

1 Functional metagenomics using *Pseudomonas putida* expands the known diversity of  
2 polyhydroxyalkanoate synthases and enables the production of novel polyhydroxyalkanoate  
3 copolymers

4

5 Jijun Cheng and Trevor C. Charles\*

6

7 Department of Biology and Centre for Bioengineering and Biotechnology, University of Waterloo,  
8 Waterloo, Ontario, Canada N2L 3G1

9

10 \*Corresponding author:

11 Dr. Trevor C. Charles

12 Department of Biology, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

13 E-mail: [tcharles@uwaterloo.ca](mailto:tcharles@uwaterloo.ca)

14

15

16

17 **Abstract**

18 Bacterially produced biodegradable polyhydroxyalkanoates with versatile properties can be achieved  
19 using different PHA synthase enzymes. This work aims to expand the diversity of known PHA  
20 synthases via functional metagenomics, and demonstrates the use of these novel enzymes in PHA  
21 production. Complementation of a PHA synthesis deficient *Pseudomonas putida* strain with a soil  
22 metagenomic cosmid library retrieved 27 clones expressing either Class I, Class II or unclassified  
23 PHA synthases, and many did not have close sequence matches to known PHA synthases. The  
24 composition of PHA produced by these clones was dependent on both the supplied growth substrates  
25 and the nature of the PHA synthase, with various combinations of SCL- and MCL-PHA. These data  
26 demonstrate the ability to isolate diverse genes for PHA synthesis by functional metagenomics, and  
27 their use for the production of a variety of PHA polymer and copolymer mixtures.

28 **Keywords:**

29 Functional metagenomics; soil metagenomic library; polyhydroxyalkanoate (PHA); bioplastics;  
30 polyhydroxyalkanoate synthase; *Pseudomonas putida*; *Sinorhizobium meliloti*

31

## 32 **Introduction**

33 Polyhydroxyalkanoates (PHAs) are natural polyesters biosynthesized by a variety of bacteria  
34 under unbalanced growth conditions. They serve as reserves of carbon and reducing power and aid  
35 in survival during starvation or stress conditions (Verlinden et al. 2007). These biodegradable and  
36 environmentally friendly polymers can be used as alternative materials to conventional  
37 petrochemical-based plastics. PHAs are classified into short-chain-length (SCL, C3-C5) and  
38 medium-chain-length (MCL,  $\geq$ C6), and copolymers (SCL + MCL) based on the number of carbon  
39 atoms per monomer. MCL-PHA is generally more useful than SCL-PHA due to it being less brittle  
40 and more flexible. To produce PHA with versatile properties cost-effectively, strategies have  
41 involved mining new PHA synthase enzymes (PhaC), engineering PhaC proteins and modifying the  
42 metabolic pathways of the production strains (Keshavarz and Roy 2010; Schallmeyer et al. 2011;  
43 Cheema et al. 2012; Park et al. 2012; Tripathi et al. 2013; Meng et al. 2014).

44 PHA molecules consist of over 150 possible constituent monomers of hydroxyl fatty acids  
45 (Meng et al. 2014) (Steinbüchel and Lütke-Eversloh 2003). PHA synthase enzymes catalyze the  
46 joining of the hydroxyl group of one monomer with the carboxyl group of another by an ester bond  
47 to form PHA polymers. The monomeric composition of PHA is determined primarily by PhaC,  
48 although the available carbon source and metabolic pathways also influence the properties of PHA  
49 (Verlinden et al. 2007). PhaC proteins are grouped into four classes based on amino acid sequence,  
50 substrate specificity and subunit composition (Rehm 2003). Class I and II PhaC consist of one  
51 subunit (PhaC), but those in Class III and IV are composed of two subunits (PhaC+PhaE and  
52 PhaC+PhaR, respectively). Class I and IV PhaCs synthesize SCL-PHA whereas Class II polymerize  
53 MCL-PHA. Class III PhaCs can synthesize both SCL and MCL monomers.

54           The key Class I PHA synthesis pathway genes include *phaC* (or *phbC*), *phaA* (or *phbA*) and  
55 *phaB* (or *phbB*). The three genes are often, but not always, clustered in a single operon. The  
56 acetoacetyl-CoA reductase (EC 2.3.1.9) encoded by the *phaA* and  $\beta$ -ketothiolase (EC 1.1.1.36)  
57 encoded by the *phaB* convert acetyl-CoA to (R)-3-hydroxybutyryl-CoA (3HB-CoA) through  
58 acetoacetyl-CoA. The Class-I PhaC (EC 2.3.1.-) then polymerizes 3HB-CoA into  
59 polyhydroxybutyrate (PHB) (Rehm 2003). The Class II *pha* cluster is well conserved in  
60 *Pseudomonas* and consists of two PHA synthase genes (*phaC1* and *phaC2*) flanking a PHA  
61 depolymerase gene (*phaZ*), and *phaD* encoding a transcriptional activator of *pha* genes. Phasin-  
62 encoding *phaF* and *phaI* gene are transcribed divergently to other *pha* genes. MCL monomers ((R)-  
63 3-hydroxylacyl-CoA) are derived from  $\beta$ -oxidation of fatty acids or fatty acid *de novo* synthesis from  
64 unrelated carbon sources (Tortajada et al. 2013).

65           Traditional culture-based strategies for obtaining new biocatalysts are limited by the inability  
66 to cultivate the majority of environmental microbes. While sequence-based metagenomics can  
67 identify genes homologous to those present in available sequence databases, it is difficult to reliably  
68 predict the function of truly new genes through homology-based analysis. In contrast, functional  
69 metagenomics involves the construction of gene libraries from microbial community genomic DNA  
70 and functional screening for novel enzymes of interest or potential industrial applications (Simon  
71 and Daniel 2011). Functional metagenomics has the potential to identify truly novel sequences for a  
72 given function.

73           In previous work, functional metagenomics was used to isolate new Class I PhaC from soil  
74 metagenomic clones by Nile red staining and phenotypic screening in  *$\alpha$ -Proteobacteria*  
75 *Sinorhizobium meliloti* (Schallmeyer et al. 2011). In another study, *phaC* genes encoding both Class I  
76 and II PhaC proteins were PCR amplified from oil-contaminated soil library clones (Cheema et al.

77 2012). One of the isolated genes produced PHA copolymer when expressed in *Pseudomonas putida*.  
78 Moreover, partial *phaC* genes were also obtained from metagenomic DNA via direct PCR  
79 amplification (Foong et al. 2014; Pärnänen et al. 2015; Tai et al. 2015). Building on these previous  
80 studies, we constructed a PHA<sup>-</sup> strain of *P. putida* KT2440 to use as a surrogate host and  
81 functionally identified *phaC* genes after screening millions of agricultural wheat soil metagenomic  
82 clones. A total of 27 PHA<sup>+</sup> clones were obtained. Accumulation and monomer composition of PHA  
83 directed by seven of these clones were further examined.

84

## 85 **Methods**

### 86 **Bacterial strains, plasmids, cosmids and growth conditions**

87 Bacterial strains, plasmids and cosmids are listed in Table 1. *E. coli* strains were grown at  
88 37°C in LB (Lennox) medium (1% tryptone (w/v), 0.5% yeast extract (w/v), and 0.5% NaCl (w/v),  
89 pH 7). *Pseudomonas* strains were grown at 30°C in LB or 0.1N M63 minimal medium (Escapa et al.  
90 2011) supplemented with 0.5% sodium octanoate (w/v), 0.5% nonanoic acid (v/v) or 1% gluconic  
91 acid (w/v). *S. meliloti* was grown in LB or YM medium (Schallmey et al. 2011). Antibiotics were  
92 used at the following concentrations: streptomycin, 200 µg/ml for *S. meliloti* and 100 µg/ml for *E.*  
93 *coli*; kanamycin, 100 µg/ml for *Pseudomonas* and 50 µg/ml for *E. coli*; neomycin, 200 µg/ml;  
94 rifampicin, 100 µg/ml; gentamicin, 10 µg/ml for *E. coli* and 100 for *P. putida*; and tetracycline 20  
95 µg/ml for *E. coli* or 40 µg/ml for *P. putida*.

96

### 97 **Construction of PHA<sup>-</sup> strain PpUW2**

98 DNA oligonucleotides are listed in Table 2. A Rif<sup>R</sup> spontaneous mutant PpUW1 of *P. putida*  
99 KT2440 was generated by plating a culture of strain KT2440 on a LB Rif plate, followed by single

100 colony purification. To construct the PHA deficient strain PpUW2 (*phaC1ZC2*) of *P. putida* PpUW1,  
101 a 943-bp DNA fragment containing 5'-*phaC1* region (766 bp) was PCR amplified using *P. putida*  
102 KT2440 genomic DNA as a template and primer pair JC161-JC162, digested with HindIII-BamHI  
103 and then cloned into the same sites in pJQ200-SK (Quandt and Hynes 1993), yielding plasmid  
104 pJC63. Another 876-bp DNA fragment containing the 3'-*phaC2* gene (120 bp) was PCR amplified  
105 using primers JC163-JC164, digested with BamHI-SalI and then cloned into the same sites in pJC63  
106 to obtain plasmid pJC69. An omega-Km cassette was obtained from pHP45Ω-Km (Fellay et al. 1987)  
107 by BamHI digestion and then inserted into the same site in pJC69 to obtain pJC71. Plasmid pJC71  
108 was then conjugated into *P. putida* PpUW1 in a triparental mating using helper plasmid pRK600.  
109 Single cross-over recombination of pJC71 into the *P. putida* chromosome was selected with Rif and  
110 Gm. A double cross-over event was achieved by growing a single Rif<sup>R</sup> Gm<sup>R</sup> colony overnight in LB,  
111 making serial dilutions and then spreading on LB Km with 5% sucrose. The resulting PHA<sup>-</sup> strain  
112 PpUW2 was verified by examining Gm sensitivity (lost plasmid backbone) and absence of PHA  
113 production in LB supplemented with 0.5% octanoate (w/v), and was further confirmed by PCR  
114 amplification analysis.

115

## 116 **Phenotypic screening for *phaC* genes from metagenomic library clones**

117 Construction of the metagenomic DNA library (11AW) of agricultural wheat soil was  
118 described previously (Cheng et al. 2014). The 11AW library contains  $9 \times 10^6$  clones hosted in *E. coli*  
119 HB101. The pooled library clones (0.5 ml of 300 ml stock) were conjugated *en masse* into *P. putida*  
120 PpUW2 (PHA<sup>-</sup>) with the helper plasmid pRK600. Mating mixture was diluted serially with 0.85%  
121 NaCl and ~20,000 transconjugants recovered on each LB Tc plate (15 mm × 150 mm) supplemented  
122 with 0.5% Na octanoate (w/v). The plates were incubated at 30°C for 24 h and then kept at 22°C for

123 2-6 days. Potential PHA<sup>+</sup> clones of opaque white colour were streak purified on the LB octanoate  
124 plates, and verified on 0.1N M63 minimal medium plates (Escapa et al. 2011) supplemented with  
125 0.5% octanoate (w/v) and Nile red (0.5 µg/ml). PHA<sup>+</sup> cosmids were transferred from *P. putida* to *E.*  
126 *coli* DH5α(Rif<sup>R</sup>), mobilized by pRK600. The cosmid DNA was isolated from *E. coli* strains using a  
127 GeneJET Plasmid Miniprep Kit (Thermo Scientific), digested by EcoRI-HindIII-BamHI (Thermo  
128 Scientific), and then resolved on 1% TAE agarose gels. Cosmids with distinct electrophoretic  
129 patterns were conjugated back to *P. putida* PpUW2 (PHA<sup>-</sup>) to confirm their PHA<sup>+</sup> phenotype.

130

### 131 **Complementation of *S. meliloti* (*phbC*)**

132 11AW cosmid DNA encoding Class I and II *phaC* genes was introduced into *S. meliloti*  
133 Rm11476 (Schallmeyer et al. 2011) via triparental conjugation. Transconjugants were selected on LB  
134 SmNmTc plates, and then streaked on YM plates containing Nile red (Schallmeyer et al. 2011) for  
135 visualizing PHB production.

136

### 137 **DNA sequencing and bioinformatics**

138 KOD Xtreme DNA polymerase (Novagen) was used for all PCR. Primers are listed in Table  
139 2. PCR reactions consisted of one cycle of 94°C for 5 min, 30 cycles of 94°C for 30 s, 53°C for 30 s  
140 and 68°C for 30 s, and final extension at 68°C for 10 min. The internal regions of *phaC* genes were  
141 PCR amplified from PHA<sup>+</sup> cosmids using primer phaCF1 and phaCR4 (Sheu et al. 2000). When no  
142 PCR product of correct size was obtained, a semi-nested PCR was performed with the primers  
143 phaCF2 and phaCR4 as described previously (Sheu et al. 2000). PCR products were resolved on 2%  
144 agarose gels, isolated from the gels using EZ-10 Spin Column DNA Gel Extraction Kit (Bio Basic),

145 and then cloned into pJET1.2 vector (Thermo Scientific). Sequences of the cloned partial *phaC*  
146 genes were obtained using pJET1.2 sequencing primers (Table 2).

147 End sequences of PHA<sup>+</sup> cosmid DNA were obtained by Sanger sequencing using universal  
148 M13F and M13R primers. For high throughput sequencing, 68 cosmids of 11AW PHA<sup>+</sup> clones and  
149 28 cosmids from other research projects were grouped into 24 pools. Tagmentation of pooled cosmid  
150 DNA, PCR amplification and clean-up, and library normalization were performed with the Nextera  
151 XT DNA library and index kits (Illumina), according to the supplier's recommendation. The library  
152 was sequenced using MiSeq 500-cycle version 2 reagent (Illumina). DNA sequences were  
153 assembled using SPAdes Genome Assembler 3.5.0 (BaseSpace, Illumina). Cosmid clones were  
154 identified based on the available Sanger end sequences (Lam et al. 2014) and ORFs were annotated  
155 with MetaGeneMark (Zhu et al. 2010). Predicted protein sequences were analyzed by BLAST  
156 against the non-redundant protein databases (NCBI). Multiple-sequence alignment was performed  
157 using MUSCLE (Edgar 2004). Phylogenetic analysis was performed using MEGA6 (Tamura et al.  
158 2013). Origin of cloned metagenomic DNA was predicted by PhyloPythia (Patil et al. 2012).

159

#### 160 **Cloning of *phaCI<sub>Pp</sub>* gene of *P. putida* KT2440**

161 To construct a *phaC* expression cosmid for a positive control, the *phaCI<sub>Pp</sub>* gene (*Pp\_5003*)  
162 of *P. putida* KT2440 was PCR amplified using primers JC159-JC160 and cloned into the HindIII  
163 and BamHI sites in the broad-host-range vector pRK7813 (Cheng et al. 2014), to obtain construct  
164 pJC67 (*Plac::phaCI<sub>Pp</sub>*). A stop codon (TGA) for terminating the translation of *lacZα* upstream of the  
165 *phaCI<sub>Pp</sub>*, and a ribosome-binding site (AGGAG) were incorporated into the primer JC159 (Table 2).

166

#### 167 **Subcloning *phaC* genes of 11AW metagenomic clone 16**

168 The Class II *phaC1*<sub>16</sub> gene of 11AW clone 16 was PCR amplified with oligos JC301 and  
169 JC302, and cloned into the HindIII-EcoRI sites in pRK7813 to obtain pJC123. The *phaC2*<sub>16</sub> was  
170 amplified using primers JC337 and JC334, and inserted into the EcoRI site in pJC123, yielding  
171 pJC157. The orientation of cloned *phaC2*<sub>16</sub> was verified by restriction enzyme mapping. The  
172 *phaC2*<sub>16</sub> gene was also obtained by PCR using oligos JC333 and JC334, and cloned into the HindIII-  
173 EcoRI sites in pRK7813 to yield pJC159. A stop codon (TGA) for terminating the translation of  
174 *lacZα* gene and ribosome-binding sites were added in primers JC333 and JC337. The cloned *phaC*  
175 genes were verified by DNA sequencing.

176

### 177 **Estimation of PHA production by fluorescent spectrometry**

178 A modification of the previously described method (Wu et al. 2003) was used. *S. meliloti*  
179 cells (2.5%, v/v) were added to YM medium. All cultures were incubated at 30°C and 200 rpm for  
180 48-72 h. The OD<sub>600</sub> values of 200 µl cultures were measured in 96-well microtiter plates with the  
181 Multiskan Spectrum spectrophotometer (Thermo Labsystems). Samples (180 µl) were stained with  
182 20 µl of Nile red (2.5 µg/ml) in the dark for 1 h. Fluorescent intensity was measured at excitation  
183 (485 nm) and emission (595 nm) with the FilterMax F5 Multi-mode microplate reader (Molecular  
184 Devices). PHA content was calculated based on the equation (fluorescent intensity/OD<sub>600</sub>×0.9).

185

### 186 **PHA production, extraction and characterization**

187 *P. putida* PpUW1 (PHA<sup>+</sup>) or PpUW2 (PHA<sup>-</sup>) carrying PHA<sup>+</sup> clones were grown overnight in  
188 LB with or without Tc, washed once with 0.85% NaCl, and then subcultured (2%, v/v) in the 0.1N  
189 M63 medium with or without Tc, supplemented with 0.5% Na octanoate, 0.5% nonanoic acid or 1%  
190 gluconic acid. The cultures were grown at 30°C and at 200 rpm for 48 hrs. Cells were collected by

191 centrifugation at 20°C and at 9000 × g for 15 min, washed once with deionized water. Cell dry  
192 weight (CDW) was obtained after drying the cells at 95°C for 24 hr. For PHA methanolysis, cell  
193 pellet (~15 mg) was suspended in 2 ml of chloroform and 2 ml of methanol with 15% H<sub>2</sub>SO<sub>4</sub> (v/v),  
194 and incubated at 100°C for 5 h. After the reaction mixture was cooled down to 20°C, 1 ml of  
195 deionized water was added, and vortexed vigorously for 1 min. The chloroform phase was passed  
196 through a cotton plug in a Pasteur glass pipette to remove any cell debris. Methanolized sample (1  
197 µl) was analyzed with GC-MS (an Agilent 7975B GC equipped with Agilent 5975B inert XL EI/CI  
198 MSD and an HP-5MS capillary column). Oven temperature was run at initial 50°C for 5 min with a  
199 ramp of 20°C/min to 280°C, and then held for 10 min. The flow rate of Helium carrier gas was 1.2  
200 ml/min. Methylated PHA monomers were identified using the Agilent enhanced MSD chemstation  
201 (E.02.01.117). PHA standards were kindly provided by Dr. Bruce A. Ramsay (Polyferm Canada).

