Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges

Janine Kamke1, Alexander Sczyrba2,3, Natalia Ivanova2, Patrick Schwientek2, Christian Rinke2, Kostas Mavromatis2, Tanja Woyke2 and Ute Hentschel1

1Department of Botany II, Julius-von-Sachs Institute for Biological Sciences, University of Wuerzburg, Wuerzburg, Germany; 2Department of Energy Joint Genome Institute, Walnut Creek, CA, USA and 3Center for Biotechnology, Bielefeld University, Bielefeld, Germany

Many marine sponges are hosts to dense and phylogenetically diverse microbial communities that are located in the extracellular matrix of the animal. The candidate phylum Poribacteria is a predominant member of the sponge microbiome and its representatives are nearly exclusively found in sponges. Here we used single-cell genomics to obtain comprehensive insights into the metabolic potential of individual poribacterial cells representing three distinct phylogenetic groups within Poribacteria. Genome sizes were up to 5.4 Mbp and genome coverage was as high as 98.5%. Common features of the poribacterial genomes indicated that heterotrophy is likely to be of importance for this bacterial candidate phylum. Carbohydrate-active enzyme database screening and further detailed analysis of carbohydrate metabolism suggested the ability to degrade diverse carbohydrate sources likely originating from seawater and from the host itself. The presence of uronic acid degradation pathways as well as several specific sulfatases provides strong support that Poribacteria degrade glycosaminoglycan chains of proteoglycans, which are important components of the sponge host matrix. We therefore propose that Poribacteria are well adapted to an existence in the sponge extracellular matrix. Poribacteria may be viewed as efficient scavengers and recyclers of a particular suite of carbon compounds that are unique to sponges as microbial ecosystems.

The ISME Journal (2013) 7, 2287–2300; doi:10.1038/ismej.2013.111; published online 11 July 2013

Subject Category: Microbe-microbe and microbe-host interactions

Keywords: marine sponge; symbiont; carbohydrate degradation; extracellular matrix; single-cell genomics

Introduction

Marine sponges (phylum Porifera) are the most ancient extant metazoans with fossil records, indicating their emergence more than 600 million years ago (Love et al., 2009). These animals are sessile filter feeders with an enormous filtering capacity that is known to affect nutrient concentrations in the surrounding environment (Gili and Coma, 1998; Maldonado et al., 2005, 2012). In addition to their evolutionary and ecological significance, sponges have attracted recent scientific attention owing to their specific and unique microbiology (Hentschel et al., 2012). The microbial biomass in sponges is located in the extracellular matrix, the so-called ‘mesohyl’, and can make up 35% of the sponge body mass (Vacelet, 1975). Collectively, representatives of more than 30 bacterial phyla and both archaeal lineages have so far been found in sponges from various geographic locations (Webster et al., 2010; Schmitt et al., 2012; Simister et al., 2012). The microbial diversity of marine sponges is well investigated (Taylor et al., 2007), and the collective repertoire of ‘omics’ approaches has been instrumental to shed light on the functional genomic traits of the collective sponge microbiome (Thomas et al., 2010; Fan et al., 2012; Liu et al., 2012; Radax et al., 2012). However, community-wide approaches do not provide sufficient information about functions of specific symbiont clades. Providing a thorough understanding of symbiont function is further complicated by the fact that many sponge symbiont lineages remain uncultivated, such as for the many candidate phyla found in these animals (Schmitt et al., 2012).

One such candidate phylum, termed Candidate phylum Poribacteria, was originally discovered and...
Carbohydrate degradation in Poribacteria

J Kamke et al

Carbohydrate degradation in Poribacteria

The ISME Journal

Materials and methods

Sample collection and processing

Samples of the marine sponge A. aerophoba were collected in September 2009 by scuba diving to a depth of 5–12 m at the Coast of Rovinj, Croatia (45°08’N, 13°64’E). The animals were transported to the University of Wuerzburg (Wuerzburg, Germany) and kept in seawater aquaria until further processing within 1 week of collection. Fresh sponge samples were used for extraction of sponge-associated prokaryotes using an established protocol of tissue disruption, density centrifugation and filtration by Fieseler et al. (2004).

Single-cell sorting, whole genome amplification and PCR screening

Single-cell isolations were conducted with freshly extracted and purified sponge-associated prokaryotes using the fluorescence-activated cell sorting Vantage SE flow cytometer with FACSDiVa option (Becton Dickinson, Heidelberg, Germany) as described previously (Siegl and Hentschel, 2010). For cell lysis and whole genome amplification (WGA) by multiple displacement amplification, we followed the same procedure as Siegl et al. (2011). To identify phylogenetically WGA products and check for possible contamination, we screened the WGA products obtained by polymerase chain reaction (PCR) using 16S or 18S rRNA gene primers targeting Eubacteria, Archaea, Poribacteria and Eukaryotes, as described previously (Siegl and Hentschel, 2010). Poribacteria-positive WGA products were additionally screened with the degenerated PCR primer pair 27f-B and 1492r-B (Cho and Giovannoni, 2004) that covers Eubacteria more broadly and enables to obtain longer 16S rRNA gene sequences from Poribacteria. Subsequent cloning of PCR products, restriction fragment length polymorphism analysis and Sanger sequencing was conducted to confirm the presence of a single cell using the same procedures as Siegl and Hentschel (2010).

WGA products that originated from single poribacterial cells were then subjected to another round of multiple displacement amplification under the same conditions as stated above. Before genome sequencing, we conducted a S1 nuclease treatment and DNA purification as described by Siegl et al. (2011). Five SAGs were selected in the PCR screening process and subjected to whole genome sequencing: Candidatus Poribacteria WGA-3G, WGA-4C, WGA-4CII, WGA-4E and WGA-4G (hereafter referred as 3G, 4C, 4CII, 4E and 4G, respectively). These were complemented by one poribacterial SAG sequence from an earlier study by Siegl et al. (2011), Candidatus Poribacteria WGA-A3 (hereafter referred as 3A). The existing assembly for this SAG was used for annotation and further analyses as described below.

