Title: Effects of host genetics and environment on egg-associated microbiotas in brown trout (Salmo trutta).
Authors: Wilkins L.G.E., *Fumagalli L., *Wedekind C., (*shared senior authors),

Journal: Molecular Ecology
Year: 2016

Volume: 25(19)

Pages: 4930-4945

DOI: 10.1111/mec.13798

Effects of host genetics and environment on egg-associated microbiotas in brown trout (*Salmo trutta*)

Laetitia G.E. Wilkins[*], Luca Fumagalli[a] and Claus Wedekind[a]

Department of Ecology and Evolution, Biophore, University of Lausanne, 1015 Lausanne, Switzerland

[*]Correspondence to L. Wilkins, Tel: +41 21 692 42 19; Fax: +41 21 692 42 65; email: laetitia.wilkins@unil.ch

[a]Joint senior authors

Key-words: Bacterial diversity, host genetics, microbiota, *Salmo trutta*, salmonid embryo.

Abstract

Recent studies found fish egg-specific bacterial communities that changed over the course of embryogenesis, suggesting an interaction between the developing host and its microbiota. Indeed, single-strain infections demonstrated that the virulence of opportunistic bacteria is influenced by environmental factors and host immune genes. However, the interplay between a fish embryo host and its microbiota has not been studied yet at the community level. In order to test whether host genetics affects the assemblage of egg-associated bacteria, adult brown trout (*Salmo trutta*) were sampled from a natural population. Their gametes were used for full-factorial *in vitro* fertilizations to separate sire from dam effects. In total 2,520 embryos were singly raised under experimental conditions that differently support microbial growth. High-throughput 16S rRNA amplicon sequencing was applied to characterize bacterial communities on milt and fertilized eggs across treatments. Dam and sire identity influenced embryo mortality, time until hatching, and composition of egg-associated microbiotas (but no link between bacterial communities on milt and on fertilized eggs could be found). Elevated resources increased embryo mortality and modified bacterial communities with a shift in their putative functional potential. Resource availability did not significantly affect any parental effects on embryo performance. Sire identity affected bacterial diversity that turned out to be a significant predictor of hatching time: embryos associated with high bacterial diversity hatched later. We conclude that both host genetics and the availability of resources define diversity and composition of egg-associated bacterial communities that then affect the life-history of their hosts.

Introduction

Mortality in fish is typically highest at early developmental stages (Llewellyn *et al*. 2014). Egg-associated bacteria play a crucial role here ranging from being the immediate cause of mortality (Wedekind 2002; Neff & Pitcher 2005; Wedekind *et al*. 2010) to acting as a first line of defence against pathogens (Boutin *et al*. 2012; Liu *et al*. 2014). Much is known about the interaction of fish embryos and their associated bacterial pathogens from single-strain infections (e.g., von Siebenthal *et al*. 2009; Aykanat *et al*. 2012; Clark *et al*. 2014). However, single-strain infections represent a strong simplification. Recent findings based on high-throughput sequencing show a considerable diversity of egg-associated microbiotas (Liu *et al*. 2014; Wilkins *et al*. 2015a; Wilkins *et al*. 2015b; Stephens *et al*. 2016) and at the community level the interactions of hosts and their associated bacteria are still not well understood (Llewellyn *et al*. 2014; Waldor *et al*. 2015). Here we specifically concentrate on the role of host genetics and environment during this interaction.

Analyses based on uncontrolled outbreaks of *Pseudomonas fluorescens* (an opportunistic fish pathogen) showed that not only virulence factors but also hatching time can be affected (Wedekind 2002) and that host susceptibility depends on immune genes (Wedekind *et al*. 2004). Subsequent experimental incubations confirmed these first findings (e.g., von Siebenthal *et al*. 2009; Aykanat *et al*. 2012; Clark *et al*. 2014). It could be demonstrated that during early development, the embryo relies on antimicrobial compounds that the mother allocated to the eggs before spawning (Hanif *et al*. 2004; D'alba *et al*. 2010). At later developmental stages the importance of zygotic gene expression becomes crucial for the interaction with associated bacteria (Finn 2007; Clark *et al*. 2013a; Clark *et al*. 2014). Host genetic effects could be revealed for important embryo life-history traits such as mortality in *Oncorhynchus tshawytscha* (Aykanat *et al*. 2012), *Salmo trutta* (Jacob *et al*. 2010; Clark *et al*. 2013b) or *Coregonus palaea* (von Siebenthal *et al*. 2009; Clark & Wedekind 2011; Clark *et al*. 2014), and time until hatching (Clark *et al*. 2013a; Clark *et al*. 2014), size at hatching (Clark *et al*. 2014; Clark *et al*. 2016) and immune gene expression (Clark *et al*. 2013a; Wilkins *et al*. 2015c) in *C. palaea*. The interplay of fish embryos with their associated bacteria also depends on environmental factors, which, for example, can turn opportunistic bacteria into virulent fish pathogens (Jacob *et al*. 2010; Wedekind *et al*. 2010). However, the factors that influence the assemblage and dynamics of fish embryo-associated bacterial communities are still poorly understood.

In adult fish, the composition of associated bacterial communities has been correlated to environmental factors, such as sampling location (Roeselers *et al*. 2011; Sullam *et al*. 2015), salinity (Sullam *et al*. 2012; Schmidt *et al*. 2015), diet (Wilson *et al*. 2008; Bolnick *et al*. 2014a; Sevellec *et al*. 2014), and temperature (Wilkins *et al*. 2015b). For bacterial communities in the guts, there is accumulating evidence that host genetics affects the composition of fish-associated microbiotas (Ghanbari *et al*. 2015). Specific bacterial groups could be linked to host phenotype (Sun *et al*. 2009; Li *et al*. 2013; Star *et al*. 2013), candidate loci in the host genome (Boutin *et al*. 2014), and alleles of the major histocompatibility complex (MHC; Bolnick *et al*. 2014b). Fish gut microbiotas are not only significantly different among different host species (Li *et al*. 2012; Larsen *et al*. 2013; Givens *et al*. 2014) but also between populations (Skrodenyte-Arbaciauskiene *et al*. 2006) and even families (Navarrete *et al*. 2012). In contrast to these distinct fish microbiotas, bacterial communities on fish embryos are expected to reflect mostly the surrounding water (Ghanbari *et al*. 2015) because in the aquatic environment bacteria can migrate freely between habitats and hosts (Llewellyn *et al*. 2014). Nonetheless, recent characterizations of naturally spawned

salmonid eggs have shown that there are egg-specific microbiotas that do not correspond to the surrounding water environment (Wilkins *et al*. 2015a; Wilkins *et al*. 2015b). Furthermore bacterial diversity decreased during embryogenesis suggesting a specific role of the host at this early developmental stage (Wilkins *et al*. 2015a).

Here, we investigated host genetic effects on embryo-associated bacterial communities using salmonid eggs. In salmonids, fertilization is external and females provide large numbers of eggs. Full-factorial breeding designs can therefore be applied where every female is crossed with every male ('North Carolina II design', Lynch & Walsh 1998). These designs allow for the disentanglement and estimation of environmental, dam, sire, and dam x sire interaction effects. Dam effects combine maternal genetic and maternal environmental effects (Royle *et al*. 1999; von Siebenthal *et al*. 2009; Aykanat *et al*. 2012) while sire effects directly reveal additive genetic variance for a trait because fathers only contribute genes to their offspring (Lynch & Walsh 1998). Accordingly, additive genetic effects can be estimated through paternal effects in this system. We sampled brown trout (*S. trutta*) from a natural sub-alpine river and collected their gametes for full-factorial *in vitro* fertilizations. Resulting embryos were raised individually under controlled conditions. In order to vary the host-microbiota environment, we added different concentrations of nutrient broth to the embryos. Nutrient broth provides resources for bacterial growth but is not harmful to embryos on its own (Wedekind *et al*. 2010). Naturally occurring bacteria on the eggs were characterized and quantified with a bacteria-specific, next-generation 16S rRNA amplicon sequencing protocol. We addressed the following objectives: (i) Are there genetic host effects on the diversity and composition of egg-associated bacteria? (ii) Which core bacterial taxa and functional pathways are affected? (iii) Is there a link between associated microbiota and host life-history? (iv) Are these interactions dependent on the availability of bacterial resources?

Materials and Methods

Sample acquisition and experimental protocol

Adult brown trout (*Salmo trutta*) were caught with electrofishing at their natural spawning grounds in the river Kiese (7°37'11,27"; 46°50'55,85"). They were kept at a hatchery for some days until the fish could be stripped of their gametes. Milt from each individual sire (40 μl) was immediately frozen in liquid nitrogen for later bacterial characterization. The remaining gametes were used for full-factorial *in vitro* fertilizations following the methods described in Jacob *et al*. (2007). Eight females were crossed with seven males in all possible combinations (56 families). After distributing fertilized eggs singly to 24-well plates (Falcon, BD Biosciences, Allschwil, Switzerland) in a block-wise design, eggs were incubated in a 6.5° C climate chamber in 2 ml of standard water. 'Standard water' was chemically standardized according to the OECD guidelines (OECD 1992). It had been autoclaved, temperated, and oxygenated before use. Once embryos had reached the late-eyed developmental stage (45 days after fertilization), healthy-looking embryos were re-distributed singly to new 24-well plates. Now, they were incubated at different concentrations of nutrient broth or they were sham-treated with standard water. Undiluted nutrient broth (NB) consisted of 3 g meat extract and 5 g bactopeptone (Sigma-Aldrich, Buchs, Switzerland) in 1 l of standard water. Two different nutrient broth concentrations (1:1000 and 1:500 dilution in the well) and sham-treatment were applied to 15 replicates each per family and treatment (n = 2,520 singly raised embryos).

Fourteen days after treatment, five eggs per family and treatment were sampled and stored at -80° C for later characterization of their associated bacterial communities. All other eggs (n = 10 per family and treatment) were transferred to new 24-well plates with 2 ml of standard water per well.

In a subsample of 20 wells per treatment, we measured oxygen concentrations (Firesting O2, Pyroscience, Aachen, Germany) as well as pH (bench pH meter FiveEasy® Plus, Life Sciences, Basel, Switzerland) every three days after the start of the treatment at a distance of 4 mm away from the egg membrane.

Statistical analyses of embryo performance

Embryo mortality was analysed as a binary response variable (dead before hatching or hatched) with logistic mixed-effect models by treating each embryo as an independent replicate. Treatment (nutrient broth concentrations) and bacterial alpha diversity on the eggs 14 days after incubation with nutrient broth (phylogenetic distance, see below for its calculation) were entered as fixed effects. Parental origin was entered as random dam, random sire, and random dam x sire interaction effects. While dam effects encompass both genetic and maternal environmental effects, sire effects represent one-quarter of the additive genetic variation, assuming that epistatic effects are negligible (Lynch & Walsh 1998). Separate models were fitted to investigate the interaction of treatment with dam or sire effects, or bacterial diversity, respectively. To test for the significance of an effect, a reduced model omitting the variable of interest was compared to the reference model. The goodness of fit of the different models is given by the logarithm of the approximated likelihood (ln L) and by the Akaike's information criterion (AIC). To test if models differ in their goodness of fit, the models were compared with likelihood ratio tests (LRT). The lme4 package v.1.1.7 for logistic mixed effect model analyses (Bates & Sarkar 2007) was used in R v.3.1.3 (R Development Core Team 2015). Time until hatching was analysed as a continuous response variable in analogous linear mixed models (LMM).

DNA extraction and preparation for sequencing

Five fertilized eggs per family and treatment were pooled for bacterial DNA extraction. Milt was analysed individually for each sire. This resulted in 175 samples (56 families x three treatments and seven milt samples). For detailed DNA extraction protocols and PCR conditions see Supplementary Information. Cleaned PCR products were sent in equimolar amounts for MiSeq illumina sequencing (Nextera protocol; Microsynth, Balgach, Switzerland) using four libraries of pooled amplicons (n = 48 samples each, combined with samples from another project in the fourth library) in one full sequencing run.

Bacterial diversity on fertilized eggs

The following analyses were done in the QIIME v.1.8.0 framework (Caporaso *et al*. 2010a). After a stringent quality control pipeline for raw next-generation sequencing reads, operational taxonomic units (OTUs) of bacteria were inferred. Different steps of the quality control pipeline and OTU picking algorithms are described in detail in the Supplementary Information. An open blast search was applied (Lan *et al*. 2012) with the RDP classifier v.2.2 (Wang *et al*. 2007) and the Greengenes reference database v.12.13 (McDonald *et al*. 2012).

Phylogenetic distance within bacterial communities on the eggs was calculated as alpha diversity measure that estimates the total descending branch length for all OTUs in a sample (Gotelli & Chao 2013). A linear mixed model using the lme4

package in R was applied to investigate the effects of nutrient broth treatment and the effects of dam and sire identity on bacterial diversity associated with fertilized brown trout eggs (14 days after treatment). Prior to this analysis, alpha diversity measures were tested for their normal distribution, homogeneity of variances and homoscedasticity of the error terms. There were no deviations from the assumption of equal variances ($F_{2,7} = 0.18$, $p = 0.9$). Significance of different explanatory variables was inferred analogously to the analysis on embryo performance above. The interaction effect of dam x sire was not analysed due to low sample sizes of replicates for this term. The relationship between bacterial diversity on the eggs and embryo performance (mortality and hatching time, respectively) was investigated with Pearson's product moment correlation (r) using mean values of brown trout families. Since mortality might bias hatching time, the same analysis was also applied to a subset of families with no mortality.

