Aminophosphonate mineralisation is a major step in the global oceanic phosphorus redox cycle

Andrew Murphy
University of Warwick
David Scanlan
University of Warwick https://orcid.org/0000-0003-3093-4245
Yin Chen
University of Warwick https://orcid.org/0000-0002-0367-4276
Andrew Bottrill
Medical Research Council Toxicology Unit, University of Leicester https://orcid.org/0000-0002-5182-3643
Gary Bending
University of Warwick
John Hammond
University of Reading https://orcid.org/0000-0002-6241-3551
Elizabeth Wellington
Warwick University
Ian Lidbury (l.lidbury@sheffield.ac.uk)
University of Sheffield https://orcid.org/0000-0001-7190-315X

Article

Keywords: Planktonic Synthesis, Organophosphorus, Aminophosphonate Uptake, Marine Heterotrophs, Mesopelagic and Epipelagic Waters, Regenerated Phosphate

Posted Date: December 18th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-114998/v1

Version of Record: A version of this preprint was published at Nature Communications on July 27th, 2021. See the published version at https://doi.org/10.1038/s41467-021-24646-z.

Aminophosphonate mineralisation is a major step in the global oceanic phosphorus redox cycle

Andrew R. J. Murphy[1], David J. Scanlan[1], Yin Chen[1], Andrew Bottrill[1], Gary Bending[1], John P. Hammond[2], Elizabeth M. H. Wellington[1], Ian D.E.A. Lidbury[3*]

[1] School of Life Sciences, University of Warwick, Gibbet Hill Road, Coventry, UK

[2] School of Agriculture, Policy, and Development, University of Reading, Earley Gate, Whiteknights, Reading, UK

[3]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK

***Corresponding author:** I.lidbury@sheffield.ac.uk

Abstract

The planktonic synthesis of reduced organophosphorus molecules, such as alkylphosphonates and aminophosphonates, represents one half of a vast global oceanic phosphorus redox cycle. Whilst alkylphosphonates tend to accumulate in recalcitrant dissolved organic matter, aminophosphonates do not. Thus, we hypothesised unknown pathways for the uptake of aminophosphonates must exist in seawater. Here, we identify three novel bacterial 2-aminoethylphosphonate (2AEP) transporters, named AepXVW, AepP and AepSTU, whose expression is independent of phosphate concentrations (phosphate-insensitive). AepXVW, is found in diverse marine heterotrophs and is ubiquitously distributed in mesopelagic and epipelagic waters. Unlike the archetypal phosphate-regulated phosphonate binding protein, PhnD, the newly identified AepX is heavily transcribed (~100-fold>PhnD) in the global ocean independently of phosphate and nitrogen concentrations. Collectively, our data identifies a mechanism responsible for the oxidative step in the marine phosphorus redox cycle and suggests 2AEP may be an important source of regenerated phosphate, which is required for oceanic primary production.

Introduction

Phosphonates are reduced organic phosphorus (P) molecules with a carbon (C)-P bond, as opposed to the more common C-oxygen (O)-P ester bonds found in many other organic P molecules[1]. Phosphonates are synthesised as both primary and secondary metabolites in various bacterial, archaeal and eukaryotic organisms[1-7] where they are incorporated into lipids (phosphonolipids) and glycans (phosphonoglycans)[4,8]. A significant proportion can also be released from the cell to facilitate favourable biotic interactions[9]. Thus, they are ubiquitous in terrestrial and aquatic ecosystems[10-14]. Phosphonates also represent a

major fraction of the marine organic phosphorus pool[11,15-17] and recent studies have now identified several cosmopolitan marine microorganisms capable of synthesising significant quantities of these compounds[6,7,9,18,19]. Collectively, this synthesis drives a vast global oceanic phosphorus redox cycle with reduced phosphorus input in the surface ocean estimated to be an order of magnitude greater than pre-anthropogenic riverine phosphorus input[9]. Whilst much attention has focused on the degradation of alkylphosphonates, such as (hydroxy-) methylphosphonate (MPn) and 2-hydroxyethylphosphonate (HEP), as a source of P[16,17,20,21], the potential for aminophosphonates such as 2-aminoethylphosphonate (2AEP) to serve as sources of C and N (Pi-insensitive) has been neglected. However, emerging evidence suggests that Pi-insensitive 2AEP catabolism occurs in nature[22,23]. Notably, the absence of 2AEP from otherwise phosphonate rich high molecular weight (HMW) DOM[16,17], despite its supposed ubiquitous production[9,19,24,25], suggests preferential catabolism of this molecule in comparison to alkylphosphonates.

The C-P bond is resistant to the activity of standard inorganic phosphate (Pi) liberating enzymes such as alkaline phosphatases and requires specific mechanisms to break, such as the C-P lyase[26,27]. Several 2AEP-specific phosphonate degradation systems (phosphonatases) have been characterised (Fig 1A). The C-P lyase, which is a non-specific promiscuous phosphonatase, is only induced in response to Pi-starvation, being regulated by the two-component master regulator of the Pi-stress response regulon, PhoBR[24]. In marine surface waters, genes encoding the C-P lyase are enriched in bacterial genomes found in regions typified by low Pi concentrations[20] where they are also heavily expressed[28]. Recent data has shown Pi-insensitive regulation of 2AEP degradation, facilitated by the 2AEP-specific phosphonatase systems, occurs in a few strains related to marine *Alphaproteobacteria*[22] and a terrestrial *Gammaproteobacteria*[23]. In both cases, a major consequence of Pi-insensitive

2AEP degradation was the remineralisation and release of labile Pi[22], due to the greater cellular demand for N over P and the 1:1 N:P stoichiometry of 2AEP.

To date, only two 2-AEP transport systems have been identified. Both are ATP-binding cassette (ABC) transporters which consist of a periplasmic substrate-binding protein (SBP), an ATP-binding domain protein, and a transmembrane permease. The first is located within the C-P lyase operon (*phnCDEFGHIJKLMNOP*)[29,30] with genes encoding the SBP, ATP-binding domain and transmembrane domains designated *phnD*, *phnC*, and *phnE*, respectively. The second is another ABC-transporter *phnSTUV* shown by Jiang *et al.* to complement a C-P lyase knockout mutant of *E. coli* together with *phnWX*[31,32]. PhnD has a restricted distribution in seawater whose abundance is highly correlated with regions of Pi-limitation[20]. However, no 2AEP transporter has been identified in the majority of bacteria possessing PhnWX and PhnWAY phosphonatases, which is surprising, given the fact 2AEP is a charged molecule and ubiquitous in marine and terrestrial ecosystems[1,20].

Here, we sought to identify transporters, which provide superb molecular tools for investigating the *in situ* cycling of specific environmental metabolites[28,33-35], required for 2AEP catabolism in environmental bacteria. Through combining laboratory-based molecular and genetic analyses with environmental meta-omics, we identified three novel transporters which have a role in 2AEP uptake and revealed Pi-insensitive 2AEP catabolism is widespread in the global ocean, likely representing a major step in the marine phosphorus redox cycle.

Results

Pseudomonas putida BIRD-1 possesses a novel Pi-sensitive 2-aminoethylphosphonate ABC transporter, AepXVW

We recently identified several candidate 2-AEP transporters (ABC-type) in *Pseudomonas* rhizobacteria that contain the PhnWX phosphonatase but lack both the archetypal PhnCDE transporter and PhnSTUV[36] (Fig 1A). In *Pseudomonas putida* BIRD-1 (hereafter BIRD-1), a periplasmic substrate binding protein associated with one of these putative transporters (PPUBIRD1_4925), which we hereafter refer to as AepX, was induced under Pi-deplete growth conditions in a PhoBR-dependent manner[36]. AepX belongs to the same family (pfam13343) as PhnS, iron and sulphate SBPs but is clearly distinct (Coverage = 40%, Identity = 25.09%, $1.1e^{-05}$) (Fig 1B).

BIRD-1 was capable of growth on 1.5 mM 2AEP as either a sole N, P, or N and P source, the latter resulting in mineralisation of Pi which was subsequently exported from the cell (Fig 1C, Fig S1). Mutagenesis of *phnWX* confirmed this phosphonatase was essential for 2AEP catabolism under both growth conditions in this bacterium (Fig S2A and S2B). Next, we investigated if AepX and its corresponding ABC transporter components, the ATP binding domain protein (AepV), and the permease domain protein (AepW) were essential for its growth on 2AEP. Surprisingly, deletion of *aepXVW*^{BIRD} had no effect on growth as a sole N source (Fig 1C). However, the mutant (Δ*aepXVW*^{BIRD}*::Gm*) had significantly ($P < 0.0001$) attenuated growth on 2AEP as a sole P source (Fig 1C). The growth defect observed during growth on 2AEP as a sole P source was largely restored by complementing the mutant with a plasmid-encoded native homolog (Fig. 1C). These data suggest that whilst the *aepXVW* transporter is not essential, it is involved in 2AEP uptake as a sole P source but is not involved

in its growth as an N source. Therefore, another 2AEP transport system must also exist in this bacterium.

Identification of two Pi-insensitive 2-aminoethylphosphonate transporters in *P. putida* BIRD-1

Next, by subjecting the BIRD-1 *ΔaepXVW*^{BIRD}*::gm* mutant to comparative proteomics, we identified a major facilitator-type (MFS) transporter, (PPUBIRD1_3129), hereafter referred to as AepP for 2-**a**mino**e**thyl**p**hosphonate **p**ermease, whose expression was significantly increased during growth on 2 AEP as a sole N source (Pi-insensitive) (Fig S3). AepP is related to the glycerol-3-phosphate (G3P): Pi antiporter, GlpT[37,38] (identity to GlpT = 27.75%, 9e^{-37}), a member of the organophosphate: phosphate antiporter (OPA) family of MFS transporters. AepP shares conserved residues essential for binding Pi and the phosphate moiety of G3P[39-42] with GlpT, whereas residues that impact the binding affinity to the glycerol moiety of G3P but not Pi are not conserved[41] (Fig S4). Mutation of *aepP* in either the wild type parental strain (Δ*aepP*, Fig1D) or the *aepXVW* mutant (Δ*aepXVW:Gm:aepP*, Fig 1E) led to an inability to grow on 2AEP as sole N source. Subsequent complementation of Δ*aepP* with its native homolog (Δ*aepP* +*pBB:aepP*) restored growth on 2AEP as a sole N source (Fig 1D). Interestingly, delayed but significant growth on 2AEP as a P source still occurred in the Δ*aepXVW:Gm:aepP* double mutant, revealing the presence of a third 2AEP transporter in this bacterium.

To identify the unknown 2AEP transporter, we reanalysed our proteomics data (Fig S3). Another substrate binding protein (PPUBIRD1_3891) containing the same pfam domain (pfam13343) as AepX (Fig 1B), hereafter named AepS, was constitutively expressed in all growth conditions. In order to uncover the role of AepS in the utilisation of 2AEP as a sole P source, a triple mutant Δ*aepXVW:Gm;aepP:aepSTU* was generated in BIRD-1 (Fig 1E). This

135 triple knockout mutant was unable to grow on 2AEP as a sole P source (Fig 1E), suggesting
136 AepSTU is a functional 2AEP transporter. However, generation of a single *aepSTU* knockout
137 mutant did not affect Pi-sensitive growth compared to the wild type (Fig S2C), suggesting
138 AepXVW is the major transporter involved in Pi-sensitive 2AEP uptake and AepSTU only has
139 an auxiliary role in 2AEP uptake. Despite expression of AepS during Pi-insensitive growth (Fig
140 S3), this transporter was unable to facilitate growth on 2AEP as a sole N source in the absence
141 of AepP. Together, these data reveal the presence of three novel and differentially regulated
142 2AEP transporters in BIRD-1. AepXVW is the primary Pi-sensitive 2AEP transporter with
143 AepSTU having an auxiliary role, whereas AepP is essential for Pi-insensitive 2AEP growth.

