

# *Transporter characterisation reveals aminoethylphosphonate mineralisation as a key step in the marine phosphorus redox cycle*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Murphy, A. R. J., Scanlan, D. J. ORCID: <https://orcid.org/0000-0003-3093-4245>, Chen, Y. ORCID: <https://orcid.org/0000-0002-0367-4276>, Adams, N. B. P. ORCID: <https://orcid.org/0000-0003-3080-3448>, Cadman, W. A., Bottrill, A. ORCID: <https://orcid.org/0000-0002-5182-3643>, Bending, G., Hammond, J. P. ORCID: <https://orcid.org/0000-0002-6241-3551>, Hitchcock, A., Wellington, E. M. H. ORCID: <https://orcid.org/0000-0002-4329-0699> and Lidbury, I. D. E. A. ORCID: <https://orcid.org/0000-0001-7190-315X> (2021)  
Transporter characterisation reveals aminoethylphosphonate mineralisation as a key step in the marine phosphorus redox cycle. *Nature Communications*, 12 (1). 4554. ISSN 2041-1723 doi: 10.1038/s41467-021-24646-z Available at <https://centaur.reading.ac.uk/99594/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1038/s41467-021-24646-z>

Publisher: Nature Publishing Group

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

## **CentAUR**

Central Archive at the University of Reading








Reading's research outputs online

## ARTICLE


<https://doi.org/10.1038/s41467-021-24646-z>

OPEN

# Transporter characterisation reveals aminoethylphosphonate mineralisation as a key step in the marine phosphorus redox cycle

Andrew R. J. Murphy<sup>1</sup>, David J. Scanlan <sup>1</sup>, Yin Chen <sup>1</sup>, Nathan B. P. Adams <sup>2,3</sup>, William A. Cadman<sup>2</sup>, Andrew Bottrill <sup>1</sup>, Gary Bending<sup>1</sup>, John P. Hammond <sup>4</sup>, Andrew Hitchcock<sup>2</sup>, Elizabeth M. H. Wellington <sup>1</sup> & Ian D. E. A. Lidbury <sup>5</sup>✉

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. Here, we identify three bacterial 2-aminoethylphosphonate (2AEP) transporters, named AepXVW, AepP and AepSTU, whose synthesis 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 phosphonate binding protein, PhnD, AepX has high affinity and high specificity for 2AEP (*Stappia stellulata* AepX  $K_d$   $23 \pm 4$  nM; methylphosphonate  $K_d$   $3.4 \pm 0.3$  mM). In the global ocean, *aepX* is heavily transcribed ( $\sim 100$ -fold  $> phnD$ ) independently of phosphate and nitrogen concentrations. Collectively, our data identifies a mechanism responsible for a major oxidation process in the marine phosphorus redox cycle and suggests 2AEP may be an important source of regenerated phosphate and ammonium, which are required for oceanic primary production.

<sup>1</sup>School of Life Sciences, University of Warwick, Gibbet Hill Road, Coventry, UK. <sup>2</sup>Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, UK. <sup>3</sup>Nanotemper Technologies GmbH, Flößergasse 4, Munich, Germany. <sup>4</sup>School of Agriculture, Policy, and Development, University of Reading, Earley Gate, Whiteknights, Reading, UK. <sup>5</sup>Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK. ✉email: [i.lidbury@sheffield.ac.uk](mailto:i.lidbury@sheffield.ac.uk)

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<sup>1</sup>. Phosphonates are synthesised as both primary and secondary metabolites in various bacterial, archaeal and eukaryotic organisms<sup>1–7</sup> where they are incorporated into lipids (phosphonolipids) and glycans (phosphonoglycans)<sup>4,8</sup>. A significant proportion can also be released from the cell to facilitate favourable biotic interactions<sup>9</sup>. Thus, they are ubiquitous in terrestrial and aquatic ecosystems<sup>10–14</sup>. Phosphonates also represent a major fraction of the marine organic P pool<sup>11,15–17</sup> and recent studies have now identified several cosmopolitan marine microorganisms capable of synthesising significant quantities of these compounds<sup>6,7,9,18,19</sup>. Collectively, this synthesis drives a vast global oceanic P redox cycle with reduced P input in the surface ocean estimated to be an order of magnitude greater than (non-anthropogenic) riverine P input<sup>9</sup>. Whilst much attention has focused on the degradation of alkylphosphonates, such as (hydroxy-)methylphosphonate and 2-hydroxyethylphosphonate, as a source of P<sup>16,17,20,21</sup>, the potential for aminophosphonates such as 2-aminoethylphosphonate (2AEP) to serve as sources of C and/or nitrogen (N) in the presence of inorganic phosphate (Pi), i.e. in a Pi-insensitive manner, has been neglected. However, emerging evidence suggests that Pi-insensitive 2AEP catabolism occurs in nature<sup>22,23</sup>. Notably, the fact that 2AEP is either absent from, or a minor component of, otherwise phosphonate rich high molecular weight dissolved organic matter (HMW DOM)<sup>16,17</sup>, despite its supposed ubiquitous production<sup>9,19,24,25</sup>, suggests preferential catabolism of this molecule in comparison to alkylphosphonates.

Unlike the majority of C–O–P monoester bonds, the C–P bond requires specific enzymes to break it, such as the C–P lyase<sup>26,27</sup>. Several 2AEP-specific phosphonate degradation systems have been characterised (Fig. 1a), including the 2AEP transaminase (PhnW) – phosphonoacetaldehyde hydrolase (phosphonatase-PhnX) system<sup>28,29</sup> and the PhnW – phosphonoacetaldehyde dehydrogenase (PhnY)-phosphonoacetate hydrolase (PhnA) system<sup>30–32</sup>. The phosphonate dioxygenase (PhnY\*) – phosphohydrolase (PhnZ) system has also been shown to degrade 2AEP<sup>33,34</sup>, though at least some homologs of this system are specific to (hydroxy-)methylphosphonate and cannot degrade 2AEP<sup>35,36</sup>. In addition, a gene encoding a recently characterised (R)-1-hydroxy-2-aminoethylphosphonate ammonia lyase (PbfA) is often found in *phnWX* and *phnWAY* operons<sup>37</sup>, expanding the known repertoire of aminophosphonate degrading capabilities<sup>37</sup>. The C–P lyase, which is a non-specific promiscuous phosphonate degrading enzyme complex, is only induced in response to Pi-starvation, being regulated by the two-component master regulator of the Pi-stress response regulon, PhoBR<sup>24</sup>. In marine surface waters, genes encoding the C–P lyase are enriched in bacterial genomes found in regions typified by low Pi concentrations<sup>20</sup> where they are also heavily expressed<sup>38</sup>. Recent data has shown Pi-insensitive regulation of 2AEP degradation, facilitated by the 2AEP-specific systems, occurs in a few strains of bacteria related to marine *Alphaproteobacteria*<sup>22</sup> and a terrestrial *gammaproteobacterium*<sup>23</sup>. In both cases, a major consequence of Pi-insensitive 2AEP degradation was the remineralisation and release of labile Pi<sup>22</sup>, due to the greater cellular demand for N over P and the 1:1 N:P stoichiometry of 2AEP.

When analytical methods are not sensitive enough to accurately quantify the concentration and turnover of specific environmental metabolites, screening for the expression of their respective uptake systems becomes an important tool in understanding their in situ cycling<sup>38–41</sup>. To date, only two 2-AEP transport systems have been identified. Both are ATP-binding cassette (ABC) transporters that consist of a periplasmic

substrate-binding protein, an ATP-binding domain protein, and a transmembrane permease. The first is located within the C–P lyase operon (*phnCDEFGHIJKLMNOP*)<sup>42,43</sup> with genes encoding the substrate-binding protein, 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*<sup>28,29</sup>. PhnD has a restricted distribution in seawater with an abundance that is highly correlated with regions of Pi-limitation<sup>20</sup>. However, no 2AEP transporter has been identified in the majority of bacteria possessing PhnWX and PhnWAY systems, which is surprising given the fact that 2AEP is a charged molecule and ubiquitous in marine and terrestrial ecosystems<sup>1,20</sup>.