Bacterial composition on fertilized eggs

Pairwise UniFrac distances (Lozupone & Knight 2005) were calculated to quantify differences in bacterial composition among samples. UniFrac distances quantify the fraction of unique branch lengths against the total branch length between pairs of bacterial communities from a common phylogenetic dendrogram (Lozupone & Knight 2005), where 0 indicates that two samples are identical and 1 indicates that two samples have no bacterial species in common. To build a phylogenetic dendrogram, OTUs were aligned using the PyNAST algorithm v.1.2.2 (Caporaso *et al*. 2010b). A phylogenetic tree was built using FastTree v.2.0 (Price *et al*. 2008, using default parameters). Quantitative measures (*i.e*. weighted UniFac) are appropriate for revealing community differences that are due to changes in relative taxon abundance (*e.g*. when a particular set of taxa grow well because a limiting nutrient source becomes abundant). Qualitative measures (*i.e*. unweighted UniFrac) are most informative when communities differ primarily by what can live in them (*e.g*. at high temperatures), partially because abundance information might obscure significant patterns of variation in which taxa are present (Lozupone & Knight 2007). Different nutrient broth concentrations were expected to affect mostly bacterial taxa abundance and not composition on fertilized eggs. Hence, weighted UniFrac measures were used for the analysis of embryonated eggs. However different bacterial taxa might be found on the gametes of individual, natural spawners. Accordingly, unweighted UniFrac distances were applied for the analysis of bacterial communities on milt. Pairwise permutation tests (R package 'vegan', Oksanen *et al*. 2013, with the functions 'vegdist' and 'Adonis') were applied to analyse whether treatment, dam and sire effects explain a significant part of the variation in bacterial composition on the eggs. This analysis was done in the R environment and is further explained in the Supplementary Information. Principal Coordinate Analysis (PCoA) plots (2-dimensional) were created to visualize the permutation tests using EMPeror v.0.9.3 (Vazquez-Baeza *et al*. 2013).

To calculate a core microbiome on fertilized eggs, bacterial taxa were identified that are present in 90 percent of all samples. We used the QIIME script compute_core_microbiome.py for this selection and a closed blast search had to be adopted (Lan *et al*. 2012). Based on these core microbiomes, a synthetic metagenome was generated with PICRUSt v.1.0.0 (Langille *et al*. 2013) using the online Galaxy version, analogous to Wilkins *et al*. (2015b). First, the core microbiome was normalized with respect to inferred 16S rRNA gene copy numbers. This normalized OTU table was then used to predict bacterial metagenomes. Predicted metagenomes

were analysed with STAMP v.2.1.2 to visualize presumable functions of the synthetic metagenomes in the three different treatment groups (Parks *et al*. 2014). Bacterial pathways were derived from the KEGG database (Kanehisa & Goto 2000; Kanehisa *et al*. 2016). To test for significant differences we used Kruskal-Wallis H-tests, a non-parametric method for testing whether or not the medians of more than two groups are equal. This test was used to compare (i) the bacterial composition (OTUs) on fertilized eggs among different sires and (ii) bacterial gene pathways among the three different treatment groups. These two tests were based on bacterial abundance. Only OTUs were included that were present in all comparison groups. We applied a Benjamini – Hochberg multiple testing correction controlling the FDR and the threshold was set as FDR value < 0.05. The R package 'phyloseq' v.1.7.12 (McMurdie & Holmes 2013) was used to build a heatmap of the most variable bacterial taxa among the offspring of different sires.

Bacterial composition on milt

To disentangle paternal environmental effects (*i.e*. bacterial composition on milt) and paternal genetic effects, the bacterial composition was contrasted between milt before fertilization and fertilized eggs. Bacterial communities on milt before fertilization were summarized with a heatmap of the 40 most abundant core bacterial taxa, analogously to the bacterial composition on fertilized eggs above. Analyses of variance were performed to compare bacterial communities on milt and fertilized eggs in the same ways as the analyses above using 'vegan' and 'Adonis'. Student's two-sample *t*-tests of pairwise means were calculated to determine whether means were significantly different from each other. 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. Coordinate Analysis (PCoA) plots (2-dimensional) were created to visualize the permutation tests.

Results

Embryo mortality

Embryo mortality increased with elevated bacterial resources (Fig. 1a, Table 1) and was different among the offspring of different mothers and fathers (Table 1). We found significant effects of dam (but not sire) identity on the susceptibility to the stress treatment (models 5 and 6 in Table 1). Bacterial diversity and possible dam x sire interactions did not seem to affect embryo mortality (models 2, 7 and 8 in Table 1).

Time until hatching

Time until hatching was not significantly different among treatments (Fig. 1b, Table 2). However, embryos differed in their average hatching time depending on their mother and their father (Table 2). Time until hatching could be predicted by the bacterial diversity on the eggs 14 days after treatment (Table 2, Fig. 2), but there were no significant interaction effects either between treatments and parental origin or between dams and sires (Table 2).

Oxygen and pH

pH measurements were not significantly different among treatments ($F_{2,59} = 0.05$, $p = 0.98$; Fig. S1). Oxygen concentrations decreased steadily during embryogenesis, reached a minimum before hatching, and recovered after hatching. Oxygen

concentrations were not significantly different among treatments ($F_{2,59}$ = 1.5, p = 0.19).

Bacterial diversity on fertilized eggs

MiSeq sequencing of the four libraries resulted in a total of 45,588,998 reads before quality control. After the splitting of reads according to primer sequences and barcodes a total of 19,347,049 reads could be retained (μ = 105,722±51,880 per sample; Table S1 and S2). After quality control 8,916,790 reads remained (μ = 48,726±24,295 per sample) including a mean of 1,874 (±747) OTUs per sample. Rarefaction curves of alpha diversities are shown in Fig. S2. Bacterial diversities on fertilized eggs were not significantly affected by the nutrient broth treatment (Fig. 1c, Table 3 and S3). Bacterial diversity was significantly different among the offspring of different sires while maternal effects played no significant role (Table 3).

We found no significant relationship between bacterial diversity on the eggs and embryo mortality within the three treatments (control: r = 0.02, t_{54} = 0.15, p = 0.88; NB 1:000: r = 0.14, t_{54} = 1.02, p = 0.31; NB 1:500: r = 0.21, t_{54} = 1.52, p = 0.14). However, bacterial diversity correlated with time until hatching (r = 0.53, t_{54} = 4.6, p < 0.0001; Fig. 2). This was still true when only families were analysed with no mortality (r = 0.61, t_{28} = 4.4, p < 0.0001; Fig. S3).

Bacterial composition on fertilized eggs

Bacterial communities on fertilized eggs differed significantly in composition among the three nutrient broth concentrations ($F_{2,55}$ = 10.4, R^2 = 0.07, p < 0.001; Fig. 1d) and among different parents (dam: $F_{7,6}$ = 7.3, R^2 = 0.18, p < 0.001; sire: $F_{6,7}$ = 9.4, R^2 = 0.21, p < 0.001; Fig. 3 and 4, Table 4).

We found 95 core bacterial taxa on fertilized eggs 14 days after treatment (Fig. S4). For graphical representations of egg-associated bacterial communities, heatmaps were plotted at the genus level (Fig. S5a, due to readability only the 40 most common core bacteria are shown). Contrasts in the putative, functional potential of bacterial communities among nutrient broth concentrations are listed in Table S4.

Bacterial composition on milt

Bacterial communities on fertilized eggs (Fig. S5a) and milt (Fig. S5b) were significantly different from each other (Adonis: p < 0.001, R^2 = 89.2%, $F_{1,6}$ = 13.9; two-sided Student's two-sample t-test: t_1 = 8.9, p < 0.001). Ordering the axes according to bacterial community distances of milt *vs*. fertilized eggs explained 57% of the variance in bacterial composition in the PCoA analysis (Fig. 3).

Discussion

We found that embryo genetics defines egg-associated bacterial diversity and composition. When we tested for potential paternal carry-over effects, we discovered a great diversity of bacteria in milt that did, however, not correspond to egg-associated bacterial communities at later developmental stages. Bacterial diversity on the eggs was correlated with time until hatching. This correlation could be driven by attributes of the egg-associated bacteria or by consequences of the switch in host life-history. These two hypotheses are non-exclusive but empirical data supports the first hypothesis. Changes in life-history traits of salmonid embryos; *e.g*. pathogen effects on the timing of hatching in particular, have been observed in similar experimental set-ups (Thorpe *et al*. 1998; Hale 1999; Pompini *et al*. 2013; Clark *et al*. 2014).

Embryo mortality

Embryo mortality increased with elevated bacterial resources, confirming previous findings in *C. palaea* (Wedekind *et al*. 2010; Clark *et al*. 2013a) and *S. trutta* (Jacob *et al*. 2010). Mortality could not be directly linked to oxygen depletion or to a change in acidity in the wells. Moreover, oxygen levels in the wells never reached a lower limit known to obstruct embryo development (Wedekind & Müller 2004). This supports the hypotheses that either a change in bacterial life-history or in bacterial community composition has detrimental effects on embryo hosts. Contrary to previous observations in whitefish (Clark *et al*. 2013a; Clark *et al*. 2014), we did not find any treatment effects on the timing of hatching.

Environmental effects

We found that the composition of bacterial communities and their putative functional pathways were dependent on the availability of bacterial resources. Increased host mortality at elevated nutrient concentrations suggests that changes in microbial communities affected bacterial virulence. Bacteria with a pathogenic potential could be found on milt, as well as on fertilized eggs. Some of the bacterial species found on fertilized eggs are known to harm their salmonid hosts. For example, *Aeromonas* spp. can cause septicaemia or ulcer disease in many freshwater fishes (Austin & Austin 2007), and several studies on various salmonids have reported pathogenic effects of infection by *Lactococcus garvieae* and *Streptococcus iniae* (Eldar & Ghittino 1999), *Pseudomonas fluorescens* (von Siebenthal *et al*. 2009; Clark *et al*. 2013a), *Stenotrophomonas maltophilia* (Looney *et al*. 2009), or species of the genus *Staphylococcus* (Craig & Pilcher 1966; Gil *et al*. 2000). Nutrient broth might affect bacterial communities by creating competition for increased resources (Rasche *et al*. 2011). Bacterial species have shown to differ in their efficiencies of converting nutrients into growth (Weintraub *et al*. 2007). A change in density of particular groups of bacteria could lead to a transition in their life-history strategies that turns benign bacteria into virulent ones (Diggle *et al*. 2007; Wedekind *et al*. 2010). Another strategy that some bacteria use to outcompete others is the production of antibiotic compounds (Cordero *et al*. 2012). Antibiotic production comes with a cost that may only be affordable if resources are prevalent (Morlon 2012). Antimicrobial compounds or their metabolites can cause harm to the embryo hosts as a side effect (Hill *et al*. 2005).

We did not find evidence for an interaction between environmental effects and host genetics but changes in bacterial community composition were accompanied by switches in the bacteria's functional potential. Putative functional gene pathways of bacteria that were enriched at high nutrient broth concentrations included, for example, the biosynthesis of bacterial toxins. The involvement of this pathway aligns well with competition among bacteria and detrimental effects to the embryo host (Chow *et al*. 2014; Evans & Wallenstein 2014). Further harmful combinations for the embryo host comprised gene pathways of pertussis, cancer, influenza or toxoplasmosis. This list of putative bacterial gene pathways can be used for future hypothesis testing and experimental designs using the salmonid embryo host system.

Paternal effects

Dam effects were found for treatment-induced mortality, time until hatching, and bacterial community composition. Dam effects represent a mix of maternal genetic and maternal environmental effects (Royle *et al*. 1999; von Siebenthal *et al*. 2009; Aykanat *et al*. 2012). For example, maternal effects can include the expression of

specific zygotic genes as well as supplements, such as proteins or mRNA that the mother allocated to the eggs before spawning (D'alba *et al*. 2010).

Sire effects were found for embryo mortality and time until hatching; *i.e*. there was additive genetic variance for viability under our experimental conditions. Such sire effects can be interpreted as the immune-competence of the embryo (Evans *et al*. 2010; Clark *et al*. 2013a; Wilkins *et al*. 2015c), its genetic load or overall genetic variability (Neff & Pitcher 2005), and its genetic quality (Neff & Pitcher 2005; Wedekind *et al*. 2008; Jacob *et al*. 2010). We also found significant sire effects on egg-associated bacterial diversity and composition. Here, the host may either actively influence its associated bacteria through immune gene expression (Finn 2007; Clark *et al*. 2013a; Wilkins *et al*. 2015c) or passively through secondary metabolites (Milligan-Myhre *et al*. 2011; Stephens *et al*. 2016) and surface receptors coded by the host genotype (Llewellyn *et al*. 2014). Forty-nine bacterial taxa (OTUs) differed in abundance among the offspring of different sires. This included many environmental bacteria with known origins in freshwater systems, such as *Brevundimonas diminuta* (Han & Andrade 2005), members of the family of Caulobacteraceae (Newton *et al*. 2011), taxa in the genera of *Comamonas, Rhodoferax, Hydrogenophaga, Limnohabitans, Delftia* and *Devosia* (Madigan *et al*. 2010; Zhang *et al*. 2012), as well as *Flavobacterium* (Newton *et al*. 2011) and the family of Hyphomonadaceae (Madigan *et al*. 2010; Newton *et al*. 2011). Nevertheless, the distribution of these bacteria on salmonid eggs could be attributed to significant genetic effects, a finding that underlines the relevance of our results.

Our experiment demonstrates that environmental effects and host genetic factors characterize the fish egg-bacteria ecosystem. Assessing the causal role of host genetic variation on bacterial community dynamics will help us understand the mechanisms of colonization and the correlation to specific functions of host-associated bacterial communities.

Acknowledgements

We thank S. Blaser, G. Brazzola, E. Clark, M. dos Santos, P. Engel, C. Heiniger, A. Liesch, L. Marques da Cunha, Y. Poirier, M. Pompini, T. Reusch, N. Reusser, A. Rogivue, J. van der Meer, and F. Witsenburg for assistance and/or discussion; U. Gutmann, B. Bracher, and M. Schmid (Fischereistützpunkt Reutigen) for catching the fish and stripping them; and C. Küng (Fisheries Inspectorate Berne) for permissions. This study complied with the relevant ethical regulations imposed by the University of Lausanne and the canton in which it was carried out. Additional authorizations from the cantonal veterinary offices were not required as all manipulations on embryos were performed prior to yolk sac absorption. The Swiss National Science Foundation provided funding for this study. L. Wilkins was supported by a fellowship in Life Sciences (Faculty of Biology and Medicine, University of Lausanne).

References

Austin B, Austin DA (2007) *Bacterial fish pathogens, disease 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**, 691-703.

Bates D, Sarkar D (2007) lme4: Linear mixed-effects models using S4 classes. In: *R package version 1.1.7*.

Bolnick DI, Snowberg LK, Hirsch PE*, et al.* (2014a) Individuals' diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch). *Ecology Letters* **17**, 979-987.

Bolnick DI, Snowberg LK, Caporaso JG*, et al.* (2014b) Major Histocompatibility Complex class IIb polymorphism influences gut microbiota composition and diversity. *Molecular Ecology* **23**, 4831-4845.

