Insights Obtained by Culturing Saccharibacteria With Their Bacterial Hosts

B. Bor, A.J. Collins, P.P. Murugkar, S. Balasubramanian, T.T. To, E.L. Hendrickson, J.K. Bedree, F.B. Bidlack, C.D. Johnston, W. Shi, J.S. McLean, X. He and F.E. Dewhirst

APPENDIX

Additional Materials and Methods

Phase contrast imaging.

All bacteria were cultured to early stationary phase under standard growth conditions. Cells were harvested, washed and resuspended in fresh medium to OD 600 of ~2.5. One µL of cell resuspension was spotted on a glass slide and covered with a cover slip. Cells were viewed using Nikon Eclipse E400 microscope equipped with a Nikon Plan Fluor × 100/1.30 oil immersion objective (Bor et al. 2016).

Fluorescence In Situ Hybridization (FISH) imaging.

FISH was carried out as previously described (Bor et al. 2016) with modifications. Briefly, the early stationary phase cells were fixed using 4% formaldehyde for 3 hours and permeabilized using 2 mg/mL lysozyme in 20 mM Tris pH 7.0 for 9 min at 37°C. Fixed cells were resuspended in 500 µL of hybridization buffer (20 mM Tris•Cl, pH 8.0, 0.9 M NaCl, 0.01% SDS, 30% deionized formamide) and incubated at 37°C for 30 minutes. A TM7-specific (TM7-567: Cy5-5′-CCTACGCAACTCTTTACGCC-3′) probe was hybridized with the cells for 3 hours at 42°C. Cells were then washed three times for 15 minutes each with 0.1x saline-sodium citrate buffer. During the second wash, the universal DNA stain SYTO®9 (Invitrogen) was added to the wash in 1:1000 dilution. Cells were mounted on the cover slip with SlowFade Gold antifade reagent (Invitrogen) and visualized with a Leica SPE I inverted confocal microscope equipped with an ACS APO 100x/1.15 oil CS immersion objective.

Scanning Electron Microscopy (SEM) imaging.

Round coverslips (12 mm) coated with poly-L-Lysine were placed in 24 well plates containing 1 mL TSBY broth in each well and 100 µL of 24-hour old host monocultures or binary cultures of hosts with their associated Saccharibacteria strains added. The coverslip was removed after 24 hours, fixed overnight in a solution of 4% glutaraldehyde, 3% paraformaldehyde, and 0.05% ruthenium red in 0.1 M Cacodylate buffer. The cover slips were then washed three times for 10 min each with 0.1M Cacodylate buffer and postfixed with 1% osmium tetroxide and 0.05% ruthenium red in 0.1M Cacodylate buffer for 1 hour. The coverslip was then washed twice for 10 minutes with distilled water and dehydrated for 10 minutes per step in an ethanol series (50%, 70%, 80%), ending with 95% ethanol, three times for 10 minutes each. The coverslip was critical point dried to maintain the shape of the cells, coated with 5 nm platinum-palladium and imaging data collected in a field emission SEM (Zeiss Ultra 55 FE-SEM) at 2kV at a working distance between 3 and 5mm using either secondary electron or in lens-detector.

PCR detection of Saccharibacteria.

A master mix was created using primers specific for Saccharibacteria (Brinig et al. 2003; Hugenholtz et al. 2001) and 1 µL of infected culture was used as template. Thermocycling conditions were: 95°C for 5 min, followed by 30 cycles of 95°C for 30s, 60°C for 30s and 72°C for 30s, followed by a final elongation step of 72°C for 2 min and an infinite hold at 4°C. Products were analyzed on a 1% agarose gel. Hosts were determined to be stably infected if they were PCR positive for Saccharibacteria through the end of the 5th passage in broth culture. Once the presence of the Saccharibacteria was confirmed, 100 µL of broth culture was serially diluted and plated onto agar medium to obtain single infected colonies. Multiple colonies were picked and cultivated in corresponding broth medium and the presence of Saccharibacteria in the coculture (containing host and single Saccharibacteria strain) was confirmed via full length 16S RNA sequencing. Once confirmed, frozen stocks of the binary culture containing the Saccharibacteria and its host were made and kept at -80°C for later use.

Genome assembly and annotation.

Ilumina sequences were assembled using SPAdes (version 3.9.0) (Nurk et al. 2013). The BBDuk (Bushnell 2015) tool for Geneious (v 9.1.8) was used to quality trim and filter Illumina adapters, artifacts, and phiX from reads. Paired reads with quality scores averaging less than 6 before trimming or with a length under 20 bp after trimming were discarded. PacBio sequencing reads were processed using Pacific Biosciences' SMRTlink pipeline version-6 with the HGAP assembly tool standard protocol. Single contigs generated through HGAP were also processed through Circlator version 1.5.5 using default settings to assign the start site of each sequence to dnaA. Hybrid assemblies were created using HybridSPAdes as part of the SPAdes package (version 3.9.0) and closed with manual curation using the assembled scaffolds. Gene predictions and annotations were determined using Prokka (prokka 1.12-beta). In order to investigate shared and unique genome functions, annotations were derived from Prokka (Seemann 2014), GhostKOALA (Kanehisa et al. 2016), and Roary (Page et al. 2015). GhostKOALA determined KEGG identifier numbers for the protein sequences. The comparative genomic analysis was performed using Roary, which clusters proteins using MCL-edge (20% amino acid identity cutoff) (Heberle et al. 2015). Graphical representations of the shared proteins were produced with Interactivenn (www.interactivenn.net). Whole genome alignments were generated using Mauve (Darling et al. 2004).

Acquiring growth curves and subsequent analysis.

Monoculture of bacterial host and coculture of bacterial host with its Saccharibacteria parasite were recovered from frozen stocks and passaged twice in culture medium to ensure homogeneity. Three to five replicate samples (250 µL) of each culture were loaded into 96-well plates with starting optical density (OD600) of 0.05. The analogous image-based cell density measurement (TANormalized) was acquired using OcelloScope (BioSense Solutions) every 60 minutes for the indicated amount of time (Bor et al. 2018). The growth curve is graphed by plotting time versus arbitrary unit (a.u.) of the OcelloScope measurement which is equivalent to cell density measurement at 600 nm.

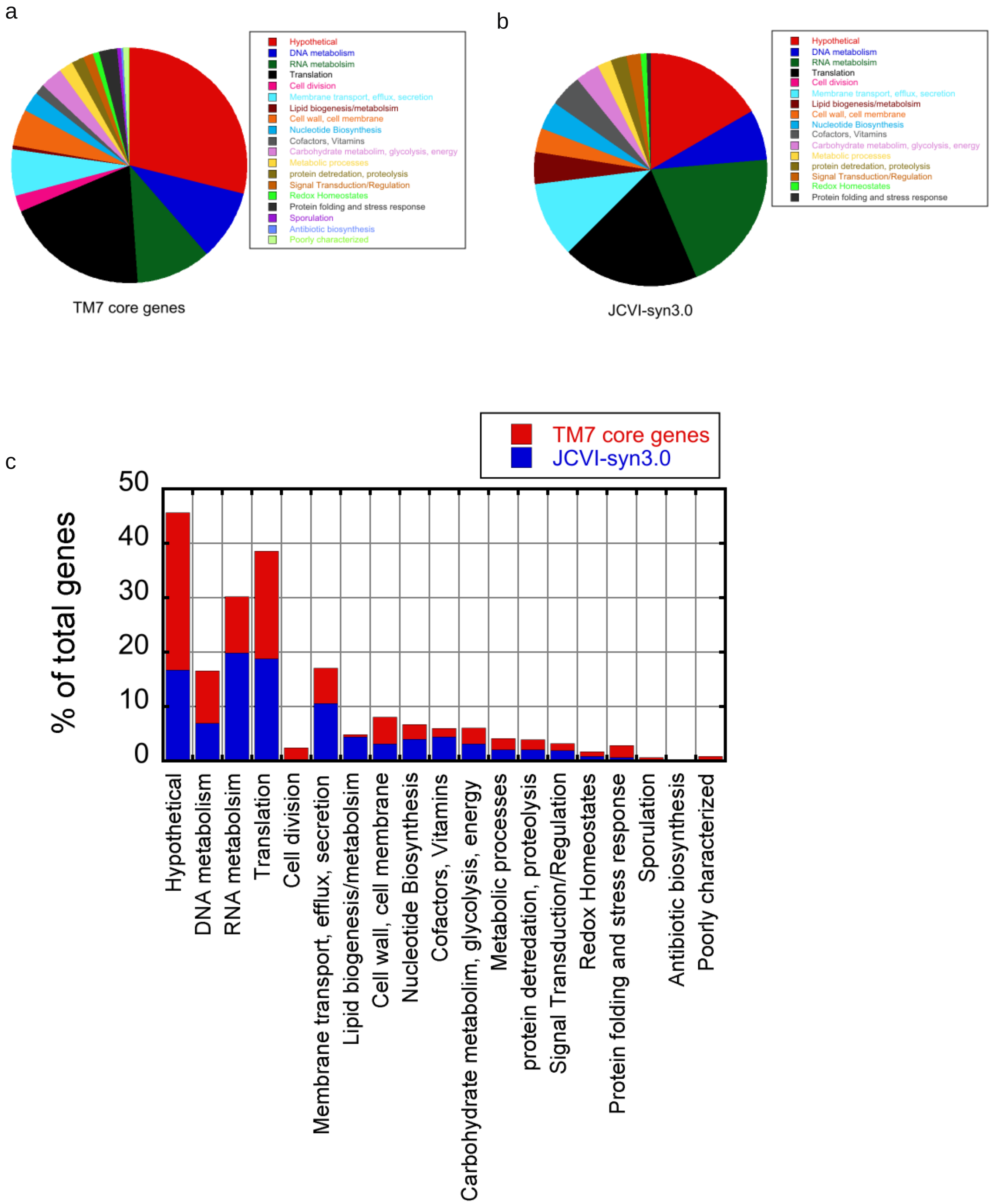
Generation times were calculated from cell densities from a 2-hour window of exponential growth in each curve with an a.u. minimum of 0.1. Cell densities were used in the equation $\log N_t - \log N_0 / 0.301 * t$ to derive the exponential growth rate, where N_t is the final cell density, N_0 is the initial cell density and t is time in hours. Generation times are the reciprocal of the exponential growth rate.

Testing Saccharibacteria species host ranges.

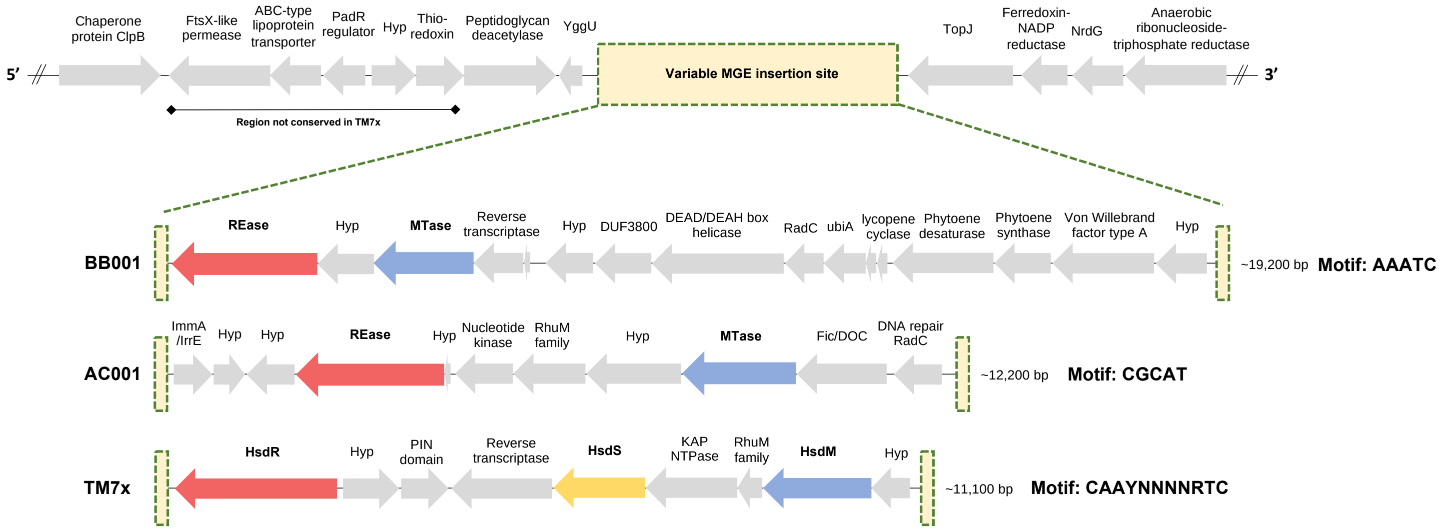
Binary cocultures of Saccharibacteria strains and their hosts were grown in TSBY (AC001 and PM004) or BHI (TM7x and BB001). The resultant growth was vortexed vigorously for one minute to detach loosely attached Saccharibacteria cells from their host and break up aggregates. Within the cocultures, there are also planktonic Saccharibacteria cells not attached to host. Twenty mL of the vortexed culture was passed through a 0.2-micron track-etched polycarbonate filter (Millipore) to collect just the Saccharibacteria cells. Filters were washed with 10 mL of buffer (Maximal Recovery Diluent, MRD). The flow-through and the wash fractions were collected in a 26 mL ultracentrifuge bottles and spun at 60,000 x g for 1 hour in Ti-70 rotor using a Beckman-Coulter Optima L-100 XP ultracentrifuge. The nearly invisible pellets were resuspended in 2 mL of buffer to create a concentrated suspension of Saccharibacteria cells independent of their host.

