Supplementary Materials
to **"An update on clostridial phylogeny: Gram-negative spore-formers and other misplaced clostridia"** by **Natalya Yutin** and **Michael Y. Galperin**

Phylogenetic tree construction

The phylogenetic tree in Fig. 1 and Fig. S1 was built from a concatenated alignment of 50 ribosomal proteins (L1-L7, L10, L11, L13-L24, L27-L29, L31-L36, S2-S20) with a total of 6,164 unambiguously aligned positions, essentially as described in (Yutin *et al.*, 2012). The ribosomal protein sequences used in (Yutin *et al.*, 2012) were supplemented with those from additional species, extracted from the RefSeq database (Pruitt *et al.*, 2012) and aligned using MUSCLE (Edgar, 2004). Positions including gaps in more than one-third of the sequences and positions with low information content were removed prior to tree computation, as described in Yutin *et al*. (Yutin *et al.*, 2008), which resulted in 6,164 unambiguously aligned positions. Preliminary maximum-likelihood trees were constructed from these alignments using the FastTree program (Price *et al.*, 2010) with default parameters (JTT evolutionary model, discrete gamma model with 20 rate categories). The preliminary tree and the alignment were then used in ProtTest (Darriba *et al.*, 2011) to determine the best substitution matrix. The final maximum-likelihood tree was constructed using TreeFinder [1,000 replicates, Search Depth 2 (Jobb *et al.*, 2004)], with the LG+G substitution matrix (which was found to be the best for a given alignment). The Expected-Likelihood Weights (ELW) of 1,000 local rearrangements were used as confidence values of TreeFinder tree branches.

The maximum-likelihood phylogenetic trees for RpoB (Fig. S2) and GyrB (Fig. S3) protein sequences have been constructed in the same way, from the alignments containing 1,154 and 631 unambiguously aligned positions, respectively. The only difference was the use of the LG+G+F substitution matrix, which was selected by ProTest as the best for both RpoB and GyrB alignments.

References

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2011) ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**: 1164-1165.

Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 113.

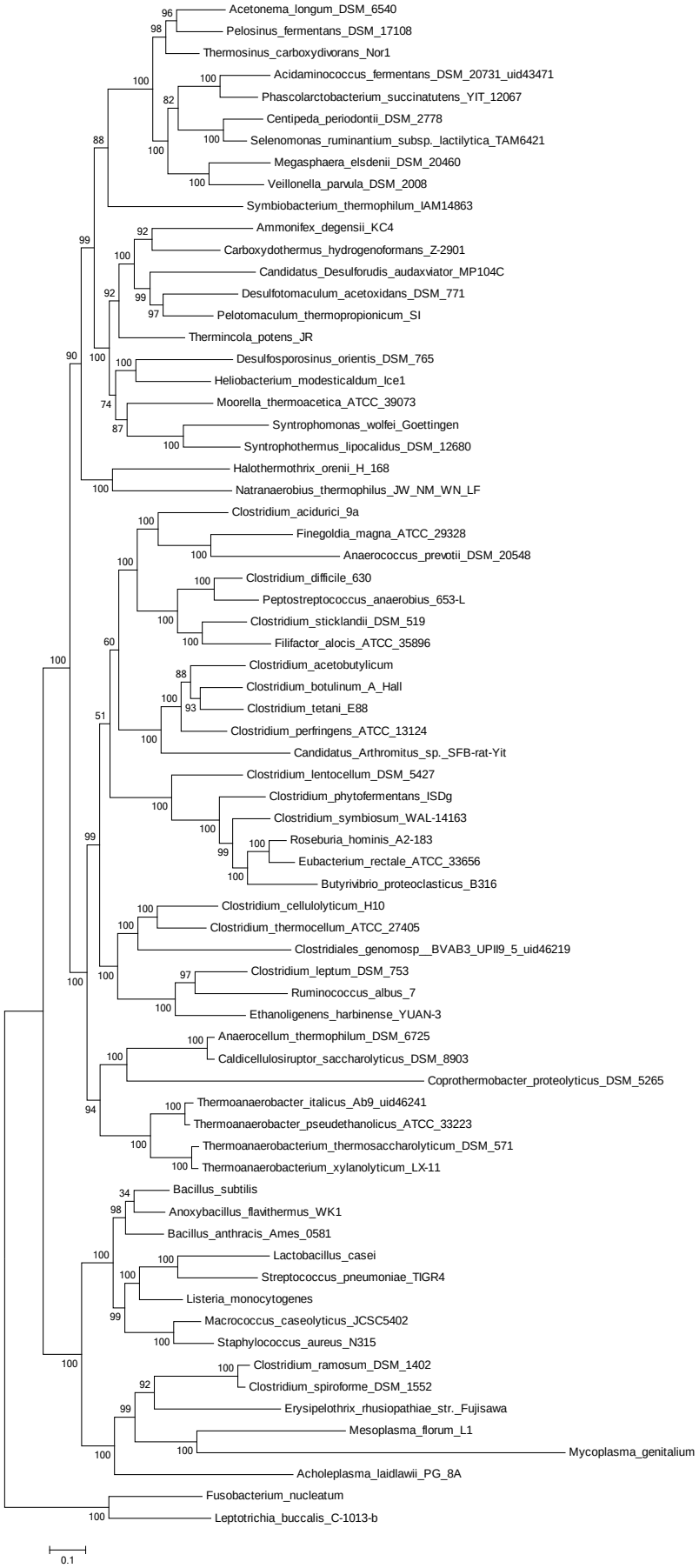
Jobb, G., von Haeseler, A., and Strimmer, K. (2004) TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol Biol* **4**: 18.

