University of Glasgow

Comparative phenotypic, proteomic and genomic approaches to assess lipopolysaccharide and outer membrane protein diversity among isolates of *Yersinia ruckeri* recovered from Atlantic salmon and rainbow trout

A thesis submitted to the University of Glasgow for the degree of Doctor of Philosophy

Michael J Ormsby
BSc (Hons)

Submitted September, 2015

Institute of Infection, Immunity and Inflammation
College of Medical, Veterinary and Life Sciences
University of Glasgow

Declaration

I hereby declare that this thesis is the result of my own work and has been composed for the degree of PhD at the University of Glasgow. This work has not been submitted for any other degree at this or any other institution. All work presented was performed by myself unless otherwise stated. All sources of information and contributions to the work have been specifically acknowledged in the text.

Michael J Ormsby

September 2015

Acknowledgments

Firstly, I would like to express my sincerest gratitude to my supervisor Dr. Robert Davies. His supervision, support, guidance and encouragement over the last four years has given me a solid training as a research scientist and made the entire experience an enjoyable one. Secondly, I wish to thank Dr. Richard Burchmore, my secondary supervisor, for the help and advice he offered throughout my PhD. Thank you to my assessors Dr. Dónal Wall and Dr. Dan Walker for their useful advice and discussions regarding my project over the years, also.

Over the course of my PhD I have been fortunate enough to share a research group with several people. I would like to thank Dr. Chonchanok Theethakaew (Mod), Dr. Teerasak E-Komon (Art), Mr. Edward Grahame, Miss Rezheen Abdulrahman, Miss Nurshahira Sulaiman, Dr. Nicky O'Boyle, Miss Erin Sutherland and Mr. Daniel Cozens for their support and guidance – but mostly for putting up with me!

All current members of the level 2 GBRC and those that have come and gone over the years have influenced my development as a researcher - I am forever grateful for that. However, I must give special mention to several people that have helped keep me sane during the writing of my thesis and constantly reassured me that 'it's not that bad!' Thank you to Dr. Khuram Ashraf, Miss Síle Johnson, Dr. James Connolly, Mr. Edward Grahame and Dr. Nicky O'Boyle. Similarly, thank you to all members of the Douce group for putting up with me in the office (and trying to talk about football and rugby with me…).

I must thank my industrial collaborators for the useful discussions we have had over the years, and their encouragement. Particularly Dr. Tim Wallis (Ridgeway Biologicals) for his support and Dr. David Verner-Jeffreys for making me so welcome whenever I travel down to CEFAS. Special mention to Bob Hatfield for housing me and making my time in Weymouth unforgettable!

Furthermore, I would like to acknowledge Miss Suzanne Eadie and Ms. Julie Galbraith of Glasgow polyomics for running proteomic and genomic samples. I would also like to thank Dr. Neil Inglis and Mr. Kevin Mclean at the Moredun research institute in Edinburgh for their useful proteomic advice.

Thank you to the University of Glasgow and to the BBSRC for their award of my industrial CASE PhD studentship.

My parents, Nicholas and Alison, and my brother and sister, Ross and Siân, have supported me throughout my studies and encouraged me to pursue whatever avenue makes me happy. Supporting me in my decision to move to Scotland turned out all right in the end!

Finally, I owe my biggest gratitude to my wife, Sarah. You provide unrivalled support for me always, and encourage me to be the best I can be. Your patience and support throughout the highs and lows is more important than you know, and I am eternally grateful.

Presentations and publications

1. Ormsby, M. J., Verner-Jeffreys, D. W., Burchmore, R. J., Wallis, T. and Davies, R. L. **Characterisation of *Yersinia ruckeri* isolates from UK Atlantic salmon and rainbow trout. Oral presentation** at the Aquaculture UK conference, 23rd – 24th May 2012, Aviemore, United Kingdom.

2. Ormsby, M. J., Verner-Jeffreys, D. W., Hopewell, R., Wallis, T., Burchmore, R. J., and Davies, R. L. ***Yersinia ruckeri* isolates recovered from UK Atlantic salmon exhibit greater diversity than rainbow trout isolates. Poster presentation** at the CEFAS student open day, 2nd – 3rd June 2013, Lowestoft, United Kingdom.

3. Ormsby, M. J., Verner-Jeffreys, D. W., Hopewell, R., Wallis, T., Burchmore, R. J., and Davies, R. L. **Characterising *Yersinia ruckeri*: The emergence of a novel serotype. Poster presentation** at the Young microbiologist symposium (YMS), 2nd – 3rd June 2014, Dundee, United Kingdom.

4. Ormsby, M. J., Verner-Jeffreys, D. W., Hopewell, R., Wallis, T., Burchmore, R. J., and Davies, R. L. **Emergence of a novel isolate of the fish pathogen *Yersinia ruckeri* in farmed Atlantic salmon. Oral presentation** at the CEFAS Weymouth laboratory, 8th January 2015, Weymouth, United Kingdom.

5. Ormsby, M. J., Verner-Jeffreys, D. W., Hopewell, R., Wallis, T., Burchmore, R. J., and Davies, R. L. **Emergence of a novel O-serotype of *Yersinia ruckeri* in Atlantic salmon and rainbow trout. Oral and poster presentation** at SGM annual conference, 23rd – 24th March 2015, Birmingham, United Kingdom.

6. Ormsby, M. J., Verner-Jeffreys, D. W., Hopewell, R., Wallis, T., Burchmore, R. J., and Davies, R. L. **Emergence of a novel O-serotype of *Yersinia ruckeri* in Atlantic salmon and rainbow trout. Oral presentation** at the CEFAS student open day, 30th April – 1st May 2015, Lowestoft, United Kingdom.

7. Ormsby, M. J., Verner-Jeffreys, D. W., Hopewell, R., Wallis, T., Burchmore, R. J., and Davies, R. L. **Emergence of a novel O-serotype of *Yersinia ruckeri* in Atlantic salmon and rainbow trout. Poster presentation** at the ASM annual conference, 30th May – 2nd June 2015, New Orleans, USA.

Summary

Yersinia ruckeri is the causative agent of enteric redmouth (ERM) disease in farmed salmonids. ERM disease is traditionally associated with rainbow trout (*Oncorhynchus mykiss*¸ Walbaum), but the incidence of the disease in Atlantic salmon (*Salmo salar*) has increased in recent years. Historically, motile (biotype 1), serotype O1 isolates of *Y. ruckeri* have been mostly responsible for ERM in rainbow trout worldwide but non-motile (biotype 2), serotype O1 isolates have become increasingly prevalent in this species over wide geographic areas since their first isolation in the UK in the 1980s. *Yersinia ruckeri* isolates responsible for infection of salmon have been less well characterised than those from rainbow trout and little is known about their diversity. The emergence of new pathogenic strains, together with vaccine breakdown in the field, has emphasised the need for greater knowledge about strain diversity which may lead to the development of improved vaccines for both species. In the present study, a unique and extensive strain collection encompassing 135 isolates of *Y. ruckeri* were characterised using complementary phenotypic, proteomic and genomic approaches.

In the initial part of this thesis, 135 isolates recovered over a 14 year period in the UK from infected Atlantic salmon (109 isolates) and rainbow trout (26 isolates) were phenotypically characterised through biotype, serotype, and outer membrane protein (OMP) -type analysis. Atlantic salmon isolates represented a wider range of O-serotypes and associated lipopolysaccharide (LPS) types, and had more diverse OMP profiles than those from rainbow trout. Most significantly, a new O-serotype/LPS type was identified in 56 Atlantic salmon isolates; other Atlantic salmon isolates were represented by serotypes O1 (five isolates), O2 (34 isolates) and O5 (14 isolates). This new LPS type comprises a core polysaccharide region similar to that of serotype O1 but has a unique, previously unidentified O-antigen region. Atlantic salmon isolates could be assigned to one of four major OMP-types and to one of 11 OMP-sub types. Isolates recovered from rainbow trout were represented by the same non-motile clone that is responsible for the majority of ERM outbreaks in this species within the UK. This clone was not associated with any infected salmon. However, two isolates of the novel serotype O1/O5 were recovered from rainbow trout in 2010 and 2011. These data suggest that different *Y. ruckeri* strains are specifically

adapted to cause disease in either Atlantic salmon or rainbow trout. The efficacy of current vaccine formulations against different clonal groups must be examined.

Subsequently, an in-depth characterisation of the outer membrane (OM) proteome of isolates recovered from Atlantic salmon and rainbow trout was conducted. Outer membrane proteins are at the interface of host pathogen interactions, with important roles in adherence, evasion of host immune response, and transport. Using a bioinformatic prediction pipeline and four publicly available genomes, 141 proteins were confidently predicted to be OM associated. Subsequently, the OM proteomes of eight representative isolates (four from rainbow trout; four from Atlantic salmon) were analysed using a combination of gel-based and gel-free proteomic approaches. In total, 66 OMPs were identified through this combined approach, of which 28 were unique to the gel-free approach and 13 were unique to the gel-based approach. Further to this, the OM proteomes of these eight representative isolates were examined when cells were grown under conditions that aimed to mimic the *in vivo* and environmental conditions of *Y. ruckeri*. These included growing cells aerobically at 22°C and 28°C, anaerobically, under iron-depletion and in an artificial seawater medium at 22°C. In total, 76 OMPs were identified in all eight isolates under these growth conditions.

Finally, a phylogenetic study was undertaken whereby the genomes of 16 representative isolates encompassing a range of biotypes, serotypes, host species (eight from rainbow trout, seven from Atlantic salmon and one from European eel), geographic locations and dates of isolation were considered. A phylogenetic species tree based on the concatenated sequences of 19 housekeeping genes revealed host specific lineages suggesting an earlier host-associated evolutionary split within *Y. ruckeri*. Subsequent analysis of the presence, absence and variation of the nucleotide and amino acid sequences of the 141 predicted OMPs revealed high levels of conservation (with 120 OMPs showing less than 1% nucleotide variation). One hundred and thirty proteins were identified in all 16 genomes examined. However, 11 proteins were not, and these included invasins, OmpE and proteins involved in pilus biogenesis. Further examination of the OMPs OmpA and OmpF, which were identified in the

genomes of all 16 isolates, revealed variation in the surface exposed loop regions which may play a role in pathogenicity and/or host specificity.

This study represents a comprehensive characterisation of *Y. ruckeri* isolates recovered from Atlantic salmon and rainbow trout using a range of molecular techniques, and reveals important adaptations that the bacteria may make in order to survive both inside and outside of the host. Importantly, this study provides comprehensive support for future work involving this fish pathogen.

Table of Contents

List of Figures

List of Tables

Abbreviations

°C	Degrees Celsius
µg	Microgram
µl	Microlitre
µM	Micromolar
µm	Micrometre
1-D	One-Dimensional
2-D	Two-Dimensional
A	Adenine
Å	Angstrom
ABC	ATP-binding cassette
ACN	Acetonitrile
$AgNO_3$	Silver nitrate
APS	Ammonium persulphate
ATCC	American type culture collection
ATP	Adenosine triphosphate
BAM	β-barrel assembly machinary
Bis	N,N'-methylenebisacrylamide
bp	Base pairs
BSA	Bovine serum albumin
C	Cytosine
$CaCl_2$	Calcium chloride
c.f.u's	Colony forming units
Da	Dalton(s)
dH_2O	Distilled water
DNA	Deoxyriboucleic acid
DTT	Dithiothreitol
ECA	Enterobacterial common antigen
ECPs	Extracellular products
ELISA	Enzyme linked immunosorbant assay
emPAI	Exponentially modified protein abundance index
ERM	Enteric redmouth
ESI	Electrospray ionisation
et al.	*et alios* (and others)
FAO	Food and Agriculture Organisation
Fur	Ferric uptake regulator
g	Gram(s)
G	Guanine
h	Hour(s)
HCl	Hydrochloric acid
HGT	Horizontal gene transfer
HMM	Hidden Markov Model
HMW	High molecular weight
HRP	Horseradish-peroxidase
HV	Hypervariable
i.m.	Intramuscular
i.p.	Intraperitoneal
IAA	Iodoacetamide
IFAT	Immunofluoresence antibody technique
IHNV	Infectious haematopoetic necrosis virus
IM	Inner membrane
IMP	Inner membrane protein
ISA	Infectious salmon anaemia
Kb	Kilo base pairs
KCl	Potassion chloride
kDa	Kilo Dalton
l/L	Litre(s)
LAMP	Loop mediated isothermal amplification assay
LC	Liquid chromatography
LC-MS/MS	LC tandem Mass spectrometry
LMW	Low molecular weight

Lol	Lipoprotein localisation machinary
LOS	Lipooligosaccharide
LPS	Lipopolysaccharide
Lpt	LPS assembly machinary
M	Molar
m/z	Mass-to-charge
mA	milliamps
MALDI	Matrix assisted laser desorption/ionisation
Mbp	Megabase Pair
MDa	Mega Dalton
MEGA	Molecular Evolutionary Genetics Analysis
MFP	Membrane fusion protein
mg	Milligram
mg L^{-1}	Milligrams per litre
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulphate
min	Minute(s)
ml	Millilitre
MLEE	Multilocus enzyme electrophesis
MLST	Multilocus sequence type
mM	Millimolar
mm	Millimetre(s)
MOWSE	Molecular weight search
MS	Mass Spectrometry
MS222	Tricaine methane sulphonate
MW	Molecular weight
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NCBI	National Centre for Biotechnology information
ng	Nanogram
ng L^{-1}	Nanograms per litre
NH_4OH	Ammonium hydroxide
O-antigen	Lipopolysaccharide antigen
OD_{600}	Optical density at 600 nm
OM	Outer membrane
OMP	Outer membrane protein
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PFGE	Pulsed field gel electrophoresis
pH	A measure of solution acidity
PMF	Proton motive force
POTRA	Polypeptide transport-associated
Pt	Pulsotype
PTS	Phosphate-starvation inducible transport system
RAST	Rapid Annotation Subsystem Technology
RFLP	Restriction fragment-length polymorphism
RNA	Ribonucleic acid
ROD	Ribose ornithine deoxycholate
rpm	Revolutions per minute
RPS	Relative percentage survival
rRNA	Ribosomal RNA
s	Second(s)
Sarkosyl	N-Lauroylsarcosine
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
Sec	General secretory pathway
STs	Sequence types
T	Thymidine
T1SS	Type one secretion system
T2SS	Type two secretion system
TTSS	Type three secretion system
T4SS	Type four secretion system

T5SS	Type five secretion system
TAM	Translocation and assembly module
Tat	Twin arginine translocation pathway
TBS	Tris buffered saline
TCA	Trichloroacetic acid
TEM	Transmission electron microscopy
TEMED	N,N,N′,N′-Tetramethylethane-1,2-diamine
TFA	Trifluoroacetic acid
TLR	Toll like receptor
TOF	Time-of-flight
TPS	Two-partner secretion
Tris	2-amino-2-(hydroxymethyl) 1,3-propandiol
TSA	Tryptone soya agar
TSB	Tryptone soya broth
TTBS	0.05% Tween 20 in TBS
V	Volt(s)
v/v	Volume per volume
w/v	Weight per volume
WBC	White blood cell
WGS	Whole genome sequencing
x *g*	Centrifugal force
Yops	Yersinia outer proteins
Δ	Deletion

Chapter 1 Introduction

1.1 Aquaculture

Aquaculture refers to the breeding, rearing and harvesting of plants and animals in all types of water environments, including the "farming" of freshwater and marine species from hatchery to market. Aquaculture is emerging as the fastest growing food production industry in the world. With an ever increasing human population, space to enhance land-based food production is limited. Almost 70% of the area on earth is underwater meaning the expansion of aquaculture produced resources is inevitable (Ma *et al.*, 2011).

Marine aquaculture is the culturing of species that live in the ocean; while this can occur in the ocean (in cages, on the seafloor, or suspended in the water column), land-based manmade systems such as ponds or tanks are also commonly used. Freshwater aquaculture produces species that are native to streams, rivers and lakes, taking place primarily in ponds and in on-land, manmade systems such as recirculating aquaculture systems.

1.1.1 **World Aquaculture**

Worldwide, the aquaculture industry has grown at an average rate of 8.9% per annum since 1970, compared with a growth of 2.8% in terrestrial farmed meat in the same period (FAO, 2014). The value of aquaculture production worldwide in 2012 was estimated at $137 billion, with Asia generating substantially more aquaculture produce than the rest of the world combined (Fig. 1.1A). Freshwater aquaculture vastly out produces its marine counterpart also (Fig. 1.1B) (FAO, 2014).

While carps, tilapias, shrimps/prawns, oysters, clams and some seaweeds are farmed in greater quantities than salmonids, this group is still produced to very high levels (FAO, 2014). In salmonid aquaculture, Atlantic salmon and rainbow trout represent the most farmed species, with over 2000 million tonnes and 855 million tonnes of each species produced in 2012, respectively. The top three producers of Atlantic salmon are Norway, Chile and the UK (Fig. 1.2A), while rainbow trout production is dominated by Chile, Iran and Turkey (Fig. 1.2B).

The UK is currently the third largest Atlantic salmon producer, and the eleventh largest producer of rainbow trout. The majority of Atlantic salmon and rainbow trout aquaculture in the UK takes place in Scotland (FAO, 2014).

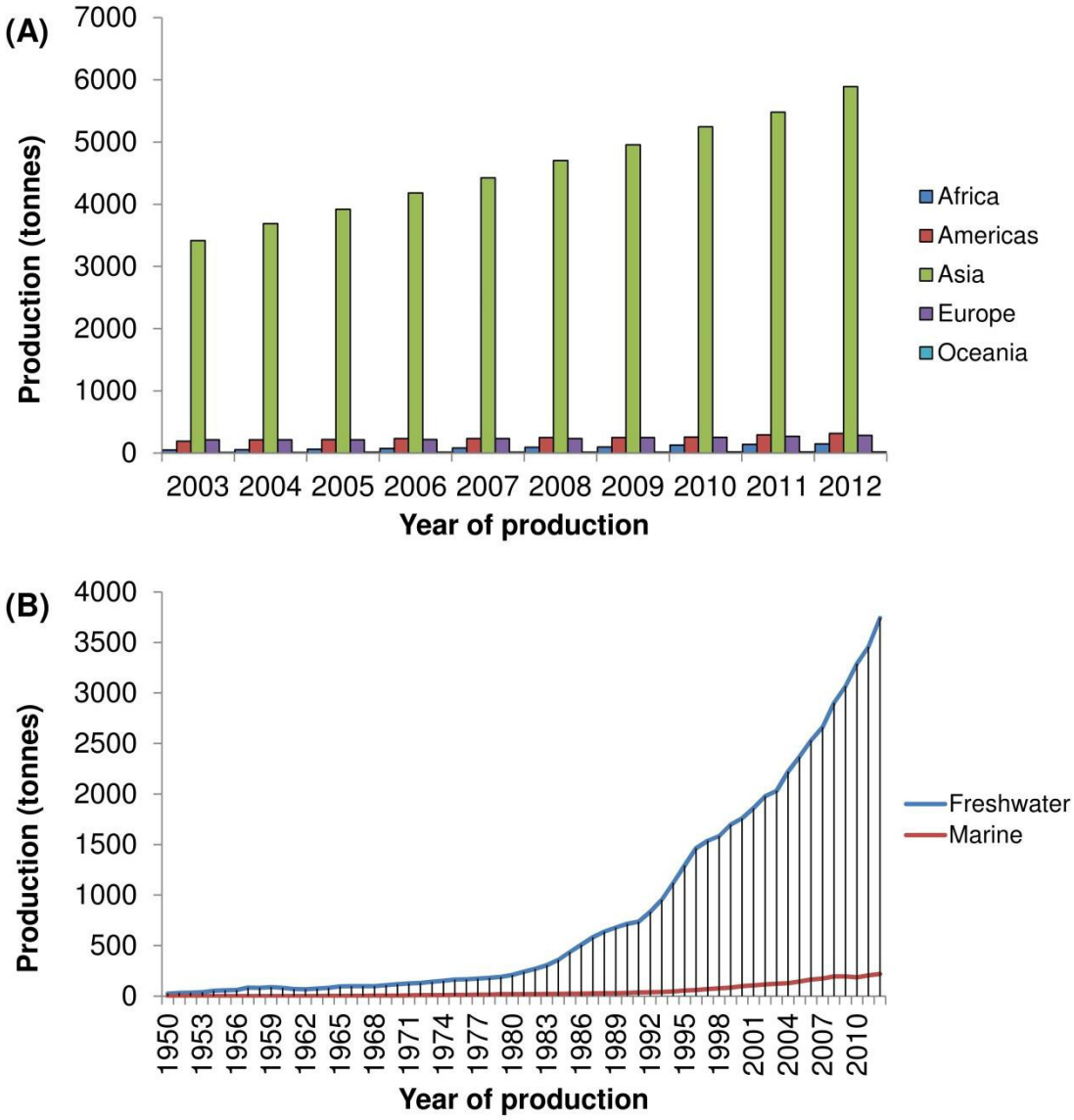


Fig. 1.1 Worldwide aquaculture production statistics.
The total quantity of all aquaculture production between 2003 and 2012 from different continents is shown **(A)**. Asia is the dominant aquaculture producer, increasing steadily from 2003-2012 as compared to other continents. Production of aquaculture in marine and freshwater environments between 1950 and 2012; freshwater production has now vastly outgrown marine aquaculture worldwide **(B)**. Statistics were gathered from (FAO, 2014).

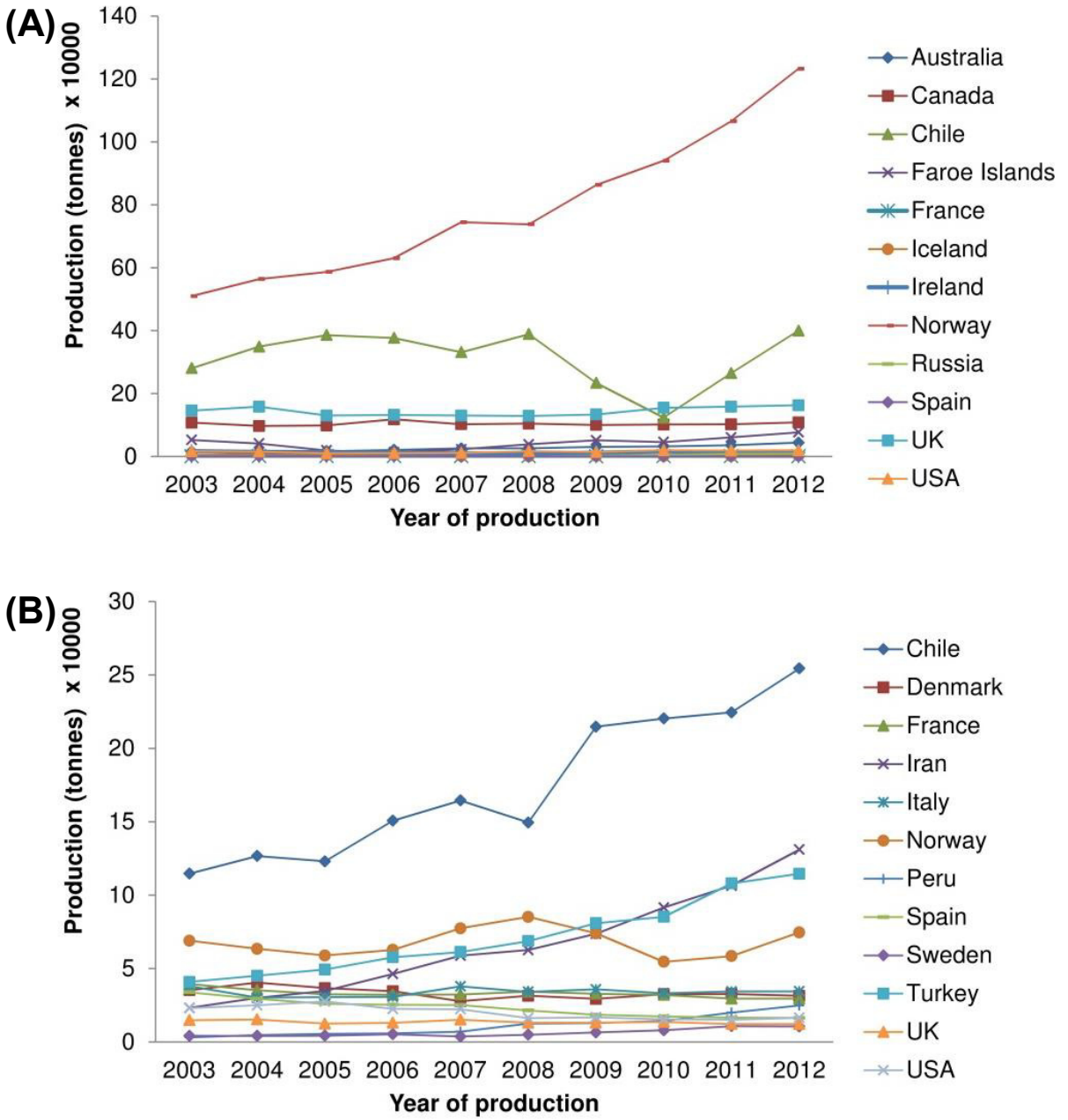


Fig. 1.2 Major producers of Atlantic salmon and rainbow trout worldwide.
The top 12 countries producing both Atlantic salmon **(A)** and rainbow trout **(B)** between 2003 and 2012 are listed. Norway has consistently been the top producer of Atlantic salmon, while Chile has produced the largest amount of rainbow trout. Statistics were gathered from (FAO, 2014).

1.1.2 **Scottish Aquaculture**

Aquaculture is a growing and increasingly important food production industry in the UK, particularly Scotland, helping to underpin sustainable economic growth in rural and coastal communities. The vast majority of aquaculture production in Scotland, by both weight and value, consists of the production of finfish in the sea; this being composed almost entirely of farmed Atlantic salmon. Scotland is currently the largest Atlantic salmon producer in the EU (163,234 tonnes in 2013

with an estimated value of £677 million), exporting to over 50 countries. Rainbow trout (5,611 tonnes), brown trout (44 tonnes) and halibut (56 tonnes) are also produced, but at much lower levels. The quantity (Fig. 1.3A and B) and value (Fig. 1.3C and D) of both Atlantic salmon and rainbow trout being produced worldwide has increased dramatically over the last decade. However, while the production of Atlantic salmon is increasing (Fig. 1.3E), rainbow trout production in the UK is decreasing (Fig. 1.3F) (FAO, 2014).

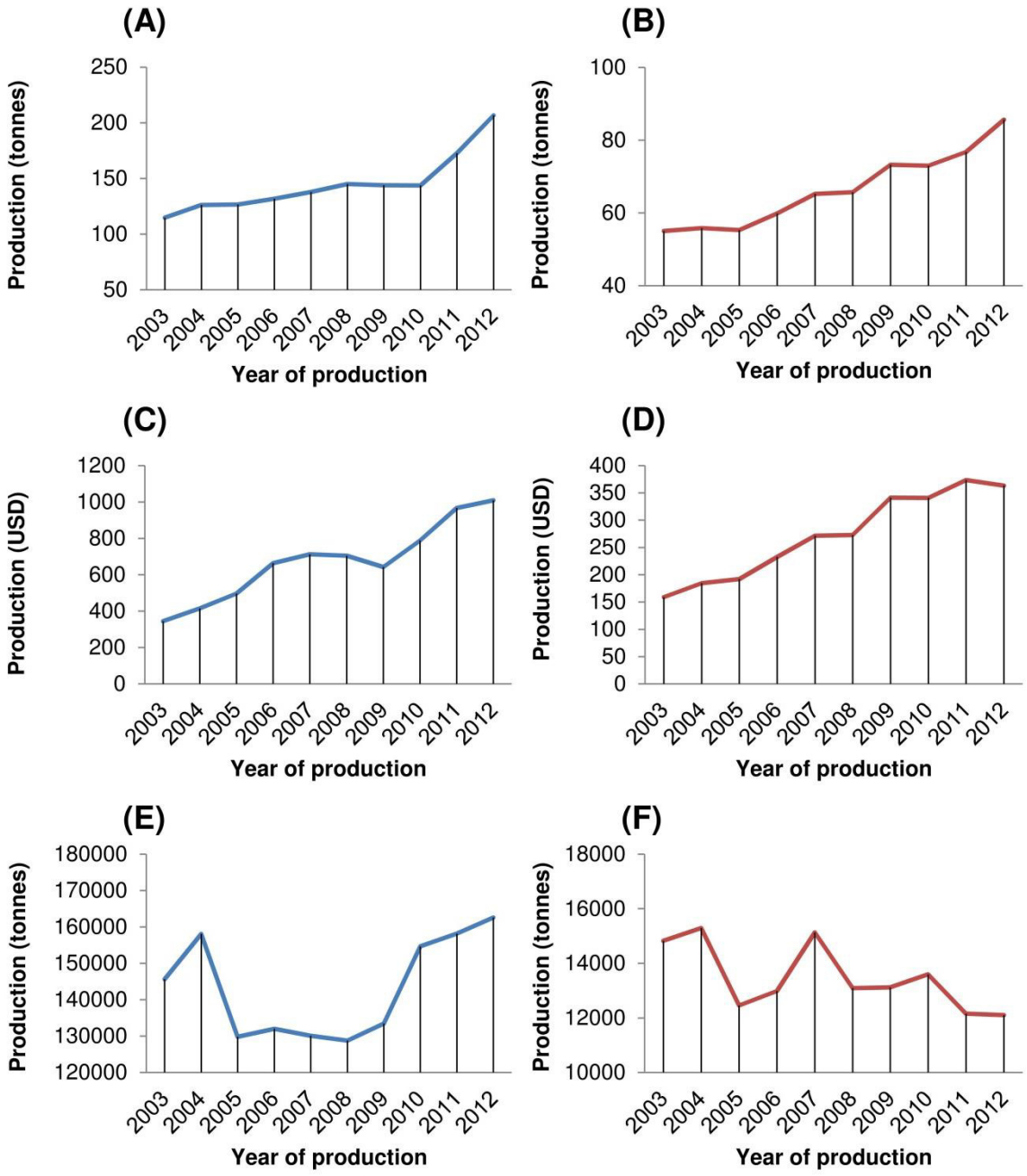


Fig. 1.3 Aquaculture production figures of Atlantic salmon and rainbow trout. Worldwide Atlantic salmon (blue lines) and rainbow trout production (red lines) figures in quantities (tonnes x 10,000; **[A & B]**) and value (USD x 10,000; **[C&D]**) are shown. UK production in tonnes (x 10,000) is also shown for both species **(E & F)**.

The marine farms where fish are grown are mainly situated on the West and North Coasts of the Scottish mainland and in the Western Isles, Orkney and Shetland (Fig. 1.4).

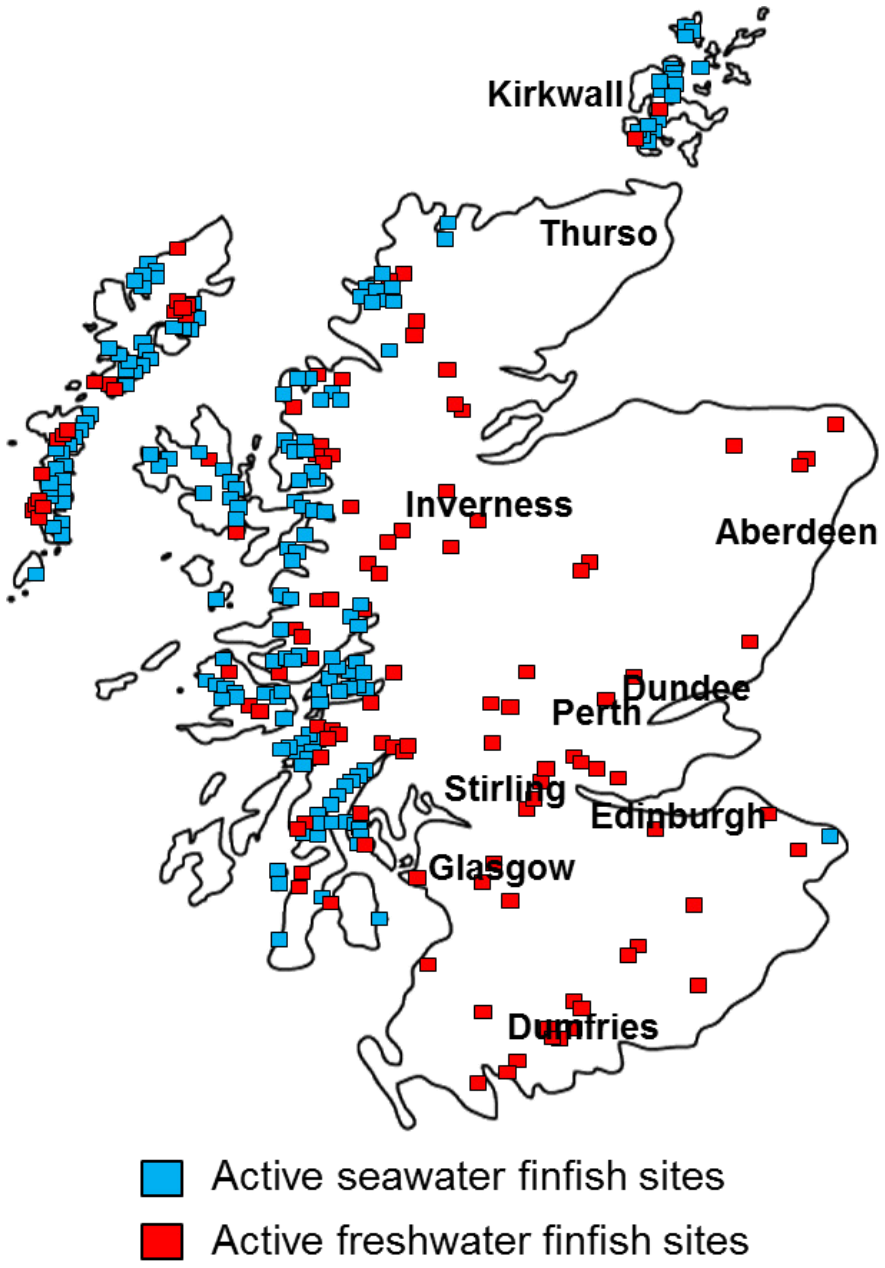


Fig. 1.4 Location of freshwater and marine finfish farms in Scotland.
Figure reproduced from www.aquaculture.scotland.gov.uk (2012).

Freshwater aquaculture is an important component of the aquaculture sector, both as a supplier of fish to marine aquaculture and as a producer in its own right. Although dominated by Atlantic salmon production, freshwater aquaculture in Scotland also includes farm production of brown trout, rainbow trout and Arctic char, for table and fishery purposes (FAO, 2014).

1.1.3 **The life cycle of farmed Atlantic salmon and rainbow trout**

The life cycle of farmed Atlantic salmon (Fig. 1.5) and rainbow trout begins with the stripping of eggs from the female broodstock, and the extraction of sperm ('milt') from the male broodstock. The broodstock are selected in order to achieve the highest quality product to produce a continuous supply of fast-growing, large, economically valuable fish. Once the eggs have been fertilised they are transferred to incubator trays and placed along a trough with a continuous supply of flowing water. After hatching, the trays are removed to allow the young fish to develop within the trough. At this stage they are called alevins, feeding for several weeks on the contents of their yolk sac. Once this has been consumed, the fish swim from the bottom of the tank and eat feed on the water surface. Now known as fingerlings, the fish at this stage are then transferred for on-growing in a variety of freshwater systems. Rainbow trout are usually reared in freshwater cages, earth ponds or raceways with a continuous flow of water, with many slaughtered as 'portion-size' trout between 280 and 450 g; although, some are left to grow on to greater weights of between 450 and 900 g. Others are slaughtered at over 900 g. It is not uncommon for some rainbow trout to be reared in freshwater initially, before transfer to seawater for on-growing, usually in cages. After only one year at sea these fish can reach weights of between 2.5 and 4 kg. Young Atlantic salmon, known as parr, are commonly raised in freshwater cages or tanks and raceways, growing rapidly through the winter before undergoing smoltification when the water temperatures and day length increase. This physiological change is a pre-adaptation to life in seawater. These fish are known as smolts. On fish farms, using artificial light and temperature manipulation means that parr can be made to smolt six months early, substantially reducing the length of growth time. Atlantic salmon smolts are then transferred to seawater sites in lorries, helicopters or well boats. A series of cages are connected to each other and to a floating metal walkway, with a net suspended from each cage containing the fish. Up to 75,000 salmon may be confined in the largest sea cages. Once transferred to seawater, Atlantic salmon begin to grow rapidly, increasing in size from only a few hundred grams to almost 6 kg in less than two years (Marine Harvest Handbook, 2012; Stevenson, 2007).

Fig. 1.5 The life cycle of farmed Atlantic salmon.
Initially, eggs are stripped from the female broodstock and mixed with milt from the male **(1)**. The fertilised eggs begin to show eyes after approximately 25 days **(2)** before hatching into alevins **(3)**. Over the next few months the alevins develop into fry (of between 60 and 100 g) and subsequently parr (as they develop a banding pattern along the flanks ['parr marks']; **4)**. As they become adapted for seawater as smolts they are transferred to artificial seawater tanks to test their readiness to effectively osmoregulate **(5)**. Once in seawater, the smolts develop into adult fish of between 4.5 and 6 kg over a period of almost two years **(6 & 7)**. At an optimal size, fish are transported to harvesting plants where they are processed for market. The life cycle of farmed rainbow trout is similar, although rainbow trout are not always transferred to seawater – instead continuing their growth in freshwater lochs.

On farms, fish are often crowded prior to grading, counting, transport and slaughter. Crowding the fish enables easier access, although of course leads to abnormally high stocking densities which is ultimately stressful and can lead to the fish becoming damaged. Many farm activities including vaccination and loading prior to transport and unloading involve handling the fish – this is also stressful for the fish. Transportation also poses a significant risk of spreading disease, and using the same machinery for different host species provides a possible reservoir for bacterial pathogens to be transmitted across farms and species.

1.1.4 **Diseases of aquaculture**

Annually, the aquaculture industry loses billions of dollars as pathogens including bacteria, fungi, parasites and viruses ravage stocks and plague farms (Verschuere *et al.*, 2000). Relatively little is known about how bacterial pathogens cause disease within aquaculture, as different strains, serotypes, genotypes and biotypes of strains vary in their ability to cause disease (Austin *et al.*, 2003; Esteve *et al.*, 2004, 2007). However, infectious diseases can be detrimental when developing and sustaining an aquaculture facility; implemented trade restrictions and poor quality fish can be damaging to the continuing operation of a farm (Verschuere *et al.*, 2000). Interactions between host, pathogen and the environment are paramount in determining the impact of disease outbreaks (Snieszko, 1974).

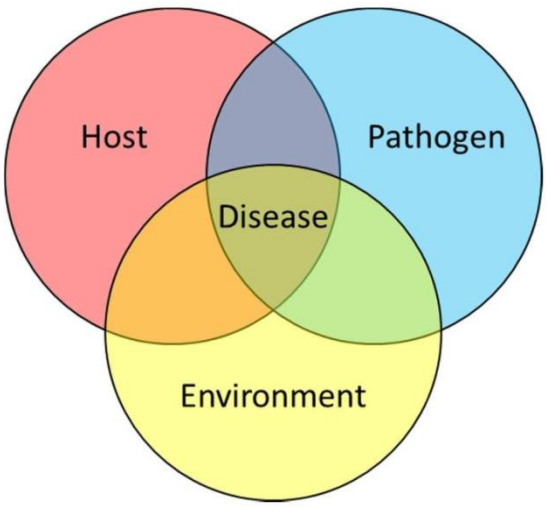


Fig. 1.6 The interaction between host, pathogen and environment results in disease. This Venn diagram was adapted from Snieszko (1974).

It is well documented that poor husbandry, stocking densities and other environmental stressors can encourage the spread of infection. Diseases may be spread more easily when populations are stocked densely due to increased interactions between infected and uninfected fish (Owens, 2012).

Bacteria are well characterised for their ability to survive within aquatic environments, particularly within biofilms (Coquet *et al.*, 2002). Many Gram-positive and Gram-negative species having been recovered from fish farms worldwide; the major bacterial species reported are summarised in Table 1.1. Bacterial fish diseases are caused by opportunistic pathogens (Austin & Austin, 2007). These occur naturally in the environment or inside the host, only causing disease once the host immune system is compromised by a stressful event or other disease process. However, some fish pathogens including *Vibrio anguillarum* and *Aeromonas salmonicida* are able to elicit an infection without the host being immunocompromised (Roberts, 2012). The annual economic loss associated with three major bacterial pathogens (*Aeromonas hydrophila, Yersinia ruckeri* and *Vibrio fluvialis*) was estimated to be more than $120 million in China alone, between 1990 and 1992 (Qi, 2002). However, the growing dependence on aquaculture in recent years has led to a corresponding increase in the incidence of bacterial infections. Toranzo, (2004) gathered disease information from 54 laboratories over a three year period from 1997 until 2000; Enteric redmouth (ERM) disease was highlighted as the third most common bacterial disease in fish, after vibriosis and pasteurellosis, with ERM occurring most commonly in trout. Bravo & Kojagura, (2004) identified that the number of hatcheries reporting *Y. ruckeri* infections in Peru increased from nine in 1998 to 22 in 2000. Recently, the emergence of non-motile variant isolates of *Y. ruckeri* in rainbow trout that had previously been vaccinated against the disease has presented a problem to the aquaculture industry. Outbreaks of this new phenotype were associated with poor water quality and husbandry practises (Austin *et al.*, 2003).

Table 1.1 Principal bacterial pathogens of fish.
This table was adapted from (Austin & Austin, 2007; Roberts, 2012).

Bacterium	Disease	Host affected	Geographical distribution
Gram-negative gliding bacteria			
Flavobacterium branchiophilum	Bacterial gill disease: fin rot	Salmonids	Europe, Korea, Japan, USA
Flavobacterium columnare	Columnaris disease	Many freshwater fish	Worldwide
Flavobacterium psychrophilum	Cold water disease: Rainbow trout fry syndrome (RTFS)	Salmonids	Australia, Europe, Japan, N. America
Tenacibaculum maritimum	Bacterial gill disease: fin rot	Many marine fish	Europe, Japan, USA
Sporocytophaga sp.	Seawater columnaris	Salmonids	Scotland, USA
Gram-negative facultatively anaerobic rods			
Edwardsiella tarda	Edwardsiella septicaemia	Many freshwater fish	Japan, Spain, USA
Edwardsiella ictaluri	Enteric septicaemia of catfish	Many species	Indonesia, USA, Vietnam
Yersinia ruckeri	**Enteric redmouth disease**	**Salmonids**	**Worldwide**
Vibrio anguillarum	Vibriosis	Most marine fish species	Worldwide
Vibrio ordalii	Vibriosis	Most marine fish species	Worldwide
Vibrio salmonicida	Hitra disesae: cold water vibriosis	Atlantic salmon	Norway, Canada, Scotland
Vibrio vulnificus	Septicaemia	Eel	Europe, Japan, China, USA
Aeromonas salmonicida	Furunculosis	Salmonids, marine species	Worldwide
atypical *Aeromonas salmonicida*	-	Ballan wrasse cleaner fish	-
Aeromonas hydrophila	Septicaemia	Many freshwater fish	Worldwide
Aeromonas sobria	Septicaemia	Perch, gizzard shad	Switzerland, USA
Moritella viscosa	Winter ulcer disease	Atlantic salmon	Iceland, Norway, Scotland
Photobacterium damselae subsp. *piscicida*	Pasteurellosis	Many species	Europe, Japan, USA
Photobacterium damselae subsp. *Damselae*	Haemorrhagic septicaemia	Many species	Europe, USA

Table 1.1 continued.

Bacterium	Disease	Host affected	Geographical distribution
Gram-negative aerobic rods			
Pseudomonas fluorescens	Septicaemia	Many species	Worldwide
Pseudomonas anguilliseptica	*Sekiten-byo*, red spot	Many species	Europe, Japan
Gram-negative obligate bacteria			
Francisellaceae spp.	Granulomatosis	Cod, Bass, Atlantic salmon, Talipia, Three-line grunt, Abalone and Cichlids	Japan, Norway, North America, South America and Central America
Piscirickettsia salmonis	Rickettsiosis	Salmon, Sea Bass	Worldwide
Gram-positive aerobic rods			
Renibacterium salmoninarum	Bacterial kidney disease (BKD)	Salmonids	Europe, Japan, N. and S. America
Carnobacterium piscicola	Similar to BKD	Salmonids	N. America, UK
Lactococcus garviae	Septicaemia	Many species	Worldwide
Streptococcus iniae	Septicaemia	Various freshwater species	Worldwide
Streptococcus faecalis subsp. *liquefaciens*	Septicaemia	Rainbow trout, Catifsh	Italy
Streptococcus agalactiae	Septicaemia	Carp, rainbow trout, Tilapia	Israel, Kuwait, USA
Streptococcus dysgalactiae	Septicaemia	Amberjack, Yellowtail	Japan
Gram-positive anaerobic rods			
Clostridium botulinum	Type E botulism	Redfish/Salmonids	Denmark, England, USA
Acid fast rods and filaments			
Mycobacterium marinum	Mycobacteriosis	Many species	Worldwide
Mycobacterium fortuitum	Mycobacteriosis	Many species	Worldwide
Mycobacterium chelonae	Mycobacteriosis	Many species	Worldwide
Nocardia asteroides	Nocardiosis	Many species	Worldwide
Nocardia seriolae	Nocardiosis	Many species	Worldwide

1.2 *Yersinia ruckeri*

Yersinia ruckeri is a Gram-negative enteric bacterium and the causative agent of ERM. The genus *Yersinia* is composed of 19 different species members, of which three (*Yersinia pestis, Yersinia pseudotuberculosis* and *Yersinia enterocolitica*) are known to be human pathogens and have been extensively characterised. A summary of the genus members and their disease characteristics is shown in Table 1.2. *Yersinia ruckeri* was first isolated from farmed rainbow trout (*Oncorhynchus* mykiss, Walbaum) in the Hagerman valley, Idaho, in the early 1950's (Ross & Klontz, 1965; Rucker, 1966).

Subsequently, the organism was isolated in many states across the USA and Canada (Bullock *et al.*, 1978; Busch, 1978; Ross *et al.*, 1966; Stevenson & Daly, 1982).

1.2.1 **Taxonomic classification**

Yersinia ruckeri is the oldest of the non-human pathogenic *Yersinia* species, being designated a member in 1978 (Ewing *et al.*, 1978). However, its taxonomic classification has been in question since with many believing that an entirely new genus be created in which to house it (Farmer *et al.*, 1985; Kotetishvili *et al.*, 2005; Méndez *et al.*, 2009). A multi locus sequence typing (MLST) study conducted to determine the relatedness amongst *Yersinia* species showed that based on the 16S rRNA gene (Fig. 1.7A) and the concatenation of four housekeeping enzymes (*glnA*, *gyrB*, *recA* and *Y-HSP60*; Fig. 1.7B), *Y. ruckeri* is "genetically the most distant species within the genus" (Kotetishvili *et al.*, 2005). The bacterium was housed amongst the Enterobacteriaceae owing to its antigenic similarity to *Salmonella arizonae* (Green & Austin, 1983), but the differences in biochemical characteristics between the two led many to question the placement. The guanine:cytosine (G:C) ratio determined *Y. ruckeri* as an Enterobacteriaceae family member with a ratio between 47.5 and 48.5% G:C (De Grandis *et al.*, 1988); this is different from *Serratia* species (52-60%) and closer to other members of the *Yersinia* (46-50%). The name '*Yersinia ruckeri*' was designated by Ewing *et al.* (1978) in memorandum of the original isolator of the bacterium, Robert R. Rucker.

Table 1.2 Properties of the nineteen described members of the *Yersinia* genus.
The original source of the *Yersinia* species, the disease it is known to cause, the susceptible species and the original reference are indicated.

Species	Subspecies	Original source	Disease caused	Affects	References
Yersinia pestis	n/a	Human, Hong Kong	Bubonic plague	Humans	(Yersin 1894) Van Loghem *et al.,* 1943
Yersinia philomiragia	n/a	Muskrat	Not known	Not known	Jensen *et al.,* 1969
Yersinia ruckeri	**n/a**	**Rainbow trout**	**Enteric redmouth disease**	**Fish**	**Ewing *et al.,* 1978**
Yersinia frederiksenii	n/a	Sewage sample	Gastrointestinal infections	Humans and fish	Ursing *et al.,* 1981
Yersinia intermedia	n/a	Human faeces	Not clinically relevant	Humans and fish	Brenner *et al.,* 1981
Yersinia kristensenii	n/a	Sheep	Bacteriocin secretion	Mice	Bercovier *et al.,* 1981
Yersinia pseudotuberculosis	*pestis*	Turkeys	Gastrointestinal infections	Humans	Bercovier *et al.,* 1981
	pseudotuberculosis				Bercovier *et al.,* 1981
Yersinia aldovae	n/a	Drinking water	Not known	Not known	Bercovier *et al.,* 1984
Yersinia rohdei	n/a	Dog faeces	Not known	Not known	Aleksic *et al.,* 1987
Yersinia bercovieri	n/a	Human faeces	Not known	Humans	Wauters *et al.,* 1988
Yersinia mollaretii	n/a	Soil	Not known	Not known	Wauters *et al.,* 1988
Yersinia enterocolitica	*enterocolitica*	Human tissue	Yersiniosis	Humans, cattle, deer, pigs, fish and birds	Neubauer *et al.,* 2000
	palearctica				Neubauer *et al.,* 2000
Yersinia aleksiciae	n/a	n/a	Not known	Not known	Sprague and Neubauer 2005
Yersinia similis	n/a	Rabbit	Not known	Not known	Sprague *et al.,* 2008
Yersinia massiliensis	n/a	Fresh water (France)	Not known	Not known	Merhej *et al.,* 2008
Yersinia entomophaga	n/a	Diseased grass grub	Not known	Not known	Hurst *et al.,* 2011
Yersinia nurmii	n/a	Broiler meat	Not known	Not known	Murros-Kontiainen *et al.,* 2011
Yersinia pekkanenii	n/a	Water, soil and lettuce	Not known	Not known	Murros-Kontiainen *et al.,* 2011
Yersinia Wautersii	n/a	Otter	Not known	Not known	Savin *et al.,* 2014

In 2009, the fully sequenced genome of *Y. ruckeri* strain ATCC (American type Culture Collection) 29473 was published. This strain represents the 'Hagerman' serotype O1 reference type isolate, and has a genome size of 3.7 Mbp (NCBI [National Center for Biotechnology Information], naval medical research centre, 2009). This strain was sequenced to contig level, and remained the reference point for several years. However, the continued reduction in whole genome sequencing (WGS) costs has led to the release of several genomes more recently, including an update to the original genome which is now at scaffold assembly level. Genomes from strains recovered from Atlantic salmon (Navas & Bohle, 2014), rainbow trout (Nelson *et al.*, 2015), and a channel catfish (Wang *et al.*, 2015) have been released, with completely closed rainbow trout (Johnson *et al.*, 2015) and Atlantic salmon (Stevenson *et al.*, 2015) genome sequences now publicly available in 2015.

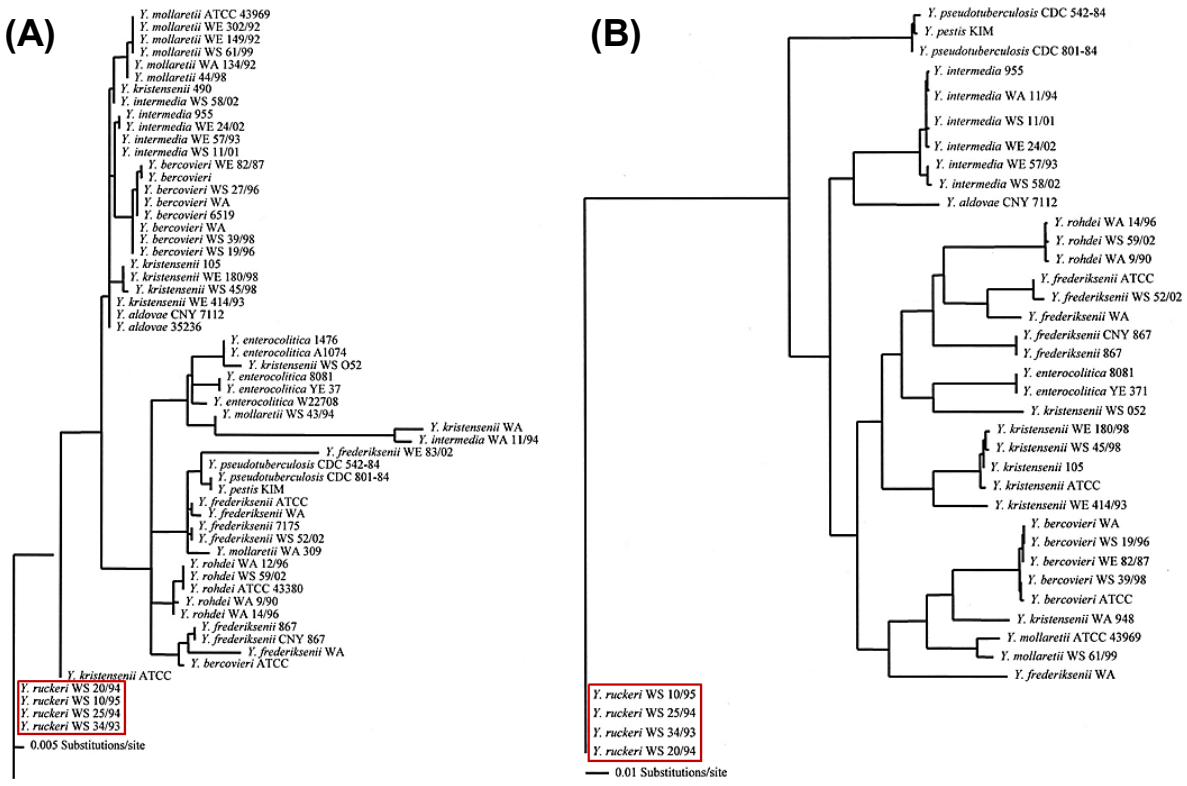


Fig. 1.7 Genetic relatedness amongst *Yersinia* species.
A phylogenetic relatedness tree of 16S rRNA **(A)** and four housekeeping enzymes (*glnA*, *gyrB*, *recA* and *Y-HSP60;* **B)** amongst members of the genus *Yersinia*. The position of *Y. ruckeri* is highlighted on each phylogenetic tree with a red box, and in both cases is the most distant member of the genus. This figure was adapted from (Kotetishvili *et al.*, 2005).

1.2.2 **Morphological and biochemical characteristics**

1.2.2.1 Cell and cultural morphology

Yersinia ruckeri is a non-spore forming, un-encapsulated variably motile bacterium (Ross *et al.*, 1966). The cells can be divided into two main groups; very short coccobacillary (0.5 µm x 0.5-10 µm), or medium sized rods (0.5 µm x 1.0-2.0 µm). On occasion, slightly longer rods (3.0 µm) with a higher frequency of chain formation are recovered (Davies & Frerichs, 1989). Colonies of *Y. ruckeri* are smooth and off white/cream in colour, and approximately 3 mm in diameter after 48 h growth at 22°C on Tryptone Soya agar (TSA; Oxoid, UK).

The majority of characterisation studies of the bacterium have taken place at 22°C (Davies & Frerichs, 1989; Davies, 1991a, b, c; Verner-Jeffreys *et al.*, 2011; Wheeler *et al.*, 2009), however the bacteria will grow at up to 37°C (Carson & Wilson, 2009). Several studies have also used 20°C (Secades & Guijarro, 1999), 25°C (Coquet *et al.*, 2002, 2005) and 28°C (Fernández *et al.*, 2007a) as the optimal growth temperature.

1.2.2.2 Motility

Discrete differences in the numbers of flagella present around cells at 22°C have been shown (Austin & Rodgers, 1982) (Fig. 1.8). Longer cells appeared to have approximately 12 peritrichous flagella, whilst shorter rods had only 3 or 4 at this temperature. Tryptone Soya agar supplemented with 3% (w/v) NaCl resulted in a marked difference in cell size, from 0.5 – 4 µm.

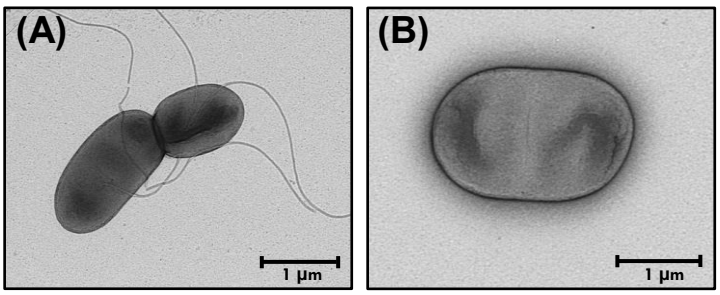


Fig. 1.8 Motile and non-motile strains of *Y. ruckeri.*
Isolates used include RD370 (motile) **(A)** and RD360 (non-motile) **(B)**. Several pertrichous flagella are visible around the motile bacterial cells in **(A)**. Images were acquired at x5000 magnification using Transmission electron microscopy (TEM).

Examination of the arrangement of flagellar genes within motile and non-motile isolates demonstrated that the emergence of the non-motile phenotype is associated with mutational loss of flagellin secretion, eliminating the production of flagella, and confirmed that loss of motility is associated with loss of secreted lipase activity. However, challenge studies have indicated that neither of these activities are necessary for virulence in rainbow trout (Evenhuis *et al.*, 2009). The loss of motility associated with the non-motile phenotype is caused by unique mutations in the flagella biosynthesis/export genes *fliR*, *flhA* and *flhB*. British and North American non-motile isolates contain the same single nucleotide deletion in *fli*R (*fliRΔ1*), suggesting that these isolates originate from a common ancestor. However, non-motile isolates recovered from Denmark and Finland have a 10 bp deletion 350 bp downstream of *fliRΔ1*, called *fliRΔ2*. Non-motile isolates from Finland have a nonsynonymous substitution within *flh*A, and non-motile isolates recovered from Spain contain a 10 bp deletion in *flh*B (Fig. 1.9) (Welch *et al.*, 2011).

Significantly, non-motile isolates recovered from fish previously vaccinated with a motile form of the disease have become increasingly problematic from a disease perspective (Arias *et al.*, 2007; Austin *et al.*, 2003; Bastardo *et al.*, 2011; Welch *et al.*, 2011; Wheeler *et al.*, 2009). Furthermore, a correlation between vaccine failure and the emergence of non-motile isolates suggests that flagella production may be a liability itself within a vaccinated host (Evenhuis *et al.*, 2009; Scott *et al.*, 2013).

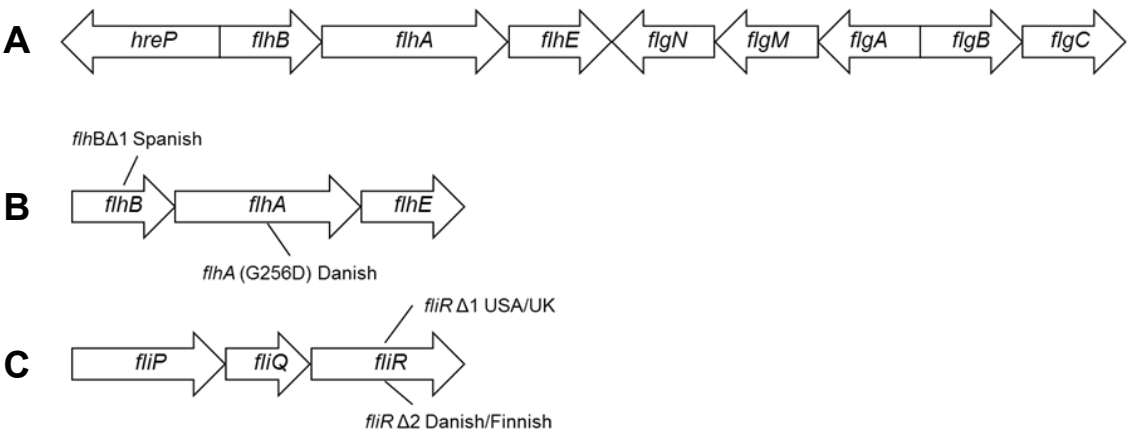


Fig. 1.9 Arrangement of flagellar genes and mutation sites in *Y. ruckeri*.
The arrangement of the flagellar genes in *Y. ruckeri* **(A)**. This figure was adapted from (Evenhuis *et al.*, 2009). The sites of mutation in *flhA* and *flhB* of Danish and Spanish strains respectively **(B)**. The sites of mutation in *fliR* of both the USA/UK and Danish/Finnish strains **(C)**. **(B)** & **(C)** were adapted from (Welch *et al.*, 2011)

1.2.2.3 Biochemical characterisation

1.2.2.3.1 Biotyping

A detailed analysis of the biochemical characteristics of 147 global *Y. ruckeri* isolates from rainbow trout, Atlantic salmon and other species revealed the presence of non-motile isolates in the UK (Davies & Frerichs, 1989). These isolates also lacked the ability to hydrolyse Tween 20 and Tween 80. Conversely, motile isolates were able to hydrolyse both Tween 20 and Tween 80. On this basis, the motile, lipase positive isolates were designated biotype 1 and the non-motile, lipase negative isolates were designated biotype 2 (Davies & Frerichs, 1989).

Since the identification of biotype 2 isolates in the UK, non-motile variants have become more widespread and have been described in the USA (Arias *et al.*, 2007), Chile (Bastardo *et al.*, 2011), Australia (Carson & Wilson, 2009), Spain (Fouz *et al.*, 2006) and other European countries (Wheeler *et al.*, 2009).

1.2.2.3.2 Serotyping/LPS analysis

The confusion surrounding the serotyping schemes of *Y. ruckeri* has been a continuous issue, with many different authors attempting to clear the waters. With the emergence of new serotypes since the initial typing scheme in 1977 (O'Leary, 1977), attempts using whole cell agglutination reactions led to the identification of seven different serovars (Bullock *et al.*, 1978; Daly *et al.*, 1986; De Grandis *et al.*, 1988; Pyle & Schill, 1985; Pyle *et al.*, 1987; Stevenson & Airdrie, 1984; Stevenson & Daly, 1982), one of which was subsequently removed when no longer recognised as *Y. ruckeri* (De Grandis *et al.*, 1988). Davies (1990) subsequently used heat stable O-antigens to identify 5 different O-serotypes. This method was based on the Kauffmann-White method of typing in *Escherichia coli*, and was more appropriate than the previous methods of differentiation, which used whole cells. Utilising heat stable O-antigens removes cross reactive proteins and is still the preferred choice, commonly used today. These serotypes were named O1 (previously serovar I), O2 (previously serovar II), O5 (previously serovar V), O6 (previously serovar VI) and O7 (this was not tested in the Stevenson and Airdrie [1984], or Daly *et al.*, [1986] studies; however, it represented serovars 4 and 3 in Pyle and Schill [1985] and Pyle *et al.*, [1987],

respectively), to aid with the transition from previous typing schemes. Serovar III was not recognised as an O-serotype in the Davies (1990) study. The original serotype III isolate described by Bullock *et al.*, (1978) has a rough type LPS structure and was therefore thought to be a rough type mutant of *Y. ruckeri*. The isolates which had been identified as serotype III had identical LPS profiles to those of other serotype O1 isolates. Combining examination of extracellular products (ECPs) with O-antigen serotyping, Romalde *et al.*, (1993) created a typing scheme consisting of four distinguished O-serovars. Serovar O1 is subdivided into two groups; O1a (previously serovar I) and O1b (previously serovar III). Serovar O2 (previously serovar II) is subdivided into three subgroups (O2a, O2b and O2c), while serovar O3 (previously serovar V) and O4 (previously serovar VI) complete the scheme (Romalde *et al.*, 1993). However, utilising ECPs to further supplement a serological differentiation technique that is based on heat stable O-antigens has added to the confusion surrounding these typing schemes. The serotyping schemes are summarised below (Table 1.3).

Table 1.3 Summary of the serotyping schemes of *Y. ruckeri.*
Between 1977 and 1993 the serotyping classification of *Y. ruckeri* has been altered 10 times. Although Romalde *et al*., 1993 is the most recent adaptation, the scheme instigated by Davies (1990) remains the most commonly used.

Year	1977	1978	1982	1984	1984	1985	1986	1987	1990	1993
Author	**P. J. O'Leary**	**Bullock, Stuckey & Shotts**	**Stevenson & Daly**	**Stevenson & Airdrie**	**De Grandis**	**Pyle & Schill**	**Daly, Lindvik & Stevenson**	**Pyle *et al.***	**Davies**	**Romalde *et al.***
Based on	**Whole cell**								**O-antigen**	**O-antigen and OMP**
Designations	I	I	I	I (Hagerman)		nt	I	1	O1	O1a
	II	II	II	II (Big creek)		2	II	2	O2	O2a, b, c
		III	III	III		nt	III	6	O1	O1b
				IV	Removed IV	-	-	-	-	-
				V		6	V	5	O5	O3
				VI		5	VI	4	O6	O4
						4	nt	3	O7	nt

1.2.2.3.3 Outer membrane protein (OMP) typing

Following the same patterns of outer membrane (OM) profiling that had been described for many other pathogenic bacteria, a typing scheme based on molecular mass variation of the major OMPs of *Y. ruckeri* (i.e., OmpA, OmpC and OmpF) was developed. This allowed for the identification of nine different OMP-types (Fig. 1.10) (Davies, 1989, 1991c).

Fig. 1.10 Outer membrane protein profile typing scheme of *Y. ruckeri*.
The OMP-typing scheme of *Y. ruckeri* is based on variations in molecular mass of the major OMPs OmpA, OmpC and OmpF (labelled A, C and F respectively). Nine different OMP-types have been described. Estimated molecular weights of proteins in each OMP-type are indicated beside each band. Adapted from (Davies, 1989, 1991c).

The scheme relies on the known heat modifiable properties of OmpA, which undergoes a characteristic shift in molecular mass when heated to higher solubilisation temperatures in the presence of Sodium dodecyl sulphate (SDS) prior to gel electrophoresis, in combination with variation in expression of the major porin proteins OmpF and OmpC when isolates are grown under anaerobic conditions; OmpC is substantially upregulated under anaerobic growth compared to standard conditions, while OmpF is down regulated. Davies identified the major OMP-types 1, 2, 3, 4 and 5, which were further subtyped based on subtle molecular mass differences of these three major proteins. OMP-type 1 was subdivided into types 1a and 1b; OMP-type 2 into subtypes 2a, 2b and 2c; OMP-type 3 into subtypes 3a and 3b; OMP-types 4 and 5 were not subtyped as further variations were not identified (Davies, 1989, 1991c).

1.2.2.4 Genetic characterisation

1.2.2.4.1 Multilocus enzyme electrophoresis

Multilocus enzyme electrophoresis (MLEE) has been used previously in population genetics studies of plants, animals, fungi, and bacteria. Allelic variants of

enzymes are detected by shifts in electrophoretic mobility as a consequence of one or more amino acid substitutions and the resultant change in net charge. Schill *et al*. (1984) conducted a study comprising 47 *Y. ruckeri* isolates recovered from the USA, Canada and Australia, examining the activity and mobility of 21 enzymes. Identical isoenzyme patterns were observed at 15 enzyme loci with only five exceptions; four of the electrophoretic types differed by less than 2 of 15 loci. This led to the designation of four clones, with clone I the most common electrophoretic type (42 of 46 isolates). This study highlighted the distinct lack of genetic diversity within this species (Schill *et al.*, 1984).

1.2.2.4.2 Pulsed field gel electrophoresis (PFGE)

Wheeler *et al*. (2009) reported the development of a PFGE method for the typing of *Y. ruckeri*. The study included 160 isolates recovered from around the world and resulted in the identification of 44 different pulsotypes (pt). Isolates collected from rainbow trout and Atlantic salmon were examined in greater detail. Within rainbow trout, isolates from the same clonal group had identical (or similar) pulsotypes, while UK, European and US virulent serotype O1 isolates also had similar pulsotypes (pt31–pt39). The European isolates (pt34–pt38) were distinctly different from UK isolates (pt31 [18 isolates representing the Hagerman ATCC 29473 strain] and pt32 [24 isolates representing the non-motile variant strain). Isolates of pt31 and pt32 were almost identical, with only an additional band of 240 kb in pt32 isolates. In isolates recovered from Atlantic salmon, a much more diverse range of pts were recovered. Of 34 strains examined (mainly of serotypes O1, O2 and O5), several clusters became apparent. The O5 isolates clustered together and showed greater homology to each other than to the O1 or O2 isolates, while the major O1 cluster was separated into two (1a [motile, pt17, pt18 or pt19] and 1b [non-motile, pt28, pt29 or pt32]). The non-motile pulsotypes pt28 and pt29 shared high homology (94%) with pt31 and pt32 of rainbow trout (Wheeler *et al.*, 2009).

UK and European isolates were more homologous to US isolates than each other, confirming previous suggestions that *Y. ruckeri* was introduced separately from the US and the transfer of strains from the UK to Europe (and vice versa) has been kept to a minimum (Wheeler *et al.*, 2009).

1.2.2.4.3 Multilocus sequence typing

Multilocus sequence typing has been described as the gold standard method of genetic typing, with advantages over MLEE and PFGE (Maiden, 2006). The aim of MLST is to provide an accurate and highly discriminating typing system, based on nucleotide sequences that can be used for most bacteria and some other organisms. *Y. ruckeri* has been included in two MLST schemes to date. Kotetishvilli *et al.* (2005) examined variation within the *Yersinia* genus, including four isolates of *Y. ruckeri*. Using both 16S rRNA, and four concatenated housekeeping enzymes (*glnA, gyrB, recA* and *Y-HSP60*), *Y. ruckeri* was shown to be genetically different from all other members of the *Yersinia*. More recently, Bastardo *et al.,* (2012a) produced a comprehensive study comprising 103 *Y. ruckeri* strains recovered from diverse geographic locations and host/environmental origins. This study utilised internal fragments of six housekeeping enzymes (*glnA, gyrB, recA, Y-HSP60, dnaJ* and *thrA*). Thirty different sequence types (STs) were identified, of which 21 were represented by single isolates. Sequence type 2 was representative of nearly a third of all isolates, and was composed almost entirely of isolates recovered from trout. The relatively few Atlantic salmon isolates in the study were represented by a more diverse range of STs. This study also supported the previous work by Wheeler *et al.*, (2009), suggesting that isolates of biotype 2 recovered from Europe and the UK were more closely related to each other than to those from the USA (Bastardo *et al.*, 2012a).

1.2.3 Virulence factors of *Y. ruckeri*

1.2.3.1 Extracellular products (ECPs)

Extracellular products play crucial roles in virulence, and have been well studied in several fish pathogenic bacteria, including *A. hydrophila* (Allan & Stevenson, 1981), *F. psychrophilum* (Ostland *et al.*, 2000) and members of the *Vibrionaceae* (Kodama *et al.*, 1984). Within *Y. ruckeri*, Romalde & Toranzo, (1993) demonstrated that injecting ECPs into rainbow trout elicited the appearance of symptoms related to yersiniosis. Several ECPs have been described in detail in *Y. ruckeri*, including the protease Yrp1 and the hemolysin YhlA.

1.2.3.2 Yrp1

In fish-pathogenic bacteria, indirect experiments have indicated the participation of proteolytic enzymes in pathogenesis; a theory supported by genetic approaches. The truncation of the encoding proteolytic genes has shown that proteolytic enzymes are involved in virulence. In *Y. ruckeri*, extracellular proteolytic activity has been related to virulence; namely a 47 kDa protease (Secades & Guijarro, 1999) produced at the end of the exponential growth phase - Yrp1. The protease is secreted by a type I ATP binding cassette (ABC) exporter system composed of three genes - *yrpD, yrpE* and *yrpF* and a protease inhibitor *inh*. Fernandez *et al.*, 2003 cloned *yrp1*, and examined its role in virulence. The presence of this protease has no relationship with serotype; some strains of serotype O1, the most virulent and common, have the protease, while other strains of the same serotype do not. This led to the designation of Azo^+ and Azo^- strains, according to the presence or absence of Yrp1 proteolytic activity respectively (Secades & Guijarro, 1999). All strains have been shown to encode *yrp1*, although, the Azo^- phenotype is due to a transcriptionally inactive *yrp1* operon (Fernández *et al.*, 2003). The protease contributes to the virulence of the bacterium, and has roles in colonisation and invasion of different tissues. Yrp1 has been shown to digest a variety of extracellular matrix and muscle proteins, and may lead to alterations and pores in the capillary vessels. The typical haemorrhages around the mouth and intestine may be as a result of leaking blood caused by this damage (Fernández *et al.*, 2003). The expression of Yrp1 has also been related to environmental growth conditions, with production sensitive to repression by carbon and nitrogen sources. Glucose and fructose were shown to be the greatest production inhibitors, while glycerol, mannitol and maltose had a potent repressive effect. Ammonium in the culture medium also reduced the presence of Yrp1 (Secades & Guijarro, 1999). Expression was higher at 18°C, and reduced at 28°C – an apparent adaptation to life inside the host, while increasing osmotic pressure decreased expression also (Fernández *et al.*, 2003). Characterisation of the protease showed that it requires Mg^{2+} and Ca^{2+} cations for maximal activity (Secades & Guijarro, 1999).

1.2.3.3 YhlA

The hemolysin YhlA has been found to be an important factor in the pathogenesis of *Y. ruckeri*. It is thought that two genes are involved in the production of YhlA - *yhlA* and *yhlB*. Fernández *et al*. (2007a) observed that the upstream gene *yhlB* is involved in the secretion and activation of the haemolysin, which is encoded by *yhlA*. It was found through genomic analysis that there are similar haemolysins in human pathogenic *Yersinia*.

YhlA was able to lyse erythrocytes in cultured fish cells, showing similarities in invasive properties with *Serratia*-type toxins. YhlA expression was higher at 18°C, the host temperature, rather than the optimal bacterial growth temperature. The production of YhlA also increased when grown under iron-limited conditions, suggesting a possible role in the acquisition of iron from the host (Fernández *et al.*, 2007a; Tobback, 2009).

1.2.3.4 Adhesins and invasins

Adhesion and invasion of pathogenic bacteria is an important initial step for the infection and subsequent colonisation. Bacterial pathogens use surface associated adhesins and invasins to create a specific interaction with host cell receptors (Niemann *et al.*, 2004).

1.2.3.4.1 Invasin, Ail, and YadA

Two of the three human pathogenic *Yersinia* species (*Y. pseudotuberculosis* and *Y. enterolitica*) are known to produce at least three proteins involved in adherence and invasion (Invasin, Ail and YadA). Invasin and Ail are encoded on the chromosome, whilst the YadA protein is encoded on a 70 kb virulence plasmid. Invasin and YadA are OM adhesins involved in the binding of ß1 integrin receptors, an important attachment in promoting internalisation into macrophages. Ail (OmpX in *E. coli*) is an OMP with an important involvement in cell invasion, through an unknown binding receptor. Kawula *et al*. (1996) examined the presence of *inv* or *ail* homologues in *Y. ruckeri*, finding no evidence of them (Fig. 1.11). However, using polymerase chain reaction (PCR) and sequencing, a possible *inv* homologue was identified in *Y. ruckeri* which could be involved in adhesion and invasion of the bacterium into the host cells

(Fernández *et al.*, 2007b). Evenhuis *et al.* (2009) however, observed that *inv*, encoded between *flhE* and *flgN* is present in the human pathogenic *Yersinia* species but is absent from *Y. ruckeri* (Evenhuis *et al.*, 2009).

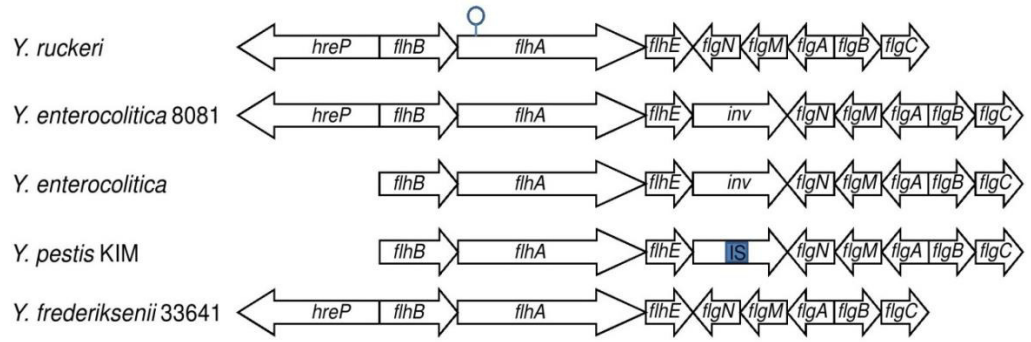


Fig. 1.11 Arrangement of the flagellar operon in different species of *Yersinia*.
The arrangement of the *flhBAE-flgNMABC* gene cluster in *Yersinia* strains. The *inv* gene is absent from non-human pathogenic species *Y. ruckeri* and *Y. frederiksenii*. A HreP region was absent from the human pathogenic species also. This figure was reproduced from (Evenhuis *et al.*, 2009).

1.2.3.5 Immune avoidance

Serum and phagocyte mediated killing represent the two major defence mechanisms of the innate immune response in fish, and resistance to these mechanisms by the pathogen contributes to infection (Blazer, 1991; Dalmo *et al.*, 1997; Ryckaert *et al.*, 2010).

The innate immune response is mediated largely by white blood cells (WBCs) including neutrophils and macrophages. These are cells involved in phagocytosis and killing of foreign infective cells, and elicit additional host responses by synthesising a range of inflammatory mediators and cytokines. In the macrophage, the bacteria is killed and broken down within a maturing phagosome. Components of the pathogen are presented to T-cells, bringing about the activation of the adaptive immune response. Ryckaert *et al.* (2010) examined the ability of *Y. ruckeri* to persist within trout macrophages, performing experiments to measure the uptake, intracellular survival, respiratory burst response, and the viability of the macrophage post infection. *Yersinia ruckeri* induced the production of reactive oxygen species (ROS) within the host. However, the bacteria were able to survive and replicate in the macrophage for 24 h. It was documented that the bacteria shifted from a predominantly extracellular existence during the first week post infection, to an intracellular one thereafter (Ryckaert *et al.*, 2010). Horne and Barnes (1999) observed that serotype O1 isolates produce superoxide dismutase and catalase

enzymes. These enzymes could allow the bacteria to survive inside the macrophages by avoiding phagocytic killing (Horne & Barnes, 1999). Similarly, all three human pathogenic *Yersinia* can survive and multiply in macrophages, being of critical importance during the initial stages of colonisation (Pujol & Bliska, 2005).

Davies (1991) identified a relationship between serum resistance and virulence in *Y. ruckeri*. Virulent serotype O1 isolates were serum resistant, whereas avirulent serotype O1 isolates were mostly serum sensitive (Davies, 1991d).

1.2.3.6 Ruckerbactin

Iron is an essential nutrient for bacterial growth, survival and virulence. Although host tissues are abundant in iron, it is generally contained internally, and bacteria are not readily able to harvest this. Similarly extracellular iron is bound by high affinity compounds including transferrin and lactoferrin making scavenging difficult. Bacterial cells have evolved certain techniques to acquire iron in order to survive inside the host (discussed later in section Fig. 1.20), with some bacteria able to excrete small compounds that can chelate iron with an extremely high affinity, called siderophores (Faraldo-Gómez & Sansom, 2003). Siderophores are low molecular mass iron chelating compounds often regarded as virulence factors due to the importance of iron to bacterial growth, colonisation and virulence. Siderophores are divided into three major classes; catecholates, hydroxamates and heterocyclic compounds (Crosa *et al.*, 2004; Wooldridge & Williams, 1993). Once bound to iron, the siderophore-Fe^{3+} complex recognises a specific bacterial OM receptor and is translocated into the cytosol where the iron is discharged from its siderophore and utilised for different metabolic pathways.

It had been suggested that *Yersinia* spp. do not produce siderophores, instead gaining their supply of iron from direct cell contact (Perry & Brubaker, 1979). Davies (1991b) observed the effect of iron limitation on the OM profiles of *Y. ruckeri*. In the absence of iron, four additional OMPs (72, 69.5, 68 and 66 kDa) were produced; however, the production of siderophores was not detected in any of the 36 isolates examined. The catechol siderophore ruckerbactin has subsequently been identified by Fernandez *et al.* (2004) in *Y. ruckeri* and has

been shown to be involved in iron acquisition under iron-depleted conditions both *in vitro* and *in vivo*. A high degree of homology was found between the ruckerbactin receptor (RupA) and the ferrichrysobactin receptor (Fct) from *Erwinina chrysanthemi*, indicating more similarity with hydroxamate receptors than with other catecholate receptor proteins such as FepA from *E. coli*. This suggests that ruckerbactin is more similar to chrysobactin than to enterobactin in chemical structure. The LD50 of strains deficient in siderophore production was higher than that of the wild-type strain (Fernández *et al.*, 2004). It has been shown that genes involved in the siderophore pathway in *Y. ruckeri* are upregulated during the infection of fish, with greater induction at 18°C than at 28°C. Therefore, the ability to scavenge iron from the environment is an essential requirement for the development of the infection process (Fernández *et al.*, 2004, 2007b).

1.2.3.7 Lipopolysaccharide (LPS)

Lipopolysaccharide (LPS), also termed endotoxin, is the major component of the OM of Gram-negative bacteria, serving as a physical barrier providing protection for the bacteria against its surroundings. In general, most Gram-negative bacteria contain LPS, covering more than 90% of the outer side of the OM. Two distinct forms of LPS exist in Gram-negative bacteria; the smooth form consisting of a O-antigen polysaccharide region, and the rough form, lacking this. In the smooth form, three well defined regions of LPS have been described; (1) Lipid A, composed of sugars and fatty acids anchoring the LPS molecule to the OM. Lipid A is responsible for the toxic effects of infections. (2) The core oligosaccharide region is composed of approximately ten monosaccharides connecting the lipid A to the O-antigen region, and (3) the O-specific antigen, which consists of repeating monosaccharides (Fig. 1.12) (Hitchcock *et al.*, 1986). LPS is recognised by Toll-like receptor 4 (TLR4) in the host as an invading bacterial pathogen, stimulating an immune response.

Y. ruckeri has a smooth type LPS structure typical of enteric bacteria, with different patterns apparent depending on the serotype of the strain (Davies, 1989; Romalde & Toranzo, 1993). It has been demonstrated previously that isolates of serotype O1, O2, O5, O6, and O7 have different core-polysaccharide and O-antigen ladder patterns from each other, allowing them to be

distinguished by western-blot analysis. It was observed that isolates of serotype O1 and O7 gave a weak agglutination with both anti-O5 and anti-O6 antiserum also.

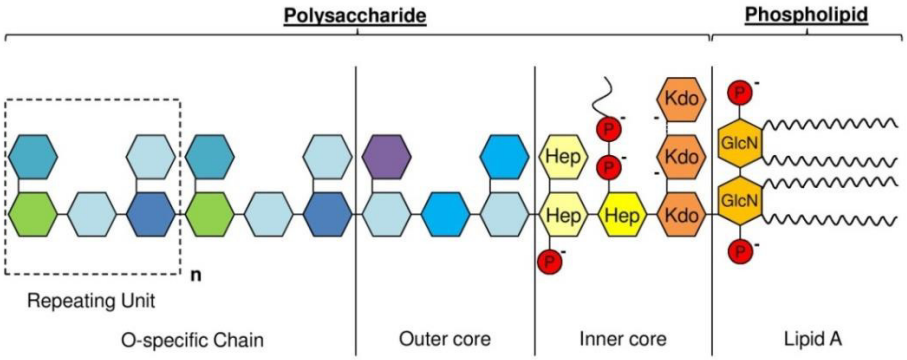


Fig. 1.12 Schematic representing the chemical structure of LPS in Gram-negative Enterobacteriaceae.
LPS consists of a membrane-anchoring lipid A domain and a covalently linked polysaccharide or oligosaccharide portion. In smooth type LPS, the polysaccharide domain is composed of the lipid A proximal core region and the terminal O-specific chain formed by up to 50 repeating units. An inner and an outer core region are commonly distinguished. The lipid A domain represents the primary immunostimulatory centre of LPS determining endotoxicity in mammalian species. GlcN, glucosamine; Kdo, '2-keto-3-deoxyoctulosonic acid' (3-deoxy-D-manno-octulosonic acid); Hep, D-glycero-D-manno-heptose. This diagram was reproduced from (Alexander & Rietschel, 2001).

Having conducted virulence and serum killing assays with a range of isolates representing different serotypes, biotypes and OMP-types, Davies (1991a) identified that isolates belonging to different clonal groups of serotype O1 showed different serum susceptibility. Davies showed that all serum sensitive isolates were avirulent. This suggests that differences in the composition of the LPS molecules of the different serotype O1 clonal groups may account for differences in serum susceptibility and virulence. Variations in the LPS profiles of both serum-sensitive and –resistant isolates have been observed also (Davies, 1989, 1991a).

1.2.3.8 Role of plasmids in virulence

Several plasmids have been described in *Y. ruckeri* as authors try to establish common links with similar virulence plasmids in the human pathogenic *Yersinia* strains. These species harbour a 70 kb virulence plasmid encoding a TTSS, which is important in counteracting the immune response of the host and ensuring survival (Tobback *et al.*, 2007). However, *Y. ruckeri* does not encode this plasmid. Garcia *et al.*, (1998) investigated the plasmid profile of 183 *Y. ruckeri*

isolates from a wide variety of sources, and identified eight different profiles; a large 75 MDa plasmid was common to all strains. The presence of a 75 MDa plasmid in most *Y. ruckeri* strains agrees with the work of other authors (Guilvout *et al.*, 1988; Romalde & Toranzo, 1993); large plasmids have been recovered from serotype O1 strains, but not from other serogroups. Small plasmids have been found in different serotypes as well as in serotypes O1, suggesting a role in virulence for the larger plasmid (De Grandis & Stevenson, 1982; Tobback *et al.*, 2007). Guilvout *et al.*, (1988) examined the plasmid profiles in *Y. ruckeri, Y. pestis*, *Y. pseudotuberculosis* and *Y. enterocolitica*. The large plasmid of *Y. ruckeri* appears to be significantly different from the virulence plasmid associated with human pathogenic *Yersinia* species.

1.2.3.9 Role of environmental factors in virulence

Environmental factors have a significant involvement in how bacteria interact with their hosts, and subsequently cause infection. In aquaculture environments, the effect is slightly heightened, as the bacterium and the hosts are more susceptible to alterations. It has been shown that factors such as temperature and salinity can influence the establishment and the severity of *Y. ruckeri* infections (Altinok & Grizzle, 2001; Altinok, 2004). Salinity can also have an effect on the pathogenesis of the bacteria, with most isolates having been obtained from fresh water samples; *Y. ruckeri* has been found to survive for four months in untreated water, whereas its survival time notably decreases in aquatic environments with high salinity (Thorsen *et al.*, 1992).

Understanding the variety of factors that are responsible for causing and sustaining infection within a host is of crucial importance in the control of a pathogen. Many of the virulence factors discussed are found at the OM, and examining this component can enhance our understanding of how the bacteria operate.

1.3 Enteric redmouth (ERM) disease

Yersinia ruckeri is the causative agent of ERM disease in salmonids. Since the original isolation from rainbow trout, the bacterium has been responsible for economic losses in the aquaculture industry, particularly in salmonid

aquaculture. The disease has since become prevalent in many countries and is found widely spread across Europe and the USA, with its first isolation in the UK coming in the 1980's (Davies & Frerichs, 1989).

1.3.1 **Epidemiology**

While *Y.* ruckeri was initially isolated from rainbow trout in North America, the bacterium has since been recovered from many different species, including salmonid and non-salmonid fish, as well as other unrelated species including birds, insects and mammals (Table 1.4). Of most economic importance however are the salmonid fish, particularly Atlantic salmon and rainbow trout.

Table 1.4 Species from which *Y. ruckeri* has been isolated.
This table was reproduced from (Horne & Barnes, 1998).

