environmental microbiology

Environmental Microbiology (2022) **00**(00), 00–00 doi:10.1111/1462-2920.15901

Prevalence of genetically similar *Flavobacterium columnare* phages across aquaculture environments reveals a strong potential for pathogen control

Anniina Runtuvuori-Salmela,[1] Heidi M. T. Kunttu,[1] Elina Laanto,[1,2] Gabriel M. F. Almeida,[1,3] Kati Mäkelä,[1] Mathias Middelboe[4] and Lotta-Riina Sundberg[1*]

[1]*Department of Biological and Environmental Science and Nanoscience Center, University of Jyväskylä, Jyväskylä, Finland.*

[2]*Faculty of Biological and Environmental Sciences, Molecular and Integrative Biosciences Research Programme, University of Helsinki, Helsinki, Finland.*

[3]*Faculty of Biosciences, Fisheries and Economics, Norwegian College of Fishery Science, UiT The Arctic University of Norway, Tromsø, Norway.*

[4]*Department of Biology, Marine Biological Section, University of Copenhagen, Helsingør, Denmark.*

Summary

Intensive aquaculture conditions expose fish to bacterial infections, leading to significant financial losses, extensive antibiotic use and risk of antibiotic resistance in target bacteria. *Flavobacterium columnare* causes columnaris disease in aquaculture worldwide. To develop a bacteriophage-based control of columnaris disease, we isolated and characterized 126 *F. columnare* strains and 63 phages against *F. columnare* from Finland and Sweden in 2017. Bacterial isolates were virulent on rainbow trout (*Oncorhynchus mykiss*) and fell into four previously described genetic groups A, C, E and G, with genetic groups C and E being the most virulent. Phage host range studied against a collection of 227 bacterial isolates (from 2013 to 2017) demonstrated modular infection patterns based on host genetic group. Phages infected contemporary and previously isolated bacterial hosts, but bacteria isolated most recently were generally resistant to previously isolated phages. Despite large differences in geographical origin, isolation year or host range of the phages, whole-genome sequencing of 56 phages showed high level of genetic similarity to previously isolated *F. columnare* phages (*Ficleduovirus*, *Myoviridae*). Altogether, this phage collection demonstrates a potential for use in phage therapy.

Received 25 September, 2020; accepted 11 January, 2022. *For correspondence. E-mail lotta-riina.sundberg@jyu.fi; Tel. (+358) 40 805 3931.

Introduction

During the past 20 years, aquaculture has been the fastest-growing food production sector (FAO, 2014), providing an important source of protein for human consumption. Intensive aquaculture production is based on monocultures of certain species, which are reared in high population densities, resulting in increased transmission of infections (Pulkkinen *et al*., 2010; Oidtmann *et al*., 2011) and antibiotic use (FAO, 2014). Approximately 70%–80% of antibiotics in aquaculture may end up in the environment (Cabello *et al*., 2013; Watts *et al*., 2017), where they select for antibiotic resistance also in the environmental bacteria (Tamminen *et al*., 2011; Yang *et al*., 2013). The World Health Organization (WHO) has declared antibiotic resistance as a major risk for global health and food security, and means to control diseases without antibiotics are therefore urgently needed.

Due to increased issues with antibiotic resistance and lack of efficient vaccines, application of lytic bacteriophages (phages) has been suggested as an alternative for controlling pathogenic bacteria (reviewed by e.g. Watts, 2017; Kortright *et al*., 2019). Their host specificity makes phages strong candidates as tools for targeted eradication of pathogenic bacteria. Indeed, the use of phage as therapeutics has a long history in medicine (Almeida and Sundberg, 2020), and recently interest towards other types of applications has increased. Phages can be used, for example to extend the food shelf life (Moye *et al*., 2018), and are already in use against *Listeria* in salad, salmon and meat packages (Sulakvelidze, 2013; Lone *et al*., 2016). Interest towards using phages in aquaculture has been steadily increasing the past decade, including studies on phage–bacterium

interactions for important fish pathogens such as Flavobacteria, Vibrio and Aeromonas (Castillo *et al*., 2012; Silva *et al*., 2014; Tan *et al*., 2014).

Flavobacterium columnare (Bacteroidetes) is a Gram-negative bacterium (Bernardet *et al*., 1996) causing columnaris epidemics at freshwater fish farms worldwide (Declercq *et al*., 2013). Conditions such as water temperature over 18°C and high fish density promote columnaris epidemics, which spread rapidly in the rearing units, and may lead to high mortality if not treated by antibiotics (Suomalainen *et al*., 2005; Karvonen *et al*., 2010; Pulkkinen *et al*., 2010). Moreover, the intensive production of rainbow trout (*Oncorhynchus mykiss*) has been shown to select for highly virulent *F. columnare* strains in the aquaculture environment (Sundberg *et al*., 2016), with different genetic groups (genotypes hereafter) being connected to different host species at the global scale (LaFrentz *et al*., 2018).

However, how the pathogen population structure in aquaculture systems influences the genetic and phenotypic (especially host range) patterns in phages is not understood. Phages infecting *F. columnare* described so far have been relatively host specific with a narrow host range (Laanto *et al*., 2011), and phage addition during experimental columnaris infection has shown promising results on the survival of rainbow trout fry (Laanto *et al*., 2015; Almeida *et al*., 2019). However, for development of successful phage therapy approach, it is essential to obtain an overview of the diversity and spatial distribution of both phage and bacterial populations, and establish a phage collection that covers this diversity. This requires isolation, and subsequent genetic and functional characterization of phage and host communities, and description of the phage–host interactions.

In this study, we isolated new *F. columnare* strains and their phages from 10 different aquaculture locations in Finland and Sweden during summer 2017. Highly virulent bacterial strains occurred at several farms. Phage infection patterns were studied on 227 different bacterial strains isolated during 2013–2017. Bacterial isolates were genetically characterized and their virulence was determined on rainbow trout. Morphology of the phages was confirmed with transmission electron microscopy (TEM), and whole-genome sequencing was performed for 56 of the isolated phages. We show that geographically distant *F. columnare* phages have very similar genomic composition and cluster according to the genetic groups of the host bacteria. The phages were able to infect bacteria isolated from different fish farms; however, the impaired capacity of phages isolated earlier to infect bacteria in a later time point suggests that bacteria evolve resistance against phage in the aquaculture conditions.

Results

Isolation and genetic characterization of F. columnare *strains*

We isolated 111 *F. columnare* strains from water samples from 10 different locations in Finland and Sweden (Supplementary Table S1). In addition, 15 Swedish *F. columnare* strains were obtained from National Veterinary Institute, Sweden. Globally, *F. columnare* strains can be classified into six genomovars (I, I/II, II, II-A, II-B and III) by restriction fragment length polymorphism (RFLP) of 16S rDNA (Hadisaputro and Wakabayashi, 1999; LaFrentz *et al*., 2014; LaFrentz *et al*., 2017; García *et al*., 2018). In this study, all isolated bacterial strains were classified as genomovar I strains, except for the previously isolated Swedish strains F310, F383 and F514, which were classified as genomovar I/II (Table 1, Supplementary Table S1).

From altogether 126 bacterial isolates, 121 Finnish and Swedish strains could be assigned into previously defined genetic groups (A, C, G, E, Suomalainen *et al*., 2006) using RFLP analysis of the internal transcribed spacer (ITS) region between 16S rRNA and 23S rRNA genes. Most of the strains fell into genetic group C (73 strains) or E (24 strains) (Fig. 1; Supplementary Table S1), whereas eight strains belonged to group G and 16 to group A. Bacterial strains isolated from Sweden fell into two genetic groups (A and E), but during 2017, only genetic group A strains were isolated (Supplementary Table S1). Five Swedish strains (Supplementary Table S1) could not be assigned to any of the previously defined genetic groups and were designated as ND1 (F397), ND2 (F512) and ND3 (F310, F383 and F514).

In general, the isolates from each fish farm represented a specific genomic group (Fig. 1) suggesting that specific genomic groups are dominating the *F. columnare* communities at the individual fish farm. Only from the most frequently sampled farm (Farm 1), we isolated bacteria belonging to two genetic groups (A and C) (Fig. 1, Supplementary Table S1). *Flavobacterium columnare* was not isolated from farms 4 and 9.

Virulence experiment

Virulence of 34 selected *F. columnare* isolates representing different genetic groups was studied on rainbow trout fry. Of the isolates, 17 were categorized as high virulence (estimated survival time <15 h), 16 as medium virulence (estimated survival time >15 h) and one as low virulence (no difference to uninfected control: $p = 0.662$, Kaplan–Meier survival analysis) isolate. The virulence observed among isolates belonging to genetic groups E and C (Table 1) were significantly higher than for groups A and G isolates ($p < 0.001$) (Fig. 2), with

Table 1. Virulence of 34 *Flavobacterium columnare* isolates in rainbow trout fry (*Oncorhynchus mykiss*) in a 24-h experiment.

Isolate	Isolation year	Country	Farm n:o	Genomovar	Genetic group	Mortality (%)	Mean estimated survival time (h)	Virulence
FCO-F16	2017	Finland	3	I	E	100.0	12.87	High
FCO-F81	2017	Finland	6	I	E	100.0	13.00	High
FCO-F88	2017	Finland	6	I	E	100.0	13.07	High
FCO-F2	2017	Finland	1	I	C	100.0	13.13	High
FCO-F13	2017	Finland	3	I	E	100.0	13.53	High
FCO-F118	2017	Finland	7	I	C	100.0	13.53	High
FCO-F11	2017	Finland	3	I	E	100.0	13.60	High
FCO-F22	2017	Finland	1	I	C	100.0	13.60	High
FCO-F33	2017	Finland	1	I	C	100.0	13.67	High
FCO-F58	2017	Finland	1	I	C	100.0	13.73	High
FCO-F98	2017	Finland	5	I	C	100.0	13.80	High
FCO-F50	2017	Finland	1	I	C	100.0	13.93	High
FCO-F86	2017	Finland	6	I	E	100.0	14.00	High
FCO-F30	2017	Finland	1	I	C	100.0	14.13	High
FCO-F78	2017	Finland	5	I	C	100.0	14.20	High
FCO-F41	2017	Finland	1	I	C	100.0	14.33	High
F310	2000	Sweden	10	I/II	ND3	100.0	14.93	High
F387	2002	Sweden	10	I	E	93.3	15.33	Medium
F383	2002	Sweden	10	I/II	ND3	100.0	15.40	Medium
FCO-F45	2017	Finland	8	I	A	100.0	15.53	Medium
F514	2013	Sweden	10	I/II	ND3	100.0	15.87	Medium
FCO-F42	2017	Finland	8	I	A	100.0	15.87	Medium
FCO-S1	2017	Sweden	10	I	A	100.0	16.00	Medium
FCO-F32	2017	Finland	1	I	A	100.0	16.21	Medium
FCO-F40	2017	Finland	1	I	C	100.0	16.27	Medium
FCO-F47	2017	Finland	8	I	A	100.0	16.33	Medium
3/3449	2017	Sweden	10	I	A	100.0	16.40	Medium
FCO-F5	2017	Finland	2	I	G	100.0	16.53	Medium
FCO-F3	2017	Finland	2	I	G	100.0	16.60	Medium
FCO-F9	2017	Finland	2	I	G	100.0	17.00	Medium
FCO-F8	2017	Finland	2	I	G	100.0	17.13	Medium
F524	2014	Sweden	10	I	A	100.0	19.87	Medium
F512	2013	Sweden	10	I	ND2	100.0	20.67	Medium
F397	2002	Sweden	10	I	ND1	20.0	22.33	Low
Neg. control						30.0	23.90	

Virulence levels (high, medium or low) were based on estimated survival time, Kaplan–Meier survival analysis.

