



# Metagenomic Shotgun Analyses Reveal Complex Patterns of Intra- and Interspecific Variation in the Intestinal Microbiomes of Codfishes

Even Sannes Riiser,<sup>a</sup> Thomas H. A. Haverkamp,<sup>a\*</sup> Srinidhi Varadharajan,<sup>a</sup> Ørnulf Borgan,<sup>b</sup> Kjetill S. Jakobsen,<sup>a</sup> Sissel Jentoft,<sup>a</sup> Bastiaan Star<sup>a</sup>

<sup>a</sup>Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, Oslo, Norway

<sup>b</sup>Department of Mathematics, University of Oslo, Oslo, Norway

**ABSTRACT** The relative importance of host-specific selection or environmental factors in determining the composition of the intestinal microbiome in wild vertebrates remains poorly understood. Here, we used metagenomic shotgun sequencing of individual specimens to compare the levels of intra- and interspecific variation of intestinal microbiome communities in two ecotypes (NEAC and NCC) of Atlantic cod (*Gadus morhua*) that have distinct behavior and habitats and three Gadidae species that occupy a range of ecological niches. Interestingly, we found significantly diverged microbiomes among the two Atlantic cod ecotypes. Interspecific patterns of variation are more variable, with significantly diverged communities for most species' comparisons, apart from the comparison between coastal cod (NCC) and Norway pout (*Trisopterus esmarkii*), whose community compositions are not significantly diverged. The absence of consistent species-specific microbiomes suggests that external environmental factors, such as temperature, diet, or a combination thereof, comprise major drivers of the intestinal community composition of codfishes.

**IMPORTANCE** The composition of the intestinal microbial community associated with teleost fish is influenced by a diversity of factors, ranging from internal factors (such as host-specific selection) to external factors (such as niche occupation). These factors are often difficult to separate, as differences in niche occupation (e.g., diet, temperature, or salinity) may correlate with distinct evolutionary trajectories. Here, we investigate four gadoid species with contrasting levels of evolutionary separation and niche occupation. Using metagenomic shotgun sequencing, we observed distinct microbiomes among two Atlantic cod (*Gadus morhua*) ecotypes (NEAC and NCC) with distinct behavior and habitats. In contrast, interspecific patterns of variation were more variable. For instance, we did not observe interspecific differentiation between the microbiomes of coastal cod (NCC) and Norway pout (*Trisopterus esmarkii*), whose lineages underwent evolutionary separation over 20 million years ago. The observed pattern of microbiome variation in these gadoid species is therefore most parsimoniously explained by differences in niche occupation.

**KEYWORDS** Atlantic cod, *Gadus morhua*, codfishes, intestinal microbiome, metagenomics

Significant research effort has focused on the importance of external, environmental factors (e.g., habitat, geography, microbial biodiversity, diet, water temperature, or salinity) and internal, host-related factors (e.g., genetics, physiology or immunity) in driving the composition of the intestinal microbiome in fish (1, 2). That external factors play an important role is well established. For instance, bacterial diversity in the

**Citation** Riiser ES, Haverkamp THA, Varadharajan S, Borgan Ø, Jakobsen KS, Jentoft S, Star B. 2020. Metagenomic shotgun analyses reveal complex patterns of intra- and interspecific variation in the intestinal microbiomes of codfishes. *Appl Environ Microbiol* 86:e02788-19. <https://doi.org/10.1128/AEM.02788-19>.

**Editor** Christopher A. Elkins, Centers for Disease Control and Prevention

**Copyright** © 2020 Riiser et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Even Sannes Riiser, [e.s.riiser@ibv.uio.no](mailto:e.s.riiser@ibv.uio.no), or Bastiaan Star, [bastiaan.star@ibv.uio.no](mailto:bastiaan.star@ibv.uio.no).

\* Present address: Thomas H. A. Haverkamp, Department of Epidemiology, Norwegian Veterinary Institute, Oslo, Norway.

**Received** 3 December 2019

**Accepted** 6 January 2020

**Accepted manuscript posted online** 17 January 2020

**Published** 2 March 2020

surrounding water influences the intestinal microbiome in fish larvae and fry (3, 4); water temperature is the main driver for the gut microbiome composition in farmed Tasmanian Atlantic salmon (*Salmo salar*) (5); and diet influences the intestinal composition in both experimental (6–9) as well as wild fish populations (10–13). Yet internal factors also influence the composition of these bacterial communities. For instance, observations of a shared (core) microbiome between wild and laboratory-raised zebrafish suggest that distinct selective pressures determine the composition of the microbial communities (14). Moreover, an association between host phylogeny and intestinal microbiome composition has been observed for a range of fishes, marine animals, and terrestrial mammals (15–19).

The adaptive immune system appears especially important for host selection. Individual variation of the major histocompatibility complex (MHC) II correlates with the gut microbiome composition in stickleback (20); mucosal IgT depletion causes dysbiosis in rainbow trout (*Oncorhynchus mykiss*) (21); and lack of a functional adaptive immune system reduces the strength of host selection in knockout zebrafish models (22). Among bony fish, gadoid fishes have an unusual adaptive immune system in which there is loss of MHC II, CD4, and invariant chain (Ii) and in which there is a range of innate (TLR) and MHC I immune-gene expansions (23, 24). Moreover, Atlantic cod has high levels of IgM (25) and a minimal antibody response after pathogen exposure (25–27). Gadoids therefore provide an interesting ecological system to study host-microbiome interactions (28).

Studies that specifically integrate internal and external influences support a role for both factors driving the microbial community composition (13, 29). Such studies, however, remain restricted in both the level of taxonomy of fishes (30) as well as taxonomical resolution of the microbial analyses (16S rRNA) (13, 29, 31–33). Importantly, it often remains difficult to separate the correlated effects of distinct behavior (e.g., diet) and niche occupation with interspecific selection. Also, no comparative studies have used metagenomic shotgun sequencing to investigate fish populations with profound differences in behavior within a single species. It therefore remains unclear whether the microbial composition for a range of wild fish species is characterized by intra- or interspecific divergence.

Here, we study intra- and interspecific divergence of intestinal microbial communities within the widespread family of Gadidae using a metagenomic shotgun data set. We compare the microbiomes from Norway pout (*Trisopterus esmarkii*), poor cod (*Trisopterus minutus*), northern silvery pout (*Gadiculus thori*), and two ecotypes of Atlantic cod (*Gadus morhua*). These four species have overlapping geographical distributions, are dietary generalists in typically feeding over sandy and muddy bottoms on pelagic or benthic crustaceans, polychaetas, and (small) fish (34, 35), and evolutionarily diverged approximately 20 million years ago (24). Norway pout is benthopelagic, distributed from the English Channel, around Iceland, and up to the Southwest Barents Sea. It is mostly found at depths between 100 and 200 m. Poor cod is also benthopelagic, distributed from the Trondheim Fjord in Norway to the Mediterranean Sea, and mostly found between 15 and 200 m. Northern silvery pout (*Gadiculus thori*) is meso- to bathypelagic (36), distributed in the North Atlantic Ocean, along the coast of Norway, and around Iceland and Greenland. It forms large schools that are usually found between 200 and 400 m (34, 36, 37). Finally, Atlantic cod has a trans-Atlantic distribution, from the Bay of Biscay to the Barents Sea, the Baltic Sea, around Iceland and Greenland, in the Hudson Bay, and along the North American coast (34, 38–42). Atlantic cod comprises various subpopulations and “ecotypes” with distinct adaptations, migratory behavior, and feeding behavior. For instance, northeast Arctic cod (NEAC) performs typical spawning migrations from the Barents Sea to the Norwegian coast, whereas the Norwegian coastal cod (NCC) remains more stationary (34, 43). These ecotypes have increased genomic divergence in several large chromosomal inversions (43–47), suggestive of local adaptation. The environments that these two ecotypes encounter are different, and they feed on distinct types of food. NEAC consumes mostly capelin and herring while NCC feeds on a wide range of crustaceans,

**TABLE 1** Species collected and sample locations

Species	Latin name	Ecotype	Sampling location	<i>n</i>	Abbreviation
Atlantic cod	<i>Gadus morhua</i>	Northeast Arctic cod	Lofoten	10	NEAC
Atlantic cod	<i>Gadus morhua</i>	Norwegian coastal cod	Lofoten	10	NCC
Atlantic cod	<i>Gadus morhua</i>	Norwegian coastal cod	Oslo fjord	2	NCC_Oslo
Poor cod	<i>Trisopterus minutus</i>		Oslo fjord	5	PC
Norway pout	<i>Trisopterus esmarkii</i>		Oslo fjord	4	NP
Northern silvery pout	<i>Gadiculus thori</i>		Oslo fjord	3	NSP

fish, and seaweed (34, 39, 48). During spawning, these ecotypes spatially co-occur, and long-term gene flow between ecotypes is supported by low overall estimates of divergence in most genomic regions, apart from the chromosomal rearrangements (43).

We hypothesize that if interspecific selection (indicative of host selection) is the main driver for the intestinal communities in the Gadidae, most differences will be found between the different species, and not between the different ecotypes within Atlantic cod. In contrast, if environmental factors are the main drivers for the intestinal communities, we expect significant compositional differences between the ecotypes of Atlantic cod, as well as various levels of differentiation between the species. We used taxonomic profiling of metagenomic shotgun reads to classify these microbiomes (obtained from various locations around the Norwegian coast [Table 1]) at order- and species-level resolution and analyzed within-species differentiation of the most abundant members by genome-wide single nucleotide variation. Finally, differences in gut bacterial community composition among the species and ecotypes were assessed using multivariate statistics.

## RESULTS

**Taxonomical composition of the intestinal microbiomes.** We analyzed a data set of 422 million paired-end reads, with a median sample size of 11.9 million reads (8.0 to 19.6 million reads per sample) (Table 2, Table S7 in the supplemental material). Following filtering, order-level classification could be obtained for 93% of all sequences (Table 2). Based on nonnormalized order-level sequence counts, we observed clear patterns of separation between species and ecotypes in a multivariate nonmetric multidimensional scaling (NMDS) plot (Fig. 1b), with NEAC and northern silvery pout forming distinct clusters, whereas the NCC populations encompass the Norway pout and poor cod populations. *Vibrionales* was the most abundant order detected in the intestinal microbiomes of NCC specimens at both coastal locations (mean relative abundance [MRA]: 76%) as well as Norway pout (MRA: 79%) and poor cod (MRA: 44%) (Table 3, Fig. 2a), with the remainder of each gut community consisting of a mix of orders with low relative abundance. The intestinal microbiomes of the NEAC and northern silvery pout specimens had a significantly more diverse community composition (Fig. 3, Fig. 2a). NEAC was dominated by *Bacteroidales* (MRA: 21%), *Vibrionales* (MRA: 17%), *Clostridiales* (MRA: 12%), and *Brevinematales* (MRA: 7%), while northern silvery pout had a high relative abundance of orders *Brachyspirales* (MRA: 16%) and *Clostridiales* (MRA: 14%). Distinct from the gut communities of the other fish populations, northern silvery pout had a low abundance of *Vibrionales*. Finally, the amount of sequences in the “Others” category, as well as sequences classified above order level (mean of all samples was 7.8%), varied slightly between the fish species (Table S8). A species-level classification was obtained for 66% of all sequences. Overall, species of the genus *Photobacterium* comprised on average 40.6% of the classified sequences, ranging from 0.2% in northern silvery pout to 74.3% in Norway pout (Fig. 2b). In particular, *P. kishitanii* and *P. iliopiscarium* represented on average 43% and 36% of all *Photobacterium* species, respectively, although the ratio differed in the different fish species (e.g., 49% versus 41% in NCC; 16% versus 56% in NEAC; and 55% versus 12% in Norway pout).