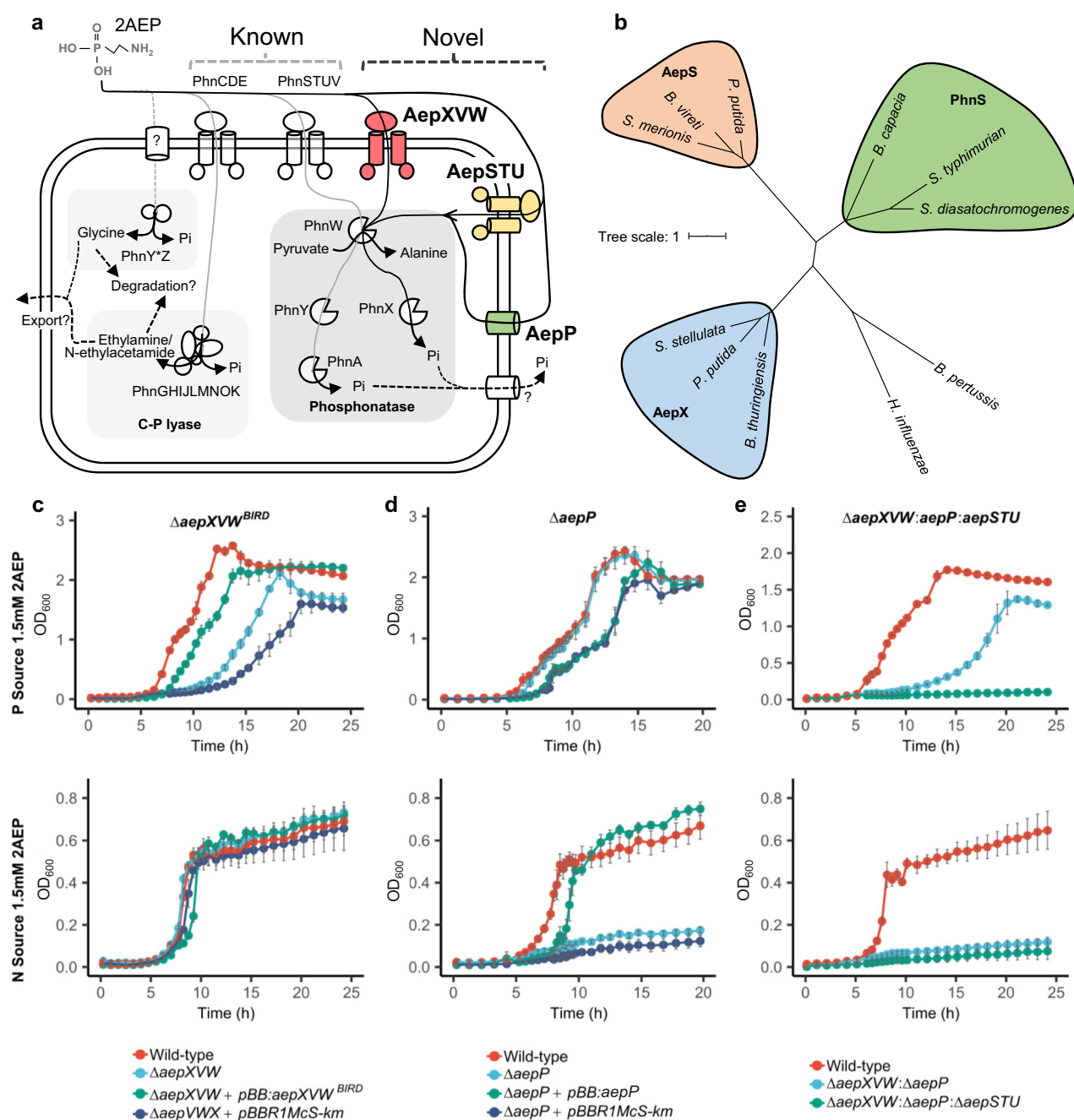
Here, we sought to identify transporters required for 2AEP catabolism in environmental bacteria lacking *PhnCDE* or *PhnSTU*. Through combining laboratory-based molecular and genetic analyses with environmental meta-omics, we identified three transporters that have a role in 2AEP uptake and revealed Pi-insensitive 2AEP catabolism is widespread in the global ocean, likely representing a major step in the marine P redox cycle.

## Results

***Pseudomonas putida* BIRD-1 possesses a 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<sup>44</sup> (Fig. 1a). In *Pseudomonas putida* BIRD-1 (hereafter BIRD-1), a periplasmic substrate-binding protein associated with one of these putative transporters was induced under Pi-deplete growth conditions in a PhoBR-dependent manner<sup>44</sup>. We hereafter refer to this substrate-binding protein (PPUBIRD1\_4925) as 2-aminoethylphosphonate X (AepX). AepX belongs to the same family (pfam13343) as PhnS, iron and sulphate substrate-binding proteins, but is clearly distinct (Coverage = 40%, Identity = 25.09%, 1.1e<sup>−05</sup>) (Fig. 1b).

BIRD-1 was capable of growth on 1.5 mM 2AEP as either a sole P, sole N or sole N and sole P source, the latter two resulting in mineralisation of excess Pi that was subsequently exported from the cell (Fig. 1c, Supplementary Fig. 1). Mutagenesis of *phnWX* confirmed this phosphonatase was essential for 2AEP catabolism under all growth conditions in this bacterium (Supplementary Fig. 2a, b). 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*<sup>BIRD</sup> had no effect on growth as a sole N source (Fig. 1c). However, the mutant ( $\Delta$ *aepXVW*<sup>BIRD</sup>) 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 in this bacterium but is not involved in 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  $\Delta$ *aepXVW*<sup>BIRD</sup> mutant to comparative proteomics, we identified a major facilitator-type transporter, (PPUBIRD1\_3129), hereafter referred to as AepP for 2-aminoethylphosphonate permease, whose expression was significantly increased during



**Fig. 1** 2AEP transport and catabolism in *P. putida* BIRD-1. **a** Schematic representation of candidate (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 phosphonate system comprising a 2AEP-pyruvate transaminase (PhnW) and either a phosphonoacetaldehyde hydrolase (PhnX)<sup>28,29</sup> or a NAD<sup>+</sup>-dependent phosphonoacetaldehyde dehydrogenase (PhnY)<sup>32</sup> and a phosphonoacetate hydrolase (PhnA)<sup>30,31</sup>, and (ii) the PhnY\*Z system comprising phosphohydrolase (PhnZ)<sup>33,34</sup> and a 2-oxoglutarate dioxygenase (PhnY\*)<sup>33</sup>. In addition, the promiscuous multi-subunit enzyme C-P lyase (PhnGHIJLMNOK) can also act on 2AEP<sup>26,27</sup>, as well as alkylphosphonates. 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<sup>3+</sup> 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*, *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,  $\Delta aepXVW$ , and the complemented mutant. **d**, Growth ( $n = 4$ ) of *P. putida* BIRD-1 wild type,  $\Delta aepP$  and the complemented mutant, and **e**, Growth ( $n = 4$ ) of *P. putida* BIRD-1 wild type, the double mutant,  $\Delta aepXVW:\Delta aepP$ , and the 2AEP null transporter mutant,  $\Delta aepXVW:\Delta aepP:\Delta aepSTU$ . All strains used 2AEP as a sole P (top panel) or N (bottom panel) source. Error bars denote standard deviation of the mean.



growth on 2AEP as a sole N source (Pi-insensitive) (Supplementary Fig. 3, Supplementary Table 3a). AepP is related to the glycerol-3-phosphate (G3P): Pi antiporter, GlpT<sup>45,46</sup> (identity = 27.75%,  $9e^{-37}$ ), a member of the organophosphate: Pi antiporter family of major facilitator transporters. AepP shares conserved residues essential for binding Pi and the Pi moiety of G3P<sup>47–50</sup> with GlpT, whereas residues that impact the binding affinity to the glycerol moiety of G3P but not Pi are not conserved<sup>49</sup> (Supplementary Fig. 4). Mutation of *aepP* in either the wild type parental strain ( $\Delta aepP$ , Fig. 1d) or the *aepXVW* mutant ( $\Delta aepXVW \Delta aepP$ , Fig. 1e) led to an inability to grow on 2AEP as sole N source. Subsequent complementation of  $\Delta aepP$  with its native homolog ( $\Delta aepP + pBB:aepP$ ) 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  $\Delta aepXVW \Delta aepP$  double mutant, revealing the presence of a third 2AEP transporter in this bacterium.

To identify the unknown 2AEP transporter, we reanalysed our proteomics data (Supplementary Fig. 3). Another substrate binding protein (PPUBIRD1\_3891) containing the same pfam domain (pfam13343) as AepX (Fig. 1b), hereafter named AepS, was constitutively synthesised 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  $\Delta aepXVW \Delta aepP \Delta aepSTU$  was generated in BIRD-1 (Fig. 1e). This triple knockout mutant was unable to grow on 2AEP as a sole P source (Fig. 1e), suggesting AepSTU is a functional 2AEP transporter. However, generation of a single  $\Delta aepSTU$  knockout mutant did not affect Pi-sensitive growth compared to the wild type (Supplementary Fig. 2c), suggesting AepXVW is the major transporter involved in Pi-sensitive 2AEP uptake and AepSTU only has an auxiliary role in 2AEP uptake. Despite production of AepS during Pi-insensitive growth (Supplementary Fig. 3), this transporter was unable to facilitate growth on 2AEP as a sole N source in the absence of AepP. Together, these data reveal the presence of three differentially regulated 2AEP transporters in BIRD-1. AepXVW is the primary Pi-sensitive 2AEP transporter with AepSTU having an auxiliary role, whereas AepP is essential for Pi-insensitive growth on 2AEP.

