

Modelagem Computacional Da Maturação E Clonalidade De Linfócitos B No Adenocarcinoma De Cólon

Computational Modeling Of The Maturation And Clonality Of B Lymphocytes In Colon Adenocarcinoma

Luiz Filipi Duarte Sarmento¹, Fabio Albuquerque Marchi², Carina Mucciolo Melo¹

¹ Centro Universitário Saúde ABC (FMABC), Faculdade de Medicina do ABC, Santo André–SP, Brasil

² Instituto do Câncer do Estado de São Paulo – Octavio Frias de Oliveira (ICESP), Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP), São Paulo–SP, Brasil

Luiz Filipi Duarte Sarmento

Centro Universitário Saúde ABC (FMABC) – Faculdade de Medicina do ABC

Av. Príncipe de Gales, 821 – Vila Príncipe de Gales

Santo André–SP, CEP 09060-650, Brasil

Telefone: +55 (11) 93717-6576; E-mail: luiz.sarmento@aluno.fmabc.net

RESUMO

Introdução: O microambiente tumoral (TME) do adenocarcinoma de cólon (COAD) abriga linfócitos B e estruturas linfoïdes terciárias (TLS), que atuam como nichos de ativação imune local. Embora a infiltração de células B e a presença de TLS estejam associadas ao prognóstico, ainda são pouco compreendidas as dinâmicas de maturação, diversidade clonal e plasticidade funcional das células B no COAD. **Objetivo:** Caracterizar a diferenciação e a clonalidade de células B no COAD, comparando tecido tumoral e adjacente por meio da integração de scRNA-seq e scBCR-seq. **Métodos:** A reanálise observacional e computacional utilizou amostras de COAD extraídas de um atlas unicelular multicâncer. A projeção UMAP específica para COAD foi gerada no Seurat, preservando as anotações originais. As análises incluíram subtipagem de células B, estimativas de diversidade e clonalidade do repertório (índices de Shannon e Simpson), intensidade de hipermutação somática (SHM; v_identity/mu_freq) e distribuição isotípica para avaliar a troca de classe (CSR; c_call). **Resultados:** A coorte incluiu cerca de 19 mil células B (15.285 no tumor e 3.714 no tecido adjacente, de 8 pacientes). A projeção UMAP identificou três agrupamentos principais e um continuum de diferenciação que se estendeu de células naïve a plasmablastos/plasmócitos, passando por estados de memória/ativação e centro germinativo. O tumor apresentou enriquecimento de subpopulações pós-CSR (ITGB1 /SwBm; NR4A2 /ACB2), enquanto o tecido adjacente reteve maiores proporções de células naïve e plasmócitos. O repertório foi policlonal, com tendência a maior diversidade no tumor (índice de Shannon), enquanto a dominância permaneceu elevada e semelhante entre os grupos (índice de Simpson; Wilcoxon pareado, p-ajust.=0,354). A SHM foi mais intensa em

IgG/IgA, ao passo que IgM/IgD mantiveram maior identidade V, compatível com estados naïve. A CSR mostrou maior frequência relativa de isótipos pós-CSR no tumor, com aumento de IgA (IGHA1) e tendência a IgG; as diferenças globais foram significativas ($\chi^2=198,4$; gl=6; $p<0,0001$). **Conclusão:** O TME do COAD apresenta enriquecimento de células B ativadas/memória pós-CSR e maior intensidade de SHM/CSR no tumor, compatível com ativação antígeno-dependente em TLS. As principais limitações incluem o número reduzido de pacientes, o desequilíbrio entre tumor e adjacente e a análise restrita à cadeia pesada de imunoglobulina (IGH).

Palavras-Chave: COAD; linfócitos B; scRNA-seq; scBCR-seq; hipermutação somática; estruturas linfoïdes terciárias.

ABSTRACT

Introduction: The tumor microenvironment (TME) of colorectal adenocarcinoma (COAD) harbors B lymphocytes and tertiary lymphoid structures (TLS), which act as local immune activation niches. Although B-cell infiltration and TLS presence have been linked to prognosis, the dynamics of maturation, clonal diversity, and functional plasticity of B cells in COAD remain poorly understood. **Objective:** To characterize B-cell differentiation and clonality in COAD by comparing tumor and adjacent tissues through integrated scRNA-seq and scBCR-seq analyses. **Methods:** An observational and computational reanalysis was performed using COAD samples extracted from a pan-cancer single-cell atlas. A COAD-specific UMAP projection was generated in Seurat while preserving original annotations. Analyses included B-cell subtyping, repertoire diversity and clonality estimates (Shannon and Simpson indices), somatic hypermutation (SHM; v_identity/mu_freq), and isotype distribution to evaluate class-switch recombination (CSR; c_call). **Results:** The cohort comprised ~19,000 B cells (15,285 from tumor and 3,714 from adjacent tissue, across 8 patients). The UMAP projection identified three main clusters and a differentiation continuum extending from naïve B cells to plasmablasts/plasma cells, passing through memory/activated and germinal center states. Tumor tissue showed enrichment of post-CSR subpopulations (ITGB1 /SwBm; NR4A2 /ACB2), whereas adjacent tissue retained higher proportions of naïve B cells and plasma cells. The repertoire was polyclonal, with a trend toward greater diversity in tumor (Shannon index), while dominance remained high and comparable between groups (Simpson index; paired Wilcoxon, adjusted p=0.354). SHM was more pronounced in IgG/IgA, whereas IgM/IgD retained higher V-gene identity, consistent with naïve states. CSR analysis revealed a higher relative

frequency of post-CSR isotypes in tumor, with increased IgA (IGHA1) and a trend toward IgG; overall differences were significant ($\chi^2=198.4$; df=6; p<0.0001).

Conclusion: The COAD TME is enriched with post-CSR activated/memory B cells and displays intensified SHM and CSR in tumor tissue, consistent with antigen-dependent activation within TLS. Main limitations include the small patient cohort, imbalance between tumor and adjacent tissues, and restriction of analyses to immunoglobulin heavy chains (IGH).