144

145 **AepXVW is found in several marine bacteria capable of Pi-insensitive mineralisation and is**
146 **a functional 2-aminoethylphosphonate transporter**

147 Using PPUBIRD1_4925 (AepX) as the query, we scrutinised the genomes of several
148 isolates related to marine *Rhodobacteraceae* (*Stappia* spp., *Terasakiella* spp.,
149 *Falsirhodobacter* spp.) capable of Pi-insensitive phosphonate catabolism[22]. ORFs encoding
150 orthologs of AepX were identified in the genomes of *Stappia stellulata* DSM 5886 and
151 *Terasakiella pusilla* DSM 6293, in addition to several other marine roseobacter strains:
152 *Aliiroseovarius crassostreae* (DSM 16950), *Aliiroseovarius sediminilitoris* (DSM 29439), *Shimia*
153 *marina* (DSM 26895), and *Thalassobius aestuarii* (DSM 15283) (Fig 2A). We also found an
154 orthologous ORF in the model rhizosphere alphaproteobacterium *Sinorhizobium meliloti*
155 strain 1021 that is capable of 2AEP catabolism via a phosphonatase[43] (Fig 2A). In all cases,
156 ORFs encoding AepXVW were located adjacent to ORFs encoding the phosphonatases,
157 PhnWAY or PhnWX, strongly suggesting a role in 2AEP transport. *S. stellulata* DSM 5886, *A.*

crassostreae DSM 16950, *A. sediminilitoris* DSM 29439, *S. marina* DSM 26895, and *T. aestuarii* DSM 15283 were all capable of growth on 2AEP as either the sole N or P source (Table S1). Indeed, both *Aliiroseovarius* strains lack other characterised 2AEP transport and degradation systems (Table S2). As previously reported, Pi was exported from cells and accumulated in the medium during growth on 2AEP as a sole N source (Fig S5A).

To confirm that *Stappia* AepXVW can take up 2AEP, we complemented the BIRD-1 null 2AEP transporter mutant (Δ*aepXVW:aepP:aepSTU:gm*) with this transporter fused with the $aepXVW^{BIRD\text{-}1}$ promoter (Fig 2B). This duly restored growth of the triple mutant confirming *S. stellulata* AepXVW is also a functional 2AEP transporter.

AepXVW is highly expressed in the marine bacterium *Stappia stellulata* during Pi-sensitive and Pi-insensitive growth on 2-aminoethylphosphonate

In BIRD-1, AepXVW was only involved in Pi-sensitive growth whilst AepP was induced during Pi-insensitive metabolism (Fig 1). However, *S. stellulata* lacks AepP but is still capable of Pi-insensitive growth and Pi export (Fig S5A). In addition to *aepXVW* and genes encoding the 2AEP phosphonatase (*phnWAY*), *S. stellulata* also possesses genes (*phnCDEFGHIJKLMN*) encoding the P-regulated C-P lyase operon and we also confirmed this strain grew on several other alkylphosphonates as sole P source (Fig S5B). Therefore, to determine which transport and degradation systems were upregulated during growth on 2AEP as either a sole N or P source, we performed comparative proteomics. Unlike BIRD-1, AepX was abundantly expressed during growth on 2AEP as either a sole N or P source, as was the phosphonatase (PhnWAY), whilst the C-P lyase operon was not (Fig 3). Importantly, whilst we detected several general nitrogen stress-response proteins induced under Pi-insensitive growth, we

didn't identify any other potential 2AEP transporters (Table S3). Therefore, unlike in BIRD-1, AepXVW likely represents the major route for 2AEP uptake in this bacterium.

aepX **and** ***aepP*** **are found in distantly related and cosmopolitan bacterial taxa**

Using the Integrated Microbial Genomes/Microbiomes from the Joint Genome Institute (IMG/M/JGI) database, we identified ORFs encoding AepX and AepP (but not AepS) in genomes retrieved from both taxonomically divergent isolates as well as single amplified genomes (SAGs) and metagenome assembled genomes (MAGs), which revealed an unexpected diversity for these substrate binding proteins (Fig 4). For AepX, this included cosmopolitan marine *Alphaproteobacteria* other than *Rhodobacteraceae*, marine *Deltaproteobacteria*, as well as marine *Vibrio* spp. AepX was also found in terrestrial *Betaproteobacteria*, *Firmicutes*, and other gram-positive bacteria (Fig 4). AepX was partitioned into several subclades, with AepX^{Stappia} and AepX^{BIRD} well separated (Fig 4). Many taxonomically divergent AepX ORFs were co-localised with ORFs encoding the various phosphonatase systems or the C-P lyase, supporting a role in 2AEP transport (Fig 4).

AepP was also found in a wide range of phylogenetically divergent taxa, such as *Acidobacteria* (*Granuliella mallensis*) and *Bacteroidetes* (*Kriegella aquimaris*), *Actinobacteria* (*Streptomyces albulus* CCRC 11814) and *Verrucomicrobia* (*Haloferula* sp. BvORR071 and *Verrucomicrobia sp.* SGGC AC-337 J20) (Fig S6). However unlike AepX, AepP was not found in cosmopolitan marine bacteria. Again, for all of these strains ORFs encoding AepP were co-localised with ORFs encoding phosphonatases (Fig S6). Notably, AepP was found in fewer marine isolates compared to AepX.

aepX gene and transcript abundance is far greater than the archetypal *phnD/phnS* in the global ocean

Using the TARA oceans OM-RGCv2+G metagenome (MG) and OM-RGCv2+T metatranscriptome (MT) datasets[44], we calculated the abundance of our newly-identified *aepX*, *aepS* and *aepP* transporters and compared this with *phnS* and the archetypal phosphonate transporter *phnD*, whose gene abundance in seawater was recently calculated[20]. We analysed data from both the epipelagic and mesopelagic zones where phosphonate mineralisation is believed to occur[15]. Across all oceanic sampling sites in both the epipelagic and mesopelagic, *aepX* gene and transcript abundance was significantly greater (MG; post-hoc Dunn's test $z = 10.4$, $p<0.001$ and $z = 4.8$, $p<0.001$, respectively) than *phnD* (Fig 5A and B). On average, in the mesopelagic almost 10% of bacterial cells possess *aepX* whilst only ~0.3% and 0.5% possess *phnD* and *aepP*, respectively (Fig 5A). *aepX* transcription was 40-fold and 350-fold greater than *phnD* in the epipelagic and mesopelagic, respectively (Fig 5B). The majority of *aepX* sequences were related to the cosmopolitan *Alphaproteobacteria* and *Deltaproteobacteria* (Fig 4). We confirmed that these abundant environmental sequences were also co-localised with phosphonate degradation genes (Fig 4). In broad agreement with *aepX*, the cumulative transcription of the two phosphonatase markers *phnA* and *phnX* is significantly greater than the C-P lyase marker *phnJ* (Kruskal-wallis $X^2 = 206.6$, $p<0.001$) strengthening this observation that 2AEP mineralisation is a major oceanic process.

The gene abundances of *aepP, aepS* and *phnS* were all significantly lower than both *aepX* and *phnD* (post-hoc Dunn's test $z = 13.1$ and 9.0, $p<0.001$, $z = 14.5$ and 12.6, $p<0.001$, and $z = 9.9$ and 8.9, $p<0.001$ respectively) in the epipelagic, whilst only *aepS and phnS* were

224 significantly lower in the mesopelagic (post-hoc Dunn's test z = 10.1 and 10.8, p<0.001, and z
225 = 4.6 and 5.5, p<0.001 respectively) (Fig 5A and S7).

226 Unlike *phnD* and *aepP*, *aepX* abundance was comparable across all oceanic regions
227 within both MG and MT at each depth suggesting 2AEP mineralisation is a ubiquitous process
228 in seawater (Fig 5C &D). For all sites at each depth, the relative abundance of *aepX* transcripts
229 was always significantly greater (Wilcoxon rank sum W=3771, p<0.001, estimated log2
230 difference = 2.24 (95CI 1.97-2.52))) than its own gene abundance. For *phnD*, we observed
231 significantly greater transcript abundance compared to its own gene abundance only in the
232 Mediterranean Sea, a region typified by Pi-depletion (Wilcoxon rank sum W=0, p<0.001,
233 estimated log2 difference = 3.11 (95CI 1.90-4.42)). Finally, in agreement with previous work[20],
234 *phnD* gene abundance was inversely correlated (R^2 = 0.340, p<0.001) with standing stock
235 concentrations of Pi (Fig 5E) as was *phnD* transcript abundance (Fig 5F). In contrast, *aepX* and
236 *aepP* gene abundance were positively correlated (R^2 = 0.098, p<0.001 and R^2 = 0.291, p<0.001,
237 respectively) with Pi concentration (Fig 5E), whilst no significant relationship between Pi
238 concentration and *aepX*/*aepP* transcription was found (Fig 5F) suggesting their activation is
239 independent of Pi in seawater globally.

240 To better understand the parameters controlling 2AEP catabolism in the ocean, we
241 compared both gene abundance and transcript in relation to R*, a measure of N vs P limitation
242 calculated as $[NO_2]+[NO_3]- 16[PO_4]$ (adapted from Smith *et al*.[45]). As expected, *phnD* gene and
243 transcript abundance were positively correlated with R* (R^2 = 0.168, p<0.001 and R^2 = 0.197,
244 p<0.001, respectively (Fig 5G and 5H), i.e. regions typified with Pi depletion, though notably
245 Pi concentration alone was a better predictor of *phnD* gene abundance (Fig 5E). However,
246 *aepX* and *aepP* gene abundance was (weakly) inversely correlated with R* (R^2 = 0.029, p<0.05

and R^2 = 0.108, p<0.001, respectively) (Fig 5G) and no significant relationship was found between R* and *aepX/aepP* transcript abundance (Fig 5H). These data are consistent with the proteomic response of *S. stellulata* under laboratory conditions and suggests AepX is induced in the presence of 2AEP (substrate-inducible) and not in response to nutrient limitation.

Discussion

Both phosphonate biosynthesis[1,3,5,46] and catabolic[1,23,25,43,47-52] genes are ubiquitous in marine, soil and gut microbiomes, suggesting phosphonate cycling is widespread in nature. In contrast, the uptake of these molecules is comparatively understudied, with only two characterised ABC transport systems confirmed, both of which are linked solely to P-acquisition[29-31]. The abundance of these Pi-sensitive transporters in marine systems is not equivalent to the abundance of catabolic genes[20], especially those (*phnWAY*) recently shown to be involved in Pi-insensitive catabolism[22]. This would suggest our knowledge of the microbial uptake of phosphonates, particularly 2AEP, is incomplete. Using transporter expression as a proxy for the cycling of specific nutrients has significantly advanced our understanding of *in situ* biogeochemical cycling[33,34,53]. This molecular approach helped resolve the biogenesis of the climate-active gas methane in oxygenated surface waters, driven through the uptake and degradation of methylphosphonate[21,28]. Thus, a gap in mechanistic knowledge on 2AEP metabolism impairs our ability to survey the *in situ* cycling of reduced organophosphorus compounds, especially when high resolution separation of such compounds is difficult[9,54]. Here, identification of novel Pi-insensitive 2AEP transporters allowed us to develop molecular markers to investigate the cycling of 2AEP on a global scale and compare these with previously characterised Pi-sensitive markers.

Genomic and biochemical analyses have revealed 2AEP, MPn and HEP are ubiquitously synthesised in the marine environment in relatively large quantitities[1,3,9,25,54]. However, several studies have shown 2AEP is absent in 'semi-labile' DOM whilst alkylphosphonates, such as MPn and HEP, tend to accumulate[16,17,55]. Collectively, this would suggest 2AEP is more susceptible to microbial mineralisation[22] and thus shorter residence times. Here, we reveal pathways for the Pi-insensitive uptake (*aepX*) and catabolism (*phnA*) are expressed at significantly higher levels across the global ocean than the Pi-repressible *phnD* and *phnJ*, providing a clear mechanism for this phenomenon. Our data also suggests 2AEP is preferentially mineralised independently of both N and P status explaining why phosphonates are metabolised in regions where phosphate concentrations are high enough to repress C-P lyase-expression[20,21,56]. Together, this adds further weight to the notion that phosphonates are rapidly cycled between reduced and oxidised forms[9,22]

Substrate inducible expression of 2AEP catabolic genes irrespective of nutrient status has previously been shown to mineralise organic N to ammonium which can cross feed into another bacterium[57]. Assuming the link between Pi-insensitive expression of AepXVW, PhnWAY and subsequent cellular export of mineralised Pi observed in our laboratory cultures and as demonstrated by Chen et al.[22] is comparable to *in situ* metabolism, we provide a clear mechanism to support the hypothesis that phosphonates are a source of regenerated Pi throughout the water column[15,22], especially at depths where AepP is also found in higher abundances. Thus, aminophosphonates likely represent another source of regenerated Pi from DOP, a mechanism which is important for maintaining biological production in Pi-deplete regions of the ocean[58,59].

