

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, dark green).

Protein Variant	Approx. PC1 (19% variance)	Approx. PC2 (19% variance)	Family
PANRDC-03A-02R	0.05	0.95	PANK
PAPVTA-09A-02R	-0.15	0.85	PANK
PANKAK-09A-02R	-0.10	0.75	PANK
PASCIU-03A-02R	0.10	0.80	PANK
PANRDC-02A-01R	0.25	0.85	PANK
PAPVTA-09A-01R	0.05	0.65	PANK
PAPNNX-09A-02R	0.15	0.65	PANK
PANRDC-05A-02R	-0.15	0.60	PANK
PANRDC-06A-02R	-0.15	0.55	PANK
PANRDC-07A-02R	-0.15	0.50	PANK
PANRDC-08A-02R	-0.15	0.45	PANK
PANKAK-09A-01R	0.05	0.55	PANK
PANRDC-09A-01R	0.25	0.55	PANK
PANRDC-10A-01R	0.40	0.55	PANK
PANRDC-11A-01R	0.35	0.45	PANK
PANRDC-12A-01R	0.40	0.35	PANK
PANRDC-13A-01R	0.40	0.25	PANK
PANRDC-14A-01R	0.40	0.15	PANK
PANRDC-15A-01R	0.40	0.05	PANK
PANRDC-16A-01R	0.40	0.00	PANK
PANRDC-17A-01R	0.40	0.05	PANK
PANRDC-18A-01R	0.40	0.15	PANK
PANRDC-19A-01R	0.40	0.25	PANK
PANRDC-20A-01R	0.40	0.35	PANK
PANRDC-21A-01R	0.40	0.45	PANK
PANRDC-22A-01R	0.40	0.55	PANK
PANRDC-23A-01R	0.40	0.65	PANK
PANRDC-24A-01R	0.40	0.75	PANK
PANRDC-25A-01R	0.40	0.85	PANK
PANRDC-26A-01R	0.40	0.95	PANK
PANRDC-27A-01R	0.40	1.05	PANK
PANRDC-28A-01R	0.40	1.15	PANK
PANRDC-29A-01R	0.40	1.25	PANK
PANRDC-30A-01R	0.40	1.35	PANK
PANRDC-31A-01R	0.40	1.45	PANK
PANRDC-32A-01R	0.40	1.55	PANK
PANRDC-33A-01R	0.40	1.65	PANK
PANRDC-34A-01R	0.40	1.75	PANK
PANRDC-35A-01R	0.40	1.85	PANK
PANRDC-36A-01R	0.40	1.95	PANK
PANRDC-37A-01R	0.40	2.05	PANK
PANRDC-38A-01R	0.40	2.15	PANK
PANRDC-39A-01R	0.40	2.25	PANK
PANRDC-40A-01R	0.40	2.35	PANK
PANRDC-41A-01R	0.40	2.45	PANK
PANRDC-42A-01R	0.40	2.55	PANK
PANRDC-43A-01R	0.40	2.65	PANK
PANRDC-44A-01R	0.40	2.75	PANK
PANRDC-45A-01R	0.40	2.85	PANK
PANRDC-46A-01R	0.40	2.95	PANK
PANRDC-47A-01R	0.40	3.05	PANK
PANRDC-48A-01R	0.40	3.15	PANK
PANRDC-49A-01R	0.40	3.25	PANK
PANRDC-50A-01R	0.40	3.35	PANK
PANRDC-51A-01R	0.40	3.45	PANK
PANRDC-52A-01R	0.40	3.55	PANK
PANRDC-53A-01R	0.40	3.65	PANK
PANRDC-54A-01R	0.40	3.75	PANK
PANRDC-55A-01R	0.40	3.85	PANK
PANRDC-56A-01R	0.40	3.95	PANK
PANRDC-57A-01R	0.40	4.05	PANK
PANRDC-58A-01R	0.40	4.15	PANK
PANRDC-59A-01R	0.40	4.25	PANK
PANRDC-60A-01R	0.40	4.35	PANK
PANRDC-61A-01R	0.40	4.45	PANK
PANRDC-62A-01R	0.40	4.55	PANK
PANRDC-63A-01R	0.40	4.65	PANK
PANRDC-64A-01R	0.40	4.75	PANK
PANRDC-65A-01R	0.40	4.85	PANK
PANRDC-66A-01R	0.40	4.95	PANK
PANRDC-67A-01R	0.40	5.05	PANK
PANRDC-68A-01R	0.40	5.15	PANK
PANRDC-69A-01R	0.40	5.25	PANK
PANRDC-70A-01R	0.40	5.35	PANK
PANRDC-71A-01R			

