Evolution of patchily distributed proteins shared between eukaryotes and prokaryotes: *Dictyostelium* as a case study.

Jan O. Andersson

Supplementary Table 1

Complete list of the frequency of occurrence of different genera in the 49 patchily distributed protein families.

#datasets	Domain[a]	Genus	Group
25	E	*Naegleria*	Heterolobosea (Excavata)
12	E	*Malawimonas*	Malawimonas (Excavata)
11	B	*Bacteroides*	Bacteroidetes/Chlorobi group
11	B	*Burkholderia*	β-Proteobacteria
11	B	*Clostridium*	Firmicutes
11	B	*Pseudomonas*	γ-Proteobacteria
10	B	*Myxococcus*	δ-Proteobacteria
10	E	*Tetrahymena*	Ciliophora (Chromalveolata)
9	B	*Bacillus*	Firmicutes
9	B	*Mycobacterium*	Actinobacteria
9	B	*Stigmatella*	δ-Proteobacteria
9	B	*Streptomyces*	Actinobacteria
9	B	*Vibrio*	γ-Proteobacteria
9	E	*Gibberella*	Fungi (Opisthokonta)
9	E	*Monosiga*	Choanomonada (Opisthokonta)
8	B	*Frankia*	Actinobacteria
8	B	*Herpetosiphon*	Chloroflexi
8	B	*Plesiocystis*	δ-Proteobacteria
8	B	*Pseudoalteromonas*	γ-Proteobacteria
8	B	*Ralstonia*	β-Proteobacteria
8	E	*Aspergillus*	Fungi (Opisthokonta)
8	E	*Phaeosphaeria*	Fungi (Opisthokonta)
7	B	*Flavobacterium*	Bacteroidetes/Chlorobi group
7	B	*Hahella*	γ-Proteobacteria
7	B	*Legionella*	γ-Proteobacteria
7	B	*Saccharopolyspora*	Actinobacteria
7	B	*Shewanella*	γ-Proteobacteria
7	E	*Hartmannella*	Tubulinea (Amoebozoa)
6	B	*Bradyrhizobium*	α-Proteobacteria
6	B	*Chromobacterium*	β-Proteobacteria
6	B	*Planctomyces*	Planctomycetes
6	B	*Rhodopirellula*	Planctomycetes
6	E	*Chaetomium*	Fungi (Opisthokonta)
6	E	*Histiona*	Jakobida (Excavata)
6	V	*Paramecium*	Phycodnaviridae
5	B	*Acidobacteria*	Fibrobacteres/Acidobacteria group
5	B	*Candidatus*	Planctomycetes
5	B	*Escherichia*	γ-Proteobacteria
5	B	*Geobacter*	δ-Proteobacteria

5	B	*Lactobacillus*	Firmicutes
5	B	*marine*	Actinobacteria
5	B	*Microscilla*	Bacteroidetes/Chlorobi group
5	B	*Nocardioides*	Actinobacteria
5	B	*Photobacterium*	γ-Proteobacteria
5	B	*Rhodopseudomonas*	α-Proteobacteria
5	B	*Saccharophagus*	γ-Proteobacteria
5	B	*Salinispora*	Actinobacteria
5	B	*uncultured*	environmental samples
5	E	*Acanthamoeba*	Acantamoebidae (Amoebozoa)
5	E	*Arabidopsis*	Land plants (Archaeplastida)
5	E	*Chlamydomonas*	Green algae (Archaeplastida)
5	E	*Coccidioides*	Fungi (Opisthokonta)
5	E	*Jakoba*	Jakobida (Excavata)
5	E	*Leishmania*	Euglenozoa (Excavata)
5	E	*Neurospora*	Fungi (Opisthokonta)
5	E	*Ostreococcus*	Green algae (Archaeplastida)
5	E	*Paramecium*	Ciliophora (Chromalveolata)
5	E	*Physarum*	Eumycetozoa (Amoebozoa)
5	E	*Phytophthora*	Oomycetes (Chromalveloata)
5	E	*Reclinomonas*	Jakobida (Excavata)
4	A	*Methanococcus*	Euryarchaeota
4	B	*Aeromonas*	γ-Proteobacteria
4	B	*Arthrobacter*	Actinobacteria
4	B	*Blastopirellula*	Planctomycetes
4	B	*Caulobacter*	α-Proteobacteria
4	B	*Fervidobacterium*	Thermotogae
4	B	*Francisella*	γ-Proteobacteria
4	B	*Methylobacterium*	α-Proteobacteria
4	B	*Roseobacter*	α-Proteobacteria
4	B	*Ruminococcus*	Firmicutes
4	B	*Silicibacter*	α-Proteobacteria
4	B	*Solibacter*	Fibrobacteres/Acidobacteria group
4	B	*Synechococcus*	Cyanobacteria
4	B	*unidentified*	Bacteroidetes/Chlorobi group
4	B	*Xanthomonas*	γ-Proteobacteria
4	E	*Ajellomyces*	Fungi (Opisthokonta)
4	E	*Hyperamoeba*	Eumycetozoa (Amoebozoa)
4	E	*Magnaporthe*	Fungi (Opisthokonta)
4	E	*Neosartorya*	Fungi (Opisthokonta)
4	E	*Thalassiosira*	Diatoms (Chromalveolata)
4	E	*Trichomonas*	Parabasalids (Excavata)
4	E	*Trypanosoma*	Euglenozoa (Excavata)
3	A	*Methanosarcina*	Euryarchaeota
3	A	*Natronomonas*	Euryarchaeota
3	B	*Alteromonas*	γ-Proteobacteria
3	B	*Anaeromyxobacter*	δ-Proteobacteria
3	B	*Azoarcus*	β-Proteobacteria
3	B	*Bdellovibrio*	δ-Proteobacteria
3	B	*Campylobacter*	ε-Proteobacteria
3	B	*Coxiella*	γ-Proteobacteria
3	B	*Deinococcus*	Deinococcus-Thermus
3	B	*Desulfitobacterium*	Firmicutes
3	B	*Dorea*	Firmicutes
3	B	*Enterobacter*	γ-Proteobacteria
3	B	*Erythrobacter*	α-Proteobacteria

