouracil, leucovorin, and irinotecan (FOLFIRI) compared with FOLFIRI alone as second-line treatment in patients with metastatic colorectal cancer. **Journal of Clinical Oncology**, v. 28, n. 31, p. 4706–4713, 2010.

PEKIN, D. et al. Droplet-based microfluidics for the quantitative detection of rare mutations. **Proceedings of MicroTAS2010**, n. October, p. 58–60, 2010.

PEREZ-TORALLA, K. et al. [Digital PCR compartmentalization I. Single-molecule detection of rare mutations]. **Médecine sciences : M/S**, v. 31, n. 1, p. 84–92, 2015.

RACHIGLIO, A. M. et al. Limits and potential of targeted sequencing analysis of liquid biopsy in patients with lung and colon carcinoma. **Oncotarget, Advanced Publications**, v. 7, n. 41, 2016.

RADOVICH, M. et al. The Integrated Genomic Landscape of Thymic Epithelial Tumors. **Cancer Cell**, v. 33, n. 2, p. 244–258.e10, 2018.

ROMA, C. et al. Detection of EGFR Mutations by TaqMan Mutation Detection Assays Powered by Competitive Allele-Specific TaqMan PCR Technology. **BioMed Reserach International**, v. 2013, p. 1–9, 2013.

ROOCK, W. DE et al. KRAS, BRAF, PIK3CA, and PTEN mutations: Implications for targeted therapies in metastatic colorectal cancer. **The Lancet Oncology**, v. 12, n. 6, p. 594–603, 2011.

ROSELL, R. et al. Screening for Epidermal Growth Factor Receptor Mutations in Lung Cancer. **New England Journal of Medicine**, v. 361, n. 10, p. 958–967, 2009.

- ROTHBERG, J. M.; LEAMON, J. H. The development and impact of 454 sequencing. **Nature Biotechnology**, v. 26, n. 10, p. 1117–1124, 2008.
- RUTSAERT, S. et al. Digital PCR as a tool to measure HIV persistence. **Retrovirology**, v. 15, n. 1, p. 16, 2018.
- SAIKI, R. et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. **Science**, v. 239, n. 4839, p. 487–491, 1988.
- SALVIANTI, F.; PINZANI, P. The diagnostic potential of mutation detection from single circulating tumor cells in cancer patients. **Expert Review of Molecular Diagnostics**, v. 0, n. 0, p. 1–7, 2017.
- SAMOWITZ, W. S. et al. Relationship of Ki- ras Mutations in Colon Cancers to Tumor Location , Stage , and Survival : A Population-based Study Relationship of Ki- ras Mutations in Colon Cancers to Tumor Location ,. **Cancer Epidemiology, Biomarkers & Prevention**, v. 9, n. November, p. 1193–1197, 2000.
- SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences**, v. 74, n. 12, p. 5463–5467, 1977.
- SEPULVEDA, A. R. et al. Molecular biomarkers for the evaluation of colorectal cancer: Guideline from The American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and the American Society of Clinical Oncology. **Journal of Clinical Oncology**, v. 35, n. 13, p. 1453–1496, 2017.
- SUDMANT, P. H. ET AL. et al. An integrated map of structural variation in 2 , 504 human genomes. **Nature**, v. 526, n. 7571, p. 75–81, 2015.
- SYKES, P. J. et al. **Quantitation of targets for PCR by use of limiting dilutionBioTechniques**, 1992.
- TATEMATSU, T. et al. The detectability of the pretreatment EGFR T790M mutations in lung adenocarcinoma using CAST-PCR and digital PCR. **Journal of Thoracic Disease**, v. 9, n. 8, p. 2397–2403, 2017.
- THRESS, K. S. et al. EGFR mutation detection in ctDNA from NSCLC patient plasma: A cross-platform comparison of leading technologies to support the clinical development of AZD9291. **Lung Cancer**, v. 90, n. 3, p. 509–515, 2015.
- TOL, J. et al. High sensitivity of both sequencing and real-time PCR analysis of KRAS mutations in colorectal cancer tissue. **Journal of Cellular and Molecular Medicine**, v. 14, n. 8, p. 2122–2131, 2010.
- TONG, Y. et al. Application of Digital PCR in Detecting Human Diseases Associated

- Gene Mutation. **Cellular Physiology and Biochemistry**, v. 43, n. 4, p. 1718–1730, 2017.
- TORRE, L. A. et al. Global Cancer Incidence and Mortality Rates and Trends--An Update. **Cancer Epidemiology Biomarkers & Prevention**, v. 25, n. 1, p. 16–27, 2016.
- VAN KRIEKEN, J. H. J. et al. RAS testing in metastatic colorectal cancer: advances in Europe. **Virchows Archiv**, v. 468, n. 4, p. 383–396, 2016.
- VOELKERDING, K. V.; DAMES, S. A.; DURTSCHI, J. D. Next-generation sequencing:from basic research to diagnostics. **Clinical Chemistry**, v. 55, n. 4, p. 641–658, 2009.
- VOGELSTEIN, B.; KINZLER, K. W. Digital PCR. **Proceedings of the National Academy of Sciences USA**, v. 96, n. August, p. 9236–9241, 1999.
- WHALE, A. S. et al. Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation. **Nucleic Acids Research**, v. 40, n. 11, 2012.
- WHALE, A. S. et al. Detection of Rare Drug Resistance Mutations by Digital PCR in a Human Influenza A Virus Model System and Clinical Samples. **Journal of Clinical Microbiology**, v. 54, n. 2, p. 392–400, 2016.
- WHIFFIN, N. et al. Identification of susceptibility loci for colorectal cancer in a genome-wide meta-analysis. **Human Molecular Genetics**, v. 23, n. 17, p. 4729–4737, 2014.
- YEH, P. et al. DNA-Mutation Inventory to Refine and Enhance Cancer Treatment (DIRECT): A Catalog of Clinically Relevant Cancer Mutations to Enable Genome-Directed Anticancer Therapy. **Clinical Cancer Research**, v. 19, n. 7, p. 1894–1901, 2013.
- YU, B.; O'TOOLE, S. A.; TRENT, R. J. Somatic DNA mutation analysis in targeted therapy of solid tumours. **Translational pediatrics**, v. 4, n. 2, p. 125–38, 2015.
- YUNG, T. K. F. et al. Single-Molecule Detection of Epidermal Growth Factor Receptor Mutations in Plasma by Microfluidics Digital PCR in Non-Small Cell Lung Cancer Patients. **Clinical Cancer Research**, v. 15, n. 6, p. 2076–2084, 2009.
- ZHANG, B. et al. Comparison of droplet digital PCR and conventional quantitative PCR for measuring EGFR gene mutation. **Experimental and Therapeutic Medicine**, v. 9, n. 4, p. 1383–1388, 2015.
- ZONTA, E. et al. Multiplex Detection of Rare Mutations by Picoliter Droplet Based Digital PCR: Sensitivity and Specificity Considerations. **Plos One**, v. 11, n. 7, p. 1–20,

