SUPPLEMENTAL DATA:

Characterization of a Metal-Independent CAZy Family 6 Glycosyltransferase from Bacteroides ovatus

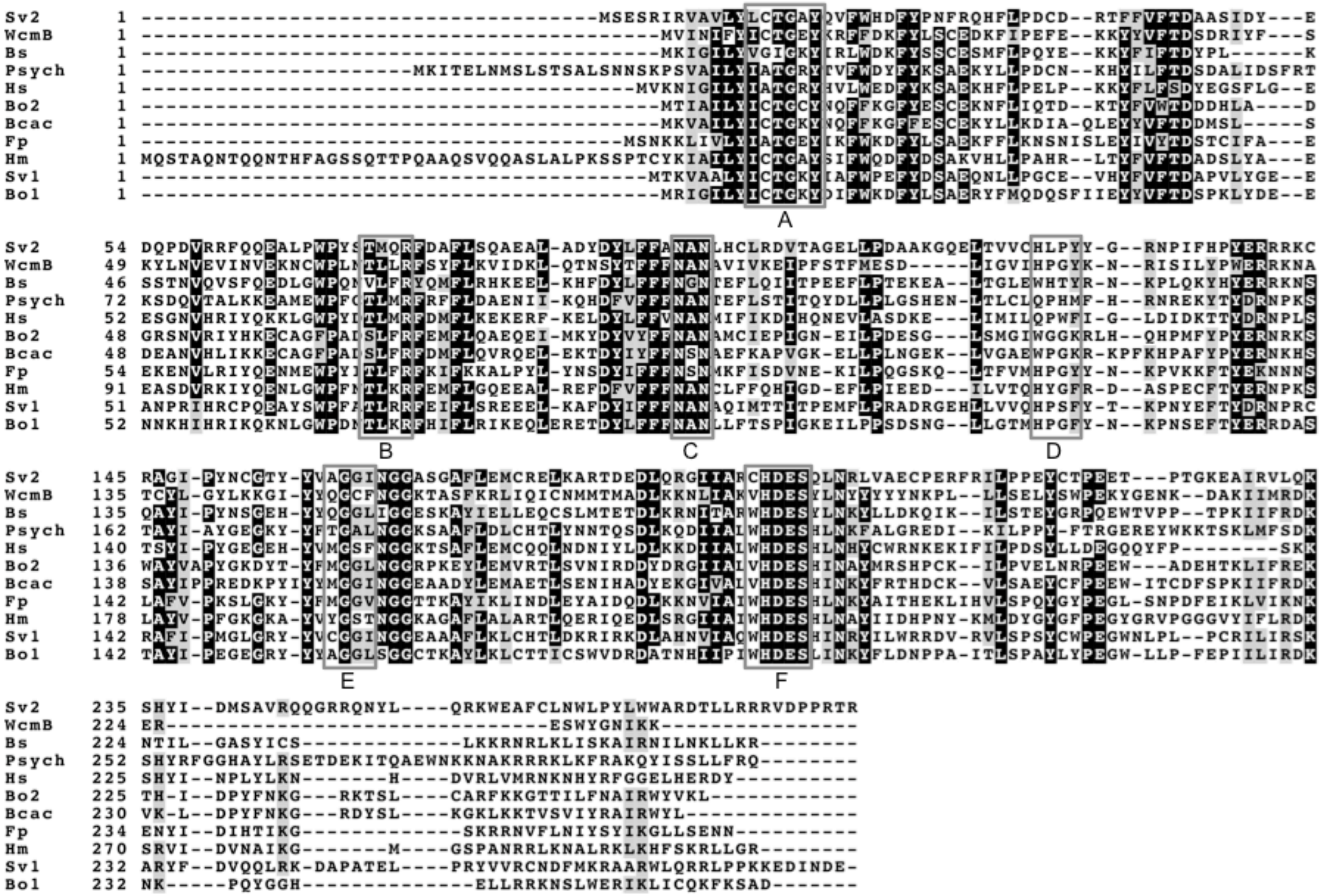
Percy Tumbale and Keith Brew

Supplemental figure legends:

