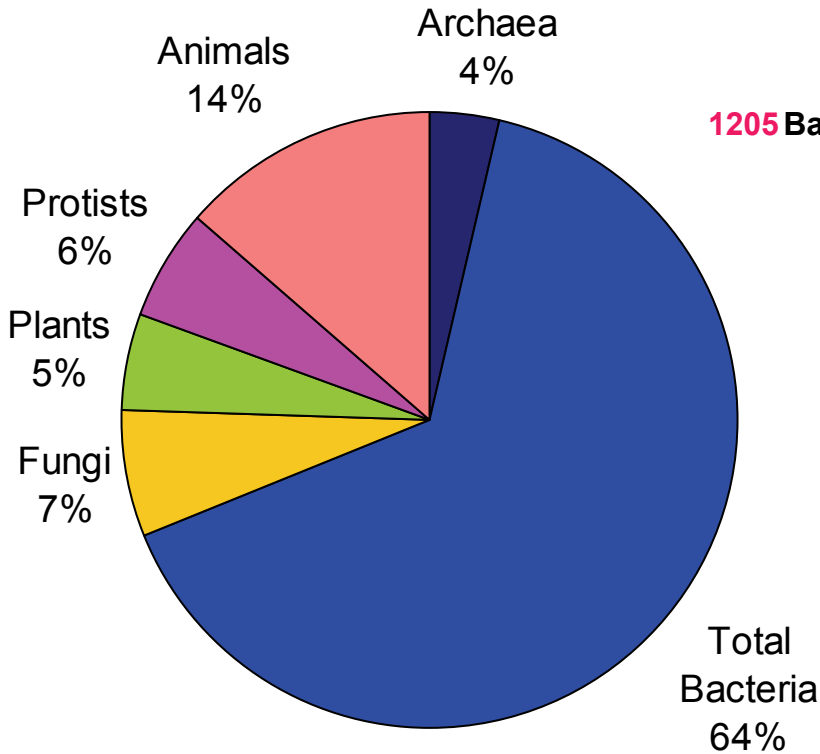


Dataset 2 (KEGG)

a) Number of protein representatives

b) Number of species representatives



97	Archaea	47	Fungi
30	Crenarchaeota	37	Ascomycetes
63	Euryarchaeota	9	Basidiomycetes
1	Korarchaeota	1	Microsporidians
1	Nanoarchaeota		
2	Thaumarchaeota	14	Plants
1205	Bacteria		
5	Acidobacteria	5	Eudicots
122	Actinobacteria	4	Green algae
137	Alphaproteobacteria	3	Monocots
46	Bacteroidetes	1	Mosses
92	Betaproteobacteria	1	Red algae
16	Chlamydiae	31	Protists
1	Chrysiogenetes	16	Alveolates
40	Cyanobacteria	1	Amoeboflagellate
12	Deinococcus	3	Amoebozoa
39	Deltaproteobacteria	1	Choanoflagellates
2	Elusimicrobia	2	Diatoms
35	Epsilonproteobacteria	1	Diplomonads
1	Fibrobacteres	5	Euglenozoa
253	Firmicutes	1	Oomycetes
5	Fusobacteria	1	Parabasalids
275	Gammaproteobacteria	48	Animals
1	Gemmatimonadetes	21	Arthropods
15	Green nonsulfur bacteria	1	Ascidians
11	Green sulfur bacteria	2	Cnidarians
28	Hyperthermophilic bacteria	1	Echinoderms
1	Other proteobacteria	1	Flatworms
5	Planctomycetes	1	Lancelets
22	Spirochaetes	3	Nematodes
2	Synergistetes	1	Placozoans
35	Tenericutes	17	Vertebrates
4	Verrucomicrobia		