生物統計学/疫学・予防保健学　抄読会

2025 年 1 月 28 日(水)
生物統計学分野 博士課程 2 年 田中晴菜

Investigating the applicability of Uniform Manifold Approximation and Projection (UMAP) to metabolomics analysis

Cancer is a complex and heterogeneous disease characterized by extensive metabolic reprogramming, which plays a central role in tumor progression, therapeutic resistance, and patient prognosis[1]. Metabolomics provides a comprehensive functional readout of cellular states and has become an essential tool for exploring cancer biology, biomarker discovery, and disease stratification. However, metabolomics datasets are typically high-dimensional and noisy, with far more variables than samples, posing substantial challenges for conventional statistical analysis.

Unsupervised dimension reduction methods are widely used for exploratory analysis and visualization of such data. Principal component analysis (PCA) has been the most commonly applied technique, but its linear nature and focus on global variance may limit its ability to capture subtle nonlinear structures and local heterogeneity in complex metabolic profiles[2]. Recently, Uniform Manifold Approximation and Projection (UMAP), a nonlinear manifold learning method, has gained increasing attention for preserving local neighborhood structure while maintaining global relationships[3].

In this presentation, I would like to introduce UMAP, which has recently attracted attention as a promising method for exploratory analysis of metabolomics data, and discuss its applicability. After briefly reviewing the principles and limitations of PCA, I will introduce the basic concept of UMAP and illustrate its performance in visualizing high-dimensional metabolomic profiles. Finally, I will discuss how UMAP may facilitate the discovery of latent metabolic patterns and biological heterogeneity and outline its potential advantages and limitations in metabolomics research.

References

1. Liu S, Zhang X, Wang W, et al. Metabolic reprogramming and therapeutic resistance in primary and metastatic breast cancer. *Mol Cancer*. 2024;23(1):261–57.

2. Walsh MC, Brennan L, Malthouse JPG, Roche HM, Gibney MJ. Effect of acute dietary standardization on the urinary, plasma, and salivary metabolomic profiles of healthy humans. *The American Journal of Clinical Nutrition*. 2006;84(3):531–539.

3. Idkowiak J, Dehairs J, Schwarzerová J, et al. Best practices and tools in R and python for statistical processing and visualization of lipidomics and metabolomics data. *Nat Commun*. 2025;16(1):8714–19.