

Project Proposal

Proteome-wide Screen for RNA-dependent Proteins

Team 3: HeLa Cells Synchronized in Interphase

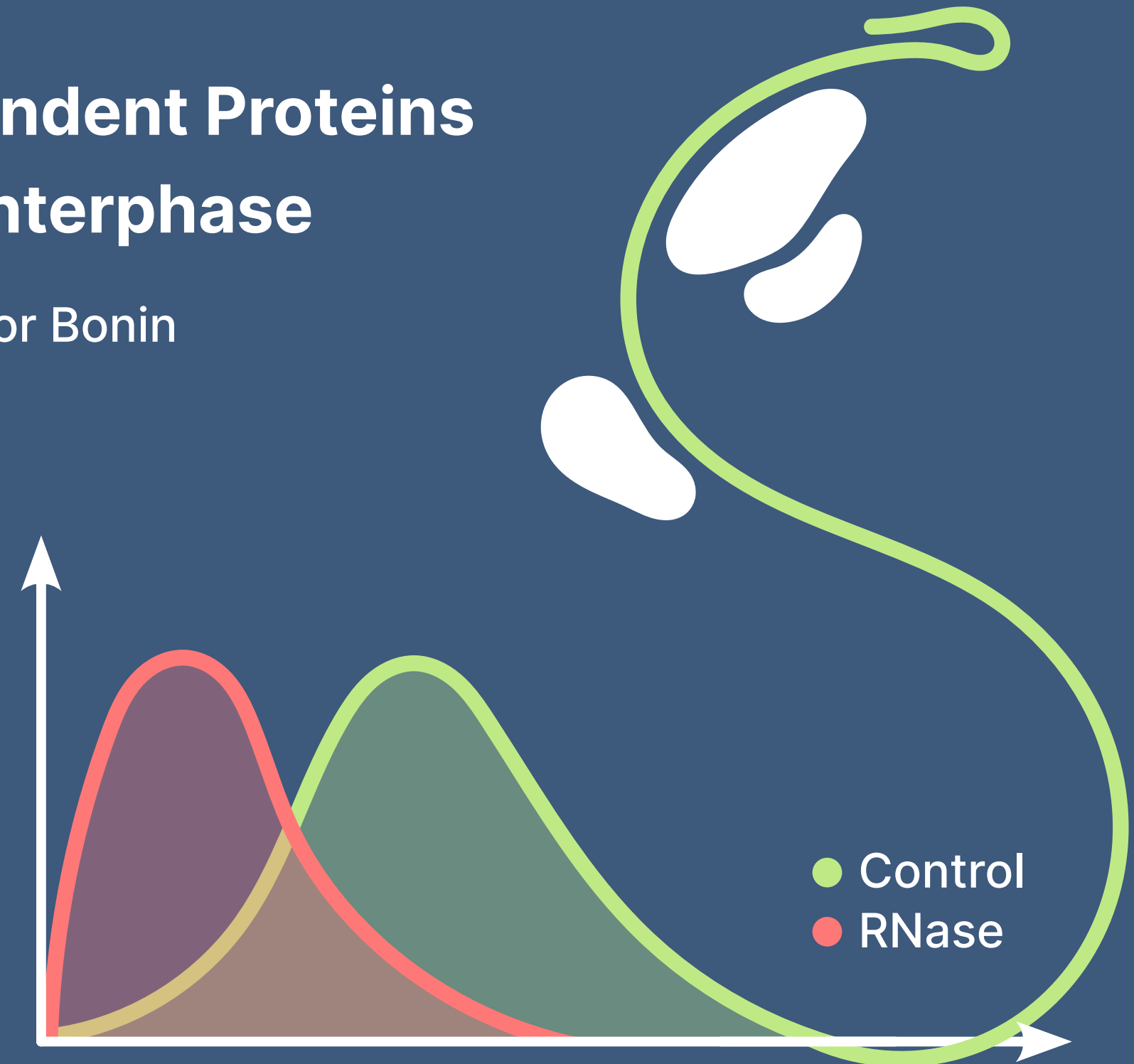
Hannah Brehm, Johann Blakytty, Kira Hoffmann, Viktor Bonin