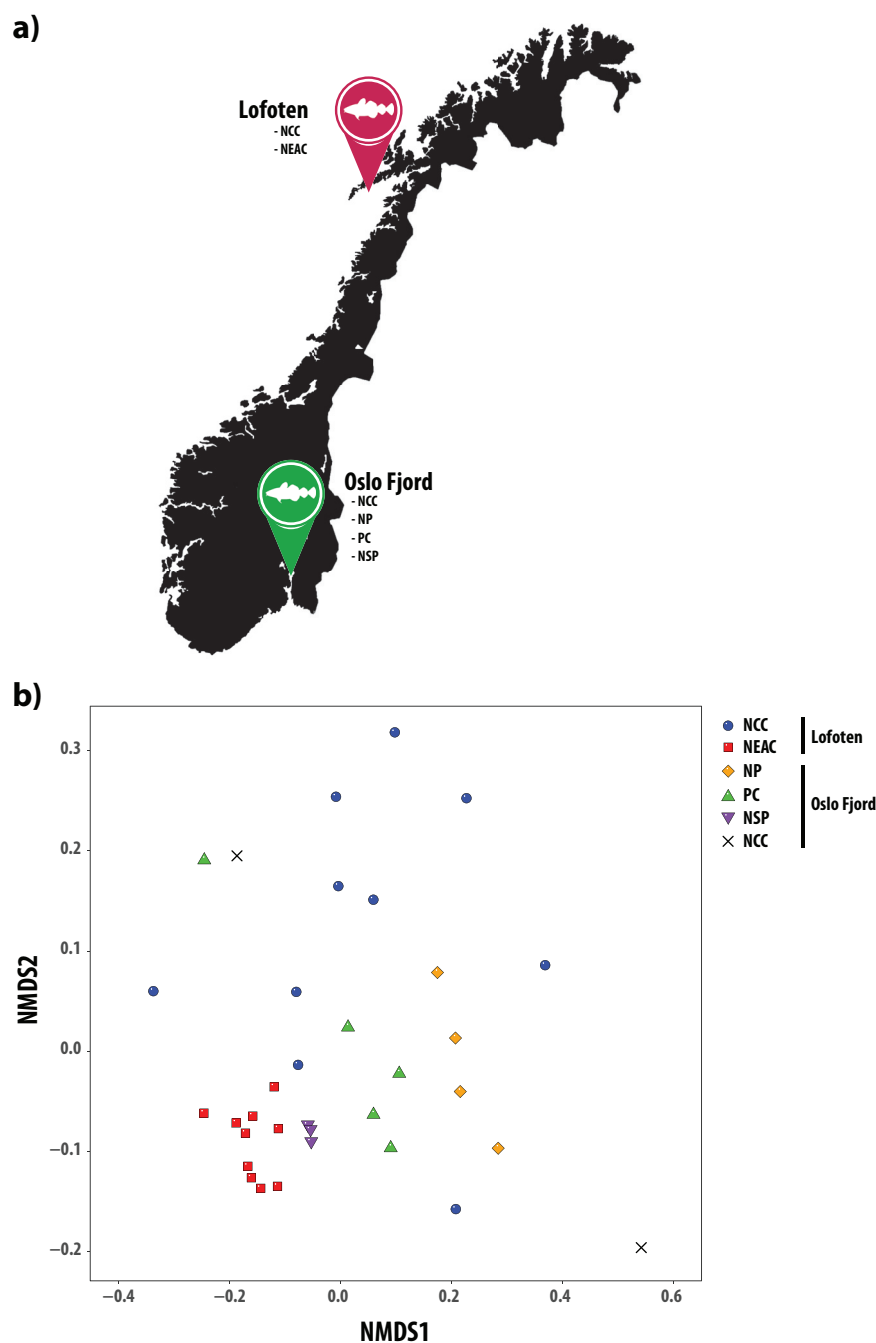
**TABLE 2** Overview of individual metagenomic sequence data from gadoid intestines<sup>a</sup>

Sample	Raw reads	After quality trimming/ filtering (%)	Host DNA (%)	Bacterial DNA (%)	Final reads
NCC_01	10,883,740	85.9	87.3	12.7	1,187,649
NCC_02	11,140,950	87.9	62.2	37.8	3,699,538
NCC_03	9,891,322	90.2	41.2	58.8	5,249,515
NCC_04	10,587,865	86.9	85.2	14.8	1,364,663
NCC_05	8,423,091	89.1	57.7	42.3	3,171,737
NCC_06	10,879,319	89.6	30.5	69.5	6,772,948
NCC_07	10,082,237	91.8	31.3	68.7	6,361,506
NCC_08	9,114,703	87.3	80.5	19.5	1,549,210
NCC_09	11,105,189	89.1	62.2	37.8	3,733,846
NCC_10	11,140,743	84.7	86.0	14.0	1,320,875
NEAC_01	13,120,072	89.7	53.6	46.4	5,463,098
NEAC_02	12,119,926	89.6	56.8	43.2	4,687,565
NEAC_03	11,981,093	89.2	54.4	45.6	4,869,722
NEAC_04	12,618,529	91.1	33.5	66.5	7,646,256
NEAC_05	12,154,047	87.6	74.2	25.8	2,747,042
NEAC_06	13,883,762	88.4	59.5	40.5	4,971,507
NEAC_07	12,149,049	89.1	56.9	43.1	4,666,533
NEAC_08	11,861,852	88.7	64.0	36.0	3,787,155
NEAC_09	11,131,413	85.7	75.1	24.9	2,378,591
NEAC_10	15,483,018	83.7	82.9	17.1	2,221,214
OO_cod_01	8,047,125	83.7	85.9	14.1	949,039
IO_cod_01	9,716,392	90.2	43.7	56.3	4,937,260
NP_01	15,621,734	84.7	51.6	48.4	6,400,120
NP_02	16,548,297	90.2	18.1	81.9	12,224,903
NP_03	16,608,312	78.3	72.6	27.4	3,568,206
NP_04	13,459,929	83.0	61.5	38.5	4,300,178
PC_01	10,743,586	87.7	30.5	69.5	6,550,868
PC_02	18,982,339	81.0	29.6	70.4	10,833,201
PC_03	9,420,298	84.1	54.9	45.1	3,568,861
PC_04	9,623,591	87.7	30.0	70.0	5,908,283
PC_05	19,630,680	77.7	55.0	45.0	6,861,474
NSP_01	14,283,994	73.2	67.5	32.5	3,396,962
NSP_02	14,527,770	76.3	63.8	36.2	4,008,926
NSP_03	15,261,446	80.9	66.6	33.4	4,123,131
<b>Total:</b>	<b>422,227,413</b>				<b>155,481,582</b>
<b>Mean:</b>	<b>12,418,453</b>	<b>86.0</b>	<b>57.8</b>	<b>42.2</b>	

<sup>a</sup>PhiX- and human-derived DNA sequences represented a negligible proportion and were excluded from the table. On average, 42.2% of the quality filtered reads per sample were used for microbiome analysis. (For further details, see Table S7.)

The NCC Lofoten intestinal microbiome was dominated by *P. iliopiscarium* (MRA: 21%) and *P. kishitanii* (MRA: 20%), followed by different species of *Aliivibrio* (*A. wodanis*, *A. logei*, and *A. fischeri*) (MRA: 13%) (Fig. 2b). Similarly, the bacterial gut community of Norway pout was also dominated by *Photobacterium* species, in particular *P. kishitanii* (MRA: 17%). The intestinal microbiome of poor cod was dominated by *Photobacterium* species (MRA: 18%), followed by different *Vibrio* spp. (MRA: 8%). The gut bacterial community of NEAC was more diverse, with high relative abundance of a *Brevinema* sp. (MRA: 31%) and different species in the genera *Photobacterium* (MRA: 34%), *Clostridium* (MRA: 12%), and *Aliivibrio* (MRA: 9%). The high abundance of *Bacteroidales* observed at the order level (Fig. 2a) was not reflected at the species level, as this order represented a high number of *Bacteroidales* species with low abundance. Consequently, no *Bacteroidales* species were among the 15 most abundant species in the NEAC intestinal microbiome (Fig. 2b). The NEAC samples also contained a *Mucispirillum* sp. (MRA: 4%) and two *Brachyspira* spp. (MRA: 2%). In northern silvery pout, the gut microbiome was quite evenly distributed between the *Brevinema* sp., the *Mucispirillum* sp., *Brachyspira pilosicoli*, *Brachyspira* sp. CAG:700, and a group of different *Clostridium* species in two of three samples. The third sample contained the same species but had an even higher relative abundance of the *Brevinema* sp. (64%) (Fig. 2b).