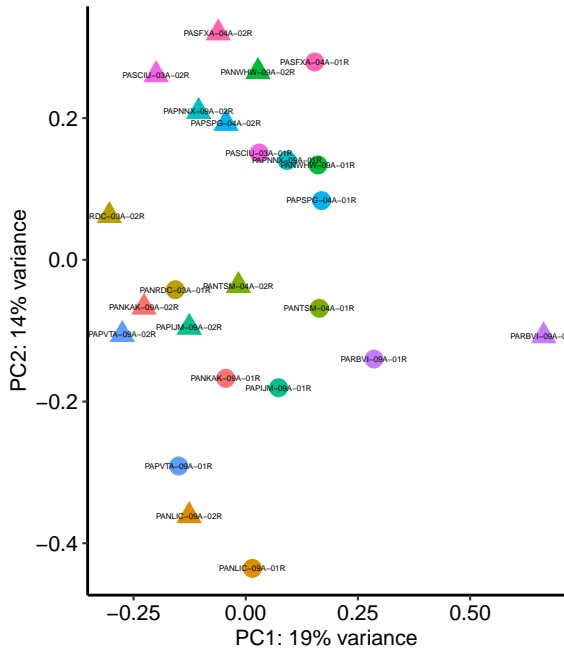
[illegible]

●

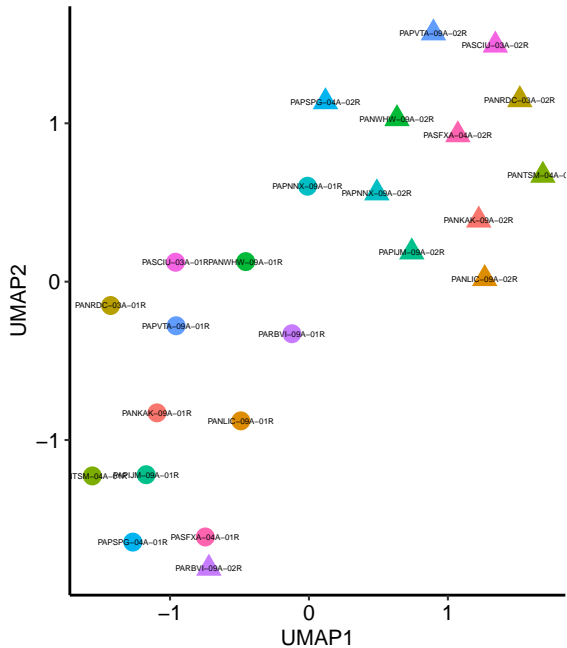
poly-A stranded

▲ stranded

PCA: RUVg output ( $k = 2$ )

