The host phenotype and microbiome varies with infection status, parasite origin and parasite microbiome composition

Megan Hahn
Stony Brook University  https://orcid.org/0000-0001-9266-8232

Agnes Piecyk
Max Planck Institute for Evolutionary Biology: Max-Planck-Institut fur Evolutionsbiologie

Fatima Jorge
University of Otago

Robert Cerrato
Stony Brook University

Martin Kalbe
Max Planck Institute for Evolutionary Biology: Max-Planck-Institut fur Evolutionsbiologie

Nolwenn M. Dheilly (✉ nolwenn.dheilly@stonybrook.edu)
Stony Brook University  https://orcid.org/0000-0002-3675-5013

Research

Keywords: helminths, tripartite interactions, parasite, microbes, hosts

DOI: https://doi.org/10.21203/rs.3.rs-323107/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
The host phenotype and microbiome varies with infection status, parasite origin and parasite microbiome composition

Running title: Host-helminth interaction and the microbiome

Megan A Hahn\textsuperscript{a,1}, Agnes Piecyk\textsuperscript{b,c,2}, Fátima Jorge\textsuperscript{d,3}, Robert Cerrato\textsuperscript{a,4}, Martin Kalbe\textsuperscript{b,c}, Nolwenn M Dheilly\textsuperscript{a,e,5*}

\textsuperscript{a} School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook NY, USA
\textsuperscript{b} Evolutionary Ecology, Max Planck Institute for Evolutionary Biology, Plön, Germany
\textsuperscript{c} Evolutionary Ecology of Marine Fishes, GEOMAR Helmholtz, Centre for Ocean Research Kiel, Germany
\textsuperscript{d} Department of Zoology, University of Otago, P.O. Box 56, Dunedin 9054, New Zealand
\textsuperscript{e} ANSES, Agence Nationale de Sécurité Sanitaire de l’Alimentation, de l’Environnement et du Travail - Laboratoire de Ploufragan-Plouzané, Unité Génétique Virale de Biosécurité, Ploufragan, France

\textsuperscript{1}mhahn@health.nyc.gov, \textsuperscript{2}apiecyk@evolbio.mpg.de, \textsuperscript{3}fatima.esperanca@gmail.com, \textsuperscript{4}rcerrato@stonybrook.edu, \textsuperscript{5}nolwenn.dheilly@anses.fr

Corresponding author*
1 Abstract

Background: A growing literature demonstrates the impact of helminths on their host gut microbiome. However, there is now a need to investigate helminth associated microbes and the complex tripartite interactions between parasite, microbes, and hosts.

Methods: We investigated whether the stickleback host microbiome depends on eco-evolutionary variables by testing the impact of exposure to the parasite *Schistocephalus solidus*, infection success, host genotype, parasite genotype, and parasite microbiome composition.

Results: We observed constitutive differences in the microbiome of stickleback of different origin that increased when sticklebacks exposed to the parasite resisted infection. In contrast, the microbiome of successfully infected sticklebacks varies with parasite genotype. More specifically, we reveal that the association between microbiome and immune gene expression increases in infected individuals, and varies with parasite genotype. In addition, we showed that *S. solidus* hosts a complex endo-microbiome and that the abundance and prevalence of an unknown *Chloroflexi* in the parasite correlate with expression of host immune genes including *foxp3*, *tnfr1*, *cd97*, *stat6* and *marco*.

Conclusions: Within this first comprehensive analysis of a cestode’s interaction with bacteria, we demonstrate that (i) regardless of infection success, parasites contribute to modulating the host microbiome, (ii) when infection is successful, the host microbiome varies with parasite genotype due to genotype-dependent variation in parasite immunomodulation, and (iii) the parasite-associated microbiome is distinct from its host’s and contribute to the host immune response to infection. While a growing number of studies focus on determining the genetic and environmental factors contributing to host microbiome composition, this study reveals that parasites, parasite
genetic factors, and parasite microbiomes can contribute regardless of whether the infection was successful.
2 Declarations

Ethics approval and consent to participate
Work with fish was carried out in full accordance with laws governing use of animals for research in Germany and care and well being of these animals was in compliance with EU legislation and directives.

Consent for publication
Not applicable.

Availability of Data and Materials
All data is being made available through the NCBI database under BioSample accessions SAMN08800217 to SAMN08800225, SAMN08805229 and SAMN08805231 for S. solidus field samples and BioSample accessions SAMN08800337 to SAMN08800360, SAMN08805227, SAMN08805228, SAMN08805230 and SAMN08805232 for G. aculeatus field samples. Cross infection samples are available under the BioSample accessions SAMN16838912 to SAMN16839059 for G. aculeatus samples and SAMN16839060 to SAMN16839129 for S. solidus samples. Kit controls for the cross infection experiment can be found under accession numbers SAMN16839130 to SAMN16839135. Code for all analyses is being made available through Github at https://github.com/megan-hahn/stickleback-ssolidus-microbiome.

Competing Interests
The authors declare that they have no competing interests.

Funding
This project was supported by grants from the Eppley Foundation for Research and the Laurie Landeau Foundation LLC awarded to NMD. This project is part of the Parasite Microbiome Project.
Authors Contributions

NMD designed the study. MH, AP, MK and NMD designed the experimental infection. MH, AP and MK collected fish and parasites and conducted the experimental infection. AP conducted the RT-qPCR. MH executed the microbiome experiments and analyses. MH and NMD analyzed the data with insight from FJ and RC. MH and NMD wrote the paper. All authors reviewed and contributed to the final version of the manuscript.

3 Introduction

The past few years have been marked by the growing recognition that microbes contribute actively to modulating the phenotype of their hosts, with applications in health, agriculture, environmental sciences and evolutionary theory [1-5]. Studies have sought to determine a core microbiome associated with a given host and to find the roles that these microbes play [6-8]. Others investigate the impact of host genotype, diet, the environment, and diseases on the microbiome as an attempt to infer consequences for the host [9-14]. But, how could two interacting organisms, each with their own assemblage of microbes, impact each other’s holobiont [15, 16]? Parasitism is one of the best examples of how species interactions are integrated and are ideal to start answering this question.

The association between helminth infections and changes in the composition and diversity of the host gut microbiota has been investigated in many systems [14, 17-20], in part as a mean to better understand the link between parasite load, the microbiome, and inflammatory bowel diseases [20-23]. Results differ greatly, with some studies showing an increase in microbial diversity associated with infection [24-26], others showing a decrease in microbial diversity [13, 17, 27, 28], and some showing no changes [18, 29, 30]. Specific bacterial taxa were at times found to be impacted such as *Hymenolepis diminuta* infection which is associated with a reduced
abundance in *Bacilli* [17], *Trichuris suis* infection that is associated with an increased abundance in *Lactobacilli* [13], and the liver fluke *Opisthorchis viverrini* that promotes *Helicobacter* growth [14].

The differences that have been reported likely result from profound differences in host and parasite eco-evolutionary history. Indeed, both host and parasite contribute to determining the impact of infection on the gut epithelium, host immune response, or mucus production, all of which could result in variations in host microbiota composition [31]. Profound differences between parasite species exist depending on whether or not they reside in the gut and depending on their impact on different host immune pathways [31]. Within a given host species, geographic variation in parasite-mediated natural selection can result in geographic variation in immune response mechanisms [32, 33]. Conversely, geographic variation in host immune response imposes selection pressures on parasites that locally adapt to their hosts and display different virulence mechanisms [34, 35]. However, it remains to be determined whether this fine-tuned crosstalk between host immune response and parasite virulence impacts host microbiota [36]. The finding that a single parasite species can differentially impact the immune system and microbiota of different host species [37], and conversely, that a single host species host different bacterial communities depending on the parasite species it is infected with [38], support this hypothesis.

Disentangling whether the parasite itself impacts the host microbiome, or whether it is the parasite interaction with the host immune system that leads in changes in microbiome composition presents another challenge [39]. Parasites have been associated with direct modulation of their host immune system via the production of virulence factors, breaching of host barrier surfaces such as skin or intestinal epithelium, and production of excrements that interact with host immune cells [40-45]. However, recent evidence has also shown that the host microbiome plays key roles in
determining the host immune response [46, 47]. In particular, anti-inflammatory interleukins, foxp3 and tgf-β appear to be at the heart of the interplay between bacteria and the immune system [48, 49]. For example, the microbiome modulates the innate immune response of mice exposed to influenza [50]. In this study, a microbiota-induced expression of IL-1β and IL-18 was found to be associated with better outcome, and a distal inoculation of LPS to the colon was sufficient to restore the immune response to influenza virus in the lung. Thus, to understand the impact a parasite has on its host holobiont, it is essential to consider both immune function and microbiome composition, and to investigate how they interact with each other.

