

PNAS

www.pnas.org

Supplementary Information for

Glycoconjugate pathway connections revealed by sequence similarity network analysis of the monotopic phosphoglycosyl transferases

Katherine H. O'Toole¹, Barbara Imperiali^{2,3*}, and Karen N. Allen^{1*}

*Corresponding authors: Karen N. Allen, drkallen@bu.edu; Barbara Imperiali, imper@mit.edu

This PDF file includes:

Supplementary text
Figures S1 to S4
Tables S1 to S4
SI References

Other supplementary materials for this manuscript include the following:

Datasets S1 to S4

InterPro or Pfam families included in outlier classification: Radical SAM (IPR007197), Phosphoglucose isomerase (IPR001672), Methyltransferase FkbM (IPR006342), Cytidyltransferase-like domain (IPR004821), HpcH/Hpal aldolase/citrate lyase domain (IPR005000), VanZ-like (IPR006976), Surfeit locus 1/Shy1 (IPR002994), Aminoacyl-tRNA synthetase, class Ic (IPR002305), Domain of unknown function DUF4422 (IPR025536), Exosortase (IPR019127), Polysaccharide biosynthesis protein (PF13440), Domain of unknown function DUF4062 (IPR025139), Uncharacterised protein family YfhO (IPR018580), ATP-grasp in the biosynthetic pathway with Ter operon (PF15632), Inositolphosphotransferase Aur1/Ipt1 (IPR026841), O-antigen ligase-related (IPR007016), Methyltransferase 23 (PF13489), Glycoside hydrolase family 10 domain (IPR001000), Protein of unknown function DUF4012 (IPR025101), Methyltransferase type 11 (IPR013216), PEP-CTERM protein-sorting domain (IPR013424), Molybdopterin oxidoreductase (IPR006656), Acidic fibroblast growth factor intracellular-binding protein (IPR008614), Glycosyl transferase, family 4 (IPR000715), and Haloacid dehalogenase-like hydrolase (IPR041492).

Table S1. Sugar-modifying domain families

InterPro Family	Highest ID (%) predicted homolog	Predicted function	Location of domain
CapD-like domain (IPR003869)	CapD (37-41%)	C2',4' dehydratase	C-term
Hexapeptide repeat (IPR001451)	EpsM (38-39%) & PglD (38-40%)	acetyltransferase	C-term
NAD-dependent epimerase/dehydratase (IPR001509)	Inconclusive	epimerase/ dehydratase	both observed
PglD, N-terminal (IPR041561)	EpsM (38-39%) & PglD (39-40%)	acetyltransferase	C-term
GNAT domain (IPR000182)	<i>E. coli</i> RimI	acetyltransferase	both observed
NAD(P)-binding domain (IPR016040)	CapD, PglF, PseB (15-20%)	dehydratase	N-term
DegT/DnrJ/EryC1/StrS aminotransferase (IPR000653)	WbpE (30-37% ID)	aminotransferase	both observed
Acyltransferase 3 (IPR002656)	WechH (21-22%)	acyltransferase	C-term
RmlD-like substrate binding domain (IPR029903)	<i>S. enterica</i> RmlD (19%)	reductase	C-term
UDP-glucose/GDP-mannose dehydrogenase, C-terminal (IPR014027)	<i>B. subtilis</i> TuaD (25%)	dehydrogenase	both observed
Nucleotidyl transferase (IPR005835)	GlmU (15-20%)	acetyltransferase & uridylyltransferase	C-term