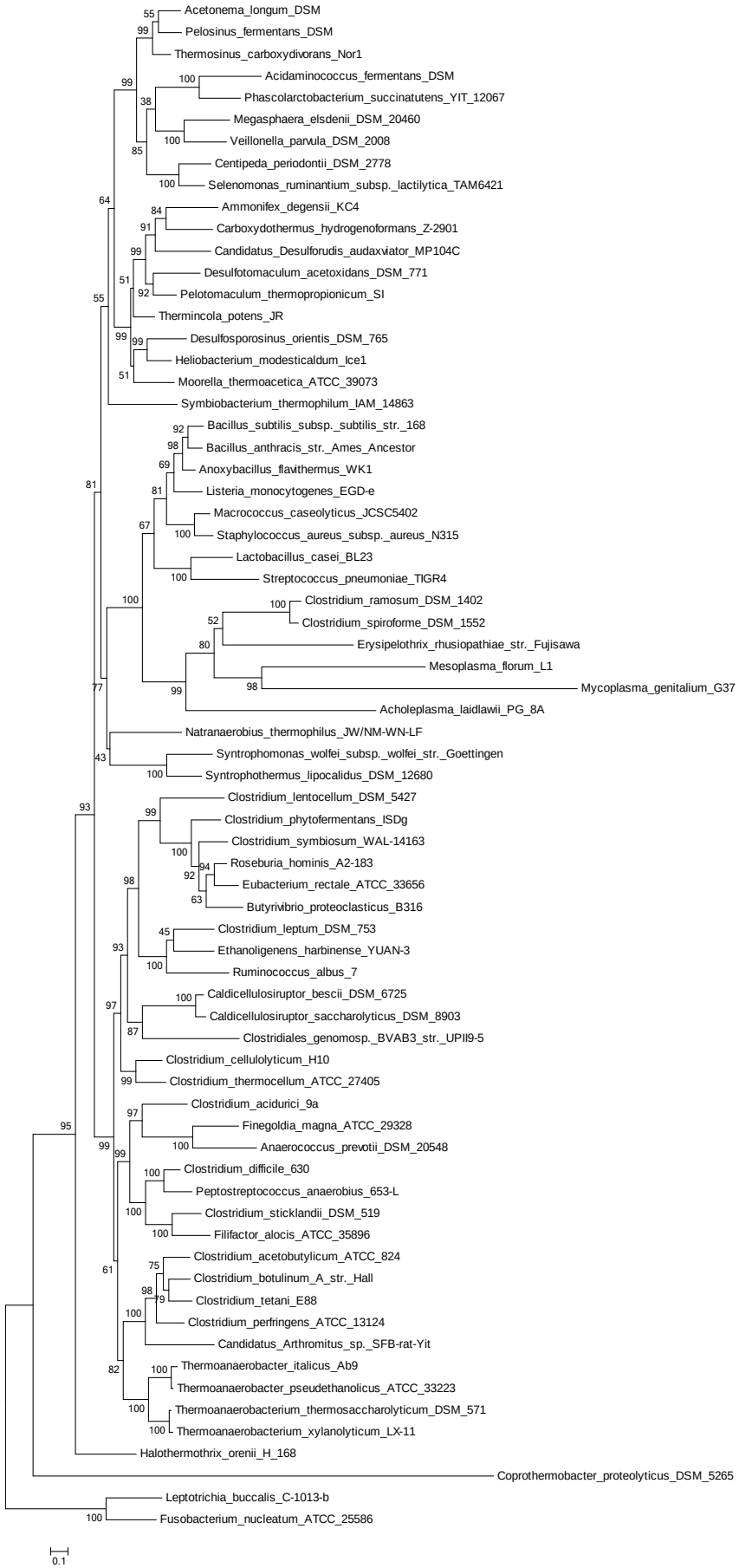
Price, M.N., Dehal, P.S., and Arkin, A.P. (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* **5**: e9490.

