

Analysis of psychiatric disorders using single nuclei data from post-mortem human brain

MAX PLANCK INSTITUTE
OF PSYCHIATRY



Nathalie Gerstner^{1,2,3}, Anna S. Fröhlich^{1,2}, Michael J. Ziller^{1,4}, Elisabeth B. Binder¹, Janine Knauer-Arloth^{1,3}

¹ Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, 80804 Munich, Germany ² International Max Planck Research School for Translational Psychiatry, 80804 Munich, Germany ³ Institute of Computational Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany

⁴ Department of Psychiatry, University of Münster, Münster, Germany

1 Research Goals

The project aims to identify **key genes and regulatory elements** that drive **psychiatric disorders** in cortical cell types of the human brain. We are going to analyze and integrate cell type specific **differential expression and chromatin accessibility** patterns between cases and controls as well as between donors with high and low genetic risk for psychiatric disorders. Furthermore, we are interested in **allele-specific regulatory and transcriptional effects** and cell-type specific **gene regulatory networks**.

2 Data

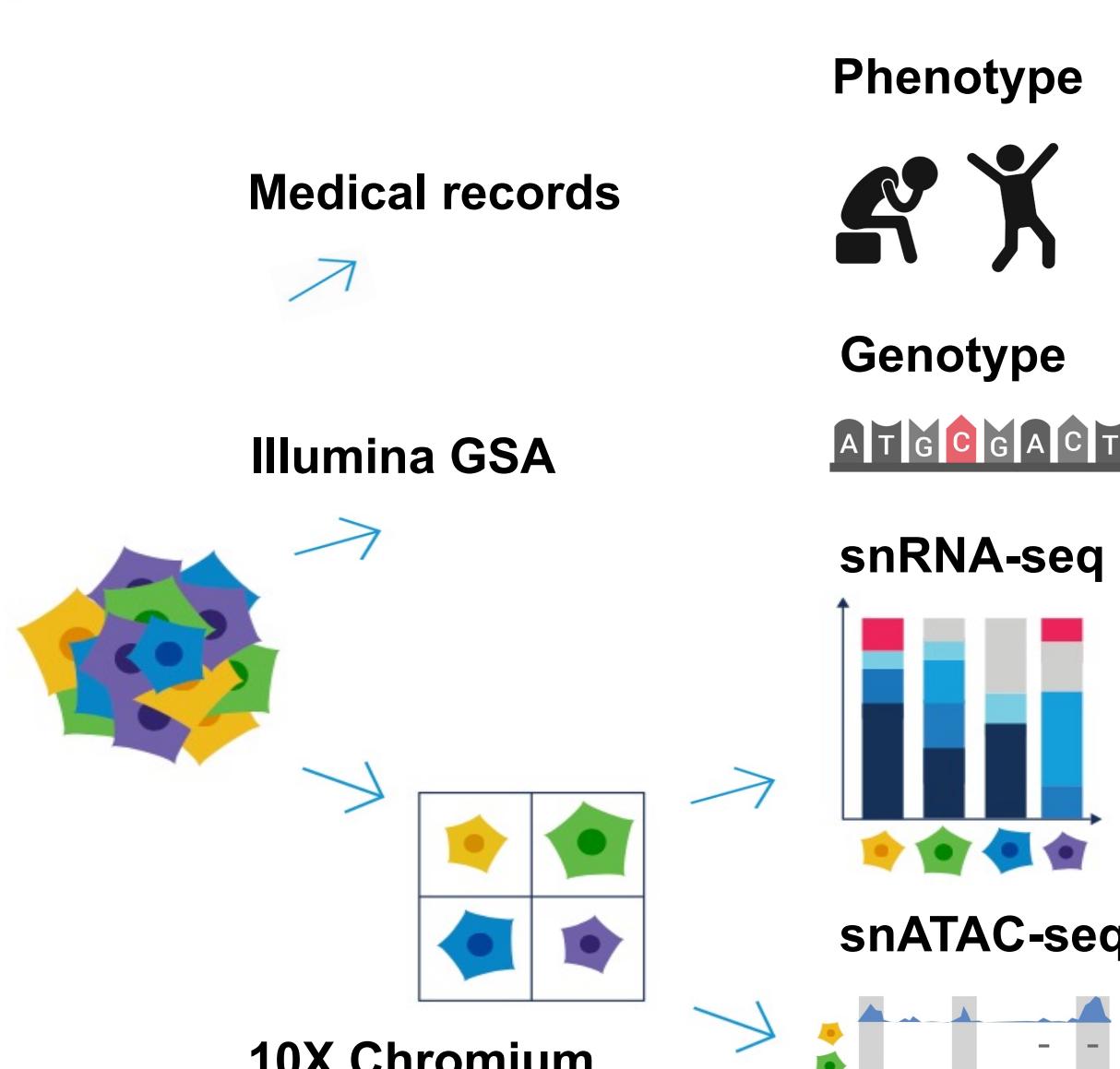


Table 1:
Overview of cohort

	Controls (n=35)	Cases (n=57)
Diagnoses	-	SCZ (38), SCA (7), MDD (7), BIP (5)
Sex	13 Female	22 Female
Age (mean ± s.e.m)	55.8 ± 2.3	53.3 ± 1.8
PMI (h) (mean ± s.e.m)	31.8 ± 1.9	35.2 ± 2.2
pH (mean ± s.e.m)	6.66 ± 0.04	6.57 ± 0.03
Cause of death	Accident (1), Natural (33), Unknown (1)	Accident (3), Natural (36), Suicide (16), Unknown (2)

Figure 1:
Overview of data modalities

The dataset contains **phenotype**, **genotype**, single nuclei transcriptomic (**snRNA-seq**) and open chromatin (**snATAC-seq**) data from the **postmortem orbitofrontal cortex** tissue of 92 donors.

3 Methods

- Calculation of **polygenic risk scores (PRS)** with PRS-CS¹ on multiple diseases and stress-related traits
- Demultiplexing and **alignment** of reads with *CellRanger*²
- Downsampling** of reads with *DropletUtils*³
- Processing of snRNA-seq** data with *Scanpy*⁴, *DoubletDetection*⁵, *sctransform*⁶ and *scArches*⁷
- Processing of snATAC-seq** data with *ArchR*⁸
- Cell-type label transfer with data from *Allen Brain Atlas*⁹ as **reference data** and manual refinement using marker genes