Keywords: COAD; B lymphocytes; scRNA-seq; scBCR-seq; somatic hypermutation; tertiary lymphoid structures.

INTRODUÇÃO

O adenocarcinoma de cólon é o subtipo mais prevalente do câncer colorretal, que figura entre as principais causas de incidência e mortalidade por câncer em todo o mundo. Estimativas recentes apontam mais de 1,9 milhão de novos casos e 930 mil óbitos anuais, tornando-o um dos maiores problemas de saúde pública global (1). A ocorrência aumenta com a idade e é influenciada por fatores de risco ambientais e comportamentais, como dieta rica em gordura e carne processada, consumo de álcool, tabagismo, sedentarismo e obesidade, além de condições hereditárias, como síndrome de Lynch e polipose adenomatosa familiar (2,3). Nas últimas décadas, o microambiente tumoral emergiu como determinante da progressão e do prognóstico, sendo composto por células neoplásicas, estroma (fibroblastos, vasos e matriz extracelular) e infiltrado imune (4). A composição imune do TME apresenta forte valor prognóstico e fundamentou o desenvolvimento do Immunoscore, baseado na densidade de linfócitos T infiltrantes (5).

Embora as terapias imunológicas tenham historicamente se concentrado nos linfócitos T, linfócitos B infiltrados em tumores (TIL-B) e plasmócitos vêm ganhando destaque na resposta antitumoral. Esses TIL-B apresentam抗ígenos a células T auxiliares, secretam citocinas e originam plasmócitos produtores de anticorpos intratumoriais (6). Dessa forma, podem amplificar tanto respostas adaptativas quanto inatas, por meio da opsonização tumoral, ativação do complemento e recrutamento de células NK. Um elemento central nesse processo são as estruturas linfoides terciárias, agregados ectópicos de células B, T e dendríticas foliculares que mimetizam linfonodos secundários e constituem nichos locais de ativação e maturação imune (7). Evidências multicêntricas demonstram que TLS maduros contendo centros germinativos associam-

se a melhor sobrevida e maior resposta a bloqueio de checkpoints imunes em melanoma, carcinoma renal e câncer colorretal (8).

Em câncer colorretal, TLS maduros associam-se a maior infiltração de CD8 e melhor resposta à terapia neoadjuvante (9). As células B têm papel ambivalente: Bregs podem favorecer o tumor via IL-10 e outros mediadores imunossupressores (10). Em contraste, TLS maduros geram células B de memória e plasmócitos de longa duração, com anticorpos de alta afinidade que se ligam às células tumorais, podem induzir apoptose e potencializar a resposta à imunoterapia (11). Assim, o impacto das células B no COAD depende do fenótipo, do estágio de maturação e da organização em TLS.

A diversidade funcional dos linfócitos B decorre de clonalidade, SHM e CSR. Nos centros germinativos (CG) há proliferação intensa e SHM em genes de Ig; em paralelo, a CSR troca IgM por IgG/IgA (12). Esses processos selecionam clones de alta afinidade e geram plasmócitos e células B de memória. No tumor, TLS maduros recapitulam esse circuito, com SHM em anticorpos intratumorais e CSR associada a maior atividade efetora (13). Em tumores epiteliais, análises single-cell + transcriptômica espacial mostram TLS com reação de CG que amadurece plasmócitos, integra-os ao leito tumoral e se associa a melhor resposta a anti-PD-1 (14). Além disso, variantes de IgG1 podem remodelar o TME e inibir a tumorigênese no cólon (15).

Um estudo recente publicado na Science integrar scRNA-seq e scBCR-seq pareados em escala pan-câncer (474.718 transcritomas; 166.733 células com cadeias pesada e leve), estabelecendo o primeiro mapa sistemático da clonalidade, diversidade, hipermutação somática e troca de classe (CSR) em células B infiltradas em tumores (16). O COAD destacou-se pelo predomínio de programas de centro germinativo (GC-dominant) e por TLS mais desenvolvidas, enquanto populações extrafoliculares (EF)

exibiram baixa CSR/SHM, sugerindo ativação majoritariamente germinativa no COAD. Persistia, contudo, uma lacuna: como maturação e arquitetura clonal de células B diferem entre Tumor e Adjacente no COAD.

Diante desse contexto, nosso objetivo foi caracterizar a maturação e a clonalidade de células B no COAD por meio da integração de scRNA-seq e scBCR-seq pareados filtrados de um atlas pan-câncer. Após isolar os casos de COAD e segmentar tecidos tumorais e adjacentes, investigamos subtipos de células B, diversidade clonal, SHM e CSR, buscando compreender como esses eixos refletem a ativação antígeno-dependente e o papel das TLS no microambiente tumoral.

METODOLOGIA

Desenho do estudo

Foi realizado um estudo observacional-computacional de célula única no adenocarcinoma de cólon, comparando tecidos tumorais e adjacentes, integrando identidade celular (scRNA-seq) e repertório de receptores de células B (scBCR-seq). As análises investigaram trajetórias de diferenciação de células B, diversidade e clonalidade do repertório, intensidade de hipermutação somática, ocorrência de troca de classe (CSR) e marcadores associados a estruturas linfoides terciárias.

Fonte dos dados e critérios de inclusão

Foram utilizados dados processados do atlas pan-câncer (16) disponibilizado em repositórios públicos (Zenodo DOI: 10.5281/zenodo.10709105; NGDC: PRJCA020880), contendo scRNA-seq, scBCR-seq e scATAC-seq, com pipelines documentados no GitHub. A partir desse recurso, foram extraídos os casos de COAD, segmentando tecidos tumorais e adjacentes e retendo células B com metadados consistentes. Em sua totalidade, o atlas comprehende aproximadamente 474 mil transcriptomas de células tumorais infiltrantes, incluindo cerca de 167 mil células com scBCR pareado, provenientes de 477 amostras de 269 pacientes abrangendo 20 tipos de câncer, o que oferece robustez para análises de clonalidade, SHM e CSR.