2016.

ANEXOS

ANEXO 1. OLIGONUCLEOTÍDEOS SINTÉTICOS UTILIZADOS COMO
CONTROLES POSITIVOS

Gene sintético	Tamanho (pb)	Sequência 5'-3'	Mutações
1	277	TGGAGAAGCTCCCAACCAAGCGGACTCTGGATCCCAAGAAAGGTGAGCTGCAGTGTCTGGGCTCCGGTCAAGGAATTAAGAGAAGCAACATCTCCGCATGGCCAAACAGGAAATCCTCGATGTGAGGGTCCCTGGCACAGGCCTGAGGTTGAGAGCCATGGTCTGCCTCACCTCCACCCGACCATGCGAAGCCACACTGACAGCTCATACGCAGCTCATGGTGTGATGCCAGCGTGGACAACCCCTTCGGCTGCCTCTGGAGGTGAGGCAGATGCC	EGFR não mutado KRAS não mutado
2	330	TGGAGAAGCTCCCAACCAAGCGGACTCTGGATCCCAAGAAAGGTGAGAGTGTGCCCTCCGGTCTGCCCTCACCTCCACCGAATAAATCATAAGTGATTAACCTTATGTGAACCATGCGAAGCCACACTGAAGTCTGCCCTCCGGTGGCCAAACAGGAAATCCTCGATGTGAGCAGCTCATCATGCAGCTCATGTAGTTGGAGCTAGTGGCGTAGGCTGGTGTGATGGCCATCGTGGACAACCGTCCCTGGCACAGGCCTGAGGTTGAGAGCCATGGCCCTTCGGCTGCCTCCACGATACAGCTAAATTCAGATGGAGGTGAGGCAGATGCC	G719A ALELO C G719S ALELO A T790M ALELO T S768I ALELO T 19 DEL G12S ALELO A
3	358	CCCTCCCTCCAGGAAGCCCTGGCAGCCGGAACGTACTGGAGAAGCTCCCAACCAAGCGCAGCATGTCAAGATCACAGATTTTGGGAAATAATCATAAGTGTATTAACCTTATGTGTGATGATGCCAGCGTGGACAAGATTTTGGCGGGCCAAACTGCAAAAGTGCTGAGCTCCGGTGTGGCCAAACAGCTGGGTGCTCTGCACACACCAGTTGAGGGTGGCTTTAGGTCAGCCGACAGTTGGAGCTTGTGGCGTAGGGGTCCCTGGCACAGGCCTGAGGTTGAGAGCCATACCATGCAGAAGGAGGACGATACAGCTAAATTCAGAGGATTCAGATGGAGGTGAGGCAGATGCC	L858R ALELO G L861Q ALELO A G12C ALELO T
4	204	TGGAGAAGCTCCCAACCAAGCGGTGCACCTGTAAATAATCCAGACTGAATAAAATCATAAAGTGTATTAACCTTATGTGTGACAAAAGTGTGTGCTCCGGTGCACACAGCAGGTAAGAGGAGTACAGAGTTGGAGCTCGTGGCGTAGGGTCCCTGGCACAGGCCCTCATGTACTGGTCCCTCATTTGCACGATACAGCTAAATTCAGA	G719C ALELO T G12R ALELO C
5	233	GGCCGATAAATCATAAAGTGTATTAACCTTATGTGTGAGGTGCACCTGTAAATAATCCAGAGGCGGATAGTGTATTAACCTTATGTGTAGTCACATTTGGCGTCAGGCCCTGCTGAAAATGACTGAATATAACTTGTGGTAGTTGGAGCTGATGGCGTAGGCCAACAACAGCAGGTGAAGAGGAGTACAAGAGTGCCTTGACGATACAGCTAAATTCAGA	G12D ALELO A
6	145	GGCCGATAAATCATAAAGTGTATTAACCTTATGTGTGAGGCGATAGTAAAGTTGGAGCTGTGGCGTAGGCCACAGGCCTGCTGAAAATGACTGAATATAACTTGTGGTAGTTGGAGAGTGCCTTGACGATACAGCTAAATTCAGA	G12V ALELO T
7	147	GGCCGATAAATCATAAAGTGTATTAACCTTATGTGTGAATAGCCGGGTGATTGCAACCTGTCAGCCTGCTGAAAATGACTGAATATAAAGTGTGGTAGTTGGAGCTGCTGGCGTAGGAGAGTGCCTTGACGATACAGCTAAATTCAGA	G12A ALELO C
8	149	GGCCGATAAATCATAAAGTGTATTAACCTTATGTGTGAATAGCCGGGTGATTGCAACCTGTCAGCCTGCTGAAAATGACTGAATATAAAGTGTGGTAGGGAGCTGGTGACCTAGGCCAAGAGAGTGCCTTGACGATACAGCTAAATTCAGA	G13D ALELO A

ANEXO 2. OLIGONUCLEOTÍDEOS INICIADORES E SONDAS DE HIDRÓLISE DOS SISTEMAS DE DETECÇÃO DE MUTAÇÕES DO GENE *EGFR*.

