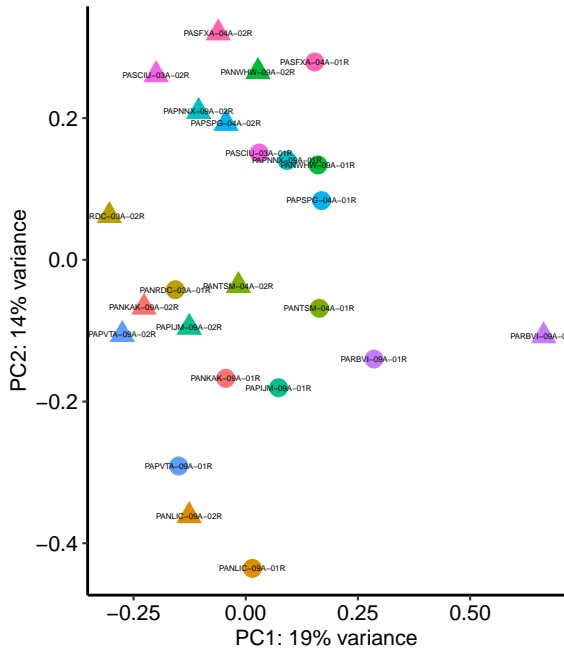


PCA plot showing the first principal component (PC1) accounting for 19% of the variance. The plot displays 15 data points, each representing a specific protein variant. The x-axis is labeled 'PC1: 19% variance' and ranges from -0.2 to 0.4. The y-axis is unlabeled. The data points are color-coded by protein family: PANK (blue, orange, green, pink, purple) and PAPS (cyan, yellow, red, magenta).

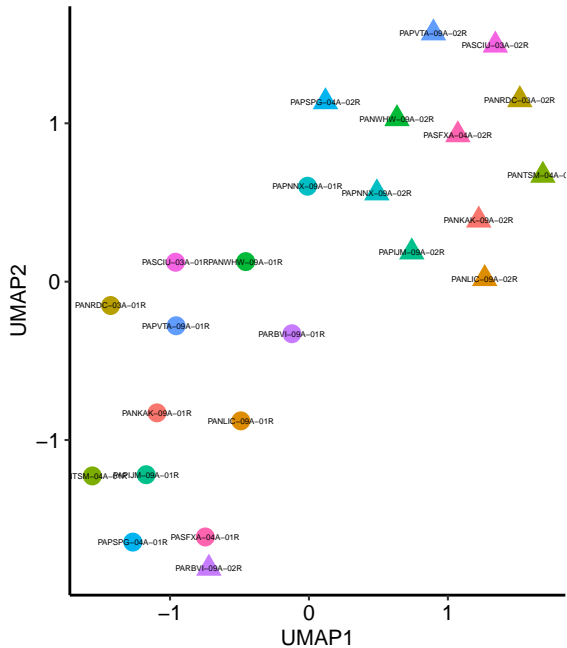
Protein Variant	Approx. PC1 (19% variance)	Approx. PC2 (19% variance)	Color	Shape
PANRDC-03A-02R	0.05	0.95	Yellow	Triangle
PAPVTA-09A-02R	-0.15	0.85	Blue	Triangle
PANKAK-09A-02R	-0.10	0.75	Red	Triangle
PASCIU-03A-02R	0.10	0.80	Magenta	Triangle
PANRDC-02A-01R	0.25	0.75	Yellow	Circle
PAPVTA-09A-01R	0.05	0.65	Blue	Circle
PAPNNX-09A-02R	0.15	0.65	Cyan	Triangle
PAPVTA-09A-02R	-0.15	0.60	Orange	Triangle
PANRDC-02A-02R	-0.15	0.55	Orange	Triangle
PANKAK-09A-01R	0.05	0.55	Red	Circle
PANWHH-03A-02R	0.25	0.50	Green	Triangle
PASCIU-03A-01R	0.40	0.50	Magenta	Circle
PAPNNX-09A-01R	0.35	0.45	Cyan	Circle
PANWHH-03A-01R	0.45	0.40	Green	Circle
PANLIC-09A-01R	-0.05	0.40	Orange	Circle
PAPLUN-09A-01R	-0.05	0.35	Green	Circle
PASFXA-04A-01R	-0.15	0.25	Pink	Circle
PANRDC-04A-01R	-0.10	0.25	Blue	Circle
PANRDC-04A-02R	-0.15	0.15	Purple	Triangle
PANRDC-04A-01R	-0.15	0.10	Pink	Triangle
PANRDC-04A-02R	-0.15	0.05	Blue	Triangle
PANRDC-04A-01R	-0.15	0.00	Pink	Triangle
PANRDC-04A-02R	-0.15	-0.05	Blue	Triangle
PANRDC-04A-01R	-0.15	-0.10	Pink	Triangle
PANRDC-04A-02R	-0.15	-0.15	Blue	Triangle
PANRDC-04A-01R	-0.15	-0.20	Pink	Triangle
PANRDC-04A-02R	-0.15	-0.25	Blue	Triangle
PANRDC-04A-01R	-0.15	-0.30	Pink	Triangle
PANRDC-04A-02R	-0.15	-0.35	Blue	Triangle
PANRDC-04A-01R	-0.15	-0.40	Pink	Triangle
PANRDC-04A-02R	-0.15	-0.45	Blue	Triangle
PANRDC-04A-01R	-0.15	-0.50	Pink	Triangle
PANRDC-04A-02R	-0.15	-0.55	Blue	Triangle
PANRDC-04A-01R	-0.15	-0.60	Pink	Triangle
PANRDC-04A-02R	-0.15	-0.65	Blue	Triangle
PANRDC-04A-01R	-0.15	-0.70	Pink	Triangle
PANRDC-04A-02R	-0.15	-0.75	Blue	Triangle
PANRDC-04A-01R	-0.15	-0.80	Pink	Triangle
PANRDC-04A-02R	-0.15	-0.85	Blue	Triangle
PANRDC-04A-01R	-0.15	-0.90	Pink	Triangle
PANRDC-04A-02R	-0.15	-0.95	Blue	Triangle