202

## 203 **Results**

### 204 **Isolation of metagenomic clones for PHA production**

205 To employ *P. putida* as a surrogate host to isolate 11AW (agricultural wheat soil)  
206 metagenomic library clones encoding functional PhaC, it was first necessary to construct a PHA  
207 synthesis mutant. Expression of the contiguous three genes encoding PHA synthase PhaC1<sub>Pp</sub>  
208 (Pp\_5003), PHA depolymerase PhaZ<sub>Pp</sub> (Pp\_5004) and PHA synthase PhaC2<sub>Pp</sub> (Pp\_5005) in *P.*  
209 *putida* KT2440 Rif<sup>R</sup> derivative PpUW1 was disrupted by deletion of 3437-bp comprised of the 3'  
210 region (911 bp) of *phaC1<sub>Pp</sub>* (1677 bp), *phaZ<sub>Pp</sub>* and the 5' region (1560 bp) of *phaC2<sub>Pp</sub>* (1680 bp) and  
211 replacement with an omega-Km kanamycin resistance insert, resulting in the PHA<sup>-</sup> strain PpUW2.  
212 Transfer of the negative control cosmid vector pJC8 to *P. putida* PpUW2 did not result in detectable  
213 PHA production. However, introduction of pJC67 (*Plac::phaC1<sub>Pp</sub>*) *in trans* restored the PHA<sup>+</sup>

214 phenotype (Fig. 1). These data suggested that the strain PpUW2 could be used for screening of  
215 PhaC-encoding metagenomic clones. The 11AW library clones (Cheng et al. 2014) were transferred  
216 to *P. putida* PpUW2 via en masse triparental conjugation. Following selection of *P. putida* PpUW2  
217 transconjugants (~4 million) on LB Km Tc plates supplemented with 0.5% Na octanoate, we  
218 obtained 72 clones that exhibited greater opacity than the PpUW2 PHA<sup>-</sup> recipient strain. The PHA-  
219 producing phenotype of those clones was verified by visualizing the fluorescence of Nile red-stained  
220 PHA in 0.1N M63 minimal medium (Escapa et al. 2011) supplemented with Na octanoate as the sole  
221 carbon source (Fig. 1). Restriction digest of DNA from the 72 clones demonstrated 68 distinct  
222 restriction patterns, suggesting the presence of a broad diversity of DNA origin in those PHA<sup>+</sup> clones.  
223

#### 224 **Identification of *phaC* genes encoding PHA synthases**

225 Internal regions (~500 bp) of the *phaC* genes of 18 distinct cosmid clones were initially  
226 obtained by PCR amplification using degenerate primers PhaCF1 and PhaCR4 (Sheu et al. 2000).  
227 For additional 4 clones, *phaC* fragments (PhaC<sub>3</sub>, PhaC<sub>7</sub>, PhaC<sub>10</sub>, and PhaC<sub>15</sub>) of ~400 bp were  
228 generated by nested-PCR using primer pair PhaCF2-PhaCR4 as described previously (Sheu et al.  
229 2000). BLASTP analysis of the cloned gene products indicated that 13 PhaC proteins could be  
230 grouped into Class I while the other 9 were categorized as Class II PHA synthases.

231 To identify all *pha* genes on the isolated clones, DNA sequences were obtained by high  
232 throughput sequencing, and additional *pha* gene loci were identified in partially and fully assembled  
233 clones (Table 3; Supplementary 1). The identified PHA synthases could be classified into 3 groups  
234 based on amino acid sequences (Fig. 2). Class I *phaC* genes were annotated in 17 clones. The *phaC*  
235 and *phaAB* genes were adjacently located in 12 of these clones, whereas the *phaB* gene was located  
236 distantly downstream of *phaCA* genes in clone 25 and clone P1N3 (Supplementary 1). A *phaR* gene

237 was located immediately downstream of *phaCB* genes in partially assembled clone P11N2. The  
238 *phaC* and *phaB* genes flanked a *phaZ* gene in the partially assembled clone P2N8. The metagenomic  
239 DNA in these Class I clones was predicted to originate from *Gemmatimonas*,  $\alpha$ -*Proteobacteria*  
240 (*Sphingomonadaceae*),  $\beta$ -*Proteobacteria* (*Leptothrix*, *Rubrivivax*, *Janthinobacterium*) and  $\gamma$ -  
241 *Proteobacteria* (*Xanthomonadaceae*) (Table 3).

242 Class II PHA genes are commonly clustered with *phaZ* flanked by two PHA synthase  
243 encoding genes, *phaC1* and *phaC2*. Nine clones carried Class II *phaC* genes (Table 3). The *pha* gene  
244 locus (*phaC1<sub>1</sub>-phaZ-phaC2<sub>1</sub>-phaD-phaF-phaI*) in clone 1 was similar to the canonical locus in  
245 *Pseudomonas* including *P. putida* KT2440 (de Eugenio et al. 2010), except the presence of *orf7*  
246 encoding a PHA granule associated protein (Pfam09650) (Fig. 3A). Clone 2 and clone 15 were  
247 identical to the clone 1, except that they contained an insertion of the 17,013-bp transposon Tn4652  
248 at 13,740 nt and 16,362 nt respectively (Supplementary 2). The Tn4652 duplicated target sequences  
249 were AACTC in clone 2 and TAGGA in clone 15. The transposon insertions in these clones are in  
250 opposite orientation. The same Tn4652 is present in *P. putida* KT2440 genome (3,366,550 -  
251 3,383,562 nt, GenBank: AE015451), located distant from the *pha* gene operon. These insertions  
252 likely occurred following introduction of the cosmid clones into *P. putida*.

253 Class II clones 5, 16, 20 and 50 had no genes encoding *phaD* or *phaFI* homologs. A *phaJ*  
254 gene encoding R-specific enoyl-CoA hydratase was identified in those clones (Fig. 3A;  
255 Supplementary 1). The *phaJ* gene was located immediately upstream of the *phaC* genes in clones 16,  
256 20 and 50, but downstream of *phaC* in clone 5. In addition, a gene encoding PHA depolymerase  
257 (PhaZ) was located downstream of the *phaC* genes in clones 16, 20 and 50. ORFs downstream of the  
258 *phaZ* genes in clones 16 and 50 encoded proteins (PhaC2<sub>16</sub> and PhaC2<sub>50</sub>) homologous to unclassified

259 PHA synthases (Fig. 2; Supplementary 1). The proteins shared conserved regions in PhaC enzymes,  
260 but were ~30% shorter in N-terminal regions than in those of Class I/II PhaC (Supplementary 2).

261 Two ORFs encoding putative PhaC and PhaE proteins were identified in partially assembled  
262 clone P2N10 (Table 3; Supplementary 1). BLASTP showed the P2N10 PhaC protein best matched  
263 to the Class III PhaC [GenBank: WP\_002708071] (43% identity) from *γ-Proteobacteria Thiiothrix*  
264 *nivea* (Table 3). The available C-terminal sequence of P2N10 PhaE (234 amino acids) only exhibited  
265 26% identity to the Class III PhaE [GenBank: WP\_002708072] of *T. nivea*. In addition, the P2N10  
266 PhaC protein was only 29% identical to the Class IV PhaC [GenBank: AAD05620] of *Bacillus*  
267 *megaterium*. Phylogenetic analysis suggested that P2N10 PHA synthase was distant from Class  
268 III/IV PhaC and might represent a new subclass of PHA synthases (Fig. 2).

269 Multiple alignment of amino acid sequences of each of the identified 11AW PhaC proteins  
270 showed the conserved catalytic triad (C287, D445 and H473 in PhaC<sub>116</sub>), a tryptophan essential for  
271 dimerization (W391 in PhaC<sub>116</sub>), and the lipase box GXCXGG (Jia et al. 2000) except that Ala  
272 replaced the first Gly residue in PhaC<sub>20</sub> and second Gly in clone P8N7 (Fig. 3B; Supplementary 3).  
273 The serine residue (S229 in PhaC<sub>116</sub>) required for PhaC activity (Hoppensack et al. 1999) was also  
274 conserved in all 11AW PhaC proteins. We chose the PhaC proteins in the Class I clones 14, 18 and  
275 25, and the Class II clones 1, 16 and 20, for further study. We examined the nature of PHA produced  
276 using different carbon sources that result in substrate production either through fatty acid synthesis  
277 or  $\beta$ -oxidation.

278

### 279 **Class II clone 1 synthesizes MCL PHA**

280 The cloned metagenomic DNA in clone 1 contained 42,747 bp (Table 3) [GenBank:  
281 KT944254] and most likely originated from *γ-Proteobacteria Pseudomonas*. The PhaC<sub>11</sub>, PhaZ,

282 PhaC<sub>2</sub><sub>1</sub> and PhaD proteins [GenBank: ALV86243, ALV86281, ALV86274 and ALV86275] were  
283 82%, 92%, 73% and 78% identical to the corresponding orthologs (PhaC1, PhaZ, PhaC2 and PhaD)  
284 of *P. putida* KT2440. Clone 1 PhaC<sub>2</sub><sub>1</sub> was phylogenetically related to the PHA synthase 2  
285 [GenBank: BAA36202] of *Pseudomonas* sp. 61-3 (Fig. 2). In contrast to other Class II PhaC, the  
286 conserved amino acid at position 129 of PhaC<sub>2</sub><sub>1</sub> was Asp rather than Glu (Fig. 3B). It has been  
287 previously demonstrated that substitution of the conserved Glu with Asp in Class II PhaC improves  
288 PHA yield with an increase of 3-hydroxybutyrate monomer (Matsumoto et al. 2005).

289 PHA synthesis in clone 1 was likely contributed solely by PhaC<sub>2</sub><sub>1</sub> because the 5'-region of  
290 *phaC1<sub>1</sub> was absent in the cloned DNA. Expression of the functional *phaZ* and *phaC2<sub>1</sub> might be  
291 driven by the promoters upstream of the individual genes, as occurs in *P. putida* KT2440 (de  
292 Eugenio et al. 2010). *P. putida* PpUW2 (PHA<sup>-</sup>) carrying clone 1 synthesized 3-hydroxyhexanoate  
293 (3HHx, C6) and 3-hydroxyoctanoate (3HO, C8) copolymer of ~95% C8 monomer when grown with  
294 octanoate, similar to the PHA produced by wild-type *P. putida* PpUW1 (Table 4). When PpUW2  
295 (clone 1) was grown with nonanoic acid, 3-hydroxynonanoate (3HN, C9) and 3-hydroxyheptanoate  
296 (3HP, C7) were incorporated into the PHA with greater 3HP than that synthesized by wild type  
297 PpUW1 (Table 4). These data suggest that those monomers were derived from  $\beta$ -oxidation of  
298 octanoate and nonanoic acid.**

299 When gluconic acid was used as the sole carbon source, 3HO and 3-hydroxydecanoate  
300 (3HD, C10) copolymer was produced in PpUW2 (PHA<sup>-</sup>) carrying clone 1, very similar to the  
301 production of the parental strain PpUW1 (PHA<sup>+</sup>) (Table 4). The monomer composition of PHA was  
302 similar in both strains, but total PHA accumulation was ~4 fold higher in gluconic acid-grown  
303 PpUW2 with clone 1 (Table 4). These results suggest that Class II PhaC<sub>1</sub><sub>2</sub> was able to synthesize  
304 MCL (C6-C10) PHA.

305

306 **Class II clone 16 synthesizes SCL-MCL copolymer**

307 Clone 16 contains 33,900-bp metagenomic DNA [GenBank: KT944263] probably originated  
308 from *β-Proteobacteria Methylibium* (Table 3). A *phaJ* gene (*phaJ<sub>16</sub>*) is predicted to encode a R-  
309 specific enoyl-CoA hydratase [GenBank: ALV86416] (Fig. 3A), which converts the  $\beta$ -oxidation  
310 intermediate 2-enoyl-CoA to (R)-3-hydroxyacyl-CoA used as PhaC substrate (Fukui et al. 1998).  
311 The PhaJ<sub>16</sub> was phylogenetically related to the PhaJ4a<sub>Re</sub> and PhaJ4b<sub>Re</sub> of *Ralstonia eutropha*  
312 (Kawashima et al. 2012), and PhaJ4<sub>Pp</sub> of *P. putida* (Sato et al. 2011) (Supplementary 4A). The  
313 amino acid residues Asp39 and His44 conserved at the active site of the dehydratases were identified  
314 in PhaJ<sub>16</sub> (Supplementary 3B), except that the Ser residue was replaced by Pro71 in PhaJ<sub>16</sub>, same as  
315 in the PhaJ orthologs of *R. eutropha* and *P. putida* PhaJ4.

316 The three ORFs downstream of the *phaJ<sub>16</sub>* gene encoded PhaC1<sub>16</sub>, PhaZ and PhaC2<sub>16</sub> (Fig.  
317 3A). PhaC1<sub>16</sub> of Class II PHA synthase most closely matched the *Pseudomonas* sp. F15 PhaC1  
318 [GenBank: WP\_021487788], exhibiting 48% amino acid identity. PhaC1<sub>16</sub> was also 42.5% and  
319 43.4% identical to the PhaC1 and PhaC2 proteins of *P. putida* KT2440. Thr318 and Val478 in the  
320 PhaC1<sub>16</sub> might favour the 3HB incorporation into PHA as occurred in the engineered PhaC proteins  
321 (Takase et al. 2003; Chen et al. 2014). Another putative PHA synthase encoded by the *phaC2<sub>16</sub>* gene  
322 downstream of the *phaC1<sub>16</sub>phaZ* genes was homologous to unclassified PHA synthases (Fig. 2).

323 The quantity of PHA accumulated in *P. putida* UW2 carrying clone 16 was comparable to that  
324 of parental PpUW1 when grown with octanoate, nonanoic acid or gluconic acid (Table 4). In  
325 contrast to the MCL PHA (C6-C8) present in octanoate-grown PpUW2 (clone 1) or wild-type  
326 PpUW1, about 20% 3HB was detected in the PHA (C4-C6-C8) accumulated in PpUW2 carrying  
327 clone 16 under the same conditions (Table 4). When PpUW2 (clone 16) was grown with nonanoic

328 acid (C9), 3-hydroxyheptanoate (3HP, C7), 3-hydroxyvalerate (3HV, C5) and 3HB (C4) were  
329 polymerized, which was in contrast to the 3HP-3HN PHA in nonanoic acid-grown PpUW with  
330 PhaC1 or wild type PpUW1. Additionally, equal amounts of C6 and C8 monomers were detected in  
331 the PHA from gluconic acid-grown PpUW2 with clone 16. However, PHA with C8 and C10  
332 monomers was accumulated in PpUW2 (clone 1) and PpUW1 under the same growth conditions.  
333 These data implied that the PhaC<sub>16</sub> and/or PhaC<sub>26</sub> prefer substrates with shorter carbon chains than  
334 the PhaC<sub>21</sub> in clone 1 and both PhaC proteins of *P. putida* KT2440.

335 To further elucidate the activity of PhaC<sub>16</sub> and PhaC<sub>26</sub>, the genes were cloned  
336 downstream of the constitutively active *Plac* promoter. When *P. putida* PpUW2 carrying pJC123  
337 (*Plac::phaC16*) was grown with octanoate, SCL-MCL copolymer was of similar composition of C4-  
338 C6-C8 monomers to that synthesized by clone 16 (Table 4). However, PHA was not detected in  
339 strain PpUW2 carrying pJC159 (*Plac::phaC26*) under the same growth conditions. Coexpression of  
340 *phaC16* and *phaC26* in pJC157 (*Plac::phaC16-phaC26*) in octanoate-grown PpUW2 resulted in  
341 synthesis of PHA with the similar monomer composition as those in the polymers produced in  
342 PpUW2 carrying clone 16 or pJC123 (*Plac::phaC16*) (Table 4). These data indicated that only the  
343 PhaC<sub>16</sub> was involved in SCL-MCL PHA biosynthesis in clone 16.

344 The *pha* gene locus of clone 16 was very similar to that of clone 50 (Supplementary 1), but  
345 the DNA in clone 50 probably originated from *γ-Proteobacteria Hydrocarboniphata* (Table 3). Both  
346 the Class II PhaC1 and unclassified PhaC2 in clones 16 and 50 shared 84% identity respectively.  
347 Expression of the *pha* genes was probably driven by a promoter upstream of *phaJ* genes.

348

349 **Class II clone 20 synthesizes MCL or SCL-MCL PHA**

350 PHA<sup>+</sup> clone 20 contained 30,964-bp metagenomic DNA [GenBank: KT944266], which  
351 probably originated from *Salinibacter* within the *Bacteroidetes* (Table 3). The cloned DNA  
352 contained only a single *phaC*<sub>20</sub> gene (Fig. 3A). A putative fatty acid CoA ligase/ synthetase  
353 [GenBank: ALV86493] (ORF9; Supplementary 1) might be involved in degradation of fatty acids to  
354 acyl-CoA (the first step of  $\beta$ -oxidation cycle). The acyl dehydratase PhaJ<sub>20</sub> [GenBank: ALV86492]  
355 was 56% identical to the PhaJ<sub>16</sub>. Conserved amino acid residues Asp41 and His46 at the active sites  
356 of the acyl dehydratases were present (Supplementary 4B) except that the Ser residue was replaced  
357 by Pro73 in the PhaJ<sub>20</sub> protein. The *phaJ*<sub>20</sub>, *phaC*<sub>20</sub> and *phaZ* genes in clone 20 were similar to those  
358 of the corresponding genes in clones 16 and 50 (Fig. 3A; Supplementary 1). In addition, the *pha*  
359 gene cluster of clone 20 was similar to that in *Burkholderia* sp. 383 (1,576,365 -1,581,635 nt)  
360 [GenBank: NC\_007511], but a gene encoding a phasin protein is located between the *phaC* and  
361 *phaZ* genes in the *Burkholderia* genome.

362 The single PhaC<sub>20</sub> [GenBank: ALV86493] best matched to its ortholog from *Burkholderia lata*  
363 at 59% amino acid sequence identity [GenBank: WP\_011354994]. PhaC<sub>20</sub> was 46-53% identical to  
364 the PhaC<sub>16</sub>, PhaC<sub>2</sub> and PhaC<sub>5</sub> as well as PhaC<sub>1</sub> and PhaC<sub>2</sub> of *P. putida* KT2440. Presence of amino  
365 acid residues Thr353 and Val516 in PhaC<sub>20</sub> (Fig. 3B) might favour substrates with fewer carbons  
366 than those preferred by clone 1 and *P. putida* PhaC proteins, as previously demonstrated with  
367 engineered Class II PHA synthases (Takase et al. 2003; Chen et al. 2014).

368 Clone 20 was able to synthesize 3HHx-3HO copolymer (C6-C8) with octanoate (Table 4), as  
369 occurred in PpUW2 (clone 1) and *P. putida* wild type. However, the proportion of C6 monomer  
370 increased 4-35 fold in the PHA isolated from PpUW2 (clone 20), compared to that in PpUW2 (clone  
371 1) and wild-type PpUW1 (Table 4). When nonanoic acid was supplied, PpUW2 (clone 20) produced  
372 PHA with 8% 3-hydroxyvalerate (3HV), which was absent in the PHA synthesized by clone 1 and

373 PpUW1 (Table 4). In contrast to the SCL-MCL PHA produced by clone 16, 3HB was not detected in  
374 the nonanoic acid or octanoate-grown PpUW2 with clone 20.

