

Single-Cell RNA-Seq Analysis: EDA

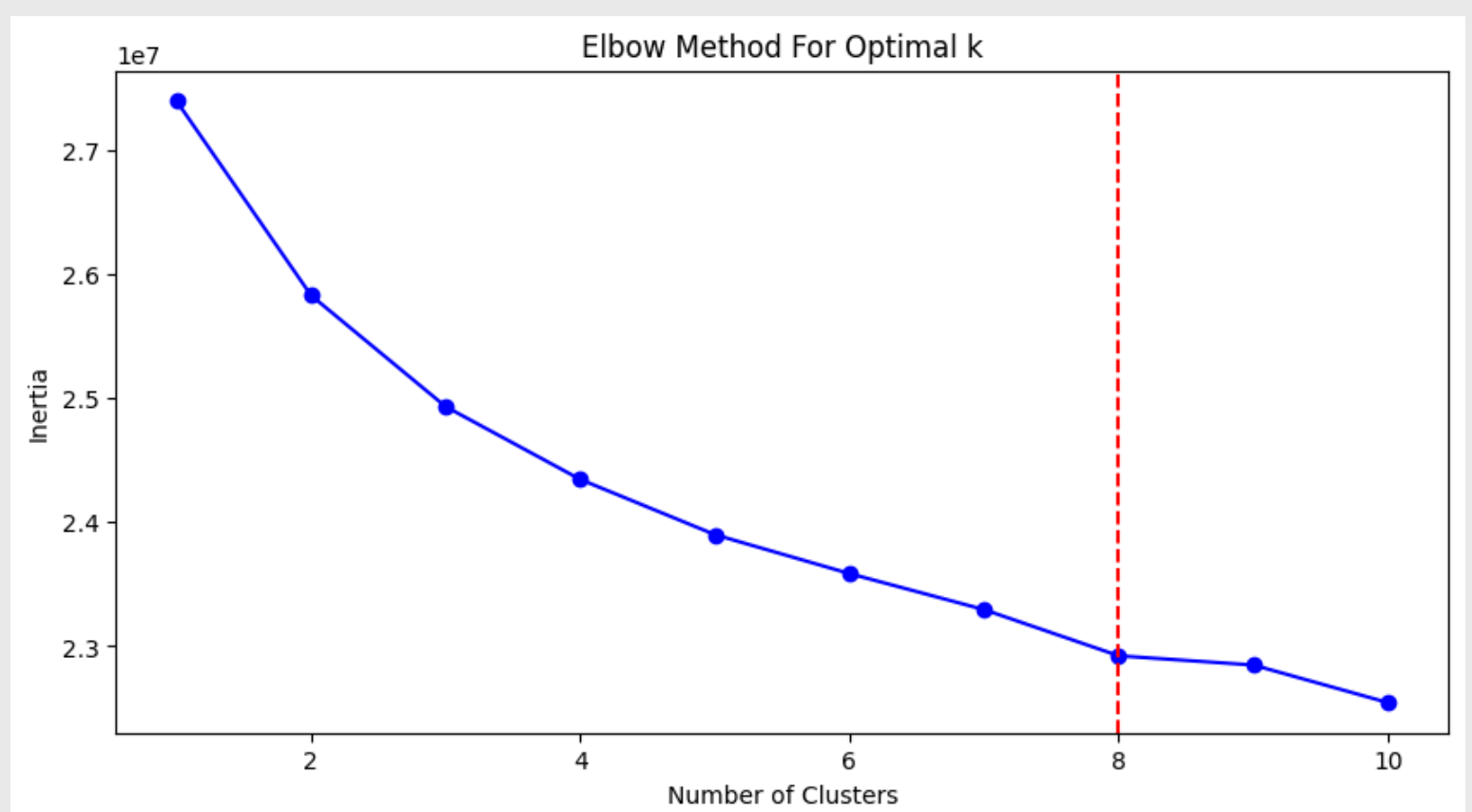
