Phylogenetic and genomic analysis of *Methanomassiliicoccales* in wetlands and animal intestinal tracts reveals clade-specific habitat preferences.

Andrea Söllinger¹, Clarissa Schwab², Thomas Weinmaier³, Alexander Loy³, Alexander T. Tveit⁴, Christa Schleper¹, Tim Urich¹,⁵*

¹ Department of Ecogenomics and Systems Biology, University of Vienna, Austria
² Institute of Food, Nutrition and Health, ETH Zürich, Switzerland
³ Department of Microbiology and Ecosystem Science, University of Vienna, Austria
⁴ Department of Arctic and Marine Biology, UiT The Arctic University of Norway, Tromsø, Norway
⁵ Institute for Microbiology, Ernst-Moritz-Arndt University Greifswald, Germany

*corresponding author, Institute for Microbiology, Ernst-Moritz-Arndt University Greifswald,
Friedrich-Ludwig-Jahn-Str. 15, 17489 Greifswald, Germany, T: +49-3834-864203,
tim.urich@uni-greifswald.de
Abstract

Methanogenic *Thermoplasmata* of the novel order *Methanomassiliicoccales* were recently discovered in human and animal gastro-intestinal tracts (GITs). However their distribution in other methanogenic environments has not been addressed systematically. Here we surveyed *Methanomassiliicoccales* presence in wetland soils, a globally important source of methane emissions to the atmosphere, and in the GITs of different animals by PCR targeting their 16S rRNA and methyl:coenzyme M reductase (α-subunit) genes. We detected *Methanomassiliicoccales* in all 16 peat soils investigated, indicating their wide distribution in these habitats. Additionally, we detected their genes in various animal feces.

*Methanomassiliicoccales* were subdivided in two broad phylogenetic clades designated ‘environmental’ and ‘GIT’ clades based on differential, although non-exclusive, habitat preferences of their members. A well-supported cluster within the environmental clade comprised more than 80 % of all wetland 16S rRNA gene sequences. Metagenome assembly from bovine rumen fluid enrichments resulted in two almost complete genomes of both *Methanomassiliicoccales* clades. Comparative genomics revealed that members of the environmental clade contain larger genomes and a higher number of genes encoding anti-oxidative enzymes than animal GIT clade representatives. This study highlights the wide distribution of *Methanomassiliicoccales* in wetlands, which suggests that they contribute to methane emissions from these climate-relevant ecosystems.

Keywords: methanogens, archaea, peat soil, methylotrophic methanogenesis, enrichments, metagenomics

Introduction
Methanogenic archaea (methanogens) produce the bulk of biologically generated methane (\(\text{CH}_4\)) on Earth, and are thus the largest global source of this potent greenhouse gas (Ciais et al. 2013). Methanogens occur in a wide range of predominantly anoxic environments including wetlands, marine and freshwater environments, landfills, and intestinal tracts of humans and various animals (Ciais et al. 2013). Natural wetlands such as peat soils cover less than 9% of the global land surface (Zedler and Kercher 2005), but are the major natural \(\text{CH}_4\) source, contributing approximately 20 - 30% to global \(\text{CH}_4\) emissions to the atmosphere (Ciais et al. 2013). However, 50 - 65% of all \(\text{CH}_4\) emitted to the atmosphere results from human activities such as animal husbandry. Enteric fermentation by ruminant animals is one major \(\text{CH}_4\) source, contributing approximately 30% to all anthropogenic \(\text{CH}_4\) emissions (Ciais et al. 2013).

All known methanogens belong to the phylum of Euryarchaeota and were until lately classified into six orders, the Methanomicrobiales, Methanobacteriales, Methanosarcinales, Methanococcales, Methanocellales and Methanopyrales. An additional order of methanogens was recently proposed within the class Thermoplasma (Paul et al. 2012) and was confirmed by the isolation of Methanomassiliicoccus luminyensis (Dridi et al. 2012), the first and still only isolate from this 7th methanogen order, the Methanomassiliicoccales (Oren and Garrity 2013). Several recent studies showed that Methanomassiliicoccales have an energy metabolism distinct from other methanogens. All currently published enrichment cultures and the sole isolate M. luminyensis were obtained on methanol and \(\text{H}_2\) (Borrel et al. 2012; Dridi et al. 2012; Paul et al. 2012; Borrel et al. 2013a; Iino et al. 2013). More recently, methylamines were identified as electron acceptors besides methanol (Poulsen et al. 2013; Brugere et al. 2013; Lang et al. 2014). Genomic analyses showed that all Methanomassiliicoccales genomes analysed so far encode a truncated methanogenesis pathway (Lang et al. 2014) via \(\text{H}_2\)-
dependent methylotrophic methanogenesis (Borrel et al. 2013b; Borrel et al. 2014; Lang et al. 2014).

*Methanomassiliicoccales* affiliated 16S rRNA gene sequences have been recovered from natural and man-made anoxic habitats, e.g. in rice paddy fields (Großkopf et al. 1998), anaerobic digestors (Godon et al. 1997) and in gastro-intestinal tracts (GIT) of various ruminant and non-ruminant animals, such as cattle (Tajima et al. 2001), sheep (Wright et al. 2004), reindeer (Sundset et al. 2009), yak (Huang et al. 2012) and wallaby (Evans et al. 2009). Paul et al. (2012) were the first to note an environment-specific clustering of *Methanomassiliicoccales* 16S rRNA gene sequences, with the observation of an animal-associated ‘intestinal cluster’ distinct from environmental sequences. Curiously, the type strain *M. luminyensis*, although obtained from human feces, does not belong to the ‘intestinal cluster’ but is associated with the environmental group, leaving the ‘intestinal cluster’ poorly characterised due to the lack of isolates. *Methanomassiliicoccales* of the ‘intestinal cluster’ have been shown to constitute a large proportion of the archaeal population in GITs of ruminant animals (Gu et al. 2011; St-Pierre and Wright 2013) as revealed by 16S rRNA gene-based clone library analyses.

