

A BAYESIAN FRAMEWORK FOR SPATIAL PROTEOMICS

Oliver M. Crook

Cambridge Centre for Proteomics

Computational Proteomics Unit

MRC Biostatistics Unit

SPATIAL PROTEOMICS

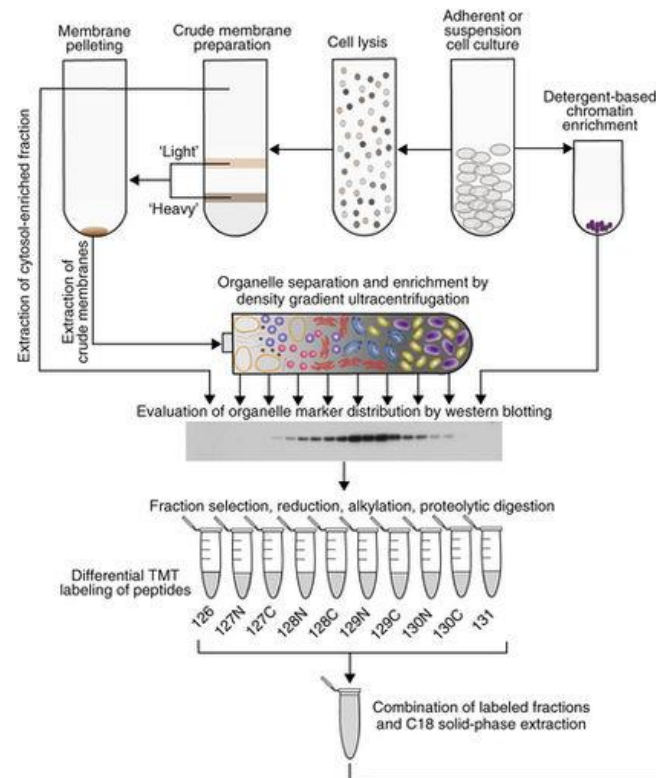
Where is a protein's subcellular location?



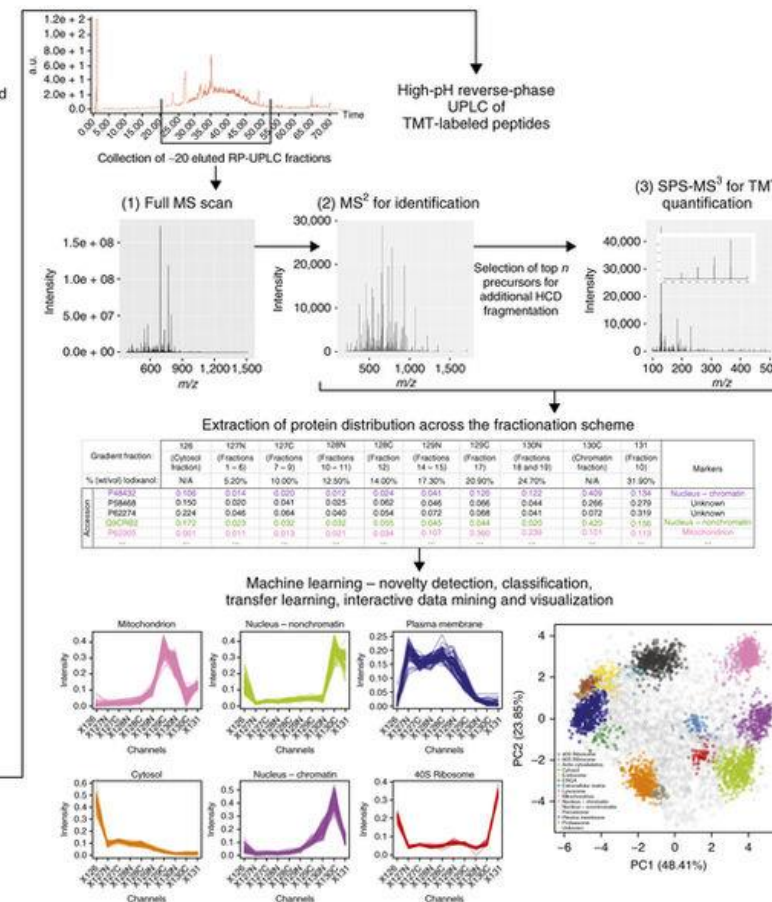
What is protein's function?

