MOLECULAR ECOLOGY

Molecular Ecology (2014) **23**, 4498–4510 doi: 10.1111/mec.12885

The *Sphagnum* microbiome supports bog ecosystem functioning under extreme conditions

ANASTASIA BRAGINA,*[1] LISA OBERAUNER-WAPPIS,*†[1] CHRISTIN ZACHOW,* BETTINA HALWACHS,†‡ GERHARD G. THALLINGER,†‡ HENRY MÜLLER* and GABRIELE BERG*

**Institute of Environmental Biotechnology, Graz University of Technology, Petersgasse 12, 8010 Graz, Austria, †Austrian Centre of Industrial Biotechnology (ACIB GmbH), Petersgasse 14, 8010 Graz, Austria, ‡Institute for Genomics and Bioinformatics, Graz University of Technology, Petersgasse 14, 8010 Graz, Austria*

Abstract

***Sphagnum*-dominated bogs represent a unique yet widely distributed type of terrestrial ecosystem and strongly contribute to global biosphere functioning. *Sphagnum* is colonized by highly diverse microbial communities, but less is known about their function. We identified a high functional diversity within the *Sphagnum* microbiome applying an Illumina-based metagenomic approach followed by *de novo* assembly and MG-RAST annotation. An interenvironmental comparison revealed that the *Sphagnum* microbiome harbours specific genetic features that distinguish it significantly from microbiomes of higher plants and peat soils. The differential traits especially support ecosystem functioning by a symbiotic lifestyle under poikilohydric and ombrotrophic conditions. To realise a plasticity–stability balance, we found abundant subsystems responsible to cope with oxidative and drought stresses, to exchange (mobile) genetic elements, and genes that encode for resistance to detrimental environmental factors, repair and self-controlling mechanisms. Multiple microbe–microbe and plant–microbe interactions were also found to play a crucial role as indicated by diverse genes necessary for biofilm formation, interaction via quorum sensing and nutrient exchange. A high proportion of genes involved in nitrogen cycle and recycling of organic material supported the role of bacteria for nutrient supply. 16S rDNA analysis indicated a higher structural diversity than that which had been previously detected using PCR-dependent techniques. Altogether, the diverse *Sphagnum* microbiome has the ability to support the life of the host plant and the entire ecosystem under changing environmental conditions. Beyond this, the moss microbiome presents a promising bio-resource for environmental biotechnology – with respect to novel enzymes or stress-protecting bacteria.**

Keywords: bog ecosystem, FISH–CLSM, illumina-based metagenomics, plant microbiome, *Sphagnum* moss

Received 29 January 2014; revision received 7 August 2014; accepted 8 August 2014

Introduction

Bog ecosystems belong to the oldest vegetation forms on earth; they have a high value for biodiversity conservation, are a reservoir for freshwater and play an extraordinary role in carbon sequestration to benefit both human welfare and our world climate (Succow & Joosten 2001; Raghoebarsing *et al.* 2005; Dise 2009). However, these long-existing ecosystems are extremely sensitive to changing abiotic factors connected with climate change (Strack 2008; Dise 2009). For example, degraded peatlands release their stored carbon in the form of greenhouse gases, and drainage of peat soils results in CO_2 and N_2O global emissions of 2–3 Gt CO_2-eq per year (Joosten & Couwenberg 2009). Mosses of the genus *Sphagnum* are among the most abundant and cosmopolitan in

Correspondence: Anastasia Bragina, Fax: +43 316 873 8819; E-mail: anastasia.krikovtseva@tugraz.at

[1]Equal contributors.

bog vegetation in the Northern hemisphere and greatly contribute to both global carbon turnover and global climate (Jassey *et al.* 2011). The ecological significance of bogs is directly related to the physical, morphological and chemical characteristics of *Sphagnum* peat mosses, which belong to the poikilohydric plants that undergo repetitive desiccation and oxidative stress (Daniels & Eddy 1985). Moreover, *Sphagnum* mosses are able to change their environments: living Sphagna have extraordinarily high cation exchange capacity and therefore acidify their environment by exchanging tissue-bound protons for basic cations in surrounding water (Soudzilovskaia *et al.* 2010).

Recently, the plant microbiome was identified as one of the key determinants of plant health and productivity (rev. in Berg *et al.* 2013b; Bulgarelli *et al.* 2013; Philippot *et al.* 2013). *Sphagnum* mosses are interesting models to study plant, especially phyllosphere microbiomes, and an enormous associated bacterial diversity was already detected (Opelt *et al.* 2007a; ¸ Bragina *et al.* 2012b, 2013), including different methanotrophic communities (Raghoebarsing *et al.* 2005; Larmola *et al.* 2010; Kip *et al.* 2011; Putkinen *et al.* 2012). As these phylogenetically old plants have no roots, the leaf-associated bacteria fulfil important functions such as nutrient supply and pathogen defence for moss growth and health (Opelt *et al.* 2007b). Host specificity of moss-associated microbiomes was detected independent of geographic region at both structural and functional levels (Bragina *et al.* 2012a, 2013). Additionally, the degree of host specificity varied between distant and closely related moss species and corresponded to spectra of secondary metabolites produced by plants (Bragina *et al.* 2012b). Moreover, environmental factors such as acidity and nutrient richness were defined as the main ecological drivers for microbial diversity, and plant specificity of functional bacterial groups is determined by their role within the ecosystem (Bragina *et al.* 2013). A core microbiome of *Sphagnum* mosses not only contained mostly potential beneficials, but was also shared between the moss generations and transferred within the spore capsules that emphasize the importance of the microbiome for mosses as the oldest phylogenetic land plants on earth (Bragina *et al.* 2012a). Although this high diversity of the *Sphagnum* microbiome is now well studied, less is known about its functional diversity. Omics technologies significantly contribute to a functional understanding of microbial ecosystems (Gilbert *et al.* 2011), but very little is known for plants (Knief *et al.* 2012; Sessitsch *et al.* 2012).

The objective of this study was to unravel the functional diversity associated with *Sphagnum* mosses. We applied an Illumina-based metagenomic approach, and through *de novo* assembly and MG-RAST annotation, we revealed specific biochemical pathways and adaptive strategies within the moss metagenome (Meyer *et al.* 2008). We analysed the *Sphagnum* microbiome with a special focus on plasticity, stability and interactions and performed a comparison with other published metagenomes of plants, peat soils, as well as aquatic systems to discover unique features and potential differences.

Methods

Sampling procedure

For this metagenomic study, we selected peat moss *Sphagnum magellanicum* BRID. (section *Sphagnum*), a typical and widespread vegetation component of the acidic peat bogs (Daniels & Eddy 1985) illustrated in Fig. S1 (Supporting information). Gametophyte samples of *S. magellanicum* were collected from the Alpine bog Pirker Waldhochmoor (N46°37′38.66″ E14°26′5.66″) in Austria in December 2011. Four independent replicates consisting of the living moss plants were collected from the sampling points situated at a distance of at least 150 m. The collected samples were placed into sterile plastic bags, cooled (4–8 °C) and transported to the laboratory.

Total community DNA isolation

To isolate the total community DNA of the *S. magellanicum* microbiome, 200 g of each sample was transferred into Stomacher bags (20 g/bag) and supplied with 0.85% NaCl solution (50 mL/bag). The diluted samples were shaken in a Stomacher laboratory blender (BagMixer, Interscience, St. Nom, France) for 3 min. To remove the plant debris, the suspension was subsequently strained through two sieves (500 and 63 μm) and the resulting liquid was centrifuged in 50-ml tubes at low speed (5000 ***g***, 5 min, 4 °C). The supernatant was discarded, and the pellets were resuspended in 1.5 mL 0.85% NaCl. After centrifugation at high speed (10 000 ***g***, 20 min, 4 °C), the obtained pellets were stored at −70 °C. The total community DNA was extracted using the FastDNA Spin Kit for Soil (BIO 101, Carlsbad, USA) according to the manufacturer's protocol and quantified using spectrophotometer NanoDrop 2000c (Thermo Scientific, Waltham, MA, USA). Finally, DNA aliquots from all samples were pooled together in an equimolar ratio and sent to Eurofins MWG Operon (Ebersberg, Germany, http://www.eurofinsgenomics.eu/) for Illumina sequencing.