Boutin S, Bernatchez L, Audet C, Derome N (2012) Antagonistic effect of indigenous skin bacteria of brook charr (*Salvelinus fontinalis*) against *Flavobacterium columnare* and *F. psychrophilum*. *Veterinary Microbiology* **155**, 355-361.

Boutin S, Sauvage C, Bernatchez L, Audet C, Derome N (2014) Inter individual variations of the fish skin microbiota: host genetics basis of mutualism? *PLoS One* **9**, e102649.

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

Caporaso JG, Bittinger K, Bushman FD*, et al.* (2010b) PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**, 266-267.

Chow CET, Kim DY, Sachdeva R, Caron DA, Fuhrman JA (2014) Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *ISME Journal* **8**, 816-829.

Clark E, Wedekind C (2011) Additive genetic effects on embryo viability in a whitefish (Salmonidae) influenced by the water mould *Saprolegnia ferax*. *Journal of Bacteriology and Parasitology* **10**, 2-6.

Clark ES, Wilkins LGE, Wedekind C (2013a) MHC class I expression dependent on bacterial infection and parental factors in whitefish embryos (Salmonidae). *Molecular Ecology* **22**, 5256-5269.

Clark ES, Stelkens RB, Wedekind C (2013b) Parental influences on pathogen resistance in brown trout embryos and effects of outcrossing within a river network. *PLoS One* **8**, e57832.

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. *Functional Ecology* **28**, 714-723.

Clark ES, Pompini M, Uppal A, Wedekind C (2016) Genetic correlations and little genetic variance for reaction norms may limit potential for adaptation to pollution by ionic and nanoparticulate silver in a whitefish (Salmonidae). *Ecology and Evolution*, in press.

Cordero OX, Wildschutte H, Kirkup B*, et al.* (2012) Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. *Science* **337**, 1228-1231.

Craig JM, Pilcher KS (1966) *Clostridium botulinum* Type F: Isolation from salmon from the Columbia River. *Science* **153**, 311-312.

D'alba L, Shawkey MD, Korsten P*, et al.* (2010) Differential deposition of antimicrobial proteins in blue tit (*Cyanistes caeruleus*) clutches by laying order and male attractiveness. *Behavioral Ecology and Sociobiology* **64**, 1037-1045.

Diggle SP, Griffin AS, Campbell GS, West SA (2007) Cooperation and conflict in quorum-sensing bacterial populations. *Nature* **450**, 411-417.

Eldar A, Ghittino C (1999) *Lactococcus garvieae* and *Streptococcus iniae* infections in rainbow trout *Oncorhynchus mykiss*: similar, but different diseases. *Diseases of Aquatic Organisms* **36**, 227-231.

Evans ML, Neff BD, Heath DD (2010) MHC-mediated local adaptation in reciprocally translocated Chinook salmon. *Conservation Genetics* **11**, 2333-2342.

Evans SE, Wallenstein MD (2014) Climate change alters ecological strategies of soil bacteria. *Ecology Letters* **17**, 155-164.

Finn RN (2007) The physiology and toxicology of salmonid eggs and larvae in relation to water quality criteria. *Aquatic Toxicology* **81**, 337-354.

Ghanbari M, Kneifel W, Domig KJ (2015) A new view of the fish gut microbiome: Advances from next-generation sequencing. *Aquaculture* **448**, 464-475.

Gil P, Vivas J, Gallardo CS, Rodriguez LA (2000) First isolation of *Staphylococcus warneri,* from diseased rainbow trout, *Oncorhynchus mykiss* (Walbaum), in Northwest Spain. *Journal of Fish Diseases* **23**, 295-298.

Givens CE, Ransom B, Bano N, Hollibaugh JT (2014) A fish tale: comparison of the gut microbiomes of 12 finfish and 3 shark species. *Marine Ecology Progress Series* **518**, 209-223.

Gotelli NJ, Chao A (2013) Measuring and estimating species richness, species diversity, and biotic similarity from sampling data. In: *Encyclopedia of Biodiversity* (ed. Levin SA), pp. 195-211. Academic Press, Waltham, MA.

Hale ME (1999) Locomotor mechanics during early life history: Effects of size and ontogeny on fast-start performance of salmonid fishes. *Journal of Experimental Biology* **202**, 1465-1479.

Han XY, Andrade RA (2005) *Brevundimonas diminuta* infections and its resistance to fluoroquinolones. *Journal of Antimicrobial Chemotherapy* **55**, 853-859.

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

Hill AJ, Teraoka H, Heideman W, Peterson RE (2005) Zebrafish as a model vertebrate for investigating chemical toxicity. *Toxicological Sciences* **86**, 6-19.

Jacob A, Nusslé S, Britschgi A*, et al.* (2007) Male dominance linked to size and age, but not to 'good genes' in brown trout (*Salmo trutta*). *BMC Evolutionary Biology* **7**, 207.

Jacob A, Evanno G, von Siebenthal BA, Grossen C, Wedekind C (2010) Effects of different mating scenarios on embryo viability in brown trout. *Molecular Ecology* **19**, 5296-5307.

Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **28**, 27-30.

Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research* **44**, D457-462.

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**, e3249.

Langille MGI, Zaneveld J, Caporaso JG*, et al.* (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology* **31**, 814-821.

Larsen A, Tao Z, Bullard SA, Arias CR (2013) Diversity of the skin microbiota of fishes: evidence for host species specificity. *FEMS Microbiology Ecology* **85**, 483-494.

Li XM, Yu YH, Feng WS, Yan QY, Gong YC (2012) Host species as a strong determinant of the intestinal microbiota of fish larvae. *Journal of Microbiology* **50**, 29-37.

Li XM, Yan QY, Xie SQ*, et al.* (2013) Gut microbiota contributes to the growth of fast-growing transgenic Common Carp (*Cyprinus carpio* L.). *PLoS One* **8**, e64577.

Liu Y, de Bruijn I, Jack AL*, et al.* (2014) Deciphering microbial landscapes of fish eggs to mitigate emerging diseases. *ISME Journal* **8**, 2002-2014.

Llewellyn MS, Boutin S, Hoseinifar SH, Derome N (2014) Teleost microbiomes: the state of the art in their characterization, manipulation and importance in aquaculture and fisheries. *Frontiers in Microbiology* **5**, Article 207.

Looney WJ, Narita M, Muhlemann K (2009) *Stenotrophomonas maltophilia*: an emerging opportunist human pathogen. *Lancet Infectious Diseases* **9**, 312-323.

Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology* **71**, 8228-8235.

Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 11436-11440.

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

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

McDonald D, Price MN, Goodrich J*, et al.* (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME Journal* **6**, 610-618.

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

Milligan-Myhre K, Charette JR, Phennicie RT*, et al.* (2011) Study of host-microbe interactions in zebrafish. *Zebrafish: Disease models and chemical screens, 3rd edn.* **105**, 87-116.

Morlon H (2012) Microbial Cooperative Warfare. *Science* **337**, 1184-1185.

Navarrete P, Magne F, Araneda C*, et al.* (2012) PCR-TTGE analysis of 16S rRNA from Rainbow trout (*Oncorhynchus mykiss*) gut microbiota reveals host-specific communities of active bacteria. *PLoS One* **7**, e31335.

Neff BD, Pitcher TE (2005) Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Molecular Ecology* **14**, 19-38.

Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S (2011) A guide to the natural history of freshwater lake bacteria. *Microbiology and Molecular Biology Reviews* **75**, 14-49.

OECD (1992) OECD Guideline for the testing of chemicals 203 (fish, acute toxicity test), p. 9.

Oksanen J, Blanchet FG, Kindt R*, et al.* (2013) vegan: Community Ecology Package. R package version 2.2.1.

Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* **30**, 3123-3124.

Pompini M, Clark ES, Wedekind C (2013) Pathogen-induced hatching and population-specific life-history response to waterborne cues in brown trout (*Salmo trutta*). *Behavioral Ecology and Sociobiology* **67**, 649-656.

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

R Development Core Team R (2015) A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rajaram S, Oono Y (2010) NeatMap - non-clustering heat map alternatives in R. *BMC Bioinformatics*, 11:45.

Rasche F, Knapp D, Kaiser C*, et al.* (2011) Seasonality and resource availability control bacterial and archaeal communities in soils of a temperate beech forest. *ISME Journal* **5**, 389-402.

Roeselers G, Mittge EK, Stephens WZ*, et al.* (2011) Evidence for a core gut microbiota in the zebrafish. *ISME Journal* **5**, 1595-1608.

Royle NJ, Surai PF, McCartney RJ, Speake BK (1999) Parental investment and egg yolk lipid composition in gulls. *Functional Ecology* **13**, 298-306.

Schmidt VT, Smith KF, Melvin DW, Amaral-Zettler LA (2015) Community assembly of a euryhaline fish microbiome during salinity acclimation. *Molecular Ecology* **24**, 2537-2550.

Sevellec M, Pavey SA, Boutin S*, et al.* (2014) Microbiome investigation in the ecological speciation context of lake whitefish (*Coregonus clupeaformis*) using next-generation sequencing. *Journal of Evolutionary Biology* **27**, 1029-1046.

Skrodenyte-Arbaciauskiene V, Sruoga A, Butkauskas D (2006) Assessment of microbial diversity in the river trout *Salmo trutta fari*o L. intestinal tract identified by partial 16S rRNA gene sequence analysis. *Fisheries Science* **72**, 597-602.

Star B, Haverkamp THA, Jentoft S, Jakobsen KS (2013) Next generation sequencing shows high variation of the intestinal microbial species composition in Atlantic cod caught at a single location. *BMC Microbiology* **13**.

Stephens WZ, Burns AR, Stagaman K*, et al.* (2016) The composition of the zebrafish intestinal microbial community varies across development. *ISME Journal* **10**, 644-654.

Sullam KE, Essinger SD, Lozupone CA*, et al.* (2012) Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Molecular Ecology* **21**, 3363-3378.

Sullam KE, Rubin BER, Dalton CM*, et al.* (2015) Divergence across diet, time and populations rules out parallel evolution in the gut microbiomes of Trinidadian guppies. *ISME Journal* **9**, 1508-1522.

Sun YZ, Yang HL, Ling ZC, Chang JB, Ye JD (2009) Gut microbiota of fast and slow growing grouper *Epinephelus coioides*. *African Journal of Microbiology Research* **3**, 713-720.

Thorpe JE, Mangel M, Metcalfe NB, Huntingford FA (1998) Modelling the proximate basis of salmonid life-history variation, with application to Atlantic salmon, *Salmo salar* L. *Evolutionary Ecology* **12**, 581-599.

Vazquez-Baeza Y, Pirrung M, Gonzalez A, Knight R (2013) EMPeror: a tool for visualizing high-throughput microbial community data. *Gigascience* **2**, 16.

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 and Shellfish Immunology* **26**, 531–535.

Waldor MK, Tyson G, Borenstein E*, et al.* (2015) Where next for microbiome research? *PLoS Biology* **13**, e1002050.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**, 5261-5267.
Wedekind C (2002) Induced hatching to avoid infectious egg disease in whitefish. *Current Biology* **12**, 69-71.
Wedekind C, Walker M, Portmann J*, et al.* (2004) MHC-linked susceptibility to a bacterial infection, but no MHC-linked cryptic female choice in whitefish. *Journal of Evolutionary Biology* **17**, 11-18.
Wedekind C, Müller R (2004) The experimental rearing of large salmonid eggs in Petri dishes. *Functional Ecology* **18**, 138-140.
Wedekind C, Jacob A, Evanno G, Nusslé S, Müller R (2008) Viability of brown trout embryos positively linked to melanin-based but negatively to carotenoid-based colours of their fathers. *Proceedings of the Royal Society B-Biological Sciences* **275**, 1737-1744.
Wedekind C, Gessner MO, Vazquez F, Märki M, Steiner D (2010) Elevated resource availability sufficient to turn opportunistic into virulent fish pathogens. *Ecology* **91**, 1251-1256.
Weintraub MN, Scott-Denton LE, Schmidt SK, Monson RK (2007) The effects of tree rhizodeposition on soil exoenzyme activity, dissolved organic carbon, and nutrient availability in a subalpine forest ecosystem. *Oecologia* **154**, 327-338.
Wilkins LGE, Rogivue A, Fumagalli L, Wedekind C (2015a) Declining diversity of egg-associated bacteria during development of naturally spawned whitefish embryos (*Coregonus spp.*) *Aquatic Sciences* **77**, 481-497.
Wilkins LGE, Rogivue A, Schütz F, Fumagalli L, Wedekind C (2015b) Increased diversity of egg-associated bacteria on brown trout (*Salmo trutta*) at elevated temperatures *Scientific Reports* **5**, 17084.
Wilkins LGE, Clark ES, Farinelli L, Wedekind C, Fumagalli L (2015c) Embryonic gene expression of *Coregonus palaea* (whitefish) under pathogen stress as analyzed by high-throughput RNA-sequencing. *Fish and Shellfish Immunology* **47**, 130-140.
Wilson B, Danilowicz BS, Meijer WG (2008) The diversity of bacterial communities associated with Atlantic Cod *Gadus morhua*. *Microbial Ecology* **55**, 425-434.
Zhang DC, Redzic M, Liu HC*, et al.* (2012) *Devosia psychrophila* sp. nov. and *Devosia glacialis* sp. nov., from alpine glacier cryoconite, and an emended description of the genus *Devosia*. *International Journal of Systematic and Evolutionary Microbiology* **62**, 710-715.

Data Accessibility

Data on embryo performance, OTU abundance and identity for each sample, and a map file as used for the QIIME analysis will the deposited on the Dryad repository after acceptance of the MS.

Author Contributions: L.F., C.W., and L.W. designed the project. L.W. and C.W. executed the experiment. L.W. performed the molecular genetic and the statistical analyses. All authors participated in the discussion of the results. L.W. and C.W. wrote the manuscript that was then critically revised L.F.

Tables and Figures

Table 1: The influence of treatment, dam and sire effects, and bacterial diversity on embryo survival.
Four main logistic mixed effect models were compared to a reference model (in bold) to test whether treatment (T), dam (D) and sire (S) effects, and bacterial alpha diversity (A) on the eggs explain a significant part of the variance in embryo survival. Four additional models were fitted to investigate the interaction terms of treatment with dam, sire, or bacterial alpha diversity respectively; as well as the interaction of dam and sire effects.