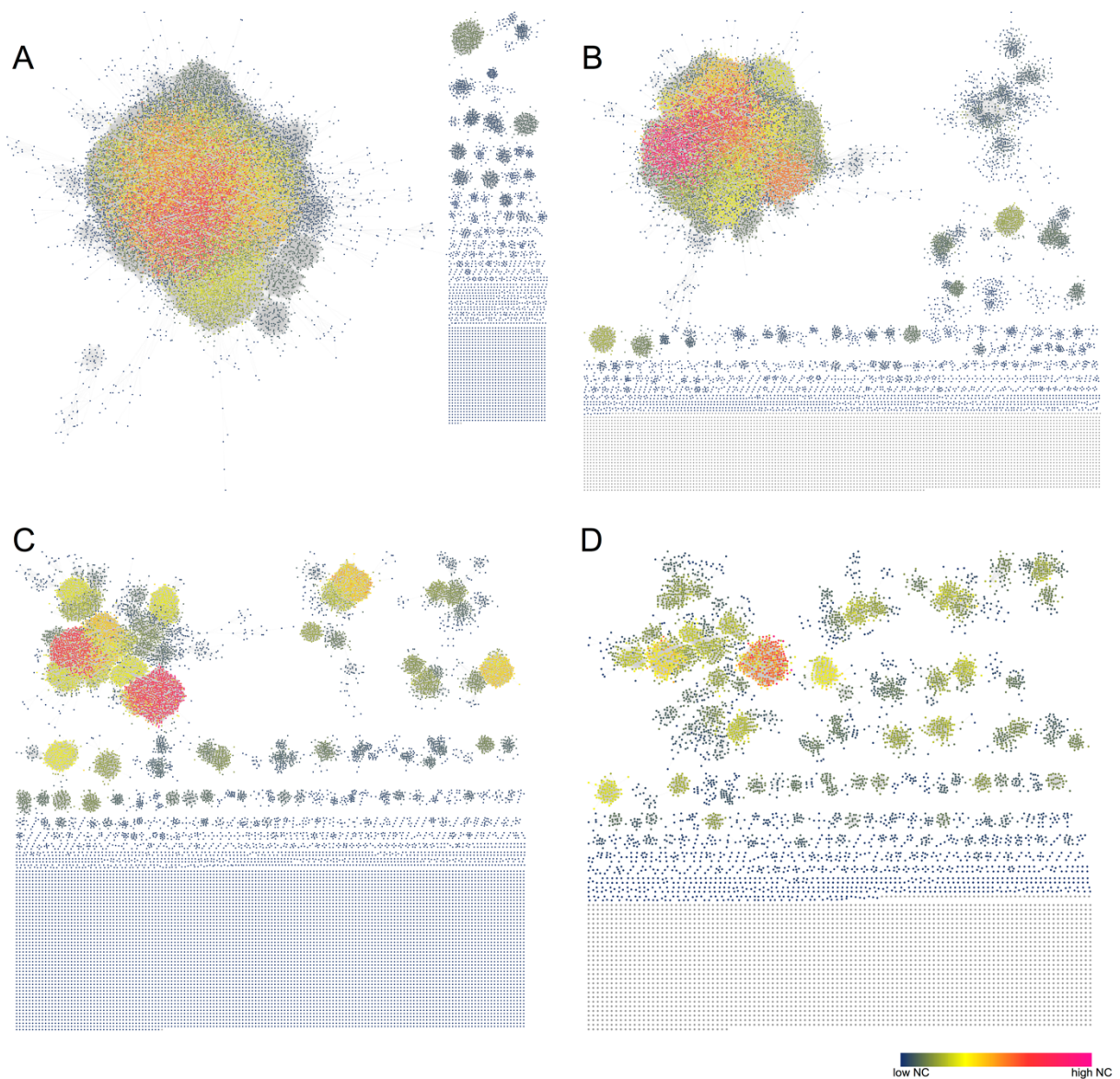


Figure S1. Sequence similarity network of the monoPGT superfamily varying E -value cutoff and % identity, colored by neighborhood connectivity (NC) from blue (low NC) to pink (high NC). A. 70% ID representative node network, E -value cutoff of 1×10^{-60} with NC scale from 1-1657, B 70% ID representative node network, E -value cutoff of 1×10^{-75} with NC scale from 1-746, C. 70% ID representative node network, E -value cutoff of 1×10^{-90} with NC scale from 1-505, D. 40% ID representative node network, E -value cutoff of 1×10^{-90} with NC scale from 1-189.

