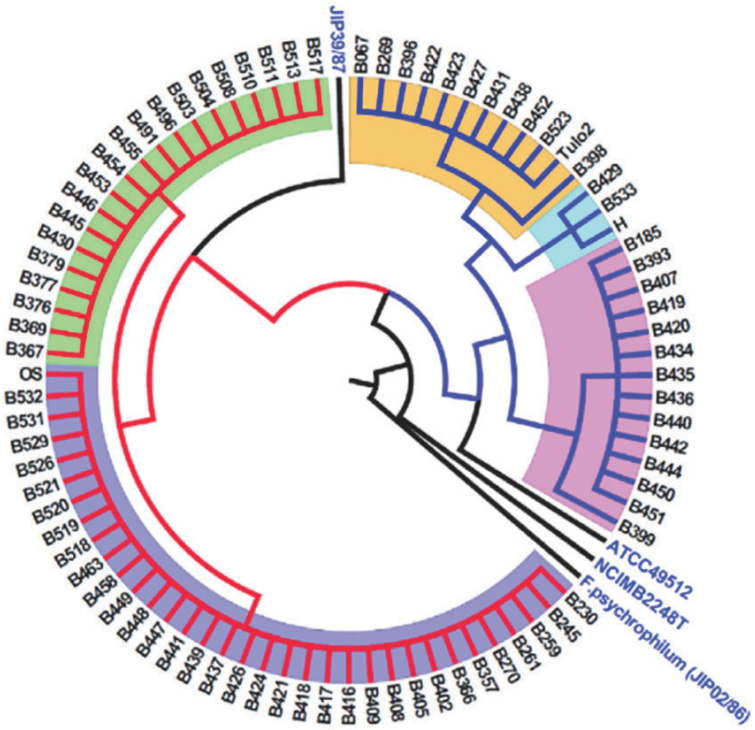
JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE
328

Roghaieh Ashrafi

Exploring Evolutionary Responses to Increasing Temperature in an Environmental Opportunistic Pathogen



JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE 328

Roghaieh Ashrafi

Exploring Evolutionary Responses to Increasing Temperature in an Environmental Opportunistic Pathogen

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Ambiotica-rakennuksen salissa YAA303,
huhtikuun 7. päivänä 2017 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Mathematics and Science of the University of Jyväskylä,
in building Ambiotica, hall YAA303, on April 7, 2017 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2017

Exploring Evolutionary Responses to Increasing Temperature in an Environmental Opportunistic Pathogen

JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE 328

Roghaieh Ashrafi

Exploring Evolutionary Responses to Increasing Temperature in an Environmental Opportunistic Pathogen

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2017

Editors
Jari Haimi
Department of Biological and Environmental Science, University of Jyväskylä
Pekka Olsbo, Ville Korkiakangas
Publishing Unit, University Library of Jyväskylä

Jyväskylä Studies in Biological and Environmental Science

Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-7007-9

URN:ISBN:978-951-39-7007-9
ISBN 978-951-39-7007-9 (PDF)

ISBN 978-951-39-7006-2 (nid.)
ISSN 1456-9701

Jyväskylä University Printing House, Jyväskylä 2017

ABSTRACT

Ashrafi, Roghaieh
Exploring evolutionary responses to increasing temperature in an environmental opportunistic pathogen
Jyväskylä: University of Jyväskylä, 2017, 45 p.
(Jyväskylä Studies in Biological and Environmental Science
ISSN 1456-9701; 328)
ISBN 978-951-39-7006-2 (nid.)
ISBN 978-951-39-7007-9 (PDF)
Yhteenveto: Lämpötilan vaikutukset ympäristössä kasvavan kalatautibakteerin evoluutioon.
Diss.

Environmentally growing opportunistic pathogens can grow outside the host. This makes them very sensitive to selection pressures outside the host, which in turn, could further alter disease severity. In this thesis, I used novel techniques and analyses to investigate the effect of environmental selection pressures, in particular temperature, on thermal performance, virulence, and intraspecific interactions of *Flavobacterium columnare*, a waterborne opportunistic bacterial pathogen of fish. The phylogenetic analysis of 83 strains from eight different geographic locations in Finland over ten years (2003-2012) were carried out using the multilocus sequence analysis (MLST/MLSA). This analysis demonstrated that strains were clearly separated into two main lineages, (I and II) which can be further divided into five genotypes. Meanwhile, isolates clustered independent from the geographic origin, host, and year with a semi-clonal and epidemic population structure (where frequent recombination occurs among isolates, but time to time, one or a few fit genotypes can arise and increase rapidly to produce an epidemic clone). *F. columnare* isolates showed variation in their thermal tolerances and in virulence, with virulence increasing over the years. Virulence was negatively correlated with a broader thermal performance curve with a right-shift towards warmer temperatures. Most likely climate change driven changes in the disease epidemics of this bacterium are associated with increased length of growing season, extending the suitable period for disease outbreaks. Finally, competition outcome in liquid cultures seemed to be well predicted by growth and yield of isolated strains at two different temperatures. However, interference competition results measured in inhibition tests on solid agar were inconsistent with head-to-head competition, suggesting that benefits of interference while growing on a surface are relaxed in liquid culture.
Keywords: *Flavobacterium columnare*; multi-locus sequence analysis; thermal performance curve; virulence; selection; and high resolution melting curve.

Roghaieh Ashrafi, University of Jyväskylä, Department of Biological and Environmental Science, P.O. Box 35, FI-40014 University of Jyväskylä, Finland.

Author's address Roghaieh Ashrafi
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland
roghaieh.ashrafi@jyu.fi

Supervisors Docent Tarmo Ketola
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland

Docent Lotta-Riina Sundberg
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland

Docent Katja Pulkkinen
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland

Reviewers Dr. Freya Harrison
Life Sciences, University of Warwick
Coventry, CV4 7AL
UK

Dr. Otto Eerikki Seppälä
Ueberlandstrasse 133
8600 Dübendorf
Switzerland

Opponent Professor Anna-Liisa Laine
Department of Biosciences
PO-Box 65 (Viikinkaari 1), 5413
FI-00014 University of Helsinki
Finland

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

The thesis is based on the following original papers:

I Ashrafi R., Pulkkinen K., Sundberg L. R., Pekkala N., & Ketola T. 2015. A multi-locus sequence analysis scheme for characterization of *Flavobacterium columnare* isolates. *BMC Microbiology*, 15(1): 243.

II Ashrafi R., Bruneaux M., Sundberg L. R., Pulkkinen K., Valkonen J. & Ketola T. 2017. Thermal tolerance at high temperature is linked with reduced virulence in a fish pathogen. Submitted manuscript.

III Ashrafi R., Bruneaux M., Sundberg L. R., Pulkkinen K. & Ketola T. 2017. Application of high resolution melting assay (HRM) to study temperature-dependent intraspecific competition in a pathogenic bacterium. Submitted manuscript.

The table shows the contributions to the original papers.

	I	**II**	**III**
Original idea	RA, TK, LS	RA, TK, LS	RA, TK
Experimental work	RA, NP	RA, LS, KP, TK	RA
Statistical analysis	RA, TK	RA, TK, MB, JV	RA, TK, MB
Manuscript	RA, LS, KP, NP, TK	RA, TK, LS, KP, MB, JV	RA, TK, LS, KP, MB

RA = Roghaieh Ashrafi, LS = Lotta-Riina Sundberg, KP = Katja Pulkkinen, NP = Nina Pekkala, MB = Matthieu Bruneaux, JV = Janne Valkonen, TK = Tarmo Ketola

1 INTRODUCTION

1.1 Opportunistic lifestyle

The term "pathogen" was introduced in 1880 by Robert Koch and can be defined as an infectious agent, such as a virus, bacterium, or fungus that causes diseases in other organisms (hosts). In terms of life style, bacterial pathogens have been classified into two broadly defined categories. Obligate pathogens require a host to grow and reproduce. However, opportunistic facultative pathogens can live either as a parasite or, in different conditions, as a nonpathogen, with the ability to survive without a host (van Baarlen *et al.* 2007, Brown *et al.* 2012). Opportunist infection is often thought of as being associated with perturbed (i.e. impaired or weakened immune system) hosts (Von Graevenitz 1977). However, opportunistic pathogens can also infect healthy hosts, as for example in the case of *Salmonella spp.* and *Flavobacterium columnare*. Recently, Brown *et al*. (2012) defined opportunistic bacteria as pathogens that are not obligate and/or specialist pathogens of a focal host. With this new definition, opportunistic pathogens have been now divided into commensal opportunists, environmental opportunists and parasite opportunists.

Environmentally growing opportunistic pathogens are pathogens that not only utilize host resources, as an alternative reproduction strategy, but also are able to grow and reproduce outside the host independently by using outside-host resources (Brown *et al.* 2012, Merikanto *et al.* 2014). It has been shown that soil and water bodies harbour diverse pathogens many of which have been identified as opportunistic pathogens such as *Legionella pneumophila*, *Pseudomonas aeruginosa, Listeria monocytogenes*, *Vibrio cholerae*, *Serratia marcescens* and others (Friedman *et al.* 2002, Hall-Stoodley and Stoodley 2005, Freitag *et al.* 2009, Toledo-Arana *et al.* 2009, Mahlen 2011, Murugaiah 2011). In this thesis, I will solely focus on opportunistic bacterial pathogens. Compared to obligate bacteria, opportunistic bacteria have been characterized with large and versatile genomes, which allow them to thrive in diverse environments and be more resistant to the effects of antimicrobial agents (Sadikot *et al.* 2005, Aujoulat *et al.*

2012).This ability makes them highly disposed to horizontal gene transfer in environment (Tamas *et al.* 2002, Sadikot *et al.* 2005, Aujoulat *et al.* 2012). Since environmental bacteria interact with other bacteria and phages (Brussow *et al.* 2004), opportunistic pathogens living in the environment may become virulent by exchanging virulence factors and/or genes through horizontal gene transfer (Pallen and Wren 2007, Johnson and Nolan 2009, Aujoulat *et al.* 2012). Therefore, environment can play an important role in the virulence of environmentally growing opportunistic bacteria in becoming a big threat to the potential hosts of these pathogens.

1.2 Virulence evolution

Since virulence is a complex, dynamic, and changeable trait that includes host and microbial factors (Casadevall and Pirofski 2001, Gupte *et al.* 2015), a uniform definition of virulence has not yet been accomplished. In general, however, virulence is defined as the capacity of a microorganism to cause disease. The origin of virulence in environmental pathogens has been explained by accidental virulence that exists by chance and independently of a host. Alternative explanations that cannot be ruled out include cryptic pathogenesis, targeted towards hosts that have not yet been discovered or towards hosts that are now extinct. In the latter case, virulence factors are maintained, because they enable the bacteria to survive better in the environmental reservoirs (Casadevall and Pirofski 2007).

In virulence theories, the basic premise has been that genotypes that are too virulent are doomed to go extinct. Host death due to high virulence reduces the time for the pathogens to be picked up by the other hosts required for their transmission (Ebert and Herre 1996, Frank 1996). On the other hand, highly restrained pathogens could be eliminated by host immune system, or by more aggressive strains that divert more host resources away from them. Therefore, virulence will evolve to a level at which virulence and transmission is balanced to optimise the spread of the pathogen (Alizon *et al.* 2009). However, there will be no single virulence optimum if the pathogen can survive periods outside the host where host death does not affect transmission of environmental pathogens (Walther and Ewald 2004, Bull and Ebert 2008, Sundberg *et al.* 2014).

Furthermore, the virulence of environmental pathogens may have evolved for some purpose other than to provide pathogens an advantage within a host or its transmission to other hosts (coincidental theory) (Adiba *et al.* 2010, Zhang *et al.* 2014b). The coincidental theory of virulence predicts that bacterial pathogenicity could be a by-product of selection acting on that pathogen in a different ecological niche. For example, synthesis of Shiga toxins of *Escherichia coli* O157:H7 are deployed against natural predators like (*Tetrahymena pyriformis*) and are also toxic to humans (Brandl 2006, Steinberg and Levin 2007). Or, *Legionella pneumophila*, which causes Legionnaires' disease originally

evolved to resist digestion of amoebas, being unrelated to human infection (Rowbotham 1980).

Due to the pathogens' short generation time and high density (compared to the host), traits that confer a survival or replication advantage on the pathogens can mutate and evolve over the course of an infection. These evolved traits are assumed to be correlated with virulence and enhance pathogens virulence in the host. This process is called short-sighted virulence or dead-end evolution because their ability to be transmitted to a new host is reduced (Martínez 2015).

1.3 Virulence across multiple environments – trade-offs

The ability of environmental opportunistic bacteria to switch between outside-host and within-host environments in their life cycles (Brown *et al.* 2012) has led to the hypothesis that selection forces in the both inside and outside-host environment are likely to influence the evolution of pathogenicity and virulence in environmentally growing opportunists (Brown *et al.* 2012). Walther and Ewald's "sit and wait" hypothesis (2004) predicts that a long survival time without replication in the environment (durability of the pathogen) may allow the development of high virulence. Nevertheless, environmental opportunistic bacteria with outside-host growth and replication are subjected to abiotic and abiotic stresses, and must adapt to rapidly changing, harsh conditions of outside host environments (Casadevall 2008, Brown *et al.* 2012). These adaptations to one context may be maladaptive to inside host environment and *vice versa*.

Survival and growth of environmental opportunistic bacteria in the outside-host environment is often traded off with virulence (infect or replicate) in the within-host environment (Casadevall 2008, Friman *et al.* 2009, Sturm *et al.* 2011, Mikonranta *et al.* 2012) but see also (Mikonranta *et al.* 2015). For example, there are trade-offs in terms of survival that, on the one hand, may increase the likelihood of reproductive success by improving coping with adverse condition of outside host environment but, on the other hand, may impair efficient host exploitation by lowering the virulence capacity of a pathogen (II). Sundberg *et al.* (2014) showed that changing from rhizoid into a rough morphotype in *F. columnare* has trade-offs in making bacteria less virulent and unable to exploit the host, but may ensure bacterial survival under unpredictable conditions in the outside host environment. It has also been shown that once *P. aeruginosa* strains have adapted for growing in the lungs of cystic fibrosis patients are released to original environment, they will present a decreased resistance to natural phage and protist predators (Friman *et al.* 2013). Moreover, the type III secretion system 1 (TTSS-1) that acts as an important virulence factor in *Salmonella typhimurium* imposes limits to bacteria growth in a non-host environment (Sturm *et al.* 2011).

Biotic factors such as predators and natural enemies in outside host environments can indirectly affect pathogen virulence through trade-offs (Friman *et al.* 2011, Mikonranta *et al.* 2012, Zhang *et al.* 2014a). Exposing the opportunistic bacterial pathogen *Serratia marcescens* to predation by protozoan *Tetrahymena thermophila* has been shown to decrease virulence in host, which has been linked to trade-offs between the bacterial anti-predatory traits and bacterial virulence factors, such as motility or resource use efficiency (Friman *et al.* 2009). Antagonistic interactions between bacterial pathogens and their parasitic phages can also favour low bacterial virulence under natural conditions (Flyg *et al.* 1980, Heierson *et al.* 1986, Santander and Robeson 2007, Friman *et al.* 2011). For example, work by Laanto *et al.* (2012) showed that phage resistant *F. columnare* strains have lost their gliding motility, which directly correlates with a decline in bacterial virulence.

1.4 Friend or foe - temperature and virulence

Disease development is determined primarily by the interactions among three major factors: the presence of a susceptible host, the presence of a virulent pathogen, and a conductive environment (Agrios 2005). The interactions among these factors have been traditionally conceptualized in the form of a disease triangle. The disease triangle shows that the environment influences disease development by interacting with both the host and pathogen. Since bacteria exist in diverse environments, from arctic oceans to hot springs, it is not surprising that specific bacteria are adapted to grow best at different temperatures. For example, bacteria adapted to the human body grow well at body temperature (37 ◦C) but a plant pathogen may be killed at that temperature.

Bacteria constantly sense the environment and adjust their physiology accordingly. A recent review by Shapiro and Cowen (2012) proposes that almost all cellular molecules, including proteins, lipids, RNA, and DNA, can act as thermosensors that detect changes in environmental temperature and initiate relevant cellular responses. Moreover, of the most important abiotic factors, temperature is a critical environmental signal that influences the virulence of many bacterial species such as *Shigella* species (Maurelli *et al.* 1984, Hromockyj *et al.* 1992), and *Yersinia* species (Guijarro *et al.* 2015). It has been shown that virulence factors involved in motility, host degradation, secretion, antimicrobial resistance and transcriptional regulation are upregulated at higher temperatures (Kimes *et al.* 2012). For example, high temperature (37°C) has been shown to be an important factor in enhancing bacterial motility and pathogenicity indirectly in *S. marcescens* bacteria (Friman *et al.* 2011).

1.5 Problems ahead - climate change and disease

Climate change has been defined as changes in the earth's weather, including changes in temperature, wind patterns and rainfall. Global warming - a gradual increase in temperatures - is caused by the increase of particular gases, especially carbon dioxide (Patz *et al.* 2005, McMichael *et al.* 2006). Effects of climate change will not be evenly distributed, with higher latitudes warming faster than lower latitudes (Hassan *et al.* 2005). Global warming affects the interactions between parasites and their hosts, depending on the thermal performance curves of both the parasites and their hosts (Thomas and Blanford 2003). If the performance curves of the host and the parasite deviate from each other, the course of infection could be strongly influenced by temperature and, particularly, by how temperature affects the respective operative ranges of both host and parasite (Thomas and Blanford 2003, Wang *et al.* 2012). However, vulnerability to rising temperature will depend mostly on the organisms' thermal tolerance, and how close they are currently to their optimal temperature. Overall, climate warming is likely to (i) increase pathogen development and transmission; (ii) increase overwinter survival and/or (iii) change the susceptibility of hosts to pathogens (Harvell *et al.* 2002). On the other hand, climate warming will disproportionately affect pathogens with complex life cycles. Pathogens with free-living stages and ectothermic hosts are very susceptible to changes in temperature because, clearly, temperature can exhibit complex and opposing effects on different parts of the pathogens life cycle (Harvell *et al.* 2002).

1.6 Climate change and the geographic distribution of infectious diseases

As climate change progresses, populations may no longer be able to cope with new conditions, and they may follow three possible fates: migrate to track optimal conditions (i.e. climate envelop), adapt locally to the new conditions, or become extinct (Chen *et al.* 2011, Box 2012, Watling *et al.* 2012). Predicting the degree to which species will respond to climate warming requires an understanding of how temperature shapes current species distributions. To this end, we need to consider the following factors: (i) thermal tolerance, (ii) phylogenetic constraints, and (iii) presence of competitors.

i Thermal tolerance capacity

Changes in environmental temperatures will affect basic physiological functions, which are usually described by continuous nonlinear reaction norms, called thermal performance curves (TPCs) (Huey and Kingsolver 1989, Angilletta 2009). The main eco-physiological parameters that characterize TPCs

are, a minimum critical temperature (CTmin) at which performance is minimized, then performance gradually increases with temperature and reaches an optimum temperature (Topt) at which performance is maximized, finally it falls where performance is minimized again at a higher temperature (CTmax). The shape and position of thermal performance curves can vary by genetic and acclimation effects (Schulte *et al.* 2011, Geerts *et al.* 2015). Ecological and evolutionary physiologists have proposed three directions or modes for changes in TPCs: height (through a vertical shift of the curve; also called faster-slower mode), the position of the Topt (through a horizontal shift of the curve; also called hotter-colder mode), and the width (or breadth; also called generalist-specialist trade-off) of the curve (Izem and Kingsolver 2005, Angilletta 2009, Kingsolver *et al.* 2009)(see Fig. 1). These patterns of variation in TPCs can be also explored and visualized by using principal components analysis (PCA) and loading values (Izem and Kingsolver 2005).

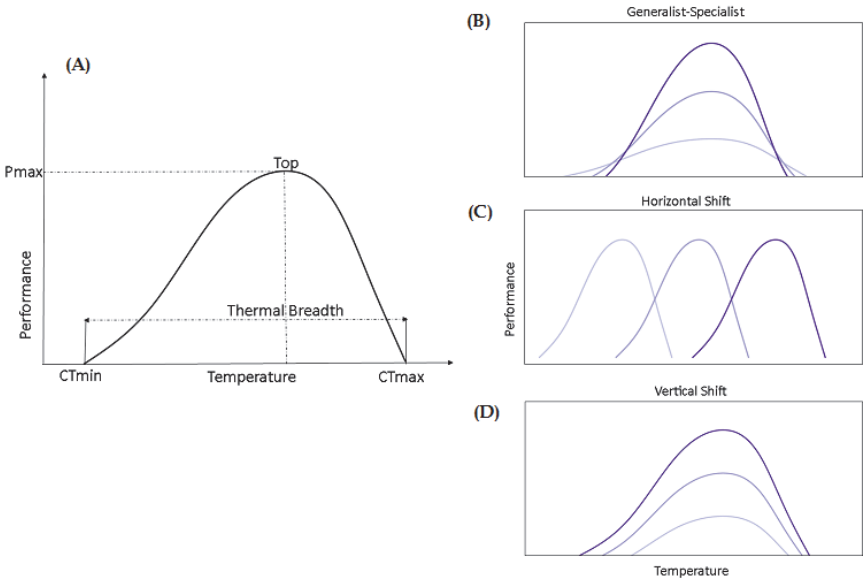


FIGURE 1 Hypothetical thermal performance curves with the commonly used descriptors of a performance curve (A). Patterns of variation for nonlinear thermal performance curves. (B) Vertical shift, (C) Horizontal shift, (D) Generalist-Specialist (Re-drawn from Kingsolver *et al.* 2004).

ii Phylogenetic constraints

Functional traits tend to be more similar in closely related species than in distant relatives (i.e. phylogenetic signal), which can result in similar responses to climate warming (Wiens and Graham 2005). However, phylogenetic signal may be attributed to similarity resulting from spatial proximity (Freckleton and Jetz 2009, Kellermann *et al.* 2012). Since phylogenetic constraints might play a significant role in determining the responses of related species/individuals to climate warming, the distinction between phylogenetically structured and environmentally driven trait variation is critical for predicting climate change impacts. Therefore, phylogeny provides useful insights into understanding the pattern of differential species' response to climate change (Willis *et al.* 2008, Hoffmann *et al.* 2013, Comte *et al.* 2014). Moreover, phylogeny can be a useful tool for generalizing predictions in less well-studied members of related clades because less well-studied members may be inferred to have similar response to environmental changes (Willis *et al.* 2008).

iii Presence of competitors

An increase in habitat suitability does not necessarily result in an increase in geographic distribution, since other factors besides climate, such as barriers to dispersal, competition, and predation, affect the realized niche (Lafferty 2009, Alexander *et al.* 2015). For example, access to climatically suitable habitat could be limited due to superior competitors (Anderson *et al.* 2002, Case *et al.* 2005, McQuillan and Rice 2015). However, multiple genotypes infecting the same host are a common in the wild. For example, mosquitoes can be infected with multiple strains of the malaria parasite (Babiker *et al.* 1999). The occurrence of multiple mixed genotype infections is expected to influence the dynamics of disease epidemics via within-host competition, which favours higher parasite virulence. deRoode *et al.* (2005) showed that within-host competition does indeed select for high parasite virulence in rodent malaria. In addition, wild *Plantago lanceolata* populations had more devastating epidemics in local populations where co-infection by multiple genotypes of the powdery mildew *Podosphaera plantaginis* was more prevalent (Susi *et al.* 2015). On the other hand, in a co-infection of the sleeping sickness parasite *Trypanosoma brucei,* the less virulent strain had an indirect positive effect on the host by suppressing the more virulent strain (Balmer *et al.* 2009). Interference competition between conspecific genotypes via toxins is a well-known mechanism in bacterial systems which exclude competitors through the release of toxins (Hibbing *et al.* 2010), and lead to decreased virulence during bacterial infection (Massey *et al.* 2004).

1.7 Flavobacterium columnare

The causative agent of columnaris disease; *F. columnare*, previously known as *Bacillus columnaris*, *Chondrococcus columnaris*, *Cytophaga columnaris* and *Flexibacter columnaris* (Bernardet and Grimont 1989) is a Gram-negative bacterium belonging to Flavobacteriaceae (Bacteroidetes). *F. columnare* is distributed in fresh water sources worldwide and it is capable of causing an infection in cold and warm waters (Schneck and Caslake 2006). Kunttu *et al.* (2011) reported that this bacterium can form three colony types: rhizoid and flat, non-rhizoid and hard or soft. The temperature range in which *F. columnare* can grow actively is approximately 15 to 35°C (Declercq *et al.* 2013). Changes in climate are likely to lengthen the favourable growth period, which in turn leads to enhanced transmission and more frequent columnaris disease occurrences (Pulkkinen et al, 2010; Karvonen 2010). It has been suggested that adhesion of the high virulence strain is boosted by higher temperature (Decostere *et al.* 1999). Temperature may also promote the activity of the virulence factors in this pathogen, such as the chondroitin AC lyase activity (Teska 1993, Suomalainen *et al.* 2006a). Because of high phenotypic homogeneity, phenotype determination by using standard biochemical tests or chemotaxonomic markers is not an appropriate method for *F. columnare* strain characterization. However, based on earlier studies, it seems clear that *F. columnare* presents enough genetic variability to be defined using different molecular markers. Based on the analysis of 16S rDNA-RFLP, Triyanto and Wakabayashi (1999) split this species to three different genomovars (i.e. strains which are phylogenetically differentiable, but are phenotypically indistinguishable) that can be further subdivided using methods with higher resolution power. Finnish isolates studied thus far belong to genomovar I. Genomovars two and three thrive at relatively high temperatures but not at 15 °C (Triyanto and Wakabayashi 1999), thus they may not be able to survive the colder temperatures in Finland.

So far, our knowledge about the genetic diversity of *F. columnare* has been limited to a few DNA based methods such as several molecular typing approaches, including restriction fragment length polymorphism (16s-RFLP) (Triyanto and Wakabayashi 1999), Amplified Fragment Length Polymorphisms (AFLP) (Arias *et al.* 2004), pulsed-field gel electrophoresis (PFGE) (Soto *et al.* 2008), Single-strand conformation polymorphism (SSCP) (Olivares-Fuster *et al.* 2007) and automated ribosomal intergenic spacer analysis (ARISA) (Suomalainen *et al.* 2006b). However, these techniques have less utility in investigating underlying phylogenetic relationships in long-term or global epidemiology (fine-scale epidemiological investigations) because their indexed genetic variation has accumulated rapidly (Enright and Spratt 1999). Moreover, these methods lack reproducibility and cannot be readily transferred between laboratories. While the discriminatory power of the aforementioned methods is high for this organism, they have several disadvantages, including the low number of bands in each profile, making it less discriminatory (PFGE), lack of

standard set of primers designated for any bacteria species (AFLP), and they are time-consuming and expensive to conduct . Designing species-specific primers based on the complete genome of *F. columnare* (Tekedar *et al.* 2012) makes it possible to adapt new techniques, such as multi-locus sequence typing (MLST) and multi-locus sequence analysis (MLSA) schemes, to detect the genetic diversity and population structure of this bacterium (for more information see methods and material). These methods have shown that different strains can co-occur in a single region, differ in virulence, and compete (Sundberg *et al.* 2016), implying the potential for intraspecific competition among genotypes, which would change the outcome of *F. columnare* infection in the fish host.

1.8 Aims of the study

The ultimate aim of this thesis is to explore different approaches of investigating how climate change will affect environmental opportunistic bacteria, with a primary focus on temperature. The results will contribute broadly to understanding the sensitivity of *F. columnare* to future temperatures, and the factors which maintain high virulence across large geographic distributions of a single species.

In order to infer possible processes that shape the evolution of thermal adaptation and epidemiology of *F. columnare* at a finer geographic scale in Finland, one needs to study the phylogeny, genetic diversity, and population structure of a *F. columnare* population that has been sampled across Finland. A multilocus sequence typing (MLST) and analysis (MLSA) study was developed to characterize *F. columnare* isolates and answer the following questions (I):

- Is the species geographically differentiated? For instance, are northern and southern isolates different enough genetically to form two separate clusters?
- Are certain genotypes replaced by others during the years?
- Is there any recombination among housekeeping genes and how clonal is the population structure of *F. columnare?*

Temperature varies considerably between geographic regions, over the seasons, throughout the course of the day and between the sun and shade (Bradshaw and Holzapfel 2006). With climatic warming, the population size and distribution of organisms will be a function of naturally-selected thermal tolerance. Global warming may benefit many bacterial species, since they will face milder winter months resulting in greater overwintering success, increased numbers of generations and, thus, higher pathogen densities to damage hosts (Burdon and Chilvers 1982, Coakley *et al.* 1999). Environmentally growing opportunistic pathogens, in contrast to obligate (fully host-dependent) pathogens, can utilize outside host resources, making them very sensitive to selection pressures outside the host. Therefore, predicting the effect of climate

warming on environmental opportunistic bacteria with life cycles both outside and inside the host present a particular challenge because pathogen fitness in both environments may be differentially affected by temperature. Not only mean increase in temperature was expected to increase but also temperature extremes and both may have different effects on pathogen virulence. To identify variation in thermal tolerance among *F. columnare* isolates, I performed growth experiments in five different temperatures in order to characterize the temperature dependence of maximum biomass (yield) in *F. columnare* isolates collected across Finland (II). Two alternative approaches were used to analyse the thermal performance data: (I) principal component analysis (PCA) and (II) determination of thermal performance curve parameters (CTmin, CTmax, μmax, T_{opt}) for each strain after curve fitting. Moreover, a virulence experiment was conducted, using the zebra fish (*Danio rerio*) bath immersion infection model (II). Using the data on thermal tolerance and from the virulence experiment, I examined the link between thermal tolerance and the virulence, and focused specifically on the following questions (II):

- Is thermal tolerance associated with virulence?
- Has virulence increased during the years?
- How do other variables, rather than temperature, such as location and bacterial genotype affect virulence?
- Does an increase in mean temperature and temperature fluctuation affect virulence?