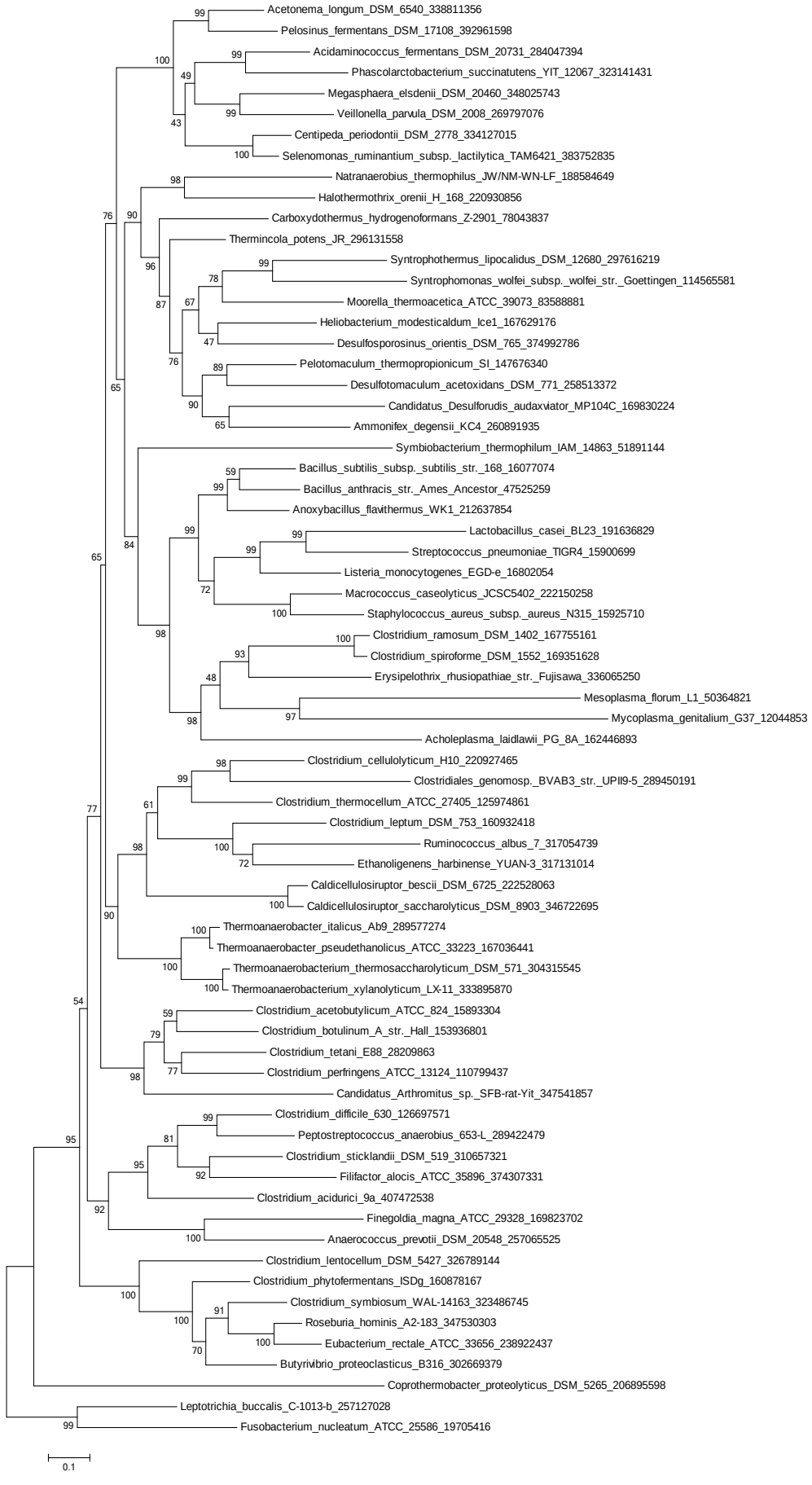
Pruitt, K.D., Tatusova, T., Brown, G.R., and Maglott, D.R. (2012) NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res* **40**: D130-D135.