Supplemental Figure 1. An alignment of bacterial GT6 glycosyltransferases. The alignment was generated using MUSCLE and shaded using BOXSHADE. Positions with shaded backgrounds are identical in 70% of the sequences. The boxes denote regions of the active site that discussed in the text. The sequences are Wcmb, *Escherichia coli* Wcmb (AAV85960); Psych, hypothetical protein PsycPRwf_0985 from *Psychrobacter sp. PRwf-1* (YP_001279885.1); Hs, putative glycosyltransferase from *Haemophilus somnus* 2336 (YP_001783501); Bs, *Bacteroides stercoris* (B0NSM3); Bo1, *Bacteroides ovatus* GT6a (A7LVT2); Bo2, *Bacteroides ovatus* GT6b (A7M0P3); Bc, *Bacteroides caccae* (A5ZC71); Fp, *Francisella philomiragia subsp. philomiragia ATCC 25015* (cont1.16, whole genome shotgun sequence) ; Hm, *Helicobacter mustelae* (from the Wellcome Trust Sanger Institute Microbial Pathogens database (www.sanger.ac.uk); Sv1 and Sv2, GT6 sequences from *Subdoligranulum variabile* (NCBI Whole-Genome Shotgun Reads database).

Supplemental Figure 2. An alignment of mammalian, cyanophage PSSM-2 and selected putative phage GT6s from the Environmental Samples database. The sequences are : bova, bovine α-1,3 galactosyltransferase (GGTA1_BOVIN); dogFS, canine Forssman glycolipid synthase (GBGT1_CANFA); huA, human histo-blood group A synthase (BGAT_HUMAN); PSSM2, cyanophage PSSM-2 (Q58M87). The sequences designated GOS are from the Marine Metagenome: GOS_A, ctg_1101668274631; GOS_B, 1096626241734; GOS_C, 1096626359338; GOS_D, 1096626608311; GOS_E, 1096626802277; GOS_F, ctg_1101667133928; GOS_G, ctg_1101668196376; GOS_H, ctg_1096626132935; GOS_I, ctg_1096626203656.