Data Analysis Project Summer Term 2023

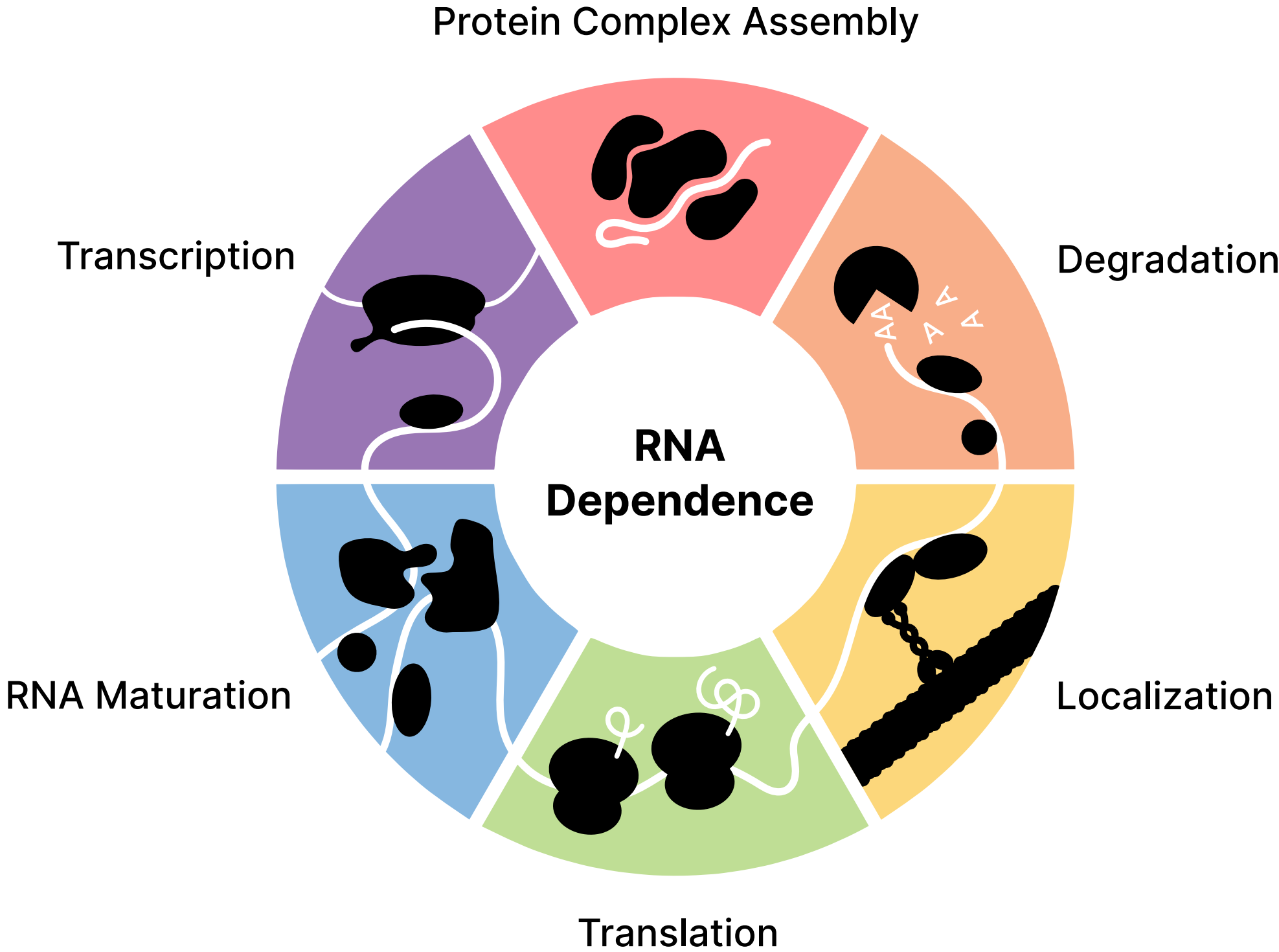
Supervisor: Dr. Maiwen Caudron-Herger

Tutor: Fabio Rauscher

Wednesday, 17.05.2023

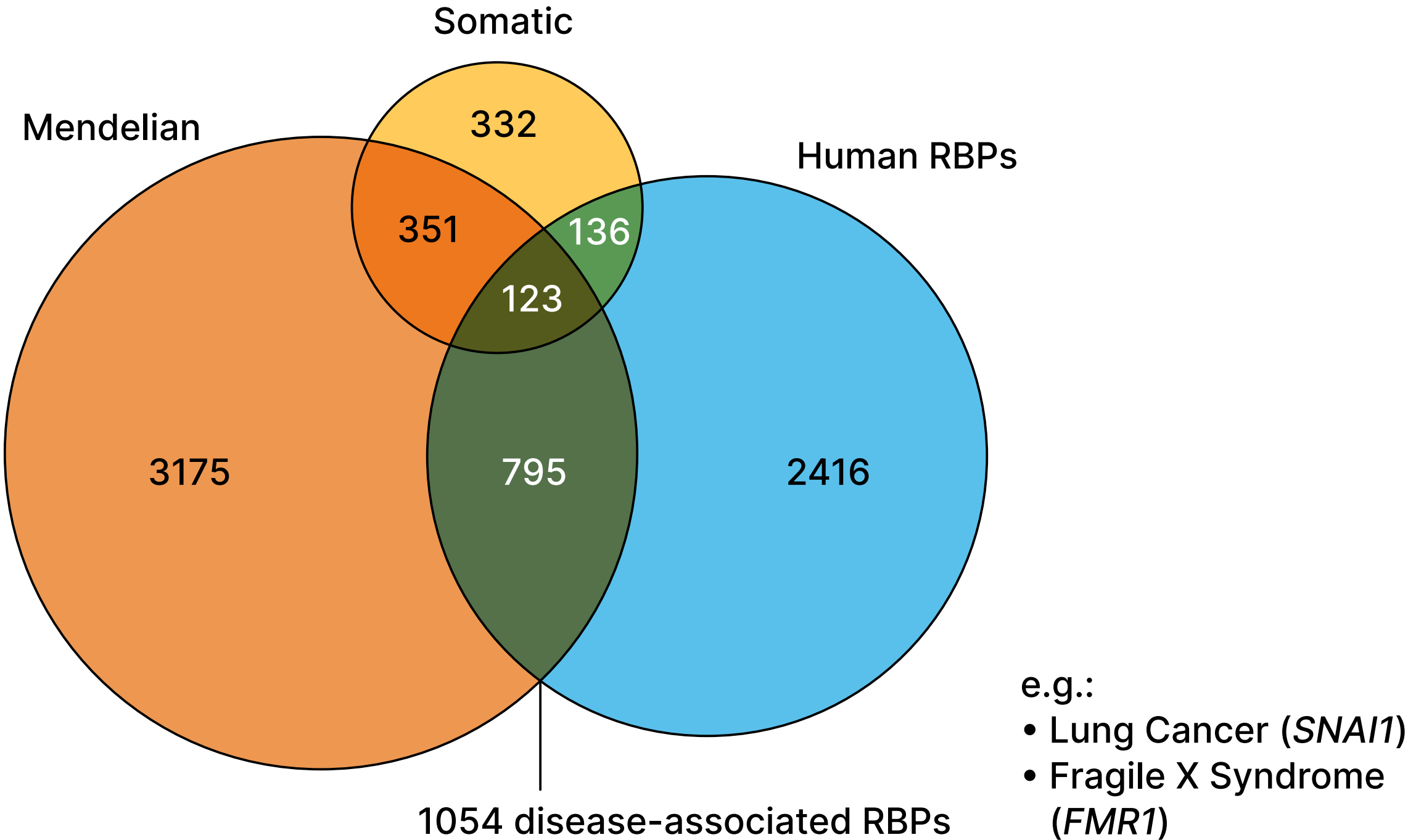


Functions of RNA-dependent Proteins



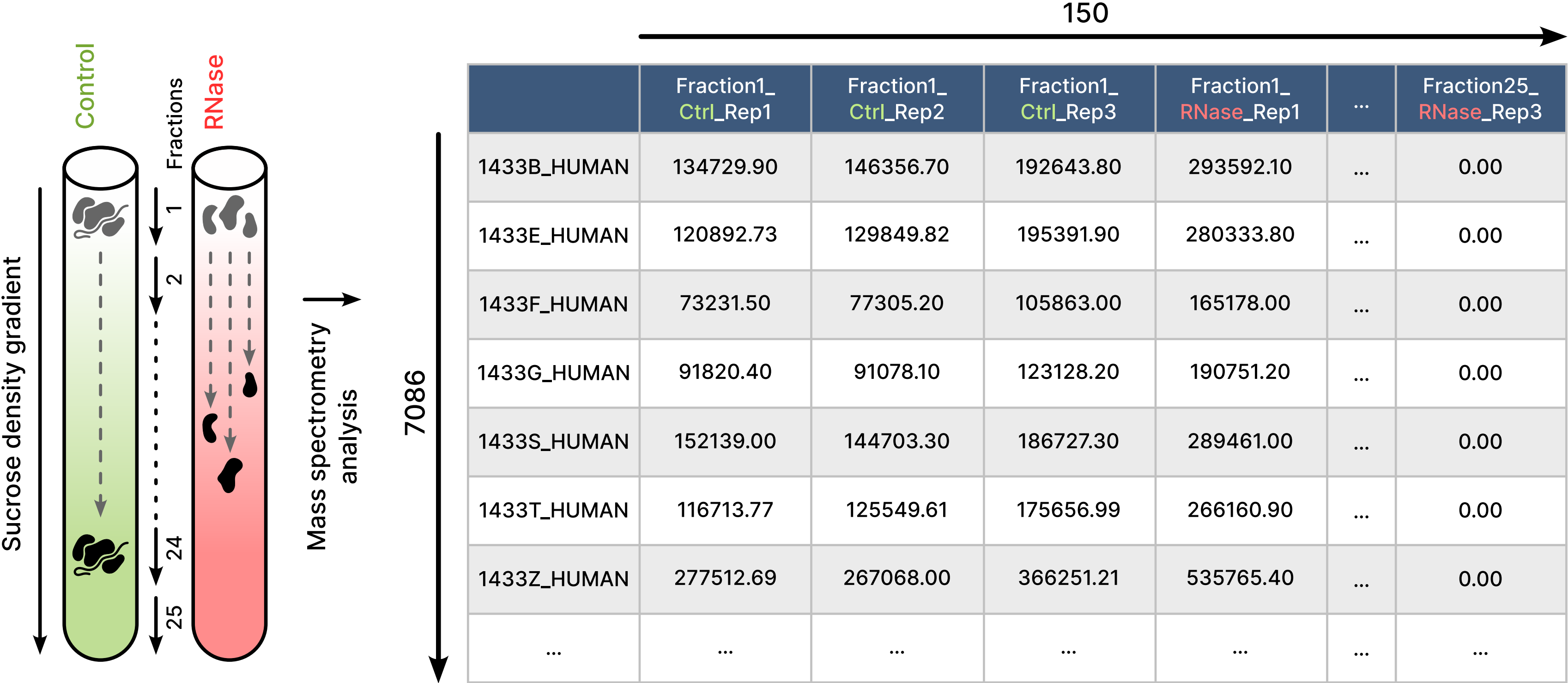
Adapted from: Gebauer, F., Schwarzl, T., Valcárcel, J., and Hentze, M.W. (2021). RNA-binding proteins in human genetic disease. Nat. Rev. Genet. 22, 185-198.