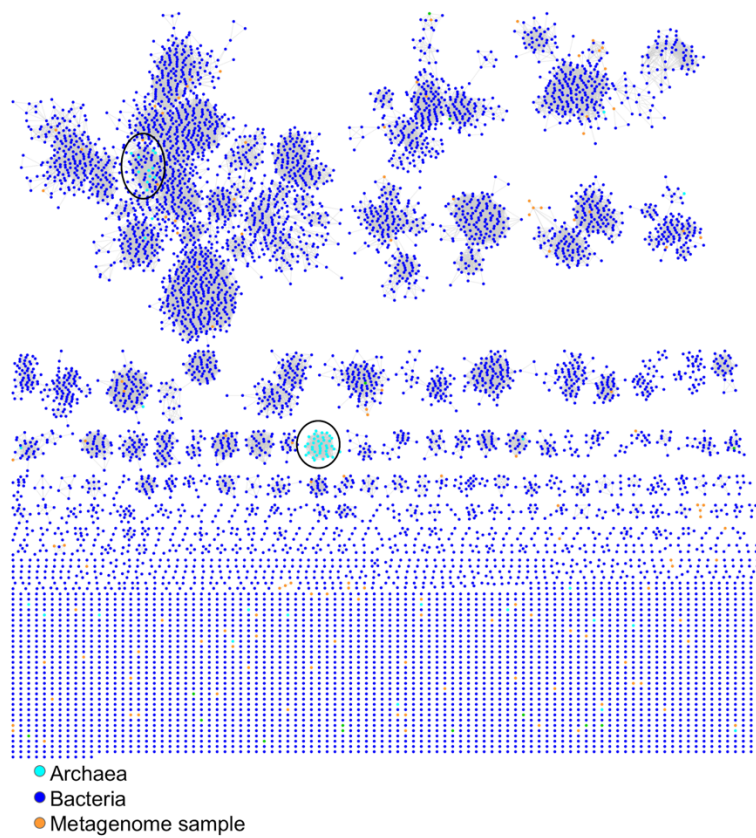
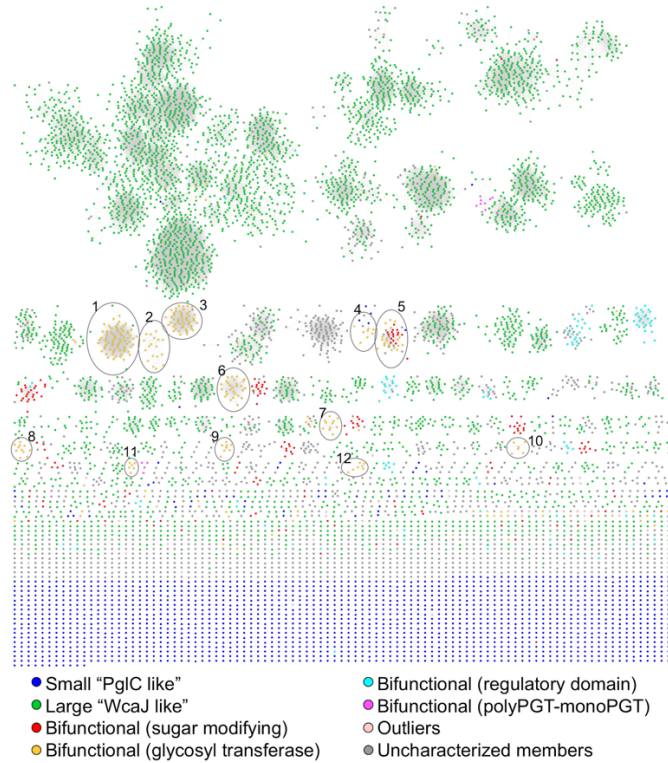


Figure S2. Sequence similarity network of the monoPGT superfamily colored by domain of life. The circled regions indicate locations of archaeal sequences. The network is a 40% representative node network with edges representing an E -value cutoff of 1×10^{-90} .



GT cluster	InterPro or Pfam family	Highest ID (%) predicted homolog	Location of domain
1	GT-2 like (IPR001173) or Galactosyltransferase (IPR027791)	WbbL (23-38%)	N-term
2	GT WecB/TagA/CpsF (IPR004629)	TagA (28-30%)	C-term
3	GT WecB/TagA/CpsF (IPR004629)	TagA (25-31%)	N-term
4	GT-2 like (IPR001173)	PglI (11-25%)	C-term
5	GT-2 like (IPR001173) or GT, family 1 (IPR001296)	PglA (22-28%)	N-term
6	GT like family 2 (PF13641)	GfT2 (15-20%)	N-term
7	GT-2 like (IPR001173)	PglI (18-23%)	N-term
8	GT, family 1 (IPR001296) or GT1 subfamily 4 (IPR028098)	inconclusive	N-term
9	GT1 subfamily 4 (IPR028098)	inconclusive	N-term
10	GT, family 1 (IPR001296) or GT1 subfamily 4 (IPR028098)	PglA (24-32%)	N-term
11	GT WecB/TagA/CpsF (IPR004629)	TagA (28-41%)	C-term
12	GT, family 1 (IPR001296) or GT1 subfamily 4 (IPR028098)	inconclusive	N-term

Figure S3. Sequence similarity network of the monoPGT superfamily colored by structural family (40% ID representative node network, E -value cutoff of 1×10^{-90}). Glycosyl transferase bifunctional enzyme clusters are circled, numbered, and annotated in the table based on highest predicted SwissProt verified homolog.