Sequencing and bioinformatic analysis

The sequencing was performed with an Illumina HiSeq 2000 system (2 × 100 bp). Prior to sequencing, the total

community DNA was split into two aliquots. The first aliquot was sequenced untreated, and the second aliquot underwent a normalization treatment that allowed removal of the most dominant sequence patterns for deeper ecological analysis. The normalization was achieved through one cycle of denaturation and reassociation of the DNA, followed by separation of the reassociated dsDNAs from the remaining ssDNAs (normalized DNA) by passing the mixture over a hydroxylapatite column. After hydroxylapatite chromatography, the ssDNAs were sequenced according to the Eurofins MWG Operon protocol. The generated paired-end reads of the normalized metagenome were *de novo* assembled using the CLC GENOMIC WORKBENCH version 4.5.1 (CLC Bio, Aarhus, Denmark) and default settings. The complete metagenome, which resulted from untreated DNA sequencing, was used for abundance-based analyses, while normalized metagenome was used to study ecologically-relevant functional patterns.

The functional composition of the *S. magellanicum* microbiome was analysed using the Metagenomic RAST (MG-RAST) server (Meyer *et al.* 2008). For this purpose, both the complete and the normalized metagenomes were uploaded to the server and initially processed with default parameters: filtered for artificial replicate sequences (Gomez-Alvarez *et al.* 2009), low quality (Cox *et al.* 2010) and short sequences, and sequences containing ambiguous bases. The annotation was done using hierarchical classification with default parameters: SEED subsystems as an annotation source, a maximum *e*-value of 10^{-5}, a minimum identity of 60% and a minimum alignment length of 15 measured in aa for protein and bp for RNA databases. Within the annotated metagenomes, each single subsystem represented a group of sequences that encode for a specific biological process or structural complex as defined by Overbeek *et al.* (2005). For the normalized metagenome, sequences from the single subsystems were aligned against a nonredundant protein sequences (nr) database using BLASTX algorithm to check their affiliation. Distribution of the functional subsystems within the normalized metagenome was visualized using Krona plot (Ondov *et al.* 2011). Enzymes involved in nitrogen metabolism from the complete and normalized metagenomes were visualized using KEEG mapper tool of the MG-RAST server with default parameters.

The interenvironmental comparison of the complete *S. magellanicum* metagenome with publicly available metagenomes was performed using the principal coordinate analysis (PCoA) tool of the MG-RAST server. Relevant publicly available metagenomes obtained from peat soils, freshwater habitats, plant tissues and human bodies are summarized in Table S1 (Supporting information). PCoA analysis was performed for the metagenomic data sets that were annotated using hierarchical classification with default parameters. For each data set, sequence counts were normalized and scaled according to the algorithm, which is specified at the MG-RAST server (http://blog.metagenomics.anl.gov/howto/mg-rast-analysis-tools/). The distance matrix for PCoA analysis was calculated using Bray–Curtis as a distance metric (Bray & Curtis 1957). The interenvironmental comparison of the metagenomes was expanded by constructing a heatmap of the complete *S. magellanicum*, higher plant, and peat soil metagenomes and their functional subsystems using the MG-RAST heatmap tool. The selected metagenomes (Table S2, Supporting information) were grouped using complete linkage clustering with Bray–Curtis distance. For these metagenomes, Kolmogorov–Smirnov test (Massey 1951) was applied on the raw abundances to test probability distributions of each subsystem (Table S2, Supporting information). Scale normalization factors were calculated to scale the raw library sizes prior to significance analysis. To make the count data ready for linear modelling, raw counts were transformed using the voom function (Law *et al.* 2014). The probability distribution of each group was visualized before and after data transformation using density plots (Fig. S2, Supporting information). Changes of the subsystems included in the heatmap between the different groups were assessed by statistical analysis using the linear modelling approach implemented by the R Bioconductor package LIMMA (version 3.16.8) (Smyth 2004). Significance analysis within LIMMA was performed by the moderated *t*-statistic, which was computed for each probe and each contrast. To account for multiple comparisons, *P*-values were adjusted by the method described by Benjamini & Hochberg (1995). Adjusted *P*-values of <0.05 were considered as statistically significant.

The taxonomic structure of the *Sphagnum*-associated bacterial community was determined on the basis of 16S rRNA genes derived from total metagenomic quality reads of the complete metagenome. Prior to taxonomic assignment, reads that comprised exclusively partial 16S rRNA genes were extracted after alignment to references of the whole 16S rRNA gene by a homology-based approach using BLASTN algorithm. Only reads which consisted of 16S rRNA gene sequences covering a length between 80 and 100 bp were retained and processed using QIIME pipeline with default parameters (release 1.7.0, Caporaso *et al.* 2010). In detail, sequence clustering was performed at 97% similarity using UCLUST algorithm and a predesigned taxonomy map (Edgar 2010) implemented in the QIIME workflow *pick_open_reference_otus.py*. Taxonomic assignment of representative sequences was done using RDP naïve Bayesian rRNA classifier (Wang *et al.* 2007) based on the reference database Greengenes release 13_5

(DeSantis *et al.* 2006). In addition, taxonomic hits distribution was deduced from the complete metagenome for both the sequences with predicted protein-coding regions and ribosomal rRNA genes using all reference databases available at the MG-RAST server.

Fluorescent in situ hybridization and confocal laser scanning microscopy

Single gametophytes of *S. magellanicum* were fixed with 4% paraformaldehyde/phosphate-buffered salt (3:1, v/v) and stained by in-tube FISH (Grube *et al.* 2009). The samples were consequently hybridized with rRNA-targeting probes (genXpress, Wiener Neudorf, Austria) specific for Alphaproteobacteria (ALF968) (Loy *et al.* 2007) and with a set of universal bacterial probes (EUB338/EUB338II/EUB338III) (Amann *et al.* 1990; Daims *et al.* 1999). Hybridization was carried out at 41 °C using hybridization buffer with 35% and 15% formamide, respectively. Negative control was hybridized with nontarget NON-EUB probe (Amann *et al.* 1990) at the same stringency conditions applied for the positive FISH probes. Confocal laser scanning microscopy (CLSM) was performed with a Leica TCS SPE confocal microscope (Leica Microsystems, Mannheim, Germany) as previously described (Bragina *et al.* 2012a) followed by volume rendering of confocal stacks and three-dimensional modelling using the software IMARIS 7.3 (Bitplane, Zurich, Switzerland).

Results

The Sphagnum metagenomic data set

Illumina HiSeq 2x100 paired-end sequencing resulted in 172 590 841 reads (41.8 Gbps in total) and 141 411 216 reads (32.0 Gbps) from the untreated and the normalized metagenomic DNA of *Sphagnum* moss, respectively (Table S3, Supporting information). *De novo* assembly of the normalized metagenome yielded 1 115 029 scaffolded contigs totalling 558 360 453 bps with an average length of 501 bps. For both metagenomes, all sequences passed the quality control (QC) pipeline during MG−RAST statistical analysis. Of the 172.6 M sequences, 153.8 M (89.1%) produced a total of 151.7 M predicted protein-coding regions. The assembled data set derived from the normalized library contained 1 115 029 contigs, of which 1 075 645 (96.5%) were translated into 1 430 118 protein fragments that encoded 1 411 717 predicted protein-coding regions. Based on their best *e*-value scores (Fig. S3, Supporting information), SEED subsystems were selected as an annotation source for functional analysis of the moss metagenome (Overbeek *et al.* 2005). The subsystems approach allowed us to precisely assign metagenomic sequences to the groups with known or hypothetical biological functions with the exception of clustering-based and miscellaneous categories.

