Extremophiles (2014) 18:853–863
DOI 10.1007/s00792-014-0676-3

SPECIAL ISSUE: ORIGINAL PAPER

10th International Congress on Extremophiles

Carbohydrate-active enzymes identified by metagenomic analysis of deep-sea sediment bacteria

Barbara Klippel · Kerstin Sahm · Alexander Basner · Sigrid Wiebusch · Patrick John · Ute Lorenz · Anke Peters · Fumiyoshi Abe · Kyoma Takahashi · Olaf Kaiser · Alexander Goesmann · Sebastian Jaenicke · Ralf Grote · Koki Horikoshi · Garabed Antranikian

Received: 28 May 2014 / Accepted: 8 July 2014 / Published online: 10 August 2014

Abstract Subseafloor sediment samples derived from a sediment core of 60 m length were used to enrich psychrophilic aerobic bacteria on cellulose, xylan, chitin, and starch. A variety of species belonging to *Alpha-* and *Gammaproteobacteria* and to *Flavobacteria* were isolated from sediment depths between 12 and 42 mbsf. Metagenomic DNA purified from the pooled enrichments was sequenced and analyzed for phylogenetic composition and presence of genes encoding carbohydrate-active enzymes. More than 200 open reading frames coding for glycoside hydrolases were identified, and more than 60 of them relevant for enzymatic degradation of lignocellulose. Four genes encoding β-glucosidases with less than 52 % identities to characterized enzymes were chosen for recombinant expression in *Escherichia coli*. In addition one endomannanase, two endoxylanases, and three β-xylosidases were produced recombinantly. All genes could be actively expressed. Functional analysis revealed discrepancies and additional variability for the recombinant enzymes as compared to the sequence-based predictions.

Keywords Biodiversity · Psychrophiles

Communicated by H. Atomi.

This article is part of a special issue based on the 10th International Congress on Extremophiles held in Saint Petersburg, Russia, September 7–11, 2014.

Electronic supplementary material The online version of this article (doi:10.1007/s00792-014-0676-3) contains supplementary material, which is available to authorized users.

B. Klippel · K. Sahm · A. Basner · S. Wiebusch · P. John · U. Lorenz · A. Peters · R. Grote · G. Antranikian (✉)
Institute of Technical Microbiology, Hamburg University of Technology, Kasernenstr. 12, 21073 Hamburg, Germany
e-mail: antranikian@tuhh.de

K. Sahm
e-mail: sahm@tuhh.de

F. Abe · K. Horikoshi
Extremobiosphere Research Center (XBR), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15 Natsushima-cho, Yokosuka, Kanagawa 237-0061, Japan

F. Abe
Department of Chemistry and Biological Science, College of Science and Engineering, Aoyama Gakuin University, 5-10-1 Fuchinobe, Chuo-ku, Sagamihara 252-5258, Japan

K. Takahashi
Center for Deep Earth Exploration (CDEX), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 3173-25, Showa-machi, Kanazawa-ku, Yokohama, Kanagawa 236-0001, Japan

O. Kaiser
Roche Diagnostics GmbH, Nonnenwald 2, 82377 Penzberg, Germany

A. Goesmann · S. Jaenicke
Center for Biotechnology, CeBiTec, Universitaetsstrasse 25, 33615 Bielefeld, Germany

Introduction

Approximately 70 % of the Earth's surface is covered by the oceans and represents one of the world's largest ecosystem. Nearly 95 % of the seabed is located in water depths characterized by low light intensity, temperatures close to freezing, low oxygen concentration, and high

hydrostatic pressure. Psychrophilic and barophilic or barotolerant bacteria thrive under these extreme conditions found in the deep sea- and subseafloor (Parkes et al. 2000; Qin et al. 2011). Marine sediments are a huge reservoir of organic carbon of biogenic, terrigenous, and volcanogenic origin and up to 10 billion tons of particulate organic matter are accumulated at the seafloor (Jørgensen and Boetius 2007; Fry et al. 2008). Sedimentary material containing polysaccharides is an important source of organic carbon in marine environments. The degradation is carried out by hydrolytic enzymes but the main microorganisms present at the seafloor and the enzymes involved in organic matter hydrolysis are mainly unknown (Jørgensen and Boetius 2007; Edwards et al. 2010). In order to gain further insight into the diversity and metabolic processes of microbial communities in marine environments, metagenomics has become more and more important. Two approaches are currently used to identify functional genes: activity- and sequence-based screening. In activity-based screening, metagenomic libraries are screened for target enzyme activities. An advantage of this screening method is that the enzyme activity is functionally guaranteed. However, the successful identification of active biocatalysts depends on the nature and quality of DNA, cloning, and expression. Sequence-based screening relies on known conserved sequences and as a result of the increased generation of sequence data, the prediction of correct functions of novel genes will be a challenge (Nyyssönen et al. 2013). The metagenomic approach is an efficient tool for the identification of the biodiversity and novel biocatalysts from many different environments such as the marine ecosystem (Berlemont et al. 2009; Glöckner and Joint 2010).

Enzymes of psychrophilic marine microorganisms are promising biocatalysts for industrial applications, since they are characterized by a higher stability and activity at low temperatures compared with enzymes derived from mesophilic counterparts. The utilization of cold-active enzymes could reduce energy costs for low-temperature operations (Zhang and Kim 2010; Wierzbicka-Wos et al. 2013). Due to the fact that marine environments harbor a variety of polysaccharides they are expected to be a valuable source for novel glycoside hydrolases (Edwards et al. 2010).

Glycoside hydrolases cleave the glycosidic bonds between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety (Henrissat and Coutinho 2001; Edwards et al. 2010). Based on amino acid sequence similarity and similar three-dimensional structure, glycoside hydrolases are classified into families and up to date 133 glycoside hydrolase families are known (Henrissat 1991; Cantarel et al. 2009; CAZy database www.cazy.org). The complete degradation of complex carbohydrates is facilitated by the synergistic action of a large variety of carbohydrate-active enzymes. Although several glycoside hydrolase producing bacteria were isolated from marine environments (Glöckner and Joint 2010), little is known about carbohydrate-active enzymes produced by subseafloor sediment bacteria.

In this study, we focused on the identification of glycoside hydrolases from subseafloor bacteria for the degradation of cellulosic and hemicellulosic biomass. Deep-sea sediments taken from Suruga Bay, Japan, were incubated with complex carbohydrates to facilitate the enrichment of polysaccharide degrading bacteria. Metagenomic DNA was isolated, sequenced and the obtained sequences were analyzed for the presence of genes encoding carbohydrate-active enzymes. Selected glycoside hydrolases were produced recombinantly and investigations were carried out to determine activity towards different cellulosic and hemicellulosic substrates.

Materials and methods

Sample collection

Deep-sea sediment samples were collected in Suruga Bay during CK09-01 expedition 903 in March 2009 by the drilling vessel "Chikyu" using the hydraulic piston coring system. Samples were obtained from site C9006A (34°52.4646′N, 138°34.1639′E) and site C9008A (34°51.3846′N, 138°33.7650′E) at a water depth of 755.86 and 730.76 m, respectively. Cores were sliced in sections with an average length of 9 m and sediment samples were taken aseptically from the center of each section. Samples were used directly onboard as inocula for enrichment cultures. Shipping and storage of enrichment cultures and sediments for subsequent DNA extraction and chemical analysis were carried out at 4 °C.

