AMERICAN SOCIETY FOR MICROBIOLOGY | mSphere®

RESEARCH ARTICLE
Applied and Environmental Science

A Multifunctional Polysaccharide Utilization Gene Cluster in *Colwellia echini* Encodes Enzymes for the Complete Degradation of κ-Carrageenan, ι-Carrageenan, and Hybrid β/κ-Carrageenan

Line Christiansen,[a]* Duleepa Pathiraja,[b] Pernille Kjersgaard Bech,[a]* Mikkel Schultz-Johansen,[a]* Rosanna Hennessy,[a] David Teze,[c] In-Geol Choi,[b] Peter Stougaard[a]*

[a]Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark
[b]Department of Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul, South Korea
[c]DTU Bioengineering, The Technical University of Denmark, Lyngby, Denmark

Line Christiansen and Duleepa Pathiraja contributed equally to the article. Author order was determined according to timing of work in the project.

ABSTRACT Algal cell wall polysaccharides constitute a large fraction in the biomass of marine primary producers and are thus important in nutrient transfer between trophic levels in the marine ecosystem. In order for this transfer to take place, polysaccharides must be degraded into smaller mono- and disaccharide units, which are subsequently metabolized, and key components in this degradation are bacterial enzymes. The marine bacterium *Colwellia echini* A3^T is a potent enzyme producer since it completely hydrolyzes agar and κ-carrageenan. Here, we report that the genome of *C. echini* A3^T harbors two large gene clusters for the degradation of carrageenan and agar, respectively. Phylogenetical and functional studies combined with transcriptomics and *in silico* structural modeling revealed that the carrageenolytic cluster encodes furcellaranases, a new class of glycoside hydrolase family 16 (GH16) enzymes that are key enzymes for hydrolysis of furcellaran, a hybrid carrageenan containing both β- and κ-carrageenan motifs. We show that furcellaranases degrade furcellaran into neocarratetraose-43-*O*-monosulfate [DA-(α1,3)-G4S-(β1,4)-DA-(α1,3)-G], and we propose a molecular model of furcellaranases and compare the active site architectures of furcellaranases, κ-carrageenases, β-agarases, and β-porphyranases. Furthermore, *C. echini* A3^T was shown to encode κ-carrageenases, ι-carrageenases, and members of a new class of enzymes, active only on hybrid β/κ-carrageenan tetrasaccharides. On the basis of our genomic, transcriptomic, and functional analyses of the carrageenolytic enzyme repertoire, we propose a new model for how *C. echini* A3^T degrades complex sulfated marine polysaccharides such as furcellaran, κ-carrageenan, and ι-carrageenan.

IMPORTANCE Here, we report that a recently described bacterium, *Colwellia echini*, harbors a large number of enzymes enabling the bacterium to grow on κ-carrageenan and agar. The genes are organized in two clusters that encode enzymes for the total degradation of κ-carrageenan and agar, respectively. As the first, we report on the structure/function relationship of a new class of enzymes that hydrolyze furcellaran, a partially sulfated β/κ-carrageenan. Using an *in silico* model, we hypothesize a molecular structure of furcellaranases and compare structural features and active site architectures of furcellaranases with those of other GH16 polysaccharide hydrolases, such as κ-carrageenases, β-agarases, and β-porphyranases. Furthermore, we describe a new class of enzymes distantly related to GH42 and GH160 β-galactosidases and show that this new class of enzymes is active only on hybrid β/κ-carrageenan oligosaccharides. Finally, we propose a new model for how the carrageenolytic enzyme repertoire enables *C. echini* to metabolize β/κ-, κ-, and ι-carrageenan.

Citation Christiansen L, Pathiraja D, Bech PK, Schultz-Johansen M, Hennessy R, Teze D, Choi I-G, Stougaard P. 2020. A multifunctional polysaccharide utilization gene cluster in *Colwellia echini* encodes enzymes for the complete degradation of κ-carrageenan, ι-carrageenan, and hybrid β/κ-carrageenan. mSphere 5:e00792-19. https://doi.org/10.1128/mSphere.00792-19.

Editor Steven J. Hallam, University of British Columbia

Address correspondence to Peter Stougaard, pst@envs.au.dk.

* Present address: Line Christiansen, ALK-Abelló Nordic a/S, Hørsholm, Denmark; Pernille Kjersgaard Bech, DTU Bioengineering, the Technical University of Denmark, Lyngby, Denmark; Mikkel Schultz-Johansen, University of Bremen, Center for Marine Environmental Sciences, Bremen, Germany; Peter Stougaard, Department of Environmental Science, Aarhus University, Aarhus, Denmark.

Received 6 November 2019
Accepted 4 December 2019
Published 8 January 2020

KEYWORDS furcellaran, carrageenan, marine bacteria, algal polysaccharides, glycoside hydrolases, metabolic pathway

Primary production in the marine environment adds up to approximately half of the primary production on Earth (1). Thus, degradation of marine phytoplankton and of macrophytes (seaweed and sea grasses) is important in order for nutrients to be transferred between trophic levels in the marine food web; in particular, cell wall polysaccharides constitute a large fraction of the biomass of marine primary producers (2, 3). The compositions of cell walls from marine algae share several features with those of land plants: Both plant groups have cellulose and hemicellulose in their cell walls, but in contrast to land plants, marine algae contain a variety of sulfated polysaccharides, e.g., agar, porphyran, and carrageenan in red algae; ulvan in green algae; and fucan in brown algae (4). Consequently, in order for organisms to degrade algal polysaccharides, they must produce enzymes that are active on the respective sulfated polysaccharides (5).

Like terrestrial herbivores, marine herbivores harbor endosymbiotic microorganisms that produce enzymes needed for the hydrolysis of plant cell wall polysaccharides, notably, cellulose and hemicellulose but also more-complex sulfated carbohydrates. In marine iguanas, the fecal microbiota produce enzymes specific for utilization of marine polysaccharides (6), turban shell (*Batillus cornutus*) feeding on brown algae harbor intestinal bacteria producing cellulases, alginate lyases, laminarinases, and "kelp lyases" (7), and alginolytic bacteria may be found in the gut of sea urchins (*Strongylocentrotus* sp.) and abalones (*Haliotis* sp.) (8).

We have previously isolated a bacterium, *Colwellia echini* A3^T, from the intestine of the sea urchin *Strongylocentrotus droebachiensis* and found that this bacterium solubilizes agar and carrageenan plate media as a result of enzymatic hydrolysis (9). Here, we demonstrate that *C. echini* harbors two large multifunctional polysaccharide utilization loci (PULs) encoding genes for the total catabolism of not only agar and κ-carrageenan but also ι-carrageenan and the hybrid β/κ-carrageenan furcellaran. Furthermore, we show that furcellaran degradation by *C. echini* is catalyzed by enzymes belonging to a new GH16_13 subfamily that are similar to those previously reported for another marine bacterium, *Paraglaciecola hydrolytica* S66^T (10, 11). Using transcriptomics, *in silico* analyses, and recombinant enzyme technology, we provide more information about the reaction mechanism of GH16_13 furcellarases and characterize the oligosaccharide products that they release. Finally, we present a model for how *C. echini* A3^T degrades furcellaran, κ-carrageenan, and ι-carrageenan. The results presented here not only improve our understanding of degradation of sulfated marine polysaccharides but also may promote the generation of sulfated oligosaccharides to be used in biotechnology and pharma.

RESULTS

Polysaccharide degradation potential of *Colwellia echini* A3^T. Bioinformatic analyses showed that the genome of *C. echini* encodes a large number of carbohydrate-active enzymes (CAZymes) (12). In comparisons to phylogenetically related *Colwellia* species, *C. echini* A3^T encoded the largest amount of CAZymes (see Table S1A in the supplemental material). In accordance with the agarolytic nature of *C. echini* A3^T and *C. agarivorans* QM50^T (13), these two bacterial species encoded enzymes belonging to GH families with predominantly agar-degrading representatives such as GH50 (β-agarases), GH86 (β-agarases/porphyranases), and GH117 (α-3,6-anhydrogalactosidases), which are keystone enzymes in and indicative of agar catabolism (14). In contrast, such genes were absent in genome sequences of the nonagarolytic relatives *C. psychrerythraea* 34H^T, *C. piezophila* Y223G^T, and *C. chukchiensis* BCw111^T (Table S1B). Unique for *C. echini* was the presence of GH96 α-agarases, which are rarely reported and include only a few characterized members (15–18). Of all the *Colwellia* genomes analyzed, that of *C. echini* encoded the largest amount of GH16 CAZymes, comprising β-agarases,

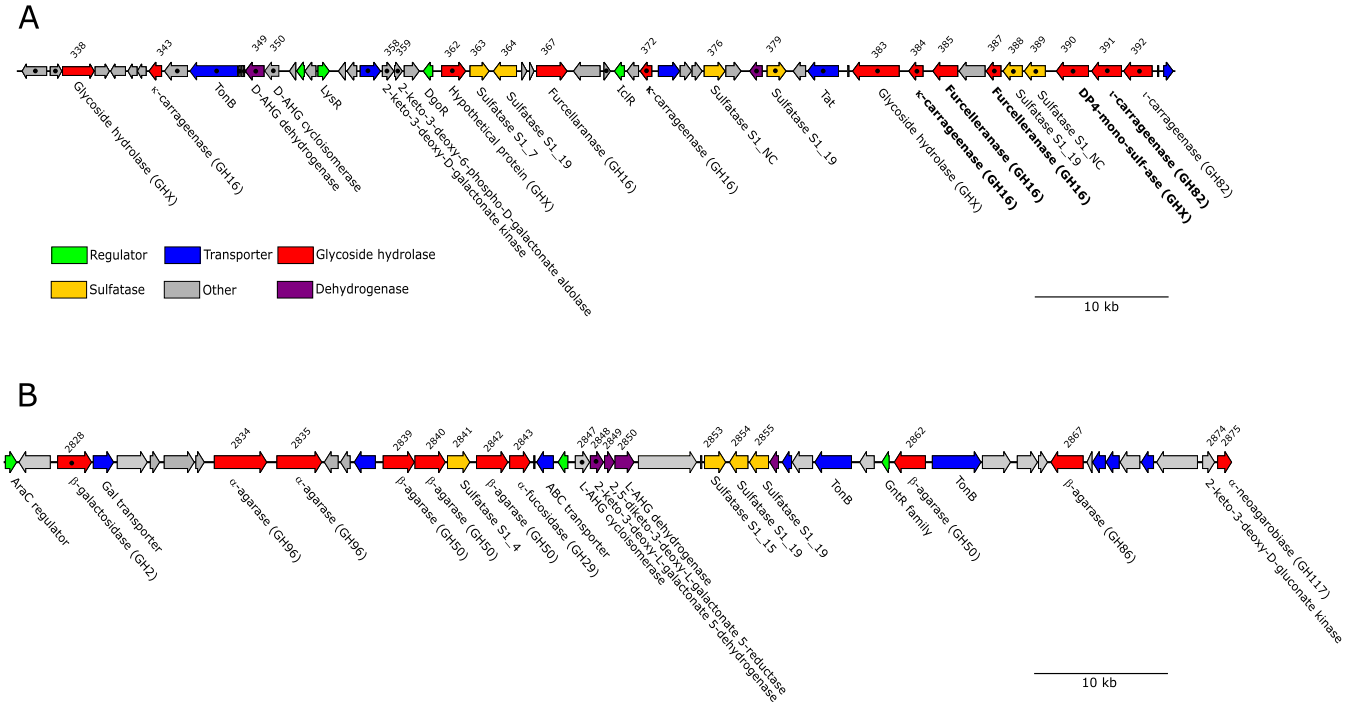

FIG 1 Map of the carrageenolytic CAR PUL (A) and of the agarolytic AGA PUL (B). Arrows indicate open reading frames (ORFs), and the color code shows the predicted or documented function of the ORFs. Gene functions documented in this work are shown in bold, and genes where the function was predicted by bioinformatic analyses are shown in regular font. ORF numbers are indicated above the gene map. Black dots inside ORFs indicate that the gene was upregulated more than 2 log in transcriptomics analysis when *C. echini* A3^T was cultivated with κ-carrageenan.

