

従来の医学研究では時系列が扱われることは少なく、個別の状態を様々な分子生物学・生化学的手法により切り取り、局所的なスナップショットとして生命現象・疾患を理解してきた。また、疾患の数理解析としては小規模なシグナル伝達系に限局した分子レベルの詳細なモデル化が中心であった。生命は局所での変化を代償する機構（ロバストネス）を持っているのに加えて、ミクロにおける分子反応はミリ秒単位の時間スケールで起こるため、日・月単位で変化する生命システム全体の変動とはタイムスケールが違いすぎるといった問題が存在しており、ミクロモデルの総和としてマクロな変化を説明することは難しい。したがって、疾患発症、病態進行の予測を行うためには、①マクロレベルの疾患状態の適切な表現と層別化、②散発的でプロセスモデルが未知の臨床時系列データに適した数理手法の整備、が必要である。

システム生物学においては、ネットワークを用いることで複雑な分子間相互作用をモデル化するという手法が用いられる。私たちは今までに、オミクスデータとネットワークに基づいて細胞・組織の制御因子を推定し、状態を表現する手法を開発してきた (Kawakami et al. *npj Syst. Biol. App.* 2016, Kawakami et al. *Nucleic Acids Res.* 2016)。一方、システム生物学の問題として、散発的な時系列を扱う方法論に乏しく、観測データに基づいたデータ駆動型の解析手法が確立されていない点が挙げられる。私たちは、機械学習および最新の数理科学手法を導入することで、網羅的測定データに基づく疾患の多様性の理解、および状態遷移の規則・秩序を抽出と生命現象・疾患の予測モデルの構築を目指して研究を行っている。今回の発表では、最近私たちが取り組んでいる、「機械学習に基づいた疾患の層別化と複合バイオマーカー探索」および「不規則で散発的な医学生物学時系列データ解析」を紹介し、複雑で多様な疾患・生命現象の理解に向けた取り組みを議論したい。

参考文献：

1. Kawakami E*, Nakaoka S, Ohta T, Kitano H., Weighted enrichment method for prediction of transcription regulators from transcriptome and global chromatin immunoprecipitation data., *Nucleic Acids Res.*, 44, 5010-21, 2016
2. Kawakami E, Singh VK, Matsubara K, Ishii T, Matsuoka Y, Hase T, Kulkarni P, Siddiqui K, Kodilkar J, Danve N, Subramanian I, Katoh M, Shimizu-Yoshida Y, Ghosh S, Jere A, Kitano H., Network analyses based on comprehensive molecular interaction maps reveal robust control structures in yeast stress response pathways., *npj Syst. Biol. App.*, 2, 15018, 2016

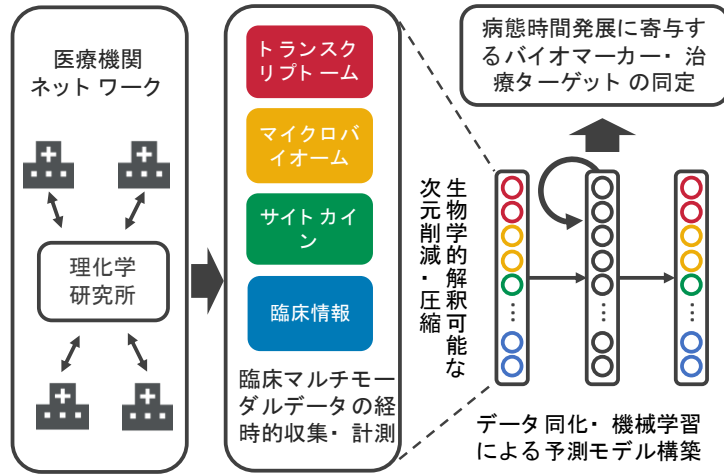


図1. 機械学習・数理科学に基づく疾患予測モデル構築

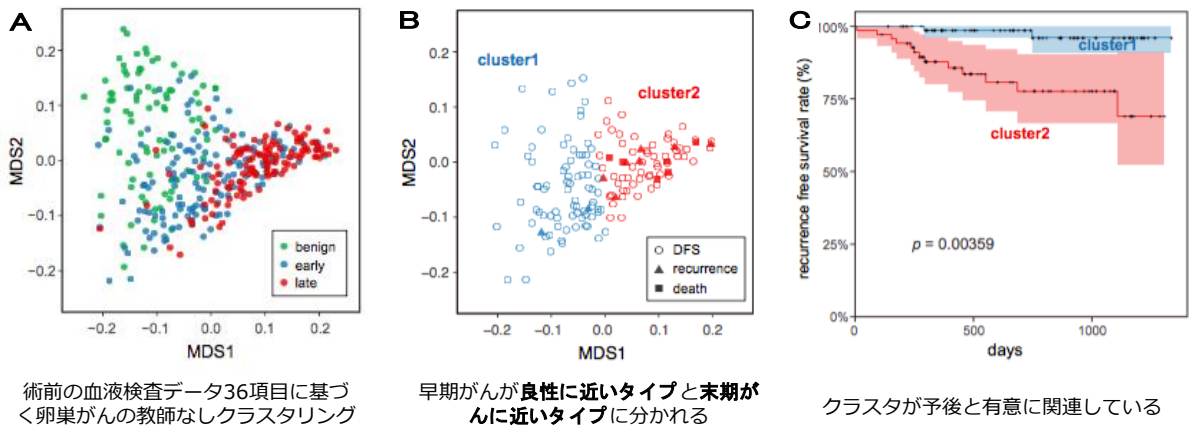


図2. 術前血液検査データに基づく卵巣がん層別化