Modular Structure of Enhancer Networks in Mixed-Phenotype Acute Leukemia Predicts Disease Genes

Mia Zhou, Hongli Lin, Jialiang Huang*
State Key Laboratory of Cellular Stress Biology, School of Life Sciences, Xiamen University, Fujian 361102, China

INTRODUCTION

Mixed-phenotype Acute Leukemias (MPAL) are characterized by the presence of blast cells that express markers of multiple lineage, often leading to diagnostic and therapeutic challenges. The regulatory mechanism driving abnormal gene expression in MPAL involve complex genetic and epigenetic interactions, with enhancers playing a pivotal role.
Advances in single-cell sequencing technologies such as scATAC-seq and scRNA-seq have enabled detailed exploration of chromatin accessibility and gene expression at the single-cell level. These technologies facilitate the identification of enhancer networks regulating each genes, providing insights into MPAL pathogenesis.
This study investigates enhancer networks in MPAL by analyzing the enhancer accessibility and gene expression data. We identify putative enhancer clusters, determine predicted enhancer interactions (PEIs) and enhancer networks to understand their regulatory roles. These networks are classified based on their modularity, and their impacts on cell identity genes is examined.
Our research aims to elucidate the epigenetic regulatory mechanisms in MPAL, identifying potential therapeutic targets to improve treatment strategies.

METHODOLOGY

Data preparation: Extract enhancer accessibility and gene expression matrices from scATAC-seq and scRNA-seq data from MPAL patients.
Node (enhancer-gene): Identify enhancer cluster regulating specific genes based on gene enhancer correlation.

- TSS range identification: locate active transcription regions.
- Gene Overlap Correlation: Find potential enhancers near genes and calculate correlation.

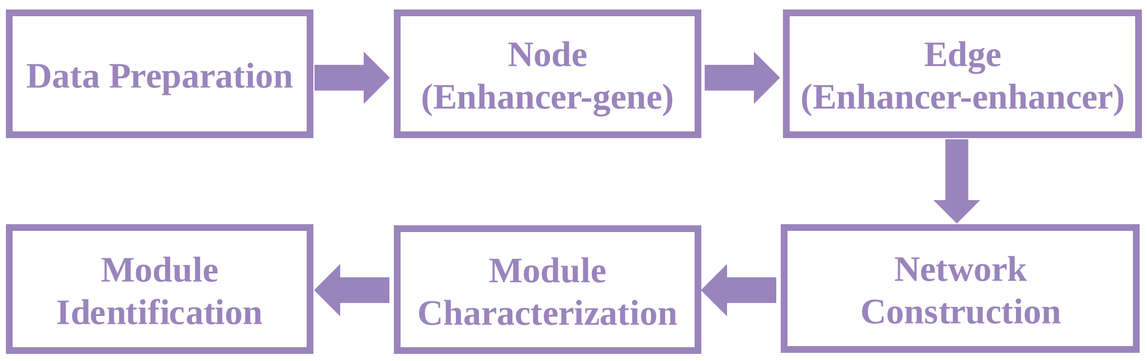
Edge(enhancer-enhancer): Identify enhancer pairs with high co-accessibility scores. The workflow enables detailed chromatin interaction analysis specific to MPAL.
Network Construction:

- Filter gene-peak pairs and co-accessibility data with significant interactions.
- Utilize parallel processing for network construction representing enhancer and promoter interactions.

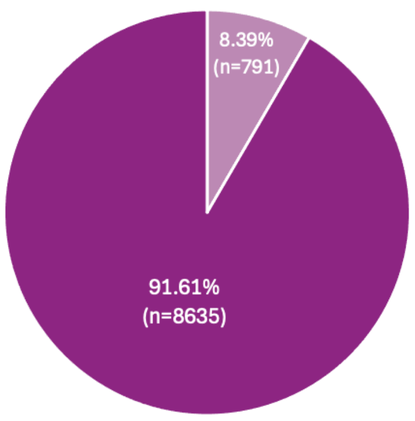
Module Identification:

- Modes: classify enhancer networks as Complex, Multiple, or Simple.
- Identify and visualize modular enhancer networks.
- Marker analysis: pinpoint genes regulated by enhancer network with significant roles in disease progression.