●

PCA: RUVg output (k = 2)

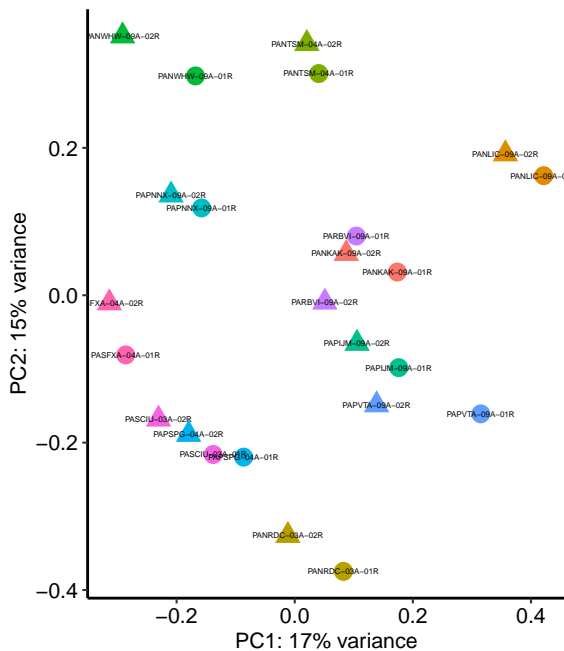


UMAP: RUVg output (k = 2)

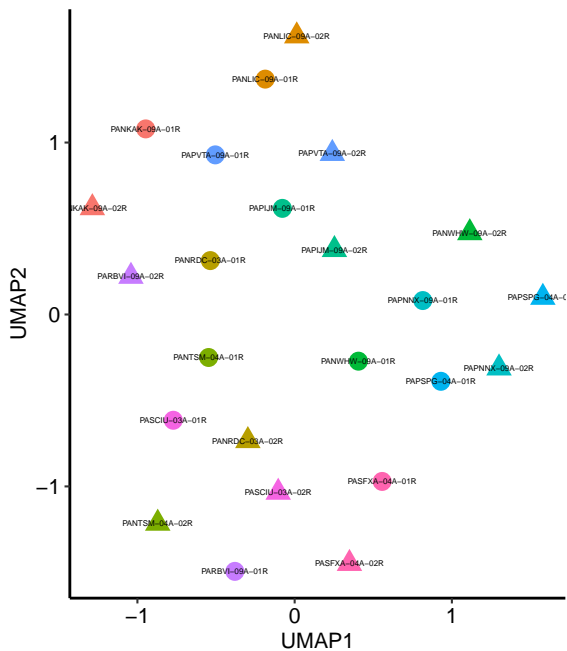


rna_library ● poly-A stranded ▲ stranded

PCA: RUVg output (k = 3)

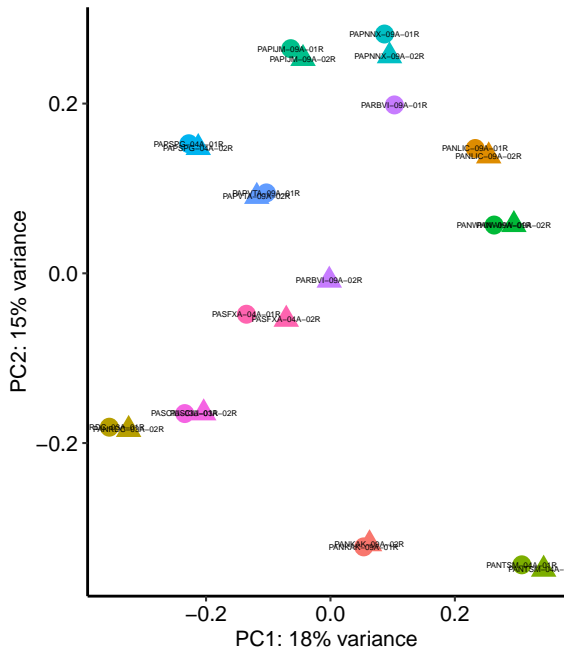


UMAP: RUVg output (k = 3)