The occurrence of *Methanomassiliicoccales* in wetlands, the major natural CH$_4$ source, has not been systematically assessed yet. This study aimed to reveal the environmental distribution of *Methanomassiliicoccales* in a range of different wetland types, spanning acidic and neutral peatlands from temperate and arctic regions. For this purpose a PCR screening for *Methanomassiliicoccales* 16S rRNA genes and *mcrA*, encoding the α-subunit of the methanogenesis key enzyme methyl:coenzyme M reductase (Mcr), as phylogenetic and functional marker genes was conducted. The screening was complemented with samples from animal intestinal tracts to enable the assessment of environment-specific patterns of *Methanomassiliicoccales* occurrence. We show that distinct *Methanomassiliicoccales* clades
dominate in wetlands compared to animals. Furthermore, we report the enrichment of rumen
*Methanomassiliicoccales* to approximately 30% relative abundance and an accompanying
enrichment bias that favours close relatives of *M. luminyensis*. We additionally provide
insights into clade-specific genomic adaptations by comparative genomics, including two new
draft genomes that were reconstructed from metagenomic data of the rumen enrichment
cultures.

**Materials and methods**

**Wetland and animal samples**

Origin and characteristics of wetland samples are given in Table 1. The two arctic peat
samples were from Svalbard (Tveit *et al.* 2012) and DNA extracts from 14 temperate wetland
samples were from Germany, Austria and Italy (Steger *et al.* 2011). Kangaroo, elephant, giant
tortoise, tortoise and rat fecal samples were collected on the 15th of February 2010 from the
Viennese Zoo (Tiergarten Schönbrunn), flash-frozen in liquid nitrogen and stored at -80°C
until DNA extraction. A bovine rumen fluid sample (AUT RF) was obtained from the
University of Veterinary Medicine Vienna. After centrifugation (5 min, 95 g) to remove
particulate material, microbial cells were harvested by centrifugation (5 min, 16,500 g) and
frozen at -20 °C until usage. A rumen fluid sample (DK RF) was received from Aarhus
University, Denmark. The rumen fluid was supplemented with 15 mM trimethylamine,
incubated anoxically in serum bottles for 24 h at 38 °C and shipped at 4 °C for 2 d. Cells were
harvested by centrifugation (5 min, 16,500 g) and frozen at -20 °C. A third rumen fluid
sample (cow) derived from a previously described 24 h incubation experiment (Poulsen *et al.*
2013).

**DNA extraction, clone library construction and RFLP**
DNA from arctic peat and animal feces samples was extracted using a phenol/chloroform/isoamyl alcohol bead beating protocol (Urich et al. 2008). QIAamp DNA Stool Mini Kit (QIAGEN, Valencia, CA, USA) was used to extract DNA from Austrian and Danish bovine rumen fluid (AUT/DK RF) according to the manufacturer’s instructions. While the usage of different DNA extraction methods likely affects the broad resulting bacterial and archaeal community composition, the reported dominant *Methanomassiliicoccales* phylotypes are likely unaffected. Their cell wall structure consisting of two membranes and no murein sacculus or S-layer should be disrupted even with the latter protocol not containing a bead beating step. DNA extracts were quality checked and quantified by agarose gel electrophoresis and NanoDrop® (ND-1000, Peqlab, Erlangen, Germany). Three different primer pairs were employed to amplify *Methanomassiliicoccales* 16S rRNA genes and *mcrA* of *Methanomassiliicoccales* and other methanogens (see Tab. 2 for details). PCR amplification was performed using a T Professional Thermocycler (biometra, Göttingen, Germany). Reaction mixtures (50 µL) contained 10 µL 5x Green GoTaq® Flexi Buffer, 4 µL MgCl (2 mM), 1 µL dNTPs (0.2 mM), 0.25 µL bovine serum albumine (0.2 mgmL⁻¹), 1 µL of each primer (0.5 µM), 0.25 µL DreamTaq (0.025 U µL⁻¹) and 1 µL DNA template. PCRs with AS1/AS2 and mlas-mod For/ mcrA-rev primer pairs were performed with 34 cycles and 45 seconds of annealing (60 °C) and elongation (72 °C).  PCR using Tp-mcrA-F/Tp-mcrA-R primers were performed with 36 cycles and 60 seconds of annealing (60 °C) and 45 seconds of elongation (72 °C). Amplified products were quality checked via agarose gel electrophoresis and purified (QIAquick PCR Purification Kit, QIAGEN, Valencia, CA, USA). PCR products were cloned using pGEM®-T Vector System I (Promega, Madison, WI, USA) and chemically competent *E. coli* cells (One Shot® TOP10 Chemically Competent *E. coli*, Invitrogen, Carlsbad, CA, USA). Ligation reactions were incubated one hour at room temperature or overnight at 4 °C until transformation. Subsequently, the cells were plated on LB agar plates containing ampicillin (100 mgL⁻¹), IPTG (0.5 mM) and X-Gal (80 mgmL⁻¹),
and incubated overnight at 37 °C. Recombinant *E. coli* colonies were identified by blue-white screening. Colony PCR using T7/SP6 primers was done to identify correct insert sizes with 32 cycles of PCR and 45 seconds of annealing (55 °C) and elongation (72 °C). More than 200 16S rRNA gene clones and nearly 200 mcrA clones (127 and 71 obtained with the universal and the specific mcrA primer, respectively) were screened (8 to 16 16S rRNA and 8 to 32 mcrA clones per wetland and animal sample). Quality check and purification were done as described above. RFLP analysis of 16S rRNA and mcrA PCR products was done with the enzyme HpyCH4V (New England BioLabs, Ipswich, MA, USA). The restriction reaction was performed at 37 °C (3 h) and resulting fragments were visualised by electrophoresis on 3 % agarose gels.

**Sequencing and sequence analysis**

Sanger sequencing of clones that were representative for the different RFLP patterns was performed by LGC genomics (Berlin, Germany). Contaminating vector sequences were removed using NCBI VecScreen (http://www.ncbi.nlm.nih.gov/tools/vecscreen/). 16S rRNA and mcrA gene sequences were verified by megablast against NCBI Nucleotide collection (nr/nt) and by blastx against NCBI Non-redundant protein (nr) database with default settings, respectively (Altschul et al. 1990; Camacho et al. 2009; Sayers et al. 2011).

