The 'zoo' accompanying *Botryococcus braunii*: Unraveling the fundamentals of algae-bacteria biocoenosis

Running title: Fundamentals of algae-bacteria biocoenosis

Olga Blifernez-Klassen[1,2], Viktor Klassen[1,2], Daniel Wibberg[2], Swapnil Chaudhari[1,2], Christian Henke[2,3], Anika Winkler[2], Arwa Al-Dilaimi[2], Oliver Rupp[4], Jochen Blom[4], Alexander Goesmann[4], Alexander Sczyrba[2,3], Andrea Bräutigam[2,5], Jörn Kalinowski[2], Olaf Kruse[1,2*]

[1]Algae Biotechnology and Bioenergy, Faculty of Biology, Bielefeld University, Universitätsstrasse 27, 33615 Bielefeld, Germany

[2]Center for Biotechnology (CeBiTec), Bielefeld University, Universitätsstrasse 27, 33615 Bielefeld, Germany

[3]Computational Metagenomics, Faculty of Technology, Bielefeld University, Universitätsstrasse 25, 33615 Bielefeld, Germany

[4]Bioinformatics and Systems Biology, Justus-Liebig-University, Heinrich-Buff-Ring 58, 35392 Gießen, Germany

[5]Computational Biology, Faculty of Biology, Bielefeld University, Universitätsstrasse 27, 33615 Bielefeld, Germany

*Address correspondence to: Olaf Kruse, Algae Biotechnology and Bioenergy, Faculty of Biology, Center for Biotechnology (CeBiTec), Bielefeld University, Universitätsstrasse 27, 33615 Bielefeld, Germany; phone: +49 521 106-12258; email: olaf.kruse@uni-bielefeld.de

Competing interests

The authors declare that they have no competing interests.

This work was supported by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement 311956 and by the BMBF-funded project "Bielefeld-Gießen Center for Microbial Bioinformatics" – BiGi (grant 031A533).

Abstract

The interactions between microalgae and bacteria represent a fundamental ecological relationship within aquatic environments, by controlling nutrient cycling and biomass production at the food web base. However, the extent of these interactions and their metabolic significance are poorly understood. To elucidate the complexity of microalga-bacteria interactions, *Botryococcus braunii* consortia were chosen as a suitable model system with an organic carbon-rich phycosphere. Here we present a comparative metagenome analysis in conjunction with physiological data that allow a deep insight into the community relationship. The reconstruction of ten high-quality draft and two full genomes of the abundant associates clarify consortia's basic functional characteristics based on the bacterial genetic repertoire. Implying a strong mutualistic algae-bacteria relationship, most abundant members of the *Botryococcus* phycosphere belong to a distinct niche of B-vitamin auxotrophs, essentially depending on the microalga to provide these important cofactors. This high dependency allows the microalga to effectively control its own phycosphere and impacts the bacteria-bacteria interactions within the consortium. Overall, our results provide highly valuable insights into the fundamentals of microbial ecology where metabolic networks cross species boundaries and define complex interactions. The knowledge of these interactions could enable the design of functional synthetic ecosystems consisting of photosynthetic and heterotrophic microorganisms.

Introduction

Microalgae are the dominant primary producers and the base of the food web within aquatic ecosystems and sustain in the natural environment in close relationship with multiple associated microorganisms [1, 2]. Algae-bacteria interactions, still poorly understood, are multifaceted and complex, and often take place in the microenvironment of individual algal cells, the phycosphere [2, 3]. The spectrum of ecological interrelations within the specific phycosphere ranges from cooperative to competitive [1], where algae and bacteria 'live together' within a framework of symbiosis [4]. The nature of the exchange of micro- and macronutrients, diverse metabolites including complex algal products such as polysaccharides, and infochemicals defines the relationship of the symbiotic partners [5, 6], which span mutualism, commensalism and parasitism [2]. Within a parasitic association, the bacteria are known to adversely affect the algae (*e.g*. cell wall degradation [7]), while in commensalism the partners can survive independently, with the commensal feeding on algae-provided substrates, but without harming the host [2]. Mutualistic interactions include *i.a.* the obligate relationships between vitamin-synthesizing bacteria and vitamin-auxotrophic phytoplankton species [8]. Many microalgae cannot synthesize several growth-essential vitamins (like B_{12} and B_1) [8–10] and obtain them through a symbiotic relationship with bacteria in exchange for organic carbon [11]. Not only the microalgae require these vitamins, but also many bacteria have to compete with other organisms for the B-vitamins [12].

However, in microbial environments, where the distinction between host and symbionts is less clear, the identification of the associated partners and the nature of their relationship is challenging. Looking for a suitable model system to elucidate the complex interactions between algae and bacteria, we focussed on the lifestyle of the colony-forming and hydrocarbon producing microalga *Botryococcus braunii*. This green microalga is known to co-exist with heterotrophic microorganisms [13]. The alga-bacteria consortium may represent an example for an association already existing during the early stages of life development on

earth, since it is generally considered that a progenitor of *Botryococcus* impacted the development of today's oil reserves and coal shale deposits [14]. Nowadays, this cosmopolitan microalga is widely considered of enormous biotechnological importance because of its unique ability to produce and secrete large quantities of aliphatic (C_{20}-C_{40}) hydrocarbons and exo-polysaccharides [15]. Due to the excretion of large quantities of organic carbon into its phycosphere, naturally occurring *B. braunii* cultures are generally not axenic but are associated with a variety of microorganisms [13]. The interactions between *B. braunii* and the associated bacteria are not fully understood, but are known to influence the microalgal growth performance and product formation capacity [16]. The associated bacteria may stimulate algal growth and product formation [17] or, conversely restrict algal growth by competing for nutrients, degrading algal exopolysaccharides and hydrocarbons, releasing anti-algal substances, or affecting algal morphology [13, 15].

To understand the complex nature of the unique habitat of a microalgal phycosphere and to unravel the food web between alga and bacteria, we investigated the consortium accompanying the hydrocarbon/carbohydrate-producing green microalga *Botryococcus braunii*. Two *B. braunii* races (A and B) were sequenced at different time points (linear and stationary growth) for the analysis and identification of the bacterial key players within communities and their ecological impact on the microalga.

Materials and Methods

Strains, cultivation conditions and growth determination parameters. Liquid, xenic cultures of *Botryococcus braunii* race A (CCALA778, isolated in Serra da Estrela (Baragem da Erva da Fome) Portugal) and race B (AC761, isolated from Paquemar, Martinique, France) as well as the axenic *B. braunii* race A strain (SAG30.81, Peru, Dpto.Cuzco, Laguna Huaypo) used for co-cultivation studies. The cultures were phototrophically cultivated under 16:8 light-dark illumination of 350±400 μmol photons m^{-2} s^{-1} white light in modified Chu-13 media

(without the addition of vitamin and organic carbon source) as described before [18]. Carbon supply and agitation were achieved by bubbling the cultures with moisture pre-saturated carbon dioxide-enriched air (5%, v/v) with a gas flow rate of 0.05 vvm. The axenic strain was regularly checked for contaminations.

The bacterial isolates were obtained by plating highly diluted xenic *B. braunii* cultures (CCALA778 and AC761) on LB agar plates. After two weeks single cell colonies were picked, separately plated and analyzed *via* 16S rDNA Sanger-sequencing by using the primer pair 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1385R (5'-CGGTGTGTRCAAGGCCC-3'). Two isolates, deriving from algal races A and B, were classified based on full length 16S rDNA similarity as *Mycobacterium* sp. and *Pimelobacter* sp., respectively and were used for further analysis.

Cultures growth was monitored by measurements of biomass organic dry weight (DW) by drying 10 mL culture per sample (in three biological and technical replicates) following the protocol according to Astals et al [19]. For the co-cultivation experiments, 500 mL axenic cultures (~0.25 gL^{-1}) were inoculated with the isolates (*Mycobacterium* sp. BbA or *Pimelobacter* sp. Bb-B (~$2x10^6$ cell/mL)) and cultured under photoautotrophic conditions for 21 days. The bacterial cell density was determined by manual cell counting using hemocytometer. Additionally, microalgae and bacteria cells were regularly checked by optical microscopy.

Hydrocarbon extraction and analysis. Hydrocarbon extraction was performed according to the protocol published previously [18]. In brief, 3.0 mL of culture was mixed with equal volume of acetone and *n*-hexane and vortexed for 2-3 min. After a heating (5 min at 85 °C) and centrifugation (5 min at 3020 ×*g*) step the upper *n*-hexane phase was collected and dried in a N_2-atmosphere. The dry extracted hydrocarbons were resuspended in 500 µL of *n*-hexane, containing an internal standard *n*-hexatriacontane ($C_{36}H_{74}$) and analysed *via* GC-MS and GC-FID. GC-MS analysis was performed with a TraceGC gas chromatograph and a ITQ

ion trap mass spectrometer equipped with an AS2000 auto sampler (ThermoFinnigan, Dreieich, Germany) using a 30-m×0.25-mm VF-5ms column with a 0.25 μm 5% diphenyl and 95% dimethylsiloxane coating (Varian Deutschland GmbH, Darmstadt, Germany). The method consisted of a temperature gradual ramp from 80 to 300 °C at 6 °C min^{-1}; at the final temperature of 300 °C, 1 μL sample was splitlessly injected. Mass spectra were recorded at 20 scans s^{-1} with a scanning range of 50–750 m/z. The evaluation of the chromatograms was accomplished with the Xcalibur software (ThermoFinnigan). Additionally, all samples were analyzed in a Shimadzu GC-2010 Plus (Shimadzu, Kyoto, Japan) unit with a FID detector as described previously [20]. Briefly, the separation was performed with a Macherey-Nagel OPTIMA FFAPplus column (30 m×0.25 mm ID, 0.40 mm OD and 0.50 μm film) and a 10 m pre-column (carbowax-deactivated, 0.25 mm ID, 0.40 mm OD, Macherey-Nagel). The sample injection volume was 1 μL working in split mode (1:5) and injected with an auto-injector (Shimadzu AOC-20i) at a constant pressure-mode of with 160 kPa with helium as carrier gas and nitrogen as makeup gas (Linde, Munich, Germany). The method consisted of a temperature ramp from 55 to 232 °C at 6 °C min^{-1} and to 250 °C at 3 °C min^{-1}, with keeping the final temperature of 250 °C constant for 29.5 min. For identification and quantification of all hydrocarbons external standards (alkane standard solution C_{21}-C_{40}, Sigma-Aldrich/Fluka) were used.

