# Multiscale Transcriptomic Architecture of the ALS Motor Cortex

Aadya Goel

Mentors: Prof. Gil Alterovitz and Dr. Shaojun Pei
Harvard Medical School / MIT PRIMES Computational Biology

MIT PRIMES Research Conference 2026

1 Introduction

2 Subtype Discovery

3 Comparision to Published Frameworks

4 Landscape and Network Architecture

5 Clinical and Genetic Associations

6 Conclusion

# Amyotrophic Lateral Sclerosis (ALS)

- Progressive neurodegenerative disease
- Defined by degeneration of upper and lower motor neurons

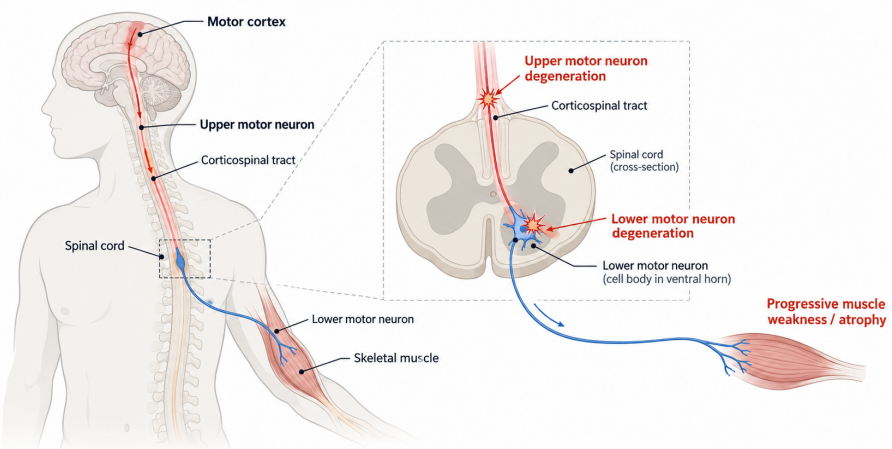
# Amyotrophic Lateral Sclerosis (ALS)

- Progressive neurodegenerative disease
- Defined by degeneration of upper and lower motor neurons
- Progressive motor-neuron loss leads to muscle weakness, paralysis, and ultimately respiratory failure.

# Amyotrophic Lateral Sclerosis (ALS)

- Progressive neurodegenerative disease
- Defined by degeneration of upper and lower motor neurons
- Progressive motor-neuron loss leads to muscle weakness, paralysis, and ultimately respiratory failure.
- Patients can severely vary in onset, progression, and cognitive/behavorial symptoms

# Amyotrophic Lateral Sclerosis (ALS)



# Defining ALS Molecular Subtypes

## ALS is a Heterogeneous Disease

Patients with ALS differ in their molecular and clinical features, motivating efforts to identify biologically meaningful subtypes. Molecular subtyping can help reveal distinct biological states within the disease.

# Defining ALS Molecular Subtypes

## ALS is a Heterogeneous Disease

Patients with ALS differ in their molecular and clinical features, motivating efforts to identify biologically meaningful subtypes. Molecular subtyping can help reveal distinct biological states within the disease.

## Existing subtype studies have two key limitations

1. Pooling multiple CNS regions can confound subtype structure: Subtyping across brain and spinal cord samples may reflect anatomical origin rather than true disease biology

# Defining ALS Molecular Subtypes

## ALS is a Heterogeneous Disease

Patients with ALS differ in their molecular and clinical features, motivating efforts to identify biologically meaningful subtypes. Molecular subtyping can help reveal distinct biological states within the disease.

## Existing subtype studies have two key limitations

1. Pooling multiple CNS regions can confound subtype structure: Subtyping across brain and spinal cord samples may reflect anatomical origin rather than true disease biology
2. Rigid subtype labels can oversimplify the biology: One label per sample may hide how subtypes relate to each other

# Research Focus and Central Question

## Why the Motor Cortex?

We focus on the motor cortex because it has been proposed that ALS progression originates in the motor cortex.

# Research Focus and Central Question

## Why the Motor Cortex?

We focus on the motor cortex because it has been proposed that ALS progression originates in the motor cortex.

## Research Question

How are ALS motor cortex subtypes organized? Are they isolated categories? Are they overlapping states? Or is it something more complex?