Age determination and analysis of chemical parameters

Deep-sea sediments from sampling site C9006A were used for radiocarbon dating and chemical analysis. Age determination was carried out at Leibniz-Laboratory for Radiometric Dating and Isotope Research, Christian-Albrechts-University (Kiel, Germany) using an accelerator mass spectrometer. Radiocarbon ages were determined for sediment samples taken from a depth of 12 and 41 m below seafloor (mbsf).

Concentration of total organic carbon (TOC) was measured with a TOC analyzer (highTOC, Elementar Analysensysteme GmbH, Hanau, Germany) at the Central Laboratory of Analytical Chemistry, Hamburg University of Technology (Hamburg, Germany).

Enrichment and cultivation of polysaccharide degrading bacteria

Media containing artificial seawater (per liter: 20 g NaCl, 0.34 g KCl, 4 g $MgCl_2 \times 6\ H_2O$, 3.45 g $MgSO_4 \times 7\ H_2O$, 0.25 g NH_4Cl, 0.14 g $CaCl_2 \times 2\ H_2O$, 0.14 g K_2HPO_4, 2.38 g 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), 10 ml trace elements (DSMZ medium 141 and additionally per liter: 1 g KNO_3, 90 mg KBr, 7.5 mg $SrCl_2 \times 6\ H_2O$, 0.1 mg $Na_2WO_4 \times 2\ H_2O$), 10 ml vitamin solution (DSMZ medium 141), 0.2 g $NaHCO_3$, pH 6.9) were used for enrichments. Mixtures of different carbohydrates in equal proportions were utilized as sole carbon source with a final concentration of 1 % (w/v): (1) cellulose, xylan, chitin, (2) starch, glucose, (3) cellulose, xylan, chitin combined with 0.01 % (w/v) yeast extract, and (4) starch and glucose combined with 0.01 % (w/v) yeast extract, respectively. Media were inoculated with 1 % (w/v) of each core section from sampling site C9008A and incubated under aerobic conditions at 15 °C and 130 rpm. After 5 days, 100 µl of the samples were plated on artificial seawater agar containing the appropriate carbon sources. Plates were incubated at 15 °C for 2 weeks before colonies were selected on the basis of different morphological characteristics. Repeated transfers were performed until pure cultures of the strains were obtained. Isolates were routinely cultivated on artificial seawater agar plates at 15 °C and stored at 4 °C.

Stability of enrichments was investigated by transfer of the microbial community to fresh artificial seawater and incubation for 1 week at 15 °C. Inoculation was repeated three times and composition of enrichment cultures was verified by denaturing gradient gel electrophoresis (DGGE).

Polymerase chain reaction and denaturing gradient gel electrophoresis

For phylogenetic analysis, 16S rRNA gene fragments were amplified using primers 27f (5′-AGAGTTTGATCCTGGCTCAG-3′) and 1492r (5′-GGTTACCTTGTTACGACTT-3′) with cell suspension or biomass from colonies as template. PCR reaction mixture contained 50 pmol of each primer, 1–50 ng template DNA, 0.2 mM of each dNTP, 20 mM Tris–HCl, 50 mM KCl, 1.5 mM $MgCl_2$, and 1 U of *Taq* polymerase (Fermentas). The PCR cycling conditions were as follows: 5 min at 94 °C and 30 amplification cycles of 30 s at 94 °C, 1 min at 50 °C, 1.5 min at 72 °C, and a final extension for 10 min at 72 °C.

For DGGE approach, the Bacteria-specific 16S rRNA gene primers GM5f GC clamp (5′-CGCCCGCCGCGCCCCGCGCCCGTCCCGCCGCCCCCGCCCGCCTACGGGAGGCAGCAG-3′) and 907r (5′-CCGTCAATTCMTTTGAGTTT-3′) were used (Muyzer et al. 1998). 100 µl of PCR mixtures contained 10 µl *Taq* polymerase buffer, 1.5 mM $MgCl_2$, 300 µg bovine serum albumin, 0.2 mM of each dNTP, 0.5 µM of each primer, 50 µl template, and 1.5 U *Taq* polymerase (Fermentas). The PCR cycling conditions were as follows: 94 °C for 5 min, 36 cycles of 30 s at 94 °C, 1.5 min at 48 °C, 1 min at 72 °C, and a final extension at 72 °C for 10 min. DGGE was performed as described by Muyzer et al. (1998) using the Dcode™ System (Bio-Rad Laboratories GmbH). Electrophoresis was carried out using a 6 % (v/v) polyacrylamide gel with a denaturing gradient of 20–80 % (v/v) at 55 °C and 200 V for 5 h. PCR products and DGGE gels were visualized by agarose gel electrophoresis using ethidium bromide and photographed on a UV transilluminator.

16S rRNA sequencing and phylogenetic analysis

PCR products were purified using the GeneJet™ Gel Extraction Kit (Fermentas) and quantified spectrophotometrically. Sequencing was carried out by Eurofins MWG Operon (Ebersberg, Germany). The basic local alignment search tool (BLAST) from the National Center for Biotechnology Information (NCBI) was used for sequence similarity analysis (Altschul et al. 1990).

DNA extraction

Original enrichment cultures grown on cellulose, xylan, and chitin and on starch and glucose were pooled and used for total DNA isolation according to Zhou et al. (1996). The method is based on lysis with a high-salt extraction buffer and extended heating in the presence of sodium dodecyl sulfate (SDS). The pellets obtained by centrifugation of the samples at $3{,}939 \times g$ for 1 h were resuspended in DNA extraction buffer, containing 1 % hexadecylmethylammonium bromide (CTAB). Prior to DNA isolation, three freeze-and-thaw cycles in liquid nitrogen and a water bath at 65 °C were conducted. Purity and concentration of DNA were analyzed by agarose gel electrophoresis and spectrophotometrically at 260/280 nm.

Metagenomic sequencing and sequence analysis

Sequencing of total community DNA was performed using 454 GS-FLX Titanium pyrosequencing (Roche Applied Science). Library preparation, sequencing and assembly were carried out by Roche Applied Science (Penzberg, Germany). Annotation of metagenomic sequences was conducted by CeBiTec (Bielefeld University, Germany)

using the genome annotation system GenDB (Meyer et al. 2003). The annotation of selected genes was verified manually using BLAST, InterProScan, CAZYmes Analysis Toolkit (CAT), and the databases for automated Carbohydrate-active enzyme ANnotation (dbCAN) and Carbohydrate-Active enZYmes (CAZy) (Zdobnov and Apweiler 2001; Cantarel et al. 2009; Park et al. 2010; Yin et al. 2012).