		Model parameters				**Likelihood ratio tests**[a]		
Model	**Effects**	**Fixed**	**Random**	**AIC**	**ln L**	**δAIC**	**χ^2**	***p***
Reference		**T, A**	**D, S**	**1115.7**	**-551.9**			
1	T	A	D, S	1176.3	-584.1	60.6	64.5	< 0.0001
2	A	T	D, S	1113.7	-551.9	2	0.01	0.91
3	D	T, A	S	1226.4	-608.2	110.7	112.7	< 0.0001
4	S	T, A	D	1124.9	-557.4	9.2	11.1	< 0.001
5	T x D	T, A	S	1112.1	-545.1	3.6	13.8	0.02
6	T x S	T, A	D	1125.2	-551.6	9.5	0.5	0.99
7	T x A	T, A	D, S	1117.1	-550.6	1.4	2.6	0.27
8	D x S	T, A	D, S	1160.3	-540.2	44.6	23.4	0.91

[a] Degrees of freedom = 1.

Table 2: The influence of treatment, dam and sire effects, and bacterial diversity on time until hatching.
Different linear mixed effect models were compared to a reference model (in bold) as in Table 1.

		Model parameters				Likelihood ratio tests[a]		
Model	**Effects**	**Fixed**	**Random**	**AIC**	**ln L**	**δAIC**	**χ^2**	***p***
Reference		**T, A**	**D, S**	**5609.5**	**-2798.8**			
1	T	A	D, S	5608	-2799	1.5	0.51	0.47
2	A	T	D, S	5611.5	-2800.2	2	3.1	0.04
3	D	T, A	S	5663.1	-2826.6	53.6	55.6	< 0.0001
4	S	T, A	D	5669	-2829.5	59.5	61.5	< 0.0001
5	T x D	T, A	S	5613.4	-2798.8	3.9	0.09	0.95
6	T x S	T, A	D	5611.9	-2797.9	2.4	1.65	0.43
7	T x A	T, A	D, S	5610.4	-2798.2	0.9	1.1	0.3
8	D x S	T, A	D, S	5647.1	-2783.6	37.6	30.4	0.65

[a] Degrees of freedom = 1.

Table 3: The influence of treatment, dam and sire effects on bacterial diversity. Different linear mixed effect models were compared to a reference model (in bold) as in Table 1

		Model parameters				Likelihood ratio tests[a]		
Model	**Effects**	**Fixed**	**Random**	**AIC**	**ln L**	**δAIC**	**χ^2**	***p***
Reference		**T**	**D, S**	**1225.9**	**-606.95**			
1	T	-	D, S	1222.3	-607.17	3.6	0.42	0.81
2	D	T	S	1225.3	-607.66	0.6	1.41	0.23
3	S	T	D	1240.5	-615.27	14.6	16.64	< 0.0001
4	T x D	T	S	1227.4	-606.72	1.5	0.56	0.75
5	T x S	T	D	1227.7	-606.85	1.8	0.29	0.86

[a] Degrees of freedom = 1.

Table 4: Analyses of variance in bacterial community distances 14 days after treatment.

The effects of treatment, parental origin and their interactions were tested in one multivariate analysis of variance using weighted UniFrac distance matrices among bacterial communities on eggs as the dependent variable with permutation tests for significance (function 'Adonis' in R package 'vegan').

Model terms	**d.f.**	**SS**	**F**	**R^2**	***p***
Treatment	2	0.22	10.36	0.07	< 0.001
Dam	7	0.53	7.33	0.18	< 0.001
Sire	6	0.59	9.37	0.21	< 0.001
Treatment x Dam	14	0.18	1.22	0.06	0.09
Treatment x Sire	12	0.11	0.91	0.03	0.66
Residuals	126	1.31		0.44	
Total	167	2.94		1	

Figure 1: Effects of treatment on embryo life-history and associated microbiota.
Effects of different dilutions of nutrient broth (NB; control = sham treated) on embryo mortality (a), time until hatching (b), egg-associated bacterial diversity (c; measured as phylogenetic distance), and bacterial community distances (d; weighted UniFrac distances). Means and 95% confidence intervals; n.s. = not significant; ** $p < 0.01$; *** $p < 0.001$; see text for statistics.

Fig. 1

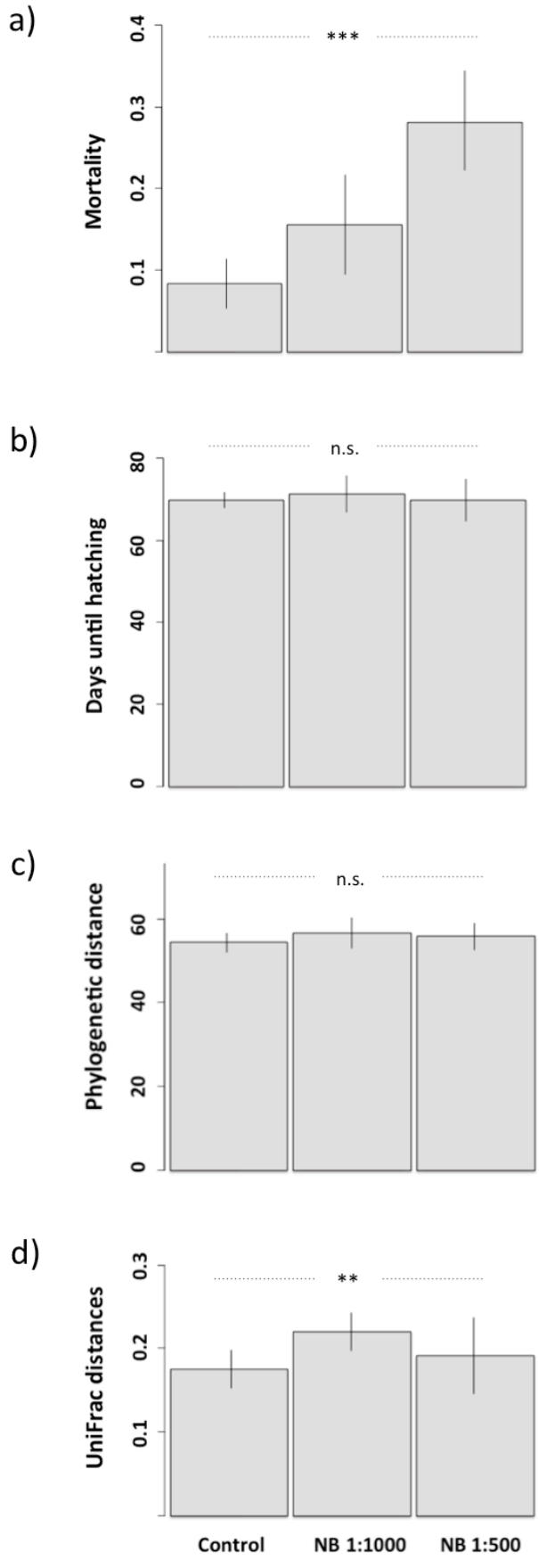

a)
Mortality
0.1
0.2
0.3
0.4
***
b)
Days until hatching
0
20
40
60
80
n.s.
c)
Phylogenetic distance
0
20
40
60
n.s.
d)
UniFrac distances
0.1
0.2
0.3
**
Control
NB 1:1000
NB 1:500

Figure 2: Relationship of bacterial diversity and hatching time.
Bacterial alpha diversity was estimated as phylogenetic distance. Means per sire and 95% confidence intervals.

Fig. 2

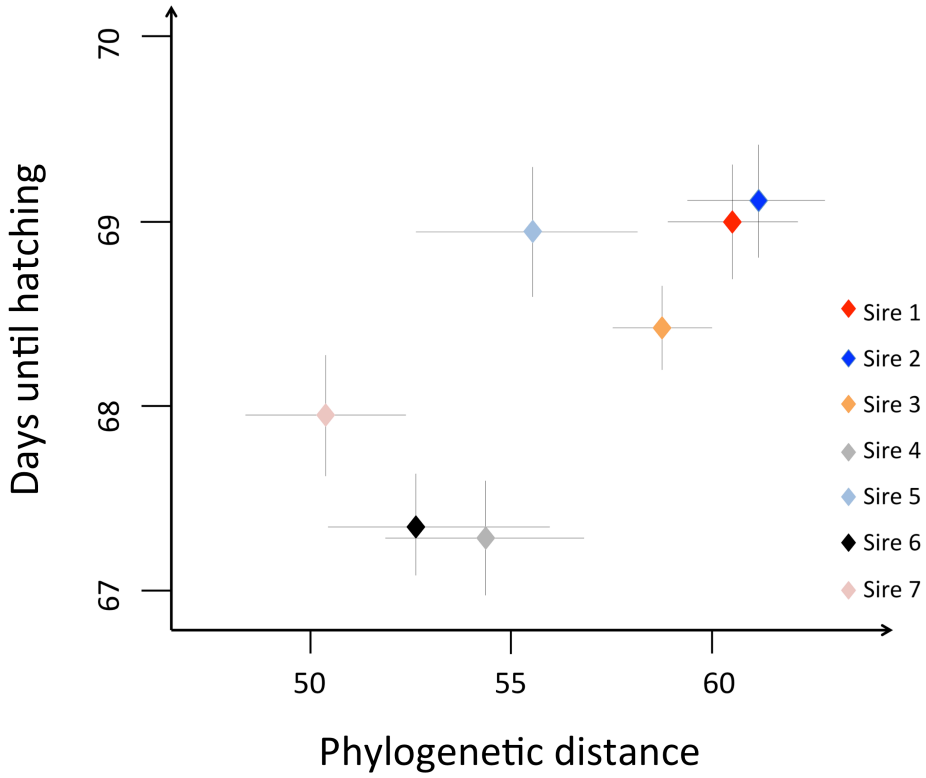


Figure 3: Bacterial community distances on brown trout milt and fertilized eggs. Two-dimensional principal coordinate analysis showing the two axes that explain most of the variance in bacterial composition (percentages explained are given at each axis) on milt (circles) versus on fertilized eggs (squares; a-c); and on embryonated eggs (d-i). Sham controls (panels a, d, g), and nutrient broth at a concentration of 1:1000 (b, e, h) and of 1:500 (c, f, i). Colours and symbols correspond to the ones used in Fig. 2.

Fig. 3

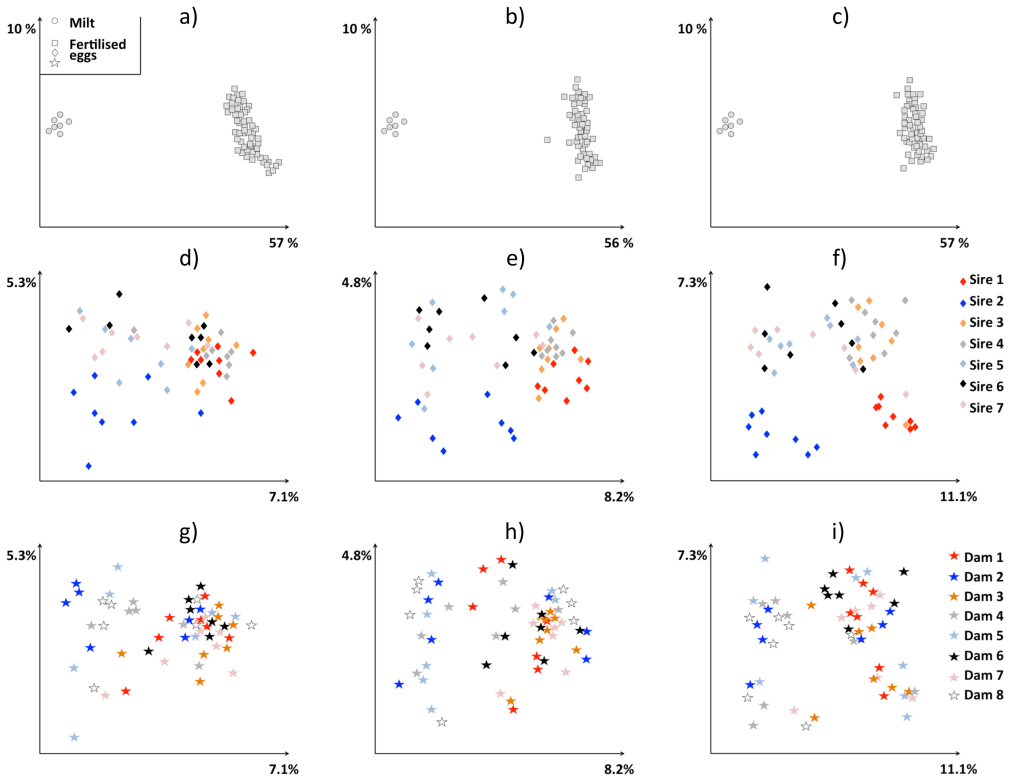
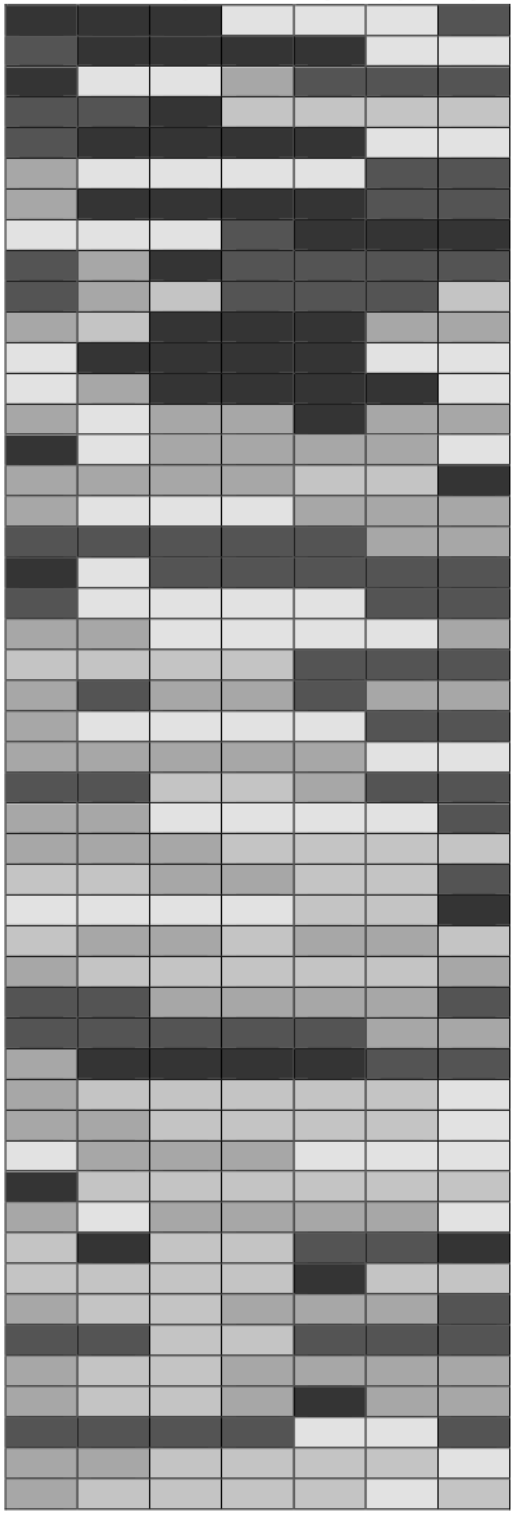


Figure 4: Differences in bacterial community composition among the offspring of different sires.
Bacterial taxa (OTUs: operational taxonomic units) that differed significantly in abundance among the offspring of different sires 14 days after treatment. The abundance of OTUs is given in shades of grey (dark = high abundance, light = low abundance). Samples are ordered according to multidimensional scaling (Rajaram & Oono 2010). *p*-values refer to a significant difference in abundance of bacteria (a Benjamini – Hochberg multiple testing correction was applied controlling the false discovery rate). Effect sizes of medians among the offspring of different sires are shown.