Module Characterization: Compare enhancer pair distance and co-accessibility within intra and inter groups to understand regulatory implication.



RESULTS AND DISCUSSION



Graph.1

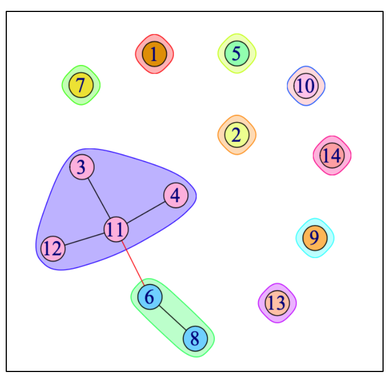


Figure.1.1. *WT1(Modular)*

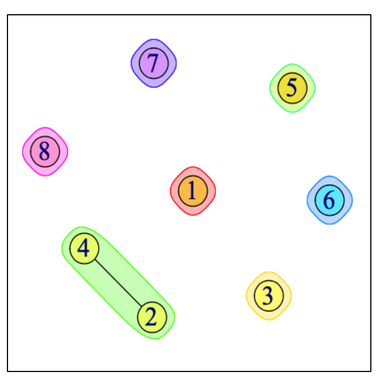
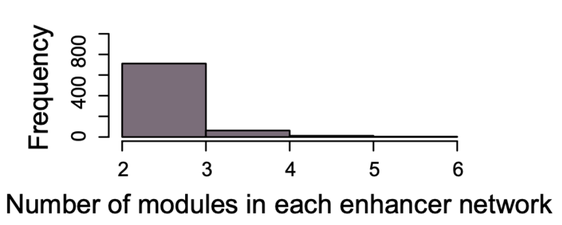
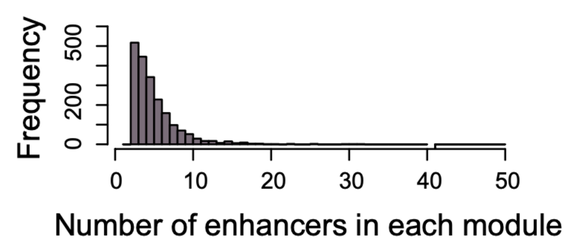


Figure.1.2. *PADI3(non-Modular)*

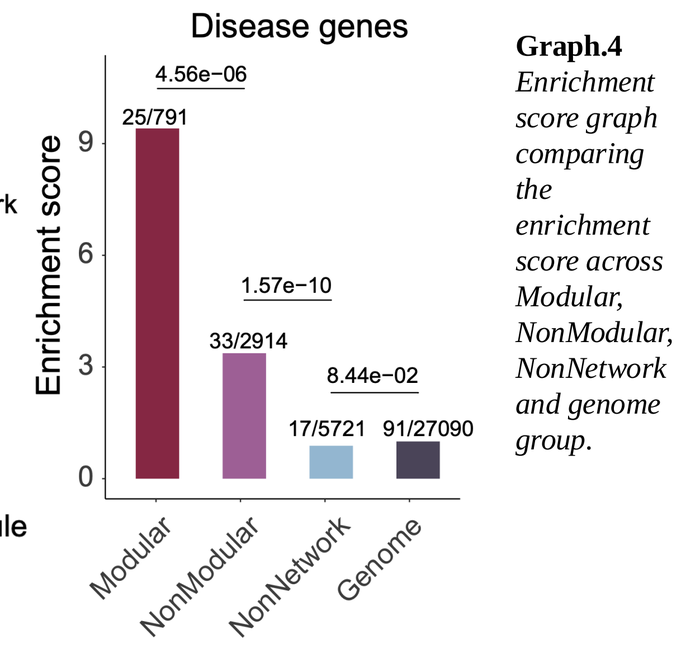
Graph1 suggests that only 8.39% of enhancer networks are classified as Modular in MPAL, despite being fewer in number, their modular structure are more likely to indicate the necessity for complex regulatory mechanism for controlling key gene expressions in MPAL. For example, *WT1*(Wilms' tumor 1) is a well-documented transcription factor involved in hematopoiesis and leukemogenesis, while *PADI3* (Peptidyl Arginine Deiminase 3) is less understood in the context in Leukemia. By comparing Figure.1.1. and Figure.1.2. *WT1*-associated networks demonstrate more complex modular structure compared to *PADI3,* correlating with higher gene impact.