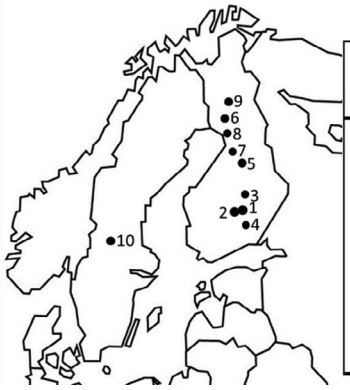
significantly faster cumulative mortality caused by group E than group C isolates ($p < 0.001$). Mortality caused by genetic groups A and G, on the other hand, did not differ from each other ($p = 0.865$).

In addition to the genetic group, initial isolation source of the bacterium (fish farm water or directly from fish, Supplementary Table S1) had a significant effect on bacterial virulence, with isolates from tank water being more virulent than bacteria isolated from fish ($p < 0.001$). There were also differences in mortalities caused by bacteria isolated from different fish species/rearing tanks of different species ($p < 0.001$), the isolates from salmon (*Salmo salar*) being the most and the isolates from trout (*S. trutta*) the least virulent.

Characterization of bacteriophages

Sixty-three bacteriophages were isolated from the water samples originating from seven different fish farms (Fig. 1; Table 2). The majority (52 out of 63) of the phages were isolated against hosts from genetic group C, while the rest of the phages were isolated against genetic groups A or G. No phages were isolated from mixed bacterial enrichment cultures or with genetic group E bacteria. Phages infecting hosts from different genetic groups [FCOV-S1 (A), FCOV-F13 (G) and FCOV-F27 (C), Supplementary Fig. S1] were morphologically similar under TEM and displayed typical characteristics of myoviruses, icosahedral capsid with a rigid, relatively thick tail (Supplementary Fig. S1). Interestingly, phages against genetic group C were isolated from fish farm 4, even though no *F. columnare* was detected from the farm during the sampling.

Host range of 71 *F. columnare* phages isolated in this study and previously isolated phages (Table 2) was tested in total with 227 different bacterial strains (Supplementary Table S4). Of these, 51 Finnish and eight Swedish bacterial strains were susceptible to one or more of the phages (Supplementary Table S4). In cases where clear infections and plaques were not



Farm number	Number of isolated *Flavobacterium columnare* bacteria (*Fc*) and their phages per genetic group													
	A		C		E		G		ND1		ND2		ND3	
	Fc	Phage	*Fc*	Phage	*Fc*	Phage	Fc	Phage	*Fc*	Phage	*Fc*	Phage	*Fc*	Phage
1	1	1	50	23	-	-	-	6	-	-	-	-	-	-
2	-	-	-	1	-	-	8	2	-	-	-	-	-	-
3	-	-	-	5	6	-	-		-	-	-	-	-	-
4	-	-	-	13	-	-	-	-	-	-	-	-	-	-
5	-	-	20	3	-	-	-	-	-	-	-	-	-	-
6	-	-	-	-	16	-	-	-	-	-	-	-	-	-
7	-	-	3	7	-	-	-	-	-	-	-	-	-	-
8	7	-	-	-	-	-	-	-	-	-	-	-	-	-
9	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10	8	2	-	-	2	-	-	-	1	-	1	-	3	-

Fig. 1. Sampling locations of fish farms in Finland and Sweden. On the left: Map of Northern Europe, where each number indicates a farm, where the water and fish samples were collected. Exact number and locations of Swedish farms are not known, and they are jointly marked as farm 10. On the right: Number of bacterial and bacteriophage isolates from individual fish farms. A, C, E, G, ND1, ND2 and ND3 indicate the different genetic groups of the isolated bacteria and the isolation hosts of the phage. ND = genetic group not determined.

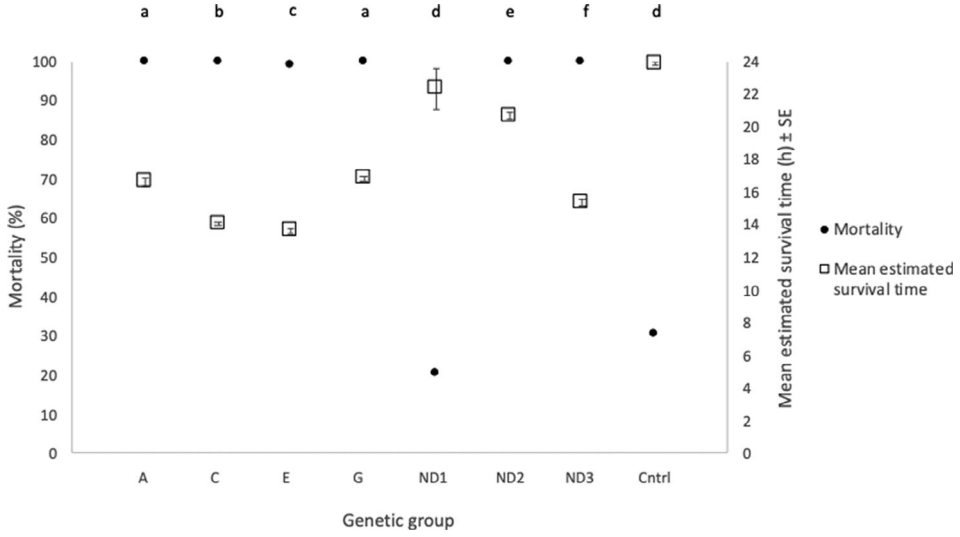


Fig. 2. Mortality percent and mean estimated survival time (±SE) of rainbow trout (*Oncorhynchus mykiss*) during 24-h experimental infection with *Flavobacterium columnare* isolates representing genetic groups A, C, E and G. ND = genetic group not determined, Cntrl = control with no bacterial infection. Different lowercase letters indicate statistical difference in cumulative mortality (Kaplan–Meier Survival Analysis) between the genetic groups.

observed, majority of the phages caused growth inhibition on bacterial lawn (Supplementary Table S4).

Bacteria isolated in 2017 were generally resistant to phages isolated earlier (Fig. 3, Supplementary Table S4). The phages characterized in this study did not, in general, infect bacteria isolated in the USA or other tested Flavobacterium species.

The phages isolated in 2017 infected both contemporary and previously isolated *F. columnare* strains within the genetic cluster, regardless of the isolation origin of the bacteria. Network analysis of the phage infection patterns revealed clustered interactions defined by the host genetic group (Fig. 3, see also Supplementary Table S4). In other words, phages isolated against a host from a specific genetic group infected generally bacteria from the same genetic group. For example FCOV-S1 isolated against host from genetic group A, infected generally bacteria from genetic group A.

Few phages deviated from these cluster patterns, showing cross-infection to hosts from a different genetic group. C-genetic group infecting phages FCOV-F25, F26, and F27 were able to infect bacterial isolates from genetic groups E (isolates FCO-F13 and FCO-F14) and G (isolate B442). FCOV-F28 was able to infect bacteria from genetic group G (isolate B442) in addition to group C bacteria (Supplementary Table S4). The pattern differentiated also when isolation year was regarded, as the phages isolated earlier (V183 and V184, V188, V189 isolated in 2015 from farm 1) were able to infect only earlier isolated bacterial strains (B245 in 2009 and B526 in 2012).

Four phages isolated with G host B185 (FCOV-F13, F14, F15 and F16) were able to also infect genetic group

Table 2. Bacteriophages isolated in this study.