RNA-binding Proteins in Human Diseases



Gebauer, F., Schwarzl, T., Valcárcel, J., and Hentze, M.W. (2021). RNA-binding proteins in human genetic disease. Nat. Rev. Genet. 22, 185-198.

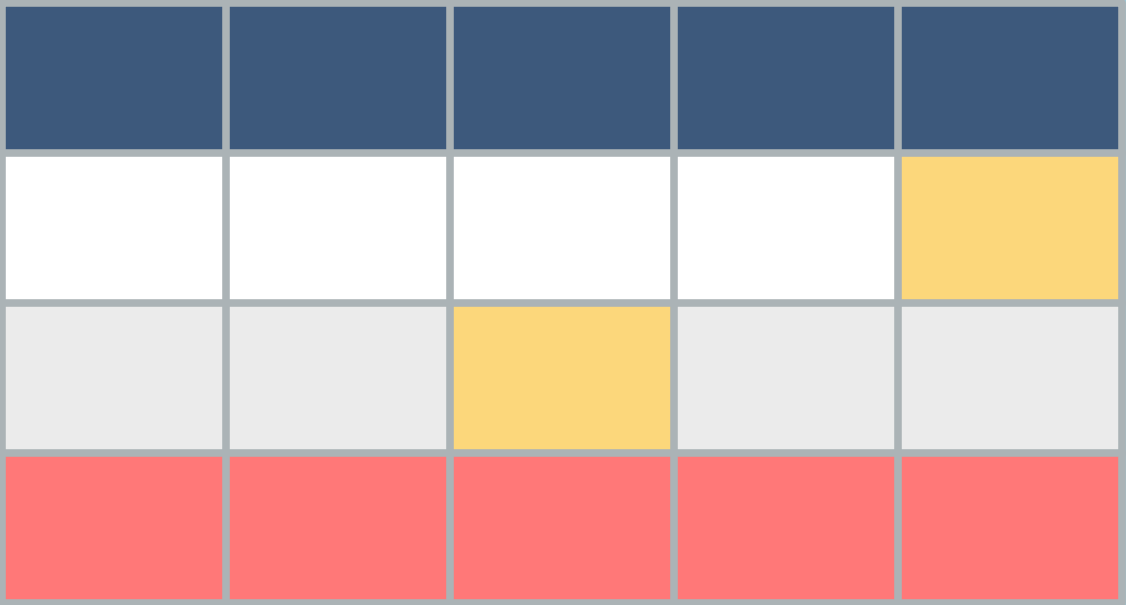
Dataset for HeLa Cells Synchronized in Interphase



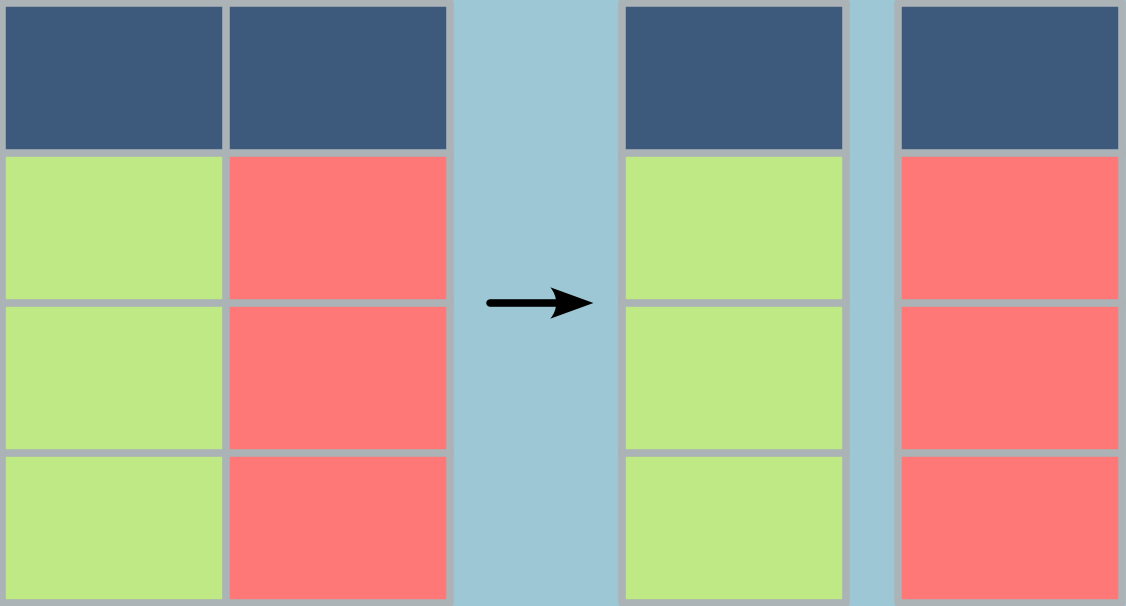
Caudron-Herger, M., et al. (2019). R-DeeP: Proteome-wide and Quantitative Identification of RNA-Dependent Proteins by Density Gradient Ultracentrifugation. Mol. Cell 75, 184-199.

