Aquat Sci (2015) 77:481–497
DOI 10.1007/s00027-015-0392-9

Aquatic Sciences

RESEARCH ARTICLE

Declining diversity of egg-associated bacteria during development of naturally spawned whitefish embryos (*Coregonus* spp.)

Laetitia G. E. Wilkins[1] · Aude Rogivue[1,2] · Luca Fumagalli[1] · Claus Wedekind[1]

Received: 28 May 2014 / Accepted: 17 March 2015 / Published online: 24 March 2015

Abstract Fish eggs are associated with microbes, whose roles range from mutualism to parasitism. Recent laboratory experiments have shown that the taxonomic composition of associated microbial communities on the egg influences embryonic development. Host genetics also plays an important role in determining the consequences for embryonic growth and survival in this interaction. Moreover, it has been found that the importance of host genetics increases during embryogenesis. These findings suggest that during embryogenesis, the host increasingly influences the composition of its associated microbial community. However, little is known about the composition of microbial communities associated with naturally spawned eggs in the wild. We sampled fertilized whitefish eggs (*Coregonus* spp.) of different developmental stages from six sub-Alpine lakes and used a universal primer pair and 454 pyrosequencing in order to describe the taxonomic composition of egg-associated bacterial communities. We found bacterial communities on early embryos to be very diverse and to resemble the bacterial composition of the surrounding water environment. The bacterial communities on late embryos were significantly less diverse than on early embryos and displayed a clear shift in taxonomic composition that corresponded poorly with the bacterial composition of the surrounding water environment. The main bacterial components on whitefish eggs in this study were *Proteobacteria*, *Actinobacteria*, and *Firmicutes*, while the five most common families were *Leuconostocaceae*, *Streptococcaceae*, *Comamonadaceae*, *Oxalobacteraceae* and *Moraxellaceae*. Their putative relationships with the host are discussed. We conclude that natural symbiotic bacterial communities become more specialized during embryogenesis because of specific interactions with their embryo host.

Keywords Bacterial community · *Coregonus* · Metagenomics · Microbiome · Salmonids · Whitefish

Introduction

Animals and plants are associated with microbes that typically outnumber host cells (Douglas 1994; Rosenberg et al. 2007). Various kinds of interactions between the host and its associated microbes are possible (Casadevall and Pirofski 2000), including: (i) commensalistic microbes that profit from their host while not harming it (Hooper and Gordon 2001), (ii) mutualistic symbionts where both host and microbes profit from the association with each other (Casadevall and Pirofski 2000), (iii) opportunistic microbes where commensalistic or mutualistic symbionts can turn into pathogens and harm the host depending on environmental conditions and the host's immune defense, and (iv) pathogenic microbes that harm the host and lead to infectious diseases (Casadevall and Pirofski 1999). Host-associated microbiota can therefore play important roles in the growth, development, and phenotypic adaptation of their hosts (Wisz et al. 2013). Various types of animals have been used to study

L. Fumagalli and C. Wedekind are joint senior authors.

Electronic supplementary material The online version of this article (doi:10.1007/s00027-015-0392-9) contains supplementary material, which is available to authorized users.

✉ Laetitia G. E. Wilkins
laetitia.wilkins@unil.ch

[1] Department of Ecology and Evolution, Biophore, University of Lausanne, 1015 Lausanne, Switzerland

[2] WSL Swiss Federal Research Institute, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

symbiotic interactions in the wild, such as insects (Dillon and Dillon 2004; Vasanthakumar et al. 2008), mice (*Mus musculus*) (Schloss et al. 2012; Pantoja-Feliciano et al. 2013), and cattle (*Bos primigenius*) (Hooper et al. 2002). These studies provided important insights into various biological contexts, including, for example, host invasion (Schloss et al. 2012; Pantoja-Feliciano et al. 2013), the evolution of new traits (Ley et al. 2008), speciation (Sevellec et al. 2014), biocatalyzation processes for biomass deconstruction, and pest management (Dillon and Dillon 2004; Shi et al. 2010). Here, we focus on a question that has received comparatively little attention, namely whether and how early host development affects the composition of associated bacterial communities under natural conditions. We used naturally spawned whitefish eggs (*Coregonus* spp.) of different origin and different developmental stages as our model.

Metagenomics is the study of genetic material recovered directly from environmental samples (Schloss and Handelsman 2005). Traditionally, molecular approaches such as denaturing gradient gel electrophoresis (DGGE), restriction fragment length polymorphism (RFLP) and fluorescence in situ hybridization (FISH) have been applied in metagenomic studies to investigate the general composition, diversity and dynamics of symbiotic microbiota (Shi et al. 2010). The field of metagenomics benefited from recent developments in next-generation sequencing techniques and advances in bioinformatics approaches (Xu 2006). Methods are now available that allow characterizing very high numbers of bacteria simultaneously and thereby greatly advancing the study of microbial communities and their interactions, including non-culturable microbes (Schloss and Handelsman 2005). Next-generation sequencing techniques no longer limit the number of characterized host-associated microbes and can be used to estimate relative abundances of different taxa (Lozupone and Knight 2005). This approach has been used, for example, to describe patterns of microbial community composition on different substrates (Lauber et al. 2009) or different host species (Ley et al. 2008; Johnson et al. 2012; Brucker and Bordenstein 2013). Microbial communities turned out to be unexpectedly diverse, even on a microscale, e.g., on different body parts of the same host (e.g., Ley et al. 2008; Fierer et al. 2010; Salonen et al. 2010; Ravel et al. 2011). Bacterial symbionts often seem to co-diversify with their hosts (Nicholson et al. 2012). On fish, metagenomic characterizations of microbial communities in guts (Mouchet et al. 2012; Ye et al. 2013), on skin (Larsen et al. 2013), or on kidneys (Boutin et al. 2012; Sevellec et al. 2014) have been used in the study of piscine taxonomic relationships.

Whitefish (*Coregonus* spp.) are typically keystone species in their respective freshwater ecosystems and can be of significant economic value (Nusslé et al. 2009). Despite the relative simplicity with which salmonid eggs can be studied under controlled laboratory conditions (Finn 2007; Wedekind et al. 2007; von Siebenthal et al. 2009), existing understanding of their associated microbial communities has been patchy at best. The chorion of fish eggs (i.e., the acellular coat surrounding the mature egg) has been shown to represent a substrate where bacteria thrive (Wedekind 2002; Treasurer et al. 2005; Wedekind et al. 2010). Eggs of many fishes, including whitefish, are spawned over some type of substrate to remain there immobile and in principle open for colonization by microbes. However, typically they are equipped with general antimicrobial compounds that the mother allocated to the eggs before spawning (Hanif et al. 2004; Løvoll et al. 2006; Finn 2007). Sensitive life-history responses of the host to changes in its symbiotic microbial community composition have been revealed during specific experimental exposure of whitefish embryos to *Pseudomonas fluorescens* (von Siebenthal et al. 2009; Clark et al. 2014) and *Saprolegnia ferax* (Clark and Wedekind 2011), as well as during non-specific experimental changes of microbial communities on whitefish eggs at varying nutrient broth concentrations (Wedekind et al. 2010). These life-history responses include changes in hatching time (Clark and Wedekind 2011; Clark et al. 2013), growth and survival (Clark et al. 2014). Some of the effects of host genetics could be linked to the embryo's MHC (major histocompatibility complex), i.e., its allelic diversity (Wedekind et al. 2004) and its expression patterns (Clark et al. 2013). While maternal effects are important during early developmental stages of the embryo, there is a transition to the importance of paternal effects during later stages of development (Clark et al. 2014). This transition reveals the increasing importance of the embryo's own genetics because paternal effects in species with external fertilization and without parental care (such as whitefish) can only be genetic (Aykanat et al. 2012).

We sampled naturally fertilized whitefish eggs from different sub-Alpine lakes to compare bacterial communities and to test for effects of host developmental stage on bacterial community composition in the wild. In order to characterize bacterial communities, a highly informative gene segment of the 16S rRNA gene was used. The 16S rRNA is a component of the 30S small subunit of prokaryotic ribosomes (Amann and Ludwig 2000). The genes coding for it are referred to as 16S rDNA and are used for phylogenetic studies of bacteria, as they provide highly conserved primer binding sites and hyper-variable regions that provide species-specific signature sequences (Huse et al. 2008). Standard databases exist for the classification and identification of 16S rDNA genes (Cole et al. 2005). To control for environmental effects on bacterial community composition and to separate them from specific host effects, we used water samples collected simultaneously with egg samples. Freshwater microbial community composition has been shown to vary with environmental factors such as

temperature, pH, or environmental stability (Martin et al. 2012; Shade et al. 2012). Applying the statistical procedure of re-sampling with replacement, we investigated whether fish egg samples are associated with bacterial species that represent a random sample of the bacterial species found in the surrounding water or whether there were specific host effects shaping their associated bacterial communities. For this statistical procedure, bacterial sequences were sampled with replacement from the water samples in our dataset in order to simulate bacterial communities of fish egg samples. These simulated communities were then compared to the observed egg-associated bacterial communities.

Material and methods

Sample acquisition

Naturally spawned whitefish eggs (*Coregonus* spp.) were collected from Lake Geneva, Lake Neuchâtel, Lake Biel, Lake Thun, Lake Lucerne, and Lake Hallwyl (all sub-Alpine lakes of Switzerland; Fig. 1) at depths of 3–12 m by scuba diving. They were sampled during the spawning seasons of the respective whitefish populations (13.12.2010–27.12.2010) and some weeks later (11.1.2011–28.2.2011) in order to obtain early embryos (blastoderm visible atop the yolk and oil droplets) and late embryos (late-eyed stage, i.e., developmental stages 3 and 6 in Luczynski and Kirklewska (1984), (Table 1; Fig. 1). At Lake Geneva, Lake Neuchâtel and Lake Biel, only dead eggs could be found at the second time point. Dead eggs were not used for further analysis. At Lake Hallwyl, eggs were not collected by scuba diving, but rather with a dredge as in Müller and Stadelmann (2004); therefore, only eggs at the late stage were sampled.