Representative 16S rRNA gene sequences were clustered into operational taxonomic units (OTUs) at ≥98 % sequence identity using cd-hit-est of the CD-HIT Suite (Huang et al. 2010). *McrA* gene sequences were translated into amino acid sequences using BioEdit (Hall 1999) and clustered at ≥98 % amino acid identity. 16S rRNA gene sequences were aligned with a set of representative reference sequences of *Thermoplasmata* and *Methanomassiliicoccales* 16S rRNA genes using SINE online (Pruesse et al. 2012). The reference sequences were selected form the SILVA Ref NR SSU r123 database (Quast et al. 2013); high quality sequences with a sequence length of >1250 nt were clustered at 94.5 % sequence identity using cd-hit-est of
the CD-HIT Suite (Huang et al. 2010) to obtain representative sequences of potential

*Methanomassiliicoccales* and *Thermoplasmata* genera (Yarza et al. 2014). Inferred McrA amino acid sequences were aligned with a set of reference McrA amino acid sequences of methanogens derived from Poulsen et al. (2013) and other studies on

*Methanomassiliicoccales* (minimum sequence length of 549 AA) using MAFFT v6.864, L-INS-i (Katoh and Toh 2008). The McrA amino acid sequence alignment was improved by gap removal with Gblocks (Talavera and Castresana 2007) using the least stringent parameters to avoid losing phylogenetic information. Phylogenetic reference trees (16S rRNA gene and McrA) were constructed using PhyML (Guindon and Gascuel 2003), a maximum likelihood method implemented in ARB (Ludwig et al. 2004), and bootstrap values were calculated (500 replicates each) to verify branch support. To select the best suited nt and AA substitution model the respective alignments were uploaded to the Model Selection tool of the IQ-TREE web server (Minh et al. 2013). The 16S rRNA reference tree was then calculated based on the GTR model and the McrA reference tree was constructed based on the Dayhoff model. Afterwards, the shorter *Methanomassiliicoccales* 16S rRNA gene and McrA protein sequences obtained in this study were placed into the reference trees using the ARB Maximum Parsimony tool (Ludwig et al. 2004). FigTree (http://tree.bio.ed.ac.uk/software/figtree/) was used for visualizing phylogeny.

**Methanomassiliicoccales enrichments from cow rumen**

Rumen *Methanomassiliicoccales* enrichments were initiated with rumen fluid of a fistulated Braunvieh (Brown Swiss) fed on hay and dairy concentrate. After collection, cow rumen fluid was immediately centrifuged 5 minutes at 1000 rpm to remove large particles. Initially, serum bottles containing modified *Methanobrevibacter* medium (Dridi et al. 2012) were inoculated with 20 % rumen fluid and incubated under an atmosphere containing N₂, CO₂ and H₂ (70:20:10) at 37 °C in the dark. Trimethylamine (TMA) was supplemented as electron
acceptor (10 mM). Rumen *Methanomassiliicoccales* enrichments were obtained after successive transfers (weekly; 10% inocula) on a medium containing (L⁻¹) 0.5 g KH₂PO₄, 0.4 g MgSO₄ · 7H₂O, 5 g NaCl, 1 g NH₄Cl, 0.05 g CaCl₂ · 2H₂O, 1.6 g sodium acetate, 0.5 g cysteine-HCl, 1 mL trace element solution, 2 mL tungstate-selenite solution, 4 g NaHCO₃, 0.5 g Na₂S · 9H₂O, 2 g sodium formate, 1 mL vitamin solution. This adjusted medium was prepared similar to the medium used by Dridi *et al.* (2012), supplemented with 10 - 60 mM TMA and sterile filtered rumen fluid (7.5% v/v), and incubated as described above.

*Methanomassiliicoccales* growth was monitored via quantitative PCR (qPCR) on a Mastercycler ep realplex (Eppendorf, Hamburg, Germany), using the same protocol and standards as employed by Poulsen and co-workers (2013).

**IonTorrent sequencing – (meta)genomic analysis**

DNA from two enrichment cultures (RumEn_MG1 from 7th transfer culture [Ion 314™ Chip Kit]; RumEn_MG2 from 13th transfer culture [Ion 316™ Chip Kit]) was single-end sequenced using IonTorrent PGM (life technologies, Carlsbad, CA, USA) and 200 bp chemistry. 16S rRNA gene fragments contained in the two metagenomes were taxonomically classified with CREST (Lanzén *et al.* 2012). Assembly of RumEn_MG2 was done with Newbler v2.9 (454 Life Sciences, Branford, CT, USA). Taxonomic binning of RumEn_MG2 contigs was done with PhymmBL (Brady and Salzberg 2009), after training on all available complete RefSeq genomes as of May 2013 and the two available *Methanomassiliicoccales* genomes *Candidatus* Methanomethylophilus alvus Mx1201 and *M. luminyensis* B10. The PhymmBL classification resulted in two *Methanomassiliicoccales* genome bins (referred to as RumEn M1 and RumEn M2 from now on). The partial genomes were functionally annotated with RAST using default settings (Aziz *et al.* 2008). RAST annotations of genes of interest (e.g. methanogenesis, energy and carbon acquisition and environmental adaptation) were verified using the BLAST tool implemented in RAST and by comparison with annotated
genes from other *Methanomassiliicoccales*, consulting different databases, e.g. Pfam (Finn et al. 2013), UniProt (UniProt Consortium 2014), and TCDB (Saier et al. 2013). tRNAs were predicted using tRNAscan-SE (Lowe and Eddy 1997) and ARAGORN (Laslett and Canback 2004). Genome completeness and contamination of both *Methanomassiliicoccales* genome bins was assessed using checkM with default settings (Parks et al. 2014).

**Growth experiments with glycine betaine**

*Methanomassiliicoccales* growth on glycine betaine (Sigma-Aldrich, St. Louis, MO, USA) was tested on *M. luminyensis* (DSMZ; DSM no. 25720). *M. luminyensis* was grown in a bicarbonate buffered medium based on Paul et al. (2012) under an atmosphere containing 20 % CO$_2$ and 80 % H$_2$. Serum bottles containing medium supplemented either with 16 mM glycine betaine, 60 mM glycine betaine or 50 mM methanol (positive control) as electron acceptors were inoculated with 10 % (v/v) of stationary phase *M. luminyensis* cultures that were grown on 50 mM methanol. These cultures and three negative controls (w/o electron acceptor, w/o inoculum but 50 mM methanol, w/o inoculum but 60 mM glycine betaine) were incubated at 37$^\circ$ C in the dark (three replicates each). Growth was regularly monitored during 46 days of incubation by OD measurements (DU 800 Spectrophotometer, Beckman Coulter, Pasadena, CA, USA) at 578 nm and microscopic cell counts.