However, no organism exists in isolation, and it is important to note that the effects of climate change invoke factors other than temperature, particularly biotic interaction within and among species (Davis *et al.* 1998, Tylianakis *et al.* 2008, Höckendorff *et al.* 2015). For example, (Luhring and DeLong 2016) showed that the overall shape of TPCs in *Paramecium aurelia* changed in response to predation, indicating that the effect of predation on growth performance is temperature dependent. Biotic interactions, such as competition, are likely to become increasingly important in determining climatic projections as both climate and human induced range changes bring differentially adapted populations of the same species into contact. Hence, the prediction of temperature driven changes in intraspecific competition in pathogens, which in turn would affect the outcome of the interactions between the host and the pathogen, is essential in disease prevention. A recent study by Sundberg *et al.* (2016) showed that competition between genotypes of *F. columnare* is not restricted to resource competition, but that strains have a high capacity for interference competition. Commonly, due to technical challenges in distinguishing and tracking individual strains, experimental evidence on intraspecific competitions has been limited until now. I developed a cost effective high-resolution melting (HRM) assay to study head-to-head competition where two closely related genotypes of *F. columnare* were competed in liquid cultures, at two different growth temperatures (optimum and elevated

temperature, 26°C and 31°C, respectively) (III). I also studied interference competition between *F. columnare* strains grown on a solid medium. Consequently, the aim of this study was to experimentally investigate:

- The effect of temperature on exploitation competition
- The effect of temperature on interference competition
- If head to head competition and interference competition in two temperatures will lead to the same or different outcome.

2 MATERIALS AND METHODS

2.1 Study isolates

A total of 83 *F. columnare* isolates were collected from eight geographic locations in Finland, comprising four northern and four southern locations over 10 years (2003-2012). This collection contains isolates from nine different fish species and water samples (Fig. 2) (I). Two *F. columnare* reference strains JIP39/87 and ATCC49512, both isolated in France, and type strain NCIMB 2248T, isolated in the USA, were also included in the sample collection.

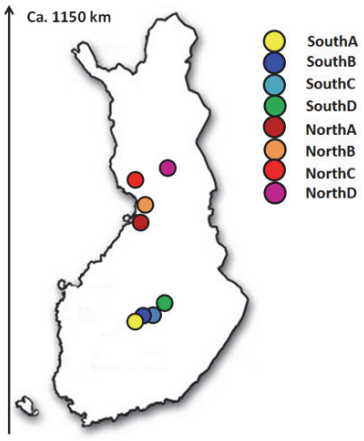


FIGURE 2 Location of sample sites in Finland for the 83 *F. columnare* isolates used in this thesis (Study I).

2.2 Phylogeography of thermal tolerance in *F. columnare*

Multi-locus sequence typing (MLST) is a typing method based on profiles and sequence type designations to estimate relatedness among isolates. MLSA relies directly on nucleotide sequences; it uses concatenated sequences of fragments of housekeeping genes to determine genus-wide phylogenetic relationships and is outlined in detail in Fig. 3 (I). 23 candidate genes were chosen based on the previously published MLST scheme for *Flavobacterium psychrophilum* by Nicolas *et al.* (2008). After screening these 23 candidate housekeeping genes, a panel of 6 genes (*trpB, tuf, atpA, rpoD, gyrB and dnaK*) was chosen for developing the MLST/MLSA scheme. I designed the primers based on the corresponding sequences derived from the whole genome sequence of *F. columnare* ATCC 49512 (= CIP 103533 = TG 44/87) (Tekedar *et al.* 2012). In order to determine the levels of inter-lineage recombination, mutation and the underlying population structure present in our data, I used both Structure (Pritchard *et al.* 2000) and ClonalFrame (Didelot and Falush 2007) programs.

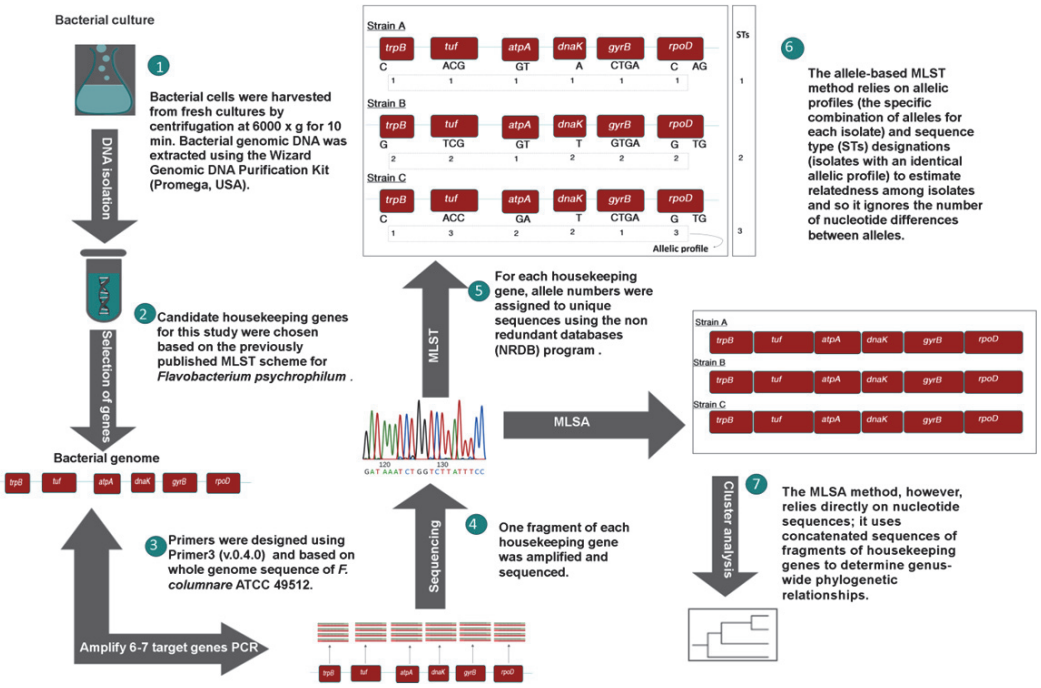


FIGURE 3 Schematic diagram of the steps involved in MLST/MLSA analysis.

2.3 Thermal tolerance of *F. columnare*

To illustrate thermal tolerances as well as the capability for thermal adaptation of *F. columnare*, performance of the bacterium (growth yield**)** was measured over five different temperatures. Bacterial growth was measured with a Bioscreen C® temperature controlled spectrophotometer (Growth Curves AB Ltd, Finland) as increase in turbidity (OD at wide band 420–580 nm) at 5 min intervals. Next, six different functions were used to determine the temperature-dependent growth yield, using the TableCurve 2D software (version 5.01, Systat Software Inc). To correctly describe the thermal performance of this bacterium and to avoid overfitting in models, I used a weighted-average model based on their AIC values for each strain. The thermal performance curve is a prerequisite to calculate key ecophysiological characteristics CTmin, CTmax, μmax, and Top (Figure 1). Next, I used these estimated key characteristics to compare the current thermal tolerances of *F. columnare* isolates of differing geographic origin, isolation year, and genotypic groups. I also used principal component analysis (PCA) to quantify and visualize different patterns of variation in thermal performance curves (TPCs) (Izem and Kingsolver 2005).

To explore if there is a link between thermal tolerance and bacterial virulence, an infection experiment was performed using the zebra fish *(Danio rerio)* infection model. The virulence experiment followed the method described by Laanto *et al*. (2014). Briefly, ten fish per strain type were challenged in individual aquaria containing 0.5 l of water. $4x10^5$ colony-forming units (CFU) mL^{-1} of overnight grown bacterial dose that were adjusted in preliminary experiments were suspended in 500 μl Shieh medium, and added into aquaria at 25.5 °C. After challenging, the fish were monitored for clinical signs of disease and mortality for 2 to 11 days depending on the progression of the disease. For the first 48 hours, the fish were monitored at intervals of one or two hours, when progress of the disease slowed down, the monitoring interval was reduced accordingly, but the fish were checked at least twice a day.

2.4 Mixed genotype competition by using HRM analysis

For the last study, *F. columnare* isolates were competed together in four competition pairs, each containing two genotypes in two different experimental temperatures (26 °C and 31 °C) for a fixed period of 24 h in liquid cultures. After 24 h, the competition mixture was diluted 100-fold into fresh medium, and this was repeated every 24 hours for 14 days. DNA samples were taken on days 0, 1, 2, 4, 6, 7, and 15 and are outlined in Fig. 4. By amplifying genomic DNA containing a single SNP (T > G substitution) at position 222 in the tryptophan synthase gene, *trpB*, I could accurately determine the proportions of the different strains present within mixture samples.

High-resolution melting analysis is a combination of a polymerase chain reaction (PCR) and a post-PCR DNA melting curve method that can be used to distinguish between very similar sequences, due to its capacity to point out a single nucleotide polymorphism. First, PCR is performed in the presence of a dye that binds to a double-stranded DNA (dsDNA) to amplify target sequences. This dye shows low levels of fluorescence when unbound, but is highly fluorescent when it bonds to double-stranded DNA. The PCR amplicons are then gradually denatured by increasing temperature in steps of ca. 0.01 °C to 0.2 °C; this is termed melting analysis. During melting analysis, dsDNA denaturates to single-stranded DNA and releases the fluorescent dye. A melting curve can be generated from the decreasing fluorescence plotted against increasing temperature. Comparison of the melting curve profiles and the specific melting temperature (Tm: the temperature at which 50% of the DNA sample is double stranded and 50% is single-stranded) of each profile, enabled us to identify and characterize DNA samples based on the amplicon's length, sequence, and GC content (Garritano *et al.* 2009). In order to quantify the proportions of two different genotypes (G and A) (I) in a mixed culture, a linear model: $meltingTemp = a + b * genotypeProp$ was fitted to the calibration data from reference samples where purified DNA from genotypes G and A pooled in known proportions (13 different proportions of the two genotypes: 100%, 95%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5% and 0%) were included in duplicate in the plate. Genotype proportions in the experimental samples were quantified by using a piecewise-defined function using the relationship $genotypeProp = 1/b * (meltingTemp - a)$.

Moreover, interactions among *F. columnare* strains are not restricted to competition for resources but strains can interfere with competitors by directly inhibiting growth. The inhibitory activity of *F. columnare* strains was tested reciprocally using an inhibition zone method with four replicates per assay, and a double layer method (Sundberg *et al.* 2016). Briefly, in inhibition zone method a small volume of overnight-grown 'recipient' bacterial cultures are mixed in about 3 ml of molten, 'soft', agar at 47°C, gently homogenized, and poured into a petri dish previously prepared with 10 ml of solid agar at the bottom. The plates were gently swirled and dried for 10 min at room temperature. Unfiltered supernatant of the 'donor' cultures (i.e. toxin production in *F. columnare* require a direct contact between bacterial cells) were spotted onto a lawn of 'recipient' bacteria on the soft agar and then inverted and incubated at two experimental temperatures (26 °C and 31 °C).

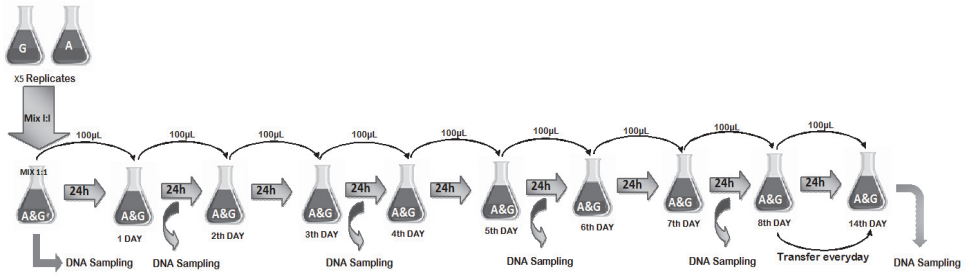


FIGURE 4 Competition experiment and DNA sampling process for high resolution melting analysis.

3 RESULTS AND DISCUSSION

3.1 MLST/MLSA analysis identifies epidemic clonal population structure for *F. columnare* in Finland (I)

Despite being important fish pathogen, little is known about the genetic diversity and evolutionary history of *F. columnare*. All isolates in this study had originally been assigned to *F. columnare*, but MLSA revealed that they represent two distinct lineages. Lineage I includes two clusters, cluster C and cluster E. Lineage II includes three clusters, clusters G, A and H. Due to high sequence similarity (almost 100 %), out of 23 candidate genes, only six loci provided a higher level of discrimination for *F. columnare*. It should, therefore, be noted that the level of discrimination with six loci is different than for other pathogens with high diversity (Feil *et al.* 2004).

The phylogenetic and Bayesian analyses clearly indicated that *F. columnare* strains are genetically distinct from *F. psychrophilum*. In addition, while the reference strains ATCC 49512 and JIP 39/87, both from France, were phylogenetically close to isolates from Finland and clustered with the Finnish strains, whereas the type strain (NCIMB 2248T) from the USA appears to be highly divergent from the Finnish *F. columnare* strains. The overall IA^S (standardized index of association) value was 0.198 (P = 0.03), meaning that recombination was not sufficient to break down linkage disequilibrium. However, results from STRUCTURE corresponded well to the results from the ClonalFrame analysis with respect to the fact that recombination events have occurred both within and across lineages. Recombination happened less frequently than mutation ($\rho/\theta \approx 0.14$; with 95% confidence interval of 0.03-0.9). Nevertheless, recombination events and horizontal gene transfer (HGT) were/have been effective in introducing polymorphisms by the *r/m* value of approximately three (with 95% confidence interval of 0.44-6.6). This indicated that recombination happened six times less often than mutation, but recombination overall caused three times more substitutions than mutation, confirming the importance of recombination in this bacterium. With this evidence, the population structure of *F. columnare* is most likely following a

semiclonal/epidemic clonality model where there is a background level of frequent recombination with consecutive clonal expansion of one or a few fit genotypes (Smith *et al.* 1993, Smith *et al.* 2000, Feil *et al.* 2004, Tibayrenc and Ayala 2015); similar results have been reported for *Vibrio parahaemolyticus* (Gonzalez-Gonzalez *et al.* 2013) and *F. psychrophilum* (Nilsen *et al.* 2014).

The observed clonality in the whole population could be explained by the dominance of clonal lineages (lineages I and II) relatively unbroken by horizontal gene transmission as a result of niche specialization (i.e., host adaptation), ecological barriers, or recent introduction (Nilsen *et al.* 2014). For example, the presence of different genetic elements such as CRISPR-Cas systems has been shown to prevent successful recombination between the lineages (Budroni *et al.* 2011, Wielgoss *et al.* 2016). *F. columnare* has two different CRISPR systems (Laanto *et al.* 2017) that have the potential to alter the patterns of phage infection and transduction. This could potentially explain the decreased homologous recombination between the clades. Alternatively, it is possible that two lineages inhabiting a spatially and/or temporally distinct niche within the same habitat could preclude close contact favouring genetic exchange (Vos 2011, Shapiro and Polz 2014, Wielgoss *et al.* 2016).

Strains from each genotypic cluster were isolated across the years and from different origins, indicating that columnaris outbreaks in Finland are caused by continuous co-circulation of *F. columnare* strains. The same pattern, lack of association between STs and their geographic origins/isolation year/host, has also been reported with this bacterium elsewhere (Arias *et al.* 2004, Schneck and Caslake 2006). A possible explanation for the occurrence of genetically similar populations in geographically distinct areas is the transportation of fish stocks between fish farms and/or natural (e.g. wild birds and bird migration) exchanges of *F. columnare* strains from other regions (Mohammed and Arias 2014).

3.2 *F. columnare* exhibits variation in thermal tolerance (II)

Karvonen *et al.* (2010) study analysing more than 20 years' worth of data shows a significant positive effect of mean water temperature on the prevalence of columnaris disease at two farms, indicating that the disease increased along the water temperature increase. Therefore, it is important to understand if climate change will change the thermal tolerance and virulence of this bacterium, especially when Finland's average annual temperature will rise nearly twice as fast as the average temperature for the whole globe (2–7°C) (Ruosteenoja *et al.* 2016).

I did a detailed characterization of temperature-dependent maximum biomass (i.e. yield) of 50 isolates of *F. columnare* that were collected from eight different areas located across Finland between 2003 and 2012; and both model based and principal component analysis gave similar biological interpretations of the data. These results revealed that Finnish *F. columnare* shows a typical

thermal performance curve with a rather high optimum temperature between 23.7 °C and 27.9°C and an upper critical temperature for yield between 30.1 and 34.7 °C. Since Finnish lakes form predominantly closed and shallow basins (average depth about 7 metres), surface waters warm up fast in spring and may reach high temperatures in summer. Therefore, strains inhabiting these lakes may exhibit high optimum temperatures and evolve high upper critical temperature during warmer summers. Although different genotypes of Finnish *F. columnare* did not clearly correspond to geographic regions, genotypes differed in their key ecophysiological characteristics (CTmin, Topt, OPperformance, TTB, and CTmax); indicating the importance of phylogenetic constraints on thermal adaptation of *F. columnare*.

It is expected that individuals in colder environments will exhibit lower critical minimum temperatures (CTmin). I showed that isolates from colder, higher latitudes display greater cold tolerance and broader tolerance ranges than those from more southern localities. However, despite the lower optimum temperatures of northern isolates in Finland, northern isolates may compensate for the short growing season by reproducing more rapidly to ensure successful reproduction and transmission (Yang *et al.* 2011).

This study also demonstrated that improved heat tolerance was accompanied by reduced cold tolerance in *F. columnare*, which was in line with horizontal shift of the curve; also, called the hotter-colder mode (Figure 1). Altogether, these results show that environmental thermal conditions could play an important role in determining the diversity and prevalence of *F. columnare*.

The first principal component for performance measurements, accounting for the largest component of variation (46%), had a large positive loading at both lower and higher temperatures but loadings at the other temperatures were comparatively small and close to zero. This pattern of loadings is qualitatively consistent with the generalist–specialist hypothesis. However, opposing the classic generalist–specialist trade off hypothesis, I did not find that generalist genotypes with broader performance breadth would have lower biomass. Yet, further analysis showed that strains with higher upper limit temperatures have smaller maximum biomass, explaining that strains with broader thermal performance still would carry a cost. Moreover, these strains with right-skewed and broader thermal performance curve were less virulent, demonstrating that the cost of generalism is not always found in the selective environment, but could instead have a trade-off with some other traits, such as virulence (Huey and Hertz 1984, Sturm *et al.* 2011, Ketola *et al.* 2013).

I think my findings form a base for future work with population genetics and thermal adaptation in this intriguing bacterium. This would be especially interesting because the genomovar I strains are only found in colder environments such as Finland, whereas other genomovars can infect also warm water fish. For example, increasing sample size with isolates from different parts of the world could produce a significantly clearer picture on worldwide population genetic structure and ecology of *F. columnare*. Based on disease triangle hypothesis, environment influences disease development by interacting

with both the host and the pathogen. The main focus of my thesis work was on bacteria and temperature; what is needed further is to evaluate the effect of temperature on both the pathogen and the host simultaneously. Previous work by Penttinen *et al.* (2016) has showed that high nutrient concentration can induce virulence factor expression and cause higher virulence in *F. columnare.* As climate warming has been suggested to increase aquatic eutophication via elevated nutrient supply (Heino *et al.* 2009) studying interactions between nutrients and temperature will be crucial for developing realistic predictions about associated ecological responses in environmentally transmitting opportunistic bacterial pathogens.

3.3 Thermal tolerance at high temperature is linked with reduced virulence (II)

In theory, due to co-evolutionary shifts in both host and pathogen, virulence is expected to decrease over time to the point of optimized fitness of both host and pathogens. However, virulence is context–dependent, as both biotic factors such as host condition (Pulkkinen and Ebert 2004) and host density, (Bieger and Ebert 2009) and abiotic factors, such as temperature (Guijarro *et al.* 2015), can influence virulence. In line with previous studies (Sundberg *et al.* 2016), I found that the virulence of Finnish *F. columnare* strains was increasing over the past 10 years. It has been shown that high environmental temperature could select for increased bacterial pathogenicity (Friman *et al.* 2011). Interestingly, not only did I find compelling evidence that high temperatures are likely to cause strong selection on virulence, but also that bacterial virulence decreased when the upper end of performance curve shifted towards higher temperatures. Therefore, climate warming is unlikely to exert strong positive selection on highly virulent strains via correlated selection; however, increases in mean temperatures associated with global change could impose further selection by increasing the length of growing seasons (Figure 5). Consequently, extended growing seasons could significantly increase the prevalence of disease epidemics.

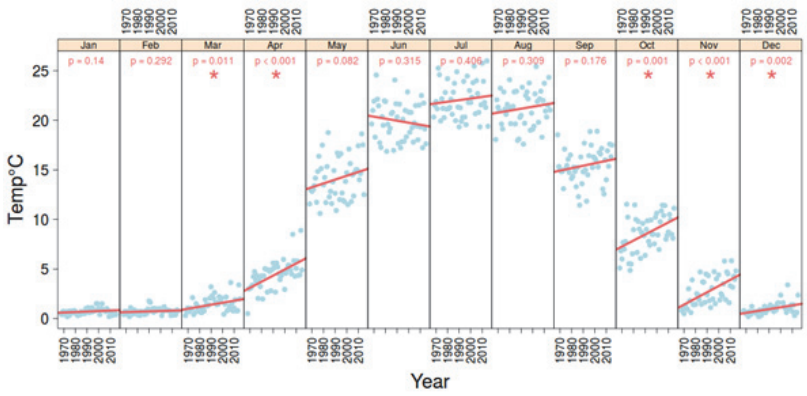


FIGURE 5 Temporal water temperature changes in a Finnish fish farm (Tyyrinvirta fish farm) (study III). The graphs show the evolution of averages of three highest temperatures per month. Red lines are fitted using a linear model within each month. The p-value for the significance of the year effect on the monthly values is reported for each month.

3.4 Temperature affects intraspecific interaction in opportunistic bacteria differently in different environments (III)

Most experimental research so far has focused on how temperature impacts organisms' survival directly, but little is known about how climate change can invoke factors other than temperature, particularly changes in biotic interactions within and among species. Monitoring co-cultured bacterial strains in different temperatures provides a perfect playground to understand the effect of changing temperatures on intraspecific competition, which would help us to address questions about the effects of biotic interactions in determining the outcome of climate change projections. However, such questions have been difficult to answer to date because of technical challenges in distinguishing and tracking individual strains. A broad range of methods are available, which have been used to discriminate strains and determine the relative abundance of multiple strains that interact within a mixed culture, including phenotypic markers based on genetic engineering (e.g. $lacZ^+$ vs $lacZ^-$, fluorescent markers, antibiotic resistance), and genotyping (Orvos *et al.* 1990, Minter *et al.* 2015). However, these markers may provide misleading results, as they may incur fitness costs or benefits in bacteria (Blot *et al.* 1994, Ramadhan and Hegedus 2005, O'Keefe *et al.* 2006). Since methods based on real-time PCR and HRM analysis do not require any alteration in the phenotype or genotype of the wild-type strain, this technique allowed us to study ecological fitness and competitive interactions between natural isolates for the first time. By amplifying genomic DNA containing a natural mutation in the tryptophan synthase gene, *trpB*, I was able to accurately determine the proportions of different strains present within mixture samples. More importantly, the results

showed that the HRM method enables us to accurately quantify the relative abundance of different strains in mixed cultures, and suggest that HRM offers a simple and cost effective means to study head-to-head competition between mixed genotype in a single culture tube; even if co-occurred samples differ from each other only by one sequence variant or SNP. I was able to show that this method is extendable to simultaneous quantitative separation of several genotypes by multiplexing different primers used in qPCR, identifying 4 genotypes of Finnish *F. columnare* (A, C, E, and G) simultaneously.

The competition outcome of two strains in mixed culture at two temperatures were associated with changes in bacterial growth rate in the given temperature; in line with faster growing organisms having an advantage during early colonization (Leveque 2003, de Muinck *et al.* 2013) which can translate into a competitive advantage in exploiting resources efficiently. However, head-to-head competition results were mostly inconsistent with interference competition on solid agar. No growth inhibition between strain pairs was detected at the higher temperature (26 vs. 31°C), in line with findings from other bacterial species such as *Yersinia* (Bottone *et al.* 1979, Aasen *et al.* 2000, Mataragas *et al.* 2003) and *Lactobacillus sakei* Lb706 (Diep *et al.* 2000), possibly indicating degradation or inactivation of the inhibitory compounds in *F. columnare* at high temperatures.

Bacterial activities such as motility, attachment, coordinated behaviour, antibiotic production and secretion of extracellular polymers can change the outcomes of competition to differ from those predicted in planktonic cultures (Hibbing *et al.* 2010).

Particularly, in biofilms in which bacterial populations are dense and nutrient-limited, cohabitation could result in strong intra- and interspecific exploitative competition. Hence, strong natural selection resulting from exploitative competition among different genotypes is predicted to generate interference competition that serves to kill competitors and divert resources to their own reproduction (Majeed *et al.* 2011, Cornforth and Foster 2013). It has been shown that bacteriocin expression in biofilms can be two- to three- folds higher than in planktonic cultures, suggesting that bacteriocin production could change competition outcome in different environments (Majeed *et al.* 2015). Alternatively, the production of inhibitory compounds can be costly (Senadheera *et al.* 2012). Therefore, production of these toxic compounds is not without costs, and producer strains may grow slower than sensitive strains (Riley and Wertz 2002). In liquid cultures, toxins are subject to dilution and sensitive strains with higher growth may win the competition. Furthermore, limited dispersal can favour toxin production, suggesting that toxins provide an advantage in competition for sessile cells on surfaces rather than the planktonic cells (Wloch-Salamon *et al.* 2008). Therefore, proxies of fitness such as individual growth and yield might not be predictive enough of which of the genotypes will excel in competition depending on the environmental setting.

4 CONCLUSION

Knowledge of the phylogenetic relationships among pathogen populations located in different parts of the distribution area is important to understand a population's current thermal tolerance and adaptation to future climate warming. Phylogenies based on housekeeping genes are highly valuable for investigating the evolutionary history and population genetic structure of bacteria, because nucleotide changes in housekeeping genes (1) are constrained by the essential biochemical functions of the proteins which they encode, (2) are not normally exposed to other selective pressures, and (3) most of the variation that is observed is neutral and slowly evolving (Maiden 1999). Implementation of this MLST/MLSA scheme with a reproducible, powerful, methodology and comparable data between laboratories will enable us to apply and extend this method to more *F. columnare* strains from other geographical locations across the world, which will facilitate the production of a global picture of the epidemiology and evolution of this pathogen. Both MLST and MLSA analyses revealed no clear clustering based on geography, host species, and year of isolation, in line with the semi-clonal population structure of this bacterium which diversifies with a moderate, but variable, recombination rate. The persistence of the same sequence type originating from different locations and over extended periods is indicative of continuous co-circulation of *F. columnare* clones. Furthermore, the recovery of genetically divergent clones from the same location (fish farms) indicates the possibility of multiple infections by different *F. columnare* genotypes that could have important implications for disease control.

However, *F. columnare* as an opportunistic bacterium occurs in different habitats, such as water, sediment, host (fish) skin, and host body, where nutrient contents, temperatures, and pH differ. Any changes in these factors require frequent bacterial adaptation. I found that *F. columnare* populations with different genotypes differ in their responses to temperature. Therefore, this would render this bacterium well suited to colonizing different locations, and it may partly explain the existence of sympatric clones with no uniform response to temperature. Moreover, thermal tolerance results indicate the northern *F.*

columnare strains seemed to be adapted to lower temperatures. However, as strains from colder environments optimize their fitness by having greater yield in lower optimum temperatures than strains from warmer environments, it appears that faster growth and higher yield may also have evolved to compensate short growing seasons. *F. columnare* strains showed a combination of two key modes of variation in their thermal performance curves, (TPCs) reflecting shifts in both thermal breadth (generalist–specialist) and optimum temperature (hotter –colder). Interestingly, generalists did not show an obvious trade-off between tolerance range (TBP) and performance at optimum conditions compared to specialists, but strains with higher upper critical temperature (CTMax) had lower maximum biomass in optimum temperatures. Virulence results showed that *F. columnare* strains have become more virulent over time. Yet, higher temperatures *per se* might not be the causative factor for selecting high virulence. Nevertheless, these strains seem to potentially benefit from future increasing temperatures via extended growing seasons and consequently cause more problems at fish farms.

