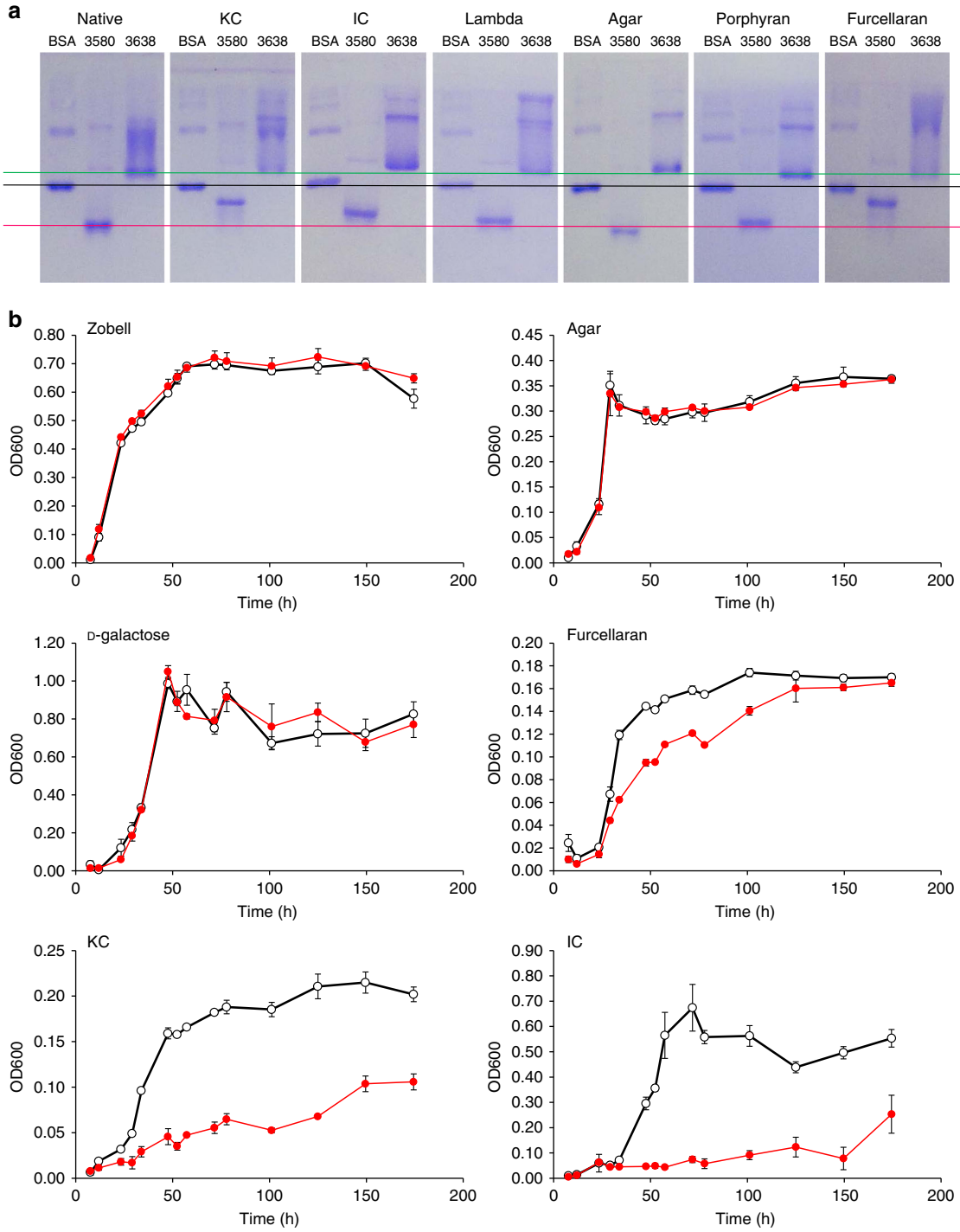

Fig. 8 Affinity gel electrophoresis for the SusD-like proteins and phenotyping for the *susCD*-like mutant Δ*cgtA-cgtB*. **a** BSA and recombinant ZGAL_3580 (CgtB, SusD-like) and ZGAL_3638 (SusD-like) migrated at 60 V over 5 h in native-PAGE gels or in gels containing different red algal cell wall polysaccharides. The delay of migration for ZGAL_3580 (red line) and ZGAL_3638 (green line) was compared between the native gels and the polysaccharide containing gels by comparison with BSA (black line). **b** Growth curves for wild-type *Z. galactanivorans* (black curves) and Δ*cgtA-cgtB* (red curves) in rich Zobell media and in minimal media containing different carbohydrate substrates

phosphorylative variant of the Entner-Doudoroff pathway to degrade 2-keto-3-deoxy-D-galactonate, as observed in the archaeon *Picrophilus torridus*[37]. The *susD*-like genes are unique to the *Bacteroidetes*[8] and transfer of the PUL to species belonging to other phyla has resulted in the loss of the *susCD*-like pair. This phenomenon is observed in the clade 3 Gammaproteobacteria (Fig. 9) and was previously identified in the case of alginolytic operons[19]. However, this does not mean that TBDT are absent in the gene clusters of other phyla. For instance, all clade 3 Gammaprotobacteria species have a TBDT gene in their cluster (CATDS2_v1220055 in *Catenovulum agarivorans* DS-2; JRKG01_v1_110122 in *Pseudoalteromonas* sp. PLSV; Patl_0887 in *P. atlantica* T6c; H978DRAFT_1909 in *Alteromonas* sp. ALT199). They are only distantly related to ZGAL_3581, but their location within the cluster strongly suggests that they play a similar role in the import of the carrageenan degradation product.

