

### 1. Introduction

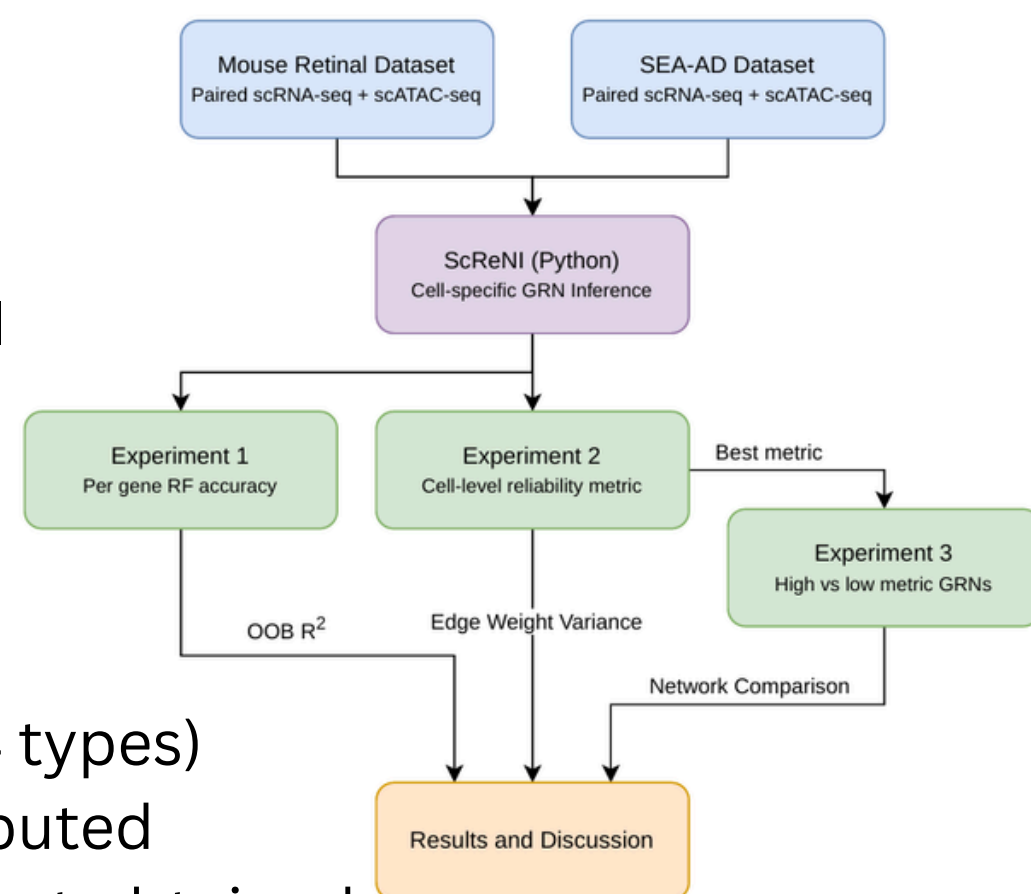
- Problem: The molecular mechanisms of Alzheimer's Disease remain poorly understood. Gene regulatory networks (GRNs) describe how transcription factors control gene expression and understanding cell-type-specific regulatory programs is essential for uncovering disease mechanisms.
- Solution: ScReNI [1] is a recently developed algorithm that infers a unique GRN for every individual cell by using gene expression and chromatin accessibility data.
- Gap: The predictive performance of the random forest model used by ScReNI has not been extensively investigated, raising the question: If the predictive model is weak, can we trust the inferred GRN?

