Inositol lipid synthesis is widespread in host-associated Bacteroidetes

S. L. Heaver [1], H. H. Le [2], P. Tang [3], A. Baslé [4], J. Marles-Wright [5], E. L. Johnson[2], D. J. Campopiano [3], R. E. Ley [1]

[1] Department of Microbiome Science, Max Planck Institute for Developmental Biology, Tübingen 72076, Germany

[2] Division of Nutritional Sciences, Cornell University, Ithaca, NY 14853, USA

[3] School of Chemistry, University of Edinburgh, Edinburgh, Scotland, UK

[4] Newcastle University Biosciences Institute, Newcastle University, UK

[5] School of Natural and Environmental Sciences, Newcastle University, UK

Abstract

Ubiquitous in eukaryotes, inositol lipids have finely tuned roles in cellular signaling and membrane homeostasis. In Bacteria, however, inositol lipid production is rare. Recently, the prominent human gut bacterium *Bacteroides thetaiotaomicron* (BT) was reported to produce inositol lipids, including inositol sphingolipids, but the pathways remain ambiguous and their prevalence unclear. Here, we investigated the gene cluster responsible for inositol lipid synthesis in BT using a novel strain with inducible control of sphingolipid synthesis. We characterized the biosynthetic pathway from *myo*-inositol-phosphate (MIP) synthesis to phosphoinositol-dihydroceramide, including structural and kinetic studies of the enzyme MIP synthase (MIPS). We determined the crystal structure of recombinant BT MIPS with bound NAD cofactor at 2.0 Å resolution, and identified the first reported phosphatase for the conversion of bacterially-derived phosphatidylinositol phosphate (PIP) to phosphatidylinositol (PI). Transcriptomic analysis indicated inositol production is nonessential but its loss alters BT capsule expression. Bioinformatic and lipidomic comparisons of Bacteroidetes species revealed a novel second putative pathway for bacterial PI synthesis without a PIP intermediate. Our results indicate that inositol sphingolipid production, via one of the two

pathways, is widespread in host-associated Bacteroidetes, and may be implicated in host interactions

both indirectly via the capsule and directly through inositol lipid provisioning.

1   **Introduction**

2   Inositol, a carbocyclic sugar abundant in eukaryotes, forms the structural basis for diverse

3   phosphorylated secondary messenger inositol phosphates and inositol lipids. At their simplest,

4   inositol lipids have inositol as their polar headgroup, as is the case with phosphatidylinositol (PI; on a

5   glycerophospholipid backbone) or inositol phosphorylceramide (on a sphingolipid backbone; Fig. 1A).

6   The inositol headgroup can be subject to further modifications, including the phosphorylation of PI at

7   multiple sites to form bioactive phosphoinositides, or the addition of a mannose on inositol

8   sphingolipids to form the mannosylinositol phosphorylceramides (MIPCs) abundant in yeast [1,2].

9   Inositol derivatives control key processes of eukaryotic cell physiology. For instance, although

10   phosphoinositides constitute a small fraction of overall phospholipids, they are ubiquitous and

11   essential in roles such as marking organelle identity, regulating cytoskeleton-membrane interactions,

12   and controlling cell division and autophagy [3,4].

13   Despite the widespread distribution of inositol lipids in eukaryotes, relatively little is known

14   about the structure and distribution of inositol lipids in Bacteria. Bacterial inositol lipids are

15   comparatively rare across bacterial species, with production previously thought to be limited largely

16   to phosphatidylinositol (PI) synthesis in members of the phylum Actinobacteria (*e.g.*, *Mycobacteria*,

17   *Corynebacteria*, and *Streptomyces*) [5–7] and Spirochaetes (*Treponema*) [8]. In *Mycobacterium*

18   *tuberculosis,* an obligate intracellular human pathogen, the inositol headgroups of PI are the outer

19   membrane molecules to which surface oligosaccharide virulence factors are linked [9].

20   In eukaryotes, both sphingolipids (SLs, lipids with a sphingosine, or long-chain base

21   backbone) and inositol lipids are involved in the regulation of cell fate and differentiation,

22   inflammation, protein trafficking, and gene regulation in central metabolic pathways, with

23   imbalances linked with the pathologies of a growing inventory of diseases [4,10–12]. These two lipid

24   types intersect in the inositol sphingolipids, such as the glycosylinositol phosphorylceramides

25   abundant in yeast and plants. Inositol SLs are known to be produced by the periodontal pathogen

26   *Tannerella forsythia* [13], a member of the Bacteroidetes. Furthermore, Brown *et al.* recently reported

27   inositol SLs in the common human gut commensal *Bacteroides thetaiotaomicron* [14]. Ceramide

28   phosphoryl-*myo*-inositol has been reported in *Sphingobacterium spiritivorum* [15], a free-living member

29   of the Bacteroidetes, and more recently in *Myxococcus xanthus* [16], of the Proteobacteria phylum. In

30   contrast to the well-studied inositol lipids, including inositol SLs, in plants and fungi, bacterial inositol

31   sphingolipid synthesis has been largely overlooked.

32
33   The discovery of inositol lipids in a several more species suggests that these lipids may be

34   more widespread in bacteria than previously thought and may use novel pathways. Across kingdoms,

35   *de novo* inositol synthesis begins with the formation of inositol phosphate from glucose 6-phosphate

36   (G6P) by a *myo*-inositol phosphate synthase (MIPS, EC 5.5.1.4) [11]. From here, a bacterial pathway for

37   inositol glycerophospholipid synthesis is mostly known (*e.g.*, in *Mycobacteria*), and differs from the

38   eukaryotic pathway by the direct use of inositol-phosphate, not its dephosphorylated inositol form,

39   as a substrate in the formation of PI. This leads first to the synthesis of phosphatidylinositol-

40   phosphate (PIP) from CDP-diacylglycerol (CDP-DAG) and MIP, which is subsequently

41   dephosphorylated to PI [17]. Though the PIP synthase has been well characterized in bacteria [18], the

42   phosphatase responsible for the conversion of bacterial PIP to PI has not yet been identified [18]. In

43   addition, though the gene cluster for bacterial inositol SL synthesis has been predicted in *B.*

44   *thetaiotaomicron* (hereafter BT) [14], the functions of these enzymes remain to be confirmed.

45   Here, we combine genomic and biochemical approaches to functionally characterize the

46   predicted inositol lipid metabolism gene cluster in BT from the initial synthesis of *myo*-inositol-

47   phosphate (MIP) to its addition as a headgroup to glycerophospholipids and SLs. Together with the

48   description of a novel putative alternative gene cluster, common in the *Prevotella,* this work

49   broadens the understanding of how gut bacteria synthesize complex lipids, and reveals an extensive

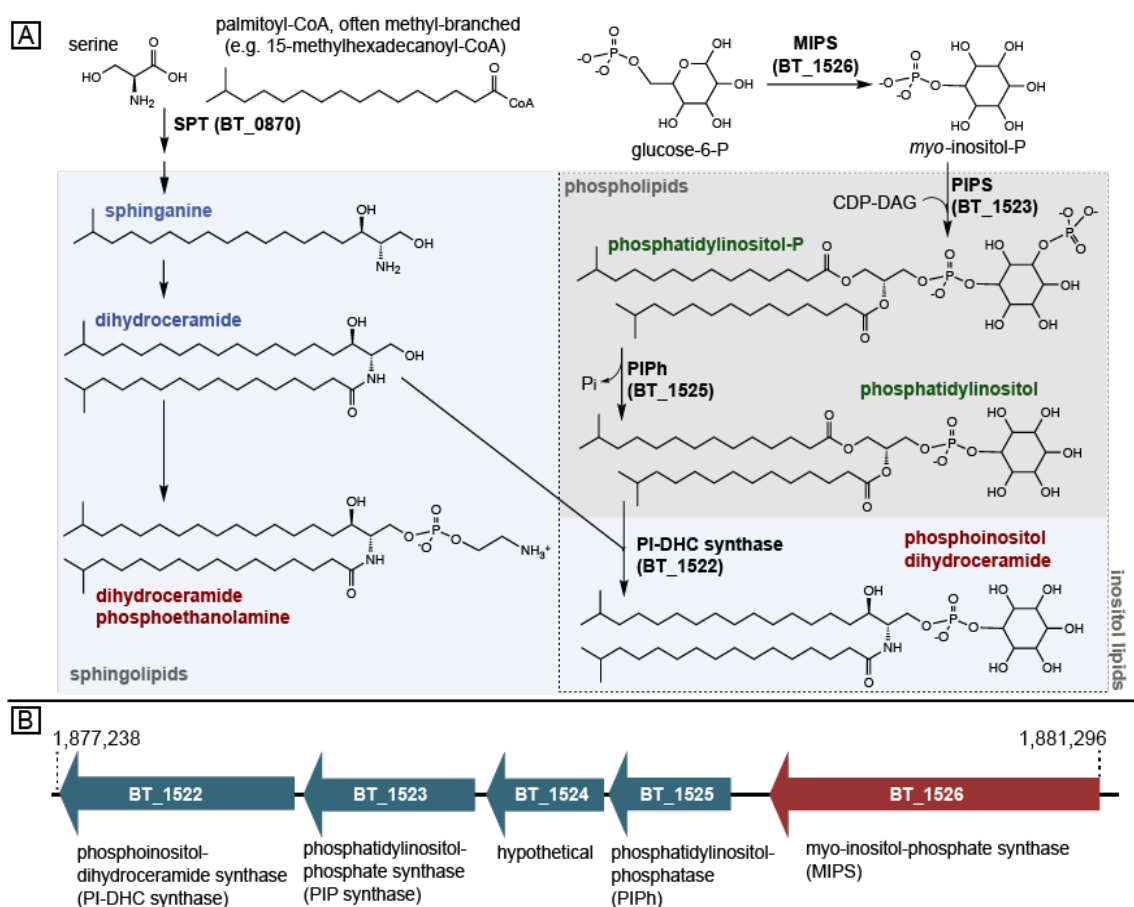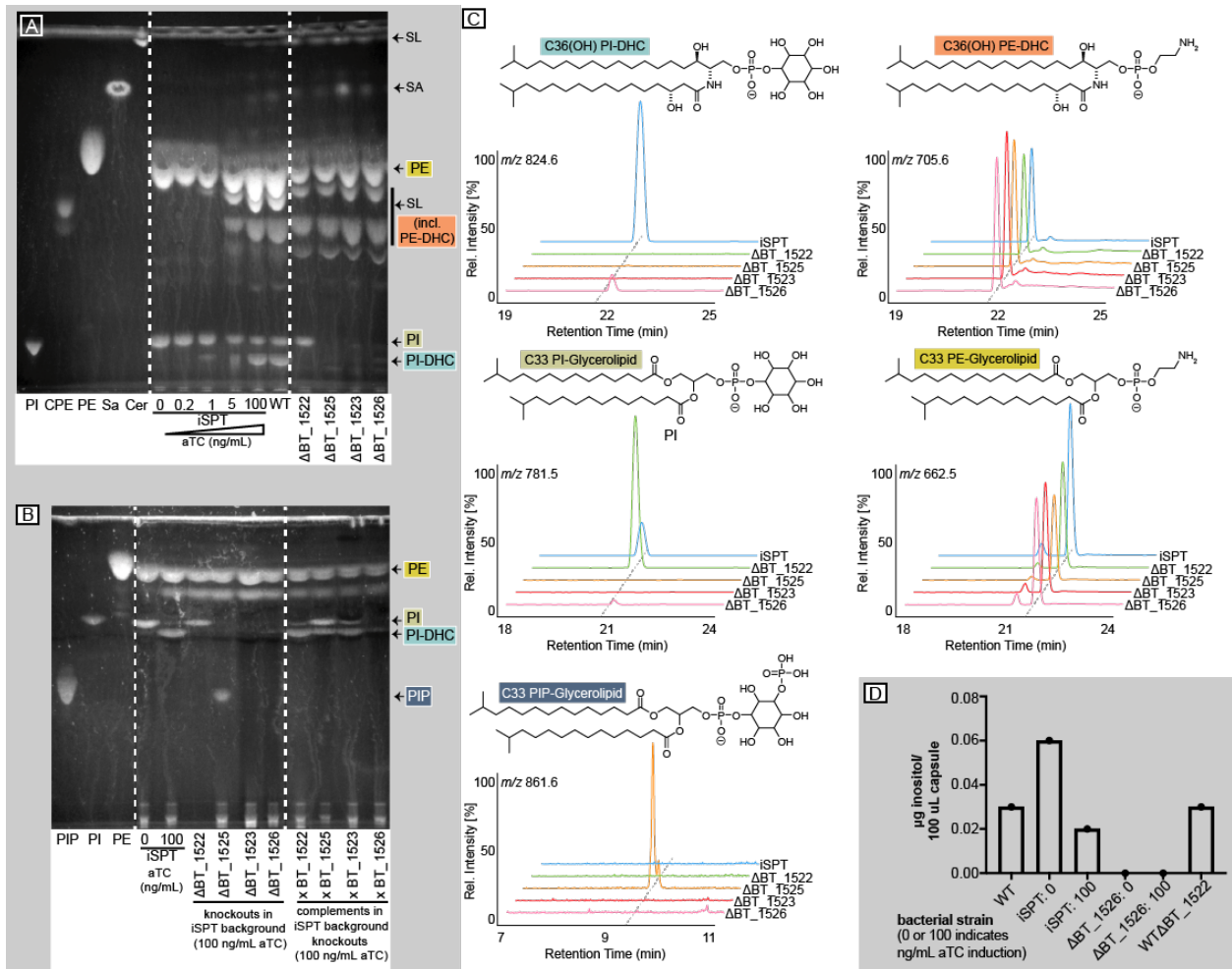50   capacity for inositol lipid synthesis among gut-associated Bacteroidetes.

51

52

53

4

54 **Results and Discussion**

55   We first identified genes responsible for inositol lipid metabolism in BT (Fig. 1A). We

56 identified BT_1522 by NCBI Blast-P as having high homology to the yeast enzyme inositol

57 phosphorylceramide synthase (IPC synthase, also known as AUR1) that catalyzes the attachment of

58 the phophorylinositol group onto ceramide (query cover 50%, e-value 1e-15, percent identity 26%).