**Data submission**

Representative sequences of OTUs were submitted to GenBank and are deposited under the accession number KT225396 - KT225423. The *Methanomassiliicoccales* mcrA gene sequences are available under the accession number KT225424 - KT225458. Annotated draft genomes are available under the accession numbers LJKK00000000 (RumEn M1) and LJKL00000000 (Rumen M2). Raw metagenomics sequencing data were deposited at the NCBI Sequence Read Archive (accession number SRP064292).
Results

Distribution of *Methanomassiliicoccales* in wetlands and animals

A screening of 16 arctic and temperate wetland samples with *Methanomassiliicoccales*-specific PCR detected *Methanomassiliicoccales* 16S rRNA genes in all samples (Fig. 1 & 2). *Methanomassiliicoccales* 16S rRNA genes were additionally detected in faeces from elephant, giant tortoise, tortoise and rat and in bovine rumen fluid, but not in kangaroo feces. Between 8 and 16 clones of each of the 22 clone libraries were subjected to RFLP analysis and clones with representative RFLP patterns were sequenced. The analysis revealed a low diversity of *Methanomassiliicoccales* with e.g. 10 out of 16 peat clone libraries containing only 1 or 2 species-level OTUs (<98 % sequence identity). Since our aim was the identification of dominant *Methanomassiliicoccales* OTUs in many samples and not an exhaustive characterisation of diversity patterns, we did not expand this analysis by using deep sequencing of *Methanomassiliicoccales* 16S rRNA gene amplicons. The 138 representative Sanger sequences that were obtained in total formed 28 distinct OTUs. Twenty-seven OTUs were affiliated with *Methanomassiliicoccales*, while one OTU belonged to *Thermoplasmata* of the marine benthic group D (MBG-D, Fig. 1). Seven OTUs, which were dominated by wetland clones, formed a monophyletic group with *M. luminyensis*, *Ca*. *M. intestinalis*, and 16S rRNA gene sequences from different aquatic and terrestrial environments, and is hereafter referred to as ‘environmental clade’ (Fig. 1). OTUs 1, 3 and 10 formed a narrow cluster that encompassed approximately 80 % of all wetland 16S rRNA gene sequences (‘wetland cluster’, Fig. 1), thus typifying the *Methanomassiliicoccales* that inhabit a wide range of peat environments. The sequences retrieved from animals were all but one from cow assigned to a second distinct sequence cluster comprising *Ca*. *M. alvus*, *Ca*. *M. caenicola*, *Ca*. *M. termitum* and sequences obtained from other ruminant and non-ruminant animals (Fig. 1).
This group corresponded to the animal associated clusters identified previously by Paul et al. (2012) and Borrel et al. (2013b) and is here referred to as ‘gastro-intestinal tract (GIT)’ clade.

Only twelve percent (n=15) of the wetland Methanomassiliicoccales sequences were members of the GIT clade. The observed preferential, although non-exclusive, environmental distribution of members of the two clades is illustrated in Figure 2 that displays the relative 16S rRNA clone abundance of both clades in each peat and animal sample. The Methanomassiliicoccales populations in all peat samples, independent of wetland type, latitude and pH, were dominated by the environmental clade, with the exception of sample AUT peat 7 (Fig. 2). The latter was an atypical peat sample, since it originated from a ferrous biofilm (Steger et al. 2011). All animal samples were dominated by GIT clade Methanomassiliicoccales and only one animal sample (cow) contained an environmental clade clone (Fig. 2).

A PCR screening for the functional marker gene of methanogens detected Methanomassiliicoccales affiliated mcrA sequences in both arctic peat soils (Fig. 1), in four temperate peat soils and in five animals (elephant, giant tortoise, tortoise, rat feces and cow rumen fluid). Again, RFLP analysis was applied and the translated 53 representative Methanomassiliicoccales McrA amino acid sequences clustered in 24 distinct OTUs (< 98 % amino acid identity). Analogous to the 16S rRNA gene tree, the McrA tree showed a well-supported separation of Methanomassiliicoccales McrA sequences into environmental and GIT clades (Fig. 1). These results are in congruence with previous literature (Paul et al. 2012, Borrel et al. 2013b). Furthermore, a cluster exclusively comprised of wetland McrA sequences was observed containing 9 out of 14 wetland Methanomassiliicoccales McrA sequences, indicative of a wetland cluster. However due to the lack of reference sequences bootstrap support for this cluster could not be given. To obtain insights into the relative abundance of Methanomassiliicoccales among methanogen communities in wetlands we
analysed metatranscriptome libraries from the arctic wetlands Knudsenheia and Solvatn (Tveit et al. 2014). We screened for rRNA transcripts of MCC and methanogens and found a low relative abundance of MCC (between 0 and 4.7 % of methanogen rRNA transcripts; Tab. 3).

Rumen Methanomassiliicoccales enrichment cultures

We conducted enrichment trials with bovine rumen fluid and TMA to obtain novel Methanomassiliicoccales isolates of the GIT clade. Quantitative PCR showed that bacteria dominated the prokaryotic community in the rumen fluid (>10⁹ bacterial 16S rRNA gene copies mL⁻¹ rumen fluid), while methanogens were less abundant (approximately 2 x 10⁷ 16S rRNA gene copies mL⁻¹), see Figure 3A. Approximately 40 % of all 16S rRNA gene copies from methanogens belonged to the Methanomassiliicoccales. All Methanomassiliicoccales 16S rRNA clones obtained from the initial rumen fluid (AUT RF; 26 clones) were assigned to the GIT clade (Fig. 1; Fig. 3A). Stable enrichments of up to 32 % Methanomassiliicoccales were obtained after successive transfers of subcultures (see material and methods section) with maximal relative abundance of 51 % (data not shown), while other rumen methanogens were not enriched (data not shown). Despite the employment of various anaerobic isolation approaches such as serial dilutions, filtrations, antibiotic treatments and ‘roll-tube method’, no pure culture of rumen Methanomassiliicoccales was obtained.

Two metagenomes, RumEn_MG1 (7th transfer) and RumEn_MG2 (13th transfer), were sequenced from enrichments, consisting of 54 Mbp (482,803 sequences of 114 bp mean length) and 239 Mbp (1,871,707 sequences of 128 bp mean length) sequence information, respectively. Taxonomic classification of the metagenomic 16S rRNA gene fragments showed Methanomassiliicoccales 16S rRNAs to be 20 % (RumEn_MG1) and 24 % (RumEn_MG2) of all 16S rRNA gene fragments. The classification also revealed differences in bacterial and Methanomassiliicoccales community composition between the enrichments (Fig. 3).
Remarkably, the initially prevalent GIT clade decreased in relative abundance to 31 % of *Methanomassiliicoccales* and the environmental clade became dominant in the enrichments (Fig. 3).