4 Genetic risk for psychiatric diseases

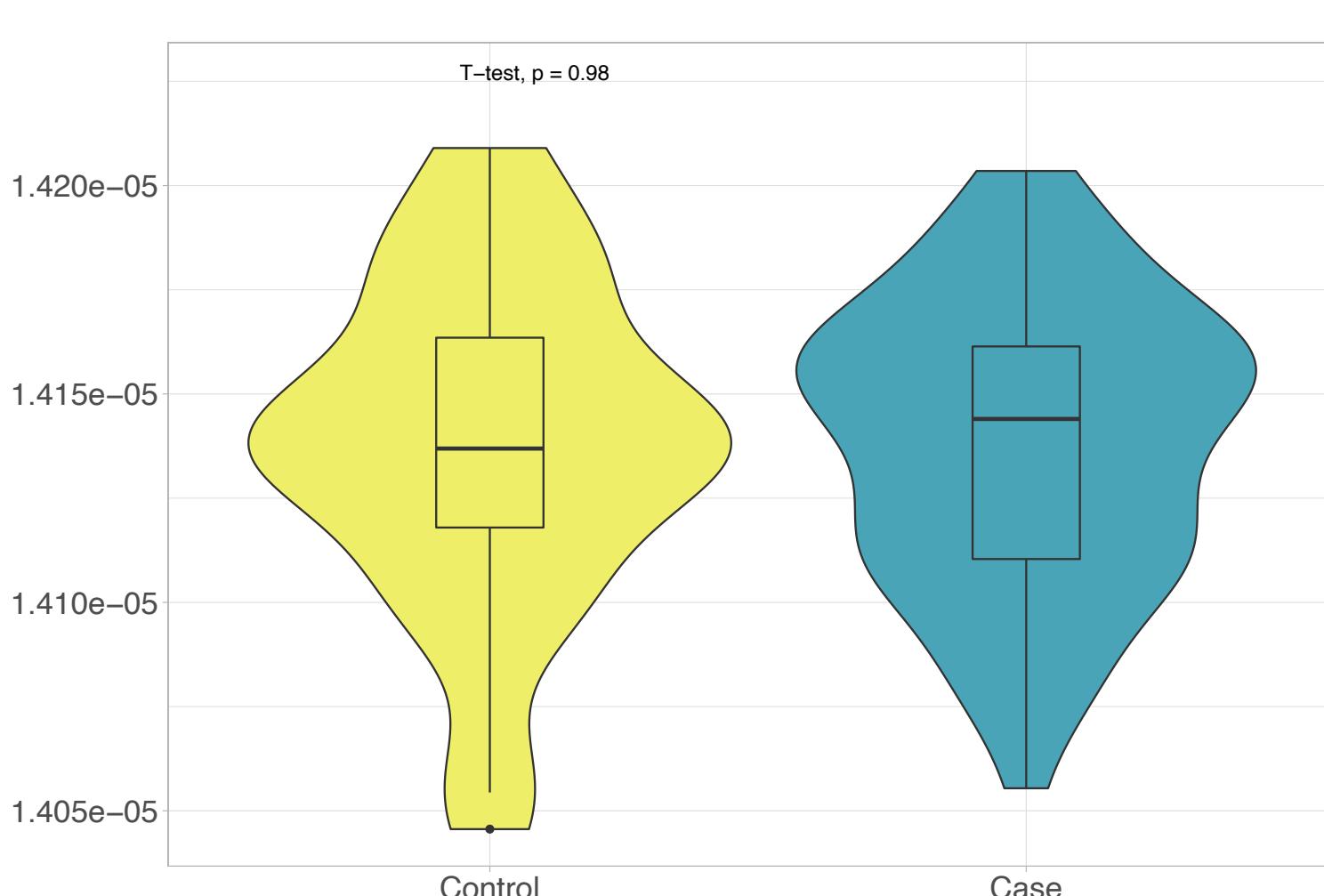


Figure 2: Distribution of PRS based on cross disorder GWAS for groups of controls and cases.

Polygenic risk scores (PRS) calculated for multiple diseases and stress-related traits are **not always correlated with diagnoses** for psychiatric disorders. The mean of PRS for cases, based on the cross disorder GWAS¹⁰, is only slightly higher than the mean for controls. It is therefore important to not only **identify molecular differences** between cases and controls, but also individuals with **high and low genetic risk**.

