

Uso do K-médias para agrupamento de células malignas e benignas no câncer gástrico

GIULIA LEONEL, JÚLIA FERNANDES, LARA RAMPIM,
LUCAS MÜLLER, NICOLLI ROSA

Câncer gástrico

- Também chamado de câncer de estômago
- Quinta malignidade mais comum
- Quarta principal causa de morte relacionada ao câncer ao redor do mundo
- Desafios: aplicações clínicas do tratamento cirúrgico e da quimioterapia



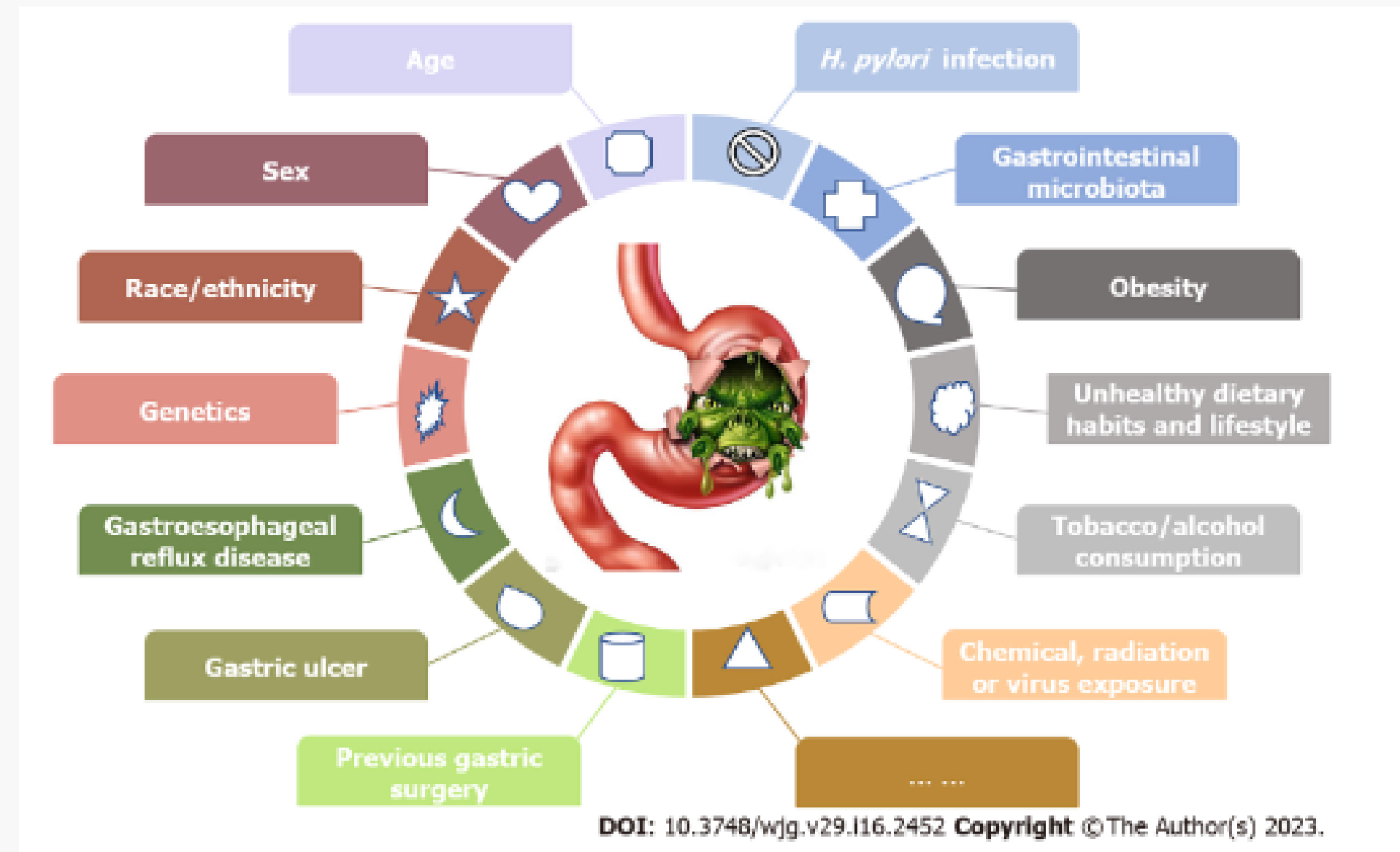
Logo, a busca por tratamentos mais eficazes e seguros para o paciente é essencial

I INTRODUÇÃO

Câncer gástrico

Fatores de risco

- Genéticos
- Epigenéticos
- Ambientais



I INTRODUÇÃO

npj | precision oncology

www.nature.com/npjprecisiononcology

ARTICLE OPEN



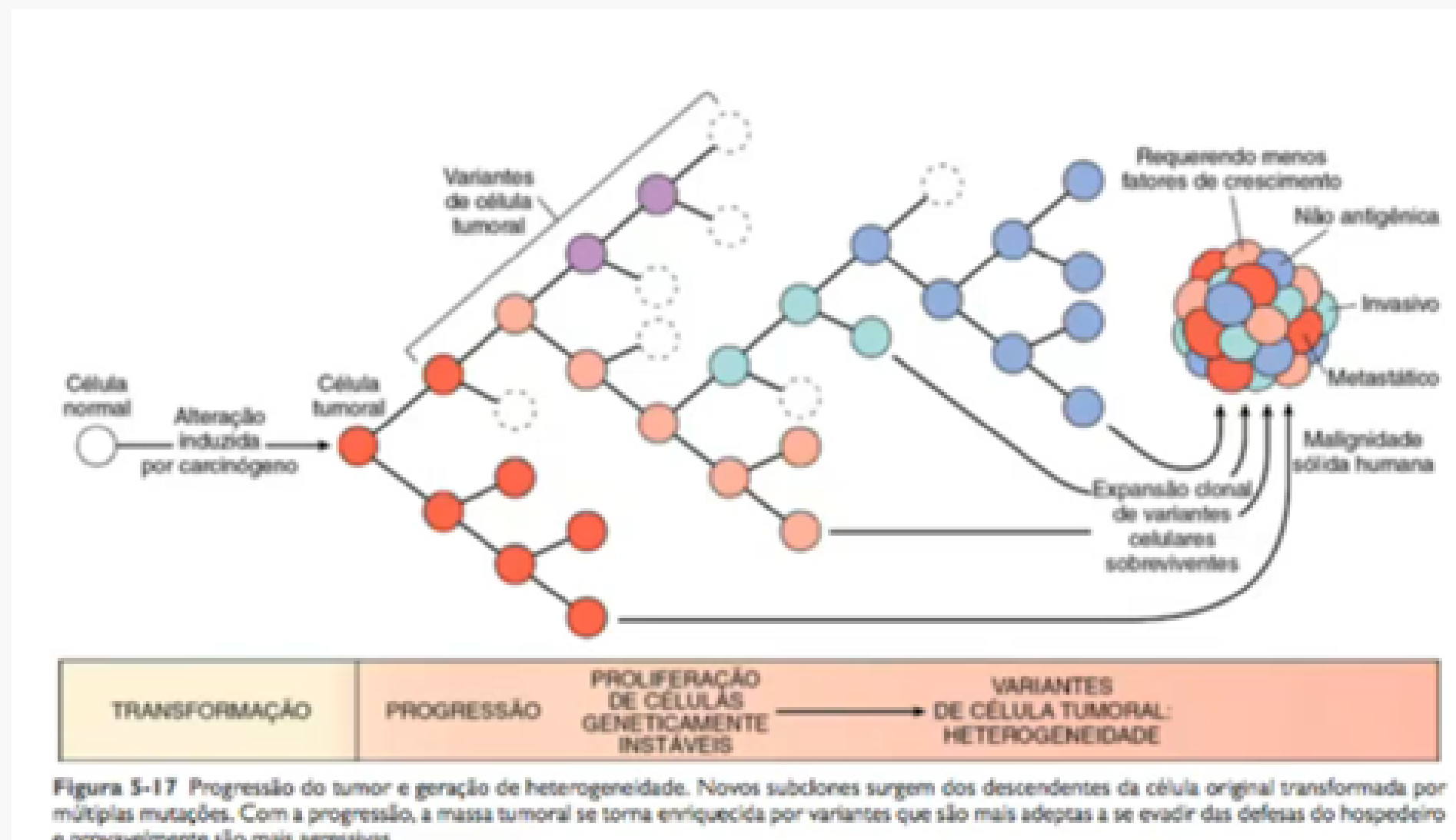
The regulatory role of cancer stem cell marker gene CXCR4 in the growth and metastasis of gastric cancer