Parasites can also host microbes and for some parasites, the role of bacterial symbionts in virulence and pathogenesis has been thoroughly investigated. Many nematodes depend on Wolbachia for normal development and fertility, and the bacteria also contributes to inflammation and adverse reaction to anti-filarial drugs [51]. Similarly, the bacteria Neorickettsia has high prevalence among digenean trematodes and is often transmitted to parasitized hosts causing associated diseases [52-55]. But, for most parasite taxa, their associated microbiota remains unknown, and in fact, the presence of other bacteria in trematodes and nematodes has not yet been thoroughly investigated [56]. This might be because most helminths reside in the microbe-rich gut of their hosts, and studies would have to disentangle the parasite microbiota from contaminants in the sampled material [57]. Recent studies have shown that both nematodes and trematodes can host a complex microbial community distinct from that of parasitized hosts and from the environment, and evidence suggests that some of these bacteria might be vertically transmitted [58-61]. However, the presence of bacteria in cestodes, the only taxa of helminths that do not have a gut and absorb nutrients through the skin, has not yet been investigated.
Herein, we characterized the microbiomes of the cestode *Schistocephalus solidus* and of its threespine stickleback (*Gasterosteus aculeatus*) fish host using 16S amplicon rDNA sequencing. This parasite system is particularly valuable because upon ingestion, *S. solidus* egresses through the intestinal wall within 24 hours, which means that it is in direct contact with the host gut microbiome for a very limited time [62]. The parasite then grows to sexual maturity in the sterile peritoneal cavity of the fish which limits chances of contamination from the host microbiota and allows a reliable assessment of the microbial communities associated with the cestode. Thus, after an initial investigation of the microbiome of wild mature *S. solidus* and of differences in the microbiome of infected and non-infected wild-caught sticklebacks, we conducted experimental infections in control conditions. Threespine sticklebacks from two different lakes in Alaska (Wolf and Walby lake), and from one lake in Germany (Lake Grosser Ploner, GSP), were experimentally exposed, and infected with *S. solidus* from Wolf and Walby lake in Alaska, and from one lake in Norway (Skogseidvatnet, SKO) (Figure 1). We tested the effect of exposure to the parasite and infection on stickleback microbiome composition. Then, we determined the role of host and parasite genetic factors in the effect of parasitism on the stickleback microbiome. We also characterized *S. solidus* microbiome and tested whether the parasite microbiome is shaped by environmental (the host microbiome) or genetic factors. Finally, using stickleback genes involved in innate and adaptive immune responses, we identified genes whose expression correlated with bacterial abundance in either the host or the parasite, and tested the impact of infection, host genotype and parasite genotype, thus revealing immune pathways at the crosstalk of *S. solidus*-bacteria interaction.
Figure 1: Summary of experimental design. (A) The map displays all sampling locations. The field study was conducted on sticklebacks and *S. solidus* from one location in Alaska (Cheney). The three genotypes of stickleback hosts (Walby (W), Wolf (w) and GPS(G)) and three genotypes of *S. solidus* (Walby (W), Wolf (w) and SKO (S)) used for experimental infection were collected from lakes in Alaska and Europe. (B) Design matrix showing the organization of rounds and tanks based on fish genotype (denoted by letters) and parasite genotype (denoted by numbers). A total of 36 tanks were used. One example is squared in red. (C) Each of the tanks contained three sham control fish (left) from Wolf (Red Tail), Walby (Green Tail), and GPS (Blue Tail) separated by a mesh divider from 5 Wolf fish, 5 GPS fish, and 4 Walby fish that had been exposed to parasites (right). (D) Upon dissection, the success of infection was assessed, and exposed individuals were classified as exposed non-infected (ENI) or exposed successfully infected (ESI). All successfully infected fish were processed, and corresponding ENI and control non-infected fish from the same tanks and fish origin were processed as controls.
4 Results

The microbiome of wild-caught stickleback hosts and S. solidus parasites

As an initial approach to investigate S. solidus interaction with microbes, we field caught threespine sticklebacks from Cheney lake, in Alaska to characterize and compare the microbiome of the gut of six non-infected fish, six infected fish, and corresponding nine S. solidus plerocercoids extracted from their body cavity. We used a quantitative PCR (qPCR) approach with two sets of universal primers targeting the bacteria 16S rRNA gene as a proxy for bacterial load and 16S amplicon-based sequencing to characterize the microbiome composition. Comparisons between non-infected and infected sticklebacks revealed profound differences in microbial abundance and diversity (Log Cell p<0.01; PCoA axes 1 and 2 unpaired T-test p<0.05; PERMANOVA = p<0.05; PERMDISP = p<0.05; Figure S1 and S2). Sequencing of the surface microbiome of S. solidus, and of the body cavity of sticklebacks yielded results similar to negative controls (Figure S3). In contrast, S. solidus endomicrobiome was as diverse as the stickleback microbiome, but with distinct compositional differences (PCoA axis 1 paired T-test p<0.05; PERMANOVA = p<0.05, Figures S3-5). Actually, bacterial load and alpha diversity in S. solidus was not significantly different from bacterial load in non-infected fish (p<0.05, Figure S1-2 and S4). Confirming this result, the presence of cultivable bacteria in homogenates of plerocercoids plated on agar, but not on plerocercoid surface confirmed that the parasite harbors an endomicrobiome (Figure S6).

Cross infection experiment

Following experimental infections with hosts and parasites of different origin (Figure 1), we quantified and sequenced the 16S genes of a total of 42 control non-infected sticklebacks (CNI), 35 exposed but non-infected sticklebacks (ENI), and 71 exposed and successfully infected
sticklebacks (ESI), and corresponding *S. solidus* (Ss). Our results confirmed that *S. solidus* and *G. aculeatus* harbor a distinct microbiome, and that exposure and infection alters the host microbiome (Figure S7-8).

The microbiome of threespine sticklebacks varied with exposure, infection, host origin, and parasite origin (Figure S9, Figure 2). Comparisons of the microbiome composition of CNI fish revealed constitutive differences between Alaskan and European sticklebacks (Figure 2A). Exposure to *S. solidus* was associated with small changes in beta diversity (Figure 2B), but resulted in an increase in differences in microbiome diversity metrics among fish of all three origins (Figure 2C). Parasite origin played a less profound role and limited differences in diversity were found between fish exposed to Walby and SKO parasites (Figure 2C). Successful infection with *S. solidus* was associated with an increase in bacterial load that varied with parasite origin (Figure 2B and 2D). The microbiome of infected fish was dominated by more taxa than non-infected fish (Figure 2B). Finally, parasite origin, but not host origin, was associated with differences in microbiome composition among infected sticklebacks (Figure 2D), characterized by a distinct clustering of fish infected by European and Alaskan parasites. Finally, regarding the microbiome of *S. solidus* plerocercoids, parasite origin, but not fish origin was associated with subtle differences in microbiome beta diversity (Figures 2E and S9).
Figure 2: Stickleback bacterial community composition vary with exposure, infection status, host genotype and parasite genotype as revealed by LMMs. (A) Among the control non-infected fish, the microbiome composition varies with fish genotype. Differences were found between Walby (n=17) and GPS fish (n=15) (Weighted Unifrac, Axes 1 LMM p= 0.038; and Axis 3, Unweighted Unifrac, Axis 3, p=0.028) and between Wolf (n=10) and GPS fish (Unweighted Unifrac Axis 1, p=0.024). (B) Exposure and infection impact the microbial diversity and
composition. Exposure impacts beta diversity (Weighted Unifrac, Axis 3, p=0.031), whereas successful infection is associated with changes in bacterial load (p=0.01) and alpha diversity (Shannon p=0.039, Simpson p=0.013). (C) Among exposed but non-infected fish, the microbiome diversity and composition vary with host genotype and parasite genotype. Differences were found between Wolf (n=9) and GPS (n=14) fish (Fisher p=0.008, Unweighted Unifrac Axis 1 p=0.005, and 3 p=0.013), between Walby (n=10) and Wolf fish (Shannon p=0.034, Simpson p=0.020, Fisher p=0.004, Unweighted Unifrac Axis 1 p=0.006) and between Walby and GPS fish (Unweighted Unifrac Axis 3 p=0.041). Differences were found between alpha diversity of fish exposed to Walby and SKO parasites (Fisher p=0.018). (D) Among infected sticklebacks, microbiome composition varies with parasite origin and host-parasite co-evolutionary history. Fish infected with SKO (n=18) parasites had higher bacterial load than fish infected with Walby (n=27, p=0.010) and Wolf parasites (n=25, p=0.001). Unweighted Unifrac distance revealed distinct clustering of fish infected with SKO parasites and fish infected with Walby or Wolf parasites (Axis 2, p=0.046 and p=0.047 respectively). (E) Within parasites, microbiome composition varies with parasite origin. Weighted Unifrac distance revealed distinct clustering of the microbiome of SKO (n=18) and Walby (n=27) parasites (Axis 3, p=0.035). The bottom and top edges of the boxes mark the 25th and 75th percentiles. The horizontal line denotes the median. Whiskers mark the range of the data excluding outliers (dots).