MUTAÇÃO	dbSNP	SEQUÊNCIA	Tm
		Primer direto	60,4°C
		Primer reverso	60,0°C
DEL19	rs121913424	Sonda referência	69,8°C
		Sonda para o alelo mutado	67,6°C
		Primer direto	60,4°C
		Primer reverso	60,8°C
p.L858R	rs121434568	Sonda para o alelo mais comum	70,9°C
		Sonda para o alelo mutado	69,8°C
		Primer direto	59,6°C
		Primer reverso	59,1°C
p.T790M	rs121434569	Sonda para o alelo mais comum	68,7°C
		Sonda para o alelo mutado	66,4°C
		Primer direto	59,8°C
		Primer reverso	59,5°C
		Sonda para o alelo mais comum	70,9°C
		Sonda p.G719A	70,9°C
		Sonda p.G719S	71,3°C
p.G719X	rs28929495	Sonda p.G719C	71,7°C
		Sonda G719X	70,9°C
		Primer direto	59,6°C
		Primer reverso	59,8°C
p.S768I	rs121913465	Sonda para o alelo mais comum	71,3°C
		Sonda para o alelo mutado	68,7°C
		Primer direto	60,2°C
		Primer reverso	60,1°C
p.L861Q	rs121913444	Sonda para o alelo mais comum	71,7°C
		Sonda para o alelo mutado	71,7°C

Legenda: Em negrito, o nucleotídeo variante. Tm = temperatura de acoplamento dos oligonucleotídeos.

ANEXO 3. OLIGONUCLEOTÍDEOS INICIADORES E SONDAS DE HIDRÓLISE DOS SISTEMAS DE DETECÇÃO DE MUTAÇÕES DO GENE KRAS.

MUTAÇÃO	dbSNP	SEQUÊNCIA	T _m
	rs121913529	Primer direto 5'AATAAATCATAAAGTGTATTAACCTTATGTGTGA3'	53,7°C
		Primer reverso 5'AATAAATCATAAATCTGAATTAGCTGTATCGT3'	53,4°C
		Sondas para o alelo mais comum 5' FAM AGTTGGAGCTGGTGGCGTAGG MGB-NFQ 3' 5' VIC AGTTGGAGCTGGTGGCGTAGG MGB-NFQ 3'	71,7°C 71,7°C
	p.G12S	5' VIC TAGTTGGAGCTAGTGGCGTAGGC MGB-NFQ 3'	70,4°C
	p.G12C	5' FAM AGTTGGAGCTTGTGGCGTAGG MGB-NFQ 3'	69,7°C
	p.G12R	5' FAM AGTTGGAGCTCGTGGCGTAGG MGB-NFQ 3'	71,5°C
	p.G12D	5' VIC AGTTGGAGCTGATGGCGTAGGCA MGB-NFQ 3'	72,8°C
	p.G12V	5' VIC AGTTGGAGCTGTTGGCGTAGGCA MGB-NFQ 3'	73,2°C
	p.G12A	5' FAM TTGGAGCTGCTGGCGTAGG MGB-NFQ 3'	70,2°C
		Primer direto 5'AATAAATCATAAAGTGTATTAACCTTATGTGTGA3'	53,7°C
		Primer reverso 5'ATAAATCATAAATCTGAATTAGCTGTATCGT3'	53,4°C
		Sonda para o alelo mais comum 5' FAM AGTTGGAGCTGGTGGCGTAGG MGB-NFQ 3'	71,7°C
p.G13D	rs121913535	Sonda para o alelo mutado 5' VIC GGAGCTGGTACCGTAGGCAAG MGB-NFQ 3'	70,1°C

Legenda: em negrito o nucleotídeo variante. T_m = temperatura de acoplamento dos oligonucleotídeos.

Legenda: As amostras estão dispostas na primeira coluna de 1 a 41, e cada linha corresponde a uma das mutações testadas pelas diferentes técnicas. Em rosa, as mutações detectadas nas amostras por PCR digital. Em azul, as mutações detectadas por *castPCR*. Em laranja, as mutações detectadas por sequenciamento.

Nota: A correlação entre a detecção de mutação pelas diferentes técnicas foi fraca, visto que a PCR digital detectou mais amostras mutadas.

APÊNDICES

APÊNDICE 1. LAUDO REPRESENTATIVO DE PACIENTE COM AUSÊNCIA DE MUTAÇÃO NO GENE *KRAS*.

Paciente: ABCDE

Número de requisição: 12345

Dr.(a) XXXXXXXX

Data da coleta: 19/03/2018

Data da liberação: 23/03/2018

Data de Nascimento: 00/00/0000

Convênio: Particular

Responsável Técnico: XXXXXXXX

ANÁLISE DE MUTAÇÕES SOMÁTICAS NO GENE *KRAS*

Material: tecido tumoral

Método: PCR digital

Região alvo: exon 2 do gene *KRAS*

RESULTADO: MUTAÇÃO NÃO DETECTADA

GENÓTIPO: SELVAGEM

Interpretação: a literatura corrente sugere que a eficácia das terapias com alvo em EGFR em pacientes com câncer colorretal restringe-se a pacientes em cujos tumores não possuem mutações no gene *KRAS*. Portanto, atentando-se ao limite de detecção da técnica, a não detecção de mutação em *KRAS* nesse espécime de tumor sugere que esse paciente possa responder a tais terapias.

Limite de detecção da técnica: 1 molécula de DNA mutante em 100 moléculas de DNA selvagem.