In the last chapter, I explored temperature-dependent biotic interactions that are still rarely incorporated in determining the outcome of climate change projections because of the technical challenges in distinguishing and tracking individual strains. I found that competition outcome in liquid cultures seemed to be well predicted by the growth yield of isolated strains, but was mostly inconsistent with interference competition results measured in inhibition tests on solid agar, especially as no growth inhibition between strain pairs was detected at the higher temperature. These results suggest that, for a given temperature, the factors driving competition outcome differ between liquid and solid environments. The efficiency of the high-resolution melting (HRM) assays in quantifying the relative proportions of strain genotypes makes it a powerful tool to characterize intraspecific competition in microbes in more detail. These results suggest that temperature-dependent variation in interaction strength could also play important role in species and community responses to climate warming. However, the environmental setting (i.e. depending on if bacteria cells are sessile or free swimming) may also translate into significant changes in interactions, making it more difficult to predict species and community responses to climate warming. Moreover, evidence of the close match between growth and competition is also important to understand possible outbreak scenarios under changing environmental conditions in environments where multiple pathogen strains co-occur.

Acknowledgements

There comes a time in life when you'll have to leave everything behind and start something new but never forget the ones who stood by your side, especially the ones who never gave up on you.

I cannot thank Tarmo Ketola enough for all his help and support over the past few years. His positivity and enthusiasm for science and my project never failed to amaze me, and kept me going despite many setbacks!! I have learnt so much from my time at your lab. Thanks, you were always there to help, give advice and support no matter how busy you were. You taught me self-worth, the importance of hard work, and to love working with bacteria! You are an inspiration.

I am forever indebted to my supervisors, Dr Lotta-Riina Sundberg and Dr Katja Pulkkinen, for their enthusiasm, guidance, and constant support throughout this process. They have routinely gone beyond their duties to instil great confidence in both myself and my work. They have always been there for me, kept me in check and making me enjoy every second of working with this fascinating creature. Lotta, you have been a tremendous supervisor for me. Your advice on both research as well as on my career have been priceless. Thank you for trusting me. Katja, it has been a pleasure working with you. Talking to you always helped me to calm down and relax.

I would like to thank Prof. Johanna Mappes, for the patient guidance, encouragement and advice she has provided throughout all these years. I have been extremely lucky to have you who cared so much about my work and encouraged me to be the best that I could be. Thank you for believing in me and making me believe in me. No amount of words can ever express my gratitude to you. Thanks for letting me follow my dreams. Thanks for everything Johanna.

Working on my thesis would be a lot different without Matthieu Bruneaux who helped me particularly in completing analysis with R. You know that your contribution in coming up with this thesis is indispensable. I will forever be grateful for having you as my friends. Elina, thank you for all your comments during my experiments and kindness! You are the best. A big thanks to the past and present members of the experimental evolution journal club and particularly, Ilkka who made it happens and keeps this club alive. I would like to thank all my roomies and friends Anbu Poosakkannu, Kati Saarinen, Hanna Kinnula, Venera Tyukmaeva, Aigi Margus, and David Hopkins. You have been great!

Aigi, thanks for always listening to me, supporting me, and encouraging me. You're a true friend, and I want you to know how much I appreciate your friendship.

I am also very grateful to my external reviewers, Dr. Freya Harrison and Dr. Otto Eerikki Seppälä for insightful comments on this thesis. Also big thanks

to my support group people Teppo Hiltunen, Juan Galarza and scientific editor Jari Haimi. Thank you for your support and encouragement.

Furthermore, I am indebted to many people for making the time working on my master and Ph.D. an unforgettable experience; thank you Mikko Mönkkönen, Leena Lindström, Anneli Hoikkala, Emily Knott, Janne Valkonen, Nina Pekkala, Juho Niva, Sami Kyröläinen, Tiina Hakanen, Paula Sarkkinen and Anne Lyytinen.

I am also very thankful to our technical assistant, Sari Viinikainen who not only introduced me to various techniques in the lab but also helped me with every problem concerning my practical work. Thanks Sari!

Finally, I would express a deep sense of gratitude to my parents, especially to my dearest darling mom, who has always stood by me like a pillar in times of need and to whom I owe my life for her constant love, encouragement, moral support and blessings. Mom, dad, you have always encouraged me to explore my potential and pursue my dreams. To my dear sister and three brothers who always encouraged me and never lost faith in me, a big, warm Thank you! I would like to thank my husband, Hamid, for his supporting and understanding during the thesis writing. And, thank you for being such a great father, even in the wake of my absence!

Sina, my son, your eyes give me the will to fight back, even when everything goes off-track. Your hugs give me a reason to smile, even when problems pile. Your love is what keeps me going, it is what keeps my heart beating. Thanks darling for being such perfect son! You are mom's everything.

YHTEENVETO (RÉSUMÉ IN FINNISH)

Lämpötilan vaikutukset ympäristössä kasvavan kalatautibakteerin evoluutioon

Luonnossa elävistä bakteereista osa on potentiaalisia taudinaiheuttajia, jotka aika ajoin aiheuttavat tautiepidemioita. Näiden opportunististen bakteerien kahtiajakoinen elämäntyyli (välillä isännässä ja välillä isännän ulkopuolella ympäristössä) altistaa ne myös ympäristön valintapaineille. Tällä hetkellä elämme keskellä voimakasta ympäristömuutosta, jossa maapallon keskilämpötila nousee ja lämpötilojen vaihtelut lisääntyvät dramaattisesti ihmisten toiminnan seurauksena - eikä muutoksessa näy hidastumisen merkkejä. Lämpötila säätelee voimakkaasti eliöiden kasvua ja selviytymistä. Varsinkin luonnossa ja vaihtolämpöisissä eliöissä elävät bakteerit ovat herkkiä lämpötilan muutoksille, mutta myös lajinsisäiselle kilpailulle. Molemmat näistä tekijöistä voivat vaikuttaa bakteerien taudinaiheutuskykyyn (virulenssi) ja sen evoluutioon.

Flavobacterium columnare – bakteeri on taudinaiheuttaja, joka sairastuttaa ja tappaa kaloja. Tautia esiintyy kalanviljely-ympäristöissä ympäri maapalloa ja se aiheuttaa merkittäviä taloudellisia tappioita. Tautiongelmat ovat lisääntyneet myös Suomessa 1990-luvulta lähtien. *F. columnare* – bakteeri on opportunistibakteeri, eli sillä on kyky elää sekä kalassa että sen ulkopuolella. Ympäristön lämpötilanvaihtelut voivat vaikuttaa merkittävästi taudin yleisyyteen ja virulenssiin. Virulenssin ja lämmönsiedon yhteyttä ei ole kuitenkaan tutkittu.

Tutkin väitöskirjassani *F. columnare* – bakteerin populaatiogenetiikkaa, kilpailukykyä ja lämpötilansietoa. Tutkimusaineistona käytin 83 bakteerikantaa, jotka on eristetty sairastuneista kaloista tai vedestä. Kannat oli eristetty kymmenen vuoden aikana (2003 - 2012) eri puolilta Suomea. Tutkimuksissani selvisi, että nämä kannat jakautuvat geneettisesti kahteen päälinjaan ja viiteen eri genotyyppiin. Käytin tässä työssä menetelmää, jossa kuuden muuntelua sisältävän geenin (*trpB*, *tuf*, *atpA*, *rpoD*, *gyrB* ja *dnaK*) sekvenssejä verrattiin keskenään. Bakteerin eristämisajankohta, -paikka tai kalalaji ei vaikuttanut genotyyppien esiintymiseen. Genotyyppien välillä oli merkittäviä eroja lämmönsietokyvyssä. Pohjois-Suomesta eristetyillä kannoilla oli parempi kylmänsieto kuin Etelä-Suomesta eristetyillä kannoilla. Lisäksi parempi kylmänsieto oli yhteydessä huonompaan lämmönsietoon. Vaikka nämä löydökset osoittavat, että lämpötila voi vaikuttaa tämän lajin esiintymiseen ja evoluutioon, lämpötilasta johtuvat evolutiiviset muutokset eivät kuitenkaan yksin pysty vakuuttavasti selittämään taudinaiheutuskyvyn muutoksia.

Kun vertailin eri kantojen kykyä sairastuttaa seeprakaloja, havaitsin, että taudinaiheutuskyky oli kasvanut vuosien saatossa. Tulokseni eivät kuitenkaan tue olettamusta, että lämpötila olisi virulenssin muutoksiin johtava valintatekijä. Tähän päätelmään on kaksi syytä: 1) aiheuttaakseen valintaa virulenssiin, lämpötilojen tulisi olla kalanviljelylaitoksessa huomattavasti havaittuja lämpötiloja korkeampia, ja 2) epidemiat puhkeavat melko alhaisissa lämpötiloissa ja bakteereita pyritään hävittämään antibiooteilla heti taudinpurkauksen alkaessa.

Tämä tarkoittaa, että ainakaan kalanviljelyolosuhteissa, joissa bakteerit on hävitetty jo ennen korkeiden lämpötilojen saavuttamista, valinta ei voi kohdistua korkeiden lämpötilojen sietoon. Havaittu epidemioiden vakavuuden lisääntyminen voikin yksinkertaisesti selittyä pidentyneellä kasvukaudella, joka mahdollistaa suuremmat bakteeribiomassat, sekä muilla kalanviljely-ympäristössä vaikuttavilla valintatekijöillä.

F. columnare -kantojen on havaittu kilpailevan voimakkaasti keskenään erittämällä myrkkyjä, jotka estävät muiden kantojen kasvua. Periaatteessa lajin sisäinen kilpailu voi sekoittaa lämpötilan aiheuttamaa valintaa ympäristössä, jos hyvin lämpöä sietävä bakteerikanta onkin huono sietämään kilpailua. Tällöin bakteerin kasvu tietyssä lämpötilassa ei välttämättä selittäisi sen selviämistä tässä lämpötilassa. Tutkin näitä olettamuksia menetelmällä, jolla pystyin seuraamaan muuten hankalasti toisistaan erotettavia bakteerikantoja. Tämä menetelmä perustuu *trpB*-geenin sekvenssissä olevien pienien erojen havaitsemiseen. Tämä tutkimus paljasti, että mitattu kasvu nesteviljelmässä ennusti myös kilpailukykyä vedessä sekä matalassa että korkeassa lämpötilassa. Sen sijaan Agar-maljojen pinnalla mitattu kantojen keskinäinen kilpailu ennusti huonosti kilpailukykyä nesteessä, ja korkeassa lämpötilassa (31 °C) kilpailua ei havaittu lainkaan. Lämpötila siis vaikuttaa voimakkaasti kilpailullisiin vuorovaikutuksiin riippuen siitä, tapahtuuko valinta pinnoilla vai nesteessä, eikä kasvumittauksilla tai kilpailutesteillä voida aukottomasti selvittää, mikä kannoista pärjää missäkin kilpailutilanteessa.

Tutkimukseni perusteella on selvää, että lämpötila vaikuttaa *Flavobacterium columnare* - bakteerin kantojen väliseen kilpailuun, evoluutioon ja kantojen esiintymiseen. Lämpötila ei kuitenkaan selitä muutoksia virulenssissa. Voimistuneiden epidemioiden syynä lienevätkin kasvukauden pidentyminen sekä muut valintatekijät.

REFERENCES

Aasen I., Møretrø T., Katla T., Axelsson L. & Storrø I. 2000. Influence of complex nutrients, temperature, and pH on bacteriocin production by *Lactobacillus sakei CCUG 42687*. *Appl. Microbiol. Biotechnol.* 53: 159-166.

Adiba S., Nizak C., Van Baalen M., Denamur E. & Depaulis F. 2010. From grazing resistance to pathogenesis: the coincidental evolution of virulence factors. *PLoS ONE* 5: e11882.

Agrios G.N. 2005. Plant Pathology. 5th eds.

Alexander J.M., Diez J.M. & Levine J.M. 2015. Novel competitors shape species/'responses to climate change. *Nature*. 525: 515-518

Alizon S., Hurford A., Mideo N. & Van Baalen M. 2009. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *J. Evol. Biol.* 22: 245-259.

Anderson R.P., Peterson A.T. & Gómez-Laverde M. 2002. Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos* 98: 3-16.

Angilletta M.J. 2009. *Thermal adaptation: a theoretical and empirical synthesis.* Oxford University Press.

Arias C.R., Welker T.L., Shoemaker C.A., Abernathy J.W. & Klesius P.H. 2004. Genetic fingerprinting of Flavobacterium columnare isolates from cultured fish. *J. Appl. Microbiol.* 97: 421-428.

Aujoulat F., Roger F., Bourdier A., Lotthé A., Lamy B., Marchandin H. & Jumas-Bilak E. 2012. From environment to man: genome evolution and adaptation of human opportunistic bacterial pathogens. *Genes* 3: 191-232.

Babiker H.A., Ranford-Cartwright L.C. & Walliker D. 1999. 3. Genetic structure and dynamics of *Plasmodium falciparum* infections in the Kilombero region of Tanzania. *Trans R. Soc. Trop. Med. Hyg*. 93: 11-14.

Balmer O., Stearns S.C., Schötzau A. & Brun R. 2009. Intraspecific competition between co-infecting parasite strains enhances host survival in African trypanosomes. *Ecology* 90: 3367-3378.

Bernardet J. & Grimont P.A. 1989. Deoxyribonucleic acid relatedness and phenotypic characterization of *Flexibacter columnaris* sp. nov., nom. rev., *Flexibacter psychrophilus* sp. nov., nom. rev., and *Flexibacter maritimus* Wakabayashi, Hikida, and Masumura 1986. *Int. J. Syst. Bacteriol.* 39: 346-354.

Bieger A. & Ebert D. 2009. Expression of parasite virulence at different host population densities under natural conditions. *Oecologia* 160: 247-255.

Blot M., Hauer B. & Monnet G. 1994. The Tn5 bleomycin resistance gene confers improved survival and growth advantage on *Escherichia coli*. *Molecular and General Genetics MGG* 242: 595-601.

Bottone E.J., Weitzman I. & Hanna B.A. 1979. *Rhizopus rhizopodiformis*: emerging etiological agent of mucormycosis. *J. Clin. Microbiol.* 9: 530-537.

Box E.E. 2012. *Macroclimate and plant forms: an introduction to predictive modeling in phytogeography*. Springer Science & Business Media.

Bradshaw W.E. & Holzapfel C.M. 2006. Evolutionary response to rapid climate change. *Science(Washington)* 312: 1477-1478.

Brandl M.T. 2006. Fitness of human enteric pathogens on plants and implications for food safety 1. *Annu. Rev. Phytopathol.* 44: 367-392.

Brown S.P., Cornforth D.M. & Mideo N. 2012. Evolution of virulence in opportunistic pathogens: generalism, plasticity, and control. *Trends. Microbiol.* 20: 336-342.

Brussow H., Canchaya C. & Hardt W.D. 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* 68: 560-602, table of contents.

Budroni S., Siena E., Dunning Hotopp J.C., Seib K.L., Serruto D., Nofroni C., Comanducci M., Riley D.R., Daugherty S.C., Angiuoli S.V., Covacci A., Pizza M., Rappuoli R., Moxon E.R., Tettelin H. & Medini D. 2011. *Neisseria meningitidis* is structured in clades associated with restriction modification systems that modulate homologous recombination. *Proc. Natl. Acad. Sci. U.S.A.* 108: 4494-4499.

Bull J.J. & Ebert D. 2008. Invasion thresholds and the evolution of non-equilibrium virulence. *Evol. Appl.* 1: 172-182.

Burdon J. & Chilvers G. 1982. Host density as a factor in plant disease ecology. *Annu. Rev. Phytopathol.* 20: 143-166.

Casadevall A. 2008. Evolution of intracellular pathogens. *Annu. Rev. Microbiol.* 62: 19-33.

Casadevall A. & Pirofski L.A. 2001. Host-pathogen interactions: the attributes of virulence. J. Infect. Dis. 184: 337-344.

Casadevall A. & Pirofski L.A. 2007. Accidental virulence, cryptic pathogenesis, martians, lost hosts, and the pathogenicity of environmental microbes. *Eukaryot Cell* 6: 2169-2174.

Case T.J., Holt R.D., McPeek M.A. & Keitt T.H. 2005. The community context of species' borders: ecological and evolutionary perspectives. *Oikos* 108: 28-46.

Chen I.C., Hill J.K., Ohlemuller R., Roy D.B. & Thomas C.D. 2011. Rapid range shifts of species associated with high levels of climate warming. *Science* 333: 1024-1026.

Coakley S.M., Scherm H. & Chakraborty S. 1999. Climate change and plant disease management. *Annu. Rev. Phytopathol.* 37: 399-426.

Comte L., Murienne J. & Grenouillet G. 2014. Species traits and phylogenetic conservatism of climate-induced range shifts in stream fishes. *Nature communications* 5.

Cornforth D.M. & Foster K.R. 2013. Competition sensing: the social side of bacterial stress responses. *Nature Rev. Microbiol.* 11: 285-293.

Davis A.J., Jenkinson L.S., Lawton J.H., Shorrocks B. & Wood S. 1998. Making mistakes when predicting shifts in species range in response to global warming. *Nature* 391: 783-786.

de Muinck E.J., Stenseth N.C., Sachse D., vander Roost J., Rønningen K.S., Rudi K. & Trosvik P. 2013. Context-dependent competition in a model gut bacterial community. *PLoS ONE* 8: e67210.

de Roode J.C., Helinski M.E., Anwar M.A. & Read A.F. 2005. Dynamics of multiple infection and within-host competition in genetically diverse malaria infections. *Am. Na.t* 166: 531-542.

Declercq A.M., Haesebrouck F., Van den Broeck W., Bossier P. & Decostere A. 2013. Columnaris disease in fish: a review with emphasis on bacterium-host interactions. *Vet. Res.* 44: 10.1186.

Decostere A., Haesebrouck F., Van Driessche E., Charlier G. & Ducatelle R. 1999. Characterization of the adhesion of *Flavobacterium columnare* (*Flexibacter columnaris*) to gill tissue. *J. Fish. Dis.* 22: 465-474.

Didelot X. & Falush D. 2007. Inference of bacterial microevolution using multilocus sequence data. *Genetics* 175: 1251-1266.

Diep D.B., Axelsson L., Grefsli C. & Nes I.F. 2000. The synthesis of the bacteriocin sakacin A is a temperature-sensitive process regulated by a pheromone peptide through a three-component regulatory system. *Microbiology* 146: 2155-2160.

Ebert D. & Herre E.A. 1996. The evolution of parasitic diseases. *Parasitology today* 12: 96-101.

Enright M.C. & Spratt B.G. 1999. Multilocus sequence typing. *Trends Microbiol.* 7: 482-487.

Feil E.J., Li B.C., Aanensen D.M., Hanage W.P. & Spratt B.G. 2004. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J. Bacteriol.* 186: 1518-1530.

Flyg C., Kenne K. & Boman H.G. 1980. Insect pathogenic properties of *Serratia marcescens*: phage-resistant mutants with a decreased resistance to *Cecropia* immunity and a decreased virulence to *Drosophila*. *Microbiology* 120: 173-181.

Frank S.A. 1996. Models of parasite virulence. *Q. Rev. Biol.* 71: 37-78.

Freckleton R.P. & Jetz W. 2009. Space versus phylogeny: disentangling phylogenetic and spatial signals in comparative data. *Proc. Biol. Sci.* 276: 21-30.

Freitag N.E., Port G.C. & Miner M.D. 2009. *Listeria monocytogenes*—from saprophyte to intracellular pathogen. *Nature Rev. Microbiol.* 7: 623-628.

Friedman H., Yamamoto Y. & Klein T.W. 2002. *Legionella pneumophila* pathogenesis and immunity. *Semin. Pediatr. Infect. Dis.* 13: 273-279.

Friman V., Ghoul M., Molin S., Johansen H.K. & Buckling A. 2013. *Pseudomonas aeruginosa* adaptation to lungs of cystic fibrosis patients leads to lowered resistance to phage and protist enemies. *PLoS ONE* 8: e75380.

Friman V., Lindstedt C., Hiltunen T., Laakso J. & Mappes J. 2009. Predation on multiple trophic levels shapes the evolution of pathogen virulence. *PLoS ONE* 4: e6761.

Friman V., Hiltunen T., Jalasvuori M., Lindstedt C., Laanto E., Ormala A., Laakso J., Mappes J. & Bamford J. 2011. High temperature and

bacteriophages can indirectly select for bacterial pathogenicity in environmental reservoirs. *PLoS ONE* 6: e17651.

Garritano S., Gemignani F., Voegele C., Nguyen-Dumont T., Le Calvez-Kelm F., De Silva D., Lesueur F., Landi S. & Tavtigian S.V. 2009. Determining the effectiveness of High Resolution Melting analysis for SNP genotyping and mutation scanning at the TP53 locus. *BMC Genet.* 10: 1.

Geerts A., Vanoverbeke J., Vanschoenwinkel B., Van Doorslaer W., Feuchtmayr H., Atkinson D., Moss B., Davidson T.A., Sayer C.D. & De Meester L. 2015. Rapid evolution of thermal tolerance in the water flea *Daphnia*. *Nat. Clim. Change* 5: 665-668.

Gonzalez-Gonzalez A., Sanchez-Reyes L.L., Delgado Sapien G., Eguiarte L.E. & Souza V. 2013. Hierarchical clustering of genetic diversity associated to different levels of mutation and recombination in *Escherichia coli*: a study based on Mexican isolates. *Infect. Genet. Evol.* 13: 187-197.

Guijarro J.A., Cascales D., García-Torrico A.I., García-Domínguez M. & Méndez J. 2015. Temperature-dependent expression of virulence genes in fish-pathogenic bacteria. *Front. Microbiol.* 6:700.

Gupte S., Kaur T. & Kaur M. 2015. Virulence Factors of Environmental Microbes in Human Disease. *J Trop Dis 3:161*.

Hall-Stoodley L. & Stoodley P. 2005. Biofilm formation and dispersal and the transmission of human pathogens. *Trends Microbiol.* 13: 7-10.

Harvell C.D., Mitchell C.E., Ward J.R., Altizer S., Dobson A.P., Ostfeld R.S. & Samuel M.D. 2002. Climate warming and disease risks for terrestrial and marine biota. *Science* 296: 2158-2162.

Hassan R., Scholes R. & Ash N. 2005. Ecosystems and human well-being: current state and trends, vol 1. Findings of the condition and trends working group of the Millennium Ecosystem Assessment. University of Pretoria Council for Science and Industrial Research UNEP World Conservation South Africa, South Africa Monitoring Centre,United Kingdom, Island Press Washington, D.C.

Heierson A., Siden I., Kivaisi A. & Boman H.G. 1986. Bacteriophage-resistant mutants of *Bacillus thuringiensi*s with decreased virulence in pupae of Hyalophora cecropia. *J. Bacteriol.* 167: 18-24.

Heino J., Virkkala R. & Toivonen H. 2009. Climate change and freshwater biodiversity: detected patterns, future trends and adaptations in northern regions. *Biological Reviews* 84: 39-54.

Hibbing M.E., Fuqua C., Parsek M.R. & Peterson S.B. 2010. Bacterial competition: surviving and thriving in the microbial jungle. *Nature Rev. Microbiol.* 8: 15-25.

Höckendorff S., Früh D., Hormel N., Haase P. & Stoll S. 2015. Biotic interactions under climate warming: temperature-dependent and species-specific effects of the oligochaete *Chaetogaster limnaei* on snails. *Freshwater Science* 34: 1304-1311.

Hoffmann A.A., Chown S.L. & Clusella-Trullas S. 2013. Upper thermal limits in terrestrial ectotherms: how constrained are they? *Funct. Ecol.* 27: 934-949.

Hromockyj A.E., Tucker S.C. & Maurelli A.T. 1992. Temperature regulation of *Shigella* virulence: identification of the repressor gene virR, an analogue of hns, and partial complementation by tyrosyl transfer RNA (tRNA1Tyr). *Mol. Microbiol.* 6: 2113-2124.

Huey R.B. & Hertz P.E. 1984. Is a jack-of-all-temperatures a master of none? *Evolution* 38: 441-444.

Huey R.B. & Kingsolver J.G. 1989. Evolution of thermal sensitivity of ectotherm performance. *Trends Ecol. Evol.* 4: 131-135.

Izem R. & Kingsolver J.G. 2005. Variation in continuous reaction norms: quantifying directions of biological interest. *Am. Nat.* 166: 277-289.

Johnson T.J. & Nolan L.K. 2009. Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* 73: 750-774.

Karvonen A., Rintamäki P., Jokela J. & 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.

Kellermann V., Loeschcke V., Hoffmann A.A., Kristensen T.N., Fløjgaard C., David J.R., Svenning J. & Overgaard J. 2012. Phylogenetic constraints in key functional traits behind species' climate niches: Patterns of desiccation and cold resistance across 95 *Drosophila* speies. *Evolution* 66: 3377-3389.

Ketola T., Mikonranta L., Zhang J., Saarinen K., Örmälä A., Friman V., Mappes J. & Laakso J. 2013. Fluctuating temperature leads to evolution of thermal generalism and preadaptation to novel environments. *Evolution* 67: 2936-2944.

Kimes N.E., Grim C.J., Johnson W.R., Hasan N.A., Tall B.D., Kothary M.H., Kiss H., Munk A.C., Tapia R. & Green L. 2012. Temperature regulation of virulence factors in the pathogen *Vibrio coralliilyticus*. *ISME* 6: 835-846.

Kingsolver J. G., Izem R., & Ragland G. J. 2004. Plasticity of Size and Growth in Fluctuating Thermal Environments: Comparing Reaction Norms and Performance Curves1. Integr. Comp. Biol. 44(6), 450-460.

Kingsolver J.G., Ragland G.J. & Diamond S.E. 2009. Evolution in a constant environment: thermal fluctuations and thermal sensitivity of laboratory and field populations of *Manduca sexta*. *Evolution* 63: 537-541.

Kunttu H., Jokinen E., Valtonen E. & Sundberg L. 2011. Virulent and nonvirulent *Flavobacterium columnare* colony morphologies: characterization of chondroitin AC lyase activity and adhesion to polystyrene. *J. Appl. Microbiol.* 111: 1319-1326.

Laanto E., Bamford J.K., Laakso J. & Sundberg L. 2012. Phage-driven loss of virulence in a fish pathogenic bacterium. *PLoS ONE* 7: e53157.

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

Laanto E., Hoikkala V., Ravantti J. & Sundberg L. 2017. Long-term genomic coevolution of host-parasite interaction in natural environment. *bioRxiv* : 101576.

Lafferty K.D. 2009. The ecology of climate change and infectious diseases. *Ecology* 90: 888-900.

Leveque C. 2003. Dynamics of communities and ecosystems. *Ecology From Ecosystem to Biosphere. EnfieldNH: Science Publishers Inc* : 216-221.

Luhring T.M. & DeLong J.P. 2016. Predation changes the shape of thermal performance curves for population growth rate. *Curr. Zool.:* zow045.

Mahlen S.D. 2011. *Serratia* infections: from military experiments to current practice. *Clin. Microbiol. Rev.* 24: 755-791.

Majeed H., Ghazaryan L., Herzberg M. & Gillor O. 2015. Bacteriocin expression in sessile and planktonic populations of Escherichia coli. *J. Antibiot.* 68: 52.

Majeed H., Gillor O., Kerr B. & Riley M.A. 2011. Competitive interactions in *Escherichia coli* populations: the role of bacteriocins. *ISME* 5: 71-81.

Martínez J.L. 2014. Short-sighted evolution of bacterial opportunistic pathogens with an environmental origin. *Front. Microbiol. 5:239*

Massey R.C., Buckling A. & ffrench-Constant R. 2004. Interference competition and parasite virulence. *Proc. Biol. Sci.* 271: 785-788.

Mataragas M., Metaxopoulos J., Galiotou M. & Drosinos E. 2003. Influence of pH and temperature on growth and bacteriocin production by *Leuconostoc mesenteroides L124* and *Lactobacillus curvatus L442*. *Meat Sci.* 64: 265-271.

Maurelli A.T., Blackmon B. & Curtiss R.,3rd. 1984. Temperature-dependent expression of virulence genes in *Shigella* species. *Infect. Immun.* 43: 195-201.

McMichael A.J., Woodruff R.E. & Hales S. 2006. Climate change and human health: present and future risks. *The Lancet* 367: 859-869.

McQuillan M.A. & Rice A.M. 2015. Differential effects of climate and species interactions on range limits at a hybrid zone: potential direct and indirect impacts of climate change. *Ecol. Evol.* 5: 5120-5137.

Merikanto I., Laakso J.T. & Kaitala V. 2014. Invasion ability and disease dynamics of environmentally growing opportunistic pathogens under outside-host competition. *PLoS ONE* 9: e113436.

Mikonranta L., Friman V. & Laakso J. 2012. Life history trade-offs and relaxed selection can decrease bacterial virulence in environmental reservoirs. *PLoS ONE*7: e43801.

Mikonranta L., Mappes J., Laakso J. & Ketola T. 2015. Within-host evolution decreases virulence in an opportunistic bacterial pathogen. *BMC Evol. Biol.* 15: 165.

Minter E.J., Lowe C.D., Brockhurst M.A. & Watts P.C. 2015. A rapid and cost-effective quantitative microsatellite genotyping protocol to estimate intraspecific competition in protist microcosm experiments. *Methods Ecol. EvoL.* 6: 315-323.

Mohammed H.H. & Arias C.R. 2014. Epidemiology of columnaris disease affecting fishes within the same watershed. *Dis. Aquat. Organ.* 109: 201-211.

Murugaiah C. 2011. The burden of cholera. *Crit. Rev. Microbiol.* 37: 337-348.

Nicolas P., Mondot S., Achaz G., Bouchenot C., Bernardet J.F. & Duchaud E. 2008. Population structure of the fish-pathogenic bacterium *Flavobacterium psychrophilum*. *Appl. Environ. Microbiol.* 74: 3702-3709.