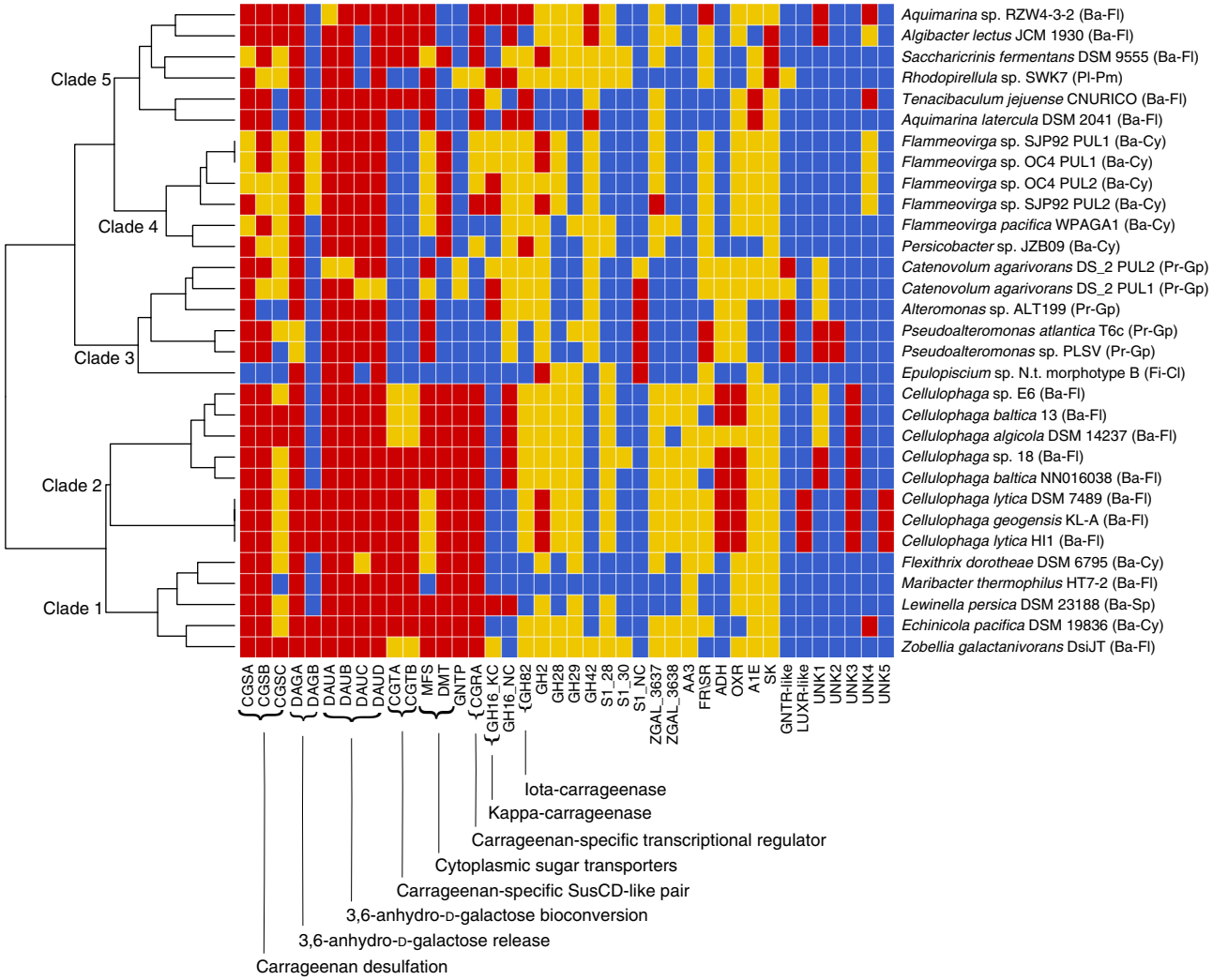

Fig. 9 Conservation of the carrageenolytic PUL organization among bacterial genomes. Putative homologs of selected carrageenan-induced genes from *Z. galactanivorans* were found in 29 other bacterial genomes with the conservation of these genes evaluated by a heatmap. Based on these conservation profiles, the bacterial species cluster into 5 main clades. Red coloring indicates the genes are conserved in a PUL, yellow indicates the genes are conserved elsewhere in the genome and blue indicates the genes are absent in the genome. Gene names for the *Z. galactanivorans* carrageenan-utilization locus are found in Table 1 and the legend of Fig. 1. Abbreviations: CGTA, CarraGeenan Transport A (SusC-like, TonB-dependent transporter); CGTB, CarraGeenan Transport B (SusD-like lipoprotein); GH, glycoside hydrolase; S1, family 1 sulfatase; AA3, auxiliary activity 3; FR/SR, fumarate reductase/succinate dehydrogenase; ADH, aldehyde dehydrogenase; OXR, isoquinoline 1-oxidoreductase alpha subunit; A1E, aldose 1-epimerase; SK, sugar kinase; GNTR-like and LUXR-like, transcription factor families GNTR and LUXR; UNK, unknown function. The phylum and class are indicated beside each organism in parentheses: Ba, *Bacteroidetes*; Pl, *Planctomycetes*; Pr, *Proteobacteria*; Fi, *Firmicutes*; Fl, *Flavobacteriia*; Pm, *Planctomycetia*; Cy, *Cytophagia*; Gp, *Gammaproteobacteria*; Cl, *Clostridia*; Sp, *Sphingobacteriia*

Such clusters containing polysaccharide-related TBDT genes have already been described in Gammaproteobacteria[38]. The presence of carrageenan-specific PULs in gut bacteria from marine herbivorous animals is also reminiscent of the horizontal acquisition of porphyran/agar-related genes in animal and human intestinal symbionts[39]. This is quite clear for the surgeon fish symbiont *Epulopiscium* sp. (Fig. 9, clade 3), which is also known to have horizontally acquired GH16 porphyranase and GH117 genes from marine flavobacteria[40]. Phylogenetic analysis of the GH127 family unraveled another likely HGT case (Supplementary Fig. 13, Supplementary Data 12). Whereas ZGAL_3148 (DagA2, GH127-2) and ZGAL_3150 (DagA3, GH127-3) cluster only with homologs from marine bacteria harboring the carrageenan-specific PUL, ZGAL_3147 (DagA1, GH127-1) is at the root of a clade including GH127 enzymes from human gut *Bacteroides* species. These *dagA*1-like genes are within PULs including GH genes at first sight unrelated to carrageenans (e.g. GH78, GH95, GH105; Supplementary Fig. 14). Thus, these *Bacteroides* GH127 genes have a marine origin, but have most likely evolved in specificity after their horizontal acquisition.

Conclusion. Here we have shown that PUL-like structures (lacking the *susCD*-like pair but maintaining other carrageenan-related genes) are found in bacterial phyla other than *Bacteroidetes*. Furthermore, the carrageenan utilization system is not static and can be characterized by gene losses and gene acquisitions with a dedicated core 3,6-anhydro-D-galactose metabolism that is conserved within carrageenolytic bacteria. This core system is essential but not sufficient for carrageenan utilization. Missing functions (e.g., carrageenases, some specific sulfatases) may be assumed by non-orthologous genes in different bacterial species. Therefore, polysaccharide utilization pathways are not always conferred by a single locus, even in *Bacteroidetes*, and may consist

of a complex regulon. Moreover, our work experimentally strengthens the recent proposition[41] that the PUL definition should not be restricted to the presence of a *susCD*-like pair and that this PUL notion should be extended to other bacterial phyla.

Methods

Materials. All materials were obtained from Sigma-Aldrich unless stated otherwise. Bacterial strains and plasmids used for phenotyping and complementation studies are listed in Supplementary Table 4.

Cloning of target genes. All the *Z. galactanivorans* genes were cloned and overexpressed as previously described[42]. Briefly, genes were PCR-amplified using the NEB Q5 High-Fidelity DNA Polymerase system (Supplementary Table 2). PCR reactions were done with 30 cycles (denaturation: 95 °C; annealing: 60 °C; elongation: 72 °C) using 0.5 units of enzyme in a total reaction of 50 µL using the primers shown in Supplementary Table 2. Amplicons were cleaned up using the QIAquick PCR Purification Kit (Qiagen) and digested with the appropriate restriction endonucleases. All ligations were done in the linearized T7 system vector pFO4 (a MCS-modified pET15b) except for ZGAL_3151 which was cloned into pET20b.

Protein production and purification. In general, *Escherichia coli* BL21(DE3) cells were transformed with the plasmids containing the gene fragment of interest then grown in the autoinduction Zyp-5052 medium[43] (200 µg mL^{-1} ampicillin, 20 °C, 72 h). ZGAL_3152 (SeMet) was similarly produced in PASM-5052 medium[43]. The sulfatases were produced in 1 L Luria-Bertani medium supplemented with 100 µg/mL ampicillin, until reaching an OD_{600} of about 0.7. Sulfatase gene expression was induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (overnight, 20 °C). In all cases, cells were collected by centrifugation at 3063×*g* for 30 min. After chemical cell lysis[44], the lysate was clarified at 13,865×*g* for 45 min at 4 °C. Using an ÄKTA FPLC, the supernatant was loaded onto a 5 mL GE His Trap HP column, washed with 20 mM Tris pH 8.0 and 100 mM NaCl and eluted with an increasing gradient of 1–100% 20 mM Tris pH 8.0, 100 mM NaCl, and 1 M imidazole. Fractions containing the protein of interest were pooled and concentrated (MWCO 5 kDa) and then loaded onto a Superdex S200 column in 20 mM Tris pH 8.0 and 100 mM NaCl. Fractions were again pooled and concentrated. The ExPASy ProtParam tool[45] was used to generate an extinction coefficient for calculation of protein concentration using the A280.