1 Introduction

2 Subtype Discovery

3 Comparision to Published Frameworks

4 Landscape and Network Architecture

5 Clinical and Genetic Associations

6 Conclusion

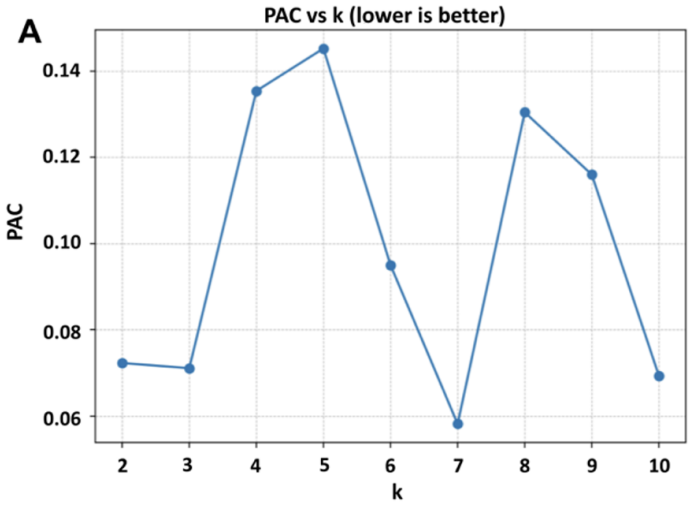
# How Many Subtypes?

- Consensus clustering: re-cluster random subsets of the samples many times
- Samples that keep landing together truly belong together

# How Many Subtypes?

- Consensus clustering: re-cluster random subsets of the samples many times
- Samples that keep landing together truly belong together
- PAC measures how often the clustering was ambiguous, so lower is better

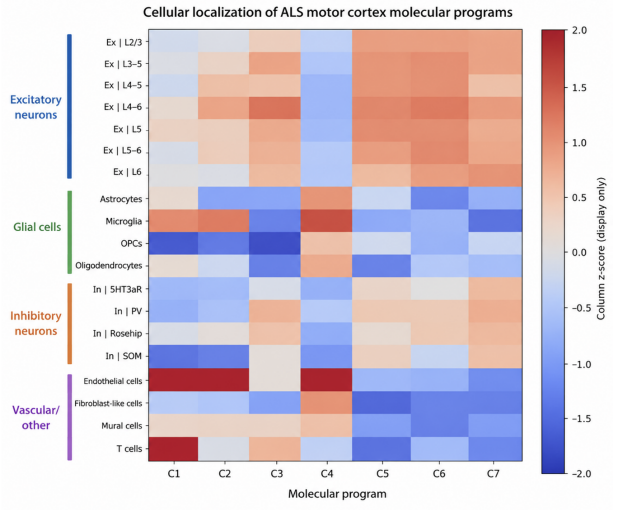
# How Many Subtypes?



# Seven Subtypes, Seven Biological Programs

- C1: inflammatory lipid and membrane stress
- C2: translation and RNA-processing stress
- C3: cell damage and apoptosis (cell death)
- C4: glial and iron-lipid remodeling (glia are support cells in the brain)
- C5: transporter and lysosomal remodeling
- C6: mitochondrial and oxidative metabolism
- C7: neuronal and synaptic signaling

# Cell Type Matters, But It's Not the Whole Story



# Do the Subtypes Hold Up in New Patients?

- Split donors into 80% training and 20% test sets
- Train a centroid-based subtype classifier using pathway profiles from the training donors

# Do the Subtypes Hold Up in New Patients?

- Split donors into 80% training and 20% test sets
- Train a centroid-based subtype classifier using pathway profiles from the training donors
- Use the trained model to classify samples from unseen test donors

# Do the Subtypes Hold Up in New Patients?

- Split donors into 80% training and 20% test sets
- Train a centroid-based subtype classifier using pathway profiles from the training donors
- Use the trained model to classify samples from unseen test donors
- Repeat the train/test split 20 times to test robustness

# Do the Subtypes Hold Up in New Patients?

- Split donors into 80% training and 20% test sets
- Train a centroid-based subtype classifier using pathway profiles from the training donors
- Use the trained model to classify samples from unseen test donors
- Repeat the train/test split 20 times to test robustness
- Balanced accuracy $= 0.939$
- Correct subtype in the top two guesses for 98.5% of hidden samples

