environmental microbiology

Environmental Microbiology (2017) **19**(6), 2320–2333 doi:10.1111/1462-2920.13726

Aquatic adaptation of a laterally acquired pectin degradation pathway in marine gammaproteobacteria

Jan-Hendrik Hehemann,[1,2*†] Le Van Truong,[3,4†] Frank Unfried,[2,3,5] Norma Welsch,[5] Johannes Kabisch,[5,6] Stefan E. Heiden,[5] Sabryna Junker,[7] Dörte Becher,[7] Andrea Thürmer,[8] Rolf Daniel,[8] Rudolf Amann[2] and Thomas Schweder[3,5]**

[1]*MARUM, Center for Marine Environmental Sciences at the University of Bremen, Leobener Strasse, Bremen, D-28359, Germany.*

[2]*Max Planck Institute for Marine Microbiology, Celsiusstrasse 1, Bremen, D-28359, Germany.*

[3]*Institute of Marine Biotechnology, W.-Rathenau-Str. 49a, Greifswald, D-17489, Germany.*

[4]*Institute of Biotechnology, Vietnamese Academy of Science and Technology, 18 Hoang Quoc Viet, Hanoi, Vietnam.*

[5]*Pharmaceutical Biotechnology, Institute of Pharmacy, Ernst-Moritz-Arndt-University, Felix-Hausdorff-Str. 3, Greifswald, D-17487, Germany.*

[6]*Department of Biology, Computer-aided Synthetic Biology, Technische Universität Darmstadt, Schnittspahnstr. 10, Darmstadt, D-64287, Germany.*

[7]*Institute of Microbiology, Ernst-Moritz-Arndt-University, Friedrich-Ludwig-Jahn-Str. 15, Greifswald, D-17487, Germany.*

[8]*Göttingen Genomics Laboratory (G2L), Institute of Microbiology and Genetics, University of Göttingen, Grisebachstr. 8, Göttingen, D-37077, Germany.*

Summary

Mobile genomic islands distribute functional traits between microbes and habitats, yet it remains unclear how their proteins adapt to new environments. Here we used a comparative phylogenomic and proteomic approach to show that the marine bacterium *Pseudoalteromonas haloplanktis* ANT/505 acquired a genomic island with a functional pathway for pectin catabolism. Bioinformatics and biochemical experiments revealed that this pathway encodes a series of carbohydrate-active enzymes including two multi-modular pectate lyases, PelA and PelB. PelA is a large enzyme with a polysaccharide lyase family 1 (PL1) domain and a carbohydrate esterase family 8 domain, and PelB contains a PL1 domain and two carbohydrate-binding domains of family 13. Comparative phylogenomic analyses indicate that the pathway was most likely acquired from terrestrial microbes, yet we observed multi-modular orthologues only in marine bacteria. Proteomic experiments showed that *P. haloplanktis* ANT/505 secretes both pectate lyases into the environment in the presence of pectin. These multi-modular enzymes may therefore represent a marine innovation that enhances physical interaction with pectins to reduce loss of substrate and enzymes by diffusion. Our results revealed that marine bacteria can catabolize pectin, and highlight enzyme fusion as a potential adaptation that may facilitate microbial consumption of polymeric substrates in aquatic environments.

Received 11 June, 2015; revised 3 March, 2017; accepted 5 March, 2017. Correspondence: *E-mail: jhhehemann@marum.de; Tel. +49 421 218 65775; Fax +49 421 218 65715. **E-mail: schweder@uni-greifswald.de; Tel. +49 3834 4204212; Fax +49 3834 4204238. †These authors contributed equally to this work.

Introduction

The transfer of functional traits between microbes from different ecosystems provides recipient cells with the molecular tools to venture into new ecological niches (Popa and Dagan, 2011; Hehemann *et al.*, 2012). Understanding how transferred pathways reliably and efficiently manifest themselves in genomes of recipient cells will help to understand evolution of microorganisms (Gogarten and Townsend, 2005) and may facilitate the identification of suitable bioengineering strategies for beneficial microbial applications (Wargacki *et al.*, 2012). The frequency of transfers with stable integration is thought to depend not only on biotic factors, such as metabolic compatibility, gene expression systems, gene-transfer mechanisms, but also on ecological factors such as proximity between interacting cells and adaptations to abiotic environments (Gogarten and Townsend, 2005). Sharing a habitat enables contact and exchange of genetic material and this appears essential in promoting transfers between cells. Accordingly, progress in microbial genome sequencing,

computational analysis and public access to microbial sequence data (Alm *et al*., 2005; Smillie *et al*., 2011) showed that gene transfers are most frequent between taxa that occupy the same habitat, while transfers between different ecosystems remain scarce (Popa and Dagan, 2011). However, we recently demonstrated that polysaccharide-degrading pathways have been transferred from marine bacteria to Japanese human gut bacteria (Hehemann *et al*., 2010), exemplifying that these transfers between different ecosystems occur in nature. Moreover, subsequent studies showed that these pathways are functional in their novel microbial host (Hehemann *et al*., 2012). These results raise the question of how transferred pathways and bacteria adapt when they are translocated into new ecosystems.

In this study, we describe a pectin-specific pathway that was transferred from terrestrial microbes to marine planktonic bacteria, therefore representing a valuable model system to pinpoint and dissect adaptive events at the protein level. This pathway was first identified in the marine bacterium *Pseudoalteromonas haloplanktis* ANT/505 (Truong *et al*., 2001), which was intriguing because pectin is synthesized in abundance by terrestrial plants (Mohnen, 2008), but is not well known as a marine polysaccharide. Accordingly, pectinolytic enzymes have been extensively found in, and were characterized from, microbes isolated from soils and from other terrestrial habitats. Pectins are anionic homogalacturonan polysaccharides of methylester-bearing α-(1,4)-linked D-galacturonate. They are depolymerized by plant pathogens, intestinal symbionts (colonic & rumenal) and soil microbes into accessible oligogalacturonides, monosaccharides and 5-keto-4-deoxyuronate by carbohydrate-active enzymes (CAZymes; CAZy database: http://www.cazy.org (Lombard *et al*., 2014)). These enzymes include polygalacturonases (GH28), pectinases and endolytic pectate lyases (e.g. PL1, PL2, PL3, PL9, PL10), carbohydrate methylesterases (CE8), exolytic pectate lyases (e.g. PL2), oligogalacturonate lyases (PL22) and the recently discovered unsaturated rhamnogalacturonyl hydrolases (GH105) (Abbott and Boraston, 2008; Collen *et al*., 2014; Lombard *et al*., 2014). We previously showed that *P. haloplanktis* ANT/505 digests pectin with the two pectate lyases PelA and PelB (Truong *et al*., 2001), both of which have catalytic domains belonging to family PL1 subfamily 5 (Lombard *et al*., 2010). We now recognized that homologues of these two pectate lyases were absent from most genomes of closely and distantly related marine *Pseudoalteromonadales*, which indicates a dynamic evolutionary history of this trait in this genus and suggests it may have derived from horizontal gene transfer (HGT).

To shed light on the genomic structure of the pectinolytic pathway we sequenced the genome of *P. haloplanktis* ANT/505. Comparative genome analyses revealed an island of ~35,000 bp. This genomic island holds the previously characterized pectate lyase PelA (Truong *et al*., 2001), encodes further essential proteins for pectin catabolism (Richard and Hilditch, 2009) and is located in the direct vicinity of genes that could play a role in the mobilization of genomic islands (Permina *et al*., 2002). A proteomic analysis showed that the different proteins of the genomic island, including PelA, a TonB-dependent receptor for oligosaccharide import (TBDR) (Reeves *et al*., 1996) and the key-enzymes for galacturonate catabolism were synthesized by *P. haloplanktis* ANT/505 when it encountered pectin. The phylogenetic analysis showed that multi-modular orthologues of PelA and PelB are only present in marine bacteria and absent from genomes of terrestrial bacteria. This may suggest that fused enzymes represent an adaptation to the dilute marine environment where they can increase interaction with the substrate and reduce loss of enzyme and substrate by diffusion. In aggregate, our results contribute to the understanding of microbial evolution, but may also prove useful for the rational design of glycan-degrading pathways, for example to bioengineer synthetic microbial processes adapted to dilute systems.

Results

Acquisition of a mobile genomic island with a pectinolytic pathway in marine bacteria

We compared the genomic location involved in pectin utilization in the pectinolytic strain *P. haloplanktis* ANT/505 (Truong *et al*., 2001) with two other marine *Pseudoalteromonadales* strains, *P. haloplanktis* TAC125 (Médigue *et al*., 2005) and *P. haloplanktis* ATCC 14393 (Gauthier *et al*., 1995; Darling *et al*., 2010). The alignment shows that the strains ANT/505 (on contig 16, accession number ADOP01000016) and ATCC 14393 (NZ_AHCA01000001) contain a genomic island with the recently detected *pelA* gene (locus tag: PH505_ap00520) (Truong *et al*., 2001) on chromosome II. This genomic island is absent in TAC125, while in ANT/505 and ATCC 14393 it appears to be inserted into an otherwise homologous genomic region, between putative iron transport (PH505_ap00410) and luciferase (PH505_ap00690) genes (Fig. 1). Furthermore, the *pelB* gene, which is not part of the genomic island, was only found in the genomes of ANT/505 (ADOP01000004; locus tag PH505_ad00690) and ATCC 14393 (NZ_AHCA01000002; locus tag PHAL_RS0106180), but not in TAC125. ANT/505 could grow with pectin as sole carbon source while TAC125 did not grow on this substrate (Supporting Information Fig. S1). The genome of TAC125 lacks not only this island but also any other known CAZyme families involved in pectin catabolism. In contrast, the genome of the related strain *P. haloplanktis* ATCC 14393 shares most of the genes, synteny and the insertion site between