Species	Common name
Salmonids	
Onchorhynchus kisutch	Coho salmon
Onchorhynchus mykiss	**Rainbow trout**
Onchorhynchus nerka	Sockeye salmon
Onchorhynchus tschawytscha	Chinook salmon
Salmo clarkii	Cutthroat trout
Salmo salar	**Atlantic salmon**
Salmo trutta	Brown trout
Salveneins alpinus	Arctic char
Salvenlinue fontinalis	Brook trout
Non-salmonids	
Acipenser baeri	Sturgeon
Anguilla anguilla	Eel
Aristichthys nobilis	Bighead carp
Carassius auratus	Goldfish
Coregonus artegii	Cisco
Coregonus peled	Whitefish
Coregonus muksun	Whitefish
Cyprinus carpio	Common carp
Hypophthalmichthys molitrix	Silver carp
Ictalurus punctatus	Channel catfish
Lota lota	Burbot
Notemigonus atherinoides	Emerald shiner
Pimephales promelas	minnow
Scophthalmus maximus	Turbot
Solea solea	Sole
Other animals	
Homo sapiens	Human
Lutrinae	Otter
Ondrata zibethica	Muskrat
Falco tinnunculus	Kestrel
Cathartes aura	Turkey vulture
Eisenia foetida	Earthworm
Laridae	Gulls

Yersinia ruckeri has been recovered from many locations around the world, highlighted in Fig. 1.13 and Table 1.5. The geographical distribution of isolates is particularly striking, with isolates recovered from almost every continent (except Antarctica). Studies that have conducted intensive characterisation of isolates from different geographical locations have reported a certain degree of specificity to each location.

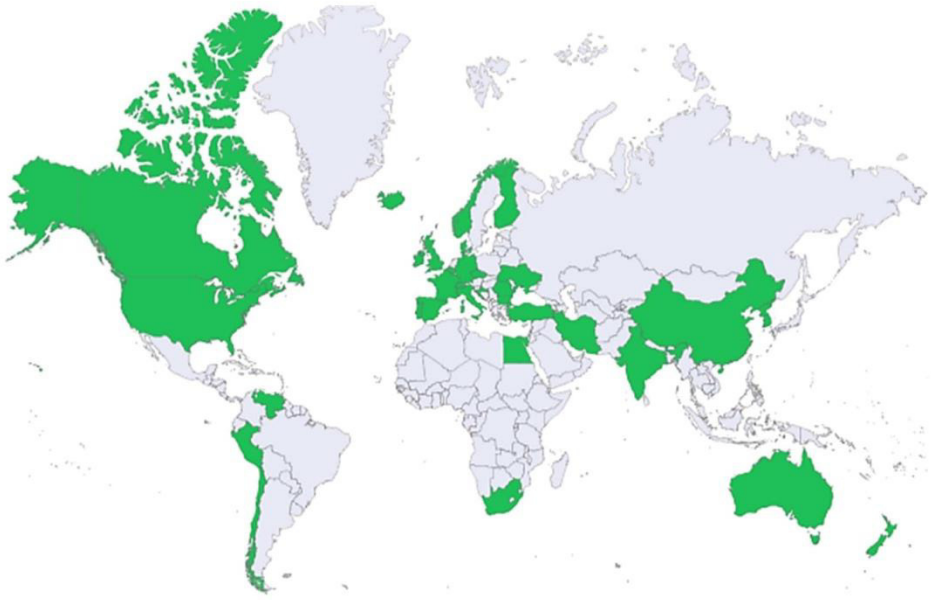


Fig. 1.13 Geographic locations of *Y. ruckeri* recovered worldwide.
The bacterium has now been recovered from almost every continent in the world. Countries highlighted in green correspond to countries which have reported *Y. ruckeri* recovery worldwide.

Table 1.5 Reported incidence of *Y. ruckeri* isolates recovered worldwide.

Country	Reference	Country	Reference
Europe		***North America***	
UK	Davies *et al.*, 1990	USA	Arias *et al.*, 2007
Belgium	De Keukeleire *et al.*, 2014	Canada	Stevenson & Daly 1982
Bulgaria	Gelev *et al.*, 1984	***South America***	
Switzerland	Meler 1986	Chile	Troncoso *et al.*, 1994
Czech Republic	Vladik & Prouza 1990	Peru	Bravo & Kojagura, 2004
Germany	Fuhrmann *et al.*, 1984	Venezuela	Alvarez *et al.*, 1987
Denmark	Dalsgaard & Madsen 2000	***Africa***	
Spain	Romalde *et al.*, 1994	Egypt	Eissa *et al.*, 2008
Finland	Rintamaki *et al.*, 1986	South Africa	Bragg & Henton, 1986
Croatia	Oraic *et al.*, 2002	***Asia***	
Ireland	McCormick *et al.*, 1993	China	Wang *et al.*, 2015
Iceland	Gudmundsdottir *et al.*, 2014	India	Manna *et al.*, 2003
Italy	Battisti *et al.*, 2008	Iran	Akhlaghi & Yazdi, 2008
Norway	Sparobe *et al.*, 1986	Korea	Joh *et al.*, 2010
Portugal	Sousa *et al.*, 2001	Singapore	McArdle *et al.*, 1985
Romania	Dascalescu *et al.*, 2003	***Oceania***	
Turkey	Onuk *et al.*, 2011	Australia	*Llewellyn* 1980
Isle of Man	Verner-Jeffreys *et al.*, 2011	New Zealand	Anderson *et al.*, 1994

The study conducted by Davies and Frerichs (1989) examined 147 isolates from Australia, North America, South Africa and Europe identified variation in isolates recovered from these locations in terms of motility and other biochemical reactions. More recent studies have uncovered clonal groups specific to each location. Davies (1990) combined biotype and OMP-type to identify six different clonal groups. Clonal groups one (biotype 1, OMP-type 1), two (biotype 2, OMP-type 1) and four (biotype 2, OMP-type 2) were unique to Australia, the UK and Norway, clonal groups three (biotype 1, OMP-type 2), five (biotype 1, OMP-type 3) and six (biotype 1, OMP-type 4) were represented by numerous countries (both European and North American). Clonal groups 2 and 5 were identified as the major virulent groups however. Wheeler *et al.*, (2009) used PFGE (discussed previously, 1.2.2.4.2) to decipher clonal groups apparent in rainbow trout and Atlantic salmon populations from Europe and North America. They showed that UK, European and North American isolates recovered from rainbow trout were distinctly different from one another. In isolates recovered from Atlantic salmon, European and UK isolates were both more similar to North American isolates than each other.

The rapid expansion of the aquaculture industry has in turn led to increased global movement of aquatic animals and their products. This in turn may have led to the emergence of many new diseases in fish. Several examples exist where the natural or unintentional movement of infected hosts has resulted in the emergence of diseases in different countries. Infectious haematopoietic necrosis (IHNV) was originally endemic in North America among salmon, emerging in the 1970s as an important pathogen of farmed rainbow trout. Subsequently, the virus was spread by the movement of contaminated eggs to several countries of Western Europe and East Asia, where it emerged to cause severe losses in farmed rainbow trout (Bootland & Leong, 1999). Infectious salmon anaemia (ISA) is a disease of Atlantic salmon, and was initially identified as the causative agent of outbreaks and high rates of mortality among Atlantic salmon in Norway (Nylund *et al.*, 1995). Since, the virus has caused losses in other areas of Western Europe where Atlantic salmon are farmed, and was also confirmed to be the cause of an emerging haemorrhagic kidney disease of farmed Atlantic salmon in Canada and the USA (Kibenge *et al.*, 2004; Lovely *et al.*, 1999; Mjaaland *et al.*, 2002). More recently, ISA has caused very extensive

losses in the Atlantic salmon industry in Chile. Genetic analysis has revealed that the Chilean isolates group with those from Norway and that the virus was likely transferred to Chile sometime around 1996 by the movement of infected eggs (Kibenge *et al.*, 2009). This suggests that the emergence of ISA appears to relate to transmission via the movement of fish or eggs used in aquaculture. The emergence of viral disease in aquatic systems can also been driven by anthropogenic factors unrelated to aquaculture such as the movement of pathogens or hosts via ballast water in ships, movement of bait by anglers and unintentional movement in other biotic or abiotic vectors (Walker & Winton, 2010).

1.3.2 **Clinical signs of ERM disease**

Enteric redmouth disease is so called due to the common feature of a reddening in the mouth and throat, caused by a subcutaneous haemorrhaging (Fig. 1.14A). However, subcutaneous haemorrhaging is not confined to the throat and mouth as it is seen at the base of the pectoral and pelvic fins, and in the gill filaments (Fig. 1.14B). Several identifying factors can lead to ERM diagnosis including lethargy, loss of equilibrium, loss of appetite and petechial haemorrhages occurring on the liver, pancreas, pyloric caecae, and swim bladder. Externally haemorrhaging around the eye (Fig. 1.14C) and a uni- or bilateral exophthalmos is common (Fig. 1.14D). Haemorrhaging along the flanks can also be seen. The spleen can become enlarged in tandem with inflammation of the lower intestine as it becomes filled with a viscous yellow fluid (Rodgers, 1992).

1.3.3 **Pathology and histopathology**

Histological examination of tissues shows a general septicaemia with an inflammatory response in virtually all tissues. Bacterial colonisation occurs in tissues such as kidney, spleen, heart, liver and gills and in areas of petechial haemorrhage (Rucker, 1966). Pathological changes in the gills, including hyperaemia, oedema and desquamation of the epithelial cells in the secondary lamellae, have been described in experimentally infected fish. Focal necrosis was observed in the liver, kidney and spleen of some fish after intraperitoneal (i.p.) injection with *Y. ruckeri* (Avci & Birincioglu, 2005; Berc *et al.*, 1999). Prominent peritonitis probably caused by long-lasting inflammation in the

peritoneal cavity and enteritis have been described in carp that survived i.p. injection with *Y. ruckeri* (Berc *et al.*, 1999).

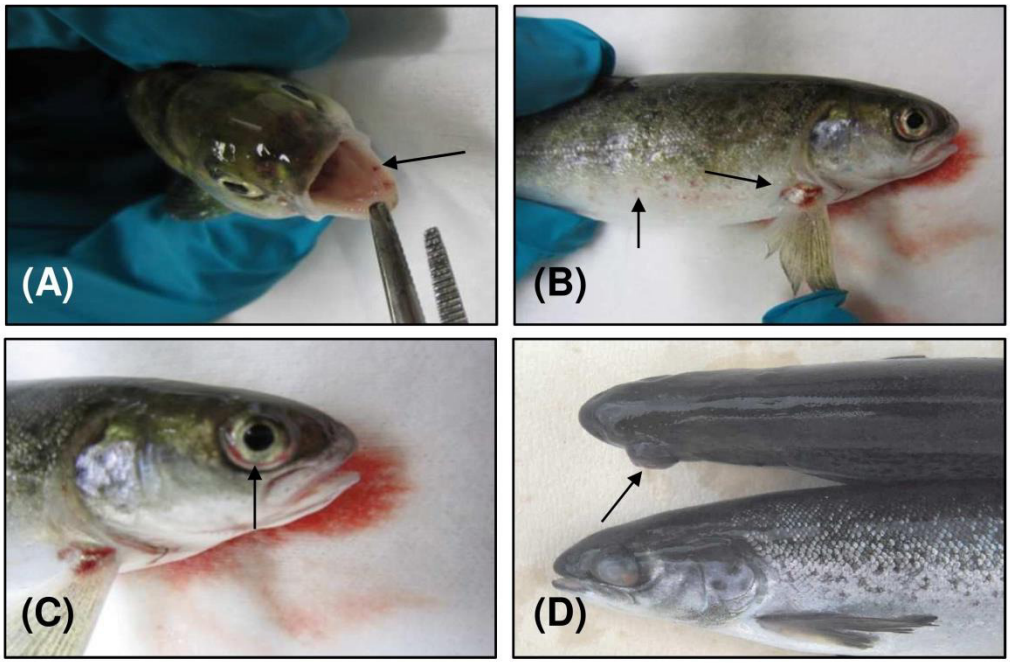


Fig. 1.14 Characteristic symptoms of enteric redmouth disease.
(A) haemorrhaging (arrow) of the oral cavity; **(B)** haemorrhaging (arrow) at the base of the pectoral fin and on the flanks; **(C)** haemorrhaging (arrow) of the eye; **(D)** exopthalmus of the eye. The images shown are taken from an infected Atlantic salmon.

Bacteria are readily detected free in the blood and in circulating and sequestered macrophages; tissue localisation of bacteria at sites of haemorrhage may also be evident. Histologically, salmon fry may contain overwhelming numbers of bacteria with high concentrations detectable in macrophages of kidney and liver sinusoids. These changes are most evident in the gills, which may show blood stasis and bacterial clumps. *Yersinia ruckeri* infection in Atlantic salmon smolt during acclimatisation stress is characterised by fewer bacteria in blood, with congestion, haemorrhage and tissue localisation more apparent than acute inflammation. Localisation in the choroid and meninges is common, and encephalitis may be seen (Carson & Wilson, 2009).

1.3.4 **Transmission**

Yersinia ruckeri infections are spread between fish by direct contact with either infected animals or carriers. Initially, Rucker (1966) recognised a carrier state for *Y. ruckeri* as the bacterium were isolated from surviving fish two months after exposure, while Busch and Lingg (1975) recovered the bacterium from

different organs of an asymptomatic trout; up to 25% of the fish in a rainbow trout population could carry *Y. ruckeri* in their lower intestines. The intestinal shedding of the pathogen was shown to cause recurrent infection and mortality within the population on a cyclical basis, with this periodic shedding crucial in the spread of disease (Busch & Lingg, 1975; Rucker, 1966). The pathogen has been isolated from the faeces of healthy fish two months after infection (Rodgers 1992). The impact of stress on the spread of infection within fish populations is well documented; increasing the water temperature from ambient to 25°C prompted carriers to transmit *Y. ruckeri* to clinically healthy fish, while unstressed fish did not (Hunter et al 1980). It has also been noted previously that survivors of the disease can become withdrawn from the population (Rucker, 1966; Busch & Lingg, 1975), while moribund and deceased fish can shed vast quantities of bacteria into the water remaining infective for extended periods in the aquatic environment (Rodgers, 1992); the bacteria has been shown to survive in fresh or brackish water for up to 100 days, with survival greatly reduced in higher salinities (Thorsen *et al.*, 1992). *Y. ruckeri* also has a biofilm-forming capacity (Coquet *et al.*, 2002), being found to produce biofilms on materials commonly found within fish farms. This may be a source of recurrent infection. The transmission mode of *Y. ruckeri* has been related to other putative vectors, including birds, wild fish (Willumsen, 1987), muskrat (*Ondrata ziebethica*) (Stevenson & Daly, 1982) and humans (Farmer *et al.*, 1985). Sauter *et al.*, (1985) recovered *Y. ruckeri* from disinfected non-fertilised eggs of Chinook salmon whose offspring experienced low mortality rates from fertilisation to 12 weeks, suggesting that infection with *Y. ruckeri* is vertically transmissible.

Yersinia ruckeri was isolated from minnows (*Pimephales promelas*) in Belgium and France, which had been imported as bait fish from Missouri and Arkansas, USA, respectively (Michel *et al.*, 1986). Upon observing the transfer of bacteria from asymptomatic to uninfected fish, Hunter *et al.*, (1980) observed that the organism was shed in the faeces over 36-40 day cycles, correlating with seasonal variations of water temperature, along with crowding, handling and other stressors.

1.3.5 **Diagnosis**

Several selective media have been developed for the identification of *Y. ruckeri*, including Waltman-Shotts (Waltman & Shotts, 1984) and ROD (Ribose ornithine deoxycholate) agar (Rodgers, 1992). Isolation using solid media has proven useful due to the rapid growth rate of the bacterium. Growth on TSA produces round cream colonies within 48 h, while growth on blood agar plates will produce white, opaque colonies. Identification can be confirmed using several biochemical tests.

1.3.5.1 Biochemical tests

Extensive biochemical characterisation studies of *Y. ruckeri* have been conducted, the results of which are summarised in Table 1.6. The optimal growth temperature of *Y. ruckeri* is between 20 - 25°C, although the bacteria will grow at 37°C. Most phenotypic characterisations of this bacterium however have occurred at 22°C (Davies & Frerichs, 1989; Davies, 1990, 1991a, b, c, d; Verner-Jeffreys *et al.*, 2011; Wheeler *et al.*, 2009).

Biochemical tests can be used to distinguish *Y. ruckeri* from other species. As with other members of the Enterobacteriaceae, *Y. ruckeri* is glucose-fermentative, oxidase-negative and nitrate-reductive (Ross *et al.*, 1966). The presence of catalase, ß-galactosidase, Lysine decarboxylase and Ornithine decarboxylase are distinguishing characteristics in conjunction with the absence of Arginine dihydrolase, H_2S, Indole and Oxidase-phenylalanine deaminase. *Yersinia ruckeri* is able to ferment glucose and mannitol, although is unable to utilise inositol, rhamnose, sucrose, melibiose or arabinose. Biochemical test kits including API 20E have also proved useful in the identification of *Y. ruckeri* (Topic Popovic *et al.*, 2007).

1.3.5.2 Serological tests

Utilising the identification of antibodies, several serological techniques have been applied for diagnosis of *Y. ruckeri* including enzyme-linked immunosorbant assay (ELISA) (Cossarini-Dunier, 1985), agglutination tests (Davies, 1990) and immunofluorescence antibody technique (IFAT) (Smith *et al.*, 1987).

Table 1.6 Biochemical characteristics of *Y. ruckeri*.
This table was adapted from (Austin & Austin, 1993) and (Akhlaghi & Yazdi, 2008).

Characteristic	Result
Fermentative Metabolism	+
Production of:	
Arginine dihydrolase	-
Catalase	+
β-galactosidase	+
H2S	-
Indole	-
Lysine decarboxylase	+
Ornithine decarboxylase	+
Oxidase-phenylalanine deaminase	-
Phosphatase	-
Methyl red test	+
Nitrate reduction	+
Voges-Proskauer (VP) reaction	-
Degradation of :	
Aesculin	-
Chitin	-
DNA	-
Elastin	-
Gelatin	+
Pectin	-
Tributyrin	-
Tween 20	Variable
Tween 40	+
Tween 60	+
Tween 80	Variable
Urea	-
Growth in 0% (w/v) NaCl	+
Growth in 0-3% (w/v) NaCl	+
Growth in 6.5% (w/v) NaCl	-
Utilization of sodium citrate	+
Acid production from:	
Fructose	+
Glucose	+
Inositol	-
Lactose	-
Maltose	+
Mannitol	+
Raffinose	-
Salicin	-

1.3.5.3 Molecular tests

Molecular techniques including restriction fragment-length polymorphism (RFLP) (Garcia *et al.*, 1998) and PCR (Altinok *et al.*, 2001; Gibello *et al.*, 1999) are often used. Gibello *et al.* (1999) developed a PCR assay based on the selective

amplification of the 16S rRNA gene for *Y. ruckeri* identification in infected trout tissues. This was sensitive enough to detect low levels of bacteria in asymptomatic carriers. Altinok *et al.* (2001) described a PCR method for the detection of *Y. ruckeri* in the blood of rainbow trout, while PCR amplification of *yruR* and *yruI*, the genes responsible for the quorum sensing system, have also proved to be highly specific (Altinok *et al.*, 2001; Temprano *et al.*, 2001). Saleh *et al.* (2008) designed a loop-mediated isothermal amplification (LAMP) assay for the detection of *Y. ruckeri*, also utilising *yruR* and *yruI*, which proved particularly useful due to the relatively little equipment needed (Saleh *et al.*, 2008). Other nonlethal methods for the diagnosis of *Y. ruckeri* are the culture of faeces from the posterior intestine using inoculating loop stabs (Busch & Lingg, 1975; Rodgers, 1992) and the biopsy of head kidney (Noga *et al.*, 1988).

1.3.6 **Host specificity**

As Atlantic salmon and rainbow trout represent the two major farmed species of salmonid aquaculture worldwide, understanding the degree of host specificity of isolates that can cause infection in each is of vital importance. Variation in isolates affecting both Atlantic salmon and rainbow trout has been discussed previously, although usually in phenotypic and genetic characterisation studies rather than challenge studies (section 1.2.2.3 & 1.2.2.4).

Studies exploring virulence are relatively limited, however major differences have been observed in the pathogenicity of isolates recovered from Atlantic salmon towards rainbow trout, and vice versa (Haig *et al.*, 2011). The organisms that typically cause ERM in rainbow trout represent a closely-related subgroup of serotype O1 (biotype 1, OMP-type 3; biotype 2, OMP-type 1). However, challenge studies identified that these isolates are also able to infect Atlantic salmon, while isolates that were recovered from and are pathogenic towards Atlantic salmon are not necessarily virulent towards Rainbow trout. Haig *et al.*, (2011) showed significant mortalities in salmon fry after challenge with serotype O2 and O5 isolates, although there was reduced virulence in older salmon (presumably due to an improved immune response). This is consistent with reports from hatcheries that *Y. ruckeri* outbreaks, particularly those associated with serotype O2 and O5 isolates, normally affect fry rather than older fish. It is suggested therefore, that serotype O2 and O5 isolates genetically very similar to

those tested have been circulating in UK salmon hatcheries for more than 15 years (Haig *et al.*, 2011). Interestingly, isolates were more virulent when tested at 16°C than 12°C, which is consistent with reports that farmers typically see *Y. ruckeri* problems in salmon hatcheries when temperatures are greater than 15°C. A serotype O1 isolate from the study was highly virulent to salmon but not to trout, even though the isolate was resistant to killing by naive serum collected from both species (Haig *et al.*, 2011).

An earlier study by Davies (1991a) demonstrated the virulence of isolates of different serotypes of *Y. ruckeri* towards rainbow trout, and drew comparisons to previous studies that used brook trout (Cipriano *et al.*, 1987). While Davies identified specific groups of isolates that were virulent towards rainbow trout, isolates that had shown virulence in brook trout were not virulent in rainbow trout (Davies, 1991a).

1.3.7 **Treatment and prevention**

1.3.7.1 Treatment

1.3.7.1.1 Antimicrobials

Antimicrobial compounds are often used in the treatment of *Y. ruckeri* infections in fish. Rucker (1966) described treating fish for five days with sulphamethazine, followed by three days of chloramphenicol or oxytetracycline administration. Rodgers and Austin (1983) used oxolinic acid for prophylaxis and therapy of ERM in rainbow trout. Although *Y. ruckeri* is sensitive to many antibiotics, acquired resistance of stains to various antimicrobial agents has been reported. Post, (1987) highlighted the complete resistance of some isolates in the USA to therapeutic levels of both sulphamerazine and oxytetracycline. Resistance to both tetracyclines and sulphonamides has also been demonstrated by other authors (De Grandis & Stevenson, 1985).

1.3.7.1.2 Probiotics

Probiotics are living microorganisms, which provide beneficial symbiosis to the host. Administration of certain probiotics to cultured trout has been shown to enhance their survival when exposed to *Y. ruckeri*. These live microbial feed

supplements are defined as beneficially affecting the host by the production of inhibitory compounds, competition for chemicals and adhesion sites, immune modulation and stimulation, and improving the microbial balance within the host (Tobback, 2009).

Food supplemented with spores of *Bacillus subtilis* and *Bacillus licheniformis* provided improved resistance in rainbow trout against infection with *Y. ruckeri* (Raida *et al.*, 2003). Similarly, food supplemented with *Carnobacterium maltaromaticum* or *Carnobacterium divergens* exhibited enhanced protection against *Y. ruckeri;* the competing cultures enhanced cellular and humoral immune responses within the fish (Kim & Austin, 2006). Rainbow trout fed with *Enterobacter cloacae* and *Bacillus mojavensis* for a total of 60 days showed an increased survival following *Y. ruckeri* challenge (Capkin & Altinok, 2009).

1.3.7.2 Prevention

1.3.7.2.1 Vaccination

Ross and Klontz (1965) saw the need for a vaccine against ERM after outbreaks at US fish farms. Their approach involved feeding rainbow trout pellets containing killed whole cells of *Y. ruckeri*; they saw an increased survival in comparison to control fish (Ross & Klontz, 1965).

Vaccines for the prevention of ERM disease were the first fish vaccines to be commercialised in 1976, composed of formalin-killed whole bacterial cells (Welch *et al.*, 2011). Administered by immersion, spray, injection or oral routes, they provide good levels of protection against disease.

The current vaccination strategy for the prevention of ERM involves immersion vaccination through bathing, or injection (i.p. or intramuscular [i.m.]) at the fry stage. It is standard practice to administer a booster vaccination six months post immersion vaccination to provide continued protection (Tatner & Horne, 1985). Intraperitoneal vaccination can provide better levels of protection. However, as ERM can affect fish at the hatchery stage, dip vaccination is more suitable (Håstein *et al.*, 2005). Although vaccination will prevent the onset of infection, the carrier state discussed previously will not be eliminated; fluctuations in temperature or stressful situations may still cause carrier fish to

shed bacteria into the water, infecting fish that have not been successfully vaccinated (Bruno & Munro, 1989). It has been demonstrated that intensive vaccination strategies may impart a selective pressure on bacteria, in turn driving the evolution of novel phenotypic and immunogenic characteristics - this could be responsible for the onset of disease in fish that have been previously vaccinated (Bachrach *et al.*, 2001).

Currently, two vaccines are primarily used for the prevention of ERM in rainbow trout. The first (AquaVac ERM™; MSD Animal health) contains inactivated cells of a biotype 1, serotype O1 isolate (the 'Hagerman' strain). The second (AquaVac RELERA™; MSD Animal health) contains inactivated cells of a biotype 2, serotype O1 isolate (EX5 strain, Austin *et al.*, 2003). However, with the emergence of new serotypes of *Y. ruckeri* in UK Atlantic salmon populations which are not present in mass produced vaccines, bespoke autogenous vaccines are used.

1.3.7.2.2 Toxoids

Fernández *et al.* (2003) found that the protease Yrp1 prepared by heat inactivation could confer relatively high survival rates in treated fish. The protease Yrp1 contributes to bacterial virulence through involvement in the invasion of different tissues and host colonisation (discussed in section 1.2.3.2). The extracellular location of the Yrp1 protease and its role in virulence led to reasonable expectations that some protection against ERM disease would be provided when a toxoid of the protein was used as an immunogen. Similar levels of protection to using whole cell vaccines were seen with this toxoid (Fernández *et al.*, 2003).

1.3.7.2.3 Attenuated vaccines

Interest in the use of live attenuated vaccines against bacterial pathogens in fish has increased, as they can provide better protection. In general, they elicit a stronger cell-mediated response than whole cell vaccines. Temprano *et al.* (2005) utilised *aroA*, (encoding the enzyme 5-enolpyruvylshikimate-3-phosphate synthase) whose function is essential for bacteria to survive and grow *in vivo* and cause disease, in an attenuated vaccine trial. The authors found that mutations of *aroA* lead to an inability of the bacterium to synthesise the organic

compounds necessary for its growth. The bacterium will therefore be unable to grow in fish tissues where there is a lack of these components, and becomes less virulent. Vaccination using this mutant conferred 90% relative percentage survival (RPS) against the biotype 1 variant of *Y. ruckeri* (Temprano *et al.*, 2005).

1.4 The Gram-negative cell envelope

The majority of bacterial typing schemes including biotype, serotype and OMP-type rely on characteristics found at the bacterial surface. The structure of the cell, particularly the cell envelope, can be crucial in the pathogenesis of the organism.

The cell envelope of Gram-negative bacteria comprises two membranes; the inner membrane (IM) and OM separated by the periplasm. The IM forms the boundary between the cytoplasm and the periplasm, while the OM separates the periplasm from the external environment. Although these two membranes are both lipid bilayers which contain proteins, their structure and composition are very different, due to their diverse functions and surrounding environments (Fig. 1.15) (Ruiz *et al.*, 2006).

1.4.1 **Inner membrane (IM)**

The lipid component of the IM is composed exclusively of phospholipids, mainly phosphatidylethanolamine (70-80%), phosphatidylglycerol and cardiolipin, equally distributed among the inner and outer leaflets, interspersed with proteins (Koebnik *et al.*, 2000). These proteins can be of two types; (1) integral membrane proteins or (2) lipoproteins. Integral inner membrane proteins (IMPs) span the IM with α-helical transmembrane domains, whereas IM lipoproteins are anchored to both the cytoplasmic and periplasmic sides of the membrane by lipid modifications at the N-terminus of the prolipoprotein. Inner membrane proteins are primarily involved in cellular processes including energy production, signal transduction, cell division and transport across the membrane.

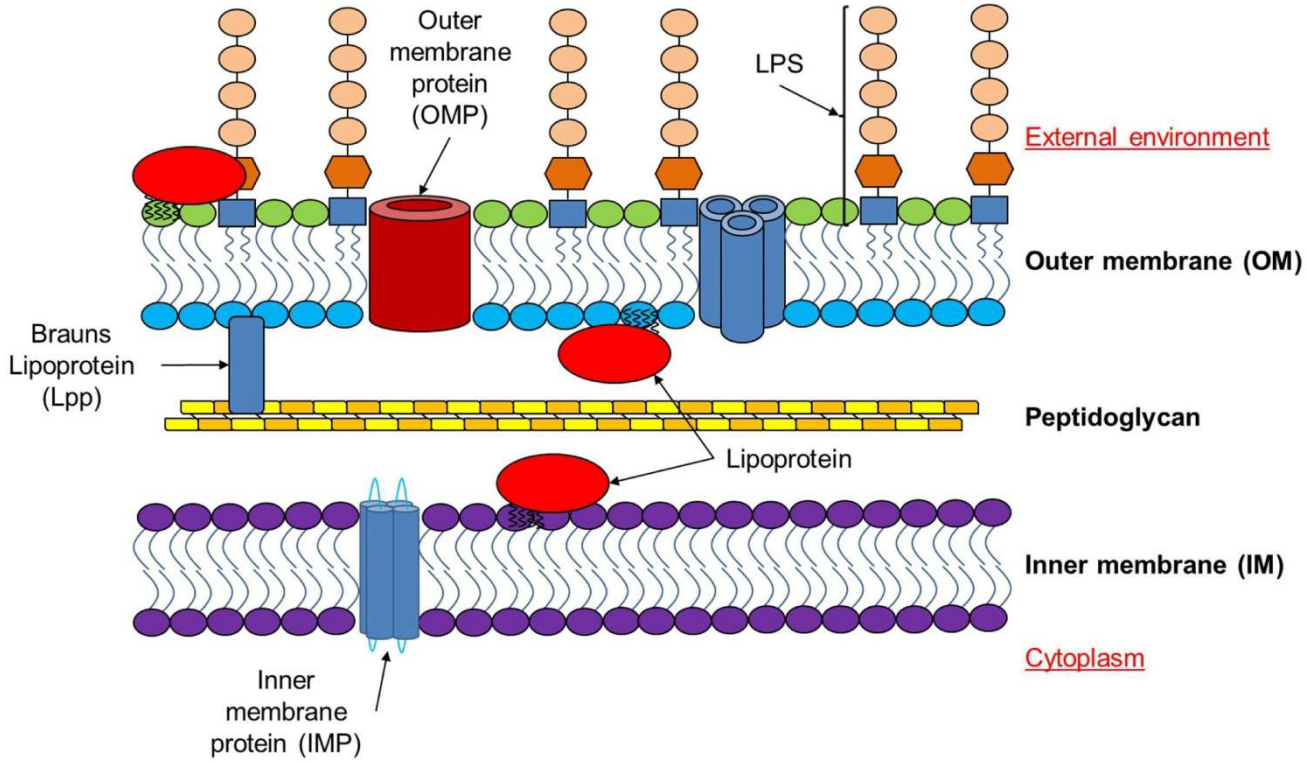


Fig. 1.15 The Gram-negative cell envelope.
Gram-negative cells have an IM, a peptidoglycan layer and an OM. The outer leaflet of the OM is composed mostly of LPS, anchored to the OM by the LPS lipid A domain (section 1.2.3.7). The inner leaflet of the OM and the entire IM are composed of phospholipids, with both bilayers containing a range of different types of membrane protein. Outer membrane proteins (OMPs) are dispersed throughout the OM, forming channels, acting as receptors and with other specialised functions. Lipoproteins are found on both the IM and OM, either integrally or tethered periplasmically or extracellularly to the membrane. Between the IM and OM, a thin peptidoglycan layer, made up of repeating units of the disaccharide N-acetyl glucosamine-N-actyl muramic acid, which are cross-linked by pentapeptide side chains. The OM is basically stapled to the underlying peptidoglycan by a lipoprotein called Braun's lipoprotein (Lpp).

1.4.2 **Periplasm**

The periplasm of Gram-negative bacteria contains the peptidoglycan layer serving as an extracytoplasmic cytoskeleton contributing to cell shape and rigidity, and protection against osmotic stresses. It also provides an oxidising environment where protein structures can be stabilised by disulphide (S-S) bonds, a crucial process for the folding and stability of proteins secreted through the cell envelope (Merdanovic *et al.*, 2011). Owing to the porous character of the OM, the periplasm is more directly exposed to the environment compared with the cytoplasm, which is shielded by the IM. Therefore, bacterial cells have established effective mechanisms of protein quality control that monitor the folded state of proteins and support protein folding, repair, and degradation in both the cytoplasm and the periplasm.

1.4.3 **Outer membrane (OM)**

The OM of Gram-negative bacteria is a highly specialised structure that lies outside the IM and peptidoglycan layer, forming a physical and functional external barrier between the cell and its environment. The major function of the OM is to control the passage of nutrients to the cell while providing protection against antibiotics, detergents and toxins. The phospholipid composition is slightly different to that of the IM, as the OM is asymmetric, and enriched with saturated fatty acids and phosphotidylethanolamine (Lugtenberg & Peters, 1976). The outer leaflet of the OM is composed mainly of LPS, which functions as an effective permeability barrier due to its low fluidity and strong lateral interactions between molecules (Nikaido, 2003). The OM contains both integral proteins and lipoproteins, similar to the IM. However, these integral proteins are predominantly amphipathic antiparallel ß-strands adopting a ß-barrel conformation; this allows many OMPs to serve as channels (Koebnik *et al.*, 2000). Lipoproteins located at the OM can be anchored either to the periplasmic side or the extracellular side of the OM (Tokuda & Matsuyama, 2004). About 50% of the OM mass consists of integral OMPs and lipoproteins (Koebnik *et al.*, 2000).

1.4.3.1 Outer membrane biogenesis (BAM complex)

OMPs are synthesised in the cytoplasm, transported across the IM into the periplasm by the Sec translocon (Pugsley, 1993). Molecular chaperones then

escort OMPs to the inner surface of the OM where they are recognised by the ß-barrel assembly machinery (BAM) complex (Fig. 1.16). The BAM complex is composed of five proteins; BamA, BamB, BamC, BamD and BamE (Noinaj *et al.*, 2013). BamA (or Omp85) plays a central role in OMP assembly, and is found in all Gram-negative bacteria. Orthologs of BamA are found in the mitochondira and chloroplasts of eukaryotes (Hagan *et al.*, 2011); bacterial BamA contains five polypeptide transport-associated (POTRA) domains, while chloroplast and mitochondrial orthologs contain three and one respectively. In bacteria, these five POTRA domains are located periplasmically, and receive substrates translocated across the IM through binding of four BAM accessory lipoproteins. The four BAM accessory lipoproteins (BamB, C, D and E) form a tight complex with BamA, although BamB associates separately from BamCDE with BamA. All of the BAM lipoproteins have roles in OMP biogenesis, as the deletion of any one leads to varying levels of OMP assembly defects, but only BamA and BamD are crucial for cell viability and OMP biogenesis (Malinverni *et al.*, 2006).

A number of possible mechanisms exist as to how the BAM complex folds and inserts proteins into the OM. (1) The pore-folding model; the ß-barrel of BamA offers its pore for insertion of the OMP into the membrane, and the POTRA domains and/or accessory components act to thread the OMP into the pore.

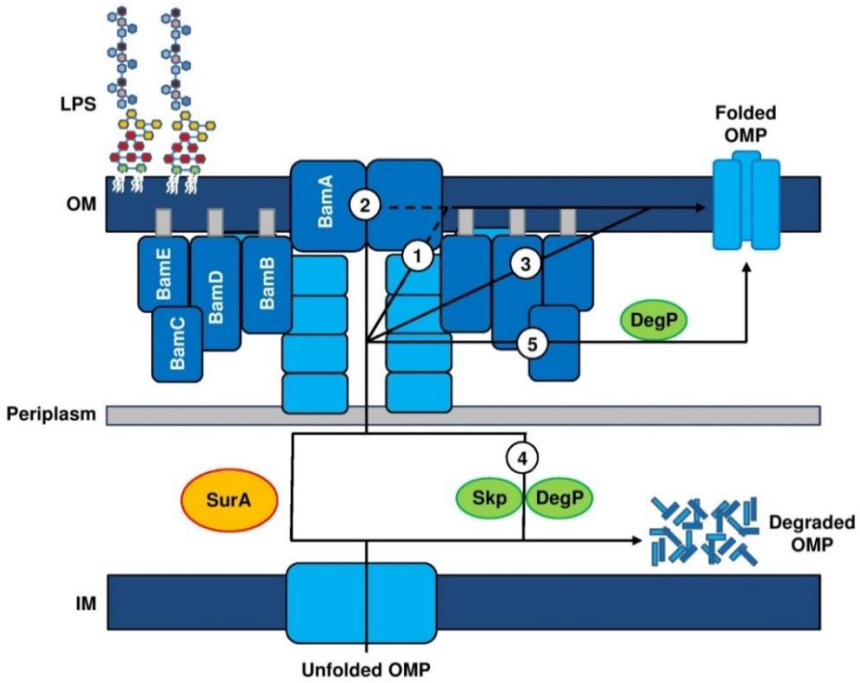


Fig. 1.16 OMP biogenesis by the BAM Complex.
(1) The pore-folding model; **(2)** The complex pore-folding model; **(3)** The barrel-folding model; **(4)** The chaperone-folding model; **(5)** The accessory folding model. This schematic was adapted from Knowles *et al.* (2009).

(2) The complex pore-folding model; a central core is formed by a multimeric BAM complex acting as the point of insertion into the membrane. (3) The barrel-folding model; the ß-barrel of BamA provides a template for barrel folding in the vicinity of the BAM complex. (4) The chaperone-folding model; the periplasmic chaperones, in particular DegP, act to fold the protein and protect it from degradation during passage through the periplasm. The BAM complex functions only to insert the protein into the membrane. (5) Finally, the accessory folding model; the BAM complex functions to fold OMPs but does not have a function in membrane insertion. The folded OMP is then released to DegP in a quality-control led mechanism to remove incorrectly folded OMPs. The protein is then inserted into the membrane either by DegP or by another mechanism that could involve the BAM complex (Fig. 1.16) (Knowles *et al.*, 2009). Interestingly, *Neisseria meningitidis* lacks both BamB and LPS, instead having a lipooligosaccharide (LOS) surface. It has therefore been suggested that BamB may help coordinate LPS assembly or the assembly of LptD, a protein crucial for LPS assembly at the OM (Hagan *et al.*, 2011).

Lipooligosaccharides are the major glycolipids expressed on mucosal Gram-negative bacteria, including members of the genera *Neisseria*, *Haemophilus*, and *Bordetella*. Lipooligosaccharide is analogous to LPS, sharing similar lipid A structures with an identical array of functional activities as those in LPS. However, LOSs lack O-antigen units. The structure and role of LPS is discussed more thoroughly in section (1.2.3.7).

1.4.3.2 The LolABCDE complex (lipoprotein assembly)

The majority of lipoproteins are located on the inner side of the OM , and are transported there by a specific lipoprotein OM localisation pathway (Lol) that is composed of five proteins; LolA, B, C, D and E (Fig. 1.17). LolCDE is the IM transmembrane protein complex, while LolA is the periplasmic chaperone and LolB is the OM receptor.

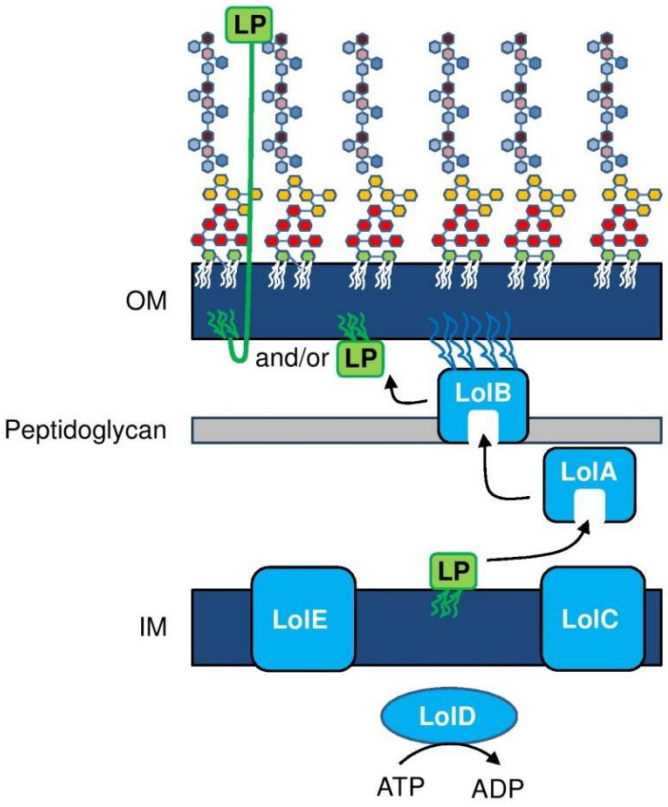


Fig. 1.17 The LolABCDE machinery for lipoprotein assembly.
The complex LolCDE forms an ABC transporter which releases OM specific lipoproteins (LP) from the IM. ATP hydrolysis by LolD transfers energy to LolC and LolE, which in turn promotes a conformational change in LolA allowing a complex between the lipoprotein and the periplasmic chaperone LolA to be formed. LolA interacts with the OM receptor LolB, transferring the protein to LolB and subsequently to the OM (Narita *et al.*, 2004). This figure was adapted from Buddelmeijer *et al.* (2015).

The mature lipoprotein first interacts with LolE, and is then transferred to LolC upon Adenosine triphosphate (ATP) hydrolysis by LolD. The lipoprotein is transferred from LolC to LolA upon their interaction. The contact between LolA and the lipoprotein is weaker than the interaction between LolB and the lipoprotein, driving transfer from LolA to LolB and preventing back-transfer to LolCE (Okuda & Tokuda, 2009; Taniguchi *et al.*, 2010). LolC and LolE share a similar membrane topology and their periplasmic domains share sequence identity with each other as well as with LolB (Okuda & Tokuda, 2009). LolA interacts specifically with LolC and not with LolE.

1.4.3.3 LPS transport and assembly machinery (LptABCDEFG complex)

Lipopolysaccarhide is typically organised into three structural domains: lipid A, a core oligosaccharide, and a highly variable O-antigen region, consisting of repeating oligosaccharide units (Polissi & Sperandeo, 2014) (Fig. 1.18; inset).

Biosynthesis of core-lipid A occurs at the inner leaflet of the IM and is flipped across the IM by the ABC transporter MsbA. O-antigen repeat units are synthesised separately on a lipid carrier by enzymes encoded by the *rfb* cluster in the cytoplasm. Once complete, the O-antigen is flipped across to the periplasmic face of the IM and covalently linked to core-lipid A, to form mature LPS. This ligation effectively terminates O-antigen synthesis (Whitfield, 1995).

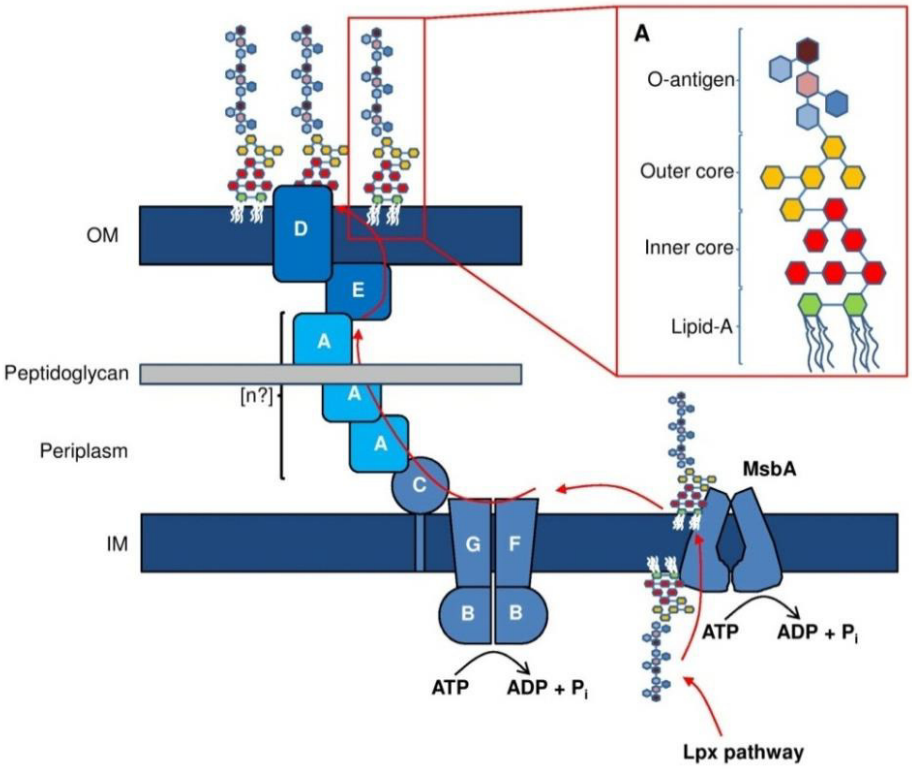


Fig. 1.18 The LptABCDEFG machinery for LPS transport.
LPS is synthesised in the cytoplasm by the Lpx pathway, flipped across the IM by the IMP MsbA, and subsequently transported across the periplasm by the chaperone LptA. The LPS is transported across the OM by the complex LptDE. The sequence of events is highlighted by red arrows. The composition of an assembled LPS molecule **(A)** is shown inset (although, described in greater detail in section 1.2.3.7). This diagram was adapted from Polissi and Sperandeo (2014).

Transport of mature LPS to the OM requires the LPS transport (Lpt) machinery (Polissi & Sperandeo, 2014). This protein complex consists of seven different proteins spanning the entirety of the cell envelope. The Lpt machinery can be divided into three subassemblies; LptBFGC, LptA and LptDE, which are located at the IM, periplasm and OM respectively. LptBFG constitutes an IM ABC transporter that is associated to an atypical subunit, the IM protein LptC, whose function as of yet remains undetermined. LptB is the ATP binding domain of this transporter, while LptF and LptG represent the transmembrane subunits. The integral ß-barrel protein LptD and lipoprotein LptE form a complex in the OM, and are required for the final stage of LPS assembly at the cell surface (Fig.

1.18). Recently, it has been shown in *E. coli* that LptE is required for the correct insertion of LptD into the OM. The periplasmic protein LptA is responsible for connecting the IM to LptDE OM components, chaperoning LPS across the periplasm (Fig. 1.18) (Dong *et al.*, 2014; Polissi & Sperandeo, 2014).
The mechanism of LPS insertion into the OM in *E. coli* has recently been elucidated (Gu *et al.*, 2015). It is suggested that one LptC molecule, four LptA molecules, and one N-terminal domain of LptD may form a ~360° rotation slide, allowing LPS transport across the periplasm. The complex LptBFG extracts LPS from the IM, before transferring it to LptC. LptC then delivers the LPS molecule to LptA. LptC, LptA, and the N-terminal domain of LptD transport LPS through their hydrophobic cores, for which the required energy is provided by LptB. Once in the N-terminal domain of LptD, the lipid A component of LPS is delivered to the OM through a intramembrane hole, while LptE allows the O-antigen to pass through the barrel of LptD. Insertion of lipid A into the OM prompts the lumen gate to open, allowing for entry of the core oligosaccharide into the barrel of LptD. A rearrangement of the ß-strands of LptD then allows the core oligosaccharide to slide to the surface of the bacteria.

1.4.3.4 Outer membrane proteins (OMPs)

Due to the broad range of ecological niches that bacteria inhabit, proteins are continually turned over to ensure the right complement of proteins are primed for action. Recent studies have shown the OM to be highly organised, and spatially aware, with newly synthesised OMPs being primarily focused to the centre of the bacterial cell and old OMPs displaced to the poles of growing cells allowing new OMPs to take their place (Rassam *et al.*, 2015). The functions of OMPs are diverse, and allow the bacterial cell to survive a variety of situations.

In *E. coli*, OMPs have been classified into six categories based on their functions. These included (1) general porins, (2) passive transporters, (3) active transporters, (4) enzymes, (5) defensive proteins and (6) structural proteins (Tamm *et al.*, 2004). In *P. multocida*, a similar categorisation scheme was instigated whereby six different groups were suggested; (1) structural proteins, (2) transport proteins, (3) binding proteins, (4) adhesins, (5) protein assembly machines and (6) membrane –associated enzymes (Hatfaludi *et al.*, 2010).

1.4.3.4.1 OM Biogenesis and integrity

Proteins involved in maintenance of the stability of the cell include OmpA and Brauns lipoprotein (Lpp), amongst others. The major OM protein OmpA is one of the best studied and most well characterised OMPs. OmpA is a multifunctional proteins, with roles in maintaining the structure of the OM and cell shape (Sonntag *et al.*, 1978), acts as a porin (Sugawara & Nikaido, 1992), bacteriophage receptor (Morona *et al.*, 1985), an adhesin/invasin (Torres & Kaper, 2003), in immune evasion for macrophage survival (Sukumaran *et al.*, 2003), and in biofilm formation (Orme *et al.*, 2006). OmpA was first purified in 1977 from *E.coli* and shown to have a molecular mass of approximately 33 kDa, although the variation in molecular mass due to the heat modifiable properties of the protein is now well documented (Beher *et al.*, 1980; Chai & Foulds, 1977). OmpA forms an eight stranded ß-barrel in the OM of Gram-negative bacteria, with four surface-exposed loops and three periplasmic turns (Pautsch & Schulz, 2000). The surface-exposed loops have recently been shown to extrude beyond the capsule in *M. haemolytica* (Hounsome *et al.*, 2011), and have had roles in host adaptation elucidated (Davies & Lee, 2004). Molecular mass heterogeneity has been observed among bovine and ovine *M. haemolytica* isolates correlating with the host species of origin (Davies & Donachie, 1996). Due to the presence of OmpA in such a high copy number (Koebnik *et al.*, 2000) and its location at the OM, the numerous functions elicited are not surprising.

The most abundant lipoprotein in Gram-negative bacteria is Braun's lipoprotein (Lpp). This protein anchors the OM to the peptidoglycan, interacting with Pal (a peptidoglycan associated protein) together contributing to the integrity of the cell wall (Kovacs-Simon *et al.*, 2011).

1.4.3.4.2 Transport and receptor

(a) Porins (Non-specific channels)

Porins are by definition, non-specific transmembrane ß-barrel structures forming channels across the bacterial membrane (Nikaido, 2003). Porins allow the diffusion of hydrophilic molecules trough aqueous pores (<600 Da) and show no particular substrate specificity, despite some selectivity for cations or anions (Koebnik *et al.*, 2000; Nikaido, 2003). The first porin structure to be resolved

was a general porin from *Rhodobacter capsulatus* in 1991, soon followed by OmpF and PhoE (OmpE) from *E. coli*. These structures revealed that porins form homotrimers in the OM, with each monomer a ß-barrel containing 16 membrane spanning antiparallel strands. Unlike the other loops, the third loop is not exposed at the cell surface, instead folding back into the barrel, forming a constriction zone and contributing significantly to the permeability properties of the pore. The second loop bends over the wall of the neighbouring barrel subunit, playing a role in trimer stabilisation (Nikaido, 2003). Porins are able to transition between open or closed states depending on the transmembrane potential across the OM (Koebnik *et al.*, 2000).

The major porins OmpF (Fig. 1.19) and OmpC can be altered by several environmental factors including osmolarity, temperature, pH, nutrient availability, aeration and various toxins (Forst *et al.*, 1989; Liu & Ferenci, 2001; Matsubara *et al.*, 2000; Pratt *et al.*, 1996). These porins are under the control of a two component system; the histidine kinase EnvZ, and the response regulator OmpR. Outer membrane protein F forms a larger pore than OmpC, and so OmpC is preferentially expressed at lower osmolarity than OmpF. The larger size of OmpF is attributed to improving efficiency of nutrient uptake in a nutritionally poor environment (Pratt *et al.*, 1996; Yoshida *et al.*, 2006).

The OMP typing scheme described previously (1.2.2.3.3) relies on two of the major porins, OmpF and OmpC, in combination with OmpA. Variations in molecular mass of these porins in different strains allow them to be used in characterisation of isolates (Davies, 1991c).

(a) Substrate specific channels

The OM also contains pores which are selective for certain substrates. The best characterised example of this is the maltose channel, LamB. LamB is present as a trimer similar to that of the non-specific porins, with each monomer a 47 kDa ß-barrel. While OmpF and OmpC contain 16 antiparallel ß-strands, LamB has 18. The pore size of LamB is also much more constrictive than the general porins, at between 0.5 and 0.6 nm in diameter (Schirmer *et al.*, 1995). Another example of a substrate specific channel is the nucleoside channel, Tsx. While this channel has specific nucleoside binding sites, it has also been shown to act as a

receptor for colicin K, bacteriophage T6 and a number of other lytic T-phages (Fsihisdli *et al.*, 1993).

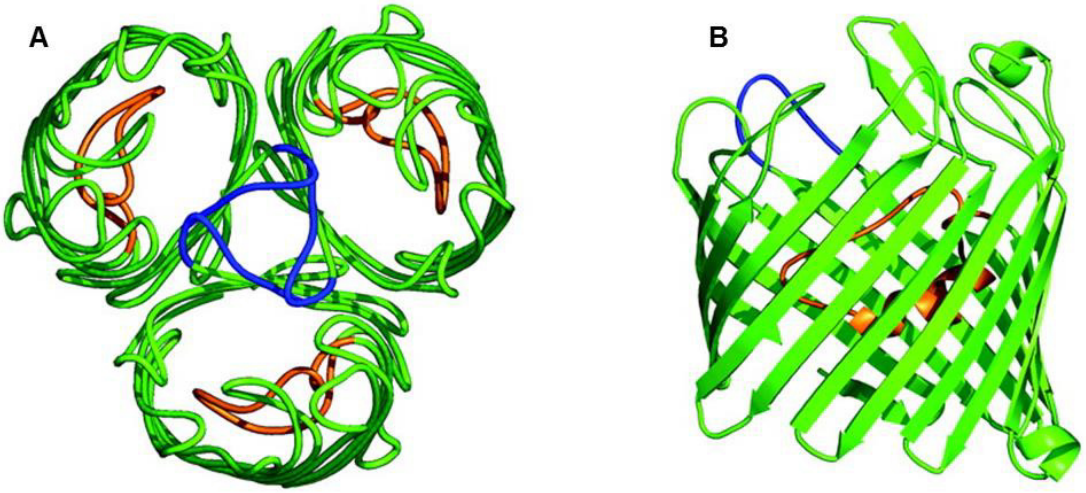


Fig. 1.19 Crystal structure of the major outer membrane protein OmpF.
The structure **(A)** is formed by three interacting subunits, known as a trimer. The monomer **(B)** structure of OmpF is very similar to OmpC, although the pore size of OmpF is slightly larger (1.2 nm compared to 1.1 nm of OmpC). Green structures represent β-strands while orange structures represent α-helices. The blue loop in **(B)** indicates the region of interaction between the three monomers, to form the trimer **(A)**. This figure was adapted from (Nikaido, 2003), in which the presented diagrams are based on Protein databank (PDB) file 2OMF, solved to 2.4Å.

In *Pseudomonas*, there are three well-characterised specific channels, namely OprB, OprP, and OprD. The specific channel OprB is the closest homolog of LamB, although while LamB is induced by maltose, OprB is induced by glucose as the sole carbon source. OprD was initially identified as a protein that was lost when *P. aeruginosa* isolates became resistant to the broad-spectrum antibiotic imipenem. It has been demonstrated that OprD is a specific porin that binds basic amino acids, dipeptides containing a basic residue and imipenem and related zwitterionic carbapenems. OprD is the closest *P. aeruginosa* homolog of the *E. coli* nonspecific porin OmpF. OprP is induced under conditions of low phosphate (< 0.15mM) and is involved in the high-affinity, phosphate-starvation inducible transport system (PTS).

(b) Iron acquisition

Iron is one of the most essential nutrients for bacteria; needed for growth, virulence and the catalysis of enzymatic reactions. Iron has been found ranging from ng L^{-1} concentrations in marine environments (Barbeau et al., 2001) to

mg L^{-1} concentrations in freshwater environments that receive acid mine drainage (Winterbourn et al., 2000). Unfortunately, although iron is one of the most abundant chemical elements it is generally in short supply to the bacteria, as it is usually bound to many host ligands including transferrin and haemoglobin, or rapidly oxidised from Fe^{2+} to Fe^{3+} (which is the more tightly bound form).

Bacteria have developed various ways to overcome this. Some bacteria are able to reduce the concentration of iron at the cell surface, creating a concentration gradient encouraging iron to be transported into the cytoplasm (Wooldridge & Williams, 1993). Many microbial pathogens have developed complex high-affinity iron transport systems (Fig. 1.20).

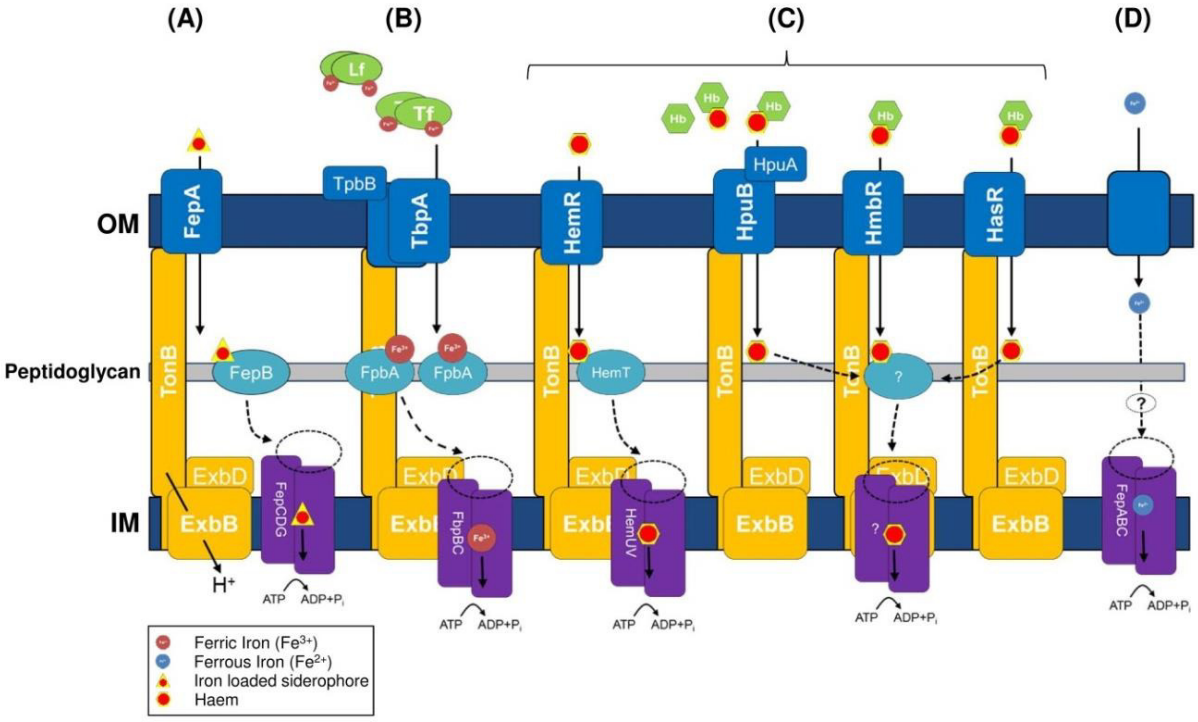


Fig. 1.20 Iron uptake systems in Gram-negative bacteria.
Bacteria rely on high affinity surface receptors to bind proteins that contain ferric iron (Fe^{3+}) such as siderophores or haem, and subsequently transport it to the periplasmic space utilising the Ton-system. *Escherichia coli* imports the siderophore enterobactin through the FepA receptor and the FepBCDG system **(A)**. Pathogenic bacteria can also use iron sources from the host; for example, in *Neisseria* sp. Fe^{3+} is removed from transferrin (Tf) and lactoferrin (Lf) at the OM before transport into the cell by the TbpAB–FbpABC system **(B)**. Haem is sequestered from haemoglobin (Hb) and haemoglobin–haptoglobin (Hb–Hp) by OMPs including HpuAB and HmbR. Haem can be also taken up directly, for example by the HemRTUV system in *Y. enterocolitica*, or be delivered by haemophores such as HasA in *S. marcensis* through HasR **(C)**. Under anaerobic conditions, ferrous iron (Fe^{2+}) can diffuse across OM porins, and is subsequently imported by energy-dependent systems such as the FeoABC **(D)**. This figure was adapted from (Faraldo-Gómez & Sansom, 2003).

In Gram-negative bacteria, iron transport systems are supplied with energy through the Ton system; comprising the integral membrane protein ExbB and the membrane anchored periplasmic proteins ExbD and TonB. TonB undergoes a conformational change, driven by the proton motive force (PMF), allowing a mechanical interaction with the OM receptors. Three Ton-B dependent siderophore receptors have been described in *E. coli* – FhuA, FepA and FecA, mediating the uptake of ferrichrome, enterobactin and ferric citrate, respectively.

As an enteric pathogen, *Y. ruckeri* primarily resides in the host gut where iron is in extremely limited supply. Previously, four OMPs were identified to be upregulated in the presence of a low iron environment (Davies, 1991b), although no further work has been conducted to identify these proteins fully. The expression of iron-regulated OMPs was not found to be an important factor in determining virulence in *Y.* ruckeri, as there was no variation according to biotype, serotype or OMP-type. Therefore it is suggested that the genes coding for these iron-regulated OMPs are most likely located on the chromosomes, as they were preserved amongst strains (Davies, 1991b). However, the hemolysin YhlA was shown to be significantly upregulated under low iron conditions (1.2.3.3) (Fernández *et al.*, 2007a).

Another Ton-B dependent receptor described in the literature, with no direct roles in iron acquisition, is the vitamin B12 transport protein, BtuB. BtuB has also been shown to act as a colicin receptor in *E. coli* (Nikaido, 2003).

(c) Secretion systems

In order to successfully infect and colonise host cells, Gram-negative bacteria have evolved complex mechanisms for the transport of proteins to the host cell cytoplasm. Secretion systems play important roles in mediating processes such as proteolysis, haemolysis, cytotoxicity and phosphorylation or dephosphorylation of host cell proteins (Beeckman & Vanrompay, 2010). Protein secretion by Gram-negative bacteria is necessary for numerous physiological processes such as adhesion, invasion, toxicity, degradation/hydrolysis, movement, provision of nutrients, cell to cell communication, detoxification of the environment and cell wall biosynthesis, as well as protein quality control (Francis, 2011).

Often, general secretory pathway (Sec) or twin-arginine translocation (Tat) systems transport unfolded and/or folded proteins first to the periplasmic space, where they are processed (for example, via the addition of disulphide bridges) in order to obtain their final three-dimensional structure, before being transported across the OM. Sec-dependent export of proteins across the IM is seen universally across all living organisms, with the pathway being studied most extensively in *E. coli*. A key component of this system being the SecB molecular chaperone, which is involved in maintaining a newly synthesised pre-polypeptide in an unfolded state; SecB binds to it as it leaves the ribosome, and pilots it to the membrane spanning SecYEG translocon, interacting specifically with SecA (an ATPase that is involved in energising the initial translocation of proteins through the translocation channel). All sec-dependent substrates have an N-terminal secretion signal, which is cleaved during transport through the IM (Fig. 1.21A) (Francis, 2011).

Independent of the sec pathway, the Tat dependent pathway is responsible for the secretion of fully folded proteins (Fig. 1.21B). These proteins are synthesised as precursors with an N-terminal signal sequence containing characteristic twin-Arg residues.

The core components of this secretion mechanism are the integral membrane proteins TatA, TatB and TatC. The human pathogenic strains of *Yersinia* are known to possess a functional Tat system, with components having different roles in virulence, suggesting a customisation to different *Yersinia* (Francis, 2011).

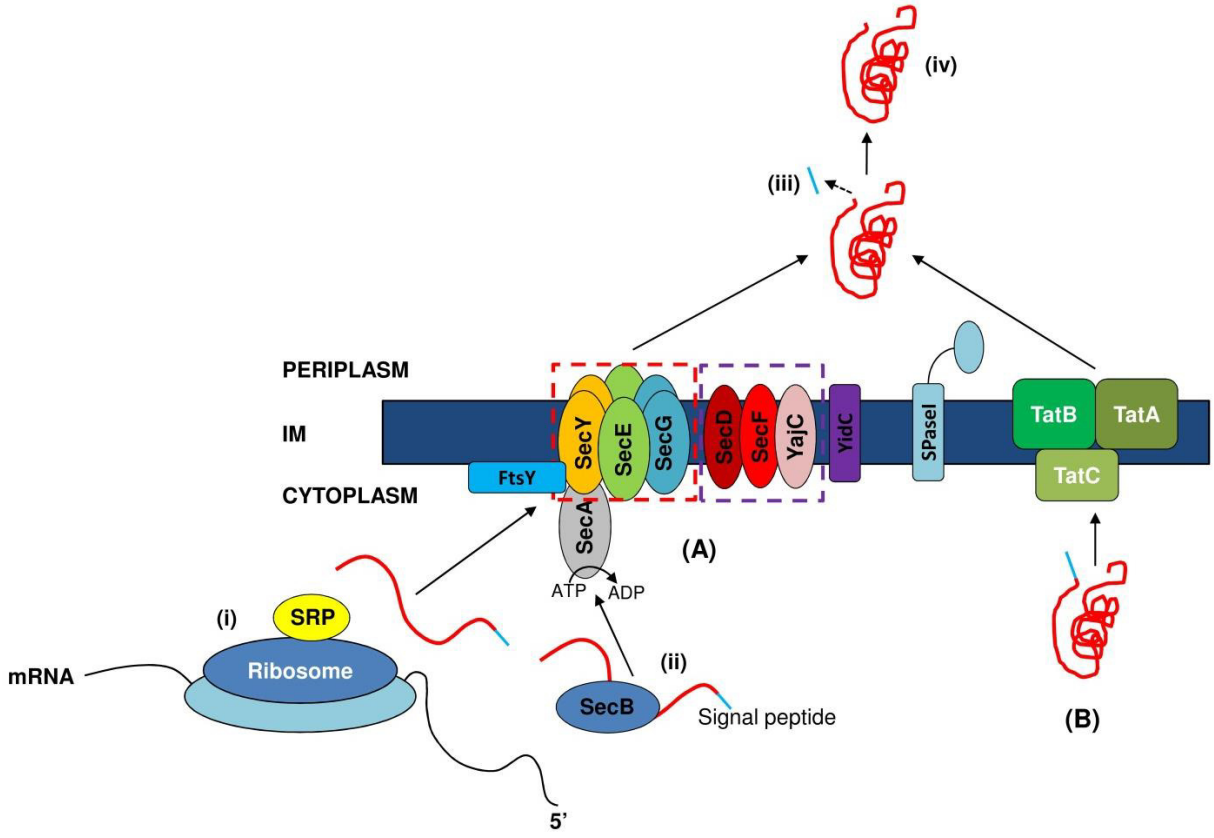


Fig. 1.21 Transport of proteins across the IM by the Sec and Tat pathways.
Preproteins are routed from the site of synthesis (cytoplasm) to the membrane by the secretion **(A; Sec)** or the twin-arginine translocation **(B; Tat)** pathway. The Sec system consists of a membrane-embedded, protein-conducting channel comprising three proteins (SecY, SecE and SecG), which form the Sec translocase and a peripherally associated, ATP-driven motor protein, SecA. SecD, SecF and YajC form the translocon-associated complex and YidC is involved in membrane protein integration and folding. In the Sec pathway, the transport of unfolded proteins occurs either post- or co-translationally. In post-translational transport **(i)**, the fully synthesised preprotein detaches from the ribosome and is directed to the Sec translocase, with the help of a chaperone, SecB. In co-translational targeting **(ii)**, the signal recognition particle (SRP) binds to the signal sequence of the protein while it emerges from the ribosome, with the complex of SRP/ribosome/protein chain targeted to the Sec translocase with the aid of the SRP receptor (FtsY). SecA accepts the unfolded proteins and threads them through the transmembrane channel.
The Tat system however, is dedicated to the transport of folded proteins, consisting of three membrane-integrated subunits (TatA, TatB and TatC). TatC and TatB form a complex involved in recognition of Tat signal sequences and their insertion into the membrane, while TatA mediates the actual translocation event. During or shortly after translocation of the proteins, the hydrophobic signal peptide of most nonlipoproteins is cleaved off **(iii)** by type I signal peptidase, resulting in the release of the mature protein **(iv)**.

Transport across the OM is an energy requiring process. In Gram-negative bacteria there is no ATP present at the OM, therefore secretion systems must be self-energised or harness energy at the cytoplasmic side of the IM.

Type I secretion systems (T1SS) or ATP-binding cassette (ABC) protein exporters are employed by a wide range of Gram-negative bacteria for the secretion of toxins, proteases and lipases. Secretion of α-hemolysin by *E. coli* represents the prototypical type I exporter. The type I pathway is sec-independent and

secretes proteins directly from the cytoplasm across the OM without a periplasmic intermediate. The type I export apparatus consists of three proteins: an IM ABC exporter, an IM-anchored protein that spans the periplasm, termed a membrane fusion protein (MFP), and an OMP. The OM channel protein involved in α-hemolysin export, TolC, assembles as a trimeric complex in the OM and is predicted to consist of a porin-like ß-barrel membrane domain with a carboxy-terminal hydrophilic region that extends into the periplasm. The periplasmic MFP component also assembles as a trimer and interacts with both the OMP and ABC exporter, and is believed to either span or anchor to the IM. The bulk of the MFP spans the periplasm to contact the OM, facilitating substrate secretion forming a closed channel (Fig. 1.22D) (Thanassi & Hultgren, 2000).

The Type II secretion system (T2SS) pathway is sec-dependent. This pathway is responsible for secretion of extracellular enzymes and toxins by a wide variety of Gram-negative bacteria. The type II secretory pathway is closely related to biogenesis of type IV pili. These are long, polarised pili important for the virulence of many bacterial pathogens. Secretion across the OM by the T2SS is quite complex, requiring between 12 and 16 accessory proteins, collectively referred to as the secreton. The T2SS secreton includes two OM components: GspD (an integral OMP – secretin superfamily) and GspS (small lipoprotein required for proper targeting and insertion of GspD into the OM) (Fig. 1.22A).

Type III secretion system (TTSS) pathways are capable of translocating anti-host factors into the cytosol of target eukaryotic cells, and have been identified in a number of animal and plant pathogens. Secretion of *Yersinia* outer proteins (Yops) by *Yersinia* spp. represents the prototypical TTSS. Translocation of Yops into the cytosol of target cells disrupts signalling pathways, allowing the bacteria to evade host defence mechanisms. Type III secretion is sec-independent, may take place without a periplasmic intermediate, and requires approximately 20 secretion components that assemble into a large structure spanning both bacterial membranes and possibly the host cell membrane as well.

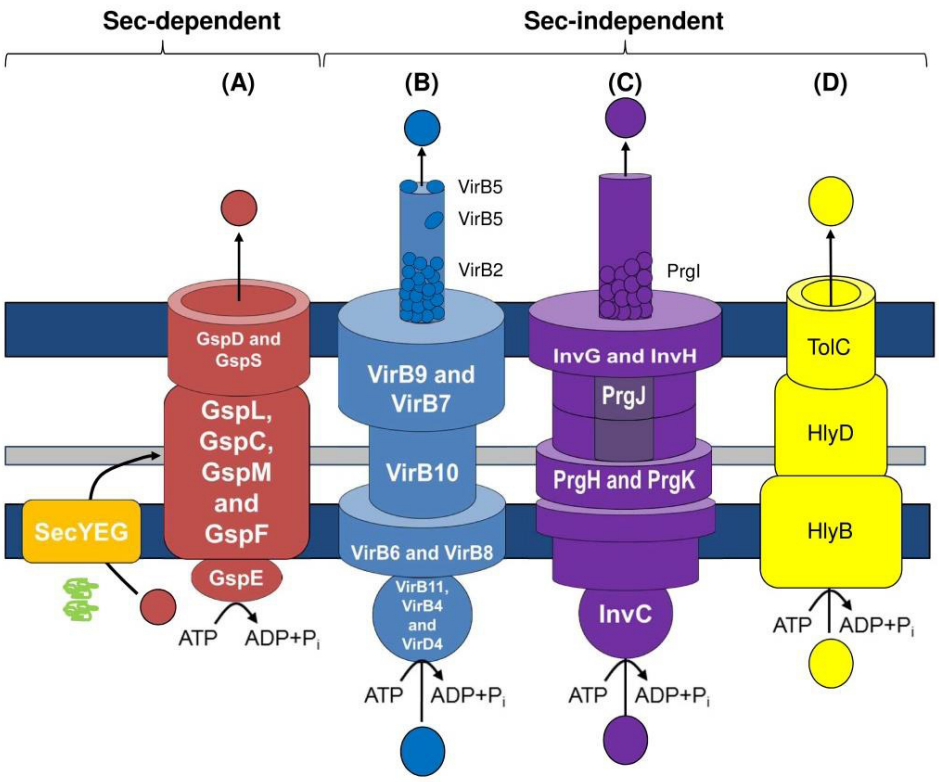


Fig. 1.22 Bacterial secretion systems.
Bacterial secretion systems are either Sec-dependent **(A)** or Sec-independent **(B, C and D)**. T1SSs **(D;** haemolysin secretion system OF *E. coli***)**, consist of an ABC transporter or a proton-antiporter (HlyB), an adaptor protein that bridges the IM and OM (HlyD), and an OMP (TolC). T2SSs **(A;** general secretion pathway (Gsp) system**)** use a two-step mechanism for translocation. Initially, the effector protein is translocated through the IM. Once in the periplasm, it is translocated through the OM. The T2SS translocon consists of 12–16 protein components that are found in both bacterial membranes, the cytoplasm and the periplasm. TTSSs **(C;** *S. enterica* subsp. enterica serovar Typhimurium system**)** mediate a single-step secretion mechanism and are used by many plant and animal pathogens, including *Salmonella* spp., *Shigella* spp., *Yersinia* spp., *E. coli* and *P. aeruginosa*. The TTSS illustrated uses the invasion (Inv) and Prg proteins. TTSSs are genetically, structurally and functionally related to bacterial flagella. T4SSs **(B;** *Agrobacterium tumefaciens* VirB/D system**)** are found in Gram-negative and Gram-positive bacteria and secrete a wide range of substrates, from single proteins to protein–protein and protein–DNA complexes. This figure was adapted from (Costa *et al.*, 2015; Fronzes *et al.*, 2009).

TTSS share one homologous component with the T2SS pathway: secretion across the OM requires a member of the secretin family of channel-forming proteins. The *Yersinia* and *Salmonella typhimurium* secretins YscC and InvG, respectively, have been imaged as ring-shaped complexes with large central pores. Proper targeting of InvG to the OM requires the InvH lipoprotein, analogous to the role of GspS in the TTSS (Büttner & Bonas, 2002; Costa *et al.*, 2015) (Fig. 1.22C).

The majority of TTSS components are thought to localise to the IM and are closely related to components of the flagellar basal body. The flagellar basal body spans the IM and OM, providing an anchor for the flagellar filament. The cytoplasmic face of the basal body is thought to contain the machinery that

drives secretion and assembly of flagella via ATP hydrolysis. This machinery exports flagellin monomers through a central channel within the basal body and filament for assembly at the distal end of the growing flagellum. Interestingly, the flagellar apparatus was recently shown to function as a protein secretion system. The *S. typhimurium* TTSS structure has been shown to be similar to the flagellar basal body, with a hollow projection extending out from the bacterial surface, termed the needle (Costa *et al.*, 2015).

A TTSS different from the human pathogenic *Yersinia* species has been identified in *Y. ruckeri* (Gunasena *et al.*, 2003). Sequence analysis of the type III ATPase gene revealed a significant homology with the chromosomally encoded *Ysa* (*Yersinia* secretion apparatus) of *Y. enterocolitica* biovar 1B, suggesting the presence of a *Yersinia*-like type III secretion system in *Y. ruckeri* (Gunasena *et al.*, 2003).

Traditionally, type IV secretion (T4SS) is used for the spread of plasmids, involving the translocation of a DNA transfer intermediate. Previous research has suggested adaptations of T4SS's in Gram-negative bacteria where virulence factors are passed to the host (Méndez *et al.*, 2009). The prototypic T4SS is the VirB complex of *Agrobacterium tumefaciens,* which uses this system to introduce the T-DNA portion of the Ti plasmid into the plant host (Cascales & Christie, 2003) (Fig. 1.22B). Bacteria including *Helicobacter pylori*, *Bordetella pertussis*, and *Legionella pneumophila*, utilise T4SS's for the translocation of numerous effector proteins into their eukaryotic host. A T4SS within *Y. ruckeri* has been identified, involving a cluster of eight genes referred to as the *tra* operon; *traHIJKCLMN* (Méndez *et al.*, 2009). This operon is absent from human pathogenic *Yersinia* strains, although was recovered from *Y. ruckeri* isolated in different geographic locations during different outbreaks. This is suggestive of an important role in virulence. Expression of the operon was temperature-dependent as has been seen for other virulence factors, being promoted at 18°C and less so at 28°C (Fernández *et al.*, 2002; Méndez *et al.*, 2009). T4SSs of intracellular pathogens such as *L. pneumophila*, *Brucella* sp., and *Bartonella* sp. have been shown to play roles in the transfer of different effector molecules into target cells, with involvement in survival within macrophages or red blood cells (Méndez *et al.*, 2009).

1.4.3.4.3 Adherence

Successful establishment of infection by bacterial pathogens requires adhesion to host cells, colonisation of tissues and invasion. Bacteria use a variety of tools for cell adhesion including pili and fimbriae, autotransporters, and OMPs (Pizarro-Cerdá & Cossart, 2006).

(a) Pili and fimbriae

Pili and fimbriae are bacterial adhesive surface appendages contributing to bacterial adherence to host tissues, colonisation and biofilm formation. These are hair-like structures protruding from the cell surface that are critical for bacterial virulence. Present in both Gram-negative and Gram-positive bacteria, pili are involved in many processes such as conjugation, adherence, twitching motility and biofilm formation. Pathogenic *N. gonorrhoeae* adheres specifically to the human cervical or urethral epithelium; enterotoxigenic *E. coli* adhere to the mucosal epithelium of the intestine; the M-protein and associated fimbriae of *Streptococcus pyogenes* are involved in adherence and to resistance to engulfment by phagocytes. In *E. coli*, a specialised pilus termed the F or sex pilus provides stability to mating bacteria during the process of conjugation. This is the transfer of DNA from one bacterial cell to another, as a way to increase genetic variability (Allen *et al.*, 2012; Pizarro-Cerdá & Cossart, 2006; Wall & Kaiser, 1999).

(b) Type V secretion (the autotransporter pathway)

The type V secretion system (T5SS; Fig. 1.23) uniquely spans only a single membrane, driving its own secretion through the OM – deriving the term autotransporter. The autotransporter family contains more than 700 proteins, and is the largest family of Gram-negative extracellular proteins (Cotter *et al.*, 2005). Secreting mainly virulence factors, the T5SS also participates in cell-to-cell adhesion and biofilm formation. The T5SS apparatus exists in three distinct forms; (1) classical autotransporters (type Va), (2) two-partner secretion (TPS) systems (type Vb) and (3) trimeric autotransporters (type Vc). All autotransporters are composed of a secreted domain, called the 'passenger' domain, that is either semi-unfolded or fully unfolded in the periplasm, and a

transmembrane ß-barrel domain, which will traverse the OM. The passenger domain secretes through the pore of the translocator domain, and contains the effector function of the protein (Henderson *et al.*, 2004).

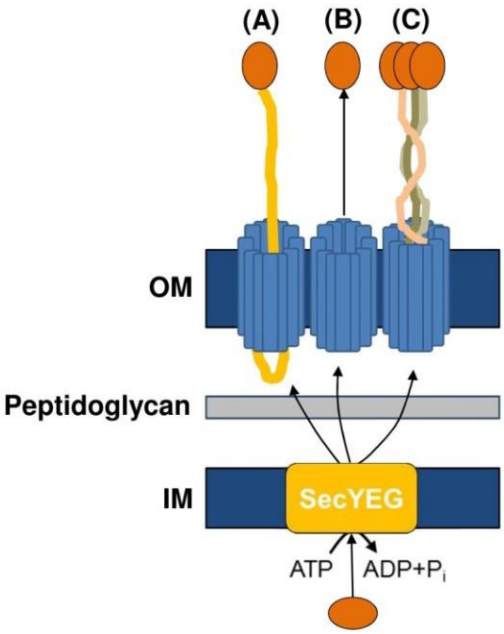


Fig. 1.23 The Type V secretion system (T5SS).
The T5SS apparatus exists in three distinct forms; **(A)** classical autotransporters (type Va), **(B)** two-partner secretion systems (type Vb) and **(C)** trimeric autotransporters (type Vc).

In *Y. pestis*, 10 conventional autotransporters have been identified. Seven of these proteins are exported to the OM while three are dependent on the OM protease Pla (YapA, YapG and YapE) and are secreted into the supernatant (Lawrenz *et al.*, 2009; Lenz *et al.*, 2011). The autotransporter YapE contributes to virulence in *Y. pestis,* although the gene encoding YapE is not present in *Y. ruckeri* (Lawrenz *et al.*, 2013). No conventional autotransporter or type V secretion system present in *Y. ruckeri* has yet been described in the literature.

In the TPS pathway (type Vb), the passenger and translocation domain are secreted as two separate proteins, namely TpsA and TpsB. The best characterised examples of type Vb systems are that of the *B. pertussis* filamentous haemagglutinin protein, FhaB (Willems *et al.*, 1994), the haemolytic ShlA/B system of *S. marcescens* (Braun *et al.*, 1993) and the high molecular weight adhesins HMW1 and HMW2 from *Haemophilus influenzae* (St. Geme & Yeo, 2009). Filamentous haemagglutinin is transported through the OM by its TpsB protein FhaC, ShlA by ShlB, whereas HMW1B and HMW2B are responsible for the transport of HMW1 and HMW2, respectively (Leo *et al.*, 2012).

The prototypical members of the trimeric autotransporters (type Vc) are the *H. influenzae* adhesin Hia and the YadA adhesin of *Y. enterocolitica* (Cotter *et al.*, 2005). The first to be discovered was YadA (*Yersinia* adhesin protein A), which is found in all three human pathogenic *Yersinia* species (Fig. 1.24). This autotransporter is encoded on a 70 kb plasmid, which is absent from *Y. ruckeri* (discussed in section 1.2.3.8). Trimeric autotransporters contain six distinct domains: (1) an N-terminal signal sequence, (2) head, (3) neck, (4) stalk, (5) linker region and (6) a C-terminal region consisting of four ß-strands (Henderson *et al.*, 2004). The head, neck and stalk make up the passenger domain, while the linker and C-terminal regions comprise the translocator domain (Hoiczyk *et al.*, 2000).

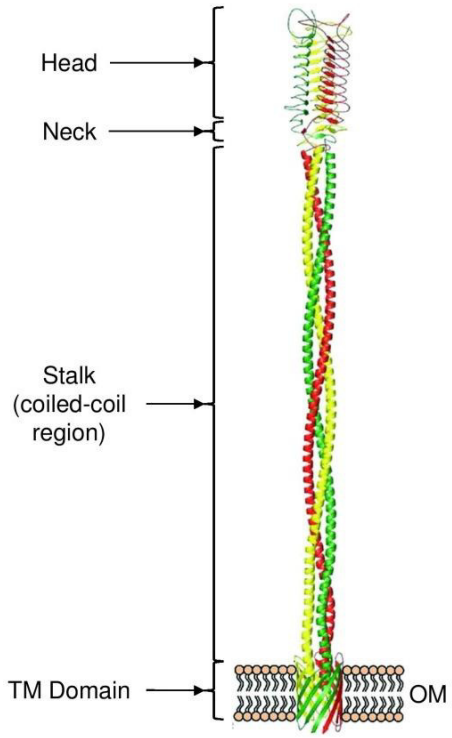


Fig. 1.24 Trimeric autotransporter YadA of *Y. enterocolitica*.
The autotransporter shown in the figure, YadA, is a trimer. Each monomer is coloured green, yellow and red respectively. YadA encompasses six domains; The N-terminal signal sequence, head, neck and stalk (coiled coil structure) protrude from the bacterial surface, while the linker region and C-terminal domains anchor the structure to the OM. This figure was adapted from (Singh *et al.*, 2012).

The T5SS requires the SecYEG translocon to transfer an unfolded autotransporter polypeptide through the IM to the periplasm. After transport through the IM, the autotransporter polypeptide is maintained in an unfolded state by the dedicated periplasmic chaperone, SurA. Recently, work has suggested that the BAM complex in combination with the translocation and assembly module (TAM)

promotes efficient secretion of autotransporters, suggesting that autotransporters are not autonomous secretion systems but instead involve larger assemblies (Webb *et al.*, 2012).

(c) OmpX/Ail

Outer membrane protein X (OmpX) from *E. coli* belongs to a family of highly conserved bacterial proteins promoting bacterial adhesion. Other membrane proteins homologous with OmpX have been described in *Klebsiella pneumonia* (OmpK17), *S. typhimurium* (PagC and Rck) and *Y. enterocolitica* (Ail). The protein Ail has been found in all three human pathogenic *Yersinia*, although not in *Y. ruckeri* (Kawula *et al.*, 1996).

(d) PqiB/Mam7

The adhesion factor PqiB/MAM7 primes the bacteria for immediate attachment when a host cells is encountered. PqiB/MAM7 is capable of binding both fibronectin and phosphatidic acid on the host cell surface, and both interactions are required for bacterial adhesion to host cells (Krachler & Orth, 2011; Krachler *et al.*, 2011; Lim *et al.*, 2014).

1.4.3.4.4 Enzymatic activity

Proteins with enzymatic activities at the OM include OmpLA and phospholipase A1. OmpLA is an OM phospholipase of *E. coli* that was the first OM enzyme to have its structure solved. OmpLA is predicted to have roles in the hydrolysis of phospholipids on the external surface of the OM (Hatfaludi *et al.*, 2010). Phospholipase A1 has been shown to be involved in colicin release from *E. coli*, and has been implicated in the virulence of *Campylobacter* and *Helicobacter* strains, by destabilising the membrane bilayer and allowing the release of colicins or virulence factors (Koebnik *et al.*, 2000).

1.4.3.4.5 Hypothetical

A hypothetical protein is a protein whose existence has been predicted, but for which there is no experimental evidence that it is expressed *in vivo*.

1.5 Bioinformatics and proteomics

Once a protein has been synthesised in the cytoplasm it is targeted to a specific subcellular location. The protein remains cytoplasmic, is transported to the IM, the periplasm, or the OM, or is secreted as an extracellular protein. Proteins are sorted to their specific compartment based on several properties, several of which have been well characterised – including the transmembrane α-helical secondary structures that are found in integral IMPs, and the ß-barrel structures common to OMPs. The improved availability of genomes has made the process of bioinformatic prediction much more accessible.

1.5.1 Bioinformatic prediction tools for the discrimination of OMPs

The identification of OMPs is an initial step in exploring the proteins involved in virulence and host pathogen interactions. Bioinformatic predictors have been used successfully in several Gram-negative bacterial species for the identification of OMPs; however prediction disagreements are often common. A bioinformatic prediction scheme developed and utilised by E-Komon *et al.* (2011) moved to address this, employing a system comprising the consensus prediction of ten bioinformatic programmes; predicting likely structure, subcellular location and lipoprotein likelihood. Having applied the workflow to *Pasteurella multocida*, the authors were able to confidently predict 98 OMPs from an avian strain (Pm70; GenBank: AE004439), and 107 OMPs from a porcine strain (3480; Project ID: 32177).

1.5.1.1 Prediction of transmembrane β-barrel proteins

Several online tools for the bioinformatic prediction of transmembrane ß-barrel proteins have been developed, including TMB-Hunt (Garrow *et al.*, 2005), TMBETADISC-RBF (Ou *et al.*, 2008), MCMBB (Bagos *et al.*, 2004) and BOMP (Berven *et al.*, 2004). Briefly, these programs rely on the structural differences between α-helices and ß-barrels. Structurally, α-helices found at the IM consist of 15-20 mostly non-polar membrane spanning amino acids, while OM ß-barrels have several characteristics which the programs can identify. These include; amino acids alternating between polar and non-polar, a flanking layer of aromatic amino acid residues, and a C-terminal signature sequence.

1.5.1.2 Prediction of lipoproteins

Several online tools for the bioinformatic prediction of lipoproteins have been developed, including LipoP (Juncker *et al.*, 2003), and LIPO (Berven *et al.*, 2006). These prediction tools rely on the differentiation of signal sequences between lipoproteins and other proteins. In the periplasm, lipoprotein signal sequences are cleaved by an enzyme called signal peptidase II (SPaseII), while other proteins are cleaved by signal peptidase I (SPaseI) (Juncker *et al.*, 2003). Outer membrane lipoproteins are further distinguishable from IM lipoproteins upon examination of the amino acids. Lipoproteins containing an aspartate residue at position +2 from the cleavage site are positioned at the IM, while the remainder are translocated to the OM (Yamaguchi *et al.*, 1988).