Fig. 4



OTUs	Sire 1	7	6	2	3	5	8	p-values (corrected)	Effect size
Alphaproteobacteria								< 0.0001	0.43
Hyphomonadaceae								< 0.0001	0.41
Devosia								< 0.0001	0.4
Oxalobacteraceae								< 0.0001	0.38
Paucibacter								< 0.0001	0.35
Mycoplana								< 0.0001	0.33
Comamonadaceae								< 0.0001	0.33
Comamonadaceae								< 0.0001	0.28
Novosphingobium								< 0.0001	0.27
Hydrogenophaga								< 0.0001	0.26
Caulobacteraceae								< 0.0001	0.25
Oxalobacteraceae								< 0.0001	0.24
Comamonadaceae								< 0.0001	0.24
Gallionella								< 0.0001	0.23
Comamonadaceae								< 0.0001	0.21
Dechloromonas								< 0.0001	0.2
Comamonadaceae								< 0.0001	0.19
Comamonadaceae								< 0.0001	0.18
Nevskia ramosa								0.0001	0.18
Rhodocyclaceae								0.0001	0.18
Dechloromonas								0.0002	0.17
Sphingomonadaceae								0.0003	0.17
Brevundimonas diminuta								0.0003	0.16
Sphingobium								0.0004	0.16
Comamonadaceae								0.0005	0.16
Methylotenera mobilis								0.0006	0.16
Novosphingobium								0.0006	0.15
Sphingomonas								0.0009	0.15
Sphingobium								0.002	0.14
Caulobacteraceae								0.002	0.14
Sphingopyxis alaskensis								0.004	0.13
Mycoplana								0.008	0.12
Rhodoferax								0.008	0.12
Devosia								0.009	0.12
Pseudomonadaceae								0.01	0.12
Acidovorax								0.01	0.12
Oxalobacteraceae								0.01	0.11
Sediminibacterium								0.01	0.11
Rhodoferax								0.01	0.11
Pseudomonas								0.01	0.11
Corynebacterium								0.02	0.11
Oxalobacteraceae								0.02	0.1
Arthrospira								0.02	0.1
Mycoplana								0.02	0.1
Caulobacteraceae								0.03	0.1
Caulobacteraceae								0.04	0.09
Caulobacter								0.04	0.09
Pseudomonas								0.04	0.08
Azoarcus								0.04	0.08

Supplementary Material

Effects of host genetics and environment on egg-associated microbiota in brown trout (*Salmo trutta*)

Laetitia G.E. Wilkins[*], Luca Fumagalli[a] and Claus Wedekind[a]

Department of Ecology and Evolution, Biophore, University of Lausanne, 1015 Lausanne, Switzerland

[*]Correspondence to L. Wilkins, Tel: +41 21 692 42 19; Fax: +41 21 692 42 65; email: laetitia.wilkins@unil.ch

[a]Joint senior authors

Supplementary Tables and Figures

Table S1: Number of reads after the different steps of quality control.

'Splitted' gives the number of reads after grouping reads according to primer sequences with individual barcodes, while 'Clustered' gives the remaining number of reads after filtering out all sequences that had failed quality control. The number of different Operational Taxonomic Units 'Nr. Of OTUs' is given that were assigned with the RDP classifier 2.2 (Wang *et al*. 2007) and the Greengenes reference database version 12.13 (McDonald *et al*. 2012). The identity of each Sample ID is explained in Table S2.

Sample ID	Splitted	Clustered	Nr. of OTUs
1	131,251	62,186	2,362
2	89,195	41,635	1,654
3	90,545	42,812	1,715
4	93,183	43,811	1,678
5	96,538	48,532	1,875
6	90,819	42,370	1,719
7	109,786	48,861	1,890
8	118,756	59,241	2,522
9	109,478	48,705	2,158
10	87,929	39,805	1,141
11	108,324	49,283	2,091
12	72,250	32,474	1,400
13	89,928	43,180	1,868
14	87,763	40,026	2,282
15	85,989	36,986	1,655
16	115,813	55,958	2,087
17	116,269	51,298	2,141
18	83,931	36,811	1,754
19	91,635	40,613	1,731
20	104,960	45,913	1,961
21	97,539	45,720	1,663
22	86,292	38,421	1,499
23	97,638	42,181	1,607
24	104,153	49,527	2,196
25	113,912	52,593	1,884
26	89,315	41,022	1,598
27	92,193	43,173	1,909
28	123,399	56,395	2,095
29	97,484	47,197	2,339
30	115,547	53,704	2,508
31	80,662	35,553	1,446
32	112,663	55,149	2,360
33	113,157	52,990	2,091
34	77,737	35,635	1,441
35	91,304	43,066	1,974
36	88,969	40,980	1,879
37	100,161	49,746	2,259

38	114,104	53,757	2,638
39	111,716	50,623	2,116
40	109,644	54,710	2,149
41	123,497	54,810	2,379
42	70,360	31,212	1,506
43	107,780	48,553	2,178
44	104,529	45,932	1,999
45	100,218	46,450	1,983
46	106,499	47,528	1,469
47	104,504	44,434	1,598
48	129,167	62,142	2,452
49	112,552	53,478	2,366
50	76,525	34,212	1,676
51	96,642	44,681	1,902
52	99,671	46,495	2,032
53	108,694	54,509	2,432
54	111,655	53,219	2,380
55	238,608	109,220	3,402
56	124,244	62,096	2,851
57	92,165	40,870	1,865
58	60,179	25,959	1,021
59	75,426	33,590	1,591
60	82,896	36,664	1,674
61	93,239	43,767	1,998
62	76,769	34,325	1,696
63	95,732	41,392	1,773
64	88,678	42,180	2,073
65	93,896	41,820	2,146
66	56,291	23,921	1,393
67	79,082	34,907	1,479
68	77,381	34,153	1,616
69	89,529	42,025	1,912
70	80,364	35,978	1,566
71	93,272	39,810	1,867
72	93,438	43,990	2,003
73	107,165	49,911	1,765
74	70,250	30,780	1,267
75	88,112	40,450	1,416
76	96,505	44,266	1,118
77	93,517	45,429	2,136
78	94,898	44,185	1,947
79	87,376	38,790	1,549
80	100,549	49,197	1,272
81	125,794	58,457	2,354
82	70,739	31,434	1,494
83	80,380	36,949	1,486

84	90,699	42,026	1,438
85	91,631	45,880	1,790
86	95,305	45,424	1,662
87	101,510	46,381	1,738
88	105,511	53,067	1,533
89	130,767	57,352	1,989
90	64,604	26,798	1,108
91	90,468	39,459	1,456
92	93,498	40,302	994
93	95,914	44,039	1,985
94	80,197	35,010	1,677
95	94,917	39,586	1,795
96	85,131	39,209	1,168
97	134,836	64,143	2,190
98	59,173	27,667	1,188
99	70,981	33,658	1,511
100	93,629	43,550	1,186
101	113,994	57,193	2,155
102	77,068	36,744	1,012
103	108,445	50,000	1,514
104	179,091	90,003	2,700
105	89,995	41,514	1,690
106	43,927	20,347	585
107	60,855	28,043	1,347
108	521,132	239,578	5,857
109	60,731	29,479	1,470
110	87,386	40,632	1,135
111	109,006	48,443	1,680
112	89,388	43,561	1,776
113	272,483	123,621	3,183
114	39,360	17,635	465
115	54,679	25,024	1,192
116	66,829	30,191	1,579
117	63,934	30,508	1,447
118	77,142	35,165	1,070
119	114,820	49,618	1,555
120	102,484	48,680	1,767
121	65,364	30,095	1,222
122	62,964	28,551	949
123	77,721	36,205	1,495
124	71,768	32,944	1,386
125	52,764	25,824	1,044
126	41,371	19,145	640
127	87,728	39,211	1,177
128	84,770	41,504	1,400
129	87,889	41,376	1,331

130	61,227	28,691	834
131	53,100	24,964	1,104
132	73,875	34,811	1,232
133	62,702	31,281	1,096
134	97,357	46,155	1,138
135	55,352	25,191	1,017
136	63,938	32,121	1,269
137	61,771	28,324	1,161
138	35,531	16,020	504
139	35,744	16,540	818
140	58,102	26,422	1,185
141	60,613	29,293	1,078
142	37,084	17,047	546
143	66,984	29,351	1,173
144	49,522	23,875	944
145	254,139	119,169	3,453
146	147,386	66,747	1,586
147	153,736	72,413	2,466
148	253,387	119,365	3,013
149	214,432	108,001	2,935
150	257,645	123,624	2,566
151	273,107	123,346	2,843
152	225,363	113,672	3,068
153	146,753	64,106	1,989
154	118,651	50,405	1,389
155	156,621	69,086	2,068
156	162,313	71,633	1,492
157	156,850	73,805	2,098
158	194,436	87,610	2,299
159	83,973	35,115	1,161
160	107,123	50,983	1,205
161	134,285	58,684	1,892
162	108,914	46,796	1,564
163	120,976	53,367	1,648
164	125,580	55,169	1,056
165	135,096	63,191	2,928
166	139,161	69,886	2,399
167	119,241	57,142	1,515
168	132,787	59,721	1,896
169	144,389	66,478	3,872
170	117,473	51,825	3,825
171	146,716	59,883	3,705
172	153,274	70,566	4,090
173	115,321	50,436	3,203
174	164,661	72,009	3,849
175	112,970	50,698	3,114

Table S2: Samples and their barcodes used in this study.
Different 8 bp long barcodes were added to each forward and reverse primer to allow the simultaneous tagging of 48 samples (dual indexing). Barcodes added to the forward primers 'Forw. Tag' and to the reverse primers 'Rev. Tag' are shown for each sample, including the treatment of each sample, identity of its mother 'Dam', its father 'Sire' and their interaction 'Family'. Barcodes had no effect on bacterial community distances (results not shown).

SampleID	Treatment	Dam	Sire	Family	Forw. Tag	Rev. Tag
1	Control	A	1	A1	TAGATCGC	TAAGGCGA
2	Control	A	2	A2	CTCTCTAT	TAAGGCGA
3	Control	A	3	A3	TATCCTCT	TAAGGCGA
4	Control	B	1	B1	AGAGTAGA	TAAGGCGA
5	Control	B	2	B2	GTAAGGAG	TAAGGCGA
6	Control	B	3	B3	ACTGCATA	TAAGGCGA
7	Control	C	1	C1	AAGGAGTA	TAAGGCGA
8	Control	C	2	C2	CTAAGCCT	TAAGGCGA
9	Control	C	3	C3	TAGATCGC	CGTACTAG
10	Control	D	1	D1	CTCTCTAT	CGTACTAG
11	Control	D	2	D2	TATCCTCT	CGTACTAG
12	Control	D	3	D3	AGAGTAGA	CGTACTAG
13	1:1000	A	1	A1	GTAAGGAG	CGTACTAG
14	1:1000	A	2	A2	ACTGCATA	CGTACTAG
15	1:1000	A	3	A3	AAGGAGTA	CGTACTAG
16	1:1000	B	1	B1	CTAAGCCT	CGTACTAG
17	1:1000	B	2	B2	TAGATCGC	TCCTGAGC
18	1:1000	B	3	B3	CTCTCTAT	TCCTGAGC
19	1:1000	C	1	C1	TATCCTCT	TCCTGAGC
20	1:1000	C	2	C2	AGAGTAGA	TCCTGAGC
21	1:1000	C	3	C3	GTAAGGAG	TCCTGAGC
22	1:1000	D	1	D1	ACTGCATA	TCCTGAGC
23	1:1000	D	2	D2	AAGGAGTA	TCCTGAGC
24	1:1000	D	3	D3	CTAAGCCT	TCCTGAGC
25	1:500	A	1	A1	TAGATCGC	GGACTCCT
26	1:500	A	2	A2	CTCTCTAT	GGACTCCT
27	1:500	A	3	A3	TATCCTCT	GGACTCCT
28	1:500	B	1	B1	AGAGTAGA	GGACTCCT
29	1:500	B	2	B2	GTAAGGAG	GGACTCCT
30	1:500	B	3	B3	ACTGCATA	GGACTCCT
31	1:500	C	1	C1	AAGGAGTA	GGACTCCT
32	1:500	C	2	C2	CTAAGCCT	GGACTCCT
33	1:500	C	3	C3	TAGATCGC	TAGGCATG
34	1:500	D	1	D1	CTCTCTAT	TAGGCATG
35	1:500	D	2	D2	TATCCTCT	TAGGCATG
36	1:500	D	3	D3	AGAGTAGA	TAGGCATG
37	Control	H	1	H1	GTAAGGAG	TAGGCATG
38	Control	H	2	H2	ACTGCATA	TAGGCATG
39	Control	H	3	H3	AAGGAGTA	TAGGCATG
40	Control	I	1	I1	CTAAGCCT	TAGGCATG
41	Control	I	2	I2	TAGATCGC	CTCTCTAC
42	Control	I	3	I3	CTCTCTAT	CTCTCTAC