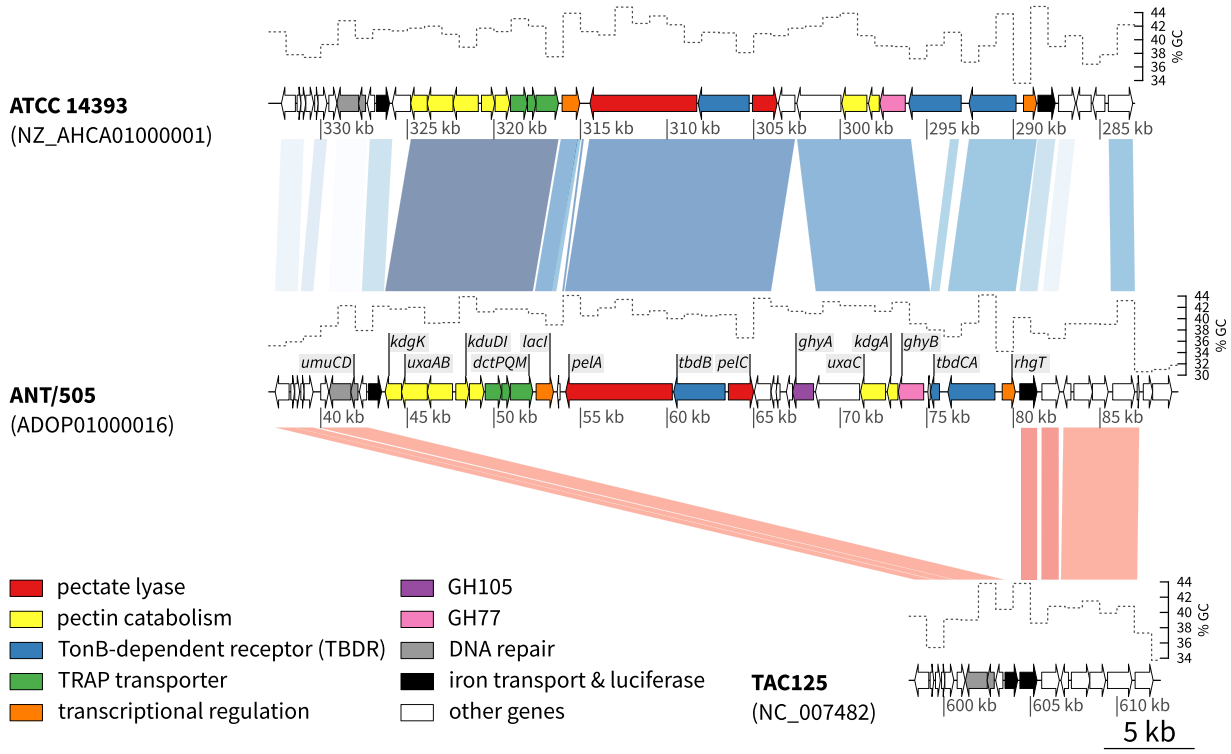

Fig. 1. Presence of a genomic island for pectin catabolism in marine *Pseudoalteromonadales*. The *P. haloplanktis* strains ANT/505 and ATCC 14393 contain a genomic island for pectin catabolism that encodes, among other genes, the pectate lyase gene *pelA* (red) for pectin degradation, TonB-dependent receptors (TBDR) (blue) for oligosaccharide import, and auxiliary catabolic genes for the catabolism of monosaccharides (yellow). The PUL is absent in an otherwise very similar genetic region in the genomes of closely related *P. haloplanktis* strains such as strain TAC125 (shown), which may be explained by gene loss or horizontal gene transfer. The black arrows signal the genes between which the island was inserted. The sequence comparisons were performed with Bl2seq (BLASTn, E-value 1e-5). Sequence similarities are depicted by red hues for direct comparisons and blue hues for reversed comparisons. Darker colours correspond to higher identities. The GC content was calculated using a non-overlapping sliding window of size 1000 bp and is shown above the genes.

a putative iron transport gene and a luciferase gene (Fig. 1, black arrows) with the strain ANT/505. It is interesting to note that we identified putative mobility genes that could play a role in the mobilization and the transfer of genomic islands such as the error prone repair proteins UmuC (PH505_ap00380) and UmuD (PH505_ap00390) (Permina *et al.*, 2002) and a phage integrase family protein. The *umuCD* genes are located immediately upstream of the island (Fig. 1, grey arrows). The gene encoding the phage integrase family protein is located about 20 genes upstream of the PelA cluster in ATCC 14393 and can be found in ANT/505 on contig061 with a similar genomic context as in ATCC 14393.

In addition to the pectate lyase gene *pelA* (Truong *et al.*, 2001), the island encodes 25 genes, many of which are involved in pectin utilization (Table 1). Seven genes code for enzymes that catalyse the different steps in the catabolism of D-galacturonic acid, the monomeric building block of pectin (Richard and Hilditch, 2009). The island contains two other putative CAZymes for glycan depolymerization, a GH105 family protein with putative unsaturated rhamnogalacturonyl hydrolase activity (Collen *et al.*, 2014) and a putative GH77 amylomaltase. Furthermore, we found a transcriptional regulator, accessory catabolic genes, and genes for a putative tripartite ATP-independent periplasmic (TRAP) transporter as well as TonB-dependent receptors (TBDR, SusC-like), with one located directly next to the *pelA* gene (Table 1).

P. haloplanktis *ANT/505 encodes two large, multi-modular pectate lyases*

CAZyme annotation (Lombard *et al.*, 2014) of the recently detected enzymes PelA and PelB (Truong *et al.*, 2001) of *P. haloplanktis* ANT/505 revealed that their pectate lyase domains are classified into the PL1 family and subfamily 5. Our bioinformatic analysis showed that the PL1 pectate lyase domain of PelB is connected with two C-terminal carbohydrate-binding modules (CBM) of family 13 (CBM13), representing to our knowledge the first example of CBM13 being involved in pectin degradation. PelB has 658 amino acids and a molecular weight of 70 kDa. PelA on the other hand is a remarkably large CAZyme with 2052 amino acids and a molecular weight of 220 kDa. The sequence

Table 1. Genes and putative protein functions of the PelA genomic island of *P. haloplanktis* ANT/505 including accession number and locus tag.

Accession number	Locus tag	Gene	Putative function
EGI73591.1	PH505_ap00420	*kdgK*	2-dehydro-3-deoxygluconate kinase (EC 2.7.1.45)
EGI73592.1	PH505_ap00430	*uxaA*	altronate dehydratase (EC 4.2.1.7)
EGI73593.1	PH505_ap00440	*uxaB*	altronate oxidoreductase (EC 1.1.1.58)
EGI73594.1	PH505_ap00450	*kduD*	2-deoxy-D-gluconate 3-dehydrogenase (EC 1.1.1.125)
EGI73595.1	PH505_ap00460	*kduI*	4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase (EC 5.3.1.17)
EGI73596.1	PH505_ap00470	*dctP*	TRAP-type C4-dicarboxylate transport system, periplasmic component
EGI73597.1	PH505_ap00480	*dctQ*	TRAP-type C4-dicarboxylate transport system, small permease component
EGI73598.1	PH505_ap00490	*dctM*	TRAP-type C4-dicarboxylate transport system, large permease component
EGI73599.1	PH505_ap00500	*lacI*	transcriptional regulator, LacI family
EGI73600.1	PH505_ap00510	*ypuA*	hypothetical protein
EGI73601.2	PH505_ap00520	*pelA*	pectate lyase, methylesterase (CE8, PL1)
EGI73602.2	PH505_ap00530	*tbdB*	TBDR (TonB-dependent receptor)
EGI73604.2	PH505_ap00550	*pelC*	pectate lyase (PL3)
EGI73605.1	PH505_ap00560	*estP*	esterase/lipase/thioesterase
EGI73606.2	PH505_ap00570	*ypuB*	hypothetical protein
EGI73607.1	PH505_ap00580	*xylB*	xylosidase/arabinosidase (cryptic gene)
EGI73608.2	PH505_ap00590	*xylA*	xylosidase/arabinosidase (cryptic gene)
EGI73609.1	PH505_ap00600	*ghyA*	unsaturated rhamnogalacturonyl hydrolase (GH105)
EGI73610.1	PH505_ap00610	*rpoA*	DNA-directed RNA polymerase, alpha subunit
EGI73611.1	PH505_ap00620	*uxaC*	uronate isomerase (EC 5.3.1.12)
EGI73612.1	PH505_ap00630	*kdgA*	2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14)
EGI73613.2	PH505_ap00640	*ghyB*	amylomaltase (GH77)
EGI73614.1	PH505_ap00650	*ypuC*	hypothetical protein
EGI73615.2	PH505_ap00660	*tbdC*	TonB-dependent receptor (cryptic gene)
EGI73616.1	PH505_ap00670	*tbdA*	TonB-dependent receptor
EGI73617.1	PH505_ap00680	*rhgT*	transcriptional regulator (GNTR)

analysis indicated that this protein combines two enzymatic functions, a putative carbohydrate methylesterase domain of family 8 (CE8) and the C-terminal PL1 domain. The combination of these two activities into one protein suggested that the CE8 domain could generate the de-esterified substrate (pectate), which represents the substrate for the PL1 domain. Because we have previously confirmed the pectate lyase activity (Truong *et al*., 2001), we here extended the biochemical analysis and tested the hypothesis that PelA is a bifunctional enzyme.