Graph.2 *Distribution of the number of each module in each enhancer network*

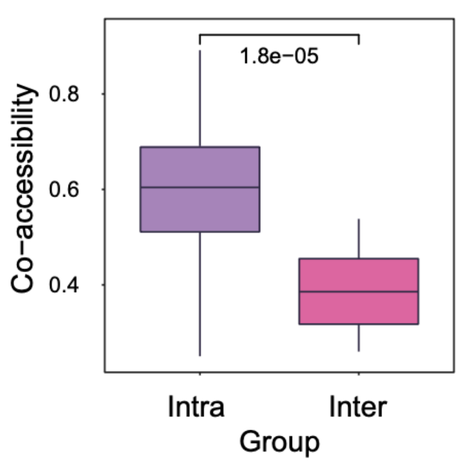


Graph.3 *Distribution of number of enhancers in each modules*

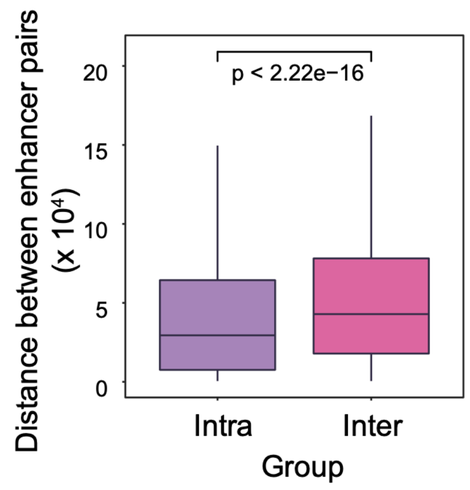


Graph.4 *Enrichment score graph comparing the enrichment score across Modular, NonModular, NonNetwork and genome group.*

In the context of MPAL, most enhancer network are relatively simple, consisting just 2 modules shown in Graph. 2. The steep drop-off in the number of network with more than 2 module suggests that more complex networks are less common. In Graph.4, the modular group exhibits a significantly higher enrichment scores compared to the other groups, suggesting that the regulatory influence of enhancer network on cell identity is predominantly facilitated through modular structure. In MPAL, the disruption of these modular networks can lead to abnormal gene expression, contributing to disease pathogenesis.



Graph.5. *Box plot showing enhancer pair distance within intra and inter groups.*



Graph.6. *Box plot showing enhancer pair co-accessibility within intra and inter groups.*

Graph 5 and Graph 6 compare the distance and co-accessibility between enhancers within Intra and Inter groups respectively, revealing that enhancers within the same group(intra) are not only closer in terms of physical distance but also exhibits higher co-accessibility compared to those between different groups. Higher co-accessibility means that enhancers in intra group are more likely to be simultaneously accessibly and functionally connected. Besides, the closer physical distance of intra group enhancers suggests that they are likely part of the same regulatory module, working together to regulate target gene expression. Vice versa.

CONCLUSION AND IMPLICATION

Although limited in number, targeting enhancer networks with modular structure could offer novel therapeutic avenues for effectively predicting and treating disease genes like in Mixed-phenotype Acute Leukemia (MPAL). Besides, the modular structure exhibit biological significance. Higher co-accessibility(intra group) and closer distance between enhancer pairs result in the formation of cohesive modules due to stronger interactions.

REFERENCE

1. Hong Danni, et al., Complexity of enhancer networks predicts cell identity and disease genes revealed by single-cell multi-omics analysis. Briefings in Bioinformatics, vol. 24, no. 1, 3 Dec. 2022, https://doi.org/10.1093/bib/bbac508.
2. Hongli Lin, et al., "Modular organisation of enhancer network provides transcriptional robustness in mammalian development." (Submitted)