Phage	Fish farm n:o	Isolation year/date	Source	Isolation host/ genetic group	Number of predicted ORFs	Genome length (kbp)	TEM	Accession number	Genetically identical isolates
FCL-1*	2	2007		B076/A					
FCL-2*	2	2008	–	B185/G		47.1	x	NC_027125	c
FCV-1*	1	2009	–	C1/C		46.5	x	NC_041845	h
V182*	1	2014	–	B245/C		49.1		KY979242	i
V183	1	2015	–	B245/C	76	49.1		MT585311	p
V184	1	2015	–	B245/C	76	49.1		MT585312	p
V186	1	2015	–	B067/A	74	46.5		MT585313	b
V188	1	2015	–	B245/C	76	49.1		MT585314	p
V189	1	2015	–	B245/C	76	49.1		MT585315	p
FCOV-S1	10	3.8.2017	Tank water	B534/A	74	46.5	x	MK756094	a
FCOV-S2	10	3.8.2017	Tank water	B067/A	74	46.4	x	MK756095	a
FCOV-F1	1	6.7.2017	Tank water	B537/C	76	49.1		MT585273	j
FCOV-F2	1	6.7.2017	Tank water	B537/C	76	49.1	x	MK756083	j
FCOV-F3	1	6.7.2017	Tank water	B537/C	76	49.1	x	MT585274	j
FCOV-F4	1	6.7.2017	Tank water	B537/C	76	49.1		MT585275	j
FCOV-F5	3	24.7.2017	Tank water	B537/C			x		
FCOV-F6	3	24.7.2017	Tank water	B537/C	76	49.1		MK756084	j
FCOV-F7	3	24.7.2017	Tank water	B537/C	76	49.1		MT585276	j
FCOV-F8	1	7.8.2017	Farm outlet water	B537/C	76	49.1		MT585277	k
FCOV-F9	1	7.8.2017	Tank water	B537/C	76	49.1	x	MK756085	k
FCOV-F10	1	7.8.2017	Tank water	FCO-F2/C	76	49.1		MT585278	j
FCOV-F11	1	7.8.2017	Tank water	FCO-F2/C	76	49.1	x	MT585279	j
FCOV-F12	4	8.8.2017	Farm outlet water	FCO-F2/C	76	49.1		MT585280	n
FCOV-F13	1	7.8.2017	Tank water	B185/G	74	47.2	x	MK756086	d
FCOV-F14	1	7.8.2017	Tank water	B185/G	74	47.2		MT585281	e
FCOV-F15	1	7.8.2017	Tank water	B185/G	74	47.2		MT585282	d
FCOV-F16	1	7.8.2017	Tank water	B185/G	74	47.2	x	MK756087	e
FCOV-F17	1	7.8.2017	Tank water	B537/C	76	49.1		MT585283	k
FCOV-F18	1	7.8.2017	Tank water	FCO-F2/C	76	49.1		MT585284	k
FCOV-F19	7	23.8.2017	Tank water	B537/C	76	49.1		MT585285	o
FCOV-F20	5	18.8.2017	Tank water	B537/C	76	49.1		MT585286	o
FCOV-F21	5	18.8.2017	Tank water	B537/C	76	49.1		MT585287	o
FCOV-F22	7	23.8.2017	Farm outlet water	FCO-F2/C	76	49.1		MT585288	o
FCOV-F23	1	7.8.2017	Tank water	B537/C	76	49.1		MT585289	k
FCOV-F24	1	7.8.2017	Tank water	FCO-F2/C	76	49.1		MT585290	j
FCOV-F25	1	7.8.2017	Tank water	B537/C	76	49.1	x	MK756088	j
FCOV-F26	1	7.8.2017	Tank water	FCO-F2/C	76	49.1		MT585291	j
FCOV-F27	1	7.8.2017	Tank water	FCO-F2/C	76	49.2		MK756089	j
FCOV-F28	1	7.8.2017	Tank water	B537/C	76	49.1		MT585292	m
FCOV-F29	1	7.8.2017	Tank water	B537/C	76	49.1		MT585293	l
FCOV-F30	1	7.8.2017	Tank water	B537/C	76	49.1		MT585294	k
FCOV-F31	1	7.8.2017	Tank water	FCO-F2/C	76	49.1		MT585295	l
FCOV-F32	1	7.8.2017	Tank water	FCO-F2/C	76	49.1		MT585296	j
FCOV-F33	4	8.8.2017	Tank water	B537/C	76	49.1		MT585297	n
FCOV-F34	4	8.8.2017	Tank water	B537/C	76	49.1		MT585298	n
FCOV-F35	4	8.8.2017	Tank water	B537/C	76	49.1		MT585299	n
FCOV-F36	4	8.8.2017	Tank water	B537/C	76	49.1		MT585300	n
FCOV-F37	4	8.8.2017	Tank water	B537/C	76	49.1		MK756090	n
FCOV-F38	4	8.8.2017	Tank water	B537/C	76	49.1		MT585301	n
FCOV-F39	4	8.8.2017	Tank water	FCO-F2/C	76	49.1		MT585302	n
FCOV-F40	4	8.8.2017	Tank water	FCO-F2/C	76	49.1		MT585303	n
FCOV-F41	4	8.8.2017	Tank water	FCO-F2/C	76	49.1		MT585304	n
FCOV-F42	4	8.8.2017	Tank water	FCO-F2/C	76	49.1		MT585305	n
FCOV-F43	4	8.8.2017	Tank water	FCO-F2/C	76	49.1		MT585306	n
FCOV-F44	4	8.8.2017	Tank water	FCO-F2/C	76	49.1		MT585307	n
FCOV-F45	2	15.8.2017	Tank water	B185/G	74	47.2		MK756091	f
FCOV-F46§	2	28.8.2017	Tank water	B185/G	74	47.2		MK756092	g
	5	18.8.2017	Tank water	B537/C	76	49.1		MT585308	o

(Continues)

Table 2. Continued

Phage	Fish farm n:o	Isolation year/date	Source	Isolation host/ genetic group	Number of predicted ORFs	Genome length (kbp)	TEM	Accession number	Genetically identical isolates
FCOV-F47§									
FCOV-F48§	2	15.8.2017	Tank water	B537/C	76	49.1		MT585309	p
FCOV-F49§	7	23.8.2017	Tank water	B537/C					
FCOV-F50§	7	23.8.2017	Tank water	B537/C					
FCOV-F51§	3	24.7.2017	Tank water	B537/C					
FCOV-F52§	7	7.8.2017	Tank water	B537/C					
FCOV-F53§	7	7.8.2017	Tank water	B537/C					
FCOV-F54§	1	3.9.2017	Tank water	B185/G	74	47.2		MK756093	d
FCOV-F55§	1	3.9.2017	Tank water	B185/G					
FCOV-F56§	1	3.9.2017	Tank water	B067A	74	46.5		MT585310	b
FCOV-F58§	7	7.8.2017	Tank water	B537/C					
FCOV-F59§	1	7.8.2017	Tank water	B537/C					
FCOV-F60§	1	7.8.2017	Tank water	B537/C					
FCOV-F61§	1	7.8.2017	Tank water	B537/C					
FCOV-F62§	3	24.7.2017	Tank water	B537/C					

Phages marked with asterisks have been isolated and characterized in previous studies. Small letters in the last column indicate genetically identical genomes (same letter = identical genome). ORFs were predicted with GenemarkS. For details of isolation hosts, see Supplementary Table S2.

§ phages isolated with mucin supplement (Almeida *et al*., 2019).

* Phages characterized in Laanto *et al*. (2011) and Laanto *et al*. (2017).

A (3/3449, 5/3451, isolated from Sweden in 2017) bacteria (Fig. 5, Supplementary Table S4). In addition, FCOV-F16 was able to infect also ND1 (isolate F397) and ND3 (isolate F310) strains from Sweden.

Interestingly, these extended host ranges were associated with minimal or no genetic differences in phage genomes (see below).

Genetic characterization of bacteriophages

Sequencing of the phage genomes revealed highly similar genomes despite the different host range of the phages (Fig. 4). Comparison of phage genomes across infection clusters (i.e. across host genomic groups, VIRIDIC) showed a nucleotide level identity between representative phage genomes of C- and G-phages of 84.4%, whereas A and G phages shared 93.41% identity, and A and C phages 88.3% (Fig. 4, Supplementary Fig. S2).

Length of the linear phage genomes varied from 46 kbp with 74 open reading frames (ORFs) (Cluster A phages) and 47 kbp with 74 ORFs (Cluster G) up to 49 kbp with 76 ORFs (Cluster C) (Supplementary Table S2). The ends of the linear DNA were Sanger sequenced here for two A cluster phages and the sequences were identical to the previously determined genome ends. The ends of these linear phage genomes have a 13 nt long 5′ cohesive ends with base-pairing complementary. This has previously been experimentally verified by ligation following sequencing of the overlap in the circularized genomes (data not shown). Most of the differences between phages infecting A-, C- and G -hosts were located at the left side of the genome, whereas approximately the first 29 kbps of the genomes were more conserved (Fig. 4, Supplementary Table S5). The conserved area consists of the predicted, e.g. packaging and structural genes. None of the phages had ORFs that could be directly associated with lysogeny, virulence, horizontal gene transfer or antibiotic resistance.

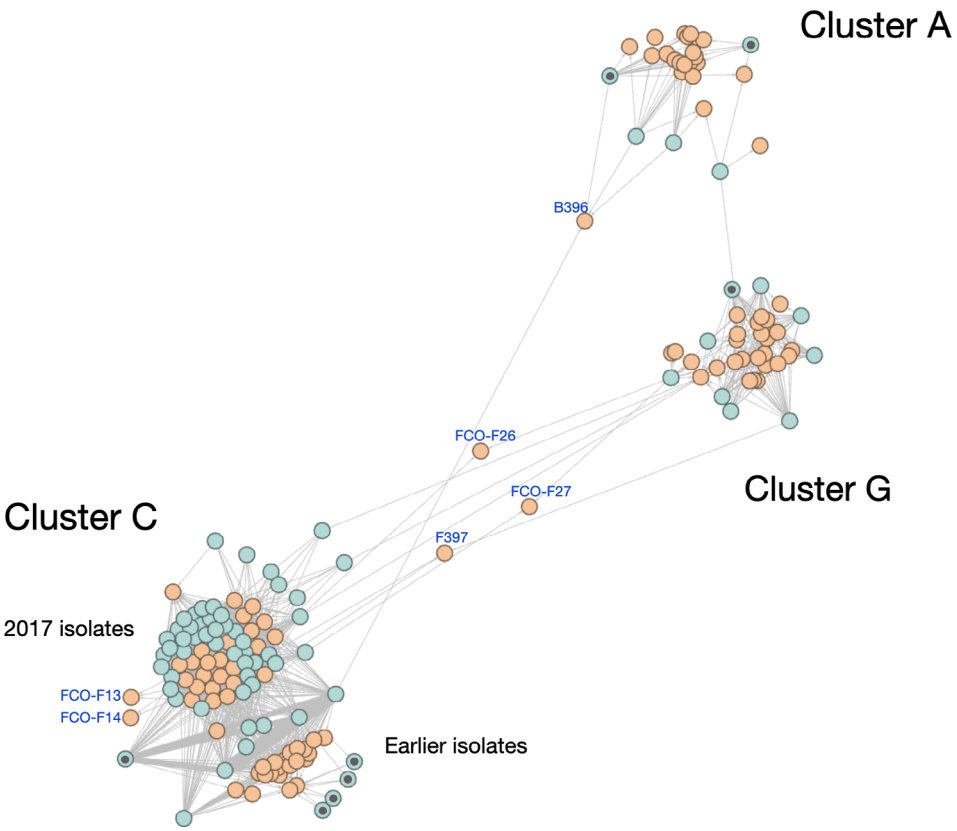


Fig. 3. Phage–bacterium interaction network. Infection patterns of phages (light blue) in bacterial hosts (orange) are clustered based on the genetic group of the bacterial host (A, C and G). Dark dot within phage circles indicates phages isolated earlier. Some key bacterial species infected by two different phage groups (FCO-F26, FCO-F27, F397, B396) or belonging to genetic group E (FCO-F13 and FCO-14) are indicated. Network was visualized using Force Atlas 2 algorithm in Gephi, where modularity of the community is interpreted by comparing the nodes with each other.

Phylogenetic tree based on complete phage genomes indicated similar clustering as the phage infection patterns (Fig. 5). Phylogenomic Genome-BLAST Distance Phylogeny (GBDP) tree inferred using the formula D6 yielded average support of 70%. The OPTSIL clustering yielded two species clusters (C-phages and G + A phages), and one genus level. All the phages characterized here can be assigned to the unclassified Ficleduoviruses (*Myoviridae*), but represent two different species.

Within the individual infection and species clusters, the phages had very similar genomes. The nucleotide level identities between phage genomes within infection clusters A, C and G (Fig. 3) were high: 99.7%–100% between A-phages, 98.2%–100% between G phages (including FCL-2) and 94.9%–100% between C-phages (Supplementary Fig. S2). It should also be noted that genetically identical phages were isolated from different fish farms, e.g. cluster C phages FCOV-F4 and FCOV-F6 (from Farms 1 and 3 respectively) and FCOV-F20 and FCOV-F22 (from Farms 5 and 7), and Cluster G phages FCOV-F45 and FCOV-F54 (Farms 2 and 1) (Table 1).