Step 1: Data Cleanup and Normalization

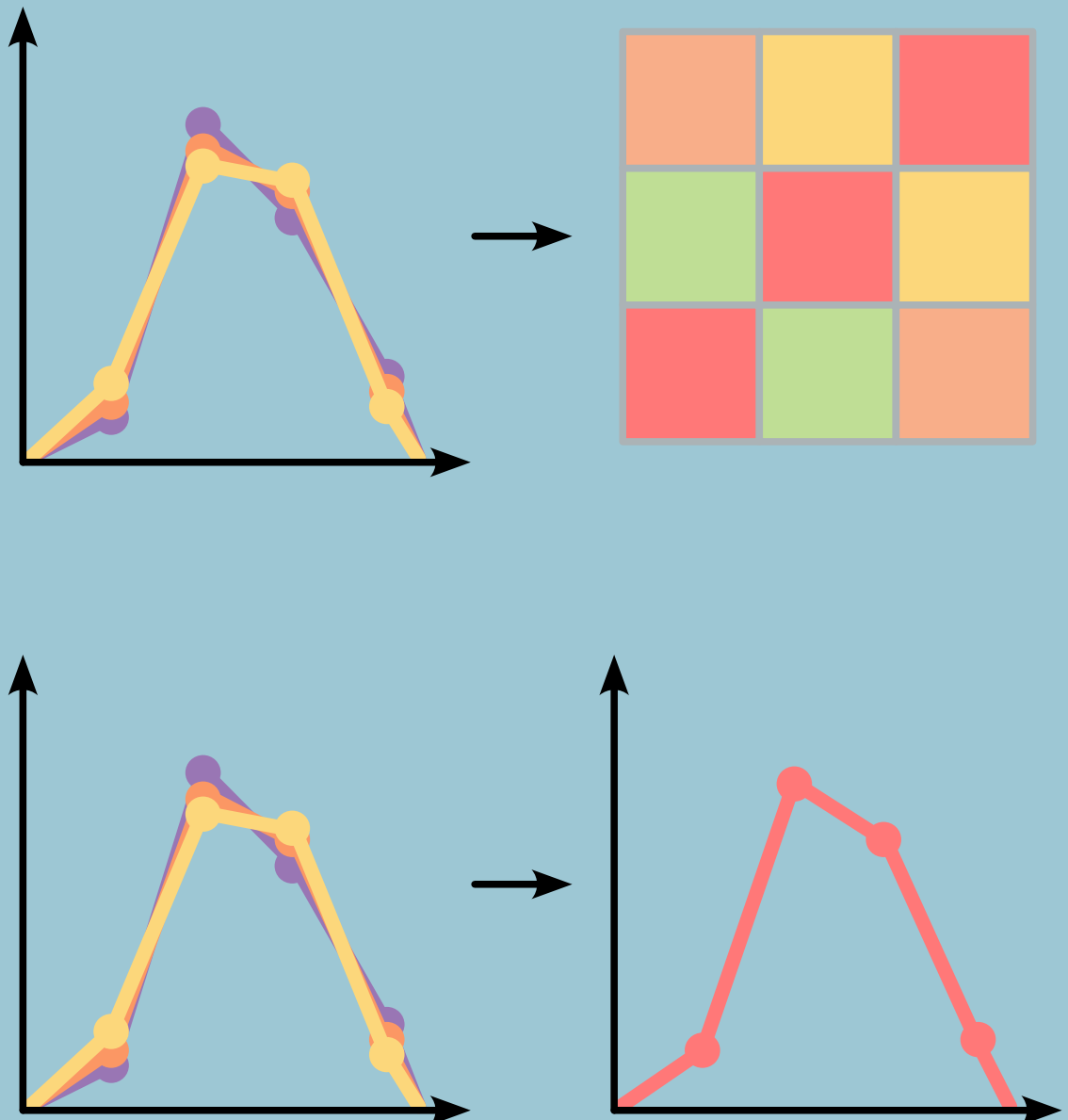
1.1 NA Values and Zero Rows



1.2 Split Dataset into Control and RNase



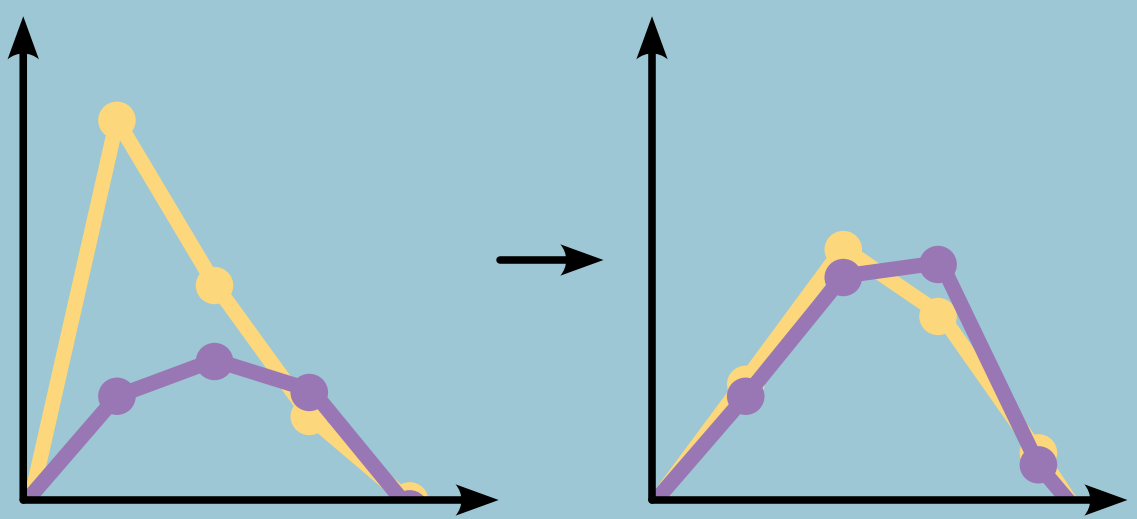
1.3 Pearson and Mean of Replicates



1.4 Normalization

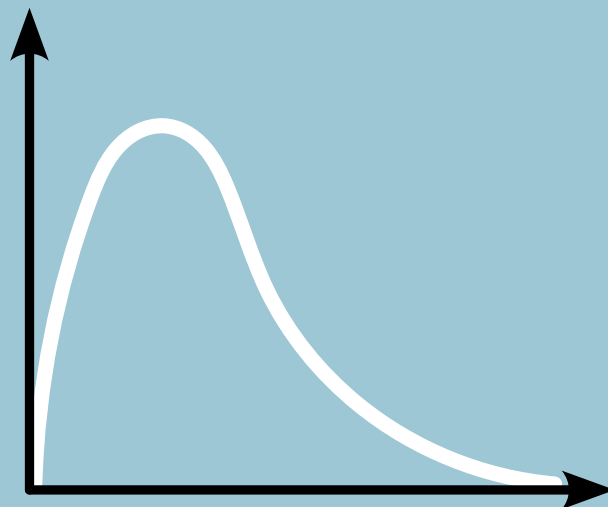
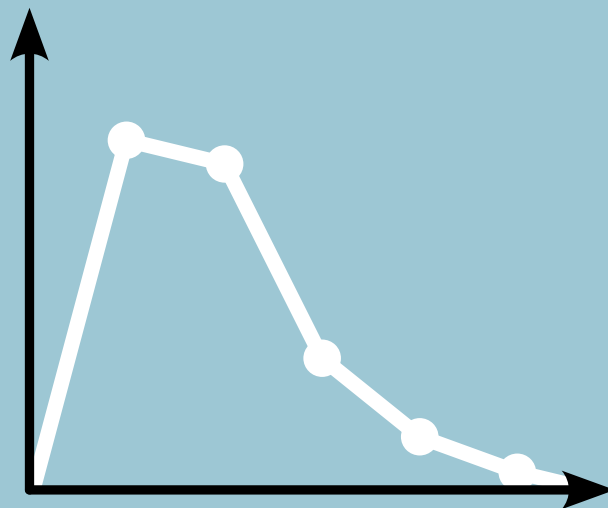
				Σ
				100
				100
				100

