

UNIVERSIDADE FEDERAL DO PARANÁ

BRUNA DA SILVA OLIVEIRA

PROSPECÇÃO DE SONDAS CROMOSSÔMICAS POR OLIGONUCLEOTÍDEOS
PARA APLICAÇÃO NOS EXAMES DE HIBRIDIZAÇÃO *IN SITU* FLUORESCENTE
(FISH) EM PACIENTES COM MIELOMA MÚLTIPLO

CURITIBA

2023

BRUNA DA SILVA OLIVEIRA

PROSPECÇÃO DE SONDAS CROMOSSÔMICAS POR OLIGONUCLEOTÍDEOS
PARA APLICAÇÃO NOS EXAMES DE HIBRIDIZAÇÃO *IN SITU* FLUORESCENTE
(FISH) EM PACIENTES COM MIELOMA MÚLTIPLO

Trabalho de Conclusão de Curso apresentado
como requisito parcial à obtenção do título de
Bacharel em Biomedicina, Setor de Ciências
Biológicas - Genética, Universidade Federal do
Paraná.

Orientador: Professor Dr. Daniel Pacheco Bruschi

CURITIBA

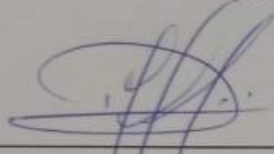
2023

TERMO DE APROVAÇÃO

BRUNA DA SILVA OLIVEIRA

PROSPECÇÃO DE SONDAS CROMOSSÔMICAS POR
OLIGONUCLEOTÍDEOS PARA APLICAÇÃO NOS EXAMES DE HIBRIDIZAÇÃO IN
SITU FLUORESCENTE (FISH) EM PACIENTES COM MIELOMA MÚLTIPLO

Trabalho de Conclusão de Curso apresentado no Setor de Ciências
Biológicas - Genética, da Universidade Federal do Paraná, como requisito parcial à
obtenção do título de Bacharel em Biomedicina.



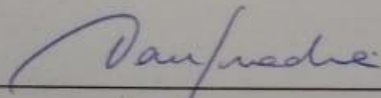
Prof. Dr. Daniel Pacheco Bruschi

Orientador – Departamento de Genética, UFPR



Dra. Carolina Mathias

Departamento de Genética, UFPR



Profa. Dra. Daniela Fiori Gradia

Departamento de Genética, UFPR

Curitiba, 06 de julho de 2023.

Dedico este trabalho à Silmara Pires de Oliveira, na tentativa de materializar a minha profunda admiração, amor e gratidão por tudo o que já fez por mim. Devo a você, tia, uma dívida de gratidão que nunca poderei realmente retribuir. Nunca foi fácil, mas você tornou tudo possível, obrigada!

AGRADECIMENTOS

Primeiramente, agradeço às três mulheres da minha vida: Silmara Pires de Oliveira, Ingrid da Silva Oliveira e Helena Oliveira da Costa Luz. Tia, irmã e filha, obrigada por serem meu tripé de sustentação, sem vocês não teria meios e nem motivos para chegar até aqui.

Agradeço também aos “Labcequers”, os amigos que fiz no Laboratório de Citogenética e Conservação Animal (LabCECA), onde fui muito bem recebida e acolhida desde o primeiro dia. Obrigada pela troca de ensinamentos e momentos compartilhados que ultrapassaram as barreiras físicas do “Labinho” e tornaram essa etapa acadêmica mais humana e positiva.

Ao Prof. Dr. Daniel Pacheco Bruschi, obrigada por me receber na equipe e pela sua dedicação em me orientar não apenas na pesquisa, mas também na minha jornada pessoal de crescimento. Ainda, agradeço pela paciência, empatia e por ter aceitado (e incentivado) o desafio de continuar como meu orientador durante o Mestrado no Programa de Pós-Graduação em Genética da UFPR.

Por fim, agradeço à Universidade Federal do Paraná, aos meus amigos e aos professores que tive ao longo da vida: vocês acreditaram, encorajaram e me ajudaram a traçar minha história até aqui. Obrigada.

"Don't forget to breathe."
(Chester Bennington, 2012)

RESUMO

O mieloma múltiplo (MM) é uma doença hematológica complexa e diversificada que afeta diversas pessoas. Seu diagnóstico preciso requer avaliações clínicas e métodos diagnósticos avançados devido às suas variadas manifestações e características. A identificação de anormalidades cromossômicas específicas visa realizar a estratificação de risco dos pacientes e fornecer informações importantes para o diagnóstico, prognóstico e monitoramento da doença. Atualmente, a metodologia de hibridização *in situ* fluorescente (FISH) é utilizada como exame padrão-ouro para essa detecção, mas, sua acessibilidade financeira ainda é um empecilho para que seja empregada em todos os pacientes com suspeita de MM, principalmente dentro do Sistema Único de Saúde. No entanto, a tecnologia OligoFISH surge como uma alternativa promissora, pois possui o potencial de se tornar aliada ao diagnóstico e possibilitar a nacionalização da tecnologia, superando as barreiras financeiras e identificando alterações cromossômicas específicas, como a deleção 17p, diretamente relacionada a um pior prognóstico em pacientes com mieloma múltiplo. Em nosso trabalho, utilizamos a análise de bancos de dados (iFISH4U) para prospectar conjuntos de sondas para a detecção da presença do cromossomo 17 bem como, a perda da região 17p.13.1. Nosso trabalho elegeu oito possíveis conjuntos de sondas de oligonucleotídeos com potencial para prosseguimento em experimentos de validação experimental. Esses passos iniciais são fundamentais para o desenvolvimento desta nova estratégia de construção de sondas cromossômicas para exames de FISH e , espera-se que, sua implantação possa, no futuro, ampliar o acesso a mais pacientes que possam se beneficiar de um diagnóstico preciso que contribuirá para um melhor gerenciamento da doença e abordagem terapêutica mais direcionada.

Palavras-chave: mieloma múltiplo, anormalidades cromossômicas, estratificação de risco, hibridização *in situ* fluorescente (FISH), OligoFISH.

ABSTRACT

Multiple myeloma (MM) is a complex and diverse hematological disease that affects many individuals. Its precise diagnosis requires clinical evaluations and advanced diagnostic methods due to its varied manifestations and characteristics. The identification of specific chromosomal abnormalities aims to stratify patients' risk and provide important information for diagnosis, prognosis, and disease monitoring. Currently, fluorescent in situ hybridization (FISH) methodology is used as the gold standard for this detection, but its financial accessibility remains a hindrance, especially within the Unified Health System. However, OligoFISH technology emerges as a promising alternative, as it has the potential to become a diagnostic ally and enable the nationalization of the technology, overcoming financial barriers and identifying specific chromosomal alterations, such as the 17p deletion, directly related to a worse prognosis in multiple myeloma patients. In our study, we used database analysis (iFISH4U) to prospect sets of probes for detecting the presence of chromosome 17, as well as the loss of the 17p.13.1 region. Our work identified eight potential sets of oligonucleotide probes for further experimental validation.

These initial steps are crucial for the development of this new strategy for constructing chromosomal probes for FISH exams, and it is hoped that its implementation may, in the future, broaden access to more patients who could benefit from accurate diagnosis, contributing to better disease management and more targeted therapeutic approaches.

Keywords: multiple myeloma, chromosomal abnormalities, risk stratification, fluorescence in situ hybridization (FISH), OligoFISH.