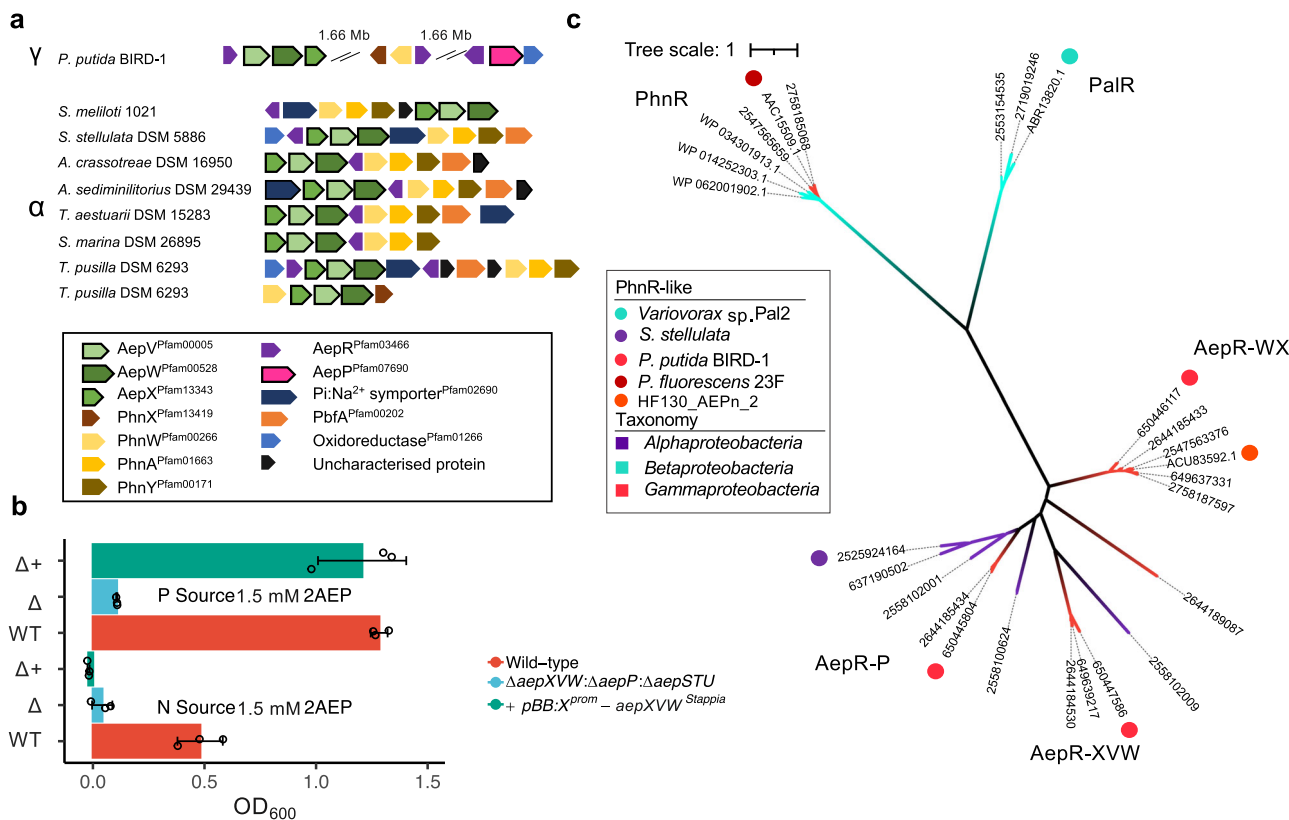
**AepXVW is found in several marine bacteria capable of Pi-insensitive mineralisation and is a functional 2-aminoethylphosphonate transporter.** Using PPUBIRD1\_4925 (AepX) as the query, we scrutinised the genomes of several isolates related to marine *Rhodobacteraceae* (*Stappia* spp., *Terasakiella* spp., *Falsirhodobacter* spp.) capable of Pi-insensitive phosphonate catabolism<sup>22</sup>. ORFs encoding orthologs of AepX were identified in the genomes of *Stappia stellulata* DSM 5886 and *Terasakiella pusilla* DSM 6293, in addition to several other marine Roseobacter strains: *Aliiroseovarius crassostreae* (DSM 16950), *Aliiroseovarius sediminilitoris* (DSM 29439), *Shimia marina* (DSM 26895), and *Thalassobius aestuarii* (DSM 15283) (Fig. 2a). We also found an orthologous ORF in the model rhizosphere alphaproteobacterium *Sinorhizobium meliloti* strain 1021 that is capable of 2AEP catabolism via PhnWAY<sup>31</sup> (Fig. 2a). In all cases, ORFs encoding AepXVW were located adjacent to ORFs encoding 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 (Supplementary Table 1, Supplementary Fig. 5). Indeed, both *Aliiroseovarius* strains lack other characterised 2AEP transport and degradation systems (Supplementary Table 2). As previously reported, Pi was exported from cells and accumulated in the medium during growth on 2AEP as

a sole N source (Supplementary Fig. 6a). To confirm that *S. stellulata* AepXVW can take up 2AEP, we complemented the BIRD-1 null 2AEP transporter mutant ( $\Delta aepXVW:\Delta aepP:\Delta aepSTU$ ) with this transporter fused with the *aepXVW*<sup>BIRD</sup> promoter (Fig. 2b). This duly restored growth of the triple mutant confirming *S. stellulata* AepXVW is also a functional 2AEP transporter.

In many 2AEP gene clusters, we identified LysR-type regulators, which we refer to as AepR. A homolog of AepR has been shown to be essential for complementation of an *Escherichia coli*  $\Delta phnHIJ KLMNOP$  mutant with a *Pseudomonas* PhnWX<sup>25</sup>, implying substrate inducible regulation. Additionally, PhnA activity has been shown to be induced by 2AEP in a marine *Falsirhodobacter* isolate even under nutrient replete conditions<sup>22</sup>, though unfortunately no sequenced *Falsirhodobacter* strain possesses an aminophosphonate operon so it is not clear if AepR is responsible for this regulation. BIRD-1 and other *Pseudomonas*, whose 2AEP operons are often fragmented throughout the genome, possess up to three distinct genes encoding LysR-type regulators (Fig. 2a). In contrast, most *Alphaproteobacteria* only possess a single gene, located upstream of the *aepXVW-phnWAY* operon. These newly identified forms are phylogenetically distinct from either the archetypal PhnR<sup>51</sup> or PalR<sup>52</sup> found in *P. fluorescens* sp. 23F and *Variovorax* sp. PAL2, respectively (Fig. 2c). The three BIRD-1 LysR-like forms were clearly distinct from each other with the AepP-associated form being closely related to the single LysR-type regulator found in *Alphaproteobacteria*.

**AepXVW is highly synthesised 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 synthesised during Pi-insensitive and Pi-sensitive metabolism (Fig. 1). *S. stellulata* lacks AepP but is still capable of Pi-insensitive growth and Pi export when grown on 2AEP as a sole N, or sole N and sole P source (Supplementary Fig. 5a). In addition to *aepXVW* and genes encoding the 2AEP degradation system (*phnWAY*), *S. stellulata* also possesses genes (*phnCDEFGHIJ KLMNO*) encoding the P-regulated C-P lyase operon and we also confirmed this strain grew on several other alkylphosphonates as sole P source (Supplementary Fig. 5b). 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 synthesised during growth on 2AEP as either a sole N or P source, as was PhnWAY and PbfA, the (R)-1-hydroxy-2AEP ammonia lyase, whilst the C-P lyase encoded proteins (including PhnD) were not (Fig. 3). Importantly, whilst we detected several general N stress-response proteins induced under Pi-insensitive growth, we did not identify any other potential 2AEP transporters (Supplementary Table 3b). Therefore, unlike in BIRD-1, AepXVW likely represents the major route for 2AEP uptake in this bacterium. These data are consistent with the hypothesis that *S. stellulata* 2AEP catabolism is regulated by a LysR-type regulator solely through substrate-induction, which will be investigated in a future study.

Finally, we determined the substrate specificity of recombinant *S. stellulata* AepX towards 2AEP and other (alkyl)phosphonates using microscale thermophoresis<sup>53,54</sup>. Unlike the relatively promiscuous phosphonate binding protein PhnD<sup>43</sup>, AepX was highly specific for 2AEP with a  $K_d$  in the nanomolar range (Table 1, Supplementary Fig. 7), consistent with the observation that *aepXVW* is typically co-localised with either *phnWX* or *phnWAY* that encode 2AEP-specific degradation systems (Fig. 2a).