43	Control	J	1	J1	TATCCTCT	CTCTCTAC
44	Control	J	2	J2	AGAGTAGA	CTCTCTAC
45	Control	J	3	J3	GTAAGGAG	CTCTCTAC
46	Control	L	1	L1	ACTGCATA	CTCTCTAC
47	Control	L	2	L2	AAGGAGTA	CTCTCTAC
48	Control	L	3	L3	CTAAGCCT	CTCTCTAC
49	1:1000	H	1	H1	TAGATCGC	TAAGGCGA
50	1:1000	H	2	H2	CTCTCTAT	TAAGGCGA
51	1:1000	H	3	H3	TATCCTCT	TAAGGCGA
52	1:1000	I	1	I1	AGAGTAGA	TAAGGCGA
53	1:1000	I	2	I2	GTAAGGAG	TAAGGCGA
54	1:1000	I	3	I3	ACTGCATA	TAAGGCGA
55	1:1000	J	1	J1	AAGGAGTA	TAAGGCGA
56	1:1000	J	2	J2	CTAAGCCT	TAAGGCGA
57	1:1000	J	3	J3	TAGATCGC	CGTACTAG
58	1:1000	L	1	L1	CTCTCTAT	CGTACTAG
59	1:1000	L	2	L2	TATCCTCT	CGTACTAG
60	1:1000	L	3	L3	AGAGTAGA	CGTACTAG
61	1:500	H	1	H1	GTAAGGAG	CGTACTAG
62	1:500	H	2	H2	ACTGCATA	CGTACTAG
63	1:500	H	3	H3	AAGGAGTA	CGTACTAG
64	1:500	I	1	I1	CTAAGCCT	CGTACTAG
65	1:500	I	2	I2	TAGATCGC	TCCTGAGC
66	1:500	I	3	I3	CTCTCTAT	TCCTGAGC
67	1:500	J	1	J1	TATCCTCT	TCCTGAGC
68	1:500	J	2	J2	AGAGTAGA	TCCTGAGC
69	1:500	J	3	J3	GTAAGGAG	TCCTGAGC
70	1:500	L	1	L1	ACTGCATA	TCCTGAGC
71	1:500	L	2	L2	AAGGAGTA	TCCTGAGC
72	1:500	L	3	L3	CTAAGCCT	TCCTGAGC
73	Control	A	4	A5	TAGATCGC	GGACTCCT
74	Control	A	5	A6	CTCTCTAT	GGACTCCT
75	Control	A	6	A7	TATCCTCT	GGACTCCT
76	Control	A	7	A8	AGAGTAGA	GGACTCCT
77	Control	B	4	B5	GTAAGGAG	GGACTCCT
78	Control	B	5	B6	ACTGCATA	GGACTCCT
79	Control	B	6	B7	AAGGAGTA	GGACTCCT
80	Control	B	7	B8	CTAAGCCT	GGACTCCT
81	Control	C	4	C5	TAGATCGC	TAGGCATG
82	Control	C	5	C6	CTCTCTAT	TAGGCATG
83	Control	C	6	C7	TATCCTCT	TAGGCATG
84	Control	C	7	C8	AGAGTAGA	TAGGCATG
85	Control	D	4	D5	GTAAGGAG	TAGGCATG
86	Control	D	5	D6	ACTGCATA	TAGGCATG
87	Control	D	6	D7	AAGGAGTA	TAGGCATG
88	Control	D	7	D8	CTAAGCCT	TAGGCATG
89	1:1000	A	4	A5	TAGATCGC	CTCTCTAC
90	1:1000	A	5	A6	CTCTCTAT	CTCTCTAC
91	1:1000	A	6	A7	TATCCTCT	CTCTCTAC
92	1:1000	A	7	A8	AGAGTAGA	CTCTCTAC

93	1:1000	B	4	B5	GTAAGGAG	CTCTCTAC
94	1:1000	B	5	B6	ACTGCATA	CTCTCTAC
95	1:1000	B	6	B7	AAGGAGTA	CTCTCTAC
96	1:1000	B	7	B8	CTAAGCCT	CTCTCTAC
97	1:1000	C	4	C5	TAGATCGC	TAAGGCGA
98	1:1000	C	5	C6	CTCTCTAT	TAAGGCGA
99	1:1000	C	6	C7	TATCCTCT	TAAGGCGA
100	1:1000	C	7	C8	AGAGTAGA	TAAGGCGA
101	1:1000	D	6	D7	GTAAGGAG	TAAGGCGA
102	1:1000	D	7	D8	ACTGCATA	TAAGGCGA
103	1:500	A	4	A5	AAGGAGTA	TAAGGCGA
104	1:500	A	5	A6	CTAAGCCT	TAAGGCGA
105	1:500	A	6	A7	TAGATCGC	CGTACTAG
106	1:500	A	7	A8	CTCTCTAT	CGTACTAG
107	1:500	B	4	B5	TATCCTCT	CGTACTAG
108	1:500	B	5	B6	AGAGTAGA	CGTACTAG
109	1:500	B	6	B7	GTAAGGAG	CGTACTAG
110	1:500	B	7	B8	ACTGCATA	CGTACTAG
111	1:500	C	4	C5	AAGGAGTA	CGTACTAG
112	1:500	C	5	C6	CTAAGCCT	CGTACTAG
113	1:500	C	6	C7	TAGATCGC	TCCTGAGC
114	1:500	C	7	C8	CTCTCTAT	TCCTGAGC
115	1:500	D	4	D5	TATCCTCT	TCCTGAGC
116	1:500	D	5	D6	AGAGTAGA	TCCTGAGC
117	1:500	D	6	D7	GTAAGGAG	TCCTGAGC
118	1:500	D	7	D8	ACTGCATA	TCCTGAGC
119	Control	H	4	H5	AAGGAGTA	TCCTGAGC
120	Control	H	5	H6	CTAAGCCT	TCCTGAGC
121	Control	H	6	H7	TAGATCGC	GGACTCCT
122	Control	H	7	H8	CTCTCTAT	GGACTCCT
123	Control	I	4	I5	TATCCTCT	GGACTCCT
124	Control	I	5	I6	AGAGTAGA	GGACTCCT
125	Control	I	6	I7	GTAAGGAG	GGACTCCT
126	Control	I	7	I8	ACTGCATA	GGACTCCT
127	Control	J	4	J5	AAGGAGTA	GGACTCCT
128	Control	J	5	J6	CTAAGCCT	GGACTCCT
129	Control	J	6	J7	TAGATCGC	TAGGCATG
130	Control	J	7	J8	CTCTCTAT	TAGGCATG
131	Control	L	4	L5	TATCCTCT	TAGGCATG
132	Control	L	5	L6	AGAGTAGA	TAGGCATG
133	Control	L	6	L7	GTAAGGAG	TAGGCATG
134	Control	L	7	L8	ACTGCATA	TAGGCATG
135	1:1000	H	4	H5	AAGGAGTA	TAGGCATG
136	1:1000	H	5	H6	CTAAGCCT	TAGGCATG
137	1:1000	H	6	H7	TAGATCGC	CTCTCTAC
138	1:1000	H	7	H8	CTCTCTAT	CTCTCTAC
139	1:1000	I	4	I5	TATCCTCT	CTCTCTAC
140	1:1000	I	5	I6	AGAGTAGA	CTCTCTAC
141	1:1000	I	6	I7	GTAAGGAG	CTCTCTAC
142	1:1000	I	7	I8	ACTGCATA	CTCTCTAC

143	1:1000	J	4	J5	AAGGAGTA	CTCTCTAC
144	1:1000	J	5	J6	CTAAGCCT	CTCTCTAC
145	1:1000	J	6	J7	TAGATCGC	TAAGGCGA
146	1:1000	J	7	J8	CTCTCTAT	TAAGGCGA
147	1:1000	L	4	L5	TATCCTCT	TAAGGCGA
148	1:1000	L	5	L6	AGAGTAGA	TAAGGCGA
149	1:1000	L	6	L7	GTAAGGAG	TAAGGCGA
150	1:1000	L	7	L8	ACTGCATA	TAAGGCGA
151	1:500	H	4	H5	AAGGAGTA	TAAGGCGA
152	1:500	H	5	H6	CTAAGCCT	TAAGGCGA
153	1:500	H	6	H7	TAGATCGC	CGTACTAG
154	1:500	H	7	H8	CTCTCTAT	CGTACTAG
155	1:500	I	4	I5	TATCCTCT	CGTACTAG
156	1:500	I	7	I8	AGAGTAGA	CGTACTAG
157	1:500	J	4	J5	GTAAGGAG	CGTACTAG
158	1:500	J	5	J6	ACTGCATA	CGTACTAG
159	1:500	J	6	J7	AAGGAGTA	CGTACTAG
160	1:500	J	7	J8	CTAAGCCT	CGTACTAG
161	1:500	L	4	L5	TAGATCGC	TCCTGAGC
162	1:500	L	5	L6	CTCTCTAT	TCCTGAGC
163	1:500	L	6	L7	TATCCTCT	TCCTGAGC
164	1:500	L	7	L8	AGAGTAGA	TCCTGAGC
165	1:1000	D	4	D5	AGAGTAGA	TAGGCATG
166	1:1000	D	5	D6	GTAAGGAG	TAGGCATG
167	1:500	I	5	I6	ACTGCATA	TAGGCATG
168	1:500	I	6	I7	AAGGAGTA	TAGGCATG
169	unfertilized	-	1	-	GTAAGGAG	GGACTCCT
170	unfertilized	-	2	-	ACTGCATA	GGACTCCT
171	unfertilized	-	3	-	AAGGAGTA	GGACTCCT
172	unfertilized	-	4	-	CTAAGCCT	GGACTCCT
173	unfertilized	-	5	-	TAGATCGC	TAGGCATG
174	unfertilized	-	6	-	CTCTCTAT	TAGGCATG
175	unfertilized	-	7	-	TATCCTCT	TAGGCATG

Table S3: Summary of mean bacterial diversities.
Bacterial diversities (phylogenetic distance) were calculated for each male and female separately. This includes the bacterial diversity on milt before fertilization as well as the mean bacterial diversity on fertilized eggs 14 days after treatment with water (Control), nutrient broth at a dilution of 1:1000 or 1:500. Mean bacterial diversities on fertilized eggs 14 days after treatment are also given overall for each male and female (All), as well as for each treatment (Total).

Male IDs	**Before fertilization**	**All**	**Control**	**NB 1:1000**	**NB 1:500**
1	58.8	60.7	60.6	62.5	59.2
2	52.0	62.1	63.8	63.7	60.1
3	50.5	59.3	59.7	59.9	59.5
4	50.4	54.4	54.9	55.2	53.5
5	52.8	55.8	52.9	52.1	62.8
6	54.8	52.6	49.3	55.3	53.0
7	66.6	51.0	52.0	51.9	48.9
Female IDs					
A	-	55.2	56.2	54.9	54.6
B	-	61.2	60.1	60.9	62.7
C	-	59.6	63.5	57.5	57.6
D	-	56.4	54.9	63.0	51.2
H	-	56.3	58.7	49.9	60.2
I	-	53.2	52.7	50.5	56.4
J	-	56.5	53.0	61.7	54.7
L	-	55.4	50.2	59.5	56.4
Total					
Control	56.0				
NB 1:1000	57.2				
NB 1:500	56.7				

Table S4: Differences in predictive, functional pathways of core bacterial communities on fertilized eggs among the three different treatment groups.
'Observed pathways' were derived from the KEGG database using STAMP v.2.1.3 (Parks *et al*. 2014). Effect sizes of medians among the three different treatment groups as well as corrected *p*-values are shown (a Benjamini – Hochberg multiple testing correction was applied controlling the false discovery rate and the threshold was set as < 0.05). Percentages are given as median relative frequencies within each treatment group. Differences in median percentage between control treatment and nutrient broth at a concentration of 1:500 are presented.