To infect new hosts, 2 mL of fresh media (BHI for *Actinomyces spp.*, TSOY/BHI/RPMI for all others tested) was inoculated with 200 µL of an overnight culture of a candidate host organism. To this, 200 µL of Saccharibacteria cell suspension prepared as described above was added. Cultures were incubated microaerophilically (2% O_2, 5% CO_2, 93% N_2) in a Coy Hypoxic Chamber at 37°C for 48 hours. Cultures were then passaged into 2 mL of fresh media with added 200 µL of an overnight culture of candidate host cells with their Saccharibacteria. Infections were screened by PCR (as described above) for the first 5 passages to determine if infections were stable.

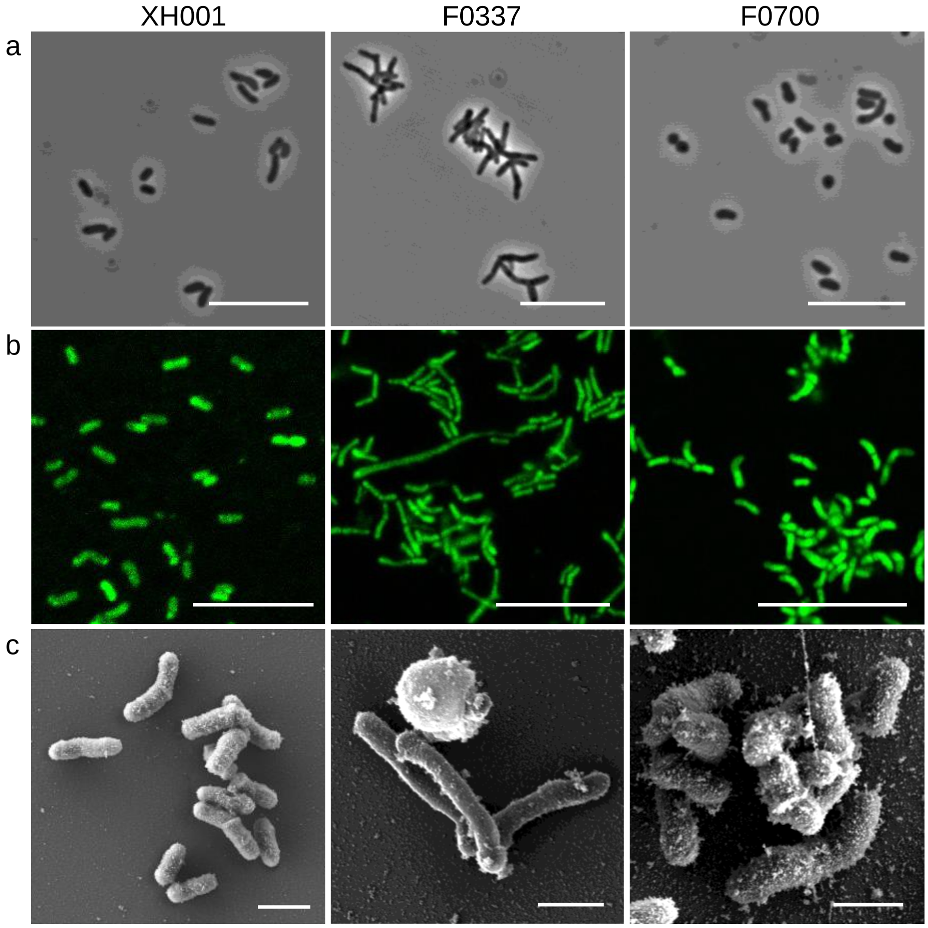
APPENDIX FIGURES



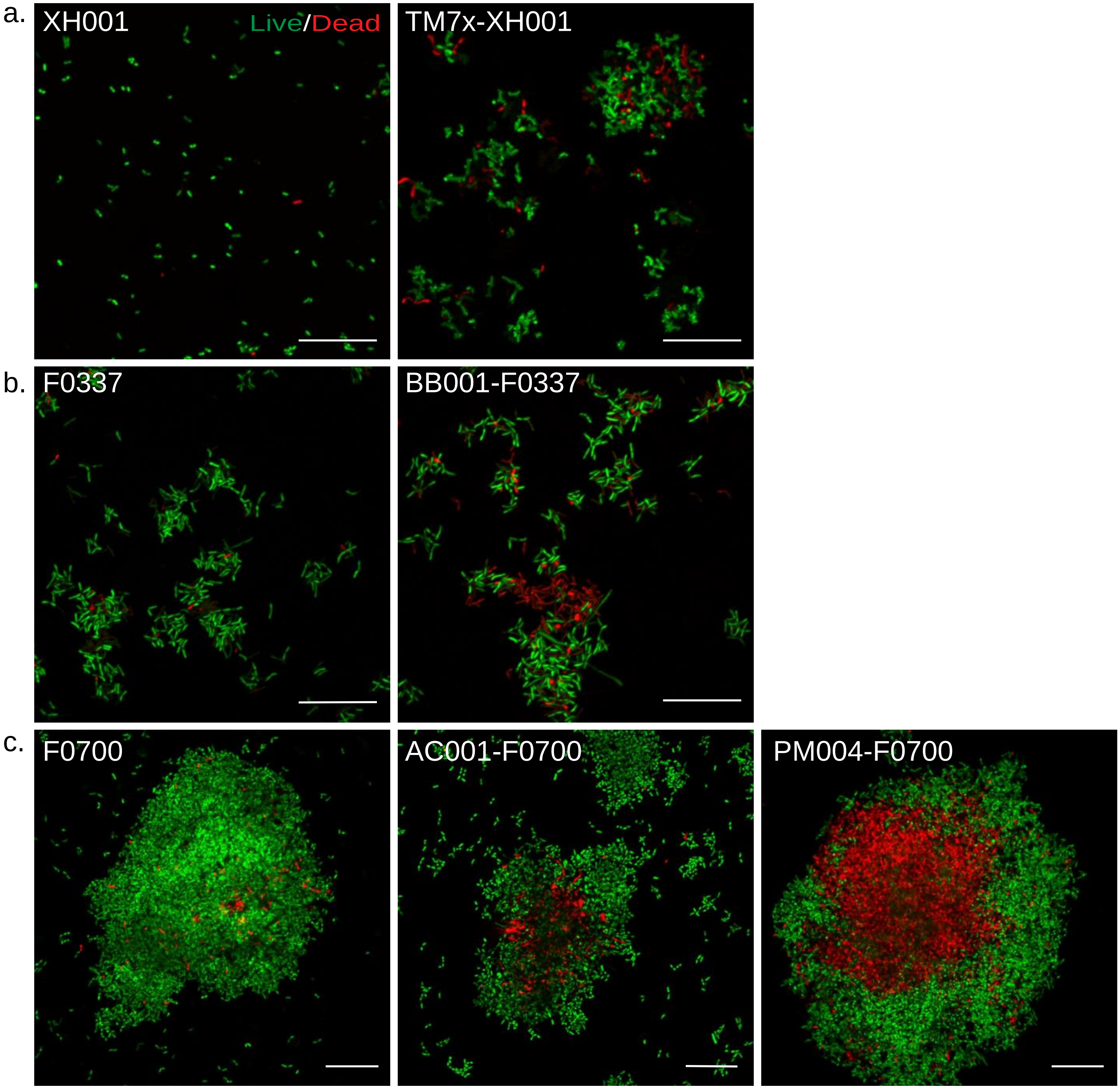
Appendix Figure 1. Core gene function. (a) 446 core genes from four oral Saccharibacteria genomes were functionally categorized into different groups. A majority of the genes were either hypothetical, DNA metabolism, RNA metabolism or translation genes. (b) Streamlined core gene categories for the JCVI-syn3.0 strain that has the minimal core genes needed to survive on its own (Hutchison et al. 2016). (c) Comparing the percent of certain functional groups between core genes of four Saccharibacteria strains and JCVI-syn3.0. Both had very similar functional gene category except Saccharibacteria had few additional genes.



Appendix Figure 2. Diagram of conserved insertion site for mobile genetic elements in TM7 strains. In each strain, a mobile genetic element harboring an RM system has inserted within a conserved genomic locus flanked by a 5’ prime 1.7 kbp region containing a Peptidoglycan-N-acetylmuramic acid deacetylase gene and a 3’ prime 5.2 kbp region containing genes associated with ferredoxin reductase and an anaerobic ribonucleoside-triphosphate reductase. Putative restriction endonuclease and methyltransferase genes are shown in red and blue, respectively, and specificity hsdS subunits are shown in yellow. Proteins not identified as part of the RM system or those with currently unknown function are shown in grey.



Appendix Figure 3. Imaging analysis of the host monocultures. Host monocultures were grown in appropriate media for 24 hours and imaged using (a) phase contrast, (b) FISH and (c) SEM. The coculture images with the Saccharibacteria are shown in main Figure 4. Scale bars are 10 μm for Phase contrast and FISH images while 1 μm for the SEM.



Appendix Figure 4. Live-Dead staining of mono and cocultures. Live (green) and Dead (red) staining of (a) TM7x-XH001, (b) BB001-F0337 and (c) AC001-F0700 and PM004-F0700 were carried out by following the manufacturer's guidelines (Invitrogen Cat#L34960). Stained cells were visualized on the Zeiss LSM 780 confocal microscope using 63x oil immersion objective. All cells grew for 24 hours and were imaged at the stationary phase. All scale bars are 20 μm.

APPENDIX TABLES

Table 1. “Baiting” host summary list.

Phylum	Species	HMT	Strain	Isolation results Bor	Isolation results Collins	Isolation results Murugkar
Actinobacteria	*Actinomyces bovis*	nonoral	ATCC 13683	negative		
Actinobacteria	*Actinomyces georgiae*	617	F0490	negative		
Actinobacteria	*Actinomyces graevenitzii*	866	F0530	negative		
Actinobacteria	*Actinomyces johnsonii*	849	F0330	negative		
Actinobacteria	*Actinomyces massiliensis*	852	F0489	negative		
Actinobacteria	*Actinomyces meyeri*	671	W712	negative		
Actinobacteria	*Actinomyces naeslundii*	176	ATCC 12104	negative		
Actinobacteria	*Actinomyces odontolyticus*	701	XH001	negative		
Actinobacteria	*Actinomyces sp.* HMT-170	170	F0386	negative		
Actinobacteria	*Actinomyces sp.* HMT-171	171	F0337	BB001		negative
Actinobacteria	*Actinomyces sp.* HMT-169	169	F0496		negative	
Actinobacteria	*Actinomyces sp.* HMT-448	448	F0400	negative		
Actinobacteria	*Actinomyces sp.* HMT-897	897	F0715		negative	
Actinobacteria	*Actinomyces sp.* HMT-180	180	ICM58	negative		
Actinobacteria	*Corynebacterium matruchotii*	666	ATCC 33806		negative	
Actinobacteria	*Pseudopropionibacterium propionicum*	739	F0230		AC001	PM004
Firmicutes	*Lachnoanaerobaculum saburreum*	494	F0468		negative	
Fusobacteria	*Leptotrichia wadei*	222	F0279			negative

Table 2. Percent Identity of cultivated Saccharibacteria strains by 16S rRNA molecular phylogeny.

	TM7x	BB001	AC001	PM004
TM7x	100	99.4	98.6	94
BB001	99.4	100	98.2	94.6
AC001	98.6	98.2	100	93.9
PM004	94	94.6	93.9	100

Table 3. Saccharibacteria genome summary.

	Genome (bp)	gene numbers	Number of RNAs	Contigs	GC content
TM7x	705,138	707	44	1	44.5
BB001	779,575	776	45	1	47.9
AC001	890,309	908	46	1	50.6
PM004	842,372	834	46	1	46.8
S. aal	1,013	1040	49	1	49.1
GWC2	1,039	1083	50	1	47.6
RAAC3	845	887	49	1	49.4

Table 4. Core genes shared among full/complete genomes of TM7x, BB001, AC001 and PM004 were compared resulting in 446 core genes (129/446 hypothetical).