Nilsen H., Sundell K., Duchaud E., Nicolas P., Dalsgaard I., Madsen L., Aspan A., Jansson E., Colquhoun D.J. & Wiklund T. 2014. Multilocus sequence

typing identifies epidemic clones of *Flavobacterium psychrophilum* in Nordic countries. *Appl. Environ. Microbiol.* 80: 2728-2736.

O'Keefe K.J., Morales N.M., Ernstberger H., Benoit G. & Turner P.E. 2006. Laboratory-dependent bacterial ecology: a cautionary tale. *Appl. Environ. Microbiol.*72: 3032-3035.

Olivares-Fuster O., Shoemaker C.A., Klesius P.H. & Arias C.R. 2007. Molecular typing of isolates of the fish pathogen, *Flavobacterium columnare,* by single-strand conformation polymorphism analysis. *FEMS Microbiol. Lett.* 269: 63-69.

Orvos D.R., Lacy G.H. & Cairns J. 1990. Genetically Engineered Erwinia carotovora: Survival, Intraspecific Competition, and Effects upon Selected Bacterial Genera. *Appl. Environ. Microbiol.* 56: 1689-1694.

Pallen M.J. & Wren B.W. 2007. Bacterial pathogenomics. *Nature* 449: 835-842.

Patz J.A., Campbell-Lendrum D., Holloway T. & Foley J.A. 2005. Impact of regional climate change on human health. *Nature* 438: 310-317.

Pritchard J.K., Stephens M. & Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

Pulkkinen K. & Ebert D. 2004. Host starvation decreases parasite load and mean host size in experimental populations. *Ecology* 85: 823-833.

Ramadhan A.A. & Hegedus E. 2005. Survivability of vancomycin resistant enterococci and fitness cost of vancomycin resistance acquisition. *J. Clin. Pathol.* 58: 744-746.

Riley M.A. & Wertz J.E. 2002. Bacteriocins: evolution, ecology, and application. *Ann. Rev. MicrobioL.* 56: 117-137.

Rowbotham T.J. 1980. Preliminary report on the pathogenicity of *Legionella pneumophila* for freshwater and soil amoebae. *J. Clin. Pathol.* 33: 1179-1183.

Ruosteenoja K., Jylhä K. & Kämäräinen M. 2016. Climate Projections for Finland Under the RCP Forcing Scenarios. *Geophysica* 51: 17-50.

Sadikot R.T., Blackwell T.S., Christman J.W. & Prince A.S. 2005. Pathogen–host interactions in *Pseudomonas aeruginosa* pneumonia. *Am. J. Respir. Crit. Care. Med.* 171: 1209-1223.

Santander J. & Robeson J. 2007. Phage-resistance of *Salmonella enterica* serovar *Enteritidis* and pathogenesis in *Caenorhabditis elegans* is mediated by the lipopolysaccharide. *EJB* 10: 627-632.

Schneck J. & Caslake L. 2006. Genetic diversity of *Flavobacterium columnare* isolated from fish collected from warm and cold water. *J. Fish. Dis.* 29: 245-248.

Schulte P.M., Healy T.M. & Fangue N.A. 2011. Thermal performance curves, phenotypic plasticity, and the time scales of temperature exposure. *Integr. Comp. Biol.* 51: 691-702.

Senadheera D.B., Cordova M., Ayala E.A., Chavez de Paz L.E., Singh K., Downey J.S., Svensater G., Goodman S.D. & Cvitkovitch D.G. 2012. Regulation of bacteriocin production and cell death by the VicRK signaling system in *Streptococcus* mutans. *J. Bacteriol.* 194: 1307-1316.

Shapiro B.J. & Polz M.F. 2014. Ordering microbial diversity into ecologically and genetically cohesive units. *Trends. Microbiol.* 22: 235-247.

Shapiro R.S. & Cowen L.E. 2012. Thermal control of microbial development and virulence: molecular mechanisms of microbial temperature sensing. *MBio* 3(5):e00238-12.

Smith J.M., Smith N.H., O'Rourke M. & Spratt B.G. 1993. How clonal are bacteria? *Proc. Natl. Acad. Sci. U.S.A.* 90: 4384-4388.

Smith J.M., Feil E.J. & Smith N.H. 2000. Population structure and evolutionary dynamics of pathogenic bacteria. *Bioessays* 22: 1115-1122.

Soto E., Mauel M.J., Karsi A. & Lawrence M.L. 2008. Genetic and virulence characterization of *Flavobacterium columnare* from channel catfish (Ictalurus punctatus). *J. Appl. Microbiol.* 104: 1302-1310.

Steinberg K.M. & Levin B.R. 2007. Grazing protozoa and the evolution of the *Escherichia coli O157:H7* Shiga toxin-encoding prophage. *Proc. Biol. Sci.* 274: 1921-1929.

Sturm A., Heinemann M., Arnoldini M., Benecke A., Ackermann M., Benz M., Dormann J. & Hardt W. 2011. The cost of virulence: retarded growth of *Salmonella Typhimurium* cells expressing type III secretion system 1. *PLoS Pathog.* 7: e1002143.

Sundberg L.R., Kunttu H.M. & Valtonen E.T. 2014. Starvation can diversify the population structure and virulence strategies of an environmentally transmitting fish pathogen. *BMC Microbiol.* 14: 1.

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

Suomalainen L., Tiirola M. & Valtonen E. 2006a. Chondroitin AC lyase activity is related to virulence of fish pathogenic *Flavobacterium columnare*. *J. Fish. Dis.* 29: 757-763.

Suomalainen L., Kunttu H., Valtonen E., Hirvela-Koski V. & Tiirola M. 2006b. Molecular diversity and growth features of *Flavobacterium columnare* strains isolated in Finland. *Dis. Aquat. Org.* 70: 55-61.

Susi H., Barrès B., Vale P.F. & Laine A. 2015. Co-infection alters population dynamics of infectious disease. *Nature communications* 6:5975.

Tamas I., Klasson L., Canback B., Naslund A.K., Eriksson A.S., Wernegreen J.J., Sandstrom J.P., Moran N.A. & Andersson S.G. 2002. 50 Million Years of Genomic Stasis in Endosymbiotic Bacteria. *Science* 296: 2376-2379.

Tekedar H.C., Karsi A., Gillaspy A.F., Dyer D.W., Benton N.R., Zaitshik J., Vamenta S., Banes M.M., Gulsoy N., Aboko-Cole M., Waldbieser G.C. & Lawrence M.L. 2012. Genome sequence of the fish pathogen *Flavobacterium columnare ATCC 49512*. *J. Bacteriol.* 194: 2763-2764.

Teska J.D. 1993. Assay to evaluate the reaction kinetics of chondroitin AC lyase produced by *Cytophaga columnaris*. *J. Aquat. Anim. Health* 5: 259-264.

Thomas M.B. & Blanford S. 2003. Thermal biology in insect-parasite interactions. *Trends Ecol. Evol.* 18: 344-350.

Tibayrenc M. & Ayala F.J. 2015. How clonal are *Neisseria* species? The epidemic clonality model revisited. *Proc. Natl. Acad. Sci. U.S.A.* 112: 8909-8913.

Toledo-Arana A., Dussurget O., Nikitas G., Sesto N., Guet-Revillet H., Balestrino D., Loh E., Gripenland J., Tiensuu T. & Vaitkevicius K. 2009.

The Listeria transcriptional landscape from saprophytism to virulence. *Nature* 459: 950-956.

Triyanto A. & Wakabayashi H. 1999. Genotypic diversity of strains of *Flavobacterium columnare* from diseased fishes. *Fish Pathol. 34*: 65-71.

Tylianakis J.M., Didham R.K., Bascompte J. & Wardle D.A. 2008. Global change and species interactions in terrestrial ecosystems. *Ecol. Lett.* 11: 1351-1363.

van Baarlen P., van Belkum A., Summerbell R.C., Crous P.W. & Thomma B.P. 2007. Molecular mechanisms of pathogenicity: how do pathogenic microorganisms develop cross-kingdom host jumps? *FEMS Microbiol. Rev.* 31: 239-277.

Von Graevenitz A. 1977. The role of opportunistic bacteria in human disease. *Ann. Rev. Microbiol.* 31: 447-471.

Vos M. 2011. A species concept for bacteria based on adaptive divergence. *Trends Microbiol.* 19: 1-7.

Walther B.A. & Ewald P.W. 2004. Pathogen survival in the external environment and the evolution of virulence. *Biological Rev.* 79: 849-869.

Wang X., Levy K., Son Y., Johnson M.W. & Daane K.M. 2012. Comparison of the thermal performance between a population of the olive fruit fly and its co-adapted parasitoids. *Biological Control* 60: 247-254.

Watling J.I., Romanach S.S., Bucklin D.N., Speroterra C., Brandt L.A., Pearlstine L.G. & Mazzotti F.J. 2012. Do bioclimate variables improve performance of climate envelope models? *Ecol. Model* 246: 79-85.

Wielgoss S., Bergmiller T., Bischofberger A.M. & Hall A.R. 2016. Adaptation to Parasites and Costs of Parasite Resistance in Mutator and Nonmutator Bacteria. *Mol. Biol. Evol.* 33: 770-782.

Wiens J.J. & Graham C.H. 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. *Annu. Rev. Ecol. Evol. Syst.* 519-539.

Willis C.G., Ruhfel B., Primack R.B., Miller-Rushing A.J. & Davis C.C. 2008. Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. *Proc. Natl. Acad. Sci. U. S. A.* 105: 17029-17033.

Wloch-Salamon D.M., Gerla D., Hoekstra R.F. & de Visser J.A. 2008. Effect of dispersal and nutrient availability on the competitive ability of toxin-producing yeast. *Proc. Biol. Sci.* 275: 535-541.

Yang L. N., Zhu W., Wu E., Yang C., Thrall P. H., Burdon J.J., Jin L.P., Shang L.P. and Zhan J. 2016. Trade-offs and evolution of thermal adaptation in the Irish potato famine pathogen *Phytophthora infestans.* Mol. Ecol. 25(16), 4047-4058.

Zhang J., Ketola T., Örmälä-Odegrip A., Mappes J. & Laakso J. 2014a. Coincidental loss of bacterial virulence in multi-enemy microbial communities. *PLoS ONE* 9: e111871.

Zhang J., Laakso J., Mappes J., Laanto E., Ketola T., Bamford J.K., Kunttu H. & Sundberg L.R. 2014b. Association of colony morphotypes with virulence, growth and resistance against protozoan predation in the fish pathogen *Flavobacterium columnare*. *FEMS Microbiol. Ecol.* 89: 553-562.

ORIGINAL PAPERS

I

A MULTILOCUS SEQUENCE ANALYSIS SCHEME FOR CHARACTERIZATION OF FLAVOBACTERIUM COLUMNARE ISOLATES

by

Roghaieh Ashrafi, Katja Pulkkinen, Lotta-Riina Sundberg, Nina Pekkala, Tarmo Ketola 2015

BMC Microbiology: 15:243.

Ashrafi et al. BMC Microbiology (2015) 15:243
DOI 10.1186/s12866-015-0576-4

BMC Microbiology

RESEARCH ARTICLE Open Access

A multilocus sequence analysis scheme for characterization of *Flavobacterium columnare* isolates

Roghaieh Ashrafi[1*], Katja Pulkkinen[2], Lotta-Riina Sundberg[1], Nina Pekkala[2] and Tarmo Ketola[1]

Abstract

Background: Columnaris disease caused by *Flavobacterium columnare* is a serious problem in aquaculture, annually causing large economic losses around the world. Despite considerable research, the molecular epidemiology of *F. columnare* remains poorly understood.

Methods: We investigated the population structure and spatiotemporal changes in the genetic diversity of *F. columnare* population in Finland by using a multilocus sequence typing (MLST) and analysis (MLSA) based on DNA sequence variation within six housekeeping genes. A total of 83 strains of *F. columnare* were collected from eight different areas located across the country between 2003 and 2012.

Results: Partial sequencing of six housekeeping genes (*trpB*, *tuf*, *atpA*, *rpoD*, *gyrB* and *dnaK*) revealed eight sequence types and a moderate level of genetic diversity (H = 0.460). Phylogenetic analysis of the concatenated protein-encoding gene sequence data (ca. 3,509 nucleotides) formed two lineages, which could be further divided into five clusters. All analysed *F. columnare* strains appeared to have a genetic origin distinct from that of another important fish pathogen form the genus Flavobacterium, *F. psychrophilum*. Although the value of the index of association between alleles, 0.292 ($P < 0.001$), supports some degree of clonality for this species in Finland, recombination has introduced molecular diversity to the population almost three times more than mutation.

Conclusion: The results suggest that Finnish *F. columnare* strains have an epidemic population structure followed by clonal expansion of successful genotypes. Our study with reproducible methodology and comparable results establishes a robust framework for the discrimination and phylogenetic analysis of *F. columnare* isolates, which will help to improve our understanding about geographic distribution and epidemiology of columnaris disease.

Keywords: Flavobacterium columnare, MLST/MLSA scheme, ClonalFrame, Recombination rate, Clonality

Background

Flavobacterium columnare is a Gram-negative bacterium belonging to the family *Flavobacteriaceae* (phylum Bacteroidetes) [1]. Columnaris disease caused by *F. columnare* represents a continuous threat to the growing aquaculture industry worldwide. It has been ranked as the second most common disease of the channel catfish (*Ictalurus punctatus*) industry in the United States [2, 3]. The bacterium is capable of causing infections in both warm and cold water species of fish [4], and it infects fish species around the world, including carp, channel catfish, goldfish, eel, perch, tilapia, pike perch, rainbow trout, brown trout, salmon, tiger muskellunge and walleye [2, 5]. *F. columnare* causes epidermal infections affecting gills, skin and fins of the fish, producing either acute or chronic infections, depending on the virulence and genetic group (genomovar) of the strain, as well as on environmental [6] and host-related factors [2, 5, 7].

F. columnare has high phenotypic homogeneity, therefore strain characterization by standard biochemical tests is not appropriate [8, 9]. However, *F. columnare* has been divided into three genomovars (I, II, III) using analysis of 16S rDNA by restriction-fragment length polymorphism (16S rDNA-RFLP) [10]. A recent study further increased

* Correspondence: roghaieh.ashrafi@jyu.fi
[1]Centre of Excellence in Biological Interactions, Department of Biological and Environmental Science, University of Jyvaskyla, P.O. Box 35, Jyväskylän yliopisto FI-40014, Finland
Full list of author information is available at the end of the article

the resolution of the method to identify a new genomovar (II/III) [11]. Of these, genomovars I and II have been reported in Europe as either common European (genomovar I) or likely imported (genomovar II or Asian type strains) [12]. To obtain higher resolution on genetic diversity of *F. columnare,* several other molecular typing approaches have been used, including single-strand conformation polymorphism (SSCP) [13], amplified fragment length polymorphism (AFLP) [14], pulsed-field gel electrophoresis (PFGE) [3], automated ribosomal intergenic spacer analysis (ARISA) [8] and internal spacer region-single strand conformation polymorphism analysis (ISR-SSCP) [15]. Although the overall discriminatory power of these methods is high, they can suffer from poor inter-laboratory interpretability, and they are not suitable for population structure studies. Moreover, these genetic markers accumulate genetic variation rapidly, which can interfere with investigating evolutionary phylogenetic relationships or global epidemiology between closely related species of bacteria [9, 16] .

In 2012, the complete genome of *F. columnare* was published [17], making it possible to compare genes from individual isolates for developing multilocus sequence typing (MLST) and multilocus sequence analysis (MLSA) schemes. MLST/MLSA schemes provide portable, universal, highly discriminatory and unambiguous data [18–20]. Because this method indexes variation in housekeeping genes that have a relatively slow evolutionary rate, it has been widely used to infer population genetic structure of several different bacterial groups [19, 21–23].

The MLST method typically uses variation in four to seven housekeeping gene sequences to characterize isolates of bacterial species. The allele-based MLST method relies on allelic profiles (the specific combination of alleles for each isolate) and sequence type designations (isolates with an identical allelic profile) to estimate relatedness among isolates and so it ignores the number of nucleotide differences between alleles. The MLSA method, however, relies directly on nucleotide sequences; it uses concatenated sequences of fragments of housekeeping genes to determine genus-wide phylogenetic relationships. Nevertheless, it has been shown that MLSA can also provide robust resolution at the intra-specific level, especially when inadequate phylogenetic resolution prevents MLST from distinguishing phylogenetically closely related strains [24, 25].

Since the first columnaris outbreak in Finland in 1984, *F. columnare* has been reported as a major threat to salmonid fish farming, particularly rainbow trout (*Oncorhynchus mykiss*) [7]. Despite its importance as a fish pathogen, the genetic diversity and population structure of *F. columnare* are poorly known. The genetic characterization of *F. columnare* is essential not only to develop appropriate management strategies to minimize the risk of columnaris disease in Finnish fish farms but also to better understand host specificity, pathogenicity, and distributional pattern of this bacterium, which is critical for understanding the emergence of columnaris outbreaks worldwide. This prompted us to develop the first MLST/MLSA scheme for this species in order to investigate the population structure of *F. columnare* strains isolated from different geographic areas in Finland.

Materials and Methods

Bacterial strains and culture conditions

From 2003 to 2012, 83 *F. columnare* strains were obtained from nine different fish species ($n = 59$) and water samples ($n = 28$), in eight geographic locations in Finland comprising four northern and four southern locations (Additional file 1 and Fig. 1). It is worth noting that all Finnish *F. columnare* isolates have been assigned to genomovar I, [8]. The *F. columnare* type strain NCIMB 2248^{T} isolated in the USA and two reference strains JIP39/87 and ATCC49512, both isolated in France, were also included in the sample collection. The sequences for strain ATCC 49512 has been retrieved from GenBank using their accession numbers. All the Finnish strains were originally isolated using standard culture methods including Shieh medium [26], Shieh medium supplemented with tobramycin [27], or AO-agar [28].

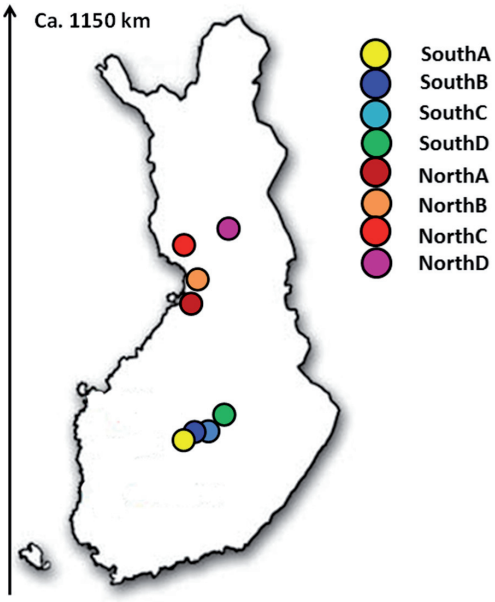


Fig. 1 Location of sample sites in Finland for the 83 *F. columnare* isolates used in this study

Pure cultures were stored frozen at −80 °C in Shieh medium containing 10 % of glycerol and 10 % of fetal calf serum. Before genomic DNA extraction, the strains were revived from the stocks in 3 ml of Shieh medium for 24 h while shaking (200 rpm) at 25.5 °C. Overnight cultures were diluted into fresh Shieh medium (1:10) and allowed to regrow at 25.5 °C on the shaker for another 24 hours.

DNA extraction, PCR amplification and sequencing of housekeeping genes

Bacterial cells were harvested from fresh cultures by centrifugation at 6000 x g for 10 min. Bacterial genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega, USA). The candidate genes for this study (*trpB*, *gyrB*, *dnaK*, *fumC*, *murG*, *rplB*, *recA*, *tuf*, *atpA*, *glyA*, *rpoD*, 16S rDNA and *fstQ*) were chosen based on the previously published MLST scheme for *Flavobacterium psychrophilum* by Nicolas et al. [22]. We used the corresponding sequences derived from the whole genome sequence of *F. columnare* ATCC 49512 (= CIP 103533 = TG 44/87) [17] to redesign the primers for *F. columnare*. The primers were designed using Primer3 (v.0.4.0) [29, 30] and checked for specificity using the NCBI Primer-BLAST tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). All primers used in this study are listed in Additional file 2.

PCR reactions were performed in a total volume of 20 μl containing 1X Phusion Flash PCR Master Mix (Thermo Scientific), 0.5 μM forward primer, 0.5 μM reverse primer, and 100 ng of genomic DNA. PCR reactions were performed on a BioRad 1000C thermal cycler, under the following conditions: 98 °C for 30 s, followed by 30 cycles at 98 °C for 10 s, 62 °C for 20 s, and 72 °C for 15 s, and a final extension at 72 °C for 5 min. Five microliters of the PCR products were run on a 1.5 % agarose gel to verify correct amplification. The PCR products were purified using 10 U of Exonuclease I and 1 U of FastAP^{TM} Thermosensitive Alkaline Phosphatase (Fermentas GmbH, Germany) for 15 min at 37 °C, followed by enzyme inactivation for 15 min at 85 °C.

The purified PCR products were then sequenced with the same primers used in amplification using Big Dye Terminator (v3.1) Cycle Sequencing Kit (Applied Biosystems). Briefly, each 20 μl sequencing reaction mixture contained 2 μl of PCR amplicon, 0.16 μM of either forward or reverse PCR primer, 0.5 μl of BigDye Ready Reaction Mix, and 1 X sequencing buffer. The sequencing reaction conditions were as follows: 30 cycles of denaturing at 96 °C for 10 s, annealing at 50 °C for 5 s, and extension at 60 °C for 4 min. The sequencing products were purified using ethanol/EDTA/sodium acetate precipitation. Sequencing was performed on an ABI 3130xl 16-capillary automated genetic analyzer.

MLST data treatment

The raw sequences were manually inspected and corrected using Sequencher 5.0.1 (Gene Codes, Ann Arbor, MI). The consensus sequence for each gene was determined by alignment of the forward and reverse sequences using BioEdit Sequence Alignment Editor Version 7.0.5.3 [31] (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). The consensus sequences were aligned with Clustal W implemented in MEGA v5.2 [32]. For each housekeeping gene, allele numbers were assigned to unique sequences using the non-redundant databases (NRDB) program (http://www.mlst.net/). For each bacterial isolate, an allelic profile was determined as the combination of alleles at the six loci selected for final analyses. Finally, strains with an identical allelic profile were assigned to the same sequence type (ST) (Additional file 1).

Phylogenetic analysis (MLSA)

A Neighbour-Joining dendrogram was constructed using individual and concatenated sequences of the six genes that were selected for final analyses. For comparison, the *F. columnare* type strain NCIMB 2248^{T}, ATCC49512 and strain JIP39/87 were also included in the phylogenetic analysis. F. *psychrophilum* JIP 02/86 strain from France was used as an outgroup strain. MEGA v5.2 [32] was used to evaluate the model for nucleotide substitutions at each protein-coding locus and to construct a phylogenetic tree. The best model, having the lowest Bayesian Information Criterion (BIC) value, was used to generate the Neighbour-Joining tree based on 1000 replicates. The Tamura-Nei model plus a gamma distribution (T93 + G) model was used to infer the dendrogram for the concatenated sequences. A Chi-square test was employed to evaluate whether the presence of MLSA phylogenetic clusters is explained by regional (north and south) separation. Logistic regression was used to explore trends in outbreak likelihood over time for each clade. Moreover, the distribution of MLSA clusters (genotypes) across sampling origin (water or fish) were determined using Pearson's chi-square statistic, or, where sample sizes were small, Fisher's exact test. A P-value less than or equal to 0.05 was defined as statistically significant.

Computational analysis

Analysis of DNA sequence variation of the housekeeping gene loci, including the number of alleles for each locus, GC (guanine + cytosine) content, the number of segregating sites (S), allelic diversity, and the nucleotide diversity (Pi), was carried out using DNAsp genetic software v5.10.01 [33]. MEGA v5.2 was used to perform Tajima's D test [34]. The START 2 package [35] was used to determine the ratios of non-synonymous to synonymous

substitutions (dN/dS) for each locus. The range of intraspecific sequence similarity (%) for each gene was resolved using BioEdit program [31].

Population genetic and recombination analyses

Evidence for clonal or recombining population structures can be estimated by assessing the level of linkage between alleles at different loci. To test the null hypothesis, i.e. whether alleles of the six MLST loci used in the analyses are independent (linkage equilibrium), the IAS (standardized index of association) values were calculated with the START2 program. The test was performed first for the entire data set of 83 isolates and then for only eight STs to avoid biased results due to unequal sample sizes in different STs. IAS values significantly different from 0 indicate that a population is clonal (linkage disequilibrium), whereas non-significant values indicate a recombining population structure [36, 37].

The concatenated sequence data (for the six core genes) were formatted using xmfa2struct (http://www.xavierdidelot.xtreemhost.com/ClonalFrame.htm). Bayesian approach model with STRUCTURE version 2.3 [38] was used to determine the levels of inter-lineage recombination and the underlying population structure present in our data. Ten independent runs were performed for each value of the number of ancestral populations (*K)* ranging from 2 to 6. STRUCTURE was run for 500,000 Markov Chain Monte Carlo (MCMC) iterations following 250,000 burn-in iterations. The linkage model that reconstructs ancestral populations from DNA polymorphism data was used. The STRUCTURE Harvester [39], which implements the Evanno method [40], was used to identify the most probable groups (*K*) that best fit the data.

In order to estimate the mutation and recombination rates in *F. columnare*, we also performed recombination analysis using ClonalFrame v1.1 [41]. Three independent runs of ClonalFrame were performed, each consisting of 500,000 MCMC iterations and 250,000 burn-in iterations. ClonalFrame was also used to compare independent runs by the method of Gelman and Rubin [42]. ClonalFrame estimates ρ/θ, which measures the relative frequency of occurrence of recombination and mutation in the history of the lineage, and r/m which measures the relative impact of recombination and mutation in the genetic diversification of the lineage. The values of ρ/θ and r/m for all 83 isolates and for each lineage were calculated by extracting the numbers of mutation events, recombination events, and substitutions introduced by recombination from the ClonalFrame output.

Results

MLST

Thirteen housekeeping genes were successfully amplified for the 83 *F. columnare* isolates from Finland and for both the type strain NCIMB 2248^{T} and the isolate JIP39/87. The housekeeping genes *ftsQ, glyA, murG, recA, fumC, 16S rDNA*, and *rplB* showed identical sequences among all 83 Finnish strains and were excluded from further analyses. Six remaining housekeeping genes, *trpB*, *tuf*, *atpA*, *rpoD*, *gyrB* and *dnaK* were selected to develop the MLST/MLSA scheme. The range of intraspecific sequence similarity (%) calculated for each gene showed that *trpB* had the highest level of sequence polymorphism between the strains (94.5 % similarity), followed by *rpoD* (97.4 %). Whereas the lowest levels of inter-strain sequence variation were found for *tuf* (99.5 %), *atpA* (99.5 %), gyrB (99.6 %), and *dnaK* (99.8 %). The GC content for individual genes exhibited very little variation between the strains (≤ ±0.5 %) (Table 1). The average overall GC content for the six genes in all 83 Finnish strains was 39 %, slightly higher than the overall GC content (32.5 %) for the *F. columnare* ATCC 49512 genome [17].

Using the MLST method, the *F. columnare* strains were divided into 8 different sequence types (STs). The number

Table 1 Characteristics of the six loci used in the analyses for the 83 *F. columnare* strains

Locus	Fragment size(bp)	No. of alleles	Allelic diversity	GC	S	dN/dS	∏	Tajima's D
trpB	626	5	0.72	0.441	49	0.046	0.0283	2.45[a]
rpoD	386	3	0.53	0.348	10	0.000	0.0117	3.38[a]
tuf	608	4	0.29	0.399	3	0.126	0.0005	−0.85[b]
atpA	645	4	0.56	0.425	3	0.000	0.0015	0.52[b.]
dnak	643	2	0.27	0.362	1	0.000	0.0004	0.43[b.]
gyrB	598	3	0.37	0.369	2	0.000	0.0006	−0.07[b.]
Concatenated	3509	9	0.74	0.390	68	-	-	-
Average	-	-	0.4604	-	-	-	-	-

GC Guanine + Cytosine content; *S* number of segregating sites; *dN/dS* the ratio of non-synonymous to synonymous substitutions; *∏* nucleotide diversity; *Tajima's D* neutrality test values

[a]statistically significant, and [b]non-significant value

of alleles at each locus ranged from 2 (*dnaK)* to 5 (*trpB*), and the number of variable sites varied between 1 (*dnaK*) and 49 (*trpB*) (Table 1). No insertions or deletions were detected within the loci. The allelic diversity ranged from 0.27 for *dnaK* to 0.72 for *trpB*, with an average value of 0.46 (Table 1). The average pairwise nucleotide diversity π for all genes was 4.2 %, with π-values for individual loci ranging from 0.0004 for *dnaK* to 0.0283 for *trpB* (Table 1). Tajima's D values for loci *tuf*, *atpA*, *dnaK*, and *gyrB* were nonsignificant, supporting neutrality of the alleles of these genes (i.e., evolution by random processes) (Table 1). In contrast, *trpB* and *rpoD* both showed a significantly positive value for Tajima's D, indicating a decrease in population size, balancing selection [43], or subdivision of the population [44].