DNA sample collection and preparation. The samples from xenic *B. braunii* race A and B cultures (three biological replicates) were taken in the linear and stationary growth phases (after 9 and 24 days). Total community genomic DNA was extracted as previously described by Zhou et al [21]. The quality of the DNA was assessed by gelelectrophoresis and the quantity was estimated using the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen) and a Tecan Infinite 200 Microplate Reader (Tecan Deutschland GmbH).

High-throughput 16S rDNA amplicon sequencing. To get insight into the community composition of *B. braunii* Race A and B, high-throughput 16S rDNA amplicon sequencing

was performed as described recently by Maus et al [22]. The primer pair Pro341F (5′-CCTACGGGGNBGCASCAG-3′) and Pro805R (5′-GACTACNVGGGTATCTAATCC-3′) was used to amplify the hypervariable regions V3 and V4 of diverse bacterial and archaeal 16S rRNAs [23]. In addition, the primers also cover the 16S rDNA gene of algal chloroplasts and eukaryotic mitochondrial genomes. Furthermore, in a two PCR steps based approach, multiplex identifier (MID) tags and Illumina-specific sequencing adaptors were added to each amplicon. Only amplicons featuring a size of ~460 bp were purified using AMPureXP® magnetic beads (Beckman Coulter GmbH, Brea, California, USA). Resulting amplicons were qualitatively and quantitatively analyzed using the Agilent 2100 Bioanalyzer system (Agilent Inc., Santa Clara, California, USA) and pooled in equimolar amounts for paired-end sequencing on the Illumina MiSeq system (Illumina, San Diego, California USA). This sequencing approach provided ~150,000 reads per sample. An in-house pipeline as described previously [24] was used for adapter and primer trimming of all samples. For amplicon processing, a further in house amplicon processing pipeline including FLASH [25], USEARCH [26] (v8.1), UPARSE [27] and the RDP [28] classifier (v2.9) was applied as described recently [29, 30]. In summary, unmerged sequences resulting by FLASH [25] (default settings +-M 300) were directly filtered out. Furthermore, sequences with >1Ns (ambiguous bases) and expected errors >0.5 were also discarded. Resulting data was further processed and operational taxonomic units (OTUs) were clustered by applying USEARCH [26] (v8.1). The resulting OTUs were taxonomically classified using the RDP [28] classifier (v2.9) in 16S modus (Threshold >0.8) and compared to the nt database by means of BLASTN [31]. In the last step, raw sequence reads were mapped back onto the OTU sequences in order to get quantitative assignments (Supplementary data 1).

(Meta)Genome sequencing, assembly, binning and annotation. For each of the four samples (*B. braunii* race A and B from linear (T1) and stationary (T2) growth phase), the genomic DNA from three replicates was pooled. To obtain the metagenome sequence, four

whole-genome-shotgun PCRfree sequencing libraries (Nextera DNA Sample Prep Kit; Illumina, Munich, Germany) were generated based on the manufacturer's protocol representing different time points of the cultivation and different algae communities. The four libraries were sequenced in paired-end mode in a MiSeq run (2 x 300 bp) resulting in 11,963,558 and 11,567,552 reads for race A T1 and T2; 6,939,996 and 12,031,204 reads for race B T1 and T2, respectively (total 12.75 Gb). The *de novo* metagenome assembly was performed using the Ray Meta assembler [32] (v2.3.2) using a k-mer size of 41 and default settings. For contigs with length larger than 1 kb (n=127,914), the total assembly size was 338.5 Mb, the N50=2,884 and the largest contig 1.59 Mb.

All processed raw sequence reads of the four samples were aligned to the assembled metagenome contigs by means of Bowtie 2 [33] (v2.2.4). By using SAMtools [34] (v1.6), the SAM files were converted to BAM files. These files were sorted and read mapping statistics were calculated. The following portions of reads could be assembled to the draft metagenomes: 70.1% and 77.5% for race A (T1 and T2); 51.0% and 62.9% for race B (T1 and T2). For binning of the metagenome assembled genomes (MAGs), the tool MetaBAT [35] (v0.21.3) was used with default settings. Completeness and contamination of the resulting MAGs were tested with BUSCO [36, 37] (v3.0.), using the Bacteria and Metazoa data set (Table S1).

For sequencing of the isolated species, 4 µg of purified chromosomal DNA was used to construct PCR-free sequencing library and sequenced applying the paired protocol on an Illumina MiSeq system, yielding ~1.71 Mio Reads (517 Mbp in 3 Scaffolds, 58 Contigs) for *Mycobacterium* sp. Bb-A and ~1.70 Mio Reads (503 Mbp in 6 Scaffolds, 65 Contigs) for *Pimelobacter* sp. Bb-B. Obtained sequences were *de novo* assembled using the GS *de novo* Assembler software (v2.8, Roche). An *in silico* gap closure approach was performed as described previously [38]. Annotation of the draft chromosomes was performed within the PROKKA [39] (v1.11) pipeline and visualized using the GenDB 2.0 system [40].

Gene prediction, taxonomic assignment and functional characterization. Gene prediction on the assembled contigs of the MAGs was performed with Prodigal [41] (v2.6.1) using the metagenome mode ("-p meta"). All genes were annotated and analyzed with the in-house EMGB annotation system using the databases NCBI NR [31], Pfam [42] and KEGG [43]. Based on DIAMOND [44] (v0.8.36) hits against the NCBI-NR [31] database, all genes were subject to taxonomic assignment using MEGAN [45]. Relationships to metabolic pathways were assigned based on DIAMOND [44] hits against KEGG [43]. For Pfam [42] annotations, pfamscan was used with default parameters.The genomes were manually inspected using a E-value cutoff of $1x10^{-10}$ for several specific sequences listed in Table 1 and Supplementary data 4.

The profiling of the carbohydrate-active enzymes (CAZy) encoded by the *B. braunii* community members was accomplished using dbCAN2 [46] metaserver (Supplementary data 5). The identification, annotation and analysis of the secondary metabolite biosynthesis gene clusters within the MAGs and complete genomes were accomplished *via* the antiSMASH [47] (v4.0) web server with default settings (Supplementary data 7).

Phylogenetic analyses. To assign and phylogenetically classify the MAGs, a core genome tree was created by applying EDGAR [48] (v2.0). For calculation of a core-genome-based phylogenetic tree, the core genes of all MAGs and the two isolates *Mycobacterium* sp. Bb-A and *Pimelobacter* sp. Bb-B, corresponding to selected reference sequences were considered. The implemented version of FASTtree [49] (using default settings) in EDGAR [48] (v2.0) was used to create a phylogenetic Maximum-Likelihood tree based on 53 core genes (see Supplementary data 2) of all genomes. Additionally, the MAGs were taxonomically classified on species level by calculation of average nucleic and amino acid identities (ANI, AAI) by applying EDGAR [48] (v2.0) (Figure S3). For phylogenetic analysis of the 16S rDNA amplicon sequences, MEGA [50] (v7.0.26) was used. Phylogenetic tree was constructed by

means of Maximum-Likelihood Tree algorithm using Tamura 3-parameter model. The robustness of the inferred trees was evaluated by bootstrap (1000 replications).

Availability of data and materials

The high-throughput 16S rDNA amplicon sequencing datasets obtained during the present work are available at the EMBL-EBI database under BioProjectID PRJEB21978. The metagenome datasets of *Botryococcus braunii* as well as the resulting MAGs have been deposited under the BioProjectIDs PRJEB26344 and PRJEB26345, respectively. The communities most abundant ten high-quality draft genomes are accessible via EMGB platform (https://emgb.cebitec.uni-bielefeld.de/Bbraunii-bacterial-consortium/). The genomes of the isolates *Mycobacterium* sp. BbA or *Pimelobacter* sp. Bb-B are accessible under BioProjectIDs PRJEB28031 und PRJEB28032, respectively.

Results

Effect of the bacterial community on the *B. braunii* host

We started our investigations with the analysis of an axenic *Botryococcus braunii* strain to assess algal capacities for growth and hydrocarbon production under strict phototrophic conditions (*e.g.* without the supplementation of vitamins or organic carbon). Axenic *B. braunii* showed a continuous increase in cell biomass and hydrocarbon content, yielding 2.1±0.2 gL^{-1} and 0.7±0.1 gL^{-1}, respectively (Figure 1a).

In a next step, we co-cultivated the axenic strain with two isolated actinobacterial associates from *B. braunii* (*Mycobacterium* sp. Bb-A and *Pimelobacter* sp. Bb-B, see Methods). *B. braunii* in bacteria-free conditions (axenic culture) accumulated hydrocarbons up to 38% of dry weight (DW) during the cultivation, while the cultures containing bacteria (xenic cultures) accumulated only very limited hydrocarbons during the bacterial exponential growth phase (up to 5% of DW for co-cultures with bacteria, Figure 1). In the stationary phase of bacterial growth, the alkadiene $C_{31}H_{60}$ accumulated again to a low level. Of

particular note was that the algal cultures changed their morphology during bacteria co-cultivation by reducing colony formation when compared to axenic growth.