1 Introduction

2 Subtype Discovery

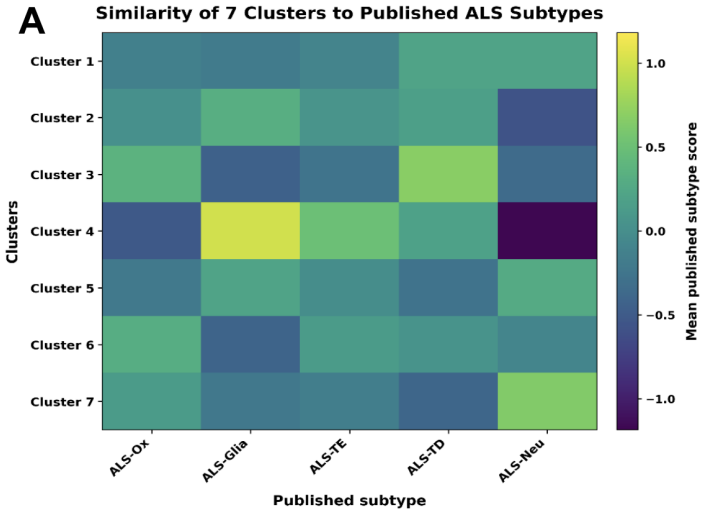
3 Comparision to Published Frameworks

4 Landscape and Network Architecture

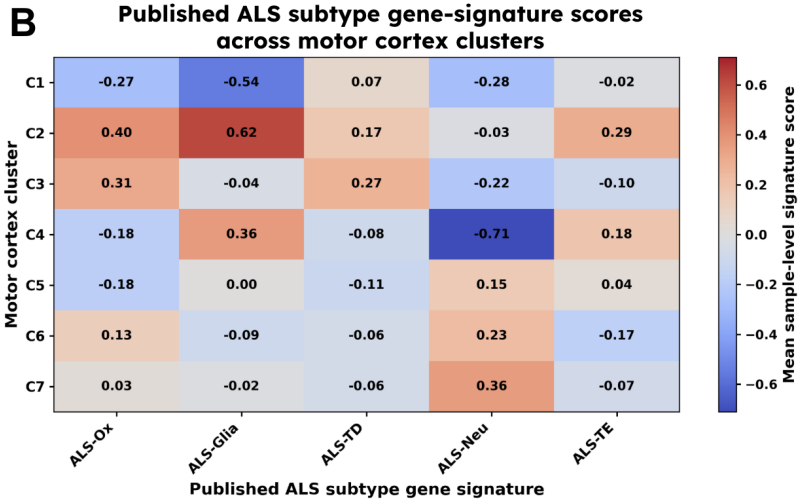
5 Clinical and Genetic Associations

6 Conclusion

# Comparison to Published ALS Subtypes Using Pathway Activity



# Comparison to Published ALS Subtypes Using Gene Signatures



# Takeaways

- Our subtypes recover major published programs but do not match them one-to-one

# Takeaways

- Our subtypes recover major published programs but do not match them one-to-one
- A single published subtype can split into several finer motor cortex states

