



UNIVERSIDADE ESTADUAL PAULISTA

“Júlio de Mesquita Filho”

FACULDADE DE MEDICINA DE BOTUCATU

Camilla Donato Contiero

**Perfil imune do adenocarcinoma de pulmão de pacientes
não fumantes**

Dissertação apresentada à Faculdade de Medicina, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Botucatu, para obtenção do título de Mestre(a) em Cirurgia e Medicina Translacional.

Orientadora: Profa. Dra. Patricia Pintor dos Reis

Coorientadora: Dra. Iael Weissberg Minutentag

BOTUCATU -SP

2026

Camilla Donato Contiero

Perfil imune do adenocarcinoma de pulmão de pacientes não fumantes

Dissertação apresentada à Faculdade de
Medicina, Universidade Estadual Paulista
“Júlio de Mesquita Filho”, Câmpus de
Botucatu, para obtenção do título de
Mestre(a) em Cirurgia e Medicina
Translacional.

Orientadora: Profa. Dra. Patrícia Pintor dos Reis

Coorientadora: Dra. Iael Weissberg Minutentag

BOTUCATU -SP

2026

C762p	Contiero, Camilla Donato Perfil imune do adenocarcinoma de pulmão de pacientes não fumantes / Camilla Donato Contiero. -- Botucatu, 2026 92 p. Dissertação (mestrado) - Universidade Estadual Paulista (UNESP), Faculdade de Medicina, Botucatu Orientadora: Patricia Pintor dos Reis Coorientadora: Iael Weissberg Minutentag 1. Genética molecular. 2. Transcriptoma. 3. Adenocarcinoma. 4. Pulmões Câncer. I. Título.
-------	---

Sistema de geração automática de fichas catalográficas da Unesp.
Dados fornecidos pelo autor(a).

Impacto Social

Esta pesquisa investiga o perfil molecular e o sistema de defesa (imunológico) do câncer de pulmão em indivíduos que nunca fumaram, visando compreender como o tumor se comporta e contorna as defesas do organismo. Alinhado ao ODS 3 (Saúde e Bem-Estar), o estudo busca identificar padrões que viabilizem diagnósticos mais precisos e precoces, fundamentais para elevar a sobrevida e a qualidade de vida dos pacientes.

ATA DA DEFESA PÚBLICA DA DISSERTAÇÃO DE MESTRADO DE CAMILLA DONATO CONTIERO, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM CIRURGIA E MEDICINA TRANSLACIONAL, DA FACULDADE DE MEDICINA - CÂMPUS DE BOTUCATU.

Aos 27 de março de 2026, às 9h, por meio de Videoconferência, realizou-se a defesa de DISSERTAÇÃO DE MESTRADO de CAMILLA DONATO CONTIERO, intitulada "Perfil imune do adenocarcinoma de pulmão de pacientes não fumantes.

". A Comissão Examinadora foi constituída pelos seguintes membros: Profa. Dra. PATRICIA PINTOR DOS REIS (Orientador(a) - Participação Virtual) do(a) Assessoria de Relações Externas / UNESP / Reitoria - RUNESP, Profa. Dra. CAMILA FERREIRA BANNWART CASTRO (Participação Virtual) do(a) Instituto de Biociências de Botucatu/IBB, Profa. Dra. CLÁUDIA MALHEIROS COUTINHO CAMILLO (Participação Virtual) do(a) Centro Internacional de Pesquisa (CIPE) / A.C.Camargo Cancer Center, Após a exposição pela mestrande e arguição pelos membros da Comissão Examinadora que participaram do ato, de forma presencial e/ou virtual, a discente recebeu o conceito final: APROVADA. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelo(a) Presidente(a) da Comissão Examinadora.

Documento assinado digitalmente
 PATRICIA PINTOR DOS REIS
Data: 01/04/2026 15:34:43-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. PATRICIA PINTOR DOS REIS

Agradecimentos

Agradeço, primeiramente, aos meus pais, José Antonio Contiero e Doraine F. Donato Contiero, ao meu namorado, Nathan Bortolotti, e a toda a minha família, por me apoiarem e encorajarem durante esses dois anos. Sem o amor, o suporte e o incentivo constante de vocês, eu não teria chegado até aqui.

À minha orientadora, Prof^a. Dr^a. Patrícia Pintor dos Reis, expresso minha gratidão especial por acreditar no meu potencial desde o início (muito antes do mestrado, ainda na iniciação científica) e por me incentivar tanto no âmbito profissional quanto pessoal. Sua orientação foi fundamental para o meu amadurecimento e crescimento.

À minha coorientadora, Dr^a. Iael Weissberg Minutentag, meu muito obrigada por compartilhar seus conhecimentos, pelos conselhos valiosos, pelo incentivo e, sobretudo, pela paciência em me ensinar cada detalhe.

Aos amigos que fiz na UNIPEX (Unidade de Pesquisa Experimental), lugar que foi minha segunda casa durante o mestrado, agradeço pelo acolhimento e por cada momento partilhado. Obrigada por dividirem comigo almoços, cafés da tarde, conselhos sinceros, risadas (sinceras e nervoso), choros. Muito obrigada por tornarem essa jornada mais leve.

Aos meus colegas de laboratório, agradeço pela convivência diária e pela ajuda mútua. Obrigada por estarem ao meu lado em todos os momentos, celebrando as vitórias e enfrentando as dificuldades. Vocês contribuíram imensamente para a minha formação, tanto profissional quanto pessoal.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de Mestrado Acadêmico, que possibilitou a realização deste trabalho.

Por fim, agradeço a todo suporte técnico pela disponibilidade e pela ajuda no dia a dia do laboratório. A colaboração de vocês foi muito importante para a realização deste trabalho.

Apesar de todos os percalços, me sinto feliz e realizada por concluir este mestrado, olhando para trás e percebendo que conquistei muito mais do que um título. Levo todos comigo, pois cada um deixou a sua marca em mim.

SUMÁRIO

1. INTRODUÇÃO	7
1.1. Câncer de Pulmão: Relevância Global e Epidemiologia	7
1.2. O Câncer de Pulmão em Não Fumantes (CPNF): Um Desafio Emergente	7
1.3. Perfil Biológico, Molecular e Imunológico do Câncer de Pulmão em Não Fumantes (CPNF)	8
1.4. O Cenário Terapêutico e o Problema do Diagnóstico Tardio em CPNF	9
1.5. Lacunas no Rastreamento e a Justificativa da Pesquisa em CPNF	11
2. OBJETIVOS	14
2.1. Objetivo geral	14
2.2. Objetivos específicos	14
2.2.1. Objetivo 1:	14
2.2.2. Objetivo 2:	14
2.2.3. Objetivo 3:	14
3. MATERIAL E MÉTODOS	15
3.1. Considerações éticas	15
3.2. Seleção de dados e amostras de pacientes	15
3.3. Processamento das amostras clínicas	18
3.3.1. Extração de RNA	18
3.3.2. Preparação de bibliotecas e sequenciamento	20
3.3.3. Análise em NanoString	21
3.4. ANÁLISE DE DADOS	21
3.4.1. Expressão gênica	21
3.4.2. Perfil Imune	24
4. RESULTADOS	27
4.1. Expressão gênica e enriquecimento funcional dos dados de rna-seq obtidos do tgca	27
4.2. Expressão gênica e enriquecimento funcional dos dados obtidos do rna-seq illumina	30

4.3.	Comparação dados de rna-seq (tcga) com dados de rna-seq illumina.....	33
4.4.	Infiltrado imune.....	35
4.4.1.	Banco de dados (tgca)	35
4.4.2.	Abundância celular nanostring	40
4.5.	Expressão gênica do nanostring	41
5.	DISCUSSÃO	44
5.1.	Resultados do sequenciamento illumina: perfil trnascriptomico e imune de tumores primários e metástase cerebral de adenocarcinoma pulmonar de pacientes não fumantes.....	44
5.2.	Análise nanostring: assinatura imune e diferenciação entre tumor primário e metástase cerebral de adenocarcinoma pulmonar de pacientes não fumantes	53
5.3.	Comparação de técnicas e robustez da metodologia	58
6.	CONCLUSÃO	61
7.	REFERÊNCIAS BIBLIOGRÁFICAS	63

RESUMO

O câncer de pulmão é um dos mais incidentes e a principal causa de mortalidade por câncer no mundo, tendo o tabagismo como fator de risco central. Entretanto, os casos em não fumantes (CPNF) têm aumentado, com diferenças biológicas, epidemiológicas e clínicas, configurando uma entidade distinta ainda pouco explorada. Diante da escassez de estudos nessa população (principalmente acerca da biologia molecular), este trabalho buscou caracterizar o perfil imune do adenocarcinoma pulmonar (LUAD) em não fumantes por meio de abordagens integradas: (1) análise computacional *in silico* de dados públicos de sequenciamento de alto desempenho (NGS-RNA-Seq) obtidos do banco de dados público The Cancer Genome Atlas (TCGA) e (2) análise de amostras clínicas do HCFMB/UNESP, incluindo tumores primários e metástases cerebrais de LUAD e histórico conhecido de tabagismo, por meio de NGS-RNA-Seq e NanoString nCounter®. Esta última, baseada em um painel de genes específicos relacionados à imunologia tumoral, permitiu refinar e detalhar os achados, com foco em genes associados à resposta imune tumoral. Considerando a limitação do TCGA em relação a amostras de pacientes brasileiros, a integração das duas abordagens de NGS-RNA-Seq e Nanostring nCounter® possibilitou comparar perfis transcriptômicos de fumantes e não fumantes, identificar genes diferencialmente expressos e relacioná-los a distintos microambientes imunes. O uso de múltiplas ferramentas de análise bioinformática conferiu maior robustez às análises, ampliando a compreensão dos resultados e maximizando a relevância translacional da pesquisa. Entre os resultados esperados, pretendemos contribuir para a elucidação de mecanismos imunológicos associados a essa fração de pacientes com adenocarcinoma pulmonar e sem histórico de tabagismo, oferecendo subsídios para estratégias terapêuticas direcionadas.

Palavras-chave: adenocarcinoma de pulmão, perfil imune, transcriptoma.

Abstract

Lung cancer is one of the most prevalent malignancies and the leading cause of cancer-related mortality worldwide, with tobacco smoking being its primary risk factor. However, the incidence of lung cancer in never-smokers (LCNS) has been increasing, presenting distinct biological, epidemiological, and clinical characteristics that define it as a separate disease entity, which remains insufficiently explored. Given the limited number of studies focusing on this population, particularly regarding its molecular biology, this study aimed to characterize the immune profile of lung adenocarcinoma (LUAD) in never-smokers through two integrated approaches: (1) an *in silico* computational analysis of publicly available next-generation RNA sequencing (RNA-Seq) data from The Cancer Genome Atlas (TCGA) and (2) the analysis of clinical samples from HCFMB/UNESP, including primary tumors and brain metastases from LUAD patients with documented smoking history, using RNA-Seq and the NanoString nCounter® platform. The latter, based on a targeted panel of genes related to tumor immunology, enabled a more refined and detailed characterization of the findings, with emphasis on genes associated with the antitumor immune response. Considering the limited representation of Brazilian patients in the TCGA dataset, the integration of RNA-Seq and NanoString nCounter® analyses allowed the comparison of transcriptomic profiles between smokers and never-smokers, the identification of differentially expressed genes, and their association with distinct immune microenvironments. The application of multiple bioinformatics analysis tools increased the robustness of the results, providing a more comprehensive understanding of the data and enhancing the translational relevance of the study. Ultimately, this work aims to contribute to the elucidation of immunological mechanisms associated with lung adenocarcinoma in patients without a history of tobacco smoking, thereby providing insights that may support the development of targeted therapeutic strategies.

Keywords: lung adenocarcinoma; immune profile; transcriptome.

1. INTRODUÇÃO

1.1. Câncer de Pulmão: Relevância Global e Epidemiologia

O câncer é um dos principais problemas de saúde pública mundialmente, destacando-se como uma das principais causas de morte e, consequentemente, uma das barreiras para o aumento da expectativa de vida em todo o mundo (Câncer, 2023). Segundo estimativas do Global Cancer Observatory (Globocan), em 2022 no mundo ocorreram cerca de 20 milhões de casos novos de câncer, com aproximadamente 9,7 milhões de mortes pelos vários tipos de câncer em todas as idades e em ambos os sexos (Bray et al., 2024). Ainda segundo as estimativas mundiais (Bray et al., 2024), o câncer de pulmão foi o câncer mais diagnosticado e é o segundo câncer mais comum com cerca de 2,5 milhões de casos novos e o líder em mortalidade com 1,8 milhões de óbitos no mundo. No Brasil, segundo as estimativas recentes (2023) do Instituto Nacional de Câncer (INCA), o câncer de pulmão é o terceiro mais comum em homens (18.020 casos novos/ano) e o quarto em mulheres (14.540 casos novos/ano) (Câncer, 2023).

1.2. O Câncer de Pulmão em Não Fumantes (CPNF): Um Desafio Emergente

Embora as taxas de tabagismo e, consequentemente, de câncer de pulmão venham diminuindo, tem-se observado um aumento expressivo da incidência da doença em indivíduos que nunca fumaram. Em 2023, nos Estados Unidos, estimou-se que mais de 20 mil mortes por câncer de pulmão ocorreram em não fumantes (Siegel et al., 2023). No Reino Unido, esse grupo foi responsável pela oitava principal causa de morte por câncer (Hill et al., 2023). Esse fenômeno tem sido especialmente notado à medida que programas de controle do tabagismo avançam e a proporção de casos de câncer de pulmão de células não pequenas (NSCLC) em não fumantes aumenta por razões ainda não completamente elucidadas (Klemperer et al., 2021).

Estudos globais apontam que o câncer de pulmão em indivíduos que nunca fumaram (CPNF) corresponde a aproximadamente 10% a 25% de todos os casos diagnosticados. Dentro desse grupo, o adenocarcinoma (LUAD) é o subtipo histológico mais prevalente, caracterizando-se também como o tipo mais comum entre os carcinomas de pulmão de células não pequenas (NSCLC), respondendo por cerca de 40% dos casos de câncer pulmonar na população geral (Boeckx et al., 2020; Cho

et al., 2017). Esse padrão reforça a predominância do adenocarcinoma como a principal forma de câncer de pulmão entre não fumantes (Kuśnierczyk, 2023).

O adenocarcinoma pulmonar (LUAD) costuma ser identificado tardiamente, quando já há disseminação regional ou à distância, incluindo o cérebro — um fator que contribui de forma decisiva para sua elevada letalidade. Entre os tumores pulmonares, é o subtipo com maior probabilidade de originar metástases cerebrais (Souza et al., 2024). Aproximadamente 20% dos pacientes desenvolvem esse tipo de metástase ao longo da evolução clínica da doença. No cenário das metástases cerebrais de origem tumoral, o LUAD é o principal responsável, representando cerca de 37% dos casos (Lin et al., 2021). Estudos recentes também indicam que o desenvolvimento de metástase no sistema nervoso central envolve mecanismos biológicos complexos, que vão além de alterações genéticas e incluem influências significativas do microambiente imunológico tumoral (TIME) (Wang et al., 2024).

1.3. Perfil Biológico, Molecular e Imunológico do Câncer de Pulmão em Não Fumantes (CPNF)

O CPNF exibe um perfil demográfico, biológico e epidemiológico diferenciado, com maior incidência em mulheres, pessoas de origem asiática e indivíduos mais jovens (Devarakonda et al., 2021). Embora os mecanismos moleculares ainda não estejam plenamente compreendidos, fatores ambientais e genéticos têm sido associados ao desenvolvimento da doença. Entre os fatores ambientais e ocupacionais, destacam-se a exposição a amianto, cromo, arsênio, poluição do ar e fumaça oriunda da queima de lenha e frituras, agentes reconhecidamente carcinogênicos, com maior impacto em países em desenvolvimento e na população feminina (Casal-Mouriño et al., 2019). No âmbito genético, o risco de câncer de pulmão é significativamente maior em indivíduos com histórico familiar da doença, especialmente em casos de início precoce (antes dos 50 anos) (Cheng et al., 2022) ou com diagnóstico de adenocarcinoma (Daylan et al., 2023), possivelmente devido à combinação de fatores ambientais compartilhados e predisposição genética (Kerrigan et al., 2021)- associada à presença de genes dominantes raros ou mutações germinativas ou somáticas que afetam diretamente a vulnerabilidade ao tumor (Carrot-Zhang et al., 2021). Ainda que esses fatores venham sendo relatados com frequência na literatura, não se pode afirmar com certeza que configuram causas diretas da doença (Dubin; Griffin, 2020).

Complementarmente, evidências moleculares vêm demonstrando que o CPNF possui assinaturas genéticas distintas em comparação com tumores associados ao tabagismo. Estudo conduzido com 232 amostras revelou que tumores de CPNF apresentam baixa carga mutacional e alterações mutuamente exclusivas nos genes EGFR e KRAS, sendo a mutação em EGFR a mais prevalente, com incidência entre 30% e 54% dos casos, enquanto mutações em KRAS ocorrem entre 5,6% e 12,5% (Zhang et al., 2021). Além disso, alterações nos genes ALK foram encontradas em até 13,9% dos casos (Devarakonda et al., 2021). Essas mutações afetam diretamente vias críticas de sinalização celular, como VEGF, ERBB, GnRH, TGF- β e o ciclo celular (Hamouz et al., 2023), sendo classificadas como alterações "driver" com relevância terapêutica, encontradas em até 92% dos casos (Devarakonda et al., 2021). Esses achados corroboram a noção de que o CPNF constitui uma entidade clínica e molecular distinta, com implicações diagnósticas e terapêuticas próprias (Wang et al., 2023a).

Fatores imunológicos também parecem influenciar a fisiopatologia do CPNF. Indivíduos não fumantes podem apresentar padrões imunológicos específicos que contribuem para o escape imune tumoral. Evidências sugerem envolvimento da via de sinalização MAPK no controle da proliferação celular e evasão imune, bem como alterações funcionais em linfócitos T, B e células NK, comprometendo as respostas inata e adaptativa (Guo et al., 2020; Hamouz et al., 2023). Apesar da menor carga mutacional, esses tumores parecem ser mais propensos à transformação maligna que os tumores de fumantes, cujas mutações muitas vezes são passageiras. O CPNF ainda tende a ocorrer em pacientes mais jovens e a apresentar melhor prognóstico e resposta ao tratamento, sobretudo em mulheres (Dubin; Griffin, 2020). Neste sentido, o desenvolvimento de biomarcadores tumorais eficazes é um desafio, visto que é comprovada a heterogeneidade do tumor, tanto intratumoral quanto entre tumores de diferentes pacientes. Quanto à heterogeneidade intratumoral, devemos considerar a complexa interação do clone tumoral com o microambiente, com a ocorrência de comunicação entre as células tumorais e outras células do microambiente tumoral (TME), incluindo as células imunes (Rahal et al., 2025).

1.4. O Cenário Terapêutico e o Problema do Diagnóstico Tardio em CPNF

Nos últimos anos, metodologias avançadas de biologia molecular, como o

sequenciamento de alto desempenho (NGS), foram desenvolvidas para solucionar esses problemas. Essas tecnologias podem revolucionar as abordagens clínicas, pois com elas seria possível identificar os marcadores moleculares de risco (em indivíduos que não possuem câncer - etapa pré-câncer), detecção e prognóstico precoces (no câncer em estágio inicial) e sensibilidade ao tratamento (em casos de estágio inicial e avançado) (Lee et al., 2024).

Apesar dos avanços na detecção e tratamento do câncer de pulmão, ainda na prática clínica o diagnóstico é realizado na maioria das vezes em uma fase tardia da doença, sendo notória a baixa sobrevida dos pacientes 5 anos após o diagnóstico (Association, [S.d.]). As opções de tratamento dos pacientes diagnosticados com carcinomas de pulmão de não pequenas células são limitadas (Siegel et al., 2023) e, de forma geral, envolvem abordagens multidisciplinares, incluindo a cirurgia, quimioterapia, radioterapia, terapias com alvos moleculares e, com destaque crescente, a imunoterapia. As modalidades de tratamento são determinadas com base no diagnóstico, no estadiamento da doença e nos critérios de elegibilidade.

As estatísticas de estadiamento no momento do diagnóstico são alarmantes: cerca de 23% dos pacientes com câncer de pulmão são diagnosticados com a doença localizada apenas no sítio primário. Em aproximadamente 21% dos casos a doença acomete os linfonodos regionais (metástase regional), e em cerca de 52% dos pacientes a doença já se encontra em fase metastática à distância, a qual acomete principalmente o cérebro e os ossos. A sobrevida é impactada drasticamente pelo estágio, sendo a cirurgia a maior chance de cura. Para a doença localizada, a sobrevida em 5 anos é de aproximadamente 65%, já na doença metastática à distância (Estágio IV), esta taxa permanece baixa, em torno de 9%, apesar dos progressos na área (SEER, 2025). Nesses pacientes, o tratamento sistêmico tem primordialmente intenção paliativa. Os esquemas de quimioterapia mais comumente usados para doença avançada são esquemas combinados como Paclitaxel/Carboplatina ou Cisplatina (Riely et al., 2024).

Com a alta frequência de mutações oncogênicas em NSCLC, como as alterações em *EGFR*, *ALK*, *ROS1*, *KRAS*, *HER2*, *BRAF*, *MET*, *RET* e *NTRK1-3*, diversas opções de terapias direcionadas foram desenvolvidas. Para a maioria desses alvos, a terapia de primeira linha consiste no uso de medicamentos específicos, com exceção de mutações em *KRAS* e *HER2*, onde o tratamento direcionado costuma ser indicado apenas em segunda linha (Hendriks et al., 2023). Esses fármacos, em sua

maioria inibidores de tirosina quinase (TKIs), são administrados por via oral, o que proporciona maior comodidade e impacto positivo na sobrevida livre de progressão da doença e na qualidade de vida, quando comparados à quimioterapia convencional (Yu et al., 2024).

Fatores imunológicos também impulsionaram os avanços recentes. O mecanismo de escape imunológico nos tumores pulmonares é bastante complexo, envolvendo os *checkpoints* imunológicos PD-1/PD-L1 e CTLA-4, que funcionam como reguladores essenciais na comunicação entre células tumorais e células do sistema imune (Tang et al., 2022). Nesse contexto, a imunoterapia tem como objetivo potencializar o reconhecimento dos antígenos tumorais, favorecendo a ativação da resposta imune. Atualmente, diferentes estratégias vêm sendo utilizadas, desde monoterapias com anticorpos monoclonais voltados contra PD-1/PD-L1 — como Nivolumabe, Pembrolizumabe, Atezolizumabe, Avelumabe, Durvalumabe e Cemiplimabe - até terapias combinadas, que associam imunoterapia à quimioterapia, radioterapia, dupla imunoterapia ou a terapias direcionadas (Tang et al., 2022). Esses avanços inovadores, marcados pelas terapias direcionadas e imunoterapias, iniciaram uma nova era da medicina de precisão, permitindo que os tratamentos sejam adaptados aos perfis moleculares e imunológicos exclusivos dos pacientes, melhorando assim a eficácia e os resultados de sobrevida (Jeon et al., 2025).

1.5. Lacunas no Rastreamento e a Justificativa da Pesquisa em CPNF

Quanto ao diagnóstico precoce, atualmente existem diretrizes de rastreamento do câncer de pulmão, sugeridas pela US Preventive Services Task Force (2013), somente para pessoas de 55 a 80 anos de idade que tenham histórico de tabagismo com consumo de 20 maços-ano e/ou fumantes atuais ou aqueles que pararam de fumar nos últimos 15 anos (US Preventive Services Task Force et al., 2021). Logo, para indivíduos não fumantes não há nenhuma política de recomendação de rastreamento, tanto no Brasil como no mundo todo. Isso se dá principalmente devido aos sintomas inespecíficos de doença nesses pacientes e aos sintomas específicos, tais como dor torácica e hemoptise, os quais confundem-se com outras doenças (Dubin; Griffin, 2020), além de não serem relatadas evidências científicas de que o rastreamento do câncer de pulmão na população em geral traga mais benefícios do que riscos (Care, 2016; US Preventive Services Task Force et al., 2021). O que se

tem conhecimento, de acordo com alguns estudos, é que a realização de tomografia de baixa dose de radiação em fumantes (indivíduos fumantes com histórico de tabagismo de um maço de cigarro/dia por 30 anos) e com idade acima de 55 anos, é capaz de reduzir a mortalidade pelo câncer de pulmão (Care, 2016; US Preventive Services Task Force et al., 2021).

Assim, se faz necessário o estudo e pesquisa mais profundos acerca do câncer de pulmão em não fumantes, porque esses pacientes possuem características epidemiológicas, clínicas patológicas, genéticas, de resposta ao tratamento e de sobrevida distintas do grupo de pacientes com histórico de tabagismo (Vasudevan; Krishna; Mehta, [S.d.]). Conhecendo-se o perfil molecular, é possível desenvolver marcadores preditivos e, assim, orientar o tratamento e otimizar o desfecho clínico desses pacientes, podendo detectar o câncer em estágio inicial (Daylan et al., 2023). Além disso, conhecendo o microambiente complexo e dinâmico dos CPNF, é possível desenvolver novas estratégias terapêuticas para beneficiar esses pacientes. Neste sentido, a quantificação das diferentes populações de células imunes no TME deve auxiliar no entendimento dos mecanismos que coordenam a resposta imune antitumoral, com vistas ao desenvolvimento de novas imunoterapias. Os métodos tradicionais de quantificação do infiltrado imune são imunohistoquímica, imunofluorescência e citometria de fluxo (Finotello; Trajanoski, 2018). Ultimamente, ferramentas de bioinformática têm sido usadas para quantificar as subpopulações celulares do infiltrado inflamatório no tecido tumoral a partir de dados de expressão gênica global (Hackl et al., 2016), conforme o proposto neste projeto de pesquisa.

Algoritmos como CIBERSORT (Newman et al., 2015), TIMER (Li et al., 2017), MCP-counter (Becht et al., 2016), xCell (Aran; Hu; Butte, 2017) e EPIC (Racle et al., 2017) são capazes de quantificar indiretamente subpopulações de células imunes infiltrantes do tumor, auxiliando na compreensão do contexto imunológico dos tumores. O método mais comumente aplicado para análise do perfil imune é a análise de enriquecimento de grupos de genes. O resultado é avaliado por meio de escores de enriquecimento, onde os extremos (escore alto e baixo) indicam genes mais ou menos expressos, respectivamente, para cada população celular investigada. Um uso alternativo desse método é a análise de enriquecimento para amostra única, onde escores individuais de enriquecimento são obtidos para cada amostra avaliada. Esse método foi utilizado pelo algoritmo xCell (Aran; Hu; Butte, 2017), cuja principal vantagem é não exigir um número elevado de amostras clínicas, tornando-se aplicável

em um número maior de estudos.

A análise de células imunes infiltrantes tumorais pode servir para o delineamento de biomarcadores prognósticos e preditivos em CPNF. Os linfócitos T, populações celulares heterogêneas, além de outras células do sistema imune, como células *natural killer* (NK), linfócitos B, células dendríticas e macrófagos, podem ter um papel relevante no desenvolvimento e progressão dos CPNF. Assim, estudos são necessários para melhor compreensão das interações tumor-sistema imune, bem como para sua aplicação clínica como biomarcadores. Portanto, esse projeto propõe aplicar métodos de análise bioinformática para a identificação do perfil imune associado ao adenocarcinoma de pulmão como uma estratégia de conhecer o TME e sua relevância para o tratamento dos pacientes.

2. OBJETIVOS

2.1. Objetivo geral

Determinar o perfil molecular do adenocarcinoma de pulmão em pacientes não fumantes.

2.2. Objetivos específicos

2.2.1. Objetivo 1:

Analisar e comparar a expressão gênica diferencial entre pacientes fumantes e não fumantes com tumor primário de adenocarcinoma pulmonar, bem como identificar, por meio de análises *in silico*, genes candidatos e vias moleculares associadas aos tumores de pacientes não fumantes.

2.2.2. Objetivo 2:

Integrar dados de RNA-Seq de amostras clínicas metastáticas (de pacientes fumantes e não fumantes), comparando com os achados do tumor primário e identificar assinaturas moleculares e vias biológicas relevantes no contexto da metástase cerebral.

2.2.3. Objetivo 3:

Caracterizar e comparar o perfil imune de amostras clínicas de adenocarcinoma pulmonar primário e metastático cerebral de pacientes não fumantes, utilizando a tecnologia NanoString nCounter® com um painel imunológica específico, visando à identificação de assinaturas imunológicas associadas à progressão tumoral.

3. MATERIAL E MÉTODOS

Este projeto foi desenvolvido com base em três abordagens complementares: (1) uma análise computacional (*in silico*) utilizando dados previamente gerados e disponibilizados em bancos de dados públicos do *The Cancer Genome Atlas* (TCGA); (2) a geração de dados transcriptômicos inéditos a partir de amostras clínicas provenientes do Hospital de Clínicas da Faculdade de Medicina de Botucatu (HCFMB) – UNESP, por meio da tecnologia de sequenciamento de RNA de nova geração (RNA-Seq); e (3) caracterização imunológica de amostras clínicas (metástases e tumores primários), por meio de um painel imune disponível na tecnologia NanoString nCounter®, visando a validação e o refinamento de assinaturas identificadas no RNA-seq.

3.1. Considerações éticas

Este projeto foi aprovado pelo HCFMB - Faculdade de Medicina de Botucatu, UNESP, Botucatu, SP (CEP #87/2024; Protocolo CAAE 78956524.2.0000.5411). A obtenção de amostras clínicas foi realizada a partir do Biobanco do HCFMB, aprovado pelo CONEP (Registro CONEP B-047-CAAE 59676422.3.0000.5411) e o qual já contempla o Termo de Consentimento Livre e Esclarecido (TCLE) assinado pelos pacientes.

3.2. Seleção de dados e amostras de pacientes

No banco de dados públicos TCGA (<https://portal.gdc.cancer.gov/>), há dados de RNA-Seq disponíveis de 192 casos de adenocarcinoma pulmonar de tumor primário, sendo 74 amostras de tumores de pacientes não fumantes e 118 de pacientes fumantes. Para seleção de dados adotamos os seguintes critérios: excluímos amostras que estavam em duplicatas, amostras de tecido sólido normal, de tumor recorrente e que não apresentavam dados clínicos completos.

Quanto às amostras do HCFMB, foram selecionados pacientes que apresentavam dados clínicos completos, incluindo subtipo histológico confirmado de adenocarcinoma pulmonar, histórico de tabagismo (fumante ou não fumante) e sítio tumoral (tumor primário ou metástase). Essas amostras foram analisadas por RNA-Seq e NanoString nCounter®. O RNA-Seq foi realizado em sete amostras clínicas, todas provenientes de tecido congelado armazenado no Biobanco do HCFMB, enquanto a análise por NanoString incluiu 9 amostras clínicas, quatro de tecido

congelado, disponíveis no Biobanco e cinco provenientes de blocos de parafina.

