

GENETIME CONFERENCE 2026

11 a 14 de agosto 2026

TÍTULO: Introdução à análise downstream de dados de Single-Cell RNA-Seq em R

EMENTA (conteúdo abordado):

Pequeno Resumo:

Este minicurso tem como objetivo capacitar pesquisadores e estudantes na análise primária e secundária de dados de transcriptômica de célula única (scRNA-seq). O curso abordará o fluxo de trabalho de bioinformática downstream utilizando a linguagem R, partindo de matrizes de contagem até a extração de insights biológicos. Os participantes aprenderão as etapas críticas de controle de qualidade, normalização, redução de dimensionalidade e anotação de tipos celulares, adquirindo autonomia para aplicar o *pipeline* em seus próprios projetos, independentemente da plataforma de captura inicial (como 10x Genomics ou BD Rhapsody).

Tópicos a serem abordados:

Módulo 1: Fundamentos e Preparação dos Dados

- Breve introdução dos conceitos e as tecnologias de scRNA-seq (diferenças entre captura baseada em gotículas vs. poços).
- O que é uma matriz de contagem e como importá-la para o ambiente R.
- Apresentação dos principais pacotes do ecossistema R (ex: Seurat ou SingleCellExperiment).

Módulo 2: Controle de Qualidade (QC)

- Por que o QC é fundamental em single-cell?
- Mapeamento de métricas essenciais: contagem de UMI, número de genes detectados e porcentagem de transcritos mitocondriais.
- Estratégias para identificar e remover células de baixa qualidade, *doublets* (gotículas/poços com mais de uma célula) e gotículas vazias.

Módulo 3: Normalização e Identificação de Variáveis

- Métodos para normalizar os dados de expressão.
- Seleção de genes altamente variáveis (Highly Variable Genes - HVGs) para guiar a análise.
- Escalonamento dos dados (Scaling).

Módulo 4: Redução de Dimensionalidade e Clusterização

- Análise de Componentes Principais (PCA): reduzindo a complexidade dos dados.
- Visualização não-linear: introdução prática ao UMAP (e/ou t-SNE).

- Algoritmos de clusterização baseados em grafos para agrupar células com perfis transcricionais semelhantes.

Módulo 5: Anotação Biológica

- Como identificar genes marcadores (expressão diferencial entre os clusters).
- Estudo de caso prático: anotação manual de um dataset de células imunes (identificando populações de linfócitos T, células B, células NK e linhagem mielóide a partir dos perfis de expressão).

Módulo 6: Análise de Expressão Diferencial (DEA) entre Condições

- O conceito de expressão diferencial em dados de célula única: comparando o mesmo tipo celular em cenários diferentes (ex: Controle vs. Condição/Tratamento).
- Utilização de funções nativas do R para extrair genes diferencialmente expressos.
- Como interpretar os resultados: entendendo os conceitos de *Log2 Fold Change* (Log2FC) e *p-value* ajustado.
- Visualização de dados: gerando e interpretando gráficos essenciais como *Volcano Plots* e *Violin Plots* para ilustrar a mudança na expressão gênica.

Metodologia:

O minicurso será de natureza teórico-prática. Os conceitos fundamentais serão apresentados brevemente, seguidos imediatamente pela aplicação prática em um script de R (estilo hands-on). Será utilizado um dataset público simplificado (GEO) para garantir que todos os participantes consigam rodar as análises em seus próprios computadores durante o curso.

Pré-requisitos (para os alunos):

Conhecimento básico em lógica de programação e linguagem R.

Objetivos do curso:

- Compreender a estrutura de dados gerados por sequenciamento de célula única.
- Realizar o controle de qualidade (QC) e a filtragem de dados brutos.
- Aplicar métodos de normalização e redução de dimensionalidade (PCA, UMAP).
- Executar a clusterização de células e identificar genes diferencialmente expressos para anotação de populações celulares.
- Realizar análise de expressão diferencial entre condições e interpretação biológica dos resultados.

CARGA HORÁRIA: () 4H (x) 8H