Genome sequencing, assembly and annotation

A detailed description of all steps of genome sequencing, assembly, annotation and quality checks can be found in Supplementary Text S1. Briefly, a combination of Illumina and 454 pyrosequencing was conducted for double displacement amplification products of SAGs 4C, 4E and 4G at LGC Genomics (Berlin, Germany) and the DOE Joint Genome Institute (JGI, Walnut Creek, CA, USA). SAGs 3G and 4CII were also sequenced at JGI using Illumina HiSeq2000 technology only. Illumina sequences were normalized using DUK, a filtering program developed at JGI, and used for assembly including 454 reads (if available). For Illumina/454 hybrid assemblies, a combination of Velvet (Zerbino and Birney, 2008), Allpaths-LG (Zerbino and Birney, 2008) and the 454 Newbler assembler (Roche/454 Life Sciences, Branford, CT, USA) was used. For Illumina assemblies, we used the programs Velvet and Allpaths-LG. All assemblies were submitted to the IMG/ER annotation pipeline (Markowitz et al., 2009) for gene prediction and automatic

described in marine sponges (Fieseler et al., 2004). Poribacteria are widely distributed and highly abundant in sponge species around the world (Fieseler et al., 2004; Lafi et al., 2009; Schmitt et al., 2011, 2012), and also occur freely in seawater, albeit at very low abundances (Pham et al., 2008; Webster et al., 2010; Taylor et al., 2013). They were shown to be affiliated with the Planctomycetes, Verrucomicrobia, Chlamydiae (PVC) superphylum (Wagner and Horn, 2006). Poribacteria are vertically transmitted via reproductive stages (Schmitt et al., 2004; Lafi et al., 2011, 2012), and also occur freely in seawater, even from high diversity environments. Here we used single-cell genomics to analyze five poribacterial cells from the Mediterranean sponge Aplysina aerophoba, expanding the existing data set from one (Siegl et al., 2011) to a total of six poribacterial single-amplified genomes (SAGs). We provide an in-depth genomic analysis of one of the main symbiont lineages in the complex microbiota of marine sponges. The property of carbohydrate degradation emerged as the most common feature among the analyzed genomes. We therefore focused on carbohydrate degradation potential of Poribacteria in this study and discussed the results in context of a nutritional basis of the sponge–microbe symbiosis.

Poribacteria are vertically transmitted via reproductive stages (Schmitt et al., 2004; Lafi et al., 2011, 2012), and also occur freely in seawater, even from high diversity environments. Here we used single-cell genomics to analyze five poribacterial cells from the Mediterranean sponge Aplysina aerophoba, expanding the existing data set from one (Siegl et al., 2011) to a total of six poribacterial single-amplified genomes (SAGs). We provide an in-depth genomic analysis of one of the main symbiont lineages in the complex microbiota of marine sponges. The property of carbohydrate degradation emerged as the most common feature among the analyzed genomes. We therefore focused on carbohydrate degradation potential of Poribacteria in this study and discussed the results in context of a nutritional basis of the sponge–microbe symbiosis.

Materials and methods

Sample collection and processing

Samples of the marine sponge A. aerophoba were collected in September 2009 by scuba diving to a depth of 5–12 m at the Coast of Rovinj, Croatia (45°08’N, 13°64’E). The animals were transported to the University of Wuerzburg (Wuerzburg, Germany) and kept in seawater aquaria until further processing within 1 week of collection. Fresh sponge samples were used for extraction of sponge-associated prokaryotes using an established protocol of tissue disruption, density centrifugation and filtration by Fieseler et al. (2004).

Single-cell sorting, whole genome amplification and PCR screening

Single-cell isolations were conducted with freshly extracted and purified sponge-associated prokaryotes using the fluorescence-activated cell sorting Vantage SE flow cytometer with FACSDiVa option (Becton Dickinson, Heidelberg, Germany) as described previously (Siegl and Hentschel, 2010). For cell lysis and whole genome amplification (WGA) by multiple displacement amplification, we followed the same procedure as Siegl et al. (2011). To identify phylogenetically WGA products and check for possible contamination, we screened the WGA products obtained by polymerase chain reaction (PCR) using 16S or 18S rRNA gene primers targeting Eubacteria, Archaea, Poribacteria and Eukaryotes, as described previously (Siegl and Hentschel, 2010). Poribacteria-positive WGA products were additionally screened with the degenerated PCR primer pair 27f-B and 1492r-B (Cho and Giovannoni, 2004) that covers Eubacteria more broadly and enables to obtain longer 16S rRNA gene sequences from Poribacteria. Subsequent cloning of PCR products, restriction fragment length polymorphism analysis and Sanger sequencing was conducted to confirm the presence of a single cell using the same procedures as Siegl and Hentschel (2010).

WGA products that originated from single poribacterial cells were then subjected to another round of multiple displacement amplification under the same conditions as stated above. Before genome sequencing, we conducted a S1 nuclease treatment and DNA purification as described by Siegl et al. (2011). Five SAGs were selected in the PCR screening process and subjected to whole genome sequencing: Candidatus Poribacteria WGA-3G, WGA-4C, WGA-4CII, WGA-4E and WGA-4G (hereafter referred as 3G, 4C, 4CII, 4E and 4G, respectively). These were complemented by one poribacterial SAG sequence from an earlier study by Siegl et al. (2011), Candidatus Poribacteria WGA-A3 (hereafter referred as 3A). The existing assembly for this SAG was used for annotation and further analyses as described below.

Genome sequencing, assembly and annotation

A detailed description of all steps of genome sequencing, assembly, annotation and quality checks can be found in Supplementary Text S1. Briefly, a combination of Illumina and 454 pyrosequencing was conducted for double displacement amplification products of SAGs 4C, 4E and 4G at LGC Genomics (Berlin, Germany) and the DOE Joint Genome Institute (JGI, Walnut Creek, CA, USA). SAGs 3G and 4CII were also sequenced at JGI using Illumina HiSeq2000 technology only. Illumina sequences were normalized using DUK, a filtering program developed at JGI, and used for assembly including 454 reads (if available). For Illumina/454 hybrid assemblies, a combination of Velvet (Zerbino and Birney, 2008), Allpaths-LG (Zerbino and Birney, 2008) and the 454 Newbler assembler (Roche/454 Life Sciences, Branford, CT, USA) was used. For Illumina assemblies, we used the programs Velvet and Allpaths-LG. All assemblies were submitted to the IMG/ER annotation pipeline (Markowitz et al., 2009) for gene prediction and automatic
functional annotation. All data sets were quality checked and screened for possible contamination (Supplementary Table S1). Manual curation and functional analyses were conducted within the IMG/MER system (Markowitz et al., 2012a), unless stated otherwise.