β-porphyranases, laminarinases, and κ-carrageenases, and only *C. echini* encoded GH82 ι-carrageenases.

The genes that encode putative agarases and carrageenases were localized in two large gene clusters: one ~92,000-bp region dedicated to agar degradation and one ~86,000-bp region dedicated to carrageenan degradation (Fig. 1). TonB-dependent receptors, previously proposed to be analogous to the *Bacteroidetes* detection system (10, 19, 20), were located in both gene clusters (Ce2863 and Ce345; Fig. 1). Furthermore, as putative transporters were similarly located in the gene clusters, we hypothesize that the two regions could be analogous to the polysaccharide utilization loci (PULs) originally described in *Bacteriodetes* (19). Thus, we named the gene cluster encoding agarolytic enzymes AGA PUL and the cluster with carrageenolytic genes CAR PUL (Fig. 1).

Glycoside hydrolases in the CAR PUL gene cluster. An automatic annotation showed that CAR PUL encoded three putative GH16 β-porphyranases (Ce367, Ce385, and Ce387), three putative GH16 κ-carrageenases (Ce343, Ce372, and Ce384), and two putative GH82 ι-carrageenases (Ce391 and Ce392) (Fig. 1A; see also Table S2). However, the level of identity between the putative glycoside hydrolases encoded by CAR PUL and sequences in the Protein Data Bank database (PDB) or Swiss-Prot database was low (Table S2). Therefore, a phylogenetic analysis of the glycoside hydrolases was initiated. GH16 sequences with known function from the PDB database and sequences from recently characterized, novel furcellaranases from *P. hydrolytica* S66^T (10) were compared to GH16 sequences from CAR PUL (Fig. 2). Enzymes Ce367, Ce385, and Ce387 grouped with the newly characterized GH16_13 furcellaranases Ph1656, Ph1663, and Ph1675 from *P. hydrolytica* S66^T (10). Ce387 showed 57% identity to Ph1656 and 46% identity to Ph1675, whereas Ce367 and Ce385 displayed less than 20% identity to the furcellaranases from *P. hydrolytica* S66^T. Also shown in Fig. 2 is the clustering of Ce343 and Ce384 with known κ-carrageenases. Ce372 clustered distantly with κ-carrageenases, but as the bootstrap value was very low (21%) and

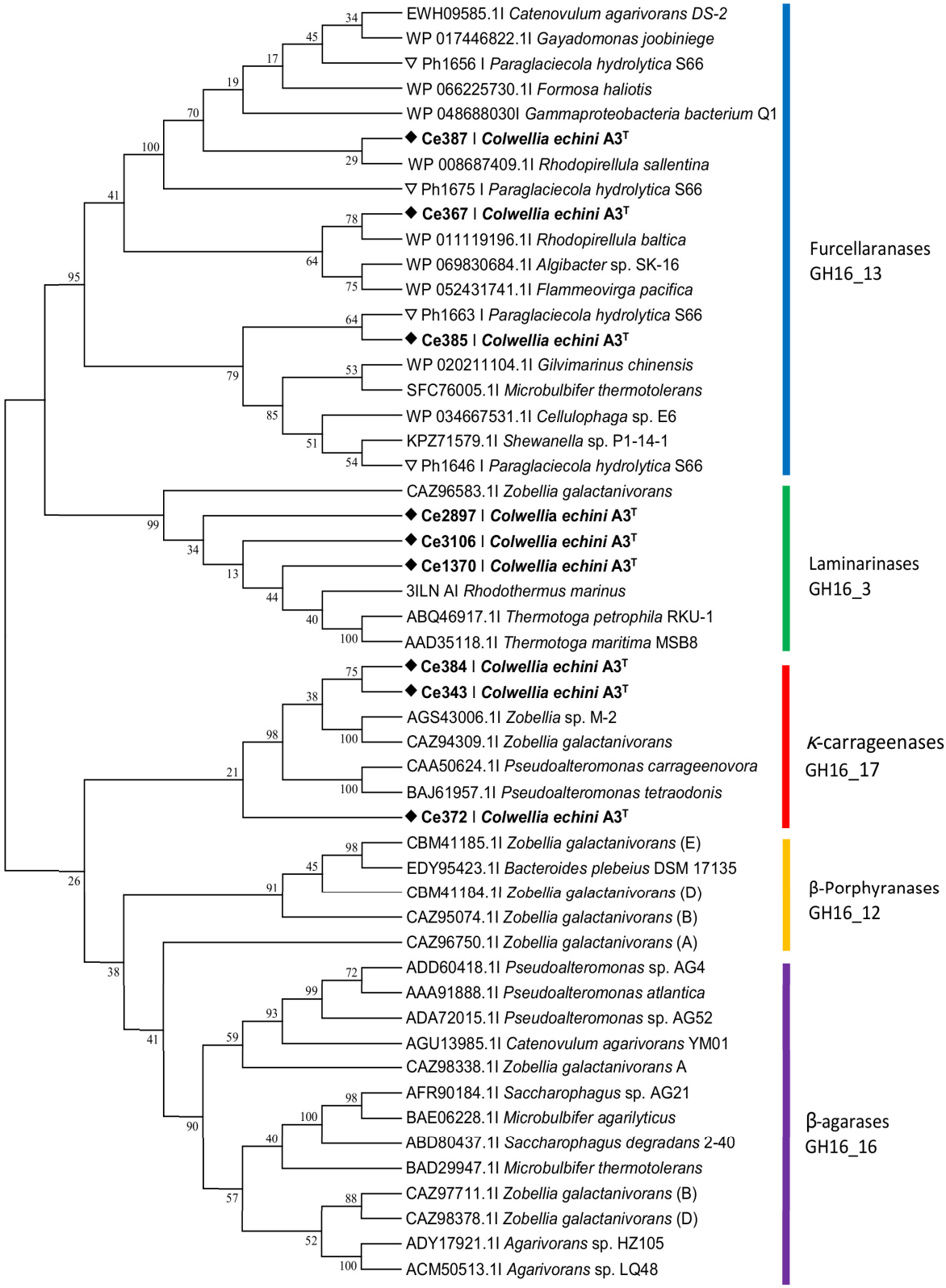

FIG 2 Phylogenetic tree of GH16 enzymes from *C. echini* A3ᵀ. Annotated β-agarases, κ-carrageenases, laminarinases, and β-porphyranases from *C. echini* A3ᵀ (black diamonds) were analyzed together with characterized proteins from the CAZy database. Furcellaranases from *C. echini* A3ᵀ were compared to characterized furcellaranases from *P. hydrolytica* (white triangles) and sequence homologs derived from the NCBI database. The tree was constructed by using the neighbor-joining method. Bootstrap values represent percentages of 1,000 replications.

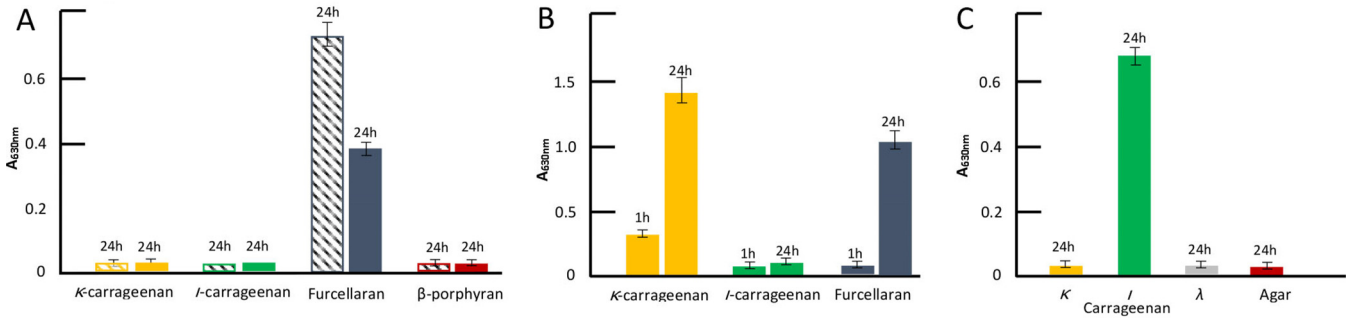

FIG 3 Substrate specificity of (A) furcellaranases (Ce385, dashed; Ce387, solid), (B) κ-carrageenase (Ce384), and (C) ι-carrageenase (Ce391). Agar, furcellaran, porphyran, κ-carrageenan, ι-carrageenan, and/or λ-carrageenan were used as substrates. The substrate concentration was 0.1% (wt/vol) in all enzyme reactions. The amount of reducing ends was determined by 3-methyl-2-benzothiazolinone hydrazone (MBTH) assay, and the number of aldehydes in the reducing end of carbohydrates was measured as absorbance at 630 nm (A_{630nm}). All assays were carried out in triplicate at 20°C and terminated after 1 h (1h) and/or 24 h (24h).

since we could not produce active enzyme, no activity could be assigned to Ce372. In addition, Ce1370, Ce2897, and Ce3106, located outside CAR PUL, clustered with known laminarinases. No enzyme sequences from strain A3^T grouped with sequences of characterized GH16_12 β-porphyranases or GH16_16 β-agarases (Fig. 2). A closer analysis of the AGA PUL region showed the presence of GH50 and GH86 β-agarases and the rare GH96 α-agarases, which could be the reason for the observed agarolytic activity; this activity is to be described in detail elsewhere.

