

PCA plot showing the first principal component (PC1) accounting for 19% of the variance. The x-axis is labeled 'PC1: 19% variance' and ranges from -0.2 to 0.4. The y-axis is unlabeled. Data points are represented by colored triangles and circles, each labeled with a protein name and a variant identifier. The points are distributed across the plot, with some clustering and others isolated.

Protein Variant	PC1 (approx.)	PC2 (approx.)
PANRDC-03A-02R	-0.05	0.45
PAPVTA-09A-02R	-0.15	0.40
PANKAK-09A-02R	-0.10	0.35
PASCJU-03A-02R	0.05	0.35
PANRDC-03A-01R	0.25	0.35
PAPVTA-09A-01R	0.05	0.25
PAPNNX-09A-02R	0.10	0.25
PANRDC-03A-02R	-0.15	0.20
PANKAK-09A-01R	0.05	0.20
PANWHH-03A-02R	0.20	0.20
PASCJU-03A-01R	0.35	0.20
PAPNNX-09A-01R	0.30	0.15
PANWHH-03A-01R	0.40	0.10
PANLIC-09A-01R	-0.05	0.15
PAPLUN-09A-01R	0.00	0.10
PASFYA-04A-01R	-0.15	0.05
PANRDC-03A-01R	-0.10	0.05
PANWHH-03A-01R	0.40	0.05
PARBVI-09A-02R	-0.15	-0.10
PARBVI-09A-01R	0.20	-0.15

●

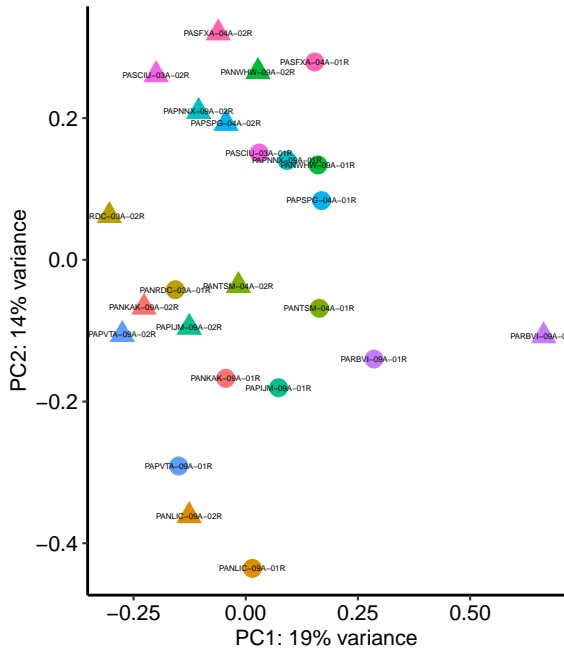
UMAP2

UMAP1

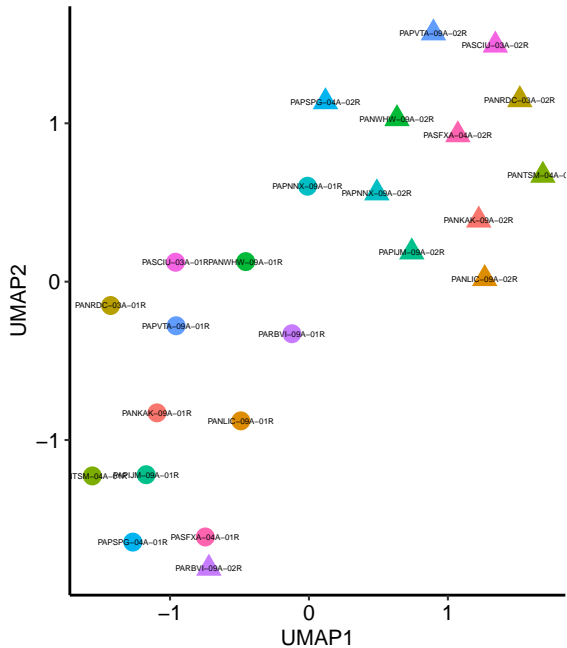
Cell lines and conditions shown in the UMAP plot:

- PANKAK-09A-01R
- PANKAK-09A-02R
- PAPLUM-09A-01R
- PAPLUM-09A-02R
- PAPVTA-09A-01R
- PAPVTA-09A-02R
- PASCIL-03A-01R
- PASCIL-03A-02R
- PANRDC-03A-01R
- PANRDC-03A-02R
- PASFXA-04A-01R
- PASFXA-04A-02R
- PAPSFG-04A-01R
- PAPSFG-04A-02R
- PANWVH-03A-01R
- PANWVH-03A-02R
- PANLIL-05A-01R
- PANLIL-05A-02R
- PARBVI-09A-01R
- PARBVI-09A-02R
- PAPNIX-09A-01R
- PAPNIX-09A-02R

### PCA: RUVg output ( $k = 2$ )

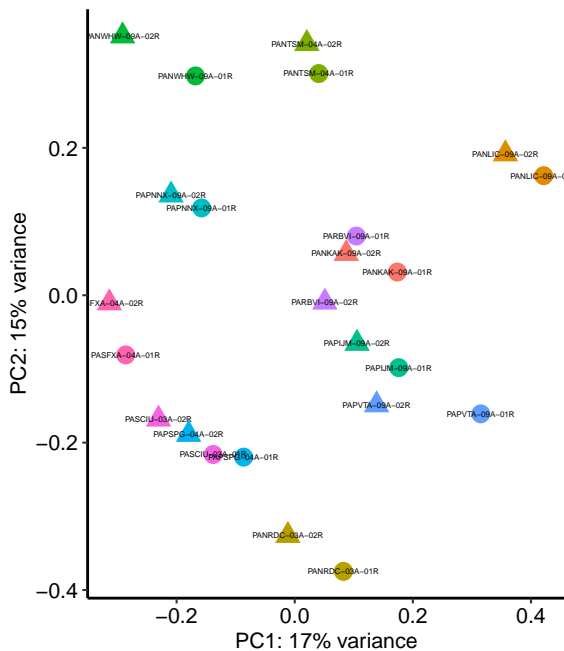


UMAP: RUVg output ( $k = 2$ )

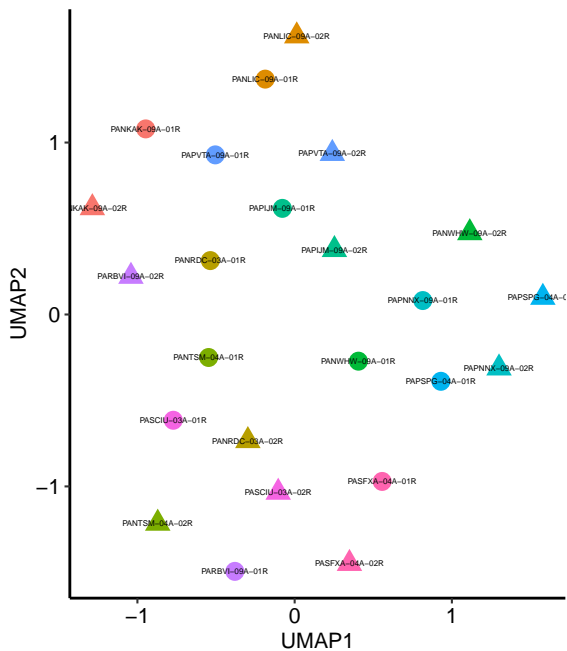


rna\_library   ●   poly-A stranded   ▲   stranded

PCA: RUVg output (k = 3)

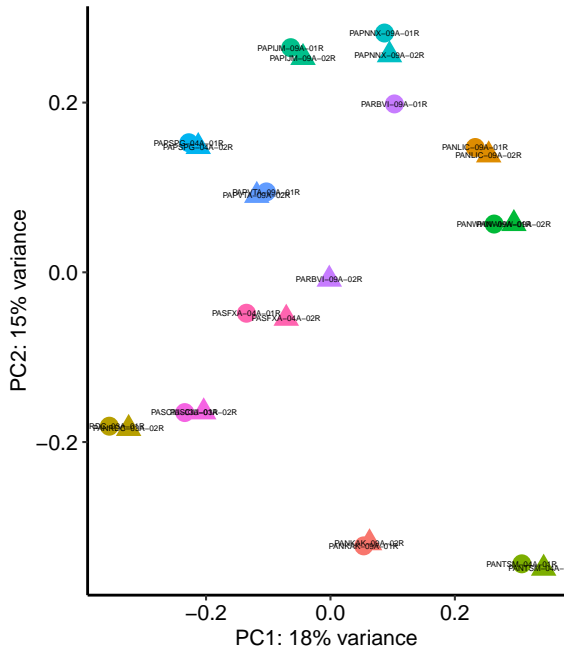


UMAP: RUVg output (k = 3)