HYPERLOPIT

Subcellular Fractionation



Isobaric tagging along density gradient

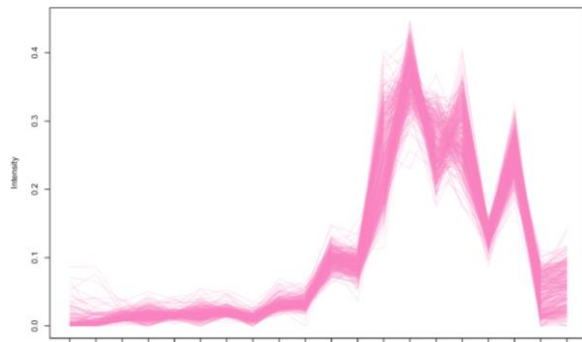


SPS-MS3

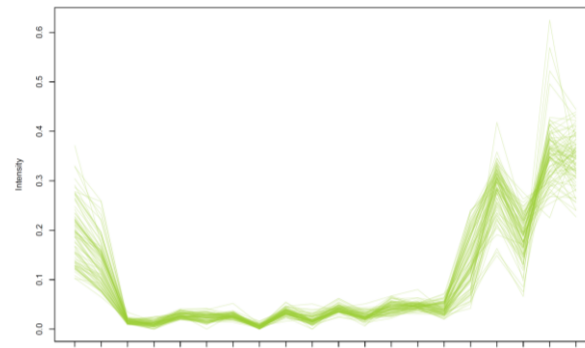
Machine learning

SVM CLASSIFICATION

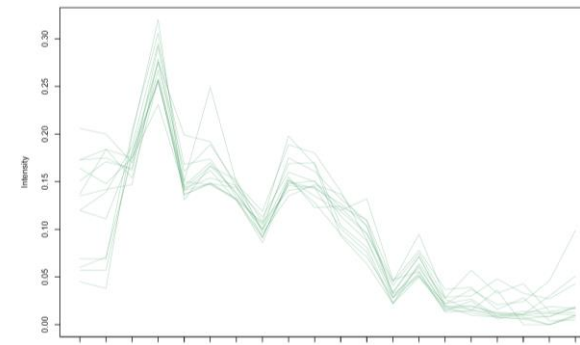
Mitochondrion



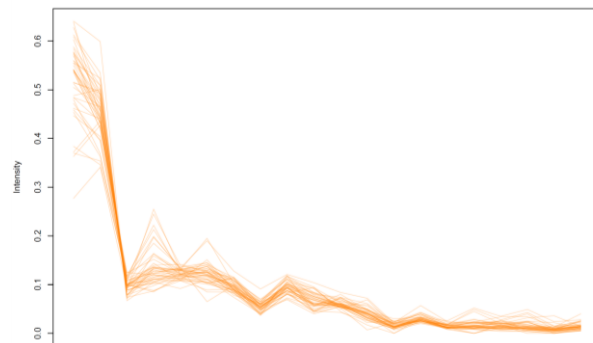
Nucleus - Non-chromatin



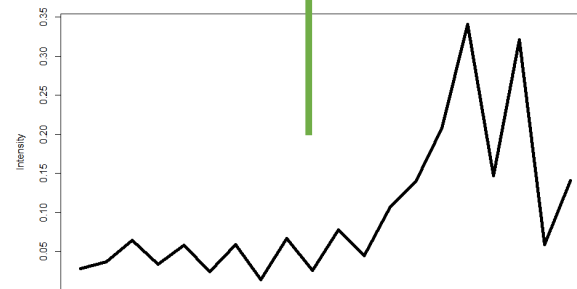
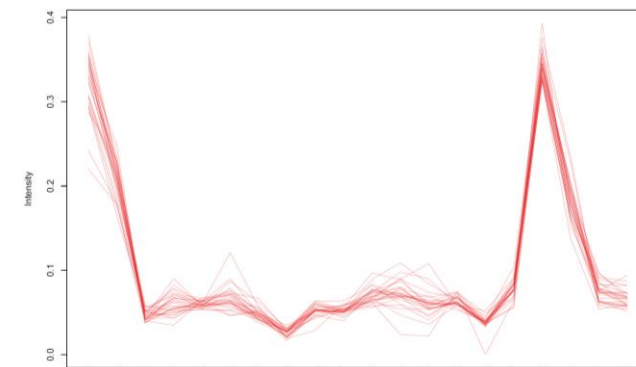
Actin cytoskeleton