1.5 Z-Scaling

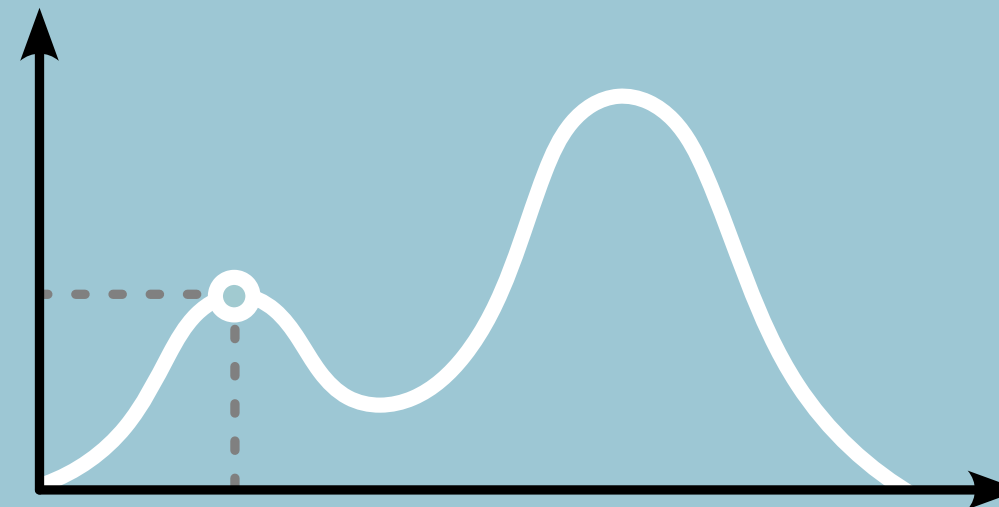
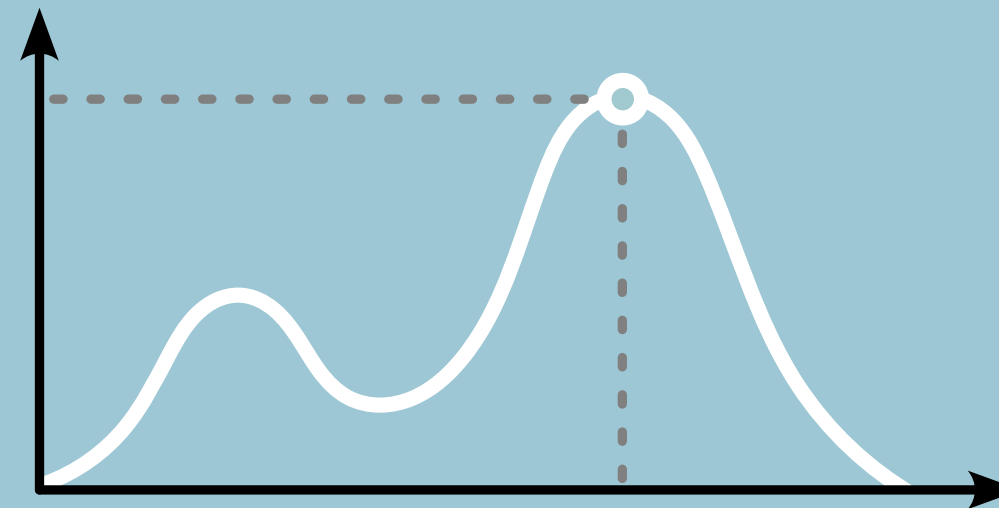


Step 2: Finding Maxima and Gaussian Fitting

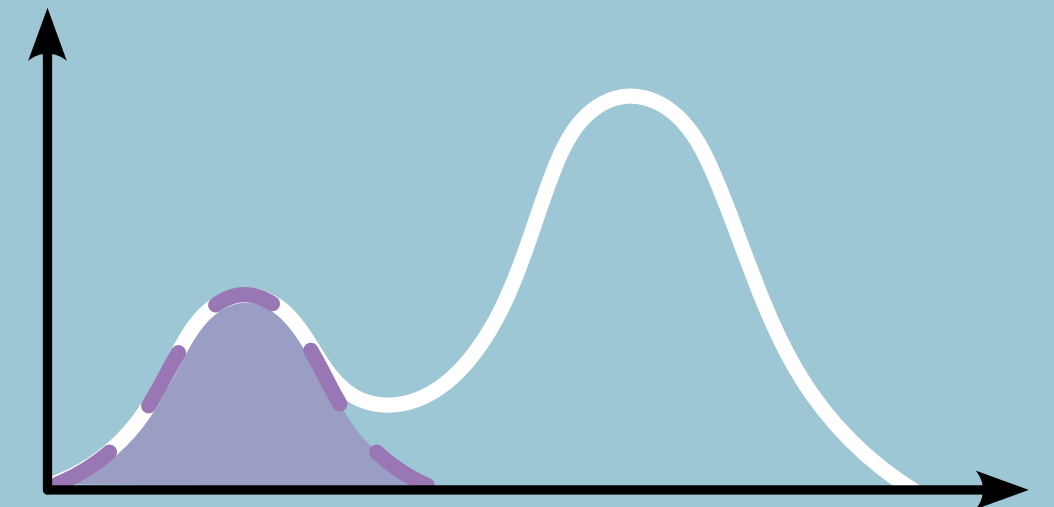
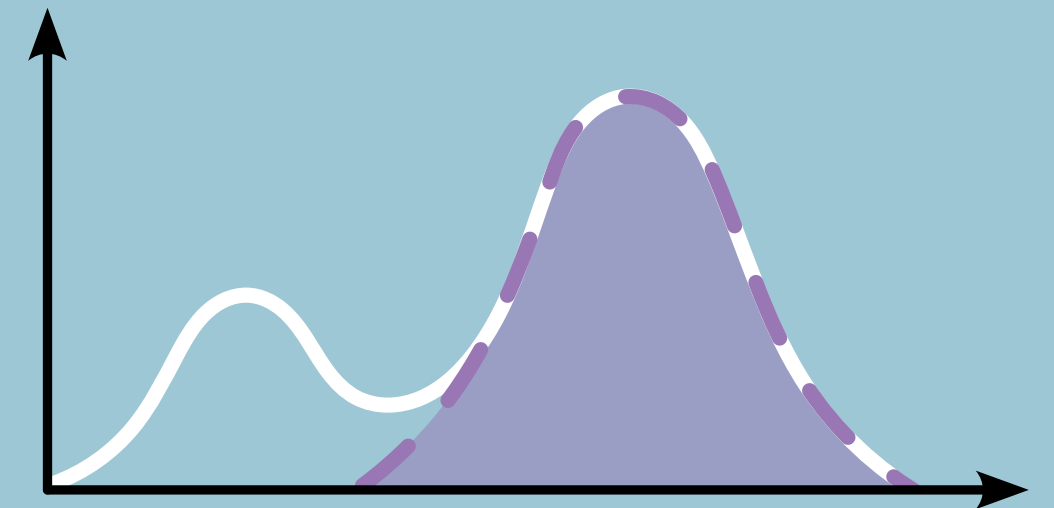
2.1 Smoothing