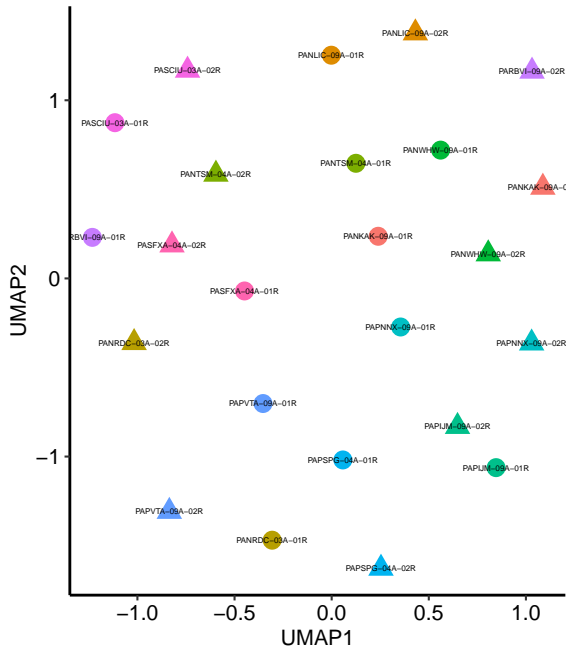


rna\_library ● poly-A stranded ▲ stranded

PCA: RUVg output (k = 4)

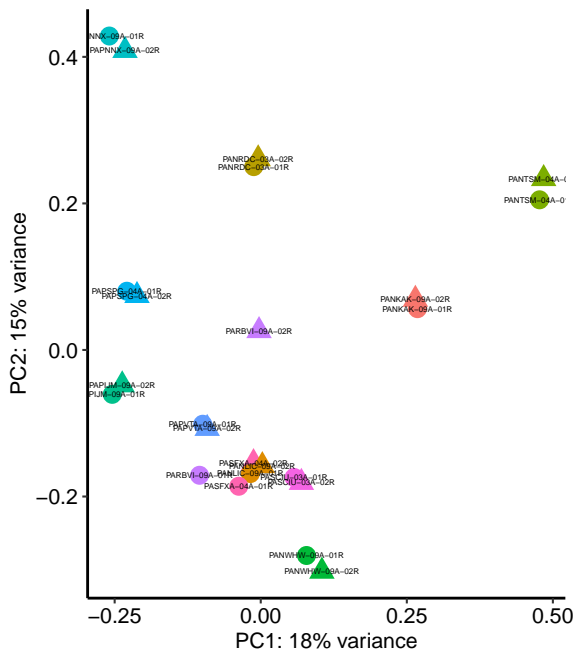


UMAP: RUVg output (k = 4)

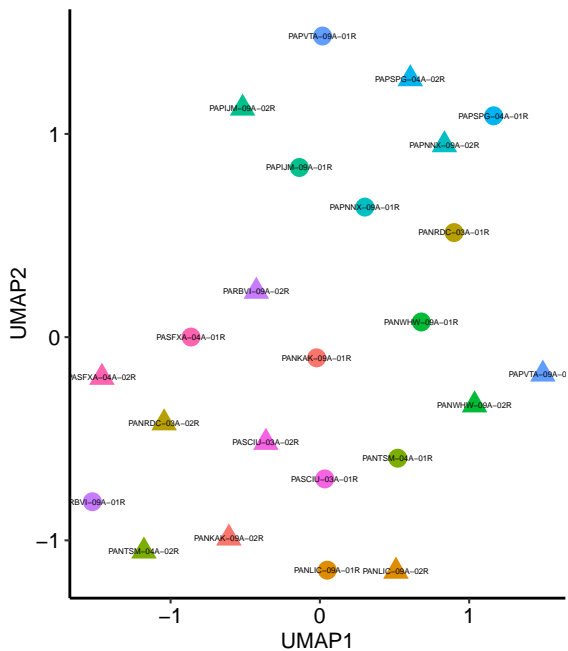


rna\_library    ● poly-A stranded    ▲ stranded

PCA: RUVg output ( $k = 5$ )




UMAP: RUVg output ( $k = 5$ )



rna\_library

- poly-A stranded

 stranded