Functional analyses were carried out in order to confirm the substrate specificity of Ce367, Ce385, and Ce387. Recombinant, His-tagged Ce385 and Ce387 were produced in *Escherichia coli* BL21, purified with nickel-nitrilotriacetic acid (Ni-NTA) chromatography, and assayed with furcellaran, porphyran, and κ- and ι-carrageenan as the substrates; Ce367 could not be expressed as a soluble and active enzyme in this study. Substrate specificities of the recombinant enzymes were analyzed in a reducing sugar assay (Fig. 3A) and by fluorophore-assisted carbohydrate electrophoresis (FACE) (Fig. 4). The results showed that Ce385 and Ce387 hydrolyzed furcellaran, a partially sulfated hybrid β/κ-carrageenan (21, 22), but not (more highly sulfated) κ-or ι-carrageenan or porphyran. No activity was detected with agar or agarose, and no activity on any substrates was detected in extracts from *E. coli* with empty vector (not shown). Matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) analysis showed that the hydrolysis products seen with furcellaran as the substrate were neocarrate-

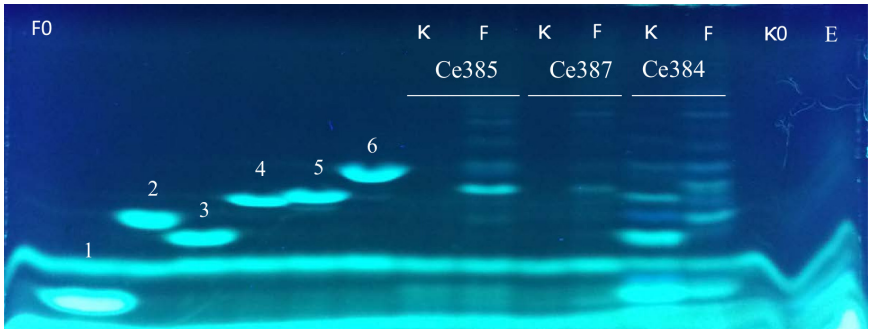

FIG 4 FACE analysis of enzymatic hydrolysis of 0.1% (wt/vol) κ-carrageenan (K) and 0.1% (wt/vol) furcellaran (F). Furcellaranases Ce385 and Ce387 and κ-carrageenase Ce384 were incubated with furcellaran or κ-carrageenan at 20°C overnight. Reference standards without enzyme as indicated as follows: 1, neocarrabiose-monosulfate (DP2-monoS); 2, neocarratetraose-monosulfate (DP4-monoS); 3, neocarratetraose-disulfate (DP4-diS); 4, neocarrahexaose-trisulfate (DP6-triS); 5, neocarrahexaose-tetrasulfate (DP6-tetraS); 6, neocarraoctaose-tetrasulfate (DP8-tetraS). F0, furcellaran without enzymes; κ0, κ-carrageenan without enzymes; E, mixture of all enzymes and no carbohydrate substrate.

TABLE 1 Key residues in the −1 subsites of algal cell wall-degrading GH16 enzymes

Enzyme-activity	Organism	PDB code(s) or name	Platform	Catalytic residues	−1 sugar	4 interacting residues
κ-Carrageenase	*Z. galactanivorans*	5OCR	Trp^{141}	Glu^{159}, Asp^{161}, Glu^{164}	G4S	Trp^{74}, Arg^{263}
	P. carrageenovora	5OCQ, 1DYP	Trp^{144}	Glu^{163}, Asp^{165}, Glu^{168}	G4S	Trp^{67}, Arg^{260}
	C. echini	Ce343 (model)	Trp^{155}	Glu^{180}, Asp^{182}, Glu^{185}	G4S	Trp^{78}, Arg^{286}
	C. echini	Ce384 (model)	Trp^{189}	Glu^{214}, Asp^{216}, Glu^{219}	G4S	Trp^{110}, Arg^{324}
β-Agarase	*M. thermotolerans*	3WZ1	Trp^{138}	Glu^{147}, Asp^{149}, Glu^{152}	Gal	Glu^{257}
	Z. galactanivorans	1O4Y	Trp^{138}	Glu^{147}, Asp^{149}, Glu^{152}	Gal	Glu^{254}
β-Porphyranase	*Z. galactanivorans*	3JUU (PorB)	Trp^{139}	Glu^{156}, Asp^{158}, Glu^{161}	Gal	Glu^{256}
	Z. galactanivorans	3ILF (PorA)	Trp^{131}	Glu^{139}, Asp^{141}, Glu^{144}	Gal	Glu^{234}
	B. plebeius	4AWD	Trp^{149}	Glu^{173}, Asp^{175}, Glu^{178}	Gal	Glu^{284}
Furcellaranase	*C. echini*	Ce385 (model)	Trp^{139}	Glu^{146}, Asp^{148}, Glu^{151}	Gal	Glu^{271}
	C. echini	Ce387 (model)	Trp^{202}	Glu^{209}, Asp^{211}, Glu^{214}	Gal	Glu^{320}
	P. hydrolytica	Ph1631 (model)	Trp^{137}	Glu^{146}, Asp^{148}, Glu^{151}	Gal	Glu^{254}
	P. hydrolytica	Ph1663 (model)	Trp^{149}	Glu^{156}, Asp^{158}, Glu^{161}	Gal	Glu^{255}
	P. hydrolytica	Ph1675 (model)	Trp^{176}	Glu^{193}, Asp^{195}, Glu^{198}	Gal	Glu^{305}

traose monosulfate and neocarrahexaose monosulfate (see Fig. S3 in the supplemental material).

***In silico* specificity study of furcellaranases.** From the structure of substrates and products of furcellaranase-catalyzed reactions, a cleavage event leaving either a G4S or a galactose (Gal) at the reducing end could be envisioned. Remarkably, the GH16 family comprises enzymes that are able to catalyze reactions leading to products presenting a sulfate at the fourth hydroxyl of the reducing-end sugar (e.g., κ-carrageenases) or an unsulfated sugar (e.g., β-agarases and β-porphyranases), and the molecular determinants allowing these two distinct specificities have previously been highlighted. Indeed, in known κ-carrageenase structures (PDB 5OCR, *Zobellia galactanivorans*; PDB 5OCQ and 1DYP, *P. carrageenovora*), an arginine and a tryptophan are responsible for binding to the sulfate of G4S in the –1 subsite; while a glutamate was reported to be a critical residue for recognition of the fourth hydroxyl of nonsulfated D-galactosidase (D-Gal) in the –1 subsite of both β-porphyranases (3JUU and 3ILF, *Z. galactanivorans*; 4AWD, *Bacteroides plebeius*) and β-agarases (3WZ1, *Marinactinospora thermotolerans*; 1O4Y, *Z. galactanivorans*) (23–25). To assess whether furcellaranases presented one of these residue sets, 2,226 GH16 sequences, sharing 10% to 80% pairwise identity and including characterized representatives for β-agarase, κ-carrageenase, porphyranase, and furcellaranase activities, were aligned by iterative multiple-sequence alignment. The characteristic GH16 catalytic triad ExDxxE was found to be largely conserved (99.5%, 97.2%, and 99.2%, respectively, for each catalytic residue). The glutamate recognizing nonsulfated Gal, together with a tryptophan in an ExxxW motif, was found in 65% of the sequences, including all the β-agarase, porphyranase, and furcellaranase sequences (Table 1). As GH16 catalysis proceeds through important conformational changes of the sugar in the –1 subsite, binding a distorted $^{1}S_{3}$ skew-boat conformation (26), it requires particularly strong interaction network at this site. Therefore, it is likely that furcellaranases accommodate a nonsulfated Gal in their –1 subsite. *In silico* folding of the *C. echini* furcellaranases and κ-carrageenase over known structures of GH16 showed that in all cases, the catalytic triad ExDxxE and the platform Trp that precedes it, as well as the pairs of residues pinpointed above as recognizing either Gal or G4S, superimposed well (Fig. S1).

Functional characterization of carrageenan degradation. Phylogenetic analysis indicated that Ce343 and Ce384 are representatives of GH16_17 κ-carrageenases (Fig. 2). Functional studies showed that recombinant Ce384 produced in *E. coli* hydrolyzed κ-carrageenan and the hybrid β/κ-carrageenan furcellaran but not ι-carrageenan (Fig. 3B; see also Fig. 4) or agar or agarose (not shown). MALDI-TOF analysis showed that Ce384 degrades κ-carrageenan into neocarrabiose monosulfate (DA-G4S). Oligo-

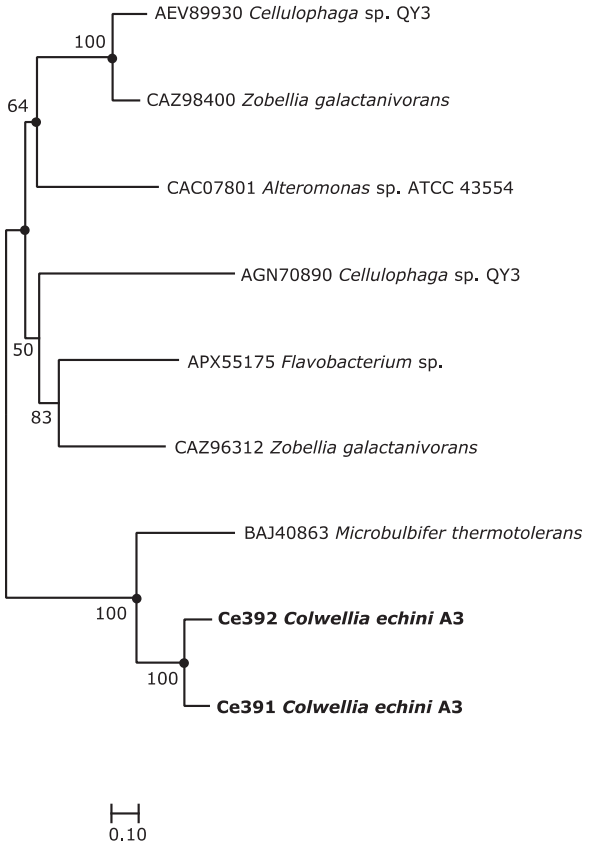

FIG 5 Neighbor-joining tree based on amino acid sequences of ι-carrageenases Ce391 and Ce392 from *C. echini* A3^T and related sequences retrieved from the NCBI database. Numbers at nodes are bootstrap values and represent percentages of 1,000 replicates; only values of >50% are shown. Filled circles indicate that the corresponding nodes were also recovered in the tree generated with the maximum likelihood algorithm. Bar, 0.1 changes per position.

saccharide products released from furcellaran were neocarrabiose monosulfate (DA-G4S), neocarratetraose monosulfate (DA-G4S-DA-G or DA-G-DA-G4S), and neocarrahexaose monosulfate (DA-G4S-DA-G-DA-G or DA-G-DA-G4S-DA-G or DA-G-DA-G-DA-G4S) (Fig. S3). *In silico* modeling showed that κ-carrageenases Ce343 and Ce384 superimposed with κ-carrageenase experimental structures from *P. carrageenovora* (1DYP) and *Z. galactanivorans* (5OCR) (Fig. S1).