The Whole Genome Shotgun projects were deposited at DDBJ/EMBL/GenBank under the accession nos. ASZN00000000 (3G), APGO00000000 (4C), ASZM00000000 (4CII), AQTV00000000 (4G) and AQPC00000000 (4G). The versions described in this paper are versions ASZN01000000 (3G), APGO01000000 (4C), ASZM01000000 (4CII), AQTV01000000 (4G) and AQPC01000000 (4G). Raw data, genome assemblies and annotations can also be accessed under the IMG software system (http://img.jgi.doe.gov) under genome IDs 2 265 129 006–2 265 129 011. Additionally, 16S rRNA gene sequences for phylogenetic analysis from PCR screenings were submitted to GenBank under accession numbers KC713965–KC713966.

ANI and tetranucleotide frequency analysis
Average nucleotide identities based on BLAST (ANIb) and tetranucleotide frequencies were estimated using the JSpecies software (v.1.2.1; http://www.imedea.uib.es/jspecies/about.html) with default parameters (Richter and Rosselló-Móra, 2009).

Genome completeness estimation
Genome size and completeness were estimated using two conserved single copy gene sets that have been determined from all bacterial (n = 1516) and all archaeal (n = 111) finished genome sequences in the IMG database (Markowitz et al., 2012b). The sets consist of 138 bacterial and 162 archaeal conserved single copy genes that occurred only once in at least 90% of all genomes by analysis of an abundance matrix based on hits to the protein family database (Punta et al., 2012). HMMSs of the identified protein families were used to search both, all SAG assemblies and all combined assemblies by means of the HMMER3 software (Finn et al., 2011; http://hmmer.janelia.org/help). Resulting best hits above precalculated cutoffs were counted and the completeness was estimated as the ratio of conserved single-copy gene to total conserved single-copy genes in the set after normalization to 90%. Thereafter, the estimated complete genome size was calculated by division of the estimated genome coverage by the total assembly size.

Phylogenetic analysis
Poribacterial 16S rRNA gene sequences were aligned using the SINA aligner (Pruesse et al., 2012) and manually checked in the ARB software package (v.5.3; http://www.arb-home.de/) (Ludwig et al., 2004). The SILVA 16S rRNA database version 111 (Pruesse et al., 2012) was used for selection of reference sequences plus additional poribacterial 16S rRNA gene sequences in the SILVA database in January 2013. Only sequences ≥1100 bp were used to construct a maximum-likelihood bootstrap tree with 1000 resamplings with the RAxML software (v.7.2.8; http://www.exelixis-lab.org/) (Stamatakis, 2006). The resulting tree was reimported into ARB and short poribacterial sequences (≤1099 bp) were added without changing tree topology using the parsimony interactive tool in ARB.

Screening for carbohydrate-active enzymes
Protein sequences of poribacterial genomes were screened against the HMM profile-based database of carbohydrate-active enzymes obtained from dbCAN (Yin et al., 2012) in December 2012 using hmmsearch in the HMMER software package (v.3.0; http://hmmer.janelia.org/help) (Finn et al., 2011). Results were filtered using an e-value cutoff <10^-5. In addition, all returned hits were manually evaluated based on their functional annotation in IMG/MER and excluded in case of conflicting results. Comparison between glycoside hydrolase (GH)-encoding genes (E.C.: 3.2.1.x) between Poribacteria and all free-living planktonic organisms available in the IMG software system in May 2013 was conducted additionally using functional annotation tools in IMG.

Discussion

Results and discussion

General genomic features

SAG sequencing. Final genome assembly sizes for the poribacterial cells ranged from 0.19 to 5.44 Mbp (Table 1). For genomes 3G, 4C and 4E, genome recovery was large enough to estimate genome coverage of 98.54%, 38.36% and 58.20%, respectively, whereas the largely fragmented assemblies of 3A, 4CII and 4G did not permit for genome size estimation. The estimated poribacterial genome sizes ranged from 4.25 to 6.27 Mb (Table 1) and do not suggest genome size reduction. The guanine–cytosine content ranged from 47% to 50%, with the exception of genome 4C (41%) (Table 1). Protein coding genes accounted for 95.5–99.6% of the retrieved genomes (Table 1). Approximately 30% of these could not be functionally assigned (50% for genome 3A). The investigated genomes encoded for only one copy of the 16S rRNA gene, which is consistent with previous reports (Fieseler et al., 2006; Siegl et al., 2011).

Definition of phylotypes. Phylogenetic analysis of nearly full-length 16S rRNA gene sequences of Poribacteria showed that three of the six analyzed SAGs (3A, 3G and 4CII) clustered closely together, with genome 4G also in close proximity (>97%
sequence similarity), whereas the other two genomes (4C and 4E) each fell separate from this group (Figure 1). Average nucleotide identity and tetranucleotide frequency analysis confirmed a closer relationship between SAGs 3A, 3G, 4CII and 4G than to the other two SAGs (Supplementary Table S2 and Supplementary Figure S1), and they were therefore defined as one phylotype named group I (Figure 1). The other two SAGs each represent a separate phylotype. Group I represents, however, a ‘composite phylotype’ as the values for tetranucleotide frequency and average nucleotide identity analysis are under the defined thresholds of 0.99 for tetranucleotide frequency and 95–96% ID for average nucleotide identity (Richter and Rossello-Mora, 2009).

GHs and other CAZymes

The ability of Poribacteria to degrade and transform complex carbohydrates was assessed by screening genome data against the dbCAN (Yin et al., 2012) and classified according to the carbohydrate-active enzymes (CAZy) database (Cantarel et al., 2009). Most poribacterial hits matched GHs and glycosyl transferase, whereas carbohydrate binding modules, carbohydrate esterases and polysaccharide lyases made up for a smaller proportion (Supplementary Table S3, Supplementary Figure S2). GH and glycosyl transferase also showed the highest diversity with up to 18 different GHs and 18 glycosyl transferases as detected in SAG 3G. We detected 17, 76, 21, 14 and 63 hits to GH families in SAGs 3A, 3G, 4C, 4CII and 4E, respectively (fasta file available in Supplementary Information). No GHs were found in 4G, which is however largely incomplete. GH frequencies of 1.6%, 1.3% and 1.9% genes/genome were estimated for 3G, 4C, and 4E, respectively. In a recent study by Martinez-Garcia et al. (2012), marine Verrucomicrobia genomes contained approximately 0.9–1.2% genes encoding for GHs as compared with the average of 0.2% in other bacteria. Considering slight differences in screening methods between our study and that by Martinez-Garcia et al. (2012), the poribacterial GH frequency was found to be similar to that of Verrucomicrobia. Poribacterial genomes also show an almost linear correlation between the number of GHs and genome coverage with \( R^2 = 0.9525 \) (Supplementary Figure S3). Similar to Verrucomicrobia, these findings could indicate a specialization of Poribacteria towards carbohydrate degradation.