375 In contrast to the absence of detectable C6 monomer in the PHA produced in PpUW2 with  
376 clone 1 or *P. putida* wild type, 3HHx (~10%) was incorporated into the MCL PHA (C6-C8-C10) in  
377 PpUW2 (clone 20) (Table 4) when gluconic acid was used as the sole carbon source. About half the  
378 quantity of monomers was C10 in the PHA accumulated in PpUW2 (clone 20) grown with gluconic  
379 acid. These results suggested that the carbon chain lengths of favourite monomers of PhaC<sub>20</sub> were  
380 between those of PhaC<sub>16</sub> and PhaC<sub>21</sub>.

381

### 382 **Class I clone 14 synthesizes SCL-MCL PHA**

383 The DNA in clone 14 probably originated from *β-Proteobacteria Methylibium* (Table 3). The  
384 PhaC<sub>14</sub>, PhaA and PhaB [GenBank: ALV86397, ALV86398 and ALV86399] were 65%, 76% and  
385 74% identical to the PhbC1 (A16\_1437), PhbA (A16\_1438) and PhbB1 (A16\_1439) of *C. necator*  
386 H16, respectively. In addition, PhaC<sub>14</sub> was 48% identical to the PhbC protein (Smc002960) of *S.*  
387 *meliloti* Rm1021. The best match of PhaC<sub>14</sub> was the PhaC protein of *Burkholderia* JOSHI\_001 at  
388 74% (Table 3). The *phaC<sub>14</sub>AB* locus of clone 14 is similar to that of clones 8, 21 and 23  
389 (Supplementary 1). Expression of the *phaC<sub>14</sub>*, *phaA* and *phaB* genes in clone 14 was most likely  
390 driven by a promoter upstream of the *orf9* encoding a hypothetical protein.

391 When PpUW2 (clone 14) was grown with nonanoic acid, 7% 3HB was incorporated into  
392 3HB-3HV copolymer (Table 4). 3HB was the primary monomer of PHA in octanoate- or gluconic  
393 acid-grown PpUW2. Both C6 and C8 monomer were present in the PHA with gluconic acid as  
394 carbon source. However, C8 monomer was absent in the PHA when octanoate was supplied (Table  
395 4). *P. putida* PpUW2 carrying clone 14 produced the highest quantity of PHA among all the PHA<sup>+</sup>

396 clones (Table 4). These results indicated that PhaC<sub>14</sub> could synthesize SCL-MCL copolymer though  
397 3HB was the dominant monomer.

398

### 399 **Class I clone 18 synthesizes SCL and SCL-MCL PHA**

400 Clone 18 contained 37,818-bp metagenomic DNA [GenBank: KT944264], which probably  
401 originated from *β-Proteobacteria Rubrivirax* (Table 3). The annotated *phaC<sub>18</sub>*, *phaA* and *phaB* were  
402 organized in one operon (Fig. 3A), as also occurred in clones 27 and 51 (Supplementary 1). PhaC<sub>18</sub>  
403 [GenBank: ALV86462] shared 100% and 99% identity to PhaC<sub>27</sub> [GenBank: ALV86602] and  
404 PhaC<sub>51</sub> [GenBank: ALV86651] respectively. The PhaC<sub>18</sub> protein best matched to its homolog in  
405 *Azohydromonas australica* (Table 3), and was 71%, 49% and 61% identical to the PhaC<sub>14</sub>, PhbC<sub>Sm</sub>  
406 and PhbC<sub>1Re</sub> respectively.

407 SCL-MCL PHA (C4-C6-C8) was produced when PpUW2 (clone 18) was grown with  
408 octanoate (Table 4), as occurred in PpUW2 with clone 16. However, the fraction of C6-C8  
409 monomers was only 3% in the PHA from clone 18, in contrast to the ~83% in the copolymer from  
410 clone 16. The yield of PHA was ~7 fold less than that produced by clone 14, but C8 monomer was  
411 absent (Table 4). PHA of 3HB-3HV monomers was accumulated in nonanoic acid-grown PpUW2  
412 (clone 18). The ratio of 3HB to 3HV increased 5 fold compared to that in PpUW2 (clone 14) under  
413 the same growth conditions (Table 4). When PpUW2 carrying clone 18 was grown on gluconic acid,  
414 only PHB was produced. The yield of polymer increased 5 fold compared to that in octanoate-grown  
415 cells, and also the second highest quantity of PHA produced among all strains grown on gluconic  
416 acid. These results implied that PhaC<sub>18</sub> could synthesize SCL- and SCL-MCL PHA dependent of the  
417 available carbon source.

418

419 **Class I clone 19 synthesizes SCL and SCL-MCL PHA**

420 Clone 19 contained 25,085-bp of DNA that might originate from *β-Protoebacteria Leptothrix*  
421 (Table 3) [GenBank: KT944265]. Three ORFs encoded PhaC<sub>19</sub>, PhaA and PhaB for PHA  
422 biosynthesis (Fig. 3A). PhaC<sub>19</sub> shared 99% identity with PhaC<sub>6</sub> and PhaC<sub>3</sub> (Fig. 2). The best match  
423 for PhaC<sub>19</sub> is the ortholog of *A. australica* at 73% (Table 3). The *phaC<sub>19</sub>AB* genes in clone 19  
424 probably consisted of one operon, and the promoter was located upstream of the *phaC<sub>19</sub>*.

425 Clone 19 was able to synthesize SCL-MCL PHA (C4-C6-C8) with octanoate but only PHB  
426 (C4) in gluconic acid-grown PpUW2 (Table 4), as occurred in PpUW2 carrying clone 18. However  
427 ~5 fold more PHA was synthesized by clone 19 than was produced by clone 18 in octanoate medium  
428 (Table 4). In contrast, the quantity of PHA was lower in gluconic acid-grown PpUW2 with clone 19  
429 than it was in the same host carrying clone 18. Moreover, PHA composed of 3HB-3HV monomers  
430 was synthesized in noanoic acid-grown PpUW2 (clone 19), as occurred in PpUW2 (clone 18). These  
431 data suggested that PhaC<sub>19</sub> was able to synthesize both SCL and SCL-MCL PHA.

432

433 **Class I clone 25 synthesizes SCL and SCL-MCL PHA**

434 Clone 25 contained 41,012-bp of DNA [GenBank: KT944269], which was predicted to  
435 originate from *β-Proteobacteria Variovorax* (Table 3). The *phaC<sub>25</sub>* and *phaA* genes were apparently  
436 located in a single operon (Fig. 3A). A *phaB* gene was located 9 open reading frames downstream of  
437 the *phaA*. The *phaC<sub>25A</sub>* and *phaB* loci in clone 25 are similar to those in clone P1N3 (Supplementary  
438 1), though the *phaB* gene in clone P1N3 was four open reading frames away from *phaC<sub>P1N3A</sub>*. PhaC  
439 proteins of clones P1N3 and 25 shared 77% identity. In addition, PhaC<sub>25</sub> was 35% and 61% identical  
440 to PhbC<sub>Sm</sub> and PhbC<sub>1Re</sub>. The conserved amino acid Asp130 affecting PhaC substrate specificity  
441 (Matsumoto et al. 2005) was replaced by Ala152 (Fig. 3B).

442 SCL-MCL PHA (C4-C6-C8) was produced in PpUW2 harbouring clone 25 when grown on  
443 octanoate, and only PHB (C4) was isolated from gluconic acid-grown cells (Table 4). Similar results  
444 were obtained in PpUW2 with clone 18 or clone 19 under the same conditions. However, the  
445 quantity of C6-C8 monomers was greater with clone 25 than those with clone 18 or 19. Nonanoic  
446 acid-grown PpUW2 (clone 25) produced 3HB-3HV copolymer similar to that from PpUW2 with  
447 clone 18 or clone 19 (Table 4). These data suggested that PhaC<sub>25</sub> had similar properties to those of  
448 PhaC<sub>18</sub> and PhaC<sub>29</sub>, but different from that of PhaC<sub>14</sub>.

449

#### 450 **Complementation of *S. meliloti* phaC mutant with 11AW PHA<sup>+</sup> clones**

451 In an earlier study, *S. meliloti* Rm11476 (*exoY::Tn5 phaC::Tn5-233*) had been employed for  
452 functionally harvesting metagenomic library clones pCX92, pCX9M1, pCX9M3 and pCX9M5  
453 expressing Class I PHA synthases (Schallmeyer et al. 2011). The PhaA/PhbB and PhaB/PhbB proteins  
454 encoded by the cloned metagenomic DNA and/or *S. meliloti* genome could supply (R)-3-  
455 hydroxybutyrate for PHB biosynthesis catalyzed by the PhaC/PhbC enzymes. In order to examine  
456 the possibility that the PHA<sup>+</sup> clones of 11AW metagenomic library were able to complement the  
457 PHA<sup>-</sup> phenotype of *S. meliloti* Rm11476, the Class I clones 14, 18, 19 and 25 (Table 3) were  
458 introduced into *S. meliloti* Rm11476. In each case, complementation was observed, indicating PhaC  
459 function in the *S. meliloti* background (Fig. 4), with greatest accumulation of PHA observed with  
460 clone 14. However, when clones pCX92, pCX9M1, pCX9M3 and pCX9M5 were transferred to *P.*  
461 *putida* PpUW2 (PHA<sup>-</sup>), PHA was not detected in the recombinant strains grown with either  
462 octanoate or gluconic acid (data not shown). These results strengthen the argument that multiple  
463 hosts should be employed to screen for novel *phaC* genes.

464 We then asked whether Class II *phaC* could function in the *S. meliloti* background. First, we  
465 observed no detection of PHA when *phaC1<sub>pp</sub>* (*Plac::phaC1<sub>pp</sub>*) was expressed in Rm11476 even  
466 though the *lac* promoter is functional and 3-hydroxybutyrate is supplied by the PhbAB<sub>Sm</sub>. Similarly,  
467 PHA<sup>+</sup> clones 1 and 20 carrying Class II *phaC* genes also failed to complement the PHB<sup>-</sup> phenotype  
468 of *S. meliloti* Rm11476 (Fig. 4), implying that the *phaC2<sub>1</sub>* and *phaC2<sub>20</sub>* genes were not expressed  
469 and/or the substrate (R)-3-hydroxyalkanoic acids were absent in *S. meliloti*. However, when clone 16  
470 was introduced into *S. meliloti* Rm11476, PHA was accumulated to a similar level as with  
471 Rm11476 containing PHB<sup>+</sup> clones pCX92, pCX9M1, pCX9M3 or pCX9M5 (Fig. 4). These data  
472 indicated that the *phaC1<sub>16</sub>* gene was expressed and its product was functional. These results suggest  
473 that *P. putida* is a more permissive surrogate host than *S. meliloti* for screening of novel PHA  
474 synthases.

475

## 476 Discussion

477 Biosynthesis of SCL-MCL PHA requires PHA synthases having a broad range of substrate  
478 specificity and SCL and MCL precursors (R-3-hydroxyacyl-CoAs) must be available within the cell.  
479 Building on earlier work that has sought to mine metagenomic DNA for modification of bacterial  
480 PHA production, we have successfully isolated cosmid clones that are able to functionally  
481 complement a *P. putida* PHA synthesis mutant. DNA sequence analysis revealed that the isolated  
482 metagenomic DNA originated from a broad diversity of bacteria, and encoded either Class I or Class  
483 II PHA synthases enzymes. Of note is that DNA from most of the isolated clones did not very  
484 closely match known sequences. Also of interest is the influence that the clones had on the quality  
485 and quantity of PHA produced in the *P. putida* surrogate host background.

486           A major appeal of PHA is that, as alternatives to conventional fossil fuel-derived plastics,  
487 they can be produced from renewable resources. The potential to improve the efficiency and cost of  
488 production, and expand the range of polymers and copolymers that are available for production,  
489 should impact on competitiveness and adoption of these materials in the marketplace. In this context,  
490 it is important to understand how the properties of PHA are determined by the PHA synthase  
491 enzyme and the substrate that is available to those enzymes, which is in turn influenced by the  
492 culture conditions and the metabolic pathways leading to substrate formation (Meng et al. 2014).  
493 PHA consisting of both short-chain-length (SCL,  $\leq C5$ ) and medium-chain-length (MCL,  $\geq C6$ )  
494 monomers have great flexibility, decreased breakage and reduced melting point (Noda et al. 2005).  
495 Most naturally occurring PhaC proteins polymerize either SCL- or MCL-monomers, but there are a  
496 few known examples such as the Class I PhaC of *Aeromonas caviae* (Fukui and Doi 1997) and some  
497 *Pseudomonas* Class II PhaC (Matsusaki et al. 1998) that can synthesize SCL-MCL copolymers.

498           In the present work, heterologous complementation of the *P. putida* PHA<sup>-</sup> strain facilitated  
499 efficient simultaneous screening of millions of 11AW metagenomic clones, resulting in recovery of  
500 a greater number of novel Class I, Class II, and unclassified PHA synthases than from previous  
501 reports (Schallmeyer et al. 2011; Cheema et al. 2012). Additionally, the clones carrying Class I *phaC*  
502 isolated in *P. putida* were able to complement the PHA<sup>-</sup> phenotype of *S. meliloti*, but the previously  
503 isolated Class I clones isolated in *S. meliloti* (Schallmeyer et al. 2011) failed to synthesize PHA in *P.*  
504 *putida*. These results support the development of multi-host systems to increase the chances of the  
505 successful expression of *pha* genes of interest.

506           Biosynthesis of SCL-MCL PHA of various forms by Class I clones 14, 18, 19 and 25 and  
507 Class II clones 16 and 20 suggested availability of a full range of R-3-hydroxyacyl-CoAs derived  
508 from  $\beta$ -oxidation of fatty acids such as nonanoic acid and octanoate. The differential monomer

509 compositions in PHA synthesized by the Class II enzymes PhaC<sub>16</sub> and PhaC<sub>20</sub> are of particular  
510 interest, especially since most Class II PHA synthases are unable to incorporate SCL monomers such  
511 as 3HB and 3HV. Changes of conserved amino acid residues Ser325Thr, Leu484Val, and  
512 Glu508Leu (conserved positions in *P. putida* PhaC1) have been identified to be involved in PhaC  
513 substrate specificity and PHA yield (Takase et al. 2003; Shozui et al. 2010; Chen et al. 2014).  
514 Presence of the amino acids Thr353 (corresponding to *P. putida* PhaC1 position 325), Val516  
515 (corresponding to *P. putida* PhaC1 position 484) and Ala540 (corresponding to *P. putida* PhaC1  
516 position 508) PhaC<sub>20</sub> might contribute to the enzyme's ability to incorporate SCL monomers  
517 compared to those of 11AW PhaC<sub>21</sub> and PHA synthases of *P. putida* KT2440. Similarly, the  
518 presence of three residues (Thr317, Val478 and Leu502) at these positions in PhaC<sub>16</sub>, most likely  
519 resulted in increased 3HB content as demonstrated previously with engineered Class II PHA  
520 synthases (Takase et al. 2003; Shozui et al. 2010; Chen et al. 2014).

521

## 522 **Conclusions**

523 We obtained 27 clones encoding Class I, II and unclassified PHA synthases by functional  
524 screening of soil metagenomic cosmid clones in a *P. putida* PHA<sup>-</sup> strain. Seven clones that were  
525 characterized in more detail were able to produce a broad range of polymers and copolymers,  
526 including SCL-MCL mixtures, depending on carbon source. Through this work we have  
527 demonstrated the potential for using metagenome-derived clones for production of a variety of PHAs  
528 of possible industrial utility. The collection of PHA metabolism genes from uncultivated organisms  
529 provides not only a resource for production strain development, but also a series of sequence  
530 templates that could prove useful in enzyme engineering efforts directed towards generation of PHA  
531 products with desired properties.

532

533 **Nucleotide sequence accession numbers**

534 Complete sequences of 11AW metagenomic PHA<sup>+</sup> DNA have been deposited in GenBank  
535 nucleotide sequence database [GenBank: KT944254-KT944278, KU728995 and KU728996].

536

537 **Abbreviations**

538 *P. putida*, *Pseudomonas putida*; PHA, polyhydroxyalkanoate; PhaA/PhbA,  $\beta$ -ketothiolase;  
539 PhaB/PhbB, acetoacetyl-CoA reductase; PhaC/PhbC, PHA synthase; PhaZ, PHA depolymerase;  
540 PhaJ, R-specific enoyl-CoA hydratase; 3HA, 3-hydroxyalkanoate; 3HB, 3-hydroxybutyrate; 3HV,  
541 3-hydroxyvalerate; 3HHx, 3-hydroxyhexanoate; 3HP, 3-hydroxyheptanoate; 3HO, 3-  
542 hydroxyoctanoate; 3HN, 3-hydroxynonanoate; 3HD, 3-hydroxydecanoate; SCL, short chain length;  
543 MCL, medium chain length; *S. meliloti*, *Sinorhizobium meliloti*; GC-MS, gas chromatography-mass  
544 spectrometry; CDW, cell dry weight.

545

546 **Authors's contributions**

547 JC and TCC conceived and designed this study, carried out all the experiments, and prepared  
548 the manuscript. Both authors read and approved the final manuscript.

549

550 **Acknowledgements**

551 We are very grateful to Ricardo Nordeste for assisting with GC-MS analysis of PHA. We  
552 wish to thank Dr. Bruce Ramsay (Polyferm Canada) for providing PHA standards, and Dr. Josh  
553 Neufeld for use of the FilterMax F5 Multi-mode microplate reader.

554

555 **Funding**

556 This work was financially supported by a New Directions grant from Ontario Ministry of  
557 Agriculture, Food, and Rural Affairs (award number 381646-09), by a Strategic Projects grant from  
558 the Natural Sciences and Engineering Research Council of Canada (award number ND2012-1679),  
559 and by Genome Canada, through the Applied Genomics Research in Bioproducts or Crops (ABC)  
560 program for the grant titled “Microbial Genomics for Biofuels and Co-Products from Biorefining  
561 Processes” awarded to David Levin and Richard Sparling.

562

563 **Conflict of interest**

564 Trevor Charles received partial financial support for presenting some of this material at the  
565 Society for Industrial Microbiology and Biotechnology Annual Meeting in 2014 and 2015. JiuJun  
566 Cheng declares no conflict of interest.

567

568 **Ethical statement**

569 The authors certify that this manuscript has not been published previously, and not under  
570 consideration for publication elsewhere, in whole or in part. No data have been fabricated or  
571 manipulated (including images), and no data, text, or theories by others are presented as if they were  
572 the authors’ own. Consent to submit has been received explicitly from all the authors listed. And  
573 authors whose names appear on the submission have contributed sufficiently to the scientific work  
574 and therefore share collective responsibility and accountability for the results. This article does not  
575 contain any studies with human participants or animals performed by any of the authors.