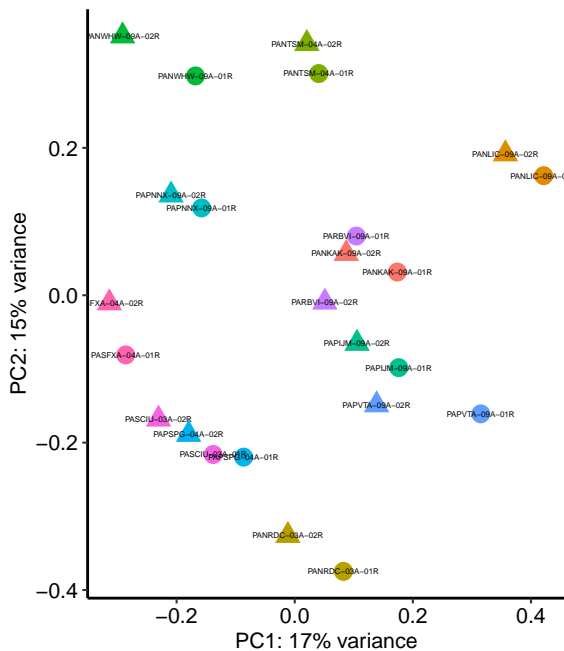


UMAP: RUVg output ( $k = 2$ )

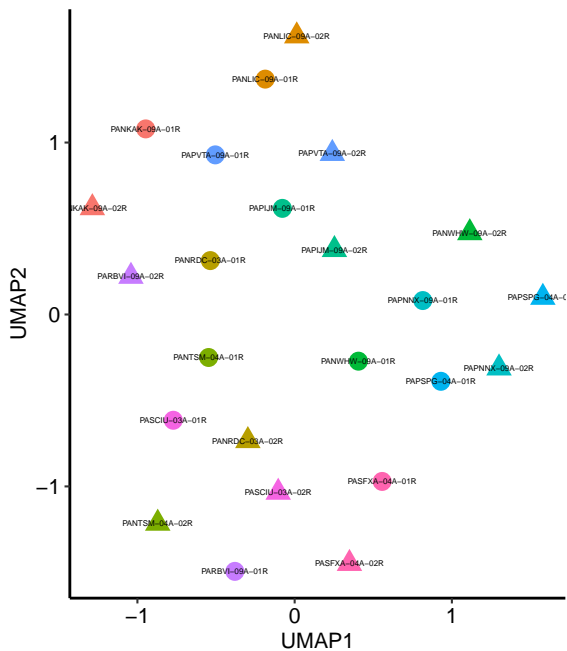


rna\_library   ●   poly-A stranded   ▲   stranded

PCA: RUVg output (k = 3)

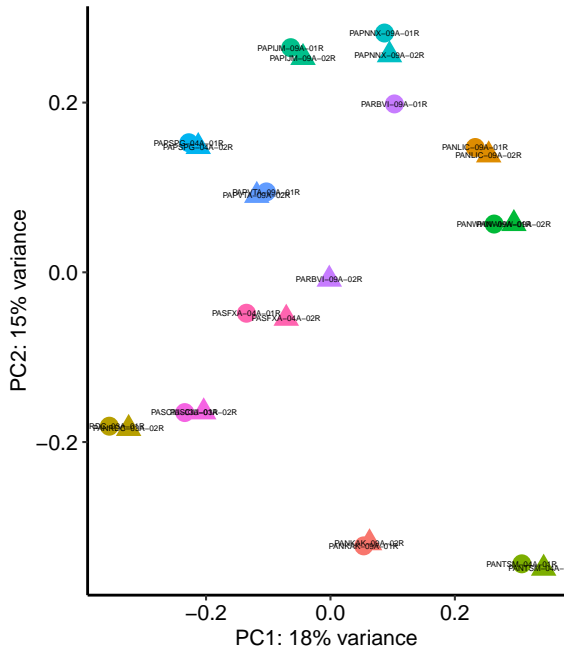


UMAP: RUVg output (k = 3)