1.5.1.3 Prediction of subcellular localisation

Several online tools for the bioinformatic prediction of subcellular localisation have been developed, including PSORTb (Yu *et al.*, 2010), CELLO (Yu *et al.*, 2006) and SOSUI-GramN (Imai *et al.*, 2008). These programs essentially rely on the same parameters discussed previously. However, they have the advantage of distinguishing multiple location sites of predicted proteins.

1.5.2 Outer membrane proteomics

Studying the variety of proteins that an organism possesses has several advantages over studying the individual roles of specific proteins. While examining an individual protein can elucidate the specialist function of that protein, identifying the entire proteome of a system provides an insight into how that system functions, while simultaneously removing the bias associated with a targeted approach (Aebersold & Mann, 2003).

Mass spectrometry (MS) is an analytical tool used for measuring the molecular mass of a sample by determining the mass-to-charge ratio (m/z) and abundance of gas-phase ions of a sample mixture. It is the method chosen for the identification of proteins in simple or complex mixtures, usually analysing whole-cell proteomes, or the composition of proteins in a particular sub compartment (sub-proteome). With the ever increasing availability of gene and genome sequences, the improvements in MS technologies and the advancements in

computing capabilities, MS-based proteomics has become a powerful tool. A mass spectrometer consists of an ion source, a mass analyser that measures the *m/z* of the ionised peptide mixture, and the detector that recognises the number of ions at each *m/z* value. Protein determination is carried out by a search engine program (MASCOT), using peptide sequences to identify proteins from a sequence database. A comprehensive review of the principles and applications of mass-spectrometry-based proteomics for the analysis of complex protein samples is provided by (Aebersold & Mann, 2003).

In order to examine the OM sub-proteome, the OM must first be isolated from the remainder of the proteins in the cell, and suitably purified to reduce contamination. Utilising properties of the different sub-cellular compartments, the OM can be separated by several different techniques, of which selective detergent solubilisation, spheroplasting and sucrose density gradient centrifugation are the most commonly used.

The study of the OM proteome in a bacterial species can reveal important information about the function, pathogenicity and adaptation to environmental conditions of OMPs (Kao *et al.*, 2009; Kumar *et al.*, 2010; Williams *et al.*, 2007; Zhang *et al.*, 2014). A variety of techniques have been utilised; namely, gel-based and gel-free proteomic technologies, combined with matrix assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI) or liquid chromatography tandem mass spectrometry (LC-MS/MS) (Boyce *et al.*, 2006; Carranza *et al.*, 2010; Hauberg *et al.*, 2010). Several studies have utilised gel-free proteomic techniques for the identification of proteins in OM vesicles, including *V. cholerae* (Altindis *et al.*, 2014), *C. jejuni* (Jang *et al.*, 2014), and *P. aeruginosa* (Choi *et al.*, 2011). However, identification of OMPs solely is less common and this technique is usually reserved for whole cells. Gel-based identification of OMPs is more frequent, with studies including *E. tarda* (Wang *et al.*, 2012), *Coxiella burnetti* (Flores-Ramirez *et al.*, 2014), *P. multocida* (Boyce *et al.*, 2006) and *M. haemolytica* (Ayalew *et al.*, 2010) identifying 19, 22, 35 and 55 proteins, respectively. Few studies have combined both approaches, although those that have done so reported overlapping coverage, and importantly, proteins identified solely by individual techniques (Aistleitner *et al.*, 2015; Cordwell *et al.*, 2008; Gesslbauer *et al.*, 2012; Kulkarni *et al.*, 2014; Tanca *et al.*, 2013; Zhou *et al.*, 2012). The results of these studies suggest that

in order to gain a true insight into the proteins encompassed in the bacterial OM, complementary techniques are necessary. Examples of studies utilising gel-free, gel-based or a combination of techniques are included in Table 1.7, highlighting the disparity between the different techniques and the range of proteins and techniques that have been employed.

One previous study has examined the OM proteome of *Y. ruckeri,* comparing OMP expression under stationary growth with that of cells immobilised in an artificial biofilm. However, this study had limitations; only a single isolate was analysed by gel-based proteomic approaches, with only differences in protein abundance noted. Also, no *Y. ruckeri* genomes were available on NCBI at the time of publication meaning the data was searched inefficiently (Coquet *et al.*, 2005). In total, 55 proteins were identified in this study (from the genomes of *Y. pestis*, *E. coli* and *S. typhimurium*).

1.5.3 **Membrane protein identification**

1.5.3.1 Gel-based approach

The gel-based (or GelC) approach of protein identification relies on the initial separation of proteins by molecular mass (one-dimensional [1D] SDS-PAGE) and charge (two-dimensional [2D] SDS-PAGE) prior to in-gel trypsin digestion and MS. Gel electrophoresis is a widely used technique allowing for the identification of proteins based on molecular mass, of which SDS-PAGE is the most commonly used (Garfin, 2003). Polyacrylamide gel polymerisation occurs by copolymerization of acrylamide and N,N'-methylenebisacrylamide (bis). The pore size of the gel can be altered by changing the concentration of polyacrylamide. Gel formation is catalysed by the addition of ammonium persulphate (APS) and N,N,N'N'-tetramethylethylene diamine (TEMED). During protein electrophoresis, proteins migrate through the gel according to both their size and electrical charge, with smaller polypeptides migrating at a faster rate. The molecular mass of proteins is determined by comparing the positions of proteins to those of a known molecular standard. Proteins of similar molecular weights (MWs) are difficult to distinguish from one another, without the use of 2D SDS-PAGE, or subsequent proteomic identification.

Table 1.7 Example studies that have utilised proteomics to identify OMPs.
WC - Whole Cell; SP - Surface Proteome; CE - Cell Envelope; OMP - Outer membrane protein; IMP - Inner membrane protein; CL - Cell lysate; GF – gel-free; GB – gel-based.

Year	Author	Bacterial species	Method	Mass Spectrometry method	Examined	Proteins identified
2015	(Kanaujia *et al.*, 2015)	*Y. enterolitica*	GB	2D - MALDI TOF MS/MS	WCs	28 differentially expressed proteins
2015	(Buján *et al.*, 2015)	*E. tarda*	GB	2D - MALDI TOF MS/MS	WCs	103 differentially expressed proteins
2015	(Wurpel *et al.*, 2015)	*E. coli*	GF	nanoLC-MS/MS	OMV	47 OMPs
2015	(Aistleitner *et al.*, 2015)	*Chlamydiae sp.*	GF+GB	MS/MS Linear Trap Quadropole	WC	approx. 1000 proteins (4 methods)
2014	(Watson *et al.*, 2014)	*Lawsonia intracellularis*	GB	LC-ESI-MS/MS	OMP	19 OMPs
2014	(Kulkarni *et al.*, 2014)	*P. syringae*	GF+GB	MALDI-TOF/TOF and ESI-MS/MS	OMVs	429 proteins (42 OMPs)
2014	(Flores-Ramirez *et al.*, 2014)	*Coxiella burnetii*	GB	LC-MS/MS analysis	WC & OMP	39 OMPs
2014	(Bai *et al.*, 2014)	*S. enterica*	GB	LC-MS/MS analysis	OMV	167 proteins
2014	(Folsom *et al.*, 2014)	*E. coli*	GB	Q-TOF LC/MS	CE	approx. 100 proteins
2014	(Altindis *et al.*, 2014)	*V. cholerae*	GF	LC-MS/MS analysis	OMV	90 proteins
2014	(Martorana *et al.*, 2014)	*E. coli*	GB	2DC-MS/MS	CE	123 proteins
2014	(Pförtner *et al.*, 2014)	*S. aureus*	GF	LC-MS/MS analysis	WC	1302 proteins
2014	(Jang *et al.*, 2014)	*Campylobacer jejuni*	GF	NanoLC-MS/MS analysis	OMV	134 proteins
2014	(Stekhoven *et al.*, 2014)	*Bartonella henselae*	GF	LTQ-ICR-FT-Ultra MS	WC	74 OMPs
2014	(Zhang *et al.*, 2014)	*Haemophilus parasuis*	GB	2D - MALDI TOF MS/MS	CE	Undisclosed (only 19 listed in paper)
2014	(Nabu *et al.*, 2014)	*N. gonorrhoeae*	GB	MALDI-TOF-MS	WC	58 proteins
2014	(Diep *et al.*, 2014)	*S. aureus*	GF	LC-MS/MS	SP	342 (Total)
2013	(Tanca *et al.*, 2013)	*E. coli*	GF+GB	LC-MS/MS analysis	WC	approx. 1000 proteins (4 methods)
2013	(Pförtner *et al.*, 2013)	*S. aureus*	GF	Nano-LC-LTQ-Orbitrap Ms	WC	600 proteins
2012	(Gu *et al.*, 2012)	*Y. enterolitica*	GB	2D - MALDI TOF MS/MS	WC+OMP	Approx. 30 proteins
2012	(Jagannadham *et al.*, 2011)	*P. syringae*	GB	1D&2D MALDI & LC-ESI MS/MS	Mem. Proteins	Approx. 100 proteins
2012	(Gesslbauer *et al.*, 2012)	*Borrelia sp.*	GF+GB	(LC)-MS/MS or (2D)-LC-MS/MS	Mem. Proteins	286 proteins
2012	(Zhou *et al.*, 2012)	*Y. pestis*	GF+GB	2D-LC-LTQ-FT	WC	1926 proteins
2012	(Wang *et al.*, 2012)	*E. tarda*	GB	2-DE, MALDI-TOF/TOF	OMP	19 by PMF
2011	(Papadioti *et al.*, 2011)	*Coxiella burnetti*	GB	1D&2D MALDI TOF/TOF	OMP	22 OMPs
2011	(Choi *et al.*, 2011)	*P. aeruginosa*	GF	LC-MS/MS	OMV	64 OMPs
2011	(Sherry *et al.*, 2011)	*S. enterica*	GB	LC-ESI/MS/MS Sequencing	SP	49 proteins
2011	(Zhang *et al.*, 2011a)	*Actinobacillus pleropneumoniae*	GB	MALDI-TOF MS	CL	42 immunoreactive proteins
2011	(Chiu *et al.*, 2011)	*V. parahaemolyticus*	GB	2D-LC-ESI-Q-TOF MS/MS	CE	Expression changes
2011	(Shevchuk *et al.*, 2011)	*Legionella Pneumophilla*	GB	2D-MALDI TOF MS/MS	WC	157 (out of 439 spots used)
2010	(Chooneea *et al.*, 2010)	*S. enterica*	GB	2D-LC-MS/MS analysis	OMVs	54 OMPs

Year	Author	Bacterial species	Method	Mass Spectrometry method	Examined	Proteins identified
2010	(Heinz *et al.*, 2010)	*Protochlamydia amoebophila*	GB	1D & 2D MALDI-TOF/TOF & nanoLC-ESI	OMP	38 OMPs
2010	(Ayalew *et al.*, 2010)	*M. hemolytica*	GB	2DE MALDI-TOF & LC-MS/MS	OMP	55 proteins
2010	(Carranza *et al.*, 2010)	*Cronobacter turicensis*	GF	2-D-LC-MALDI-TOF/TOF-MS	WC & SP	891 proteins
2010	(Kruh *et al.*, 2010)	*Mycobacterium tuberculosis*	GF	LC-MS/MS analysis	WC	545 proteins
2009	(Kao *et al.*, 2009)	*V. anguillarum*	GB	LC-nano ESI-Q-TOF MS/MS	OMP	6 proteins
2010	(Kumar *et al.*, 2010)	*E. tarda*	GB	2D-MALDI-TOF MS	OMP	21 proteins
2009	(Wheeler, 2009)	*P. multocida*	GB	MALDI-TOF-MS	OMP	39 proteins
2008	(Cordwell *et al.*, 2008)	*C. jejuni*	GF	2D-MALDI-TOF MS & 2DLC MS/MS	SP	77 Proteins & 432 proteins
2008	(Liu *et al.*, 2008)	*F. columnare*	GB	RP-HPLC MS/MS	OMP	36 proteins
2008	(Kim *et al.*, 2008)	*H. pylori*	GB	Q-TOF MS/MS	OMP	42 OMPs
2008	(Nakatsuji *et al.*, 2008)	*Propionibacterium acnes*	GF	Nano-LC-MS/MS analysis	WC	Not discussed
2008	(Al Dahouk *et al.*, 2008)	*Brucella suis*	GB	2D DIGE / MALDI-TOF MS	WC	168 intramacrophagic
2007	(Anaya *et al.*, 2007)	*Porphyromonas gingivalis*	GF	LC-MS/MS	WC	1223 / 987 (replicate 1 / replicate 2)
2007	(Lee *et al.*, 2007)	*E. coli*	GB	Nano-LC-ESI-MS/MS	OMV	141 proteins
2007	(Williams *et al.*, 2007)	*N. meningitidis*	GB	LC-MS/MS	OMP & OMV	236 proteins
2006	(Connolly *et al.*, 2006)	*Brucella abortus*	GF+GB	MALDI-TOF MS & LC-MS/MS	CE	163 proteins
2006	(Boyce *et al.*, 2006)	*P. multocida*	GB	2D-MALDI-TOF MS & 1D nanoLC-MS/MS	OMP	35 OMPs
2006	(Xu *et al.*, 2006)	*E. coli*	GB	2D-MALDI-TOF MS	OMP	30 OMPs
2005	(Coquet *et al.*, 2005)	*Y. ruckeri*	GB	2D-nanoLC/MS/MS	OMP	55 OMPs
2005	(Ying *et al.*, 2005)	*Shigella flexneri*	GB	MALDI-TOF-MS (PMF)	OMP	55 OMPs
2005	(Xu *et al.*, 2005)	*Vibrio alginolyticus*	GB	MALDI-TOF/MS (PMF) & NanoESI-MS/MS	OMP	138 proteins (not all OMPs)
2005	(Ebanks *et al.*, 2004)	*A. salmonicida*	GB	2-DE, LC-MS/MS	OMP	76 proteins (not all OMPs)
2005	(Peng *et al.*, 2005)	*P. aeruginosa*	GB	2D-MALDI-TOF MS	OMP	15 proteins
2004	(Lai *et al.*, 2004)	*E. coli*	GB	2D-MALDI-TOF MS	Mem. Proteins	173 proteins (41 OMPs)
2004	(Rhomberg *et al.*, 2004)	*Bartonella henselae*	GB	1 and 2D-MALDI-TOF-MS	OMP	53 OMPs
2004	(Baik *et al.*, 2004)	*H. pylori*	GB	MALDI-TOF-MS	OMP	62 proteins (not all OMPs)
2003	(Guina *et al.*, 2003)	*P. aeruginosa*	GB	2D-ESI-MS/MS + GPF	WC	1331 proteins
2001	(Phadke *et al.*, 2001)	*Caulobacter crescentus*	GB	2D-MALDI-TOF-MS	OMP	54 proteins (41 OMPs)
2000	(Molloy *et al.*, 2000)	*E. coli*	GB	MALDI-TOF-MS	OMP	25 OMPs

SDS-PAGE allows for efficient solubilisation of integral membrane proteins, through the use of SDS as a detergent - SDS can solubilise proteins but will also denature proteins (Cordwell, 2006).

Identification by peptide mass fingerprint (PMF) is common, although when more than one protein is likely to be present in the sample, MS/MS analysis is often employed. One-dimensional SDS-PAGE is particularly successful in the separation of membrane proteins as they are readily solubilised in SDS sample buffer (Cordwell, 2006). Many samples can be separated on a single gel, allowing for an initial comparison in protein expression and profile appearance between samples. However, as the separation is only in a single dimension, proteins of a similar molecular mass can be difficult to distinguish from one another. Relative quantification of proteins within a single gel band can however be estimated using exponentially modified protein abundance index (emPAI) score (Ishihama *et al.*, 2005). One-dimensional SDS-PAGE has several advantages over 2-D SDS-PAGE however, as although samples are run through two dimensions using the latter, this approach does not lend itself well to the separation of membrane proteins. Many membrane proteins are not solubilised in the non-detergent isoelectric focusing sample buffer due to their hydrophobic nature, with many that are precipitating at their isoelectric point (Wu & Yates, 2003). Severe protein losses have also been reported with 2-D SDS-PAGE as membrane proteins often don't enter the second dimension of separation (Adessi *et al.*, 1997; Pasquali *et al.*, 1997) One-dimensional SDS-PAGE is also much quicker, less labour intensive, and more reproducible (Görg *et al.*, 2004).

1.5.3.2 Gel-free approach

Gel-free proteomic approaches are particularly useful for the identification of proteins within complex mixtures. The technique requires an additional separation step (usually liquid chromatography [LC]) prior to mass analysis due to the complex nature of the samples. Proteins are enzymatically digested without prior gel-based separation and the resultant peptides separated by one or more steps of LC. Peptides are ionised by electrospray ionisation (ESI) before entering the mass spectrometer for analysis. Selected peaks are then further fragmented by energetic collisions with gas, and the resultant spectra utilised for protein identification.

1.6 Research objectives

The initial objective of this thesis was to examine recent isolates of *Y. ruckeri* recovered from Atlantic salmon in the UK, in order to compliment previous published work on isolates recovered from UK rainbow trout. As *Y. ruckeri* is primarily a disease of rainbow trout, the transition to becoming a major pathogen of Atlantic salmon in the UK is not fully understood, and the diversity of isolates infecting Atlantic salmon has not been previously described. An extensive and unique strain collection gathered since 2001, comprising 135 isolates collected from Atlantic salmon and rainbow trout were characterised using previously described phenotypic approaches (biotype, serotype and OMP-type), in combination with phylogenetic approaches. This has ultimately led us to understand the diversity of isolates of *Y. ruckeri* infecting both Atlantic salmon and rainbow trout. From this, a representative group of eight reference isolates were selected for further proteomic analysis.

A comprehensive bioinformatic prediction study was undertaken, whereby the genomes (available on NCBI) of four isolates recovered from Atlantic salmon and rainbow trout were analysed in order to identify OM located proteins. Within the genomes of a further 16 representative isolates of *Y. ruckeri* representing different host species (Atlantic salmon, rainbow trout and European eel) and phenotypic characteristics, the aim was to identify variation in both amino acid and nucleotide sequence of the predicted OMPs.

Variation in OM proteome composition between isolates recovered from both Atlantic salmon and rainbow trout were observed. In order to gain as deep an insight as possible, isolates were examined using a combination of proteomic techniques combining both gel-based and gel-free technologies. Subsequently, changes in OM proteome composition when isolates were grown under conditions which mimicked both environmental and *in vivo* growth were identified.

Finally, a small challenge study was performed examining the virulence of newly identified isolates towards both Atlantic salmon and rainbow trout.

This work will not only provide an understanding of the adaptations *Y. ruckeri* makes for survival in different environments, but will also provide an improved

understanding of host specificity within this bacterium. This will provide useful information for the aquaculture and scientific industries as to the diversity of isolates causing infection in different host species, and an in-depth analysis of the proteins that may be responsible for the infection.

Combining all we can learn about a bacterium using phenotypic, proteomic and genetic approaches is of critical importance when trying to provide preventative measures to avoid future infections.

Chapter 2 The diversity of *Y. ruckeri* isolates recovered from Atlantic salmon and rainbow trout, and the emergence of a novel O-serotype

2.1 Introduction

Yersinia ruckeri has primarily been a disease of rainbow trout since its first isolation in the Hagerman valley, Idaho, in 1956 (Ross & Klontz, 1965; Ross *et al.*, 1966). The bacterium has since spread worldwide, with the recovery of isolates from the USA (Arias *et al.*, 2007), Chile (Bastardo *et al.*, 2011), China (Wang *et al.*, 2015), Canada (Stevenson & Daly, 1982) and many parts of Europe (Davies, 1990; Wheeler *et al.*, 2009). *Yersinia ruckeri* has been recovered from many different species; however, the predominant recovery of isolates from Atlantic salmon and rainbow trout is of serious economic concern. The characterisation of *Y. ruckeri* was conducted using a variety of phenotypic methods discussed in detail in section 1.2.2.3 of the Introduction. Isolates of serotype O1 are responsible for most outbreaks of disease in rainbow trout aquaculture worldwide (Davies, 1990; Haig *et al.*, 2011; Stevenson, 1984; Wheeler *et al.*, 2009).

There have been very few characterisation studies of *Y. ruckeri* isolates recovered from Atlantic salmon to date, as ERM is primarily a disease of rainbow trout (Bastardo *et al.*, 2011, 2012b; Davies & Frerichs, 1989; Davies, 1991a). Little information is available as to the diversity of isolates currently circulating in Atlantic salmon populations and the emergence of strains pathogenic towards this species could become an increasing problem for the aquaculture industry. The growing economic importance of farmed Atlantic salmon means a more detailed understanding of strain diversity associated with this species will be important for the future control of infection.

The aim of this study was to compare and contrast the level of diversity within a unique and extensive strain collection comprising 135 isolates of *Y. ruckeri* recovered from Atlantic salmon and rainbow trout based on biotype, serotype and OMP-type. In particular, a major aim was to assess whether isolates associated with Atlantic salmon and rainbow trout represented host specific sub-populations. The majority of isolates (109 out of 135) were recovered from

Atlantic salmon. A smaller number of contemporary rainbow trout isolates (26 out of 135) were included to supplement previous studies and draw more recent comparisons between species.

A major finding of this study was the emergence of a novel serotype (O1/O5) which represented a unique LPS-type and was almost exclusive in the Atlantic salmon population. In order to determine the virulence of this isolate in comparison to an isolate of serotype O1 (of known virulence) in Atlantic salmon and rainbow trout, i.p. injection and bath challenge studies were conducted.

The work carried out in this chapter provides important information relevant to the aquaculture industry by revealing the phenotypic diversity of isolates that are the major cause of infection in Atlantic salmon and rainbow trout, in the UK. The identification of a novel serotype is particularly relevant as the current vaccination strategies are based on serotype O1 whole-cell preparations (section 1.3.7.2.1 of the Introduction). Therefore, these include the important O1 antigenic determinant. This work also formed the basis of the selection of representative isolates for the subsequent proteomic and sequence studies (subsequent Chapters).

2.2 Materials and Methods

2.2.1 Bacterial isolates

One-hundred and thirty five field isolates of *Y. ruckeri* recovered exclusively from Atlantic salmon (109 isolates) and rainbow trout (26 isolates) from UK fish farms are included in this study. All field isolates were collected between 2001 and 2014, and were supplied by MH (Marine Harvest), FVG (Fish Vet Group), DF (Dawnfresh Seafoods) or TW (Ridgeway biologicals). Representative reference isolates (RD6, 22, 28, 64, 124, 128, 138, 154, 158 and 168) used in this study were collected previously, and have been used extensively in other studies (Davies & Frerichs, 1989; Davies, 1991a, b, c, d; Haig *et al.*, 2011; Wheeler *et al.*, 2009). The properties of all isolates used are presented as Table 2.1.

2.2.2 Storage of isolates

Upon arrival of isolates to the laboratory, bacteria were subsequently subcultured on TSA (Oxoid, UK) and grown at 22°C for 48 h. Cultures were resuspended in 1 ml of 50% glycerol (v/v) in tryptone soya broth (TSB; Oxoid, UK) and stored at -80°C.

2.2.3 Growth conditions

2.2.3.1 Revival of glycerol stocks

Glycerol stock tubes were initially thawed at room temperature before 15 µl of bacterial suspension were streaked onto TSA and grown at 22°C for 48 h.

2.2.3.2 Revival of freeze dried stocks

Vials containing freeze dried stocks were carefully opened to avoid aerosolising the bacteria, before resuspension in 100 µl TSB within the vial. Fifteen microlitres of the suspension was then streaked onto TSA, and grown at 22°C for 48 h.

Table 2.1 Properties of isolates of *Y. ruckeri* included in this study.

Designation		Source				Phenotype			
Lab	**Previous**	**Location**	**Host**	**Date**	**Water**	**Biotype**	**Serotype**	**OMP**	**Sub-type**
RD6	-	Scotland	Rainbow trout	Pre 1990	-	2	O1	1b	ND
RD22	Bacter 3	Finland	Atlantic salmon	Pre 1990	-	1	O1	4	ND
RD28	BA2	U.K	Rainbow trout	Pre 1990	-	1	O5	2a	ND
RD64	F53.1/82	West Germany	Rainbow trout	1982	-	1	O2	2a	ND
RD124	851014	Denmark	Rainbow trout	Pre 1990	-	1	O1	3a	ND
RD128	-	U.K	Rainbow trout	Pre 1990	-	2	O1	1a	ND
RD138	MLSP.06	France	Rainbow trout	Pre 1990	-	1	O1	3b	ND
RD154	85/09/1578	Norway	Atlantic salmon	Pre 1990	-	2	O1	2c	ND
RD158	85.9919	Australia	Atlantic salmon	Pre 1990	-	1	O1	2a	ND
RD168	RS3;BC74	U.S.A	Chinook salmon	Pre 1990	-	1	O2	2b	ND
RD330	Y.ruckeri 231.99	Speyside	Rainbow trout	20/07/2001	-	2	O1	1a	1a.1
RD332	Y.ruckeri 194.99	Galloway	Rainbow trout	20/07/2001	-	2	O1	1a	1a.1
RD336	Y.ruckeri 229.99	Speyside	Atlantic salmon	20/07/2001	-	1	O2	2a	2a.1
RD338	Y.ruckeri 38501	Naver	Atlantic salmon	20/07/2001	-	1	O1	2a	2a.1
RD340	Y.ruckeri 107-02	U.K	Atlantic salmon	25/03/2002	Freshwater	1	O8	3a	3a.1
RD342	Y.ruckeri 171.3/1A	U.K	Atlantic salmon	27/03/2002	Freshwater	1	O2	2a	2a.1
RD344	Y.ruckeri 7262	Invergarry	Atlantic salmon	07/04/2003	Seawater	1	O2	2c	2c.1
RD346	Y.ruckeri 7268	Inverpolly	Atlantic salmon	07/04/2003	Seawater	1	O2	2c	2c.1
RD348	FVG 230/04 IG13	U.K	Atlantic salmon	21/05/2004	Freshwater	1	O2	2a	2a.1
RD350	FVG 253/04 B5	U.K	Atlantic salmon	21/05/2004	Freshwater	1	O2	2a	2a.1
RD352	TW59/05	Scotland	Atlantic salmon	20/05/2005	-	1	O2	2a	2a.1
RD354	TW60/05	Scotland	Atlantic salmon	20/05/2005	-	1	O2	2a	2a.1
RD356	TW85/05	Company A, site 1, West Coast Scotland	Atlantic salmon	20/07/2005	-	1	O5	2c	2c.1
RD358	TW86/05	Company A, site 2 West Coast Scotland	Atlantic salmon	20/07/2005	-	1	O1	2c	2c.1
RD360	TW87/05	Scotland	Atlantic salmon	28/07/2005	-	1	O5	2c	2c.2
RD362	TW88/05	Scotland	Atlantic salmon	28/07/2005	-	1	O5	2c	2c.2
RD364	TW89/05	Scotland	Atlantic salmon	28/07/2005	-	1	O5	2c	2c.1
RD366	TW90/05	Scotland	Atlantic salmon	20/07/2005	-	1	O5	2c	2c.1
RD368	TW91/05	Scotland	Atlantic salmon	20/07/2005	-	1	O5	2c	2c.1
RD370	TW92/05	Company A, site 2 West Coast Scotland	Atlantic salmon	20/07/2005	-	1	O2	2c	2c.1
RD372	TW93/05	Scotland	Atlantic salmon	20/07/2005	-	1	O5	2c	2c.1
RD374	TW94/05	Scotland	Atlantic salmon	20/07/2005	-	1	O1	2c	2c.3
RD376	TW130/05	Scotland	Atlantic salmon	25/10/2005	-	1	O2	2a	2a.1

Designation		Source				Phenotype			
Lab	**Previous**	**Location**	**Host**	**Date**	**Water**	**Biotype**	**Serotype**	**OMP**	**Sub-type**
RD378	FVG 248/06	U.K	Atlantic salmon	06/05/2006	Freshwater	1	O2	2a	2a.1
RD380	FVG 249/06	U.K	Atlantic salmon	06/05/2006	Freshwater	1	O5	2c	2c.2
RD382	FVG 269/06	U.K	Atlantic salmon	06/05/2006	-	1	O1	3a	3a.5
RD384	TW 184/06	Company B, outer hebridies	Atlantic salmon	17/06/2006	-	1	O2	v	2a.1
RD386	Hagerman 06018	U.K	Rainbow trout	03/07/2007	-	1	O1	3a	3a.6
RD388	TW248/07	Company C, SW Eng	Rainbow trout	24/07/2007	-	2	O1	1a	1a.1
RD390	TW 249/07	Company C, SW Eng	Rainbow trout	24/07/2007	-	2	O1	1a	1a.2
RD392	TW 250/07	Company C, SW Eng	Rainbow trout	24/07/2007	-	2	O1	1a	1a.2
RD394	TW 251/07	Company C, SW Eng	Rainbow trout	24/07/2007	-	2	O1	1a	1a.2
RD396	TW273/07	Company D, S Eng	Rainbow trout	08/07/2007	-	2	O1	1a	1a.3
RD398	TW274/07	Company D, S Eng	Rainbow trout	08/07/2007	-	2	O1	1a	1a.1
RD400	TW277/07	Company D, S Eng	Rainbow trout	08/07/2007	-	2	O1	1a	1a.1
RD402	TW278/07	Company D, S Eng	Rainbow trout	08/07/2007	-	2	O1	1a	1a.1
RD404	TW292/07	Company E, Outer Hebrides	Atlantic salmon	20/08/2007	-	1	O8	3a	3a.1
RD406	H14A	Company F, site 1 West coast of Scotland	Atlantic salmon	15/04/2008	-	1	O5	2c	2c.1
RD408	H14C	Company F, site 1 West coast of Scotland	Atlantic salmon	15/04/2008	-	1	O5	2c	2c.1
RD410	H14C	Company F, site 1 West coast of Scotland	Atlantic salmon	15/04/2008	-	1	O2	2a	2a.1
RD412	164/08	Kinlochmoidart	Atlantic salmon	05/08/2008	Seawater	1	O5	2c	2c.1
RD414	TW87/O8	Company G, outer Hebridies	Atlantic salmon	08/07/2008	-	1	O5	3a	3a.1
RD416	TW88/08	Company A, site 3 west coast of Scotland	Atlantic salmon	08/07/2008	-	1	O2	2c	2c.1
RD418	TW109/08	Company G, outer Hebridies	Atlantic salmon	08/07/2008	-	1	O2	2a	2a.2
RD420	TW110/08	Company G, outer Hebridies	Atlantic salmon	08/07/2008	-	1	O8	3a	3a.1
RD422	TW138/08	Company E, Outer Hebrides	Atlantic salmon	30/09/2008	-	1	O8	3a	3a.1
RD424	TW139/08	Company E, Outer Hebrides	Atlantic salmon	30/09/2008	-	1	O8	3a	3a.5
RD426	TW140/08	Company E, Outer Hebrides	Atlantic salmon	30/09/2008	-	1	O8	3a	3a.2
RD428	TW141/08	Company E, Outer Hebrides	Atlantic salmon	30/09/2008	-	1	O8	3a	3a.2
RD430	TW142/08	Company E, Outer Hebrides	Atlantic salmon	30/09/2008	-	1	O8	3a	3a.2
RD432	TW143/08	Company E, Outer Hebrides	Atlantic salmon	30/09/2008	-	1	O8	3a	3a.2
RD434	TW157/08	Company E, Outer Hebrides	Atlantic salmon	30/09/2008	-	1	O8	3a	3a.2
RD436	Yersinia1	Outer Hebrides	Atlantic salmon	19/06/2009	-	1	O2	2a	2a.3
RD438	Yersinia2	Outer Hebrides	Atlantic salmon	19/06/2009	-	1	O8	3a	3a.2
RD440	Yersinia3	Outer Hebrides	Atlantic salmon	19/06/2009	-	1	O8	3a	3a.1
RD442	Yersinia4	Outer Hebrides	Atlantic salmon	19/06/2009	-	1	O8	3a	3a.1
RD444	Yersinia5	Outer Hebrides	Atlantic salmon	19/06/2009	-	1	O2	2a	2a.2

Designation		Source				Phenotype			
Lab	**Previous**	**Location**	**Host**	**Date**	**Water**	**Biotype**	**Serotype**	**OMP**	**Sub-type**
RD446	Yer1	Torridon	Atlantic salmon	07/09/2009	-	1	O8	3a	3a.1
RD448	Yer2	Torridon	Atlantic salmon	07/09/2009	-	1	O8	3a	3a.1
RD450	Yer3	Torridon	Atlantic salmon	07/09/2009	-	1	O8	3a	3a.1
RD452	Yer4	Torridon	Atlantic salmon	07/09/2009	-	1	O8	3a	3a.1
RD458	690-09	U.K	Atlantic salmon	23/11/2009	Freshwater	1	O8	3a	3a.3
RD460	768-09	U.K	Atlantic salmon	23/11/2009	Freshwater	1	O8	3a	3a.3
RD462	MT3492	U.K	Atlantic salmon	12/04/2009	-	1	O2	2a	2a.2
RD464	437/09	Loch Lochy	Atlantic salmon	12/04/2009	Freshwater	1	O2	2a	2a.2
RD466	466/09	Loch Seaforth	Atlantic salmon	17/12/2009	Seawater	1	O2	2a	2a.2
RD468	TW7/10	Company H, Central Scotland	Atlantic salmon	29/01/2010	-	1	O8	3a	3a.3
RD470	TW8/10	Company H, Central Scotland	Atlantic salmon	29/01/2010	-	1	O8	3a	3a.3
RD472	TW9/10	Company H, Central Scotland	Atlantic salmon	29/01/2010	-	1	O8	3a	3a.2
RD474	TW12/10	Company H, Central Scotland	Atlantic salmon	29/01/2010	-	1	O8	3a	3a.2
RD476	Glenfinnan 6.4.10	Loch Shiel	Atlantic salmon	24/04/2010	Seawater	1	O2	2a	2a.3
RD478	Glenfinnan 7.4.10	Loch Shiel	Atlantic salmon	24/04/2010	Seawater	1	O2	2a	2a.3
RD480	Lochy 6.4.10	Loch Lochy	Atlantic salmon	24/04/2010	Freshwater	1	O2	2a	2a.3
RD482	Lochy 7.4.10	Loch Lochy	Atlantic salmon	24/04/2010	Freshwater	1	O2	2a	2a.3
RD484	TW 81/10	Company A site 1, West Coast Scotland	Atlantic salmon	28/04/2010	-	1	O2	2a	2a.2
RD486	TW 82/10	Company A site 1, West Coast Scotland	Atlantic salmon	28/04/2010	-	1	O2	2a	2a.1
RD488	TW 83/10	Company A site 3, Central Scotland	Atlantic salmon	28/04/2010	-	1	O2	2a	2a.3
RD490	TW 84/10	Company A site 3, Central Scotland	Atlantic salmon	28/04/2010	-	1	O2	2a	2a.3
RD492	M1S	U.K	Atlantic salmon	22/07/2010	-	1	O8	3a	3a.5
RD494	M2	U.K	Atlantic salmon	22/07/2010	-	1	O8	3a	3a.2
RD496	M3	U.K	Atlantic salmon	22/07/2010	-	1	O8	3a	3a.2
RD498	M4S	U.K	Atlantic salmon	22/07/2010	-	1	O8	3a	3a.2
RD500	M5S	U.K	Atlantic salmon	22/07/2010	-	1	O8	3a	3a.2
RD502	FVG 0562/10	Fortwilliam	Atlantic salmon	23/08/2010	Freshwater	1	O8	3a	3a.1
RD504	Sub3	U.K	Atlantic salmon	24/09/2010	-	1	O8	3a	3a.2
RD506	Sub4	U.K	Atlantic salmon	24/09/2010	-	1	O8	3a	3a.2
RD508	Yruk1	Loch Lochy	Atlantic salmon	10/05/2010	-	1	O8	3a	3a.1
RD510	Yruk2	Loch Lochy	Atlantic salmon	10/05/2010	-	1	O8	3a	3a.2
RD512	Yruk3	Loch Lochy	Atlantic salmon	10/05/2010	-	1	O8	3a	3a.1
RD514	Yruk4	Loch Lochy	Atlantic salmon	10/05/2010	-	1	O2	2a	2a.2
RD516	FVG0735/10 (TW117/10)	Kenmure	Rainbow trout	04/11/2010	Freshwater	2	O8	1a	1a.1

Designation		Source				Phenotype			
Lab	**Previous**	**Location**	**Host**	**Date**	**Water**	**Biotype**	**Serotype**	**OMP**	**Sub-type**
RD518	FVG 0750/10	Fortwilliam	Rainbow trout	17/11/2010	Freshwater	2	O1	1a	1a.1
RD520	FVG 205 (T1-1)	Etive	Rainbow trout	26/04/2011	Brackish	2	O8	1a	1a.1
RD522	TW 86/11	Company I, E Scotland	Rainbow trout	21/07/2011	-	2	O1	1a	1a.3
RD524	TW87/11	Company I, E Scotland	Rainbow trout	21/07/2011	-	2	O1	1a	1a.1
RD526	THM1	SW Scotland	Rainbow trout	29/07/2011	-	2	O1	1a	1a.1
RD528	THM2	SW Scotland	Rainbow trout	29/07/2011	-	2	O1	1a	1a.1
RD530	SSF627	Scotland	Atlantic salmon	28/09/2011	-	1	O8	3a	3a.1
RD532	TW119/11	Scotland	Atlantic salmon	28/09/2011	-	1	O8	3a	3a.4
RD534	MIG523K1	Scotland	Atlantic salmon	28/09/2011	-	1	O8	3a	3a.4
RD536	Barvasy	U.K	Atlantic salmon	15/07/2004	Freshwater	1	O2	2a	2a.1
RD538	0179-FVG	Etive	Rainbow trout	04/04/2012	Brackish	2	O1	1a	1a.1
RD540	0375-FVG	Kenmure	Rainbow trout	28/09/2012	Freshwater	2	O1	1a	1a.1
RD542	464/12	Scotland	Atlantic salmon	08/10/2012	-	1	O2	2a	2a.1
RD544	0455-FVG	Braevallich	Rainbow trout	28/09/2012	Freshwater	1	O1	3a	3a.6
RD546	0621-FVG	Etive	Rainbow trout	28/09/2012	Brackish	2	O1	1a	1a.1
RD548	#340	Scotland	Atlantic salmon	25/10/2012	Freshwater	1	O2	2a	2a.1
RD550	P16F2	Scotland	Atlantic salmon	20/02/2013	-	1	O5	2a	2a.1
RD552	038/13	Scotland	Atlantic salmon	02/05/2013	Seawater	1	O2	2a	2a.1
RD554	022/13	Scotland	Atlantic salmon	02/05/2013	Seawater	1	O2	2a	2a.1
RD556	2013-0771	Tervine	Rainbow trout	08/10/2013	Freshwater	2	O1	1a	1a.1
RD558	2013-0871	Scotland	Atlantic salmon	08/10/2013	Freshwater	1	O5	3a	3a.2
RD560	MT3492	Scotland	Rainbow trout	08/10/2013	Freshwater	2	O1	1a	1a.1
RD562	2013-0694	Scotland	Rainbow trout	08/10/2013	Freshwater	2	O1	1a	1a.1
RD564	2013-0944	Scotland	Atlantic salmon	08/10/2013	Freshwater	1	O1	3a	3a.2
RD566	2013-1006	Trevine	Rainbow trout	08/10/2013	Freshwater	2	O1	1a	1a.1
RD568	RI14096	Company E, Outer Hebrides	Atlantic salmon	12/06/2014	-	1	O8	3a	3a.1
RD570	RI14165	Company H, Central Scotland	Atlantic salmon	15/08/2014	-	1	O8	3a	3a.1
RD572	RI14168	Company H, Central Scotland	Atlantic salmon	15/08/2014	-	1	O8	3a	3a.1
RD574	0728-FVG	Meavaig hatchery	Atlantic salmon	20/08/2014	Freshwater	1	O8	3a	3a.1
RD576	0812-FVG	Meavaig hatchery	Atlantic salmon	20/08/2014	Freshwater	1	O8	3a	3a.1
RD578	RI14264	Company F, site 3 West coast of Scotland	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD580	RI14265	Company F, site 3 West coast of Scotland	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD588	RI14269	Company F, site 3 West coast of Scotland	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD590	RI14270	Company F, site 3 West coast of Scotland	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1

Designation		Source				Phenotype			
Lab	**Previous**	**Location**	**Host**	**Date**	**Water**	**Biotype**	**Serotype**	**OMP**	**Sub-type**
RD592	RI14271	Company F, site 1 Hebrides	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD594	RI14272	Company F, site 1 Hebrides	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD596	RI14273	Company F, site 1 Hebrides	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD598	RI14274	Company F, site 1 Hebrides	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD600	RI14275	Company F, site 2 West coast of Scotland	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD602	RI14276	Company F, site 1 Hebrides	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD604	RI14277	Company F, site 1 Hebrides	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD606	RI14278	Company F, site 1 Hebrides	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD608	RI14279	Company F, site 1 Hebrides	Atlantic salmon	10/11/2014	-	1	O8	3a	3a.1
RD610	TW224/10	Company F, Central Scotland	Atlantic salmon	18/11/2014	-	1	O8	3a	3a.1

2.2.3.3 Preparation of liquid cultures

Liquid cultures were prepared by inoculating three or four single colonies of plate grown bacteria into 10 ml volumes of TSB and incubating overnight at 22°C (with shaking at 120 rpm). For the production of OMPs, three or four single colonies were resuspended in 3 ml TSB in a bijoux, before 400 µl were inoculated into 400 ml volumes of TSB in 2-litre Erlenmeyer flasks.

2.2.3.4 Aerobic growth conditions

Bacteria were grown aerobically in 400 ml of TSB in 2-litre Erlenmeyer flasks at 22°C with shaking at 120 rpm.

2.2.4 Isolate characterisation

2.2.4.1 Motility

2.2.4.1.1 Motility assay

Semi-solid agar plates (0.3% agar; TSB) were prepared on the day of use. A single colony from a 48 h plate was touched against the surface of the motility agar. Growth distance was measured with a ruler after 24 and 48 h, and imaged using a Uvipro gold (Uvitec) gel documentation system.

2.2.4.1.2 Hanging drop method

Isolates grown overnight as liquid cultures in TSB were viewed for motility using the "hanging-drop" method (Webley, 1953) and phase contrast microscopy (Olympus CK2).

2.2.4.1.3 Transmission electron microscopy

Transmission electron microscopy was performed essentially as described by Davies and Frerichs (Davies & Frerichs, 1989). Isolates were grown for 48 h at 22°C on TSA. Bacterial colonies were carefully resuspended in 5 ml dH_2O so as not to detach the flagella. The suspensions were diluted and the optical density at 600nm (OD_{600nm}) measured of 0.7 (~10^6 colony forming units [c.f.u.] ml^{-1}). Ninety microlitres of the resultant suspension was transferred to an Eppendorf and 10 µl of 10% phosphotungstic acid were added. A drop of the suspension was

placed on a carbon-coated formvar grid and left to air dry for 5 min before being viewed under TEM on a LEO 912 AB OMEGA TEM running at 120kV. Tif images were captured on OLYMPUS ITEM Soft Imaging System.

2.2.4.2 Lipase activity

To detect lipase activity, isolates were grown on TSA supplemented with 1.0% (v/v) Tween 80 at 22°C for 48 h as previously described (Davies & Frerichs, 1989; Davies, 1991d). A positive result was recorded when clear zones were observed around the colonies (Barrow & Fektham, 1993). Lipase activity in combination with motility distinguishes two distinct biotypes. Biotype 1 isolates are motile and able to hydrolase Tween, while biotype 2 isolates are non-motile and unable to hydrolyse Tween.

2.2.4.3 Biochemical tests

2.2.4.4 Growth curve analysis

Growth curves were performed in 25 ml volumes of TSB in 100 ml Erlenmeyer flasks for all conditions except anaerobic growth. Anaerobic growth curves were achieved through growth in 100 ml TSB in a 100 ml Duran overlaid with sterile mineral oil. Aerobic cultures were grown at 22°C with shaking at 120 rpm; anaerobic cultures were grown statically. One millilitre samples were taken at two hourly intervals and the OD_{600nm} measured.

2.2.4.5 O-Serotyping

O-serotyping of isolates was conducted by slide agglutination as previously described (Davies, 1990). Isolates were grown for 48 h at 22°C on TSA. The resultant bacterial growth was resuspended in 1 ml dH_2O and heated at 100°C for 1 h in a boiling water bath. Cells were centrifuged at 13,000 x *g* for 2 min and resuspended in 1 ml dH_2O. Cells were centrifuged at 13,000 x *g* for 2 min and resuspended in 100µl dH_2O. Five microlitres of heat-inactivated O-antigen suspensions were mixed with an equal volume of O1, O2, O5, O6 or O7 antiserum on a glass slide. Agglutination was recorded as a positive result. Isolates that gave a positive reaction were recorded as +, however isolates that gave a weak cross agglutination with other antisera were recorded as ±. Isolates that did not react with the antisera were recorded as –.

2.2.4.6 Preparation of cross absorbed sera

Cross-absorbed anti-O1 and anti-O5 antisera were prepared against the cross-agglutinating isolate RD426 as previously described (Davies, 1990). Isolate RD426 was grown for 48 h at 22°C on TSA. The resultant growth was resuspended in 10 ml 0.85% (v/v) PBS and washed twice after centrifugation at 4,000 x *g* for 10 min at 4°C. The final pellet was resuspended in 10 ml PBS and 2 ml aliquots transferred to each of five Eppendorfs. Cells were pelleted by centrifugation at 10,000 x *g* for 10 min and resuspended in 1 ml of antiserum. The culture was incubated for 1 h at room temperature on a rotating mixer. The bacteria were pelleted by centrifugation at 10,000 x *g* for 10 min and the antiserum was used to resuspend the bacterial cells of the second vial. This process was repeated for all vials and the supernatant was filter sterilised (0.2 µM pore filter) and stored at -80°C.

2.2.4.7 Isolation of outer membrane

Outer membranes were isolated essentially as previously described (Davies, 1991c; Davies *et al.*, 1992, 2003, 2004). The growth of 400 ml mid-log phase cultures was halted by chilling on iced water for 5 min and the bacterial cells harvested by centrifugation at 13,000 x *g* for 30 min at 4°C. The bacterial cells were resuspended in 50 ml of 20 mM-Tris/HCl (pH 7.2) and harvested by centrifugation at 12,000 x *g* for 30 min at 4°C. The cells were resuspended in ~7 ml ice-cold 20mM Tris/HCl (pH 7.2) and sonicated, on iced water, for 5 min using a Soniprep sonicator (11 microns amplitude). The sonicated samples were brought to a final volume of 10 ml with 20mM Tris/HCl (pH 7.2), and centrifuged at 13,000 x *g* for 30 min at 4°C to remove unbroken cells. The supernatants were centrifuged at 84,000 x *g* for 1 h at 4°C in a Sorvall ultracentrifuge to pellet the cell envelopes. The gelatinous pellets were resuspended in 7 ml of 0.5% sodium N-lauroylsarcosine (Sarkosyl; Sigma) for 20 min at room temperature to solubilise the cytoplasmic membranes and centrifuged at 84,000 x *g* for 1 h at 4°C to pellet the OMs. The gelatinous pellets were resuspended in 20 mM-Tris/HCl (pH 7.2) and centrifuged at 84,000 x *g* for 1 h at 4°C. The final pellets were carefully resuspended in ~0.5 ml of 20 mM-Tris/HCl (pH 7.2). Fifty microlitre volumes of OM suspensions were removed and transferred to Eppendorf tubes for protein concentration determination by the modified Lowry

procedure (Markwell *et al.*, 1978). One hundred microlitre aliquots of the OM suspensions were adjusted to 2 mg/ml with 20 mM-Tris/HCl (pH 7.2) and stored at -20°C), and samples were stored at -80°C.

2.2.4.8 Preparation of LPS from OMs

LPS was prepared by proteinase K digestion of the OM samples as previously described (Davies & Donachie, 1996; Davies, 1989; Davies *et al.*, 1991, 1994; Hitchcock & Brown, 1983). Twenty five micrograms of proteinase K resuspended in sample buffer (Sigma, P6556) were added to 20 µl of 1 mg/ml OM sample (also in sample buffer, and previously heated at 100°C for 5 min) and incubated at 60°C for 1 h. Ten microlitres of each LPS sample were separated by SDS-PAGE.

2.2.4.9 One-dimensional SDS-PAGE

Outer membrane samples were adjusted to 1 mg/ml in 2X sample buffer (0.125 M Tris/HCl [pH6.8], 20% [v/v] glycerol, 4% [w/v] SDS, 10% [v/v] ß-mercaptoethanol, 0.0025% [w/v] bromophenol blue) and heated at 100°C for 5 min. Twenty micrograms of protein for each sample were stacked in a 4% acrylamide stacking gel and separated in a 12% acrylamide resolving gel. Electrophoresis was carried out using a discontinuous buffer system (Hoeffer SE600) at a constant current of 20 mA per gel through the stacking gel (until the dye front reached the resolving gel) and at a constant current of 30 mA through the resolving gel until the dye front reached about 1 cm from the bottom of the gel. Gels were stained with Coomassie Brilliant blue R250 (Fisher Bioreagents, UK) (Davies *et al.*, 1992, 2003, 2004; Laemmli, 1970). The apparent molecular masses of the OMPs were estimated using a low molecular mass SDS marker kit (GE Healthcare, UK). OMP-types were assigned as previously described (Davies, 1989, 1991c). To allow distinction between the major OMPs, selected OM samples were heated at 50, 60, 70, 80, 90 and 100°C for 5 min prior to separation by SDS-PAGE (Davies, 1991c). Gels were stained overnight using Coomassie brilliant blue (R250, Sigma, UK) and destained using a 30% methanol - 15% acetic acid solution.

Mini gels were prepared to the same specifications as previously (4% stacking gel, 12% resolving gel), however only ten micrograms of protein were loaded into the gel. Electrophoresis was carried out at constant current of 30 mA through

both the stacking and resolving gels, using a Biorad Mini-PROTEAN II gel electrophoresis system.

2.2.4.10 Silver-staining of LPS

Lipopolysaccharides were visualised by silver-staining (Tsai & Frasch, 1982). After electrophoresis, the gel was soaked overnight in 40% ethanol - 5% acetic acid solution. The fixing solution was replaced with 0.1% periodic acid in 40% ethanol - 5% acetic to oxidise the LPS for 5 min. The gel was then washed three times in dH_2O before 150 ml staining reagent (0.013M NH_4OH, 0.019M NaOH, 0.66% $AgNO_3$, dH_2O) were added for 10 min. The gel was washed three times in dH_2O before 100 ml developing solution (50 mg Citric acid, 0.5 ml 37% formaldehyde per litre) was added. The LPS was stained dark brown in ~5 min and the staining stopped with the addition of water. The gel was washed thoroughly and stored in dH_2O.

2.2.4.11 Western blotting

The LPSs of paired representative isolates were analysed by Western-blotting using anti-O1, O2 and O5 antisera, as well as with cross-absorbed anti-O1 and anti-O5 antisera, essentially as previously described (Ali *et al.*, 1992; Davies *et al.*, 1990; Hounsome *et al.*, 2011; McCluskey *et al.*, 1994). SDS-PAGE gels were equilibrated in transfer buffer (25 mM Tris, 192 mM glycine [pH8.3]) for 30 min and proteins transferred to nitrocellulose membranes overnight at 30 V. The nitrocellulose membranes were washed twice in Tris-buffered saline (TBS) for 5 min, blocked with 3% gelatin in TBS for 1 h and washed twice in TTBS (0.05% Tween 20 in TBS) for 5 min.

The membranes were incubated in primary rabbit antiserum diluted 1:250 in antibody buffer (1% gelatin in TTBS) for 2 h at room temperature and washed twice in TTBS for 5 min each. Membranes were then incubated in secondary horseradish peroxidase (HRP)-conjugated anti-rabbit antibody diluted 1:3,000 in antibody buffer for 2 h at room temperature followed by two 5 min washes in TTBS and one 5 min wash in TBS. Membranes were developed in a substrate solution containing 0.05% (w/v) 4-chloro-1-napthol (dissolved in 20 ml of ice-cold methanol) and 0.05% (v/v) hydrogen peroxide in 100 ml of TBS. Development

was stopped by washing the membranes in distilled water for 10 min. The membranes were dried on filter paper and subsequently photographed.

2.2.5 **Challenge studies**

2.2.5.1 Viability counts

Viable counts were performed using the method of Miles and Misra (Miles *et al.*, 1938). Eight serial ten-fold dilutions were made in phosphate buffered saline (PBS) and 20 µl of each dilution plated out on TSA. Plates were incubated at 22°C and counts made after 24 h. All plate counts were made in triplicate.

2.2.5.2 Fish stocks

Stocks of disease-free rainbow trout and Atlantic salmon, reared in-house at the CEFAS Weymouth laboratory, were used for the challenge studies. These included rainbow trout (80 g; bath challenge) and Atlantic salmon (35g - i.p.-challenge; 35g – bath challenge).

2.2.5.3 Preparation of intraperitoneal (i.p) challenge dose

Isolates RD6 and RD420 were recovered from -80°C glycerol stocks onto TSA and incubated at 22°C for 48 h. Isolated pure colonies were inoculated into 400 ml TSB and grown overnight at 22°C with shaking at 120 rpm until an OD_{600nm} of 1.5 had been reached. Cultures were harvested and resuspended in 100 ml PBS by centrifugation at 300 x *g* for 20 min at 4°C (repeated three times). Finally, samples were resuspended in 100 ml sterile PBS. The final suspension was adjusted to an OD_{600nm} of 1.0, being equivalent to a challenge dose of ~1.0 x 10^9 c.f.u's ml^{-1}. Colony forming units per millilitre were determined using viability counts, as described previously (2.2.5.1).

2.2.5.4 Preparation of bath challenge dose

Isolates recovered from the head kidneys of fish previously challenged during the i.p. experiments were used to prepare the bath challenge inoculum. These *in vivo*-passaged isolates were maintained as glycerol stocks at -80°C. Four litres of overnight culture grown at 22°C with shaking at 120 rpm (2 L in separate 5 L Erlenmeyer flasks) were prepared for both isolates RD6 and RD420. Bacteria

were harvested in 500 ml aliquots by centrifugation at 2,500 x *g* for 20 min at 4°C. Pellets were resuspended in 50 ml sterile PBS and pooled to a total volume of 200 ml. Suspensions were centrifuged at 2,500 x *g* for 20 min at 4°C and resuspended in 50 ml sterile PBS. Suspensions were then centrifuged at 2,500 x *g* for 20 min at 4°C and resuspended in 25 ml sterile PBS. The 25 ml resuspensions from each original 2 L Erlenmeyer flask were pooled to give a final volume of 50 ml and adjusted to an OD_{600nm} of 1.0 using sterile PBS; this being equivalent to a challenge dose of ~1.0 x 10^9 c.f.u's ml^{-1}. Colony forming units per millilitre were determined using viability counts, as described previously (2.2.5.1). The preparation of bath challenge doses is summarised in Fig. 2.1.

2.2.5.5 Intraperitoneal challenge of Atlantic salmon

All Atlantic salmon were fasted 24 h prior to challenge. Atlantic salmon were removed from their 50 L holding tanks (containing 20 L water) and placed in a bucket containing Tricaine methane sulphonate (MS222) at a sufficient strength to induce anaesthesia within 2-3 min (~90 mg L^{-1}). Two groups of 20 Atlantic salmon parr (approx. 35 g each) and one group of five Atlantic salmon were each challenged by i.p. injection with strain RD420, positive control RD6 and negative control PBS respectively. The virulent strain RD6 had been used as positive control in previous studies (Davies, 1991a; Haig *et al.*, 2011). Fish were injected into the peritoneal cavity with a 0.1 ml dose containing 5.7 x 10^8 and 2.18 x 10^8 c.f.u's ml^{-1} for isolates RD6 and RD420 respectively. After challenge, fish were returned to their original tank and further observed to confirm they had safely recovered from anaesthesia. Supply of feed was resumed 24 h post challenge. Test temperatures were monitored at 16°C throughout the challenge experiments, which were to last a maximum of 21 days. Throughout the challenge, fish were monitored at 2 h intervals, with moribund or dead fish removed immediately. Moribund fish were euthanised using an approved Schedule One method and swabs from both hind gut and head kidney taken and plated onto TSA and incubated at 22°C for 48 h. Isolates recovered in pure culture were confirmed as *Y. ruckeri* based on colony appearance and Mono-Yr latex agglutination testing (Bionor Laboratories, Norway).

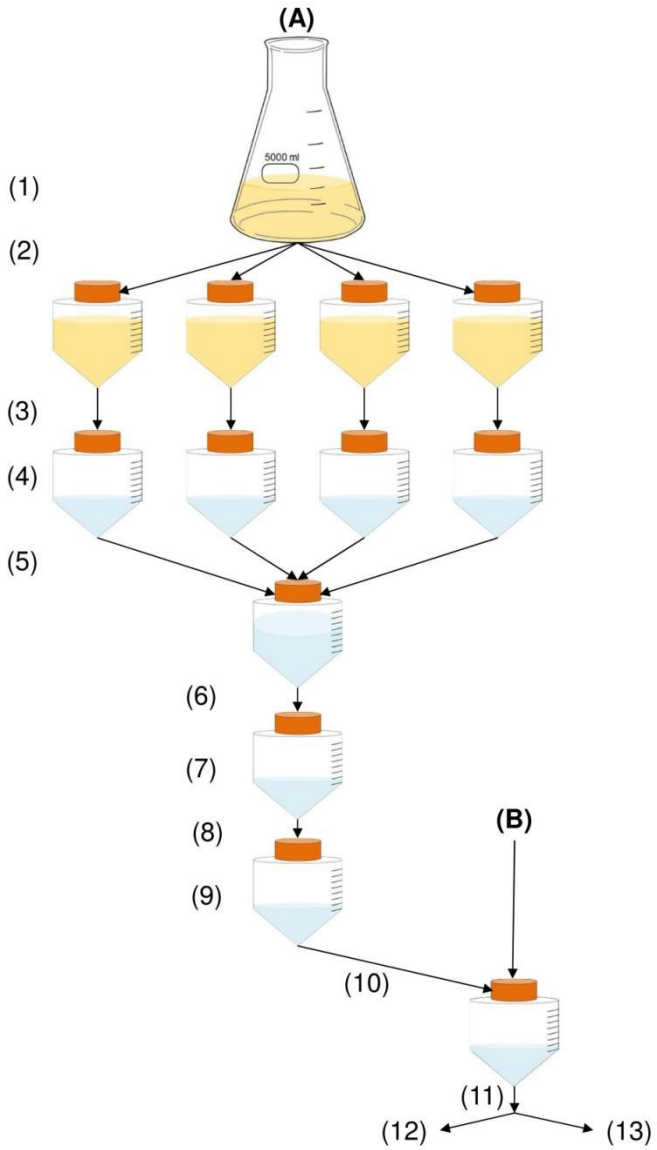


Fig. 2.1 Preparation of inoculum for bath challenge of Atlantic salmon and rainbow trout with *Y. ruckeri*.
Isolates RD6 and RD420 were prepared in the same way. Two litres of TSB in a 5 L Erlenmeyer flask were inoculated with plate grown cultures and grown overnight to an OD_{600nm} of 1.5 (**1**; this process was completed in duplicate although only **[A]** is shown from steps **[1-8]**). The overnight culture was separated into 500 ml aliquots in 500 ml falcon tubes **(2)**. The aliquots were then centrifuged at 2,500 x *g* for 20 min at 4°C **(3)** and resuspended in 50 ml sterile PBS **(4)**. Samples were pooled to 200 ml total volumes in a single 500 ml falcon tube **(5)**. The sample was centrifuged at 2,500 x *g* for 20 min at 4°C **(6)** and resuspended in 50 ml PBS **(7)**. The sample was centrifuged at 2,500 x *g* for 20 min at 4°C **(8)** and resuspended in 25 ml PBS **(9)**. At this point **(A)** and **(B)** were pooled giving a total volume of 50 ml **(10)**. The sample was adjusted to OD_{600} 1.0 equating to approximately 1 x 10^9 c.f.u's ml^{-1} and a viability check conducted **(11)**. For RD6, 200 ml challenge dose was required for Atlantic salmon and 500 ml for rainbow trout **(12)**. For RD420, 200 ml challenge dose was required for Atlantic salmon and 500 ml for rainbow trout for each replicate (400 ml for Atlantic salmon and 1000 ml for rainbow trout in total)**(13)**.

2.2.5.6 Bath challenge of Atlantic salmon and rainbow trout

Four groups of 15 Atlantic salmon (approx. 35 g) were transferred from their holding tank (20 L of water, 50 L bucket) to 10 L of water, in a 20 L aerated bucket for challenge. Four groups of 25 rainbow trout (approx. 80 g) were

transferred from their holding tank (50 L of water, 100 L bucket) to 20 L of water, in a 50 L aerated bucket for challenge. Numbers were selected to mimic natural shoaling.

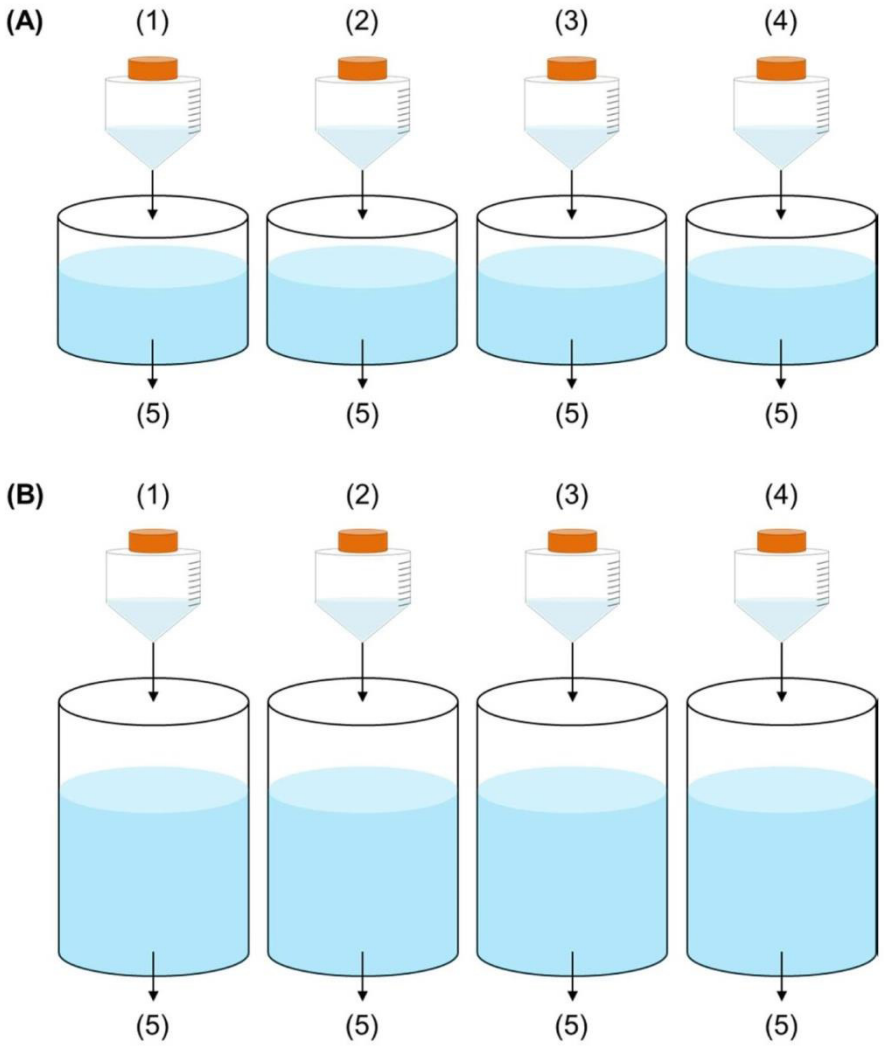


Fig. 2.2 Administration of bath challenge dose to Atlantic salmon and rainbow trout. Atlantic salmon **(A)** and rainbow trout **(B)** were challenged with isolates of *Y. ruckeri.* Isolate RD6 **(1)** was used as a positive control, RD420 **(2 & 3)** as the challenge strain (in duplicate), and PBS **(4)** was the negative control. Atlantic salmon were challenged with a 200 ml bacterial dose in groups of 15, held in 10 L of water in a 20 L bucket. Rainbow trout were challenged with a 500 ml bacterial dose in groups of 25, held in 20 L of water in a 50 L bucket. Viability counts were performed on all challenge buckets to determine actual challenge dose **(5)**.

Two hundred millilitres of bacterial suspensions were added to Atlantic salmon buckets giving final concentrations of 8.4×10^7, 1.1×10^8 and 1.21×10^8 c.f.u's ml^{-1} for RD6, RD420 tank 1 and RD420 tank 2, respectively. Two hundred millilitres of bacterial suspensions were added to rainbow trout buckets giving final concentrations of 8.0×10^7, 1.0×10^8 and 1.02×10^8 c.f.u's ml^{-1} for RD6, RD420 tank 1 and RD420 tank 2, respectively. Colony forming units per millilitre were determined using viability counts, as discussed previously. Experiment

duration was up to 21 days. All challenge experiments were conducted in compliance with the requirements of the UK's Animals (Scientific Procedures) Act 1989, under a UK Home Office Project Licence. This included prior approval by the CEFAS Weymouth Local Review of Ethical Procedures Committee.

2.3 Results

2.3.1 Biotype variation in isolates of *Y. ruckeri*

The presence or absence of flagella was confirmed using TEM in motile (Fig. 2.3A) and non-motile isolates (Fig. 2.3B). In motile isolates, one or two flagella were normally associated with each bacterial cell although in some cases the flagella had become detached, most likely during the preparation for TEM. The ability or lack thereof to hydrolyse Tween was confirmed using 1.0% Tween 80 (v/v) in all isolates (Fig. 2.3A & B). Isolates that were able to hydrolyse Tween 80 showed crystallisation around the colonies, while isolates unable to, did not.

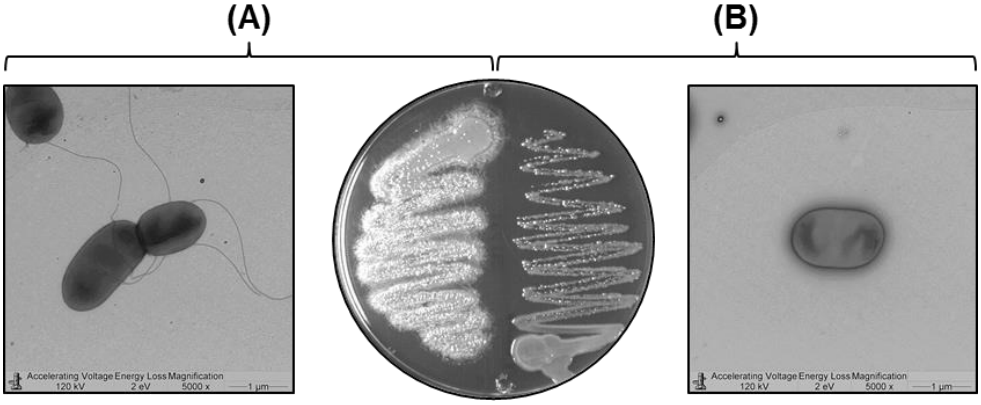


Fig. 2.3 Transmission electron microscopy and Tween hydrolysis of selected isolates of *Y. ruckeri.*
The presence or absence of flagella is shown in TEM images of motile **(A)** and non-motile isolates **(B)** RD124 and RD6 respectively, while tween hydrolysis was examined on TSA supplemented with Tween 80 (1.0% v/v; shown centre image). Positive Tween hydrolysis was observed in isolate RD124, while isolate RD6 was negative. In combination, motility and Tween hydrolysis positive isolates represent biotype 1, while non-motile Tween hydrolysis negative isolates represent biotype 2.

Of the 109 isolates recovered from Atlantic salmon, all were of biotype 1 (i.e., these were motile and positive for lipase hydrolysis). Twenty four out of the 26 rainbow trout isolates were of biotype 2 (i.e., non-motile and lipase negative) (Fig. 2.4). All isolates examined conformed to what had previously been described as the biotype 1 and biotype 2 phenotypes - motile isolates were able to hydrolyse Tween, while non-motile isolates were unable to do so.

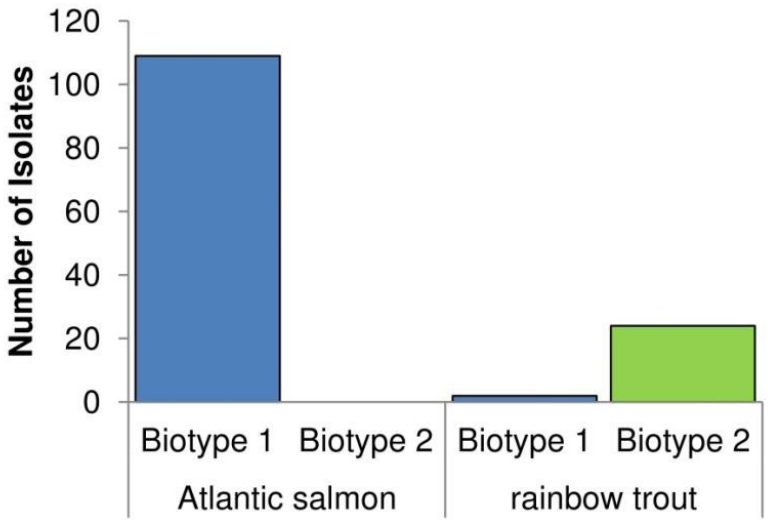


Fig. 2.4 Distribution of biotypes of *Y. ruckeri* amongst Atlantic salmon and rainbow trout.
Biotype 1 isolates are motile and positive for Tween hydrolysis, while biotype 2 isolates are non-motile and Tween hydrolysis negative. All isolates recovered from Atlantic salmon were motile and Tween hydrolysis positive, while the majority of isolates recovered from rainbow trout were non-motile and Tween hydrolysis negative.

2.3.2 **Serotype variation in isolates of *Y. ruckeri***

Serotypes were confirmed using slide-agglutination assays (Fig. 2.5), as detailed in section 2.2.4.5 of the Materials and Methods. Twenty nine isolates were of serotype O1, 34 were of serotype O2, and 14 were of serotype O5. Fifty eight isolates showed a strong positive agglutination reaction with type O1 antiserum but also showed a weak cross-agglutination with type O5 antiserum. This cross-reactivity was removed when anti-O5 antiserum was cross-absorbed with the O1/O5 isolate RD426. However, the cross-absorbed antiserum still agglutinated with serotype O5 isolates on the slide.

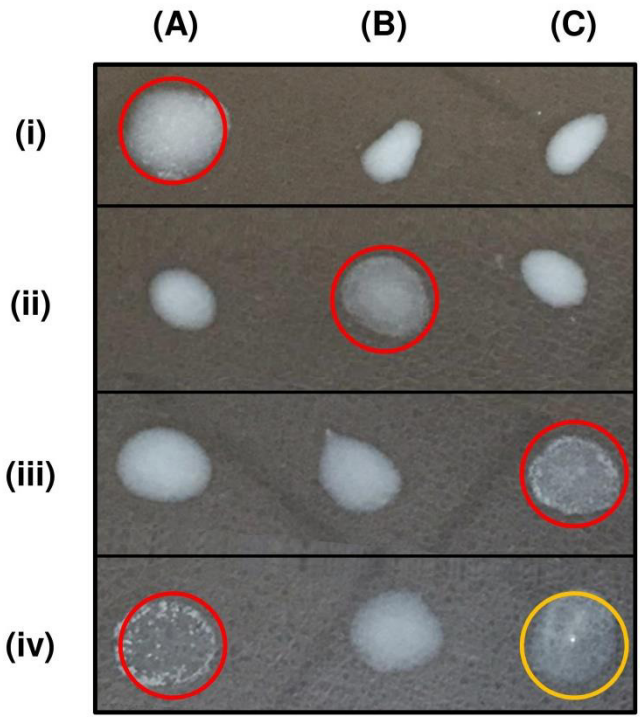


Fig. 2.5 Slide agglutination results to determine serotype of isolates of *Y. ruckeri*. Isolates of serotype O1 (**i**; RD6), O2 (**ii**; RD64), O5 (**iii**; RD28) and O1/O5 (**iv**; RD420) were reacted with anti-O1 antiserum **(A)**, anti-O2 antiserum **(B)** and anti-O5 antiserum **(C)**. Positive reactions were confirmed as agglutination within a few seconds (red circles). A weak reaction between the isolate of serotype O1/O5 **(iv)** and anti-O5 antiserum **(C)** is shown with an orange circle.

As shown in Fig. 2.6A, the 109 isolates recovered from Atlantic salmon between 2001 and 2014 exhibited substantial serotypic diversity, with isolates of serotypes O1, O2, O5 and O1/O5 being recovered. Isolates of serotype O1/O5 were the most commonly recovered (56 isolates) whereas serotype O2 (34 isolates) was the next most abundant; serotypes O1 (5 isolates) and O5 (14 isolates) were recovered less frequently. Isolates of serotype O1 have been recovered irregularly from Atlantic salmon, with isolates appearing in 2001 (one isolate), 2005 (two isolates), 2006 (one isolate) and 2010 (one isolate). Between 2001 and 2006 isolates of serotype O2 were the most frequently recovered serotype (with the exception of 2005). Isolates of serotype O2 are still commonly recovered from Atlantic salmon, and were particularly abundant in 2010. An isolate of serotype O1/O5 was recovered in 2002; this serotype was not recovered again until 2007. However, the re-emergence of this serotype in 2007 has led to this becoming the most abundantly recovered serotype to date in Atlantic salmon. Serotype O5 isolates were recovered mainly between 2005 and 2008 and represented the most common serotype in 2005. All isolates recovered from Atlantic salmon were motile (biotype 1), regardless of serotype.

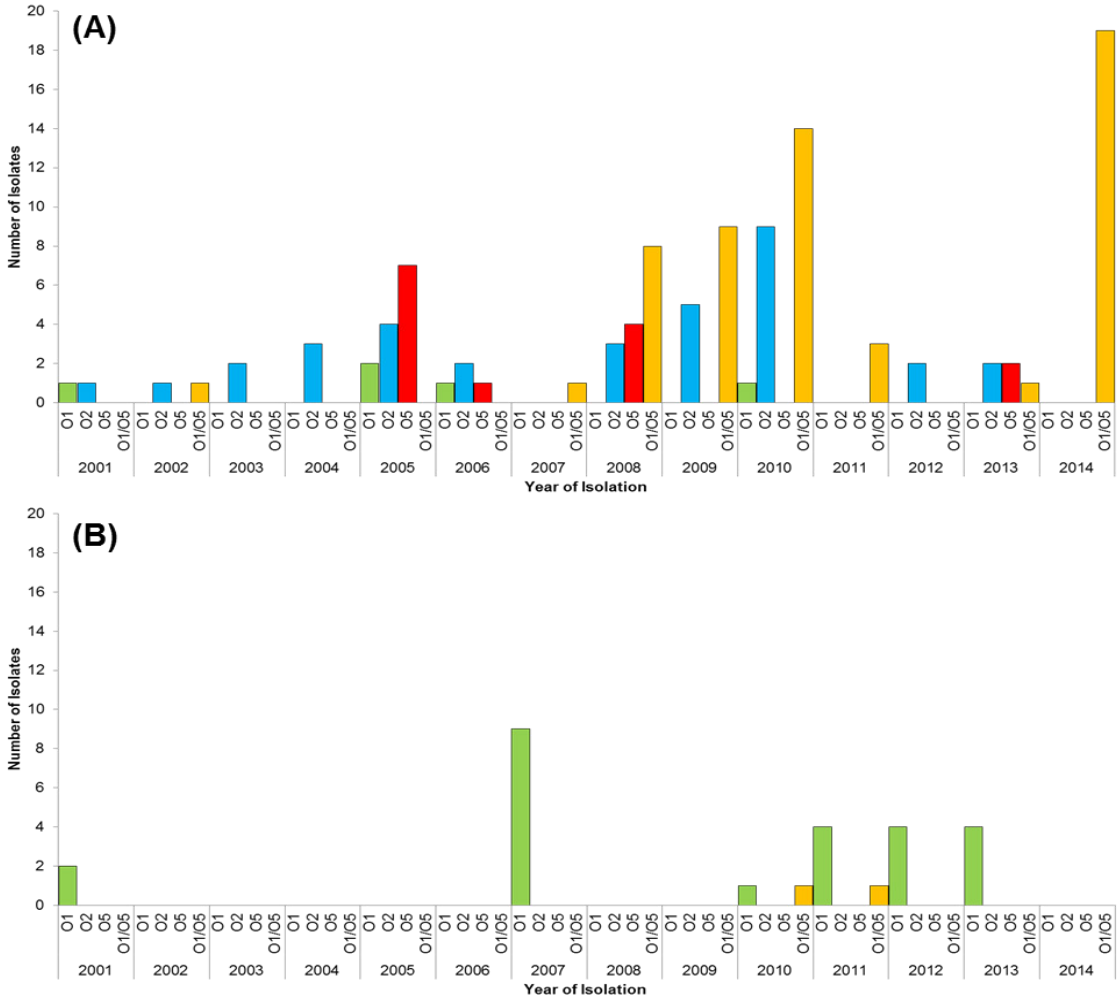


Fig. 2.6 Distribution of O-serotypes of *Y. ruckeri* recovered from Atlantic salmon and rainbow trout in the UK between 2001 and 2014.
Isolates recovered from Atlantic salmon **(A)** showed substantial serotypic diversity, with serotypes O1, O2, O5 and O1/O5 recovered. Isolates recovered from rainbow trout **(B)** were much less diverse, mostly of serotype O1.

Within the rainbow trout population, isolates recovered were almost exclusively of serotype O1 (24 isolates) (Fig. 2.6B). However, two isolates representing the novel serotype O1/O5 were recovered in 2010 and 2011. Notably, these serotype O1/O5 isolates were non-motile (biotype 2) and had not previously been represented in rainbow trout.

2.3.3 **LPS-profile analysis**

From the LPS analysis using SDS-PAGE, it was clear that all isolates possessed a smooth-type LPS structure exhibiting the characteristic ladder pattern (representing the O-antigen repeats) typical of enteric bacteria (Fig. 2.7). However, variation was identified in both the O-antigen (ladder patterns) and core polysaccharide regions of the LPS profiles representing each of the four

serotypes (O1, O2, O5 and O1/O5). The LPS of serotype O1 (Fig. 2.7, lanes 1 and 2) and O1/O5 (Fig. 2.7, lanes 3 and 4) isolates share a common core polysaccharide region but were distinguishable by having slightly different O-antigen ladder patterns (when analysed by eye). The molecular mass of the bands in the O-antigen were different in isolates of all serotypes, indicating that the composition of sugars in the O-antigen are unique to each serotype. The core polysaccharide and O-antigen regions of serotype O2 (Fig. 2.7, lanes 5 and 6) and O5 (Fig. 2.7, lanes 7 and 8) isolates were distinct both from each other and from those of serotype O1 and O1/O5 isolates. In all cases, the LPS types identified by SDS-PAGE corresponded with the O-serotypes determined by slide agglutination and were designated as LPS-types O1 (29 isolates), O1/O5 (58 isolates), O2 (34 isolates) and O5 (14 isolates).

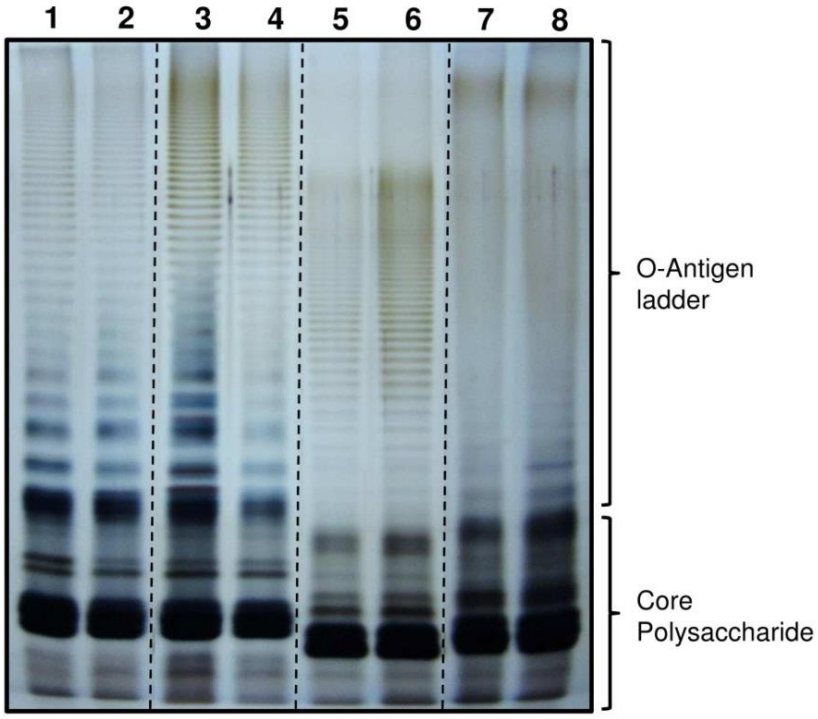


Fig. 2.7 Representative LPS profiles of different serotypes of *Y. ruckeri*.
Isolates are presented in pairs representing serotypes O1 (lanes 1 and 2), O1/O5 (lanes 3 and 4), O2 (lanes 5 and 6) and O5 (lanes 7 and 8). Lanes 1 to 8 represent isolates RD332, RD358, RD428, RD474, RD464, RD486, RD362 and RD372, respectively. The core polysaccharide and O-antigen regions of the LPS molecule are indicated.

2.3.4 **Western-blotting analysis of LPS**

Anti-O1 antiserum (raised against isolate RD6), reacted strongly with both core polysaccharide and O-antigen regions of serotype O1 LPS (Fig. 2.8A, lanes 1 and 2) but not with either component of serotype O2 (Fig. 2.8A, lanes 3 and 4) or O5 (Fig. 2.8A, lanes 5 and 6) LPS. However, this antiserum reacted strongly with

the core polysaccharide region of serotype O1/O5 isolates and weakly with the O-antigen region (Fig. 2.8A, lanes 7 and 8). Upon cross absorbing the anti-O1 antiserum with an O1/O5 isolate, reactivity against the core polysaccharide region of both the O1 (Fig. 2.8B, lanes 1 and 2) and the O1/O5 (Fig. 2.8B, lanes 7 and 8) LPS was removed. The cross-absorbed anti-O1 antiserum gave a moderate reaction with the O-antigen side chains of serotype O1 LPS (Fig. 2.8B, lanes 1 and 2), but no reaction occurred with the O-antigen region of serotype O1/O5 LPS (Fig. 2.8B, lanes 7 and 8). These observations indicated that serotype O1 and O1/O5 LPS types possess a common core polysaccharide region but have different O-antigen side chains from each other. Anti-O2 antiserum reacted strongly with both O-antigen regions and core polysaccharide regions of serotype O2 LPS (Fig. 2.8C, lanes 3 and 4); minor cross-reactivity occurred against a component of the core polysaccharide region of serotype O1/O5 LPS in one isolate (Fig. 2.8C, lane 7, arrow). Anti-O5 antiserum reacted strongly with both the O-antigen and core polysaccharide regions of serotype O5 LPS (Fig. 2.8D, lanes 5 and 6). A weak reaction occurred against the core polysaccharide regions of both serotype O1 (Fig. 2.8D, lanes 1 and 2) and serotype O1/O5 (Fig. 2.8D, lanes 7 and 8) LPS; there was also a very weak reaction against the O2 O-antigen (Fig. 2.8, lanes 3 and 4). When anti-O5 antiserum was cross-absorbed with an O1/O5 isolate, cross-reactivity against the O1 and O1/O5 core polysaccharide regions was mostly removed (Fig. 2.8E, lanes 1 and 2, and 7 and 8); there remained some cross-reactivity against a core polysaccharide component of one of the O1/O5 isolates (Fig. 2.8E, lane 7, arrow). Overall, these results demonstrate that the observed cross reactivity of serotype O1/O5 isolates with anti-O1 and anti-O5 antisera is based mainly on the core polysaccharide region. More significantly, the results demonstrate that the LPS of serotype O1/O5 isolates possesses a unique, previously unrecognised O-antigen which has minor cross-reactivity with that of serotype O1 isolates.

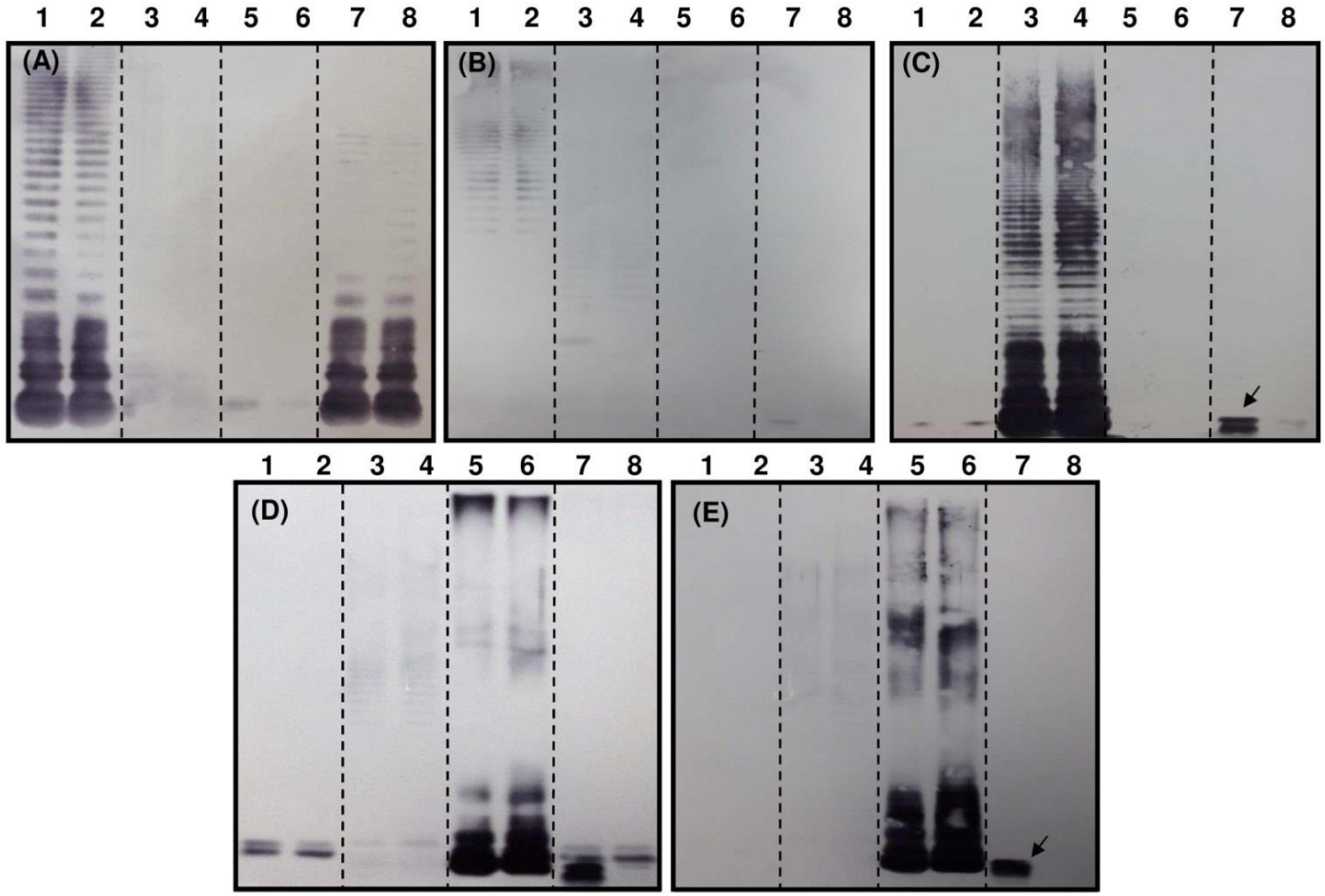


Fig. 2.8 Western blotting analysis of *Y. ruckeri* LPS.
Panels **(A)** to **(E)** were reacted with anti-O1 antiserum, anti-O1 antiserum cross absorbed with an O1/O5 isolate, anti-O2 antiserum, anti-O5 antiserum and O5 antiserum cross absorbed with an O1/O5 isolate, respectively. Lanes 1 to 8 represent isolates RD332, RD358 (serotype O1), RD464, RD486 (serotype O2), RD362, RD372 (serotype O5), RD428 and RD474 (serotype O1/O5), respectively.