576

577

578

579

580 **References**

- 581
- 582 Cheema S, Bassas-Galia M, Sarma PM, Lal B, Arias S (2012) Exploiting metagenomic diversity for  
583 novel polyhydroxyalkanoate synthases: Production of a terpolymer poly(3-hydroxybutyrate-co-  
584 3-hydroxyhexanoate-co-3-hydroxyoctanoate) with a recombinant *Pseudomonas putida* strain.  
585 *Bioresource Technology* 103:322–328. doi: 10.1016/j.biortech.2011.09.098
- 586 Chen Y-J, Tsai P-C, Hsu C-H, Lee C-Y (2014) Critical residues of class II PHA synthase for  
587 expanding the substrate specificity and enhancing the biosynthesis of polyhydroxyalkanoate.  
588 *Enzyme and Microbial Technology* 56:60–66. doi: 10.1016/j.enzmictec.2014.01.005
- 589 Cheng J, Pinnell L, Engel K, Neufeld JD, Charles TC (2014) Versatile broad-host-range cosmids for  
590 construction of high quality metagenomic libraries. *Journal of Microbiological Methods* 99:27–  
591 34. doi: 10.1016/j.mimet.2014.01.015
- 592 de Eugenio LI, Escapa IF, Morales V, Dinjaski N, Galàn B, García JL, Prieto MA (2010) The  
593 turnover of medium-chain-length polyhydroxyalkanoates in *Pseudomonas putida* KT2442 and  
594 the fundamental role of PhaZ depolymerase for the metabolic balance. *Environ Microbiol*  
595 12:207–221. doi: 10.1111/j.1462-2920.2009.02061.x
- 596 Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space  
597 complexity. *BMC Bioinformatics* 5:113–19. doi: 10.1186/1471-2105-5-113
- 598 Escapa IF, Morales V, Martino VP, Pollet E, Avérous L, García JL, Prieto MA (2011) Disruption of  
599  $\beta$ -oxidation pathway in *Pseudomonas putida* KT2442 to produce new functionalized PHAs with  
600 thioester groups. *Appl Microbiol Biotechnol* 89:1583–1598. doi: 10.1007/s00253-011-3099-4
- 601 Fellay R, Frey J, Krisch H (1987) Interposon mutagenesis of soil and water bacteria: a family of  
602 DNA fragments designed for in vitro insertional mutagenesis of gram-negative bacteria. *Gene*  
603 52:147–154.
- 604 Foong CP, Lau N-S, Deguchi S, Toyofuku T, Taylor TD, Sudesh K, Matsui M (2014) Whole  
605 genome amplification approach reveals novel polyhydroxyalkanoate synthases (PhaCs) from  
606 Japan trench and nankai trough seawater. *BMC Microbiol* 14:e11836. doi: 10.1186/s12866-014-  
607 0318-z
- 608 Fukui T, Doi Y (1997) Cloning and analysis of the poly(3-hydroxybutyrate-co-3-hydroxyhexanoate)  
609 biosynthesis genes of *Aeromonas caviae*. *J Bacteriol* 179:4821–4830. doi: 10.1186/s12866-014-  
610 0318-z
- 611 Fukui T, Shiomi N, Doi Y (1998) Expression and characterization of (R)-specific enoyl coenzyme A  
612 hydratase involved in polyhydroxyalkanoate biosynthesis by *Aeromonas caviae*. *J Bacteriol*  
613 180:667–673.
- 614 Hoppensack A, Rehm BH, Steinbüchel A (1999) Analysis of 4-phosphopantetheinylation of  
615 polyhydroxybutyrate synthase from *Ralstonia eutropha*: generation of beta-alanine auxotrophic  
616 Tn5 mutants and cloning of the *panD* gene region. *J Bacteriol* 181:1429–1435.

- 617 Jia Y, Kappock TJ, Frick T, Sinskey AJ, Stubbe J (2000) Lipases provide a new mechanistic model  
618 for polyhydroxybutyrate (PHB) synthases: characterization of the functional residues in  
619 *Chromatium vinosum* PHB synthase. *Biochemistry* 39:3927–3936. doi: 10.1021/bi9928086
- 620 Kawashima Y, Cheng W, Mifune J, Orita I, Nakamura S, Fukui T (2012) Characterization and  
621 functional analyses of R-specific enoyl coenzyme A hydratases in polyhydroxyalkanoate-  
622 producing *Ralstonia eutropha*. *Appl Environ Microbiol* 78:493–502. doi: 10.1128/AEM.06937-  
623 11
- 624 Keshavarz T, Roy I (2010) Polyhydroxyalkanoates: bioplastics with a green agenda. *Current*  
625 *Opinion in Microbiology* 13:321–326. doi: 10.1016/j.mib.2010.02.006
- 626 Lam KN, Hall MW, Engel K, Vey G, Cheng J, Neufeld JD, Charles TC (2014) Evaluation of a  
627 pooled strategy for high-throughput sequencing of cosmid clones from metagenomic libraries.  
628 *PLoS ONE* 9:e98968. doi: 10.1371/journal.pone.0098968
- 629 Matsumoto K, Takase K, Aoki E, Doi Y, Taguchi S (2005) Synergistic effects of Glu130Asp  
630 substitution in the type II polyhydroxyalkanoate (PHA) synthase: enhancement of PHA  
631 production and alteration of polymer molecular weight. *Biomacromolecules* 6:99–104. doi:  
632 10.1021/bm049650b
- 633 Matsusaki H, Manji S, Taguchi K, Kato M, Fukui T, Doi Y (1998) Cloning and molecular analysis  
634 of the Poly(3-hydroxybutyrate) and Poly(3-hydroxybutyrate-co-3-hydroxyalkanoate)  
635 biosynthesis genes in *Pseudomonas* sp. strain 61-3. *J Bacteriol* 180:6459–6467.
- 636 Meng D-C, Shen R, Yao H, Chen J-C, Wu Q, Chen G-Q (2014) Engineering the diversity of  
637 polyesters. *Curr Opin Biotechnol* 29:24–33. doi: 10.1016/j.copbio.2014.02.013
- 638 Noda I, Green PR, Satkowski MM, Schechtman LA (2005) Preparation and properties of a novel  
639 class of polyhydroxyalkanoate copolymers. *Biomacromolecules* 6:580–586. doi:  
640 10.1021/bm049472m
- 641 Park SJ, Kim TW, Kim MK, Lee SY, Lim S-C (2012) Advanced bacterial polyhydroxyalkanoates:  
642 towards a versatile and sustainable platform for unnatural tailor-made polyesters. *Biotechnology*  
643 *Advances* 30:1196–1206. doi: 10.1016/j.biotechadv.2011.11.007
- 644 Patil KR, Roune L, McHardy AC (2012) The PhyloPythiaS web server for taxonomic assignment of  
645 metagenome sequences. *PLoS ONE* 7:e38581. doi: 10.1371/journal.pone.0038581
- 646 Pärnänen K, Karkman A, Virta M, Eronen-Rasimus E, Kaartokallio H (2015) Discovery of bacterial  
647 polyhydroxyalkanoate synthase (PhaC)-encoding genes from seasonal Baltic Sea ice and cold  
648 estuarine waters. *Extremophiles* 19:197–206. doi: 10.1007/s00792-014-0699-9
- 649 Quandt J, Hynes MF (1993) Versatile suicide vectors which allow direct selection for gene  
650 replacement in Gram-negative bacteria. *Gene* 127:15–21. doi: 10.1016/0378-1119(93)90611-6
- 651 Rehm BHA (2003) Polyester synthases: natural catalysts for plastics. *Biochem J* 376:15–33. doi:  
652 10.1042/BJ20031254

- 653 Sato S, Kanazawa H, Tsuge T (2011) Expression and characterization of (R)-specific enoyl  
654 coenzyme A hydratases making a channeling route to polyhydroxyalkanoate biosynthesis in  
655 *Pseudomonas putida*. Appl Microbiol Biotechnol 90:951–959. doi: 10.1007/s00253-011-3150-5
- 656 Schallmey M, Ly A, Wang C, Meglei G, Voget S, Streit WR, Driscoll BT, Charles TC (2011)  
657 Harvesting of novel polyhydroxyalkanoate (PHA) synthase encoding genes from a soil  
658 metagenome library using phenotypic screening. FEMS Microbiol Lett 321:150–156. doi:  
659 10.1111/j.1574-6968.2011.02324.x
- 660 Sheu DS, Wang YT, Lee CY (2000) Rapid detection of polyhydroxyalkanoate-accumulating  
661 bacteria isolated from the environment by colony PCR. Microbiology 146 ( Pt 8):2019–2025.  
662 doi: 10.1099/00221287-146-8-2019
- 663 Shozui F, Sun J, Song Y, Yamada M, Sakai K, Matsumoto K, Takase K, Taguchi S (2010) A new  
664 beneficial mutation in *Pseudomonas* sp. 61-3 polyhydroxyalkanoate (PHA) synthase for  
665 enhanced cellular content of 3-hydroxybutyrate-based PHA explored using its enzyme homolog  
666 as a mutation template. Biosci Biotechnol Biochem 74:1710–1712. doi: 10.1271/bbb.100224
- 667 Simon C, Daniel R (2011) Metagenomic analyses: past and future trends. Appl Environ Microbiol  
668 77:1153–1161. doi: 10.1128/AEM.02345-10
- 669 Steinbüchel A, Lütke-Eversloh T (2003) Metabolic engineering and pathway construction for  
670 biotechnological production of relevant polyhydroxyalkanoates in microorganisms. Biochemical  
671 Engineering Journal 16:81–96. doi: 10.1016/S1369-703X(03)00036-6
- 672 Tai YT, Foong CP, Najimudin N, Sudesh K (2015) Discovery of a new polyhydroxyalkanoate  
673 synthase from limestone soil through metagenomic approach. Journal of Bioscience and  
674 Bioengineering. doi: 10.1016/j.jbiosc.2015.08.008
- 675 Takase K, Taguchi S, Doi Y (2003) Enhanced synthesis of poly(3-hydroxybutyrate) in recombinant  
676 *Escherichia coli* by means of error-prone PCR mutagenesis, saturation mutagenesis, and in vitro  
677 recombination of the type II polyhydroxyalkanoate synthase gene. J Biochem 133:139–145.
- 678 Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S (2013) MEGA6: Molecular Evolutionary  
679 Genetics Analysis version 6.0. Mol Biol Evol 30:2725–2729. doi: 10.1093/molbev/mst197
- 680 Tortajada M, da Silva LF, Prieto MA (2013) Second-generation functionalized medium-chain-length  
681 polyhydroxyalkanoates: the gateway to high-value bioplastic applications. Int Microbiol 16:1–15.  
682 doi: 10.2436/20.1501.01.175
- 683 Tripathi L, Wu L-P, Dechuan M, Chen J, Wu Q, Chen G-Q (2013) *Pseudomonas putida* KT2442 as  
684 a platform for the biosynthesis of polyhydroxyalkanoates with adjustable monomer contents and  
685 compositions. Bioresource Technology 142:225–231. doi: 10.1016/j.biortech.2013.05.027
- 686 Verlinden RAJ, Hill DJ, Kenward MA, Williams CD, Radecka I (2007) Bacterial synthesis of  
687 biodegradable polyhydroxyalkanoates. J Appl Microbiol 102:1437–1449. doi: 10.1111/j.1365-  
688 2672.2007.03335.x

- 689 Wu H-A, Sheu D-S, Lee C-Y (2003) Rapid differentiation between short-chain-length and medium-  
690 chain-length polyhydroxyalkanoate-accumulating bacteria with spectrofluorometry. *Journal of*  
691 *Microbiological Methods* 53:131–135. doi: 10.1016/S0167-7012(02)00232-4
- 692 Zhu W, Lomsadze A, Borodovsky M (2010) Ab initio gene identification in metagenomic sequences.  
693 *Nucleic Acids Research* 38:e132–e132. doi: 10.1093/nar/gkq275

694

## 695 **Figure legends**

696

697 Fig. 1. Activity-based screening of 11AW metagenomic library clones carrying PHA biosynthesis  
698 genes in *P. putida* PpUW2 (PHA<sup>-</sup>). PpUW2 recombinant strains were grown in 0.1N M63 medium  
699 containing 0.5% Na octanoate and Nile red (0.5 µg/ml). Cosmid pJC8 and pJC67 (*Plac::phaC1<sub>Pp</sub>*)  
700 were used as negative and positive controls respectively. 11AW clones 14 and 16 were streaked on  
701 the same selection plate to verify the PHA<sup>+</sup> phenotype initially screened on LB with 0.5% (w/v)  
702 octanoate.

703

704 Fig. 2. Phylogenetic analysis of the polyhydroalkanoate synthases (PhaC) of 11AW metagenomic  
705 library clones. Protein sequence alignments were performed using MUSCLE (Edgar 2004).  
706 Neighbour-joining phylogenetic trees were generated with MEGA6 (Tamura et al. 2013). GenBank  
707 accession numbers of 11AW PhaC proteins: PhaC<sub>21</sub> (ALV86274), PhaC<sub>3</sub> (ALV86289), PhaC<sub>4</sub>  
708 (ALV86299), PhaC<sub>5</sub> (ALV86308), PhaC<sub>6</sub> (ALV86351), PhaC<sub>7</sub> (partial, ALV86358), PhaC<sub>8</sub>  
709 (ALV86364), PhaC<sub>10</sub> (partial, ALV86387), PhaC<sub>14</sub> (ALV86397), PhaC<sub>16</sub> (ALV86417), PhaC<sub>216</sub>  
710 (ALV86419), PhaC<sub>18</sub> (ALV86462), PhaC<sub>19</sub> (ALV86476), PhaC<sub>20</sub> (ALV86493), PhaC<sub>21</sub>  
711 (ALV86517), PhaC<sub>23</sub> (ALV86529), PhaC<sub>25</sub> (ALV86574), PhaC<sub>27</sub> (ALV86602), PhaC<sub>150</sub>  
712 (ALV86626), PhaC<sub>250</sub> (ALV86626), PhaC<sub>51</sub> (ALV86651), PhaC<sub>P1N3</sub> (ALV86715), PhaC<sub>P2N8</sub>  
713 (ALV86750), PhaC<sub>P2N10</sub> (partial, ALV86755), PhaC<sub>P4N10</sub> (ALV86768), PhaC<sub>P8N7</sub> (ALV86771).

714  
715 Fig. 3. The *pha* genes and PhaC proteins in 11AW metagenomic DNA library clones. (A) The PHA<sup>+</sup>  
716 clones are classified into Class I and II based on the PhaC protein sequences. (B) Conserved amino  
717 acids required for the activity of PHA synthases are indicated in closed triangles. Residues affecting  
718 substrate specificity were marked by open triangles. The positions of amino acid residues were  
719 numbered based on the sequence of PhaC<sub>16</sub> [GenBank: ALV86417]. GenBank accession numbers  
720 of PhaC proteins: PhaC<sub>14</sub>, ALV86397; PhaC<sub>18</sub>, ALV86462; PhaC<sub>19</sub>, ALV86476; PhaC<sub>25</sub>,  
721 ALV86574; PhaC<sub>12</sub>, ALV86274; PhaC<sub>20</sub>, ALV86493; PhaC<sub>16</sub>, ALV86417; *C. necator* PhbC1,  
722 AAA21975; *P. putida* PhaC1, Q88D25; and *P. sp.* 61-3 PhaC1, BAA36200.

723  
724 Fig. 4. Complementation of *S. meliloti* Rm11476 (*phbC*) with metagenomic clones. The cosmid  
725 DNA of PHA<sup>+</sup> clones encoding Class I and II PHA synthases was transferred to *S. meliloti* Rm11476  
726 by conjugation. Recombinant strains were grown in YM medium. PHA production was estimated by  
727 measuring the fluorescence of Nile red stained cells. *S. meliloti* wild type Rm1021 and previously  
728 isolated Class I clones (CX92, pCX92M1, pCX92M3, and pCX92M5; (Schallmey et al. 2011)) were  
729 used as controls.

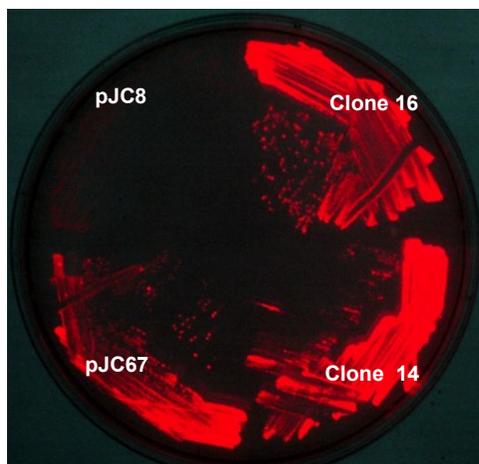


Figure 1.