**Taxa distribution**

We observed 458 bacterial families associated with *G. aculeatus* samples and 216 families associated with *S. solidus*. Of these, 202 families were found in both fish and parasites. The 253 families found in fish only, and the 14 families found in parasites only were of low abundance and
prevalence, contributing a cumulative relative abundance of 6.9% (±SE 0.02) and 0.1% (±SE 0.001), respectively. Thus, all abundant bacterial families in S. solidus were found in sticklebacks (Figure 3A). But, the most abundant and most prevalent bacteria in parasites and fish are not the same and comparison of the relative abundance of bacterial taxa between individual S. solidus and their corresponding fish host revealed an absence of relationship (Figure 3A, Figure S8). In total, 93% of the ASVs, and 9.4% of the families present in ESI sticklebacks were never found in S. solidus, including the Vibrionaceae that represented an average relative abundance of 5.5% in ESI fish. Most of the abundant ASVs in S. solidus had low abundance and prevalence in sticklebacks, including ASVs of Bacillaceae, Lactobacillus, Chloroflexi, Rhodospirillales, Actinomycetales, and Chitinophagaceae (Figure 3B). Actually, the family Bacillaceae was the most abundant in both fish and parasites (with a relative abundance of 30.6% and 7.5%, respectively, Figure 3A), but comparisons at the ASV level revealed major compositional differences. Of the 144 ASVs that compose this family, only 16 were found at least once in both S. solidus and G. aculeatus, and all 12 most abundant and prevalent ASVs of Bacillaceae were significantly differentially abundant (DESeq Adj p <0.05) in fish and parasite: four are more abundant in parasites and eight are more abundant in fish. A neighbor joining analysis of all ASVs of Bacillaceae revealed that those most prevalent and abundant in S. solidus cluster separately from those most prevalent and most abundant in G. aculeatus (Figure S10).
Figure 3: Bacterial taxa at multiple taxonomic levels vary in relative abundance and prevalence between threespine stickleback gut and *S. solidus* plerocercoids. The heatmaps depict the prevalence of the bacterial families (A) or ASVs (B) at different detection thresholds (defined in terms of minimum relative abundance) for successfully infected (ESI) sticklebacks, and corresponding *S. solidus*. A/ 30 most abundant bacterial families across all samples. B/ 30 most abundant ASVs in *S. solidus*. The phylum is provided as a color code next to the taxa name.

We used DESeq2 to identify differentially abundant bacteria phylotypes (Figure S11). Exposure and infection had a limited effect on individual bacteria. Exposure to *S. solidus* was associated with higher relative abundance of *Rhodobacteraceae*, and with a lower relative abundance of *Betaproteobacteria*, whereas successful infection resulted in a higher relative abundance in *Vibrionaceae*, an unknown *Synechococcophycideae* and another unknown Cyanobacteria.
Host genotype and parasite genotype both contributed to differences in relative abundance of bacterial families (Figure S11). In CNI, host genotype was associated with variation in *Streptococcaceae, Phormidiaceae* and *Ruminococcaceae*. Focusing on ENI, the relative abundance of an unknown Chloroflexi varied with both parasite origin and fish origin, and the relative abundance in the aforementioned *Rhodobacteraceae* (impacted by exposure to *S. solidus*) varied with parasite origin. In ESI, two *Synechococcophycideae* (family *Synechococcaceae* and an unknown family) varied in relative abundance depending on fish origin. Similarly, two families of Bacillales (*Bacillaceae* and an unknown family) and two families of Betaproteobacteria (an unknown *Burkholderiales* and another unknown family) in addition to the family *Pseudanabaenaceae* (*Synechococcophycideae*) varied in relative abundance depending on parasite origin.

In *S. solidus*, the relative abundance in bacterial families varied with the origin of the infected host and with parasite origin. Of interest, 11 families were found to vary in abundance with both host and parasite origin, including the aforementioned unknown family of Bacillales. In addition, the relative abundance in *Lactobacillaceae* varied with parasite origin, whereas the relative abundance in *Phormidiaceae* and *Rhodospirillaceae* varied with fish origin.

**Impact of *S. solidus* on the interaction between immune gene expression and the microbiota**

Finally, we investigated the impact of *S. solidus* infection on the relationship between host gene expression and the microbiome (Figure S12, Figure 4). Spearman’s rank correlation revealed an interaction between gene expression profiles and beta diversity metrics (Unweighted unifrac distance PCoA axis 2 p=0.0288 and PCoA axis 3 p=0.0017; Figure 4A). The relationship between microbiome composition and gene expression profile was more pronounced in infected
individuals, the greater slope indicating that the parasite increases the dependency between immune response and microbiome composition (Figure 4B). Among infected individuals, parasite origin, but not host origin was associated with significant differences in the strength of the correlation between immune gene expression and microbiome composition (Figure 4B). These correlations appear to be driven by a subset of bacterial families, among which some were positively correlated with gene expression whereas others were negatively correlated (Figure 4C). The most significant correlations involved Treg-inducing genes stat4, stat6 and il16, Treg associated gene foxp3, complement factor cfb, anti-microbial innate regulatory genes cd97 and marco, and the regulator of inflammation tnfr1 (Figure 4C, Figure S12). More specifically, the relative abundance of Synechococcophycideae, Lactobacillaceae and Pseudanabaenaceae were negatively correlated with immune gene expression (Figure 4E-G) whereas the relative abundance in Vibrionaceae was negatively correlated with the expression of the immune response genes (Figure 4H). Relative abundance in a different subset of bacteria, namely Aerococcaceae, Streptococcaceae and an unknown Actinomycetales were positively correlated with a higher hepatosomatic index (Figure 4D).
Figure 4: The host phenotype depends on the composition of the microbiome of the host and of its parasite. A. Correlation heatmap depicting gene expression – microbial profile correlations. Color of the squares indicate the strength of Spearman’s rank correlation. Asterisks indicate significance of the correlation (* $q < 0.05$; ** $q < 0.05$; *** $q < 0.001$). B. Comparison of correlations between gene PC1 and uwuPC3 depending on infection status, host genotype and parasite genotype. Significant differences are observed when the confidence interval at the lower and/or higher values do not overlap with the mean of the other one. C. A. Correlation heatmap between the expression of genes, and phenotypic indices (HI, Hepatosomatic index; SI, Splenosomatic index, HKI Head-kidney index; BC, Body condition) and the relative abundance in the most abundant acterial families in the host *G. aculeatus*, and its parasite *S. solidus*. Color indicate the strength of Spearman’s rank correlation. The complete figure with indication of significance is available in Figure S12. Significant Spearman’s rank correlation (with FDR <0.05) are plotted in (D) between the relative abundance in *Streptococcaceae*, an unknown Actynomicetales and *Aerococcaceae* in the host with Heptaosomatic index (N=145), in (E) between the relative abundance of *Synechoccophycidae* in fish and the relative expression of *foxp3, il16, cd97 and stat6* (N=145), in (F) between the relative abundance in *Lactobacillaceae* in fish and the relative expression of *foxp3, cd97 and stat6* (N=145), in (G) between the relative abundance in *Pseudanabaenaceae* in fish and the relative expression of *foxp3, cd97 and stat4* (N=145), in (H) between the relative abundance of *Vibrionaceae* in fish and relative expression of *foxp3, il16, cd97 and stat6* (N=145) and in (I) between the relative abundance of *Chloroflexi* in the parasite and relative abundance in *foxp3, tnfr1, cd97, stat6, and marco* (N=69). The significant Spearman correlation, after Bonferroni correction, are provided on each plot.
The putative role of *S. solidus*-associated bacteria on immune gene expression

Finally, we investigated whether bacterial families associated with *S. solidus* correlated with immune gene expression in infected individuals. Our analysis revealed a positive correlation between some of the highly prevalent bacterial families and immune gene expression in infected fish (Figure 4 C). Most specifically, the relative abundance of the family of *Chloroflexi* in *S. solidus* was correlated with a significantly lower PC1 values, and higher expression of *stat6, foxp3, tnfr1, cd97*, and *marco* demonstrating the impact of parasite-associated bacteria on host immune response to infection (Figure 4I).