**Variation in bacterial community composition among species and ecotypes.** Significant differences in within-sample diversity (alpha diversity) at the order level



**FIG 1** The intestinal microbiomes obtained from a range of gadoid species and ecotypes. (A) Map of sampling locations in Norway, Europe. Northeast Arctic cod (NEAC), and Norwegian coastal cod (NCC) were obtained from Lofoten. NCC (two individuals), poor cod (PC), Norway pout (NP), and northern silvery pout (NSP) were obtained from the Oslo Fjord. (Map data copyright 2020 Google.) (B) Nonmetric multidimensional scaling (NMDS) plot of nonnormalized, order-level sequence counts from the intestinal microbiomes of all samples. Each point represents an individual sample, and the species or ecotypes are indicated by different shapes and colors. The stress value of the NMDS plot is 0.14.

were observed among all species and within-species ecotypes, except between NCC and Norway pout (Table 4, Table S5). None of the other covariates had a significant effect on alpha diversity. Similar to the results from the within-sample diversity, significant differences in community structure (beta diversity) were observed among the gadoid species at order, genus, and species levels (Table 5, Table S6). At the order level, the NEAC intestinal community had a different structure than what was observed

**TABLE 3** Mean relative abundance (%) of the ten most abundant bacterial orders in the intestinal microbiomes of gadoid species and ecotypes<sup>a</sup>

NEAC <sup>b</sup>	NCC <sup>b</sup>	PC	NP	NSP
<i>Bacteroidales</i> (21.39)	<i>Vibrionales</i> (75.69)	<i>Vibrionales</i> (44.13)	<i>Vibrionales</i> (78.57)	<i>Brachyspirales</i> (15.88)
<i>Vibrionales</i> (16.83)	<i>Alteromonadales</i> (4.34)	<i>Clostridiales</i> (11.16)	<i>Clostridiales</i> (3.36)	<i>Clostridiales</i> (14.14)
<i>Clostridiales</i> (11.66)	<i>Clostridiales</i> (3.47)	<i>Mycoplasmatales</i> (8.92)	<i>Alteromonadales</i> (2.11)	<i>Brevinematales</i> (7.11)
<i>Brevinematales</i> (7.43)	<i>Fusobacteriales</i> (3.46)	<i>Alteromonadales</i> (5.15)	<i>Enterobacteriales</i> (2.04)	<i>Deferribacteriales</i> (4.81)
<i>Bacillales</i> (2.64)	<i>Oceanospirillales</i> (1.56)	<i>Enterobacteriales</i> (2.97)	<i>Bacteroidales</i> (1.02)	<i>Bacillales</i> (4.44)
<i>Alteromonadales</i> (2.61)	<i>Enterobacteriales</i> (1.19)	<i>Bacteroidales</i> (2.09)	<i>Mycoplasmatales</i> (0.87)	<i>Fusobacteriales</i> (1.94)
<i>Flavobacteriales</i> (2.17)	<i>Bacteroidales</i> (0.92)	<i>Bacillales</i> (1.81)	<i>Oceanospirillales</i> (0.64)	<i>Desulfovibrionales</i> (1.70)
<i>Fusobacteriales</i> (1.62)	<i>Bacillales</i> (0.60)	<i>Oceanospirillales</i> (1.18)	<i>Burkholderiales</i> (0.47)	<i>Lactobacillales</i> (1.57)
<i>Brachyspirales</i> (1.23)	<i>Pseudomonadales</i> (0.33)	<i>Lactobacillales</i> (0.95)	<i>Bacillales</i> (0.43)	<i>Rhizobiales</i> (1.40)
<i>Deferribacteriales</i> (1.23)	<i>Flavobacteriales</i> (0.27)	<i>Burkholderiales</i> (0.84)	<i>Pseudomonadales</i> (0.43)	<i>Spirochaetales</i> (1.28)

<sup>a</sup>NEAC, Northeast Arctic cod; NCC, Norwegian coastal cod; PC, poor cod; NP, Norway pout; NSP, northern silvery pout.

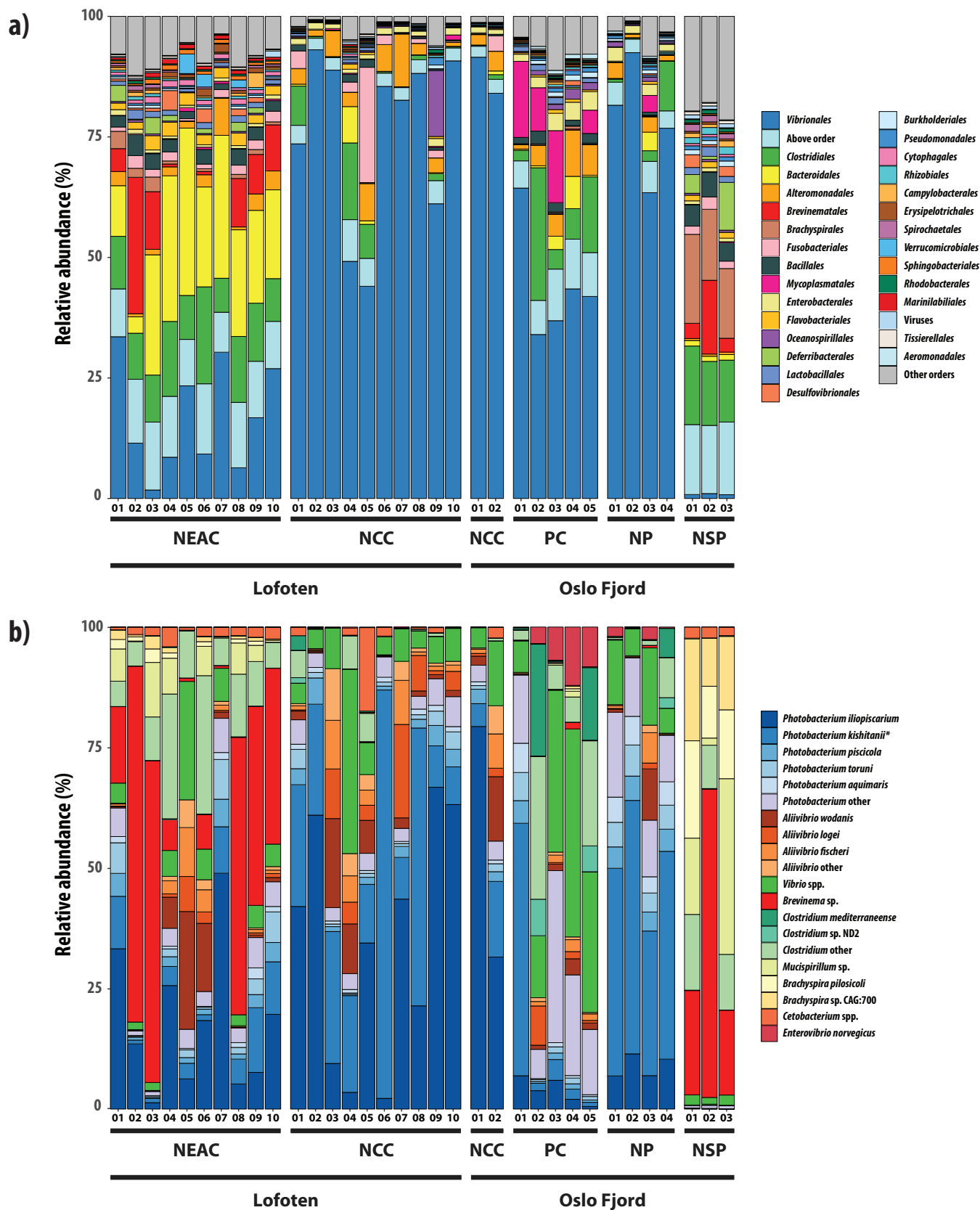
<sup>b</sup>These two ecotypes belong to the same species, *Gadus morhua*.

in all the other gadoids (at a significance level of 0.05). The NCC intestinal microbiome was also different from that of both poor cod and northern silvery pout. In agreement with results of within-sample (alpha) diversity, no differences in community structure were observed between the microbiomes of NCC and Norway pout. Finally, no differences were observed between the gut microbiome of poor cod versus Norway pout, poor cod versus northern silvery pout, or Norway pout versus northern silvery pout ( $P = 0.074$  for all). Beta diversity analysis also demonstrated that community differences at the genus and species levels were similar to those observed at the order level (Table S6).

Differences in the intestinal community composition between these gadoids are predominantly explained by changes in the relative abundance of a limited number of orders. For example, different proportions of *Vibrionales* contribute 29% to the (Bray-Curtis) dissimilarity between the NCC and NEAC ( $P = 0.001$ ), followed by differences in the relative abundance of *Bacteroidales*, explaining 10% of the dissimilarity ( $P = 0.001$ ) (Table S9). Together, 80% of the observed dissimilarity between NCC and NEAC is explained by differences in their relative abundance of the top six orders. Similarly, 60% of the dissimilarity between NCC and northern silvery pout is driven by *Vibrionales*, *Brachyspirales*, and *Clostridiales*.

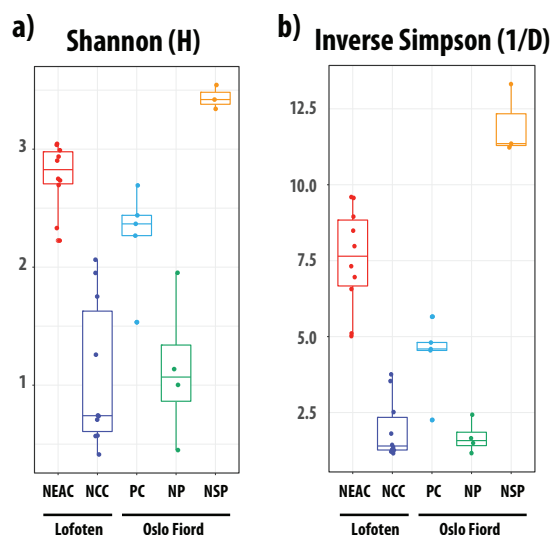
**Bacterial within-species variation of SNV heterogeneity.** We investigated bacterial within-species variation of *P. iliopiscarium* and *P. kishitanii*—with sufficient read coverage across all samples—among the different gadoids by mapping sequencing reads to their respective reference genomes (GCF\_000949935.1 and GCF\_000613045.2). In the samples used for single nucleotide variant (SNV) analysis, the mean percentage of the reference genomes with minimum 20-fold coverage (coverage breadth) after mapping were 63% for *P. iliopiscarium* and 19% for *P. kishitanii*. Hence, the variation analysis of the two species was based on different proportions of the reference genomes. The two reference genomes varied widely in the number of SNVs observed in all samples, from 84,866 in *P. iliopiscarium* to 1,229 in *P. kishitanii* (Fig. 4a). The density of variable sites within each individual sample showed various levels of heterogeneity in the bacterial populations (Fig. 4b). This heterogeneity was particularly clear in *P. kishitanii*, with site density varying from 0.5 to 45.4 variant positions per kbp per individual specimen. Further, the heat map showed gadoid-specific SNV patterns (Fig. 4c), in particular for *P. iliopiscarium*, where Norway pout contained a distinct pattern compared to the other gadoids, indicating the presence of specific *P. iliopiscarium* strain(s). Statistical analyses of SNV variation revealed that NEAC had a significantly different SNV pattern from Norway pout (Chi-square,  $P = 0.017$ ) and poor cod ( $P = 0.028$ ) for *P. kishitanii*, and from NCC ( $P = 0.033$ ) and Norway pout ( $P = 0.000$ ) for *P. iliopiscarium* (Fig. 4d, Table S10). NCC had a significantly different SNV pattern from Norway pout ( $P = 0.003$ ) for *P. iliopiscarium*. (Fig. 4d, Table S10). The relative abundance of *P. kishitanii* and *P. iliopiscarium* varied greatly among the fish specimens used in the variation analysis (Fig. 4e).