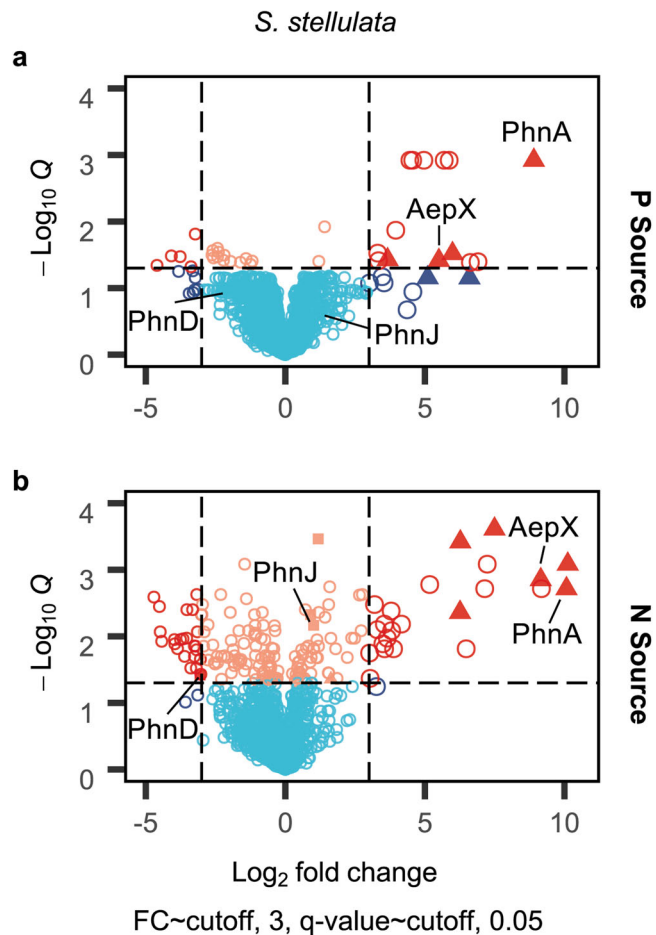
**Fig. 2** Distribution and functional characterisation of *AepXVW* in marine bacteria. **a** Genetic neighbourhoods of *aepXVW* within marine Alpha- and terrestrial Gamma-proteobacteria. Strains shown are *Pseudomonas putida* BIRD-1, *Sinorhizobium meliloti* 1021, *Stappia stellulata* DSM 5886, *Aliiroseovarius 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 ( $\Delta aepXVW:\Delta aepP:\Delta aepSTU$ ) complemented with *aepXVW*<sup>Stappia</sup> concatenated with the promoter region from *aepXVW*<sup>BIRD</sup> on 2AEP as either a sole N (60 h) or P (48 h) source. Data represent the mean of triplicates cultures. Error bars denote standard deviation. **c** Phylogeny of phosphonate-associated LysR-type regulators. Labels denote IMG/JGI gene IDs or Genbank accession numbers. Tree topology and branch lengths were calculated by maximum likelihood using the LG + G4 model of evolution for amino acid sequences based on 744 sites in IQ-TREE software<sup>84</sup>. Tree Scale represents the number of substitutions per site. A consensus tree was generated using 1000 bootstraps.

***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 and metagenome-assembled genomes, which revealed an unexpected diversity for these substrate-binding proteins (Fig. 4). For *AepX*, this included cosmopolitan marine *Alphaproteobacteria* in addition to *Rhodobacteraceae*, marine *Deltaproteobacteria*, and 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*<sup>Stappia</sup> and *AepX*<sup>BIRD</sup> well separated (Fig. 4). Many taxonomically divergent *AepX* encoding ORFs were co-localised with ORFs encoding the various 2AEP degradation systems, including the (R)-1-hydroxy-2-aminoethylphosphonate specific PbfA, the C-P lyase, or putative uncharacterised ORFs encoding potentially novel phosphonate catabolic enzymes, 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) (Supplementary Fig. 6). 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 *PhnWAY* or *PhnWX*, or putative catabolic enzymes (Supplementary Fig. 8). 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<sup>55</sup>, we calculated the abundance of *aepX*, *aepS* and *aepP* and compared this with *phnS* and the gene (*phnD*) encoding the archetypal phosphonate transporter, whose gene abundance in seawater was recently calculated<sup>20</sup>. The distribution of markers (*phnJ*, *phnA*, *phnX*) for the various phosphonate degradation mechanisms were also analysed (Supplementary Fig. 9). We analysed data from both the epipelagic and mesopelagic zones where phosphonate mineralisation is believed to occur<sup>15</sup>. 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, 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). The transcript abundance of *aepX* was 40-fold and 140-fold greater than *phnD* in the epipelagic and mesopelagic, respectively (Fig. 5b). The majority of *aepX* sequences were related to the cosmopolitan *Alphaproteobacteria*, although some deltaproteobacterial



**Fig. 3** Proteomic analysis of 2AEP-grown *S. stellulata* cells. Whole-cell protein profiles ( $n = 3$ ) for *S. stellulata* grown using either Pi or 2AEP as sole P source (**a**) or  $\text{NH}_4$  or 2AEP as the sole N source (**b**). Fold change (FC) represents the difference in  $\text{Log}_2$  Label Free Quantification (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 C-P lyase operon are shown as squares, all other proteins are shown as open circles. Data plotted represents the mean of triplicate cultures. Vertical dashed lines represent an  $\text{Log}_2$  LFQ difference  $> -3$  or  $< 3$ . The horizontal dashed line illustrates a cut off for a significant Q value ( $p < 0.05$ ). Sky blue represents proteins showing no significant difference between treatments. Red indicates proteins significantly changing in abundance  $< -3$ -fold or  $> 3$ -fold, respectively. Peach indicates significant changes less than 3-fold in either direction.

Table 1 Microscale thermophoresis determined dissociation constants ( $K_d$ ) of the <i>S. stellulata</i> AepX for selected phosphonate ligands.		
Ligand	$K_d$ ( $\mu\text{M}$ )	
2-Aminoethylphosphonate	AepX Ss $0.023 \pm 0.004$	PhnD Ec $>0.050$ – $0.1$
Methylphosphonate	$3404.15 \pm 280.99$	$1.3$ – $18.4$
Ethylphosphonate	$145.96 \pm 15.18$	$0.2$ – $1.4$
Aminomethylphosphonate	$4490.91 \pm 808.24$	$16.6$

Results are compared against those obtained for *E. coli* PhnD in previous studies<sup>42,65</sup>.

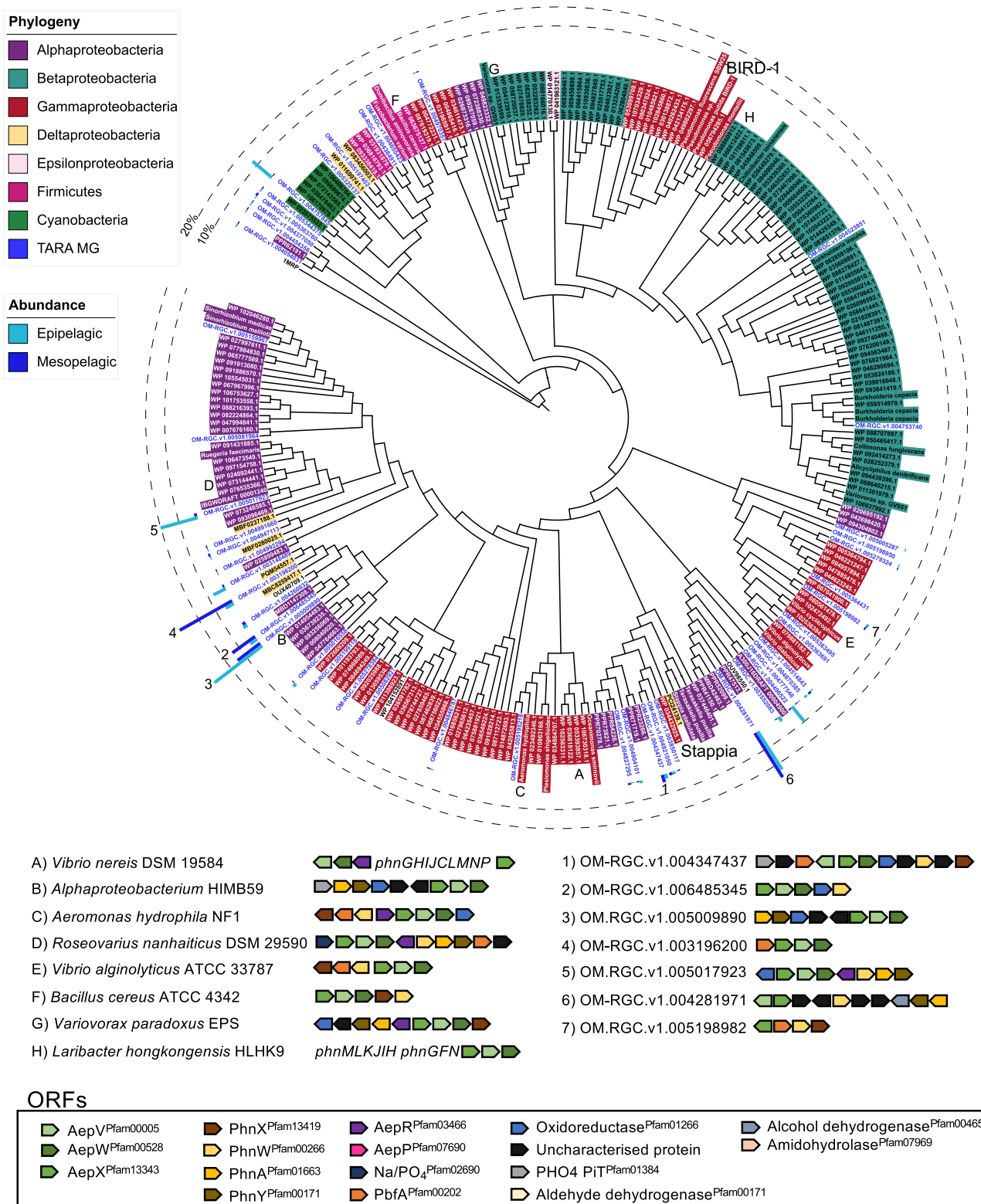
sequences fall within this clade (Fig. 4). We confirmed that these abundant environmental sequences were also co-localised with characterised and putative phosphonate degradation genes (Fig. 4). In broad agreement with *aepX*, the cumulative transcription of the two markers *phnA* and *phnX* is significantly greater than the C-P lyase marker *phnJ* (Kruskal–wallis  $X^2 = 206.6$ ,  $p < 0.001$ ), strengthening the observation that 2AEP mineralisation is a major oceanic process (Supplementary Fig. 9). The *phnZ* marker is split into several subclades, with only the original PhnY\*Z specific for 2AEP (Supplementary Fig. 10). This 2AEP-specific form was found at very few sites (MG = 9; MT = 2) and in low abundance in both the MG and MT (Supplementary Fig. 10). Homologs related to the two PhnZ clades associated with either methylphosphonate<sup>35,36</sup> or (N)-trimethyl-2-aminoethylphosphonate<sup>56</sup> degradation were found at comparable gene and transcript abundances to *phnJ* and significantly lower than *phnA* (Supplementary Fig. 11).