After collection, eggs were immediately rinsed with 2 L of autoclaved and filtered (0.2 µm, Millipore, Zug, Switzerland) water. Developmental stage was determined on a sterile glass slide under a field light microscope (Motic Microscopes 1820, Wetzlar, Germany). Eggs were then frozen in liquid nitrogen and later stored at −80 °C without a storage buffer. For all but one sampling event (Table 1), 1.5 L water was

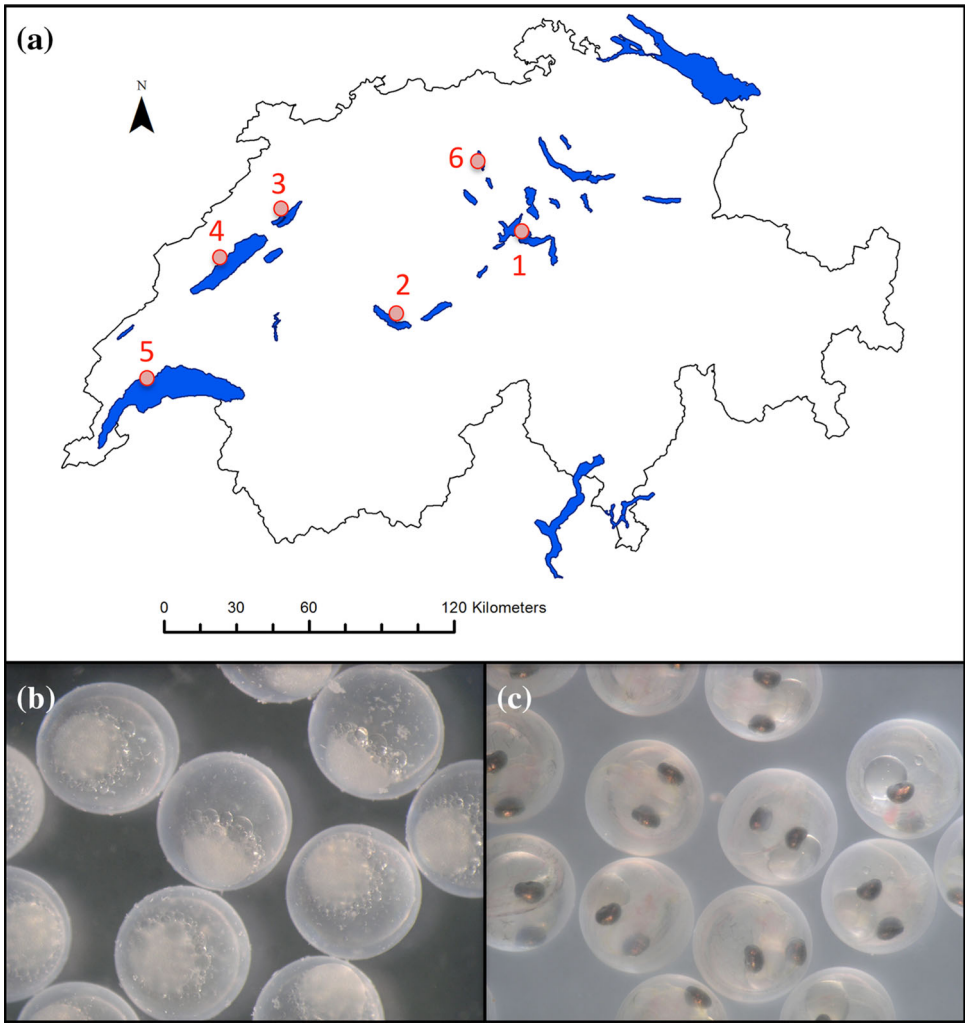

Fig. 1 Metagenomic sampling **a** Water and whitefish egg sampling locations within Switzerland (*1* Lake Lucerne, *2* Lake Thun, *3* Lake Biel, *4* Lake Neuchâtel, *5* Lake Geneva, *6* Lake Hallwyl). Fertilized whitefish eggs were sampled at **b** 'early stage' = blastoderm visible atop the oil droplets, or **c** 'late stage' = late-eyed stage

Table 1 Locations for sampling of water and naturally spawned eggs used for 454 pyrosequencing

Location[a]	GPS coordinates	Date and embryonic stage[b]	Depth (m)	Barcode[c]
Lake Lucerne	47°01′43.08″	21.12.10	3	TGTACTACTC
Coregonus suidteri	8°24′47.62″	(Early)		TAGAGACGAG
Lake Lucerne	47°01′43.08″	14.01.11	3	TCGTCGCTCG
C. suidteri	8°24′47.62″	(Late)		TGGTCGCTCG
Lake Thun	46°41′05.57″	13.12.10	5	TCTACGTAGC
C. fatioi	7°44′52.24″	(Early)		CGTAGACTAG
Lake Thun	46°41′05.57″	11.01.11	5	TACTGAGCTA
C. fatioi	7°44′52.24″	(Late)		ACATACGCGT
Lake Biel	47°04′37.20″	22.12.10	4	ATACGACGTA
C. palaea	7°08′40.13″	(Early)		TACTCTCGTG
Lake Neuchâtel[d]	46°50′10.08″	11.12.10	3	TCACGTACTA
C. palaea	6°42′14.15″	(Early)		–[d]
Lake Léman	46°30′32.44″	27.12.10	3	CGTCTAGTAC
C. palaea	6°32′59,41″	(Early)		TACGAGTATG
Lake Hallwyl	47°17′03.95″	28.02.11	12	CGAGAGATAC
C. suidteri	8°13′38.06″	(Late)		ACTGTACAGT

[a] Species names according to Kottelat and Freyhof (2007)

[b] Developmental stage of the embryo was either early (Fig. 1b) or late (Fig. 1c)

[c] Barcode for individual tagging of bacterial communities during 454 pyrosequencing: barcode of fish egg sample followed by barcode of corresponding water sample

[d] No water sample

collected in a sterile glass bottle, as close as possible to the location where eggs had been found. This water was filtered (0.2 µm, Millipore, Zug, Switzerland), and the filters were stored at −80 °C (without storage buffer).

Molecular genetic analyses

Thirty-two eggs per spawning location, i.e., 256 eggs in total, were randomly sampled for individual DNA extraction for each time point (Table 1). Each egg was homogenized in a 2-mL Eppendorf tube (Sarstedt, Nümbrecht, Germany) in 1.6 mL of buffer ASL (Qiagen, Hombrechtikon, Switzerland) with a mixer mill (MM300; Retsch, Düsseldorf, Germany) for 2 × 30 s using six tungsten beads (3 mm, Qiagen), five silica beads (1.5 mm, Qiagen), and 0.4 g silica powder (0.2 mm, Qiagen). The homogenate was heated to 95 °C for 5 min, then vortexed again for 15 s and centrifuged at 13,000 rpm for 1 min. An InhibitEX® tablet (Qiagen) was added to each supernatant in order to digest DNA extraction inhibitors (lipids and proteins from the egg); then, the sample was centrifuged again at 13,000 rpm for 3 min. The new supernatant was treated with 25 µL proteinase K (50 µg/ml, Qiagen) per sample and incubated at 70 °C for 10 min. This mix was purified with absolute ethanol and subjected to the QIAamp DNA micro kit (Qiagen) following the manufacturer's protocol. DNA was eluted in 10 µL of DNase-free water (Millipore, Zug, Switzerland). The same protocol was used for water filters, with the extra step of adding carrier RNA (supplied in the kit).

PCR was performed with a bacterial-specific primer pair, 27F and 338R (Hamady et al. 2008) that amplifies a 311-bp (base pairs) fragment of the V1-V2 hypervariable region of bacterial 16S rRNA. These primers have widely been used in metagenomic studies and therefore were chosen in an attempt to make our study comparable to other microbial characterizations of different habitats. The 338R primer included a unique 10-bp sequence tag to barcode each sample (Table 1). The barcodes had been tested in silico and empirically to show that they do not bias amplification (Hamady et al. 2008; Berry et al. 2011). Thirty-two individual bacterial DNA extractions were pooled for every time point and location, resulting in eight PCR reactions for egg samples, seven PCR reactions for water samples, and one PCR reaction for the negative control using all reagents except the DNA extraction (Table 1).

Each PCR was performed in a total volume of 25 µL consisting of 1–50 ng bacterial genomic DNA, 2.5 µL of 10× PCR buffer, 400 µM of each dNTP, 2.5 mM of $MgCl_2$, 0.6 µM of each primer, and 0.625 U of *Taq* polymerase (Life Technologies, Zug, Switzerland). The thermal profile was modified from that of Berry et al. (2011) and consisted of a two-step PCR. Step I with primer pair 27F and 338R (without tags): 3 min at 94 °C; 25 cycles of 94 °C for 30 s, 56 °C for 30 s, 72 °C for 30 s, and a final extension at 72 °C for 10 min. Step II: 1 µL of the product of step I, 2.5 µL of 10× PCR buffer, 400 µM of each dNTP, 2.5 mM of $MgCl_2$, 0.6 µM of the primers 27F and individually tagged 338R (Table 1), and 0.625 U of *Taq* polymerase (Life Technologies) in a total volume of 50 µL (1:50 dilution). The following conditions were used for step II: 3 min at 94 °C, 5 cycles of 94 °C for 30 s, 56 °C for 30 s, 72 °C for 30 s, and a final extension at 72 °C for 10 min. Amplified products were purified with the Wizard® SV Gel and PCR Clean-Up System (Promega, Dübendorf, Switzerland). For every sample, five PCR reactions were pooled from PCR step I into PCR step II to avoid amplification bias (Berry et al. 2011).

Cleaned PCR products were run on an agarose gel (1.5 %, 100 V, 45 min), cut out, purified, and quantified using a Qubit ® 2.0 Fluorometer (Life Technologies) and on a bioanalyzer using the DNA 1000 kit (Agilent, Morges, Switzerland). Equimolar PCR reactions of 15 ng/µL DNA were pooled in equimolar amounts in a 10 mM Tris–HCl buffer at pH 8.5 and sent for 454 pyrosequencing (Microsynth, Balgach, Switzerland) on a Genome Sequencer FLX System (Roche, Basel, Switzerland).

Prior to 454 pyrosequencing, 10 randomly selected samples (all egg samples and two water samples from Lake

Lucerne) were cloned to confirm the amplification of only bacterial 16S sequences and to ensure the incorporation of the 10-bp sequence tag and adaptor sequences to the primer pair 27F and 338R. Cloning was performed with the TOPO TA Cloning® Kit (Life Technologies) using the pCR®2.1-TOPO vector and One Shot ®TOP10 chemically competent cells. Twenty positive clones per individual were selected and amplified with M13 forward and reverse primers (Life Technologies) with the following amplification profile: 94 °C for 2 min, followed by 33 cycles of 94 °C for 15 s, 54 °C for 15 s, 72 °C for 15 s, and a final 10 min extension at 72 °C. Reactions were performed in a total volume of 15 µL and contained: 10 µM of each primer, 1.5 µL of 10× PCR buffer, 0.5 µM of each dNTP, 0.08 µL *Taq* DNA polymerase, 10 ng of cloned DNA, and water. Blank PCR reactions were included as controls. Amplified products were purified with the Wizard® SV Gel and PCR Clean-Up System (Promega) and sequenced in the reverse direction with M13 on an ABI Prism 3100 genetic analyzer (Applied Biosystems, Life Technologies).