Hongying Zhao¹✉, Rongke Jiang¹, Chunmei Zhang², Zhijing Feng² and Xue Wang¹

II FUNDAMENTAÇÃO TEÓRICA

As células-tronco cancerosas (CSCs)

- Células dentro de um tumor que possui a capacidade de se autorrenovar e de gerar as linhagens heterogêneas de células cancerosas que compõem o tumor



Impulsionadoras principais do crescimento e metástase do CG

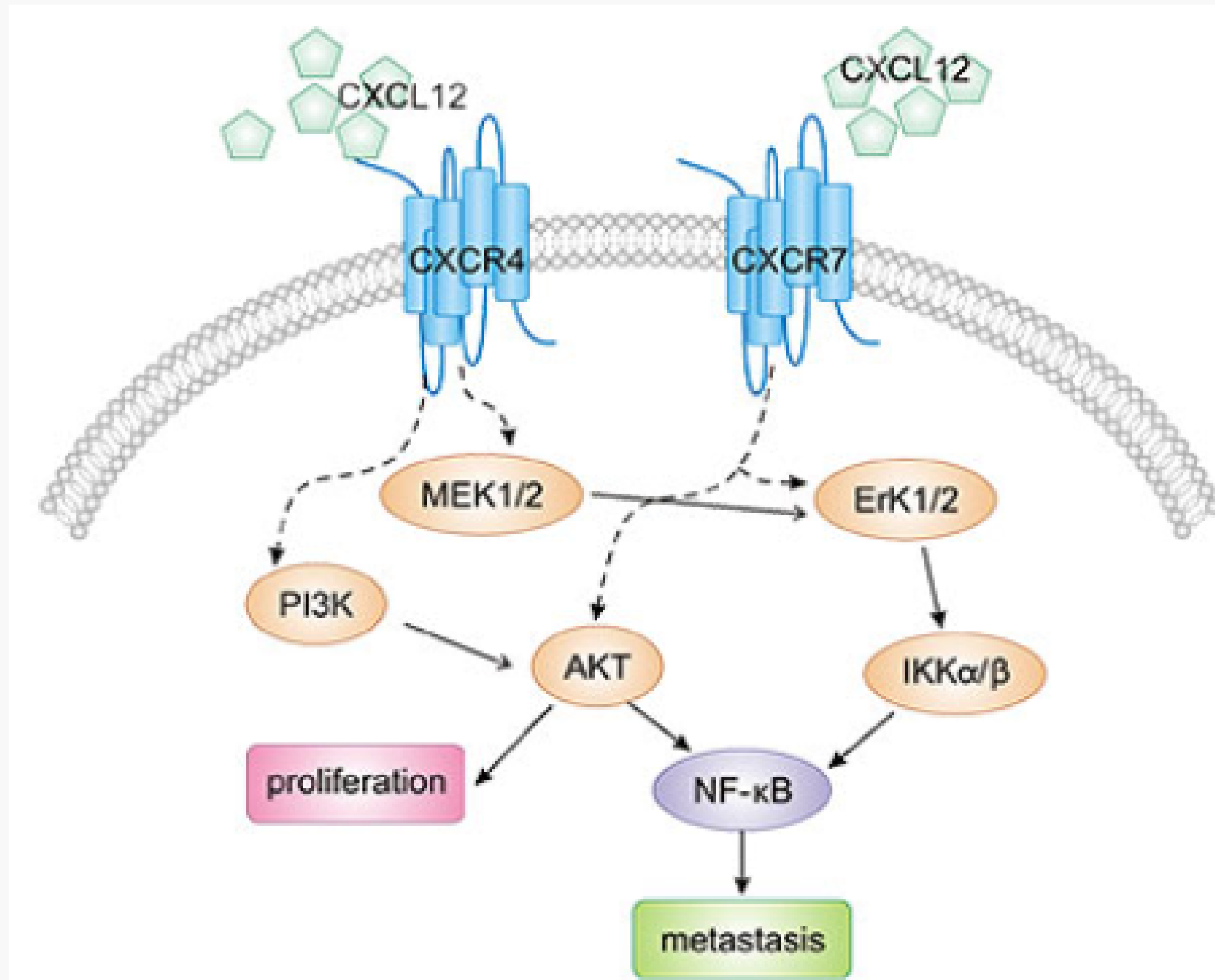
- Alvo de tratamentos
- Diagnóstico
- Prognóstico

II FUNDAMENTAÇÃO TEÓRICA

O gene CXCR4

- Codifica para uma proteína chamada receptor de quimiocina CXCR4, que é um receptor acoplado à proteína G
- O CXCR4 é o receptor específico para o CXCL12
- **Funções:** direcionamento de células do sistema imunológico, desenvolvimento de órgãos na embriogênese
- **Metástase do câncer:** Ele ajuda células tumorais a migrar para outros tecidos
 1. CXCL12 se liga ao CXCR4
 2. Ativação de vias de sinalização intracelular
 3. biológicos, incluindo o crescimento celular, a sobrevivência, a invasão e a migração

II FUNDAMENTAÇÃO TEÓRICA



II FUNDAMENTAÇÃO TEÓRICA

Tratamentos e terapias

- Inibição da expressão de CXCR4 pode suprimir fenótipos malignos das células-tronco cancerígenas -> atividade antitumoral
- Terapias que visam o eixo CXCR4-CXCL12 mostram a repressão do crescimento metastático

Diagnóstico e prognóstico

- O gene CXCR4 tem emergido como um importante marcador molecular relacionado à metástase e a uma menor sobrevida global

III DESCRIÇÃO DA ABORDAGEM COMPUTACIONAL

Clustering

- Medida de Distância
- Critério de Clustering



Buscar entre todas as possíveis partições do conjunto de dados para encontrar aquela que maximize o critério



K-médias

1. Escolher $K \rightarrow$ Centros
2. Cada centro é substituído pelas médias dos seus clusters
3. Novos centros \rightarrow Novos clusters \rightarrow Novas médias
4. O resultado se dá quando converge com o melhor critério de agrupamento

Extração

Base de Dados GEO

- Repositório abrangente de sequências de RNA de célula única (scRNA-seq)
- Utilizado para extrair 7 amostras de RNA de célula única de tumor gástrico:
 - 3 de tumor in situ (PT1, PT2 e PT3)
 - 2 de tumor metastático em linfonodo (LN1 e LN2)
 - 2 de tumor metastático em fígado (Li1 e Li2)