We additionally compared the poribacterial GH frequencies with those of all available finished genomes of marine planktonic bacteria in IMG (\( n=102, \text{ May 2013} \)) and also the nearly closed Verrucomicrobia single-cell genome AAA168-F10 by Martinez-Garcia et al. (2012) based on annotation

### Table 1 Summary of assembly and genome statistics of poribacterial genomes

<table>
<thead>
<tr>
<th></th>
<th>Group I</th>
<th>3A</th>
<th>3G</th>
<th>4C</th>
<th>4E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assembly size (bp)</td>
<td>414 219</td>
<td>5 441 554</td>
<td>543 453</td>
<td>189 191</td>
<td>1 629 923</td>
</tr>
<tr>
<td>Estimated genome size (bp)</td>
<td>N/A</td>
<td>5 521 899</td>
<td>N/A</td>
<td>N/A</td>
<td>4 249 040</td>
</tr>
<tr>
<td>Estimated genome recovery (%)</td>
<td>N/A</td>
<td>98.34</td>
<td>N/A</td>
<td>N/A</td>
<td>38.36</td>
</tr>
<tr>
<td>Number of contigs</td>
<td>157</td>
<td>286</td>
<td>44</td>
<td>15</td>
<td>276</td>
</tr>
<tr>
<td>Largest contig size (bp)</td>
<td>13 447</td>
<td>227 865</td>
<td>92 961</td>
<td>33 129</td>
<td>76 460</td>
</tr>
<tr>
<td>Sequencing effort (Mbp)</td>
<td>454 FLX</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>454 FLX Titanium</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>97.1</td>
</tr>
<tr>
<td>Illumina GA IIx</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>5 800</td>
</tr>
<tr>
<td>Illumina Hi Seq 2000</td>
<td>N/A</td>
<td>1410</td>
<td>780</td>
<td>N/A</td>
<td>2 300</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>49</td>
<td>48</td>
<td>48</td>
<td>47</td>
<td>41</td>
</tr>
<tr>
<td>Protein CDs</td>
<td>503</td>
<td>4772</td>
<td>473</td>
<td>170</td>
<td>1 618</td>
</tr>
<tr>
<td>%</td>
<td>99.60</td>
<td>99.00</td>
<td>98.95</td>
<td>95.51</td>
<td>99.02</td>
</tr>
<tr>
<td>Protein coding genes with function prediction</td>
<td>No 256</td>
<td>3228</td>
<td>305</td>
<td>112</td>
<td>950</td>
</tr>
<tr>
<td>%</td>
<td>50.89</td>
<td>6.74</td>
<td>64.38</td>
<td>65.88</td>
<td>58.71</td>
</tr>
<tr>
<td>Protein coding genes without function prediction</td>
<td>No 247</td>
<td>1544</td>
<td>168</td>
<td>58</td>
<td>668</td>
</tr>
<tr>
<td>%</td>
<td>49.11</td>
<td>32.36</td>
<td>35.72</td>
<td>34.12</td>
<td>41.29</td>
</tr>
<tr>
<td>rRNAs</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>%</td>
<td>0.20</td>
<td>0.04</td>
<td>0.63</td>
<td>0.56</td>
<td>0.18</td>
</tr>
<tr>
<td>tRNAs</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>%</td>
<td>0.20</td>
<td>0.89</td>
<td>0.42</td>
<td>3.93</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Abbreviations: CD, coding genes; GC, guanine–cytosine content; NA, not applicable; rRNA, ribosomal RNA; tRNA, transfer RNA.
Carbohydrate degradation in Poribacteria

J Kamke et al

Figure 1 Phylogenetic maximum-likelihood tree based on 16S rRNA genes of the candidate phylum Poribacteria. Sequences from poribacterial SAGs are shown in bold and gray shading. The tree was constructed based on long sequences (≥ 1100 nucleotides), shorter sequences were added without changing tree topology and are indicated by dashed lines. Bootstrap support (1000 resamplings) of ≥ 90% is shown by filled, and ≥ 75% by open circles. The outgroup consisted of several Spirochaetes sequences. Scale bar represents 4% sequence divergence.

The ISME Journal
of EC numbers 3.2.1.x (see CAZY database). GH frequencies based on this analysis were on average 0.32% for planktonic bacteria, 0.24% for *Verrucomicrobia* AAA168-F10 and 0.08–0.42% for poribacterial genomes. A *Poribacteria*-specific set of enzymes was however not identified, which may be because of the fact that the most dominant GH families in *Poribacteria* are not accessible in the IMG system because of the lack of EC number annotations available for these families.