Cloning and expression of glycoside hydrolase encoding genes

Genes encoding different glycoside hydrolases were amplified using the primers *bgl3A*.f (5′-ATGATTGGCAACTATTACGGCATCTCTG-3′), *bgl3A*.r (5′-TTAAAATTGTATTTCTTGACGCTGGTGAGC-3′), *bgl3B*.f (5′-ATGAATGACCAACCCTGGTTTGATACC-3′), *bgl3B*.r (5′-TTAGATAACGAACTTGGCTATTTGTGG-3′), *bgl3C*.f (5′-ATGAGCGATTACCAAGACGCG-3′), *bgl3C*.r (5′-CTATTGCACACTCACATTCGTCAG-3′), *bgl3D*.f (5′-ATGAATGTAACCAACACGGCTTCG-3′), *bgl3D*.r (5′-TTACTCAGTCGGGCAACCTATATC-3′), xyn8A.f (5′-ATGGGGCCTAACACAAACTTCGGGTAC-3′), xyn8A.r (5′-TTACTCATTTTCATTGGGTTTGTAAATTTTGAACTCGCC-3′), xyn10A.f (5′-ATGTCAGTCCCTTACGTATTCGAAGC-3′), xyn10A.r (5′-TCAGTCTACAGGCAGCGCAATTAC-3′), man25A.f (5′-ATGGGTGAGAGTCTCAATACAGAATCGGC-3′), man26A.r (5′-TTATTGGGTAAACTTCACATCGTCGATAAAAAATG-3′), xyl43A.f (5′-ATGACTCAGGCCAATCAAAAACAAAGCC-3′), xyl43A.r (5′-CTAGTTACTGTTTTGAAAAGTGCCAGCG-3′), xyl43B.f (5′-ATGGATAACCCTCTTGTTACTCATATG-3′), xyl43B.r (5′-CTACTTAAAGCTGACCCAATCGATGTTG-3′), xyl43C.f (5′-ATGTGTACAGAACAAGTTAGTCATAAAG-3′), xyl43C.r (5′-TTACTTGTAGTAAGCCTGATAAGGC-3′), and genomic DNA of *Glaciecola* sp. 4H-3-7+YE-5 as template (Klippel et al. 2011). For AT cloning the pETBlue-1 AccepTor™ Vector Kit (Novagen) was used according to the instructions of the manufacturer. Gene expression was induced with 1 mM IPTG after cells reached an optical density of 0.6 nm at 37 °C, and cells were harvested after cultivation over night at 15 °C. Cell disruption was carried out by French press using 20 mM sodium phosphate buffer, pH 6.8 (5 ml per 1 g cell wet weight).

Determination of enzymatic activity

Activity of endoxylanase and endomannanase was investigated using 0.5 % (w/v) beechwood xylan or locust bean gum as substrates and 20 mM sodium phosphate buffer (pH 7.0). Incubation was started after addition of crude extract and carried out for 30 min at 30 °C. Reaction was stopped by the addition of 3,5-dinitrosalicylic acid (DNS) and boiling of the samples for 5 min (Bernfeld 1955). The formation of reducing sugar ends was measured spectrophotometrically at 546 nm.

β-Glucosidase and β-xylosidase activity were measured using different α- and β-linked aryl-glycosides as substrates. The reaction mixture contained 2 mM substrate, 20 mM sodium phosphate buffer (pH 7.0), and crude extract and was incubated for 25 min at 30 °C. After the addition of 10 mM Na_2CO_3 for stopping the reaction, the release of *p*-nitrophenol was measured at 410 nm.

Accession number

The complete metagenomic sequence has been submitted to GenBank under the accession number SRP041194.

Results and discussion

Sample characteristics

Deep-sea sediment samples were taken in Suruga Bay, the deepest bay of Japan with a maximum depth of about 2,500 m at the bay mouth (Matsuyama et al. 1993). The sampling took place at a water depth of about 731 and 756 m and two cores of an overall length of 52 and 66 m were collected using the deep-sea drilling vessel “Chikyu”. Drilling cores were sliced in sections of approximately 9 m and sediment samples were taken for various analyses. Based on ^{14}C dating sediment age was between 5,155 (±158) years before present (BP) at a depth of 12 mbsf and 6,321 (±548) years BP for samples taken at 41 mbsf. These data are consistent with the results obtained by Nakamura et al. (1990), who have measured radiocarbon ages of different fractions of a sediment core taken from Suruga Trough. They determined a ^{14}C age of 2,270 (±90) years BP for a sandy mud fraction from a depth of 0.220–0.270 mbsf.

Analysis of the TOC content revealed a concentration of about 0.5 % (w/w), showing little variation along the depth profile. Equal results with an average TOC content of 0.4–0.5 % were measured for sediment samples from the Nankai Trough, which is closely located to Suruga Bay (Reed et al. 2002). Furthermore, TOC concentrations <1 % were observed in deep-sea sediments from the southwestern Gulf of Mexico, the Arabian Sea and the continental margin of Chile, whereas a decrease of the organic carbon content occurs with an increase in sediment depth and distance from land (Littke et al. 1995; Schnetger et al. 2000; Jørgensen and Boetius 2007; Escobar-Briones and Garcia-Villalobos 2009).

Table 1 Phylogenetic affiliation of isolates obtained on complex carbon sources

Depth (mbsf)	Closest BLAST hit	Carbon source
11.7–13.1	*Rheinheimera perlucida*	SG CXC
	Pseudomonas sp.	
21.2–22.6	*Colwellia* sp.	SG CXC
	Maribacter sp.	
30.7–32.1	*Phaeobacter arcticus*	SG
	Pseudomonas sp.	SG
	Colwellia sp.	CXC
	Krokinobacter diaphorus	CXC
30.7–32.1	*Pseudomonas* sp.	SG + 0.01 % YE
	Glaciecola agarilytica	CXC + 0.01 % YE
	Sulfitobacter sp.	CXC + 0.01 % YE
40.3–41.7	*Lacinutrix algicola*	CXC

YE yeast extract

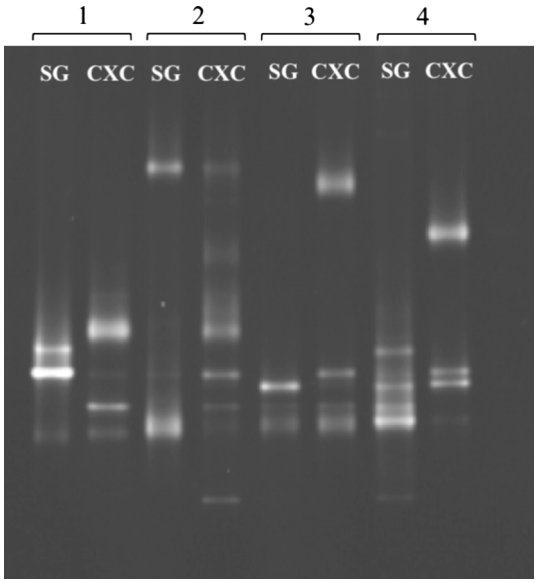

Fig. 1 Diversity of enrichment cultures grown on starch and glucose (SG) or on cellulose, xylan, and chitin (CXC) along the depth profile. *1* 11.7–13.1 mbsf, *2* 21.2–22.6 mbsf, *3* 30.7–32.1 mbsf, *4* 30.7–32.1 mbsf + 0.01 % yeast extract

Enrichment on complex carbon sources

Aerobic enrichment cultures were started on board ship. Complex carbon sources such as cellulose, xylan, and chitin (CXC), and a starch-glucose mixture (SG) were used as substrates. Samples from different sediment depths from hole C9008A were inoculated. The influence of 0.01 % of yeast extract in the enrichment medium was tested for two selected depths. After 1 week of incubation at 15 °C, growth was observed in enrichments from sediment depths between 12 and 42 mbsf. At 42 mbsf growth was only detected using CXC as carbon source. Independent of the substrate, no growth was observed in flasks inoculated with sediment derived from 3 to 4 mbsf and in anaerobic enrichments (data not shown).