In summary, this study identified three novel 2AEP transporters in marine and terrestrial bacteria capable of Pi-insensitive 2AEP catabolism. One of these, AepXVW, is the most abundant phosphonate transporter in seawater and is ubiquitously transcribed at high levels suggesting 2AEP mineralisation is a major process in the marine organic C, N and P cycles and may present a significant source of regenerated Pi available for oceanic production. This conclusion is strengthened by two key observations: 1) AepX transcription is not repressed by standing stock concentrations of Pi, and 2) extracellular Pi export occurs during Pi-insensitive 2AEP metabolism. Thus, we provide further evidence for the role of low molecular weight phosphonates acting as a phosphorus currency between autotrophic and heterotrophic microbes[9].

Materials and Methods

Bacterial strains and growth conditions

Pseudomonas strains used in this work were maintained on Luria Bertani (LB) agar (1.5% w/v) medium at 30°C. *Stappia stellulata* and the *Roseobacter* strains were maintained on Marine Broth agar (1.5% w/v) medium at 30°C. *Pseudomonas* mutants and complemented mutants were maintained on similar plates containing the appropriate antibiotic. For all growth and proteomics experiments cultures were grown in an adapted Minimal A medium[36] using Na-Succinate (20 mM) as the sole carbon source and, where applicable, 10mM NH_4Cl was added as the sole N source. 2AEP and KH_2PO_4 were added to a final concentration of 100 µM or 1.5mM as specified in the text and figure legends. *Pseudomonas* strains were pre-cultured in minimal medium A containing 100 µM Pi and 1.5 mM NH_4Cl to ensure adequate growth while minimising the potential for carryover of residual nutrients into experimental cultures. Co-culture experiments were carried out according to the protocol described in[60]. Culture experiments were performed using a FLUOStar Omega 96-well plate reader using Sarstedt 96-well plates incubated at 30°C, shaking at 200 rpm. For 2AEP as N source proteomics, *S. stellulata* was grown in modified defined marine ammonium mineral salt (MAMS) media, lacking ammonium, and where HEPES replaced the phosphate buffer[61]. For all other growth and proteomics experiments these marine bacteria were grown in Sea Salts media[62], where HEPES replaced the phosphate buffer.

Generation and complementation of *Pseudomonas* mutants

Mutants were generated and complemented via the protocols outlined in[36,60]. A full list of strains, plasmids, and primers used in this study is presented in Table S4. Briefly, regions of genomic DNA at the 5' and 3' end of each deletion target were amplified, along with in some cases the gentamicin resistance cassette from p34S-Gm[63]. Other mutants were

constructed to be marker-less, and so lacked this resistance cassette. DNA fragments were ligated into linearised pk18mob*sacB*[64] using the HiFi DNA Assembly Kit (New England Biolabs, Hitchin, UK) according to the manufacturer's instructions. For the construction of plasmids for complementation, genes and their corresponding promoters were cloned into linearised pBBR1McS-km again using the HiFi DNA Assembly Kit.

Plasmids were subsequently electroporated into *Escherichia coli* S17.1 and mobilized into *Pseudomonas* by conjugation. Transconjugants were selected via gentamicin (50 µg ml^{-1}) or kanamycin (50 µg ml^{-1}) selection and chloramphenicol (10 µg ml^{-1}) counter-selection. Single crossover transconjugants were identified by PCR, and double crossover mutants selected via plating on LB containing either gentamicin or no antibiotic, and 10% (w/v) sucrose. Homologous recombination was confirmed by PCR and Sanger sequencing. Complemented mutants were selected using kanamycin with chloramphenicol counter-selection.

Proteomics preparation and analysis

To identify proteins involved in 2AEP uptake and catabolism in *S. stellulata*, total protein (n=3 for each treatment), was retrieved by sampling cell cultures (OD_{540} 0.8-1.0) and pelleting cells (centrifugation at 16200 x g for 5 mins). Cell lysis was achieved via boiling in 100µl LDS buffer (Expedeon) prior to loading 20µl onto a 4-20% Bis-Tris SDS precast gel (Expedeon). For enrichment of the membrane protein fraction of the *P. putida* BIRD-1 *aepXVW::gm* mutant, we adapted the methods outlined in[36]. Briefly, 40ml cells were grown to an OD_{600} of 0.8-1.0, a volume equivalent to 5 OD_{600} units was centrifuged at 2000 x g for 10 mins, resuspended in 1ml 50mM Tris-HCl, pH 7.6, and centrifuged at 2000 x g for 10 mins. Pellets were resuspended in 500µl 200mM $MgCl_2$, 50mM Tris-HCl, pH 7.6 and incubated for 30 mins at 30°C with gentle shaking. Cells were cooled on ice for 5 mins and incubated at

room temperature for 15 mins, then centrifuged at 8000 x g for 10 mins at 4°C. Pellets were washed in 1ml 50mM Tris-HCl, pH 7.6, and again centrifuged at 8000 x g for 10 mins at 4°C. Pellets were resuspended in 0.5ml 50mM Tris-HCl, pH 7.6, and sonicated for 30 s twice, on ice. The solution was centrifuged at 2000 x g for 15 mins at 4°C, and supernatants were then ultracentrifuged at 120000 x g for 45 mins at 4°C. Following this, the supernatant was discarded and the pellet resuspended in 50µl LDS buffer (Expedeon) prior to loading 20µl onto a 4-20% Bis-Tris SDS precast gel (Expedeon). Gel sections were de-stained with 50mM ammonium bicarbonate in 50% ethanol, dehydrated with 100% ethanol, reduced and alkylated with Tris-2-carboxyethylphosphine (TCEP) and iodoacetamide (IAA), washed with 50mM ammonium bicarbonate in 50% ethanol and dehydrated with 100% ethanol prior to overnight digestion with trypsin. Peptides were extracted and analysed using an Orbitrap Fusion Ultimate 3000 RSLCNano System (Thermo Scientific) in electrospray ionization mode at the Warwick Proteomics Research Technology Platform.

Resulting MS/MS files were searched against the relevant protein sequence database (*P. putida* BIRD-1 , NC_017530.1, *S. stellulata*, GCF_000423715.1) using MaxQuant[65] with default settings and quantification was achieved using Label Free Quantification (LFQ). The proteomics analysis software Perseus (1.6.12)[66] was used to identify differentially expressed proteins based on LFQ values, using a False Discovery Rate (FDR) of 0.01. Identified proteins were retained if they were present in at least two biological replicates within a treatment. Missing (N/A) values were imputed from a normal distribution using the default parameters. Differential expression was identified by two-sample Student's *t*-test, using an s0 constant of 0.1, or ANOVA, where appropriate.

Bioinformatics analyses of *aepXVW*/*aepP*

AepX homologs were identified using the IMG/JGI (https://img.jgi.doe.gov/) 'Customized Homolog Display' search tool. Strains containing homologs were identified (cut-off values: e^{-70}, min. identity 40%), preferentially from type strains. In addition, representative strains from soil and marine environments were added to this list (see Table S2). Protein sequences were aligned using ClustalOmega[67] and profile Hidden Markov Models (pHMMs) were constructed from these sequences using the hmmbuild function of hmmer 3.3 (http://hmmer.org)[68]. The previously characterised *Escherichia coli* K-12 PhnD[69] and the SAR11 clade isolate *Pelagibacter* sp. HTCC7211 PhnD[21] showed surprisingly low sequence homology (BLAST %ID 28.46%, query coverage 76%, e-value 2e-25). We therefore developed two pHMMS for PhnD to reflect this. There was no overlap in environmental sequences retrieved from each search using either hmm model. Therefore, abundance counts for each PhnD form were combined together as a collective PhnD group. These pHMMs were used to search the TARA ocean metagenome (OM-RGC_v2_metaG) and metatranscriptome (OM-RGC_v2_metaT) via the Ocean Gene Atlas web interface[44], using a stringency of $1E^{-80}$. Sequence abundances were expressed as average percentage of genomes containing a copy by dividing the percentage of total mapped reads by the median abundance (as a percentage of total mapped reads) of 10 single-copy marker genes[70] for both MG and MT. The pHMMs were used to search the soil MG via hmmsearch[68] using the same stringency as above. Similarly, abundances were calculated as average percentage of genomes containing a copy as above.

Phylogenetic analyses were performed using IQ-TREE 2[71] using the following parameters: -m TEST -bb 1000 -alrt 1000. Evolutionary relationships were inferred by maximum-likelihood analysis, and visualised using the Interactive Tree of Life (iTOL) v5.6.3 online platform (https://itol.embl.de/)[72]

Statistical analysis

Unless specified above, all statistical analysis was performed using R (version 4.02)[73], within the RStudio programme (version 1.3)[74].

Acknowledgements

We thank the Warwick Proteomics Research Facility, namely Dr. Cleidiane Zampronio for her assistance in generating and processing the mass-spectrometry data. This study was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) under project codes BB/L026074/1 and BB/T009152/1 linked to The Soil and Rhizosphere Interactions for Sustainable Agri-ecosystems (SARISA) programme and a Discovery Fellowship (IL), respectively.

References

1 Villarreal-Chiu, J. F., Quinn, J. P. & McGrath, J. W. The genes and enzymes of phosphonate metabolism by bacteria, and their distribution in the marine environment. *Front Microbiol* **3**, 19, doi:10.3389/fmicb.2012.00019 (2012).

2 Mukhamedova, K. S. & Glushenkova, A. I. Natural Phosphonolipids. *Chem Natural Compounds* **36**, 329-341, doi:10.1023/A:1002804409503 (2000).

3 Yu, X. *et al.* Diversity and abundance of phosphonate biosynthetic genes in nature. *Proc Natl Acad Sci U S A* **110**, 20759-20764, doi:10.1073/pnas.1315107110 (2013).

4 Ju, K.-S., Doroghazi, J. R. & Metcalf, W. W. Genomics-enabled discovery of phosphonate natural products and their biosynthetic pathways. *J Industrial Microbiol & Biotechnol* **41**, 345-356, doi:10.1007/s10295-013-1375-2 (2014).

5 Peck, S. C. & van der Donk, W. A. Phosphonate biosynthesis and catabolism: a treasure trove of unusual enzymology. *Curr Opin Chem Biol* **17**, 580-588, doi:https://doi.org/10.1016/j.cbpa.2013.06.018 (2013).

6 Metcalf, W. W. *et al.* Synthesis of methylphosphonic acid by marine microbes: a source for methane in the aerobic ocean. *Science* **337**, 1104-1107, doi:10.1126/science.1219875 (2012).

7 Born, D. A. *et al.* Structural basis for methylphosphonate biosynthesis. *Science* **358**, 1336, doi:10.1126/science.aao3435 (2017).

8 Hildebrand, R. *The Role of Phosphonates in Living Systems*. (CRC Press, 1983).

9 Van Mooy, B. A. S. *et al.* Major role of planktonic phosphate reduction in the marine phosphorus redox cycle. *Science* **348**, 783, doi:10.1126/science.aaa8181 (2015).