**FIG 2** Taxonomic composition of the fish intestinal microbiomes. (a) Relative abundance of metagenomic shotgun sequences classified at the order level (93%). Colors represent the 28 orders with highest relative abundance, sequences assigned to other orders or viruses, and sequences classified above order level.

(Continued on next page)



**FIG 3** Within-sample microbial diversity in the gadoid species and ecotypes. Boxplots of Shannon (a) and inverse Simpson (b) diversity in the fish species/ecotypes. Each individual is represented by a point, and the individuals are grouped and colored by species and ecotype. The middle band represents the median, while the upper and lower bands show the 75<sup>th</sup> and 25<sup>th</sup> percentiles. The boxplots include the minimum and maximum alpha diversity values.

## DISCUSSION

Using metagenomic shotgun sequencing, we show the composition of the intestinal microbiomes of two Atlantic cod ecotypes (NEAC, NCC) to be at least as divergent as those found between the different codfish species investigated here. Our findings have several implications for our understanding of the composition of the intestinal microbiome in wild fish populations.

Although species-specific selection has been proposed as a factor driving the composition of the intestinal community in fish in a variety of settings (13, 14, 16–19, 29, 33), our results show that this may not be the most important driver among gadoid species in wild populations. First, we observed highly significant differences in the intestinal microbiomes at order, species, and within-species bacterial levels between the NEAC and NCC ecotypes. Despite showing different migratory behaviors, these ecotypes co-occur during seasonal spawning in northern Norway (Lofoten), from where most of the samples were collected (43–45). Second, we observed no significant bacterial order- or species-level differences in the intestinal microbiome between different gadoid species, Atlantic cod (ecotype NCC), and Norway pout, which were sampled from different geographical locations (Lofoten and Oslo Fjord). We did not observe differentiation between the NCC sampled from Lofoten and the Oslo Fjord (although statistical certainty was low), which reflects an earlier observed lack of geographical structure for this ecotype (30). The similarity of the microbial compositions of the NCC and Norway pout is striking, as these are distinctly different genetic lineages with an evolutionary separation of at least 20 million years (24). These results suggest that NCC and Norway pout occupy an environmental niche that allows bacterial members with a broad geographical distribution to colonize their intestinal communities. Overall, the observation of a significant differentiation between microbiomes from ecotypes of the same species and a lack of differentiation between

## FIG 2 Legend (Continued)

Numbers along the x axis indicate the individual samples of the different species/ecotypes. (b) Relative abundance of metagenomic shotgun sequences classified at the species level (66%). The plot includes the most highly abundant species and other members of their parent bacterial genera ("other" categories) in the different fish species/ecotypes. Numbers along the x axis indicate the individual samples of the different species/ecotypes. The asterisk denotes the *P. kishitanii* strain that was reclassified from *P. phosphoreum*.



**TABLE 4** Effects of covariates on the intestinal microbial diversity (alpha diversity) of gadoid species and ecotypes<sup>a,b</sup>

Species or ecotype	Shannon		Simpson		Inv. Simpson	
	Estimate	P value	Estimate	P value	Estimate	P value
Intercept	1.08	<b>0.0000</b>	0.38	<b>0.0000</b>	1.93	<b>0.0001</b>
NEAC	1.69	<b>0.0000</b>	0.48	<b>0.0000</b>	5.62	<b>0.0000</b>
NP	0.06	0.8367	-0.01	0.8784	-0.25	0.7476
PC	1.18	<b>0.0001</b>	0.37	<b>0.0002</b>	2.44	<b>0.0017</b>
NSP	2.36	<b>0.0000</b>	0.53	<b>0.0000</b>	10.03	<b>0.0000</b>

<sup>a</sup>Results from the optimal linear regression models used in testing for significant effects of covariates on within-sample (alpha) diversity based on nonnormalized, order-level sequence counts. Population (species/ecotype) is the only covariate with a significant effect, and estimates are given relative to NCC. Significant effects ( $P < 0.05$ ) are indicated in bold.

<sup>b</sup>NEAC, Northeast Arctic cod; PC, poor cod; NP, Norway pout; NSP, northern silvery pout.

microbiomes from two distinct species suggests that the intestinal microbiome in these gadoid species and ecotypes is not driven by species-specific selection alone.

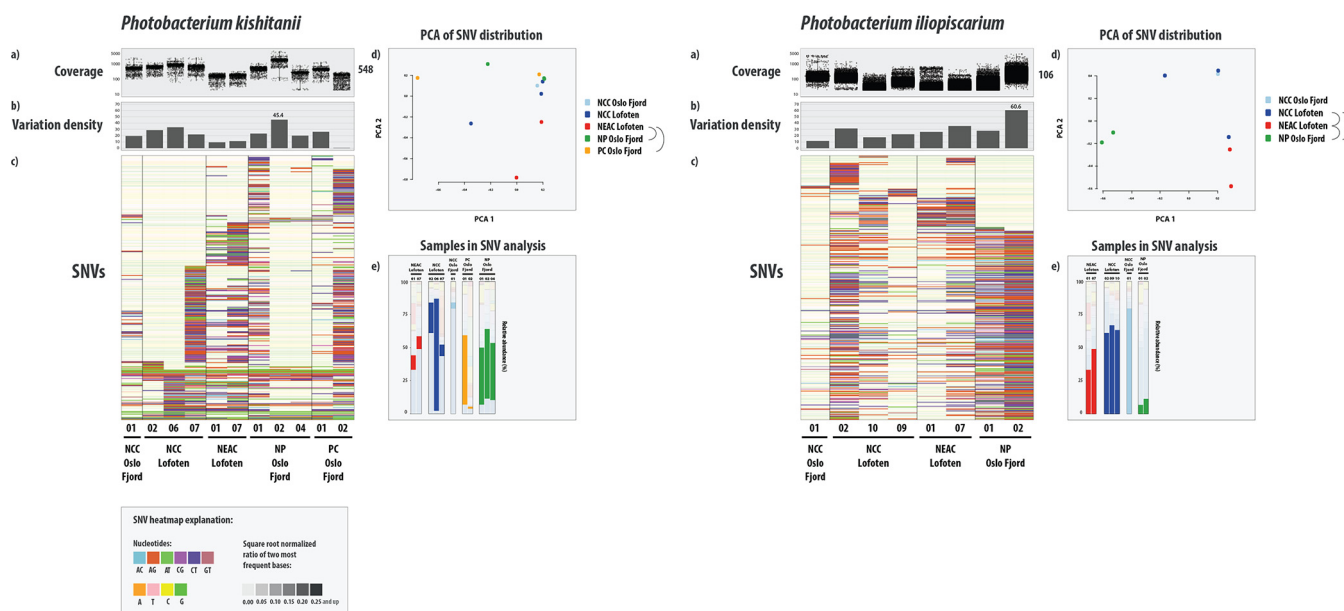
There are several factors that may underlie the compositional differences in the NCC and NEAC intestinal microbiomes. First, for more than 10 months during the year, the two populations encounter different habitats, as the NEAC ecotype is distributed in the pelagic waters of the Barents Sea while NCC remains more stationary in coastal waters (49). Although several 16S rRNA-based studies have reported limited effects of geographic location on the composition and diversity of the fish intestinal microbiome (32, 50), the Barents Sea has significantly lower temperatures (51) than Norwegian coastal waters (52). Temperature has been shown to have a significant impact on the intestinal microbiome in several studies, e.g., Senegalese sole (*Solea senegalesis*), Tasmanian Atlantic Salmon (*Salmo salar*), and mummichog (*Fundulus heteroclitus*) (5, 53, 54), but not in all cases (e.g., Atlantic salmon) (55). Second, the ecotypes were sampled during different seasons: NCC Lofoten during summer (August) and NEAC during winter/early spring (March). Nonetheless, a lack of difference between NCC Lofoten (August) and NCC Oslo Fjord (May) suggests that seasonality is unlikely to fully explain the observed differences between NEAC and NCC. Third, the ecotypes show different feeding behaviors; while the NEAC may perform vertical movements down to 500 m during foraging and spawning migrations from the Barents Sea (42, 56, 57), NCC mainly occupy shallow and warmer coastal and fjord waters (58). These behaviors expose the two ecotypes to different sources of food, with NEAC predominantly eating capelin and herring (48) and NCC living on a more diverse diet, including crustaceans, fish, and even seaweeds (34, 39). Diet has been shown to influence the composition of the intestinal

**TABLE 5** PERMANOVA analysis of intestinal microbial diversity from gadoid species and ecotypes<sup>a</sup>

Populations <sup>b</sup>	R <sup>2</sup>	P value	Adjusted P value
NEAC vs. NCC	0.71	<b>0.0001</b>	<b>0.0005</b>
NEAC vs. PC	0.53	<b>0.0002</b>	<b>0.0018</b>
NEAC vs. NP	0.70	<b>0.0010</b>	<b>0.0076</b>
NEAC vs. NSP	0.50	<b>0.0041</b>	<b>0.0207</b>
NCC vs. PC	0.42	<b>0.0026</b>	<b>0.0182</b>
NCC vs. NP	0.04	0.7138	0.7138
NCC vs. NSP	0.80	<b>0.0035</b>	<b>0.0207</b>
PC vs. NP	0.53	<b>0.0233</b>	0.0740
PC vs. NSP	0.80	<b>0.0185</b>	0.0740
NP vs. NSP	0.94	<b>0.0286</b>	0.0740

<sup>a</sup>R<sup>2</sup> values, P values, and adjusted P values for pairwise comparisons of community composition (beta diversity) between the different species or ecotypes using PERMANOVA. The tests are based on Bray-Curtis dissimilarity calculated from order-level, normalized sequence counts. P values are adjusted for multiple testing by the Holm method. Significant differences ( $P < 0.05$ ) are indicated in bold. (Genus- and species-level results can be found in Table S6.)

<sup>b</sup>NEAC, Northeast Arctic cod; NCC, Norwegian coastal cod; PC, poor cod; NP, Norway pout; NSP, northern silvery pout.