Phylogenetic analysis based on sequencing of the almost complete 16S rRNA-encoding gene revealed that isolates shared 97–100 % sequence similarity to species belonging to *Gammaproteobacteria* (*Pseudomonas* sp., *Glaciecola* sp., *Oleispira* sp.), *Flavobacteria* (*Krokinobacter* sp., *Polaribacter* sp., *Maribacter* sp., *Flavobacteria* sp.), and *Alphaproteobacteria* (*Sulfitobacter* sp., *Loktanella* sp., *Phaeobacter* sp.) (Table 1). Most of these species are widely distributed in marine sediments and are well characterized (Nogi et al. 2004; Cho et al. 2008; Li et al. 2009, 2011; Srinivas et al. 2013). The dominance of members of *Gammaproteobacteria* and *Bacteroidetes* was also observed for a biofilm developed on a cotton string incubated in the Irish Sea. The utilization of cellulose containing yarn as carbon source resulted in the enrichment of species of *Sulfitobacter* and *Glaciecola* (Edwards et al. 2010). The ability to degrade polymers such as cellulose and chitin was shown for a strain related to *Sulfitobacter pontiacus*, which was isolated from organic particles of the upper water column of the Equatorial Atlantic (Berkenheger and Fischer 2004). Furthermore, the degradation of xylan is known for *Glaciecola mesophila* producing two active endoxylanases (Guo et al. 2009, 2013).

Microbial diversity within enrichments was investigated by denaturing gradient gel electrophoresis (DGGE) to compare the influence of sediment depth, and parameters such as carbon source and availability of yeast extract. Figure 1 depicts the DGGE analysis of PCR amplified 16S rRNA-encoding genes from the different enrichment cultures along the depth profile. While some bands could be detected in almost all samples, others were only detected at specific depths and carbon sources. Diversity was higher in enrichments using CXC than in those using SG, if no yeast extract was added. Addition of yeast extract enhanced diversity in case of SG and lowered it in case of CXC. The microbial consortium that had established in the first enrichment step was stable during three transfers as could be seen by DGGE analysis, shown for the enrichment on CXC (Fig. 2). The influence of different carbon sources on the diversity of enrichment cultures was also shown by Toffin et al. (2004). They enriched sediment samples from Nankai Trough on media with varying carbon sources and concentrations, yielding a higher number and intensity of bands on DGGE profile with increasing carbon concentrations. It is assumed that fast-growing bacteria were dominant in these enrichment cultures, since mostly monomeric sugars were used as carbon source. However,

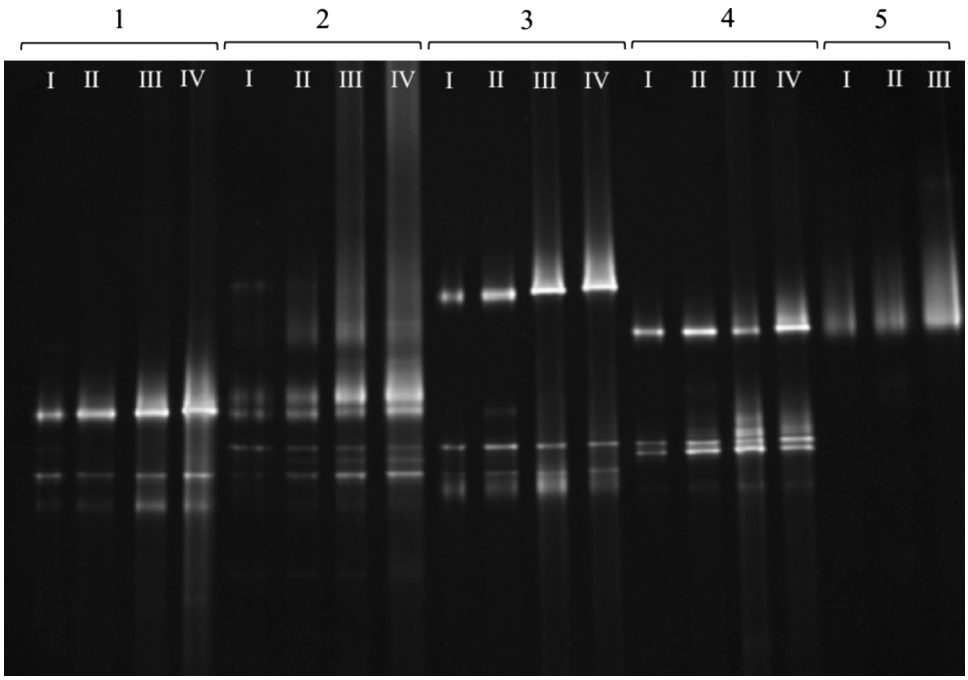

Fig. 2 Microbial diversity in enrichment cultures grown on cellulose, xylan, and chitin during transfer. *I*, *II*, *III*, *IV* first, second, third and fourth transfer. *1* 11.7–13.1 mbsf, *2* 21.2–22.6 mbsf, *3* 30.7–32.1 mbsf, *4* 30.7–32.1 mbsf + 0.01 % yeast extract, *5* 40.3–41.7 mbsf

for the degradation of polysaccharides such as cellulose and chitin the production of extracellular enzymes is necessary, resulting in longer generation times of enrichment cultures from deep-sea sediments (Boetius and Lochte 1996; Kanzog et al. 2008). In order to prevent the dominance of fast-growing bacteria and to enhance the growth of polymeric carbohydrate-degrading bacteria, we have used complex carbohydrates as sole carbon source in almost all enrichment cultures.

Metagenomic sequencing and analyses

Cultures from the first enrichment step were pooled for DNA extraction and metagenomic sequencing. Using 454-GS FLX Titanium technique approximately 1.694 Gbp of sequence data were generated from four runs. Metagenomic sequencing resulted in 4,264,168 high quality reads with an average sequence length of about 489 bp. Data could be assembled into 45,769 contigs (>100 bp) with 27,106 large contigs (>500 bp) by using the GS De Novo Assembler. Annotation was done automatically by the GenDB annotation system for prokaryote genomes and a total of 52,660 coding sequences (CDS) could be obtained.