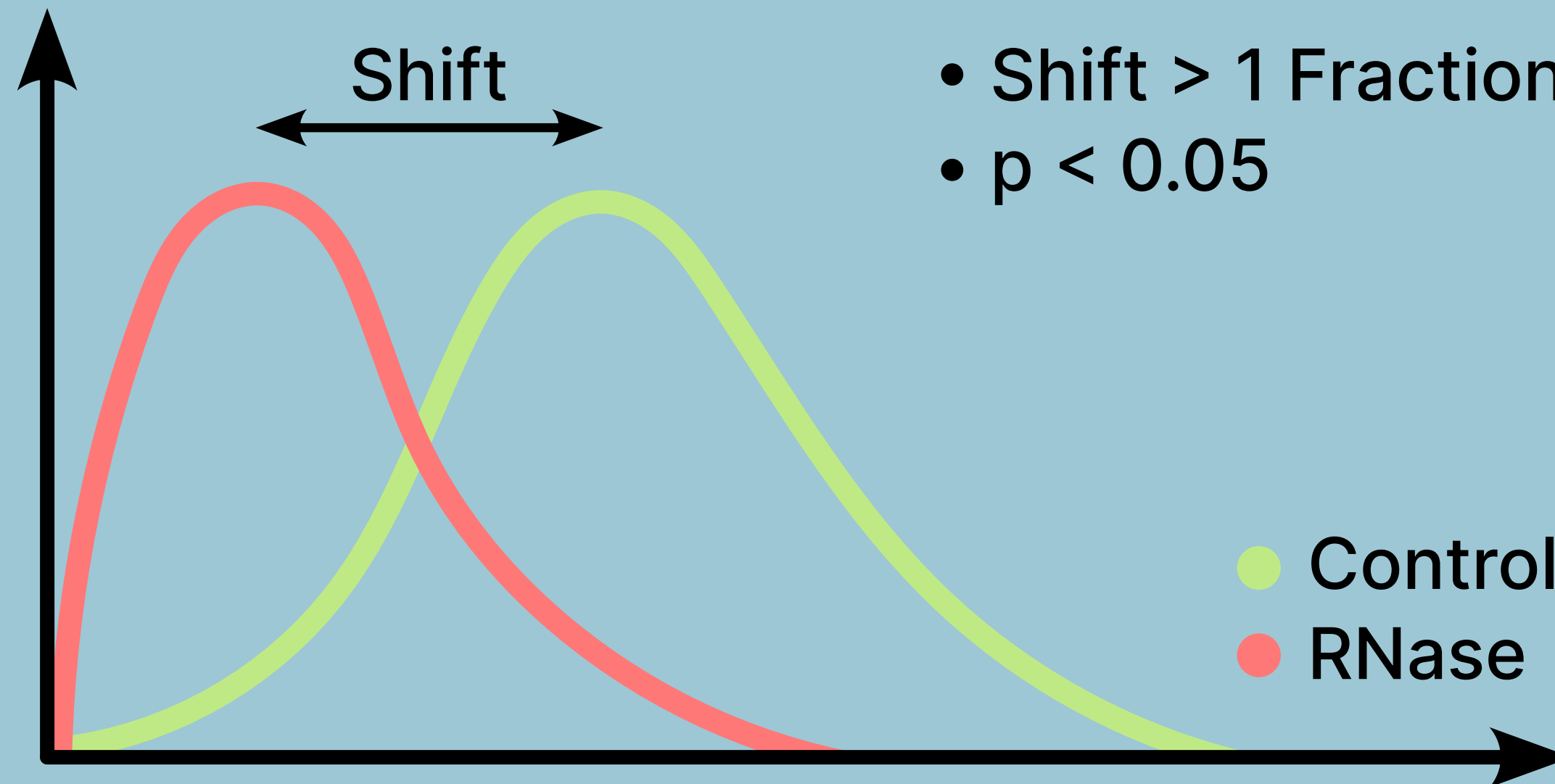
2.2 Global and Local Maxima



2.3 Gaussian Fitting



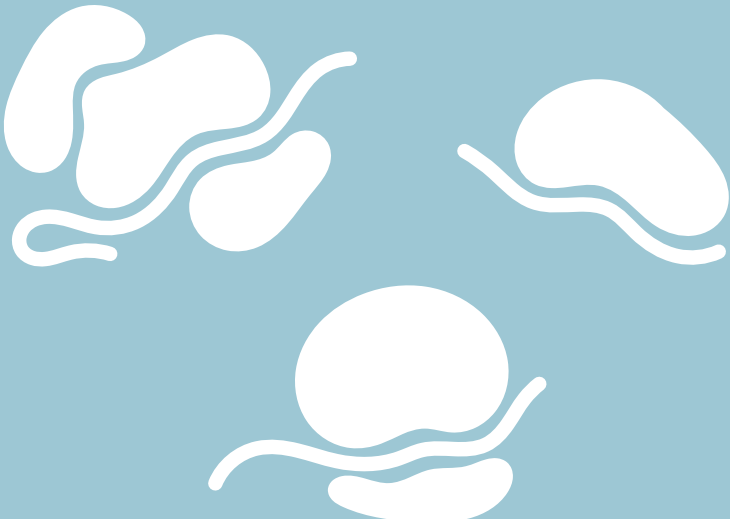