Marine polysaccharide and oligosaccharide substrates. Carrageenan polysaccharides were obtained commercially from CP-Kelco. The kappa/mu- and kappa-carrageenans were extracted from *Kappaphycus alvarezii*, the iota/nu- and iota-carrageenans from *Euchema denticulatum* and the lambda-carrageenan from tetrasporophytes of *Gigartina skottsbergii*. Furcellaran, a beta-carrageenan from *F. lumbricalis* consisting primarily of beta- and kappa-carrabiose motifs, was obtained as a generous gift from CP Kelco (Brian Rudolph). The exception is alpha-carrageenan which was produced using the native sulfatase from *Pseudoalteromonas atlantica* which is active on IC using a previously developed protocol[6]. The oligo-iota-carrageenans were incubated with pure sulfatase ZGAL_3145 to produce oligo-iota/alpha-carrageenans. One volume of oligo-iota-carrageenans (0.5% w/v in H_20 mQ) was incubated at 37 °C over 48 h with the same volume of the sulfatase at 1.0 mg mL^{-1} in 50 mM Tris pH 8.0, 200 mM NaCl, 1 mM $CaCl_2$. Oligosaccharides of furcellaran, kappa- and iota-carrageenans were produced using the recombinant kappa-carrageenase from *Pseudoalteromonas carrageenovora*[46] and iota-carrageenase from *Alteromonas fortis*[47]. Carrageenans (0.25%, 5 mL) were treated 48 h at 37 °C with kappa-carrageenase or iota-carrageenase (0.3 mg mL^{-1}, 50 µL) in 100 mM HEPES pH 7.5 and 25 mM NaCl. After checking the total hydrolysis, oligosaccharides mixtures were fractionated by size-exclusion chromatography (SEC). To this end, 5 mL of hydrolysate (concentrated by rotary evaporation at about 5% w/v) were filtered (0.2 µm, Millipore), and injected on 3 GE Healthcare Superdex 30 prep-grade columns (600 × 26 mm i.d.) mounted in series. The elution was conducted at a flow rate of 1 mL/min at 20 °C using 50 mM $(NH_4)_2CO_3$ as the eluent. Oligosaccharides were detected by differential refractometry (Spectra System RI-50, Thermo Separation products) and fractions of 5 mL were collected.

Sulfatase activity assays. *HPLC*. Carrageenan-sulfatase activity was measured by high-performance anion-exchange chromatography (HPAEC), according to a protocol adapted from Préchoux et al.[7]. Carrageenan solutions (0.5 % w/v in H_2O mQ) were incubated in presence of purified sulfatases (0.5 mg mL^{-1}) over 20 h at 37 °C in 25 mM Tris pH 8.0, 0.1 M NaCl, 0.5 mM $CaCl_2$. For each reaction, a control sample was realized in similar conditions but with inactivated enzyme (100 °C, 10 min). Reactions were filtered (10 kDa, Amicon Ultra, *Millipore*) then injected (AS-AP Autosampler) onto an AG11-HC guard column (4 × 50 mm) mounted in series with an AS11-HC anion-exchange column (4 × 250 mm) using an ICS5000 system (*Thermo Scientific Dionex*). Elutions for the detection of sulfate were performed with isocratic 12 mM NaOH at a flow rate of 1 mL min^{-1} (Single Pump-5), and the detection of anions was done by an Analytical CD Conductivity Detector associated to a suppressor (ASRS 500, 4 mm) running at 50 mA. Using a standard curve of sulfate and through integration of the peaks, the concentration of sulfate released by the enzymatic reaction was calculated from the difference in the amount of free sulfate (retention time at 4 min) between samples and their associated blanks. For oligosaccharide detection HPAEC analyzes were conducted on the same system described for sulfate quantification. Elutions were performed at a flow rate of 0.5 mL min^{-1} using a NaOH multistep gradient from 8 to 280 mM (40 min). Oligosaccharides were detected by conductivity mode under a current suppression of 50–300 mA.

^{1}H NMR. 800 µL of carrageenan (1% w/v in H_2O mQ) was incubated at 37 °C with the purified sulfatase (0.5 mg/ml), over several days. After checking for sufficient desulfation by HPAEC, incubation media were freeze-dried, exchanged twice then dissolved in D_2O (99.9%) to 10 mg mL^{-1}. Samples were transferred into 5-mm NMR tubes and ^{1}H-NMR spectra were recorded at 70 °C, on 500 MHz Bruker AVANCE III HD spectrometer, equipped with an indirect 5-mm gradient probehead TBI ^{1}H/(BB)/^{13}C. Chemical shifts were referenced with respect to trimethylsilylpropionic acid (TSP) used as an external standard, and expressed in ppm. For these experiments, the spectra required 64 scans.

Glycoside hydrolase activity assays. Small scale GH127 and GH129-like (ZGAL_3152) reactions were set up in two formats. Volume of 100 mL of Zyp5052 autoinduction medium was inoculated with *E. coli* BL21(DE3) transformed with the pFO4 vector alone as a control or with the GH clones and grown 72 h at 20 °C. Cells were collected from 2 mL of the cultures, resuspended in 2 mL 50 mM HEPES pH 7.5 and 150 mM NaCl and lysed using a French pressure cell. The lysate was clarified at 13,865×*g* for 45 min and the supernatant was used for the enzyme activity tests which were set up as follows in a 20 µL reaction volume: 0.5% furcellaran hexasaccharide, 150 mM NaCl, 3.75 mM $CaCl_2$, 100 mM HEPES pH 7.5 and 2 µL soluble cell lysate. Alternatively, instead of soluble cell lysate, 1 µL of purified enzyme (1 mg mL^{-1}) was added to the same reaction mixture described above. Reactions were left overnight at room temperature. For the GH2 enzymes, the sequential digest reactions were set up with 150 µL 0.25% furcellaran oligosaccharides that had been predigested with the purified GH129-like enzyme, as described above, and then heat denatured for 10 mins in a water bath at 90 °C. The 200 µL reactions were incubated overnight at 37 °C in the following conditions: 5 µg purified GH2, 1 mM $CaCl_2$, 1 mM $MgCl_2$ and 20 mM HEPES pH 7.5.

Fluorophore assisted carbohydrate electrophoresis. A speed vacuum was used to dry ~50 µg of oligo- or polysaccharide reaction volume to be destined for FACE[31]. Volume of 2 µL of 0.15 M ANTS (8-aminonaphthalene-1,3,6-trisulfonic acid) in a solution of acetic acid and water (3:17) was added to the dried reaction followed by 5 µL of freshly made 1 M sodium cyanoborohydride in DMSO. The samples were incubated overnight at 37 °C in the dark then resuspended in 20 µL of 20% glycerol. Between 2 and 10 µL of sample was loaded onto a 27% polyacrylamide gel and migrated at 200 V for 2 h in the dark at 4 °C. The gels were visualized under UV light.

3,6-Anhydro-D-galactose dehydrogenase activity assay. The enzymatic activity of ZGAL_3155 (DauA, 3,6-Anhydro-D-galactose dehydrogenase) was determined spectrophotometrically as a function of the reduction of NAD^+ or $NADP^+$ using a Spark 10 M microplate reader (Tecan, France). The reaction was performed in 25 mM Tris-HCl (pH 7.5) containing 100 mM NaCl and the reaction mixture (250 µL) contained 0.8–4.0 µg of pure recombinant dehydrogenase, 10 mM D-AnG (Dextra) and 1.5 mM NAD^+ or $NADP^+$. A negative control was performed by adding 4 µg pure ZGAL_3155 denatured for 10 min at 100 °C, all other conditions were the same within the reaction. An additional control using 10 mM D-galactose instead of D-AnG was also included. The absorbance at 340 nm was followed as a function of time until the absorbance reached a plateau. The activity of the enzyme is expressed as the number of µmoles of NADH produced min^{-1} mg^{-1} assuming the $\varepsilon_{340\ nm}$ NADH = 6220 M^{-1} cm^{-1}. Each reaction was performed in triplicate. At the end of the reaction, the reaction mixture was denatured for 10 min at 100 °C. The reaction mixture was then centrifuged for 10 min at 29,000×*g* and the supernatant was used as the substrate for the next enzyme, ZGAL_3156.