SUMÁRIO

1. INTRODUÇÃO	7
1.1. CONTEXTO E PROBLEMA	7
1.2. OBJETIVOS	7
1.2.1. Objetivo geral	7
1.2.2. Objetivos específicos.....	7
1.3. JUSTIFICATIVA	8
2. REVISÃO TEÓRICO-EMPÍRICA	9
2.1. MIELOMA MÚLTIPLO.....	9
2.1.1. ALTERAÇÕES GENÉTICAS EM PACIENTES COM MM.....	10
2.2. OLIGO-FISH	11
3. METODOLOGIA	13
3.2. PROSPECÇÃO DAS SEQUÊNCIAS DE OLIGONUCLEOTÍDEOS.....	14
4. APRESENTAÇÃO DOS RESULTADOS	17
4.1. SONDA PARA VERIFICAR A DELEÇÃO 17p13.1	17
4.2. SONDA PARA VERIFICAR A PRESENÇA DO CROMOSSOMO 17	22
5. CONSIDERAÇÕES FINAIS	26
REFERÊNCIAS.....	28

1. INTRODUÇÃO

1.1. CONTEXTO E PROBLEMA

O Mieloma Múltiplo é uma forma de câncer incurável que afeta os plasmócitos na medula óssea, responsáveis pela produção de anticorpos e se caracteriza pelo crescimento desordenado das células. É uma doença multifatorial, progressiva e os sintomas podem ser inespecíficos (dor óssea, fadiga, anemia, infecções recorrentes e problemas renais), o que contribui para a dificuldade no diagnóstico e estadiamento da doença, já que pode ser facilmente confundida com outras condições médicas.

A fim de determinar a presença e fase do MM, indica-se o uso de testes citomoleculares, como a técnica de hibridização *in situ* fluorescente (FISH), para detectar alterações genéticas específicas, prognóstico e, consequentemente, proporcionar uma terapia mais eficiente para o paciente. Essa técnica é considerada padrão-ouro para identificar alterações cromossômicas nos clones malignos presentes nas neoplasias onco-hematológicas, porém, tal tecnologia não é implementada no Sistema Único de Saúde (SUS) para cobrir as suspeitas diagnósticas devido ao alto custo de implementação.

1.2. OBJETIVOS

1.2.1. Objetivo geral

Investigar regiões cromossômicas para detecção de alterações genômicas de alto risco para pacientes com mieloma múltiplo para construção de sondas baseadas em oligonucleotídeos para uso em exames de FISH.

1.2.2. Objetivos específicos

- Identificar e selecionar regiões cromossômicas relevantes associadas ao mieloma múltiplo em bancos de dados genômicos.
- Verificar o potencial da região 17p13.1 em gerar sequência de DNA propícias ao desenho de sondas de oligonucleotídeos para experimentos de OligoFISH.

- Avaliar se as sequências de oligonucleotídeos potenciais se adequam a critérios mínimos para seu uso em experimentos de OligoFISH.
- Contribuir para o desenvolvimento de sondas nacionais para o diagnóstico da deleção 17p em pacientes com mieloma múltiplo.

1.3. JUSTIFICATIVA

O mieloma múltiplo é uma doença complexa e multifatorial caracterizada pela proliferação anormal e descontrolado dos plasmócitos, considerada progressiva e, infelizmente, até o momento, não possui cura. No entanto, embora não seja possível eliminar o MM completamente, existem várias opções de tratamento disponíveis que visam garantir uma melhor sobrevida ao controlar a doença, aliviar os sintomas e melhorar a qualidade de vida do paciente - a escolha e o uso adequado da estratégia terapêutica são essenciais para isso.

A determinação do tratamento mais adequado é personalizada e baseada em vários fatores, incluindo o diagnóstico preciso da doença, seu estágio e a estratificação de risco em que ela se encontra, que é dividido em baixo, médio e alto risco. No estágio inicial ou em pacientes de baixo risco, o tratamento envolve uma abordagem de vigilância ativa, para monitorar a progressão da doença e intervir somente quando necessário. Já nos casos mais avançados ou pacientes de alto risco, são indicados tratamentos mais agressivos, como quimioterapia, imunoterapia, terapia alvo e, em alguns casos, transplante de medula óssea.

Dentro do espectro de alto risco, encontra-se o diagnóstico da doença associado à presença da alteração cromossômica caracterizada pela deleção 17p, uma vez que está diretamente associada ao gene *TP53*, localizado nessa região do cromossomo. O gene *TP53* desempenha um papel crucial na supressão de tumores e na regulação do ciclo celular, tornando-se um dos principais guardiões da integridade genômica. Portanto, a deleção 17p, que envolve a perda do gene *TP53*, está associada a um prognóstico desfavorável e a uma resposta menos favorável aos tratamentos convencionalmente utilizados. Para realizar o diagnóstico que identifica essa deleção parcial no cromossomo 17, é essencial o uso de técnicas específicas de diagnóstico molecular, como a hibridização *in situ* por fluorescência (FISH), que permite visualizar e mapear as alterações cromossômicas ao utilizar sondas de DNA

marcadas com fluoróforos que se ligam especificamente a sequências genéticas de interesse.

Atualmente, um dos desafios enfrentados pelo Sistema Único de Saúde (SUS) frente ao diagnóstico e, conseqüentemente, ao tratamento adequado de pacientes com MM, está relacionado à falta de acesso a sondas FISH. As sondas comerciais existentes possuem custo muito elevado, limitando a capacidade de cobrir todas as suspeitas, diagnosticar e estratificar os pacientes de forma adequada.

Reconhecendo a importância do diagnóstico preciso, especificamente o que envolve a deleção 17p, o presente trabalho visa contribuir para resolver o problema da escassez de sondas nacionais, apresentando novas maneiras de design de sondas utilizando a tecnologia OligoFISH, sendo mais flexível e acessível em comparação com as sondas tradicionais. Ao possibilitar a produção local, haverá queda no custo de acesso à tecnologia e conseqüente melhora na detecção da deleção 17p, contribuindo para a personalização do tratamento, melhores resultados clínicos e qualidade de vida para os pacientes, mesmo dentro do contexto do SUS.

2. REVISÃO TEÓRICO-EMPÍRICA

2.1. MIELOMA MÚLTIPLO

O Mieloma Múltiplo (MM) é a segunda neoplasia hematológica mais comum no mundo e é classificado como uma doença complexa, por causar sinais e sintomas locais, sistêmicos e muitas vezes inespecíficos (van de Donk et al., 2021). Considerando as principais manifestações clínicas da doença, o MM também pode ser descrito como síndrome “CRAB”, acrônimo do inglês para hipercalcemia (Calcium elevation), insuficiência renal (Renal failure), anemia (Anemia) e doença óssea (Bone disease) (CONITEC 2022).

Essas manifestações da doença são desencadeadas pelo acúmulo de plasmócitos (células de origem de linfócitos B) anômalos na medula óssea (Albagoush et al., 2023). Os plasmócitos são células do sistema imunológico responsáveis pela produção de anticorpos no organismo e, quando em malignidade como no MM, proliferam-se descontroladamente e passam a secretar suas proteínas monoclonais (também conhecidas como proteína M ou paraproteínas) no sangue e na urina. Elas,

então, acumulam-se nos diversos tecidos do organismo resultando nos sintomas e complicações característicos da doença (Cowan et al., 2022).

De acordo com a American Cancer Society, existem alguns fatores de risco conhecidos para o mieloma múltiplo, como idade, sexo, etnia, histórico familiar, obesidade e exposição à radiação. Como é uma doença multifatorial, sua causa exata ainda não é totalmente compreendida, mas aponta para uma forte interação entre fatores genéticos e ambientais determinantes para seu surgimento e progresso (Dominik et al., 2007).

O MM é uma doença incurável, mas possui tratamento multidisciplinar com objetivo de controlar a progressão da doença, minimizar sintomas, promover qualidade de vida e prolongar a sobrevida dos pacientes. Existem diferentes subtipos de mieloma múltiplo e a estratificação de risco é um processo usado para determinar o prognóstico e o tratamento mais adequado para cada paciente (Rajkumar, 2022). A linha de escolha depende do estágio e agressividade da doença e envolve o uso de medicamentos quimioterápicos, radioterapia, imunomoduladores, inibidores de proteassoma e, em alguns casos, transplante de medula óssea (Soekojo, 2022).

2.1.1. ALTERAÇÕES GENÉTICAS EM PACIENTES COM MM

Grande variedade de alterações cromossômicas numéricas e estruturais são observadas em pacientes diagnosticados com MM. Autores têm buscado distingui-las primárias e secundárias (KUMAR & RAJKUMAR, 2018; RAJKUMAR, 2019) como forma de entender melhor a heterogeneidade de cariótipos acometidos por portadores de MM.