**Assembly and analysis of Methanomassiliicoccales draft genomes**

Newbler assembly of RumEn_MG2 resulted in 8,854 contigs (>500bp) of 10.71 Mbp of average contig length 1210 bp with the largest contig being 202,459 bp in length. Binning with PhymmBL resulted in two partial *Methanomassiliicoccales* genome bins, RumEn M1 (182 contigs, 2.21 Mbp) and RumEn M2 (18 contigs, 1.28 Mbp). Table 4 shows some characteristics of RumEn M1 and RumEn M2 and published *Methanomassiliicoccales* genomes. Analysis of their full-length 16S rRNA genes revealed 89 % 16S rRNA gene sequence identity confirming the distant relationship of M1 and M2. M1 belonged to the environmental clade (Fig. 1). The 96 % sequence identity to the 16S rRNA gene of *M. luminyensis* and *Ca. M. intestinalis* suggested that M1 represents a novel member of the family *Methanomassiliicoccaceae*, while M2 represents a novel member of the GIT clade (95 % seq. id. to *Ca. M. termitum* and 91 % seq. id. to *Ca. M. alvus*; Fig. 1). Both were with 87 % and 92 % 16S rRNA gene sequence identity also distantly related to BRNA1, another *Methanomassiliicoccales* genome obtained from a rumen enrichment culture (GenBank acc. no.: CP002916).

Genome completeness estimates with checkM based on presence/absence of single-copy and multi-copy marker genes revealed that both genomes were almost complete (M1: 96.6 % and M2: 94.1 %) and showed no or only low level of contamination (M1: 1.6 %; M2: 0 %). The completeness was also reflected by the number and type of encoded tRNA and tRNA synthetase genes (Tab. 4). No cysteine and tryptophan specific tRNAs were found in M1.

Both genomes contained a nearly full set of 21 aminoacyl tRNA synthetase genes
(20 common and one pyrrolysine specific), with tRNA$^\text{Ala}$ synthetase missing in M1 and tRNA$^\text{Pyl}$ synthetase missing in M2.

Both draft genomes possessed all genes for enzymes involved in the last step of methanogenesis, the reduction of methyl-coenzym M (CH$_3$-S-CoM) to CH$_4$ by methyl:coenzyme M reductase (mcrABGCD). Furthermore, both genomes encoded a soluble heterodisulfide reductase (HdrABC) and the associated methyl viologen-dependent hydrogenase (MvhADG), required for the regeneration of coenzyme M. Like all other Methanomassiliicoccales, both genomes encoded HdrD, one subunit of a membrane-bound heterodisulfide reductase (HdrDE), and a 11-subunit Fpo-like complex, homologue to the 11 core subunits of a membrane-bound F$_{420}$-methanophenazine oxidoreductase complex found in other methanogens (Moparthi and Hägerhäll 2011; Lang et al. 2014). However, one subunit (FpoA) was missing in M2.

M1, contrary to M2, encoded mtaBCA, the complete gene set for reducing methanol to methyl-coenzym M. Gene sets necessary for the utilisation of other methylated substrates were entirely missing in the M2 genome and M1 encoded only an incomplete gene set (mtmBC) for the reduction of monomethylamine to methyl-coenzym M. Like all methylamine methyltransferases (Krzycki 2004), the monomethylamine methyltransferase of M1 is predicted to contain pyrrolysine. Consistently, M1 contained all genes necessary for the biosynthesis of pyrrolysine (pylBCD) as well as pylS and pylT, which encode the enzyme catalysing the ligation of pyrrolysine to its specific tRNA (PylS) and the specific tRNA (tRNA$^\text{Pyl}$). The pyrrolysine-associated gene set (pylBCDST) was entirely missing in M2.

M1 and M2 lacked all genes necessary for the reduction of CO$_2$ to CH$_4$ and both encoded an ADP-forming acetyl-CoA synthetase (AscA) homologue, which allows heterotrophic growth on acetate. These are two common features of all Methanomassiliicoccales genomes known so far (Borrel et al. 2014; Lang et al. 2014).
RAST analysis revealed that M1 and M2 both contained a gene encoding a glycine betaine ABC transporter, a homologue to the functional glycine betaine ABC transporter of *Methanosarcina mazei*, encoded by *otaABC* (Schmidt *et al.* 2007). We tested growth of *M. luminyensis* with glycine betaine as electron acceptor instead of methanol. However, no growth, even with 60 mM glycine betaine, was observed (data not shown). Furthermore, M1 and M2 encoded several genes possibly involved in oxidative stress response. M1 encoded four peroxiredoxin and rubrerythrin genes, two rubredoxin and superoxide reductase genes as well as and one superoxide dismutase gene. In contrast, M2 encoded only two genes for peroxiredoxin, rubrerythrin, and superoxide reductase, two catalase genes, and no rubredoxin and superoxide dismutase genes.

**Discussion**

*Methanomassiliicoccales are widely distributed in wetlands*

There has been no systematic study on the distribution of the recently described 7th methanogen order *Methanomassiliicoccales* in wetlands. Our broadly conceived PCR screening for *Methanomassiliicoccales* in various wetland types, including arctic and temperate acidic and neutral peat soils, revealed the presence of *Methanomassiliicoccales* in all tested samples (Fig. 2), which suggests their ubiquitous distribution in wetlands. Thus, *Methanomassiliicoccales* contribute to the high diversity of methanogens in wetlands, that include observations of *Methanosarcinales*, *Methanobacteriales*, *Methanomicrobiales*, *Methanococcales* and *Methanocellales* in temperate wetlands (Nercessian *et al.* 1999; Upton *et al.* 2000; Yavitt *et al.* 2012; Andersen *et al.* 2013), and *Methanosarcinales*, *Methanobacteriales* *Methanomicrobiales* and *Methanocellales* in arctic wetlands (Galand *et al.* 2002; Hoj *et al.* 2005; Tveit *et al.* 2012), respectively.
There is currently few data on the relative abundance of *Methanomassiliicoccales* among methanogens in wetlands. A screening of five metatranscriptomic libraries from the arctic wetlands Knudsenheia and Solvatn (Tveit *et al.* 2014) for rRNA transcripts of *Methanomassiliicoccales* and methanogens revealed a rather low relative abundance of *Methanomassiliicoccales* among methanogens (Tab. 3). Furthermore, the general mcrA primer pair used in this study, targeting all methanogens, amplified mostly mcrA sequences associated with *Methanomicrobiales, Methanocellales* and *Methanosarcinales* (data not shown), suggesting that these methanogens dominated the wetlands while *Methanomassiliicoccales* were less abundant. This would be in line with many published studies on methanogenic archaea from wetlands (Steinberg and Regan 2008; Yrjälä *et al.* 2011; Yavitt *et al.* 2012; Deng *et al.* 2014; Mondav *et al.* 2014; He *et al.* 2015). One determinant for this might be the parallel dependency of *Methanomassiliicoccales* on two substrates (i.e. H₂ and methylated compounds), as recently suggested by (Lang *et al.* 2014).