3,6-Anhydro-D-galactonate cycloisomerase activity assay. The enzymatic activity of ZGAL_3156 (DauB, 3,6-Anhydro-D-galactonate cycloisomerase) was determined using the 2-thiobarbituric assay (TBA)[33,48]. Volume of 250 µL of the reaction mixture from the preceding enzymatic step (the conversion of D-AnG into 3,6-anhydro-D-galactonate through the action of ZGAL_3155) was incubated with 10–40 µg of pure recombinant ZGAL_3156. The reaction (final volume 300 µL) was performed at room temperature in 25 mM MES pH 6.5 containing 100 mM NaCl (pH 7.5) over 20 min. To measure the activity, aliquots (25 µL) were withdrawn at the 20 min time point and the reaction was stopped with 1/10 volume of 12% trichloroacetic acid (TCA). Samples were then centrifuged for 10 min at 29,000×*g*. Volume of 25 µL of the supernatant was then incubated for 20 min at room temperature in the dark with 62.5 µL 25 mM periodic acid in 250 mM H_2SO_4 to oxidize 2-keto-3-deoxy-D-galactonate. Oxidation was terminated by the addition of 125 µL 2% (w/v) sodium arsenite in 500 mM HCl. 500 µL 0.3 % (w/v) TBA was then added and the reaction mixture was incubated for 10 min in a boiling water bath. After cooling down to room temperature, a sample of the solution was

removed. The color was intensified by the addition of an equal volume of dimethylsulfoxide and the absorbance was measured at 549 nm. To produce the substrate for the next reaction, the incubation with ZGAL_3156 was performed for 1 h with 50 µg of enzyme at room temperature. The enzyme was then inactivated for 10 min at 100 °C and centrifuged for 10 min at 29,000×*g* and 150 µL of supernatant was used as the substrate for ZGAL_3154.

2-keto-3-deoxy-D-galactonate kinase activity assay. The enzymatic activity of ZGAL_3154 (DauC, 2-keto-3-deoxy-D-galactonate kinase) was determined indirectly as a function of the oxidation of NADH using an NADH coupled assay. Reactions were performed in 25 mM MES buffer (pH 6.5). A typical reaction mixture (total volume 193 µL) contained 150 µL of reaction mixture from the preceding step, 2 µg pure recombinant ZGAL_3154, 0.97 mM adenosine-5-triphosphate (ATP), 9.7 mM $MgCl_2$ (Acros organics), 0.8 mM Phospho(enol) pyruvic acid tri(cyclohexylammonium) salt, 0.16 mM β-NADH (Applichem), 8 mM KCl, and 0.95 µL of mix of pyruvate kinase and lactic dehydrogenase enzymes from rabbit muscle (1.1 Units of pyruvate kinase and 0.8 Units of lactate dehydrogenase, respectively). The reactions were performed at room temperature in 96-wells UV-Star plates (Greiner) and the decrease of absorbance at 340 nm was read every 5 s with a Spark 10 M microplate reader (Tecan, France). The absorbance at 340 nm was followed as a function of time until it stabilized. A blank was performed using 2 µg pure ZGAL_3154 denatured for 10 min at 100 °C with all other reaction components the same.

2-keto-3-deoxy-D-galactonate aldolase activity assay. ZGAL_3153 (DauD, 2-keto-3-deoxy-D-galactonate aldolase) activity was verified using the TBA assay (as described above) to measure cleavage or synthesis of 2-keto-3-deoxy-6-phospho-D-galactonate (D-KDPGA). Degradation was measured using 15 µl of product of ZGAL_3154 incubated 10 min with 5 µg of ZGAL_3153. Second, synthesis of D-KDPGA from pyruvate and D-Glyceraldehyde-3-phosphate or D-glyceraldehyde was tested as follows: 30 µl of MES 25 mM pH 6.5, pyruvate 50 mM and D-glyceraldehyde-3-phosphate or D-glyceraldehyde 20 mM were incubated with 5 µg of ZGAL_3153 for 10 min. Both reactions were done in triplicate at room temperature and stopped by addition of 10% (v/v) of TCA 12%. The negative control was done with the same quantity of enzyme, previously boiled for 10 min.

Affinity gel electrophoresis. Interaction of the SusD-like proteins ZGAL_3580 and ZGAL_3638 with polysaccharides was tested by affinity gel electrophoresis. Kappa-carrageenan (from *Kappaphycus alvareizii*), iota-carrageenan (*Eucheuma denticulum*), lambda-carrageenan (Dupont batch 2321914837), agar (Sigma), porphyran (water extraction from *Porphyra* sp.) and furcellaran (CP Kelco) were incorporated at a final concentration of 0.21% in native 10% acrylamide gels. ZGAL_3580 (4 µg), ZGAL_3638 (8 µg) and bovine serum albumin (BSA, 6 µg) were loaded into the wells and migrated at 60 V over 5 h. Gels were stained with Coomassie Blue solution.

Protein crystallization. Initial hits were obtained using hanging drop vapor diffusion in the JCSG + screen (Qiagen) for ZGAL_3156 and the PACT screen (Qiagen) for ZGAL_3152. Optimized conditions were a 1:1 ratio of 0.14 M Na/K-tartrate, 12% PEG 3350 with 9 mg mL^{-1} ZGAL_3152 and 0.2 M tri potassium citrate, 20% PEG 3350 with 25 mg mL^{-1} ZGAL_3156.

X-ray crystallography data collection and processing. SeMet ZGAL_3152 diffraction data were collected at the ESRF on beamline ID-14-4 at the selenium peak with a wavelength of 0.97936 Å ($f' = -7.40$, $f'' = 6.38$) at 100 K. The cryoprotectant used was 15% MPD in mother liquor. The data were processed using MOSFLM[49,50], Pointless[51] was used to determine the spacegroup and the data were scaled using SCALA[52] within the CCP4 suite of programs[53]. PrepHAData was used to convert the mtz to SHELXS format and SHELX_CDE[54] was used to identify the Se subsites through SAD. Finally, CRANK[55] was used to construct the initial model of ZGAL_3152 through substructure refinement and model building. This model was then used to solve the structure of the native ZGAL_3152 by molecular replacement using MOLREP[56] (pdb id 5opq, Supplementary Table 1). The ZGAL_3152 structure went through an iterative process of refinement using REFMAC5[57] and model building using WinCoot[58]. Ramachandran statistics were 96.6% most favored, 3.2% additional allowed and 0.2% disallowed.

Native diffraction data for ZGAL_3156 was collected on the Proxima2 beamline at SOLEIL using an EIGER detector (Dectris), with a wavelength of 0.98 Å and an oscillation range of 0.1° at 100 K. The crystals were cryoprotected by short immersion in mother liquor supplemented with 15% Glycerol. The data were processed and scaled using XDS and XSCALE[59,60], Pointless was used to determine the spacegroup[51]. The structure (pdb id 5olc, Supplementary Table 1) was solved by molecular replacement using PHASER[61] and pdb id 4hpn using the biological assembly (octamer) as the primary model. The ZGAL_3156 structure went through an iterative process of refinement using BUSTER[62] and model building using WinCoot[58]. Ramachandran statistics were 94.4% most favored, 5.3% additional allowed and 0.3% disallowed.