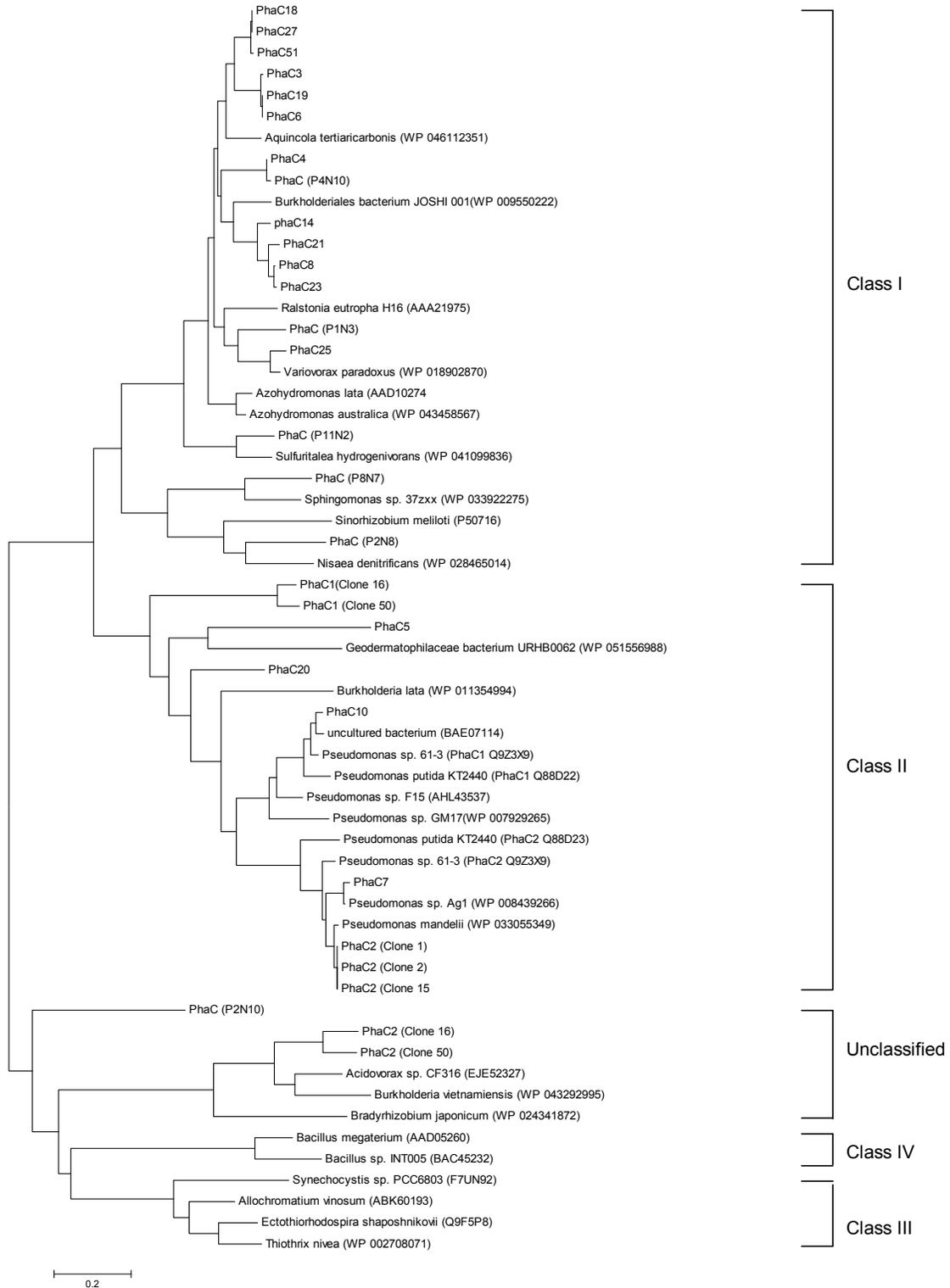


Fig. 2

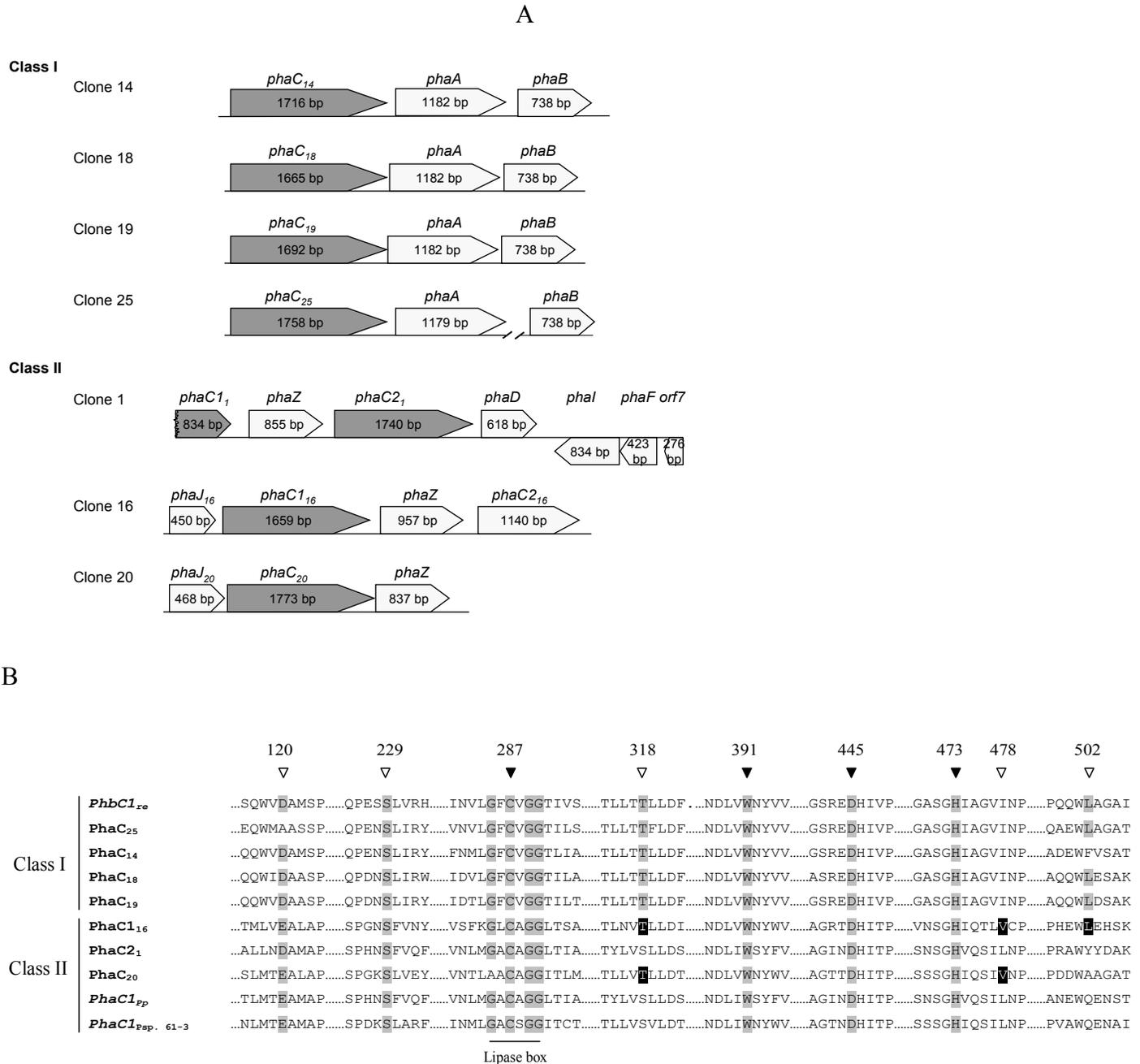


Fig. 3

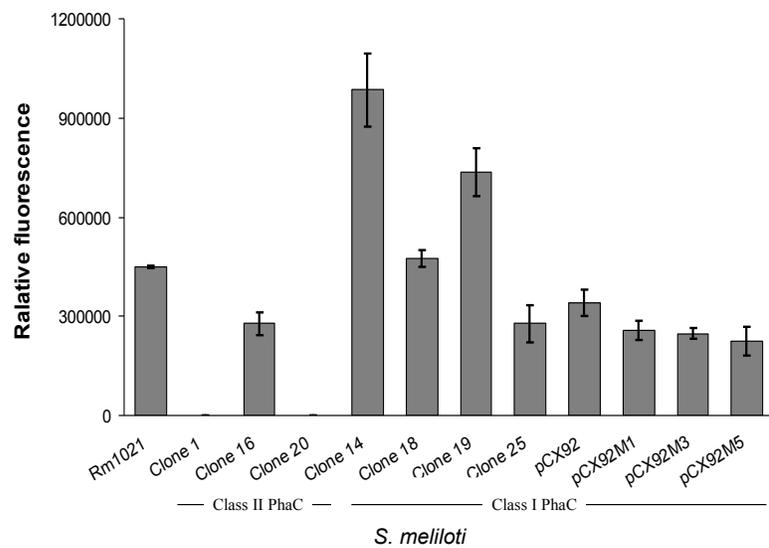


Fig. 4

Table 1. Bacterial strains, plasmids and cosmids

Strain, plasmids and cosmids	Relevant characteristics*	Reference
<i>E. coli</i>		
DH5α	F <sup>-</sup> φ80 <i>lacZAM15Δ(lacZYA-argF)U169 recA1 endA1 hsdR17 phoA supE44 thi-1 gyrA96 relA1</i>	Life Technologies
DH5α (Rif <sup>R</sup> )	A spontaneous Rif <sup>r</sup> mutant of DH5α	(Cheng et al. 2014)
HB101	F <sup>-</sup> <i>thi-1 hsdS20 supE44 recA13 ara-14 leuB6 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1</i> , Sm <sup>R</sup>	Life Technologies
<i>S. meliloti</i>		
Rm1021	Wild type, Sm <sup>R</sup>	Lab collection
Rm11476	Rm1021, <i>exoY::Tn5 phbC::Tn5-233</i> , Sm <sup>R</sup> Nm <sup>R</sup>	(Schallmeyer et al. 2011)
<i>P. putida</i>		
KT2440	Wild type	Lab collection
PpUW1	A spontaneous Rif <sup>R</sup> mutant of KT2440, Rif <sup>R</sup>	This study
PpUW2	PpUW1 ( <i>phaC1-phaZ-phaC2::ΩKm<sup>R</sup></i> ), Rif <sup>R</sup> Km <sup>R</sup>	This study
Plasmids/cosmids		
pRK600	pRK2013 <i>tra</i> Nm <sup>R</sup> ::Tn9, Cm <sup>R</sup>	Lab collection
pRK7813	IncP <i>oriT cos lacZα</i> , Tc <sup>R</sup>	Lab collection
pJC8	Broad-host-range Gateway <sup>®</sup> entry cosmid, Tc <sup>R</sup>	(Cheng et al. 2014)
pJQ200-SK	Cloning vector, <i>sacB</i> , Gm <sup>R</sup>	(Quandt and Hynes 1993)
pHP45-ΩKm <sup>R</sup>	Ω-Km <sup>R</sup> cassette, Km <sup>R</sup>	(Fellay et al. 1987)
pJC63	pJQ200-SK carrying 5'- <i>phaC1</i> of <i>P. putida</i> KT2440, Gm <sup>R</sup>	This study
pJC69	pJC63 carrying 3'- <i>phaC2</i> of <i>P. putida</i> KT2440, Gm <sup>R</sup>	This study
pJC71	pJC69 carrying Ω-Km <sup>R</sup> between the <i>phaC1'</i> and <i>phaC2'</i> fragments, Gm <sup>R</sup> Km <sup>R</sup>	This study
pJC67	pRK7813 carrying <i>Plac::phaC1<sub>pp</sub></i> , Tc <sup>R</sup>	This study
pJC123	pRK7813 carrying <i>Plac::phaC1<sub>16</sub></i> , Tc <sup>R</sup>	This study
pJC157	pRK7813 carrying <i>Plac::phaC1<sub>16</sub>::phaC2<sub>16</sub></i> , Tc <sup>R</sup>	This study
pJC159	pRK7813 carrying <i>Plac::phaC2<sub>16</sub></i> , Tc <sup>R</sup>	This study
pCX92	pRK7813 carrying <i>phaC</i> and <i>phaP</i> , Tc <sup>R</sup>	(Schallmeyer et al. 2011)
pCX9M1	pRK7813 carrying <i>phaC</i> , <i>phaA</i> and <i>phaB</i> , Tc <sup>R</sup>	(Schallmeyer et al. 2011)
pCX9M3	pRK7813 carrying <i>phaC</i> , Tc <sup>R</sup>	(Schallmeyer et al. 2011)
pCX9M5	pRK7813 carrying <i>phaC</i> and <i>phaB</i> , Tc <sup>R</sup>	(Schallmeyer et al. 2011)

\*Antibiotic resistances: Cm<sup>R</sup>, chloramphenicol; Gm<sup>R</sup>, gentamycin; Km<sup>R</sup>, kanamycin; Nm<sup>R</sup>, neomycin; Sm<sup>R</sup>, streptomycin; Tc<sup>R</sup>, tetracycline.

Table 2. DNA oligonucleotides

Oligo ID	Sequence
JC159	CGCCAAGCTTGTG <u><b>AGGAG</b></u> CGTCGTAGATGAGTAAC
JC160	GCGCGGATCCTCAACGCTCGTGAACGTAGGT
JC161	GCGCTCTAGACACGAAAGTCGGGTAACCAAT
JC162	GGTCGGATCCCGCCAGCTGATGATGAAGGTC
JC163	GCTGGGATCCGTTGTGGCTGGAGTGGATCAC
JC164	CGCCCTCGAGCACCGGTACCCTGATCTGATA
phaCF1	ATCAACAARTWCTACRTCYTSGACCT
phaCR2	GTSTTCRTRSRTSWSCTGGCGCAACCC
phaCR4	AGGTAGTTGTYGACSMRTAGKTCCA
JC202	CCGGATGCATGAT <u><b>AGGA</b></u> CGCCGGATGACAGCAGAGAAGGCTGAG
KC228	CCCAGATCTGCAGTATTCTAGATTCAGGCGCGCACGCGCACGTA
JC204	CCGGTCTAG <u><b>AGGA</b></u> ACGTCCCATGAGCAATCC
JC205	GGGGATGCATGCCTCAGGCGAAGTACTGGCCGCCATTG
JC301	CGCGAAGCTTCTG <u><b>AGGAG</b></u> AGAACACATGGCGAAGACCA
JC302	GGGGAATTCTGAATATCAGCGTCGGAGCATG
JC333	CGCGAAGCTTGTG <u><b>AGGAG</b></u> AGATCTATGAGCTCCCTCCCCGTGGCCGAACGTTC
JC334	CGCTGAATTCTCTACGCCGCTGACCCACGAT
JC337	ACCAGAATTC <u><b>AGGAG</b></u> ATCTATATGAGCTCCCTCCCCGTGGCCGAAC
pJET forward	CGACTCACTATAGGGAGAGCGGC
pJET reverse	AAGAACATCGATTTCCATGGCAG

Restriction sites are underlined. Translation stop codon is in bold, and ribosomal binding site is

double underlined.

Table 3. 11AW metagenomic clones encoding PHA synthases (PhaC).

PHA <sup>+</sup> Clones	DNA (bp)	GC content (%)	Possible origin of cloned DNA	Best match of PHA synthases	Identity (%)	Cosmid Genbank ID
<b>Class I</b>						
3	>16,077	67.5	<i>Leptothrix</i> ( <i>̑-Proteobacteria</i> )	Hypothetical protein (WP_046112351), <i>Aquicola tertiarycarbonis</i>	74	KT944255
4	40,585	65.7	<i>Leptothrix</i> ( <i>̑-Proteobacteria</i> )	Polyhydroxyalkanoate synthase (AAD10274), <i>Azohydromonas lata</i>	72	KT944256
6	>11,479	67.7	<i>Leptothrix</i> ( <i>̑-Proteobacteria</i> )	Polyhydroxyalkanoate synthase (AAD10274), <i>Azohydromonas lata</i>	73	KT944258
8	34,220	67.8	<i>Leptothrix</i> ( <i>̑-Proteobacteria</i> )	Poly-(R)-hydroxyalkanoic acid synthase (WP_009550222), Burkholderiales bacterium JOSHI_001	74	KT944260
14	29,565	69.1	<i>Methylibium</i> ( <i>̑-Proteobacteria</i> )	Poly-(R)-hydroxyalkanoic acid synthase (WP_009550222), Burkholderiales bacterium JOSHI_001	74	KT944262
18	37,718	69.4	<i>Rubrivivax</i> ( <i>̑-Proteobacteria</i> )	Hypothetical protein (WP_043458567), <i>Azohydromonas australica</i>	73	KT944264
19	25,085	68.1	<i>Leptothrix</i> ( <i>̑-Proteobacteria</i> )	Polyhydroxyalkanoate synthase (AAD10274), <i>Azohydromonas lata</i>	73	KT944265
21	33,064	69.3	<i>Gemmatimonas</i> ( <i>Gemmatimonadetes</i> )	Poly-(R)-hydroxyalkanoic acid synthase (WP_009550222), Burkholderiales bacterium JOSHI_001	73	KT944267
23	42,865	69.6	<i>Leptothrix</i> ( <i>̑-Proteobacteria</i> )	Poly-(R)-hydroxyalkanoic acid synthase (WP_009550222), Burkholderiales bacterium JOSHI_001	74	KT944268
25	41,012	66.8	<i>Variovorax</i> ( <i>̑-Proteobacteria</i> )	Poly-(R)-hydroxyalkanoic acid synthase (WP_018902870), <i>Variovorax paradoxus</i>	93	KT944269
27	37,156	67.0	<i>Rubrivivax</i> ( <i>̑-Proteobacteria</i> )	Hypothetical protein (WP_043458567), <i>Azohydromonas australica</i>	73	KT944270
51	38,794	69.9	<i>Rubrivivax</i> ( <i>̑-Proteobacteria</i> )	Hypothetical protein (WP_043458567), <i>Azohydromonas australica</i>	73	KT944272
P2N8	22,625	66.3	<i>Xanthomonadaceae</i> ( <i>̑-Proteobacteria</i> )	Poly(3-hydroxyalkanonate) synthase (WP_028465014), <i>Nisaea denitrificans</i>	55	KT944272
P11N2	40,293	58.4	<i>Janthinobacterium</i> ( <i>̑-Proteobacteria</i> )	Poly(3-hydroxyalkanonate) synthase (WP_041099836), <i>Sulfuritalea hydrogenivorans</i>	73	KT944278
P1N3	32,797	68.6	<i>Variovorax</i> ( <i>̑-Proteobacteria</i> )	Poly(3-hydroxyalkanonate) synthase (WP_028253507), <i>Variovorax</i> sp. URB0020	86	KT944273
P4N10	>2,724	64.2	<i>Leptothrix</i> ( <i>̑-Proteobacteria</i> )	Polyhydroxyalkanonate synthase (AAD10274), <i>Azohydromonas lata</i>	72	KT944276
P8N7	>5,619	67.6	<i>Sphingomonadaceae</i> ( <i>̑-Proteobacteria</i> )	Poly-(R)-hydroalkanoic acid synthase (WP_033922275), <i>Sphingomonas</i> sp. 37zxx	67	KT944277
<b>Class II</b>						
16 (PhaC <sub>1a</sub> )	33,900	66.1	<i>Methylibium</i> ( <i>̑-Proteobacteria</i> )	PhaC1 synthase (AHL43537), <i>Pseudomonas</i> sp.F15	48	KT944263
50 (PhaC <sub>1a</sub> )	22,965	67.0	<i>Hydrocarboniphaga</i> ( <i>̑-Proteobacteria</i> )	Polyhydroxyalkanoic acid synthase (WP_007929265), <i>Pseudomonas</i> sp. GM17	46	KT944271
1 (PhaC <sub>1c</sub> )	42,747	60.0	<i>Pseudomonas</i> ( <i>̑-Proteobacteria</i> )	Poly(R)-hydroxyalkanoic acid synthase (WP_033055349), <i>Pseudomonas mandelii</i>	98	KT944254
2	59,765	59.1	<i>Pseudomonas</i> ( <i>̑-Proteobacteria</i> )	Poly(R)-hydroxyalkanoic acid synthase (WP_033055349), <i>Pseudomonas mandelii</i>	98	KU729001
15	59,765	59.1	<i>Pseudomonas</i> ( <i>̑-Proteobacteria</i> )	Poly(R)-hydroxyalkanoic acid synthase (WP_033055349), <i>Pseudomonas mandelii</i>	98	KU728996
5	37473	66.3	<i>Burkholderia</i> ( <i>̑-Proteobacteria</i> )	Poly-beta-hydroxybutyrate polymerase (WP_051556988), <i>Geodermatophilaceae bacterium</i> URHB0062	53	KT944257
7	>407	65.6	<i>Pseudomonadaceae</i> ( <i>̑-Proteobacteria</i> )	Poly(R)-hydroxyalkanoic acid synthase (WP_008439266), <i>Pseudomonas</i> sp. Ag1	100	KT944259
10	>404	63.9	<i>Pseudomonas</i> ( <i>̑-Proteobacteria</i> )	PHA synthase (BAE07114), uncultured bacterium	98	KT944261
20	30,964	66.8	<i>Salinibacter</i> ( <i>Bacteroidetes</i> )	Poly(R)-hydroxyalkanoic acid synthase (WP_011354994), <i>Burkholderia lata</i>	59	KT944266
<b>Unclassified</b>						
P2N10	>13,508	66.6	<i>Alcagenaceae</i> ( <i>̑-Proteobacteria</i> )	Poly-R-hydroxyalkanoic acid synthase PhaE (EGV16635), <i>Thiocapsa marina</i> 5811	27	KT944275
				Poly-beta-hydroxybutyrate polymerase PhaC (WP_002708071), <i>Thiothrix nivea</i>	43	KT944275
16 (PhaC <sub>2a</sub> )	33,900	66.1	<i>Methylibium</i> ( <i>̑-Proteobacteria</i> )	Easterase (WP_043292995), <i>Burkholderia vietnamiensis</i>	53	KT944263
50 (PhaC <sub>2a</sub> )	22,965	67.0	<i>Hydrocarboniphaga</i> ( <i>̑-Proteobacteria</i> )	Easterase (WP_043292995), <i>Burkholderia vietnamiensis</i>	53	KT944271

The PhaC proteins were classified based on their amino acid sequences. Possible origin of cloned metagenomic DNA was predicted by PhyloPythia.