Ce391 and Ce392 displayed low-level phylogenetic relationships (14% to 19% identity) to known ι-carrageenases in the PDB and Swiss-Prot databases (Fig. 5; see also Table S2). Biochemical analyses showed that recombinant Ce391 degrades ι-carrageenan but not agar, furcellaran, or κ- or λ-carrageenan (Fig. 3C; see also Fig. 6). Negative controls with extracts from empty vector containing *E. coli* showed no degradation of any of the polysaccharides. The band pattern in the FACE gel indicated that the products were neocarrabiose disulfate (DA2S-G4S) and possibly neocarratetraose tetrasulfate (DA2S–G4S-DA2S–G4S).

Miscellaneous CAZymes in CAR PUL. In addition to GH16 and GH82 enzymes, bioinformatic analyses identified four additional open reading frames (ORFs; *Ce338*, *Ce362*, *Ce383*, and *Ce390*) that encode putative glycoside hydrolases with low similarity to known CAZymes. Phylogenetic analyses showed that Ce338 and Ce390 protein sequences were distantly related (less than 17% identity) to those of characterized lactose-specific β-galactosidases from GH42 and of newly characterized GH160 enzymes (Fig. 7). However, a BLASTp search revealed that sequences similar to those of Ce338 and Ce390 were discovered in several marine bacteria, including *P. hydrolytica* S66^T (10). Biochemical analysis showed that recombinant Ce390 hydrolyzed neocarrao-

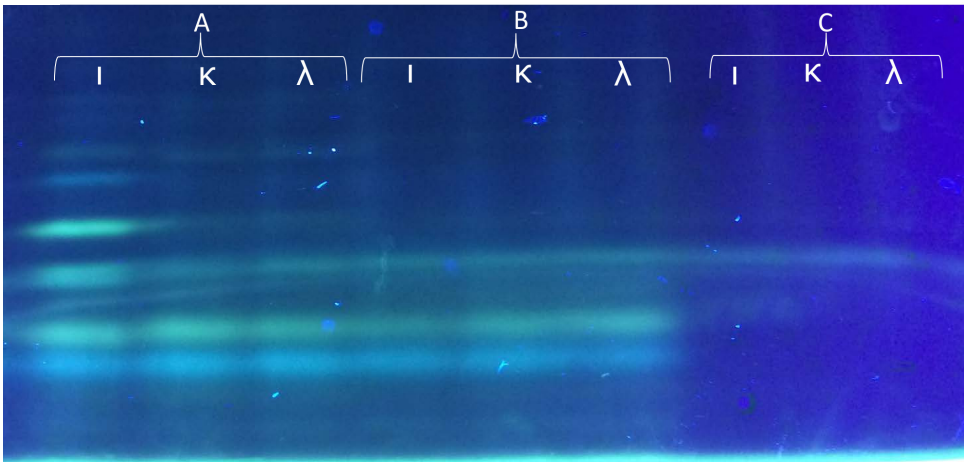

FIG 6 ι-Carrageenase activity visualized by FACE. 0.1% (wt/vol) ι-, λ-, or κ-carrageenan (ι, λ, or κ) incubated at 20°C overnight (A) with recombinant Ce391 enzyme extracts or (B) with extracts from *E. coli* with empty vector or (C) without enzyme extracts.

ligosaccharides such as neocarratetraose 41-*O*-monosulfate (DA-G-DA-G4S), neocarratetraose 41,43-*O*-disulfate (DA-G4S-DA-G4S), and neocarrahexaose 24,41,43,45-*O*-tetrasulfate (DA-G4S-DA2S-G4S-DA-G4S) to the corresponding unsulfated and sulfated neocarrabioses (Fig. 8; see also Fig. S2); lactose was not hydrolyzed (data not shown). As argued by Schultz-Johansen et al. (10), this GH42/GH160-like enzyme could be the enzyme that was described by McLean and Williamson (27), who published a similar enzyme activity from *Pseudoalteromonas carrageenovora* Psc^T and named the enzyme "neocarratetraose 4-*O*-monosulphate β-hydrolase." Although BLASTP indicated a distant relationship with GH42 β-galactosidases (Table S2), the level of identity between GH42 (and GH160) enzymes and *Colwellia* enzymes Ce338 and Ce390 was low. Homologs to Ce338 and Ce390 and to Ce362 and Ce383 were discovered in other marine bacteria (Fig. 7), but since no enzyme activity could be ascribed to Ce362 or Ce383 and since no structures were known for any of the enzymes, more-detailed phylogenetic analyses were not possible.

Model for carrageenan utilization by *C. echini* A3^T. In addition to genes encoding furcellaranases, ι- and κ-carrageenases, and GH42/GH160-like enzymes, the CAR PUL gene cluster contained six genes encoding putative sulfatases (28) and enzymes expected to be involved in the final catabolism of hydrolysis products of carrageenan breakdown. Transcriptomics analysis showed that several of these sulfatases and auxiliary enzymes were upregulated under conditions of *C. echini* A3^T cultivation in the presence of κ-carrageenan (Fig. 1; see also Table S3). By combining functional studies, phylogenetical analyses, and transcriptomics, we were able to propose a model for how *C. echini* A3^T degrades hybrid β/κ-carrageenan, κ-carrageenan, and ι-carrageenan (Fig. 9). Furcellaranases Ce385 and Ce387 and possibly Ce367 degrade furcellaran to neocarratetraose-mono-sulfate (DA-G4S-DA-G/DA-G-DA-G4S) and neocarrahexaose-mono-sulfate (DA-G4S-DA-G-DA-G/DA-G-DA-G4S-DA-G/DA-G-DA-G-DA-G4S). Ce390 (and probably Ce338), upregulated in transcriptomics, subsequently degrades neocarratetraose-mono-sulfate to neocarrabiose (DA-G) and neocarrabiose-mono-sulfate (DA-G4S). Sulfatase Ce379 is upregulated during cultivation of strain A3^T on carrageenan and shows 62% identity to the S1_19 sulfatase PCAR9_p0034 (Table S4) that removes sulfate from galactose-4-sulfate in κ-type carrageenans in *P. carrageenovora* 9^T (29). Thus, we propose that Ce379 has a similar function in strain A3^T and converts DA-G4S to unsulfated carrabiose, namely, DA-G.

The κ-carrageenase Ce384 and Ce372 were upregulated in transcriptomics (log fold change [logFC] values of 3.69 and 3.03, respectively), and Ce384 was shown to degrade κ-carrageenan to DA-G4S, which would subsequently be desulfated to DA-G by the G4S sulfatase Ce379 as described above.

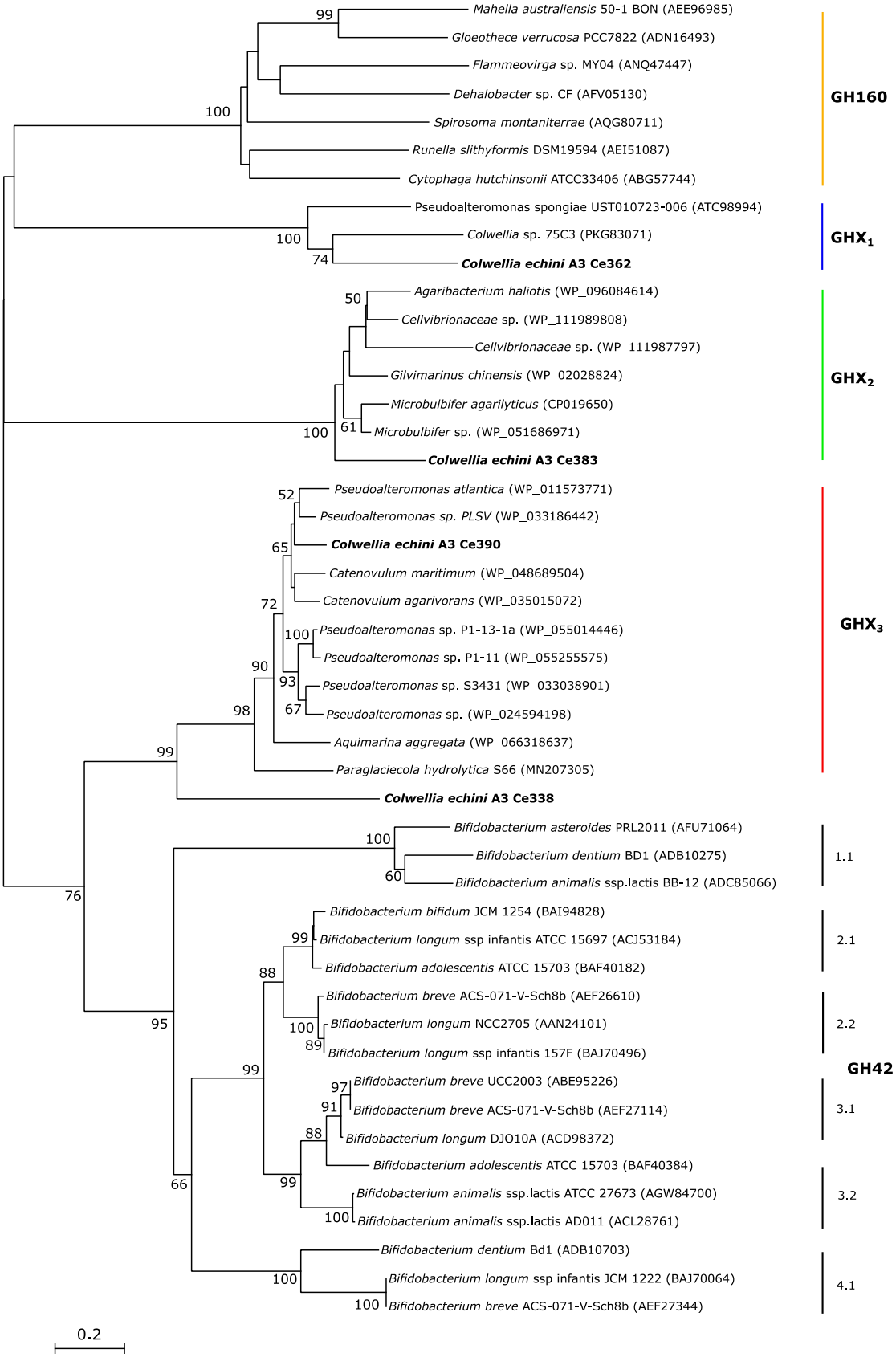

FIG 7 Neighbor-joining tree showing phylogenetical relationships between Ce338, Ce362, Ce383, and Ce390 from *C. echini* A3^T and characterized GH42 and GH160 β-galactosidases. Characterized sequences of Ph1657 furcellaranases from *P. hydrolytica* S66 that