Pré-processamento do scRNA-seq (R/Seurat)

As análises foram realizadas no ambiente R (<https://www.r-project.org/>) em versão ≥4.3, utilizando o pacote Seurat em suas versões v4 e v5 para manipulação de dados, normalização, redução de dimensionalidade e visualização (17,18). A partir do objeto filtrado do atlas, apenas os casos de COAD foram extraídos, segmentando amostras tumorais e adjacentes. A qualidade das células foi avaliada com base no número de genes detectados, na contagem total de RNA e na proporção de genes mitocondriais. Após a normalização e a padronização de escala, a redução de dimensionalidade foi realizada por análise de componentes principais (PCA), seguida da construção de uma projeção UMAP específica para COAD (19), preservando integralmente as anotações celulares do atlas.

Anotação e análise de repertório BCR

As análises de repertório BCR utilizaram os campos padronizados do atlas, sem reanotação de rearranjos V(D)J. Foram empregados v_identity e mu_freq como métricas de SHM, c_call para a identificação do isótipo de cadeia pesada (IGH) e clonotipos definidos pelo estudo original. A interpretação seguiu os padrões de dados da plataforma [Immcantation](https://immcantation.readthedocs.io/en/stable/datastandards.html) (<https://immcantation.readthedocs.io/en/stable/datastandards.html>) e a lógica de anotação de repertório estabelecida no pacote Change-O (20).

Subtipos celulares e heterogeneidade

O objeto COAD_Tumor_Adjacente_QC.rds foi empregado para gerar projeções UMAP específicas de subtipos de células B previamente anotados. A composição

relativa desses subtipos foi comparada entre tumor e tecido adjacente por meio de tabelas de contingência, sendo representada em gráficos de barras empilhadas e agrupadas. Essa etapa forneceu o contexto necessário para a interpretação das análises de repertório, relacionando clonalidade, SHM e CSR à heterogeneidade celular do COAD.

Clonalidade do repertório

A frequência de clonótipos foi derivada de `clone_id`, agregando células por tecido. A distribuição clonal foi descrita por frequências absolutas, destacando os 10 clonótipos mais abundantes por condição. A diversidade do repertório foi estimada pelos índices de Shannon e Simpson. As comparações Tumor vs Adjacente foram pareadas por paciente com teste de Wilcoxon (`stats::wilcox.test`) e *p* ajustados por Benjamini–Hochberg (21).

Hipermutação Somática

A SHM foi avaliada em cadeia pesada produtiva (IGH) com `c_call`, `v_identity` e `mu_freq` válidos. Construímos boxplots de `v_identity` por isótipo × tecido (valores menores = maior SHM) e um heatmap da `mu_freq` média por paciente × isótipo, padronizado por linha e agrupado por similaridade, apresentado para as amostras tumorais.

Troca de classe (CSR)

A CSR foi inferida a partir do campo `c_call`, considerando isótipos pré-CSR (IgM/IgD) em contraste com isótipos pós-CSR (principalmente IgG e IgA). As distribuições isotípicas foram representadas em contagens absolutas e em proporções relativas dentro de cada tecido, ordenadas segundo a hierarquia biológica (IGHD,

IGHM, IGHG1–4, IGHA1–2). Isótipos pouco frequentes foram agrupados na categoria “Outros”. As diferenças globais entre tumor e adjacente foram avaliadas pelo teste do qui-quadrado (`stats::chisq.test`), considerando a distribuição total de células B por amostra e aplicando agregações quando frequências esperadas eram baixas.

RESULTADOS

Preparação, filtragem e panorama da coorte

O objeto filtrado de COAD foi carregado a partir do atlas, restringindo-se às condições tumorais (grupo “Cancer”) e adjacentes (grupo “Adjacent”). Foram inspecionados os metadados associados a cada célula, incluindo tipo de amostra (type), paciente (patient), clonótipo (clone_id), isótipos e métricas de hipermutação somática.

A coorte final compreendeu oito pacientes (COAD2, COAD3, COAD14, COAD15, COAD16, COAD20, COAD21 e COAD24), totalizando aproximadamente 19 mil células B, das quais 15.285 foram provenientes de tecido tumoral (80,5%) e 3.714 de tecido adjacente (19,5%).

A Tabela S1 apresenta a distribuição global das células B entre tumor e tecido adjacente, enquanto a Tabela S2 detalha a contribuição individual de cada paciente, evidenciando variações interindividuais tanto no número absoluto de células quanto na proporção relativa entre compartimentos teciduais.

A Tabela S2 apresenta o ranqueamento dos pacientes de acordo com o número absoluto de células B recuperadas. No compartimento tumoral, destacaram-se COAD20 (3.724 células), COAD15 (3.331 células) e COAD14 (3.058 células) como os principais contribuintes da coorte. No tecido adjacente, as maiores contribuições foram observadas em COAD16 (1.091 células), COAD15 (929 células) e COAD24 (754 células), refletindo heterogeneidade interindividual na composição celular entre os tecidos analisados.

Esses resultados confirmam que a coorte apresenta volume amostral adequado e relativamente balanceado entre as condições tumor e adjacente, permitindo comparações robustas apesar da heterogeneidade interindividual (Tabelas S1–S2).

Pré-processamento e projeção UMAP (Tumor vs. Adjacente)

Após a filtragem, foram analisadas 15.285 células B de tecido tumoral e 3.714 de tecido adjacente, distribuídas entre oito pacientes. A UMAP revelou três agrupamentos e sobreposição parcial, com regiões enriquecidas no compartimento adjacente, sugerindo a presença de subpopulações não associadas diretamente ao microambiente tumoral. Essa análise funcionou como verificação de consistência, indicando que as diferenças entre tumor e adjacente seriam mais bem exploradas nas análises subsequentes de subtipagem, clonalidade e SHM.