59 This led us to hypothesize that this enzyme is responsible for phosphoinositol SL synthesis in BT,

60 though BT SLs have predominantly dihydroceramide (not ceramide) backbones, leading instead to

61 the synthesis of phosphoinositol dihydroceramide (PI-DHC). *BT_1522* and its gene cluster (Fig. 1B)

62 were previously predicted to be involved in inositol lipid metabolism [14]. Adjacent predicted genes in

63 the cluster include BT_1523 (annotated as a CDP-diacylglycerol-inositol 3-phosphatidyltransferase),

64 BT_1524 (hypothetical protein), BT_1525 (currently annotated as phosphatidylglycerophosphatase A,

65 PgpA), and BT_1526 (*myo*-inositol phosphate synthase, "MIPS").

66

67 **Figure 1. Enzymatic pathway for inositol lipid synthesis in BT.** (A) The *de novo* sphingolipid synthesis
68 metabolic pathway in relation to inositol lipid synthesis, with BT enzymes investigated in this study
69 (bolded black) and representative lipid structures. Sphingolipid structures are on a light blue
70 background; phospholipid structures are on a gray background. (B) The genomic region of BT inositol
71 and inositol lipid synthesis. Gene color (blue or red) indicates membership in an operon predicted by
72 BioCyc. Annotations are of the enzyme functions elucidated in this study (due to the lack of a lipid
73 phenotype in its knockout strain, BT_1524 was not investigated further and remains hypothetical).
74

75
76
77 **Figure 2. BT produces inositol phospholipids and sphingolipids.** (A) Thin layer chromatography (TLC)
78 of left to right: first section indicated by dashed lines includes five standards: PI = 16:0
79 phosphatidylinositol; PIP = 18:1 PI(3)P; CPE = ceramide phosphoethanolamine; PE = egg yolk
80 phosphatidylethanolamine, Sa = d18:0 sphinganine, Cer = d18:1/18:0 ceramide; second section
81 includes six standard (non-acidic) lipid extracts from the iSPT BT strain (used as a background for
82 knockout generation) at 0, 0.2, 1, 5 and 100 ng/mL aTC induction of SPT, and WT BT VPI-5482; third
83 section includes standard lipid extraction from ΔBT_1522, ΔBT_1523, ΔBT_1525, and ΔBT_1526
84 knockout strains in the iSPT background at 100 ng/mL aTC induction of SPT. (B) TLC of: standards
85 (PIP, PI, PE) as in panel A; second section shows PIP lipid extractions of iSPT strains at 0 and 100
86 ng/mL aTC followed by ΔBT_1522, ΔBT_1523, ΔBT_1525, ΔBT_1526, and each of their respective
87 complementations (all strains in the iSPT background) at 100 ng/mL aTC induction of SPT. (C)
88 Predicted structures and ion chromatograms demonstrating detection of inositol lipids and
89 sphingolipids in iSPT, ΔBT_1522, ΔBT_1523, ΔBT_1525, and ΔBT_1526 (all strains in the iSPT
90 background) at 100 ng/mL aTC induction. (D) Quantification of inositol in the capsule of BT strains

6

91 with or without sphingolipids (iSPT 100 vs. 0 ng/mL aTC induction), PI-DHC synthase (WTΔBT_1522),

92 and MIPS (ΔBT_1526, at 0 and 100 ng/mL aTC induction).

93

94 We next constructed a BT strain with tunable SL synthesis. As inositol lipids in BT include

95 both glycerophospholipids (PI) and SLs (PI-DHC), we created a strain of BT with inducible control of

96 the first enzyme in the *de novo* SL synthesis pathway, serine palmitoyltransferase (SPT; BT_0870; Fig.

97 1) [19]. This inducible-SPT ("iSPT") strain enables precise control over BT synthesis to produce both PI

98 and PI-DHC, or solely PI. As expected, in the absence of SPT, we detected no SLs by thin-layer

99 chromatography (TLC) analysis (Fig. 2A). However, in iSPT, SL synthesis gradually increased with

100 increasing levels of the anhydrotetracycline (aTC) inducer, to an approximate wild-type (WT) SL

101 abundance with 100 ng/mL aTC induction (Fig. 2A). At full induction, SLs composed roughly half of

102 the extracted lipids measured by TLC densitometry (47±7%, n=6). These SLs include PI-DHC and

103 phosphoethanolamine dihydroceramide (PE-DHC), among others (Fig. 2C).

104 To uncover the function of each predicted enzyme in the putative inositol lipid metabolism

105 pathway, we knocked out the individual genes (BT_1522 to BT_1526) in the iSPT background by

106 scarless deletion [20] (denoted ΔBT_1522 to ΔBT_1526). BT_1522 was also knocked out in the WT

107 background (indicated by WTΔBT_1522). We examined the lipid content of the resulting knockout

108 strains (with SL synthesis fully induced) using TLC and HPLC-MS. Consistent with the predicted role

109 for BT_1522 as a PI-DHC synthase, the ΔBT_1522 strain failed to produce PI-DHC, but production of

110 PI and non-inositol SLs, including PE-DHC, was unaltered (Fig. 2A-C). Similarly, the ΔBT_1526 strain

111 (lacking the predicted MIPS) failed to produce both PI and PI-DHC, in accordance with the loss of the

112 *myo*-inositol-phosphate substrate.

113 Interestingly, both the ΔBT_1523 and ΔBT_1525 strains also failed to produce both PI and PI-

114 DHC (Fig. 2A). As the synthesis of other glycerophospholipids did not appear to be affected in the

115 ΔBT_1525 strain, this observation was not in agreement with the annotated function of BT_1525 as a

116 PgpA [21]. We hypothesized that BT may use a two-step process to synthesize PI similar to that found

117 in *Mycobacteria*, which uses a PIP intermediate [18]. In accord, comparison of the functional protein

7

118   motifs in BT_1523 and BT_1525 with those in the characterized *Renobacterium salmoninarum* PIP

119   synthase (PIPS) [22] revealed the same conserved catalytic residues (DX$_2$DGX$_2$AR…GX$_3$DX$_3$D) in BT_1523.

120   This observation supports the notion that BT_1523 functions as a PIPS in the biosynthesis of PIP.

121   To our knowledge, PIP has not been reported in BT. This is likely due, in part, because PIP

122   extractions under non-acidic conditions may be low yielding. Also, PIP abundance may be inherently

123   low in BT, similar to low phosphoinositide abundances in eukaryotes [3]. Following a PIP-optimized

124   lipid extraction of BT and the knockout strains, we detected high levels of PIP in ΔBT_1525,

125   detectable by both TLC and HPLC-MS (Fig. 2B-C). PIP accumulation in ΔBT_1525 suggests BT_1525 is

126   most likely a phosphatidylinositolphosphatase (PIPh), responsible for the rapid downstream

127   conversion of PIP to PI. Though BT_1525 has homology to the phosphatidylglycerophosphatase A

128   protein family (Pfam), this sequence similarity could reflect an expansion of the functional role of this

129   protein motif from the dephosphorylation of a phospholipid glycerophosphate headgroup to an

130   inositolphosphate headgroup. Previous work has shown that transcriptomic expression of *BT_1525* is

131   higher than *BT_1523* and *BT_1522* in every growth phase of BT [23], likely enabling the rapid

132   conversion of PIP to PI and preventing accumulation of PIP in BT.

133   Despite the central location of *BT_1524* in the inositol lipid metabolism gene cluster, inositol

134   lipids in ΔBT_1524 phenocopied BT by TLC. The *BT_1524* gene is predicted to encode an integral

135   membrane protein with a Gtr-A motif (Pfam); other Gtr-A family proteins are involved in cell surface

136   polysaccharide or exopolysaccharide synthesis [24–26], suggesting BT_1524 may be involved in the

137   membrane transfer of an inositol-linked lipid. Due to a lack of a detectable lipid phenotype in this

138   mutant, we did not investigate it further.

139   To confirm that the loss of inositol lipids in knockout strains was not due to off-target effects,

140   the native BT sequence of each gene was integrated genomically into knockout strains in the iSPT

141   background (ΔBT_1522, ΔBT_1523, ΔBT_1525, ΔBT_1526), paired with a constitutive promoter

142   optimized for BT [27]. The complementation was successful for three of the four strains (ΔBT_1522,

143   ΔBT_1523, and ΔBT_1525), fully restoring the capacity for both PI and PI-DHC synthesis (Fig. 2B).
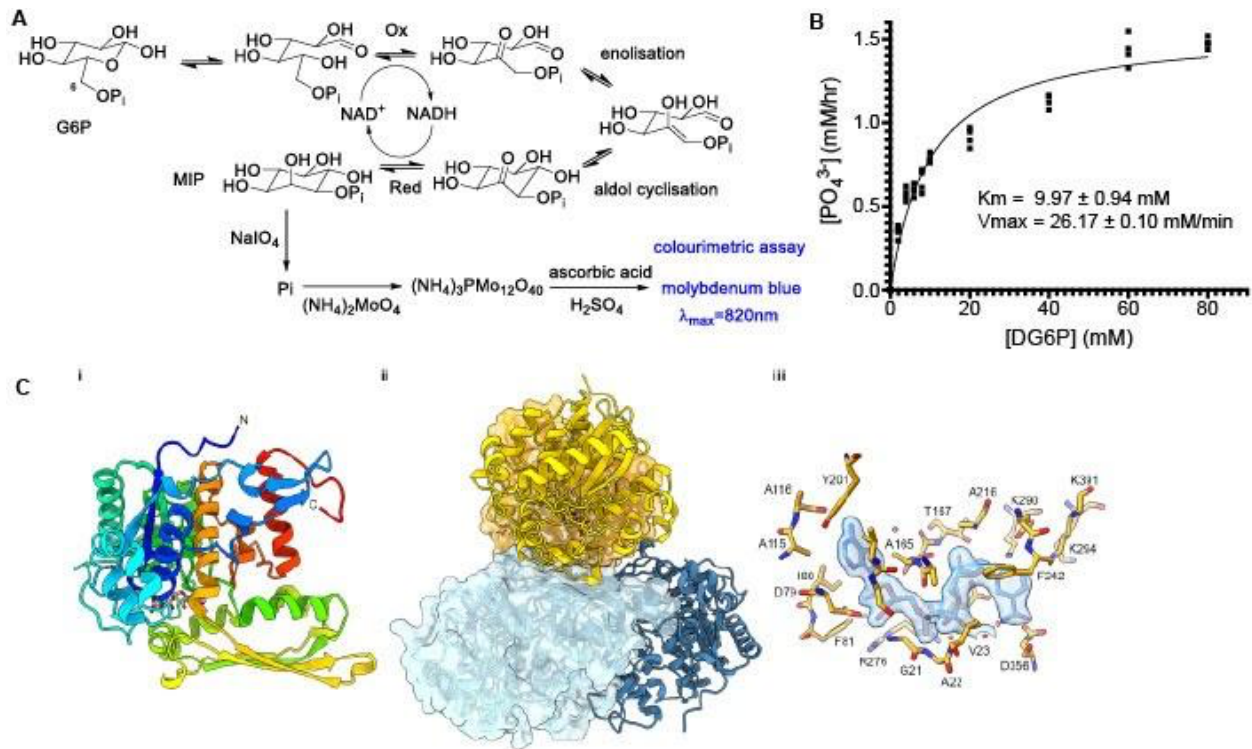
Figure 3. BT_1526 produces *myo*-inositol-phosphate *in vitro*. (A) Proposed mechanism for the MIPS-catalyzed NAD-dependent/redox-neutral conversion of G6P to MIP. (B) Molybdenum blue assay for detection of MIP. Kinetic analysis of recombinant BT_1526 MIPS using G6P as substrate. (C) The crystal structure of BT_1526 MIPS: (i) The monomer subunit, (ii) the tetramer, (iii) the structure of the MIPS:NAD complex.

To confirm the predicted function of BT_1526 as a redox-neutral, $NAD^+/NADH$-dependent MIPS, we cloned the gene and heterologously overexpressed the protein in *E. coli* (Fig. 3 and Supporting Information). The N-terminally His-tagged BT_1526 expressed well in a highly soluble form (~50 kDa in size, observed by SDS-PAGE and confirmed by electrospray ionization mass spectrometry, see SI), and was purified to homogeneity by standard immobilized metal affinity chromatography methods. The MIPS activity of BT_1526 was confirmed using a colorimetric endpoint assay which monitors the appearance of the inorganic phosphate released from the MIP product and not the G6P substrate [28] (Fig. 3A). Kinetic analysis of BT_1526 MIPS operating on G6P are as follows: K_m = 9.97 ± 0.94 mM, V_{max} = 26.17 ± 0.10 µM/min, specific activity = 0.513 µmol/min/mg (Fig. 3B). This activity is in the range of the published specific activities of MIPS from *S. cerevisiae* (0.41 µmol/min/mg) [29,30], *Synechocystis* sp. (0.02 µmol/min/mg) [29], *A. fulgidus* (11.8 µmol/min/mg) [31], and *A. thaliana* (~0.1 µmol/min/mg) [32].

9

163　　　　　We further characterized the BT_1526 protein by determining the X-ray crystal structure to

164　　2.0 Å resolution by molecular replacement using a model derived from the *Archaeoglobus fulgidus*

165　　MIPS structure (PDBID: 3QVT) (refinement statistics in Supplementary Table 1). The overall 3D fold of

166　　BT_1526 is consistent with other members of this family with a Rossman fold-like nucleotide binding

167　　domain (residues 4-244, 355-429) with an intercalated catalytic/dimerization domain (residues 245-

168　　354) [33–36]. The protein adopts a tetrameric dimer of dimers quaternary structure of approximately

169　　188 kDa, with the dimerization domain forming an extended beta-sheet between monomers to

170　　create a saddle-like interface for the two dimers within the tetramer. Though the protein was

171　　purified without the addition of any cofactors or substrates, strong electron density consistent with

172　　the NAD^+ cofactor was observed in the initial maps calculated after molecular replacement and the

173　　final structure contains an NAD^+ molecule associated with each chain modelled at unit occupancy.

174　　Given the redox neutrality of the MIPS enzyme (*i.e.*, it catalyzes substrate oxidation, then reduction)

175　　there is a clear benefit to the retention of the NAD^+ in the active site for the lifetime of the protein, in

176　　agreement with NAD^+ retention in MIPS from other species [37]. The catalytic active site region of the

177　　protein is well conserved among members of the MIPS family for which a structure has been

178　　determined, with a cluster of lysine and aspartic acid residues responsible for binding and orienting

179　　the G6P substrate for isomerization to the MIP product, highlighting the importance of these

180　　residues for the correct activity of the enzyme. Outside of the highly conserved ligand binding site,

181　　the overall fold and quaternary structure of the representatives of the family in the PDB is highly

182　　conserved, although eukaryotic MIPS proteins have an N-terminal extension which appears to

183　　further stabilize the quaternary structure of the protein (Supp. Fig. 4-5). Overall, the enzyme

184　　described here as the first MIPS representative in the Bacteroidetes retains the key functional

185　　elements of other MIPS enzymes, underscoring its conservation across biological kingdoms.

186　　　　　To determine the role of MIPS in BT, we compared gene transcription of the ΔBT_1526 strain

187　　to its iSPT strain at full SPT induction. Lacking MIPS, ΔBT_1526 cells lack not only inositol lipids

188　　(PIP/PI/PI-DHC), but also any other molecules for which inositol may be used as a substrate, such as

189   cell surface polysaccharides. Only 29 genes were differentially expressed with greater than 1.5

190   absolute log_2-fold-change in the ΔBT_1526 strain compared to the background iSPT strain at 100

191   ng/mL induction of SPT. These genes were almost entirely involved in capsule biosynthesis (Table S3).