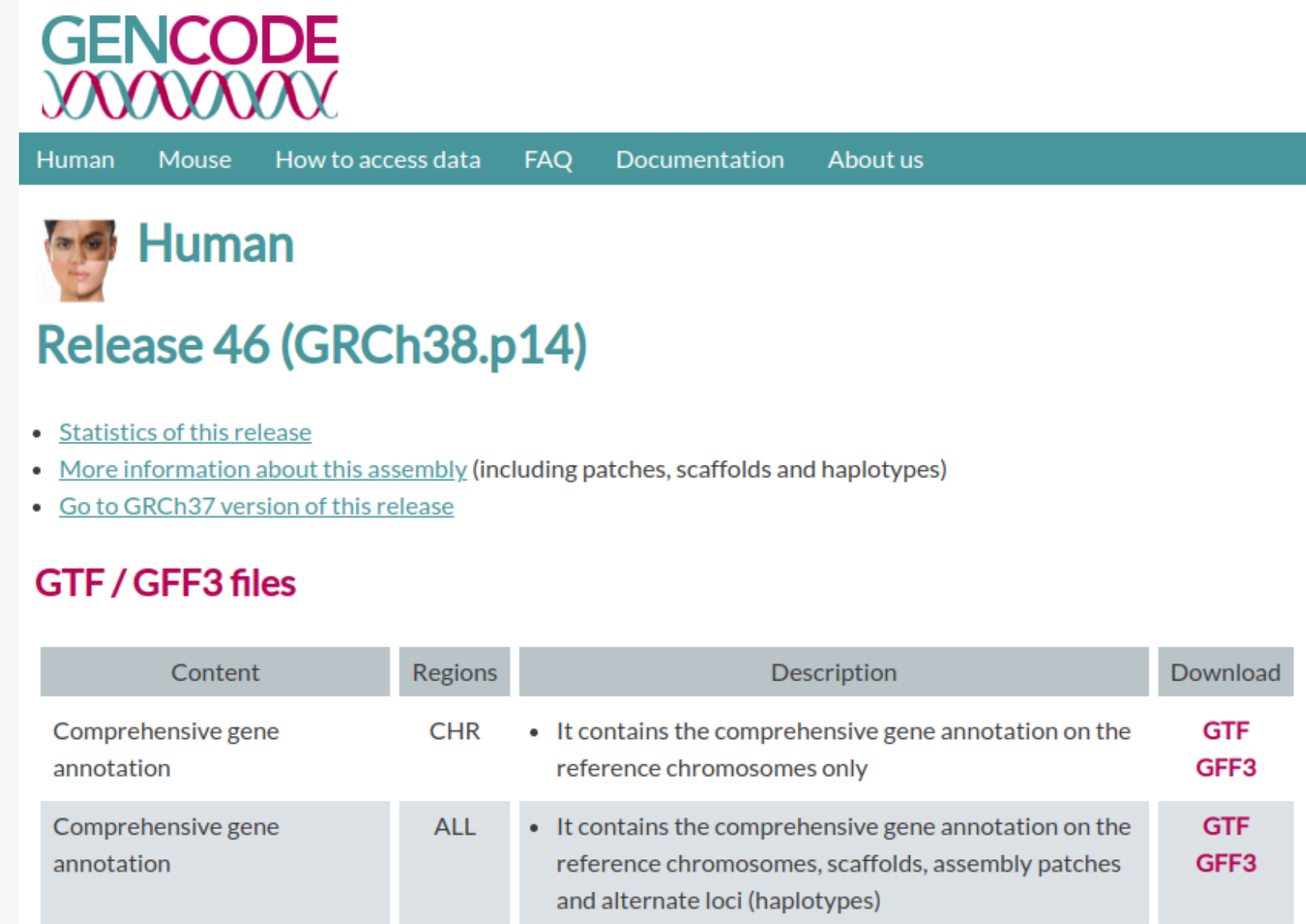
The screenshot shows the NCBI GEO Accession Display page for GSE163558. The page header includes the NCBI logo and the GEO logo (Gene Expression Omnibus). The navigation bar contains links for HOME, SEARCH, SITE MAP, GEO Publications, FAQ, MIAME, and Email GEO. The main content area displays the accession number GSE163558 and a search bar. Below the search bar, the series information is shown, including the title, organism, experiment type, summary, overall design, and contributor(s).

Series GSE163558	
Status	Public on Feb 22, 2022
Title	Transcriptional heterogeneity in primary and metastatic gastric cancer revealed using single-cell RNA sequencing
Organism	Homo sapiens
Experiment type	Expression profiling by high throughput sequencing
Summary	Deciphering intra- and inter-tumoral heterogeneity is essential for understanding gastric cancer (GC) biology and identifying effective therapeutic targets. We used single-cell RNA sequencing (scRNA-seq) to reveal the transcriptional heterogeneity of diverse cells within primary and metastatic GC and their roles in disease progression.
Overall design	We performed unbiased scRNA-seq analysis of 42,968 cells from 10 fresh human tissue samples from six patients. Primary tumor and adjacent non-tumoral samples and six metastases from different organs or tissues (liver, peritoneum, ovary, lymph node) were evaluated. Validation experiments were performed using histological assays and bulk transcriptomic datasets.
Contributor(s)	Jiang H, Yu D, Yang P
Citation(s)	Jiang H, Yu D, Yang P, Guo R et al. Revealing the transcriptional heterogeneity of organ-specific metastasis in human gastric cancer using single-cell RNA Sequencing. <i>Clin Transl Med</i> 2022 Feb;12(2):e730. PMID: 35184420

IV FERRAMENTAS


Base de Dados GENCODE

- Repositório abrangente que mapeia e identifica todas as características genéticas no genoma humano
- Utilizado para classificar as células das diferentes amostras em diferentes tipos celulares (NK, T, B, Epitelial, Stromal e Myeloid)



GENCODE

Human Mouse How to access data FAQ Documentation About us

 **Human**

Release 46 (GRCh38.p14)

- [Statistics of this release](#)
- [More information about this assembly](#) (including patches, scaffolds and haplotypes)
- [Go to GRCh37 version of this release](#)

GTF / GFF3 files

Content	Regions	Description	Download
Comprehensive gene annotation	CHR	• It contains the comprehensive gene annotation on the reference chromosomes only	GTF GFF3
Comprehensive gene annotation	ALL	• It contains the comprehensive gene annotation on the reference chromosomes, scaffolds, assembly patches and alternate loci (haplotypes)	GTF GFF3

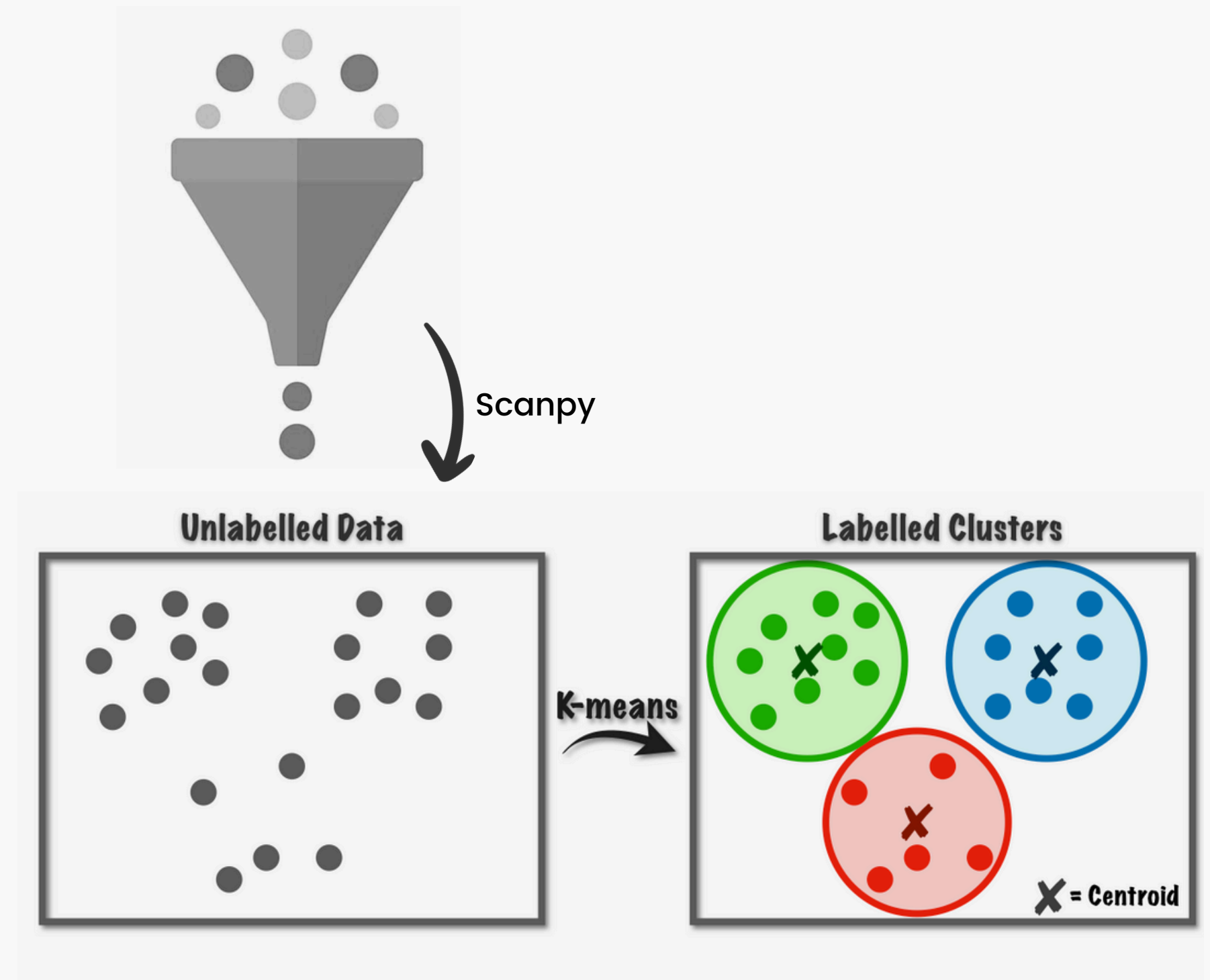
Tratamento

Scanpy

- Utilizado para análise de dados de RNA de célula única, incluindo normalização, filtragem, e redução de dimensionalidade