1 Introduction

2 Subtype Discovery

3 Comparision to Published Frameworks

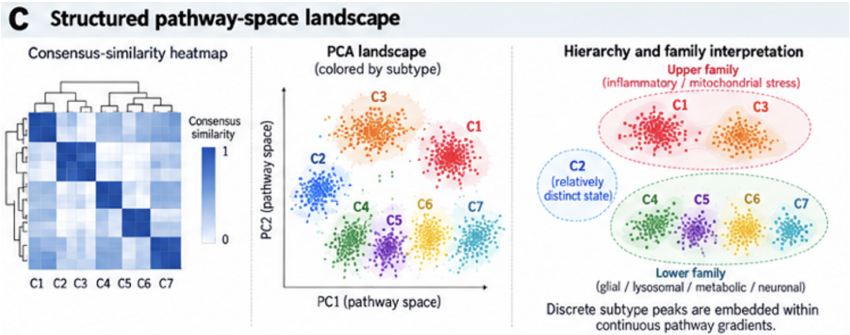
4 Landscape and Network Architecture

5 Clinical and Genetic Associations

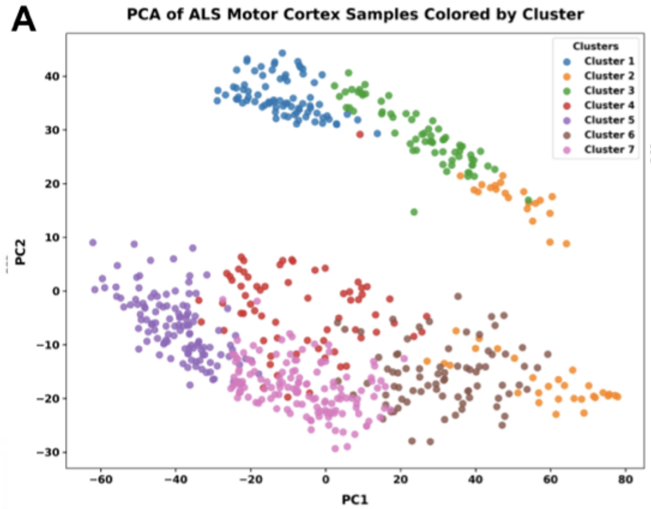
6 Conclusion

# Islands or a Landscape?

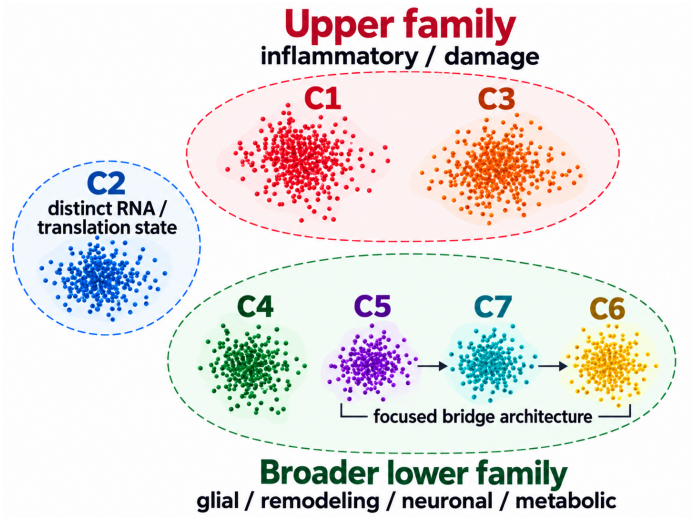
- Clusters tell us which group a sample belongs to, not how the groups relate
- Are the subtypes separate islands, or points on one continuous landscape?



# A PCA Map of the Subtypes



# A Multiscale Architecture



# Do Genes Agree with the Map?

- Build co-expression networks: link genes whose activity rises and falls together
- A module is a tight group of linked genes with a shared function

# Do Genes Agree with the Map?

- Build co-expression networks: link genes whose activity rises and falls together
- A module is a tight group of linked genes with a shared function
- In the lower family, modules form three poles: C5 (transport/remodeling), C6 (metabolic), C7 (neuronal)
- C7 links to both C5 and C6, but C5 and C6 do not link directly

# Important Genes Across Multiple Subtypes

- We next looked for genes that appeared to play important roles across more than one subtype or subtype relationship

# Important Genes Across Multiple Subtypes

- We next looked for genes that appeared to play important roles across more than one subtype or subtype relationship
- Six genes: PRKACB, RPL3, COX4I1, UBQLN1, G3BP2, HK1
- They connect to protein-making, mitochondria, signaling, protein quality control, and metabolism