Observed pathway	*p*-values (corrected)	Effect size	Control	NB 1:1000	NB 1:500	Difference
Bacterial toxins	< 0.0001	0.251	0.93	0.99	0.97	0.045
Aminoacyl-tRNA biosynthesis	< 0.0001	0.152	0.95	1.00	0.99	0.034
Lipopolysaccharide biosynthesis proteins	< 0.0001	0.223	0.60	0.61	0.63	0.034
Inositol phosphate metabolism	< 0.0001	0.421	0.25	0.27	0.27	0.026
Transcription machinery	< 0.0001	0.189	0.80	0.82	0.82	0.023
Lysine biosynthesis	< 0.0001	0.161	0.61	0.63	0.63	0.023
Homologous recombination	< 0.0001	0.126	0.68	0.70	0.70	0.021
Amino sugar and nucleotide sugar metabolism	0.002	0.080	0.82	0.84	0.84	0.019
DNA replication proteins	0.001	0.092	0.90	0.93	0.91	0.018
Two-component system	0.001	0.087	3.10	3.05	3.12	0.017
Protein export	0.0001	0.117	0.51	0.54	0.53	0.017
Lipopolysaccharide biosynthesis	< 0.0001	0.165	0.46	0.48	0.48	0.017
Mismatch repair	0.001	0.087	0.61	0.63	0.62	0.013
DNA replication	0.0002	0.114	0.51	0.53	0.52	0.013
Peptidoglycan biosynthesis	0.0002	0.110	0.63	0.65	0.65	0.013
Alanine, aspartate and glutamate metab.	< 0.0001	0.222	0.90	0.92	0.92	0.012
Pertussis	< 0.0001	0.128	0.09	0.09	0.10	0.012
Cell cycle - Caulobacter	0.03	0.045	0.65	0.67	0.66	0.012
Glutamatergic synapse	< 0.0001	0.210	0.08	0.09	0.10	0.011
Riboflavin metabolism	< 0.0001	0.324	0.30	0.31	0.31	0.010
Fatty acid biosynthesis	0.003	0.077	0.60	0.61	0.61	0.009
Pathways in cancer	< 0.0001	0.220	0.08	0.09	0.09	0.009
Translation factors	< 0.0001	0.122	0.40	0.42	0.41	0.009
Lipid biosynthesis proteins	0.0002	0.109	0.96	0.96	0.97	0.008
Glycerophospholipid metabolism	0.0004	0.102	0.55	0.55	0.55	0.008
Glycosyltransferases	0.0008	0.094	0.34	0.34	0.35	0.008
One carbon pool by folate	0.02	0.053	0.49	0.50	0.50	0.007
Nucleotide excision repair	0.002	0.082	0.30	0.30	0.30	0.007
Plant-pathogen interaction	< 0.0001	0.138	0.19	0.20	0.19	0.006
Isoquinoline alkaloid biosynthesis	< 0.0001	0.178	0.07	0.07	0.08	0.006
Amyotrophic lateral sclerosis (ALS)	0.0003	0.108	0.06	0.07	0.07	0.005
Biosynthesis of siderophore group	< 0.0001	0.185	0.04	0.04	0.05	0.005
Cyanoamino acid metabolism	0.02	0.054	0.29	0.29	0.29	0.005
Phenyltransferases	0.002	0.081	0.27	0.28	0.27	0.005
Tropane, piperidine, pyridine biosynthesis	0.008	0.064	0.14	0.14	0.14	0.004
Biotin metabolism	0.0002	0.115	0.14	0.15	0.14	0.003
D-Glutamine and D-glutamate metabolism	0.0002	0.114	0.12	0.13	0.13	0.003
Lipoic acid metabolism	< 0.0001	0.140	0.06	0.07	0.07	0.003
Pentose and glucuronate interconversions	0.02	0.053	0.36	0.37	0.36	0.002
p53 signaling pathway	0.005	0.071	0.03	0.04	0.03	0.002
Viral myocarditis	0.005	0.071	0.03	0.04	0.03	0.002
Influenza A	0.005	0.071	0.03	0.04	0.03	0.002
Toxoplasmosis	0.005	0.071	0.03	0.04	0.03	0.002
Phenylpropanoid biosynthesis	0.03	0.047	0.11	0.11	0.11	0.002
Phosphonate and phosphinate metabolism	0.0005	0.099	0.09	0.08	0.09	0.002
Progesterone-mediated oocyte maturation	0.0001	0.117	0.02	0.02	0.02	0.002
Antigen processing and presentation	0.0001	0.117	0.02	0.02	0.02	0.002

Figure S1: Average oxygen concentrations and pH measurements in a subsample of experimental cells across treatments

Oxygen concentration in the wells, as well as pH were measured at a distance of 4 mm away from the brown trout egg membrane or hatchling in a subsample of 20 wells per treatment (control, nutrient broth at a concentration of 1:1000 and 1:500 in the wells) with a Firesting O2 device (Pyroscience) and a bench pH meter (Life Sciences). Oxygen was measured in mV corresponding to oxygen concentrations of mg/L. Different linings show different treatments (control, nutrient broth at a concentration of 1:1000, and nutrient broth at a concentration of 1:500), while different colours delineate their 95% confidence intervals.

Figure S2: Rarefaction curves of alpha diversity measure

Rarefaction curves of bacterial alpha diversity on fertilized eggs measured as phylogenetic distance are shown by family (different colours).

Figure S3: Relationship of bacterial diversity and hatching time in brown trout embryos

Bacterial alpha diversity was estimated as phylogenetic distance and is based on phylogenetic relationships among bacterial taxa (Gotelli & Chao 2013). Time until hatching was measured as days since *in vitro* fertilizations. Mean values per family are shown and coloured according to sire identity, correspondingly to Fig. 2. a = all trout families and b = only trout families with no mortality were included.

Figure S4: Number of bacterial taxa that could be found at different fractions of samples

To determine a core microbiome of fertilized brown trout eggs 14 days after treatment, OTUs (Operational Taxonomic Units) were ordered according to their presence at different fractions of samples. The QIIME script compute_core_microbiome.py was used for this selection and a closed blast search had to be adopted (Lan *et al*. 2012). For downstream analysis OTUs had to be observed in 90% of all samples.

Figure S5: Heatmap of core bacterial taxa on fertilized brown trout eggs and milt

Heatmaps were plotted at the lowest taxonomic level possible (family, genus, species) using a closed blast search (Lan *et al*. 2012) with the RDP classifier 2.2 (Wang *et al*. 2007) and the Greengenes reference database v.12.13 (McDonald *et al*. 2012). The abundance of OTUs (operational taxonomic units) is given in shades of blue (dark blue = low abundance, light blue = high abundance). Samples are ordered according to multidimensional scaling (Rajaram & Oono 2010). Shown are the most abundant 40 core bacterial taxa on a = fertilized brown trout eggs according to treatment and b = on milt.

Fig. S1

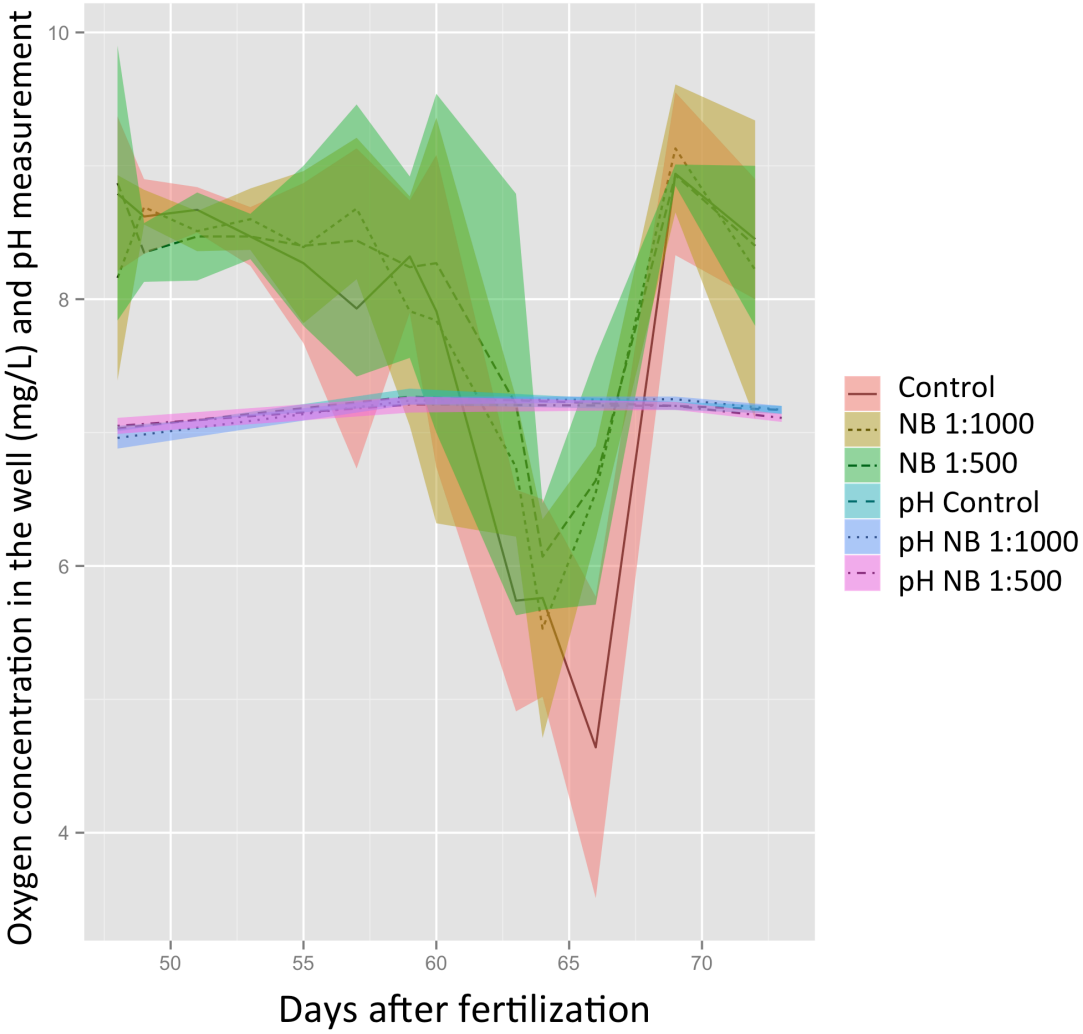

Oxygen concentration in the well (mg/L) and pH measurement
10
8
6
4
50
55
60
65
70
Days after fertilization
Control
NB 1:1000
NB 1:500
pH Control
pH NB 1:1000
pH NB 1:500

Fig. S2

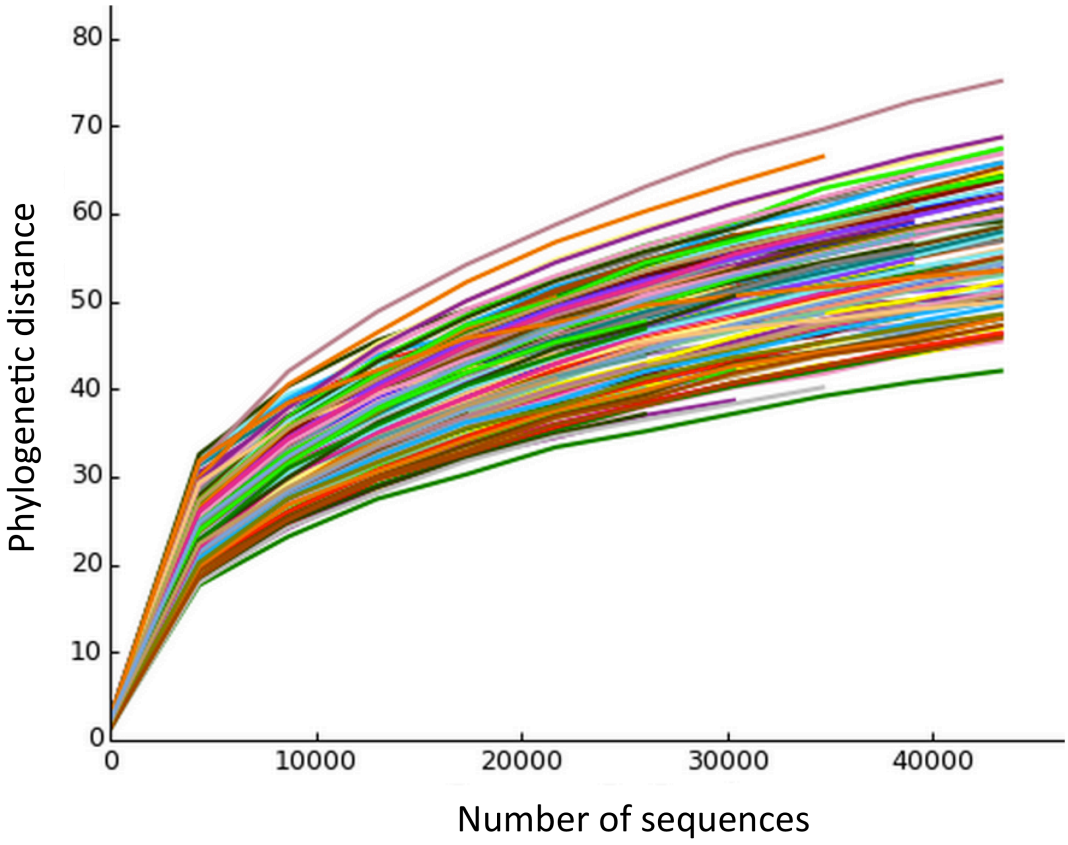
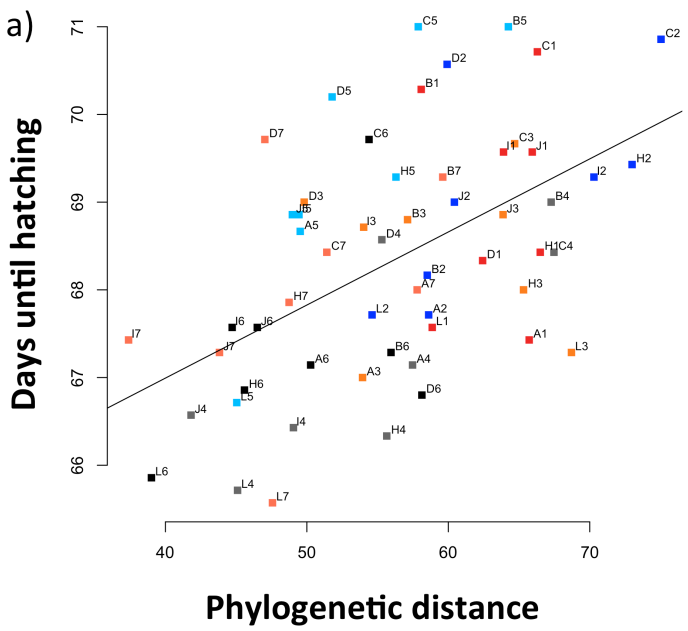

Phylogenetic distance
0
10
20
30
40
50
60
70
80
0
10000
20000
30000
40000
Number of sequences