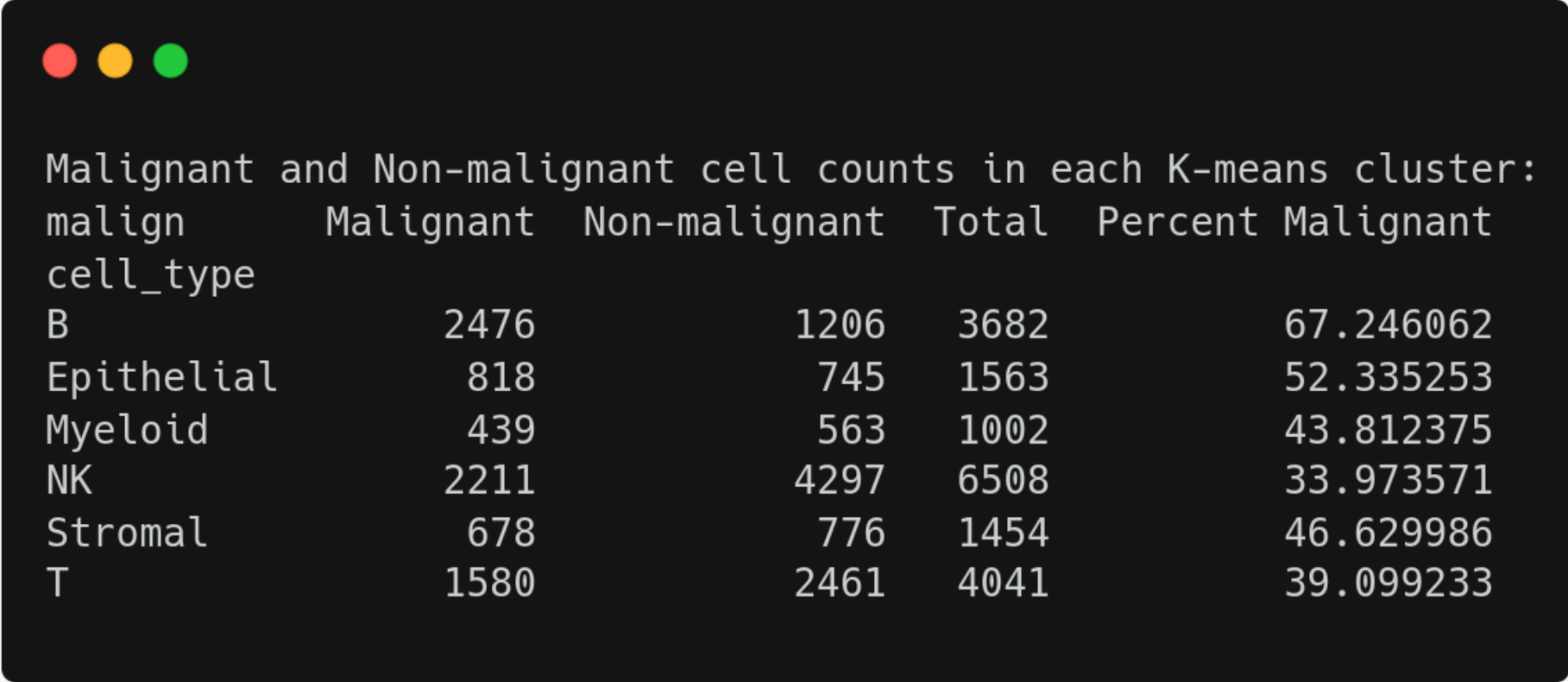
K-means

- Algoritmo foi aplicado para a clusterização dos dados de expressão gênica normalizados



Infercnvpy

- Infere cópias anormais em amostragem de células únicas com base no algoritmo CNV (Copy Number Variation)
- Utilizado para análise de variação do número de cópias anormais para cada tipo celular, identificando possíveis células malignas

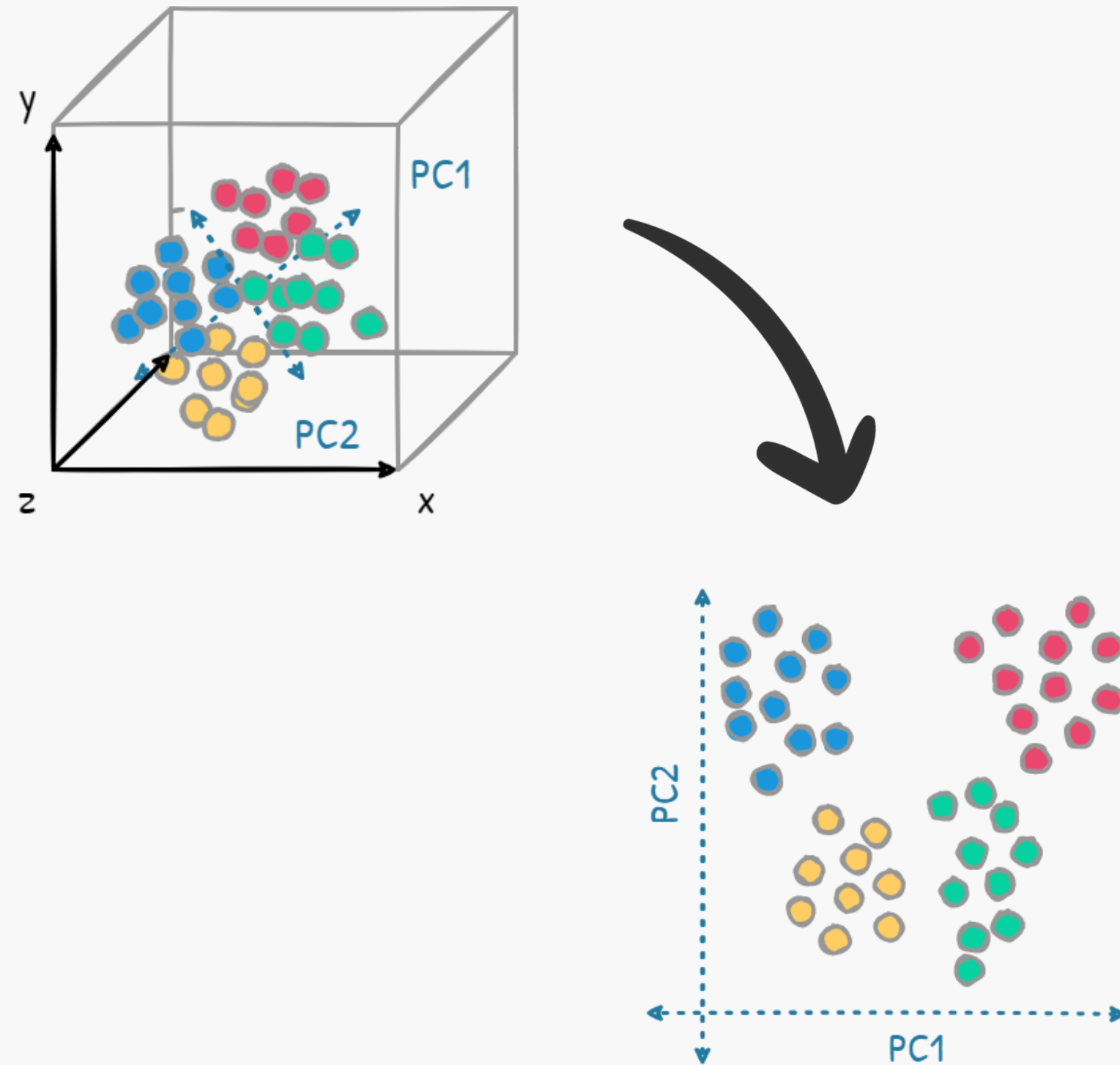


```
Malignant and Non-malignant cell counts in each K-means cluster:
malign      Malignant  Non-malignant  Total  Percent Malignant
cell_type
B            2476        1206          3682        67.246062
Epithelial   818          745           1563        52.335253
Myeloid      439          563           1002        43.812375
NK           2211        4297          6508        33.973571
Stromal      678          776           1454        46.629986
T            1580        2461          4041        39.099233
```

V RESULTADOS

UMAP (Uniform Manifold Approximation and Projection)

- Método de visualização que permite uma interpretação intuitiva da estrutura dos dados de RNA de célula única.
- Utilizado para todas as renderizações gráficas neste projeto.