No total, 12 amostras foram incluídas no estudo, das quais quatro foram avaliadas por ambas as metodologias. Entre as amostras de tecido congelado (n=7), todas correspondiam a metástases cerebrais, incluindo quatro de pacientes não fumantes e três de pacientes fumantes, ambos com diagnóstico de adenocarcinoma pulmonar. Já as amostras provenientes de blocos de parafina (n=5) eram exclusivamente de pacientes não fumantes com adenocarcinoma pulmonar. O detalhamento do grupo amostral, bem como as características clínicas e demográficas e o delineamento experimental do estudo, estão apresentados na Figura 1 e na Tabela 1, sendo informações mais detalhadas encontram-se nas tabelas suplementares (Tabela Suplementar 1 e Tabela Suplementar 2).

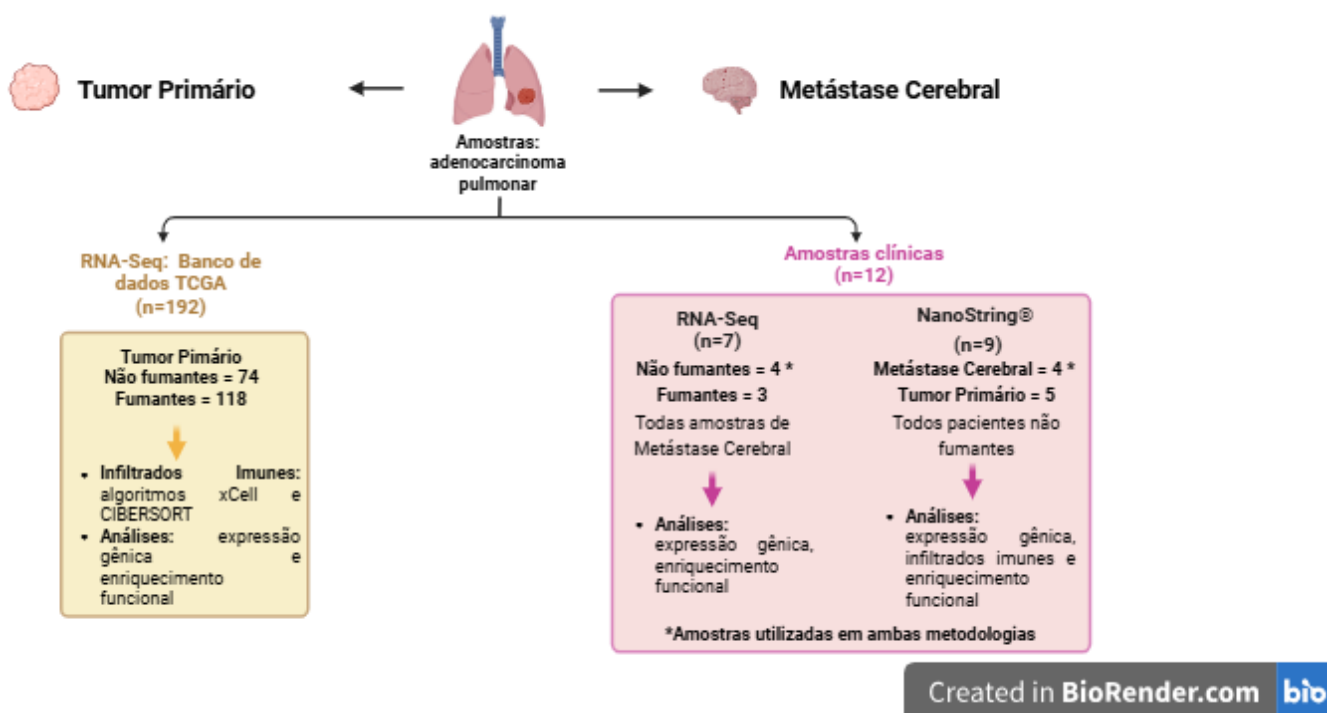


Figura 1. Fluxograma de detalhamento dos grupos amostrais e do delineamento experimental do projeto, incluindo a caracterização clínica e a indicação das análises realizadas.

Tabela 1. Características clínicas e demográficas das amostras clínicas obtidas no HCFMB e do banco de dados TCGA, separadas por histórico de tabagismo. São apresentadas a idade média com desvio padrão, sexo e presença de metástase à distância.

Variáveis	Amostras Clínicas (N total=12)		TCGA LUAD (N total=192)	
	Fumantes	Não fumantes	Fumantes	Não fumantes
Amostragem	3	9	118	74
Idade Média ± DP (anos)	60.67 ± 11.24	61.22 ± 10.13	61,21 ± 9,75	65,82 ± 9,69
Sexo				
Masculino	2	3	67	21
Feminino	1	6	50	53
Tamanho Tumoral (T)				
T1	1	4	41	17
T2	1	2	58	44
T3	-	-	14	8
T4	1	3	4	1
Tx	-	-	-	4
Status de Linfonodos (N)				
Positivo (N1, N2 e N3)	1	3	41	20
Negativo (N0)	2	6	76	48
Indeterminado (Nx)	-	-	1	6
Metástase a Distância (M)				
Positivo (M1)	3	5	6	5
Negativo (M0)	-	4	76	50
Indeterminado (Mx)	-	-	36	19
Estadiamento Patológico				
I	-	2	57	37
II	-	-	34	20
III	-	2	18	10
IV	3	5	6	6
ND	-	-	3	1

3.3. PROCESSAMENTO DAS AMOSTRAS CLÍNICAS

3.3.1. Extração de RNA

Amostras de tecido congelado (-80°C) de adenocarcinoma pulmonar de não fumantes e fumantes foram obtidas do Biobanco do HCFMB ($n=8$). A extração do RNA foi realizada através do protocolo híbrido, utilizando protocolo inicial do TRIzol (Invitrogen, MA, EUA) (Rio et al., 2010), seguido da purificação em coluna com o RNeasy Mini Kit (Qiagen, Hilden, Alemanha). Os procedimentos foram realizados segundo as etapas: (1) As amostras foram homogeneizadas utilizando o homogeneizador Precellys 24 (Bertin Technologies, França). Cada amostra recebeu 1 mL de Trizol por 0,1 g de tecido, em tubos contendo beads de cerâmica de 2,8 mm (CK28), e foi homogeneizada a 6.000 rpm por 30 segundos, em 2 ciclos de moagem. O material homogeneizado foi então incubado à temperatura ambiente por 5 minutos; (2) As amostra foram centrifugada por 10 minutos a $12.000\times g$ a 4°C para eliminar detritos e polissacarídeos; (3) O clorofórmio foi adicionado ao homogeneizado (0,2 ml de clorofórmio por ml de Trizol utilizado), agitado vigorosamente por 20 segundos. As amostras ficaram em repouso à temperatura ambiente por 2 minutos e centrifugada a $12.000\times g$ por 15 minutos a 4°C ; (4) A fase aquosa foi removida cuidadosamente e transferida para um novo tubo estéril, ao qual foi adicionado um volume equivalente de etanol 100% isento de RNA; (5) As amostras foram colocadas nas colunas RNeasy posicionadas em um tubo de coleta e centrifugadas por 30 segundos a $8.000\times g$ (descartar o líquido eluído); (6) 700 μL do tampão RW1 foi adicionado a coluna e centrifugado por 30 segundos a $8.000\times g$ (descartar o líquido eluído); (7) Foi adicionado 500 μL do tampão RPE a coluna e centrifugado por 30 segundos a $8.000\times g$ (descartar o líquido eluído); (8) O tampão RPE foi adicionado novamente às amostras, que foram centrifugadas por 2 minutos a $8.000 \times g$. Em seguida, após o descarte do eluído, a coluna foi centrifugada por 1 minuto em velocidade máxima para remover quaisquer resíduos de tampão; (9) A coluna foi transferida para um novo tubo de coleta de 1,5 mL, e 50 μL de água livre de RNase foram pipetados diretamente sobre a membrana da coluna. A amostra foi deixada em eluição à temperatura ambiente por 1–2 minutos e, em seguida, centrifugada por 1 minuto a $8.000 \times g$.

Após a extração, as amostras de RNA obtidas foram avaliadas quanto à quantidade de material genético e à qualidade da extração, utilizando o equipamento Nanodrop 8000 (Thermo Scientific). A integridade do RNA foi avaliada por meio do

sistema Agilent 2100 Bioanalyzer, empregando o kit RNA 6000 Nano (Agilent Technologies, CA, EUA), conforme as instruções do fabricante. Apenas as amostras com valores de RIN (RNA Integrity Number) superiores a 3 foram selecionadas para as etapas subsequentes de construção de biblioteca para RNA-seq.

As amostras de RNA selecionadas foram requantificadas utilizando o Qubit 4 Fluorometer, utilizando o kit RNA BR Assay (Invitrogen, MA, EUA) e tratadas com DNase I, RNase-free (Thermo Fisher Scientific, MA, EUA) para remover possíveis contaminantes de DNA, antes da montagem da biblioteca. Após tratamento, as amostras foram quantificadas novamente utilizando Qubit 4 Fluorometer e diluídas para a concentração de 20 ng/μl.

Já as amostras de tecido em bloco de parafina de adenocarcinoma pulmonar de não fumantes foram obtidas no Departamento de Patologia da FMB (n=5) e o RNA extraído utilizando o protocolo do kit RecoverAll Total Nucleic Acid Isolation (Invitrogen™). Os procedimentos foram realizados segundo as etapas: (1) Os tecidos foram cortados em seções finas e colocado em microtubos de 1,5ml; (2) Adicionamos 1ml de xilol 100% em cada microtubo; (3) As amostras foram vortexadas e incubadas por 3 minutos a 50°C; (4) As amostras foram centrifugadas por 2 minutos a 25°C a 14000rpm; (5) O xilol foi retirado e descartado; (6) Foi adicionado 1 ml de etanol 100% a temperatura ambiente e vortexado; (7) As amostras foram centrifugadas novamente por 2 minutos a 25°C a 14000rpm; (8) O etanol foi retirado e descartado e uma nova lavagem com etanol foi realizada e as etapas 6 e 7 foram feitas novamente; (9) O etanol foi removido e descartado e o pellet ficou a 37°C por 30 min; (10) Para a digestão de proteína, foi adicionado 200 μl do Digestion Buffer em cada amostra; (11) Foi adicionado 4 μl da protease e homogeneizado gentilmente; (12) As amostras ficaram incubadas por 14 horas a 50°C; (13) Para o isolamento dos ácidos nucléicos, foi adicionado 790 μl do mix composto por 240 μl do Isolation Additive e 550 μl do etanol 100%; (14) Foi realizada a homogeneização por inversão com a pipeta; (15) 700 μl das amostras foram adicionados no centro do filtro de cada coluna; (16) As amostras foram centrifugadas por 30 segundos a 10000g; (17) O líquido foi descartado e as colunas mantidas; (18) A etapa 15, 16 e 17 foram repetidas; (19) Foi adicionado 700 μl do Wash 1 no centro do filtro e centrifugado por 30 segundos a 10000g; (20) O líquido foi descartado e as colunas mantidas; (21) 500 μl do Wash 2/3 foi adicionado no centro do filtro e uma nova etapa de centrifugação foi realizada; (22) O líquido foi descartado e as colunas mantidas; (23) Foi feito um spin para remover o líquido

residual, o líquido foi descartado e as colunas mantidas; (24) Para a digestão e purificação final do RNA, foi adicionado 60 µl do mix composto por 6 µl do DNase Buffer 10X, 4 µl da DNase e 50 µl de água nuclease free; (25) As amostras foram incubadas por 30 minutos em temperatura ambiente; (26) Foi adicionado 700 µl do Wash 1 no centro do filtro e incubado por 30 segundos a temperatura ambiente; (27) As amostras foram centrifugadas por 30 segundos a 10000g; (28) O líquido foi descartado e as colunas mantidas; (29) As etapas 21 e 22 foram repetidas mais duas vezes; (30) Foi realizada uma centrifugação por 1 minuto a 10000g; (31) As colunas foram transferidas para um novo tubo coletor e 60 µl da água nuclease free foi adicionada ao centro da coluna; (32) As amostras foram incubadas por 1 minuto a temperatura ambiente e centrifugada por 1 minuto a 14000g; (33) As colunas foram descartadas e o RNA obtido foi quantificado no equipamento Nanodrop 8000 (Thermo Scientific) e posteriormente estocado em -80°C.

3.3.2. Preparação de bibliotecas e sequenciamento

As bibliotecas de cDNA foram construídas através do kit Illumina Stranded Total RNA Prep, ligation with Ribo-Zero Plus (Illumina, CA, EUA), utilizando 100ng de input de RNA tratado e seguindo as instruções do fabricante. Após a montagem, as bibliotecas foram analisadas quanto a qualidade e a quantidade, utilizando Agilent 2100 Bioanalyzer- DNA 1000 kit e o Qubit 4 Fluorometer- dsDNA BR Assay Kit. A concentração da biblioteca foi ajustada para 0,5 nM seguindo a orientação do fabricante. As amostras foram sequenciadas em pares no equipamento NovaSeq 6000 Sequencing System (Illumina), na Faculdade de Medicina de Ribeirão Preto-USP. A avaliação inicial da qualidade dos dados brutos de sequenciamento foi conduzida por meio da ferramenta FastQC (Andrews, 2010), a qual possibilitou detectar possíveis inconsistências técnicas, como regiões com baixa qualidade de bases, contaminação por adaptadores e níveis excessivos de duplicação de leituras. Posteriormente, o processamento dos dados foi conduzido com o pipeline nf-core/rnaseq (v3.8.1), assegurando um fluxo de análise padronizado, transparente e totalmente reprodutível. A remoção dos adaptadores foi executada com o Trim Galore (Krueger et al., 2023), e as leituras filtradas com qualidade satisfatória foram alinhadas ao genoma de referência GRCh38 utilizando o STAR, um alinhador sensível a junções intrônicas e amplamente otimizado para dados de RNA-seq. A etapa de quantificação gênica foi conduzida com o Salmon, que permite a estimativa precisa das abundâncias

de transcritos, corrigindo possíveis vieses inerentes ao processo de sequenciamento.

Por fim, os arquivos resultantes da quantificação foram processados com o pacote tximport (v1.22.0), com o objetivo de converter e agregar as estimativas em nível de transcrito em contagens consolidadas por gene, preparando os dados para as análises estatísticas subsequentes.

3.3.3. Análise em NanoString

Esta metodologia foi escolhida por permitir a quantificação confiável e sem o viés de amplificação da expressão gênica de múltiplos genes associados à resposta imune tumoral em diferentes tipos de amostras, incluindo tecidos de blocos de parafina e amostras frescas congeladas de pacientes não fumantes com adenocarcinoma pulmonar (5 amostras de metástase e 5 amostras de tumor primário).

Essas amostras foram submetidas à análise de expressão gênica através do painel nCounter PanCancer Immune Profiling, disponível no Nanostring Technologies, e seguindo as recomendações do fabricante do kit Gene Expression CodeSet RNA Hybridization Protocol, que inclusive recomenda a utilização desses 2 tipos de amostras. Este painel foi projetado para estudos em imuno-oncologia voltados ao entendimento dos processos de evasão imunológica no TME e desenvolvimento de assinaturas com aplicação prognóstica, preditivas ou de mecanismo de ação para imunoterapias. O equipamento Nanostring nCounter® está disponível para uso na UNIPLEX-FMB. Utilizamos o painel nCounter PanCancer Immune Profiling, que realiza análise multiplex de expressão gênica de 770 genes relacionados a diferentes tipos de células imunológicas, inibidores de checkpoint comuns, antígenos câncer-testes e genes relacionados às respostas imunes adaptativa e inata. Ao analisar simultaneamente diversos elementos da resposta imune, o painel facilita o desenvolvimento rápido de perfis de expressão gênica clinicamente relevantes, que podem ser úteis para abordagens de imunoterapia.

3.4. ANÁLISE DE DADOS

3.4.1. Expressão gênica

Os dados de RNA-Seq das amostras clínicas foram utilizados para a identificação de vias moleculares relacionadas ao perfil imune tumoral e sua associação com mecanismos de sinalização celular. A análise de expressão

diferencial foi realizada utilizando o pacote DESeq2 (Love; Huber; Anders, 2014), o qual emprega um modelo estatístico fundamentado na distribuição binomial negativa para a avaliação de dados de contagem. Os valores de expressão gênica foram submetidos à transformação rlog, procedimento que promove a estabilização da variância entre as amostras e reduz o impacto de diferenças de profundidade de sequenciamento. As estatísticas de expressão diferencial foram então estimadas no ambiente analítico integrado do DESeq2, assegurando consistência e rigor nos resultados obtidos.

O fator *smoker status* foi definido como variável explicativa no modelo estatístico, sendo o grupo *smoker* utilizado como condição de referência. Genes diferencialmente expressos foram definidos exclusivamente com base no critério estatístico de significância (p ajustado $< 0,05$), sem a aplicação de limiar arbitrário de fold change, de acordo com práticas atuais recomendadas na literatura. Os 50 genes diferencialmente expressos mais significativos, selecionados com base no menor valor de p , foram utilizados para a construção de um heatmap. A anotação funcional dos genes diferencialmente expressos foi realizada no ambiente R (versão 4.4.3) utilizando pacotes do Bioconductor (Gentleman et al., 2004). Os identificadores gênicos foram padronizados e anotados com informações de símbolo gênico, nome do gene e identificador Entrez por meio do pacote org.Hs.eg.db (Carlson, 2023), que fornece mapeamentos curados para genes humanos. A classificação dos genes quanto ao seu biotipo (por exemplo, *protein_coding*, *rRNA*, *lncRNA*) foi obtida de forma offline utilizando o banco EnsDb.Hsapiens.v86 (Rainer, 2017), garantindo reprodutibilidade da anotação.

A análise de enriquecimento funcional foi conduzida com o pacote clusterProfiler (Yu et al., 2012) empregando identificadores Entrez como entrada. Foram realizadas análises de enriquecimento de termos e categorias do Gene Ontology (GO) — *Biological Process*. O universo de fundo foi definido como o conjunto completo de genes testados na análise de expressão diferencial, assegurando correção estatística e evitando vieses de enriquecimento. A correção para múltiplos testes foi realizada pelo método de Benjamini–Hochberg, e os resultados foram visualizados por meio de gráficos de enriquecimento (*dotplots*) gerados com os pacotes enrichplot (Yu, 2024) e ggplot2 (Wickham, 2016).

Adicionalmente, dados de RNA-Seq do TCGA (<https://portal.gdc.cancer.gov/>) foram processados utilizando o pipeline limma/voom (Ritchie et al., 2015), adequado

para modelagem estatística de dados de contagem. A organização e o pré-processamento dos dados foram feitos com o pacote tidyverse (Wickham et al., 2019) enquanto a normalização das contagens foi conduzida com edgeR (Robinson; McCarthy; Smyth, 2010). O método voom (limma) foi aplicado para transformar as contagens e estimar a variância média–dependente, permitindo o uso de modelos lineares.

A comparação estatística foi realizada entre não fumantes e fumantes, utilizando fumantes como grupo de referência, de modo que valores positivos de fold change indicam maior expressão em não fumantes. Os genes diferencialmente expressos foram definidos com base apenas no p-valor ajustado (p ajustado $< 0,05$), sem aplicação de cutoff de fold change. A anotação funcional dos genes foi realizada com os pacotes AnnotationDbi (Pagès et al., 2025), org.Hs.eg.db (Carlson, 2023) e EnsDb.Hsapiens.v86 (Rainer, 2017), permitindo a obtenção de identificadores gênicos, nomes oficiais e classificação por biotipo de forma offline. A análise de enriquecimento funcional foi conduzida com o pacote clusterProfiler (Yu et al., 2012), utilizando como universo todos os genes testados no modelo estatístico. Foi avaliado o termo *Biological Process* do Gene Ontology, com correção para múltiplos testes. Por fim, os resultados foram visualizados por meio de heatmaps e gráfico de enriquecimento utilizando os pacotes pheatmap (Kolde, 2025), ggplot2 (Wickham, 2016) e enrichplot (Yu, 2024). Os resultados dessas análises foram posteriormente comparados para caracterização do perfil imune predominante em pacientes não fumantes.

A análise de expressão gênica foi realizada a partir de dados obtidos pela plataforma NanoString nCounter, comparando amostras de tumor primário e metástase. Os arquivos RCC foram importados e processados no ambiente R (versão 4.4.3), utilizando o pacote NanoTube (Class et al., 2022), que permite a leitura direta das contagens gênicas. Os metadados das amostras foram organizados com os pacotes readr (Wickham et al., 2026b), dplyr (Wickham et al., 2026a) e tidyr (Wickham et al., 2025), garantindo a correta associação entre amostras e grupos experimentais.

As contagens gênicas foram organizadas em uma matriz de expressão e estruturadas em um objeto ExpressionSet por meio do pacote Biobase (Huber et al., 2015). A normalização dos dados foi realizada de forma manual, seguindo o protocolo de normalização recomendado pela NanoString Technologies, adaptado para execução manual em R, a fim de garantir rastreabilidade das etapas. O procedimento

envolveu três fases principais: **1) Subtração de background:** média dos controles negativos subtraída das contagens de todos os genes; **2) Correção por controles positivos:** normalização das amostras pelo total de sondas positivas, ajustado à mediana global; e **3) Normalização por genes housekeeping:** cálculo da média geométrica de genes constitutivos (ex.: *GAPDH*, *GUSB*, *HPRT1*, *SDHA*, *TBP*) e aplicação de um fator de escala para equalização entre amostras.

A análise de expressão diferencial foi conduzida utilizando os pacotes edgeR (Robinson; McCarthy; Smyth, 2010) e limma (Ritchie et al., 2015). Inicialmente, o edgeR foi empregado para o cálculo de fatores de normalização, seguido da transformação voom, que estabiliza a variância dos dados. Em seguida, modelos lineares foram ajustados com o limma para identificar genes diferencialmente expressos entre os grupos (metástase *versus* tumor primário). Os valores de *p* foram corrigidos para múltiplos testes pelo método de Benjamini–Hochberg, considerando-se significativos os genes com *p* ajustado $< 0,05$ e $|\text{fold change}| > 1,5$. A escolha do valor do fold change justifica-se pela alta heterogeneidade celular intrínseca ao microambiente tumoral e pelo perfil difuso de alterações moleculares causadas pelo tabagismo. Valores maiores de FC poderiam excluir genes com modulações biológicas sutis, porém funcionalmente relevantes na carcinogênese induzida pelo tabaco, comprometendo a identificação de vias de sinalização críticas (Hammouz et al., 2020; Landi et al., 2008). A visualização dos resultados foi realizada por meio de volcano plots, heatmaps e boxplots, utilizando os pacotes ggplot2 (Wickham, 2016), pheatmap (Kolde, 2025) e ggpubr (Kassambara et al., 2026).

A interpretação funcional dos genes diferencialmente expressos foi conduzida por análises de enriquecimento de Gene Ontology (*Biological Process*) e vias do Reactome, utilizando os pacotes clusterProfiler (Yu et al., 2012), org.Hs.eg.db (Carlson, 2023), enrichplot (Yu, 2024) e ReactomePA (Yu; He, 2016). Adicionalmente, foi estimada a abundância relativa de tipos celulares com base em conjuntos de genes marcadores associados ao painel PanCancer Immune Profiling, por meio do cálculo de escores médios de expressão. As comparações entre os grupos foram realizadas utilizando o teste não paramétrico de Wilcoxon, e os resultados foram representados graficamente por boxplots.

3.4.2. Perfil Imune

Os dados obtidos do TCGA foram analisados por diferentes algoritmos

disponíveis na literatura: CIBERSORT (Newman et al., 2015), TIMER (Li et al., 2017), xCell (Aran; Hu; Butte, 2017), MCP-counter (Becht et al., 2016) e EPIC (Racle et al., 2017). O algoritmo CIBERSORT (cell-type identification by estimating relative subsets of RNA transcripts) (<http://cibersort.stanford.edu/>) é um método para caracterização da composição de tecidos complexos a partir do perfil de expressão gênica que é utilizado para estimar as proporções relativas de cada população celular de interesse. O algoritmo TIMER (“Tumor IMmune Estimation Resource”) (<http://cistrome.org/TIMER/>) é uma ferramenta que avalia o impacto clínico de seis diferentes células imunes (células B, células T CD4+, células T CD8+, neutrófilos, macrófagos e células dendríticas). A estimativa do infiltrado imune nos tumores incluídos na análise será obtida a partir do módulo Estimation da plataforma. Os genes selecionados serão correlacionados negativamente com a pureza tumoral (porcentagem de células tumorais de cada amostra) e será aplicado o método dos quadrados mínimos aos níveis de expressão dos genes de interesse para prever a abundância dos seis subtipos de células imunes. O xCell (<http://xCell.ucsf.edu/>) se baseia em análise de enriquecimento para amostra única, a fim de inferir as populações de 64 tipos de células imunes e estromais, incluindo células de resposta imune adaptativa e inata, células hematopoiéticas, células epiteliais e células da matriz extracelular. Para determinar o escore de abundância de cada tipo celular nas amostras, o xCell parte de um compêndio de 489 conjuntos de genes extraídos de diferentes estudos de larga escala. O método MCP-counter (“Microenvironment Cell Populations-counter”) (<http://cit.ligue-cancer.net/language/en/mcp-counter/>) é um pacote do R capaz de prever a abundância de 10 populações celulares (células T CD3+, células T CD8+, células T citotóxicas, células NK, células B, células originadas de monócitos, células dendríticas, neutrófilos, células endoteliais e fibroblastos) a partir do perfil de transcrição de tecidos humanos. Cada amostra é avaliada individualmente gerando escores independentes que são utilizados para comparações diretas de abundância de cada tipo celular em dados de tecido normal. O MCP-counter se baseia em marcadores transcriptômicos presentes em cada uma das populações celulares identificadas. O algoritmo EPIC é útil para estimar as frações de células imunológicas e tumorais a partir de dados de expressão gênica de tumores. Isto é feito ajustando perfis de referência de expressão gênica dos principais tipos de células não malignas e simultaneamente representando um tipo de célula não caracterizado sem conhecimento prévio sobre o mesmo (e.g., células de câncer em amostras de tumores

sólidos). Os perfis de expressão de genes de referência derivados dos principais tipos de células de infiltrado tumoral (i.e. subconjuntos imunes, células estromais e endoteliais). Após testar esses diferentes algoritmos nos dados de amostras do TCGA, os resultados de cada análise foram comparados e as vantagens e desvantagens de cada algoritmo foram avaliadas. Em seguida, o perfil imunológico dos tumores foi avaliado utilizando o(s) algoritmo(s) que apresentaram os resultados mais completos com um maior número de células imunes identificadas na etapa *in silico*. Para a comparação entre os grupos de pacientes não fumantes (CPNF) e fumantes, foi empregada a plataforma Orange Data Mining (Demsar et al., [S.d.]). As análises estatísticas foram realizadas no software GraphPad Prism, utilizando o teste de Mann-Whitney para verificar diferenças significativas entre os grupos quanto a abundância celular (considerando o valor de $p < 0,05$).

Já com os dados gerados pelo NanoString, foi possível inferir a composição imune do TME a partir da expressão gênica normalizada do mesmo painel NanoString. Essa análise foi conduzida independentemente da seleção de genes diferencialmente expressos — ou seja, considerando o perfil global de expressão imune (composição imune geral). O pipeline foi implementado no ambiente R (versão 4.4.3) com o uso dos pacotes NanoTube (Class et al., 2022), Biobase (Huber et al., 2015), dplyr (Wickham et al., 2026a), tidyr (Wickham et al., 2025), tibble (Müller; Wickham, 2016), readr (Wickham et al., 2026b), ggplot2 (Wickham, 2016) e ggpubr (Kassambara et al., 2026). Foram definidos painéis de marcadores gênicos característicos de cada tipo celular imune com base na literatura e no próprio PanCancer Immune Profiling Panel. Para cada tipo celular, foi calculado um “*Cell Type Score*”, representando a média da expressão normalizada dos genes associados àquela população. Exemplos de marcadores utilizados estão elencados na Tabela 2.

Tabela 2. Abundância relativa de tipos celulares com base em conjuntos de genes marcadores associados ao painel PanCancer Immune Profiling.

Tipo Celular	Genes representativos
CD8⁺ T cells	<i>CD8A, CD8B, GZMB, PRF1, GNLY</i>
CD4⁺ T cells	<i>CD4, ICOS, IL2RA, CCR7</i>
Tregs	<i>FOXP3, IL2RA</i>
NK cells	<i>NCR1, KLRD1, KLRK1, GNLY, PRF1</i>

Tipo Celular	Genes representativos
B cells	<i>MS4A1, CD19, CD79A, POU2AF1</i>
Macrófagos	<i>CD68, CD163, MRC1, MSR1</i>
Células dendríticas	<i>ITGAX, CD1C, IRF8</i>
Neutrófilos	<i>CEACAM8, S100A9, ELANE</i>
Endoteliais	<i>PECAM1, VWF, CDH5</i>
Fibroblastos	<i>FAP, PDGFRA</i>

Para a análise de expressão diferencial (DEG), “*Tumor Primário*” foi definido como o nível de referência do fator experimental. Logo, os valores de $FC \geq 1,5$ indicam superexpressão em amostras metastáticas, enquanto valores menores que este indicam subexpressão em amostras metastáticas. As diferenças na abundância relativa de cada tipo celular entre os grupos Tumor primário e Metástase foram avaliadas pelo teste de Wilcoxon não pareado. Os resultados foram representados por boxplots faceteados com indicação dos valores de p sobre cada gráfico. As análises e visualizações foram realizadas com os pacotes ggplot2 e ggpubr.

4. RESULTADOS

4.1. Expressão gênica e enriquecimento funcional dos dados de RNA-Seq obtidos do TCGA

A comparação entre amostras de pacientes fumantes e não fumantes diagnosticados com adenocarcinoma pulmonar resultou na identificação de 4093 genes diferencialmente expressos entre os grupos, com significância estatística de p ajustado $< 0,05$. A expressão aumenta ou diminuída desses genes referem-se à expressão em pacientes não fumantes. Dentre eles estavam 288 pseudogenes (167 processado, 49 não processado, 44 transcrito não processado, 15 transcrito processado, 7 unitário, 2 polimorfo e 4 transcrito unitário), 96 antisense, 13 imunoglobulinas (2 gene IG_C, 2 gene IG_D, 2 gene IG_V, 1 pseudogene IG_C e 6 pseudogene IG_V), 54 longo não codificante (51 lincRNA e 3 intronico sense), 34 microRNA, 22 miscâneo (miscRNA), 23 transcrito processado, 3372 genes codificadores de proteínas, 9 ribossômico (rRNA), 60 nuclear pequeno (1 scaRNA, 25 snoRNA e 34 snRNA), 2 receptor de célula T (TR_V) e 120 não identificados.