At first we cloned and overexpressed the putative methylesterase domain of PelA in *E. coli* (see Supporting Information Fig. S2). The resulting heterologous protein accumulated in inclusion bodies and could not be refolded with high yield and activity. We therefore overexpressed the full-length PelA protein in *E. coli* (Supporting Information Fig. S3) and compared the activity of the heterologous enzyme with the purified PelA protein from *P. haloplanktis* ANT/505. With pectin as substrate both showed significant carbohydrate methylesterase activity with a broad temperature range between 25°C and 45°C and a narrow pH optimum of pH 6.5–7.5 (Fig. 2). These results confirmed the role of the CE8 domain of PelA and established the different functions of this big multi-modular enzyme.

Evolution of multi-modular pectate lyases in marine microbes

The multi-modular character of PelA and PelB is illustrated in Fig. 3A. The mosaic presence of the genomic island among closely related *Pseudoalteromonas* spp. suggested a dynamic evolutionary history of pectin utilization in marine bacteria. To investigate this hypothesis we used PelA and PelB as queries to extract sequences from the non-redundant protein database at NCBI (Altschul *et al*., 1997) using BLAST (Altschul *et al*., 1990). The phylogenetic tree shows that the two PL1 domains of PelA and PelB belong to two separated clades. All multi-modular sequences that showed >90% query coverage when compared to these two enzymes are from marine bacteria (Fig. 3B) even though the PL1 family is large, with currently over 1423 bacterial homologues in the CAZy database (Lombard *et al*., 2014). The full-length homologues were found in marine Gammaproteobacteria such as *Pseudoalteromonas* spp. and *Alteromonas* spp. and other Gammaproteobacteria. Homologues were detected in strains from different geographic regions, for example, in a *P. haloplanktis* isolate from the Pacific Ocean (Hawaii) (Baumann *et al*., 1972) or *Alteromonas macleodii* strains from surface water of the Atlantic, Mediterranean and Black Sea (Lopez-Perez *et al*., 2012).

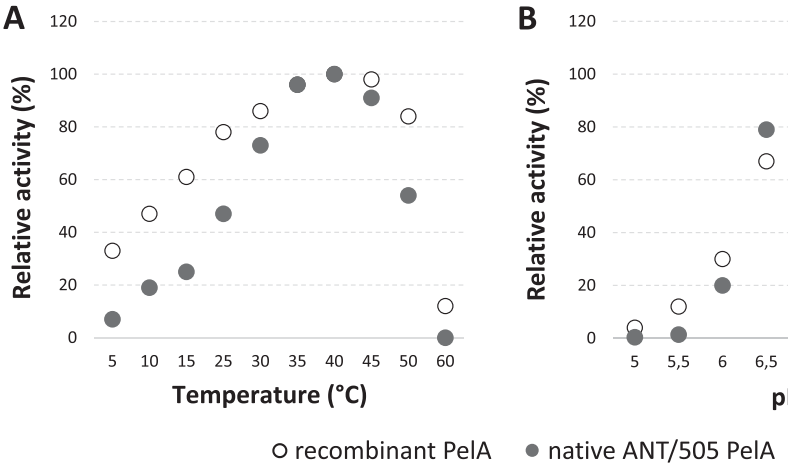

Fig. 2. Determination of the pectin methylesterase activity of the purified recombinant PelA protein overexpressed in *E. coli* (*open circles*) in comparison to the native purified PelA protein from *P. haloplanktis* ANT/505 (*closed circles*) at different temperatures and pH conditions. The data presented are relative data in percent. The maximal value was set as 100%. Values are the mean of two independent biological replicates. Replicates were run on separate days with independent samples.

We expanded the phylogenetic analysis and constructed maximum likelihood trees with ~100 sequences obtained from GenBank using PL1 and CE8 domains of PelA (PH505_ap00520) as query (Fig. 3C and D, Supporting Information Figs. S4 and S5). In addition, phylogenetic trees for PelC (a PL3 domain enzyme; PH505_ap00550), another CAZyme from the genomic island, and an iron-binding protein (PH505_ap00410), that is located just outside the island, were generated (Fig. 3E and F, Supporting Information Figs. S6 and S7). The phylogenetic analysis shows short homologues with a separate PL1 domain were present in many terrestrial bacteria for example in *Cellvibrio japonicus* Ueda107 (DeBoy *et al*., 2008), which is specialized in the degradation of plant glycans (Larsbrink *et al*., 2014). We also found shorter homologues with only the PL1 domain in some marine bacteria and in the salt marsh bacterium *Saccharophagus degradans* 2-40, which specializes in the degradation of plant glycans and a small subset of glycans from red seaweeds (Ekborg *et al*., 2005).

However, fused homologues of PelA were only found in marine bacteria suggesting a significant association of this trait with the marine environment (Fisher's exact test, $0.00001 > p$). The sequences from *Pseudoalteromonas* spp. and *Alteromonas* spp. and other marine bacteria formed a clade that is nested within sequences and part of a monophyletic clade with terrestrial bacteria. A similar pattern was observed for a maximum likelihood tree of the CE8 domain (Fig. 3D). The phylogenetic analysis of the PL3 domain-containing PelC sequence (Fig. 3E), which is also part of the island, shows similar to the PelA domains a monophyletic clade with the marine sequences. The maximum likelihood tree of the protein involved in iron binding and transport, which is located just outside of the genomic island, is almost exclusively populated with marine sequences. Iron is a trace metal in the sea, and mechanisms involved in iron acquisition are strongly selected for in marine bacteria (Bertrand *et al*., 2015), which may explain why many of the marine *Pseudoalteromonas* spp. carry these proteins (Fig. 3F).

P. haloplanktis *ANT/505 responds to pectin and secretes PelA and PelB into the environment*

In order to experimentally characterize the sequences upstream of *pelA* and *pelB* we used primer extension to identify the correct transcriptional start sites of both genes (Supporting Information Figs. S8 and S9). These experiments enabled us to verify the predicted 5′-end of the *pelA* and *pelB* transcripts and allowed us to analyse upstream sequences of both genes. Furthermore, bioinformatic analyses of the N-terminal coding sequences with PSORTb (version 3.0.2) (Yu *et al*., 2010) gave a maximum value of 10 for PelB and of 7.91 for PelA, suggesting they are secreted enzymes.

Following this analysis, we aimed to investigate the metabolic response of *P. haloplanktis* ANT/505 to pectin. We previously demonstrated that functional PelA proteins can be purified from *P. haloplanktis* ANT/505 cell extracts in the presence of pectin (Truong *et al*., 2001), which indicated a specific expression of this secreted protein. This finding, their multi-modular character and the biochemical function suggested that these enzymes operate outside of the cell. Moreover, a plate-based screening assay led to the identification of *P. haloplanktis* ANT/505 as a pectin degrader by revealing distinct clearing zones on agar plates containing pectin (Truong *et al*., 2001), which indicated a secretion of pectate lyases. We therefore analysed subproteomes of *P. haloplanktis* ANT/505 cultivated with and without pectin. We enriched the intracellular, membrane-bound and extracellular protein fraction, and subjected the respective proteins to tandem mass spectrometry (MS/MS). This analysis revealed the accumulation of PelA and PelB in the extracellular soluble secretome (Fig. 4) indicating that ANT/505 depolymerizes pectins with soluble pectate lyases. Relative quantification,

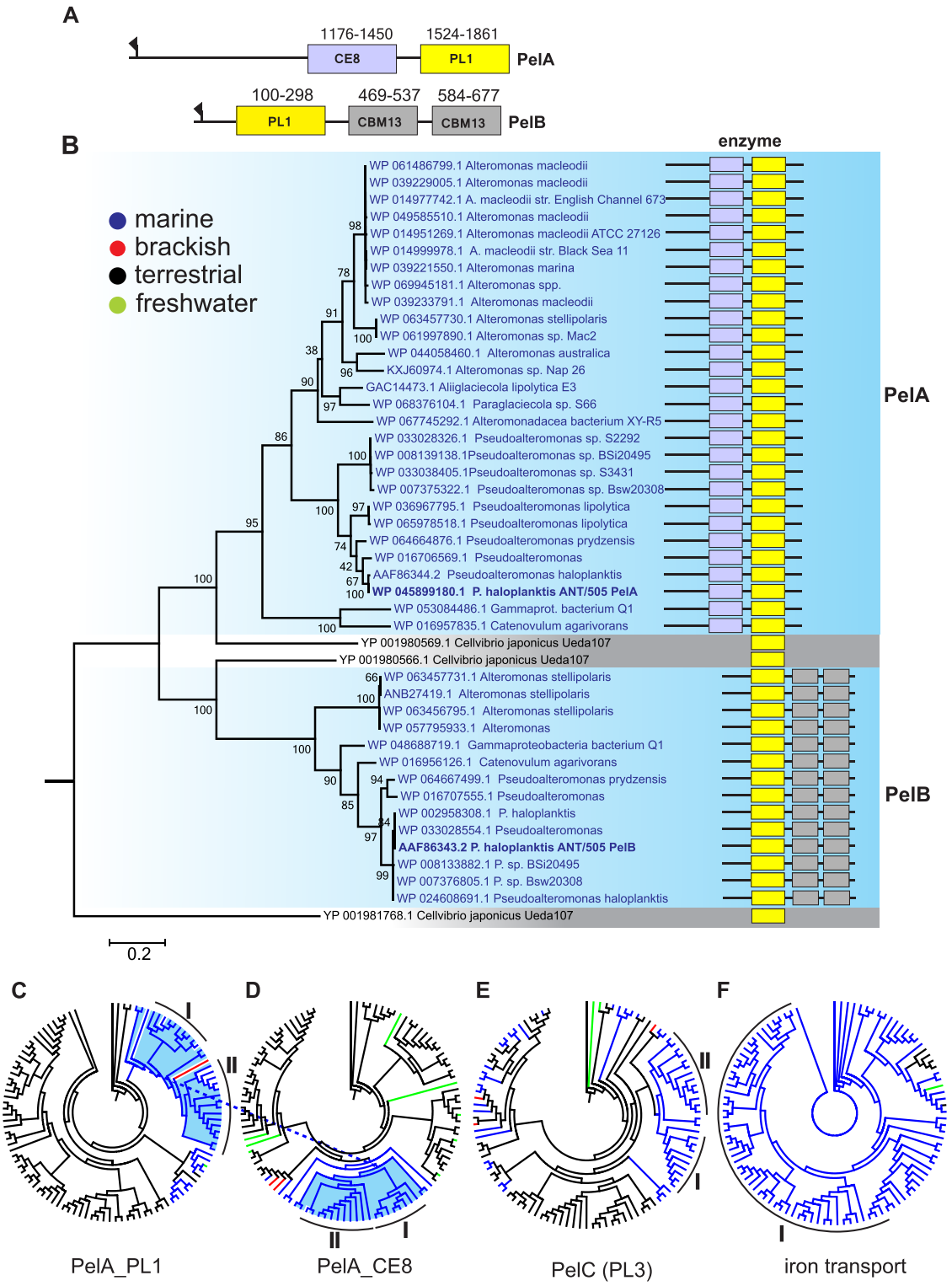

Fig. 3. Maximum likelihood phylogenies of the multi-modular pectate lyases and associated proteins located within and outside of the genomic island for pectin catabolism.

