Summary of changes made in AGORA version 1.03, released 25.02.2019

Author: Almut Heinken

Additions in the updated version

In the present update, the 818 AGORA were expanded by several pathways and further curated against newly available experimental data. Moreover, an extensive correction of predictions that disagree with experimental data as well as quality control/quality assurance of reconstruction properties was performed with the help of a test suite (publication in preparation).

Experimental data from several recent publications (Table 1) was retrieved and served as the input for the data-driven curation and expansion of pathways in AGORA. First, 54 AGORA reconstructions were refined based on defined growth media reported by Tramontano et al [1] for the corresponding 54 organisms (Supplementary Table 1). Moreover, an improved pipeline was used to resolve false positive predictions of required nutrients in all reconstructions with available experimental data (see below). Growth on the media was enabled by combining the gap-filling reactions proposed by Tramontano et al. with gap-filling performed by the AGORA pipeline. Second, a comparative genomic analysis of a recently described pathway for aromatic amino acid degradation [2] was performed and the corresponding reactions were added to AGORA. Third, comparative genomic data was retrieved from a recent study on putrefaction pathways in gut microbes [3] and the corresponding reactions were reconstructed. Fourth, the 12α-hydroxysteroid dehydrogenase reaction, and the trans-4-hydroxy-L-proline dehydratase reaction were added based on a comparative genomic analysis of the respective, recently described genes [4, 5]. Fifth, novel carbon sources were added based on recent experimental evidence [6]. Sixth, 5-aminovalerate fermentation to valerate was reconstructed for *Clostridium viride* DSM 6836 [7]. Seventh, experimental data on B-vitamin secretion as well as secretion of vitamin K and GABA was gathered [8-23] and secretion of these compounds was enabled in the corresponding reconstructions. Finally, we recently performed a comparative genomic analysis of mucin degradation pathways in the gut microbiome [24]. The corresponding pathways were reconstructed with great biochemical detail.

For all added pathways (Table 1), it was also ensured that the added reactions could carry flux in every reconstruction.

Table 1: Pathways that were added in the new version, and supporting references.

Pathway added	**Number of analyzed AGORA organisms carrying pathway**	**Reference**
Putrefaction pathways	159	[3]
Aromatic amino acid degradation	14	[2]
12α-hydroxysteroid dehydrogenase reaction	37	[4]
trans-4-hydroxy-L-proline dehydratase reaction	21	[5]

5-aminovalerate fermentation to valerate	1	[7]
Mucin degradation pathway reconstructed through comparative genomic analysis	233	[24]
Experimentally determined carbon sources	12	[6]
B-vitamin, vitamin K, and GABA secretion based on experimental evidence	124	[8-23]

Testing and quality control/quality assurance of the reconstructions

A COBRA Toolbox-based test suite for the AGORA reconstructions (publication in preparation) was created that systematically accesses the capability of each reconstruction to capture known metabolic traits of the organism and determines features of the reconstructions that indicate their quality, e.g., mass and charge balance, blocked reactions, futile cycles, and leaking metabolites. The tested features are summarized in Table 2. The comparison with experimental data (e.g., carbon sources, fermentation products) and comparative genomics (e.g., aromatic amino acid degradation) was carried out as follows: The experimental or comparative genomics data serves as the input and the capability of each corresponding AGORA reconstruction to take up or produce the corresponding metabolite is tested. True positives indicate that the strain is known to take up or produce the metabolite and the corresponding AGORA reconstruction can also take or secrete the metabolite. False negatives indicate that the strain is known to have this capability but the corresponding reconstruction does not capture the trait. For growth requirements, two types of experimental information were available: nutrients that are known to be required by the organism in question, and nutrients that are known to not be required. This allowed us to additionally determine true negatives and false positives for growth requirements (Figure 1).

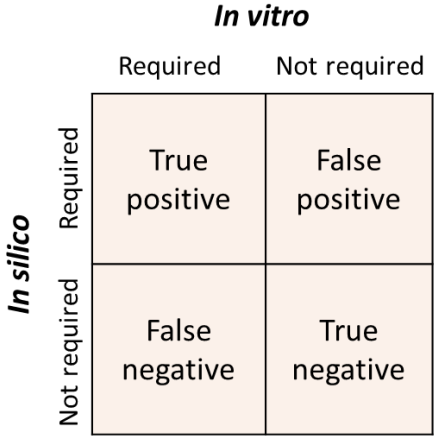

Figure 1: Schema of the four outcomes when comparing *in vivo* findings and *in silico* predictions of growth requirements for an organism.