Dos 3372 genes codificadores de proteínas, 1681 apresentaram expressão aumentada e 1691 expressão reduzida em tumores de pacientes não fumantes em comparação com fumantes (Tabela Suplementar 3). Os top 20 genes foram elencados na Tabela 3.

Tabela 3. Top 20 genes codificantes diferencialmente expressos em amostras de tumor primário de adenocarcinoma pulmonar do banco TCGA, comparando pacientes não fumantes em relação aos fumantes (grupo de referência). Os genes foram selecionados com base no valor de p ajustado, utilizando o pacote *limma*, com aplicação do teste t moderado e critério de significância estatística de $p < 0,05$. Genes com expressão aumenta e expressão diminuída referem-se à expressão em pacientes não fumantes, sendo os genes com maior expressão destacados em vermelho e os com menor expressão em azul.

Gene ID	Gene	FC	padj
<i>GPR15</i>	G Protein-Coupled Receptor 15	0.515589541	1.60E-15
<i>LRRC31</i>	Leucine Rich Repeat Containing 31	1.991796848	1.48E-08
<i>C1QTNF7</i>	C1q And TNF Related Protein 7	1.503898393	1.65E-07
<i>CD1E</i>	CD1e Molecule	1.717149056	1.75E-07
<i>CD1C</i>	CD1c Molecule	1.642760005	1.75E-07
<i>CLEC4F</i>	C-Type Lectin Domain Family 4 Member F	1.440422259	1.75E-07
<i>AHRR</i>	Aryl-Hydrocarbon Receptor Repressor	0.702873688	2.87E-07
<i>CD207</i>	CD207 Molecule (Langerin)	1.940818236	3.36E-07
<i>CA10</i>	Carbonic Anhydrase 10	1.728360786	3.54E-07
<i>SLC40A1</i>	Solute Carrier Family 40 Member 1 (Ferroportin)	1.205388729	7.20E-07
<i>PRG4</i>	Proteoglycan 4	1.694786075	7.78E-07
<i>NR3C2</i>	Nuclear Receptor Subfamily 3 Group C Member 2 (Mineralocorticoid Receptor)	1.443047016	7.78E-07
<i>ZDHHC15</i>	Zinc Finger DHHC-Type Palmitoyltransferase 15	1.393902434	7.78E-07
<i>CA5B</i>	Carbonic Anhydrase 5B, Mitochondrial	1.18239218	1.26E-06
<i>EDA</i>	Ectodysplasin A	1.479864223	1.70E-06
<i>CD1A</i>	CD1a Molecule	1.824971176	1.70E-06
<i>ZBTB7C</i>	Zinc Finger And BTB Domain Containing 7C	1.551872881	1.70E-06
<i>GTPBP4</i>	GTP Binding Protein 4	0.863241974	1.70E-06
<i>CST5</i>	Cystatin D	1.550868288	1.70E-06
<i>SHISA3</i>	Shisa Family Member 3	1.854883803	1.70E-06

A análise de enriquecimento funcional evidenciou uma predominância de processos relacionados ao ciclo celular e à divisão celular, com destaque para fissão de organelas, segregação cromossômica, divisão nuclear e transição de fase do ciclo celular mitótico. Também foram enriquecidos termos associados à replicação do DNA, segregação de cromátides irmãs e organização do citoesqueleto mitótico, indicando uma atividade proliferativa elevada (**Error! Reference source not found.**).

O alto GeneRatio e os valores altamente significativos de p ajustado sugerem uma alteração robusta de vias fundamentais da proliferação celular, compatível com

processos de progressão tumoral e desregulação do ciclo celular.

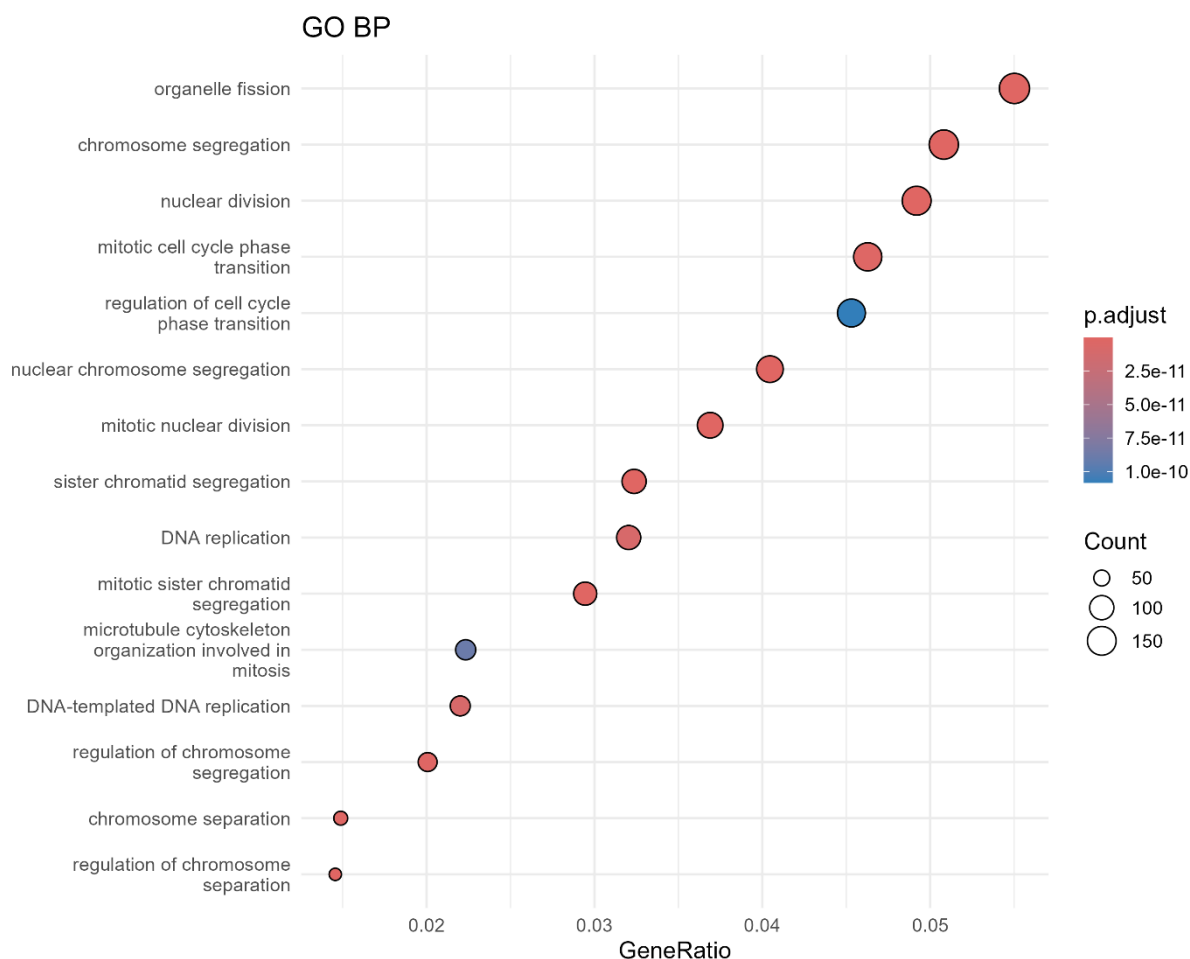


Figura 2. Principais Processos Biológicos (Gene Ontology – BP) enriquecidos a partir dos Genes Diferencialmente Expressos entre Fumantes e Não Fumantes resultantes da análise do TCGA.

4.2. Expressão gênica e enriquecimento funcional dos dados obtidos do RNA-Seq illumina

Utilizando a metodologia de RNA-Seq, foram analisadas sete amostras clínicas de metástase cerebral diagnosticadas como adenocarcinoma pulmonar, provenientes de pacientes fumantes ($n = 3$) e não fumantes ($n = 4$). A análise de expressão gênica diferencial (DEG) foi realizada considerando os pacientes fumantes como grupo de referencia e os pacientes não fumantes como grupo teste. Essa abordagem resultou na identificação de 257 genes diferencialmente expressos entre os grupos, com significância estatística de $p \text{ ajustado} < 0,05$. Dentre eles, 3 nuclear pequeno (1 snoRNA e 2 snRNA), 12 antisense, 1 imunoglobulina (IG_V), 15 longo não-codificante (lincRNAs), 23 pseudogenes (14 Pseudogenes processados, 7 Pseudogene não

processado e 1 Transcritos não processado), 5 transcrito processado, 1 ribossômico (rRNA), 2 overlapping sense, 159 genes codificadores de proteínas e 37 não identificados. Dos codificantes de proteínas, 85 estavam com expressão aumentada e 74 com expressão diminuída em tumores de pacientes não fumantes comparado com fumantes (Tabela Suplementar 4). Os top 20 genes foram elencados na Tabela 4.

Tabela 4. Top 20 genes codificantes diferencialmente expressos na análise de RNA-Seq de amostras clínicas de tumor primário de adenocarcinoma pulmonar, comparando pacientes não fumantes em relação aos fumantes (grupo de referência). Os genes foram selecionados com base no valor de p ajustado, utilizando o pacote *DESeq2*, com modelo baseado na distribuição binomial negativa e critério de significância estatística de $p < 0,05$. Genes com expressão aumentada e expressão diminuída referem-se à expressão em pacientes não fumantes, sendo os genes com maior expressão destacados em vermelho e os com menor expressão em azul.

Gene ID	Gene	FC	padj
<i>ABCF1</i>	ATP binding cassette subfamily F member 1	1.010273966	4.57003E-12
<i>RP11-577H5.5</i>	POLR2J2-UPK3BL1 readthrough	1.011145682	4.57003E-12
<i>TRIM26</i>	tripartite motif containing 26	1.010396009	4.57003E-12
<i>MUC21</i>	mucin 21, cell surface associated	1.009595702	4.57003E-12
<i>SFTA2</i>	surfactant associated 2	1.010136364	7.02221E-12
<i>UGT1A6</i>	UDP glucuronosyltransferase family 1 member A6	0.883525738	7.02221E-12
<i>ABHD16A</i>	abhydrolase domain containing 16A, phospholipase	1.010840048	7.02221E-12
<i>NEU1</i>	neuraminidase 1	1.009969297	3.88703E-11
<i>PIWIL3</i>	piwi like RNA-mediated gene silencing 3	1.010088753	9.60501E-11
<i>DDR1</i>	discoidin domain receptor tyrosine kinase 1	8.303197975	1.01903E-09
<i>UGT1A4</i>	UDP glucuronosyltransferase family 1 member A4	0.000430107	1.22162E-09
<i>UGT1A1</i>	UDP glucuronosyltransferase family 1 member A1	0.003389237	1.95741E-08
<i>PNP</i>	purine nucleoside phosphorylase	382.690612	2.05613E-08
<i>ZNF502</i>	zinc finger protein 502	85.28320583	1.02318E-07
<i>AKR1C2</i>	aldo-keto reductase family 1 member C2	0.933115503	1.55137E-06
<i>NELFE</i>	negative elongation factor complex member E	0.925245858	3.37767E-06
<i>ABC13-</i>	chromosome 13 open reading frame 46	697.6711192	5.086E-06

47488600 E17.1				
C4B_2	complement component 4B (Chido/Rodgers blood group), copy 2	337.4624607	8.72367E-06	
PCDHB11	protocadherin beta 11	3240.737727	3.74998E-05	
ABC13-47488600 E17.1	chromosome 13 open reading frame 46	0.892933015	7.57678E-05	

Além disso, foi gerado um heatmap (**Error! Reference source not found.**) que evidenciou a clusterização distinta dos dois grupos de pacientes, demonstrando características moleculares distintas no que se refere à expressão de genes específicos.

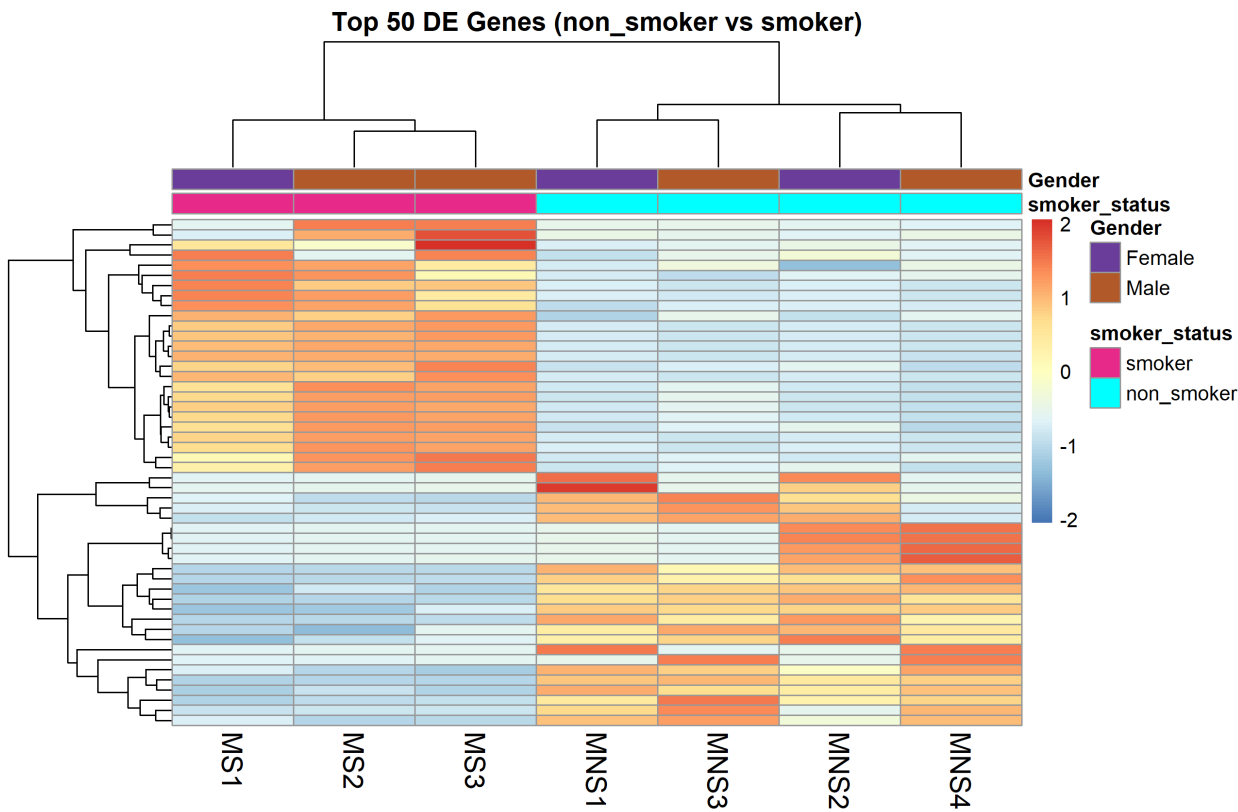


Figura 3. Heatmap mostrando a clusterização hierárquica não supervisionada das amostras de metástase cerebral de adenocarcinoma pulmonar de acordo com o histórico de tabagismo: Não Fumantes (MNS) vs. Fumantes (MS).

A análise de enriquecimento funcional para GO *Biological Process* (**Error! Reference source not found.**) demonstrou que a metástase do adenocarcinoma pulmonar em pacientes fumantes e não fumantes é marcada principalmente por alterações em vias metabólicas, com destaque para o metabolismo de pequenas

moléculas, corpos cetônicos, ácidos graxos e metabolismo secundário. Esses processos indicam uma reprogramação metabólica tumoral, sugerindo adaptação energética e metabólica das células tumorais em contexto metastático, possivelmente influenciada pelo status tabágico.

Adicionalmente, observou-se enriquecimento de vias relacionadas ao metabolismo de retinoides e ácido retinoico, bem como de isoprenoides e terpenoides, processos associados à regulação da diferenciação celular, sinalização oncogênica e progressão tumoral. A presença de termos ligados ao metabolismo de fármacos e xenobióticos também sugere diferenças na capacidade de detoxificação e potencial resposta terapêutica entre fumantes e não fumantes, reforçando a existência de uma assinatura metabólica distinta na metástase do adenocarcinoma pulmonar.

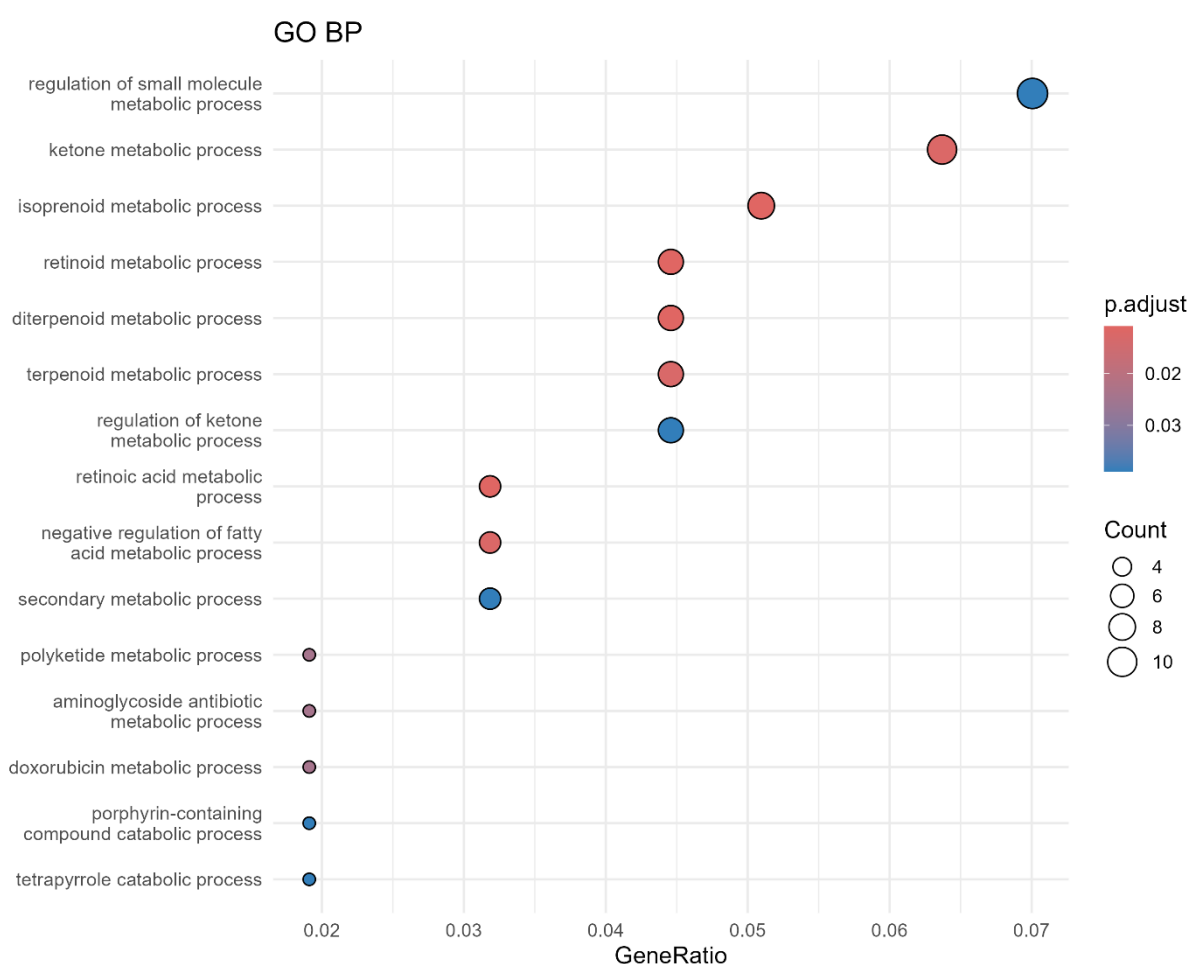


Figura 4. Principais Processos Biológicos (Gene Ontology – BP) Enriquecidos a partir dos Genes Diferencialmente Expressos entre Fumantes e Não Fumantes resultantes da análise do RNA-Seq das amostras clínicas.

4.3. Comparação dados de RNA-Seq (TCGA) com dados de RNA-Seq illumina

Comparando os genes codificadores de proteínas diferencialmente expressos entre pacientes não fumantes e fumantes obtidos da análise do banco de dados TCGA (tumor primário) com aqueles identificados por RNA-Seq nas amostras clínicas de metástase cerebral, foi observada sobreposição de 42 genes entre os conjuntos analisados. Dentre estes, 30 genes possuem padrão de expressão concordante entre tumor primário e metástase, mantendo-se aumentados ou diminuídos em tumores de pacientes não fumantes em ambas as condições. Esses achados sugerem a existência de uma continuidade biológica relevante no adenocarcinoma pulmonar deste grupo experimental, independentemente da progressão tumoral (Tabela Suplementar 5).

Além disso, a análise de sobreposição dos Genes Diferencialmente Expressos (DEGs) revelou um conjunto de 12 genes que exibem padrões de expressão opostos entre o sítio primário e o metastático (Figura 5). Estes genes evidenciam diferenças específicas entre tumor primário e metástase no grupo estudado, sugerindo que essas alterações estejam relacionadas a mecanismos que definem e direcionam a progressão tumoral do adenocarcinoma pulmonar.

No tumor primário de não fumantes, observou-se a super-expressão de marcadores de identidade tecidual e apresentação de antígenos (*SFTA2*, *HLA-DMB*). Contudo, na metástase cerebral, estes genes sofrem regulação negativa, enquanto emerge uma assinatura de adaptação ao nicho neural e resiliência citotóxica, evidenciada pelo aumento de reguladores de iniciação traducional e de vias inflamatórias (*ABCF1*), chaperonas de estresse (*HSPA1B*) e proteínas de estabilidade genômica (*MDC1*). Este padrão sugere um processo de reprogramação molecular ativa necessário para a transição do pulmão para o microambiente cerebral.

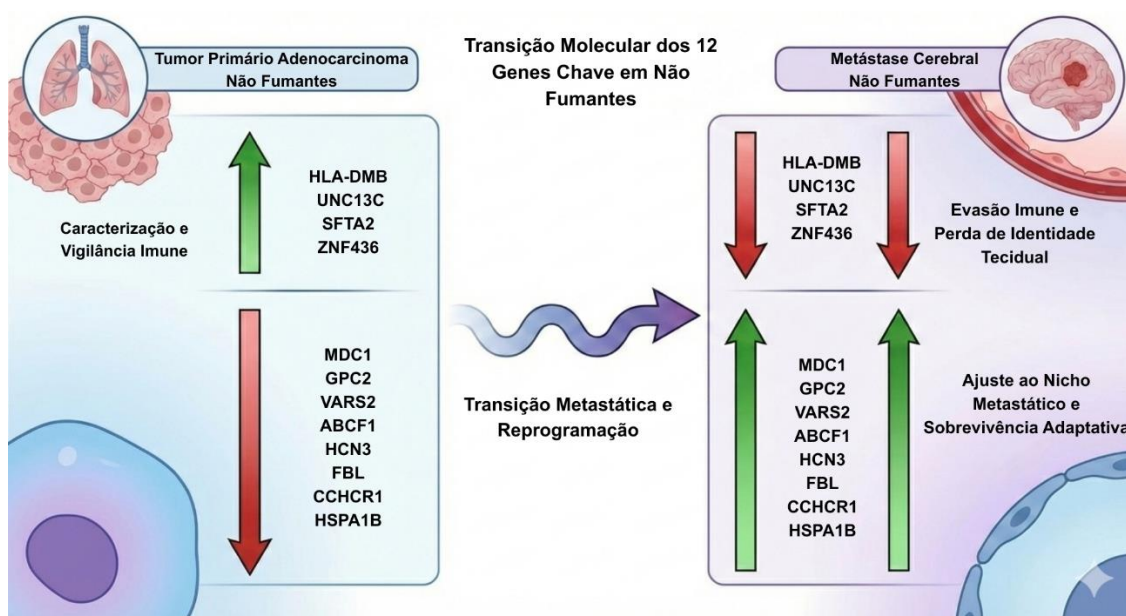


Figura 5. Dinâmica de Reprogramação Transcriptômica na Transição do Sítio Primário para o Nicho Metastático Cerebral. Figura elaborada com auxílio da ferramenta de inteligência artificial Google Gemini.

4.4. INFILTRADO IMUNE

4.4.1. Banco de dados (TGCA)

Dentre as cinco plataformas testadas, resultados significativos foram obtidos para duas plataformas: CIBERSORT e xCell. O CIBERSORT foi escolhido pois além de ser o mais citado na literatura, consegue detectar um grande conjunto de tipos de células e, como um método baseado em referência, calcula proporções também comparáveis entre os tipos de células (pelo fato de basear suas análises em uma matriz de assinatura gênica de leucócitos (LM22) que distingue 22 fenótipos de células hematopoiéticas humanas) entre elas células T-cell, células B naïve e memória, células plasmáticas, células natural killer (NK) e células mielóides (White et al., 2024). Já o xCell por utilizar em suas abordagens o enriquecimento, o que permite comparações dentro da amostra entre os tipos de células- engloba uma assinatura gênica de 64 tipos de células, que abrangem múltiplas células relacionadas a imunidade adaptativa e inata (35 tipos de células), progenitoras hematopoiéticas, células epiteliais e células da matriz extracelular-, ao mesmo tempo em que é robusto à contaminação por tipos de células desconhecidos não presentes nos dados de treinamento (White et al., 2024).

Através dos dados gerados pelo CIBERSORT, identificamos 12 populações celulares diferencialmente abundantes entre o grupo de pacientes fumantes (S) e não fumantes (NS): célula B de memória, célula T CD8+, célula T CD4+ memória de repouso, célula T CD4+ memória ativada, macrófagos M0, M1 e M2, célula mielóide dendrítica de repouso, célula mielóide dendrítica ativada, célula NK ativada, mastócito ativado e monócito (Figura 6). Sendo que as células com maior infiltração no grupo S são as células B plasmática ($p=0,02649$), célula T CD8+ ($p=7,573 \times 10^{-007}$), célula T CD4+ memória ativada ($p=0,01507$), célula NK ativada ($p=0,003153$), macrófagos M0 ($p=0,01956$) e macrófagos M1 ($p=0,0002765$). Já no grupo NS temos célula B de memória ($p=0,001310$), célula T CD4+ memória de repouso ($p=4,527 \times 10^{-005}$), macrófago M2 ($p=0,007808$), mastócito ativado ($p=3,049 \times 10^{-005}$), célula mielóide dendrítica de repouso ($p=1,130 \times 10^{-005}$), célula mielóide dendrítica ativada ($p=0,002039$).

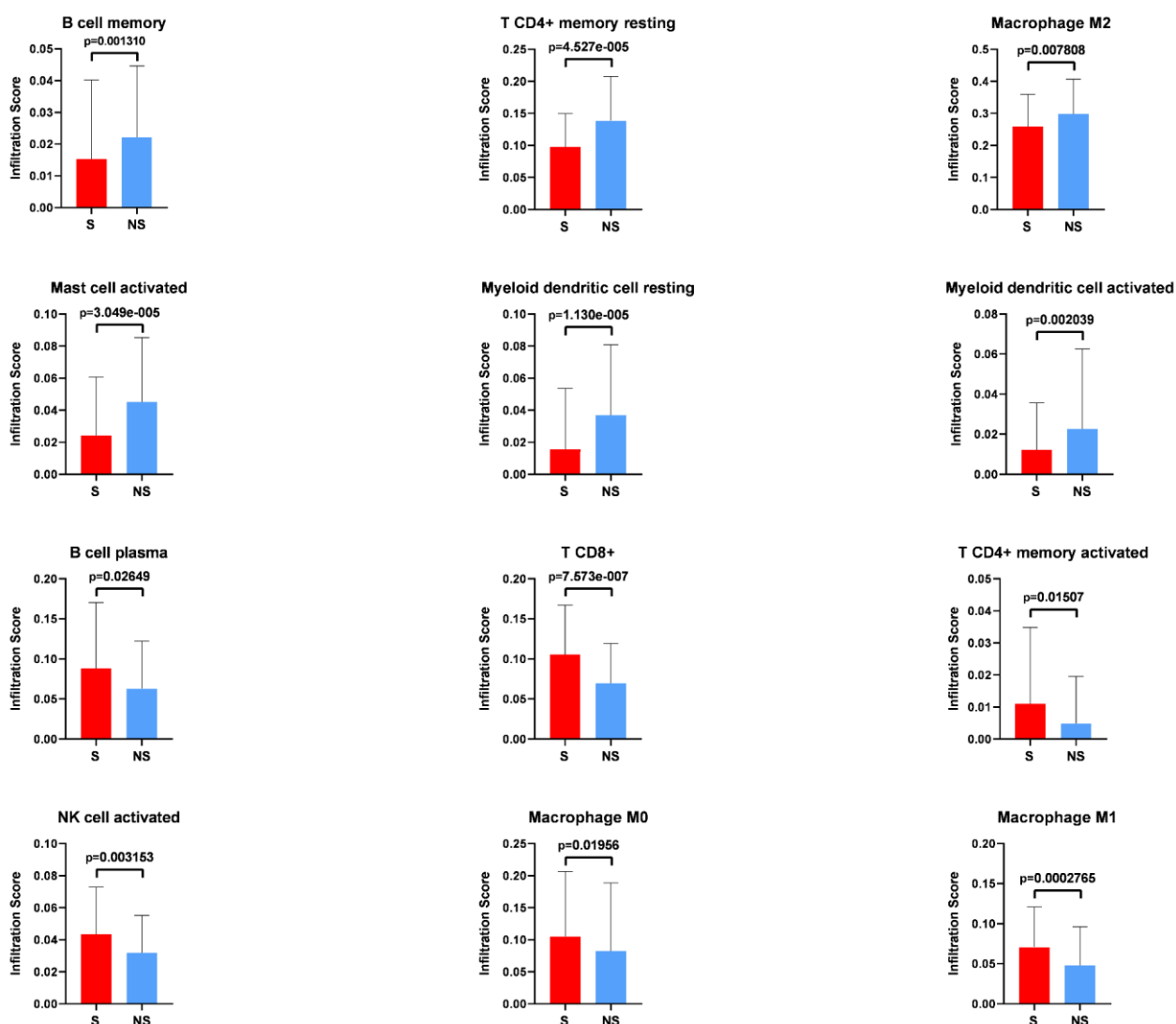


Figura 6. Estimativa da abundância relativa de diferentes tipos celulares imunes no TME de amostras de tumor primário de adenocarcinoma pulmonar, obtida por meio da deconvolução de dados de expressão gênica utilizando o algoritmo *CIBERSORT*. As barras vermelhas representam pacientes fumantes (S) e as barras azuis pacientes não fumantes (NS). Foram identificados 6 tipos celulares imunes com maior abundância em S, sugerindo um perfil mais inflamatório, enquanto 6 tipos celulares apresentaram maior abundância em NS, caracterizando um perfil mais associado à regulação da resposta imune.