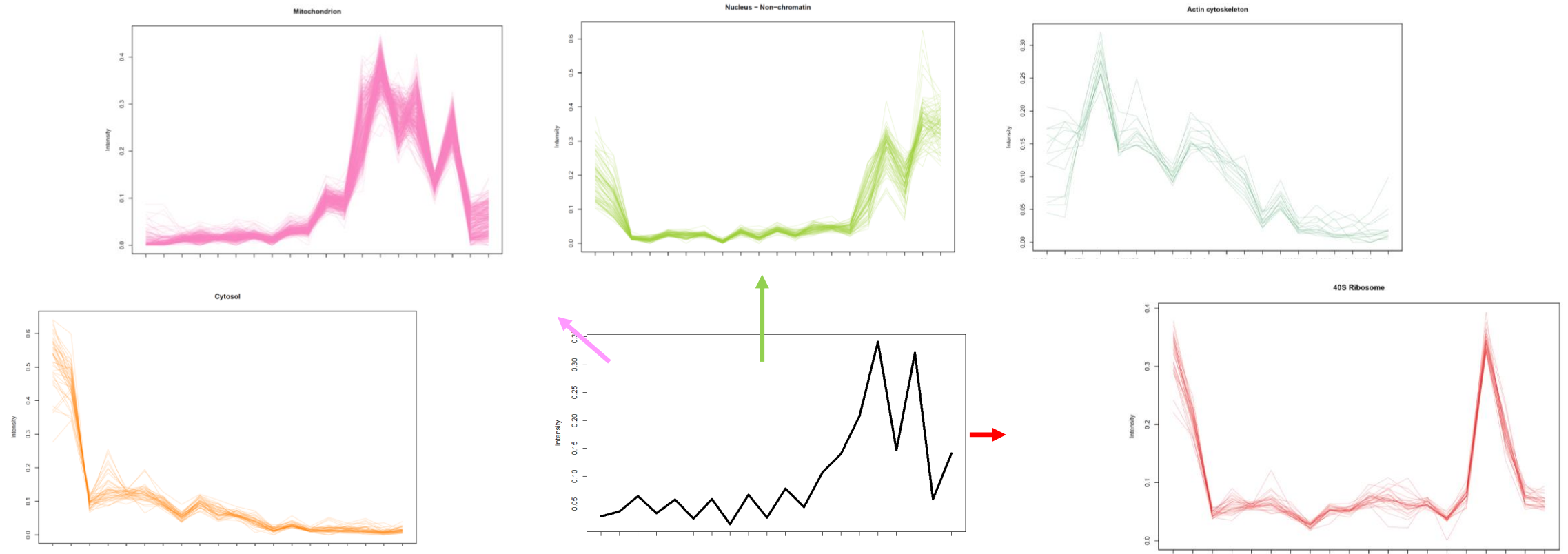
Cytosol



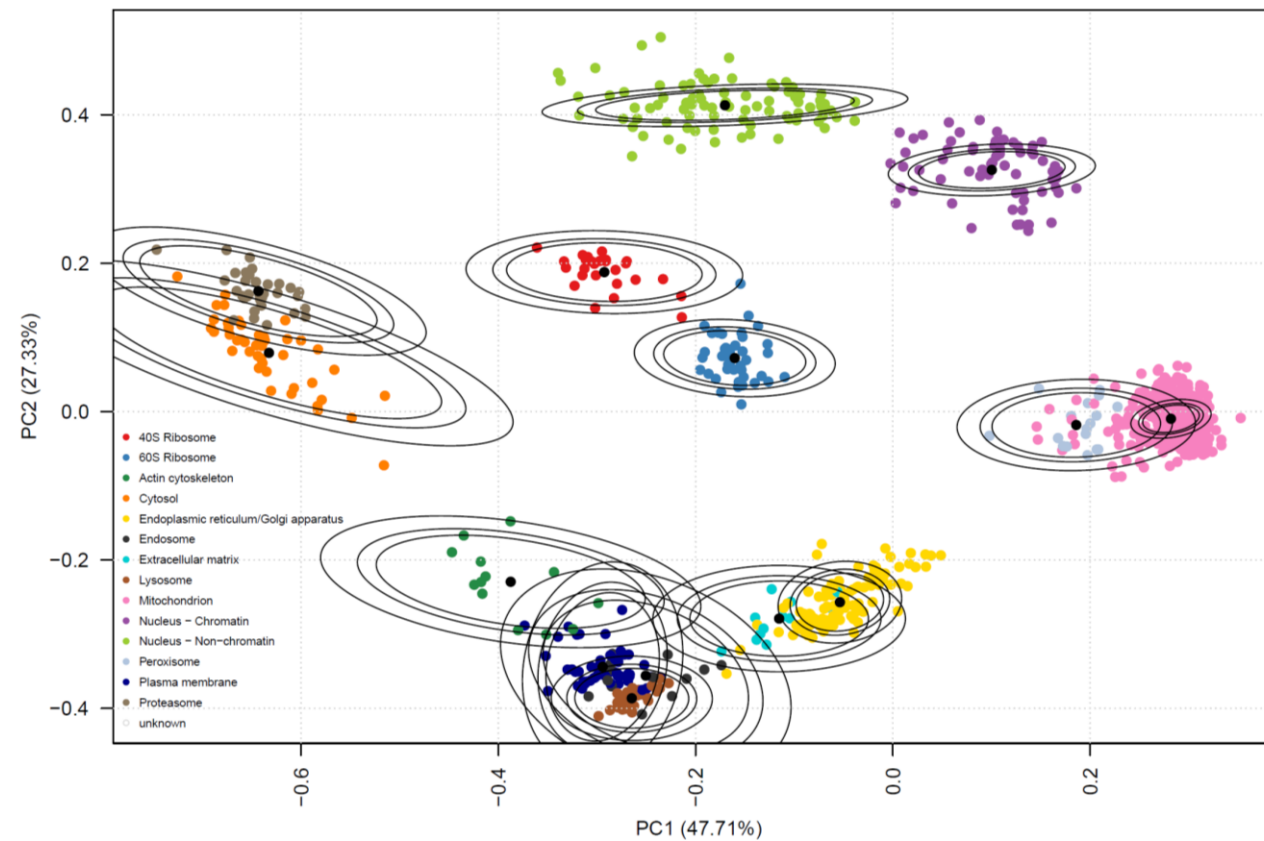
40S Ribosome