2.3.5 **Analysis of OMP profiles**

The OMP profiles of the 135 isolates (supplementary Fig. 8.1) differed slightly to those of the previously described strains in terms of the abundance of the proteins used for characterisation (Davies, 1989, 1991c). Previously, bacterial cells were grown statically under conditions of low aeration, whereas our isolates were grown in a shaking incubator under aerobic conditions. We observed expression of the three major proteins; the heat-modifiable protein OmpA, and the porins, OmpC and OmpF. To confirm that these subtle differences in aeration accounted for the observed differences in protein expression, we compared the OMP profiles of eight reference isolates from the previous study (Davies, 1989, 1991c) grown under aerobic and anaerobic conditions (Fig. 2.9). When isolates were grown aerobically, two major proteins, OmpA and OmpF, were clearly expressed in six of the eight isolates (Fig. 2.9A). In the other two isolates (RD6 and RD128), a single major protein was observed. In these cases, OmpA and OmpF most likely have the same molecular mass (Fig. 2.9A, arrows). However, under anaerobic growth conditions, a third major protein, OmpC was expressed in all eight isolates (Fig. 2.9B, arrow).

In order to confirm the identity of the three major proteins as OmpA, OmpC or OmpF, OMP profiles of four reference isolates (RD6, RD124, RD154 and RD158) were grown under aerobic and anaerobic growth conditions and compared after heating the samples at 50, 60, 70, 80, 90 and 100°C prior to SDS-PAGE (Fig. 2.10). The heat-modifiable protein, OmpA, could be differentiated from OmpC and OmpF due to the known characteristic shift in molecular mass of OmpA, from approximately 30 to 40 kDa, when heated at 90 or 100°C (Davies, 1991c). In addition, the porin proteins OmpC and OmpF only appear in the gel at the higher temperatures (60 to 80°C) because they are associated with peptidoglycan at lower temperatures. Conversely, the low-molecular-mass form of OmpA is present from 50°C. Initially, OmpC and OmpF were identified from their predicted molecular masses (40.5 kDa and 40.1 kDa, respectively) recovered from the genome sequence of isolate ATCC 29473 (Genbank: ACCC00000000)(Chen *et al.*, 2010). However, OmpC could also be distinguished from its known behaviour under different conditions of aeration; OmpC is not expressed as abundantly under aerobic growth conditions (Ni Bhriain *et al.*, 1989). Comparison of the four isolates grown aerobically and anaerobically (Fig.

2.10) confirmed that OmpC represents the higher molecular mass protein because it is up-regulated under anaerobic growth conditions. Identification of OmpA, OmpC and OmpF is most clearly seen in isolate RD158 (Fig. 2.10, G and H). In all cases, the transition of OmpA from low- to high-molecular-mass forms occurred at a lower temperature (80 to 90°C) in cells grown under anaerobic conditions (Fig. 2.10B, D, E and H) compared to cells grown under aerobic conditions (90 to 100°C) (Fig. 2.10A, C, F and G).

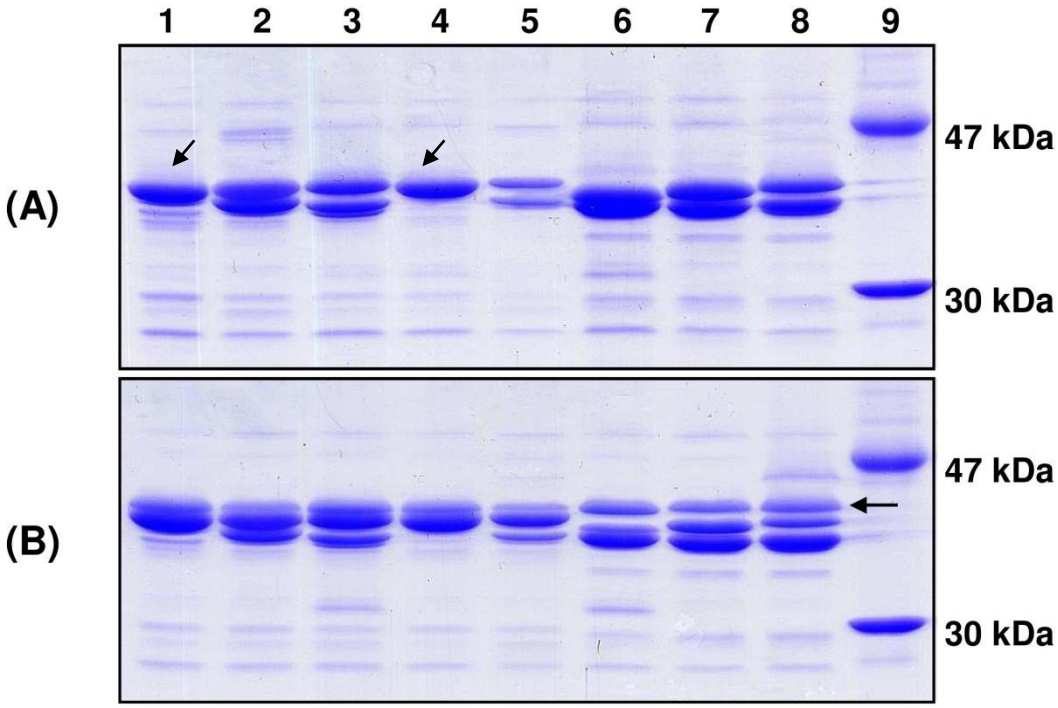


Fig. 2.9 OMP profiles of *Y. ruckeri* reference isolates grown under aerobic and anaerobic growth conditions.
Isolates were grown under aerobic **(A)** and anaerobic conditions **(B)**. Lanes 1 to 8 represent isolates RD6, RD22, RD124, RD128, RD138, RD154, RD158 RD168, respectively. Lane 9 represents a molecular mass standard (GE Healthcare, UK). In **(A)**, the arrows indicate the similar MW of proteins OmpA and OmpF in profiles of OMP-type 1. In **(B)**, the arrow indicates expression of OmpC protein under anaerobic growth conditions.

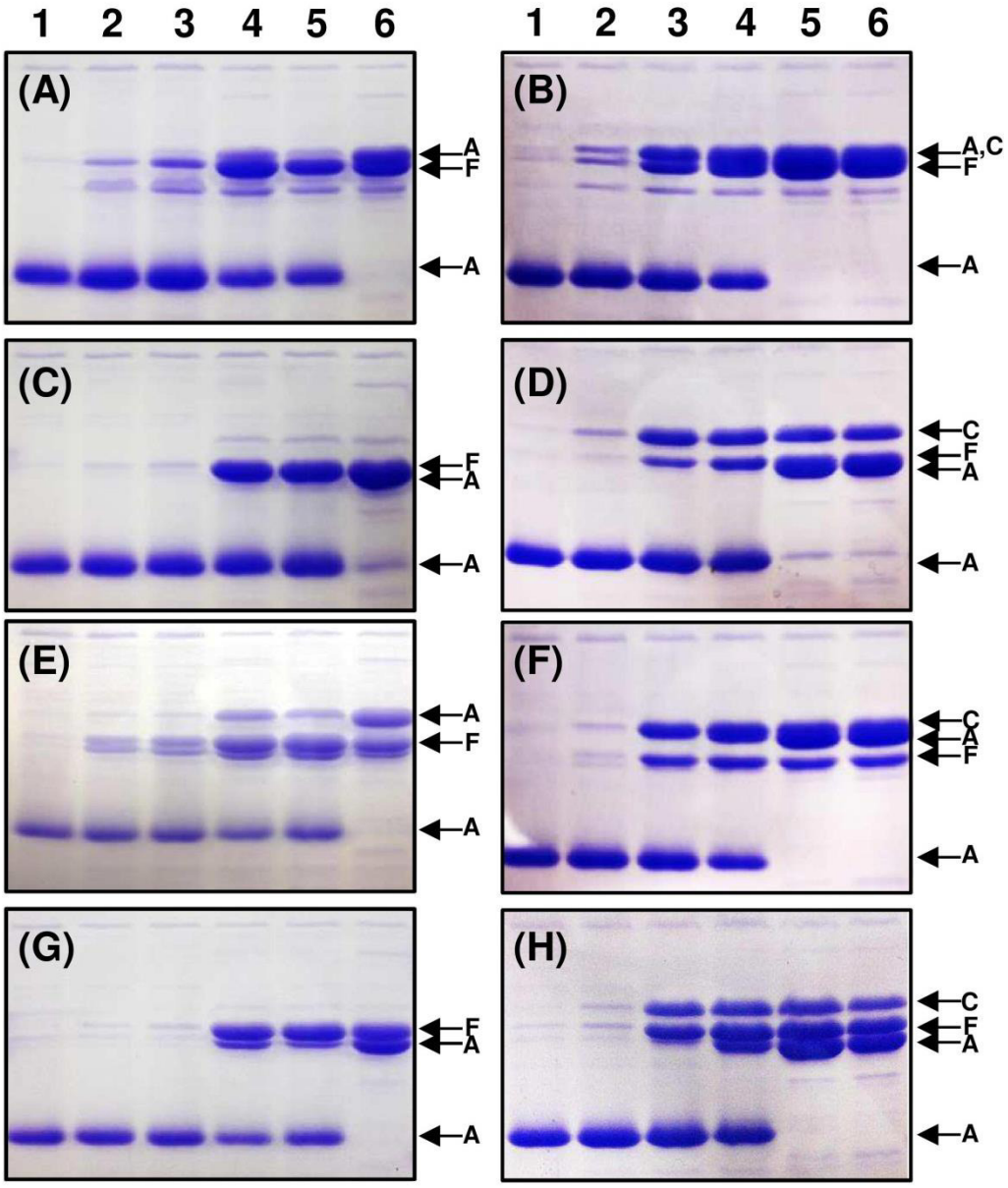


Fig. 2.10 Identification of OmpA, OmpC and OmpF using different solubilisation temperatures and aeration conditions.
Isolates RD6 **(A and B)**, RD124 **(C and D)**, RD154 **(E and F)** and RD158 **(G and H)** were grown under aerobic **(A, C, E and G)** and anaerobic **(B, D, F and H)** growth conditions. The major OMPs OmpA, OmpC and OmpF are indicated by arrows and the letters A, C and F. Lanes 1 to 6 represent solubilisation temperatures of 50, 60, 70, 80, 90 and 100°C, respectively.

Based on variation of the major OMPs, the 135 isolates were assigned to one of four distinct OMP-types using the previously described OMP typing scheme (Davies, 1989, 1991c). The isolates were assigned to OMP-types 1a, 2a, 2c or 3a (Fig. 2.11). The majority (24/26) of rainbow trout isolates were of OMP-type 1a; two isolates were of OMP-type 3a. The Atlantic salmon isolates were more diverse; 32 isolates were of OMP-type 2a, 17 were of OMP-type 2c and 60 were of OMP-type 3a. However, the isolates could be further subtyped based on variation of minor proteins (Fig. 2.11). OMP-type 1a could be subdivided into subtypes 1a.1 (19 isolates), 1a.2 (3) and 1a.3 (2) (Fig. 2.11A); OMP-type 2a could be subdivided into subtypes 2a.1 (18), 2a.2 (7) and 2a.3 (7) (Fig. 2.11B); OMP-type 2c could be subdivided into subtypes 2c.1 (13), 2c.2 (3) and 2c.3 (1) (Fig. 2.11C) and OMP-type 3a could be subdivided into subtypes 3a.1 (36), 3a.2 (16), 3a.3 (4), 3a.4 (2), 3a.5 (2) and 3a.6 (2) (Fig. 2.11D). It is highly significant that rainbow trout and Atlantic salmon isolates were represented by mutually exclusive OMP-types, i.e. rainbow trout isolates were represented almost entirely by OMP-type 1a whereas Atlantic salmon isolates were represented by OMP-types 2a, 2c and 3a. The association of isolates representing the various OMP-types with Atlantic salmon and rainbow trout is summarised in Fig. 2.12.

Each of the four OMP-types was associated almost exclusively with a specific O-serotype (Fig. 2.12). Thus, OMP-type 1a isolates were mostly (90%) of serotype O1, OMP-type 2a isolates were mostly (93.8%) of serotype O2, OMP-type 2c isolates were mostly (58.8%) of serotype O5 (serotypes O1 and O2 were also represented) and OMP-type 3a isolates were mostly (90.3%) of serotype O1/O5. When biotype, serotype and OMP-type were considered together (Fig. 2.12), the clonal groups associated with each species become clearly apparent. Atlantic salmon isolates were highly diverse encompassing 19 different clonal groups. In contrast, rainbow trout isolates were much less diverse being represented by only five distinct clonal groups. The predominant UK rainbow trout *Y. ruckeri* isolate was represented by biotype 2, serotype O1, OMP-type 1a.1 (13/26). Within the Atlantic salmon population the predominant isolates were represented by biotype 1, serotype O1/O5, OMP-types 3a.1 (36/109) and 3a.2 (13/109), and serotype O2, OMP-type 2a.1 (16/109) (Fig. 2.12).

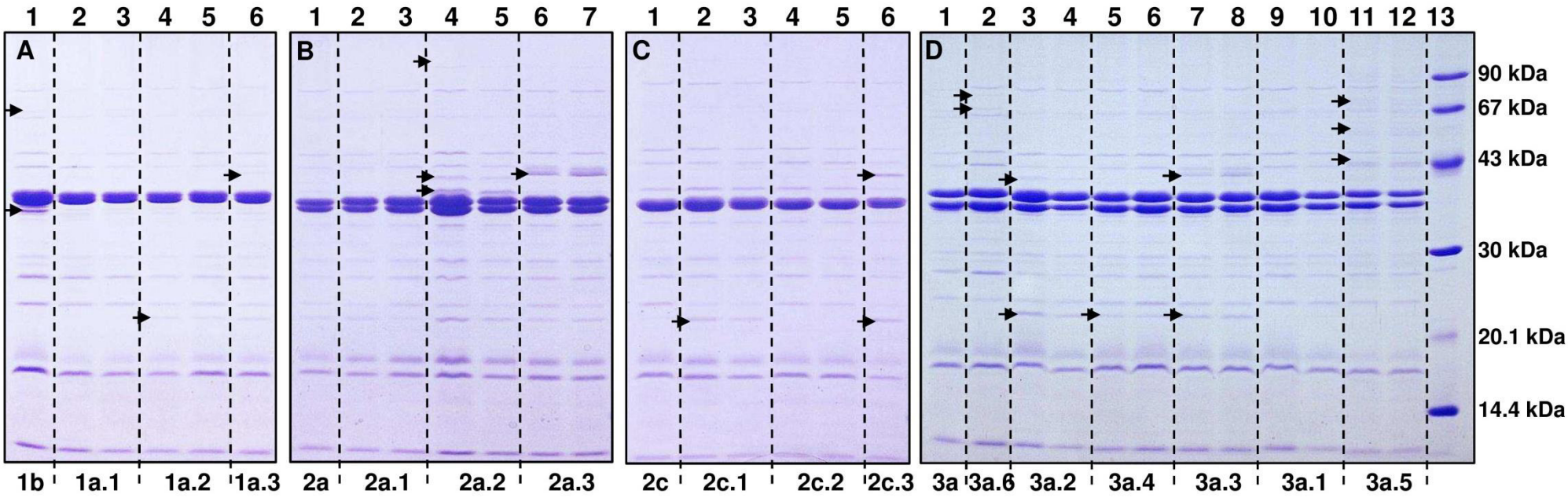


Fig. 2.11 OMP profiles and OMP-subtypes of selected paired isolates of *Y. ruckeri*.
In each panel **(A to D)**, lane 1 represents a reference isolate from the original typing scheme (Davies, 1991c). **(A)** OMP-type 1a: lanes 1 to 6 represent isolates RD6 (reference representing OMP-type 1b with an additional 36 kDa protein), RD400 and RD402 (OMP-type 1a.1), RD396 and RD522 (OMP-type 1a.2) and RD394 (OMP-type 1a.3). **(B)** OMP-type 2a: lanes 1 to 7 represent isolates RD158 (reference), RD336 and RD342 (OMP-type 2a.1), RD436 and RD480 (OMP-type 2a.2), RD418 and RD464 (OMP-type 2a.3). **(C)** OMP-type 2c: lanes 1 to 6 represent isolates RD154 (reference), RD370 and RD412 (OMP-type 2c.1), RD362 and RD380 (OMP-type 2c.2), and RD374 (OMP-type 2c.3). **(D)** OMP-type 3a: lanes 1 to 12 represent isolates RD124 (reference), RD386 (OMP-type 3a.6), RD422 and RD440 (OMP-type 3a.2), RD532 and RD534 (OMP-type 3a.4), RD458 and RD468 (OMP-type 3a.3), RD426 and RD496 (OMP-type 3a.1), RD382 and RD492 (OMP-type 3a.5), Lane 13 represents molecular mass standards. Arrows indicate differences in protein expression with respect to the reference isolate in each panel.

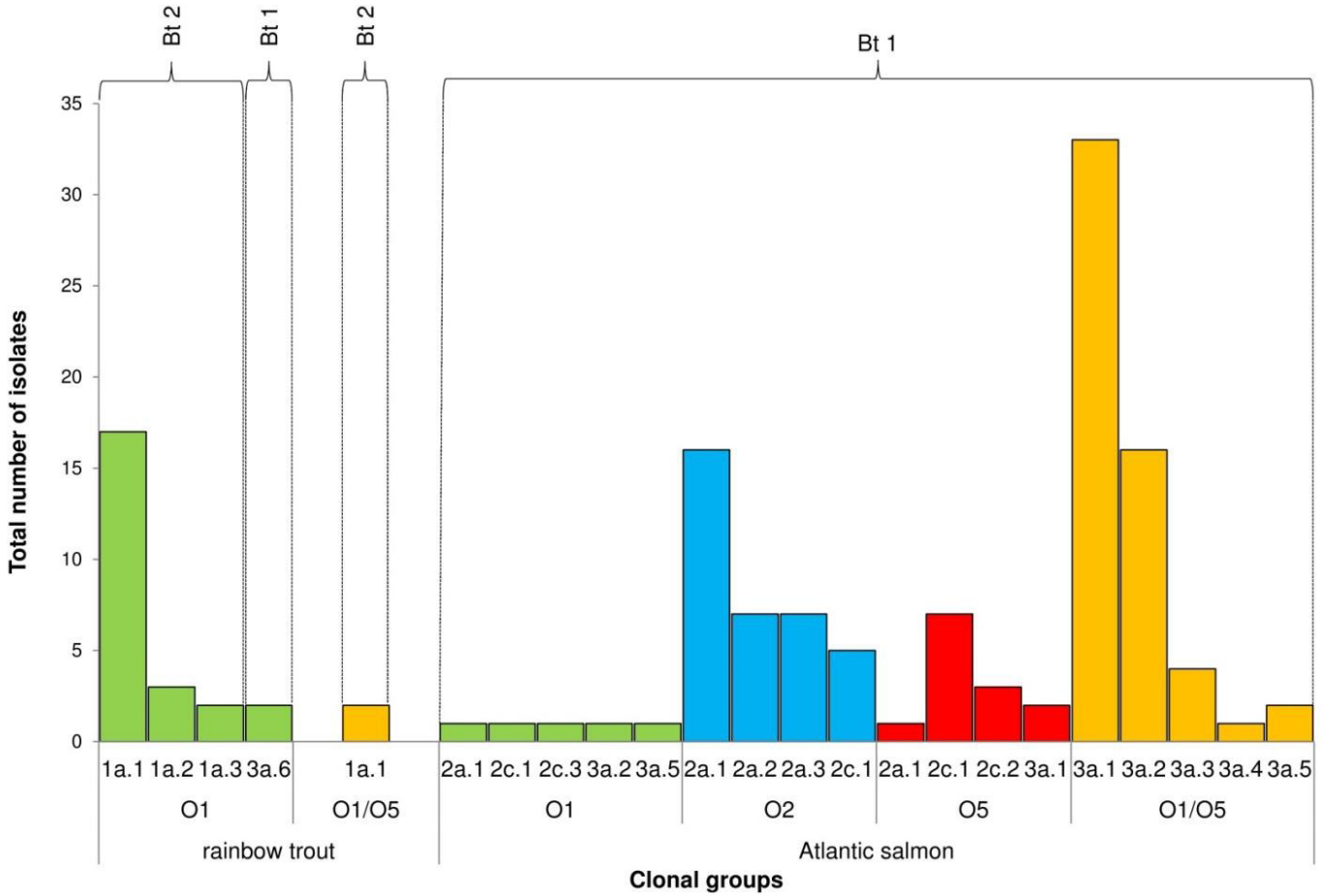


Fig. 2.12 Distribution of clonal groups amongst isolates of *Y. ruckeri* recovered from UK Atlantic salmon and rainbow trout.
Isolates recovered from rainbow trout are represented mainly by serotype O1, and OMP-type 1a. The major clonal group infecting this species is serotype O1 and OMP-type 1a.1. Isolates recovered from Atlantic salmon are more diverse with four O-serotypes recovered and 11 OMP-subtypes. Within the serotypes O2, O5 and O1/O5, the major OMP-subtypes are 2a.1, 2c.1 and 3a.1 respectively. Bt1 – Biotype 1; Bt2 – Biotype 2.

2.3.6 Pathogenicity of the novel serotype O1/O5 isolate towards Atlantic salmon and rainbow trout

2.3.6.1 Intraperitoneal challenge of Atlantic salmon with a novel serotype O1/O5 isolate

The virulence of the novel serotype O1/O5 isolate towards Atlantic salmon was examined using i.p. injection. This was performed in order to satisfy the project license that future bath challenge experiments were viable. Initially, challenge doses of isolates RD6 and RD420 were determined using viable counts. Challenge doses of ~5 x 10^8 c.f.u's ml^{-1} were obtained, equating to OD_{600nm} values of ~0.5 for isolates RD6 and RD420 (Fig. 2.13).

Twenty Atlantic salmon were i.p. challenged with a 0.1 ml suspension of isolate RD420. As a positive control, twenty Atlantic salmon were challenged with RD6, while as a negative control five were challenged with PBS. Plate counts were conducted to determine the actual challenge dose administered; fish challenged with RD6 received ~5.70 x 10^8 c.f.u's ml^{-1}, while those challenged with RD420 received 2.18 x 10^8 c.f.u's ml^{-1}. All plate counts were conducted in triplicate.

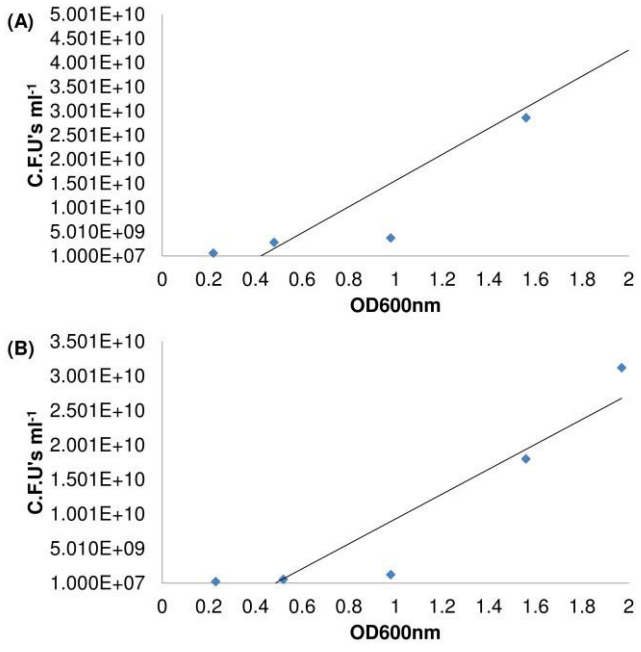


Fig. 2.13 Relationship between OD_{600nm} and colony forming units per millilitre (c.f.u's ml^{-1}) of isolates of *Y. ruckeri*.
The isolates RD6 **(A)** and RD124 **(B)** were examined to determine the relationship between c.f.u's ml^{-1} and OD_{600nm}.

Atlantic salmon challenged with isolate RD6 were killed within two days of challenge (Fig. 2.14), while isolates challenged with the novel serotype O1/O5

isolate RD426 displayed 90% mortality after three days. At this point, the experiment was terminated as a mortality rate of >40% was required by the project license to proceed to the bath challenge studies. Fish were removed when they showed signs of infection, to limit the spread of further bacteria into the water; this ensured that the main source of bacteria was our challenge dose.

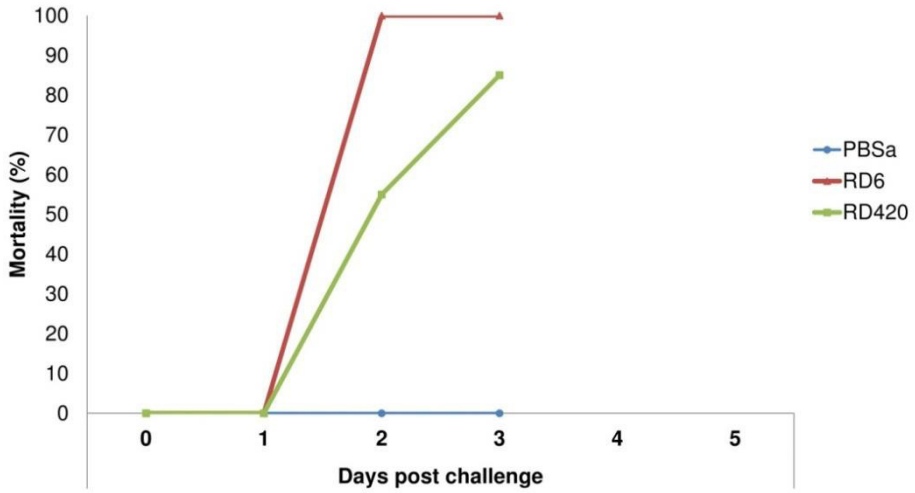


Fig. 2.14 Intraperitoneal challenge of Atlantic salmon with *Y. ruckeri*.
Isolate RD420 is of the novel serotype O1/O5. Isolate RD6 is a positive control (serotype O1) isolate, while PBS_a is a negative control. Fish challenged with isolate RD6 reached mortality after two days, while 90% of fish challenged with isolate RD420 were moribund after 3 days. No fish challenged with PBS died. Atlantic salmon were approximately 35 g at the time of challenge.

The clinical signs of ERM that all fish showed after challenge were typical and identical for both isolates RD6 and RD420. Severe haemorrhaging around the mouth (Fig. 2.15A) and eye (Fig. 2.15B) were obvious several hours prior to death, and usually combined with a unilateral exophthalmos. Atlantic salmon challenged with RD6 and RD420 did not feed well and began to lose equilibrium, swimming erratically to maintain balance. Upon dissection of fish challenged with both RD6 and RD420, haemorrhaging along the flanks of the body was obvious (Fig. 2.15C). Internally, organs generally appeared normal. However, there was haemorrhaging (Fig. 2.15D) and an ascitic fluid oozed from the belly of several fish (Fig. 2.15E).

Bacteria were isolated from both the head kidney and hind gut of all challenged fish both for confirmation and use in subsequent bath challenge experiments. All negative control fish challenged with PBS survived the challenge - no bacteria were subsequently isolated from these fish post mortem.

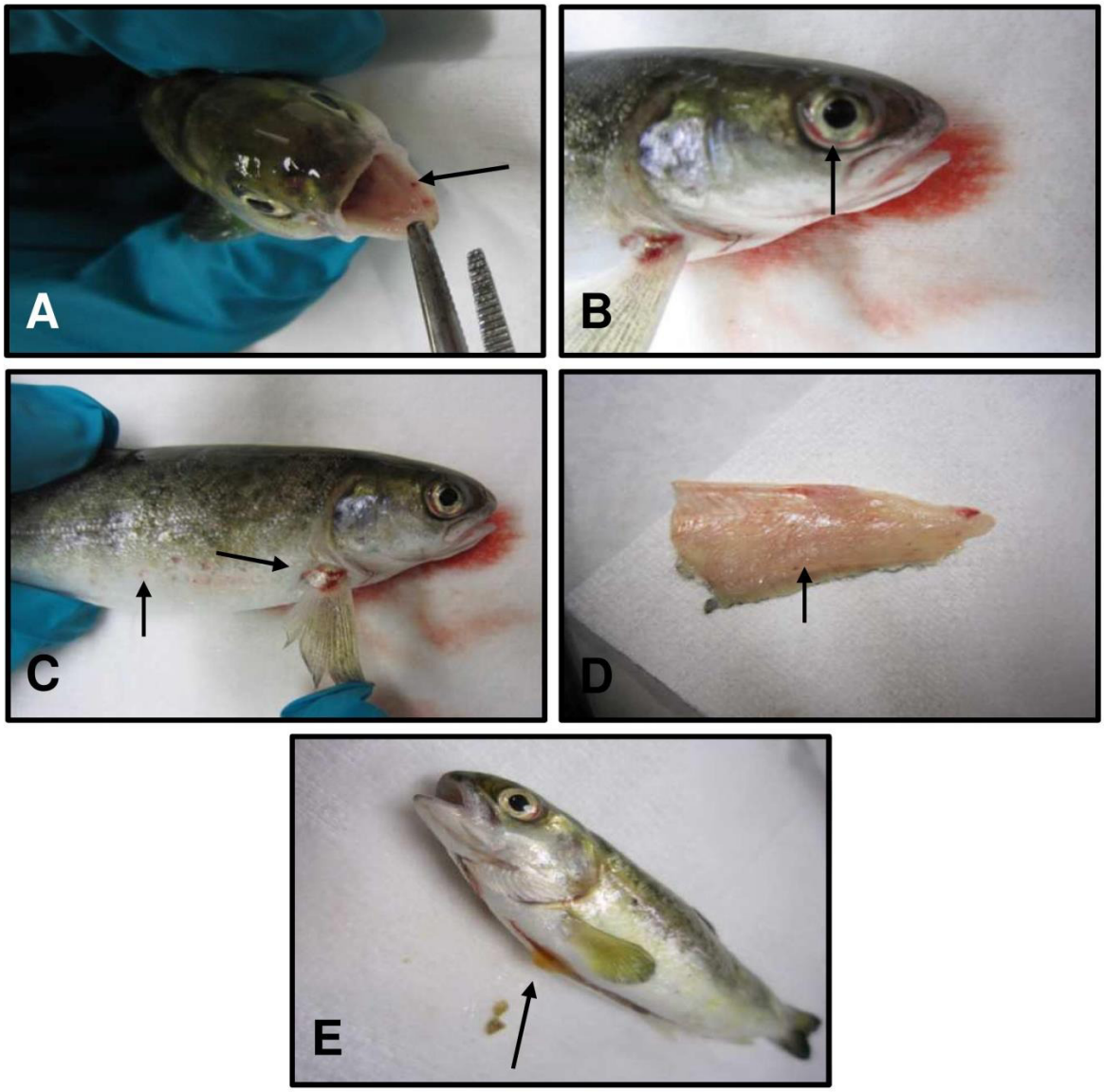


Fig. 2.15 Clinical signs following i.p. challenge of Atlantic salmon with *Y. ruckeri*. Panels **(A)** to **(E)** show the classical symptoms of ERM recovered post infection. Panels represent **(A)** haemorrhaging of the oral cavity; **(B)** haemorrhaging of the eye; **(C)** haemorrhaging at the base of the pectoral fin; **(D)** haemorrhaging of internal tissue; **(E)** oozing ascitic fluid. Atlantic salmon challenged were approximately 35 g, while the challenge dose was 0.1 ml (equating to 5.70×10^8 c.f.u's ml^{-1} and 4.18×10^8 c.f.u's ml^{-1} for RD6 and RD420, respectively).

2.3.7 Bath challenge of Atlantic salmon and rainbow trout with a novel serotype O1/O5 isolate

Bath challenge studies were conducted to (1) determine the pathogenicity of the O1/O5 serotype towards both Atlantic salmon and rainbow trout using a more natural route of exposure, and (2) compare virulence with a known virulent serotype O1 isolate (RD6).

The bath challenge experiments did not replicate the results of the i.p. challenge study, or previous experiments (Davies, 1991a; Haig *et al.*, 2011). While two Atlantic salmon challenged with isolate RD420 became infected, no fish challenged with isolate RD6 did (Fig. 2.16A). No rainbow trout were overcome with infection when challenged with either RD420 or RD6 (Fig. 2.16B).

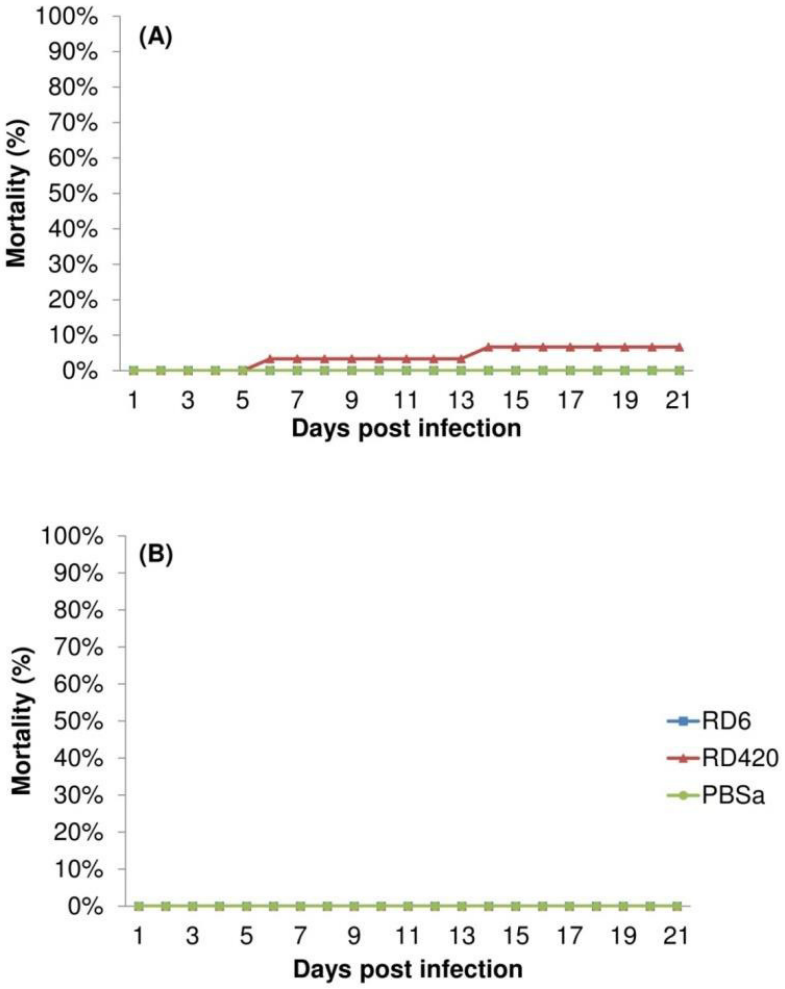


Fig. 2.16 Bath challenge of Atlantic salmon and rainbow trout with *Y. ruckeri*.
Atlantic salmon **(A)** and rainbow trout **(B)** were challenged with isolates RD420 (serotype O1/O5) and RD6 (positive control; serotype O1). PBS was used as a negative control.

Plate counts revealed that the challenge doses obtained were 8.40 x 10^7 c.f.u's ml^{-1} for strain RD6, 1.10 x 10^8 c.f.u's ml^{-1} for RD420.1 and 1.21 x 10^8 c.f.u's ml^{-1} for RD420.2 in Atlantic salmon (the RD420 challenge was completed in duplicate). Doses were determined as 8.00 x 10^7 c.f.u's ml^{-1} for strain RD6, 1.00 x 10^8 c.f.u's ml^{-1} for RD420.1 and 1.02 x 10^8 c.f.u's ml^{-1} for RD420.2 in rainbow trout. The Atlantic salmon that were overcome with infection did so on days 6 and day 14 of the experiment (Fig. 2.16A).

The challenge was terminated 21 days post infection, as per the protocol guidelines.

All fish were terminated using a Schedule one approved method. Blood was collected and pooled in order to obtain serum (for potential use in future experiments). The hind gut was sampled onto both TSA and the selective medium ROD agar and incubated at 22°C for 48 h to detect carriage, while head kidney was sampled onto TSA to determine infected survivor rates (Table 2.2). Isolates were confirmed as *Y. ruckeri* using slide agglutination tests.

Table 2.2 Pathogenicity of isolates RD6 and RD124 towards Atlantic salmon and rainbow trout.

Infected carrier described as bacteria recovered from the hind gut post mortem, and cultured on ROD agar. Subsequently sub-cultured onto TSA and confirmed with agglutination tests. Infected survivor described as bacteria recovered from the head kidney post mortem, and cultured onto TSA. Isolates subsequently confirmed with agglutination tests.

		Fish sampled	Mortality (%)	Rate of infection (%)	Rate of carriage (%)
Atlantic salmon	RD6	15	0.0	20.0	33.3
	RD420.1	14	6.7	42.9	42.9
	RD420.2	13	13.3	53.8	61.5
Rainbow trout	RD6	25	0.0	16.0	24.0
	RD420.1	25	0.0	32.0	44.0
	RD420.2	25	0.0	12.0	4.0

The rate of infection and rate of carriage were different for RD6 and RD420 in Atlantic salmon and rainbow trout. Isolate RD6 did not kill any Atlantic salmon or rainbow trout. As this is an isolate of known virulence, this indicated flaws in our experimental design. However, RD6 was recovered from the kidney of 20% of the surviving Atlantic salmon and 16% of the surviving rainbow trout. It was recovered in the hind gut of a third (33.3%) of Atlantic salmon and nearly a quarter of rainbow trout (24%). However, isolate RD420 was recovered from the kidneys of substantially more Atlantic salmon (42.9 and 53.8% for replicates 1 and 2, respectively) than RD6 was. Similarly, it was recovered in the hind gut of

more Atlantic salmon (42.9 and 61.5% for replicates 1 and 2, respectively) than RD6 also. This indicates that isolate RD420, of novel serotype O1/O5 appears to be more pathogenic towards Atlantic salmon than the serotype O1 isolate, RD6. In rainbow trout, RD6 was recovered from the kidney of 16% of the surviving fish, while in the hind gut of 24% of fish. In contrast, isolate RD420 replicate 1 was recovered in the kidney of 32% of rainbow trout, and the hind gut of 44%. Isolate RD420 replicate 2 however, was recovered from the kidneys of only 12% and hind gut of only 4% of rainbow trout. Although these results are a little conflicting, they suggest that the novel serotype O1/O5 isolate (RD420) is not as pathogenic towards rainbow trout as the serotype O1 isolate (RD6).

Together these results suggest that the novel serotype O1/O5 may be more pathogenic towards Atlantic salmon than rainbow trout. However further experiments are necessary for confirmation.

2.4 Discussion

The majority of previously published characterisation studies of *Y. ruckeri* are based on isolates recovered predominantly from rainbow trout (Bastardo *et al.*, 2012a, b; Coquet *et al.*, 2002; Davies, 1990, 1991c, d; Tinsley, 2010; Tinsley *et al.*, 2011; Wheeler *et al.*, 2009). There have been very few such studies characterising *Y. ruckeri* isolates from Atlantic salmon (Bastardo *et al.*, 2011, 2012b; Davies, 1990, 1991a) and these have included only small numbers of isolates. In the present study, 109 isolates recovered from diseased Atlantic salmon and 26 isolates from diseased rainbow trout, farmed in the UK, were characterised by biotyping, serotyping, LPS analysis and comparison of OMP profiles.

Historically, the majority of isolates responsible for disease in the UK rainbow trout population have been of biotype 2 (i.e., non-motile and unable to hydrolyse Tween 20 and 80) (Austin *et al.*, 2003; Davies & Frerichs, 1989). The results from the isolates examined in the present study were in full agreement with previous findings and confirmed that biotype 2 serotype O1 isolates remain the principal cause of disease in UK rainbow trout. These isolates also represent a distinct OMP-type (OMP-type 1). However, two biotype 2 rainbow trout isolates were represented by the novel O1/O5 serotype, and two rainbow trout isolates were of biotype 1, serotype O1, OMP-type 3a.6 (these latter isolates represent the classical Hagerman strain). In contrast, the Atlantic salmon isolates were exclusively of biotype 1 (i.e., motile and able to hydrolyse Tween 20 and 80). These observations indicate that the switch from the biotype 1 to the biotype 2 phenotype that has occurred in European and North American rainbow trout isolates has not yet taken place in isolates recovered from UK Atlantic salmon. Tinsley *et al.* (2011) observed that the biotype 2 phenotype is not confined to serotype O1 isolates, having recovered a single serotype O7 isolate of biotype 2 (host species not named) in Denmark. These authors also identified apparent non-motile, Tween-hydrolysis positive isolates of serotype O5 in France, Iceland and Chile (Tinsley *et al.*, 2011). Wheeler *et al.* (2009) documented non-motile biotype 2 isolates recovered from the UK, Norway, USA and Spain (Wheeler *et al.*, 2009). In support of previous studies (Austin *et al.*, 2003; Davies, 1990), the UK biotype 2 isolates were recovered solely from rainbow trout and were of OMP-type 1; however, the USA and Spanish biotype 2

isolates were from rainbow trout and were of OMP-type 3 and a single Norwegian isolate was from Atlantic salmon and was of OMP-type 2.

An intriguing finding of this study was that the non-motile phenotype has not yet emerged in Atlantic salmon isolates in the UK. Welch *et al.* (2011) demonstrated that the loss of motility associated with the biotype 2 phenotype is caused by unique mutations in the flagella biosynthesis/export genes *fliR*, *flhA* and *flhB* (discussed in detail in section 1.2.2.3.1 of the Introduction). Repression and mutation of flagella biosynthesis genes also occurs in other pathogenic *Yersinia* species (Minnich & Rohde, 2007); *Y. enterocolitica* and *Y. pseudotuberculosis* repress expression of the flagella genes whereas *Y. pestis* has lost the ability to express them altogether. Artificial expression of flagellin in *Y. enterocolitica* completely attenuated virulence, supporting the hypothesis that motility is a liability in the mammalian host (Minnich & Rohde, 2007). Repression or loss of flagellin may be advantageous *in vivo* as it is a potent inducer of innate immunity (Minnich & Rohde, 2007). Common molecular structures in microorganisms activate the host innate immune system via TLRs (Akira *et al.*, 2006) - in particular, Toll-like receptor 5 (TLR5) recognises flagellin in many host species, including rainbow trout and Atlantic salmon (Cullender *et al.*, 2013; Hynes *et al.*, 2011; Raida & Buchmann, 2009). Stimulation of TLR5 ultimately leads to the maturation of antigen-presenting cells and the secretion of pro-inflammatory cytokines and chemokines (Didierlaurent *et al.*, 2004; Eaves-Pyles *et al.*, 2001; Gewirtz *et al.*, 2001; Tallant *et al.*, 2004; Yu *et al.*, 2003). Clearly, repression or loss of flagellin would provide a selective advantage to *Y. ruckeri,* through down regulation of the host immune response and could explain the emergence of non-motile isolates in rainbow trout. The TLR5 flagellin recognition system of fish is more sophisticated than that of mammals and comprises membrane (TLR5M) and soluble (TLR5S) isoforms; both isoforms have been identified in Atlantic salmon (Tsoi *et al.*, 2006) and rainbow trout (Tsujita *et al.*, 2004). TLR5S recognises flagellin in the fluid phase and binds to TLR5M, amplifying the signal cascade (Basu *et al.*, 2012); it has also been suggested that TLR5S is required in fish to systemically amplify the inflammatory response generated by TLR5M (Tsujita *et al.*, 2004). Clearly, the interaction of TLR5S and TLR5M, and subsequent triggering of the host innate immune response, is highly complex and perhaps sufficient variation exists in Atlantic salmon and rainbow

trout to provide a stronger selective pressure in the latter species that is driving loss of flagella. However, an alternative and more straight-forward explanation might be due to vaccination pressure on the bacterium. It has been suggested that flagellin present in the original vaccine preparation (based on formalin killed whole cells of a biotype 1 [motile] strain) may have elicited a protective immune response towards flagellin-producing biotype 1 isolates of *Y. ruckeri* (Arias *et al.*, 2007; Austin *et al.*, 2003; Scott *et al.*, 2013; Wheeler *et al.*, 2009). Thus, it is perhaps likely that the emergence of flagellin deficient biotype 2 strains has been driven by the selection of isolates capable of avoiding the immune response of the vaccinated host. *Yersinia ruckeri* has been associated with Atlantic salmon for a much shorter period of time than it has with rainbow trout and may simply not have had time to adapt. If this is the case, it is feasible that non-motile, biotype 2 isolates may emerge in Atlantic salmon in the near future and pose a more significant threat to this species.

The discovery of a new LPS type, represented by the O1/O5 serotype, was a highly significant outcome of the present study. It is clear from the silver-stained LPS profiles (Fig. 2.7) and Western-blotting (Fig. 2.8) that serotype O1 and O1/O5 LPS have common, if not identical, core polysaccharide regions. Although the O-antigen regions of each LPS type have very similar (but not identical) silver-stained ladder patterns (Fig. 2.7) they are, nevertheless, quite distinct by Western-blotting (Fig. 2.8). Therefore, it could be hypothesised that the O1/O5 LPS type has recently emerged from the O1 LPS type as a consequence of a horizontal gene transfer (HGT) and recombination event involving the O-antigen biosynthesis genes. The core polysaccharide region and O-antigen side chains of LPS are synthesised separately and are combined during translocation across the OM (Whitfield, 1995), as discussed in section 1.2.3.7 of the Introduction. A recombination event may have occurred whereby the O-antigen gene cluster of a serotype O1 isolate has been replaced by a new gene cluster responsible for synthesis of the O1/O5 O-antigen; while the core polysaccharide biosynthesis genes have remained unchanged. Similar exchange of O-antigen biosynthesis genes has occurred in other bacterial species such as *Vibrio cholerae* (Li *et al.*, 2002) and *E. coli* (Tarr *et al.*, 2000). The epidemic *V. cholerae* O139 Bengal strain originated from a seventh-pandemic O1 El Tor strain by homologous recombination-mediated exchange of O-antigen biosynthesis

clusters (Li *et al.*, 2002). Similarly, pathogenic serotype O55 and O157 *E. coli* clones arose through recombination of O-antigen synthesis genes (Tarr *et al.*, 2000). Examining a much larger range of isolates would allow us to explore this event in more detail, uncovering how far this serotype has spread geographically, into what species, and potentially pinpointing the point of emergence. This in turn would allow us to examine the rate of evolution of this pathogen and identify trends that may indicate what is driving the evolution.

It has been well documented that disease-causing isolates within the British rainbow trout population are biotype 2 variants of serotype O1 (Austin *et al.*, 2003; Davies & Frerichs, 1989; Davies, 1991d; Wheeler *et al.*, 2009). Isolates recovered over the last 14 years have not displayed any additional variation in this trait, as 22 of the 26 isolates examined were of biotype 2, serotype O1. However, the recovery of two biotype 2 serotype O1/O5 isolates from rainbow trout in 2010 and 2011 is a significant finding. These isolates may represent the emergence of a new biotype 2 clone expressing the novel O1/O5 O-antigen. The identification of the O1/O5 LPS type in biotype 2 strains recovered from rainbow trout provides further evidence for the HGT of the O1/O5 O-antigen biosynthesis gene cluster. In this case, the O1/O5 gene cluster may have been transferred to a biotype 2, serotype O1 rainbow trout isolate from a serotype O1/O5 Atlantic salmon isolate. The emergence of the new serotype O1/O5 clone in rainbow trout is a highly significant development, potentially having a serious impact on rainbow trout aquaculture. Current rainbow trout vaccines (section 1.3.7.2.1) are based on serotype O1 isolates and may not provide adequate protection against this emergent strain. On Atlantic salmon farms vaccinated with serotype O2 or O5 whole cell vaccines, outbreaks of serotype O1/O5 have occurred (Tim Wallis [Ridgeway Biologicals], personal communication). This indicates that LPS is a protective antigen. Further experimental data is required to confirm the protection afforded by these vaccines against this novel strain, and the pathogenicity of this strain towards this species.

In contrast to the very limited serotypic variation observed in rainbow trout isolates, more extensive serotypic diversity was identified among Atlantic salmon isolates. The majority (90/109) of isolates recovered from Atlantic salmon were of the previously unrecognised O1/O5 serotype (56 isolates) or the O2 serotype (34 isolates); five serotype O1 and 14 serotype O5 isolates were also

recovered. Serotypes O1/O5 and O2 were clearly the most commonly recovered serotypes in Atlantic salmon and together were responsible for the majority of infections. Presumably, these serotypes are more virulent in Atlantic salmon than are other serotypes. However, the prevalence of serotype O1/O5 and O2 isolates was very different over the 14-year period (Fig. 2.6). Serotype O2 isolates were recovered almost every year (with the exception of 2007, 2011 and 2014) since 2001 whereas with the exception of a single isolate in 2002, serotype O1/O5 isolates were not recovered until 2007. However, serotype O1/O5 was clearly the dominant serotype from 2007 to 2011. The emergence of the O1/O5 serotype within the Atlantic salmon population from 2007 may be due to serotypic conversion as a consequence of selective pressure resulting from widespread sero-specific vaccination. Such serotypic conversion as a consequence of vaccination pressure has occurred in the fish pathogen *Streptococcus iniae* (Bachrach *et al.*, 2001) and human pathogens including *Streptococcus pneumoniae* (DiNubile, 2012; Hanage *et al.*, 2007; Weinberger *et al.*, 2009), *Shigella flexneri* (Allison & Verma, 2000) and *Vibrio parahaemolyticus* (Chen *et al.*, 2011).

A possible explanation for the emergence of the novel O-serotype is the ability of bacteria to alter the antigenic properties of their cell surface. This occurs by two mechanisms: (i) antigenic variation leading to a change in composition of the structure, and (ii) phase variation, which involves a reversible gain or loss of an antigen. *Salmonella enterica* serotype Typhimurium expresses several antigens on the cell surface whose expression is regulated by phase variation, including flagellar subunits, type I fimbriae, and long polar fimbriae (Norris & Bäumler, 1999). Phase variation is a process of change in the expression of the epitopes of the cellular surface of the bacteria, in order to evade the immune response, and generate diversity within a population. Phase variation arises through either genetic (slipped-strand mispairing, recombination) or epigenetic (DNA methylation) mechanisms (Maskell *et al.*, 1993; Reyes *et al.*, 2012; Wang *et al.*, 2000). It is possible that within *Y. ruckeri*, the novel serotype O1/O5 has arisen through phase variation, where the O-antigen biosynthesis genes encoding for this LPS-type, have been switched 'on' in Atlantic salmon strains in order to overcome the immune response. A similar phenomenon has been reported in *H. pylori*, regarding Lewis antigens. Lewis antigens are fucosylated oligosaccharide

epitopes identified on the surface of certain eukaryotic and prokaryotic cells (structurally related to the determinants of the human ABH blood group system) (Wang *et al.*, 2000). It has been suggested that the Lewis antigens are continually changing and the expression of certain antigens on the bacterial surface is beneficial during colonisation and/or infection, while other antigens are expressed to mimic the host cells and evade the immune response. The primary culture of a gastric biopsy produces multiple genetically similar *H. pylori* isolates with altered LPS profiles, providing evidence for a dynamic process of Lewis antigen variation during infection by *H. pylori* (Wang *et al.*, 2000).

When isolates were grown under anaerobic conditions, growth was considerably slower and expression of OmpC was up-regulated (Fig. 2.9). The pore size of OmpC is much smaller than that of OmpF and, in an effort to avoid environmental stress, OmpC is upregulated (while OmpF is downregulated) to reduce harmful compounds entering the cell or cell contents leaking out (Batchelor & Walthers, 2005; Chowdhury *et al.*, 1996). Taking the differences in aeration into consideration, and the effect on OmpC in particular, the OMP profiles of aerobic-grown isolates in the present study were consistent with those obtained previously (Davies, 1991c) which were grown under conditions of low aeration (Fig. 2.9). Upon examination of the OMP profiles of all isolates, it became clear that isolates recovered from Atlantic salmon exhibited much greater diversity in terms of OMP profile than isolates recovered from rainbow trout; Atlantic salmon isolates were represented by OMP-types 2a, 2c and 3a (and eleven associated sub-types), whereas rainbow trout isolates comprised only OMP-type 1a (Fig. 2.11). It was also apparent that multiple serotypes are associated with specific OMP-types and -subtypes and that this was more apparent within Atlantic salmon isolates than rainbow trout isolates (Fig. 2.12). For example, serotypes O1, O2 and O5 are associated with OMP-subtype 2c.1 and serotypes O1, O5 and O1/O5 are associated with OMP-subtype 3a.2 (Fig. 2.12). These observations provide evidence that strains of the same OMP-type have likely acquired different LPS types by HGT. Historically, the majority of disease in rainbow trout in North America and Europe has been caused by isolates representing biotype 1, OMP-type 3a (the "Hagerman" strain) or biotype 2, OMP-type 1a (Davies, 1991c, d). Presumably, these represent hyper-virulent clones

that have become widely distributed in rainbow trout. The greater strain diversity observed in Atlantic salmon suggests that *Y. ruckeri* is not a recently acquired pathogen of this species. The data also suggest that infection in Atlantic salmon is not dominated by one or two hyper-virulent clones as is the case in rainbow trout. Rather, the enhanced diversity of Atlantic salmon isolates is more typical of a commensal bacterium, and opportunistic pathogen, that has been long established within this host species. A relationship between extensive genetic diversity and opportunistic pathogenicity has been proposed for *E. coli* strains (White *et al.*, 1990; Whittam, 1995). Nevertheless, it is clear that different strains are adapted to either Atlantic salmon or rainbow trout and that these strains have unique, host-specific OMP profiles. It is highly probable that Atlantic salmon and rainbow trout isolates have subtle differences in their OM proteomes that allow them to colonise and cause disease in these two host species.

Identification of novel serotype O1/O5 isolates as the main clonal group infecting UK Atlantic salmon highlighted the need to understand how pathogenic this strain is towards both Atlantic salmon and rainbow trout. Through i.p. challenge with an O1/O5 strain (RD420) into Atlantic salmon, it became apparent that this isolate was virulent towards Atlantic salmon; 17/20 fish having succumbed to infection within 3 days. The virulence shown by our positive control isolate (RD6) agreed with previous studies (Davies, 1991a; Haig *et al.*, 2011), and highlighted the validity of selecting this isolate as our positive control. This isolate is of serotype O1 and biotype 2 and is representative of the major clonal group that has been infecting rainbow trout for the last 30 years. The isolate RD6 is known to be virulent towards Atlantic salmon also, having killed Atlantic salmon in previous studies (60% of fry by i.p. challenge; 74% of fry and 63% of parr by bath challenge)(Haig et al, 2011). The twenty Atlantic salmon challenged with this strain in our challenge had all died within two days (Fig. 2.14).

Having successfully shown the virulence of this strain towards Atlantic salmon through i.p. challenge, it became important to examine the virulence through a more natural exposure route. However, the bath challenge experiments were only partially successful. The failure of the bath challenge can be explained through several factors upon comparison to previous studies ((Davies, 1991a;

Haig *et al.*, 2011)). The length of challenge time (1 h) was comparable to the study by Davies (1991a). However, the number of fish per bucket, the size of the bucket, and the volume of water were significantly different. While Davies challenged 30 fish in 5 l of water in an aerated bucket, we challenged 15 Atlantic salmon in 20 l of water in a 50 l bucket, and 25 rainbow trout in 50 l of water in a 100 l bucket (in accordance with ethical guidelines). Also, while Davies challenged rainbow trout at an average weight of 7.7 g, the fish we used were much larger at approximately 35 g and 80 g for Atlantic salmon and rainbow trout, respectively. While the challenge dose was able to cause infection as indicated by the recovery of bacteria from the kidney of both Atlantic salmon and rainbow trout, the fish may not have been sufficiently stressed in order to be overcome with the infection. A possible way to increase the level of stress would be to alter the temperature of the water during the infection. This may cause carriers to shed bacteria into the water increasing the likelihood of other fish becoming infected.

Haig *et al.*, (2011) used fish more comparable to ours (10 g Atlantic salmon, and 250 g rainbow trout), although the challenge time, at 4 h, was significantly longer.

In conclusion, this study has demonstrated that isolates of *Y. ruckeri* recovered from Atlantic salmon display much greater heterogeneity than those recovered from rainbow trout; they represent a wider range of O-serotypes (and associated LPS types) and have more diverse OMP profiles. The isolates recovered from rainbow trout were represented almost entirely by a single non-motile clone (biotype 2, serotype O1) that is responsible for the majority of ERM outbreaks in this species within the UK. Significantly, this clone was not associated with a single case of infection of salmon. Our findings also revealed the emergence of a new LPS type (O1/O5), which is particularly prevalent in Atlantic salmon. This LPS type has a core-polysaccharide region similar to that of type O1 LPS but has a novel O-antigen side-chain. These findings suggest that different serotype-specific vaccine formulations are likely to be required to protect against *Y. ruckeri* infection in rainbow trout and Atlantic salmon. However, the identification of this novel LPS type in two biotype 2 rainbow trout isolates strongly suggests that it has been acquired by HGT; these strains could represent a newly emerged threat to rainbow trout aquaculture. Thus, our findings suggest

that different *Y. ruckeri* strains are specifically adapted to cause disease in either Atlantic salmon or rainbow trout.

The virulence of the novel serotype O1/O5 towards Atlantic salmon has been demonstrated through i.p. challenge. While the bath challenges of Atlantic salmon and rainbow trout were not entirely successful, the recovery of bacteria from the kidney of both host species was encouraging as to the infectious capabilities of this strain. Isolate RD6 did not kill any Atlantic salmon or rainbow trout. However, it was recovered from the kidney of 20% of the surviving Atlantic salmon and 16% of the surviving rainbow trout. It was recovered in the hind gut of a third (33.3%) of Atlantic salmon and nearly a quarter of rainbow trout (24%). However, the recovery of isolate RD420 from the kidneys (42.9 and 53.8% for replicates 1 and 2, respectively) and hind gut (42.9 and 61.5% for replicates 1 and 2, respectively) of Atlantic salmon in greater numbers than from the kidneys (32 and 12% for replicates 1 and 2, respectively) and hind gut (44 and 4% for replicates 1 and 2, respectively) of rainbow trout, suggests that the novel serotype O1/O5 isolate (RD420) is not as pathogenic towards rainbow trout as the serotype O1 isolate (RD6).

The virulence shown by isolate RD420 through i.p. challenge towards Atlantic salmon was also encouraging. It is important that the virulence of this novel serotype towards Atlantic salmon and rainbow trout be determined. Crucially, the cross-protective capabilities of current serotype-specific vaccines against this emerging serotype must be examined.

Chapter 3 Bioinformatic and proteomic analysis of *Y. ruckeri* isolates recovered from Atlantic salmon and rainbow trout

3.1 Introduction

The composition of the Gram-negative cell envelope is discussed in detail in section (1.3.7). The important role that OMPs play in eliciting and sustaining infection in the host has also been discussed (1.4.3.4). The OM functions as a selectively permeable barrier, controlling the passage of nutrients into the bacterial cell. The role of OMPs in colonisation and virulence is critical as they are the first point of contact between pathogen and the host, while also playing important roles in the adaptations of the bacteria to different environmental conditions and host niches (Ruiz *et al.*, 2006).

The composition of the OM in different bacterial species can be diverse, although in general, will contain proteins with important specialist roles in biogenesis and maintaining the integrity of the OM, nonspecific porin activity, energy-dependent transport, adherence and enzymatic activities (Lin *et al.*, 2002). The major OMPs OmpA, OmpC and OmpF have been used in the implementation of a typing scheme in *Y. ruckeri*, again discussed previously (1.2.2.3.3).

The identification of OMPs is an initial step in exploring the proteins involved in host pathogen interactions and virulence. Bioinformatic predictors have been used successfully in several Gram-negative bacterial species for the identification of OMPs including *P. multocida* (E-komon *et al.*, 2012), *C. burnettii* (Flores-Ramirez *et al.*, 2014), *Ehrlichia ruminantum* (Moumène *et al.*, 2015), *F. columnare* (Liu *et al.*, 2008) and *Chlamydia sp.* (Aistleitner *et al.*, 2015). The study of the OM proteome of a number of bacterial species including *V. anguillarum* (Kao *et al.*, 2009), *E. tarda* (Kumar *et al.*, 2010), *N. meningitidis* (Williams *et al.*, 2007) and *A. salmonicida* (Zhang *et al.*, 2011b) has revealed important information about the function, pathogenicity and adaptation to environmental conditions of OMPs.

In the previous section (Chapter 2), the characterisation of the OMP profiles of 135 isolates of *Y. ruckeri*, recovered from Atlantic salmon (109 isolates) and rainbow trout (26 isolates) in the UK, was described. It became apparent that a restricted number of clonal groups were responsible for the majority of infection in both Atlantic salmon and rainbow trout. The dominant clonal groups associated with each species possessed unique OMP profiles based on molecular mass heterogeneity of the major OMPs, OmpA, OmpC and OmpF. Variation of minor OMPs was also apparent. These observations suggest that differences in the OM proteome of *Y. ruckeri* strains recovered from Atlantic salmon and rainbow trout might correlate with differences in host specificity and disease pathogenesis. The identity (and function) of the proteins in the OMP profiles of *Y. ruckeri* remains largely unknown. The identification of these proteins in tandem with an improved understanding of the variation that exists between strains will represent an important step in understanding the interactions between *Y. ruckeri* and its host.

The aim of the present study was to characterise and compare the OM proteome of different strains of *Y. ruckeri* recovered from Atlantic salmon and rainbow trout, using a combination of bioinformatic and proteomic approaches. Bioinformatic approaches were used to predict the proteins that were likely to be located at the OM of the bacteria, while a combination of gel-free and gel-based proteomic approaches were used to identify OMPs in representative isolates recovered from Atlantic salmon and rainbow trout.

While bioinformatics prediction methods can reveal which genes encode likely OMPs, they cannot indicate levels of expression. Proteomic experiments are therefore required in order to elucidate protein presence or absence. Using these methodologies the aim was to explore the OM proteome of *Y. ruckeri*, identify proteins involved in pathogenesis and virulence, and uncover proteins with potential roles in adaptation to either Atlantic salmon or rainbow trout. The effectiveness of different proteomic approaches for OMP identification was assessed.

3.2 Materials and Methods

3.3 Bioinformatics and proteomics

3.3.1 Bioinformatic prediction of genome encoded OMPs

The publicly available genomes of four isolates of *Y. ruckeri* (the properties of which are summarised in Table 3.1) were used for bioinformatic analyses. These four genomes were scrutinised using bioinformatic approaches according to the workflow described by E-Komon *et al*. (2012) to predict proteins which localise to the OM, with modifications based on programme availability.

Table 3.1 Properties of four *Y. ruckeri* genomes examined bioinformatically for the prediction of OMPs.
Genomes were obtained from public records on NCBI.

Strain	Host	Phenotype		Genome				
		Serotype	Biotype	Size (Mb)	GC%	Genes	Proteins	Date
ATCC 29473	RT	O1	1	3.77281	47.4	3457	3377	2014
CSF007-82	RT	O1	1	3.83052	47.4	3483	3352	2014
37551	AS	O1	1	3.77549	47.6	3466	3406	2014
YRB	RT	n/a	n/a	3.60522	47.5	3219	3079	2015

RT – Rainbow trout; AS – Atlantic salmon

Genomes were subjected to scrutiny through three categories of bioinformatic prediction software, using a total of eight prediction tools. Subcellular localisation predictors included PSORTb (Gardy *et al.*, 2005), CELLO (Yu *et al.*, 2006) and SOSUI-GramN (Imai *et al.*, 2008); ß-barrel predictors included TMBETADISC-RBF (Ou *et al.*, 2008), MCMMB (Bagos *et al.*, 2004) and BOMP (Berven *et al.*, 2004); and OM lipoprotein predictors included LipoP (Juncker *et al.*, 2003) and LIPO (Berven *et al.*, 2006). Subcellular localisation predictor Proteome Analyst v. 3.0 PA, and ß-barrel predictor TMB-Hunt (Garrow *et al.*, 2005) were no longer available and hence omitted from this study. A consensus prediction framework was followed whereby proteins that were predicted to (a) be localised to the OM by at least two subcellular localisation predictors, (b) have a ß-barrel conformation by at least two transmembrane ß-barrel predictors or (c) be OM lipoproteins by at least one lipoprotein predictor, were considered to be putative OMPs. In several instances, two predicted proteins within a genome were determined to constitute a single functional protein and were therefore grouped as such. A list of putative OMPs within the genome was produced by integrating the results from each of the prediction categories. The

putative OMPs were further scrutinised using additional domain, homology and literature searches to assign likely functions and to predict their subcellular localisations with greater confidence. Based on this further information, each putative OMP was assigned.

3.3.2 Gel-based proteomic analysis

Twenty micrograms of each OMP preparation were separated by 1D SDS-PAGE as described previously (2.2.4.9). Individual gel bands were excised and transferred to a 48-well plate. Gel pieces were washed three times in 100 µl of 50 mM ammonium bicarbonate, 50% (v/v) methanol and then twice in 100 µl of 75% (v/v) acetonitrile (ACN) before drying. Gel pieces were rehydrated with trypsin solution (20 µg trypsin/ml [Promega, V5111] in 20 mM ammonium bicarbonate) and incubated at 37°C overnight (approx. 12 h). Peptides were extracted by washing the gel pieces twice in 100 µl of 50% (v/v) ACN, 0.1% trifluoroacetic acid (TFA), before being transferred in solution to a 'V-bottom' 96-well plate and dried. Dried peptide samples were stored at -20°C until analysed by Nanoflow HPLC Electrospray Tandem Mass Spectrometry (nLC-ESI-MS/MS).

3.3.3 Gel-free proteomic analysis (methanol aided trypsin digestion)

OM fractions were directly digested with trypsin without prior separation by 1D SDS-PAGE using the methanol-aided trypsin digestion protocol as previously described by Bridges et al., 2008. Forty microlitres of 2 mg/ml protein was resuspended in 88 µl of 50 mM ammonium bicarbonate and placed in a sonicator bath for 20 min (with regular vortexing at 5 min intervals), before being incubated at 60°C for 20 min. Samples were placed on ice for 3 min before adding 120 µl of methanol and incubating for a further 5 min in a sonicator bath with regular vortexing. Thirty two microlitres of 20 µg/ml trypsin (Promega, V5111) in 25 mM ammonium bicarbonate was added followed by 120 µl of methanol. After vortexing briefly, samples were incubated at 37°C overnight (approx. 12 h). The digested samples were dried down in an Eppendorf SpeedVac vacuum centrifuge to approximately 15 µl and stored at -20°C until analysis by mass spectrometry.

3.3.4 Nanoflow HPLC Electrospray Tandem Mass Spectrometry (nLC-ESI-MS/MS)

Peptide samples prepared by both gel-free and gel-based methods were analysed by nLC-ESI-MS/MS. Peptides were solubilised in 2% acetonitrile with 0.1% TFA and fractionated on a nanoflow uHPLC system (Thermo RSLCnano) before online analysis by ESI-MS on an Amazon ion trap MS/MS (Bruker Daltonics). Peptide separation was performed on a Pepmap C18 reversed phase column (LC Packings), desalted and concentrated for 4 min on a C18 trap column followed by an ACN gradient (in 0.1% [v/v] formic acid) (3.2 – 32% [v/v] 4 - 27 min, 32% to 80% v/v 27 - 36 min, held at 80% [v/v] 36 - 41 min and re-equilibrated at 3.2%) for a total time of 45 min. A fixed solvent flow rate of 0.3 μl / min was used for the analytical column. The trap column solvent flow rate was fixed at 25 μl / min, using 2% ACN with 0.1% (v/v) TFA. Mass spectrometric analysis was performed at Glasgow polyomics using a continuous duty cycle of survey MS scan followed by up to ten MS/MS analyses of the most abundant peptides, choosing the most intense ions with multiple charges with dynamic exclusion for 120 s. MS data were processed using Data Analysis software (Bruker) and the automated Matrix Science Mascot Daemon server (v2.4.1). Protein identifications were assigned using the Mascot search engine to interrogate protein sequences in the *Y. ruckeri* NCBI Genbank database, allowing a mass tolerance of 0.4 Da for both MS and MS/MS analyses. The MASCOT program assigned a probability based MOWSE score to each protein. The identified proteins ($p \leq 0.05$) were significant if MOWSE scores were greater than 18.

3.3.5 Integration of experimentally identified and confidently predicted OMPs

Proteins identified by both gel-based and gel-free proteomic techniques were integrated and compared to the list of confidently predicted putative OMPs. The OMPs identified from eight representative isolates of *Y. ruckeri* were compared in relation to the animal host of origin.

3.4 Results

3.4.1 Bioinformatic prediction of proteins from four *Y. ruckeri* genomes

OMPs were predicted from four publicly available genomes of *Y. ruckeri* using eight different bioinformatic programs (section 3.3.1 of the Materials and Methods). These included rainbow trout strains ATCC29473, CSF007-82, YRB and Atlantic salmon strain 37551 (Table 3.1). The use of these eight predictors in combination predicted 97, 93, 102 and 88 putative OMPs from the rainbow trout strains ATCC29473, CSF007-82, YRB and Atlantic salmon strain 37551, respectively.

In total, 141 unique proteins were identified within the four genomes of *Y. ruckeri* (Table 3.2). Of these proteins, 48 were considered to represent the predicted OM core proteome (Fig. 3.1) (34.0% of total), as they were common to all genomes examined. These included the major OMPs OmpA, OmpC and OmpF as well as all members of the BAM complex (BamABCDE). Certain predicted OMPs were unique to individual genomes, although these numbers were relatively small in some cases; ATCC29473 (13 proteins; 9.2%), CSF007-82 (3; 2.1%), YRB (4; 2.8%) and 37551 (7; 2.2%). In the rainbow trout genomes examined 50 proteins were identified as the core OM proteome. However, two of these proteins were not identified in the Atlantic salmon genome. These proteins were identified as an OM ß-barrel protein (gi|490856998) and PilQ (gi|490858758). However, four proteins were identified uniquely in the Atlantic salmon genome. These included invasin 2 (gi|238705088), YebT (gi|238705953), YiiQ (gi|238707052) and TadD (gi|238706298).

Table 3.2 Predicted OMPs from four genomes of *Y. ruckeri.*

The OMPs predicted to be encoded in four genomes of *Y. ruckeri* are presented below. Red boxes indicate that protein was not identified through the prediction scheme in the specified genome, while green boxes indicate a positive identification, with the appropriate accession number indicated. Proteins are categorised by known function based on literature reviews. Columns labelled 1, 2 and 3 represent prediction by transmembrane β-barrel, subcellular localisation and lipoprotein predictors respectively. The total number of OMPs predicted in each genome is indicated at the bottom of table.

	Accession	Protein	Function	MW (Da)	Genome				1	2	3
					ATCC29473	CSF007-82	37551	YRB			
	1. OM biogenesis and integrity										
1	C4UJ19_YERRU	BamA/YaeT/Omp85	BAM complex	88100	gi|685059551	gi|731156702	gi|669788205	gi|755307797	+	+	+
2	C4UGF1_YERRU	BamB	BAM complex	42300	gi|685060357	gi|731156544	gi|669790837	gi|755308785	-	+	+
3	C4UMN4_YERRU	BamC/NlpB	BAM complex	38300	gi|685060479	gi|731156497	gi|669790880	gi|755308479	+	+	+
4	C4UNP8_YERRU	BamD/YfiO	BAM complex	27400	gi|685058394	gi|731154737	gi|669788556	gi|755306194	-	-	+
5	C4UGL8_YERRU	BamE/SmpA	BAM complex	12500	gi|685060222	gi|731156614	gi|669790767	gi|755307760	-	-	+
6	C4UL53_YERRU	LolB/HemM	Chaperone and lipoprotein transport	24200	gi|685058190	gi|731155467	gi|669790121	-	-	-	+
7	C4UFS2_YERRU	Lpp (Brauns)	Anchors peptidoglycan to OM	8400	gi|685060290	gi|731155796	gi|669789949	gi|755306936	-	-	+
8	C4ULP1_YERRU	LptD/Imp	LPS assembly	89700	gi|685059092	gi|731154507	gi|669791214	gi|755306471	+	+	+
9	C4UJT1_YERRU	LptE/RlpB	LPS assembly	21000	gi|685057941	gi|731154992	gi|669789281	gi|755307183	-	-	+
10	C4UHR6_YERRU	OmpA	OM integrity, porin, adherence	37800	gi|685058433	gi|731155311	gi|669790685	gi|755308394	+	+	+
11	C4UMF8_YERRU	Pal	Peptidoglycan associated	18100	gi|685058131	gi|731155065	gi|669788638	gi|755308162	-	-	+
12	C4UJ96_YERRU	Slp family	Structural and OM stabilisation	20600	gi|685058965	-	gi|669790174	gi|755307084	-	-	+
13	C4UFP9_YERRU	SlyB	Membrane integrity	15300	gi|685059839	gi|731155772	gi|669789929	gi|755308355	-	-	+
14	C4UG92_YERRU	VacJ/MlaA	Phospholipid homeostasis	28800	gi|685059698	gi|731156405	gi|669788025	gi|755306269	-	-	+
15	C4UF89_YERRU	YbhG	Membrane anchor protein	35700	-	gi|731155111	gi|669788724	-	+	+	+
	2. Transport and receptor										
16	C4UMU7_YERRU	Blc	Lipid storage/transport	20600	-	gi|731154283	-	-	-	-	+
17	C4UIS2_YERRU	BtuB	Vitamin B12 transport	69900	gi|685055844	gi|731157313	gi|669788411	gi|755306703	+	+	+
18	C4UGV8_YERRU	CsuB	Fimbrial chaperone (Type I pilus)	18600	gi|685052782	-	-	-	+	-	-
19	C4UGV7_YERRU	CsuC	Fimbrial chaperone (Type I pilus)	28700	gi|685052971	-	gi|669790238	gi|755307666	-	-	+
20	C4UGV6_YERRU	CsuD	Fimbrial usher (Type I pilus)	86400	gi|685052972	-	-	gi|755308591	+	+	-
21	C4UGV4_YERRU	CsuE-like	Surface-fimbriae tip adhesin	34500	-	gi|731153986	gi|669790240	gi|755306253	+	-	+
22	C4UG91_YERRU	FadL/OmpP1	Fatty acid/Hydrophobic compound trans.	46100	gi|685059974	gi|731156404	gi|669788026	gi|755308905	+	+	+
23	C4UM94_YERRU	FhuA	Ferric hydroxymate receptor	81700	gi|685058920	gi|731156904	gi|669788351	gi|755307928	+	+	+
24	C4UJX6_YERRU	FiuA	Iron transport	78400	gi|685058140	-	gi|669788980	gi|755306155	+	+	+
25	C4UIT5_YERRU	HasR	Iron transport	93100	gi|685055851	gi|731157328	gi|669788425	gi|755308584	+	+	-

Table 3.2 continued

	Accession	Protein	Function	MW (Da)	Genome				1	2	3
					ATCC29473	CSF007-82	37551	YRB			
26	C4UIQ8_YERRU	LamB	Maltoporin	47600	gi|685052613	gi|731157113	gi|669791036	gi|755308886	+	+	+
27	C4UIR2_YERRU	MalE	Maltose export	43800	gi|685052647	-	-	-	-	-	+
28	C4UFZ3_YERRU	OmpC.1	Porin (Small molecules)	40200	gi|685058722	gi|731156302	gi|669790669	gi|755309129	+	+	+
29	C4UKD7_YERRU	OmpC.2	Porin (Small molecules)	40000	gi|685059870	gi|731155424	gi|669790081	gi|755308567	+	+	+
30	C4UFF3_YERRU	OmpE	Inorganic phosphate uptake	39900	gi|685058036	gi|731155137	gi|669788699	gi|755307304	+	+	+
31	C4UHQ0_YERRU	OmpF	Porin (Small molecules)	41400	gi|685058318	gi|731155294	gi|669788850	gi|755307191	+	+	+
32	C4UNB2_YERRU	OmpL/KdgM	Porin (oligogalacturonides)	26700	gi|685058409	-	-	gi|755308758	+	+	+
33	C4UJH6_YERRU	OmpW	Small hydrophobic compound transport	23300	-	gi|731155612	gi|669789780	gi|755307222	+	+	+
34	C4UEU8_YERRU	OmpX	Virulence related	18400	gi|685057900	gi|731155147	gi|669788689	gi|755308685	+	+	+
35	C4UH51_YERRU	OprC	Copper Receptor	81400	gi|685052923	-	-	gi|755306267	+	+	+
36	C4UM23_YERRU	OprD	Chitoporin	48400	gi|685058306	-	-	gi|755307472	+	+	+
37	C4UFE1_YERRU	PhnE	Phosphonate ABC transporter	23300	-	gi|731155879	gi|669790030	-	-	+	+
38	C4UKR5_YERRU	ShuA	Heme receptor	74400	gi|685060527	gi|731155812	gi|669789964	gi|755306307	+	+	+
39	C4UJ18_YERRU	Skp	Chaperone	18300	gi|685059544	-	-	-	-	-	+
40	C4UNK8_YERRU	TamA/YtfN	Autotransporter assembly	64800	gi|685058056	gi|731154335	gi|669790469	gi|755306125	+	+	+
41	C4UHJ5_YERRU	TctC	Tricarboxylic transport	35900	-	gi|731154869	gi|669789068	gi|755307802	-	-	+
42	C4UKY9_YERRU	TolC	OM Channel (Efflux system)	52800	gi|685059183	gi|731156939	gi|669788387	gi|755306596	+	+	+
43	C4UFB5_YERRU	TonB dependent receptor	Iron transport	76600	gi|685058700	gi|731155097	gi|669788737	gi|755307185	+	+	+
44	C4UHF4_YERRU	Tsx	Colicin/phage receptor	29100	gi|685058676	gi|731154827	gi|669789029	gi|755306503	+	+	+
	3. Adherence										
45	C4UMZ1_YERRU	AidA	Adhesin	91700	gi|685058000	gi|731156983	gi|669790909	gi|755307003	-	-	+
46	C4UNB5_YERRU	Autotransporter	Adhesin	78200	gi|685058889	gi|731154411	gi|669788065	gi|755309020	+	+	+
47	C4UGD2_YERRU	FimA	Fimbrial protein	17900	gi|685058539	-	-	-	-	-	+
48	C4UGD1_YERRU	FimD	FimA export and assembly	96800	gi|685052781	gi|731156522	gi|669789235	gi|755306175	+	+	+
49	C4UL41_YERRU	PefC	Fimbrial usher	86700	gi|685057903	-	gi|669790109	-	+	+	+
50	C4UGF6_YERRU	PilF	Minor pilin	27300	gi|685060422	gi|731156549	gi|669790832	gi|755308853	+	-	+
51	C4UKM0_YERRU	PilP	Type IV pilus bio.	20600	gi|685053779	-	-	-	+	-	+
52	C4UKM1_YERRU	PilQ	Type IV pilus biogenesis	82500	gi|685052780	gi|731154139	-	gi|755307539	+	+	-
53	C4UKM6_YERRU	PilV	Minor pilin	42700	gi|685053742	-	-	-	+	-	-
54	C4UHR2_YERRU	PqiB/Mam7	Adherence	60600	-	gi|731155305	gi|669790679	-	+	-	-
55	C4UPA8_YERRU	ShlA/HecA/FhaA family	Heme utilization and adhesin	418400	gi|685058362	gi|731156978	gi|669790905	-	+	+	+
56	C4UHW8_YERRU	SteB	Fimbral protein	98200	gi|685057922	gi|731156134	gi|669789101	-	+	-	+
57	C4UMU3_YERRU	SteJ	Fimbrial usher	79700	gi|685058085	gi|731154314	-	-	+	+	+

Table 3.2 continued

	Accession	Protein	Function	MW (Da)	Genome				1	2	3
					ATCC29473	CSF007-82	37551	YRB			
58	C4UKN0_YERRU	TraL	F Pilin formation	15800	gi|685053786	-	-	-	-	-	+
59	C4UGS5_YERRU	Fimbral family protein	Fimbrial bioproteinsis usher	31400	gi|685058155	-	-	-	+	+	-
60	C4UJG6_YERRU	SfmD	Fimbrial usher	92300	gi|685059749	-	gi|669790857	gi|755307421	+	+	+
	4. Enzymatic activity										
61	C4UNV7_YERRU	BcsC	Cellulose biosynthesis	128800	gi|685052900	gi|731154040	gi|669789714	gi|755309011	-	-	+
62	C4UL77_YERRU	Hemolysin	Hemolysin	157800	gi|685058718	-	gi|669790143	gi|755306363	+	+	+
63	C4UP06_YERRU	LpoA	Peptidoglycan synthesis	70300	gi|685052689	gi|731157007	gi|669790932	gi|755306544	+	-	+
64	C4UHW9_YERRU	LpoB	Peptidoglycan synthesis	20600	gi|685058877	gi|731155366	gi|669790737	-	-	-	+
65	C4UIZ8_YERRU	MetQ	D-methionine binding	29300	gi|685059512	gi|731156679	gi|669788182	gi|755308344	-	-	+
66	C4UJ46_YERRU	MltA	Peptidoglycan maintenance	42900	gi|685059422	gi|731156730	gi|669791052	gi|755308938	+	-	+
67	C4UGA2_YERRU	MltB	Peptidoglycan maintenance	39200	gi|685060287	gi|731156418	gi|669788012	gi|755308684	+	-	+
68	C4UIJ5_YERRU	MltC	Peptidoglycan maintenance	39900	gi|685059461	gi|731156845	gi|669791064	gi|755306330	-	+	+
69	C4UJU6_YERRU	MltD	Peptidoglycan maintenance	52800	-	gi|731154743	-	gi|755307803	+	-	-
70	C4UKM4_YERRU	MltE	Peptidoglycan maintenance	23800	-	gi|731155853	-	-	-	-	+
71	C4UFR3_YERRU	NlpC	Copper homeostasis	25300	gi|685060195	gi|731156683	gi|669788186	gi|755308825	-	-	+
72	C4UIX9_YERRU	PepM37	Enzymatic activity	50500	-	gi|731157374	gi|669788469	-	+	+	+
73	C4UN64_YERRU	Phoslipase A1	Bacteriocin secretion	33700	gi|685053812	gi|731157257	gi|669789453	gi|755308294	+	+	+
74	C4UKB8_YERRU	PrtF	Protease secretion	50800	gi|685058530	-	gi|669790062	-	+	+	+
75	C4UIR9_YERRU	ShlB	Hemolysin activator	61900	gi|685052665	-	-	gi|755306726	+	+	+
76	C4UKU0_YERRU	Spr	Enzymatic activity	21700	gi|685059998	-	-	-	+	-	+
77	C4UMN9_YERRU	YfgC	Zinc Metalloprotease	53900	-	gi|731156502	gi|669790875	-	+	-	+
78	C4UIS0_YERRU	YhlA	Hemolysin	169600	-	gi|731157126	gi|669791048	-	+	-	+
79	C4UNG7_YERRU	Endopeptidase	Enzymatic activity	16600	-	-	gi|669788084	gi|755306717	+	-	+
	5. Motility										
80	A0A094V4E9_YERRU	FlaA	Flagellin	44000	gi|685060178	gi|731156087	gi|669790307	gi|755306665	+	-	-
81	C4UKI7_YERRU	FlgD	Flagellar hook assembly	24200	gi|685060359	gi|731156121	gi|669790338	-	+	-	-
82	C4UKI6_YERRU	FlgE	Links flagella to driver	44200	-	gi|731156120	gi|669790337	-	+	-	-
83	C4UKI4_YERRU	FlgG	Distal rod protein	27900	gi|685060177	gi|731156118	gi|669790335	-	+	-	-
84	C4UKI3_YERRU	FlgH	Flagella L-ring protein	25300	gi|685060129	gi|731156116	gi|669790334	gi|755307341	+	+	-
85	C4UKI2_YERRU	FlgI	Flagella P-ring protein	36700	gi|685060451	-	-	gi|755306525	+	-	+
86	C4UKI1_YERRU	FlgK	Hook associated - filament junction	58000	gi|685060018	gi|731156113	gi|669790331	gi|755307674	+	-	+
87	C4UKI0_YERRU	FlgL	Hook associated - filament junction	35200	gi|685059813	gi|731156112	gi|669790330	-	+	-	-
88	C4UKG1_YERRU	FliD	Flagellar filament cap	49100	-	gi|731156090	gi|669790308	gi|755308536	+	-	-

Table 3.2 continued

	Accession	Protein	Function	MW (Da)	Genome				1	2	3
					ATCC29473	CSF007-82	37551	YRB			
89	C4UKH4_YERRU	FliF	Hook length control	51300	-	gi|731156104	-	-	+	-	-
	6. Unknown										
90	C4ULV1_YERRU	Lipoprotein 2	Unknown	23300	gi|685059823	-	-	gi|755306932	-	+	+
91	C4T4V0_YERIN	Lipoprotein 1	Unknown	20200	-	-	-	gi|755306431	-	-	+
92	C4UHY4_YERRU	OM Beta barrel	Unknown	22800	gi|685059281	gi|731155382	-	gi|755308603	+	+	+
93	C4UNT0_YERRU	Smp	Unknown	25600	gi|685058941	-	-	gi|755307405	-	-	+
94	C4UHF6_YERRU	YajG	Unknown	21000	gi|685058261	-	-	gi|755306254	-	-	+
95	C4UM24_YERRU	YbfN/ChiQ	Unknown	12000	gi|685059005	-	-	gi|755306621	-	-	+
96	C4UL67_YERRU	YceB	Unknown	20500	-	-	-	gi|755306592	-	-	+
97	CSF007_7410	YcfL	Unknown	14100	-	gi|731155365	gi|669790736	-	+	-	+
98	C4UFI7_YERRU	YdgH	Unknown	34000	-	gi|731155707	gi|669789869	-	+	-	+
99	C4UNP0_YERRU	YdiY-like	Salt induced OM protein	28000	gi|685058723	gi|731154730	gi|669788563	gi|755306179	+	-	+
100	C4ULU5_YERRU	YebT	Unknown	95100	-	-	gi|669790023	-	+	+	-
101	C4UG12_YERRU	YfaZ	Unknown	18700	gi|685060439	gi|731156322	gi|669788867	gi|755306209	+	-	+
102	C4UN85_YERRU	YfeY	Unknown	20900	gi|685060131	-	-	gi|755306499	-	-	+
103	C4UKZ0_YERRU	YgiB	Biofilm formation	23900	-	-	gi|669788388	gi|755306281	-	-	+
104	C4UGV3_YERRU	YiaF	Unknown	25700	-	-	-	gi|755306620	-	-	+
105	C4UIW5_YERRU	YiiQ	Unknown	21800	-	-	gi|669788454	-	+	+	+
106	C4UMT6_YERRU	YjeI	Unknown	14200	gi|685052919	-	-	-	-	+	+
	7. Other										
107	C4ULV3_YERRU	CsgG	Curli organelles and biofilm formation	24000	gi|685059979	gi|731155881	gi|669790032	-	+	+	+
108	C4UIB8_YERRU	GspD	General secretion path.	71400	gi|685058319	gi|731154705	gi|669788587	gi|755309115	+	+	+
109	C4UNV8_YERRU	HmsP	Biofilm formation	75500	-	gi|731154041	gi|669789713	-	-	-	+
110	C4UN28_YERRU	Invasin 1	Invasion	93100	gi|685052862	gi|731154053	-	-	+	+	+
111	C4UP93_YERRU	Invasin 2	Invasion	69400	-	-	gi|669791050	-	+	+	+
112	C4UJC3_YERRU	MipA	MltA interacting	28000	gi|685058574	gi|731155552	gi|669790199	gi|755308017	+	+	+
113	C4UK25_YERRU	NlpI	Cell division	33800	gi|685057898	gi|731154368	gi|669790436	-	-	-	+
114	C4UJK4_YERRU	OsmB	Osmotically inducible	6800	-	gi|731155642	gi|669789808	gi|755308649	-	-	+
115	C4UN57_YERRU	OutS	Secretion	14500	-	gi|731154694	gi|669788598		-	-	+
116	C4UKV9_YERRU	RcpA	Pilus secretion	48300	-	gi|731156910	gi|669788357	-	-	+	+
117	A0A085U2W5_YERRU	RcsF	Phosphorelay signaling path	14500	gi|685059441	gi|731156680	gi|669788183	-	-	-	+
118	C4UJS4_YERRU	RlpA	Rod shape and daughter cell seperation	37600	gi|685059034	gi|731154985	gi|669789171	gi|755307715	+	-	+
119	C4UMD8_YERRU	TTSS protein	TTSS protein	45900	-	-	-	gi|755309125	+	-	-

Table 3.2 continued

	Accession	Protein	Function	MW (Da)	Genome				1	2	3
					ATCC29473	CSF007-82	37551	YRB			
120	C4UKW6_YERRU	TadD	Pilus assembly	28500	-	-	gi|669788364	-	-	-	+
121	C4UKN3_YERRU	TraN	F Plasmid conj. Transfer	35600	gi|685053751	-	-	-	-	-	+
122	C4UFJ9_YERRU	YdgA	Serine protease	29300	-	-	-	gi|755306197	-	-	+
123	C4UGU4_YERRU	YiaD/Omp16	Multicopy suppressor of BamD	22200	-	gi|731153975	gi|669790249	-	+	+	+
124	C4UFK4_YERRU	YnbE	YnbE-like	7400	gi|685059872	-	-	-	-	-	+
125	C4UP09_YERRU	Yrap	Phospholipid biogenesis	20200	-	gi|731157010	gi|669790935	gi|755308730	+	-	+
126	C4UME3_YERRU	YscC	Type III secretion OM pore	63000	-	gi|731154668	gi|669788620	gi|755306535	+	+	+
127	C4UME8_YERRU	YscJ	Yop protein export	27500	-	gi|731154663	gi|669788625	gi|755306200	+	-	+
	8. Hypothetical										
128	C4UIN2_YERRU	Hypothetical 12	Hypothetical	10800	-	-	-	gi|755306536	-	-	+
129	C4UFK3_YERRU	Hypothetical 10	Hypothetical	12200	-	gi|731155751	gi|669789909	-	+	-	+
130	C4UFM9_YERRU	Hypothetical 9	Hypothetical	12200	-	gi|731155723	gi|669789883	-	+	-	+
131	C4UM58_YERRU	Hypothetical 14	Hypothetical	22900	-	-	gi|669790753	gi|755306447	+	+	+
132	A0A085UA61_YERRU	Hypothetical 7	Hypothetical	22900	gi|685058168	-	gi|669790491	gi|755306482	+	+	+
133	C4UMI8_YERRU	Hypothetical 4	Hypothetical	24600	gi|685060141	-	-	-	+	-	+
134	C4UIF9_YERRU	Hypothetical 6	Hypothetical	26400	-	gi|731156807	gi|669791099	gi|755306551	+	-	+
135	C4UIJ0_YERRU	Hypothetical 8	Hypothetical	26500	-	gi|731156838	gi|669791070	-	+	-	+
136	C4UJP5_YERRU	Hypothetical 1	Hypothetical	30000	-	gi|731155684	gi|669789847	gi|755306436	+	+	+
137	A0A0B6FWT8_YERRU	Hypothetical 5	Hypothetical	33500	gi|685058205	gi|731154953	gi|669789142	gi|755306461	+	+	-
138	C4UNB4_YERRU	Hypothetical 13	Hypothetical	35300	gi|685058188	gi|731154410	gi|669788064	-	+	+	+
139	C4UJQ3_YERRU	Hypothetical 11	Hypothetical	50300	-	-	-	gi|755306223	+	-	+
140	C4UGY2_YERRU	Hypothetical 3	Hypothetical	75200	gi|685052820	-	gi|669789737	-	+	+	+
141	C4UGU5_YERRU	Hypothetical 2	Hypothetical	80800	gi|685052857	-	gi|669790248	-	-	+	+
				Total	**97**	**93**	**102**	**88**	**94**	**63**	**122**

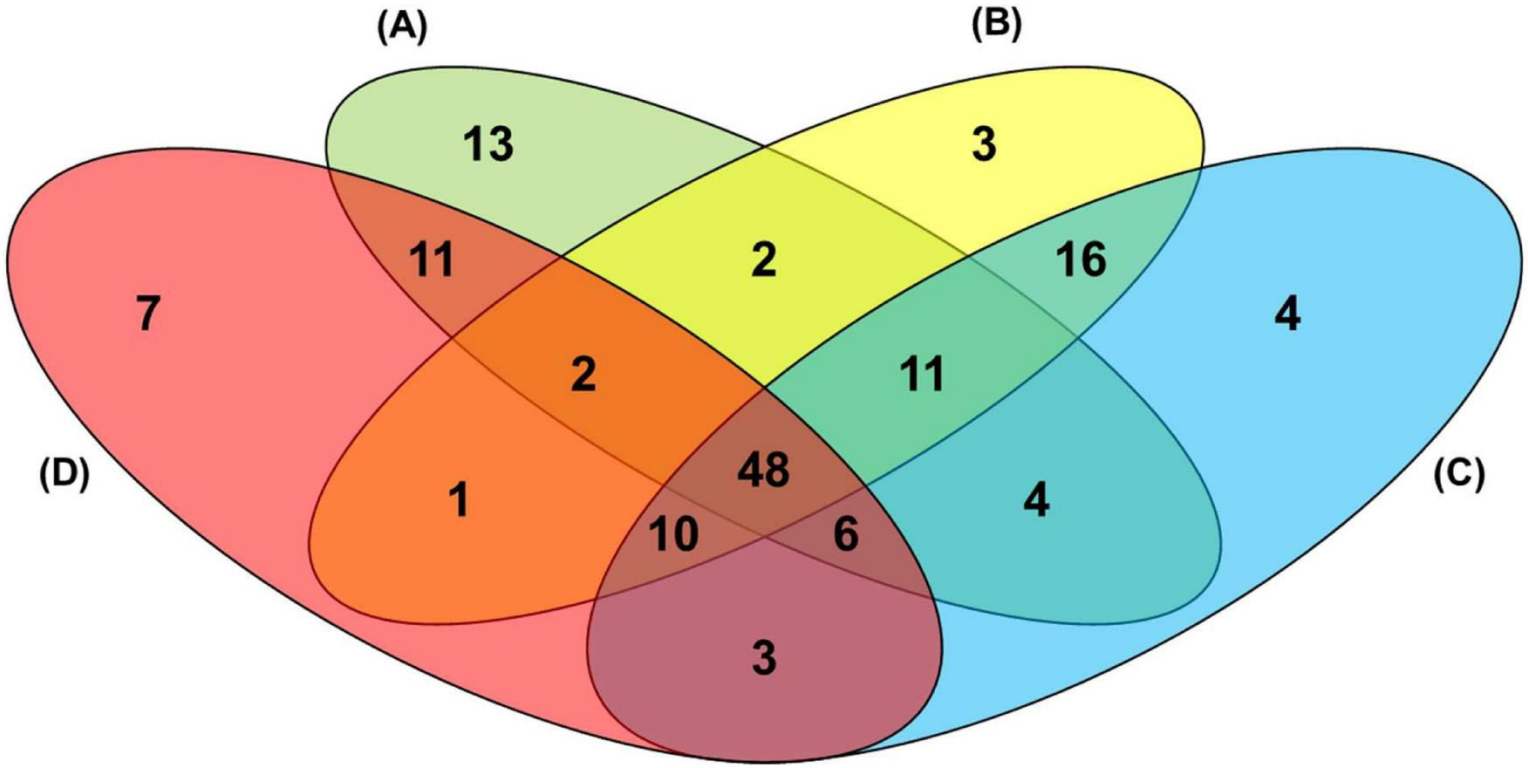


Fig. 3.1 Venn diagram comparing predicted OMPs from four *Y. ruckeri* genomes.
The genomes used were obtained from NCBI and represent strains **(A)** ATCC 29473; **(B)** CSF007-82; **(C)** 37551; **(D)** YRB. Forty eight proteins (34.0% of the total predicted proteins) were predicted to represent the core proteome, as these were identified in all isolates. In total, 141 unique OMPs were predicted within the four genomes.