Approximately 19,047 of the total reads (0.45 %) contained 16S rRNA-encoding gene fragments of which 24 % could be identified to a genus level. The metagenomic data revealed that we had isolated representatives from eight of the most common genera in our enrichment approaches using complex carbon sources (Supplementary Table 1). The majority of the 16S rRNA-encoding sequences show affiliation of the identified bacteria to *Gammaproteobacteria* (56.3 %) followed by *Alphaproteobacteria* (24.5 %) and *Flavobacteria* (17.2 %). Even though we used enrichment cultures for the determination of the community composition, a similar pattern was observed for deep-sea sediments taken from unmodified volcanic ash layers of the Sea of Okhotsk at a depth of 18.3–45.7 mbsf and sediments of Nankai Trough at 1–98 mbsf. In contrast, sediments of a deeper depth of Nankai Trough (194 mbsf) and samples of pelagic clay layers from the Sea of Okhotsk (7.5–57.8 mbsf) harbor only few or no representatives of *Gamma-* and *Alphaproteobacteria* (Inagaki et al. 2003; Kormas et al. 2003; Fry et al. 2008).

Prediction of carbohydrate-active enzymes

Due to the fact that little is known about carbohydrate-active enzymes from organisms of marine sediments, we carried out a detailed analysis of annotated genes encoding putative glycoside hydrolases. First of all, initial analyses revealed the presence of 17,458 genes coding for enzymes belonging to the six Enzyme Commission (EC) groups established by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB). This corresponds to 33.2 % of the total dataset. Most of the sequences (10.1 %) are encoding putative transferases (EC 2), followed by 9.4 % coding sequences for hydrolases (EC 3) and 6.7 % for oxidoreductases (Supplementary Table 2). A closer look at the predicted open reading frames encoding hydrolases indicated the

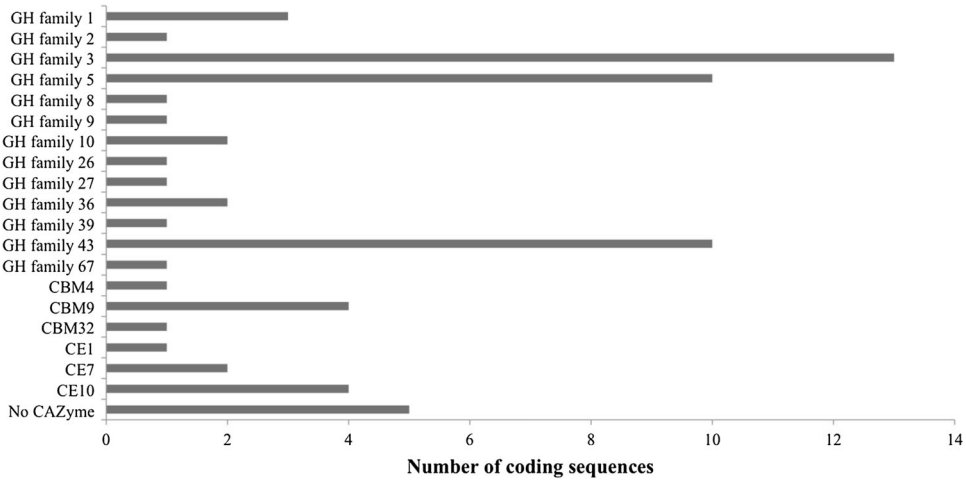

Fig. 3 Distribution of genes encoding carbohydrate-active enzymes. Genes encoding enzymes with predicted carbohydrate-degrading activity were grouped into families of glycoside hydrolases (GH), carbohydrate-binding modules (CBM), and carbohydrate esterases (CE)

presence of 324 genes coding for glycosylases (EC 3.2.), and 207 of them are encoding putative glycoside hydrolases (EC 3.2.1.). A detailed overview of the glycoside hydrolases identified in the metagenomic dataset is given in Supplementary Table 3.

Focussing on glycoside hydrolases relevant for lignocellulose degradation, we identified 65 genes encoding putative carbohydrate-active enzymes. Manual verification, CAT and dbCAN analyses confirmed the existence of catalytic domains and carbohydrate-binding modules (CBM) for 60 out of 65 of the predicted enzymes (Fig. 3). This discrepancy can be explained by the application of different databases for the annotation and verification since they are using different strategies and pursue different objectives. Therefore, different results will be obtained by the analysis carried out with one method compared to the results done with another method (Lombard et al. 2014).

Among the enzymes predicted to be involved in cellulose- and hemicellulose-degradation, members of 11 glycoside hydrolase, three CBMs and three carbohydrate esterase (CE) families were determined (Fig. 3). Most of them were grouped in families 3, 5, and 43.

Glycoside hydrolase family 3 mainly consists of exo-acting enzymes and most of the annotated β-glucosidases (6.3 % of total glycoside hydrolases) from the metagenomic sequences belong to this family. As shown in our dataset, a high abundance of β-glucosidases was also observed in metagenomic libraries prepared from decomposing leaf litter (Nyyssönen et al. 2013). Family 5 is one of the largest families and contains primarily cellulases, and is also the most common family being found for the endoglucanases (3.9 %) of this study. Most of the identified β-xylosidases (3.4 %) are grouped in family 43, a glycoside hydrolase family with known activities of α-arabinofuranidases, β-xylosidases, and arabinases (Cantarel et al. 2009; Lombard et al. 2014). As expected from the results of our enrichments, similarities of the predicted enzymes to proteins identified in the genome of *Maribacter*, *Pseudomonas*, *Glaciecola*, and further marine bacteria were observed by BLAST searches with identities ranging from 35 to 100 % to hypothetical proteins and putative glycoside hydrolases.

Even though using starch as carbon source for the enrichment, coding sequences for α-glucosidases were the only ones predicted to be involved in starch degradation. Contrary to our expectations, no genes encoding α-amylase, β-amylase and glucoamylase could be found. Although coding sequences for amylases were observed when manually searching the complete genomes of *Colwellia* and *Rheinheimera* (Methe et al. 2005; Gupta et al. 2011), it remains unclear why no genes encoding starch-degrading enzymes were identified in the metagenomic dataset. However, a total of 15 CDS for chitin-degrading β-acetylhexosaminidases were detected among the metagenomic sequences.

Cloning of selected glycoside hydrolases

For preliminary investigations of the hydrolytic activity of putative carbohydrate-active enzymes identified by metagenomic sequencing and analyses, selected genes were chosen for cloning and expression. BLAST analyses of the amino acid sequences of proteins predicted to be involved in biomass degradation revealed a high percentage of glycoside hydrolases encoded by the genome of *Glaciecola* sp. 4H-3-7+YE-5, a strain isolated from our enrichment cultures with CXC as carbon source (Klippel et al. 2011). As represented by the CAZy database, a total of 61 proteins were annotated as glycoside hydrolases (29.5 % of the total glycoside hydrolases of the metagenomic sequence), and almost one-third of the β-glucosidases and nearly half of the β-xylosidases identified in the metagenomic dataset are encoded by the genome of *Glaciecola* sp. 4H-3-7+YE-5. In order to gain more insight in the carbohydrate-active

Table 2 Identities of selected glycoside hydrolases to characterized enzymes based on BLAST analyses