**FIG 4** SNV variation analysis of the two most abundant bacterial genomes in the microbiomes of gadoid species. For each genome, the figure displays read coverage per single nucleotide variant (SNV) position in each sample from the different species/ecotypes (a) (mean coverage on right side of plot) and variation density (b) (number of variable positions per 1,000 bp reported in each individual sample, independent of coverage in the other samples) per sample (maximum value indicated). The y axes of the coverage and variation density plots are scaled across the genomes. (c) Heat map of a randomly chosen subset of 400 SNVs. In the heat map, each row represents a unique variable nucleotide position, where the color of each tile represents the two most frequent competing nucleotides in that position. The shade of each tile represents the square root-normalized ratio of the most frequent two bases at that position (i.e., the more variation in a nucleotide position, the darker the tile is). See legend at the bottom of the figure. (d) Principal-component analysis (PCA) plot of the SNV distribution (within-species variation) among the different samples. Each sample is represented by a dot, and colored according to species or ecotype membership. Half-circles to the right of the legend indicate species or ecotypes with significantly different within-species variation (i.e., different strains). (e) Relative abundance of the different samples used in variation analysis. The bars are colored according to the SNV plot in panel d.

microbiome in several fish species (9, 10, 13, 53, 59, 60). Finally, the Barents Sea has a high microbial biodiversity compared to coastal areas (61). The specific bacterial load in the surrounding waters also influences the intestinal microbiome composition in fish, including Atlantic cod (3, 4). Nonetheless, because these different environmental and behavioral factors are correlated, it is unclear which of these parameters contributes the most to the observed differences in the intestinal microbiome composition between these ecotypes.

Comparing two spatially separated coastal Atlantic cod populations, metagenomic shotgun data revealed no strain-level differentiation (30). In this study, we found specific SNV variants among the most abundant bacterial species that were associated with single species or specific Atlantic cod ecotypes. This indicates that NEAC harbor different strains of *P. iliopiscarium* than those identified in the NCC ecotype and the other gadoid species. Our current study encompasses a significantly greater geographical area and broader range of taxonomical samples than the earlier coastal comparison (30–32), and is indicative of strain-level variation at larger comparative scales. In line with Riiser et al. (30), this study shows that such strain-level differences cannot be detected using 16S rRNA techniques alone, and that metagenomic shotgun sequencing is currently the most accurate approach to detect strain-level spatial variation in the marine environment.

Most striking among the comparisons of gadoid species were the microbiome differences observed in NEAC, northern silvery pout, and poor cod compared to NCC and Norway pout. Several bacterial species that drive this differentiation are of particular interest. First, two bacterial species, *Mucispirillum* sp. and *Brevinema* sp., are almost exclusively detected in the intestinal microbiomes of NEAC and northern silvery pout. Nonetheless, these genera are represented by a single species in the *RefSeq* database (62) (accessed 10 January 2019) and hence little is known. *Brevinema andersonii* (order *Brevinematales*) was originally identified in short-tailed shrews (*Blarina brevicauda*) and

white-footed mice (*Peromyscus leucopus*) and was found to be unable to grow below 25°C (63). *Brevinema* spp. have previously been identified in Atlantic cod (32) and in Atlantic salmon (64). *Mucispirillum schaedleri* (order *Deferribacterales*) is a mucosa-associated member of the intestinal microbiome in terrestrial animals such as pigs, goats, and rodents, where it is thought to be involved in mucus production through expression of lectins, important components in the innate immune response (65, 66). Nevertheless, the distant relationship between Atlantic cod and these terrestrial hosts and the availability of only single reference genomes for *Mucispirillum* and *Brevinema* strongly suggest that the representatives found here are related but novel species with different intestinal ecologies and physiologies. Second, both NEAC and northern silvery pout contain significant fractions of *Brachyspira* spp., previously identified as dominant members in the gut of the carnivorous marine fish species mahi mahi (*Coryphaena hippurus*) (12, 67). *Brachyspira* spp. are known as intestinal pathogens in pigs and humans (68, 69), although recent studies show that *Brachyspira* spp. are more widespread in the wildlife community than previously thought, including in freshwater (70). The ecology of *Brachyspira* in the marine environment is unclear, although an association with the carnivorous diet of mahi mahi and NEAC may suggest that the diet of northern silvery pout also has a considerable carnivorous component. Third, poor cod is the only species with considerable abundance of *Enterovibrio norvegicus* (Table S11). This bacterium within the *Vibrionaceae* family was isolated from the intestines of cultured turbot (*Scophthalmus maximus*) larvae in Norway and classified as a novel species phenotypically similar to the *Vibrio* genus (71). Interestingly, poor cod also host the highest abundance of *Vibrio* spp. among the fish species in this study (Table S11). Other *Enterovibrio* species have been found in association with diseased corals (72) and internal organs of cultured fish species in the Mediterranean Sea (73–75). However, little is known about the function of this relatively novel genus in fish intestines.

Given the observations of species-specific selection for a similar microbiome in various teleosts and range of habitats (13, 14, 16–19, 29, 33), the diverse microbiomes within and among gadoid species may suggest that their intestinal communities could be more easily modulated by external factors. At this stage, limited sampling across various fish taxa and the lack of comparative approaches leave reasons for such diverse communities speculative. Nonetheless, it is interesting to note that all gadoids have an unusual adaptive immune system in which there is loss of MHC II, CD4, and invariant chain (Ii) and a range of innate (TLR) and MHC I immune-gene expansions (23, 24). There are significant correlations between immune genes and the vertebrate microbiome (76, 77), and it has been hypothesized that adaptive immunity has evolved to help maintain a complex community of beneficial commensal bacteria (78). Indeed, studies of wild-type zebrafish and knockout zebrafish without a functional adaptive immune system suggest that adaptive immunity increases the strength of host filtering of potential fish-associated microbes (22). The unusual adaptive immune system of gadoids may therefore affect the strength of coevolutionary associations within their microbiomes.

In conclusion, based on metagenomic shotgun sequencing, we here characterize the intra- and interspecific community compositions among two ecotypes of Atlantic cod and three related fish species in the Gadidae family. Several of these fish species harbor unique, and possibly novel, bacterial species. We identify a complex pattern of diversity with significant differences between the Atlantic cod ecotypes and variable interspecific patterns of variation. Although most species and ecotypes yield different communities, those found in coastal cod (NCC) and Norway pout are not significantly diverged, indicating that ecological niche plays an important role in determining the intestinal microbiomes in these gadoid species.

## MATERIALS AND METHODS

**Sample collection.** Northeast Atlantic cod (*Gadus morhua*) (NEAC, 10 individuals) were collected in Lofoten (N68.0619167, E13.5921667) in March 2014 and Norwegian coastal cod (*Gadus morhua*) (NCC, 10 individuals) at the same location in August 2014 (Fig. 1a; see Table S1 in the supplemental material). NCC (2 individuals) were also collected in the Oslo Fjord (N58.9125100, E9.9202624 and N59.8150006,

E10.5544914). Norway pout (*Trisopterus esmarkii*, 4 individuals), poor cod (*Trisopterus minutus*, 5 individuals), and northern silvery pout (*Gadiculus thori*, 3 individuals) were collected in the inner Oslo Fjord in May 2015 (Table S1). All fish specimens were collected from wild populations. A 3-cm-long part of the hindgut (immediately above the short, wider rectal chamber) was aseptically removed postmortem by scalpel and stored in 70% ethanol. The samples were frozen (−20°C) for long-term storage. Relevant metadata such as length, weight, sex, and maturity were registered. As we strive to reduce the impact of our sampling needs on populations and individuals, samples were therefore obtained as a by-product of conventional business practices. Specimens were caught by commercial vessels, euthanized by local fishermen, and intended for human consumption. Samples were taken postmortem and no scientific experiments have been performed on live animals. This sampling follows the guidelines set by the “Norwegian consensus platform for replacement, reduction and refinement of animal experiments” (79) and does not fall under any specific legislation in Norway, requiring no formal ethics approval.

**Sample preparation and DNA extraction.** Intestinal samples were split open lengthwise before the combined gut content and mucosa were gently removed using a sterile disposable spatula. Each individual sample was washed in 500  $\mu$ l 100% ethanol (EtOH) and centrifuged before the ethanol was allowed to evaporate, after which dry weight was measured before proceeding to DNA extraction. DNA was extracted from between <10 and 300 mg dry weight of gut content using the MoBio Powersoil HTP 96 Soil DNA isolation kit (Qiagen, Valencia, CA, USA) according to the DNA extraction protocol (v. 4.13) utilized by the Earth Microbiome Project (80). DNA was eluted in 100  $\mu$ l elution buffer and stored at −20°C. Due to high methodological consistency between biological replicates in previous experiments, only one sample was collected per fish (32).

**Sequence data generation and filtering.** Quality and quantity of the DNA was measured using a Qubit fluorometer (Life Technologies, Carlsbad, CA, USA) and normalized by dilution. DNA libraries were prepared using the Kapa HyperPlus kit (Roche Sequencing, Pleasanton, CA, USA) and paired-end sequenced (2  $\times$  125 base pairs) on an Illumina HiSeq2500 using the HiSeq SBS V4 chemistry with dual indexing in two independent sequencing runs. Read qualities were assessed using FastQC (81) before adapter removal, singleton read identification, deduplication, and further read quality trimming was performed using Trimmomatic (ver. 0.36) (82) and PRINSEQ-lite (ver. 0.20.4) (83) (Table S2). PhiX, host, and human sequences were removed by mapping reads to the phiX reference genome (GenBank: J02482.1), the Atlantic cod genome assembly (gadMor 2) (this applied to all the fish species) (84), and a masked version of the human genome (HG19) (85) using BWA (ver. 0.7.13) (86) or BMAP (ver. 37.53) (87) (JGI) with default parameters and discarding matching sequences using seqtk (ver. 2012.11) (88). All sequence data have been deposited in the European Nucleotide Archive (ENA) under study accession number PRJEB31095.

**Taxonomic profiling.** Taxonomic classification of quality trimmed and filtered metagenomic paired-end reads was performed using Kaiju (ver. 1.5.0) (89) (“greedy” heuristic approach, -e 5), with the NCBI nr database (release 84) (including proteins from fungal and microbial eukaryotes) as reference (62). Counts of sequences successfully assigned to orders and species were imported into RStudio (ver. 1.1.383) (90), based on R (ver. 3.4.2) (91), for further processing. Filtering of the most abundant bacterial orders for visualization was based on a minimum relative abundance threshold of 1% of the total number of sequences per library (the threshold ranged from 5,933 to 95,146 depending on the sample size). Similarly, filtering of the most abundant bacterial species was based on a minimum relative abundance threshold of 2% of the total number of sequences per library (the threshold ranged from 6,548 to 190,294 depending on the sample size). Any taxon not exceeding this threshold in at least one (order-level) or two (species-level) samples was removed. All filtering was based on the R package genefilter (ver. 1.62.0) (92). Final results were visualized using the R package ggplot (ver. 2.2.1) (93). Note that, based on a recent reclassification (94), we refer to the reference strain *Photobacterium phosphoreum* ANT-2200 (accession number GCF\_000613045.2) as *Photobacterium kishitanii* (Table S3).