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 significantly lower in the mesopelagic (post-hoc Dunn’s test  $z = 10.1$  and  $10.8$ ,  $p < 0.001$ , and  $z = 4.6$  and  $5.5$ ,  $p < 0.001$  respectively) (Fig. 5a and S12).

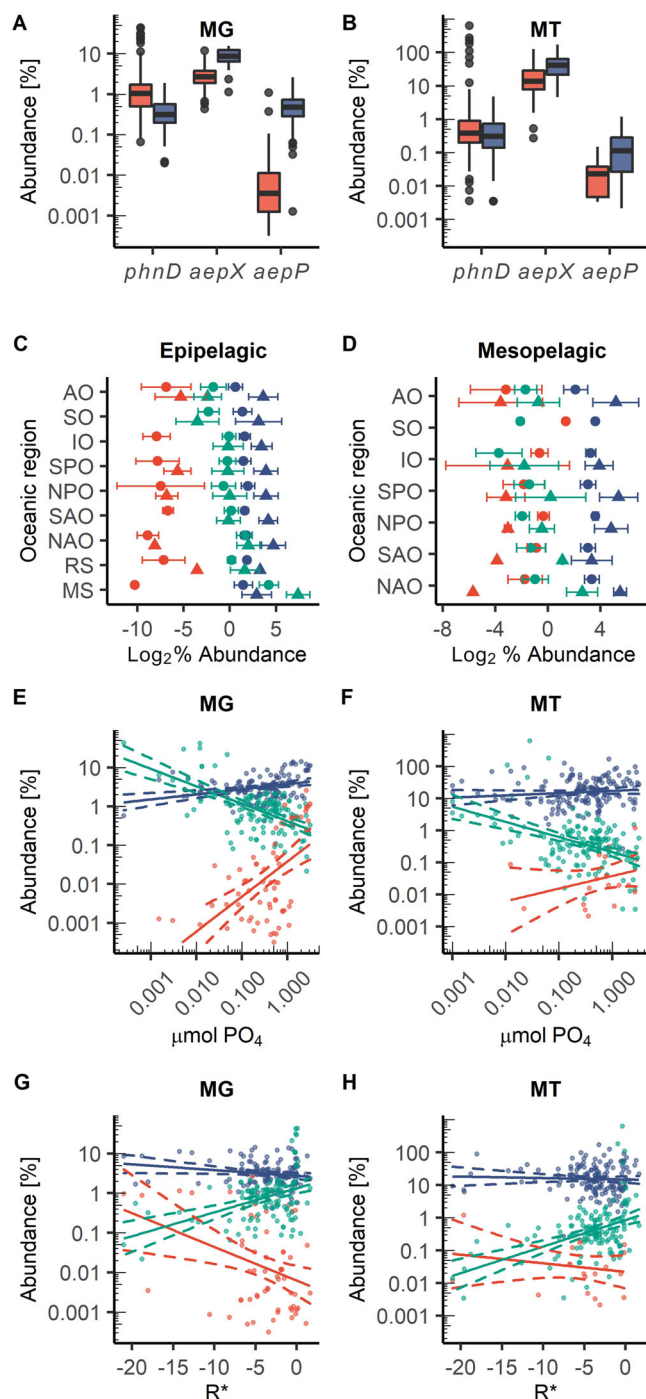
Unlike *phnD* and *aepP*, *aepX* abundance was comparable across all oceanic regions within both MG and MT at each depth suggesting 2AEP mineralisation is a ubiquitous process in seawater (Fig. 5c, d). For all sites at each depth, the relative abundance of *aepX* transcripts was always significantly greater (Wilcoxon rank sum  $W = 3771$ ,  $p < 0.001$ , estimated  $\text{log}_2$  difference =  $2.24$  (95% CI  $1.97$ – $2.52$ )) than its own gene abundance. For *phnD*, we observed significantly greater transcript abundance compared to its own gene abundance only in the Mediterranean Sea, a region typified by Pi-depletion (Wilcoxon rank sum  $W = 0$ ,  $p < 0.001$ , estimated  $\text{log}_2$  difference =  $3.11$  (95% CI  $1.90$ – $4.42$ )). Finally, in agreement with previous work<sup>20</sup>, *phnD* gene abundance was inversely correlated ( $R^2 = 0.340$ ,  $p < 0.001$ ) with standing stock concentrations of Pi (Fig. 5e), as was *phnD* transcript abundance (Fig. 5f). In contrast, *aepX* and *aepP* gene abundance were positively correlated ( $R^2 = 0.098$ ,  $p < 0.001$  and  $R^2 = 0.291$ ,  $p < 0.001$ , respectively) with Pi concentration (Fig. 5e), whilst no significant relationship between Pi concentration and *aepX/aepP* transcript abundance was found (Fig. 5f), suggesting their expression is independent of Pi in seawater globally.

To better understand the parameters controlling 2AEP catabolism in the ocean, we compared both gene and transcript abundance in relation to  $R^*$ , a measure of N vs P limitation calculated as  $[\text{NO}_2] + [\text{NO}_3] - 16[\text{PO}_4]$  (adapted from Smith et al.<sup>57</sup>). As expected, *phnD* gene and transcript abundance were positively correlated with  $R^*$  ( $R^2 = 0.168$ ,  $p < 0.001$  and  $R^2 = 0.197$ ,  $p < 0.001$ , respectively) (Fig. 5g, h), i.e. regions typified by Pi depletion, though Pi concentration alone was a better predictor of *phnD* gene abundance (Fig. 5e). However, *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.





**Fig. 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 operational taxonomic units (OTUs) and letters indicate isolates, metagenome-assembled genomes (MAGs) or single-cell amplified genomes (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<sup>84</sup>. 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 OTUs (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.