192   Expression of CPS loci was fairly uniform across varied levels of SPT induction in the iSPT strain, while

193   the ΔBT_1526 strain had notable upregulation of capsular polysaccharide synthesis loci 1 and 6 (CPS1

194   and CPS6) (Fig. 4).
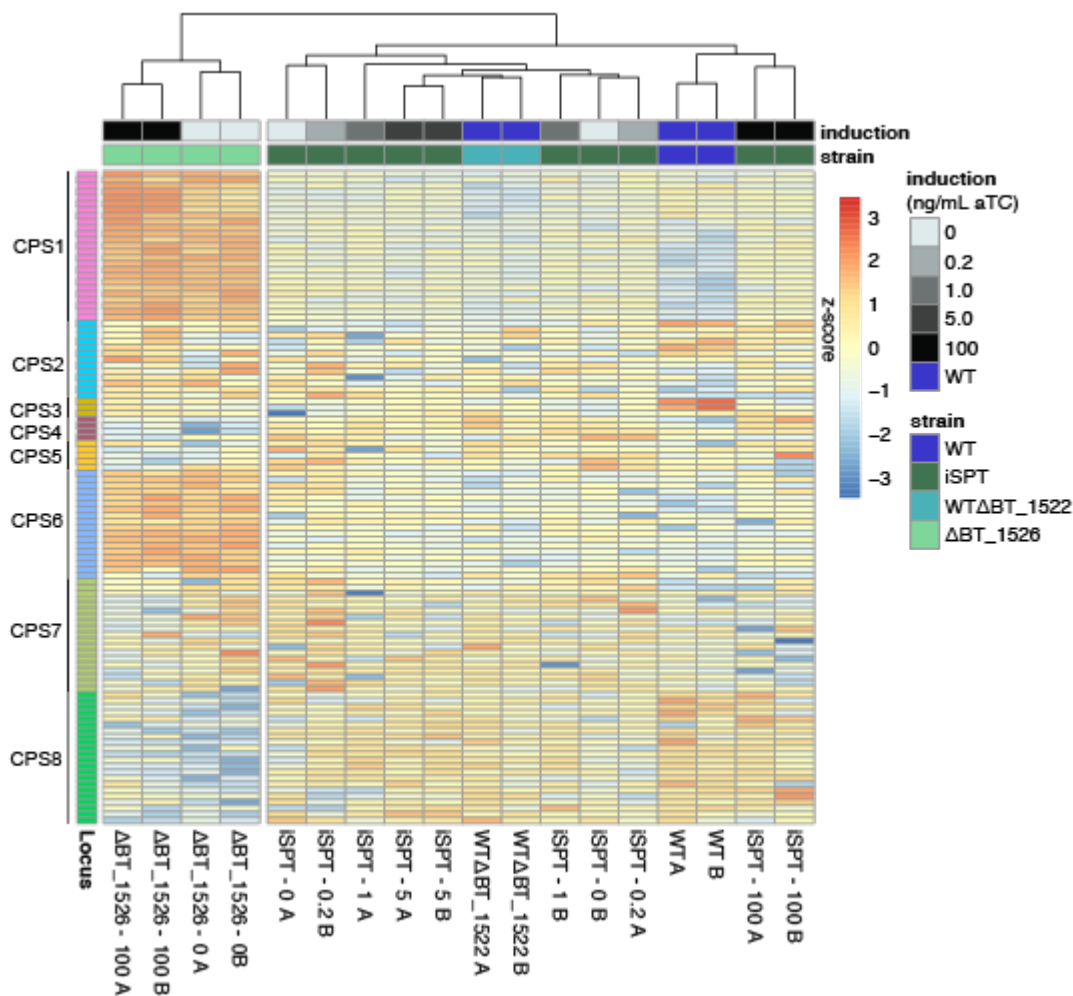
195

196   **Figure 4. Deletion of MIPS (BT_1526) alters expression of genes for capsular polysaccharide**
197   **synthesis pathway loci in BT.** Gene expression data (normalized log_2 expression values, scaled by
198   row, with Euclidean column clustering) in the 8 BT capsular polysaccharide synthesis (CPS) loci. For
199   easier visualization, genes were filtered to include those in which maximum log_2-normalized
200   expression is > 1.5 and exclude those with maximum absolute log_2-fold-change difference in
201   expression < 1.5 in all pairwise comparisons of conditions. Color in the far left column indicates gene
202   assignment to one of 8 CPS loci. Strains tested include WT BT, iSPT, WTΔBT_1522, and ΔBT_1526 in
203   the iSPT background. SPT induction in the iSPT strains at 0, 0.2, 1.0, 5.0, or 100 ng/mL aTC induction
204   is indicated in shades of grey. Labels below each column indicate strain and aTC induction level
205   (ng/mL aTC) redundantly with the color key; "A" and "B" labels represent biological replicates.
206

11

207   To assess whether PI-DHC loss alone was responsible for the capsule effect in BT, we

208   monitored the transcriptional response of WTΔBT_1522 at early stationary phase, grown in minimal

209   medium with glucose as the sole carbon source. Compared to its BT background strain, the

210   WTΔBT_1522 strain had differential expression of 37 genes above a cutoff of 1.5 absolute $\log_2$-fold-

211   change (Table S2). These included many hypothetical proteins and membrane-associated proteins,

212   including those involved in carbohydrate metabolism, such as SusD starch-binding protein homologs

213   (BT_3025 and BT_2806). Pathway enrichment analysis of these 37 genes showed an enrichment of

214   transcripts involved in sugar degradation (specifically, 5-dehydro-4-deoxy-D-glucuronate

215   degradation, p = 0.006, Fisher exact test, Benjamini-Hochberg correction), acetate and ATP formation

216   from acetyl-CoA (p = 0.098), and carboxylate degradation (p = 0.098). In plants and yeasts, which also

217   produce inositol-linked SLs, these lipids are critical for fundamental aspects of the organism's

218   physiology, for example, protein anchoring and programmed cell death in plants [38]. The yeast

219   homolog of BT_1522 is an antifungal target, highlighting its core function in yeast physiology, which

220   is inhibited by the cyclic depsipeptide natural product aureobasidin (hence the name AUR1) [39].

221   Though WTΔBT_1522 had few transcriptomic changes relative to WT controls, its affected pathways

222   appear central to carbohydrate degradation and energy synthesis. As such, other inositol derivatives

223   or PI, not PI-DHC, are implicated in the altered capsule expression of the ΔBT_1526 strain.

224   Inositol has not been previously reported as a component of BT capsule [40], perhaps due to its

225   common use as an internal standard in the HPAEC-PAD analysis of capsule components. Using an

226   alternative standard, we detected inositol in the capsular monosaccharide components of WT BT,

227   iSPT, and WTΔBT_1522 strains (Fig. 2D). However, we did not detect inositol in the capsules of SPT-

228   induced and -uninduced ΔBT_1526 strains. To assess whether the transcriptional trend of the

229   ΔBT_1526 to upregulate CPS1 and CPS6 was reflected in the visible cell capsule, we performed

230   scanning electron microscopy on iSPT and ΔBT_1526 strains with and without SL induction (Fig. 5A).

231   Capsule structures were heterogeneous in both strains, but in comparison to the iSPT background

232   strain, more of the cells of the ΔBT_1526 strain exhibited a dense structure extending from the cell

233   surface, with apparent exopolysaccharide connecting adjacent cells. This was particularly noticeable

234   when SL synthesis was not induced.

235   These results indicate that inositol lipid synthesis is tied in a regulatory fashion to capsule

236   specificity in BT. The CPS loci expressed in a BT population influence both recognition by the host

237   adaptive immune system and bacteriophage predation [40,41]. A role for SLs in mediating the

238   interaction of a microbe with environmental stresses and external threats (*e.g*. antibiotics such as

239   polymyxin and phages, respectively) has also been shown in the freshwater bacterium *Caulobacter*

240   *crescentus*, which responds to phosphate starvation by producing complex glycosphingolipids

241   [42]. Given the remodeling of the capsule observed here when inositol pathways are genetically

242   manipulated, inositol synthesis in BT could influence BT capsule detection by host immunity or

243   phage. The regulatory link between inositol lipid synthesis and capsule specificity implies an indirect

244   role of inositol in host-BT interactions.

245   We assessed how widespread the capacity for inositol lipid synthesis is in the Bacteroidetes.

246   Inositol SLs have only been described in few bacterial species to date [14–16]. We investigated the

247   extent of this biosynthetic capacity across ten representative members of the phylum Bacteroidetes

248   and in other known bacterial SL-producers. Using a homology cutoff of an e-value less than 1e-8, we

249   compared the BT amino acid sequences for MIPS, PIPS, PIPh, PI-DHC synthase, and the BT InsP6

250   phosphatase, MINPP [43] (BT_1526, BT_1523, BT_1525, BT_1522, and BT_4744, respectively) to these

251   related genera (Fig. 5B; Table S4). TLC analysis of lipids from these species revealed that most species

252   with homology to the BT MIPS and PI-DHC synthase had lipid bands consistent with PI and/or PI-DHC

253   in line with their genomically-predicted capacity (the exception was *Flectobacillus major*, a

254   Proteobacterial species with genes encoding protein homology but that did not produce inositol

255   lipids under tested conditions). However, we were surprised to also observe lipid bands consistent

256   with the synthesis of PI and PI-DHC in species lacking homology to BT_1522/23/25. HPLC-MS analysis

257   of these lipids confirmed that two species genomically predicted to lack inositol lipids (*B. vulgatus*

13

258   and *P. veroralis*) in fact produced the same PI-DHC species as those with homology to

259   BT_1522/23/25 (*B. uniformis* and *P. copri*) (Fig. 5C).
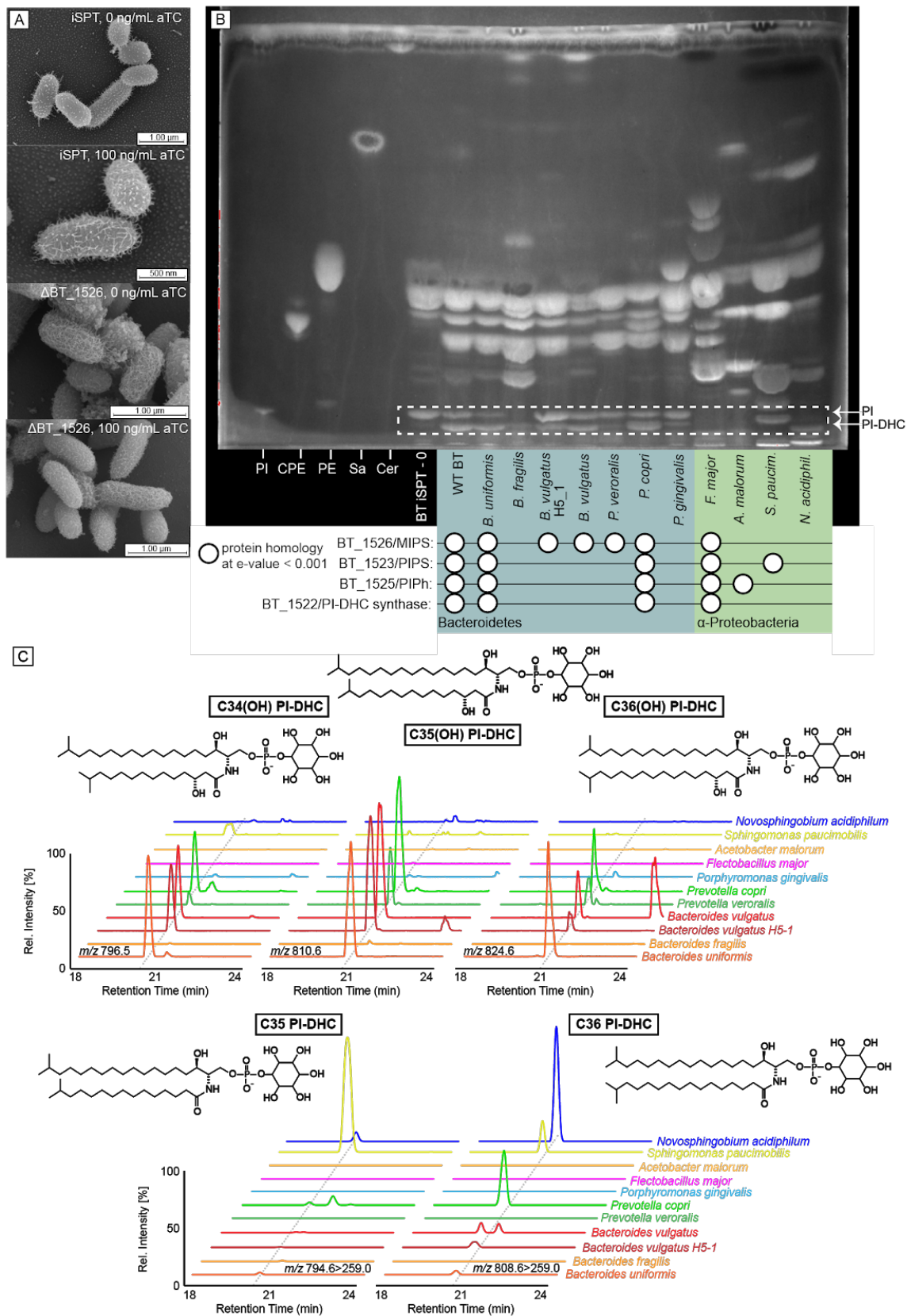
261 **Figure 5. The capacity to produce PI-DHC is widespread among sphingolipid-producing bacteria.** (A)
262 Scanning electron microscopy of the iSPT and ΔBT_1526 strains at 0 and 100 ng/mL aTC induction of
263 SPT grown in minimal medium. (B) TLC of lipid standards and lipid extractions from a diverse array of
264 sphingolipid-producing bacteria. Lanes 1-5, left to right: PI = 16:0 phosphatidylinositol; CPE =
265 ceramide phosphoethanolamine; PE = egg yolk phosphatidylethanolamine, Sa = d18:0 sphinganine,
266 Cer = d18:1/18:0 ceramide. From the sixth lane onward are standard Folch (non-acidic) lipid
267 extractions from: BT iSPT 0 ng/mL aTC induction (no SL), WT BT, *Bacteroides uniformis* (DSM 6597)*,*
268 *Bacteroides fragilis* (DSM 2151)*, Bacteroides vulgatus* H5_1 (DSM 108228)*, Bacteroides vulgatus*
269 (DSM 1447)*, Prevotella veroralis* (ATCC 33779)*, Prevotella copri* (DSM 18205)*, Porphyromonas*
270 *gingivalis* (DSM 20709)*, Flectobacillus major* (DSM 103)*, Acetobacter malorum* (DSM 14337)*,*
271 *Sphingomonas paucimobilis* (ATCC 29837)*,* and *Novosphingobium acidiphiulum* (DSM 19966).
272 Homology to BT protein sequences in the inositol lipid cluster using NCBI BlastP (at e-values below
273 0.001) are indicated below species names with a white circle. Bacteroidetes spp. are on a blue
274 background; alpha-Proteobacteria spp. are on a green background. (C) Predicted structures and ion
275 chromatograms of PI-DHC structures in lipids extracted from the diverse sphingolipid-producing
276 species shown in the same order in panel B. Phosphoinositol dihydroceramide structures include
277 C34(OH) PI-DHC, C35(OH) PI-DHC, C36(OH) PI-DHC, C35 PI-DHC, and C36 PI-DHC.
278