O teste foi baseado em PCR digital através da amplificação de um painel com 7 mutações dentro dos códons 12 e 13 do gene *KRAS*: Glicina12Serina; Glicina12Cisteína; Glicina12Arginina; Glicina12Aspartato; Glicina12Valina; Glicina12Alanina; Glicina13Aspartato. Os resultados do teste devem ser interpretados no contexto clínico, junto a amostragens tumorais e outros dados laboratoriais pertinentes. Caso os resultados obtidos não correspondam a outros achados laboratoriais ou clínicos, por gentileza, contate o laboratório para possível interpretação. Erros na interpretação dos resultados podem ocorrer se a informação provida é incorreta ou incompleta.

Esse método foi desenvolvido e validado pela equipe do Laboratório Genoprimer.

APÊNDICE 2. LAUDO REPRESENTATIVO DE PACIENTE COM AUSÊNCIA DE MUTAÇÃO NO GENE *EGFR*.

Paciente: ABCDE

Número de requisição: 12345

Dr.(a) XXXXXXXX

Data de Nascimento: 00/00/0000

Responsável Técnico: XXXXXXXX

Data da coleta: 19/03/2018

Data da liberação: 23/03/2018

Convênio: Particular

ANÁLISE DE MUTAÇÕES SOMÁTICAS NO GENE *EGFR*

Material: tecido tumoral

Método: PCR digital

Região alvo: exons 18-21 do gene *EGFR*

RESULTADO: MUTAÇÃO NÃO DETECTADA

GENÓTIPO: SELVAGEM

Interpretação: a literatura corrente sugere que a eficácia das terapias com alvo em *EGFR* em pacientes com câncer de pulmão de células não pequenas possui melhor resposta na presença de mutações ativadoras (DEL19 e p.L858R). No entanto, a presença da mutação p.T790M está associada a resistência ao tratamento. Portanto, atentando-se ao limite de detecção da técnica, a não detecção de mutações em *EGFR* nesse espécime de tumor sugere que esse paciente não possui resistência a tais terapias, e seu prognóstico não é determinado pela presença de mutações ativadoras. Limite de detecção da técnica: 1 molécula de DNA mutante em 100 moléculas de DNA selvagem.

O teste foi baseado em PCR digital através da amplificação de um painel com 7 mutações dentro dos exons 18 a 21 do gene *EGFR*: Glicina719X; possíveis deleções no exon 19; Treonina790Metionina; Serina768Isoleucina; Leucina858Arginina; Leucina861Glutamina. Os resultados do teste devem ser interpretados no contexto clínico, junto a amostragens tumorais e outros dados laboratoriais pertinentes. Caso os resultados obtidos não correspondam a outros achados laboratoriais ou clínicos, por gentileza, contate o laboratório para possível interpretação. Erros na interpretação dos resultados podem ocorrer se a informação provida é incorreta ou incompleta.

Esse método foi desenvolvido e validado pela equipe do Laboratório Genoprimer.

APÊNDICE 3. ARTIGO EM DESENVOLVIMENTO.

dPCR detection of *EGFR* somatic mutations in non-small-cell lung cancer FFPE samples.

Tamyres M. Carvalho¹, Renata M. Dourado^{1 2}, Sueli M. Nakatani², Cesar A. B. Duarte², Angelica B. W. Boldt¹.

Affiliations

¹Laboratório de Genética Molecular Humana, Universidade Federal do Paraná - Curitiba – BR.

²Laboratório Genoprimer – Diagnóstico Molecular e Medicina Personalizada – Curitiba - BR.

Contact information

Angelica B.W.Boldt, angelicaboldt@gmail.com.

Keywords: Digital PCR; *EGFR* mutations; tumor biopsy; NSCLC; Brazilian patients.

Abstract

Screening of endothelial growth factor receptor (*EGFR*) mutations is mandatory in non-small-cell lung cancer (NSCLC). Digital PCR (dPCR) is proposed to replace Sanger sequencing or real-time PCR for detection and quantification of rare mutations. In order to validate dPCR to detect *EGFR* somatic mutations in NSCLC, we extracted DNA from formalin-fixed paraffin-embedded (FFPE) samples of 57 NSCLC and 10 free-cancer Brazilian patients, collected between 2003 and 2007, screening the *EGFR* mutations in exons 18-21 using dPCR, Sanger Sequencing and Competitive Allele-Specific PCR (castPCR), and determining dPCR sensitivity, limits of detection (LOD) and quantification (LOQ). We found 27 samples with *EGFR* mutations by dPCR, versus 11 by castPCR ($p=0.014$) and 4 by Sanger Sequencing ($p=0.0002$). LOD was 100 molecules of DNA/uL and LOQ, 1%. Most of the samples (86.96%) identified by castPCR as wild-type and by dPCR as mutated, presented less than 10% mutated DNA molecules (mean 4.57%). We found no mutations in free-cancer samples. Accuracy was 94.44%, as measured with the assay recommended by the College of American Pathologist. These results indicated higher sensibility and specificity of

dPCR for screening *EGFR* mutations in NSCLC biopsies, compared with Sanger Sequencing and castPCR.

Introduction

Lung carcinoma is the most common diagnosed cancer, worldwide. It is also the leading global cause of cancer-related mortality, due to smoking and tobacco consumption (TORRE et al., 2016). The disease involves the accumulation of mutations, which provide advantages to the cell for survival and/or proliferation (DENIS et al., 2017). Among these mutations, those in the epidermal growth factor receptor gene (*EGFR*) may cause chemo-resistance and lead to therapy failure. Guidelines of the College of American Pathologists, the International Association for the Study of Lung Cancer and the Association for Molecular Pathology, thus recommend to initially screen *EGFR* mutations, in order to select the best therapeutic strategy (LINDEMAN et al., 2013). Testing for *EGFR* should be even prioritized over other molecular markers in non-small-cell lung cancer (NSCLC) (LINDEMAN et al., 2013).