Yutin, N., Makarova, K.S., Mekhedov, S.L., Wolf, Y.I., and Koonin, E.V. (2008) The deep archaeal roots of eukaryotes. *Mol Biol Evol* **25**: 1619-1630.

Yutin, N., Puigbo, P., Koonin, E.V., and Wolf, Y.I. (2012) Phylogenomics of prokaryotic ribosomal proteins. *PLoS One* **7**: e36972.


Acetonema_longum_DSM_6540
Pelosinus_fermentans_DSM_17108
Thermosinus_carboxydivorans_Nor1
Acidaminococcus_fermentans_DSM_20731_uid43471
Phascolarctobacterium_succinatutens_YIT_12067
Centipeda_periodontii_DSM_2778
Selenomonas_ruminantium_subsp._lactilytica_TAM6421
Megasphaera_elsdenii_DSM_20460
Veillonella_parvula_DSM_2008
Symbiobacterium_thermophilum_IAM14863
Ammonifex_degensii_KC4
Carboxydothermus_hydrogenoformans_Z-2901
Candidatus_Desulforudis_audaxviator_MP104C
Desulfotomaculum_acetoxidans_DSM_771
Pelotomaculum_thermopropionicum_SI
Thermincola_potens_JR
Desulfosporosinus_orientis_DSM_765
Heliobacterium_modesticaldum_Ice1
Moorella_thermoacetica_ATCC_39073
Syntrophomonas_wolfei_Goettingen
Syntrophothermus_lipocalidus_DSM_12680
Halothermothrix_orenii_H_168
Natranaerobius_thermophilus_JW_NM_WN_LF
Clostridium_acidurici_9a
Finegoldia_magna_ATCC_29328
Anaerococcus_prevotii_DSM_20548
Clostridium_difficile_630
Peptostreptococcus_anaerobius_653-L
Clostridium_sticklandii_DSM_519
Filifactor_alocis_ATCC_35896
Clostridium_acetobutylicum
Clostridium_botulinum_A_Hall
Clostridium_tetani_E88
Clostridium_perfringens_ATCC_13124
Candidatus_Arthromitus_sp._SFB-rat-Yit
Clostridium_lentocellum_DSM_5427
Clostridium_phytofermentans_ISDg
Clostridium_symbiosum_WAL-14163
Roseburia_hominis_A2-183
Eubacterium_rectale_ATCC_33656
Butyrivibrio_proteoclasticus_B316
Clostridium_cellulolyticum_H10
Clostridium_thermocellum_ATCC_27405
Clostridiales_genomosp__BVAB3_UPII9_5_uid46219
Clostridium_leptum_DSM_753
Ruminococcus_albus_7
Ethanoligenens_harbinense_YUAN-3
Anaerocellum_thermophilum_DSM_6725
Caldicellulosiruptor_saccharolyticus_DSM_8903
Coprothermobacter_proteolyticus_DSM_5265
Thermoanaerobacter_italicus_Ab9_uid46241
Thermoanaerobacter_pseudethanolicus_ATCC_33223
Thermoanaerobacterium_thermosaccharolyticum_DSM_571
Thermoanaerobacterium_xylanolyticum_LX-11
Bacillus_subtilis
Anoxybacillus_flavithermus_WK1
Bacillus_anthracis_Ames_0581
Lactobacillus_casei
Streptococcus_pneumoniae_TIGR4
Listeria_monocytogenes
Macrococcus_caseolyticus_JCSC5402
Staphylococcus_aureus_N315
Clostridium_ramosum_DSM_1402
Clostridium_spiroforme_DSM_1552
Erysipelothrix_rhusiopathiae_str._Fujisawa
Mesoplasma_florum_L1
Mycoplasma_genitalium
Acholeplasma_laidlawii_PG_8A
Fusobacterium_nucleatum
Leptotrichia_buccalis_C-1013-b
0.1