Supplemental Figure 3. A homology model of the 3D structure of BoGT6a based on the structure of bovine α3GT showing the residues subjected to mutagenesis. The coordinates were produced through the Swiss-Model website (swissmodel.expasy.org) and the structure is displayed in ribbon format using Chimera (Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004) *J. Comput. Chem.* **25** 1605-1612). The chain is colored blue to red from N-terminus to C-terminus; the side chains of residues that have major effects on activity are purple and others are dark grey.



```
Sv2      1 -------------------------------------MSESRIRVAVLYLCTGAYQVFWHDFYPNFRQHFLPDCD--RTFFVFTDAASIDY---E
WcmB     1 ------------------------------------------MVINIFYICTGEYKRFFDKFYLSCEDKFIPEFE--KKYYVFTDSDRIYF---S
Bs       1 ------------------------------------------MKIGILYVGIGKYIRLWDKFYSSCESMFLPQYE--KKYFIFTDYPL------K
Psych    1 ----------------------MKITELNMSLSTSALSNNSKPSVAILYIATGRYTVFWDYFYKSAEKYLLPDCN--KHYILFTDSDALIDSFRT
Hs       1 ----------------------------------------MVKNIGILYIATGRYHVLWEDFYKSAEKHFLPELP--KKYFLFSDYEGSFLG--E
Bo2      1 ------------------------------------------MTIAILYICTGCYNQFFKGFYESCEKNFLIQTD--KTYFVWTDDDHLA----D
Bcac     1 ------------------------------------------MKVAILYICTGKYNQFFKGFFESCEKYLLKDIA-QLEYYVFTDDMSL------S
Fp       1 ---------------------------------------MSNKKLIVLYIATGEYIKFWKDFYLSAEKFFLKNSNISLEYIVYTDSTCIFA---E
Hm       1 MQSTAQNTQQNTHFAGSSQTTPQAAQSVQQASLALPKSSPTCYKIAILYICTGAYSIFWQDFYDSAKVHLLPAHR--LTYFVFTDADSLYA---E
Sv1      1 -----------------------------------------MTKVAALYICTGKYIAFWPEFYDSAEQNLLPGCE--VHYFVFTDAPVLYGE--E
Bo1      1 ------------------------------------------MRIGILYICTGKYDIFWKDFYLSAERYFMQDQSFIIEYYVFTDSPKLYDE--E

Sv2     54 DQPDVRRFQQEALPWPYSTMQRFDAFLSQAEAL-ADYDYLFFANANLHCLRDVTAGELLPDAAKGQELTVVCHLPYY-G--RNPIFHPYERRRKC
WcmB    49 KYLNVEVINVEKNCWPLNTLLRFSYFLKVIDKL-QTNSYTFFFNANAVIVKEIPFSTFMESD-----LIGVIHPGYK-N--RISILYPWERRKNA
Bs      46 SSTNVQVSFQEDLGWPQNVLFRYQMFLRHKEEL-KHFDYLFFFNGNTEFLQIITPEEFLPTEKEA--LTGLEWHTYR-N--KPLQKYHYERRKNS
Psych   72 KSDQVTALKKEAMEWPFQTLMRFRFFLDAENII-KQHDFVFFFNANTEFLSTITQYDLLPLGSHEN-LTLCLQPHMF-H--RNREKYTYDRNPKS
Hs      52 ESGNVHRIYQKKLGWPYDTLMRFDMFLKEKERF-KELDYLFFVNANMIFIKDIHQNEVLASDKE---LIMILQPWFI-G--LDIDKTTYDRNPLS
Bo2     48 GRSNVRIYHKECAGFPADSLFRFEMFLQAEQEI-MKYDYVYFFNANAMCIEPIGN-EILPDESG---LSMGIWGGKRLH--QHPMFYPYERNRKS
