



HAL
open science

Identification and Characterization of a Halotolerant, Cold-Active Marine Endo- β -1,4-Glucanase by Using Functional Metagenomics of Seaweed-Associated Microbiota

Marjolaine Martin, Sophie Biver, Sébastien Steels, Tristan Barbeyron, Murielle Jam, Daniel Portetelle, Gurvan Michel, Micheline Vandenbol

► To cite this version:

Marjolaine Martin, Sophie Biver, Sébastien Steels, Tristan Barbeyron, Murielle Jam, et al.. Identification and Characterization of a Halotolerant, Cold-Active Marine Endo- β -1,4-Glucanase by Using Functional Metagenomics of Seaweed-Associated Microbiota. *Applied and Environmental Microbiology*, 2014, 80 (16), pp.4958-4967. 10.1128/AEM.01194-14 . hal-02138133

HAL Id: hal-02138133

<https://hal.science/hal-02138133>

Submitted on 23 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Functional screening of a metagenomic library of seaweed-associated**
2 **microbiota: identification and characterization of a halotolerant, cold-**
3 **active marine endo- β -1,4-endoglucanase**

4

5 Marjolaine Martin^{a#}, Sophie Biver^a, Sébastien Steels^a, Tristan Barbeyron^{b,c}, Murielle
6 Jam^{b,c}, Daniel Portetelle^a, Gurvan Michel^{b,c}, Micheline Vandenberg^a

7 ^a Microbiology and Genomics Unit, Gembloux Agro-Bio Tech, University of Liège,
8 Belgium

9 ^b Sorbonne Universités, UPMC Univ Paris 06, UMR 8227, Integrative Biology of
10 Marine Models, Station Biologique de Roscoff, Roscoff cedex, Bretagne, France

11 ^c CNRS, UMR 8227, Integrative Biology of Marine Models, Station Biologique de
12 Roscoff, Roscoff cedex, Bretagne, France

13

14 Running head: Functional screening of an algal biofilms metagenomic library

15

16 #Address correspondence to marjolaine.martin@ulg.ac.be

17 **Abstract**

18 A metagenomic library was constructed from microorganisms associated with the
19 brown alga *Ascophyllum nodosum*. Functional screening of this library revealed
20 thirteen novel putative esterase genes and two glycoside hydrolase genes. Sequence
21 and gene cluster analysis showed the wide diversity of the identified enzymes and
22 gave an idea of the microbial populations present during the sample collection period.
23 Lastly, an endo- β -1,4-endoglucanase having less than 50% identity to sequences of
24 known cellulases was purified and partially characterized, showing activity at low
25 temperature and after prolonged incubation in concentrated salt solutions.

26

27 **Keywords**

28 Microbial enzymes; *Ascophyllum nodosum*; metagenomics; plates tests; biomass
29 hydrolysis ; cellulase ; beta-glucosidase; esterase; functional screening, GH5

30

31

32 **Introduction**

33 Previous surveys have revealed that less than 1% of existing microorganisms can be
34 studied by traditional culturing methods. This leaves most microorganisms and their
35 by-products unknown and unexploited (1, 2) and explains why metagenomics, a
36 culture-independent approach using total microbial genomes from environment
37 samples, has met with great success over the past decade (3, 4). Sequence-based
38 metagenomics has already provided information about the composition, organization,
39 functioning, and hierarchy of microbial communities in particular habitats (5). On the
40 other hand, functional genomics applied to metagenomic libraries from diverse
41 environments has led to the discovery of many new enzymes and bioactive
42 compounds and to substantial enrichment of genome databases. To date, most of the
43 enzymes brought to light through metagenomics have been derived from soil (6–9)
44 and gut (10–13) samples.

45 Marine microorganisms represent great candidate sources of original biocatalysts, as
46 they are exposed to extreme conditions of temperature, pressure, salinity, nutrient
47 availability, etc. Hence, functional screening of marine microbial populations should
48 yield new enzymes with specific and unique physiological and biochemical
49 properties, produced by organisms far different from those usually described in
50 terrestrial environments (14). New enzymes have already been identified in marine
51 metagenomic libraries from seawater samples (15, 16) and from microorganisms in
52 symbiosis with marine organisms such as sponges and corals (17). To our knowledge,
53 however, nobody has yet performed functional screening of metagenomic libraries
54 from algal microbial communities. As algal microbial biofilms are in constant
55 interaction with algal biomass, they represent an interesting source of enzymes and
56 other active compounds (18). Sequence-based studies have already revealed the

57 importance and functions of microbial communities living on the surfaces of algae,
58 showing tight interdependence between algae and their biofilms (19–21).
59 Furthermore, many genes coding for enzymes involved in hydrolyzing algal biomass,
60 such as cellulases, xylanases, α -amylases, and specific algal polysaccharidases (e.g.
61 agarases, carrageenases, and alginate lyases) have been identified from cultivable
62 marine bacteria (22, 23). Only a few such enzymes have been found again in marine
63 metagenomes by functional screening. For example, no cellulase gene has yet been
64 identified by marine metagenomics, although cellulase genes have been identified in
65 the genomes of cultivable marine bacteria such as *Caldocellum saccharolyticum* (24),
66 *Paenibacillus* sp. BME14 (25), *Martelella mediterranea* (26), and *Marinobacter* sp.
67 MS1032 (27). Cellulases are widely used in biotechnological applications (e. g. in the
68 pulp and paper, textile, and food industries) and their capacity to convert cellulosic
69 biomass is increasingly exploited in sustainable applications and bioprocesses (28).
70 Although several cellulases are already in use, there is an increasing demand for
71 cellulases with better or special industrial properties (29). Here we describe the
72 construction and functional screening of a metagenomic library from microorganisms
73 living on the annual brown alga *Ascophyllum nodosum*. The library was screened for
74 particular hydrolytic enzymes. The screen has yielded genes encoding new esterases
75 belonging to diverse lipolytic enzyme families, a novel beta-glucosidase, and a novel
76 cellulase. These discoveries highlight the utility of seeking original biocatalysts in
77 algal biofilm populations. We have purified and characterized the identified new
78 cellulase. To our knowledge, no functional screening of a metagenomic library from
79 alga-associated microbial biofilms has been realized before, and no cellulase has
80 already been identified by functional metagenomics in a marine environment.

81 **Material and methods**

82 **Environmental Sample Collection**

83 Several plants (~ 100g) of *Ascophyllum nodosum* were collected from the foreshore
84 (48° 43' 36.07" SE, -3° 59' 22.96") in Roscoff (Brittany, France), at the start of
85 February 2012. Samples were collected in triplicate and put in sterile plastic bags. The
86 algae were rinsed three times with sterile seawater to remove loosely associated
87 microorganisms and cut into smaller sections. Washed and cut algae were freeze-dried
88 at -80°C in new sterile plastic bags until DNA extraction.

89 **Microbial Genomic DNA Extraction**

90 Bacterial DNA was extracted by swabbing the algal surfaces with sterile cotton tips
91 followed by classic DNA extraction (30). Six cotton tips were used per 25 g of algae.
92 Microbial DNA was extracted from the cotton tips with 6.75 ml lysis buffer (100 mM
93 Tris-HCl; 0.5 M EDTA; 100 mM sodium phosphate; 5 M NaCl; 10% CTAB) and
94 incubated with 50 µl proteinase K (10 mg/ml) at 37°C. After one hour, 750 µl of 20%
95 SDS was added and the cotton tips were incubated for 2 h at 65°C, with manual
96 inversion every 20 minutes. After centrifugation, 1 volume of chloroform-isoamyl
97 alcohol (v:v, 24:1) was added to the supernatant and the mixture centrifuged. The
98 aqueous phase was recovered and DNA was precipitated with 0.6 vol isopropanol for
99 one hour at room temperature. DNA was recovered by centrifugation (30 min –
100 10.000g) at 4°C, washed with 70% ethanol, and the pellet was air dried and
101 resuspended in 100 µl Tris-EDTA. The DNA extracted from triplicate samples was
102 pooled for DNA library construction. The quantity of DNA was measured with a
103 Qubit® (Invitrogen) and its quality was checked on a 0.8% agarose gel.

104 **Library Construction and functional screening**

105 The pool of DNA obtained from triplicate samples was partially cut with DpnII (New
106 England Biolabs). A PEG/NaCl size-selective precipitation was done according to

107 Biver et al. (6) in order to remove small DNA inserts. An equal volume of 6.5%
108 PEG/0.4M NaCl was added to the restriction fragments (final DNA concentration: 20
109 ng/μl) and the mixture was incubated for 2 h at RT. The DNA was purified by
110 migration through a 1% low-melting-point agarose gel (Promega) and DNA inserts
111 between 1.5 and 7 kb were recovered by AgarACE digestion (Promega). The inserts
112 were ligated into the BamHI (Roche) linearized and dephosphorylated (Roche) shuttle
113 vector YEp356 (Novagen) with T4 DNA ligase (Roche). Electrompetent Electromax
114 *E. coli* cells (Life Technologies) were finally transformed with the ligation product.
115 The average DNA insert size was estimated by isolation and purification of 20
116 randomly chosen plasmids. The transformants were pooled and the metagenomic
117 library was plated and screened for lipolytic activity on 3% tributyrin 2xYT agar
118 plates containing 100 μg/ml ampicillin. Beta-glucosidases were sought on 2xYT agar
119 containing 0.5% esculin (Sigma—Aldrich) and 0.1% ammonium iron (III) citrate
120 (Sigma—Aldrich) (12). Cellulolytic activity was detected by plating the library on
121 2xYT agar plates containing 1 g/l insoluble synthetic AZCL-HE-cellulose
122 (Megazyme) and 100μg/ml ampicillin. A blue halo was observed around colonies
123 exhibiting cellulase activity. The plasmids of positive clones were extracted and used
124 for activity confirmation in DH5α *E. coli* subcloning cells (Life Technologies). Insert
125 sizes were estimated with restriction enzymes cutting at both extremities of the DNA
126 insert (*HindIII* and *EcoRI*, Roche).