This dramatic decrease in hydrocarbon content and disturbed colony formation was not observed in the xenic *Botryococcus braunii* cultures, from which both bacterial isolates are derived (Figure S1). To test the metabolic capabilities and possible interactions in the *B. braunii* consortia of races A and B at linear and stationary algal growth, we assessed the bacteria associated with each culture using metagenomics.

Metagenomic survey of the *Botryococcus braunii* consortia

We profiled four different communities using high-throughput 16S rDNA gene amplicon data and added 12.75 Gb of paired-end metagenome data. The analysis of the 16S rDNA amplicons and a read-based metagenome approach showed that each of the *B. braunii* communities represents a distinct large consortium, consisting of a single algal and various bacterial species as well as traces of Archaea and viruses (Figure S2). Each community is comprised of four bacterial phyla, consisting of at least 25 species with the members of the phylum Proteobacteria being the most prominent in both *B. braunii* races. Of 33 different bacterial genera, four were specific to race A, seven were specific to race B and 22 were shared (Figure 2). The relative abundance of plastid 16S rDNA sequences accounted up to 49 and 59% of all detected amplicon reads in race A and B samples in the linear growth phase and declined to approximately 26% and 44% at stationary phase, respectively (Figure S2a). At the same time, the 16S rDNA contribution of the accompanying microbial community increased. The community members varied quantitatively based on both race and time of sampling, but not qualitatively.

Both *B. braunii* races were dominated by the genera *Devosia* and *Dyadobacter* (up to 26% and 22% of the bacterial community) in the logarithmic growth phase, followed by *Dokdonella* in race A (16%) and *Acidobacteria* in race B (22%) (Figure 2). During the

cultivation, the relative abundancy of the dominant species in the race A consortium such as *Devosia* and *Dokdonella* increased at the expense of *Dyadobacter* in comparison to the linear growth phase. In the race B consortium, less abundant taxa such as *Mesorhizobium*, *Brevundimonas* and *Hydrogenophaga* increased at the expense of *Dyadobacter* and *Devosia*.

The detected relative levels of both actinobacterial isolates, which had been investigated in co-cultivation experiments with the axenic algae (Figure 1) were comparatively low, although both *B. braunii* races are capable to provide a wealth of organic substrate (in form of hydrocarbons and/or carbohydrates) to ensure their proliferation (Figure S1). *Mycobacterium* sp. Bb-A isolate, obtained from the *B. braunii* race A community, was almost exclusively propagated therein and increased during stationary growth to 7% of all detected bacterial amplicons (Figure 2). *Pimelobacter* sp. Bb-B was isolated from the race B consortium, and is present in both communities in equally (≤1%) low abundances which increased slightly during stationary phase.

To test the relationships between *B. braunii*, *Mycobacterium*, *Pimelobacter* and the remaining consortium, we explored the genetic potential of the abundant representatives of algal-bacterial consortia. We combined metagenome *de novo* assembly and differential coverage and tetranucleotide signature based binning of the *B. braunii* race A and B samples. This approach resulted in the reconstruction of ten high-quality (completeness >80%, contamination <10%) and ten fragmentary metagenome assembled genomes (MAGs) as well as seven eukaryotic genome fragments (for details see Methods, Table S1). To identify taxa representing the ten high-quality draft MAGs, we constructed a phylogenetic tree built out of a core of 53 highly conserved marker genes per genome (Figure 3). All abundant taxa shared by both races were recovered, including the most abundant alphaproteobacterial MAG of the order *Rhizobiales*, designated Dev-G (similar to *Devosia* sp. 66-14, Table S2 and Figure S3). Additionally, MAGs of *Dyadobacter* and *Hydrogenophaga* were constructed (Dya-G and Hyd-G, similar to *Dyadobacter* sp. 50-39 and *Hydrogenophaga* sp. PML113, respectively).

Another two alphaproteobacterial MAGs Bre1-G and Bre2-G of the order *Caulobacteriales* (similar to *Brevundimonas* sp. DS20), clustered together within the 16S rDNA amplicon approach (Figure 2) but were separated through the binning procedure (Tables S1 and S2). We reconstructed sphingobacterial MAGs of the order *Chitinophagales*, Ter1-G and Ter2-G similar to different extent to *Terrimonas ferruginea* DSM30193. Further two MAGs, basically occurring in race A strain, included a gammaproteobacterial bacterium similar to *Dokdonella koreensis* DS-123 (Dok-G) and a member of *Corynebacteriales* similar to *Mycobacterium* sp. Root135 (Myc-G). An exclusively in *B. braunii* race B occurring member of the subdivision IV of the phyla Acidobacteria (Gp4-G) is similar to *Pyrinomonas methylaliphatogenes* K22.

We also sequenced the two actinobacterial isolates, *Mycobacterium* sp. Bb-A and *Pimelobacter* sp. Bb-B, used for the co-cultivation experiments (Figure 1). After assembly and *in silico* finishing, both genomes were completed and represented by a single chromosome, each (Figure S4). The *Mycobacterium* sp. Bb-A genome sequence was identical to the Myc-G, indicating that both represent the same species (Figure S3h), thus underlining the quality of our metagenome assembly and binning approach.

Elucidation of the genetic portfolio of the bacterial community

To identify genes that enabled the bacterial community members to contribute to and to benefit from the food web of the *B. braunii* algae based consortium, we performed a KEGG-based quantitative functional assignment of the annotated high-quality MAGs and sequenced genomes by applying the elastic metagenome browser (EMGB, https://emgb.cebitec.uni-bielefeld.de/Bbraunii-bacterial-consortium/).

Six of twelve genomes carried genes for bacterial chemotaxis and five of twelve genes for flagella synthesis, hinting at active motility and ability to move towards substrates in the medium (Figure S5). *Botryococcus braunii* synthesizes aliphatic long-chain hydrocarbons, which accumulate in large quantities in the extracellular matrix of this species [15]. Seven of

twelve genomes carried gene homologs of already characterized hydrocarbon hydroxylases (*alkB*, *ladA*, *almA* and/or cytochrome P450 (CYP153), Table 1) for the initial oxyfunctionalization of long-chain hydrocarbons [51]. While the most abundant consortia members Dev-G and Dya-G exhibited a low enzyme number, the more rare *Mycobacterium* and *Pimelobacter* genomes contained a large portfolio of hydrocarbon monooxygenase/ hydroxylase genes, and were the only genomes to carry non-heme membrane-associated monooxygenase *alkB* genes (Table 1).

Mucilaginous (exo-)polysaccharides as part of the internal fibrillar layer of the *B. braunii* cell wall represent another prospective carbon source for the bacterial community, consisting mainly of galactose but also of fucose and rhamnose [15, 18]. The profiling of the carbohydrate-active enzymes (CAZymes [52]), encoded by the members of *B. braunii* communities using dbCAN2 [46] metaserver, revealed the presence of several putative enzyme families involved in the complex carbohydrate metabolism (Table 1).

Eight out of twelve genomes coded for several carbohydrases for the degradation of poly- and oligosaccharides, generally present in the cell walls of microalgae, and contained genes for chitin disintegration and assimilation (Supplementary data 5). However, six species showed a clear prevalence in the number of enzymes involved in the biosynthesis of carbohydrates (glycosyltransferases) compared to degradation (glycoside hydrolases). Some species like the members of Bacteroidetes contained a large number of CAZyme-encoding genes in their genomes (up to 5.9% of total gene content) with a clear prevalence to hydrolytic activity of complex carbohydrates (Table 1). For instance, the Dya-G genome contains more genes encoding for the glycoside hydrolase (GH) families involved in the degradation of complex polymers (cellulose, xylan, chitin), while others, such as the genome of Dev-G, harbor enzymes acting on oligosaccharides [52] (Supplementary data 5).

Eight of twelve genomes encode genes for the synthesis of cyanophycin granule polypeptides (CGPs) [53], however only four contained genes for the degradation of this C-

and N-storage compound (Table 1). Related to this, all genomes carry genes for the synthesis but not for the degradation of polyphosphate granules (Poly-P), a polymeric reserve with a significant role in the regulation of enzymatic activities, gene expression and stress adaptation processes [54]. Suggesting an adaptation to phosphorous and iron limitation, all genomes contained genes relevant for phosphate starvation (*pst* system) or siderophore biosynthesis (*fhu* system).

Looking at the primary metabolism, all genomes encode genes for processes essential for aerobic respiration (Figure S5, Supplementary data 3). Genes that enable pyruvate fermentation to lactate, acetate and ethanol are also present, showing a versatility of the community members to switch between aerobic and anaerobic metabolisms depending on the oxygen availability. All except *Mycobacterium* genomes contained genes for both the aerobic cobalamin (vitamin B_{12})-independent form and the oxygen-independent but B_{12}-dependent type of ribonucleotide reductases [55] (Table 1), demonstrating adaptations to a facultative aerobic life style of the *B. braunii* associates.

We were unable to reconstruct the complete cobalamin (vitamin B_{12}) synthesis pathway for eight out of twelve assembled genomes. The absence of the *metE* gene for the B_{12}-independent methionine synthesis pathway in these genomes (except Dya-G, Bre1-G and Hyd-G) implies the need of exogenous supply of vitamin B_{12} for survival (Table 1). Seven of twelve genomes lack the complete gene portfolio for the biotin (vitamin B_7) synthesis, but contained the complete biosynthesis pathway of fatty acids for which biotin is essential [12]. Four of twelve genomes were incomplete in the synthesis of thiamin (vitamin B_1), which is essential for citrate cycle and pentose phosphate pathway [12]. In addition to vitamin B_1 and B_7-auxotrophy, the most abundant Dev-G species was found to comprise incomplete *de novo* biosynthesis pathways of riboflavin (vitamin B_2, essential as the FMN/FAD precursor) and pantothenate (vitamin B_5, essential for CoA synthesis) [56] (Table 1).