5 Gene expression and chromatin accessibility at the resolution of single cells

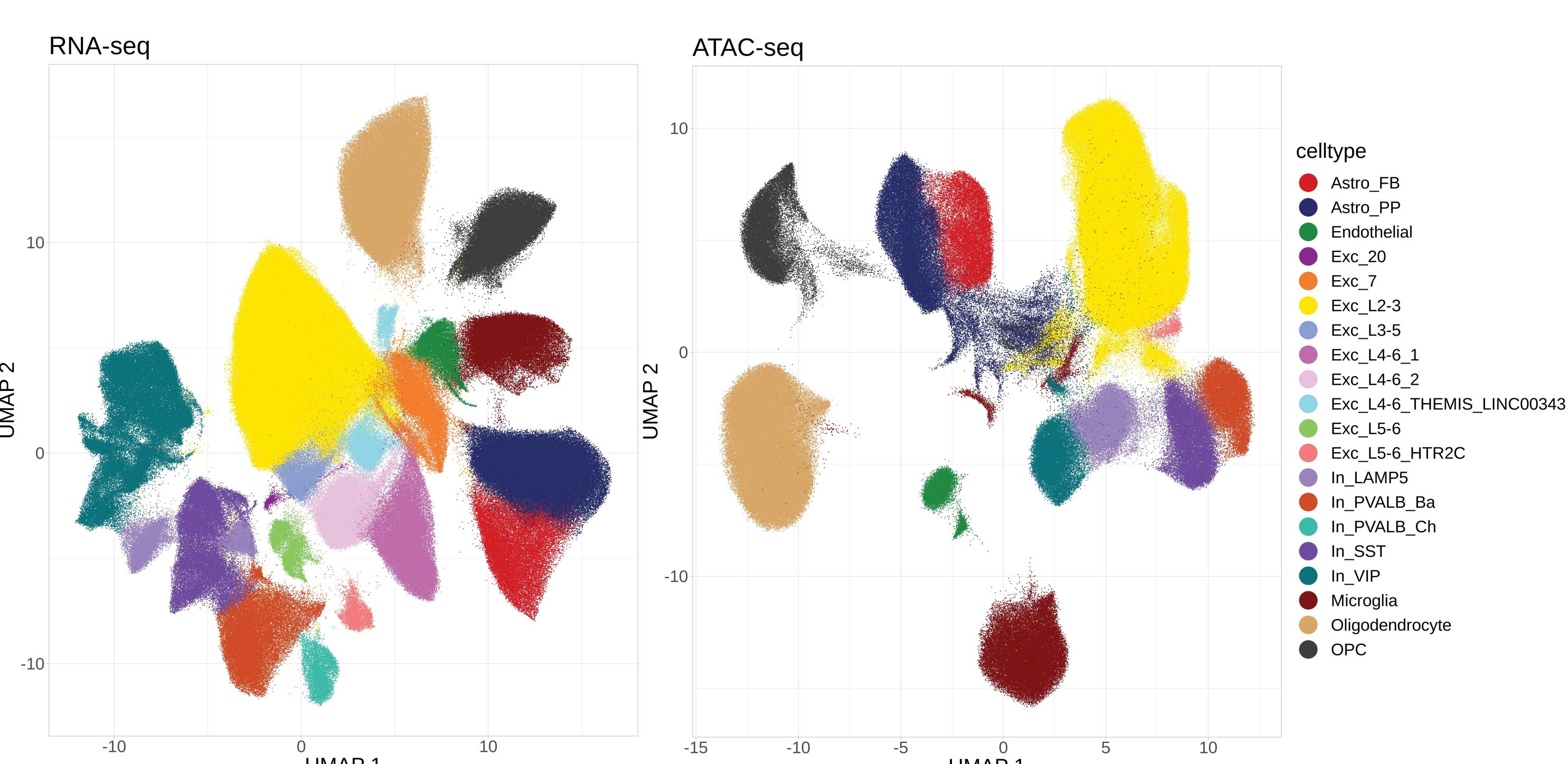


Figure 3: UMAP representation of snRNA-seq and snATAC-seq data.

After the downsampling of reads per cell, quality control, doublet removal, dimensionality reduction, clustering and cell type assignment, the **snRNA-seq data** consists of gene expression values for **26.195 genes** across **813.095 cells** originating from 87 individuals. Cells were assigned to **20 distinct cell types** that cover the major cell types of the prefrontal cortex and the different neuronal layers.

The **snATAC-seq data** consists of fragments from accessible chromatin across **459.678 cells** originating from 91 individuals after quality control, doublet removal, dimensionality reduction clustering and cell type assignment. Cells were assigned to **12 distinct cell types** - neuronal subtypes present in the RNA-seq data are missing.