279   To understand this unexpected result, we searched the genomes of related species

280 containing a BT_1526 (MIPS) homolog but lacking homology to the remainder of the BT cluster. Using

281 PHI-BLAST with the conserved catalytic residues in BT_1523 ($DX_2DGX_2AR...GX_3DX_3D$) [22], we identified

282 a predicted CDP-alcohol phosphatidyltransferase in the vicinity of the MIPS homolog in *Bacteroides*

283 *vulgatus.* Expanding the analysis to include other genomes from the Bacteroidetes, we observed that

284 almost every *Bacteroides/Prevotella* species containing a MIPS homolog had one of two clusters

285 directly in the vicinity of the *MIPS* gene - either the BT-like cluster (BT_1522/23/25), or an alternate

286 cluster including an NTP transferase (nucleotidyltransferase) domain-containing protein, CDP-

287 alcohol-phosphatidyltransferase, and haloalkanoate dehalogenase (HAD) hydrolase (Fig. 6). The NTP

288 transferase domain family protein (NCBI Conserved Domain Family cl11394) also shares homology

289 with a phosphocholine cytidyltransferase motif, suggesting this protein may synthesize cytidine 5'-

290 diphosphoinositol (CDP-inositol), similar to the synthesis of CDP-inositol as a precursor to di-*myo*-

291 inositol phosphate solutes in hyperthermophiles [44]. The HAD hydrolase superfamily is large and

292 diverse, with the majority of characterized members functioning as phosphotransferases [45]. As a lipid

293 phosphate phosphohydrolase, this HAD hydrolase may function similarity to AUR1[46], acting as a PI-

294 DHC synthase like BT_1522.

15

295   The functions of these genes (HAD hydrolase, NTP transferase domain protein, and CDP-

296   alcohol-phosphatidyltransferase) are not confirmed, but offer an alternative pathway enabling

297   synthesis of PI-DHC without a PIP intermediate (similar to PI synthesis in eukaryotes [47]), with PI

298   synthesis resembling the synthesis of phosphatidylethanolamine or phosphatidylcholine in the

299   Kennedy pathway [48,49]. Following this logic, the NTP transferase protein would first synthesize CDP-

300   inositol from *myo*-inositol phosphate and CTP. CDP-inositol and a diacylglycerol (DAG) substrate

301   would then be converted to PI, and PI would be converted to PI-DHC by the HAD hydrolase. The MIPS

302   homolog is most commonly clustered directly with these other genes, with some exceptions (*e.g.,*

303   *Prevotella copri*) (Fig. 6). Interestingly, in *P. veroralis*, the CDP-alcohol phosphatidyltransferase and

304   HAD hydrolase proteins are fused (Supp. Fig. 5), suggesting the possibility for a cohesive single-

305   enzyme conversion of CDP-inositol to PI-DHC through a PI intermediate. Some additional putative

306   enzymes are shared in the vicinity of both clusters, including genes annotated as a

307   lysylphosphatidylglycerol synthase (BT_1521 homolog) and a carboxypeptidase-regulatory-like

308   domain protein (BT_1527 homolog) (Fig. 6). This alternative pathway could explain PI-DHC synthesis

309   by *P. veroralis* and *B. vulgatus* despite lack of homology to the BT inositol lipid cluster

310   (BT_1522/23/25) .

311   Among the Proteobacteria species tested, *Sphingomonas paucimobilis* and *Novosphingobium*

312   *acidiphilum* did not make hydroxylated SLs (data not shown). Despite lacking homology to either the

313   BT-like or putative alternative inositol lipid cluster, *N. acidiphilum* produced a SL with a retention

314   time and headgroup fragmentation consistent with *Bacteroides* PI-DHC fragmentation (Supp. Fig.

315   2A). *S. paucimobilis* also produced a lipid with fragmentation similar to *Bacteroides* PI-DHC, but did

316   not produce the fragment at 241 m/z, suggesting a phosphorylated-hexose DHC unlike those

317   produced by *Bacteroides* (Supp. Fig. 2B). In addition, the TLC analysis (Fig. 5B) shows a lipid band in

318   the PI/PI-DHC region for mouth-associated *Porphyromonas gingivalis*, which is likely a

319   phosphorylglycerol-DHC (Supp. Fig. 2C) [50].
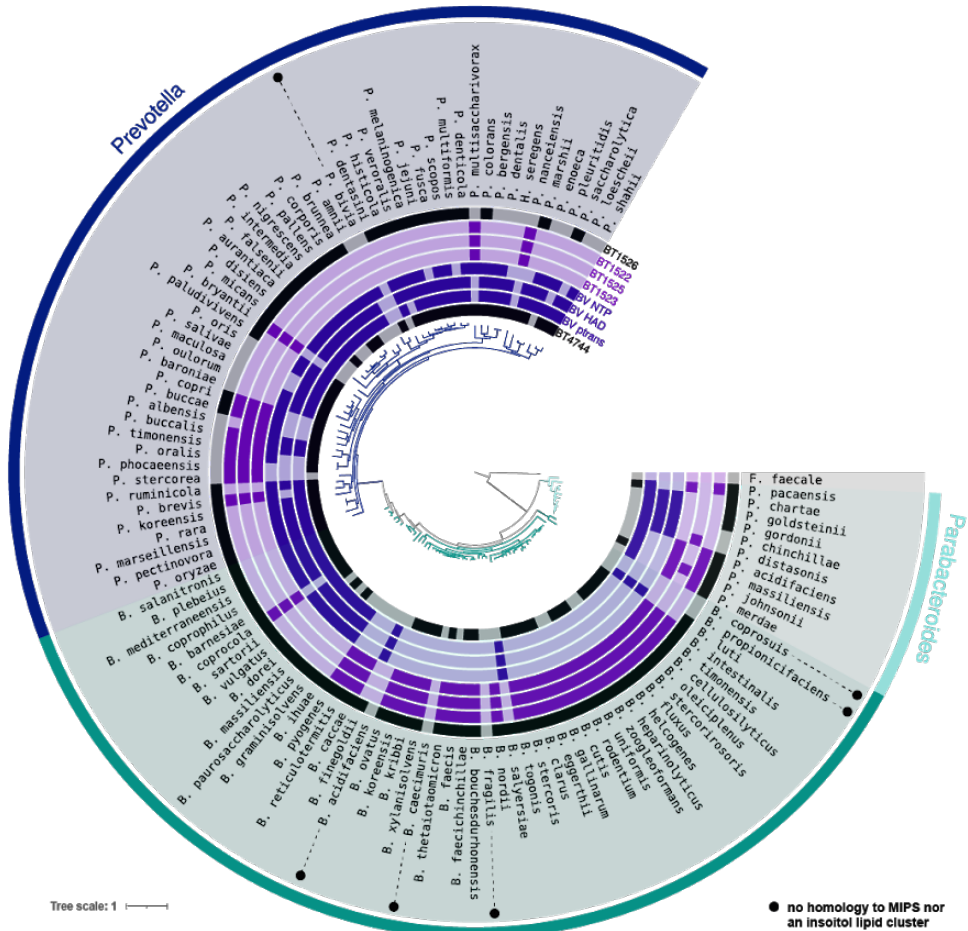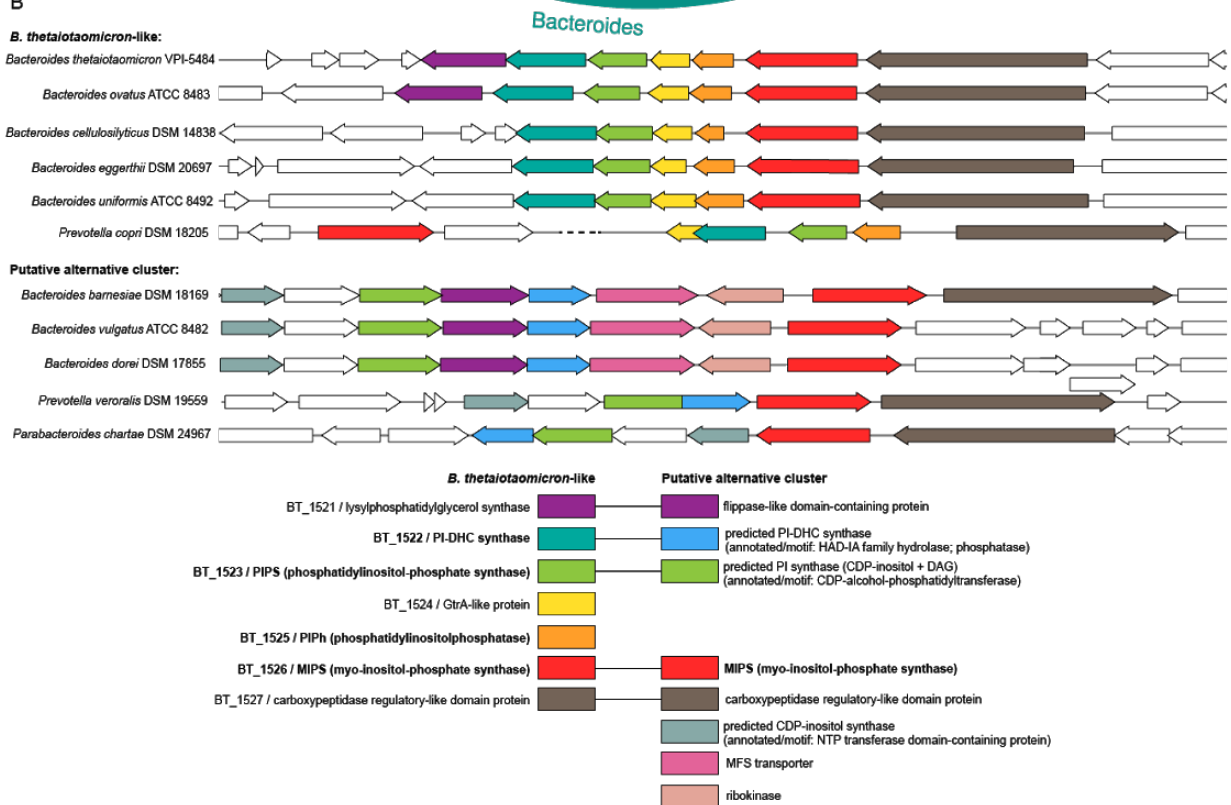
Figure 6. The capacity for inositol lipid synthesis is widespread within the Bacteroidetes. (A) Maximum likelihood based phylogeny of representative *Bacteroides, Prevotella,* and *Parabacteroides*

323 species, produced from 71 conserved single copy genes present in all genomes (identified and
324 concatenated using Anvi'o), and generated by RAxML (best tree; substitution model PROTCAT, matrix
325 name DAYHOFF, Hill-climbing algorithm, bootstrap 50); *Flavobacterium faecale* is included as an
326 outgroup. The rings surrounding the tree indicate species with genes that have NCBI BlastP homology
327 to the BT inositol lipid cluster (in light purple; BT_1522, BT_1523, BT_1525, BT_1526), the BT Minpp
328 (BT_4744), or representative proteins from the *Bacteroides vulgatus* putative alternative inositol lipid
329 cluster (in dark purple; phosphatidyltransferase: BVU_RS13105, HAD hydrolase: BVU_RS13115, NTP
330 transferase: BVU_RS13095). Homology at an e-value below 1e-8 is indicated by dark coloration in the
331 inner circles. (B) Genomic regions surrounding the BT_1526/MIPS homolog in representative
332 Bacteroidetes, compiled using the PATRIC 3.6.9 Compare Region Viewer. Protein homology
333 (determined using NCBI-BlastP) to proteins in the BT-like inositol lipid metabolism cluster (left in key)
334 or the *Bacteroides vulgatus*-like putative alternative inositol metabolism cluster (right in key) is
335 indicated by color. The functions of enzymes in bold were characterized in this study; sequences with
336 predicted redundant functions between both clusters are linked in the key.
337

338   To assess the distribution of the BT inositol lipid cluster (BT_1522/23/25) and the potential

339 alternative pathway among the Bacteroidetes, we searched for homology in 162 representative

340 species of these genera (*Bacteroides, Prevotella, Parabacteroides, Porphyromonas,*

341 *Sphingobacterium,* and *Chlorobium* spp.) (Table S4). Most strains with a homolog to BT_1522,

342 BT_1523, or BT_1525, in fact have homologs of all three of these enzymes, with a distribution that

343 does not track phylogeny, supporting the lateral exchange of this full cluster among these host-

344 associated species (Fig. 6). Roughly three-quarters of species we assessed belonging to the

345 *Bacteroides, Prevotella,* and *Parabacteroides* genera have a MIPS/BT_1526 homolog, and most

346 species with a MIPS homolog contain either the BT-like inositol lipid cluster or the putative

347 alternative cluster. One notable exception is *Bacteroides fragilis,* which does not produce inositol

348 lipids but synthesizes the bioactive glycosphingolipid α-galactosylceramide [51,52]. The BT-like inositol

349 lipid gene cluster is roughly two times more common than the alternative cluster among *Bacteroides*

350 species, while the alternative cluster is about four times more common among the *Prevotella*

351 species. Homologs to proteins in either cluster were absent or highly infrequent in the genera

352 *Porphyromonas, Sphingobacterium,* and *Chlorobium*, with the exceptions of moderate alternative

353 cluster homology in *Sphingobacteria*, and extensive BT_1525 homology in *Chlorobium* (Table S4)*,*

354 which may reflect a true phosphatidylglycerophosphatase function.

355   Our comparative genomic analyses revealed inositol lipid synthesis to be far more

356   widespread in host-associated Bacteroidetes than previously thought. Although the putative

357   alternative pathway remains to be functionally confirmed, the vast majority of species encoded one

358   of the two pathways, with the alternative pathway more common in the *Prevotella*. The extensive

359   prevalence of this function is in agreement with the widespread capacity in gut commensals for

360   phytate (InsP6) degradation, which releases phosphorylated inositol derivatives. Although InsP6

361   phosphatase (Minpp) is rare across Bacteria (present in only 2.2% of completed genomes in EMBL-

362   EBI in 2014), the majority of these enzymes are found in gut microbiome-affiliated species [43]. In

363   addition to the widespread capacity for *de novo* synthesis of inositol and its lipids reported here,

364   these observations suggest that inositol and inositol lipid cycling in the gut are fundamental

365   attributes of the gut microbiome.

366   Bacterial lipids with high structural similarity to eukaryotic bioactive lipids (*e.g.,* SLs) have

367   been shown to influence the metabolism and immune homeostasis of their hosts [14,19,51,53–55].