127 **DNA sequencing and sequence analysis**

128 DNA inserts of positive clones were sequenced at GATC Biotech (Germany).
129 Sequence similarity searches were carried out with the NCBI BlastP program.
130 Phylogenetic trees were produced by the neighbor-joining method with MEGA6.06.
131 Signal peptides were predicted with the SignalP 4.1 server (31). All putative proteins

132 corresponding to the identified ORFs in the inserts were analyzed with BlastP against
133 the GenBank non-redundant database and Swissprot (curated database). The putative
134 esterases were also analyzed with Pfam (32).

135 **Overexpression and purification of the CellMM5.1 cellulase**

136 The coding sequence of the *Cell5.1_3* gene was amplified, without its putative peptide
137 signal, with Platinum *Pfx* DNA Polymerase, 0.5x PCRx Enhancer Solution
138 (Invitrogen), 5'- gggaattcCATATGCAAACCTCTGACAGCTATGTTCCGC -3' as
139 forward primer, and 5'-

140 cgcGGATCCTCATCAGTGGTGGTGGTGGTGGTGGCGCCCAAGCAAGGCGT
141 -3' as reverse primer. The restriction sites (NdeI, BamHI) used for subcloning are
142 underlined and the histidine tag sequence appears in bold. PCR products were cloned
143 into the pET30b(+) vector (Novagen) and the resulting plasmids were introduced into
144 *E. coli* Rosetta 2 (DE3) cells (Novagen). The cellulolytic activity of the obtained
145 transformants was evaluated by spotting 5 µl overnight culture (grown at 37°C in
146 liquid 2xYT medium, 30 µg/ml chloramphenicol and 30 µg/ml kanamycin) onto solid
147 2xYT medium supplemented with 10 µM isopropyl β-D-1-thiogalactopyranoside
148 (IPTG) (Fermentas), 1g/l AZCL-HE-cellulose (Megazyme), and appropriate
149 concentrations of antibiotics (30 µg/ml chloramphenicol and 30 µg/ml kanamycin).
150 The bacteria were allowed to grow at 37°C for 3 days. To purify the recombinant
151 cellulase, an overnight culture of transformed *E. coli* Rosetta cells was diluted to
152 OD_{600nm}=0.005 in a 1-l shake flask containing 200 ml 2xYT medium and the
153 corresponding antibiotics. The cultures were grown at 37°C under shaking at 160 rpm
154 until the OD_{600nm} reached about 0.4. Overproduction was induced by addition of 1 µM
155 IPTG followed by incubation at 16°C for ~20 h. The cells were harvested by
156 centrifugation and lysed at 37°C for 30 min in 17 ml lysis buffer (20 mM Tris-HCl

157 buffer pH 8 containing 0.3 M NaCl, 1 ml/l Triton X-100, and 5 mg/ml lysozyme)
158 according to Biver et al (33). After sonication on ice (4 cycles of 5 s, 100 W, Sonifier
159 250, Branson), the lysate was centrifuged at 10,000g for 20 min (4°C) to separate
160 insoluble (I0) from soluble (S0) proteins. To solubilize the recombinant cellulase
161 found mostly in the resulting pellet (I0), the latter was incubated at 4°C for 4 h in 5 ml
162 solubilization buffer (20 mM Tris-HCl pH 8, 0.5 M NaCl, 20 ml/l Triton X-100, 2 M
163 urea). The partly resuspended fraction was then centrifuged again to collect the
164 soluble fraction (S1) now containing a sufficient amount of His-tagged cellulase,
165 which was next purified by Ni-NTA affinity chromatography (Qiagen). The soluble
166 fraction was diluted 10 times in 20 mM Tris-HCl pH 8 and 0.5 M NaCl to reduce the
167 urea and Triton X-100 concentrations to 0.2 M and 2 ml/l, respectively, and then
168 incubated overnight at 4°C with 200 µl Ni-NTA agarose pre-equilibrated with binding
169 buffer (20 mM Tris-HCl pH 8, 0.5 M NaCl, 2 ml/l Triton X-100, 0.2 M urea). The
170 resin was recovered by centrifugation at 800 g for 5 min and washed twice with 1 ml
171 washing buffer 1 (20 mM Tris-HCl pH 7.5, 0.5 M NaCl, 1 ml/l Triton X-100, 10 mM
172 imidazole) and twice with 1 ml washing buffer 2 (20 mM Tris-HCl pH 7.5, 0.5 M
173 NaCl, 1 ml/l Triton X-100, 20 mM imidazole) before elution with 2 ml of 20 mM
174 Tris-HCl pH 7.5, 0.5 M NaCl, 250 mM imidazole. After analysis by SDS-PAGE, this
175 first eluate (E1) was repurified by Ni-NTA chromatography (E2) and then dialyzed at
176 4°C against 50 mM sodium phosphate pH 6.8 containing 50 mM NaCl.

177 **Cell5.1_3 cellulase activity characterization**

178 Enzymatic activity was quantified by measuring the quantity of generated reducing
179 sugars with the 3',5'-dinitrosalicylic acid reagent (DNSA, Sigma) (34). The
180 enzymatic assay was performed with 1% (w/v) carboxymethylcellulose (CMC) (Ultra
181 low viscosity, Fluka), 50 mM phosphate buffer pH 7 for 10 minutes at 40°C in a total

182 volume of 390 μ l, to which 10 μ l suitably diluted enzyme was added. Then an equal
183 volume (400 μ l) of DNSA was added and the mixture heated for 10 minutes at 99°C.
184 After cooling on ice, the absorbance was read at 540 nm (Multiskan Go,
185 ThermoScientific). The concentration of reducing sugar was determined against a
186 glucose standard. Cell5.1_3 activity was measured in μ moles of glucose released per
187 minute. All assays were performed in triplicate. The temperature range tested was
188 from 0°C to 70°C. The tested pH ranges were from pH 4 to pH 6 (in sodium acetate
189 buffer), from pH 6 to pH 8 (in sodium phosphate buffer), from pH 8 to pH 9 (in
190 Tris/HCl buffer), and from pH 8 to pH 10 (in glycine/NaOH buffer). The temperature
191 and pH stability were estimated, respectively, by incubating the enzyme for 1 hour at
192 the appropriate temperature or for 24 h at the appropriate pH (in the appropriate
193 buffer). Residual activities were then measured under standard conditions.
194 Halotolerance was determined by measuring the residual activity after 1.5, 5, and 24 h
195 of incubation in 3 M KCl or 4 M NaCl. Substrate specificity was determined by
196 replacing CMC in the standard assay with 1% (w/v) lichenan (1-3,1-4- β -D-glucan,
197 Megazyme), laminarin (Sigma), xyloglucan (Megazyme), mannan (1,4- β -D-mannan,
198 Megazyme), or crystalline cellulose (Avicel, Sigma). For the insoluble substrates
199 mannan and Avicel, the reaction time was extended to 5 h at room temperature, with
200 automatic inversion. Effects of metal ions on the activity of the Cell5.1_3 cellulase
201 were determined under standard conditions with the metal ion of interest at 1 mM
202 final concentration. Effects of the chemicals EDTA (10 mM), SDS (1%), ethanol
203 (10%), Triton (5%), glycerol (15%), and DMSO (15%) were also tested under
204 standard conditions. Kinetic parameters (K_m and k_{cat}) were determined from a
205 Lineweaver-Burk plot obtained by varying the concentration of CMC (0.5 – 4%) in
206 assays performed under standard conditions.

207 **Accession numbers**

208 The DNA insert sequences of *LipMM5.1*, *LipMM5.3*, *LipMM5.5*, *LipMM5.7*,
209 *LipMM5.8*, *LipMM5.11*, *LipMM5.12*, *LipMM5.13*, *LipMM5.14*, *LipMM5.15*,
210 *LipMM5.19*, *CellMM5.1*, and *BglucMM5.1* have been submitted to GenBank under
211 accession numbers KF726988, KF726989, KF726990, KF726991, KF726992,
212 KF726993, KF726994, KF726995, KF726996, KF726997, KF726998, KF726999,
213 and KJ499810, respectively (see also TABLE 1).