**Fig. 5 Distribution and expression of phosphonate transporter genes in the global ocean.** Abundance (% abundance [gene or transcript] relative to the median abundance [gene or transcript] of 10 single copy core genes) of *phnD*, *aepX*, *aepP* in marine epipelagic (red) and mesopelagic (blue) waters, split by metagenome (MG) (**a**) (epipelagic: *phnD*  $n = 137$ , *aepX*  $n = 137$ , *aepP*  $n = 60$ , mesopelagic: *phnD*  $n = 43$ , *aepX*  $n = 43$ , *aepP*  $n = 42$ , where  $n$  equals the number of biologically independent sampling sites where the genes were located), and metatranscriptome (MT) (**b**) (epipelagic: *phnD*  $n = 148$ , *aepX*  $n = 154$ , *aepP*  $n = 11$ , mesopelagic: *phnD*  $n = 32$ , *aepX*  $n = 33$ , *aepP*  $n = 18$ , where  $n$  equals the number of biologically independent sampling sites where the transcripts were located). Abundance 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. Circles are mean values of  $\text{Log}_2$  abundance, error bars represent standard deviation of the mean. Epipelagic (**c**) *phnD* AO MG  $n = 29$ , MT  $n = 26$ , SO MG  $n = 3$ , MT  $n = 8$ , IO MG  $n = 21$ , MT  $n = 18$ , SPO MG  $n = 25$ , MT  $n = 35$ , NPO MG  $n = 11$ , MT  $n = 20$ , SAO MG  $n = 14$ , MT  $n = 17$ , NAO MG  $n = 16$ , MT  $n = 17$ , RS MG  $n = 6$ , MT  $n = 3$ , MS MG  $n = 12$ , MT  $n = 7$ , *aepX* AO MG  $n = 29$ , MT  $n = 28$ , SO MG  $n = 3$ , MT  $n = 8$ , IO MG  $n = 21$ , MT  $n = 19$ , SPO MG  $n = 25$ , MT  $n = 35$ , NPO MG  $n = 11$ , MT  $n = 20$ , SAO MG  $n = 14$ , MT  $n = 17$ , NAO MG  $n = 16$ , MT  $n = 17$ , RS MG  $n = 6$ , MT  $n = 3$ , MS MG  $n = 12$ , MT  $n = 7$ , *aepP* AO MG  $n = 11$ , MT  $n = 3$ , SO MG  $n = 0$ , MT  $n = 0$ , IO MG  $n = 10$ , MT  $n = 0$ , SPO MG  $n = 13$ , MT  $n = 4$ , NPO MG  $n = 5$ , MT  $n = 2$ , SAO MG  $n = 4$ , MT  $n = 0$ , NAO MG  $n = 11$ , MT  $n = 1$ , RS MG  $n = 4$ , MT  $n = 1$ , MS MG  $n = 2$ , MT  $n = 0$ . Mesopelagic (**d**) *phnD* AO MG  $n = 9$ , MT  $n = 7$ , SO MG  $n = 1$ , MT  $n = 0$ , IO MG  $n = 6$ , MT  $n = 4$ , SPO MG  $n = 9$ , MT  $n = 9$ , NPO MG  $n = 5$ , MT  $n = 5$ , SAO MG  $n = 5$ , MT  $n = 2$ , NAO MG  $n = 8$ , MT  $n = 5$ , *aepX* AO MG  $n = 9$ , MT  $n = 7$ , SO MG  $n = 1$ , MT  $n = 0$ , IO MG  $n = 6$ , MT  $n = 4$ , SPO MG  $n = 9$ , MT  $n = 9$ , NPO MG  $n = 5$ , MT  $n = 5$ , SAO MG  $n = 5$ , MT  $n = 2$ , NAO MG  $n = 8$ , MT  $n = 6$ , *aepP* AO MG  $n = 9$ , MT  $n = 6$ , SO MG  $n = 1$ , MT  $n = 0$ , IO MG  $n = 6$ , MT  $n = 2$ , SPO MG  $n = 9$ , MT  $n = 6$ , NPO MG  $n = 5$ , MT  $n = 2$ , SAO MG  $n = 5$ , MT  $n = 1$ , NAO MG  $n = 7$ , MT  $n = 1$ . The relationship between the standing stock  $\text{Pi}$  concentration and transporter abundance, analysed by linear regression of  $\text{Log}_{10}$   $\text{Pi}$  concentration and  $\text{Log}_{10}$  gene/transcript abundance, in the MG (**e**), (*aepX*  $R^2 = 0.098$ ,  $p = 1.477\text{e-}4$ , *phnD*  $R^2 = 0.340$ ,  $p = 1.349\text{e-}13$ , *aepP*  $R^2 = 0.291$ ,  $p = 1.466\text{e-}6$ ) and MT (**f**), (*aepX*  $R^2 = 0.007$ ,  $p = 0.1544$ , *phnD*  $R^2 = 0.203$ ,  $p = 4.066\text{e-}9$ , *aepP*  $R^2 = 0.058$ ,  $p = 0.1581$ ). *aepP* (red), *phnD* (green), *aepX* (blue). 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  $\text{Pi}$  concentration, and transporter abundance, analysed by linear regression of  $R^*$  and  $\text{Log}_{10}$  gene/transcript abundance, in the MG (**g**) (*aepX*  $R^2 = 0.029$ ,  $p = 0.033$ , *phnD*  $R^2 = 0.168$ ,  $p = 1.611\text{e-}6$ , *aepP*  $R^2 = 0.108$ ,  $p = 5.31\text{e-}3$ ) and MT (**h**) (*aepX*  $R^2 = -0.005$ ,  $p = 0.606$ , *phnD*  $R^2 = 0.197$ ,  $p = 1.693\text{e-}8$ , *aepP*  $R^2 = -0.014$ ,  $p = 0.402$ ). *aepP* (red), *phnD* (green), *aepX* (blue). 95% confidence intervals are shown by dashed lines. In (**a**, **b**) data are represented as boxplots, where the middle line is the median and the upper and lower hinges correspond to the first and third quartiles. The upper whisker extends from the upper hinge to the largest value that is no more than  $1.5 \times \text{IQR}$  (inter-quartile range) from the upper hinge, and the lower whisker extends from the lower hinge to the smallest value that is no further than  $1.5 \times \text{IQR}$  from the lower hinge. Data beyond the ends of the whiskers are outlying points that are plotted individually. Note: *aepP* transcripts were not detected at numerous sites which is represented by an omission of data points.

## Discussion

Both phosphonate biosynthesis<sup>1,3,5,58</sup> and catabolic<sup>1,23,25,31,59–64</sup> 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<sup>28,42,43</sup>. The abundance of these  $\text{Pi}$ -sensitive transporters in marine systems is not equivalent to the abundance of catabolic genes<sup>20</sup>, especially those (*phnWAY*) recently shown to be involved in  $\text{Pi}$ -insensitive catabolism<sup>22</sup>. 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<sup>39,40,65</sup>. 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<sup>21,38</sup>. 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<sup>9,66</sup>. Here, identification of 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.

To date, whilst several ABC transporters show preference for methylphosphonate and phosphite<sup>53,67</sup>, or bind a range of phosphonates with low micromolar or lower  $K_d$ <sup>43</sup>, no ABC transporter showing a strong preference for 2AEP has been identified. Here, we revealed AepX appears to be highly specific for 2AEP and has substantially lower affinity for methylphosphonate or ethylphosphonate than PhnD<sup>43</sup>. The occurrence of *aepXVW* adjacent to putative phosphonate catabolic genes, and the characterised PbfA<sup>37</sup>, does suggest some degree of promiscuous binding, albeit likely to related aminophosphonates. Thus, the molecular mechanisms governing the specificity of AepX towards aminophosphonates warrant further investigation.

Genomic and biochemical analyses have revealed 2AEP and the alkylphosphonates, methylphosphonate and hydroxyethylphosphonate are ubiquitously synthesised in the marine environment in relatively large quantities<sup>1,3,9,25,66</sup>. However, several studies have shown 2AEP is not detected as a significant component of 'semi-labile' DOM whilst alkylphosphonates tend to accumulate<sup>16,17,68</sup>. Collectively, this would suggest 2AEP is more susceptible to microbial mineralisation<sup>22</sup> and thus shorter residence times. Here, we reveal genes 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 suggest 2AEP is preferentially mineralised independently of both N and P status, explaining why phosphonates are metabolised in regions where Pi or ammonium concentrations are high enough to repress C-P lyase-expression<sup>20,21,69</sup>. Together, this adds further weight to the notion that phosphonates are rapidly cycled between reduced and oxidised forms and are a source of regenerated Pi throughout the water column<sup>9,15,22</sup>.

Substrate inducible expression of catabolic genes targeting organic N molecules, irrespective of nutrient status, has previously been shown to drive mineralisation of N and cross feed into surrounding microbial cells<sup>70–73</sup>. Indeed, ammonium mineralisation may also occur if 2AEP, (R)-1-hydroxy-2-aminoethylphosphonate or (N)-trimethyl-2-aminoethylphosphonate are also used as carbon and energy sources, similar to methylamines and quaternary amines<sup>70,71</sup>. In agreement with Chin et al.<sup>22</sup>, our proteomic data for *S. stellulata* and in situ environmental data strongly suggests PhnA and AepX are highly synthesised in a substrate-inducible manner that would facilitate the remineralisation of labile inorganic N and P. Even if ammonium concentration does play a role in the occurrence and regulation of 2AEP degradation genes (i.e. 2AEP is primarily a N source), our combined data clearly demonstrates the potential for the in situ mineralisation of semi-recalcitrant DOP into labile Pi, a mechanism which is important for maintaining biological production in Pi-deplete regions of the ocean<sup>74,75</sup>. It should be noted that the *Pseudomonas aepR* located adjacent to *aepP* is most similar to marine *aepR*; this could explain why AepP was more abundant in the presence of 2AEP (Fig. 1d), although the differences in growth rate and protein abundance suggest substrate induction is not the sole mechanism of regulation in *P. putida* BIRD-1.

In summary, this study identified three 2AEP transporters in marine and terrestrial bacteria capable of Pi-insensitive 2AEP catabolism. The substrate-binding protein for one of these, AepX, is the most abundant and highly transcribed of the known

phosphonate binding proteins in seawater. Therefore, 2AEP mineralisation may represent a major process in the marine organic N and P cycles and a significant source of regenerated labile Pi 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 P currency between autotrophic and heterotrophic microbes<sup>9</sup>.

## Methods

**Bacterial strains and growth conditions.** All strains used in this work were axenic. *Pseudomonas* strains 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 strains were maintained on similar plates containing the appropriate antibiotic. For all growth and proteomics experiments cultures were grown in an adapted minimal A medium<sup>44</sup> using Na-Succinate (20 mM) as the sole carbon source and, where applicable, 10 mM  $\text{NH}_4\text{Cl}$  was added as the sole N source. 2AEP and  $\text{KH}_2\text{PO}_4$  were added to a final concentration of 100  $\mu\text{M}$  or 1.5 mM as specified in the text and figure legends. *Pseudomonas* strains were pre-cultured in minimal medium A containing 100  $\mu\text{M}$  Pi and 1.5 mM  $\text{NH}_4\text{Cl}$  to ensure adequate growth while minimising the potential for carryover of residual nutrients into experimental cultures. 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 proteomics where 2AEP is used as an N source, *S. stellulata* was grown in modified defined marine ammonium mineral salt (MAMS) media lacking ammonium and with HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) replacing the Pi buffer<sup>71</sup>. For all other growth and proteomics experiments these marine bacteria were grown in Sea Salts media<sup>76</sup>, with Pi buffer replaced by HEPES buffer. The method used for quantifying extracellular Pi is detailed in the Supplementary Materials.