368   Likewise, bacteria are already known to manipulate their host through inositol and inositol lipid

369   metabolic pathways [43,56], and many bacterial and viral pathogens have also adapted to hijack the

370   host phosphoinositide system [57,58]. Thus, a precedent exists for trans-kingdom manipulation of

371   inositol levels, to either a beneficial or detrimental outcome for the host. As one of the most

372   abundant phyla within the human gut, the widespread synthesis of inositol lipids from gut-associated

373   Bacteroidetes (*Bacteroides*, *Prevotella*, and *Parabacteroides* spp.) could represent a significant

374   contribution to the lipid milieu of the gut. Of the six most prevalent and abundant *Bacteroides*

375   species in the human gut [59], five have genes with homology to the BT-like inositol lipid cluster (*B.*

376   *cellulosyliticus, B. eggerthii, B. ovatus*) or the potential alternative cluster (*B. dorei, B. vulgatus*),

377   indicating potential for inositol lipid synthesis. How these lipids interact with the human host remains

378   to be investigated.

379

380

19

381   **Conclusion**

382   Inositol lipids have only recently been reported in commensal gut bacteria. In this study, we

383   characterized the gene cluster recently hypothesized to be involved in bacterial inositol lipid

384   synthesis in BT, in the first known study to show a functional role for these genes in the

385   Bacteroidetes. BT synthesizes PI using a mycobacterial-like pathway with a PIP intermediate;

386   previously, the bacterial PI synthesis pathway lacked a PIPh, which we have identified here as

387   *BT_1525*. We also identified a putative alternative pathway for PI-DHC synthesis common among

388   *Prevotella* species that lacks a PIP intermediate resembling the eukaryotic Kennedy pathway for

389   phosphatidylethanolamine and phosphatidylcholine synthesis. The majority of Bacteroidetes encode

390   one or the other of these pathways, indicating that inositol lipid production is a fundamental trait in

391   the phylum. Together with the importance of inositol lipids in pathogen-host interactions [60], their

392   high prevalence in the host-associated Bacteroidetes suggests an unexplored role in host

393   interactions, potentially mediated both directly through provisioning and indirectly via effects on the

394   capsule.

395
396

397   **Materials and Methods**

398   Bacterial strains and culturing conditions

399   Unless otherwise stated, all liquid *B. thetaiotaomicron* VPI-5482 (BT) cultures were grown

400   anaerobically (95% N_2 and 5% CO_2 atmosphere) at 37℃ in supplemented BHI media (BHIS; 37 g/L brain-heart

401   infusion, 5 g/L yeast extract, 1 mg/L menadione, 1 mg/L resazurin, 10 mg/L hemin, 0.5 g/L cysteine-HCl). *E. coli*

402   cultures were grown aerobically at 37℃ in Luria broth with shaking. Final concentrations of antibiotics and

403   selection agents were as follows: erythromycin 25 μg/mL, gentamicin 200 μg/mL, streptomycin 100 μg/mL,

404   carbenicillin 100 μg/mL, 5-fluoro-2'-deoxyuridine 200 μg/mL. In select experiments, BT was grown in

405   Bacteroides minimal media (BMM); per liter: 13.6 g KH_2PO_4, 0.875g NaCl, 1.125 g $(NH_4)_2SO_4$, 5 g glucose, (pH to

406   7.2 with concentrated NaOH), 1 mL hemin (500 mg dissolved in 10 mL of 1M NaOH then diluted to final volume

407   of 500 mL with water), 1 mL $MgCl_2$ (0.1 M in water), 1 mL $FeSO_4x7H_2O$ (1 mg per 10 mL of water), 1 mL vitamin

408   K3 (1 mg/mL in absolute ethanol), 1 mL CaCl₂ (0.8% w/v), 250 μL vitamin B12 solution (0.02 mg/mL), 0.5 g L-

409   cysteine HCl.

410   For lipid analysis of non-BT strains: *Sphingomonas paucimobilis* (ATCC 29837) was grown aerobically at

411   30℃ in nutrient broth (per L: 5.0 g peptone, 3.0 g meat extract; pH 7.0). *Bacteroides fragilis* (DSM 2151),

412   *Porphyromonas gingivalis* (DSM 20709), *Bacteroides uniformis* (DSM 6597), *Bacteroides vulgatus* H5_1 (DSM

413   108228), *Bacteroides vulgatus* (DSM 1447), *Prevotella veroralis* (ATCC 33779), and *Prevotella copri* (DSM

414   18205) were grown anaerobically at 37℃ in BHIS. *Flectobacillus major* (DSM 103) was grown at 26℃ in DSM

415   Medium 7 (per L: 1.0 g glucose, 1.0 g peptone, 1.0 g yeast extract; pH 7.0). *Acetobacter malorum* (DSM 14337)

416   was grown at 28℃ in DSM Medium 360 (per L: 5.0 g yeast extract, 3.0 g peptone, 25.0 g mannitol).

417   *Novosphingobium acidiphilum* (DSM 19966) was grown at 28℃ in DSM Medium 1199 (per L: 1.0 g glucose, 1.0

418   g yeast extract, 1.0 g peptone; pH 5.5).

419   <u>Generation of BT knockouts and inducible SPT strain</u>

420   Genetic manipulations in the *Bacteroides thetaiotaomicron* VPI-5482 tdk ("WT") strain were

421   performed using double recombination from a suicide plasmid as previously described [20]. The generation of the

422   *BT_0870* (SPT) knockout is previously described [54]. To create the inducible SPT (iSPT) strain, three TetR

423   cassettes were inserted into the ΔBT_0870 genome with the constitutive PBT1311 promoter as previously

424   described [27], with the native SPT (*BT_0870)* sequence reintroduced under the inducible P1TDP promoter.

425   *BT_1522*, *BT_1523*, *BT_1525*, and *BT_1526* were knocked out using the same process in both the WT and iSPT

426   strains. Complements for each enzyme were created in the iSPT knockout strains, likewise using the

427   constitutive PBT1311 promoter and native BT sequences and integrated genomically, just upstream of the

428   inositol gene cluster. Plasmids, strains, primers, and gene sequences are listed in Table S5. All constructs were

429   verified by Sanger sequencing.

430   <u>Bacterial lipid extraction and thin layer chromatography</u>

431   BT strains were grown 14-20 hours in BHIS; all other strains were grown in the media and

432   temperatures described above to density. **"Standard" (non-acidic) lipid extraction:** Bacteria were pelleted at

433   3500 x g for 15 minutes, the pellet washed in PBS, and re-spun. The washed lipid pellets were lipid extracted by

434   the Folch method [61], the organic fraction dried under nitrogen, the lipid film re-suspended in 2:1 (v/v)

435   chloroform:methanol. **PIP lipid extraction:** To detect PIP, lipids were extracted according to the PI(3)P Mass

436   ELISA Kit (Echelon Biosciences Inc.) protocol. Cells from 50-ml BHIS cultures were pelleted at 3500 x g for 15

437 minutes at 4℃, resuspended in 5-mL cold 0.5 M trichloroacetic acid (TCA), incubated 5 minutes on ice, and

438 pelleted at 3500 x g for 15 minutes at 4℃. The pellets were washed twice in 3 mL 5% TCA with 1 mM EDTA,

439 then neutral lipids were extracted twice by vortexing the pellet in 3 mL 2:1 methanol:chloroform for 10

440 minutes. The resulting pellets were extracted into 2.25 mL methanol:chloroform:12 N HCl (80:40:1), 0.75 mL of

441 chloroform and 1.35 mL of 0.1 N HCl added and vortexed. The lower fraction was dried under nitrogen and

442 resuspended in 20:9:1 chloroform:methanol:water for TLC.

443   Thin layer chromatography of lipids

444   Lipid extracts were applied to a silica HPTLC plate with concentration zone (Supelco #60768), with

445 loading volumes normalized to the OD_{600} of original cultures. Plates were developed in a 62:25:4 (v/v)

446 chloroform:methanol:ammonium hydroxide system (for standard lipid extractions) or 48:40:7:5

447 chloroform:methanol:water:ammonium hydroxide (for PIP extractions), then sprayed with primuline (0.1

448 mg/mL in 4:1 v/v acetone:dH$_2$O), and imaged under UV transillumination (365 nm). Lipid standards include 16:0

449 phosphatidylinositol (Avanti #850141), 18:1 PI(3)P (Avanti #850150), ceramide phosphorylethanolamine

450 (Sigma-Aldrich #C4987), egg yolk phosphatidylethanolamine (Pharmacoepia), d18:1/18:0 ceramide (Cayman

451 #19556), and d18:0 sphinganine (Avanti #860498).

452   Sample Prep For HPLC-MS

453   Samples were frozen over liquid nitrogen and lyophilized to dryness. 1 mL of HPLC grade methanol

454 was added to the dried material and the mixture was sonicated for 3 min (on/off pulse cycles of 2 second on, 2

455 seconds off, at power 100%) using a Qsonica Ultrasonic Processor (Model Q700) with a water bath cup horn

456 adaptor (Model 431C2), with water bath flow to maintain approximately room temperature. Samples were

457 then moved to an end-over-end rotator and extractions proceeded for 12 hours. Samples were then

458 centrifuged at 18000 G for 30 minutes at 4 °C. The supernatant was transferred to a fresh centrifuge tube and

459 solvent was dried with a Thermo Scientific Savant SpeedVac SPD130DLX. The dried material was resuspended

460 in 200 μL HPLC-grade methanol, briefly sonicated, and centrifuged as before. The concentrated extract was

461 transferred to HPLC vial with a 300 μL glass insert and stored at 4 °C until further analysis.

462   HPLC-MS instrumentation

463   LC−MS analysis was performed on a ThermoFisher Scientific Vanquish Horizon UHPLC System coupled

464 with a ThermoFisher Scientific TSQ Quantis Triple Quadrupole mass spectrometer equipped with a HESI ion

465 source. All solvents and reagent for HPLC-MS were purchased as Optima LC-MS grade (Fisher Scientific).

22

466   HPLC-MS generalized method

467   Mobile phase A was 94.9% water, 5% methanol, 0.1% formic acid (v/v) with 10 mM ammonium

468   acetate. Mobile phase B was 99.9% methanol, and 0.1% formic acid (v/v). 1 μL of extract was injected and

469   separated on a mobile phase gradient with an Agilent Technologies InfinityLab Poroshell 120 EC-C18 column

470   (50 mm × 2.1 mm, particle size 2.7 μm, part number: 699775-902) maintained at 50 °C. A/B gradient started at

471   15% B for 1 min after injection and increased linearly to 100% B at 22 min and held at 100% B for 5 min, using a

472   flow rate 0.6 mL/min. Full Scan Q1 mass spectrometer parameters: spray voltage 2.0 kV for negative mode, ion

473   transfer tube temperature 350 °C, vaporizer temperature 350 °C; sheath, auxiliary, and spare gas 60, 15, and 2,

474   respectively. Tandem mass spectrum analysis was carried out with Product Ion Scan mode with the following

475   additions: collision energy: 30 V, CID gas 1.5 mTorr.

476   HPLC-MS method for phosphatidylinositol phosphates

477   The method was slightly modified from Bui et al. 2018 [62]. Mobile phase A was 99.9% water, 0.1% N,N-

478   Diisopropylethylamine (v/v) with 10 uM Disodium EDTA. Mobile phase B was 99.9% acetonitrile, 0.1% N,N-

479   Diisopropylethylamine (v/v). 3 μL of extract was injected and separated on a mobile phase gradient with an

480   Kinetex EVO C18 UHPLC column, 2.1 × 150 mm, 1.7μm (Phenomenex, CA, PN:00F-4726-AN) maintained at 60

481   °C. A/B gradient started at 38% B for 6 min after injection and increased linearly to 100% B at 12 min and held

482   at 100% B for 3 min, using a flow rate 0.35 mL/min. Full Scan Q1 mass spectrometer parameters: spray voltage

483   4.5 kV for negative mode, ion transfer tube temperature 325 °C, vaporizer temperature 350 °C; sheath,

484   auxiliary, and spare gas 50, 15, and 1, respectively. Tandem mass spectrum analysis was carried out with

485   Product Ion Scan mode with the following additions: collision energy: 30 V, CID gas 1.5 mTorr.

486   Capsule monosaccharide analysis

487   For capsule extraction, 16-hour 20 mL BHIS cultures were normalized to OD_{600}, centrifuged 3500 x g

488   for 20 min, and gently washed two times in 50 mL of PBS (16000 x g, 4 min). The pellets were shaken (900

489   RPM) in 500 uL of aqueous phenol for three hours at room temperature, centrifuged at 5,000 x g for 20 min at

490   4°C, and the aqueous phase ethanol precipitated (cold absolute EtOH added to final concentration 80% v/v for

491   2 hr at -20°C, centrifuged at 18000 x g for 20 min at 4°C, washed with cold 80% EtOH and centrifugation

492   repeated). The resulting pellet was dissolved in PBS and treated with Roth Proteinase K (1 hr at 60°C) and

493   Merck Benzonase nuclease (20 min at 37°C). Samples were dialyzed against water (1 kDa MWCO, G-Biosciences

494   Tube-O-DIALYZER) and stored at -80 deg C prior to inositol quantification by the UCSD GlycoAnalytics Core.

495   Purification and enzymatic characterization of BT_1526

496   The synthetic gene encoding the BT_1526 ORF (wild type) was ordered from Genscript cloned into a

497 pET-28a expression plasmid with a six-histidine tag at the N-terminus. The pET-BT_1526 plasmid was used to

498 transform *E. coli* BL21 (DE3) cells for overexpression. The BT_1526 MIPS protein was expressed by culturing the

499 transformed cells in LB medium supplemented with 35 ug/ml kanamycin at 37 ˚C 200 rpm shaking, until the

500 cells reached the mid-exponential growth stage (OD_{600} = 0.5). Protein expression was then induced by the

501 addition of 0.1 mM IPTG for five hours with a reduced temperature of 30 ˚C with shaking 200 rpm. Cells were

502 harvested by centrifugation 4000 x g, 4 ˚C and sonicated in 10 x v/w HisA buffer (50 mM Tris-HCl, 20 mM

503 NH4Cl, 0.2 mM DTT and 30 mM imidazole, pH 7.5) to lyse the cells. The lysate was clarified by centrifugation at

504 35,000 x g, 4˚C and the supernatant was loaded onto a 5 ml HisTrap column (Cytiva) equilibrated with HisA

505 buffer. Unbound proteins were washed off the column with 20 column volumes of HisA prior to elution of the

506 tagged BT_1526 with HisB buffer (20 mM NH_4Cl, 0.2 mM DTT and 500 mM imidazole, pH 7.5). The protein was

507 then subjected to size-exclusion chromatography for polishing and buffer exchange. A Superdex S200 column

508 was equilibrated with GF buffer (50 mM Tris-HCl and 150 mM NaCl) and sample added for isocratic elution over

509 1.2 column volumes. Fractions from size-exclusion chromatography were analyzed by 15 % SDS-PAGE and

510 fractions containing the pure BT_1526 protein were pooled and used in subsequent experiments. The yield of

511 recombinant BT_1526 was typically >20mg per litre of *E. coli* culture.