Heterogeneidade de subtipos de células B

A projeção UMAP específica para COAD (Figura 1) revelou subtipos de células B bem delimitados. Plasmócitos (B.15) e plasmablastos (B.14) foram observados segregados do contínuo principal, enquanto três regiões distintas corresponderam a fenótipos compatíveis com centro germinativo (B.11–B.13). Subtipos naïve e de memória/ativação organizaram-se em um gradiente de diferenciação, que se estendeu de células naïve até plasmablastos e plasmócitos, passando por estados intermediários de memória/ativação e centro germinativo.

Na Figura 2, o Tumor mostrou enriquecimento de subpopulações ativadas/memória pós-CSR: ITGB1 /SwBm 27,9% vs 10,3%, NR4A2 /ACB2 19,5% vs 13,5%, CCR7 /ACB3 5,0% vs 2,5% e AtM/DUSP4 4,9% vs 2,0% (Tumor vs Adjacente). O Adjacente apresentou mais células naïve (24,9% vs 16,3%) e plasmócitos maduros (22,3% vs 3,7%). Fenótipos compatíveis com GC ocorreram em ambos, com leve predominância no Adjacente (LZ 5,9% vs 3,6%; Cycling 4,7% vs 3,3%). Em conjunto, reforça-se a heterogeneidade funcional das células B no COAD, com

predomínio de subtipos pós-CSR no Tumor e manutenção de compartimentos naïve/plasmocítico no Adjacente.

Os resultados indicam que o microambiente tumoral do COAD é dominado por células B ativadas e de memória em estágio pós-CSR, enquanto o tecido adjacente concentra um repertório mais basal, caracterizado por maior proporção de células naïve e plasmócitos. Esse contraste é compatível com ativação antígeno-dependente ocorrendo no interior do tumor e sugere o envolvimento de TLS na indução de respostas de memória e de troca de classe no TME.

Clonalidade BCR: Top 10 por tecido

Evidenciou-se expansão policlonal, sem clonótipo dominante. Clones compartilhados entre Tumor e Adjacente (p.ex., 275281_217056, 290644_217057) sugerem seleção antígeno-dependente comum; já clones restritos a cada tecido (p.ex., 216222_217216 no Adjacente; 10727_207661/122490_207697 no Tumor) indicam seleção local adicional no TME (**Figura S1**).

Diversidade clonal por paciente

A análise por paciente (**Figura S2**) indicou Shannon ligeiramente maior no Tumor (~7–8) vs Adjacente (~6–6,5), enquanto Simpson manteve-se alto e semelhante (~0,997–0,999), sem clonótipos hiperdominantes. O Wilcoxon pareado não foi significativo ($p_{ajust.}=0,354$). Esse resultado pode refletir n reduzido, desequilíbrio de doadores (Tumor=8; Adjacente=5) e variabilidade interindividual.

Hipermutação Somática (SHM)

Aplicado apenas BCR produtivo de cadeia pesada (locus = IGH), mantendo células com c_call, v_identity e mu_freq anotados. As análises foram estratificadas por tecido (Tumor vs Adjacent) e por isotipo.

A análise da identidade V por isotipo evidenciou padrões distintos entre tumor e tecido adjacente. Como mostrado na Figura 4A, isotipos naïve (IGHM e IGHD) mantiveram identidade próxima a 1 em ambos os compartimentos, compatível com mínima ocorrência de hipermutação somática. Em contraste, os isotipos pós-CSR (IgA e IgG) apresentaram valores em torno de 0,90–0,95, refletindo maior acúmulo de mutações. No tumor, especialmente em IgG2/3, observou-se maior dispersão e presença de valores mínimos mais baixos, sugerindo a existência de subgrupos mais mutados sob pressão antigênica no microambiente tumoral. Já no adjacente, IGHA1/2 mostraram identidade discretamente menor, indicando SHM ligeiramente mais intensa nesse compartimento.

A distribuição isotípica corrobora os achados de SHM. Em valores absolutos (Figura 4B), IGHM foi o mais representado em ambos os tecidos (perfil naïve). No Tumor, houve aumento de IgA (IGHA1/2) e discreto de IgG (IGHG1/2), com IgM em frequências semelhantes, padrão que excede o efeito do maior N tumoral e sugere expansão pós-CSR específica do TME. Isotipos raros (IGHD, IGHG3/IGHG4) permaneceram residuais.

Considerando as proporções relativas (Figura 4C), o Adjacente predominou em IgM/IgD, indicando repertório basal/naïve e menor atividade de CSR (menor pressão antigênica local). O Tumor mostrou enriquecimento pós-CSR, com aumento marcante de IgA (sobretudo IGHA1) e tendência a IgG, achado esperado no cólon, onde IgA é

central à imunidade mucosa. Esse perfil sugere ativação antígeno-dependente sustentada, possivelmente organizada em TLS intratumorais. O acúmulo de IgA no Tumor pode refletir tanto o microambiente mucoso quanto remodelamento ativo do repertório B, compatível com a maior SHM e diversificação funcional observadas.

Para aprofundar a análise da distribuição isotípica, foi construída uma tabela de contingência com as contagens absolutas de isótipos em tecido tumoral e adjacente (**Tabela S3**). Esses dados permitem avaliar estatisticamente se as diferenças observadas visualmente nos painéis anteriores refletem uma redistribuição significativa do repertório entre os compartimentos.

DISCUSSÃO

Este estudo integra dados de scRNA-seq e scBCR-seq para caracterizar a maturação e a clonalidade de linfócitos B no adenocarcinoma de cólon, comparando tecido tumoral e adjacente. A coorte (~19 mil células B de oito pacientes) foi adequada para comparações, embora limitada pela ausência de tecido adjacente em parte dos casos. A projeção UMAP específica para COAD revelou três agrupamentos principais e um continuum de diferenciação, no qual células naïve se organizam em trajetória progressiva até estados de memória/ativação, fenótipos compatíveis com centro germinativo e, por fim, plasmoblastos e plasmócitos. O padrão, coerente com atlas de célula única, indica que o TME colônico sustenta ativação antígeno-dependente mediada por TLS.