Tests for reconstruction properties that determine biochemical and thermodynamic feasibility and quality of the reconstruction (e.g., biomass production, ATP production, mass-charge balance, blocked reactions)

were carried out using established COBRA Toolbox functions (Table 2). All models in AGORA 1.03 produced biomass and reasonable amounts of ATP on the Western Diet (Figure 2).

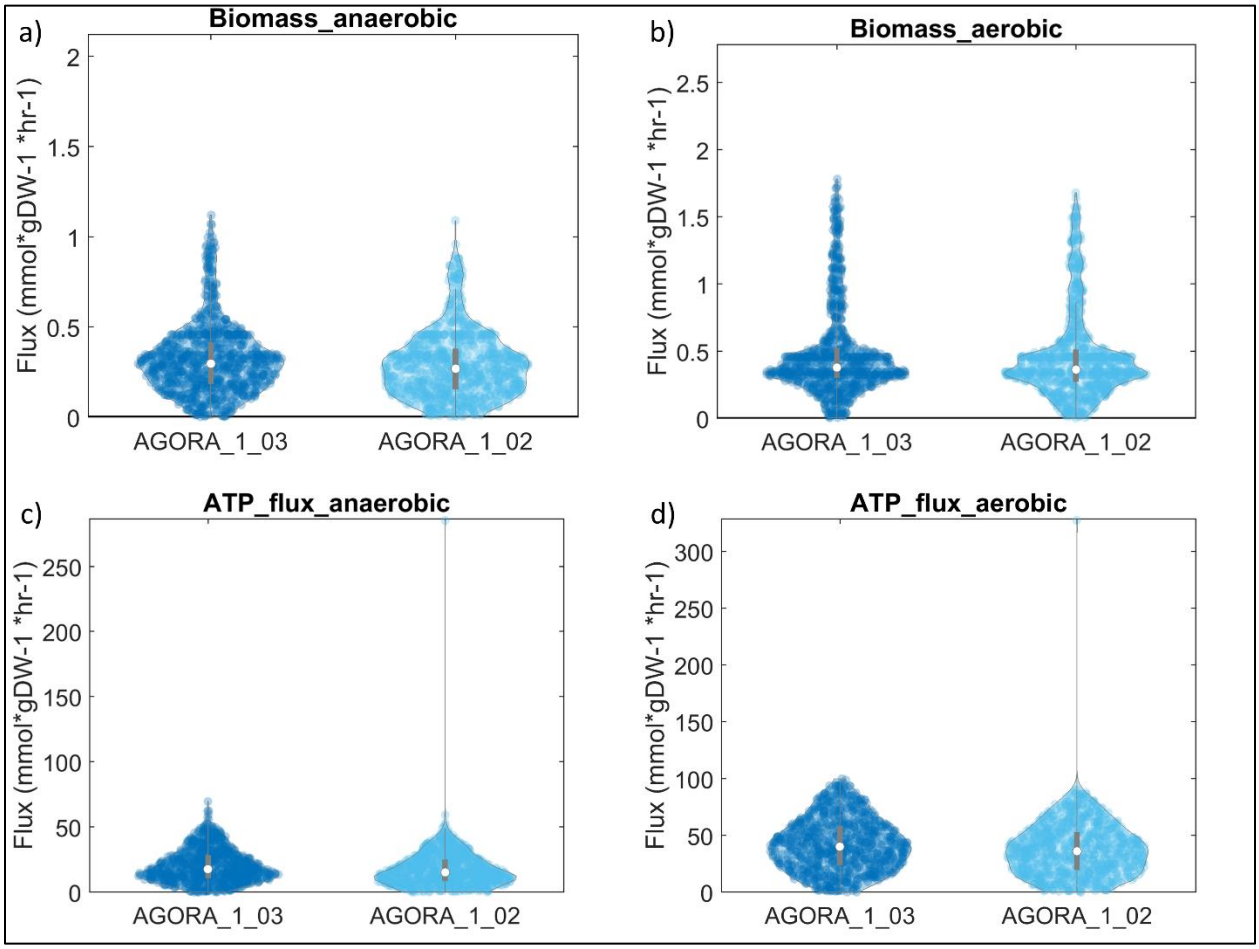

Figure 2: Model properties computed for the current version (AGORA 1.03) and the previous version (AGORA 1.02) on a simulated Western diet. a) Anaerobic biomass production, b) aerobic biomass production, c) anaerobic ATP production, d) aerobic ATP production.

