Precision Medicine: Innovative methods for advanced understanding of molecular underpinnings of disease

Yana Bromberg†
Department of Biology, Emory University, 1510 Clifton Rd
Department of Computer Science, Emory University, 400 Dowman Dr.
Atlanta, GA 30317, USA
Email: yana.bromberg@emory.edu

Hannah Carter†
Department of Medicine, University of California San Diego, 9500 Gilman Dr.
La Jolla, CA 92093, USA
Email: hkcarter@health.ucsd.edu

Steven E. Brenner
Department of Plant & Microbial Biology, University of California Berkeley, 461 Koshland Hall
Berkeley, California 94720-3102, USA
Email: brenner@compbio.berkeley.edu

Precision medicine, also often referred to as personalized medicine, targets the development of treatments and preventative measures specific to the individual's genomic signatures, lifestyle, and environmental conditions. The series of Precision Medicine sessions in PSB has continuously highlighted the advances in this field. Our 2024 collection of manuscripts showcases algorithmic advances that integrate data from distinct modalities and introduce innovative approaches to extract new, medically relevant information from existing data. These evolving technology and analytical methods promise to bring closer the goals of precision medicine to improve health and increase lifespan.

1. Introduction

Precision medicine involves tailoring medical decisions and treatments to individual patients in a data-driven manner. The accumulation of medically-relevant and, particularly, molecular data has uncovered the potential for mechanistic insight into disease processes facilitating clinical decision making. Advances in genomic techniques, e.g. spatial transcriptomics and single cell analysis, have further enabled identification of the genetic biomarkers of patient drug responses, susceptibility to diseases, and other medically-relevant outcomes. At the same time, the enormous scale of this data has stimulated use of novel computational methods, resulting in, e.g., the recent explosion in deep learning-based, biological and medical data analysis techniques.

† Corresponding Authors

While the concept of personalized medicine stretches back nearly two decades – just slightly longer than our PSB session – the implementation of precision medicine in practice (still) remains in its early stages. Novel technologies require new integrative approaches to advance the state of the art in this field. In our 2024 session we feature work from researchers across diverse domains, who integrate various omics data to provide valuable insights into disease mechanisms, diagnosis, and treatment. In this collection, we explore their cutting-edge advancements in more detail.

2. Session Contributions

2.1. *Transcriptome and Histopathology Integration*

A number of studies submitted to our session focused on integrating spatial transcriptomics and histopathology data and demonstrating the potential of this combination to enhance our understanding of tumor biology. Song et al enriched their transcriptome-driven findings by incorporating morphological features extracted from histopathology images to enable a comprehensive analysis of tumor architecture via feature clustering. Azher et al employed contrastive learning and Graph Convolutional Neural Networks (GCN) to predict disease stage, lymph node metastasis, and survival prognosis in cancer patients. Meanwhile, Srinivasan et al developed a transformer-based model to shed light on the molecular pathways involved in skin aging due to light exposure. Their findings not only contribute to our understanding of their chosen conditions but also demonstrate the potential of their approaches for studying other diseases.

2.2. *Spatial Proteomics: Revealing Tissue Microenvironments.*

Wu et al introduced innovative methods to analyze tissue microenvironments at high resolution using spatial proteomics. By measuring inferred protein polarity, they identified distinct subpopulations of immune cells within tumors, shedding light on potential markers of better prognosis. This approach holds promise for identifying patients who may respond favorably to specific treatments.

2.3. *Microbiome Analysis: A Closer Look at Gut Health.*

Sapoval et al proposed a novel metagenomic analysis pipeline that bypasses the need for genome assembly, allowing for direct comparisons between patients and healthy controls. This reference-free approach is particularly valuable for studying conditions like myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS), where the gut microbiome plays a crucial role. Understanding dysbiosis at this level can help identify potential disease markers and therapeutic targets.

2.4. *Polygenic Risk Scores: Bridging Genomics and Disease.*

Cardone et al examined the role of a lymphocyte count PRS (polygenic risk score) in predicting CD4 T-cell recovery in individuals with HIV undergoing anti-retroviral therapy. While their findings indicated limited PRS impact compared to clinical factors, they underscore the importance of considering multiple variables in precision medicine studies. Kember et al focused on improving PRS accuracy for cardiometabolic traits; their findings emphasize the need for implementing separate scoring mechanisms for diverse ancestries.

2.5. *Integrative Methods for Clustering, Meta-analysis, Deconvolution, and Network Rewiring.*

Numerous contributions aimed at enhancing integrative methods for meta-analysis, subtype detection, cell type deconvolution, and network rewiring. Zhang et al introduced nSEA, an algorithm for unsupervised clustering of low grade Gliomas, uncovering a novel subtype with clinical implications. Huang et al

proposed a multi-modal clustering approach that combines various data types to cluster tumor samples across different cancer types, offering a more holistic perspective on cancer classification.