A comparação Tumor vs Adjacente reforçou essa leitura. O Tumor mostrou enriquecimento de subpopulações ativadas/memória pós-CSR (*ITGB1* /*SwBm*, *NR4A2* /*ACB2*, *CCR7* /*ACB3*, *AtM/DUSP4*), enquanto o Adjacente concentrou mais células naïve e plasmócitos. Estados compatíveis com GC apareceram em ambos, com leve predominância no adjacente, alinhada a TLS peritumoriais no CCR. Já a predominância intratumoral de B ativadas/memória pós-CSR é consistente com TLS maduros sustentando ciclos de SHM/CSR e geração de plasmócitos no próprio TME. Evidências recentes em CCR e outros tumores sólidos indicam que TLS organizados com reação de GC associam-se a maior infiltração adaptativa, melhor prognóstico e maior resposta à imunoterapia, dando suporte biológico aos achados.

A clonalidade corroborou ativação antígeno-dependente. Identificamos clonótipos compartilhados entre Tumor e Adjacente (tráfego/respostas comuns) e clones

exclusivos por nicho (seleção local). O repertório permaneceu policlonal, sem hiperdominância (Simpson ≈ 1), e Shannon foi discretamente maior no Tumor, porém não significativo após correção múltipla. Em conjunto, indica maior diversidade intratumoral de clones secretores, sem oligoclonalidade extrema, em acordo com atlas pan-câncer.

As análises de SHM apontam TLS como nichos de centro germinativo intratumorais. Isótipos pós-CSR (IgG2, IgA1/2) exibiram SHM elevada, enquanto IgM/IgD mantiveram alta identidade V (estado naïve). No Tumor, IgG mostrou maior heterogeneidade, com subgrupos altamente mutados, indicativo de pressão抗igenética local e ciclos repetidos de seleção. No Adjacente, IgA apresentou SHM discretamente maior, coerente com o ambiente mucoso do cólon. Em conjunto, os achados sugerem TLS intra- e peritumoriais sustentando trajetórias paralelas de maturação B, com pesos distintos de IgA e IgG.

A distribuição isotípica reforça o quadro. Em valores absolutos, IGHM foi o mais representado em ambos os tecidos; no Tumor houve aumento de IgA e discreto de IgG. Nas proporções relativas, o Adjacente predominou em IgM/IgD (perfil basal), enquanto o Tumor ficou enriquecido em pós-CSR, sobretudo IgA (IGHA1), com tendência a IgG. A redistribuição foi significativa ($\chi^2=198,4$; gl=6; $p<0,0001$), em linha com estudos que apontam IgA em regiões mucosas peri-tumorais e IgG no interior tumoral. A predominância pós-CSR no TME sugere TLS maduros ativos (GC in situ), sustentando SHM/CSR e gerando plasmócitos IgG /IgA produtores de anticorpos intratumoriais.

Os resultados indicam que o TME do COAD não apenas recruta células B, mas promove sua maturação para perfis pós-CSR, sustentada por SHM/CSR, configurando circuitos antígeno-dependentes análogos aos de centros germinativos. Isso explica a associação de TLS maduros a melhor resposta imune adaptativa no câncer colorretal. Limitações incluem o uso de anotações pré-existentes restritas à IGH, a ausência de cadeias leves, n reduzido e desequilíbrio Tumor/Adjacente. Perspectivas incluem ampliar a coorte, integrar cadeias leves, associar transcriptômica espacial e validar histologicamente TLS para consolidar a compreensão da dinâmica de células B no TME do cólon.

CONCLUSÃO

Este estudo utilizou dados de scRNA-seq e scBCR-seq filtrados de um atlas multicâncer para modelar computacionalmente a maturação e a clonalidade de linfócitos B no adenocarcinoma de cólon, comparando tecido tumoral e adjacente. A integração de projeções específicas para COAD com análises do repertório BCR, abrangendo diversidade clonal, hipermutação somática e troca de classe, permitiu explorar de forma sistemática a relação entre identidade celular e uso de isótipos no contexto da ativação antígeno-dependente.

Observou-se um contínuo de diferenciação B: o Tumor apresentou enriquecimento de subpopulações ativadas/memória pós-CSR e maior diversidade por Shannon, mantendo policlonalidade sem hiperdominância (Simpson alto). O Adjacente reteve mais células naïve e plasmócitos. A SHM foi mais intensa em IgG/IgA, enquanto IgM/IgD preservaram alta identidade V. A CSR mostrou aumento intratumoral de isótipos pós-CSR, com destaque para IgA e tendência a IgG, compatível com ativação antígeno-dependente sustentada por TLS.

Limitações incluem o n pequeno, desequilíbrio entre Tumor e Adjacente, dependência de dados pré-processados e foco apenas em cadeias pesadas (IGH), sem integração de cadeias leves nem validação espacial/funcional. Perspectivas: ampliar a coorte pareada, integrar transcriptômica espacial e marcadores de TLS/CG, incluir cadeias leves na reconstrução do BCR e realizar validações funcionais de anticorpos de clones expandidos. Esses avanços devem refinar os eixos de clonalidade, SHM e CSR no TME do cólon e consolidar o papel das células B como mediadoras centrais da imunidade antitumoral.

REFERÊNCIAS

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021;71(3):209–49. doi:10.3322/caac.21660.
2. Keum NN, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat Rev Gastroenterol Hepatol.* 2019;16(12):713–32. doi.org/10.1038/s41575-019-0189-8.
3. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *Lancet.* 2019;394(10207):1467–80. doi:10.1016/S0140-6736(19)32319-0.
4. Fridman WH, Pagès F, Sautès-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer.* 2012;12(4):298–306. doi:10.1038/nrc3245.
5. Pagès F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet.* 2018;391(10135):2128–39. doi:10.1016/S0140-6736(18)30789-X.
6. Laumont CM, Nelson BH. B cells in the tumor microenvironment: Multi-faceted organizers, regulators, and effectors of anti-tumor immunity. *Cancer Cell.* 2023;41(3):466–89. doi:10.1016/j.ccr.2023.02.017.
7. Esparcia-Pinedo L, Romero-Laorden N, Alfranca A. Tertiary lymphoid structures and B lymphocytes: a promising therapeutic strategy to fight cancer. *Front Immunol.* 2023;14:1231315. doi:10.3389/fimmu.2023.1231315.

