

数学と情報科学で解き明かす多様な対象の数理構造と活用
2020 年度採択研究者

2021 年度 年次報告書

井元 佑介

京都大学 高等研究院
特定助教

多重解像度の細胞分化構造解析システムの確立

§ 1. 研究成果の概要

本研究は、細胞集団から単一細胞までの多重解像度の数理構造をシングルセル遺伝子発現データから抽出する解析システムの開発およびその数理基盤を確立し、細胞分化の原理の解明やその理解に基づく再生医療などへの臨床応用を目指すものである。2021年度(研究2年次)は主に以下の4点を実施した。

初年度に開発を進めたシングルセル遺伝子発現データのノイズ削減手法 RECODE の実データによる検証を進めた。さらに、RECODE の python/R コード、アプリケーションを公開し、生物学者からの利用を進めている。

位相的データ解析手法 Mapper の拡張手法 V-Mapper を開発した。V-Mapper はシングルセル遺伝子発現データと RNA velocity によって得られる速度場を入力とすることで、細胞分化のトポロジー構造と分化の流れを同時に取得できる。本手法をシングルセル遺伝子発現データに適用し、新たな細胞分化構造を抽出することに成功した。

代表者らは統計的因果探索手法 LiNGAM を遺伝子発現ダイナミクスの微分方程式モデルへ拡張した GRN-LiNGAM を開発している。本年度は RNA velocity から得られるシングルセル遺伝子発現の離散ベクトル場を追加入力とすることで、GRN-LiNGAM を 1 タイムポイントデータのみから GRN を推定可能になった。この GRN-LiNGAM をヒト始原生殖細胞誘導データへ適用し、既知の制御関係を含むネットワークを推定することに成功した。

生殖細胞の発生過程において、細胞分化に伴ってクロマチンが脱凝集化するという生物学的予想を裏付けるために、Hi-C データを用いてクロマチンの支配領域の推定を行った。MDS 法の誤差関数をもとに Hi-C データを距離行列に変換および 3 次元構造を推定し、その 3 次元構造の凸包の体積を計算することで、クロマチンの支配領域を定量的に示した。さらに、DNA FISH を用いてクロマチンの支配領域を実験的に調べたところ、データ解析に適合した結果が得られたため、数学と生物学の融合研究によって生物学的予想が実証された。

【代表的な原著論文情報】

- 1) Resolution of the curse of dimensionality in single-cell RNA sequencing data analysis. Y Imoto, T Nakamura, E G Escobar, M Yoshiwaki, Y Kojima, Y Yabuta, Y Katou, T Yamamoto, Y Hiraoka, M Saitou. *bioRxiv* (プレプリント), May 2, 2022.
- 2) Nucleome programming is required for the foundation of totipotency in mammalian germline development. M Nagano, B Hu, S Yokobayashi, A Yamamura, F Umemura, M Coradin, H Ohta, Y Yabuta, Y Ishikura, I Okamoto, H Ikeda, N Kawahira, Y Nosaka, S Shimizu, Y Kojima, K Mizuta, T Kasahara, Y Imoto, K Meehan, R Stocsits, G Wutz, Y Hiraoka, Y Murakawa, T Yamamoto, K Tachibana, J-M Peters, L A Mirny, B A Garcia, J Majewski, M Saitou. *The EMBO Journal*, *accepted*.
- 3) Difference between smoothed particle hydrodynamics and moving particle semi-implicit operators. Y Imoto. *Computer Methods in Applied Mechanics and Engineering*, Vol. 395, 115012, May 15, 2022.