### Table 2 Overview of glycoside hydrolase families in *Poribacteria*

<table>
<thead>
<tr>
<th>GH family</th>
<th>Known activities</th>
<th>Potential activities in <em>Poribacteria</em></th>
<th>Putative substrates</th>
<th>Group 1</th>
<th>4C</th>
<th>4E</th>
<th>3A</th>
<th>3G</th>
<th>4CH</th>
</tr>
</thead>
<tbody>
<tr>
<td>GH2</td>
<td>β-Galactosidase, β-glucuronidase, β-mannosidase, others</td>
<td>β-Galactosidase/β-glucuronidase</td>
<td>β-o-galactoside</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH4</td>
<td>α-Glucosidase, α-galactosidase, α-glucuronidase, α-galacturonidase</td>
<td>α-Galactosidase</td>
<td>β-o-glucuronic acid (glycosaminoglycans/mucopolysaccharides</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH5</td>
<td>Cellulase, β-1,3-galactosidase, many other enzymes (55 subfamilies, 20 experimentally defined enzyme functions)</td>
<td>—</td>
<td>β-Linked oligo- and polysaccharides and glycoconjugates</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH13</td>
<td>α-Amylase and related enzymes</td>
<td>1,4-α-Glucan branching enzyme (EC: 2.4.1.18), α-amylase</td>
<td>α-Glycoside linkages, (1–4)-α-glucosidic linkages in polysaccharides</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH23</td>
<td>G-type lysozyme, peptidoglycan lyase, peptidoglycan-lytic transglycosylase</td>
<td>Soluble lytic murein transglycosylase and related proteins (EC: 3.2.1.1)</td>
<td>N-acetylmuramyl and N-acetylgalactosaminyl residues in peptidoglycan</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH32</td>
<td>Levanase, invertase, others</td>
<td>β-Fructofuranosidase (EC: 2.4.1.20)</td>
<td>Sucrose</td>
<td>3</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH33</td>
<td>Sialidase or neuraminidase (EC: 3.2.1.18); trans-sialidase (EC: 2.4.1.11); 2-keto-3-deoxynononic acid sialidase (EC: 3.2.1.1)</td>
<td>Exo-α-sialidase, neuraminidase (EC: 3.2.1.18)</td>
<td>Neuraminic acids, α-(2→3)-, α-(2→6)-, α-(2→8)-glycosidic linkages of terminal sialic acid residues in oligosaccharides, glycoproteins, glycolipids, colominic acid</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH36</td>
<td>α-Galactosidase, α-N-acetylgalactosaminidase</td>
<td>α-Galactosidase</td>
<td>Terminal α-galactosyl moieties from glycolipids and glycoproteins</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH50</td>
<td>β-Agarases</td>
<td>—</td>
<td>Cleave β-1,4 glycosidic bonds of agarose, releasing neogalacto-biose, -tetrasac- and -hexaose containing (1,3)- and/or (1,5)-linkages, arabinoxylans and arabinogalactans</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH51</td>
<td>Endoglucanase, α-L-arabinofuranosidase</td>
<td>α-L-arabinofuranosidase</td>
<td>—</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH67</td>
<td>α-Glucuronidase, xylan</td>
<td>α-Glucuronidase (EC: 3.2.1.139)</td>
<td>Glucuronic acid appended to the C2-OH of xylose at the non-reducing end of xylooligosaccharides</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH74</td>
<td>Endoglucanase, oligoxylglycan reducing end-specific celllobiohydrolase, xylglucanase</td>
<td>Some with BNR (bacterial nuraminidase/xylglucanase)</td>
<td>β-1,4-Linkages of various glucans</td>
<td>5</td>
<td>25</td>
<td>7</td>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>GH76</td>
<td>1,6-α-Mannosidase</td>
<td>—</td>
<td>(1→6)-α-β-mannosidic linkages in unbranched (1→6)-β-mannans</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH88</td>
<td>β-4,5 unsaturated β-glucuronidase</td>
<td>—</td>
<td>Release of 4-deoxy-4(5)-unsaturated β-glucuronic acid from oligosaccharides produced by polysaccharide lyases</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH93</td>
<td>Exo-α-L-1,5-arabinase</td>
<td>—</td>
<td>Release of arabinobiose from the non-reducing end of α-L-1,5-arabinan</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH95</td>
<td>α-L-fucosidase, α-L-arabinofuranosidase</td>
<td>α-L-fucosidase (EC:3.2.1.51)</td>
<td>Fuco-α-L-2-Gal linkages attached at the non-reducing ends of oligosaccharides</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH105</td>
<td>Unsaturated rhamnogalacturonidase</td>
<td>—</td>
<td>Ramnosylgalacturan degradation</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH106</td>
<td>α-L-rhamnosidase</td>
<td>—</td>
<td>Hydrolysis of terminal non-reducing α-L-rhamnose residues in α-L-rhamnosides</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH109</td>
<td>α-N-acetylgalactosaminidase</td>
<td>—</td>
<td>N-acetylgalactosamine linkage in glycoproteins</td>
<td>2</td>
<td>29</td>
<td>4</td>
<td>8</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>GH127</td>
<td>β-arabinofuranosidase</td>
<td>—</td>
<td>Release of L-arabinose from specific disaccharides and glycoconjugates</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH129</td>
<td>α-N-acetylgalactosaminidase</td>
<td>—</td>
<td>Mucin-type glycoproteins</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH130</td>
<td>1-β-α-mannopyranosyl-4-α-glucopyranose: phosphate α-β-mannosyltransferase</td>
<td>—</td>
<td>Mannan catabolism</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: GH, glycoside hydrolase.
**Galactoside, fructoside, xyloside and rhamnoside degradation.** Poribacterial genomes showed the potential for degradation of galactoside polymers such as melibiose and lactose. Oxidative degradation of galactosides was supported as well as parts of the Lelior pathway (Supplementary Figure S5 and Supplementary Table S5). The potential for degradation of the fructoside polymers levan and sucrose was also encoded in poribacterial genomes (Supplementary Figure S6) and additionally genes coding for enzymes involved in degradation of D-xylene over the xylene isomerase pathway (Supplementary Figure S6). The genomic potential for rhamnoside degradation was shown by genes relevant for the L-rhamnose isomerase pathway and oxidative L-rhamnose degradation, which was strongly supported by multiple gene copies encoding for enzymes involved in this pathway (Supplementary Figure S7 and Supplementary Table S5). For a detailed description of these pathways in Poribacteria, please refer to Supplementary Text S2.

The substrates of these degradation pathways can generally be found in oligo- and polysaccharides of glycoconjugates or biopolymers of various organisms, and especially in the cell walls of many plants and bacteria (Sutherland, 1985; Rehm, 2010; Ray et al., 2011; Visnapuu et al., 2011; Singh et al., 2012). Therefore, they might be freely available in the sponge mesohyl as a result of sponge feeding on bacteria and microalgae.