**Generation and complementation of *Pseudomonas* mutants.** Mutants were generated and complemented via the protocols outlined in<sup>44,77</sup>, detailed descriptions of which are outlined in the Supplementary Materials. A full list of strains, plasmids, and primers used in this study is presented in Supplementary Table 4.

**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 ( $\text{OD}_{540}$  0.8–1.0) and pelleting cells (centrifugation at 16200  $\times g$  for 5 mins). Cell lysis was achieved via boiling in 100  $\mu\text{l}$  lithium dodecyl sulphate (LDS) buffer (Expediton) prior to loading 20  $\mu\text{l}$  onto a 4–20% Bis-Tris sodium dodecyl sulphate (SDS) precast gel (Expediton). For enrichment of the membrane protein fraction of the *P. putida* BIRD-1  $\Delta aepXVW$  mutant, we adapted the methods outlined in<sup>44</sup>. A full description of this protocol is outlined in the Supplementary Materials. Gel sections were de-stained with 50 mM ammonium bicarbonate in 50% (v/v) ethanol, dehydrated with 100% ethanol, reduced and alkylated with Tris-2-carboxyethylphosphine (TCEP) and iodoacetamide (IAA), washed with 50 mM ammonium bicarbonate in 50% (v/v) ethanol and dehydrated with 100% ethanol prior to overnight digestion with trypsin. Peptides were extracted and analysed using an Orbitrap Fusion Ultimate 3000 RSLC Nano System (Thermo Scientific) in electrospray ionisation mode at the Warwick Proteomics Research Technology Platform.

Resulting tandem mass spectrometry (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<sup>78</sup> with default settings and quantification was achieved using Label Free Quantification (LFQ). The proteomics analysis software Perseus (1.6.12)<sup>79</sup> 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  $s_0$  constant of 0.1, or ANOVA, where appropriate.

**Ligand binding affinity of recombinant *S. stellulata* AepX.** Recombinant AepX was over-produced and purified to homogeneity using methods described in<sup>53</sup>. Binding affinity was determined by microscale thermophoresis using a Monolith NT.115 instrument (NanoTemper Technologies GmbH, Germany) following the protocols described in<sup>53,54</sup>. A detailed protocol is outlined in the Supplementary Materials.

**Bioinformatics analyses of *aepXVW/aepP*.** ORFs encoding AepX homologs were identified using the IMG/JGI (<https://img.jgi.doe.gov/>) 'Customised 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 Supplementary Table 2). Protein sequences were aligned using ClustalOmega<sup>80</sup> and profile Hidden Markov Models (pHMMs) were constructed from these sequences using the hmmbuild function of hmmer 3.3 (<http://hmmer.org>). The previously characterised *E. coli* K-12 PhnD<sup>81</sup> and the SAR11 clade isolate *Pelagibacter* sp. HTCC7211 PhnD<sup>21</sup> showed surprisingly low sequence homology (BLAST ID 28.46%, query coverage 76%, e-value  $2 \times 10^{-25}$ ). We therefore, developed two pHMMs for PhnD to reflect this. There was no overlap in environmental sequences retrieved from each search using either pHMM. 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<sup>55</sup>, using a stringency of  $1 \times 10^{-80}$ . Sequence abundances were expressed as the 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<sup>82</sup> for both MG and MT. The pHMMs were used to search the soil MG via a hmmsearch<sup>83</sup> using the same stringency as above. Similarly, abundances were calculated as the average percentage of genomes containing a copy as above.

Phylogenetic analyses were performed using IQ-TREE 2<sup>84</sup> 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/>)<sup>85</sup>.

**Statistical analysis.** Unless specified above, all statistical analysis was performed using R (version 4.02)<sup>86</sup>, within the RStudio programme (version 1.3).

**Reproducibility.** All growth experiments were performed a minimum of two times. Proteomics experiments were performed once.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The single-culture proteomic data generated in this study have been deposited in the ProteomeXchange Consortium via the PRIDE database (accession code PXD026804 [10.6019/PXD026804]). The TARA Oceans metagenomic and metatranscriptomic data used in this study are available in the European Nucleotide Archive database under accession code [PRJEB7988](https://www.ebi.ac.uk/ena/record/PRJEB7988).

Received: 10 December 2020; Accepted: 29 June 2021;

Published online: 27 July 2021

## References

- 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 (2012).
- Mukhamedova, K. S. & Glushenkova, A. I. Natural Phosphonolipids. *Chem. Nat. Compd.* **36**, 329–341 (2000).
- Yu, X. et al. Diversity and abundance of phosphonate biosynthetic genes in nature. *Proc. Natl. Acad. Sci. USA* **110**, 20759–20764 (2013).
- Ju, K.-S., Doroghazi, J. R. & Metcalf, W. W. Genomics-enabled discovery of phosphonate natural products and their biosynthetic pathways. *J. Ind. Microbiol. Biotechnol.* **41**, 345–356 (2014).
- 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 (2013).
- Metcalf, W. W. et al. Synthesis of methylphosphonic acid by marine microbes: a source for methane in the aerobic ocean. *Science* **337**, 1104–1107 (2012).
- Born, D. A. et al. Structural basis for methylphosphonate biosynthesis. *Science* **358**, 1336 (2017).
- Hildebrand, R. The Role of Phosphonates in Living Systems. (CRC Press, 1983).
- Van Mooy, B. A. S. et al. Major role of planktonic phosphate reduction in the marine phosphorus redox cycle. *Science* **348**, 783 (2015).
- Kolowitz, L. C., Ingall, E. D. & Benner, R. Composition and cycling of marine organic phosphorus. *Limnol. Oceanogr.* **46**, 309–320 (2001).
- Clark, L. L., Ingall, E. D. & Benner, R. Marine phosphorus is selectively remineralized. *Nature* **393**, 426–426 (1998).
- 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 (2004).
- Tate, K. R. & Newman, R. H. Phosphorus fractions of a climosequence of soils in New Zealand tussock grassland. *Soil Biol. Biochem.* **14**, 191–196 (1982).
- Cade-Menun, B. J., Navaratnam, J. A. & Walbridge, M. R. Characterizing dissolved and particulate phosphorus in water with <sup>31</sup>P nuclear magnetic resonance spectroscopy. *Environ. Sci. Technol.* **40**, 7874–7880 (2006).
- Clark, L. L., Ingall, E. D. & Benner, R. Marine organic phosphorus cycling: novel insights from nuclear magnetic resonance. *Am. J. Sci.* **299**, 724–737 (1999).
- Repeta, D. J. et al. Marine methane paradox explained by bacterial degradation of dissolved organic matter. *Nat. Geosci.* **9**, 884–887 (2016).
- 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 (2020).
- 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. *Nat. Geosci.* **2**, 696–699 (2009).
- Acker, M. et al. Phosphonate production by marine microbes: exploring new sources and potential function. *bioRxiv*, 2020.2011.2004.368217, <https://doi.org/10.1101/2020.11.04.368217> (2020).
- 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 (2019).
- Carini, P., White, A. E., Campbell, E. O. & Giovannoni, S. J. Methane production by phosphate-starved SAR11 chemoheterotrophic marine bacteria. *Nat. Commun.* **5**, 4346 (2014).
- Chin, J. P., Quinn, J. P. & McGrath, J. W. Phosphate insensitive aminophosphonate mineralisation within oceanic nutrient cycles. *ISME J.* **12**, 973–980 (2018).
- 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 (1998).
- White, A. K. & Metcalf, W. W. Microbial metabolism of reduced phosphorus compounds. *Annu. Rev. Microbiol.* **61**, 379–400 (2007).
- 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 (2010).
- Wanner, B. L. Molecular genetics of carbon-phosphorus bond cleavage in bacteria. *Biodegradation* **5**, 175–184 (1994).
- Kononova, S. V. & Nesmeyanova, M. A. Phosphonates and their degradation by microorganisms. *Biochem. (Mosc.)* **67**, 184–195 (2002).
- 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 (1995).
- Kim, A. D. et al. The 2-aminoethylphosphonate-specific transaminase of the 2-aminoethylphosphonate degradation pathway. *J. Bacteriol.* **184**, 4134 (2002).
- Cooley, N. A. et al. Phosphonoacetate biosynthesis: In vitro detection of a novel NADP<sup>+</sup>-dependent phosphonoacetaldehyde-oxidizing activity in cell-extracts of a marine Roseobacter. *Microbiology* **80**, 335–340 (2011).
- 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 (2011).
- Agarwal, V. et al. Structure and function of phosphonoacetaldehyde dehydrogenase: The missing link in phosphonoacetate formation. *Chem. Biol.* **21**, 125–135 (2014).
- McSorley, F. R. et al. PhnY and PhnZ comprise a new oxidative pathway for enzymatic cleavage of a carbon-phosphorus bond. *J. Am. Chem. Soc.* **134**, 8364–8367 (2012).
- van Staalduinen, L. M. et al. Crystal structure of PhnZ in complex with substrate reveals a di-iron oxygenase mechanism for catabolism of organophosphonates. *Proc. Natl. Acad. Sci. USA* **111**, 5171–5176 (2014).
- Sosa, O. A., Casey, J. R. & Karl, D. M. Methylphosphonate oxidation in *Prochlorococcus* strain MIT9301 supports phosphate acquisition, formate excretion, and carbon assimilation into purines. *Appl. Environ. Microbiol.* **85**, e00289–e00219 (2019).
- Gama, S. R. et al. An oxidative pathway for microbial utilization of methylphosphonic acid as a phosphate source. *ACS Chem. Biol.* **14**, 735–741 (2019).
- Zangelmi, E. et al. Discovery of a new, recurrent enzyme in bacterial phosphonate degradation: (R)-1-hydroxy-2-aminoethylphosphonate ammonia-lyase. *Biochemistry* **60**, 1214–1225 (2021).
- Sowell, S. M. et al. Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J.* **3**, 93–105 (2008).