512   Mass spectrometry analysis of purified MIPS

513   Samples were analysed in the positive ion mode using HPLC coupled to a Waters Synapt G2 QTOF with

514 an electrospray ionisation source (ESI). 5-10 µL of 10 µM protein was injected onto a Phenomenex C4 3.6µ

515 column. The conditions for the qTOF are as follows: source temperature 120 °C, back pressure 2 mbar, and

516 sampling cone voltage 54V. The protein was eluted with a 12 minute gradient, starting at 5% acetonitrile with

517 0.1% formic acid to 95% acetonitrile. The resulting spectra were processed and the charge state distributions

518 deconvoluted using MassLynx V4.1 software.

519   Assay of BT_1526 for MIPS activity.

520   The purified BT_1526 was assayed for MIPS activity was based on a method published by Barnett [28]

521 (see Fig. 3A). The assay was as follows: 1 µM enzyme, 0-50 mM D-glucose-6-phosphate, 0.8 mM NAD+ for 1 hr

522 at 25°C. The reaction was quenched with 20% TCA then 0.2M $NaIO_4$ was added for an hour at room

523 temperature. Then 1.5M Na_2SO_3 was added to remove excess $NaIO_4$. The reagent mix was incubated for 1 hour

524   at room temperature, then the absorbance was measured on BioTek Synergy HT plate reader at 820 nm.

525   Values were determined with reference to inorganic phosphate standards.

526   Crystallization of BT_1526

527   MIPS was initially screened using commercial kits (Molecular Dimensions and Hampton Research). The

528   protein concentration was 11.4 mg/ml. The drops, composed of 0.1 ul or 0.2 ul of protein solution plus 0.1 of

529   reservoir solution, were set up using a Mosquito crystallization robot (SPT Labtech) using the sitting drop vapor

530   diffusion method. The plates were incubated at 20 °C and the initial hits were suitable for diffraction

531   experiments. The condition yielding crystals that were subjected to X-ray diffraction was PACT F6 (Molecular

532   Dimensions, 200 mM Sodium formate, 100 mM Bis Tris Propane pH 6.5, and 20 % (w/v) PEG 3350). The sample

533   was cryoprotected with the addition of 20% PEG 400 to the reservoir solution.

534   Data collection, structure solution, model building, refinement and validation of BT_1526

535   Diffraction data were collected at the synchrotron beamline I04 of Diamond light source (Didcot, UK)

536   at a temperature of 100 K. The data set was integrated with XIA2 [63] using DIALS [64] and scaled with Aimless (52).

537   The space group was confirmed with Pointless [65]. The phase problem was solved by molecular replacement

538   with Phaser [66] using PDB file 3QVT as search model. The model was refined with refmac [67] and manual model

539   building with COOT [68]. The model was validated using Coot and Molprobity [69]. Other software used were from

540   CCP4 cloud and the CCP4 suite [70]. Figures were made with ChimeraX [71].

541   Electron microscopy

542   Imaging was performed by the Electron Microscopy Core at the Max Planck Institute for

543   Developmental Biology in Tübingen, Germany. For scanning electron microscopy (SEM), cells were fixed in 2.5%

544   glutaraldehyde/4% formaldehyde in PBS for 2 hours at room temperature and mounted on poly-L-lysine-

545   coated cover slips. Cells were post-fixed with 1% osmium tetroxide for 45 minutes on ice. Subsequently,

546   samples were dehydrated in a graded ethanol series followed by critical point drying (Polaron) with CO_2.

547   Finally, the cells were sputter-coated with a 3 nm thick layer of platinum (CCU-010, Safematic) and examined

548   with a field emission scanning electron microscope (Regulus 8230, Hitachi High Technologies) at an accelerating

549   voltage of 3 kV.

550   RNA-seq of BT at varied levels of SPT induction

551   Overnight cultures were used to inoculate (in duplicate) BMM media 1:2500 uninduced, or at one of

552   five varied anhydrotetracycline (aTC) concentrations (0, 0.2, 1.0, 5.0, 100 ng/mL), and incubated at 37℃ for 14

25

553   hours to an OD600 of 0.10-0.17. Cultures were spun at 3500 x g for 15 min and RNA was extracted from the

554   bacterial pellet with QIAzol lysis reagent and the miRNeasy Mini Kit (Qiagen). rRNA was removed with the

555   Bacterial RiboMinus Transcriptome Isolation Kit (Invitrogen) and the library prepared with the TruSeq Stranded

556   Total RNA Library Kit (Illumina); libraries were pooled nine per lane and sequenced by HiSeq3000 (Illumina).

557   Quality assessment of reads was performed using FastQC pre- and post-quality filtering with bbduk

558   (quality cutoff = 20) [72]. Reads were aligned to the Ensembl *B. thetaiotaomicron* VPI-5482 genome with bowtie2

559   and assigned using htseq-count (alignment quality cutoff = 10) [73–75]. Differential expression analysis was

560   performed with EdgeR and limma [76,77]: reads assigned to rRNA genes, "ambiguous," or "no feature" were

561   removed, lowly expressed genes were filtered, and gene expression distributions were normalized (method

562   Trimmed Means of M values, "TMM"). Count data from BMM samples, which were in duplicate, were further

563   normalized by Bayes moderated variance before calculation of differential expression (adjusted p-value via

564   Benjamini-Hochberg method). Annotations were assigned from the JGI IMG database. Heatmaps were

565   generated with pheatmap using normalized log2 expression values, scaled by row with Euclidean clustering.

566   <u>Phylogenies of homology to BT inositol lipid metabolic enzymes in diverse bacteria</u>

567   For the smaller phylogeny of diverse sphingolipid-producers (Fig. 5D), homology to BT inositol and

568   inositol lipid metabolism enzymes BT_1522, BT_1523, BT_1525, and BT_1526 was identified using NCBI Blast-P

569   [78] to the indicated species. For the larger phylogeny of Bacteroidetes and related genera (Fig. 5), all

570   representative species for *Bacteroides, Prevotella, Parabacteroides, Porphyromonas, Flavobacterium,*

571   *Sphingobacterium*, and *Chlorobium* genera with nomenclature recognized in the LPSN [79] were tested for

572   homology to BT inositol metabolism enzymes BT_1522, BT_1523, BT_1525, BT_1526, and BT_4744. For

573   phylogenetic comparison in both trees, 71 single copy genes present in all genomes (HMM profile Bacteria_71)

574   were identified and concatened using Anvi'o [80], with alignment using MUSCLE [81]. RAxML [82] was used to

575   generate a maximum likelihood tree (Protcat substitution model, Dayhoff matrix, Hill-climbing algorithm, 50

576   bootstrap iterations). Strain accession numbers and Blast-P results are in Table S4.

577

26

586

587   **Data Availability**

588   The BT_1526 MIPS structure analyzed during the current study is available in the Protein

589   Data Bank repository, PDB 7NWR. Transcriptomic reads, mass spectrometry files, and all unique

590   strains generated in this study are available from the corresponding author upon request. All

591   remaining data generated during this study are included in this published article and its

592   supplementary information files.

593
594

References

595

1. Nakase, M. *et al.* Mannosylinositol phosphorylceramide is a major sphingolipid component and is required for proper localization of plasma-membrane proteins in Schizosaccharomyces pombe. *J. Cell Sci.* **123**, 1578–1587 (2010).

2. Megyeri, M., Riezman, H., Schuldiner, M. & Futerman, A. H. Making sense of the yeast sphingolipid pathway. *J. Mol. Biol.* **428**, 4765–4775 (2016).

3. Dickson, E. J. & Hille, B. Understanding phosphoinositides: rare, dynamic, and essential membrane phospholipids. *Biochem. J* **476**, 1–23 (2019).

4. Phan, T. K. *et al.* Phosphoinositides: multipurpose cellular lipids with emerging roles in cell death. *Cell Death Differ.* **26**, 781–793 (2019).

5. Haites, R. E., Morita, Y. S., McConville, M. J. & Billman-Jacobe, H. Function of phosphatidylinositol in mycobacteria. *J. Biol. Chem.* **280**, 10981–10987 (2005).

6. Yagüe, G., Segovia, M. & Valero-Guillén, P. L. Phospholipid composition of several clinically relevant Corynebacterium species as determined by mass spectrometry: an unusual fatty acyl moiety is present in inositol-containing phospholipids of Corynebacterium urealyticum. *Microbiology* **149**, 1675–1685 (2003).

7. Masayama, A. *et al.* Streptomyces phospholipase D mutants with altered substrate specificity capable of phosphatidylinositol synthesis. *Chembiochem* **9**, 974–981 (2008).

8. Belisle, J. T., Brandt, M. E., Radolf, J. D. & Norgard, M. V. Fatty acids of Treponema pallidum and Borrelia burgdorferi lipoproteins. *J. Bacteriol.* **176**, 2151–2157 (1994).

9. Nigou, J., Gilleron, M. & Puzo, G. Lipoarabinomannans: from structure to biosynthesis. *Biochimie* **85**, 153–166 (2003).

10. Olson, D. K., Fröhlich, F., Farese, R. V., Jr & Walther, T. C. Taming the sphinx: Mechanisms of cellular sphingolipid homeostasis. *Biochim. Biophys. Acta* **1861**, 784–792 (2016).

11. Michell, R. H. Inositol lipids: from an archaeal origin to phosphatidylinositol 3,5-bisphosphate faults in human disease. *FEBS J.* **280**, 6281–6294 (2013).

12. Hannun, Y. A. & Obeid, L. M. Sphingolipids and their metabolism in physiology and disease. *Nat. Rev. Mol. Cell Biol.* **19**, 175–191 (2018).

622 13. Megson, Z. A. *et al.* Inositol-phosphodihydroceramides in the periodontal pathogen Tannerella forsythia:

623   Structural analysis and incorporation of exogenous myo-inositol. *Biochim. Biophys. Acta* **1851**, 1417–1427

624   (2015).

625 14. Brown, E. M. *et al.* Bacteroides-Derived Sphingolipids Are Critical for Maintaining Intestinal Homeostasis and

626   Symbiosis. *Cell Host Microbe* **25**, 668–680.e7 (2019).

627 15. Naka, T. *et al.* Structural analysis of sphingophospholipids derived from Sphingobacterium spiritivorum, the

628   type species of genus Sphingobacterium. *Biochim. Biophys. Acta* **1635**, 83–92 (2003).

629 16. Lorenzen, W., Bozhüyük, K. A. J., Cortina, N. S. & Bode, H. B. A comprehensive insight into the lipid composition

630   of Myxococcus xanthus by UPLC-ESI-MS. *J. Lipid Res.* **55**, 2620–2633 (2014).

631 17. Morii, H., Ogawa, M., Fukuda, K., Taniguchi, H. & Koga, Y. A revised biosynthetic pathway for

632   phosphatidylinositol in Mycobacteria. *J. Biochem.* **148**, 593–602 (2010).

633 18. Belcher Dufrisne, M. *et al.* Structural and Functional Characterization of Phosphatidylinositol-Phosphate

634   Biosynthesis in Mycobacteria. *J. Mol. Biol.* **432**, 5137–5151 (2020).

635 19. Harrison, P. J., Dunn, T. M. & Campopiano, D. J. Sphingolipid biosynthesis in man and microbes. *Nat. Prod. Rep.*

636   (2018) doi:10.1039/c8np00019k.

637 20. Koropatkin, N. M., Martens, E. C., Gordon, J. I. & Smith, T. J. Starch catabolism by a prominent human gut

638   symbiont is directed by the recognition of amylose helices. *Structure* **16**, 1105–1115 (2008).

639 21. Xu, J. *et al.* A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. *Science* **299**, 2074–2076

640   (2003).

641 22. Clarke, O. B. *et al.* Structural basis for phosphatidylinositol-phosphate biosynthesis. *Nat. Commun.* **6**, 8505

642   (2015).

643 23. Ryan, D., Jenniches, L., Reichardt, S., Barquist, L. & Westermann, A. J. A high-resolution transcriptome map

644   identifies small RNA regulation of metabolism in the gut microbe Bacteroides thetaiotaomicron. *Nat. Commun.*

645   **11**, 593 (2020).

646 24. Kolly, G. S. *et al.* GtrA Protein Rv3789 Is Required for Arabinosylation of Arabinogalactan in Mycobacterium

647   tuberculosis. *J. Bacteriol.* **197**, 3686–3697 (2015).

648 25. Fukao, M. *et al.* Plasmid-encoded glycosyltransferase operon is responsible for exopolysaccharide production,

649   cell aggregation, and bile resistance in a probiotic strain, Lactobacillus brevis KB290. *J. Biosci. Bioeng.* **128**, 391–

650   397 (2019).

651 26. Mayer, M. J., D'Amato, A., Colquhoun, I. J., Le Gall, G. & Narbad, A. Identification of Genes Required for Glucan

652 Exopolysaccharide Production in Lactobacillus johnsonii Suggests a Novel Biosynthesis Mechanism. *Appl.*

653 *Environ. Microbiol.* **86**, (2020).

654 27. Lim, B., Zimmermann, M., Barry, N. A. & Goodman, A. L. Engineered Regulatory Systems Modulate Gene

655 Expression of Human Commensals in the Gut. *Cell* **169**, 547–558.e15 (2017).

656 28. Barnett, J. E., Brice, R. E. & Corina, D. L. A colorimetric determination of inositol monophosphates as an assay

657 for D-glucose 6-phosphate-1L-myoinositol 1-phosphate cyclase. *Biochem. J* **119**, 183–186 (1970).

658 29. Chatterjee, A., Majee, M., Ghosh, S. & Majumder, A. L. sll1722, an unassigned open reading frame of

659 Synechocystis PCC 6803, codes for L-myo-inositol 1-phosphate synthase. *Planta* **218**, 989–998 (2004).

660 30. Donahue, T. F. & Henry, S. A. myo-Inositol-1-phosphate synthase. Characteristics of the enzyme and

661 identification of its structural gene in yeast. *J. Biol. Chem.* **256**, 7077–7085 (1981).

662 31. Chen, L., Zhou, C., Yang, H. & Roberts, M. F. Inositol-1-phosphate synthase from Archaeoglobus fulgidus is a

663 class II aldolase. *Biochemistry* **39**, 12415–12423 (2000).

664 32. Donahue, J. L. *et al.* The Arabidopsis thaliana Myo-inositol 1-phosphate synthase1 gene is required for Myo-

665 inositol synthesis and suppression of cell death. *Plant Cell* **22**, 888–903 (2010).

666 33. Norman, R. A. *et al.* Crystal structure of inositol 1-phosphate synthase from Mycobacterium tuberculosis, a key

667 enzyme in phosphatidylinositol synthesis. *Structure* **10**, 393–402 (2002).

668 34. Jin, X. & Geiger, J. H. Structures of NAD+- and NADH-bound 1-l-myo-inositol 1-phosphate synthase. *Acta*

669 *Crystallogr. D Biol. Crystallogr.* **59**, 1154–1164 (2003).

670 35. Stieglitz, K. A., Yang, H., Roberts, M. F. & Stec, B. Reaching for mechanistic consensus across life kingdoms:

671 structure and insights into catalysis of the myo-inositol-1-phosphate synthase (mIPS) from Archaeoglobus

672 fulgidus. *Biochemistry* **44**, 213–224 (2005).