Step 3: Defining Selection Criteria



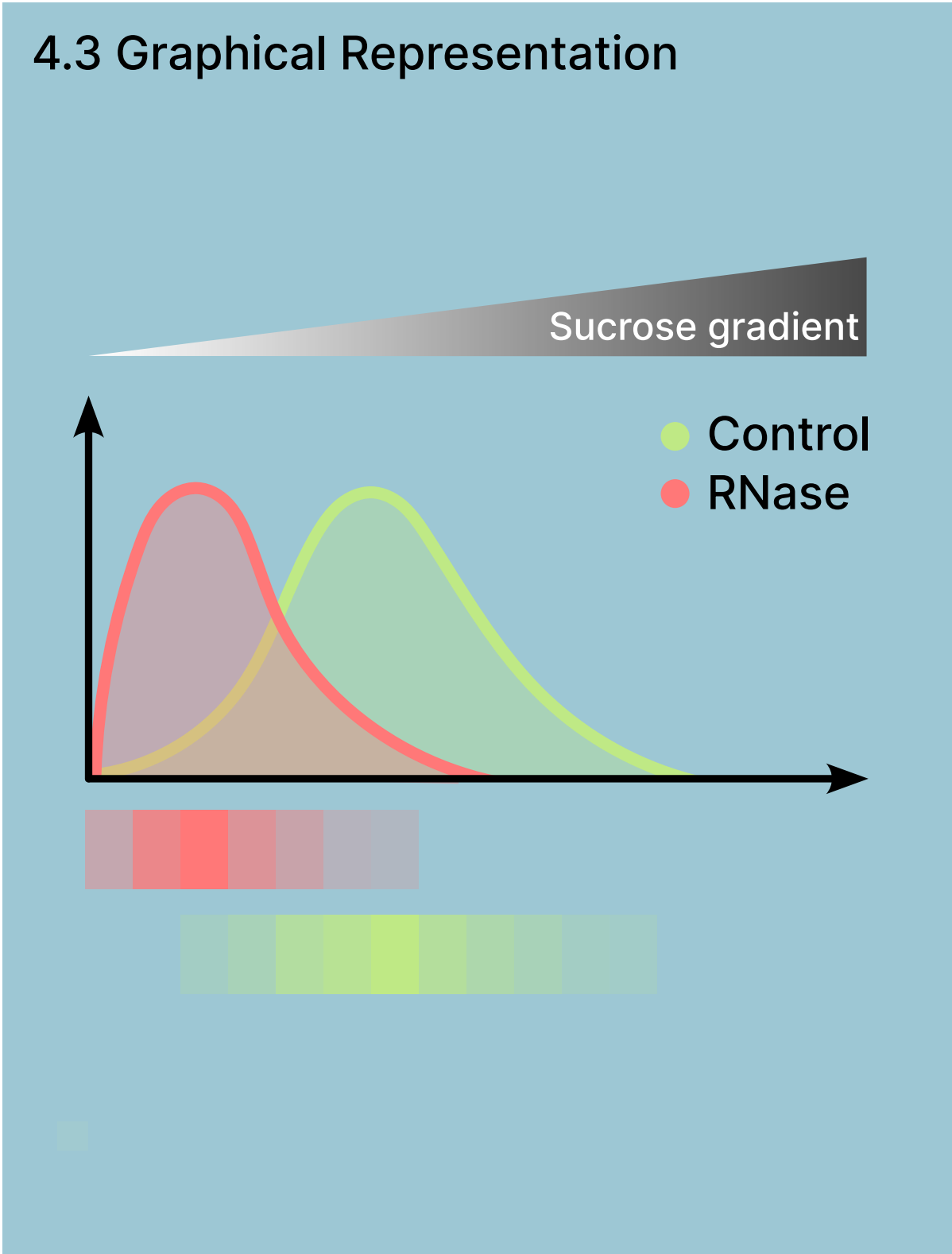
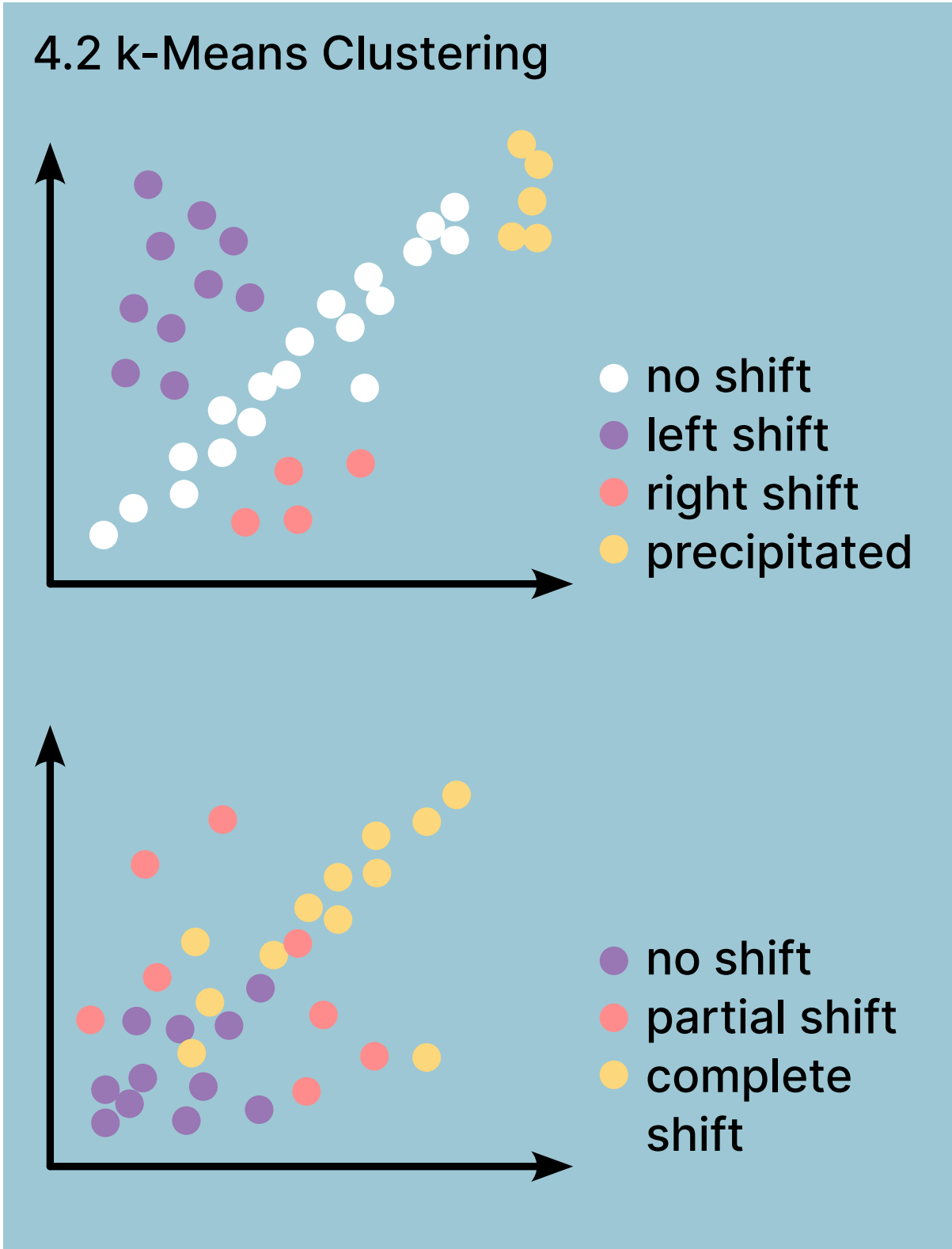
Step 4: Identification of RNA-dependent Proteins