As alterações primárias são características da origem clonal da neoplasia e ocorrem de maneira isolada no clone neoplásico. Assim, os eventos genéticos que são observados em pacientes com Gamopatia Monoclonal de Significância Indeterminada (MGUS), estágio prévio ao estabelecimento do MM, são aquelas envolvidos no estabelecimento da população clonal e desenvolvimento do tumor (FONSECA et al., 2002), englobam as translocações envolvendo o locus da cadeia pesada de imunoglobulina (IgH) no cromossomo 14q32, assim como as trissomias, principalmente envolvendo pares ímpares que constituem o cariótipo (por exemplo, homólogos dos pares 5, 7, 9, 11, 13, 15, 17) (KUMAR & RAJKUMAR, 2018; RAJKUMAR, 2019). As translocações envolvendo a região 14q32 que justapõe o gene

da IgH com oncogenes (KUEHL & BERGSAGEL, 2002) tem o potencial de gerar uma transcrição contínua e por isso, resultam em uma produção anormal de imunoglobulinas monoclonais disfuncionais.

Desta forma, algumas translocações envolvendo a região 14q32 são com frequência observada nos pacientes com MM e marcadores para diagnóstico e prognóstico de pacientes. Translocações envolvendo o cromossomo 14 incluem a t(4;14); t(6;14), t(8;14), t(11;14), t(14;16) e a t(14;20) (KUMAR & RAJKUMAR, 2018). destas, a t(4;14) é tradicionalmente considerada de alto risco (KEATS, et al., 2003) enquanto a t(11;14) é considerada de risco intermediário ao paciente (FONSECA et al., 2004).

Já os eventos que são detectados ao longo da progressão do MM, ditas como secundárias, envolvem principalmente ganho e/ou perda de regiões genômicas, como ganho(1q), del(1p), del(17p), del(13q), além de mutações *RAS*, que podem levar a vias de sinalização celular desreguladas e contribui para o crescimento descontrolado, e translocações envolvendo o gene *MYC*, localizado no cromossomo 8 (KUMAR & RAJKUMAR, 2018).

Portanto, o reconhecimento do evento molecular de origem ou envolvido na progressão do MM nos pacientes é fundamental para as estratégias terapêuticas, como proposto pelo recente Relatório (n o.116/2021) da CONITEC, visando um resultado diagnóstico mais preciso principalmente para a translocação t(4;14) e a deleção em del(17p).

2.2. OLIGO-FISH

Hibridação “*in situ*” fluorescente (FISH) é o método no qual sequências ou regiões genômicas (DNA-alvo), seja cópia única ou repetitivas, podem ser detectadas por meio de uma reação de hibridação entre a molécula de DNA-alvo e sequências de DNA complementares marcadas com moléculas de natureza fluorescente, denominadas de sondas de DNA.

Virtualmente, todas as sondas de DNA que apresentem pelo menos 70% de homologia com as sequência de DNA-alvo tem potencial para ser aplicada em um experimento de FISH. Assim, é possível construirmos sondas de DNA a partir de qualquer região genômica conhecida. A molécula de DNA-alvo pode ser detectada tanto em cromossomos metafásicos quanto em núcleos interfásicos.

Atualmente, a metodologia de FISH é empregada como exame padrão-ouro para detecção de alterações cromossômicas de clones malignos em neoplasias onco-hematológicas (Hu et al., 2014), na detecção de aneuploidias e ou de rearranjos específicos de algumas doenças constitucionais (síndrome de Prader-Willi, Síndrome de Angelman, por exemplo) (Huber et al., 2018). Para muitas destas doenças, sondas comerciais já estão disponíveis para aplicação em diagnóstico. Outras delas estão disponíveis para pesquisa e são aplicadas para o diagnóstico após períodos de validação laboratorial “in house”.

A hibridização in situ de fluorescência de oligonucleotídeos (Oligo-FISH) é uma técnica emergente que utiliza oligonucleotídeos de cópia única (oligos) como sondas FISH. Neste método, as sondas são sinteticamente produzidas utilizando bancos de dados genômicos e utilizados como sequências de oligonucleotídeos nos experimentos de hibridação. O desenho das sondas pode ser feito a partir de diferentes ferramentas computacionais que combinam *scripts* das mais diversificadas naturezas. Entre as mais conhecidas ferramentas metodológicas para prospecção de sondas de oligonucleotídeos, podemos relacionar OligoArray, PROBER, OligoMiner, iFISH, PaintSHOP e uma crescente ascensão de novos programas e *scripts* têm sido desenvolvidos recentemente (para revisão, ver Liu e Zhang 2021)

Essa ferramenta possibilita visualizar e mapear a localização de sequências específicas de DNA, para detectar variações cromossômicas e decifrar estruturas genômicas 3D (Zhang, 2021).

Ao construir sondas cromossômicas, como sondas de hibridização fluorescente *in situ* (FISH), três características importantes devem ser consideradas: tamanho (size), centralidade (centrality) e homogeneidade (homogeneity). Tais características estão relacionadas à seleção e preparação mais adequada das sondas para obtenção de resultados eficazes e confiáveis. Estes três parâmetros são baseados em características dos próprios oligonucleotídeos prospectados, como devido ao tamanho da sequência de oligonucleotídeo, propriedades termodinâmicas e sua especificidade de sequências (Dieffenbach et al., 1993; Liu e Zhang 2021).

O tamanho dos oligonucleotídeos a serem empregados como sondas deve ser controlado para evitar excessiva instabilidade, além de garantir a adequada stringência durante as etapas de lavagem no decorrer do experimento. (Dieffenbach et al., 1993). Já o conteúdo de GC, por sua vez, interfere no rigor experimental, já que um dos fatores importantes para ser levado em consideração é a temperatura de

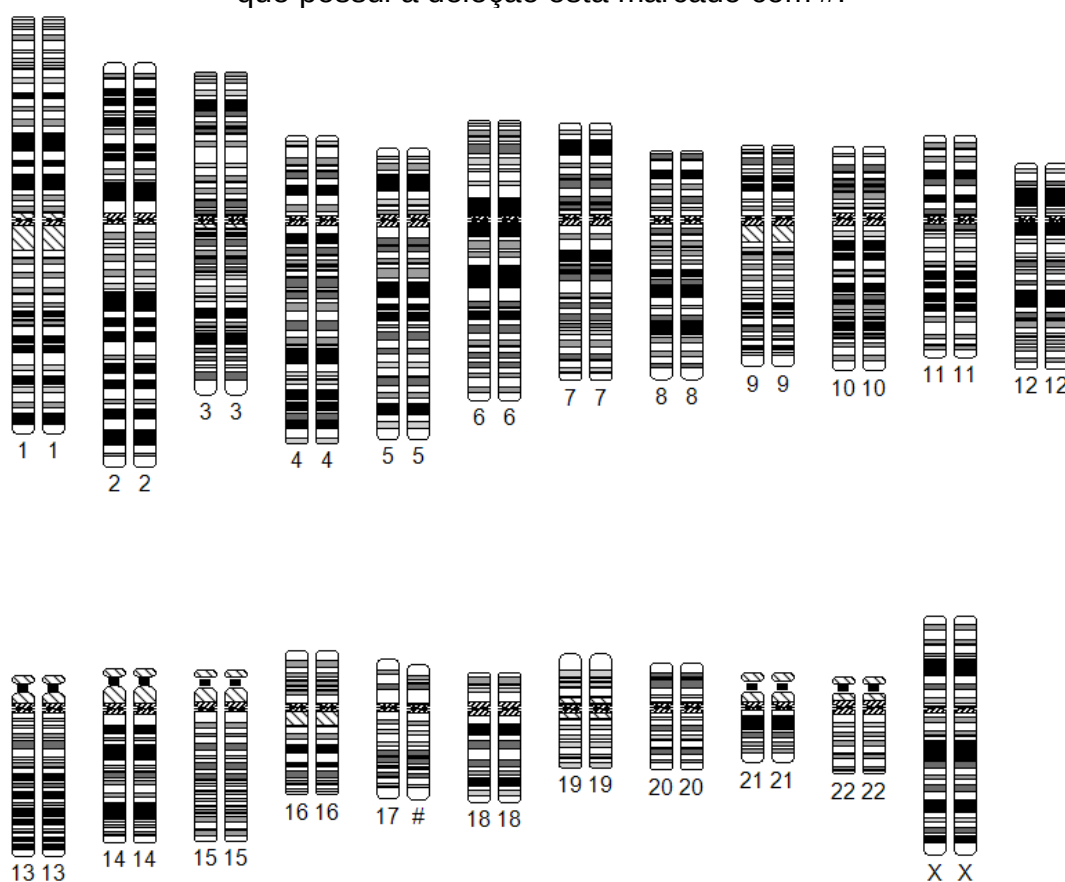
melting (processo de separação das duas cadeias de DNA, normalmente emparelhadas, pela elevação de temperatura) (T_M) dos oligonucleotídeos sintetizados, além de sofrer interferência da concentração de sódio e formamida ao qual a reação está sendo realizada (Dieffenbach et al., 1993). Por fim, a especificidade da sonda ainda é um fator limitante pensando em sequência de tamanho reduzido e de quanto de polimorfismos podem ser suportados para que a sonda possa rigorosamente se ligar à região genômica alvo e não em um locus inespecífico (Liu e Zhang, 2021).