Phylogenetic analysis (MLSA)

In total, 3,509 bp of sequence was analyzed for each strain. The phylogenetic tree based on the concatenated sequence shows two major lineages (I and II) which can be further divided into five MLSA clusters (Fig. 2 and Additional file 3). Both lineages are strongly supported by high bootstrap values. The clusters within lineages associated uniformly with the ARISA genotypes of the strains (for more information, see Additional file 4). Therefore, we decided to designate the MLSA clusters in accordance with the corresponding ARISA genotypes [8]. Lineage I includes two clusters, corresponding to ARISA genotypes C and E. Cluster C (ST2) contains 36 strains isolated from six locations: NorthA, NorthB, NorthC, SouthA, SouthC, and SouthD (see Figs. 1 and 2). Cluster E (STs 4 and 8) includes 20 strains isolated from SouthC. Lineage II includes three clusters, corresponding to ARISA genotypes G, A and H. Cluster G (STs 1 and 7) contains strains isolated from NorthB, NorthC, SouthB and SouthC. Cluster A (STs 3 and 6) includes strains isolated from NorthB, NorthD, SouthB, and SouthC. Cluster H (ST5) contains three strains isolated form NorthB and SouthC. Statistical analysis showed that there was no significant regional (north and south) clustering of *F. columnare* in Finland, but the occurrence of cluster E isolates was more frequent in more recent years (Table 2). No other trends were apparent over the study period 2003–2012 (Table 2). Cluster H strains (total of 3 isolates) were only found in years 2003 and 2012, which most likely results from random sampling of a very rare genotype, rather than any systematic

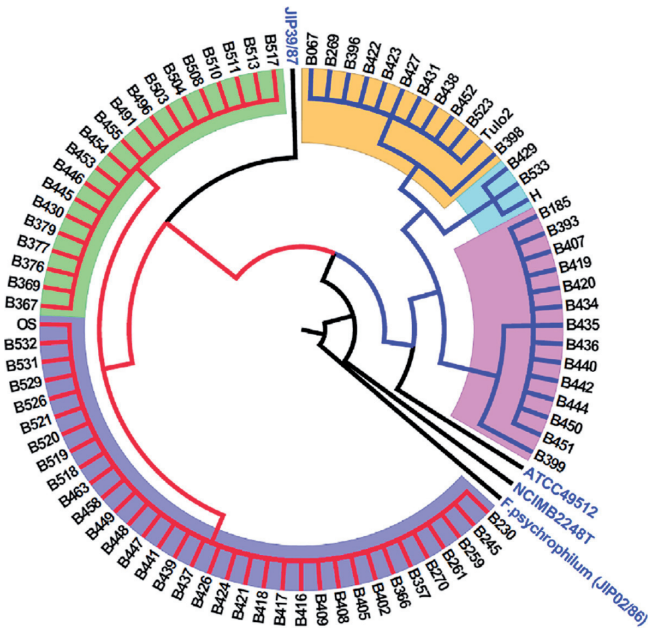


Fig. 2 Phylogenetic tree of *F. columnare* strains based on a concatenated 6-housekeeping gene (*trpB, tuf, atpA, rpoD, gyrB and dnaK*). Strict consensus Neighbour Joining tree for the 83 *F. columnare* strains from Finland and three other strains, including the type strain NCIMB 2248^{T}, ATCC 49512, and strain JIP39/87. The numbers at the nodes represent levels (%) of bootstrap support from 1000 resampled datasets. The sequence of the *Flavobacterium psychrophilum* (JIP 02/86) strain from France was used as an out-group. MLSA clusters (designated in accordance with the corresponding ARISA genotypes) C, E, A, H, and G are marked with purple, green, orange, light blue, and pink, respectively. Branches from Finnish lineage I strains are coloured red, those from lineage II are coloured blue, and those from other countries are coloured black

reasons. Furthermore, no association was found between MLSA clusters and the sampling origin including water and fish species (P = 0.52, Additional file 1).

The phylogenetic and Bayesian analyses clearly indicated that *F. columnare* strains are genetically distinct from *F. psychrophilum.* In addition, the early-branching type strain (NCIMB 2248^T) from the USA appears to be highly divergent from the Finnish *F. columnare* strains. However, the reference strains ATCC 49512 and JIP 39/87, both from France, were phylogenetically close to isolates from Finland and clustered with the Finnish strains.

Population genetic and recombination analysis

Results from the ClonalFrame analysis (based on both the 83 Finnish isolates and eight STs) showed that mutations have occurred approximately seven times more frequently than recombination ($\rho/\theta \approx 0.14$; with 95 % confidence interval of 0.03-0.9). However, recombination has had a significant effect in the evolutionary process as shown by the r/m value of approximately three (with 95 % confidence interval of 0.44-6.6). This indicates that even though recombination has been less frequent than mutation, each recombination event has introduced almost three times more substitutions than mutation. According to the results extracted from the ClonalFrame output, the role of recombination seems to be uneven across the two MLSA lineages. The relative frequency of occurrence of recombination *versus* mutation (ρ/θ) was 0.2 for lineage I, and 0.08 for lineage II. Whereas the relative impact of recombination *versus* point mutation (r/m) was 3.14 for lineage I and 1.43 for lineage II. However, even in case of high recombination rate, analysis of all 83 isolates yielded an IAS (standardized index of association) value of 0.292 ($P < 0.001$), indicating persistence of identical clonally expanded sequence types over sampling years. Moreover, IAS values calculated separately for lineages I and II were also significantly different from zero (0.81 and 0.65, respectively, $P < 0.001$ for both). When only one representative of each sequence type was included, the overall IAS value was decreased to 0.198, but remained significant (P = 0.03), meaning that recombination was not sufficient to break down linkage disequilibrium. The analysis with STRUCTURE detected two subpopulations (best K = 2), denoted as lineages I and II (Fig. 3). Lineage II forms a genetically homogeneous group. However, lineage I represent a genetically heterogeneous group with genetic material imported from lineage II.

Discussion

In this paper, we report the results of the first MLST/MLSA scheme developed to determine the genetic variability and population genetic structure of the *F. columnare* strains isolated from different locations in Finland. According to classification based on 16 s rDNA-RFLP analysis, all Finnish isolates studied thus far belong to genomovar I [8] (see Additional file 5). This is likely explained by the fact that strains from genomovars II and III have higher temperature requirements and thus may not be able to tolerate colder temperatures in Finland; it has been shown that these genomovars thrive at higher temperatures but not at 15 °C [10]. Therefore, because we focus here on the Finnish population, isolates from genomovars II and III were not included in the study. Our results demonstrated that Finnish *F. columnare* strains were separated into different STs (by MLST) and phylogenetic clusters (by MLSA) irrespective of the geographic origin, host species or year of isolation. Previous studies have reported a similar lack of relationship between genotypes and isolation source including both host and geographical origin. Based on three genotyping methods including rDNA-RFLP, ISR and AFLP, Arias *et al.* [14] analysed 30 *F. columnare* isolates from four countries and found that these strains did not cluster according to host species or geographic origin. Similarly, an analysis of ten *F. columnare* isolates originating from cold and warm waters using RFLP analysis showed no relationship between geographic origin and genomic types [4]. One possible explanation for the lack of association between STs and their geographic origin (i.e. genetically similar populations occur in geographically distinct areas) is that the transportation of fish stocks or equipment between fish farms and natural exchanges such as wild bird could contribute to the spread of *F. columnare* strains between regions [45].

Table 2 Proportion of isolates of *F. columnare* from each region (south and north) and year of isolation within the MLSA clusters (%)

MLSA clusters	No. of isolates	North	South	P value	2003	2005	2006	2007	2008	2009	2010	2011	2012	P value
C	34	58.8	41.2	0.182	0	0	23.5	14.7	8.8	20.6	17.6	2.9	11.8	0.88
E	20	0	100	-	5.0	0	0	10.0	15.0	0	25.0	5.0	40.0	0.003*
G	14	57.1	42.9	0.593	0	14.3	7.1	35.7	14.3	7.1	21.4	0	0	0.145
H	3	33.3	66.7	0.564	66.7	0	0	0	0	0	0	0	33.3	-[a]
A	12	50	50	1.00	9.1	0	18.2	27.3	0	27.3	18.2	0	0	0.118

Observations for regional sources were compared by nonparametric chi-square test (* = $P < 0.05$). Observations for year of isolation were tested with logistic regression where year was fitted as a continuous covariate

[a]Due to the small number of cluster H isolates in our dataset ($n = 3$), we did not include cluster H in logistic regression

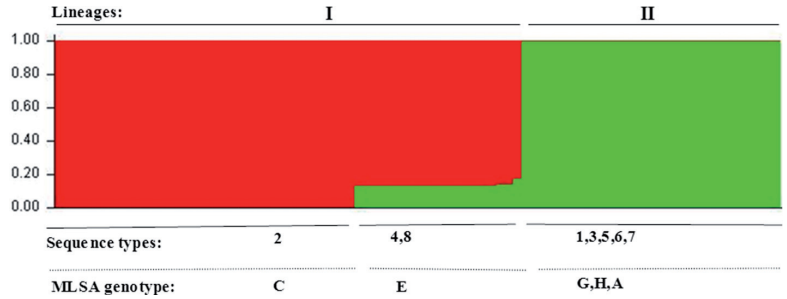


Fig. 3 Estimated *F. columnare* population structure using STRUCTURE version 2.3. Ancestral population size of 2 (K = 2) was chosen as the best fit for the current data. Upper labels indicate the ancestral populations (lineages); sequence types and MLSA clusters of the isolates are listed at the bottom

The six loci used in the analyses (*trpB*, *tuf*, *atpA*, *rpoD*, *gyrB*, *dnaK*) were successfully amplified and sequenced for all isolates. Based on MLST statistical analyses and concatenated phylogenetic analyses, all 83 isolates were grouped into eight STs and two major lineages (I and II) that were further assigned into five clusters (C, E, A, H, and G) (Fig. 1). Overall, the isolates were commonly designated to cluster C while cluster H strains were the rarest ones. We found that the environmental isolates of *F. columnare* are phylogenetically clustered with the epidemic strains, indicating that the environmental population may be the source of epidemic strains and *vice versa*. Non-synonymous base substitutions in gene sequences were rare ($dN/dS < 1$) (Table 1), indicating purifying selection against amino acid changes. This verifies that the identified sequence variability is selectively neutral at the protein level, making the sequences good candidates for multilocus sequence typing and analysis. However, an excess of polymorphisms in the *trpB* and *rpoD* sequences can also indicate balancing selection, supported also by the positive Tajima's D values. This finding is consistent with two clearly distinct lineages in the Finnish sample of *F. columnare*. It is also possible that the differences in *trpB* and *rpoD* are caused by both purifying selection (limiting the amino acid changes due to functional constraints on housekeeping genes), and potential balancing selection (maintaining the genetic polymorphism within a population).

The individual trees based on *trpB* and *rpoD* were found to provide higher phylogenetic resolution than corresponding *tuf*, *atpA*, *dnaK*, and *gyrB* gene sequences (Additional file 1). Although cluster H strains could only be resolved by using the concatenated sequences of all six housekeeping gene sequences, the tree constructed from the *trpB* gene alone could identify four major clusters (A, C, E, and G,) resolved by MLSA analysis. Moreover, to determine whether the *trpB* gene used in the concatenated sequence influenced tree topology, it was compared to the tree constructed on five MLSA gene sequences (i.e. excluding *trpB* gene). Adding the *trpB* sequence to the MLSA-concatenated sequence (*rpoD, dnaK, tuf, gyrB,* and *atpA)* thus had minor effect on the clustering of strains. Furthermore, we found that only one strain, B399, had discordant *trpB*-based identification (Additional files 3 and 6). Thereby, we propose that the *trpB* gene can serve as a reliable, cost-effective, and quick molecular marker to differentiate closely related *F. columnare* strains.

Significant linkage disequilibrium between the *F. columnare* MLST alleles suggests a non-random distribution of alleles and a clonal population structure where diversity increases by the accumulation of point mutations. Indeed, recovering identical or similar STs from different geographic origins suggests that *F. columnare* strains may undergo clonal expansion. This conclusion is also supported by the observation that the same genotypes are repeatedly isolated over the years. On the other hand, the low IAS values (although significant) suggest that recombination may have occurred among these strains, but not frequently enough to completely remove linkage disequilibrium. Finding linkage disequilibrium between the alleles is not surprising, considering that an allele must change at least 20 times more frequently by recombination than by point mutation to achieve random assortment within a bacterial population [37]. Based on both STRUCTURE and ClonalFrame analyses, however, recombination events have occurred both within and across lineages. Considering both the relative frequency of occurrence of recombination and mutation ($\rho/\theta \approx 0.14$) and the relative effect of recombination and mutation in genetic diversification ($r/m \approx 3$), the Finnish *F. columnare* strains seem to have a moderate recombination rate. The observed genetic exchange and recombination within and across the two lineages is irreconcilable with a strictly clonal population

structure. Therefore, we suggest that the *F. columnare* population structure in Finland follows an epidemic population structure where there is a background level of frequent recombination with consecutive clonal expansion of one or a few fit genotypes [46], similar to that of other bacterial species such as *Escherichia coli* [47], *Vibrio parahaemolyticus* [48] and *F. psychrophilum* [49]. Compared with other fish pathogens, the estimated recombination rate for the Finnish *F. columnare* strains seems to be slightly higher than in *Tenacibaculum maritimum* (~2.4:1) [50], and is clearly lower than in the highly recombinogenic bacterium *F. psychrophilum* (~26:1) [51]. However, as our data set consists of local samples, we cannot estimate the recombination rate of *F. columnare* at a global (species) level.

Our results further demonstrate that lineage I displays more evidence of recombination than lineage II ($r/m \approx 3.14$ and 1.43, respectively). We also show that only lineage I seems to exhibit mixed ancestries (lower than 20 %) suggesting that limited, unidirectional inter-lineage admixture has taken place (Fig. 3). Such infrequent recombination between lineages is also consistent with our results from the phylogenetic analysis, indicating divergence of the two lineages (I and II). A low level of inter-lineage recombination, even among sympatric strains, coupled with high levels of intra-lineage recombination suggests that natural barriers could prevent recombination across the two lineages. Since the statistical analysis indicates that the clustering of the lineages is not caused by geography, the barrier against inter-lineage gene transfer may be caused by other, yet undefined, factors.

Conclusions

Our MLST/MLSA scheme data demonstrate that both recombination and clonality play a role in shaping the population structure of *F. columnare* in Finland. The population structure of the Finnish *F. columnare* is probably semi-clonal which diversifies with moderate, but variable, recombination rate. Limited association of genotypes with geography or year of isolation indicates that columnaris outbreaks in Finland are caused by continuous co-circulation of *F. columnare* strains. Furthermore, recovering identical genotypes from both fish and from the environment confirms that environmentally persistent strains are also epidemiologically important.

Availability of supporting data

DNA sequences of all STs have been submitted to the European genetic database EMBL under the following accession numbers: *trpB* (LN624115, LN624121, LN624122, LN624123, and LN624124), *rpoD* (LN624116, LN624125, and LN624126), *atpA* (LN624118, LN624133, LN624134, and LN624135), *tuf* (LN624120, LN624130, LN624131, and LN624132), *gyrB* (LN624119, LN624128, and LN624129), *dnaK* (LN624117, LN624127).

Additional files

Additional file 1 Site, year of isolation, source of isolation (fish or water), location of isolation, host species, sequence type (ST), and allelic profile data for the 83 F. columnare strains from Finland analyzed by MLST. (DOCX 46 kb)

Additional file 2 Target genes and primers for the housekeeping genes of *F. columnare*. The loci used for the MLST/MLSA scheme are shown in bold font. Length refers to the length of the target sequence. * Reference for 16S rDNA primers [52]. (DOCX 20 kb)

Additional file 3 Neighbor Joining phylogenetic trees based on the individual sequences of six MLST loci (*trpB, rpoD, gyrB, dnaK, atpA*,and *tuf*). MEGA v5.2 was used to evaluate the models for nucleotide substitution for each protein-coding locus and to construct the phylogenetic trees for Finnish *F. columnare* strains. The *F. columnare* type strain NCIMB 2248^T isolated in the USA and two reference strains JIP39/87 and ATCC49512, both isolated in France, were also included in the phylogenetic analysis. The best model indicated by the lowest Bayesian Information Criterion (BIC) value was used to generate the Neighbor-Joining tree based on 1000 replicates. The T92 model was selected for *dnaK*, *tuf*, and *gyrB*, while T92 + G was selected for both *trpB* and *rpoD* and JC model was selected for *atpA*. Strains corresponding to lineage I (cluster C and E) are coloured red and strains corresponding to lineage II (cluster G, A and H) are in blue. Other sequences assigned to reference strains are in black. Few strains representative of each genotype were used for tree construction (identical sequences removed for clarity of representation). (PDF 120 kb)

Additional file 4 The fluorescence peak profiles for the ARISA genotypes analysed with ABI Prism 3130xl Genetic Analyser and the GeneMapper v.5.0 software (Applied Biosystems, Carlsbad, California, USA). For the 83 strains isolated from Finland, we also determined the ARISA (automated ribosomal intergenic spacer analysis) genotypes following the procedure described by Suomalainen et al. [8]. However, the previously published method was modified so that ABI Prism 3130xl Genetic Analyser is used instead of LI-COR 4200 automatic sequencer The analysis revealed that ARISA genotypes associate uniformly with the clusters from the MLSA scheme. Briefly, the PCR reaction mixture (total volume 10 ul) contained 1X DreamTaq Buffer (Thermo Scientific), 0.2 mM dNTPs (Thermo Scientific), 0.5 μM reverse primer (23Sr, 5'-GGGTTBCCCCATTCRG-3), 0.44 μM forward primer (rD1f, 5'-GGCTGGATCACCTCCTT-3'), 0.06 μM 6-carboxyfluorescein (6-FAM) labelled forward primer (rD1f), 1 U of DreamTaq DNA Polymerase (Thermo Scientific) and 1 μl of template DNA. The PCR reactions were carried out using Bio-Rad C1000 or S1000 thermal cyclers (Bio-Rad Laboratories, Hercules, CA, USA). The thermo-cycling conditions included initial denaturation at 95 °C for 2 min followed by 30 cycles of amplification (94 °C for 30 s, 52 °C for 30 s, 72 °C for 3 min) and a final extension at 72 °C for 15 min. The PCR products were denatured with formamide and GeneScanTM 1200 LIZ Size Standard was added. Products were separated with an ABI Prism 3130xl Genetic Analyser, and visualized with GeneMapper v.5.0 software (all Applied Biosystems, Carlsbad, California, USA). Based on the fluorescence peak profiles and using the strains from Suomalainen et al. [8] as positive controls, the 87 *F. columnare* strains were designated into ARISA genotypes A to H [8]. (PDF 222 kb)

Additional file 5 Phylogenetic tree based on the 16 s rDNA sequence data obtained from the representatives of Finnish *F .columnare* genotypes (A-H) studied in this study and other *F .columnare* sequences obtained from the GenBank. The tree was constructed by a UPGMA clustering method with a resampling of 1,000 bootstrap replicates and the Jukes Cantor model. Two strains representative of each ARISA genotype/MLSA cluster studied in this study were used for tree construction (identical sequences removed for clarity of representation). (PDF 191 kb)

Additional file 6 The concatenated tree based on six housekeeping MLSA genes (including *trpB;* right) and five MLSA gene sequences (*rpoD, dnaK, tuf, gyrB, atpA*; left). The arrow shows disagreement concerning the position of strain G2 (B399) between the two trees. (PDF 77 kb)

Abbreviations

MLSA: Multilocus sequence analysis; MLST: Multilocus sequence typing; ATCC: American type culture collection; 16S rDNA-RFLP: 16S rDNA by restriction-fragment length polymorphism; SSCP: Single-strand conformation polymorphism; AFLP: Amplified fragment length polymorphism; PFGE: Pulsed-field gel electrophoresis; ARISA: Automated ribosomal intergenic spacer analysis; ISR-SSCP: Internal spacer region-single strand conformation polymorphism analysis; MCMC: Markov Chain Monte Carlo; IAS: Standardized index of association; dn/ds: the ratios of non-synonymous to synonymous substitutions; T93 + G: Tamura-Nei model and a gamma distribution; BIC: Bayesian Information Criterion; GC: Guanine + Cytosine content; S: the number of segregating sites.

Competing interests

The authors declare no competing interests; financial or otherwise.

Authors' contributions

R.A. designed and performed experiments, analysed data and wrote the paper. K.P. participated in data collection, and helped to draft the manuscript. LR.S. participated in designing the study and helped to draft the manuscript. N.P. performed ARISA genotyping analysis and helped to draft the manuscript. T.K. supervised the project and participated in designing the study, data analysis, and helped to draft the manuscript.

Acknowledgements

We would like to thank Dr. Päivi Rintamäki, Dr. Jean-Francois Bernardet, Dr. Heidi Kunttu, MSc. Reetta Penttinen and Dr. Elina Laanto for donating bacterial isolates for this study; and R. Penttinen for help in laboratory. We also thank Dr. Juan Galarza, Dr. Andrea González and Dr. Emily Knott for providing constructive comments and help in improving the contents of this paper. This work was supported by KONE foundation (Roghaieh Ashrafi via project "Constraints of evolutionary adaptation to climate change" to Tarmo Ketola) and Academy of Finland (Lotta-Riina Sundberg #272995, Tarmo Ketola # 278751, Katja Pulkkinen and Nina Pekkala via grant # 260704 to Jouni Taskinen) and Centre of Excellence in Biological Interactions (#252411, Prof. Johanna Mappes) to Roghaieh Ashrafi, Lotta-Riina Sundberg, and Tarmo Ketola.

Author details

[1]Centre of Excellence in Biological Interactions, Department of Biological and Environmental Science, University of Jyvaskyla, P.O. Box 35, Jyväskylän yliopisto FI-40014, Finland. [2]Department of Biological and Environmental Science, University of Jyvaskyla, P.O. Box 35, Jyväskylän yliopisto FI-40014, Finland.

Received: 6 May 2015 Accepted: 16 October 2015
Published online: 30 October 2015

References

1. Bernardet J, Grimont PA. Deoxyribonucleic acid relatedness and phenotypic characterization of *Flexibacter columnaris* sp. nov., nom. rev., *Flexibacter psychrophilus* sp. nov., nom. rev., and *Flexibacter maritimus* Wakabayashi, Hikida, and Masumura. Int J Syst Bacteriol 1989. 1986;39(3):346–54.
2. Declercq AM, Haesebrouck F, Van den Broeck W, Bossier P, Decostere A. Columnaris disease in fish: a review with emphasis on bacterium-host interactions. Vet Res. 2013;44(27):10.1186.
3. Soto E, Mauel MJ, Karsi A, Lawrence ML. Genetic and virulence characterization of *Flavobacterium columnare* from channel catfish (Ictalurus punctatus). J Appl Microbiol. 2008;104(5):1302–10.
4. Schneck J, Caslake L. Genetic diversity of *Flavobacterium columnare* isolated from fish collected from warm and cold water. J Fish Dis. 2006;29(4):245–8.
5. Austin B, Austin DA. Bacterial fish pathogens: disease of farmed and wild fish: Springer. 2007.
6. Verma V, Prasad Y, Singh BR. Effect of pH and salinity on pathogenicity of *flavobacterium columnare* and myxobacterium sp. in Indian cat fish, *Clarias batrachus* (Linn.) and *Heteropneustes fossilis* (Bloch.). J Environ Biol. 2011;32(5):573–7.
7. Pulkkinen K, Suomalainen LR, Read AF, Ebert D, Rintamaki P, Valtonen ET. Intensive fish farming and the evolution of pathogen virulence: the case of columnaris disease in Finland. Proc Biol Sci. 2010;277(1681):593–600.
8. Suomalainen LR, Kunttu H, Valtonen ET, Hirvela-Koski V, Tiirola M. Molecular diversity and growth features of *Flavobacterium columnare* strains isolated in Finland. Dis Aquat Organ. 2006;70(1–2):55–61.
9. Achtman M. A phylogenetic perspective on molecular epidemiology. In: Sussman M, editor. Molecular Medical Microbiology. London: Academic; 2002. p. 485–509.
10. Triyanto A, Wakabayashi H. Genotypic diversity of strains of *Flavobacterium columnare* from diseased fishes. Fish Pathol. 1999;65–71.
11. LaFrentz BR, Waldbieser GC, Welch TJ, Shoemaker CA. Intragenomic heterogeneity in the 16S rRNA genes of *Flavobacterium columnare* and standard protocol for genomovar assignment. J Fish Dis. 2014;37(7):657–69.
12. Michel C, Messiaen S, Bernardet J. Muscle infections in imported neon tetra, *Paracheirodon innesi* Myers: limited occurrence of microsporidia and predominance of severe forms of columnaris disease caused by an Asian genomovar of *Flavobacterium columnare*. J Fish Dis. 2002;25(5):253–63.
13. Olivares-Fuster O, Shoemaker CA, Klesius PH, Arias CR. Molecular typing of isolates of the fish pathogen, *Flavobacterium columnare*, by single-strand conformation polymorphism analysis. FEMS Microbiol Lett. 2007;269:63–9.
14. Arias CR, Welker TL, Shoemaker CA, Abernathy JW, Klesius PH. Genetic fingerprinting of *Flavobacterium columnare* isolates from cultured fish. J Appl Microbiol. 2004;97(2):421–8.
15. Olivares-Fuster O, Baker JL, Terhune JS, Shoemaker CA, Klesius PH, Arias CR. Host-specific association between *Flavobacterium columnare* genomovars and fish species. Syst Appl Microbiol. 2007;30(8):624–33.
16. Enright MC, Spratt BG. Multilocus sequence typing. Trends Microbiol. 1999;7(12):482–7.
17. Tekedar HC, Karsi A, Gillaspy AF, Dyer DW, Benton NR, Zaitshik J, et al. Genome sequence of the fish pathogen *Flavobacterium columnare* ATCC 49512. J Bacteriol. 2012;194(10):2763–4.
18. Bougnoux ME, Morand S, d'Enfert C. Usefulness of multilocus sequence typing for characterization of clinical isolates of *Candida albicans*. J Clin Microbiol. 2002;40(4):1290–7.
19. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A. 1998;95(6):3140–5.
20. Maiden MC. Multilocus sequence typing of bacteria. Annu Rev Microbiol. 2006;60:561–88.
21. de Las RB, Marcobal A, Munoz R. Allelic diversity and population structure in *Oenococcus oeni* as determined from sequence analysis of housekeeping genes. Appl Environ Microbiol. 2004;70(12):7210–9.
22. Nicolas P, Mondot S, Achaz G, Bouchenot C, Bernardet JF, Duchaud E. Population structure of the fish-pathogenic bacterium *Flavobacterium psychrophilum*. Appl Environ Microbiol. 2008;74(12):3702–9.
23. Bilhere E, Lucas PM, Claisse O, Lonvaud-Funel A. Multilocus sequence typing of *Oenococcus oeni*: detection of two subpopulations shaped by intergenic recombination. Appl Environ Microbiol. 2009;75(5):1291–300.
24. Mo S, You M, Su YC, Lacap-Bugler DC, Huo YB, Smith GJ, et al. Multilocus sequence analysis of *Treponema denticola* strains of diverse origin. BMC Microbiol. 2013;13:24.
25. Rong X, Liu N, Ruan J, Huang Y. Multilocus sequence analysis of *Streptomyces griseus* isolates delineating intraspecific diversity in terms of both taxonomy and biosynthetic potential. Antonie Van Leeuwenhoek. 2010;98(2):237–48.
26. Shieh H. Studies on the nutrition of a fish pathogen. Flexibacter columnaris Microbios Letters. 1980;13(51/52):129–33.
27. Decostere A, Haesebrouck F, Devriese L. Shieh medium supplemented with tobramycin for selective isolation of *Flavobacterium columnare* (*Flexibacter columnaris*) from diseased fish. J Clin Microbiol. 1997;35:322–4.
28. Anacker R, Ordal E. Studies on the myxobacterium *Chondrococcus columnaris*. I Serological typing J Bacteriol. 1959;78:25–32.
29. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3–new capabilities and interfaces. Nucleic Acids Res. 2012;40(15), e115.
30. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. Bioinformatics. 2007;23(10):1289–91.

31. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acid Sym Ser. 1999;41:95–8.
32. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011;28(10):2731–9.
33. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25(11):1451–2.
34. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123(3):585–95.
35. Jolley KA, Feil EJ, Chan MS, Maiden MC. Sequence type analysis and recombinational tests (START). Bioinformatics. 2001;17(12):1230–1.
36. Haubold B, Hudson RR. LIAN 3.0: detecting linkage disequilibrium in multilocus data. Linkage Analysis. Bioinformatics. 2000;16(9):847–8.
37. Smith JM, Smith NH, O'Rourke M, Spratt BG. How clonal are bacteria? Proc Natl Acad Sci U S A. 1993;90(10):4384–8.
38. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155(2):945–59.
39. Earl DA. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation genetics resources. 2012;4(2):359–61.
40. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005;14(8):2611–20.
41. Didelot X, Falush D. Inference of bacterial microevolution using multilocus sequence data. Genetics. 2007;175(3):1251–66.
42. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Statistical science. 1992;7:457–72.
43. Nielsen R. Statistical tests of selective neutrality in the age of genomics. Heredity. 2001;86(6):641–7.
44. Simonsen KL, Churchill GA, Aquadro CF. Properties of statistical tests of neutrality for DNA polymorphism data. Genetics. 1995;141(1):413–29.
45. Mohammed HH, Arias CR. Epidemiology of columnaris disease affecting fishes within the same watershed. Dis Aquat Organ. 2014;109:201–11.
46. Smith JM, Feil EJ, Smith NH. Population structure and evolutionary dynamics of pathogenic bacteria. Bioessays. 2000;22(12):1115–22.
47. Gonzalez-Gonzalez A, Sanchez-Reyes LL, Delgado Sapien G, Eguiarte LE, Souza V. Hierarchical clustering of genetic diversity associated to different levels of mutation and recombination in *Escherichia coli*: a study based on Mexican isolates. Infect Genet Evol. 2013;13:187–97.
48. Gonzalez-Escalona N, Martinez-Urtaza J, Romero J, Espejo RT, Jaykus LA, DePaola A. Determination of molecular phylogenetics of *Vibrio parahaemolyticus* strains by multilocus sequence typing. J Bacteriol. 2008;190(8):2831–40.
49. Nilsen H, Sundell K, Duchaud E, Nicolas P, Dalsgaard I, Madsen L, et al. Multilocus sequence typing identifies epidemic clones of *Flavobacterium psychrophilum* in Nordic countries. Appl Environ Microbiol. 2014;80(9):2728–36.
50. Habib C, Houel A, Lunazzi A, Bernardet JF, Olsen AB, Nilsen H, et al. Multilocus sequence analysis of the marine bacterial genus Tenacibaculum suggests parallel evolution of fish pathogenicity and endemic colonization of aquaculture systems. Appl Environ Microbiol. 2014;80(17):5503–14.
51. Vos M, Didelot X. A comparison of homologous recombination rates in bacteria and archaea. ISME J. 2009;3(2):199–208.
52. Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 16S ribosomal DNA 414 amplification for phylogenetic study. J Bacteriol. 1991;173:697–703.

Additional File 1. Site , year of isolation, source of isolation (fish or water), location of isolation, host species, sequence type (ST), and allelic profile data for the 83 *F. columnare* strains from Finland analyzed by MLST.

Strain	Site	Year	Source	Location	Host	ST	*trpB*	*rpoD*	*tuf*	*atpA*	*dnaK*	*gyrB*
B531	NorthA	2012	Fish	northern	brown trout (*Salmo trutta*)	2	2	2	2	1	1	1
B532	NorthA	2012	Fish	northern	brown trout (*Salmo trutta*)	2	2	2	2	1	1	1
B529	NorthC	2012	Fish	northern	Atlantic salmon (*Salmo salar)*	2	2	2	2	1	1	1
B526	NorthC	2012	Fish	northern	Atlantic salmon (*Salmo salar*)	2	2	2	2	1	1	1
B521	NorthA	2006	Fish	northern	Atlantic salmon (*Salmo salar*)	2	2	2	2	1	1	1
B520	NorthA	2006	Fish	northern	Atlantic salmon (*Salmo salar*)	2	2	2	2	1	1	1
B519	NorthA	2006	Fish	northern	brown trout (*Salmo trutta*)	2	2	2	2	1	1	1
B518	NorthB	2006	Fish	northern	Atlantic salmon (*Salmo salar*)	2	2	2	2	1	1	1
B463	SouthD	2011	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	2	2	2	2	1	1	1
B458	NorthB	2009	Fish	northern	brown trout (*Salmo trutta*)	2	2	2	2	1	1	1
B449	SouthA	2007	Water	southern	Water	2	2	2	2	1	1	1
B448	SouthA	2007	Water	southern	Water	2	2	2	2	1	1	1
B447	SouthC	2007	Water	southern	Water	2	2	2	2	1	1	1
B441	NorthB	2007	Fish	northern	Atlantic salmon (*Salmo salar*)	2	2	2	2	1	1	1
B439	SouthD	2006	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	2	2	2	2	1	1	1
B437	NorthB	2006	Fish	northern	brown trout (*Salmo trutta*)	2	2	2	2	1	1	1
B426	NorthB	2006	Fish	northern	Atlantic salmon (*Salmo salar)*	2	2	2	2	1	1	1
B424	NorthA	2007	Fish	northern	brown trout (*Salmo trutta*)	2	2	2	2	1	1	1
B421	NorthB	2008	Fish	northern	Atlantic salmon (*Salmo salar*)	2	2	2	2	1	1	1
B418	NorthA	2009	Fish	northern	brown trout (*Salmo trutta*)	2	2	2	2	1	1	1
B417	NorthA	2008	Fish	northern	Atlantic salmon (*Salmo salar*)	2	2	2	2	1	1	1
B416	NorthA	2008	Fish	northern	brown trout (*Salmo trutta)*	2	2	2	2	1	1	1

B409	NorthA	2010	Fish	northern	sea trout (*Salmo trutta trutta*)	2	2	2	2	1	1	1
B408	SouthC	2010	Water	southern	Water	2	2	2	2	1	1	1
B405	SouthA	2010	Water	southern	Water	2	2	2	2	1	1	1
B402	SouthC	2010	Fish	southern	European whitefish(*Coregonus lavaretus*)	2	2	2	2	1	1	1
B366	SouthC	2010	Water	southern	Water	2	2	2	2	1	1	1
B357	SouthC	2010	Water	southern	Water	2	2	2	2	1	1	1
B270	NorthB	2009	Water	northern	Water	2	2	2	2	1	1	1
B261	SouthC	2009	Water	southern	Water	2	2	2	2	1	1	1
B259	SouthC	2009	Water	southern	Water	2	2	2	2	1	1	1
B245	SouthC	2009	Water	southern	Water	2	2	2	2	1	1	1
B230	SouthC	2009	Water	southern	Water	2	2	2	2	1	1	1
Os06	NorthC	2006	Fish	northern	Atlantic salmon (*Salmo salar*)	2	2	2	2	1	1	1
B367	SouthC	2010	Water	southern	Water	4	4	2	2	3	1	2
B369	SouthC	2010	Fish	southern	brown trout (*Salmo trutta*)	4	4	2	2	3	1	2
B376	SouthC	2010	Water	southern	Water	4	4	2	2	3	1	2
B377	SouthC	2010	Water	southern	Water	4	4	2	2	3	1	2
B379	SouthC	2010	Water	southern	Water	4	4	2	2	3	1	2
B430	SouthC	2003	Fish	southern	Pike perch (*Stizostedion lucioperca*)	4	4	2	2	3	1	2
B445	SouthC	2007	Water	southern	Water	4	4	2	2	3	1	2
B453	SouthC	2008	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	4	4	2	2	3	1	2
B454	SouthC	2008	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	4	4	2	2	3	1	2
B455	SouthC	2008	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	4	4	2	2	3	1	2
B491	SouthC	2012	Water	southern	Water	4	4	2	2	3	1	2
B496	SouthC	2012	Water	southern	Water	4	4	2	2	3	1	2
B503	SouthC	2012	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	4	4	2	2	3	1	2
B504	SouthC	2012	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	4	4	2	2	3	1	2
B508	SouthC	2012	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	4	4	2	2	3	1	2

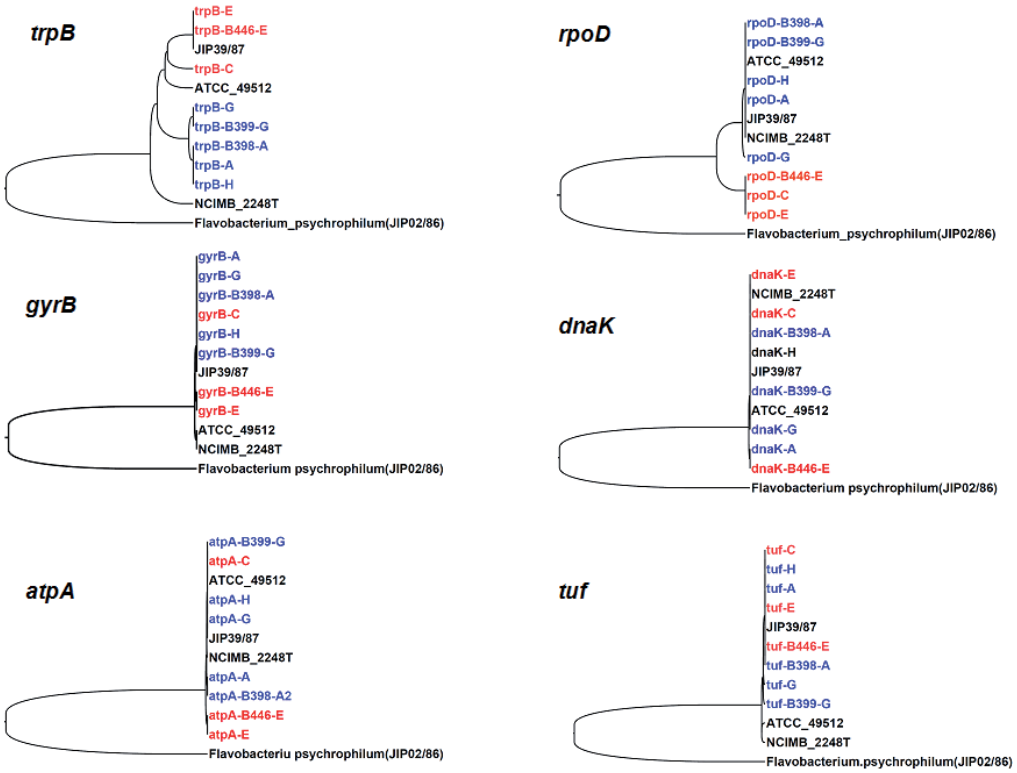
B510	SouthC	2012	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	4	4	2	2	3	1	2
B511	SouthC	2012	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	4	4	2	2	3	1	2
B513	SouthC	2012	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	4	4	2	2	3	1	2
B446	SouthC	2007	Water	southern	Water	8	4	2	2	3	2	2
B517	SouthC	2011	Water	southern	Water	4	4	2	2	3	1	2
B450	SouthB	2007	Fish	southern	Pike perch (*Stizostedion lucioperca*)	1	1	1	1	1	1	1
B451	SouthB	2007	Fish	southern	Pike perch (*Stizostedion lucioperca*)	1	1	1	1	1	1	1
B444	NorthC	2007	Fish	northern	Atlantic salmon (*Salmo salar*)	1	1	1	1	1	1	1
B442	NorthC	2007	Fish	northern	Atlantic salmon (*Salmo salar*)	1	1	1	1	1	1	1
B440	NorthB	2007	Fish	northern	Atlantic salmon (*Salmo salar*)	1	1	1	1	1	1	1
B436	NorthC	2006	Fish	northern	Atlantic salmon (*Salmo salar*)	1	1	1	1	1	1	1
B435	NorthC	2005	Fish	northern	Atlantic salmon (*Salmo salar*)	1	1	1	1	1	1	1
B434	NorthC	2005	Fish	northern	Atlantic salmon (*Salmo salar*)	1	1	1	1	1	1	1
B420	NorthB	2009	Fish	northern	Atlantic salmon (*Salmo salar*)	1	1	1	1	1	1	1
B419	NorthB	2009	Fish	northern	Atlantic salmon (*Salmo salar*)	1	1	1	1	1	1	1
B407	SouthC	2010	Water	southern	Water	1	1	1	1	1	1	1
B393	SouthC	2010	Fish	southern	bream (*Abramis brama*)	1	1	1	1	1	1	1
B185	SouthB	2009	Water	southern	Water	1	1	1	1	1	1	1
B399	SouthC	2010	Water	southern	Water	7	1	3	3	1	1	1
B429	SouthC	2003	Fish	southern	Pike perch (*Stizostedion lucioperca*)	5	3	3	2	1	1	1
B533	NorthB	2012	Fish	northern	Atlantic salmon (*Salmo salar*)	5	3	3	2	1	1	1
H	SouthC	2003	Fish	southern	rainbow trout (*Oncorhynchus mykiss*)	5	3	3	2	1	1	1
B398	SouthC	2010	Water	southern	Water	6	3	3	2	2	1	1
B269	NorthB	2009	Water	northern	Water	3	3	3	2	2	2	1
B396	SouthC	2010	Water	southern	Water	3	3	3	2	2	2	1
B422	NorthD	2009	Fish	northern	sea trout (*Salmo trutta trutta*)	3	3	3	2	2	2	1
B423	NorthD	2009	Fish	northern	lake trout (*Salmo trutta lacustris*)	3	3	3	2	2	2	1

B427	NorthD	2006	Fish	northern	brown trout (*Salmo trutta*)	3	3	3	2	2	2	1
B431	SouthB	2003	Fish	southern	grayling *Thymallus thymallus*	3	3	3	2	2	2	1
B438	NorthD	2006	Fish	northern	rainbow trout (*Oncorhynchus mykiss*)	3	3	3	2	2	2	1
B452	SouthB	2007	Fish	southern	Pike perch (*Stizostedion lucioperca)*	3	3	3	2	2	2	1
B523	NorthD	2007	Fish	northern	brown trout (*Salmo trutta*)	3	3	3	2	2	2	1
B067	SouthB	2007	Fish	southern	brown trout (*Salmo trutta*)	3	3	3	2	2	2	1
tulo2	SouthC	2010	Water	southern	Water	3	3	3	2	2	2	1

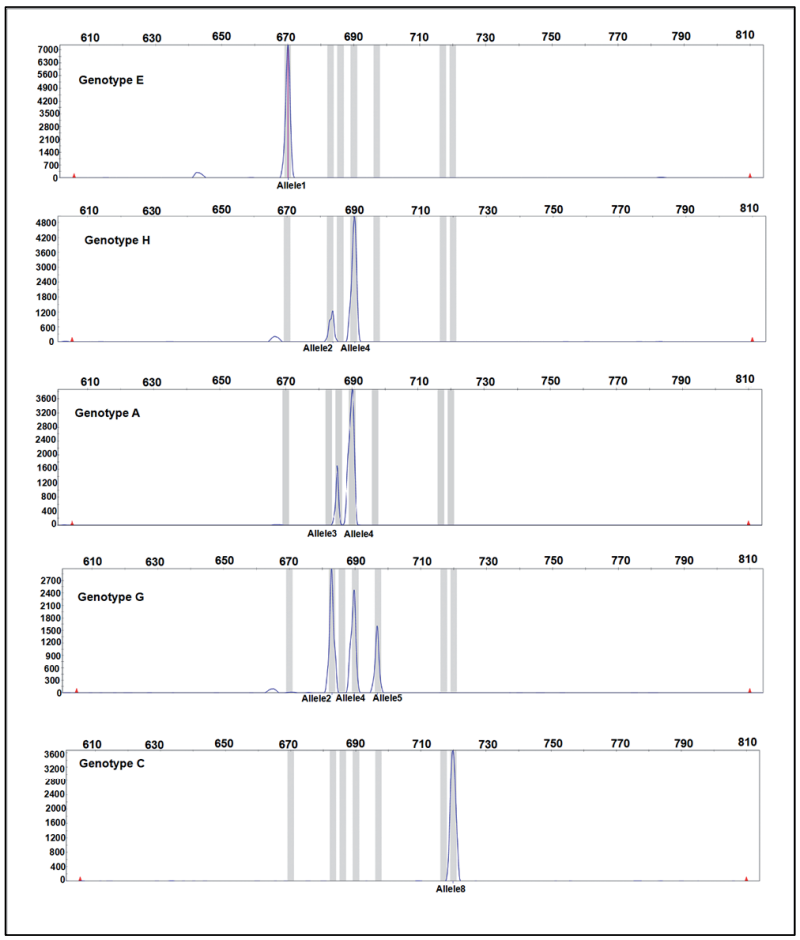
Additional File 2. Target genes and primers for the housekeeping genes of *F. columnare.* The loci used for the MLST/MLSA scheme are shown in bold font. Length refers to the length of the target sequence. .* Reference for 16S rDNA primers [52].

Target gene	Gene product	Direction	primer sequence	Length
trpB	**tryptophan synthase subunit B**	**Forward**	**5'- TGCCATACAGGTGCGCATAA -3'**	**730 bp**
		Reverse	**5'- TGGCTCTCCCGCTTTTGAAT-3**	
gyrB	**DNA gyrase B**	**Forward**	**5'-TACGCACGAAGGAGGAACAC-3'**	**704 bp**
		Reverse	**5'-GGCTCCCATCAATATCGGCA-3'**	
dnaK	**molecular chaperone DnaK**	**Forward**	**5'-GGCTACAGCTTCTGGACCAA-3'**	**727 bp**
		Reverse	**5'-AGCGGCTTTATCTGCTTCCG-3'**	
tuf	**Elongation factor Tu**	**Forward**	**5'-ACATGGTTACTGGTGCTGCT-3'**	**705 bp**
		Reverse	**5'-TATGGAATGGCGTGTGACGA-3'**	
atpA	**α-subunit of bacterial ATP synthase**	**Forward**	**5'-TGGACGTACCCCAGAGTTGA-3'**	**745 bp**
		Reverse	**5'-GCGTAAAGCACCAGGGGTAA-3'**	
rpoD	**RNA polymerase sigma factor**	**Forward**	**5'-AGCACAACGCATCAAGGCTGGT-3'**	**480 bp**
		Reverse	**5'-TGGGGCATCCATAGATAAATGGCGT-3'**	
fumC	fumarate hydratase	Forward	5'-GTTCTAACCTGCCCGCTCAA-3'	744 bp
		Reverse	5'-GCATTGGTTTGCCTGCATCT-3'	
murG	N-acetylglucosaminyltransferases	Forward	5'-GGGGAGGTACAGGAGGTCAT-3'	767 bp
		Reverse	5'-AATGACTATATCGGCGGCGG-3'	
rplB	50S ribosomal protein L2	Forward	5'-CCTATTACCCCAGGTCAGCG-3'	719 bp
		Reverse	5'-CTTGAACGTGGGTGACCTCC-3'	
recA	recombinase A	Forward	5'-GGCGGCTATCCTAAAGGTCG-3'	746 bp
		Reverse	5'-CGGCCTTGACCGAGTTTTGT-3'	
fstQ	cell division protein FtsQ	Forward	5'-ACCCAAGTGGAAATTGTGGA-3'	522bp
		Reverse	5'-TGTATCCTGGACTGCTTTTTGA-3'	
glyA	serine hydroxy methyl transferase	Forward	5'-AATGGAAGCGGCAGGTTCTT-3'	705 bp
		Reverse	5'-GGGCCTCCTTGATTACCTGG-3'	
16S rDNA	16S ribosomal DNA	Forward (fD1)*	5'-AGAGTTTGATCCTGGCTCAG-3'	500 bp
		Reverse (Prun 518)	5'-ATTACCGCGGCTGCTGG-3'	

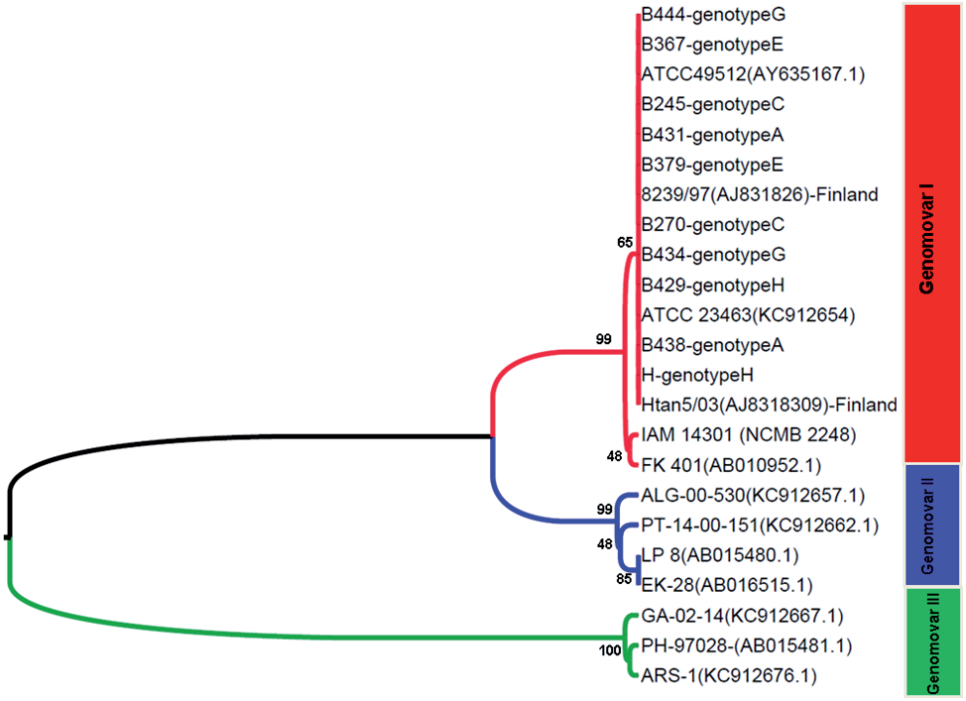
Additional File 3. Neighbor Joining phylogenetic trees based on the individual sequences of six MLST loci (*trpB, rpoD, gyrB , dnaK, atpA,*and *tuf*). MEGA v5.2 was used to evaluate the models for nucleotide substitution for each protein-coding locus and to construct the phylogenetic trees for Finnish *F. columnare* strains. The *F. columnare* type strain NCIMB 2248^{T} isolated in the USA and two reference strains JIP39/87 and ATCC49512, both isolated in France, were also included in the phylogenetic analysis. The best model indicated by the lowest Bayesian Information Criterion (BIC) value was used to generate the Neighbor-Joining tree based on 1000 replicates. The T92 model was selected for *dnaK*, *tuf,* and *gyrB*, while T92+G was selected for both *trpB* and *rpoD* and JC model was selected for *atpA*. Strains corresponding to lineage I (cluster C and E) are coloured red and strains corresponding to lineage II (cluster G, A and H) are in blue. Other sequences assigned to reference strains are in black. Few strains representative of each genotype were used for tree construction (identical sequences removed for clarity of representation).



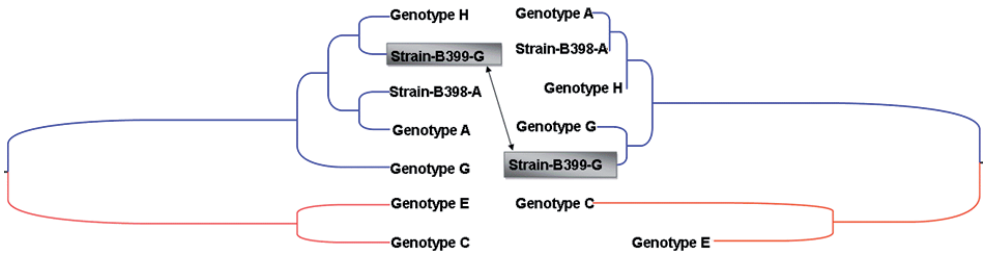
Additional File 4. The fluorescence peak profiles for the ARISA genotypes analysed with ABI Prism 3130xl Genetic Analyser and the GeneMapper v.5.0 software (Applied Biosystems, Carlsbad, California, USA). For the 83 strains isolated from Finland, we also determined the ARISA (automated ribosomal intergenic spacer analysis) genotypes following the procedure described by Suomalainen et al.[8]. However, the previously published method was modified so that ABI Prism 3130xl Genetic Analyser is used instead of LI-COR 4200 automatic sequencer The analysis revealed that ARISA genotypes associate uniformly with the clusters from the MLSA scheme. Briefly, the PCR reaction mixture (total volume 10 ul) contained 1X DreamTaq Buffer (Thermo Scientific), 0.2 mM dNTPs (Thermo Scientific), 0.5 µM reverse primer (23Sr, 5'-GGGTTBCCCCATTCRG-3), 0.44 µM forward primer (rD1f, 5'-GGCTGGATCACCTCCTT-3'), 0.06 µM 6-carboxyfluorescein (6-FAM) labelled forward primer (rD1f), 1 U of DreamTaq DNA Polymerase (Thermo Scientific) and 1 µl of template DNA. The PCR reactions were carried out using Bio-Rad C1000 or S1000 thermal cyclers (Bio-Rad Laboratories, Hercules, CA, USA). The thermo-cycling conditions included initial denaturation at 95°C for 2 min followed by 30 cycles of amplification (94°C for 30 s, 52°C for 30 s, 72°C for 3 min) and a final extension at 72°C for 15 min. The PCR products were denatured with formamide and GeneScanTM 1200 LIZ Size Standard was added. Products were separated with an ABI Prism 3130xl Genetic Analyser, and visualized with GeneMapper v.5.0 software (all Applied Biosystems, Carlsbad, California, USA). Based on the fluorescence peak profiles and using the strains from Suomalainen et al. (8) as positive controls, the 87 *F. columnare* strains were designated into ARISA genotypes A to H [8].



Additional File 5. Phylogenetic tree based on the 16s rDNA sequence data obtained from the representatives of Finnish *F .columnare* genotypes (A-H) studied in this study and other *F .columnare* sequences obtained from the GenBank. The tree was constructed by a UPGMA clustering method with a resampling of 1,000 bootstrap replicates and the Jukes Cantor model. Two strains representative of each ARISA genotype/MLSA cluster studied in this study were used for tree construction (identical sequences removed for clarity of representation).



Additional File 6. The concatenated tree based on six housekeeping MLSA genes (including *trpB;* right) and five MLSA gene sequences (*rpoD, dnaK, tuf, gyrB, atpA*; left). The arrow shows disagreement concerning the position of strain B399-G between the two trees.



II

THERMAL TOLERANCE AT HIGH TEMPERATURE IS LINKED WITH REDUCED VIRULENCE IN A FISH PATHOGEN

by

Roghaieh Ashrafi, Matthieu Bruneaux, Lotta-Riina Sundberg, Katja Pulkkinen, Janne Valkonen, Tarmo Ketola 2017

Submitted Manuscript

III

APPLICATION OF HIGH RESOLUTION MELTING ASSAY (HRM) TO STUDY TEMPERATURE-DEPENDENT INTRASPECIFIC COMPETITION IN A PATHOGENIC BACTERIUM

by

Roghaieh Ashrafi, Matthieu Bruneaux, Lotta-Riina Sundberg, Katja Pulkkinen, Tarmo Ketola 2017

Scientific Reports, in press

Application of high resolution melting assay (HRM) to study temperature-dependent intraspecific competition in a pathogenic bacterium

Roghaieh Ashrafi[1]*, Matthieu Bruneaux[1] , Lotta-Riina Sundberg[1], Katja Pulkkinen[2], Tarmo Ketola[1]

1University of Jyvaskyla, Centre of Excellence in Biological Interactions. Department of Biological and Environmental Science (and Nanoscience Center). P.O. Box 35. FI-40014 University of Jyvaskyla. FINLAND

2University of Jyvaskyla, Department of Biological and Environmental Science. P.O. Box 35. FI-40014 University of Jyvaskyla. FINLAND

Keywords: High resolution melting curve (HRM); Intraspecific competition; Temperature; Opportunistic bacteria

*Corresponding author e-mail: roghaieh.ashrafi@jyu.fi

Fax Number: +358-14-617-239

Running Title: Simultaneous identification of mixed genotypes

ABSTRACT

Studies on species' responses to climate change have focused largely on the direct effect of abiotic factors and in particular temperature, neglecting the effects of biotic interactions in determining the outcome of climate change projections. Many microbes rely on strong interference competition; hence the fitness of many pathogenic bacteria could be a function of both their growth properties and intraspecific competition. However, due to technical challenges in distinguishing and tracking individual strains, experimental evidence on intraspecific competition has been limited so far. Here, we developed a robust application of the high-resolution melting (HRM) assay to study head-to-head competition between mixed genotype co-cultures of a waterborne bacterial pathogen of fish, *Flavobacterium columnare,* at two different temperatures. We found that competition outcome in liquid cultures seemed to be well predicted by growth yield of isolated strains, but was mostly inconsistent with interference competition results measured in inhibition tests on solid agar, especially as no growth inhibition between strain pairs was detected at the higher temperature. These results suggest that, for a given temperature, the factors driving competition outcome differ between liquid and solid environments.

INTRODUCTION

Climate models predict that climate change will lead to increased temperatures and larger thermal fluctuations in the future [1]. As a consequence of climate-induced range shifts and expanded human activities leading to the transportation of species, we expect to see previously separated populations come into contact with one another, providing new opportunities for interactions. Hence, it is of the utmost important to note that the effects of climate change can invoke factors other than temperature, particularly changes in biotic interactions within and among species [2-8]. For example, winners judged by the suitability of habitat can be competitively excluded by superior competitors [9, 10]. Given the global significance of pathogens, predicting how climate warming changes the outcome of intraspecific interactions in pathogens, which in turn would affect their population dynamics such as growth and virulence, is extremely important for the prevention of pathogen spread and the diseases caused by them.

Empirical investigations of intraspecific interactions using multiple strains from a single species have limitations in tracking individual strains in mixed cultures. Traditionally, genotyping or phenotypic markers based on LacZ, fluorescence or antibiotic resistance have been used to discriminate strains and determine the relative abundance of multiple strains that interact within a mixed culture [11, 12]. However, markers based on genetic engineering may provide misleading results, as they may incur fitness costs or benefits in bacteria [13-15]. To address these problems, and in order to perform our study, we applied a method for the quantification of competition outcome for our study species, a bacterial fish pathogen, *Flavobacterium columnare*. We implemented an assay based on high-resolution melting curve (HRM) which has been used for various applications, most commonly the determination of DNA methylation status and genotyping [16-18]. Using genomic DNA containing a naturally occurring mutation in the tryptophan synthase gene *trpB* enabled us to quantitatively distinguish several *F. columnare* strains belonging to two genotypes of *trpB* from mixed-genotype competition cultures. To our knowledge, this is the first study that has used a HRM approach to monitor and quantify very closely related genotypes of one bacterial species in a single tube, and the method is extendable to multiple genotypes by multiplexing, as is exemplified in Appendix S1 (See also Fig. S4).