Gene	Name	Size (aa)	Closest hit to biochemically characterized enzyme, GenBank Accession No.	Identity (%)	References
Glaag_0892	Bgl3A	325	β-Xylosidase B, *Clostridium stercorarium*, CAD48309.1	36	Adelsberger et al. (2004)
Glaag_3503	Bgl3B	733	β-Xylosidase B, *Clostridium stercorarium*, CAD48309.1	47	Adelsberger et al. (2004)
Glaag_3721	Bgl3C	803	ND	ND	
Glaag_3726	Bgl3D	856	CelD 1,4-β-D-glucan glucohydrolase, *Cellvibrio japonicus*, CAA46499.1	51	Rixon et al. (1992)
Glaag_3338	Xyn8A	596	Family 8 endo-beta-xylanase, *Glaciecola mesophila*, AEC33258.1	85	Guo et al. (2013)
Glaag_1620	Xyn10A	423	Family 10 endo-beta-xylanase, *Glaciecola mesophila* KMM 241, ACN76857.1	76	Guo et al. (2009)
Glaag_3297	Man26A	1,063	Mannanase 26B, *Clostridium thermocellum*, BAB19050.1	39	Kurokawa et al. (2001)
Glaag_2342	Xyl43A	361	α-Neoagarobiose hydrolase, *Saccharophagus degradans* 2–40, ABD81917.1	70	Ha et al. (2011)
Glaag_3521	Xyl43B	450	ND	ND	
Glaag_3719	Xyl43C	383	β-Xylosidase, uncultured organism, AFP23142.1	60	Lee et al. (2013)

ND not detected

enzymes of *Glaciecola* sp. involved in cellulose and hemicellulose degradation, we chose 10 enzymes for preliminary activity determinations. Identities up to 99 % to glycoside hydrolases of different *Glaciecola* strains were observed, but with the exception of two xylanases of *Glaciecola mesophila* none of these proteins have been characterized biochemically (Guo et al. 2009, 2013). By recombinant expression of the respective genes we also intended to verify sequence-based prediction and functionality of the metagenomic sequences.

Three of the β-glucosidases of *Glaciecola* sp. 4H-3-7+YE-5 showed 36–51 % identity to characterized glycoside hydrolases of *Clostridium stercorarium* and *Cellvibrio japonicus*, but no characterized protein with identities >30 % was found for Bgl3C using BLAST searches (Table 2). The recombinant β-glucosidases were tested for activity toward different α- and β-linked aryl glycosides and all of them showed activity on 4-Nitrophenyl α-D-glucopyranoside (Fig. 4). Bgl3A and Bgl3D exhibited no activity against further substrates, indicating that these enzymes should be regarded as α-glucosidases rather than as β-glucosidases. Besides the α-glucosidase activity of Bgl3B and Bgl3C, β-xylosidase/α-arabinofuranosidase activity was observed for both enzymes and additionally, cellobiohydrolase and β-glucosidase activity was detected for Bgl3C. Interestingly, only one of the four predicted β-glucosidases of GH family 3 showed β-glucosidase activity.

Furthermore, one endomannanase, two endoxylanases, and three β-xylosidases were produced recombinantly and BLAST searches revealed identities ranging from 39 to 85 % to functionally active glycoside hydrolases (Table 2). The activity of the recombinant enzymes was tested toward locust bean gum, beechwood xylan, 4-Nitrophenyl β-D-xylopyranoside, and 4-Nitrophenyl α-L-arabinofuranoside. A degradation of locust bean gum was measured after incubation with Man26A, confirming the endomannanase activity of this enzyme, whereas endoxylanase activity was observed only for Xyn10A (data not shown). Among the three β-xylosidases, activity toward 4-Nitrophenyl β-D-xylopyranoside and 4-Nitrophenyl α-L-arabinofuranoside was detected for one of the investigated enzymes. In addition to the hydrolysis of the aryl glycosides, Xyl43C was able to hydrolyze beechwood-xylan (Fig. 5).

Most glycoside hydrolases listed in the NCBI and CAZy database are identified by sequencing of (meta)genomic DNA, and in the majority of cases their function and functionality is unknown. We could show in this study, that sequence-homology-based prediction was misleading in particular in the case of β-glucosidases and that additional functions could be revealed by activity determination. The fact that none of the selected genes were silent and that they could all be successfully expressed in *E. coli* is promising for future studies.

The existence of many additional genes encoding carbohydrate-active enzymes identified in the metagenome of subsurface marine sediments demonstrates the great potential of metagenomic approaches with prior enrichments utilizing selective carbon sources.

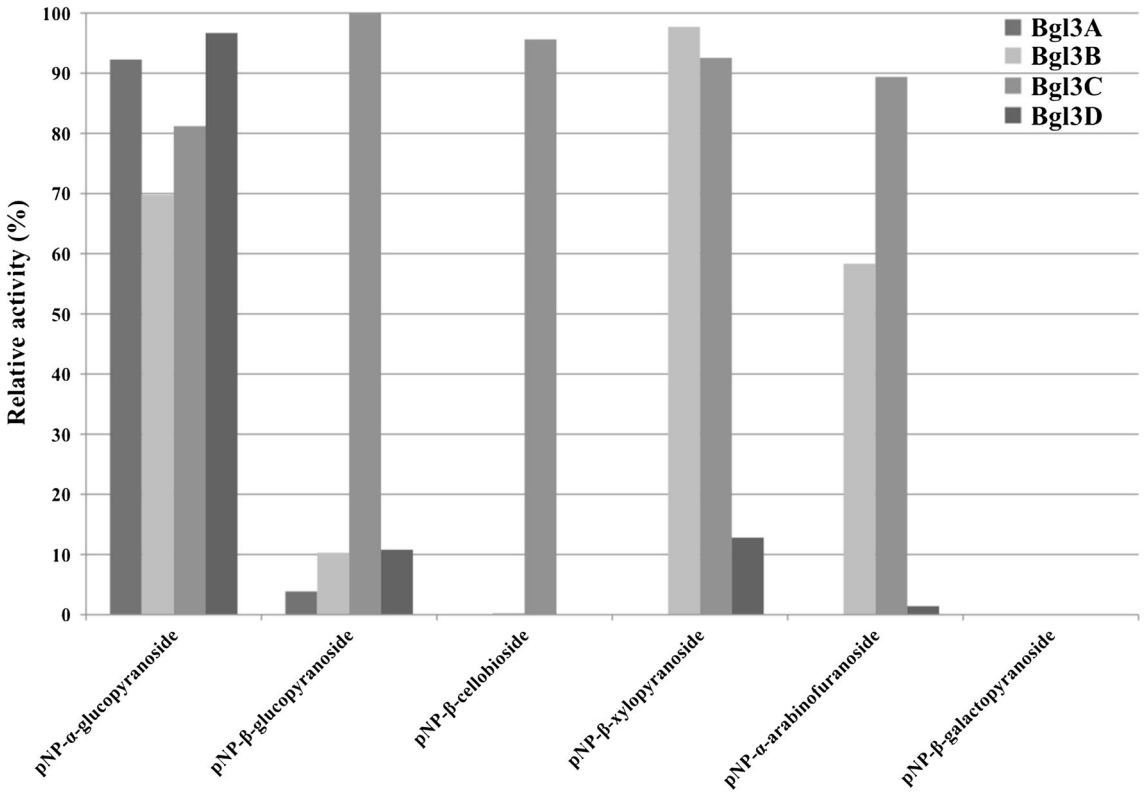

Fig. 4 Preliminary activity determination of recombinant β-glucosidases. Hydrolysis of various substrates (2 mM) was carried out for 25 min at 30 °C in 20 mM sodium phosphate buffer (pH 7.0). Release of nitrophenol was determined spectrophotometrically at 410 nm