In general, the nucleotide level differences leading to amino acid level changes were detected in the putative structural proteins, in addition to several ORFs without putative annotated function (hypothetical proteins). Detailed list of differing ORFs is provided in Supplementary Table S5. In cluster A phages, changes were seen also in the putative 3′-phosphoadenosine-5′-phosphosulfate (PAPS), putative DNA methylase (ORF 47), and in the putative replication proteins (ORFs 50 and 51). In cluster G phages, the additional ORFs with amino acid changes were the putative peptidase (ORF 16) and ssDNA-binding protein (ORF 63).

Notably, Cluster A-phages originating from different countries (Fig. 6A) shared surprisingly high level of genetic identity. Within each phage cluster the hot spots for genetical variability were found in different parts of the genomes (compared with the consensus, Fig. 6). In Cluster A-phages the hot spot was located in the area coding for putative replication proteins (after 35 kbp), while in C-phages it was in the area encoding putative tail proteins. In addition, the area around 25 kbp (from the genome start), which has been speculated to code for tail fibre proteins (e.g. in FCOV-F25 ORFs 35 and 36) (Laanto *et al*., 2017), was also characterized by variability among phages.

Within Cluster G phages, genetically identical FCOV-F45 and FCOV-F46 differed from the other Cluster G phages isolated in 2017, although only two nucleotide-level differences were detected. One change was located in non-coding area (11 837 bp) and the other one was a

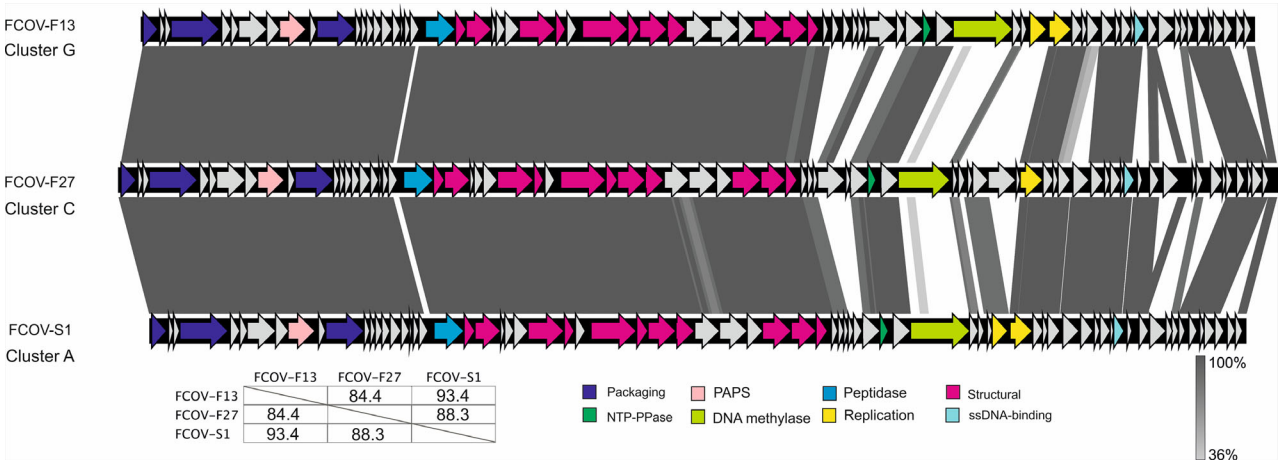


Fig. 4. Genomic comparison of representative phages of the Clusters A (FCOV-S1), C (FCOV-F27) and G (FCOV-F13). Arrows in the image indicate locations and orientations of ORFs. ORFs with unknown function are marked with grey, ORFs with putative function are marked with colours indicating the putative function as marked in the bottom. PAPS stands for 3′-phosphoadenosine-5′-phosphosulfate. The nucleotide level identity between the genomes is indicated in the box bottom left. Grey regions between the genomes indicate the level of identity (see the legend).

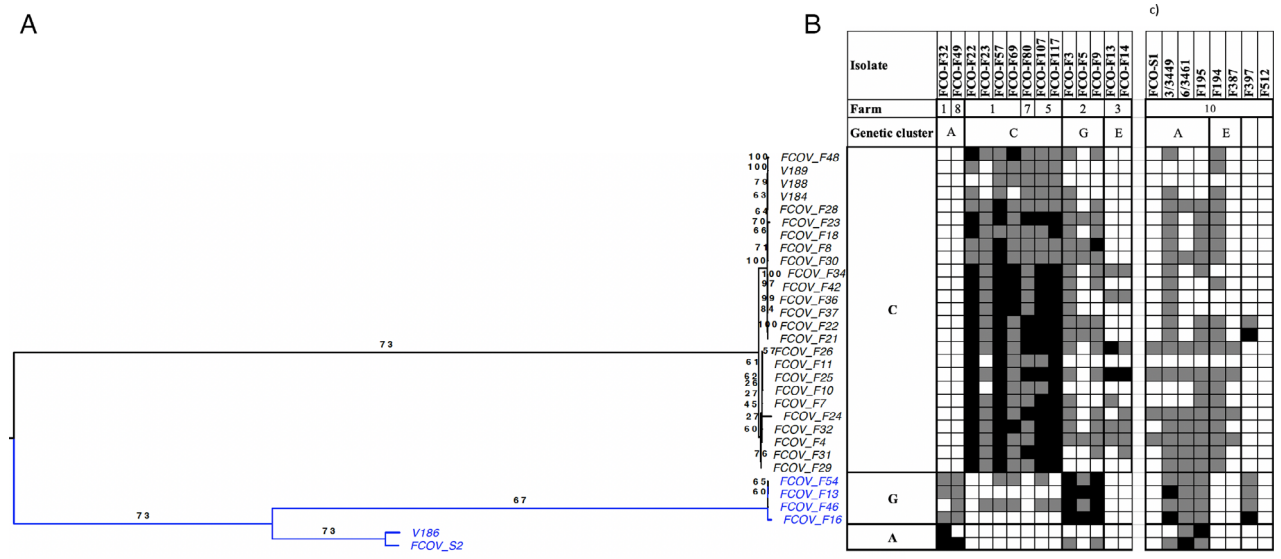


Fig. 5. Genetic distance versus host range of sequenced phages.
A. Genome BLAST Distance Phylogeny (GBDP) tree. The numbers above branches are GBDP pseudo-bootstrap support values from 100 replications. Host range of sequenced bacteriophages against representative bacterial strains isolated in (B) Finland, and (C) in Sweden. Black square indicates phage infection, grey square growth inhibition and white square bacterial resistance. Complete host range is provided in Supplementary materials.

non-synonymous change in a putative tail protein (16 932 bp). However, these phages (FCOV-F45 and FCOV-F46) differ from each other in their host range, with FCOV-F46 able to infect also eight bacterial isolates in the genetic group C. The previously isolated phage FCL-2 (from 2008) differs from other Cluster G phages, especially in the putative tail tape measure protein, although it has been isolated from the same fish farm than FCOV-F45 and FCOV-F46 (Fig. 6B). In this comparison genetic changes were also seen in the ORF coding for speculated tail fibre proteins (hypothetical proteins around 25 kbp, similarly to Cluster C phages, see below).

Cluster C-phages were isolated most frequently, allowing higher molecular resolution of phage genomes in this group. Although the genomes were highly identical (Fig. 5; Table 1), comparisons revealed seven genetically variable areas, forming five genetic variants (GV 1–5)

where nucleotide differences leading to amino acid change occurred (Fig. 6C). The variation and similarity were not dependent on bacterial strain used in phage isolation (Table 1).

Most of the genetic variance occurred in the hypothetical proteins, but in GV2 a putative tail protein was included (ORF27). In addition, GV4 includes the speculated tail fibre proteins in the 25 kbp area. Here, amino acid-level differences in ORFs 35 and 36 between FCOV-F25 and FCOV-F28 might explain differences in host range (Fig. 6D, Supplementary Table S4).

Discussion

In this study, we examined phenotypic and genetic characteristics of 126 *F. columnare* isolates and 63 phages from fish farms in Finland and Sweden. Bacterial isolates represented two previously characterized genomovars (I, I/II, based on 16S rDNA, LaFrentz *et al*., 2014) and four genetic groups (A, C, G, E, Suomalainen *et al*., 2006). Bacteria belonging to genetic groups C and E had the highest virulence in rainbow trout, but also other genetic groups caused high mortality. The isolated phages were all tailed *Myoviridae* dsDNA phages, and were genetically similar compared to previously described *F. columnare* phages (Laanto *et al*., 2017). The phages clustered into specific units of infection based on the bacterial genomic groups, with a few exceptions of phages able to cross infect to other bacterial groups (see below). Importantly, the isolated phages were able to impair the growth of the virulent bacteria, suggesting potential to be used in phage therapy against columnaris disease.

Previous studies have shown that genetically different *F. columnare* strains co-occur at fish farms (Ashrafi *et al*., 2015; Sundberg *et al*., 2016). Our results confirm the presence of previously isolated genetic groups in Finnish fish farms, suggesting that these bacterial populations have been circulating at farms during the last decades (Suomalainen *et al*., 2006; Ashrafi *et al*., 2015). Furthermore, the intensive aquaculture production in salmonid fish species in the Nordic countries may also select for certain host-associated *F. columnare* strains (LaFrentz *et al*., 2018). Due to the convenient distance from the laboratory, Farm 1 was sampled the most frequently, and thus the majority of the isolated bacteria (51) and phages (30) were obtained from this farm, likely explaining the higher diversity obtained here. Nevertheless, the sampling showed presence of virulent *F. columnare* genetic groups (C and E) at almost all farms in Finland (Fig. 1), and also in Sweden (exact farm locations for Swedish farms are not known). However, the Swedish *F. columnare* population was genetically more diverse with more genetic group isolated, some of which have not been found in Finland. Yet, the virulence of Swedish strains was lower than Finnish ones. To our knowledge, *F. columnare* isolates from Sweden have not been characterized previously.

We isolated phages infecting *F. columnare* from six farms in Finland and from one sample collected from Sweden. Interestingly, phages were also isolated when their bacterial hosts were not, suggesting that phages can be useful indicators of pathogen diversity during epidemics. Phage occurrence as a proxy for pathogen presence has been used also elsewhere in aquatic environments (Jofre *et al*., 2016; McMinn *et al*., 2017; Farkas *et al*., 2019).