Table 4. Biosynthesis of SCL- and/or -MCL PHAs by 11AW metagenomic clones in *Pseudomonas putida*.

<i>P. putida</i>	Carbon sources	PHA content (%)	PHA composition (mol%)						
			3HB	3HV	3HHx	3HP	3HO	3HN	3HD
Class II									
PpUW1	Octanoate	30.0			0.7		99.3		
	Nonanoic acid	45.0				19.1		80.9	
	Gluconic acid	1.7					21.1		76.8
PpUW2 (clone 1)	Octanoate	32.7			4.5		95.5		
	Nonanoic acid	23.4				42.0		58.1	
	Gluconic acid	8.2					20.4		79.6
PpUW2 (clone 16)	Octanoate	25.1	17.3		39.3		43.2		
	Nonanoic acid	46.5	3.4	40.9		55.4			
	Gluconic acid	4.5			47.6		52.4		
PpUW2 (pJC123)	Octanoate	17.5	13.9		30.3		51.3		
PpUW2 (pJC157)	Octanoate	32.5	14.5		33.9		49.3		
PpUW2 (clone 20)	Octanoate	17.6			17.3		82.7		
	Nonanoic acid	45.0		7.5		50.2		42.3	
	Gluconic acid	3.9			9.9		37.4		52.7
Class I									
PpUW2 (clone 14)	Octanoate	41.8	98.9		1.1				
	Nonanoic acid	59.4	6.8	93.2					
	Gluconic acid	81.2	97.9		0.8		1.3		
PpUW2 (clone 18)	Octanoate	6.4	97.9		0.5		2.4		
	Nonanoic acid	17.3	32.8	67.2					
	Gluconic acid	30.7	100.0						
PpUW2 (clone 19)	Octanoate	30.9	98.8		0.5		0.7		
	Nonanoic acid	38.1	20.9	79.2					
	Gluconic acid	12.3	100.0						
PpUW2 (PhaC25)	Octanoate	15.6	95.0		2.9		2.1		
	Nonanoic acid	43.1	25.2	74.9					
	Gluconic acid	4.1	100.0						

*P. putida* PpUW1 (PHA<sup>+</sup>) and PpUW2 (PHA<sup>-</sup>) carrying PHA<sup>+</sup> cosmids of 11AW metagenomic clones were grown at 30°C for 48 h in 0.1N M63 media supplemented with 0.5% octanoate, 0.5% nonanoic acid or 1% gluconic acid. PHA content was represented as relative to cell dry weight. Values were the averages of triplicate cultures. The standard deviations were <10% of the values. Abbreviations: 3HB, 3-hydroxybutyrate (C4); 3HHx, 3-hydroxyhexanoate (C6); 3HO, 3-hydroxyoctanoate (C8); 3HD, 3-hydroxydecanoate (C10).

## Supplementary figures

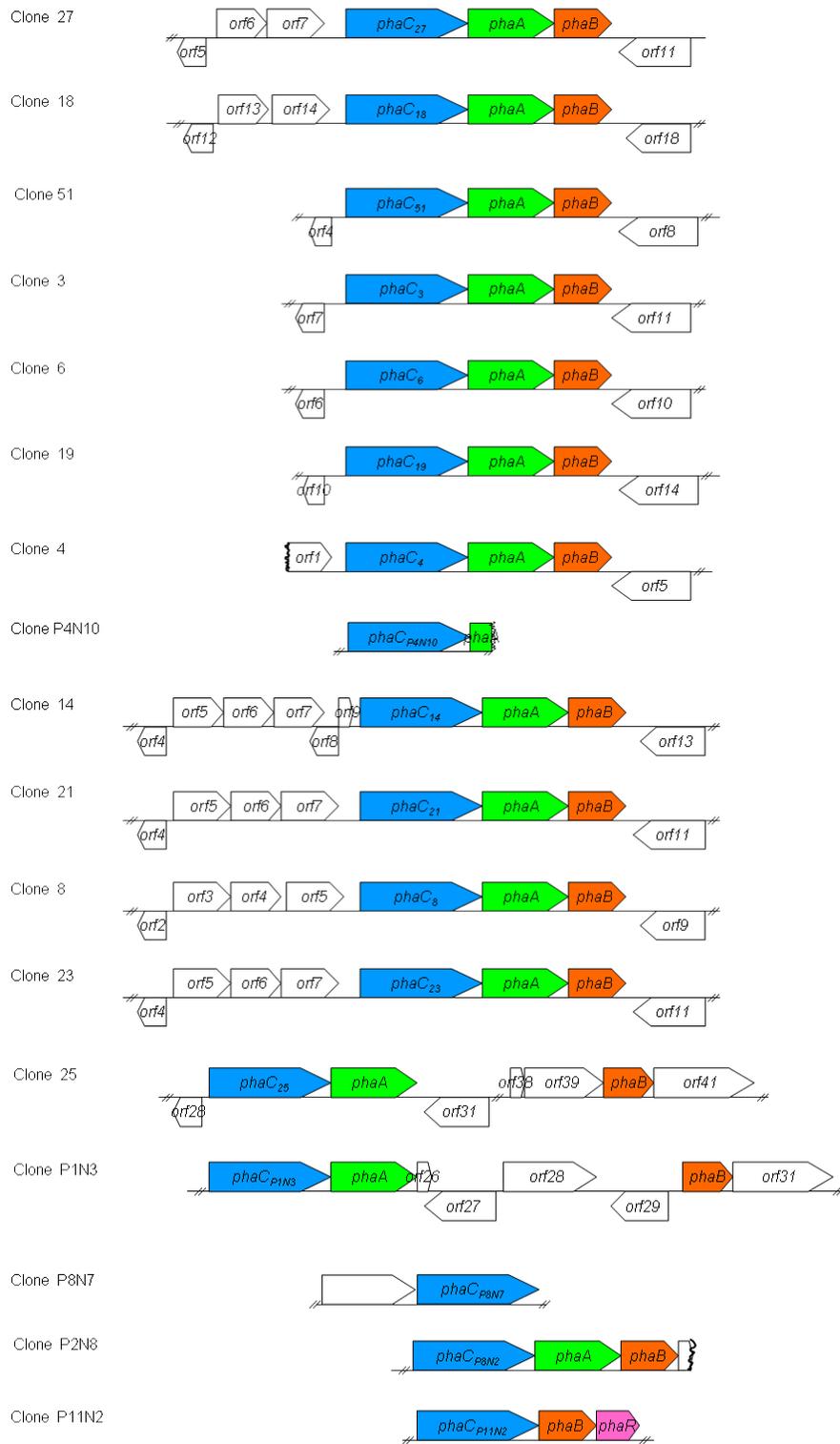
Supplementary 1. The *pha* gene locus in 11AW metagenomic DNA clones encoding PHA synthases (PhaC) were classified based on the PhaC protein sequences.

Supplementary 2. Transposon Tn4652 in 11AW PHA<sup>+</sup> clones 2 and 15. Insertion of the Tn4652 into the *orf17* of clone 1 resulted in clone 2. Clone 15 was derived from clone 1 by the transposon insertion into the *orf19*.

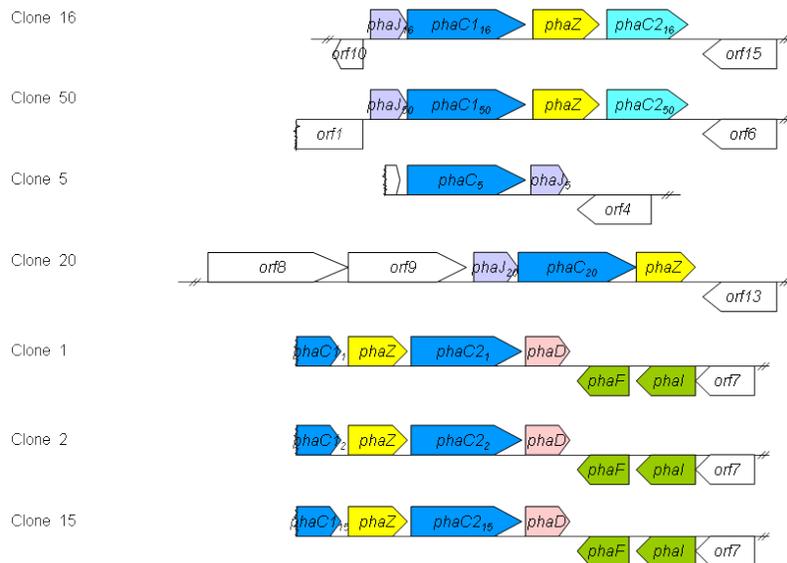
Supplementary 3. Multiple sequence alignment of PHA synthases of 11AW metagenomic DNA clones. Alignment was performed with MUSCLE. The conserved amino acid residues affecting substrate specificity in Class I and II PhaC are highlighted in purple and yellow respectively. The positions of amino acids essential for PhaC activity are highlighted in green. The position of amino acid residues are marked based on PhaC<sub>16</sub> sequence. Names of Class I PhaC are in regular font, Class II PhaC in bold, and unclassified PhaC underlined. PhaC proteins with partial sequences are marked with a star.

Supplementary 4. R-specific enoyl-CoA hydratases PhaJ. (A) Phylogenetic tree was constructed with MEGA6. The bar represents substitution of amino acid residue. (B) Conserved amino acid Asp<sup>39</sup>, His<sup>44</sup> and Ser<sup>7</sup> required for activity are marked.

## Class I



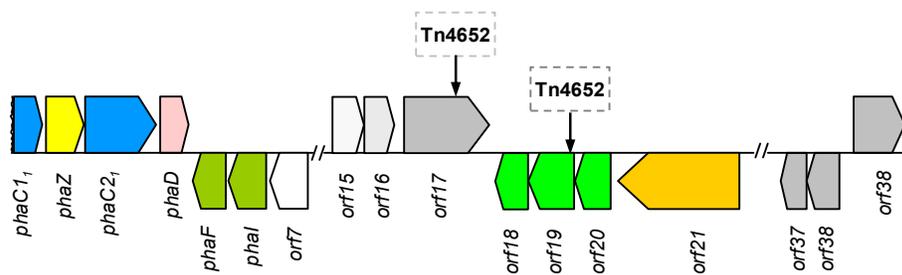
## Class II



## Unclassified



## Supplementary 1.



Supplementary 2.

```

PhaC216 -----msslpvaers-----
PhaC250 -----Mistkpvaers-----
PhaC_P2N10* -----
PhaC116 MAktkgsrrAriEsE---psrS---EmLrlAL
PhaC150 ---MaKtkApprwarlpsspsrs---ElLrlALv
PhaC5 MAkreatknvAatev-----vggGedVG
PhaC20 --mtatRrskArEde-prakkrv-----rkVaaAVGgapkpeagEArgfae---
PhaC10* -----
PhaC21 --mrektarefsEtPaafinaqS-----aitGlrgr-----
PhaC7* -----
PhaC_P11N2 -----MqyaghAlfqs-----fn-----QlLaaqLqQ-----
PhaC25 -----MmKqgAtgada---fapfq-----QALsegwnKAlesfqgSatqgasafn
PhaC_P1N3 -----MrQeqgAaada---tfapfq-----QALtgqwe-----
PhaC14 -----MkskpAAsarPtkqvtDsA-----QALGhAwG-----
PhaC8 -----MkskpAAsrP-dqvaEsA-----QALGhAwG-----
PhaC23 -----MkskpAAsrP-dqvaEsA-----QALGhAwG-----
PhaC21 -----MkskpAAsrP-dqvaEsA-----QsLgHawG-----
PhaC4 -----MpKtiASlPas-----A-----QAYGaAVG-----
PhaC_P4N10 -----MpKtiASlPas-----A-----QAYGaAVG-----
PhaC19 -----MsKntklPhs-----EqA-----dAfGhAVG-----
PhaC6 -----MsKntklPhs-----EqA-----dAfGhAVG-----
PhaC3 -----MsKntklPhs-----EqA-----dAfGhAVG-----
PhaC51 -----MmtgdA-----dAfGkAMG-----
PhaC27 -----mtgdA-----dAfGkAMG-----
PhaC18 -----mtgdA-----dAfGkAMG-----
PhaC_P8N7 MAdpdtKapeApElP---slEem-----QhwtwvMGRAgqmmmlE-----
PhaC_P2N8 -----MqtenSApIdPgavleElAqvaqqiqgqmrdaLdggwqQpvvdg-----

PhaC216 -----nLeK1L-----LAVP-----ayW---
PhaC250 -----nLeKAL-----LafP-----pyW---
PhaC_P2N10* -----
PhaC116 -----vilgrLrGn-----pgLP-----IRhiagfakef
PhaC150 -----ifgRlrrn-----pALP-----IRhiaafakel
PhaC5 -----vlSprSl-----Teaa-----tsAMNpL-----tmaRESarLygew
PhaC20 -----rvvdAlpG-----pnPfvvgfSvEdvlsaaeQlvagalqQpglVlrnqaaFaGd
PhaC10* -----
PhaC21 -----dLFsTLrs-----vaAhglr-----npvhtarhalklgGq
PhaC7* -----
PhaC_P11N2 -----ssveltng-----LSht-----gtaasasvmaL-----QeEwqarhmQLWqGm
PhaC25 vtgggtLWqipqt-----aSMp---elpkisiidpekLqsi-----QqQYVaEATELWrrqg
PhaC_P1N3 -----kaleAfqslgKApqadMPwdmprfsfsasrLqEL-----QrhYTeEATELWrrqd
PhaC14 -----eMwKsAvG-----LSLP-----tdAMSEL-----QtrYVQKQATELWNqa
PhaC8 -----mWKSvAG-----MSLP-----atAMTEL-----QSnYlKQATELWNqt
PhaC23 -----dMwKsvAG-----MSLP-----atAMTEL-----QSnYlKQATELWNqt
PhaC21 -----dMwKsvAG-----MSLP-----atAMTEL-----QSnYlKQATELWNqt
PhaC4 -----diWKSMSse-----LNLp-----1PALSEL-----QSnYVKQATElWNGa
PhaC_P4N10 -----diWKSMSse-----LNLp-----1PALSEL-----QSnYVKQATElWNGa
PhaC19 -----diWKSMSqG-----LSLP-----1PALSEL-----QtEYLKDATAlWNGa
PhaC6 -----diWKSMSqG-----LSLP-----1PALSEL-----QtEYLKDVtaLWNGa
PhaC3 -----diWKSMSqG-----LSLP-----1PALSEL-----QtEYLKDATAlWNGa
PhaC51 -----rMlqAaAG-----LTLp-----ptALTDL-----QAEYLREATEMWNGa
PhaC27 -----rMlqAaAG-----LTLp-----psALTDL-----QAEYLREATEMWNGa
PhaC18 -----rMlqAaAG-----LTLp-----psALTDL-----QAEYLREATEMWNGa
PhaC_P8N7 -----hvaRAMgetaEAkapALP---wllTlFgdpakLAQq-----QAElwtEglaIWqra
PhaC_P2N8 -----pvlnSMAG---gllqALavfnqSlFadparllQa-----QAKlwdgyaQLWqrg

PhaC216 -----piEW-----ar
PhaC250 -----piEW-----tk
PhaC_P2N10* LDaIlgasaea-----tpaagDkRFAdaaWSgNrlSgalvQvYiaasqeayaLvkeLEl-eg
PhaC116 LniLrgksdv---tpaagDkRFAdaaWSgNrlSgalvQvYiaasqeayaLvkeLEl-eg
PhaC150 LrvLrnksev---ApsagDkRFAdaaWSgNrlSgalvQvYiaasqeayaLvneLEl-Dg
PhaC5 LKiLvgrser---dvEvkDwRFAdptWhehPYkrTAQqYLaafdcaidRviDgnp---Dw
PhaC20 LaRlLsgrstl---epegDRRFQdptWkeNPFyraglQtYlLawrkgvhRLvegagl-Da
PhaC10* -----
PhaC21 LgRvlligetlh---ptnpgDnRFAdpaWS1NPfyrrSlQaYlswqevkhwiDdsnm-tp
PhaC7* -----
PhaC_P11N2 LQsnqaatEva---qslpgDRRFshpaWaesPiydylrQaYlINadyLqRLADtap---eg
PhaC25 fa-----AkEegDRRFASdaWgSNP1SAfSAavYLLNgRTLlnMADAIDa-De
PhaC_P1N3 ia-----ArthgDRRFAAeaWgSNPVAafSAavYLLNtRTLlGMAeAVEG-Da
PhaC14 LaRLqpDGakETAPAAFitDkrFAApEWSANPATAYTAQmYLLNARTLMQLAesMQG-Da
PhaC8 LsRLqpEGatETAaAtFitDkrFAApEWSANPASAYTAQmYLLNARTLMQLAesMQG-Da
PhaC23 LsRLqpEGatETAaAtFitDkrFAApEWSANPASAYTAQmYLLNARTLMQLAesMQG-Da
PhaC21 LsRLqpEGatETAPATFitDkrFAApEWSASPASAYTAQmYLLNARTLMQLAesMQG-Da
PhaC4 vERLqadGets---kkElpDRRFaanEWAANPvSALTAQiyLLNARTLLQMADAIEG-De
PhaC_P4N10 vERLqadGets---kkElpDRRFaanEWAANPvSALTAQiyLLNARTLLQMADAIEG-De
PhaC19 vERLktDesqt---AvaprDRRFaadEWSANPAAAMAaQTYLLNARTLMKMADAIEG-Dt
PhaC6 vERLktDeskt---AvaprDRRFaadEWSANPAAAMAaQTYLLNARTLMKMADAIEG-Dt
PhaC3 vERLktDeskt---AvaprDRRFaadEWSANPAAAMAaQTYLLNARTLMKMADAIEG-Dt
PhaC51 vERfsgkaPeQ---AqkigDRRFaanDWAANPATALaAqaYLLNsRTLmKMADAIEG-Da
PhaC27 vERfsgkaPeQ---ShkigDRRFaanDWAANPATALaAqaYLLNsRTLmKMADAIEG-Da
PhaC18 vERfsgkaPeQ---ShkigDRRFaanDWAANPATALaAqaYLLNsRTLmKMADAIEG-Da