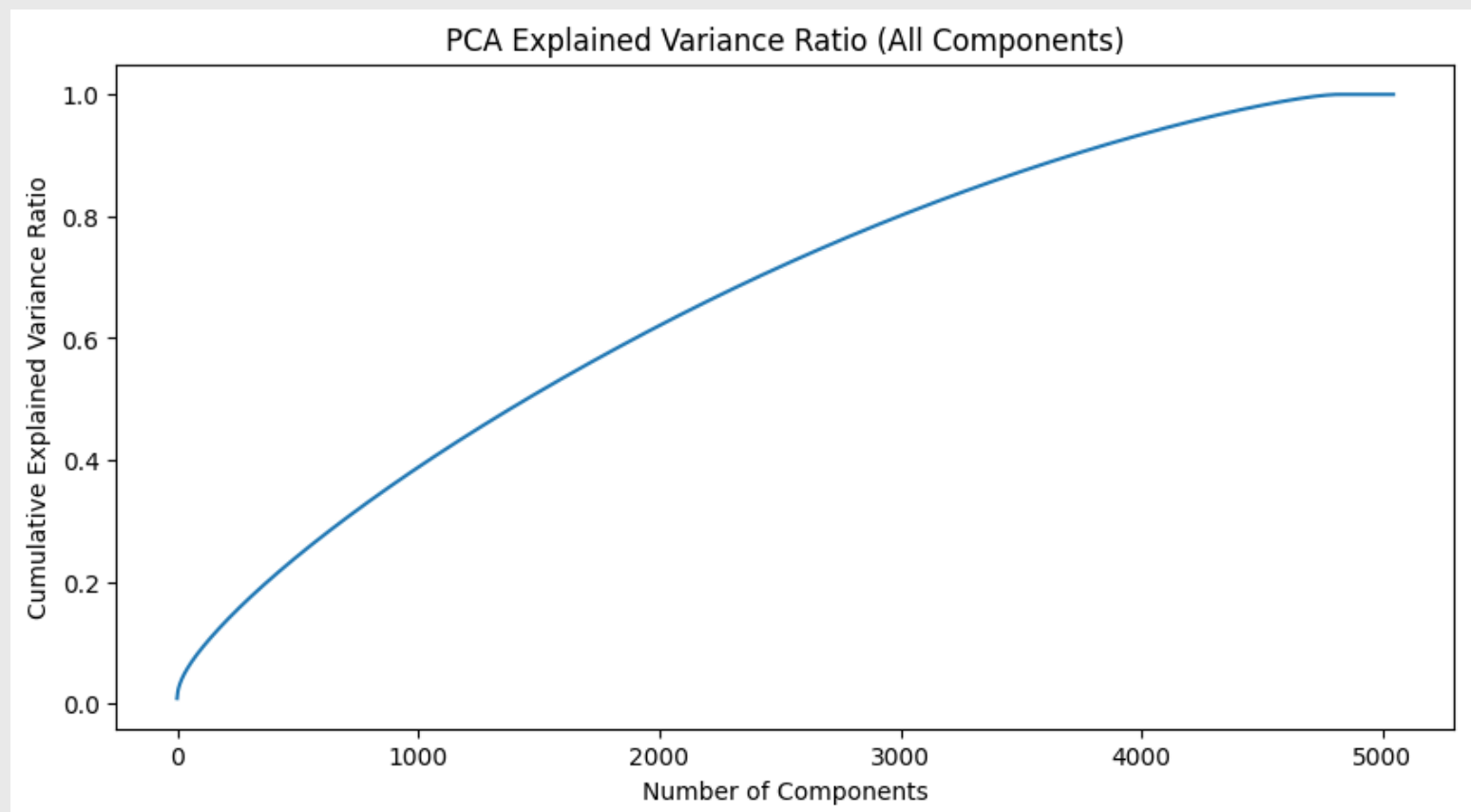
Schehat Abdel Kader • Detijon Lushaj