Within the complete metagenome, the most dominant subsystems represented carbohydrate and protein metabolism (amino acids and protein metabolism) as the most important biochemical processes for all forms of life (Fig. 1). Less dominant subsystems contained metagenomic sequences that encode pathways for biological monomers (nucleoside and nucleotides), more complex biochemical compounds (cofactors, vitamins, prosthetic groups, pigments; aromatic compounds; fatty acids, lipids and isoprenoids) and structural elements such as the cell wall and capsule. Subsystems corresponding to environmental information processing such as membrane transport, stress responses, virulence, disease and defence followed. Among the less dominant subsystems, several subsystems were crucial for processing genetic information in- and outside the cells (DNA and RNA metabolism; phages, prophages, transposable elements, plasmids). Subsystems responsible for single chemical element cycling (N, S, P, K, Fe) comprised a minor portion of all subsystems with the highest relative abundance for sulphur metabolism. Genetic features that characterize cellular processes were irregularly distributed within the annotated metagenome and found in the less dominant subsystems of cell regulation and signalling, cell division and cycle, in the minor subsystems of motility and chemotaxis, and dormancy and sporulation.

Taxonomic diversity and spatial structure of the S. magellanicum microbiome

A total of 7318 reads containing partial 16S rRNA genes were obtained from metagenomic sequences to characterize the structure of bacterial communities (Fig. 2). At phylum level, the majority of reads were assigned to *Proteobacteria* (65.8%) followed by *Acidobacteria* (11.4%), *Actinobacteria* (5.6%), *Bacteroidetes* (4.2%) and *Verrucomicrobia* (2.0%). The remaining portion of the classified reads was distributed among 13 bacterial phyla which notably contained *Planctomycetes*. At class level, Alphaproteobacteria and Betaproteobacteria were the most abundant taxa among the phylum *Proteobacteria*, while Gammaproteobacteria represented a subdominant taxon. The classes Acidobacteria, Actinobacteria and Sphingobacteria dominated the phyla *Acidobacteria*, *Actinobacteria* and *Bacteroidetes*, respectively.

The taxonomic hits distribution of metagenomic sequences with predicted protein-coding regions and ribosomal rRNA genes (Fig. S4, Supporting information) revealed highly similar dominant patterns to the 16S rRNA genes data. Within the reads assigned to domain

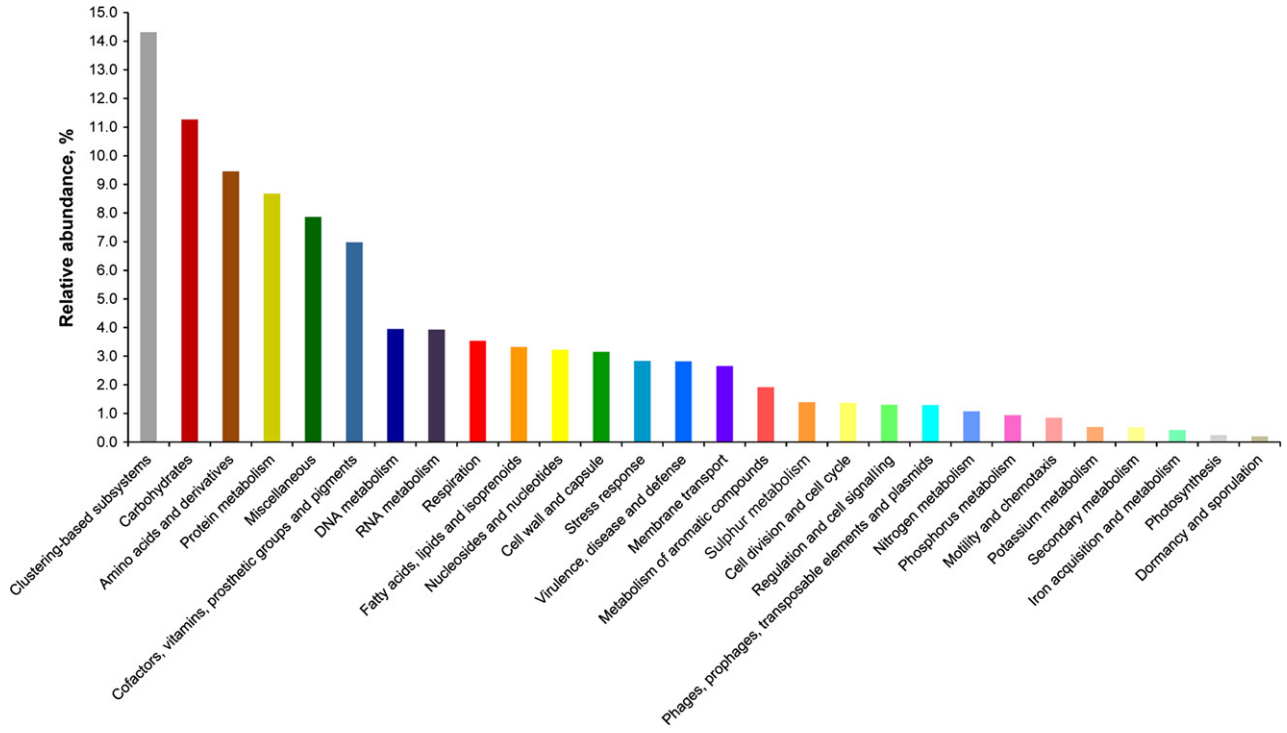

Fig. 1 Functional composition of the complete *S. magellanicum* metagenome. Distribution of 35 702 611 metagenomic sequences annotated using functional subsystems of SEED database with max. *e*-value cut-off of 10^{-5}, min. identity cut-off of 60% and min. alignment length of 15 aa (protein annotations) or bp (rRNA annotations). All functional groups are shown at the subsystems level 1.

Bacteria (61 528 765 sequences), dominant portion was composed of *Proteobacteria* (61.9%), *Acidobacteria* (13.1%), *Actinobacteria* (8.3%), *Bacteroidetes* (4.2%) and *Verrucomicrobia* (3.0%). The minor fraction of functional bacterial reads was distributed among 16 phyla that were not covered by partial 16S rRNA genes. Consequently, FISH and CLSM techniques were used to visualize Alphaproteobacteria – one of the most abundant bacterial patterns in *S. magellanicum* gametophytes. In general, *Sphagnum* mosses are characterized by unique morphology that distinguishes them from other bryophytes (Daniels & Eddy 1985). In particular, *Sphagnum* leaves are composed of a single-layer cell net of photosynthetic chlorocytes and dead hyalocytes, which contain large pores. By applying FISH–CLSM approach, we demonstrated that hyalocytes of moss leaves serve as a main colonization compartment for bacteria (Fig. S5, Supporting information). Alphaproteobacteria represented up to 31.9% of the detected bacterial cells that coincided with its relative abundance in metagenomic data sets (30.2%).