10 Kolowith, L. C., Ingall, E. D. & Benner, R. Composition and cycling of marine organic phosphorus. *Limnol Oceanogr* **46**, 309-320, doi:10.4319/lo.2001.46.2.0309 (2001).
11 Clark, L. L., Ingall, E. D. & Benner, R. Marine phosphorus is selectively remineralized. *Nature* **393**, 426-426, doi:10.1038/30881 (1998).
12 Turner, B. L., Baxter, R., Mahieu, N., Sjögersten, S. & Whitton, B. A. Phosphorus compounds in subarctic Fennoscandian soils at the mountain birch (*Betula pubescens*)—tundra ecotone. *Soil Biol Biochem* **36**, 815-823, doi:https://doi.org/10.1016/j.soilbio.2004.01.011 (2004).
13 Tate, K. R. & Newman, R. H. Phosphorus fractions of a climosequence of soils in New Zealand tussock grassland. *Soil Biol Biochem* **14**, 191-196, doi:https://doi.org/10.1016/0038-0717(82)90022-0 (1982).
14 Cade-Menun, B. J., Navaratnam, J. A. & Walbridge, M. R. Characterizing dissolved and particulate phosphorus in water with 31P nuclear magnetic resonance spectroscopy. *Environ Sci Technol* **40**, 7874-7880, doi:10.1021/es061843e (2006).
15 Clark, L. L., Ingall, E. D. & Benner, R. Marine organic phosphorus cycling; novel insights from nuclear magnetic resonance. *American J Sci* **299**, 724-737 (1999).
16 Repeta, D. J. *et al.* Marine methane paradox explained by bacterial degradation of dissolved organic matter. *Nature Geosci* **9**, 884-887, doi:10.1038/ngeo2837 (2016).
17 Sosa, O. A. *et al.* Phosphonate cycling supports methane and ethylene supersaturation in the phosphate-depleted western North Atlantic Ocean. *Limnol Oceanogr* **65**, 2443-2459, doi:10.1002/lno.11463 (2020).
18 Dyhrman, S. T., Benitez-Nelson, C. R., Orchard, E. D., Haley, S. T. & Pellechia, P. J. A microbial source of phosphonates in oligotrophic marine systems. *Nature Geoscience* **2**, 696-699, doi:10.1038/ngeo639 (2009).
19 Acker, M. *et al.* Phosphonate production by marine microbes: exploring new sources and potential function. *bioRxiv*, 2020.2011.2004.368217, doi:10.1101/2020.11.04.368217 (2020).
20 Sosa, O. A., Repeta, D. J., DeLong, E. F., Ashkezari, M. D. & Karl, D. M. Phosphate-limited ocean regions select for bacterial populations enriched in the carbon-phosphorus lyase pathway for phosphonate degradation. *Environ Microbiol* **21**, 2402-2414, doi:10.1111/1462-2920.14628 (2019).
21 Carini, P., White, A. E., Campbell, E. O. & Giovannoni, S. J. Methane production by phosphate-starved SAR11 chemoheterotrophic marine bacteria. *Nature Comms* **5**, 4346, doi:10.1038/ncomms5346 (2014).
22 Chin, J. P., Quinn, J. P. & McGrath, J. W. Phosphate insensitive aminophosphonate mineralisation within oceanic nutrient cycles. *ISME J* **12**, 973-980, doi:10.1038/s41396-017-0031-7 (2018).
23 Ternan, N. G. & Quinn, J. P. Phosphate starvation-independent 2-aminoethylphosphonic acid biodegradation in a newly isolated strain of Pseudomonas putida, NG2. *Syst Appl Microbiol* **21**, 346-352, doi:10.1016/S0723-2020(98)80043-X (1998).
24 White, A. K. & Metcalf, W. W. Microbial metabolism of reduced phosphorus compounds. *Annu Rev Microbiol* **61**, 379-400, doi:10.1146/annurev.micro.61.080706.093357 (2007).
25 Martinez, A., Tyson, G. W. & Delong, E. F. Widespread known and novel phosphonate utilization pathways in marine bacteria revealed by functional screening and metagenomic analyses. *Environ Microbiol* **12**, 222-238, doi:10.1111/j.1462-2920.2009.02062.x (2010).
26 Wanner, B. L. Molecular genetics of carbon-phosphorus bond cleavage in bacteria. *Biodegradation* **5**, 175-184 (1994).
27 Kononova, S. V. & Nesmeyanova, M. A. Phosphonates and their degradation by microorganisms. *Biochemistry (Mosc)* **67**, 184-195 (2002).
28 Sowell, S. M. *et al.* Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* **3**, 93-105, doi:http://www.nature.com/ismej/journal/v3/n1/suppinfo/ismej200883s1.html (2008).

29 Alicea, I. *et al.* Structure of the *Escherichia coli* phosphonate binding protein PhnD and rationally optimized phosphonate biosensors. *J Mol Biol* **414**, 356-369, doi:10.1016/j.jmb.2011.09.047 (2011).

30 Rizk, S. S., Cuneo, M. J. & Hellinga, H. W. Identification of cognate ligands for the *Escherichia coli phnD* protein product and engineering of a reagentless fluorescent biosensor for phosphonates. *Protein Sci* **15**, 1745-1751, doi:10.1110/ps.062135206 (2006).

31 Jiang, W., Metcalf, W. W., Lee, K. S. & Wanner, B. L. Molecular cloning, mapping, and regulation of Pho regulon genes for phosphonate breakdown by the phosphonatase pathway of *Salmonella typhimurium* LT2. *J Bacteriol* **177**, 6411, doi:10.1128/jb.177.22.6411-6421.1995 (1995).

32 Kim, A. D. *et al.* The 2-aminoethylphosphonate-specific transaminase of the 2-aminoethylphosphonate degradation pathway. *J Bacteriol* **184**, 4134, doi:10.1128/JB.184.15.4134-4140.2002 (2002).

33 Ottesen, E. A. *et al.* Pattern and synchrony of gene expression among sympatric marine microbial populations. *Proc Natl Acad Sci USA* **110**, E488-E497, doi:10.1073/pnas.1222099110 (2013).

34 Lidbury, I., Murrell, J. C. & Chen, Y. Trimethylamine N-oxide metabolism by abundant marine heterotrophic bacteria. *Proc Natl Acad Sci USA* **111**, 2710-2715, doi:10.1073/pnas.1317834111 (2014).

35 Mauchline, T. H. *et al.* Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome. *Proc Natl Acad Sci* **103**, 17933-17938, doi:10.1073/pnas.0606673103 (2006).

36 Lidbury, I. D. *et al.* Comparative genomic, proteomic and exoproteomic analyses of three *Pseudomonas* strains reveals novel insights into the phosphorus scavenging capabilities of soil bacteria. *Environ Microbiol* **18**, 3535-3549, doi:10.1111/1462-2920.13390 (2016).

37 Lemieux, M. J., Huang, Y. & Wang, d. N. Crystal structure and mechanism of GlpT, the glycerol-3-phosphate transporter from *E. coli*. *J Electron Microsc* **54 Suppl 1**, i43-46, doi:10.1093/jmicro/54.suppl_1.i43 (2005).

38 Elvin, C. M., Hardy, C. M. & Rosenberg, H. Pi exchange mediated by the GlpT-dependent sn-glycerol-3-phosphate transport system in *Escherichia coli*. *J Bacteriol* **161**, 1054-1058 (1985).

39 Enkavi, G. & Tajkhorshid, E. Simulation of spontaneous substrate binding revealing the binding pathway and mechanism and initial conformational response of GlpT. *Biochemistry* **49**, 1105-1114, doi:10.1021/bi901412a (2010).

40 Law, C. J. *et al.* Salt-bridge dynamics control substrate-induced conformational change in the membrane transporter GlpT. *J Mol Biol* **378**, 828-839, doi:https://doi.org/10.1016/j.jmb.2008.03.029 (2008).

41 Law, C. J., Enkavi, G., Wang, D.-N. & Tajkhorshid, E. Structural basis of substrate selectivity in the glycerol-3-phosphate: phosphate antiporter GlpT. *Biophysical J* **97**, 1346-1353, doi:10.1016/j.bpj.2009.06.026 (2009).

42 Moradi, M., Enkavi, G. & Tajkhorshid, E. Atomic-level characterization of transport cycle thermodynamics in the glycerol-3-phosphate:phosphate antiporter. *Nature Comms* **6**, 8393, doi:10.1038/ncomms9393 (2015).

43 Borisova, S. A. *et al.* Genetic and biochemical characterization of a pathway for the degradation of 2-aminoethylphosphonate in *Sinorhizobium meliloti* 1021. *J Biol Chem* **286**, 22283-22290, doi:10.1074/jbc.M111.237735 (2011).

44 Villar, E. *et al.* The Ocean Gene Atlas: exploring the biogeography of plankton genes online. *Nucleic Acids Res* **46**, W289-W295, doi:10.1093/nar/gky376 (2018).

45 Smith, A. F. *et al.* Elucidation of glutamine lipid biosynthesis in marine bacteria reveals its importance under phosphorus deplete growth in *Rhodobacteraceae*. *ISME J* **13**, 39-49, doi:10.1038/s41396-018-0249-z (2019).

46 Chin, J. P., McGrath, J. W. & Quinn, J. P. Microbial transformations in phosphonate biosynthesis and catabolism, and their importance in nutrient cycling. *Curr Opin Chem Biol* **31**, 50-57, doi:10.1016/j.cbpa.2016.01.010 (2016).

47 White, A. K. & Metcalf, W. W. Two C-P lyase operons in Pseudomonas stutzeri and their roles in the oxidation of phosphonates, phosphite, and hypophosphite. *J Bacteriol* **186**, 4730-4739, doi:10.1128/JB.186.14.4730-4739.2004 (2004).

48 Ermakova, I. T. *et al.* Organophosphonates utilization by soil strains of *Ochrobactrum anthropi* and *Achromobacter* sp. *Arch Microbiol* **199**, 665-675, doi:10.1007/s00203-017-1343-8 (2017).

49 Hartley, L. E., Kaakoush, N. O., Ford, J. L., Korolik, V. & Mendz, G. L. Characterisation of *Campylobacter jejuni* genes potentially involved in phosphonate degradation. *Gut Pathog* **1**, 13, doi:10.1186/1757-4749-1-13 (2009).

50 Imazu, K. *et al.* Enhanced utilization of phosphonate and phosphite by *Klebsiella aerogenes*. *Appl Environ Microbiol* **64**, 3754-3758 (1998).

51 Mendz, G. L., Mégraud, F. & Korolik, V. Phosphonate catabolism by *Campylobacter* spp. *Arch Microbiol* **183**, 113-120, doi:10.1007/s00203-004-0752-7 (2005).

52 Metcalf, W. W. & Wanner, B. L. Evidence for a fourteen-gene, *phnC* to *phnP* locus for phosphonate metabolism in *Escherichia coli*. *Gene* **129**, 27-32, doi:10.1016/0378-1119(93)90692-v (1993).

53 Hultman, J. *et al.* Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature* **521**, 208-212, doi:10.1038/nature14238 (2015).

54 Frischkorn, K. R. *et al. Trichodesmium* physiological ecology and phosphate reduction in the western tropical South Pacific. *Biogeosciences* **15**, 5761-5778, doi:10.5194/bg-15-5761-2018 (2018).

55 Young, C. L. & Ingall, E. D. Marine dissolved organic phosphorus composition: insights from samples recovered using combined electrodialysis/ reverse osmosis. *Aquatic Geochem* **16**, 563-574, doi:10.1007/s10498-009-9087-y (2010).

56 Benitez-Nelson, C. R., O'Neill, L., Kolowith, L. C., Pellechia, P. & Thunell, R. Phosphonates and particulate organic phosphorus cycling in an anoxic marine basin. *Limnol Oceanogr* **49**, 1593-1604, doi:10.4319/lo.2004.49.5.1593 (2004).

57 Lidbury, I. D. E. A., Murrell, J. C. & Chen, Y. Trimethylamine and trimethylamine N-oxide are supplementary energy sources for a marine heterotrophic bacterium: implications for marine carbon and nitrogen cycling. *ISME J* **9**, 760-769, doi:10.1038/ismej.2014.149 (2015).

58 Dyhrman, S. T. *et al.* Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. *Nature* **439**, 68-71, doi:10.1038/nature04203 (2006).

59 Karl, D. M. *et al.* Aerobic production of methane in the sea. *Nature Geosci* **1**, 473-478, doi:10.1038/ngeo234 (2008).

60 Lidbury, I. D. E. A. *et al.* Identification of extracellular glycerophosphodiesterases in *Pseudomonas* and their role in soil organic phosphorus remineralisation. *Sci Rep* **7**, 2179, doi:10.1038/s41598-017-02327-6 (2017).

61 Lidbury, I., Kimberley, G., Scanlan, D. J., Murrell, J. C. & Chen, Y. Comparative genomics and mutagenesis analyses of choline metabolism in the marine *Roseobacter* clade. *Environ Microbiol* **17**, 5048-5062 (2015).

62 Chen, Y. Comparative genomics of methylated amine utilization by marine *Roseobacter* clade bacteria and development of functional gene markers (*tmm*, *gmaS*). *Environ Microbiol* **14**, 2308-2322, doi:10.1111/j.1462-2920.2012.02765.x (2012).

63 Dennis, J. J. & Zylstra, G. J. Plasposons: modular self-cloning minitransposon derivatives for rapid genetic analysis of gram-negative bacterial genomes. *Appl Environ Microbiol* **64**, 2710-2715 (1998).

64 Schäfer, A. *et al.* Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome

of *Corynebacterium glutamicum*. *Gene* **145**, 69-73, doi:10.1016/0378-1119(94)90324-7 (1994).

65 Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* **11**, 2301-2319, doi:10.1038/nprot.2016.136 (2016).

66 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **13**, 731-740, doi:10.1038/nmeth.3901 (2016).

67 Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **47**, W636-W641, doi:10.1093/nar/gkz268 (2019).

68 Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* **39**, W29-W37, doi:10.1093/nar/gkr367 (2011).

69 Wackett, L. P., Wanner, B. L., Venditti, C. P. & Walsh, C. T. Involvement of the phosphate regulon and the *psiD* locus in carbon-phosphorus lyase activity of *Escherichia coli* K-12. *J Bacteriol* **169**, 1753-1756, doi:10.1128/jb.169.4.1753-1756.1987 (1987).