**Sequence variation analysis.** In order to assess the heterogeneity of the most abundant bacteria in the fish species, we analyzed the sequence variation in the two genomes with the highest mean relative abundance over all fish species and ecotypes, *Photobacterium kishitanii* and *Photobacterium iliopiscarium*. Paired-end reads from each individual fish were mapped to the reference genomes (Table S3) using the Snakemake workflow (95) of anvi'o (ver. 5.1) (96) with default parameters in the “all-against-all” modus (with anvi-profile -min-coverage-for-variability 20). Samples of low coverage, restricting detection of SNVs in anvi'o, were excluded from the variation analysis. For each individual sample, variable sites were identified, and the mean number of these per 1,000 bp was calculated (variation density). A variable site required a minimum coverage of 20 $\times$ . Next, variable sites with a minimum of 20 $\times$  coverage in all samples were defined as single nucleotide variants (SNVs, anvi-gen-variability-profile -min-occurrence 1 -min-coverage-in-each-sample 20). Coverage, variation density, and SNV profiles were plotted in RStudio following the R script provided by anvi'o (97). The anvi'o SNV output was converted to .vcf format using a custom-developed script ([https://github.com/srinidhi202/AnvioSNV\\_to\\_vcf](https://github.com/srinidhi202/AnvioSNV_to_vcf)), and the resulting .vcf files were used in a principal-component analysis (PCA) to test for population differences as implemented in smartpca (ver. 6.1.4) (EIGENSOFT) (98).

**Statistical analysis.** Although included in data visualization, the Oslo Fjord NCC samples were excluded from statistical analysis due to low sample size ( $n = 2$ ). Within-sample diversity (alpha diversity) was calculated using the diversity function in the R package vegan (ver. 2.4-1) (99) based on Shannon, Simpson, and inverse Simpson indices calculated from nonnormalized order-level read counts (Table S4). Differences in alpha diversity were studied using linear regression. The “optimal model” (the model that best describes the individual diversity) was identified through a “top-down” strategy, including all covariates (Table S5) except

weight, which highly correlated with length ( $r = 0.95$ ), and selected through  $t$  tests. Model assumptions were verified through plotting of residuals. Differences in bacterial community structure (beta diversity) between the fish species or ecotypes were visualized using nonmetric multidimensional scaling (NMDS) plots based on the Bray-Curtis dissimilarity calculated from order-level sequence counts. Next, pairwise differences in beta diversity between the fish species or ecotypes were tested using permutational multivariate analysis of variance (PERMANOVA) in the R package pairwise.adonis (ver. 0.1) (100), a wrapper for the adonis functions in vegan (ver. 2.4-1), based on Bray-Curtis dissimilarity calculated from order-, genus- and species-level sequence counts. The package pairwise.adonis was run with 20,000 permutations, and  $P$  values were adjusted for multiple testing using the Holm method (101). Adjusted  $P$  values of  $<0.05$  indicate statistical significance. PERMANOVA assumes the multivariate dispersion in the compared groups to be homogeneous; this was verified ( $P > 0.05$ ) using the betadisper function (vegan) (Table S6). Similarity percentage (SIMPER) procedure implemented in vegan was used to quantify the contribution of individual orders to the overall Bray-Curtis dissimilarity between the species/ecotypes. All beta diversity analyses were based on sequence counts normalized using a common scaling procedure, following McMurdie and Holmes (102). This involves multiplying the sequence count of every unit (e.g., order) in a given library with a factor corresponding to the ratio of the smallest library size in the data set to the library size of the sample in question. Normalization using this procedure effectively results in library scaling by averaging an infinite number of repeated subsamplings. We used chi-squared statistics, as implemented in smartpca (98), to test for significant differences in the distributions of SNVs per reference genome while correcting for multiple testing using sequential Bonferroni (101).

**Data availability.** The data set generated and analyzed for this study is available in the European Nucleotide Archive (ENA) under study accession number [PRJEB31095](https://www.ebi.ac.uk/ena/record/PRJEB31095).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, XLSX file, 0.6 MB.

## ACKNOWLEDGMENTS

We thank Børge Iversen and Helle Tessand Baalsrud for their kind help in sampling Atlantic cod specimens in Lofoten and Martin Malmstrøm, Paul Ragnar Berg, and Monica Hongrø Solbakken for sampling at Sørøya. We are grateful for the metagenome sequencing performed at the Norwegian Sequencing Centre (NSC: <https://www.sequencing.uio.no>).

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

S.J., B.S., and T.H.A. conceived and designed the experiments. K.S.J. provided the initial framework for the study. E.S.R. and S.J. sampled the specimens. E.S.R. performed the laboratory work. E.S.R. and T.H.A. performed data analysis. S.V. created the Python script to convert the anvi'o format to V.C.F. Ø.B., T.H.A., E.S.R., and B.S. interpreted the results. E.S.R. and B.S. wrote the paper with the input of all authors. All authors read and approved the final manuscript.

This work was funded by grants from the Research Council of Norway (project no. 222378) and the University of Oslo (Strategic Research Initiative), both to K.S.J.

## REFERENCES

- Wang AR, Ran C, Ringø E, Zhou ZG. 2017. Progress in fish gastrointestinal microbiota research. *Rev Aquac* 10:626–640. <https://doi.org/10.1111/raq.12191>.
- Romero J, Ringø E, Merrifield DL. 2014. The gut microbiota of fish, p 75–100. In Merrifield DL, Ringø E (ed), *Aquaculture nutrition: gut health, probiotics and prebiotics*. Wiley Online Library. <https://doi.org/10.1002/9781118897263.ch4>.
- Nayak SK. 2010. Role of gastrointestinal microbiota in fish. *Aquac Res* 41:1553–1573. <https://doi.org/10.1111/j.1365-2109.2010.02546.x>.
- Olafsen JA. 1983. Ingestion of bacteria by cod (*Gadus morhua* L.) larvae, p 627–643. In *The propagation of cod Gadus morhua L.: an international symposium*. Institute of Marine Research, Flodevigen Biological Station, Bergen, Norway.
- Neuman C, Hatje E, Zarkasi KZ, Smullen R, Bowman JP, Katouli M. 2016. The effect of diet and environmental temperature on the faecal microbiota of farmed Tasmanian Atlantic salmon (*Salmo salar* L.). *Aquac Res* 47:660–672. <https://doi.org/10.1111/are.12522>.
- Zarkasi KZ, Taylor RS, Abell GCJ, Tamplin ML, Glencross BD, Bowman JP. 2016. Atlantic salmon (*Salmo salar* L.) gastrointestinal microbial community dynamics in relation to digesta properties and diet. *Microb Ecol* 71:589–603. <https://doi.org/10.1007/s00248-015-0728-y>.
- Zhou Z, Olsen RE. 2012. Culturable autochthonous gut bacteria in Atlantic salmon (*Salmo salar* L.) fed diets with or without chitin. Characterization by 16S rRNA gene sequencing, ability to produce enzymes and in vitro growth inhibition of four fish pathogens. *Aquaculture* 326–329:1–8. <https://doi.org/10.1016/j.aquaculture.2011.10.016>.
- Schmidt V, Amaral-Zettler L, Davidson J, Summerfelt S, Good C. 2016. Influence of fishmeal-free diets on microbial communities in Atlantic salmon (*Salmo salar*) recirculation aquaculture systems. *Appl Environ Microbiol* 82:4470–4481. <https://doi.org/10.1128/AEM.00902-16>.
- Desai AR, Links MG, Collins SA, Mansfield GS, Drew MD, Van Kessel AG, Hill JE. 2012. Effects of plant-based diets on the distal gut microbiome of rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 350–353: 134–142. <https://doi.org/10.1016/j.aquaculture.2012.04.005>.
- Bolnick DI, Snowberg LK, Hirsch PE, Lauber CL, Org E, Parks B, Lusi AJ, Knight R, Caporaso JG, Svanba R. 2014. Individual diet has sex-