Chromosomal replication initiator protein DnaA_dnaA_	Enolase 2_eno2_	hypothetical protein_group_866_
1-deoxy-D-xylulose-5-phosphate synthase_dxs_	Exodeoxyribonuclease 7 large subunit_xseA_	hypothetical protein_group_867_
1,4-alpha-glucan branching enzyme GlgB_glgB_	Exodeoxyribonuclease I_sbcB_	hypothetical protein_group_873_
2,3-bisphosphoglycerate-dependent phosphoglycerate mutase_gpmA_	Exodeoxyribonuclease_exoA_	hypothetical protein_group_903_
23S rRNA (uracil-C(5))-methyltransferase RlmCD_rlmCD_	Ferric uptake regulation protein_fur_	hypothetical protein_group_908_
30S ribosomal protein S1_rpsA_	FK506-binding protein_fbp_	hypothetical protein_group_912_
30S ribosomal protein S10_rpsJ_	Flavohemoprotein_hmp_	hypothetical protein_group_99_
30S ribosomal protein S11_rpsK_	Formamidopyrimidine-DNA glycosylase_mutM_	hypothetical protein_iolC_
30S ribosomal protein S12_rpsL_	GDP-6-deoxy-D-mannose reductase_lnpD_	hypothetical protein_pgsA_
30S ribosomal protein S13_rpsM_	GDP-mannose-dependent alpha-(1-6)-phosphatidylinositol monomannoside mannosyltransferase_pimB_	hypothetical protein_pilE_
30S ribosomal protein S15_rpsO_	Gluconeogenesis factor_group_476_	hypothetical protein_surA_
30S ribosomal protein S16_rpsP_	Glutamate--tRNA ligase 1_gltX1_	hypothetical protein_tqsA_
30S ribosomal protein S17_rpsQ_	Glutamyl-tRNA(Gln) amidotransferase subunit A_gatA_	hypothetical protein_xseB_
30S ribosomal protein S18_rpsR_	Glyceraldehyde-3-phosphate dehydrogenase_gap_	hypothetical protein_ydeN_1_
30S ribosomal protein S19_rpsS_	Glycine--tRNA ligase_glyQS_	Inorganic pyrophosphatase_ppa_
30S ribosomal protein S2_rpsB_	Glycogen synthase_group_134_	Isoprenyl transferase_uppS_
30S ribosomal protein S20_rpsT_	GTP pyrophosphokinase_relA_	Leucine--tRNA ligase_leuS_
30S ribosomal protein S21_rpsU_	GTP-binding protein TypA/BipA_typA_	LexA repressor_lexA_
30S ribosomal protein S3_rpsC_	GTPase Der_der_	Lipopolysaccharide assembly protein B_lapB_
30S ribosomal protein S4_rpsD_	GTPase Obg_obg_	Lipoprotein-releasing system ATP-binding protein LolD_lolD_1_
30S ribosomal protein S5_rpsE_	Guanylate kinase_gmk_	Magnesium and cobalt efflux protein CorC_corC_
30S ribosomal protein S6_rpsF_	Heat-inducible transcription repressor HrcA_hrcA_	Membrane protein insertase YidC 2_yidC2_
30S ribosomal protein S7_rpsG_	Hemolysin A_tlyA_	Mercuric reductase_merA_
30S ribosomal protein S8_rpsH_	Histidine--tRNA ligase_hisS_	Metalloprotease MmpA_mmpA_
30S ribosomal protein S9_rpsI_	Holliday junction ATP-dependent DNA helicase RuvA_ruvA_	Methionine aminopeptidase 1_map_
50S ribosomal protein L1_rplA_	Holliday junction ATP-dependent DNA helicase RuvB_ruvB_	Methionine--tRNA ligase_metG_
50S ribosomal protein L10_rplJ_	hypothetical protein_aroK_	Methionyl-tRNA formyltransferase_fmt_
50S ribosomal protein L11_rplK_	hypothetical protein_bcrC_	N utilization substance protein B_nusB_
50S ribosomal protein L13_rplM_	hypothetical protein_cyaB_	N-acetylmuramoyl-L-alanine amidase sle1_sle1_
50S ribosomal protein L14_rplN_	hypothetical protein_ftsA_1_	NAD-dependent malic enzyme_group_822_
50S ribosomal protein L15_rplO_	hypothetical protein_ftsQ_	Nucleoid-associated protein_group_672_
50S ribosomal protein L16_rplP_	hypothetical protein_group_103_	Nucleoside diphosphate kinase_ndk_
50S ribosomal protein L17_rplQ_	hypothetical protein_group_192_	Oligo-1,6-glucosidase_malL_
50S ribosomal protein L18_rplR_	hypothetical protein_group_202_	Oligopeptide-binding protein AppA_appA_
50S ribosomal protein L19_rplS_	hypothetical protein_group_203_	Oligoribonuclease_orn_
50S ribosomal protein L2_rplB_	hypothetical protein_group_211_	Orotate phosphoribosyltransferase_pyrE_
50S ribosomal protein L20_rplT_	hypothetical protein_group_270_	Penicillin-binding protein 1F_pbpF_
50S ribosomal protein L21_rplU_	hypothetical protein_group_271_	Peptide chain release factor 1_prfA_
50S ribosomal protein L22_rplV_	hypothetical protein_group_285_	Peptide chain release factor 2_prfB_
50S ribosomal protein L23_rplW_	hypothetical protein_group_331_	Peptide deformylase_def_
50S ribosomal protein L24_rplX_	hypothetical protein_group_338_	Peptidyl-tRNA hydrolase_pth_
50S ribosomal protein L27_rpmA_	hypothetical protein_group_367_	Phenylalanine--tRNA ligase alpha subunit_pheS_
50S ribosomal protein L28_rpmB_	hypothetical protein_group_369_	Phenylalanine--tRNA ligase beta subunit_pheT_
50S ribosomal protein L29_rpmC_	hypothetical protein_group_371_	Phospho-N-acetylmuramoyl-pentapeptide-transferase_mraY_
50S ribosomal protein L3_rplC_	hypothetical protein_group_38_	Polyribonucleotide nucleotidyltransferase_pnp_
50S ribosomal protein L31 type B_rpmE2_	hypothetical protein_group_393_	Primosomal protein N'_priA_
50S ribosomal protein L33_rpmG_	hypothetical protein_group_401_	Proline--tRNA ligase_proS_
50S ribosomal protein L35_rpmI_	hypothetical protein_group_409_	Protein GrpE_grpE_
50S ribosomal protein L4_rplD_	hypothetical protein_group_426_	Protein RecA_recA_
50S ribosomal protein L5_rplE_	hypothetical protein_group_429_	Protein translocase subunit SecA_secA_
50S ribosomal protein L6_rplF_	hypothetical protein_group_435_	Protein translocase subunit SecE_secE_

50S ribosomal protein L7/L12_rplL_	hypothetical protein_group_445_	Protein translocase subunit SecY_secY_
6-phospho-5-dehydro-2-deoxy-D-gluconate aldolase_iolJ_	hypothetical protein_group_446_	putative ABC transporter ATP-binding protein YknY_yknY_
7-carboxy-7-deazaguanine synthase_queE_	hypothetical protein_group_452_	putative ABC transporter permease YknZ_yknZ_
8-demethyl-8-(2,3-dimethoxy-alpha-L-rhamnosyl)-tetracenomycin-C 4'-O-methyltransferase_elmMIII_	hypothetical protein_group_453_	Putative ATP-dependent DNA helicase YjcD_yjcD_
ABC transporter permease YtrF_ytrF_	hypothetical protein_group_454_	putative chromosome-partitioning protein ParB_parB_
Adenylate kinase_adk_	hypothetical protein_group_461_	putative CtpA-like serine protease_ctpB_
ADP-dependent (S)-NAD(P)H-hydrate dehydratase_nnrD_	hypothetical protein_group_484_	putative lipid II flippase MurJ_murJ_
Alanine--tRNA ligase_alaS_	hypothetical protein_group_490_	Putative lipoprotein YerB_yerB_
Alkaline phosphatase synthesis transcriptional regulatory PhoP	hypothetical protein_group_491_	putative membrane protein_group_537_
Alpha-monoglucosyldiacylglycerol synthase_mgs_	hypothetical protein_group_494_	putative MscS family protein YkuT_ykuT_
Alternate 30S ribosomal protein S14_rpsN2_	hypothetical protein_group_497_	Putative multidrug export ATP-binding/permease protein
Aminopeptidase N_pepN_	hypothetical protein_group_499_	Putative multidrug export ATP-binding/permease protein_lolD_
Anaerobic ribonucleoside-triphosphate reductase_nrdD_	hypothetical protein_group_500_	putative peptidoglycan glycosyltransferase FtsW_ftsW_
Aquaporin Z_aqpZ_	hypothetical protein_group_504_	putative PIN and TRAM-domain containing protein YacL_yacL_
Arginine transport system permease protein ArtQ_artQ_	hypothetical protein_group_514_	Putative pre-16S rRNA nuclease_yrrK_
Arginine--tRNA ligase_argS_	hypothetical protein_group_522_	Putative serine protease HtrA_htrA_
Arginine-binding extracellular protein ArtP_artP_	hypothetical protein_group_524_	Putative threonylcarbamoyl-AMP synthase_ywlC_
Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B_gatB_	hypothetical protein_group_526_	Putative transcriptional regulator YvhJ_yvhJ_
Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit C_gatC_	hypothetical protein_group_530_	putative transcriptional regulatory protein/MSMEI_2866_pmpR_
ATP synthase epsilon chain_atpC_	hypothetical protein_group_532_	putative tRNA-dihydrouridine synthase_dus_
ATP synthase gamma chain_atpG_	hypothetical protein_group_534_	Putative two-component membrane permease complex subunit SMU_747c_group_958_
ATP synthase subunit a_atpB_	hypothetical protein_group_539_	Putative zinc metalloprotease Rip2_rip2_
ATP synthase subunit alpha_atpA_	hypothetical protein_group_541_	Pyruvate kinase_pyk_
ATP synthase subunit b_atpF_	hypothetical protein_group_543_	Recombination protein RecR_recR_
ATP synthase subunit beta_atpD_2_	hypothetical protein_group_544_	Regulatory protein RecX_recX_
ATP synthase subunit c_atpE_	hypothetical protein_group_545_	Release factor glutamine methyltransferase_prmC_
ATP synthase subunit delta_atpD_1_	hypothetical protein_group_550_	Replication-associated recombination protein A_rarA_
ATP-dependent Clp protease ATP-binding subunit ClpC_clpC_	hypothetical protein_group_551_	Replicative DNA helicase_dnaC_
ATP-dependent Clp protease ATP-binding subunit ClpE_clpE_	hypothetical protein_group_554_	Ribonuclease 3_rnc_
ATP-dependent Clp protease proteolytic subunit_clpP_	hypothetical protein_group_568_	Ribonuclease HI_rnhA_
ATP-dependent DNA helicase PcrA_pcrA_	hypothetical protein_group_569_	Ribonuclease HII_rnhB_
ATP-dependent DNA helicase RecG_recG_	hypothetical protein_group_570_	Ribonuclease J1_rnjA_
ATP-dependent RecD-like DNA helicase_recD2_	hypothetical protein_group_571_	Ribonuclease P protein component_rnpA_
ATP-dependent zinc metalloprotease FtsH_ftsH_	hypothetical protein_group_581_	Ribonuclease Y_rny_
Bifunctional oligoribonuclease and PAP phosphatase NrnA_nrnA_	hypothetical protein_group_583_	Ribonucleoside-diphosphate reductase 1 subunit alpha_nrdA_
Bifunctional PGK/TIM_pgk/tpi_	hypothetical protein_group_584_	Ribonucleoside-diphosphate reductase subunit beta_nrdB_
Bifunctional protein FolD protein_folD_	hypothetical protein_group_587_	Ribose-5-phosphate isomerase B_rpiB_
Carbamate kinase 1_arcC1_	hypothetical protein_group_596_	Ribosomal large subunit pseudouridine synthase B_rluB_
Cell division ATP-binding protein FtsE_ftsE_	hypothetical protein_group_597_	Ribosomal large subunit pseudouridine synthase D_rluD_
Cell division protein FtsA_ftsA_2_	hypothetical protein_group_598_	Ribosomal RNA large subunit methyltransferase H_rlmH_
Cell division protein FtsA_ftsA_3_	hypothetical protein_group_600_	Ribosomal RNA small subunit methyltransferase A_rsmA_
Cell division protein FtsA_group_272_ftsA_1	hypothetical protein_group_601_	Ribosomal RNA small subunit methyltransferase D_rsmD_
Cell division protein FtsL_ftsL_	hypothetical protein_group_602_	Ribosomal RNA small subunit methyltransferase H_rsmH_
Cell division protein FtsX_ftsX_	hypothetical protein_group_603_	Ribosome hibernation promotion factor_hpf_
Cell division protein FtsZ_ftsZ_	hypothetical protein_group_605_	Ribosome-binding ATPase YchF_ychF_
Chaperone protein DnaJ_dnaJ_	hypothetical protein_group_607_	Ribosome-recycling factor_frr_
Chaperone protein DnaK_dnaK_	hypothetical protein_group_608_	Ribulose-phosphate 3-epimerase_rpe_
Cobyric acid synthase_cobQ_	hypothetical protein_group_612_	RNA polymerase principal sigma factor HrdA_hrdA_
ComE operon protein 3_comEC_	hypothetical protein_group_618_	RNA polymerase sigma factor SigA_sigA_
Competence protein ComM_comM_	hypothetical protein_group_621_	RNA pyrophosphohydrolase_rppH_4_