However, further studies on *Methanomassiliicoccales* abundances and activity in wetlands are required to elucidate their contribution to methanogenesis. In contrast, *Methanomassiliicoccales* comprise a large proportion of the methanogen community in ruminant animals (St-Pierre and Wright, 2013). Correspondingly, approximately 40% of all methanogens detected in the bovine rumen fluid used in this study belonged to the *Methanomassiliicoccales*. The cause for this might be the stable conditions inside the animal host and the constant supply of both H₂ and methylated compounds. Given the approximately similar magnitudes of CH₄ emitted to the atmosphere from wetlands and ruminant livestock, these preliminary data might indicate that wetland *Methanomassiliicoccales* are contributing less to global CH₄ emissions compared to their relatives in animal GITs.

**Clade-specific habitat preferences of *Methanomassiliicoccales***
The phylogenetic separation between the majority of wetland and animal associated
Methanomassiliicoccales 16S rRNA gene and McrA sequences (Fig. 1) revealed an
environmental clade representing the Methanomassiliicoccaceae and a second family-level
clad associated with the animal GIT. Our data from abundant clone library members show
that in particular one narrow cluster of OTUs within the environmental clade comprises the
majority of wetland Methanomassiliicoccales. The methanogenic nature of this wetland
cluster is supported by Methanomassiliicoccales McrA sequences forming a similarly
coherent cluster. Our wetland cluster is part of the ‘lake pavin cluster’ that was previously
mentioned by Borrel et al. (2013b). The prevalence of the wetland cluster in 15 out of 16
investigated samples, irrespective of wetland type, pH and latitude suggests that our screening
has identified the most abundant Methanomassiliicoccales in wetlands worldwide.

Methanomassiliicoccales sequences obtained from a great variety of ruminant and non-
ruminant animals form a GIT-specific Methanomassiliicoccales clade. However, the habitat
distribution of environmental and GIT Methanomassiliicoccales clades is non-exclusive, as
both Methanomassiliicoccales clades contained at least one clone obtained from the other
habitat. These results point at a co-occurrence of both clades, although environmental and
GIT clade Methanomassiliicoccales are dominating the respective habitat. This assumption is
strengthened by the results from the rumen enrichment trials where an originally low
abundant environmental clade Methanomassiliicoccales (represented by the genome bin M1)
was dominating the rumen enrichments after several passages.

Enrichment bias favours environmental clade

To our surprise, the two genome bins M1 and M2 from rumen fluid metagenomes represented
both Methanomassiliicoccales clades and the initially not detected environmental clade (M1)
became dominant. One reason for this might be an enrichment bias favouring the
environmental over the GIT clade. In fact, the composition of the enrichment medium was
very similar to the one used to isolate *M. luminyensis* from human feces, although human 
feces is reportedly dominated by members of the GIT clade (Mihajlovski *et al.* 2010; 
Vanderhaeghen *et al.* 2015; our own unpublished data). Since *M. luminyensis* does not belong 
to the GIT clade but is associated with the environmental clade, it is reasonable to assume that 
similar enrichment biases against the intestinal clade took place in both studies, enriching for 
a previously low abundant *Methanomassiliicoccales* of the *Methanomassiliicoccus* genus in 
our case. One could speculate that the smaller genome size of GIT clade 
*Methanomassiliicoccales* reflects the specialization to a very particular environment, which 
causes a lower competitiveness under the artificial enrichment conditions compared to 
*Methanomassiliicoccus*.

**Methanomassiliicoccales core metabolism**

The presence of genes for methanogenesis key enzymes (e.g. *mcrABG*, *hdrABC* and 
*mvhADG*) in both genomes confirmed the methanogenic nature of RumEn M1 and RumEn 
M2. In general our results agree with recent comparative genomic studies that have 
established a hydrogen-dependent methylotrophic methanogenesis for 
*Methanomassiliicoccales* (Borrel *et al.* 2013b; Borrel *et al.* 2014; Lang *et al.* 2014). Detailed 
metabolic schemes can be found in these earlier publications. The reoxidation of ferredoxin 
and generation of electrochemical membrane potential is possibly catalyzed by a 
ferredoxin:heterodisulfide oxidoeductase complex (Fpo-like complex + HdrD) in a recently 
proposed novel mode of energy conversion (Lang *et al.* 2014). The presence of *acsA* gene 
suggests that both M1 and M2 are likely chemolithoheterotrophs using acetate as carbon 
source.

However, the substrate spectrum of M1 and M2 for electron acceptors for methanogenesis 
remains somewhat obscure. In contrast to *M. luminyensis*, *Ca.* M. *intestinalis* and *Ca.* M. 
alvus, both partial genomes do not encode any genes necessary for the reduction of TMA to
methyl-coenzyme M, although the enrichments were incubated with TMA. In fact, no homologues of any gene involved in the reduction of methylated substrates to CH$_4$ were detected in the M2 genome. On the contrary, M1 encoded enzymes of methanol and monomethylamine reduction. However, these analyses are somewhat hampered by the fact that both genomes are not complete.

**Genome differences and possible adaptations**

The most evident difference between RumEn M1 (belonging to the environmental clade) and RumEn M2 (belonging to the GIT clade) was their divergent genome size. Despite similar genome completeness, M1 had a nearly 1 Mbp larger genome than M2. Available genome data show that the environmental clade genomes are consistently larger than the GIT clade genomes (Tab. 4), possibly indicative of massive genomic streamlining in the GIT clade. It appears that the GIT clade is better adapted to life in rather nutrient-rich, stable GIT environments at the expense of versatility and competitiveness under variable environmental conditions. In contrast, the larger genome size within the environmental clade offers this versatility for life in highly fluctuating environments such as soils and sediments.

Borrel *et al.* (2014) suggested several specific adaptations of *Methanomassiliicoccales* to soil environments, i.e. diazotrophy, osmoprotection and a greater antioxidative capacity encoded in the genome of *M. luminyensis* compared to genomes derived from animal GITs. The latter show adaptations to the GIT environment, e.g. presence of resistance genes to bile salts in *Ca. M. alvus* genome (Borrel *et al.* 2014). Neither RumEn M1 nor RumEn M2 encoded a choloylglycine hydrolase gene like *Ca. M. alvus*, which is involved in bile salt resistance. Since bile salts are secreted to the small intestine of ruminants and not to the rumen (Bauman and Lock 2006), there might be no selective pressure for rumen *Methanomassiliicoccales* to keep this gene. M1 and M2 lack 4 out of the 6 previously defined core genes for N$_2$-fixation.
conserved among all diazotrophs (Dos Santos et al. 2012), i.e. \textit{nifDEKN} are missing, while homologues of \textit{nifH} and \textit{nifB} are present, suggesting that neither M1 nor M2 are able to fix nitrogen.