Phage infection patterns clustered according to *F. columnare* genetic groups, although there are a few exceptions. For example, no phages were isolated using group E bacteria as isolation hosts. However, a few phages (FCOV-F25 - F27) had the ability to infect group E bacteria (FCO-F13 and FCO-14) in addition to their isolation host (C). Despite this, these phages did not genetically differ from some of the other C-phages (indicated by letter b in Table 1, e.g. FCOV-F1 – F7), which could only infect C type bacteria. Similarly, phage FCOV-F46 clustering to group G had the ability to infect a few group C bacteria, although sharing high genetical similarity with FCOV-F45 infecting only G bacteria. A previous study with Bacteroidetes phages showed similar results, as serial passage of phages in bacterial hosts can result in changes in host range, even without detectable genetic changes (Holmfeldt *et al*., 2016). It is also possible that epigenetic modifications play a role in host range (Hattman, 2009) but that remains to be verified in these phages. However, some differences in host range were associated with clear genetic changes. FCOV-F28 was able to infect genetic groups C and G, and comparison to FCOV-F25 revealed several non-synonymous nucleotide changes in the previously speculated tail fibre genes, ORFs35 and 36 (Laanto *et al*., 2017). In our previous study these same ORFs accumulated several mutations during a co-culture of *F. columnare* and phage FCV-1, leading to change in host range (Laanto *et al*., 2020). Previous data with other species also suggest point mutations in tail fibre proteins increase phage infectivity (Uchiyama *et al*., 2011; Boon *et al*., 2020).

Although the increased accumulation of phage genomic and metagenomics data has revealed their enormous genetic diversity both on global and local scales (Salmond and Fineran, 2015), specific phages infecting specific hosts have been isolated across large geographical areas (Kellogg *et al*., 1995; Wolf *et al*., 2003; Sonnenschein *et al*., 2017), suggesting that some groups of closely related phages may have a worldwide distribution. An important finding in this study is that similar phages exist at different fish farms and countries, although some small genetic differences occur. In

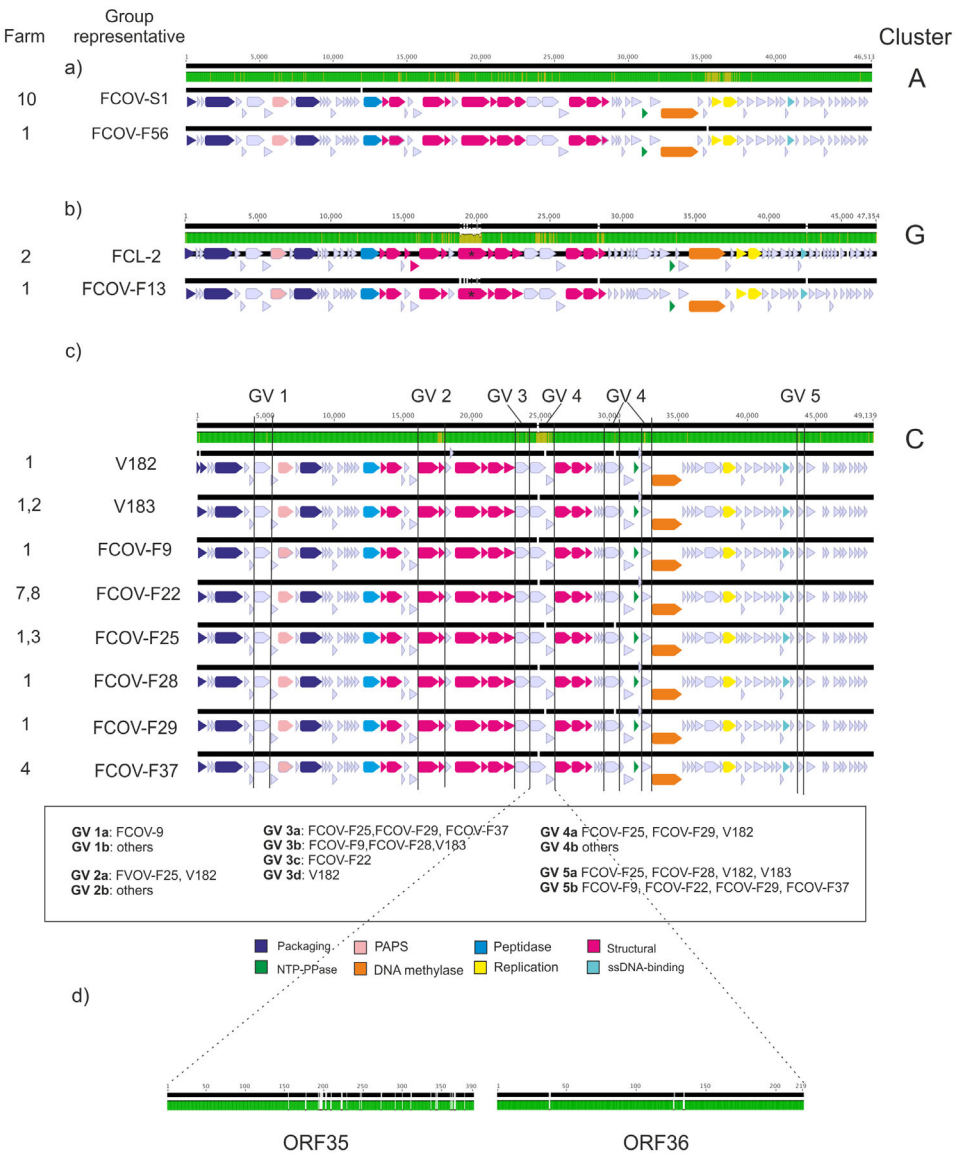


Fig. 6. Nucleotide alignments of genetically different phages infecting *Flavobacterium columnare*.
A. Cluster A (FVOV-S1 from Sweden and FCOV-F56 from Finland), (B) Cluster G (FCL-2 isolated in 2008), and (C) Cluster C phages (previously analysed V182 as a reference), and (D) consensus of the amino acid alignments of ORF 35 and ORF 36 (speculated tail fibre proteins) between FCOV-F25 and FCOV-F28, which may be linked with differences in host range. In the genome consensus (on top of each alignment) green colour indicates 100% identical sequence and yellow >30%. GV indicates genetically variable area (GV 1–5) where the differences localize. ORFs with unknown function are marked with grey, ORFs with putative function are marked with colours indicating the putative function as marked in the bottom. Letters after GV (a, b, c or d) indicate phages, which are similar in certain GV area but differ in other areas. PAPS = 3′-phosphoadenosine-5′-phosphosulfate. Asterisk marks the tape tail measure protein in case of cluster G phages. Phage V182 isolated in 2014 was included in alignments as a reference to the last time point of the phage genome evolution dataset published earlier (Laanto *et al*., 2017).

addition, the isolated phage genomes were highly similar to the previously described phages in our dataset, as also reported earlier with phages infecting the genetic group C bacteria (Laanto *et al*., 2017). Comparable findings have been reported in aquaculture-related phage-bacterium systems, for example with pathogens *F. psychrophilum* and *Vibrio anguillarum* (Castillo and Middelboe, 2016; Kalatzis *et al*., 2017), where genetically similar phages were widely distributed across large temporal and spatial scales. Furthermore, genetically similar phages have been isolated also in other types of phage-bacterium systems, such as *Salmonella*, *Synechococcus* and *Escherichia* (Gregory *et al*., 2016; Michniewski *et al*., 2019; Thanki *et al*., 2019).

Phages infecting flavobacterial species are known to regulate the genetic and phenotypic diversity of their bacterial hosts. As a co-evolutionary response, this should select for diversity also in the phage population. Yet, *F. columnare* phages with 100% nucleotide identity were isolated from different fish farms and using enrichment hosts isolated in different years (Table 1). For example, Cluster C phages FCOV-F29 (isolated using B537 host from 2013) and FCOV-F31 (isolated with FCO-F2 from 2017) are identical, as are also e.g. V183 (B245 from 2009) and FCOV-F48 (B537), and A-phages FCOV-S1 and FCOV-S2 (isolated using B534 from 2013 and B067 from 2007 respectively). A possible explanation for the low genetic diversity among *F. columnare* phages includes a potential for transfer of phages and bacteria between farms with the fish stocks or via water sources. This could explain the 100% similarity of phages from Farms 5 and 7, which are located close to each

other and share the same source of water. Similarly, the shared water source and transfer of fish fry from Farm 1 to Farm 3 may contribute to the isolation of identical phages at these farms. On the other hand, A-phages originating from Finland and Sweden shared 99.7% identity despite the geographical distance and different isolation hosts. Therefore, the low genetic diversity and the high costs of phage resistance (Laanto *et al*., 2012; Laanto *et al*., 2014) in *F. columnare* may select for a low genetic diversity of the phages. Also, the use of antibiotics may play a central role in maintaining the low genetic diversity of the host bacteria in aquaculture. Application of antibiotic treatment to control disease epidemics rapidly impact the bacterial population size, leaving less possibilities for phage–bacterium interaction at short (within-season) time span.

When looking at the temporal differences of phage isolates our results are in accordance with previous findings: the phages isolated in most recently (2017) had broader host range than the previously isolated phages [FCV-1 (2008), V183-V189 (2015), Cluster C, Supplementary Table S4] from the same fish farm, which were, mostly able to infect isolates from earlier time points. This indicates a coevolutionary response to evolution of bacterial resistance towards previously isolated phages. At the same time, the bacterial isolates were susceptible to contemporary phages, which may have evolved to overcome the resistance mechanisms of the hosts. Similar results from environmental data have been derived also in other phage–bacterium systems (Koskella and Parr, 2015). However, the low diversity both in the phage and host *F. columnare* populations seems to restrict the genetic changes to small areas in the phage genomes. Our genetic data indicate changes at the end of the genome which might explain why host range between 2015 and 2017 phages differs in Cluster C phages. Phage V182 (isolated in 2014, Farm 1) used as a reference in genome alignment (Fig. 6C) was distinct to 2015 phages from the same farm. V182 was the most recent phage isolate used in our previous study on phage–bacterium coevolution during 2007–2014 at Farm 1 (Laanto *et al*., 2017). The genetic comparisons thus suggest that bacterial resistance mechanisms cause directional selection in the phage genome over long time spans. However, this phenomenon was not detectable in isolates obtained within one outbreak season at Farm 1, as phages isolated during three time points were identical (Table 1). This might have been caused by other factors at farms (e.g. the use of antibiotics) or in the natural waters. Similar results can be derived from comparison of Cluster G phages isolated from farm 2. FCL-2 (isolated in 2008) was genetically different to phages isolated in 2017, which can also be seen in differences in the host range.

In relation to the genetic similarity, another main finding of our data is that phages isolated from certain farms were able to infect bacterial hosts from other farms. This indicates that the aquatic farming environment (probably together with fish transfers and other reasons mentioned above) does not form isolation barriers between geographic locations, which would lead to locally adapting phage populations. At the nucleotide level, some farm-specific differences between the phages were observed, but this did not impact their host range. In a broader perspective, none of the phages was able to infect *F. columnare* strains isolated from Central Europe, and only one infected isolate from the USA. Also in Swedish hosts the infectivity was limited to few strains. It is therefore likely that phage–bacterium coevolution has different trajectories in different geographic areas. This remains to be demonstrated until phage–*F. columnare* interactions have been characterized outside Nordic countries. In addition, further genomic analysis of the host bacteria could reveal the resistance mechanisms in bacteria, explaining the phage host range between farms and in time.