8. Cui X, Gu X, Li D, Wu P, Sun N, Zhang C, et al. Tertiary lymphoid structures as a biomarker in immunotherapy and beyond: Advancing towards clinical application. *Cancer Lett.* 2025;613:217491. doi:10.1016/j.canlet.2025.217491.
9. Wang Q, Zhong W, Shen X, Hao Z, Wan M, Yang X, et al. Tertiary lymphoid structures predict survival and response to neoadjuvant therapy in locally advanced rectal cancer. *npj Precis Oncol.* 2024;8:61. doi:10.1038/s41698-024-00533-w.
10. Yang SY, Long J, Huang MX, Luo PY, Bian ZH, Xu YF, et al. Characterization of organ-specific regulatory B cells using single-cell RNA sequencing. *Front Immunol.* 2021;12:711980. doi:10.3389/fimmu.2021.711980.
11. Meylan M, Petitprez F, Becht E, Bougoüin A, Pupier G, Calvez A, et al. Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. *Immunity.* 2022;55(3):527–541.e5. doi:10.1016/j.immuni.2022.02.001.
12. King HW, Orban N, Riches JC, Clear AJ, Warnes G, Teichmann SA, et al. Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. *Sci Immunol.* 2021;6(56):eabe6291. doi:10.1126/sciimmunol.abe6291.
13. Mazor RD, Nathan N, Gilboa A, Stoler-Barak L, Moss L, Solomonov I, et al. Tumor-reactive antibodies evolve from non-binding and autoreactive precursors. *Cell.* 2022;185(7):1208–1222.e21. doi:10.1016/j.cell.2022.02.012.
14. Liu Y, Ye SY, He S, Chi DM, Wang XZ, Wen YF, et al. Single-cell and spatial transcriptome analyses reveal tertiary lymphoid structures linked to tumour

- progression and immunotherapy response in nasopharyngeal carcinoma. *Nat Commun.* 2024;15:7713. doi:10.1038/s41467-024-52153-4.
15. Yang B, Zhang Z, Chen X, Wang XY, Qin S, Du L, et al. An Asia-specific variant of human IgG1 represses colorectal tumorigenesis by shaping the tumor microenvironment. *J Clin Invest.* 2022;132(6):e153454. doi:10.1172/JCI153454.
 16. Ma J, Wu Y, Ma L, Yang X, Zhang T, Song G, et al. A blueprint for tumor-infiltrating B cells across human cancers. *Science.* 2024;384(6695):eadj4857. doi:10.1126/science.adj4857.
 17. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive integration of single-cell data. *Cell.* 2019;177(7):1888–1902.e21. doi:10.1016/j.cell.2019.05.031.
 18. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell.* 2021;184(13):3573–3587.e29. doi:10.1016/j.cell.2021.04.048.
 19. McInnes L, Healy J, Melville J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv.* 2018;arXiv:1802.03426. doi:10.48550/arXiv.1802.03426.
 20. Gupta NT, Vander Heiden JA, Uduman M, Gadala-Maria D, Yaari G, Kleinstein SH. Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics.* 2015;31(20):3356–8. doi:10.1093/bioinformatics/btv359.
 21. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Methodol.* 1995;57(1):289–300. doi:10.1111/j.2517-6161.1995.tb02031.x.

FIGURAS

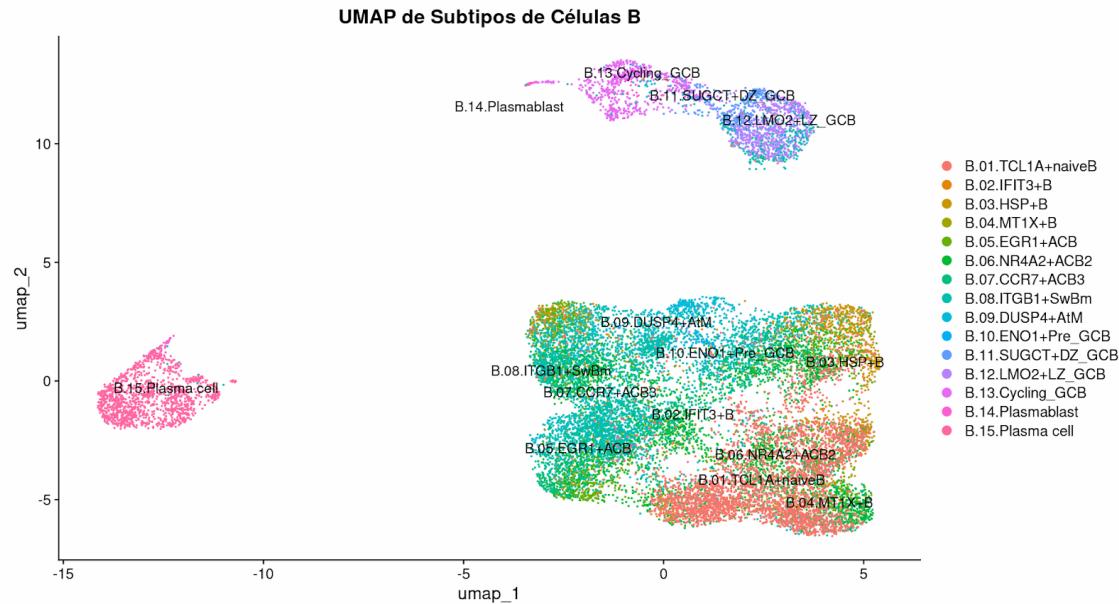


Figura 1. Projeção UMAP das células B do COAD (n=8 pacientes; 15.285 Tumor, 3.714 Adjacente), colorida por subtipo (objeto filtrado do atlas multicâncer). Pré-processamento no Seurat (NormalizeData, PCA, viz. UMAP específica de COAD).