VI CONCLUSÕES

- Aplicação bem sucedida do K-médias
- Clusters bem divididos
- Boa correlação com o artigo de base

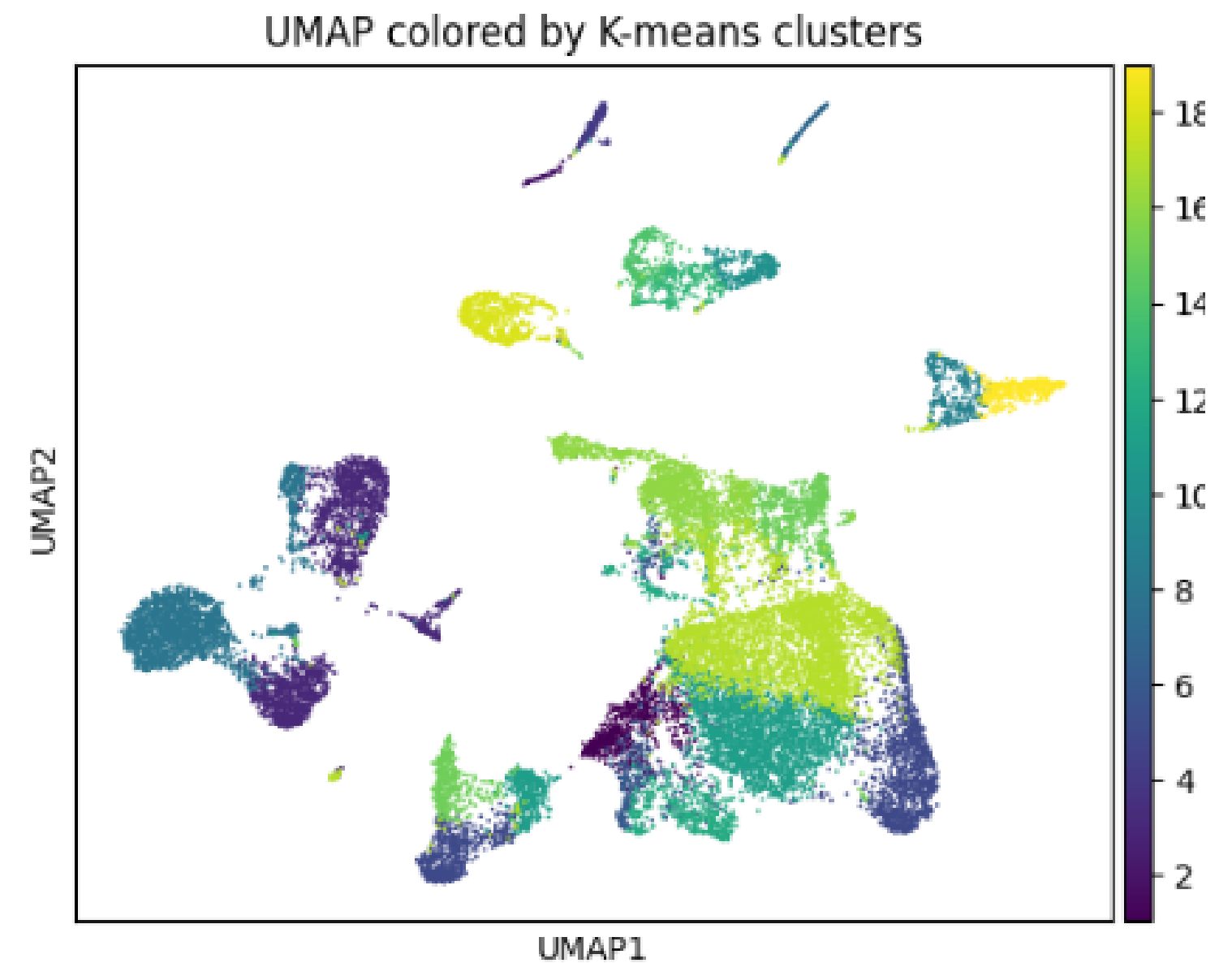
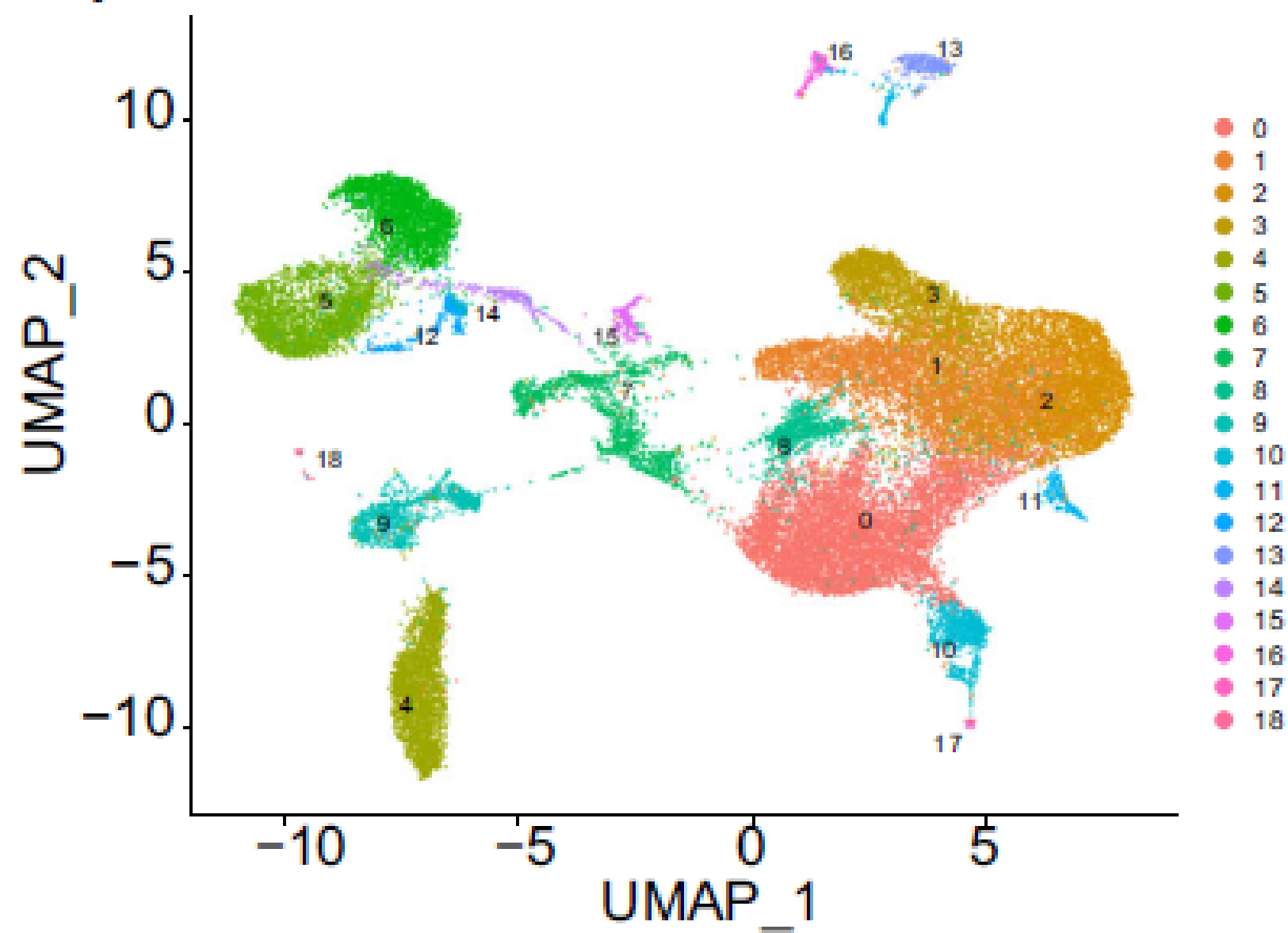
Padrões dos clusters reproduzidos diferentes do artigo de base:

- Ferramenta
- Biblioteca
- Processamento
- Filtragem

VI CONCLUSÕES

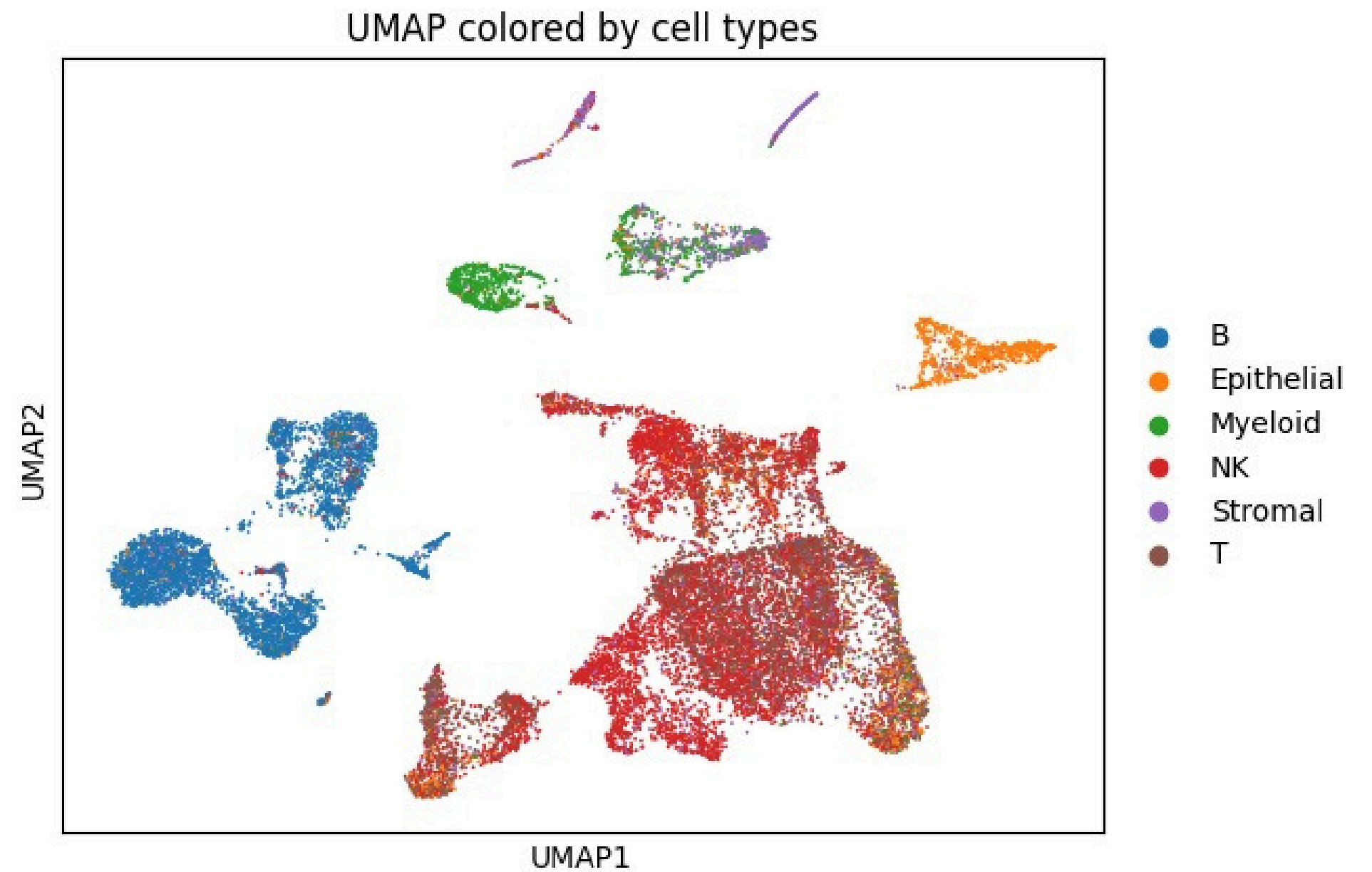
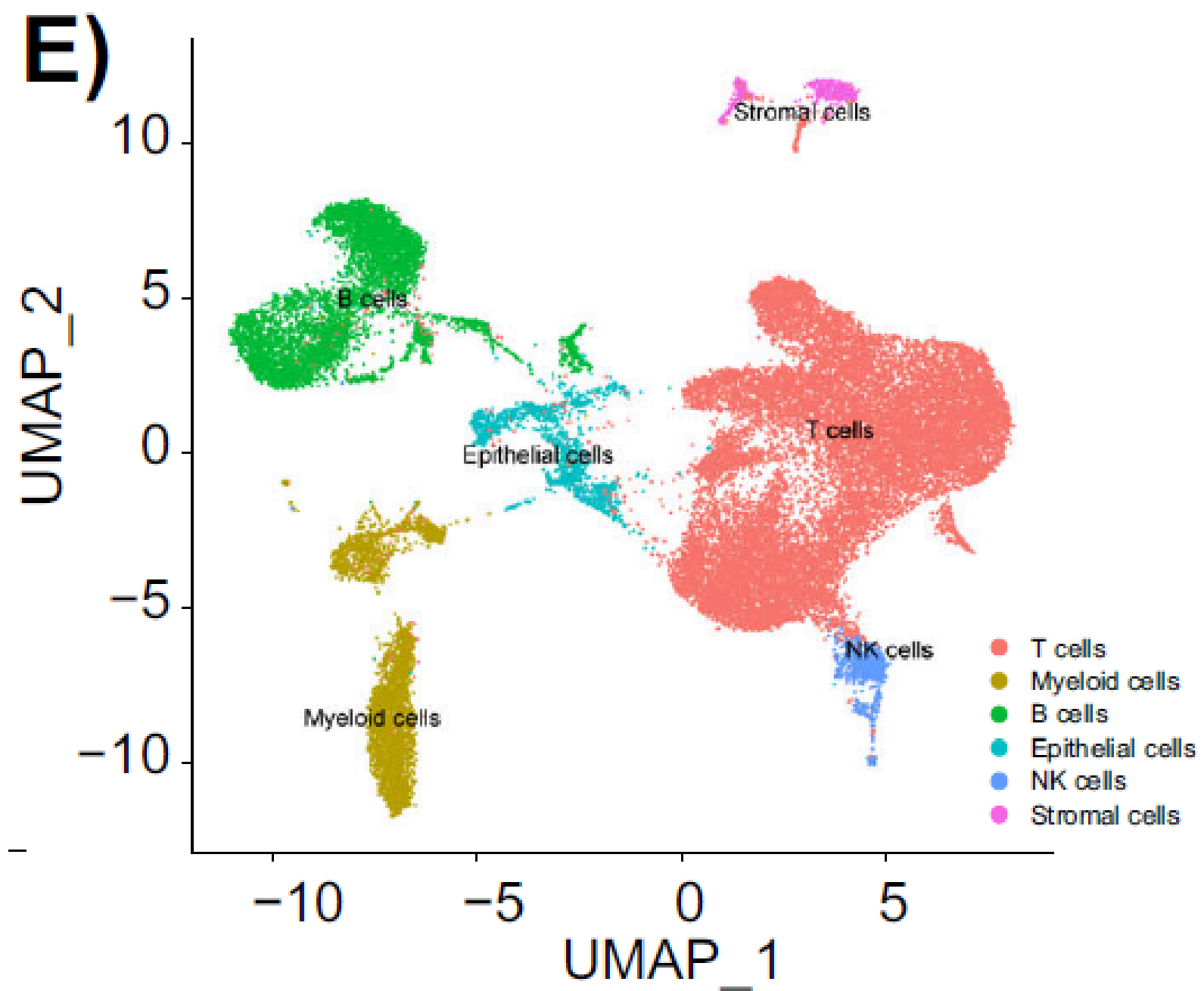
Heterogeneidade do Microambiente Tumoral

A)