Quality control

All steps of the 454 pyrosequencing data quality control were done in the QIIME framework (Caporaso et al. 2010b). Raw reads were split based on their 10-bp barcode. Forward and reverse primers were removed, and sequences produced by sequencing errors were filtered out (Pantoja-Feliciano et al. 2013). To pass, a sequence read needed to: (i) have a perfect match to the barcode and the 16S rRNA primer, (ii) be at least 300-bp long, (iii) have no more than two undetermined bases, and (iv) have at least a 60 % match to a 16S rRNA gene sequence from the Greengenes database (Ong et al. 2013). The resulting clean reads were screened for chimeras using the UCHIME algorithm (Edgar 2010). As a last step of quality control, all sequences were removed that had been detected in the blank control sample of only beads and buffers. All resulting reads (Table 2, Supplementary Tables S1 and S2) were used for further analysis.

Bacterial community comparison

For statistical analysis of bacterial communities, operational taxonomic units (OTUs) were defined using Usearch algorithm 5.2.236 (Edgar 2010). This algorithm assigns similar sequences to OTUs by clustering them based on a user-defined similarity threshold [sequence similarity was set to 0.97, roughly corresponding to species-level OTUs (Hughes et al. 2002)]. In every OTU cluster, the most abundant sequence was chosen and then assigned to a reference using an open reference BLAST search (Lan et al. 2012) with RDP classifier 2.2 (Wang et al. 2007) and Greengenes reference database version

Table 2 Number of barcoded bacterial DNA sequences before and after quality control

Location	Sampling time[a]	Type of sample	Reads (454)	Denoising	Bead-filtering	OTUs
Lake Biel	Early	Embryo	6798	6283	5911	714
Lake Biel	Early	Water	6996	6396	6198	489
Lake Geneva	Early	Embryo	6287	5855	5467	725
Lake Geneva	Early	Water	6824	6285	6083	476
Lake Hallwyl	Late	Embryo	4301	3991	3939	160
Lake Hallwyl	Late	Water	7199	6672	6590	336
Lake Lucerne	Early	Embryo	6289	5798	5540	789
Lake Lucerne	Early	Water	5758	5270	5097	345
Lake Lucerne	Late	Embryo	6514	6092	5783	532
Lake Lucerne	Late	Water	3719	3396	3343	340
Lake Neuchâtel	Early	Embryo	4919	4474	4141	726
Lake Thun	Early	Embryo	5608	5246	5133	892
Lake Thun	Early	Water	5028	4653	4592	288
Lake Thun	Late	Embryo	6070	5671	5295	608
Lake Thun	Late	Water	4137	3712	3652	286
		Beads only	5260	4837	0 (16)[b]	0

'Denoising' gives the number of reads after exclusion of PCR-artifacts and chimeras, while 'Bead-Filtering' gives the number of reads remaining after filtering out all DNA sequences that had been found in the negative control sample. The number of different Operational Taxonomic Units 'OTUs' is given that were assigned by RDP classifier 2.2 (Wang et al. 2007) and the Greengenes reference database version 12.10 (McDonald et al. 2012)

[a] State of embryonic development at time of sampling

[b] 16 different OTUs were found in the negative control, they were filtered out in all other samples

12.10 (McDonald et al. 2012). This approach enables the identification of previously undescribed bacterial sequences in a dataset.

The following alpha diversity measures for each sample were calculated: ACE, Chao1, Fisher's Diversity Index, Inverse Simpson Diversity Index, number of observed species, phylogenetic distance, Shannon Diversity Index, and Simpson Diversity Index (Gotelli and Chao 2013). Alpha diversities were visualized using rarefaction curves and boxplots. Only rarefied measures, i.e., diversity measures reaching saturation, were used for downstream analysis, as sampling effort can have a great influence on the values (Hill et al. 2003).

Types of bacterial communities were compared based on their alpha diversity matrices and on phylogenetic assignments. UniFrac distances (Lozupone and Knight 2005; Lozupone et al. 2011) were calculated to depict differences among samples (Legendre and Anderson 1999; Anderson and Willis 2003). UniFrac quantifies the fraction of unique branch lengths against the total branch length between pairs of communities from a common phylogenetic dendrogram, giving an estimate of the phylogenetic distance between those communities (Lozupone and Knight 2005). To build a phylogenetic dendrogram, OTUs were aligned using the PyNAST algorithm (Caporaso et al. 2010a), and a phylogenetic tree was built using FastTree (Price et al. 2008, using default parameters). Quantitative measures (weighted UniFrac) are ideally suited for revealing community differences that are due to changes in relative taxon abundance, e.g., when a particular set of taxa flourish, e.g., because a limiting nutrient source becomes abundant (Lozupone and Knight 2007). Qualitative measures (unweighted UniFrac) are most informative when communities differ primarily by what species can live in them, e.g., at high temperatures, in part because abundance information can obscure significant patterns of variation in which taxa are present or absent. Accordingly, unweighted UniFrac distances were applied for the downstream analysis using only OTUs' presence or absence information.

Analyses of variance were performed in the R environment version 3.0.1 (2011) in order to investigate the effect of developmental stage on bacterial composition on eggs, using the 'vegan' package (Oksanen et al. 2013) with 'adonis' (Anderson 2001; McArdle and Anderson 2001) and 'ANOSIM' (Warton et al. 2012). These analyses of variance methods are non-parametric, but they assume equal variances among groups of samples (Anderson 2001, 2006; Warton et al. 2012). The R package 'Permdisp' (Anderson 2006) was applied to test for equal variances among groups, and the R package 'phyloseq' (McMurdie and Holmes 2013) was used to read the output from QIIME into R.

Adonis is a nonparametric statistical method that takes the UniFrac distance matrix among samples and a categorical variable that groups the samples to estimate the percentage of variation (R^2) explained by the supplied categorical variable used for grouping. It creates a set by first identifying the relevant centroids of each group and then calculating the squared deviations from these points. Significance tests are performed using *F*-tests based on sequential sums of squares from permutations of all samples irrespective of the grouping. ANOSIM is another nonparametric statistical method that tests whether two or more groups of samples are significantly different from each other. It differs from adonis in how R^2s are calculated (Fierer et al. 2010). Boxplots were used for a visual comparison of UniFrac distances between developmental stages. Multiple student's two-sample *t*-tests of pairwise means (Hollander and Wolfe 1973) were calculated for all pairs of boxplots to determine whether means were significantly different from each other. Within every comparison, a Bonferroni correction was applied to correct for multiple testing [number of tests within each comparison = 5: all within-stage, all between-stage, early stage vs. early stage, late stage vs. late stage, and early stage vs. late stage, analogous to Costello et al. (2009)]. Conservative non-parametric *p*-values using 999 Monte Carlo permutations of the raw data are reported instead of parametric *p*-values from a *t*-distribution. The same tests were applied to test whether the number of unique OTUs was different between bacterial communities of egg samples at an early stage and bacterial communities of egg samples at a late stage. Analogous analyses were applied to water samples to control for environmental effects on bacterial community composition.

Because the bacterial communities in this study included numerous OTUs, an exhausting list of the bacterial taxa on whitefish egg samples and corresponding water samples was delegated to the Supplementary Material (Table S1). In order to elaborate on the implications that these bacteria might have on their host, the ten most common OTUs that were found on early whitefish embryos only, on late whitefish embryos only, and at both stages are discussed in further detail. For this analysis, bacterial sequences needed to result in an unambiguous match to a reference with RDP classifier 2.2 and Greengenes reference database version 12.10 at the family, genus or species level. Bacterial sequences from the blank control sample of only beads and buffers were not removed. Information on the ecology of these bacteria was derived from a search in Web Of Knowledge™ v. 5.12 (Thomson Reuters) or according to Austin and Austin (2007) and Madigan et al. (2010).

Eggs at an early stage had been collected at five locations, whereas eggs at a late stage had been collected at only three locations. Tests investigating the effect of sample size (number of locations) are described in the Supplementary Material. Permutation tests were used to test whether bacteria associated with eggs constituted a

random sample of the bacteria in the surrounding water environment. Permutation tests were conducted as follows: One thousand bacterial communities of egg samples were simulated for each sampling site in order to obtain a density distribution of its bacterial community composition and corresponding confidence intervals. For the simulation of egg-associated bacterial communities, bacterial sequences were randomly sampled with replacement from the corresponding water sample. The same number of sequences that had been found in the true egg sample was sampled from the corresponding water sample to simulate bacterial communities of an egg sample. This was done 1000 times for each site and time point. The probability of obtaining the observed UniFrac distance between bacterial communities of water and egg samples in the simulations was calculated. UniFrac distances between bacterial communities of the respective egg and water sample at a given location were compared to the simulated mean UniFrac distances between a given water sample and a simulated egg sample. Differences in bacterial communities between whitefish egg samples and corresponding water samples were also investigated using adonis, ANOSIM, and two-sample *t*-tests, analogous to the analyses of effects due to developmental stage in egg samples and water samples, respectively. The relationship between bacterial alpha diversities of whitefish egg samples and corresponding water samples using Pearson's product moment correlations is reported in the Supplementary Material.

Results

Quality control

Eighty-four percent of the raw bacterial sequence reads (76,764 of in total 91,707; Table 2) were retained after different steps of quality control for downstream analysis, leaving a mean of 5118 (standard deviation = 952) sequences per sample that were used further. After searching for sequences against the Greengenes database, a mean of 513 OTUs (SD = 220) could be assigned to each sample (Table 2). On whitefish egg samples, 2194 different OTUs could be differentiated (after quality control, Tables S1). Bacterial sequences that were found in the negative control are reported in the Supplementary Table S2. The majority of genera identified on whitefish eggs in this study were bacteria representing *Proteobacteria*, *Actinobacteria*, and *Firmicutes*. The five most common families were *Leuconostocaceae*, *Streptococcaceae*, *Comamonadaceae*, *Oxalobacteraceae* and *Moraxellaceae*. Rarefaction curves for the majority of samples using four different alpha diversity matrices (Chao1, number of observed species, phylogenetic distance, and Shannon index) did not approach saturation (Fig. 2), even with nearly 4000 reads. Cloning resulted in 200 sequenced reads of bacterial origin only. Cloning confirmed that all adaptors had been correctly incorporated, read and interpreted (results not shown).