Additionally, eleven out of twelve genomes encoded for bacteriocins, antimicrobial peptides with significant potency against other bacteria, which can exhibit narrow spectra of activity (targeting members of the same species) or display broader activity spectra (targeting other species and genera) [57]. Especially, Hyd-G, *Mycobacterium* sp. Bb-A and *Pimelobacter* sp. Bb-B exhibited a large portfolio of metabolic gene cluster highly similar to antibiotic compounds with a broad spectrum of activity against Gram-positive bacteria which are exceedingly rare in the communities of *B. braunii*.

Collectively, within our survey of the microbial community accompanying the *Botryococcus braunii* races A and B, we made complex observations, which hint at interconnected food webs and are summarized in Figure 4. The genetic portfolio of *Mycobacterium* sp. Bb-A and *Pimelobacter* sp. Bb-B reflect the strong metabolization of hydrocarbons (Figure 1, Table 1). However, despite the apparent vitamin-autotrophy and the pronounced proliferation in co-cultivation experiments with the microalga, both actinobacterial strains were poorly represented in the consortia of both *B. braunii* races. On the other hand, Gram-negative bacterial species such as *Devosia* (Dev-G) and *Dyadobacter* (Dya-G) that displayed several vitamin auxotrophies (especially B_7) and would not be viable on their own, dominated the algal communities. These results clearly indicate that both *Actinobacteria* are actively restrained by the abundant consortia members, possibly realized by the synthesis and secretion of secondary metabolites (*e.g.* bacteriocines).

Discussion

Algae exist in the natural environment as part of a complex microbial consortium, where the different species may exert considerable influence on each other for the exploitation of unique biological functions and/or for the metabolite exchange [1].

B. braunii bacterial communities appear to be highly conserved and illustrate relatively low species biodiversity, with most of the detected genera occurring in both communities,

albeit differentially abundant (Figure 2). The re-processing of the *B. braunii* Guadeloupe strain metagenome data [58] also shows data qualitatively but not quantitatively similar to the two *Botryococcus* consortia analyzed in this study. The direct comparison of the different metagenome datasets revealed the coinciding presence of the individual genera (especially members of Bacteroidetes, Alpha- and Betaproteobacteria), although massively varying in the abundance pattern (Figure S6). The communities did differ regarding the presence of the Gram-positive bacterial representatives: while the consortia obtained from our studies harbored the members of *Actinobacteria*, the members of the phyla *Firmicutes* were strongly represented in the Guadeloupe strain [58]. This evaluation of three *B. braunii* consortia under changing conditions has emphasized that depending on the physiological condition (*e.g.* supplementation of vitamins or antibiotics [58]), the relative abundance of certain species may widely vary, suggesting the existence of many bacterial taxa with functional redundancy [59].

Many genera identified within the *Botryococcus* communities (such as *Devosia*, *Dyadobacter*, *Mesorhizobium* etc.) also occur in other algae-bacterial consortia like biofilms [60], perhaps indicating preferential coexistence between certain bacteria species and microalgae. The interdependence of certain bacterial phyla such as Alphaproteobacteria and Bacteroidetes observed in this study, which had been frequently found to be associated with phytoplankton, underlines the proposed existence of the ‘archetypal microalga/phytoplankton-associated bacterial taxa’ [61, 62].

At first glance, the *Botryococcus* alga does not appear obligatory dependent on a bacterial community, since the axenic strains can be fully photoautotrophically cultured without vitamin supplements (Figure 1). However, *Botryococcus* readily releases huge amounts of organic carbon (up to 60% of DW in form of hydrocarbons and carbohydrates) into the extracellular milieu (Figure S1), which creates a phycosphere that naturally attracts many microorganisms, including those with potential damaging effects. Evolutionary fitness will depend on efficient management of the “bacteria zoo” by *B. braunii*.

From the perspective of a microalga, the associated bacteria can be providers of micro- and macronutrients but also competitors for limiting nutrients [6]. Based on their observed generic potential, individual species appear to be involved in the acquisition and remineralization of macronutrients and in the storage of nitrogen and phosphorous (in form of cyanophycin and Poly-P, Table 1), which benefit the microalga.

We post that the microalga supplies vitamins B_1, B_2, B_5 and especially B_7 to its resident community [12]. The abundant community members of the *B. braunii* consortia are auxotrophs for various B-vitamins making them dependent on an exogenous supply of vitamins for survival (Table 1). The consortium's most abundant bacterial species *Devosia* (Dev-G, Figure 2) is indeed auxotroph for vitamins B_1, B_2, B_5 and B_7, but can synthesize vitamin B_{12}. In view of the fact that the bacterial abundance determines their functional relevance for the community [63, 64], this species is also likely to provide both the bacterial community and the microalga with this vitamin [8], thus confirming the recent proposition that the B_{12} supply is realized only by a few certain bacteria within the alga-bacteria consortia [60]. In this context, since the B_7-auxotrophs are prevalent within the detected consortia (except for Dok-G, appearing only in race A), it is doubtful that the large required supply of vitamin B_7 is of bacterial origin. Because of its predominance within the community (Figure S2), the microalga is more likely to secrete and provide this vitamin to the bacteria, at least the secretion of the B-vitamins (*e.g.* B_5 and B_7) has already been observed in species such as *Chlamydomonas*, *Chlorella* and *Coccomyxa* [65, 66]. The high requirement [64] of Dev-G for other B-vitamins will most likely also be supplemented by the microalga. In fact, the bacterial dependence on algal vitamins has been suggested earlier in connection with marine phytoplankton communities [12]. However, it can not be ruled out that the missing cofactors may be complemented by other bacteria in exchange for *e.g.* B_{12}, suggesting also strong bacteria-bacteria relations in addition to the algae-bacteria interactions. Considering this view, specific bacteria that are requiring particular B-vitamins are deliberately promoted by

Botryococcus braunii in exchange for *e.g.* vitamin B_{12}, micro- and macronutrients as well as essential metabolites. The presence of particular antibacterial and probiotic compounds (Table 1) explains, why *B. braunii* communities consist mainly of Gram-negative bacteria and only a few low abundant Gram-positive species (*e.g. Mycobacterium* and *Pimelobacter*, Figure 2). Thus, the *B. braunii* associative consortia appears restrained by the microalgae, yet displaying strong bacteria-bacteria interactions and capable to synthesize and provide a great deal of substances to "fight" the potential competitors.

Based on the genetically encoded features, the bacterial associates of *B. braunii* can be apportioned into two roughly framed categories of *i*) B-vitamin-auxotrophic (controllable), and *ii*) B-vitamin-independent (uncontrollable) species (Figure 4). The first category includes the members of the class of Alphaproteobacteria (like Dev-G), which display several B-vitamin-auxotrophies, but are delivering B_{12} to the algae (and bacteria). Bacterial associates, which are incapable to synthesize essential cofactors (including B_{12}), rely heavily on both, the microalga and the associated bacteria. Although non-viable on their own, these phylotypes belonging to the Proteobacteria, Bacteroidetes and Acidobacteria are dominating the consortiae, indicating a very strong mutualistic relationship with the microalga. Completely autonomous species like our isolates *Mycobacterium* sp. Bb-A and *Pimelobacter sp.* Bb-B are assigned to the second category. These phylotypes are known to be hydrocarbonoclastic parasitic microorganisms targeting mainly the organic carbon source and therefore pose a potential risk to the microalga since they are uncontrollable, strongly implying a parasitic relationship. And although both isolates can in principle also supply the microalgae with vitamin B_{12}, nutrients and lots of antibacterial metabolites, they are harmful to the alga due to their destructive effect through uncontrolled proliferation and substrate consumption, as demonstrated in the co-cultivation experiments (Figure 1). Their activity causes a dramatic change/damage to the algal colony morphology, resulting in a less complex colony structure,

thus having a negative impact on survival in natural environments, *i.e.* a higher susceptibility to grazers [67].

Conclusions

While *B. braunii* can be grown axenically, naturally occurring populations exist together with their bacterial consortia. The interaction between the microalgae and bacteria within the *Botrycoccus braunii* consortiae display a complex interplay between all members, however based on our observtions the microalgae seems to be the "conductor" in this play, attracting specifically vitamin-auxotrophic bacterial species and keeping them in close proximity by controlling through the subsequent vitamin supply. Metagenome analysis indicated a "zoo" of largely supportive bacteria which supply vitamin B_{12}, and assimilated N and P in exchange for other vitamins and carbon from the alga. In addition to this benign members, autark bacteria without vitamin auxotrophies are present in small proportions, likely decimated by antibacterial agents from the supportive community, and attack the hydrocarbon/ polysaccharide coat of the alga. War and peace play out amongst this microscopic community that keeps the alga well enough to sustain collective growth.

Based on the nature of the presented observations, we unraveled only the obvious interrelations, which are encoded in the bacterial genomes of the conserved consortiae of *B. braunii*. The interactions between the bacteria and microalgae (phytoplankton in general) are far more complex and multifarious, and may involve many physiological exchanges and communications, which are not solely educible from the genetic predisposition and should be subject for future research. But, with the present work we provide a rough framework for the basic interdependence between a microalga (phytoplankton) and its associated bacteria, which greatly expands the current knowledge in this area.

Acknowledgements

We are thankful to Joao Gouveia and Douwe van der Veen (Wageningen University, The Netherlands) for providing *Botryococcus braunii* race A and B (CCALA778 and AC761) stock cultures. This work was supported by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement 311956 (relating to project "SPLASH – Sustainable PoLymers from Algae Sugars and Hydrocarbons"). The bioinformatics support by the BMBF-funded project "Bielefeld-Gießen Center for Microbial Bioinformatics" – BiGi (grant 031A533) within the German Network for Bioinformatics Infrastructure (de. NBI) is gratefully acknowledged.