Table S2. Regulatory domain families

InterPro Family	Protein with highest sequence identity	Predicted function	Location of domain
Signal transduction response regulator, receiver domain (IPR001789)	<i>E. coli</i> CpxR (~30-40%) & <i>P. aeruginosa</i> C-di-GMP phosphodiesterase (24-37%)	signal receiver, regulatory	both observed
Sulphate transporter and anti-sigma factor antagonist domain (IPR002645)	<i>B. subtilis</i> SpoIIAA (20-30%)	anti-anti sigma factor, regulatory	N-term
AAA domain (IPR025669)	<i>S. aureus</i> CapB (39%)	tyrosine kinase, regulatory	N-term

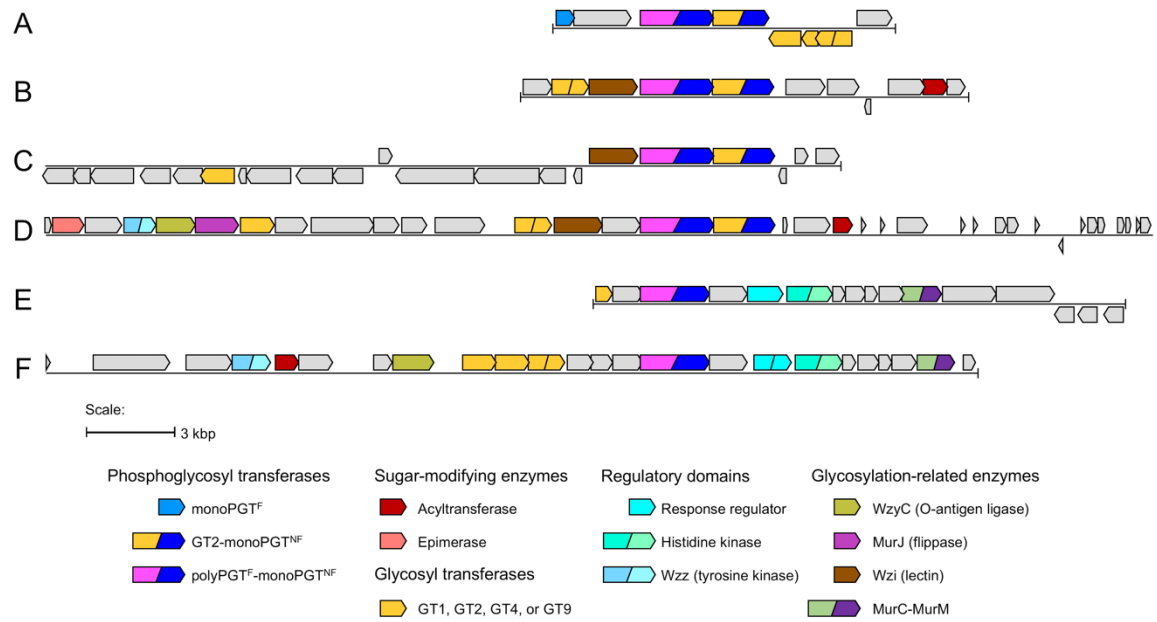


Figure S4. Genome neighborhood diagram of select polyPGT^F-monoPGT^{NF} genes. Representation of the genome neighborhood network. UniProt IDs for the polyPGT^F-monoPGT^{NF} genes selected: A0A1V4RUK6 (A), A0A3M2FN56 (B), A0A3N9MXS9 (C), H1XS79 (D), A0A1F4X332 (E), A0A1F4XPX4 (F).

Table S3. UniProt IDs included in phylogenetic reconstruction. UniProt IDs are listed starting with the starred branch and following in a clockwise manner.

