第22回 統合病理学教室セミナー



演者　丹下（たんげ）　正一朗（しょういちろう）　先生

札幌医科大学医学部
ゲノム予防医学講座ゲノム医科学分野 助教

演題

腫瘍学研究における公共データベース活用の光と影
～RNA-Seqデータのジェネレーションギャップと解析の落とし穴

2026年11月6日（金）17：00～

医学部中研究棟 3－1共通セミナー室

要旨

20年前に立ち上がったThe Cancer Genome Atlas（TCGA）プロジェクトでは、33がん種・1万人を超える患者のマルチオミクス情報を収載しており、cBioPortalやUCSC Xenaをはじめとする多彩なWebツールの普及とも相まって、現代の腫瘍学研究において重要なインフラとなっている。しかし、これらのデータベースは運用時期によって解析パイプラインやゲノムアノテーション（GTFファイル）が繰り返し更新されており、同一の臨床検体由来データであってもバージョン（世代）間で質的・量的な不整合が生じている。公共データの利便性（光）の裏には、データの変遷や仕様の限界を正しく解釈しなければ誤った研究仮説に誘導されるリスク（影）が潜んでいる。公共データベースは極めて強力な解析基盤である一方、データ算出の裏にあるパイプラインの歴史的背景、GTFアノテーションの仕様、アライメント/カウントアルゴリズムの挙動（オーバーラップ遺伝子やリードスルー転写産物の扱い）を正しく理解して利用することが不可欠である。

上述のようなデータ世代間の違いやアノテーションの差異に起因する影響を迅速に確認できるよう、我々は発現分布やがん種・サブタイプ別の発現量を並列で比較・直感的に可視化できるブラウザベースのWebツールを構築・公開した（https://5spgm.github.io/tcga-expression-explorer/）。本発表を通じて、データ解析における「罠」を回避し、ドライ解析とウェット検証を適切に組み合わせる重要性について議論したい。

参考文献

1. https://5spgm.github.io/tcga-expression-explorer/ (論文作成中)
2. Tange and Idogawa, https://doi.org/10.5281/zenodo.22670234 (2026)
3. Kamikokura et al., Cancers, 16(5), 882 (2024)
4. Tange et al., BMC Cancer, 23:85 (2023)
5. Idogawa et al., Cancer Biology and Therapy, 22, 527–528 (2022)

責任者：統合病理学教室　谷口　浩二（内線5050）
E-mail: path1@med.hokudai.ac.jp