5 Discussion

*S. solidus* hosts a diverse microbiome

Our results provide the first set of evidence of an endomicrobiome in the cestode *S. solidus*. We collected *S. solidus* plerocercoids from the body cavity of *G. aculeatus*, so that the parasite was no longer in contact with the host gut microbiota, limiting the potential for contamination [76]. We did not culture any bacteria after spreading freshly sampled plerocercoids on agar suggesting the absence of a surface microbiome. Using 16 sequencing, we further tested the presence of bacteria in the body cavity of parasitized fish and on the surface of *S. solidus*, and obtained results similar to blank controls, confirming that the fish body cavity is sterile, and that no bacteria are present on the cestode tegument. In contrast, we could grow colonies on agar and characterized a complex microbiome from plerocercoid homogenates. In both field and laboratory experiments, the microbiome of *S. solidus* was dominated by Proteobacteria and Firmicutes, and to a lesser extent with Bacteroidetes, Actinobacteria and Chloroflexi. The microbiome of the trematodes *Coitocaecum parvum* and *Opisthorchis viverrini* are also dominated by Firmicutes, Proteobacteria
and Actinobacteria [14, 60, 63, 64]. Together, these results strongly support the hypothesis that any parasitic flatworm may host a microbiome and encourage studies to investigate the structure and function of helminth microbiome.

Our results show that *S. solidus* microbiome is not a random assemblage of microbes present in the parasite environment and suggest that host-based selective forces play a role in shaping the parasite microbiome composition and diversity. Microbial community assembly can consist of vertical transmission of symbiotic members, and horizontal acquisition of microbes from the organism’s environment (through random process or host-based selection). Vertical transmission and host-based selection of environmental microbes can lead to diversification of microbial assemblages, and even phylosymbiosis, where the clustering of microbiome composition data concur with the phylogeny of the hosts [65]. In accordance, we found significant differences in our two Alaskan *S. solidus* populations, and the European population, that are all recognized as genetically distinct lineages [66, 67]. But, our results indicate that the *S. solidus* microbiome is markedly different from its host microbiome, even at the ASV level: different ASVs had high prevalence and relative abundance in parasites and fish. The presence in *S. solidus* of ASVs that were absent in threespine sticklebacks suggest that some of these bacteria could be vertically transmitted. In disagreement with this hypothesis, none of the ASVs found in *S. solidus* had high prevalence in parasites of only one origin. Yet, ASVs of Bacillaceae of *S. solidus* tend to cluster together and separately from ASVs of *Bacillaceae* associated with *G. aculeatus*. We suggest that *S. solidus* might select bacteria from its environment at previous stages of its life cycle. Given the diversity of microbial reservoirs from which this parasite may acquire microbiota throughout its complex lifecycle (definitive bird host, environmental water and soil, intermediate copepod host and intermediate stickleback host), more studies need to be completed to determine
the sources for the parasite microbiome, the mechanisms of acquisition, and the conditions that impact the maintenance of these microbial communities.

Both exposure and infection influence the host microbiome composition

We observed an impact of both exposure and infection by *S. solidus* on the fish gut microbiome even though the parasite had not been in contact with host-associated microbes for weeks at the time of sampling. Regardless of host or parasite genotype, and regardless of the lab environment, exposure and infection by *S. solidus* were associated with consistent changes in the microbiome. Indeed, herein, and within Ling et al. [36], a significant increase in Rhodobacteriales was found associated with exposure and infection by *S. solidus*. In addition, both our field study and experimental infections revealed a previously unreported increase in alpha diversity metrics and bacterial load in infected individuals. We also observed variations among studies that may be due to differences in environmental conditions, of using sticklebacks with different genetic background, and from using different strains of parasites. Indeed, while Ling et al [36] showed an increase in Lactobacilliales in infected individuals, we observed an increase in Vibrionales and Desulfovibrionales following experimental infection and field sampling, respectively. The colonization of the gut of infected fish by these bacteria may be dependent on their initial presence in the environment of the fish. Unfortunately, this is a limitation of our design, as we did not characterize the environmental water microbiome, which prevented us from testing the hypothesis that these bacteria had variable presence in these different environmental and field settings. All these bacterial families are of interest because of their pathogenic potential [68-70], and might indicate that infection with *S. solidus* increases *G. aculeatus* susceptibility to pathobionts as previously reported for other helminths [71].
The impact of exposure to *S. solidus* on the microbiome is intriguing because there is currently no evidence that exposure alone induces a sustained effect on *G. aculeatus*. In fact, until now, studies comparing threespine stickleback immune gene expression and other immunological parameters could not reveal differences between control and exposed but non-infected sticklebacks [33, 72]. However, we recently observed an impact of exposure on threespine stickleback hepatosomatic index and body condition, traditionally used as a measure of host condition [33, 73]. This reduced host condition upon exposure suggests that the cost of resistance to infection, and elimination of the parasite through cell-mediated immunity within the first two weeks of infection [74] could indirectly impact the microbiome of exposed sticklebacks.

*The interaction of host- and parasite-associated bacteria with the host immune system*

As an initial approach to characterize the underlying mechanisms, for the indirect effect of *S. solidus* presence in the body cavity on the host microbiome, we looked for the impact of exposure and infection on the correlation between immune gene expression and microbiome composition. The most significant correlations involved *foxp3*, a specific markers for Treg and Treg-inducing genes *stat4*, *stat6* and *il16*, the *cfb* central to the complement system, antigen *cd97* that promote cell migration to sites of inflammation, the scavenger receptor cysteine-rich protein *marco* that binds bacteria, and the regulator of inflammation *tnfr1* that activate the transcription factor NF-kappaB. Other genes involved in innate immunity, and the Th1 or Th2 adaptive immune response did not correlate with bacterial abundance, despite being strongly modulated following *S. solidus* infection [56]. Correlation between immune parameters and gut microbes are well documented in other species including horses [75], mice [76-78], rats [79] and humans [80, 81] and have been used to infer pathways involved in the interaction. In particular, commensal bacteria are associated with an anti-inflammatory regulatory T cells, whereas pathogens induce a pro-
inflammatory response. We found a negative correlation between pro-inflammatory immune gene Foxp3, Il16, Stat4, Stat6 and cd97 and the family Lactobacillaceae, and the Cyanobacteria Synechococcophycidae and Pseudoanabaenaceae but a positive correlation of the same genes with Vibrionaceae, indicating that these bacterial families associated with G. aculeatus are likely commensals and pathogens respectively. The strength of the interaction between microbiome composition and gene expression profile was higher in infected individuals, which is explained by the higher relative abundance in Vibrionaceae and Synechococcophycideae in infected individuals. This results indicates that S. solidus reinforce the immune communication between the host and its microbiome.

Interestingly, a very strong association was found between the family Chloroflexi, highly prevalent in S. solidus and expression of pro-inflammatory genes Foxp3, Tnfr1, Cd97, Stat6 and Marco. Although not significant with our sample size, other bacteria associated with S. solidus, such as two families of Rhodospirillales and a Chitinophagaceae, also correlate with a pathogen-like immune gene expression pattern. Even though little is known about these bacteria, Chloroflexi and Rhodospirillaceae include opportunistic pathogens [82, 83]. Helminths, including S. solidus induce a Treg response, ensuring their persistence during chronic infection, and preventing a pathological inflammatory response that would be detrimental to the host [72, 75, 84]. Thus, the observed parasite-associated bacteria could contribute, or hinder the mechanisms of parasite modulation of host immune response. While it has previously been shown that host associated microbiota can contribute to host-parasite coevolution [85], this study demonstrates the necessity of investigating the role of parasite-associated microbes.