70 Milanese, A. *et al.* Microbial abundance, activity and population genomic profiling with mOTUs2. *Nature Comms* **10**, 1014, doi:10.1038/s41467-019-08844-4 (2019).

71 Minh, B. Q. *et al.* IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* **37**, 1530-1534, doi:10.1093/molbev/msaa015 (2020).

72 Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**, W256-W259, doi:10.1093/nar/gkz239 (2019).

73 (2020)., R. C. T. *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.*, <https://www.R-project.org/> (2020).

74 (2020)., R. T. *RStudio: Integrated Development for R.*, <http://www.rstudio.com/> (2020).

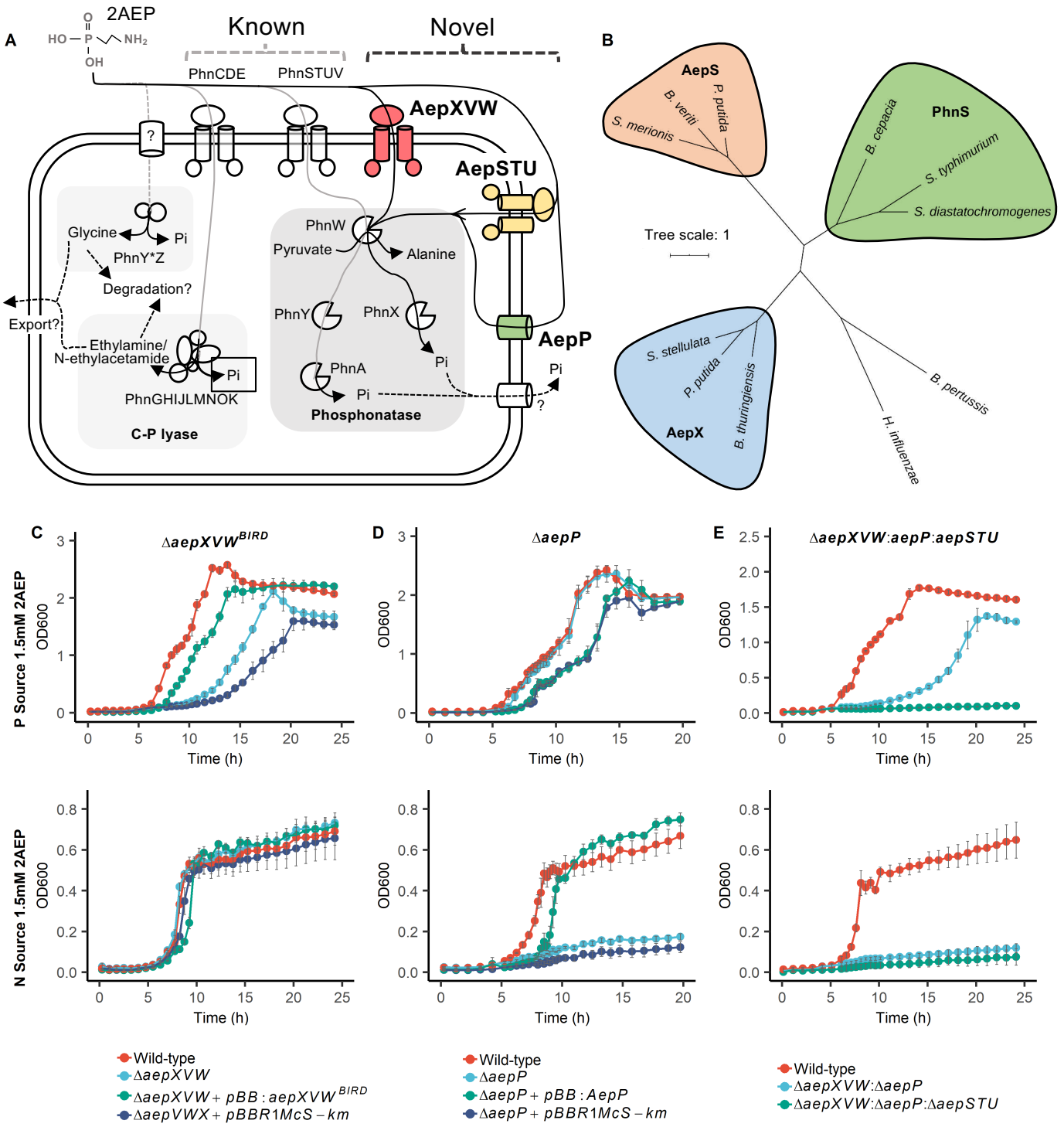


Figure 1. **2AEP transport and catabolism in *P. putida* BIRD-1. (A)** Schematic representation of novel (highlighted in bold and coloured) routes for 2AEP transport, together with existing characterised and putative 2AEP transport routes. Each catabolic system for degradation is highlighted and includes i) the phosphonatase system comprising a 2AEP-pyruvate transaminase (PhnW) and either a phosphonoacetaldehyde hydrolase (PhnX)[74,75] or a NAD^+-dependent phosphonoacetaldehyde dehydrogenase (PhnY) and a phosphonoacetate hydrolase (PhnA)[43], and ii) the PhnY*Z system comprising phosphohydrolase (PhnZ)[21,76] and a 2-oxoglutarate dioxygenase (PhnY*)[77]. In addition, the promiscuous multi subunit enzyme C-P lyase (PhnGHIJKLMN) can also act on 2AEP[78], as well as akylphosphonates[14,15,23]. Pathways found in BIRD-1 are represented by black lines, pathways absent from BIRD-1 are shaded grey. Characterised pathways are shown with solid lines, uncharacterised pathways are shown with dashed lines. Transporters found in BIRD-1 are red if Pi-sensitive, green if Pi-insensitive, and yellow if constitutive. Unknown mechanisms are denoted by a '?'. **(B)** Phylogenetic tree of AepX, PhnS, and AepS, using the characterised Fe^{3+} substrate binding protein FBPa from *Haemophilus influenzae* and *Bordetella pertussis* as an outgroup. *P. putida* = *Pseudomonas putida* BIRD-1, *S. stellulata* = *Stappia stellulata*, *B. cepacia* = *Burkholderia cepacia*, *B. vireti* = *Bacillus veriti*, *S. merionis* = *Streptococcus merionis*, *S. typhimurium* = *Salmonella typhimurium*, *S. diastatochromogenes* = *Streptomyces diastatochromogenes*, *B. thuringiensis* = *Bacillus thuringiensis*. **(C)** Growth (n=4) of *P. putida* BIRD-1 wild type, Δ*aepXVW*, and the complemented mutant. **(D)**, Δ*aepP* and the complemented mutant, and **(E),** the 2AEP null mutant, Δ*aepXVW::aepP::aepSTU.* All strains used 2AEP as a sole P (top panel) or N (bottom panel) source. Error bars denote standard deviation of the mean.

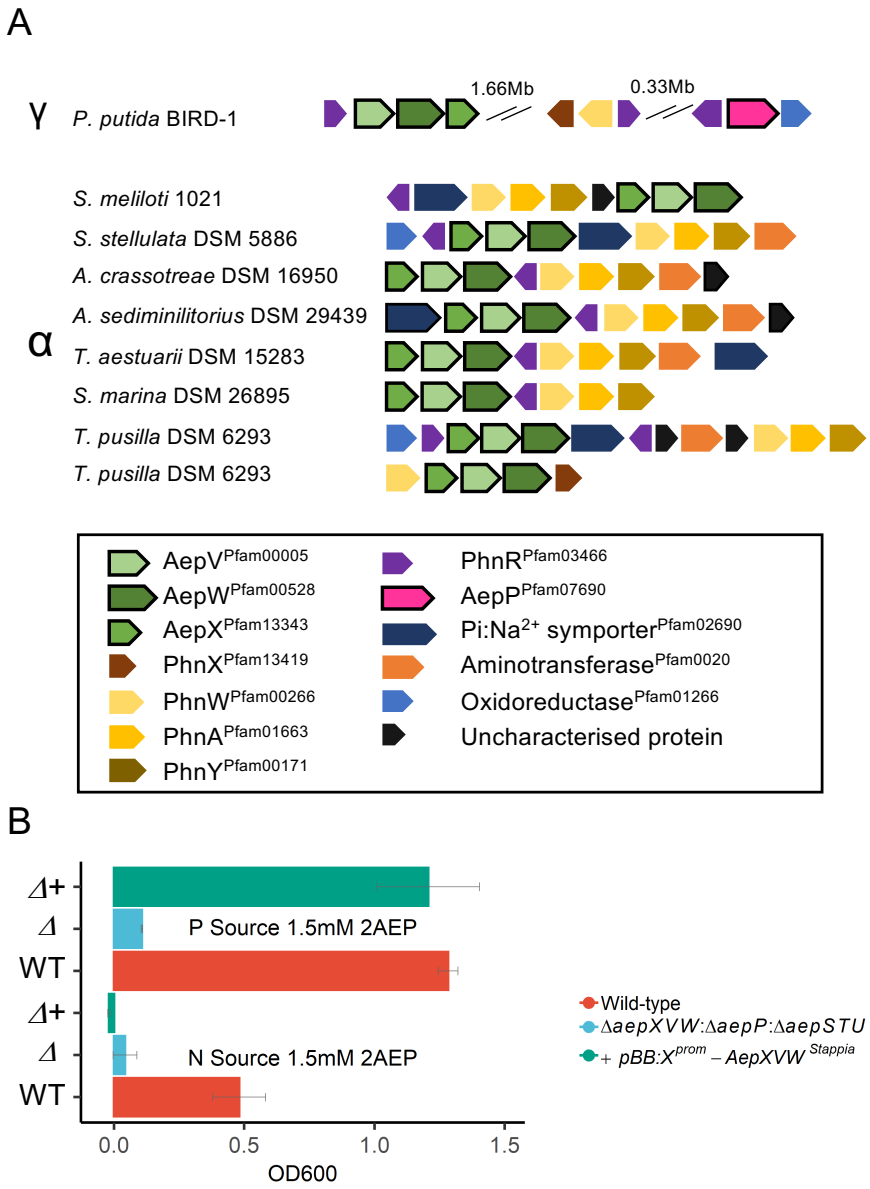


Figure 2. Distribution and functional characterisation of AepXVW in marine bacteria. (A) Genetic neighbourhoods of *aepXVW* within marine *Alpha-* and terrestrial *Gamma-proteobacteria*. Strains shown are *Pseudomonas putida* BIRD-1, *Sinorhizobium meliloti* 1021, *Stappia stellulata* DSM 5886, *Aliiroseovarius crassostreae* DSM 16950, *Aliiroseovarius sediminilitoris* DSM 29439, *Thalassobius aestuarii* DSM 15283, and *Shimia marina* DSM 26895. ORFs separated on the genome are indicated by breaks with the corresponding gap given in megabases (Mb). **(B)** Growth of the *P. putida* BIRD-1 triple 2AEP transporter mutant (Δ*aepXVW:aepP:aepSTU:gm*) complemented with *aepXVW*^{Stappia} concatenated with the promoter region from *aepXVW*^{BIRD} on 2AEP as either a sole N (60 h) or P (48 h) source. Data represents the mean of triplicates cultures. Error bars denote standard deviation.