673 36. Kudo, F., Tsunoda, T., Yamaguchi, K., Miyanaga, A. & Eguchi, T. Stereochemistry in the Reaction of the myo-

674 Inositol Phosphate Synthase Ortholog Ari2 during Aristeromycin Biosynthesis. *Biochemistry* **58**, 5112–5116

675 (2019).

676 37. Chhetri, D. R., Adhikari, J. & Mukherjee, A. K. NAD+ mediated differential thermotolerance between

677 chloroplastic and cytosolic L-myo-inositol-1-phosphate synthase from Diplopterygium glaucum (Thunb.) Nakai.

678 *Prep. Biochem. Biotechnol.* **36**, 307–319 (2006).

679 38. Gronnier, J., Germain, V., Gouguet, P., Cacas, J.-L. & Mongrand, S. GIPC: Glycosyl Inositol Phospho Ceramides, the major sphingolipids on earth. *Plant Signal. Behav.* **11**, e1152438 (2016).

681 39. Cerantola, V. *et al.* Aureobasidin A arrests growth of yeast cells through both ceramide intoxication and deprivation of essential inositolphosphorylceramides. *Mol. Microbiol.* **71**, 1523–1537 (2009).

683 40. Porter, N. T., Canales, P., Peterson, D. A. & Martens, E. C. A Subset of Polysaccharide Capsules in the Human Symbiont Bacteroides thetaiotaomicron Promote Increased Competitive Fitness in the Mouse Gut. *Cell Host Microbe* **22**, 494–506.e8 (2017).

686 41. Porter, N. T. *et al.* Phase-variable capsular polysaccharides and lipoproteins modify bacteriophage susceptibility in Bacteroides thetaiotaomicron. *Nature Microbiology* (2020) doi:10.1038/s41564-020-0746-5.

688 42. Stankeviciute, G., Guan, Z., Goldfine, H. & Klein, E. A. Caulobacter crescentus Adapts to Phosphate Starvation by Synthesizing Anionic Glycoglycerolipids and a Novel Glycosphingolipid. *MBio* **10**, (2019).

690 43. Stentz, R. *et al.* A bacterial homolog of a eukaryotic inositol phosphate signaling enzyme mediates cross-kingdom dialog in the mammalian gut. *Cell Rep.* **6**, 646–656 (2014).

692 44. Brito, J. A., Borges, N., Vonrhein, C., Santos, H. & Archer, M. Crystal structure of Archaeoglobus fulgidus CTP:inositol-1-phosphate cytidylyltransferase, a key enzyme for di-myo-inositol-phosphate synthesis in (hyper)thermophiles. *J. Bacteriol.* **193**, 2177–2185 (2011).

695 45. Huang, H. *et al.* Panoramic view of a superfamily of phosphatases through substrate profiling. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E1974–83 (2015).

697 46. Levine, T. P., Wiggins, C. A. & Munro, S. Inositol phosphorylceramide synthase is located in the Golgi apparatus of Saccharomyces cerevisiae. *Mol. Biol. Cell* **11**, 2267–2281 (2000).

699 47. Blunsom, N. J. & Cockcroft, S. Phosphatidylinositol synthesis at the endoplasmic reticulum. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1865**, 158471 (2020).

701 48. Kennedy, E. P. & Weiss, S. B. The function of cytidine coenzymes in the biosynthesis of phospholipides. *J. Biol. Chem.* **222**, 193–214 (1956).

703 49. Farine, L., Niemann, M., Schneider, A. & Bütikofer, P. Phosphatidylethanolamine and phosphatidylcholine biosynthesis by the Kennedy pathway occurs at different sites in Trypanosoma brucei. *Sci. Rep.* **5**, 16787 (2015).

705 50. Olsen, I. & Nichols, F. C. Are Sphingolipids and Serine Dipeptide Lipids Underestimated Virulence Factors of Porphyromonas gingivalis? *Infect. Immun.* **86**, (2018).

707 51. An, D. *et al.* Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells.

708   *Cell* **156**, 123–133 (2014).

709 52. Wieland Brown, L. C. *et al.* Production of α-galactosylceramide by a prominent member of the human gut

710   microbiota. *PLoS Biol.* **11**, e1001610 (2013).

711 53. Heaver, S. L., Johnson, E. L. & Ley, R. E. Sphingolipids in host-microbial interactions. *Curr. Opin. Microbiol.* **43**,

712   92–99 (2018).

713 54. Johnson, E. L. *et al.* Sphingolipids produced by gut bacteria enter host metabolic pathways impacting ceramide

714   levels. *Nat. Commun.* **11**, 2471 (2020).

715 55. Rocha, F. G. *et al.* Porphyromonas gingivalis Sphingolipid Synthesis Limits the Host Inflammatory Response. *J.*

716   *Dent. Res.* 22034520908784 (2020) doi:10.1177/0022034520908784.

717 56. Wu, S.-E. *et al.* Microbiota-derived metabolite promotes HDAC3 activity in the gut. *Nature* (2020)

718   doi:10.1038/s41586-020-2604-2.

719 57. Mücksch, F. *et al.* Quantification of phosphoinositides reveals strong enrichment of PIP2 in HIV-1 compared to

720   producer cell membranes. *Sci. Rep.* **9**, 17661 (2019).

721 58. Ledvina, H. E. *et al.* A Phosphatidylinositol 3-Kinase Effector Alters Phagosomal Maturation to Promote

722   Intracellular Growth of Francisella. *Cell Host Microbe* **24**, 285–295.e8 (2018).

723 59. Kraal, L., Abubucker, S., Kota, K., Fischbach, M. A. & Mitreva, M. The prevalence of species and strains in the

724   human microbiome: a resource for experimental efforts. *PLoS One* **9**, e97279 (2014).

725 60. Grāve, K., Bennett, M. D. & Högbom, M. Structure of Mycobacterium tuberculosis phosphatidylinositol

726   phosphate synthase reveals mechanism of substrate binding and metal catalysis. *Commun Biol* **2**, 175 (2019).

727 61. Folch, J., Lees, M. & Sloane Stanley, G. H. A simple method for the isolation and purification of total lipids

728   from animal tissues. *J. Biol. Chem.* **226**, 497–509 (1957).

729 62. Bui, H. H. *et al.* Direct analysis of PI(3,4,5)P3 using liquid chromatography electrospray ionization tandem mass

730   spectrometry. *Anal. Biochem.* **547**, 66–76 (2018).

731 63. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**,

732   186–190 (2009).

733 64. Winter, G. *et al.* DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr D Struct*

734   *Biol* **74**, 85–97 (2018).

735 65. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 282–292 (2011).

737 66. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

738 67. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).

740 68. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).

742 69. Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).

744 70. Krissinel, E., Uski, V., Lebedev, A., Winn, M. & Ballard, C. Distributed computing for macromolecular crystallography. *Acta Crystallogr D Struct Biol* **74**, 143–151 (2018).

746 71. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

748 72. Andrews, S. *et al.* FastQC. (2012).

749 73. Cunningham, F. *et al.* Ensembl 2019. *Nucleic Acids Res.* **47**, D745–D751 (2019).

750 74. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

751 75. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

753 76. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

755 77. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

757 78. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).

758 79. Parte, A. C. LPSN - List of Prokaryotic names with Standing in Nomenclature (bacterio.net), 20 years on. *Int. J. Syst. Evol. Microbiol.* **68**, 1825–1829 (2018).

760 80. Murat Eren, A. *et al.* Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**, e1319 (2015).

762 81. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

764 82. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.

765 *Bioinformatics* **30**, 1312–1313 (2014).