A. *Pseudoalteromonas haloplanktis* ANT/505 contains two multi-modular pectate lyases, PelA and PelB. PelA consists of a fused methylesterase domain of family CE8 and a pectate lyase catalytic domain of the PL1 family. PelB consists of a PL1 domain and two putative carbohydrate-binding modules (CBMs) of family 13.

B. Maximum likelihood phylogenetic tree of *P. haloplanktis* ANT/505 PelA (PH505_ap00520) and PelB (PH505_ad00690) and homologues with >90% query coverage obtained from GenBank. All sequences highlighted by a blue background are from marine bacteria, terrestrial

homologues have a grey background. No multi-modular homologues from terrestrial bacteria could be recovered at the time of this study. Three mono-modular, homologous sequences from the terrestrial Gammaproteobacterium *Cellvibrio japonicus* Ueda107 were included for comparison and one of these sequences, YP_001981768.1, was used as outgroup. The legend of the colour code indicates from which environment the corresponding bacteria were isolated (blue = marine, red = brackish, black = terrestrial including gut, green = freshwater).
C. Phylogenetic tree of best BLAST hits, shows homologues of PelA_PL1 exist in marine and terrestrial bacteria and in other habitats.
D. Phylogenetic analysis of the CE8 domain of PelA. The blue sequences underlined with a blue box and connected with the dotted line belong to one fused protein.
E. Phylogenetic analysis of PelC (PL3 domain, PH505_ap00550,). F. Phylogenetic analysis of the iron binding protein (PH505_ap00410). The sequences labelled with I belong to *Pseudoalteromonas* spp., the sequences labelled with II belong to *Alteromonas* spp. Shown are cladograms of rooted phylogenetic trees (C-E) and one unrooted tree (F). Fully annotated, high resolution versions of the phylogenetic trees with species names, bootstrap statistics and GI identifiers are given in the Supporting Information Figs. S3–S6. The colour coding shown in 3B is also relevant for the trees in 3C–F.

using normalized spectral abundance factors (NSAFs) (Fig. 4, Supporting Information Table S1) showed that the most abundant pectin degradation-related protein in the extracellular environment was PelB (%NSAFx1000 = 55), closely followed by PelA (%NSAFx1000 = 37). Both enzymes revealed a basal expression (%NSAFx1000 of PelB = 12.8 and PelA = 11.8) under control conditions without pectin. However, their abundance increased 3.2- and 4.4-fold, respectively, in the presence of pectin, confirming a substrate-specific response.

A series of additional catabolic proteins encoded by the genomic island were induced by pectin, such as KduD (2-deoxy-D-gluconate 3-dehydrogenase) and KduI (4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase). The latter was present in both conditions, but showed a 1.9-fold higher expression level in the presence of pectin. In *Escherichia coli*, galacturonate degradation is typically performed by uronate isomerase (UxaC), altronate oxidoreductase (UxaB) and altronate dehydratase (UxaA). The corresponding genes are also localized in the PelA island

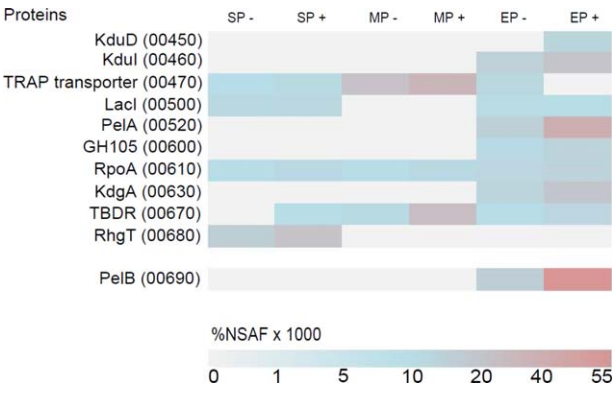

Fig. 4. *P. haloplanktis* ANT/505 reveals a substrate-dependent regulation of the pectin pathway and secretes PelA and PelB into the environment. A heat map for the relative abundances of detected proteins from the PelA island and PelB between all three different cellular protein fractions (SP = soluble proteome, MP = membrane proteome, EP = extracellular proteome) with pectin (+) and without pectin (–) in *P. haloplanktis* ANT/505 is shown. Scale color: red – highly abundant, blue – medium abundant, white – not detected. The data of this analysis are summarized in the Supporting Information Table S1. The mean values shown are based on three biological replicates.

of *P. haloplanktis* ANT/505, but none of them could be detected in our proteomic experiments. Another gene located in the genomic island, encoding the KDPG aldolase (2-dehydro-3-deoxy-phosphogluconate aldolase), was expressed at low levels in the absence of pectin but was detected with 2.4-fold higher concentration in the presence of pectin.

An important step in the bacterial enzymatic degradation of polysaccharides is the uptake of smaller oligosaccharides that are accessible for import through specific membrane proteins (Weiss *et al*., 1991). For a variety of glycans the transport is performed via TonB-dependent transport systems (Blanvillain *et al*., 2007). In our proteomic experiment the expression of one of the TonB-dependent receptors (PH505_ap00670) located in the genomic island was 6-fold higher in the membrane fraction in the presence of pectin, as compared to the pectin-free control.

The comparable expression patterns of *pelA* and *pelB* indicate a similar regulation. A “gapped local alignment” of the *pelA* and *pelB* upstream regions revealed a palindromic consensus sequence with 14 nucleotides (TGCCACC GGTGGCA), which could be relevant for the regulation of these genes. This sequence motif was also present in the upstream regions of four other genes of the *pelA* genomic island on chromosome II of *P. haloplanktis* ANT/505 (2-deoxy-D-gluconate-3-dehydrogenase (PH505_ap00450), GH77 (PH505_ap00640), hypothetical protein (PH505_ap00650), transcriptional regulator (GNTR) (PH505_ap00680)). Two of these proteins (PH505_ap00450 and PH505_ap00680) were upregulated under pectin conditions (Fig. 4, Supporting Information Table S1). The upstream sequence motif of these pectin-responsive genes does neither resemble the consensus of the KdgR repressor, which is one of the major regulators of many pectinolytic genes in other bacteria (James and Hugouvieux-Cotte-Pattat, 1996), nor is it similar to the consensus of the repressors PecS (Reverchon *et al*., 1994) and PecT (Surgey *et al*., 1996), which also play important roles in the regulation of terrestrial pectin catabolism. It is noteworthy in this context that one gene upstream of *pelA*, a LacI-type repressor gene is present (Perez-Rueda *et al*.,

1998). This repressor protein appears to be constitutively expressed, which is typical for this type of carbohydrate effector molecule that is usually defunctionalized by the availability of the substrate (Matthews and Nichols, 1998). The identified palindromic consensus sequence, mentioned above, resembled the core nucleotide sequence of the suggested cognate binding sites of LacI-type repressors (C amas *et al*., 2010). Therefore, a LacI-type repressor-like protein could be involved in the regulation of the pectin-responsive genes of *P. haloplanktis* ANT/505.

Discussion

Our results suggest that *P. haloplanktis* ANT/505 metabolizes pectin with a genomic island obtained by horizontal gene transfer. Different lines of evidence support this hypothesis. The island is absent in most of the closely related *Pseudoalteromonas* spp.; of currently 128 sequenced strains at NCBI ~10 contain orthologues of the CAZymes present in the island. The island is co-localized with genes, whose products have been associated with genetic mobilization of DNA. Rooted maximum likelihood phylogenetic trees of the PL1 domain and the CE8 domain of PelA as well as the PL3 (PelC) protein, all of which are encoded within the island, show that they form clades of sequences from marine bacteria nested within sequences from terrestrial microbes. While more dynamic exchange of PL3 domains is visible, these clades are clearly nested within sequences from terrestrial bacteria belonging to different environments (gut, soil, freshwater). The phylogenetic tree of the iron import protein, the gene of which is located just outside of the island, shows a very different evolutionary history. This tree is mainly populated with sequences from marine microbes, which is in agreement with a topology derived from speciation. The different topologies of genes located within and outside of the island are indicative of different evolutionary histories. In aggregate, these different results strongly indicate that the island was acquired by HGT.

Our genomic and proteomic analyses showed that *P. haloplanktis* ANT/505 has a functional and complete pathway to catabolize pectin (Fig. 5). It is interesting to note that the PelA island for pectin utilization of *P. haloplanktis* is reminiscent of bacteroidetal polysaccharide utilization loci (PULs; (Sonnenburg *et al*., 2010)). Beside a highly expressed TBDR gene, the genomic island encodes TRAP transport proteins and several glycoside hydrolases, which also reveal a pectin-responsive expression.