MALDI-TOF-MS analysis. MALDI-TOF-MS spectra were acquired in negative ionization mode and reflector detection in the m/z range 500–1500 on an AutoflexSpeed TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany), equipped with a Smartbeam Laser (355 nm, 1000 Hz). Samples were solubilized in water (500 µg mL^{-1}). Volume of 1 µL of the samples were mixed with 1 µL of a DMA-DHB matrix solution prepared as described[63] directly on a polished steel MALDI target plate. Acquisition parameters (laser power, pulsed ion extraction, and so on) were optimized on the sample without any enzyme process. Spectra were recorded using FlexControl 3.4 and processed using FlexAnalysis 3.4 (Bruker Daltonics, Bremen, Germany).

Construction of dagA3, dagB, dauA, and cgrA deletion mutants. Single deletion mutants for *dagA*3 (ZGAl_3150, GH127-3), *dagB* (ZGAL_3152, GH129-like), *dauA* (ZGAL_3155, 3,6-Anhydro-D-galactose dehydrogenase), and *cgrA* (ZGAL_3159, AraC family regulator), and a double deletion mutant for *dagA*3/*dagB* were constructed following the previously described method[35], based on a *sacB* system (Supplementary Tables 3, 4). To delete *dagA*3, a 2.1-kbp fragment including the first 33 bp of *dagA*3 and 2,058 bp of upstream sequence was amplified using primers OFT0001 and OFT0003. The fragment was digested with PstI and SalI and ligated into pYT313 that had been digested with the same enzymes, to generate pFT2. A 2.1 kb fragment including the final 33 bp of *dagA*3 and 2,013 bp of downstream sequence was amplified using primers OFT0002 and OFT0004. The fragment was cloned into SalI and BamHI sites of pFT2 to generate the *dagA*3 deletion construct pFT5. To delete *dagB*, a 2.1-kbp fragment including the first 48 bp of ZGAL_3152 and 2047 bp of upstream sequence was amplified using primers ORL664 and ORL665. The fragment was digested with BamHI and SalI and ligated into pYT313 that had been digested with the same enzymes, to generate pRF8. A 2.2 kb fragment including the final 3 bp of *dagB* and 2172 bp of downstream sequence was amplified using primers ORL666 and ORL667. The fragment was cloned into SalI and SphI sites of pRF8 to generate the *dagB* deletion construct pRF9. To delete *dauA*, a 2.0-kbp fragment including the first 81 bp of *dauA* and 1887 bp of upstream sequence was amplified using primers ORL670 and ORL671. The fragment was digested with BamHI and SalI and ligated into pYT313 that had been digested with the same enzymes, to generate pRF10. A 2.2 kb fragment including the final 69 bp of *dauA* and 2116 bp of downstream sequence was amplified using primers ORL672 and ORL673. The fragment was cloned into SalI and PstI sites of pRF10 to generate the *dauA* deletion construct pRF11. To delete *cgrA*, a 2.1-kbp fragment including the first 5 bp of *cgrA* and 2030 bp of upstream sequence was amplified using primers ORL676 and ORL677. The fragment was digested with BamHI and SalI and ligated into pYT313 that had been digested with the same enzymes, to generate pRF12. A 2.2 kb fragment including the final 73 bp of *cgrA* and 2064 bp of downstream sequence was amplified using primers ORL678 and ORL679. The fragment was cloned into SalI and PstI sites of pRF12 to generate the *cgrA* deletion construct pRF13. Plasmids pFT5, pRF9, pRF11 and pRF13 were introduced individually into the wild-type *Z. galactanivorans* $Dsij^T$ by conjugation from *E. coli* S17_1 strains. Conjugants with plasmids integrated in the genome were isolated on *Cytophaga*-agar containing 50 µg mL^{-1} erythromycin. Single erythromycin-resistant colonies were grown overnight in *Cytophaga* medium in the absence of antibiotics at 30 °C. The cells where a second recombination event resulted in loss of the plasmid were selected on *Cytophaga*-agar containing 5% sucrose. Isolated colonies were checked for erythromycin sensitivity. Deletions were confirmed by PCR on isolated colonies using primer pairs OFT0005-OFT0006 to identify the *dagA*3 deletion mutant (mZG_0026), ORL668-ORL669 to identify the *dagB* deletion mutant (mZG_0007), ORL674-ORL675 to identify the *dauA* deletion mutant (mZG_0008) and ORL680-ORL681 to identify the *cgrA* deletion mutant (mZG_0011). A double deletion mutant for both *dagA*3 and *dagB* (mZG_0047) was constructed by introducing pRF9 into Δ*dagA*3 by conjugation, followed by sucrose selection and PCR confirmation as described above.

Construction of cgtA-cgtB double gene deletion mutant. To delete *cgtA-cgtB* [*zgal_3581* (*susC-like*)-*zgal_3580* (*susD-like*)], a 2.1-kbp fragment including the first 54 bp of *cgtA* and 2060 bp of upstream sequence was amplified using primers 2088 and 2089 (Supplementary Tables 3, 4). The fragment was digested with SacI and SpeI and ligated into pYT354 that had been digested with the same enzymes, to generate pYT381. A 2.2 kb fragment including the final 54 bp of *cgtB* and 2,169 bp of downstream sequence was amplified using primers 2090 and 2091. The fragment was cloned into SpeI and SphI sites of pYT381 to generate the *cgtA-cgtB* deletion construct pYT382. Primers 2102 and 2103 were used to identify the *cgtA-cgtB* deletion mutant (mZG_0054).

Complementation of dauA and cgrA deletion mutants. Deletion mutants Δ*dauA* and Δ*cgrA* were complemented by ectopic plasmid integration at a neutral site, as previously described[35] (Supplementary Tables 3 and 4). To complement Δ*dauA*, promoterless *dauA* was amplified with primers 1978 and 1979 and cloned into the XbaI and SphI sites of pYT356 to generate pYT360. Plasmid pYT360 was inserted into Δ*dauA* by conjugation to obtain the complemented strain Δ*dauA* + CP (complementation plasmid) (mZG_0039). To complement Δ*cgrA*, promoterless *cgrA* was amplified with primers 1980 and 1981 and cloned into the XbaI and SphI sites of pYT356 to generate pYT361. Plasmid pYT361 was inserted into Δ*cgrA*

by conjugation to obtain the complemented strain Δ*cgrA* + CP (mZG_0043). Control strains with the empty vector pYT356 inserted into the chromosome were generated in a similar way. In all cases, cells with integration of the plasmid at the neutral site were selected on *Cytophaga*-agar containing 50 µg mL^{-1} erythromycin and screened by PCR.

Liquid growth tests. *Z. galactanivorans* strains were routinely grown from glycerol stocks in Zobell medium 2216E (5 g L^{-1} tryptone, 1 g L^{-1} yeast extract, filtered seawater)[64]. Before use, all pre-cultures were collected by centrifugation 10 min at 2844×*g*, washed in 2 volumes of sterile saline solution, and resuspended in sterile saline solution to the same OD_{600}. Fifty microliters of bacterial suspension were inoculated into triplicate 40 mL flasks containing 5 mL of Zobell medium or marine mineral medium (MMM) composed for 1 L of 24.7 g NaCl, 6.3 g $MgSO_4{\cdot}7H_2O$, 4.6 g $MgCl_2{\cdot}H_2O$, 2 g NH_4Cl, 0.7 g KCl, 0.6 g $CaCl_2$, 200 mg $NaHCO_3$, 100 mg K_2HPO_4, 50 mg yeast extract, and 20 mg $FeSO_4{\cdot}7H_2O$[19] and supplemented with 4 g L^{-1} kappa-carrageenan, iota-carrageenan, agar, D-galactose or D-AnG. Growth at 20 °C under 160 r.p.m. was followed by measuring the absorbance at 600 nm in a Spark 10 microplate reader (Tecan) on 200 µL aliquots. Erythromycin (10 µg mL^{-1}) was added to all experiments with strains mZG_0031, mZG_0037, mZG_0039, mZG_0041, and mZG_0043 (Supplementary Table 4).