F. columnare serves as an excellent model system to examine whether temperature-dependent biotic interactions, such as competition, could alter population dynamics and climate warming projections for a few reasons. First, it has been shown that high summer temperatures correlate with increased *F. columnare* outbreaks [19]. Second, multiple genotypes are commonly present in a single region [20, 21] implying the potential for intraspecific competition among genotypes. Third, competition among *F. columnare* strains, especially between the ones chosen for this experiment, is not restricted to resource competition and strains can directly interfere with competitors via growth inhibition [21, 22].

The aim of our study was to test if head to head competition, interference competition, and growth parameters in two temperatures will lead to the same or different conclusions. If for the majority of the clone pairs the competition outcomes in different temperatures can be predicted by growth assays or interference assays that will facilitate predictions of future climate change on this important fish pathogen. However, if predictions between growth assays, interference, and head to head competition drastically differ, it will cause challenges in predicting which of the *F. columnare* genotypes will excel under progressing climate change. Although we use here a species with applied importance, the same question applies to all attempts to understand the fate of species under changing conditions in the wild, where intraspecific competitive interactions are the norm rather than the exception.

RESULTS

Quantification of genotype proportions in control mixtures using HRM

A single SNP (T > G substitution) at position 222 of the *trpB* gene was used to implement the high-resolution melting analysis (HRM) for *F. columnare* genotypes, G and A, circulating in Finnish fish farms. The amplicon subjected to high resolution melting analysis could be used to distinguish between the two SNP alleles based on the melting curve shape. The melt peak of the genotype A was centered at 80°C, while that of the genotype G was centered at 80.60°C (Fig. 1). Sequencing of the amplified products in both directions confirmed the HRM results (Fig. S2).

Comparison of the melting temperature values for the reference samples across the HRM runs showed that there were some variations between the runs. For this reason, it is important to use within-plate reference samples for the calibration of each run. The within-plate calibration curves showed a good linear fit between melting temperatures and genotype proportions for the calibration samples (see R^2 values on Fig. S3). For competition group C, there was a technical issue during the preparation of the reference samples for mixes from 20% to 60%, and those calibration samples were discarded from the analysis. Our estimate of genotype proportions were consistent with the clustering results obtained using the Precision Melt Analysis software (Bio-rad); however, they enabled us to obtain a more quantitative estimate of genotype proportions in the experimental samples compared to the cluster classification obtained with Precision Melt Analysis software (Figure 2).

Competition and growth

The temporal change of genotype proportions as estimated by HRM during the competition in liquid medium is presented in Figure 3. Overall, the five replicates within each competition group showed consistent results. Group A was the only competition group for which the outcome of competition differed between 26°C and 31°C (Table 1). The strains with the faster growth and higher yield were those able to utilize resources faster and won in all tested competition pairs at both

temperatures, except in one group (C) where the losing competitor had a higher yield than the winner at 26°C. However, the results of interference competition on plates with solid agar matched the results of the competition in liquid media in only half of the cases (Table1). At 31 °C no inhibition between strains was detected (Table1).

DISCUSSION

While temperature change can affect thermal sensitivities, tolerance range, and optimal performance of populations, predicting how this change will ultimately affect the structure and distribution of species depends also on indirect effects mediated by biotic interactions [5]. However, due to technical limitations, such as tracking individual isolates in co-cultures, relatively few studies have empirically examined how joint changes in intraspecific interactions and temperature affect ecological responses to increasing temperatures [12]. Based on a diagnostic SNP, we successfully established a method based on real-time PCR and HRM analysis for the rapid differentiation and quantification of two closely related genotypes of *F. columnare* in mixed culture. Using this method, we could resolve the outcome of head-to-head competition of bacterial strains in liquid culture. Since this method does not require any alteration in the genotype of the wild-type strain, it is a suitable method to study ecological fitness and competitive interactions between natural isolates. Overall, our results confirm that the HRM method can be used to accurately quantify the relative abundance of different strains in co-culture samples, and suggest that HRM offers a simple and cost effective means for detection of these strains in monitoring intraspecific competition in a single culture tube.

We observed that the competition outcomes of two strains in mixed culture at two temperatures were associated with changes in bacterial growth rate in the given temperature, rather than with their capacity for interference (Table 1). This gives a positive indication that growth measurements could be effective in predicting the success of the clones in liquid cultures and in different temperatures, despite the interference competition between genotypes of this species. This verification of the close match between growth parameters and competition is important to model possible outbreak scenarios under changing environmental conditions in environments where multiple pathogen strains co-occur.

Bacteria often produce toxins that kill or inhibit the growth other strains [23-25]. Inconsistency between competition outcomes in liquid cultures and on plate culture observed in our study suggests that benefits of interference on a surface is relaxed in liquid culture, where the bacterial cells are more likely to have less frequent contact with competitor cells, and the toxins are subjected to dilution. This indicates different trajectories of competition strategies depending whether the cells live in biofilm or in a planktonic state [26, 27]. Toxins are, indeed, commonly released in surfaces (such as biofilm) in which bacterial populations are dense and consequently nutrient limited [28]. Furthermore, limited dispersal can favour toxin

production, suggesting that this mechanism will profit bacteria growing on surfaces rather than the planktonic cells in free waters and liquid cultures [29].

Increased global temperatures can strongly decrease the efficiency of food production by influencing pathogen species [30], such as *F. columnare* [19, 21, 31]. Prolonged infective seasons provided by the higher mean temperatures allow for more bacterium-bacterium and bacterium-host interactions, increasing the benefits of interference. Increased frequency of high and maximum temperatures, on the other hand, may select for bacterial strains with a wider temperature optimum, and change the strain composition of bacterial populations. Interestingly, however, we did not find any interference at high temperatures (31°C). Technical problems can be safely ruled out as interference tests at both temperatures were conducted at the same time, from the same original bacterial stocks. Previous studies have shown that bacteriocin (growth-inhibiting substance needed for interference) production is sensitive to high temperature in lactic acid bacteria and in *Yersinia* [32-34]. In temperate regions, the temperature of the aquatic environment rarely exceeds the +25°C used as the lower temperature in our interference experiments, which could maintain diversity and competitive interactions in bacterial populations. For example, during recent years the propensity of *F. columnare* strains isolated from disease epidemics from fish farms to inhibit the growth of competing strains has increased, suggesting increased intraspecific competition in the fish farming environment [21]. The benefits of interference in *F. columnare* might be realized in the field conditions during biofilm formation and multiple infections, which are likely to occur when several strains are present.

Most of the current research on the effects of climate change concentrates on single genotypes, or species in isolation, but in the wild the situation is not so simple. Although laboratory experiments, such as ours, can shed light on the pairwise interactions of bacterial strains and improve our understanding on traits that make certain strains more successful, conditions in the environment are more complex. In real life, species and genotypes interact and may have additive effects for factors such as virulence [35], thus proxies of fitness, such as individual growth rates, might not be predictive enough of which of the genotypes will excel in competition. The use of the HRM method with high sensitivity and specificity will likely allow us mimic more natural systems, which will in turn help in understanding the circumstances that affect pathogenicity of this environmentally growing opportunistic bacterium. Moreover, as shown in Appendix S1 (See also Fig. S4), this method is extendable to simultaneous quantitative separation of multiple genotypes. We expect this method to be useful in disentangling how global warming and other abiotic and biotic factors contribute to bacterial interactions and impact the dynamics and functionality of bacterial populations.

MATERIAL AND METHODS

Bacterial strain characteristics and culture conditions

Six *F. columnare* isolates, previously assigned to two genetic groups (genotypes G or A) with the MLSA method [20], were used in this study. Before growth yield

measurement, inhibition and competition, the strains were revived from the frozen stocks to 3 ml of Shieh medium [36] and incubated for 24 h in a shaker (200 rpm) at 26 °C. After 24 hours, cultures were diluted into fresh modified Shieh medium (1:10) and allowed to regrow at 26 °C on the shaker for another 24 hours. The six *F. columnare* isolates were distributed into four competition pairs, each containing one strain from the G- and one strain from the A-genotype, and all different growth and competition assays were performed on these strains (Table 1).

Competition

The optical density (OD, at 570 nm) of the bacterial cultures was measured with a spectrophotometer adjusted to an approximate value of 0.140 (~7.6×10^6 cells/ml) using the Shieh medium. Then, replicates made from 250 µl of pre-culture from a genotype A strain and 250 µl of pre-culture from a genotype G strain mixed in 9.5 ml of Shieh medium were put in two different temperatures (26°C and 31°C, 5 replicates per temperature). After 24 hours, the overnight competition mixture was diluted 100-fold (0.1 ml into 9.9 ml) into fresh medium, and this was repeated every 24 hours for 14 days. Total DNA was extracted immediately after mixing the pre-cultures at the start of the experiment (day 0) and from overnight cultures prior to dilution into fresh medium on day 1, 2, 4, 6, 7, and 15. The extracted DNA was stored at -20°C until further analysis with HRM assay (below).

HRM assay development

HRM analysis uses a saturating, double-stranded DNA-binding dye, which specifically intercalates with double stranded DNA. In the post-PCR step, when the dsDNA dissociates into single strands, there is no longer any double stranded DNA present and thus fluorescence is reduced. The change in fluorescence is plotted against the temperature, generating a melting curve characteristic of the amplicon [16, 17]. The six strains used in this study (Table 1) served as controls for developing the HRM assay based on real-time PCR associated with melting curve analysis in order to quantify the proportions of strains of G genotype and strains of A genotype in a mixed culture. These two genotypes exhibited three single-nucleotide polymorphisms (SNPs) in the *trpB* gene, in positions 222, 587, and 590 of the amplified DNA fragments of the gene (NCBI Reference Sequence: LN624115, LN624122) [20]. HRM assay was not able to correctly discriminate between targeted genotypes and amplicons containing two or more SNPs. In this study, therefore, we describe the development of a real-time PCR-HRM technique targeting only one SNP (position 222) for rapid discrimination between the two genotypes, as described below. Extension of this method to discriminate between several genotypes, by multiplexing, is described in Appendix S1 (See also Fig. S4).

DNA extraction

Bacterial cells were harvested from overnight cultures from the competition experiment by centrifugation at 6000 g for 10 min at room temperature, and

genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega, USA). The DNA concentration and purity after extraction was estimated using the Qubit® Fluorometer (Life Technologies, Carlsbad, USA). DNA quality was assessed on an agarose gel and the final concentration was adjusted to 5 ng μl^{-1}.

Primer design and PCR amplification

Forward (5′-CGCAGAAGTCCGTCCTG-3′) and reverse (5′-AAAGGTATCTAGCGGATTGTT-3′) primers were designed to target 97 bp of the *trpB* gene (bp 171–267), spanning a single SNP (T>G substitution) at position 222 of the gene. This primer was designed using the Primer-BLAST software (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) (Fig.S1). Optimal primer annealing temperatures were established by temperature gradient PCR and HRM.

PCR amplification and HRM analysis were performed using CFX Real Time PCR Detection System (Bio-Rad Laboratories, USA). The PCR amplifications were performed and monitored using the CFX Manager Software, and the HRM data was analyzed with the Bio-Rad Precision Melt Analysis. PCR amplification was performed in a total volume of 10 μL, containing 5 µL of 1x Precision Melt Supermix (Bio-Rad Laboratories, USA), 1 μL 200 nM of each designed primer, and 4 μL 5 ng µl-1 of the DNA template. All samples were amplified in triplicate. The PCR reaction started at 95°C for 1 minute for initial denaturation, followed by 40 cycles of 30 seconds at 95°C for denaturation, 30 seconds at 63°C for annealing and another 30 seconds at 72°C for extension. The PCR amplification was then followed by heteroduplex formation by heating at 95°C for 30 seconds and subsequent cooling at 60°C for 1 minute. The high-resolution melting analysis was performed immediately afterwards by increasing the temperature from 65°C to 95°C by steps of 0.2°C maintained for 10 seconds each.

Primer efficiency

The primer efficiency was assessed using both G and A genotypes using 10-fold serial dilutions of DNA of each genotype. The slopes of the standard curves (an indicator of PCR efficiency) were similar, ranging from −3.45 (genotype A) to −3.56 (genotype G), demonstrating that the assays should be able to detect each genotype with similar high efficiencies ranging from 94.8% (genotype A) to 90.8% (genotype G).

Sequencing of targets

Real-time PCR products were subsequently sequenced to verify the identity of the amplified sequences. Sanger sequencing was performed on the same amplicons as used for HRM analysis. 5 µl of PCR products were purified with exonuclease I and Fast-AP (Thermo Fisher Scientific, Waltham, USA) for 15 min at 37°C and 15 min by 80°C. A sequencing reaction was set up with 1 µl of purified PCR products and the BigDye® Terminator v1.1 Cycle Sequencing Kit (Life Technologies). Briefly, each 20 µl sequencing reaction mixture contained 1 µl of PCR amplicon, 0.16 µM

of either forward or reverse PCR primer, 0.5 µl of BigDye Ready Reaction Mix, and 1 X sequencing buffer. The sequencing reaction conditions were as follows: 30 cycles of denaturing at 96°C for 10 seconds, annealing at 50°C for 5 seconds, and extension at 60°C for 4 minutes. The sequencing products were purified using ethanol/EDTA/sodium acetate precipitation. Sequencing was performed on an ABI 3130xl 16-capillary automated genetic analyser.

Quantification of mixed genotypes in competition cultures

For each HRM run, reference samples consisting of purified DNA from genotypes G and A pooled in known proportions (13 different proportions of the two genotypes: 100%, 95%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5% and 0%) were included in duplicate in the plate, and a no template control (NTC) was included in triplicate in the plate. Samples from the competition experiment for which genotype proportions were to be determined were run in triplicate. For each well, melting temperature was defined as the temperature of the inflexion point of the melting curve, using normalized relative fluorescent units (RFU). This point was determined by searching for the minimum value of the first-order derivative of the melting curve. Within each plate, the reference samples were used to calibrate the relationship between genotype proportion and melting temperature and to estimate the genotype proportions in the experimental samples. A linear model of the form [$meltingTemp = a + b * genotypeProp$] was fitted to the calibration data. The genotype proportions in the experimental samples were estimated by using a piecewise-defined function using the relation [$genotypeProp = 1/b * (meltingTemp - a)$] and setting the estimated values to 0% if they were negative and 100% if they were greater than 100% (Fig. S3). Prior to genotype proportion estimation experimental sample, wells for which the melting temperature was outside the range of values observed for calibration samples by more than 10% the calibration range span, were discarded. Confidence intervals were calculated according to Lavagnini *et al.* [37].

Growth

The growth rate and yield of isolates were measured in liquid culture using a temperature-controlled spectrophotometer (Bioscreen C, Growth Curves Ltd, Helsinki, Finland) [21]. Briefly, after isolate revival at 26 °C for 24h on a shaker (120 rpm.), the optical density (OD, at 570 nm) of the bacterial cultures was measured with a spectrophotometer adjusted to same level (0.10–0.20). After OD adjustment, 40 µl of each isolate was inoculated into 400 µL of fresh Shieh liquid medium on a Bioscreen C 100 wells plates in five replicates and randomized order. Plates were placed in Bioscreen at 26°C and 31°C for 4 days until growth in all wells stopped. To find the maximum growth rate (OD460–580 nm h^{-1}) and maximum population size (yield), we estimated the biomass with a MATLAB (version 2008b; MathWorks Inc., Natick, MA) script that fits linear regressions into ln-transformed population growth data consisting of 30 datapoints' sliding time window. The MATLAB code to perform these analyses is further described Ketola *et al.* [38].

Inhibition assays

Within each of the above mentioned competition groups, the inhibitory activity of *F. columnare* strains was tested reciprocally using an inhibition zone method with four replicates per assay, and a double layer method [21]. The optical density (OD, at 570 nm) of the bacterial cultures was first adjusted to same level (0.250-0.290). Fresh overnight-grown 'recipient' bacterial cultures in about 300 µl were mixed with 3 ml of molten soft Shieh agar (47°C), which was then poured on to the surface of dried Shieh agar. Five microliters of the unfiltered supernatant of the 'donor' cultures, which has been centrifuged at 17 000g for 3 min in room temperature, were spotted on the surface of the top agar. Plates were incubated for 48 h either at 26°C or 31°C; then, the plates were checked to determine whether the 'donor' strain had caused a growth inhibition of the underlying 'recipient' bacterial lawn.

ACKNOWLEDGMENTS

We would like to thank Dr. Ilkka Kronholm and Dr. Elina Laanto for providing constructive comments and help in improving the contents of this paper; and Dorrit Hämäläinen for help with the interference assays. This work was supported by KONE foundation (Roghaieh Ashrafi via project "Constraints of evolutionary adaptation to climate change" to Tarmo Ketola), OLVI foundation (Roghaieh Ashrafi #201620393), the Jane and Aatos Erkko Foundation (Lotta-Riina Sundberg), Finnish Cultural Foundation (Katja Pulkkinen) and Academy of Finland (Lotta-Riina Sundberg #272995, Tarmo Ketola # 278751, Jouni Taskinen # 260704 to Katja Pulkkinen) and Centre of Excellence in Biological Interactions (#252411, Prof. Johanna Mappes) to Roghaieh Ashrafi, Lotta-Riina Sundberg, and Tarmo Ketola.

DATA ACCESSIBILITY

DNA sequences used to develop the primers: NCBI Reference Sequence: LN624115, LN624122. Melting temperatures and calibration curve for HRM runs in supplementary material: Figure S3. Data from HRM measurements (melting curves from controlled mixes of genotypes and melting temperatures for experimental and calibration samples from the competition experiment) will be deposited in Dryad and made publicly available.

Competing interests

The authors declare no competing interests; financial or otherwise.

Authors' contributions

R.A., T.K. conceived and designed the experiment. R.A., M.B., L.S, K.P., T.K. wrote the paper. R.A., M.B. analysed the data.

REFERENCES

1. Solomon, S. *et al*. IPCC, 2007: summary for policymakers. *Climate change*, 93-129 (2007).

2. Davis, A. J., Jenkinson, L. S., Lawton, J. H., Shorrocks, B. & Wood, S. Making mistakes when predicting shifts in species range in response to global warming. *Nature* **391**, 783-786 (1998).

3. Pearson, R. G. & Dawson, T. P. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecol. Biogeogr.* **12**, 361-371 (2003).

4. Araújo, M. B. & Luoto, M. The importance of biotic interactions for modelling species distributions under climate change. *Global Ecol. Biogeogr.* **16**, 743-753 (2007).

5. Tylianakis, J. M., Didham, R. K., Bascompte, J. & Wardle, D. A. Global change and species interactions in terrestrial ecosystems. *Ecol. Lett.* **11**, 1351-1363 (2008).

6. Kordas, R. L., Harley, C. D. & O'Connor, M. I. Community ecology in a warming world: the influence of temperature on interspecific interactions in marine systems. *J. Exp. Mar. Biol. Ecol.* **400**, 218-226 (2011).

7. Blois, J. L., Zarnetske, P. L., Fitzpatrick, M. C. & Finnegan, S. Climate change and the past, present, and future of biotic interactions. *Science* **341**, 499-504 (2013).

8. Höckendorff, S., Früh, D., Hormel, N., Haase, P. & Stoll, S. Biotic interactions under climate warming: temperature-dependent and species-specific effects of the oligochaete *Chaetogaster limnaei* on snails. *Freshwater Science* **34**, 1304-1311 (2015).

9. McQuillan, M. A. & Rice, A. M. Differential effects of climate and species interactions on range limits at a hybrid zone: potential direct and indirect impacts of climate change. *Ecol. Evol.* **5**, 5120-5137 (2015).

10. Case, T. J., Holt, R. D., McPeek, M. A. & Keitt, T. H. The community context of species' borders: ecological and evolutionary perspectives. *Oikos* **108**, 28-46 (2005).

11. Orvos, D. R., Lacy, G. H. & Cairns, J. Genetically Engineered *Erwinia carotovora*: Survival, Intraspecific Competition, and Effects upon Selected Bacterial Genera. *Appl. Environ. Microbiol.* **56**, 1689-1694 (1990).

12. Minter, E. J., Lowe, C. D., Brockhurst, M. A. & Watts, P. C. A rapid and cost-effective quantitative microsatellite genotyping protocol to estimate intraspecific competition in protist microcosm experiments. *Methods Ecol. Evol.* **6**, 315-323 (2015).

13. Blot, M., Hauer, B. & Monnet, G. The Tn5 bleomycin resistance gene confers improved survival and growth advantage on *Escherichia coli*. *Molecular and General Genetics MGG* **242**, 595-601 (1994).

14. Ramadhan, A. A. & Hegedus, E. Survivability of vancomycin resistant enterococci and fitness cost of vancomycin resistance acquisition. *J. Clin. Pathol.* **58**, 744-746 (2005).

15. O'Keefe, K. J., Morales, N. M., Ernstberger, H., Benoit, G. & Turner, P. E. Laboratory-dependent bacterial ecology: a cautionary tale. *Appl. Environ. Microbiol.* **72**, 3032-3035 (2006).

16. Gundry, C. N. *et al*. Amplicon melting analysis with labeled primers: a closed-tube method for differentiating homozygotes and heterozygotes. *Clin. Chem.* **49**, 396-406 (2003).

17. Wittwer, C. T., Reed, G. H., Gundry, C. N., Vandersteen, J. G. & Pryor, R. J. High-resolution genotyping by amplicon melting analysis using LCGreen. *Clin. Chem.* **49**, 853-860 (2003).

18. Tong, S. Y. & Giffard, P. M. Microbiological applications of high-resolution melting analysis. *J. Clin. Microbiol.* **50**, 3418-3421 (2012).

19. Karvonen, A., Rintamäki, P., Jokela, J. & Valtonen, E. T. 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 (2010).

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

21. Sundberg, L. R. *et al*. Intensive aquaculture selects for increased virulence and interference competition in bacteria. *Proc. Biol. Sci.* **283**, 20153069 (2016).

22. Anacker, R. L. & Ordal, E. J. Studies on the myxobacterium *Chondrococcus columnaris*. II. Bacteriocins. *J. Bacteriol.* **78**, 33-40 (1959).

23. Riley, M. A. & Wertz, J. E. Bacteriocins: evolution, ecology, and application. *Annual Reviews in Microbiology* **56**, 117-137 (2002).

24. Brown, S. P., Fredrik Inglis, R. & Taddei, F. Evolutionary ecology of microbial wars: within-host competition and (incidental) virulence. *Evol. Appl.* **2**, 32-39 (2009).

25. Hibbing, M. E., Fuqua, C., Parsek, M. R. & Peterson, S. B. Bacterial competition: surviving and thriving in the microbial jungle. *Nature Rev. Microbiol.* **8**, 15-25 (2010).

26. Majeed, H., Lampert, A., Ghazaryan, L. & Gillor, O. The weak shall inherit: bacteriocin-mediated interactions in bacterial populations. *PLoS ONE* **8**, e63837 (2013).

27. Majeed, H., Ghazaryan, L., Herzberg, M. & Gillor, O. Bacteriocin expression in sessile and planktonic populations of *Escherichia coli*. *J. Antibiot.* **68**, 52 (2015).

28. Cornforth, D. M. & Foster, K. R. Competition sensing: the social side of bacterial stress responses. *Nature Rev. Microbiol.* **11**, 285-293 (2013).

29. Wloch-Salamon, D. M., Gerla, D., Hoekstra, R. F. & de Visser, J. A. Effect of dispersal and nutrient availability on the competitive ability of toxin-producing yeast. *Proc. Biol. Sci.* **275**, 535-541 (2008).

30. Harvell, C. D. *et al*. Climate warming and disease risks for terrestrial and marine biota. *Science* **296**, 2158-2162 (2002).

31. Pulkkinen, K. *et al*. Intensive fish farming and the evolution of pathogen virulence: the case of columnaris disease in Finland. *Proc. Biol. Sci.* **277**, 593-600 (2010).

32. Bottone, E. J., Weitzman, I. & Hanna, B. A. Rhizopus rhizopodiformis: emerging etiological agent of mucormycosis. *J. Clin. Microbiol.* **9**, 530-537 (1979).

33. Mataragas, M., Metaxopoulos, J., Galiotou, M. & Drosinos, E. Influence of pH and temperature on growth and bacteriocin production by *Leuconostoc mesenteroides* L124 and *Lactobacillus curvatus* L442. *Meat Sci.* **64**, 265-271 (2003).

34. Aasen, I., Møretrø, T., Katla, T., Axelsson, L. & Storrø, I. Influence of complex nutrients, temperature and pH on bacteriocin production by *Lactobacillus sakei* CCUG 42687. *Appl. Microbiol. Biotechnol.* **53**, 159-166 (2000).

35. Louhi, K. R., Sundberg, L. R., Jokela, J. & Karvonen, A. Interactions among bacterial strains and fluke genotypes shape virulence of co-infection. *Proc. Biol. Sci.* **282**, 20152097 (2015).

36. Shieh, H. Studies on the nutrition of a fish pathogen, *Flexibacter columnaris*. *Microbios Letters* **13**, 129-133 (1980).

37. Lavagnini, I. & Magno, F. A statistical overview on univariate calibration, inverse regression, and detection limits: application to gas chromatography/mass spectrometry technique. *Mass Spectrom. Rev.* **26**, 1-18 (2007).

38. Ketola, T. *et al*. Fluctuating temperature leads to evolution of thermal generalism and preadaptation to novel environments. *Evolution* **67**, 2936-2944 (2013).

Figure 1. High-resolution DNA melting curve analysis results for a 97-bp amplicon containing one SNP in *trpB* gene for *Flavobacterium columnare*. (Right) Normalized melting curves, (Left) derivative plots. The melting curves depict pure genotype A, pure genotype G, and 50/50 ratio of both genotypes. Solid and dashed lines represent replicate measurements.

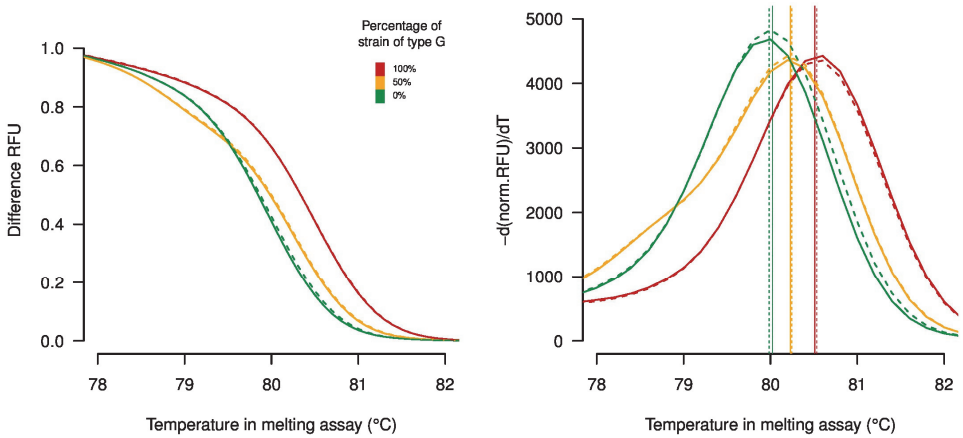


Figure 2. (Right) Normalized melting curves and (Left) difference plots of reference samples consisting of purified DNA from genotypes G and A pooled in known proportions (13 different proportions of the two genotypes: 100%, 95%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5% and 0%). Solid and dashed lines represent replicate measurements.

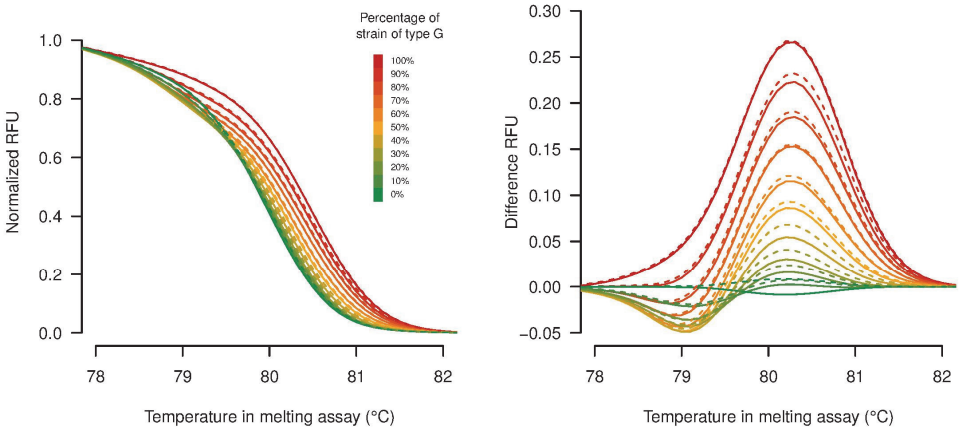


Figure 3. Genotype proportions estimated by HRM runs after 1, 2, 4, 6, 7 and 15 days of competition in liquid medium. For each day, vertical bars show the relative proportions of each competing genotype in five experimental replicates within each competition group (A, B, C and D) and black lines indicates the 90% confidence interval for each proportion estimate based on the triplicate HRM samples. The ordering of experimental replicates in the bar plots are the same from day 1 to day 15 within each group x temperature treatment. The names of the competing genotypes are indicated on the vertical axes. Hatched bars indicate replicates for which HRM data could not be used to estimate genotype proportions. For each day within each competition group, symbols between the 26°C and 31°C bar plots indicate the p-values of a Welch's t-test comparing the genotype proportions for the 26°C replicates and for the 31°C replicates (*** < 0.001 < ** < 0.01 < * < 0.05).