PelA and PelB were suggested to initiate degradation of pectin into oligosaccharides, which may be cleaved by the GH105 family protein of this island into unsaturated galacturonate (Collen *et al*., 2014). Galacturonate degradation can be performed by the uronate isomerase (UxaC) and altronate oxidoreductase (UxaB) of this gene cluster (Fig. 5). It is interesting to note that the island encodes a 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase (KduI) and a 2-deoxy-D-gluconate 3-dehydrogenase (KduD), which have recently been proposed as two enzymes of an alternative pathway for galacturonate degradation in *E. coli* (Rothe *et al*., 2013), and which can process the unsaturated galacturonate that is produced by GH105 via the *kdgF* (PH505_dn00060) encoded enzyme (Hobbs *et al*., 2016). Another pectin-responsive protein of the genomic island is the KDPG aldolase (2-dehydro-3-deoxy-phosphogluconate aldolase), which catalyzes a late step in the biodegradation of pectin by converting 2-dehydro-3-deoxy-D-gluconate 6-phosphate (KDPG) into pyruvate and D-glyceraldehyde 3-phosphate.

Our semi-quantitative proteomic analysis revealed a basal expression of the extracellular enzymes PelA and PelB in the absence of pectin. In the presence of pectin as a substrate, *P. haloplanktis* ANT/505 reveals a pectin-dependent expression of proteins of the pectin pathway and starts broadcasting the PelA and PelB enzymes. Basal expression and secretion of CAZymes can initiate the production of oligosaccharides that provide the signal to induce transcription of the rest of the pathway and can thereby enable a substrate-dependent rapid cellular response. This constitutive expression of "sentry CAZymes", in the absence of a substrate has previously been described in *Bacteroidetes* from marine habitats (Thomas *et al*., 2012) and from the human gut (Hehemann *et al*., 2012).

Our data revealed that PelA and PelB gained functions that can increase enzyme substrate interactions and reduce loss by diffusion in seawater. In PelA the new CE8 domain prunes the methyl-esterified groups, then the PL1 domain can cleave the resulting non-decorated pectate polysaccharide into oligosaccharides. The associated CBMs of the PelB enzyme can bind to and concentrate on the particulate gel-forming pectin substrate without being lost into the liquid, mobile phase. That CBMs provide this type of binding ability and increase glycan degradation in the aquatic environment has been repeatedly shown, also for CBM13 domains (Boraston *et al*., 2004). Moreover, these domains have a trefoil fold with three sugar binding sites per CBM domain (Fujimoto *et al*., 2002; Jiang *et al*., 2012). For PelB with two CBM13 this would result in six glycan binding sites with the potential of avidity effects, i.e. increased binding affinity.

In addition to these gained functions, we also identified loss of functions in the pathway. Two short putative proteins PH505_ap00580 and PH505_ap00590 of 78 (8 kDa) and 119 amino acids (13 kDa) are present in the genomic island. These proteins are related to GH43 enzymes with beta-D-xylosidase or alpha-L-arabinofuronidase activity. GH43 enzymes are five-bladed-beta-propeller of ~ 40 kDa molecular weight (Nurizzo *et al*., 2002), while the two

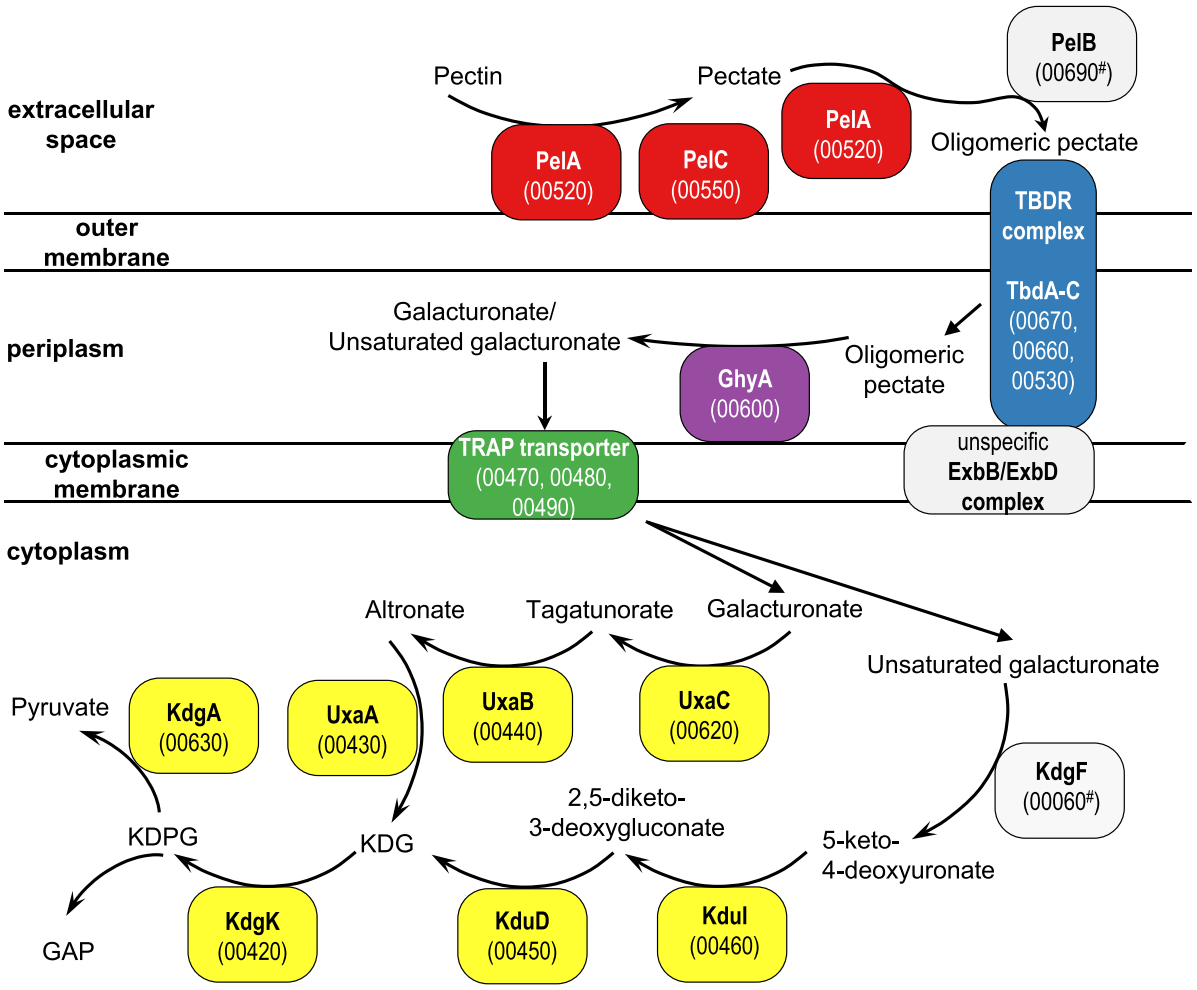

Fig. 5. Proposed pathway for the decomposition of pectins in *P. haloplanktis* ANT/505. As revealed by our proteome analyses, all genes were expressed, except for *uxaA-C*. Locus tags are indicated by the numbers given in parentheses below the protein names. Enzymatic functions of the putative pectin degradation pathway that are not encoded in the PelA genomic island are colored in light grey. PelA = pectate lyase/methylesterase, PelB = pectate lyase, TBDR = TonB-dependent receptor, GH105 = glycosyl hydrolase 105, RhgT = rhamnogalacturonan acetylesterase, KduD = 2-deoxy-D-gluconate 3-dehydrogenase, KduI = 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase, UxaA = altronate dehydratase, UxaB = altronate oxidoreductase, UxaC = uronate isomerase, KdgA = 2-dehydro-3-deoxy-phosphogluconate aldolase, KdgK = 2-dehydro-3-deoxy-glucono-kinase, KDPG = 2-keto-3-deoxy-phospho-gluconate, KDG = 2-keto-3-deoxy-gluconate. #The PelB (PH505_ad00690) and KdgF (PH505_ad00690) encoding genes are not located in the PelA encoding genomic island.

genes in the pectin-responsive island of *P. haloplanktis* ANT/505 are considerably smaller and therefore not of sufficient size to build the entire fold of a functional GH43. To exclude the possibility of sequencing artefacts we amplified and sequenced the corresponding regions. They were indeed truncated and this confirmed that the two genes are non-functional and have become cryptic genes. Arabinoxylans are abundant in cell walls of terrestrial plants where they form a network with pectins (Mohnen, 2008). The truncation of these GH43 like genes may imply they are not under positive selection possibly because arabinoxylans do not co-occur with pectin or pectin-like substrates in the marine environment, which would render the GH43 obsolete. Notably, these short gene fragments are absent in the related ATCC 14393 strain indicating that these homologous genomic islands are at different stages of adaptation to the marine environment.

Pectin-like polysaccharides have been identified in red and green algae, in seagrasses but also in marine microalgae (for Review (Popper *et al.*, 2011)). Other sources of pectin in the oceans are trees or other plants that arrive through rivers or by coastal runoff in the sea. While it remains unclear whether the *Pseudoalteromonadales* adapted to consume pectins from terrestrial plants, marine seagrasses or algae, PelA-positive cells were isolated from distant marine provinces (Baumann *et al.*, 1972; Truong *et al.*, 2001; Lopez-Perez *et al.*, 2012). This distribution may indicate a global ecological role of the pectin pathway

or that pectin-like substrates are widely dispersed carbon and energy sources in marine systems.