Because of specific fatal effect against bacteria, lytic phages have been considered and studied as antimicrobial agents against bacterial infections instead of antibiotics (Lin *et al*., 2017; Ooi *et al*., 2019). Here we characterized phage diversity against *F. columnare* and its four different genetic groups. In this study, we increased the collection of the isolated and characterized *F. columnare* strains and phages from different fish farms. Our findings suggest phages capable of infecting virulent *F. columnare* strains are present at fish farms and these phages could be used as potential antimicrobial agents in future applications.

Experimental procedures

Isolation of bacteria

Samples from >10 fish farms were collected between June and August 2017 during columnaris disease outbreaks (Fig. 1 and Supplementary Table S1; exact number and locations of the Swedish fish farms are confidential and not known by the authors, so they are jointly marked as Farm 10). *Flavobacterium columnare* was isolated from water samples and directly from infected fish, using standard culture methods on Shieh agar plates supplemented with tobramycin (Song *et al*., 1988; Decostere *et al*., 1997). Water samples (1000 ml) were collected from earthen ponds, fibreglass and plastic tanks, and from the outlet water of the farms. The obtained isolates were pure cultured and stored at −80°C with 10% glycerol and 10% fetal calf serum.

Genetic characterization of bacterial strains

Bacterial genomic DNA was extracted from overnight bacterial liquid cultures with DNeasy® Blood & Tissue Kit (Qiagen) according to the manufacturer's instructions. For some bacterial isolates, a template for PCR reaction for genomovar and genetic group classification (see below) was obtained by picking one bacterial colony into 100 μl of sterile distilled water.

Flavobacterium columnare isolates were classified to genomovars with RFLP analysis of 16S rRNA gene according to LaFrentz *et al*. (2014) with some modifications. 16S rDNA was amplified by PCR (10 min at 95°C; 40 cycles of 30 s at 94°C, 45 s at 55°C and 1 min at 72°C; 10 min at 72°C) using universal primers fD1 (Weisburg *et al*., 1991) and 1500R (Hadisaputro and Wakabayashi, 1999) with 1 μM of each primer, 1× DreamTaq Green buffer, 0.2 mM dNTP mix, 0.5 U Dream Taq DNA polymerase. PCR products (10 μl) were digested overnight at 37°C with 3.3 U of HaeIII. Restriction fragments were run in 12% acrylamide gels, which were stained with ethidium bromide and visualized under UV light.

The bacteria were further classified into genetic groups by RFLP of 16S – 23S ITS region correlating with the ARISA analysis designed for *F. columnare* (Suomalainen *et al*., 2006). ITS region was amplified by PCR (2 min at 95°C; 35 cycles of 30 s at 94°C, 45 s at 52°C and 3 min at 72°C; 15 min at 72°C) using primer pair rD1f (Weisburg *et al*., 1991) and 23Sr (Borneman and Triplett, 1997) as above. PCR products (10 μl) were double-digested overnight at 37°C with 3.3 U of both HaeIII and HinfI. Restriction fragments were run and visualized as described above.

Virulence experiment

Thirty-four bacterial isolates representing all the genetic groups were selected for virulence testing on rainbow trout fry (Supplementary Table S1). Bacteria were revived from −80°C by inoculation to 5 ml of Shieh medium and cultured overnight at 25°C under constant agitation (120 rpm). Bacteria were enriched by subculturing (1:10) and incubating for 24 h. Bacterial cell density was measured as an optical density (OD, 595 nm; Multiskan FC Thermo Scientific) and colony-forming units per ml (CFU ml^{-1}) estimated based on our previously determined OD–CFU relationship.

A total of 527 rainbow trout fry (*Oncorhynchus mykiss*), average weight 1.25 g, were randomly selected and placed individually into experimental aquaria with 500 ml of pre-aerated water (25°C). For each bacterial isolate, 14–15 individual fish (20 for negative control) were infected by pipetting into each aquarium 500 μl of bacterial solution giving a final dose of 5×10^3 CFU ml^{-1}. Shieh medium was used for negative controls. Fish morbidity and symptoms were observed at 1-h intervals for 24 h. Symptomatic fish not responding to stimuli were removed from the experiment and measured. To confirm the presence/absence of the bacterium, cultivations from gills of the dead fish were made on Shieh agar supplemented with tobramycin (Decostere *et al*., 1997). At the end of the experiment, surviving fish were euthanatized with overdose of benzocaine. Mortality data were analyzed using Kaplan–Meier survival analysis in IBM-SPSS statistics 24 SPSS. High and medium virulence of individual isolates were defined by an estimated survival time of <15 h and >15 h respectively, and low virulence was when no significant difference was detected compared to the control group.

Virulence test was performed according to the Finnish Act on Use of Animal for Experimental Purpose, under permission ESAVI/3940/04.10.07/2015 granted for Lotta-Riina Sundberg by the National Animal Experiment Board at the Regional State Administrative Agency for Southern Finland.

Isolation and characterization of bacteriophages

For bacteriophage isolation, 500 ml of water sample was filtered using rapid flow filters (PES membrane, pore size 0.45 μm, Nalgene®). 5× Shieh medium was diluted to 1× using filtered water sample. 1 ml of overnight-grown bacterial host (or mixture of hosts) was added to the diluted Shieh (total volume 21 ml). In some isolations, Shieh was diluted to 0.5× Shieh supplemented with 0.1% mucin (Almeida *et al*., 2019).

Four previously isolated and characterized *F. columnare* strains (genetic group in parenthesis) were used as enrichment hosts; B185 (G), B480 (E), B534 (A) and B537 (C). In addition, strains F514 (ND3) isolated from Sweden and FCO-F2 (C) isolated from Finland were used in some of the enrichments (see Supplementary Table S1). All the strains were used both individually and as a mixture. In mixed cultures, the total bacterial cell density (at OD 595 nm) was adjusted to the same OD level as the bacterium with the lowest OD in the individual enrichments.

After incubating for 24 h at 25°C (120 rpm), 0.5 ml samples were taken from enrichment cultures, centrifuged (3 min, 8000*g*), and supernatants were collected. Presence of phages was detected using double-layer agar method. 300 μl of fresh indicator bacterial culture and 300 μl of supernatant from the enrichment culture were mixed with 3 ml of soft Shieh agar (0.5%) tempered to 46°C and poured on Shieh agar plates. When mucin was used in the isolation procedure, also the soft agar contained 0.1% mucin. After 1–2 days of incubation at

25°C, individual plaques were transferred to a 500 μl Shieh medium and subjected to three rounds of plaque purification.

Phage stocks were prepared by adding 5 ml Shieh medium on confluent (lysis of all or almost all bacterial cells) and semi-confluent (lysis of approximately half of the bacterial cells) plates. Plates were incubated in a constant shaking (90 rpm) in cold room (7°C) for approximately 12 h. The lysate was collected with a syringe, sterile-filtered (Acrodisc® Syringe Filters with Supor® Membrane, pore size 0.45 μm), and stored at 7°C for further use.

Transmission electron microscopy

Three representative phages infecting different hosts and originating from different locations were selected for TEM imaging. TEM samples were prepared from phage lysates on Cu-grids. A drop of lysate was added to the grid and after 15–30 s the grids were dried with moist filter paper (Whatman). Dried samples were incubated with phospho-tungstic acid (1% PTA, pH 7.5) for 30–60 s and dried as above. Grids were air-dried overnight and protected from light. Imaging was done with JEOL JEM-1400 with 80 kVA.

Phage host range

Host range of 71 phages (Table 2) was tested on 227 different bacterial hosts (Supplementary Tables S1, S2 and S3) using the double-layer agar method. Two microliters of each high-titre phage lysate and their 10- and 100-fold dilutions were spotted on bacterial lawns. Results were recorded after 2 days of incubation at room temperature. Phages were considered to infect the bacterium if all phage dilutions had clear spots or if individual plaques were observed. When the drop area was not clearly lysed, but bacterial growth was impaired, phages were considered to cause growth inhibition. Bacteria were considered resistant if the phage had no effect on the growth.

Interaction map of the phage–bacterium infection patterns was done with Gephi 0.9.2 (Bastian *et al*., 2009) using the Force Atlas 2 network visualization algorithm (Jacomy *et al*., 2014).

Genetic characterization of bacteriophages

High-titre phage lysates (from 10^{10} to 10^{12} PFU ml^{-1}) were used for phage DNA extraction using the zinc chloride method (Santos, 1991) with small modifications. DNase and RNase (final concentrations 1 and 10 μg ml^{-1} respectively) were added to filtered phage lysates and incubated at 37°C for 30 min. Particles were concentrated with 0.2 M filtered $ZnCl_2$ and treated with 0.8 mg ml^{-1} Proteinase K (Thermo Scientific). Finally, DNA was extracted using phenol:chloroform:isoamyl alcohol (25:24:1) (Sambrook and Russell, 2001) and precipitated with isopropanol. DNA pellets were dissolved to 10–30 μl of dH_2O. Absence of bacterial host DNA was confirmed with PCR using universal bacterial 16S rDNA primers fD1 and rD1 (Weisburg *et al*., 1991). Genome restriction analysis with EcoRI was performed to confirm phage cluster type (data not shown).

Phage genome sequencing was performed by the Institute for Molecular Medicine Finland (FIMM). Fifty six phages were selected for sequencing based on host genetic group, isolation farm and isolation method (with and without added mucin). Phage genomes were sequenced with Illumina HiSeq2500, and PE100 sequencing produced on average 800× coverage of the genomes.

Phage genome assemblies were produced by mapping the reads to reference sequence (KY979242 for genetic group C-infecting phages, and NC_027125 for genetic group G-infecting phages) using Geneious mapper (Geneious 7.1.4 and later versions, Biomatters) with the option to find structural variants, short insertions and deletions of any size (up to 1000 bp). Minimum support for structural variant discovery was two reads, and the option to detect insertion in structural variants was included. Gaps were also allowed.

Genetic group A-infecting phages were *de novo* assembled with Velvet 1.2.10 (in Geneious). Genome ends were checked from V186 and FCOV-S1 with Sanger sequencing using primers designed for *F. columnare* phage FCV-1 genome end verification (Laanto *et al*., 2017). Briefly, primer walking was done towards the end of the genome and detection of a peak in the electropherogram (artificial base added by the sequencing polymerase) and the end of sequence suggested the physical ends of the genome.