3	B	*Exiguobacterium*	Firmicutes
3	B	*Flavobacteriales*	Bacteroidetes/Chlorobi group
3	B	*Fusobacterium*	Fusobacteria
3	B	*Gramella*	Bacteroidetes/Chlorobi group
3	B	*Janibacter*	Actinobacteria
3	B	*Lentisphaera*	Chlamydiae/Verrucomicrobia group
3	B	*Leptospira*	Spirochaetes
3	B	*Marinobacter*	γ-Proteobacteria
3	B	*Marinomonas*	γ-Proteobacteria
3	B	*Nitrobacter*	α-Proteobacteria
3	B	*Nostoc*	Cyanobacteria
3	B	*Oceanobacillus*	Firmicutes
3	B	*Oceanobacter*	γ-Proteobacteria
3	B	*Opitutaceae*	Chlamydiae/Verrucomicrobia group
3	B	*Polaribacter*	Bacteroidetes/Chlorobi group
3	B	*Prochlorococcus*	Cyanobacteria
3	B	*Psychroflexus*	Bacteroidetes/Chlorobi group
3	B	*Rhodococcus*	Actinobacteria
3	B	*Rhodoferax*	β-Proteobacteria
3	B	*Roseovarius*	α-Proteobacteria
3	B	*Serratia*	γ-Proteobacteria
3	B	*Stenotrophomonas*	γ-Proteobacteria
3	B	*Thermobifida*	Actinobacteria
3	B	*Thermotoga*	Thermotogae
3	B	*Xanthobacter*	α-Proteobacteria
3	B	*Yersinia*	γ-Proteobacteria
3	E	*Botryotinia*	Fungi (Opisthokonta)
3	E	*Cryptococcus*	Fungi (Opisthokonta)
3	E	*Euglena*	Euglenozoa (Excavata)
3	E	*Mastigamoeba*	Pelobionts (Amoebozoa)
3	E	*Oryza*	Land plants (Archaeplastida)
3	E	*Phaeodactylum*	Diatoms (Chromalveolata)
3	E	*Plasmodium*	Apicomplexa (Chromalveolata)
3	E	*Seculamonas*	Jakobida (Excavata)
3	E	*Tetraodon*	Animals (Opisthokonta)
3	E	*Vitis*	Land plants (Archaeplastida)
3	V	*Acanthamoeba*	Mimivirus
3	V	*Aeromonas*	Caudovirales
2	A	*Halobacterium*	Euryarchaeota
2	A	*Halorubrum*	Euryarchaeota
2	A	*Methanocaldococcus*	Euryarchaeota
2	A	*Pyrococcus*	Euryarchaeota
2	A	*Thermococcus*	Euryarchaeota
2	B	*Acinetobacter*	γ-Proteobacteria
2	B	*Agrobacterium*	α-Proteobacteria
2	B	*Alcanivorax*	γ-Proteobacteria
2	B	*Algoriphagus*	Bacteroidetes/Chlorobi group
2	B	*Alkaliphilus*	Firmicutes
2	B	*Alteromonadales*	γ-Proteobacteria
2	B	*Anabaena*	Cyanobacteria
2	B	*Azotobacter*	γ-Proteobacteria
2	B	*Beggiatoa*	γ-Proteobacteria
2	B	*Bordetella*	β-Proteobacteria
2	B	*Brevibacterium*	Actinobacteria
2	B	*Butyrivibrio*	Firmicutes
2	B	*Clavibacter*	Actinobacteria