UniProt ID	Organism	UniProt Protein Name*	PGT Subclass
A0A1G0P652	<i>Ignavibacteria bacterium</i>	Bac_transf domain-containing protein	large
A0A1F7QQD9	<i>Candidatus Saccharibacteria bacterium</i>	UDP-phosphate galactose phosphotransferase	large
A0A167G8C9	<i>Ulvibacter sp. LPB0005</i>	Phosphosugar transferase	small
A0A2E0W3V7	<i>Ignavibacteriae bacterium</i>	Glycosyl transferase	Bifunctional/ glycosyl transferase
A0A1I2MZI2	<i>Pontibacter chinhatensis</i>	Undecaprenyl-phosphate glucose phosphotransferase	large
A0A257J4M7	<i>Alphaproteobacteria bacterium PA2</i>	Bac_transf domain-containing protein	large
A0A063XYI9	<i>Nitricola laciaponensis</i>	Undecaprenyl-phosphate galactosephosphotransferase	large
P71241	<i>Escherichia coli (strain K12)</i>	UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferase	large
A0A1Z8TXU5	<i>Verrucomicrobia bacterium TMED56</i>	Bac_transf domain-containing protein	large
A0A2H3K9V0	<i>Flavobacterium branchiophilum</i>	Phosphosugar transferase	uncharacterized
A0A0A6PLW0	<i>Candidatus Thiomargarita nelsonii</i>	Uncharacterized protein	Bifunctional/ sugar-modifying
A0A1F4YGZ9	<i>candidate division Zixibacteria bacterium</i>	Bac_transf domain-containing protein	Bifunctional/ polyPGT-monoPGT
A0A078M711	<i>Jeotgalicoccus saudimassiliensis</i>	Putative phosphosugar transferase EpsL	small
A0A1V6I822	<i>Candidatus Marinimicrobia bacterium ADurb.Bin030</i>	Putative undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase	Bifunctional/ polyPGT-monoPGT
A0A0D1F875	<i>Vibrio parahaemolyticus 901128</i>	Uncharacterized protein	small
A0A1I2K568	<i>Clostridium cadaveris</i>	Exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	large
A0A1I0S5X7	<i>Chitinophaga arvensicola</i>	Phosphosugar transferase involved in LPS biosynthesis (Colanic, teichoic acid)	uncharacterized
A0A1E4A3B2	<i>Cytophagaceae bacterium SCN 52-12</i>	Phosphosugar transferase	Bifunctional/ regulatory
A0A1M3LES6	<i>Micrococcales bacterium 73-13</i>	Glutamate synthase subunit alpha	Bifunctional/ uncharacterized
A0A1G8UVZ4	<i>Pedobacter sp. ok626</i>	Phosphosugar transferase involved in LPS biosynthesis (Colanic, teichoic acid)	regulatory

A0A2A5DCU0	<i>Planctomycetes bacterium</i>	Bac_transf domain-containing protein	large
A0A2G6GHL8	<i>Chloroflexi bacterium</i>	Uncharacterized protein	Bifunctional/ glycosyl transferase
A0A2H6FVN3	<i>bacterium BMS3Abin10</i>	UDP-N-acetyl-alpha-D-glucosamine C6 dehydratase	Bifunctional/ regulatory
A0A0P0NFR9	<i>Pedobacter sp. PACM 27299</i>	Glycosyl transferase	small
A0A0A2DUD0	<i>Porphyromonas sp. COT-239 OH1446</i>	Glycosyl transferase	small
A0A077EHY8	<i>Elizabethkingia anophelis NUHP1</i>	Lipid carrier: UDP-N-acetylgalactosaminyltransferase	small
A0A0E9MUH9	<i>Flaviumibacter petaseus NBRC 106054</i>	Putative glycosyltransferase	large
A0A0B2JXH2	<i>Anaerovibrio lipolyticus</i>	Bac_transf domain-containing protein	small
A0A345Y3H3	<i>Crenobacter cavernae</i>	NAD-dependent epimerase/dehydratase family protein	Bifunctional/ sugar-modifying
A0A0X8X646	<i>Mucilaginibacter gotjawali</i>	UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferase	large
A0A0L6JVG9	<i>Pseudobacteroides cellulosolvens</i>	Undecaprenyl-phosphate galactose phosphotransferase	small
P95706	<i>Staphylococcus aureus</i>	Bacterial phosphosugar transferase (CapM)	small
A0A1G7HH12	<i>Halorientalis regularis</i>	Phosphosugar transferase involved in LPS biosynthesis (Colanic, teichoic acid)	large
A0A0A7PG93	<i>Sphingopyxis fribergensis</i>	Phosphosugar transferase	small
A0A2E7Q4P5	<i>Rhodobacteraceae bacterium</i>	Phosphosugar transferase	Bifunctional/ sugar-modifying
A0A1I3C877	<i>Halomonas xianhensis</i>	Exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	large
A0A0K6GZ10	<i>Idiomarina woesei</i>	Exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	large
A0A0U2WZ63	<i>Enterococcus rotai</i>	UDP-phosphate galactose phosphotransferase	large
A0A0G1ASD6	<i>Parcubacteria group bacterium</i>	Phosphosugar transferase	large
A0A0C2URJ0	<i>Sulfurovum sp. AS07-7</i>	Phosphosugar transferase	Bifunctional/ uncharacterized
A0A1H1TW74	<i>Opitutus sp. GAS368</i>	Exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	large
A0A0Q5UNY3	<i>Sphingomonas sp. Leaf198</i>	Phosphosugar transferase	large
A0A0A6XW68	<i>Bacillus ginsengihumi</i>	UDP-phosphate N-acetylgalactosaminyl-1-phosphate transferase	small