In conclusion, we have detected a pathway from terrestrial microbes in marine Gammaproteobacteria that is in the process of adapting to aquatic life by multi-modular fusions of selected key components. The suggested new principle for pectin degradation in a marine bacterium highlights the role of abiotic constraints for the function of transferred genes and pathways in marine habitats. The detection of a complete pathway for pectin degradation in *P. haloplanktis*, as reported in our study, indicates a new substrate niche for marine bacteria in the ocean.

Experimental procedures

Bacterial strains and cultivation conditions

P. haloplanktis ANT/505 was isolated from sea ice-covered surface water in the Southern ocean of the Antarctic and cultivated at 16°C on a modified Zobell medium as previously described (Truong *et al.*, 2001). For subsequent proteome and mRNA analyses but also for the purification of PelA and genomic DNA the cells were precultured for 24 h in 5 mL Zobell medium. Subsequently, 1.5 mL of the culture was transferred into 150 mL Zobell medium in a 500 mL shake flask containing 2 g/L citrus pectin (Sigma Aldrich) and the cells were cultivated for two days while shaking at 200 rpm. *Escherichia coli* DH5α cells were used for all cloning experiments. *E. coli* BL21 (DE3, pLysS) was used for the overexpression of the entire PelA enzyme and its separated domains. *E. coli* cells were cultivated under vigorous agitation at 37°C in Luria Bertani (LB) medium.

Whole genome shotgun sequencing

The genome of *P. haloplanktis* ANT/505 was sequenced by FLX pyrosequencing (454 Life Sciences/Roche). The GS De Novo Assembler (454 Life Sciences/Roche) enabled minimization of the genome fragmentation by ordering contiguous sequences into scaffolds. Automated detection of open reading frames and annotation of all draft genomes were performed by means of the RAST service pipeline (Overbeek *et al.*, 2014). This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank and can be found under the accession number ADOP01000000. To exclude the possibility of sequencing artefacts we amplified and resequenced the corresponding regions described in this paper by Sanger sequencing. The primers for these analyses are given in Supporting Information Table S3. The so updated version of the manually corrected sequences of the PelA genomic island can be found under the accession number ADOP01000016.2 and for *pelB* under the accession number ADOP01000004.1.

Proteome fractionation

Cellular protein extracts from *P. haloplanktis* cell cultures, grown either with or without pectin, were partitioned into: the intracellular soluble fraction, the membrane-bound fraction and the extracellular soluble fraction of the proteome. For purification of extracellular proteins, cells were centrifuged (6600 × *g* at 4°C for 10 min) and the resulting supernatant was precipitated at 4°C using 10% trichloroacetic acid, according to the protocol published by (Antelmann *et al.*, 2001)). The precipitated proteins were pelleted via centrifugation (40,000 × *g* at 4°C for 60 min), washed with ethanol and solubilized in a 8 M urea/2 M thiourea solution. To extract the membrane-bound and intracellular protein fractions, the cells were disrupted by sonication on ice (Bandelin sonopuls HD) followed by removal of cell debris by centrifugation (40,000 × *g* at 4°C for 20 min). Subsequent ultracentrifugation (100,000 × *g* at 4°C for 60 min) allowed for separation of intracellular soluble proteins (supernatant) and membranes and associated proteins (pellet). Purification of the membrane fraction was performed as described by (Eymann *et al.*, 2004), omitting the n-dodecyl-β-d-maltoside treatment. Accordingly, pelleted bacterial membranes were homogenized, washed for 45 min with a highly concentrated saline solution (20 mM Tris-HCl pH 7.5, 10 mM EDTA, 1 M NaCl) and ultracentrifuged (100,000 × *g* at 4°C for 60 min). Following this, the same steps were carried out using a carbonate buffer (100 mM Na_2CO_3-HCl pH 11, 10 mM EDTA, 100 mM NaCl). Finally, enriched membrane and intracellular proteins were solubilized in 50 mM triethylammonium bicarbonate buffer (Sigma Aldrich) supplemented with Complete Protease Inhibitor (Roche, Berlin, Germany).

1D-PAGE and protein measurements

From three biological replicates, 25 μg protein of each fraction was separated by 1D-PAGE (10% acrylamide, Protein ladder by Bio Rad: Precision Plus Protein™ Unstained Standards, Waltham, MA, USA), stained with Coomassie G-250 'blue silver' (Candiano *et al.*, 2004), eluted from the gel and digested using trypsin. Peptides were subsequently desalted by ZipTip columns (Millipore, Billerica, MA, USA) according to the manufacturer's guidelines. Peptide mixes were separated by reverse phase chromatography using a nano-Acquity UPLC System (Waters, Milford, MA, USA) and analysed by MS/MS in a LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). Spectra were assigned to the corresponding amino acid sequence of the *P. haloplanktis* ANT/505 database using Sorcerer-SEQUEST (SEQUEST version 2.7 revision 11, Thermo Scientific) including Scaffold Version 4.3.4. (Proteome Software Inc., Portland, OR, USA). A parent ion tolerance of 10 ppm and a fragment ion mass tolerance of 1.00 Da were used for SEQUEST searches. For protein identification, a stringent SEQUEST filter for peptides was used (Xcorr versus charge state: 1.80 for singly, 2.2 for doubly, and 3.3 for triply charged ions and deltaCn value greater than 0.10) and at least two unique peptides per proteins were required for identification. Protein fractions from both conditions were considered to be different if no peptides were detected in the alternative fraction or if a significant difference was shown by the Scaffold internal *t*-test analysis (threshold of 95%).

Pectin methylesterase enzyme assay

Pectin methylesterase activity was determined using a spectrophotometrical assay (Wojciechowski and Fall, 1996; Pilling

et al., 2000) with minor adaptations. The assay solution contained 300 μL of 0.4 M KH_2PO_4 (pH 7.0), 300 μL of 0.5% of citrus pectin (93% methylation, Sigma Aldrich), 30 μL of 375 mg/mL Floral-P (Sigma Aldrich) in distilled water, 2 μL of *Pichia pastoris* alcohol oxidase, 70 μL of enzyme solution with a final volume of 700 μL. The samples were incubated at 35°C for 1 h. After centrifugation at 9000 g for 2 min, the samples were measured spectrophotometrically at 405 nm.

Bioinformatic and phylogenetic analysis

All genes in the *pelA* gene cluster of *P. haloplanktis* ANT/505 (Fig. 1) were compared to the genomes of *P. haloplanktis* TAC125 (Médigue *et al.*, 2005) and *P. haloplanktis* ATCC 14393 (Gauthier *et al.*, 1995; Darling *et al.*, 2010) using Bl2seq (BLASTn, E-value 1e-5) (Camacho *et al.*, 2009). Sequence data and the BLAST comparison files were drawn with the R package genoPlotR (Guy *et al.*, 2010) version 0.8.4 and edited in Inkscape version 0.91. BLAST results were automatically edited, so that short hits contained in longer hits, and hits with a bitscore below 100 were removed. The GC content was calculated using a non-overlapping sliding window of size 1000 bp using the R package seqinr (Charif and Lobry, 2007) version 3.3-3. For cases where the sequence length was not a multiple of 1000, the last window was calculated from at least 500 bp (window size/2), i.e. merging windows $<$ 500 bp with the previous window.

The maximum likelihood phylogenetic tree containing sequences of PL1 from PelA (EGI73601.2) and PelB (EGI74776.1) was calculated with PL1 domains that are part of full-length homologues of PelA and PelB ($>$90% query coverage) extracted from GenBank. All alignments are based on amino acid sequences. Three sequences of *Cellvibrio japonicus* UEDA were added to the dataset. The alignment with MUSCLE (Edgar, 2004) was manually trimmed to contain only the PL1 domain by using the biochemically and structurally characterized PL1 sequence of *Xanthomonas campestris* (PDBid 2Qy1) as guide structure. The phylogeny was calculated in MEGA6 (Tamura *et al.*, 2013) with the LG replacement matrix with maximum likelihood estimated base frequencies for the amino acid partitions, 5 discrete gamma categories and a BioNJ starting tree, gaps were partially deleted with a 95% coverage cutoff. We performed 100 bootstrap iterations followed by a search for the best-scoring maximum likelihood tree.

Four proteins of *Pseudoalteromonas haloplanktis* ANT/505 were used to calculate maximum likelihood phylogenetic trees with PhyML (Guindon *et al.*, 2010). We used the PL1 and CE8 domains of PelA (EGI73601.2) and PelB (EGI74776.1) and PL3 (PelC, EGI73604.2) and the iron siderophore-binding protein (EGI73590.1) to recruit homologues by BLAST against GenBank at NCBI. When identical sequences belonged to multiple species only one of these was included in the phylogenetic analysis. PelA (EGI73601.2) of *P. haloplanktis* ANT/505 is a large multidomain protein. To identify domain boundaries, the multimodular proteins were truncated into individual catalytic domains with a guide structure of a functionally characterized protein with known 3D coordinates from the PDB database (http://www.rcsb.org/). Non-homologous parts were omitted from the phylogenetic analysis. For this the complete PelA sequence of *P. haloplanktis* ANT/505 (EGI73601.2) was analysed by BLAST by using the protein database (www.pdb.org) (Altschul *et al.*, 1990). The pectate lyase of family PL1 from *Xanthomonas campestris* (PDBid 2Qy1) (Xiao *et al.*, 2008) aligned with 44% sequence identity to residues 1603-1861 of PelA and was used to define the PL1 boundaries. PDB 1XG2 from *Lycopersicon esculentum* aligned with 27% sequence identity with residues 1183-1363 of PelA and was used to define the CE8 boundaries. Using the truncated domains as query, homologous protein sequences were retrieved by BLASTp searches against the non-redundant GenBank protein database from the National Center for Biotechnology Information (NCBI) website. The 100 best hits were recovered, excluding sequences below 90% coverage with the functional domains and were used for subsequent analysis. Sequences were aligned with the MUSCLE algorithm (Edgar, 2004) with default parameters within Mega6. Poorly aligned regions were automatically removed with the Gblocks algorithm using low stringency parameters (Castresana, 2000). Maximum likelihood phylogenetic reconstruction of proteins was performed using PhyML (Katoh and Toh, 2008), with 500 bootstrap replicates. The Le and Gascuel (LG) substitution model (Le and Gascuel, 2008) was chosen based on model testing with ProtTest (Abascal *et al.*, 2005). Maximum likelihood phylogenetic reconstruction was performed with estimated portions of invariable sites, four substitution rate categories, estimated gamma shape parameter, estimated amino acid frequencies, and a BIONJ starting tree. For the CE8 domain we chose seven sequences of eukaryotic pectinesterases as outgroup (GI: 6174913, 6093734, 2507165, 6093738, 57014096, 332278229, 20455195). For the PL1 domain we chose the five sequences At4g13210, At4g13710, At4g22080, At4g22090, At4g24780 as outgroup. For the PL3 domain-protein PelC (EGI73604.2) we chose five sequences from *Aspergillus nidulans* as outgroup (GI: 40739376, 40745491, 40738891, 40744125, 40747919). The siderophore-binding protein tree was not rooted.