Unique plant–microbe biocoenosis assessed using comparative metagenomics

To study the specificity of the *Sphagnum* microbiome, the complete *S. magellanicum* metagenome was compared with publicly available metagenomes accessible through MG-RAST. We selected metagenomes obtained from peat soils, freshwater habitats and plant tissues most relevant to the moss metagenome and metagenomes obtained from human bodies as outgroups to all tested environmental metagenomes (Table S1, Supporting information). PCoA analysis showed that the *Sphagnum* metagenome has a distinct position outside all examined groups (Fig. 3). On the PCoA plot, the closest group of metagenomes originated from higher plants, such as rice, clover, soybean and thale cress. The heatmap and statistical analyses revealed a high specificity for the functional traits that underlie the *Sphagnum*–microbe biocoenosis (Fig. S6, Table S4, Supporting information). Statistical analysis resulted in a significant difference ($P < 0.05$) for 106 functional groups that were differentially abundant between the *S. magellanicum* and higher plant metagenomes, of which 51 groups were significantly enriched in the moss metagenome. In contrast to the higher plant metagenomes, the *Sphagnum* metagenome contained significantly higher abundances of functional groups that are responsible for siderophore production and genetic plasticity of the microbiome (gene transfer agents, GTA). The *Sphagnum* metagenome was also enriched in subsystems involving genetic traits for interactions with other microbes and plant host

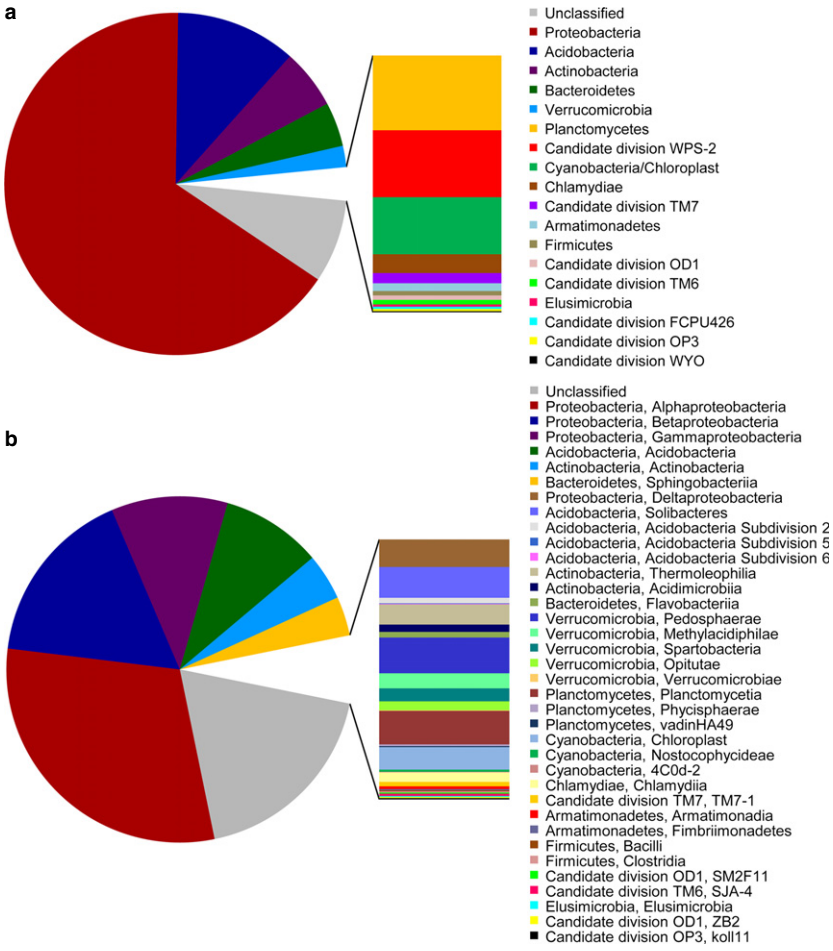

Fig. 2 Taxonomic composition of *S. magellanicum*-associated bacterial community. 16S rRNA gene sequences were retrieved from the complete metagenome and classified using RDP classifier with a confidence threshold of 80%. Pie charts represent relative abundance of bacterial taxa at phylum (a) and class (b) level. Taxa below 1% of relative abundance are shown as separate bar charts.

including various secretion systems. Comparison of moss and peat soils metagenomes revealed 37 differentially abundant functional groups, of which 20 groups were significantly enriched in the moss metagenome. The moss metagenome significantly differed from the peat soils metagenomes by higher abundances of functional groups that are responsible for respiration (reverse electron transport and sodium-ion-coupled energetics) and motility of bacteria (social motility and nonflagellar swimming of bacteria). Compared with both the plant and the peat soils metagenomes, the moss metagenome was significantly enriched in functional groups of stress response: desiccation and oxidative stress, and spore DNA protection. In conclusion, we demonstrated that the *Sphagnum* microbiome harbours highly specific genetic features that distinguish it from microbial communities of higher plants and peat soils.

Functional versatility of the moss metagenome

Functional subsystems were further studied in terms of plasticity, stability and interaction as main maintenance strategies of the *Sphagnum*–microbe biocoenosis (Table 1). For this purpose, we analysed the normalized and assembled metagenomic data set that comprised 657 466 sequences assigned to certain functional subsystems of SEED database (Fig. S7, Supporting information). Regarding plasticity traits, we detected highly abundant subsystems responsible for genetic exchange: (i) temperate bacteriophages (prophages) and their GTA analogues; (ii) plasmids likely involved in natural competence; and (iii) type IV pili and conjugative transport systems. Genetic attributes of microbiome stability were found in subsystems that encode for resistance to environmental factors, repair and self-controlling mechanisms. For instance, we identified a set of pathways that contribute to the oxidative stress response and DNA repair. These subsystems encode enzymatic responses of the cells and damage elimination caused by the oxidative stress. Notably, the highest diversity was observed among subsystems essential for bacterial interaction within the microbiome. In particular, quorum sensing was represented by autoinducer-2 (AI-2), acyl homoserine lactones (AHLs) and gamma-butyrolactones

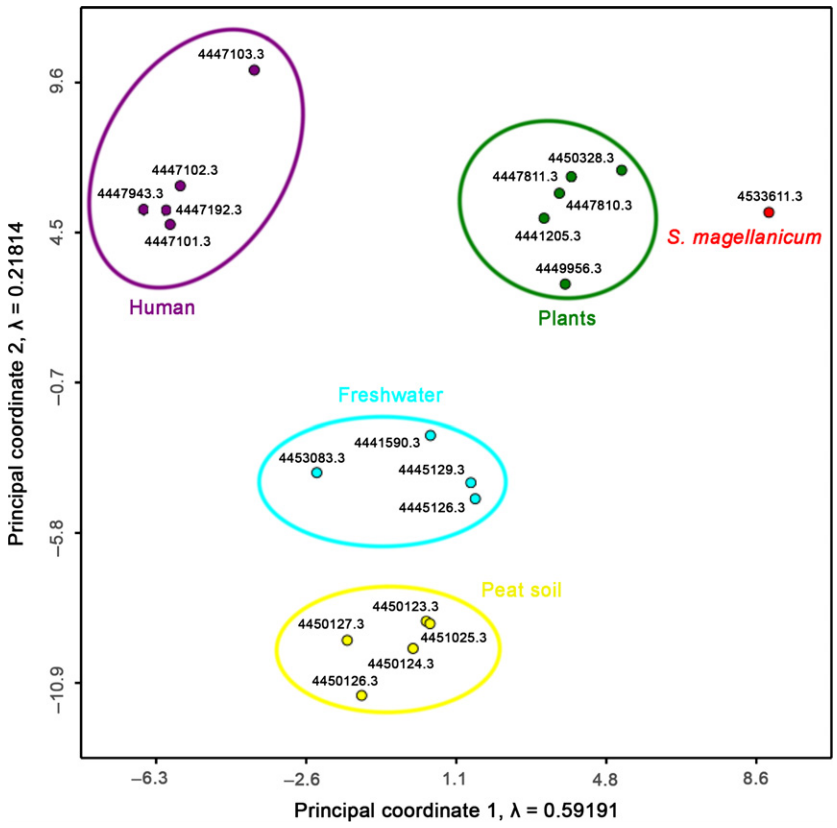

Fig. 3 Interenvironmental comparison of the complete *S. magellanicum* and publicly available metagenomes by principal coordinate analysis (PCoA). PCoA plot is based on the Bray–Curtis distance matrix of metagenomes that were assigned to functional subsystems of SEED database and normalized. Single metagenomes are shown by coloured dots with MG-RAST identical numbers (IDs) and grouped according to biome types (coloured ellipses). Eigenvalues (λ) correspond to variation explained by each principal coordinate, respectively.