As seen in Fig. 3.2, 49 proteins were identified by all three prediction methods, while 12 proteins were uniquely identified by the transmembrane ß-barrel predictors (at least two out of three programs), 40 proteins were uniquely identified by the lipoprotein predictors (either program), while no proteins were unique to the subcellular localisation programs. Seven proteins were commonly identified by the subcellular localisation and lipoprotein predictors, seven by both subcellular localisation and transmembrane ß-barrel programs, and 26 proteins were predicted by both the ß-barrel protein and lipoprotein programs. This study highlights the importance of combining different prediction methods in order to maximise the confidence of predicted proteins.

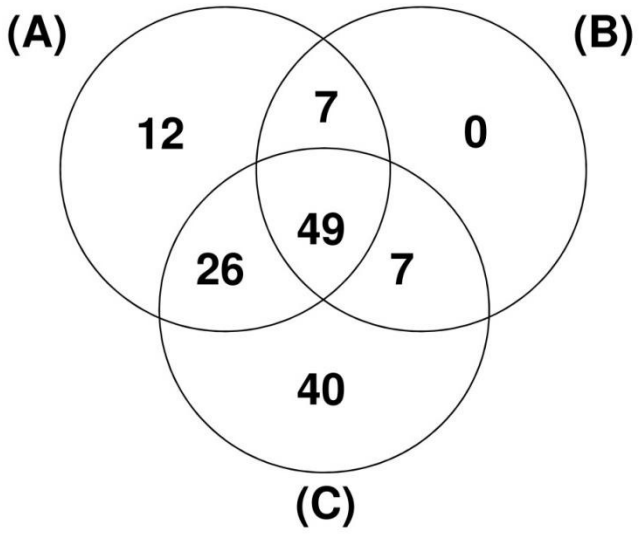


Fig. 3.2 Comparison of the number of proteins predicted by each of the three prediction methods.
Proteins were predicted by **(A)** β-barrel predictors including BOMP, TMBETADISC-RBF and MCMBB, **(B)** subcellular predictors including SOSUI GramN, CELLO and pSORTb v3.0 and **(C)** lipoprotein predictors including LIPO and LipoP. Forty nine proteins were identified by all three prediction methods, while 12 proteins where uniquely identified by transmembrane β-barrel predictors and 40 by lipoprotein predictors. No proteins were uniquely identified by subcellular localisation programs.

3.4.2 **Functional categorisation of the confidently predicted OMPs**

The specific functions of the 141 confidently predicted OMPs are shown in Table 3.2, although are summarised into functional categories in Table 3.3. The functions include roles in OM biogenesis and integrity (15 proteins; 12 core), transport and receptor (29 proteins; 15 core), adherence (16 proteins; 4 core), motility (10 proteins; 3 core), enzymatic activity (19 proteins; 8 core), motility (10 proteins; 3 core), other (21 proteins; 3 core) and proteins of unknown

function (17 proteins; 2 core). Fourteen proteins were determined to be hypothetical (14 proteins; 1 core). Hypothetical proteins represent areas of particular interest as determining the function and importance of these proteins could improve our understanding of the pathogenicity of the organism.

Table 3.3 Functional categorisation of predicted OMPs.
The number of proteins in each category is indicated, as well as the number of those proteins that were identified in all genomes (the core). The percentage of proteins considered core in each category is indicated in the final column.

Functional Category	Total	Core	Core (%)
1. OM Biogenesis and integrity	15	12	80.0
2. Transport and receptor	29	15	51.7
3. Adherence	16	4	25.0
4. Enzymatic activities	19	8	42.1
5. Motility	10	3	30.0
6. Other function	21	3	11.8
7. Unknown function	17	2	14.3
8. Hypothetical	14	1	7.1

3.4.3 Selection of representative isolates of *Y. ruckeri* for proteomic analysis

Having conducted a comprehensive characterisation study of the isolates of *Y. ruckeri* currently causing infection within UK Atlantic salmon and rainbow trout, different clonal groups that represent the major infective phenotypes in both species were identified (this is discussed in Chapter 2, previously). Within rainbow trout, two clonal groups were responsible for the majority of infection. These were isolates of serotype O1, biotype 1 and OMP-type 3a (representing the 'Hagerman' phenotype), and isolates of serotype O1, biotype 2 and OMP-type 1a (representing the non-motile variant first identified in the UK by Davies & Frerichs, 1989). The isolates causing infection in rainbow trout in our recent study did not differ from those investigated in previous studies in terms of their biotype, serotype and OMP-type (Davies & Frerichs, 1989; Davies, 1990, 1991c). Within Atlantic salmon, a much more diverse range of isolates were responsible for the infections. However, the major groups included isolates of serotype O1/O5, biotype 1 and OMP-type 3a; serotype O2, biotype 1 and OMP-type 2a; and serotype O5, biotype 1 and OMP-type 2c. This information is summarised in Table 3.4.

Table 3.4 Representative phenotypes of *Y. ruckeri* recovered from rainbow trout and Atlantic salmon populations within this study.

Biotype	Serotype	OMP-type	Number of Isolates	Representative of host species (%)
Atlantic salmon				
1	O1	2a	1	0.9
1		2c	2	1.8
1		3a	2	1.8
1	O2	2a	30	27.5
1		2c	3	2.8
1	O5	2a	1	0.9
1		2c	10	9.2
1		3a	2	1.8
1	O1/O5	3a	58	51.4
Rainbow trout				
1	O1	3a	2	7.7
2		1a	22	84.6
2	O1/O5	1a	2	7.7

As the isolates recovered from rainbow trout in our current study were phenotypically identical to those analysed in previous studies (Davies & Frerichs, 1989; Davies, 1990, 1991a, c), isolates from the previous studies were selected for further characterisation by proteomic analysis. Information regarding the virulence of these older isolates will supplement the proteomic analysis as we may identify proteins involved in virulence from known virulent isolates that are absent in known avirulent isolates.

Conversely, the isolates recovered from Atlantic salmon populations in our recent study are more appropriate for selection, and several have been used in a virulence study (Haig *et al.*, 2011).

Eight isolates (four from rainbow trout and four from Atlantic salmon) were selected as representative of the *Y. ruckeri* population and are presented in Table 3.5.

Table 3.5 Representative isolates from rainbow trout and Atlantic salmon selected for further proteomic characterisation.

Designation		Source			Phenotype		
Lab	Previous	Location	Host	Date	Biotype	Serotype	OMP type
RD6	-	Scotland	Rainbow trout	Pre 1990	2	O1	1b
RD28	BA2	U.K		Pre 1990	1	O5	2a
RD64	F53.1/82	West Germany		Pre 1990	1	O2	2a
RD124	851014	Denmark		Pre 1990	1	O1	3a
RD354	TW60/05	Scotland	Atlantic salmon	20/05/2005	1	O2	2a
RD366	TW90/05	Scotland		20/07/2005	1	O5	2c
RD382	FVG 269/06	Scotland		06/05/2006	1	O1	3a
RD420	TW110/08	Outer Hebrides		08/07/2008	1	O1/O5	3a

Isolate RD6 is representative of the non-motile biotype 2, serotype O1 variant that is the major cause of disease within UK rainbow trout populations. This isolate has been shown to display particular virulence to both rainbow trout and Atlantic salmon. Isolate RD124 is representative of the motile biotype 1, serotype O1 variant dubbed the 'Hagerman' strain, in rainbow trout. This isolate has also displayed virulence in both rainbow trout and Atlantic salmon, previously (Davies, 1991a; Haig *et al.*, 2011). Isolate RD382, recovered from Atlantic salmon, was selected to be directly comparible with isolate RD124. However, the phenotype of isolate RD382 (biotype 1, serotype O1, OMP-type 3a), did not represent a large group of isolates in the Atlantic salmon population (two isolates). Isolates RD354 and RD366 of serotypes O2 and O5, respectively, are representative of two major groups causing infection in UK Atlantic salmon currently. Isolates RD28 and RD64, recovered from rainbow trout were chosen for direct comparison with Atlantic salmon isolates of the same serotype. Finally, isolate RD420 is representative of the major clonal group found in UK Atlantic salmon, currently. This isolate is of the novel serotype O1/O5, of which very little is known.

3.4.4 Gel-free proteomic identification of the OMPs in eight representative isolates of *Y. ruckeri*

The total number of unique OMPs identified in the eight isolates using gel-free proteomic analysis was 65 (Table 3.6). The number of proteins identified in individual strains ranged from 33 in isolate RD382 to 46 in isolate RD6. Seven OMPs were identified in all three replicates of all isolates (BamB, Lpp [Brauns], Pal, OmpA, OmpF, BamA/YaeT and OsmY). These proteins were considered to represent the core OM proteome. Flagellin (FlaA) and the flagellar apparatus

Table 3.6 Proteins identified through gel-free and gel-based proteomic approaches in *Y. ruckeri*.
RT – Rainbow trout; AS – Atlantic salmon. Sixty five proteins were identified through gel-free analysis. Protein names coloured in green were identified only in RT isolates. Boxes coloured red, orange and yellow represent proteins identified in 3, 2 and 1 replicate, respectively. [1]Protein was only identified in rainbow isolates; [2]Protein was only identified in Atlantic salmon isolates.

	Accession	Protein	Function	MW (Da)	RT				AS				Prediction
					RD6	RD28	RD64	RD124	RD354	RD366	RD382	RD420	
1	gi|490855955	BamB	BAM complex	42,328	3	3	3	3	3	3	3	3	+
2	gi|490838180	Lpp (Brauns)	Anchors peptidoglycan to OM	6,687	3	3	3	3	3	3	3	3	+
3	gi|490860109	Pal	Peptidoglycan associated	18,103	3	3	3	3	3	3	3	3	+
4	gi|490856883	OmpA	OM integrity, porin, adherence	37,756	3	3	3	3	3	3	3	3	+
5	gi|300680152	OmpF	Porin (Small molecules)	40,049	3	3	3	3	3	3	3	3	+
6	gi|490857618	BamA/YaeT	BAM complex	88,077	3	3	3	3	3	3	3	3	+
7	gi|490857926	OsmY	Osmoregulation	11,017	3	3	3	3	3	3	3	3	-
8	gi|498532474	Flagellin	Flagellin	43,615	-	3	3	3	3	3	3	3	+
9	gi|490858688	FlgE	Links flagella to driver	44,281	-	2	3	3	3	3	3	3	+
10	gi|490855868	VacJ/MlaA	Phospholipid homeostasis	28,735	3	3	3	1	3	3	3	2	+
11	gi|490857575	MetQ	D-methionine binding	29,316	3	3	3	3	3	2	2	2	+
12	gi|490861349	YeeJ	Unknown	54,668	3	3	3	3	3	2	2	2	-
13	gi|490855735	TcyP	Cysteine transporter	48,096	3	3	3	2	1	2	1	1	-
14	gi|490860502	ShlB/FhaC/HecB family	Hemolysin activator	62,461	3	3	2	2	2	2	2	3	+
15	gi|490860274	BamC	BAM complex	38,311	3	2	3	3	3	1	1	1	+
16	gi|490855565	Pcp/SlyB	Membrane integrity	15,334	3	2	1	2	2	2	2	2	+
17	gi|490855768	OmpC.1	Porin (Small molecules)	40,484	3	2	-	-	-	1	-	-	+
18	gi|490855057	OmpE	Inorganic phosphate uptake	39,949	3	1	3	2	3	1	2	2	+
19	gi|490861198	BcsC	Cellulose biosynthesis	126,490	2	2	2	3	1	2	3	3	+
20	gi|490857772	Slp Family	Structural and OM stabilisation	20,645	3	-	-	3	2	1	1	1	+
21	gi|490861063	BamD	BAM complex	27,392	2	3	3	3	3	1	-	-	+
22	gi|490857817	MipA	MltA interacting	28,040	2	3	2	2	-	2	2	3	+

Table 3.6 continued

	Accession	Protein	Function	MW (Da)	RT				AS				Prediction
					RD6	RD28	RD64	RD124	RD354	RD366	RD382	RD420	
23	gi|490860153	YqjD	Unknown	26,277	2	3	-	1	1	1	2	1	-
24	gi|490856221	YiaD/Omp16	Multicopy suppressor of BamD	22,248	2	2	3	3	3	2	2	2	+
25	gi|490857673	MltA	Peptidoglycan maintenance	42,866	2	2	3	2	3	2	2	3	+
26	gi|490861316	C-terminal protease	Unknown	75,933	2	2	1	2	-	1	-	1	-
27	gi|490858151	LptE/RlpB	LPS assembly	21,089	2	1	2	2	2	1	-	3	+
28	gi|490861334	ShlA/FhaA/HecA family	Adhesin	155,563	2	1	2	-	3	-	-	-	+
29	gi|490855484	HslJ	Heat-inducible protein	15,597	2	1	1	3	2	2	1	2	-
30	gi|490858989	TolC	OM Channel (Efflux system)	53,080	2	1	1	1	3	1	1	2	+
31	gi|490856025	YfhG/QseG	Quorum sensing/Virulence	34,809	2	1	1	1	2	1	1	2	-
32	gi|490858084	Hypothetical 5	Unknown	31,358	2	1	-	2	1	1	-	-	+
33	gi|490859665	CsgG	Curli organelles/biofilm formation	24,670	2	-	2	-	1	-	-	-	+
34	gi|490860499	Hemolysin, partial[1]	Filamentous haemaglutinnin	48,764	2	-	-	2	-	-	-	-	-
35	gi|490860978	TamA/YtfN	Autotransporter assembly	63,376	2	-	-	1	3	-	1	1	+
36	gi|490858839	ShuA	Heme receptor	73,397	2	-	-	1	-	-	1	-	+
37	gi|490860076	TTSS protein[1]	Type III secretion protein	61,640	2	-	-	-	-	-	-	-	+
38	gi|490855443	Hypothetical 15[1]	Unknown	34,982	1	2	-	-	-	-	-	-	-
39	gi|490855637	HemR	Iron transport	72,952	1	1	1	1	1	1	1	-	-
40	gi|490857431	BtuB	Vitamin B12 transport	68,999	1	1	1	-	-	1	1	-	+
41	gi|490856454	OprC	Chitoporin	75,008	1	1	-	2	1	1	1	2	+
42	gi|490858586	OmpC.2	Porin (Small molecules)	39,971	1	1	-	1	-	1	2	1	+
43	gi|490856874	PqiB/MAM7	Adherence	59,670	1	-	1	-	1	-	-	-	+
44	gi|490856945	Hypothetical 16	Endonuclease/Exonuclease	56,471	1	-	1	-	-	-	-	1	-
45	gi|490861264	LpoA/LppC[1]	Peptidoglycan synthesis	74,138	1	-	-	1	-	-	-	-	+
46	gi|490859532	LptD[1]	LPS assembly	89,597	1	-	-	-	-	-	-	-	+
47	gi|490859521	Hypothetical 17[1]	Membrane associated	17,747	1	-	-	-	-	-	-	-	-

Table 3.6 continued

	Accession	Protein	Function	MW (Da)	RT				AS				Prediction
					RD6	RD28	RD64	RD124	RD354	RD366	RD382	RD420	
48	gi|490859968	FhuA[1]	Ferric hydroxymate receptor	81,661	1	-	-	-	-	-	-	-	+
49	gi|490861267	YraP	Phospholipid biogenesis	20,176	-	1	2	3	1	1	-	2	+
50	gi|490857535	Pep M37[1]	Enzymatic activity	50,517	-	1	-	-	-	-	-	-	+
51	gi|490855259	Lipoprotein 1[1]	Unknown	20,214	-	1	-	-	-	-	-	-	+
52	gi|490860414	Blc[1]	Lipid storage/transport	20,700	-	1	-	-	-	-	-	-	+
53	gi|490858640	FliD	Flagellar filament cap	49,084	-	-	2	-	2	1	-	-	+
54	gi|490857051	Hypothetical 18[1]	Unknown	54,586	-	-	1	1	-	-	-	-	-
55	gi|490856645	YajG[1]	Unknown	20,929	-	-	1	-	-	-	-	-	+
56	gi|490857925	OmpW[1]	Hydrophobic compound transport	23,880	-	-	1	-	-	-	-	-	+
57	gi|490857249	MltC[1]	Peptidoglycan maintenance	39,902	-	-	1	-	-	-	-	-	+
58	gi|490856303	Hypothetical 19[1]	Unknown	51,997	-	-	1	-	-	-	-	-	-
59	gi|46369605	RupA[2]	Unknown	17,892	-	-	-	-	-	1	-	-	-
60	gi|490860631	SecD[1]	General secretion path.	71,754	-	-	-	1	-	-	-	-	+
61	gi|490860818	DNA Circulation protein[1]	Unknown	50,502	-	-	-	1	-	-	-	-	-
62	gi|490855867	FadL[2]	Fatty acid compound transport	46,063	-	-	-	-	1	-	-	2	+
63	gi|490859171	TpsB family[2]	Hemolysin activator	61,725	-	-	-	-	-	-	1	2	-
64	gi|490858139	RlpA[2]	Rod shape cell seperation	37,557	-	-	-	-	-	-	-	1	+
65	gi|490856067	Phospholipase A1[2]	Bacteriocin secretion	32,533	-	-	-	-	-	-	-	1	+
				Total	**46**	**40**	**40**	**41**	**37**	**38**	**33**	**37**	**52**

protein FlgE were identified in the seven isolates that were motile; they were not present in the non-motile isolate RD6. Proteins identified in at least one replicate of all eight isolates increased the identified core proteome to 23 OMPs. In addition to the above, these proteins included VacJ/MlaA, MetQ, YeeJ, TcyP, ShlB/FhaC/HecB family, BamC, Pcp/SlyB, OmpE, BcsC, YiaD/Omp16, MltA, HslJ, TolC and YfhG/QseG. Furthermore, several proteins with putative roles in iron acquisition were identified including ShuA, HemR and FhuA, although not routinely in all isolates and replicates. ShuA was identified in isolate RD6 (two replicates), RD124 (one replicate) and RD382 (one replicate). These three isolates were all of serotype O1. However, while isolates RD124 and RD382 represented biotype 1 and OMP-type 3a, isolate RD6 was of biotype 2 and OMP-type 1b. HemR was identified in a single replicate of all isolates except isolate RD420, and the protein FhuA was only identified in a single replicate of isolate RD6. However, growth conditions were not optimal for the expression of these OMPs although it is still of note that these proteins should be identified in a media that is not iron-depleted.

Two forms of OmpC were identified through both bioinformatic (in all genomes) and proteomic approaches (in all isolates). These proteins have hence been labelled as OmpC.1 and OmpC.2 (with the molecular weights of these proteins being 40,484 Da [KGA49862.1] and 39,971 Da [KGA51002.1] in isolate ATCC29473, respectively). Conducting BLAST alignment of these proteins shows only a 54% identity to each other, however BLAST analysis of OmpC.1 (gi|238706508) returns a 100% sequence homology to OmpC of *Y. ruckeri* strain ATCC 29473, while OmpC.2 (gi|238707949) is 99% similar in sequence to the "porin" of strain ATCC 29473. Upon blasting the porin of ATCC 29473, it appears very similar to OmpC of other *Yersinia* species.

Seventeen proteins were identified solely in rainbow trout isolates. These included a partial hemolysin, LpoA/LppC, SecD, a DNA circulation protein, YajG, OmpW, MltC, PepM37, lipoprotein 1, Blc, a TTSS protein, LptD, FhuA, and four hypothetical proteins (hypothetical proteins 15, 17, 18 and 19). However, in many cases these were identified in only a single replicate with the exception of the TTSS protein (RD6; two replicates), a partial hemolysin (RD6 and RD124; two replicates) and hypothetical protein 15 (gi|490855443; RD6 and RD28; one and

two replicates, respectively). Five proteins were unique to isolates recovered from Atlantic salmon. These included a TpsB-family protein, FadL, RlpA, Phospholipase A1, and RupA although again these were usually present in only one replicate with the exceptions of FadL (RD354 and RD420; one and two replicates, respectively) and a TpsB family protein (RD382 and RD420; one and two replicates, respectively). The proteins RlpA and Phospholipase A1 were both identified in single replicates of isolate RD420, while RupA was identified in a single replicate of isolate RD366.

3.4.5 Gel-based proteomic identification of the OMPs in representative isolate RD366

In order to save time and resources, a single isolate, RD366, was selected for thorough gel-based proteomic analysis. This isolate was selected as the OM proteome profile comprised the greatest number of visible bands through SDS-PAGE analysis. Firstly, this isolate was separated using a mini-gel format and the entire lane excised into 17 equal fractions (3 mm each). This encompassed both the bands and 'interband' regions, and was completed in triplicate. A total of 47 proteins were identified (Fig. 3.3).

Proteins labelled in red were uniquely identified using the gel-based method.
Out of 47 proteins identified, 29 of these were identified in all replicates, while 39 proteins were identified in at least two replicates (Fig. 3.4), highlighting the reproducibility of this technique. In replicate one a Slp family protein, LptE/RlpB and NlpI were unique. In replicate two, a membrane protein, PilV, BamE and NlpC were uniquely identified. Finally, in replicate three the OMPs MltA and YdgA were uniquely identified.

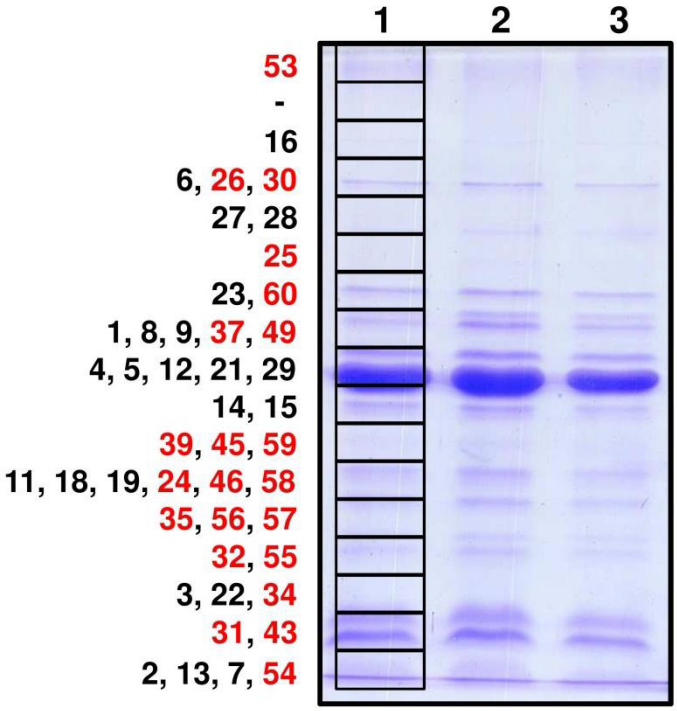


Fig. 3.3 The excised fractions from the OMP profiles of Atlantic salmon isolate RD366 after separation by 1D SDS-PAGE (mini-gel format).
The regions of gel excised are indicated by boxes in lane 1. The gel lane was separated into 17 sections (3 mm each). Identified OMPs are labelled numerically and correspond to those in Table 3.7. Proteins labelled in red were uniquely identified using gel based methods.

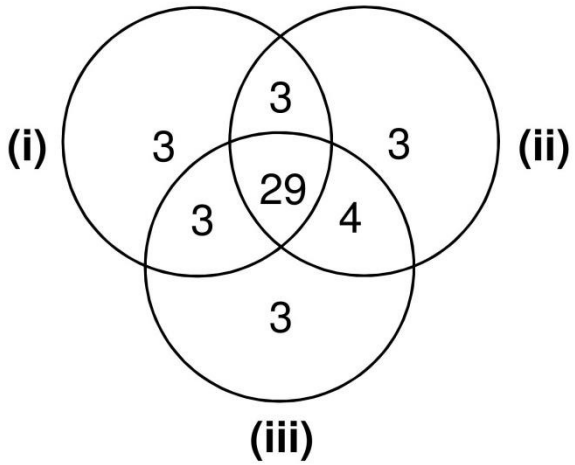


Fig. 3.4 Comparison of the OMPs identified in replicates of isolate RD366 using a mini-gel SDS-PAGE format.
Twenty nine proteins were identified in all three replicates of isolate RD366. However, 39 OMPs were identified in at least two out of three replicates. Replicates 1, 2 and 3 are represented by (i, ii and iii).

Eighteen bands were visible in each replicate of isolate RD366 when separated by the large-gel format (Fig. 3.5). This allowed for the identification of 41 individual OMPs (Table 3.7). Again, the reproducibility of this method is highlighted in Fig. 3.6, as 29 proteins were identified commonly to all 3 replicates. In replicate one hypothetical protein 13 and a membrane associated protein were unique. In replicate two a transporter and YegD/Hsp70 protein were uniquely identified. Finally, in replicate three the OMPs NlpD, MetQ and LpoB were uniquely identified.

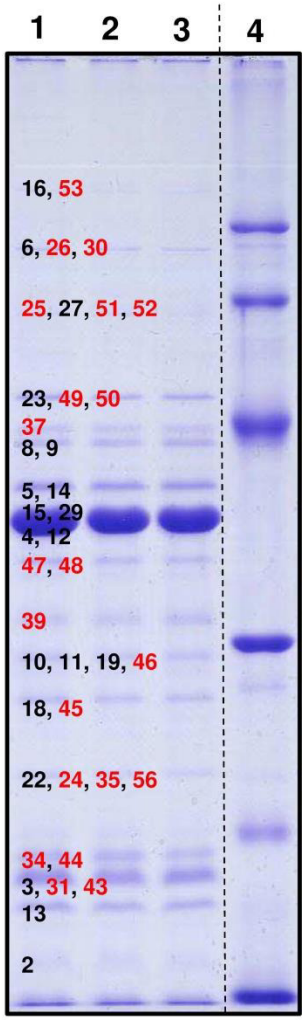


Fig. 3.5 The excised bands from the OMP profiles of Atlantic salmon isolate RD366 after separation by 1D SDS-PAGE (large-gel format).
Replicates 1, 2 and 3 are indicated by lanes 1, 2 and 3, respectively. The OMPs identified are indicated in lane 1. Identified OMPs are labelled numerically and correspond to those inTable 3.7. Proteins labelled in red were uniquely identified using gel based methods.

Combining the list of proteins identified in isolate RD366 by excising both the individual bands of the large-gel and the entire gel lane of the mini-gel allowed for a total of 54 unique proteins to be identified. Thirty five proteins were identified using both techniques. However, the proteins YegD/Hsp70, FliF,

Table 3.7 Comparison of proteins identified through gel-based proteomic approaches in *Y. ruckeri* isolate RD366.
Fifty three unique proteins were identified through combined mini-gel and large gel format gel-based proteomic analyses. 'Mini-gel' refers to isolate RD366 run on a mini gel format and the entire gel lane excised into 18 fractions (3 mm each). 'Large-gel' refers to the excised bands of isolate RD366 from the large-gel SDS-PAGE format. Boxes coloured red, orange and yellow represent proteins identified in 3, 2 and 1 replicate, respectively. Numbers in boxes correspond to position of protein on SDS-PAGE gels in Fig. 3.3 and Fig. 3.8.

	Accession	Protein	Function	MW (Da)	RD366		Prediction
					Mini-gel	Large-gel	
1	gi|490855955	BamB	BAM complex	42,328	1	-	+
2	gi|490838180	Lpp (Brauns)	Anchors peptidoglycan to OM	6,687	2	2	+
3	gi|490860109	Pal	Peptidoglycan associated	18,103	3	3	+
4	gi|490856883	OmpA	OM integrity, porin, adherence	37,756	4	4	+
5	gi|300680152	OmpF	Porin (Small molecules)	40,049	5	5	+
6	gi|490857618	BamA/YaeT	BAM complex	88,077	6	6	+
7	gi|490857926	OsmY	Osmoregulation	11,017	7	-	-
8	gi|498532474	Flagellin	Flagellin	43,615	8	8	+
9	gi|490858688	FlgE	Links flagella to driver	44,281	9	9	+
10	gi|490855868	VacJ/MlaA	Phospholipid homeostasis	28,735	10	10	+
11	gi|490857575	MetQ	D-methionine binding	29,316	11	11	+
12	gi|490860274	BamC	BAM complex	38,311	12	12	+
13	gi|490855565	Pcp/SlyB	Membrane integrity	15,334	13	13	+
14	gi|490855768	OmpC.1	Porin (Small molecules)	40,484	14	14	+
15	gi|490855057	OmpE	Inorganic phosphate uptake	39,949	15	15	+
16	gi|490861198	BcsC	Cellulose biosynthesis	126,490	16	16	+
17	gi|490857772	Slp family	Structural and OM stabilisation	20,645	17	-	+
18	gi|490861063	BamD	BAM complex	27,392	18	18	+
19	gi|490857817	MipA	MltA interacting	28,040	19	19	+
20	gi|490857673	MltA	Peptidoglycan maintenance	42,866	21	-	+
21	gi|490858151	LptE/RlpB	LPS assembly	21,089	22	22	+
22	gi|490858989	TolC	OM Channel (Efflux system)	53,080	23	23	+
23	gi|490859665	CsgG	Curli organelles/biofilm formation	24,670	24	24	+

Table 3.7 continued

	Accession	Protein	Function	MW (Da)	RD366		Prediction
					Mini-gel	Large-gel	
24	gi|490860978	TamA/YtfN	Autotransporter assembly	63,376	25	25	+
25	gi|490858839	ShuA	Heme transport	73,397	26	26	+
26	gi|490857431	BtuB	Vitamin B12 transport	68,999	27	27	+
27	gi|490856454	OprC	Chitoporin	75,008	28	-	+
28	gi|490858586	OmpC.2	Porin (Small molecules)	39,971	29	29	+
29	gi|490859532	LptD	LPS assembly	89,597	30	30	+
30	gi|490859521	Hypothetical 20	Membrane associated	17,747	31	31	-
31	gi|490855259	Lipoprotein 1	Unknown	20,214	32	-	+
32	gi|490856645	YajG	Unknown	20,929	34	34	+
33	gi|490857925	OmpW	Hydrophobic compound transport	23,880	35	35	+
34	gi|490855867	FadL	Fatty acid compound trans.	46,063	37	37	+
35	gi|490856067	Phos. A1	Bacteriocin secretion	32,533	39	39	+
36	gi|238708334	OmpX	Virulence	17,086	43	43	+
37	gi|490856978	LpoB	Peptidoglycan synthesis	20,698	-	44	+
38	gi|490861050	Membrane protein	Unknown	27,916	45	45	-
39	gi|490856641	Tsx	Nucleoside channel/Colicin rec.	29,122	46	46	+
40	gi|238707352	NlpD	Unknown	34,142	-	47	-
41	gi|685058188	Hypothetical 13	Unknown	34,471	-	48	+
42	gi|685053742	PilV	Major Pilin	48,727	49	49	+
43	gi|490855905	YegD/Hsp70	Heat shock protein	49,742	-	50	-
44	gi|490859142	Transporter	OM Transport	60,560	-	51	-
45	gi|740408485	FliF	Flagellar machinary	61,743	-	52	+
46	gi|490855624	GlpC	Iron-Sulfur protein	115,110	53	53	-
47	gi|740407770	BamE	BAM complex	12,651	54	-	+
48	gi|740408670	Lipoprotein 3	OM lipoprotein	11,336	55	-	-
49	gi|755306536	Hypothetical 12	Unknown	11,794	56	56	+
50	gi|685060129	FlgH	Flagella L-ring protein	13,149	57	-	+

Table 3.7 continued

	Accession	Protein	Function	MW (Da)	RD366		Prediction
					Mini-gel	Large-gel	
51	gi|685060195	NlpC/CutF	Copper homeostasis	25,530	58	-	+
52	gi|685057898	NlpI	Cell division	33,918	59	-	+
53	gi|755308840	YdgA	GTP-binding protein	54,395	60	-	+
				Total	**47**	**41**	**46**

hypothetical protein 13, LpoB, NlpD and a transporter were unique to excising individual bands from the large-gel, and the proteins BamB, OsmY, Slp family, MltA, OprC, OM lipoprotein (gi|490855259), BamE, OM lipoprotein (gi|740408670), FlgH, NlpC/CutF, NlpI and YdgA were unique to excising the entire lane. The proteins that were unique to the mini-gel format corresponded mainly with 'interband' regions based on their MW, so it is unlikely that these would have been identified through excision of only visible bands.

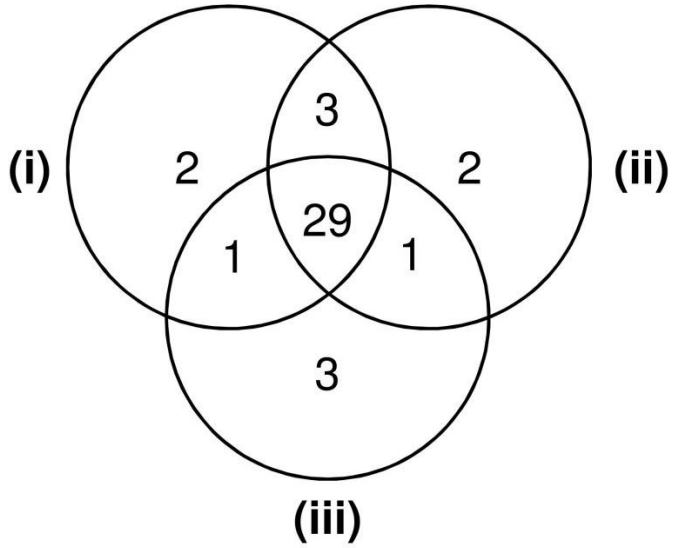


Fig. 3.6 Comparison of the OMPs identified in replicates of isolate RD366 using a large-gel SDS-PAGE format.
Twenty nine proteins were identified in all three replicates of isolate RD366. However, 34 OMPs were identified in at least two out of three replicates. Replicates 1, 2 and 3 are represented by (i, ii and iii).

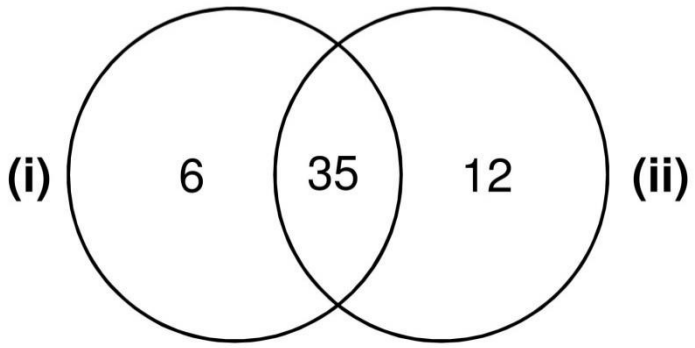


Fig. 3.7 Comparison of the OMPs identified in isolate RD366 using the mini and large gel SDS-PAGE format.
Thirty five proteins were identified in isolate RD366 by gel-based analysis of both the mini and large-gel format. Six proteins were identified solely by excision of the individual bands in the large-gel format **(i)**. Twelve proteins were identified solely by excision of the entire gel lane in the mini-gel format **(ii)**.

3.4.6 Identification of proteins in addition to the reference isolate using gel-based proteomics

In the remaining seven isolates, bands that appeared to represent different proteins to that of the reference isolate (RD366; Fig. 3.8C, lanes 4, 5 and 6) were excised from the large-format gel and subjected to proteomic identification. The major protein bands representing OmpA, OmpC and OmpF were excised in all isolates in order to provide confirmation as to the identity of each protein in the OMP-typing scheme (Davies, 1991c). The position of OmpA, OmpC and OmpF was confirmed in each isolate, endorsing the typing schemes reliability.

In total, four proteins were identified in addition to isolate RD366 (Table 3.8). These proteins were YiaD/Omp16 (Isolates RD28; Fig. 3.8A, lanes 4, 5 and 6, RD382 and RD420; Fig. 3.8D, lanes 1 to 6), RlpA (isolate RD420; Fig. 3.8D, lanes 4, 5 and 6), Surface Ag (isolate RD420; Fig. 3.8D, lanes 1, 2 and 3) and MalA (isolate RD28; Fig. 3.8A, lanes 4, 5 and 6, RD420; Fig. 3.8D, lanes 1, 2 and 3). Of these, YiaD/Omp16 and RlpA were also identified using gel-free methods. However, these proteins were not identified when examining the entire lane of isolate RD366 (Fig. 3.3).

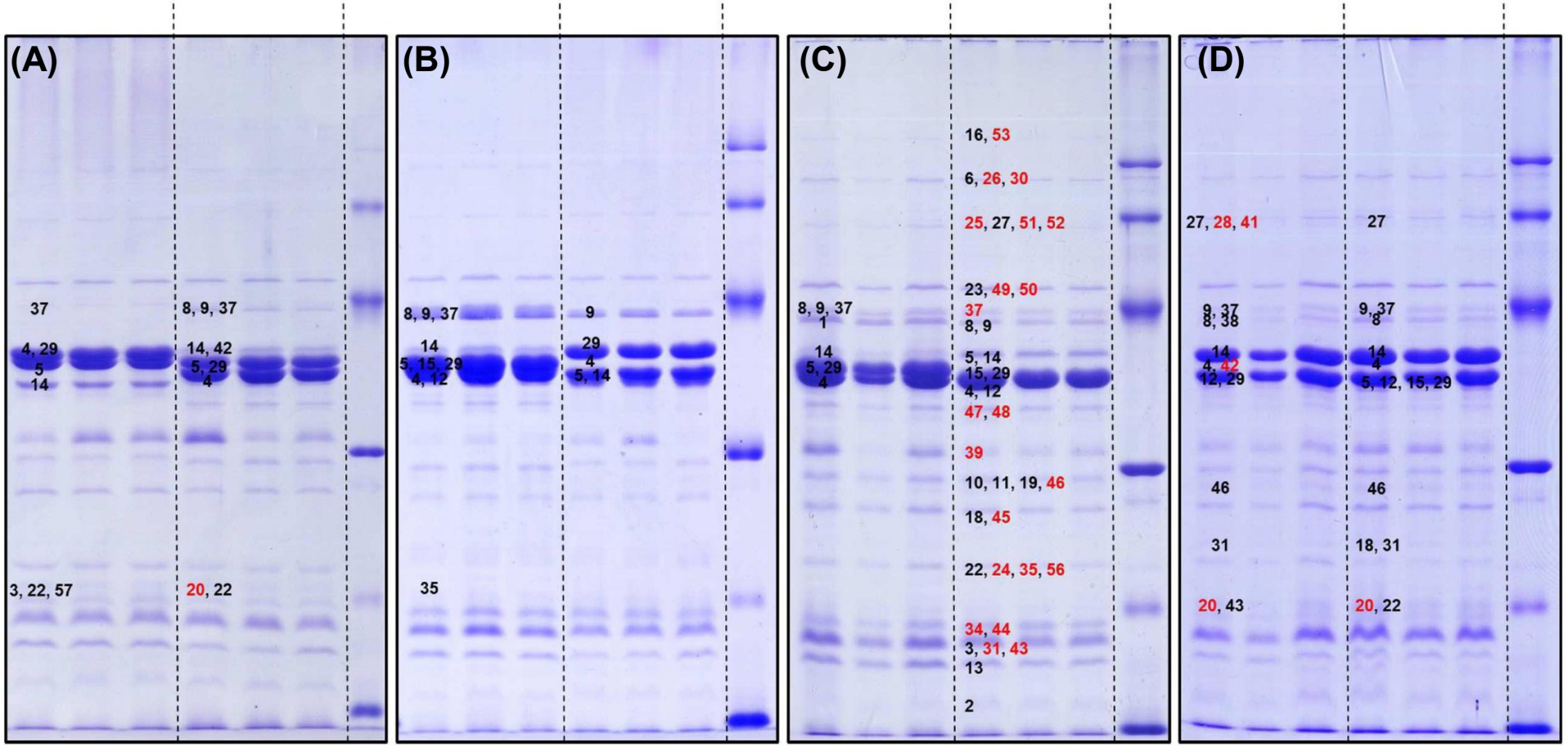


Fig. 3.8 The excised bands from OMP profiles of representative isolates of *Y. ruckeri* separated by 1D SDS-PAGE.
(A) Represents isolates RD6 (lanes 1-3) and RD28 (lanes 4-6); **(B)** represents isolates RD64 (lanes 1-3) and RD124 (lanes 4-6); **(C)** represents isolate RD354 (lanes 1-3) and reference isolate RD366 (lanes 4-6); **(D)** represents isolates RD382 (lanes 1-3) and RD420 (lanes 4-6). Lane 7 of each panel represents a MW standard (GE Healthcare, UK). Identified proteins are labelled numerically and correspond to **Table 3.6**. Proteins labelled in red were uniquely identified using gel based methods.

Table 3.8 Proteins identified through gel-based proteomics of remaining seven isolates of *Y. ruckeri*

Four proteins were identified through gel-based analysis when gel bands were excised from the remaining seven representative isolates of *Y.* ruckeri. Numbers in the 'gel position' column refer to the position of the protein in Fig. 3.8.

	Accession	Protein	Function	MW (Da)	Gel position	Prediction
1	gi|490858139	RlpA	Rod shape cell seperation	37,557	38	+
2	gi|755600220	Surface Ag	Antigen	64,555	41	-
3	gi|740407986	MalA	Macrolide transporter	40,203	42	-
4	gi|490856221	YiaD/Omp16	Multicopy suppressor of BamD	22,248	20	+
				Total	**4**	**2**

3.4.7 Comparison between OMPs of isolate RD366 identified using gel-based and gel-free proteomic approaches

Sixty six proteins were identified in isolate RD366 using a combination of gel-free and gel-based approaches (Table 3.9). Twenty eight OMPs were identified solely using gel-free approaches, 13 OMPs were identified only using gel-based approaches, while 25 OMPs were identified using both methods (Fig. 3.9).

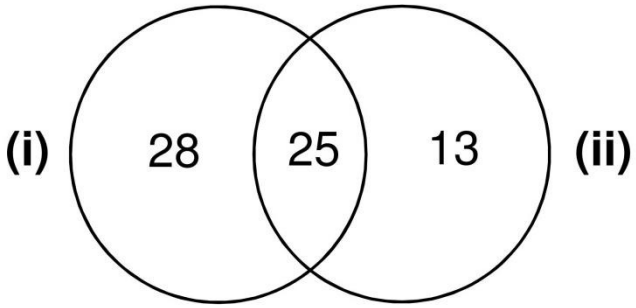


Fig. 3.9 Comparison between gel-free and gel-based proteins identified in isolate RD366
Twenty five proteins were identified in isolate RD366 by gel-free and gel-based proteomic analysis. Twenty eight proteins were identified solely by gel-based methods (i). Thirteen proteins were identified solely by gel-free methods (ii).

The 25 proteins identified by both gel-free and gel-based methods ecompassed three predicted by the transmembrane ß-barrel programs, nine predicted by the lipoprotein programs and 12 proteins predicted by both programs. However, one protein (OsmY) was not predicted by either program.

Of the 28 OMPs identified using gel-based approaches, FliF, OmpX, Phospholipase A1, TamA/YtfM, FlgH and PilV were predicted to be transmembrane ß-barrels. The proteins NlpI, NlpC/CutF, BamE, NlpD, LpoB, YajG and two hypothetical proteins (hypothetical 12 [gi|755306536] and hypothetical 20 [gi|490859521]) were predicted to be lipoproteins. Lipoprotein 3, FadL, OmpW, Lipoprotein 1, LptD, ShuA, CsgG, YdgA, a hypothetical protein (hypothetical 13 [gi|685058188]), Tsx and a membrane protein were predicted by both transmembrane ß-barrel and lipoprotein predictors. The proteins GlpC, YegD/Hsp70 and a transporter were not predicted by either program.

Table 3.9 Comparison of proteins identified in isolate RD366 using gel-based and gel-free methods.
Sixty six proteins were identified in isolate RD366 using gel-based and gel-free proteomic approaches. Fifty three proteins were identified using gel-based methods, while 38 proteins were identified through gel-free methods.

	Accession	Protein	Function	MW (Da)	Gel-based	Gel-free	Prediction	Lipo	β-barrel
1	gi|490855955	BamB	BAM complex	42,328	1	3	+	+	+
2	gi|490838180	Lpp (Brauns)	Anchors peptidoglycan to OM	6,687	2	3	+	+	-
3	gi|490860109	Pal	Peptidoglycan associated	18,103	3	3	+	+	-
4	gi|490856883	OmpA	OM integrity, porin, adherence	37,756	4	3	+	+	+
5	gi|300680152	OmpF	Porin (Small molecules)	40,049	5	3	+	+	+
6	gi|490857618	BamA/YaeT	BAM complex	88,077	6	3	+	+	+
7	gi|490857926	OsmY	Osmoregulation	11,017	7	3	-	-	-
8	gi|498532474	Flagellin	Flagellin	43,615	8	3	+	-	+
9	gi|490858688	FlgE	Links flagella to driver	44,281	9	3	+	-	+
10	gi|490855868	VacJ/MlaA	Phospholipid homeostasis	28,735	10	3	+	+	-
11	gi|490857575	MetQ	D-methionine binding	29,316	11	2	+	+	-
12	gi|490861349	YeeJ	Unknown	54,668	-	2	-	-	-
13	gi|490860274	BamC	BAM complex	38,311	12	1	+	+	+
14	gi|490855735	TcyP	Cysteine transporter	48,096	-	2	-	-	-
15	gi|490855565	Pcp/SlyB	Membrane integrity	15,334	13	2	+	+	-
16	gi|490860502	ShlB/FhaC/HecB family	Hemolysin activator	62,461	-	2	+	-	+
17	gi|490855768	OmpC.1	Porin (Small molecules)	40,484	14	1	+	+	+
18	gi|490855057	OmpE	Inorganic phosphate uptake	39,949	15	1	+	+	+
19	gi|490861198	BcsC	Cellulose biosynthesis	126,490	16	2	+	+	-
20	gi|490857772	Slp family	Structural and OM stabilisation	20,645	17	1	+	+	-
21	gi|490861063	BamD	BAM complex	27,392	18	1	+	+	-
22	gi|490857817	MipA	MltA interacting	28,040	19	2	+	+	+
23	gi|490857673	MltA	Peptidoglycan maintenance	42,866	21	2	+	+	+

Table 3.9 continued

	Accession	Protein	Function	MW (Da)	Gel-based	Gel-free	Prediction	Lipo	β-barrel
24	gi|490858151	LptE/RlpB	LPS assembly	21,089	22	1	+	+	-
25	gi|490858989	TolC	OM Channel (Efflux system)	53,080	23	1	+	+	+
26	gi|490859665	CsgG	Curli organelles/biofilm formation	24,670	24	-	+	+	+
27	gi|490860153	YqjD	Unknown	26,277	-	1	-	-	-
28	gi|490860978	TamA/YtfN	Autotransporter assembly	63,376	25	-	+	-	+
29	gi|490856221	YiaD/Omp16	Multicopy suppressor of BamD	22,248	-	2	+	+	+
30	gi|490858839	ShuA	Heme transport	73,397	26	-	+	+	+
31	gi|490861316	C-terminal protease	Unknown	75,933	-	1	-	-	-
32	gi|490857431	BtuB	Vitamin B12 transport	68,999	27	1	+	-	+
33	gi|490856454	OprC	Chitoporin	75,008	28	1	+	+	+
34	gi|490858586	OmpC.2	Porin (Small molecules)	39,971	29	1	+	+	+
35	gi|490855484	HslJ	Heat-inducible protein	15,597	-	2	-	-	-
36	gi|490859532	LptD	LPS assembly	89,597	30	-	+	+	+
37	gi|490859521	Hypothetical 20	Membrane associated	17,747	31	-	-	+	-
38	gi|490855259	Lipoprotein 1	Unknown	20,214	32	-	+	+	+
39	gi|490856025	YfhG/QseG	Quorum sensing/Virulence	34,809	-	1	-	-	-
40	gi|490858084	Hypothetical 5	Unknown	31,358	-	1	+	-	+
41	gi|490856645	YajG	Unknown	20,929	34	-	+	+	-
42	gi|490857925	OmpW	Hydrophobic compound transport	23,880	35	-	+	+	+
43	gi|490855867	FadL	Fatty acid compound trans.	46,063	37	-	+	+	+
44	gi|490856067	Phos. A1	Bacteriocin secretion	32,533	39	-	+	-	+
45	gi|238708334	OmpX	Virulence	17,086	43	-	+	-	+
46	gi|490856978	LpoB	Peptidoglycan synthesis	20,698	44	-	+	+	-
47	gi|490861050	Membrane protein	Unknown	27,916	45	-	-	+	+
48	gi|490855637	HemR	Iron transport	72,952	-	1	-	-	+
49	gi|490856641	Tsx	Nucleoside channel/Colicin rec.	29,122	46	-	+	+	+
50	gi|238707352	NlpD	Unknown	34,142	47	-	-	+	-

Table 3.9 continued

	Accession	Protein	Function	MW (Da)	Gel-based	Gel-free	Prediction	Lipo	β-barrel
51	gi|685058188	Hypothetical 13	Unknown	34,471	48	-	+	+	+
52	gi|685053742	PilV	Major Pilin	48,727	49	-	+	-	+
53	gi|490855905	YegD/Hsp70	Heat shock protein	49,742	50	-	-	-	-
54	gi|490859142	Transporter	OM Transport	60,560	51	-	-	-	-
55	gi|740408485	FliF	Flagellar machinary	61,743	52	-	+	-	+
56	gi|490855624	GlpC	Iron-Sulfur protein	115,110	53	-	-	-	-
57	gi|740407770	BamE	BAM complex	12,651	54	-	+	+	-
58	gi|740408670	Lipoprotein 3	OM lipoprotein	11,336	55	-	-	+	+
59	gi|755306536	Hypothetical 12	Unknown	11,794	56	-	+	+	-
60	gi|490861267	YraP	Phospholipid biogenesis	20,176	-	1	+	-	-
61	gi|685060129	FlgH	Flagella L-ring protein	13,149	57	-	+	-	+
62	gi|685060195	NlpC/CutF	Copper homeostasis	25,530	58	-	+	+	-
63	gi|685057898	NlpI	Cell division	33,918	59	-	+	+	-
64	gi|490858640	FliD	Flagellar filament cap	49,084	-	1	+	-	-
65	gi|755308840	YdgA	GTP-binding protein	54,395	60	-	+	+	+
66	gi|46369605	RupA[2]	Unknown	17,892	-	1	-	-	-
				Total	**53**	**38**	**56**	**41**	**36**

Of the 13 OMPs identified solely using gel-free approaches, ShlB/FhaC/HecB family protein, hypothetical protein 5 and HemR were predicted by transmembrane ß-barrel programs. YiaD/Omp16 was predicted by both transmembrane ß-barrel and lipoprotein programs, while YeeJ, TcyP, HslJ, YqjD, a C-terminal protease, YfhG/Qseg, YraP, FliD and RupA were not predicted by either.

3.4.8 Comparison between proteomic identification and bioinformatic prediction

Combining the unique proteins identified through gel-free approaches in all isolates, gel-based analysis of the mini and large-gel profiles of isolate RD366 and the gel based analysis of the remaining isolates, 86 unique proteins were identified. However, only 60 of these had been predicted (69.8%). Therefore, a further 26 OMPs were identified that had not been predicted (30.2%).

Previously, 141 proteins had been confidently predicted from four genomes of Y. ruckeri to be OM located (Table 3.2). As we had only identified 86 proteins of which 60 had been identified, a further 81 proteins that had been confidently predicted remained unidentified. Of these proteins, 10 were predicted to be transmembrane ß-barrel proteins, 27 were predicted to be OM lipoproteins and 44 as both. These proteins belonged to the functional categories OM biogenesis and integrity (two proteins), transport and receptor (14 proteins), adherence (13 proteins), enzymatic activity (nine proteins) and motility (five proteins). Thirty eight confidently predicted OMPs with other (15 proteins), hypothetical (11 proteins) or unknown (12 proteins) functions were also not identified.

3.5 Discussion

Conventional proteomics looking to characterise the bacterial OM typically involves an initial sub fractionation of cell lysates using a strong detergent followed by gel electrophoresis, LC and MS (Alteri & Mobley, 2007; Taniguchi *et al.*, 2010; Walters & Mobley, 2009; Yoon *et al.*, 2003). Studies that have investigated the OM proteome have been limited to the characterisation of single or few strains (detailed in section 1.5.3 of the Introduction), limiting the understanding of the complexity and heterogeneity of the mechanisms that govern interaction with the host (Wurpel *et al.*, 2015). In this study these limitations were addressed by utilising a combination of bioinformatic and proteomic approaches to analyse the OM proteome of eight isolates of *Y. ruckeri* recovered from two host species, Atlantic salmon and rainbow trout. The reliability of gel-free and gel-based proteomic techniques has been evaluated in combination with thorough bioinformatic prediction analyses. The identification and characterisation of proteins associated with the OM of a bacterium improves our understanding of important host-pathogen interactions, which in turn may lead to the development of new diagnostic, drug and vaccine targets (Wurpel *et al.*, 2015).

A bioinformatic workflow developed and utilised for *P. multocida* (E-komon *et al.*, 2012) was used to identify OM-localised proteins in the genomes of four *Y. ruckeri* isolates (three from rainbow trout [ATCC29473, CSF007-82 and YRB], and one from Atlantic salmon [37551]). Using this approach, 141 OMPs were confidently predicted across the four genomes, representing 2.87%, 2.77%, 2.86% and 2.99% of the genomes ATCC29473, CSF007-82, YRB and 37551, respectively. These figures are lower than those obtained for the *P. multocida* avian genome (4.8% of the proteome) (E-komon *et al.*, 2012). A similar approach was taken in the analysis of the intracellular pathogen *Ehrlichia ruminantium*, where the OM proteome comprised 5.4% of the genome (Moumène *et al.*, 2015), and several members of the *Chlamydiae* family - *Parachlamydia acanthamoebae* (2.5% of the genome), *Simkania negevensis* (3.8% of the genome) and *Waddlia chonrophila* (2.8% of the genome) (Aistleitner *et al.*, 2015). The genomes of *Y. ruckeri* examined encode substantially more proteins than *P. multocida*, *E. ruminantium*

or any of the *Chlamydiae* genomes examined. The four genomes of *Y. ruckeri* encode between 3079 and 3352 proteins, whereas the genomes of *P. multocida*, *E. ruminantium, P. acanthamoebae*, *S. negevensis* and *W. chonrophila* that were examined encode 2009, 950, 2531, 2218 and 1828 proteins respectively.

Putative functions were assigned to 121 (85.8%) of the 141 confidently predicted OMPs. Further information is required to characterise the remaining 20 proteins which were either hypothetical or of unclear function. A predicted core OM proteome of 48 proteins (Table 3.2; Fig. 3.1) identified in all four genomes is low considering the known homogeneity of this bacterium however. However, as the prediction was carried out on strains representing different phenotypes, variation in surface-expressed proteins may be expected. Examining proteins identified in at least three out of four genomes increases the number of predicted core OMPs to 78, which equates to over 55% of the total predicted proteins. Of these 78 proteins, only 50 are common to all three rainbow trout genomes examined (Fig. 3.1).

Forty nine proteins were identified by the consensus prediction of transmembrane ß-barrel, subcellular localisation and lipoprotein prediction programs. Significantly, 12 proteins were identified solely by transmembrane ß-barrel predictors and 40 proteins were identified by lipoprotein predictors. This result supports the need to utilise a variety of prediction tools to accurately predict the OM proteome, as suggested by E-Komon *et al.*, (2012) and supported by Aistleitner *et al.*, (2015). However, as only one of the genomes examined was completely closed and fully annotated, this may account for the variations in OMP numbers. Certain proteins may not have been covered in the genomes sequenced to contig or scaffold level.

An important aspect of this study was to try and uncover proteins that may be specific to either rainbow trout or Atlantic salmon that may have important roles in pathogenesis. Through the prediction analysis, 39 proteins were specific to the rainbow trout genomes, and 4 proteins were unique to the Atlantic salmon genome. However, as the prediction was carried out on three rainbow trout

genomes, and only a single Atlantic salmon genome, it is not plausible to draw conclusions from the prediction at this stage.

Upon assigning functional categorisation to the predicted OMPs, 10.6% were suggested to have roles in OM biogenesis and maintenance of cell integrity. This included the proteins LptD and LptE which have roles in LPS assembly (Fig. 1.18), the Brauns lipoprotein (Lpp) which is involved in anchoring the peptidoglycan to the OM, and members of the BAM complex (Fig. 1.16). Proteins of the BAM complex are involved in the assembly of proteins with diverse cellular functions, including solute transport, protein secretion and the assembly of protein and lipid components of the OM. They are essential for bacterial viability and function, with crucial roles in assembly of the majority of bacterial OMPs (Hagan *et al.*, 2011; Ricci & Silhavy, 2012). The insertion of proteins into the OM is dependent on the protein complex containing BamA and four associated lipoproteins (BamB, C, D and E) (discussed extensively in section 1.4.3.1 of the Introduction).

Over a fifth (20.6%) of the predicted OMPs were suggested to have roles as transporters and receptors, which include the porin proteins. Porins participate in adhesion to and invasion of host cells and evasion of host defence mechanisms. However, the main function of porin proteins is in transport and receptor activities (section 1.4.3.4.2)(Galdiero *et al.*, 2012). Within this study, several OMPs were confidently predicted with known roles as OM porins in *Y. ruckeri,* of which many were further identified through gel-based and gel-free approaches; these include OmpC, OmpF, OmpE, and OmpW.

A smaller number of predicted OMPs were involved in adherence (11.3%). Adhesins facilitate the fundamental initial step in infection, through recognition and binding to specific host tissues. Fourteen OMPs were identified through the bioinformatic prediction that had roles in pilus and fimbriae formation, with three of these proteins identified in all four genomes (FimD, SafC and PilF). These adhesins are known to play important roles in the infection of other pathogenic bacteria including *E. tarda, A. hydrophila* and uropathogenic *E. coli* (UPEC) (Niemann *et al.*, 2004; Sakai *et al.*, 2003; Sato *et al.*, 1989). The large

number of proteins predicted with roles in fimbrial production suggests that fimbriae are important factors in the initial adherence of *Y. ruckeri* to the host, although infection studies would be necessary for confirmation. The major subunit of fimbriae (FimA) was identified in the genome of rainbow trout isolate ATCC29473, although only fimbrial ushers were identified in the remaining genomes.

Proteins involved in motility (7.1% of the total) were relatively abundant across all four genomes examined. Isolate RD6 is of the biotype 2 (non-motile, lipase negative) phenotype, and so the absence of flagella or flagella apparatus proteins detected in both gel-based and gel-free proteomics was expected. The genomes of non-motile biotype 2 isolates are not currently available on NCBI. Therefore, we were unable to complete a prediction of the OM proteome of this phenotype at present. However, Welch *et al.*, (2011) demonstrated that *fliP, fliQ, fliR, flhB, flhA* and *flhE* encoding flagellar proteins were present in biotype 2 isolates, but are mutated (section 1.2.2.2). However, these genes are still likely to be predicted within the genomes.

Utilising nLC-ESI-MS/MS and gel-free proteomic analyses, a total of 65 OMPs were identified across the eight representative isolates of *Y. ruckeri*. Fifty two of these proteins were confidently predicted by the bioinformatic prediction approach. The remaining 13 proteins (YeeJ, TcyP, YqjD, carboxy terminal domain protein, HslJ, YfhG, partial hemolysin, RupA, DNA circulation protein, TspB family protein and three hypothetical proteins [hypothetical 15, 16 and 17]) were indicated to be OMPs only after subsequent BLAST, domain and literature searches. These proteins were most likely overlooked in the bioinformatic prediction search either through poor annotation within the genomes, being filtered out by the selection criteria, or absence from the genomes on which the prediction was carried out.

The OMPs YeeJ, TcyP, HslJ and YfhG were identified in all eight isolates through gel-free proteomic analysis. The functions of these proteins are diverse. YeeJ is suggested to have roles in adhesin and biofilm formation (Beloin *et al.*, 2008), TcyP is involved in the uptake of L-cysteine (Burguière *et al.*, 2004), HslJ is a

heat-inducible protein with involvement in novobiocin resistance (Jovanovic *et al.*, 2003), and YfhG/QseG is involved in the regulation of virulence and metabolism in EHEC and required for pedestal formation in host epithelial cells during infection (Reading *et al.*, 2009).

Proteins BamA, B, C, D and E were identified in all four genomes through the prediction analysis, while only BamA, B and C were identified through gel-free proteomics in all isolates. BamD was absent in isolates RD382 and RD420, while BamE was not identified in any isolate. The crucial role of BamA in OM β-barrel assembly has been well documented, while BamD is the only essential lipoprotein in the BAM complex, and is highly conserved (Malinverni *et al.*, 2006; Moumène *et al.*, 2015; Sklar *et al.*, 2007). It is therefore interesting that BamD should be absent from the Atlantic salmon strains RD382 and RD420. However, all proteomics experiments have sensitivity thresholds below which proteins may not be identified - even though it may be present at a functional level of perhaps a few molecules per cell. Proteins that are small, relatively hydrophobic or heavily modified will be more difficult to detect (Chandramouli & Qian, 2009). A possible explanation as to why BamD was absent from isolates RD382 and RD420, and BamE was not found in any isolates is that they may be only loosely associated with the OM, and the extraction process may have resulted in the loss of these proteins. Alternatively, BamD and BamE may either be expressed at very low abundance in these isolates or may not be expressed under this growth condition. The sensitivity of the mass spectrometer may have meant that these proteins were overlooked if their abundance was too low.

The role of BamA in autotransporter biogenesis has been well documented. However, a study by Rossiter *et al.*, (2011) identified the importance of BamD in this process also. It is possible therefore that the Atlantic salmon strains RD382 and RD420 are unable to produce an autotransporter protein that is present in the remaining isolates, although no autotransporters were detected under this growth condition in any isolate. It is plausible that the examined isolate did not express BamD to detectable levels - however, it is more likely that with the essential role of BamD in the BAM complex, BamD was lost during the sample preparation process. Examining the expression of proteins under different

conditions of growth may reveal an optimum condition for the expression of this protein.

Proteins involved in iron transport including HemR (all isolates [one replicate]; absent from RD420), FhuA (RD6 [one replicate]) and ShuA (RD6 [two replicates]; RD124 [one replicate]; RD382 [one replicate]) were identified through gel-free analysis. ShuA was also identified in isolate RD366 through gel-based analysis. Proteins involved in the uptake of iron are normally upregulated under conditions of iron-delpetion. As iron is essential for the bacteria (1.4.3.4.2), under conditions of iron-depletion these proteins are upregulated in order to enhance iron scavenging and uptake efficiency. Under this growth condition iron was readily available in the media; therefore conditions were not optimal for expression of proteins involved in iron transport. Thus, the identification of these proteins was not entirely expected. This makes it unfeasible to draw conclusions as to the presence or absence of these proteins in particular isolates at this time. However, we are able to conclude that these proteins may be under the control of different regulatory mechanisms in different isolates.

Both OmpA and OmpF were identified in all isolates suggesting that they are ubiquitously expressed under this growth condition. However, the proteins OmpC.1 and OmpC.2 were not expressed simultaneously, or in all isolates. OmpC.1 was identified only in isolates RD6 (three replicates), RD28 (two replicates) and RD366 (one replicate). OmpC.2 was identified in all isolates except RD354. However, it was only identified in one replicate of each except RD382, in which it was identified twice. OmpC is not routinely expressed under aerobic growth conditions in other bacteria, including *E. coli* (Batchelor & Walthers, 2005), so it is perhaps not surprising that we should not identify it. However, it is revealing that both OmpC.1 and OmpC.2 are not expressed to the same levels in all isolates. This suggests that these proteins are under the control of separate promotors, and perhaps have different functions.

The gel-free proteomic analysis did reveal proteins that were specific to each host. Seventeen proteins were found only in the rainbow trout isolates, and five proteins were unique to the Atlantic salmon isolates. However, these proteins

were found mostly in single replicates and did not appear to indicate any sort of pattern. Thus, we are unable to suggest host specific roles of these proteins at this stage. Variation may exist in the sequence of proteins found in isolates recovered from rainbow trout and Atlantic salmon, which future analysis may reveal. However, it is more plausible that the conditions under which the bacteria are grown will affect the regulation of different proteins.

Through a thorough gel-based proteomic analysis of the reference isolate RD366, 47 unique proteins were identified when the entire gel lane (encompassing both bands and interband regions) was excised into 17 fractions (Fig. 3.3). As 29 of these proteins were identified in three replicates and 39 proteins identified in at least two replicates, the reproducibility of the method is highlighted (Fig. 3.4). Excising only the visible bands of the isolate RD366 (in triplicate), resulted in the identification of 41 OMPs (Fig. 3.5). Twenty nine were common to all replicates and 34 were common to at least two replicates. Combining these two approaches resulted in 53 unique OMPs being identified, of which 47 had been predicted using the aforementioned bioinformatic approaches.

The major OMPs OmpA, OmpC and OmpF were identified in all isolates through gel-based analysis. A typing scheme based on variations in the molecular mass of these OMPs was developed in order to characterise *Y. ruckeri* isolates (Davies, 1991c). OmpA, OmpC and OmpF were excised from all eight reference isolates in order to confirm their identity. Referring to the original typing scheme by Davies (1991c), it is clear that these major proteins are identifiable in isolates of different OMP-types in the correct orientation as previously described (1.2.2.3.3). This study serves as a confirmation for the previous work of Davies (1991c), and enforces the reliability of the results within this study.

Several proteins with roles in adherence (and ultimately virulence) were identified through gel-based analysis of isolate RD366. These included PilV, OmpX, CsgG, YhlA and PqiB/MAM7. Of these, only CsgG (RD6 [two replicates]; RD64 [two replicates]; RD354 [one replicate]) and PqiB/MAM7 (RD6 [one replicate]; RD64 [one replicate]; RD354 [one replicate]) were identified through gel-free analysis.

The major pilin subunit PilV was identified through gel-based proteomics in isolate RD366, although as this was the only isolate where the entire gel lane was examined it is impossible to conclude its omission from the remaining strains as confirmation of presence, absence or host specificity. Adhesion is required for biofilm formation, and the overexpression of flagellar proteins is one of the phenotypic characteristics of bacteria that display high adhesion efficiency. In other bacteria, genetic screening analyses of mutants defective in biofilm formation has shown that the initial reaction with a surface is promoted by pili and flagella, which subsequently allow the bacteria to move and encounter other bacteria to form microcolonies (O'Toole & Kolter, 1998; Pratt & Kolter, 1998; Watnick & Kolter, 2012).

OmpX has been shown to be structurally similar to OmpA in terms of ß-sheet topology, and belongs to a family of highly conserved proteins that appear important for virulence by neutralising host defence mechanisms. The prototype OmpX protein, Ail from *Y. enterocolitica*, promotes adhesion to and entry into eukaryotic tissue culture cells (Heffernan *et al.*, 1994; Koebnik *et al.*, 2000; Vogt & Schulz, 1999). OmpX from *E. coli* belongs to a family of highly conserved bacterial proteins promoting bacterial adhesion. While Ail has been described in all three human pathogenic *Yersinia*, it was not identified in *Y. ruckeri* during previous studies (Kawula *et al.*, 1996).

CsgG is an OM lipoprotein (Loferer *et al.*, 1997) required for the secretion and stabilisation of two other proteins, CsgA and CsgB, which form curli amyloid fibres on the extracellular surface. These fibres have been implicated in a number of processes including biofilm formation, attachment to and invasion of host cells, interaction with host proteins and activation of the immune system (Barnhart & Chapman, 2006).

The hemolysin YhlA is a known virulence factor of *Y. ruckeri*, with roles in cytotoxicity and contributing to the acquisition of iron from host cells (discussed in section 1.2.3.3) (Fernández *et al.*, 2007a).

The adhesion factor PqiB/MAM7 primes bacteria for immediate attachment when a host cell is encountered. PqiB/MAM7 is capable of binding both fibronectin and phosphatidic acid on the host cell surface and both interactions are required for bacterial adhesion to host cells (Krachler & Orth, 2011; Krachler *et al.*, 2011; Lim *et al.*, 2014).

A protein with enzymatic functions, phospholipase A1, was also identified through gel-based analysis. This protein has been shown to be involved in colicin release from *E. coli* and has been implicated in the virulence of *Campylobacter* and *Helicobacter* strains by destabilising the membrane bilayer and allowing the release of colicins or virulence factors (Koebnik *et al.*, 2000).

In total, four proteins were identified in the remaining seven isolates in addition to isolate RD366 (Fig. 3.8 and Table 3.8). These proteins were YiaD/Omp16 (Isolates RD28; Fig. 3.8A, lanes 4, 5 and 6, RD382 and RD420; Fig. 3.8D, lanes 1 to 6), RlpA (isolate RD420; Fig. 3.8D, lanes 4, 5 and 6), Surface Ag (isolate RD420; Fig. 3.8D, lanes 1, 2 and 3) and MalA (isolate RD28; Fig. 3.8A, lanes 4, 5 and 6, RD420; Fig. 3.8D, lanes 1, 2 and 3). Of these, YiaD/Omp16 (all isolates) and RlpA (RD420; one replicate) were also identified using gel-free methods.

As only isolate RD366 was subjected to thorough gel-based analyses conclusions between the use of gel-based and gel-free approaches can only be drawn using this isolate. In total, 66 proteins were identified in isolate RD366 using a combination of gel-free and gel-based approaches (Table 3.9). Of note, 28 OMPs were identified solely using gel-free approaches, 13 with gel-based approaches, and 25 identified using both methods (Fig. 3.9). This highlights the importance of combining proteomic techniques to ensure the greatest numbers of proteins are identified.

Of the 28 OMPs identified solely by gel-based proteomics, six were predicted to be transmembrane ß-barrels and eight as lipoproteins. Eleven proteins were predicted by both programs. Of the 13 OMPs identified solely using gel-free approaches, three were predicted by transmembrane ß-barrel programs while

none were predicted as lipoproteins. However, YiaD/Omp16 was predicted by both transmembrane ß-barrel and lipoprotein programs.

The identification of antibodies produced by the host in response to specific proteins utilised by the bacteria could provide valuable information for the development of future vaccines. Utilising western blotting to analyse the antibodies present in fish serum could allow for the identification of immunogenic OMPs, as has been shown for other fish pathogenic bacteria including *P. salmonis* (Marshall *et al.*, 2007), *A. hydrophila* (Khushiramani *et al.*, 2012) and *Aeromonas veronii*(Vázquez-Juárez *et al.*, 2003).

In conclusion, this study has highlighted the importance of utilising complementary proteomic techniques in order to fully appreciate the OM proteome of a bacterial species. Further studies of these proteins may shed light on how these bacteria adapt to different host species and niches and cause infection in Atlantic salmon and rainbow trout. The results in this study have not enabled a simple distinction in OM proteome composition between isolates recovered from Atlantic salmon and rainbow trout. However, further examination of isolates grown under conditions mimicking *in vivo* and environmental growth would provide information into the expression of specific OMPs during relevant situations. Certain conditions of growth are known to be influential in the regulation of different OMPs, including degree of aeration (Batchelor & Walthers, 2005), temperature (Fernández *et al.*, 2007a), iron availability (Davies, 1991b; Kanaujia *et al.*, 2015; LaFrentz *et al.*, 2009) and salt concentration (Kao *et al.*, 2009). In addition to the method used to obtain the OM fractions, consideration must also be given to the growth conditions of the bacterium to allow identification of additional OMPs.

Chapter 4 Proteomic analysis of the *in vivo* and environmental growth conditions of *Y. ruckeri*

4.1 Introduction

Microbes inhabit a diverse range of environments. The conditions that define an environmental niche provide the context that stipulates bacterial gene expression. A single bacterial species may be adapted to survive in multiple environments, including ranging temperatures, oxygen availability, temperature and salinity. The protein expression of a pathogen is defined by that given set of conditions and will differ for every kind of environment that the organism encounters.

The life cycle of the Atlantic salmon and rainbow trout (Fig. 1.5) means that bacteria must be particularly resilient in order firstly to survive in a freshwater environment, before gaining entry to the host through the gills, residing in the gut as an enteric pathogen and finally causing septicaemia residing in the blood (Ohtani *et al.*, 2014; Ross *et al.*, 1966).

The bacteria also have to adapt to changing temperatures as the seasons both heat and cool the water. Transmission of ERM is known to increase as the water temperature begins to rise in the early spring and summer and fish become more stressed (Busch, 1978; Rodgers, 1991). *Yersinia ruckeri* must also survive the transition between fresh, brackish and seawater environments in accordance with the different stages of the Atlantic salmon and rainbow trout life cycle (Fig. 1.5).

Inside the host itself, there are a variety of different conditions which the bacteria must overcome to survive. *Yersinia ruckeri* must be able to survive in an anaerobic environment as little oxygen is available to the bacterium inside the host. As discussed previously (1.4.3.4.2), iron is an essential nutrient for the bacteria to survive and cause infection. However, it is scarcely available within the host where it is bound to compounds including haemoglobin and transferrin. Bacteria have developed complex systems in order to scavenge iron from the host and environment (Fig. 1.20). Understanding how the bacteria utilise

different OMPs in different environments would further our understanding of the organism.

The rate at which a bacterium grows is influenced by its surroundings. The available nutrients play a role, but similarly, the amount of energy the bacteria have to expend in order to overcome adverse conditions will limit how quickly it can replicate and to what degree. Bacteria must alter the OM in response to the conditions encountered to adapt to their surroundings. Several abiotic aspects can influence the rate at which the bacterium is able to grow. These include the ambient temperature, oxygen availability, pH and osmotic pressure. Similarly, nutrient availability will not only dictate the ability of the bacteria to grow, but also affect its pathogenicity (Shehata & Marr, 1971). In this study, the rate of growth of isolates was affected by different conditions of aeration, temperature, iron-depletion and salinity.

Increased infection is often attributed to poor husbandry conditions which increase the levels of stress within fish. These include overcrowding, altering the temperature of the water, the levels of dissolved oxygen, increased handling or other environmental factors which can affect both the bacteria and the host (Rodgers, 1991; Stevenson, 2007).

The aim of this part of the study was to examine the effect of different *in vivo* and environmental factors on the rate of growth of representative isolates of *Y. ruckeri*, before observing changes in OM proteome composition. Comparing to standard growth conditions under aerobic aeration, the OMP profiles and OM proteome were examined under anaerobic conditions, growth in an iron depleted environment, in seawater medium, in media supplemented with NaCl and at temperatures ranging from 8°C to 45°C.

Two reference isolates, RD124 and RD158, were selected to determine the optimal conditions mimicking *in vivo* and environmental growth (as above) in which the rate of growth of the bacterium was not substantially inhibited, while still allowing changes in the OMP profiles of the isolates to be observed. Once these conditions had been determined, the rate of growth and OMP-profiles of a

further seven isolates (used previously in Chapter 3), representative of the infectious clones of Atlantic salmon and rainbow trout were examined. The OM proteome was then subjected to gel-free proteomic analysis to identify differences in protein expression between these isolates under various growth conditions.

4.2 Materials and methods

4.2.1 Anaerobic growth conditions

Bacteria were grown anaerobically in 1000 ml of TSB in static 1-litre Erlenmeyer flasks overlaid with sterile mineral oil, at 22°C.

4.2.2 Temperature

Bacteria were grown at temperatures of 8, 15, 22, 28, 37 and 45°C in 400 ml of TSB in 2-litre Erlenmeyer flasks with shaking at 120 rpm. Temperature controlled shaking incubators were used, and the temperature was fine-tuned and confirmed with the aid of a thermometer.

4.2.3 Iron Depletion

Bacteria were grown under iron restricted conditions created through the addition of the iron chelator 2,2′-dipyridyl (Sigma, UK) to 400 ml TSB in a 2-litre Erlenmeyer flask. Optimal concentrations of 2,2′-dipyridyl were determined through growth curve and OMP analysis, with concentrations ranging from 170 to 230 μM.

4.2.4 Sodium chloride (NaCl)

Bacteria were grown in 400 ml of TSB in a 2-litre Erlenmeyer flask supplemented with NaCl (Sigma, UK). Isolates were grown at 22°C with shaking at 120 rpm. Optimal concentrations of NaCl were determined through growth curve and OMP analysis, with concentrations ranging from 0 to 9% examined.

4.2.5 Seawater medium

Bacteria were grown in 400 ml of TSB in a 2-litre Erlenmeyer flask supplemented with Peacock Salt SEAMIX (Peacock, SMX0015) to mimic seawater. SEAMIX is a specially produced formula, that when mixed with freshwater makes an artificial seawater that is commonly use to examine the suitability of fish (smolts) for transfer to seawater. The mix contains a complex blend of salts including NaCl,

magnesium sulphate ($MgSO_4$), magnesium chloride ($MgCl_2$), calcium chloride ($CaCl_2$) and potassium chloride (KCl), which are the major components of seawater. Isolates were grown at 22°C with shaking at 120 rpm. Optimal concentrations of artificial seawater were determined through growth curve and OMP analysis, with concentrations of 1.5, 2.5 and 3.5% used.

4.3 Results

4.3.1 Optimisation of growth conditions using reference isolates

In order to determine the optimum conditions for mimicking both *in vivo* and environmental growth, isolates RD124 and RD158 were examined initially. Isolate RD124 is a biotype 1, serotype O1, OMP-type 3a isolate recovered from rainbow trout. This isolate is representative of the 'Hagerman' type strain. Isolate RD158 is a serotype O1, biotype 1, OMP-type 2c isolate recovered from Atlantic salmon. These isolates were selected due to the substantial information available as to their phenotypic and pathogenic properties (Davies, 1991a, d).

The isolates RD124 and RD158 were both grown under conditions including different degrees of aeration, temperatures ranging from 8°C to 45°C, salinities ranging from 0% (0 ppt) to 9% (90 ppt) and in concentrations of the iron chelator 2,2'-dipyridyl ranging from 0 µM to 230 µM.

4.3.2 Effect of growth phase on the OM proteome profile

In order to determine differences in OMP profiles at different phases of growth, isolates RD124 and RD158 were grown under standard conditions of aeration (2.2.3.4) and harvested after 6, 12, 18 and 24 h. These time points were representative of lag phase, early log phase, late log phase and stationary phase, respectively (Fig. 4.1).

Upon observing the OMP profile in both isolates (Fig. 4.2), there was an up regulation of a ~39.5 kDa protein as growth increased from lag to stationary phase (Fig. 4.2, lanes 4 and 8, arrows[1]). Based on previous work and its known MW, this protein is presumed to be OmpC. Similarly, there was down regulation of a ~39 kDa and ~40 kDa protein most obvious at stationary phase in isolates RD124 and RD158 respectively (Fig. 4.2, lanes 4 and 8, arrows[2]). Two proteins of ~28 kDa and ~43 kDa were most abundant during lag phase in both isolates (Fig. 4.2, Lanes 1 and 5, arrows[3]). A protein of ~42 kDa was present in isolate RD124 at 12 h growth only (Fig. 4.2, Lane 2, arrow[4]).

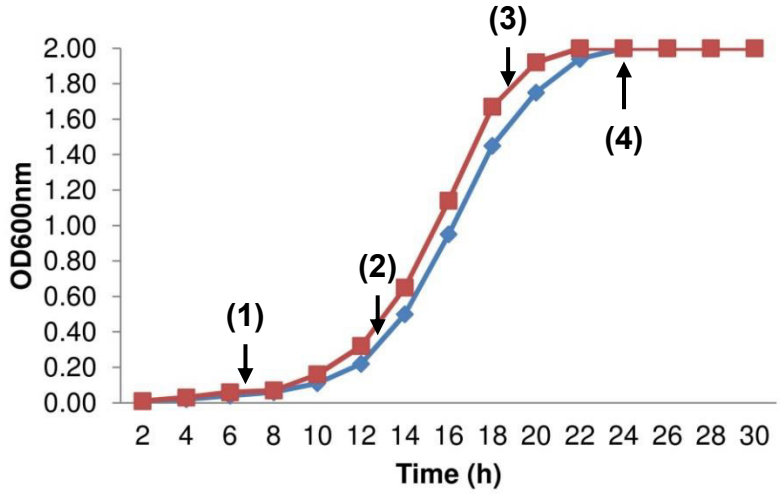


Fig. 4.1 Growth of isolates RD124 and RD158.
Isolates RD124 (red) and RD158 (blue) pass through lag phase (2-6 h), logarithmic growth phase (10-20 h) and stationary phase (22 h onwards). Isolates were harvested at points (**1**; 6 h), (**2**; 12 h), (**3**; 18 h) and (**4**; 24 h) in order to determine the effect that the period of growth had on the OMP profile.