```

PhaC\_P8N7 Lg--vdsGPrtelvekadrDRRFAApQWreNPlfdMirQsYLLvserLLgsvDsIDGLDp  
 PhaC\_P2N8 LgaLgtgqvEpaagEptaptDRRFRAqDwql-PAfdLirQwYLLtsqWLEqvtgagLDp

120

PhaC2<sub>16</sub> qIRkRgtklprdwIryVraagaIttf-PD-----Rpd--wVt  
 PhaC2<sub>50</sub> qIRqRgmrlpldwVsyVreakaIttf-PD-----KpG--wVt  
 PhaC\_P2N10\* KaR-----  
 PhaC1<sub>16</sub> rdaSaaRf1stmlVeAlaPSNnplntNFvvlKrvrETRgkS1lrGfQHfLsDrdlnaGmIS  
 PhaC1<sub>50</sub> rdaSsaRf1atmlVeAmApsNnplntNFvvmqrvrETKGrS1lrGLQHfTsDrdlnaGmIS  
 PhaC5 rkReRaRf1tgiltsAmAptNtLigNFAa1KKAyETgGkSlvrGvQnLIDDVrhnkGmpS  
 PhaC20 KgteRaRfvaslmtEAlaPtNFflgNFAaIKrALETgGgS1lrGLanLfDDvatngGmpS  
 PhaC10\* -----  
 PhaC2<sub>1</sub> ddRARahFAfallnDAMaPSNsL-LNFLaIKeifnsgGnSlvrGisHLvDdlhndGlp  
 PhaC7\* -----  
 PhaC\_P11N2 qIRnRmQf1thQaIDAMaPSNfaAtNPEfiKtAlaTeGESIrlGionLLqDlEkG--rIS  
 PhaC25 KTRARLRFAVEQWmaAsSPSNsLafNaEAQKKAIdTgGESIAkGionLLhDVKQG--H1S  
 PhaC\_P1N3 KTRARLRFAVEQWvaAsSPSNYLafNaEAQKKALESqGESIAkGionLLkDVQQG--H1S  
 PhaC14 KTRARVRFVQQWVDAMSPaNFMALNPEAQKKAldTKGESVSGiQqLLNDlRkG--aVS  
 PhaC8 KTRARVRFVQQWVDAMSPaNFMALNPEAQKKAldTKGESVAQGLQqLfnDlRkG--sVS  
 PhaC23 KTRARVRFVQQWVDAMSPaNFMALNPEAQKKAldTKGESVAQGLQqLfnDlRkG--sVS  
 PhaC21 KTRARVRFVQQWVDAMSPaNFMALNPEAQKKAldTKGESVAQGLQqLfnDlRkG--sVS  
 PhaC4 KTRGRIRFAVQQWVeAaSPSNYLALNPEAQrKALETkGESITqGLLHLWqDVQQG--HVS  
 PhaC\_P4N10 KTRGRIRFAEQWgeAarrSNYVALNPEAQrKALETkGESITqGLLHLWqDVQQG--HVS  
 PhaC19 KTRARIRFAVQQWVDAaSPSNYLALNPEAQrKALETkGESIAQGLRqLwDDIQG--HVS  
 PhaC6 KTRARIRFAVQQWVDAaSPSNYLALNPEAQrKALETkGESIAQGLRqLwDDIQG--HVS  
 PhaC3 KTRARIRFAVQQWVDAaSPSNYLALNPEAQrKALETkGESIAQGLRqLwDDIQG--HVS  
 PhaC51 KTRARVRFVQQWIDAaSPSNYLafNPEAQrKALETkGESIAQGLaHLlGDIQQG--HVS  
 PhaC27 KTRARVRFVQQWIDAaSPSNYLafNPEAQrKALETkGESIAQGLaHLlGDIQQG--HVS  
 PhaC18 KTRARVRFVQQWIDAaSPSNYLafNPEAQrKALETkGESIAQGLaHLlGDIQQG--HVS  
 PhaC\_P8N7 KqRekLRfMtrnfvDAMaPSNfaAtNPvvlrAmATRGES1lkGLEHmLDDVskG--qlt  
 PhaC\_P2N8 aTRhkLaFytrQfaDALSPSNfvAtNPEvIratvEsgGqnllhGfRHvLEDMkrGgGrTd

PhaC2<sub>16</sub> -----pntiayEsqtmLlRnfsdpvvsadtiPviidaPfaghPstIa  
 PhaC2<sub>50</sub> -----phtiaFETqtmwLrdfsdadApadaiPvividPyaghPstIa  
 PhaC\_P2N10\* -----LLcYpsLrepa-aRPLlLVyslvNr-ihvL  
 PhaC1<sub>16</sub> --mvDrtaFEVGRNVaAtpGdVVFvREvleLLqYKpqsAeVyiEiPlmvVPPqINK-YYvL  
 PhaC1<sub>50</sub> --mvDrtPFEVGRNVaAtpGdVVFvREvceVlqYKpqtAeVysiPlmvVPPqINK-fYvL  
 PhaC5 --QvkrSdFkVgEnLaAtpGAVVlRtErleLlqYqPtTAKVyaiPtLiIPppIGk-hYfi  
 PhaC20 --QvDkSaFvRgEnLaAtpGAVVfREvleivqYaPagAqVygRPLlLVPPqINK-YYvi  
 PhaC10\* -----  
 PhaC2<sub>1</sub> --QvtkqaFEVGktVATTtGsVVFvRENLEliqYKPMseKqysKPLlLVPPqINK-YYif  
 PhaC7\* -----  
 PhaC\_P11N2 --mTDEaaFEiGRNlAvTsGsVVFENElmQliqYaPlTpKVaaRPLviVPPCINK-fYIL  
 PhaC25 --mTDESaFEVGRNVATTEGAVVFENELFQLLEYKPLTpkVyERpflLVPPCINK-YYIL  
 PhaC\_P1N3 --mTDESaFEVGRNVATTEGAVVFENELFQLLEYKPLTAKVyERpflLVPPCINK-fYIL  
 PhaC14 --QTDetVFEVGRNVATsEGAVVFENELFQLiEYKPLTgKIhtRPMFLVPPCINK-YYIL  
 PhaC8 --QTDsvFEVGRNVATsEGAVVFENELFQLiEYKPLTAKVhtRPMFLVPPCINK-fYIL  
 PhaC23 --QTDsvFEVGRNVATsEGAVVFENELFQLiEYKPLTAKVhtRPMFLVPPCINK-fYIL  
 PhaC21 --QTDsvFEVGRNVATsvGAVVFENELFQLiEYKPLTAKVhtRPMFLVPPCINK-fYIL  
 PhaC4 --QTDsvFEVGRNVATTEGAVVFENELFQLiEYKPLTdkVHETpFLFVPPCINK-YYIL  
 PhaC\_P4N10 --QTDsvFEVGRNVATTEGAVVFENELFQLiEYKPLTdkVHETpFLFVPPCINK-YYIL  
 PhaC19 --QTDsvFEVGRNVATsEGAVVFENELFQLLEYKPLTaqVHERPMLFIPPCINK-YYIL  
 PhaC6 --QTDsvFEVGRNVATsEGAVVFENELFQLLEYKPLTaqVHERPMLFIPPCINK-YYIL  
 PhaC3 --QTDsvFEVGRNVATsEGAVVFENELFQLLEYKPLTaqVHERPMLFIPPCINK-YYIL  
 PhaC51 --QTDsvFEVGRNVATTEGAVVFENELFQLLEYKPLTaqVHERPMLFVPPCINK-YYIL  
 PhaC27 --QTDsvFEVGRNVATTEGAVVFENELFQLLEYKPLTaqVHERPMLFVPPCINK-YYIL  
 PhaC18 --QTDsvFEVGRNVATTEGAVVFENELFQLLEYKPLTaqVHERPMLFVPPCINK-YYIL  
 PhaC\_P8N7 --hTDeaFEVGRNiAvTpGkVvkrtpLYeLiqYsPtTetVygtpLviFPpWInR-fYIL  
 PhaC\_P2N8 prmTDPnaFEVGaNiAaTpGkVVFqNgLmQliqYtPvTdqVyrRPLlLVPPqINK-fYvm

229

PhaC2<sub>16</sub> DyadggSLvqtlkagGCKRvYvtdWksatheMrnfgidkYlsDld---QaIdAVgGRa-  
 PhaC2<sub>50</sub> DyadggSLvqtlkagGCKRlyVtdWksathaMrnfgidkYlaD--LR--daVdhVgGRa-  
 PhaC\_P2N10\* DLlarchfIerlqglG-HpVYvldwgrPdaSdAglsIaDYcgylstAVRLVrrhgcaR-  
 PhaC1<sub>16</sub> DLsPgNsfvNyaVKQG-fqVfMvSWRNpPpeLAhlsldhYlr-GleeAhnaVmeITGak-  
 PhaC1<sub>50</sub> DLsPgNsfvNyaVKQG-fqVfMvSWRNpPpeLAy1slqhYlQ-aledAhRaVmeITGad-  
 PhaC5 DLsPgNsfveYaVakG-fqVfCISWRNPqveqgdwgiDDYI-gsmLeAVdaVckITGsK-  
 PhaC20 DLsPgkSLveYaVasG-iqfFaVSWRNptvaerdwglEtYll-alseAIdaVceITGsE-  
 PhaC10\* -----FmVSWRNptkaqrewglstYIE-alkaAVGVVsAITGsK-  
 PhaC2<sub>1</sub> DLsPhNsfvqfalKnG-lqtFmISWRNPdvrhrewglssYVE-aveeAmnVcrAITGaR-  
 PhaC7\* -----VfVvSWRNpdvrhrewglstYVa-aleeAlnVtrAITGaR-  
 PhaC\_P11N2 DLQPENSliRYaVEQG-ntVfMvSWRNvqadLghlTWDDYIEqGaLKAIRVaQqIcrvp-  
 PhaC25 DLQPENSliRYaVEQG-HRVFVvSWRNpDeSLAnaTWDDYIEnaaIKAIhtVQdISGsK-  
 PhaC\_P1N3 DLQPENSliRYaVEQG-HRVFVvSWRNpDaSmSgyTWDDYIEDaaIKAIQVtreITGQdk  
 PhaC14 DLQPENSliRYTVEQG-HRVFvMvSWRNpDeSLAayTWDDYIEkGpLKAIEVVOqIGSQK-  
 PhaC8 DLQPENSliRYTVEQG-HRVFvMvSWRNpDeSLAhfTWDDYIEkGpLKAIEVVOqIGSQK-  
 PhaC23 DLQPENSliRYTVEQG-HRVFvMvSWRNpDeSLAhfTWDDYIEkGpLKAIEVVOqIGSQK-  
 PhaC21 DLQPENSliRYTVEQG-HRVFvMvSWRNpDeSLAhfTWDDYIEkGpLKAIEVVOqIGSQK-  
 PhaC4 DLQPENSliRwTVEQG-HRlFvMvSWRNpDdSVaakTWndYIEDGTIRAIeVvreITGaE-  
 PhaC\_P4N10 DLQPENSliRwTVEQG-HRlFemSWRNpDdSVaakTWndYIEDGTIRAIeVvreITGaE-  
 PhaC19 DLQPENSliRYTVAQG-HRVFVvSWRNpDpSIAaakTWDDYIDDGaIRAIReVQAISGQa-

PhaC6 DLQPDNSLIRYTVaQG-HRtFVVSWRNPdPSIAakTWDDYIDDGaIRAIReVQAISGQa-  
 PhaC3 DLQPDNSLIRYTVaQG-HRtFVVSWRNPdPSIAakTWDDYIDDGaIRAIReVQAISGQa-  
 PhaC51 DLQPDNSLIRwTVgQG-HRtFVVSWRNPdPSIAadkTWDDYIEDGaIRAIReVQISIGQK-  
 PhaC27 DLQPDNSLIRwTVgQG-HRtFVVSWRNPdPSIAadkTWDDYIEDGaIRAIReVQISIGQK-  
 PhaC18 DLQPDNSLIRwTVgQG-HRtFVVSWRNPdPSIAadkTWDDYIEDGaIRAIReVQISIGQK-  
 PhaC\_P8N7 DLNPKKsIRwaVEQG-ltVfVVSWSksaDeSLAgTLDYVlGgVdAdtIrellGvE-  
 PhaC\_P2N8 DLQPKNSLvgwwlQQG-ytVfMISWlNfpgqLAEksfeDYmiEGpLaAldaIEqacGER-

287 Lipase box 318

PhaC2<sub>16</sub> ----hlvGlcqGGwLs----AMhAARfp-gkAvSlvLagapiDthAgh-Gpvtkmaqrt  
 PhaC2<sub>50</sub> ----hlvGlcqGGwLst----MLAARyp-dkvvSlvLagapiDthAgD-Gpvtkkskrt  
 PhaC\_P2N10\* ---tIdlGlcCqGGs-LALcLAsLepRal----nrLvtLvTpvDFh--tpG--DVlAdlA  
 PhaC1<sub>16</sub> ---KVsfkGlcAgGltSAlALgryAdqGkLdtvntLTlnvTLLDiSgmEktnmgyFLtpe  
 PhaC1<sub>50</sub> ---KVNfklGlcAgGltLAlALgryAdqGkLdtvntLTlnvTLLDiAgmEktnmgyFLtpe  
 PhaC5 ---KVNisGFCaGgimttLmLsyaAAvrD-krvnaaafgvmLLDFd--teapigalhsrk  
 PhaC20 ---RVNTLaacaGgitLmlLsYLAAGD-rrvnSvTLLvTLLDte--aeGvLgLFasEe  
 PhaC10\* ---dINmLcaGsgitctalLghyAALGE-kkvnLTLLvsVLDtT--ldtqvalFVDEq  
 PhaC21 ---EVNLMGacaGgltIAalqghLqAkRQLrrvSSaTyLvsLLDsq--mdspatLFADEq  
 PhaC7\* ---EVNLMGacaGgltIAalqghLqAkRQLrrvSSaTyLvsLLDsq--idspatLFADEq  
 PhaC\_P11N2 ---QVNLaGFCVGGTILssALAVARIGE-dPvASLTLLTLLDfS--DsGeigyFVDEA  
 PhaC25 ---QVNvLGFVCGGTILsTALAVLAARGE-kPASSvTLLTfLDFs--DTGiLDIFVDEp  
 PhaC\_P1N3 ---kGgQIdTLFCVGGTILsTALAVLAARGE-QPASSvTLLTLLDfE--DTGiLDIFIDEg  
 PhaC14 ---tfnmLGFVCGGTILATALAVLAARGQ-QPAhSLTLLTLLDfK--hTGvLDVfVDEp  
 PhaC8 ---QfnmLGFVCGGTILATALAVLAARGE-QPAhSLTLLTLLDfK--hTGvLDVfVDEp  
 PhaC23 ---QfnmLGFVCGGTILATALAVLAARGE-QPAhSLTLLTLLDfK--hTGvLDVfVDEp  
 PhaC21 ---EfnmLGFVCGGTILsTALAVLAARGE-QPAhSLTLLTLLDfK--hTGvLDVfVDEp  
 PhaC4 ---RlNtLGFVCGGTILATALAVLAARGD-QPAASvTLLTaLLDfS--ETGvLDLFIIDEp  
 PhaC\_P4N10 ---RlNtLGFVCGGTILATALAVLAARGD-QPAASvTLLTaLLDfS--ETGvLDLFIIDEp  
 PhaC19 ---tIdTLGFVCGGTILtTALAVLAARGE-QPAASLTLLTLLDfA--DTGvLDIFIDEA  
 PhaC6 ---tIdTLGFVCGGTILtTALAVLAARGE-QPAASLTLLTLLDfA--DTGvLDIFIDEA  
 PhaC3 ---tIdTLGFVCGGTILtTALAVLAARGE-QPAASLTLLTLLDfG--DTGvLDIFIDEA  
 PhaC51 ---QIdvLGFVCGGTILATALAVLAARGE-QPAASLTLLTLLDfS--NTGvLDIFIDEA  
 PhaC27 ---QIdvLGFVCGGTILATALAVLAARGE-QPAASLTLLTLLDfS--NTGvLDIFIDEA  
 PhaC18 ---QIdvLGFVCGGTILATALAVLAARGE-QPAASLTLLTLLDfS--NTGvLDIFIDEA  
 PhaC\_P8N7 ---sVhaIGYCVaGtTLAatLALLEARGEadkvASatffTaqvDfS--EaGdLnLFVade  
 PhaC\_P2N8 ---EVNavGYCIGTLLlStLaymtAvnD-ErikSatTfaSLDfS--DpGdLgVFIDEA

PhaC2<sub>16</sub> pfnfyrsIvA---mGgGrmRGrfmlqgwnknhPdehyWgkyidlfndnitPqyveRARH  
 PhaC2<sub>50</sub> plsfyrsmVa---mGgGrmRGrfmlqgwnknhPdehyWgkyidlfndnitPqyvsRARH  
 PhaC\_P2N10\* rVldpallapd----GnLpGavLagLFGmLRFlrgVgaarapwldGgaEPekleRlRr  
 PhaC1<sub>16</sub> glEtsmkrskKE----GvLyChEmAkMFawLRPNdLVWNYwVnNYLmGQKPaFD----  
 PhaC1<sub>50</sub> glEksmkrskQE----GvLyChEmAkMFawLRPNdLVWNYwVnNYLmGKPaFD----  
 PhaC5 lIqVarsrsakK----GThpasaLqQvFawmRPNDLVWNYwVnNYLtgQdPPSFD----  
 PhaC20 aVaLaklnswtk----GvLpGEElgrvFawLRPNdLVWNYwVnNYLlGnaPPaFD----  
 PhaC10\* tLEaakrrsyQa----GvLEGrDmAkVFawmRPNDLi-----  
 PhaC21 tLEaakrrsyQk----GvLdGrDmAkVFawmRPNDLiWsYfVnNYLlGKEPPaFD----  
 PhaC7\* tLEaakrrsyQq----GvLdGrDmAkVFawmRPNDLi-----  
 PhaC\_P11N2 sVaaREagIG----QG-GLLhGrELATvFSaLRaNDLiWqYVVGNYLKGKPPaFD----  
 PhaC25 mVayREmqLG----kG-GLLpGmDLAsTFsFLRPNDLVWNYVVGNYLKGEPpPPFD----  
 PhaC\_P1N3 fVrFREmGmN----aG-GLLpGsELAsTFsFLRPNDLVWNYVVGNYLKGEPpPPFD----  
 PhaC14 sVQLREvTIGEQAPnGpGLLKGkELATTFsFLRPNDLVWNYVVGNYLKGEPpPPFD----  
 PhaC8 aVQmREmTLGEQAPnGPaLLKGkELATTFsFLRPNDLVWNYVVGNYLKGEPpPPFD----  
 PhaC23 tVQmREmTLGEQAPnGPaLLKGkELATTFsFLRPNDLVWNYVVGNYLKGEPpPPFD----  
 PhaC21 sVQmREmTIGEQAPnGPaLLKGkELATTFsFLRPNDLVWNYVVGNYLKGEPpPPFD----  
 PhaC4 aVQmREaTIGpQsPnGcGLLKGQELATTFsFLRPNDLVWNYVVGNYLKGEPpPPFD----  
 PhaC\_P4N10 aVQmREaTIGpQsPnGcGLLKGQELATTFsFLRPNDLVWNYVVGNYLKGEPpPPFD----  
 PhaC19 aVQLREmTLGEQAPQGPaLLKGQELATTFsFLRPNDLVWNYVVGNYLKGnKPPpPPFD----  
 PhaC6 aVQLREmTLGEQAPQGPaLLKGQELATTFsFLRPNDLVWNYVVGNYLKGnKPPpPPFD----  
 PhaC3 aVQLREmTLGEQAPQGPaLLKGQELATTFsFLRPNDLVWNYVVGNYLKGnKPPpPPFD----  
 PhaC51 qVQLREaTIGldAPQGPGLLKGQELATTFsFLRPNDLVWNYVVGNYLKGDKpPPFD----  
 PhaC27 qVQLREaTIGldAPQGPGLLKGQELATTFsFLRPNDLVWNYVVGNYLKGDKpPPFD----  
 PhaC18 qVQLREaTIGldAPQGPGLLKGQELATTFsFLRPNDLVWNYVVGNYLKGDKpPPFD----  
 PhaC\_P8N7 tLQLVeqisaDk----GyLdGrymAAtnLrgrDLiWsYVtnNYLmGEdyaPFD----  
 PhaC\_P2N8 qTaalEgkMa----GaGyLdsadMaTafnLrANdLVwsffVnNYLlGkDpPaFD----