Both genomes encode a functional transporter for glycine betaine, a common osmoprotectant in plants and a known precursor of TMA in the rumen (Mitchell et al. 1979). Very recently, glycine betaine was identified as a direct substrate for methanogenesis in \textit{Methanococcioides} ssp. (Watkins et al. 2014), prompting us to test this possibility with \textit{Methanomassiliicoccales}. However, no growth of \textit{M. luminyensis} with glycine betaine as electron acceptor was observed, suggesting that it is solely used as osmoprotectant. The \textit{Methanomassiliicoccales}-related, environmental clade genome M1 encodes a higher antioxidative capacity, in terms of higher diversity and redundancy of genes encoding enzymes for detoxification of reactive oxygen species compared to M2. Neither RumEn M1 nor RumEn M2 encoded a choloylglycine hydrolase gene like \textit{Ca. M. alvus}, which is involved in bile salt resistance.

Since bile salts are secreted to the small intestine of ruminants and not to the rumen (Bauman and Lock 2006), there might be no selective pressure for rumen \textit{Methanomassiliicoccales} to keep this gene. In summary, our comparative analysis of these two new \textit{Methanomassiliicoccales} genomes could confirm only some of the proposed \textit{Methanomassiliicoccales} adaptations to GIT and soil environments as suggested by Borrel et al. (2014).

**Conclusions**

This study establishes the wide distribution of a wetland-specific \textit{Methanomassiliicoccales} cluster in many peat soils across Europe. Preliminary data suggest that \textit{Methanomassiliicoccales} are present at low relative abundances among wetland methanogens, and therefore are probably less important contributors to atmospheric CH\textsubscript{4} emissions than...
their relatives residing in the GIT of animal. Compared to the GIT clade,

*Methanomassiliicoccales* of the environmental clade appear physiologically more flexible to cope with variable soil environment with its fluctuating moisture contents, temperatures and oxygen gradients. Furthermore, this study illustrates the well-known shortcomings of isolation attempts, by enriching for *Methanomassiliicoccales* that are not representative of the majority in situ. The cause for this might lie in a yet unidentified auxotrophy of the GIT clade, that has until now prevented their isolation in pure culture, despite many attempts. Nevertheless, cultivation attempts to obtain *Methanomassiliicoccales* isolates of the wetland cluster and GIT clade are indispensable to obtain a detailed understanding of their biology and habitat-specific adaptations.

**Funding**

This work was financially supported by a University of Vienna ‘Focus of Excellence’ grant to TU, by a scholarship of the University of Vienna for doctoral candidates (uni:docs) to AS and by the Austrian Science Fund (FWF, P23117-B17 to AL and FWF P25369-B22 to ChS).

**Acknowledgements**

We thank Ricardo Alves (University of Vienna, Austria) for help with phylogenetic tree constructions. We thank Michael Pester (University of Vienna, Austria) for kindly providing DNA extracts from temperate wetland samples. We thank the Viennese Zoo for providing animal fecal samples, and Morten Poulsen (University of Aarhus, Denmark), Kathrin Deckardt and Johannes Kohl (University of Veterinary Medicine, Vienna, Austria) for providing rumen fluid samples.
References


Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 2004; **32**: 11-16.


Table 1 Origin and characteristics of arctic and temperate wetland samples as reported by Tveit et al. (2012) and Steger et al. (2011).

<table>
<thead>
<tr>
<th>wetland type</th>
<th>Origin/ID</th>
<th>$pH$</th>
<th>samples/ID</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>arctic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fen</td>
<td>Knudsenheia, Norway</td>
<td>5.5 ± 0.2</td>
<td>Knudsenheia</td>
</tr>
<tr>
<td>fen</td>
<td>Solvatn, Norway</td>
<td>5.2 ± 0.1</td>
<td>Solvatn</td>
</tr>
<tr>
<td>fen</td>
<td>Roßbrand, Austria</td>
<td>4.1 ± 0.1</td>
<td>AUT 1, 2</td>
</tr>
<tr>
<td>fen</td>
<td>Roßbrand, Austria</td>
<td>4.9 ± 0.2</td>
<td>AUT 3</td>
</tr>
<tr>
<td>fen</td>
<td>Roßbrand, Austria</td>
<td>4.9 ± 0.2</td>
<td>AUT 3</td>
</tr>
<tr>
<td>fen</td>
<td>Roßbrand, Austria</td>
<td>ferrous biofilm</td>
<td>n.a.</td>
</tr>
<tr>
<td>bog</td>
<td>Große Heide, Austria</td>
<td>5.3 ± 0.1</td>
<td>AUT 4, 5</td>
</tr>
<tr>
<td>bog</td>
<td>Schremser Hochmoor, Austria</td>
<td>4.7 ± 0.4</td>
<td>AUT 6</td>
</tr>
<tr>
<td>fen</td>
<td>Schallhof, Austria</td>
<td>7.3 ± 0.1</td>
<td>AUT 8</td>
</tr>
<tr>
<td>fen</td>
<td>Berndorf, Austria</td>
<td>7.6 ± 0.0</td>
<td>AUT 9, 10</td>
</tr>
<tr>
<td>bog</td>
<td>Krähmoos, Italy</td>
<td>4.0 ± 0.1</td>
<td>IT 1</td>
</tr>
<tr>
<td>fen</td>
<td>Rasner Müser, Italy</td>
<td>4.3 ± 0.1</td>
<td>IT 2, 3</td>
</tr>
<tr>
<td>fen</td>
<td>Schlöppnerbrunnen, Germany</td>
<td>acidic</td>
<td>GER 1</td>
</tr>
</tbody>
</table>

**Table 2** Primers used in this study.

<table>
<thead>
<tr>
<th>primer</th>
<th>sequence</th>
<th>fragment length</th>
<th>target gene and taxon (reference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS1</td>
<td>CAG CAG TCG CGA AAA CTT C</td>
<td>485 nt</td>
<td>16S rRNA of MMC (Mihajlovski et al 2010)</td>
</tr>
<tr>
<td>AS2</td>
<td>AAC AAC TTC TCT CCG GCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tp-mcrA-F</td>
<td>GAY RAC ATC CTB GAR GAY TA</td>
<td>360 nt</td>
<td>mcrA of MMC (Petersen et al 2014)</td>
</tr>
<tr>
<td>Tp-mcrA-R</td>
<td>RTC GWA WCC RTA GAA TCC GAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mlas-mod For</td>
<td>GGY GGT GTM GGD TTC ACM CAR TA</td>
<td>470 nt</td>
<td>mcrA of methanogens (Steinberg and Regan 2009)</td>
</tr>
<tr>
<td>mcrA-rev</td>
<td>CGT TCA TBG CGT AGT TVG GRT AGT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3** Relative abundances (% of total methanogenic community) of methanogens in arctic peat soil from Knudsenheia (Ka, Kb, Kc) and Solvatn (Sa, Sb).