Bcac    48 DEANVHLIKKECAGFPADSLFRFDMFLQVRQEL-EKTDYIYFFNSNAEFKAPVGK-ELLPLNGEK--LVGAEWPGKR-KPFKHPAFYPYERNKHS
Fp      54 EKENVLRIYQENMEWPYITLFRFKIFKKALPYL-YNSDYIFFFNSNMKFISDVNE-KILPQGSKQ--LTFVMHPGYY-N--KPVKKFTYEKNNNS
Hm      91 EASDVRKIYQENLGWPFNTLKRFEMFLGQEEAL-REFDFVFFFNANCLFFQHIGD-EFLPIEED---ILVTQHYGFR-D--ASPECFTYERNPKS
Sv1     51 ANPRIHRCPQEAYSWPFATLRRFEIFLSREEEL-KAFDYIFFFNANAQIMTTITPEMFLPRADRGEHLLVVQHPSFY-T--KPNYEFTYDRNPRC
Bo1     52 NNKHIHRIKQKNLGWPDNTLKRFHIFLRIKEQLERETDYLFFFNANLLFTSPIGKEILPPSDSNG--LLGTMHPGFY-N--KPNSEFTYERRDAS

Sv2    145 RAGI-PYNCGTY-YVAGGINGGASGAFLEMCRELKARTDEDLQRGIIARCHDESQLNRLVAECPERFRILPPEYCTPEET---PTGKEAIRVLQK
WcmB   135 TCYL-GYLKKGI-YYQGCFNGGKTASFKRLIQICNMMTMADLKKNLIAKVHDESYLNYYYYYNKPL--LLSELYSWPEKYGENK--DAKIIMRDK
Bs     135 QAYI-PYNSGEH-YYQGGLIGGESKAYIELLEQCSLMTETDLKRNITARWHDESYLNKYLLDKQIK--ILSTEYGRPQEWTVPP--TPKIIFRDK
Psych  162 TAYI-AYGEGKY-YFTGALNGGKSAAFLDLCHTLYNNTQSDLKQDIIALWHDESHLNKFALGREDI--KILPPY-FTRGEREYWKKTSKLMFSDK
Hs     140 TSYI-PYGEGEH-YVMGSFNGGKTSAFLEMCQQLNDNIYLDLKKDIIALWHDESHLNHYCWRNKEKIFILPDSYLLDEGQQYFP--------SKK
Bo2    136 WAYVAPYGKDYT-YFMGGLNGGRPKEYLEMVRTLSVNIRDDYDRGIIALVHDESHINAYMRSHPCK--ILPVELNRPEEW---ADEHTKLIFREK
Bcac   138 SAYIPPREDKPYIYYMGGINGGEAADYLEMAETLSENIHADYEKGIVALVHDESHINKYFRTHDCK--VLSAEYCFPEEW-ITCDFSPKIIFRDK
Fp     142 LAFV-PKSLGKY-YFMGGVNGGTTKAYIKLINDLEYAIDQDLKKNVIAIWHDESHLNKYAITHEKLIHVLSPQYGYPEGL-SNPDFEIKLVIKNK
Hm     178 LAYV-PFGKGKA-YVYGSTNGGKAGAFLALARTLQERIQEDLSRGIIAIWHDESHLNAYIIDHPNY-KMLDYGYGFPEGYGRVPGGGVYIFLRDK
Sv1    142 RAFI-PMGLGRY-YVCGGINGGEAAAFLKLCHTLDKRIRKDLAHNVIAQWHDESHINRYILWRRDV-RVLSPSYCWPEGWNLPL--PCRILIRSK
Bo1    142 TAYI-PEGEGRY-YYAGGLSGGCTKAYLKLCTTICSWVDRDATNHIIPIWHDESLINKYFLDNPPA-ITLSPAYLYPEGW-LLP-FEPIILIRDK

Sv2    235 SHYI--DMSAVRQQGRRQNYL----QRKWEAFCLNWLPYLWWARDTLLRRRVDPPRTR
WcmB   224 ER-------------------------------ESWYGNIKK----------------
Bs     224 NTIL--GASYICS-------------LKKRNRLKLISKAIRNILNKLLKR--------
Psych  252 SHYRFGGHAYLRSETDEKITQAEWNKKNAKRRRKLKFRAKQYISSLLFRQ--------
Hs     225 SHYI--NPLYLKN-------H----DVRLVMRNKNHYRFGGELHERDY----------
Bo2    225 TH-I--DPYFNKG---RKTSL----CARFKKGTTILFNAIRWYVKL------------
Bcac   230 VK-L--DPYFNKG---RDYSL----KGKLKKTVSVIYRAIRWYL--------------
Fp     234 ENYI--DIHTIKG-------------SKRRNVFLNIYSYIKGLLSENN----------
Hm     270 SRVI--DVNAIKG-------M----GSPANRRLKNALRKLKHFSKRLLGR--------
Sv1    232 ARYF--DVQQLRK-DAPATEL----PRYVVRCNDFMKRAARWLQRRLPPKKEDINDE-
Bo1    232 NK-----PQYGGH--------------ELLRRKNSLWERIKLICQKFKSAD-------
```