signalling pathways. Other mechanisms coupled with biofilm formation were the production of surface adhesins and extracellular polysaccharides, motility, and chemotaxis. To restrict our search of bacterial interactions based on the food web, we focused on the nitrogen cycle as one of the bottlenecks in bog ecosystems. Interestingly, subsystems of nitrogen acquisition and conservation strongly prevailed over subsystems of nitrogen release and efflux from the ecosystem (Fig. S8, Supporting information). Moreover, bacterial protein degradation, which corresponds to organic nitrogen recycling (mineralization), was shown among highly abundant subsystems.

Discussion

We revealed a versatile genetic potential by analysing our metagenomic data set for the *Sphagnum* moss microbiome: a model for the first and phylogenetically oldest land plants. Moreover, we found a unique combination of the functional traits in terms of plasticity, stability, and microbe–microbe and plant–microbe interactions.

By employing an interenvironmental comparison, we demonstrated that the examined moss microbiome is distinct from microbial communities of higher plants and peat soils by its genetic content. The *Sphagnum* microbiome contained significantly higher abundances of genetic systems encoding for siderophores that are produced by antagonistic bacteria associated with *Sphagnum* mosses (Opelt & Berg 2004; S. Maier, A. Bragina, G. Berg, unpublished) and bacteria supporting plant nutrition under low concentrations of bio-available iron (Jin *et al.* 2010), which is characteristic for ombrotrophic peatlands. The moss microbiome differed from the peat soils microbiomes by increased abundance of several genetic systems responsible for bacterial respiration, which differ along the vertical oxygen gradient (Tveit *et al.* 2013), and bacterial motility that plays an important role for establishing of plant-associated microbiome (Delmotte *et al.* 2009). Moreover, the *Sphagnum* microbiome was significantly enriched in stress response systems, especially those coupling with oxidative stress, which was also shown as an essential functional trait of rice root endophytes in waterlogged paddy soils (Sessitsch *et al.* 2012).

These differences indicate the specific interactions established between *Sphagnum* mosses and their microbiome. Previous research proposed that the *Sphagnum* microbiome intimately cooperated with the host plants via nutrient supply and defence against pathogens (Raghoebarsing *et al.* 2005; Opelt *et al.* 2007b; Bragina *et al.* 2013), but Illumina sequencing of the moss

Table 1 Ecologically relevant functional subsystems of *S. magellanicum* metagenome

Category	Function	Level 2 subsystems	Sequence, abundances*	Details	References
Plasticity	Genetic exchange	Phages, prophages†	15 076	Temperate bacteriophages (prophages) responsible for horizontal gene transfer	Canchaya *et al.* (2003)
		Protein and nucleoprotein secretion system, type IV†	12 187	Plasmids carrying type IV secretion system genes, type IV pili and conjugative transport systems	
		Genetic transfer agents (GTA)†	2047	Phage-like elements in Bacteria	Lang *et al.* (2012)
Stability	Stress tolerance	Oxidative stress†	15 355	Glutathione-, mycothiol-, rubrerythrin-mediated reactions, etc.	
		Cold shock	1207		
	Repair	DNA repair	7967	Base excision and mismatch repair, nonhomologous end joining, homologous recombination and SOS-response systems	
	Resistance	Resistance to antibiotics and toxic compounds	18 512	Cobalt–zinc–cadmium resistance, multidrug resistance efflux pumps	
		CRISPRs	652	Clustered regularly interspace short palindromic repeats (CRISPRs) – resistance to exogenous genetic elements	Horvath & Barrangou (2010)
	Self-control	Programmed cell death and toxin–antitoxin systems	5045	Various toxin–antitoxin (programmed cell death) systems: PhD-Doc, YdcED, MazEF, etc.	Van Melderen (2010)
Interaction	Motility and chemotaxis	Flagellar motility in Prokaryota†	7788	Flagellar biosynthesis proteins and transcription initiation factors	
		Social motility and nonflagellar swimming in bacteria†	307	Rhamnolipids (biosurfactants) in *Pseudomonas*	D'aes *et al.* (2010)
	Quorum sensing, biofilm formation and signalling	Quorum sensing and biofilm formation†	7313	Biofilm adhesion biosynthesis, autoinducer-2 and acyl homoserine lactone biosynthesis and processing, symbiotic colonization and sigma-dependent biofilm formation, etc.	Nadell *et al.* (2009)
		Capsular and extracellular polysaccharides	6087	Biosynthesis of rhamnosylated glycans	Mäki & Renkonen (2004)
		Proteolytic pathway	3341	Regulatory intramembrane proteolysis in Bacteria‡	Wiegert 2010;
		Bacterial cytostatics, differentiation factors and antibiotics†	1726	γ-butyrolactones and other morphogens	Kato *et al.* (2007)
	Attachment	Adhesion†	2465	Adhesins from nonpathogenic bacteria‡	Danhorn & Fuqua (2007)
		Desiccation stress†	562	O-antigen capsule important for plants	Barak *et al.* (2007)

Table 1 *Continued*

Category	Function	Level 2 subsystems	Sequence, abundances*	Details	References
				colonization (Yih family proteins)	
	Nutrition (N-cycling)	Nitrogen metabolism, –	14 751	Dominant: nitrate and nitrite ammonification, nitrogen fixation and ammonium assimilation; Minor: dissimilatory nitrite reductase, allantoin utilization, nitric oxide synthase, cyanate hydrolysis, denitrification and nitrilase	Rydin & Jeglum (2006)
		Protein degradation†	9114	Proteolysis in bacteria, eukaryotic and bacterial proteasomes	Schimel & Bennett (2004)

*Sequence abundances correspond to the normalized metagenome that accounts for 657 455 assembled metagenomic sequences.
†Differentially abundant subsystems show statistically significant difference ($P < 0.05$) between metagenomes of *S. magellanicum*, higher plants and/or peat soils.
‡According to the BLASTX alignment.

metagenome obtained a much higher functional diversity than previously reported. To elucidate this profound diversity, we developed a framework in the form of plasticity–stability interaction that integrates genetic signatures of symbiosis (Gilbert *et al.* 2012) within the plant–microbe biocoenosis (Fig. 4). Specifically, the moss metagenome contained a relatively high number of mobile elements which were also found in the metagenomes of symbiotic bacterial consortia and considered to play an important role in the evolution of bacterial genomes for symbiosis with their hosts (Ochman & Moran 2001; Thomas *et al.* 2010). Furthermore, *Sphagnum* mosses belong to the poikilohydric plants that undergo repetitive desiccation and oxidative stress (Daniels & Eddy 1985; Scheibe & Beck 2011). Due to the high diversity and abundance of genes responsible for the oxidative stress response in the studied metagenome, we proposed that the bacterial capacity to tolerate oxidative stress may determine the effective and stable colonization of the *Sphagnum* mosses. In regard to interaction traits, vegetation in peatland ecosystems is strongly limited by nitrogen availability and therefore requires prokaryotic associates for nitrogen supply (Rydin & Jeglum 2006). Since Granhall & Hofsten (1976) observed nitrogen-fixing symbiotic *Cyanobacteria* in *Sphagnum* for the first time, diazotrophic communities of Sphagna have been characterized by a high taxonomic diversity and shown to transfer fixed nitrogen to the host plants (Bragina *et al.* 2012b, 2013; Berg *et al.* 2013a). In the current study, we observed and determined the entire nitrogen turnover of the moss microbiome. The pathways for bacterial nitrogen acquisition and conservation strongly prevailed over those for nitrogen release and efflux processes within the metagenome. These data are consistent with a recent study by Lin *et al.* (2014) that showed a high abundance and diversity of nitrogen fixation genes and suggested active degradation of organic nitrogen in the upper zone of an ombrotrophic peatland. Overall, we provided evidence that the *Sphagnum* microbiome carries essential genetic potential for sustainable functioning in association with the host plants and within the peatland ecosystem.