Crossover junction endodeoxyribonuclease RuvC_ruvC_	hypothetical protein_group_622_	Sensor histidine kinase ResE_walK_2_
Cysteine desulfurase IscS_iscS_	hypothetical protein_group_626_	Sensor histidine kinase WalK_walK_1_
Cysteine--tRNA ligase_cysS_	hypothetical protein_group_630_	Sensor histidine kinase WalK_walK_3_
D-aminoacyl-tRNA deacylase_tatD_	hypothetical protein_group_631_	Sensor protein KdpD_kdpD_
dCTP deaminase_dcd_	hypothetical protein_group_636_	Sensor protein kinase WalK_group_83_walK_1
Dephospho-CoA kinase_coaE_	hypothetical protein_group_643_	Serine hydroxymethyltransferase_glyA_
Diadenosine hexaphosphate hydrolase_ndx1_	hypothetical protein_group_647_	Serine--tRNA ligase_serS_
Dihydrofolate reductase_dfrA_	hypothetical protein_group_649_	Signal peptidase I T_sipT_
Dihydroorotate dehydrogenase (quinone)_pyrD_	hypothetical protein_group_664_	Single-stranded DNA-binding protein_ssb_
Disulfide bond formation protein D_bdbD_	hypothetical protein_group_679_	Single-stranded-DNA-specific exonuclease RecJ_recJ_
DNA base-flipping protein_atl_	hypothetical protein_group_688_	Spore protein SP21_hspA_
DNA gyrase subunit A_gyrA_	hypothetical protein_group_702_	Sporulation initiation inhibitor protein Soj_soj_
DNA gyrase subunit B_gyrB_	hypothetical protein_group_711_	SsrA-binding protein_smpB_
DNA ligase_ligA_	hypothetical protein_group_713_	Stage V sporulation protein D_spoVD_
DNA polymerase I, thermostable_polA_	hypothetical protein_group_714_	Teichoic acid translocation permease protein TagG_tagG_
DNA polymerase III subunit alpha_dnaE_	hypothetical protein_group_716_	Teichoic acids export ATP-binding protein TagH_tagH_
DNA polymerase III subunit beta_dnaN_	hypothetical protein_group_717_	Thioredoxin reductase_trxB_
DNA polymerase III subunit gamma/tau_dnaX_2_	hypothetical protein_group_723_	Threonine--tRNA ligase_thrS_
DNA polymerase III subunit tau_dnaX_1_	hypothetical protein_group_729_	Thymidylate kinase_tmk_
DNA primase_dnaG_	hypothetical protein_group_730_	Thymidylate synthase_thyA_
DNA processing protein DprA_dprA_	hypothetical protein_group_732_	Transcription elongation factor GreA_greA_1_
DNA repair protein RadA_radA_	hypothetical protein_group_734_	Transcription termination/antitermination protein NusA_nusA_
DNA repair protein RecO_recO_	hypothetical protein_group_741_	Transcription termination/antitermination protein NusG_nusG_
DNA replication and repair protein RecF_recF_	hypothetical protein_group_743_	Transcriptional regulatory protein OmpR_ompR_
DNA topoisomerase 1_topA_	hypothetical protein_group_752_	Transcriptional regulatory protein SrrA_srrA_
DNA translocase SpoIIIE_spoIIIE_	hypothetical protein_group_753_	Transcriptional repressor NrdR_nrdR_2_
DNA-binding response regulator MtrA_mtrA_	hypothetical protein_group_754_	Transketolase 2_tktB_
DNA-directed RNA polymerase subunit alpha_rpoA_	hypothetical protein_group_762_	Translation initiation factor IF-1_infA_
DNA-directed RNA polymerase subunit beta_rpoB_	hypothetical protein_group_768_	Translation initiation factor IF-2_infB_
DNA-directed RNA polymerase subunit beta'_rpoC_	hypothetical protein_group_769_	Translation initiation factor IF-3_infC_
Elongation factor G_fusA_	hypothetical protein_group_771_	Trigger factor_tig_
Elongation factor P_efp_	hypothetical protein_group_775_	Triosephosphate isomerase_tpiA_
Elongation factor Ts_tsf_	hypothetical protein_group_780_	tRNA (guanine-N(1)-)-methyltransferase_trmD_
Elongation factor Tu_tuf_	hypothetical protein_group_786_	tRNA (guanine-N(7)-)-methyltransferase_trmB_
Endolytic murein transglycosylase_mltG_	hypothetical protein_group_796_	tRNA (guanosine(18)-2'-O)-methyltransferase_trmH_1_
UDP-N-acetylmuramoylalanine--D-glutamate ligase_murD_2_	hypothetical protein_group_800_	tRNA (guanosine(18)-2'-O)-methyltransferase_trmH_2_
Undecaprenyl-diphosphatase_uppP_	hypothetical protein_group_804_	tRNA dimethylallyltransferase_miaA_
UTP--glucose-1-phosphate uridylyltransferase_gtaB_2_	hypothetical protein_group_830_	tRNA N6-adenosine threonylcarbamoyltransferase_tsaD_
UvrABC system protein A_uvrA_	hypothetical protein_group_835_	tRNA pseudouridine synthase B_truB_
UvrABC system protein B_uvrB_	hypothetical protein_group_862_	tRNA threonylcarbamoyladenosine biosynthesis protein TsaB_tsaB_
UvrABC system protein C_uvrC_	UDP-N-acetylenolpyruvoylglucosamine reductase_murB_	tRNA threonylcarbamoyladenosine biosynthesis protein TsaE_tsaE_
Valine--tRNA ligase_valS_	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1_murAA_	tRNA-specific 2-thiouridylase MnmA_mnmA_
Type II secretion system protein F_epsF_2_	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase_murG_	tRNA(Ile)-lysidine synthase_tilS_
Tyrosine recombinase XerD_xerD_	UDP-N-acetylmuramate--L-alanine ligase_murC_	Tryptophan--tRNA ligase_trpS_
Tyrosine--tRNA ligase_tyrS_	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase_murE_	Twitching mobility protein_pilT_
UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferase_wcaJ_	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase_murF_2_	Type 4 prepilin-like proteins leader peptide-processing enzyme_comC_
Type II secretion system protein F_epsF_1_	UDP-N-acetylmuramoylalanine--D-glutamate ligase_murD_1_	Type II secretion system protein E_xpsE_1_
Type II secretion system protein E_xpsE_3_	Type II secretion system protein E_xpsE_2_	

Table 5. Oral Saccharibacteria core gene functional category.

Hypothetical (129 proteins)	**DNA metabolism (repair, replication, uptake, topology, chromosome segregation)**	**RNA metabolism**	**Translation**	**Cell division**	**Membrane transport, efflux, secretion**	**Lipid biogenesis/ metabolism**	**Cell wall, cell membrane**	**Nucleotide Biosynthesis**	**Cofactors transport and salvage, Vitamins**	**Secretion system**	**Carbohydrate metabolim, glycolysis, energy production**	**Proton motive force generation (under transport)**	**Metabolic processes**	**Protein detredation, proteolysis**	**Signal Transduction/ Regulation**	**Redox Homeostates**	**Protein folding and stress response**	**Poorly characterized**	**Sporulation**
hypothetical protein_group_367_	Exodeoxyribonuclease 7 large subunit_xseA_	RNA polymerase sigma factor SigA_sigA_	Elongation factor G_fusA_	Cell division protein FtsA_group_272_ftsA_1	putative ABC transporter permease YknZ_yknZ_	Undecaprenyl-diphosphatase_uppP_	Teichoic acids export ATP-binding protein TagH_tagH_	Polyribonucleotide nucleotidyltransferase_pnp_	ATP-dependent zinc metalloprotease FtsH_ftsH_	Type II secretion system protein E_xpsE_1_	Ribulose-phosphate 3-epimerase_rpe_	ATP synthase epsilon chain_atpC_	Thioredoxin reductase_trxB_	Signal peptidase I T_sipT_	Sensor histidine kinase ResE_walK_2_	7-carboxy-7-deazaguanine synthase_queE_	Chaperone protein DnaJ_dnaJ_	N utilization substance protein B_nusB_	Spore protein SP21_hspA_
hypothetical protein_group_369_	Exodeoxyribonuclease_exoA_	RNA polymerase principal sigma factor HrdA_hrdA_	Elongation factor P_efp_	Cell division protein FtsZ_ftsZ_	putative ABC transporter ATP-binding protein YknY_yknY_	Alpha-monoglucosyldiacylglycerol synthase_mgs_	Teichoic acid translocation permease protein TagG_tagG_	Ribonucleoside-diphosphate reductase subunit beta_nrdB_	Dihydrofolate reductase_dfrA_	Type II secretion system protein F_epsF_1_	Transketolase 2_tktB_	ATP synthase gamma chain_atpG_	1-deoxy-D-xylulose-5-phosphate synthase_dxs_	ATP-dependent Clp protease ATP-binding subunit ClpC_clpC_	Sensor histidine kinase WalK_walK_1_	Disulfide bond formation protein D_bdbD_	Chaperone protein DnaK_dnaK_	Nucleoid-associated protein_group_672_	Sporulation initiation inhibitor protein Soj_soj_
hypothetical protein_group_435_	Exodeoxyribonuclease I_sbcB_	DNA-directed RNA polymerase subunit alpha_rpoA_	Elongation factor Ts_tsf_	Cell division ATP-binding protein FtsE_ftsE_	ABC transporter permease YtrF_ytrF_	**Lipid biogenesis/metabolsim**	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase_murE_	Nucleoside diphosphate kinase_ndk_	Dephospho-CoA kinase_coaE_	Type II secretion system protein E_xpsE_2_	NAD-dependent malic enzyme_group_822_	ATP synthase subunit a_atpB_	6-phospho-5-dehydro-2-deoxy-D-gluconate aldolase_iolJ_	ATP-dependent Clp protease ATP-binding subunit ClpE_clpE_	Sensor histidine kinase WalK_walK_3_	Ferric uptake regulation protein_fur_	Protein GrpE_grpE_	putative membrane protein_group_537_	Stage V sporulation protein D_spoVD_
hypothetical protein_group_211_	Crossover junction endodeoxyribonuclease RuvC_ruvC_	DNA-directed RNA polymerase subunit beta'_rpoC_	Elongation factor Tu_tuf_	Cell division protein FtsX_ftsX_	Arginine transport system permease protein ArtQ_artQ_	Undecaprenyl-diphosphatase_uppP_	UDP-N-acetylmuramoylalanine--D-glutamate ligase_murD_1_	Anaerobic ribonucleoside-triphosphate reductase_nrdD_	Cobyric acid synthase_cobQ_	Type II secretion system protein E_xpsE_3_	Pyruvate kinase_pyk_	ATP synthase subunit alpha_atpA_	ADP-dependent (S)-NAD(P)H-hydrate dehydratase_nnrD_	ATP-dependent Clp protease proteolytic subunit_clpP_	Sensor protein kinase WalK_group_83_walK_1	Mercuric reductase_merA_	Aquaporin Z_aqpZ_	putative PIN and TRAM-domain containing protein YacL_yacL_	**Sporulation**
hypothetical protein_group_371_	Holliday junction ATP-dependent DNA helicase RuvB_ruvB_	DNA-directed RNA polymerase subunit beta_rpoB_	30S ribosomal protein S20_rpsT_	Cell division protein FtsL_ftsL_	Magnesium and cobalt efflux protein CorC_corC_	Alpha-monoglucosyldiacylglycerol synthase_mgs_	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase_murG_	Ribonucleoside-diphosphate reductase 1 subunit alpha_nrdA_	Cysteine desulfurase IscS_iscS_	Type II secretion system protein F_epsF_2_	Ribose-5-phosphate isomerase B_rpiB_	ATP synthase subunit b_atpF_	Ribulose-phosphate 3-epimerase_rpe_	Metalloprotease MmpA_mmpA_	GTP-binding protein TypA/BipA_typA_	**Redox Homeostates**	FK506-binding protein_fbp_	**Poorly characterized**	Spore protein SP21_hspA_
hypothetical protein_group_103_	Holliday junction ATP-dependent DNA helicase RuvA_ruvA_	23S rRNA (uracil-C(5))-methyltransferase RlmCD_rlmCD_	30S ribosomal protein S15_rpsO_	Cell division protein FtsA_ftsA_2_	Protein translocase subunit SecA_secA_	**Lipid biogenesis/metabolsim**	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase_murF_2_	Adenylate kinase_adk_	Serine hydroxymethyltransferase_glyA_	Twitching mobility protein_pilT_	Enolase 2_eno2_	ATP synthase subunit beta_atpD_2_	Bifunctional protein FolD protein_folD_	putative CtpA-like serine protease_ctpB_	Sensor protein KdpD_kdpD_	7-carboxy-7-deazaguanine synthase_queE_	Flavohemoprotein_hmp_	N utilization substance protein B_nusB_	Sporulation initiation inhibitor protein Soj_soj_
hypothetical protein_group_99_	ATP-dependent RecD-like DNA helicase_recD2_	Putative pre-16S rRNA nuclease_yrrK_	30S ribosomal protein S6_rpsF_	Cell division protein FtsA_ftsA_3_	Protein translocase subunit SecE_secE_	Undecaprenyl-diphosphatase_uppP_	UDP-N-acetylenolpyruvoylglucosamine reductase_murB_	Thymidylate kinase_tmk_	Carbamate kinase 1_arcC1_	Type 4 prepilin-like proteins leader peptide-processing enzyme_comC_	Glyceraldehyde-3-phosphate dehydrogenase_gap_	ATP synthase subunit c_atpE_	GDP-6-deoxy-D-mannose reductase_lnpD_	Putative serine protease HtrA_htrA_	**Signal Transduction/Regulation**	Disulfide bond formation protein D_bdbD_	Gluconeogenesis factor_group_476_	Nucleoid-associated protein_group_672_	Stage V sporulation protein D_spoVD_
hypothetical protein_group_568_	Putative ATP-dependent DNA helicase YjcD_yjcD_	tRNA dimethylallyltransferase_miaA_	30S ribosomal protein S18_rpsR_	putative chromosome-partitioning protein ParB_parB_	Protein translocase subunit SecY_secY_	Alpha-monoglucosyldiacylglycerol synthase_mgs_	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1_murAA_	Thymidylate synthase_thyA_	**Cofactors transport and salvage, Vitamins**		2,3-bisphosphoglycerate-dependent phosphoglycerate mutase_gpmA_	ATP synthase subunit delta_atpD_1_	Guanylate kinase_gmk_	Putative zinc metalloprotease Rip2_rip2_	Sensor histidine kinase ResE_walK_2_	Ferric uptake regulation protein_fur_	GTP pyrophosphokinase_relA_	putative membrane protein_group_537_	**Sporulation**
hypothetical protein_group_569_	Replicative DNA helicase_dnaC_	tRNA (guanosine(18)-2'-O)-methyltransferase_trmH_1_	50S ribosomal protein L28_rpmB_	putative peptidoglycan glycosyltransferase FtsW_ftsW_	Arginine-binding extracellular protein ArtP_artP_	**Lipid biogenesis/metabolsim**	UDP-N-acetylmuramoylalanine--D-glutamate ligase_murD_2_	dCTP deaminase_dcd_	ATP-dependent zinc metalloprotease FtsH_ftsH_		Bifunctional PGK/TIM_pgk/tpi_								
hypothetical protein_group_570_	ATP-dependent DNA helicase RecG_recG_	tRNA (guanosine(18)-2'-O)-methyltransferase_trmH_2_	50S ribosomal protein L20_rplT_	N-acetylmuramoyl-L-alanine amidase sle1_sle1_	Hemolysin A_tlyA_	Undecaprenyl-diphosphatase_uppP_	UDP-N-acetylmuramate--L-alanine ligase_murC_	Inorganic pyrophosphatase_ppa_	Dihydrofolate reductase_dfrA_		Triosephosphate isomerase_tpiA_								
hypothetical protein_group_571_	ATP-dependent DNA helicase PcrA_pcrA_	tRNA (guanine-N(1)-)-methyltransferase_trmD_	50S ribosomal protein L35_rpmI_		putative MscS family protein YkuT_ykuT_	Alpha-monoglucosyldiacylglycerol synthase_mgs_	UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferase_wcaJ_	Dihydroorotate dehydrogenase (quinone)_pyrD_	Dephospho-CoA kinase_coaE_		UTP--glucose-1-phosphate uridylyltransferase_gtaB_2_								
hypothetical protein_group_581_	DNA ligase_ligA_	tRNA (guanine-N(7)-)-methyltransferase_trmB_	30S ribosomal protein S1_rpsA_		Putative multidrug export ATP-binding/permease protein_lolD_	**Lipid biogenesis/metabolsim**	putative lipid II flippase MurJ_murJ_	Orotate phosphoribosyltransferase_pyrE_	Cobyric acid synthase_cobQ_		Diadenosine hexaphosphate hydrolase_ndx1_								
hypothetical protein_group_903_	DNA polymerase III subunit beta_dnaN_	Ribonuclease HI_rnhA_	30S ribosomal protein S21_rpsU_		Putative multidrug export ATP-binding/permease protein_group_20_	Undecaprenyl-diphosphatase_uppP_	Lipopolysaccharide assembly protein B_lapB_	**Nucleotide Biosynthesis**	Cysteine desulfurase IscS_iscS_		Oligo-1,6-glucosidase_malL_								
hypothetical protein_group_583_	DNA polymerase I, thermostable_polA_	Oligoribonuclease_orn_	50S ribosomal protein L21_rplU_		Putative two-component membrane permease complex subunit SMU_747c_group_958_	Alpha-monoglucosyldiacylglycerol synthase_mgs_	Lipoprotein-releasing system ATP-binding protein LolD_lolD_1_	Polyribonucleotide nucleotidyltransferase_pnp_	Serine hydroxymethyltransferase_glyA_										
hypothetical protein_group_584_	DNA polymerase III subunit tau_dnaX_1_	Ribonuclease J1_rnjA_	50S ribosomal protein L31 type B_rpmE2_																
hypothetical protein_group_587_	DNA polymerase III subunit gamma/tau_dnaX_2_	Bifunctional oligoribonuclease and PAP phosphatase NrnA_nrnA_	30S ribosomal protein S2_rpsB_																
hypothetical protein_group_445_	DNA polymerase III subunit alpha_dnaE_	Ribonuclease HII_rnhB_	50S ribosomal protein L27_rpmA_																
hypothetical protein_group_446_	Single-stranded-DNA-specific exonuclease RecJ_recJ_	Ribonuclease 3_rnc_	50S ribosomal protein L7/L12_rplL_																
hypothetical protein_ftsA_1_	Formamidopyrimidine-DNA glycosylase_mutM_	Ribonuclease Y_rny_	50S ribosomal protein L10_rplJ_																
hypothetical protein_group_596_	Single-stranded DNA-binding protein_ssb_	Ribonuclease P protein component_rnpA_	50S ribosomal protein L1_rplA_																
hypothetical protein_group_597_	DNA replication and repair protein RecF_recF_	tRNA-specific 2-thiouridylase MnmA_mnmA_	50S ribosomal protein L11_rplK_																
hypothetical protein_group_598_	DNA repair protein RecO_recO_	Ribosomal large subunit pseudouridine synthase B_rluB_	50S ribosomal protein L33_rpmG_																

