

Supplemental Table S9. Linear regression results for each dataset. Asterisks (*) next to Coefficient P-value indicate a significant relationship between genome size and either gene content or tandem repeats using $\alpha = 0.1$. However, we note that in many cases, even when significant relationships are determined, the data are not well fit to the model (low r^2 value). PIC=phylogenetically independent contrasts.

Sampling	Variable	Dataset	p-value with PIC	r^2 with PIC	p-value w/o PIC	r^2 w/o PIC
Without outliers	Gene content (%)	Brassicales	0.00643*	0.1173	0.000818*	0.169
		Brassicaceae	0.00758*	0.1774	0.00192*	0.2262
		Capparaceae	0.506	0.49	0.282	0.5148
		Cleomaceae	0.00153*	0.6139	0.000818*	0.6208
	Tandem repeats (%)	Brassicales	0.281	0.01937	0.834	0.0007296
		Brassicaceae	0.454	0.01524	0.688	0.004299
		Capparaceae	0.409	0.6409	0.226	0.5985
		Cleomaceae	0.0241*	0.3833	0.000622*	0.6371
With outliers	Gene content (%)	Brassicales	0.0000000237*	0.4196	0.00000225*	0.2855
		Brassicaceae	0.0000000728*	0.4942	0.000000141*	0.4709
		Capparaceae	0.506	0.49	0.282	0.5148
		Cleomaceae	0.00153*	0.6139	0.000818*	0.6208
	Tandem repeats (%)	Brassicales	0.266	0.0187	0.406	0.01033
		Brassicaceae	0.422	0.01508	0.656	0.004543
		Capparaceae	0.409	0.6409	0.226	0.5985
		Cleomaceae	0.0241*	0.3833	0.000622*	0.6371