**Inositol degradation.** A nearly complete inositol dehydrogenase pathway was present in group I genomes, as well as in genome 4E, and partially in 4C. This pathway degrades myoinositol to glyceraldehyde-3-phosphate, which is further used in the central metabolism (Figure 2). Inositol phosphates are cell wall compounds in all eukaryotes and archaea, and are rarely found in bacteria (Michell, 2011). In addition, phosphorylated inositol is a
precursor for several important lipid molecules including sphingolipids, ceramides and glycosylphosphatidylinositol anchors (Reynolds, 2009), as well as many stress-protective solutes of eukaryotes (Michell, 2011). Inositol phosphate is part of the signal transduction in sponges, as shown in Geodia cydonium, where production of inositol triphosphate increases after sponge cell aggregation (Müller et al., 1987). It seems therefore likely that either the sponge itself or eukaryotic microorganisms could be a source of inositol for Poribacteria. Inositol degradation has been reported from a variety of other bacteria (Fry et al., 2001; Yoshida et al., 2008, 2012; Kohler et al., 2010) including Sinorhizobium symbionts of soybean where this pathway was shown to provide a competitive advantage in the plant rhizosphere (Galbraith et al., 1998). We hypothesize that Poribacteria not only use myoinositol as a carbon source but also as an agent for regulating metabolic functions involved in sponge–microbe symbiosis.

Uronic acid degradation. Analysis of the poribacterial genomes further revealed the presence of several genes connected to uronic acid degradation (Figure 3 and Supplementary Table S5). The degradation of polymers such as pectin or pectin-like glycoconjugates likely occurs in all investigated phylotypes as indicated by the presence of genes encoding for polygalacturonase (EC: 3.2.1.15) in genome 4E and pectate lyase (EC: 4.2.2.2) in group

![Figure 2](attachment:image2.png)  
Figure 2 Schematic reconstruction of inositol degradation as encoded on poribacterial SAGs. Numbers within circles represent the number of genomes encoding for the corresponding enzyme.

![Figure 3](attachment:image3.png)  
Figure 3 Schematic reconstruction of uronic acid degradation as encoded on poribacterial SAGs. Numbers within circles represent the genomes encoding for the corresponding enzyme. Dashed circles represent manually annotated genes.
I genome 3G and genome 4C. This was further supported by the presence of genes encoding for enzymes participating in 5-dehydro-4-deoxy-D-glucuronate degradation such as oligogalacturonide lyase (EC: 4.2.2.6), 2-deoxy-D-glucurononate 3-dehydrogenase (EC: 1.1.1.125) and 2-dehydro-3-desoxy-D-glucokinase (EC: 2.7.1.45) in SAG group I genomes and partially in genome 4C. Furthermore, glucurononoside polymers appear to be degradable by Poribacteria as indicated by the presence of α- and β-D-glucuronosidases (EC: 3.2.1.139; EC: 3.2.1.31) in group I genomes 3G and 4CII, respectively. The occurrence of genes encoding for galacturonate isomerase (EC: 5.3.1.12), alteronate hydrolase (EC: 4.2.1.7) and mannnonate dehydratase (EC: 4.2.1.8) in high copy number on several poribacterial genomes of the phylotypes represented by group I and 4E points towards the possibility of galacturonate and glucuronate catabolism. The products of these degradation steps could then enter the ED pathway via 2-dehydro-3-desoxyphosphogluconate aldolase (EC: 4.1.2.14; 4.1.3.16), which was found in group I genomes and in genome 4E.

Uronic acids are sugar acids that can be found in various biopolymers of plant, animal or bacterial origin (Sutherland, 1985; Rehm, 2010). The occurrence of uronic acids in glycosaminoglycans (GAGs) is especially worth noting, because the extracellular matrix of sponges is largely constructed by these polymers (Fernandez-Busquets and Burger, 2003). Therefore, Poribacteria should literally be submerged in uronic acid-containing substances. It is known that GAGs of sponges are different from those of higher animals (Misevic and Burger, 1993), with the main components being fucose, glucuronic acid, mannose, galactose, N-acetylglucosamine and sulfate in sponge GAGs (Misevic and Burger, 1986, 1993; Misevic et al., 1987).

### Table 3 Transporters classes on poribacterial genomes

<table>
<thead>
<tr>
<th>Function ID</th>
<th>Transporter classification</th>
<th>Group I</th>
<th>4C</th>
<th>4E</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC: 1.A.30</td>
<td>H⁺ or Na⁺-translocating bacterial flagellar Mot/Exb superfamily</td>
<td>0</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>TC: 2.A.55</td>
<td>Nram family</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>TC: 2.A.56</td>
<td>TRAP-T family</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>TC: 2.A.64</td>
<td>Tat family</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>TC: 2.A.66</td>
<td>MOP flippase superfamily</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>TC: 3.A.1</td>
<td>ABC superfamily</td>
<td>12</td>
<td>174</td>
<td>19</td>
</tr>
<tr>
<td>TC: 3.A.15</td>
<td>Outer membrane protein secreting MTB family</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>TC: 3.A.5</td>
<td>Sec family</td>
<td>0</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>TC: 3.A.7</td>
<td>Type IV (conjugal DNA–protein transfer or VirB) secretory pathway (IVSP) family</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>TC: 4.C.1</td>
<td>Proposed FAT family</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>TC: 5.A.4</td>
<td>Prokaryotic SDH family</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>TC: 9.B.22</td>
<td>Leukotoxin secretion MorC family</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Abbreviations:** ABC, ATP-binding cassette; FAT, fatty acid transporter; MOP, m/oligosaccharidyl-lipid/poly saccharide; MorC, morphogenesis protein C; Mot/Exh, Motor/ExbBD outer membrane transport energizer; MTB, main terminal branch; Nram, metal ion [(Mn²⁺-iron] transporter; RhtB, resistance to homoserine/threonine; SDH, succinate dehydrogenase; Sec, general secretory pathway; Tat, twin arginine targeting; TRAP-T, tripartite ATP-independent periplasmic transporter.

**Table 3** Transporters classes on poribacterial genomes

**Transporters**

The poribacterial genomes code for a range of transporters representing different families (Table 3 and Figure 4), and here we concentrate on those involved in carbohydrate metabolism. Typical sugar transport systems/phosphotransferase systems were missing from all poribacterial genomes. Genes coding for proteins of the tripartite ATP-independent periplasmic transporter family were found on all three phylotypes, which are often involved in organic acid transport such as C4-dicarboxylates, keto-acids and sugar acids (N-acetyl neuraminic acids, sialic acid). The most dominant transporter family in all poribacterial genomes was the ATP-binding cassette superfamily. ATP-binding cassette transporters were found for a variety of broader substrate categories such as amino acids, di- and oligopeptides, carbohydrates, lipoproteins, metal ions and systems involved in cell protection and competition with other organisms (Supplementary Table S6). Most of the detected carbohydrate ATP-binding cassette transporters were simple or multiple sugar transporters and did not show any further specification. However, we detected a system of α-xylene transport on group I genome 3G and transporters for maltose, malto-oligosaccharides, arabinose and lactose on several other genomes besides 3G.