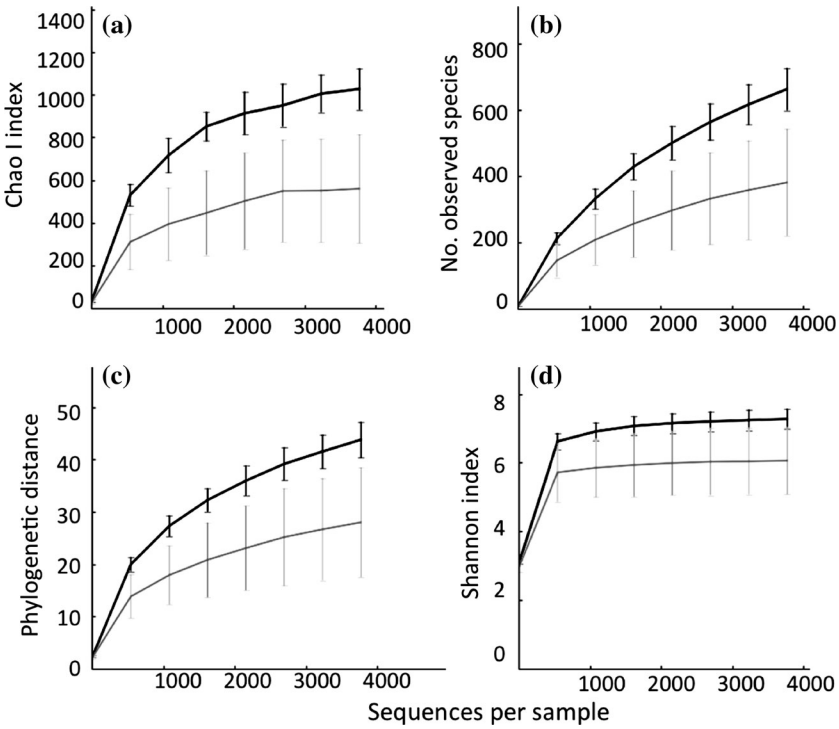

Fig. 2 Rarefaction curves for alpha diversity measures of bacterial communities. Rarified mean alpha diversity (±1 SD) for bacterial communities on whitefish eggs at an early stage (blastoderm visible atop the oil droplets; *thick line*) and whitefish eggs at a late stage (late-eyed stage; *thin line*). The alpha diversities measured are **a** Chao 1 index, **b** number of observed species, **c** phylogenetic distance, and **d** Shannon index

Effect of developmental stage

To test for an effect of early and late developmental stage on bacterial community composition on egg samples, non-parametric testing was applied with adonis ($p = 0.018$, $R^2 = 30.1$ %) and ANOSIM ($p = 0.019$, $R^2 = 69.2$ %). Difference in mean bacterial diversity was visualized using boxplots (Supplementary Fig. S1a), and tests of significant pairwise differences among means were performed using a two-sided Student's two-sample t test of the two means ($t = 2.9$, d.f. $= 5$, $p = 0.018$). As a control, the same analysis was applied to water samples: with adonis ($p = 0.67$, $R^2 = 31.2$ %) and ANOSIM ($p = 1.0$, $R^2 = 25.0$ %; Supplementary Fig. S1b). Two-sided Student's two-sample t-test of the two means was non-significant ($t = 0.21$, d.f. $= 4$, nonparametric $p = 0.84$). Notably, whitefish eggs at an early stage showed a significantly higher alpha diversity of associated bacteria than whitefish eggs at a late stage consistently across all seven alpha diversity measures (Fig. 3). This decrease in bacterial diversity is also reflected in the significantly different number of unique OTUs at the two stages (only early n = 809, only late n = 175, shared = 750, d.f. $= 5$, $p < 0.001$; Supplementary Table S3) that did not change by splitting the data into different subsets of equal sample size (Supplementary Table S3).

To test whether the assumptions of equal variances for non-parametric testing with adonis and ANOSIM were met, Permdisp was applied to the following groupings of data: Whitefish dataset testing the effect of developmental stage ($F = 0.21$, d.f. $= 5$, $p = 0.67$) and water samples testing for an effect of developmental stage ($F = 0.68$, d.f. $= 4$, $p = 0.46$). There were no deviations from the assumption of equal variances.

The ten most frequent bacterial OTUs that were found on whitefish eggs at the early stage only accounted for 23 % of total bacterial sequences within this group (Fig. 4). Only eight sequences resulted in an unambiguous match to a reference with RDP classifier 2.2 and Greengenes reference database version 12.10 at the family, genus or species level, which accounts for 20.1 % of total bacterial sequences within this group. At the late stage, the ten most frequent bacterial OTUs accounted for 64 % of total bacteria within this group, and the ten most frequent bacterial OTUs that were found at both stages accounted for 35 %. Presumable significance of these most common bacterial species on whitefish embryos is further discussed below.

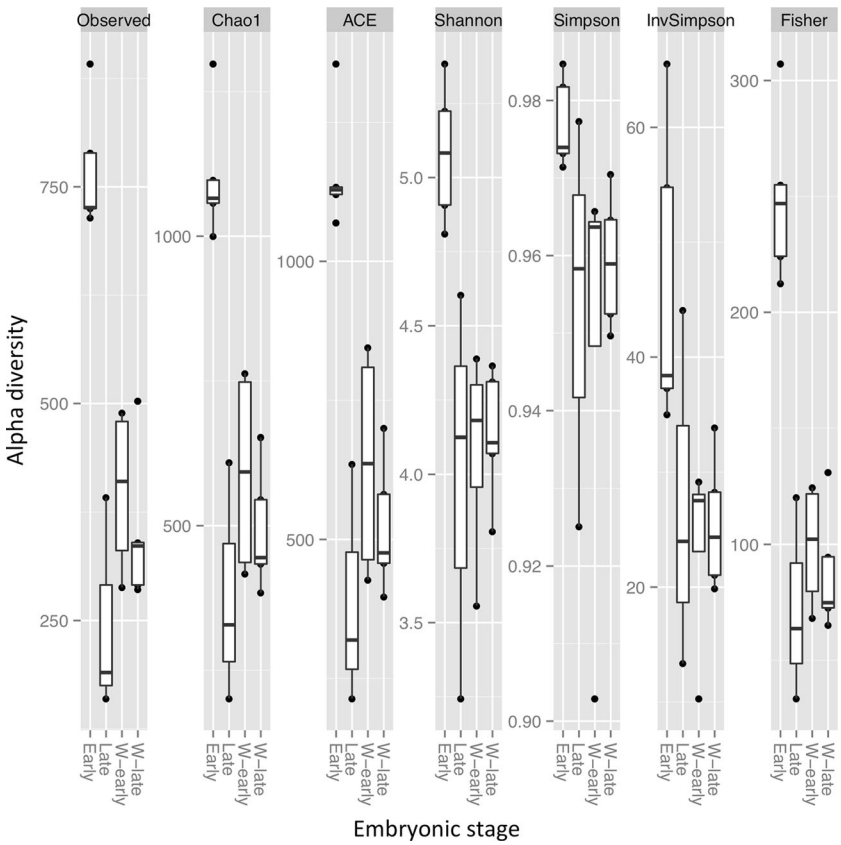

Fig. 3 Alpha diversity measures of bacterial communities on different types of samples. Seven alpha diversity measures [number of observed species, Chao 1, ACE, Shannon diversity index, Simpson diversity index, inverse Simpson diversity index, and Fisher's diversity index (Gotelli and Chao 2013)] were calculated separately for the bacterial communities on the two types of hosts ('Early' = blastoderm visible atop the oil droplets, 'Late' = late-eyed stage) and in the water samples that had been collected simultaneously with egg samples ('W-early' and 'W-late'). The *boxes* show medians, *quartiles*, *whiskers* and *outliers*

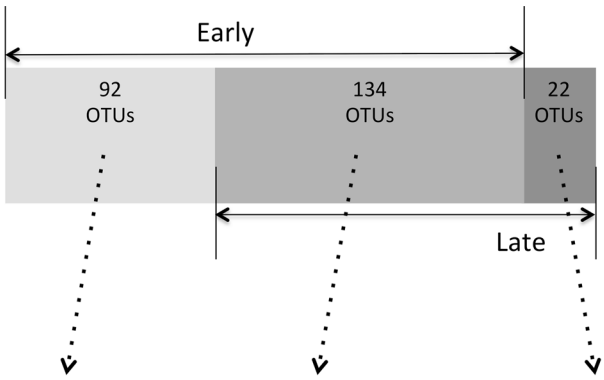

Fig. 4 Whitefish egg-associated bacterial taxa. Number of operational taxonomic units (OTUs) that were found on whitefish eggs at the early stage only, at the late stage only, and at both stages. Ten most abundant bacterial sequences that matched unambiguously to a reference with RDP classifier 2.2 and Greengenes reference database (version 12.10) at the family, genus or species level are listed for each of these groups, including their abundance. [a]Only eight bacterial sequences at the early developmental stage of whitefish eggs matched unambiguously to a reference. [b]*Mycoplana*, *Citrobacter*, and *Pseudomonas* sequences were also found in the blank control sample of only beads and buffers

Most frequent 8 taxa[a] of early only reads:

1. *Polynucleobacter* (6.6%)
2. *Sulfuritalea* (5.3%)
3. Isosphaeraceae (1.9%)
4. Hyphomonadaceae (1.4%)
5. *Mycoplana* (1.4%)[b]
6. Saprospiraceae (1.2%)
7. *Asticcacaulis* (1.2%)
8. *Hydrogenophaga* (1.1%)

Most frequent 10 taxa of shared reads:

1. Xanthomonadaceae (9.1%)
2. Desulfohalobiaceae (5.2%)
3. Nitrospinaceae (5.1%)
4. *Bdellovibrio* (3.9%)
5. *Gemmatimonas* (3.8%)
6. *Zoogloea* (2.1%)
7. *Stenotrophomonas* (1.8%)
8. Rhodocyclaceae (1.7%)
9. Alteromonadaceae (1.7%)
10. *Citrobacter* (0.7%)[b]

Most frequent 10 taxa of late only reads:

1. *Streptobacillus* (18%)
2. *Oxalobacter* (9.6%)
3. *Thermodesulfo-vibrio* (5.9%)
4. *Microvirgula* (5.9%)
5. *Pseudomonas* (5.1%)[b]
6. Fusobacteriaceae (4.9%)
7. *Candidatus methylomirabilis* (4.4%)
8. *Deefgea* (3.6%)
9. *Aquitalea* (3.6%)
10. *Nitrospira* (2.6%)

Resampling strategy

Water samples had been collected at seven locations simultaneously with egg samples (Table 1). Non-parametric testing was applied to determine whether bacterial communities in the water samples were significantly different from those on egg samples with adonis ($p < 0.001$, $R^2 = 44.1$ %) and ANOSIM ($p < 0.001$, $R^2 = 96.2$ %), while the assumptions of the two tests were met and variances between the two groups were not significantly different from each other ($F = 0.65$, d.f. $= 4$, $p = 0.77$). The difference in mean bacterial diversity was visualized in boxplots (Supplementary Fig. S1c), and the tests of significant pairwise differences among means were performed using a two-sided Student's two-sample *t*-test of the two means ($t = 15.7$, d.f. $= 4$, $p = 0.001$). Bacterial communities on four of the seven egg samples turned out to be significantly different from those of the corresponding simulated egg samples using sampling with replacement from the associated water samples (Table 3). Bacterial communities on the three egg samples that did reflect a random sample of bacterial communities in their corresponding water sample were all eggs at an early stage. All bacterial species of this study are documented in the Supplementary Tables S1 and S2. Alpha diversities of bacteria from whitefish egg samples were not correlated with alpha diversities of those from the corresponding water samples (Chao1: correlation $= -0.19$, $p = 0.59$, number of observed species: correlation $= -0.37$, $p = 0.38$, phylogenetic distance: correlation $= -0.094$, $p = 0.79$; Supplementary Fig. S2).