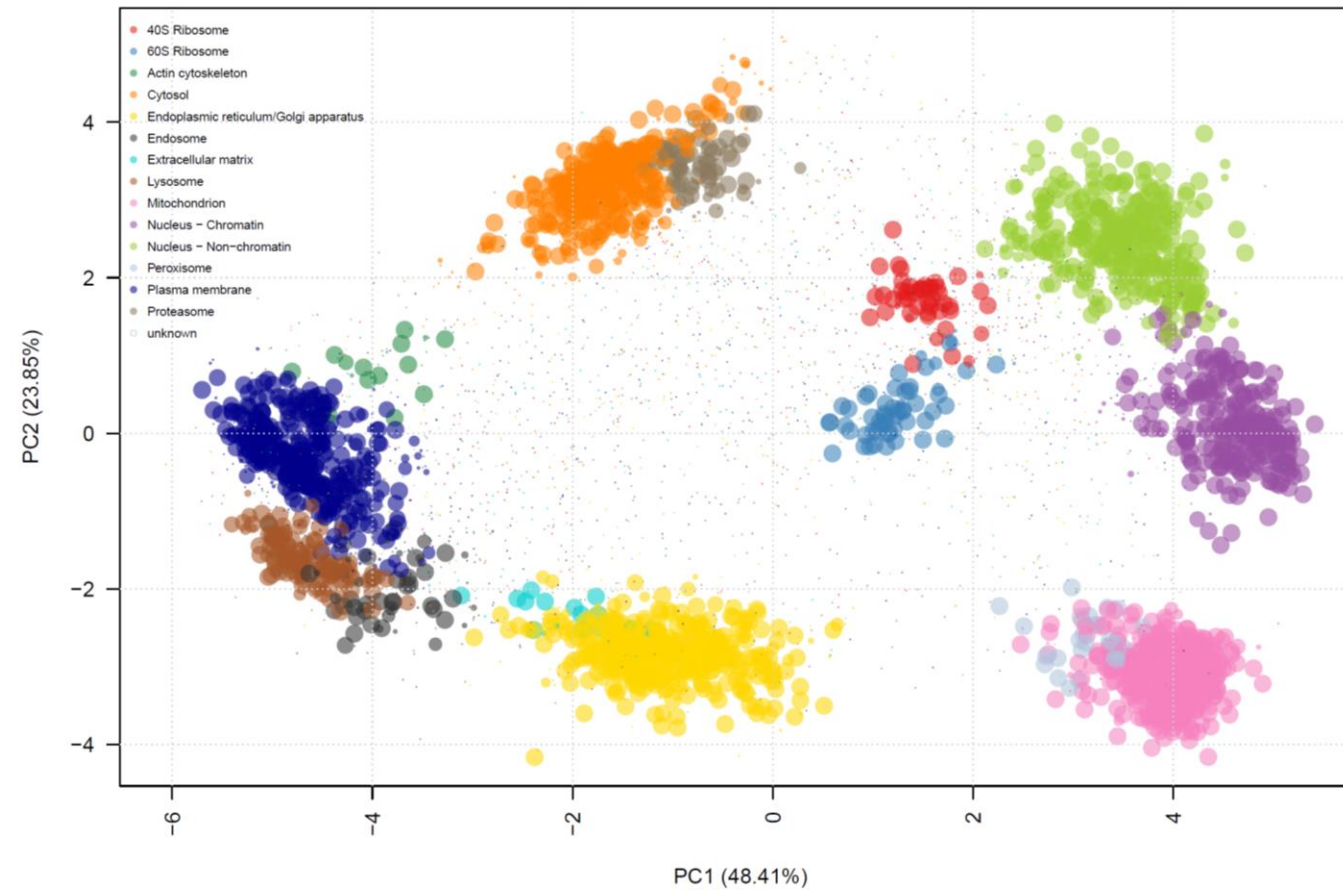
DISTRIBUTION OVER LOCALISATIONS



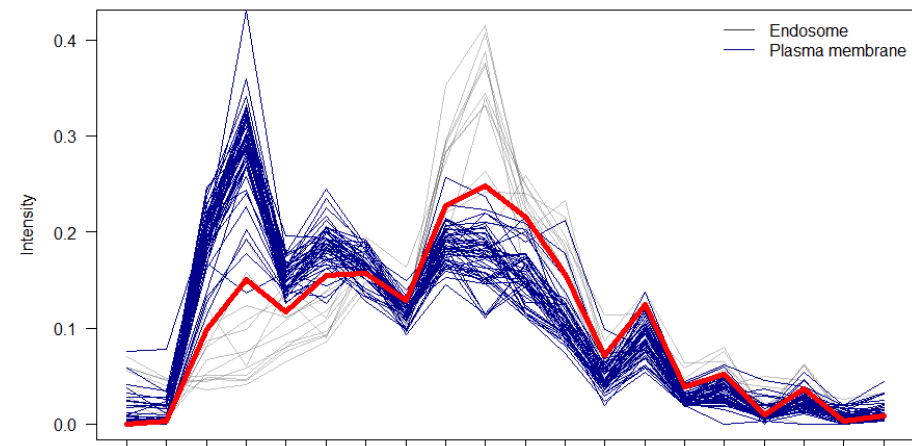