6 Outlook

The next steps in our analyses include the **differential analysis of gene expression and chromatin accessibility** between cases and controls and genetic risk exposures. We are planning to identify **cell-type specific eQTLs** by taking allelic imbalance into account and thereby increasing the power of the analysis. **Gene regulatory networks** will be inferred for the different cortical cell types.

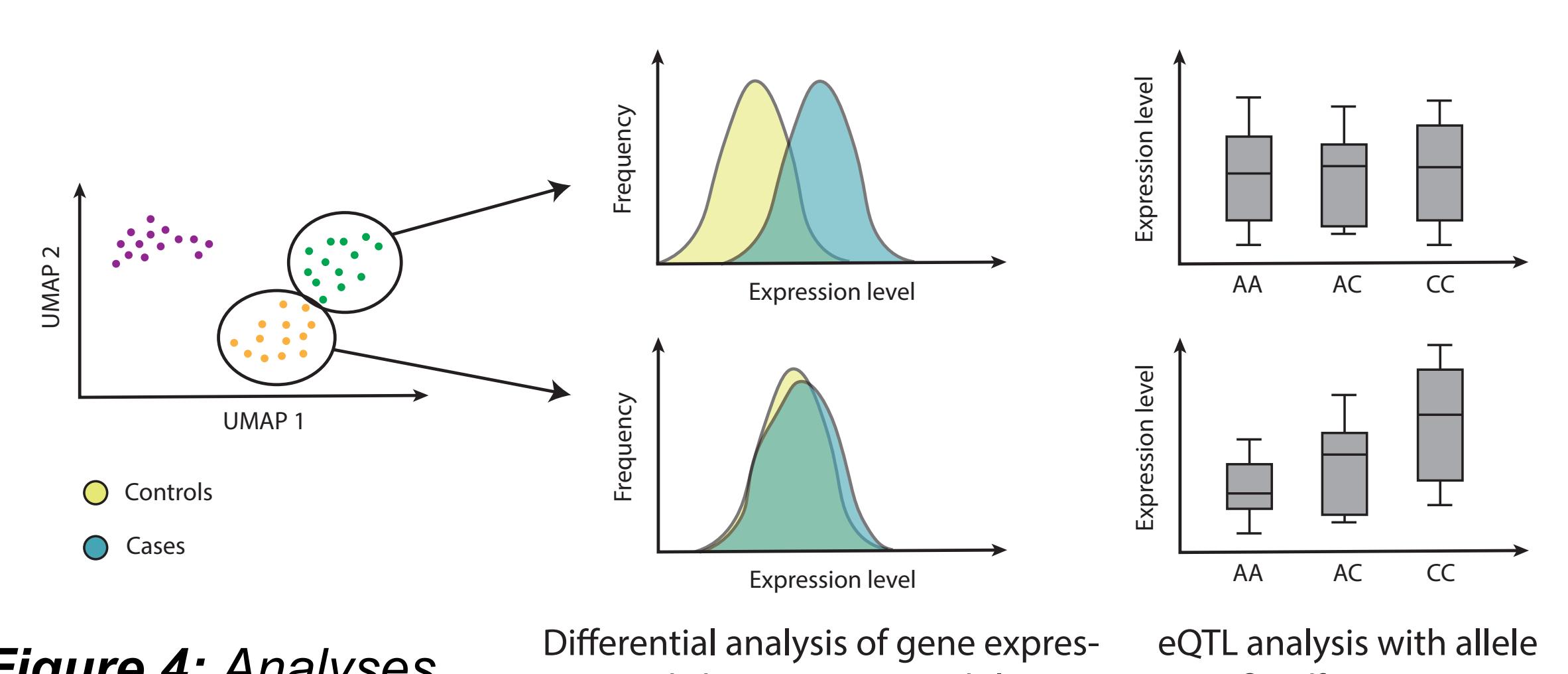


Figure 4: Analyses planned to perform on the data modalities.

References

Please scan QR code for references