**Sulfatases**

Genome analysis of the six poribacterial genomes revealed a total of 103 genes coding for sulfatases (Table 4). Individually, group I genomes coded for 57, genome 4C for 25 and genome 4E for 21 sulfatase genes. Genes coding for choline sulfatase (bettC; EC: 3.1.6.6) were identified in group I genomes, and 4E. This enzyme transforms choline sulfate to choline.
and then to glycine betaine, which is often used as an osmoprotectant (Le Rudulier et al., 1984) or can be degraded via glycine and serine to pyruvate. This degradative pathway is largely present on genome 4E, whereas glycine degradation was encoded in all three phylotypes. As choline-O-sulfate is synthesized by different microorganisms (Spencer and Harada, 1960; Fitzgerald and Luschinski, 1977; Rivoal and Hanson, 1994), it is likely available as carbon and sulfur substrate for Poribacteria.

Two genes coding for arylsulfatase (EC: 3.1.6.1) were detected in group I genomes 3G and 4E. General substrates for this type of enzyme are phenol sulfates, but the exact kind of phenol sulfate is difficult to determine because of high similarities between different types of substrates. Genome 3G

![Diagram of Poribacteria cell](image)

**Figure 4** Schematic overview of a poribacterial cell in the sponge extracellular matrix illustrating pathways of carbohydrate metabolism and glycosaminoglycan degradation by poribacterial enzymes. The dashed arrow represents glycolysis that is not supported by the dominant poribacterial phytype group I.

**Table 4** Sulfatase genes on poribacterial genomes

<table>
<thead>
<tr>
<th>Sulfatase type</th>
<th>KO term</th>
<th>KO name</th>
<th>Function</th>
<th>Number of genes on genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arylsulfatase</td>
<td>KO01130</td>
<td>aslA</td>
<td>EC: 3.1.6.1</td>
<td>8</td>
</tr>
<tr>
<td>Cerebroside-sulfatase</td>
<td>KO01134</td>
<td>ARSA</td>
<td>EC: 3.1.6.8</td>
<td>45</td>
</tr>
<tr>
<td>Choline-sulfatase</td>
<td>KO01133</td>
<td>betC</td>
<td>EC: 3.1.6.6</td>
<td>3</td>
</tr>
<tr>
<td>Iduronate-2-sulfatase</td>
<td>KO01136</td>
<td>IDS</td>
<td>EC: 3.1.6.13</td>
<td>1</td>
</tr>
<tr>
<td>N-acetyl-galactosamine-4-sulfatase</td>
<td>KO01135</td>
<td>ARSB</td>
<td>EC: 3.1.6.12</td>
<td>1</td>
</tr>
<tr>
<td>N-sulfogalactosamine-sulfohydrodase</td>
<td>KO01565</td>
<td>SGSH</td>
<td>EC: 3.10.1.1</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized</td>
<td>KO01138</td>
<td></td>
<td>EC: 3.1.6.-</td>
<td>1</td>
</tr>
<tr>
<td>Not determined</td>
<td>N/A</td>
<td></td>
<td></td>
<td>7</td>
</tr>
</tbody>
</table>

**Table 4** Sulfatase genes on poribacterial genomes

<table>
<thead>
<tr>
<th>Sulfatase type</th>
<th>KO term</th>
<th>KO name</th>
<th>Function</th>
<th>Number of genes on genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arylsulfatase</td>
<td>KO01130</td>
<td>aslA</td>
<td>EC: 3.1.6.1</td>
<td>8</td>
</tr>
<tr>
<td>Cerebroside-sulfatase</td>
<td>KO01134</td>
<td>ARSA</td>
<td>EC: 3.1.6.8</td>
<td>45</td>
</tr>
<tr>
<td>Choline-sulfatase</td>
<td>KO01133</td>
<td>betC</td>
<td>EC: 3.1.6.6</td>
<td>3</td>
</tr>
<tr>
<td>Iduronate-2-sulfatase</td>
<td>KO01136</td>
<td>IDS</td>
<td>EC: 3.1.6.13</td>
<td>1</td>
</tr>
<tr>
<td>N-acetyl-galactosamine-4-sulfatase</td>
<td>KO01135</td>
<td>ARSB</td>
<td>EC: 3.1.6.12</td>
<td>1</td>
</tr>
<tr>
<td>N-sulfogalactosamine-sulfohydrodase</td>
<td>KO01565</td>
<td>SGSH</td>
<td>EC: 3.10.1.1</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized</td>
<td>KO01138</td>
<td></td>
<td>EC: 3.1.6.-</td>
<td>1</td>
</tr>
<tr>
<td>Not determined</td>
<td>N/A</td>
<td></td>
<td></td>
<td>7</td>
</tr>
</tbody>
</table>

Abbreviations: KO, KEGG orthology; N/A, not applicable.

![Diagram of glycaminoglycan degradation by Poribacteria](image)
codes also for a cerebroside-sulfatase (EC: 3.1.6.8), which also hydrolyzes phenol sulfates, ascorbate 2-sulfate and galactose-3-sulfate residues in lipids. These enzymes might provide *Poribacteria* with an organic sulfur source from hydrolyzed sulfate esters in the absence of inorganic sulfate. *Poribacteria* have the genomic potential for assimilatory sulfate reduction over adenosine-5’-phosphosulfate and 3’-phosphoadenosin-5’-phosphosulfate and subsequent cysteine synthesis (data not shown). The use of organic sulfur compounds under sulfur-limiting conditions has been shown in various bacteria (Kertesz, 2000). The presence of sulfated lipids and polysaccharides in the sponge extracellular matrix is well documented (Zierer and Mourao, 2000; Vilanova et al., 2009), and it may thus serve as a possible carbon and sulfur source.

Group I and genome 4C both encode for N-acetylgalactosamine-4-sulfatase (EC: 3.1.6.12), which lyses the sulfate groups of N-acetylgalactosamine-4-sulfate from chondroitin and dermatan sulfate. Chondroitin is known to be part of sponge GAG chains, and *N*-acetylgalactosamine was found in GAG chains and glycoproteins of sponges (Fernandez-Busquets and Burger, 2003). In addition, genome 4C codes for idurionate-2-sulfatase (EC: 3.1.6.13) and SAG group I genome 3G for N-sulfoglucosamine sulfohydrolase (EC: 3.10.1.1); however, their substrates are yet to be identified in sponges.