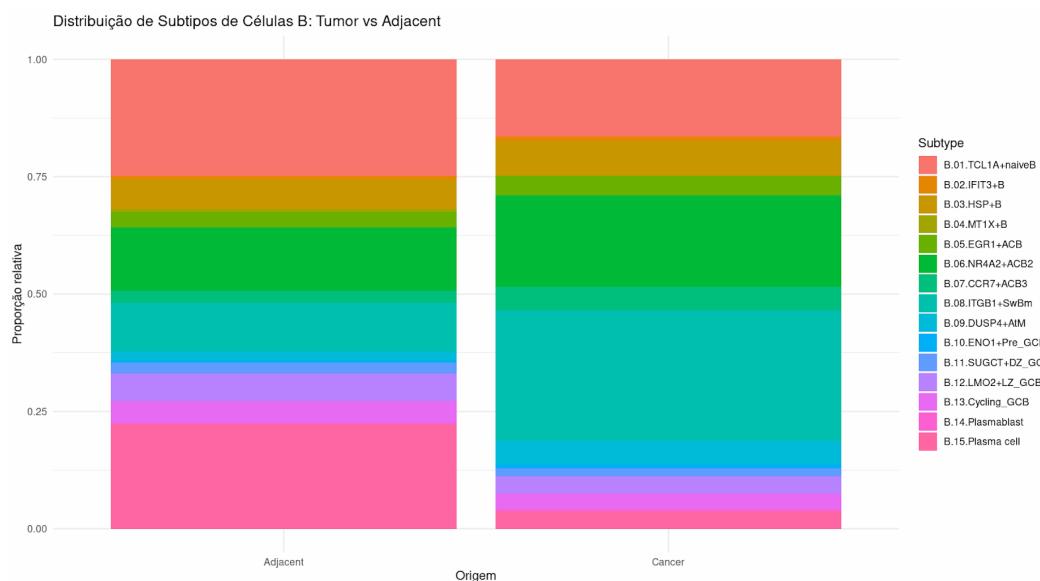


Figura 2. Proporção relativa de subtipos de células B no adenocarcinoma de cólon, (≈ 19.000 células válidas), Tumor vs Adjacente. O Tumor está enriquecido em subpopulações ativadas/memória pós-CSR (p.ex., ITGB1 /SwBm, NR4A2 /ACB2),

enquanto o Adjacente concentra mais células naïve e plasmócitos. Padrão compatível com ativação antígeno-dependente e organização em TLS.

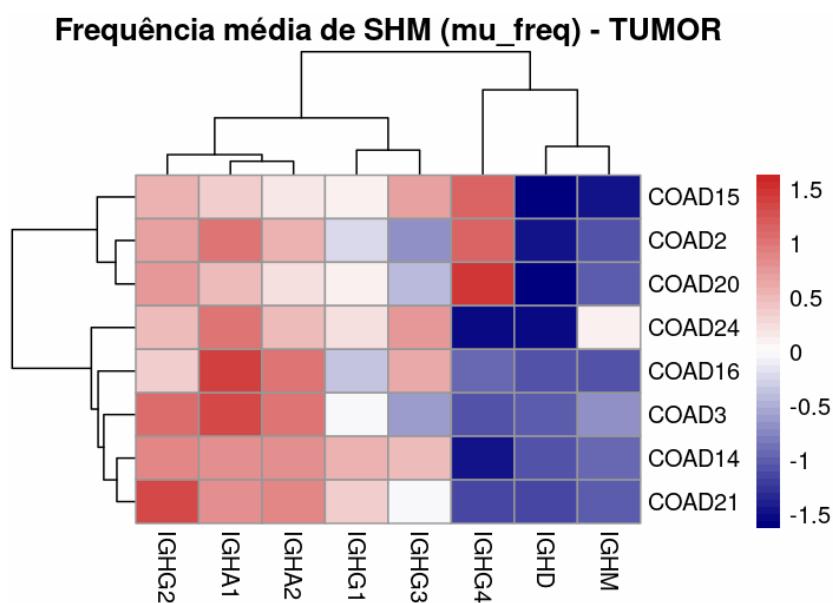
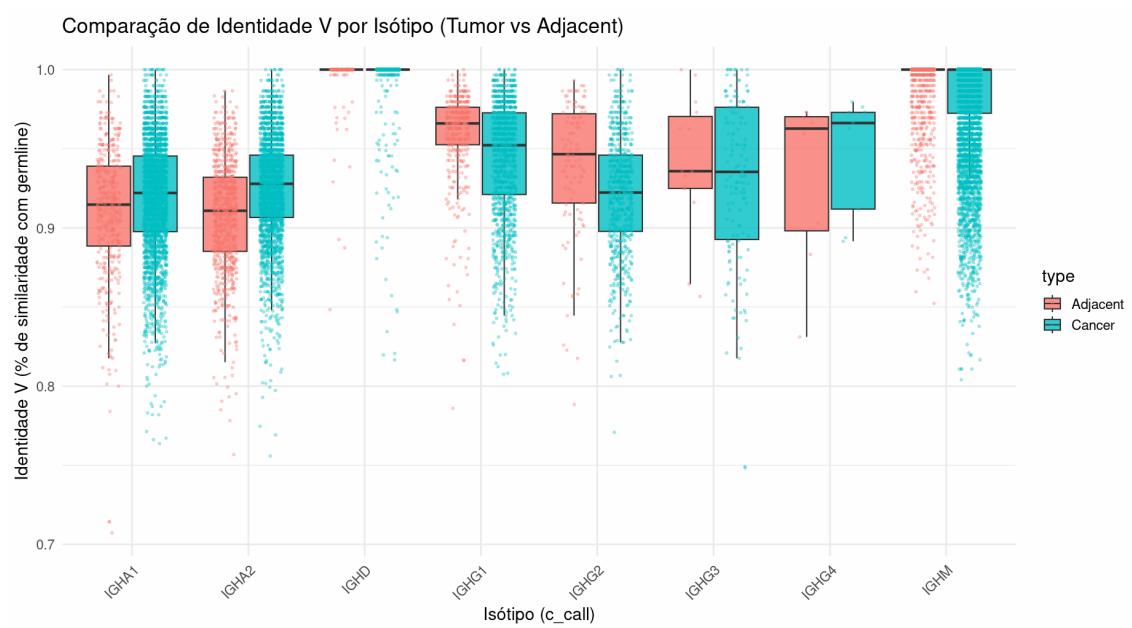