(Continued on next page)

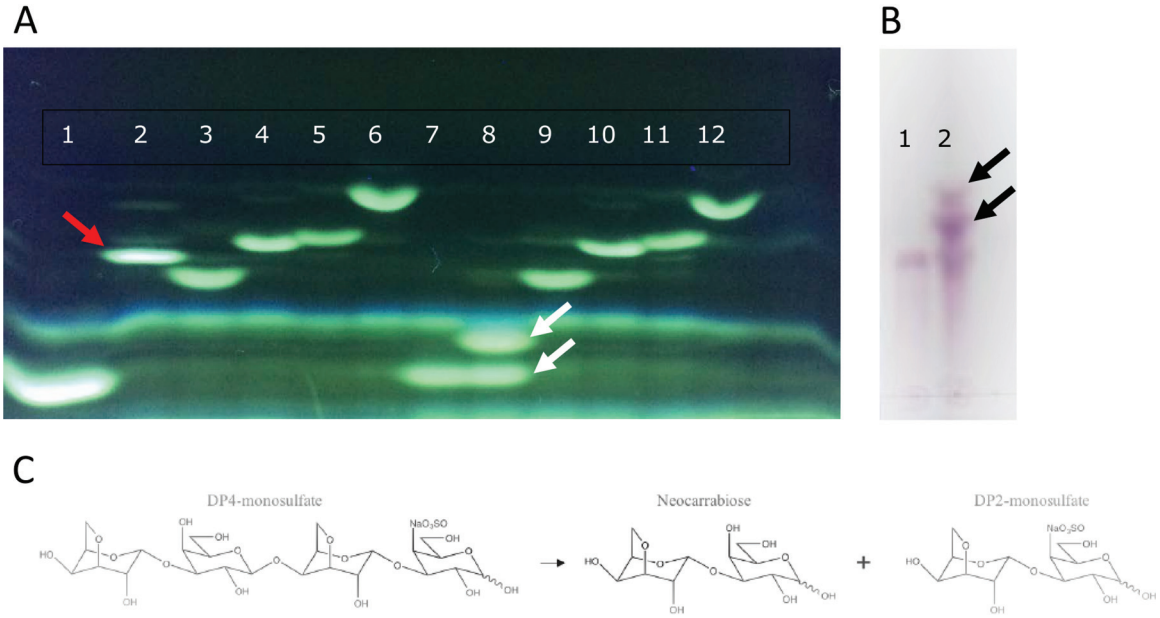

FIG 8 (A) FACE analysis showing 0.1% (wt/vol) neocarraoligosaccharides before and after digestion with Ce390. Lanes 1 and 7, DP2 monosulfate (monoS); lanes 2 and 8, DP4 monoS; lanes 3 and 9, DP4 diS; lanes 4 and 10; DP6 triS; lanes 5 and 11, DP6 tetraS; lanes 6 and 12, DP8 tetraS. Lanes 1 to 6 represent results obtained without enzyme; lanes 7 to 12 represent results obtained with Ce390 enzyme. Arrows indicate the position of DP4 monosulfate before (red) and after (white) digestion with Ce390. (B) TLC analysis of DP4 monosulfate before (lane 1) and after (lane 2) digestion with Ce390. Black arrows point to hydrolysis products. (C) Interpretation of activity of Ce390 on DP4 monosulfate.

A pathway for the degradation of ι-carrageenan in strain A3^T showed similarity to the pathways in *P. carrageenovora* 9^T (29) and *Z. galactanivorans* (20). The ι-carrageenase Ce391 and possibly Ce392 hydrolyze ι-carrageenan to ι-neocarrabiose (DA2S-G4S), which subsequently may be desulfated by a DA2S sulfatase and a G4S sulfatase. A search for these sulfatases in the genome sequence of A3^T showed that the S1_19 sulfatase Ce388 displayed 56% and 50% identities to the G4S sulfatases from *P. carrageenovora* 9^T (PCAR_p0023, S1_19) and *Z. galactanivorans* (ZGAL_3145, S1_19), respectively (Table S4), which previously have been reported to cleave the 4-sulfate in ι-carrageenan. Subsequent desulfation of the 2-sulfate may be carried out by Ce376 (S1_NC) and Ce389 (S1_NC), which display 58% and 61% identity to a ι-specific DA2S sulfatase from *P. carrageenovora* 9^T (PCAR_p0022; S1_NC) (29). Both ι-carrageenases (Ce391 and Ce392) and the two sulfatases Ce388 (S1_19) and Ce389 (S1_NC) were upregulated in transcriptomics analysis. The CAR PUL encoded two more putative sulfatases, Ce363 (S1_7) and Ce364 (S1_19), with less than 31% and 35% identity to known sulfatases (Table S4).

Thus, the net result of degradation of furcellaran and κ- and ι-carrageenan is hypothesized to be neocarrabiose (DA-G). In order for neocarrabiose to enter the central metabolism, this disaccharide must be hydrolyzed further to galactose and D-AHG (3,6-anhydro-D-galactose). Ficko-Blean et al. (20) showed that *Z. galactanivorans* encodes GH127 and GH129 enzymes capable of hydrolysing D-AHG from neocarraoligosaccharides, and Schultz-Johansen et al. (10) reported that a GH127 enzyme similar to those of *Z. galactanivorans* was found in *P. hydrolytica* S66^T. A search in the *C. echini* A3^T genome sequence for GH127 and GH129 enzymes produced negative results. However, as *C. echini* A3^T was able to grow with κ-carrageenan as the sole carbon

FIG 7 Legend (Continued)
hydrolyze neocarratetraose monosulfate and related sequences were retrieved from the NCBI database. Numbers at nodes represent bootstrap values and are shown as percentages of 1,000 replicates; only values of >50% are shown. Vertical bars to the right indicate the GH160 clade and the six subgroups of GH42, 1.1 to 4.1, according to Viborg et al. (50). Data representing GHX_1 to GHX_3 indicate glycoside hydrolase sequences that may represent novel GH families. Ce338, Ce362, Ce383, and Ce390 from this work were found in separate clades together with other uncharacterized proteins retrieved from the NCBI database. Bar, 0.2 changes per position.

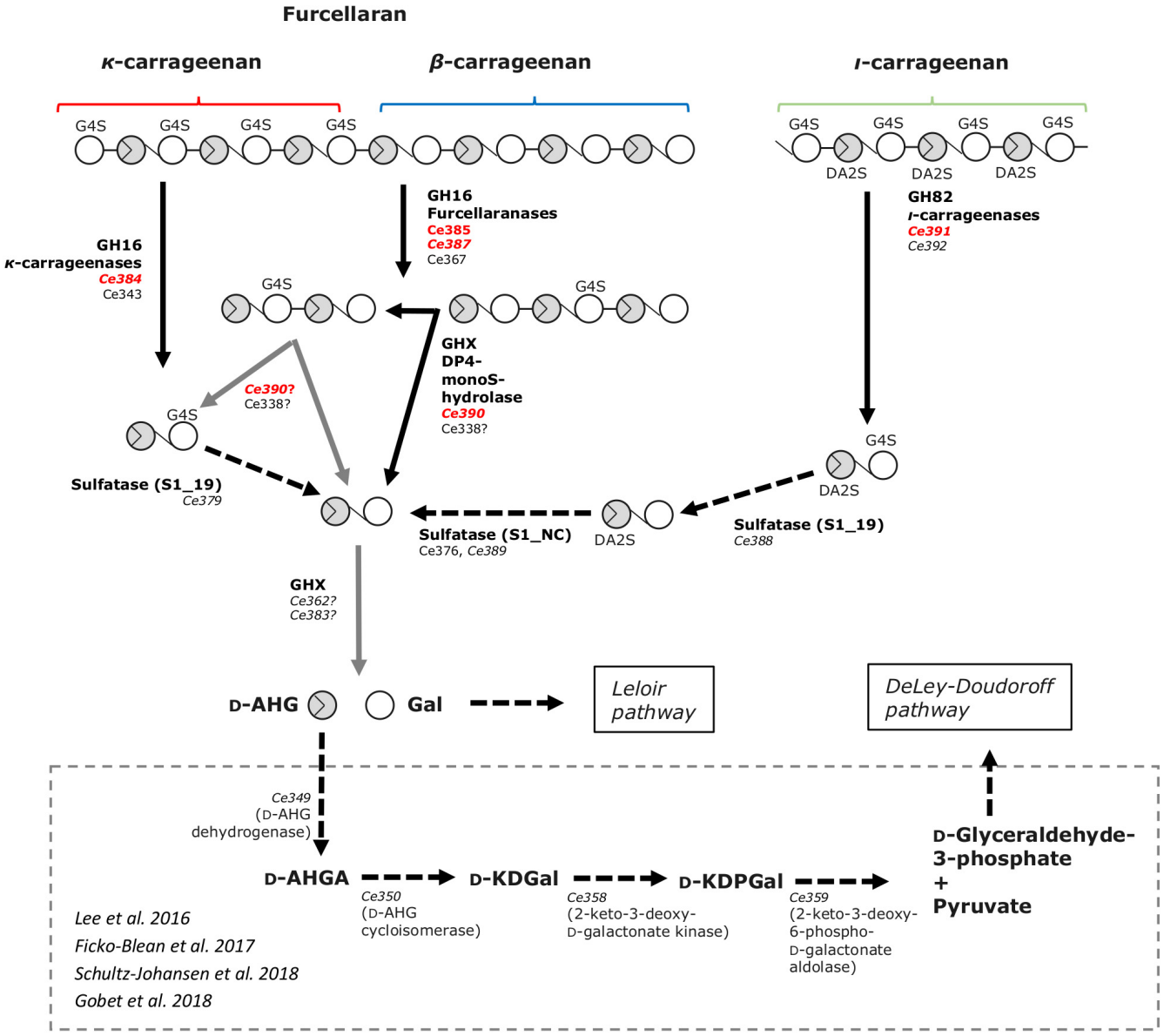

FIG 9 Pathway for degradation of β/κ-carrageenan (furcellaran), κ-carrageenan, and ι-carrageenan by *C. echini* A3^T. Functionally characterized enzymes are shown in red, whereas italics indicate enzymes that were upregulated in native *C. echini* A3^T cultivated with κ-carrageenan. Black arrows indicate reactions that were documented through functional and/or bioinformatics analyses. Gray arrows indicate hypothesized enzyme activities that are yet to be confirmed. Data from Lee et al. (30), Ficko-Blean et al. (20), Schultz-Johansen et al. (10), and Gobet et al. (29) are indicated.

source, the bacterium must contain a gene encoding an enzyme that hydrolyzes neocarrabiose to D-AHG and D-galactose.