Table 2: Summary of reconstruction features analyzed in the test suite for AGORA that was used to access the predictive potential of the reconstructions.

Feature	Input data	COBRA Toolbox function
Mass and charge balance	Reconstructions	checkMassChargeBalance
Leaking metabolites	Reconstructions	fastLeakTest
Blocked reactions	Reconstructions	identifyBlockedRxns
ATP production on Western diet	Reconstructions	optimizeCbModel
Biomass on Western diet	Reconstructions	optimizeCbModel
Carbon source usage	Reconstructions, experimental data	In preparation
Fermentation products	Reconstructions, experimental data	In preparation
Growth requirements	Reconstructions, experimental data	In preparation

Growth on defined medium according to experimental data	Reconstructions, experimental data	In preparation
B-vitamin biosynthesis	Reconstructions, comparative genomics, experimental data	In preparation
B-vitamin secretion	Reconstructions, experimental data	In preparation
4-hydroxyproline dehydration	Reconstructions, comparative genomics	In preparation
Bile acid deconjugation and conversion	Reconstructions, comparative genomics	In preparation
Putrefaction	Reconstructions, comparative genomics	In preparation
Aromatic amino acid degradation	Reconstructions, comparative genomics, experimental data	In preparation
Mucin degradation	Reconstructions, comparative genomics	In preparation