EGFR mutations are commonly detected by Sanger sequencing, real-time quantitative polymerase chain reaction (qPCR), and single-base extension genotyping (SBE), each with different advantages (LINDEMAN et al., 2013). Sanger sequencing of PCR-amplified *EGFR* fragments may detect every kind of mutation, including insertions/deletions (indels) (VAN KRIEKEN et al., 2016). Nevertheless, detection is expensive and limited by a minimum of 10 - 25% of cells carrying the mutation (CAROTENUTO et al., 2010; PANDEY et al., 2016). In contrast, qPCR relies on previous knowledge of the mutations, in order to design mutation-specific hydrolysis probes that will allow relative yet more sensitive quantification of mutated fragments, compared with a non-mutated generic sequence (BUSTIN; HUGGETT, 2017; TOL et al., 2010).

Digital PCR (dPCR) was conceptually idealized in the nineties (SYKES et al., 1992). It is designed to occur in droplets (ddPCR) or wells (chip-based). The term “digital” means that each partition should ideally have only one of two values: 0 (no target) or 1 (amplification and consequent detection of a fluorescent signal, corresponding to either mutant or wild-type sequence) (DENIS et al., 2017; VOGELSTEIN; KINZLER, 1999). It allows nanofluidic multiple real-time PCR amplifications of the same kind to occur simultaneously, each on a single DNA molecule and completely isolated from each other. Mutation detection is given by fluorescent hydrolysis probes, probing in

identical compartments the fragments that result from each amplification round of an original single DNA template. Thus, dPCR may be less susceptible to inhibitors (HUGGETT et al., 2013).

The strategy of limiting DNA dilution allows random distribution of templates into individual dPCR partitions and eliminates the need for external calibrators and standard curves, necessary in classical qPCR, to quantify the absolute and relative copy number of target DNA (BHAT et al, 2009; BHAT et al, 2010). Instead, for a given proportion of positive dPCR partitions, it is possible to estimate the amount of mutated DNA molecules through Poisson statistics (MAJUMDAR et al., 2017; SYKES et al., 1992). Random distribution of the templates in identical compartments also allows fair detection of even a single mutated DNA molecule in a given amount of genomic DNA. In other words, dPCR seems perfectly suited for the assessment of low-abundance mutations, since it allows both absolute quantification of tiny amounts of mutated DNA molecules and accurate determination of the mutant to wild-type DNA ratio of a given sample (DENIS et al., 2017).

Although dPCR has been validated to screen p.T790M in formalin-fixed paraffin-embedded (FFPE) (MASAGO et al., 2018) and liquid biopsies (THRESS et al., 2015), its performance for the detection of other mutations, especially in FFPE samples, has not been characterized. Precise identification of these mutations is highly relevant for epidemiological retrospective studies. In order to fill in this gap, we designed the probes, validated and analyzed the performance of dPCR test to detect *EGFR* somatic mutations in samples of NSCLC Brazilian patients.

Methods

The study was approved by the Ethics Commission of Research of the “Liga Paranaense de Combate ao Câncer do Hospital Erasto Gaertner”, protocol 885.643/2014.

Samples

FFPE tumor samples from fifty-seven patients with pulmonary adenocarcinoma, collected between 2003 and 2007, were retrospectively included in this study. All patients were from Hospital Nossa Senhora das Graças and Hospital Erasto Gaertner. We also included ten free-cancer FFPE samples as negative controls. DNA was extracted with QIAamp DNA Micro Kit (Qiagen, Hilden, DE), and diluted to 1ng/μL.

Analytical and Clinical Validation

We used three methods to detect *EGFR* mutations: dPCR, Sanger Sequencing and castPCR. Synthetic DNA (Integrated DNA Technologies) carrying the *EGFR* mutations and with wild-type DNA sequences were used to determine the sensitivity and limit of detection (LoD) of dPCR assays. We produced DNA mixtures with different fractional concentration of mutant DNA (50%, 10%, 5% and 1%) in a range of wild-type DNA sequences (also synthetic) to determine limit of quantification (LOQ) of p.T790M and p.L861Q dPCR assays. Accuracy of dPCR assays was measured by a test developed by the College of American Pathologists (CAP) laboratory accreditation program.

dPCR

dPCR primers and hydrolysis probes designed for detecting assay *EGFR* mutations are listed in Table 1. One probe was specific for the wild-type allele, and the other one, for the mutated variant (or for the most common variant in the case of multiple alleles). We performed the reactions on a *QuantStudio™ 3D Digital PCR System*, using dPCR reagents from Thermo Fisher Scientific (USA). The final 15 µL of PCR reaction mixture was made up in the following way: 1X *QuantStudio™ 3D Digital PCR Master Mix*, 900nM of each primer and 200nM of each probe, 3 ng of diluted DNA and ultrapure water, and then loaded into the *QuantStudio™ 3D Digital PCR Chip*, which has 20000 mini-chambers. To perform the PCR using the *ProFlex™ PCR System*, thermal cycling profile was 10 minutes of incubation at 96°C, followed by 39 cycles of 60°C for 2 minutes, 98°C for 30 seconds, 60°C for 2 minutes, and then 4°C hold. We used the *QuantStudio™ 3D Digital PCR* instrument to read the chip and the *QuantStudio 3D Analysis Suite* Software for subsequent analysis.

Sanger sequencing

Analysis of *EGFR* mutations in tumor tissues by sequencing was done by PCR amplification of tumor DNA using primers specific for the exons 18, 19, 20 and 21 of the *EGFR* gene (Table 1). PCR products were purified and sequenced using the *BigDye Terminator Cycle Sequencing Ready Reaction Kit* (Applied Biosystems) and analyzed by *Mutation Surveyor* (*StataSoft*s).

castPCR

Competitive allele-specific TaqMan® PCR (castPCR) *EGFR* mutation assay was performed following the manufacturer's instructions (Thermo Fisher Scientific, Waltham, EUA). This method is based on allelic-specific primers for the mutant variant, in addition to a minor groove binding (MGB) blocker for the most common allele, on

the same PCR reaction. In addition, for mutation detection assays, a gene reference was added to confirm the specificity of the amplification.