**Symbiotic heterotrophy in sponges**

Symbioses frequently have a nutritional basis, such as nitrogen fixation in *Rhizobium*–legume symbioses (Lodwig et al., 2003), supplementation of amino acids in *Buchnera* symbionts of aphids (Ramsey et al., 2010; Hansen and Moran, 2011), chemoaerotrophy in marine mussels and worms (Woyke et al., 2006; Petersen et al., 2011) or photosynthesis in symbionts of corals or ascidians (Weis and Allemann, 2009; Schnitzler and Weis, 2010; Donia et al., 2011). Here we present bacterial heterotrophy, that is, the ability to utilize diverse carbon sources, as a potential functional basis for the interaction of microbial symbionts with sponges. The role of heterotrophy in symbiosis in general has so far been underestimated. Only recently, studies of gut microbiomes of humans, ruminants or termites, and other terrestrial hosts (Warnecke et al., 2007; Hess et al., 2011; Zhu et al., 2011; Schloissnig et al., 2013) have received major attention. Here we show that symbiotic heterotrophy could also be of relevance in a marine host, using poribacterial symbionts of marine sponges as an example. Whether the host gains benefit from the heterotrophic metabolism of its symbionts is an interesting question for future investigations.

The carbohydrate degradation potential of *Poribacteria* opens up various nutritional sources that are taken up by the host’s extensive filtration activities and transported into the sponge interior. Sponges feed on dissolved and particulate organic matter (Yahel et al., 2003; De Goeij et al., 2008) as well as heterotrophic bacteria and eukaryotes (Maldonado et al., 2012; Perea-Blázquez et al., 2012). It has further been shown that sponge feeding on dissolved and particulate organic matter from bacterial and algal sources can be mediated by bacterial symbionts (De Goeij et al., 2008). Here we provide the genomic background and detailed carbon degradation pathways behind this ecological observation. We could show that *Poribacteria* have the potential to catabolize various substrates that can be found in, for example, cell wall components, glycoproteins and polysaccharides from marine algae and prokaryotes (Rehm, 2010; Jiao et al., 2011).

The carbon degradation repertoire of *Poribacteria* would further be consistent with degradation of compounds from the extracellular matrix of the sponge host (Figure 4). Our hypothesis is supported by (i) the high abundance of GH families with acetylgalactosaminidase or sialidase activities (and several GHs consistent with this hypothesis), (ii) the ability to degrade uronic acids, frequent components of glycoconjugates and especially GAGs, as well as (iii) the presence of sulfatases with GAGs as the specific substrate. Remarkably, *Poribacteria* are significantly enriched in high microbial abundance over low microbial abundance of sponges or seawater (Schmitt et al., 2012; Taylor et al., 2013) One major difference between high microbial abundance and low microbial abundance sponges (apart from the amount and diversity of microorganisms itself) is that high microbial abundance sponges contain a notably expanded mesohyl matrix as compared with their low microbial abundance counterparts (Hentschel et al., 2003; Weisz et al., 2008). The extracellular matrix may thus provide both habitat and nutrient source to *Poribacteria*. As sponges are known to continuously remodel matrix components (Bond, 1992), it is likely that polymers that become available in this process, serve as nutrient substrates for *Poribacteria*. This process is unlikely harmful to the host as no signs of tissue destruction in healthy *A. aerophoba* sponges can be observed neither in the natural environment nor by transmission electron microscopy of the mesohyl matrix (Vacelet, 1975; Friedrich et al., 2001).

The bacteria–sponge interaction may go beyond nutrition and may have also a mechanistic basis. In this context, the so-called ‘sponge aggregation factor’ is of relevance, which shows a proteoglycan-like structure and has a major role in cell-specific aggregation cell–matrix connections and adhesion processes (Müller and Zahn, 1973; Misievc and Burger, 1993; Fernandez-Busquets and Burger, 2003). It might be conceivable that *Poribacteria* influence the function of the sponge aggregation factor through glycosylase, glycotransferase activities, or by modification of sulfate groups on the
mucopolysaccharide chains. The poribacterial potential to influence sialylation by GH family 33 enzymes might also affect this process. It has been shown that sialic acid residues are an important component of glycoproteins and mucopolysaccharides on sponge cell surfaces (Garrone et al., 1971) and that these also have a major role in sponge cell aggregation (Müller et al., 1977). It might thus be possible that poribacterial enzymes interfere with mechanistic processes related to adhesion, aggregation and self–non-self recognition.

Conflict of Interest
The authors declare no conflict of interest.

Acknowledgements
We gratefully acknowledge the marine operations personnel at the Ruder Boskovic Institute (Rovinj/Croatia) for the help during sponge collection, Kristina Bayer for logistical support, C Linden (University of Würzburg) for FACS analysis of sponge symbiont cells and Michael Richter (Ribocon GmbH, Bremen) for useful bioinformatic advice. LGC Genomics (Berlin) is acknowledged for excellent customer services. Financial support to UH was provided by the US Department of Energy Joint Genome Institute, Office of Science of the Energy Department of California Technology Center (BaCaTeC). TW, CR, PS, NI and KM were funded by the US Department of Agriculture, and that these also have a major role in sponge cell aggregation (Müller et al., 1977). It might thus be possible that poribacterial enzymes interfere with mechanistic processes related to adhesion, aggregation and self–non-self recognition.

Conflict of Interest
The authors declare no conflict of interest.

Acknowledgements
We gratefully acknowledge the marine operations personnel at the Ruder Boskovic Institute (Rovinj/Croatia) for the help during sponge collection, Kristina Bayer for logistical support, C Linden (University of Würzburg) for FACS analysis of sponge symbiont cells and Michael Richter (Ribocon GmbH, Bremen) for useful bioinformatic advice. LGC Genomics (Berlin) is acknowledged for excellent customer services. Financial support to UH was provided by the US Department of Energy Joint Genome Institute, Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231.

References


Schloissinig S, Arumugam M, Sunagawa S, Mitrova M, Tap J, Zhu A et al. (2013). Genomic variation...


Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)