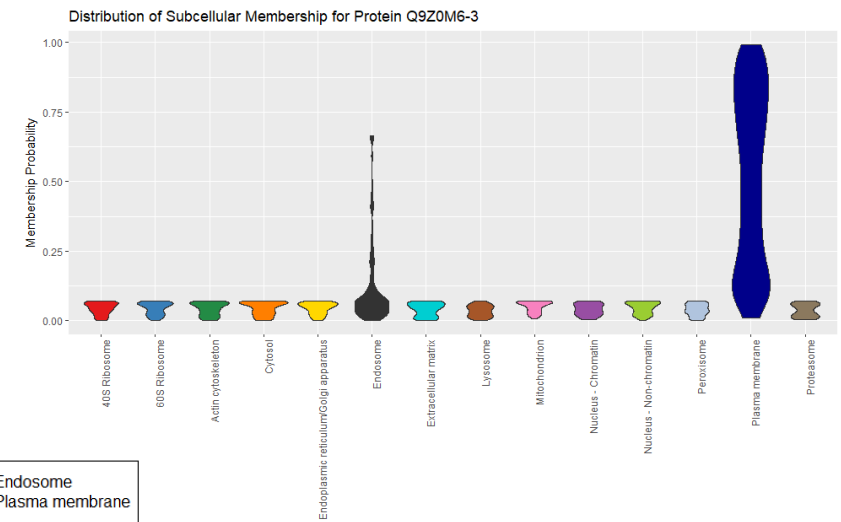
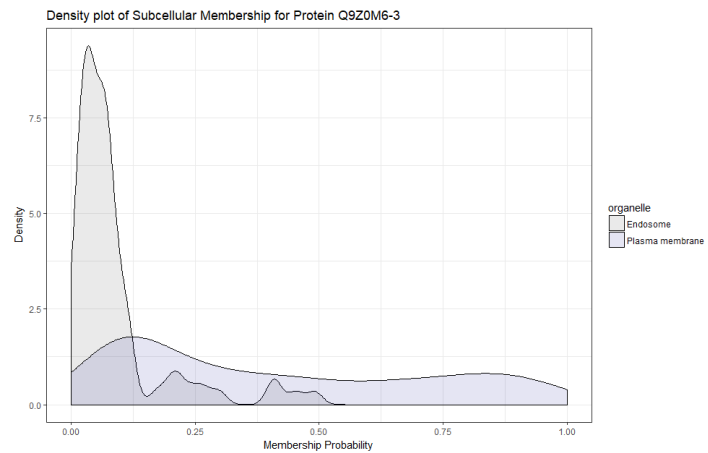
GENERATIVE MODEL