214 **Results**

215 **DNA extraction, library construction, and functional screening**

216 Bacterial DNA was first extracted according to Burke et al. (35) with a tri-enzymatic
217 cleaner (ANIOS PlaII, France). As this procedure resulted in co-extraction of
218 polyphenols, present in brown algal cell walls (36), the DNA could not be purified
219 (data not shown). Therefore, a method described by Lachnit et al. (30), consisting of
220 swabbing the algal surfaces with sterile cotton tips followed by classic DNA
221 extraction, was tried. DNA quality was checked by migration through an agarose gel.
222 High-molecular-weight DNA, without any visible degradation smear, was observed
223 (data not shown). An average of 0.11 µg DNA per gram of *Ascophyllum nodosum* was
224 recovered with this method. As 75 g of algae were swabbed, approximately 8 µg of
225 microbial DNA was recovered and used to construct the metagenomic library. The
226 DNA was restricted and small DNA fragments were removed by PEG/NaCl
227 precipitation followed by agarose purification. DNA fragments with a size between 3
228 and 10 kb were ligated into the YEp356 cloning vector and the resulting recombinant
229 plasmids were inserted into electrocompetent DH10b *E. coli* cells. The generated
230 metagenomic library was plated, resulting in approximately 40,000 clones (180 Mb)
231 with an average DNA insert size of 4.5 kb (estimated on 20 randomly chosen

232 recombinant plasmids). The recombinant clones were pooled and screening was done
233 by plating the library several times on 2xYT agar plates containing tributyrin,
234 esculetin, or AZCL-HE-cellulose (for detection of lipase (clear halo), beta-glucosidase
235 (brown precipitate), or cellulase (blue halo) activity, respectively). The screens
236 yielded respectively 20, 1, and 1 candidate(s). Activity was confirmed by extraction of
237 their recombinant plasmids and transformation of DH5 α *E. coli* cells. The novel
238 recombinant clones were then plated again on selective media: from the 20 candidates
239 showing lipase activity, 11 plasmids were found to confer lipolytic activity
240 (*LipMM5.1*, *LipMM5.3*, *LipMM5.5*, *LipMM5.7*, *LipMM5.8*, *LipMM5.11*, *LipMM5.12*,
241 *LipMM5.13*, *LipMM5.14*, *LipMM5.15*, *LipMM5.19*), and the plasmids assumed to
242 carry a cellulase gene (*CellMM5.1*) or a beta-glucosidase gene (*BglucMM5.1*)
243 accordingly conferred the corresponding activity.

244 **DNA Sequence analysis**

245 The DNA inserts of the positive clones were sequenced. ORFs on the DNA inserts
246 were sought with ORF Finder. On the DNA insert of each positive candidate, 3 to 7
247 ORFs were indentified. Each ORF gene product was analyzed with BlastP against the
248 GenBank non-redundant database and Swissprot (a curated database). The putative
249 esterases were also analyzed with Pfam (32) (Supplementary Data, DATASET S1).
250 In each cluster, we identified at least one ORF whose predicted product shows
251 sequence similarity to the amino acid sequence of a known esterase, beta-glucosidase,
252 or cellulase (shown in TABLE 1). Putative esterase genes were assigned to four
253 carboxyesterase families by amino acid sequence alignment (data not shown).
254 Lip5.8_1 shows sequence similarity to members of bacterial esterase family II (37,
255 38). Lip5.1_5, Lip5.3_3, Lip5.5_4, Lip5.13_3, Lip5.14_3, and Lip5.19_3 belong to
256 family IV, Lip5.13_1 to family V, and Lip5.11_3 and Lip5.15_3 to family VIII. The

257 gene products of *Lip5.7_2* and *Lip5.12_2* could not be assigned to any esterase family
258 described by Arpigny (37) nor to one of the novel families EstA (39) and LipG (16).
259 However, these gene products could be assigned to the Pfam alpha/beta-hydrolase
260 family (subfamilies 6 and 5, respectively), which contains numerous esterases. In fact,
261 the esterase families IV and V correspond to alpha/beta-hydrolase subfamilies 3 and
262 6, respectively (32).

263 On the DNA insert of the beta-glucosidase-positive clone, the following were
264 identified: an incomplete ORF encoding a putative glycoside hydrolase (GH), an
265 incomplete ORF encoding a putative xylulokinase, and a complete ORF encoding a
266 putative esterase (GDSL). The incomplete glycoside hydrolase gene (*Bgluc5.1_1*) was
267 probably responsible for the beta-glucosidase activity and was assigned to family
268 GH3 of the CAZyme database (40). The esterase gene (*Bgluc5.1_2*) encodes a
269 putative GDSL lipolytic protein assigned to family II by amino acid sequence
270 alignment (37, 38).

271 On the *CellMM5.1* candidate, we identified an ORF (*Cell5.1_3*) of 359 amino acids
272 showing 42% sequence identity to a glycoside hydrolase of *Hirschia baltica* ATCC
273 49814 (GenBank: YP_003060354.1). A GH5-family protein signature was also
274 identified. Phylogenetic analysis showed that the enzymes most similar to the
275 *Cell5.1_3* putative cellulase are subfamily-GH5_25 enzymes of proteobacteria such as
276 *Hirschia baltica*, *Caulobacter sp.*, and *Ramlibacter sp.* (41) (FIG 1). A signal peptide
277 was also predicted. As no other cellulase has been identified to date by functional
278 metagenomics in a marine environment, we chose to purify and characterize the
279 putative cellulase gene (*Cell5.1_3*) product.

280 It is difficult to attribute a specific biological function to the different gene clusters on
281 the basis of gene annotations, especially for the clusters containing esterases. In

282 contrast, curated gene annotations give insights into the taxonomic affiliation of the
283 *Ascophyllum*-associated bacteria, at least at the level of the class or order. In most of
284 the clusters, the different genes have a consistent taxonomic affiliation (based on the
285 bacterial origin of the closest homologous proteins). The *LipMM5.1* insert is an
286 exception: its affiliation is ambiguous, with homologous genes found among the
287 *Firmicutes*, *Actinobacteria*, and *Gammaproteobacteria*. The *LipMM5.7* and
288 *LipMM5.13* inserts seem to come from representatives of the *Betaproteobacteria* and
289 *Gammaproteobacteria*, respectively. The remaining inserts originate from
290 *Alphaproteobacteria* members, mainly of the order *Rhodobacterales*. Interestingly,
291 several inserts seem to belong to bacteria close to *Robiginitomaculum antarcticum*, a
292 marine *Rhodobacterales* species isolated from Antarctic seawater (42).

293 **Cellulase purification and activity characterization**

294 The recombinant Cell5.1_3 cellulase, without its signal peptide, was produced in *E.*
295 *coli* Rosetta 2 (DE3) from the pET30b vector and purified by Ni-chromatography.
296 SDS-PAGE of the eluted fractions showed a clear band between 30 and 40 kDa
297 (Figure 2), in accordance with the predicted molecular mass (38 kDa). The activity of
298 the Cell5.1_3 enzyme was tested by the DNSA method on CMC. Its optimal
299 temperature was found to be about 40°C, although the enzyme proved unstable after a
300 1-hour incubation at this temperature (FIG 3A). At 40°C, the optimal pH was pH 7,
301 and the cellulase appeared stable over a wide pH range when pre-incubated for 24 h at
302 4°C in different buffers (FIG 3C, D). The kinetic parameters determined for the
303 reaction with CMC, at optimal pH and temperature, were $K_m = 8.37 \mu\text{M}$ and $k_{cat} =$
304 10.35 s^{-1} . Halotolerance was also observed: the activity of the enzyme remained
305 unchanged, under standard conditions, after a 24-hour pre-incubation with 4 M NaCl
306 or 3 M KCl (FIG 3B). The specificity of the enzyme was tested in the presence of

307 different substrates (TABLE 2). The Cell5.1_3 cellulase failed to hydrolyze the
308 insoluble substrates crystalline cellulose (Avicel) and mannan, even when incubated
309 for 5 h at 20°C. Among the soluble substrates tested, it showed no activity against
310 xyloglucan, low activity against laminarin, and higher activity against CMC and
311 lichenan (β -1,3;1,4-glucan). The activity of the enzyme was enhanced by Mn^{2+} and
312 Cd^{2+} but clearly inhibited by Cu^{2+} , Zn^{2+} , and Fe^{2+} (Figure 3E). Triton X-100 was
313 found to increase the activity of the enzyme to 350%, while the other tested additives
314 were found to reduce (EDTA, glycerol) or totally inhibit (SDS, EtOH, DMSO) its
315 activity (FIG 3E).

316 **Discussion**

317 Microorganisms on algae, in the context of their numerous interactions with the host,
318 are known to produce diverse hydrolytic enzymes. Furthermore, the complexity of the
319 marine environment in which these microorganisms live leads to the production of
320 specific enzymes with original biochemistry (18). Here we have constructed a
321 metagenomic library from the genomic DNA of the bacterial communities associated
322 in February 2012 to the brown alga *Ascophyllum nodosum* (*Fucales* order) and have
323 screened it for diverse hydrolytic enzymes. From this library we have identified a
324 range of new esterase genes and two new glycoside hydrolase genes, thus
325 demonstrating the pertinence of seeking new enzymes in such microbial communities
326 by functional metagenomics. All these hydrolase genes can most likely be assigned to
327 the phylum *Proteobacteria* and essentially to the class *Alphaproteobacteria*, and a few
328 to the classes *Betaproteobacteria* and *Gammaproteobacteria*. This is consistent with
329 previous findings on the bacterial community associated in January 2007 with another
330 *Fucales* species, *Fucus vesiculosus*, as this bacterial community appeared dominated
331 by *Alphaproteobacteria* and *Gammaproteobacteria* species (20). Nonetheless,

332 *Bacteroidetes* and *Planctomycetes* species are also relatively abundant on brown algal
333 surfaces (20, 43), and one can be surprised that our functional metagenomic screen
334 yielded no genes identified as originating from these phyla. An explanation might be
335 that our *E. coli* expression strain (*Gammaproteobacteria*) does not readily recognize
336 gene promoters of *Bacteroidetes* and *Planctomycetes* species.