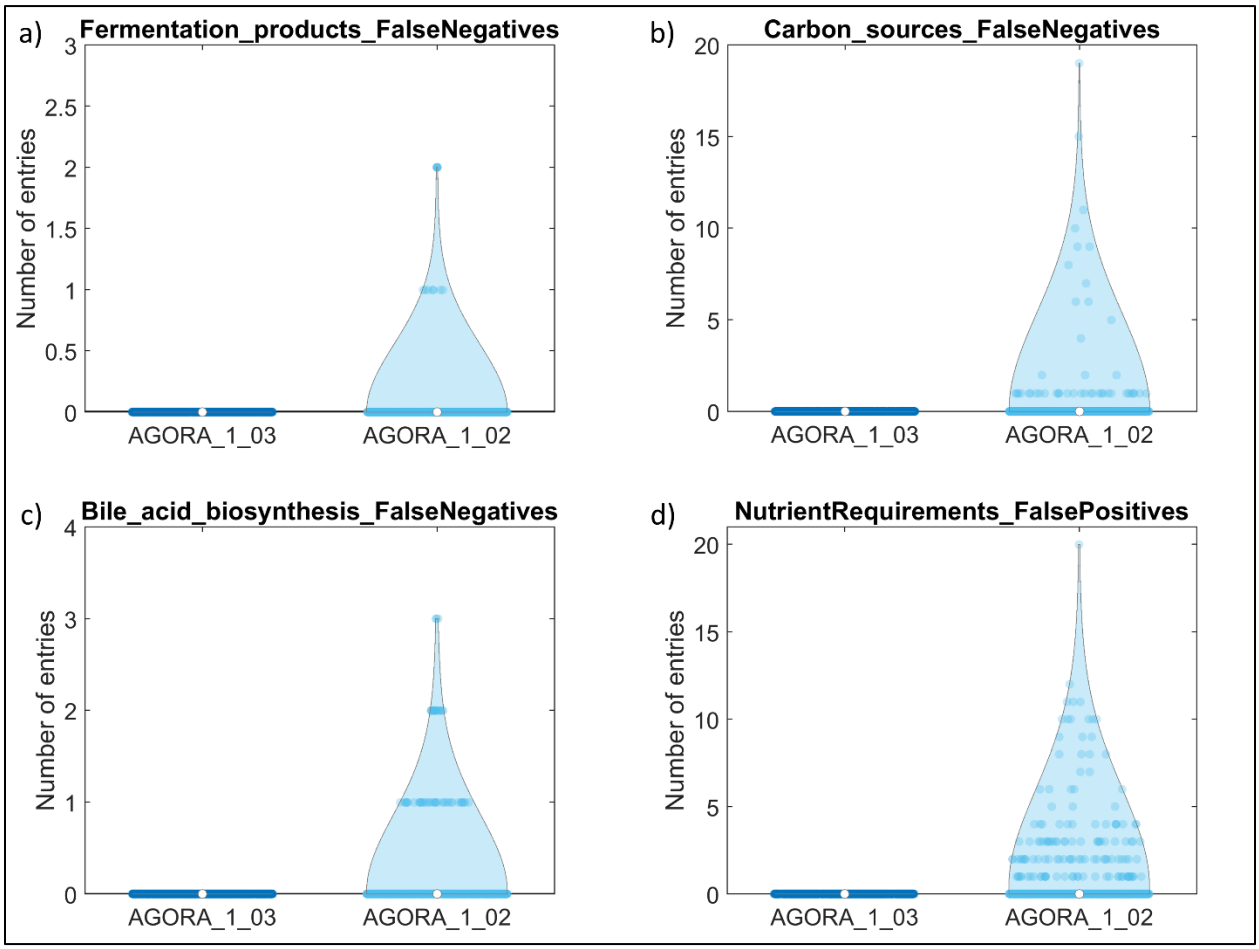

Figure 3: Comparison of false negative predictions for a) fermentation products, b) carbon sources, c) bile acids synthesized, and d) false negative predictions nutrient requirements in the current version (AGORA 1.03) and the previous version (AGORA 1.02).

The tests were carried out for the 818 reconstructions in the recent version (1.03) and the previous version (1.02). When performing the tests, a number of remaining false negative predictions for carbon sources and fermentation products, and false positives/ false negatives for growth requirements were found. Extensive curation of the corresponding AGORA reconstructions was performed to ensure that the corresponding reconstructions could take up all known carbon sources, produce all known fermentation products, and match known growth requirements. It was noted that many false negatives for growth requirements could not be corrected. This is due to an organism requiring a nutrient despite the biosynthesis pathway for the nutrient being present in its genome. Such discrepancies are challenging to curate against even in fully manually curated reconstructions [25, 26]. As a result of the extensive curation of growth requirements, all 279 AGORA 1.03 models with available experimental data were able to grow on the respective experimentally determined media compared with only 129 models in AGORA 1.02.

Overall, as a result of the additional curation, the number of true positives was increased and the number of false negatives was decreased in version 1.03 compared with version 1.02 (Figure 3, Table 3). The sensitivity for all pathways was 1 (Table 3) demonstrating that all false negative predictions with the exception of false negative growth requirements were eliminated in AGORA 1.03.

Table 3: Comparison in predictive potential in the current version (1.03) compared with the previous version (1.02) of AGORA. n.d.=not determined.

Feature	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
	AGORA 1.03 (01/2019)			AGORA 1.02 (02/2018)		
Aromatic amino acid degradation	1.00	n.d.	n.d.	0.20	n.d.	n.d.
Bile acid biosynthesis	1.00	n.d.	n.d.	0.87	n.d.	n.d.
Carbon sources	1.00	n.d.	n.d.	0.98	n.d.	n.d.
Fermentation products	1.00	n.d.	n.d.	0.99	n.d.	n.d.
Nutrient requirements	0.54	1.00	0.99	0.55	0.97	0.96
Putrefaction pathways	1.00	n.d.	n.d.	0.24	n.d.	n.d.
Vitamin secretion	1.00	n.d.	n.d.	0.07	n.d.	n.d.

Availability of AGORA version 1.03

Due to the large number of reactions added with the mucin degradation subsystem, two versions of the refined reconstructions (AGORA 1.03) are provided: one with and one without the mucin degradation subsystem. Both are available in SBML format at https://www.vmh.life/#downloadview).

Both versions are provided without dietary constraints. To enable users to simulate fluxes in AGORA on a diet, a tutorial has been created

(https://github.com/opencobra/COBRA.tutorials/tree/develop/analysis/simulateAGORAGrowthInDiets).