Discussion

We investigated the diversity of bacterial communities on naturally spawned whitefish eggs sampled at two different developmental stages and from six different lakes. We found a large number of OTUs associated with these eggs (as listed in the Supplementary Table S1). We also found that the bacterial diversity significantly decreased over the course of host development. At an early developmental stage, bacterial communities on eggs were similar to the bacterial communities in the surrounding water environment, whereas at a later stage they constituted a significantly different community with fewer bacterial taxa. This decrease suggests a transition of the egg as substrate for bacteria from unselective to more selective throughout the course of embryonic development. Our field study therefore supports conclusions derived from recent laboratory experiments (Clark and Wedekind 2011; Clark et al. 2013, 2014).

Table 3 Simulated UniFrac bacterial community distances and probability of randomly obtaining the observed distances by sampling with replacement

Location	UniFrac-Dist	sim. Mean	sim. CI	Probability
Whitefish-Lucerne-early	0.494	0.506	0.489; 0.524	0.59
Whitefish-Biel-early	0.578	0.576	0.560; 0.592	0.51
Whitefish-Thun-early	0.479	0.469	0.452; 0.484	0.45
Whitefish-Geneva-early	0.745	0.56	0.543; 0.577	<0.001
Whitefish-Thun-late	0.802	0.555	0.537; 0.574	<0.001
Whitefish-Lucerne-late	0.759	0.511	0.493; 0.529	<0.001
Whitefish-Hallwyl-late	0.806	0.474	0.456; 0.493	<0.001

Egg samples and water samples were collected simultaneously at seven different locations. UniFrac distances (UniFrac-Dist) between the respective egg and water sample at a given location were compared to the simulated mean UniFrac distances (sim. Mean) between a given water sample and a simulated egg sample. Egg samples were simulated by randomly sampling with replacement from the OTUs found in the respective water samples. One thousand egg samples were simulated in order to obtain a density distribution and corresponding confidence intervals (sim. CI). The probability was calculated of obtaining the given UniFrac distance between water and egg sample in the simulations

Quality control

We performed a stringent quality control including the exclusion of all the sequences that had been found in the negative controls. It is likely that we thereby eliminated not only systematic contaminations in all samples, but also some true bacterial species collected at the sampling sites. We nevertheless preferred this conservative approach because a substantial amount of bacterial sequences could be detected in the negative controls. It may be generally assumed that negative controls do not amplify anything when there is no band detected during the electrophoresis of PCR products (Git et al. 2010). Indeed, we did not notice any bands after PCR amplification, but found reads in the negative controls after 454 pyrosequencing. Some of these reads seem to represent bacteria that are typically associated with freshwater environments, such as species of the genus *Limnohabitans* (Simek et al. 2010) or *Leuconostoc* (Arahal et al. 2008). Two of the sequences detected in the negative controls may even represent known fish pathogens, *Aeromonas* spp. (Austin and Austin 2007) and *Carnobacterium* spp. (Leisner et al. 2007). However, it is possible that these reads originated from the background noise of the 454 pyrosequencing process and did not represent contaminations in beads and buffers (Xu 2006; Schloss et al. 2011). Background noise sequences can arise due to the specificity of the primers or incomplete sequence products and polymerase errors during the PCR (von Wintzingerode et al. 1997; Junier et al. 2008; Lighten et al. 2014). 454 pyrosequencing is based on several steps of PCR amplification. We therefore decided to filter out all OTUs that could be found in the negative controls and to then report them.

Fish embryos and their associated bacteria

In freshwater systems, the composition of microbial communities is often driven by strong host effects (Qi et al. 2009; Mouchet et al. 2012), and could be linked to host diet (Nicholson et al. 2012), host respiration rates (Gattuso et al. 2002), and host adaptation to different ecological conditions (Pavey et al. 2013; Sevellec et al. 2014). Whether and how early host development in the wild affects the composition of symbiotic bacterial communities has not been studied, neither in freshwater systems nor in other environments (Romero and Navarrete 2006; Gilbert et al. 2010). Host effects on natural bacterial communities mostly have been studied using host species at juvenile stage or later (e.g., Evans and Neff 2009; Shi et al. 2010; Pantoja-Feliciano et al. 2013; Sevellec et al. 2014). Here, we looked at the relationship of hosts and their associated bacteria during whitefish embryogenesis. Under controlled laboratory conditions, fish embryos have been used to study, for example, colonization by pathogenic bacteria (Milligan-Myhre et al. 2011; Rendueles et al. 2012), maternal microbial transmission (Funkhouser and Bordenstein 2013), or the effect of nutrient concentration on virulence of associated bacteria (Jacob et al. 2010; Wedekind et al. 2010). Complementary field studies are necessary to put such findings into their natural context and to assess their ecological relevance.

The aims of the present field study were to describe the diversity of bacterial communities associated with whitefish at different embryonic developmental stages and to compare them to the bacterial communities in the surrounding environment, i.e., to test whether the bacterial communities simply reflect the bacterial species found in the surrounding water or if there are sorting effects of the

host on egg-associated bacteria. Whitefish eggs may be ideal for this type of study because they represent an immobile host that goes through embryogenesis while fully exposed to the bacterial communities found in the water environment. More than 40 % of DNA sequences on the egg samples did not result in a perfect match to a bacterial sequence in the reference database (after quality control). This outcome is possible because: (i) bacteria associated with fish eggs have not been studied much, and therefore are not represented in the reference database, and (ii) we applied a de novo OTU picking process (Schloss et al. 2011; Ong et al. 2013). For this process, all reads that passed quality control are first clustered against one another without any external reference. Afterwards, the most abundant DNA sequence from every cluster is searched against a reference database. New DNA sequences can be found that have not been described previously using this approach. These new DNA sequences were assigned to the most similar reference available in the reference database. If there are many new and undescribed DNA sequences in a dataset, it is possible that many different sequences match to the same reference sequence. They might represent bacterial species or strains specifically associated with freshwater fish eggs (Ong et al. 2013).

We found natural bacterial communities on whitefish eggs to be diverse, ranging from 160 to 892 different bacterial OTUs per sample. Within this range, there was a clear pattern: the diversity of bacterial communities on whitefish eggs decreased significantly from an early to a late host developmental stage. This decrease persisted among different lakes and species and was consistently reflected in different alpha diversity measures.

Time effects

Which factors would result in a decreased diversity of bacterial communities associated with whitefish embryos in a natural freshwater environment? The embryo is likely to be under strong pathogen pressures throughout its development, as the egg surface provides a substrate for bacterial growth, and the embryo is unable to escape (Treasurer et al. 2005; Wedekind et al. 2010). The bacterial diversity on whitefish eggs is expected to decrease if some substrates, such as for example different glycoproteins, become exhausted (Brivio et al. 1991). Some bacteria are able to outcompete the others (Burke et al. 2011; Johnson et al. 2012) or produce antibiotics to which they themselves are resistant (Cordero et al. 2012). Temperature also can have an effect on microbial community composition (Gilbert et al. 2010). It has been shown that a decrease in temperature can diminish the diversity of natural microbial communities (Zogg et al. 1997; Coulon et al. 2007; Urakawa et al. 2008). The temperature in sub-Alpine Swiss lakes decreases during the incubation of naturally spawned whitefish eggs (Hari et al. 2006). Nevertheless, temperature alone is unlikely to be solely responsible for our findings because in our dataset: (i) only the diversities of bacterial communities on whitefish eggs decreased significantly with time while the simultaneously collected bacterial communities of water samples did not, and (ii) later species compositions on eggs no longer matched those of the respective water samples. Alternatively, bacterial communities might become more specialized during embryogenesis because of embryo defense mechanisms. Fish embryos seem to complement and specifically increase their immunity through time (Romero and Navarrete 2006). Indeed, towards the end of embryogenesis, important immune genes are expressed (Clark et al. 2013) and the immune system can be shown to influence the embryos' resistance to microbial challenges (Ellis 2001; Wedekind et al. 2004; Pitcher and Neff 2006). In-vitro fertilized whitefish eggs have been used as models to study genetic and maternal effects during bacterial pathogen incubation (von Siebenthal et al. 2009) using full factorial breeding designs ('North Carolina II design', Lynch and Walsh 1998). This design allows researchers to separately estimate the maternal effects and paternal effects on embryo performance. Such studies have shown that the embryos' defense is first dominated by maternally transmitted factors, and that the embryos' own genetics and specific immunities become increasingly important during later embryogenesis (Clark et al. 2014). This transition of host immunity could lead to the observed reduction in bacterial diversity.

Properties of whitefish egg-associated bacteria

In order to discuss possible consequences of egg-associated bacteria on their whitefish embryo host, the most abundant bacterial sequences found on whitefish eggs exclusively were studied. To be considered, these bacterial sequences needed to result in a perfect match to a reference at the family, genus or species level. Bacterial taxa on whitefish eggs were split up into three groups: early developmental stage, late developmental stage, and shared between stages of whitefish embryos. For the first of these three groups, only eight bacterial sequences matched unambiguously, e.g., 84 sequences could not be assigned to a reference at the family or lower level. Accordingly, bacteria on early whitefish eggs most probably present a bacterial community that is not well represented in the reference database, i.e., early whitefish eggs were colonized by a community of bacteria that is generally uncharacterized. The most common bacterial taxa on early whitefish eggs that matched to a previously reported reference according to a search in Web Of Knowledge™ comprised mainly free-living

freshwater bacteria, such as the genera of planktonic *Hydrogenophaga* (Willems et al. 1989), *Polynucleobacter* (Hahn et al. 2010), or the facultatively autotrophic genus *Sulfuritalea* (Kojima and Fukui 2011). Four ubiquitous bacterial groups that contributed with many sequences to the bacterial communities on early whitefish eggs have previously mostly been found in soil samples: *Asticcacaulis* (Kim et al. 2013), Hyphomonadaceae (Lee et al. 2005), Isosphaeraceae (Joseph et al. 2003), and *Mycoplana* (Urakami et al. 1990). Another family that was fairly common was the family of Saprospiraceae, which contains many genera of aquatic cosmopolitans (Schauer et al. 2005). All of these bacterial taxa exemplify common environmental bacteria without any previously reported association to fish (Madigan et al. 2010).