Whereas D-galactose enters directly into the Leloir pathway, D-AHG must be processed further before entering into the central metabolism. Genes similar to those encoding the four D-AHG-converting enzymes that have been associated with catabolism of D-AHG (20, 29, 30) were identified in strain A3^T; thus, it is proposed that strain A3^T is able to utilize both D-AHG and D-galactose. In *Cellulophaga lytica* LIM-21, *Pseudoalteromonas atlantica* T6c, *Epulopiscium* sp., *Z. galactanivorans*, and *P. hydrolytica* S66^T, the four D-AHG-converting enzymes were organized in an operon-like structure (Fig. S4A). However, in *C. echini* A3^T and another agarolytic *Colwellia* species, *C. agarivorans* QM50^T, the first two enzymes, 3,6-anhydro-D-galactose dehydrogenase and 3,6-anhydro-D-galactonate cycloisomerase, were encoded by two overlapping reading frames (in strain A3^T *Ce349* and *Ce350*), indicating that the genes were cotranscribed.

The last two enzymes, 2-keto-3-deoxy-D-galactonate kinase and 2-keto-3-deoxy-6-phospho-D-galactonate aldolase (in A3^T *Ce359* and *Ce358*), were similarly encoded by two ORFs separated by only four nucleotides, indicating that the two genes could also be cotranscribed. Transcriptomics analysis showed that all four genes were upregulated when strain A3^T was grown on carrageenan (Table S3). Thus, the net result of the reaction of enzymes Ce349, Ce350, Ce358, and Ce359 would be the conversion of D-AHG to pyruvate and D-glyceraldehyde-3-phosphate, which enter the central metabolism through the DeLey-Doudoroff pathway (Fig. 9).

DISCUSSION

The marine bacterium *C. echini* A3^T isolated from the gut of sea urchin utilizes agar and κ-carrageenan for growth (9). In this study, we performed bioinformatic, structural, and biochemical analyses of the enzymatic carrageenan degradation system in *C. echini* A3^T. Among all of the sequenced genomes of *Colwellia* species, that of *C. echini* A3^T has the greatest number of CAZymes and other enzymes involved in degradation of agars and carrageenans. Likewise, the genome of *C. agarivorans* QM50^T encodes many agarolytic enzymes, which correlates with its ability to hydrolyze agar (13). Thus, the CAZyme repertoires of these two *Colwellia* species reflect their physiological function as red alga degraders.

Algal polysaccharide-degrading bacteria are common in the marine environment: several reports previously described CAZyme-encoding *Bacteroidetes* (20, 31–34) and gammaproteobacteria (10, 29, 35). However, the genetic organizations of CAZymes and other enzymes involved in algal polysaccharide degradation may differ considerably. In *P. hydrolytica* S66^T, all genes involved in utilization of agar, furcellaran, and κ-carrageenan are localized in one large ~167-kb PUL (10). In *P. carrageenovora* 9^T and *Alteromonas* sp. 76-1, algal hydrolyzing enzymes are encoded on large plasmids (29, 35), and sometimes CAZymes are scattered on the chromosomes; e.g., agarases are scattered across four contigs in *C. agarivorans* (data not shown). Here, we show that genes necessary for degradation of different types of carrageenans are located in one ~86,000-bp region, CAR PUL, and that agarolytic genes are found on another ~92,000-bp region (AGA PUL). As discussed by Gobet et al. (29) and Schultz-Johansen et al. (10), the PUL structure of gammaproteobacteria shows a resemblance to the PULs described from *Bacteroidetes*. Putative transporters and TonB-dependent receptors, proposed to be analogous to the canonical SusC/SusD sensor/regulator system in *Bacteroidetes* PUL structures (36, 37), were found in both *C. echini* PULs, and transcriptomics showed that genes in CAR PUL were jointly regulated when *C. echini* was cultivated with κ-carrageenan as the sole carbon source. Gene activity in the AGA PUL was similarly shown to be jointly regulated (data are to be presented elsewhere).

The degradation of κ- and ι-carrageenan to neocarrabiose (DA-G) is catalyzed through the combined activities of κ- and ι-carrageenases (GH16_17, GH82, GH42-like, etc.) and of κ- and ι-specific sulfatases (S1_19) as shown for *P. hydrolytica* S66^T (10) and *Z. galactanivorans* Dsij^T (20) and as hypothesized for *P. carrageenovora* 9^T (29). Degradation of the β/κ-carrageenan furcellaran is initiated by specific GH16 endolytic enzymes that recognize only partially sulfated carrageenan. The products released from hydrolysis of furcellaran were previously proposed to be DP4-monosulfate and DP6-monosulfate (10). Here, we show that enzymes Ce385 and Ce387 from strain A3^T are such GH16_13 furcellaranases. We here show by mass spectrometry (MS) that the products are neocarratetraose-mono-sulfate and neocarrahexaose-mono-sulfate.

Neocarrabiose is a common key metabolite seen in the degradation of carrageenans in all bacteria investigated so far. However, bacteria metabolize neocarrabiose differently among different species. In *Z. galactanivorans* Dsij^T, GH127 and GH129 enzymes hydrolyze neocarrabiose to D-AHG and galactose, which then enter the central metabolism (20). Mining the genome of *P. hydrolytica* S66^T revealed a similar (60% to 62% identity) GH127-encoding gene (10), but previous searches for GH127 and GH129 enzymes in *P. carrageenovora* 9^T and *C. echini* A3^T gave negative results. This led Gobet et al. (29) to conclude that *P. carrageenovora* is unable to release 3,6-anhydro-D-

galactose from carrageenan degradation products. We observed that *C. echini* A3^T was able to grow with κ-carrageenan as the sole carbon source, and the fact that all genes necessary for degradation of κ-carrageenan to neocarrabiose and for the subsequent catabolism of D-AHG are found in CAR PUL led us to look for novel neocarrabiose-hydrolyzing enzymes in this gene cluster. Candidates for this activity could be the putative hydrolases Ce362 and Ce383 (separately or together). Both enzymes were shown to have carbohydrate binding domains by HHpred analysis, and they displayed low similarity to known GH enzymes, indicating that they may represent new GH families. BLAST and phylogenetical analyses indicated that both enzymes might be distantly related to known β-galactosidases affiliated with GH42 and GH160 families and to the GHX neocarratetraose monosulfate hydrolase described here, indicating that Ce362 and/or Ce383 might hydrolyze galactose-containing sugars such as neocarrabiose. Results of Pred-Lipo, Signal P, and Lipo P analyses indicated that Ce383 could be cytoplasmic whereas Ce362 might be anchored in the membrane. The predicted intracellular localization supports the hypothesis that Ce362 and/or Ce383 might be the missing 3,6-anhydro-D-galactosidase, as neocarrabiose is believed to be transported into the periplasm or cytoplasm in analogy with neoagarobiose (38, 39). However, functional analysis was not possible because expression of both genes in *E. coli* failed to give soluble, active enzymes. Once neocarrabiose has been transported into the cell and is hydrolyzed to D-AHG and galactose, galactose enters the general metabolism via the Leloir pathway, and D-AHG is converted to pyruvate and to D-glyceraldehyde-3-phosphate, which enters the metabolism via the DeLey-Doudoroff pathway as reported previously (10, 20, 29, 30).

MATERIALS AND METHODS

Bacterial strains and growth conditions. *C. echini* A3^T was isolated and cultivated as described by Christiansen et al. (9). *Escherichia coli* TOP10 (Novagen) was used for cloning, and *E. coli* BL21(DE3) (Novagen) and *E. coli* BL21(DE3) Δ*lac*Z, kindly provided by Jin-Ho Seo from Seoul National University, South Korea, was used for expression studies.

DNA isolation, gene cloning, and expression of recombinant enzymes. Extraction of genomic DNA, PCR with gene-specific primers (see Table S5 in the supplemental material), and cloning of genes into plasmids pET9a-USER-1 and pET9a-USER-2 were carried out as described previously (10). Expression of recombinant enzymes was carried out in ZYP-5052 autoinduction medium (40) supplemented with appropriate antibiotics, and cells were disrupted in a FastPrep-24 5G bead beater (MP Biomedicals) as described by Schultz-Johansen et al. (10). Recombinant, His-tagged enzymes were subsequently applied to a HisTrap FF column (GE Healthcare, Uppsala, Sweden) charged with 100 mM $NiSO_4$. After a washing step performed with 50 mM imidazole, the bound proteins were eluted with a linear gradient of imidazole ranging from 50 to 700 mM. Final protein concentration and purity were determined with a NanoDrop spectrophotometer (Thermo Scientific, Illkirch, France) and by SDS/PAGE, respectively. Extracts from *E. coli* cells treated with an empty vector were processed and analyzed in parallel as negative controls.

Enzyme reactions and activity visualization. Crude cell lysates were analyzed in a 3-methyl-2-benzothiazolinone hydrazine (MBTH) reducing end sugar assay (41). κ-Carrageenan, ι-carrageenan, λ-carrageenan, agar (Sigma-Aldrich), porphyran (prepared as described previously [23]), furcelleran (Est-Agar, Estonia), and agarose (Invitrogen) were used as the substrates (0.1% [wt/vol]) in a 50-μl reaction mixture with 1 μl crude cell lysate.

Thin-layer chromatography (TLC) analysis of enzyme activity was carried out in total volumes of 50 μl containing 10 μl crude cell lysate and 0.1% (wt/vol) polysaccharide. Assays performed with neocarrao-ligosaccharides (Dextra, United Kingdom) were carried out volumes of 1 μl crude enzyme and 2 μg/μl (wt/vol) of the respective oligosaccharide substrates in a total reaction volume of 12 μl. All reaction mixtures were incubated overnight. The enzyme reaction mixtures (6 μl) were spotted onto a silica gel 60 TLC plate (Merck). Plates were run twice in *n*-butanol-acetic acid-water (2:1:1 [vol/vol/vol]). Visualization of the plates was performed with 5-methylresorcinol monohydrate–5% (wt/vol) H_2SO_4, and development on a heating plate was performed for approximately 5 min.