3. METODOLOGIA

3.1. DEFINIÇÃO DAS REGIÕES GENÔMICAS

A escolha da região para a prospecção da sonda baseou-se em uma das alterações cromossômicas de mieloma múltiplo com valor prognóstico e com estratificação de risco alto (desfavorável), no caso, a deleção em (17p13.1), conforme (KUMAR & RAJKUMAR, 2018; RAJKUMAR, 2019) (FIGURA 1).

FIGURA 1 - Representação de um cariótipo 46,XX,del(17)(p13.1). O cromossomo que possui a deleção está marcado com #.



FONTE: CyDAS Online Analysis Site.

A região genômica 17p13.1 é caracterizada por conter vários genes e elementos regulatórios necessários para o funcionamento celular, dentre eles, o gene *TP53* (chr17:7571720-7590868) que é classificado como supressor de tumor. Tal gene é responsável por codificar a proteína p53 que atua no ciclo celular controlando os eventos de proliferação celular, apoptose e reparo do DNA, fatores que implicam diretamente no prognóstico desfavorável ao influenciar na resistência ao tratamento, progressão da doença e probabilidade de recidiva.

Adjacente à essa localização, encontra-se o *WRAP53* (chr17:7589389-7606820), gene que codifica para um RNA não-codificante que está envolvido na alongação dos telômeros, manutenção dos corpúsculos de Cajal e também na estabilização e transporte do mRNA de *TP53* para o citoplasma, possibilitando a tradução da proteína p53 e também foi incluída nesta prospecção de oligonucleotídeos.

No design das sondas, prospectou-se oligonucleotídeos de uma outra região do cromossomo para funcionar como região marcadora da presença do cromossomo 17 na amostra a ser analisada. Para isso, a região de escolha foi um intervalo da região pericentromérica do cromossomo 17, especificamente na região 17p11.1 (FIGURA 2).

FIGURA 2: Ideograma do cromossomo 17. Localização dos genes *TP53* e *WRAP53* na região 17p13.1 marcada em vermelho e região pericentromérica em verde.



FONTE: UCSC Genome Browser.

3.2. PROSPECÇÃO DAS SEQUÊNCIAS DE OLIGONUCLEOTÍDEOS

O uso de ferramentas de bioinformática é fundamental para a prospecção de sequências de oligonucleotídeos. Existem várias ferramentas disponíveis que auxiliam na análise e seleção de oligonucleotídeos, como o Primer3, OligoAnalyzer, iFISH4u, entre outras. Tais ferramentas podem ser utilizadas para realizar análises de similaridade, avaliar parâmetros termodinâmicos e verificar a especificidade das sequências-alvo desejadas.

A partir da análise e definição das regiões de interesse descritas em 3.1, utilizou-se a plataforma iFISH4u para desenhar sondas de interesse. A principal motivação para escolha foi devido a essa plataforma ser um recurso disponível publicamente para o estudo da arquitetura do genoma por meio da técnica de FISH e disponibiliza um *software* para projetar sondas de DNA específicas a partir de um banco de dados de sequências genômicas. Através da plataforma é possível utilizar as informações dos seguintes bancos de dados: iFISH-40mer (Gelali et al., 2019), OligoMiner-balanced, OligoMiner-coverage ou OligoMiner-stringent (Beliveau, et al., 2018) para a prospecção de sondas, cada um com suas particularidades.

No iFISH4u, o banco de dados utilizado como padrão (iFISH-40mer) foi criado pela própria plataforma, porém, ele é baseado em sondas de 40 nucleotídeos e composto por sondas já pré-determinadas, o que acaba por limitar a detecção de certas alterações cromossômicas, devido à complexidade estrutural das regiões de interesse e também restringir a personalização e adaptação das sondas para fins específicos, respectivamente.

A partir disso, analisou-se os três banco de dados disponibilizados pela OligoMiner (OM), um por vez, para direcionar a produção das sondas e então avaliar qual seria a mais adequada para atingir os objetivos do projeto. O OM-balanced é um banco de dados projetado para fornecer uma coleção equilibrada de oligonucleotídeos visando oferecer uma visão abrangente das sequências. O OM-coverage se concentra na cobertura de regiões específicas do genoma para fornecer uma ampla cobertura de sequências em áreas de interesse específicas. Já o OM-stringent é uma versão mais rigorosa, visando uma maior especificidade e considera fatores como potenciais hibridizações cruzadas ou regiões repetitivas no genoma (Beliveau, et al., 2018).

Para avaliar deleção de 17p13.1, considerou-se como referência de valores extremos para a sonda o início da coordenada gênica de *TP53* (*Start position*) e o fim de *WRAP53* (*End position*), para abranger toda a região de interesse (FIGURA 3). Já para avaliar a presença do cromossomo 17 na amostra, utilizou-se as coordenadas cromossômicas de uma região pericentromérica, rica em regiões satélites (Shannon, 2017), delimitada pelas coordenadas chr17:22200000-22900000.

FIGURA 3 - etapa de estabelecer os seguintes critérios para prospecção de sondas: o banco de dados a ser utilizado (Database), o cromossomo de interesse (Chromosome) e as coordenadas mínimas de início (Start position) e fim (End position) da região-alvo.

The image shows a web interface titled "Where" with four input fields for searching probes. The fields are labeled "Database", "Chromosome", "Start position", and "End position". The "Database" field has a dropdown menu with "OligoMiner-ba" selected. The "Chromosome" field has a dropdown menu with "chr17" selected. The "Start position" field has a text input with "7571720" and a small up/down arrow icon. The "End position" field has a text input with "76068020" and a small up/down arrow icon.

FONTE: Site iFISH4U (<http://ifish4u.com/>)

A partir dos critérios previamente estabelecidos, é possível organizar a ordem de prioridade (em relação à centralidade, tamanho e homogeneidade) para

classificação das sondas, caso haja mais de uma possibilidade viável como retorno (FIGURA 4).

FIGURA 4 - onfigurações avançadas para estabelecer a ordem de prioridade para classificação das sondas desenhadas, considerando o tamanho, centralidade e homogeneidade, respectivamente.

Advanced settings

Feature	Use	Size (minimize)	Centrality (maximize)	Homogeneity (maximize)
First ¹	Filter probe candidates.	☒	☐	☐
Second	Rank probe candidates.	☐	☐	☒
Third	<i>Not used.</i>	☐	☒	☐

¹ Candidate probes are selected based on the first feature, in the range `best_f1_value ± (best_f1_value * threshold)`.

FONTE: Imagem retirada do site iFISH4U (<http://ifish4u.com/>).

Para se obter o valor da cobertura em kb, é necessário considerar o tamanho do genoma-alvo e o número de sondas utilizadas. A fórmula geral utilizada para calcular a cobertura foi: Cobertura (em kb) = (Tamanho do genoma-alvo (tamanho) / Número de sondas (noligo) * 1000.

4. APRESENTAÇÃO DOS RESULTADOS

4.1. SONDA PARA VERIFICAR A DELEÇÃO 17p13.1

A partir da inserção dos dados da região de interesse na plataforma iFISH4u, obteve-se 4365 possibilidades de sondas, das quais, apenas cinco conjuntos (set) abrangiam toda a região de interesse pré-estabelecida. As coordenadas e o banco de dados usado como referência encontram-se na Tabela 1.