Solid growth tests. Degradation of carrageenans was tested on either ZoBell or MMM medium solidified with 10 g L^{-1} kappa-carrageenan or 20 g L^{-1} iota-carrageenan. Two microliters of cell suspension prepared as described above were spotted in the center of the Petri dish, and incubated at 30 °C. Degradation was evidenced by the formation of a hole for kappa-carrageenan or liquefaction for iota-carrageenan.

RNA-seq expression profiling. *Bacterial strain and culture conditions*. The type strain Dsij^T of *Z. galactanivorans* was routinely grown in Zobell medium 2216E (Difco) at 28 °C, 170 r.p.m. For transcriptome profiling, cells were cultivated in synthetic Marine Mineral Medium (MMM)[65]. MMM was supplemented with different carbon sources: D-galactose (Sigma-Aldrich #G0750), iota-carrageenan (Danisco, 2544-88-20), kappa-carrageenan (Coffoni ×6913) or D-AnG (Dextra Laboratories #G0002). Briefly, overnight cultures performed in Zobell medium were diluted at OD_{600} 0.05 in triplicate MMM containing 0.5% glucose as C source and incubated at 28 °C until reaching the stationary phase. These cultures were then used to inoculate 10 mL of MMM containing 0.4 g L^{-1} D-galactose, D-AnG, kappa- or iota-carrageenan. When cell density reached an OD_{600} of 0.7 (±0.1) bacteria were collected by centrifugation for 3 min at 4 °C after addition of ½ volume of frozen killing buffer (20 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 20 mM NaN_3) to the culture sample. The cell pellets were frozen in liquid nitrogen and stored at −80 °C until RNA extraction.

RNA extraction. The cell pellets were resuspended into 800 µL of lysis buffer (4 M guanidine thiocyanate, 25 mM sodium acetate pH 5.2, 5 g L^{-1} N-laurylsarcosinate), immediately mixed to 1 mL hot acid phenol (Sigma #P4682) and incubated during 5 min at 65 °C for efficient cell lysis. The aqueous phases were recovered after addition of 1 mL chloroform and centrifugation at 16,000×*g* during 10 min at room temperature. The samples were extracted at least four times with an equal volume of acid phenol:chloroform:IAA (25:24:1, (pH 4.5)) and once with chloroform. Total RNAs were ethanol precipitated at room temperature and the pellets resuspended in RNase-free water. The RNA concentration of samples was measured using a NanoDrop 1000 Spectrophotometer (NanoDrop Technologies, Inc.). The quality of RNA preparations was assessed by capillary electrophoresis using RNA Nano chips with a Bioanalyzer Agilent 2100 (Agilent Technologies, Palo Alto, USA). Total RNA extracts were treated using DNase I (Qiagen) to remove residual genomic DNA.

mRNA enrichment, RNA-seq library preparation and sequencing. mRNA enrichment was performed using the Illumina Ribo-Zero rRNA Removal Kit (Bacteria). The rRNA-depleted samples were purified using a RNeasy MinElute Cleanup Kit (Qiagen) and efficient rRNA depletion was confirmed by Bioanalyzer 2100. Strand-specific RNA-seq libraries were generated with the ScriptSeq V2 RNA-Seq Library Preparation Kit (Illumina), according to the manufacturer's instructions. Briefly, 2–4 µg of RNA samples were fragmented, di-tagged cDNAs were synthetized by successive random priming with terminal-tagging oligos and then purified with the Ampure bead XP system. Enrichment of purified cDNAs was done by no more than 14 PCR cycles. The sequencing was performed with a NextSeq 500 platform using a single-end 75 bases run. A total of 3.5–11 million passing filters reads were obtained per sample.

Read mapping and differential expression analysis. Sequencing reads were pre-processed for trimming of adapter sequences with cutadapt-1.9.1, then using PRINSEQ[66] for read quality with filtering (-min_len 20 -min_qual_mean 20 -min_qual_score 10) and trimming (trim_qual_right 10) options. Mapping was performed on *Z. galactanivorans* Dsij^T reference genome (retrieved from MicroScope "zobellia_gal_DsiJT_v2"; Refseq NC_015844.1) by Bowtie2 software with the very-sensitive option[67]. Alignments files were then converted to BAM files using SAMtools[68], log2-genome coverage files were computed as described[69] and these expression profiles can be visualized in the Artemis viewer[70]. The number of reads mapping to each predicted CDS was calculated for the 12 data sets by HTSeq-count[71] with the -m union option. Differential expression analysis (normalization and statistical tests) was performed using the SARTools with DESeq2 software[72] and a bonferroni *p*-value adjustment was used to correct for multiple testing. Genes with an adjusted *p*-value <0.05 were considered as differentially expressed. RNA-seq data are presented in Table 1, Supplementary Data 2.

PUL identification and comparison. Similar PULs were detected in 29 organisms using synteny results of the MicroScope platform[73] (Supplementary Data 10–11, Supplementary Fig. 12). The proteins from identified PULs and from *Z. galactanivorans* regulon were then manually grouped in homolog clusters based on sequence similarity. From these 40 clusters, presence/absence of homolog proteins encoded outside the PULs was determined by blastP alignments with at least 35% of amino acid identity and 80% coverage threshold. Results were then gathered in a matrix to indicate, for each organism, if a protein cluster homolog is encoded in the PUL (value 2), elsewhere on the genome (value 1) or absent (value 0). From this matrix, a heat map and a hierarchical classification of the organisms were made using heatmap.2 function and the ward.D2 algorithm with Manhattan distances from gplots and hclust R packages, respectively (Fig. 9).

GH127 phylogeny reconstruction. 549 proteins having a GH127 catalytic module, which aligned on more than 60% of the DBCAN domain, were extracted from the MicroScope platform. Protein sequences were aligned using MAFFT v7.307[74], then ambiguous and saturated regions were removed with BMGE v1.12 (with the gap rate parameter set to 0.5)[75]. The best fitting model of amino acid substitution for this data set was selected with ProtTest v3.4.2[76]. A Maximum-Likelihood phylogenetic tree was generated with the alignment using PhyML 3.1.0.2[77] using the LG amino acid substitution model with gamma-distributed rate variation (four categories), estimation of the proportion of invariable sites and exploring tree topologies. 100 bootstrap replicates were performed. The phylogenetic tree was displayed and annotated using the interactive tree of life (iTOL) online tool[78] (Supplementary Figs. 13, 14, Supplementary Data 12).

Data availability. The coordinates and structure factors for the proteins described above have been deposited in the Protein Data Bank (pdb id: 5opq and 5olc). The RNA-seq transcriptomic data have been deposited in the GEO database (GEO accession number: GSE101142). The sequence of the *Tenacibaculum jejuense* genome (used in the comparative genomic analysis) has been deposited at EMBL (accession: GCA_900198195). All other relevant data are available in this article and its Supplementary Information files, or from the corresponding author upon request.