### 2. Research Questions

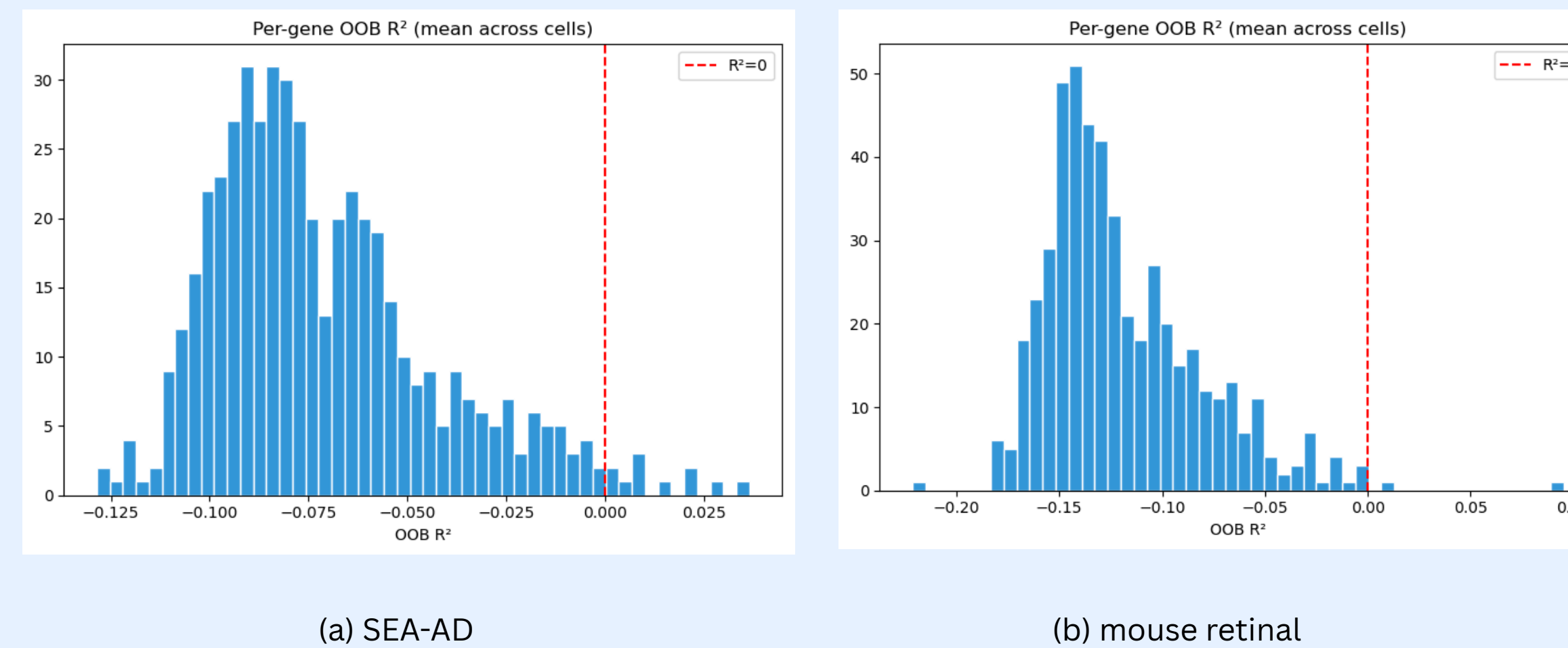
1. Do some genes have higher accuracy than others and how is that linked to the gene's role in the GRN?
2. Can we define a metric that measures the reliability of each cell-specific GRN and how does the metric affect the resulting networks?