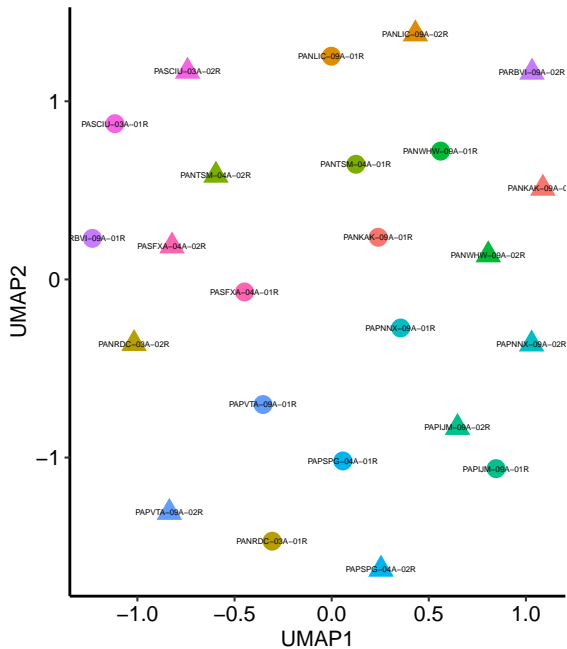


rna_library ● poly-A stranded ▲ stranded

PCA: RUVg output (k = 4)

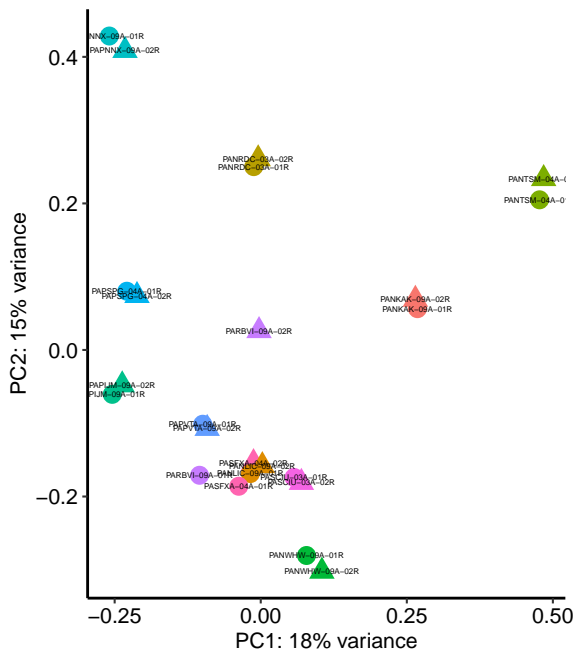


UMAP: RUVg output (k = 4)

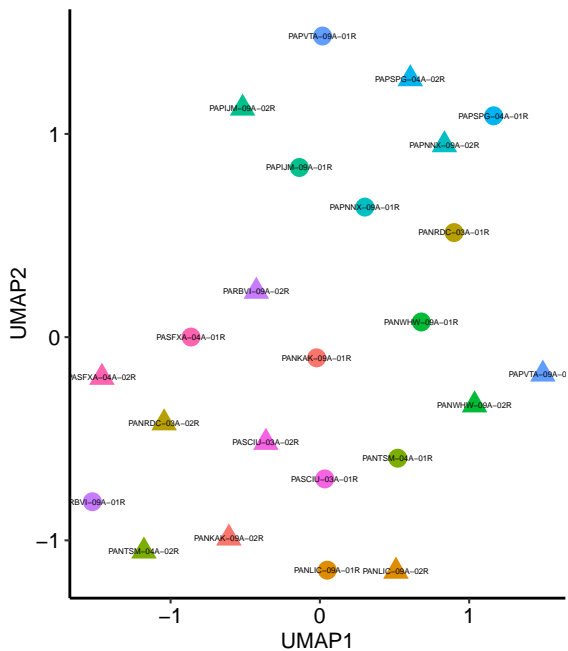


rna_library ● poly-A stranded ▲ stranded

PCA: RUVg output ($k = 5$)



UMAP: RUVg output ($k = 5$)



rna_library ● poly-A stranded ▲ stranded