Statistical analysis

To evaluate the efficiency of the elaborated detection systems we used a linear regression test with the Ct values obtained by the linearity experiment. Descriptive statistics and correlation between the results obtained with the different methods were calculated with GraphPad Prism v.5.01.

Results and Discussion (not finished)

LoD was 100 molecules of DNA/mL, and LoQ was 1% of mutant DNA in a range of wild-type DNA. Mutant DNA concentrations higher than 1% were found with homogeneous and efficiency distribution ($r^2 = 0.9906$) (Figure 2).

Among the three methods used to detect *EGFR* mutations in the FFPE samples, we found a higher number of *EGFR* gene mutations by dPCR, compared with castPCR or Sanger Sequencing (Tables 2 and 3 and Figure 3), suggesting that dPCR was more effective in detecting rare mutations present among the predominantly wild-type sequences. We also found a trend for higher dPCR-detection, compared to castPCR and Sanger Sequencing for *EGFR* p.L858R ($p=0.09$) (Table 3). As expected, we found no mutations in free-cancer samples, proving the dPCR specificity.

We observed relatively weak, yet highly significant correlations for presence/absence of mutations detected with dPCR and with *castPCR* ($p<0.0001$, $r^2=0.53$), as well as with dPCR and with Sanger Sequencing ($p=0.019$, $r^2=0.30$). This reflects the higher specificity of dPCR, compared to the other two methods. There was no difference between the number of DNA molecules carrying *EGFR* mutations/sample detected by either dPCR and castPCR ($p=0.1297$).

After restricting the analysis to samples judged as wild-type by *castPCR*, but with detectable mutations by dPCR, we found that most of these samples (86.96%) presented a mutation frequency lower than 10% (mean 4.16 ± 6.03) (Figure 4). This data leads us to suggest that dPCR presents advantages over *castPCR* to detect samples with a lower range of mutated DNA. Accuracy was 94.44%, as measured with the assay recommended by the College of American Pathologist (CAP). This result was expected, since compared with qPCR, dPCR is indeed more effective for detecting rare mutations (Huggett & Whale, 2013).

CONCLUSION

Regarding detection of *EGFR* gene mutations in FFPE samples of NSCLC Brazilian patients, our results lead us to recommend dPCR for its superior sensitivity and specificity than other methods, as Sanger Sequencing and castPCR.

Abbreviations

NSCLC: non-small-cell lung cancer; EGFR: epidermal growth factor receptor; FFPE: formalin-fixed and paraffin-embedded; dPCR: Digital PCR; castPCR: Competitive Allele-Specific TaqMan® PCR; LoD: limit of detection; LoQ: limit of quantification.

Author Contributions

T.M.C, R.M.D., S.M.N and C.A.B.D. designed the study. T.M.C selected case samples, prepared them for experiments and perform most of the Digital PCR experiments. R.M.D perform some Digital PCR experiments, and also assist T.M.C with castPCR and Sanger sequencing experiments. T.M.C made bioinformatics and statistical analyses with R.M.D., S.M.N, C.A.B.D and A.B.W.B assistance. T.M.C drafted the manuscript. All authors reviewed and contributed to the manuscript.

Acknowledgments

This study would not have been possible without the contributions of Laboratório Genoprimer, and also Laboratório de Citogenética Humana e Oncogenética from Universidade Federal do Paraná and Hospital Erasto Gaertner that kindly donated us the FFPE samples.

Conflicts of Interest

The authors has no conflict of interest.

Funding

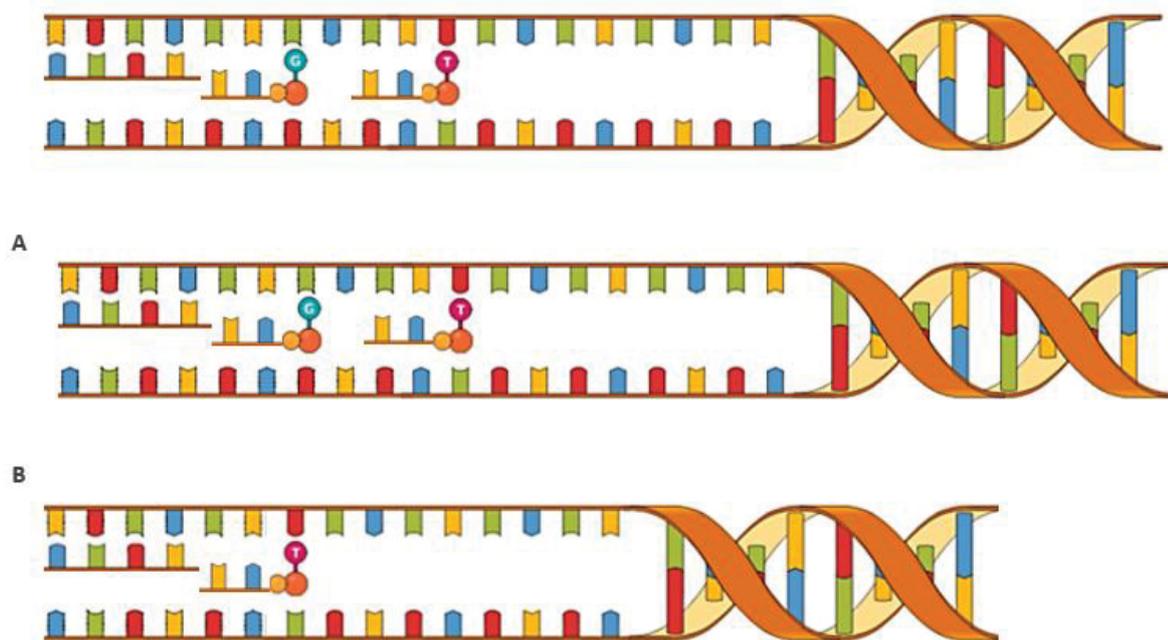
Support for this study was provided by Financiadora de Estudos e Projetos (FINEP) in a project named TECNOVA-PR from Fundação Araucária and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Table 1. Primers and probes using sequences for the 3D digital PCR assays.