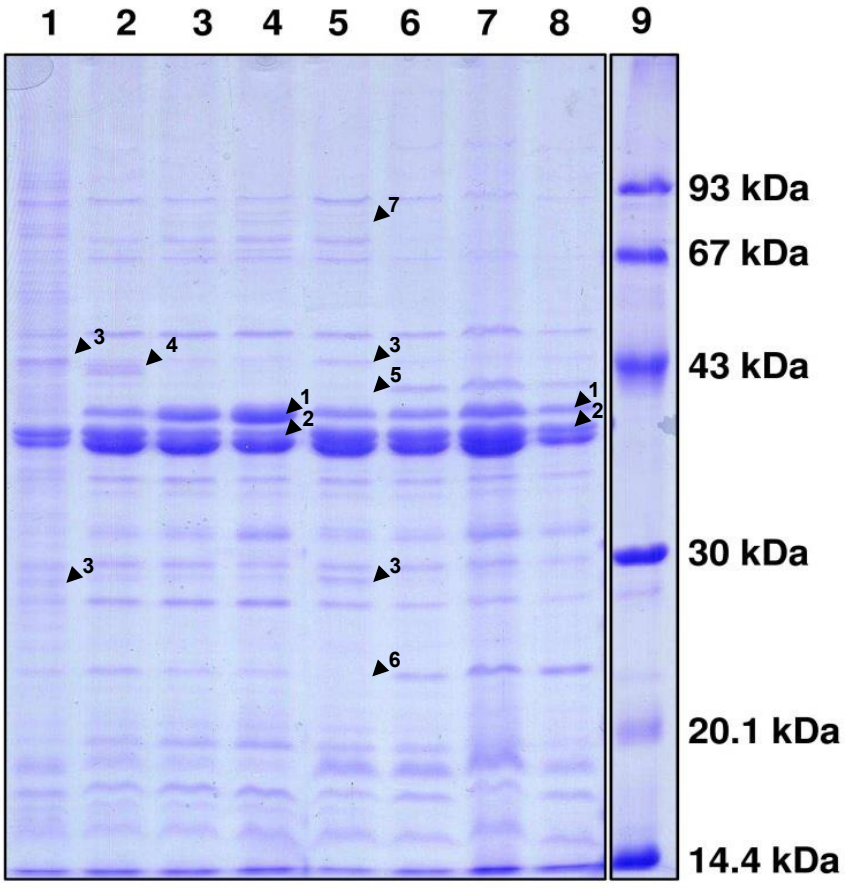


Fig. 4.2 Variation in the OMP profiles of *Y. ruckeri* at different stages of growth.
Isolates RD124 (lanes 1-4) and RD158 (lanes 5-8) were harvested after 6 h (lag phase; Lane 1 and 5), 12 h (early log phase; Lane 2 and 6), 18 h (late log phase; Lane 3 and 7) and 24 h (stationary phase; Lane 4 and 8). Lane 9 is a MW standard (GE healthcare, UK). Arrows represent differences in protein expression at different stages of growth, discussed further in the text.

A ~41 kDa protein was present in isolate RD158 at all stages of growth except lag phase (Fig. 4.2, Lane 5, arrow[5]). A protein of ~23 kDa was absent from both isolates during lag phase. This is seen most clearly in isolate RD158 (Fig. 4.2, Lane 5, arrows[6]). A protein of ~72 kDa was present under all conditions in isolate RD124. However, it was only found during lag phase in isolate RD158 (Fig. 4.2, lane 5, arrow[7]).

It appears that there is not an enormous degree of variation between OMP profiles and that growth phase has only a minimal effect on the proteins expressed in the OM of the bacteria. In order to standardise the analysis of the remaining OMP profiles, isolates were grown to mid/late log phase before the isolates were harvested and the OM proteome extracted. This will remove variation in the OMP profile based on the point of growth, ensuring that any variation observed was as a result of the condition examined.

4.3.3 **Effect of aeration on the growth rate and OMP profiles**

Isolates were grown under varying conditions of aeration, including high aeration (created using baffled Erlenmeyer flasks, with shaking at 180 rpm), moderate aeration (normal flask, 120 rpm), low aeration (normal flask, static) and anaerobic growth (100 ml TSB, 100 ml Duran overlaid with sterile mineral oil, grown statically).

Similar patterns of growth were seen for both isolate RD124 (Fig. 4.3A) and RD158 (Fig. 4.3B), with isolates grown under conditions of high and moderate aeration taking longer to enter logarithmic growth (~10 h), before growing rapidly and reaching stationary phase after 16 h. There was not a great deal of difference between isolates grown under low aeration or anaerobic conditions; both isolates entered logarithmic growth after 6 h, before entering stationary phase at 16 h. Growth under lower and anaerobic conditions was to a much lower OD_{600} than when grown under moderate and high aeration. Mid log-phase cultures were at OD_{600nm} 1.0 and 0.6 for higher and lower aerations, respectively.

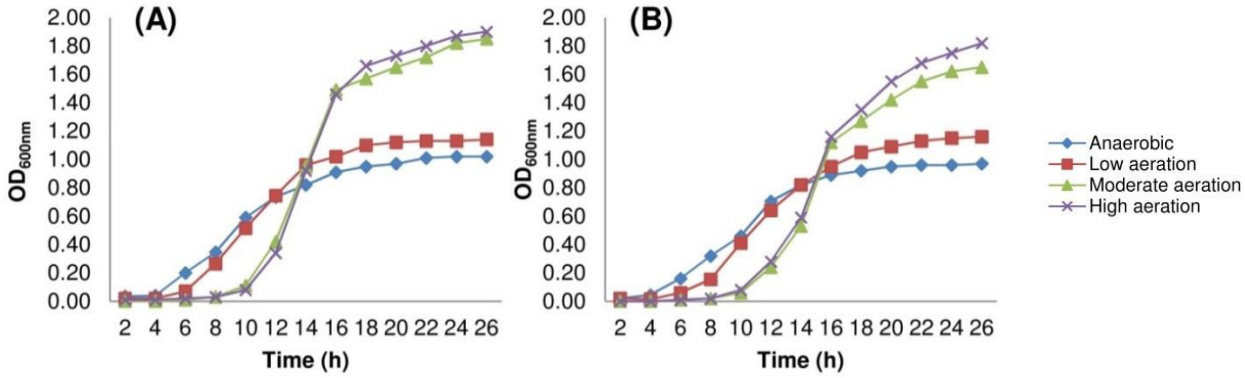


Fig. 4.3 Effect of aeration on the growth of representative isolates of *Y. ruckeri*.
Isolates RD124 **(A)** and RD158 **(B)** were grown under anaerobic, low, moderate and high aeration, and OD_{600nm} measured at two hourly intervals.

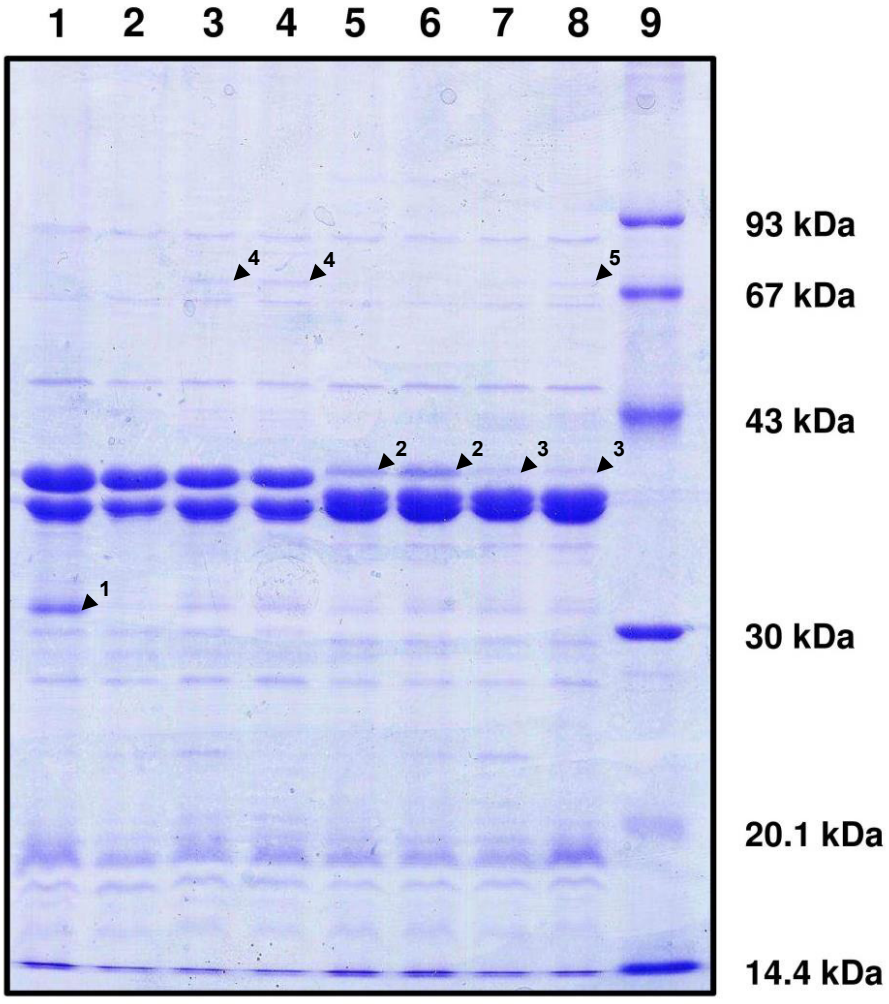


Fig. 4.4 The effect of aeration on the OMP profiles of *Y. ruckeri*.
Isolates RD124 (lanes 1-4) and RD158 (lanes 5-8) were grown under anaerobic (lanes 1 and 5), low (lanes 2 and 6), moderate (lanes 3 and 7) and high (lanes 4 and 8) aeration. Lane 9 is a MW standard (GE Healthcare, UK).

Isolates grown under high and moderate aeration also had similar profiles to each other (Fig. 4.4, lanes 3 and 4, 7 and 8). Under anaerobic growth, isolate

RD124 had one additional band at ~31 kDa, which was not present under the subsequent aerations (Fig. 4.4, lane 1, arrow[1]). Anaerobic conditions and low aeration produced a more abundant band of ~39.5 kDa, which was more obvious in isolate RD158 (Fig. 4.4, lanes 5 and 6, arrows[2]). This band was drastically down-regulated under moderate and high aeration (Fig. 4.4, lanes 7 and 8, arrows[3]). In isolate RD124, an additional band of ~69 kDa was apparent in under moderate and high aeration (Fig. 4.4, lanes 3 and 4, arrows[4]). This band was also apparent in isolate RD158 under high aeration (Fig. 4.4, lane 8, arrow[5]).

There is little difference in the profiles of anaerobically grown isolates and isolates grown at low aeration. Similarly, isolates grown at moderate and higher degrees of aeration are similar to one another. However, as the profiles of isolates grown to anaerobic/low aeration and moderate/high aeration are different from one another, future representative isolates will be examined under both anaerobic conditions and moderate degrees of aeration to examine the variations in protein expression in more detail.

4.3.4 Effect of growth temperature on the growth rate and OMP profile

Isolates RD124 and RD158 were grown at temperatures of 8, 15, 22, 28, 37 and 45°C. This was to examine the effect of a large range of temperatures, including extremes of 8 and 45°C. While 15°C is a temperature more likely to be associated with the host, 22 and 28°C have both been used extensively within the characterisation of *Y. ruckeri*. The human pathogenic *Yersinia* sp. grow optimally at 37°C so this was chosen as a comparison.

Growth was most rapid at 28°C for both RD124 (Fig. 4.5A) and RD158 (Fig. 4.5B). Isolate RD124 entered logarithmic growth after 8 h, while RD158 entered after 6 h. Both isolates grew very quickly, reaching stationary phase after 12 h. Mid-log phase was at OD_{600nm} 1.2. At 22°C, the temperature at which the majority of characterisation studies have taken place, isolates entered logarithmic phase after ~10 h, before reaching stationary phase at 16 h. Mid-log phase was at OD_{600nm} 1.2. Isolates RD124 and RD158 grew at slightly different rates at 37°C. Isolate RD124 entered logarithmic growth after 8 h, while RD158 enterered after

4 h. RD124 and RD158 had reached stationary phase after 18 and 12 h, respectively. The OD_{600nm} at stationary phase was 1.6, with mid-log phase growth at approximately OD_{600nm} 0.9. Isolates grown at both 15°C and 8°C grew much more slowly, taking 16 h and 18 h to enter logarithmic growth, respectively. Isolates were unable to grow at 45°C.

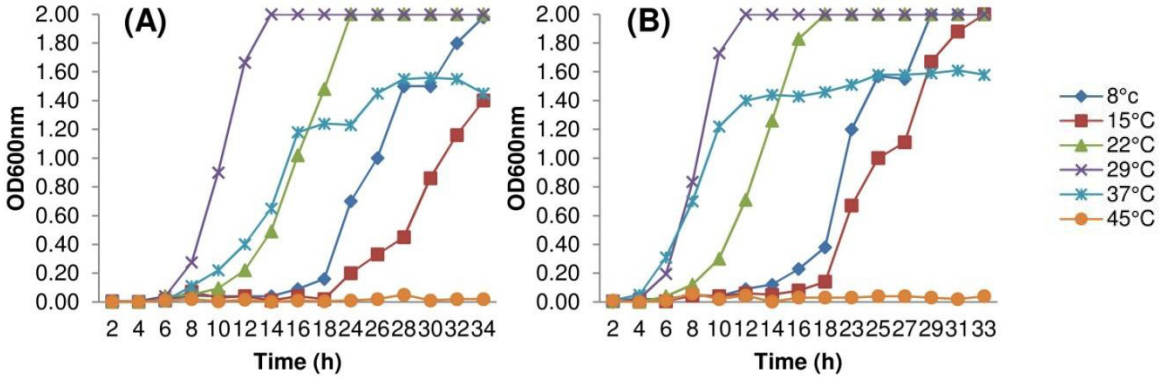


Fig. 4.5 Effect of temperature on the growth in representative isolates of *Y. ruckeri*. Isolates RD124 **(A)** and RD158 **(B)** were grown at 8, 15, 22, 28 and 37° C. OD_{600nm} was measured at two hourly intervals.

When isolates were grown at 8°C (Fig. 4.6, Lanes 1 and 6), 15°C (Fig. 4.6, Lanes 2 and 7) and 22°C (Fig. 4.6, Lanes 3 and 8) there were no apparent differences in the OM profiles from each other. However, at 28°C (Fig. 4.6, Lane 4) RD124 showed a marked decrease in expression of a protein of apparent molecular weight (MW) 37 kDa (presumed OmpF; Fig. 4.6, lane 4, arrow[1]) and up-regulation of a 39.5 kDa protein (presumed OmpC; Fig. 4.6, lane 4, arrow[2]). In isolate RD158, where the up-regulation of the 39.5 kDa protein (Fig. 4.6, lane 9, arrow[3]) and down regulation of a 37.5 kDa protein (Fig. 4.6, lane 9, arrow[4]) was more obvious. This differential expression was even more pronounced at 37°C in both isolates (Fig. 4.6, Lanes 5 and 10, arrows[5]), where the lower MW protein was not present in isolate RD124, and barely visible in isolate RD158.

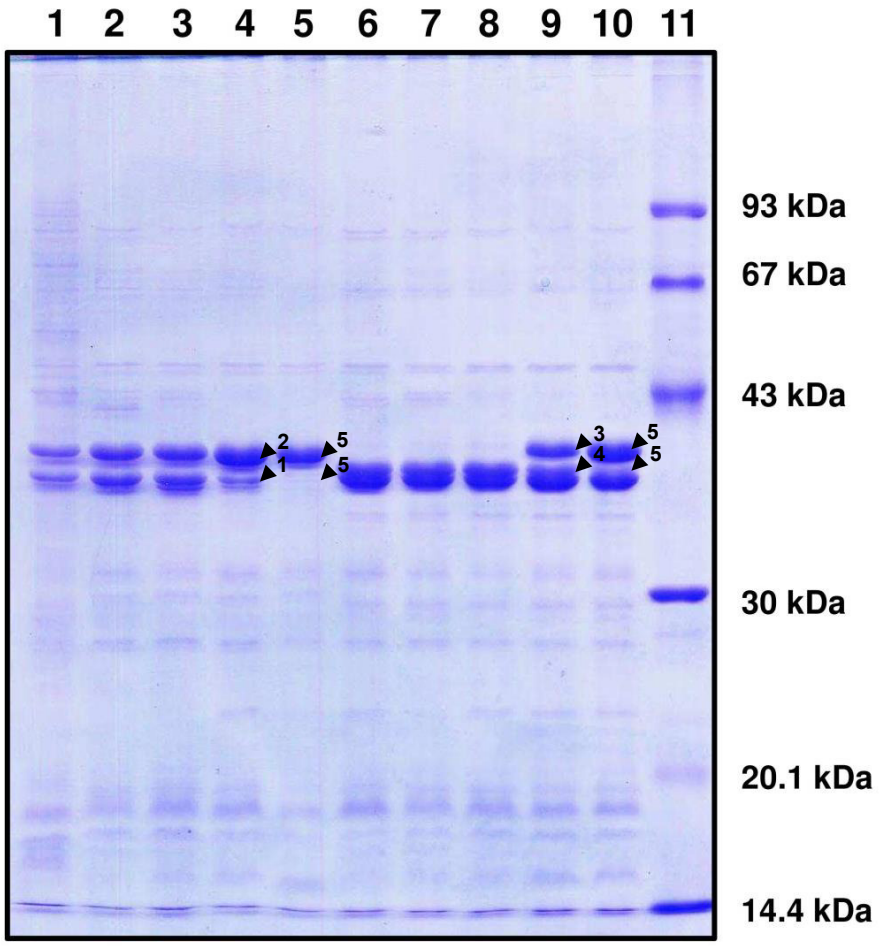


Fig. 4.6 The effect of temperature on the OMP profiles of *Y. ruckeri.*
Isolates RD124 (lanes 1-4) and RD158 (lanes 5-8) were grown at 8, 15, 22, 28 and 37° C in lanes 1 to 5, and 6 to 10 respectively. Lane 11 is a MW standard (GE Healthcare, UK).

As the profiles were not drastically altered at lower temperatures between 8 and 22°C, and several studies have used 28°C for the growth of the bacteria, both 22°C and 28°C will be used for further proteomic analysis of the selected representative isolates.

4.3.5 Effect of sodium chloride (NaCl) on the growth rate and OMP profile

Isolates RD124 and RD158 were grown under different concentrations of NaCl to determine the effect of salt on the rate of growth of *Y. ruckeri*. NaCl is the major salt component of seawater. Isolates were grown in TSB supplemented with 0% (0 ppt), 1% (10 ppt), 2% (20 ppt), 3% (30 ppt), 6% (60 ppt) and 9% (90 ppt) NaCl in order to determine the optimal concentration of NaCl that would induce alterations in the OMP profile without inhibiting growth substantially. In

isolates RD124 (Fig. 4.7A) and RD158 (Fig. 4.7B), 1% NaCl did not affect the rate of growth noticeably. Entry to logarithmic growth was disrupted by approximately 1 h from 0% NaCl. However, isolates still reached stationary phase by 15 h in both cases. Growth in 2% NaCl had a more noticeable effect, delaying logarithmic growth phase by approximately 2 h. Growth in 3% NaCl delayed the growth rate dramatically, with isolates not entering logarithmic growth until after 20 h. The growth was much reduced also, as isolates reached a maximum OD_{600nm} of 1.6, with mid-log phase at approximately OD_{600nm} 1.0 (after 25 h). Both RD124 and RD158 were unable to grow at 6% or 9% NaCl.

The OMP profiles were examined after growth at 0%, 1%, 2% and 3% NaCl in isolates RD124 and RD158 (Fig. 4.8). Concentrations 6% and 9% were not examined, due to the inability of the bacteria to grow at these concentrations. As the concentration of NaCl increased, there was a noticeable change in the expression of the major porins presumed to be OmpC and OmpF, in both isolates. There was an up-regulation of a protein at 39.5 kDa (presumably OmpC; Fig. 4.8, lanes 4 and 8, arrows[1]) in tandem with down-regulation of 37 kDa protein in isolate RD124 and 37.5 kDa protein in RD158 (presumably OmpF; Fig. 4.8, lanes 4 and 8, arrows[2]). An additional protein of ~50 kDa is apparent after growth in 0% NaCl (Fig. 4.8, lanes 1 and 5, arrows[3]). However, this band is absent in the remaining samples.

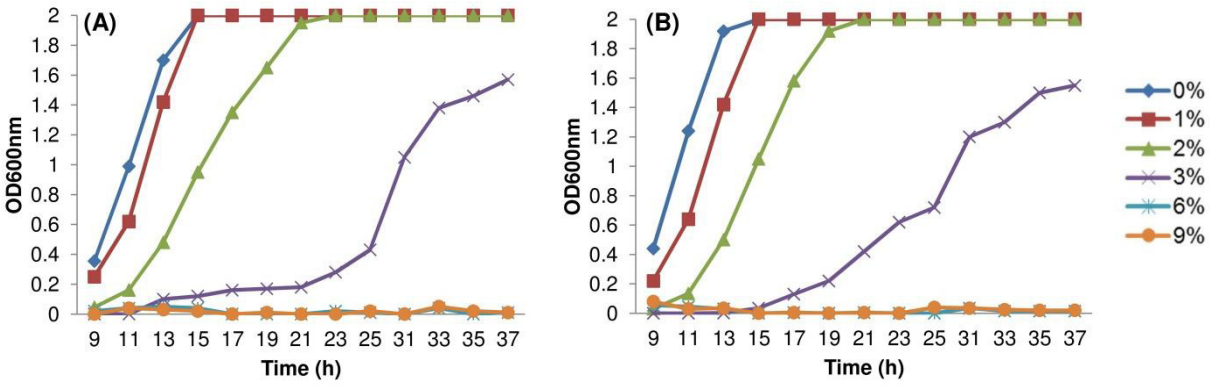


Fig. 4.7 Effect of NaCl on the growth in representative isolates of *Y. ruckeri*.
Isolates RD124 **(A)** and RD158 **(B)** were grown in 0%, 1%, 2%, 3%, 6% and 9% NaCl in TSB. Rate of growth was measured using OD_{600nm} at two hourly intervals.

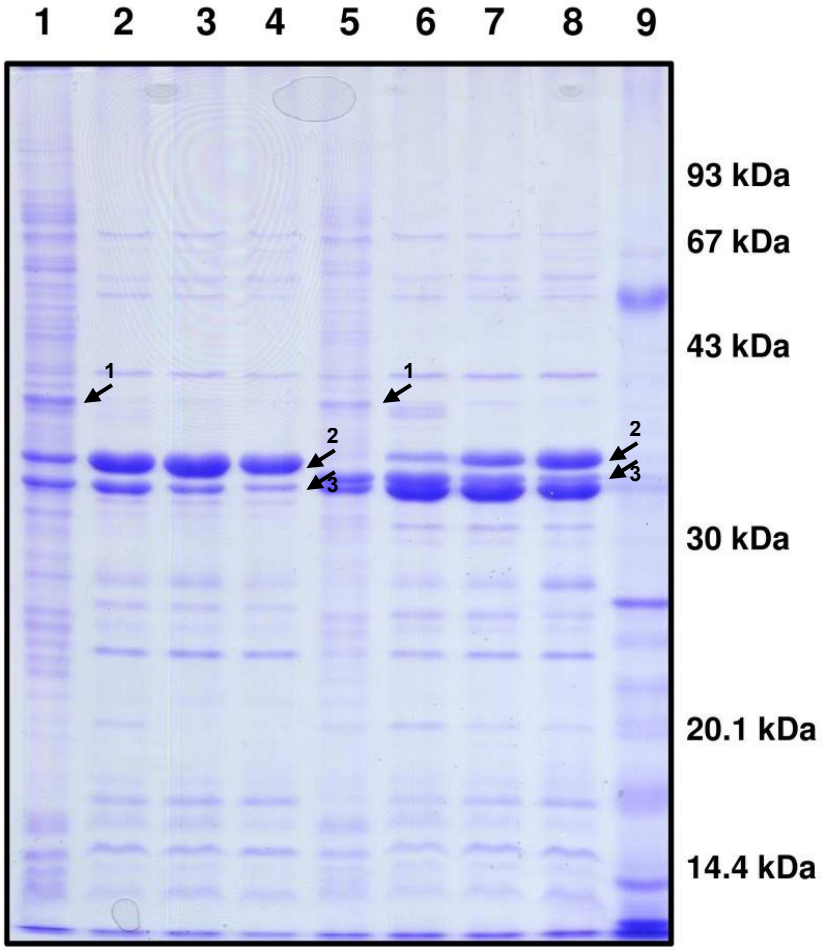


Fig. 4.8 Effect of NaCl on the OMP profiles of *Y. ruckeri.*
Isolates RD124 (lanes 1-4) and RD158 (lanes 5-8) were grown in 0%, 1%, 2% and 3% NaCl added in lanes 1 to 4, and 5 to 8 respectively. Lane 9 is a MW standard (GE Healthcare, UK).

4.3.6 **Effect of seawater medium on the growth rate**

In view of the effects of NaCl on the growth of bacteria, we subsequently decided to examine the effects of growing *Y. ruckeri* in seawater. The average salinity of seawater is approximately 3.5% (35 ppt) although this does vary depending on the ocean and the temperature. The rate of growth of *Y. ruckeri* was examined in seawater medium. Concentrations of 0%, 1.5%, 2.5% and 3.5% seawater medium were prepared and the rate of growth of isolates RD124 and RD158 examined.

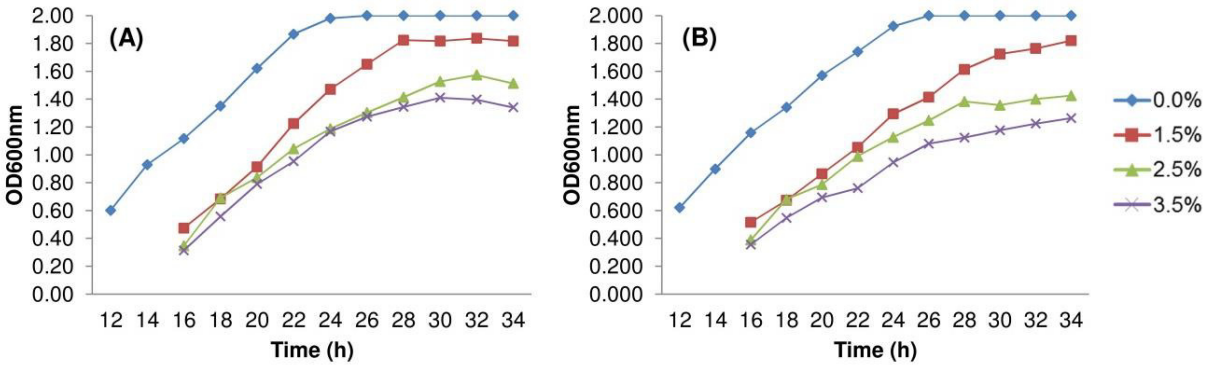


Fig. 4.9 Effect of seawater concentration on the growth of representative isolates of *Y. ruckeri*.
Isolates RD124 **(A)** and RD158 **(B)** were grown in seawater medium at concentrations of 1.5, 2.5 and 3.5%. Rate of growth was measured using OD_{600nm} at two hourly intervals.

In a similar response to growth in NaCl, the rate of growth was reduced as the salinity increased. In isolates RD124 (Fig. 4.9A) and RD158 (Fig. 4.9B), growth in 1.5% seawater medium altered the rate of growth slightly, with entry to logarithmic growth disrupted by approximately 3 h. However, isolate RD124 and RD158 reached stationary phase after 28 and 30 h, respectively. Both isolates achieved a final OD_{600nm} of 1.85. Growth in 2.5% seawater medium had a more noticeable effect, delaying logarithmic growth phase by approximately 4 h, and with a final OD_{600nm} of ~1.5 in both isolates RD124 and RD158. Growth in 3.5% seawater medium delayed the growth rate further in relation to the 0% seawater medium, with isolates not entering logarithmic growth until after 16 h, which was slightly quicker than in NaCl. The growth was much reduced also, as isolates reached a maximum OD_{600nm} of 1.3, with mid-log phase at OD_{600nm} ~0.8 (after 25 h).

Based on the results of the NaCl growth curve and OMP profile analysis, combined with the seawater medium growth curves, 3.5% seawater medium will be selected as the optimum growth media for proteomic analysis of the remaining representative isolates. This mimics the conditions of seawater more accurately, in terms of both salt concentration and composition and is more representative than NaCl alone.

4.3.7 **Effect of iron-depletion on the growth rate and OMP profile**

The effect of iron on bacterial growth rate has been well documented, and is discussed in detail elsewhere (1.4.3.4.2). Isolates RD124 and RD158 were grown in concentrations of 2,2'-dipyridyl (Sigma, UK) ranging from 170 μM to 230 μM (Fig. 4.10). There was a noticeable decrease in the growth capabilities of both isolates when the iron concentration in the media was reduced. In the iron-replete medium (not containing 2,2'-dipyridyl), both isolates entered logarithmic growth after approximately 8 h, reached mid-log phase after 12 h, and stationary phase after 20 h. As the concentration of 2,2'-dipryidyl increased, the rate of growth of both RD124 and RD158 was disrupted, similarly. At a concentration of 170 μM 2,2'-dipyridyl entry to logarithmic growth was at 12 h, mid-log phase was at 20 h and entry to stationary phase at 26 h for isolate RD124. For isolate RD158 entry to logarithmic growth was at 10 h, mid-log phase was at 14 h and entry to stationary phase at 18 h. At a concentration of 230 μM 2,2'-dipyridyl entry to logarithmic growth was at 16 h, mid-log phase was at 24 h and entry to stationary phase at 32 h for isolate RD124. For isolate RD158 entry to logarithmic growth was at 12 h, mid-log phase was at 20 h and entry to stationary phase at 24 h.

In order to determine the effect on the OMP profiles that this alteration in iron availability was having, bacteria were grown in TSB supplemented with concentrations of 170, 200 and 230 μM 2,2'-dipyridyl (Fig. 4.11). It was apparent that in the presence of 2,2'-dipyridyl there was up-regulation of four additional protein bands of ~72, ~74, ~76 and ~78 kDa respectively (Fig. 4.11, lanes 4 and 8, arrows[1]). There was also up-regulation of a ~39.5 kDa protein, which has been seen under other conditions of stress, and is presumed to be OmpC (Fig. 4.11, lanes 4 and 8, arrows[2]).

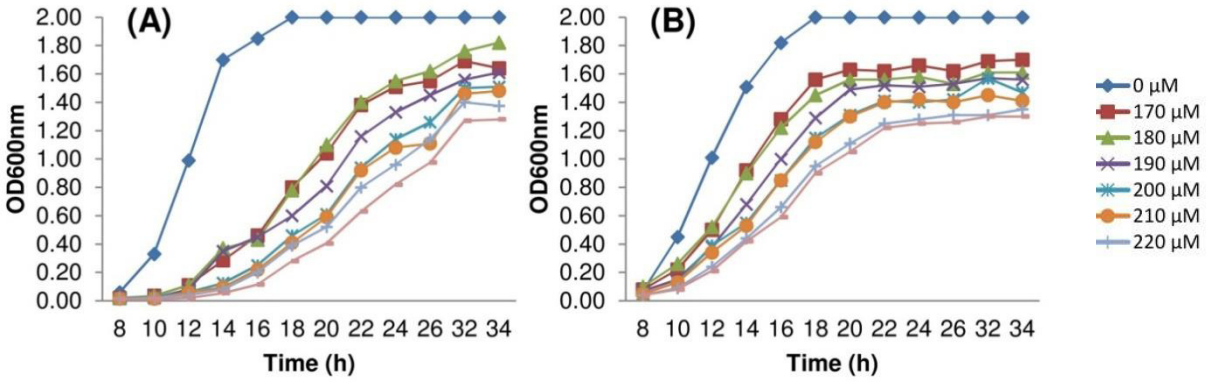


Fig. 4.10 Effect of 2,2'-dipyridyl on the growth of representative isolates of *Y. ruckeri*. Isolates RD124 **(A)** and RD158 **(B)** were grown in concentrations ranging from 170 µM to 230 µM 2,2'-dipyridyl in TSB (Oxoid, UK). Rate of growth was measured using OD_{600nm} at two hourly intervals.

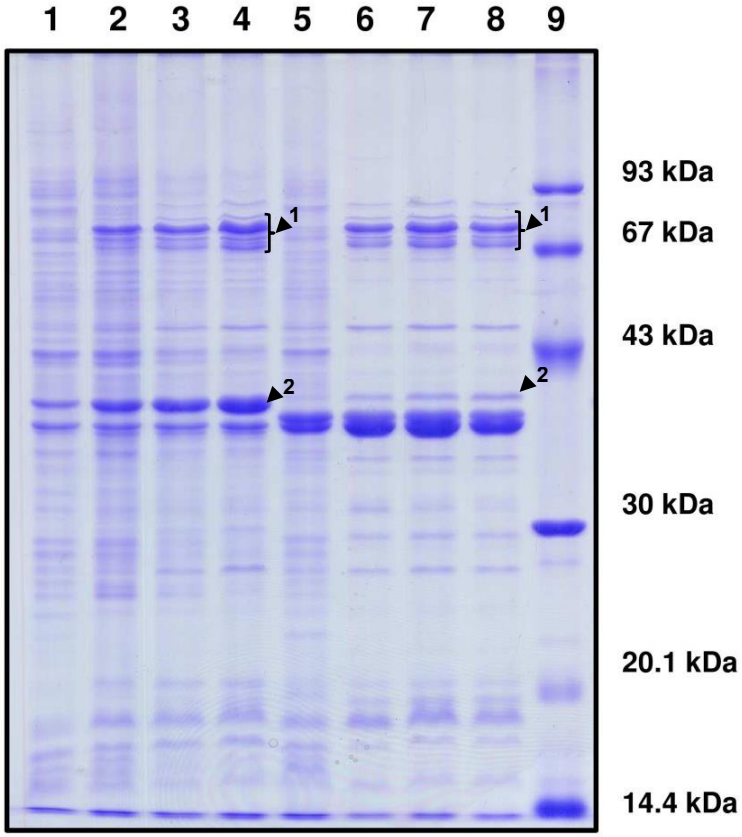


Fig. 4.11 Effect of 2,2'-dipyridyl on the OMP profiles of *Y. ruckeri*. Isolates RD124 (lanes 1-4) and RD158 (lanes 5-8) were grown in TSB supplemented with 2,2'-dipyridyl at concentrations of 0, 170, 200 and 230 µM in lanes 1 to 4, and 5 to 8 respectively. Lane 9 is a MW standard (GE Healthcare, UK).

Since there is little to no difference in the abundance of these iron-regulated proteins after growth in the presence of different concentrations of 2,2'-dipyridyl, 200 µM will be chosen as the optimum concentration for proteomic analysis of the representative reference isolates. This concentration did not

inhibit the growth rate of the bacteria extensively, while allowing high expression of additional proteins.

4.3.8 Effect of the optimised growth conditions on the growth rate of representative isolates of *Y. ruckeri*.

Eight isolates that were representative of the major clonal groups that cause infection in Atlantic salmon and rainbow trout were grown aerobically at 22 and 28°C, anaerobically, in iron-depleted media (200 μM 2,2'-dipyridyl) and in an artificial seawater medium (35 ppt). The details of these isolates have been discussed previously (section 3.4.3). Initially, the effect of these different conditions on the rate of growth was examined.

4.3.8.1 Aerobic growth of representative isolates at 22°C.

There was no discernible difference in the rate of growth between the majority of isolates (Fig. 4.12). Isolates RD124 and RD382 grew slightly slower, entering logarithmic growth after 10 h, as opposed to 8 h for the remaining isolates, and stationary phase by 26 h, as opposed to 18 h. This may simply have been due to an unintentionally lower starter inoculum.

4.3.8.2 Aerobic growth of representative isolates at 28°C.

Several studies have grown the bacteria at 28°C, while others have used 22°C. It is suggested that certain virulence factors are more readily expressed under temperatures more closely associated with the host, while others are up-regulated at 28°C (Section 1.2.3). In order to examine this, isolates were grown at 28°C and the OMP profiles examined. The majority of isolates grew more rapidly at 28°C than at 22°C (Fig. 4.13). Isolates RD28, RD64, RD354, RD366 and RD420 began to grow much more rapidly, entering logarithmic growth before 6 h and stationary phase after 12 h.

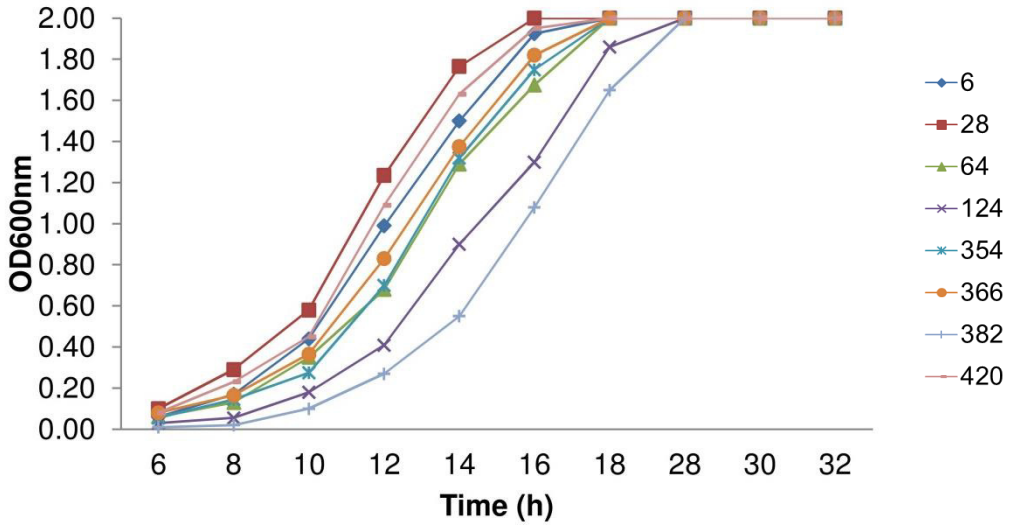


Fig. 4.12 Growth of representative isolates of *Y. ruckeri* at 22°C.
Isolates were grown aerobically at 22°C. Isolates RD6, RD28, RD64 and RD124 were recovered from rainbow trout, while isolates RD354, RD366, RD382 and RD420 were recovered from Atlantic salmon. Growth was measured at OD_{600nm} at two hourly intervals.

However, isolates RD6, RD124 and RD382 grew slower, entering logarithmic growth after 14 h, 10 h, and 8 h respectively. While isolates RD124 and RD382 still reached stationary phase by 16 h and 14 h respectively, isolate RD6 did not until after 18 h. This was substantially slower than when grown at 22°C.

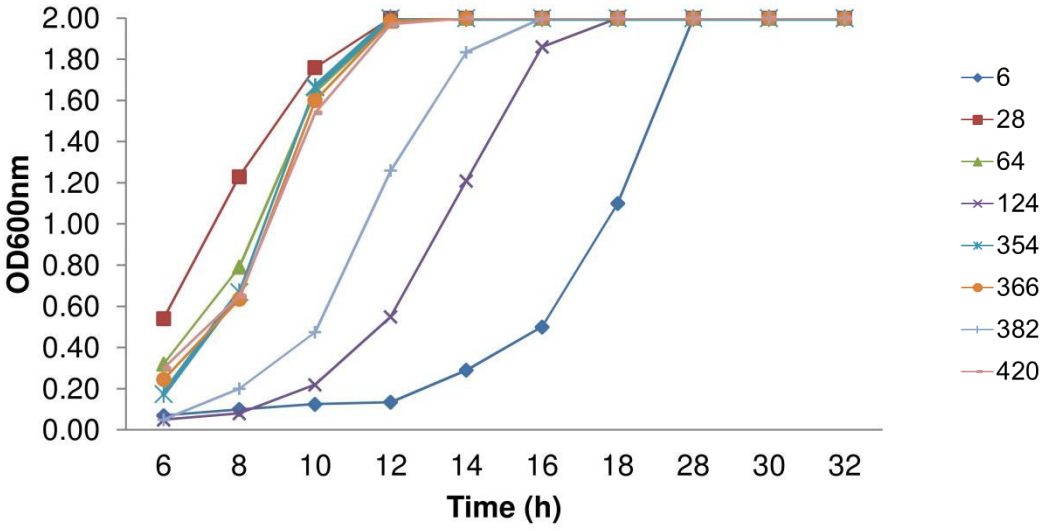


Fig. 4.13 Growth of representative isolates of *Y. ruckeri* at 28°C.
Isolates were grown aerobically at 28°C. Isolates RD6, RD28, RD64 and RD124 were recovered from rainbow trout, while isolates RD354, RD366, RD382 and RD420 were recovered from Atlantic salmon. Growth was measured at OD_{600nm} at two hourly intervals.

4.3.8.3 Anaerobic growth of representative isolates.

When compared to aerobic growth rates (Fig. 4.14), the growth was substantially reduced under anaerobic conditions (Fig. 4.14). The maximum OD_{600nm} achieved was ~1.0 in all isolates, when compared to an OD_{600nm} 2.0 under aerobic conditions. Initially, isolates grew relatively quickly, entering logarithmic growth after 6 h. However, they had achieved stationary phase by 12 h presumably as available oxygen was consumed.

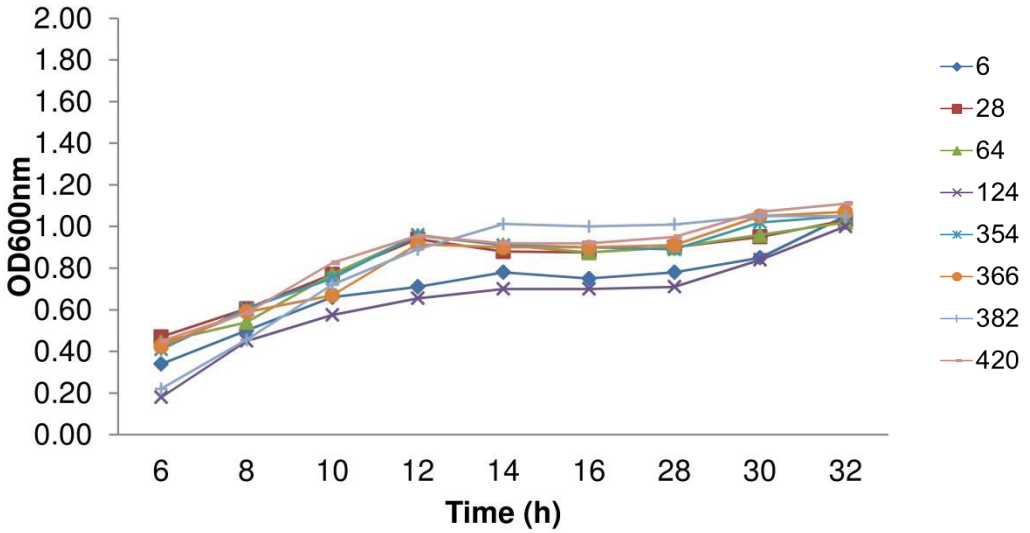


Fig. 4.14 Growth of representative isolates of *Y. ruckeri* grown anaerobically. Isolates were grown anaerobically at 22°C. Isolates RD6, RD28, RD64 and RD124 were recovered from rainbow trout, while isolates RD354, RD366, RD382 and RD420 were recovered from Atlantic salmon. Growth was measured at OD_{600nm} at two hourly intervals.

4.3.8.4 Iron-depleted growth of representative isolates.

Under iron-depleted growth conditions, isolates took much longer to reach log phase (~12 h) in comparison to iron-replete conditions, before reaching stationary phase at ~28 h (Fig. 4.15). Isolates RD124 and RD382, recovered from rainbow trout and Atlantic salmon respectively, did not grow as efficiently in iron-depleted media as the remaining isolates. By 28 h, RD124 and RD382had reached a maximum growth OD_{600nm} of 1.20, while the remainder had achieved 1.60.

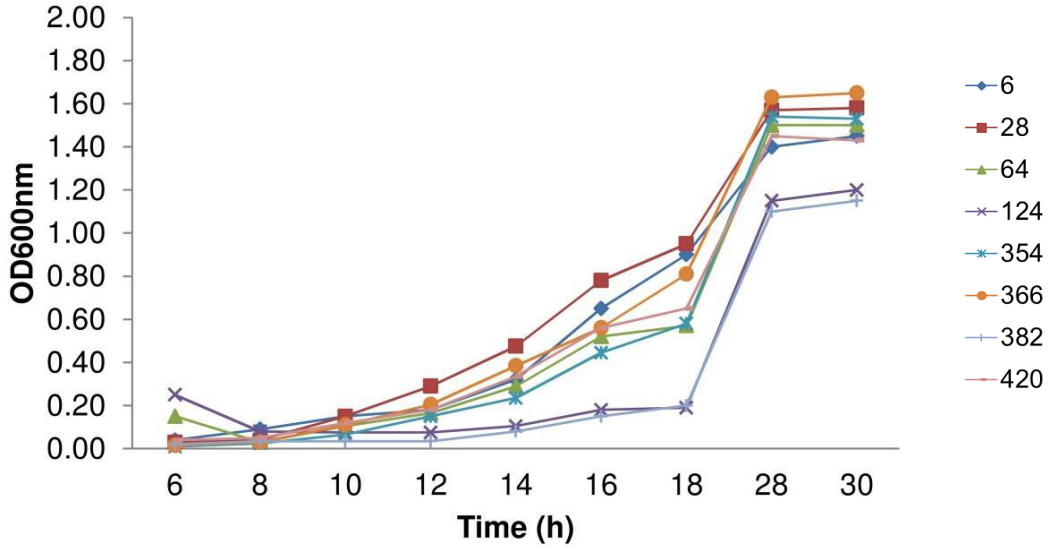


Fig. 4.15 Growth of representative isolates of *Y. ruckeri* grown under iron-depletion. Isolates were grown in TSB supplemented with 2,2'-dipyrdidyl to a final concentration of 200 µM. Isolates RD6, RD28, RD64 and RD124 were recovered from rainbow trout, while isolates RD354, RD366, RD382 and RD420 were recovered from Atlantic salmon. Growth was measured at OD_{600nm} at two hourly intervals.

4.3.9 Effect of optimal growth conditions on the OMP profiles of representative reference isolates

In order to directly compare variation in the OMP profiles within individual isolates caused by the different growth conditions, OMP samples representing each isolate were run alongside each other on a single gel. Examining the OMP profiles directly against one another allows initial comparison of the OMP presence, absence and abundance under different growth conditions.

4.3.9.1 Isolate RD6

Isolate RD6 is a biotype 2, serotype O1, OMP-type 1b isolate recovered from rainbow trout. This isolate is representative of the majority of isolates recovered from rainbow trout in the UK. Under aerobic conditions (Fig. 4.16A, lane 1), the major proteins used to characterise this OMP-type are difficult to distinguish from one another, without heat solubilisation experiments prior to SDS-PAGE (discussed previously). It is also quite challenging to identify the up-regulation of the major 39.5 kDa protein (OmpC) under the variant growth conditions (Fig. 4.16A, lanes 2 to 5, arrow[1]).

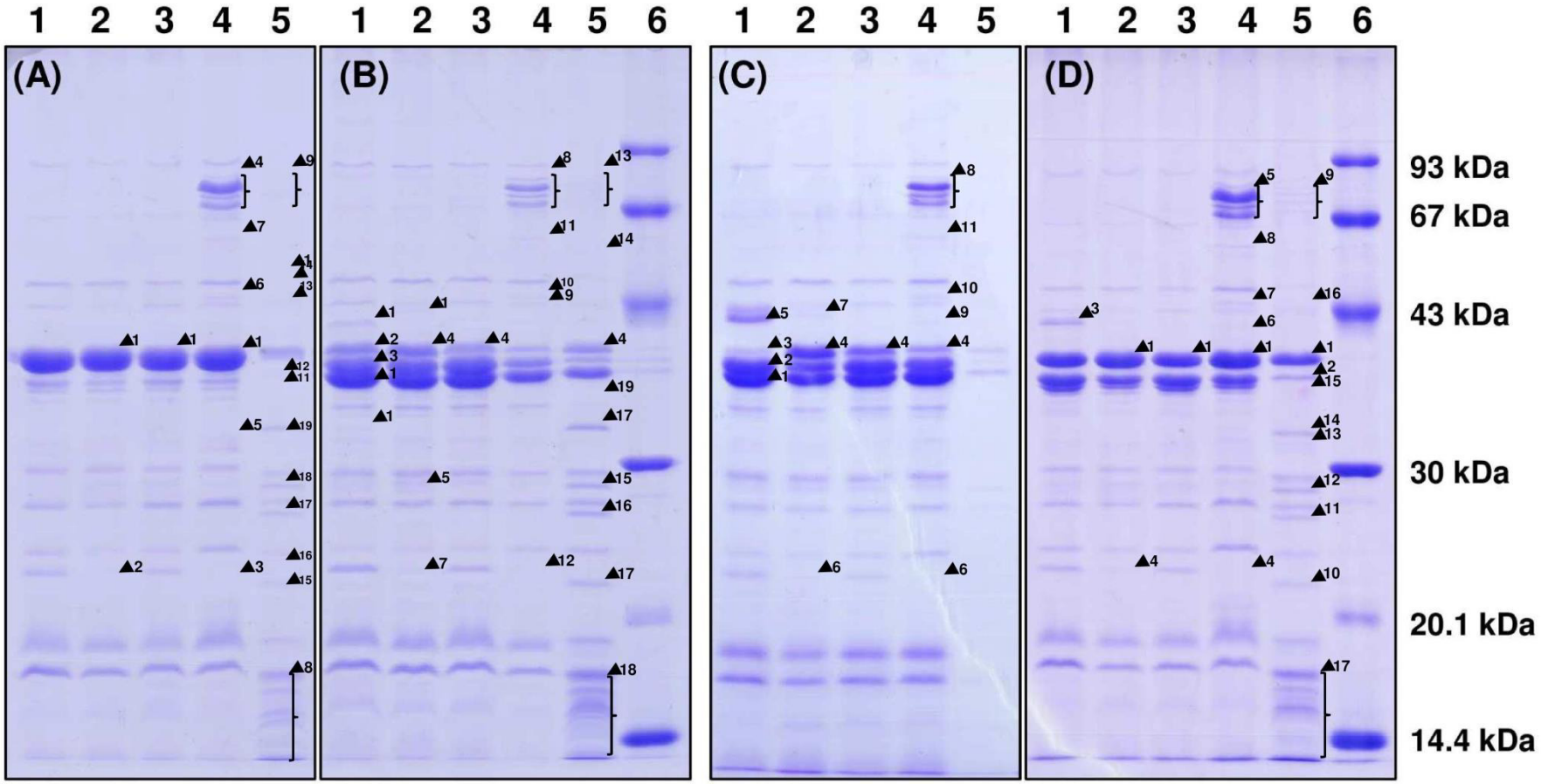


Fig. 4.16 OMP profiles of representative isolates of *Y. ruckeri* recovered from rainbow trout grown under conditions mimicking environmental and *in vivo* growth.

Isolates **(A)** RD6, **(B)** RD28, **(C)** RD64 and **(D)** RD124 were grown under varying conditions. Lanes 1 to 5 in each panel represent OMP profiles of isolates grown under aerobic (22°C), anaerobic, aerobic (28°C), iron-depleted (170 µM) and in seawater medium (SEAMIX; 3.5%) conditions. Lane 6 is a MW standard (GE Healthcare, UK).

The OMP profiles are almost identical under aerobic (22°C) (Fig. 4.16A, lane 1), anaerobic (Fig. 4.16A, lane 2) and aerobic (28°C) (Fig. 4.16A, lane 3) growth conditions. However, a protein of ~23 kDa is absent under anaerobic growth (Fig. 4.16A, lane 2, arrow[2]). This protein is also absent under iron-depleted and seawater medium growth (Fig. 4.16A, lanes 4 and 5, arrow[3]).

The most noticeable changes in OMP profile occur during growth in iron-depleted medium and in seawater medium (Fig. 4.16A, lanes 4 and 5). When cells were grown in an iron-depleted medium there was up-regulation of four high molecular weight (HMW) proteins ranging from ~70 to ~80 kDa (Fig. 4.16A, lane 4, arrows[4]). Additionally, proteins of ~33 kDa (Fig. 4.16A, lane 4, arrow[5]), ~45 kDa (Fig. 4.16A, lane 4, arrow[6]) and ~55 kDa (Fig. 4.16A, lane 4, arrow[7]) are present. In seawater medium several additional proteins are present in the low molecular weight (LMW) range, which are absent from the other growth conditions (Fig. 4.16A, lane 5, arrow[8]). Proteins in the HMW range similar to those seen during iron-restricted growth are also apparent (Fig. 4.16A, lane 5, arrows[9]). Proteins including major bands at ~23 kDa (Fig. 4.16A, lane 5, arrows[10]), ~34 kDa (Fig. 4.16A, lane 5, arrows[11]), ~36 kDa (Fig. 4.16A, lane 5, arrows[12]), ~50 kDa (Fig. 4.16A, lane 5, arrows[13]) and ~50 kDa (Fig. 4.16A, lane 5, arrows[14]) are now absent from the OMP profiles. Three additional protein bands at ~20 kDa (Fig. 4.16A, lane 5, arrows[15]), ~24 kDa (Fig. 4.16A, lane 5, arrows[16]), ~26 kDa (Fig. 4.16A, lane 5, arrows[17]), ~28 kDa (Fig. 4.16A, lane 5, arrows[18]) and 31 kDa (Fig. 4.16A, lane 5, arrows[19]) are also present in isolate RD6 grown in seawater medium.

4.3.9.2 Isolate RD28

Isolate RD28 is a biotype 1, serotype O5, OMP-type 2a isolate recovered from rainbow trout. This isolate was included for comparison to isolates of the same OMP-type recovered from Atlantic salmon. However, this clonal group has not been a major cause of disease in rainbow trout in the UK. Under aerobic growth conditions (Fig. 4.16B), the major proteins OmpA (Fig. 4.16B, lane 1, arrow[1]), OmpC (Fig. 4.16B, lane 1, arrow[2]) and OmpF (Fig. 4.16B, lane 1, arrow[3]) used to characterise this OMP-type are visible. However, OmpC (39.5 kDa) is not abundant. Under the other growth conditions, there was an increase in

expression of OmpC, particularly under anaerobic, aerobic (28°C) and growth in seawater medium (Fig. 4.16B, lanes 2, 3 and 5, arrow[4]). Under anaerobic growth conditons, two additional proteins of ~29 kDa (Fig. 4.16B, lane 2, arrow[5]) and ~93 kDa (Fig. 4.16B, lane 2, arrow[6]) were apparent. A protein of ~23 kDa was less abundant than when grown aerobically (Fig. 4.16B, lane 2, arrow[7]). When this isolate was grown at 28°C (Fig. 4.16B, lane 3) there was very little difference in the OMP profile compared to aerobic growth (Fig. 4.16B, lane 2). However, the abundance of OmpC was much greater at 28°C.

The most noticeable changes in OMP profile occur during growth in iron-deplete medium (Fig. 4.16B, lane 4) and in seawater medium (Fig. 4.16B, lane5). The presence of four additional HMW proteins under iron-depleted growth conditions was the most obvious change (Fig. 4.16B, lane 4, arrow[8]), while several other proteins were also present at ~47 kDa (Fig. 4.16B, lane 4, arrow[9]), ~56 kDa (Fig. 4.16B, lane 4, arrow[10]) and ~60 kDa (Fig. 4.16B, lanes 4 and 5, arrow[11]). A protein of ~23 kDa, which was significantly down regulated under anaerobic growth conditions was now absent in cells grown under iron-depleted growth conditions (Fig. 4.16B, lane 4, arrow[12]). In growth in seawater medium, three of the four HMW proteins (Fig. 4.16B, lane 5, arrow[13]) and the 47 kDa protein (Fig. 4.16B, lane 5, arrow[14]) seen under iron-restricted growth were also expressed. The ~29 kDa protein present after growth under anaerobic conditions was also significantly upregulated (Fig. 4.16B, lane 5, arrow[15]), while a ~26 kDa protein (Fig. 4.16B, lane 5, arrow[16]) and a ~33 kDa protein (Fig. 4.16B, lane 5, arrow[17]) were also expressed. Several LMW proteins were expressed under growth in seawater medium also (Fig. 4.16B, lane 5, arrow[18]). A protein of ~32 kDa, present under all growth conditions, was absent when the cells were grown in seawater medium (Fig. 4.16B, lane 5, arrow[19]).

4.3.9.3 Isolate RD64

Isolate RD64 is a biotype 1, serotype O2, OMP-type 2a isolate recovered from rainbow trout. This isolate was included for comparison to isolates of one of the major clonal groups recovered from Atlantic salmon. However, this clonal group has not been a major cause of disease in rainbow trout in the UK.

Under aerobic growth (22°C) (Fig. 4.16C, lane 1), the major proteins OmpA (Fig. 4.16C, lane 1, arrow[1]) and OmpF (Fig. 4.16C, lane 1, arrow[2]) are abundant, while OmpC (Fig. 4.16C, lane 1, arrow[3]) is less so. OmpC is upregulated under anaerobic, aerobic (28°C) and iron-depleted growth (Fig. 4.16C, lanes 2, 3 and 4, arrow[4]). A ~43 kDa protein is present only under aerobic (22°C) growth conditions (Fig. 4.16C, lane 1, arrow[5]). As has been seen in other isolates, under anaerobic and iron-depleted growth conditions (Fig. 4.16C, lanes 2 and 4, respectively) growth condtions, a protein of 23 kDa is absent (Fig. 4.16C, lanes 2 and 4, arrow[6]). A protein of ~43 kDa is absent under anaerobic growth, also (Fig. 4.16C, lane 2, arrow[7]).

Again, under iron-depleted growth conditions the OMP profile is altered substantially. The presence of four additional HMW proteins was the most obvious change (Fig. 4.16C, lane 4, arrow[8]), while several other proteins were also present at ~47 kDa (Fig. 4.16B, lane 4, arrow[9]), ~56 kDa (Fig. 4.16B, lane 4, arrow[10]) and ~60 kDa (Fig. 4.16B, lanes 4 and 5, arrow[11]). The OMP profile from cells grown in seawater medium was not apparent enough to draw conclusions as to the apparent presence or absence of proteins at this stage.

4.3.9.4 Isolate RD124

Isolate RD124 is a biotype O1, serotype O1, OMP-type 3a isolate recovered from rainbow trout. This isolate is representative of the 'Hagerman' type strain, which is the major disease causing strain worldwide, until the emergence of the biotype 2 variant.

There is up-regulation of a 39.5 kDa (presumably OmpC) under anaerobic, aerobic (28°C), iron depleted and growth in seawater medium. However, this is difficult to distinguish due to the similarity in MW between that and OmpA (Fig. 4.16D, lanes 2 to 5, arrow[1]). A protein of ~36 kDa (presumably OmpF) is down regulated under growth in seawater medium (Fig. 4.16D, lane 5, arrow[2]). A protein of ~41 kDa is present only under aerobic (22°C) growth (Fig. 4.16D, lane 1, arrow[3]). As has been seen in other isolates, under anaerobic and iron-depleted growth (Fig. 4.16D, lanes 2 and 4, respectively) condtions, a protein of 23 kDa is absent (Fig. 4.16D, lanes 2 and 4, arrow[6]).

The most noticeable changes in OMP profile occur during growth in iron-depleted medium and in seawater medium. Under iron-depleted growth conditions four additional HMW proteins were upregulated (Fig. 4.16D, lane 4, arrow[5]), while several other proteins were also present at ~41 kDa (Fig. 4.16D, lane 4, arrow[6]), ~47 kDa (Fig. 4.16D, lane 4, arrow[7]) and ~60 kDa (Fig. 4.16D, lane 4, arrow[8]). In artificial seawater, the four HMW proteins upregulated under iron-deplete conditions are not as obvious as they are in other isolates, although there is still an indication that they may be slightly up regulated (Fig. 4.16D, lane 5, arrow[9]). Proteins of ~21 kDa (Fig. 4.16D, lane 5, arrow[10]), ~26 kDa (Fig. 4.16D, lane 5, arrow[11]), ~29 kDa (Fig. 4.16D, lane 5, arrow[12]), ~33 kDa (Fig. 4.16D, lane 5, arrow[13]) and ~34 kDa (Fig. 4.16D, lane 5, arrow[14]) are upregulated. Proteins of ~37 kDa (Fig. 4.16D, lane 5, arrow[15]) and 44 kDa (Fig. 4.16D, lane 5, arrow[16]) are absent under this condition. The several LMW proteins present under this growth condition in other isolates are not as abundant but still present (Fig. 4.16D, lane 5, arrow[17]).

4.3.9.5 Isolate RD354

Isolate RD354 is a biotype 1, serotype O2, OMP-type 2a isolate recovered from Atlantic salmon in the UK. This isolate is representative of 27.5% of isolates recovered from Atlantic salmon in our collection and has been a major clonal group in Atlantic salmon over the previous 14 years.

Under aerobic growth (22°C) (Fig. 4.17A, lane 1), the major proteins OmpA (Fig. 4.17A, lane 1, arrow[1]) and OmpF (Fig. 4.17A, lane 1, arrow[2]) are abundant, while OmpC (C, lane 1, arrow[3]) is less so. OmpC is upregulated under anaerobic, aerobic (28°C), iron-depleted growth and growth in seawater medium (Fig. 4.17A, lane2, 3, 4 and 5, arrow[4]). Proteins of ~31 kDa (Fig. 4.17A, lane 1, arrow[5]) and ~43 kDa (Fig. 4.17A, lane 1, arrow[5]) are present only under aerobic (22°C) growth conditions.

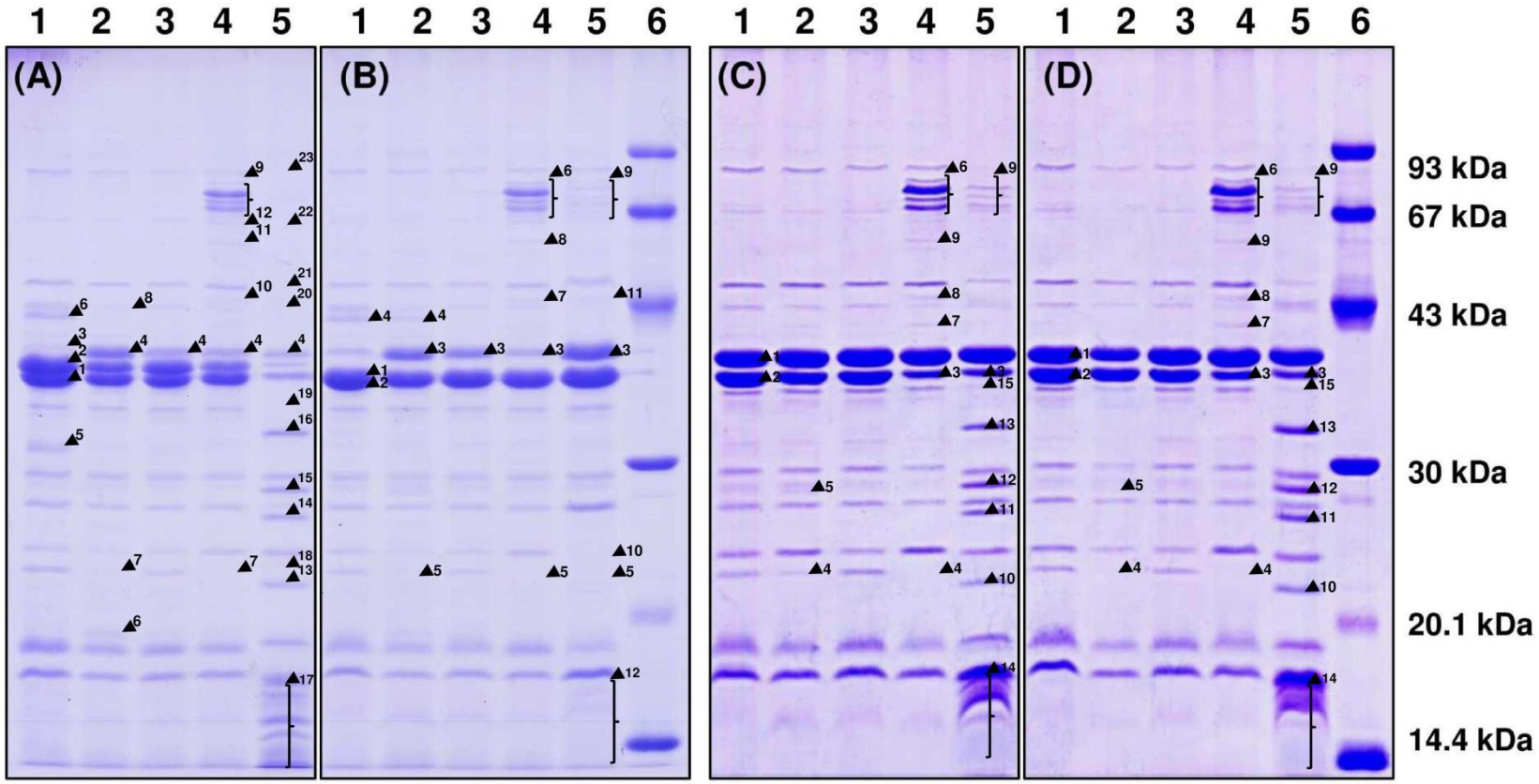


Fig. 4.17 OMP profiles of representative isolates of *Y. ruckeri* recovered from Atlantic salmon grown under conditions mimicking environmental and *in vivo* growth.

Isolates RD354 **(A)**, RD366 **(B)**, RD382 **(C)** and RD420 **(D)** were grown under varying conditions. Lanes 1 to 5 in each panel represent OMP profiles of isolates grown under aerobic (22°C), anaerobic, aerobic (28°C), iron-depleted (170 µM) and in seawater medium (SEAMIX; 3.5%) conditons. Lane 6 is a MW standard (GE Healthcare, UK).

As has been seen in other isolates, under anaerobic and iron-depleted growth (Fig. 4.17A, lanes 2 and 4, respectively) conditions, a protein of 23 kDa is absent (Fig. 4.17A, lanes 2 and 4, arrow[7]). A protein of ~43 kDa is absent under anaerobic growth conditions, also (Fig. 4.17A, lane 2, arrow[8]).

Under iron-depleted growth conditions the OMP profile is altered substantially. The presence of four additional HMW proteins was the most obvious change (Fig. 4.17A, lane 4, arrow[9]), while several other proteins were also present at ~47 kDa (Fig. 4.17A, lane 4, arrow[10]), ~56 kDa (Fig. 4.17A, lane 4, arrow[11]) and ~60 kDa (Fig. 4.17A, lanes 4 and 5, arrow[12]). The OMP profile from cells grown in seawater medium was substantially different from the other growth conditions (Fig. 4.17A, lane 5). Proteins of ~21 kDa (Fig. 4.17A, lane 5, arrow[13]), ~26 kDa (Fig. 4.17A, lane 5, arrow[14]), ~28 kDa (Fig. 4.17A, lane 5, arrow[15]), ~32 kDa (Fig. 4.17A, lane 5, arrow[16]) and several LMW proteins were present in this profile Fig. 4.17A, lane 5, arrow[17]), while proteins of ~23 kDa (Fig. 4.17A, lane 5, arrow[18]), ~34 kDa (Fig. 4.17A, lane 5, arrow[19]), ~44 kDa (Fig. 4.17A, lane 5, arrow[20]), ~49 kDa (Fig. 4.17A, lane 5, arrow[21]), ~62 kDa (Fig. 4.17A, lane 5, arrow[22]) and ~88 kDa (Fig. 4.17A, lane 5, arrow[23]) were now absent.

4.3.9.6 Isolate RD366

Isolate RD366 is an OMP-type 2c, serotype O5 isolate, recovered from Atlantic salmon in the UK. This isolate is again representative of a large number of isolates in our collection, and has been responsible for a large number of infections in Atlantic salmon over the previous 14 years.

Under aerobic growth (22°C) (Fig. 4.17B, lane 1), the major proteins OmpA (Fig. 4.17B, lane 1, arrow[1]) and OmpF (Fig. 4.17B, lane 1, arrow[2]) are abundant. The protein OmpC is upregulated under anaerobic, aerobic (28°C), iron-depleted growth and growth in seawater medium (Fig. 4.17B, lane 2, 3, 4 and 5, arrow[3]). A protein of ~43 kDa (Fig. 4.17B, lane 1, arrow4) is present only under aerobic (22°C) growth conditions. Under anaerobic, iron-depleted growth and growth in seawater medium (Fig. 4.17B, lanes 2, 4 and 5, respectively), a protein of ~23 kDa is absent (Fig. 4.17B, lanes 2, 4 and 5, arrow[5]).

Under iron-depleted growth conditions four additional HMW proteins are up-regulated (Fig. 4.17B, lane 4, arrow[6]), while other proteins were present at ~47 kDa (Fig. 4.17B, lane 4, arrow[7]) and ~56 kDa (Fig. 4.17B, lane 4, arrow[8]). Under growth in seawater medium (Fig. 4.17B, lane 5) several LMW proteins were present (Fig. 4.17B, lane 5, arrow[12]), while proteins of ~26 kDa (Fig. 4.17B, lane 5, arrow[10]), ~41 kDa (Fig. 4.17B, lane 5, arrow[11]) and four additional HMW proteins (Fig. 4.17B, lane 5, arrow[9]) were now present.

4.3.9.7 Isolate RD382

Isolate RD382 is a biotype 1, serotype O1, OMP-type 3a isolate recovered from Atlantic salmon in the UK. This isolate has not been recovered extensively in the UK Atlantic salmon populations. However, it is a typical 'Hagerman' type isolate and is included for comparison to the major disease causing clonal group in rainbow trout represented by isolate RD124.

The major proteins OmpA (Fig. 4.17C, lane 1, arrow[1]) and OmpF (Fig. 4.17C, lane 1, arrow[2]) are clearly visible and abundant under aerobic (22°C). However, it is difficult to distinguish OmpC under any growth condition, due to the similarity in MW between that and OmpA. OmpF is downregulated under both iron-depleted growth and growth in seawater medium (Fig. 4.17C, lanes 4 and 5, arrow[3]).

A protein of ~41 kDa is present in abundance under aerobic (22°C) growth, and less so under anaerobic growth conditions (Fig. 4.17C, lanes 1 and 2, arrow[4]). Under anaerobic, iron-depleted growth and growth in seawater medium (Fig. 4.17C, lanes 2, 4 and 5, respectively), a protein of ~23 kDa is absent (Fig. 4.17C, lanes 2, 4 and 5, arrow[5]).

The most noticeable changes in OMP profile occur during growth in iron-deplete medium and in seawater medium. Under iron-depleted growth conditions four additional HMW proteins were upregulated (Fig. 4.17C, lane 4, arrow[6]), while several other proteins were also present at ~41 kDa (Fig. 4.17C, lane 4, arrow[7]) and ~47 kDa (Fig. 4.17C, lane 4, arrow[8]). In artificial seawater, the four HMW proteins upregulated under iron-deplete conditions are also slightly up regulated (Fig. 4.17C, lane 5, arrow[9]). A protein of ~24 kDa (Fig. 4.17C, lane 5, arrow[10]) is

down-regulated, while a ~44 kDa (Fig. 4.17C, lane 5, arrow[11]) and several LMW proteins are up-regulated (Fig. 4.17C, lane 5, arrow[12]).

4.3.9.8 Isolate RD420

Isolate RD420 is a biotype 1, serotype O1/O5, OMP-type 3a isolate recovered from Atlantic salmon in the UK. This isolate is now the most extensively recovered isolate in UK Atlantic salmon populations, representing the major clonal group infecting UK Atlantic salmon.

The major proteins OmpA (Fig. 4.17D, lane 1, arrow[1]) and OmpF (Fig. 4.17D, lane 1, arrow[2]) are clearly visible and abundant under aerobic (22°C) conditions. However, it is difficult to distinguish OmpC under any growth condition, due to the similarity in MW between that and OmpA. OmpF is downregulated under both iron-depleted growth and growth in seawater medium (Fig. 4.17D, lanes 4 and 5, arrow[3]).

A protein of ~41 kDa is present in abundance under aerobic (22°C) growth, and less so under anaerobic growth conditions (Fig. 4.17D, lanes 1 and 2, arrow[4]). Under anaerobic conditions, iron-depleted growth and growth in seawater medium (Fig. 4.17D, lanes 2, 4 and 5, respectively), a protein of 23 kDa is absent (Fig. 4.17D, lanes 2, 4 and 5, arrow[5]).

The most noticeable changes in OMP profile occur during growth in iron-depleted medium and in seawater medium. Under iron-depleted growth conditions four additional HMW proteins were upregulated (Fig. 4.17D, lane 4, arrow[6]), while several other proteins were also present at ~41 kDa (Fig. 4.17D, lane 4, arrow[7]) and ~47 kDa (Fig. 4.17D, lane 4, arrow[8]). In seawater medium the four HMW proteins upregulated under iron-deplete conditions are also slightly up regulated (Fig. 4.17D, lane 5, arrow[9]). A protein of ~24 kDa (Fig. 4.17D, lane 5, arrow[10]) is down-regulated, while a ~44 kDa (Fig. 4.17D, lane 5, arrow[11]) and several LMW proteins are up-regulated (Fig. 4.17D, lane 5, arrow[12]).

4.3.10 Gel-free proteomic analysis

Due to the extensive work involved in analysing the OM proteome of eight isolates grown under five distinct growth conditions (in triplicate), gel-based proteomic approaches were unfeasible. Therefore, gel-free technologies were used to examine variation in OM proteome composition under conditions which mimicked the *in vivo* and *in vitro* environments that *Y. ruckeri* may face. Preparing and analysing triplicate samples allowed examination of the reproducibility of the technique, while increasing the number of proteins that we were able to identify.

The number of proteins identified in each isolate under each growth condition was quite variable (Table 4.1). In total, 73 unique proteins were identified in all isolates and growth conditions (Table 4.2). When considering the core proteome, only three proteins were identified in all replicates of all isolates under every growth condition. These were OmpA, Braun's lipoprotein (Lpp) and Pal (peptidoglycan associated lipoprotein). These proteins are all of crucial importance to the bacterial cell as they play vital roles in maintenance of cell structure and viability. When the column representing growth in seawater medium is excluded from the table the core proteome increases to seven proteins now including several members of the BAM complex (BamA, BamB and BamC) and SlyB. These proteins are all important in OM biogenesis. If the core proteome is considered to be any protein that is identified in an isolate under all growth conditions (irrespective of the number of replicates the protein was identified in), the core proteome increases to four, as the porin OmpF is included. Excluding growth in seawater medium from the equation results in 13 proteins identified as the core proteome. These proteins include BamD, MetQ, OsmY, LptD and LptE.

Table 4.1 The total number of proteins identified through gel-free proteomic analysis in representative isolates of *Y. ruckeri* grown under variant growth conditions.

Condition	Rainbow trout				Atlantic salmon			
	RD6	RD28	RD64	RD124	RD354	RD366	RD382	RD420
Aerobic (22°C)	40	36	38	37	43	35	33	38
Anaerobic	31	36	31	31	36	28	26	28
Aerobic (28°C)	39	34	40	38	38	37	34	36
Iron-depleted	39	40	40	44	41	36	39	36
Seawater medium	17	33	25	34	30	36	14	15
Average:	**33**	**36**	**35**	**37**	**38**	**34**	**29**	**31**

Table 4.2 OMPs identified in representative isolates of *Y. ruckeri* under conditions mimicking *in vivo* and environmental growth by gel-free proteomics.

Proteins identified in three replicates are indicated by a red 3, in two replicates by an orange 2, and in one replicate by a yellow 1. Proteins not identified in the isolate in any replicate are indicated by a dash.

	Accession	Protein	MW (Da)	RD6 Replicates					RD28 Replicates					RD64 Replicates					RD124 Replicates					RD354 Replicates					RD366 Replicates					RD382 Replicates					RD420 Replicates				
				1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1. OM Biognesis and integrity																																											
1	gi|238708193	Brauns (Lpp)	6740	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2	gi|238707374	OmpA	37846	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3	gi|238705740	Pal	18149	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	gi|238708170	SlyB	15382	3	3	3	3	3	3	3	3	3	1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	3	3	3	3	-
5	gi|238706031	LptD	89769	3	3	3	2	1	2	2	3	3	1	3	3	3	3	1	1	3	3	2	2	3	3	3	3	1	1	1	3	2	1	2	1	3	3	-	1	2	3	2	-
6	gi|238706768	LptE	21133	3	3	3	3	-	2	3	3	2	-	2	3	3	3	-	3	3	3	3	2	3	3	3	3	1	1	2	3	2	1	2	2	2	1	-	1	2	3	2	-
7	gi|238706888	Slp Family	20689	3	3	3	3	-	3	1	3	3	1	3	3	3	3	-	3	2	3	3	2	3	3	3	3	-	2	2	3	3	1	2	1	3	3	-	2	2	3	3	-
8	gi|238706959	BamA/YaeT	88022	3	3	3	3	-	3	3	3	3	2	3	3	3	3	1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	-	3	3	3	3	1
9	gi|238707860	BamB	42358	3	3	3	3	-	3	3	3	3	1	3	3	3	3	-	3	3	3	3	3	3	3	3	3	2	3	3	3	3	2	3	3	3	3	-	3	3	3	3	3
10	gi|238705699	BamC	38345	3	3	3	3	-	3	3	3	3	1	3	3	3	3	1	3	3	3	3	3	3	3	3	3	2	3	3	3	3	2	3	3	3	3	-	3	3	3	3	-
11	gi|238705321	BamD	27432	3	3	3	3	-	3	2	3	3	1	3	3	3	3	1	3	3	3	3	3	3	3	3	3	2	3	2	3	3	2	2	2	3	3	-	3	3	3	3	-
12	gi|238707927	BamE	9943	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-
13	gi|238708048	VacJ	28773	3	3	-	3	-	3	3	-	3	1	3	3	-	3	1	3	3	-	3	3	3	3	-	3	2	3	3	-	3	2	3	3	-	3	-	3	3	-	3	-
14	gi|238706986	MltA	42895	3	1	-	-	-	1	-	1	-	-	3	-	1	2	-	2	-	1	-	2	3	1	1	2	-	1	-	1	-	1	1	-	-	-	-	-	-	1	-	-
15	gi|238707253	MltC	39935	-	-	-	-	-	-	-	-	-	-	3	-	1	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16	gi|238706640	MltD	48282	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2. Transport and receptor																																											
17	gi|23870735	OmpF	40024	3	3	3	3	1	2	3	3	3	1	1	3	3	3	3	2	3	3	3	2	2	3	3	3	1	2	2	3	2	1	2	3	3	2	3	2	3	3	3	2
18	gi|23870632	TolC	53048	1	3	3	3	2	2	3	3	3	2	3	3	3	3	3	3	3	3	2	3	2	3	3	3	3	2	3	3	3	3	1	3	3	3	3	-	3	3	3	2
19	gi|23870650	OmpC.1	39947	3	3	3	2	-	-	1	2	-	-	2	2	3	-	-	3	2	3	3	-	3	1	2	-	-	2	-	2	-	-	-	-	-	-	-	-	1	1	-	-
20	gi|23870794	OmpC.2	40459	3	3	3	3	2	-	2	3	1	3	-	3	3	2	3	-	3	3	3	3	-	3	3	2	1	-	3	3	3	3	-	2	3	3	3	-	3	3	1	3
21	gi|23870682	OsmY	11124	3	3	2	2	-	3	3	2	2	1	3	3	3	3	-	1	3	2	2	3	3	3	3	3	2	3	3	3	3	2	3	3	3	3	-	2	3	2	3	3
22	gi|23870693	MetQ	29356	2	2	3	3	1	3	3	3	3	3	1	2	2	3	2	3	2	3	3	3	1	1	1	2	2	2	2	3	3	3	1	3	3	3	-	3	2	3	3	-
23	gi|23870536	TamA/YtfN	63394	2	3	3	2	-	1	3	2	3	1	2	3	3	3	-	2	3	3	3	3	2	3	3	3	2	1	3	3	3	2	-	2	3	3	-	3	3	3	3	-
24	gi|23870634	ShuA	73465	2	-	3	3	3	3	-	3	3	3	3	-	3	3	2	2	-	3	3	3	3	-	1	3	3	2	-	2	3	2	2	-	3	3	3	3	-	3	3	2
25	gi|23870820	HemR	72964	1	-	-	3	2	-	-	-	3	2	-	-	-	3	3	-	-	-	3	3	-	-	-	3	3	-	-	-	3	2	-	-	-	3	3	-	-	-	3	1

Table 4.2 continued

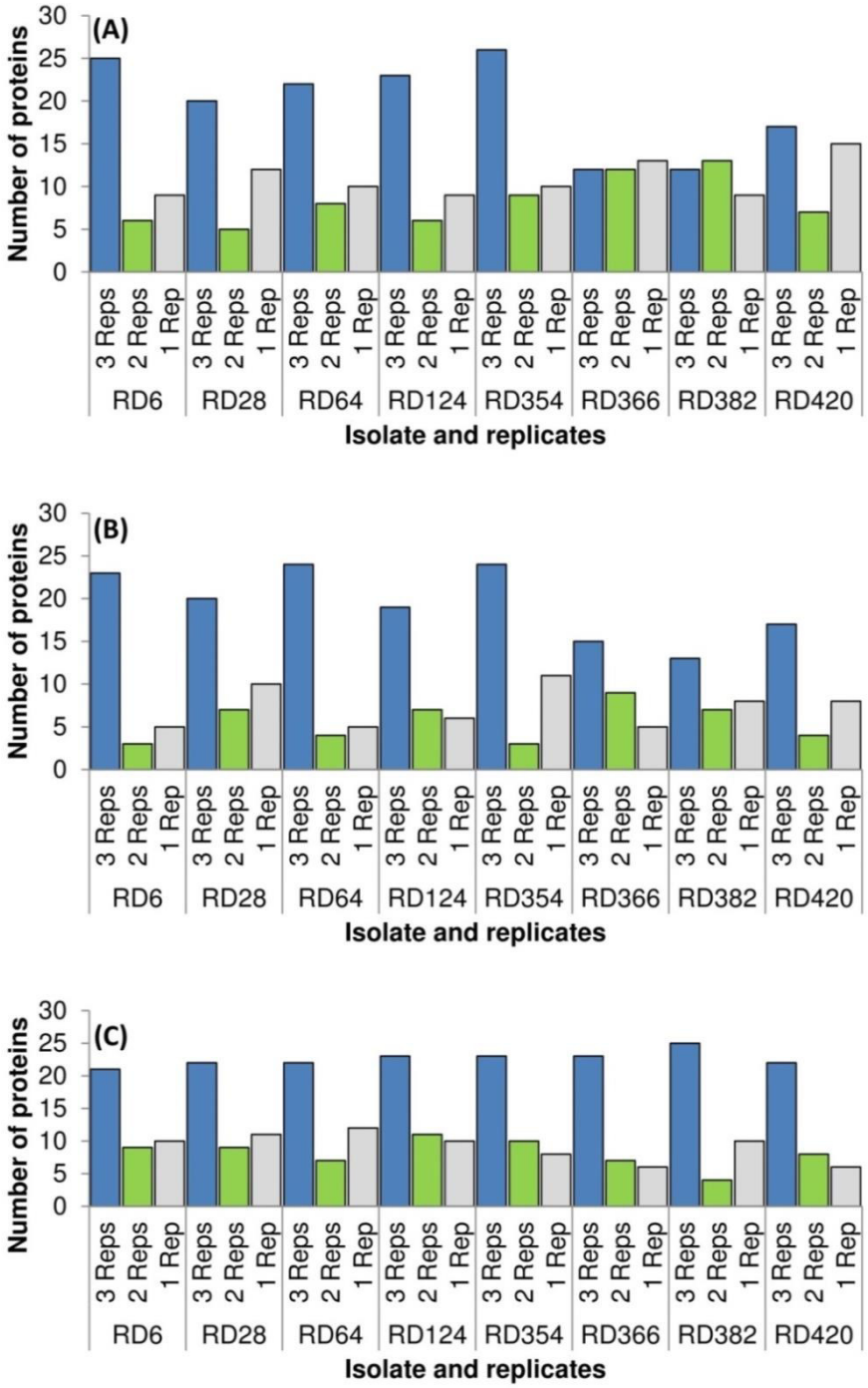
	Accession	Protein	MW (Da)	RD6					RD28					RD64					RD124					RD354					RD366					RD382					RD420				
				Replicates					Replicates					Replicates					Replicates					Replicates					Replicates					Replicates					Replicates				
				1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
26	gi|23870585	FhuA	81724	1	-	1	3	1	-	-	-	3	2	-	-	-	3	3	1	-	1	3	3	1	-	1	3	3	2	-	-	3	3	2	-	1	3	3	2	-	-	3	1
27	gi|23870667	FiuA	77361	-	-	1	3	1	-	-	-	3	1	-	-	-	1	1	-	-	2	3	2	-	-	-	1	2	-	-	-	3	2	-	-	-	3	1	1	-	-	3	-
28	gi|23870828	Ton-B dependent rec.	76573	-	-	-	2	-	-	-	-	2	-	-	-	-	2	1	-	-	-	3	-	-	-	-	3	-	-	-	-	2	-	-	-	-	3	-	-	-	-	-	-
29	gi|23870702	HasR	89969	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
30	gi|23870802	OmpE	39925	3	3	1	-	-	-	1	1	1	-	-	-	2	1	-	-	2	2	2	1	-	1	1	2	-	-	-	3	-	-	-	-	-	1	-	-	2	-	-
31	gi|23870763	OprC	75132	2	-	-	2	-	1	-	-	2	1	1	-	2	2	2	-	-	-	2	3	3	-	-	2	2	-	-	1	3	2	1	-	1	2	-	1	-	1	3	-
32	gi|23870682	OmpW	23865	-	-	3	-	-	-	-	3	-	-	1	1	2	-	-	-	-	2	-	1	2	1	3	-	-	-	-	2	-	-	-	-	3	-	-	-	-	2	-	-
33	gi|23870700	BtuB	68957	1	-	1	-	-	-	-	-	-	-	-	-	2	-	-	1	1	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
34	gi|23870804	FadL	46035	-	1	1	-	-	-	-	-	-	-	-	-	3	1	-	-	-	2	-	-	1	3	3	1	-	-	1	2	-	-	-	-	1	-	-	-	-	-	-	-
3. Adherence																																											
35	gi|23870511	ShlA/FhaA/HecA fam.	155469	1	-	-	1	-	1	-	-	-	1	-	-	-	-	-	-	-	-	1	-	3	-	-	3	-	2	-	-	-	-	2	-	-	1	-	1	-	-	2	-
36	gi|23870737	PqiB/Mam7	59633	1	-	1	1	-	1	1	-	1	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	2	1	-	-	1	1	-	1	-	-	-	-
37	gi|23870556	AidA	44266	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
38	gi|23870562	Fim. Usher	87773	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4. Enzymatic activity																																											
39	gi|23870524	BcsC	126527	3	-	1	1	-	3	3	3	3	-	3	3	3	3	-	3	2	2	2	-	3	3	3	3	-	2	2	3	3	-	2	3	3	3	-	3	3	3	3	-
40	gi|23870555	ShlB/FhaC/HecB fam.	62594	2	2	1	1	-	-	-	1	-	-	-	-	2	1	-	1	-	1	1	-	3	-	1	2	-	-	-	-	-	-	1	-	-	1	-	1	-	1	-	-
41	gi|23870555	Hemolysin, partial	48963	3	-	-	-	-	2	-	2	1	-	2	-	2	-	-	2	1	3	-	-	2	3	3	-	-	-	-	1	-	-	2	1	1	-	-	1	1	2	2	-
42	gi|23870624	Hemolysin activator	61801	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-
43	gi|23870706	Pep M37	50544	-	-	-	-	-	1	-	-	-	1	3	1	-	-	-	2	-	-	-	-	2	-	-	-	-	1	-	-	-	-	-	-	-	-	-	1	-	-	-	-
44	gi|23870548	Phos. A1	32689	-	3	1	-	-	-	2	1	-	-	-	1	3	-	-	-	2	2	-	-	-	1	1	-	-	-	3	1	-	-	-	1	-	-	-	1	-	-	-	-
45	gi|23870597	Protease II	77905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-
46	gi|11000733	YhlB	62424	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-
5. Motility																																											
47	gi|49853247	Flagellin (FlaA)	43589	-	-	-	-	-	3	3	1	3	-	3	3	1	1	-	3	1	-	-	-	3	3	-	-	-	3	3	-	-	-	3	1	-	-	-	3	1	-	-	-
48	gi|23870646	FlgE	44254	-	-	-	-	-	1	3	-	2	-	2	3	-	-	-	3	-	-	-	-	1	2	-	-	-	1	2	-	-	-	-	-	-	-	-	2	-	-	-	-
49	gi|23870643	FliD	49054	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6. Other																																											
50	gi|23870776	YiaD	22348	3	3	3	3	2	3	3	3	2	1	3	3	3	3	-	3	3	3	3	3	3	1	3	3	1	-	-	3	2	2	3	2	3	3	3	3	3	3	2	-
51	gi|23870789	YfhG	35016	3	3	3	1	-	3	3	3	2	1	3	3	3	2	-	3	3	3	2	2	3	3	3	2	-	2	-	3	1	1	2	1	3	2	-	2	3	3	1	-

Table 4.2 continued

	Accession	Protein	MW (Da)	RD6					RD28					RD64					RD124					RD354					RD366					RD382					RD420				
				Replicates					Replicates					Replicates					Replicates					Replicates					Replicates					Replicates					Replicates				
				1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
52	gi|23870519	YraP	20221	-	2	3	-	1	1	2	1	2	1	2	3	1	2	2	3	-	2	2	2	1	2	2	2	2	2	1	2	1	2	2	-	-	2	-	3	1	1	2	2
53	gi|23870596	CsgG	24712	2	1	1	3	-	1	-	-	1	-	2	-	-	1	1	1	-	-	-	-	3	-	1	1	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	1
54	gi|23870717	SecD	71766	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1	-	1	-	-	-	-	-	-	-
55	gi|23870578	TTSS protein 2	61602	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	1	-	-	-	-	-	-	-
56	gi|23870578	TTSS protein 1	45918	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-
57	gi|23870585	AmiB	62746	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-
58	gi|23870563	Blc	20801	-	-	-	1	-	1	-	-	1	-	1	-	-	2	-	-	-	-	-	3	1	-	-	2	1	-	-	-	2	1	-	-	-	3	-	-	-	-	3	-
59	gi|23870595	Mce-Like	91347	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	1	-
60	gi|23870691	MipA	28023	1	1	-	2	-	-	1	-	1	1	-	-	1	1	3	-	-	2	1	2	-	1	2	1	2	-	-	2	1	2	-	-	-	-	2	-	-	1	-	-
61	gi|23870742	NagZ	99335	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	1	-	-	-	-	-	-	1	-	-	-	-	2	-
62	gi|23870735	NlpD	34028	1	1	3	-	-	-	2	2	-	1	-	2	3	1	1	1	1	3	2	3	1	3	3	-	1	2	2	2	-	2	1	-	3	1	-	1	-	2	1	-
63	gi|23870694	NlpE/CutF	25245	-	3	1	1	-	-	2	1	2	1	-	-	-	-	-	-	3	1	3	3	-	-	-	2	1	-	2	3	3	2	-	1	2	3	-	-	-	1	2	-
64	gi|23870833	OmpX	17086	-	-	-	-	-	-	1	-	-	-	-	1	1	-	2	-	1	-	-	-	-	1	3	-	-	-	-	-	-	1	-	-	-	-	-	-	-	1	-	-
65	gi|16388385	PtrA	107905	-	-	-	-	1	-	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
66	gi|23870676	RlpA	37591	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-
67	gi|46369605	RupA	17881	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-
68	gi|23870508	YeeJ	54634	3	-	-	2	-	-	1	-	-	-	-	-	-	-	-	3	-	-	2	-	2	-	1	-	-	1	-	-	1	1	1	-	-	-	-	1	1	1	-	-
7. Hypothetical																																											
69	gi|23870602	Hypothetical 17	17964	3	-	1	2	-	3	1	2	-	1	2	1	3	1	2	3	1	3	1	2	2	1	3	1	1	1	1	2	1	1	2	-	3	-	-	2	1	2	-	-
70	gi|23870525	Hypothetical 22	40804	-	-	-	-	-	-	1	-	-	-	-	-	1	1	-	-	-	2	-	-	-	-	1	-	-	-	-	1	-	-	-	2	1	-	-	-	-	2	-	-
71	gi|23870673	Hypothetical 5	31452	3	-	-	1	-	3	-	-	1	-	1	-	1	-	-	-	-	-	-	-	3	1	-	1	-	1	-	1	1	-	-	-	-	1	-	-	-	-	-	-
72	gi|23870618	Hypothetical 20	17544	-	-	1	-	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	1	-	-	2	1	-	-	-	-	-	-
73	gi|23870610	Hypothetical 21	106480	-	-	-	-	-	-	-	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
74	gi|23870740	Hypothetical 16	56778	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-
75	gi|23870744	Hypothetical 14	22794	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
76	gi|23870729	Hypothetical 18	54666	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Total:		40	31	39	39	17	36	36	34	40	33	38	31	40	40	25	37	31	38	44	34	43	36	38	41	30	35	28	37	36	36	33	26	34	39	14	38	28	36	36	15

Aerobic (22°C), anaerobic, aerobic (28°C), iron-depleted and seawater medium growth conditions are indicated by 1 – 5, respectively.