Tabela 1. Conjunto de sondas numeradas de acordo com a ordem em que foram sendo geradas.

Número da	Posição	Posição	Tamanh	OligoMiner	nOligos	Cobertura
-----------	---------	---------	--------	------------	---------	-----------

Sonda	de início	de término	o (pb)			(em kb)
Conjunto 1	7560299	7607786	47487	coverage	279	170204
Conjunto 2	7560299	7607754	47455	stringent	248	191350
Conjunto 3	7560327	7607772	47445	coverage	279	170053
Conjunto 4	7571194	7606760	35566	stringent	206	172650
Conjunto 5	7570764	7607794	37030	stringent	126	293888

As imagens da coluna da direita (FIGURA 5) representam um histograma de densidade da distância entre os oligonucleotídeos consecutivos de cada conjunto gerado. A distância no eixo x refere-se ao número de nucleotídeos entre dois oligonucleotídeos consecutivos no conjunto e a frequência das diferentes distâncias é registrada e representada no eixo y do gráfico. A presença de um pico proeminente em determinada distância, indica que os oligonucleotídeos estão bem posicionados e espaçados de forma regular, o que indica boa especificidade e capacidade de se ligar com precisão à sequência alvo.

Assim, os cinco conjuntos são homogêneos, em termos de densidade de oligonucleotídeos, e a alta densidade nas faixas de tamanho determinadas indicam que os oligonucleotídeos teriam boa eficiência de hibridização.

Já nos gráficos da esquerda (FIGURA 5), está representada a identificação de cada oligonucleotídeo (pontos vermelhos) no eixo x e suas coordenadas genômicas no eixo y, para cada um dos cinco conjuntos. Essa representação permite visualizar a distribuição espacial dos oligonucleotídeos ao longo da região genômica e a linha preta representa a distância entre eles. Caso tivessem a mesma distância entre si, todos os pontos vermelhos estariam sobre a linha preta, resultando em uma distribuição homogênea dos oligonucleotídeos ao longo da região genômica de interesse.

A homogeneidade na distribuição dos oligonucleotídeos é um ponto importantíssimo a ser considerado porque permite uma cobertura uniforme da região genômica em análise, ou seja, cada parte da região genômica terá a mesma probabilidade de ser alvo de um oligonucleotídeo.

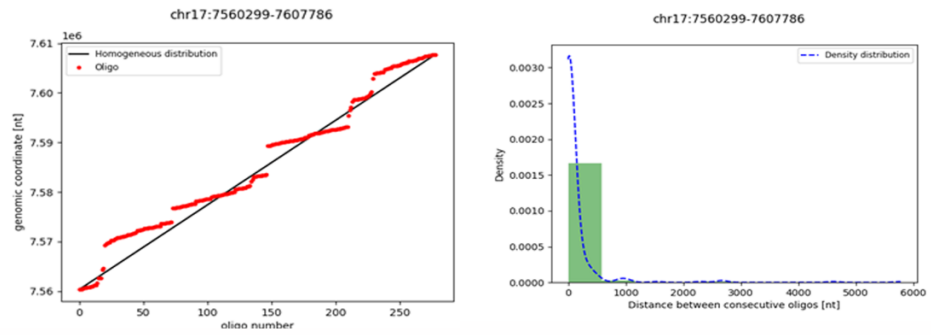
Consequentemente, a falta de homogeneidade pode resultar em lacunas de cobertura, o que torna o conjunto 3, dentre as opções, o mais inviável para prosseguir

no estudo de prospecção de sondas, por ser o que mais apresentou afastamento dos pontos da linha central.

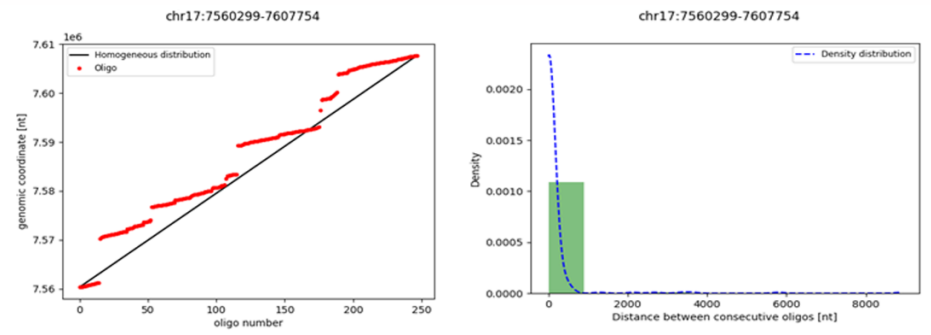
Ainda, para cada um dos cinco conjuntos, a região genômica abrangida pelas sondas (FIGURA 6) também influencia no seu sucesso. A falta de oligonucleotídeos em algumas regiões indica uma baixa densidade ou até mesmo a ausência de mapeamento de oligonucleotídeos na região genômica.

FIGURA 5: Gráficos de homogeneidade (à esquerda) e histograma de densidade (à direita) para cada um dos cinco conjuntos analisados.