Competing interests

The authors declare that they have no competing interests

Authors' contributions

O.B.K., V.K, and O.K. conceived this study. O.B.K., A.S., J.K. and O.K. supervised experiments and analyses. O.B.K., V.K. and S.C. accomplished all physiological studies. O.B.K., D.W., A.S., C.H., A.B., V.K., A.A.D., A.W., J.K., O.R., J.B. and A.G. performed the sequencing and the bioinformatic analyses. O.B.K., V.K. and O.K. wrote the original draft manuscript with contributions from all authors. O.B.K., V.K., A.B and O.K. reviewed and edited the manuscript. All authors read and approved the final manuscript.

Supplementary information

Supplementary information is available.

References

1. Seymour JR, Amin SA, Raina JB, Stocker R. Zooming in on the phycosphere: The ecological interface for phytoplankton-bacteria relationships. *Nat Microbiol* . 2017. , **2**

2. Ramanan R, Kim B-H, Cho D-H, Oh H-M, Kim H-S. Algae–bacteria interactions: Evolution, ecology and emerging applications. *Biotechnol Adv* 2016; **34**: 14–29.

3. Bell W, Mitchell R. Chemotactic and Growth Responses of Marine Bacteria to Algal Extracellular Products. *Biol Bull* 1972; **143**: 265–277.

4. Cooper MB, Smith AG. Exploring mutualistic interactions between microalgae and bacteria in the omics age. *Curr Opin Plant Biol* . 2015. , **26**: 147–153

5. Amin SA, Hmelo LR, Van Tol HM, Durham BP, Carlson LT, Heal KR, et al. Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature* 2015; **522**: 98–101.

6. Cole JJ. Interactions Between Bacteria and Algae in Aquatic Ecosystems. *Annu Rev Ecol Syst* 1982; **13**: 291–314.

7. Wang X, Li Z, Su J, Tian Y, Ning X, Hong H, et al. Lysis of a red-tide causing alga, Alexandrium tamarense, caused by bacteria from its phycosphere. *Biol Control* 2010; **52**: 123–130.

8. Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG. Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature* 2005; **438**: 90–93.

9. Croft MT, Warren MJ, Smith AG. Algae need their vitamins. *Eukaryot Cell* . 2006. , **5**: 1175–1183

10. Tang YZ, Koch F, Gobler CJ. Most harmful algal bloom species are vitamin B1 and B12 auxotrophs. *Proc Natl Acad Sci* 2010; **107**: 20756–20761.

11. Grant MAA, Kazamia E, Cicuta P, Smith AG. Direct exchange of vitamin B12 is demonstrated by modelling the growth dynamics of algal–bacterial cocultures. *ISME J* 2014; **8**: 1418–1427.

12. Sañudo-Wilhelmy SA, Gómez-Consarnau L, Suffridge C, Webb EA. The Role of B Vitamins in Marine Biogeochemistry. *Ann Rev Mar Sci* 2014; **6**: 339–367.

13. Chirac C, Casadevall E, Largeau C, Metzger P. Bacterial Influence Upon Growth and Hydrocarbon Production of the Green Alga Botryococcus Braunii. *J Phycol* . 1985. , **21**: 380–387

14. McKirdy DM, Cox RE, Volkman JK, Howell VJ. Botryococcane in a new class of Australian non-marine crude oils. *Nature* 1986; **320**: 57–59.

15. Banerjee A, Sharma R, Chisti Y, Banerjee UC. Botryococcus braunii: A renewable source of hydrocarbons and other chemicals. *Crit Rev Biotechnol* 2002; **22**: 245–279.

16. Fuentes JL, Garbayo I, Cuaresma M, Montero Z, González-Del-Valle M, Vílchez C. Impact of microalgae-bacteria interactions on the production of algal biomass and

associated compounds. *Mar Drugs* . 2016. , **14**

17. Tanabe Y, Okazaki Y, Yoshida M, Matsuura H, Kai A, Shiratori T, et al. A novel alphaproteobacterial ectosymbiont promotes the growth of the hydrocarbon-rich green alga Botryococcus braunii. *Sci Rep* 2015; **5**.

18. Blifernez-Klassen O, Chaudhari S, Klassen V, Wördenweber R, Steffens T, Cholewa D, et al. Metabolic survey of Botryococcus braunii: Impact of the physiological state on product formation. *PLoS One* 2018; **13**: e0198976.

19. Astals S, Nolla-Ardèvol V, Mata-Alvarez J. Anaerobic co-digestion of pig manure and crude glycerol at mesophilic conditions: Biogas and digestate. *Bioresour Technol* 2012; **110**: 63–70.

20. Schwarzhans JP, Cholewa D, Grimm P, Beshay U, Risse JM, Friehs K, et al. Dependency of the fatty acid composition of Euglena gracilis on growth phase and culture conditions. *J Appl Phycol* 2014; **27**: 1389–1399.

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

22. Maus I, Kim YS, Wibberg D, Stolze Y, Off S, Antonczyk S, et al. Biphasic Study to Characterize Agricultural Biogas Plants by High-Throughput 16S rRNA Gene Amplicon Sequencing and Microscopic Analysis. *J Microbiol Biotechnol* 2017; **27**.

23. Takahashi S, Tomita J, Nishioka K, Hisada T, Nishijima M. Development of a prokaryotic universal primer for simultaneous analysis of bacteria and archaea using next-generation sequencing. *PLoS One* 2014; **9**: e105592.

24. Wibberg D, Andersson L, Rupp O, Goesmann A, Pühler A, Varrelmann M, et al. Genome analysis of the sugar beet pathogen Rhizoctonia solani AG2-2IIIB revealed high numbers in secreted proteins and cell wall degrading enzymes. *BMC Genomics* 2016; 245.

25. Magoč T, Salzberg SL. FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011; **27**: 2957–2963.

26. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010; **26**: 2460–2461.

27. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 2013; **10**: 996–8.

28. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007; **73**: 5261–5267.

29. Liebe S, Wibberg D, Winkler A, Puehler A, Schlueter A, Varrelmann M. Taxonomic analysis of the microbial community in stored sugar beets using high-throughput sequencing of different marker genes. *FEMS Microbiol Ecol* 2016; **92**: 1–12.

30. Klassen V, Blifernez-Klassen O, Wibberg D, Winkler A, Kalinowski J, Posten C, et al. Highly efficient methane generation from untreated microalgae biomass. *Biotechnol*

Biofuels 2017; **10**: 186.

31. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 2016; **44**: D733–D745.

32. Boisvert S, Raymond F, Godzaridis E, Laviolette F, Corbeil J. Ray Meta: scalable de novo metagenome assembly and profiling. *Genome Biol* 2012; **13**: R122.

33. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012; **9**: 357–359.

34. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.

35. Kang DD, Froula J, Egan R, Wang Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 2015; **3**: e1165.

36. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V., Zdobnov EM. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 2015; **31**: 3210–3212.

37. Waterhouse RM, Seppey M, Simao FA, Manni M, Ioannidis P, Klioutchnikov G, et al. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol* 2018; **35**: 543–548.

38. Wibberg D, Blom J, Jaenicke S, Kollin F, Rupp O, Scharf B, et al. Complete genome sequencing of Agrobacterium sp. H13-3, the former Rhizobium lupini H13-3, reveals a tripartite genome consisting of a circular and a linear chromosome and an accessory plasmid but lacking a tumor-inducing Ti-plasmid. *J Biotechnol* 2011; **155**: 50–62.

39. Seemann T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30**: 2068–2069.

40. Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, et al. GenDB - An open source genome annotation system for prokaryote genomes. *Nucleic Acids Res* 2003; **31**: 2187–2195.

41. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010; **11**.

42. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: The protein families database. *Nucleic Acids Res* . 2014. , **42**

43. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 2017; **45**: D353–D361.

44. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2015; **12**: 59–60.

45. Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, et al. MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput Biol* 2016; **12**.

46. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, et al. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 2018; 1–7.

47. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, et al. antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* 2017; **45**: W36–W41.

48. Blom J, Kreis J, Spänig S, Juhre T, Bertelli C, Ernst C, et al. EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Res* 2016; **44**: W22–W28.

49. Price MN, Dehal PS, Arkin AP. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* 2010; **5**.

50. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 2016; **33**: 1870–1874.

51. Rojo F. Enzymes for Aerobic Degradation of Alkanes. In: Timmis KN (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. 2010. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 781–797.

52. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 2014; **42**.

53. Rehm BHA. Bacterial polymers: Biosynthesis, modifications and applications. *Nat Rev Microbiol* . 2010. , **8**: 578–592

54. Kulakovskaya T V., Vagabov VM, Kulaev IS. Inorganic polyphosphate in industry, agriculture and medicine: Modern state and outlook. *Process Biochem* . 2012. , **47**: 1–10

55. Torrents E. Ribonucleotide reductases: essential enzymes for bacterial life. *Front Cell Infect Microbiol* 2014; **4**.

56. Webb ME, Marquet A, Mendel RR, Rébeillé F, Smith AG. Elucidating biosynthetic pathways for vitamins and cofactors. *Nat Prod Rep* . 2007. , **24**: 988–1008

57. Cotter PD, Ross RP, Hill C. Bacteriocins-a viable alternative to antibiotics? *Nat Rev Microbiol* . 2013. , **11**: 95–105

58. Sambles C, Moore K, Lux TM, Jones K, Littlejohn GR, Gouveia JD, et al. Metagenomic analysis of the complex microbial consortium associated with cultures of the oil-rich alga Botryococcus braunii. *Microbiologyopen* 2017; **6**.

59. Allison SD, Martiny JBH. Resistance, resilience, and redundancy in microbial communities. *Proc Natl Acad Sci* 2008; **105**: 11512–11519.