A host-genotype dependent constitutive response to exposure impacts the microbiome
Our results indicate that the microbiome composition of fish exposed to *S. solidus* strongly depends on host genotype. In control individuals, we observed host genotype-specific constitutive differences in microbiome composition that confirm that host-based selective processes shape the organism microbiota [86-88], probably due to constitutive differences in immune gene expression pathways between populations [33, 73, 89, 90]. In agreement, we found host-genotype dependent differences in the relationship between key immune genes and abundance of bacterial families, indicating that fish of different origins have developed different sensitivity to different bacterial families. Interestingly, we found that exposure to *S. solidus* increased differences based on host genotypes. In contrast, limited differences between fish populations were found after infection by *S. solidus*. The same observation was recently made concerning the immune system of threespine sticklebacks exposed to *S. solidus*: constitutive differences in immune gene expression between populations increased upon exposure to *S. solidus*, whereas gene expression profiles converged upon infection [33]. These results indicate that a non-specific, but a host genotype dependent response to *S. solidus* modulates the Threespine stickleback microbiome composition in exposed individuals. In infected individuals, the decrease variance is consistent with parasite manipulation of host phenotype, homogenizing responses towards an optimum that favors parasite fitness [91].

**The impact of infection on the host microbiome is parasite genotype dependent**

In addition to demonstrating that *S. solidus* - that resides in the stickleback body cavity and is physically separated from the gut microbiome - can nevertheless impact its host gut microbiome composition, we showed that genetic variations in parasitic worms influences the ways the microbiota changes. This parasite-genotype effect was observed in all exposed threespine sticklebacks, but it was stronger in successfully infected fish. Related species of helminth parasites have been shown to modulate their host microbiomes differently as is the case with the cestodes...
Hymenolepis dimunita and Hymenolepis nana [38], and with the nematodes Trichuris suis and
Trichuris muris [12, 13]. However, our study is the first to demonstrate that different genotypes of
the same parasite species have different impacts on host microbiome composition and diversity.
The observed limited parasite genotype effect on the microbiome of fish that resisted the parasite,
together with the fact that the parasite has a very limited time to directly interact with gut microbes
supports the hypothesis of a differential modulation of the immune response by different parasite
genotypes that indirectly impact the host microbiome. Variations in the immune response the host
mount during infection can either result from locally adapted parasites that differentially
manipulate the host immune system [90, 92], or from the host mounting a defense mechanism that
differentiate between parasite genotypes [33, 73, 93, 94]. In support of parasite local adaptation,
we showed that parasites of different origin differentially impact the correlation between host
microbiome composition and immune gene expression. Our results suggest that variations in the
microbiome of infected hosts vary with the mechanisms of local host-parasite counter adaptation
due to the tight association between microbiome composition and immune gene expression.

Conclusions and perspectives

Tapeworms represent an important problem for health authorities with over 40 species
infecting humans as definitive hosts, and 15 during the larval stage, many of which are included
in the list of neglected zoonotic diseases by the World Health Organization due to their impact on
human health and morbidity in impoverished nations [95]. But helminths in general are also under
study because of their beneficial regulatory role in autoimmune diseases, and inflammatory
diseases including atopic dermatitis, inflammatory bowel diseases, asthma and allergies.
Helminthes typically induce a Th2 and Treg immune response. There is a growing recognition of
the role of microbes in helminths ability to modulate the host immune system and of the imperative
to study host-parasite-microbe interactions [96, 97]. Herein, we showed the presence of a complex microbiome within *S. solidus* and provided evidence for non-random selection of microbes from the parasite environment. We also provide evidence that some of the parasite-associated bacteria contribute to modulating the host immune system. Recently, we also discovered that viruses are associated with *S. solidus* and that some of these are transmitted to parasitized hosts, which means that viruses could also impact the host immune response [98]. Further studies are needed to test whether *S. solidus* relies on these microbes for infection, or whether they are detrimental for the parasite. Herein, we also demonstrate that the host microbiome, which varies with host genotype in healthy individuals, is impacted by infection and depends on the parasite genotype in infected individuals. The relative abundance of key members of the microbiome correlated with expression of Treg, complement factor, and innate gene involved in recruiting cells to the site of inflammation, and was impacted by infection status and parasite origin. This result suggest that infection by different parasites differentially impact the immune communication between the host and its microbiome, and that constitutive microbiome differences among hosts do not contribute to differences in immune responsiveness to parasite infection. Further studies need to be conducted to assess the role of host-associated and parasite-associated microbes in parasite fitness, and to determine if microbes could be leveraged for developing new treatment strategies [99].

6 Methods

*Ethics statements*

The sampling was conducted under Fish sampling permit #SF2015-263 and #P17-025 and fish transport permits #15A-0084 and 17A-0024 provided by the State of Alaska Department of Fish and Game to Michael Bell. Animal sampling and experiments in Germany and Norway were approved by the Ministry of Energy Transition, Agriculture, Environment and Rural Areas of
Schleswig-Holstein under reference number V 312–7224.123-34. All experiments were performed in accordance with relevant guidelines and regulations in the Public Health Service Policy (PHS) on Humane Care and Use of Laboratory Animals.

**Processing of field collected specimens**

In September 2015, threespine stickleback (*Gasterosteus aculeatus*) were collected from Cheney Lake in Anchorage Alaska (61° 12’ 17” N, 149° 45’ 33” W) to conduct an initial field study. Fish were caught using un-baited minnow traps placed along the north shoreline of the lake between 0.25 and 2 m deep, and at least 2 m apart. Fish were shipped to Stony Brook University, kept in 20-gallon tanks of 6% seawater at 5°C at the Flax Pond Marine Lab, and fed mysid shrimp twice per day. Following an acclimation period of 3 weeks, and 12 hours of starvation, fish were anaesthetized and euthanized in tricaine methane sulphonate solution (MS222), decapitated, and dissected in sterile conditions. Prior to and in between dissections all surfaces were cleaned with 80% ethanol and betadine to remove contaminants. Before dissection, fish were brushed with betadine to prevent contamination of the body cavity with microbes from the surface of the fish. For dissection, an incision was made along the lateral line of the fish body, around the bony pelvis. The cut extended from the pectoral fins to just anterior of the anus to avoid cutting the intestine. The sex of the fish was assessed by visual inspection of gonads at the time of dissection and then confirmed using PCR with sex specific primers as described in [100]. The presence of *S. solidus* was recorded, and stomach and intestines were flash frozen in liquid nitrogen for future DNA extraction. Swab samples of the body cavity of two non-infected and two infected individuals were also collected. Two parasites were also vigorously shaken in sterile phosphate buffered saline (PBS) and the resulting PBS was stored. These swab and PBS samples were used as confirmation
that any bacteria found in *S. solidus* were indeed part of an endomicrobiome and not contamination from the fish body cavity or surface of the parasite. All samples were stored at -80°C until use.

**Fish and Parasite collection for the cross-infection experiment**

In Spring of 2016, threespine sticklebacks were caught, as described above, from Wolf lake (WO, 61°38'36" N, 149°16'32" W) and Walby Lake (WA, 64°37'12" N, 149°12'36" W) in Alaska, Grosser Plöner See (GPS, 54°9'16"N, 10°25'14"E) in Germany, and Lake Skogseidvatnet (SKO, 60°14'44"N, 5°55'03"E) in Norway in preparation for the cross infection experiment.

For Wolf, Walby and GPS, *G. aculeatus* eggs were dissected out of gravid females euthanized in MS-222. Eggs were fertilized with sperm suspension collected from euthanized male sticklebacks and then washed with acriflavine (50µL; 30 sec), methylene blue (500µL per L from stock: 1 g/L methylene blue; 30 sec), and 3ppt artificial seawater. Eggs were kept at 4°C and shipped to the Max Planck Institute for Evolutionary Biology (MPI), Plon, Germany. Fin clips were taken from the parents of each family to enable microsatellite analyses post experiment. Upon maturation, fish were kept in a flow through system and fed *ad libitum* bloodworms every other day.