Acetonema_longum_DSM
Pelosinus_fermentans_DSM
Thermosinus_carboxydivorans_Nor1
Acidaminococcus_fermentans_DSM
Phascolarctobacterium_succinatutens_YIT_12067
Megasphaera_elsdenii_DSM_20460
Veillonella_parvula_DSM_2008
Centipeda_periodontii_DSM_2778
Selenomonas_ruminantium_subsp._lactilytica_TAM6421
Ammonifex_degensii_KC4
Carboxydothermus_hydrogenoformans_Z-2901
Candidatus_Desulforudis_audaxviator_MP104C
Desulfotomaculum_acetoxidans_DSM_771
Pelotomaculum_thermopropionicum_SI
Thermincola_potens_JR
Desulfosporosinus_orientis_DSM_765
Heliobacterium_modesticaldum_Ice1
Moorella_thermoacetica_ATCC_39073
Symbiobacterium_thermophilum_IAM_14863
Bacillus_subtilis_subsp._subtilis_str._168
Bacillus_anthracis_str._Ames_Ancestor
Anoxybacillus_flavithermus_WK1
Listeria_monocytogenes_EGD-e
Macrococcus_caseolyticus_JCSC5402
Staphylococcus_aureus_subsp._aureus_N315
Lactobacillus_casei_BL23
Streptococcus_pneumoniae_TIGR4
Clostridium_ramosum_DSM_1402
Clostridium_spiroforme_DSM_1552
Erysipelothrix_rhusiopathiae_str._Fujisawa
Mesoplasma_florum_L1
Mycoplasma_genitalium_G37
Acholeplasma_laidlawii_PG_8A
Natranaerobius_thermophilus_JW/NM-WN-LF
Syntrophomonas_wolfei_subsp._wolfei_str._Goettingen
Syntrophothermus_lipocalidus_DSM_12680
Clostridium_lentocellum_DSM_5427
Clostridium_phytofermentans_ISDg
Clostridium_symbiosum_WAL-14163
Roseburia_hominis_A2-183
Eubacterium_rectale_ATCC_33656
Butyrivibrio_proteoclasticus_B316
Clostridium_leptum_DSM_753
Ethanoligenens_harbinense_YUAN-3
Ruminococcus_albus_7
Caldicellulosiruptor_bescii_DSM_6725
Caldicellulosiruptor_saccharolyticus_DSM_8903
Clostridiales_genomosp._BVAB3_str._UPII9-5
Clostridium_cellulolyticum_H10
Clostridium_thermocellum_ATCC_27405
Clostridium_aciduricii_9a
Finegoldia_magna_ATCC_29328
Anaerococcus_prevotii_DSM_20548
Clostridium_difficile_630
Peptostreptococcus_anaerobius_653-L
Clostridium_sticklandii_DSM_519
Filifactor_alocis_ATCC_35896
Clostridium_acetobutylicum_ATCC_824
Clostridium_botulinum_A_str._Hall
Clostridium_tetani_E88
Clostridium_perfringens_ATCC_13124
Candidatus_Arthromitus_sp._SFB-rat-Yit
Thermoanaerobacter_italicus_Ab9
Thermoanaerobacter_pseudethanolicus_ATCC_33223
Thermoanaerobacterium_thermosaccharolyticum_DSM_571
Thermoanaerobacterium_xylanolyticum_LX-11
Halothermothrix_orenii_H_168
Coprothermobacter_proteolyticus_DSM_5265
Leptotrichia_buccalis_C-1013-b
Fusobacterium_nucleatum_ATCC_25586
0.1