hypothetical protein_group_38_	DNA repair protein RadA_radA_	Ribosomal large subunit pseudouridine synthase D_rluD_	30S ribosomal protein S9_rpsI_																
hypothetical protein_group_600_	Recombination protein RecR_recR_	tRNA pseudouridine synthase B_truB_	50S ribosomal protein L13_rplM_																
hypothetical protein_group_601_	Regulatory protein RecX_recX_	putative tRNA-dihydrouridine synthase_dus_	50S ribosomal protein L17_rplQ_																
hypothetical protein_group_602_	Protein RecA_recA_	tRNA threonylcarbamoyl adenosine biosynthesis protein TsaB_tsaB_	30S ribosomal protein S4_rpsD_																
hypothetical protein_group_603_	Chromosomal replication initiator protein DnaA_dnaA_	tRNA threonylcarbamoyl adenosine biosynthesis protein TsaE_tsaE_	30S ribosomal protein S11_rpsK_																
hypothetical protein_group_605_	DNA primase_dnaG_	tRNA(Ile)-lysidine synthase_tilS_	30S ribosomal protein S13_rpsM_																
hypothetical protein_surA_	DNA translocase SpoIIIE_spoIIIE_	tRNA N6-adenosine threonylcarbamoyltransferase_tsaD_	50S ribosomal protein L15_rplO_																
hypothetical protein_group_607_	DNA base-flipping protein_atl_	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B_gatB_	30S ribosomal protein S5_rpsE_																
hypothetical protein_group_608_	DNA gyrase subunit B_gyrB_	Glutamyl-tRNA(Gln) amidotransferase subunit A_gatA_	50S ribosomal protein L18_rplR_																
hypothetical protein_group_270_	DNA gyrase subunit A_gyrA_	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit C_gatC_	50S ribosomal protein L6_rplF_																
hypothetical protein_group_612_	DNA-binding response regulator MtrA_mtrA_	Transcription termination/antitermination protein NusG_nusG_	30S ribosomal protein S8_rpsH_																
hypothetical protein_group_618_	DNA topoisomerase 1_topA_	Transcription termination/antitermination protein NusA_nusA_	Alternate 30S ribosomal protein S14_rpsN2_																
hypothetical protein_group_202_	DNA processing protein DprA_dprA_	Transcriptional repressor NrdR_nrdR_2_	50S ribosomal protein L5_rplE_																
hypothetical protein_group_452_	UvrABC system protein B_uvrB_	Putative transcriptional regulator YvhJ_yvhJ_	50S ribosomal protein L24_rplX_																
hypothetical protein_group_192_	UvrABC system protein C_uvrC_	Transcriptional regulatory protein OmpR_ompR_	50S ribosomal protein L14_rplN_																
hypothetical protein_group_621_	UvrABC system protein A_uvrA_	Alkaline phosphatase synthesis transcriptional regulatory protein PhoP_phoP_2_	30S ribosomal protein S17_rpsQ_																
hypothetical protein_group_453_	ComE operon protein 3_comEC_	Transcriptional regulatory protein SrrA_srrA_	50S ribosomal protein L29_rpmC_																
hypothetical protein_group_454_	Competence protein ComM_comM_	putative transcriptional regulatory protein/MSMEI_2866_pmpR_	50S ribosomal protein L16_rplP_																
hypothetical protein_group_622_	Primosomal protein N'_priA_	Heat-inducible transcription repressor HrcA_hrcA_	30S ribosomal protein S3_rpsC_																
hypothetical protein_group_331_	Replication-associated recombination protein A_rarA_	D-aminoacyl-tRNA deacylase_tatD_	50S ribosomal protein L22_rplV_																
hypothetical protein_group_271_	Tyrosine recombinase XerD_xerD_	Ribosome hibernation promotion factor_hpf_	30S ribosomal protein S19_rpsS_																
hypothetical protein_group_626_		RNA pyrophosphohydrolase_rppH_4_	50S ribosomal protein L2_rplB_																
hypothetical protein_group_630_		Aminopeptidase N_pepN_	50S ribosomal protein L23_rplW_																
hypothetical protein_group_631_		Release factor glutamine methyltransferase_prmC_	50S ribosomal protein L4_rplD_																
hypothetical protein_group_636_			50S ribosomal protein L3_rplC_																
hypothetical protein_group_461_			30S ribosomal protein S10_rpsJ_																
hypothetical protein_group_338_			30S ribosomal protein S7_rpsG_																
hypothetical protein_group_647_			30S ribosomal protein S12_rpsL_																
hypothetical protein_group_780_			50S ribosomal protein L19_rplS_																

hypothetical protein_xseB_			30S ribosomal protein S16_rpsP_																
hypothetical protein_group_649_			Valine--tRNA ligase_valS_																
hypothetical protein_group_401_			Arginine--tRNA ligase_argS_																
hypothetical protein_ftsQ_			Threonine--tRNA ligase_thrS_																
hypothetical protein_group_203_			Proline--tRNA ligase_proS_																
hypothetical protein_group_490_			Serine--tRNA ligase_serS_																
hypothetical protein_group_491_			Glycine--tRNA ligase_glyQS_																
hypothetical protein_group_285_			Glutamate--tRNA ligase 1_gltX1_																
hypothetical protein_group_494_			Cysteine--tRNA ligase_cysS_																
hypothetical protein_group_664_			Tryptophan--tRNA ligase_trpS_																
hypothetical protein_group_497_			Histidine--tRNA ligase_hisS_																
hypothetical protein_tqsA_			Tyrosine--tRNA ligase_tyrS_																
hypothetical protein_group_499_			Phenylalanine--tRNA ligase beta subunit_pheT_																
hypothetical protein_group_500_			Phenylalanine--tRNA ligase alpha subunit_pheS_																
hypothetical protein_group_484_			Leucine--tRNA ligase_leuS_																
hypothetical protein_group_504_			Alanine--tRNA ligase_alaS_																
hypothetical protein_aroK_			Methionine--tRNA ligase_metG_																
hypothetical protein_group_679_			Peptidyl-tRNA hydrolase_pth_																
hypothetical protein_group_743_			Ribosomal RNA small subunit methyltransferase D_rsmD_																
hypothetical protein_group_702_			Ribosomal RNA small subunit methyltransferase A_rsmA_																
hypothetical protein_group_514_			Ribosomal RNA large subunit methyltransferase H_rlmH_																
hypothetical protein_group_741_			Ribosomal RNA small subunit methyltransferase H_rsmH_																
hypothetical protein_group_732_			GTPase Obg_obg_																
hypothetical protein_group_734_			Ribosome-binding ATPase YchF_ychF_																
hypothetical protein_iolC_			Transcription elongation factor GreA_greA_1_																
hypothetical protein_ydeN_1_			Translation initiation factor IF-3_infC_																
hypothetical protein_group_711_			Translation initiation factor IF-2_infB_																
hypothetical protein_group_713_			Translation initiation factor IF-1_infA_																
hypothetical protein_group_714_			Peptide chain release factor 1_prfA_																
hypothetical protein_bcrC_			Peptide chain release factor 2_prfB_																
hypothetical protein_group_522_			SsrA-binding protein_smpB_																
hypothetical protein_group_524_			Methionyl-tRNA formyltransferase_fmt_																
hypothetical protein_group_723_			Methionine aminopeptidase 1_map_																
hypothetical protein_group_534_			Peptide deformylase_def_																
hypothetical protein_group_409_			Ribosome-recycling factor_frr_																
hypothetical protein_group_729_			GTPase Der_der_																

hypothetical protein_group_730_			Putative threonylcarbamoyl-AMP synthase_ywlC_																
hypothetical protein_group_526_																			
hypothetical protein_group_530_																			
hypothetical protein_group_532_																			
hypothetical protein_group_716_																			
hypothetical protein_group_717_																			
hypothetical protein_group_539_																			
hypothetical protein_group_426_																			
hypothetical protein_group_541_																			
hypothetical protein_group_754_																			
hypothetical protein_group_753_																			
hypothetical protein_group_752_																			
hypothetical protein_group_554_																			
hypothetical protein_group_771_																			
hypothetical protein_group_545_																			
hypothetical protein_group_762_																			
hypothetical protein_group_550_																			
hypothetical protein_group_768_																			
hypothetical protein_group_551_																			
hypothetical protein_group_830_																			
hypothetical protein_group_769_																			
hypothetical protein_group_543_																			
hypothetical protein_group_544_																			
hypothetical protein_group_775_																			
hypothetical protein_group_786_																			
hypothetical protein_group_393_																			
hypothetical protein_group_643_																			
hypothetical protein_cyaB_																			
hypothetical protein_group_800_																			
hypothetical protein_group_804_																			
hypothetical protein_group_688_																			
hypothetical protein_group_796_																			
hypothetical protein_group_912_																			
hypothetical protein_group_429_																			
hypothetical protein_group_862_																			
hypothetical protein_pgsA_																			
hypothetical protein_group_866_																			

hypothetical protein_group_835_																			
hypothetical protein_group_873_																			
hypothetical protein_pilE_																			
hypothetical protein_group_908_																			
hypothetical protein_group_867_																			

Table 6. Unique genes comparing full/near complete genomes of TM7x, BB001, AC001 and PM004.