References

1. Tramontano M, Andrejev S, Pruteanu M, Klunemann M, Kuhn M, Galardini M, Jouhten P, Zelezniak A, Zeller G, Bork P, et al: **Nutritional preferences of human gut bacteria reveal their metabolic idiosyncrasies.** *Nat Microbiol* 2018, **3:**514-522.
2. Dodd D, Spitzer MH, Van Treuren W, Merrill BD, Hryckowian AJ, Higginbottom SK, Le A, Cowan TM, Nolan GP, Fischbach MA, Sonnenburg JL: **A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites.** *Nature* 2017, **551:**648-652.
3. Kaur H, Das C, Mande SS: **In Silico Analysis of Putrefaction Pathways in Bacteria and Its Implication in Colorectal Cancer.** *Front Microbiol* 2017, **8:**2166.
4. Doden H, Sallam LA, Devendran S, Ly L, Doden G, Daniel SL, Alves JMP, Ridlon JM: **Metabolism of Oxo-Bile Acids and Characterization of Recombinant 12alpha-Hydroxysteroid Dehydrogenases from Bile Acid 7alpha-Dehydroxylating Human Gut Bacteria.** *Appl Environ Microbiol* 2018, **84**.
5. Levin BJ, Huang YY, Peck SC, Wei Y, Martinez-Del Campo A, Marks JA, Franzosa EA, Huttenhower C, Balskus EP: **A prominent glycyl radical enzyme in human gut microbiomes metabolizes trans-4-hydroxy-l-proline.** *Science* 2017, **355**.
6. Desai MS, Seekatz AM, Koropatkin NM, Kamada N, Hickey CA, Wolter M, Pudlo NA, Kitamoto S, Terrapon N, Muller A, et al: **A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility.** *Cell* 2016, **167:**1339-1353 e1321.
7. Buckel W: **Unusual enzymes involved in five pathways of glutamate fermentation.** *Appl Microbiol Biotechnol* 2001, **57:**263-273.
8. Burgess CM, Smid EJ, Rutten G, van Sinderen D: **A general method for selection of riboflavin-overproducing food grade micro-organisms.** *Microb Cell Fact* 2006, **5:**24.
9. Morishita T, Tamura N, Makino T, Kudo S: **Production of menaquinones by lactic acid bacteria.** *J Dairy Sci* 1999, **82:**1897-1903.
10. Burgess C, O'Connell-Motherway M, Sybesma W, Hugenholtz J, van Sinderen D: **Riboflavin production in Lactococcus lactis: potential for in situ production of vitamin-enriched foods.** *Appl Environ Microbiol* 2004, **70:**5769-5777.
11. LeBlanc JG, Laino JE, del Valle MJ, Vannini V, van Sinderen D, Taranto MP, de Valdez GF, de Giori GS, Sesma F: **B-group vitamin production by lactic acid bacteria--current knowledge and potential applications.** *J Appl Microbiol* 2011, **111:**1297-1309.
12. Santos F, Wegkamp A, de Vos WM, Smid EJ, Hugenholtz J: **High-Level folate production in fermented foods by the B12 producer Lactobacillus reuteri JCM1112.** *Appl Environ Microbiol* 2008, **74:**3291-3294.
13. Arena MP, Russo P, Capozzi V, Lopez P, Fiocco D, Spano G: **Probiotic abilities of riboflavin-overproducing Lactobacillus strains: a novel promising application of probiotics.** *Appl Microbiol Biotechnol* 2014, **98:**7569-7581.
14. Pacheco Da Silva FF, Biscola V, LeBlanc JG, Gombossy de Melo Franco BD: **Effect of indigenous lactic acid bacteria isolated from goat milk and cheeses on folate and riboflavin content of fermented goat milk.** *LWT - Food Science and Technology* 2016, **71:**155-161.
15. Russo P, Capozzi V, Arena MP, Spadaccino G, Duenas MT, Lopez P, Fiocco D, Spano G: **Riboflavin-overproducing strains of Lactobacillus fermentum for riboflavin-enriched bread.** *Appl Microbiol Biotechnol* 2014, **98:**3691-3700.
16. Salvucci E, LeBlanc JG, Pérez G: **Technological properties of Lactic acid bacteria isolated from raw cereal material.** *LWT - Food Science and Technology* 2016, **70:**185-191.