<table>
<thead>
<tr>
<th>order</th>
<th>Ka</th>
<th>Kb</th>
<th>Kc</th>
<th>Sa</th>
<th>Sb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methanomassiliicoccales</td>
<td>4.7</td>
<td>0.5</td>
<td>0.7</td>
<td>0.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Methanobacteriales</td>
<td>66.3</td>
<td>3.1</td>
<td>3.4</td>
<td>10.0</td>
<td>2.6</td>
</tr>
<tr>
<td>Methanosarcinales</td>
<td>20.9</td>
<td>84.7</td>
<td>87.0</td>
<td>70.0</td>
<td>87.9</td>
</tr>
<tr>
<td>Methanomicrobiales</td>
<td>8.1</td>
<td>11.7</td>
<td>8.9</td>
<td>20.0</td>
<td>7.9</td>
</tr>
<tr>
<td>Methanocellales</td>
<td>2.3</td>
<td>2.7</td>
<td>11.6</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Table 4 Characteristics of the two partial rumen MMC genomes, RumEn M1 (182 contigs) and RumEn M2 (18 contigs) and genomes of other members of MMC\textsuperscript{a}. The dashed line separates the environmental clade and the GIT clade MMC.

<table>
<thead>
<tr>
<th>genome</th>
<th>accession no.</th>
<th>G+C content (mol%)</th>
<th>size [Mbp]</th>
<th>tRNA genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>RumEn M1</td>
<td>LJKK00000000</td>
<td>62.5</td>
<td>~2.21</td>
<td>43</td>
</tr>
<tr>
<td>\textit{M. luminyensis}\textsuperscript{b}</td>
<td>CAJE01000001-26</td>
<td>60.5</td>
<td>&gt;2.62</td>
<td>43</td>
</tr>
<tr>
<td>Ca. M. intestinalis\textsuperscript{b}</td>
<td>CP005934</td>
<td>41.3</td>
<td>1.93</td>
<td>46</td>
</tr>
<tr>
<td>RumEn M2</td>
<td>LJKL00000000</td>
<td>54.8</td>
<td>~1.28</td>
<td>44</td>
</tr>
<tr>
<td>Ca. M. termitum\textsuperscript{c}</td>
<td>CP010070</td>
<td>49.2</td>
<td>1.49</td>
<td>46</td>
</tr>
<tr>
<td>Ca. M. alvus</td>
<td>CP004049</td>
<td>55.6</td>
<td>1.67</td>
<td>45</td>
</tr>
<tr>
<td>BRNA1\textsuperscript{d}</td>
<td>CP002916</td>
<td>58.3</td>
<td>1.46</td>
<td>44</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Data are derived from the latest MMC comparative genome analysis (Lang et al 2014) and GenBank.

\textsuperscript{b} Closest relatives of M1; based on a nearly full length 16S rRNA gene sequence (1467 nt), with 96\% 16S rRNA sequence identity to both \textit{Methanomassiliicoccus} species.

\textsuperscript{c} Closest relative of M2; based on a 16S rRNA gene sequence (1272 nt), with 95\% 16S rRNA sequence identity to \textit{Ca. M. termitum}.

\textsuperscript{d} BRNA1 was also obtained from bovine rumen but is only distantly related to M1 and M2, with 87\% and 92\% 16S rRNA sequence identity, respectively. Number of tRNAs were obtained as for M1 and M2.
**Figure 1.** Phylogenetic trees showing the relationship among 16S rRNA gene and McrA protein sequences obtained from animal fecal samples, bovine rumen fluid and wetland soils. Operational taxonomic units (OTUs) with <98 % sequence identity and sequences derived from rumen *Methanomassiliicoccales* genomes (RumEn M1 and RumEn M2) are indicated in bold fonts. The positions of 16S rRNA gene and McrA protein sequences obtained from the same organism or enrichment cultures are shown in the same colour. Maximum likelihood trees of near full-length references sequences were calculated with ARB. The partial *Methanomassiliicoccales* 16S rRNA gene and McrA protein sequences were placed into these
trees using the maximum parsimony tool implemented in ARB. Numbers in brackets refer to clone numbers obtained from the same sample. AUT, Austrian; DK, Danish; RF, rumen fluid; GER, German; IT, Italian. Substitutions per site are indicated with scale bars.
**Figure 2.** Distribution of environmental and GIT clade *Methanomassiliicoccales* 16S rRNA gene clones in animal and wetland samples. Colour code indicates the percentages of 16S rRNA gene clones belonging to one of these two *Methanomassiliicoccales* family-level clades or to related *Thermoplasmata*.
Figure 3. Enrichment of *Methanomassiliicoccales* clades from bovine rumen fluid. Panel (A) community composition of rumen fluid used for initial inoculation of *Methanomassiliicoccales* enrichment cultures. Left column: relative abundance of bacterial and archaeal 16S rRNA gene copy numbers obtained from quantitative PCR assays (see materials and methods for details). Central column: relative abundance of *Methanomassiliicoccales* and other methanogenic archaea obtained from quantitative PCR assays. Right column: community composition of *Methanomassiliicoccales* derived from 16S rRNA gene clone library (AUT RF; 26 clones), showing the dominance of the GIT clade. Panel (B) community composition in metagenome RumEn_MG1 based on the classification of 234 16S rRNA gene fragments. Panel (C) community composition in metagenome RumEn_MG2 based on the classification of 1360 16S rRNA gene fragments. The left
columns of (B) and (C) show the 16S rRNA gene fragment abundances of archaeal and bacterial orders. The right columns of (B) and (C) depict the relative abundance of 16S rRNA gene fragments of the *Methanomassiliicoccales* GIT and environmental clades.