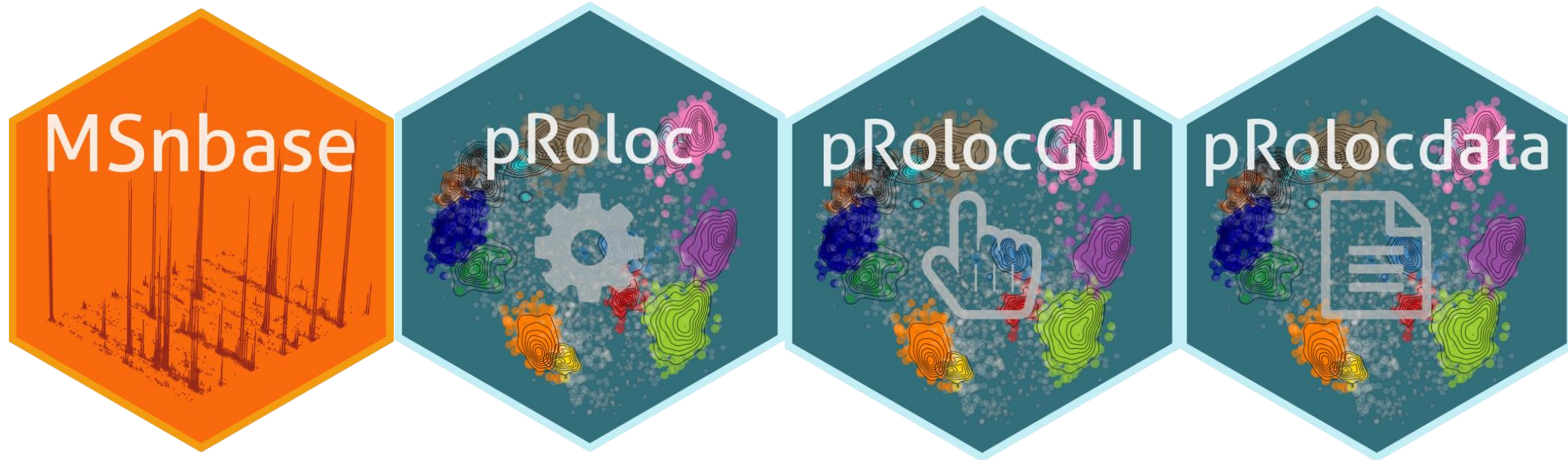
FULL BAYESIAN INFERENCE USING MCMC



UNCERTAINTY



SOFTWARE



CONCLUSIONS AND FUTURE WORK

- Interpretable model of spatial proteomics data
- Probabilistic allocations of proteins to subcellular niches
- Quantified allocation uncertainty
- Extended to model proteins with multiple localizations
- Model protein which move dynamically after perturbation
- Integrate multiple datasets

THANK YOU

Supervisors:

Laurent Gatto

Kathryn Lilley

Paul Kirk

Useful discussions

Claire Mulvey, Sean Holden, Sylvia Richardson

Contact:

omc25@cam.ac.uk

oliver.crook@mrc-bsu.cam.ac.uk

<https://github.com/ococrook>

@OllyMCrook