In the realm of meta-analysis, Fukutani et al overcame batch effects in gene expression data, highlighting the importance of robust analytical techniques in large-scale studies. Sufriyana et al employed data-driven ontology inference to uncover novel gene sets relevant to subtypes of preeclampsia, showcasing the power of meta-analysis in identifying novel biological processes.

Deconvolution, a vital tool for deciphering cellular composition from omics data, faced challenges in understanding nascent RNA. Maas et al introduced an adaptation for nascent RNA sequencing, addressing the nuances of this emerging field.

Finally, Dannenfelser et al explored how alternative splicing rewires protein interaction networks in cancer. Their Splitpea method provides insights into the complex interplay between alternative splicing and disease, offering a novel perspective on cancer biology.

References

1. Azher, Z., Fatemi, M., Lu, Y., Srinivasan, G., Diallo A., Christensen, B., Salas, L., Kolling IV, F., Perreard, L., Palisoul, S., Vaickus, L., Levy, J. (2024) Spatial Omics Driven Crossmodal Pretraining Applied to Graph-based Deep Learning for Cancer Pathology Analysis. *Pacific Symposium on Biocomputing 2024.*
2. Dannenfelser, R., Yao, V., (2024) Splitpea: uncovering cancer patient-specific protein interaction network rewiring. *Pacific Symposium on Biocomputing 2024.*
3. Fukutani, K.F., Hampton, T.H., Bobak, C.A., MacKenzie, T.A. , Stanton, B.A. (2024) Application of Quantile Discretization and Bayesian Network Analysis to Publicly Available Cystic Fibrosis Data Sets. *Pacific Symposium on Biocomputing 2024.*
4. Maas, Z., Sigauke, R., Dowell, R. (2024) Deconvolution of Nascent Sequencing Data Using Transcriptional Regulatory Elements. *Pacific Symposium on Biocomputing 2024.*
5. Sapoval, N., Tanevski, M., Treangen, T.J. (2024) KombOver: Efficient k-core and K-truss based characterization of chronic disease impact on the human gut microbiome. *Pacific Symposium on Biocomputing 2024.*
6. Song, J., Ramaswamy, V.G., Lamstein, J., Webb, M., Zada, G., Finkbeiner, S., Craig, D.W. (2024) Enhancing Spatial Transcriptomics Analysis by Integrating Image-Aware Deep Learning Methods. *Pacific Symposium on Biocomputing 2024.*
7. Srinivasan, G., Davis, M., LeBoeuf, M., Fatemi, M., Azher, Z., Lu, Y., Diallo, A., Saldias Montivero, M., Kolling IV, F., Perrard, L., Salas, L., Christensen, B., Palisoul, S., Tsongalis, G. Vaickus, L., Preum, S. Levy, J. (2024) Potential to Enhance Large Scale Molecular Assessments of Skin Photoaging through Virtual Inference of Spatial Transcriptomics from Routine Staining. *Pacific Symposium on Biocomputing 2024.*
8. Sufriyana, H., Wu, Y., Su, E.C.Y. (2024) Transcriptome and interactome connecting endometrial-decidua-placental origin of preeclampsia subtypes: A preliminary study. *Pacific Symposium on Biocomputing 2024.*

9. Wu, E., Wu, Z., Mayer, A.T., Trevino, A.E., Zou, J. (2024) Polarity measurements from spatial proteomics imaging suggest immune cell engagement. *Pacific Symposium on Biocomputing 2024.*
10. Zhang, Z., Wang, C., Zhao, Z., Yi, Z., Durmaz, A., Yu, J., Bebek, G. (2024) nSEA: n-Node Subnetwork Enumeration Algorithm Identifies Lower Grade Glioma Subtypes with Altered Subnetworks and Distinct Prognostics. *Pacific Symposium on Biocomputing 2024.*
11. Cardone, K.M., Dudek, S., Keat, K., Bradford, Y., Cindi, Z., Daar, E.S., Gulick, R., Riddler, S.A., Lennox, J.L., Sinxadi, P., Haas, D.W., Ritchie, M.D. (2024) Lymphocyte Count Derived Polygenic Score and Interindividual Variability in CD4 T-cell Recovery in Response to Antiretroviral Therapy. *Pacific Symposium on Biocomputing 2024.*
12. Huang, C., Kuan, P.F. (2024) intCC: An efficient weighted integrative consensus clustering of multimodal data. *Pacific Symposium on Biocomputing 2024.*
13. Kember, R.L., Verma, A., Verma, S.S., Xiao, B., Lucas, A., Kripke, C.M., Judy, R., Chen, J., Damrauer, S.M., Rader, D.J., Ritchie, M.D. (2024) Polygenic risk scores for cardiometabolic traits demonstrate importance of ancestry for predictive precision medicine. *Pacific Symposium on Biocomputing 2024.*