Já com os dados do xCell, identificamos 19 populações celulares diferencialmente abundantes entre o grupo de pacientes fumantes (S) e não fumantes (NS): células T CD4+ de memória efetora, CD4+ Th1, CD4+ Th2, células mieloides dendríticas, células mielóides dendríticas ativadas, células dendríticas plasmocitoides, T CD8+ Naive, células NKT, células plasmáticas, Monócitos, Macrófagos M2, Tumor-associated fibroblast infiltration (CAFs), Progenitor linfóide comum, Progenitor mielóide comum, Progenitor de granulócitos-monócitos, Célula tronco hematopoiética, score imunológico, estromal e do microambiente. Como o foco do trabalho é a parte

imunológica, analisamos somente 12 dessas células (Figura 7), sendo que no grupo NS temos maior abundância de células mielóides dendríticas ativadas ($p=0,1261$), células T CD4+ de memória efetora ($p=8,522 \times 10^{-005}$), célula mielóide dendrítica ($p=1,059 \times 10^{-006}$), macrófagos M2 ($p=0,002364$), monócitos ($p=0,007551$), células NKT ($p=0,0008804$), CAFs ($p=2,314 \times 10^{-006}$). Já no grupo S temos células T CD8+ Naive ($p=0,004277$), células B plasmáticas ($p=0,01557$), células dendríticas plasmocitoides ($p=0,0008778$), células CD4+ Th1 ($p=1,170 \times 10^{-005}$) e CD4+ Th2 ($p=1,666 \times 10^{-006}$).

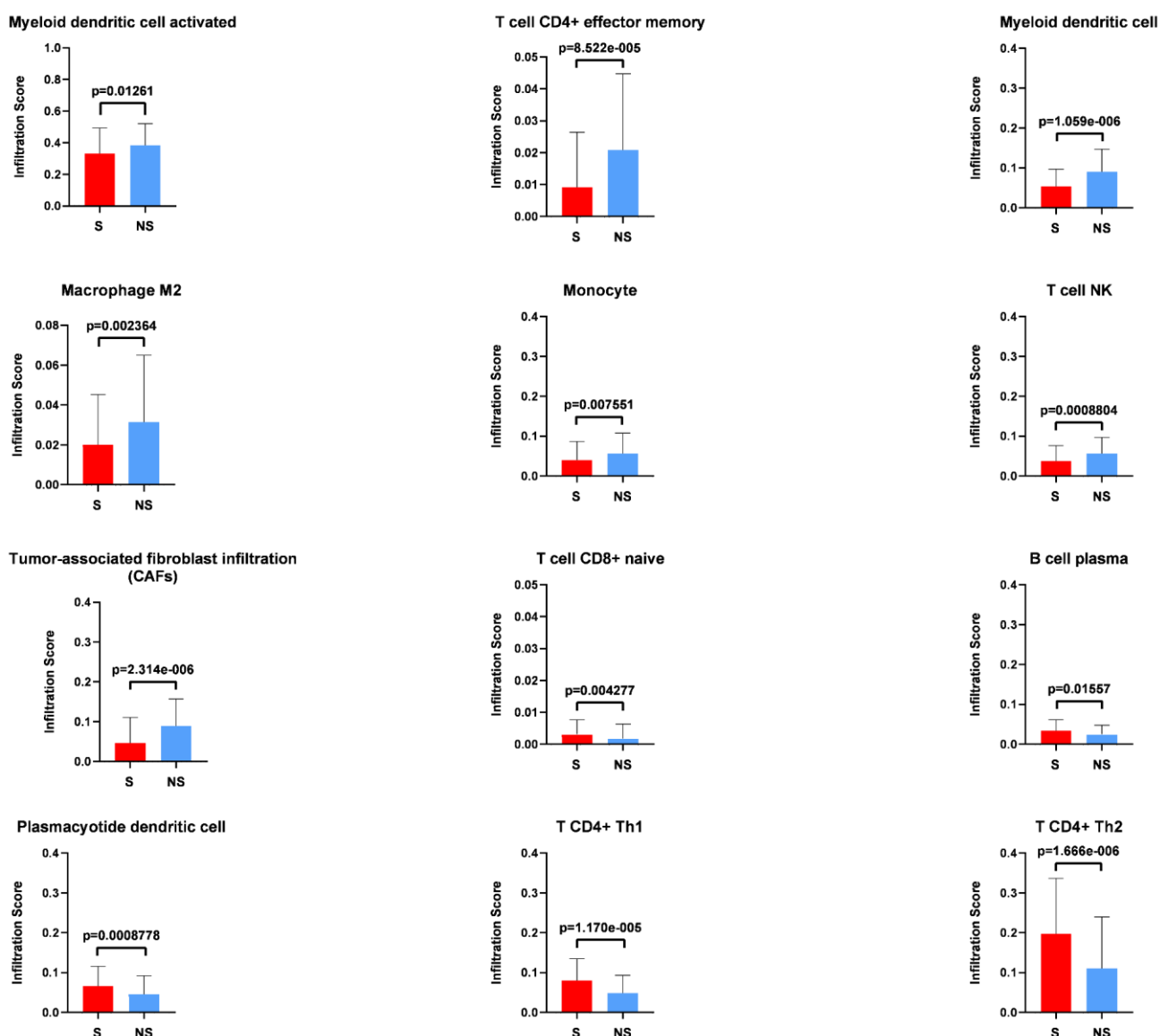


Figura 7. Estimativa da abundância relativa de diferentes tipos celulares imunes no TME de amostras de tumor primário de adenocarcinoma pulmonar, calculada pelo algoritmo xCell a partir de dados de expressão gênica. As barras vermelhas representam os pacientes fumantes (S) e as barras azuis os pacientes não fumantes (NS). Foram identificados 5 tipos celulares com maior abundância em S, associados a maior infiltração de células da imunidade adaptativa e indicando um perfil mais imunologicamente ativo, enquanto 7 tipos celulares apresentaram maior abundância em NS, sugerindo

um microambiente pró-tumoral, imunossupressor e relacionado ao remodelamento de matriz.

Comparando os 2 algoritmos, as células mais abundantes em pacientes NS são os Macrófagos M2 e as Células mieloides dendríticas ativadas, já as menos abundantes são as células B plasmáticas (Figura 8).

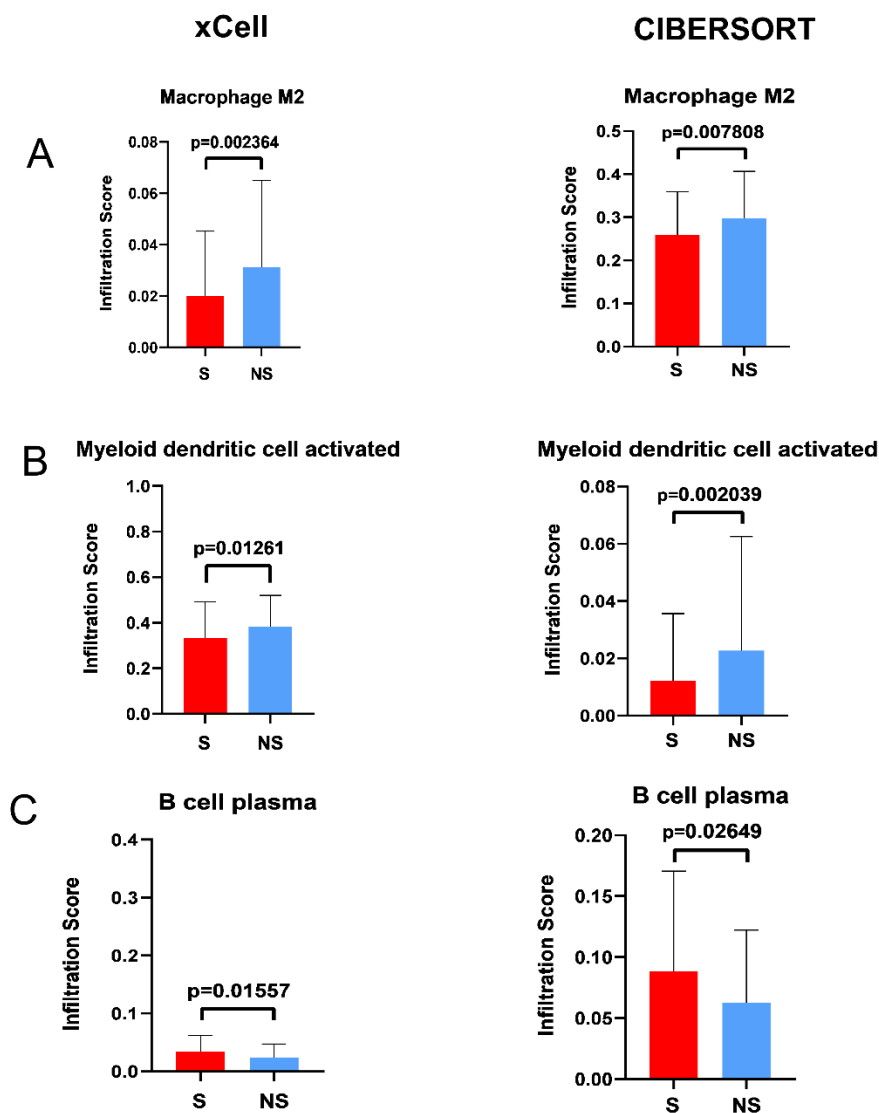


Figura 8. Comparação do Escore de Infiltração de células entre os dois algoritmos xCell e CIBERSORT, sendo que em A e B temos os tipos celulares mais abundantes em pacientes não fumantes (Macrófagos M2 e células mieloides dendríticas ativadas), e em C temos o menos abundante (células B plasmáticas).

Fizemos uma análise de predição na plataforma Orange Data Mining utilizando os dados obtidos com os dois algoritmos imunes (CIBERSORT e xCell). Analisando os dados do CIBERSORT (através do classificador probabilístico Naive Bayes), obtivemos uma matriz de confusão que indicou uma precisão de 68,9% e 73,7% para a correta identificação de pacientes não fumantes e fumantes, respectivamente (Figura 9).

		Predicted		Σ
		Non Smoker	Smoker	
Actual	Non Smoker	68.9 %	31.1 %	74
	Smoker	26.3 %	73.7 %	118
Σ		82	110	192

Figura 9. Representação da análise de predição. Matriz de confusão demonstrando que em 68,9% e 73,7% das vezes houve identificação correta de pacientes não fumantes e fumantes, respectivamente.

Já com os dados do xCell, obtivemos uma matriz de confusão em que os pacientes não fumantes foram identificados corretamente em 67,6% das vezes e pacientes fumantes foram identificados corretamente em 74,6% das vezes (Figura 10).

		Predicted		Σ
		Non Smoker	Smoker	
Actual	Non Smoker	67.6 %	32.4 %	74
	Smoker	25.4 %	74.6 %	118
Σ		80	112	192

Figura 10. Representação da análise de predição. Matriz de confusão demonstrando que em 67,6% e 74,6% das vezes houve identificação correta de pacientes não fumantes e fumantes, respectivamente.

4.4.2. Abundância celular nanostring

A deconvolução dos tipos celulares, baseada no NanoString nCounter® PanCancer Immune Profiling Panel, revelou diferenças marcantes na composição imune entre as amostras de adenocarcinoma pulmonar primário (LUAD) e metástases cerebrais de pacientes não fumantes. Dentre os 14 tipos celulares imunes avaliados, três populações apresentaram diferenças estatisticamente significativas entre os grupos no teste não paramétrico de Wilcoxon (p ajustado $< 0,05$): linfócitos T CD8⁺, células dendríticas e células NK (Figura 11). As amostras de tumores primários exibiram escores de linfócitos T CD8⁺, células NK e células dendríticas significativamente mais elevados em comparação às metástases ($p = 0.016$), sugerindo um TME mais citotóxico e imunologicamente ativo no sítio pulmonar primário.

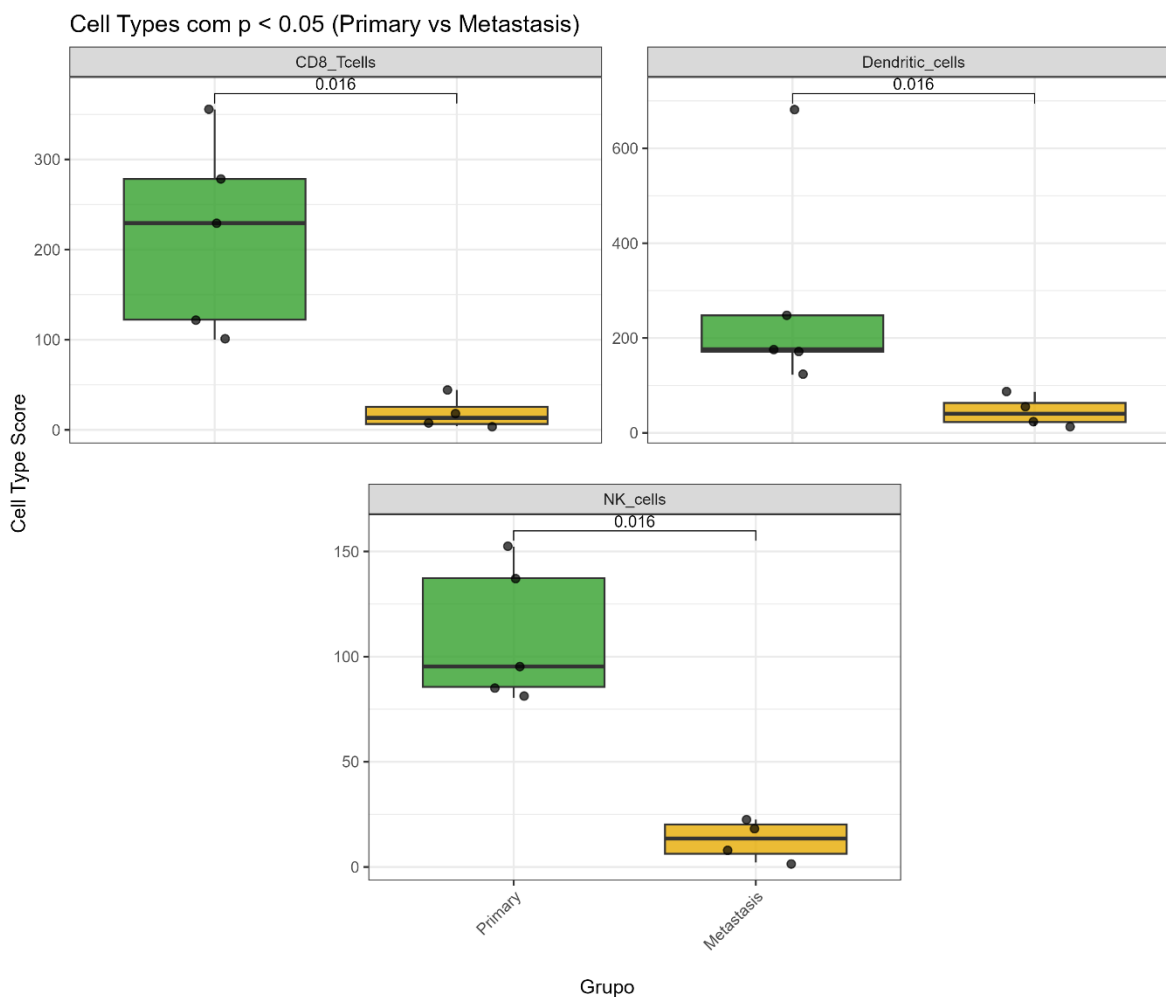


Figura 11. Populações celulares com diferenças estatisticamente significativas entre os grupos tumor primário e metástase ($p < 0,05$; teste de Wilcoxon): linfócitos T CD8⁺, células dendríticas e células NK.

4.5. Expressão gênica do nanostring

Utilizando o pacote de análise Nanotube, foi realizado um pipeline incluindo controle de qualidade, correção de background normalização e análise de expressão gênica diferencial. A comparação entre adenocarcinoma pulmonar primário e metástase cerebral, ambos provenientes de pacientes não fumantes, resultou na identificação de 17 genes diferencialmente expressos entre os dois estágios tumorais (Tabela 5). Desses genes, 1 apresentou expressão aumentada e 16 apresentaram expressão diminuída em metástases cerebrais de pacientes não fumantes, tendo como referência as amostras de tumor primário, considerando significância estatística de p ajustado $< 0,05$ e *fold change* $\geq 1,5$. Os processos biológicos enriquecidos indicam predominância de vias associadas à ativação e regulação da resposta imune adaptativa, incluindo ativação de linfócitos, imunidade mediada por células T e B e

sinalização por receptores de superfície. Esses achados são coerentes com a redução da expressão de genes observada nas metástases cerebrais em relação ao tumor primário em pacientes não fumantes, especialmente aqueles relacionados à ativação e coestimulação de células T (*ITK*, *CD27*, *ICOS*, *TNFRSF4*), citotoxicidade imune (*PRF1*, *GNLY*, *CD8B*) e sinalização por citocinas (*IL2RB*). Adicionalmente, genes envolvidos na imunidade inata e no recrutamento celular (*CCR2*, *NOD2*, *FPR2*, *LTF*) e na ativação de células B (*BTK*, *CD70*) também contribuíram para o enriquecimento de vias inflamatórias e de resposta imune. Em contraste, o aumento da expressão de *CDH1* nas metástases sugere a ativação de processos relacionados à adesão celular e plasticidade tumoral. Em conjunto, esses resultados apontam para um enfraquecimento da resposta imune efetora nas metástases cerebrais, acompanhado por mecanismos associados à adaptação tumoral ao microambiente do sistema nervoso central (Figura 12).

Tabela 5. Genes diferencialmente expressos em pacientes não fumantes, comparando metástases cerebrais em relação ao tumor primário (grupo de referência). A análise diferencial foi realizada utilizando o pipeline *limma-voom*, considerando significância estatística com p ajustado $< 0,05$ e *fold change* $\geq |1,5|$. Foram identificados 17 genes, sendo que 1 apresentou expressão aumentada (*CDH1*) e 16 expressão diminuída nas metástases cerebrais em comparação ao tumor primário.

Gene ID	Gene	Linear fold change	adj.P.Val
<i>RRAD</i>	RRAD, Ras Related Glycolysis Inhibitor And Calcium Channel Regulator	0.034725777	0.014049677
<i>ITK</i>	IL2-Inducible T-Cell Kinase	0.034963177	0.028004094
<i>CD27</i>	CD27 Molecule	0.045073196	0.028004094
<i>LTF</i>	Lactotransferrin (Lactoferrin)	0.041100786	0.041544541
<i>CD70</i>	CD70 Molecule	0.052026046	0.041544541
<i>IL2RB</i>	Interleukin 2 Receptor Subunit Beta	0.074206568	0.041544541
<i>NOD2</i>	Nucleotide-Binding Oligomerization Domain Containing 2	0.067403335	0.041550356
<i>CD8B</i>	CD8b Molecule	0.045954544	0.047074564
<i>CCR2</i>	C-C Motif Chemokine Receptor 2	0.053858679	0.047074564
<i>BTK</i>	Bruton Tyrosine Kinase	0.089613712	0.047074564
<i>IL11</i>	Interleukin-11	0.061952618	0.047074564
<i>TNFRSF4</i>	TNF Receptor Superfamily Member 4	0.096413168	0.048951421
<i>GNLY</i>	Granulysin	0.074859956	0.048951421
<i>FPR2</i>	Formyl Peptide Receptor 2	0.064794052	0.048951421
<i>ICOS</i>	Inducible T-Cell Costimulator	0.063424372	0.048951421
<i>PRF1</i>	Perforin	0.103979863	0.048951421
<i>CDH1</i>	Cadherin 1 (E-Cadherin)	7.58301298	0.048951421

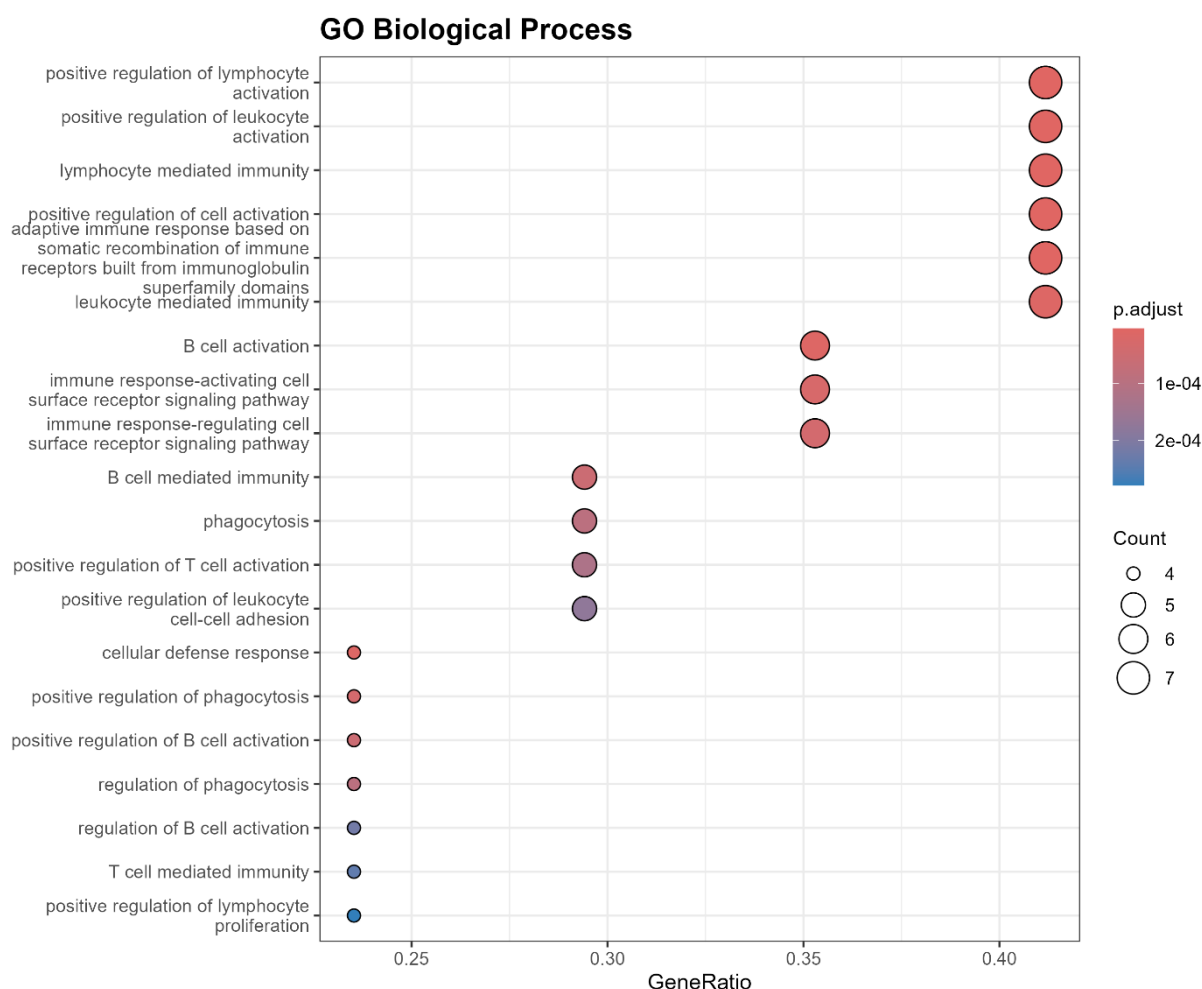


Figura 12. Principais Processos Biológicos (Gene Ontology – BP) Enriquecidos a partir dos Genes Diferencialmente Expressos entre Metástase e Tumor Primário em pacientes não fumantes com adenocarcinoma pulmonar resultantes da análise do NanoString.

O Heatmap (Figura 13) evidenciou a clusterização distinta dos dois estágios tumorais: metástase e tumor primário. Além do sítio tumoral (Primário vs. Metástase Cerebral), o estágio clínico do tumor primário (IA2, IIIA e IV) também influenciou o padrão de clusterização. Observou-se que tumores primários em estágios mais avançados apresentaram maior proximidade com as amostras metastáticas, sugerindo uma convergência molecular associada à progressão tumoral. Embora amostras de diferentes estágios clínicos estejam representadas, a clusterização ocorreu predominantemente de acordo com o sítio tumoral (Primário vs. Metástase Cerebral), sugerindo que as alterações da progressão metastática exercem maior influência na separação dos grupos.

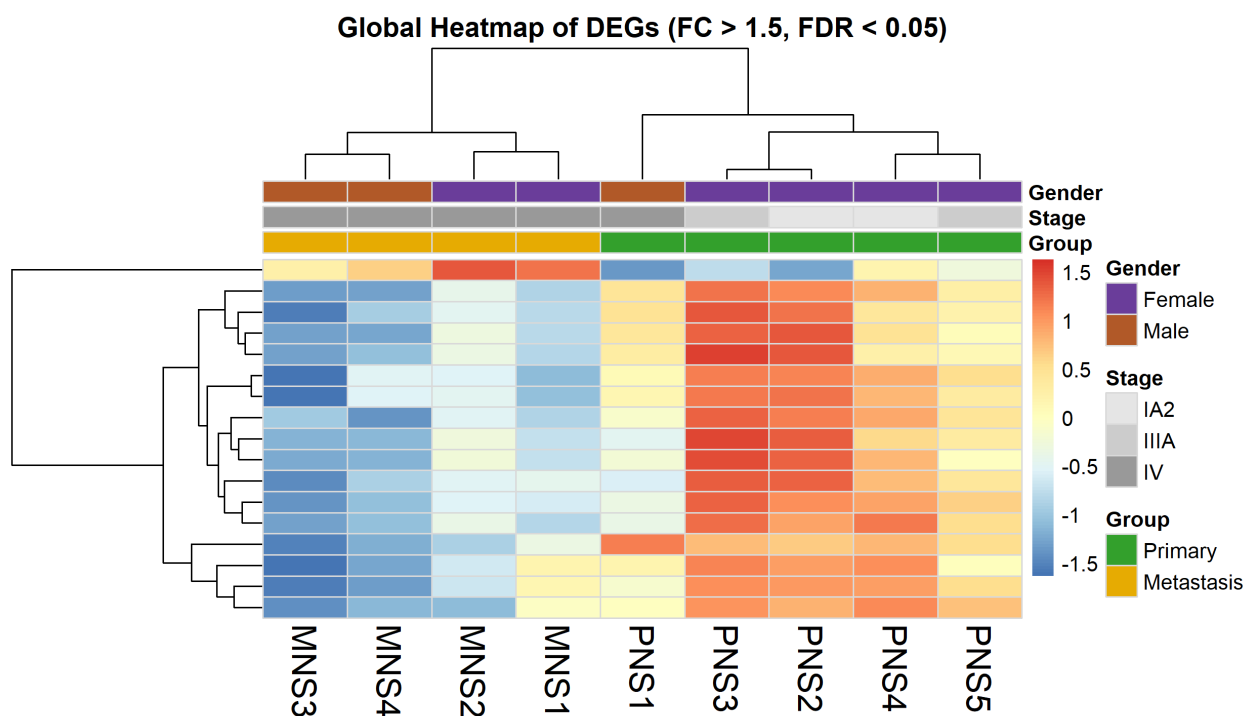


Figura 13. Heatmap mostrando a clusterização hierárquica não supervisionada das amostras de adenocarcinoma pulmonar de pacientes não fumantes de acordo com o sítio: Tumor Primário (PNS) vs. Metástase (MNS).

5. DISCUSSÃO

5.1. Resultados do sequenciamento illumina: perfil transcriptômico e imune de tumores primários e metástase cerebral de adenocarcinoma pulmonar de pacientes não fumantes

A análise da expressão gênica é uma abordagem fundamental para decifrar os mecanismos moleculares do câncer, e neste estudo, a investigação do perfil transcriptômico do adenocarcinoma de pulmão em não fumantes (CPNF) revelou assinaturas moleculares específicas que refletem adaptações funcionais cruciais para a biologia tumoral. A análise transcriptômica foi conduzida a partir de dados públicos do The Cancer Genome Atlas (TCGA-LUAD), incluindo tumores primários de adenocarcinoma pulmonar de pacientes fumantes (S, n=118) e não fumantes (NS, n=74), além de amostras clínicas obtidas de metástase cerebral (NS = 4 e S =3) submetidas a sequenciamento de RNA pela plataforma Illumina.

Foram identificados 4.093 genes diferencialmente expressos (DEGs) (p ajustado < 0,05), com predomínio de genes codificadores de proteínas (3.372 genes), sendo que 1.681 apresentaram expressão aumentada e 1.691 expressão reduzida em

tumores de pacientes não fumantes em comparação com fumantes de acordo com os dados do TCGA. Os top 20 identificados mostraram-se em sua maioria relacionados a processos imunes, regulação metabólica e controle da proliferação celular, três eixos centrais na biologia do adenocarcinoma pulmonar. No grupo de pacientes não fumantes, observamos a expressão aumentada de genes associados à apresentação de antígenos lipídicos (*CD1A*, *CD1C*, *CD1E*), bem como marcadores de células dendríticas especializadas (*CD207*), sugere enriquecimento de populações dendríticas no TME (Gunbin et al., 2024). No entanto, o aumento de macrófagos M2 e fibroblastos associados ao tumor, apontam para um fenótipo imuno regulador, no qual a apresentação de antígenos podem não estar associados a uma resposta efetora eficiente. Em contraste, os genes *GPR15* e *AHRR*, apresentam expressão reduzida em não fumantes, reforçando sua natureza como biomarcadores clássicos da exposição ao tabaco. Enquanto o *GPR15* atua no recrutamento de linfócitos T, o *AHRR* funciona como um marcador da resposta a xenobióticos via mecanismos epigenéticos, ambos consolidando uma assinatura inflamatória induzida pelo cigarro no contexto dos tumores de fumantes (Andersen et al., 2021; Tantoh et al., 2019). Por fim, o eixo metabólico, revela adaptações essenciais à sobrevivência tumoral. A regulação positiva, em tumor primário de não fumantes, de genes ligados ao equilíbrio do pH celular e à homeostase do ferro, como as anidrases carbônicas (*CA5B*, *CA10*) e o transportador *SLC40A1*, apontam para adaptações metabólicas distintas entre os grupos, revelando estratégias distintas de reprogramação bioenergética e resistência ao estresse oxidativo, além de favorecer um TME permissivo, contribuindo para imunossupressão e remodelação estromal (Pastorekova; Gillies, 2019; Zhang et al., 2023, 2024a).