MUT	dbSNP	Primer/probe	SEQUENCE
		Primer forward	5' G G A C T C T G G A T C C C A G A A G G T G A G 3'
		Primer reverse	5' C T G A G G T T C A G A G C C A T G G 3'
DEL19	rs121913424	Reference probe	5'VIC GCCAACAAGGAAATCCTCGATGTGAG MGB-NFQ 3'
		Mutation probe	5'FAM CAAGGAATTAAGAGAAGCAACATCTCCG MGB-NFQ 3'
		Primer forward	5' C T G G C A G C C A G G A A C G T A C T G 3'
		Primer reverse	5' G T C G G C T G A C C T A A G C C A C C 3'
p.L858R	rs121434568	Wild-type probe	5'FAM AGATTTTGGGC <u>I</u> GGCCAAACTGC MGB-NFQ 3'
		Mutation probe	5' VIC GATTTTGGGC <u>G</u> GGCCAAACTG MGB-NFQ 3'
		Primer forward	5'TCTGCCTCACCTCCACCG3'
		Primer reverse	5' GGAGGCAGCCGAAGGG3'
p.T790M	rs121434569	Wild-type probe	5' FAM CAGCTCATCA <u>C</u> GCAGCTCATG MGB-NFQ 3'
		Mutation probe	5' VIC CAGCTCATCA <u>T</u> GCAGCTCATG MGB-NFQ 3'
		Primer forward	5' T G G A G A A G C T C C C A A C C A A G C 3'
		Primer reverse	5' G C C T G T G C C A G G G A C C 3'
		Wild-type probe	5' FAM AGTGCTGG <u>G</u> CTCCGGTG MGB-NFQ 3'
		Mutation probe p.G719A	5' VIC AGTGCTGG <u>C</u> CTCCGGTG MGB-NFQ 3'
		Mutation probe p.G719S	5' FAM CAAAGTGCTG <u>A</u> GCTCCGGTGC MGB-NFQ 3'
p.G719X	rs28929495	Mutation probe p.G719C	5' VIC CAAAGTGCTG <u>I</u> GCTCCGGTGC MGB-NFQ 3'
		Primer forward	5'ACCATGCGAAGCCCACTGA3'
		Primer reverse	5'TGGAGGTGAGGCAGATGCC3'
p.S768I	rs121913465	Wild-type probe	5'FAM GTGATGGCCA <u>G</u> CGTGGACAAC MGB-NFQ 3'
		Mutation probe	5' VIC GTGATGGCCA <u>I</u> CGTGGACAAC MGB-NFQ 3'
		Primer forward	5'GCAGCATGTCAAGATCACAGATTTTGGG'
		Primer reverse	5'CCTCCTTCTGCATGGTATTCTTTCTTCC3'
		Wild-type probe	5' FAM TGGCCAAAC <u>I</u> GCTGGGTGC MGB-NFQ 3'
p.L861Q	rs121913444	Mutation probe	5' VIC TGGCCAAAC <u>A</u> GCTGGGTGC MGB-NFQ 3'

In bold and underlined: variant nucleotides

Figure 1. Detection of *EGFR* exon 19 deletion.



Legend: (A) detection of a no-mutated sample. The wild-type specific probe annealed the *EGFR* exon 19 where deletion most frequently occurs. The other probe annealed a nearby region where mutation/deletion; (B) when the region was deleted, the wild-type specific probe would not be able to anneal to the template, giving no fluorescence signal. The other probe annealed a nearby region where mutation/deletion and acted as a reference signal to the presence of a DNA molecule regardless of the mutational status of the molecule.

Table 2. Cases of lung cancer samples carrying *EGFR* gene mutations, detected by dPCR, *castPCR* or Sanger Sequencing.

	All	Only 1 mutation	Complex Mutations	p value*
dPCR	24	18	6	--
<i>castPCR</i>	11	8	3	0,014
Sanger	4	4	0	0,00002

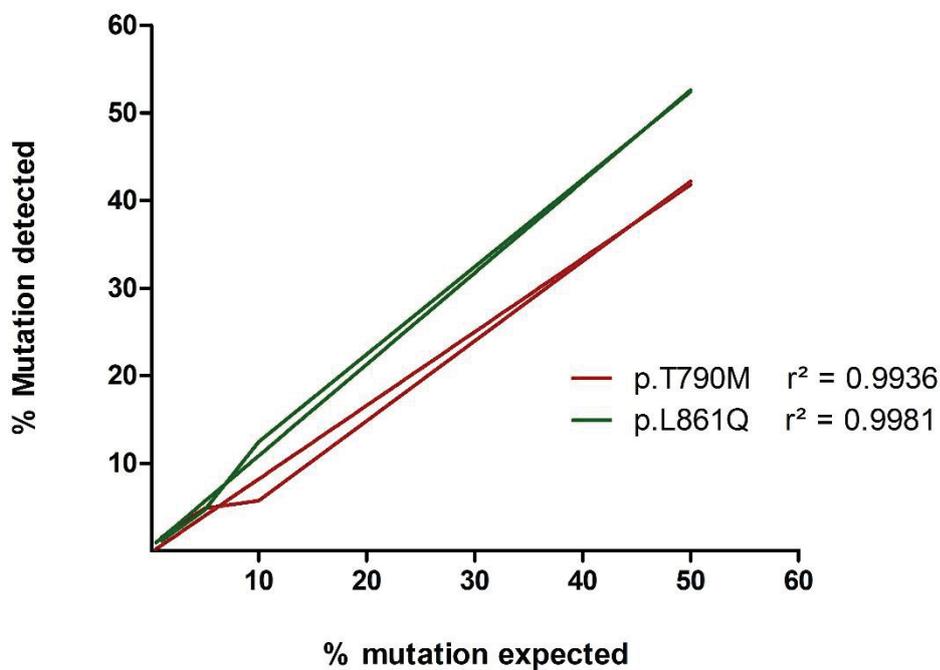
Legend: the values are absolute numbers of samples that showed some *EGFR* mutation. (*) the statistical value was measured compared to dPCR.