VI CONCLUSÕES

Anotação dos Tipos Celulares



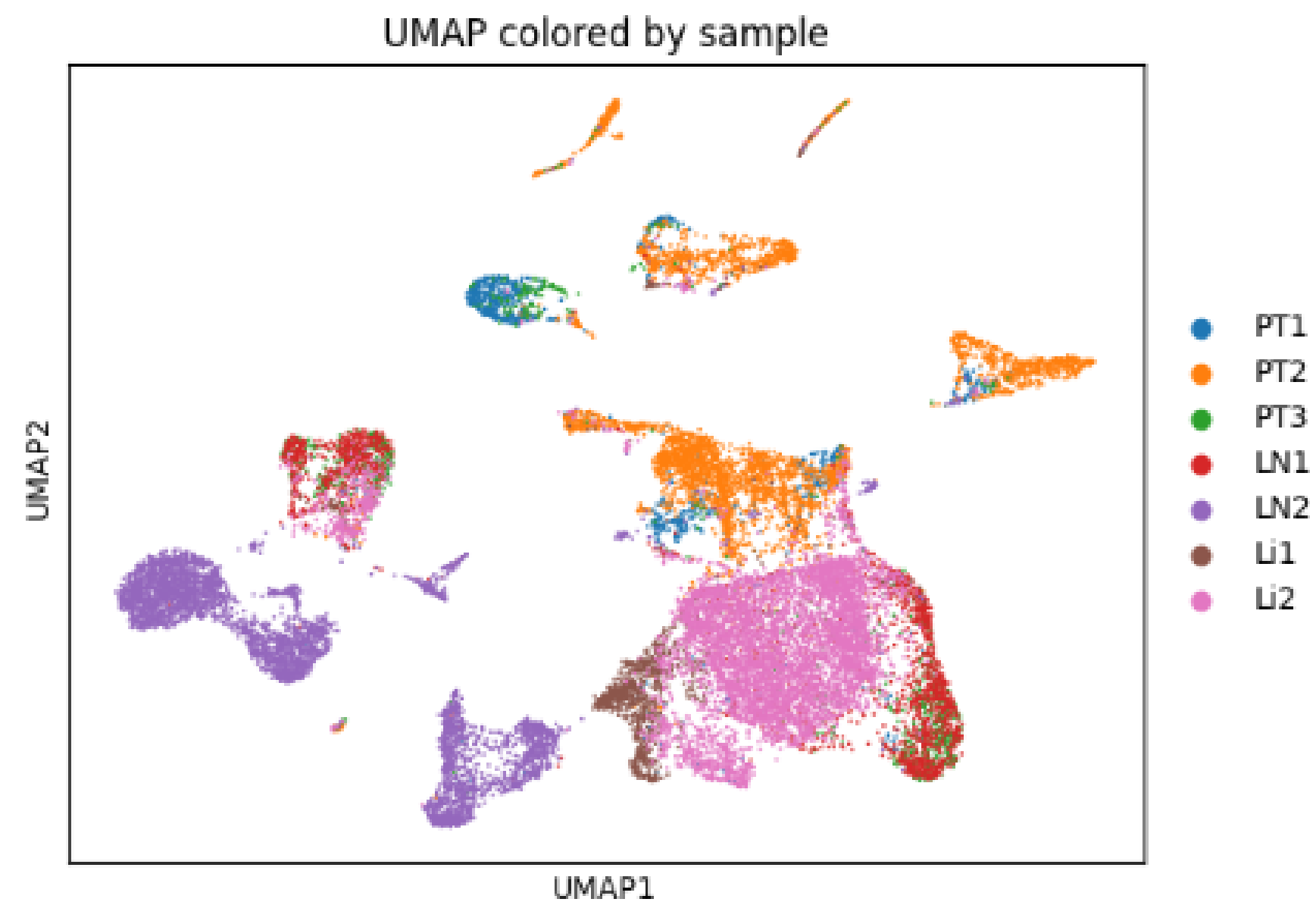
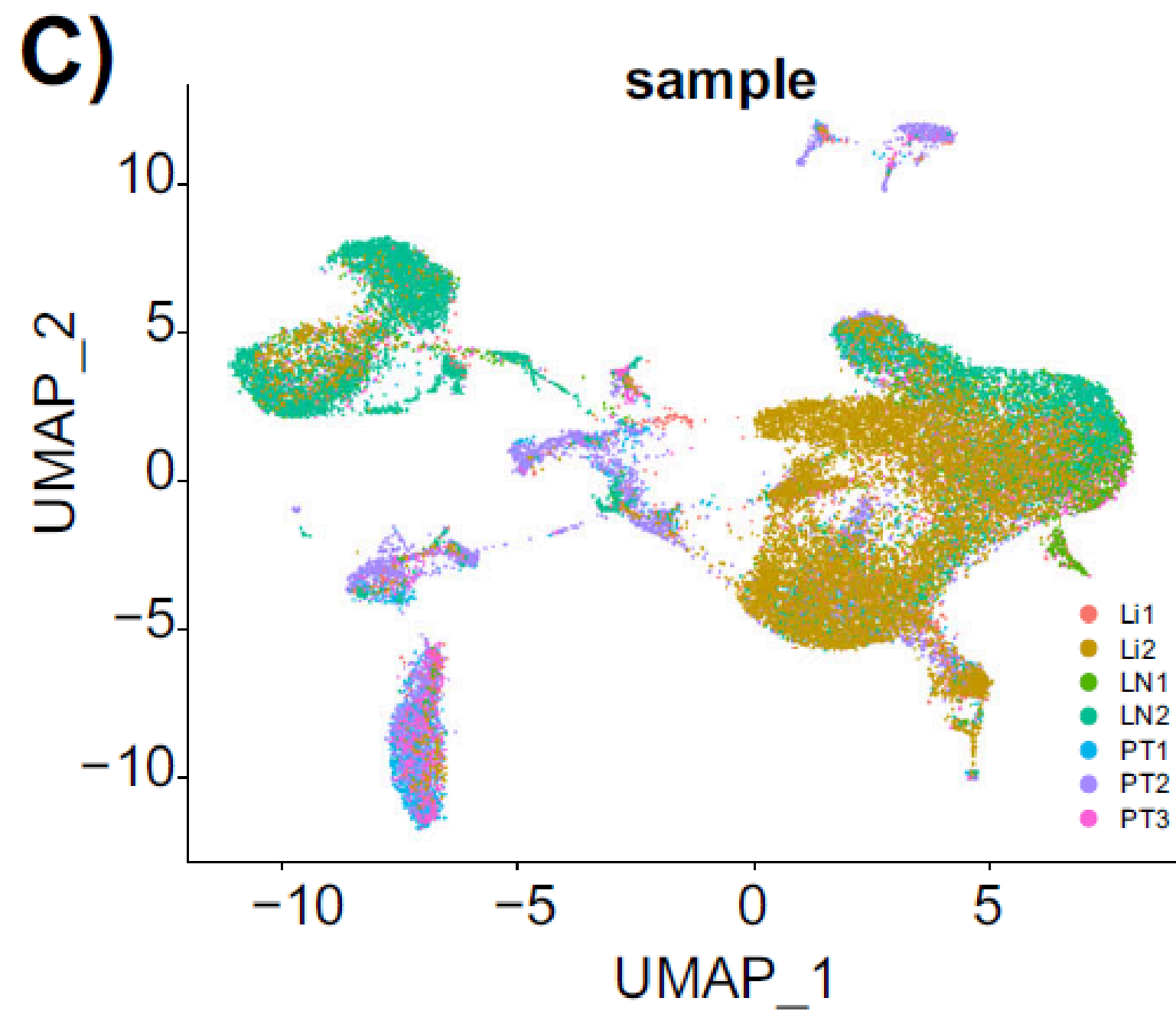
VI CONCLUSÕES

Heterogeneidade do Microambiente Tumoral

- Clara segregação entre clusters distintos
- scRNA-seq: genes altamente variáveis dentro de células homogêneas
- Cada tipo celular desempenha um papel específico na progressão do câncer e na resposta e mecanismos de escape imune
- Heterogeneidade indica maior resistência a terapias