766

Supplementary Figures
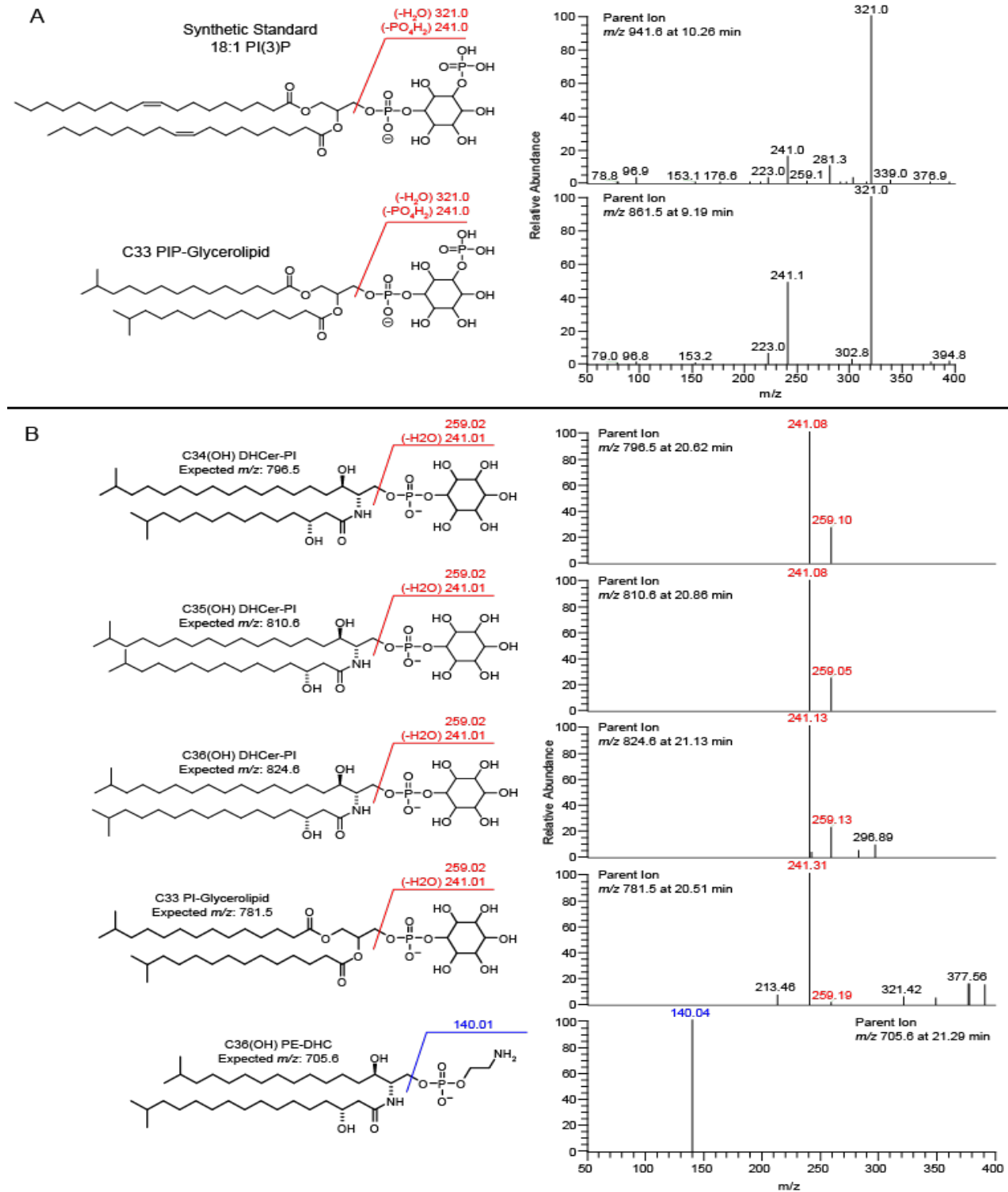
Supplementary Figure 1. Lipid structures and fragmentation patterns of BT inositol and ethanolamine lipids (A) Comparison of LC-MS/MS fragmentation patterns of BT-derived PIP with the synthetic standard, 18:1 PI(3)P. (B) LC-MS/MS fragmentation patterns of lipid structures present in iSPT BT at 100 ng/mL aTC induction, including PI-DHC lipids (C34(OH)DHCer-PI, C35(OH)DHCer-PI, C36(OH)DHCer-PI), C33 PI-glycerolipid, and C36(OH) PE-DHC. Loss of the phosphoinositol head group is indicated at mass 259. Fragments characteristic for lipids with phosphoinositol-based headgroups are in red; those for phosphoethanolamine-based headgroups are in blue.
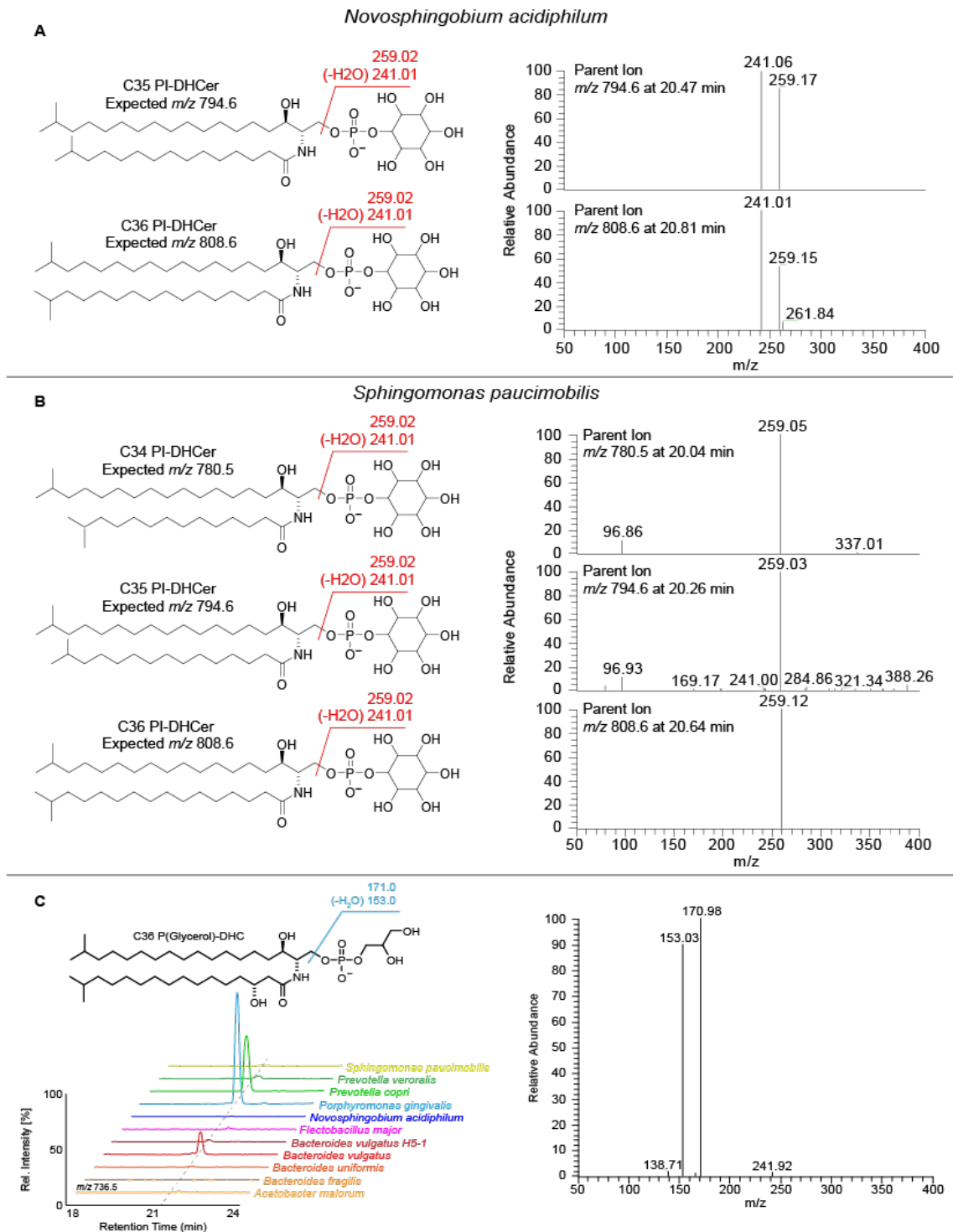
Supplementary Figure 2. Inositol-like lipid structures in diverse sphingolipid-producing species (A) LC-MS/MS fragmentation patterns of lipids extracted from *Novosphingobium acidiphilum* consistent with the synthesis of C35 and C36 PI-DHC. (B) LC-MS/MS fragmentation pattern of lipids extracted from *Sphingomonas paucimobilis*, demonstrating the presence of a headgroup with the same mass as inositol phosphate (259) but lacking the characteristic fragment of this group (241). (C) LC-MS/MS spectra and fragmentation pattern of a C36 P(Glycerol)-DHC structure present in *Prevotella copri*, *Porphyromonas gingivalis,* and *Bacteroides vulgatus.*
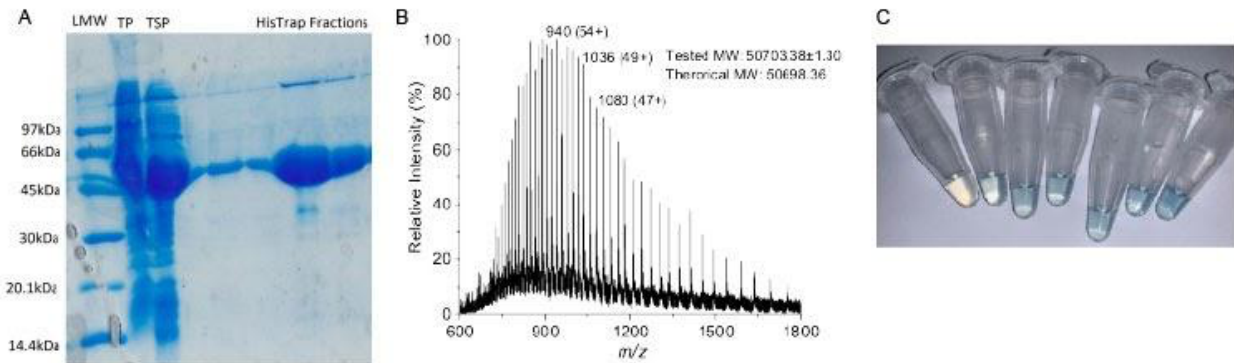
Supplementary Figure 3. Purification of BT_1526 (A) Expression and purification of recombinant BT_1526 MIPS purified from *E. coli*. The SDS-PAGE analysis shows the purity of the samples isolated by immobilised metal affinity chromatography (IMAC). The band between the 66k and 45 kDa corresponds to BT_1526 MIPS. (B) Electrospray ionisation mass spectrometry (ESI-MS) analysis of the purified BT_1526 MIPS. Positive mode ion envelope with charge states annotated on particular masses. The predicted molecular weight matches well with the theoretical mass without the initial Met residue. (C) Typical colour observed with the molybdenum blue assay of MIPS activity at 820 nm.
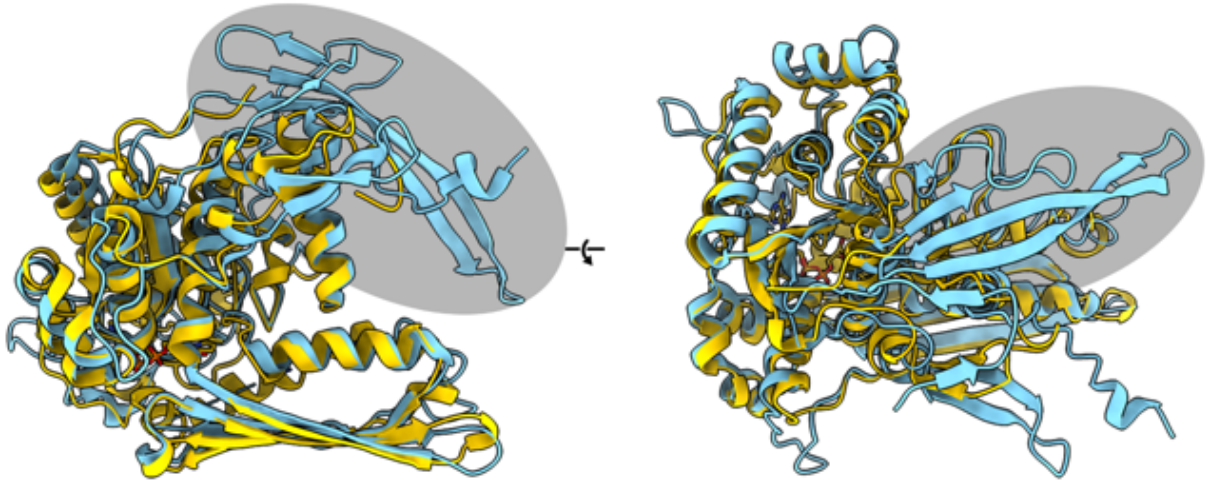
Supplementary Figure 4. Structural comparison of prokaryotic and eukaryotic MIPS proteins.
Secondary structure alignment of MIPS BT_1526 (in yellow) and *Saccharomyces cerevisiae* MIPS (PDBID: 1P1i) (in cyan). The N-terminal extension present in eukaryotic MIPS structures is highlighted with a grey oval.
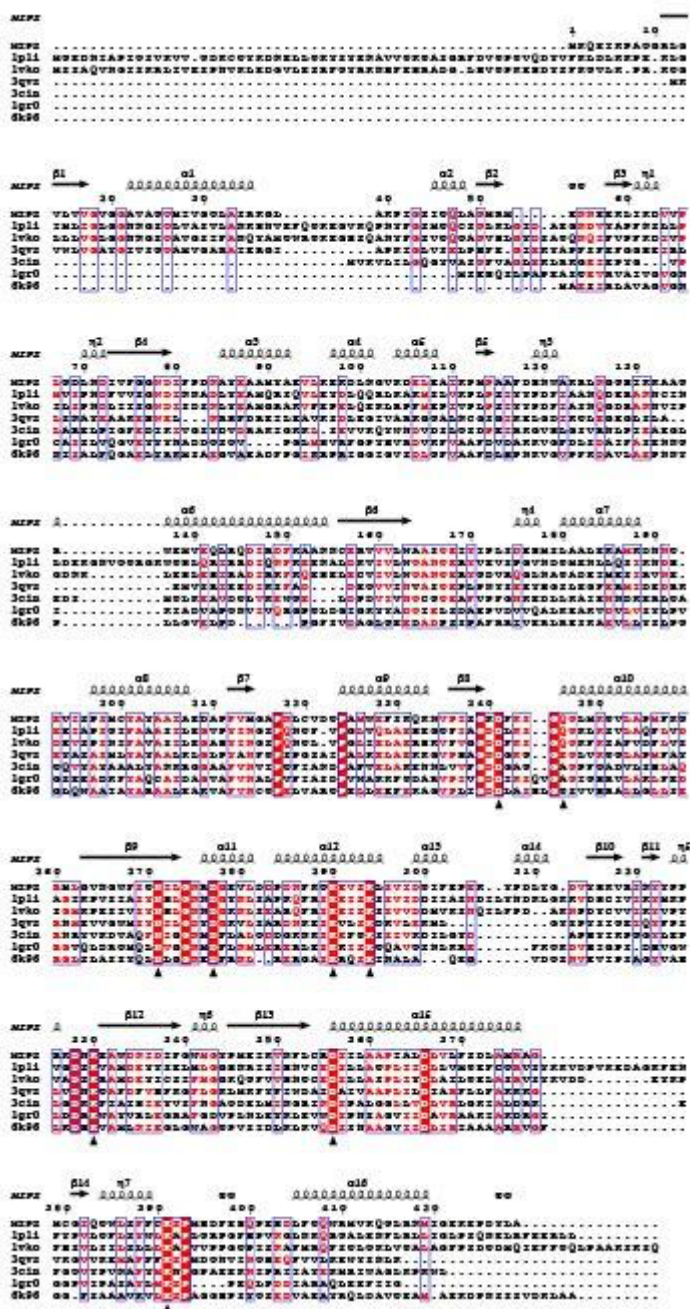
Supplementary Figure 5. MIPS structure and sequence comparison with active site residues. The sequences are: **MIPS (BT_1526), 1p1i** (*Saccharomyces cerevisiae*), **1vko** (*Caenorhabditis elegans*), **3qvs** (*Archaeoglobus fulgidus*), **3cin** (*Thermotoga maritima* MSB8), **1gr0** (*Mycobacterium tuberculosis*) and **6k96** (*Streptomyces citricolor* Ari2). Secondary structural elements are annotated, active site residues are marked with an arrow and conserved residues marked with red shading (fully conserved) or pink shading (similar).
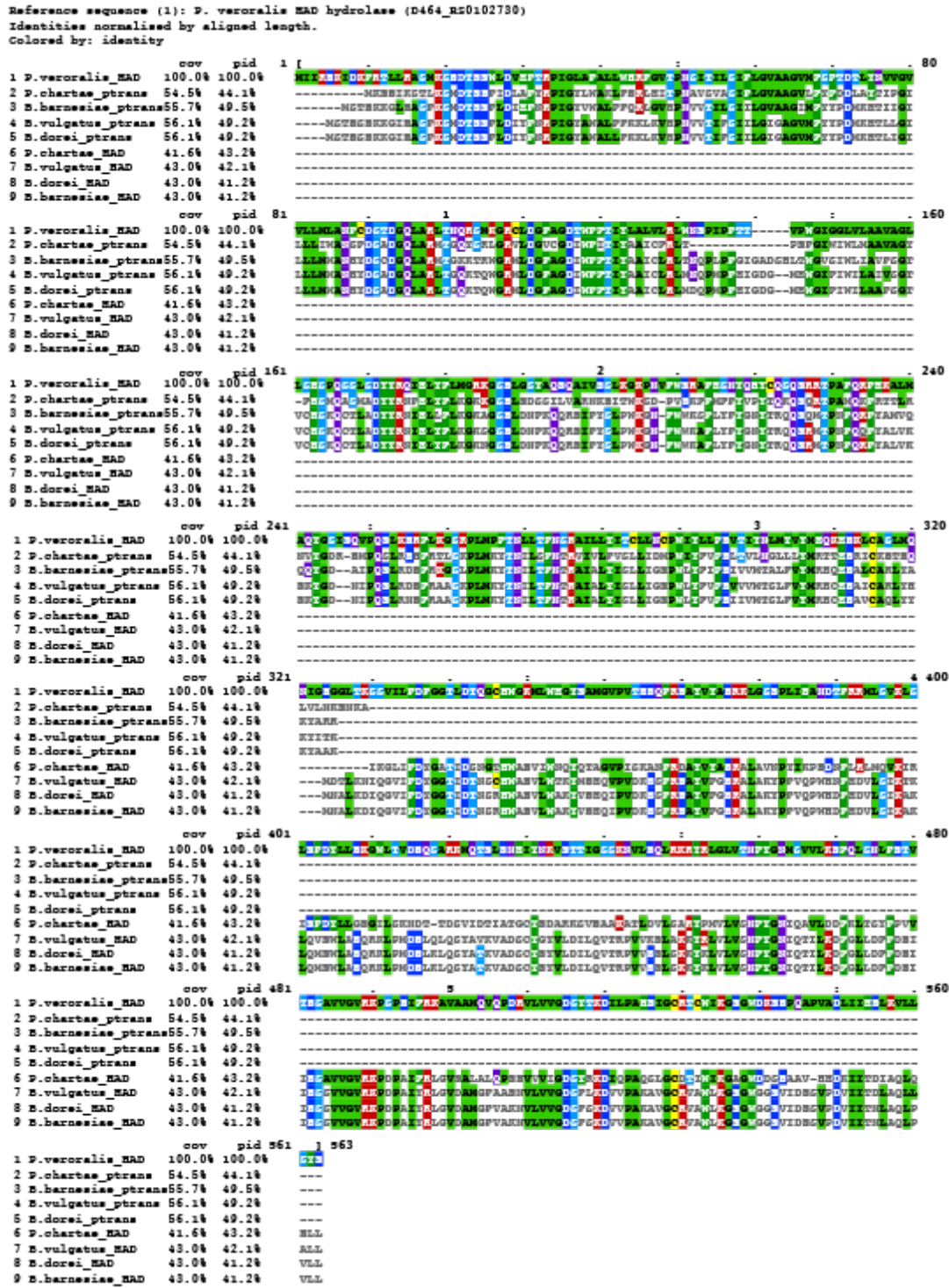
Supplementary Figure 6. Multiple sequence alignment of predicted proteins with putative involvement in an alternative inositol lipid metabolism cluster. Amino acid sequences from representative species containing the putative alternative inositol lipid metabolism cluster (*Parabacteroides chartae, Bacteroides barnesiae, Bacteroides vulgatus,* and *Bacteroides dorei*) HAD hydrolase and CDP-alcohol phosphatidyltransferase, aligned to the *Prevotella veroralis* predicted fusion protein with homology to both of these proteins (D464_RS0102730). Alignment was performed using Clustal Omega with visualization by MView.

Supplementary Tables

Supplementary Table 1: BT_1526 (MIPS) X-ray data collection and refinement statistics.

Supplementary Table 2: Top differentially expressed genes in wild-type *B. theta* compared to the WTΔBT_1522 strain in minimal medium with > 1.5 log2FC. Adjusted-P-value is Benjamini-Hochberg corrected.

Supplementary Table 3: Top differentially expressed genes in the iSPT strain compared to the ΔBT_1526 strain, both at 100 ng/mL aTC induction in minimal medium, with > 1.5 log2FC. Adjusted-P-value is Benjamini-Hochberg corrected. "CPS" column indicates the capsular polysaccharide synthesis (CPS) locus to which the gene belongs, when applicable.

Supplementary Table 4: E-values of BLAST-P homology to the *B. thetaiotaomicron* inositol lipid cluster, or the putative alternative pathway (using *B. vulgatus* sequences: phosphatidyltransferase BVU_RS13105, HAD hydrolase BVU_RS13115, NTP transferase BVU_RS13095).

Supplementary Table 5: Strains and plasmids used in this study; primers used in the amplification of genomic regions prior to plasmid assembly via restriction digest or Gibson cloning; TetR cassette components, insertion locations, primers, and gene fragment for the generation of the inducible SPT BT strain. Components and assembly were inspired by Lim et al. 2017.

Supplementary Table 6: Quantification of inositol in surface polysaccharides (raw values).

Supplementary Table 7: Kinetic data of BT_1526 MIPS (raw values).