Received: 16 March 2017 Accepted: 17 October 2017
Published online: 22 November 2017

References

1. Anderson, N. S., Dolan, T. C. & Rees, D. A. Evidence for a common structural pattern in the polysaccharide sulphates of the *Rhodophyceae*. *Nature* **205**, 1060–1062 (1965).
2. Martin, M., Portetelle, D., Michel, G. & Vandenbol, M. Microorganisms living on macroalgae: diversity, interactions, and biotechnological applications. *Appl. Microbiol. Biotechnol.* **98**, 2917–2935 (2014).
3. Michel, G. et al. The kappa-carrageenase of *P. carrageenovora* features a tunnel-shaped active site: a novel insight in the evolution of Clan-B glycoside hydrolases. *Structure* **9**, 513–525 (2001).
4. Michel, G. et al. The iota-carrageenase of *Alteromonas* fortis. A beta-helix fold-containing enzyme for the degradation of a highly polyanionic polysaccharide. *J. Biol. Chem.* **276**, 40202–40209 (2001).
5. Guibet, M. et al. Degradation of lambda-carrageenan by *Pseudoalteromonas* carrageenovora lambda-carrageenase: a new family of glycoside hydrolases unrelated to kappa- and iota-carrageenases. *Biochem. J.* **404**, 105–114 (2007).
6. Prechoux, A., Genicot, S., Rogniaux, H. & Helbert, W. Controlling carrageenan structure using a novel formylglycine-dependent sulfatase, an endo-4S-iota-carrageenan sulfatase. *Mar. Biotechnol.* **15**, 265–274 (2013).
7. Prechoux, A., Genicot, S., Rogniaux, H. & Helbert, W. Enzyme-assisted preparation of furcellaran-like kappa-/beta-carrageenan. *Mar. Biotechnol.* **18**, 133–143 (2016).
8. Thomas, F., Hehemann, J. H., Rebuffet, E., Czjzek, M. & Michel, G. Environmental and gut bacteroidetes: the food connection. *Front. Microbiol.* **2**, 93 (2011).
9. Teeling, H. et al. Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**, 608–611 (2012).
10. Barbeyron, T. et al. Habitat and taxon as driving forces of carbohydrate catabolism in marine heterotrophic bacteria: example of the model algae-associated bacterium *Zobellia galactanivorans* DsijT. *Environ. Microbiol.* **18**, 4610–4627 (2016).

11. Anderson, K. L. & Salyers, A. A. Genetic evidence that outer membrane binding of starch is required for starch utilization by *Bacteroides thetaiotaomicron*. *J. Bacteriol.* **171**, 3199–3204 (1989).
12. Anderson, K. L. & Salyers, A. A. Biochemical evidence that starch breakdown by *Bacteroides thetaiotaomicron* involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes. *J. Bacteriol.* **171**, 3192–3198 (1989).
13. Bjursell, M. K., Martens, E. C. & Gordon, J. I. Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period. *J. Biol. Chem.* **281**, 36269–36279 (2006).
14. Terrapon, N., Lombard, V., Gilbert, H. J. & Henrissat, B. Automatic prediction of polysaccharide utilization loci in *Bacteroidetes species*. *Bioinformatics* **31**, 647–655 (2015).
15. Sonnenburg, E. D. et al. Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell* **141**, 1241–1252 (2010).
16. Larsbrink, J. et al. A discrete genetic locus confers xyloglucan metabolism in select human gut *Bacteroidetes*. *Nature* **506**, 498–502 (2014).
17. Cuskin, F. et al. Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. *Nature* **517**, 165–169 (2015).
18. Ndeh, D. et al. Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* **544**, 65–70 (2017).
19. Thomas, F. et al. 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 (2012).
20. Hehemann, J. H., Kelly, A. G., Pudlo, N. A., Martens, E. C. & Boraston, A. B. Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. *Proc. Natl Acad. Sci. USA* **109**, 19786–19791 (2012).
21. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–D495 (2014).
22. Barbeyron, T., Gerard, A., Potin, P., Henrissat, B. & Kloareg, B. The kappa-carrageenase of the marine bacterium *Cytophaga drobachiensis*. Structural and phylogenetic relationships within family-16 glycoside hydrolases. *Mol. Biol. Evol.* **15**, 528–537 (1998).
23. Barbeyron, T., Michel, G., Potin, P., Henrissat, B. & Kloareg, B. iota-Carrageenases constitute a novel family of glycoside hydrolases, unrelated to that of kappa-carrageenases. *J. Biol. Chem.* **275**, 35499–35505 (2000).
24. Rebuffet, E. et al. Identification of catalytic residues and mechanistic analysis of family GH82 iota-carrageenases. *Biochemistry* **49**, 7590–7599 (2010).
25. Craigie J. Cell Walls. in Biology of the red algae (eds Cole K., Sheath R.). (Cambridge University Press, Cambridge, 1990).
26. Barbeyron, T. et al. Matching the diversity of sulfated biomolecules: creation of a classification database for sulfatases reflecting their substrate specificity. *PLoS ONE* **11**, e0164846 (2016).
27. Fujita, K., Takashi, Y., Obuchi, E., Kitahara, K. & Suganuma, T. Characterization of a novel beta-L-arabinofuranosidase in *Bifidobacterium longum*: functional elucidation of a DUF1680 protein family member. *J. Biol. Chem.* **289**, 5240–5249 (2014).
28. Kiyohara, M. et al. Alpha-N-acetylgalactosaminidase from infant-associated bifidobacteria belonging to novel glycoside hydrolase family 129 is implicated in alternative mucin degradation pathway. *J. Biol. Chem.* **287**, 693–700 (2012).
29. Popper, Z. A. et al. Evolution and diversity of plant cell walls: from algae to flowering plants. *Annu. Rev. Plant Biol.* **62**, 567–590 (2011).
30. Correc, G. et al. Comparison of the structures of hybrid κ-/β-carrageenans extracted from *Furcellaria lumbricalis* and *Tichocarpus crinitus*. *Carbohydr. Polym.* **88**, 31–36 (2012).
31. Jackson, P. The use of polyacrylamide-gel electrophoresis for the high-resolution separation of reducing saccharides labelled with the fluorophore 8-aminonaphthalene-1,3,6-trisulphonic acid. Detection of picomolar quantities by an imaging system based on a cooled charge-coupled device. *Biochem. J.* **270**, 705–713 (1990).
32. Yun, E. J. et al. The novel catabolic pathway of 3,6-anhydro-L-galactose, the main component of red macroalgae, in a marine bacterium. *Environ. Microbiol.* **17**, 1677–1688 (2015).
33. Lee, S. B., Kim, J. A. & Lim, H. S. Metabolic pathway of 3,6-anhydro-D-galactose in carrageenan-degrading microorganisms. *Appl. Microbiol. Biotechnol.* **100**, 4109–4121 (2016).
34. Vetting, M. W., Bouvier, J. T., Gerlt, J. A. & Almo, S. C. Purification, crystallization and structural elucidation of D-galactaro-1,4-lactone cycloisomerase from *Agrobacterium tumefaciens* involved in pectin degradation. *Acta. Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **72**, 36–41 (2016).
35. Zhu, Y. et al. Genetic analyses unravel the crucial role of a horizontally acquired alginate lyase for brown algal biomass degradation by *Zobellia galactanivorans*. *Environ. Microbiol.* **19**, 2164–2181 (2017).
36. Hehemann, J. H. et al. Biochemical and structural characterization of the complex agarolytic enzyme system from the marine bacterium *Zobellia galactanivorans*. *J. Biol. Chem.* **287**, 30571–30584 (2012).
37. Reher, M., Fuhrer, T., Bott, M. & Schonheit, P. The nonphosphorylative Entner-Doudoroff pathway in the thermoacidophilic euryarchaeon *Picrophilus torridus* involves a novel 2-keto-3-deoxygluconate- specific aldolase. *J. Bacteriol.* **192**, 964–974 (2010).
38. Blanvillain, S. et al. Plant carbohydrate scavenging through tonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. *PLoS ONE* **2**, e224 (2007).
39. Hehemann, J. H. et al. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* **464**, 908–912 (2010).
40. Rebuffet, E. et al. Discovery and structural characterization of a novel glycosidase family of marine origin. *Environ. Microbiol.* **13**, 1253–1270 (2011).
41. Grondin J. M., Tamura K., Dejean G., Abbott D. W. & Brumer H. Polysaccharide utilization loci: fuelling microbial communities. *J. Bacteriol.* doi:10.1128/JB.00860-16.
42. Groisillier, A. et al. MARINE-EXPRESS: taking advantage of high throughput cloning and expression strategies for the post-genomic analysis of marine organisms. *Microb. Cell Fact.* **9**, 45 (2010).
43. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
44. Ficko-Blean, E. et al. Biochemical and structural investigation of two paralogous glycoside hydrolases from *Zobellia galactanivorans*: novel insights into the evolution, dimerization plasticity and catalytic mechanism of the GH117 family. *Acta Crystallogr. D* **71**, 209–223 (2015).
45. Gasteiger E., et al. Protein identification and analysis tools on the ExPASy Server. in The Proteomics Protocols Handbook (ed. Walker J. M.) (Humana Press, Totowa, New Jersey, USA, 2005).
46. Michel, G. et al. Expression, purification, crystallization and preliminary x-ray analysis of the kappa-carrageenase from *Pseudoalteromonas carrageenovora*. *Acta Crystallogr. D* **55**, 918–920 (1999).
47. Michel, G. et al. Expression, purification, crystallization and preliminary X-ray analysis of the iota-carrageenase from Alteromonas fortis. *Acta Crystallogr. D* **56**, 766–768 (2000).
48. Buchanan, C. L., Connaris, H., Danson, M. J., Reeve, C. D. & Hough, D. W. An extremely thermostable aldolase from Sulfolobus solfataricus with specificity for non-phosphorylated substrates. *Biochem. J.* **343**, 563–570 (1999). Pt 3.
49. Powell, H. R. The Rossmann Fourier autoindexing algorithm in MOSFLM. *Acta Crystallogr. D* **55**, 1690–1695 (1999).
50. Leslie A. G. W. Recent changes to the MOSFLM package for processing film and image plate data. *Jnt. CCP4/ESF–EACBM Newslett. Protein Crystallogr.* **26**, (1992).
51. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D* **67**, 282–292 (2011).
52. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D* **62**, 72–82 (2006).
53. Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
54. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D* **58**, 1772–1779 (2002).
55. Ness, S. R., de Graaff, R. A., Abrahams, J. P. & Pannu, N. S. CRANK: new methods for automated macromolecular crystal structure solution. *Structure* **12**, 1753–1761 (2004).
56. Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30**, 1022–1025 (1997).
57. Vagin, A. A. et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D* **60**, 2184–2195 (2004).
58. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
59. Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125–132 (2010).
60. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D* **66**, 133–144 (2010).
61. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
62. Smart, O. S. et al. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D* **68**, 368–380 (2012).
63. Ropartz, D. et al. Performance evaluation on a wide set of matrix-assisted laser desorption ionization matrices for the detection of oligosaccharides in a high-throughput mass spectrometric screening of carbohydrate depolymerizing enzymes. *Rapid Commun. Mass. Spectrom.* **25**, 2059–2070 (2011).
64. ZoBell, C. Studies on marine bacteria I The cultural requirements of heterotrophic aerobes. *J Mar Res* **4**, 42–75 (1941).
65. Thomas, F., Barbeyron, T. & Michel, G. Evaluation of reference genes for real-time quantitative PCR in the marine flavobacterium Zobellia galactanivorans. *J. Microbiol. Methods* **84**, 61–66 (2011).
66. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864 (2011).

67. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
68. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
69. Toffano-Nioche, C. et al. Transcriptomic profiling of the oyster pathogen *Vibrio splendidus* opens a window on the evolutionary dynamics of the small RNA repertoire in the Vibrio genus. *RNA* **18**, 2201–2219 (2012).
70. Rutherford, K. et al. Artemis: sequence visualization and annotation. *Bioinformatics* **16**, 944–945 (2000).
71. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–1669 (2015).
72. Varet, H., Brillet-Gueguen, L., Coppee, J. Y. & Dillies, M. A. SARTools: A DESeq2- and Edger-based R pipeline for comprehensive differential analysis of RNA-Seq data. *PLoS ONE* **11**, e0157022 (2016).
73. Vallenet, D. et al. MicroScope in 2017: an expanding and evolving integrated resource for community expertise of microbial genomes. *Nucleic Acids Res.* **45**, D517–D528 (2016).
74. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
75. Criscuolo, A. & Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**, 210 (2010).
76. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**, 1164–1165 (2011).
77. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
78. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).

Acknowledgments

We are grateful to Dr Diane Jouanneau for critical discussions on the development on this project and to Dr Lionel Cladière and Ms. Anaïs Naretto for technical assistance. We thank Jonas Collén for taking and sharing with us an image that we submitted for consideration as a featured image on the journal's homepage. We thank the local contact support on the beamlines ID-14-4 and Proxima2 at the ESRF (Grenoble, France) and SOLEIL (Paris, France) with special thanks to Gavin Fox for his help with data treatment. This work has benefited from the platforms and expertize of the High-throughput Sequencing Platform of I2BC (Paris, France). We thank the INRA MIGALE bioinformatics platform (http://migale.jouy.inra.fr) for providing computational resources, help and support. G.M., E.D, and C.M. acknowledge support from the Agence National de la Recherche (ANR) with regard to the "Blue Enzymes" project (reference ANR-14-CE19-0020-01). G.M. and M.C. are also grateful to ANR for its support with regards to the investment expenditure program IDEALG (http://www.idealg.ueb.eu/, grant agreement No. ANR-10-BTBR-04). M.J.M. thanks the National Science Foundation for the grant MCB-1516990. K.A.S. thanks the Australian Research Council (FT100100291) and the Centre for Microscopy, Characterisation and Analysis at The University of Western Australia.

Author Contributions

E.F.-B. purified and characterized the ZGAL_3633, ZGAL_4655 (GH2), ZGAL_3147, ZGAL_3148, ZGAL_3150 (DagA1-3, GH127-1-3), and ZGAL_3152 (DagB, GH129-like) enzymes, crystallized and solved the structure of ZGAL_3152, and wrote the paper. A.P. and R.L. cloned and expressed all target-genes. A.P. purified and characterized the sulfatases and GH2 enzymes and assisted in writing the paper. M.J. characterized the SusD-like proteins. F.T., R.L., Y.T., T.B. and M.J.McB. performed the genetic experiments and the phenotyping of the *Zobellia* mutant strains. F.T. assisted in writing the paper. T.R., D. P.-P. and E.D. performed all the transcriptomics experiments and analyzes. T.R. assisted in writing the paper. M.S., A.C., B.V., C.M. and D.V. performed all the bioinformatics analyzes. M.S, A.C. and D.V. assisted in writing the paper. S.G. and M.J. purified and characterized ZGAL_3155 (DauA), ZGAL_3156 (DauB), ZGAL_3154 (DauC), and ZGAL_3153 (DauD) and assisted in writing the paper. S.G. was involved in the characterization of the GH2 enzymes. D.R. and H.R. performed the mass spectrometry experiments. G.C. produced and purified the furcellaran oligosaccharides. M.M. performed the site-directed mutagenesis of ZGAL_3152. K.A.S. synthesized initial potential substrates for ZGAL_3152 and assisted in writing the paper. M.J. crystallized ZGAL_3156 and A.J. assisted with the crystallization experiments. M.C. collected crystallography data and assisted in writing the paper. G.M. designed research, determined the structure of ZGAL_3156 and wrote the paper with E.F.-B.

Additional information

Supplementary Information accompanies this paper at doi:10.1038/s41467-017-01832-6.

Competing interests: The authors declare no competing financial interests.