2	B	*Corynebacterium*	Actinobacteria
2	B	*Croceibacter*	Bacteroidetes/Chlorobi group
2	B	*Crocosphaera*	Cyanobacteria
2	B	*Cyanothece*	Cyanobacteria
2	B	*Cytophaga*	Bacteroidetes/Chlorobi group
2	B	*Desulfotalea*	δ-Proteobacteria
2	B	*Dokdonia*	Bacteroidetes/Chlorobi group
2	B	*Enterococcus*	Firmicutes
2	B	*Flavobacteria*	Bacteroidetes/Chlorobi group
2	B	*Halothermothrix*	Firmicutes
2	B	*Herminiimonas*	β-Proteobacteria
2	B	*Hyphomonas*	α-Proteobacteria
2	B	*Kineococcus*	Actinobacteria
2	B	*Klebsiella*	γ-Proteobacteria
2	B	*Leeuwenhoekiella*	Bacteroidetes/Chlorobi group
2	B	*Limnobacter*	β-Proteobacteria
2	B	*Lyngbya*	Cyanobacteria
2	B	*Magnetococcus*	Proteobacteria
2	B	*Mariprofundus*	Proteobacteria
2	B	*Mesorhizobium*	α-Proteobacteria
2	B	*Methylococcus*	γ-Proteobacteria
2	B	*Nitrosococcus*	γ-Proteobacteria
2	B	*Nocardia*	Actinobacteria
2	B	*Oceanicaulis*	α-Proteobacteria
2	B	*Pediococcus*	Firmicutes
2	B	*Pedobacter*	Bacteroidetes/Chlorobi group
2	B	*Polaromonas*	β-Proteobacteria
2	B	*Porphyromonas*	Bacteroidetes/Chlorobi group
2	B	*Robiginitalea*	Bacteroidetes/Chlorobi group
2	B	*Roseiflexus*	Chloroflexi
2	B	*Rubrobacter*	Actinobacteria
2	B	*Salinibacter*	Bacteroidetes/Chlorobi group
2	B	*Stappia*	α-Proteobacteria
2	B	*Streptococcus*	Firmicutes
2	B	*Syntrophobacter*	δ-Proteobacteria
2	B	*Thermosinus*	Firmicutes
2	B	*Thermosipho*	Thermotogae
2	B	*Thermus*	Deinococcus-Thermus
2	B	*Trichodesmium*	Cyanobacteria
2	B	*Vibrionales*	γ-Proteobacteria
2	B	*Victivallis*	Chlamydiae/Verrucomicrobia group
2	B	*Xylella*	γ-Proteobacteria
2	E	*Batrachochytrium*	Fungi (Opisthokonta)
2	E	*Bigelowiella*	Cerceozoa (Rhizaria)
2	E	*Capsaspora*	Ichthyosporea (Opisthokonta)
2	E	*Coprinopsis*	Fungi (Opisthokonta)
2	E	*Cryptosporidium*	Apicomplexa (Chromalveolata)
2	E	*Danio*	Animals (Opisthokonta)
2	E	*Entamoeba*	Entamoebidae (Amoebozoa)
2	E	*Guillardia*	Cryptomonads (Chromalveolata)
2	E	*Medicago*	Land plants (Archaeplastida)
2	E	*Mus*	Animals (Opisthokonta)
2	E	*Pavlova*	Haptophytes (Chromalveolata)
2	E	*Rattus*	Animals (Opisthokonta)
2	E	*Sawyeria*	Heterolobosea (Excavata)
2	E	*Sclerotinia*	Fungi (Opisthokonta)