337 Esterases are the category of enzymes most frequently isolated by functional
338 metagenomics. Here we have identified 13 esterase-encoding genes by screening a
339 relatively low number of Mb: approximately 180 Mb, which makes on the average 1
340 esterase per 14 Mb screened. This is much higher than in other metagenomic studies
341 (see review of Steele et al. (44)). The 13 esterase genes identified here can be assigned
342 to four esterase families. Six of them (*Lip5.1_5*, *Lip5.3_3*, *Lip5.5_4*, *Lip5.13_3*,
343 *Lip5.14_3*, *Lip5.19_3*) have been assigned to family IV. This esterase family appears
344 highly represented in marine metagenomes, as out of 34 esterases identified in 8
345 screens of marine metagenomes, 27 have been classified as family-IV esterases (39,
346 45–51). We have assigned one of the esterase genes identified here (*Lip5.13_1*) to
347 esterase family V, which appears to count relatively few members. Esterases of this
348 family have been found in microorganisms displaying very different growth-
349 temperature ranges, such as *Sulfolobus* (thermophilic archaea), *Psychrobacter*
350 (psychrophiles), and *Moraxella* (mesophiles) species (37, 38). Only two esterases of
351 marine origin (FJ483459, FJ483468), identified in a metagenomic library from marine
352 sediments, have been assigned to this family previously (48). The *Lip5.11_3* and
353 *Lip5.15_3* esterases have been classified as family-VIII esterases. Members of family
354 VIII show no typical α/β -hydrolase fold but are very similar to β -lactamases.
355 Furthermore, they have a typical molecular weight around 42 kDa (38). The predicted
356 *Lip5.11_3* esterase shows sequence identity to β -lactamases and has an estimated

357 molecular weight of 41.4 kDa. The predicted Lip5.15 esterase has an estimated
358 molecular weight of 41.5 kDa. To our knowledge, no marine-metagenome-derived
359 esterase has previously been assigned to this family. We have identified conserved
360 domains of the SGNH superfamily (family-II esterases) in the esterase gene
361 (*Bgluc5.1_2*) found on the DNA insert of the beta-glucosidase candidate *BglucMM5.1*
362 and in the esterase gene (*Lip5.8_1*) of *LipMM5.8*. Only one marine-metagenome-
363 derived esterase (AB432912), identified in a sponge-associated bacterial metagenome
364 (50), has been classified previously as a member of this family. SGNH hydrolases are
365 a subgroup of the GDSL family II esterases. GDSL esterases have thioesterase,
366 protease, lysophospholipase, and arylesterase activities (52). In the BLAST results for
367 both *Bgluc5.1_2* and *Lip5.8_1*, we found sequence identities to thioesterases and
368 lysophospholipases.

369 Finally, the two thioesterase genes identified on *LipMM5.11* (*Lip5.11_4* and *_5*) seem
370 to be in an operon. The first gene (*Lip5.11_4*) has sequence identity to a thioesterase
371 of *Caulobacter* sp. K31 (Caul_4896, YP_001686513.1), located with another
372 thioesterase in a 3'-5' oriented operon (Caul_4895, YP_001686512.1). The other
373 thioesterase gene (*Lip5.11_5*) has sequence identity to a thioesterase gene of
374 *Parvibaculum lavamentivorans* DS-1 (Plav_0440, YP_001411720.1), which is also in
375 a 3'-5' oriented operon with another thioesterase (Plav_0441, YP_001411721.1). The
376 thioesterases and carboxylesterases constitute separate enzyme groups, although a few
377 proteins classified as carboxylesterases also show thioesterase activity. Despite careful
378 annotation of the esterase-containing gene clusters, it is difficult to predict the exact
379 biological function of these esterases on the basis of the genomic context.

380 Because (i) we have found thirteen genes encoding esterases, (ii) we have assigned
381 these genes to four esterase families, and (iii) most of the predicted protein sequences

382 are less than 50% identical to known esterases, we conclude that this library
383 constructed from algal biofilms can be considered rich and diverse in novel microbial
384 genes and enzymes.

385 As compared to esterase genes, few cellulase genes have been identified by functional
386 metagenomics. To our knowledge no cellulase has been identified by metagenomics
387 in a marine environment before. All of the cellulases identified by analysis of soil
388 (53–56) or rumen (57, 58) metagenomes display more than 50% sequence identity to
389 known cellulases, in contrast to our cellulase, which is only 42% identical to the
390 closest known cellulase. Hence, our cellulase appears new and only distantly related
391 to known cellulases. We have assigned the Cell5.1_3 cellulase to endo- β -1,4-
392 glucanase family GH5, subfamily 25. Although most cellulases identified by
393 metagenomics have been classified as GH5-family enzymes (8, 54, 57–59), only one
394 of them (ABE60714.1), identified in a pulp sediment metagenome from a paper mill
395 effluent, belongs to subfamily 25 (60). The other members of this subfamily were
396 identified in cultivable microorganisms such as *Caulobacter*, *Clostridium*,
397 *Dictyoglomus*, *Fervidobacterium*, and *Thermotoga* species isolated from diverse
398 environments (CAZyme database, www.cazy.org). Only two subfamily-25 enzymes
399 have been characterized so far: *Tm_Cel5A* (Q9X273) of *Thermotoga maritima* (61,
400 62) and *FnCel5A* (A7HNC0) of *Fervidobacterium nodosum* (63). Both of these
401 appear thermostable. Cell5.1_3, in contrast, shows low thermostability: when pre-
402 incubated for 1 hour at various temperatures, it began to show decreased CMC-
403 hydrolyzing activity when the pre-incubation temperature exceeded 25°C, and when
404 the pre-incubation temperature exceeded 40°C, no activity was observed. Under our
405 assay conditions, its activity was highest at 40°C, which is relatively low as compared
406 to the enzymes *Tm_Cel5A* and *FnCel5A*, identified in the thermophile *Thermotogales*

407 (around 80°C for both). On the other hand, our enzyme retained 11.8% of its
408 maximum activity at 0°C and 28.7% at 10°C. According to Asperborg et al. (64), most
409 of the endoglucanases of this subfamily have been derived from thermophiles, but
410 interestingly, the here-identified cold-active cellulase Cell5.1_3 would appear to come
411 from a mesophilic (or even psychrophilic) bacterium. A cold-active cellulase (CelX)
412 identified in *Pseudoalteromonas sp.* DY3 shows very similar optimal temperature and
413 thermostability (65). However, Cell5.1_3 is active over a broad pH range, from 5 to 8,
414 and is stable at pH values from 4 to 10 (it is still active after a 24-hour pre-incubation
415 in this pH range), which is not the case of the cold-active CelX (65). What's more,
416 other metagenome-derived GH5 cellulases show stability only at acidic (4-6.6, (59))
417 or alkaline (6-10, (55)) pH. Another advantage of Cell5.1_3 is its halotolerance, as it
418 retains 93% of its activity after a 24-hour pre-incubation in 3 M KCl, and 97% of its
419 activity after a pre-incubation in 4 M NaCl. These remaining activities are much
420 higher than those reported for other halotolerant GH5 enzymes. For example, Cel5A,
421 isolated from a soil metagenome, retains approximately 87% activity after a 20-h pre-
422 incubation in 3M NaCl or 4M KCl (54), and a GH5 endoglucanase of the
423 thermophilic eubacterium *Thermoanaerobacter tengcongensis* MB4 retains less than
424 15% of its activity after a 12-hour pre-incubation in 4 M NaCl (66). As GH5
425 endoglucanases hydrolyze a wide range of cellulose substrates (64), we have tested
426 the substrate specificity of Cell5.1_3. The enzyme appears to degrade CMC and
427 mixed glucans (lichenan (β 1,3- β 1,4) and laminarin (β 1,3- β 1,6)), but not β -1,4 linked
428 xyloglucan. Other metagenome-derived cellulases either fail to degrade laminarin
429 (53–55, 58) or, like the GH5 enzyme described by Juan et al., show very low specific
430 activity (approximately 0.002 U/mg) (59). In agreement with its classification as an
431 endoglucanase, Cell5.1_3 failed, in our assay, to degrade the insoluble substrates

432 Avicel and mannan. Triton X-100 enhances significantly the activity of Cell5.1_3.
433 Zheng et al (67) reported three main hypothesis explaining the enhancement of
434 enzymatic cellulose hydrolysis in the presence of non-ionic surfactants: 1) they
435 stabilize the enzyme by reducing thermal and/or mechanical shear forces; 2) they
436 change the substrate structure, enhancing the substrate accessibility; and 3) they affect
437 enzyme-substrate interaction, preventing enzymes inactivation due to non-productive
438 adsorption when hydrolyzing for example lignocellulosic substrates. This last
439 hypothesis doesn't concern the substrate we used (CMC) and as, generally, no
440 significant (positive or negative) effect of Triton X-100 has been observed on other
441 bacterial cellulase activities characterized with CMC (9, 54, 63, 68), the second
442 hypothesis must not be the reason why Cell5.1_3 activity increases to 350% in the
443 presence of Triton X-100. However, Cell5.1_3 is probably more stable, and therefore
444 more active on CMC, in the presence of Triton X-100.

445 **Competing interests**

446 The authors declare no competing interests.

447 **Acknowledgments**

448 This project was funded by Gembloux Agro-Bio Tech (ULg) and by Wallonie-
449 Bruxelles International (WBI) and the Fonds Scientifique de la Recherche (F.R.S-
450 F.N.R.S) in the framework of the Collaboration Program Hubert Curien. MM was
451 supported by ASSEMBLE grant agreement no. 227799 from the European Union to
452 the Station Biologique de Roscoff (FR). SB was a Postdoctoral Researcher of the Fonds
453 National de la Recherche Scientifique (F.R.S-FNRS). GM, MJ and TB were also
454 supported by the French National Research Agency with regard to the Investment
455 Expenditure Program IDEALG (<http://www.idealg.ueb.eu/>, grant agreement no. ANR-
456 10-BTBR-04).