Acknowledgements

We gratefully acknowledge the mass spectrometric analyses of the PelA and PelB proteins by Dirk Albrecht. Jan-Hendrik Hehemann was supported by an HFSP fellowship LT000465/2012 and acknowledges current funding by the German Research Foundation (DFG) through Emmy Noether Grant HE 7217/1-1. The work of Thomas Schweder was supported by the DFG in the framework of the priority program (SPP 1158) "Antarctic Research with comparative investigations in Arctic ice areas" by a grant (SCHW 595/9-1). The work of Frank Unfried was financially supported by scholarships from the Institute of Marine Biotechnology e.V. and the Ph.D. graduate program of the International Max Planck Research School of Marine Microbiology (MarMic). We thank Wade Abbott for his constructive criticism of the manuscript.

References

Abascal, F., Zardoya, R., and Posada, D. (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21:** 2104–2105.

Abbott, D.W., and Boraston, A.B. (2008) Structural biology of pectin degradation by Enterobacteriaceae. *Microbiol Mol Biol Rev* **72:** 301–316.

Alm, E.J., Huang, K.H., Price, M.N., Koche, R.P., Keller, K., Dubchak, I.L., and Arkin, A.P. (2005) The MicrobesOnline Web site for comparative genomics. *Genome Res* **15:** 1015–1022.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215:** 403–410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389–3402.

Antelmann, H., Tjalsma, H., Voigt, B., Ohlmeier, S., Bron, S., van Dijl, J.M., and Hecker, M. (2001) A proteomic view on genome-based signal peptide predictions. *Genome Res* **11:** 1484–1502.

Baumann, L., Baumann, P., Mandel, M., and Allen, R.D. (1972) Taxonomy of aerobic marine eubacteria. *J Bacteriol* **110:** 402–429.

Bertrand, E.M., McCrow, J.P., Moustafa, A., Zheng, H., McQuaid, J.B., Delmont, T.O., *et al*. (2015) Phytoplankton-bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proc Natl Acad Sci U S A* **112:** 9938–9943.

Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denance, N., *et al*. (2007) Plant carbohydrate scavenging through tonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. *PLoS One* **2:** e224.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., and Davies, G.J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* **382:** 769–781.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* **10:** 421.

Camas, F.M., Alm, E.J., and Poyatos, J.F. (2010) Local gene regulation details a recognition code within the LacI transcriptional factor family. *PLoS Comput Biol* **6:** e1000989.

Candiano, G., Bruschi, M., Musante, L., Santucci, L., Ghiggeri, G.M., Carnemolla, B., *et al*. (2004) Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* **25:** 1327–1333.

Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* **17:** 540–552.

Charif, D., and Lobry, J.R. (2007) SeqinR 1.0-2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In *Structural Approaches to Sequence Evolution: Molecules, Networks, Populations*. Bastolla, U., Porto, M., Roman, H.E., and Vendruscolo, M. (eds). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 207–232.

Collen, P.N., Jeudy, A., Sassi, J.F., Groisillier, A., Czjzek, M., Coutinho, P.M., and Helbert, W. (2014) A novel unsaturated beta-glucuronyl hydrolase involved in ulvan degradation unveils the versatility of stereochemistry requirements in family GH105. *J Biol Chem* **289:** 6199–6211.

Darling, A.E., Mau, B., and Perna, N.T. (2010) progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* **5:** e11147.

DeBoy, R.T., Mongodin, E.F., Fouts, D.E., Tailford, L.E., Khouri, H., Emerson, J.B., *et al*. (2008) Insights into plant cell wall degradation from the genome sequence of the soil bacterium Cellvibrio japonicus. *J Bacteriol* **190:** 5455–5463.

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

Ekborg, N.A., Gonzalez, J.M., Howard, M.B., Taylor, L.E., Hutcheson, S.W., and Weiner, R.M. (2005) Saccharophagus degradans gen. nov., sp. nov., a versatile marine degrader of complex polysaccharides. *Int J Syst Evol Microbiol* **55:** 1545–1549.

Eymann, C., Dreisbach, A., Albrecht, D., Bernhardt, J., Becher, D., Gentner, S., *et al*. (2004) A comprehensive proteome map of growing Bacillus subtilis cells. *Proteomics* **4:** 2849–2876.

Fujimoto, Z., Kuno, A., Kaneko, S., Kobayashi, H., Kusakabe, I., and Mizuno, H. (2002) Crystal structures of the sugar complexes of Streptomyces olivaceoviridis E-86 xylanase: sugar binding structure of the family 13 carbohydrate binding module. *J Mol Biol* **316:** 65–78.

Gauthier, G., Gauthier, M., and Christen, R. (1995) Phylogenetic analysis of the genera Alteromonas, Shewanella, and Moritella using genes coding for small-subunit rRNA sequences and division of the genus Alteromonas into two genera, Alteromonas (emended) and Pseudoalteromonas gen. nov., and proposal of twelve new species combinations. *Int J Syst Evol Microbiol* **45:** 755–761.

Gogarten, J.P., and Townsend, J.P. (2005) Horizontal gene transfer, genome innovation and evolution. *Nat Rev Microbiol* **3:** 679–687.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59:** 307–321.

Guy, L., Kultima, J.R., and Andersson, S.G. (2010) genoPlotR: comparative gene and genome visualization in R. *Bioinformatics* **26:** 2334–2335.

Hehemann, J.H., Kelly, A.G., Pudlo, N.A., Martens, E.C., and Boraston, A.B. (2012) Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proc Natl Acad Sci U S A* **109:** 19786–19791.

Hehemann, J.H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., and Michel, G. (2010) Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* **464:** 908–912.

Hobbs, J.K., Lee, S.M., Robb, M., Hof, F., Barr, C., Abe, K.T., *et al*. (2016) KdgF, the missing link in the microbial metabolism of uronate sugars from pectin and alginate. *Proc Natl Acad Sci U S A* **113:** 6188–6193.

James, V., and Hugouvieux-Cotte-Pattat, N. (1996) Regulatory systems modulating the transcription of the pectinase genes of Erwinia chrysanthemi are conserved in Escherichia coli. *Microbiology* **142:** 2613–2619.

Jiang, D., Fan, J., Wang, X., Zhao, Y., Huang, B., Liu, J., and Zhang, X.C. (2012) Crystal structure of 1,3Gal43A, an exo-