Fig. S3


a)
Days until hatching
Phylogenetic distance
66
67
68
69
70
71
40
50
60
70

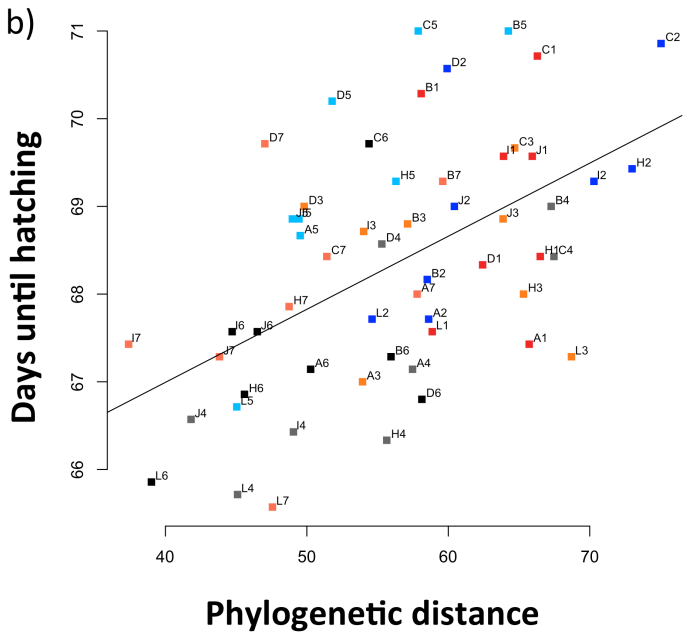

b)
Days until hatching
Phylogenetic distance
66
67
68
69
70
71
40
50
60
70

Fig. S4

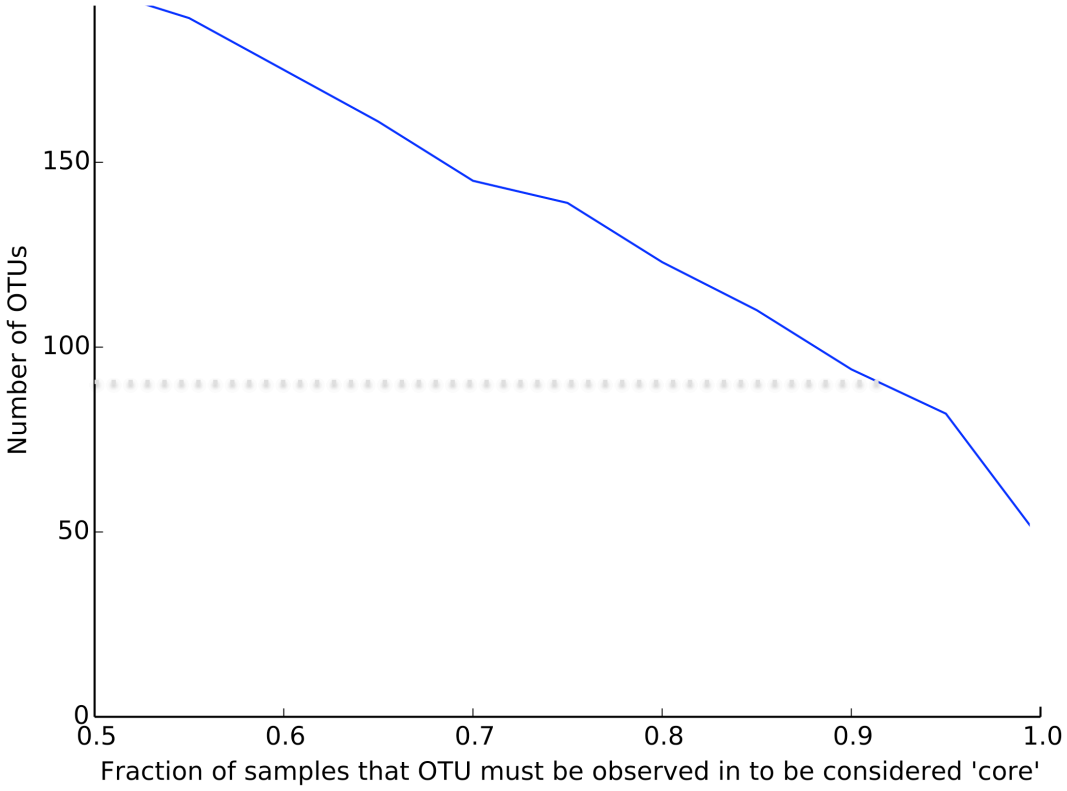
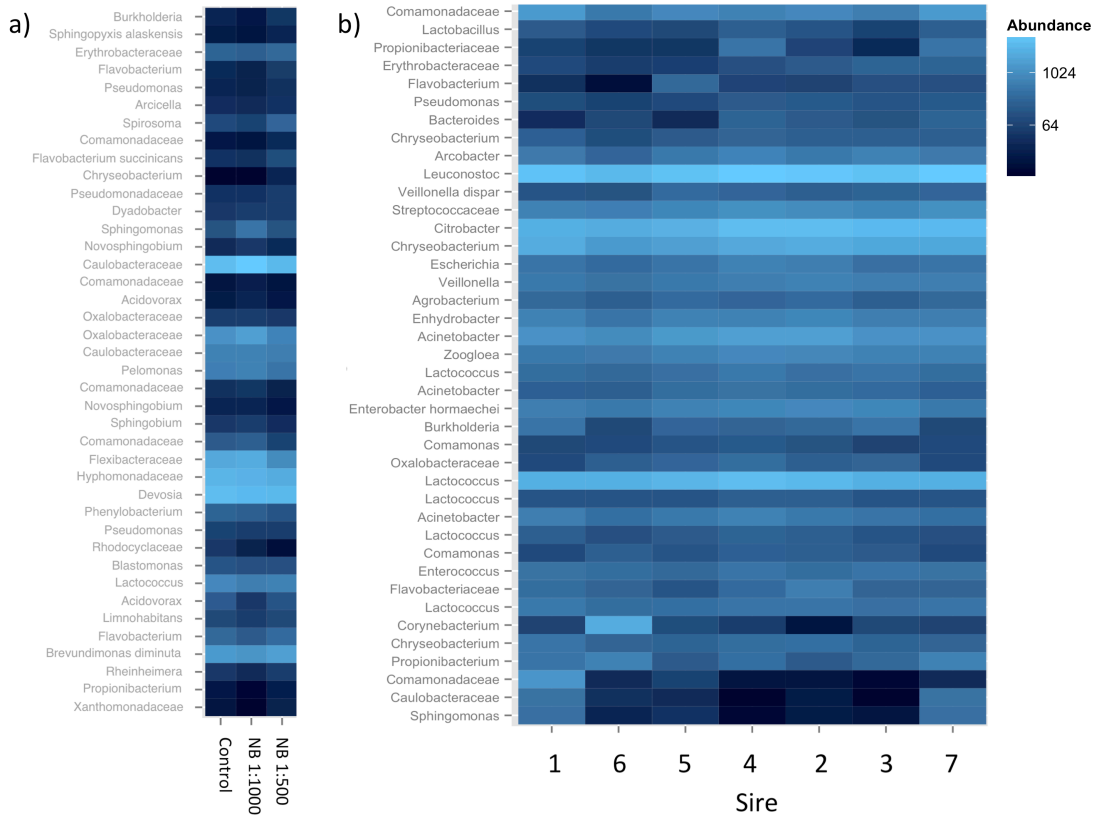

Number of OTUs
150
100
50
0
0.5
0.6
0.7
0.8
0.9
1.0
Fraction of samples that OTU must be observed in to be considered 'core'

Fig. S5


a)
Burkholderia
Sphingopyxis alaskensis
Erythrobacteraceae
Flavobacterium
Pseudomonas
Arcicella
Spirosoma
Comamonadaceae
Flavobacterium succinicans
Chryseobacterium
Pseudomonadaceae
Dyadobacter
Sphingomonas
Novosphingobium
Caulobacteraceae
Comamonadaceae
Acidovorax
Oxalobacteraceae
Oxalobacteraceae
Caulobacteraceae
Pelomonas
Comamonadaceae
Novosphingobium
Sphingobium
Comamonadaceae
Flexibacteraceae
Hyphomonadaceae
Devosia
Phenylobacterium
Pseudomonas
Rhodocyclaceae
Blastomonas
Lactococcus
Acidovorax
Limnohabitans
Flavobacterium
Brevundimonas diminuta
Rheinheimera
Propionibacterium
Xanthomonadaceae
Control
NB 1:1000
NB 1:500
b)
Comamonadaceae
Lactobacillus
Propionibacteriaceae
Erythrobacteraceae
Flavobacterium
Pseudomonas
Bacteroides
Chryseobacterium
Arcobacter
Leuconostoc
Veillonella dispar
Streptococcaceae
Citrobacter
Chryseobacterium
Escherichia
Veillonella
Agrobacterium
Enhydrobacter
Acinetobacter
Zoogloea
Lactococcus
Acinetobacter
Enterobacter hormaechei
Burkholderia
Comamonas
Oxalobacteraceae
Lactococcus
Lactococcus
Acinetobacter
Lactococcus
Comamonas
Enterococcus
Flavobacteriaceae
Lactococcus
Corynebacterium
Chryseobacterium
Propionibacterium
Comamonadaceae
Caulobacteraceae
Sphingomonas
1
6
5
4
2
3
7
Sire
Abundance
1024
64

Supplementary Materials and Methods

Bacterial DNA extraction and PCR protocols

Brown trout eggs and milt were homogenized in a 2 ml Eppendorf tube (safe-lock; Sarstedt, Nümbrecht, Germany) in 1.6 ml of buffer ASL (Qiagen, Hombrechtikon, Switzerland) with a mixer mill (MM300; Retsch, Düsseldorf, Germany) for 2 x 30 sec 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 sec and centrifuged at 13,000 rpm for 1 min. An InhibitEX® tablet (Qiagen) was added to each supernatant (vortexed until the tablet was completely suspended) 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 (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 (Milipore, Zug, Switzerland).

PCR was performed with a bacterial-specific primer pair, 27F and 338R (Hamady *et al*. 2008) that amplifies a 311 bp fragment of the V1-V2 hypervariable region of bacterial 16S rRNA. These primers have widely been used in environmental sequencing studies and were therefore chosen in an attempt to make our study comparable to other microbial characterizations of different habitats. Eight, respectively six, different 8 bp long barcodes were added to each forward and reverse primer to allow the simultaneous tagging of 48 different samples (dual indexing, Table S2). Using four different libraries allows the simultaneous sequencing of 196 samples.

Each PCR was performed in a total volume of 25 μl consisting of 1-50 ng bacterial genomic DNA, 2.5 μl of 10X PCR buffer, 400 μM of each dNTP, 2.5 mM of $MgCl_2$, 0.6 μM of each primer, and 0.625U of Taq polymerase (High Fidelity, Life Technologies, Zug, Switzerland). The thermal profile was modified from 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 sec, 56° C for 30 sec, 72° C for 30 sec, and a final extension at 72° C for 10 min. Step II: 1 μl of the product of step I, 2.5 μl of 10X PCR buffer, 400 μM of each dNTP, 2.5 mM of $MgCl_2$, 0.6 μM individually tagged fusion primers 27F and 338R, and 0.625U of Taq polymerase (High Fidelity, 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 sec, 56° C for 30 sec, 72° C for 30 sec, 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 miSeq Illumina sequencing (Roche, Basel) using four libraries of pooled amplicons (n = 48 samples each) in one full run.

MiSeq data quality control

All steps of the Illumina sequencing raw data quality control were done in the QIIME framework v.1.8.0 (Caporaso *et al*. 2010). Raw reads were split based on their dual barcode combinations (8bp barcode on forward and reverse primer, each) with one mismatch allowed per sequence. Raw reads needed to (i) be at least 300bp long, (ii) have a Phred score of > 25 over a window of 50bp, and (iii) have a minimal abundance of > 20 reads per bacterial reference sequence. Reads were screened for chimeras using the UCHIME algorithm (Edgar 2010). Since paired-end MiSeq Illumina sequencing was applied, a sequence only passed if both forward and reverse reads passed quality control. For statistical analysis of bacterial communities operational taxonomic units (OTUs) were picked using the Usearch algorithm 5.2.236 (Edgar 2010) with default parameters. 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 blast search (Lan *et al*. 2012) with the RDP classifier 2.2 (Wang *et al*. 2007) and the Greengenes reference database version 12.13 (McDonald *et al*. 2012). This approach enables the identification of previously undescribed bacterial sequences in a dataset. All resulting reads (Table S1) were used for further analysis.

Non-parametric analyses of variances in bacterial community composition

Analyses of variance were performed in the R environment using the 'vegan' package (Oksanen *et al*. 2013) with 'Adonis' (Anderson 2001; McArdle & Anderson 2001). These analyses of variance methods are non-parametric, but they assume equal variances among groups of samples. The R package 'Permdisp' (Anderson 2006) was applied to test for equal variances among groups, and the R package 'phyloseq' (McMurdie & Holmes 2013) was used to read the output from QIIME into R. These tests are non-parametric multivariate analyses of variance (MANOVA) using pairwise UniFrac distance matrices as the dependent variable. They partition distance matrices among sources of variation and fit linear models using a permutation test with *pseudo*-F ratios. The observed variance of bacterial community composition is partitioned into components attributable to different sources of variation (here treatment, dam and sire effects). Then the test determines whether the grouping explains a significant part of the variance in bacterial community composition. Because 'Adonis' and other multivariate permutation methods do not allow testing for fixed and random effect models we used all categorical factors as fixed independent variables.

Supplementary References:

Anderson, M.J. (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecology,* **26,** 32-46.

Anderson, M.J. (2006) Distance-based tests for homogeneity of multivariate dispersions. *Biometrics,* **62,** 245-253.

Berry, D., Ben Mahfoudh, K., Wagner, M. & Loy, A. (2011) Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. *Applied and Environmental Microbiology,* **77,** 7846-7849.

Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L. & Knight, R. (2010) PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics,* **26,** 266-267.

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

Gotelli, N.J. & Chao, A. (2013) Measuring and estimating species richness, species diversity, and biotic similarity from sampling data. *Encyclopedia of Biodiversity* (ed. S.A. Levin), pp. 195-211. Academic Press, Waltham, MA.

Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J. & Knight, R. (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nature Methods,* **5,** 235-237.

Hughes, J.B., Hellmann, J.J., Ricketts, T.H. & Bohannan, B.J.M. (2002) Counting the uncountable: Statistical approaches to estimating microbial diversity. *Applied and Environmental Microbiology,* **68,** 448-448.

Lan, Y.M., Wang, Q., Cole, J.R. & Rosen, G.L. (2012) Using the RDP classifier to predict taxonomic novelty and reduce the search space for finding novel organisms. *PLoS One,* **7,** e3249.

McArdle, B.H. & Anderson, M.J. (2001) Fitting multivariate models to community data: A comment on distance-based redundancy analysis. *Ecology,* **82,** 290-297.

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R. & Hugenholtz, P. (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME Journal,* **6,** 610-618.

McMurdie, P.J. & Holmes, S. (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One,* **8,** e61217.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stenvens, M.H.H. & Wagner, H. (2013) vegan: Community Ecology Package. R package version 2.0-8.

Parks, D.H., Tyson, G.W., Hugenholtz, P. & Beiko, R.G. (2014) STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics,* **30,** 3123-3124.

Rajaram, S. & Oono, Y. (2010) NeatMap - non-clustering heat map alternatives in R. *BMC Bioinformatics***,** 11:45.

Wang, Q., Garrity, G.M., Tiedje, J.M. & Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology,* **73,** 5261-5267.