VI CONCLUSÕES

Heterogeneidade Entre os Tecidos



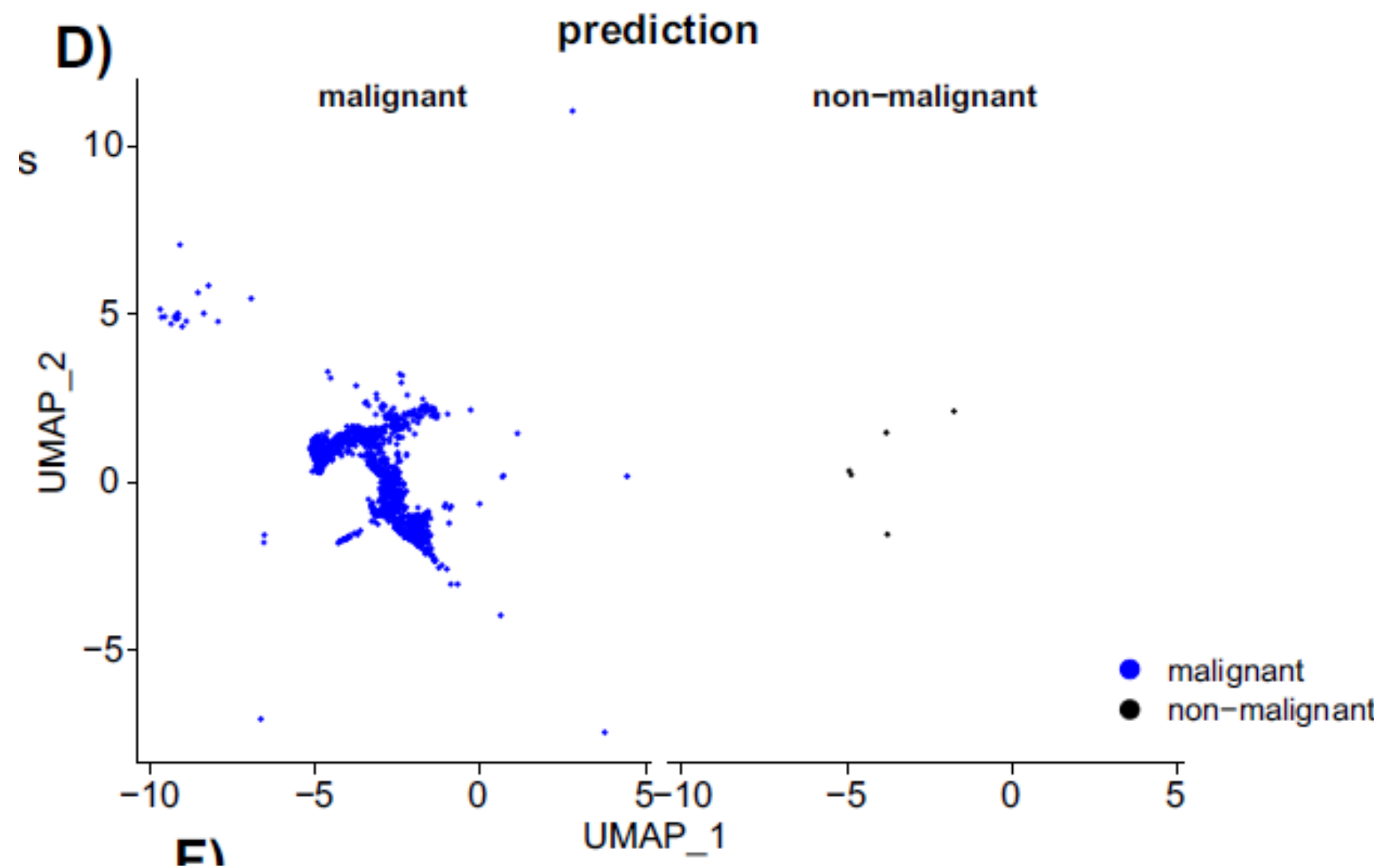
VI CONCLUSÕES

Heterogeneidade Entre os Tecidos

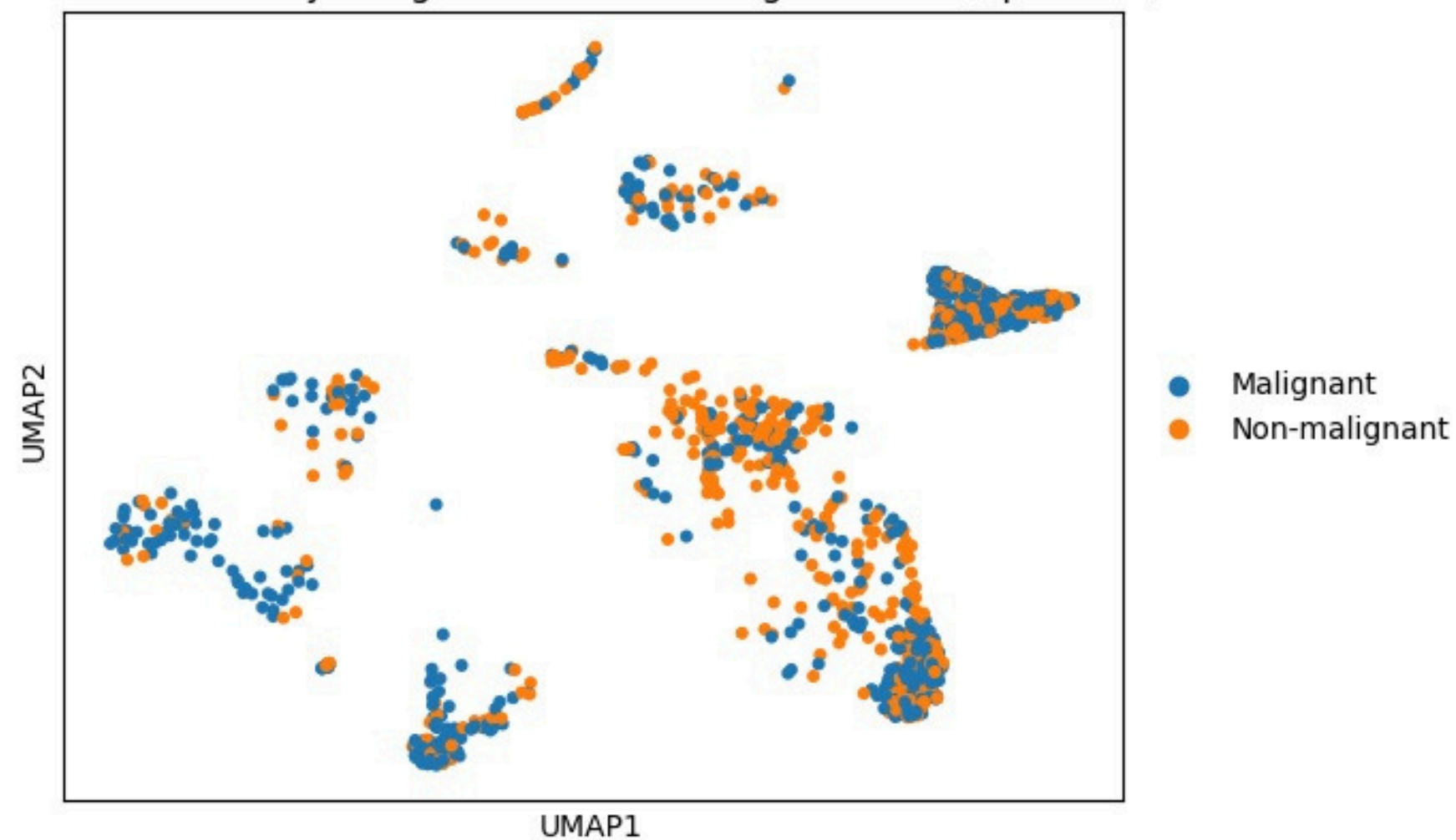
- Pode influenciar diretamente a heterogeneidade das células e nos níveis de expressão gênica
- Como as células tumorais evoluem e se adaptam ao longo da progressão do câncer
- Linfonodos são frequentemente caminhos para a metástase
- Fígado é um importante destino de metástase no CG

VI CONCLUSÕES

Células Epiteliais Malignas x Não Malignas



UMAP colored by Malignant and Non-malignant cells (Epithelial)



VI CONCLUSÕES

Heterogeneidade Entre os Tecidos

- O CG se origina do epitélio glandular da mucosa gástrica
- Proporção a partir de K-médias
- Avaliação de características e marcadores que as diferenciam
- CXCR4: alta capacidade metastática, resistência a quimioterápicos e pior prognóstico
- Marca a presença de células-tronco cancerosas = alvos importantes para o tratamento

AGRADECEMOS A ATENÇÃO!