Data

- 91205 cells
- 7353 control cells
- 83852 pertubated cells
- 5045 genes
- Shows how different pertubations affect gene expression
- All instances are from A549 cell
- Cell is used to test the effects of drugs and treatments on lungcancer cells

PCA & K-Means Clustering

- PC1 and PC2 explain 1.3% of the variance
- PC1 - PC128 explain 10%
- > use for clustering



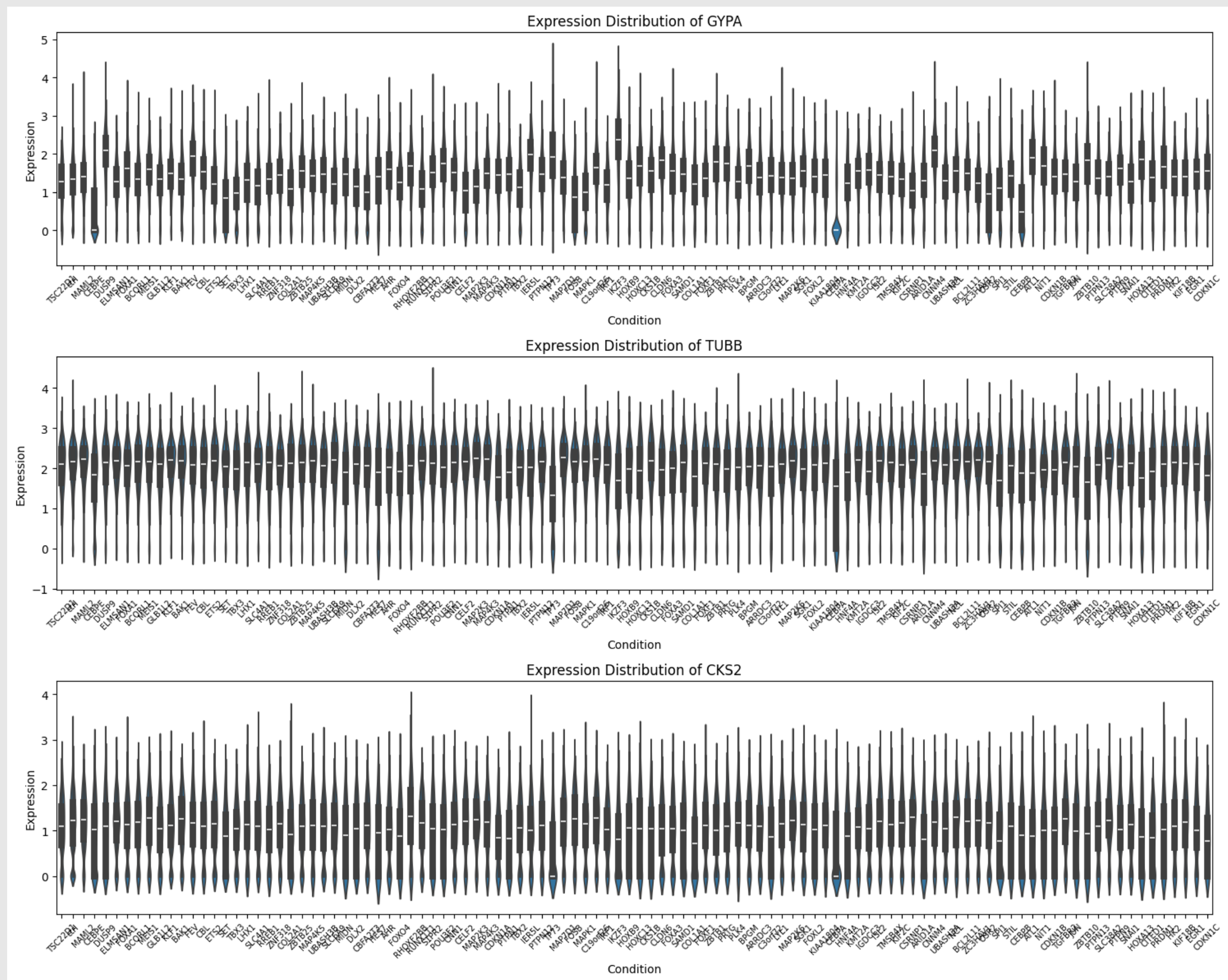
- In C0, C1, C4 & C7 highest expression gene MALAT1
- C5 one cell by condition TBX2
- High expression of gene PLCG2
- Suppresses gene MT-ATP6
- C6 one control cell

	C0	C1	C2	C3	C4	C7
Highest variance gene	HBZ	HBZ	CLC	HBZ	GDF15	HIST1H4C
Caused by condition	IKZF3	KLF1	CEBPA	SET	OSR2	KLF1
Difference to control	0.51	0.32	1.60	0.45	1.33	-0.58