In contrast to the main bacterial groups on early-developed whitefish embryos, the most abundant bacterial taxa at a later developmental stage comprised a more specialized community in terms of substrate utilization. The egg-associated bacterial community at this stage was dominated by a bacterial sequence that matched the genus *Streptobacillus* that has the potential to act as a bacterial fish pathogen (Maher et al. 1995). The second most common bacterial group was *Oxalobacter*, a common genus of commensalists in vertebrates, which usually inhabit the gastrointestinal tract (Allison et al. 1985). These beneficial bacteria can degrade calcium oxalate. Two genera belonged to the guild of denitrifying bacteria, i.e., *Aquitalea* (Lee et al. 2012) and *Microvirgula* (Patureau et al. 1998). The habitat for this guild of bacteria is not specified. They can reduce both oxygen and nitrogen simultaneously when oxygen is present. There was also a nitrite-oxidizing genus *Nitrospira* (Ehrich et al. 1995; Watson et al. 1986) and a sulfate-reducing genus *Thermodesulfovibrio* (Henry et al. 1994). Furthermore, we found sequences that matched *Candidatus methylomirabilis* that is capable of coupling methane oxidation to denitrification (Wu et al. 2012). The last two members of the bacterial community on fertilized whitefish eggs at the late-eyed stage listed here have previously been described to associate with fish: *Pseudomonas* is a bacterial genus that harbors several species that can act as fish pathogens (Austin and Austin 2007; Eissa et al. 2010), such as *P. fluorescens*, an opportunistic whitefish pathogen that can significantly reduce embryonic survival and performance (von Siebenthal et al. 2009; Clark et al. 2014) and different members of the *Deefgea* genus, which have also been involved in opportunistic infections of aquatic organisms (Jung and Jung-Schroers 2011). In summary, it appears that egg-associated bacteria at a later stage of whitefish embryonic development are better characterized in the literature and present more discrete bacterial guilds.

The bacterial community that was shared between stages harbored a mix of environmental and specialized bacteria. Xanthomonadaceae with the main genus *Stenotrophomonas* are very abundant and colonize mainly soil and freshwater habitats (Denton and Kerr 1998). Some species in this family have been shown to act as plant pathogens (Madigan et al. 2010). Nitrospinaceae and Desulfohalobiaceae present a microbial community that is not well characterized. The latter family contains several thermophilic sulfate-reducing species (Roychoudhury et al. 2013). The genus *Bdellovibrio* includes small parasites of bacteria, which parasitize preferably *Pseudomonas* and *Enterobacteria* (Rendulic et al. 2004). *Gemmatimonas* is a recently described bacterial group whose members are widespread in the environment, particularly in soil habitats (Joseph et al. 2003; Zhang et al. 2003). *Zoogloea* are abundant in sludge (Shin et al. 1993). They can be used as an indicator of nitrogen or phosphorus deficiency. Alteromonadaceae is a bacterial family that contains many genera that belong to bacterioplankton (Fodelianakis et al. 2014) and can resist low temperatures (Ivanova et al. 2004). Citrobacter, the last group that was characterized for this group of whitefish eggs, are ubiquitous bacteria that contain many species that may use citrate as a sole carbon source (Madigan et al. 2010). They are found in vertebrate intestines and can live in symbiosis with invertebrates (Behar et al. 2005).

The bacterial guilds in the three groups of whitefish eggs listed above represent only bacterial sequences that matched perfectly to a reference in the database. The majority of the remaining sequences, especially at the early developmental stage, did not result in an unambiguous match, i.e., they have not been deposited in the reference before. Bacterial sequences in our dataset are confined to a length of 311-bp long reads, which also limited our ability to determine bacteria to a species level. Nevertheless, we set a starting point for the investigation of an interaction between whitefish embryos and their associated bacterial communities. We characterized the bacterial community upon developing whitefish eggs collected from six sub-Alpine lakes and showed that the bacterial community is strikingly rich and changes over time.

Conclusions

Members of the salmonid family cover a wide range of different life-history strategies (Hutchings and Morris 1985; Crete-Lafreniere et al. 2012) and ecological conditions (e.g., Maret et al. 1997; Evans et al. 2010, 2012; Hecht et al. 2012). Symbiotic microbes may to a large degree reflect these ecological conditions. Such microbes can significantly affect host development and life history (Wedekind 2002; von Siebenthal et al. 2009; Wedekind et al. 2010), while hosts have various kinds of defense

mechanisms (e.g., Ellis 2001; Evans and Neff 2009; Aykanat et al. 2012). Salmonid eggs are therefore promising models for studying host-symbiont co-evolution in different ecological conditions. Here, we provide a detailed protocol on how to collect and extract bacterial DNA from whitefish eggs in the wild and how to analyze next-generation sequencing reads in the QIIME framework (Caporaso et al. 2010b). We report a large number of OTUs found on eggs from different lakes and at different embryonic developmental stages. We describe changes in bacterial diversity during embryogenesis that are consistent over sampling sites and host whitefish species, and that support predictions from recent experimental work.

Acknowledgments The fieldwork was permitted by the fisheries authorities of the cantons of Berne, Lucerne, and Vaud. We thank Philippe Amiet, Joachim Guthruf, Frédéric Hofmann, Christoph Küng, Joseph Muggli, Rudolph Müller, Patrick Porchet, Rolf Schneider, and Gérard Zürcher for permissions, organizational support, or help in finding the spawning places and collecting the eggs. Cathérine Berney, Gregory Brazzola, Emily Clark, Hélène Collin, Lucas Marques da Cunha, Miguel dos Santos, Simon Gingins, Marianne Johner, Catalina Luca, Yves Poirier, Manuel Pompini, Thorsten Reusch, Adin Ross-Gillespie, Vladimir Sentchilo, Rike Stelkens, Pierre Taberlet, Anshu Uppal, Jan Roelof van der Meer, Beat von Siebenthal, Donny Wilkins, Fardo Witsenburg, and Erika Yashiro are thanked for assistance or discussion, and the Swiss National Science Foundation for financial support. We also thank Stuart Findlay and two anonymous reviewers for comments on the manuscript. L. Wilkins was supported by a fellowship in Life Sciences from the Faculty of Biology and Medicine of the University of Lausanne.

References

Allison MJ, Dawson KA, Mayberry WR, Foss JG (1985) *Oxalobacter formigenes* gen. nov., sp. nov., oxalate degrading anaerobes that inhabit the gastrointestinal tract. Arch Microbiol 141(1):1–7

Amann R, Ludwig W (2000) Ribosomal RNA-targeted nucleic acid probes for studies in microbial ecology. FEMS Microbiol Rev 24(5):555–565

Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Aust Ecol 26(1):32–46

Anderson MJ (2006) Distance-based tests for homogeneity of multivariate dispersions. Biometrics 62(1):245–253

Anderson MJ, Willis TJ (2003) Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. Ecology 84(2):511–525

Arahal DR, Sanchez E, Macian MC, Garay E (2008) Value of *recN* sequences for species identification and as a phylogenetic marker within the family 'Leuconostocaceae'. Int Microbiol 11(1):33–39

Austin B, Austin DA (2007) Bacterial fish pathogens: diseases of farmed and wild fish, 5th edn. Springer, London

Aykanat T, Heath JW, Dixon B, Heath DD (2012) Additive, non-additive and maternal effects of cytokine transcription in response to immunostimulation with *Vibrio* vaccine in Chinook salmon (*Oncorhynchus tshawytscha*). Immunogenetics 64(9):691–703

Behar A, Yuval B, Jurkevitch E (2005) Enterobacteria-mediated nitrogen fixation in natural populations of the fruit fly *Ceratitis capitata*. Mol Ecol 14(9):2637–2643

Berry D, Ben Mahfoudh K, Wagner M, Loy A (2011) Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. Appl Environ Microb 77(21):7846–7849

Boutin S, Sevellec M, Pavey SA, Bernatchez L, Derome N (2012) A fast, highly sensitive double-nested PCR-based method to screen fish immunobiomes. Mol Ecol Resour 12(6):1027–1039

Brivio MF, Bassi R, Cotelli F (1991) Identification and characterization of the major components of the *Oncorhynchus mykiss* egg chorion. Mol Reprod Dev 28(1):85–93

Brucker RM, Bordenstein SR (2013) The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus *Nasonia*. Science 341(6146):667–669

Burke C, Steinberg P, Rusch D, Kjelleberg S, Thomas T (2011) Bacterial community assembly based on functional genes rather than species. Proc Natl Acad Sci USA 108(34):14288–14293

Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R (2010a) PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics 26(2):266–267

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Tumbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010b) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7(5):335–336

Casadevall A, Pirofski LA (1999) Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. Infect Immun 67(8):3703–3713

Casadevall A, Pirofski LA (2000) Host-pathogen interactions: basic concepts of microbial commensalism, colonization, infection, and disease. Infect Immun 68(12):6511–6518

Clark E, Wedekind C (2011) Additive genetic effects on embryo viability in a whitefish (Salmonidae) influenced by the water mould *Saprolegnia ferax*. J Bacteriol Parasitol 10:2–6

Clark ES, Wilkins LGE, Wedekind C (2013) MHC class I expression dependent on bacterial infection and parental factors in whitefish embryos (Salmonidae). Mol Ecol 22(20):5256–5269

Clark ES, Pompini M, Marques da Cunha L, Wedekind C (2014) Maternal and paternal contributions to pathogen resistance dependent on development stage in a whitefish. Funct Ecol 28:714–723

Cole JR, Chai B, Farris RJ, Wang Q, Kulam SA, McGarrell DM, Garrity GM, Tiedje JM (2005) The ribosomal database project (RDP-II): sequences and tools for high-throughput rRNA analysis. Nucleic Acids Res 33(Database issue):D294–D296

Cordero OX, Wildschutte H, Kirkup B, Proehl S, Ngo L, Hussain F, Le Roux F, Mincer T, Polz MF (2012) Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. Science 337(6099):1228–1231

Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R (2009) Bacterial community variation in human body habitats across space and time. Science 326(5960):1694–1697

Coulon F, McKew BA, Osborn AM, McGenity TJ, Timmis KN (2007) Effects of temperature and biostimulation on oil-degrading microbial communities in temperate estuarine waters. Environ Microbiol 9(1):177–186

Crete-Lafreniere A, Weir LK, Bernatchez L (2012) Framing the Salmonidae family phylogenetic portrait: a more complete picture from increased taxon sampling. PLoS One 7(10):e46662

Denton M, Kerr KG (1998) Microbiological and clinical aspects of infection associated with *Stenotrophomonas maltophilia*. Clin Microbiol Rev 11(1):57–80

Dillon RJ, Dillon VM (2004) The gut bacteria of insects: non-pathogenic interactions. Annu Rev Entomol 49:71–92

Douglas AE (1994) Symbiotic interactions. Oxford University Press, Oxford

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19):2460–2461

Ehrich S, Behrens D, Lebedeva E, Ludwig W, Bock E (1995) A new obligately chemolithoautotrophic, nitrite-oxidizing bacterium, *Nitrospira moscoviensis* sp. nov. and its phylogenetic relationship. Arch Microbiol 164(1):16–23

Eissa NME, Abou El-Ghiet EN, Shaheen AA, Abbas A (2010) Characterization of *Pseudomonad* species isolated form Tilapia "*Oreochromis niloticus*" in Qaroun and Wadi-El-Rayan Lakes. Egypt. Global Veterinaria 5(2):116–121