Acetonema_longum_DSM_6540_338811356
Pelosinus_fermentans_DSM_17108_392961598
Acidaminococcus_fermentans_DSM_20731_284047394
Phascolarctobacterium_succinatutens_YIT_12067_323141431
Megasphaera_elsdenii_DSM_20460_348025743
Veillonella_parvula_DSM_2008_269797076
Centipeda_periodontii_DSM_2778_334127015
Selenomonas_ruminantium_subsp._lactilytica_TAM6421_383752835
Natranaerobius_thermophilus_JW/NM-WN-LF_188584649
Halothermothrix_orenii_H_168_220930856
Carboxydothermus_hydrogenoformans_Z-2901_78043837
Thermincola_potens_JR_296131558
Syntrophothermus_lipocalidus_DSM_12680_297616219
Syntrophomonas_wolfei_subsp._wolfei_str._Goettingen_114565581
Moorella_thermoacetica_ATCC_39073_83588881
Heliobacterium_modesticaldum_Ice1_167629176
Desulfosporosinus_orientis_DSM_765_374992786
Pelotomaculum_thermopropionicum_SI_147676340
Desulfotomaculum_acetoxidans_DSM_771_258513372
Candidatus_Desulforudis_audaxviator_MP104C_169830224
Ammonifex_degensii_KC4_260891935
Symbiobacterium_thermophilum_IAM_14863_51891144
Bacillus_subtilis_subsp._subtilis_str._168_16077074
Bacillus_anthracis_str._Ames_Ancestor_47525259
Anoxybacillus_flavithermus_WK1_212637854
Lactobacillus_casei_BL23_191636829
Streptococcus_pneumoniae_TIGR4_15900699
Listeria_monocytogenes_EGD-e_16802054
Macrococcus_caseolyticus_JCSC5402_222150258
Staphylococcus_aureus_subsp._aureus_N315_15925710
Clostridium_ramosum_DSM_1402_167755161
Clostridium_spiroforme_DSM_1552_169351628
Erysipelothrix_rhusiopathiae_str._Fujisawa_336065250
Mesoplasma_florum_L1_50364821
Mycoplasma_genitalium_G37_12044853
Acholeplasma_laidlawii_PG_8A_162446893
Clostridium_cellulolyticum_H10_220927465
Clostridiales_genomosp._BVAB3_str._UPII9-5_289450191
Clostridium_thermocellum_ATCC_27405_125974861
Clostridium_leptum_DSM_753_160932418
Ruminococcus_albus_7_317054739
Ethanoligenens_harbinense_YUAN-3_317131014
Caldicellulosiruptor_bescii_DSM_6725_222528063
Caldicellulosiruptor_saccharolyticus_DSM_8903_346722695
Thermoanaerobacter_italicus_Ab9_289577274
Thermoanaerobacter_pseudethanolicus_ATCC_33223_167036441
Thermoanaerobacterium_thermosaccharolyticum_DSM_571_304315545
Thermoanaerobacterium_xylanolyticum_LX-11_333895870
Clostridium_acetobutylicum_ATCC_824_15893304
Clostridium_botulinum_A_str._Hall_153936801
Clostridium_tetani_E88_28209863
Clostridium_perfringens_ATCC_13124_110799437
Candidatus_Arthromitus_sp._SFB-rat-Yit_347541857
Clostridium_difficile_630_126697571
Peptostreptococcus_anaerobius_653-L_289422479
Clostridium_sticklandii_DSM_519_310657321
Filifactor_alocis_ATCC_35896_374307331
Clostridium_acidurici_9a_407472538
Finegoldia_magna_ATCC_29328_169823702
Anaerococcus_prevotii_DSM_20548_257065525
Clostridium_lentocellum_DSM_5427_326789144
Clostridium_phytofermentans_ISDg_160878167
Clostridium_symbiosum_WAL-14163_323486745
Roseburia_hominis_A2-183_347530303
Eubacterium_rectale_ATCC_33656_238922437
Butyrivibrio_proteoclasticus_B316_302669379
Coprothermobacter_proteolyticus_DSM_5265_206895598
Leptotrichia_buccalis_C-1013-b_257127028
Fusobacterium_nucleatum_ATCC_25586_19705416
0.1