2	E	*Trimastix*	Preaxostyla (Excavata)
2	V	*Antheraea*	Baculoviridae
2	V	*Enterobacteria*	Caudovirales
2	V	*Listeria*	Caudovirales
2	V	*Spodoptera*	Baculoviridae
2	V	*Vibrio*	Caudovirales
1	A	*Archaeoglobus*	Euryarchaeota
1	A	*Ferroplasma*	Euryarchaeota
1	A	*Haloarcula*	Euryarchaeota
1	A	*Methanobrevibacter*	Euryarchaeota
1	A	*Methanocorpusculum*	Euryarchaeota
1	A	*Methanoculleus*	Euryarchaeota
1	A	*Methanopyrus*	Euryarchaeota
1	A	*Methanosphaera*	Euryarchaeota
1	A	*Methanospirillum*	Euryarchaeota
1	A	*Methanothermobacter*	Euryarchaeota
1	A	*Methanothermus*	Euryarchaeota
1	A	*Picrophilus*	Euryarchaeota
1	A	*Sulfolobus*	Crenarchaeota
1	A	*Thermoplasma*	Euryarchaeota
1	B	*Acidiphilium*	α-Proteobacteria
1	B	*Actinobacillus*	γ-Proteobacteria
1	B	*Actinomyces*	Actinobacteria
1	B	*alpha*	α-Proteobacteria
1	B	*Aster*	Firmicutes
1	B	*Aurantimonas*	α-Proteobacteria
1	B	*Bifidobacterium*	Actinobacteria
1	B	*Caldibacillus*	Firmicutes
1	B	*Caldicellulosiruptor*	Firmicutes
1	B	*Caminibacter*	ε-Proteobacteria
1	B	*Cellvibrio*	γ-Proteobacteria
1	B	*Chloroflexus*	Chloroflexi
1	B	*Chromohalobacter*	γ-Proteobacteria
1	B	*Collinsella*	Actinobacteria
1	B	*Colwellia*	γ-Proteobacteria
1	B	*Comamonas*	β-Proteobacteria
1	B	*Congregibacter*	γ-Proteobacteria
1	B	*Dehalococcoides*	Chloroflexi
1	B	*Delftia*	β-Proteobacteria
1	B	*Desulfotomaculum*	Firmicutes
1	B	*Desulfovibrio*	δ-Proteobacteria
1	B	*Erwinia*	γ-Proteobacteria
1	B	*Eubacterium*	Firmicutes
1	B	*Ewingella*	γ-Proteobacteria
1	B	*Geobacillus*	Firmicutes
1	B	*Gloeobacter*	Cyanobacteria
1	B	*Gluconobacter*	α-Proteobacteria
1	B	*Haemophilus*	γ-Proteobacteria
1	B	*Helicobacter*	ε-Proteobacteria
1	B	*Idiomarina*	γ-Proteobacteria
1	B	*Janthinobacterium*	β-Proteobacteria
1	B	*Lactococcus*	Firmicutes
1	B	*Lawsonia*	δ-Proteobacteria
1	B	*Leifsonia*	Actinobacteria
1	B	*Leuconostoc*	Firmicutes