60. Krohn-Molt I, Alawi M, Förstner KU, Wiegandt A, Burkhardt L, Indenbirken D, et al. Insights into Microalga and bacteria interactions of selected phycosphere biofilms

using metagenomic, transcriptomic, and proteomic approaches. *Front Microbiol* 2017; **8**.

61. Buchan A, LeCleir GR, Gulvik CA, González JM. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat Rev Microbiol* . 2014. , **12**: 686–698

62. Ramanan R, Kang Z, Kim B-H, Cho D-H, Jin L, Oh H-M, et al. Phycosphere bacterial diversity in green algae reveals an apparent similarity across habitats. *Algal Res* 2015; **8**: 140–144.

63. Reed H, Martiny J. Testing the functional significance of microbial composition in natural communities. *FEMS Microbiol Ecol* 2007; **62**: 161–170.

64. Rivett DW, Bell T. Abundance determines the functional role of bacterial phylotypes in complex communities. *Nat Microbiol* 2018; **3**: 767–772.

65. Bednar TW, Holm-Hansen O. Biotin liberation by the lichen aga Coccomyxa sp. and by Chlorella pyrenoidos. *Plant Cell Physiol* 1964; **5**: 297–303.

66. Aaronson S, Dhawale SW, Patni NJ, Deangelis B, Frank O, Baker H. The cell content and secretion of water-soluble vitamins by several freshwater algae. *Arch Microbiol* 1977; **112**: 57–59.

67. Pančić M, Kiørboe T. Phytoplankton defence mechanisms: traits and trade-offs. *Biol Rev* 2018; **93**: 1269–1303.

Figure legends

Figure 1: Co-cultivation of axenic *B. braunii* with bacterial associates *Mycobacterium* sp. Bb-A and *Pimelobacter* sp. Bb-B. Determination of algal growth and product formation performance of **a.** axenic cultures and **b.** xenic cultures supplemented with bacterial isolates. The progress of the bacterial growth is represented by the number of cells on a logarithmic (log_{10}) scale.The error bars represent standard error of mean values of three biological and three technical replicates (SE; n=9).

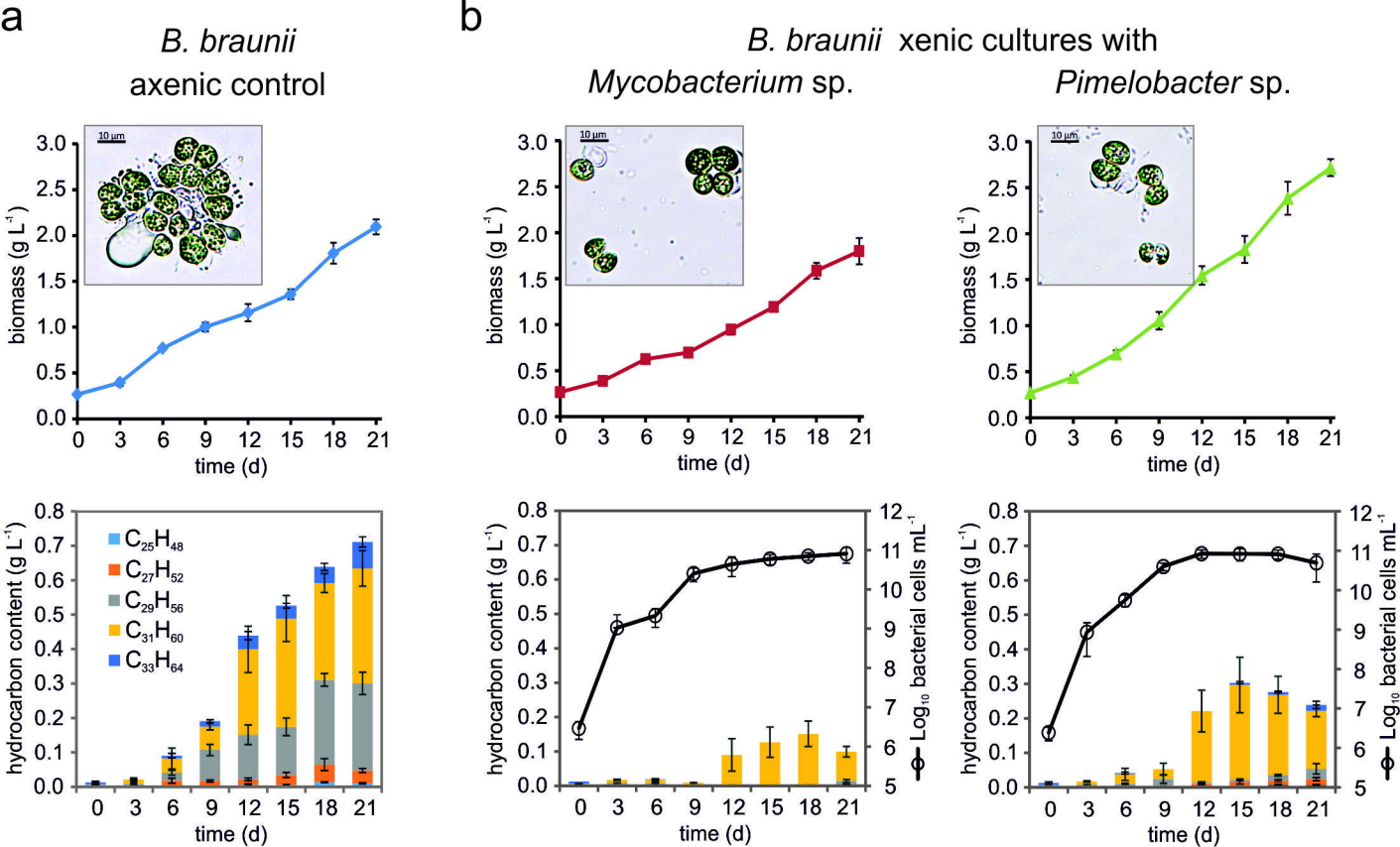
Figure 2: Bacterial representatives accompanying the *B. braunii* consortiae. Illustration of the relative abundance and phylogenetic distribution of bacterial genera detected *via* high-throughput 16S rDNA amplicon sequencing approach (see Supplementary data 1). Bacterial genera labeled green or red are represented only in the consortia of race A or B, respectively. The abundant members among the communities are marked bold. Black, grey and white

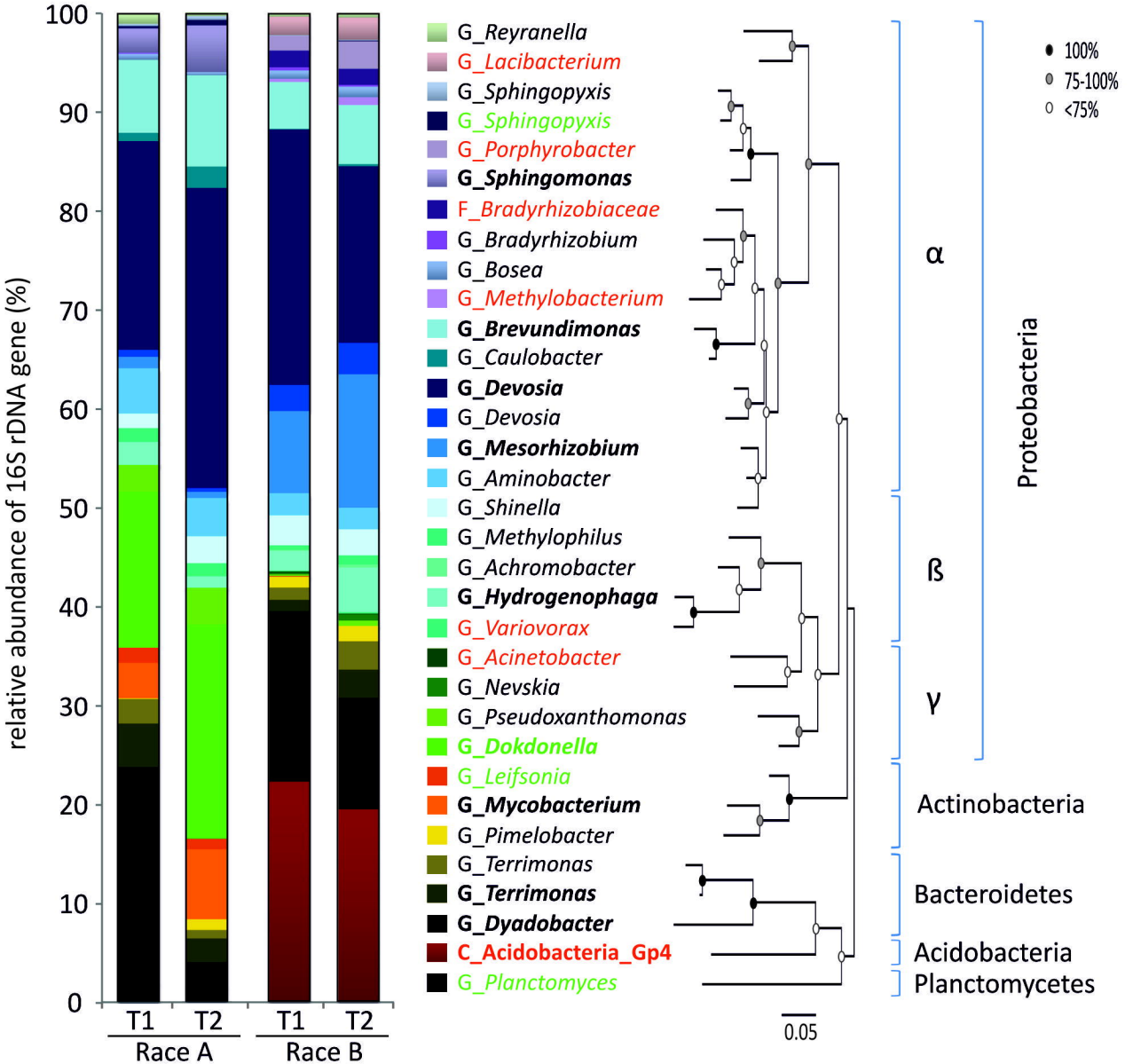
circles represent the nodes with calculated bootstrap values of 100, 75-100 and <75% (1,000 replicates).