Mature *S. solidus* plerocercoids from Wolf, Walby and SKO were shipped to the MPI in freshly caught conspicuously parasitized fish that were dissected immediately upon arrival. Pairs of *S. solidus* plerocercoids were weight-matched and bred *in vitro* [101]. The eggs were then kept at 4°C in the dark until use.

**Cross infection experiment**

We conducted experimental infections of threespine sticklebacks with *S. solidus* in the spring of 2017. Briefly, for each locality, three families of fish and four families of parasites were used (Wolf, Walby, and GPS for fish and Wolf, Walby and SKO for parasites). *Macrocyclops albidus* copepods from a laboratory stock were singly infected with *S. solidus* procercoïds as
previously described [102]. Fish were starved for 24 hours before being fed either one singly infected copepod or one non-exposed copepod (sham control). After 2 days, fish were transferred into 16L aquaria. Fish exposed to a given parasite family were held together in the same tank. Each tank held five exposed fish from Wolf, five exposed fish from GPS and four exposed fish from Walby, in addition to one control fish per fish population (17 fish/tank). The common garden design aims at controlling for tank effects observed in flow through systems (Figure 2) [103, 104]. Fish were fed with frozen Chironomids larvae three times a week. Due to the large number of tanks used, we conducted three rounds of experimental infection. Each round, different families of fish, and different families of parasites from each of the three populations were used.

Nine weeks after exposure, fish were euthanized with MS-222 and had morphometric data recorded including fish standard length, fish total length, and fish weight. Upon dissection, masses of the spleen, liver, and head kidney were weighed, and these tissues were preserved in RNAlater (Thermofisher CAT# AM7020). Intestine and plerocercoid (from successfully infected fish) samples were flash frozen in liquid nitrogen for DNA extraction. A fin clip was collected and preserved in ethanol to be used for microsatellite analysis and to determine fish origin. DNA was extracted from fin clips collected from parents in June of 2016 for each of the different fish genotypes used in the experiment. DNA was then extracted from preserved fin clips from experimental fish and microsatellite analysis was used to assign fish in the common garden design to the genotypes they originated from.

Using recorded morphometrics, we determined the fish’s condition (condition factor, CF), hepatosomatic index (HIS), and immunological activation (splenosomatic index, SSI; head kidney index, HKI) [105, 106]. The parasite index (PI, the relative weight of the parasite in the host) was also calculated if the fish was infected [107].
Quantitative real-time PCR (RT-qPCR) of immune response genes

Head kidney RNA was extracted with a NucleoSpin® 96 kit (Macherey-Nagel) following the manufacturer’s protocol. RNA concentration and purity were determined spectrophotometrically (NanoDrop1000; Thermo Scientific). We used the Omniscript RT kit (Qiagen) according to the manual but used 0.2 µl of a 4-unit RNase inhibitor (Qiagen) per reaction. Differences in gene transcription levels were tested using 96.96 Dynamic Array IFCs on a Biomark™ HD system (Fluidigm) with EvaGreen as DNA intercalating dye on all the threespine sticklebacks used for this study as previously described. Briefly, we pre-amplified the cDNA samples by using TaqMan PreAmp Master Mix (Applied Biosystems) according to the manufacturer’s protocol (14 cycles). All targets for a given sample were included in the same run and measured in technical triplicates. Inter-run calibrators and negative controls were included on each IFC. Melting curves were analysed with the Fluidigm Analysis software v.4.5.1.

Bacterial microbiome composition and bacterial load

DNA was extracted from intestines (In) and stomach (St) collected from 6 non-infected and 6 infected field caught individuals from Cheney lake. In addition, DNA was extracted from the intestine of 148 individuals from Walby (56 individuals), Wolf (38 individuals), and GPS (54 individuals) that were either (i) sham-exposed (42 individuals), (ii) exposed but non-infected by *S. solidus* from SKO (10 individuals), Walby (13 individuals) and Wolf (13 individuals), or (iii) exposed and successfully infected by *S. solidus* from SKO (18 individuals), Walby (28 individuals) and Wolf (24 individuals) (Table S1). The DNA of *S. solidus* worms from SKO (18 individuals), Walby (28 individuals) and Wolf (24 individuals) collected from ESI sticklebacks from Walby (27 individuals), Wolf (19 individuals), and GPS (24 individuals) was also extracted. No Wolf Lake
fish were successfully infected by SKO parasites, so that neither fish intestines nor S. solidus worms were collected for this host-parasite combination.

DNA extraction was performed in the same manner for both field sampled and cross infection samples using a Powerlyzer Powersoil Kit (Cat#12855) with the following modifications to the manufacturer’s protocol. First, all tissue samples were placed in tough tubes (Qiagen CAT# 19301) with 3mm glass beads (Sigma CAT# S72683). To each tube, 750μL of bead beating solution supplied with the extraction kit was added. Samples underwent three rounds of 20 sec bead beating cycles and were kept on ice between each round. The homogenate was transferred to the Qiagen supplied 0.1mm bead tube and incubated with 60μL of C1 solution for 10 min at 60°C. After vortexing, centrifugation, and filtration with solution C5, 75μL of solution C6 was added to the spin filter in a clean 2mL tube and incubated at room temperature for 5 min. Two rounds of centrifugation at 10,000g for 30 sec were used to recover 150μL final DNA extract. Six negative controls of molecular biology grade water were processed together with samples to identify kit contaminants.

**Illumina Sequencing**

Sequencing was conducted by the Mr DNA lab in Shallowater, TX. Briefly, the 16S rDNA V4-V5 hypervariable region was amplified with *E. coli* 515f and 806r primers (with barcode on the forward primer) as specified by the Earth Microbiome [108-110]. After amplification, PCR products were checked in 2% agarose gel to determine the success of amplification and the relative intensity of bands. PCR products were pooled together in equal proportions based on their molecular weight and DNA concentrations. Pooled samples were purified before Illumina DNA library preparation. Illumina sequencing was employed to obtain 250bp paired end reads on a
MiSeq following the manufacturer’s guidelines. Sequence data were initially processed to join forward and reverse reads and remove barcodes.

Sequencing of field collected samples yielded an average of 43,645 (±SE 16,027) and 10,800 (±SE 8,824) high quality paired end reads for *G. aculeatus* and *S. solidus* tissue samples respectively. Sequencing depth of control samples was higher and yielded an average of 83,875 (±SE 10,080) high quality reads. Sequencing of cross infection samples yielded an average of 53,827 (±SE 24,004) high quality paired end reads for fish samples, 55,308 (±SE 28,426) for *S. solidus* samples, and 69,273 (±SE 16,277) high quality reads for kit control samples. Due to the large number of samples sequenced, multiple sequencing runs were processed. In order to extract the same Amplicon Sequence Variants (ASVs) for all sequencing runs, each run was demultiplexed, sequences quality scores were controlled, and sequences were trimmed at the same quality threshold to obtain sequences of a uniform 200 bp length. ASVs were resolved to avoid applying the arbitrary 97% threshold that defines molecular OTUs but has limited biological meaning [111]. Taxonomy was assigned to each of the ASVs using a pretrained Naïve-Bayes Classifier trained on the GreenGenes 13_8 99% OTUs.

Swabs of the body cavity of *G. aculeatus* and surface bacteria from *S. solidus* used as controls had a significantly different microbiome from fish and parasite samples (Figure S4). Similarly, blank controls used for the cross-infection experiments had a different microbiome from fish and parasite samples. ASVs abundant in all control samples were extremely similar to common bacterial contaminants of the MoBio kit used for bacterial DNA extraction and were removed from the analysis [112-114]. ASVs classified only to the domain level or listed as unidentified were blasted against the GenBank reference database and sequences that aligned to non-bacterial organisms were removed. We further decontaminated samples using the Decontam
package in R which has been found to be successful at decontaminating data sets with low biomass samples [115,116].