O conjunto das alterações transcricionais identificadas revela que a distinção entre pacientes fumantes e não fumantes é estruturada em eixos funcionais divergentes de resposta imune e regulação metabólica, os quais influenciam diretamente as etapas da progressão tumoral. Sugerem que o tumor primário de não fumantes apresenta um microambiente com atividade de apresentação de antígenos, porém essa pode não se traduzir em uma resposta imune efetora eficiente. Apesar desse perfil imunologicamente robusto, sua manutenção depende de mecanismos de plasticidade celular e controle da homeostase iônica para garantir adaptação a microambientes metabolicamente adversos. Em contraste, o contexto do tabagismo é marcado por inflamação crônica e resposta exacerbada a xenobióticos, favorecendo

instabilidade genômica e remodelação agressiva do TME. Enquanto nos fumantes o processo é classicamente impulsionado por dano genotóxico e alta carga mutacional (Alexandrov et al., 2020), o adenocarcinoma em não fumantes parece evoluir sob uma arquitetura distinta, na qual reprogramação bioenergética e evasão imune moldam a progressão. Assim, as diferenças entre os grupos vão além de mutações driver, refletindo configurações específicas de vias imunes e metabólicas que sustentam o crescimento e a resiliência tumoral em cada cenário clínico.

O enriquecimento funcional do TCGA (Gene Ontology - Biological Process) demonstrou que os processos mais significativamente representados entre os genes diferencialmente expressos envolvem segregação cromossômica, fissão de organelas, replicação de DNA e organização do citoesqueleto mitótico. Estudos prévios de Dasari et al. (2024) e Ernst et al. (2023) revelam que tumores de adenocarcinoma pulmonar associados ao tabagismo apresentam ativação exacerbada desses programas e maior instabilidade cromossômica. Consequentemente, esses tumores operam em um regime de alta taxa proliferativa e podem manifestar maior agressividade clínica observada em pacientes fumantes (Dasari; Alfaro-Murillo; Townsend, 2024; Ernst et al., 2023). Embora esses programas estejam mais pronunciados nos tumores de fumantes, sua menor representação nos tumores de não fumantes sugere que estes últimos podem apresentar menor dependência de vias clássicas de alta taxa proliferativa e instabilidade cromossômica (Díaz-Gay et al., 2025). Além disso, a ativação persistente desses eixos proliferativos está associada à deriva clonal e à heterogeneidade intratumoral (Fu et al., 2025; Wang et al., 2020), reforçando a hipótese de que o adenocarcinoma em não fumantes não constitui apenas uma variante clínica, mas sim uma entidade molecularmente distinta, possivelmente sustentada por mecanismos alternativos de progressão tumoral, como alterações driver específicas, remodelação do TME ou vias de sinalização oncogênica independentes da carga mutacional associada ao tabaco (Chen; Huang; Li, 2026; LoPiccolo et al., 2024).

A integração desses dados com a análise dos algoritmos imunes, o CIBERSORT e o xCell, demonstrou uma contradição no TME. Segundo a LV et al. (2022) o TME é povoado por diversos tipos de células imunes, que podem ser classificadas em células imunes de defesa tumoral e de crescimento tumoral (LV et al., 2022). As células de defesa tumoral incluem citotóxicas promotoras de citocinas (como IFN- γ , TNF- α , IL-6, IL-8 e IL-12), tendo por integrantes as células TCD8⁺/CD4⁺

efetoras, células NK, dendríticas (DCs) e macrófagos M1. Já as células associadas ao crescimento tumoral são representadas pelas Tregs e pelos macrófagos M2, que promovem a homeostase e tolerância periférica, inibindo a ação das células antitumorais, favorecendo o escape imune (Böttcher; Reis e Sousa, 2018; Cózar et al., 2021; Farhood; Najafi; Mortezaee, 2019; Goswami; Bose; Baral, 2021; Sakaguchi et al., 2008). Nesse contexto, faz-se necessário o estudo do TME principalmente em tumores de pacientes não fumantes, pois eles possuem diferenças genéticas e microambientais que podem influenciar a progressão tumoral (Sui et al., 2020).

Analisando as amostras de tumores primários de indivíduos que nunca fumaram (TCGA) observamos os seguintes subtipos celulares: no xCell foram identificados células mieloides dendríticas ativadas, mieloides dendríticas, T CD4+ de memória efetora, macrófagos M2, monócito, células NKT e fibroblastos associados ao câncer (CAFs); já no CIBERSORT foram observadas células B plasmáticas de memória, T CD4+ de memória em repouso, macrófagos M2, células mieloides dendríticas ativadas e mastócitos. De maneira geral, os subtipos celulares identificados em não fumantes são compatíveis com um ambiente imunossupressor e estromal. Já nos pacientes fumantes, encontramos os seguintes subtipos celulares: no xCell células dendríticas plasmocitoides, células B plasmáticas, células TCD8+ naive, TCD4+ Th1 e Th2; já no CIBERSORT células NK ativadas, macrófagos M1 e M0, TCD4+ de memória ativada, TCD8+ e células B plasmáticas. Assim, os dados sugerem que tumores de pacientes fumantes apresentam um perfil imune mais ativo e efetor, caracterizado por células NK ativadas, macrófagos M1, linfócitos TCD8+ e TCD4+ (Th1 e Th2)- possivelmente impulsionado pela alta carga mutacional induzida pelo tabaco, que gera neoantígenos capazes de ativar o sistema imune adaptativo (Rizvi et al., 2015). Contudo, após a eliminação das células mais imunogênicas, há equilíbrio e evasão tumoral (Mellman et al., 2023; Thommen; Schumacher, 2018).

Em contrapartida, o tumor de não fumante exibe um perfil imune mais remodelado. Nossos achados identificaram um aumento significativo de macrófagos M2, células dendríticas ativadas e fibroblastos associados ao câncer (CAFs), acompanhado por uma redução de células B plasmáticas em tumores de pacientes não fumantes (NS). A polarização de macrófagos para o fenótipo M2 a presença de CAFs são marcadores clássicos de um microambiente imunossupressor e pró-tumoral, facilitando tolerância imune e crescimento tumoral em contextos menos proliferativos (Binnewies et al., 2018; Cao et al., 2024; Chen; McAndrews; Kalluri,

2021). Essa configuração imune sugerida, com acúmulo de macrófagos M2 e CAFs, está associada a um TME permissivo que favorece a evasão imune, sobrevivência celular e progressão tumoral, particularmente em contextos de baixa atividade proliferativa, como observado em adenocarcinomas de não fumantes. Estudos recentes mostram que a interação CAF–macrófago M2 contribui resulta em nichos imunossupressores, por meio de secreção de fatores como *TGF- β* , *CXCL12* e *IL-6*, que suprimem a resposta efetora de linfócitos T e facilitam a tolerância imune (Cao et al., 2024; Khalaf et al., 2021; Liu et al., 2025).

Já as células dendríticas exercem papel de mediadoras entre a imunidade inata e adaptativa, reconhecendo e processando os antígenos e apresenta-os para os linfócitos T (nos órgãos periféricos e no caso do câncer, para linfócitos TCD8+), desencadeando então respostas específicas por contatos célula-célula e produção de citocinas (Alattas; Alkhatabi; Alzahrani, 2020; Wculek et al., 2020). Elas dividem-se em dois principais subtipos: células dendríticas mieloides (mDCs) e as células dendríticas plasmocitóides (pDCs). O equilíbrio entre os subtipos é importante para a homeostase imunológica, já que as mDCs atuam na ativação de células T e no desencadeamento de respostas imunes efectoras contra patógenos e células tumorais, enquanto as pDCs produzem grandes quantidades de interferon tipo I, especialmente o IFN- α , induzindo tolerância e suprimindo respostas inflamatórias excessivas (Alattas; Alkhatabi; Alzahrani, 2020). Assim, alterações na proporção desses subtipos podem indicar mudanças funcionais importantes no sistema imune e estão associadas a diferentes processos patológicos, incluindo o câncer. Segundo Alattas et al. 2020, no tabagismo há um aumento de pDCs e diminuição relativa de mDCs, e isso gera uma maior imunossupressão e um risco aumentado de câncer (Alattas; Alkhatabi; Alzahrani, 2020).

A presença de células dendríticas ativadas concomitantemente a macrófagos M2 indica uma tentativa de vigilância imune que é, no entanto, modulada por sinais supressores do TME (Lu et al., 2025). Este achado é particularmente relevante para a compreensão da progressão metastática. De acordo com Mughal et al. (2025) e Lu et al. (2025), há uma transição gradual de um fenótipo imuno-ativo no tumor primário para um microambiente imune-frio em sítios metastáticos, particularmente no cérebro, caracterizado pela depleção de linfócitos T CD8⁺ e células dendríticas, concomitante ao aumento de populações imunossupressoras, como macrófagos associados ao tumor (TAMs) e células T regulatórias (Tregs) (Ahmad et al., 2023; Lu et al., 2025;

Mughal et al., 2025). Adicionalmente, abundância de mastócitos em repouso e células TCD4⁺ de memória, além de uma redução de células B plasmáticas, indicam um microambiente menos exposto a estímulos antigênicos crônicos e mais propenso à manutenção de nichos linfoides estáveis (Hao et al., 2022; Phad et al., 2022). Em fumantes, a alta taxa de mutações gera uma diversidade antigênica que estimula a resposta humoral e a formação de estruturas linfoides terciárias, enquanto nos não fumantes, a progressão tumoral parece depender de uma comunicação mais sutil entre as células tumorais e o sistema imune inato.

Assim, diferentemente do contexto de dano genômico intenso induzido pelo tabaco, o tumor primário de não fumantes tende a evoluir por mecanismos mais sutis de plasticidade tumoral e remodelação imune, favorecendo a evasão imune e a sobrevivência celular em ambientes menos proliferativos. Essa adaptação imunológica parece crucial para o sucesso metastático e pode explicar a menor resposta à imunoterapia observada em metástases cerebrais, reforçando a importância de estratégias terapêuticas que revertam o estado imune-frio e restaurem a atividade antitumoral efetora.

A análise de RNA-Seq das metástases cerebrais reforçou essa interpretação ao revelar um conjunto de 257 genes diferencialmente expressos entre não fumantes e fumantes, sendo que 85 estavam com expressão aumentada e 74 com expressão diminuída. Entre os genes mais diferencialmente expressos (Top 20) observou-se um predomínio de funções relacionadas a regulação imune, adesão celular, interação com a matriz extracelular, metabolismo e controle epigenético e metabolismo. Nas metástases cerebrais de pacientes não fumantes, a maior expressão de genes associados à resposta imune e ao sistema complemento (como *ABCF1*, *TRIM26*, *C4B_2* e *NEU1*) sugere modulação ativa do TME no sistema nervoso central (Garlanda; Dambra; Magrini, 2025), enquanto a regulação positiva de genes envolvidos na adesão e remodelamento da matriz extracelular (como *MUC21* e *DDR1*) é compatível com mecanismos de plasticidade e adaptação ao nicho metastático (Lambert; Pattabiraman; Weinberg, 2017). Adicionalmente, a expressão aumentada de genes relacionados ao controle transcricional e epigenético, como *PIWIL3* e *ZNF502*, reforça a importância da reprogramação molecular na colonização cerebral (Flavahan; Gaskell; Bernstein, 2017; Powell et al., 2025). Em contraste, a menor expressão de genes ligados à detoxificação de xenobióticos, como os da família *UGT1A*, *AKR1C2*, *NELFE*, pode refletir reprogramação metabólica adaptativa

característica de tumores em progressão (Faubert; Solmonson; DeBerardinis, 2020).

O enriquecimento funcional dos DEGs oriundos da análise de amostras clínicas de metástase cerebral (RNA-Seq), destacou uma profunda reprogramação metabólica nas amostras de metástase cerebral de não fumantes. As vias identificadas apontam para um uso estratégico do metabolismo de ácidos graxos, corpos cetônicos e compostos secundários, refletindo a plasticidade necessária para que as células neoplásicas sobrevivam em um nicho com disponibilidade limitada de nutrientes e alta pressão seletiva (Tyagi; Wu; Watabe, 2022). Além disso, o envolvimento das vias de retinoides e isoprenoides reforça a existência de programas de diferenciação celular e sinalização oncogênica adaptados ao SNC, os quais podem conferir vantagens de sobrevivência e resistência terapêutica ao tumor metastático (Chen et al., 2022). Em resumo, todo esse perfil sugere que, em não fumantes, as metástases cerebrais apresentam menor ativação de vias clássicas de biotransformação associadas ao tabagismo e maior ênfase em mecanismos de adaptação imune e estrutural ao microambiente do SNC.

A integração das análises de tumor primário (TCGA) e metástase cerebral (RNA-Seq clínico) revelou, em pacientes não fumantes, uma assinatura central de 42 genes comuns, sendo que 30 desses genes mantiveram padrão de expressão conservado entre esses diferentes sítios, sugerindo a existência de um programa biológico característico do adenocarcinoma pulmonar nesse grupo. Esses genes organizam-se em eixos funcionais que refletem processos biológicos centrais para a adaptação tumoral e a interação com o microambiente, especialmente no contexto do sistema nervoso central (Bergers; Fendt, 2021; Patel; Kolundzic; Abedalthagafi, 2025).

Entre os genes consistentemente mais expressos, destaca-se *FOS*, cuja maior expressão nos dois contextos reforça seu papel na plasticidade fenotípica na regulação de eventos associados à transição epitélio-mesenquimal, e à agressividade biológica em carcinomas pulmonares (Manios et al., 2020; Song; Lian; Zhang, 2023). De forma complementar, a expressão aumentada de *DUSP6*, um importante gene modulador do eixo MAPK/ERK tradicionalmente descrito como regulador negativo da via em adenocarcinoma pulmonar (Moncho-Amor et al., 2019), sugere uma arquitetura regulatória particular no adenocarcinoma de não fumantes, possivelmente relacionada à modulação fina da progressão tumoral. Além desses reguladores centrais, a maior expressão de genes ligados à interação imune, como *HLA-DPB1* e *CX3CL1*, e ao remodelamento da matriz extracelular, como *HAS3*, reforça um eixo funcional voltado

à comunicação tumor–microambiente (Binnewies et al., 2018; Michalczyk et al., 2022). A presença concomitante de genes associados a canais iônicos e adaptação ao microambiente neural, como *KCNJ15* e *SLC1A7*, sugere ainda que características compatíveis com competência metastática cerebral podem estar incorporadas ao perfil molecular dos tumores de não fumantes, em consonância com modelos atuais que descrevem a progressão metastática ao sistema nervoso central como um processo seletivo e adaptativo (Gritsch; Brastianos, 2025; Peinado et al., 2017). Em resumo, a conservação direcional desses genes sustenta que tais diferenças refletem um traço intrínseco da biologia tumoral associada ao tumor em pacientes não fumantes, não exclusivamente dependente do sítio tumoral.

Embora a maioria dos genes compartilhados entre as análises apresente padrão de expressão concordante, sugerindo a existência de uma assinatura intrínseca do adenocarcinoma em pacientes não fumantes, um subconjunto de 12 genes demonstrou inversão de direção de expressão entre o tumor primário e a metástase cerebral. Diferentemente dos genes com padrão conservado, esse subconjunto sugere reprogramação molecular dependente do sítio tumoral. Estudos recentes têm identificado diversos genes com potencial envolvimento no risco e na progressão do câncer de pulmão, seja por elevada expressão no epitélio pulmonar, como *SFTA2*, seja por participação em processos como modulação do TME, reparo de DNA, biogênese ribossômica, adaptação ao estresse e metabolismo mitocondrial, incluindo *HLA-DMB*, *MDC1*, *ABCF1*, *GPC2*, *FBL*, *HSPA1B*, *VAR2*, *HCN3* e *CCHCR1*. Entre esses genes, destaca-se *HLA-DMB* por seu papel central na via de apresentação de antígenos mediada por MHC classe II. Esse gene codifica uma molécula reguladora do complexo MHC-II, fundamental para a edição e o carregamento de peptídeos antigênicos no MHC-II, promovendo a remoção do peptídeo CLIP e permitindo a apresentação eficiente a linfócitos T CD4⁺ (Suarez-Trujillo et al., 2023). Estudos demonstram que sua expressão diferencial em adenocarcinoma pulmonar está associada a alterações no microambiente imune e a desfechos clínicos (Xi et al., 2025). Em nossos resultados, a maior expressão de *HLA-DMB* no conjunto de tumores primários de pacientes não fumantes quando comparada à expressão observada nas metástases cerebrais do mesmo grupo sugere eficiência da via MHC-II no SNC. Considerando que a evasão imune constitui um dos pilares da progressão metastática (Hanahan, 2022), a modulação negativa desse gene na metástase de não fumantes pode representar um mecanismo adaptativo que favorece

o escape da vigilância imunológica e o estabelecimento tumoral no sistema nervoso central (Guan; Guan; Song, 2020; Tang et al., 2024; Zhang et al., 2024b).

De forma semelhante, *SFTA2*, associado à identidade epitelial pulmonar, apresentou, em nossos achados, uma maior expressão tumor primário de pacientes não fumantes em comparação à metástase cerebral de não fumantes. Esse gene codifica uma proteína secretada predominantemente expressa no epitélio respiratório, relacionada à homeostase pulmonar e à manutenção do fenótipo alveolar (Li et al., 2022; Mittal et al., 2012). Evidências recentes sugerem ainda que sua expressão pode estar associada a características prognósticas nesse subtipo tumoral (Luyapan et al., 2023). Em nossos resultados, sua redução na metástase cerebral de não fumantes é compatível com perda parcial da identidade tecidual e com processos de desdiferenciação celular, fenômenos frequentemente associados à plasticidade tumoral e à progressão metastática (Li et al., 2022; Lin et al., 2025). A diminuição de marcadores específicos do tecido de origem, observada nesse contexto, reforça a hipótese de reprogramação transcricional necessária para adaptação ao microambiente cerebral, biologicamente distinto do pulmão em termos celulares, metabólicos e imunológicos (Lin et al., 2025).

Em contrapartida, a maior expressão de genes como *ABCF1* e *HSPA1B* na metástase cerebral de pacientes não fumantes evidencia a pressão adaptativa imposta pelo nicho do sistema nervoso central. *HSPA1B*, membro da família HSP70, atua como chaperona molecular, promovendo dobramento proteico adequado e prevenindo apoptose sob condições de estresse celular, incluindo hipóxia e agressões citotóxicas (Zhao et al., 2023). Já *ABCF1*, embora não funcione como transportador clássico de efluxo, participa da regulação da iniciação traducional e de vias inflamatórias, podendo contribuir para a manutenção da síntese proteica e da viabilidade celular em contextos adversos (Cao et al., 2020). Em conjunto, a superexpressão desses genes pode conferir maior resiliência celular durante o processo de colonização do sistema nervoso central. Paralelamente, o aumento de genes associados ao reparo de DNA, como *MDC1*, reforça a hipótese de ativação da resposta ao dano genômico, uma vez que essa proteína atua como mediadora central no recrutamento e na amplificação de fatores da via de DNA damage response (DDR), contribuindo para a manutenção da estabilidade cromossômica e da integridade genômica (Ruff et al., 2020). A elevação de *FBL*, componente nucleolar essencial para a metilação de rRNA e biogênese ribossômica, sugere incremento da capacidade

biossintética celular (Elhamamsy et al., 2022). Esse conjunto de alterações pode refletir um estado compatível com preparação para reentrada proliferativa após a fase inicial de adaptação ao microambiente cerebral, favorecendo o estabelecimento progressivo da metástase.

Em conjunto, esses dados indicam que, no contexto dos pacientes não fumantes analisados, o cenário imunológico e funcional do adenocarcinoma pulmonar não se restringe à composição celular do infiltrado tumoral, mas envolve uma reprogramação molecular dinâmica, evidenciada pela modulação coordenada de vias imunes, metabólicas e estruturais (Khalaf et al., 2021). Essa arquitetura sugere que, nesse grupo, a progressão tumoral resulta da coexistência entre uma identidade transcricional parcialmente conservada e transições fenotípicas contextuais que favorecem imunoevasão e adaptação metastática. Embora tumores associados ao tabagismo frequentemente apresentem microambiente mais inflamatório e maior carga mutacional, refletindo maior potencial de reconhecimento antigênico, a dinâmica adaptativa observada nos não fumantes pode seguir trajetórias distintas de interação com o sistema imune. A compreensão dessas assinaturas funcionais torna-se fundamental para o delineamento de estratégias terapêuticas mais precisas, particularmente no contexto da imunoterapia, cuja eficácia depende não apenas da carga mutacional, mas também do estado funcional e adaptativo do TME.

5.2. Análise NanoString: assinatura imune e diferenciação entre tumor primário e metástase cerebral de adenocarcinoma pulmonar de pacientes não fumantes

A caracterização do TME (TME) é um eixo central para compreender a progressão do adenocarcinoma pulmonar (LUAD), especialmente em pacientes não fumantes, cujo perfil genômico e imunológico difere substancialmente dos fumantes (Hamouz et al., 2023). Estudos demonstram que o TME é diretamente influenciado pelas células imunes, que podem promover o crescimento e a progressão das células cancerígenas (Anderson; Simon, 2020). Além disso, o TME não apenas apresenta variações entre indivíduos com o mesmo subtipo tumoral, mas também sofre profundas alterações ao longo da progressão espacial da doença, especialmente na transição do tumor primário pulmonar para a metástase cerebral (Bremnes et al., 2016). Um estudo recente envolvendo análise de sequenciamento de RNA de célula

única (scRNA-seq) de adenocarcinoma de pulmão metastático (LUAD), um subtipo principal de NSCLC, mostrou que tanto as células cancerígenas e quanto as células do TME passam por extensa reprogramação molecular e celular para criar um ambiente pró-tumoral e imunossupressor propício à metástase (Kim et al., 2020). Da mesma forma que, análises multi-ômicas de 15 cânceres metastáticos revelaram que o microambiente cerebral de diferentes cânceres são caracterizados por um TME evasivo imunológico com a presença de macrófagos associados a metástases (MAMs) e respostas heterogêneas de células T (Gonzalez et al., 2022). Apesar de o microambiente do cérebro permanecer distintamente diferente de outros órgãos, inclusive o pulmão (Wang et al., 2022), nossos dados demonstram que amostras de estágios avançados tendem a se agrupar por perfil de expressão, independentemente de sua origem (primária ou metastática). Isso indica que, no nível transcricional, a gravidade da progressão pode aproximar perfis biológicos de tecidos distintos.

No presente estudo, a análise realizada pela plataforma NanoString nCounter® PanCancer Immune Profiling Panel permitiu caracterizar diferenças imunológicas marcantes entre tumores primários pulmonares e metástases cerebrais de pacientes não fumantes. Após controle de qualidade, normalização e correção de background, identificamos 17 genes diferencialmente expressos entre os dois grupos (p ajustado $< 0,05$; $FC \geq 1,5$), refletindo alterações significativas regulação de vias associadas à ativação imune, citotoxicidade e inflamação, sustentando a transição de um tumor primário com perfil imunologicamente ativo (“imune-quente”) para uma metástase cerebral caracterizada por um estado imunologicamente suprimido (“imune-frio”).

Essa interpretação foi reforçada pela análise de deconvolução celular, que revelou maior abundância de linfócitos T $CD8^+$, células NK e células dendríticas ($p = 0.016$) nos tumores primários pulmonares quando comparados às metástases cerebrais. Essas três populações celulares formam o eixo central da imunidade antitumoral, sendo que as células dendríticas atuam na linha de frente da apresentação de antígenos, enquanto os linfócitos T $CD8^+$ e as células NK são os principais efetores da citotoxicidade (Anderson; Simon, 2020). A maior abundância dessas células no sítio pulmonar (primário) sugere um TME mais imunologicamente ativo, no qual o sistema imune ainda exerce pressão seletiva e vigilância sobre as células tumorais. De forma consistente com essa interpretação, a presença de linfócitos T citotóxicos no TME tem sido associada a melhor prognóstico em pacientes com câncer (Anderson; Simon, 2020). Além disso, estudos demonstram que células

TCD8⁺ podem apresentar o subtipo predominante em tumores primários, sendo o único subtipo identificado nos estágios II e IV e o mais abundante no estágio III (Souza et al., 2024), reforçando seu papel central na resposta imune antitumoral. Em conjunto, esses achados sugerem que, em adenocarcinomas pulmonares de pacientes não fumantes, o tumor primário mantém um microambiente mais favorável à ativação de respostas imunes efetoras, enquanto a metástase cerebral reflete um contexto de menor infiltração imune e maior permissividade à progressão tumoral, possivelmente associado às características imunológicas do sistema nervoso central, que tende a limitar respostas inflamatórias exacerbadas para preservar a integridade tecidual, além da atuação da barreira hematoencefálica na restrição do recrutamento de células imunes periféricas.

Em contrapartida, as metástases cerebrais exibiram um percentual diminuído consistente dessas populações celulares, o que pode ser parcialmente explicado pelo fato de que lesões metastáticas no cérebro podem comprometer o recrutamento e o sequestro de células T, um mecanismo que tem sido apontado como uma possível direção de investigação para o desenvolvimento de novas estratégias terapêuticas no tratamento de metástases cerebrais (Wang et al., 2023b). Esses achados corroboram estudos que descrevem o sistema nervoso central (SNC) como um ambiente de privilégio imune, no qual mecanismos ativos de exclusão e supressão imunológica são necessários para a colonização metastática bem-sucedida (Schreurs et al., 2024). Além disso, as metástases cerebrais de LUAD são caracterizadas por baixa infiltração linfocitária, acompanhada por um aumento compensatório de macrófagos associados à metástase (MAMs) com fenótipo imunossupressor M2, que inibem ativamente a ativação de células T (Ahmad et al., 2023; Mughal et al., 2025). Nesse contexto, a redução observada de linfócitos T CD8⁺, células NK e células dendríticas nas metástases cerebrais de pacientes não fumantes sugere a existência de um TME mais permissivo à progressão metastática, caracterizado por menor vigilância imunológica e maior capacidade de evasão tumoral. Assim, a transição do tumor primário pulmonar para o sítio metastático cerebral parece envolver não apenas alterações intrínsecas das células tumorais, mas também uma reprogramação do microambiente imune, favorecendo um estado mais imunossuprimido que sustenta a sobrevivência e expansão das células metastáticas no SNC.

No nível transcricional, 17 genes diferencialmente expressos entre tumores primários pulmonares e metástases cerebrais em pacientes não fumantes foram

identificados, sendo que 16 desses genes estão com expressão reduzida nas metástases cerebrais de não fumantes, enquanto o *CDH1* foi o único gene com expressão aumentada nesse sítio metastático. De maneira geral, os genes suprimidos estão majoritariamente associados à citotoxicidade imune e à ativação da resposta imune adaptativa, fornecendo uma explicação para o fenótipo imune-frio observado nas metástases cerebrais.

Entre os genes com expressão reduzida, destacam-se *PRF1* e *GNLY*, que codificam respectivamente a perforina e a granulisina, componentes centrais dos grânulos citolíticos presentes em linfócitos T CD8⁺ e células NK. A perforina atua criando poros na membrana da célula alvo, permitindo a entrada das granzimas que desencadeiam a ativação da cascata de caspases e a fragmentação do DNA, culminando na morte da célula tumoral por apoptose (Voskoboinik; Whisstock; Trapani, 2015). Dessa forma, a diminuição da expressão desses genes nas metástases cerebrais de não fumantes sugere comprometimento funcional direto da resposta citotóxica efetora, favorecendo a sobrevivência tumoral nesse nicho metastático. Paralelamente, observou-se a redução na expressão de genes envolvidos na sinalização e co-estimulação de linfócitos T, como *ITK*, *CD27*, *CD70*, *IL2RB*, *ICOS*, *TNFRSF4* (*OX40*) e *CD8B*. Esses genes participam de etapas essenciais da ativação, expansão clonal e manutenção da resposta de células T mediada pelo receptor de célula T (TCR) (Chang et al., 2024; Chen et al., 2026). Essa redução sugere que, mesmo que linfócitos residuais estejam presentes, eles permanecem em estado de anergia ou exaustão, sendo incapazes portanto de montar uma resposta citotóxica eficaz contra as células tumorais (Thommen; Schumacher, 2018).

Além disso, genes associados à imunidade inata e à comunicação entre compartimentos imunes, como *NOD2*, *CCR2*, *BTK*, *LTF* e *FPR2*, também apresentaram expressão reduzida nas metástases cerebrais de não fumantes. Esses genes participam do reconhecimento de padrões moleculares associados a patógenos ou dano celular, recrutamento de células imunes e modulação de respostas inflamatórias, sugerindo que a supressão imunológica no sistema nervoso central ocorre de maneira multifatorial, afetando simultaneamente componentes da imunidade inata e adaptativa (Mellman et al., 2023).

Em contraste com o declínio generalizado das assinaturas imunes, observamos regulação positiva de *CDH1* (E-cadherina) nas metástases cerebrais. A E-cadherina

é uma molécula de adesão celular fundamental para a manutenção do fenótipo epitelial e da arquitetura tecidual, estando também envolvida em processos de plasticidade tumoral e colonização metastática. Estudos sugerem que a re-expressão ou manutenção de E-cadherina pode favorecer a colonização de sítios metastáticos por meio da transição mesenquimal-epitelial (MET), etapa importante para o estabelecimento e crescimento de metástases em órgãos distantes (Lambert; Pattabiraman; Weinberg, 2017).

Esse padrão transcricional sugere que, durante a progressão metastática para o cérebro, o tumor passa a priorizar programas de adaptação celular, sobrevivência, plasticidade fenotípica em detrimento da imunogenicidade (O'Donnell; Teng; Smyth, 2019). Esse fenômeno é particularmente relevante no sistema nervoso central, um ambiente com características imunológicas e metabólicas singulares que podem favorecer a adaptação e a reprogramação metabólica de células tumorais metastáticas (Cheng et al., 2022; Jin; Jin, 2020).

Em conjunto, esses achados são consistentes com o conceito de imunoeedição do câncer, no qual a interação dinâmica entre o tumor e o sistema imune ocorre nas fases de eliminação, equilíbrio e escape. Ao longo desse processo, a pressão seletiva exercida pelo sistema imune favorece a sobrevivência de clones tumorais menos imunogênicos, culminando em um TME permissivo à evasão imunológica e à progressão metastática (Liu; Sun; Ren, 2023). Nesse contexto, a colonização do sistema nervoso central parece favorecer subpopulações tumorais capazes de prosperar em um nicho onde as respostas imunes adaptativas são naturalmente mais restritas (Zhang et al., 2022).

