Bayesian Rhythmic Model for Jointly Detecting Circadian Biomarkers and Predicting Molecular Circadian Time in Human Post-Mortem Brain Transcriptome

Xiangning Xue[1], Liangliang Zhang[2], **George Tseng[1]**

[1]*Department of Biostatistics and Health Data Science, University of Pittsburgh*

[2]*Department of Population and Quantitative Health Sciences, Case Western Reserve University.*

ABSTRACT

Transcriptomic circadian analysis in human post-mortem brain has provided an unprecedented opportunity to decipher in vivo zzhowmolecular circadian rhythms in different human brain regions that are known to play essential roles in aging and psychiatric disorders. Detecting circadian biomarkers is often the first objective in the statistical analysis. Confounded with the task is the common issue that the true molecular circadian clock time within a human subject is usually inconsistent with the record, which could come from observation or recording errors, or the intrinsic biological variation. In the literature, many methods have been proposed for detecting rhythmic biomarkers or for predicting true molecular circadian time, respectively. To date, no method is developed to achieve both objectives simultaneously. In this paper, we propose a Bayesian model for simultaneous Circadian marker detection and molecular circadian Time estimation (BayCT). The model is further extended to repeated measure of multiple tissues or brain regions. We adopt Von Moses prior distribution for angular data with slice sampling and reversible jump sampling in the Markov chain Monte Carlo (MCMC) procedure for Bayesian inference. We demonstrate the method by extensive simulations and two applications in 12 tissues in mouse brain and three brain regions in human. The result shows superior performance in both statistical objectives with large margin in both circadian marker detection and circadian time detection, as well as the advantage of integrating multiple brain regions.

Keywords: circadian rhythm, omics analysis, Bayesian model

Adjusting Transcript Leakage in Spatial Transcriptomic Data

Christina Huan Shi[1,#], Yibo Zhai[2,#], Savio Ho-Chit Chow[1], Liangbang Li[2], Peter D. Adams[1], Bing Ren[3], Marissa Schafer[4], **Yingying Wei[2,*]**, Kevin Y. Yip[1,*]

[1]*Sanford Burnham Prebys Medical Discovery Institute.*

[2]*Department of Statistics and Data Science, The Chinese University of Hong Kong*

[3]*Department of Systems Biology, Columbia University*

[4]*Mayo Clinic*

[#]*Contributed equally*

ABSTRACT

Spatial transcriptomics offers unprecedented opportunities to investigate the tissue organization, cell-cell interactions and the tumour microenvironment. Compared with spot-based spatial transcriptomic technologies, which typically assay multiple cells within a single regular-shaped spot and hence yield bulk-level data, recent spatial transcriptomic platforms can achieve cellular or even subcellular resolution. However, recent benchmarking studies of spatial transcriptomic platforms have reported that certain genes appear to be expressed in cell types where such expression is biologically unexpected, even after applying highly stringent cell segmentation and filtering procedures. Our data analysis suggests that this phenomenon is likely attributable to transcript leakage from neighboring cells. However, despite active research on the deconvolution of spot-level spatial transcriptomic data, statistical methods for decontaminating spatial transcriptomic data affected by transcript leakage are lacking. Unfortunately, if the transcript leakage problem is not properly adjusted, it can lead to serious consequences for the quantification of expression levels, cell type annotation, differential gene expression detection and identification of spatially variable genes. Here, we have developed a Bayesian hierarchical model to adjust gene expression contamination resulting from transcript leakage. We have proven the model identifiability, which shows that contamination due to transcript leakage can be distinguished from the true underlying gene expression. Application to real data shows that our model successfully adjusts transcript leakage in spatial transcriptomic data.

Keywords: Spatial Transcriptomics; Deconvolution; Identifiability; Bayesian Hierarchical Model.

Computational Intelligence from Omics to Medicine

Ka-Chun Wong

Department of Computer Science, City University of Hong Kong

ABSTRACT

In recent years, the integration of Artificial Intelligence (AI) in scientific research has revolutionized the field of molecular biology and medicine. This speech aims to explore three significant aspects of computational intelligence in molecular biology and medicine. In particular, I will navigate my research group efforts from omics to medicine: bioinformatics, medical informatics, and clinical informatics. By leveraging computational intelligence, my research group has enabled multiple advances in DNA motif analysis, cancer detection, gene editing, and small-molecule drug discovery.

1. In bioinformatics, pattern recognition algorithms on DNA motifs are presented and demonstrated to be instrumental in identifying and understanding the intricate patterns within DNA sequences. These algorithms aid in deciphering gene regulatory elements, enabling researchers to unravel the complexity of genetic networks and their impact on cellular processes, enabling insights into the fundamental mechanisms governing gene expression and regulation.

2. In medical informatics, machine learning algorithms are presented to demonstrate accuracy in analysing complex medical data. By training models on vast datasets, the proposed algorithms can identify subtle patterns indicative of cancerous cells, assisting in early detection and precise localization of tumours, leading to improved patient outcomes and personalized treatment strategies.

3. In clinical informatics, several computational intelligence approaches are presented for gene editing and small-molecule drug discovery. This enables precise gene editing, offering potential therapeutic interventions for genetic disorders. Last but not least, drug docking techniques are proposed to accelerate the identification of small-molecule compounds with diffusion modelling.

Keywords: AI for Science; Bioinformatics; Medical Informatics; Gene Editing; Drug Discovery