A0A0A2GWW3	<i>Dokdonia donghaensis</i> DSW-1	Phosphosugar transferase	large
A0A1Z4RGQ1	<i>Calothrix</i> sp. NIES-4101	Phosphosugar transferase	large
A0A081NKA5	<i>Endozoicomonas numazuensis</i>	Phosphosugar transferase	small
A0A0S2K1K2	<i>Pseudoalteromonas phenolica</i>	Intercellular adhesion protein A	Bifunctional/ glycosyl transferase
A0A1F2RMI6	<i>Acidobacteria bacterium</i>	Bac_transf domain-containing protein	regulatory
A0A3A4R061	<i>Anaerolineaceae bacterium</i>	Bac_transf domain-containing protein	large
A0A081GQD0	<i>Cyanobium</i> sp. CACIAM 14	Anti-sigma-factor antagonist (STAS) and Phosphosugar transferase	Bifunctional/ regulatory
A0A2E2VCH6	<i>Euryarchaeota archaeon</i>	Bac_transf domain-containing protein	large
A0A1H9DX23	<i>Streptomyces radiopugnans</i>	Exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	large
Q608Q3	<i>Methylococcus capsulatus</i>	Bacterial Phosphosugar transferase/glycosyl transferase, WecB/TagA/CpsF family protein	Bifunctional/ glycosyl transferase
A0A099W939	<i>Listeriaceae bacterium</i> FSL A5-0209	Bac_transf domain-containing protein	small
A0A087C7E0	<i>Bifidobacterium mongoliense</i> DSM 21395	Exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	large
A0A0F2QCV2	<i>Clostridiaceae bacterium</i> BRH_c20a	Multidrug MFS transporter	small
A0A1E5AIQ7	<i>Rhodobacteraceae bacterium</i>	UDP-phosphate galactose phosphotransferase	Bifunctional/ glycosyl transferase
A0A0S3U064	<i>Leptolyngbya</i> sp. NIES-3755	Phosphosugar transferase	large
A0A0C2QYQ1	<i>Jeotgalibacillus campisalis</i>	Multidrug MFS transporter	small
A0A060M6D2	<i>Bacillus lehensis</i> G1	Galactosyl transferase CpsE	small
A0A0K9GF90	<i>Bacillus</i> sp. FJAT-27916	Multidrug MFS transporter	small
A0A075LME8	<i>Terribacillus goriensis</i>	Multidrug MFS transporter	small
A0A0A3HUK8	<i>Lysinibacillus sinduriensis</i> BLB-1 = JCM 15800	Multidrug MFS transporter	small
A0A0D0END6	<i>Bacillus thermoamylovorans</i>	Undecaprenyl-phosphate galactosephosphotransferase	small
A0A098EJU0	<i>Planococcus massiliensis</i>	Putative phosphosugar transferase EpsL	small
A0A2D5YSD6	<i>Planctomycetaceae bacterium</i>	Bac_transf domain-containing protein	uncharacterized