For enzymatic reactions analyzed by fluorophore-assisted carbohydrate electrophoresis (FACE) (42), 5-μl volumes of crude cell lysates were used with a final concentration of 0.1% (wt/vol) polysaccharide. Enzyme reactions were performed with oligosaccharides as described above. Reaction mixtures were incubated overnight, and reactions were terminated at 90°C for 10 min followed by centrifugation (17,000 × g, 10 min). The supernatants were dried in a speed-vacuum centrifuge. Samples were labeled with 2 μl 8-aminonaphthalene-1,3,6-trisulfonate (ANTS) solution (0.15 M ANTS dissolved in acetic acid-water [3:17 {vol/vol}]) and with 5 μl freshly prepared 1 M sodium cyanoborohydride–dimethyl sulfoxide (DMSO) (38). After incubation overnight at 37°C, 25 μl glycerol (25%) was added and 5 to 10 μl was loaded onto a 6% stacking and 27% running polyacrylamide gel. The gel was subjected to electrophoresis at 4°C at 200V for 2 h. Hydrolysis was visualized under UV light.

DNA sequence analyses and *in silico* analysis. Automatic annotation was performed online in the Rapid Annotation using Subsystem Technology (RAST) server (43). Correlations between gene and protein names reported in this work and locus tags in GenBank are shown in Table S6. A more thorough search and identification of CAZymes and sulfatases were carried out using hidden Markov model searches as described previously by Schultz-Johansen et al. (10). Selected gene sequences from RAST were imported into CLC Main Workbench 7.0 (Qiagen) for primer design. Similarity searches and predictions of conserved domains were performed by the use of BLASTp searches against the NCBI (https://www.ncbi.nlm.nih.gov) Protein Data Bank database (PDB), Swiss-Prot, and the nonredundant protein database (nr) and the Conserved Domain Database (CDD). Prediction of signal peptides and lipoproteins was performed with Pred-Lipo (44), SignalP 4.1 (45, 46), and LipoP 1.0 (47). All alignments were made in CLC Main Workbench 7.0, and phylogenetic trees were constructed in MEGA7 (48). Sequences were retrieved from the nonredundant database using *C. echini* A3^T genes or sequences from structurally determined proteins. *In silico* folding was performed using SWISS-MODEL (https://swissmodel.expasy.org) (51), and model quality was assessed using the QMEAN server (http://swissmodel.expasy.org/qmean) (52); models with QMEANDisCo values of <0.6 were discarded (51). Additionally, local quality estimate was used and only residues with a local quality score of >0.75 are discussed. Structural models were superimposed using the "super" function in PyMOL (The PyMOL Molecular Graphics System, Version 2.0; Schrödinger, LLC).

RNA-seq analysis. *C. echini* A3^T was grown in marine minimal broth (MMB) containing 2.3% (wt/vol) aquarium sea salt mix (Instant Ocean Sea Salts; Aquarium Systems, Mentor, OH, USA), 0.1% (wt/vol) yeast extract, 0.05% (wt/vol) NH_4Cl, and 10 mM Tris-HCl buffer (pH 7.4). D-Glucose (DaeJung, South Korea) or κ-carrageenan (Tokyo Chemical Industries, Japan) (0.2% [wt/vol]) was supplemented as the sole carbon source. All experiments were carried out in triplicate. Cells for transcriptome sequencing (RNA-seq) analysis were cultivated in MMB at 20°C and 180 rpm. Overnight cultures were diluted in MMB containing either 0.2% (wt/vol) D-glucose or 0.2% (wt/vol) κ-carrageenan, and when the optical density at 600 nm (OD_{600}) reached 0.8, cells were harvested by centrifugation at 4°C. Cell pellets were resuspended in 1 ml of TRIzol reagent (Invitrogen) and incubated 2 min at room temperature for cell lysis. After complete lysis, cell lysates were mixed with 200 μl of chloroform and subjected to brief vortex mixing. Phase separation was facilitated by centrifugation at 12,500 × *g* for 5 min, and the aqueous phase was separated. A half-volume of isopropanol (Sigma-Aldrich, St. Louis, MO, USA) was added to the separated aqueous phase and loaded onto an RNeasy minicolumn (Qiagen, Inc.). The column was washed with 70% ethanol, and RNA was eluted with RNase-free water. A Qubit RNA BR assay kit (Invitrogen) was used to determine the RNA quantity. The quality of the RNA was determined using a Qsep1 Bio-fragment analyzer (BiOptic Inc.) equipped with an RNA cartridge. Extracted RNA was stored at −80°C until further use. Before the library preparation, mRNA enrichment was performed using a Ribo-Zero rRNA removal kit for bacteria (Illumina, Inc.). Enriched mRNA samples were purified using Agencourt AMPure XP beads (Beckman Coulter, Inc.). The RNA-seq library was prepared using a NEBNext Ultra RNA library preparation kit for Illumina (New England Biolabs, USA) according to the instructions of the manufacturer. Sequencing was performed using an Illumina MiSeq platform (Illumina, Inc.) and Illumina MiSeq reagent kit V3 (300 bp by 2 paired end).

Quality control and trimming of Illumina sequenced reads were performed with TrimGalore ver. 0.5.0 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Illumina universal adapter sequences and the index sequences were trimmed. Reads shorter than 40 bp were discarded, and the bases with Q values of <20 were trimmed from the 3′ and 5′ ends of reads. Reads were mapped to the *C. echini* A3^T genome (GenBank accession no. PJAI00000000.2) by the use of Bowtie2 software. SAMtools was used to convert the alignment files into BAM files. The numbers of reads mapped to the predicted coding DNA sequence (CDS) were analyzed using the Bioconductor package (49). Differential gene expression analysis was performed with the edgeR package in Bioconductor. Genes with a *P* value of <0.05 and a |logFC| value of >2 were considered to be differentially expressed (Table S3).

Qualitative analysis of reaction products by MALDI-TOF. For qualitative analysis of reaction products, MALDI-TOF MS was carried out using 2,5-dihydroxybenzoic acid (DHB) as the matrix. DHB was dissolved in 50:50 (vol/vol) acetonitrile and water containing 0.5% trifluoroacetic acid (TFA) at a concentration of 10 mg/ml. The aliquot of the enzymatic reaction products (2 μl) was placed on the target plate and mixed with DHB. The mass spectra were obtained using a Bruker UltrafleXtreme MALDI MS instrument (Bremen, Germany) equipped with a SmartBeam II laser.

Data availability. The draft genome sequence of *C. echini* A3T is available from GenBank (accession number PJAI00000000.2).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 0.3 MB.

FIG S2, PDF file, 0.1 MB.

FIG S3, PDF file, 0.1 MB.

FIG S4, PDF file, 0.1 MB.

TABLE S1, DOCX file, 0.01 MB.

TABLE S2, DOCX file, 0.01 MB.

TABLE S3, DOCX file, 0.04 MB.

TABLE S4, DOCX file, 0.01 MB.

TABLE S5, DOCX file, 0.01 MB.
TABLE S6, DOCX file, 0.01 MB.

ACKNOWLEDGMENTS

This work was supported by grants from the Novo Nordisk Foundation (grant no. NNF12OC0000797) and The Danish Council for Independent Research, Technology and Production Sciences (0602-02399B) to P.S.