Table 3. *EGFR* gene mutations detected on lung cancer samples by dPCR, *castPCR* or Sanger Sequencing.

Mutation	dPCR			<i>castPCR</i>				Sanger			
	All	Simple	Complex	All	Simple	Complex	P value	All	Simple	Complex	P value
DEL19	10	6	3	4	2	2	<u>0,04</u>	4	4	0	<u>0.04</u>
p.L858R	8	6	2	2	2	0	0,09	0	0	0	<u>0.04</u>
p.L861Q	1	0	1	0	0	0	ns	0	0	0	ns
p.G719X	4	0	4	2	0	2	ns	0	0	0	0.04
p.T790M	8	6	2	5	3	2	ns	0	0	0	<u>0.002</u>
p.S768I	1	0	1	1	1	0	ns	0	0	0	ns
Total	32			14				4			

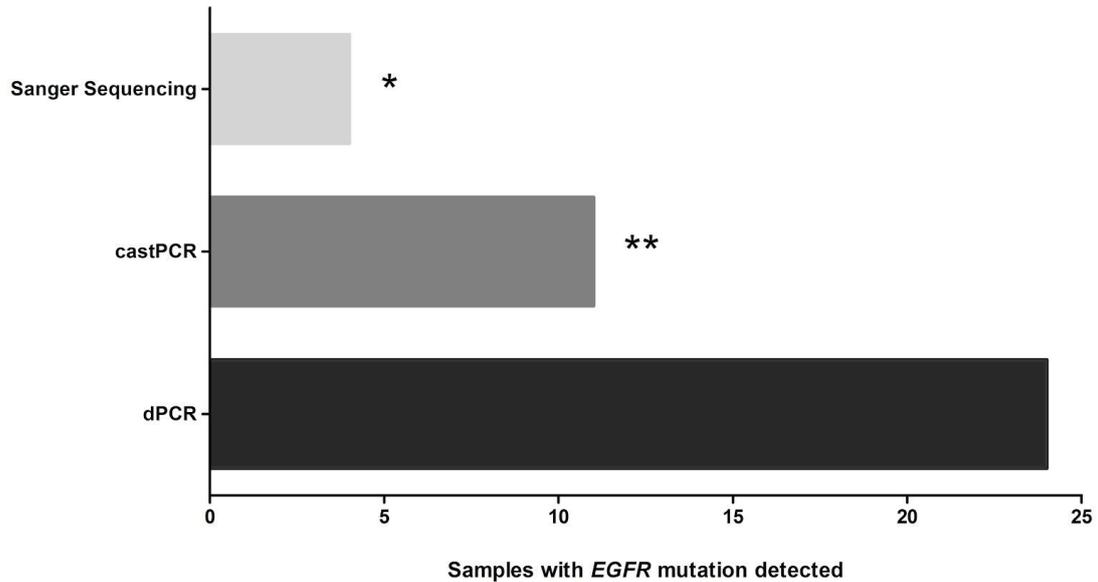
Legend: ns = no statistical significance.

Figure 2. Limit of quantification (LoQ) of Digital PCR assay.



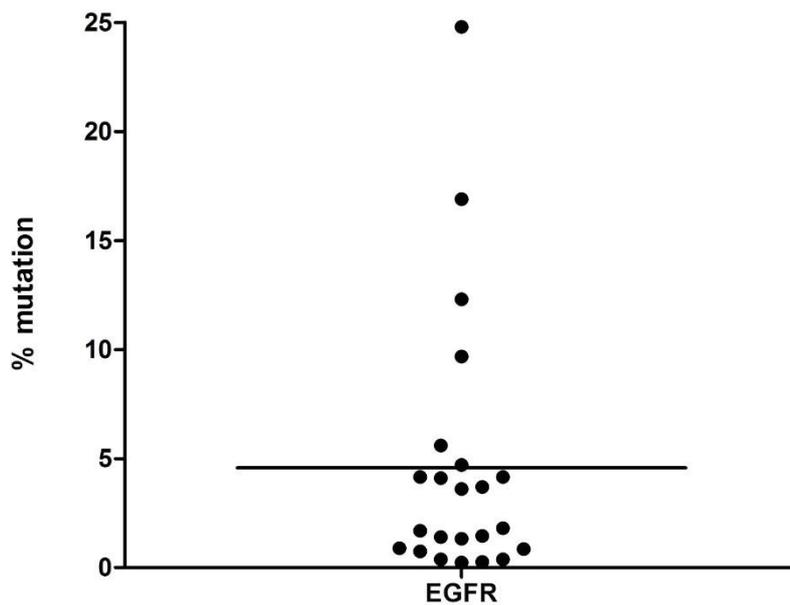
Legend: Linear regression to determine limit of quantification (LoQ) of p.T790M and p.L861Q dPCR assays. On X axis, DNA mixtures with different fractional concentration of mutant DNA in a range of wild-type DNA sequences On Y axis, DNA concentration that was found using dPCR technique.

Figure 3. Cases of lung cancer samples with *EGFR* mutations, detected by dPCR, castPCR or Sanger sequencing.



Legend: (**) dPCR was capable to detect a higher number of mutated cases compared to *castPCR* ($p=0.014$), and also to (*) Sanger sequencing ($p=0.00002$).

Figure 4. Frequency of dPCR-detected *EGFR* mutations in lung cancer samples.



Legend: On Y axis, DNA concentration that was found using dPCR technique. Each point reflect one measured sample.