457 We thank Cécile Herve for help with sample collection.

458 **References**

459

- 460 1. **Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl D a.** 1990.
461 Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for
462 analyzing mixed microbial populations. *Appl. Environ. Microbiol.* **56**:1919–25.
- 463 2. **Pace NR.** 1997. A Molecular View of Microbial Diversity and the Biosphere. *Science*
464 (80-.). **276**:734–740.
- 465 3. **Handelsman J.** 2004. Metagenomics: Application of Genomics to Uncultured
466 Microorganisms **68**:669–685.
- 467 4. **Sleator RD, Shortall C, Hill C.** 2008. Metagenomics. *Lett. Appl. Microbiol.* **47**:361–
468 6.
- 469 5. **Edwards JL, Smith DL, Connolly J, McDonald JE, Cox MJ, Joint I, Edwards C,**
470 **McCarthy AJ.** 2010. Identification of Carbohydrate Metabolism Genes in the
471 Metagenome of a Marine Biofilm Community Shown to Be Dominated by
472 Gammaproteobacteria and Bacteroidetes. *Genes (Basel).* **1**:371–384.
- 473 6. **Biver S, Vandebol M.** 2012. Characterization of three new carboxylic ester
474 hydrolases isolated by functional screening of a forest soil metagenomic library. *J. Ind.*
475 *Microbiol. Biotechnol.*
- 476 7. **Sharma S, Khan FG, Qazi GN.** 2010. Molecular cloning and characterization of
477 amylase from soil metagenomic library derived from Northwestern Himalayas. *Appl.*
478 *Microbiol. Biotechnol.* **86**:1821–8.
- 479 8. **Geng A, Zou G, Yan X, Wang Q, Zhang J, Liu F, Zhu B, Zhou Z.** 2012. Expression
480 and characterization of a novel metagenome-derived cellulase Exo2b and its

- 481 application to improve cellulase activity in *Trichoderma reesei*. *Appl. Microbiol.*
482 *Biotechnol.* **96**:951–62.
- 483 9. **Feng Y, Duan C-J, Pang H, Mo X-C, Wu C-F, Yu Y, Hu Y-L, Wei J, Tang J-L,**
484 **Feng J-X.** 2007. Cloning and identification of novel cellulase genes from uncultured
485 microorganisms in rabbit cecum and characterization of the expressed cellulases. *Appl.*
486 *Microbiol. Biotechnol.* **75**:319–328.
- 487 10. **Wang T-Y, Chen H-L, Lu M-YJ, Chen Y-C, Sung H-M, Mao C-T, Cho H-Y, Ke**
488 **H-M, Hwa T-Y, Ruan S-K, Hung K-Y, Chen C-K, Li J-Y, Wu Y-C, Chen Y-H,**
489 **Chou S-P, Tsai Y-W, Chu T-C, Shih C-C a, Li W-H, Shih M-C.** 2011. Functional
490 characterization of cellulases identified from the cow rumen fungus *Neocallimastix*
491 *patriciarum* W5 by transcriptomic and secretomic analyses. *Biotechnol. Biofuels* **4**:24.
- 492 11. **Ferrer M, Ghazi A, Beloqui A, Vieites JM, López-Cortés N, Marín-Navarro J,**
493 **Nechitaylo TY, Guazzaroni M-E, Polaina J, Waliczek A, Chernikova TN, Reva**
494 **ON, Golyshina O V, Golyshin PN.** 2012. Functional metagenomics unveils a
495 multifunctional glycosyl hydrolase from the family 43 catalysing the breakdown of
496 plant polymers in the calf rumen. *PLoS One* **7**:e38134.
- 497 12. **Mattéotti C, Thonart P, Francis F, Haubruge E, Destain J, Brasseur C, Bauwens**
498 **J, De Pauw E, Portetelle D, Vandenbol M.** 2011. New glucosidase activities
499 identified by functional screening of a genomic DNA library from the gut microbiota
500 of the termite *Reticulitermes santonensis*. *Microbiol. Res.* **166**:629–42.
- 501 13. **Tasse L, Bercovici J, Pizzut-Serin S, Robe P, Tap J, Klopp C, Cantarel BL,**
502 **Coutinho PM, Henrissat B, Leclerc M, Doré J, Monsan P, Remaud-Simeon M,**
503 **Potocki-Veronese G.** 2010. Functional metagenomics to mine the human gut
504 microbiome for dietary fiber catabolic enzymes. *Genome Res.* **20**:1605–12.
- 505 14. **Kennedy J, O’Leary ND, Kiran GS, Morrissey JP, O’Gara F, Selvin J, Dobson**
506 **ADW.** 2011. Functional metagenomic strategies for the discovery of novel enzymes

- 507 and biosurfactants with biotechnological applications from marine ecosystems. *J.*
508 *Appl. Microbiol.* **111**:787–799.
- 509 15. **Lee D-G, Jeon JH, Jang MK, Kim NY, Lee JH, Lee J-H, Kim S-J, Kim G-D, Lee**
510 **S-H.** 2007. Screening and characterization of a novel fibrinolytic metalloprotease from
511 a metagenomic library. *Biotechnol. Lett.* **29**:465–72.
- 512 16. **Lee M-H, Lee C-H, Oh T-K, Song JK, Yoon J-H.** 2006. Isolation and
513 characterization of a novel lipase from a metagenomic library of tidal flat sediments:
514 evidence for a new family of bacterial lipases. *Appl. Environ. Microbiol.* **72**:7406–9.
- 515 17. **Sim S-J, Baik KS, Park SC, Choe HN, Seong CN, Shin T-S, Woo HC, Cho J-Y,**
516 **Kim D.** 2012. Characterization of alginate lyase gene using a metagenomic library
517 constructed from the gut microflora of abalone. *J. Ind. Microbiol. Biotechnol.* **39**:585–
518 93.
- 519 18. **Martin M, Portetelle D, Michel G, Vandebol M.** 2014. Microorganisms living on
520 macroalgae: diversity, interactions, and biotechnological applications. *Appl.*
521 *Microbiol. Biotechnol.* **98**:2917–2935.
- 522 19. **Burke C, Steinberg P, Rusch D, Kjelleberg S, Thomas T.** 2011. Bacterial
523 community assembly based on functional genes rather than species. *PNAS*
524 **108**:14288–14293.
- 525 20. **Lachnit T, Meske D, Wahl M, Harder T, Schmitz R.** 2011. Epibacterial community
526 patterns on marine macroalgae are host-specific but temporally variable. *Environ.*
527 *Microbiol.* **13**:655–65.
- 528 21. **Fisher MM, Wilcox LW, Graham LE, Wilcox LEEW.** 1998. Molecular
529 Characterization of Epiphytic Bacterial Communities on Charophycean Green Algae
530 Molecular Characterization of Epiphytic Bacterial Communities on Charophycean
531 Green Algae **64**.

- 532 22. **Debnath M, Paul AK, Bisen PS.** 2007. Natural Bioactive Compounds and
533 Biotechnological Potential of Marine Bacteria. *Curr. Pharm. Biotechnol.* **1**:253–260.
- 534 23. **Michel G, Czjzek M.** 2013. Polysaccharide-degrading enzymes from marine bacteria,
535 p. 429–464. *In* Trincone, A (ed.), *Marine enzymes for biocatalysis: sources, biocatalytic*
536 *characteristic and bioprocesses of marine enzymes.*
- 537 24. **Gibbs MD, Saul DJ, Lüthi E, Bergquist PL.** 1992. The beta-mannanase from
538 “*Caldocellum saccharolyticum*” is part of a multidomain enzyme. *Appl. Environ.*
539 *Microbiol.* **58**:3864–7.
- 540 25. **Fu X, Liu P, Lin L, Hong Y, Huang X, Meng X, Liu Z.** 2010. A Novel
541 Endoglucanase (Cel9P) From a Marine Bacterium *Paenibacillus* sp. BME-14. *Appl.*
542 *Biochem. Biotechnol.* **160**:1627–1636.
- 543 26. **Dong J, Hong Y, Shao Z, Liu Z.** 2010. Molecular cloning, purification, and
544 characterization of a novel, acidic, pH-stable endoglucanase from *Marteella*
545 *mediterranea*. *J Microbiol.* **48**:393–398.
- 546 27. **S. Shanmughapriya, Kiran GS, Selvin J, Thomas TA, Rani C.** 2010. Optimization,
547 Purification, and Characterization of Extracellular Mesophilic Alkaline Cellulase from
548 Sponge-Associated *Marinobacter* sp. MSI032. *Appl. Biochem. Biotechnol.* **162**:625–
549 640.
- 550 28. **Kuhad RC, Gupta R, Singh A.** 2011. Microbial cellulases and their industrial
551 applications. *Enzyme Res.* **2011**:280696.
- 552 29. **Duan C-J, Feng J-X.** 2010. Mining metagenomes for novel cellulase genes.
553 *Biotechnol. Lett.* **32**:1765–75.
- 554 30. **Lachnit T, Blümel M, Imhoff J, Wahl M.** 2009. Specific epibacterial communities
555 on macroalgae: phylogeny matters more than habitat. *Aquat. Biol.* **5**:181–186.
- 556 31. **Petersen TN, Brunak S, von Heijne G, Nielsen H.** 2011. SignalP 4.0: discriminating
557 signal peptides from transmembrane regions. *Nat. Methods* **8**:785–6.