17. Walther B, Karl JP, Booth SL, Boyaval P: **Menaquinones, bacteria, and the food supply: the relevance of dairy and fermented food products to vitamin K requirements.** *Adv Nutr* 2013, **4:**463-473.
18. Holasova M, Fiedlerova V, Roubal P, Pechacova M: *Biosynthesis of folates by lactic acid bacteria and propionibacteria in fermented milk.* 2004.
19. Pompei A, Cordisco L, Amaretti A, Zanoni S, Matteuzzi D, Rossi M: **Folate production by bifidobacteria as a potential probiotic property.** *Appl Environ Microbiol* 2007, **73:**179-185.
20. Linares DM, Gomez C, Renes E, Fresno JM, Tornadijo ME, Ross RP, Stanton C: **Lactic Acid Bacteria and Bifidobacteria with Potential to Design Natural Biofunctional Health-Promoting Dairy Foods.** *Front Microbiol* 2017, **8:**846.
21. LeBlanc JG, Chain F, Martin R, Bermudez-Humaran LG, Courau S, Langella P: **Beneficial effects on host energy metabolism of short-chain fatty acids and vitamins produced by commensal and probiotic bacteria.** *Microb Cell Fact* 2017, **16:**79.
22. Qing GP, L.: **Biosynthesis of Vitamins by Probiotic Bacteria, Probiotics and Prebiotics in Human Nutrition and Health.** *IntechOpen* 2016.
23. Strandwitz P, Kim KH, Terekhova D, Liu JK, Sharma A, Levering J, McDonald D, Dietrich D, Ramadhar TR, Lekbua A, et al: **GABA-modulating bacteria of the human gut microbiota.** *Nat Microbiol* 2018.
24. Ravcheev DA, Thiele I: **Comparative Genomic Analysis of the Human Gut Microbiome Reveals a Broad Distribution of Metabolic Pathways for the Degradation of Host-Synthetized Mucin Glycans and Utilization of Mucin-Derived Monosaccharides.** *Front Genet* 2017, **8:**111.
25. Teusink B, van Enckevort FH, Francke C, Wiersma A, Wegkamp A, Smid EJ, Siezen RJ: **In silico reconstruction of the metabolic pathways of Lactobacillus plantarum: comparing predictions of nutrient requirements with those from growth experiments.** *Appl Environ Microbiol* 2005, **71:**7253-7262.
26. Heinken A, Khan MT, Paglia G, Rodionov DA, Harmsen HJ, Thiele I: **A functional metabolic map of Faecalibacterium prausnitzii, a beneficial human gut microbe.** *J Bacteriol* 2014.

Supplementary Table 1: List of the 54 AGORA reconstructions that were curated against defined media determined for the corresponding organisms by Tramontano et al. [1].

AGORA Reconstruction ID	Strain	NCBI Taxonomy ID
Akkermansia_muciniphila_ATCC_BAA_835	Akkermansia muciniphila ATCC BAA-835	349741
Actinomyces_odontolyticus_ATCC_17982	Actinomyces odontolyticus ATCC 17982	411466
Alistipes_putredinis_DSM_17216	Alistipes putredinis DSM 17216	445970
Alistipes_shahii_WAL_8301	Alistipes shahii WAL 8301	717959
Bifidobacterium_adolescentis_ATCC_15703	Bifidobacterium adolescentis ATCC 15703	367928
Bifidobacterium_animalis_lactis_Bi_07	Bifidobacterium animalis lactis Bi-07	742729
Bifidobacterium_animalis_lactis_Bl_04_ATCC_SD5219	Bifidobacterium animalis lactis Bl-04, ATCC SD5219	580050
Bacteroides_caccae_ATCC_43185	Bacteroides caccae ATCC 43185	411901
Bacteroides_clarus_YIT_12056	Bacteroides clarus YIT 12056	762984
Bacteroides_coprocola_M16_DSM_17136	Bacteroides coprocola M16, DSM 17136	310298
Butyrivibrio_crossotus_DSM_2876	Butyrivibrio crossotus DSM 2876	511680
Bacteroides_dorei_DSM_17855	Bacteroides dorei DSM 17855	483217
Bacteroides_eggerthii_DSM_20697	Bacteroides eggerthii DSM 20697	483216
Bacteroides_fragilis_NCTC_9343	Bacteroides fragilis NCTC 9343	272559
Bacteroides_fragilis_3_1_12	Bacteroides fragilis 3_1_12	457424
Blautia_hansenii_VPI_C7_24_DSM_20583	Blautia hansenii VPI C7-24, DSM 20583	1322
Bifidobacterium_longum_infantis_ATCC_15697	Bifidobacterium longum infantis ATCC 15697	1682
Blautia_obeum_ATCC_29174	Blautia obeum ATCC 29174	411459
Bacteroides_ovatus_ATCC_8483	Bacteroides ovatus ATCC 8483	411476
Bacteroides_thetaiotaomicron_VPI_5482	Bacteroides thetaiotaomicron VPI-5482	226186
Bacteroides_uniformis_ATCC_8492	Bacteroides uniformis ATCC 8492	411479
Bacteroides_vulgatus_ATCC_8482	Bacteroides vulgatus ATCC 8482	435590
Collinsella_aerofaciens_ATCC_25986	Collinsella aerofaciens ATCC 25986	411903
Clostridium_bolteae_ATCC_BAA_613	Clostridium bolteae ATCC BAA-613	411902