A0A1W1HNW2	<i>Nitrospira sp. ND1</i>	Putative Mannose-1-phosphate guanylyltransferase	Bifunctional/ sugar-modifying
A0A0P1IMQ4	<i>Thalassobius activus</i>	Putative phosphosugar transferase EpsL	small
A0A0C1GQU2	<i>Ruegeria sp. ANG-R</i>	phosphosugar transferase	small
A0A0J5T177	<i>Puniceibacterium sp. IMCC21224</i>	Glycosyl transferase	small
A0A0H4KXY5	<i>Marinovum algicola DG 898</i>	Phosphosugar transferase involved in lipopolysaccharide synthesis	small
A0A1E4EEC8	<i>Methylobacterium sp. SCN 67-24</i>	Bac_transf domain-containing protein	large
A0A0C2UHH6	<i>Bacillus badius</i>	Protein-export membrane protein SecG	small
A0A031I0P7	<i>Sphingomonas sp. RIT328</i>	Undecaprenyl-phosphate galactosephosphotransferase	small
A0A011UU28	<i>Aquamicrobium defluvii</i>	Exopolysaccharide production protein ExoY (phosphosugar transferase)	small
A0A1U7J8H2	<i>Phormidium tenue NIES-30</i>	UDP-phosphate galactose phosphotransferase	large
A0A0F9AFY9	<i>marine sediment metagenome</i>	Bac_transf domain-containing protein	small
A0A017H830	<i>Limimanicola hongkongensis DSM 17492</i>	Undecaprenyl-phosphate galactosephosphotransferase	small
A0A1B0ZPV5	<i>Phaeobacter gallaeciensis</i>	Exopolysaccharide biosynthesis protein	small
P26406	<i>Salmonella enterica</i>	Undecaprenyl-phosphate galactose phosphotransferase (WbaP)	large
A0A011Q8T3	<i>Alkalibacterium sp. AK22</i>	Lipid carrier: UDP-N-acetylgalactosaminyltransferase	small
A0A1F9UCC6	<i>Elusimicrobia bacterium</i>	Bac_transf domain-containing protein	large
A0A098BX95	<i>Fermentimonas caenicola</i>	Bac_transf domain-containing protein	small
A0A372SQY9	<i>Absiella sp. AM09-45</i>	Exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	large
A0A073IUV1	<i>Synergistes jonesii</i>	Bac_transf domain-containing protein	large
A0A0F6AJN0	<i>Thermotoga sp. (strain RQ2)</i>	Exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	large
A0A1G7EX33	<i>Sphingomonas carotiniaciens</i>	Phosphosugar transferase involved in LPS biosynthesis (Colanic, teichoic acid)	large
A0A146AWZ9	<i>Bordetella ansorpii</i>	Exopolysaccharide biosynthesis glycosyl transferase	small
A0A0N1F5M8	<i>Bosea vaviloviae</i>	Bac_transf domain-containing protein	large
A0A099WFB8	<i>Listeria booriae</i>	UDP-galactose phosphate transferase	small

A0A365P052	<i>Flavobacterium tibetense</i>	Uncharacterized protein	Bifunctional/ sugar-modifying
A0A085L597	<i>Schleiferia thermophila str. Yellowstone</i>	Phosphosugar transferase	small
A0A090VNM3	<i>Algibacter lectus</i>	UDP-N-acetylgalactosaminyltransferase	small
A0A101AVP0	<i>Mycobacterium sp. GA-2829</i>	UDP-phosphate galactose phosphotransferase	large
A0A077M629	<i>Tetrasphaera jenkinsii Ben 74</i>	Putative phosphotransferase involved in extracellular matrix synthesis	small
A0A0Q9KUR6	<i>Knoellia sp. Soil729</i>	UDP-galactose phosphate transferase	small
A0A0U4CG29	<i>Hymenobacter sedentarius</i>	Bac_transf domain-containing protein	small
J6UDJ7	<i>Rhodovulum sp. PH10</i>	Lipid carrier: UDP-N-acetylgalactosaminyltransferase	large
A0A0P6XNN7	<i>Thermanaerotherix daxensis</i>	Bac_transf domain-containing protein	small
A0A158M4S6	<i>Bordetella holmesii CDC-H585-BH</i>	Phosphosugar transferase	small
A0A1A9DHV8	<i>Streptomyces sp. OspMP-M45</i>	Exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	large
A0A099DKB0	<i>Exiguobacterium mexicanum</i>	Phosphosugar transferase	small
A0A031I293	<i>Sphingomonas sp. RIT328</i>	Phosphosugar transferase	small
A0A1C5T7N5	<i>uncultured Blautia sp.</i>	Putative colanic biosynthesis UDP-glucose lipid carrier transferase	small
A0A317RJM9	<i>Eubacterium limosum</i>	Lipopolysaccharide/colanic/teichoic acid biosynthesis glycosyltransferase	Bifunctional/ glycosyl transferase
A0A0A3IMY7	<i>Lysinibacillus manganicus DSM 26584</i>	Phosphosugar transferase	small
A0A0P0DN04	<i>Sphingopyxis macrogoltabida</i>	Bac_transf domain-containing protein	small
A0A0C2YWP4	<i>Sulfurovum sp. AS07-7</i>	Bac_transf domain-containing protein	small
A7ZET4	<i>Campylobacter concisus</i>	N,N'-diacetylbaucillosaminyl-1-phosphate transferase	small
Q0P9D0	<i>Campylobacter jejuni</i>	Undecaprenyl phosphate N,N'-diacetylbaucillosamine 1-phosphate transferase	small
A0A0K1Q5Q4	<i>Labilithrix luteola</i>	Lipid carrier: UDP-N-acetylgalactosaminyltransferase	small
A0A3T1D1W1	<i>Cohnella abietis</i>	Uncharacterized protein	Bifunctional/ glycosyl transferase
A0A081BXT8	<i>Candidatus Vecturithrix granuli</i>	Undecaprenyl-phosphate galactose phosphotransferase	small
D6HAT8	<i>Neisseria gonorrhoeae DGI2</i>	Pilin glycosylation protein PglB	Bifunctional/ sugar-modifying

