Software development in single-cell RNA data analysis considering pseudotime

Department of RNA Biology and Neuroscience, Graduate School of Medicine

Associate Professor Yuki Kato

Researchmap https://researchmap.jp/read0207242?lang=en

Abstract

Single-cell RNA sequencing data provide information on the RNA population of just a single cell from a large population of mixed cells, which enable us to perform cell trajectory analysis for clarifying cellular transitional processes. Early methods for comparing cell trajectories, however, require selection of a pair of linear paths in branching trajectories using prior knowledge. Our team has developed CAPITAL, whose innovation means that complex branching trajectories can now be accurately and automatically aligned and compared using a method knows as tree alignment.

Background & Results

A population of cells taken from a tissue at a certain time can be considered as a snapshot that contains cells ranging from undifferentiated cells to matured cells. These heterogeneous cells with their expression profiles are grouped and placed in a time series, resulting in a cell trajectory along a hypothetical path that reflects their progress through the transitional process. This approach has a major advantage of short time and low cost, which can "see the wood for the trees." If two cell trajectories from different experiments are aligned, potential regulatory genes varying across the conditions can be identified.

Here our team has developed CAPITAL (comparative analysis of pseudotime trajectory inference with tree alignment), where two cell trajectory trees that consist of cell types are aligned, and then cell-cell alignment for each pair of linear paths is performed. After developing the algorithm, we tested it on synthetic datasets with multiple branches. The results demonstrated that CAPITAL is statistically more accurate and robust than the data integration methods that existed previously, showing major advances over these methods. We next tested CAPITAL on authentic datasets from human and mouse bone marrow cells, indicating that CAPITAL can reveal the existence of different molecular patterns across different species even when the expression patterns are similar and appear to be conserved.

Significance of the research and Future perspective

CAPITAL can compare complex datasets from single cells undergoing transitional processes. For one promising example, consider a population of blood cells, which are known to differentiate with multiple branches. We can first obtain a single-cell dataset of a diseased patient and that of a healthy control, and then run CAPITAL to identify novel disease-associated genes. We aim at characterizing disease states using multi-omics data, including transcriptome, epigenome, and proteome. Our powerful new technique is not limited to just the type of transcriptome data but applicable to a wide range of those high-throughput datasets. Thus, CAPITAL represents a significant advance in the field of single-cell biology.

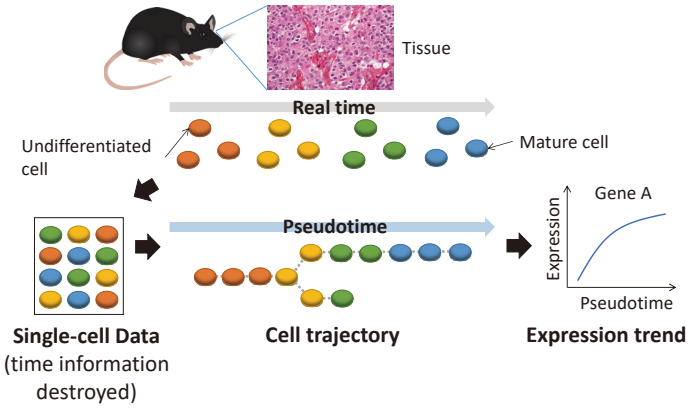

Fig. 1: Overview of cell trajectory analysis using single-cell data

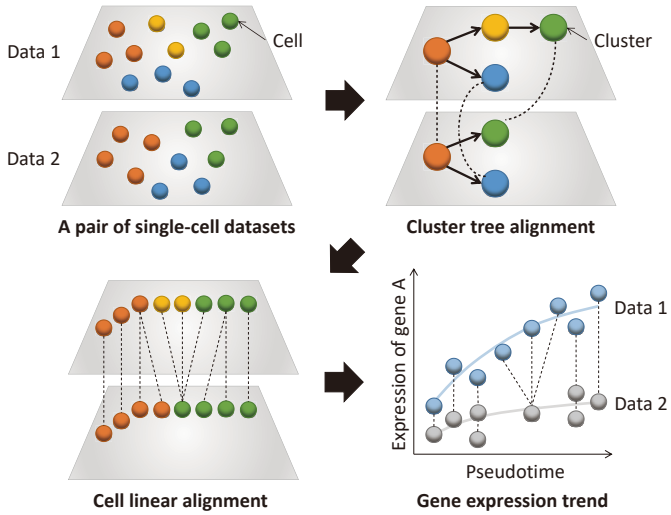

Fig. 2: Overview of our algorithm for aligning branching cell trajectories

Patent

Treatise Sugihara, Reiichi; Kato, Yuki et al. Alignment of single-cell trajectory trees with CAPITAL. Nature Communications. 2022, 13, 5972. doi: 10.1038/s41467-022-33681-3

URL https://www.med.osaka-u.ac.jp/pub/rna/ykato/en/research.html

Keyword single cell, gene expression, trajectory inference, alignment