Figure 1

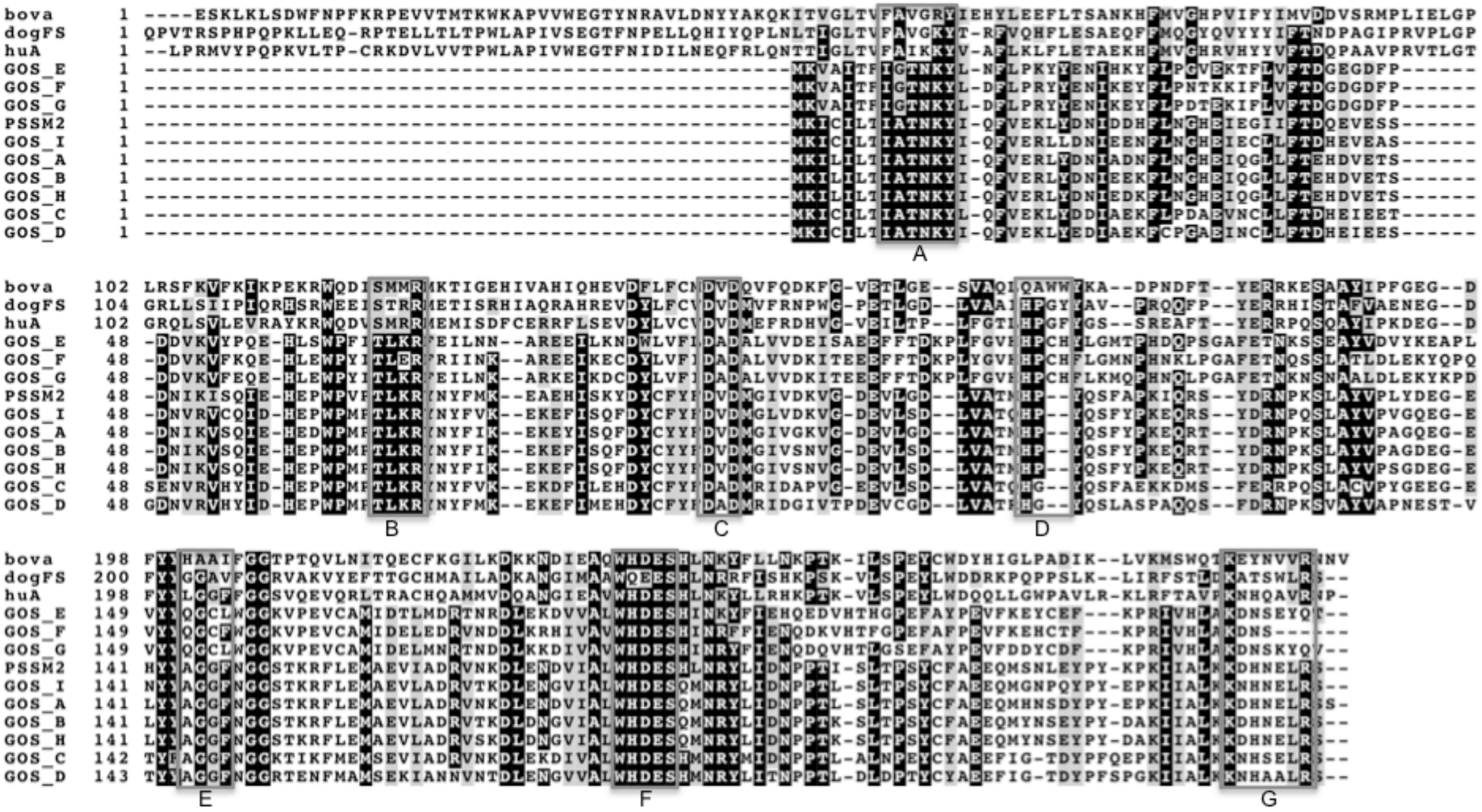

bova
dogFS
huA
GOS_E
GOS_F
GOS_G
PSSM2
GOS_I
GOS_A
GOS_B
GOS_H
GOS_C
GOS_D
A
B
C
D
E
F
G

Figure 2

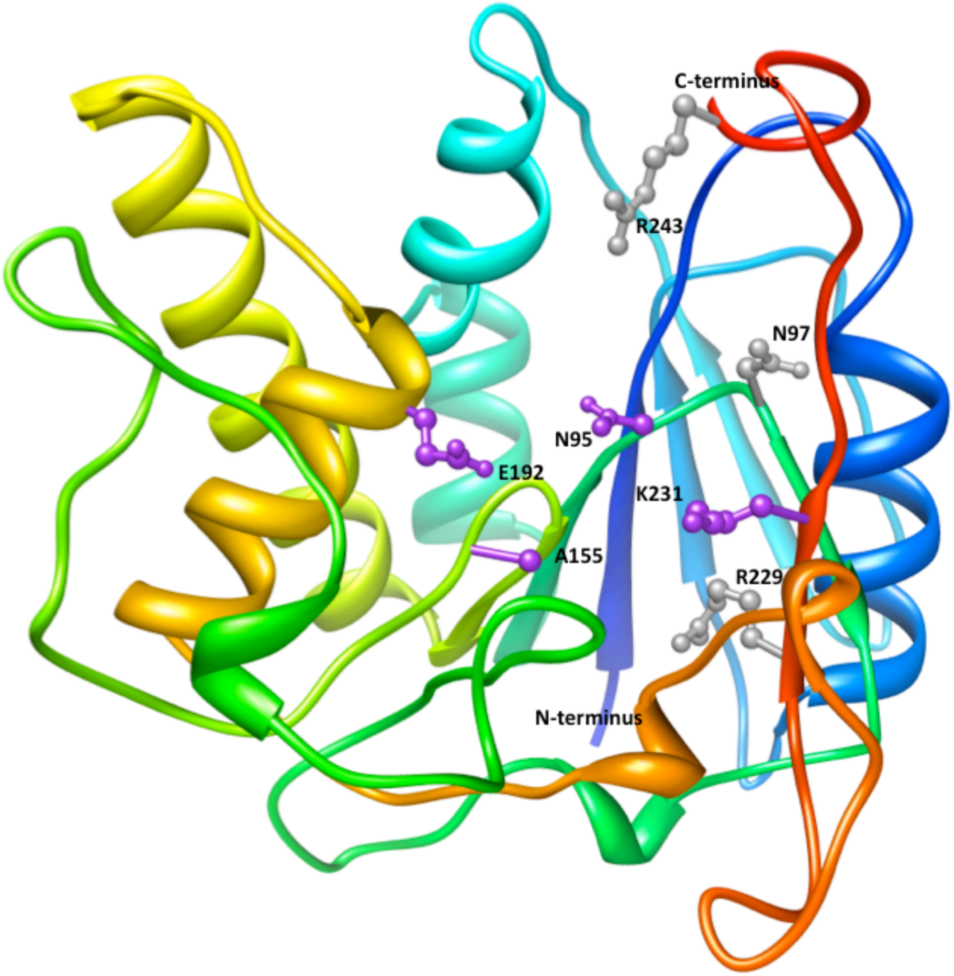


Figure 3