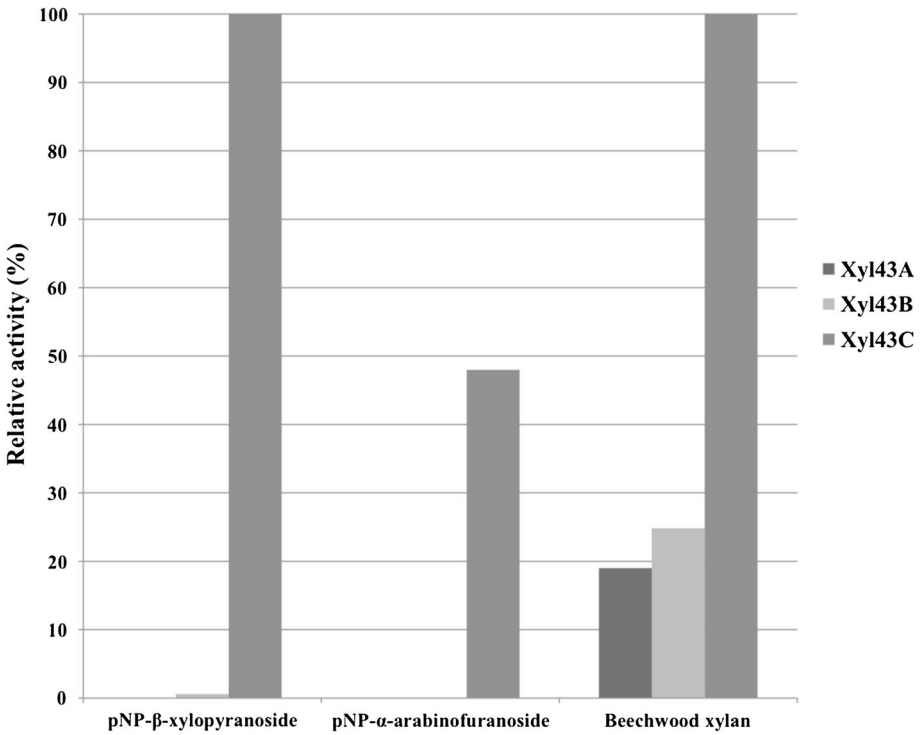

Fig. 5 Preliminary activity determination of recombinant β-xylosidases. Hydrolysis of aryl glycosides (2 mM) and beechwood xylan was carried out for 1 h and 30 min at 30 °C in 20 mM sodium phosphate buffer (pH 7.0), respectively. Release of nitrophenol and reducing sugar ends was determined spectrophotometrically at 410 and 546 nm

References

Adelsberger H, Hertel C, Glawischnig E, Zverlov VV, Schwarz WH (2004) Enzyme system of *Clostridium stercorarium* for hydrolysis of arabinoxylan: reconstitution of the in vivo system from recombinant enzymes. Microbiology 150(Pt 7):2257–2266

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410

Berkenheger I, Fischer U (2004) Competition for polymers among heterotrophic bacteria, isolated from particles of the equatorial Atlantic. Int Microbiol 7:13–18

Berlemont R, Delsaute M, Pipers D, D'Amico S, Feller G, Galleni M, Power P (2009) Insights into bacterial cellulose biosynthesis by functional metagenomics on antarctic soil samples. ISME J 3(9):1070–1081

Bernfeld P (1955) Amylases, alpha and beta. Methods Enzymol 1:149–158

Boetius A, Lochte K (1996) Effect of organic enrichments on hydrolytic potentials and growth of bacteria in deep-sea sediments. Mar Ecol Prog Ser 140:239–250

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res 37(Database issue):D233–D238

Cho KH, Hong SG, Cho HH, Lee YK, Chun J, Lee HK (2008) *Maribacter arcticus* sp. nov., isolated from arctic marine sediment. Int J Syst Evol Microbiol 58(Pt 6):1300–1303

Edwards JL, Smith DL, Connolly L, McDonald JE, Cox MJ, Joint I, Edwards C, McCarthy AJ (2010) Identification of carbohydrate metabolism genes in the metagenome of a marine biofilm community shown to be dominated by *Gammaproteobacteria* and *Bacteroidetes*. Genes 1:371–384

Escobar-Briones E, Garcia-Villalobos FJ (2009) Distribution of total organic carbon and total nitrogen in deep-sea sediments from the southwestern gulf of Mexico. Boletin de la Sociedad Geol Mexicana 61(1):73–86

Fry JC, Parkes RJ, Cragg BA, Weightman AJ, Webster G (2008) Prokaryotic biodiversity and activity in the deep subseafloor biosphere. FEMS Microbiol Ecol 66(2):181–196

Glöckner FO, Joint I (2010) Marine microbial genomics in Europe: current status and perspectives. Microb Biotechnol 3(5):523–530

Guo B, Chen XL, Sun CY, Zhou BC, Zhang YZ (2009) Gene cloning, expression and characterization of a new cold-active and salt-tolerant endo-beta-1,4-xylanase from marine *Glaciecola mesophila* KMM 241. Appl Microbiol Biotechnol 84(6):1107–1115

Guo B, Li PY, Yue YS, Zhao HL, Dong S, Song XY, Sun CY, Zhang WX, Chen XL, Zhang XY, Zhou BC, Zhang YZ (2013) Gene cloning, expression and characterization of a novel xylanase from the marine bacterium, *Glaciecola mesophila* KMM241. Mar Drugs 11(4):1173–1187

Gupta HK, Gupta RD, Singh A, Chauhan NS, Sharma R (2011) Genome sequence of *Rheinheimera* sp. strain A13l, isolated from Pangong Lake, India. J Bacteriol 193(20):5873–5874

Ha SC, Lee S, Lee J, Kim HT, Ko HJ, Kim KH, Choi IG (2011) Crystal structure of a key enzyme in the agarolytic pathway, alpha-neoagarobiose hydrolase from *Saccharophagus degradans* 2–40. Biochem Biophys Res Commun 412(2):238–244

Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J 280(Pt 2):309–316

Henrissat B, Coutinho PM (2001) Classification of glycoside hydrolases and glycosyltransferases from hyperthermophiles. Methods Enzymol 330:183–201

Inagaki F, Suzuki M, Takai K, Oida H, Sakamoto T, Aoki K, Nealson KH, Horikoshi K (2003) Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk. Appl Environ Microbiol 69(12):7224–7235

Jørgensen BB, Boetius A (2007) Feast and famine-microbial life in the deep-sea bed. Nat Rev Microbiol 5(10):770–781

Kanzog C, Ramette A, Queric NV, Klages M (2008) Response of benthic microbial communities to chitin enrichment: an in situ study in the deep Arctic Ocean. Polar Biol 32:105–112