- dependent effects on vertebrate gut microbiota. *Nat Commun* 5:4500. <https://doi.org/10.1038/ncomms5500>.
11. Bolnick DI, Snowberg LK, Hirsch PE, Lauber CL, Knight R, Caporaso JG, Svanbäck R. 2014. Individuals' diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch). *Ecol Lett* 17:979–987. <https://doi.org/10.1111/ele.12301>.
12. Givens CE, Ransom B, Bano N, Hollibaugh JT. 2015. Comparison of the gut microbiomes of 12 bony fish and 3 shark species. *Mar Ecol Prog Ser* 518:209–223. <https://doi.org/10.3354/meps11034>.
13. Miyake S, Ngugi DK, Stingl U. 2015. Diet strongly influences the gut microbiota of surgeonfishes. *Mol Ecol* 24:656–672. <https://doi.org/10.1111/mec.13050>.
14. Roeselers G, Mittge EK, Stephens WZ, Parichy DM, Cavanaugh CM, Guillemin K, Rawls JF. 2011. Evidence for a core gut microbiota in the zebrafish. *ISME J* 5:1595–1608. <https://doi.org/10.1038/ismej.2011.38>.
15. Ley RE, Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JL. 2008. Evolution of mammals and their gut microbes. *Science* 320:1647–1651. <https://doi.org/10.1126/science.1155725>.
16. Zhang M, Sun Y, Liu Y, Qiao F, Chen L, Liu WT, Du Z, Li E. 2016. Response of gut microbiota to salinity change in two euryhaline aquatic animals with reverse salinity preference. *Aquaculture* 454:72–80. <https://doi.org/10.1016/j.aquaculture.2015.12.014>.
17. Li X, Yu Y, Feng W, Yan Q, Gong Y. 2012. Host species as a strong determinant of the intestinal microbiota of fish larvae. *J Microbiol* 50:29–37. <https://doi.org/10.1007/s12275-012-1340-1>.
18. Li J, Ni J, Li J, Wang C, Li X, Wu S, Zhang T, Yu Y, Yan Q. 2014. Comparative study on gastrointestinal microbiota of eight fish species with different feeding habits. *J Appl Microbiol* 117:1750–1760. <https://doi.org/10.1111/jam.12663>.
19. Navarrete P, Magne F, Arana C, Fuentes P, Barros L, Opazo R, Espejo R, Romero J. 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. <https://doi.org/10.1371/journal.pone.0031335>.
20. Bolnick DI, Snowberg LK, Caporaso JG, Lauber C, Knight R, Stutz WE. 2014. Major histocompatibility complex class IIb polymorphism influences gut microbiota composition and diversity. *Mol Ecol* 23:4831–4845. <https://doi.org/10.1111/mec.12846>.
21. Salinas I, Casadei E, Takizawa F, Shibasaki Y, Sunyer OJ. 2018. Interactions between microbiota and the teleost immune system in health and disease. *J Immunol* 200(1 Suppl):53.19.
22. Stagaman K, Burns AR, Guillemin K, Bohannon B. 2017. The role of adaptive immunity as an ecological filter on the gut microbiota in zebrafish. *ISME J* 11:1630–1639. <https://doi.org/10.1038/ismej.2017.28>.
23. Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmström M, Gregers TF, Rounge TB, Paulsen J, Solbakken MH, Sharma A, Wetten OF, Lanzén A, Winer R, Knight J, Vogel J-H, Aken B, Andersen O, Lagesen K, Tooming-Klunderud A, Edvardsen RB, Tina KG, Espelund M, Nepal C, Previti C, Karlén BO, Moum T, Skage M, Berg PR, Gjølén T, Kuhl H, Thorsen J, Malde K, Reinhardt R, Du L, Johansen SD, Searle S, Lien S, Nilsen F, Jonassen I, Omholt SW, Stenseth NC, Jakobsen KS. 2011. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477:207–210. <https://doi.org/10.1038/nature10342>.
24. Malmström M, Matschner M, Tørresen OK, Star B, Snipen LG, Hansen TF, Baalsrud HT, Nederbragt AJ, Hanel R, Salzburger W, Stenseth NC, Jakobsen KS, Jentoft S. 2016. Evolution of the immune system influences speciation rates in teleost fishes. *Nat Genet* 48:1204–1210. <https://doi.org/10.1038/ng.3645>.
25. Magnadottir B, Gudmundsdottir S, Gudmundsdottir BK, Helgason S. 2009. Natural antibodies of cod (*Gadus morhua* L.): specificity, activity and affinity. *Comp Biochem Physiol Part B* 154:309–316. <https://doi.org/10.1016/j.cbpb.2009.07.005>.
26. Pilström L, Warr G, Strömberg S. 2005. Why is the antibody response of Atlantic cod so poor? The search for a genetic explanation. *Fish Sci* 71:961–971. <https://doi.org/10.1111/j.1444-2906.2005.01052.x>.
27. Solem ST, Stenvik J. 2006. Antibody repertoire development in teleosts—a review with emphasis on salmonids and *Gadus morhua* L. *Dev Comp Immunol* 30:57–76. <https://doi.org/10.1016/j.dci.2005.06.007>.
28. Star B, Jentoft S. 2012. Why does the immune system of Atlantic cod lack MHC II? *Bioessays* 34:648–651. <https://doi.org/10.1002/bies.201200005>.
29. Sullam KE, Essinger SD, Lozupone CA, O'Connor MP, Rosen GL, Knight R, Kilham SS, Russell JA. 2012. Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Mol Ecol* 21:3363–3378. <https://doi.org/10.1111/j.1365-294X.2012.05552.x>.
30. Riiser ES, Haverkamp THA, Varadharajan S, Borgan Ø, Jakobsen KS, Jentoft S, Star B. 2019. Switching on the light: using metagenomic shotgun sequencing to characterize the intestinal microbiome of Atlantic cod. *Environ Microbiol* 21:2576–2594. <https://doi.org/10.1111/1462-2920.14652>.
31. Star B, Haverkamp TH, 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 Microbiol* 13:248. <https://doi.org/10.1186/1471-2180-13-248>.
32. Riiser ES, Haverkamp THA, Borgan Ø, Jakobsen KS, Jentoft S, Star B. 2018. A single Vibrionales 16S rRNA oligotype dominates the intestinal microbiome in two geographically separated Atlantic cod populations. *Front Microbiol* 9:1561. <https://doi.org/10.3389/fmicb.2018.01561>.
33. Sevellec M, Laporte M, Bernatchez A, Derome N, Bernatchez L. 2019. Evidence for host effect on the intestinal microbiota of whitefish (*Coregonus* sp.) species pairs and their hybrids. *Ecol Evol* 9:11762–11774. <https://doi.org/10.1002/ece3.5676>.
34. Cohen DM, Inada T, Iwamoto T, Scialabba N. 1990. FAO species catalogue, vol 10. Gadiform fishes of the world (Order Gadiformes). Food and Agriculture Organization of the United Nations, Rome, Italy.
35. Froese R, Pauly D. 2019. fishbase.org, version (06/2018). <https://www.fishbase.org/search.php>.
36. Gaemers PAM, Poulsen JY. 2017. Recognition and distribution of two North Atlantic Gadid species, *G. argenteus* and *G. thori* (Gadidae), based on otolith morphology, larval pigmentation, molecular evidence, morphometrics and meristics. *Fishes* 2:15–24. <https://doi.org/10.3390/fishes2030015>.
37. Albert OT. 1993. Distribution, population structure and diet of silvery pout (*Gadiculus argenteus thori* J. Schmidt), poor cod (*Trisopterus minutus minutus* (L.)), four-bearded rockling (*Rhinonemus cimbrius* (L.)), and Vahl's eelpout (*Lycodes vahllei gracilis* Reinhardt) in the Norwegian Deep. *Sarsia* 78:141–154. <https://doi.org/10.1080/00364827.1993.10413531>.
38. Froese R, Pauly D. 2012. Species fact sheets: *Gadus morhua* (Linnaeus, 1758). Food and Agriculture Organization of the United Nations, Rome, Italy.
39. Link JS, Bogstad B, Sparholt H, Lilly GR. 2009. Trophic role of Atlantic cod in the ecosystem. *Fish Fish* 10:58–87. <https://doi.org/10.1111/j.1467-2979.2008.00295.x>.
40. Link JS, Garrison LP. 2002. Trophic ecology of Atlantic cod (*Gadus morhua*) on the Northeast US continental shelf. *Mar Ecol Prog Ser* 227:109–123. <https://doi.org/10.3354/meps227109>.
41. Michalsen K, Johannesen E, Bogstad B. 2008. Feeding of mature cod (*Gadus morhua*) on the spawning grounds in Lofoten. *ICES J Mar Sci* 65:571–580. <https://doi.org/10.1093/icesjms/fsn019>.
42. Godø OR, Michalsen K. 2000. Migratory behaviour of North-east Arctic cod, studied by use of data storage tags. *Fish Res* 48:127–140. [https://doi.org/10.1016/S0165-7836\(00\)00177-6](https://doi.org/10.1016/S0165-7836(00)00177-6).
43. Berg PR, Star B, Pampoulie C, Sodeland M, Barth JMI, Knutsen H, Jakobsen KS, Jentoft S. 2016. Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Sci Rep* 6:23246–23212. <https://doi.org/10.1038/srep23246>.
44. Berg PR, Star B, Pampoulie C, Bradbury IR, Bentzen P, Hutchings JA, Jentoft S, Jakobsen KS. 2017. Trans-oceanic genomic divergence of Atlantic cod ecotypes is associated with large inversions. *Heredity* (Edinb) 119:418–428. <https://doi.org/10.1038/hdy.2017.54>.
45. Star B, Boessenkool S, Gondek AT, Nikulina EA, Hufthammer AK, Pampoulie C, Knutsen H, André C, Nistelberger HM, Dierking J, Petereit C, Heinrich D, Jakobsen KS, Jentoft S, Stenseth NC, Barrett JH. 2017. Ancient DNA reveals the Arctic origin of Viking Age cod. *Proc Natl Acad Sci U S A* 114:9152–9157. <https://doi.org/10.1073/pnas.1710186114>.
46. Kirubakaran TG, Grove H, Kent MP, Sandve SR, Baranski M, Nome T, De Rosa MC, Righino B, Johansen T, Otterå H, Sonesson A, Lien S, Andersen Ø. 2016. Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. *Mol Ecol* 25:2130–2143. <https://doi.org/10.1111/mec.13592>.
47. Sodeland M, Jorde PE, Lien S, Jentoft S, Berg PR, Grove H, Kent MP, Arnyasi M, Olsen EM, Knutsen H. 2016. "Islands of divergence" in the Atlantic cod genome represent polymorphic chromosomal rearrangements. *Genome Biol Evol* 8:1012–1022. <https://doi.org/10.1093/gbe/ewv057>.
48. Holt RE, Bogstad B, Durant M, Dolgov AV, Ottersen G. 2019. Barents Sea