39. Ottesen, E. A. et al. Pattern and synchrony of gene expression among sympatric marine microbial populations. *Proc. Natl Acad. Sci. USA* **110**, E488–E497 (2013).
40. Lidbury, I., Murrell, J. C. & Chen, Y. Trimethylamine *N*-oxide metabolism by abundant marine heterotrophic bacteria. *Proc. Natl Acad. Sci. USA* **111**, 2710–2715 (2014).
41. Mauchline, T. H. et al. Mapping the *Sinorhizobium meliloti* 1021 solute-binding protein-dependent transportome. *Proc. Natl Acad. Sci. USA* **103**, 17933–17938 (2006).
42. Alicea, I. et al. Structure of the *Escherichia coli* phosphonate binding protein PhnD and rationally optimized phosphonate biosensors. *J. Mol. Biol.* **414**, 356–369 (2011).
43. 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 (2006).
44. 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 (2016).
45. Lemieux, M. J., Huang, Y. & Wang, D. N. Crystal structure and mechanism of GltP, the glycerol-3-phosphate transporter from *E. coli*. *J. Electron Microsc.* **54**, i43–i46 (2005). Suppl. 1.
46. Elvin, C. M., Hardy, C. M. & Rosenberg, H. Pi exchange mediated by the GltP-dependent sn-glycerol-3-phosphate transport system in *Escherichia coli*. *J. Bacteriol.* **161**, 1054–1058 (1985).
47. Enkavi, G. & Tajkhorshid, E. Simulation of spontaneous substrate binding revealing the binding pathway and mechanism and initial conformational response of GltP. *Biochemistry* **49**, 1105–1114 (2010).
48. Law, C. J. et al. Salt-bridge dynamics control substrate-induced conformational change in the membrane transporter GltP. *J. Mol. Biol.* **378**, 828–839 (2008).
49. Law, C. J., Enkavi, G., Wang, D.-N. & Tajkhorshid, E. Structural basis of substrate selectivity in the glycerol-3-phosphate: phosphate antiporter GltP. *Biophysical J.* **97**, 1346–1353 (2009).
50. Moradi, M., Enkavi, G. & Tajkhorshid, E. Atomic-level characterization of transport cycle thermodynamics in the glycerol-3-phosphate: phosphate antiporter. *Nat. Comm.* **6**, 8393 (2015).
51. Kulakova, A. N. et al. Structural and functional analysis of the phosphonoacetate hydrolase (*phnA*) gene region in *Pseudomonas fluorescens* 23F. *J. Bacteriol.* **183**, 3268 (2001).
52. Kulakova, A. N. et al. Expression of the phosphonoalanine-degradative gene cluster from *Variovorax* sp. Pal2 is induced by growth on phosphonoalanine and phosphonopyruvate. *FEMS Microbiol. Lett.* **292**, 100–106 (2009).
53. Bisson, C. et al. The molecular basis of phosphite and hypophosphite recognition by ABC-transporters. *Nat. Comm.* **8**, 1746 (2017).
54. Adams, N. B. P., Robertson, A. J., Hunter, C. N., Hitchcock, A. & Bisson, C. Phosphite binding by the HtxB periplasmic binding protein depends on the protonation state of the ligand. *Sci. Rep.* **9**, 10231–10231 (2019).
55. Villar, E. et al. The Ocean Gene Atlas: exploring the biogeography of plankton genes online. *Nucleic Acids Res.* **46**, W289–W295 (2018).
56. Rajakovich, L. J. et al. A new microbial pathway for organophosphonate degradation catalyzed by two previously misannotated non-heme-iron oxygenases. *Biochemistry* **58**, 1627–1647 (2019).
57. 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 (2019).
58. 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 (2016).
59. 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 (2004).
60. Ermakova, I. T. et al. Organophosphonates utilization by soil strains of *Ochrobactrum anthropi* and *Achromobacter* sp. *Arch. Microbiol.* **199**, 665–675 (2017).
61. 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 (2009).
62. Imazu, K. et al. Enhanced utilization of phosphonate and phosphite by *Klebsiella aerogenes*. *Appl. Environ. Microbiol.* **64**, 3754–3758 (1998).
63. Mendz, G. L., Mégraud, F. & Korolik, V. Phosphonate catabolism by *Campylobacter* spp. *Arch. Microbiol.* **183**, 113–120 (2005).
64. 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 (1993).
65. Hultman, J. et al. Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature* **521**, 208–212 (2015).
66. Frischkorn, K. R. et al. *Trichodesmium* physiological ecology and phosphate reduction in the western tropical South Pacific. *Biogeosciences* **15**, 5761–5778 (2018).
67. Feingersch, R. et al. Potential for phosphite and phosphonate utilization by *Prochlorococcus*. *ISME J.* **6**, 827–834 (2012).
68. Young, C. L. & Ingall, E. D. Marine dissolved organic phosphorus composition: insights from samples recovered using combined electrodialysis/reverse osmosis. *Aquat. Geochem.* **16**, 563–574 (2010).
69. Benitez-Nelson, C. R., O'Neill, L., Kolowitz, L. C., Pellechia, P. & Thunell, R. Phosphonates and particulate organic phosphorus cycling in an anoxic marine basin. *Limnol. Oceanogr.* **49**, 1593–1604 (2004).
70. 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 (2015).
71. 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).
72. Zecher, K., Hayes, K. R. & Philipp, B. Evidence of interdomain ammonium cross-feeding from methylamine- and glycine betaine-degrading *Rhodobacteraceae* to diatoms as a widespread interaction in the marine phycosphere. *Front Microbiol.* **11**, 2431 (2020).
73. Christie-Oleza, J. A., Sousoni, D., Lloyd, M., Armengaud, J. & Scanlan, D. J. Nutrient recycling facilitates long-term stability of marine microbial phototroph–heterotroph interactions. *Nat. Microbiol.* **2**, 17100 (2017).
74. Dyhrman, S. T. et al. Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. *Nature* **439**, 68–71 (2006).
75. Karl, D. M. et al. Aerobic production of methane in the sea. *Nat. Geosci.* **1**, 473–478 (2008).
76. 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 (2012).
77. 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 (2017).
78. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
79. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
80. Madeira, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
81. Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29–W37 (2011).
82. Wackett, L. P., Wanner, B. L., Venditti, C. P. & Walsh, C. T. Involvement of the phosphate regulon and the *psdI* locus in carbon-phosphorus lyase activity of *Escherichia coli* K-12. *J. Bacteriol.* **169**, 1753–1756 (1987).
83. Milanese, A. et al. Microbial abundance, activity and population genomic profiling with mOTUs2. *Nat. Comm.* **10**, 1014 (2019).
84. 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 (2020).
85. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259 (2019).
86. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/> (2020).

## 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, BB/T009152/1 and NE/S013539/1 linked to The Soil and Rhizosphere Interactions for Sustainable Agri-ecosystems (SARISA) programme and a Discovery Fellowship (IL) and NERC Environmental 'Omics Synthesis Grant (IL and EW), respectively.

## Author contributions

A.M. and I.L. wrote the manuscript with comprehensive feedback and guidance from Y.C. and D.S. A.M. performed the experimental work, and A.M. and I.L. analysed the data, except for the binding affinity assays, performed and analysed by N.A., W.C., A.H. and I.L. I.L. and A.M. conceptualised and developed the research, respectively. A.B. provided input and support for proteomic data generation and analyses. All authors contributed to revisions of the manuscript.

## Competing interests

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-24646-z>.

**Correspondence** and requests for materials should be addressed to I.D.E.A.L.

**Peer review information** *Nature Communications* thanks P. Dreux Chappell, Juan Villarreal Chiu, and other, anonymous, reviewers for their contributions to the peer review of this work. Peer review reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021