ORFs were predicted using GenemarkS (Besemer *et al*., 2001) and Glimmer (Kelley *et al*., 2012). BLASTP (Altschul *et al*., 1990) and HHPred (Söding *et al*., 2005) were employed for annotating the putative function of the ORFs. Genomes were aligned using MUSCLE (Edgar, 2004) using default settings suggested by Geneious 7.1.4 (Biomatters). Phage V182 isolate from farm 1 in 2014 was included in alignments as a reference to the last time point of the phage genome evolution dataset published earlier (Laanto *et al*., 2017). Genome comparison of the three phages (FCOV-S1, FCOV-F13 and FCOV-F27) infecting different host genetic groups was created with Easyfig (Sullivan *et al*., 2011) employing BLASTX. All versus all comparisons for genomes were done using VIRIDIC (Moraru *et al*., 2020). Putative transposases, virulence factors and antibiotic-resistant genes were manually checked using BLASTP.

Phage genetic distances were analyzed with Victor (Meier-Kolthoff and Göker, 2017). All pairwise comparisons of the nucleotide sequences were conducted using the GBDP method (Meier-Kolthoff *et al*., 2013) under settings recommended for prokaryotic viruses (Meier-Kolthoff and Göker, 2017).

The resulting intergenomic distances were used to infer a balanced minimum evolution tree with branch support via FASTME including SPR postprocessing (Lefort *et al*., 2015) for each of the formulas D0, D4 and D6 respectively. Branch support was inferred from 100 pseudo-bootstrap replicates each. Trees were rooted at the midpoint (Farris, 1972) and visualized with FigTree (Rambaut, 2016). Taxon boundaries at the species, genus and family level were estimated with the OPTSIL program (Göker *et al*., 2009), the recommended clustering thresholds (Meier-Kolthoff and Göker, 2017) and an *F* value (fraction of links required for cluster fusion) of 0.5 (Meier-Kolthoff *et al*., 2014).

Acknowledgements

This work resulted from the BONUS FLAVOPHAGE project supported by BONUS (Art 185), funded jointly by the EU and Academy of Finland and Innovation Fund Denmark. We also acknowledge funding from the Academy of Finland (grants #314939 and #321985), the Jane and Aatos Erkko Foundation and the Kone Foundation. The authors wish to thank Dr. Päivi Rintamäki (University of Oulu), Eva Jansson (National Veterinary Institute of Sweden), Mr Yrjö Lankinen and Natural research Institute Finland for donating water samples and bacterial isolates, and Ms. Hannah Kempf, Dr. Roghaieh Ashrafi and special laboratory technician Mr. Petri Papponen for technical assistance. We also thank Jean-Francois Bernardet (INRA, France), Mark McBride (UWM, USA) and Attila Karsi (Mississippi State University, USA) for providing bacterial strains for phage host range testing. L.-R.S., G.M.F.A., E.L., and the University of Jyväskylä are responsible for a patent application covering the commercial use of purified mucin for production, quantification and isolation of bacteriophages. It is titled 'Improved methods and culture media for production, quantification and isolation of bacteriophages' and was filed with the Finnish Patent and Registration Office under patent no. FI20185086 (PCT/FI2019/050073) on 31 January 2018.

References

Almeida, G.M.d.F., and Sundberg, L.R. (2020) The forgotten tale of Brazilian phage therapy. *Lancet Infect Dis* **20**: e90–e101.

Almeida, G.M.F., Laanto, E., Ashrafi, R., and Sundberg, L.R. (2019) Bacteriophage adherence to mucus mediates preventive protection against pathogenic bacteria. *MBio* **10**: 1–12.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.

Ashrafi, R., Pulkkinen, K., Sundberg, L.-R., Pekkala, N., and Ketola, T. (2015) A multilocus sequence analysis scheme for characterization of Flavobacterium columnare isolates. *BMC Microbiol* **15**: 210–243.

Bastian, M., Heymann, S., and Jacomy, M. (2009) Gephi: An Open Source Software for Exploring and Manipulating Networks. Third International AAAI Conference on Weblogs and Society Media.

Bernardet, J.F., Segers, P., Vancanneyt, M., Berthe, F., Kersters, K., and Vandamme, P. (1996) Cutting a gordian knot: emended classification and description of the genus Flavobacterium, emended description of the family Flavobacteriaceae, and proposal of Flavobacterium hydatis nom nov (basonym, Cytophaga aquatilis Strohl and Tait 1978). *Int J Syst Bacteriol* **46**: 128–148.

Besemer, J., Lomsadze, A., and Borodovsky, M. (2001) GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* **29**: 2607–2618.

Boon, M., Holtappels, D., Lood, C., van Noort, V., and Lavigne, R. (2020) Host range expansion of Pseudomonas virus LUZ7 is driven by a conserved tail fiber mutation. *PHAGE* **1**: 87–90.

Borneman, J., and Triplett, E.W. (1997) Molecular microbial diversity in soils from eastern Amazonia: evidence for unusual microorganisms and microbial population shifts associated with deforestation. *Appl Environ Microbiol* **63**: 2647–2653.

Cabello, F.C., Godfrey, H.P., Tomova, A., Ivanova, L., Dölz, H., Millanao, A., and Buschmann, A.H. (2013) Antimicrobial use in aquaculture re-examined: its relevance to antimicrobial resistance and to animal and human health. *Environ Microbiol* **15**: 1917–1942.

Castillo, D., Higuera, G., Villa, M., Middelboe, M., Dalsgaard, I., Madsen, L., and Espejo, R.T. (2012) Diversity of Flavobacterium psychrophilum and the potential use of its phages for protection against bacterial cold water disease in salmonids. *J Fish Dis* **35**: 193–201.

Castillo, D., and Middelboe, M. (2016) Genomic diversity of bacteriophages infecting the fish pathogen Flavobacterium psychrophilum. *FEMS Microbiol Lett* **363**: 272–276.

Declercq, A.M., Haesebrouck, F., den Broeck, W., Bossier, P., and Decostere, A. (2013) Columnaris disease in fish: a review with emphasis on bacterium-host interactions. *Vet Res* **44**: 27.

Decostere, A., Haesebrouck, F., and Devriese, L.A. (1997) Shieh medium supplemented with tobramycin for selective isolation of Flavobacterium columnare (Flexibacter columnaris) from diseased fish. *J Clin Microbiol* **35**: 322–324.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.

FAO. (2014) *The State of World Fisheries and Aquaculture*. Rome, FAO: Food and Agriculture Organization of the United Nations, p. 243.

Farkas, K., Adriaenssens, E.M., Walker, D.I., McDonald, J. E., Malham, S.K., and Jones, D.L. (2019) Critical evaluation of CrAssphage as a molecular marker for human-derived wastewater contamination in the aquatic environment. *Food Environ Virol* **11**: 113–119.

Farris, J.S. (1972) Estimating phylogenetic trees from distance matrices. *Am Nat* **106**: 645–668.

García, J.C., LaFrentz, B.R., Waldbieser, G.C., Wong, F.S., and Chang, S.F. (2018) Characterization of atypical *Flavobacterium columnare* and identification of a new genomovar. *J Fish Dis* **41**: 1159–1164.

Göker, M., García-Blázquez, G., Voglmayr, H., Tellería, M. T., and Martín, M.P. (2009) Molecular taxonomy of phytopathogenic fungi: a case study in Peronospora. *PLoS One* **4**: e6319.

Gregory, A.C., Solonenko, S.A., Ignacio-Espinoza, J.C., LaButti, K., Copeland, A., Sudek, S., *et al*. (2016) Genomic differentiation among wild cyanophages despite widespread horizontal gene transfer. *BMC Genomics* **17**: 930.

Hattman, S. (2009) The first recognized epigenetic signal: DNA glucosylation of T-even bacteriopages. *Epigenetics* **4**: 150–151.

Holmfeldt, K., Solonenko, N., Howard-Varona, C., Moreno, M., Malmstrom, R.R., Blow, M.J., and Sullivan, M.B. (2016) Large-scale maps of variable infection efficiencies in aquatic *Bacteroidetes*phage-host model systems. *Environ Microbiol* **18**: 3949–3961.

Jacomy, M., Venturini, T., Heymann, S., and Bastian, M. (2014) ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One* **9**: e98679.

Jofre, J., Lucena, F., Blanch, A.R., and Muniesa, M. (2016) Coliphages as model organisms in the characterization and management of water resources. *Water (Switzerland)* **8**: 199.

Kalatzis, P.G., Rørbo, N., Castillo, D., Mauritzen, J.J., Jørgensen, J., Kokkari, C., *et al*. (2017) Stumbling across the same phage: comparative genomics of widespread temperate phages infecting the fish pathogen vibrio anguillarum. *Viruses* **9**: 122.

Karvonen, A., Rintamäki, P., Jokela, J., and Valtonen, E.T. (2010) Increasing water temperature and disease risks in aquatic systems: climate change increases the risk of some, but not all, diseases. *Int J Parasitol* **40**: 1483–1488.

Kelley, D.R., Liu, B., Delcher, A.L., Pop, M., and Salzberg, S.L. (2012) Gene prediction with Glimmer for metagenomic sequences augmented by classification and clustering. *Nucleic Acids Res* **40**: e9.

Kellogg, C.A., Rose, J.B., Jiang, S.C., and Thurmond, J.M. (1995) Genetic diversity of related vibriophages isolated from marine environments around Florida and Hawaii, USA. *Mar Ecol Prog Ser* **120**: 89–98.

Kortright, K.E., Chan, B.K., Koff, J.L., and Turner, P.E. (2019) Phage therapy: a renewed approach to combat antibiotic-resistant bacteria. *Cell Host Microbe* **25**: 219–232.

Koskella, B., and Parr, N. (2015) The evolution of bacterial resistance against bacteriophages in the horse chestnut phyllosphere is general across both space and time. *Philos Trans R Soc B Biol Sci* **370**: 20140297.

Laanto, E., Bamford, J.K.H., Laakso, J., and Sundberg, L.-R. (2012) Phage-driven loss of virulence in a fish pathogenic bacterium. *PLoS One* **7**: e53157.

Laanto, E., Bamford, J.K.H., Ravantti, J.J., and Sundberg, L.-R. (2015) The use of phage FCL-2 as an alternative to chemotherapy against columnaris disease in aquaculture. *Front Microbiol* **6**: 829.

Laanto, E., Hoikkala, V., Ravantti, J., and Sundberg, L.-R. (2017) Long-term genomic coevolution of host-parasite interaction in the natural environment. *Nat Commun* **8**: 111.

Laanto, E., Mäkelä, K., Hoikkala, V., Ravantti, J.J., and Sundberg, L.R. (2020) Adapting a phage to combat phage resistance. *Antibiotics* **9**: 1–12.

Laanto, E., penttinen, R.K., Bamford, J.K.H., and Sundberg, L.-R. (2014) Comparing the different morphotypes of a fish pathogen – implications for key virulence factors in Flavobacterium columnare. *BMC Microbiol* **14**: 170.