- 558 32. **Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A,**
559 **Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M.** 2014.
560 Pfam: the protein families database. *Nucleic Acids Res.* **42**:D222–30.
- 561 33. **Biver S, Portetelle D, Vandenbol M.** 2013. Characterization of a new oxidant-stable
562 serine protease isolated by functional metagenomics. *Springerplus* **2**:410.
- 563 34. **Miller GL.** 1959. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing
564 Sugar. *Anal. Chem.* **31**:426–428.
- 565 35. **Burke C, Kjelleberg S, Thomas T.** 2009. Selective extraction of bacterial DNA from
566 the surfaces of macroalgae. *Appl. Environ. Microbiol.* **75**:252–6.
- 567 36. **Popper Z a, Michel G, Hervé C, Domozych DS, Willats WGT, Tuohy MG,**
568 **Kloareg B, Stengel DB.** 2011. Evolution and diversity of plant cell walls: from algae
569 to flowering plants. *Annu. Rev. Plant Biol.* **62**:567–90.
- 570 37. **Arpigny JL, Jaeger KE.** 1999. Bacterial lipolytic enzymes: classification and
571 properties. *Biochem. J.* **343 Pt 1**:177–83.
- 572 38. **Hausmann S, Jaeger KE.** 2010. Lipolytic enzymes of bacteria, p. 1099–1126. *In*
573 Timmis, KN (ed.), *Handbook of Hydrocarbon and Lipid Microbiology.* Springer
574 Berlin Heidelberg, Berlin, Heidelberg.
- 575 39. **Chu X, He H, Guo C, Sun B.** 2008. Identification of two novel esterases from a
576 marine metagenomic library derived from South China Sea. *Appl. Microbiol.*
577 *Biotechnol.* **80**:615–25.
- 578 40. **Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B.**
579 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for
580 Glycogenomics. *Nucleic Acids Res.* **37**:D233–8.
- 581 41. **Aspeborg H, Coutinho PM, Wang Y, Brumer H, Henrissat B.** 2012. Evolution,
582 substrate specificity and subfamily classification of glycoside hydrolase family 5
583 (GH5). *BMC Evol. Biol.* **12**:186.

- 584 42. **Lee K, Lee HK, Choi T-H, Cho J-C.** 2007. *Robiginitomaculum antarcticum* gen.
585 nov., sp. nov., a member of the family Hyphomonadaceae, from Antarctic seawater.
586 *Int. J. Syst. Evol. Microbiol.* **57**:2595–9.
- 587 43. **Bengtsson MM, Sjoetun K, Lanzen A, Oevreas L.** 2013. Bacterial diversity in
588 relation to secondary production and succession on surfaces of the kelp *Laminaria*
589 *hyperborea*. *ISME J.* **6**:2188–2198.
- 590 44. **Steele HL, Jaeger K-E, Daniel R, Streit WR.** 2009. Advances in recovery of novel
591 biocatalysts from metagenomes. *J. Mol. Microbiol. Biotechnol.* **16**:25–37.
- 592 45. **Fu C, Hu Y, Xie F, Guo H, Ashforth EJ, Polyak SW, Zhu B, Zhang L.** 2011.
593 Molecular cloning and characterization of a new cold-active esterase from a deep-sea
594 metagenomic library. *Appl. Microbiol. Biotechnol.* **90**:961–70.
- 595 46. **Jeon JH, Kim J-T, Kang SG, Lee J-H, Kim S-J.** 2009. Characterization and its
596 potential application of two esterases derived from the arctic sediment metagenome.
597 *Mar. Biotechnol. (NY).* **11**:307–16.
- 598 47. **Jeon JH, Lee HS, Kim JT, Kim S-J, Choi SH, Kang SG, Lee J-H.** 2012.
599 Identification of a new subfamily of salt-tolerant esterases from a metagenomic library
600 of tidal flat sediment. *Appl. Microbiol. Biotechnol.* **93**:623–31.
- 601 48. **Hu Y, Fu C, Huang Y, Yin Y, Cheng G, Lei F, Lu N, Li J, Ashforth EJ, Zhang L,**
602 **Zhu B.** 2010. Novel lipolytic genes from the microbial metagenomic library of the
603 South China Sea marine sediment. *FEMS Microbiol. Ecol.* **72**:228–37.
- 604 49. **Jiang X, Xu X, Huo Y, Wu Y, Zhu X, Zhang X, Wu M.** 2012. Identification and
605 characterization of novel esterases from a deep-sea sediment metagenome. *Arch.*
606 *Microbiol.* **194**:207–14.
- 607 50. **Okamura Y, Kimura T, Yokouchi H, Meneses-Osorio M, Katoh M, Matsunaga T,**
608 **Takeyama H.** 2010. Isolation and characterization of a GDSL esterase from the

- 609 metagenome of a marine sponge-associated bacteria. *Mar. Biotechnol. (NY)*. **12**:395–
610 402.
- 611 51. **Oh K-H, Nguyen G-S, Kim E-Y, Kourist R, Bornscheuer U, Oh T-K, Yoon J-H.**
612 2012. Characterization of a novel esterase isolated from intertidal flat metagenome and
613 its tertiary alcohols synthesis. *J. Mol. Catal. B Enzym.* **80**:67–73.
- 614 52. **Akoh CC, Lee G-C, Liaw Y-C, Huang T-H, Shaw J-F.** 2004. GDSL family of serine
615 esterases/lipases. *Prog. Lipid Res.* **43**:534–52.
- 616 53. **Nacke H, Engelhaupt M, Brady S, Fischer C, Tautzt J, Rolf D.** 2012. Identification
617 and characterization of novel cellulolytic and hemicellulolytic genes and enzymes
618 derived from German grassland soil metagenomes. *Biotechnol. Lett.* **34**:663–675.
- 619 54. **Voget S, Steele HL, Streit WR.** 2006. Characterization of a metagenome-derived
620 halotolerant cellulase. *J. Biotechnol.* **126**:26–36.
- 621 55. **Pang H, Zhang P, Duan C-J, Mo X-C, Tang J-L, Feng J-X.** 2009. Identification of
622 cellulase genes from the metagenomes of compost soils and functional
623 characterization of one novel endoglucanase. *Curr. Microbiol.* **58**:404–8.
- 624 56. **Soo-Jin K, Chang-Muk L, Bo-Ram H, Min-Young K, Yun-Soo Y, Sang-Hong Y,**
625 **Bon-Sung K, Hong-Ki J.** 2008. Characterization of a gene encoding cellulase from
626 uncultured soil bacteria. *FEMS Microbiol. Lett.* **282**:44–51.
- 627 57. **Rashamuse KJ, Visser DF, Hennessy F, Kemp J, Roux-van der Merwe MP,**
628 **Badenhorst J, Ronneburg T, Francis-Pope R, Brady D.** 2013. Characterisation of
629 two bifunctional cellulase-xylanase enzymes isolated from a bovine rumen
630 metagenome library. *Curr. Microbiol.* **66**:145–51.
- 631 58. **Bao L, Huang Q, Chang L, Zhou J, Lu H.** 2011. Screening and characterization of a
632 cellulase with endocellulase and exocellulase activity from yak rumen metagenome. *J.*
633 *Mol. Catal. B Enzym.* **73**:104–110.

- 634 59. **Liu J, Liu W, Zhao X, Shen W, Cao H, Cui Z.** 2011. Cloning and functional
635 characterization of a novel endo- β -1,4-glucanase gene from a soil-derived
636 metagenomic library. *Appl. Microbiol. Biotechnol.* **89**:1083–1092.
- 637 60. **Xu Y, Duan C, Zhou Q, Tang J, Feng J.** 2006. Cloning and identification of cellulase
638 genes from uncultured microorganisms in pulp sediments from paper mill effluent.
639 *Wei Sheng Wu Xue Bao.* **46**:783–788.
- 640 61. **Pereira JH, Chen Z, McAndrew RP, Sapra R, Chhabra SR, Sale KL, Simmons B**
641 **a, Adams PD.** 2010. Biochemical characterization and crystal structure of
642 endoglucanase Cel5A from the hyperthermophilic *Thermotoga maritima*. *J. Struct.*
643 *Biol.* **172**:372–9.
- 644 62. **Chhabra SR, Shockley KR, Ward DE, Kelly RM.** 2002. Regulation of Endo-Acting
645 Glycosyl Hydrolases in the Hyperthermophilic Bacterium *Thermotoga maritima*
646 Grown on Glucan- and Mannan-Bases Polysaccharides. *Appl. Environ. Microbiol.*
647 **68**:545–554.
- 648 63. **Wang Y, Wang X, Tang R, Yu S, Zheng B, Feng Y.** 2010. A novel thermostable
649 cellulase from *Fervidobacterium nodosum*. *J. Mol. Catal. B Enzym.* **66**:294–301.
- 650 64. **Aspeborg H, Coutinho PM, Wang Y, Brumer H, Henrissat B.** 2012. Evolution,
651 substrate specificity and subfamily classification of glycoside hydrolase family 5
652 (GH5). *BMC Evol. Biol.* **12**:186.
- 653 65. **Zeng R, Xiong P, Wen J.** 2006. Characterization and gene cloning of a cold-active
654 cellulase from a deep-sea psychrotrophic bacterium *Pseudoalteromonas* sp. DY3.
655 *Extremophiles* **10**:79–82.
- 656 66. **Liang C, Xue Y, Fioroni M, Rodríguez-Ropero F, Zhou C, Schwaneberg U, Ma Y.**
657 2011. Cloning and characterization of a thermostable and halo-tolerant endoglucanase
658 from *Thermoanaerobacter tengcongensis* MB4. *Appl. Microbiol. Biotechnol.* **89**:315–
659 26.