1	B	*Listonella*	γ-Proteobacteria
1	B	*Magnetospirillum*	α-Proteobacteria
1	B	*Maricaulis*	α-Proteobacteria
1	B	*Mesoplasma*	Firmicutes
1	B	*Moritella*	γ-Proteobacteria
1	B	*Mycoplasma*	Firmicutes
1	B	*Nitrosomonas*	β-Proteobacteria
1	B	*Nitrosospira*	β-Proteobacteria
1	B	*Nodularia*	Cyanobacteria
1	B	*Oceanicola*	α-Proteobacteria
1	B	*Oenococcus*	Firmicutes
1	B	*Parabacteroides*	Bacteroidetes/Chlorobi group
1	B	*Paracoccus*	α-Proteobacteria
1	B	*Parvibaculum*	α-Proteobacteria
1	B	*Parvularcula*	α-Proteobacteria
1	B	*Pelobacter*	δ-Proteobacteria
1	B	*Pelotomaculum*	Firmicutes
1	B	*Petrotoga*	Thermotogae
1	B	*Photorhabdus*	γ-Proteobacteria
1	B	*Psychromonas*	γ-Proteobacteria
1	B	*Reinekea*	γ-Proteobacteria
1	B	*Rhodobacter*	α-Proteobacteria
1	B	*Rhodobacterales*	α-Proteobacteria
1	B	*Rickettsia*	α-Proteobacteria
1	B	*Rickettsiella*	α-Proteobacteria
1	B	*Salinivibrio*	γ-Proteobacteria
1	B	*Salmonella*	γ-Proteobacteria
1	B	*Sodalis*	γ-Proteobacteria
1	B	*Sphingopyxis*	α-Proteobacteria
1	B	*Staphylococcus*	Firmicutes
1	B	*Synechocystis*	Cyanobacteria
1	B	*Syntrophus*	δ-Proteobacteria
1	B	*Thermoanaerobacter*	Firmicutes
1	B	*Thermoanaerobacterium*	Firmicutes
1	B	*Thermomonospora*	Actinobacteria
1	B	*Thiobacillus*	β-Proteobacteria
1	B	*Treponema*	Spirochaetes
1	B	*Wolinella*	ε-Proteobacteria
1	B	*Zymomonas*	α-Proteobacteria
1	E	*Aedes*	Animals (Opisthokonta)
1	E	*Anopheles*	Animals (Opisthokonta)
1	E	*Antonospora*	Fungi (Opisthokonta)
1	E	*Bos*	Animals (Opisthokonta)
1	E	*Canis*	Animals (Opisthokonta)
1	E	*Cyanophora*	Glaucophyta (Archaeplastida)
1	E	*Equus*	Animals (Opisthokonta)
1	E	*Gallus*	Animals (Opisthokonta)
1	E	*Giardia*	Diplomonads (Excavata)
1	E	*Glycine*	Land plants (Archaeplastida)
1	E	*Homo*	Animals (Opisthokonta)
1	E	*Isochrysis*	Haptophytes (Chromalveolata)
1	E	*Karlodinium*	Dinoflagellates (Chromalveolata)
1	E	*Macaca*	Animals (Opisthokonta)
1	E	*Micromonas*	Green algae (Archaeplastida)
1	E	*Monodelphis*	Animals (Opisthokonta)

1	E	*Mortierella*	Fungi (Opisthokonta)
1	E	*Mythimna*	Animals (Opisthokonta)
1	E	*Neocallimastix*	Fungi (Opisthokonta)
1	E	*Ornithorhynchus*	Animals (Opisthokonta)
1	E	*Oxytricha*	Ciliophora (Chromalveolata)
1	E	*Polysphondylium*	Eumycetozoa (Amoebozoa)
1	E	*Saccharum*	Land plants (Archaeplastida)
1	E	*Strongylocentrotus*	Animals (Opisthokonta)
1	E	*Tribolium*	Animals (Opisthokonta)
1	E	*Ustilago*	Fungi (Opisthokonta)
1	E	*Xenopus*	Animals (Opisthokonta)
1	E	*Yarrowia*	Fungi (Opisthokonta)
1	V	*Agrotis*	Baculoviridae
1	V	*Anomala*	Poxviridae
1	V	*Anticarsia*	Baculoviridae
1	V	*Autographa*	Baculoviridae
1	V	*Bombyx*	Baculoviridae
1	V	*Chlorella*	Phycodnaviridae
1	V	*Choristoneura*	Poxviridae
1	V	*Chrysodeixis*	Baculoviridae
1	V	*Clanis*	Baculoviridae
1	V	*Cydia*	Baculoviridae
1	V	*Ecotropis*	Baculoviridae
1	V	*Emiliania*	Phycodnaviridae
1	V	*Epiphyas*	Baculoviridae
1	V	*Erwinia*	Caudovirales
1	V	*Helicoverpa*	Baculoviridae
1	V	*Hyphantria*	Baculoviridae
1	V	*Leucania*	Baculoviridae
1	V	*Lymantria*	Baculoviridae
1	V	*Mamestra*	Baculoviridae
1	V	*Mannheimia*	Caudovirales
1	V	*Maruca*	Baculoviridae
1	V	*Orgyia*	Baculoviridae
1	V	*Rhodothermus*	Caudovirales
1	V	*Trichoplusia*	Ascoviridae
1	V	*Xestia*	Baculoviridae

[a]) A: Archaea; B: bacteria; E: Eukaryota; V: Viruses.

Evolution of patchily distributed proteins shared between eukaryotes and prokaryotes: *Dictyostelium* as a case study.

Jan O. Andersson

Supplementary Table 2

The properties of the 49 patchily distributed genes compared to the 9952 *D. discoideum* genes coding for proteins more than 200 amino acids in length.

Property	49 selected genes	All genes
Average length (aa)	451	669
Median length (aa)	424	516
Introns/gene	1,41	1,41
Fraction of intronless genes	36,7%	29,5%
Average G+C content	28,5%	27,6%
Median G+C content	28,4%	27,4%
Average GC3s content	10,9%	12,4%
Median GC3s content	10,6%	11,0%