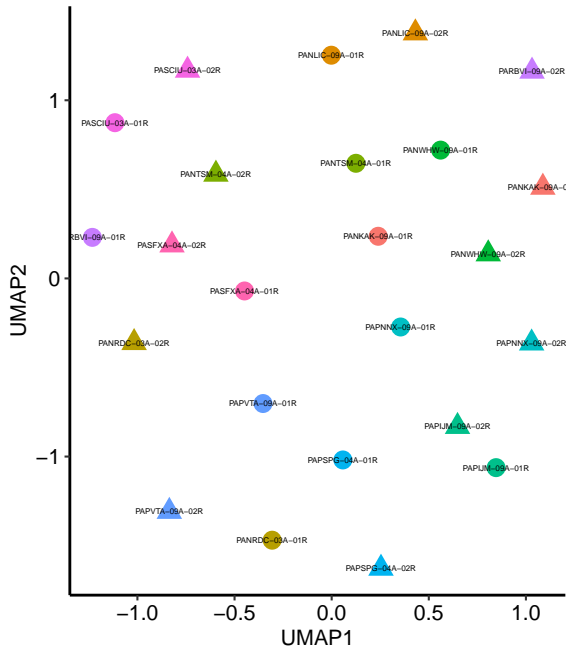


rna\_library ● poly-A stranded ▲ stranded

PCA: RUVg output ( $k = 4$ )




UMAP: RUVg output ( $k = 4$ )

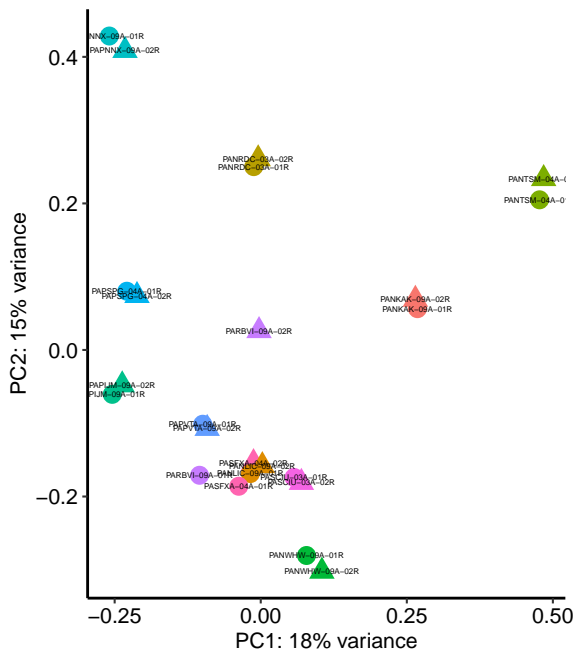


rna\_library

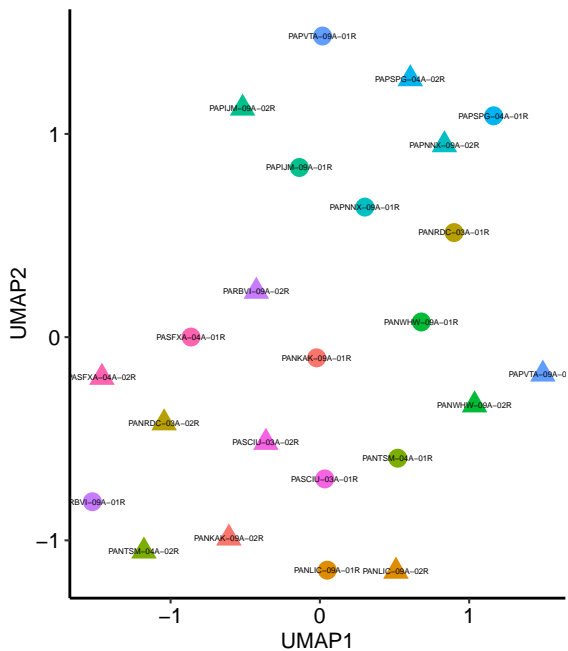
● poly-A stranded

 stranded

PCA: RUVg output (k = 5)



UMAP: RUVg output (k = 5)



rna\_library ● poly-A stranded ▲ stranded