A0A081NCB0	<i>Endozoicomonas montiporae</i>	Phosphosugar transferase	small
A0A081N3Q4	<i>Endozoicomonas numazuensis</i>	Phosphosugar transferase	small

*UniProt protein name for monoPGT core domain corrected from sugar transferase to phosphosugar transferase

Table S4. Characterized members of the monoPGT superfamily

UniProt ID (reference)	Representative ID*	Protein	PGT Family	Organism	CHO	Glycoconjugate	Structure
A7ZET4(1)	A0A1X9SUU7	PglC	small	<i>Campylobacter concisus</i>	diNAcBac	N-glycoprotein	5W7L
P95706(2)	A0A133Q0D2	CapM	small	<i>Staphylococcus aureus</i>	FucNAc	Type 5 polysaccharide	N/A
P71241(3)	A0A1B7HPD9	WcaJ	large	<i>Escherichia coli (K12)</i>	Glc	Colanic acid	N/A
Q0P9D0(4, 5)	A0A099WFB8	PglC	small	<i>Campylobacter jejuni</i>	diNAcBac	N-glycoprotein	N/A
Q56770(6)	Q5GXZ0	GumD	large	<i>Xanthomonas campestris</i>	Glc	xanthan	N/A
B3FN88(7)	A0A3N6XNY4	WecP	large	<i>Aeromonas hydrophila</i>	GalNAc	O-antigen/LPS	N/A
Q9ABR0(3)	N/A	PssY	small	<i>Caulobacter vibriodes</i>	Glc	colanic acid	N/A
P26406(8, 9)	A0A379W6Q6	WbaP	large	<i>Salmonella enterica</i>	Gal	O-antigen/LPS	N/A
D6HAT8(10)	B4RQ09	PglB	bifunctional/ sugar-modifying	<i>Neisseria gonorrhoeae</i>	diNAcBac	O-glycoprotein	N/A

* Representative ID for 70% ID CD-HIT clustering

SI References

1. Das D, Kuzmic P, & Imperiali B, Analysis of a dual domain phosphoglycosyl transferase reveals a ping-pong mechanism with a covalent enzyme intermediate. *Proc Natl Acad Sci U S A* 114, 7019-7024 (2017).
2. Rausch M, *et al.*, Coordination of capsule assembly and cell wall biosynthesis in *Staphylococcus aureus*. *Nat Commun* 10, 1404 (2019).
3. Patel KB, *et al.*, Functional characterization of UDP-glucose:undecaprenyl-phosphate glucose-1-phosphate transferases of *Escherichia coli* and *Caulobacter crescentus*. *J Bacteriol* 194, 2646-2657 (2012).
4. Glover KJ, Weerapana E, Chen MM, & Imperiali B, Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry* 45, 5343-5350 (2006).
5. Lukose V, *et al.*, Conservation and Covariance in Small Bacterial Phosphoglycosyltransferases Identify the Functional Catalytic Core. *Biochemistry* 54, 7326-7334 (2015).
6. Katzen F, *et al.*, *Xanthomonas campestris* pv. *campestris* gum mutants: effects on xanthan biosynthesis and plant virulence. *J Bacteriol* 180, 1607-1617 (1998).
7. Merino S, *et al.*, A UDP-HexNAc:polyprenol-P GalNAc-1-P transferase (WecP) representing a new subgroup of the enzyme family. *J Bacteriol* 193, 1943-1952 (2011).
8. Patel KB, Ciepichal E, Swiezewska E, & Valvano MA, The C-terminal domain of the *Salmonella enterica* WbaP (UDP-galactose:Und-P galactose-1-phosphate transferase) is sufficient for catalytic activity and specificity for undecaprenyl monophosphate. *Glycobiology* 22, 116-122 (2012).
9. Saldías MS, *et al.*, Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. *Microbiology* 154, 440-453 (2008).
10. Hartley MD, *et al.*, Biochemical characterization of the O-linked glycosylation pathway in *Neisseria gonorrhoeae* responsible for biosynthesis of protein glycans containing N,N'-diacetylbacillosamine. *Biochemistry* 50, 4936-4948 (2011).