REFERENCES

1. Behrenfeld MJ, Falkowski PG. 1997. Photosynthesis rates from satellite-based chlorophyll concentration. Limnol Oceanogr 42:1–20. https://doi.org/10.4319/lo.1997.42.1.0001.
2. Buchan A, LeCleir GR, Gulvik CA, González JM. 2014. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. Nat Rev Microbiol 12:686–698. https://doi.org/10.1038/nrmicro3326.
3. Kraan S. 2012. Algal polysaccharides, novel applications and outlook. *In* Chang C-F (ed), Carbohydrates - comprehensive studies on glycobiology and glycotechnology. InTechOpen, London, United Kingdom. https://doi.org/10.5772/51572.
4. Popper ZA, Michel G, Hervé C, Domozych DS, Willats WGT, Tuohy MG, Kloareg B, Stengel DB. 2011. Evolution and diversity of plant cell walls: from algae to flowering plants. Annu Rev Plant Biol 62:567–590. https://doi.org/10.1146/annurev-arplant-042110-103809.
5. Barbeyron T, Thomas F, Barbe V, Teeling H, Schenowitz C, Dossat C, Goesmann A, Leblanc C, Glöckner FO, Czjzek M, Amann R, Michel G. 2016. Habitat and taxon as driving forces of carbohydrate catabolism in marine heterotrophic bacteria: example of the model algae-associated bacterium *Zobellia galactanivorans* Dsij^T. Environ Microbiol 18:4610–4627. https://doi.org/10.1111/1462-2920.13584.
6. Hong PY, Mao Y, Ortiz-Kofoed S, Shah R, Cann I, Mackie RI. 2015. Metagenomic-based study of the phylogenetic and functional gene diversity in Galápagos land and marine iguanas. Microb Ecol 69:444–456. https://doi.org/10.1007/s00248-014-0547-6.
7. Gomare S, Kim HA, Ha JH, Lee MW, Park JM. 2011. Isolation of the polysaccharidase-producing bacteria from the gut of sea snail, *Batillus cornutus*. Korean J Chem Eng 28:1252–1259. https://doi.org/10.1007/s11814-010-0506-y.
8. Sawabe T, Oda Y, Shiomi Y, Ezura Y. 1995. Alginate degradation by bacteria isolated from the gut of sea urchins and abalones. Microb Ecol 30:193–202. https://doi.org/10.1007/BF00172574.
9. Christiansen L, Bech PK, Schultz-Johansen M, Martens HJ, Stougaard P. 2018. *Colwellia echini* sp. nov., an agar and carrageenan solubilizing bacterium isolated from sea urchin. Int J Syst Evol Microbiol 68:687–691. https://doi.org/10.1099/ijsem.0.002568.
10. Schultz-Johansen M, Bech PK, Hennessy RC, Glaring MA, Barbeyron T, Czjzek M, Stougaard P. 3 May 2018, posting date. A novel enzyme portfolio for red algal polysaccharide degradation in the marine bacterium *Paraglaciecola hydrolytica* S66T encoded in a sizeable polysaccharide utilization locus. Front Microbiol https://doi.org/10.3389/fmicb.2018.00839.
11. Viborg AH, Terrapon N, Lombard V, Michel G, Czjzek M, Henrissat B, Brumer H. 9 September 2019, posting date. A subfamily roadmap for functional glycogenomics of the evolutionarily diverse glycoside hydrolase family 16 (GH16). J Biol Chem https://doi.org/10.1074/jbc.RA119.010619.
12. CAZypedia Consortium. 2018. Ten years of CAZypedia: a living encyclopedia of carbohydrate-active enzymes. Glycobiology 28:3–8. https://doi.org/10.1093/glycob/cwx089.
13. Xu ZX, Zhang HX, Han JR, Dunlap CA, Rooney AP, Mu DS, Du ZJ. 2017. *Colwellia agarivorans* sp. nov., an agar-digesting marine bacterium isolated from coastal seawater. Int J Syst Evol Microbiol 67:1969–1974. https://doi.org/10.1099/ijsem.0.001897.
14. Hehemann JH, Correc G, Thomas F, Bernard T, Barbeyron T, Jam M, Helbert W, Michel G, Czjzek M. 2012. Biochemical and structural characterization of the complex agarolytic enzyme system from the marine bacterium *Zobellia galactanivorans*. J Biol Chem 287:30571–30584. https://doi.org/10.1074/jbc.M112.377184.
15. Potin P, Richard C, Rochas C, Kloareg B. 1993. Purification and characterization of the α-agarase from *Alteromonas agarlyticus* (Cataldi) comb. nov., strain GJ1B. Eur J Biochem 214:599–607. https://doi.org/10.1111/j.1432-1033.1993.tb17959.x.
16. Ohta Y, Hatada Y, Miyazaki M, Nogi Y, Ito S, Horikoshi K. 2005. Purification and characterization of a novel alpha-agarase from a *Thalassomonas* sp. Curr Microbiol 50:212–216. https://doi.org/10.1007/s00284-004-4435-z.
17. Flament D, Barbeyron T, Jam M, Potin P, Czjzek M, Kloareg B, Michel G. 2007. Alpha-agarases define a new family of glycoside hydrolases, distinct from beta-agarase families. Appl Environ Microbiol 73:4691–4694. https://doi.org/10.1128/AEM.00496-07.
18. Zhang W, Xu J, Liu D, Liu H, Lu X, Yu W. 2018. Characterization of an α-agarase from *Thalassomonas* sp. LD5 and its hydrolysate. Appl Microbiol Biotechnol 102:2203–2212. https://doi.org/10.1007/s00253-018-8762-6.
19. Grondin JM, Tamura K, Déjean G, Abbott DW, Brumer H. 11 July 2017, posting date. Polysaccharide utilization loci: fueling microbial communities. J Bacteriol https://doi.org/10.1128/JB.00860-16.
20. Ficko-Blean E, Préchoux A, Thomas F, Rochat T, Larocque R, Zhu Y, Stam M, Génicot S, Jam M, Calteau A, Viart B, Ropartz D, Pérez-Pascual D, Correc G, Matard-Mann M, Stubbs KA, Rogniaux H, Jeudy A, Barbeyron T, Médigue C, Czjzek M, Vallenet D, McBride MJ, Duchaud E, Michel G. 2017. Carrageenan catabolism is encoded by a complex regulon in marine heterotrophic bacteria. Nat Commun 8:1685. https://doi.org/10.1038/s41467-017-01832-6.
21. Yang B, Yu G, Zhao X, Ren W, Jiao G, Fang L, Wang Y, Du G, Tiller C, Girouard G, Barrow CJ, Ewart HS, Zhang J. 2011. Structural characterisation and bioactivities of hybrid carrageenan-like sulphated galactan from red alga *Furcellaria lumbricalis*. Food Chem 124:50–57. https://doi.org/10.1016/j.foodchem.2010.05.102.
22. Lv Y, Yang B, Zhao X, Zhang J, Yu G. 2015. Structural characterization of a hybrid carrageenan-like sulfated galactan from a marine red alga *Furcellaria lumbricalis*. Methods Mol Biol 1308:325–346. https://doi.org/10.1007/978-1-4939-2684-8_21.
23. Hehemann JH, Correc G, Barbeyron T, Helbert W, Czjzek M, Michel G. 2010. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. Nature 464:908–912. https://doi.org/10.1038/nature08937.
24. Matard-Mann M, Bernard T, Leroux C, Barbeyron T, Larocque R, Préchoux A, Jeudy A, Jam M, Nyvall Collén P, Michel G, Czjzek M. 2017. Structural insights into marine carbohydrate degradation by family GH16 κ-carrageenases. J Biol Chem 292:19919–19934. https://doi.org/10.1074/jbc.M117.808279.
25. Michel G, Chantalat L, Duee E, Barbeyron T, Henrissat B, Kloareg B, Dideberg O. 2001. 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. https://doi.org/10.1016/s0969-2126(01)00612-8.
26. Biarnés X, Nieto J, Planas A, Rovira C. 2006. Substrate distortion in the Michaelis complex of *Bacillus* 1,3-1,4-beta-glucanase. Insight from first principles molecular dynamics simulations. J Biol Chem 281:1432–1441. https://doi.org/10.1074/jbc.M507643200.
27. McLean MW, Williamson FB. 1979. kappa-Carrageenase from *Pseudomonas carrageenovora*. Eur J Biochem 93:553–558. https://doi.org/10.1111/j.1432-1033.1979.tb12854.x.
28. Barbeyron T, Brillet-Guéguen L, Carré W, Carrière C, Caron C, Czjzek M, Hoebeke M, Michel G. 2016. Matching the diversity of sulfated biomolecules: creation of a classification database for sulfatases reflecting their substrate specificity. PLoS One 11:e0164846. https://doi.org/10.1371/journal.pone.0164846.
29. Gobet A, Barbeyron T, Matard-Mann M, Magdelenat G, Vallenet D, Duchaud E, Michel G. 2018. Evolutionary evidence of algal polysaccha-

ride degradation acquisition by *Pseudoalteromonas carrageenovora* 9T to adapt to macroalgal niches. Front Microbiol 9:2740. https://doi.org/10.3389/fmicb.2018.02740.
30. Lee SB, Kim JA, Lim HS. 2016. Metabolic pathway of 3,6-anhydro-D-galactose in carrageenan-degrading microorganisms. Appl Microbiol Biotechnol 100:4109–4121. https://doi.org/10.1007/s00253-016-7346-6.
31. Bauer M, Kube M, Teeling H, Richter M, Lombardot T, Allers E, Würdemann CA, Quast C, Kuhl H, Knaust F, Woebken D, Bischof K, Mussmann M, Choudhuri JV, Meyer F, Reinhardt R, Amann RI, Glöckner FO. 2006. Whole genome analysis of the marine *Bacteroidetes* '*Gramella forsetii*' reveals adaptations to degradation of polymeric organic matter. Environ Microbiol 8:2201–2213. https://doi.org/10.1111/j.1462-2920.2006.01152.x.
32. Kabisch A, Otto A, König S, Becher D, Albrecht D, Schüler M, Teeling H, Amann RI, Schweder T. 2014. Functional characterization of polysaccharide utilization loci in the marine *Bacteroidetes* '*Gramella forsetii*' KT0803. ISME J 8:1492–1502. https://doi.org/10.1038/ismej.2014.4.
33. Tang K, Lin Y, Han Y, Jiao N. 2017. Characterization of potential polysaccharide utilization systems in the marine bacteroidetes Gramella flava JLT2011 using a multi-omics approach. Front Microbiol 8:220. https://doi.org/10.3389/fmicb.2017.00220.
34. Valdehuesa KNG, Ramos KRM, Moron LS, Lee I, Nisola GM, Lee WK, Chung WJ. 2018. Draft genome sequence of newly isolated agarolytic bacteria *Cellulophaga omnivescoria* sp. nov. W5C carrying several gene loci for marine polysaccharide degradation. Curr Microbiol 75:925–933. https://doi.org/10.1007/s00284-018-1467-3.
35. Koch H, Freese HM, Hahnke RL, Simon M, Wietz M. 2019. Adaptations of *Alteromonas* sp. 76–1 to polysaccharide degradation: a CAZyme plasmid for ulvan degradation and two alginolytic systems. Front Microbiol 10:504. https://doi.org/10.3389/fmicb.2019.00504.
36. Bjursell MK, Martens EC, Gordon JI. 2006. 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. https://doi.org/10.1074/jbc.M606509200.
37. Martens EC, Koropatkin NM, Smith TJ, Gordon JI. 2009. Complex glycan catabolism by the human gut microbiota: the *Bacteroidetes* Sus-like paradigm. J Biol Chem 284:24673–24677. https://doi.org/10.1074/jbc.R109.022848.
38. Pluvinage B, Grondin JM, Amundsen C, Klassen L, Moote PE, Xiao Y, Thomas D, Pudlo NA, Anele A, Martens EC, Inglis GD, Uwiera RER, Boraston AB, Abbott DW. 2018. Molecular basis of an agarose metabolic pathway acquired by a human intestinal symbiont. Nat Commun 9:1043. https://doi.org/10.1038/s41467-018-03366-x.
39. Hehemann JH, Smyth L, Yadav A, Vocadlo DJ, Boraston AB. 2012. Analysis of keystone enzyme in agar hydrolysis provides insight into the degradation of a polysaccharide from red seaweeds. J Biol Chem 287:13985–13995. https://doi.org/10.1074/jbc.M112.345645.
40. Studier FW. 2005. Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41:207–234. https://doi.org/10.1016/j.pep.2005.01.016.
41. Anthon GE, Barrett DM. 2002. Determination of reducing sugars with 3-methyl-2-benzothiazolinonehydrazone. Anal Biochem 305:287–289. https://doi.org/10.1006/abio.2002.5644.
42. Starr CM, Masada RI, Hague C, Skop E, Klock JC. 1996. Fluorophore-assisted carbohydrate electrophoresis in the separation, analysis, and sequencing of carbohydrates. J Chromatogr A 720:295–321. https://doi.org/10.1016/0021-9673(95)00749-0.
43. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.
44. Bagos PG, Tsirigos KD, Liakopoulos TD, Hamodrakas SJ. 2008. Prediction of lipoprotein signal peptides in Gram-positive bacteria with a hidden Markov model. J Proteome Res 7:5082–5093. https://doi.org/10.1021/pr800162c.
45. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–786. https://doi.org/10.1038/nmeth.1701.
46. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol 37:420–423. https://doi.org/10.1038/s41587-019-0036-z.
47. Juncker AS, Willenbrock H, von Heijne G, Nielsen H, Brunak S, Krogh A. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci 12:1652–1662. https://doi.org/10.1110/ps.0303703.
48. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054.
49. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T, Gottardo R, Hahne F, Hansen KD, Irizarry RA, Lawrence M, Love MI, MacDonald J, Obenchain V, Oleś AK, Pagès H, Reyes A, Shannon P, Smyth GK, Tenenbaum D, Waldron L, Morgan M. 2015. Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods 12:115–121. https://doi.org/10.1038/nmeth.3252.
50. Viborg AH, Katayama T, Abou Hachem M, Andersen MC, Nishimoto M, Clausen MH, Urashima T, Svensson B, Kitaoka M. 2014. Distinct substrate specificities of three glycoside hydrolase family 42 β-galactosidases from *Bifidobacterium longum* subsp. infantis ATCC 15697. Glycobiology 24:208–216. https://doi.org/10.1093/glycob/cwt104.
51. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46:W296–W303. https://doi.org/10.1093/nar/gky427.
52. Benkert P, Biasini M, Schwede T. 2011. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27:343–350. https://doi.org/10.1093/bioinformatics/btq662.