This metagenome study provided also new insights into the taxonomic diversity of the *Sphagnum*-associated microbiome. Our approach allowed for a deep analysis of the 16S rRNA gene diversity without PCR bias. Although the dominant bacterial taxa detected using Illumina sequencing were similar to those revealed by PCR-dependent approaches (Bragina *et al.* 2012a), their relative abundance considerably differed. As such, we observed a low number of *Planctomycetes* 16S rRNA genes that contrasts with their relatively high abundance in the Northern peat bogs and Arctic peat soils (Serkebaeva *et al.* 2013; Tveit *et al.* 2013). Despite these differences, we were able to prove the dominance of Alphaproteobacteria in *Sphagnum* microbiome by FISH–CLSM analysis. Furthermore, the microbiome composition was complemented with several taxa through sequencing of the metagenome that were not observed

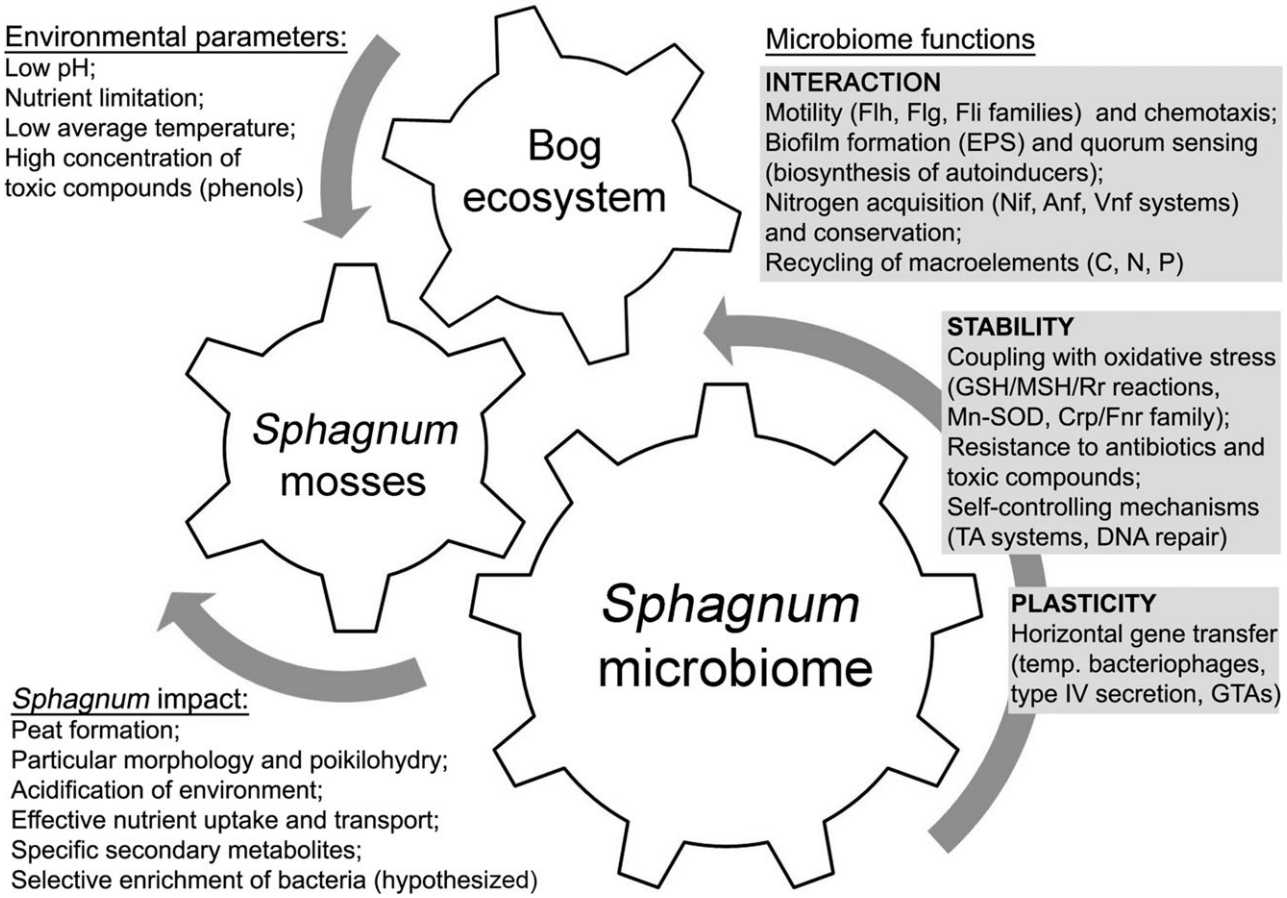

Fig. 4 Model for *Sphagnum*–microbiome biocoenosis. Interaction, stability and plasticity traits of microbiome were deduced from metagenomic sequences that were annotated using functional subsystems of SEED database. Examples in the brackets are the most pronounced and differentially abundant genetic signatures. EPS: extracellular polysaccharides; GSH: glutathione; MSH: mycothiol; Rr: rubrerythrin; TA: toxin–antitoxin; GTAs: genetic transfer agents.

in previous studies (Bragina *et al.* 2012a; Serkebaeva *et al.* 2013), for instance subdivisions 5 and 6 (*Acidobacteria*) and Phycisphaerae (*Planctomycetes*). Additionally, we found evidence of the dominant functional groups (subsystems) of *Proteobacteria*, which were reported as the most abundant nitrogen-fixing bacteria associated with *Sphagnum* mosses (Bragina *et al.* 2012b, 2013) and subsurface peat layers (Lin *et al.* 2014). Moreover, genes that encode for autoinducers produced by Proteobacteria for quorum sensing (Miller & Bassler 2001) were shown among the dominant functional groups of the *S. magellanicum* metagenome.

For the interpretation of metagenomic data, several limitations have to be considered (Committee on Metagenomics 2007; Thomas *et al.* 2012). For example, high-throughput sequencing of the metagenome provides only a partial DNA sampling, which, however, might have to be used to predict general features rather than analyse the total functional diversity of the sample (Prakash & Taylor 2012). Furthermore, automatic *in silico* annotation is characterized by a relatively high error rate and disregards proteins of unknown function as well (Teeling & Glöckner 2012). However, through the combination of the newly discovered genetic features and knowledge of ecological ontology of the samples, we can cautiously interpret the metagenomic data in terms of microbiome biodiversity and functioning. For the inter-environmental comparison, we used publically available and *S. magellanicum* metagenomes that were generated using Roche 454 and Illumina technologies, respectively. Although these technologies vary in sequencing depth and reads length, they provide comparable view of the sampled communities (Luo *et al.* 2012). Despite this fact, this technical source of error cannot be completely excluded. Moreover, sampling strategies, DNA isolation procedure and library preparation can be potential confounding factors of the analysis.