4.1 Application of Our Selection Criteria

↓

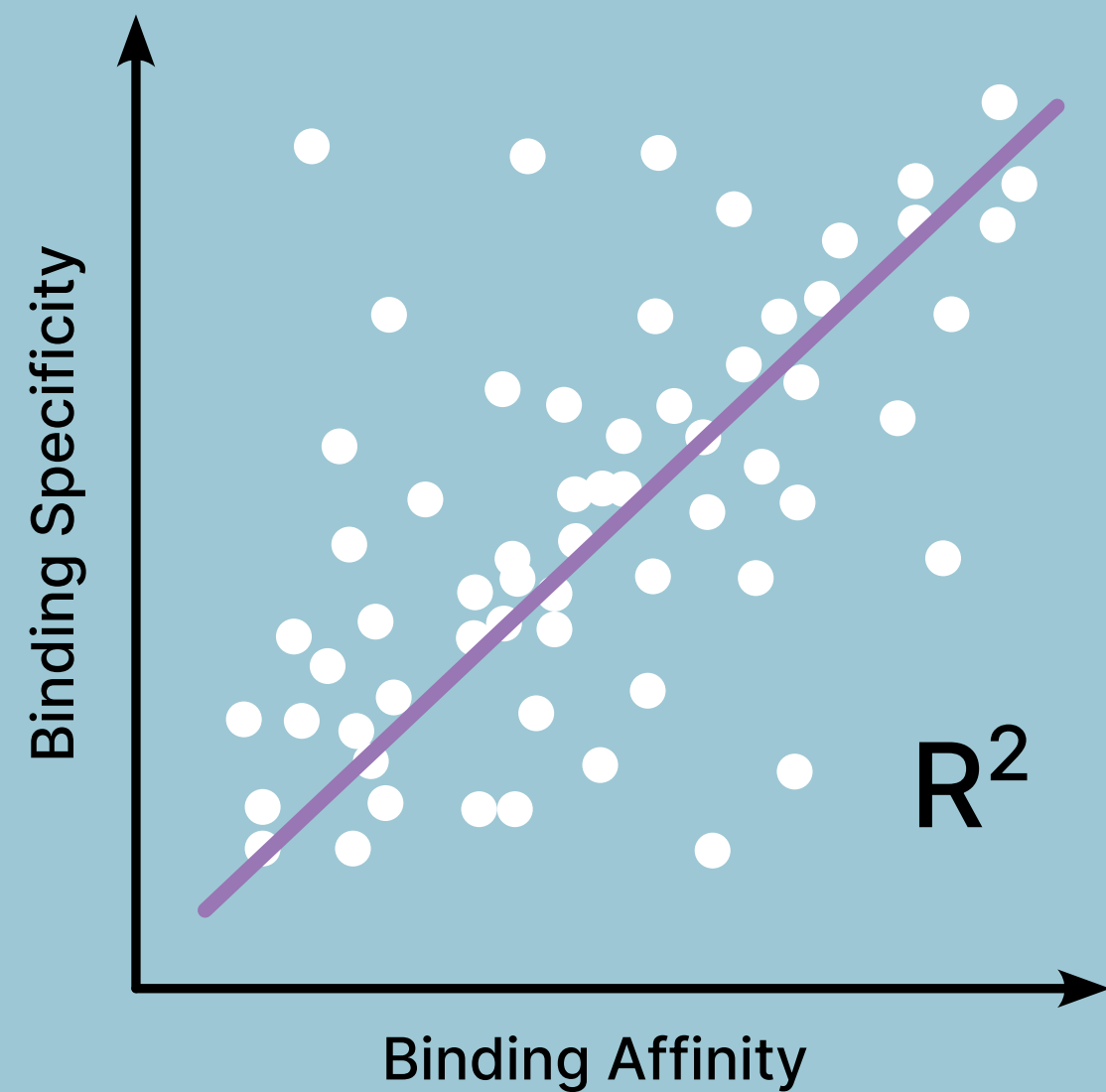


The diagram shows three protein-RNA complexes. The top two are composed of multiple subunits, while the bottom one is a single subunit with a large, rounded domain.

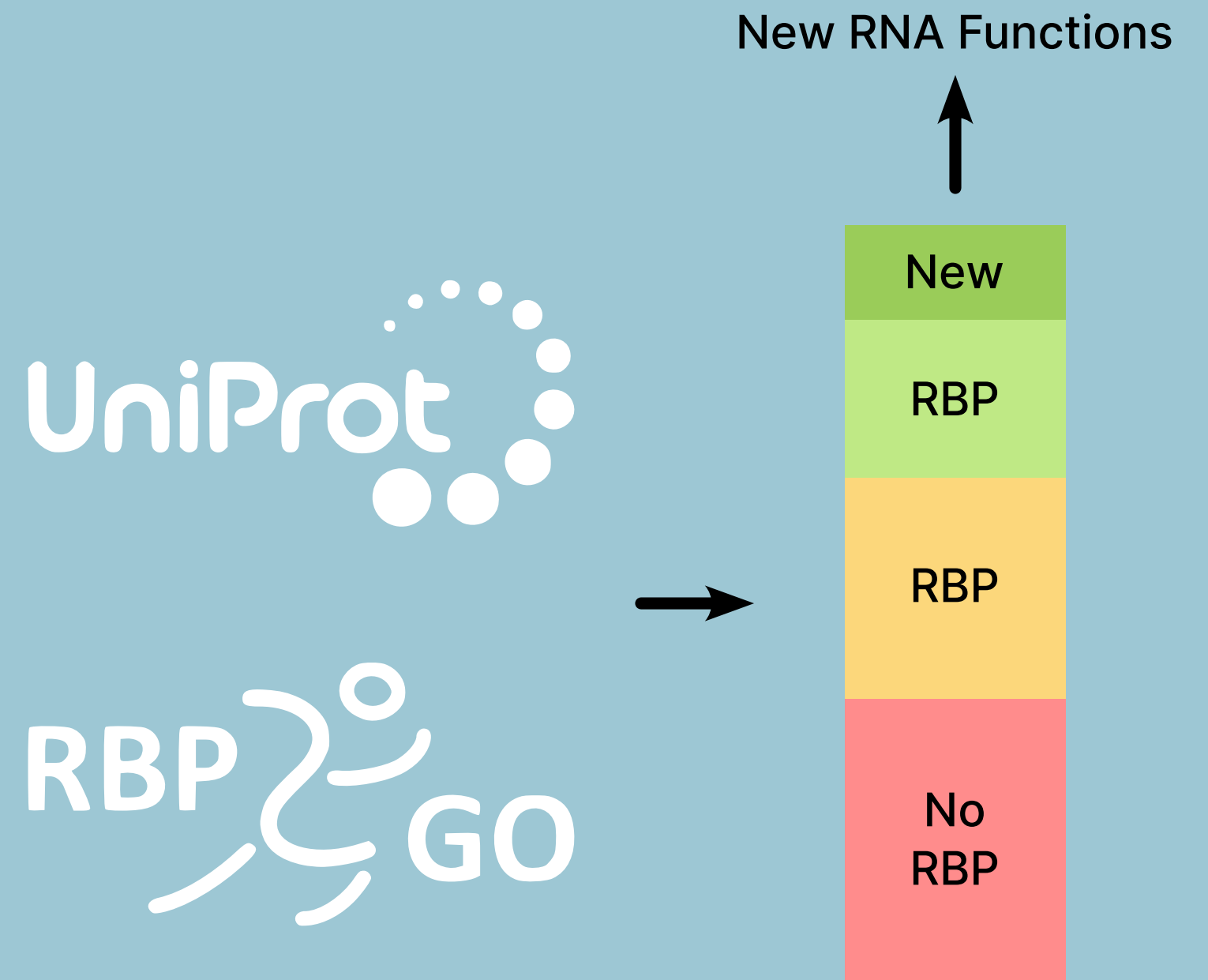


Step 5: Linear Regression and Further Analysis

5.1 Linear Regression



5.2 Further Analysis



Timeline