TM7x_110 genes (84 (76.4%) hypothetical)*	**BB001_150 genes** (111 (74%) hypothetical)*	**AC001_279 genes** (216 (77.4%) hypothetical)*	**PM004_273 genes** (190 (69.6%) hypothetical)**
Superoxide dismutase [Mn]_sodA_	hypothetical protein_group_1014_	hypothetical protein_group_1110_	hypothetical protein_group_210_
2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase_lip1_	(2E,6E)-farnesyl diphosphate synthase_sdsA_	hypothetical protein_group_947_	hypothetical protein_group_104_
ABC transporter ATP-binding protein uup_uup_	**ATP-dependent RNA helicase RhlE_rhlE_**	hypothetical protein_group_1092_	putative HIT-like protein_group_518_
Acetyltransferase YpeA_ypeA_	hypothetical protein_group_1292_	hypothetical protein_group_965_	Lipid II:glycine glycyltransferase_femX_2_
Arginine deiminase_group_358_arcA	hypothetical protein_group_1273_	hypothetical protein_group_1081_	hypothetical protein_group_1173_
ATP-dependent zinc metalloprotease FtsH 1_ftsH4_	hypothetical protein_group_1306_	Type II secretion system protein G_epsG_	hypothetical protein_group_1256_
Autolysin_lytA_	hypothetical protein_group_1307_	Phosphatase NudJ_nudJ_	hypothetical protein_group_1033_
Energy-dependent translational throttle protein EttA_group_27_ettA	Phytoene synthase_crtB_	hypothetical protein_rpf2_	Cytidylate kinase_cmk_
Glucose-1-phosphate thymidylyltransferase 1_rfbA_	zeta-carotene-forming phytoene desaturase_carA2_	hypothetical protein_group_971_	Thymidine kinase_tdk_
Glycerol-3-phosphate cytidylyltransferase_tagD_	hypothetical protein_group_1303_	**Modification methylase HhaI_hhaIM_**	putative MscS family protein YkuT_ykuT_1_
hypothetical protein_glcA_	hypothetical protein_group_1302_	hypothetical protein_group_1101_	hypothetical protein_group_1176_
hypothetical protein_group_1125_	4-hydroxybenzoate octaprenyltransferase_ubiA_	hypothetical protein_group_204_	hypothetical protein_group_1037_
hypothetical protein_group_1137_	Adenosine deaminase_add_	hypothetical protein_group_977_	hypothetical protein_group_1142_
hypothetical protein_group_1148_htrA	Argininosuccinate synthase_argG_	Large cysteine-rich periplasmic protein OmcB_group_195_omcB	hypothetical protein_group_352_
hypothetical protein_group_1323_	Argininosuccinate lyase_argH_	Folylpolyglutamate synthase_fpgS_	hypothetical protein_group_1041_
hypothetical protein_group_1357_	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase_arnC_	Divalent-cation tolerance protein CutA_cutA_	Cardiolipin synthase B_clsB_
hypothetical protein_group_1444_	Beta-glucanase_bglA_	putative ABC transporter ATP-binding protein YbhF_ybhF_3_	hypothetical protein_group_1044_
hypothetical protein_group_1445_	Copper-exporting P-type ATPase A_copA_2_	Glutamine--fructose-6-phosphate aminotransferase [isomerizing]_glmS_	S-adenosylmethionine synthase_metK_
hypothetical protein_group_1446_	Bifunctional hemolysin/adenylate cyclase_cya_	Internalin-A_inlA_1_	Zinc transporter ZupT_zupT_
hypothetical protein_group_1447_	HTH-type transcriptional regulator CymR_cymR_	hypothetical protein_group_1247_	hypothetical protein_group_1045_
hypothetical protein_group_2629_	hypothetical protein_group_1000_	Staphylococcal secretory antigen SsaA_ssaA_	hypothetical protein_group_932_
hypothetical protein_group_2630_	hypothetical protein_group_1001_	hypothetical protein_group_368_	hypothetical protein_group_724_
hypothetical protein_group_2631_	hypothetical protein_group_1009_	hypothetical protein_group_370_	Xaa-Pro aminopeptidase_pepP_
hypothetical protein_group_2632_	hypothetical protein_group_1010_	hypothetical protein_group_1249_	50S ribosomal protein L25_rplY_
hypothetical protein_group_2633_	hypothetical protein_group_110_	hypothetical protein_group_97_	Type II secretion system protein G_pulG_9_
hypothetical protein_group_2634_	Riboflavin biosynthesis protein_group_1122_	hypothetical protein_group_1250_	hypothetical protein_group_1156_
hypothetical protein_group_2635_	hypothetical protein_group_1127_	HTH-type transcriptional activator HxlR_hxlR_1_	hypothetical protein_group_1049_
hypothetical protein_group_2636_	hypothetical protein_group_1128_	Phosphoglucomutase_pgcA_	hypothetical protein_group_145_
hypothetical protein_group_2637_	Daunorubicin/doxorubicin resistance ATP-binding protein DrrA_group_12_ybhF	hypothetical protein_group_1090_	hypothetical protein_group_1182_
hypothetical protein_group_2638_	hypothetical protein_group_1270_	Multidrug resistance ABC transporter ATP-binding and permease protein_lmrA_	hypothetical protein_group_1158_
hypothetical protein_group_2639_	hypothetical protein_group_1276_	hypothetical protein_group_979_	DNA polymerase IV_dinB1_
hypothetical protein_group_2640_	hypothetical protein_group_1277_	Nicotinamidase_pncA_	hypothetical protein_group_931_
hypothetical protein_group_2641_	hypothetical protein_group_1278_	Nicotinate phosphoribosyltransferase_pncB_	hypothetical protein_group_1160_
hypothetical protein_group_2642_	hypothetical protein_group_1279_	Non-canonical purine NTP pyrophosphatase_rdgB_1_	hypothetical protein_group_1054_
hypothetical protein_group_2643_	hypothetical protein_group_1282_	hypothetical protein_group_1223_	Xylulose-5-phosphate phosphoketolase_xpkA_
hypothetical protein_group_2644_	hypothetical protein_group_1283_	hypothetical protein_group_1084_	putative adenylyltransferase/sulfurtransferase MoeZ_moeZ_
hypothetical protein_group_2645_	hypothetical protein_group_1293_	Response regulator MprA_mprA_	hypothetical protein_group_843_

hypothetical protein_group_2646_	hypothetical protein_group_1296_	Signal-transduction histidine kinase senX3_senX3_	hypothetical protein_group_1061_
hypothetical protein_group_2647_	hypothetical protein_group_1297_	Bis(5'-nucleosyl)-tetraphosphatase, symmetrical_apaH_	Polyprenol monophosphomannose synthase_group_191_epsH
hypothetical protein_group_2648_	hypothetical protein_group_1299_	Arginine deiminase_arcA_	N-acetyl-alpha-D-glucosaminyl L-malate synthase_tagE_
hypothetical protein_group_2649_	hypothetical protein_group_1300_	N-acetyl-alpha-D-glucosaminyl L-malate synthase_bshA_	GDP-mannose 4,6-dehydratase_gmd_
hypothetical protein_group_2651_	hypothetical protein_group_1585_	Competence protein ComM_comM_2_	hypothetical protein_group_1168_
hypothetical protein_group_2653_	hypothetical protein_group_1586_	Cyanophycin synthetase_cphA_	Mannosylfructose-phosphate synthase_mshA_3_
hypothetical protein_group_2654_	hypothetical protein_group_1587_	Peptide deformylase_defB_	hypothetical protein_group_841_
hypothetical protein_group_2655_	hypothetical protein_group_1588_	Toxin Doc_doc_	Murein DD-endopeptidase MepM_mepM_
hypothetical protein_group_2656_	hypothetical protein_group_1590_	ADP-ribosyl-[dinitrogen reductase] glycohydrolase_draG_	Deoxyuridine 5'-triphosphate nucleotidohydrolase_dut_
hypothetical protein_group_2657_	hypothetical protein_group_1593_	Energy-dependent translational throttle protein EttA_ettA_	hypothetical protein_group_1315_
hypothetical protein_group_2658_	hypothetical protein_group_1594_	Ethanolamine utilization protein EutD_eutD_	hypothetical protein_group_757_
hypothetical protein_group_2660_	hypothetical protein_group_1595_	3-oxoacyl-[acyl-carrier-protein] reductase FabG_fabG_	putative sulfate transporter_group_791_
hypothetical protein_group_2661_	hypothetical protein_group_1596_	Adenosine monophosphate-protein transferase SoFic_fic_	putative ABC transporter ATP-binding protein YknY_group_1_macB_2
hypothetical protein_group_2662_	hypothetical protein_group_1597_	Glycogen synthase_glgA_	Protease HtpX_htpX_1_
hypothetical protein_group_2663_	hypothetical protein_group_1598_	hypothetical protein_group_1077_	Lipase 3_lip3_1_
hypothetical protein_group_2664_	hypothetical protein_group_1599_	hypothetical protein_group_1079_	Putative carboxypeptidase YodJ_yodJ_
hypothetical protein_group_2665_	hypothetical protein_group_1600_	hypothetical protein_group_1094_	Fimbrial protein_pilE1_2_
hypothetical protein_group_2666_	hypothetical protein_group_1601_	hypothetical protein_group_1096_	hypothetical protein_group_1149_
hypothetical protein_group_2667_	hypothetical protein_group_1602_	hypothetical protein_group_1118_	Glycerophosphodiester phosphodiesterase_glpQ_
hypothetical protein_group_2668_	hypothetical protein_group_1603_	hypothetical protein_group_1192_	hypothetical protein_group_1347_
hypothetical protein_group_2669_	hypothetical protein_group_1605_	hypothetical protein_group_1195_	D-inositol 3-phosphate glycosyltransferase_mshA_1_
hypothetical protein_group_2670_	hypothetical protein_group_1606_	hypothetical protein_group_1196_	hypothetical protein_group_1047_
hypothetical protein_group_2671_	hypothetical protein_group_1607_	hypothetical protein_group_1197_	hypothetical protein_group_1371_
hypothetical protein_group_2672_	hypothetical protein_group_1608_	hypothetical protein_group_1198_	hypothetical protein_group_1375_
hypothetical protein_group_2673_	hypothetical protein_group_1609_	hypothetical protein_group_1199_	Succinate--CoA ligase [GDP-forming] subunit beta_sucC_
hypothetical protein_group_2674_	hypothetical protein_group_1610_	hypothetical protein_group_1200_	Succinate--CoA ligase [ADP-forming] subunit alpha_sucD_
hypothetical protein_group_2675_	hypothetical protein_group_1611_	hypothetical protein_group_1202_	Phosphatidyl-myo-inositol mannosyltransferase_pimA_2_
hypothetical protein_group_2676_	hypothetical protein_group_1612_	hypothetical protein_group_1203_	50S ribosomal protein L34_rpmH_
hypothetical protein_group_2677_	hypothetical protein_group_1613_	hypothetical protein_group_1204_	hypothetical protein_group_2009_
hypothetical protein_group_2678_	hypothetical protein_group_1614_	hypothetical protein_group_1205_	hypothetical protein_group_2010_
hypothetical protein_group_2679_	hypothetical protein_group_1615_	hypothetical protein_group_1207_	hypothetical protein_group_106_
hypothetical protein_group_2680_	hypothetical protein_group_1616_	hypothetical protein_group_1208_	hypothetical protein_group_1384_
hypothetical protein_group_2681_	hypothetical protein_group_1617_	hypothetical protein_group_1209_	hypothetical protein_group_100_
hypothetical protein_group_2682_	hypothetical protein_group_1618_	hypothetical protein_group_1214_	hypothetical protein_group_2011_
hypothetical protein_group_2683_	hypothetical protein_group_1620_	hypothetical protein_group_1215_	hypothetical protein_group_2012_
hypothetical protein_group_2684_	hypothetical protein_group_1621_	hypothetical protein_group_1216_	hypothetical protein_group_1378_
hypothetical protein_group_2685_	hypothetical protein_group_1622_	hypothetical protein_group_1218_	hypothetical protein_group_2013_
hypothetical protein_group_2686_	hypothetical protein_group_1623_	hypothetical protein_group_1220_	hypothetical protein_group_2014_
hypothetical protein_group_2688_	hypothetical protein_group_1624_	hypothetical protein_group_1221_	Ribosome-associated ATPase_rbbA_1_
hypothetical protein_group_2689_	hypothetical protein_group_1625_	hypothetical protein_group_1227_	hypothetical protein_group_1385_
hypothetical protein_group_2690_	hypothetical protein_group_1626_	hypothetical protein_group_1230_	hypothetical protein_group_1346_
hypothetical protein_group_2691_	hypothetical protein_group_1627_	hypothetical protein_group_1232_	hypothetical protein_group_2015_