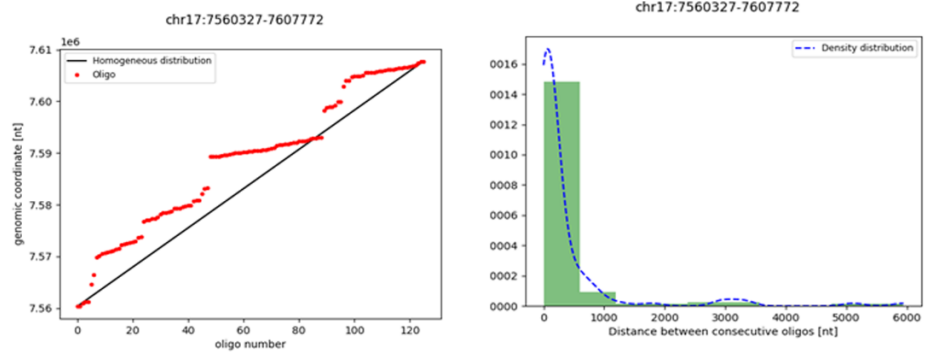
Conjunto 1



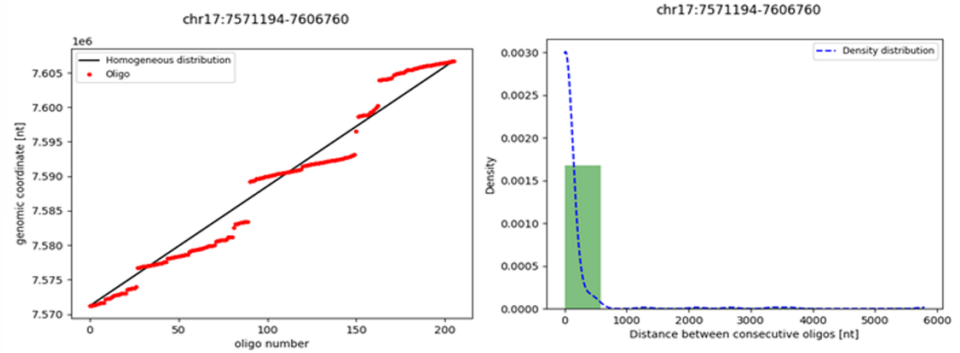
Conjunto 2



Conjunto 3



Conjunto 4



Conjunto 5

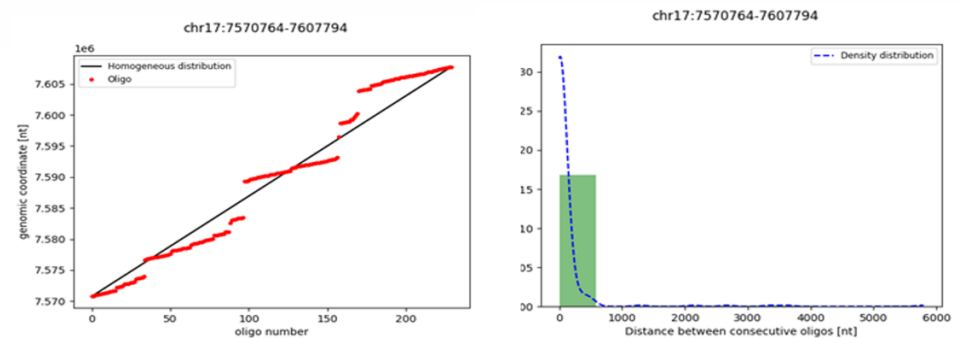
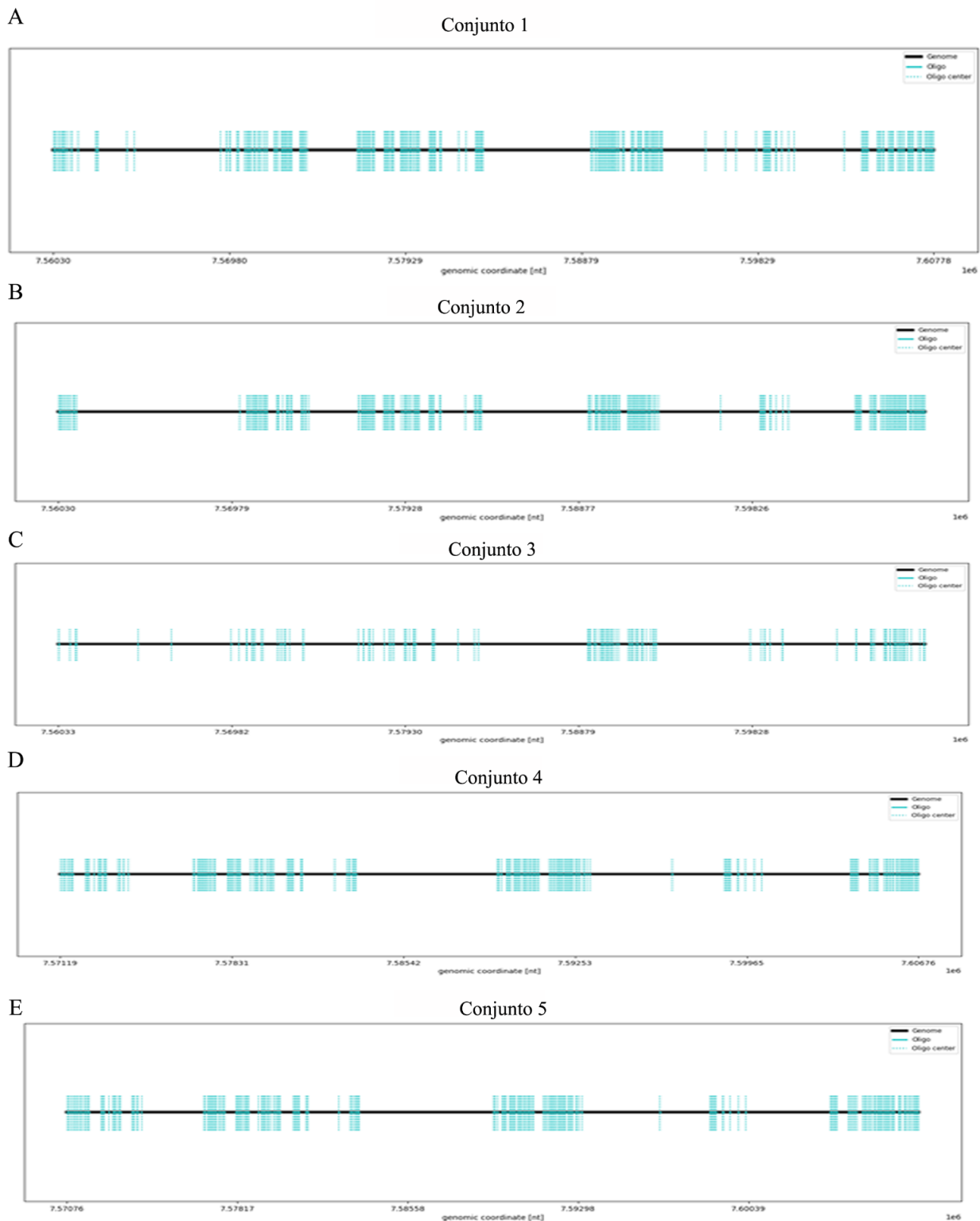


FIGURA 6: Gráficos da região genômica abrangida pela sonda representada em preto, enquanto a localização de cada oligonucleotídeo é mostrada em azul e as linhas verticais pontilhadas indicam o centro de cada oligonucleotídeo.



A partir da análise dos resultados, nós optamos por realizar uma análise comparativa entre os parâmetros que elencamos como fundamentais para a escolha

do melhor conjunto de sondas para esta região genômica. Para isso, atribuiu-se arbitrariamente a pontuação de 0 a 4, onde 4 representa a melhor qualidade ou adequação e 0 a pior, para ranquear os cinco conjuntos (Tabela 2), levando em consideração o gráfico de homogeneidade, a região genômica abrangida pela sonda e cobertura em kb que cada conjunto possui. O histograma de densidade não foi utilizado nesta etapa devido à similaridade de resultados dentre os cinco conjuntos; a pontuação final é calculada através da soma das notas atribuídas a cada critério.

Tabela 2. Análise dos critérios estabelecidos para o ranqueamento dos conjuntos de sondas baseados em parâmetros de homogeneidade, abrangência e cobertura das sondas.

Conjunto	Homogeneidade	Abrangência	Cobertura	Resultado
1	4	4	1	9
2	1	0	3	4
3	0	1	0	1
4	2	2	2	6
5	3	3	4	10

Assim, o conjunto que apresenta, até então, os melhores resultados e será o primeiro a iniciar a validação prática para identificar a deleção 17p13.1 é o Conjunto 5 e, a depender dos resultados, os outros conjuntos serão testados de acordo com a ordem classificatória que se encontra na Tabela 3.

Tabela 3. Posição classificatória dos conjuntos de acordo com os critérios de avaliação estabelecidos em 4.1.

Classificação	Número da sonda
1	Conjunto 5
2	Conjunto 1
3	Conjunto 4
4	Conjunto 2
5	Conjunto 3

4.2. SONDA PARA VERIFICAR A PRESENÇA DO CROMOSSOMO 17

A partir da inserção dos dados da região de interesse na plataforma iFISH4u, obteve-se 3195 possibilidades de sondas, porém, apenas três conjuntos (set) abrangiam toda a região pré-estabelecida. As coordenadas e o banco de dados usado como referência encontram-se na Tabela 4.

Tabela 4. Conjunto de sondas numeradas de acordo com a ordem em que foram sendo geradas.

Número da Sonda	Posição de início	Posição de término	Tamanho (pb)	OligoMiner	nOligos	Cobertura (em kb)
Conjunto 1	22202943	22203434	491	stringent	7	70142
Conjunto 2	22203069	22204995	1926	stringent	22	87545
Conjunto 3	22203069	22216712	13643	stringent	57	239350

Assim como para avaliar os conjuntos referentes à deleção de 17p13.1, utilizou-se os gráficos de homogeneidade e histograma de densidade (FIGURA 7) e a região genômica abrangida pelas sondas (FIGURA 8) de cada conjunto a fim de validar a presença do cromossomo 17 nos testes.

Por se tratar de uma região com sequências repetitivas de DNA, observa-se nos histogramas dos três conjuntos, mais de um pico de densidade, caracterizando uma distribuição multimodal.

FIGURA 7: Gráficos de homogeneidade (à esquerda) e histograma de densidade (à direita) para cada um dos cinco conjuntos analisados.