- cod (*Gadus morhua*) diet composition: long-term interannual, seasonal, and ontogenetic patterns. *ICES J Mar Sci* 76:1641–1652. <https://doi.org/10.1093/icesjms/fsz082>.
49. Nordeide JT, Johansen SD, Jørgensen TE, Karlsen BO, Moum T. 2011. Population connectivity among migratory and stationary cod *Gadus morhua* in the Northeast Atlantic—a review of 80 years of study. *Mar Ecol Prog Ser* 435:269–283. <https://doi.org/10.3354/meps09232>.
  50. Llewellyn MS, McGinnity P, Dionne M, Letourneau J, Thonier F, Carvalho GR, Creer S, Derome N. 2016. The biogeography of the Atlantic salmon (*Salmo salar*) gut microbiome. *ISME J* 10:1280–1284. <https://doi.org/10.1038/ismej.2015.189>.
  51. Furevik T. 2001. Annual and interannual variability of Atlantic Water temperatures in the Norwegian and Barents Seas: 1980–1996. *Deep Sea Res* 48:383–404. [https://doi.org/10.1016/S0967-0637\(00\)00050-9](https://doi.org/10.1016/S0967-0637(00)00050-9).
  52. Eilertsen HC, Skarðhamar J. 2006. Temperatures of north Norwegian fjords and coastal waters: variability, significance of local processes and air-sea heat exchange. *Estuar Coast Shelf Sci* 67:530–538. <https://doi.org/10.1016/j.ecss.2005.12.006>.
  53. Givens CE. 2012. A fish tale: comparison of the gut microbiome of 15 fish species and the influence of diet and temperature on its composition. Ph.D. dissertation. The University of Georgia, Athens, GA.
  54. Martin-Antonio B, Manchado M, Infante C, Zerolo R, Labella A, Alonso C, Borrego JJ. 2007. Intestinal microbiota variation in Senegalese sole (*Solea senegalensis*) under different feeding regimes. *Aquac Res* 38: 1213–1222. <https://doi.org/10.1111/j.1365-2109.2007.01790.x>.
  55. Hovda MB, Fontanillas R, McGurk C, Obach A, Rosnes JT. 2012. Seasonal variations in the intestinal microbiota of farmed Atlantic salmon (*Salmo salar* L.). *Aquac Res* 43:154–159. <https://doi.org/10.1111/j.1365-2109.2011.02805.x>.
  56. Karlsen BO, Klingan K, Emblem Å, Jørgensen TE, Jueterbock A, Furmanek T, Hoarau G, Johansen SD, Nordeide JT, Moum T. 2013. Genomic divergence between the migratory and stationary ecotypes of Atlantic cod. *Mol Ecol* 22:5098–5111. <https://doi.org/10.1111/mec.12454>.
  57. Stensholt BK. 2001. Cod migration patterns in relation to temperature: analysis of storage tag data. *ICES J Mar Sci* 58:770–793. <https://doi.org/10.1006/jmsc.2001.1067>.
  58. World Wildlife Fund. 2004. The Barents Sea cod. <http://www.wwf.eu/?12983/Executive-Summary-The-Barents-Sea-Cod>.
  59. Merrifield DL, Olsen RE, Myklebust R, Ringø E. 2011. Dietary effect of soybean (*Glycine max*) products on gut histology and microbiota of fish. *IntechOpen* <https://doi.org/10.5772/20101>.
  60. Ringø E, Zhou Z, Olsen RE, Song SK. 2012. Use of chitin and krill in aquaculture—the effect on gut microbiota and the immune system: a review. *Aquac Nutr* 18:117–131. <https://doi.org/10.1111/j.1365-2095.2011.00919.x>.
  61. Sakshaug E, Bjørge A, Gulliksen B, Loeng H, Mehlum F. 1994. Structure, biomass distribution, and energetics of the pelagic ecosystem in the Barents Sea: a synopsis. *Polar Biol* 14:405–411. <https://doi.org/10.1007/BF00240261>.
  62. O'Leary NA, Wright MW, Brister JR, Ciufio S, Haddad D, McVeigh R, Rajput B, Robertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badredin A, Bao Y, Blinkova O, Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey KM, Murphy MR, O'Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C, Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE, Vatsan AR, Wallin C, Webb D, Wu W, Landrum MJ, Kimchi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD, Pruitt KD. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 44:D733–D745. <https://doi.org/10.1093/nar/gkv1189>.
  63. Defosse DL, Johnson RC, Paster BJ, Dewhirst FE, Fraser GJ, Haven W. 1995. *Brevinema andersonii* gen. nov., sp. nov., an infectious spirochete isolated from the short-tailed shrew (*Blarina brevicauda*) and the white-footed mouse (*Peromyscus leucopus*). *Int J Syst Bacteriol* 45:78–84. <https://doi.org/10.1099/00207713-45-1-78>.
  64. Gajardo K, Rodiles A, Kortner TM, Kroghdal Å, Bakke AM, Merrifield DL, Sørum H. 2016. A high-resolution map of the gut microbiota in Atlantic salmon (*Salmo salar*): a basis for comparative gut microbial research. *Sci Rep* 6:30893. <https://doi.org/10.1038/srep30893>.
  65. Loy A, Pfann C, Steinberger M, Hanson B, Herp S, Brugiroux S, Gomes C, Rattei T, Stecher B. 2017. Lifestyle and horizontal gene transfer-mediated evolution of *Mucispirillum schaedleri*, a core member of the murine gut microbiota. *mSystems* 2:e00171-16. <https://doi.org/10.1128/mSystems.00171-16>.
  66. Brinchmann MF, Rajan B, Fernandes JMO, Caipang CMA, Rombout J, Kiron V. 2013. Atlantic cod (*Gadus morhua*) skin mucus proteins—focus on lectins. *Fish Shellfish Immunol* 34:1641. <https://doi.org/10.1016/j.fsi.2013.03.021>.
  67. Palko BJ, Beardsley GL, Richards WJ. 1982. Synopsis of the biological data on dolphins-fishes, *Coryphaena hippurus* Linnaeus and *Coryphaena equiselis* Linnaeus. NOAA Technical Report NMFS Circular 443. National Oceanic and Atmospheric Administration, Washington, DC.
  68. Hampson DJ, Ahmed N. 2009. Spirochaetes as intestinal pathogens: lessons from a *Brachyspira* genome. *Gut Pathog* 1:10–13. <https://doi.org/10.1186/1757-4749-1-10>.
  69. Westerman LJ, de Boer RF, Roelfsema JH, Friesema IHM, Kortbeek LM, Wagenaar JA, Bonten MJM, Kusters JG. 2013. *Brachyspira* species and gastroenteritis in humans. *J Clin Microbiol* 51:2411–2413. <https://doi.org/10.1128/JCM.01069-13>.
  70. Oxberry SL, Trott DJ, Hampson DJ. 1998. *Serpulina pilosicoli*, water-birds and water: potential sources of infection for humans and other animals. *Epidemiol Infect* 121:219–225. <https://doi.org/10.1017/S0950268898008863>.
  71. Thompson FL, Hoste B, Thompson CC, Goris J, Gomez-Gil B, Huys L, De Vos P, Swings J. 2002. *Enterovibrio norvegicus* gen. nov., sp. nov., isolated from the gut of turbot (*Scophthalmus maximus*) larvae: a new member of the family *Vibrionaceae*. *Int J Syst Evol Microbiol* 52: 2015–2022. <https://doi.org/10.1099/00207713-52-6-2015>.
  72. Thompson FL, Thompson CC, Naser S, Hoste B, Vandemeulebroecke K, Munn C, Bourne D, Swings J. 2005. *Photobacterium rosenbergii* sp. nov. and *Enterovibrio coralli* sp. nov., vibrios associated with coral bleaching. *Int J Syst Evol Microbiol* 55:913–917. <https://doi.org/10.1099/ijs.0.63370-0>.
  73. Pujalte MJ, Sitja-Bobadilla A, Alvarez-Pellitero P, Garay E. 2003. Carriage of potentially fish-pathogenic bacteria in *Sparus aurata* cultured in Mediterranean fish farms. *Dis Aquat Organ* 54:119–126. <https://doi.org/10.3354/dao054119>.
  74. Company R, Sitja-Bobadilla A, Pujalte MJ, Garay E, Alvarez-Pellitero P, Perez-Sanchez J. 2002. Bacterial and parasitic pathogens in cultured common dentex, *Dentex dentex* L. *J Fish Dis* 22:299–309. <https://doi.org/10.1046/j.1365-2761.1999.00182.x>.
  75. Pascual J, Macia MC, Arahal DR, Garay E, Pujalte MJ. 2009. Description of *Enterovibrio nigricans* sp. nov., reclassification of *Vibrio calviensis* as *Enterovibrio calviensis* comb. nov. and emended description of the genus *Enterovibrio* Thompson et al. 2002. *Int J Syst Evol Microbiol* 59:698–704. <https://doi.org/10.1099/ijs.0.001990-0>.
  76. Org E, Parks BW, Joo JWJ, Emert B, Schwartzman W, Kang EY, Mehra-bian M, Pan C, Knight R, Gunsalus R, Drake TA, Eskin E, Lusis AJ. 2015. Genetic and environmental control of host-gut microbiota interactions. *Genome Res* 25:1558–1569. <https://doi.org/10.1101/gr.194118.115>.
  77. Goodrich JK, Davenport ER, Waters JL, Clark AG, Ley RE. 2016. Cross-species comparisons of host genetic associations with the microbiome. *Science* 352:532–535. <https://doi.org/10.1126/science.aad9379>.
  78. McFall-Ngai M. 2007. Care for the community. *Nature* 445:153. <https://doi.org/10.1038/445153a>.
  79. Norecopa. 2020. Norecopa guidelines for animal experiments. <https://norecopa.no/prepare>.
  80. Gilbert JA, Meyer F, Jansson J, Gordon J, Pace N, Ley R, Fierer N, Field D, Kyrpides N, Glöckner F. 2010. The Earth Microbiome Project: meeting report of the “1st EMP meeting on sample selection and acquisition” at Argonne National Laboratory October 6th 2010. *Stand Genomic Sci* 3:249–253. <https://doi.org/10.4056/aigs.1443528>.
  81. Andrews S. 2010. FastQC. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
  82. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
  83. Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>.
  84. Tørresen OK, Star B, Jentoft S, Reinart WB, Grove H, Miller JR, Walenz BP, Knight J, Ekholm JM, Peluso P, Edvardsen RB, Tooming-Klunderud A, Skage M, Lien S, Jakobsen KS, Nederbragt AJ. 2017. An improved genome assembly uncovers prolific tandem repeats in Atlantic cod. *BMC Genomics* 18:123. <https://doi.org/10.1186/s12864-016-3448-x>.

85. Genome Reference Consortium. 2009. Genome Reference Consortium Human Build 37 (GRCh37). <https://www.ncbi.nlm.nih.gov/grc/human/data?asm=GRCh37>.
86. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
87. Bushnell B. 2014. BBMap: a fast, accurate, splice-aware aligner. <https://www.osti.gov/servlets/purl/1241166>.
88. Li H. 2012. *Seqtk*. <https://github.com/lh3/seqtk>.
89. Menzel P, Ng KL, Krogh A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7:11257. <https://doi.org/10.1038/ncomms11257>.
90. Racine JS. 2010. Rstudio: a platform-independent ide for R and SWEAVE. *Financ Dev* 47:36–37. <https://doi.org/10.1002/jae.1278>.
91. R Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
92. Gentleman R, Carey V, Huber W, Hahne F. 2019 Package ‘genefilter.’ <https://bioconductor.org/packages/release/bioc/html/genefilter.html>.
93. Wickham H. 2016. *ggplot2: elegant graphics for data analysis*. Springer-Verlag, New York, NY. <https://doi.org/10.1007/978-3-319-24277-4>.
94. Machado H, Gram L. 2017. Comparative genomics reveals high genomic diversity in the genus *Photobacterium*. *Front Microbiol* 8:1204. <https://doi.org/10.3389/fmicb.2017.01204>.
95. Köster J, Rahmann S. 2012. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* 28:2520–2522. <https://doi.org/10.1093/bioinformatics/bts480>.
96. Eren AM, Esen C, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi’o: an advanced analysis and visualization platform for ‘omics data. *PeerJ* 3:e1319. <https://doi.org/10.7717/peerj.1319>.
97. Eren AM, Esen C, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Visualizing SNV profiles using R. *merenlab.org*. <http://merenlab.org/tutorials/infant-gut>.
98. Patterson N, Price AL, Reich D. 2006. Population structure and eigenanalysis. *PLoS Genet* 2:e190. <https://doi.org/10.1371/journal.pgen.0020190>.
99. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O’Hara RB, Simpson GL, Solymos P, Stevens MHH, Scoecs E, Wagner H. 2017. *vegan: Community Ecology Package*. R package version 2.4.3. <https://rdrr.io/cran/vegan/man/vegan-package.html>.
100. Martinez Arbizu P. 2019 pairwiseAdonis: pairwise multilevel comparison using adonis. R package version 0.1. <https://github.com/pmartinezarbizu/pairwiseAdonis>.
101. Holm S. 1979. A simple sequentially rejective multiple test procedure. *Scand J Stat* 6:65–70.
102. McMurdie PJ, Holmes S. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol* 10:e1003531. <https://doi.org/10.1371/journal.pcbi.1003531>.