hypothetical protein_group_2692_	hypothetical protein_group_1628_	hypothetical protein_group_1235_	Transcription elongation factor GreA_group_2016_greA_1
hypothetical protein_group_2693_	hypothetical protein_group_1629_	hypothetical protein_group_1237_	Non-canonical purine NTP pyrophosphatase_group_251_
hypothetical protein_group_2694_	hypothetical protein_group_1630_	hypothetical protein_group_1239_	Anti-sigma-I factor RsgI2_group_2017_rsgI2
hypothetical protein_group_2695_	hypothetical protein_group_1631_	hypothetical protein_group_1240_	hypothetical protein_group_2018_
hypothetical protein_group_2696_	hypothetical protein_group_1632_	hypothetical protein_group_1242_	hypothetical protein_group_2019_
hypothetical protein_group_2697_	hypothetical protein_group_1633_	hypothetical protein_group_1243_	hypothetical protein_group_2020_
hypothetical protein_group_2699_	hypothetical protein_group_1634_	hypothetical protein_group_1245_	hypothetical protein_group_1386_
hypothetical protein_group_2700_	hypothetical protein_group_1636_	hypothetical protein_group_1251_	hypothetical protein_group_1387_
hypothetical protein_group_354_	hypothetical protein_group_1637_	hypothetical protein_group_1252_	hypothetical protein_group_1388_
hypothetical protein_group_413_	hypothetical protein_group_1638_	hypothetical protein_group_1259_	hypothetical protein_group_2021_
hypothetical protein_group_414_	hypothetical protein_group_1639_	hypothetical protein_group_1260_	hypothetical protein_group_2022_
hypothetical protein_group_415_	hypothetical protein_group_1640_	hypothetical protein_group_1263_	hypothetical protein_group_2023_
hypothetical protein_group_562_	hypothetical protein_group_1641_	hypothetical protein_group_1264_	hypothetical protein_group_1389_
hypothetical protein_group_574_	hypothetical protein_group_1642_	hypothetical protein_group_1265_	hypothetical protein_group_2024_
hypothetical protein_group_836_	hypothetical protein_group_1643_	hypothetical protein_group_1266_	hypothetical protein_group_859_
Non-canonical purine NTP pyrophosphatase_group_249_	hypothetical protein_group_1644_	hypothetical protein_group_1267_	hypothetical protein_group_2025_
Polyamine aminopropyltransferase_group_2652_speE	hypothetical protein_group_1645_	Alpha-monoglucosyldiacylglycerol synthase_group_128_mgs	hypothetical protein_group_2026_
putative ABC transporter ATP-binding protein_group_15_	hypothetical protein_group_1646_	hypothetical protein_group_1448_	hypothetical protein_group_2027_
putative glycosyltransferase_group_184_	hypothetical protein_group_1648_	hypothetical protein_group_1449_	hypothetical protein_group_2028_
putative type I restriction enzymeP M protein_group_1167_	hypothetical protein_group_1649_	hypothetical protein_group_1450_	hypothetical protein_group_1390_
putative type I restriction enzymeP M protein_hsdM_	hypothetical protein_group_1650_	hypothetical protein_group_1451_	hypothetical protein_group_2029_
Putative type-1 restriction enzyme specificity protein MG438_group_2698_	hypothetical protein_group_1651_	hypothetical protein_group_1452_	Lipoprotein-releasing system ATP-binding protein LolD_group_31_lolD_1
Ribosomal protein S6--L-glutamate ligase_gshB_	hypothetical protein_group_1652_	hypothetical protein_group_1453_	hypothetical protein_group_872_
RNA pyrophosphohydrolase_group_141_rppH_2	hypothetical protein_group_1653_	hypothetical protein_group_1454_	hypothetical protein_group_2030_
S-adenosylmethionine decarboxylase proenzyme_speH_	hypothetical protein_group_1656_	hypothetical protein_group_1455_	hypothetical protein_group_2031_
Sensor kinase CusS_group_77_cusS	hypothetical protein_group_1657_	hypothetical protein_group_1456_	hypothetical protein_group_392_inlA
Sensory transduction protein regX3_group_86_regX3	hypothetical protein_group_1658_	hypothetical protein_group_1457_	RNA pyrophosphohydrolase_group_868_rppH_3
Transcriptional regulator WhiD_whiB4_	hypothetical protein_group_1659_	hypothetical protein_group_1458_	hypothetical protein_group_2032_
Type I restriction enzyme EcoR124II R protein_hsdR_2_	hypothetical protein_group_1660_	hypothetical protein_group_1459_	**CRISPR-associated endoribonuclease Cas2_cas2_**
Type-1 restriction enzyme R protein_hsdR_1_	hypothetical protein_group_1661_	hypothetical protein_group_1460_	**CRISPR-associated endonuclease Cas1_cas1_**
Tyrosine recombinase XerC_xerC_	hypothetical protein_group_1663_	hypothetical protein_group_1461_	**CRISPR-associated endonuclease Cas9_cas9_**
	hypothetical protein_group_1664_	hypothetical protein_group_1462_	hypothetical protein_group_2036_
	hypothetical protein_group_1666_	hypothetical protein_group_1464_	Chromosome partition protein Smc_smc_
	hypothetical protein_group_1667_	hypothetical protein_group_1465_	hypothetical protein_group_2038_
	hypothetical protein_group_1668_	hypothetical protein_group_1466_	hypothetical protein_group_2039_
	hypothetical protein_group_1669_	hypothetical protein_group_1467_	hypothetical protein_group_2040_
	hypothetical protein_group_1670_	hypothetical protein_group_1468_	hypothetical protein_group_2041_
	hypothetical protein_group_1671_	hypothetical protein_group_1469_	hypothetical protein_group_2042_
	hypothetical protein_group_1672_	hypothetical protein_group_1470_	hypothetical protein_group_2043_
	hypothetical protein_group_1673_	hypothetical protein_group_1471_	Adenylosuccinate synthetase_purA_
	hypothetical protein_group_1674_	hypothetical protein_group_1472_	hypothetical protein_group_2045_

	hypothetical protein_group_1675_	hypothetical protein_group_1473_	hypothetical protein_group_2046_
	hypothetical protein_group_1676_	hypothetical protein_group_1474_	Protease HtpX_htpX_2_
	hypothetical protein_group_1678_	hypothetical protein_group_1475_	hypothetical protein_group_2047_
	hypothetical protein_group_1680_	hypothetical protein_group_1476_	hypothetical protein_group_2048_
	putative ABC transporter ATP-binding protein_group_17_	hypothetical protein_group_1477_	hypothetical protein_group_2049_
	hypothetical protein_group_194_	hypothetical protein_group_1478_	Phosphoserine phosphatase SerB1_serB1_2_
	Glucose-1-phosphate thymidylyltransferase 1_group_259_rfbA	hypothetical protein_group_1479_	hypothetical protein_group_2050_
	Putative carboxypeptidase YodJ_group_302_yodJ_1	hypothetical protein_group_1480_	PGL/p-HBAD biosynthesis glycosyltransferase_group_186_
	putative FAD-linked oxidoreductase_group_317_	hypothetical protein_group_1481_	hypothetical protein_group_1392_
	Lipoprotein-releasing system ATP-binding protein LolD_group_32_lolD_2	hypothetical protein_group_1482_	hypothetical protein_group_1393_
	hypothetical protein_group_333_	hypothetical protein_group_1483_	hypothetical protein_group_2051_
	hypothetical protein_group_573_	hypothetical protein_group_1484_	Type II secretion system protein G_xcpT_1_
	Signal transduction histidine-protein kinase ArlS_group_75_arlS	hypothetical protein_group_1485_	Acetate kinase_ackA_
	hypothetical protein_group_845_	hypothetical protein_group_1486_	rRNA adenine N-6-methyltransferase_ermA1_
	Response regulator MprA_group_85_mprA	hypothetical protein_group_1487_	hypothetical protein_group_1395_
	GMP synthase [glutamine-hydrolyzing]_guaA_	hypothetical protein_group_1488_	hypothetical protein_group_2052_
	Ycf48-like protein_hcf136_	hypothetical protein_group_1489_	hypothetical protein_group_2053_
	Hemolysin secretion protein D, chromosomal_hlyD_	hypothetical protein_group_1490_	hypothetical protein_group_2054_
	Tyrosine recombinase XerC_Int-Tn_	hypothetical protein_group_1491_	N-acetylmuramoyl-L-alanine amidase sle1_group_340_sle1_1
	Penicillin-binding protein 4*_pbpE_	hypothetical protein_group_1492_	hypothetical protein_group_1396_
	Serine/threonine-protein kinase pkn5_pkn5_	hypothetical protein_group_1493_	hypothetical protein_group_1397_
	dTDP-glucose 4,6-dehydratase_rmlB_	hypothetical protein_group_1494_	hypothetical protein_group_2055_
	Anti-sigma-I factor RsgI2_rsgI2_	hypothetical protein_group_1495_	hypothetical protein_group_2056_
	Ribosomal RNA small subunit methyltransferase F_rsmF_	hypothetical protein_group_1496_	hypothetical protein_group_2057_
	Serine 3-dehydrogenase_sdh_	hypothetical protein_group_1497_	Urease accessory protein UreG_ureG_
	tRNA-specific adenosine deaminase_tadA_	hypothetical protein_group_1498_	hypothetical protein_group_2059_
	dTDP-6-deoxy-L-talose 4-dehydrogenase (NAD( ))_tll_	hypothetical protein_group_1499_	Inorganic triphosphatase_group_2060_
	putative transcriptional regulatory protein YedW_yedW_2_	hypothetical protein_group_1500_	hypothetical protein_group_2061_
	ABC transporter ATP-binding protein YtrB_ytrB_	hypothetical protein_group_1501_	hypothetical protein_group_2062_
	ESX secretion system protein YueB_yueB_	hypothetical protein_group_1502_	hypothetical protein_group_2063_
		hypothetical protein_group_1503_	hypothetical protein_group_2064_
		hypothetical protein_group_1504_	hypothetical protein_group_1398_
		hypothetical protein_group_1505_	hypothetical protein_group_2065_
		hypothetical protein_group_1506_	Thymidylate kinase_tmk_2_
		hypothetical protein_group_1507_	hypothetical protein_group_57_
		hypothetical protein_group_1508_	hypothetical protein_group_2067_
		hypothetical protein_group_1509_	hypothetical protein_group_1399_
		hypothetical protein_group_1510_	hypothetical protein_group_2068_
		hypothetical protein_group_1511_	hypothetical protein_group_1177_
		hypothetical protein_group_1512_	hypothetical protein_group_1178_
		hypothetical protein_group_1513_	hypothetical protein_group_2069_
		hypothetical protein_group_1514_	hypothetical protein_group_2070_
		hypothetical protein_group_1515_	TVP38/TMEM64 family inner membrane protein YdjZ_ydjZ_