Do ponto de vista translacional, o fenótipo imune-frio das metástases cerebrais, evidenciado pela diminuição da infiltração de células efectoras e pela diminuição da expressão de genes relacionados à citotoxicidade e ativação imune, pode ajudar a explicar a menor responsividade às imunoterapias baseadas em bloqueio de checkpoints (anti-PD-1/PD-L1) nesse contexto. Esse cenário é consistente com estudos que demonstram menor eficácia dessas terapias em pacientes não fumantes com adenocarcinoma pulmonar, possivelmente devido à menor carga mutacional tumoral e à reduzida imunogenicidade do tumor (Devarakonda et al., 2021). Assim, estratégias terapêuticas futuras podem se beneficiar de abordagens capazes de recondicionar o TME, promovendo o recrutamento e a ativação de células imunes efectoras (Lu et al., 2025).

De forma integrada, os resultados deste estudo demonstram que tumores primários pulmonares de pacientes não fumantes apresentam um microambiente mais imunologicamente ativo, caracterizado por maior abundância de linfócitos T CD8⁺, células NK e células dendríticas, bem como maior expressão de genes associados à citotoxicidade e ativação imune. Em contraste, as metástases cerebrais exibem redução dessas populações celulares e repressão transcricional de genes envolvidos na resposta imune, acompanhadas pela regulação positiva de genes associados à adaptação tumoral. Esses achados indicam que a progressão do adenocarcinoma pulmonar para o cérebro envolve uma reprogramação simultânea do microambiente celular e do perfil transcricional, culminando em um nicho metastático imunologicamente suprimido e favorável à evasão tumoral.

5.3. Comparação de técnicas e robustez da metodologia

A abordagem metodológica deste estudo combinou a utilização de dados públicos (TCGA), a geração de dados inéditos por RNA-Seq (NGS) e a caracterização imunológica de amostras clínicas brasileiras com a plataforma NanoString®, visando a validação e o refinamento de assinaturas identificadas no RNA-seq e consequentemente conferindo robustez e aplicabilidade aos resultados.

O RNA-Seq é uma das principais tecnologias para análise do transcriptoma, permitindo a quantificação precisa da expressão gênica global, a identificação de transcritos alternativos e a descoberta de novos RNAs, incluindo isoformas e RNAs não codificantes. Sua sensibilidade, abrangência e capacidade de gerar dados em nível de base tornam-no uma ferramenta poderosa tanto para estudos exploratórios quanto para análises funcionais aprofundadas. No entanto, assim como ocorre nas abordagens genômicas convencionais, o RNA-Seq realizado a partir de populações celulares heterogêneas apresenta uma limitação fundamental: a perda de informações específicas de cada tipo celular (Cilento; Sweeney; Butler, 2024). Ao capturar o RNA total de um tecido ou amostra mista, os sinais de expressão gênica refletem uma média ponderada de todos os tipos celulares presentes, o que pode mascarar diferenças críticas em subpopulações celulares minoritárias ou biologicamente relevantes. Isso dificulta a identificação de assinaturas transcriptômicas específicas de um tipo celular, reduz a capacidade de detectar variações sutilmente localizadas e compromete a interpretação funcional em contextos biológicos complexos, como microambientes tumorais ou tecidos imunologicamente ativos.

Já o sistema NanoString nCounter® apresenta características metodológicas que o tornam especialmente vantajoso em determinadas aplicações clínicas e experimentais: ele realiza uma quantificação direta e sem amplificação de transcritos pré-definidos, o que evita vieses comuns decorrentes da amplificação por PCR e da conversão de RNA em cDNA (Bondar et al., 2020). Isso confere ao NanoString® alta reprodutibilidade técnica e maior tolerância a variações na qualidade das amostras, o que o torna particularmente útil em contextos nos quais o RNA se encontra fragmentado ou degradado, como em tecidos fixados em formalina e embebidos em parafina (FFPE) (Veldman-Jones et al., 2015). A plataforma permite a análise simultânea de até 800 genes em uma única reação, por meio de sondas específicas denominadas CodeSets. Esses painéis são altamente customizáveis e direcionados, o que possibilita a investigação de vias moleculares conhecidamente associadas à imunidade tumoral e à evasão do câncer, em contraste com abordagens exploratórias de transcriptoma global. Dessa forma, o uso do NanoString® complementa as análises de RNA-Seq, permitindo ampliar a resolução imunológica por meio da validação e do refinamento das assinaturas gênicas previamente identificadas (Reis et al., 2011). No presente estudo, utilizamos um painel direcionado à resposta imune no TME, com foco no câncer de pulmão. Essa abordagem possibilitou investigar de maneira mais detalhada os genes diretamente associados ao adenocarcinoma pulmonar em não fumantes, fortalecendo a relevância translacional dos achados. A aplicabilidade dessa estratégia já foi demonstrada em estudos recentes na área de câncer de pulmão, como o LungTS (“LungTS: A new gene expression signature for prediction of response to checkpoint inhibitors in non-small cell lung cancer. | Journal of Clinical Oncology”, [S.d.]), o PEOPLE trial (Lo Russo et al., 2023) e a análise de marcadores de exaustão imunológica em NSCLC (Sellmer et al., 2022).

Contudo, apesar das vantagens de robustez e confiabilidade, o painel nCounter também apresenta limitações. A principal delas é seu design direcionado: como os genes são previamente selecionados e especificados, a tecnologia não permite a descoberta de novos genes ou vias moleculares relevantes para o câncer (Eastel et al., 2019). Isso faz com que o método seja ideal para a validação de hipóteses, mas não para a exploração de alvos moleculares desconhecidos, o que reforça a complementaridade das duas técnicas. Tanto o RNA-Seq quanto o NanoString® apresentam limitações na detecção de genes de baixa abundância. No caso do NanoString®, essa restrição é mais acentuada, pois a tecnologia apresenta uma faixa

dinâmica mais estreita, o que dificulta a quantificação de genes com expressão extremamente baixa. Além disso, pela limitação da tecnologia dependente de sondas específicas, o NanoString pode mascarar variações genéticas sutis e comprometer a detecção de alterações geneticamente relevantes (Bergbower; Pierson; Azimzadeh, 2020). Essa questão é agravada pelo fato de o nCounter ser otimizado para a quantificação de RNA mensageiro (mRNA), que representa apenas uma pequena fração do RNA total da amostra. Em consequência, a plataforma pode exigir uma quantidade maior de RNA de entrada para garantir uma sensibilidade adequada, o que se torna um desafio em estudos que utilizam material biológico escasso, como biópsias finas ou amostras raras (Chilimoniuk et al., 2024).

O NanoString® disponibiliza o nSolver®, um software proprietário com interface gráfica voltado especificamente para a análise dos dados gerados pelo sistema nCounter. Embora facilite o controle de qualidade, a normalização e a análise de expressão gênica, seu uso é limitado aos métodos propostos pelo fabricante, restringindo a flexibilidade analítica. Em contrapartida, a comunidade científica tem desenvolvido alternativas mais robustas e versáteis utilizando a linguagem R, amplamente adotada em bioinformática. Com o suporte do projeto Bioconductor, o R oferece uma ampla gama de pacotes estatísticos e gráficos para análise reprodutível de dados de expressão gênica, incluindo ferramentas específicas para o processamento de dados do nCounter. Essa abordagem proporciona maior liberdade metodológica, personalização das análises e integração com pipelines mais avançados (Canouil et al., 2019; Chilimoniuk et al., 2024).

Estudos comparativos demonstram que o nCounter e o RNA-Seq apresentam alta correlação nos níveis de expressão gênica para alvos definidos, mas o nCounter se destaca em contextos clínicos onde a integridade do RNA é comprometida ou onde se busca reprodutibilidade técnica em painéis específicos, em detrimento de uma análise exploratória ampla. Portanto, embora o RNA-Seq continue sendo a escolha padrão para estudos de transcriptoma global, o NanoString nCounter® representa uma solução estratégica e confiável para validação de alvos, estudos direcionados e aplicação clínica, desde que suas limitações sejam adequadamente consideradas no desenho experimental.

6. CONCLUSÃO

A análise de RNA-Seq de tumores primários de adenocarcinoma pulmonar em pacientes não fumantes, provenientes do banco TCGA, revelou um perfil transcriptômico marcado pela ativação de vias relacionadas ao microambiente tumoral, especialmente aquelas envolvidas em resposta imune, inflamação e interação entre células tumorais e o sistema imunológico. Esses achados indicam que, no tumor primário, há participação ativa de componentes imunológicos, sugerindo um contexto biologicamente dinâmico e potencialmente responsivo a estímulos imunológicos.

A integração dos dados de RNA-Seq de tumores primários com amostras de metástases cerebrais permitiu identificar um conjunto de genes compartilhados entre esses dois contextos, sugerindo a existência de assinaturas moleculares associadas à progressão metastática em pacientes não fumantes. Esses genes possivelmente refletem mecanismos comuns envolvidos na capacidade de disseminação tumoral e adaptação das células neoplásicas a diferentes microambientes, incluindo o sistema nervoso central.

A caracterização imunológica por meio da plataforma NanoString evidenciou diferenças marcantes entre os sítios tumorais. Os tumores primários apresentaram maior abundância de células imunes efetoras, como linfócitos T CD8⁺, células NK e células dendríticas, além de maior expressão de genes associados à citotoxicidade e ativação da resposta imune. Em contraste, as metástases cerebrais mostraram redução dessas populações celulares, acompanhada pela diminuição da expressão de genes relacionados à resposta imune adaptativa e inata (como *PRF1*, *GNLY*, *CD8B*, *ICOS* e *CCR2*), além do aumento da expressão de *CDH1*, sugerindo adaptação tumoral, maior plasticidade celular e possível favorecimento da colonização no sistema nervoso central.

Em conjunto, esses achados sugerem que, em adenocarcinomas pulmonares de pacientes não fumantes, o tumor primário mantém um microambiente mais favorável à ativação de respostas imunes efetoras, enquanto a metástase cerebral reflete um contexto de redução da infiltração imune e maior permissividade à progressão tumoral. Assim, a transição do tumor primário pulmonar para o sítio

metastático cerebral parece envolver não apenas alterações intrínsecas das células tumorais, mas também uma reprogramação do microambiente imune, favorecendo um estado mais imunossuprimido que sustenta a sobrevivência e expansão das células metastáticas no sistema nervoso central. Esses fatores podem impactar diretamente o desenvolvimento de estratégias terapêuticas mais direcionadas para pacientes não fumantes.

7. REFERÊNCIAS BIBLIOGRÁFICAS

AHMAD, Aatiya *et al.* Immunotherapy: an emerging modality to checkmate brain metastasis. **Molecular Cancer**, v. 22, p. 111, 15 jul. 2023.

ALATTAS, Amal A.; ALKHATTABI, Nuha A.; ALZAHRANI, Maryam H. Alterations associated in dendritic cells sub-types in smokers and non-smokers. 2020.

ALEXANDROV, Ludmil B. *et al.* The repertoire of mutational signatures in human cancer. **Nature**, v. 578, n. 7793, p. 94–101, fev. 2020.

ANDERSEN, Allan M. *et al.* Inflammatory biomarker relationships with helper T cell GPR15 expression and cannabis and tobacco smoking. **Journal of Psychosomatic Research**, v. 141, p. 110326, fev. 2021.

ANDERSON, Nicole M.; SIMON, M. Celeste. Tumor Microenvironment. **Current biology : CB**, v. 30, n. 16, p. R921–R925, 17 ago. 2020.

ANDREWS, Simon. **FastQC: A Quality Control tool for High Throughput Sequence Data**. Disponível em: <<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>>. Acesso em: 9 mar. 2026.

ARAN, Dvir; HU, Zicheng; BUTTE, Atul J. xCell: digitally portraying the tissue cellular heterogeneity landscape. **Genome Biology**, v. 18, p. 220, 15 nov. 2017.

ASSOCIATION, American Lung. **State of Lung Cancer | Key Findings**. Disponível em: <<https://www.lung.org/research/state-of-lung-cancer/key-findings>>. Acesso em: 16 out. 2025.

BECHT, Etienne *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. **Genome Biology**, v. 17, p. 218, 20 out. 2016.

BERGBOWER, Emily AS; PIERSON, Richard N.; AZIMZADEH, Agnes M. Multi-Gene Technical Assessment of qPCR and Nanostring n-Counter Analysis Platforms in Cynomolgus Monkey Cardiac Allograft Recipients. **Cellular immunology**, v. 347, p. 104019, jan. 2020.

BERGERS, Gabriele; FENDT, Sarah-Maria. The metabolism of cancer cells during metastasis. **Nature reviews. Cancer**, v. 21, n. 3, p. 162–180, mar. 2021.

BINNEWIES, Mikhail *et al.* Understanding the tumor immune microenvironment (TIME) for effective therapy. **Nature medicine**, v. 24, n. 5, p. 541–550, maio 2018.

BOECKX, Bram *et al.* The genomic landscape of nonsmall cell lung carcinoma in never smokers. **International Journal of Cancer**, v. 146, n. 11, p. 3207–3218, 1 jun. 2020.

BONDAR, Galyna *et al.* Comparing NGS and NanoString platforms in peripheral blood mononuclear cell transcriptome profiling for advanced heart failure biomarker development. **Journal of Biological Methods**, v. 7, n. 1, p. e123, 3 jan. 2020.

BÖTTCHER, Jan P.; REIS E SOUSA, Caetano. The Role of Type 1 Conventional Dendritic Cells in Cancer Immunity. **Trends in Cancer**, v. 4, n. 11, p. 784–792, nov. 2018.

BRAY, Freddie *et al.* Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: A Cancer Journal for Clinicians**, v. 74, n. 3, p. 229–263, 2024.

BREMNES, Roy M. *et al.* The Role of Tumor-Infiltrating Lymphocytes in Development, Progression, and Prognosis of Non–Small Cell Lung Cancer. **Journal of Thoracic Oncology**, v. 11, n. 6, p. 789–800, 1 jun. 2016.

CÂNCER, Instituto Nacional de. **Estimativa 2023: incidência de câncer no Brasil**. Rio de Janeiro, RJ: Instituto Nacional De Câncer, 2023.

CANOUIL, Mickaël *et al.* NACHO: an R package for quality control of NanoString nCounter data. **Bioinformatics**, v. 36, n. 3, p. 970–971, 23 ago. 2019.

CAO, Liren *et al.* Macrophage heterogeneity and its interactions with stromal cells in tumour microenvironment. **Cell & Bioscience**, v. 14, p. 16, 1 fev. 2024.

CAO, Quynh T. *et al.* ABCF1 Regulates dsDNA-induced Immune Responses in Human Airway Epithelial Cells. **Frontiers in Cellular and Infection Microbiology**, v. 10, p. 487, 16 set. 2020.

CARE, Canadian Task Force on Preventive Health. Recommendations on screening for lung cancer. **CMAJ**, v. 188, n. 6, p. 425–432, 5 abr. 2016.

CARLSON, Marc. **org.Hs.eg.db: Genome wide annotation for Human**. Disponível em: <<http://bioconductor.org/packages/org.Hs.eg.db/>>. Acesso em: 9 mar. 2026.

CARROT-ZHANG, Jian *et al.* Genetic ancestry contributes to somatic mutations in lung cancers from admixed Latin American populations. **Cancer discovery**, v. 11, n. 3, p. 591–598, mar. 2021.

CASAL-MOURIÑO, Ana *et al.* Lung cancer survival among never smokers. **Cancer Letters**, v. 451, p. 142–149, 1 jun. 2019.

CHANG, Shi-Rou *et al.* The expression of immune co-stimulators as a prognostic predictor of head and neck squamous cell carcinomas and oral squamous cell carcinomas. **Journal of Dental Sciences**, v. 19, n. 3, p. 1380–1388, 1 jul. 2024.

CHEN, Po-Ming; HUANG, Yu-Han; LI, Chia-Ying. Lung Cancer in Never-Smokers: Risk Factors, Driver Mutations, and Therapeutic Advances. **Diagnostics**, v. 16, n. 2, p. 245, 12 jan. 2026.

CHEN, Shuting *et al.* Retinoids in cancer chemoprevention and therapy: Meta-analysis of randomized controlled trials. **Frontiers in Genetics**, v. 13, p. 1065320, 9 nov. 2022.

CHEN, Yang; MCANDREWS, Kathleen M.; KALLURI, Raghu. Clinical and therapeutic relevance of cancer-associated fibroblasts. **Nature Reviews. Clinical Oncology**, v. 18, n. 12, p. 792–804, dez. 2021.

CHEN, Yuanyuan *et al.* CD27 and ICOS as Targets of PD-1/PD-L1 Signaling to Regulate Resident Memory CD8⁺ T-Cell-Mediated Pulmonary Protection and Pathology. **Advanced Science**, v. 13, n. 7, p. e12452, fev. 2026.

CHENG, Elvin S. *et al.* Evaluating risk factors for lung cancer among never-smoking individuals using two Australian studies. **Journal of Cancer Research and Clinical Oncology**, v. 148, n. 10, p. 2827–2840, 27 maio 2022.

CHILIMONIUK, Jarosław *et al.* Challenges and opportunities in processing NanoString nCounter data. **Computational and Structural Biotechnology Journal**, v. 23, p. 1951–1958, 30 abr. 2024.

CHO, Jaeyoung *et al.* Proportion and clinical features of never-smokers with non-small cell lung cancer. **Chinese Journal of Cancer**, v. 36, p. 20, 8 fev. 2017.

CILENTO, Michael A.; SWEENEY, Christopher J.; BUTLER, Lisa M. Spatial transcriptomics in cancer research and potential clinical impact: a narrative review. **Journal of Cancer Research and Clinical Oncology**, v. 150, n. 6, p. 296, 8 jun. 2024.

CLASS, Caleb A. *et al.* Easy NanoString nCounter data analysis with the NanoTube. **Bioinformatics**, v. 39, n. 1, p. btac762, 28 nov. 2022.

CÓZAR, Beatriz *et al.* Tumor-infiltrating natural killer cells. **Cancer discovery**, v. 11, n. 1, p. 34–44, 1 jan. 2021.

DASARI, Krishna; ALFARO-MURILLO, Jorge A.; TOWNSEND, Jeffrey P. **Tobacco smoke alters the landscape of mutation, selection, and epistasis in lung adenocarcinoma.** bioRxiv, , 2 dez. 2024. Disponível em: <<https://www.biorxiv.org/content/10.1101/2024.11.27.625765v1>>. Acesso em: 12 fev. 2026

DAYLAN, Ayse Ece Cali *et al.* Lung Cancer in Never Smokers: Delving into Epidemiology, Genomic and Immune Landscape, Prognosis, Treatment, and Screening. **Lung**, v. 201, n. 6, p. 521–529, 1 dez. 2023.

DEMSˆAR, Janez *et al.* Orange: Data Mining Toolbox in Python. [S.d.].

DEVARAKONDA, Siddhartha *et al.* Genomic Profiling of Lung Adenocarcinoma in Never-Smokers. **Journal of Clinical Oncology**, v. 39, n. 33, p. 3747–3758, 20 nov. 2021.

DÍAZ-GAY, Marcos *et al.* The mutagenic forces shaping the genomes of lung cancer in never smokers. **Nature**, v. 644, n. 8075, p. 133–144, ago. 2025.

DUBIN, Sarah; GRIFFIN, Daniel. Lung Cancer in Non-Smokers. **Missouri Medicine**, v. 117, n. 4, p. 375–379, 2020.

EASTEL, Jennifer Mary *et al.* Application of NanoString technologies in companion diagnostic development. **Expert Review of Molecular Diagnostics**, v. 19, n. 7, p. 591–598, 3 jul. 2019.

ELHAMAMSY, Amr R. *et al.* Ribosome Biogenesis: A Central Player in Cancer

Metastasis and Therapeutic Resistance. **Cancer Research**, v. 82, n. 13, p. 2344–2353, 5 jul. 2022.

ERNST, Sophie M. *et al.* Tobacco Smoking-Related Mutational Signatures in Classifying Smoking-Associated and Nonsmoking-Associated NSCLC. **Journal of Thoracic Oncology**, v. 18, n. 4, p. 487–498, 1 abr. 2023.

FARHOOD, Bagher; NAJAFI, Masoud; MORTEZAEE, Keywan. CD8+ cytotoxic T lymphocytes in cancer immunotherapy: A review. **Journal of Cellular Physiology**, v. 234, n. 6, p. 8509–8521, 2019.

FAUBERT, Brandon; SOLMONSON, Ashley; DEBERARDINIS, Ralph J. Metabolic reprogramming and cancer progression. **Science (New York, N.Y.)**, v. 368, n. 6487, p. eaaw5473, 10 abr. 2020.

FINOTELLO, Francesca; TRAJANOSKI, Zlatko. Quantifying tumor-infiltrating immune cells from transcriptomics data. **Cancer Immunology, Immunotherapy : CII**, v. 67, n. 7, p. 1031–1040, 14 mar. 2018.

FLAVAHAN, William A.; GASKELL, Elizabeth; BERNSTEIN, Bradley E. Epigenetic plasticity and the hallmarks of cancer. **Science**, v. 357, n. 6348, p. eaal2380, 21 jul. 2017.

FU, Yue-Chun *et al.* Intratumoral heterogeneity and drug resistance in cancer. **Cancer Cell International**, v. 25, n. 1, p. 103, 18 mar. 2025.

GARLANDA, Cecilia; DAMBRA, Monica; MAGRINI, Elena. Interplay between the complement system and other immune pathways in the tumor microenvironment. **Seminars in Immunology**, v. 78, p. 101951, 1 jun. 2025.

GENTLEMAN, Robert C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. **Genome Biology**, v. 5, n. 10, p. R80, 15 set. 2004.

GONZALEZ, Hugo *et al.* Cellular Architecture of Human Brain Metastases. **Cell**, v. 185, n. 4, p. 729–745.e20, 17 fev. 2022.

GOSWAMI, Kuntal Kanti; BOSE, Anamika; BARAL, Rathindranath. Macrophages in tumor: An inflammatory perspective. **Clinical Immunology**, v. 232, p. 108875, 1 nov. 2021.

GRITSCH, David; BRASTIANOS, Priscilla K. Molecular evolution of central nervous system metastasis and therapeutic implications. **Trends in Molecular Medicine**, v. 31, n. 3, p. 240–251, 1 mar. 2025.

GUAN, Xiaoqing; GUAN, Zhiyuan; SONG, Chunli. Expression profile analysis identifies key genes as prognostic markers for metastasis of osteosarcoma. **Cancer Cell International**, v. 20, p. 104, 30 mar. 2020.

GUNBIN, Konstantin V. *et al.* Features of the CD1 gene family in rodents and the uniqueness of the immune system of naked mole-rat. **Biology Direct**, v. 19, p. 58, 29 jul. 2024.

GUO, Dina *et al.* A new immune signature for survival prediction and immune checkpoint molecules in lung adenocarcinoma. **Journal of Translational Medicine**, v. 18, n. 1, p. 123, 6 mar. 2020.

HACKL, Hubert *et al.* Computational genomics tools for dissecting tumour–immune cell interactions. **Nature Reviews Genetics**, v. 17, n. 8, p. 441–458, ago. 2016.

HAMMOUZ, Raneem Y. *et al.* Differential expression of lung adenocarcinoma transcriptome with signature of tobacco exposure. **Journal of Applied Genetics**, v. 61, n. 3, p. 421–437, set. 2020.

HAMOUZ, Mohammad *et al.* A Functional Genomics Review of Non-Small-Cell Lung Cancer in Never Smokers. **International Journal of Molecular Sciences**, v. 24, n. 17, p. 13314, 28 ago. 2023.

HANAHAHAN, Douglas. Hallmarks of Cancer: New Dimensions. **Cancer Discovery**, v. 12, n. 1, p. 31–46, jan. 2022.

HAO, Dapeng *et al.* The Single-Cell Immunogenomic Landscape of B and Plasma Cells in Early-Stage Lung Adenocarcinoma. **Cancer discovery**, v. 12, n. 11, p. 2626–2645, 2 nov. 2022.

HENDRIKS, L. E. *et al.* Oncogene-addicted metastatic non-small-cell lung cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up☆. **Annals of Oncology**, v. 34, n. 4, p. 339–357, 1 abr. 2023.

HILL, William *et al.* Lung adenocarcinoma promotion by air pollutants. **Nature**, v. 616, n. 7955, p. 159–167, 6 abr. 2023.

HUBER, Wolfgang *et al.* Orchestrating high-throughput genomic analysis with Bioconductor. **Nature Methods**, v. 12, n. 2, p. 115–121, fev. 2015.

JEON, Hyein *et al.* Update 2025: Management of Non-Small-Cell Lung Cancer. **Lung**, v. 203, n. 1, p. 53, 2025.

JIN, Ming-Zhu; JIN, Wei-Lin. The updated landscape of tumor microenvironment and drug repurposing. **Signal Transduction and Targeted Therapy**, v. 5, p. 166, 25 ago. 2020.

KASSAMBARA, Alboukadel *et al.* **ggpubr: “ggplot2” Based Publication Ready Plots**. , 24 fev. 2026. Disponível em: <<https://cran.r-project.org/web/packages/ggpubr/index.html>>. Acesso em: 9 mar. 2026

KERRIGAN, Kathleen *et al.* Real World Characterization of Advanced Non–Small Cell Lung Cancer in Never Smokers by Actionable Mutation Status. **Clinical Lung Cancer**, v. 22, n. 4, p. 260- 267.e2, 1 jul. 2021.

KHALAF, Khalil *et al.* Aspects of the Tumor Microenvironment Involved in Immune Resistance and Drug Resistance. **Frontiers in Immunology**, v. 12, p. 656364, 27 maio 2021.

KIM, Nayoung *et al.* Single-cell RNA sequencing demonstrates the molecular and

cellular reprogramming of metastatic lung adenocarcinoma. **Nature Communications**, v. 11, n. 1, p. 2285, 8 maio 2020.

KLEMPERER, Elias M. *et al.* Tobacco and Nicotine Use Among US Adult “Never Smokers” in Wave 4 (2016–2018) of the Population Assessment of Tobacco and Health Study. **Nicotine & Tobacco Research**, v. 23, n. 7, p. 1199–1207, 1 jul. 2021.

KOLDE, Raivo. **pheatmap: Pretty Heatmaps**. , 5 jun. 2025. Disponível em: <<https://cran.r-project.org/web/packages/pheatmap/index.html>>. Acesso em: 9 mar. 2026

KRUEGER, Felix *et al.* **FelixKrueger/TrimGalore: v0.6.10 - add default decompression path**. Zenodo, , 2 fev. 2023. Disponível em: <<https://zenodo.org/records/7598955>>. Acesso em: 19 fev. 2026

KUŚNIERCZYK, Piotr. Genetic differences between smokers and never-smokers with lung cancer. **Frontiers in Immunology**, v. 14, p. 1063716, 2 fev. 2023.

LAMBERT, Arthur W.; PATTABIRAMAN, Diwakar R.; WEINBERG, Robert A. EMERGING BIOLOGICAL PRINCIPLES OF METASTASIS. **Cell**, v. 168, n. 4, p. 670–691, 9 fev. 2017.

LANDI, Maria Teresa *et al.* Gene expression signature of cigarette smoking and its role in lung adenocarcinoma development and survival. **PloS One**, v. 3, n. 2, p. e1651, 20 fev. 2008.

LEE, Po-Hsin *et al.* Benefits of NGS in Advanced Lung Adenocarcinoma Vary by Populations and Timing of Examination. **International Journal of Molecular Sciences**, v. 25, n. 13, p. 6949, 25 jun. 2024.

LI, Na *et al.* Transcriptomic and immunologic implications of the epithelial–mesenchymal transition model reveal a novel role of SFTA2 in prognosis of non-small-cell lung carcinoma. **Frontiers in Genetics**, v. 13, p. 911801, 26 ago. 2022.

LI, Taiwen *et al.* TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. **Cancer research**, v. 77, n. 21, p. e108–e110, 1 nov. 2017.

LIN, Yen-Yu *et al.* Gene Expression Profile in Primary Tumor Is Associated with Brain-Tropism of Metastasis from Lung Adenocarcinoma. **International Journal of Molecular Sciences**, v. 22, n. 24, p. 13374, 13 dez. 2021.

LIN, Zhaofan *et al.* Single cell and machine learning identify type II pneumocyte-derived biomarkers HN1/OCIAD2/SFTA2 for non-small cell lung cancer prognosis and immune regulation. **European Journal of Medical Research**, v. 30, p. 1184, 26 nov. 2025.

LIU, Shaochuan; SUN, Qian; REN, Xiubao. Novel strategies for cancer immunotherapy: counter-immunoediting therapy. **Journal of Hematology & Oncology**, v. 16, n. 1, p. 38, 13 abr. 2023.

LIU, Xinyan *et al.* Crosstalk between cancer-associated fibroblasts and inflammation in tumor microenvironment: A novel perspective in cancer therapy (Review). **Oncology**

Reports, v. 54, n. 2, p. 93, 12 jun. 2025.

LO RUSSO, Giuseppe *et al.* PEOPLE (NTC03447678), a phase II trial to test pembrolizumab as first-line treatment in patients with advanced NSCLC with PD-L1 <50%: a multiomics analysis. **Journal for Immunotherapy of Cancer**, v. 11, n. 6, p. e006833, 7 jun. 2023.

LOPICCOLO, Jaclyn *et al.* Lung cancer in patients who have never smoked — an emerging disease. **Nature reviews. Clinical oncology**, v. 21, n. 2, p. 121–146, fev. 2024.

LOVE, Michael I.; HUBER, Wolfgang; ANDERS, Simon. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. **Genome Biology**, v. 15, n. 12, p. 550, 2014.

LU, Yu *et al.* Cancer brain metastasis: molecular mechanisms and therapeutic strategies. **Molecular Biomedicine**, v. 6, n. 1, p. 12, 25 fev. 2025.

LungTS: A new gene expression signature for prediction of response to checkpoint inhibitors in non-small cell lung cancer. | Journal of Clinical Oncology. Disponível em: <https://ascopubs.org/doi/10.1200/JCO.2022.40.16_suppl.e21143>. Acesso em: 19 ago. 2025.

LUYAPAN, Jennifer *et al.* Candidate pathway analysis of surfactant proteins identifies CTSH and SFTA2 that influences lung cancer risk. **Human Molecular Genetics**, v. 32, n. 18, p. 2842–2855, 20 jul. 2023.

LV, Bingzhe *et al.* Immunotherapy: Reshape the Tumor Immune Microenvironment. **Frontiers in Immunology**, v. 13, p. 844142, 6 jul. 2022.

MANIOS, Konstantinos *et al.* c-Fos/ c-Jun transcription factors in non-small cell lung carcinoma. **Journal of B.U.ON.: official journal of the Balkan Union of Oncology**, v. 25, n. 5, p. 2141–2143, 2020.

