

Table 2. TF repertoires in the three main superkingdoms of life: Bacteria, Archaea, and Eukaryota

	Bacteria	Archaea	Eukaryota	Cellular organisms
Proteins				
Proteins per species	3140	1966	14 141	3885
Length of all proteins (residues)	322	289	465	328
Domains assigned per protein	1.41	1.30	1.53	1.42
Distinct domain families per protein	1.33	1.25	1.29	1.32
Length of protein domains (residues)	180	171	161	177
TFs				
TFs per species	131	60	325	155
Distinct architectures per species	39	19	45	39
Length of TFs (residues)	242	196	560	253
DBDs per TF	1.04	1.00	1.41	1.05
Distinct DBD families per TF	1.00	1.00	1.01	1.00
Length of DBDs (residues)	62	60	64	62
Partner domains per TF	0.58	0.25	0.24	0.49
Distinct partner domain families per TF	0.57	0.24	0.21	0.48
Length of partner domains (residues)	153	97	85	139
TF content in genome (%)	4.39	2.94	2.91	3.59
DBDs				
DBD families	61	15	77	131
Superkingdom-specific DBD families	43	0	69	
Partner domain families	228	55	795	938
Superkingdom-specific partner domain families	116	12	693	
Distinct domain architectures	605	118	2209	2779
DBDs per species	109	42	206	122
Distinct DBD families per species	23	12	27	24

Domain assignments are from Pfam. Median values of all species in each lineage are displayed. Mean values and their SDs for each property are described in ‘Supplementary Data’.