Coprococcus_comes_ATCC_27758	Coprococcus comes ATCC 27758	470146
Clostridium_leptum_DSM_753	Clostridium leptum DSM 753	428125
Clostridium_perfringens_ATCC_13124	Clostridium perfringens ATCC 13124	195103
Clostridium_ramosum_VPI_0427_DSM_1402	Clostridium ramosum VPI 0427, DSM 1402	1547
Dorea_formicigenerans_ATCC_27755	Dorea formicigenerans ATCC 27755	411461
Desulfovibrio_piger_ATCC_29098	Desulfovibrio piger ATCC 29098	411464
Escherichia_coli_UTI89_UPEC	Escherichia coli UTI89 (UPEC)	364106
Eubacterium_eligens_ATCC_27750	Eubacterium eligens ATCC 27750	515620
Eggerthella_lenta_DSM_2243	Eggerthella lenta DSM 2243	479437
Eubacterium_siraeum_DSM_15702	Eubacterium siraeum DSM 15702	428128
Lactobacillus_acidophilus_NCFM	Lactobacillus acidophilus NCFM	272621
Lactobacillus_fermentum_ATCC_14931	Lactobacillus fermentum ATCC 14931	525325
Lactobacillus_gasseri_ATCC_33323	Lactobacillus gasseri ATCC 33323	324831
Lactococcus_lactis_subsp_lactis_Il1403	Lactococcus lactis subsp. lactis Il1403	272623
Lactobacillus_plantarum_WCFS1	Lactobacillus plantarum WCFS1	220668
Lactobacillus_salivarius_HO66_ATCC_11741	Lactobacillus salivarius HO66, ATCC 11741	1624
Lactobacillus_vaginalis_ATCC_49540	Lactobacillus vaginalis ATCC 49540	1423814
Odoribacter_splanchnicus_1651_6_DSM_20712	Odoribacter splanchnicus 1651/6, DSM 20712	28118
Pseudoflavonifractor_capillosus_strain_ATCC_29799	Pseudoflavonifractor capillosus strain ATCC 29799	411467
Prevotella_copri_CB7_DSM_18205	Prevotella copri CB7, DSM 18205	165179
Parabacteroides_distasonis_ATCC_8503	Parabacteroides distasonis ATCC 8503	435591
Prevotella_melaninogenica_ATCC_25845	Prevotella melaninogenica ATCC 25845	553174
Parabacteroides_merdae_ATCC_43184	Parabacteroides merdae ATCC 43184	411477
Ruminococcus_gnavus_ATCC_29149	Ruminococcus gnavus ATCC 29149	411470
Roseburia_hominis_A2_183	Roseburia hominis A2-183	585394
Roseburia_intestinalis_L1_82	Roseburia intestinalis L1-82	536231

Ruminococcus_torques_ATCC_27756	Ruminococcus torques ATCC 27756	411460
Salmonella_enterica_enterica_sv_Typhimurium_LT2	Salmonella enterica enterica sv Typhimurium LT2	1457319
Veillonella_parvula_Te3_DSM_2008	Veillonella parvula Te3, DSM 2008	29466
Yersinia_pseudotuberculosis_YPIII	Yersinia pseudotuberculosis YPIII	502800