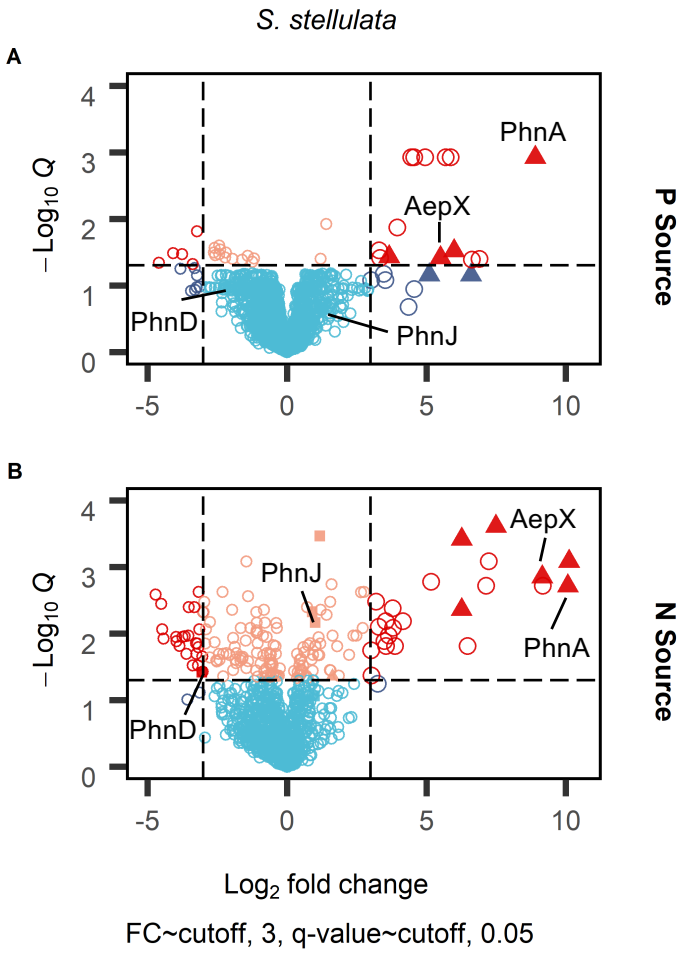
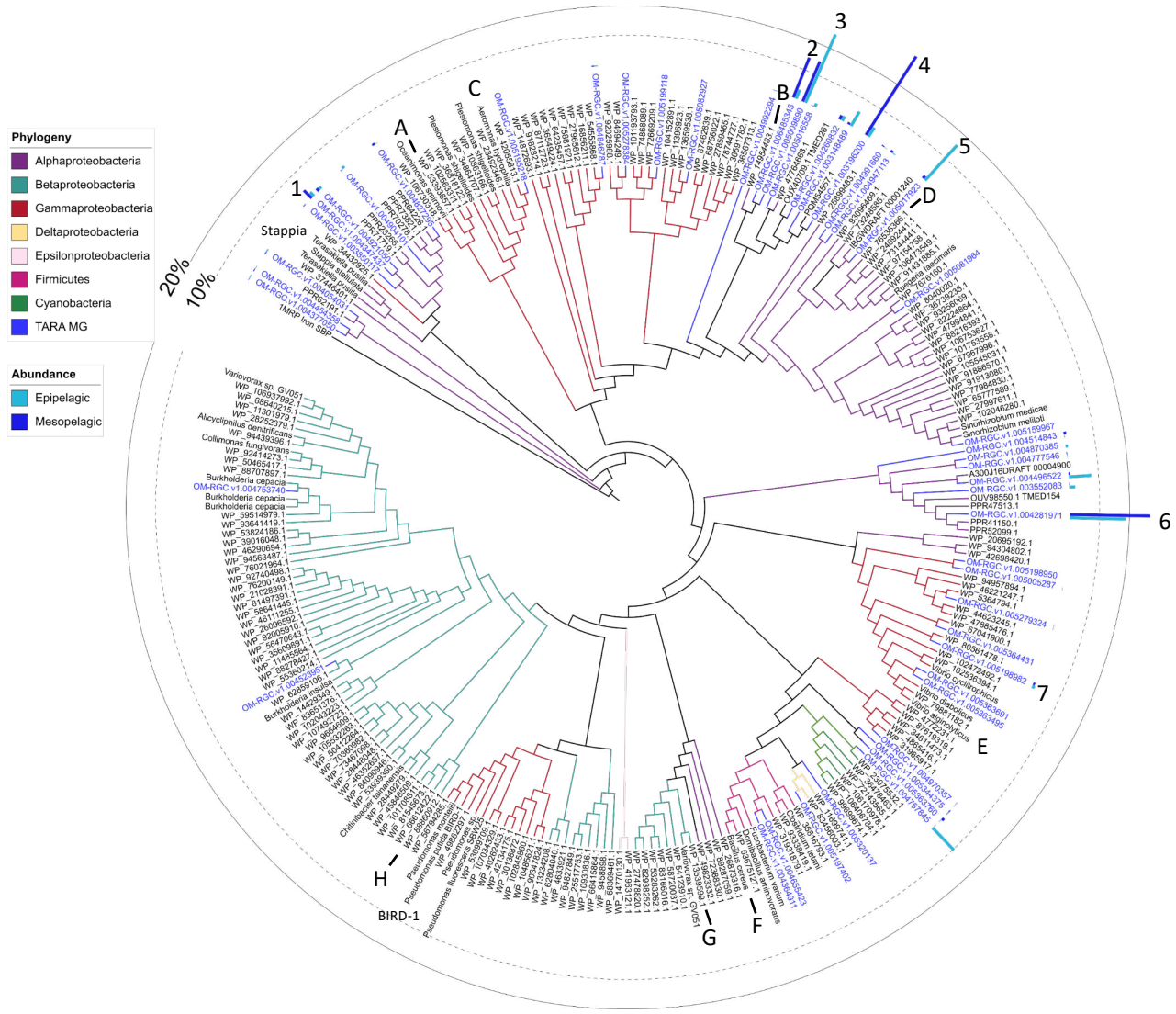


Figure 3. Proteomic analysis of 2AEP-grown *S. stellulata* cells. Whole-cell protein expression profiles (n=3) for *S. stellulata* grown using either Pi or 2AEP as sole P source **(A)** or NH_4 or 2AEP as the sole N source **(B)**. Fold change represents the difference in Log2 LFQ values between each treatment and the statistical value on the Y axis is generated from Q values (FDR corrected P values). Members of the *aepXVW-phnWAY* operon are shown as triangles, members of the CP lyase operon are shown as squares, all other proteins are shown as open circles. Data plotted represents the mean of triplicate cultures.



A) *Vibrio nereis* DSM 19584 — PhnGHIJCLMNP

B) *Alphaproteobacterium* HIMB59

C) *Aeromonas hydrophila* NF1

D) *Roseovarius nanhaiticus* DSM 29590

E) *Vibrio alginolyticus* ATCC 33787

F) *Bacillus cereus* ATCC 4342

G) *Variovorax paradoxus* EPS

H) *Laribacter hongkongensis* HLHK9 — PhnMLKJIH PhnGFN

1) OM-RGC.v1.004347437

2) OM-RGC.v1.006485345

3) OM.RGC.v1. 005009890

4) OM-RGC.v1.003196200

5) OM.RGC.v1. 005017923

6) OM-RGC.v1. 004281971

7) OM.RGC.v1. 005198982

ORFs

- AepV^{Pfam00005}
- AepW^{Pfam00528}
- AepX^{Pfam13343}
- PhnX^{Pfam13419}
- PhnW^{Pfam00266}
- PhnA^{Pfam01663}
- PhnY^{Pfam00171}
- PhnR^{Pfam03466}
- AepP^{Pfam07690}
- AepY^{Pfam02690}
- AepZ^{Pfam00202}
- Oxidoreductase^{Pfam01266}
- Uncharacterised protein
- PHO4 PiT^{Pfam01384}
- Aldehyde dehydrogenase^{Pfam00171}
- Alcohol dehydrogenase^{Pfam00465}
- Amidohydrolase^{Pfam07969}

Figure 4. Phylogenetic and genomic analyses of aepX in marine and terrestrial bacteria. Genetic neighbourhoods for selected aepX homologs are presented adjacent to trees. Numbers indicate environmental OTUs and letters indicate isolates or MAGs/SAGs. Tree topology and branch lengths were calculated by maximum likelihood using the LG+F+I+G4 model of evolution for amino acid sequences based on 744 sites in IQ-TREE software 70. A consensus tree was generated using 1000 bootstraps. Branches representing isolates or MAGs/SAGs are colour coded based on their phylogenetic affiliation (see legends). Branches and identifiers for representative environmental OTU sequences (clustered at 0.8) retrieved from the TARA Oceans database are highlighted blue. The outer ring denotes the relative abundance of environmental AepX OTUs using the same colour scheme; 10% (dashed line) and 20% (filled line) thresholds are shown for scale. S. stellulata DSM 5886 and P. putida BIRD-1 aepX are labelled.

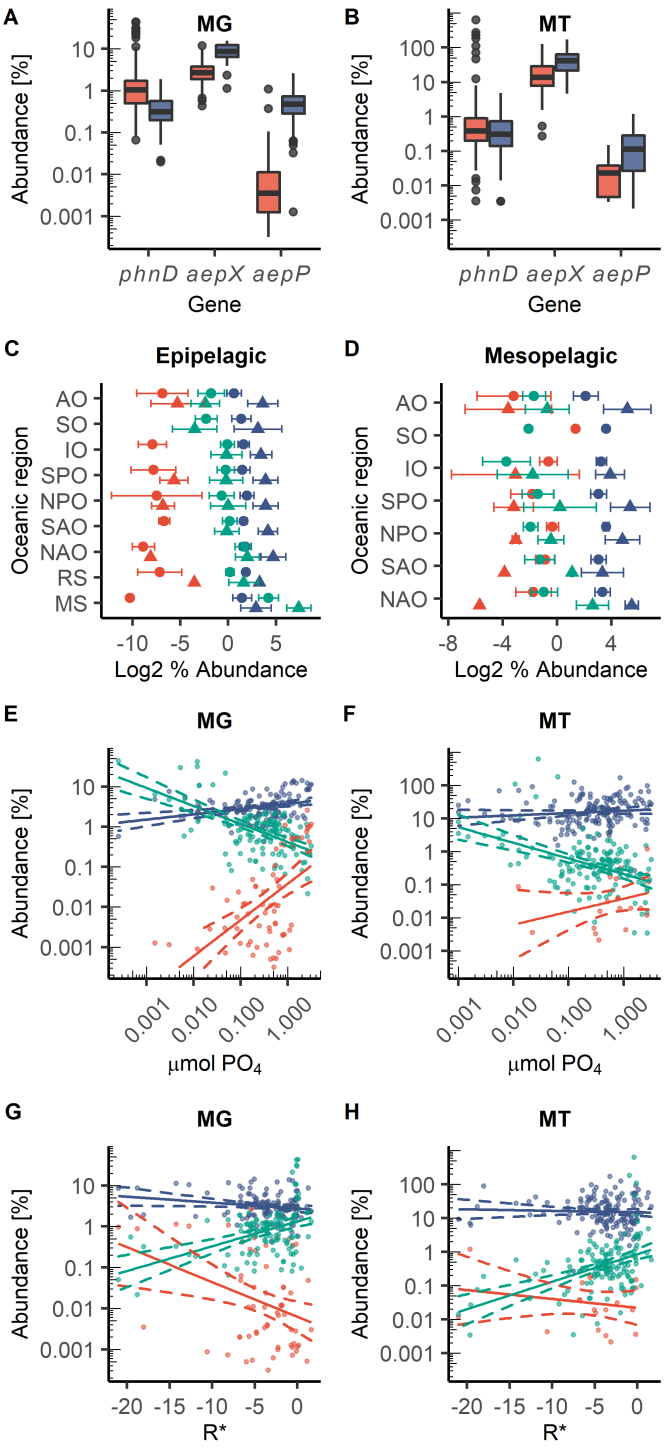


Figure 5. Distribution and transcriptional regulation of phosphonate transporter genes in the global ocean. Abundance of *phnD*, *aepX*, *aepP* in marine epipelagic (red) and mesopelagic (blue) waters, split by metagenome (MG) (**A**), and metatranscriptome (MT) (**B**). Abundance (Log2 % abundance related to single copy core genes) of *phnD*, *aepX*, *aepP* in MG (circles) and MT (triangles) in epipelagic (**C**) and mesopelagic (**D**) waters, split by oceanic region. *aepP* (red), *phnD* (green), *aepX* (blue). AO = Arctic Ocean, SO = Southern Ocean, IO = Indian Ocean, SPO = South Pacific Ocean, NPO = North Pacific Ocean, SAO = South Atlantic Ocean, NAO = North Atlantic Ocean, RS = Red Sea, MS = Mediterranean Sea. The relationship between the standing stock Pi concentration and transporter abundance in the MG (**E**), (*aepX* $R^2 = 0.098^{***}$, *phnD* $R^2 = 0.340^{***}$, *aepP* $R^2 = 0.291^{***}$) and MT (**F**), (*aepX* $R^2 = 0.007^{ns}$, *phnD* $R^2 = 0.203^{***}$, *aepP* $R^2 = 0.058^{ns}$). *aepP* (red), *phnD* (green), *aepX* (blue), ns = not significant, *** = $p < 0.001$. The relationship between R*, a measure of N vs P limitation defined as the sum of standing stock nitrate plus nitrite concentration minus 16x standing stock Pi concentration, and transporter abundance in the MG (**G**) (*aepX* $R^2 = 0.029^{*}$, *phnD* $R^2 = 0.168^{***}$, *aepP* $R^2 = 0.108^{***}$) and MT (**H**) (*aepX* $R^2 = -0.005^{ns}$, *phnD* $R^2 = 0.197^{***}$, *aepP* $R^2 = -0.014^{ns}$). *aepP* (red), *phnD* (green), *aepX* (blue), ns = not significant, * = $p < 0.05$, *** = $p < 0.001$.