Ellis AE (2001) Innate host defense mechanisms of fish against viruses and bacteria. Dev Comp Immunol 25(8–9):827–839

Evans ML, Neff BD (2009) Major histocompatibility complex heterozygote advantage and widespread bacterial infections in populations of Chinook salmon (*Oncorhynchus tshawytscha*). Mol Ecol 18(22):4716–4729

Evans ML, Neff BD, Heath DD (2010) MHC-mediated local adaptation in reciprocally translocated Chinook salmon. Conserv Genet 11(6):2333–2342

Evans ML, Praebel K, Peruzzi S, Bernatchez L (2012) Parallelism in the oxygen transport system of the lake whitefish: the role of physiological divergence in ecological speciation. Mol Ecol 21(16):4038–4050

Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R (2010) Forensic identification using skin bacterial communities. Proc Natl Acad Sci USA 107(14):6477–6481

Finn RN (2007) The physiology and toxicology of salmonid eggs and larvae in relation to water quality criteria. Aquat Toxicol 81(4):337–354

Fodelianakis S, Papageorgiou N, Pitta P, Kasapidis P, Karakassis I, Ladoukakis ED (2014) The Pattern of change in the abundances of specific bacterioplankton groups is consistent across different nutrient-enriched habitats in Crete. Appl Environ Microb 80(13):3784–3792

Funkhouser LJ, Bordenstein SR (2013) Mom knows best: the universality of maternal microbial transmission. PLoS Biol 11(8):e1001631

Gattuso JP, Peduzzi S, Pizay MD, Tonolla M (2002) Changes in freshwater bacterial community composition during measurements of microbial and community respiration. J Plankton Res 24(11):1197–1206

Gilbert JA, Field D, Swift P, Thomas S, Cummings D, Temperton B, Weynberg K, Huse S, Hughes M, Joint I, Somerfield PJ, Muhling M (2010) The taxonomic and functional diversity of microbes at a temperate coastal site: a 'multi-omic' study of seasonal and diel temporal variation. PLoS One 5(11):e15545

Git A, Dvinge H, Salmon-Divon M, Osborne M, Kutter C, Hadfield J, Bertone P, Caldas C (2010) Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. RNA 16(5):991–1006

Gotelli NJ, Chao A (2013) Measuring and estimating species richness, species diversity, and biotic similarity from sampling data. In: Levin SA (ed) Encyclopedia of Biodiversity, vol 5, 2nd edn. Academic Press, Waltham, pp 195–211

Hahn MW, Lang E, Brandt U, Lünsdorf H, Wu QL, Stackebrandt E (2010) *Polynucleobacter cosmopolitanus* sp. nov., free-living planktonic bacteria inhabiting freshwater lakes and rivers. Int J Syst Evol Microbiol 60:166–173

Hamady M, Walker JJ, Harris JK, Gold NJ, Knight R (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nat Methods 5(3):235–237

Hanif A, Bakopoulos V, Dimitriadis GJ (2004) Maternal transfer of humoral specific and non-specific immune parameters to sea bream (*Sparus aurata*) larvae. Fish Shellfish Immun 17(5):411–435

Hari RE, Livingstone DM, Siber R, Burkhardt-Holm P, Guttinger H (2006) Consequences of climatic change for water temperature and brown trout populations in Alpine rivers and streams. Global Change Biol 12(1):10–26

Hecht BC, Thrower FP, Hale MC, Miller MR, Nichols KM (2012) Genetic architecture of migration-related traits in rainbow and steelhead trout, *Oncorhynchus mykiss*. G3-Genes Genomes Genet 2(9):1113–1127

Henry EA, Devereux R, Maki JS, Gilmour CC, Woese CR, Mandelco L, Schauder R, Remsen CC, Mitchell R (1994) Characterization of a new thermophilic sulfate-reducing bacterium—*Thermodesulfovibrio yellowstonii*, gen. nov. and sp. nov., its phylogenetic relationship to *Thermodesulfobacterium commune* and their origins deep within the bacterial domain. Arch Microbiol 161(1):62–69

Hill TCJ, Walsh KA, Harris JA, Moffett BF (2003) Using ecological diversity measures with bacterial communities. FEMS Microbiol Ecol 43(1):1–11

Hollander M, Wolfe DA (1973) Nonparametric statistical methods, 2nd edn. Wiley, New York

Hooper LV, Gordon JI (2001) Commensal host-bacterial relationships in the gut. Science 292(5519):1115–1118

Hooper LV, Midtvedt T, Gordon JI (2002) How host-microbial interactions shape the nutrient environment of the mammalian intestine. Annu Rev Nutr 22:283–307

Hughes JB, Hellmann JJ, Ricketts TH, Bohannan BJM (2002) Counting the uncountable: statistical approaches to estimating microbial diversity. Appl Environ Microb 68(1):448

Huse SM, Dethlefsen L, Huber JA, Welch DM, Relman DA, Sogin ML (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. PLoS Genet 4(11):e1000255

Hutchings JA, Morris DW (1985) The influence of phylogeny, size and behavior on patterns of covariation in salmonid life histories. Oikos 45(1):118–124

Ivanova EP, Flavier S, Christen R (2004) Phylogenetic relationships among marine *Alteromonas*-like proteobacteria: emended description of the family Alteromonadaceae and proposal of Pseudoalteromonadaceae fam. nov., Colwelliaceae fam. nov., Shewanellaceae fam. nov., Montellaceae fam. nov., Ferrimonadaceae fam. nov., Idiomarinaceae fam. nov and Psychromonadaceae fam. nov. Int J Syst Evol Microbiol 54:1773–1788

Jacob A, Evanno G, von Siebenthal BA, Grossen C, Wedekind C (2010) Effects of different mating scenarios on embryo viability in brown trout. Mol Ecol 19(23):5296–5307

Johnson DR, Goldschmidt F, Lilja EE, Ackermann M (2012) Metabolic specialization and the assembly of microbial communities. ISME J 6(11):1985–1991

Joseph SJ, Hugenholtz P, Sangwan P, Osborne CA, Janssen PH (2003) Laboratory cultivation of widespread and previously uncultured soil bacteria. Appl Environ Microbiol 69(12):7210–7215

Jung A, Jung-Schroers V (2011) Detection of *Deefgea chitinilytica* in freshwater ornamental fish. Lett Appl Microbiol 52(5):497–500

Junier P, Kim OS, Hadas O, Imhoff JF, Witzel KP (2008) Evaluation of PCR primer selectivity and phylogenetic specificity by using amplification of 16S rRNA genes from betaproteobacterial ammonia-oxidizing bacteria in environmental samples. Appl Environ Microbiol 74(16):5231–5236

Kim S, Gong G, Park TH, Um Y (2013) *Asticcacaulis solisilvae* sp. nov., isolated from forest soil. Int J Syst Evol Microbiol 63:3829–3834

Kojima H, Fukui M (2011) *Sulfuritalea hydrogenivorans* gen. nov., sp. nov., a facultative autotroph isolated from a freshwater lake. Int J Syst Evol Microbiol 61:1651–1655

Kottelat M, Freyhof J (2007) Handbook of European Freshwater Fishes. Kottelat/Freyhof, Cornol/Berlin

Lan YM, Wang Q, Cole JR, Rosen GL (2012) Using the RDP classifier to predict taxonomic novelty and reduce the search space for finding novel organisms. PLoS One 7(3):e3249

Larsen A, Tao Z, Bullard SA, Arias CR (2013) Diversity of the skin microbiota of fishes: evidence for host species specificity. FEMS Microbiol Ecol 85(3):483–494

Lauber CL, Hamady M, Knight R, Fierer N (2009) Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl Environ Microbiol 75(15):5111–5120

Lee KB, Liu CT, Anzai Y, Kim H, Aonol T, Oyaizu H (2005) The hierarchical system of the 'Alphaproteobacteria': description of Hyphomonadaceae fam. nov., Xanthobacteraceae fam. nov., and Erythrobacteraceae fam. nov. Int J Syst Evol Microbiol 55:1907–1919

Lee CM, Weon HY, Kim YJ, Son JA, Yoon SH, Koo BS, Kwon SW (2012) *Aquitalea denitrificans* sp. nov., isolated from a Korean wetland. Int J Syst Evol Micrbiol 62:259

Legendre P, Anderson MJ (1999) Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. Ecol Monogr 69(1):1–24

Leisner JJ, Laursen BG, Prevost H, Drider D, Dalgaard P (2007) *Carnobacterium*: positive and negative effects in the environment and in foods. FEMS Microbiol Rev 31(5):592–613

Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JI (2008) Evolution of mammals and their gut microbes. Science 320(5883):1647–1651

Lighten J, van Oosterhout C, Paterson IG, McMullan M, Bentzen P (2014) Ultra-deep Illumina sequencing accurately identifies MHC class IIb alleles and provides evidence for copy number variation in the guppy (*Poecilia reticulata*). Mol Ecol Resour (in press)

Løvoll M, Kilvik T, Boshra H, Bogwald J, Sunyer JO, Dalmo RA (2006) Maternal transfer of complement components C3-1, C3-3, C3-4, C4, C5, C7, Bf, and Df to offspring in rainbow trout (*Oncorhynchus mykiss*). Immunogenetics 58(2–3):168–179

Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microb 71(12):8228–8235

Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. Proc Natl Acad Sci USA 104(27):11436–11440

Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R (2011) UniFrac: an effective distance metric for microbial community comparison. ISME J 5(2):169–172

Luczynski M, Kirklewska A (1984) Dependence of *Coregonus albula* embryogenesis rate on the incubation temperature. Aquaculture 42(1):43–55

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates Inc., Sunderland

Madigan MT, Martinko JM, Stahl D, Clark DP (2010) Brock biology of microorganisms, 13th edn. Benjamin Cummings, San Francisco

Maher M, Palmer R, Gannon F, Smith T (1995) Relationship of a novel bacterial fish pathogen to *Streptobacillus moniliformis* and the Fusobacteria group, based on 16s Ribosomal RNA analysis. Syst Appl Microbiol 18(1):79–84

Maret TR, Robinson CT, Minshall GW (1997) Fish assemblages and environmental correlates in least-disturbed streams of the upper Snake River basin. Trans Am Fish Soc 126(2):200–216

Martin G, Guggiari M, Bravo D, Zopfi J, Cailleau G, Aragno M, Job D, Verrecchia E, Junier P (2012) Fungi, bacteria and soil pH: the oxalate-carbonate pathway as a model for metabolic interaction. Environ Microbiol 14(11):2960–2970

McArdle BH, Anderson MJ (2001) Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology 82(1):290–297

McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 6(3):610–618

McMurdie PJ, Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8(4):e61217

Milligan-Myhre K, Charette JR, Phennicie RT, Stephens WZ, Rawls JF, Guillemin K, Kim CH (2011) Study of host-microbe interactions in zebrafish. Method Cell Biol 105:87–116