Klippel B, Lochner A, Bruce DC, Davenport KW, Detter C, Goodwin LA, Han J, Han S, Land ML, Mikhailova N, Nolan M, Pennacchio L, Pitluck S, Tapia R, Woyke T, Wiebusch S, Basner A, Abe F, Horikoshi K, Keller M, Antranikian G (2011) Complete genome sequence of the marine cellulose- and xylan-degrading bacterium *Glaciecola* sp. strain 4H-3-7 + YE-5. J Bacteriol 193(17):4547–4548

Kormas KA, Smith DC, Edgcomb V, Teske A (2003) Molecular analysis of deep subsurface microbial communities in nankai trough sediments (ODP leg 190, site 1176). FEMS Microbiol Ecol 45(2):115–125

Kurokawa J, Hemjinda E, Arai T, Karita S, Kimura T, Sakka K, Ohmiya K (2001) Sequence of the *Clostridium thermocellum* mannanase gene *man26b* and characterization of the translated product. Biosci Biotechnol Biochem 65(3):548–554

Lee CC, Braker JD, Grigorescu AA, Wagschal K, Jordan DB (2013) Divalent metal activation of a GH43 beta-xylosidase. Enzyme Microb Technol 52(2):84–90

Li H, Yu Y, Luo W, Zeng Y, Chen B (2009) Bacterial diversity in surface sediments from the Pacific Arctic Ocean. Extremophiles 13(2):233–246

Li HJ, Zhang XY, Zhang YJ, Zhou MY, Gao ZM, Chen XL, Dang HY, Zhang YZ (2011) *Rheinheimera nanhaiensis* sp. nov., isolated from marine sediments, and emended description of the genus *Rheinheimera* Brettar et al. 2002 emend. Merchant et al. 2007. Int J Syst Evol Microbiol 61(Pt 5):1016–1022

Littke R, Disko U, Rullkötter J (1995) Characterization of organic matter in deep-sea sediments on the Chile continental margin with special emphasis on maturation in an area of high geothermal heat flow. In: Lewis SD, Behrmann JH, Musgrave RJ, Cande SC (eds) Proceedings of the Ocean Drilling Program, Scientific Results. http://www-odp.tamu.edu/publications/141_SR/141TOC.HTM

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42(Database issue):D490–D495

Matsuyama M, Ohta S, Hibiya T, Yamada H (1993) Strong tidal currents observed near the bottom in the Suruga Trough, Central Japan. J Oceanogr 49:683–696

Methe BA, Nelson KE, Deming JW, Momen B, Melamud E, Zhang X, Moult J, Madupu R, Nelson WC, Dodson RJ, Brinkac LM, Daugherty SC, Durkin AS, DeBoy RT, Kolonay JF, Sullivan SA, Zhou L, Davidsen TM, Wu M, Huston AL, Lewis M, Weaver B, Weidman JF, Khouri H, Utterback TR, Feldblyum TV, Fraser CM (2005) The psychrophilic lifestyle as revealed by the genome sequence of *Colwellia psychrerythraea* 34H through genomic and proteomic analyses. Proc Natl Acad Sci USA 102(31):10913–10918

Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, Kalinowski J, Linke B, Rupp O, Giegerich R, Puhler A (2003) GenDB-an open source genome annotation system for prokaryote genomes. Nucleic Acids Res 31(8):2187–2195

Muyzer G, Brinkhoff T, Nübel U, Santegoeds C, Schäfer H, Wawer C (1998) Denaturing gradient gel electrophoresis (DGGE) in microbial ecology. In: Akkermans ADL, van Elsas JD, de Bruijn FJ (eds) Molecular microbial ecology manual. Kluwer Academic Publishers, London, pp 3.4.4.1–3.4.4.27

Nakamura T, Shiki T, Nakai N (1990) Variations in ^{14}C ages of various organic fractions in a turbidite sediment core from Suruga Trough. Geochem J 24:47–56

Nogi Y, Hosoya S, Kato C, Horikoshi K (2004) *Colwellia piezophila* sp. nov., a novel piezophilic species from deep-sea sediments of the Japan Trench. Int J Syst Evol Microbiol 54(Pt 5):1627–1631

Nyyssönen M, Tran HM, Karaoz U, Weihe C, Hadi MZ, Martiny JB, Martiny AC, Brodie EL (2013) Coupled high-throughput functional screening and next generation sequencing for identification of plant polymer decomposing enzymes in metagenomic libraries. Front Microbiol 4:282

Park BH, Karpinets TV, Syed MH, Leuze MR, Uberbacher EC (2010) CAZymes analysis toolkit (CAT): web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. Glycobiology 20(12):1574–1584

Parkes RJ, Cragg BA, Wellsbury P (2000) Recent studies on bacterial populations and processes in marine sediments. Hydrogeol J 8:11–28

Qin QL, Li Y, Zhang YJ, Zhou ZM, Zhang WX, Chen XL, Zhang XY, Zhou BC, Wang L, Zhang YZ (2011) Comparative genomics reveals a deep-sea sediment-adapted life style of *Pseudoalteromonas* sp. SM9913. ISME J 5(2):274–284

Reed DW, Fujita Y, Delwiche ME, Blackwelder DB, Sheridan PP, Uchida T, Colwell FS (2002) Microbial communities from methane hydrate-bearing deep marine sediments in a forearc basin. Appl Environ Microbiol 68(8):3759–3770

Rixon JE, Ferreira LM, Durrant AJ, Laurie JI, Hazlewood GP, Gilbert HJ (1992) Characterization of the gene *celD* and its encoded product 1,4-beta-D-glucan glucohydrolase D from *Pseudomonas fluorescens* subsp. *cellulosa*. Biochem J 285(Pt 3):947–955

Schnetger B, Brumsack H-J, Schale H, Hinrichs J, Dittert L (2000) Geochemical characteristics of deep-sea sediments from the Arabian Sea: a high-resolution study. Deep Sea Res Part II 47:2735–2768

Srinivas TN, Prasad S, Manasa P, Sailaja B, Begum Z, Shivaji S (2013) *Lacinutrix himadriensis* sp. nov., a psychrophilic bacterium isolated from a marine sediment, and emended description of the genus *Lacinutrix*. Int J Syst Evol Microbiol 63(Pt 2):729–734

Toffin L, Webster G, Weightman AJ, Fry JC, Prieur D (2004) Molecular monitoring of culturable bacteria from deep-sea sediment of the Nankai Trough, Leg 190 Ocean Drilling Program. FEMS Microbiol Ecol 48(3):357–367

Wierzbicka-Wos A, Bartasun P, Cieslinski H, Kur J (2013) Cloning and characterization of a novel cold-active glycoside hydrolase family 1 enzyme with beta-glucosidase, beta-fucosidase and beta-galactosidase activities. BMC Biotechnol 13:22

Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 40(Web Server issue):W445–W451

Zdobnov EM, Apweiler R (2001) Interproscan—an integration platform for the signature-recognition methods in interpro. Bioinformatics 17(9):847–848

Zhang C, Kim SK (2010) Research and application of marine microbial enzymes: status and prospects. Mar Drugs 8(6):1920–1934

Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. Appl Environ Microbiol 62(2):316–322