- 660 67. **Zheng Y, Pan Z, Zhang R, Wang D, Jenkins B.** 2008. Non-ionic surfactants and
661 non-catalytic protein treatment on enzymatic hydrolysis of pretreated Creeping Wild
662 Ryegrass. *Appl. Biochem. Biotechnol.* **146**:231–48.
- 663 68. **Duan C-J, Xian L, Zhao G-C, Feng Y, Pang H, Bai X-L, Tang J-L, Ma Q-S, Feng**
664 **J-X.** 2009. Isolation and partial characterization of novel genes encoding acidic
665 cellulases from metagenomes of buffalo rumens. *J. Appl. Microbiol.* **107**:245–56.
- 666
- 667

668 **FIG 1** Neighbor-joining phylogenetic tree of subfamily GH5_25 and classification of
669 the cellulase gene *Cell5.1_3*. Bootstrap values higher than 50% are indicated at
670 branch points (1,000 replicates). Reference sequences from the GenBank database are
671 added.

672

673 **FIG 2** SDS-PAGE analysis of the recombinant cellulase. **MW**: Molecular weight, **T**:
674 Total extract. **I0** and **S0** show the insoluble and soluble fractions obtained by
675 sonication, followed by centrifugation, of *E. coli* cells overexpressing the cellulase
676 gene *Cell5.1_3*. As the GH was mostly present in the insoluble fraction (I0), we
677 resuspended it for 4 h in a solubilization buffer and centrifuged it, aiming to obtain a
678 higher, and sufficient, amount of our cellulase in the soluble fraction (S1). This
679 fraction was then purified twice by chromatography (E1 and E2)

680

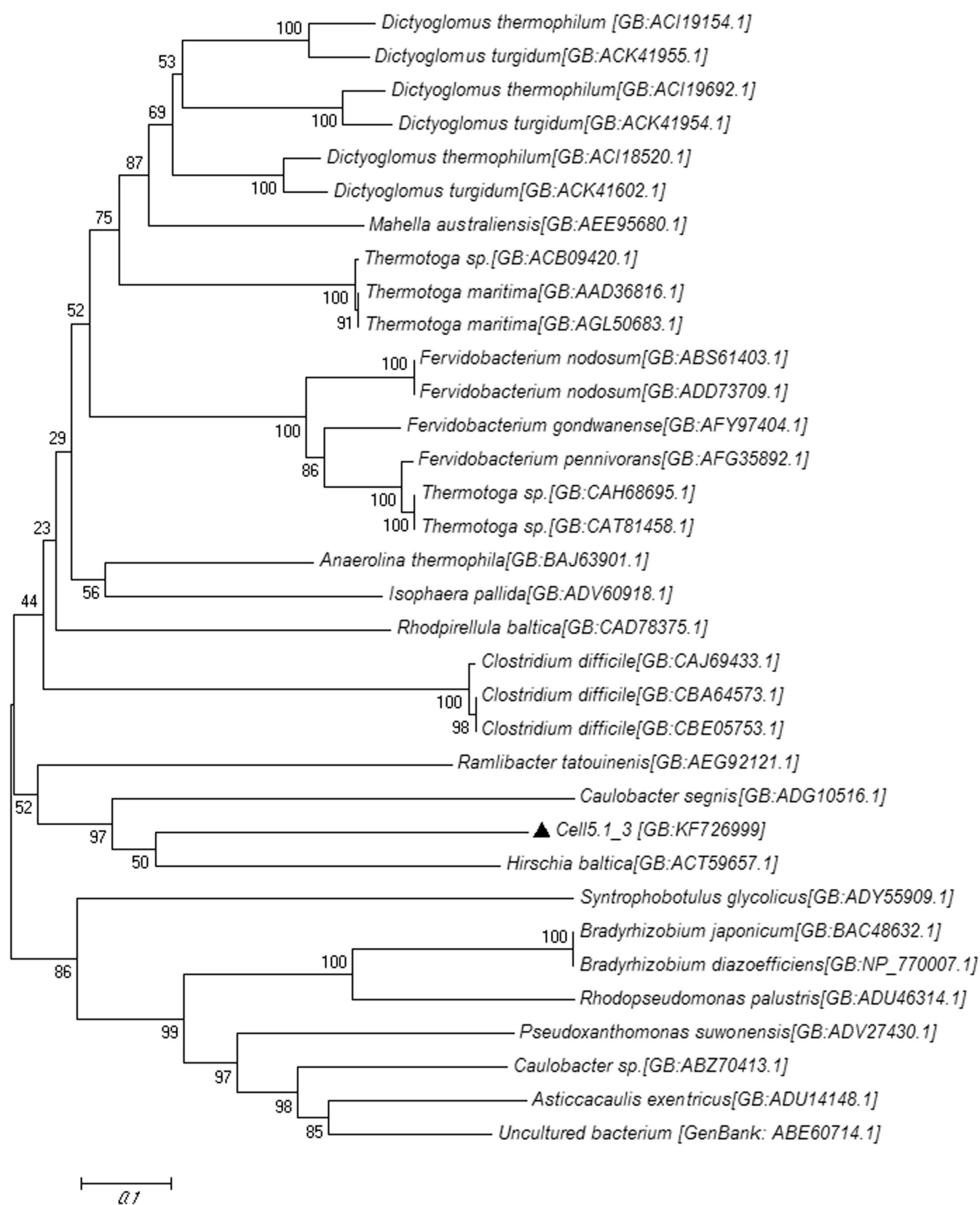
681 **FIG 3** Characterization of the activity of Cell5.1_3. (A) Influence of temperature on
682 Cell5.1_3 activity (●) and thermostability after one hour incubation at indicated
683 temperature (■) (B) Halotolerance of Cell5.1_3 (C) Influence of pH on Cell5.1_3
684 activity (D) Residual activity of Cell5.1_3 after a 24-hour incubation at the indicated
685 pH, (Ac : acetate Buffer, Ph: phosphate buffer, Tris : Tris-HCl buffer, Gly : glycine-
686 NaOH buffer) (E) Effects of ions and other additives on the activity of the Cell5.1_3.

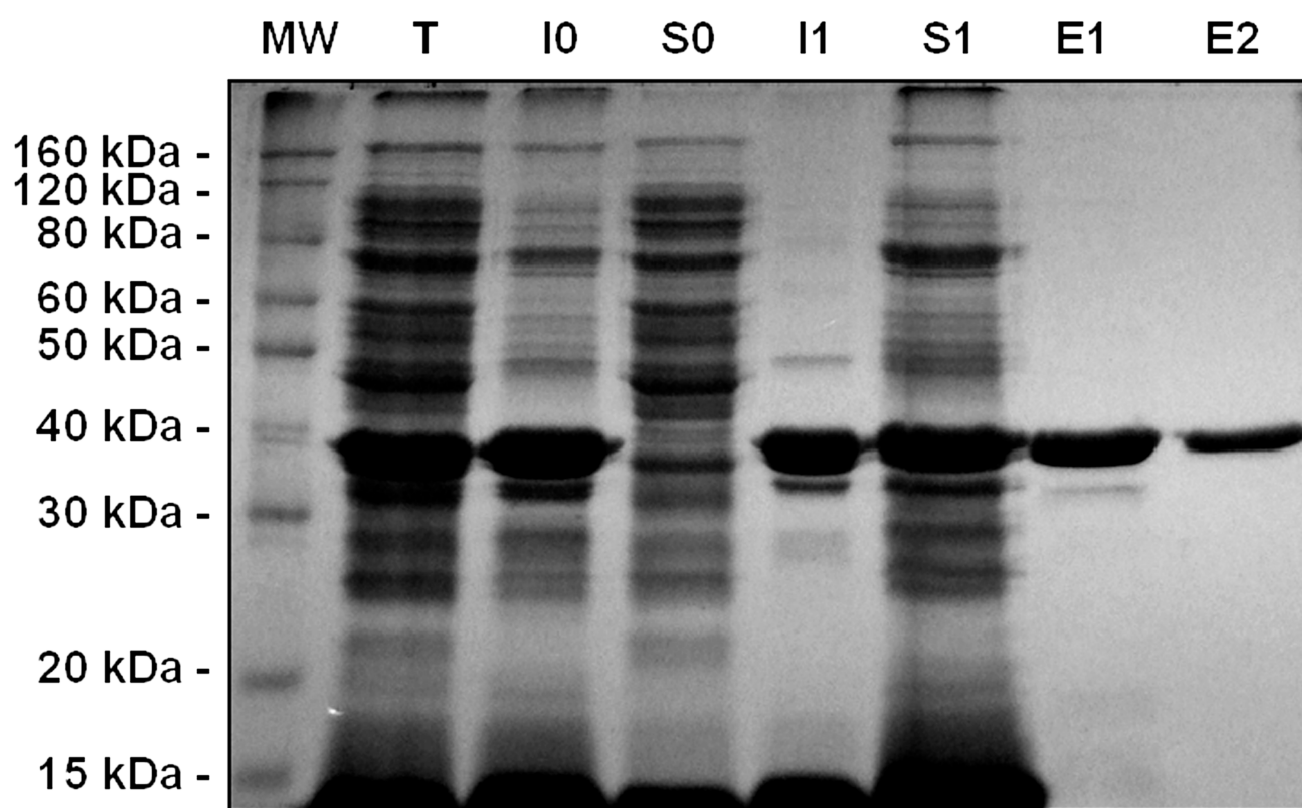
687

688 **TABLE 1** ORFs on the esterase-, cellulase-, and beta-glucosidase-encoding plasmids,
689 showing sequence identities to known esterases and glycoside hydrolases.