Laanto, E., Sundberg, L.-R., and Bamford, J.K.H. (2011) Phage specificity of the freshwater fish pathogen Flavobacterium columnare. *Appl Environ Microbiol* **77**: 7868–7872.

LaFrentz, B.R., García, J.C., Dong, H.T., Waldbieser, G.C., Rodkhum, C., Wong, F.S., and Chang, S.F. (2017) Optimized reverse primer for 16S-RFLP analysis and genomovar assignment of Flavobacterium columnare. *J Fish Dis* **40**: 1103–1108.

LaFrentz, B.R., García, J.C., Waldbieser, G.C., Evenhuis, J. P., Loch, T.P., Liles, M.R., *et al*. (2018) Identification of four distinct phylogenetic groups in Flavobacterium columnare with fish host associations. *Front Microbiol* **9**: 1–13.

LaFrentz, B.R., Waldbieser, G.C., Welch, T.J., and Shoemaker, C.A. (2014) Intragenomic heterogeneity in the 16S rRNA genes of *Flavobacterium columnare*and standard protocol for genomovar assignment. *J Fish Dis* **37**: 657–669.

Lefort, V., Desper, R., and Gascuel, O. (2015) FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol Biol Evol* **32**: 2798–2800.

Lin, D.M., Koskella, B., and Lin, H.C. (2017) Phage therapy: an alternative to antibiotics in the age of multi-drug resistance. *World J Gastrointest Pharmacol Ther* **8**: 162–173.

Lone, A., Anany, H., Hakeem, M., Aguis, L., Avdjian, A.C., Bouget, M., *et al*. (2016) Development of prototypes of bioactive packaging materials based on immobilized bacteriophages for control of growth of bacterial pathogens in foods. *Int J Food Microbiol* **217**: 49–58.

McMinn, B.R., Ashbolt, N.J., and Korajkic, A. (2017) Bacteriophages as indicators of faecal pollution and enteric virus removal. *Lett Appl Microbiol* **65**: 11–26.

Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.P., and Göker, M. (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**: 60.

Meier-Kolthoff, J.P., and Göker, M. (2017) VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics* **33**: 3396–3404.

Meier-Kolthoff, J.P., Hahnke, R.L., Petersen, J., Scheuner, C., Michael, V., Fiebig, A., *et al*. (2014) Complete genome sequence of DSM 30083T, the type strain (U5/41T) of Escherichia coli, and a proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci* **9**: 2.

Michniewski, S., Redgwell, T., Grigonyte, A., Rihtman, B., Aguilo-Ferretjans, M., Christie-Oleza, J., *et al*. (2019) Riding the wave of genomics to investigate aquatic coliphage diversity and activity. *Environ Microbiol* **21**: 2112–2128.

Moraru, C., Varsani, A., and Kropinski, A.M. (2020) VIRIDIC – a novel tool to calculate the intergenomic similarities of prokaryote-infecting viruses. *Viruses* **12**: 1286.

Moye, Z.D., Woolston, J., and Sulakvelidze, A. (2018) Bacteriophage applications for food production and processing. *Viruses* **10**: 2015.

Oidtmann, B.C., Crane, C.N., Thrush, M.A., Hill, B.J., and Peeler, E.J. (2011) Ranking freshwater fish farms for the risk of pathogen introduction and spread. *Prev Vet Med* **102**: 329–340.

Ooi, M.L., Drilling, A.J., Morales, S., Fong, S., Moraitis, S., Maclas-Valle, L., *et al*. (2019) Safety and tolerability of bacteriophage therapy for chronic rhinosinusitis due to staphylococcus aureus. *JAMA Otolaryngol - Head Neck Surg* **145**: 723–729.

Pulkkinen, K., Suomalainen, L.-R., Read, A.F., Ebert, D., Rintamaki, P., and Valtonen, E.T. (2010) Intensive fish farming and the evolution of pathogen virulence: the case of columnaris disease in Finland. *Proc R Soc B Biol Sci* **277**: 593–600.

Rambaut, A. (2016) *FigTree. version 1.4.3*. Edinburgh, Scotland: Institute of Evolutionary Biology, University of Edinburgh.

Salmond, G.P.C., and Fineran, P.C. (2015) A century of the phage: past, present and future. *Nat Rev Microbiol* **13**: 777–786.

Sambrook, J., and Russell, D.W. (2001, 2001) *Molecular Cloning: A Laboratory Manual (3-volume set)*. Cold Spring Harbour, New York: Cold Spring Harbor Laboratory Press, p. 2100.

Santos, M.A. (1991) An improved method for the small scale preparation of bacteriophage DNA based on phage precipitation by zinc chloride. *Nucleic Acids Res* **19**: 5442.

Silva, Y.J., Costa, L., Pereira, C., Mateus, C., Cunha, Â., Calado, R., *et al*. (2014) Phage therapy as an approach to prevent Vibrio anguillarum infections in fish larvae production. *PLoS One* **9**: e114197.

Söding, J., Biegert, A., and Lupas, A.N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**: W244–W248.

Song, Y.L., Fryer, J.L., and Rohovec, J.S. (1988) Comparison of six media for the cultivation of Flexibacter columnaris. *Fish Pathol* **23**: 81–94.

Sonnenschein, E.C., Nielsen, K.F., D'Alvise, P., Porsby, C. H., Melchiorsen, J., Heilmann, J., *et al*. (2017) Global occurrence and heterogeneity of the Roseobacter-clade species Ruegeria mobilis. *ISME J* **11**: 569–583.

Sulakvelidze, A. (2013) Using lytic bacteriophages to eliminate or significantly reduce contamination of food by foodborne bacterial pathogens. *J Sci Food Agric* **93**: 3137–3146.

Sullivan, M.J., Petty, N.K., and Beatson, S.A. (2011) Easyfig: A genome comparison visualizer. *Bioinformatics* **27**: 1009–1010.

Sundberg, L.-R., Ketola, T., Laanto, E., Kinnula, H., Bamford, J.K.H., Penttinen, R., and Mappes, J. (2016) Intensive aquaculture selects for increased virulence and interference competition in bacteria. *Proc R Soc B Biol Sci* **283**: 20153069.

Suomalainen, L.-R., Kunttu, H., Valtonen, E.T., Hirvela-Koski, V., and Tiirola, M. (2006) Molecular diversity and growth features of Flavobacterium columnare strains isolated in Finland. *Dis Aquat Organ* **70**: 55–61.

Suomalainen, L.R., Tiirola, M.A., and Valtonen, E.T. (2005) Influence of rearing conditions on Flavobacterium columnare infection of rainbow trout, Oncorhynchus mykiss (Walbaum). *J Fish Dis* **28**: 271–277.

Tamminen, M., Karkman, A., Lõhmus, A., Muziasari, W.I., Takasu, H., Wada, S., *et al*. (2011) Tetracycline resistance genes persist at aquaculture farms in the absence of selection pressure. *Environ Sci Technol* **45**: 386–391.

Tan, D., Gram, L., and Middelboe, M. (2014) Vibriophages and their interactions with the fish pathogen Vibrio anguillarum. *Appl Environ Microbiol* **80**: 3128–3140.

Thanki, A.M., Brown, N., Millard, A.D., and Clokie, M.R.J. (2019) Genomic characterization of jumbo *salmonella* phages that effectively target United Kingdom pig-associated *Salmonella* serotypes. *Front Microbiol* **10**: 1491.

Hadisaputro, T., and Wakabayashi, H. (1999) Genotypic diversity of strains of Flavobacterium columnare from diseased fishes. *Fish Pathol* **34**: 65–71.

Uchiyama, J., Takemura, I., Satoh, M., Kato, S.-i., Ujihara, T., Akechi, K., *et al*. (2011) Improved adsorption of an enterococcus faecalis bacteriophage φEF24C with a spontaneous point mutation. *PLoS One* **6**: e26648.

Watts, G. (2017) Phage therapy: revival of the bygone antimicrobial. *Lancet* **390**: 2539–2540.

Watts, J.E.M., Schreier, H.J., Lanska, L., and Hale, M.S. (2017) The rising tide of antimicrobial resistance in aquaculture: sources, sinks and solutions. *Mar Drugs* **15**: 158.

Weisburg, W.G., Barns, S.M., Pelletier, D.A., and Lane, D.J. (1991) 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* **173**: 697–703.

Wolf, A., Wiese, J., Jost, G., and Witzel, K.P. (2003) Wide geographic distribution of bacteriophages that lyse the same indigenous freshwater isolate (Sphingomonas sp. strain B18). *Appl Environ Microbiol* **69**: 2395–2398.

Yang, J., Wang, C., Shu, C., Liu, L., Geng, J., Hu, S., and Feng, J. (2013) Marine sediment Bacteria Harbor antibiotic resistance genes highly similar to those found in human pathogens. *Microb Ecol* **65**: 975–981.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1: Supporting Information.

Appendix S2: Supporting Information.

Supplementary Fig. S1. Transmission electron microscopy (TEM) images of representative F. columnare phages infecting different genetic groups of the host (A, C and G). A). FCOV-S1, B). FCOV-F13 and C). FCOV-F27. Scale bar in A) and C) is 200 nm and B) 500 nm.

Supplementary Table S1. Bacterial strains isolated and/or characterized in this study. Swedish fish farms are confidential and not known by the authors, so they are jointly marked as Farm 10.

Supplementary Table S2. Previously isolated *F. columnare* bacterial strains used in host range studies. *F. columnare* strains from H to B533 have been previously characterized (Ashrafi *et al*., 2015). Strains marked with asterisks were used in phage isolation. Strains collected from farms 5-9 were isolated and kindly donated by Dr. Päivi Rintamäki.

Supplementary Table S3. Previously isolated *F. columnare* strains from USA and France, and other bacterial species used in host range studies.

Supplementary Table S4. Host range of phages in *F. columnare* strains isolated in this study, in previously isolated *F. columnare* strains, and in other flavobacterial species. Black square indicates infection (clear lysis in three consequent phage dilutions), grey square indicates growth inhibition, and white square no effect (i.e. bacterial resistance). Each column represents a phage isolate and each row represents a host isolate.

Supplementary Table S5. List of open reading frames (ORFs) that displayed differences between phage genomes in phages infecting genetic group C, A and G hosts.

Supplementary Fig. S2. A heatmap of all *Flavobacterium columnare* phage genomes in this study (including also previously described phages FCL-2 and V182) generated with VIRIDIC incorporating intergenomic similarity values (right) and alignment indicators (left). Numbers on the right half represent the similarity values for each genome pair and in the left three values represent (from top to bottom): aligned fraction for the genome that is found in this row (top), genome length ratio for the genomes found in this pair (middle) and aligned fraction for the genome found in this column (bottom). On the right, the darker the colours the more closely-related the genomes. On the left the darker colours indicate low values and white colours indicate genome pairs with higher similarity values.