Cluster	#Cells
0	5382
1	14651
2	947
3	21409
4	466
5	1
6	1
7	12903

Dynamic Gene Expression Patterns Across Experimental Conditions



GYPA

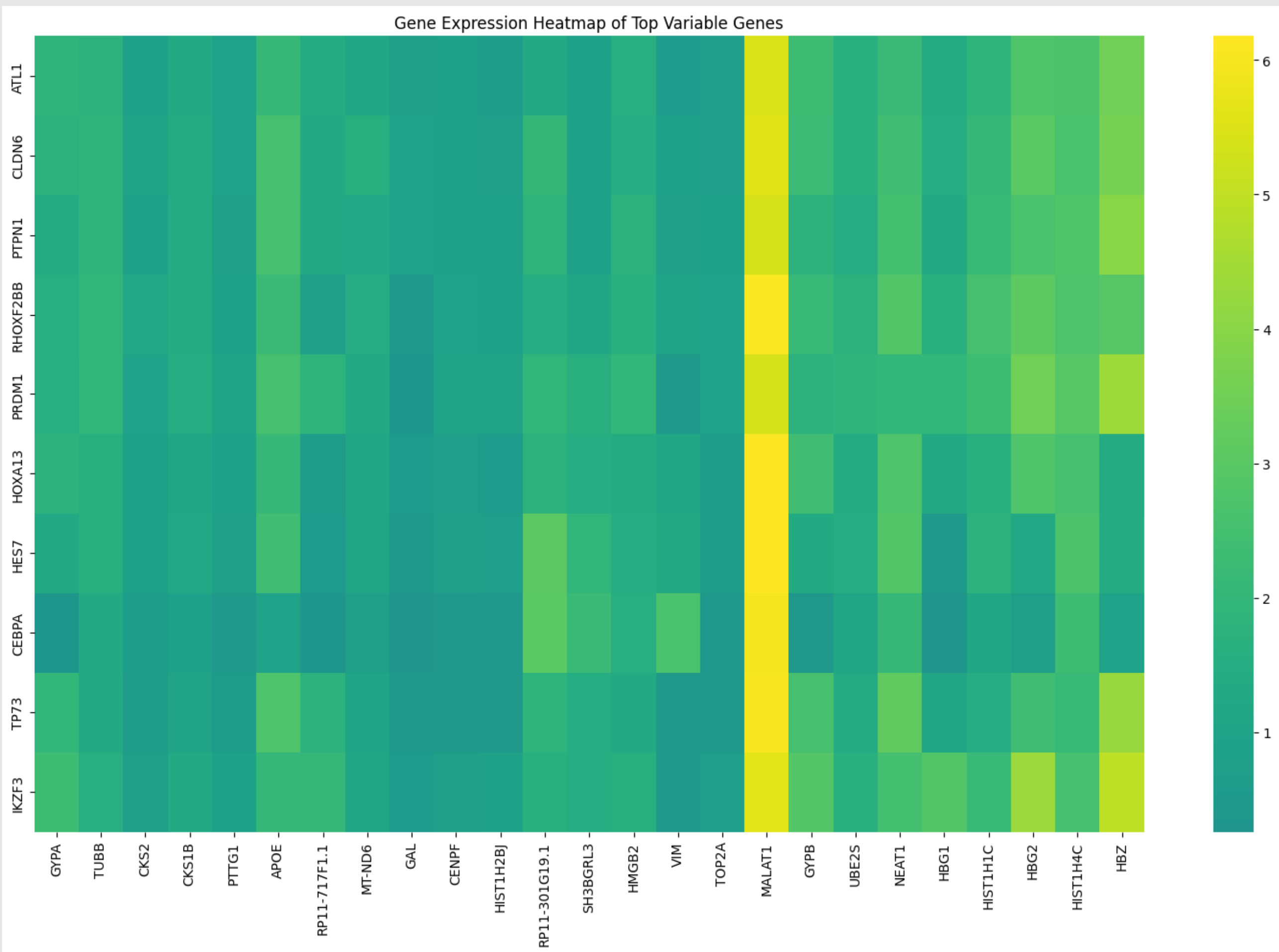
- Low expression, but high variability

TUBB

- Moderate, consistent expression

CKS2

- Low, stable expression across all conditions



MALAT1

- High, robust expression
- Likely key in regulatory pathways