MELLMAN, Ira *et al.* The cancer-immunity cycle: Indication, genotype, and immunotype. **Immunity**, v. 56, n. 10, p. 2188–2205, 10 out. 2023.

MICHALCZYK, Monika *et al.* Hyaluronic Acid as a Modern Approach in Anticancer Therapy-Review. **International Journal of Molecular Sciences**, v. 24, n. 1, p. 103, 21 dez. 2022.

MITTAL, Rashmi A. *et al.* SFTA2—A Novel Secretory Peptide Highly Expressed in the Lung—Is Modulated by Lipopolysaccharide but Not Hyperoxia. **PLoS ONE**, v. 7, n. 6, p. e40011, 29 jun. 2012.

MONCHO-AMOR, Verónica *et al.* Role of Dusp6 Phosphatase as a Tumor Suppressor in Non-Small Cell Lung Cancer. **International Journal of Molecular Sciences**, v. 20, n. 8, p. 2036, 25 abr. 2019.

MUGHAL, Sadaf S. *et al.* Identification and characterization of tertiary lymphoid structures in brain metastases. **Acta Neuropathologica Communications**, v. 13, n.

1, p. 91, 3 maio 2025.

MÜLLER, Kirill; WICKHAM, Hadley. **tibble: Simple Data Frames**. , 23 mar. 2016. Disponível em: <<https://CRAN.R-project.org/package=tibble>>. Acesso em: 19 fev. 2026

NEWMAN, Aaron M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. **Nature methods**, v. 12, n. 5, p. 453–457, maio 2015.

O'DONNELL, Jake S.; TENG, Michele W. L.; SMYTH, Mark J. Cancer immunoediting and resistance to T cell-based immunotherapy. **Nature Reviews. Clinical Oncology**, v. 16, n. 3, p. 151–167, mar. 2019.

PAGÈS, Hervé *et al.* **AnnotationDbi**. Disponível em: <<http://bioconductor.org/packages/AnnotationDbi/>>. Acesso em: 9 mar. 2026.

PASTOREKOVA, Silvia; GILLIES, Robert J. The role of carbonic anhydrase IX in cancer development: links to hypoxia, acidosis, and beyond. **Cancer Metastasis Reviews**, v. 38, n. 1–2, p. 65–77, jun. 2019.

PATEL, Lalit; KOLUNDZIC, Nikola; ABEDALTHAGAFI, Malak. Progress in personalized immunotherapy for patients with brain metastasis. **npj Precision Oncology**, v. 9, n. 1, p. 31, 29 jan. 2025.

PEINADO, Héctor *et al.* Pre-metastatic niches: organ-specific homes for metastases. **Nature Reviews. Cancer**, v. 17, n. 5, p. 302–317, maio 2017.

PHAD, Ganesh E. *et al.* Clonal structure, stability and dynamics of human memory B cells and circulating plasmablasts. **Nature Immunology**, v. 23, n. 7, p. 1076–1085, 2022.

POWELL, Aoibhín M. *et al.* The epigenetic landscape of brain metastasis. **Oncogene**, v. 44, n. 27, p. 2227–2239, 2025.

RACLE, Julien *et al.* Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. **eLife**, v. 6, p. e26476, 13 nov. 2017.

RAHAL, Zahraa *et al.* Tumour and microenvironment crosstalk in NSCLC progression and response to therapy. **Nature reviews. Clinical oncology**, v. 22, n. 7, p. 463–482, jul. 2025.

RAINER, Johannes. **EnsDb.Hsapiens.v86**. Disponível em: <<http://bioconductor.org/packages/EnsDb.Hsapiens.v86/>>. Acesso em: 9 mar. 2026.

REIS, Patricia P. *et al.* mRNA transcript quantification in archival samples using multiplexed, color-coded probes. **BMC Biotechnology**, v. 11, n. 1, p. 46, 9 maio 2011.

RIELY, Gregory J. *et al.* Non-Small Cell Lung Cancer, Version 4.2024, NCCN Clinical Practice Guidelines in Oncology. **Journal of the National Comprehensive Cancer Network: JNCCN**, v. 22, n. 4, p. 249–274, maio 2024.

RIO, Donald C. *et al.* Purification of RNA using TRIzol (TRI reagent). **Cold Spring**

Harbor Protocols, v. 2010, n. 6, p. pdb.prot5439, jun. 2010.

RITCHIE, Matthew E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. **Nucleic Acids Research**, v. 43, n. 7, p. e47, 20 abr. 2015.

RIZVI, Naiyer A. *et al.* Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. **Science**, v. 348, n. 6230, p. 124–128, 3 abr. 2015.

ROBINSON, Mark D.; MCCARTHY, Davis J.; SMYTH, Gordon K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics**, v. 26, n. 1, p. 139–140, 1 jan. 2010.

RUFF, Sophie *et al.* Roles for MDC1 in cancer development and treatment. **DNA repair**, v. 95, p. 102948, nov. 2020.

SAKAGUCHI, Shimon *et al.* Regulatory T Cells and Immune Tolerance. **Cell**, v. 133, n. 5, p. 775–787, 30 maio 2008.

SCHREURS, Luca D. *et al.* The immune landscape in brain metastasis. **Neuro-Oncology**, v. 27, n. 1, p. 50–62, 15 out. 2024.

SEER. **Cancer of the lung and bronchus**. [S.l.]: Surveillance, Epidemiology, and End Results Program, 2025. Disponível em: <<https://seer.cancer.gov/statfacts/html/lungb.html>>. Acesso em: 9 mar. 2025.

SELLMER, Laura *et al.* Markers of Immune Cell Exhaustion as Predictor of Survival in Surgically-Treated Early-Stage NSCLC. **Frontiers in Immunology**, v. 13, p. 858212, 27 jun. 2022.

SIEGEL, Rebecca L. *et al.* Cancer statistics, 2023. **CA: A Cancer Journal for Clinicians**, v. 73, n. 1, p. 17–48, jan. 2023.

SONG, Dandan; LIAN, Yan; ZHANG, Lin. The potential of activator protein 1 (AP-1) in cancer targeted therapy. **Frontiers in Immunology**, v. 14, p. 1224892, 6 jul. 2023.

SOUZA, Vanessa G. P. *et al.* Comprehensive Analysis of Lung Adenocarcinoma and Brain Metastasis through Integrated Single-Cell Transcriptomics. **International Journal of Molecular Sciences**, v. 25, n. 7, p. 3779, 28 mar. 2024.

SUAREZ-TRUJILLO, Fabio *et al.* HLA-DMB alleles and haplotypes in Ecuador (Cuenca) Amerindians: Importance for HLA and disease studies. **Human Immunology**, v. 84, n. 4, p. 261–262, 1 abr. 2023.

SUI, Qihai *et al.* Genetic and microenvironmental differences in non-smoking lung adenocarcinoma patients compared with smoking patients. **Translational Lung Cancer Research**, v. 9, n. 4, p. 1407–1421, ago. 2020.

TANG, Keke *et al.* Identification of the molecular subtype and prognostic characteristics of lung adenocarcinoma based on CD8+ T cell-related gene signature. **Cancer Biomarkers**, v. 41, n. 3–4, p. 18758592241296764, 1 dez. 2024.

TANG, Shengjie *et al.* Immune Checkpoint Inhibitors in Non-Small Cell Lung Cancer: Progress, Challenges, and Prospects. **Cells**, v. 11, n. 3, p. 320, 19 jan. 2022.

TANTOH, Disline Manli *et al.* Methylation at cg05575921 of a smoking-related gene (AHRR) in non-smoking Taiwanese adults residing in areas with different PM2.5 concentrations. **Clinical Epigenetics**, v. 11, p. 69, 6 maio 2019.

THOMMEN, Daniela S.; SCHUMACHER, Ton N. T Cell Dysfunction in Cancer. **Cancer cell**, v. 33, n. 4, p. 547–562, 9 abr. 2018.

TYAGI, Abhishek; WU, Shih-Ying; WATABE, Kounosuke. Metabolism in the progression and metastasis of brain tumors. **Cancer letters**, v. 539, p. 215713, 28 jul. 2022.

US PREVENTIVE SERVICES TASK FORCE *et al.* Screening for Lung Cancer: US Preventive Services Task Force Recommendation Statement. **JAMA**, v. 325, n. 10, p. 962, 9 mar. 2021.

VASUDEVAN, Smreti; KRISHNA, Vidya; MEHTA, Anurag. Lung Cancer in Non-Smokers: Clinicopathological and Survival Differences from Smokers. **Cureus**, v. 14, n. 12, p. e32417, [S.d.].

VELDMAN-JONES, Margaret H. *et al.* Evaluating Robustness and Sensitivity of the NanoString Technologies nCounter Platform to Enable Multiplexed Gene Expression Analysis of Clinical Samples. **Cancer Research**, v. 75, n. 13, p. 2587–2593, 30 jun. 2015.

VOSKOBOINIK, Ilia; WHISSTOCK, James C.; TRAPANI, Joseph A. Perforin and granzymes: function, dysfunction and human pathology. **Nature Reviews Immunology**, v. 15, n. 6, p. 388–400, jun. 2015.

WANG, Peiyao *et al.* New insights into the biology and development of lung cancer in never smokers—implications for early detection and treatment. **Journal of Translational Medicine**, v. 21, n. 1, p. 585, 31 ago. 2023a.

WANG, Siwei *et al.* Tumor evolutionary trajectories during the acquisition of invasiveness in early stage lung adenocarcinoma. **Nature Communications**, v. 11, p. 6083, 27 nov. 2020.

WANG, Xin *et al.* Genetic Intratumor Heterogeneity Remodels the Immune Microenvironment and Induces Immune Evasion in Brain Metastasis of Lung Cancer. **Journal of Thoracic Oncology**, v. 19, n. 2, p. 252–272, 1 fev. 2024.

WANG, Yuchang *et al.* Tumor Immune Microenvironment and Immunotherapy in Brain Metastasis From Non-Small Cell Lung Cancer. **Frontiers in Immunology**, v. 13, p. 829451, 17 fev. 2022.

WANG, Zihao *et al.* Single-cell transcriptomic analyses provide insights into the cellular origins and drivers of brain metastasis from lung adenocarcinoma. **Neuro-Oncology**, v. 25, n. 7, p. 1262–1274, 6 jul. 2023b.

WCULEK, Stefanie K. *et al.* Dendritic cells in cancer immunology and immunotherapy.

Nature Reviews Immunology, v. 20, n. 1, p. 7–24, jan. 2020.

WHITE, Brian S. *et al.* Community assessment of methods to deconvolve cellular composition from bulk gene expression. **Nature Communications**, v. 15, n. 1, p. 7362, 27 ago. 2024.

WICKHAM, Hadley. **Create Elegant Data Visualisations Using the Grammar of Graphics**. Disponível em: <<https://ggplot2.tidyverse.org/>>. Acesso em: 9 mar. 2026.

WICKHAM, Hadley *et al.* Welcome to the Tidyverse. **Journal of Open Source Software**, v. 4, n. 43, p. 1686, 21 nov. 2019.

WICKHAM, Hadley *et al.* **tidyr: Tidy Messy Data**. , 19 dez. 2025. Disponível em: <<https://cran.r-project.org/web/packages/tidyr/index.html>>. Acesso em: 9 mar. 2026

WICKHAM, Hadley *et al.* **dplyr: A Grammar of Data Manipulation**. , 3 fev. 2026a. Disponível em: <<https://cran.r-project.org/web/packages/dplyr/index.html>>. Acesso em: 9 mar. 2026

WICKHAM, Hadley *et al.* **readr: Read Rectangular Text Data**. , 19 fev. 2026b. Disponível em: <<https://cran.r-project.org/web/packages/readr/index.html>>. Acesso em: 9 mar. 2026

XI, Xiaoyu *et al.* HLA-DMB correlates with antitumor immunity and an improved prognosis in endometrial carcinoma tumors. **Frontiers in Oncology**, v. 14, 23 jan. 2025.

YU, Guangchuang *et al.* clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. **OMICS: A Journal of Integrative Biology**, v. 16, n. 5, p. 284–287, 1 maio 2012.

YU, Guangchuang. **enrichplot: Visualization of Functional Enrichment Result**. 2024.

YU, Guangchuang; HE, Qing-Yu. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. **Molecular BioSystems(MBS)**, v. 12, n. 2, p. 477–479, 1 fev. 2016.

YU, Le *et al.* Targeted therapy of non-small cell lung cancer: mechanisms and clinical trials. **Frontiers in Oncology**, v. 14, p. 1451230, 2024.

ZHANG, Huan *et al.* CTRP family in diseases associated with inflammation and metabolism: molecular mechanisms and clinical implication. **Acta Pharmacologica Sinica**, v. 44, n. 4, p. 710–725, abr. 2023.

ZHANG, Qi *et al.* The spatial transcriptomic landscape of non-small cell lung cancer brain metastasis. **Nature Communications**, v. 13, n. 1, p. 5983, 10 out. 2022.

ZHANG, Tongwu *et al.* Genomic and evolutionary classification of lung cancer in never smokers. **Nature Genetics**, v. 53, n. 9, p. 1348–1359, set. 2021.

ZHANG, Yan *et al.* SLC40A1 in iron metabolism, ferroptosis, and disease: A review. **WIREs mechanisms of disease**, v. 16, n. 4, p. e1644, 2024a.

ZHANG, Zhanshuo *et al.* Five-gene prognostic model based on autophagy-dependent cell death for predicting prognosis in lung adenocarcinoma. **Scientific Reports**, v. 14, p. 26449, 2 nov. 2024b.

ZHAO, Kejia *et al.* HSP70 Family in Cancer: Signaling Mechanisms and Therapeutic Advances. **Biomolecules**, v. 13, n. 4, p. 601, 27 mar. 2023.

Tabela Suplementar 1. Dados demográficos e características patológicas de amostras primárias de adenocarcinoma pulmonar provenientes da coorte TCGA, incluindo status tabágico, idade ao diagnóstico, tipo tumoral e estadiamento patológico baseado na classificação TNM da AJCC.

Sample ID	Gender Demographic	Sample Type	Smoker Status	Age (Years)	Pathologic Diagnoses (AJCC) - TNM			Pathological Classification (Stages)
					T	N	M	
TCGA-50-7109-01A	Male	Primary tumor	Yes	60	T1	N0	M0	Stage IA
TCGA-MP-A4TJ-01A	Female	Primary tumor	Yes	62	T1	N0	M0	Stage IA
TCGA-80-5608-01A	Female	Primary tumor	Yes	No information	T1	N0	M0	Stage IA
TCGA-78-7162-01A	Male	Primary tumor	Yes	75	T1	N0	M0	Stage IA
TCGA-99-AA5R-01A	Female	Primary tumor	Yes	70	T1a	N0	M0	Stage IA
TCGA-97-A4M2-01A	Male	Primary tumor	Yes	66	T1a	N0	M0	Stage IA
TCGA-99-8032-01A	Male	Primary tumor	Yes	61	T1a	N0	M0	Stage IA
TCGA-55-8089-01A	Male	Primary tumor	Yes	56	T1a	N0	M0	Stage IA
TCGA-64-1676-01A	Male	Primary tumor	Yes	58	T1a	N0	M0	Stage IA
TCGA-95-7944-01A	Male	Primary tumor	Yes	71	T1a	N0	M0	Stage IA
TCGA-69-8255-01A	Male	Primary tumor	Yes	71	T1a	N0	M0	Stage IA
TCGA-55-8090-01A	Male	Primary tumor	Yes	80	T1a	N0	M0	Stage IA
TCGA-86-8669-01A	Male	Primary tumor	Yes	64	T1b	N0	M0	Stage IA
TCGA-MN-A4N4-01A	Male	Primary tumor	Yes	57	T1b	N0	M0	Stage IA
TCGA-44-7672-01A	Female	Primary tumor	Yes	52	T1b	N0	M0	Stage IA
TCGA-55-8203-01A	Female	Primary tumor	Yes	69	T1b	N0	M0	Stage IA
TCGA-55-8085-01A	Male	Primary tumor	Yes	64	T1b	N0	M0	Stage IA
TCGA-55-8208-01A	Female	Primary tumor	Yes	73	T1b	N0	M0	Stage IA
TCGA-78-7535-01A	Male	Primary tumor	Yes	45	T2	N0	M0	Stage IB
TCGA-75-6211-01A	Female	Primary tumor	Yes	No information	T2	N0	M0	Stage IB
TCGA-05-4420-01A	Male	Primary tumor	Yes	41	T2	N0	M0	Stage IB
TCGA-MP-A4SV-01A	Male	Primary tumor	Yes	67	T2	N0	M0	Stage IB
TCGA-44-2666-01A	Male	Primary tumor	Yes	43	T2	N0	M0	Stage IB
TCGA-75-7027-01A	Male	Primary tumor	Yes	No information	T2	N0	M0	Stage IB

TCGA-71-6725-01A	Female	Primary tumor	Yes	48	T2	N0	M0	Stage IB
TCGA-78-8662-01A	Female	Primary tumor	Yes	53	T2	N0	M0	Stage IB
TCGA-86-8673-01A	Male	Primary tumor	Yes	61	T2	N0	M0	Stage IB
TCGA-55-6972-01A	Male	Primary tumor	Yes	72	T2	N0	M0	Stage IB
TCGA-78-7542-01A	Male	Primary tumor	Yes	56	T2	N0	M0	Stage IB
TCGA-62-A46O-01A	Female	Primary tumor	Yes	65	T2	N0	M0	Stage IB
TCGA-38-4625-01A	Female	Primary tumor	Yes	66	T2a	N0	M0	Stage IB
TCGA-55-A493-01A	Female	Primary tumor	Yes	54	T2a	N0	M0	Stage IB
TCGA-55-7576-01A	Male	Primary tumor	Yes	54	T2a	N0	M0	Stage IB
TCGA-86-8073-01A	Male	Primary tumor	Yes	58	T2a	N0	M0	Stage IB
TCGA-86-8358-01A	Male	Primary tumor	Yes	44	T2a	N0	M0	Stage IB
TCGA-55-A48Y-01A	Male	Primary tumor	Yes	69	T2b	N0	M0	Stage IIA
TCGA-86-A4D0-01A	Male	Primary tumor	Yes	48	T2b	N0	M0	Stage IIA
TCGA-78-8640-01A	Male	Primary tumor	Yes	59	T1	N1	M0	Stage IIA
TCGA-44-7670-01A	Female	Primary tumor	Yes	47	T1b	N1	M0	Stage IIA
TCGA-86-6851-01A	Female	Primary tumor	Yes	73	T1b	N1	M0	Stage IIA
TCGA-86-8074-01A	Female	Primary tumor	Yes	62	T1b	N1	M0	Stage IIA
TCGA-86-7711-01A	Male	Primary tumor	Yes	70	T2a	N1	M0	Stage IIA
TCGA-86-8055-01A	Male	Primary tumor	Yes	79	T2a	N1	M0	Stage IIA
TCGA-MN-A4N1-01A	Male	Primary tumor	Yes	60	T2a	N1	M0	Stage IIA
TCGA-L9-A743-01A	Male	Primary tumor	Yes	56	T2a	N1	M0	Stage IIA
TCGA-44-8119-01A	Male	Primary tumor	Yes	73	T3	N0	M0	Stage IIB
TCGA-62-A472-01A	Male	Primary tumor	Yes	70	T3	N0	M0	Stage IIB
TCGA-50-8459-01A	Male	Primary tumor	Yes	68	T3	N0	M0	Stage IIB
TCGA-78-8648-01A	Female	Primary tumor	Yes	58	T3	N0	M0	Stage IIB
TCGA-73-A9RS-01A	Male	Primary tumor	Yes	41	T3	N0	M0	Stage IIB
TCGA-49-4510-01A	Female	Primary tumor	Yes	51	T2	N1	M0	Stage IIB
TCGA-49-4506-01A	Female	Primary tumor	Yes	68	T2	N1	M0	Stage IIB
TCGA-78-7148-01A	Male	Primary tumor	Yes	71	T2	N1	M0	Stage IIB
TCGA-78-7147-01A	Female	Primary tumor	Yes	67	T2	N1	M0	Stage IIB

TCGA-55-1596-01A	Male	Primary tumor	Yes	55	T2	N1	M0	Stage IIB
TCGA-05-4397-01A	Male	Primary tumor	Yes	65	T2	N1	M0	Stage IIB
TCGA-05-4432-01A	Male	Primary tumor	Yes	66	T2	N1	M0	Stage IIB
TCGA-MP-A4T4-01A	Female	Primary tumor	Yes	68	T2	N1	M0	Stage IIB
TCGA-86-8054-01A	Male	Primary tumor	Yes	61	T2b	N1	M0	Stage IIB
TCGA-64-5775-01A	Male	Primary tumor	Yes	71	T4	N0	M0	Stage IIIA
TCGA-64-1679-01A	Female	Primary tumor	Yes	58	T1	N2	M0	Stage IIIA
TCGA-44-6774-01A	Female	Primary tumor	Yes	56	T1	N2	M0	Stage IIIA
TCGA-86-7714-01A	Female	Primary tumor	Yes	61	T1b	N2	M0	Stage IIIA
TCGA-75-6214-01A	Female	Primary tumor	Yes	No information	T2	N2	M0	Stage IIIA
TCGA-64-1677-01A	Female	Primary tumor	Yes	77	T2	N2	M0	Stage IIIA
TCGA-78-7220-01A	Female	Primary tumor	Yes	53	T2	N2	M0	Stage IIIA
TCGA-55-1594-01A	Male	Primary tumor	Yes	68	T2	N2	M0	Stage IIIA
TCGA-44-3396-01A	Female	Primary tumor	Yes	74	T2	N2	M0	Stage IIIA
TCGA-50-5941-01A	Female	Primary tumor	Yes	55	T2a	N2	M0	Stage IIIA
TCGA-44-5643-01A	Male	Primary tumor	Yes	53	T2b	N2	M0	Stage IIIA
TCGA-05-4418-01A	Male	Primary tumor	Yes	69	T3	N2	M0	Stage IIIA
TCGA-86-8359-01A	Male	Primary tumor	Yes	52	T3	N2	M0	Stage IIIA
TCGA-78-7149-01A	Male	Primary tumor	Yes	71	T4	N0	M0	Stage IIIB
TCGA-05-4398-01A	Female	Primary tumor	Yes	47	T4	N3	M0	Stage IIIB
TCGA-38-4626-01A	Female	Primary tumor	Yes	57	T2b	N0	M0	No information
TCGA-55-5899-01A	Male	Primary tumor	Yes	58	T1a	N1	M0	No information
TCGA-55-6968-01A	Male	Primary tumor	Yes	61	T1	N0	M1	Stage IV
TCGA-78-7167-01A	Male	Primary tumor	Yes	77	T2	N0	M1	Stage IV
TCGA-53-7624-01A	Female	Primary tumor	Yes	40	T2	N0	M1	Stage IV
TCGA-05-4434-01A	Female	Primary tumor	Yes	67	T4	N1	M1	Stage IV
TCGA-97-8171-01A	Male	Primary tumor	Yes	81	T2a	N2	M1a	Stage IV
TCGA-55-8512-01A	Male	Primary tumor	Yes	41	T1a	N1	M1b	Stage IV
TCGA-49-AAR0-01A	Male	Primary tumor	Yes	57	T1	N0	MX	Stage IA
TCGA-97-7553-01A	Female	Primary tumor	Yes	58	T1	N0	MX	Stage IA

TCGA-J2-A4AD-01A	Female	Primary tumor	Yes	61	T1a	N0	MX	Stage IA
TCGA-55-7570-01A	Male	Primary tumor	Yes	60	T1a	N0	MX	Stage IA
TCGA-55-8507-01A	Male	Primary tumor	Yes	53	T1a	N0	MX	Stage IA
TCGA-L9-A444-01A	Female	Primary tumor	Yes	60	T1a	N0	MX	Stage IA
TCGA-49-AARO-01A	Female	Primary tumor	Yes	39	T1a	N0	MX	Stage IA
TCGA-91-8499-01A	Female	Primary tumor	Yes	76	T1b	N0	MX	Stage IA
TCGA-55-7903-01A	Male	Primary tumor	Yes	64	T1b	N0	MX	Stage IA
TCGA-44-7659-01A	Male	Primary tumor	Yes	70	T1b	N0	MX	Stage IA
TCGA-44-7660-01A	Male	Primary tumor	Yes	72	T2	N0	MX	Stage IB
TCGA-55-6642-01A	Male	Primary tumor	Yes	63	T2	N0	MX	Stage IB
TCGA-55-8302-01A	Male	Primary tumor	Yes	54	T2	N0	MX	Stage IB
TCGA-91-6836-01A	Female	Primary tumor	Yes	52	T2	N0	MX	Stage IB
TCGA-55-8510-01A	Female	Primary tumor	Yes	55	T2a	N0	MX	Stage IB
TCGA-55-8511-01A	Female	Primary tumor	Yes	73	T2a	N0	MX	Stage IB
TCGA-55-8301-01A	Male	Primary tumor	Yes	58	T2a	N0	MX	Stage IB
TCGA-L9-A8F4-01A	Female	Primary tumor	Yes	64	T2a	N0	MX	Stage IB
TCGA-55-8096-01A	Female	Primary tumor	Yes	67	T2a	N0	MX	Stage IB
TCGA-44-7662-01A	Male	Primary tumor	Yes	61	T2a	N0	MX	Stage IB
TCGA-69-7979-01A	Female	Primary tumor	Yes	71	T2a	N0	MX	Stage IB
TCGA-44-7669-01A	Male	Primary tumor	Yes	59	T1b	N1	MX	Stage IIA
TCGA-55-7914-01A	Female	Primary tumor	Yes	71	T1b	N1	MX	Stage IIA
TCGA-55-8508-01A	Female	Primary tumor	Yes	60	T2a	N1	MX	Stage IIA
TCGA-49-AAR9-01A	Male	Primary tumor	Yes	61	T3	N0	MX	Stage IIB
TCGA-55-7994-01A	Male	Primary tumor	Yes	81	T3	N0	MX	Stage IIB
TCGA-55-8506-01A	Female	Primary tumor	Yes	62	T3	N0	MX	Stage IIB
TCGA-44-7667-01A	Female	Primary tumor	Yes	49	T3	N0	MX	Stage IIB
TCGA-49-6767-01A	Female	Primary tumor	Yes	46	T3	N0	MX	Stage IIB
TCGA-95-7039-01A	Female	Primary tumor	Yes	54	T3	N0	MX	Stage IIB
TCGA-69-7978-01A	Male	Primary tumor	Yes	59	T2b	N1	MX	Stage IIB
TCGA-55-7727-01A	Male	Primary tumor	Yes	70	T1a	N2	MX	Stage IIIA

TCGA-49-AAR4-01A	Male	Primary tumor	Yes	51	T2	N2	MX	Stage IIIA
TCGA-55-8615-01A	Male	Primary tumor	Yes	67	T3	N2	MX	Stage IIIA
TCGA-44-5644-01A	Female	Primary tumor	Yes	51	T2a	N0	No information	Stage IB
TCGA-91-6831-01A	No information	Primary tumor	Yes	No information	No information	No information	No information	No information
TCGA-50-6673-01A	Female	Primary tumor	No	84	T1	N0	M0	Stage I
TCGA-55-6980-01A	Male	Primary tumor	No	56	T1	N0	M0	Stage IA
TCGA-86-8076-01A	Male	Primary tumor	No	42	T1	N0	M0	Stage IA
TCGA-44-3919-01A	Female	Primary tumor	No	71	T1	N0	M0	Stage IA
TCGA-97-A4M6-01A	Female	Primary tumor	No	45	T1a	N0	M0	Stage IA
TCGA-67-6216-01A	Female	Primary tumor	No	57	T1a	N0	M0	Stage IA
TCGA-86-8668-01A	Female	Primary tumor	No	61	T1b	N0	M0	Stage IA
TCGA-86-7953-01A	Female	Primary tumor	No	69	T1b	N0	M0	Stage IA
TCGA-55-8206-01A	Male	Primary tumor	No	56	T1b	N0	M0	Stage IA
TCGA-55-A57B-01A	Female	Primary tumor	No	80	T1b	N0	M0	Stage IA
TCGA-50-5066-01A	Male	Primary tumor	No	72	T2	N0	M0	Stage IB
TCGA-71-8520-01A	Female	Primary tumor	No	60	T2	N0	M0	Stage IB
TCGA-49-4501-01A	Female	Primary tumor	No	67	T2	N0	M0	Stage IB
TCGA-35-3615-01A	Male	Primary tumor	No	57	T2	N0	M0	Stage IB
TCGA-67-3772-01A	Female	Primary tumor	No	82	T2	N0	M0	Stage IB
TCGA-78-7143-01A	Female	Primary tumor	No	62	T2	N0	M0	Stage IB
TCGA-38-4630-01A	Female	Primary tumor	No	75	T2	N0	M0	Stage IB
TCGA-55-6986-01A	Female	Primary tumor	No	74	T2	N0	M0	Stage IB
TCGA-50-6597-01A	Female	Primary tumor	No	79	T2	N0	M0	Stage IB
TCGA-78-7540-01A	Female	Primary tumor	No	66	T2	N0	M0	Stage IB
TCGA-86-8075-01A	Female	Primary tumor	No	66	T2	N0	M0	Stage IB
TCGA-75-5147-01A	Female	Primary tumor	No	No information	T2	N0	M0	Stage IB
TCGA-73-7499-01A	Female	Primary tumor	No	81	T2a	N0	M0	Stage IB
TCGA-75-6205-01A	Female	Primary tumor	No	No information	T2a	N0	M0	Stage IB
TCGA-86-8585-01A	Male	Primary tumor	No	57	T2a	N0	M0	Stage IB

TCGA-97-8177-01A	Female	Primary tumor	No	59	T2a	N0	M0	Stage IB
TCGA-86-7955-01A	Male	Primary tumor	No	62	T2a	N0	M0	Stage IB
TCGA-67-6215-01A	Female	Primary tumor	No	52	T2a	N0	M0	Stage IB
TCGA-86-A4P7-01A	Female	Primary tumor	No	63	T2a	N0	M0	Stage IB
TCGA-05-5715-01A	Female	Primary tumor	No	69	T2a	N0	M0	Stage IB
TCGA-86-8280-01A	Female	Primary tumor	No	54	T2b	N0	M0	Stage IIA
TCGA-86-7713-01A	Male	Primary tumor	No	70	T2b	N0	M0	Stage IIA
TCGA-86-8279-01A	Male	Primary tumor	No	46	T2a	N1	M0	Stage IIA
TCGA-86-6562-01A	Male	Primary tumor	No	52	T2a	N1	M0	Stage IIA
TCGA-75-7030-01A	Male	Primary tumor	No	No information	T3	N0	M0	Stage IIB
TCGA-86-8672-01A	Male	Primary tumor	No	59	T3	N0	M0	Stage IIB
TCGA-62-8397-01A	Female	Primary tumor	No	70	T3	N0	M0	Stage IIB
TCGA-69-7760-01A	Male	Primary tumor	No	73	T3	N0	M0	Stage IIB
TCGA-62-8395-01A	Female	Primary tumor	No	80	T3	N0	M0	Stage IIB
TCGA-55-6982-01A	Female	Primary tumor	No	79	T2	N1	M0	Stage IIB
TCGA-86-8278-01A	Female	Primary tumor	No	63	T2	N1	M0	Stage IIB
TCGA-75-6212-01A	Female	Primary tumor	No	No information	T2	N1	M0	Stage IIB
TCGA-44-2665-01A	Female	Primary tumor	No	55	T2	N1	M0	Stage IIB
TCGA-38-4628-01A	Female	Primary tumor	No	65	T2	N1	M0	Stage IIB
TCGA-86-8671-01A	Female	Primary tumor	No	72	T2b	N1	M0	Stage IIB
TCGA-62-8402-01A	Female	Primary tumor	No	73	T2	N2	M0	Stage IIIA
TCGA-62-A46Y-01A	Female	Primary tumor	No	70	T2	N2	M0	Stage IIIA
TCGA-75-6203-01A	Female	Primary tumor	No	No information	T2	N2	M0	Stage IIIA
TCGA-05-5429-01A	Male	Primary tumor	No	60	T3	N2	M0	Stage IIIA
TCGA-62-8394-01A	Female	Primary tumor	No	65	T4	N2	M0	Stage IIIB
TCGA-50-6591-01A	Female	Primary tumor	No	63	T2	N0	M1	Stage IV
TCGA-86-7701-01A	Male	Primary tumor	No	66	T2	N0	M1	Stage IV
TCGA-05-4402-01A	Female	Primary tumor	No	57	T2	NX	M1	Stage IV
TCGA-99-8033-01A	Female	Primary tumor	No	74	TX	NX	M1	Stage IV
TCGA-93-A4JP-01A	Male	Primary tumor	No	64	TX	NX	M1b	Stage IV

TCGA-97-8552-01A	Female	Primary tumor	No	55	T1a	N0	MX	Stage I
TCGA-95-8039-01A	Male	Primary tumor	No	72	T1	N0	MX	Stage IA
TCGA-91-8497-01A	Female	Primary tumor	No	75	T1a	N0	MX	Stage IA
TCGA-J2-A4AE-01A	Female	Primary tumor	No	77	T1a	N0	MX	Stage IA
TCGA-55-8087-01A	Female	Primary tumor	No	59	T2a	N0	MX	Stage IB
TCGA-55-7815-01A	Male	Primary tumor	No	76	T2a	N0	MX	Stage IB
TCGA-91-8496-01A	Female	Primary tumor	No	63	T2a	NX	MX	Stage IB
TCGA-49-AAQV-01A	Female	Primary tumor	No	63	T1	N1	MX	Stage II
TCGA-55-6978-01A	Male	Primary tumor	No	81	T2b	N0	MX	Stage IIA
TCGA-J2-8192-01A	Female	Primary tumor	No	65	T2a	N1	MX	Stage IIA
TCGA-55-8619-01A	Female	Primary tumor	No	72	T3	N0	MX	Stage IIB
TCGA-55-8513-01A	Female	Primary tumor	No	77	T3	N0	MX	Stage IIB
TCGA-55-8505-01A	Male	Primary tumor	No	62	T1a	N2	MX	Stage IIIA
TCGA-86-A4P8-01A	Female	Primary tumor	No	59	T1b	N2	MX	Stage IIIA
TCGA-49-4512-01A	Female	Primary tumor	No	69	T2	N2	MX	Stage IIIA
TCGA-97-8547-01A	Female	Primary tumor	No	78	T2a	N2	MX	Stage IIIA
TCGA-55-7816-01A	Female	Primary tumor	No	49	TX	NX	MX	Stage IV
TCGA-38-6178-01A	Female	Primary tumor	No	70	T2b	N2	No information	Stage IIIA
TCGA-44-2661-01A	No information	Primary tumor	No	No information	No information	No information	No information	No information

Tabela Suplementar 2. Características clínicas e patológicas de tumores pulmonares primários e amostras metastáticas analisadas por NanoString® e sequenciamento de nova geração (NGS), incluindo status tabágico, idade ao diagnóstico e estadiamento baseado na classificação TNM.