Figures

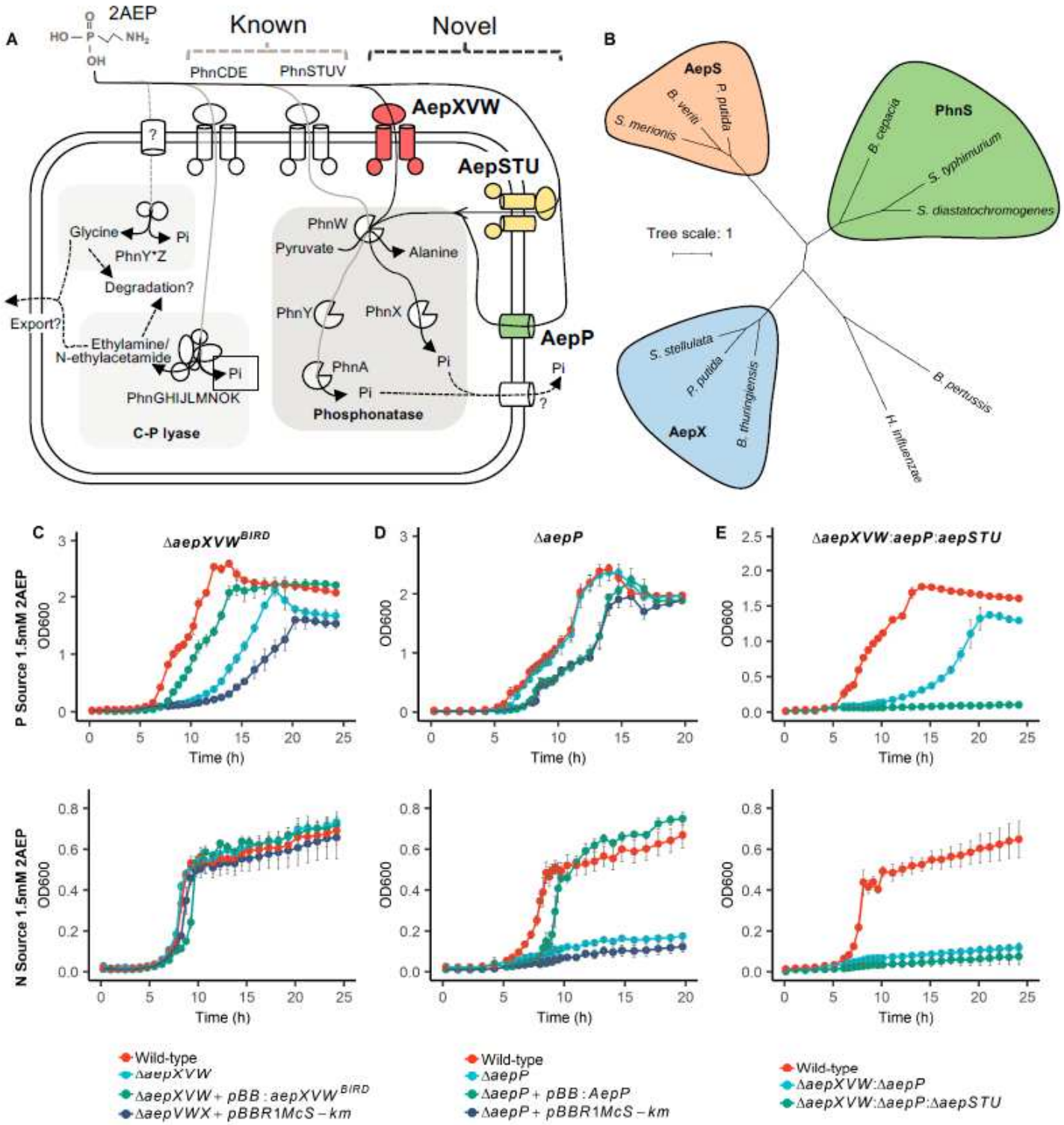


Figure 1

2AEP transport and catabolism in P. putida BIRD-1. (A) Schematic representation of novel (highlighted in bold and coloured) routes for 2AEP transport, together with existing characterised and putative 2AEP transport routes. Each catabolic system for degradation is highlighted and includes i) the phosphonatase

system comprising a 2AEP-pyruvate transaminase (PhnW) and either a phosphonoacetaldehyde hydrolase (PhnX)74,75 or a NAD+-dependent phosphonoacetaldehyde dehydrogenase (PhnY) and a phosphonoacetate hydrolase (PhnA)43, and ii) the PhnY*Z system comprising phosphohydrolase (PhnZ)21,76 and a 2-oxoglutarate dioxygenase (PhnY*)77. In addition, the promiscuous multi subunit enzyme C-P lyase (PhnGHIJKLMN) can also act on 2AEP78, as well as akylphosphonates14,15,23. Pathways found in BIRD-1 are represented by black lines, pathways absent from BIRD-1 are shaded grey. Characterised pathways are shown with solid lines, uncharacterised pathways are shown with dashed lines. Transporters found in BIRD-1 are red if Pisensitive, green if Pi-insensitive, and yellow if constitutive. Unknown mechanisms are denoted by a '?'. (B) Phylogenetic tree of AepX, PhnS, and AepS, using the characterised Fe3+ substrate binding protein FBPa from Haemophilus influenzae and Bordetella pertussis as an outgroup. P. putida = Pseudomonas putida BIRD-1, S. stellulata = Stappia stellulata, B. cepacia = Burkholderia cepacia, B. vireti = Bacillus veriti, S. merionis = Streptococcus merionis, S. typhimurium = Salmonella typhimurium, S. diastatochromogenes = Streptomyces diastatochromogenes, B. thuringiensis = Bacillus thuringiensis. (C) Growth (n=4) of P. putida BIRD-1 wild type, ΔaepXVW, and the complemented mutant. (D), ΔaepP and the complemented mutant, and (E), the 2AEP null mutant, ΔaepXVW::aepP::aepSTU. All strains used 2AEP as a sole P (top panel) or N (bottom panel) source. Error bars denote standard deviation of the mean.

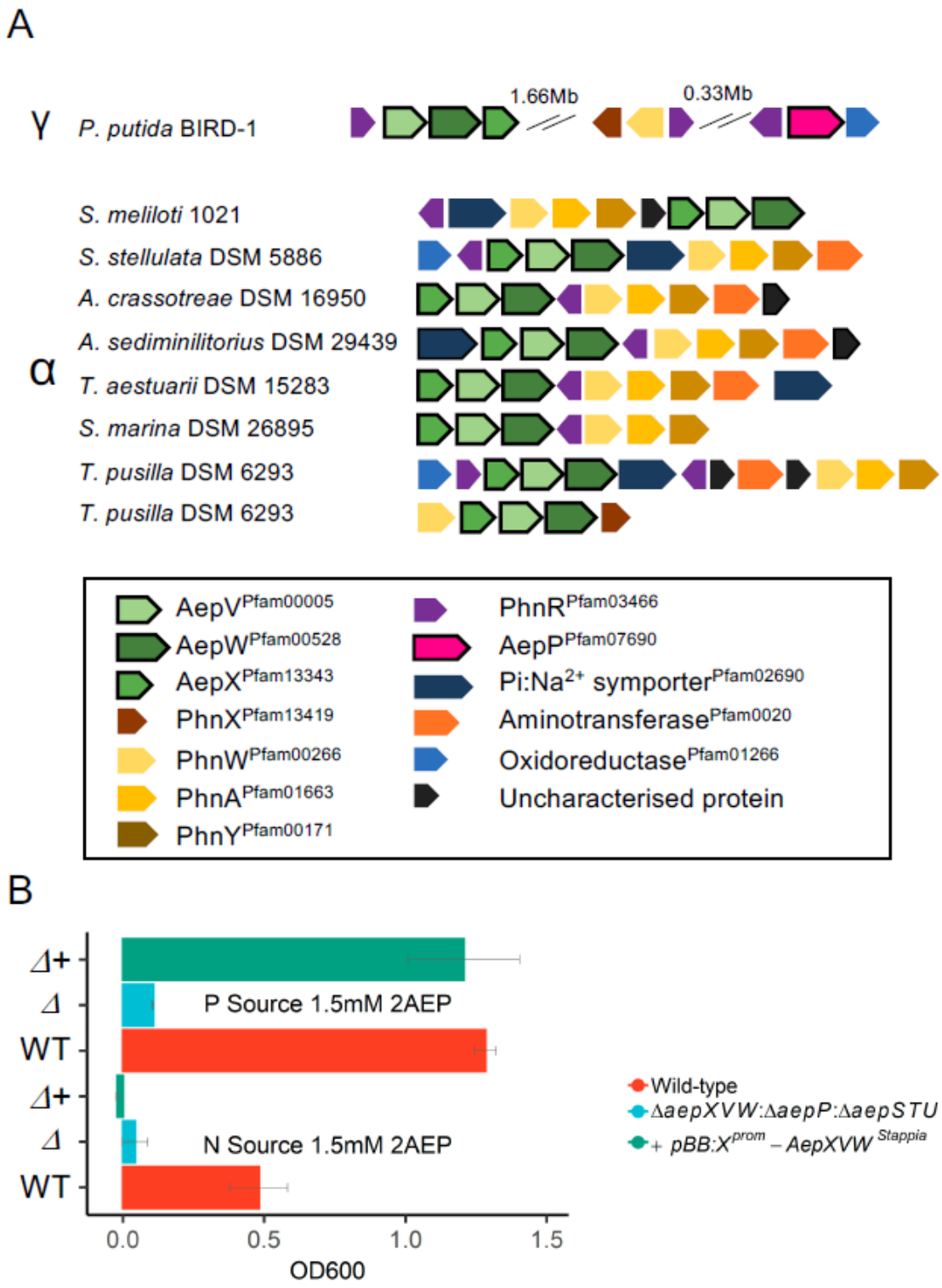


Figure 2

Distribution and functional characterisation of AepXVW in marine bacteria. (A) Genetic neighbourhoods of aepXVW within marine Alpha- and terrestrial Gamma-proteobacteria. Strains shown are Pseudomonas putida BIRD-1, Sinorhizobium meliloti 1021, Stappia stellulata DSM 5886, Aliiroseovarius crassostreae DSM 16950, Aliiroseovarius sediminilitoris DSM 29439, Thalassobius aestuarii DSM 15283, and Shimia marina DSM 26895. ORFs separated on the genome are indicated by breaks with the corresponding gap

given in megabases (Mb). (B) Growth of the P. putida BIRD-1 triple 2AEP transporter mutant (ΔaepXVW:aepP:aepSTU:gm) complemented with aepXVWStappia concatenated with the promoter region from aepXVWBIRD on 2AEP as either a sole N (60 h) or P (48 h) source. Data represents the mean of triplicates cultures. Error bars denote standard deviation.