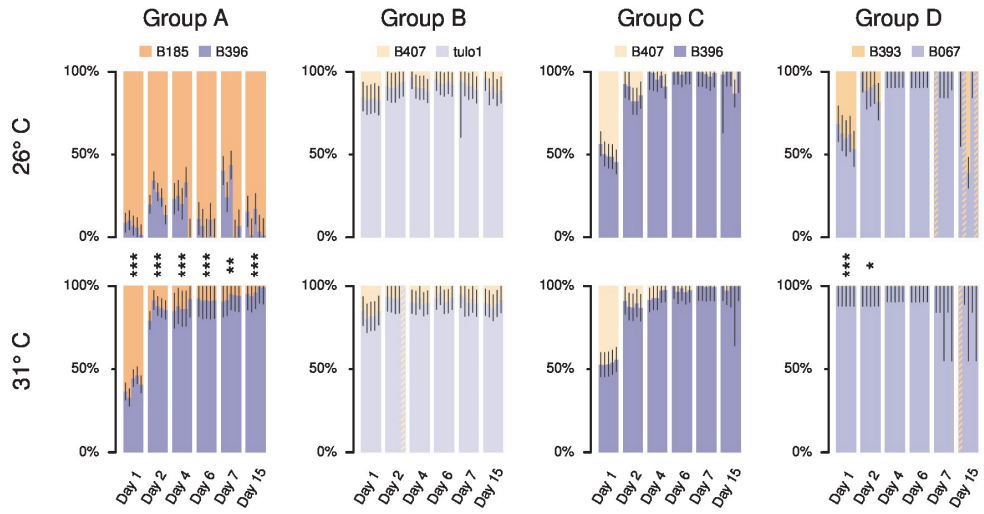


Table 1. Details of *Flavobacterium columnare* strains used in this study, and their performance in competition pairs (A-D, + indicates winner strain in competition). Growth and yield equals growth rate and yield measured in liquid medium, and inhibition (+) refers to which strain inhibited which in competition pairs. **In bold**: indicates which strains of a given pair has the fastest growth rate or highest yield at a given temperature.

Strains	Genotype			Competition Group	26°C				31°C			
		Year	Origin		Comp.	Growth	Yield	Inhibition	Comp.	Growth	Yield	Inhibition
B185	G	2009	Fish farm	A	+	**0.199± 0.020 (5)**	**0.788 ± 0.088 (5)**	-	-	0.199 ± 0.012 (4)	0.198 ± 0.006 (4)	No inhib.
B396	A	2010	Fish farm		-	0.158 ± 0.028 (5)	0.699 ± 0.052 (5)	+	+	**0.200 ± 0.010 (4)**	**0.298 ± 0.038 (4)**	No inhib.
B407	G	2010	Wild	B	-	0.138 ± 0.038 (5)	0.717 ± 0.054 (5)	-	-	0.184 ± 0.044 (5)	0.137 ± 0.022 (5)	No inhib.
TULO1	A	2010	Fish farm		+	**0.220 ± 0.022 (3)**	**0.779 ± 0.106 (3)**	+	+	**0.208 ± 0.022 (5)**	**0.285 ± 0.068 (5)**	No inhib.
B407	G	2010	Wild	C	-	0.138 ± 0.038 (5)	**0.717 ± 0.054 (5)**	-	-	0.184 ± 0.044 (5)	0.137 ± 0.022 (5)	No inhib.
B396	A	2010	Fish farm		+	**0.158 ± 0.028 (5)**	0.699 ± 0.052 (5)	+	+	**0.200 ± 0.010 (4)**	**0.298 ± 0.038 (4)**	No inhib.
B393	G	2010	Wild	D	-	0.189 ± 0.030 (5)	0.622 ± 0.046 (5)	+	-	0.190 ± 0.022 (5)	0.166 ± 0.006 (5)	No inhib.
B067	A	2007	Fish farm		+	0.189 ± 0.010 (5)	**0.773 ± 0.064 (5)**	-	+	**0.216 ± 0.016 (5)**	**0.347 ± 0.070 (5)**	No inhib.

SUPPLEMENTARY MATERIAL

Figure S1. Alignment of Forward and reverse primers that were designed to target 97 bp of the trpB gene (bp 171–267), spanning a single SNP (T>G substitution) at position 222 of *trpB* gene.

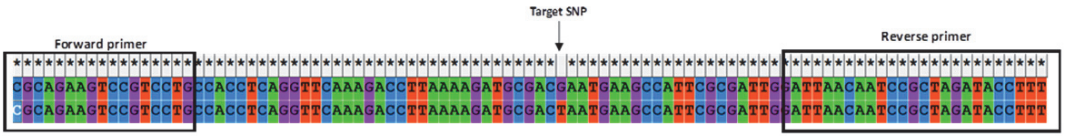


Figure S2. Chromatogram of reference samples by Sanger sequencing. 13 different proportions of the two genotypes: 100%, 95%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5% and 0%.

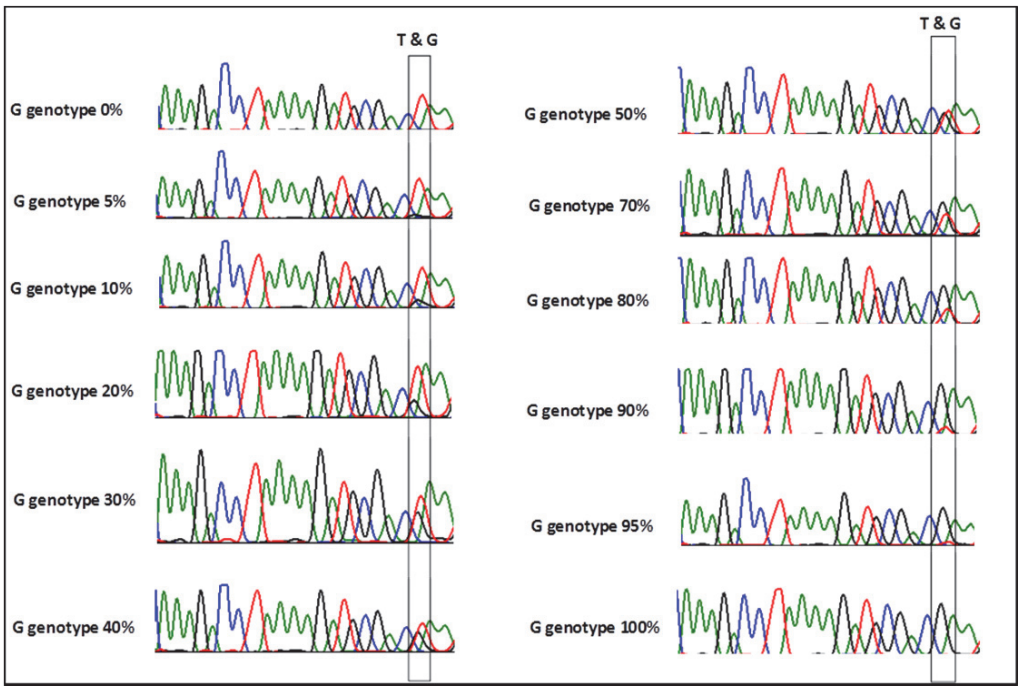
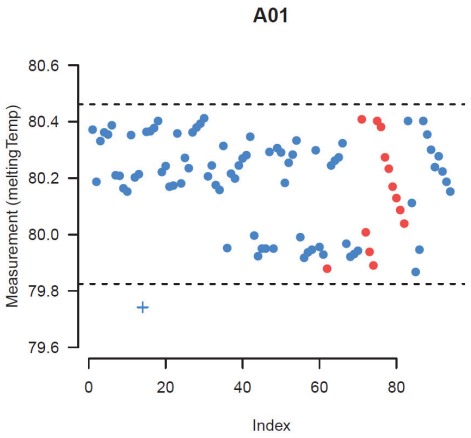
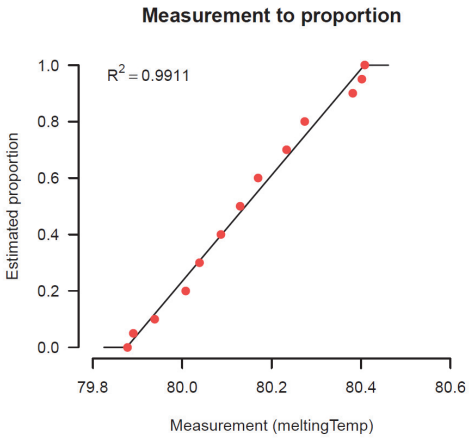
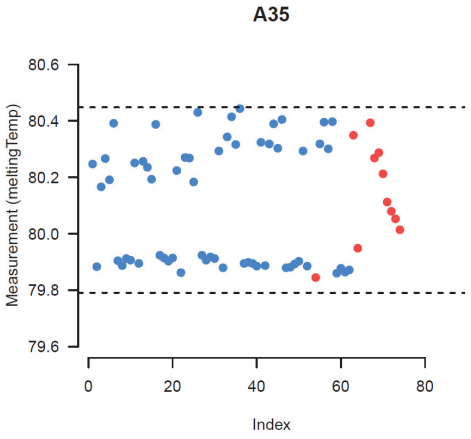
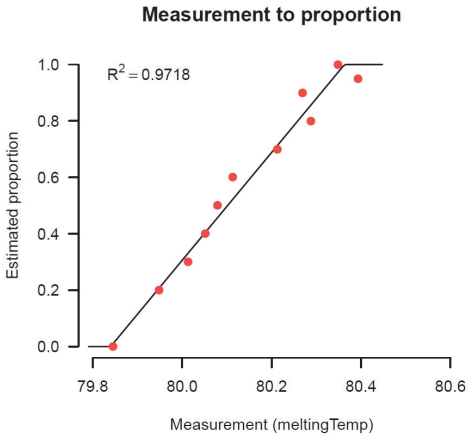


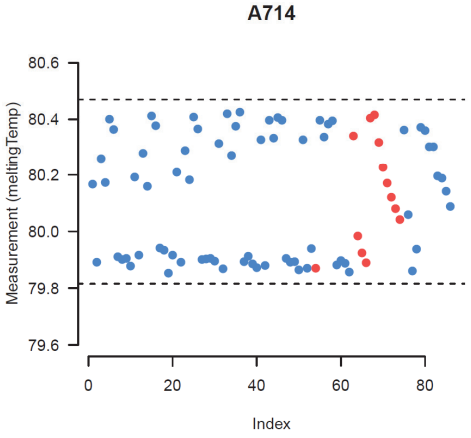
Figure S3. Melting temperatures and calibration curve for HRM runs. Each horizontal pair of panels corresponds to one run. Plate run name is indicated as the title of left panels ("A01", "A35," ...). Left panels, melting temperature values observed for each plate well. Experimental samples are marked with blue and calibration sample of known genotype proportions marked with red. Dashed lines indicate the range of melting temperature values observed for calibration samples +/- 10% of the range span. Wells for which melting temperature values were outside this range (blue crosses) were discarded from the analysis. Right panels, calibration curve used to estimate genotype proportions in experimental samples. Red dots are calibration samples. R^2 values of the linear model *proportion ~ meltingTemp* fitted on the calibration sample data are reported. Estimated genotype proportions for experimental samples which were below 0 or above 1 according to the linear fit model were set to 0 or 1, respectively, as shown by the plotted piecewise function.

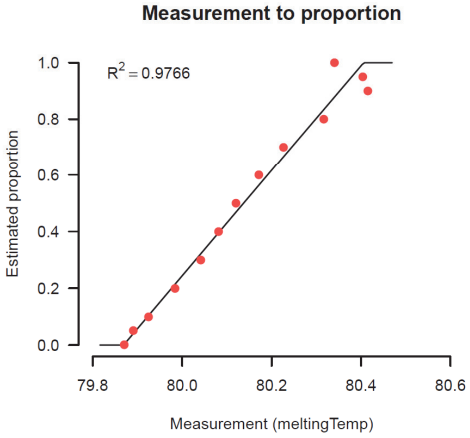

A01
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index

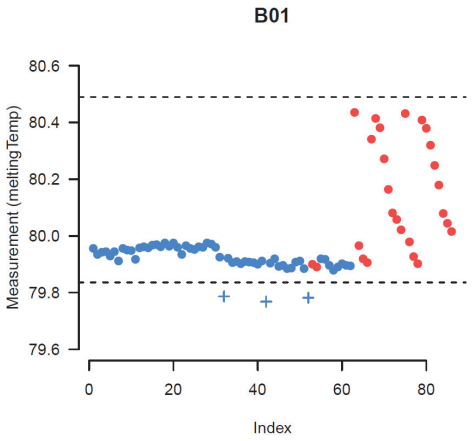

Measurement to proportion
$R^2 = 0.9911$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)

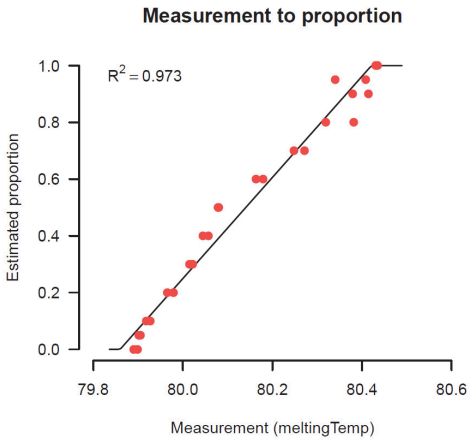

A35
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index

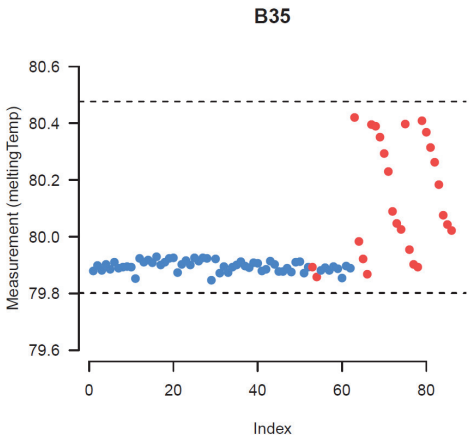

Measurement to proportion
$R^2 = 0.9718$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)

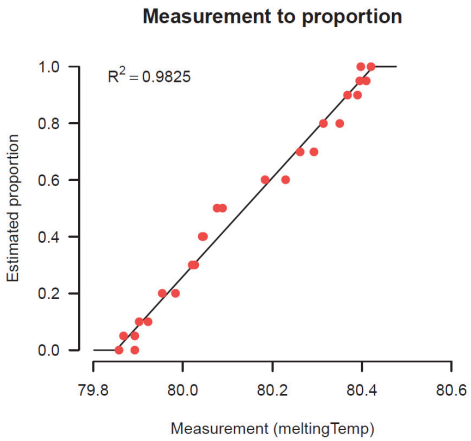

A714
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index

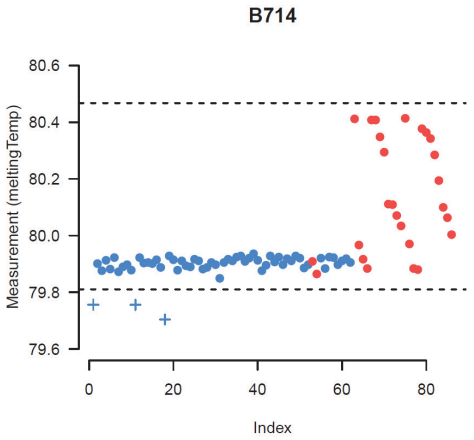

Measurement to proportion
$R^2 = 0.9766$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)

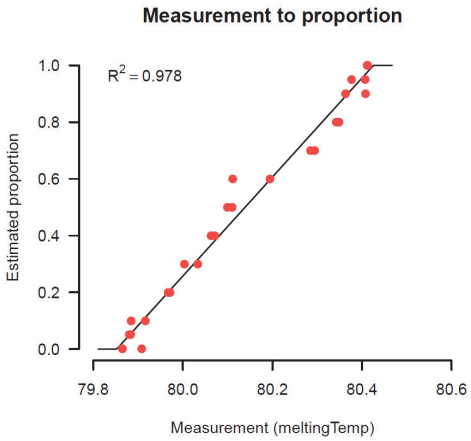

B01
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index

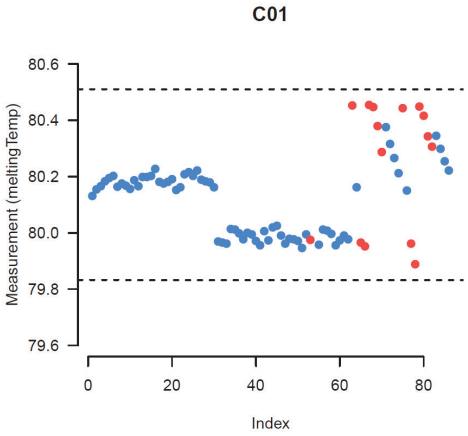

Measurement to proportion
$R^2 = 0.973$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)

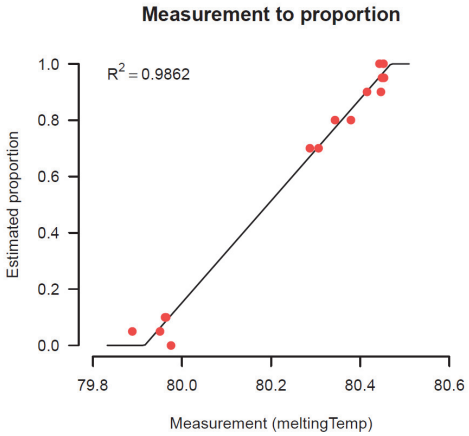

B35
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index

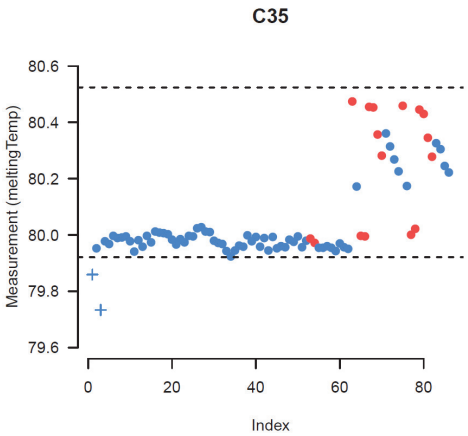

Measurement to proportion
$R^2 = 0.9825$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)

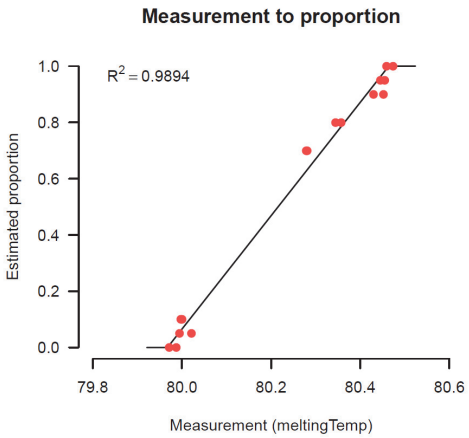

B714
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index

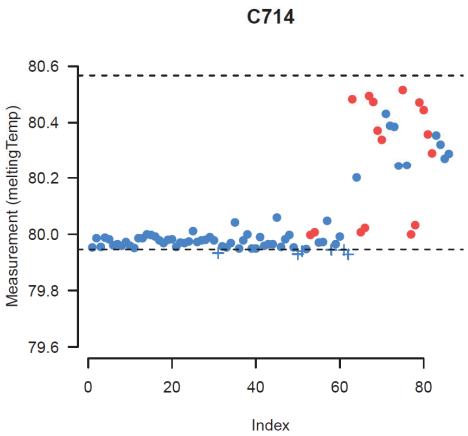

Measurement to proportion
$R^2 = 0.978$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)

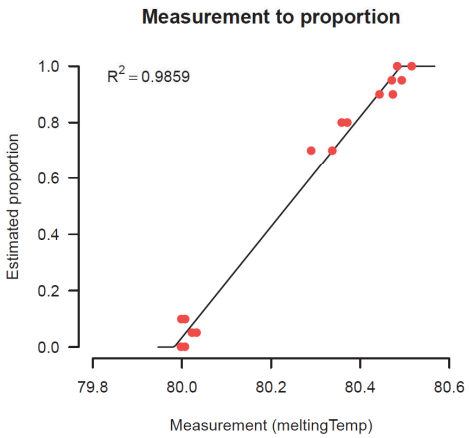

C01
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index

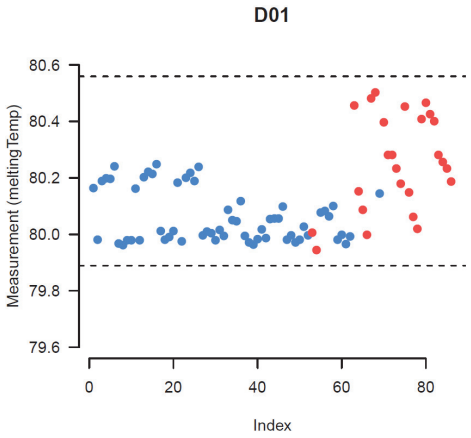

Measurement to proportion
$R^2 = 0.9862$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)

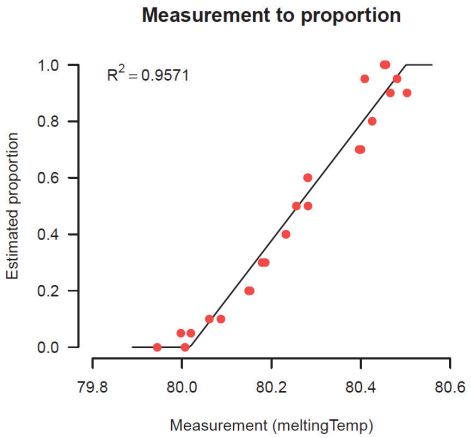

C35
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index

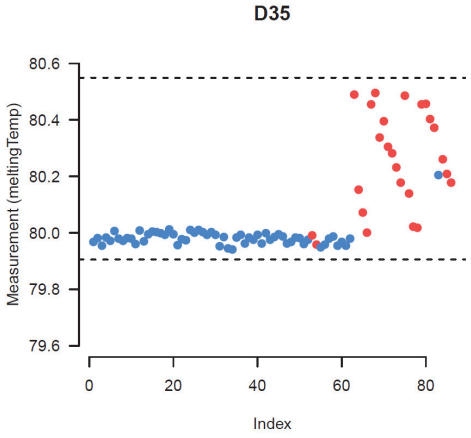

Measurement to proportion
$R^2 = 0.9894$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)

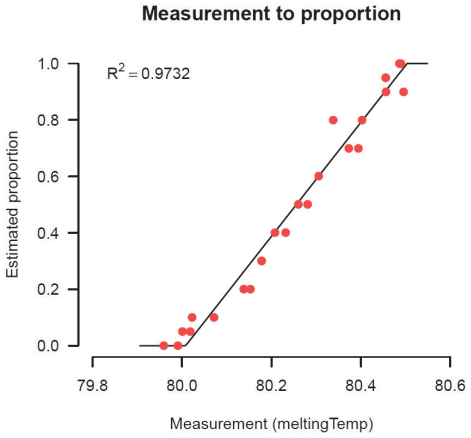

C714
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index

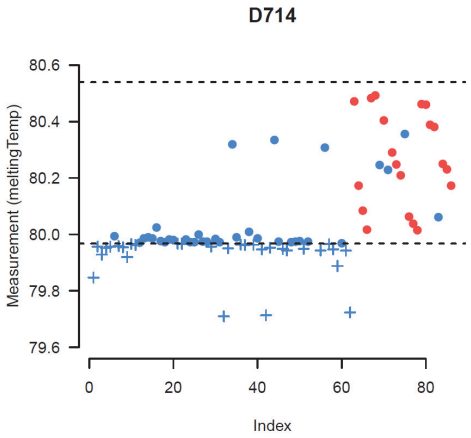

Measurement to proportion
$R^2 = 0.9859$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)


D01
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index


Measurement to proportion
$R^2 = 0.9571$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)


D35
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index


Measurement to proportion
$R^2 = 0.9732$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)


D714
Measurement (meltingTemp)
80.6
80.4
80.2
80.0
79.8
79.6
0
20
40
60
80
Index

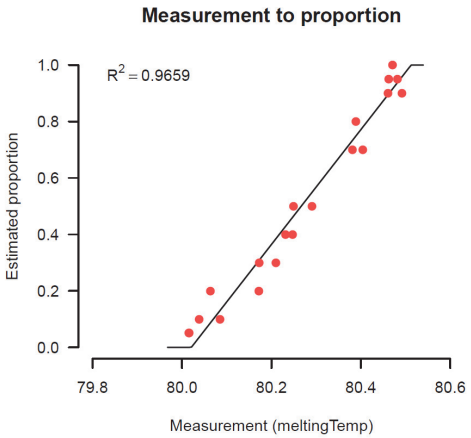
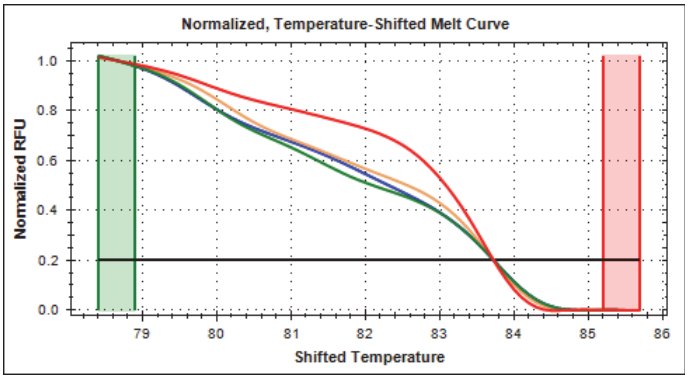
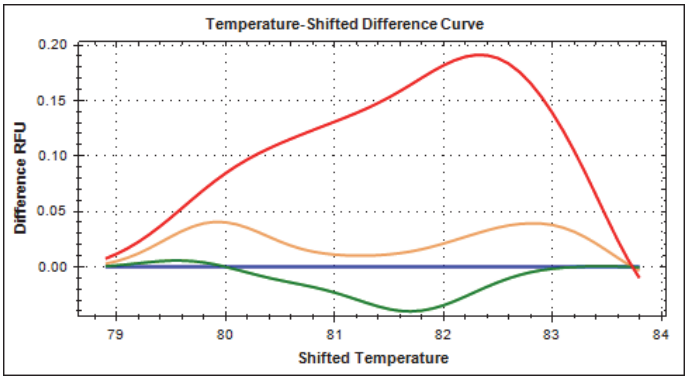

Measurement to proportion
$R^2 = 0.9659$
Estimated proportion
1.0
0.8
0.6
0.4
0.2
0.0
79.8
80.0
80.2
80.4
80.6
Measurement (meltingTemp)

Figure S4. (A) High-resolution melt normalization and **(B)** differential plots depicting four *Flavobacterium columnare* genotypes by HRMA (See the method in **Appendix S1.**)

(A)



(B)



Appendix S1.

The duplex HRM assay developed for supplementry use in identification of *F. columnare* genotypes (C, E, G, A) that has been found in Finland so far. The PCR/HRM primers were designed to individually and simultaneously detect above genotypes. Primer sets have been designed that target different variable regions of *trpB* gene in *F. columnare*. The sequence of used primers was 5-CCGCATCATAGCCGAAACA-3 (CE-forward) and 5-GACGTTTGATGTCGATTTCGC-3 (CE-reverse reverse) for genotypes C and E amplification, and 5-CGCAGAAGTCCGTCCTG-3 (GA1-forward) and 5-AAAGGTATCTAGCGGATTGTT-3 (GA1-reverse) for genotypes G and A amplification. Primers amplified 113- and 97-bp products that corresponded to nucleotides 24 to 137 and 171 to 267, respectively, in *trpB* gene (GenBank accession numbers: LN624115 for genotype G, LN624121for genotype C, LN624122 for genotype A, LN624123 for genotype E)

Briefly, PCR amplification was performed in a total volume of 20 µL, containing 10ul of 1x Precision Melt Supermix (Bio-Rad Laboratories, USA), 1.2 µl of CE primer mix (final concentration 200 µm) primer, 0.8 µl of CE primer mix (final concentration 200nm), 3 µl of PCRgrade water and 5 µl 5 ng the DNA template . All samples were performed in triplicate. The PCR condition was started at 95°C for 1 minute for initial denaturation, followed by 40 cycles of 30 seconds at 95°C for denaturation, 30 seconds at 63°C for annealing and another 30 seconds at 72°C for extension. The PCR amplification was then followed by heteroduplex formation at 95°C for 30 seconds and subsequently 60°C for 1 minute. The high-resolution melting analysis was performed immediately afterwards by increasing the temperature from 65°C to 95°C by steps of 0.2°C maintained for 10 seconds each.

HRM analysis clearly separated 4 of aimed *F. columnare* genotypes with distinctive melting curves Supplementary_Figure 4.