Besides the importance of the *Sphagnum* microbiome for ecosystem function in association with the host plants – seen as meta-organisms – and within the

peatland ecosystem, this microbiome presents a promising bio-resource for environmental biotechnology. For example, drought resistance is one of the major challenges for sustainable agriculture influenced by climate change (Berg *et al*. 2013b). Stress-protecting bacteria that have co-evolved in association with the poikilohydric *Sphagnum* moss can contribute to solve these problems as already shown by Zachow *et al*. (2013).

Acknowledgements

We thank Christian Berg (Graz) and Wilfried R. Franz (Klagenfurt) for their bryological expertise, sampling and helpful discussions. We would also like to acknowledge Meg Starcher (Graz/Washington) for English revision of the manuscript. We are grateful to Massimiliano Cardinale (Graz) for assistance with computer reconstruction of FISH–CLSM data. This work has been supported by a grant to G.B. (K2 grant 824186) from the Federal Ministry of Economy, Family and Youth (BMWFJ), the Federal Ministry of Traffic, Innovation and Technology (bmvit), the Styrian Business Promotion Agency SFG, the Standortagentur Tirol and ZIT – Technology Agency of the City of Vienna through the COMET-Funding Program managed by the Austrian Research Promotion Agency FFG. We also gratefully acknowledge support from NAWI Graz.

References

Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Applied and Environmental Microbiology*, **56**, 1919–1925.

Barak JD, Jahn CE, Gibson DL, Charkowski AO (2007) The role of cellulose and O-antigen capsule in the colonization of plants by *Salmonella enterica*. *Molecular Plant-Microbe Interactions*, **20**, 1083–1091.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **57**, 289–300.

Berg A, Danielsson Å, Svensson BH (2013a) Transfer of fixed-N from N-fixing cyanobacteria associated with the moss *Sphagnum riparum* results in enhanced growth of the moss. *Plant and Soil*, **362**, 271–278.

Berg G, Zachow C, Müller H, Philipps J, Tilcher R (2013b) Next-generation bio-products sowing the seeds of success for sustainable agriculture. *Agronomy*, **3**, 648–656.

Bragina A, Berg C, Cardinale M, Shcherbakov A, Chebotar V, Berg G (2012a) *Sphagnum* mosses harbor highly specific bacterial diversity during their whole lifecycle. *The ISME Journal*, **6**, 802–813.

Bragina A, Maier S, Berg C *et al*. (2012b) Similar diversity of Alphaproteobacteria and nitrogenase gene amplicons on two related *Sphagnum* mosses. *Frontiers in Microbiology*, **2**, 275.

Bragina A, Berg C, Müller H, Moser D, Berg G (2013) Insights into functional bacterial diversity and its effect on Alpine bog ecosystem functioning. *Scientific Reports*, **3**, 1955.

Bray JR, Curtis JT (1957) An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs*, **27**, 325–349.

Bulgarelli D, Schlaeppi K, Spaepen S, Ver Loren van Themaat E, Schulze-Lefert P (2013) Structure and functions of the bacterial microbiota of plants. *Annual Review Plant Biology*, **64**, 807–838.

Canchaya C, Fournous G, Chibani-Chennoufi S, Dillmann ML, Brüssow H (2003) Phage as agents of lateral gene transfer. *Current Opinion in Microbiology*, **6**, 414–424.

Caporaso JG, Kuczynski J, Strombauh J *et al*. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, **7**, 335–336.

Committee on Metagenomics, National Research Council (2007) *Data Management and Bioinformatics Challenges of Metagenomics. The New Science of Metagenomics. Revealing the Secrets of Our Microbial Planet*, pp. 85–97. The National Academies Press, Washington, District of Columbia.

Cox MP, Peterson DA, Biggs PJ (2010) SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics*, **11**, 485.

D'aes J, De Maeyer K, Pauwelyn E, Höfte M (2010) Biosurfactants in plant-*Pseudomonas* interactions and their importance to biocontrol. *Environmental Microbiology Reports*, **2**, 359–372.

Daims H, Brühl A, Amann R, Schleifer KH, Wagner M (1999) The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. *Systematic and Applied Microbiology*, **22**, 434–444.

Daniels RE, Eddy A (1985) *Handbook of European Sphagna*. Institute of Terrestrial Ecology, Aberystwyth, UK.

Delmotte N, Knief K, Chaffron S *et al*. (2009) Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 16428–16433.

DeSantis TZ, Hugenholtz P, Larsen N *et al*. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, **72**, 5069–5072.

Dise NB (2009) Environmental science. Peatland response to global change. *Science*, **326**, 810–811.

Danhorn T, Fuqua C (2007) Biofilm formation by plant-associated bacteria. *Annual Review of Microbiology*, **61**, 401–422.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**, 2460–2461.

Gilbert JA, O'Dor R, King N, Vogel TM (2011) The importance of metagenomic surveys to microbial ecology: or why Darwin would have been a metagenomic scientist. *Microbial Informatics and Experimentation*, **1**, 5.

Gilbert SF, Sapp J, Tauber AI (2012) A symbiotic view of life: we have never been individuals. *The Quarterly Review of Biology*, **87**, 325–341.

Gomez-Alvarez V, Teal TK, Schmidt TM (2009) Systematic artifacts in metagenomes from complex microbial communities. *The ISME Journal*, **3**, 1314–1317.

Granhall U, Hofsten AV (1976) Nitrogenase activity in relation to intracellular organisms in *Sphagnum* mosses. *Physiologia Plantarum*, **36**, 88–94.

Grube M, Cardinale M, de Castro JV Jr, Müller H, Berg G (2009) Species-specific structural and functional diversity of

bacterial communities in lichen symbioses. *The ISME Journal*, **3**, 1105–1115.

Horvath P, Barrangou R (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science*, **327**, 167–170.

Jassey VEJ, Gilbert D, Binnet P, Toussaint ML, Chiapusioa G (2011) Effect of temperature gradient on *Sphagnum fallax* and its associated microbial communities: a study under controlled conditions. *Canadian Journal of Microbiology*, **57**, 226–235.

Jin CW, Li GX, Yu XH, Zheng SJ (2010) Plant Fe status affects the composition of siderophore-secreting microbes in the rhizosphere. *Annals of Botany*, **105**, 835–841.

Joosten H, Couwenberg J (2009) *Are Emissions Reductions from Peatlands MRV-able?*. Wetlands International, Ede, the Netherlands.

Kato JY, Funa N, Watanabe H, Ohnishi Y, Horinouchi S (2007) Biosynthesis of gamma-butyrolactone autoregulators that switch on secondary metabolism and morphological development in *Streptomyces*. *Proceedings of the National Academy of Sciences*, **104**, 2378–2383.

Kip N, Ouyang W, van Winden J *et al.* (2011) Detection, isolation and characterization of acidophilic methanotrophs from *Sphagnum* mosses. *Applied and Environmental Microbiology*, **77**, 5643–5654.

Knief C, Delmotte N, Chaffron S *et al.* (2012) Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *The ISME Journal*, **6**, 1378–1390.

Lang AS, Zhaxybayeva O, Beatty JT (2012) Gene transfer agents: phage-like elements of genetic exchange. *Nature Reviews Microbiology*, **10**, 472–482.

Larmola T, Tuittila ES, Tiirola M *et al.* (2010) The role of *Sphagnum* mosses in the methane cycling of a boreal mire. *Ecology*, **91**, 2356–2365.

Law CW, Chen Y, Shi W, Smyth GK (2014) Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology*, **15**, R29.

Lin X, Tfaily MM, Green SJ *et al.* (2014) Microbial metabolic potential for carbon degradation and nutrient (nitrogen and phosphorus) acquisition in an ombrotrophic peatland. *Applied and Environmental Microbiology*, **80**, 3531–3540.