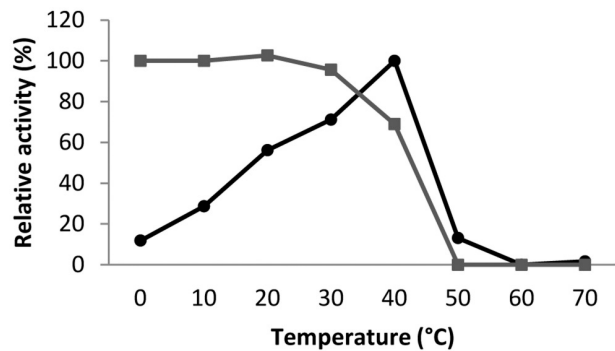
690

691 **TABLE 2** Substrate specificity of Cell5.1_3.

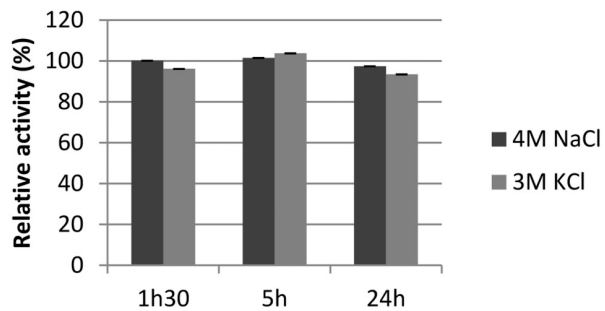




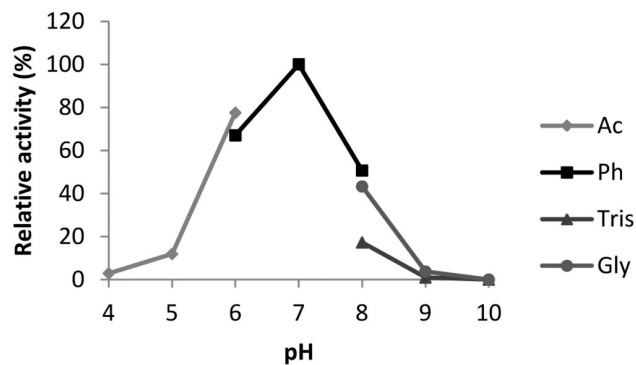
(A)



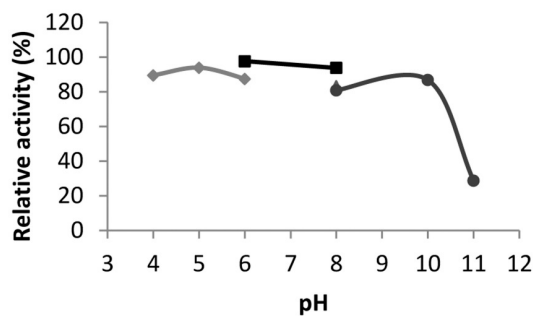
(B)



(C)



(D)



(E)

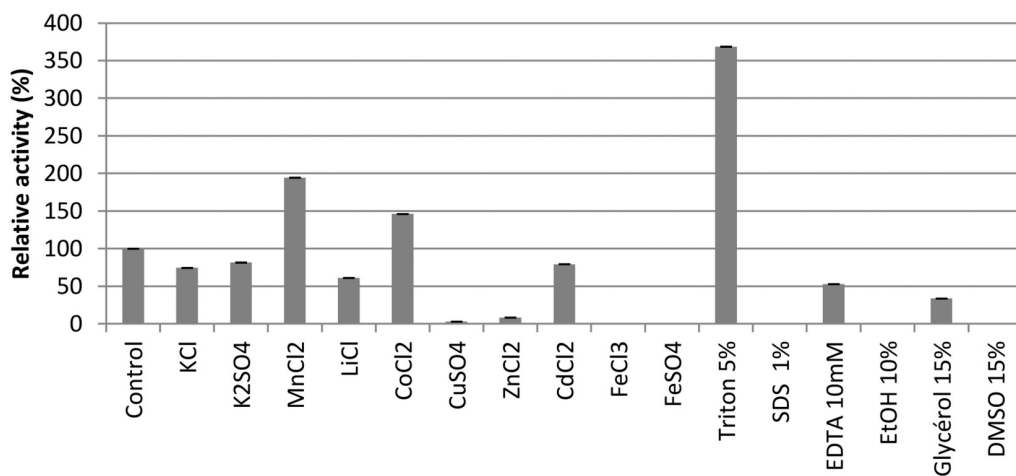


TABLE 1 ORFs on the esterase-, cellulase-, and beta-glucosidase-encoding plasmids, sequences showing sequence identities to known esterases and glycoside hydrolases.

Activity	Candidate name	Gene name	ORF size (amino acids)	Best hit against GenBank (Source organisms)	Sequence Identity (%)	Family assignment	Accession number	
Carboxyesterase	<i>LipMM5.1</i>	<i>Lip5.1_5</i>	295	Lipase/esterase [uncultured bacterium] ACZ16567	58	Est IV	KF726988	
	<i>LipMM5.3</i>	<i>Lip5.3_3</i>	314	Alpha/beta hydrolase [<i>Hirschia baltica</i> ATCC 49814] YP_003058522	41	Est IV	KF726989	
	<i>LipMM5.5</i>	<i>Lip5.5_4</i>	328	Alpha/beta hydrolase [<i>Maricaulis maris</i>] YP_757536.1	36	Est IV	KF726990	
	<i>LipMM5.7</i>	<i>Lip5.7_2</i>	414	Lysophospholipase [<i>Burkholderia</i> sp. KJ006] YP_006332790	33	/	KF726991	
	<i>LipMM5.8</i>	<i>Lip5.8_1</i>	237	Hydrolase GDSL [<i>Ponticaulis koreensis</i>] WP_006609299.1	46	Est II	KF726992	
		<i>Lip5.8_2</i>	221	ABC Transporter [<i>Erythrobacter litoralis</i> HTCC2594] YP_458689.1	57	/		
		<i>Lip5.8_3</i>	844	Putative ABC-type transport system involved in lysophospholipase L1 biosynthesis, permease component [<i>Caulobacter vibrioides</i>] WP_004617736.1	41	/		
	<i>LipMM5.11</i>	<i>Lip5.11_3</i>	378	Penicillin-binding protein, beta-lactamase class C [<i>Caulobacter vibrioides</i>] WP_004621811	51	Est VIII	KF726993	
		<i>Lip5.11_4</i>	153	Thioesterase superfamily protein [<i>Caulobacter</i> sp. K31] YP_001686513	49	/		
		<i>Lip5.11_5</i>	138	Thioesterase superfamily protein [<i>Parvibaculum lavamentivorans</i> DS-1] YP_001411720.1	52	/		
	<i>LipMM5.12</i>	<i>Lip5.12_2</i>	288	Esterase [uncultured bacterium] ACF49126.1	38	/	KF726994	
	<i>LipMM5.13</i>	<i>Lip5.13_1</i>	332	Alpha/beta hydrolase [<i>Pseudomonas pseudoalcaligenes</i>] WP_004422344.1	34	Est V	KF726995	
		<i>Lip5.13_3</i>	348	Lipase/esterase [uncultured sludge bacterium] ADC79145.1	47	Est IV		
	<i>LipMM5.14</i>	<i>Lip5.14_3</i>	314	Alpha/beta hydrolase [<i>Hirschia baltica</i> ATCC 49814] WP_017932623.1	41	Est IV	KF726996	
	<i>LipMM5.15</i>	<i>Lip5.15_3</i>	383	Esterase [<i>Glaciecola nitratreducens</i> FR1064] YP_004870252.1	57	Est VIII	KF726997	
	<i>LipMM5.19</i>	<i>Lip5.19_3</i>	283	Esterase [<i>Erythrobacter litoralis</i> HTCC2594] YP_458833.1	48	Est IV	KF726998	
	Beta-glucosidase	<i>BglucMM5.1</i>	<i>Bgluc5.1_1</i>	749*	1,4-beta-D-glucan glucohydrolase [<i>Xanthomonas campestris</i>] WP_017932269	56	GH3	KJ499810
			<i>Bgluc5.1_2</i>	244	G-D-S-L family lipolytic protein [<i>Hirschia baltica</i> ATCC 49814] YP_003059731.1	37	Est II	
			<i>Bgluc5.1_4</i>	521*	Xylulokinase [<i>Novosphingobium</i> sp. PP1Y] YP_004534688	60	/	
Cellulase	<i>CellMM5.1</i>	<i>Cell5.1_3</i>	359	Glycoside hydrolase family protein [<i>Hirschia baltica</i> ATCC 49814] YP_003060354.1	42	GH5	KF726999	

* Incomplete ORF

TABLE 2 Substrate specificity of *Cell5.1_3*

Substrates	Main linkage type	Specific activity ^a (U/mg)
<u>Soluble substrates</u>		
Carboxymethylcellulose	β -1,4-glucan	10.22 \pm 0.002
Lichenan	β -1,3/1,4-glucan	12.61 \pm 0.008
Laminarin	β -1,3/1,6-glucan	5.08 \pm 0.004
Xyloglucan	β -1,4-xyloglucan	UD
<u>Insoluble substrates</u> ^b		
Mannan	β -1,4-mannan	UD
Avicel	β -1,4-glucan	UD

UD, Undetectable

^a One unit of enzyme activity (U) corresponds to 1 μ mole of glucose released in 1 min. Measurements were performed in triplicate and standard deviations were calculated.

^b For these substrates, the reaction time was extended to 5 h at 20°C, with automatic inversion.