391 445

PhaC2<sub>16</sub> farYehvvdLPGrMYleavkHlflkNlRlar-GEfvalGERIsLksIavPVYllAgadD  
 PhaC2<sub>50</sub> fasYehvvdLPGrLYleavkHlflkNlRlar-GEfvalGERIsLksItvPVYllAgadD  
 PhaC\_P2N10\* LvaYggDypdqaGrawLefvtacYreNrLlk-GtLvldGhtVdLRlRrlPvlnvfaRaDH  
 PhaC1<sub>16</sub> vLYNSDSTRLPaaLHhdfc-dmvakNELngvevyqLpdaqVDLsRvDvesfVAgRtDH  
 PhaC1<sub>50</sub> vLYNSDaTRLPaaLHhdfc-ellktNELgrgevyqVedaKVDLRvDvdsfVAgRtDH  
 PhaC5 lLawsVdGTNLPGkLHqgfL-difeaNaLpkkGsinlGkplDLesIkvetlVtggttDH  
 PhaC20 vLYNNDtTRLPaLHqgfL-dflLlNpfrNpdaLellGkpiDvanvsodaYVAgttDH  
 PhaC10\* -----  
 PhaC21 iLYNNDnTrLPaaLHgdll-dffkhNpLshPGgLeVCGtpiDLqkvNvdsfsvAginDH  
 PhaC7\* -----  
 PhaC\_P11N2 LLYNSDSTNLPGFflTWYLRnmYLeNnLriPGKLamCGvKaDLGhvDmPsfVvAcREDH  
 PhaC25 LLYNSDaTNLPGFfYTWYLRnTYheNkLakPnaLTVCGEKiDLGKIDiPaYIYGsREDH  
 PhaC\_P1N3 LLYNSDaTNLPGPMYCWYLRnTYheNkLakPnaLTVaGecqVDLGRIEAPVYlYgSREDH  
 PhaC14 LLYNNGDSTNLPGPMYCWYLRHTYLqNELrqpGKLTVCGEKiDLGaIDAPVYIYGsREDH

PhaC8 LLYWNgDSTNLPgPMYCWYLRHTYltNELRqPGKLTVCGEKiDLGaIDAPVYIYgSREdH  
 PhaC23 LLYWNgDSTNLPgPMYCWYLRHTYltNELRqPGKLTVCGEKiDLGaIDAPVYIYgSREdH  
 PhaC21 LLYWNgDSTNLPgPMYCWYLRHTYltNELRqPGKLTVCGEKiDLGaIDAPVYIYgSREdH  
 PhaC4 LLYWNSDSTNLPgFflCWYLRnTYLeNkLkDpGavTVCGEKVDLGAIEAPVYIYsSREdH  
 PhaC\_P4N10 LLYWNSDSTNLPgFflCWYLRnTYLeNkLkDpGavTVCGEKVDLGAIEAPVYIYsSREdH  
 PhaC19 LLYWNSDSTNLPgMFCWYLRHTYlqNELRvPGKLVVCGEKVDLGIkAPVfVYgSREdH  
 PhaC6 LLYWNSDSTNLPgMFCWYLRHTYlqNELRvPGKLVVCGEKVDLGIkAPVfVYgSREdH  
 PhaC3 LLYWNSDSTNLPgMFCWYLRHTYlqNELRvPGKLVVCGEKVDLGIkAPVfVYgSREdH  
 PhaC51 LLYWNSDSTNLPgMFCWYLRHTYlqNELRqPGKLTVCgQKVDLGIkAPVfVYASREdH  
 PhaC27 LLYWNgDSTNLPgMFCWYLRHTYlNnELRqPGKLTVCgQKVDLGIkAPVfVYASREdH  
 PhaC18 LLYWNgDSTNLPgMFCWYLRHTYlNnELRqPGKLTVCgQKVDLGIkAPVfVYASREdH  
 PhaC\_P8N7 LLhVNSDStTNLPakwHraYLRdFYrdNkLvraGELvVdGtpiDihKvktPtYVqAgREdH  
 PhaC\_P2N8 LLfVNaDaTrmPaaMHsfYLRnmYLnNrLRePGgiTLaGvpinLsKItvPcYfvAtvEDH

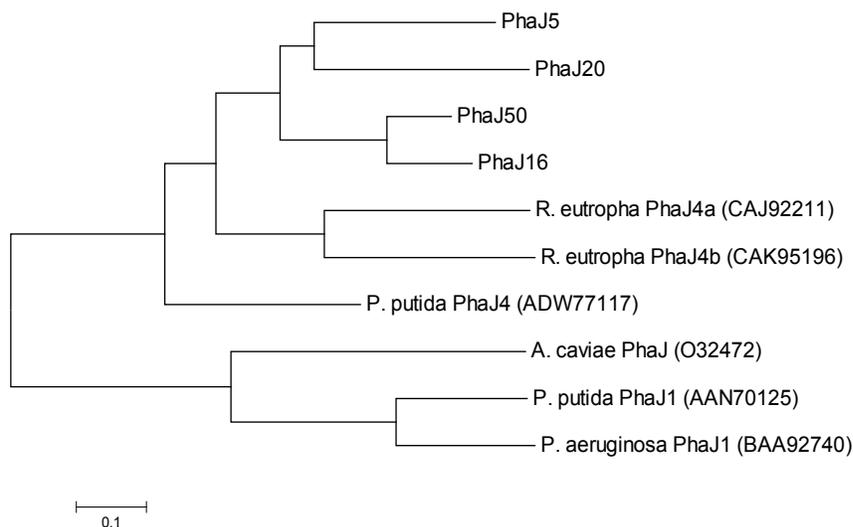
473 478

PhaC2<sub>16</sub> IttpeqvfnaenliGTPKdqRkeLapGChIglfmgaktlKqg-----  
 PhaC2<sub>50</sub> IttpeqvfnaenlvGTPRdqRkeLapGChIglfmgaktlneq-----  
 PhaC\_P2N10\* lVPaaAsqalgRai---gsriryqLaieGgIvGV-----  
 PhaC1<sub>16</sub> ItPWdAcYKsvnLL---gGksqFVLvnSGHlqtlvcPPgKqKaSyqT--aDvLFeDphEW  
 PhaC2<sub>150</sub> ItPWdgcYKsvnLL---gGesqFVLvnSGHlqtlvcPPgKqKaSyqT--gEtLFeDphEW  
 PhaC5 ItPWkcyrtTQLL---gGpstFVLsnaGvVaslNPPgnpKatyWl--gpKpddDpEaW  
 PhaC20 ItPWktcYatTQLL---tGereFVLSSSGHIqsIvNPPgnpKarfiT--NpsfPaDpDDW  
 PhaC10\* -----  
 PhaC2<sub>1</sub> ItPWdAvYrSTlLL---gGerRFVLSnSGHvqsIvNPPsnpKanfve--stKlsgDpraW  
 PhaC7\* -----  
 PhaC\_P11N2 lVPWrsYrgrRLL---gGksRFVLGASGHlAGVINPPAKgKRnyWvngDagkasADEW  
 PhaC25 lVPiggaYaSTQLL---pGKkRFVmgASGHlAGVINPPAKKRSHWiRddgKfPktqaEW  
 PhaC\_P1N3 lVPiggaYaSTQLL---KGKkRFVmgASGHlAGVINPPAKKRSHWiRddgKfPktqaEW  
 PhaC14 lVPWHAAYESTTil---ntKrRFVLGASGHlAGVINPaAKnKRSHWv--NDKLPpiADEW  
 PhaC8 lVPWHAAYESTTil---nsKrRFVLGASGHlAGVINPasKnKRSHWT--NDKLPasADDW  
 PhaC23 lVPWHAAYESTTil---nsKrRFVLGASGHlAGVINPasKnKRSHWT--NDKLPasADDW  
 PhaC21 lVPWQAAYESTTil---nsKrRFVLGASGHlAGVINPasKnKRSHWT--NDKLEpsADDW  
 PhaC4 lVPWtgAYaSTaLL---RGKkRFVLGASGHlAGVINPPAKKRSHWT--NtttPkDpqrW  
 PhaC\_P4N10 lVPWtgAYaSpaLL---RGKkRFVLGASGHlAGVINPPAKKRSHWT--NtttPkDpqrW  
 PhaC19 lVPWQAAYESTKil---KGKvRFVLGASGHlAGVINPPAaKKRSHWv--dgKlGatAgQW  
 PhaC6 lVPWQAAYESTKil---KGKvRFVLGASGHlAGVINPPAaKKRSHWv--dgKlGatAgQW  
 PhaC3 lVPWQAAYESTKil---KGKvRFVLGASGHlAGVINPPAaKKRSHWv--dgKlGatAgQW  
 PhaC51 lVPWQAAYESTKil---KGKvRFVLGASGHlAGVINPPAaKKRSHWv--dgKlGatAgQW  
 PhaC27 lVPWQAAYESTKil---KGKvRFVLGASGHlAGVINPPAaKKRSHWv--dgKlGatAgQW  
 PhaC18 lVPWQAAYESTKil---KGKvRFVLGASGHlAGVINPPAaKKRSHWv--dgKlGatAgQW  
 PhaC\_P8N7 lAaPaasvWkiThyf---QGplRFVLagSGHlAGVINPPeAQkyqyWT--NDgkaetlDQf  
 PhaC\_P2N8 lAaPwrsvYagaRlp---gGKtRFVLagSGHlAGVINPPAahKyqHWT--cDsLftspDEW

PhaC2<sub>16</sub> -----WariSAWivqgAa-----  
 PhaC2<sub>50</sub> -----WariSAWivqgg-----  
 PhaC\_P2N10\* -----fAGR-----hglaiVpPAisRflTAtsiKKpr  
 PhaC1<sub>16</sub> LehsKattGSWWemWIKWaserhGpkkPA--PKApGs-arfpSlgaAPGeYVrA-----  
 PhaC2<sub>150</sub> LeqsKptiGSWWdlvkwWasqrsGakrPA--PKApGs-akfppIgaAPGdYVrt-----  
 PhaC5 lqeAtkhtGtWwvWvdWagkrsGRqkLA--PKAlGN-atYKArEkAPGTYVmerA-----  
 PhaC20 aAgAtphtGSWWehWtdWLatrsGerkaA--PKSlGs-ERHqplaaAPGTYVhg-----  
 PhaC10\* -----  
 PhaC2<sub>1</sub> yydAKQvdGSWWTqWlgwiQersGalKet--hmAlGN-QnYppmEaAPGTYVrvr-----  
 PhaC7\* -----  
 PhaC\_P11N2 LeTAsEvsGSWWplWaeWlGgfgGkkVtA--rrrlGs-adsfplEPAPGRYVKeKA-----  
 PhaC25 lAgAtEqPGSWWTDWAcqWlKgHAGKqVPA--PKAYGdGKKYKAIEPAPGRYVKArA-----  
 PhaC\_P1N3 lAgAtEhPGSWWTDWsnWlKtHAGKqIPA--PKSYGkGsaYKAIEsAPGRYVKArA-----  
 PhaC14 fvSAtErPGSWWpDWSAWLKaqqGAmIaA--PKgYGN-KtFKaIEPAPGRYVKqKA-----  
 PhaC8 MASAIekPGSWSDWSAWLktqGpMVaA--PKgYGN-KKlKAIEPAPGRYVKqKA-----  
 PhaC23 MAgAiEkPGSWSDWSAWLktqGpMVaA--PKgYGN-KmlKAIEPAPGRYVKqKA-----  
 PhaC21 MASAtEkPGSWSDWSAWLktqGpMVaA--PKgYGN-KKlKAIEPAPGRYVKqKA-----  
 PhaC4 lAgAtEsPGSWWTDWAAWlLaplsGKeIaA--PKAlGs-RKYKSIEPAPGRYVKqKA-----  
 PhaC\_P4N10 lAgAtEsPGSWWTDWAAWlLaplsGKeIaA--PKAlGs-RKYKSIEPAPGRYVKqKA-----  
 PhaC19 lDSAKEvPGSWWTDWAWLkPHAGKqVPA--PKtYGN-KsHKvIEPAPGRYVKAKA-----  
 PhaC6 lDSAKEvPGSWWTDWAWLkPHAGKqVPA--PKtYGN-KsHKvIEPAPGRYVKAKA-----  
 PhaC3 lDSAKEvPGSWWTDWAWLkPHAGKqVPA--PKtYGN-KsHKvIEPAPGRYVKAKA-----  
 PhaC51 LeSAKdVPGSWSDWAAWlKpHAGKMVPA--PKAYGd-RsHKAIEaAPGRYVKAKA-----  
 PhaC27 LeSAKdVPGSWSDWAAWlKpHAGKMVPA--PKAYGd-RsHKAIEaAPGRYVKAKA-----  
 PhaC18 LeSAKdVPGSWSDWAAWlKpHAGKMVPA--PKAYGd-RsHKAIEaAPGRYVKAKA-----  
 PhaC\_P8N7 vAgAtEhkGSWWpDWieWigaqdGKrVPAkgarvpGk-gKlKAIEaAPGSYVKAr-----  
 PhaC\_P2N8 qAgAaahaGSWWpDWkfwLgrrAGgkVMA---rtPgS-ggltAIEdAPGSYVrmtshQE--

### Supplementary 3

A.



B.

	39	44
R. eutropha PhaJ4a	-MrtiaSLeeLEglqGQEvavSDWIEvtQqqvngFAdaTGDhQWIIHIDVERAKKESPyGG	
R. eutropha PhaJ4b	-MktyeNiadLqplVGevIGtSEWLaLDQARIntFAdaTGDhQWIIHIDVERAKn-gPFGa	
P. putida PhaJ4	--mphvpvteLsqyVGkELGhSEWLKIDQqRInlFAeaTGDfQfIHVDpEKAAK-tPFGG	
PhaJ20	mteEaySvrtIpqfaGrELGvSDWLRIDQeRIdrFAACTdDrQWIIHIDVERAeRESPFGt	
PhaJ50	--mstlkMatLtekVQELGttrWvalDQSRIQaFAdCTGDgQWIIHIDVERAKKESPFGG	
PhaJ16	--msaltMqnIaakIGeELGvtgWttLDQARIQaFAdCTGDgQWIIHIDVERAKKESPFGG	
PhaJ5	-MadaltLatIEsfVGrELGeSgWvRmDQqRIdaFAeCTGDrQWIIHIDVERAAREgPFGG	
A. caviae PhaJ	-----MsaqsleVGQkarlSk--RfgaAevaaFAalseDfnplHlDpafAAt-taFer	
P. putida PhaJ1	-MsQvtNtpyealeVGQkaeyk--sVEerdIqlFAAmsGDhnpvHlDaEfaAK-SmFre	
P. aeruginosa PhaJ1	-MsQvqNipyaleVGQkaeyts--sIaerdIqlFAAvsGDnrvpHlDaayAAt-tqFke	
	71	
R. eutropha PhaJ4a	PIAHGfLTLSSLpKfmbHNalhmPskI--GVNYGLnRVRfTAPVpVGSklRaRIKLLkver	
R. eutropha PhaJ4b	PIAHGfLTLSSLpafthsayrirms-stGVNYGLDKVRFpAPVpVdSllRaqfKLMSyea	
P. putida PhaJ4	tIAHGfLTLSSLiPklieDilvlPqgLkMvVNYGLDsVRFiqPVKvdsrVrLkVKLgvevte	
PhaJ20	tvAHGLLvLSSLlErfgfEVglvPpgVssalNYGfdgVRFvSPVkaGARVRdRVtLLeatd	
PhaJ50	PvAHGMLTLSSLlPmwlfDLpaaPddagailNYGfdKVRFlAPVksGakVRaRIKLLAatP	
PhaJ16	PvAHGMLTLSSLlPwlfELpaaPldagailNYGfdKVRFlAPVkaGsrVRgRIKLLAaaP	
PhaJ5	tIAHGfLTLSiigpaqlDVwiaPagIgtahYGLDKVRFflAPVpaGrnVrtrIKLLaAvea	
A. caviae PhaJ	PIvHGML-LSLfsgllgqqlpgkgsIyLG-----qslsFklPVfVGdeVtaeVevtAlre	
P. putida PhaJ1	rIAHGmfsgaLisaaavactlpgPgtIyLG-----qQmsFqkPVKIGdltlvRleiLeklP	
P. aeruginosa PhaJ1	rIAHGMLsgaLisaaiaatVlpgPgtIyLG-----qt1RFtrPVKLGddlkvelevLeklP	
R. eutropha PhaJ4a	ldplpkspelvgqaqstwevTVEREGsdrPvcVAESIttryg-----	
R. eutropha PhaJ4b	len-----GgaqFkvemmVERQGGSKPvcIAESILrrfP-----	
P. putida PhaJ4	Kkp-----GqwLlkaiaTlEIEGEGEKPAyIAESLsLcfv-----	
PhaJ20	KgE-----GrllvkarhTIEIEGESKPALVAEmLamLItg-----	
PhaJ50	KeK-----GrvLlTqeyTVEIEnEtKPALIAELLVmlIPkaa-----	
PhaJ16	KdK-----GrlllTqeyTVEIEnEtKPALIAELLVmlvPka-----	
PhaJ5	Kgs-----GrLlVtTenvVEIEGhGKPALIAAtaLaMiMP-----	
A. caviae PhaJ	dkp---iatLltrIFTgggalaVtGE-----VvkLP-----	
P. putida PhaJ1	KfK----vRiatnVYnqndelvVaGEAeilaprkkqtVeLvspPnfvas	
P. aeruginosa PhaJ1	KnR----vRmatrVFngagkqvVDGEAeimapeKkLsVeLaelPpisig	

## Supplementary 4