The reproducibility of the gel-free method is highlighted in Fig. 4.18, as the majority of proteins were identified in more than one replicate in each isolate under each growth condition. In most cases, the number of proteins identified under aerobic growth (Fig. 4.18A) was the highest. In isolates RD366 and RD382, the number of proteins identified in all three replicates under aerobic growth conditions was lower than in the remaining isolates. Thirteen proteins were identified in all replicates of both RD366 and in RD382, while in the remaining isolates the total number of proteins identified in all three replicates ranged from 16 to 25.



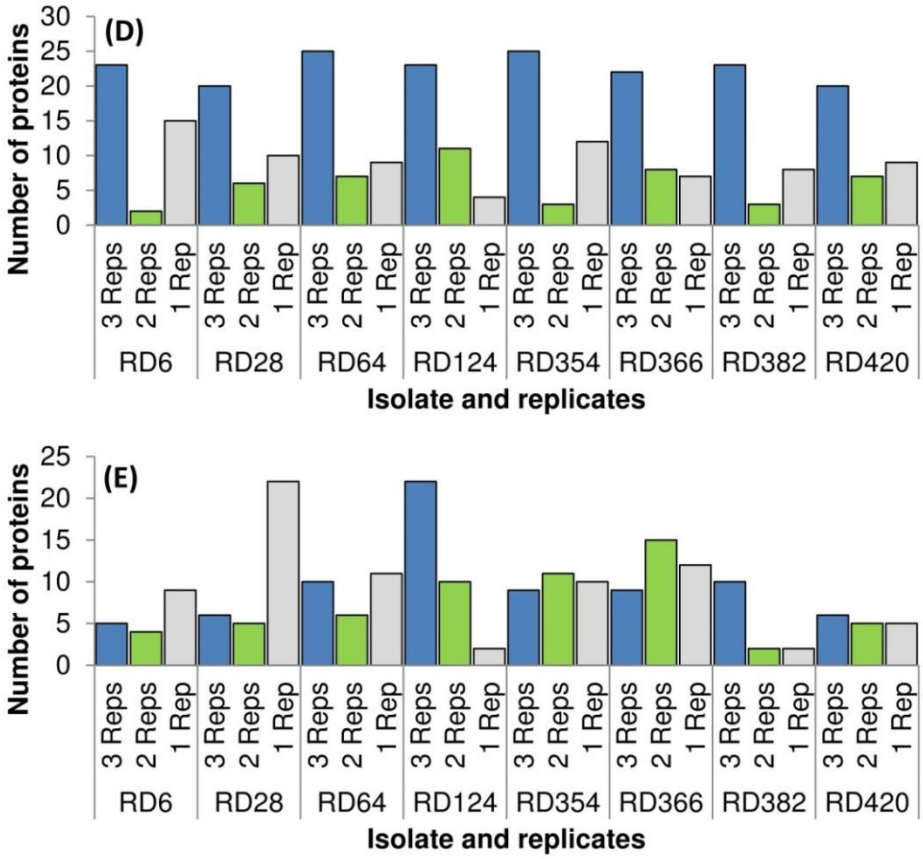


Fig. 4.18 The number of OMPs identified in three, two and one replicate through gel-free proteomic analysis of isolates of *Y. ruckeri* under different growth conditions. Isolates were grown aerobically **(A)**, anaerobically **(B)**, aerobically at 28°C **(C)**, under iron depletion **(D)** and in a seawater medium **(E)**. Isolates RD6, RD28, RD64 and RD124 were recovered from rainbow trout, while isolates RD354, RD366, RD382 and RD420 were recovered from Atlantic salmon. In most cases, the majority of proteins were identified in all three replicates under each growth condition. However, under growth in seawater medium the results were more varied.

When isolates were grown anaerobically (Fig. 4.18B), the total number of proteins identified was generally lower than that for aerobically grown isolates. The total proteins identified ranged from 30 in RD382 to 41 in both RD28 and RD354 with the OMPs identified in all three replicates the highest for all isolates. However, under aerobic (28°C) (Fig. 4.18C) and iron-depleted growth (Fig. 4.18D), the total number of OMPs identified was generally higher than under aerobic growth. For isolates grown aerobically at 28°C the number of OMPs identified ranged from 40 in RD420 to 51 in RD124 with the OMPs identified in all three replicates the highest for all isolates. In isolates grown under iron-depletion the number of OMPs identified ranged from 41 in RD28 to 46 in RD64 with the OMPs identified in all three replicates the highest for all isolates. The range of OMPs identified when isolates were grown in seawater medium (Fig. 4.18E) was much more diverse, with 16 proteins identified in isolate RD382 to 39

proteins identified in RD366. When examining the proteins that were identified in all three replicates, growth in seawater medium was again the most diverse. Only in isolates RD124 and RD420 were OMPs identified in all three replicates the most common. In the remaining isolates, OMPs were usually identified only in single replicates.

When cells were grown aerobically at 22°C, the core proteome included eight proteins. These were OmpA, Pal, SlyB, Braun's lipoprotein (Lpp), VacJ, BamA, BamB and BamC. Flagellin was also identified in all isolates except the non-motile isolate, RD6. Similarly, the protein YiaD (role) was identified in all isolates except RD6. When proteins present in any replicate of each isolate are considered, the core OM proteome increases by an additional 10 proteins. These include BamD, BcsC, ShuA, OsmY, Slp family protein, YfhG, hypothetical 17, OmpF, LptE and a partial haemolysin. Six proteins were unique to the Atlantic salmon strains examined. These included FiuA (one replicate; RD420), Phospholipase A1 (one replicate; RD420), RlpA (one replicate; RD420), Haemolysin activator (one replicate; RD382), MltD (one replicate; RD366) and RupA (one replicate; RD366). Eight proteins were unique to the rainbow trout strains examined. These included Hypothetical 21 (one replicate; RD64), Hypothetical 18 (one replicate; RD64), Hypothetical 20 (one replicate; RD28), OmpC.2 (three replicates; RD6), OmpE (three replicates; RD6), HemR (one replicate; RD6), TTSS protein 2 (one replicate; RD6) and MipA (one replicate; RD6).

When cells were grown anaerobically, the core proteome included ten proteins. These were OmpA, Pal, SlyB, Braun's lipoprotein (Lpp), VacJ, BamA, BamB, BamC, TolC and OsmY. When proteins present in any replicate of each isolate are considered, the core OM proteome increases by an additional eight proteins. These include OmpF, TamA/YtfM, OmpC.2, BamD, MetQ, LptE, Slp family protein and LptD. Flagellin was also identified in all isolates except the non-motile isolate, RD6. Similarly, the protein BcsC was identified in all isolates except RD6. The proteins YiaD and YfhG were identified in all isolates except RD366. Three proteins were unique to the Atlantic salmon strains examined. These included FiuA (one replicate; RD420), Phospholipase A1 (one replicate; RD420),

RlpA (one replicate; RD420), Haemolysin activator (one replicate; RD382), MltD (one replicate; RD366) and RupA (one replicate; RD366). Six proteins were unique to the rainbow trout strains examined. These included BtuB (one replicate; RD124), Peptidase M37 (one replicate; RD64), PqiB/Mam7 (one replicate; RD28), AidA (one replicate; RD28), ShlB/FhaC/HecB family protein (two replicates; RD6) and CsgG (one replicate; RD6).

When cells were grown aerobically at 28°C, the core proteome included 15 proteins. These were Braun's lipoprotein (Lpp), OmpA, Pal, SlyB, LptD, Slp family protein, BamA, BamB, BamC, BamD, OmpF, TolC, OmpC.2, YiaD and YfhG. When proteins present in any replicate of each isolate are considered, the core OM proteome increases by an additional nine proteins. These include TamA/YtfM, BcsC, MetQ, ShuA, LptE, OsmY, NlpD, Hypothetical 17 and OmpW. OmpC.1 was identified in all isolates except RD366. Similarly, the protein YraP was identified in all isolates except RD366. The partial hemolysin protein was not identified in isolate RD6. Three proteins were unique to the Atlantic salmon strains examined. These included SecD (one replicate; RD382), TTSS protein 2 (one replicate; RD382) and YeeJ (one replicate; RD354 and RD420). Five proteins were unique to the rainbow trout strains. These included BtuB (one replicate – RD6; two replicates – RD64 and RD124), FiuA (one replicate – RD6; two replicates – RD124), Flagellin (one replicate; RD28 and RD64), MltC (one replicate; RD64) and MltD (one replicate; RD6).

When cells were grown under iron-depletion, the core proteome included 13 proteins. These were Braun's lipoprotein (Lpp), OmpA, Pal, SlyB, Slp family protein, BamA, BamB, BamC, BamD, VacJ, ShuA, HemR and FhuA. When proteins present in any replicate of each isolate are considered, the core OM proteome increases by an additional 13 proteins. These include TamA/YtfM, TolC, BcsC, OsmY, MetQ, FiuA, OprC, OmpF, YiaD, LptD, LptE, OmpC.2 and YfhG. The proteins Blc, NlpE/CutF, YraP and a Ton-B dependent receptor were absent from strains RD124, RD64, RD6 and RD420, respectively. Three proteins were unique to the Atlantic salmon strains examined. These included Mce-like (one replicate; RD382 and RD420), protease II (one replicate; RD354 and RD420) and BamE (one replicate; RD420). Ten proteins were unique to the rainbow

trout strains. These included OmpC.1 (two replicates – RD6; three replicates – RD124), HasR (one replicate – RD6; one replicate – RD124), FliD (one replicate; RD124), RlpA (one replicate; RD124), Hypothetical 14 (one replicate; RD124), Flagellin (three replicates – RD28; one replicate – RD64), RupA (one replicate; RD64), Hypothetical 22 (one replicate; RD64), FlgE (two replicates; RD28) and Hypothetical 21 (one replicate; RD28).

When cells were grown in seawater medium the core proteome included three proteins. These were Braun's lipoprotein (Lpp), OmpA and Pal. When proteins present in any replicate of each isolate are considered, the core OM proteome increased by an additional six proteins. These include OmpC.2, TolC, ShuA, OmpF, FhuA and HemR. The protein SlyB was absent from isolate RD420. Four proteins were unique to the Atlantic salmon strains examined. These included PqiB/Mam7 (one replicate; RD366), Hypothetical 20 (one replicate; RD366), YeeJ (one replicate; RD366) and a Fimbral usher protein (one replicate; RD354). Seven proteins were unique to the rainbow trout strains. These included OmpW (one replicate – RD124), Ton-B dependent receptor (one replicate; RD64), PtrA (one replicate – RD6; two replicates – RD28), Peptidase M37 (one replicate; RD28), ShlA/FhaA/HecA family protein (one replicate; RD28), TTSS protein 1 (one replicate; RD28) and RlpA (one replicate; RD6).

It was difficult to identify patterns in the presence of OMPs under the different growth conditions. However, several proteins did indicate a degree of condition specificity. Proteins involved in OM biogenesis and maintenance of cell integrity are usually well conserved due to the important role that they play in the cells viability. As described previously (section 1.4.3.1), the BAM complex is involved in the assembly of ß-barrel proteins in the OM, and is therefore important for the bacterial cell as the majority of proteins at the OM are ß-barrel in structure. The proteins BamA, BamB, BamC and BamD were identified in nearly all replicates of each isolate grown aerobically at 22°C, anaerobically, aerobically at 28°C and under iron-deplete conditions. However, these proteins were rarely identified under growth in seawater medium. No members of the BAM complex were identified in isolates RD6 or RD382 when cells were grown in seawater medium. BamA Was identified in isolates RD28 (two replicates), RD64 (one

replicate) and RD420 (one replicate), although was identified in all three replicates of isolates RD124, RD354 and RD366. BamB Was identified in isolates RD28 (one replicate), RD354 (two replicates) and RD366 (two replicates), although was identified in all three replicates of isolates RD124 and RD420. BamB was not identified in isolate RD64 under growth in seawater medium. BamC Was identified in isolates RD28 (one replicate), RD64 (one replicate), RD354 (two replicates) and RD366 (two replicates), although was identified in all three replicates of isolate RD124. BamC was not identified in isolates RD382 or RD420 under growth in seawater medium. BamD Was identified in isolates RD28 (one replicate), RD64 (one replicate), RD354 (two replicates) and RD366 (two replicates), although was identified in all three replicates of isolate RD124. BamD was not identified in isolates RD382 or RD420 under growth in seawater medium. BamE was only identified in a single replicate of isolate RD354 grown aerobically (22°C) and a single replicate of isolate RD420 grown aerobically (28°C)

The OM lipoprotein VacJ, with roles in actively preventing phospholipid accumulation at the cell surface and maintaining lipid asymmetry in the OM, was identified in all replicates of all isolates grown aerobically (22°C), anaerobically, and aerobically at 28°C. However, it was downregulated under iron-depleted growth and in seawater medium, not identified in any isolate under iron-depleted growth or in isolates RD6, RD382 and RD420 grown in seawater medium. VacJ was identified in all replicates of isolate RD124 in artificial seawater, albeit in low abundance (Supplementary Table 8.1).

The Slp-family OM lipoprotein has structural responsibilities at the OM and was substantially downregulated under growth in seawater medium, present in one replicate of isolates RD28 and RD366 and in two replicates of isolate RD124, only.

The major porins OmpF and OmpC allow the transport of molecules in and out of the cell (discussed in detail in section 1.4.3.4.2). These proteins are under the control of the two component system EnvZ/OmpR; while OmpF expression is increased, OmpC expression is decreased, and vice versa. OmpF was identified in all isolates under all growth conditions. However, it was not identified in

every replicate. It was identified in every replicate of isolates grown under iron-depletion. When cells were grown in seawater medium, the presence of OmpF appeared greatly reduced as it was identified in only two replicates of isolates RD124 and RD420, and in single replicates of the isolates RD6, RD28, RD354 and RD366. Similarly, the emPAI number (Supplementary Table 8.1) was reduced under all conditions in comparison to standard aerobic growth. Two forms of OmpC, OmpC.1 and OmpC.2, were identified. Protein OmpC.2 was absent from all isolates under standard aerobic growth, apart from in isolate RD6 (where it was found in three replicates). OmpC.2 was identified under all growth conditions in the remaining isolates. However, OmpC.1 was identified less frequently. While it was present in all strains under aerobic growth (in all replicates), it was absent under growth in artificial seawater, and absent under all conditions in isolate RD382.

Under conditions of iron-depletion four additional HMW proteins were observed previously (Fig. 4.11). However, through gel-free analysis of the OM proteome five proteins were upregulated under iron-depleted growth and were identified in almost all replicates, of all isolates. These proteins included HemR, ShuA, a Ton-B dependent receptor, FiuA and FhuA and have predicted MWs of 73.5 kDa, 81.7 kDa, 73.0 kDa, 77.4 kDa and 76.6 kDa, respectively. The Ton-B dependent receptor was not identified in isolate RD420 under any growth condition. Upon comparison to the SDS-PAGE gels (Fig. 4.16 and Fig. 4.17, lane 4 of each panel), FiuA and HemR correspond to the MWs of the most abundant bands. These proteins were also observed in the SDS-PAGE profiles of isolates grown in artificial seawater in most cases (Fig. 4.16 and Fig. 4.17, lane 5 of each panel). However, while HemR, ShuA, FiuA and FhuA were identified through gel-free proteomics of cells grown in seawater medium, the Ton-B dependent receptor at 76.6 kDa was absent.

The protein BcsC, involved in cellulose biosynthesis, was present in all isolates under most growth conditions (this protein was absent from isolate RD6 under anaerobic growth). BcsC was not identified under growth in seawater medium in any isolate.

Autotransporters are important for bacterial adhesion to host cells (described in section 1.2.3.4). The TAM complex is important in the assembly of autotransporters into the bacterial OM. TamA (the OM component) was identified in most isolates, under all conditions except in seawater medium. TamA was not identified in isolate RD382 under aerobic (22°C) growth, however was identified under anaerobic, aerobic (28°C) and iron-depleted growth in the same isolate. However, autotransporter proteins were not identified in any isolate, under any condition. Other proteins involved in adherence including PqiB/Mam7 and a possible adhesin (gi|238705110) were identified although only in single replicates and no apparent pattern based on the growth conditions.

The OM lipoprotein CsgG (Loferer et al., 1997) is required for the secretion and stabilisation of two other proteins, CsgA and CsgB which form curli amyloid fibres on the extracellular surface. These fibres have been implicated in a number of processes including biofilm formation, attachment and invasion of host cells, interaction with host proteins and activation of the immune system (Barnhart and Chapman, 2006). The protein CsgG was identified by gel-free proteomics most abundantly in isolate RD6, being particularly expressed under iron-depleteD conditions.

The absence of flagellin (FlaA) and flagellar machinery in cells grown aerobically at 28 °C, under iron-depletion and in seawater medium is of high significance. Isolate RD6 is of biotype 2 and lacks flagella. No proteins with involvement in flagella production or flagellin itself were identified in this isolate under any growth condition. Under aerobic growth conditions, flagellin (FlaA) and FlgE were identified in the remaining seven isolates. Growth aerobically at 28°C, under iron-depletion and in seawater medium reduced flagella production, with observations being reduced generally to in single replicates of isolates. Flagellin was not identified in any isolate grown in seawater medium, and only in isolates RD28 and RD64 grown aerobically at 28°C and under iron-depletion.

4.3.11 Effect of iron depletion on motility

When grown aerobically at 28°C, under iron-depletion or in seawater medium, there was an absence of flagellin and flagellar apparatus proteins. As iron-

depletion mimics conditions inside the host, this indicates that the bacteria may be down-regulating flagella during infection.

The motility of isolate RD124 was examined under increasing concentrations of the iron-chelator 2,2'-dipyridyl. Using motility plates, the diameter of motility was measured over 24 and 48 h (Fig. 4.19). The non-motile isolate RD6 was included as a control.

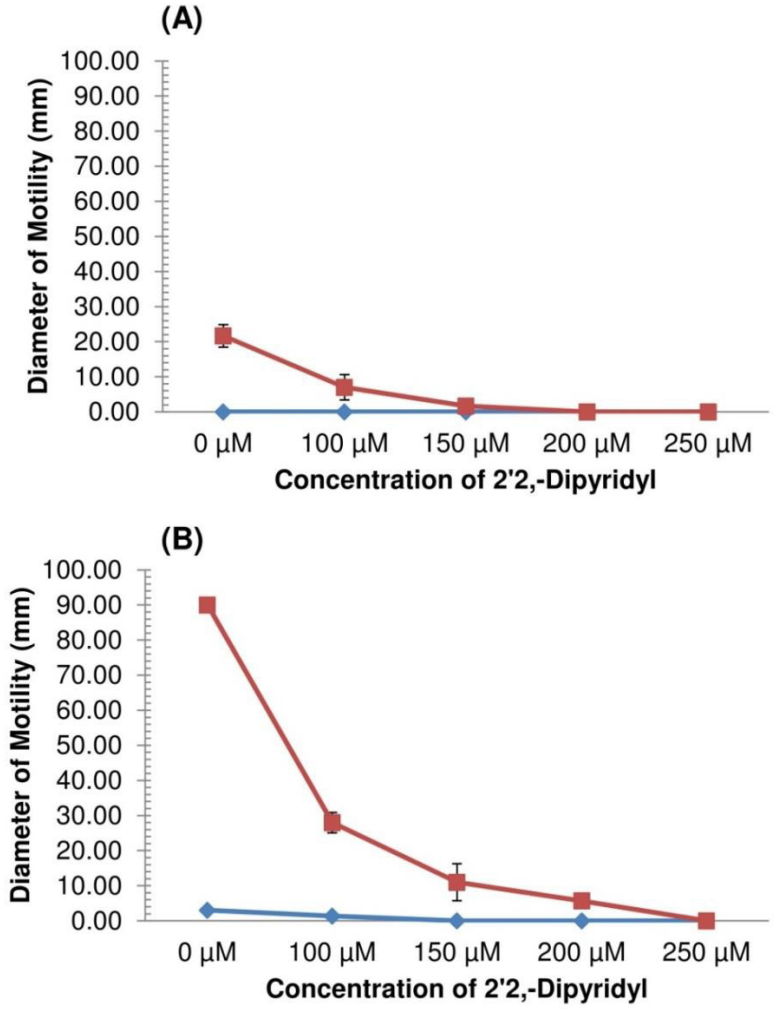


Fig. 4.19 Effect of 2,2'-dipyridyl on *Y. ruckeri* motility.
The rate of motility in motile isolate RD124 (red) and non-motile isolate RD6 (blue) were recorded after **(A)** 24 h and **(B)** 48 h. Isolates were grown under increasing concentrations of the iron chelator 2'2,-dipyridyl. Concentrations ranged from 0 µM to 250 µM.

As the concentration of 2,2'-dipyridyl increased from 0 µM to 250 µM, the motility of isolate RD124 decreased. At 0 µM 2,2'-dipyridyl, isolate RD124 spread to a diameter of ~22 mm at 24 h (Fig. 4.19A), and had completely filled the plate (90 mm diameter) by 48 h (Fig. 4.19B). As the concentration of 2,2'-dipyridyl increased to 100 µM, the diameter of motility decreased to 7 mm at 24

and 30 mm 48 h, respectively. At a concentration of 150 µM, the diameter was 2 mm at 24 h and 12 mm at 48 h. At 200 µM 2,2'-dipyridyl, there was no visible motility at 24 h, and less than 10 mm after 48 h. There was no motility whatsoever at 250 µM 2,2'-dipyridyl concentration. The decrease in motility as the concentration of 2,2'-dipyridyl increases is clearly seen in Fig. 4.20. Isolate RD6 did not dislay motility at any concentration of 2,2'-dipyridyl.

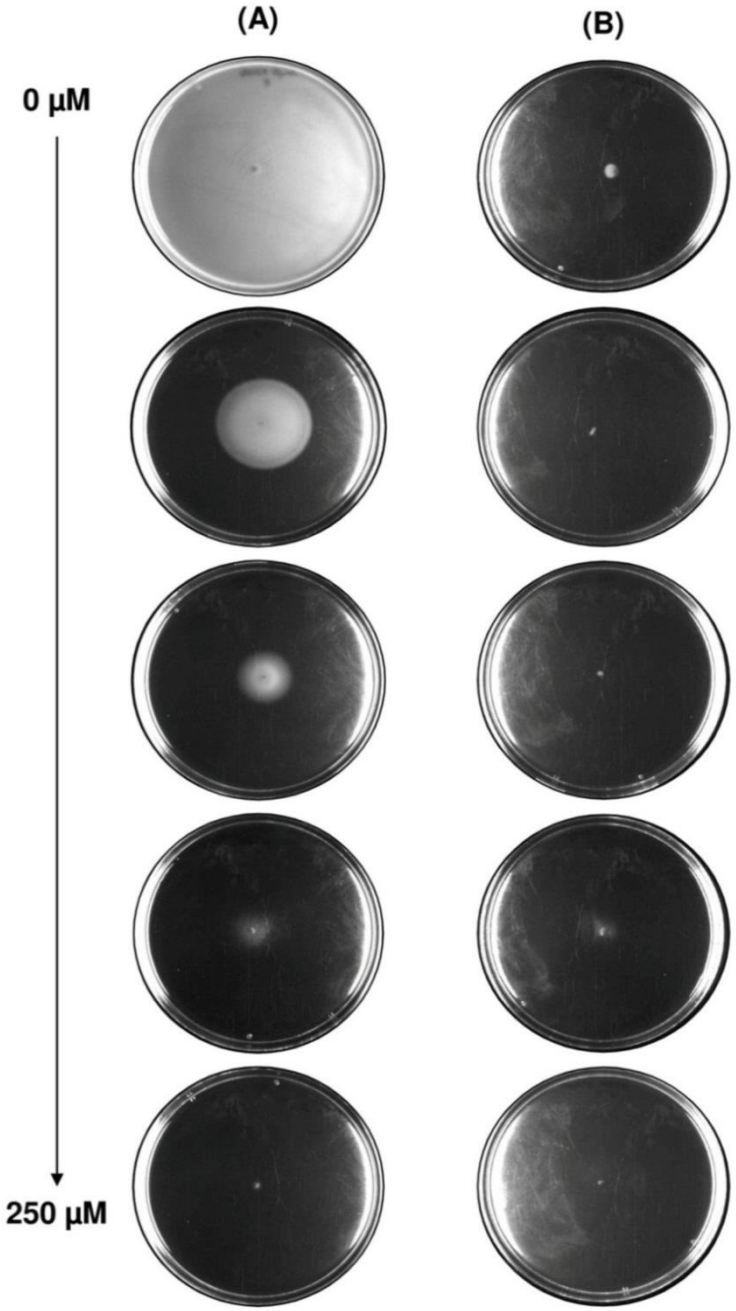


Fig. 4.20 The effect of 2,2'-dipyridyl on the motility of isolates of *Y. ruckeri*.
The motile isolate RD124 **(A)** and the non-motile isolate RD6 **(B)** were examined under increasing concentrations of the iron chelating agent, 2,2'-dipyridyl. As the concentration of 2,2'-dipyridyl increased from 0 µM to 250 µM (in 50 µM increments [50 µM itself was not examined]), the motility of isolate RD124 decreased. Non-motile isolate RD6 was not affected by the changing concentrations of 2,2'-dipyridyl. Images were acquired after 48 h.

4.4 Discussion

The OM proteome of *Y. ruckeri* was examined under situations which aimed to mimic the *in vivo* and environmental conditions the bacteria is likely to encounter inside and outside of the host. The rate of bacterial growth and alterations in the OM proteome were examined under different conditions of aeration, temperature, iron availability and salinity.

(a) Aeration. *Yersinia ruckeri* is a facultative anaerobe and is able to survive in both aerobic and anaerobic environments (Daligault *et al.*, 2014). Examining the rate of growth of isolates RD124 and RD158 under anaerobic conditions and low, moderate and high degrees of aeration showed that the rate of growth decreases as oxygen availability decreases (Fig. 4.3). Under standard conditions of aeration (22°C), there was very little difference in the rate at which the eight reference isolates grew (Fig. 4.12). All eight isolates grew at similar rates under anaerobic conditions (Fig. 4.13). However, the rate of growth under anaerobic conditions was inhibited in comparsion to aerobic growth. The bacteria were still able to proliferate in anaerobic and low oxygen environments – this mimics conditions inside the host, as oxygen is not freely available within the intestinal tract (Sengupta *et al.*, 2014) and is bound to haemoglobin in the blood (red blood cells; oxy-haemoglobin) where the bacteria causes septicaemia. Although the bacteria grew rapidly initially under anaerobic conditions, presumably using residual oxygen in the media, their rate of growth and maximum OD_{600nm} obtained declined over time (Fig. 4.14). This was most likely due to the bacteria switching to fermentation or anaerobic respiration for ATP production, which is less efficient than aerobic respiration (Shi *et al.*, 2005). The bacteria are most likely preserving energy rather than commiting to proliferation, as available oxygen and nutrients become depleted.

The OMP profiles of isolates grown under anaerobic conditions and low aeration were very similar. The OMP profiles of isolates grown under moderate and high degrees of aeration were also similar to each other (Fig. 4.3). The most noticeable change in OMP-profile caused by differences in aeration was the upregulation of a ~39.5 kDa protein, OmpC, in tandem with down regulation of

OmpF. This occured when the cells were grown under low aeration and anaerobic conditions (Fig. 4.3). This has been observed for other bacteria including *E. coli* (Batchelor & Walthers, 2005; Lan & Igo, 1998), *S. marcescens* (Begic & Worobec, 2006) and *A. hydrophila* (Yadav *et al.*, 2014), and is due to the differential pore size of these non-specific porins. Under stressful conditions (i.e. low aeration), the bacterial cell limits the loss of nutrients from the cell by expressing a pore with a smaller channel (OmpC).

Through gel-free proteomic analysis of isolates grown aerobically or under anaerobic conditions, differences in OMP expression were apparent. The proteins OmpC.1 and OmpC.2 are of predicted MWs 39.9 kDa and 40.1 kDa, respectively (in strain ATCC29473). These have been discussed previously (3.4.4). OmpC.2 was not expressed under aerobic growth conditions (with the exception of isolate RD6) but was abundant during all other growth conditions, including cells grown anaerobically. OmpC.1 was not identified during aerobic growth in isolates RD28, RD382 or RD420 but was identified in all three replicates of isolates RD6, RD124, RD354 and RD366. Having two forms of the same protein can be advantageous to bacteria. Having two copies that are transcribed together would enable the bacteria to quickly express this protein when needed. However, as these proteins do not appear to be expressed simultaneously this suggests that these proteins have different promoters. Having a second copy of a protein ensures that an unchanged copy of the original protein remains. Variations and mutations appear in the second copy, conserving the structure and function of the first. In *P. multocida*, three variants of the major OMP OmpH have been described; OmpH1, OmpH2 and OmpH3. While OmpH1 and OmpH2 are present in all strains, OmpH3 is not. With respect to amino acid sequences, OmpH2 is also seen to be more highly conserved than OmpH1 (Johnson *et al.*, 2013).

A ~73.5 kDa OM siderophore receptor was identified in all isolates under all growth conditions, except anaerobic growth. Siderophores are high affinity iron compounds that sequester iron from the host (discussed in detail in section 1.4.3.4.2). It is peculiar that this protein should be expressed constitutively under most growth conditions, while the other proteins involved in iron transport

that were identified (FiuA, HemR, FhuA and a Ton-B dependent receptor) were most abundant during iron-depleted growth, and were not identified under anaerobic growth. It is possible that this protein is expressed in order to identify when iron is available in the environment. For the bacteria to produce four additional proteins for iron transport when they are not required would be of huge energy expenditure and detrimental to the bacteria. However, this is unlikely. Described in *E. coli* initially, and in since in many other bacteria, the regulation of proteins involved in iron-transport is mediated by the ferric-uptake regulator protein (Fur). Fur is responsible for the iron-dependent expression of more than 90 proteins in *E. coli*, acting as a repressor, repressing transcription upon interaction with its co-repressor, Fe^{2+}, and causing de-repression in its absence (Andrews *et al.*, 2003; Hantke, 2001). The main function of Fur is to repress the genes encoding proteins involved in iron acquisition when iron is plentiful. This allows the iron transport systems to be induced by iron-restriction. However, Fur also regulates the expression of genes that are not involved in iron transport. These include genes with roles in respiration, flagella chemotaxis, the TCA cycle, glycolysis, methionine biosynthesis, phage-DNA packaging, DNA synthesis, purine metabolism and redox stress resistance (Andrews *et al.*, 2003). This makes Fur a global regulator and it is unlikely that one gene involved in iron uptake should be expressed while others are not. Ferrous iron (Fe^{2+}) is the more abundant form of iron in anaerobic environments, meaning that Fur transcription would be restricted. However, a siderophore secretion system that is expressed independently of Fur has been described in *B. subtilis* (Miethke *et al.*, 2008). It is therefore possible that *Y. ruckeri* may produce a siderophore receptor system in an anaerobic environment.

(b) Temperature. The bacteria were able to grow at 8, 15, 22, 28 and 37°C (Fig. 4.5). However, at 45°C *Y. ruckeri* was not. The majority of characterisation studies have been conducted at 22°C (Davies & Frerichs, 1989; Davies, 1990, 1991a, b, c, d; Verner-Jeffreys *et al.*, 2011; Wheeler *et al.*, 2009). However, growth was most rapid at 28°C for both RD124 (Fig. 4.5A) and RD158 (Fig. 4.5B). This suggests that 28°C is the 'optimal' growth temperature for the bacterium, and has been used in several studies (Fernández *et al.*, 2004, 2007a). Isolates were able to grow at both 8°C and 15°C however; they grew much more

slowly than at 22, 28 or 37°C. Both 8 and 15°C are more comparable with those inside the host.

Alterations in the OMP profiles were apparent between cells grown to the lower temperatures of 8, 15 and 22°C and the higher temperatures of 28 and 37°C (Fig. 4.6). This suggests that although isolates will grow faster at higher temperatures, this is not necessarily representative of the OM proteome during infection and should perhaps not be used as the 'optimum' temperature for studies of the bacterium.

When isolates were grown at 8°C (Fig. 4.6, Lanes 1 and 6), 15°C (Fig. 4.6, Lanes 2 and 7) and 22°C (Fig. 4.6, Lanes 3 and 8) there were no apparent differences in the OM profiles from each other. This suggests that OMP characterisation at 22°C is representative of the environmental and host temperature. At 28°C (Fig. 4.6) there was a marked decrease in expression of a protein; presumed OmpF. There was up-regulation of a 39.5 kDa protein; presumed OmpC.

Temperature-dependent expression of virulence genes has been reported in numerous bacterial species, including several important fish pathogenic species (Table 4.3). While the molecular mechanisms governing the expression of temperature-regulated virulence factors in mammalian pathogenic bacteria have been well studied, little is known about the temperature-regulated virulence factors in fish-associated bacteria. A key environmental stressor involved in outbreaks in fish farms is water temperature. A drop in water temperature has resulted in disease outbreaks, including cold water vibriosis (Enger et al, 1991), cold water disease (Cipriano and Holt, 2005) and ERM (Fernandez et al 2004). A rise in water temperature has been reported to be responsible for outbreak of disease in other cases, for example Lactococcosis (Venderall et al 2006), haemorrhagic septicaemia (Austin and Austin 2007), ERM (Busch, 1978; Rodgers, 1991) and Edwardsiellosis (Mohanty and Sahoo, 2007). In most cases, however, the majority of outbreaks occur at temperatures sub-optimal for the growth of the bacterial species involved.

Table 4.3 Examples of fish pathogenic bacteria affected by temperature changes. Table adapted from (Guijarro *et al.*, 2015).

Bacterial species	Disease caused	Optimal bacterial growth temperature	Optimal infection temperature
Y. ruckeri	ERM	28°C	Approx. 18°C
F. psychrophilum	Cold water disease/RTFS	20°C	<14°C
Lactococcus garvieae	Lactococcosis	18-37°C	18-37°C
A. hydrophila	Haemorrhagic septicaemia	37°C	20/25°C
E. tarda	Edwardsiellosis	37°C	25°C
V. salmonicida	Cold water vibrosis	15°C	<10°C

In isolates grown aerobically at 28°C, there was upregulation of the major protein OmpW (Table 4.2). OmpW belongs to the same family of proteins as OmpA, OmpX and other eight stranded ß-barrel proteins (Albrecht et al., 2006 and Hong et al., 2006). Previous studies have demonstrated that OmpW may be involved in the protection of bacteria against various forms of environmental stress, including osmosis (Xu *et al.*, 2005), oxidation (Gil *et al.*, 2007), temperature and the unavailability of nutrients and oxygen (Nandi *et al.*, 2005). In addition, proteomic methods demonstrate some correlation between the presence of OmpW in cells and bacterial resistance to antibiotics, including ampicillin, tetracycline and ceftriaxone (Goel & Jiang, 2010). The upregulation of OmpW is also correlated with an increase in bacterial virulence (Goel & Jiang, 2010). The expression of the OmpW protein of *E. coli* K-12 is modified by temperature changes compatible with moving from the environment to warm-blooded hosts. While OmpW is present in the OM of *E. coli* grown at 37°C, it disappears at 23°C (Brambilla *et al.*, 2014). In *V. cholerae*, the expression of OmpW is governed by environmental conditions. Expression is enhanced under nutrient-limiting conditions, elevated temperature, high salt concentration and low aeration (Nandi *et al.*, 2005). The role of OmpW in biofilm formation has been observed Upon examination of the OM of the marine bacterium *Pseudoalteromonas* sp. strain D41, four OMPs were strongly induced within biofilms; a TonB-dependent receptor, OmpW, OmpA and PilF (Ritter *et al.*, 2012). Furthermore, *P. aeruginosa* mutants unable to produce any of the OmpW, OmpA, and PilF homologues yielded smaller, weaker biofilms, confirming the involvement of these proteins in biofilm formation (Ritter *et al.*, 2012). This

suggests that at higher temperatures, isolates of *Y. ruckeri* may be up-regulating OmpW in order to produce a biofilm.

The OM lipoprotein VacJ has important roles in preventing phospholipid accumulation at the cell surface and maintaining lipid asymmetry in the OM. The role of VacJ in intracellular spreading has been highlighted among *Shigella* and enteroinvasive *E. coli*. VacJ mutants were unable to spread intracellularly, indicating that *vacJ* is essential for the pathogenicity of *Shigella*. This protein appeared to be downregulated under both growth aerobically at 28°C and in seawater medium. VacJ could therefore have a role in the intracellular spread of *Y. ruckeri* also. *Yersinia ruckeri* is not generally regarded as an intracellular pathogen, instead residing in the gut before causing a septicaemia. However, this protein may have a secondary function within the host. Growth aerobically at 28°C and in seawater medium are conditions not found inside the host, suggesting that the absence of this protein under these conditions could be an adaptation to survival outside the host. Further analysis of this through creating a *vacJ* knockout mutant and examining *Y. ruckeri* virulence through challenge experiments would be necessary.

The protein YfhG was identified as part of the OM core proteome in cells grown at 28°C. The function of YfhG is not known, meaning further characterisation work is necessary to uncover the role of this protein, and potential reasons for its upregulation at higher temperatures in *Y. ruckeri*.

(c) Iron-depletion. The rate of growth was extensively altered under iron-limited conditions. All isolates grew much slower in the presence of 2'2,-dipridyl (Fig. 4.10). This was to be expected, as iron is known to be essential for the growth of the bacteria (1.4.3.4.2) and limiting its availability within the media will hinder the rate of growth.

The upregulation of proteins with roles in iron acquisition under this condition was also expected, as the bacteria search for any available iron in order to continue their growth. Proteins involved in iron transport are transcribed simultaneously, as part of the Fur regulon (described previously). Davies (1991b)

described the presence of four additional HMW proteins of 72 kDa, 69.5 kDa, 68 kDa and 66 kDa expressed under iron-limited growth conditions. While SDS-PAGE analysis of the eight reference isolates yielded four additional proteins (Fig. 4.16 and Fig. 4.17), the MWs were slightly higher at 81.7 kDa, 77.4 kDa, 76.6 kDa, 73.0 kDa. However, these are highly likely to be the same proteins. Estimation of protein MW from SDS-PAGE gels is not exact.

However, through gel-free proteomic analysis five proteins of HMW with roles in iron acquistion were upregulated. As previously no identity had been assigned to the individual proteins present on SDS-PAGE gels, it is difficult to distinguish which protein is likely to have been unidentified. However, as the MWs of the proteins are close to one another, it is possible that the bands may overlap. The five proteins identified included HemR, ShuA, a Ton-B dependent receptor, FiuA and FhuA and have MWs of 72.9 kDa, 73.5 kDa, 76.6 kDa, 77.4 kDa and 81.7 kDa, respectively. However, the Ton-B dependent receptor was not identified in isolate RD420 (under any growth condition). Upon comparison to the SDS-PAGE gels (Fig. 4.16 and Fig. 4.17, lane 4 of each panel), FiuA and HemR correspond to the MWs of the most abundant bands.

The down-regulation of the flagellar proteins flagellin and FlgE (the OM components of the flagellar apparatus) under iron-depleted conditions was particularly striking. Members of the *Yersinia* have been shown to switch to a non-motile state upon gaining entry to the host (*Y. enterocolitica* and *Y. pseudotuberculosis*), while *Y. pestis* no longer expresses flagella at all inside or outside of the host (Minnich & Rohde, 2007). The emergence of non-motile biotype 2 isolates in *Y. ruckeri* has been discussed previously (section 2.4). In order to confirm that the absence of flagellar proteins in motile isolates grown under iron-depleted conditions was related to the growth conditions, the motility of isolate RD124 was examined in an environment with decreasing iron availability. As the concentration of the iron chelator 2,2'-dipyridyl increased from 0 μM to 250 μM, the motility of isolate RD124 was reduced, thus confirming that in an iron-depleted environment motile *Y. ruckeri* isolates do not produce flagella. *Yersinia ruckeri* has primarily been a disease of rainbow trout, and has been associated with this species for a much longer period of time than with

Atlantic salmon. It is therefore highly likely that isolates have evolved without flagella, due to flagella being a liability within the host (Evenhuis *et al.*, 2009; Scott *et al.*, 2013). Alternatively, the vaccination pressure evoked by using whole cell vaccines containing flagella have prepared the host for entry of motile isolates, meaning the non-motile phenotype may have emerged in rainbow trout as a vaccine avoidance strategy. Given enough time, the main clonal groups circulating with the Atlantic salmon populations may revert to a non-motile phenotype, also. This has been discussed in more detail, previously (section 2.4).

Iron availability in the ocean has been shown to affect the growth and production of phytoplankton and free-living bacteria. Tang and Grossart (2007) examined the effect of iron availability on particle colonisation behaviour, motility, and enzymatic activities of four strains of marine bacteria. They showed that iron depletion greatly reduced the bacterial particle colonisation rate, attributing it to reduced swimming speeds. They suggest that the limited availability of iron in the ocean could reduce bacterial motility and colonisation rate (Tang & Grossart, 2007). In clinical studies, it has been observed that bacterial biofilm formation is supressed under iron-limiting conditions also (Weinberg, 2004).

Flagella have been shown to play important roles in early biofilm formation in several Gram-negative bacteria, including *P. aeruginosa, V. cholerae* and *E. coli* (O'Toole *et al.*, 2000). Flagellar motility is thought to be necessary to overcome surface repulsion, thereby allowing initial surface contact. Evidence has been provided that flagella can act as cell-to-surface adhesins also for some bacteria. It was demonstrated in *E. coli* that cells lacking flagella were severely hindered in the initial stages of biofilm formation, indicating that motility is the key factor in early biofilm formation (Pratt & Kolter, 1998). It would therefore be interesting to examine variations in the ability of isolates of different biotypes of *Y. ruckeri* to form biofilms.

The OMP lipoprotein CsgG (involved in secretion of curli fibres) was most prominently identified in isolate RD6 (non-motile, biotype 2), in all three

replicates when grown under iron-depleted conditions. Curli are biologically important amyloid fibres that have been associated with biofilm formation, host cell adhesion and invasion, and immune system activation. It may therefore be possible that the non-motile isolate RD6 is overexpressing CsgG as a compensatory mechanism for the lack of flagella, in order to enhance biofilm formation (Barnhart & Chapman, 2006; Loferer *et al.*, 1997).

(c) Salinity. Initially, the effect of NaCl on the rate of growth was examined. Concentrations of 0% (0 ppt), 1% (10 ppt), 2% (20 ppt), 3% (30 ppt), 6% (60 ppt) and 9 % (90 ppt) NaCl were examined, with isolates RD124 and RD158 able to grow in concentrations up to 3% which mimicked that of natural seawater. However, the rate of growth was slowed as the concentration of NaCl in the media increased (Fig. 4.7). Subsequently, in view of the effects of NaCl on the growth of bacteria, an artificial seawater medium was used to be more representative of seawater than using NaCl alone. Concentrations of 0%, 1.5%, 2.5% and 3.5% seawater medium were prepared and the rate of growth of isolates RD124 and RD158 examined (Fig. 4.9). In a similar response to growth in NaCl, the rate of growth was reduced as the salinity increased suggesting that the bacteria were struggling to adapt to the new environment and spending more energy on osmoregulation.

The OMP profiles of isolates grown in seawater medium showed the greatest change from isolates grown under all other conditions (Fig. 4.16 and Fig. 4.17). However, isolates were problematic to grow under this condition with cells difficult to retrieve, thus the profiles are not as well resolved as the other growth conditions. There was drastic up- and down-regulation of many proteins on the gels, discussed previously in the results section. There was an up-regulation of proteins of similar MW to those up-regulated under iron-depleted conditions (Fig. 4.16 and Fig. 4.17, lane 5 of each panel). The major proteins of OmpC and OmpF appeared to be affected by the increasing salinites, as OmpC was up-regulated as the salinity increased, and OmpF was down-regulated. This is seen most clearly when isolates were grown in NaCl as the concentration of NaCl in the media increased (Fig. 4.8). Several proteins of LMW were identified

in all isolates when grown in seawater medium (Fig. 4.16 and Fig. 4.17, lane 5 of each panel), also.

Gel-free proteomic analysis revealed that the OM of cells grown in seawater medium had been dramatically altered (Table 4.2). OmpC.1 was not identified in any isolate during growth in seawater medium. However, OmpC.2 was observed in all three replicates of each isolate (only in two replicates of RD6) under this condition. This could indicate that OmpC.2 may have a more important role in maintaining the osmolarity of the cell than OmpC.1. Correlating with the HMW proteins observed through SDS-PAGE analysis, HemR, ShuA, FiuA and FhuA were identified through gel-free proteomics of cells grown in seawater medium. However, the Ton-B dependent receptor at ~76.6 kDa was absent. The upregulation of these proteins in this environment is unlikely to be due to a lack of iron in the media, as there are no chelating agents in the seawater mix added to TSB.

The BAM complex is essential for the assembly of ß-barrel proteins into the OM of the bacteria (section 1.4.3.1). Under growth in seawater medium the proteins BamA, BamB, BamC and BamD were not identified routinely. BamE was not generally identified through this proteomic technique in any growth condition. No members of the BAM complex were identified in any replicates of isolates RD6 and RD382, and rarely in isolates RD28, RD64 and RD420. However, BamA, BamB, BamC and BamD were identified in almost all replicates of isolates RD124, RD354 and RD366. Without these proteins, the growth rate of the bacteria would be severely limited. It is possible that these proteins were downregulated as the bacteria controlled the proteins that were incorporated into the OM under this growth condition. Many of the proteins assembled by this complex are porins and specific channels, which the bacteria may have restricted in the stressful environment to avoid cell contents leaking out, or toxic compounds coming in.

An interesting correlation was the absence of BAM complex proteins and the LPS assembly proteins, LptD and LptE. In isolates RD6, RD28, RD64, RD382 and RD420, very few BAM complex proteins were identified when cells were grown in

seawater medium and, similarly, LptD and LptE were mostly absent (Table 4.2). LptD and LptE were not identified in isolates RD382 and RD420. LptD was present in single replicates of the remaining isolates (except RD124 in which it was in two replicates). LptE was only identified in isolates RD124 (two replicates), RD354 and RD366 (one replicate). The Gram-negative bacterium *N. meningitidis* lacks both BamB and LPS (instead producing LOS; discussed previously in section 1.4.3.1). Therefore, it has been suggested that BamB may help coordinate LPS assembly or the assembly of LptD, (a protein crucial for LPS assembly at the OM) (Hagan *et al.*, 2011). Under growth in seawater medium a similar absence of LptD and LptE in the majority of isolates was observed. The LptABCDEFG complex is involved in LPS transport and assembly and has been discussed previously (Section 1.4.3.3). The OM components LptD and LptE serve as the final step in the transport of fully formed LPS molecules to the bacterial cell surface. Without these proteins, LPS cannot be transported across the OM, instead building up internally. This has been demonstrated using *ΔLptD* knockout mutants (Chng *et al.*, 2010). There is a build-up of LPS molecules inside the cell, and increased OM blebbing due to the important structural role LPS plays at the OM. It is possible that LptD and LptE are reduced to levels below the identification stringency of the MS, and that LptD and LptE do not need to be present in a large copy number to efficiently transport LPS across the membrane, meaning that isolates grown in seawater still have a fully formed LPS outer leaflet. Alternatively, as the bacteria appear to be in a particularly stressful situation demonstrated by the dramatic alteration of the OM proteome, they may have reverted to a 'survival' state and only express proteins necessary to continue metabolic processes and maintain cell integrity. In *Y. enterocolitica* transcription of the O-antigen gene cluster is affected by the phase of bacterial growth, low pH, iron concentration, oxygen tension and ionic strength (Skurnik & Bengoechea, 2009). It was suggested that since several of these factors represent *in vivo* conditions, *Y. enterocolitica* may not express O-antigen *in vivo* (Gort & Miller, 2000). This indicates that the expression of O-antigens is highly regulated and sensitive to environmental conditions. Therefore, it is possible that *Y. ruckeri* operates in the same way, and regulates the production of LPS under certain conditions, hence why LptD and LptE were less abundant under

growth in seawater medium. In order to examine this, it would be interesting to observe variations in LPS abundance under this growth condition.

In *E. coli* a surprisingly small number of genes were identified that when mutated significantly affect seawater sensitivity (Rozen & Belkin, 2001). The ability of bacteria to overcome the osmotic challenges in the marine environments largely influences their survival (Gauthier *et al.*, 1987; Munro *et al.*, 1989). Upon an osmotic upshift, bacterial cells accumulate or synthesise specific osmoprotectant molecules (trehalose, glycine betaine, glutamic acid), in order to equalise osmotic pressure and avoid drastic loss of water from the cytoplasm (Munro *et al.*, 1987). The essential role of osmoregulatory mechanisms in enteric bacterial survival in seawater has been demonstrated. Munro *et al.* (1989) and Gauthier *et al.* (1987) found that cells preadapted to high osmolarity are more readily available to survive to seawater.

As mentioned in the discussion of Chapter 3, the identification of antibodies produced by the host in response to specific proteins utilised by the bacteria could provide valuable information for the development of future vaccines. Utilising western blotting to analyse the antibodies present in fish serum could allow for the identification of immunogenic OMPs. Combining this with the proteomic analysis of different conditions conducted in this chapter, would allow for the understanding of which surface characteristics are immunogenic and under which condition they are produced.

In conclusion, the conditions in which the bacteria are grown play important roles in controlling the proteins that constitute the OM proteome. Seventy-six unique proteins were identified through gel-free proteomic analysis of eight representative isolates of *Y. ruckeri* grown under aerobic (22°C), anaerobic, aerobic (28°C), iron-depleted and seawater medium conditions. However, only three proteins were identified as the core OM proteome observed in all replicates, of all isolates under every growth condition. These included the major OMPs involved in OM biogenesis and integrity, OmpA, Braun's lipoprotein (Lpp) and Pal (peptidoglycan associated lipoprotein).

Most proteins identified in this study were recognised in three replicates highlighting the reproducibility of these methods. However, when proteins were identified in only one replicate, it was usually with a relatively low MOWSE score and few peptides (supplementary Table 8.1). The most abundant protein in all isolates and all growth conditions was Braun's lipoprotein (Lpp). This protein is important structurally for the bacteria, and is unlikely to be downregulated under any condition. Although gel-free proteomics has several advantages over gel-based proteomics, including the ability to analyse more complex samples, the technique has been shown to be more useful in the identification of lipoproteins over transmembrane proteins. As work (previous section 3.4.6) has shown the success of combining two complementary methods in increasing the number of proteins identified, the OM proteomes of isolates grown under the varying conditions would now be analysed using gel-based proteomic methods. However, the enormity of the study would render this unfeasible.

The major finding of this work is that the growth conditions that the bacteria is subjected to can greatly affect the rate of growth and induce variations in the OM proteome through up- and down-regulation of OMPs. Growing *Y. ruckeri* under aerobic aeration, anaerobic conditions, temperatures ranging from 8°C to 45°C, iron-depletion and salinites ranging from 0% to 9% all provided stressful situations for the bacteria to overcome, allowing a better understanding of the mechanisms that the bacteria have developed to survive both inside and outside of the host. Previously, poor husbandry has been blamed for the onset of ERM in both Atlantic salmon and rainbow trout (Austin & Austin, 1993). While increasing the stress of fish through handling and overcrowding will increase the onset of infection, this is most likely due to enhanced shedding of the bacteria from the fish. However, increasing the water temperature and transferring the fish to water of different salinity, will also alter the OM of the bacteria, which may in turn increase the pathogenicity of the organism. Having identified several proteins with potential roles in the virulence of *Y. ruckeri*, this work will be valuable for future vaccination studies, as the proteins that are upregulated under these conditions could be incorporated into recombinant vaccine formulations, or alternatively, whole cell vaccines prepared under a specific condition such as iron-depletion could prove valuable.

Chapter 5 Evolutionary relationships of representative isolates of *Y. ruckeri* based on housekeeping genes, LPS and OMPs

5.1 Introduction

Analysis of the genetic sequence of a bacterium can reveal important information about phylogenetic and evolutionary relationships. Further analysis of the nucleotide and amino acid sequences of genes and proteins with potential roles in host-pathogen interactions and bacterial survival can reveal important host specific properties and identify vaccine targets. The process of identifying proteins with potential roles in bacterial virulence using genetic information has been termed reverse vaccinology (Rappuoli, 2000), and has been applied to several pathogens, including *N. meningitidis*, *S. agalactiae*, *S. pyogenes*, *Streptococcus pneumoniae*, and *E. coli* (Seib *et al.*, 2012).

Proteins are composed of polypeptides, which are in turn made up of a series of amino acids in a specific order. Protein-coding DNA is divided into codons three bases long. Insertions and deletions can alter a gene so that its message is no longer correct. Variation can arise within these genes in several different ways; (i) a substitution mutation (Fig. 5.1A) exchanges one nucleotide for another, which may encode for a different amino acid or create a premature stop codon, (ii) an insertion mutation (Fig. 5.1B) in which extra base pairs are inserted into the DNA, (iii) a deletion mutation (Fig. 5.1C) in which a section of DNA is lost or deleted, and finally, (iv) a frameshift mutation (Fig. 5.1D) in which one or two nucleotides are inserted causing the reading frame to shift; this can occur as a result of insertion or deletion events.

When variation arises within a gene, it can cause changes in the amino acid sequence of the corresponding protein, resulting in alterations in structure and/or function. Even a single nucleotide change can render a protein non-functional, while the introduction of a premature stop codon may create a truncated gene.

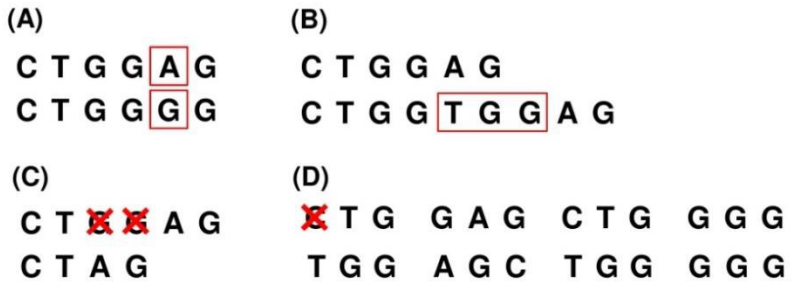


Fig. 5.1 Examples of genetic mutations that may alter the structure or function of a protein.
(A) Substitution mutation in which an individual base is substituted for another; **(B)** Insertion mutation in which additional nucleotides are inserted; **(C)** Deletion mutation were nucleotides are deleted and **(D)** frameshift mutation in which the order of the protein coding codons is altered by a deletion or insertion. The letters A, T, C and G represent the nucleotides adenine, thymine, cytosine and guanine respectively.

Previous studies examining the genetic variation of *Y. ruckeri* have utilised MLEE (Schill *et al.*, 1984), PFGE (Wheeler *et al.*, 2009) and MLST (Bastardo *et al.*, 2012a; Kotetishvili *et al.*, 2005). These studies have highlighted the homogenous nature of *Y. ruckeri*, with very little variation identified in terms of the genetic make-up of the organism. Variation within nucleotide and amino acid sequences of the proteins that compose the OM of *Y. ruckeri* is yet to be studied in detail.

Identification of the presence or absence of a protein within a specific bacterial strain can be crucial in understanding the pathogenicity of that bacterium. While the presence and absence of proteins is important, there may still be variation within the nucleotide or amino acid sequence of the protein. This in turn may render the protein host specific. In several pathogenic bacteria the roles of surface-exposed proteins in eliciting host specific infection have been observed (Pan *et al.*, 2014). These include the immunoglobulin A1 (IgA1) protease, type IV pili, complement factor H binding proteins (FHBP), gonococcal porin, transferrin-binding proteins and lactoferrin-binding proteins of *Neisseria*; the metallo-type IgA1 protease of *S. pneumoniae*; and the avian-specific AC/I pili and lamb-specific K99 fimbriae from septicemic *E. coli*. In *M. haemolytica* and *M. glucosida* the genetic diversity of *ompA* revealed variation in the hypervariable loop regions of the protein sequence, wherein isolates recovered from different host species were distinctly different from each other. This suggests that this major protein has a role in host specificity (Davies & Lee, 2004).

The aim of this section of the study was firstly to explore the phylogenetic relationship of 16 representative isolates of *Y. ruckeri* based on the variation within 19 housekeeping genes. Subsequently, the LPS encoding genes were examined in an attempt to determine the basis of the emergence of the novel serotype O1/O5 identified previously (Chapter 2). Finally, the presence, absence and variation of the OMPs that had been previously predicted (Chapter 3) was determined. We examined whether the 16 reference isolates encode the predicted OMPs or not, before examining variation within both the nucleotide and amino acid sequences of these proteins.

5.2 Materials and Methods

5.3 Sequence analysis

5.3.1 Selection of representative isolates

Sixteen isolates of *Y. ruckeri* recovered from Atlantic salmon (7 isolates), rainbow trout (8 isolates) and European eel (1 isolate) from diverse geographic locations are included in this study. Isolates collected pre 1990 are described in previous studies (Davies & Frerichs, 1989; Davies, 1989). The remaining isolates were collected between 2001 and 2014, and were supplied by MH (Marine Harvest), FVG (Fish Vet Group), DF (Dawnfresh Seafoods) or TW (Ridgeway Biologicals Ltd), and were extensively characterised in Chapter 2 (Table 2.1). The properties of all isolates used are presented as Table 5.1.

These isolates were included to represent a range of serotypes, biotypes, OMP-types and host species, as well as allowing isolates of similar phenotypes recovered in different time periods to be compared.

5.3.2 Extraction of genomic DNA

Genomic DNA was extracted using the PurElute Bacterial Genomic Kit (Edgebio; 85171). Five millilitres of an overnight culture were harvested by centrifugation at 4,000 x *g* for 5 min at 4°C. The supernatant was discarded and the bacterial pellet resuspended in 2 ml of sterile PBS and centrifuged at 13,000 x *g* for 1 min at 4°C. Four hundred microlitres of spheroplast buffer were added and the sample vortexed at high speed for 10 s to resuspend the pellet. The samples were incubated at 37°C for 10 min in a water bath. One hundred microlitres of Lysis solution 1 were added and mixed by vortexing for 10 s. One hundred microlitres of Lysis solution 2 were added and mixed by vortexing for a further 10 s. The sample was incubated at 65°C for 5 min in a water bath. One hundred microlitres of extraction buffer were added and the sample vortexed vigorously for 10 s and centrifuged at 13,000 x *g* for 3 min at 4°C. The supernatant was transferred to a clean Eppendorf tube and 100 µl of Advamax 2 beads added. The sample was inverted ten times to mix. The beads were pelleted by centrifugation. Supernatants were transferred to clean Eppendorfs, an equal volume of isopropanol added, and the sample inverted ten times to mix. The

Table 5.1 Properties of isolates of *Y. ruckeri* included in the sequencing study.

Lab Designation	Previous Designation	Source		Biotype	Serotype	OMP type	Date Isolated
		Location	Host				
RD6	-	Scotland	Rainbow trout	2	O1	1b	Pre 1990
RD10	-	Scotland	Rainbow trout	2	O1	1a	1985
RD28	BA2	U.K.	Rainbow trout	1	O5	2a	Pre 1990
RD64	F53.1/82	West Germany	Rainbow trout	1	O2	2a	1982
RD84	5710/83	Italy	Rainbow trout	1	O1	3b	1983
RD124	851014	Denmark	Rainbow trout	1	O1	3a	Pre 1990
RD150	38/85	Denmark	European eel	1	O7	1a	1985
RD162	-	Finland	Atlantic salmon	1	O6	2c	Pre 1990
RD290	-	Scotland	Atlantic salmon	1	O5	2c	Pre 1990
RD354	TW60/05	Scotland	Atlantic salmon	1	O2	2a	20/05/2005
RD366	TW90/05	Scotland	Atlantic salmon	1	O5	2c	20/07/2005
RD382	FVG 269/06	U.K.	Atlantic salmon	1	O1	3a	06/05/2006
RD420	TW110/08	Outer Hebrides	Atlantic salmon	1	O1/O5	3a	08/07/2008
RD520	FVG 205 (T1-1)	Scotland	Rainbow trout	2	O1/O5	1a	26/04/2011
RD524	TW87/11	E Scotland	Rainbow trout	2	O1	1a	21/07/2011
RD532	TW119/11	Scotland	Atlantic salmon	1	O1/O5	3a	28/09/2011

DNA was pelleted by centrifugation at 13,000 x *g* for 3 min at 4°C and the supernatant carefully removed. The DNA was washed by adding 750 µl 70% ethanol, inverted two to three times, and centrifuged at 13,000 x *g* for 3 min at 4°C. The supernatant was removed and the sample left to air dry for 30 min. Finally, the DNA was resuspended in 100 µl nuclease free water and stored at -20°C.

5.3.3 **Whole genome sequencing (WGS)**

Whole genome sequencing was conducted using the Illumina Mi-seq system at Glasgow Polyomics, employing 300-bp paired-end sequencing. Reads were trimmed of Illumina adapter sequences and low quality bases then *de novo* assembled and scaffolded using the CLC Genomics Workbench (Version 7.5.1, Qiagen). Assembled scaffolds were then annotated with the Rapid Annotation using Subsystem Technology (RAST) resource (Aziz *et al.*, 2008; Overbeek *et al.*, 2014).

5.3.4 **Nucleotide and amino acid sequence analysis of housekeeping enzymes and predicted OMPs of *Y. ruckeri***

Template sequences where obtained from the reference genome ATCC 29743 for comparison with the 16 genomes of *Y. ruckeri* sequenced (Table 5.1). Matching sequences of both housekeeping enzymes (Table 5.2) and predicted OMPs (Table 3.2) were extracted for phylogenetic and molecular evolutionary analysis using MEGA (version 5.0), in conjunction with alignment programs written by T. S. Whittam (Michigan State University). An internal Blast database was created in CLC genomics workbench (Version 7.5.1, Qiagen) and gene sequences were extracted from the 16 genomes. Sequences were formatted using Microsoft word (Microsoft, Office 2010). Finalised nucleotide sequences were firstly converted into amino acid sequences using AASeq (T.S. Whittam) and the amino acid sequences aligned using ClustalX (Conway Institute, UCD). Using RealigX (T.S. Whittam) the amino acid sequences were realigned based on the original nucleotide sequence. The aligned sequences were characterised using MEGA (Molecular Evolution and Genetic Analysis, Version 5.0; Tamura *et al.*, 2007) and a neighbour-joining phylogenetic tree (Jukes-Cantor model) generated to determine sequence variation. Boot-strap analysis (1000 replications) analysis

was conducted using Molecular Evolutionary Genetic Analysis (MEGA) software (version 5.0).

5.3.5 **Phylogeny of *Y. ruckeri***

A phylogenetic relatedness tree based on nineteen (Table 5.2) housekeeping genes was constructed using CLC genomics workbench (version 7), in conjunction with MEGA (version 5) and alignment programs written by T. S. Whittam (Michigan State University). The sequences of nineteen housekeeping genes were obtained from NCBI using reference strain ATCC 29743. The properties of these are summarised in Table 5.2. These genes were selected based on previous MLST studies used for *Y. ruckeri* (Bastardo *et al.*, 2012a), other members of the *Yersinia* genus (Hall *et al.*, 2015; Hurst *et al.*, 2011) and *P. multocida* (Davies *et al.*, 2004). Genes were extracted from the genomes of the 16 representative isolates of *Y. ruckeri* using the CLC genomics 'internal Blast' tool and concatenated using CLC genomics 'join sequences' tool (Version 7.5.1, Qiagen) generating a sequence of 22,375 nucleotides in length for each isolate. The sequences were then analysed as described above (5.3.4).

5.3.6 **Structural analysis of selected OMPs**

Secondary structure predictions were performed with the Psipred secondary structure prediction method (http://bioinf.cs.ucl.ac.ul/psipred/) and the SAM-T99 sequence alignment and modelling system (http://www.cse.ucsc.edu/research/compio/HMM-apps).

Table 5.2 The housekeeping enzyme genes used to construct phylogeny of *Y. ruckeri*.

Gene	Name	Position		Direction	Length (nt)	Study used in (and species)
		Start	Stop			
aarF	Ubiquinone biosynthesis	1853520	1855154	Complement	1635	*Yersinia* - Hall *et al*., 2015
adk	Adenylate kinase (Nucleotide biosynthesis)	3172292	3172936	Forward	645	*P. multocida* - Davies *et al.,* 2004
aroA	3-phosphoshikimate 1-carboxyvinyltransferase (Amino acid biosynthesis)	3571559	3572845	Forward	1284	*P. multocida* - Davies *et al.,* 2004
deoD	Purine nucleoside phosphorylase (Nucleotide biosynthesis)	2713635	2714360	Forward	726	*P. multocida* - Davies *et al.,* 2004
dfp	Phosphopantothenate--cysteine ligase	2094362	2095576	Forward	546	*Yersinia* - Hall *et al*., 2015
dnaJ	Chaperone protein	2747109	2748242	Forward	1134	*Y. ruckeri* - Bastardo *et al.,* 2012
g6pd	Glucose-6-phosphate 1-dehydrogenase (Energy metabolism: Pentose phosphate pathway)	212504	213979	Forward	1476	*P. multocida* - Davies *et al.,* 2004
gdhA	Quinoprotein glucose dehydrogenase A (Amino acid biosynthesis)	2313006	2314349	Complement	1344	*P. multocida* - Davies *et al.,* 2004
glnA	Glutamine synthetase	2125330	2126739	Forward	596	*Y. ruckeri* - Bastardo *et al.,* 2012
glnS	Glutamine--tRNA ligase	3298889	3300556	Forward	1668	*Yersinia* - Hall *et al*., 2015
groEL	60 kDa chaperonin	2476907	2478550	Forward	1644	*Y. enterocolitica* - Hurst *et al*., 2011
gyrB	DNA gyrase subunit B	2204186	2206600	Forward	458	*Y. ruckeri* - Bastardo *et al.,* 2012
hemA	Glutamyl-tRNA reductase	160404	161666	Forward	1263	*Yersinia* - Hall *et al*., 2015
mdh	Malate dehyrdogenase (Energy metabolism: TCA cycle)	2549973	2550908	Forward	939	*P. multocida* - Davies *et al.,* 2004
pgi	Glucose-6-phosphate 1-isomerase (Energy metabolism: Glycolysis)	1736929	1738575	Complement	1647	*P. multocida* - Davies *et al.,* 2004
recA	Recombinase A (DNA repair and recombination)	2990962	2992032	Forward	501	*Y. ruckeri* - Bastardo *et al.,* 2012
rfaE	Bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase	1512212	1513642	Complement	1431	*Yersinia* - Hall *et al*., 2015
speA	Biosynthetic arginine decarboxylase (Amino acid biosynthesis)	1410853	1412832	Complement	1980	*Yersinia* - Hall *et al*., 2015
thrA	Bifunctional aspartokinase/homoserine dehydrogenase 1 (Amino acid biosynthesis)	2732281	2734740	Forward	2460	*Y. ruckeri* - Bastardo *et al.,* 2012

5.4 Results

5.4.1 Phylogenetic relationships of representative strains of *Y. ruckeri*

Concatenated sequences of 19 housekeeping genes (Table 5.2) of *Y. ruckeri* were used to determine the phylogenetic relatedness among the 16 strains. The positions of these genes within the circularised genome are indicated (Fig. 5.2). Selecting genes spread throughout the chromosome removed bias, and avoided selecting a single area which may have undergone extensive recombination. The concatenated sequence represented 22,375 nucleotides encoding for 7458 amino acids. A phylogenetic species tree based on variation in the concatenated sequence of the representative isolates is shown in Fig. 5.3. Distinct host specific lineages are apparent, as the Atlantic salmon, rainbow trout and European eel isolates fall onto different clades of the tree, with the exception of the rainbow trout isolates RD28 and RD64, which more closely resemble the Atlantic salmon isolates (Fig. 5.3).

Rainbow trout isolates RD84 and RD124 are typical 'Hagerman' isolates; these are biotype 1, serotype O1 strains, and have grouped together on the tree. Similarly, the non-motile biotype 2 variant found in rainbow trout represented by strains RD6, RD10 and RD524 have also grouped together and are closely related to serotype O1 isolates RD84 and RD124. Within the Atlantic salmon branch isolates of serotype O5, RD290 and RD366, have grouped together and notably, with isolate RD28; this a serotype O5 isolate recovered from rainbow trout. Isolates of serotype O2, RD354, and O6, RD162, have formed their own branches, although appear more closely related to each other than to the other serotypes; serotype O2 rainbow trout isolate, RD64, is particularly closely related to these isolates.

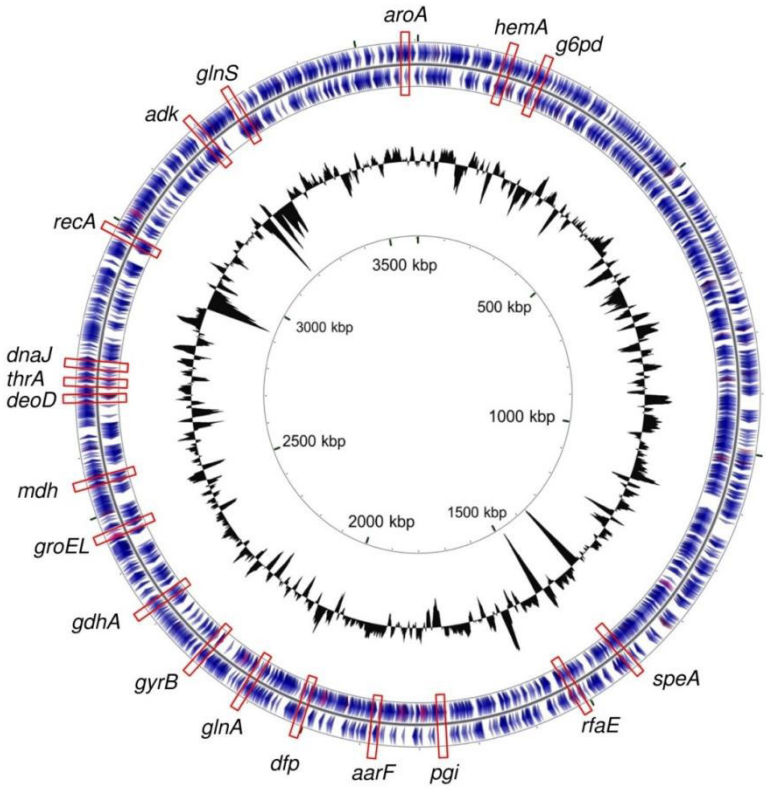


Fig. 5.2 Location of housekeeping genes on the closed chromosome of *Y. ruckeri* strain YRB.
The positions of the nineteen housekeeping enzymes used to construct phylogeny of *Y. ruckeri* within the closed genome of strain YRB (GenBank: CP009539.1) are shown. The functions, precise locations and studies from which these genes have previously been used are presented in Table 5.2 (Materials and methods). These genes are dispersed around the chromosome. The circular genome was generated using CGView Server software (V 1.0) (Grant & Stothard, 2008), and annotated using Microsoft PowerPoint.

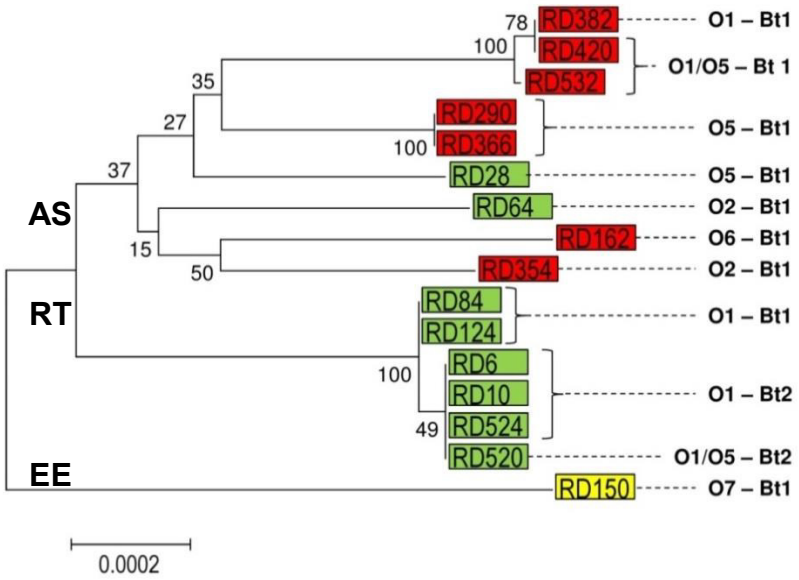


Fig. 5.3 Neighbour joining tree representing the phylogenetic relationships of 16 *Y. ruckeri* strains based on the concatenated sequences of nineteen housekeeping genes.
Isolates were recovered from Atlantic salmon (AS; red); rainbow trout (RT; green); and European eel (EE; yellow). The 19 housekeeping enzymes code for 22,375 nucleotides (7458 amino acids). Parameters assessed with a bootstrap method (1000 tests). The tree was constructed using the Jukes-Cantor correction for synonymous changes.

The isolates recovered from Atlantic salmon of serotype O1/O5, RD420 and RD532, are grouped together and are almost identical. Isolate RD382, a serotype O1 isolate recovered from Atlantic salmon is identical to RD420. This is suggestive of a common ancestral origin of these strains, and that the O1/O5 serotype has arisen from a serotype O1 isolate in Atlantic salmon. Within the rainbow trout lineage, the novel serotype O1/O5 isolate, RD520, is identical to the serotype O1 biotype 2 isolates RD6, RD10 and RD524. Again this suggests that this isolate has potentially emerged from a serotype O1, biotype 2 rainbow trout isolate. The European eel isolate, RD150, forms a distinct lineage on the tree. As isolates recovered from Atlantic salmon, rainbow trout and European eel appear to represent distinct phylogenetic lineages; this suggests an early host–associated evolutionary split within *Y. ruckeri*.

5.4.2 **Analysis of the LPS encoding operon**

The arrangement of 12 genes functioning in the synthesis of the lipid A, core and O-antigen regions of *Y. ruckeri* LPS were obtained from NCBI reference strain ATCC29473 and concatenated using CLC Genomics workbench (v7.0) (Fig. 5.4). Details of these genes are summarised in Table 5.3.

A neighbour joining phylogenetic relatedness tree was constructed using the concatenated sequence of the 12 genes encoding the LPS operon (12, 991 nucleotides) (Fig. 5.5). Isolates recovered from the same host species branched together. Isolates of serotype O1 and O1/O5 recovered from rainbow trout were identical to each other; similarly, isolates of serotype O1 and O1/O5 from Atlantic salmon were identical to each other. However, the isolates of serotype O1 and O1/O5 from rainbow trout and Atlantic salmon were located on separate branches. The serotype O5 isolates (RD290 and RD366) recovered from Atlantic salmon were grouped together, but separately from the serotype O5 (RD28) isolate recovered from rainbow trout. Isolate RD64 (a rainbow trout serotype O2 strain) was grouped most closely with the rainbow trout serotype O1 (RD6, RD10, RD524, RD84 and RD124) and O1/O5 (RD520) isolates and the Atlantic salmon serotype O1 (RD382) and O1/O5 strains (RD420 and RD532).

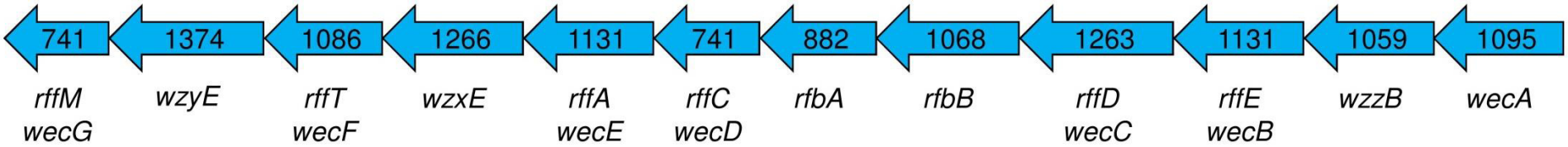


Fig. 5.4 The LPS biosynthesis cluster of *Y. ruckeri*.
The genes involved in LPS biosynthesis identified in 16 representative strains of *Y. ruckeri*. The size of the 12 individual genes is included within the arrows of the schematic.

Table 5.3 Details of LPS biosynthesis genes of *Y. ruckeri*.
Nucleotide (DNA) and amino acid (AA) variation of the individual genes identified to have roles in LPS biosynthesis in 16 strains is shown. The variation seen in the concatenated operon is also included.

Name	Gene	Synonym	Size (nt)	DNA variation	AA variation
Undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase	WecA	Rfe	1095	5	1
Chain length determinant protein	WzzB	-	1059	1	-
UDP-N-acetylglucosamine 2-epimerase	WecB	RffE	1131	4	-
UDP-glucose dehydrogenase	WecC	RffD	1263	3	1
dTDP-glucose 4,6-dehydratase	RfbB	-	1068	4	1
Glucose-1-phosphate thymidylyltransferase	RfbA	-	882	5	-
TDP-D-fucosamine acetyltransferase	RffC	WecD	741	5	3
TDP-4-oxo-6-deoxy-D-glucose aminotransferase	RffA	WecE	1131	1	-
ECA Polymerase (O-antigen translocase)	WzxE	-	1266	1	-
4-alpha-L-fucosyltransferase	RffT	WecF	1086	2	1
Putative ECA polymerase	WzyE	-	1374	6	-
UDP-N-acetyl-D-mannosaminuronic acid transferase	RffM	WecG	741	3	2
Concatenated operon	LPS		12991	46	28

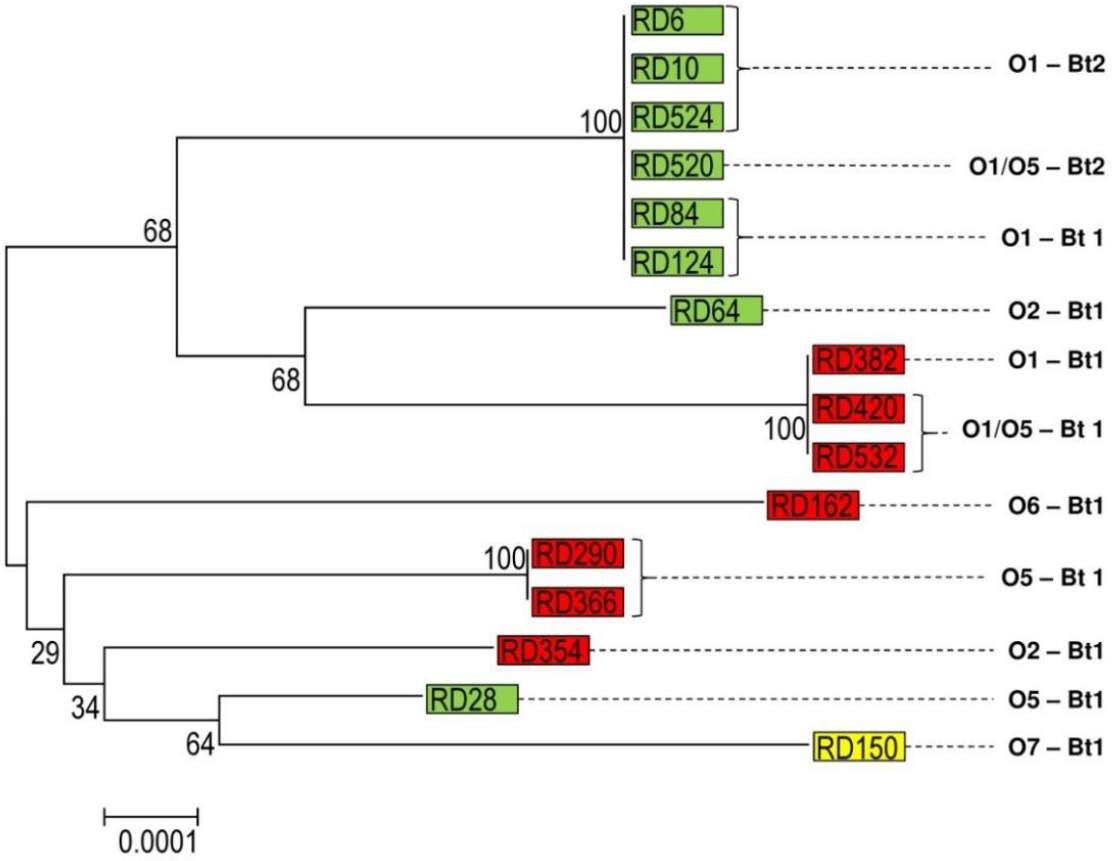


Fig. 5.5 Neighbour joining tree representing the phylogenetic relationships of 16 *Y. ruckeri* strains based on the concatenated sequence of the genes encoding LPS biosynthesis.
Isolates were recovered from Atlantic salmon (AS; red); rainbow trout (RT; green); and European eel (EE; yellow). The 12 LPS biosynthesis genes code a total of 12,991 nucleotides. Parameters assessed with a bootstrap method (1000 tests). The tree was constructed using the Jukes-Cantor correction for synonymous changes.

Overall, the diversity present within this operon is minimal. Within individual genes isolates of different serotypes differ only by a maximum of six nucleotides, while across the entire operon only 46 nucleotides were variable (Table 5.3). Within individual genes, only the mutations within *rfe, rffD, rfbB, rffC, rffT* and *rffM* were non-synonymous. The greatest variation occured within *wzyE* (encoding a putative ECA [Enterobacterial common antigen] polymerase) although this differed by only six nucleotides resulting in no amino acid alterations. This suggests that the variation in LPS structure between isolates of different serotypes is not based on these genes (Fig. 5.4). Instead, additional genes involved in glycosylation are likely to be present elsewhere in the genome.

5.4.3 **Presence, absence and variation of predicted OMPs**

The presence or absence of 141 OMPs that had been previously predicted from four genomes of *Y. ruckeri* (Chapter 3, section 3.4.1; Table 3.2) was examined in the genomes of the 16 representative isolates.

Of the 141 predicted OMPs, 130 were shown to be encoded in all 16 genomes (Supplementary Table 8.2). This represented 92.2% of the predicted OM proteome and suggests that variation in proteome composition within isolates recovered from different host species is limited. These were also relatively highly conserved with only minor variation in the nucleotide and amino acid sequences of several proteins (selected examples are shown in Table 5.4). The majority of proteins (120; 85% of the total predicted OMPs) exhibited less than 1% nucleotide variation (Fig. 5.6A). However, upon examination of the amino acid sequence (Fig. 5.6B), 57 proteins (40% of the predicted OMPs) were 100% conserved, and 116 proteins showed less than 1% variation (82% of the predicted OMPs).

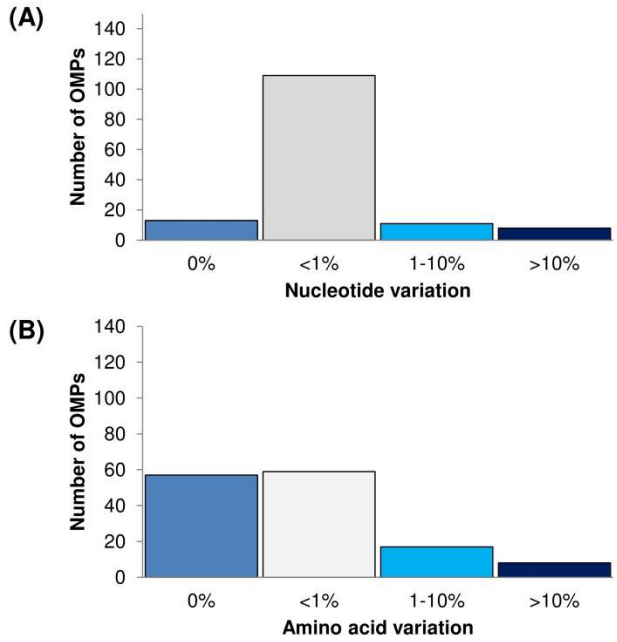


Fig. 5.6 The range of variation within the nucleotide and amino acid sequences of the predicted OMPs.
The range of variation with the nucleotide **(A)** and amino acid **(B)** sequences of the 141 predicted OMPs of *Y. ruckeri*. Within nucleotide variation, 13 genes were completely conserved in the sixteen genomes examined. One hundred and nine genes exhibited less than 1% nucleotide variation. Eleven genes varied by between 1 and 10%, while eight genes showed greater than 10% nucleotide variation. Within amino acid variation, 57 proteins were completely conserved in the sixteen genomes examined. Fifty nine proteins exhibited less than 1% amino acid variation. Seventeen proteins varied by between 1 and 10%, while eight proteins showed greater than 10% amino acid variation.

Table 5.4 Examples of variation within the OMPs identified in all sixteen genomes of *Y. ruckeri*.