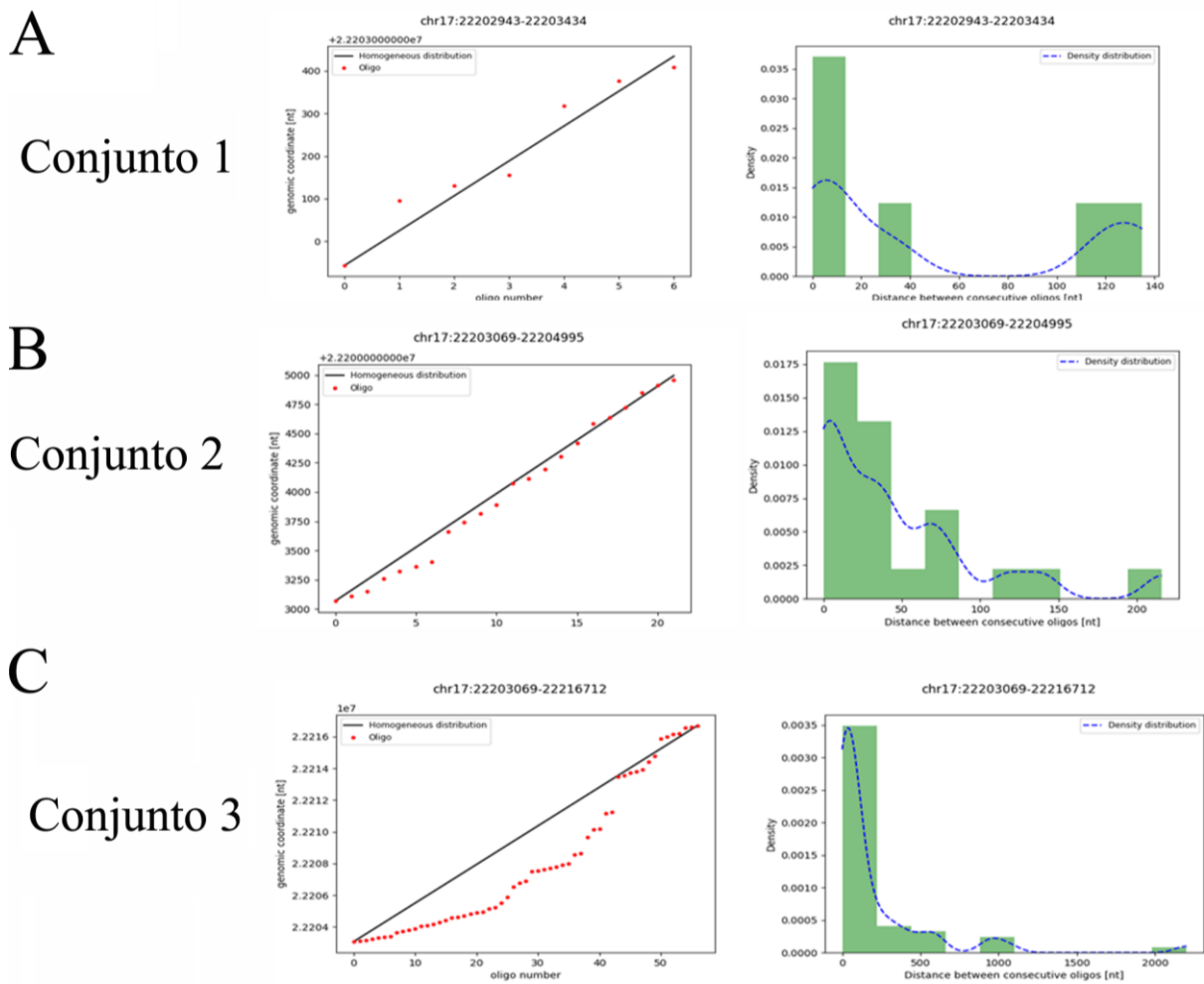
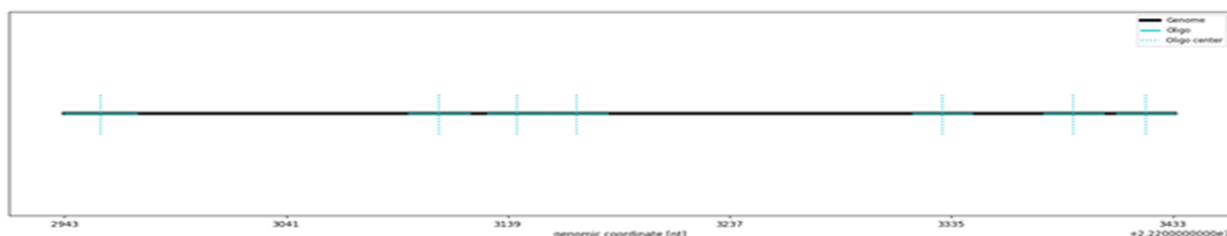


FIGURA 8: Gráficos da região genômica abrangida pela sonda representada em preto, enquanto a localização de cada oligonucleotídeo é mostrada em azul e as linhas verticais pontilhadas indicam o centro de cada oligonucleotídeo.

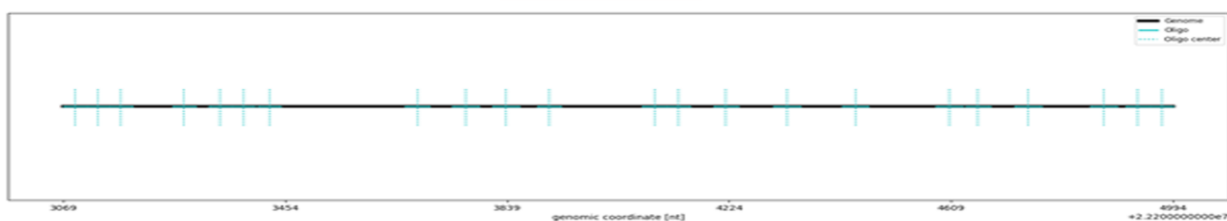
A

Conjunto 1



B

Conjunto 2



C

Conjunto 3



A partir da análise dos resultados, como só houve 3 conjuntos passíveis de ser utilizados, a pontuação atribuída foi de 0 a 2, onde 2 representa a melhor qualidade ou adequação e 0 a pior, para ranqueá-los (Tabela 5), levando em consideração o gráfico de homogeneidade, a região genômica abrangida pela sonda e cobertura em kb que cada conjunto possui.

O histograma de densidade não foi utilizado nesta etapa devido à característica da região-alvo, pois ao se tratar de regiões satélite, espera-se essa variação, independente das características da sonda em si; a pontuação final é calculada através da soma das notas atribuídas a cada critério.

Tabela 5. Análise dos critérios estabelecidos para o ranqueamento dos conjuntos de sondas baseados em parâmetros de homogeneidade, abrangência e cobertura das sondas.

Conjunto	Homogeneidade	Abrangência	Cobertura	Resultado
1	1	1	0	2
2	2	2	1	5
3	0	0	2	2

Assim, o conjunto que apresenta, até então, os melhores resultados e será o primeiro a iniciar a validação prática para confirmar a presença do cromossomo 17 é o Conjunto 2 e, a depender dos resultados, os outros conjuntos serão testados de acordo com a ordem classificatória que se encontra na Tabela 6. Como houve empate entre dois conjuntos, a ordenação foi estabelecida priorizando o que obteve menor pontuação zero dentre os critérios avaliados.

Tabela 6. Posição classificatória dos conjuntos de acordo com os critérios de avaliação estabelecidos em 4.2.

Classificação	Número da sonda
1	Conjunto 2
2	Conjunto 1
3	Conjunto 3

5. CONSIDERAÇÕES FINAIS

O projeto buscou investigar regiões cromossômicas com o objetivo de detectar alterações genômicas de alto risco em pacientes com mieloma múltiplo, primeiramente focando na deleção 17p13.1, visando a construção de sondas baseadas em oligonucleotídeos para utilização em exames de FISH com custo reduzido.

Ao verificar o potencial da região 17p13.1 em gerar sequências de DNA adequadas ao desenho de sondas de oligonucleotídeos, confirmou-se que essa região apresenta as características necessárias para a produção das sondas desejadas, como reduzido conteúdo de DNA repetitivo. Nossos dados sugerem que a aplicação dos conjuntos de sondas para a detecção da região 17p13.1 podem satisfatoriamente detectar esta região cromossômica em substituição ao uso das sondas comerciais baseadas em BACs (Bacterial Artificial Chromosome). Isso porque as atuais sondas baseadas em BAC apresentam uma cobertura de 161000 kb desta região enquanto nosso conjunto 5 de sondas consegue cobrir uma região de 293888 kb.