1 Introduction

2 Subtype Discovery

3 Comparision to Published Frameworks

4 Landscape and Network Architecture

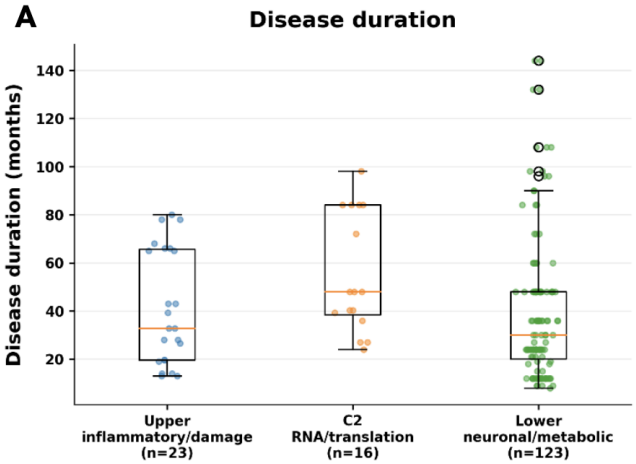
5 Clinical and Genetic Associations

6 Conclusion

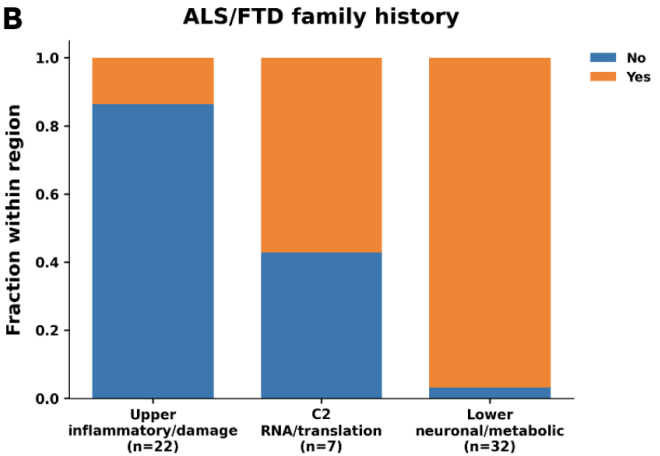
# What Biological Features Define the Subtype Structure?

- Is the subtype structure associated with disease progression or genetic background?
- Compare the upper (C1/C3), C2, and lower (C4–C7) regions against clinical and genetic variables

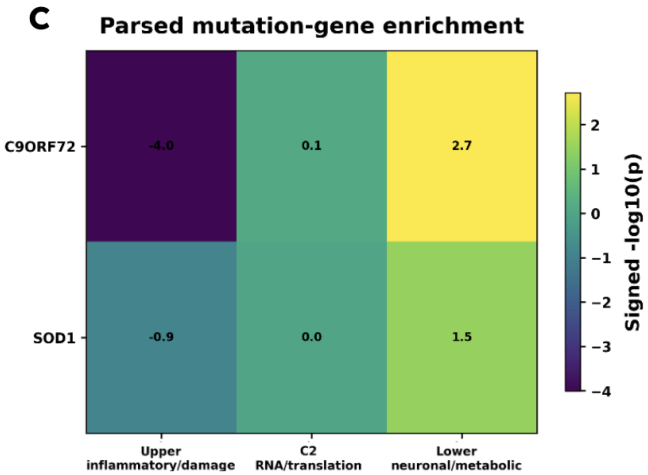
# Disease Duration Does Not Significantly Differ



# Family History Differs Between Groups



# ALS-Associated Mutations Are Enriched in the Lower Landscape



1 Introduction

2 Subtype Discovery

3 Comparision to Published Frameworks

4 Landscape and Network Architecture

5 Clinical and Genetic Associations

6 Conclusion

# Conclusion

- ALS motor cortex heterogeneity is a multiscale transcriptomic architecture of seven subtypes
- Gene networks support it and nominate six candidate genes
- It refines published ALS subtypes and tracks genetic background rather than disease duration

# Future Works

- Validate in independent cohorts (postmortem data are a single late-stage snapshot)
- Test candidate genes for biomarkers and drug targets

# Acknowledgments


- Thank you to the MIT PRIMES program for making this project possible.
- Thank you to my mentors, Prof. Gil Alterovitz and Dr. Shaojun Pei, for guiding my research, and supporting me in every way.
- Thank you to Target ALS Postmortem Tissue and Genomics Core for providing me with the data used in this project.

# References


- Mejzini et al., “ALS Genetics, Mechanisms, and Therapeutics: Where Are We Now?” Front Neurosci 2019
- Jammal et al., “Classification of ALS molecular subtypes: a literature review on machine learning applications and their clinical value,” BMC Med 2026
- Tam et al., “Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia,” Cell Rep 2019
- Monti et al., “Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data,” Machine Learning 2003
- Ziff et al., “Integrated transcriptome landscape of ALS identifies genome instability linked to TDP-43 pathology,” Nat Commun 2023

# Appendix: Each Pathway Peaks in a Subtype & Fades Gradually Into Neighbors