**qPCR of 16S gene**

Real-time qPCR was conducted using the QuantStudio 6 Flex RT PCR System (Fisher Scientific). The qPCR mixture (20μL) was composed of 10μL of SYBR green master mix (Thermofisher CAT# 4309155), 4.8μL molecular biology grade water (Thermofisher CAT#46000), 0.6μL forward primer, 0.6μL reverse primer, and 4μL DNA. Two 16S primers were used (Forward: TCCTACGGGAGGCAGCAGT, Reverse: GACTACCAGGATATCTAATCTGT [108] and Forward: GTGSTGAYGGYTGTCGTCA, Reverse: ACWGRTCCMCCACCTTCC [109]) and compared to an *E. coli* standard for which cell count was obtained using a hemocytometer. After quantitative PCR, CT values were averaged and used to calculate the relative number of cells using the slope formula generated with the *E. coli* standard curve. Even though the number of 16S gene copies per bacterial cell cannot be controlled, the qPCR assay provides a good proxy for bacterial abundance [117-118].

**Diversity metrics and statistical analyses**

**Diversity calculations and field sample analysis**

All data analyses were performed using QIIME2 and the Phyloseq, and lmertest packages in R [119-122]. First, Shannon, Simpson, and Fisher alpha diversity metrics were used to estimate within sample diversity and compared using Tukey comparison of means. Beta diversity was calculated with Weighted and Unweighted Unifrac Distances. Principal coordinate analysis (PCoA) on Unifrac distances was performed to look for clustering of samples in unconstrained
Coordinates along PCoA, PERMANOVA, and PERMDISP were used to compare differences between sample groups. Heatmaps were generated using Excel.

**Cross infection microbiome models**

In order to account for all potential factors, statistical analyses for the cross infection experiment were conducted using linear mixed models using the lmertest package in R using each of the different microbiome indices (Shannon, Simpson, Fisher, PCoA1, PCoA2, PCoA3, and bacterial load) as the dependent variable for a given model [120]. We began by testing the impact on fish population in control non-exposed fish (CNI) (Figure 2A). We conducted model selection starting with the most complex model which included fish population, sex and their interactions as fixed factors and tank and round as random factors down to the simplest model which included only fish population as a fixed factor and tank and round as random factors. We then calculated the AIC for each of the models and chose the model with the lowest AIC to run. If the AIC values differed by less than a value of two for the best fit models, we then conducted ANOVA to see if there was significant variance between the best fit model. If there were significant differences, the model with the lowest AIC was selected. If there were no significant differences the simplest model was selected. P-values were derived from the selected model using least squared means (LSM) calculated with the lsmeans package [122].

To investigate the impact of infection status on these indices, we used infection status, which included the levels CNI (control non-infected), ENI (exposed non-infected), and ESI (exposed and successfully infected), sex, and their interaction as potential fixed factors, and random factors were fish population, parasite population, tank, and round. Model selection was performed as described above and the best-fit model was chosen to calculate p-values (Figure 2).
Following this we tested the impact of fish and parasite population in exposed non-infected fish (ENI), in successfully infected fish (ESI), and parasites (Ss). After conducting model selection as described above, we tested for the role of fish population and parasite population separately (Figure 2). When testing fish population effects, we used fish population, sex, and their interaction as fixed factors, and random factors were parasites population, tank, and round. When testing parasite population effects, we used parasite population, sex, and their interaction as fixed factors and random factors were fish population, tank, and round. Model selection was conducted each time and the best fit model was chosen to calculate p-values.

**Differentially abundant bacteria**

We used the DESeq2 package in R [123] to identify differentially abundant bacteria. From field samples, DESeq2 was used to compare bacterial abundance between threespine sticklebacks and *S. solidus*, between intestine and stomach, and between non-infected and infected fish. From the cross-infection experiment, DESeq2 was used to test variations in bacterial abundance between threespine sticklebacks and *S. solidus*, depending on infection status (CNI, ENI and ESI), fish population (for Ss, CNI, ENI and ESI), parasite population (for Ss, ENI and ESI) while controlling for tank, round effects, and fish and parasite population as appropriate by including these variables in the models.

**Correlation between gene expression and microbiome parameters**

We assessed whether microbiome diversity metrics and the relative abundance of most abundant bacterial families correlated with phenotypic data and gene expression data. First, we conducted a principal component analysis (PCA) with on 25 genes including three regulatory
genes (*abtb1, kat2a, mapk13*) and 22 immune related genes from innate immunity (*marco, mst1ra, mif, tnfr1, saal1, tlr2, csf3r, p22phox, nkef-b, sla1, cd97*), adaptive immunity (*stat4, stat6, igm, cd83, foxp3, il-16, mhcII, tcr-β*) and the complement system (*cfb, c7, c9*) for all samples for which the microbiome composition had been assessed. Next, we conducted a correlation analysis including all microbiome diversity metrics (Log Cells, Fisher, Shannon, Simpson, weighted unifrac PC1 to 3 and unweighted unifrac PC 1 to 3) and the 30 most abundant bacterial families, together with the three top gene PCA coordinates and relative expression of all 25 tested genes. We used Spearman’s rank correlation because it performs better with log transformed data and microbiome relative abundance and allows comparisons between datasets with different distribution profiles. Due to multiple testing, p values were corrected for multiple comparisons with a Bonferroni correction.

7 Acknowledgements

Not applicable.

8 References


28. Houlden, A., et al., Chronic Trichuris muris infection in C57BL/6 mice causes significant changes in host microbiota and metabolome: effects reversed by pathogen clearance. 2015. 10(5).


56. Hahn, M.A. and N.M. Dheilly, Experimental Models to Study the Role of Microbes in Host-Parasite Interactions. Frontiers in Microbiology, 2016. 7(1300).


Figure 1: Summary of experimental design. (A) The map displays all sampling locations. The field study was conducted on sticklebacks and *S. solidus* from one location in Alaska (Cheney).
The three genotypes of stickleback hosts (Walby (W), Wolf (w) and GPS(G)) and three genotypes of *S. solidus* (Walby (W), Wolf (w) and SKO (S)) used for experimental infection were collected from lakes in Alaska and Europe. (B) Design matrix showing the organization of rounds and tanks based on fish genotype (denoted by letters) and parasite genotype (denoted by numbers). A total of 36 tanks were used. One example is squared in red. (C) Each of the tanks contained three sham control fish (left) from Wolf (Red Tail), Walby (Green Tail), and GPS (Blue Tail) separated by a mesh divider from 5 Wolf fish, 5 GPS fish, and 4 Walby fish that had been exposed to parasites (right). (D) Upon dissection, the success of infection was assessed, and exposed individuals were classified as exposed non-infected (ENI) or exposed successfully infected (ESI). All successfully infected fish were processed, and corresponding ENI and control non-infected fish from the same tanks and fish origin were processed as controls.

**Figure 2: Stickleback bacterial community composition vary with exposure, infection status, host genotype and parasite genotype as revealed by LMMs.** (A) Among the control non-infected fish, the microbiome composition varies with fish genotype. Differences were found between Walby (n=17) and GPS fish (n=15) (Weighted Unifrac, Axes 1 LMM p=0.038; and Axis 3, Unweighted Unifrac, Axis 3, p=0.028) and between Wolf (n=10) and GPS fish (Unweighted Unifrac Axis 1, p=0.024). (B) Exposure and infection impact the microbial diversity and composition. Exposure impacts beta diversity (Weighted Unifrac, Axis 3, p=0.031), whereas successful infection is associated with changes in bacterial load (p=0.01) and alpha diversity (Shannon p=0.039, Simpson p=0.013). (C) Among exposed but non-infected fish, the microbiome diversity and composition vary with host genotype and parasite genotype. Differences were found between Wolf (n=9) and GPS (n=14) fish (Fisher p=0.008, Unweighted Unifrac Axis 1 p=0.005,
and 3 p=0.013), between Walby (n=10) and Wolf fish (Shannon p=0.034, Simpson p=0.020, Fisher
p=0.004, Unweighted Unifrac Axis 1 p=0.006) and between Walby and GPS fish (Unweighted
Unifrac Axis 3 p=0.041). Differences were found between alpha diversity of fish exposed to Walby
and SKO parasites (Fisher p=0.018). (D) Among infected sticklebacks, microbiome composition
varies with parasite origin and host-parasite co-evolutionary history. Fish infected with SKO
parasites (n=18) had higher bacterial load than fish infected with Walby (n=27, p=0.010) and Wolf
parasites (n=25, p=0.001). Unweighted Unifrac distance revealed distinct clustering of fish
infected with SKO parasites and fish infected with Walby or Wolf parasites (Axis 2, p=0.046 and
p=0.047 respectively). (E) Within parasites, microbiome composition varies with parasite origin.
Weighted Unifrac distance revealed distinct clustering of the microbiome of SKO (n=18) and
Walby (n=27) parasites (Axis 3, p=0.035). The bottom and top edges of the boxes mark the 25th
and 75th percentiles. The horizontal line denotes the median. Whiskers mark the range of the data
excluding outliers (dots).