### 3. Methodology

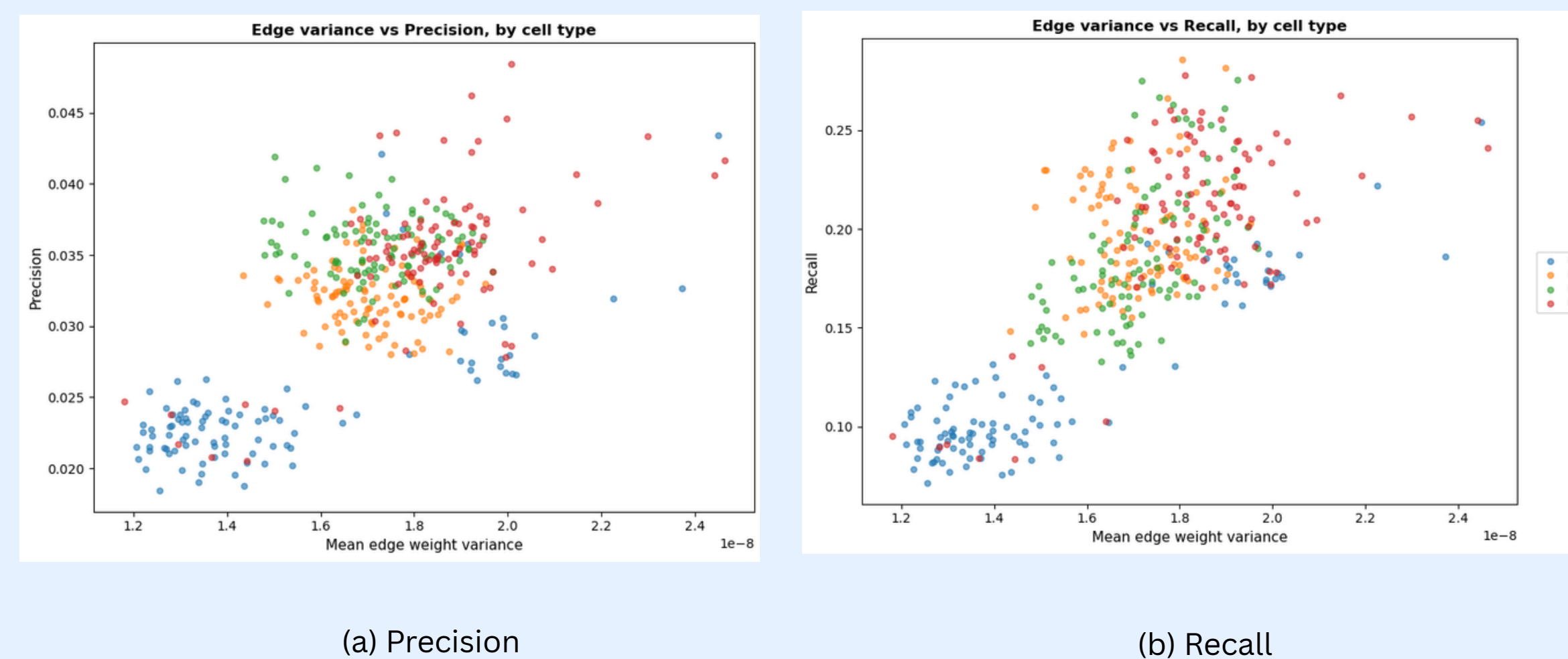
- Python version of ScReNI
- 1200 cell subset from the SEA-AD dataset (50/type × 24 types)
- 400 cell subset of mouse retinal data (100/type × 4 types)
- Precision and recall computed based on a ground truth set obtained from CHIP-Atlas
- Per gene OOB  $R^2$  scores used as random forest accuracy
- Cell reliability metrics computed as mean R per cell and edge weight variance
- Edge weight variance is computed across 5 runs of the network inference