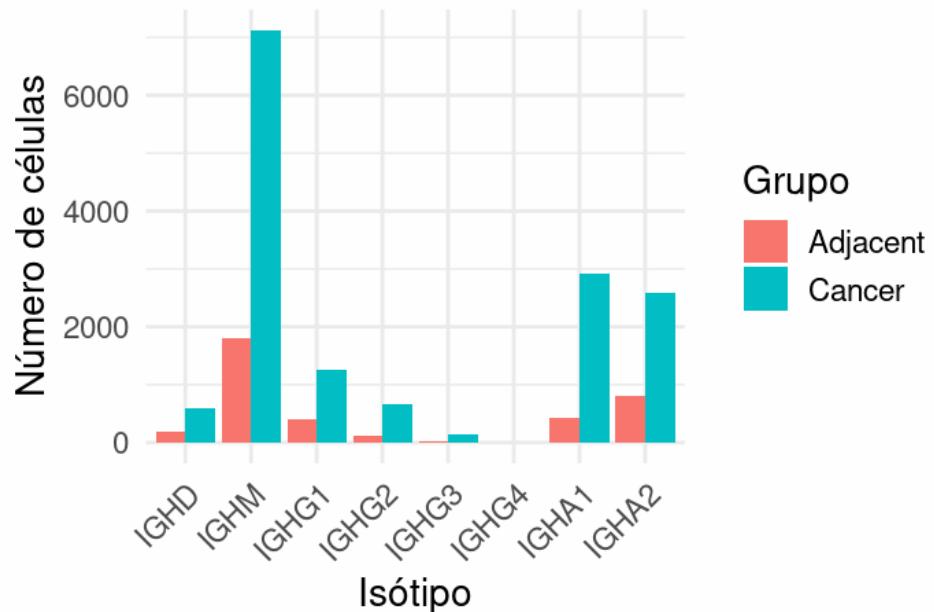
Figura 3. Heatmap da média de mu_freq (SHM em IGH) nas amostras tumorais, por paciente × isótipo. Valores padronizados por linha (z-score por isótipo) e clusterizados por hierarquia (Ward.D2). Isótipos pós-CSR (IGHG2, IGHA1, IGHG2) exibem SHM mais alta (vermelho), enquanto IGHM/IGHD mostram baixa mutação (azul), compatível com fenótipos naïve.

A) Comparação de Identidade V por Isótipo (Tumor vs Adjacent)



B) Frequência Absoluta de Isótipos

Frequência absoluta de isótipos



C) Frequência Relativa de Isótipos

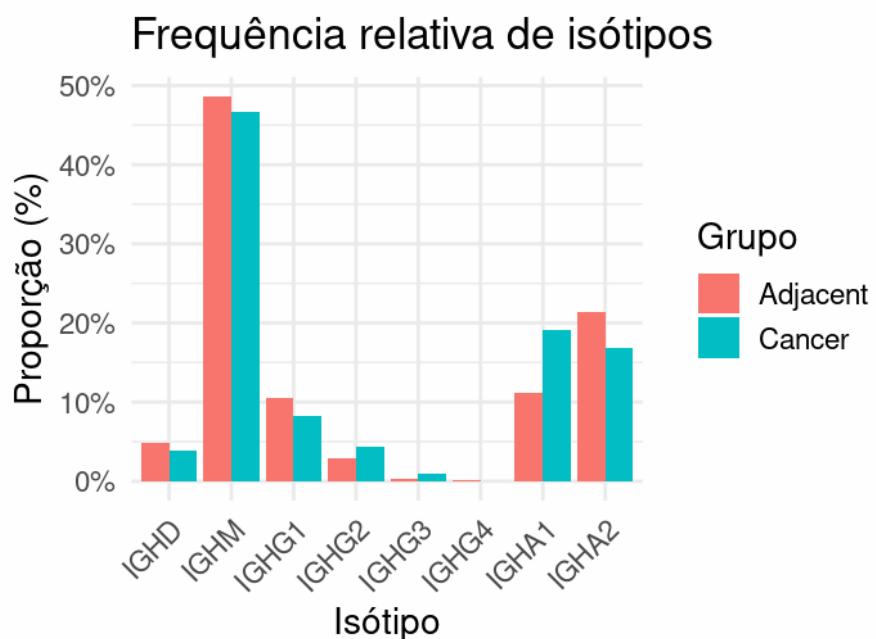


Figura 4. Identidade V e composição isotípica de células B no COAD (Tumor vs Adjacente). (A) v_identity por isótipo em IGH produtivo. $\text{IGHM}/\text{IGHD} \approx 1$ (SHM mínima); $\text{IgA}/\text{IgG} \approx 0,90-0,95$ (SHM maior). No Tumor, heterogeneidade mais alta, com maior dispersão em IGHG2/3 , sugerindo subconjuntos mais mutados sob pressão antigênica. (B) Frequência absoluta: IGHM é o mais abundante em ambos; o Tumor apresenta aumento de IgA (IGHA1/2) e discreto de IgG (IGHG1/2), indicando expansão pós-CSR além do efeito do maior N tumoral. (C) Frequência relativa: Adjacente com perfil basal (IgM/IgD); Tumor enriquecido em pós-CSR, sobretudo IgA (IGHA1) e tendência a IgG. Padrão compatível com ativação antígeno-dependente em TLS e com a relevância de IgA na mucosa do cólon.