beta-1,3-galactanase from Clostridium thermocellum. *J Struct Biol* **180:** 447–457.
Kabisch, A., Otto, A., Konig, S., Becher, D., Albrecht, D., Schuler, M., *et al*. (2014) Functional characterization of polysaccharide utilization loci in the marine Bacteroidetes 'Gramella forsetii' KT0803. *ISME J* **8:** 1492–1502.
Katoh, K., and Toh, H. (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform* **9:** 286–298.
Larsbrink, J., Rogers, T.E., Hemsworth, G.R., McKee, L.S., Tauzin, A.S., Spadiut, O., *et al*. (2014) A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. *Nature* **506:** 498–502.
Le, S.Q., and Gascuel, O. (2008) An improved general amino acid replacement matrix. *Mol Biol Evol* **25:** 1307–1320.
Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42:** D490−D495.
Lombard, V., Bernard, T., Rancurel, C., Brumer, H., Coutinho, P.M., and Henrissat, B. (2010) A hierarchical classification of polysaccharide lyases for glycogenomics. *Biochem J* **432:** 437–444.
Lopez-Perez, M., Gonzaga, A., Martin-Cuadrado, A.B., Onyshchenko, O., Ghavidel, A., Ghai, R., and Rodriguez-Valera, F. (2012) Genomes of surface isolates of Alteromonas macleodii: the life of a widespread marine opportunistic copiotroph. *Sci Rep* **2:** 696.
Matthews, K.S., and Nichols, J.C. (1998) Lactose repressor protein: functional properties and structure. *Prog Nucleic Acid Res Mol Biol* **58:** 127–164.
Médigue, C., Krin, E., Pascal, G., Barbe, V., Bernsel, A., Bertin, P.N., *et al*. (2005) Coping with cold: the genome of the versatile marine Antarctica bacterium Pseudoalteromonas haloplanktis TAC125. *Genome Res* **15:** 1325–1335.
Mohnen, D. (2008) Pectin structure and biosynthesis. *Curr Opin Plant Biol* **11:** 266–277.
Nurizzo, D., Turkenburg, J.P., Charnock, S.J., Roberts, S.M., Dodson, E.J., McKie, V.A., *et al*. (2002) Cellvibrio japonicus alpha-L-arabinanase 43A has a novel five-blade beta-propeller fold. *Nat Struct Biol* **9:** 665–668.
Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., *et al*. (2014) The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* **42:** D206–D214.
Perez-Rueda, E., Gralla, J.D., and Collado-Vides, J. (1998) Genomic position analyses and the transcription machinery. *J Mol Biol* **275:** 165–170.
Permina, E.A., Mironov, A.A., and Gelfand, M.S. (2002) Damage-repair error-prone polymerases of eubacteria: association with mobile genome elements. *Gene* **293:** 133–140.
Pilling, J., Willmitzer, L., and Fisahn, J. (2000) Expression of a Petunia inflata pectin methyl esterase in Solanum tuberosum L. enhances stem elongation and modifies cation distribution. *Planta* **210:** 391–399.
Popa, O., and Dagan, T. (2011) Trends and barriers to lateral gene transfer in prokaryotes. *Curr Opin Microbiol* **14:** 615–623.
Popper, Z.A., Michel, G., Herve, C., Domozych, D.S., Willats, W.G., Tuohy, M.G., *et al*. (2011) Evolution and diversity of plant cell walls: from algae to flowering plants. *Annu Rev Plant Biol* **62:** 567–590.
Reeves, A.R., D'elia, J.N., Frias, J., and Salyers, A.A. (1996) A Bacteroides thetaiotaomicron outer membrane protein that is essential for utilization of maltooligosaccharides and starch. *J Bacteriol* **178:** 823–830.
Reverchon, S., Nasser, W., and Robert-Baudouy, J. (1994) pecS: a locus controlling pectinase, cellulase and blue pigment production in Erwinia chrysanthemi. *Mol Microbiol* **11:** 1127–1139.
Richard, P., and Hilditch, S. (2009) D-galacturonic acid catabolism in microorganisms and its biotechnological relevance. *Appl Microbiol Biotechnol* **82:** 597–604.
Rothe, M., Alpert, C., Loh, G., and Blaut, M. (2013) Novel insights into E. coli's hexuronate metabolism: KduI facilitates the conversion of galacturonate and glucuronate under osmotic stress conditions. *PLoS One* **8:** e56906.
Smillie, C.S., Smith, M.B., Friedman, J., Cordero, O.X., David, L.A., and Alm, E.J. (2011) Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **480:** 241–244.
Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N., and Sonnenburg, J.L. (2010) Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell* **141:** 1241–1252.
Surgey, N., Robert-Baudouy, J., and Condemine, G. (1996) The Erwinia chrysanthemi pecT gene regulates pectinase gene expression. *J Bacteriol* **178:** 1593–1599.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30:** 2725–2729.
Thomas, F., Barbeyron, T., Tonon, T., Genicot, S., Czjzek, M., and Michel, G. (2012) Characterization of the first alginolytic operons in a marine bacterium: from their emergence in marine Flavobacteriia to their independent transfers to marine Proteobacteria and human gut Bacteroides. *Environ Microbiol* **14:** 2379–2394.
Truong, L.V., Tuyen, H., Helmke, E., Binh, L.T., and Schweder, T. (2001) Cloning of two pectate lyase genes from the marine Antarctic bacterium Pseudoalteromonas haloplanktis strain ANT/505 and characterization of the enzymes. *Extremophiles* **5:** 35–44.
Wargacki, A.J., Leonard, E., Win, M.N., Regitsky, D.D., Santos, C.N., Kim, P.B., *et al*. (2012) An engineered microbial platform for direct biofuel production from brown macroalgae. *Science* **335:** 308–313.
Weiss, M.S., Abele, U., Weckesser, J., Welte, W., Schiltz, E., and Schulz, G.E. (1991) Molecular architecture and electrostatic properties of a bacterial porin. *Science* **254:** 1627–1630.
Wojciechowski, C.L., and Fall, R. (1996) A continuous fluorometric assay for pectin methylesterase. *Anal Biochem* **237:** 103–108.
Xiao, Z., Bergeron, H., Grosse, S., Beauchemin, M., Garron, M.L., Shaya, D., *et al*. (2008) Improvement of the thermostability and activity of a pectate lyase by single amino acid substitutions, using a strategy based on melting-temperature-guided sequence alignment. *Appl Environ Microbiol* **74:** 1183–1189.
Yu, N.Y., Wagner, J.R., Laird, M.R., Melli, G., Rey, S., Lo, R., *et al*. (2010) PSORTb 3.0: improved protein subcellular

localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26:** 1608–1615.

Supporting information

Additional Supporting Information can be found in the online version of this article at the publisher's web-site:

Table S1. Summary of the proteome data. Listed are all identified proteins in the respective enriched proteome fractions and the calculated NSAFs.
Table S2. GC content of the genes of the PelA containing genomic island. The average GC-content of the complete genome of ANT/505 is 39.72%.
Table S3. Primers used for the resequencing of the corresponding genes of the PelA region to exclude the possibility of sequencing artefacts.
Table S4. List of primers used for cloning experiments and transcriptional analyses. *The T7 promoter sequence is underlined.
Fig. S1. Growth of *Pseudoaltermonas haloplanktis* on pectin. The growth of the strains ANT/505 and TAC125 was tested in a marine mineral salt medium (Kabisch *et al*., 2014) with 0.2% citrus pectin as the only carbon and energy source. The cultivations of each strain were performed in triplicates (n=3), standard deviations are indicated by error bars.
Fig. S2. SDS-PAGE of the inclusion body (IB) protein fraction of *E. coli* BL21(DE3) (pLysS) (pETpme) after overexpression of the recombinant gene fragment of the Pme domain of PelA. Lane 1 and 2: with pETpme; Lane 3 and 4: control with the empty plasmid pET20b(+) without *pme* gene. Odd numbers show the IB protein fraction after cultivation at 20°C and even numbers at 30°C. The Pme protein with a calculated molecular weight of about 50 kD is indicated with an arrow.
Fig. S3. SDS-PAGE of the cytoplasmic protein fraction of *E. coli* BL21(DE3) (pLysS) after overproduction of PelA with the vector pRSETpelA (lane 2). Control: pRSET plasmid without *pelA* (lane 1); M: protein standard; The PelA protein band at about 220 kD is indicated by an arrow.
Fig. S4. Maximum likelihood phylogenetic tree of the PL1 domain of PelA (PH505_ap00520) from ANT/505. The alignment was prepared with MUSCLE (Edgar, 2004) and the tree with PhyML (Guindon *et al*., 2010). The protocol is described with additional details in the Material and Methods section. Numbers given on the branches are bootstrap proportions as a percentage of 500 replicates for values ≥10%. The numbers before the species names are gene identifier (GI number) except for the one sequence from *Xanthomonas campestris* for which the pdb-id is given. The scale bar represents the number of amino acid substitutions per site. The colors indicate from which environment the host microbes were isolated (blue = marine, red = brackish, black = terrestrial including gut, green = freshwater).
Fig. S5. Maximum likelihood phylogenetic tree of the CE8 domain of PelA (PH505_ap00520). The alignment was prepared with MUSCLE (Edgar, 2004) and the tree with PhyML (Guindon *et al*., 2010). The protocol is described with more details in the Material and Methods section. Numbers given on the branches are bootstrap proportions as a percentage of 500 replicates for values ≥10%. The numbers before the species names are gene identifier (GI number). The scale bar represents the number of amino acid substitutions per site. The colors indicate from which environment the host microbes were isolated (blue = marine, red = brackish, black = terrestrial including gut, green = freshwater).
Fig. S6. Maximum likelihood phylogenetic tree of the PelC enzyme (PH505_ap00550), which contains only a PL3 domain. The alignment was prepared with MUSCLE (Edgar, 2004) and the tree with PhyML (Guindon *et al*., 2010). The protocol is described with more detail in the Material and Methods section. Numbers given on the branches are bootstrap proportions as a percentage of 500 replicates for values ≥10%. The numbers before the species names are gene identifier (GI number). The scale bar represents the number of amino acid substitutions per site. The colors indicate from which environment the host microbes were isolated (blue = marine, red = brackish, black = terrestrial including gut, green = freshwater).
Fig. S7. Unrooted maximum likelihood phylogenetic tree of the iron import protein (PH505_ap00410). The alignment was prepared with MUSCLE (Edgar, 2004) and the tree with PhyML (Guindon *et al*., 2010). The protocol is described with additional details in the Material and Methods section. Numbers given on the branches are bootstrap proportions as a percentage of 500 replicates for values ≥10%. The numbers before the species names are gene identifier (GI number). The scale bar represents the number of amino acid substitutions per site. The colors indicate from which environment the host microbes were isolated (blue = marine, black = terrestrial, green = freshwater).
Fig. S8. Determination of the transcriptional start site of *pelA*.
A. Primer extension analysis for *pelA*. Total RNA was isolated from ANT/505 at different times of growth corresponding to an optical density at 600 nm of 1.6 (lane 1); 1.8 (lane 2) and 2.1 (lane 3). The transcriptional start sites are indicated by arrows.
B. Nucleotide and deduced N-terminal amino acid sequences of the *pelA* gene of *P. haloplanktis* strain ANT/505. The mRNA starting point is indicated by "+1". The potential promoter region at positions −35 and −10 and the ribosomal binding site are underlined.
Fig. S9. Determination of the transcriptional start site of *pelB*.
A. Primer extension analysis for *pelB*. Total RNA was isolated from ANT/505 at different times of growth corresponding to an optical density at 600 nm of 1.6 (lane 1); 1.8 (lane 2) and 2.1 (lane 3). The transcriptional start sites are indicated by arrows.
B. Nucleotide and deduced N-terminal amino acid sequences of the *pelB* gene of *P. haloplanktis* strain ANT/505. The mRNA starting point is indicated by "+1". The potential promoter region at positions −35 and −10 and the ribosomal binding site are underlined.