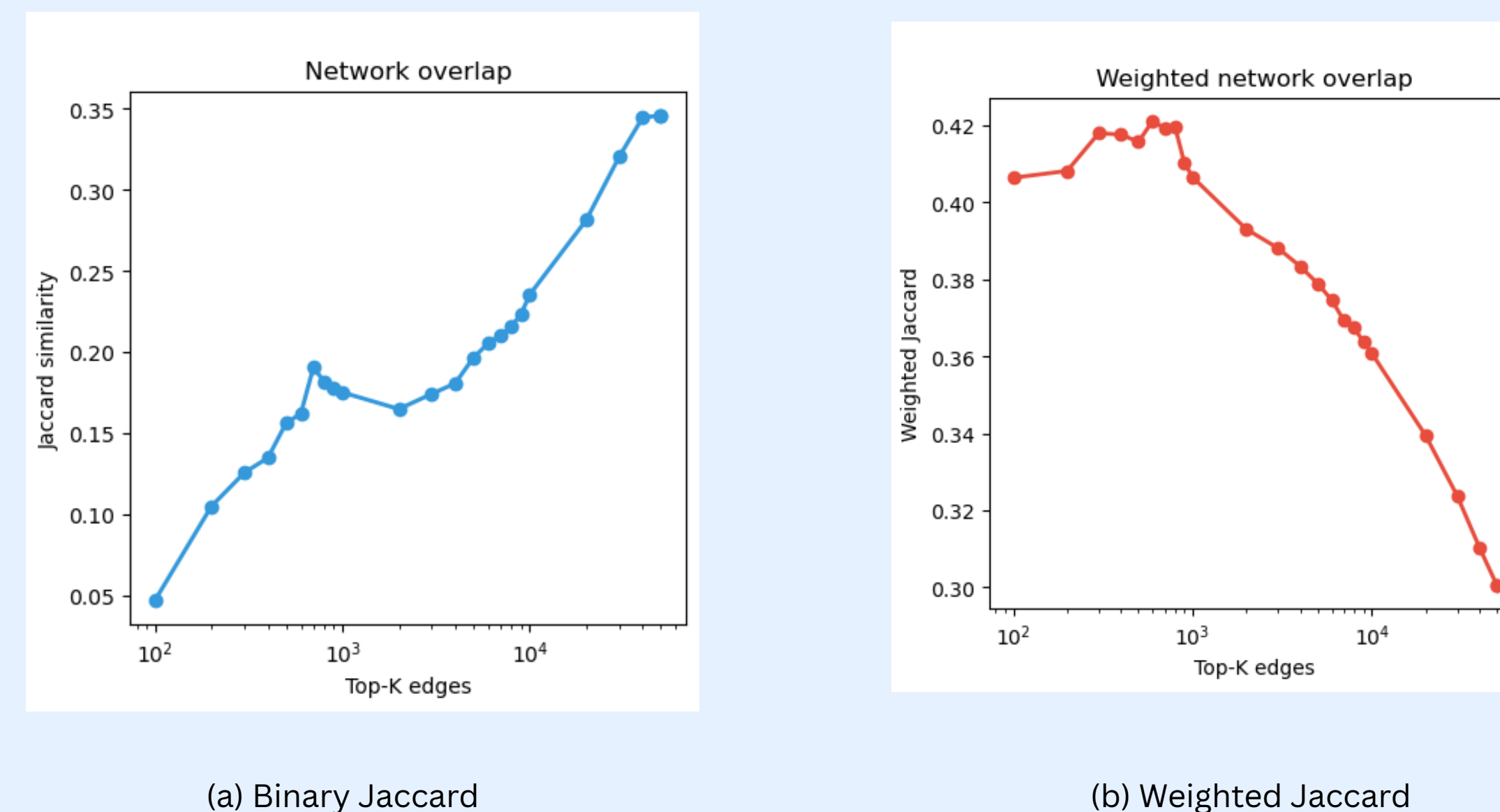
### 4. Results



**Fig. 1:** Histograms showing the distribution of OOB  $R^2$  for each gene for the (a) SEA-AD and (b) mouse retinal datasets.



**Fig. 2:** Correlation of edge weight variance with precision and recall



**Fig. 3:** Network overlap between high and low edge weight variance groups

Property	SEA-AD	Mouse retinal
Mean expression	0.673	0.725
Expression variance	0.592	0.720
Mean outgoing edge weight	0.520	0.437
Mean incoming edge weight	0.442	0.518
In-degree	0.412	0.428
PageRank	0.409	0.393
Betweenness centrality	0.287	0.234
Out-degree	0.032 <sup>†</sup>	-0.234

<sup>†</sup> Not significant ( $p = 0.475$ ).

**Tab. 1:** Spearman correlation between gene-level OOB  $R^2$  and network/expression properties

### 5. Discussion

- Only 11/500 genes have positive OOB  $R^2$  for SEA-AD and only 2/500 for mouse retinal
- Accuracy is low but varies across genes
- Best predicted genes are highly regulated and network-central
- Expression level is the dominant driver and is consistent across datasets
- OOB  $R^2$  has no significant precision signal and recall signal is mostly explained by network density
- Edge weight variance maintains correlation with precision and recall even after controlling for network density and expression, proving to be a promising metric candidate
- Edge weight variance is inconsistent to precision within cell subtypes
- No significant difference in edge consistency between high and low edge variance groups per cell type
- High and low variance cells share network topology, but disagree on edge strength as more edges are considered.
- TFs regulate more genes in high variance cells, but this effect could be a byproduct of expression level

### 6. References

[1] Xu X. et al., "ScReNI: Single-cell Regulatory Network Inference Through Integrating scRNA-seq and scATAC-seq Data", Genomics, Proteomics Bioinformatics, 2025