Sample ID	Gender Demographic	Sample Type	Smoker Status	Age (years)	Methodology	TNM Clinical Classification			Pathological Classification
						T	N	M	Stages
PNS1	Male	Primary Tumor	No	71	NanoString Technology only	T4	N0	M1c	IV
PNS2	Female	Primary Tumor	No	69	NanoString Technology only	T1b	N0	M0	IA2
PNS3	Female	Primary Tumor	No	63	NanoString Technology only	T4	N0	M0	IIIA
PNS4	Female	Primary Tumor	No	59	NanoString Technology only	T1b	N0	M0	IA2
PNS5	Female	Primary Tumor	No	71	NanoString Technology only	T1c	N2	M0	IIIA
MNS1	Female	Metastasis	No	65	NanoString Technology and NGS	T2	N0	M1b	IVA
MNS2	Female	Metastasis	No	48	NanoString Technology and NGS	T2	N2	M1c	IV
MNS3	Male	Metastasis	No	42	NanoString Technology and NGS	T1b	N0	M1c	IV
MNS4	Male	Metastasis	No	63	NanoString Technology and NGS	T4	N1	M1c	IV
MS1	Female	Metastasis	Yes	58	NGS only	T1c	N0	M1c	IV
MS2	Male	Metastasis	Yes	73	NGS only	T2	N0	M1b	IVA
MS3	Male	Metastasis	Yes	51	NGS only	T4	N2	M1c	IV

Tabela Suplementar 3. Genes codificantes diferencialmente expressos em pacientes não fumantes em relação aos fumantes (grupo de referência), a partir de dados do TCGA, identificados pelo pacote *limma*. A análise utilizou o teste t moderado, considerando significância estatística de $p < 0,05$.

https://docs.google.com/spreadsheets/d/1tnOskh_qFdVdfZmIUqVc0XXLwO1t1u7s/edit?usp=drive_link&oid=105021159038736184705&rtpof=true&sd=true

Tabela Suplementar 4. Genes codificantes diferencialmente expressos em pacientes não fumantes em relação aos fumantes (grupo de referência), a partir de dados de NGS-RNA-Seq de amostras clínicas, identificados pelo *DESeq2*. A análise utilizou modelo baseado na distribuição binomial negativa, considerando significância estatística de $p < 0,05$.

Gene ID	Gene	FC	padj	Gene ID	Gene	FC	padj
<i>ABCF1</i>	ATP binding cassette subfamily F member 1	1.010273966	4.57003E-12	<i>CPLX2</i>	Complexin 2	0.94349	0.011588
<i>RP11-577H5.5</i>	POLR2J2-UPK3BL1 readthrough	1.011145682	4.57003E-12	<i>CPAMD8</i>	C3 and PZP like alpha-2-macroglobulin domain containing 8	4.270716	0.011972
<i>TRIM26</i>	Tripartite motif containing 26	1.010396009	4.57003E-12	<i>BFAR</i>	Bifunctional apoptosis regulator	2.467642	0.012263
<i>MUC21</i>	Mucin 21, cell surface associated	1.009595702	4.57003E-12	<i>DUSP6</i>	Dual specificity phosphatase 6	5.174485	0.012761
<i>SFTA2</i>	Surfactant associated 2	1.010136364	7.02221E-12	<i>TSPAN7</i>	Tetraspanin 7	0.83904	0.013144
<i>UGT1A6</i>	UDP glucuronosyltransferase family 1 member A6	0.883525738	7.02221E-12	<i>HLA-DMB</i>	Major histocompatibility complex, class II, DM beta	0.903836	0.013328
<i>ABHD16A</i>	Abhydrolase domain containing 16A, phospholipase	1.010840048	7.02221E-12	<i>FGFBP1</i>	Fibroblast growth factor binding protein 1	27.07065	0.013328
<i>NEU1</i>	Neuraminidase 1	1.009969297	3.88703E-11	<i>FBL</i>	Fibrillarin	18.7011	0.013414
<i>PIWIL3</i>	Piwi like RNA-mediated gene silencing 3	1.010088753	9.60501E-11	<i>TMEM63C</i>	Transmembrane protein 63C	9.415338	0.013812
<i>DDR1</i>	Discoidin domain receptor tyrosine kinase 1	8.303197975	1.01903E-09	<i>ONECU T2</i>	One cut homeobox 2	0.935511	0.014038
<i>UGT1A4</i>	UDP glucuronosyltransferase family 1 member A4	0.000430107	1.22162E-09	<i>PSCA</i>	Prostate stem cell antigen	0.969823	0.014089
<i>UGT1A1</i>	UDP glucuronosyltransferase family 1	0.003389237	1.95741E-	<i>SAGE1</i>	Sarcoma antigen 1	0.969038	0.01443

	member A1		08				
<i>PNP</i>	Purine nucleoside phosphorylase	382.690612	2.05613E-08	<i>CXCL2</i>	C-X-C motif chemokine ligand 2	2.784666	0.014582
<i>ZNF502</i>	Zinc finger protein 502	85.28320583	1.02318E-07	<i>TMEM47</i>	Transmembrane protein 47	0.380315	0.015073
<i>AKR1C2</i>	Aldo-keto reductase family 1 member C2	0.933115503	1.55137E-06	<i>SFTA2</i>	Surfactant associated 2	0.973568	0.015102
<i>NELFE</i>	Negative elongation factor complex member E	0.925245858	3.37767E-06	<i>DDR1</i>	Discoidin domain receptor tyrosine kinase 1	46.946	0.015572
<i>ABC13-47488600 E17.1</i>	Chromosome 13 open reading frame 46	697.6711192	5.086E-06	<i>SNCB</i>	Synuclein beta	0.962839	0.016209
<i>C4B_2</i>	Complement component 4B (Chido/Rodgers blood group), copy 2	337.4624607	8.72367E-06	<i>COL21A1</i>	Collagen type XXI alpha 1 chain	8.996714	0.016468
<i>PCDHB11</i>	Protocadherin beta 11	3240.737727	3.74998E-05	<i>ITM2C</i>	Integral membrane protein 2C	0.291146	0.017887
<i>ABC13-47488600 E17.1</i>	Chromosome 13 open reading frame 46	0.892933015	7.57678E-05	<i>CKMT1B</i>	Creatine kinase, mitochondrial 1B	0.197639	0.018893
<i>RP11-321N4.5</i>	Synaptotagmin binding cytoplasmic RNA interacting protein	0.89032643	9.98077E-05	<i>PTCHD1</i>	Patched domain containing 1	0.136278	0.019105
<i>UGT1A7</i>	UDP glucuronosyltransferase family 1 member A7	0.893346384	0.000144599	<i>CATSPER2</i>	Cation channel sperm associated 2	0.360284	0.019105
<i>GPC4</i>	Glypican 4	7.004103434	0.000248007	<i>GABRG2</i>	Gamma-aminobutyric acid type A receptor subunit gamma2	0.953905	0.019519
<i>CDX2</i>	Caudal type homeobox 2	0.951571182	0.000248007	<i>GABRA1</i>	Gamma-aminobutyric acid type A receptor subunit alpha1	0.965687	0.022511

<i>ACKR2</i>	Atypical chemokine receptor 2	34.5200852	0.00032044	<i>GPC2</i>	Glypican 2	3.593419	0.023103
<i>IFT88</i>	Intraflagellar transport 88	0.278529015	0.00032044	<i>HLA-DPB1</i>	Major histocompatibility complex, class II, DP beta 1	1.023053	0.023234
<i>SATB2</i>	SATB homeobox 2	0.290679563	0.000338264	<i>NEURL1</i>	Neuralized E3 ubiquitin protein ligase 1	0.961602	0.023258
<i>DDR1</i>	Discoidin domain receptor tyrosine kinase 1	0.96169011	0.000396006	<i>CXCL8</i>	C-X-C motif chemokine ligand 8	7.343279	0.024246
<i>OR2A4</i>	Olfactory receptor family 2 subfamily A member 4	276.6568293	0.000422527	<i>NDFIP2</i>	Nedd4 family interacting protein 2	0.418537	0.024757
<i>SNAP25</i>	Synaptosome associated protein 25	0.829170448	0.000442026	<i>DDC</i>	Dopa decarboxylase	0.966787	0.024757
<i>MTUS2</i>	Microtubule associated scaffold protein 2	0.073175542	0.000442026	<i>CYP4F11</i>	Cytochrome P450 family 4 subfamily F member 11	0.964796	0.024866
<i>NBPF26</i>	NBPF member 26	10.60625829	0.000564309	<i>TNC</i>	Tenascin C	7.624164	0.024866
<i>SLC1A7</i>	Solute carrier family 1 member 7	108.3710414	0.000569767	<i>EPHA6</i>	EPH receptor A6	0.938039	0.025418
<i>C14orf80</i>	Tubulin epsilon and delta complex 1	4.238016776	0.000665851	<i>PPP1R11</i>	Protein phosphatase 1 regulatory inhibitor subunit 11	0.981715	0.026022
<i>EVA1A</i>	Eva-1 homolog A, regulator of programmed cell death	9.631234593	0.00068233	<i>CTAGE9</i>	CTAGE family member 9	1.034777	0.026022
<i>TSC22D3</i>	TSC22 domain family member 3	3.555449346	0.00097205	<i>P2RY8</i>	P2Y receptor family member 8	1.019783	0.026278
<i>TMC4</i>	Transmembrane channel like 4	45.13020241	0.00097205	<i>LILRB4</i>	Leukocyte immunoglobulin like receptor B4	0.909088	0.026958
<i>AKR1C1</i>	Aldo-keto reductase family 1 member C1	0.941712082	0.001104807	<i>PTPRN2</i>	Protein tyrosine phosphatase receptor type N2	0.885131	0.027042

<i>TCF7L1</i>	Transcription factor 7 like 1	79.00171146	0.00117155 4	<i>HTR3A</i>	5-hydroxytryptamine receptor 3A	13.82185	0.027042
<i>NPC2</i>	NPC intracellular cholesterol transporter 2	4.012290839	0.00150068 7	<i>POTEE</i>	POTE ankyrin domain family member E	0.950481	0.02724
<i>NOTCH4</i>	Notch receptor 4	1060.757673	0.00150068 7	<i>TMC4</i>	Transmembrane channel like 4	0.182051	0.027323
<i>PRSS51</i>	Serine protease 51	0.891383143	0.00150068 7	<i>SLITRK6</i>	SLIT and NTRK like family member 6	25.21532	0.027323
<i>AKR1C3</i>	Aldo-keto reductase family 1 member C3	0.937076757	0.00150068 7	<i>STAC</i>	SH3 and cysteine rich domain	5.474451	0.027972
<i>GPX2</i>	Glutathione peroxidase 2	0.925083333	0.00170038 9	<i>STXBP5L</i>	Syntaxin binding protein 5L	0.914447	0.028196
<i>TCF19</i>	Transcription factor 19	1447.953024	0.00233923 3	<i>CXCL6</i>	C-X-C motif chemokine ligand 6	23.30433	0.028196
<i>TRIM39</i>	Tripartite motif containing 39	831.6819771	0.00279310 3	<i>CFB</i>	Complement factor B	21.71472	0.02837
<i>CFAP47</i>	Cilia and flagella associated protein 47	0.933614359	0.00279310 3	<i>PBX2</i>	PBX homeobox 2	0.959378	0.033289
<i>CES1</i>	Carboxylesterase 1	0.938541452	0.00281529 7	<i>SLC7A11</i>	Solute carrier family 7 member 11	0.893763	0.033469
<i>CEP72</i>	Centrosomal protein 72	0.91704454	0.00289968 6	<i>MAP2</i>	Microtubule associated protein 2	0.346738	0.033469
<i>HCN3</i>	Hyperpolarization activated cyclic nucleotide gated potassium channel 3	23.11161733	0.00291079 6	<i>RP4-583P15.14</i>	NA	1.026254	0.033768
<i>NOTCH4</i>	Notch receptor 4	6.791698306	0.00298633 1	<i>NEIL2</i>	Nei like DNA glycosylase 2	0.217516	0.034612

<i>HSPA1B</i>	Heat shock protein family A (Hsp70) member 1B	39.25606033	0.00298633 1	<i>AKAP11</i>	A-kinase anchoring protein 11	0.502181	0.034746
<i>KCNJ15</i>	Potassium inwardly rectifying channel subfamily J member 15	49.61817336	0.00298633 1	<i>HSPA1B</i>	Heat shock protein family A (Hsp70) member 1B	7.110752	0.035343
<i>XPO4</i>	Exportin 4	0.431599872	0.00355992 6	<i>GGT5</i>	Gamma-glutamyltransferase 5	3.516342	0.035343
<i>FAM181A</i>	Family with sequence similarity 181 member A	0.928719795	0.00374368 8	<i>AC012314.6</i>	NADH:ubiquinone oxidoreductase subunit A3	0.980019	0.035343
<i>HAS3</i>	Hyaluronan synthase 3	5.536075998	0.00411701 7	<i>FOS</i>	Fos proto-oncogene, AP-1 transcription factor subunit	3.441081	0.03577
<i>ZNF546</i>	Zinc finger protein 546	67.81622557	0.00440924 5	<i>TPD52L1</i>	TPD52 like 1	0.336085	0.036096
<i>UCHL1</i>	Ubiquitin C-terminal hydrolase L1	0.150302484	0.00451662	<i>ACO1</i>	Aconitase 1	2.230522	0.03675
<i>HLA-DMB</i>	Major histocompatibility complex, class II, DM beta	1.022062625	0.00486439	<i>MICA</i>	MHC class I polypeptide-related sequence A	2.752116	0.03918
<i>ALDH1A1</i>	Aldehyde dehydrogenase 1 family member A1	0.120533051	0.00493301 5	<i>TMEM64</i>	Transmembrane protein 64	0.230686	0.039213
<i>ZMYM2</i>	Zinc finger MYM-type containing 2	0.430486346	0.00522861	<i>BLVRA</i>	Biliverdin reductase A	2.543697	0.039706
<i>ADAM23</i>	ADAM metallopeptidase domain 23	0.200600036	0.00535370 1	<i>KIAA1456</i>	tRNA methyltransferase 9B (putative)	0.927453	0.042866
<i>RPP21</i>	Ribonuclease P/MRP subunit p21	437.5058207	0.00668957 8	<i>SYN2</i>	Synapsin II	0.953954	0.043001
<i>TFF1</i>	Trefoil factor 1	0.978419873	0.00714311 7	<i>SYT12</i>	Synaptotagmin 12	7.833364	0.043001
<i>CRABP2</i>	Cellular retinoic acid binding protein 2	6.040715972	0.00726307	<i>UNC13C</i>	Unc-13 homolog C	0.958305	0.043001

<i>ARHGDIB</i>	Rho GDP dissociation inhibitor beta	2.777915395	0.00726307	<i>SYNE4</i>	Spectrin repeat containing nuclear envelope family member 4	3.06567	0.043543
<i>MYO3A</i>	Myosin IIIA	0.97209773	0.00739447 3	<i>SLCO2A1</i>	Solute carrier organic anion transporter family member 2A1	4.12112	0.04454
<i>TUBB</i>	Tubulin beta class I	0.978969736	0.00755759 4	<i>MAGIX</i>	MAGI family member, X-linked	3.542051	0.044672
<i>CACNA1B</i>	Calcium voltage-gated channel subunit alpha1 B	0.949496629	0.00811813 7	<i>TFPI2</i>	Tissue factor pathway inhibitor 2	8.794237	0.044672
<i>ENPP3</i>	Ectonucleotide pyrophosphatase/phosphodiesterase 3	15.37269858	0.00867708 8	<i>VAR2</i>	Valyl-tRNA synthetase 2, mitochondrial	22.20468	0.046556
<i>RCN3</i>	Reticulocalbin 3	6.214952193	0.00884675 4	<i>EVPLL</i>	Envoplakin like	1.022742	0.04663
<i>CYTH3</i>	Cytohesin 3	2.21172118	0.00899103 9	<i>TAP2</i>	Transporter 2, ATP binding cassette subfamily B member	1.026451	0.046649
<i>TSSK2</i>	Testis specific serine kinase 2	111.8568535	0.00911801	<i>F2RL3</i>	F2R like thrombin or trypsin receptor 3	6.489994	0.046649
<i>TFF3</i>	Trefoil factor 3	0.94850217	0.00934758	<i>PAPPA2</i>	Pappalysin 2	0.959459	0.046677
<i>NEFL</i>	Neurofilament light chain	0.957224886	0.01078070 4	<i>CUTA</i>	CutA divalent cation tolerance homolog	2.015809	0.046677
<i>ZNF436</i>	Zinc finger protein 436	0.980406467	0.01090625 7	<i>RB1</i>	RB transcriptional corepressor 1	0.389754	0.046731
<i>MDC1</i>	Mediator of DNA damage checkpoint 1	1.023887875	0.01105234 3	<i>KCNV1</i>	Potassium voltage-gated channel modifier subfamily V member 1	0.959771	0.04892
<i>SLCO5A1</i>	Solute carrier organic anion transporter family member 5A1	7.515792866	0.01122973 3	<i>CCHCR1</i>	Coiled-coil alpha-helical rod protein 1	21.19161	0.049506
<i>C14orf39</i>	Chromosome 14 open reading frame 39	170.0736286	0.01148471 2	<i>CX3CL1</i>	C-X3-C motif chemokine ligand 1	5.351215	0.049852

<i>CYP8B1</i>	Cytochrome P450 family 8 subfamily B member 1	37.00126256	0.01148471 2
---------------	---	-------------	-----------------

Tabela Suplementar 5. Lista dos 42 genes concordantes entre as análises de tumor primário (TCGA) e metástase cerebral (NGS-RNA-Seq clínico), considerando a comparação entre pacientes não fumantes em relação aos fumantes (grupo de referência). Foram incluídos apenas genes codificadores de proteínas com valor de p ajustado < 0,05. Os 30 primeiros genes apresentam padrão de expressão na mesma direção entre os dois conjuntos de dados, enquanto os 12 últimos exibem inversão de expressão, estando destacados em negrito.

Gene	Gene ID	FC TCGA	FC RNA-Seq
<i>KCNJ15</i>	Potassium Inwardly Rectifying Channel Subfamily J Member 15	1.573547253	5.63279672
<i>ENPP3</i>	Ectonucleotide Pyrophosphatase/Phosphodiesterase 3	1.640622126	3.942298538
<i>HAS3</i>	Hyaluronan Synthase 3	1.448873774	2.468863748
<i>ARHGDIB</i>	Rho GDP Dissociation Inhibitor Beta	1.104937888	2.777915395
<i>STXBP5L</i>	Syntaxin Binding Protein 5 Like	0.895464979	0.914446519
<i>NEURL1</i>	Neuralized E3 Ubiquitin Protein Ligase 1	0.726449315	0.9616019
<i>TSC22D3</i>	TSC22 Domain Family Member 3	1.109542704	3.555449346
<i>HLA-DPB1</i>	Major Histocompatibility Complex, Class II, DP Beta 1	1.113551335	1.023052612
<i>FOS</i>	Fos Proto-Oncogene, AP-1 Transcription Factor Subunit	1.133155376	3.441081259
<i>EVA1A</i>	EVA-1 Homolog A	1.221707231	9.631234593
<i>NPC2</i>	Niemann-Pick Disease, Type C2	1.09749774	4.012290839
<i>CEP72</i>	Centrosomal Protein 72	0.87024565	0.91704454
<i>KCNV1</i>	Potassium Voltage-Gated Channel Modifier Subfamily V Member 1	0.885674259	0.959770579
<i>GPC4</i>	Glypican 4	1.157973176	7.004103434
<i>UCHL1</i>	Ubiquitin C-Terminal Hydrolase L1	0.716753097	0.150302484
<i>STAC</i>	SH3 and Cysteine Rich Domain Containing	1.296353086	5.47445136
<i>SLC1A7</i>	Solute Carrier Family 1 Member 7	1.446200451	108.3710414
<i>DUSP6</i>	Dual Specificity Phosphatase 6	1.132218987	5.174485006
<i>BLVRA</i>	Biliverdin Reductase A	1.066989848	2.543697112
<i>HTR3A</i>	5-Hydroxytryptamine Receptor 3A	1.390113431	13.82185177
<i>CYTH3</i>	Cytohesin 3	1.111019289	2.21172118

<i>BFAR</i>	Bifunctional Apoptosis Regulator	1.061870563	2.467642123
<i>NELFE</i>	Negative Elongation Factor Complex Member E	0.951559959	0.925245858
<i>ZMYM2</i>	Zinc Finger MYM-Type Containing 2	0.934205061	0.430486346
<i>UGT1A4</i>	UDP Glucuronosyltransferase Family 1 Member A4	0.971776196	0.000430107
<i>CPAMD8</i>	C3 and PZP Like Alpha-2-Macroglobulin Domain Containing 8	1.263784529	4.270715781
<i>CX3CL1</i>	C-X3-C Motif Chemokine Ligand 1	1.184206961	5.351214629
<i>CPLX2</i>	Complexin 2	0.755898852	0.943490172
<i>UGT1A7</i>	UDP Glucuronosyltransferase Family 1 Member A7	0.908052631	0.893346384
<i>AKR1C2</i>	Aldo-Keto Reductase Family 1 Member C2	0.670640219	0.933115503
<i>HLA-DMB</i>	Major Histocompatibility Complex, Class II, DM Beta	1.188448553	0.90383635
<i>MDC1</i>	Mediator of DNA Damage Checkpoint 1	0.900645279	1.023887875
<i>GPC2</i>	Glypican 2	0.757031727	3.593419155
<i>UNC13C</i>	Unc-13 Homolog C	1.061883077	0.95830536
<i>VARS2</i>	Valyl-tRNA Synthetase 2, Mitochondrial	0.919964022	22.20467695
<i>ABCF1</i>	ATP Binding Cassette Subfamily F Member 1	0.951697411	1.010273966
<i>HCN3</i>	Hyperpolarization Activated Cyclic Nucleotide Gated Potassium Channel 3	0.875941964	23.11161733
<i>SFTA2</i>	Surfactant Associated 2	1.228367983	0.973567538
<i>FBL</i>	Fibrillarin	0.959020056	18.70109709
<i>CCHCR1</i>	Coiled-Coil Alpha-Helical Rod Protein 1	0.932049275	21.19161197
<i>ZNF436</i>	Zinc Finger Protein 436	1.062096769	0.980406467
<i>HSPA1B</i>	Heat Shock Protein Family A (Hsp70) Member 1B	0.935004619	7.11075239

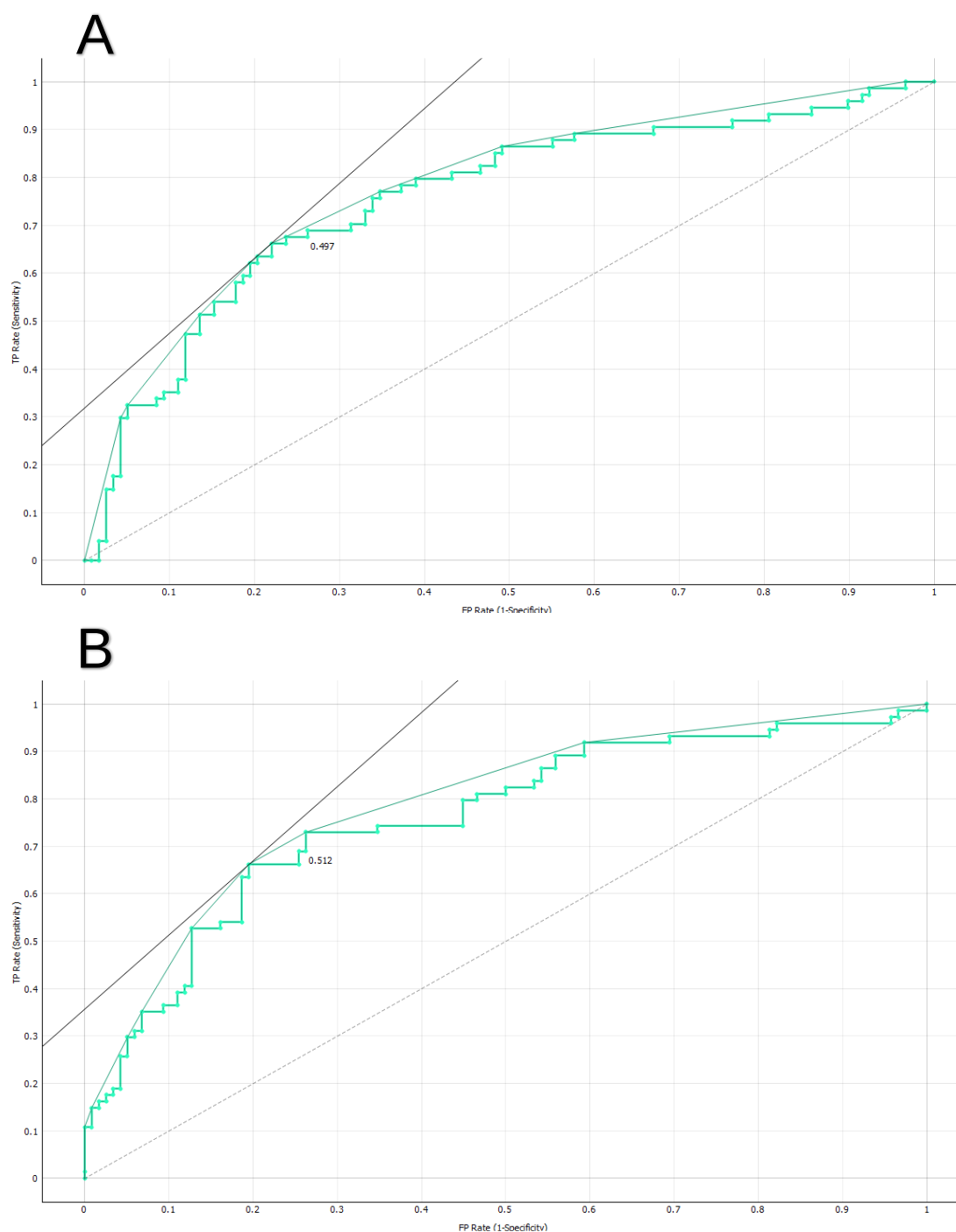


Figura Suplementar 1. Curva ROC gerada pelo programa Orange Data Mining. (A) Curva ROC dos dados do algoritmo xCell (AUC de 0.497), (B) Curva ROC dos dados do algoritmo CIBERSORT (AUC de 0.512).