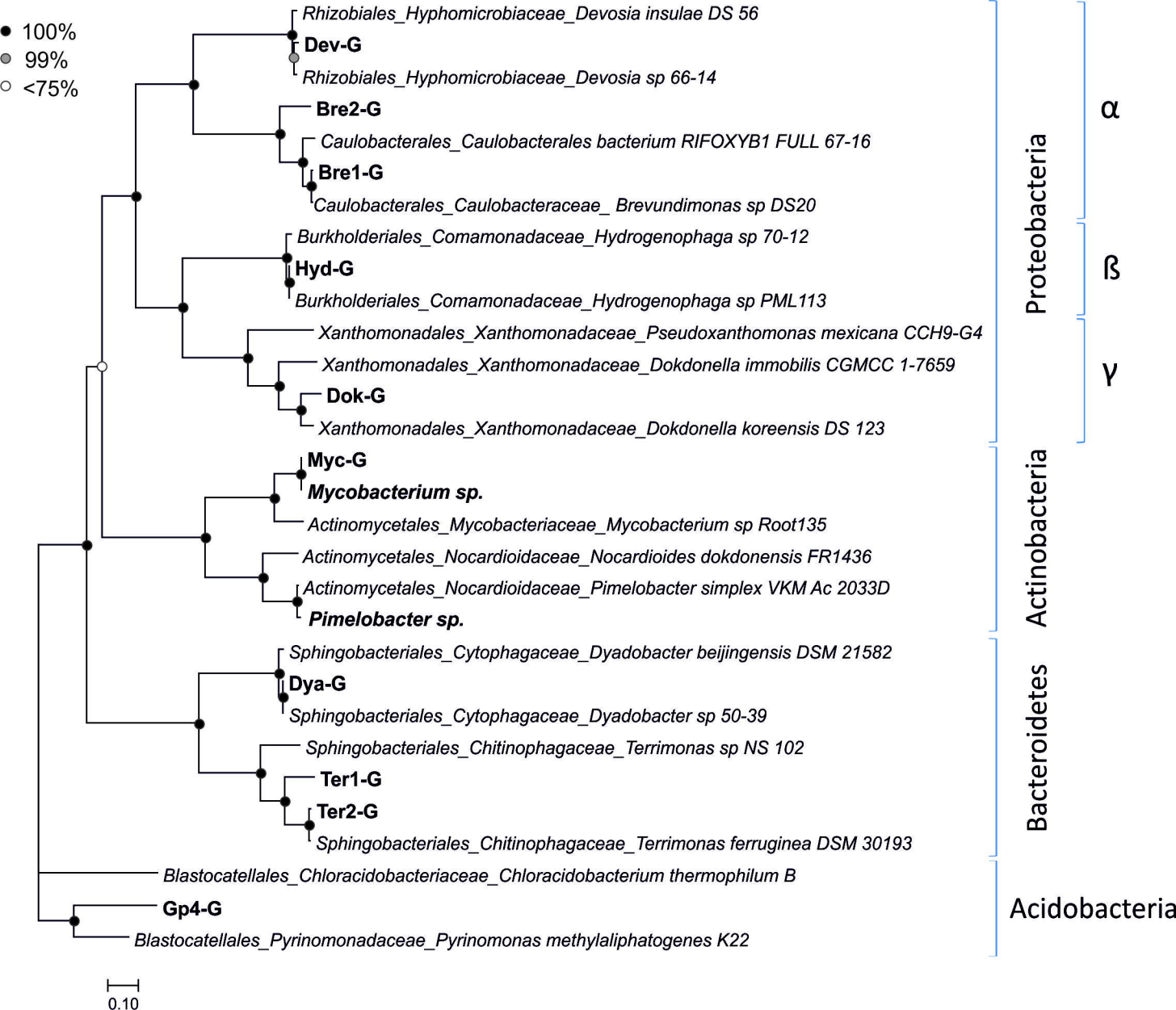
Figure 3: Phylogenetic characterization of the high-quality draft and complete genomes of *B. braunii* accompanying bacterial community. Maximum-likelihood-based phylogenetic tree built out of a core of 53 conserved marker genes (Supplementary data 2) per analyzed genome (bold). Black, grey and white circles represent the nodes with calculated bootstrap values of 100, 99 and <75%, respectively (1,000 replicates).

Figure 4: Alga-bacteria interactions within the *Botryocccus braunii* consortia. Summarizing view on the observed generic potential encoded in the high-quality draft and complete genomes of the associates (Table 1, Figure S5), and the resultant interactions between alga and bacteria. The size of the illustrated bacteria indicates the relative abundance (averaged) of the respective species within the consortia. The species marked with stripes, such as Dok-G and Gp4-G, indicate that these occur only within *B. braunii* race A or B, respectively. The incomplete B-vitamin *de novo* biosynthesis pathways are shown by the red capital letter B with the number of the subsequent vitamin. The arrows indicate the possible external supply of B-vitamins within the consortia. The apparent ability of the particular associates to store nitrogen and phosphorous as polymeric reserves is indicated by N in the green square (cyanophycin granule polypeptides) and P in the yellow circle (PolyP, Polyphosphate granules), respectively. The estimated potential for degradation of hydrocarbons and carbohydrates of the bacterial associates is illustrated based on the identified enzyme number (Table 1) in conjunction with the relative abundance of the respective species via yellow and striped pacmans, respectively. The genetic predisposition to the synthesis of bacteriocins and antibiotics encoded by the secondary metabolite biosynthesis gene clusters, is indicated by the blue filled or empty triangles, respectively. *M*.Bb-A, *Mycobacterium* sp. Bb-A; *P*.Bb-B, *Pimelobacter* sp. Bb-B; HC, Hydrocarbons; CH, carbohydrates; B1, Thiamin; B2, Riboflavin; B5, Pantothenate; B7, Biotin; B12, Cobalamin.

Table 1: Selected genetic features of the *B. braunii* bacterial associates. The the high-quality draft and complete bacterial genomes were manually inspected using a E-value cutoff of $1x10^{-10}$ for specific gene groups (see Supplementary data 4). Carbohydrate-active enzymes (CAZymes) were profiled using dbCAN2 [46] metaserver (Supplementary data 5). The values in parentheses represent the results acquired with two tools. The key enzymes involved in B-vitamin *de novo* biosynthesis were searched in the EMGB database (https://emgb.cebitec.uni-bielefeld.de/Bbraunii-bacterial-consortium/) and manually verified via BLASTp (Supplementary data 6). The analysis of the secondary metabolite biosynthesis gene clusters within the high-quality draft and complete genomes was accomplished via antiSMASH [47] 4.0 web server (Supplementary data 7). *M*. Bb-A, *Mycobacterium* sp. Bb-A; *P*.Bb-B, *Pimelobacter* sp. Bb-B; RNR, Ribonucleotide reductase; GH, Glycoside hydrolases; GT, Glycosyltransferases.


a
B. braunii
axenic control
b
B. braunii xenic cultures with
Mycobacterium sp.
Pimelobacter sp.
10 µm
biomass (g L-1)
time (d)
hydrocarbon content (g L-1)
Log10 bacterial cells mL-1
C25H48
C27H52
C29H56
C31H60
C33H64


relative abundance of 16S rDNA gene (%)
100
90
80
70
60
50
40
30
20
10
0
T1
T2
T1
T2
Race A
Race B
G_Reyranella
G_Lacibacterium
G_Sphingopyxis
G_Sphingopyxis
G_Porphyrobacter
G_Sphingomonas
F_Bradyrhizobiaceae
G_Bradyrhizobium
G_Bosea
G_Methylobacterium
G_Brevundimonas
G_Caulobacter
G_Devosia
G_Devosia
G_Mesorhizobium
G_Aminobacter
G_Shinella
G_Methylophilus
G_Achromobacter
G_Hydrogenophaga
G_Variovorax
G_Acinetobacter
G_Nevskia
G_Pseudoxanthomonas
G_Dokdonella
G_Leifsonia
G_Mycobacterium
G_Pimelobacter
G_Terrimonas
G_Terrimonas
G_Dyadobacter
C_Acidobacteria_Gp4
G_Planctomyces
0.05
100%
75-100%
<75%
α
ß
γ
Proteobacteria
Actinobacteria
Bacteroidetes
Acidobacteria
Planctomycetes


100%
99%
<75%
Rhizobiales_Hyphomicrobiaceae_Devosia insulae DS 56
Dev-G
Rhizobiales_Hyphomicrobiaceae_Devosia sp 66-14
Bre2-G
Caulobacterales_Caulobacterales bacterium RIFOXYB1 FULL 67-16
Bre1-G
Caulobacterales_Caulobacteraceae_ Brevundimonas sp DS20
Burkholderiales_Comamonadaceae_Hydrogenophaga sp 70-12
Hyd-G
Burkholderiales_Comamonadaceae_Hydrogenophaga sp PML113
Xanthomonadales_Xanthomonadaceae_Pseudoxanthomonas mexicana CCH9-G4
Xanthomonadales_Xanthomonadaceae_Dokdonella immobilis CGMCC 1-7659
Dok-G
Xanthomonadales_Xanthomonadaceae_Dokdonella koreensis DS 123
Myc-G
Mycobacterium sp.
Actinomycetales_Mycobacteriaceae_Mycobacterium sp Root135
Actinomycetales_Nocardioidaceae_Nocardioides dokdonensis FR1436
Actinomycetales_Nocardioidaceae_Pimelobacter simplex VKM Ac 2033D
Pimelobacter sp.
Sphingobacteriales_Cytophagaceae_Dyadobacter beijingensis DSM 21582
Dya-G
Sphingobacteriales_Cytophagaceae_Dyadobacter sp 50-39
Sphingobacteriales_Chitinophagaceae_Terrimonas sp NS 102
Ter1-G
Ter2-G
Sphingobacteriales_Chitinophagaceae_Terrimonas ferruginea DSM 30193
Blastocatellales_Chloracidobacteriaceae_Chloracidobacterium thermophilum B
Gp4-G
Blastocatellales_Pyrinomonadaceae_Pyrinomonas methylaliphatogenes K22
Proteobacteria
α
ß
γ
Actinobacteria
Bacteroidetes
Acidobacteria
0.10