		hypothetical protein_group_1516_	hypothetical protein_group_2072_
		hypothetical protein_group_1517_	hypothetical protein_group_2073_
		hypothetical protein_group_1518_	hypothetical protein_group_2074_
		hypothetical protein_group_1519_	S-adenosylmethionine:tRNA ribosyltransferase-isomerase_queA_
		hypothetical protein_group_1520_	hypothetical protein_group_2075_
		hypothetical protein_group_1521_	hypothetical protein_group_1401_
		hypothetical protein_group_1522_	hypothetical protein_group_2076_
		hypothetical protein_group_1523_	hypothetical protein_group_2077_
		hypothetical protein_group_1524_	hypothetical protein_group_1062_
		hypothetical protein_group_1525_	hypothetical protein_group_1064_
		hypothetical protein_group_1526_	hypothetical protein_group_1065_
		hypothetical protein_group_1527_	hypothetical protein_group_1063_
		hypothetical protein_group_1528_	hypothetical protein_group_2078_
		hypothetical protein_group_1529_	hypothetical protein_group_1402_
		hypothetical protein_group_1530_	hypothetical protein_group_2079_
		hypothetical protein_group_1531_	Histidine--tRNA ligase_hisS_2_
		hypothetical protein_group_1532_	hypothetical protein_group_2080_
		hypothetical protein_group_1533_	hypothetical protein_group_2081_
		hypothetical protein_group_1534_	hypothetical protein_group_1403_
		hypothetical protein_group_1535_	hypothetical protein_group_2082_
		hypothetical protein_group_1536_	Bifunctional (p)ppGpp synthase/hydrolase RelA_relA_2_
		hypothetical protein_group_1537_	hypothetical protein_group_2083_
		hypothetical protein_group_1538_	hypothetical protein_group_2084_
		hypothetical protein_group_1539_	hypothetical protein_group_2085_
		hypothetical protein_group_1540_	hypothetical protein_group_906_
		hypothetical protein_group_1541_	Proline iminopeptidase_pip_
		hypothetical protein_group_1542_	Nucleoid occlusion factor SlmA_slmA_
		hypothetical protein_group_1543_	Putative aminoacrylate hydrolase RutD_rutD_2_
		Cardiolipin synthase_group_1544_	hypothetical protein_group_2089_
		hypothetical protein_group_1545_	Putative mutator protein MutT4_mutT4_2_
		hypothetical protein_group_1546_	hypothetical protein_group_2090_
		hypothetical protein_group_1547_	hypothetical protein_group_2091_
		hypothetical protein_group_1548_	hypothetical protein_group_1377_
		hypothetical protein_group_1549_	hypothetical protein_group_2092_
		hypothetical protein_group_1550_	hypothetical protein_group_2093_
		hypothetical protein_group_1551_	dTDP-fucosamine acetyltransferase_wecD_
		hypothetical protein_group_1552_	hypothetical protein_group_2094_
		hypothetical protein_group_1553_	hypothetical protein_group_2095_
		hypothetical protein_group_1554_	D-alanyl-D-alanine carboxypeptidase_group_2096_
		hypothetical protein_group_1555_	hypothetical protein_group_2097_
		hypothetical protein_group_1556_	hypothetical protein_group_2098_
		hypothetical protein_group_1558_	hypothetical protein_group_2099_
		hypothetical protein_group_1559_	Type II secretion system protein G_xcpT_2_
		hypothetical protein_group_1560_	hypothetical protein_group_576_
		hypothetical protein_group_1561_	hypothetical protein_group_998_
		hypothetical protein_group_1562_	hypothetical protein_group_792_
		hypothetical protein_group_1563_	Septum formation protein Maf_maf_
		hypothetical protein_group_1564_	hypothetical protein_group_2101_
		hypothetical protein_group_1565_	hypothetical protein_group_2102_
		hypothetical protein_group_1566_	hypothetical protein_group_2103_
		hypothetical protein_group_1567_	hypothetical protein_group_2104_
		Lysine--tRNA ligase_group_1568_lysS	hypothetical protein_group_2105_

		hypothetical protein_group_1569_	hypothetical protein_group_999_
		hypothetical protein_group_1570_	hypothetical protein_group_2106_
		hypothetical protein_group_1571_	hypothetical protein_group_2107_
		hypothetical protein_group_1572_	hypothetical protein_group_1404_
		hypothetical protein_group_1573_	RNA pyrophosphohydrolase_group_2108_rppH_4
		hypothetical protein_group_1574_	hypothetical protein_group_2109_
		hypothetical protein_group_1575_	hypothetical protein_group_2110_
		hypothetical protein_group_1577_	Fluorothreonine transaldolase_group_423_
		hypothetical protein_group_1578_	hypothetical protein_group_2111_
		hypothetical protein_group_1579_	hypothetical protein_group_2112_
		hypothetical protein_group_1580_	hypothetical protein_group_2113_
		Undecaprenyl-phosphate mannosyltransferase_group_1581_	hypothetical protein_group_2114_
		hypothetical protein_group_1582_	hypothetical protein_group_2115_
		hypothetical protein_group_1583_	hypothetical protein_group_327_
		hypothetical protein_group_1584_	hypothetical protein_group_284_
		Lipase 1_group_220_	hypothetical protein_group_2116_
		hypothetical protein_group_222_	hypothetical protein_group_2117_
		hypothetical protein_group_223_	hypothetical protein_group_2118_
		Lipase_group_226_	hypothetical protein_group_2119_
		hypothetical protein_group_254_	hypothetical protein_group_2120_
		Aspartate--tRNA ligase_group_278_aspS	Magnesium and cobalt efflux protein CorC_group_289_corC_1
		hypothetical protein_group_311_	hypothetical protein_group_2121_
		putative oxidoreductase_group_398_	Glucose-1-phosphate adenylyltransferase_glgC_
		hypothetical protein_group_402_	ADP-ribose pyrophosphatase_group_475_act
		hypothetical protein_group_430_	Elongation factor 4_group_117_lepA_1
		Isoleucine--tRNA ligase 2_group_439_ileS2	hypothetical protein_group_2123_
		hypothetical protein_group_560_	hypothetical protein_group_2124_
		hypothetical protein_group_575_	hypothetical protein_group_2125_
		hypothetical protein_group_578_	hypothetical protein_group_2126_
		hypothetical protein_group_81_	hypothetical protein_group_2127_
		hypothetical protein_group_817_	hypothetical protein_group_2128_
		hypothetical protein_group_818_	ATP-dependent zinc metalloprotease FtsH_group_347_ftsH
		hypothetical protein_group_819_	hypothetical protein_group_2129_
		hypothetical protein_group_860_	hypothetical protein_group_1379_
		aspartoacylase_group_864_	hypothetical protein_group_2130_
		hypothetical protein_group_969_	hypothetical protein_group_1380_
		hypothetical protein_group_973_	hypothetical protein_group_2131_
		hypothetical protein_group_974_	hypothetical protein_group_2132_
		hypothetical protein_group_98_	Inner membrane protein YrbG_yrbG_
		hypothetical protein_group_986_	Macrolide export ATP-binding/permease protein MacB_macB_3_
		D-psicose 3-epimerase_group_996_	HTH-type transcriptional regulator ImmR_immR_
		Metallopeptidase ImmA_immA_	hypothetical protein_group_2134_
		Internalin-A_inlA_2_	Sensor histidine kinase ResE_resE_2_
		Lipoprotein-releasing system ATP-binding protein LolD_lolD_3_	Response regulator MprA_group_89_mprA
		Diaminopimelate decarboxylase_lysA_	Phosphoglycolate phosphatase_gph_
		Mercuric resistance operon regulatory protein_merR_	Fatty acid metabolism regulator protein_fadR_
		Mycothiol acetyltransferase_mshD_	hypothetical protein_group_291_
		putative 8-oxo-dGTP diphosphatase 1_mutT1_	hypothetical protein_group_869_

		Putative 8-oxo-dGTP diphosphatase 2_mutT2_	hypothetical protein_group_2136_
		hypothetical protein_mutX_	hypothetical protein_group_1405_
		NH(3)-dependent NAD( ) synthetase_nadE_	Lipase 2_group_215_
		endonuclease 4_nfo_	hypothetical protein_group_2137_
		NADH pyrophosphatase_nudC_	hypothetical protein_group_2138_
		Nucleoside triphosphatase NudI_nudI_	hypothetical protein_group_2139_
		Oligoribonuclease_orn_2_	hypothetical protein_group_2140_
		dITP/XTP pyrophosphatase_rdgB_	GDP-perosamine synthase_perA_
		Uridine kinase_udk_	hypothetical protein_group_133_
		dTDP-fucosamine acetyltransferase_wecD_2_	hypothetical protein_group_2141_
		Tyrosine recombinase XerC_xerC_2_	
		Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase_yfdH_	_
		Putative NAD(P)H nitroreductase_yfkO_	
		putative ABC transporter ATP-binding protein YknY_yknY_2_	
		Putative carboxypeptidase YodJ_yodJ_2_	
		ABC transporter permease YtrF_ytrF_2_	

*Restriction-modification genes are bolded.

**CRISPR-Cas9 genes are bolded.

Table 7. Summary of restriction-modification and orphan methyltransferase systems of Saccharibacteria isolates.

Strain	Type	System	Predicted Recognition Sequence*	MTases active in the genome	Genes	Name	Coordinates
TM7x	I	R-M	TT**A**NNNNNNN**T**ARC	Active	S	S1.SorTM7II	327806-328825
					S	S2.SorTMIIP	329836-330966c
	I	R-M	CA**A**YNNNNR**T**C	Active	R	SorTM7IP	655481-658615c
					M	M.SorTM7I	661559-662668c
					S	S.SorTM7I	664329-665816c
BB001	III	R-M	AA**A**TC	Active	R	SspTM7IP	262014-264926c
					M	M.SspTM7I	265685-267694c
AC001	II	R-M	Unknown	Unknown	R	SstACORFBP	183419-185092 c
					M	M.SstACORFBP	185089-186144 c
	III	R-M	CGC**A**T	Active	R	SstACIP	354085-356841 c
					M	M.SstACI	360026-361870 c

*Bolded nucletotides are putative methylation sites by PacBio sequencing analysis

Table 8. PM004 CRISPR RNA array

Number of spacers/repeat	Position	Repeat Sequence	Spacer Sequence
1	2065	ATTATACCAATATCTAAATTGTCAGGGAACTATAAC	TTTATTGACGTATGGTGGGTCATTGGTTAA
2	1999	. . . .	ATGTAATACCTGCTTCCTTTCTACAAAAAT
3	1933	. . . .	TGTCGGCTACGCTTACTCCATAAGAGTAAA
4	1867	. . . .	TTTGAAAAGGTCTATGAGGTGCGCGGAGAT
5	1801	. . . .	TCAGTTAGTTTGCTCAACTGTTCTTCAACT
6	1735	. . . .	TTTTAAACAGTCAATGGTTTTATACGTTTG
7	1669	. . . .	GAGATAAATACAGACTGCATCGTTTCCTCT
8	1603	. . . .	TTTGCCCGTGAACAAAACCATTCCGATAGG
9	1537	. . . .	CCCGACAACTTGATATCACGCTTCAGCCGT
10	1471	. . . .	GCTCGATCAAGCAGGTGGCTCTATCTAAGG
11	1405	. . . .	CAAGCACCGATCGTTGACCGCGACGGTGAT
12	1339	. . . .	GAGATAAATACAGACTGCATCGTTTCCTCT
13	1273	. . . .	TTTGCCCGTGAACAAAACCATTCCGATAGG
14	1207	. . . .	CCCGACAACTTGATATCACGCTTCAGCCGT
15	1141	. . . .	GCTCGATCAAGCAGGTGGCTCTATCTAAGG
16	1075	. . . .	CAAGCACCGATCGTTGACCGCGACGGTGAT
17	1009	. . . .	TCTGTTCAACTGGTAACACCAACGCACCTT
18	943	. . . .	ATCGCAATCTATTGGGTCGTACGATTTTTA
19	877	. . . .	GTCAAATATGTCGCAAAATAAAGTATATTT
20	811	. . . .	CCACATTTCTTACACAGCTCTTTAGCTTTA
21	745	. . . .	ACTAGCCAATTGCACGAGAAGGACGAGTAT
22	679	. . . .	TAGTTTTCCTTCAGTAGTTAAATTCTCTAA
23	613	. . . .	ACTAGCCAATTGCACGAGAAGGACGAGTAT
24	547	. . . .	TAGTTTTCCTTCAGTAGTTAAATTCTCTAA
25	481	ATTATACCAATATCTAAATTGTCAGGGAACTATGGC	

Table 9. Calculated generation/doubling time for each bulk assay measurement/replicate shown in minutes.

Cells types/Replicates	1	2	3	4	5	Average	Std deviation
XH001	138.583	120.15	102.7224			120.4851333	17.93264882
XH001+ TM7x	144.76	159.06176	157.07557			153.6324433	7.747672277
F0337	110.5257	136.0748	121.3825	113.3766	132.3474	122.7414	11.27856602
F0337+BB001	143.2952	134.4476	129.6742	140.9998	142.1557	138.1145	5.834824884

APPENDIX REFRENCES

Bor B, Poweleit N, Bois JS, Cen L, Bedree JK, Zhou ZH, Gunsalus RP, Lux R, McLean JS, He XS et al. 2016. Phenotypic and physiological characterization of the epibiotic interaction between tm7x and its basibiont actinomyces. Microbial Ecology. 71(1):243-255.

Brinig MM, Lepp PW, Ouverney CC, Armitage GC, Relman DA. 2003. Prevalence of bacteria of division tm7 in human subgingival plaque and their association with disease. Applied and Environmental Microbiology. 69(3):1687-1694.

Bushnell B. 2015. BBDuk: Adapter/quality trimming and filtering, Software, https://www.Geneious.Com/plugins/bbduk/

Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14(7):1394-1403.

Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. 2015. Interactivenn: A web-based tool for the analysis of sets through venn diagrams. BMC Bioinformatics. 16:169.

Hugenholtz P, Tyson GW, Webb RI, Wagner AM, Blackall LL. 2001. Investigation of candidate division tm7, a recently recognized major lineage of the domain bacteria with no known pure-culture representatives. Appl Environ Microbiol. 67(1):411-419.

Hutchison CA, 3rd, Chuang RY, Noskov VN, Assad-Garcia N, Deerinck TJ, Ellisman MH, Gill J, Kannan K, Karas BJ, Ma L et al. 2016. Design and synthesis of a minimal bacterial genome. Science. 351(6280):aad6253.

Kanehisa M, Sato Y, Morishima K. 2016. Blastkoala and ghostkoala: Kegg tools for functional characterization of genome and metagenome sequences. J Mol Biol. 428(4):726-731.

Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics. 31(22):3691-3693.