Figure 3: Bacterial taxa at multiple taxonomic levels vary in relative abundance and
prevalence between threespine stickleback gut and S. solidus plerocercoids. The heatmaps
depict the prevalence of the bacterial families (A) or ASVs (B) at different detection thresholds
(defined in terms of minimum relative abundance) for successfully infected (ESI) sticklebacks,
and corresponding S. solidus. A/ 30 most abundant bacterial families across all samples. B/ 30
most abundant ASVs in S. solidus. The phylum is provided as a color code next to the taxa name.

Figure 4: The host phenotype depends on the composition of the microbiome of the host and
Color of the squares indicate the strength of Spearman’s rank correlation. Asterisks indicate significance of the correlation (* q<0.05; ** q<0.05; *** q<0.001). B. Comparison of correlations between gene PC1 and uwuPC3 depending on infection status, host genotype and parasite genotype. Significant differences are observed when the confidence interval at the lower and/or higher values do not overlap with the mean of the other one. C. A. Correlation heatmap between the expression of genes, and phenotypic indices (HI, Hepatosomatic index; SI, Splenosomatic index, HKI Head-kidney index; BC, Body condition) and the relative abundance in the most abundant acterial families in the host *G. aculeatus*, and its parasite *S. solidus*. Color indicate the strength of Spearman’s rank correlation. The complete figure with indication of significance is available in Figure S12. Significant Spearman’s rank correlation (with FDR <0.05) are plotted in (D) between the relative abundance in *Streptococcaceae*, an unknown Actynomicetales and *Aerococcaceae* in the host with Heptaosomatic index (N=145), in (E) between the relative abundance of *Synechoccophycidae* in fish and the relative expression of *foxp3, il16, cd97 and stat6* (N=145), in (F) between the relative abundance in *Lactobacillaceae* in fish and the relative expression of *foxp3, cd97 and stat6* (N=145), in (G) between the relative abundance in *Pseudanabaenaceae* in fish and the relative expression of *foxp3, cd97 and stat4* (N=145), in (H) between the relative abundance of *Vibrionaceae* in fish and relative expression of of *foxp3, il16, cd97 and stat6* (N=145) and in (I) between the relative abundance of *Chloroflexi* in the parasite and relative abundance in *foxp3, tnfr1, cd97, stat6, and marco* (N=69). The significant Spearman correlation, after Bonferroni correction, are provided on each plot.
Summary of experimental design. (A) The map displays all sampling locations. The field study was conducted on sticklebacks and S. solidus from one location in Alaska (Cheney). The three genotypes of stickback hosts (Walby (W), Wolf (w) and GPS(G)) and three genotypes of S. solidus (Walby (W), Wolf (w) and SKO (S)) used for experimental infection were collected from lakes in Alaska and Europe. (B) Design matrix showing the organization of rounds and tanks based on fish genotype (denoted by letters) and parasite genotype (denoted by numbers). A total of 36 tanks were used. One example is squared in red. (C) Each of the tanks contained three sham control fish (left) from Wolf (Red Tail), Walby (Green Tail), and GPS (Blue Tail) separated by a mesh divider from 5 Wolf fish, 5 GPS fish, and 4 Walby fish that had been exposed to parasites (right). (D) Upon dissection, the success of infection was assessed, and exposed individuals were classified as exposed non-infected (ENI) or exposed successfully infected (ESI). All successfully infected fish were processed, and corresponding ENI and control non-infected fish from the same tanks and fish origin were processed as controls. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.
Figure 2

Stickleback bacterial community composition vary with exposure, infection status, host genotype and parasite genotype as revealed by LMMs. (A) Among the control non-infected fish, the microbiome composition varies with fish genotype. Differences were found between Walby (n=17) and GPS fish (n=15) (Weighted Unifrac, Axes 1 LMM p= 0.038; and Axis 3, Unweighted Unifrac, Axis 3, p=0.028) and between Wolf (n=10) and GPS fish (Unweighted Unifrac Axis 1, p=0.024) . (B) Exposure and infection impact the microbial diversity and composition. Exposure impacts beta diversity (Weighted Unifrac, Axis 3, p=0.031), whereas successful infection is associated with changes in bacterial load (p=0.01) and
alpha diversity (Shannon p=0.039, Simpson p=0.013). (C) Among exposed but non-infected fish, the microbiome diversity and composition vary with host genotype and parasite genotype. Differences were found between Wolf (n=9) and GPS (n=14) fish (Fisher p=0.008, Unweighted Unifrac Axis 1 p=0.005, and 3 p=0.013), between Walby (n=10) and Wolf fish (Shannon p=0.034, Simpson p=0.020, Fisher p=0.004, Unweighted Unifrac Axis 1 p=0.006) and between Walby and GPS fish (Unweighted Unifrac Axis 3 p=0.041). Differences were found between alpha diversity of fish exposed to Walby and SKO parasites (Fisher p=0.018). (D) Among infected sticklebacks, microbiome composition varies with parasite origin and host-parasite co-evolutionary history. Fish infected with SKO (n=18) parasites had higher bacterial load than fish infected with Walby (n=27, p=0.010) and Wolf parasites (n=25, p=0.001). Unweighted Unifrac distance revealed distinct clustering of fish infected with SKO parasites and fish infected with Walby or Wolf parasites (Axis 2, p=0.046 and p=0.047 respectively). (E) Within parasites, microbiome composition varies with parasite origin. Weighted Unifrac distance revealed distinct clustering of the microbiome of SKO (n=18) and Walby (n=27) parasites (Axis 3, p=0.035). The bottom and top edges of the boxes mark the 25th and 75th percentiles. The horizontal line denotes the median. Whiskers mark the range of the data excluding outliers (dots).

Figure 3

Bacterial taxa at multiple taxonomic levels vary in relative abundance and prevalence between threespine stickleback gut and S. solidus plerocercoids. The heatmaps depict the prevalence of the bacterial families (A) or ASVs (B) at different detection thresholds (defined in terms of minimum relative abundance) for successfully infected (ESI) sticklebacks, and corresponding S. solidus. A/ 30 most abundant bacterial
families across all samples. B/ 30 most abundant ASVs in S. solidus. The phylum is provided as a color code next to the taxa name.

Figure 4

The host phenotype depends on the composition of the microbiome of the host and of its parasite. A. Correlation heatmap depicting gene expression – microbial profile correlations. Color of the squares indicate the strength of Spearman’s rank correlation. Asterisks indicate significance of the correlation (*
q< 0.05; ** q<0.05; *** q<0.001). B. Comparison of correlations between gene PC1 and uwuPC3 depending on infection status, host genotype and parasite genotype. Significant differences are observed when the confidence interval at the lower and/or higher values do not overlap with the mean of the other one. C. A. Correlation heatmap between the expression of genes, and phenotypic indices (HI, Hepatosomatic index; SI, Splenosomatic index, HKI Head-kidney index; BC, Body condition) and the relative abundance in the most abundant acterial families in the host G. aculeatus, and its parasite S. solidus. Color indicate the strength of Spearman's rank correlation. The complete figure with indication of significance is available in Figure S12. Significant Spearman's rank correlation (with FDR <0.05) are plotted in (D) between the relative abundance in Streptococcaceae, an unknown Actynomicetales and Aerococcaceae in the host with Heptaosomatic index (N=145), in (E) between the relative abundance of Synechococophycidae in fish and the relative expression of foxp3, il16, cd97 and stat6 (N=145), in (F) between the relative abundance in Lactobacillaceae in fish and the relative expression of foxp3, cd97 and stat6 (N=145), in (G) between the relative abundance in Pseudanabaenaceae in fish and the relative expression of foxp3, cd97 and stat4 (N=145), in (H) between the relative abundance of Vibrionaceae in fish and relative expression of of foxp3, il16, cd97 and stat6 (N=145) and in (I) between the relative abundance of Chloroflexi in the parasite and relative abundance in foxp3, tnfr1, cd97, stat6, and marco (N=69). The significant Spearman correlation, after Bonferroni correction, are provided on each plot.