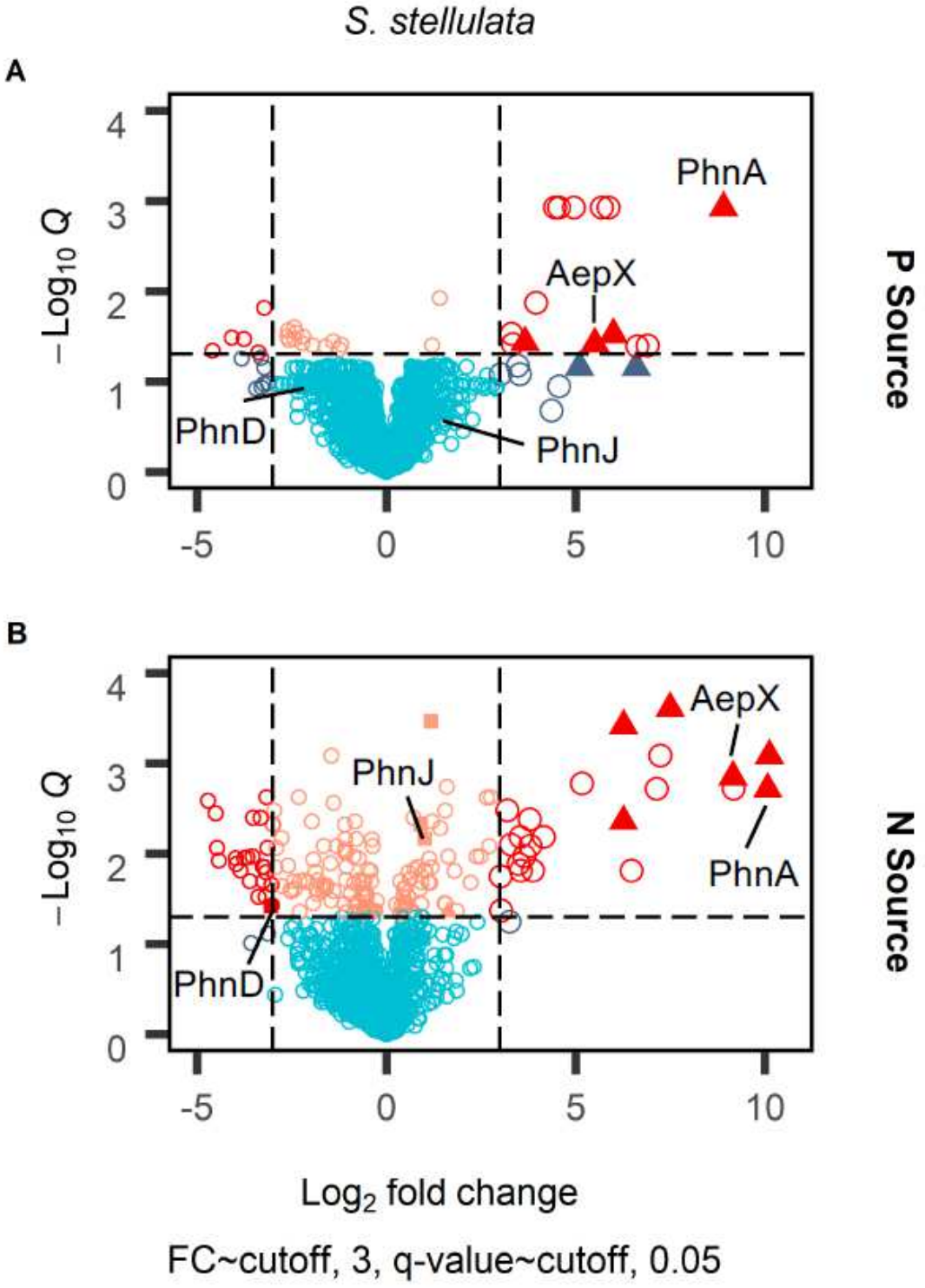


Figure 3

Proteomic analysis of 2AEP-grown S. stellulata cells. Whole-cell protein expression profiles (n=3) for S. stellulata grown using either Pi or 2AEP as sole P source (A) or NH4 or 2AEP as the sole N source (B). Fold change represents the difference in Log2 LFQ values between each treatment and the statistical value on the Y axis is generated from Q values (FDR corrected P values). Members of the aepXVW-phnWAY operon are shown as triangles, members of the CP lyase operon are shown as squares, all other proteins are shown as open circles. Data plotted represents the mean of triplicate cultures.

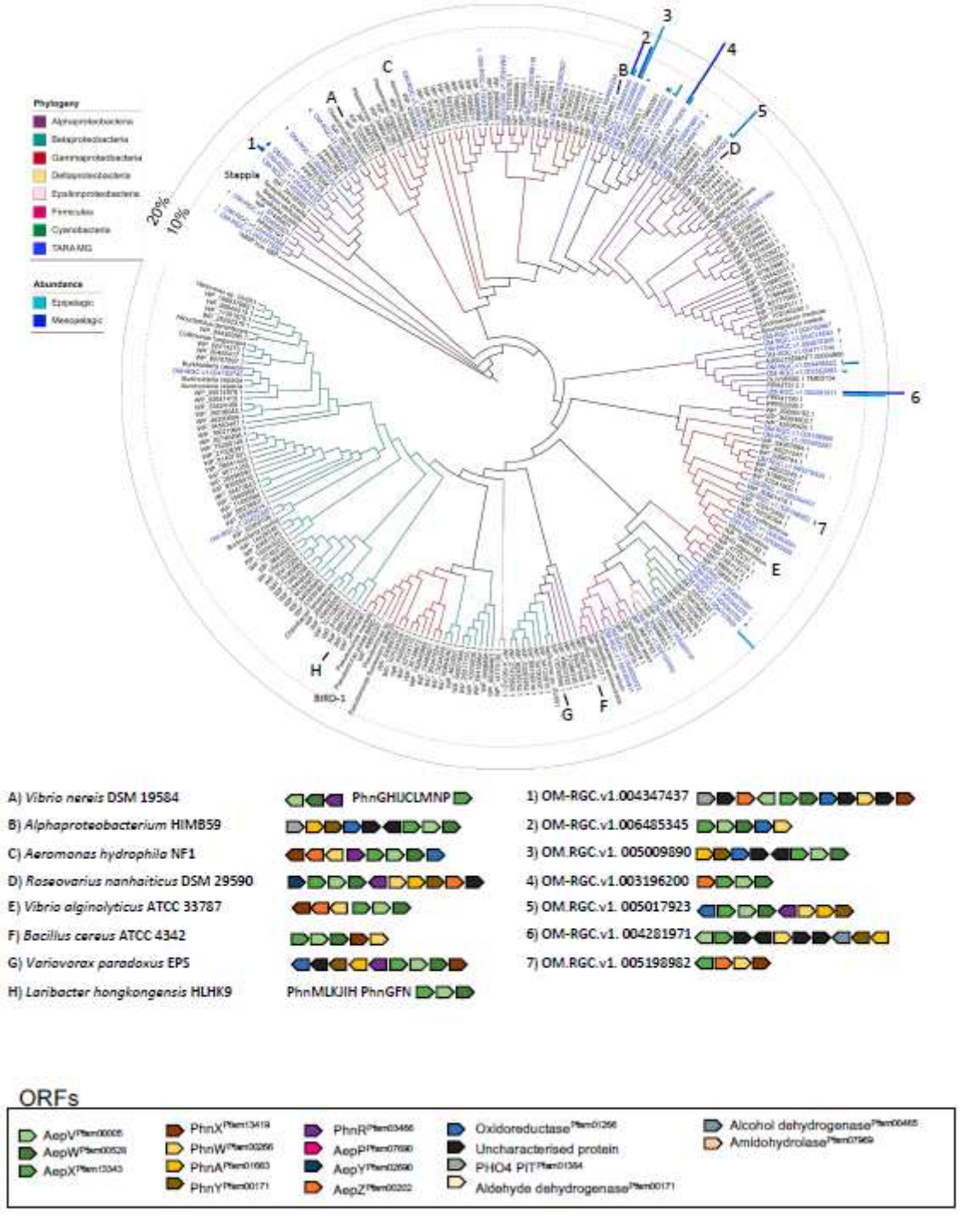


Figure 4

Phylogenetic and genomic analyses of aepX in marine and terrestrial bacteria. Genetic neighbourhoods for selected aepX homologs are presented adjacent to trees. Numbers indicate environmental OTUs and letters indicate isolates or MAGs/SAGs. Tree topology and branch lengths were calculated by maximum likelihood using the LG+F+I+G4 model of evolution for amino acid sequences based on 744 sites in lQ-TREE software 70. A consensus tree was generated using 1000 bootstraps. Branches representing isolates or MAGs/SAGs are colour coded based on their phylogenetic affiliation (see legends). Branches and identifiers for representative environmental OTU sequences (clustered at 0.8) retrieved from the TARA Oceans database are highlighted blue. The outer ring denotes the relative abundance of environmental AepX OTUs using the same colour scheme; 10% (dashed line) and 20% (filled line) thresholds are shown for scale. S. stellulata DSM 5886 and P. putida BIRD-1 aepX are labelled.

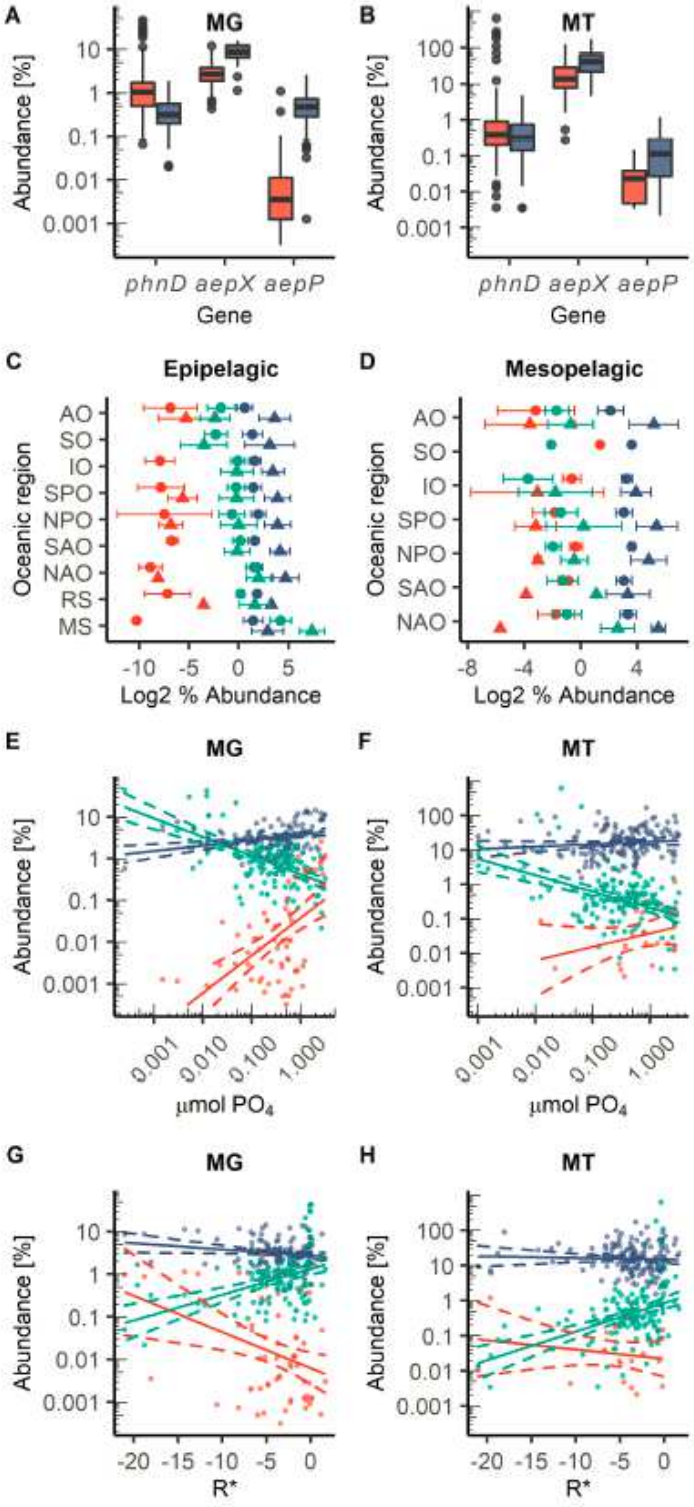


Figure 5

Distribution and transcriptional regulation of phosphonate transporter genes in the global ocean. Abundance of phnD, aepX, aepP in marine epipelagic (red) and mesopelagic (blue) waters, split by metagenome (MG) (A), and metatranscriptome (MT) (B). Abundance (Log2 % abundance related to single copy core genes) of phnD, aepX, aepP in MG (circles) and MT (triangles) in epipelagic (C) and mesopelagic (D) waters, split by oceanic region. aepP (red), phnD (green), aepX (blue). AO = Arctic Ocean,

SO = Southern Ocean, IO = Indian Ocean, SPO = South Pacific Ocean, NPO = North Pacific Ocean, SAO = South Atlantic Ocean, NAO = North Atlantic Ocean, RS = Red Sea, MS = Mediterranean Sea. The relationship between the standing stock Pi concentration and transporter abundance in the MG (E), (aepX R2 = 0.098***, phnD R2 = 0.340***, aepP R2 = 0.291***) and MT (F), (aepX R2 = 0.007ns, phnD R2 = 0.203***, aepP R2 = 0.058ns). aepP (red), phnD (green), aepX (blue), ns = not significant, *** = $p < 0.001$. The relationship between R*, a measure of N vs P limitation defined as the sum of standing stock nitrate plus nitrite concentration minus 16x standing stock Pi concentration, and transporter abundance in the MG (G) (aepX R2 = 0.029*, phnD R2 = 0.168***, aepP R2 = 0.108***) and MT (H) (aepX R2 = -0.005ns, phnD R2 = 0.197***, aepP R2 = -0.014ns). aepP (red), phnD (green), aepX (blue), ns = not significant, * = $p < 0.05$, *** = $p < 0.001$.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- MurphyPhosphonatesSuppFigsNov20Final.pptx
- SuppFigure4.docx
- MurphyPhosphonatesSuppTablesNov20Final.docx
- SupplementaryproteomicstableS3.xlsx