Mouchet MA, Bouvier C, Bouvier T, Troussellier M, Escalas A, Mouillot D (2012) Genetic difference but functional similarity among fish gut bacterial communities through molecular and biochemical fingerprints. FEMS Microbiol Ecol 79(3):568–580

Müller R, Stadelmann P (2004) Fish habitat requirements as the basis for rehabilitation of eutrophic lakes by oxygenation. Fish Manag Ecol 11(3–4):251–260

Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, Pettersson S (2012) Host-gut microbiota metabolic interactions. Science 336(6086):1262–1267

Nusslé S, Bornand CN, Wedekind C (2009) Fishery-induced selection on an Alpine whitefish: quantifying genetic and environmental effects on individual growth rate. Evol Appl 2:200–208

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stenvens MHH, Wagner H (2013) vegan: Community Ecology Package. R package version 2.0-8

Ong SH, Kukkillaya VU, Wilm A, Lay C, Ho EXP, Low L, Hibberd ML, Nagarajan N (2013) Species identification and profiling of complex microbial communities using shotgun Illumina sequencing of 16S rRNA amplicon sequences. PLoS One 8(4):e60811

Pantoja-Feliciano IG, Clemente JC, Costello EK, Perez ME, Blaser MJ, Knight R, Dominguez-Bello MG (2013) Biphasic assembly of the murine intestinal microbiota during early development. ISME J 7(6):1112–1115

Patureau D, Bernet N, Bouchez T, Dabert P, Delgenes JP, Moletta R (1998) Biological nitrogen removal in a single aerobic reactor by association of a nitrifying ecosystem to an aerobic denitrifier, *Microvirgula aerodenitrificans*. J Mol Catal B-Enzym 5(1–4):435–439

Pavey SA, Sevellec M, Adam W, Normandeau E, Lamaze FC, Gagnaire PA, Filteau M, Hebert FO, Maaroufi H, Bernatchez L (2013) Nonparallelism in MHC II diversity accompanies nonparallelism in pathogen infection of lake whitefish (*Coregonus clupeaformis*) species pairs as revealed by next-generation sequencing. Mol Ecol 22(14):3833–3849

Pitcher TE, Neff BD (2006) MHC class IIB alleles contribute to both additive and nonadditive genetic effects on survival in Chinook salmon. Mol Ecol 15(9):2357–2365

Price MN, Dehal PS, Arkin AP (2008) FastBLAST: homology relationships for millions of proteins. PLoS One 3(10):e3589

Qi WH, Nong G, Preston JF, Ben-Ami F, Ebert D (2009) Comparative metagenomics of *Daphnia* symbionts. BMC Genomics 10(1):172

Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, Forney LJ (2011) Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci USA 108:4680–4687

Rendueles O, Ferrieres L, Fretaud M, Begaud E, Herbomel P, Levraud JP, Ghigo JM (2012) A new zebrafish model of oro-intestinal pathogen colonization reveals a key role for adhesion in protection by probiotic bacteria. PLoS Pathog 8(7):e1002815

Rendulic S, Jagtap P, Rosinus A, Eppinger M, Baar C, Lanz C, Keller H, Lambert C, Evans KJ, Goesmann A, Meyer F, Sockett RE, Schuster SC (2004) A predator unmasked: life cycle of *Bdellovibrio bacteriovorus* from a genomic perspective. Science 303(5658):689–692

Romero J, Navarrete P (2006) 16S rDNA-based analysis of dominant bacterial populations associated with early life stages of coho salmon (*Oncorhynchus kisutch*). Microb Ecol 51(4):422–430

Rosenberg E, Koren O, Reshef L, Efrony R, Zilber-Rosenberg I (2007) The role of microorganisms in coral health, disease and evolution. Nat Rev Microbiol 5(5):355–362

Roychoudhury AN, Cowan D, Porter D, Valverde A (2013) Dissimilatory sulphate reduction in hypersaline coastal pans: an integrated microbiological and geochemical study. Geobiology 11(3):224–233

Salonen A, Nikkila J, Jalanka-Tuovinen J, Immonen O, Rajilic-Stojanovic M, Kekkonen RA, Palva A, de Vos WM (2010) Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. J Microbiol Meth 81(2):127–134

Schauer M, Kamenik C, Hahn MW (2005) Ecological differentiation within a cosmopolitan group of planktonic freshwater bacteria (SOL cluster, Saprospiraceae, Bacteroidetes). Appl Environ Microb 71(10):5900–5907

Schloss PD, Handelsman J (2005) Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. Genome Biol 6(8):229

Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One 6(12):e27310

Schloss PD, Schubert AM, Zackular JP, Iverson KD, Young VB, Petrosino JF (2012) Stabilization of the murine gut microbiome following weaning. Gut Microbes 3(4):383–393

Sevellec M, Pavey S, Boutin S, Filteau M, Derome N, Bernatchez L (2014) Microbiome investigation in the ecological speciation context of lake whitefish (*Coregonus clupeaformis*) using next generation sequencing. J Evol Biol 27(6):1029–1046

Shade A, Read JS, Youngblut ND, Fierer N, Knight R, Kratz TK, Lottig NR, Roden EE, Stanley EH, Stombaugh J, Whitaker RJ, Wu CH, McMahon KD (2012) Lake microbial communities are resilient after a whole-ecosystem disturbance. ISME J 6(12):2153–2167

Shi WB, Syrenne R, Sun JZ, Yuan JS (2010) Molecular approaches to study the insect gut symbiotic microbiota at the 'omics' age. Insect Sci 17(3):199–219

Shin YK, Hiraishi A, Sugiyama J (1993) Molecular systematics of the genus *Zoogloea* and emendation of the genus. Int J Syst Bacteriol 43(4):826–831

Simek K, Kasalicky V, Jezbera J, Jezberova J, Hejzlar J, Hahn MW (2010) Broad habitat range of the phylogenetically narrow R-BT065 cluster, representing a core group of the Betaproteobacterial genus *Limnohabitans*. Appl Environ Microb 76(3):631–639

R Development Core Team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Treasurer JW, Cochrane E, Grant A (2005) Surface disinfection of cod *Gadus morhua* and haddock *Melanogrammus aeglefinus* eggs with bronopol. Aquaculture 250(1–2):27–35

Urakami T, Oyanagi H, Araki H, Suzuki K-I, Komagata K (1990) Recharacterization and emended description of the genus *Mycoplana* and description of two new species, *Mycoplana ramosa* and *Mycoplana segnis*. Int J Syst Evol Micr 40(4):434–442

Urakawa H, Tajima Y, Numata Y, Tsuneda S (2008) Low temperature decreases the phylogenetic diversity of ammonia-oxidizing archaea and bacteria in aquarium biofiltration systems. Appl Environ Microb 74(3):894–900

Vasanthakumar A, Handelsman J, Schloss PD, Bauer LS, Raffa KF (2008) Gut microbiota of an invasive subcortical beetle, *Agrilus planipennis* Fairmaire, across various life stages. Environ Entomol 37(5):1344–1353

von Siebenthal BA, Jacob A, Wedekind C (2009) Tolerance of whitefish embryos to *Pseudomonas fluorescens* linked to genetic and maternal effects, and reduced by previous exposure. Fish Shellfish Immun 26:531–535

von Wintzingerode F, Gobel UB, Stackebrandt E (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol Rev 21(3):213–229

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microb 73(16):5261–5267

Warton DI, Wright ST, Wang Y (2012) Distance-based multivariate analyses confound location and dispersion effects. Methods Ecol Evol 3(1):89–101

Watson SW, Bock E, Valois FW, Waterbury JB, Schlosser U (1986) *Nitrospira marina* gen. nov., sp. nov., a chemolithotrophic nitrite-oxidizing bacterium. Arch Microbiol 144(1):1–7

Wedekind C (2002) Induced hatching to avoid infectious egg disease in whitefish. Curr Biol 12(1):69–71

Wedekind C, Walker M, Portmann J, Cenni B, Müller R, Binz T (2004) MHC-linked susceptibility to a bacterial infection, but no MHC-linked cryptic female choice in whitefish. J Evol Biol 17(1):11–18

Wedekind C, von Siebenthal B, Gingold R (2007) The weaker points of fish acute toxicity tests and how tests on embryos can solve some issues. Environ Pollut 148(2):385–389

Wedekind C, Gessner MO, Vazquez F, Märki M, Steiner D (2010) Elevated resource availability sufficient to turn opportunistic into virulent fish pathogens. Ecology 91(5):1251–1256

Willems A, Busse J, Goor M, Pot B, Falsen E, Jantzen E, Hoste B, Gillis M, Kersters K, Auling G, Deley J (1989) *Hydrogenophaga*, a new genus of hydrogen-oxidizing bacteria that includes *Hydrogenophaga flava* comb. nov. (formerly *Pseudomonas flava*), *Hydrogenophaga palleronii* (formerly *Pseudomonas palleronii*), *Hydrogenophaga pseudoflava* (formerly *Pseudomonas pseudoflava* and *Pseudomonas carboxydoflava*), and *Hydrogenophaga taeniospiralis* (formerly *Pseudomonas taeniospiralis*). Int J Syst Bacteriol 39(3):319–333

Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, Damgaard CF, Dormann CF, Forchhammer MC, Grytnes JA, Guisan A, Heikkinen RK, Hoye TT, Kuhn I, Luoto M, Maiorano L, Nilsson MC, Normand S, Ockinger E, Schmidt NM, Termansen M, Timmermann A, Wardle DA, Aastrup P, Svenning JC (2013) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. Biol Rev 88(1):15–30

Wu ML, van Teeseling MCF, Willems MJR, van Donselaar EG, Klingl A, Rachel R, Geerts WJC, Jetten MSM, Strous M, van Niftrik L (2012) Ultrastructure of the denitrifying methanotroph "*Candidatus Methylomirabilis oxyfera*," a novel polygon-shaped bacterium. J Bacteriol 194(2):284–291

Xu JP (2006) Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. Mol Ecol 15(7):1713–1731

Ye L, Amberg J, Chapman D, Gaikowski M, Liu WT (2013) Fish gut microbiota analysis differentiates physiology and behavior of invasive Asian carp and indigenous American fish. ISME J 8(10):541–551

Zhang H, Sekiguchi Y, Hanada S, Hugenholtz P, Kim H, Kamagata Y, Nakamura K (2003) *Gemmatimonas aurantiaca* gen. nov., sp. nov., a gram-negative, aerobic, polyphosphate-accumulating micro-organism, the first cultured representative of the new bacterial phylum Gemmatimonadetes phyl. nov. Int J Syst Evol Micr 53:1155–1163

Zogg GP, Zak DR, Ringelberg DB, MacDonald NW, Pregitzer KS, White DC (1997) Compositional and functional shifts in microbial communities due to soil warming. Soil Sci Soc Am J 61(2):475–481