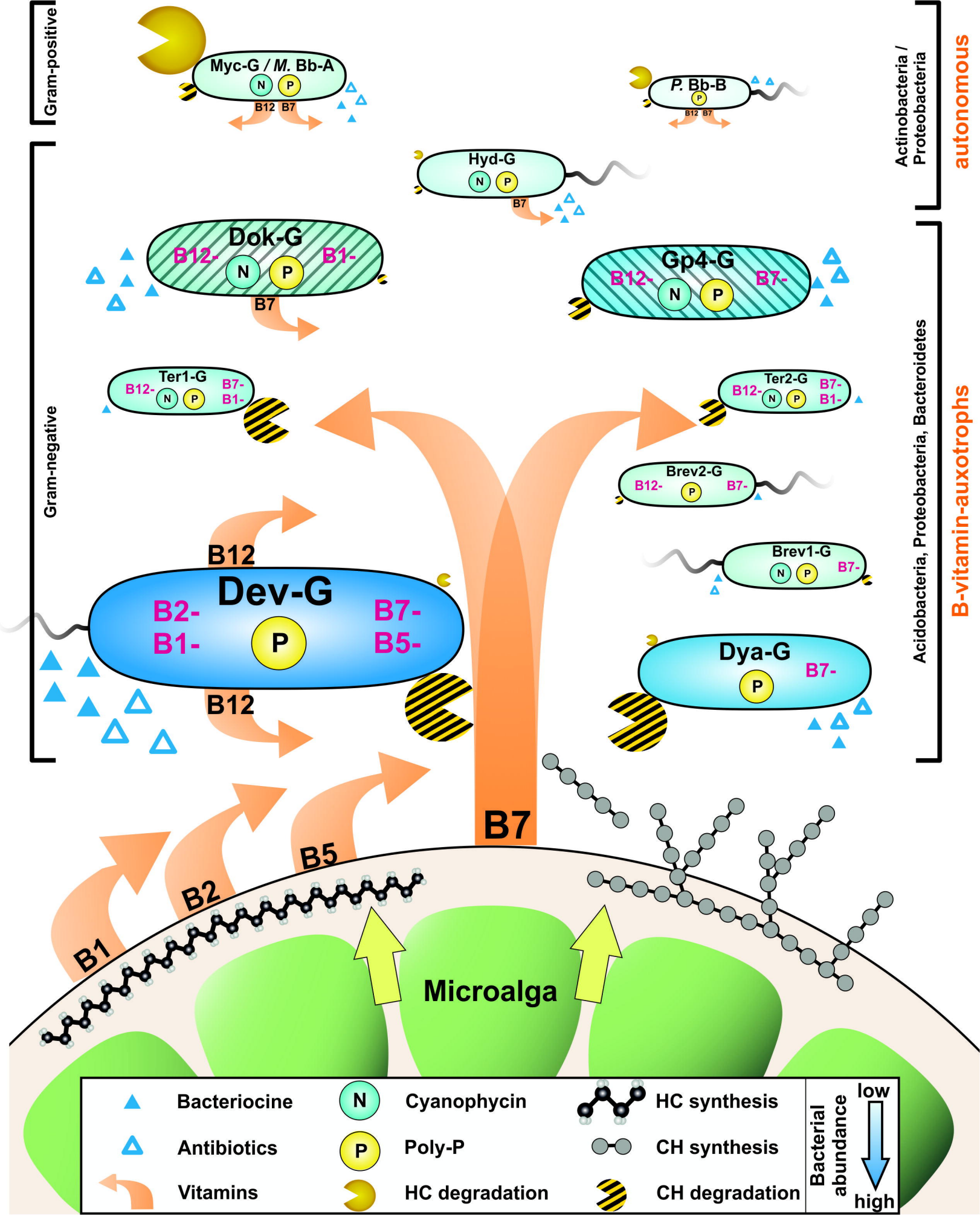

Gram-positive
Myc-G / M. Bb-A
N
P
B12
B7
P. Bb-B
Actinobacteria / Proteobacteria
autonomous
Hyd-G
Dok-G
B12-
B1-
Gp4-G
B7-
Gram-negative
Ter1-G
Ter2-G
Brev2-G
Brev1-G
Dev-G
B2-
B1-
B5-
Dya-G
Acidobacteria, Proteobacteria, Bacteroidetes
B-vitamin-auxotrophs
B1
B2
B5
Microalga
Bacteriocine
Antibiotics
Vitamins
Cyanophycin
Poly-P
HC degradation
HC synthesis
CH synthesis
CH degradation
Bacterial abundance
low
high

Table 1: Selected genetic features of *B. braunii* bacterial associates.

	Dev-G	Dok-G	Gp4-G	Dya-G	Ter1-G	Ter2-G	Bre1-G	Bre2-G	Hyd-G	Myc-G	*M*.Bb-A	*P*.Bb-B
Enzymes oxidizing long-chain hydrocarbons												
AlkB	-	-	-	-	-	-	-	-	-	3	3	2
AlmA	2	-	-	1	-	-	-	-	1	17	17	11
LadA	3	-	-	-	-	-	-	-	-	9	11	8
CYP153	1	-	-	2	-	-	1	-	1	40	40	12
CAZymes (Enzymes involved in carbohydrate metabolism)												
total number	164 (98)	109 (59)	111 (50)	395 (264)	240 (170)	195 (125)	86 (61)	68 (45)	96 (47)	183 (127)	182 (127)	131 (79)
GHs class	79 (56)	19 (10)	40 (26)	183 (146)	138 (117)	91 (73)	32 (27)	26 (21)	24 (14)	51 (43)	50 (42)	36 (28)
GTs class	47 (23)	55 (44)	39 (14)	103 (77)	46 (30)	52 (31)	27 (22)	20 (17)	47 (24)	60 (47)	60 (47)	53 (38)
% of total gene content	3.1 (1.9)	3.0 (1.7)	2.7 (1.2)	5.9 (3.9)	5.2 (3.7)	4.7 (3.0)	2.7 (1.9)	2.2 (1.4)	2.0 (1.0)	3.2 (2.2)	3.2 (2.2)	2.3 (1.4)
Methionine synthases												
MetH [a]	√	√	√	√	√	√	√	√	√	√	√	√
MetE	√			√			√		√	√	√	
Ribonucleoside diphosphate reductase												
RNR class I [b]	√	√	√	√	√	√	√	√	√	√	√	√
RNR class II [a]	√	√	√	√	√	√	√	√	√			√
Nutrient storage												
cyanophycin synthase		√	√		√	√	√		√	√	√	
cyanophycinase		√	√		√	√						
polyphosphate kinase	√	√	√	√	√	√	√	√	√	√	√	√
exopolyphosphotase	√	√		√		√	√	√	√	√	√	√
Vitamin *de novo* synthesis pathways												
Cobalamin (vitamin B_{12})	√									√	√	√
Vitamin B_6	√	√	√	√	√	√	√	√	√	√	√	√
Riboflavin (vitamin B_2)		√	√	√	√	√	√	√	√	√	√	√
Biotin (vitamin B_7)		√							√	√	√	√
Thiamin (vitamin B_1)			√	√			√	√	√	√	√	√
Pantothenate (vitamin B_5)		√	√	√	√	√	√	√	√	√	√	√
Nicotinate (vitamin B_3)	√		√	√	√	√	√	√	√	√	√	√
Secondary metabolite gene cluster with putative												
herbicidal activity	√									√	√	
antiviral activity												√
antifungal activity									√	√	√	√
insecticidal activity									√			
antibacterial activity	√	√	√	√			√		√	√	√	√
against $Gram^+$-bacteria									√	√	√	√
bacteriocins	√	√	√	√	√	√	√	√	√	√	√	
total cluster number	37	34	31	60	22	23	18	15	42	91	93	72

[a]Vitamin B_{12} (Cobalamin) dependent enzyme
[b]*Mycobacterium* sp. Bb-A (Myc-G and *M*.Bb-A) encodes genes for RNR class Ib and all other genomes for RNR class Ia

The high-quality draft and complete bacterial genomes were manually inspected using a E-value cutoff of $1x10^{-10}$ for specific gene groups (see Supplementary data 4). Carbohydrate-active enzymes (CAZymes) were profiled using dbCAN2 [46] metaserver (Supplementary data 5). The values in parentheses represent the results acquired with two tools. The key enzymes involved in

B-vitamin *de novo* biosynthesis were searched in the EMGB database (https://emgb.cebitec.uni-bielefeld.de/Bbraunii-bacterial-consortium/) and manually verified via BLASTp (Supplementary data 6). The analysis of the secondary metabolite biosynthesis gene clusters within the high-quality draft and complete genomes was accomplished via antiSMASH [47] 4.0 web server (Supplementary data 7). *M*. Bb-A, *Mycobacterium* sp. Bb-A; *P*.Bb-B, *Pimelobacter* sp. Bb-B; RNR, Ribonucleotide reductase; GH, Glycoside hydrolases; GT, Glycosyltransferases.

References

46. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, et al. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 2018; 1–7.

47. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, et al. antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* 2017; **45**: W36–W41.