

Aeropyrum pernix	Mycoplasma genitalium
Pyrobaculum aerophilum	Agrobacterium tumefaciens
Archaeoglobus fulgidus	Neisseria meningitidis
Haloarcula marismortui	Azoarcus sp
Methanobacterium thermoautotrophicum	Bdellovibrio bacteriovorus
Methanococcus jannaschii	Desulfotalea psychrophila
Methanosarcina acetivorans	Campylobacter jejuni
Methanopyrus kandleri	Helicobacter pylori
Pyrococcus abyssi	Acinetobacter calcoaceticus
Picrophilus torridus	Escherichia coli
Thermoplasma acidophilum	Leptospira interrogans
Nanoarchaeum equitans	Borrelia burgdorferi
Mycobacterium bovis	Plasmodium falciparum
Symbiobacterium thermophilum	Trypanosoma cruzi
Aquifex aeolicus	Saccharomyces cerevisiae
Bacteroides fragilis	Schizosaccharomyces pombe
Chlamydia pneumoniae	Anopheles gambiae
Chlorobium tepidum	Drosophila melanogaster
Dehalococcoides ethenogenes	Homo sapiens
Synechococcus sp	Mus musculus
Anabaena sp	Caenorhabditis briggsae
Deinococcus radiodurans	Caenorhabditis elegans
Bacillus subtilis	Cyanidioschyzon merolae
Clostridium acetobutylicum	Thalassiosira pseudonana
Enterococcus faecalis	Arabidopsis thaliana

Table S1. Fully sequenced genomes in the COGENT database used in the analysis. The total number of genomes was 50: 12 (24%) from Archaea, 25 (50%) from Bacteria and 13 (26%) from Eukaryota.