Loy A, Maixner F, Wagner M, Horn M (2007) probeBase – an online resource for rRNA-targeted oligonucleotide probes: new features 2007. *Nucleic Acids Research*, **35**, D800–D804.

Luo C, Tsementzi D, Kyrpides N, Read T, Konstantinidis KT (2012) Direct comparison of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. *PLoS One*, **7**, e30087.

Mäki M, Renkonen R (2004) Biosynthesis of 6-deoxyhexose glycans in bacteria. *Glycobiology*, **14**, 1R–15R.

Massey FJ Jr (1951) The Kolmogorov–Smirnov test for goodness of fit. *The Journal of American Statistical Association*, **46**, 68–78.

Meyer F, Paarmann D, D'Souza M *et al.* (2008) The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, **9**, 386.

Miller MB, Bassler BL (2001) Quorum sensing in bacteria. *Annual Review in Microbiology*, **55**, 165–199.

Nadell CD, Xavier JB, Foster KR (2009) The sociobiology of biofilms. *FEMS Microbiology Reviews*, **33**, 206–224.

Ochman H, Moran NA (2001) Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. *Science*, **292**, 1096–1099.

Ondov BD, Bergman NH, Phillippy AM (2011) Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics*, **12**, 385.

Opelt K, Berg G (2004) Diversity and antagonistic potential of bacteria associated with bryophytes from nutrient-poor habitats of the Baltic Sea coast. *Applied and Environmental Microbiology*, **70**, 6569–6579.

Opelt K, Berg C, Schönmann S, Eberl L, Berg G (2007a) High specificity but contrasting biodiversity of *Sphagnum*-associated bacterial and plant communities in bog ecosystems independent of the geographical region. *The ISME Journal*, **1**, 502–516.

Opelt K, Chobot V, Hadacek F, Schönmann S, Eberl L, Berg G (2007b) Investigation of the structure and function of bacterial communities associated with *Sphagnum* mosses. *Environmental Microbiology*, **91**, 2795–2809.

Overbeek R, Begley T, Butler RM *et al.* (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Research*, **33**, 5691–5702.

Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH (2013) Going back to the roots: the microbial ecology of the rhizosphere. *Nature Reviews Microbiology*, **11**, 789–799.

Prakash T, Taylor TD (2012) Functional assignment of metagenomic data: challenges and applications. *Briefings in Bioinformatics*, **13**, 711–727.

Putkinen A, Larmola T, Tuomivirta T *et al.* (2012) Water dispersal of methanotrophic bacteria maintains functional methane oxidation in *Sphagnum* mosses. *Frontiers in Microbiology*, **3**, 15.

Raghoebarsing AA, Smolders AJP, Schmid MC *et al.* (2005) Methanotrophic symbionts provide carbon for photosynthesis in peat bogs. *Nature*, **436**, 1153–1156.

Rydin H, Jeglum JK (2006) Nutrients, light and temperature. In: *The Biology of Peatlands* (eds Rydin H, Jeglum JK), pp. 164–186. Oxford University Press, Oxford.

Scheibe R, Beck E (2011) Drought, desiccation, and oxidative stress. In: *Plant Desiccation Tolerance* (eds Lüttge U, Beck E & Bartels D), pp. 209–232. Springer-Verlag, Berlin, Heidelberg.

Schimel JP, Bennett J (2004) Nitrogen mineralization: challenges of a changing paradigm. *Ecology*, **85**, 591–602.

Serkebaeva YM, Kim Y, Liesack W, Dedysh SN (2013) Pyrosequencing-based assessment of the Bacteria diversity in surface and subsurface peat layers of a Northern wetland, with focus on poorly studied phyla and candidate divisions. *PLoS One*, **8**, e63994.

Sessitsch A, Hardoim P, Weilharter A, Krause A, Woyke T, Mitter B (2012) Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. *Molecular Plant-Microbe Interactions*, **25**, 28–36.

Smyth GK (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, **3**, article 3.

Soudzilovskaia NA, Cornelissen JHC, During HJ, van Logtestijn RSP, Lang SI, Aerts R (2010) Similar cation exchange capacities among bryophyte species refute a presumed mechanism of peatland acidification. *Ecology*, **91**, 2716–2726.

Strack M (2008) *Peatlands and Climate Change*. International Peat Society, Jyväskylä, Finnland.

Succow M, Joosten H (2001) *Landschaftsökologische Moorkunde*. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart, Germany.

Teeling H, Glöckner FO (2012) Current opportunities and challenges in microbial metagenome analysis - a bioinformatic perspective. *Briefings in Bioinformatics*, **13**, 728–742.
Thomas T, Rusch D, DeMaere MZ *et al.* (2010) Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. *The ISME Journal*, **4**, 1557–1567.
Thomas T, Gilbert J, Meyer F (2012) Metagenomics – a guide from sampling to data analysis. *Microbial Informatics and Experimentation*, **2**, 3.
Tveit A, Schwacke R, Svenning MM, Urich T (2013) Organic carbon transformation in high-Arctic peat soils: key functions and microorganisms. *The ISME Journal*, **7**, 299–311.
Van Melderen L (2010) Toxin-antitoxin systems: why so many, what for? *Current Opinion in Microbiology*, **13**, 781–785.
Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, **73**, 5261–5267.
Wiegert T (2010) Regulated intramembrane proteolysis in bacterial transmembrane signaling. In: *Bacterial Signaling* (eds Krämer R & Jung K), pp. 229–246. Willey-VCH, Weinheim.
Zachow C, Müller H, Tilcher R, Donat C, Berg G (2013) Catch the best: novel screening strategy to select stress protecting agents for crop plants. *Agronomy*, **3**, 794–815.

G.B., H.M. and C.Z. designed the research. G.B. collected the samples. L.O.-W. and C.Z. performed the research, B.H. and G.G.T. assembled the contigs and performed the statistical analysis. A.B. and H.M. analysed the bioinformatic data. A.B., L.O.-W., H.M. and G.B. wrote the manuscript. All authors discussed the results and commented on the manuscript at all stages.

Data accessibility

The complete metagenome of *S. magellanicum* is publicly available at the MG-RAST server under the accession no. 4533611.3. Partial 16S rRNA gene sequences from the complete *S. magellanicum* metagenome and partial gene sequences encoding for nitrogen metabolism from the complete and normalized metagenomes were deposited in the DRYAD repository (http://datadryad.org/pages/repository) under the accession identifier doi:10.5061/dryad.9r816.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1. Description of publicly available metagenomes used in this study.

Table S2. Summary of the Kolmogorov–Smirnov test results.

Table S3. Overview of sequencing data after CLC genomic workbench *de novo* assembly and MG-RAST analysis.

Table S4. Statistical data for differentially abundant functional subsystems of the complete *S. magellanicum*, higher plants and peat soils metagenomes.

Fig. S1 This photograph shows *S. magellanicum* plants (red) as the dominant vegetation component and higher plants such as *Calluna vulgaris* and *Andromeda polifolia* in the bog ecosystem.

Fig. S2 Density plots of the statistically analysed metagenomes.

Fig. S3 Annotation of the normalised *S. magellanicum* metagenome using various databases.

Fig. S4 Taxonomic hits distribution of the complete *S. magellanicum* metagenome.

Fig. S5 Colonisation pattern of *S. magellanicum* microbiome visualised by FISH-CLSM.

Fig. S6 Functional heatmap of the complete *S. magellanicum*, higher plants, and peat soils metagenomes.

Fig. S7 Dominant functional groups of the normalised *S. magellanicum* metagenome.

Fig. S8 KEGG map of enzymes involved in nitrogen metabolism in the *S. magellanicum* metagenomes.