Na prospecção de marcadores utilizada aqui, incluímos a detecção de marcadores para a região pericentromérica de 17p para servir como controle positivo experimental da presença do cromossomo 17 em experimentos de FISH interfásico. Em nossa prospecção, a escolha dos nucleotídeos foi baseada em sequências cópia única, ao invés de marcadores de sequências repetitivas como os aplicados tradicionalmente nas sondas comerciais. Isso se deve ao fato do iFISH ser um banco de dados que evita a inclusão de sondas de DNA repetitivo, já que essas sequências poderiam gerar anelamento inespecífico em experimentos de FISH.

Apesar de alguns avanços na escolha das sondas para aplicação em experimentos de OligoFISH, ainda é necessário alguns ajustes para lapidar as sondas aqui descritas, como (i) avaliação do nível de polimorfismo nas regiões de sondas avaliadas, (ii) verificação sobre cobertura das regiões, especialmente nas regiões que são *hotspot* para quebras envolvendo esta região de 17p e (iii) avaliar, em experimentos de FISH, se a densidade dos marcadores é capaz de gerar um sinal de hibridação compatível com detecção clara e precisa em microscopia de fluorescência. O curto prazo de desenvolvimento do projeto associado à falta de conhecimento prévio das ferramentas de bioinformática necessárias, além da própria complexidade

da doença combinada à análise das regiões-alvo são alguns dos limitantes deste trabalho.

Por fim, apresentamos os passos iniciais de um trabalho de desenvolvimento tecnológico que poderá permitir a nacionalização da tecnologia de diagnóstico diminuindo seu custo e, ampliando, assim, o acesso ao exame de FISH para uma parcela maior de pacientes diagnosticados com doenças onco hematológicas, em especial, no Mieloma múltiplo. Isso não viabilizará apenas o diagnóstico e a estratificação de risco no contexto do mieloma múltiplo, mas também permitirá estender o uso dessa tecnologia para identificação de outras alterações cromossômicas em diversas doenças que empregam o FISH como método diagnóstico e de acompanhamento.

REFERÊNCIAS

Van de Donk NWCJ, Pawlyn C, Yong KL. Multiple myeloma. *Lancet*. 2021 Jan 30;397(10272):410-427. doi: 10.1016/S0140-6736(21)00135-5. PMID: 33516340.

Albagoush SA, Shumway C, Azevedo AM. Multiple Myeloma. 2023 Jan 30. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan–. PMID: 30521185.

Cowan AJ, Green DJ, Kwok M, Lee S, Coffey DG, Holmberg LA, Tuazon S, Gopal AK, Libby EN. Diagnosis and Management of Multiple Myeloma: A Review. *JAMA*. 2022 Feb 1;327(5):464-477. doi: 10.1001/jama.2022.0003. PMID: 35103762.

Dominik D. Alexander, Mink, P.J., Adami, H.-O., Cole, P., Mandel, J.S., Oken, M.M. and Trichopoulos, D. (2007), Multiple myeloma: A review of the epidemiologic literature. *Int. J. Cancer*, 120: 40-61. <https://doi.org/10.1002/ijc.22718>

Rajkumar SV. Multiple myeloma: 2022 update on diagnosis, risk stratification, and management. *Am J Hematol*. 2022 Aug;97(8):1086-1107. doi: 10.1002/ajh.26590. Epub 2022 May 23. PMID: 35560063; PMCID: PMC9387011.

Soekojo CY, Chng WJ. Treatment horizon in multiple myeloma. *Eur J Haematol*. 2022 Nov;109(5):425-440. doi: 10.1111/ejh.13840. Epub 2022 Sep 18. PMID: 35880395.

Rajkumar SV. Multiple myeloma: 2018 update on diagnosis, risk- stratification, and management. *Am J Hematol*. 2018 Aug 16;93(8):981-1114. doi: 10.1002/ajh.25117. PMID: 30400719; PMCID: PMC6223128.

Rajkumar SV. Multiple myeloma: Every year a new standard? *Hematol Oncol*. 2019 Jun;37 Suppl 1(Suppl 1):62-65. doi: 10.1002/hon.2586. PMID: 31187526; PMCID: PMC6570407.

Fonseca R, Blood EA, Oken MM, Kyle RA, Dewald GW, Bailey RJ, Van Wier SA, Henderson KJ, Hoyer JD, Harrington D, Kay NE, Van Ness B, Greipp PR. Myeloma and the t(11;14)(q13;q32); evidence for a biologically defined unique subset of patients. *Blood*. 2002 May 15;99(10):3735-41. doi: 10.1182/blood.v99.10.3735. PMID: 11986230.

Kuehl WM, Bergsagel PL. Multiple myeloma: evolving genetic events and host interactions. *Nat Rev Cancer*. 2002 Mar;2(3):175-87. doi: 10.1038/nrc746. PMID: 11990854.

Keats JJ, Reiman T, Maxwell CA, Taylor BJ, Larratt LM, Mant MJ, Belch AR, Pilarski LM. In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood*. 2003 Feb 15;101(4):1520-9. doi: 10.1182/blood-2002-06-1675. Epub 2002 Oct 3. PMID: 12393535.

Fonseca R, Barlogie B, Bataille R, Bastard C, Bergsagel PL, Chesi M, Davies FE, Drach J, Greipp PR, Kirsch IR, Kuehl WM, Hernandez JM, Minvielle S, Pilarski LM, Shaughnessy JD Jr, Stewart AK, Avet-Loiseau H. Genetics and cytogenetics of

multiple myeloma: a workshop report. *Cancer Res.* 2004 Feb 15;64(4):1546-58. doi: 10.1158/0008-5472.can-03-2876. PMID: 14989251.

BRASIL. Comissão Nacional de Incorporação de Tecnologias no Sistema Único de Saúde (CONITEC). Relatório de recomendação: teste citogenético por Hibridização in Situ por Fluorescência (FISH) para detecção de alterações citogenéticas de alto risco em pacientes com mieloma múltiplo. BRASÍLIA, DF. DEZEMBRO, 2021.

Hu L, Ru K, Zhang L, Huang Y, Zhu X, Liu H, Zetterberg A, Cheng T, Miao W. Fluorescence in situ hybridization (FISH): an increasingly demanded tool for biomarker research and personalized medicine. *Biomark Res.* 2014 Feb 5;2(1):3. doi: 10.1186/2050-7771-2-3. PMID: 24499728; PMCID: PMC3917523.

Liu G, Zhang T. Single Copy Oligonucleotide Fluorescence In Situ Hybridization Probe Design Platforms: Development, Application and Evaluation. *Int J Mol Sci.* 2021 Jul 1;22(13):7124. doi: 10.3390/ijms22137124. PMID: 34281175; PMCID: PMC8268824.

Dieffenbach CW, Lowe TM, Dveksler GS. General concepts for PCR primer design. *PCR Methods Appl.* 1993 Dec;3(3):S30-7. doi: 10.1101/gr.3.3.s30. PMID: 8118394.

Hiller B, Bradtke J, Balz H and Rieder H (2004): "CyDAS: a cytogenetic data analysis system", *Bioinformatics* 2005 21(7):1282-1283; doi:10.1093/bioinformatics/bti146. *Bioinformatics Advance Access* originally published online on November 16, 2004

UCSC Genome Browser: Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res.* 2002 Jun;12(6):996-1006.

Gelali E, Girelli G, Matsumoto M, Wernersson E, Custodio J, Mota A, Schweitzer M, Ferenc K, Li X, Mirzazadeh R, Agostini F, Schell JP, Lanner F, Crosetto N, Bienko M. iFISH is a publically available resource enabling versatile DNA FISH to study genome architecture. *Nat Commun.* 2019 Apr 9;10(1):1636. doi: 10.1038/s41467-019-09616-w. PMID: 30967549; PMCID: PMC6456570.

Beliveau, Brian J., et al. "OligoMiner provides a rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes." *Proceedings of the National Academy of Sciences* (2018): 201714530.

Shannon M. McNulty, Lori L. Sullivan, Beth A. Sullivan, Human Centromeres Produce Chromosome-Specific and Array-Specific Alpha Satellite Transcripts that Are Complexed with CENP-A and CENP-C, *Developmental Cell*, Volume 42, Issue 3, 2017, Pages 226-240.e6, ISSN 1534-5807, <https://doi.org/10.1016/j.devcel.2017.07.001>.