Protein					Amino Acid		Nucleotide			Genome presence														
Gene locus	Name	Function	MW (kDa)	Length (aa)	Variable	%	Length (nt)	Variable	%	RD6	RD10	RD28	RD64	RD84	RD124	RD520	RD524	RD162	RD290	RD354	RD366	RD382	RD420	RD532	RD150
C4UJ19_YERRU	BamA	BAM complex	88,077	795	1	0.13%	2385	4	0.17%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UGF1_YERRU	BamB/YfgL	BAM complex	42,328	393	2	0.51%	1179	6	0.51%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UMN4_YERRU	BamC	BAM complex	38,311	349	1	0.29%	1047	1	0.10%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UNP8_YERRU	BamD/YfiO	BAM complex	27,392	243	0	0.00%	729	1	0.14%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UGL8_YERRU	BamE/SmpA	BAM complex	9,949	88	1	1.14%	264	1	0.38%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
A0A094V4E9_YERRU	FlaA	Flagellin	44,042	369	125	33.88%	1107	332	29.99%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UKI7_YERRU	FlgD	Flagellar hook assembly	23,597	229	0	0.00%	687	3	0.44%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UKI6_YERRU	FlgE	Links flagella to driver	44,281	427	4	0.94%	1281	6	0.47%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UKI4_YERRU	FlgG	Distal rod protein	27,892	260	0	0.00%	780	3	0.38%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UKI3_YERRU	FlgH	Flagella L-ring protein	23,206	219	0	0.00%	657	5	0.76%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UKI2_YERRU	FlgI	Flagella P-ring protein	38,327	369	2	0.54%	1107	7	0.63%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UKI1_YERRU	FlgK	Hook associated	58,039	547	0	0.00%	1641	6	0.37%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UKI0_YERRU	FlgL	Hook associated	34,501	322	0	0.00%	966	4	0.41%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UKG1_YERRU	FliD	Hook associated	49,084	469	0	0.00%	1407	3	0.21%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UKH4_YERRU	FliK	Hook length control	41,575	403	1	0.25%	1209	5	0.41%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UHR6_YERRU	OmpA	OM integrity, porin, adherence	37,756	355	7	1.97%	1065	18	1.69%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UFZ3_YERRU	OmpC.1	Porin (Small molecules)	40,484	368	1	0.27%	1104	9	0.82%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UKD7_YERRU	OmpC.2	Porin (Small molecules)	39,971	365	0	0.00%	1095	3	0.27%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UHQ0_YERRU	OmpF	Porin (Small molecules)	40,198	366	146	39.89%	1137	329	28.94%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UNB2_YERRU	OmpL/KdgM	Oligogalacturonate-specific porin	25,157	217	0	0.00%	651	4	0.61%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UJH6_YERRU	OmpW	Hydrophobic compound trans.	23,880	215	0	0.00%	645	0	0.00%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
C4UEU8_YERRU	OmpX	Virulence related	17,097	161	1	0.62%	483	1	0.21%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Phylogenetic trees representing the relatedness of all predicted OMPs are presented in supplementary Fig. 8.2. The majority of major proteins with roles in OM biogenesis and transport, including OmpA, OmpC and OmpF, were identified in all isolates. The BAM complex was highly conserved across all genomes examined which was not unexpected due to its importance in assembling ß-barrel proteins into the OM. Similarly, the flagellar apparatus components were generally well conserved in all isolates, including the non-motile biotype 2 rainbow trout strains RD6, RD10, RD520 and RD524.

5.4.4 **The flagellar hook associated protein FlgL**

Having constructed phylogenetic trees based on the sequence variation of the proteins flagellin (FlaA), FlgD, FlgE, FlgG, FlgH, FlgI, FlgK, FlgL, FliD and FliF (Fig. 5.7; Supplementary Fig. 8.2), the non-motile biotype 2 isolates group together only in the case of the protein FlgL (which is a hook filament junction protein) (Fig. 5.8). The variation between the motile biotype 1 and non-motile biotype 2 isolates recovered from rainbow trout was the result of a single nucleotide substitution at position 284. While isolates ATCC29473, RD84 and RD124 had a guanine residue at this position, isolates RD6, RD10, RD520 and RD524 had a cysteine. This in turn coded for an amino acid change from alanine to glycine in non-motile isolates at position 95.

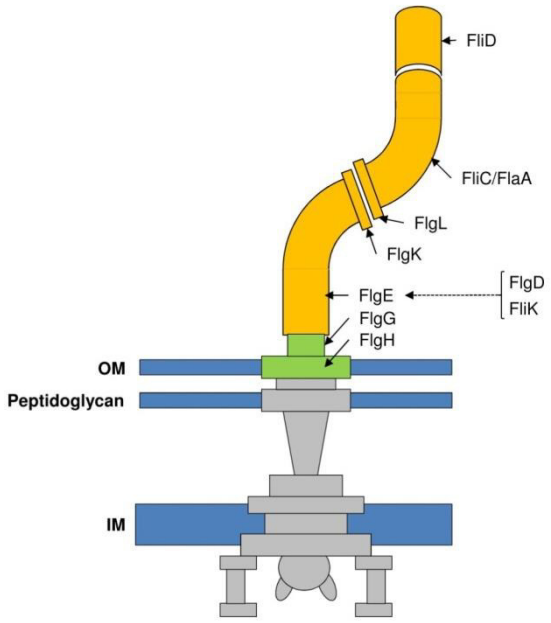


Fig. 5.7 The structure of the flagellar apparatus in Gram-negative bacteria.
Only the OM and exterior components are labelled. This figure was reproduced from the KEGG pathway (http://www.genome.jp/kegg-bin/show_pathway?eco02040).

In the study by Welch *et al.*, (2011) that identified the basis of non-motile biotype 2 isolates, the variation was shown to have arisen in *fliP, fliQ, fliR, flhB, flhA* and *flhE*. The proteins that these genes encode are not located at the OM of the bacteria (Fig. 5.7). The variation identified in biotype 2 isolates, is in the protein FlgL. Although the variation is minor, it would be interesting to examine more isolates that have been recovered from around the world and uncover further geographic variation based on this protein also.

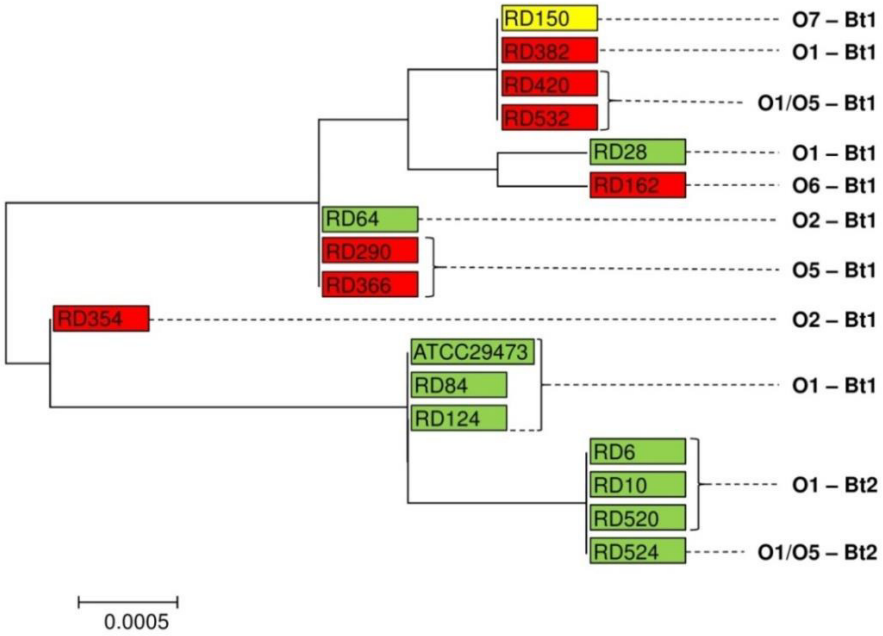


Fig. 5.8 Phylogenetic relatedness of *flgL* in sixteen strains of *Y. ruckeri*.
Strain ATCC 29473 is a reference from which the original sequence was obtained. Isolates were recovered from rainbow trout (green), Atlantic salmon (red) and European eel (yellow). The UniPROT accession number for protein FlgL is C4UK10_YERRU.

5.4.5 The major OMPs OmpA and OmpF

5.4.5.1 OmpA

The sequence variation of OmpA in all sixteen genomes was minimal, differing by a maximum of only 18 nucleotides out of an encoded 1065 (1.69%) and seven out of an encoded 355 amino acids (1.97%). Through the construction of a phylogenetic relatedness tree (Fig. 5.9), there did appear to be a host specific and serotype specific trend to the grouping. However, the conserved nature of OmpA was highlighted as the variation between lineages was minimal. The rainbow trout serotype O1 and O1/O5 isolates were identical with respect to their OmpA sequence, and very similar to the Atlantic salmon serotype O1 and

O1/O5 isolates. The Atlantic salmon serotype O5 isolates were also identical to each other, as were the isolates of serotype O2. The OmpA of serotype O5 rainbow trout isolate, RD28, was identical to those of serotype O2 isolates, RD64 and RD354 recovered from rainbow trout and Atlantic salmon respectively. However, these isolates were very closely related to the Atlantic salmon serotype O5 strains, RD290 and RD366. The serotype O6 strain, RD162, was closely related to all serotype O2 and O5 isolates, while the OmpA sequence of serotype O7 isolate, RD150, was most similar to the isolates of serotype O1 and O1/O5.

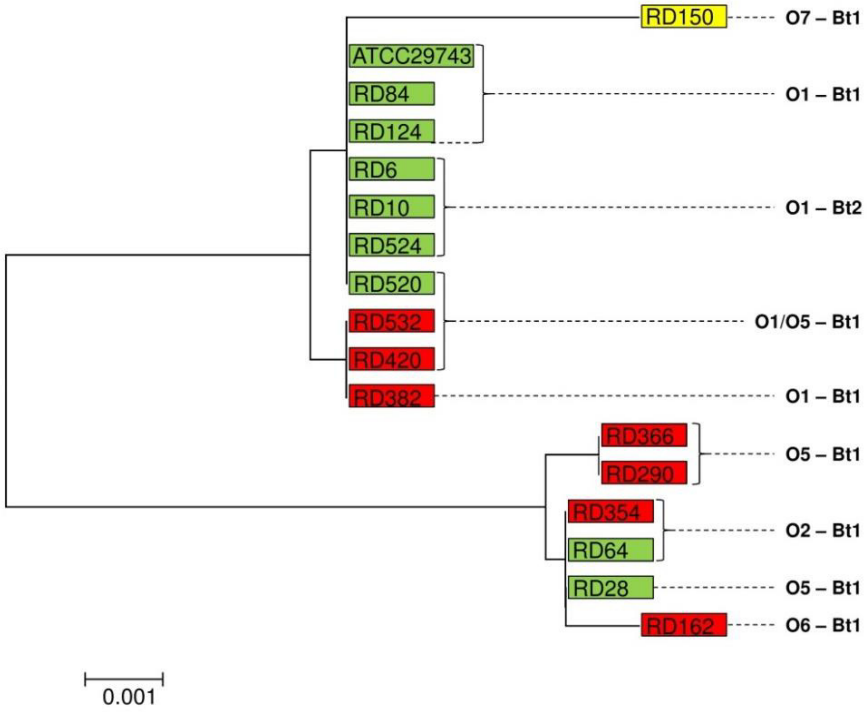


Fig. 5.9 Phylogenetic relatedness of *ompA* from *Y. ruckeri*.
Strain ATCC 29473 is a reference from which the original sequence was obtained. Isolates were recovered from rainbow trout (green), Atlantic salmon (red) and European eel (yellow). The UniPROT accession number for protein OmpA is C4UHR6_YERRU.

When comparing the phylogenetic relatedness tree of *ompA* (Fig. 5.9) to the phylogenetic tree produced based on the concatenated sequence of nineteen housekeeping genes (Fig. 5.8), it appears as if the variation in *ompA* has not arisen through a recent HGT event. The isolates have branched together in the same clusters on both trees, suggesting that the divergence of *ompA* was an early event and that the isolates have a common evolutionary ancestor.

Upon inspection of the amino acid sequence and identification of the loop and ß-barrel transmembrane domains in the protein, three distinct alleles were

identified (Fig. 5.10). These were designated *ompA.1 to ompA.3*. The association of *ompA* alleles within *Y. ruckeri* was more dependent on serotype than host species. All serotype O1 and O1/O5 isolates were represented by *ompA.1*, while serotypes O2, O5 and O6 were represented by *ompA.2*. The serotype O7 isolate recovered from European eel was of *ompA.3*.

The variation predominantly appeared to be in the loop regions of the protein, specifically in loop three (Fig. 5.10). While isolate RD150 had identical loop regions to that of the serotype O1 and O1/O5 rainbow trout and Atlantic salmon isolates, a single amino acid variation in the transmembrane domain immediately before loop three was identified.

Construction of a predicted secondary structure of OmpA indeed reveals that the variable region is in loop three (Fig. 5.11), with the variation occurrng at the tip of the loop. This is potentially the point of adherence for this protein to host tissue. However, it does not appear to be host specific as there is no variation in OmpA structure between isolates recovered from Atlantic salmon or rainbow trout. Historically, serotype O1 isolates are particularly pathogenic. The variation in loop three may be partially responsible for this.

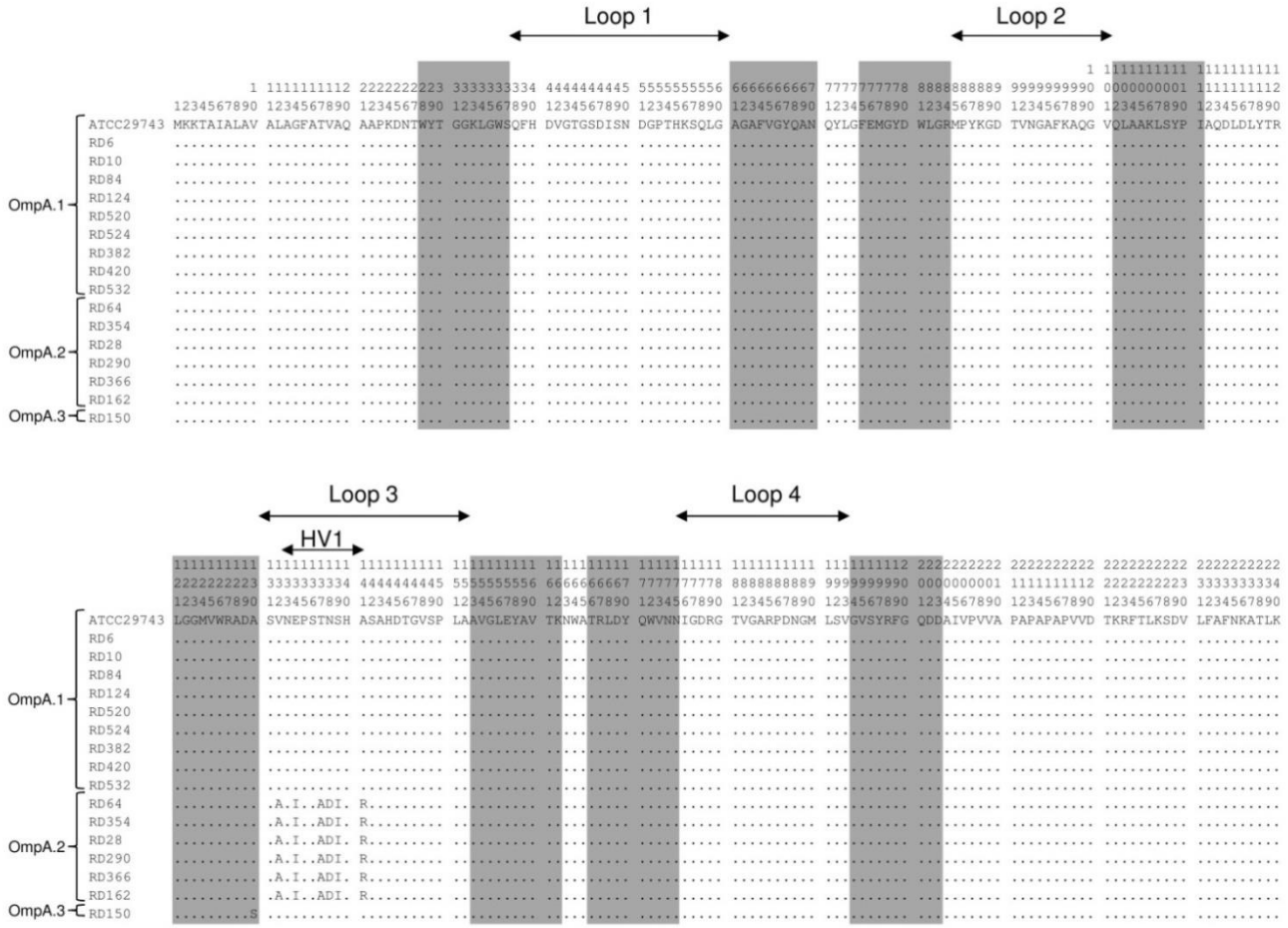


Fig. 5.10 Distribution of variable inferred amino acid sites in the N-terminal transmembrane domains of 16 OmpA proteins of *Y. ruckeri*.
The numbers above the sequences (read vertically) represent amino acid positions. The dots represent sites where the amino acids match those of the first (topmost sequence). HV1 represents the hypervariable domain within surface-exposed loops 1 to 4. Shaded regions represent predicted membrane–spanning β-strand structures. Distinct alleles are represented by OmpA.1, OmpA.2 and OmpA.3.

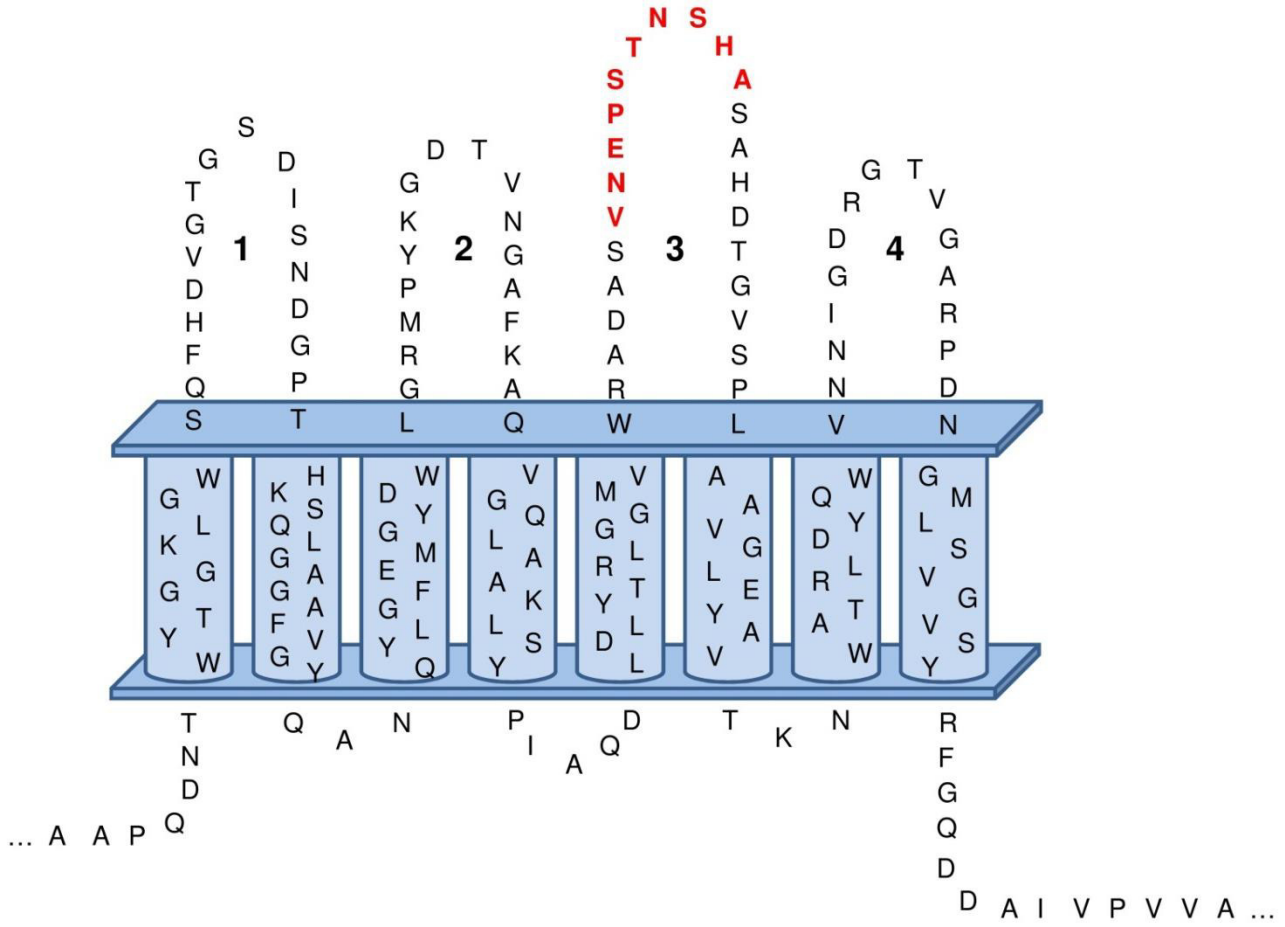


Fig. 5.11 Proposed secondary structure of N-terminal transmembrane domain of the OmpA protein of *Y. ruckeri.*
The sequence is based on OmpA of strain ATCC29473. The hyper variable domains are shown in bold red. Domains were predicted using Pred TMBB based on Hidden Markov Models (HMM). Sequences at both N- and C-terminals are truncated.

5.4.5.2 OmpF

Substantially more variation was observed in the major protein OmpF than in OmpA. This protein differed in 329 out of 1137 nucleotides (29%), and 146 out of 366 amino acids (40%). The variation was evident in the phylogenetic relatedness tree (Fig. 5.12), as there appeared to be a host specific and serotype specific trend to the grouping. Within the serotype O1 isolates recovered from rainbow trout, isolates of biotype 1 were identical to each other. However, the biotype 2 isolates were represented on a separate phylogenetic lineage. The Atlantic salmon serotype O1 and O1/O5 isolates were identical to each other. The serotype O5 isolates recovered from Atlantic salmon, RD290 and RD366, were very closely related to both the serotype O1 and O1/O5 Atlantic salmon groups. The OmpF protein from isolates of the remaining serotypes were more removed from the cluster, although isolates tended to remain more closely associated to one another in terms of the host species they were recovered from.

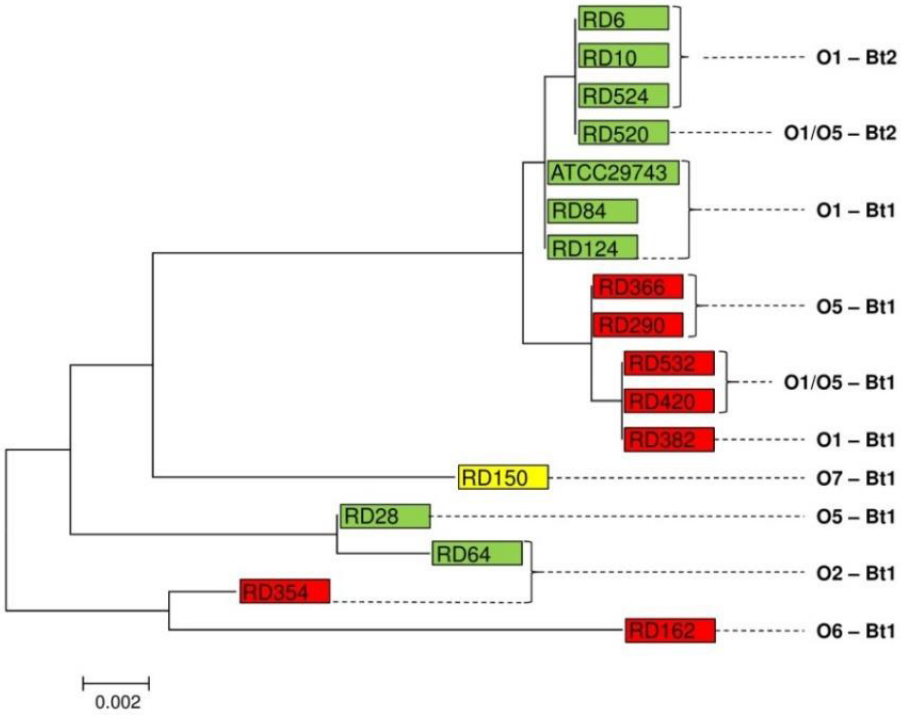


Fig. 5.12 Phylogenetic relatedness of *ompF* from *Y. ruckeri*.
Strain ATCC 29473 is a reference from which the original sequence was obtained. Isolates were recovered from rainbow trout (green), Atlantic salmon (red) and European eel (yellow). The UniPROT accession number for protein OmpF is C4UHQ0_YERRU.

When comparing the phylogenetic relatedness tree of *ompF* (Fig. 5.12) to the phylogenetic tree produced based on the concatenated sequence of nineteen

housekeeping genes (Fig. 5.8), it appears as if the variation in *ompF* has not arisen through a recent HGT event. The isolates have branched together in the same clusters on both trees, suggesting that the divergence of *ompF* was an early event and that the isolates have a common evolutionary ancestor.

Upon inspection of the amino acid sequences and identification of the loop and ß-barrel transmembrane domains in the protein, individual sequences assigned as distinct alleles displayed much greater diversity than in OmpA; seven major groups were distinguished (ompF.1 to ompF.7) (Fig. 5.13).

OmpF alleles correlated with host species and to some degree OMP-type. The first group, *ompF.1* was composed of three rainbow trout isolates, all of serotype O1 and OMP-type 3. Isolates of *ompF.2* were all from Atlantic salmon, although serotypes O1, O5 and O1/O5 were all present. These isolates were either of OMP-type 2c or 3a. Isolates of *ompF.3* were all non-motile rainbow trout isolates, of serotype O1 or O1/O5 and OMP-type 1. The single isolate of *ompF.4* was recovered from European eel, while both isolates of *ompF.5* were recovered from rainbow trout (these were serotype O2 and O5, and OMP-type 2a). The isolates of *ompF.6* and *ompF.7* were recovered from Atlantic salmon, and of OMP-types 2a and 2c respectively.

The variation again appeared predominantly in the loop regions of the protein. OmpF has eight surface exposed loops as opposed to the four of OmpA (Fig. 5.10). Variation was observed in all loops except loop 3, which displayed minimal variation; the variation in loop three was that of a single amino acid substitution at site 141 in strains RD354 and RD162. This led to the identification of six hypervariable domains (HV1-HV6). The most variable loops were loops four (HV2) and five (HV3). In HV2 (loop 4) several amino acid variations were evident in strains RD28, RD64, RD150, RD162 and RD354. In HV3 (loop 5), 13 additional amino acids had been inserted from positions 228 to 240 in the non-motile biotype 2 isolates RD6, RD10, RD520 and RD524. Construction of a predicted secondary structure of OmpF reveals that the variable regions are again most prominent at the tips of the loops (Fig. 5.14).

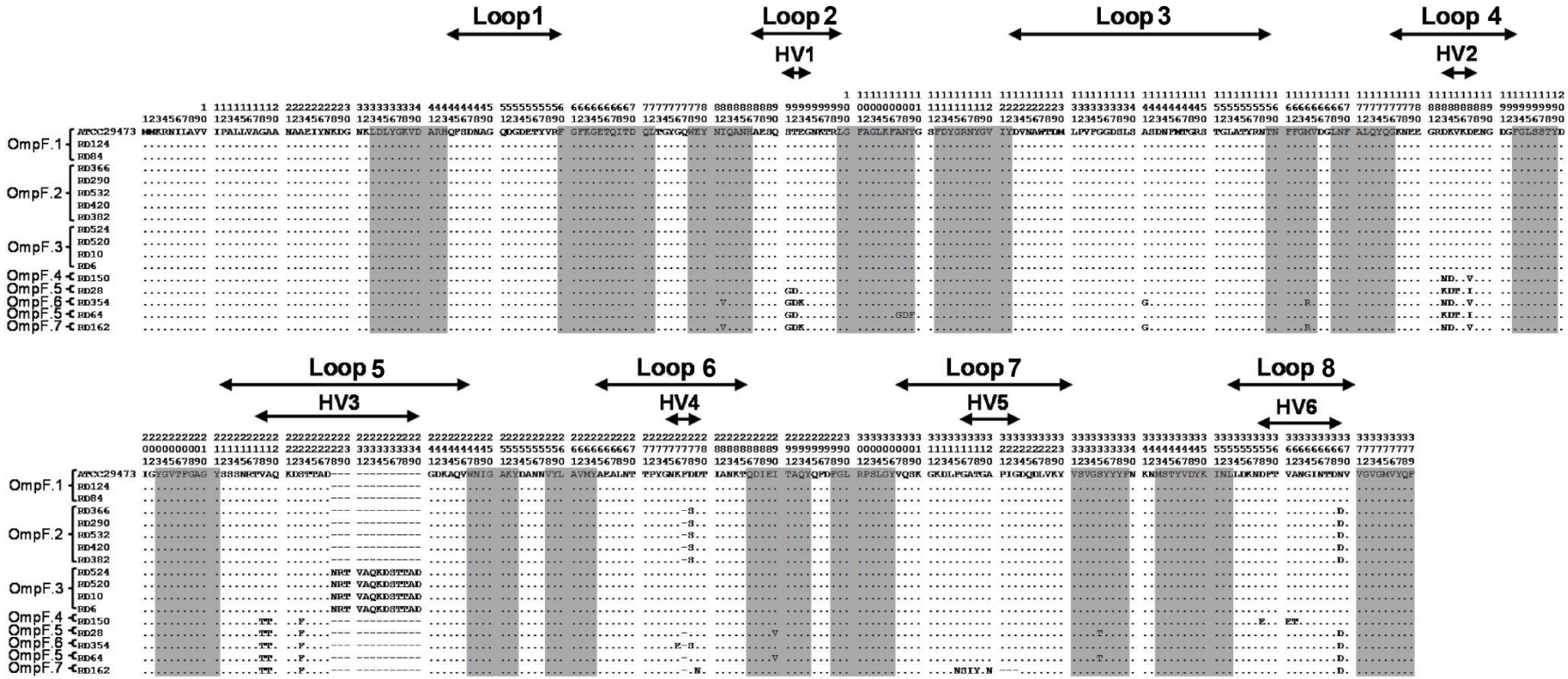


Fig. 5.13 Distribution of variable inferred amino acid sites in the N-terminal transmembrane domains of 17 OmpF proteins of *Y. ruckeri*.
The numbers above the sequences (read vertically) represent amino acid positions. The dots represent sites where the amino acids match those of the first (topmost sequence). HV1-6 represents the hypervariable domains within surface-exposed loops 1 to 8. Shaded regions represent predicted membrane–spanning β-strand structures. Distinct alleles are represented by OmpF.1 to OmpF.7.

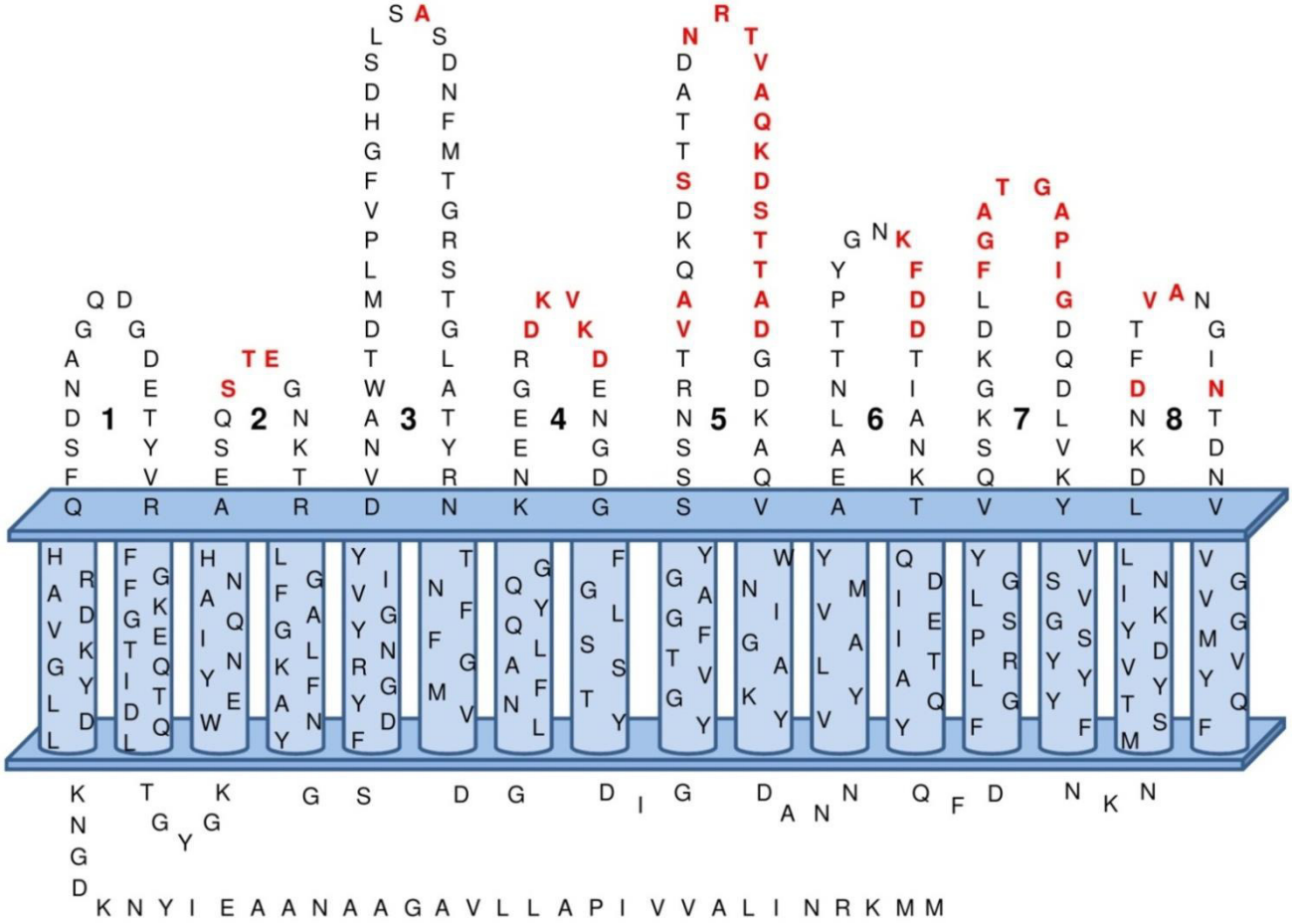


Fig. 5.14 Proposed secondary structure of N-terminal transmembrane domain of the OmpF protein of *Y. ruckeri.*
The sequence is based on OmpA of strain ATCC29473. The hyper variable domains are shown in bold red. Domains were predicted using Pred TMBB based on HMM.

A high d_S/d_N ratio (greater than one) indicates that natural selection is acting to restrict mutations that result in amino acid alteration and replacement, i.e. selective constraint. Conversely, a low d_S/d_N ratio (less than one) indicates that natural selection is actively driving amino acid alteration and replacement, i.e. diversifying selection. OmpA was divided into nine domains (with each loop representing an individual domain, and the inter-loop regions another), while OmpF was divided into 16 domains. Within *ompA*, the d_S/d_N ratios were not calculable for all the hypervariable extracellular loop domains (L1, L2, L3 and L4), due to the conserved nature of the domains in this protein. In loop 3, where the most variation existed the ratio was less than 1 (0.46). This provides evidence that natural selection is driving the diversification in this hypervariable loop. The lack of diversity seen in the transmembrane domains is expected, as these are mostly involved in maintaining OM structural integrity and cannot tolerate amino acid change.

Table 5.5 The d_S/d_N ratios for the domains and loop regions of the major porin OmpA in *Y. ruckeri*

	OmpA		
	d_S	d_N	d_S/d_N
Domain 1	0.000	0.000	0.00
Loop 1	0.000	0.000	0.00
Domain 3	0.000	0.000	0.00
Loop 2	0.000	0.000	0.00
Domain 5	0.118	0.000	n/a
Loop 3	1.934	4.243	0.46
Domain 7	0.008	0.000	0.00
Loop 4	0.000	0.000	0.00
Domain 9	0.004	0.000	0.00
All loops	1.934	4.243	0.46
All conserved	0.000	0.000	0.00

Within *ompF*, the d_S/d_N ratios were calculable for all the hypervariable extracellular loop domains (except loop 1). All loops had a d_S/d_N ratio below 1 again providing evidence that natural selection is driving the diversification in the hypervariable loop regions of this protein. The d_S/d_N ratio calculated for all loop regions in this protein was 0.42, which was substantially lower than the 0.88 obtained for the conserved domains. This is expected, as the conserved domains act to restrict amino acid replacement in the non-loop regions.

Table 5.6 **The d_S/d_N ratios for the domains and loop regions of the major porin OmpF in *Y. ruckeri***

	OmpF		
	d_S	d_N	d_S/d_N
Domain 1	0.000	0.000	n/a
Loop 1	0.000	0.000	n/a
Domain 3	0.000	0.221	0.00
Loop 2	0.691	1.368	0.51
Domain 5	0.118	0.353	0.33
Loop 3	0.000	0.221	0.00
Domain 7	0.000	0.221	0.00
Loop 4	1.835	2.647	0.69
Domain 9	0.750	0.000	n/a
Loop 5	0.382	1.765	0.22
Domain 11	0.000	0.000	n/a
Loop 6	0.309	1.206	0.26
Domain 13	0.221	0.221	1.00
Loop 7	0.353	1.059	0.33
Domain 15	0.000	0.221	0.00
Loop 8	0.294	0.941	0.31
All loops	3.882	9.206	0.42
All conserved	1.088	1.235	0.88

5.4.6 **Proteins not present in all genomes**

Eleven proteins (7.8%) out of the 141 predicted to be OM located were not identified in all isolates (Table 5.7). The proteins PilP, PilQ and PilV (which have roles in pilus biogenesis) were only identified in isolates recovered from rainbow trout, and significantly only in the serotype O1 and O1/O5 strains. PilP was not however identified in rainbow trout isolate RD6 (serotype O1, biotype 2). Similarly, proteins TraL and TraN were found only in isolates of serotype O1 and O1/O5; the sequences of these proteins were also significantly 100% conserved.

The major fimbral subunit, FimA, was identified in all isolates, except the serotype O1 and O1/O5 Atlantic salmon strains. The variation within this protein however, was very high; 406 nucleotides out of 1050 (39%) were variable with all strains, equating to 186 out of 350 amino acids (53%). Isolates did tend to group together in terms of serotype (Supplementary Fig. 8.2). The rainbow trout serotype O1 and O1/O5 isolates were on the same branch of the phylogenetic tree, as were the serotype O5 Atlantic salmon isolates. The rainbow trout serotype O5 isolate, RD28, was closely related to the O1 and O1/O5 group, while the serotype O2 isolates from both species were similar to the serotype O5 isolates. The serotype O6 and O7 isolates from Atlantic salmon and European eel respectively were more removed from the general cluster.

Table 5.7 Variation in selected OMPs in sixteen genomes of *Y. ruckeri*.

Protein					Amino Acid		Nucleotide			Genome presence															
Gene locus	Name	Function	MW (kDa)	Length (aa)	Variable	%	Length (nt)	Variable	%	RD6	RD10	RD28	RD64	RD84	RD124	RD520	RD524	RD162	RD290	RD354	RD366	RD382	RD420	RD532	RD150
C4UN28_YERRU	Invasin 1	Cell adhesion	93,186	842	0	0.00%	2526	1	0.04%	+	+	-	-	+	+	+	+	-	-	-	-	+	+	+	+
C4UP93_YERRU	Invasin 2	Cell adhesion	54,668	501	1	0.20%	1503	2	0.13%	+	+	-	+	+	+	+	+	-	+	+	+	+	+	+	-
C4T4V0_YERIN	Lipoprotein 1	Unknown	20,092	174	0	0.00%	522	1	0.19%	-	-	+	+	-	-	-	-	+	+	+	+	-	-	-	+
C4UFF3_YERRU	OmpE	Inorganic phosphate uptake	39,949	365	0	0.00%	900	2	0.22%	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+	-
C4UGD2_YERRU	FimA	Fimbrial protein	38,619	350	186	53.14%	1050	406	38.67%	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+
C4UKM4_YERRU	MltE	Peptidoglycan maintenance	18,138	161	0	0.00%	483	0	0.00%	+	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
C4UKM0_YERRU	PilP	Type IV pilus bio.	18,959	181	0	0.00%	543	0	0.00%	-	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
C4UKM1_YERRU	PilQ	Type IV pilus biogenesis	51,393	420	0	0.00%	1260	0	0.00%	+	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
C4UKM6_YERRU	PilV	Minor pilin	46,734	420	0	0.00%	1260	0	0.00%	+	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
C4UKN0_YERRU	TraL	F Pilin formation	9,748	94	0	0.00%	282	0	0.00%	+	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
C4UKN3_YERRU	TraN-Like	F Plasmid conj. Transfer	27,705	255	0	0.00%	765	0	0.00%	+	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-

The phosphoporin OmpE was absent from the more recently recovered rainbow trout isolates, RD520 and RD524, and similarly the European eel isolate RD150. The isolates in which OmpE was identified tended to group together by host species or serotype. The Atlantic salmon O1 and O1/O5 isolates were identical, and closely related to the rainbow trout serotype O1 isolates. Isolates of serotype O2, O5 and O6 were identical to each other, but separate from the O1 and O1/O5 isolates. In general, this protein was well conserved, with only 2 nucleotides out of 900 varying (0.22%).

A lipoprotein of unknown function (lipoprotein 1) was not identified in any of the serotype O1 or O1/O5 isolates from either host species, while in the remaining isolates of serotype O2, O5, O6 and O7 lipoprotein 1 was highly conserved (except the serotype O2 rainbow trout isolate RD64, which differed only by one nucleotide [0.19%]).

An invasin (invasin 1) was identified only in O1 and O1/O5 isolates from Atlantic salmon and rainbow trout. However, it was identified in the serotype O7 isolate recovered from European eel, RD150. This invasin was highly conserved varying only by a single nucleotide in isolates RD382, RD420 and RD532 (the Atlantic salmon serotype O1 and O1/O5s respectively). A second invasin (invasin 2) was absent from the serotype O5 rainbow trout strain (RD28), the serotype O6 Atlantic salmon isolate (RD162) and the serotype O7 European eel isolate (RD150), however was highly conserved amongst the remaining isolates.

5.5 Discussion

Previously (Chapter 2), a novel LPS-type was identified which represented the serotype responsible for the majority of ERM disease infections within the Atlantic salmon population. Crucially, this serotype has been recovered from rainbow trout in more recent years, albeit only on two occasions. In order to understand the phylogenetic relationships between strains of different serotypes recovered from different host species (primarily Atlantic salmon and rainbow trout), a small scale phylogenetic study was conducted. The construction of a phylogenetic tree based on the concatenated sequences of 19 selectively neutral housekeeping genes (Fig. 5.2) in 16 strains, provides support to our previous suggestion (section 2.4) that the O1/O5 LPS gene cluster has been horizontally transferred between strains of *Y. ruckeri* recovered from Atlantic salmon and rainbow trout (Fig. 5.3). The serotype, O1/O5, is represented by Atlantic salmon isolates RD420 and RD532 and rainbow trout isolate RD520. As the Atlantic salmon and rainbow trout isolates are located on divergent branches of the phylogenetic tree; this suggests that the O1/O5 O-antigen biosynthesis gene cluster has been horizontally transferred between these host species. The O1/O5 serotype most likely arose first in Atlantic salmon because it was first identified, and is far more abundant in, Atlantic salmon isolates (Fig. 2.6, Chapter 3). The serotype O1/O5 isolate has only been recovered in rainbow trout since 2010, while they appeared in Atlantic salmon initially in 2002, before becoming predominant in 2006. A second more recent event involving the HGT of the O1/O5 gene cluster between Atlantic salmon isolates and rainbow trout isolates is likely to have followed. However, it is also plausible that this serotype may have emerged independently in rainbow trout. Although our data is based on 19 housekeeping genes and 16 isolates it reveals useful information about the population structure and evolution of *Y. ruckeri*. The inclusion of an isolate recovered from European eel is significant as it provides evidence for an earlier host associated evolutionary split within *Y. ruckeri*. Three distinct branches are apparent on the tree with isolates recovered from Atlantic salmon, rainbow trout and European eel mostly residing on individual branches (Fig. 5.5).

In order to identify if the differences in LPS-profiles is brought about as a result of nucleotide or amino acid variation in the sequences of the genes and proteins encoded on the LPS-biosynthesis cluster, a neighbour joining phylogenetic

relatedness tree (Fig. 5.5) was constructed using the concatenated sequences (12,991 nucleotides) of the entire LPS operon (Fig. 5.4). Isolates recovered from the same host species branched more closely together. However, isolates did not branch entirely by serotype. Isolates of serotype O1 and O1/O5 recovered from rainbow trout were identical, while similarly isolates of serotype O1 and O1/O5 from Atlantic salmon were identical to each other. However, the isolates of serotype O1 and O1/O5 from rainbow trout and Atlantic salmon were on separate branches. This indicates that while the O1 and O1/O5 LPS-biosynthesis clusters are identical within each host species, the variation in O-antigen structure must be brought about by additional genes that are encoded elsewhere in the genome. However, this does not rule out a HGT between isolates. It is possible that a gene involved in glycosylation has been horizontally transferred to an Atlantic salmon serotype O1 isolate, possibly from a serotype O5 isolate, resulting in the emergence of the O1/O5 LPS-type. This gene could in turn have been horizontally transferred to a rainbow trout serotype O1 isolate resulting in the emergence of this serotype in this species. Genes involved in glycosylation may alter the O-antigen structure, resulting in differences in LPS profile. In *Salmonella*, phase variation of the genes encoding the glycosyltransferases (the *gtr* operon) has been shown to elicit modifications in the O-antigen that contribute to the diversity of the serotypes. The *gtr* operon consists of three genes; gtrA (a membrane protein), gtrB (a glycosyl translocase) and gtrC (a glycosyltransferase) which mediates the bonding of glucose to the O antigen (Reyes *et al.*, 2012). In *S. flexneri*, the glycosyltransferases are encoded within the genomes of temperate phages. When the phage lyses, the glycosyltransferase genes are transcribed resulting in serotypic conversion (Allison & Verma, 2000). Therefore, it is possible that a glycosyltransferase gene (or cluster) has been horizontally acquired by a serotype O1 strain of *Y. ruckeri* resulting in the emergence of a new LPS-type, serotype O1/O5.

Isolates of serotype O5 recovered from Atlantic salmon, RD290 and RD366, were identical to each other. However, the rainbow trout serotype O5 isolate, RD28, was not identical to these. The serotype O2 isolates, RD64 and RD354, recovered from rainbow trout and Atlantic salmon respectively, were located separately from one another on the phylogenetic tree. Also, isolates of serotype O6, RD162, and O7, RD150, were located on independent branches.

Overall, the diversity present within the concatenated LPS operon is minimal, with no individual gene appearing responsible for the differences in serotypic profiles.

The presence or absence of a protein in a specific isolate can drastically impact the pathogenic capabilities of that strain. If a protein with a known role in survival or virulence is absent from a particular strain, then the ability of that isolate to persist in the environment, invade the host, survive within the host, and ultimately outcompete other bacterial species for resources will be diminished. By examining the presence or absence of proteins that were predicted to be located in the OM of *Y. ruckeri*, and examining the variation within these proteins both at a nucleotide and amino acid level, proteins that may be host specific can be identified.

Having examined the 141 proteins that had been predicted to be OM located, the vast majority of these proteins were identified to be present in the genomes of all 16 representative isolates. As these were selected from the previous characterisation study as representative of the wider collection, we can hypothesise that most OMPs are generally well conserved within *Y. ruckeri*, at least in terms of presence or absence. This was to be expected due to the previously discussed genetic homogeneity of this organism (Bastardo *et al.*, 2012a; Kotetishvili *et al.*, 2005; Schill *et al.*, 1984; Wheeler *et al.*, 2009).

Within the 130 proteins that were identified in all genomes (supplementary Table 8.2), the degree of variation of both nucleotide and amino acid sequence was limited. Thirteen OMP-encoding genes had a 100% conserved nucleotide sequence, while 109 had less than 1% variation, 11 had less than 10% variation and only eight OMP-encoding genes showed greater than 10% nucleotide variation. The majority of mutations were synonymous substitutions with no protein alterations (supplementary Table 8.2). This therefore would not alter the function, structure, or host specificity of these proteins. Several proteins that have important functions in maintaining the structure of the cell, either through synthesis of new proteins (BAM complex) or providing important structural support to the cell (Pal, Braun's lipoprotein [Lpp] and OmpA) were well conserved; this is most likely due to the importance of these proteins in

maintaining cell shape and viability meaning that these proteins are generally resistant to mutations through selective pressure.

Proteins that comprise the OM and extracellular components of the flagellar apparatus were identified in all isolates. Isolates RD6, RD10 and RD524 were of biotype 2 (non-motile) and do not have a functional flagella. However, all components of the flagellar apparatus were identified in these strains. Welch *et al.*, (2011) described mutations in components of the apparatus that render the flagella non-functional. These are primarily located in the IM components (*fliP, fliQ, fliR, flhB, flhA* and *flhE*). However, the OM or extracellular components were not examined (Welch *et al.*, 2011).

While most of the proteins were well conserved amongst isolates, both flagellin (FlaA) and FlgL were less so. Flagella mediate bacterial motility; adhesion and virulence are sometimes enhanced by flagellar expression and motility. Flagellin is also a pathogen-associated molecular pattern (PAMP) recognised by the innate immune system through TLR5 (discussed previously 2.4). A comprehensive study involving many flagellated bacteria observed that unique flagellin sequences of different species contain amino acid variations in the TLR5 recognition site that permit TLR5 evasion. These strains also possess compensatory mutations that preserve bacterial motility (Andersen-Nissen *et al.*, 2005). This suggests that as flagellin is immunogenic, the variation may be driven by immue evasion. However, futher analysis would be necessary for confirmation of this in *Y. ruckeri*.

Variation in the hook associated protein FlgL allowed the non-motile biotype 2 isolates to be distinguished from the remaining motile biotype 1 isolates. The mutation in FlgL was a single nucleotide substitution at position 284, which in turn coded for an amino acid change from alanine to glycine in non-motile isolates at position 95. As Welch *et al.*, (2011) demonstrated that motility could be restored through genetic complementation with IM components; this suggests that the mutation in FlgL does not limit flagellar activity. It would be interesting however, to determine if the same mutation in FlgL that has been identified in these UK isolates, is present in isolates from elsewhere in the world.

Upon examination of the sequence variation in the major protein OmpA, a noticeable lack of diversity was apparent (Fig. 5.9). Sequence diversity to account for the subtle variations in molecular mass observed through SDS-PAGE was expected. However, the protein appeared remarkably well conserved amongst isolates of different OMP-types, serotypes, biotypes and those recovered from different host species. While subtle variation was identified between isolates of serotypes O2, O5, O6 and O7 with those of serotype O1 and O1/O5, this was only in seven amino acids across the entire protein. Six of these amino acids were located in surface exposed loop three (Fig. 5.10 & Fig. 5.11). Natural selection may be acting to restrict amino acid replacement in the non-loop regions of OmpA. The membrane spanning and periplasmic domains of the protein are mostly involved in maintaining OM structural integrity (Koebnik *et al.*, 2000), and may not tolerate amino acid change. Calculation of the d_S/d_N ratio of extracellular loop three (0.46) provides strong evidence that natural selection is driving diversification in the hypervariable extracellular loop region.

OmpA is highly conserved among the enterobacteriaceae family (Jeannin *et al.*, 2002). The multifunctional nature of OmpA has been described previously (section 1.4.3.4.1 of the Introduction). The physiological function of OmpA is to provide a physical link between the OM and the underlying peptidoglycan layer. Mutations of OmpA and Braun's lipoprotein (Lpp) lead to spherical cells that can only survive under well-balanced osmotic conditions (Sonntag et al., 1978). OmpA also acts as a porin (Sugawara & Nikaido, 1992), bacteriophage receptor (Morona *et al.*, 1985), an adhesin/invasin (Torres & Kaper, 2003), in immune evasion for macrophage survival (Sukumaran *et al.*, 2003), and in biofilm formation (Orme *et al.*, 2006). The surface-exposed loops of OmpA have been shown to extrude beyond the capsule in *M. haemolytica* (Hounsome *et al.*, 2011), and have had roles in host adaptation elucidated (Davies & Lee, 2004).

The most abundant lipoprotein in Gram-negative bacteria is Braun's lipoprotein (Lpp). This protein anchors the OM to the peptidoglycan, interacting with Pal (a peptidoglycan associated protein) together contributing to the integrity of the cell wall (Kovacs-Simon *et al.*, 2011).

Isolates of serotype O1 have historically been particularly virulent in rainbow trout worldwide (Austin *et al.*, 2003; Davies & Frerichs, 1989; Davies, 1991c, d), while isolates of serotype O1/O5 are now the most commonly recovered serotype in UK Atlantic salmon (Chapter 2). The virulence of this serotype to both species is unclear at present, however as we hypothesise that the serotype O1/O5 isolates have arisen from serotype O1 strains, they are most likely virulent towards both species, and were recovered from the kidneys of both Atlantic salmon and rainbow trout during the previous bath challenge experiments (section 2.3.7). It is therefore possible, that the conserved amino acid sequence we see in OmpA of serotype O1 and O1/O5 isolates is related to the virulence of these particular clones. The similarity in OmpA sequence between isolates of serotypes O2, O5, O6 and O7, and the comparative lack of virulence in isolates of these serotypes suggests that this protein may have an important role in the pathogenicity of this organism. Furthermore, the identical OmpA sequences shared by isolates of serotype O1 and O1/O5 is further evidence that the novel O1/O5 isolates have arisen from a serotype O1 strain. The homogeneity of OmpA is displayed as only three distinct alleles were identified amongst the 16 representative strains (Fig. 5.10). These alleles were associated more with serotypes than OMP-types or host species. All serotype O1 and O1/O5 isolates were represented by *ompA.1*, while serotypes O2, O5 and O6 were represented by *ompA.2*. The serotype O7 isolate recovered from European eel was of *ompA.3*. This suggests that OmpA may have more of a role in pathogenesis than host specificity, as isolates of serotype O1 and O1/O5 are recovered more often in rainbow trout and Atlantic salmon than serotypes O2, O5, O6 or O7. Serotypes O1 and O1/O5 are the predominant clones recovered from rainbow trout and Atlantic salmon, respectively. Serotype O1 has been shown to be highly virulent in rainbow trout and Atlantic salmon (Davies, 1991a; Haig *et al.*, 2011), while serotype O1/O5 was shown to be highly virulent towards Atlantic salmon through i.p. challenge, previously (2.2.5.5).

Conversely to OmpA, substantial variation was identified in the major porin OmpF. The major porin OmpF can be altered by several environmental factors including osmolarity, temperature, pH, nutrient availability, aeration and various toxins (Forst *et al.*, 1989; Liu & Ferenci, 2001; Matsubara *et al.*, 2000; Pratt *et al.*, 1996). OmpF is under the control of a two component system (with

OmpC); the histidine kinase EnvZ, and the response regulator OmpR. The larger pore size of OmpF (in relation to OmpC) is attributed to improving efficiency of nutrient uptake in a nutritionally poor environment (Pratt *et al.*, 1996; Yoshida *et al.*, 2006).

Isolates grouped together on the phylogenetic tree (Fig. 5.12) in terms of host species and serotype. All isolates of serotype O1 and O1/O5 recovered from rainbow trout were similar. However, the biotype 2 isolates had a 39 base pair region from position 682 to 720 that was not present in the biotype 1 genomes examined (Fig. 5.13). This region occurs in loop seven of the predicted secondary structure of OmpF (Fig. 5.14). Calculation of the d_S/d_N ratio of all loops in OmpF (0.42) suggests that these regions are particularly susceptible to mutation through natural selection. The conservation of loop 3 was in agreement with the findings in other members of the *Enterobacteriaceae* (Luckey, 2014). This loop is internal to the ß-barrel and plays an important role in constricting the channel, resulting in the selectivity of the pore. Seven alleles were identified in *ompF* amongst the 16 strains examined. These displayed a greater degree of host and OMP-type specificity, compared to OmpA. The major rainbow trout alleles were *ompF.1* and *ompF.3*, which represented the motile OMP-type 3 ('Hagerman') and non-motile OMP-type 1 clones, respectively. The major Atlantic salmon groups were represented by the allele's *ompF.2*, *ompF.5*, *ompF.6* and *ompF.7*, although only single isolates were associated with *ompF.6* and *ompF.7*, representing OMP-types 2c and 2a, respectively. This suggests that virulent isolates recovered from rainbow trout are distinguishable based on the sequence of OmpF, which could form the basis of a future diagnostic test. Further isolates should be examined to determine the reliability of this trait in isolates.

The proteins PilP, PilQ and PilV (involved in pilus formation [Boyd *et al.*, 2008]) were only identified in isolates only of serotype O1 and O1/O5. Isolates of serotype O1 are highly virulent towards both rainbow trout and Atlantic salmon, while the virulence of the novel O1/O5 serotype towards both species requires further examination. Type IV pili have been shown to have roles in twitching motility (Wall & Kaiser, 1999), biofilm formation (O'Toole & Kolter, 1998), attachment (Pujol *et al.*, 1999) and HGT (Filloux, 2010). In twitching motility the ends of the pili attach to a surface or to other bacteria, contract, and pull

the bacteria forward (Wall & Kaiser, 1999). Many pathogenic and non-pathogenic bacterial species are known to produce type IV pili including *N. gonorrhoeae*, *N. meningitidis*, *P. aeruginosa*, *L. pneumophila*, enteropathogenic and enterohemorrhagic *E. coli*, and *V. cholerae* (Mattick, 2002). The major subunit of type IV pili is the protein PilA. This protein was identified in all isolates examined except the serotype O1 and O1/O5 Atlantic salmon strains. The PilA protein of isolates recovered from rainbow trout of serotype O1 and O1/O5 were identical to each other. However, there was substantial variation in PilA compared to isolates of other serotypes. The serotype O5 Atlantic salmon isolates, RD290 and RD366, had identical PilA sequences, although that of the rainbow trout isolate RD28 was very different. It is interesting that isolates that did not encode the other proteins involved in pilus production (PilP, PilQ or PilV) should still encode PilA. It is possible that within these isolates, the PilA gene has not been lost but the ability to express functional pili has.

Yersinia ruckeri is known to encode a T4SS system (Méndez *et al.*, 2009) (the *tra* operon; section 1.4.3.4.2), which is absent from the human pathogenic *Yersinia* species. The proteins TraL and TraN were predicted to be OM located (Table 3.2). Interestingly, these proteins were only identified in strains of serotype O1 and O1/O5 in both rainbow trout and Atlantic salmon. The suspected roles in virulence of this secretion system has been well documented previously (Méndez *et al.*, 2009), and its absence from the strains of serotype O2, O5, O6 and O7, which are not particularly virulent, is supportive of its role in virulence.

Invasin is a protein that allows enteric bacteria to penetrate host cells, causing the cell to polymerise and depolymerise actin filaments. The rearrangement of the cytoskeleton enables phagocytes and non-phagocytic cells to engulf bacteria into a vacuole. The bacteria subsequently escape the vacuole and replicate in the cytoplasm of the host cell. However, other bacteria including *Salmonella* remain inside the vacuole allowing evasion of the immune response (Steele-mortimer, 2009). In this study, two putative invasins were identified through bioinformatic prediction. However, these proteins were not identified in the genomes of all 16 isolates. Invasins have been shown to play important roles in the pathogenesis of other members of the *Yersinia*, including *Y. enterocolitica* (Uliczka *et al.*, 2011), and *Y. pseudotuberculosis* (Uliczka *et al.*, 2009).

Invasin 1 was present only in O1 and O1/O5 isolates in rainbow trout and Atlantic salmon, and in the serotype O7 European eel isolate, RD150, while invasin 2 was absent from the serotype O5 rainbow trout strain, RD28, the serotype O6 Atlantic salmon isolate, RD162, and the serotype O7 European eel isolate, RD150. In the isolates in which they were present, both invasins were highly conserved, displaying only single nucleotide polymorphisms. Since invasin 1 was present only in serotype O1 and O1/O5 isolates, it may be particularly important in their pathogenicity. Invasin 2 was present in most strains, although absent from the serotype O5 rainbow trout strain, RD28, the serotype O6 Atlantic salmon strain, RD162, and the serotype O7 strain from European eel, RD150. In the study by Davies (1991a) examining the virulence of selected isolates towards rainbow trout, isolates RD28, RD162 and RD150 were not particularly virulent. This suggests that invasin 2 may play a role in the virulence of this bacterium also.

Lipoprotein 1 was not identified in any of the serotype O1 or O1/O5 isolates from either host species. While the function of this protein remains unknown, it was highly conserved in the remaining isolates. This protein is unlikely to be host specific or play a crucial role in virulence, as it was identified in isolates recovered from both rainbow trout and Atlantic salmon, and in isolates that are known to be avirulent, and has been lost from the pathogenic serotype O1 and O1/O5 isolates.

The phosphoporin OmpE is involved in the uptake of inorganic phosphate, phosphorylated compounds, and some other negatively charged solutes (Koebnik *et al.*, 2000). OmpE was absent in isolates RD520 and RD524. These represent rainbow trout isolates that have been recovered more recently (2011). This protein was present in the older rainbow trout strains (pre 1990) and all Atlantic salmon strains. It is possible that OmpE is no longer required in *Y. ruckeri* and has been lost from the rainbow trout strains more recently due to the longer association of the bacteria with this host.

Through phylogenetic relatedness it has been shown that isolates recovered from Atlantic salmon and rainbow trout represent distinct phylogenetic lineages from each other, suggesting a host-associated evolutionary split within this bacterial species. Isolates of serotypes O1 and O1/O5 are very similar to each other,

supporting the original hypothesis that serotype O1/O5 isolates have emerged from serotype O1 strains. The majority of OMPs that are present in serotype O1 strains are also present in the serotype O1/O5 isolates, while proteins absent from the O1 strains are also missing in the O1/O5 isolates. Furthermore, certain proteins present in serotype O2, O5, O6 and O7 strains are not found in the O1 or O1/O5 isolates either.

In conclusion, isolates recovered from Atlantic salmon and rainbow trout represent distinct phylogenetic lineages based on the concatenated sequences of 19 housekeeping enzymes. The orignial hypothesis that the novel serotype O1/O5 has arisen as a consequence of HGT event between Atlantic salmon and rainbow trout is supported, as this serotype is present on both the Atlantic salmon and rainbow trout lineages of the phylogenetic tree. Analysis of the LPS-biosynthesis gene cluster has shown that serotypic diversity is not a result of variation within this operon, and suggests that it may be due to glycosylation genes found elsewhere in the genome. High levels of conservation have been identified in the majority of predicted OMPs, suggesting that *Y. ruckeri* is indeed a highly genetically homogenous species. However, several proteins were not encoded in the genomes of all isolates examined. This study should form the basis of future research whereby the precise role of these proteins in pathogenicity of both Atlantic salmon and rainbow trout is deduced and further examination on the basis for serotypic diversity in *Y. ruckeri* is explored.

Chapter 6 Final discussion

Yersinia ruckeri is the causative agent of ERM disease in salmonids. Infections caused in Atlantic salmon and rainbow trout are of particular concern due to the high economic value placed on these species. The aim of this study was to utilise comparative phenotypic, proteomic and genomic approaches to further our understanding of the diversity of *Y. ruckeri* isolates recovered from Atlantic salmon and rainbow trout, in the UK.

The LPS and OMP diversity of an extensive and unique strain collection comprising 135 isolates of *Y. ruckeri*, 109 of which were recovered from Atlantic salmon and 26 from rainbow trout, was assessed. These isolates were supplemented with representative isolates from an earlier study for which further information regarding the virulence characteristics was also available (Davies, 1991a). Isolates representative of the major clonal groups infecting both Atlantic salmon and rainbow trout were selected and subjected to proteomic analyses to examine variation in OM proteome composition. Subsequently, the evolutionary relationships between 16 isolates representative of different biotypes, serotypes and OMP-types recovered from Atlantic salmon and rainbow trout were examined. Variation in the nucleotide and amino acid sequences of the genes encoding OMPs was examined in order to identify proteins which may be specific to certain clonal groups, or have regions with potential roles in host specificity.

The main disease causing isolates infecting UK rainbow trout are of biotype 2, serotype O1 and OMP-type 1a. This clonal group has been the major cause of infection in rainbow trout in the UK since these isolates were first described in the late 1980's (Austin *et al.*, 2003; Davies & Frerichs, 1989). This clonal group has since been recovered from elsewhere in the world including the USA (Arias *et al.*, 2007), Chile (Bastardo *et al.*, 2011), Australia (Carson & Wilson, 2009), Spain (Fouz *et al.*, 2006) and other European countries (Wheeler *et al.*, 2009).

The main disease-causing strain affecting Atlantic salmon is biotype 1, serotype O1/O5 and OMP-type 3a, and represents 53% of isolates recovered from Atlantic salmon in this study. The emergence of the novel O1/O5 serotype was a significant finding of this study. This LPS-type comprised a core polysaccharide

region identical to that of serotype O1 isolates combined with a unique O-antigen subunit (identified through silver stained SDS-PAGE and Western-blot analysis). While this serotype is representative of the major LPS-type of isolates currently found in Atlantic salmon, it was also recovered from rainbow trout in 2010 (one isolate) and 2011 (one isolate). This is of significance as the pathogenicity of this isolate towards rainbow trout is yet to be determined. The protection afforded by current commercial vaccines which are whole cell vaccines prepared using serotype O1, biotype 1 and 2 variants, must be determined. While some cross protection may be afforded by these vaccines against isolates of different serotypes, this study has shown that isolates of serotypes O1, O2, O5 and O1/O5 have different LPS profiles, and that each serotype is primarily associated with a different OMP-type.

Overall, the isolates infecting Atlantic salmon and rainbow trout are different. While rainbow trout isolates were almost exclusively represented by the same non-motile biotype 2 clone, isolates recovered from Atlantic salmon were represented by four different serotypes and three different OMP-types. These OMP-types could be subdivided further, based on variation in minor proteins, into 11 OMP sub-types. This suggests that as LPS and OMPs are potentially important antigenic determinants (Koppe & Harms, 1994), different vaccine formulations may be necessary to provide adequate protection against different clonal groups of *Y. ruckeri*. However, bath challenge studies of Atlantic salmon and rainbow trout were not completely successful. Repeating the challenge studies conducted in this work to first determine the pathogenicity of serotype O1/O5 isolates towards both species, before examining the protection afforded by the current commercial vaccines against this strain, is essential.

A thorough bioinformatic prediction study identified 141 proteins that were predicted to be OM located within the genomes of four isolates of *Y. ruckeri* (three from rainbow trout and one from Atlantic salmon). The four genomes examined encoded an average of 3406 proteins. Thus, the predicted OM proteome equates to 4.1% of the entire genome. Upon comparison to other species including *P. multocida* (E-komon *et al.*, 2012) and *E. ruminantium* (Moumène *et al.*, 2015) of which the predicted OM proteome represented 4.8 and 5.8% of the genomes, respectively, *Y. ruckeri* appears to encode relatively

fewer OMPs. Forty eight proteins were identified as the predicted core OM proteome, which represented 34% of the total predicted OMPs.

Subsequently, the OM proteome of eight representative isolates grown aerobically (22°C) was elucidated using complementary gel-based and gel-free proteomic techniques. This resulted in the confident identification of 70 OMPs, of which 52 had been predicted (74.3%). The OM proteome of isolates grown aerobically (22°C) was compared to those of isolates grown under conditions which mimicked the *in vivo* and *in vitro* environments that *Y. ruckeri* may encounter. These included examining cells grown anaerobically, aerobically at 28°C, under iron-depleted conditions and in an artificial seawater medium. In total, 76 OMPs were identified in eight isolates grown under five distinct growth conditions by gel-free proteomics (triplicate samples). Proteins which appeared to be host, strain and/or serotype specific were identified. These may have important roles in the pathogenesis of this bacterium. These included proteins involved in OM biogenesis and cell integrity (OmpA, Braun's lipoprotein [Lpp] and SlyB), transport (OmpF, OmpC, FhuA, ShuA, FiuA and HemR), adherence (OmpA, OmpX and PqiB/Mam7), enzymatic activity (phospholipase A1), proteins with other roles (CsgG, Invasin 1, Invasin 2) and hypothetical proteins. However, the majority of proteins which were identified were common to all strains. It became clear that conducting replicate experiments and combining both gel-based and gel-free proteomic techniques was important to ensure optimal protein recovery and identification. Of the few studies that have combined both approaches, those that have done so reported overlapping coverage and, importantly, proteins identified solely by individual techniques (Aistleitner *et al.*, 2015; Cordwell *et al.*, 2008; Gesslbauer *et al.*, 2012; Kulkarni *et al.*, 2014; Tanca *et al.*, 2013; Zhou *et al.*, 2012).

One previous study examined the OM proteome of *Y. ruckeri,* comparing OMP expression under stationary growth with that of cells immobilised in an artificial biofilm (Coquet *et al.*, 2005). In total, 55 OMPs were identified in this study. However, the data was searched on NCBI against the genomes of *Y. pestis*, *E. coli* and *S. typhimurium*. Only a single isolate was analysed by gel-based proteomic approaches with only differences in protein abundance noted. Thus, this study was not optimal for the characterisation of *Y. ruckeri* OMPs.

Examining both the nucleotide and amino acid sequence variation amongst the genes encoding the 141 predicted OMPs of *Y. ruckeri*, supported previous studies which highlighted the homogeneity of the bacterium (Bastardo *et al.*, 2012a; Kotetishvili *et al.*, 2005; Schill *et al.*, 1984; Wheeler *et al.*, 2009). Of the 141 proteins that had been predicted, 130 were present in the genomes of all 16 representative isolates examined. Of these, the majority were well conserved. However, variation in several important proteins including OmpA and OmpF was identified. These variations may account for virulence properties or host specific adaptations of certain isolates. Within isolates of serotype O1 and O1/O5, the sequences of OmpA were identical to each other. However, the OmpA sequences in isolates of serotype O2, O5 and O6 were different from those of the serotype O1 and O1/O5 isolates. As the serotype O1 and O1/O5 isolates represent the hyper-virulent clones in rainbow trout and Atlantic salmon, respectively, OmpA may play an important role in virulence. Roles in adherence, invasion, intracellular survival, toxicity, pro-inflammatory responses and evasion of the host defence system have been described in many bacteria (Confer & Ayalew, 2013). However, further work is necessary to confirm if OmpA has a role in these in *Y. ruckeri*.

Extensive variation was identified in loop 5 of the major porin protein, OmpF. A 39 bp region was identified which allowed biotype 2 isolates to be distinguished from those of biotype 1. Upon comparison to a phylogenetic tree produced using the concatenated sequence of 19 housekeeping enzymes, it is suggested that this region has been inserted into the biotype 2 isolates rather than lost from the biotype 1 isolates. This information could prove useful in the development of improved rapid diagnostic methods for the identification of non-motile variants of *Y. ruckeri*.

Eleven predicted OMPs were not identified in all 16 genomes examined; this included two invasins (invasin 1 and invasin 2). Invasin 1 was present only in the serotype O1 and O1/O5 isolates, thus suggesting that this protein may play an important role in the virulence of *Y. ruckeri*. Proteins involved in pilus biogenesis were also only identified in isolates of serotype O1 and O1/O5, in both Atlantic salmon and rainbow trout. This suggests that initial attachment to host cells using pili may be an important factor in the pathogenicity of *Y.*

ruckeri. This has been seen in other fish pathogenic bacteria, including *A. hydrophila* (Lallier & Daigneault, 1984) and *A. salmonicida* (Boyd *et al.*, 2008).

The work conducted in this thesis has provided important information as to the diversity of *Y. ruckeri* isolates currently circulating within the Atlantic salmon and rainbow trout populations. Also, the variety of proteins encompassing the OM proteome and the conditions under which they are upregulated has been explored.

Biotyping and serotyping are two of the most important phenotypic characteristics that need to be determined for epidemiological surveillance, outbreak assessment and the design of vaccination strategies. However, biotyping and serotyping are time-consuming processes taking many days; furthermore, serotyping can be performed only in a limited number of laboratories that possess the required reference antisera. For these reasons, the development of a rapid diagnostic tool that is able to rapidly confirm species identity and identify the biotypes and serotypes of clinical specimens would be beneficial. Real time (RT) PCR is now widely used in clinical and veterinary microbiology. Furthermore, the accessibility of different fluorescent probes has allowed multiplex RT PCR approaches to be developed to distinguish between different species and strains (Fratamico *et al.*, 2011; Pinsky & Banaei, 2008; Reddington *et al.*, 2011). Indeed, this method has been developed for detecting various pathogens in fish tissue, including *Y. ruckeri* and *F. psychrophilum* (Bastardo *et al.*, 2012c; Glenn *et al.*, 2011; Keeling *et al.*, 2012; Strepparava *et al.*, 2014). Welch *et al.* (2011) demonstrated that loss of motility in *Y. ruckeri* is caused by mutations or deletions in the flagella biosynthesis/export genes *fliR*, *flhA* and *flhB*. Thus, with the availability of genomic sequence data from multiple strains of *Y. ruckeri*, a multiplex RT PCR assay could be designed that makes use of these genomic mutations and deletions.

The genes encoding proteins involved in O-antigen biosynthesis are located on an O-antigen gene cluster (the *rfb* cluster) (Samuel & Reeves, 2003). Standard PCR-based methods, based on variation of specific genes within the O-antigen gene cluster, have been developed for the rapid determination of O-genotypes for various bacterial species including *Salmonella enterica* (Fitzgerald *et al.*, 2003), *Leptospira* (Cai *et al.*, 2010), and *Y. pestis* and *Y. pseudotuberculosis*

(Bogdanovich *et al.*, 2003). However, multiplex RT PCR can be used to differentiate between different O-antigen types within bacterial species (Debroy *et al.*, 2005). With the availability of genomic sequence data from multiple *Y. ruckeri* strains, this approach could easily be adopted for the development of a multiplex RT PCR assay to differentiate between the six O-serotypes of *Y. ruckeri*. However, based on the relative homogeneity of the LPS biosynthesis genes examined in this study, further analysis to identify serotype specificity is required.

As several proteins with potential roles in the pathogenesis of *Y. ruckeri* have been identified (discussed earlier), this should prompt future vaccine development studies. Examination of the efficacy of whole cell vaccines prepared using cells grown under anaerobic, aerobic (28°C), iron-depleted and artificial seawater medium conditions in comparison to vaccines prepared with isolates grown under aerobic (22°C) conditions could improve vaccine design. Rainbow trout immunised against furunculosis (caused by *A. salmonicida*) using a whole cell vaccine grown under iron-depleted conditions, showed enhanced protection (Durbin *et al.*, 1999). It would also be interesting to examine the variations in LPS expression and profile under the variant conditions. In *Y. enterocolitica*, the temperature at which the cells were grown influenced the structure of the LPS molecule. Cells grown at lower temperatures had a smooth type LPS, while cells grown at higher temperatures had a partial smooth-rough LPS (Kawaoka *et al.*, 1983).

Similarly, growing the bacteria under a further range of conditions and examining the OM proteome could reveal more information as to the diversity of isolates recovered from both Atlantic salmon and rainbow trout, and the role of OMPs both inside and outside of the host. Particularly, growing the bacteria in serum, bile salts or mixed microbial communities would have been highly relevant. Serum factors include complement, antibodies, hormones and other host proteins which may interact with the bacterial cell surface. Serum components have been demonstrated to interact with and modulate the expression of OMPs in other bacteria including *E. coli* and *Actinobacillus actinomycetemcomitans* (Hellman *et al.*, 2000; Johansson *et al.*, 2003). Growth in the presence of bile can have a dramatic effect on the OM of bacteria. For example, in *Salmonella sp.* genes involved in TTSS are downregulated as the

bacteria senses bile in the intestine. However, as the bacteria moves through the mucus layer and reaches the epithelial cell surface, the bile concentration decreases and the invasion machinery is upregulated (Prouty & Gunn, 2000). Conversely, in *Shigella sp.* genes involved in TTSS are upregulated in the presence of bile (Pope *et al.*, 1995). In *V. parahaemolyticus*, a hemolysin is upregulated in the presence of bile (Miyamoto *et al.*, 1969), while in *V. cholerae*, cholera toxin and toxin co-regulated pili are down-regulated (Gunn, 2000; Schuhmacher & Klose, 1999).

Examination of the presence or absence of the predicted OMPs in a much larger range of strains would be useful, as it could reveal more information as to the host specificity of these proteins and in turn relate to the pathogenic capabilities of individual isolates. Similarly, mining the entire genome could reveal important differences in isolates pathogenic towards Atlantic salmon and rainbow trout, and may be important in future vaccination and control strategies as reverse vaccinology (Rappuoli, 2000) becomes a more appropriate technique. Reverse vaccinology is a bioinformatic technique that has been applied to several pathogens, including *N. meningitidis*, *S. agalactiae, S. pyogenes, S. pneumoniae* and *E. coli* in order to identify proteins that may have roles in pathogenicity (Seib *et al.*, 2012).

Creating knockout mutant strains deficient in proteins that have suggested roles in the pathogenicity of *Y. ruckeri*, would allow us to understand the pathogenic properties of isolates in greater depth. While this work could be conducted using live animals, the constant strive to 'reduce, replace and refine' would support the development of a model which would reduce the cost and ethical implications involved with using animals. The use of an infection model would allow the proteins involved in host interaction to be elucidated and studied in much greater detail. In other Gram-negative bacteria, *in vivo* proteomics have been used in order to ascertain a greater understanding of host-pathogen interactions. For example, in *E. coli*, dialysis tubing was used to examine the bacterial proteome of isolates grown in the bovine rumen, with adaptive bacterial responses observed (Kudva *et al.*, 2014); the OM proteome of *P. multocida* recovered from the bloodstream of infected chickens has been examined (Boyce *et al.*, 2006); the OMPs upregulated in *A. salmonicida* under *in vivo* growth conditions have been examined utilising an *in vivo* growth chamber

model (Ebanks *et al.*, 2004). As the numbers of bacteria obtained from *in vivo* cultures may be small, conventional proteomic techniques may not be appropriate for use. Filter-aided sample preparation (FASP) (Wisniewski & Zougman, 2009) is a gel-free proteomic technique which is applicable to a much lower starting cell number (Sharma *et al.*, 2012) than the gel-based and gel-free techniques employed in this thesis. Filter-aided sample preparation has been employed for protein identification in several species, including *E. coli* (Sharma *et al.*, 2012; Tanca *et al.*, 2013), *C. jejuni* (Jang *et al.*, 2014), *Shewanella oneidensis, Shewanella putrefaciens* and *Pseudomonas putida* (Sharma *et al.*, 2012).

Utilising *in vivo* proteomics to examine the OMPs upregulated during infection would also be an appropriate technique, as demonstrated by Poobalane *et al.*, (2008) who implanted dialysis tubing containing *A. hydrophila* into the peritoneal cavity of common carp (*Cyprinus carpio L.*). Through this approach the authors were able to identify OMPs, whole cell proteins and ECPs.

In conclusion, the work presented in this thesis represents a study of *Y. ruckeri* isolates causing infection in Atlantic salmon and rainbow trout, in the UK. The diversity of isolates causing infection in these two hosts has been observed, with clonal groups specific to each species identified. The identification of the novel O1/O5 LPS-type, predominant in Atlantic salmon is highly significant; the recovery of this serotype from rainbow trout in more recent years is also highly relevant. Utilising bioinformatic analyses in combination with gel-free and gel-based proteomic approaches has allowed the OM proteome of *Y. ruckeri* to be elucidated, and a much greater number of OMPs to be identified than using these techniques individually. The identification of OMPs upregulated under conditions which specifically mimic the *in vivo* and *in vitro* environment of *Y. ruckeri* has provided enhanced information as to how the bacteria may elicit infection. Through sequence analysis, it was observed that many of the OMPs were well conserved. However, 11 OMPs were not present in all genomes examined. Further work is necessary to identify the role of these OMPs in the pathogenesis of *Y. ruckeri*.

This project has addressed important questions' regarding the diversity of isolates of the fish pathogen *Y. ruckeri* infecting Atlantic salmon and rainbow

trout, and in turn paved the way for future studies utilising the information generated in this study as a starting point.

Chapter 7 References

Adessi, C., Miege, C., Albrieux, C. & Rabilloud, T. (1997). Two-dimensional electrophoresis of membrane proteins: A current challenge for immobilized pH gradients. *Electrophoresis* **18**, 127–135.

Aebersold, R. & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature* **422**, 198–207.

Aistleitner, K., Anrather, D., Schott, T., Klose, J., Bright, M., Ammerer, G. & Horn, M. (2015). Conserved features and major differences in the outer membrane protein composition of *Chlamydiae*. *Environ Microbiol* **17**, 1397–1413.

Akhlaghi, M. & Yazdi, S. H. (2008). Detection and identification of virulent *Yersinia ruckeri* : the causative agent of enteric redmouth disease in rainbow trout (*Oncorhynchus mykiss*) cultured in Fars province, Iran. *Iran J Vet Res* **9**.

Akira, S., Uematsu, S. & Takeuchi, O. (2006). Pathogen recognition and innate immunity. *Cell* **124**, 783–801.

Alexander, C. & Rietschel, E. T. (2001). Bacterial lipopolysaccharides and innate immunity. *J Endotoxin Res* **7**, 167–202.

Ali, Q., Davies, R. L., Parton, R., Coote, J. G. & Gibbs, H. A. (1992). Lipopolysaccharide heterogeneity in *Pasteurella haemolytica* isolates from cattle and sheep. *J Gen Microbiol* **138**, 2185–95.

Allan, B. J. & Stevenson, R. M. (1981). Extracellular virulence factors of *Aeromonas hydrophila* in fish infections. *Can J Microbiol* **27**, 1114–1122.

Allen, W. J., Phan, G. & Waksman, G. (2012). Pilus biogenesis at the outer membrane of Gram-negative bacterial pathogens. *Curr Opin Struct Biol* **22**, 500–6. Elsevier Ltd.

Allison, G. E. & Verma, N. K. (2000). Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. *Trends Microbiol* 17–23.

Alteri, C. J. & Mobley, H. L. T. (2007). Quantitative profile of the Uropathogenic *Escherichia coli* outer membrane proteome during growth in human urine. *Infect Immun* **75**, 2679–2688.

Altindis, E., Fu, Y. & Mekalanos, J. J. (2014). Proteomic analysis of *Vibrio cholerae* outer membrane vesicles. *Proc Natl Acad Sci U S A* **111**, E1548–56.

Altinok, I. (2004). The infectious route of *Yersinia ruckeri* is affected by salinity. *Bull Eur Assoc Fish Pathol* **24**, 253–259.

Altinok, I. & Grizzle, J. M. (2001). Effects of low salinities on *Flavobacterium columnare* infection of euryhaline and freshwater stenohaline fish. *J Fish Dis* **24**, 361–367.

Altinok, I., Grizzle, J. M. & Liu, Z. (2001). Detection of *Yersinia ruckeri* in rainbow trout blood by use of the polymerase chain reaction. *Dis Aquat Organ* **44**, 29–34.

Anaya, C., Church, N. & Lewis, J. P. (2007). Detection and identification of bacterial cell surface proteins by fluorescent labeling. *Proteomics* **7**, 215–9.

Andersen-Nissen, E., Smith, K. D., Strobe, K. L., Barrett, S. L. R., Cookson, B. T., Logan, S. M. & Aderem, A. (2005). Evasion of Toll-like receptor 5 by flagellated bacteria. *Proc Natl Acad Sci U S A* **102**, 9247–9252.

Andrews, S. C., Robinson, A. K. & Rodríguez-Quiñones, F. (2003). Bacterial iron homeostasis. *FEMS Microbiol Rev* **27**, 215–237.

Arias, C. R., Olivares-Fuster, O., Hayden, K., Shoemaker, C. A., Grizzle, J. M. & Klesius, P. H. (2007). First report of *Yersinia ruckeri* biotype 2 in the USA. *J Aquat Anim Health* **19**, 35–40.

Austin, B. & Rodgers, C. J. (1982). Morphological diversity among strains of *Yersinia ruckeri*. *Aquaculture* **27**, 73–78.

Austin, B. & Austin, D. A. (1993). *Bacterial Fish Pathogens: Diseases in farmed and wild fish*, 2nd edn. (B. Austin & D. A. Austin, Eds.). Horwood, Ellis.

Austin, B. & Austin, D. A. (2007). *Bacterial fish pathogens: Diseases of farmed and wild fish*, 4th edn. Springer-Praxis.

Austin, D. A., Robertson, P. A. W. & Austin, B. (2003). Recovery of a new biogroup of *Yersinia ruckeri* from diseased rainbow trout (*Oncorhynchus mykiss*, Walbaum). *Syst Appl Microbiol* **131**, 127–131.

Avci, H. & Birincioglu, S. S. (2005). Pathological findings in rainbow trout (*Oncorhynchus mykiss*, Walbaum, 1792) experimentally infected with *Yersinia ruckeri*. *Turkish J Vet Anim Sci* **29**, 1321–1328.

Ayalew, S., Confer, A. W., Hartson, S. D. & Shrestha, B. (2010). Immunoproteomic analyses of outer membrane proteins of *Mannheimia haemolytica* and identification of potential vaccine candidates. *Proteomics* **10**, 2151–64.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M. & other authors. **(2008).** The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9**, 75.

Bachrach, G., Zlotkin, A., Hurvitz, A., Evans, D. L. & Eldar, A. (2001). Recovery of *Streptococcus iniae* from diseased fish previously vaccinated with a *streptococcus* vaccine. *Appl Environ Microbiol* **67**, 3756–8.

Bagos, P. G., Liakopoulos, T. D., Spyropoulos, I. C. & Hamodrakas, S. J. (2004). PRED-TMBB: a web server for predicting the topology of beta-barrel outer membrane proteins. *Nucleic Acids Res* **32**, W400–4.

Bai, J., Kim, S. I., Ryu, S. & Yoon, H. (2014). Identification and characterization of outer membrane vesicle-associated proteins in *Salmonella enterica* serovar *Typhimurium*. *Infect Immun* **82**, 4001–10.

Baik, S., Kim, K., Song, S. & Kim, D. (2004). Proteomic analysis of the sarcosine-insoluble outer membrane fraction of *Helicobacter pylori* strain 26695. *J Bacteriol* **186**, 949–955.

Barnhart, M. M. & Chapman, M. R. (2006). Curli biogenesis and function. *Annu Rev Microbiol* **60**, 131–147.

Barrow, G. I. & Fektham, R. K. A. (1993). Cowan and Steel's manual for the identification of medical bacteria. In *Cowan Steel's Man Identif Med Bact*, p. 329.

Bastardo, A., Bohle, H., Ravelo, C., Toranzo, A. E. & Romalde, J. L. (2011). Serological and molecular heterogeneity among *Yersinia ruckeri* strains isolated from farmed Atlantic salmon (*Salmo salar*) in Chile. *Dis Aquat Organ* **93**, 207–14.

Bastardo, A., Ravelo, C. & Romalde, J. L. (2012a). Multilocus sequence typing reveals high genetic diversity and epidemic population structure for the fish pathogen *Yersinia ruckeri*. *Environ Microbiol* **14**, 1888–97.

Bastardo, A., Ravelo, C. & Romalde, J. L. (2012b). A polyphasic approach to study the intraspecific diversity of *Yersinia ruckeri* strains isolated from recent outbreaks in salmonid culture. *Vet Microbiol* **160**, 176–82. Elsevier B.V.

Bastardo, A., Ravelo, C. & Romalde, J. L. (2012c). Highly sensitive detection and quantification of the pathogen *Yersinia ruckeri* in fish tissues by using real-time PCR. *Appl Microbiol Biotechnol* **96**, 511–20.

Basu, M., Swain, B., Maiti, N. K., Routray, P. & Samanta, M. (2012). Inductive expression of toll-like receptor 5 (TLR5) and associated downstream signaling molecules following ligand exposure and bacterial infection in the Indian major carp, mrigal (*Cirrhinus mrigala*). *Fish Shellfish Immunol* **32**, 121–31. Elsevier Ltd.

Batchelor, E. & Walthers, D. (2005). The *Escherichia coli* CpxA-CpxR envelope stress response system regulates expression of the porins ompF and ompC. *J Bacteriol* **187**, 5723–5731.

Beeckman, D. S. A. & Vanrompay, D. C. G. (2010). Bacterial secretion systems with an emphasis on the chlamydial Type III secretion system. *Curr Issues Mol Biol* **12**, 17–41.

Begic, S. & Worobec, E. A. (2006). Regulation of *Serratia marcescens* ompF and ompC porin genes in response to osmotic stress, salicylate, temperature and pH. *Microbiology* **152**, 485–491.

Beher, M. G., Schnaitman, C. A. & Pugsley, A. P. (1980). Major heat-modifiable outer membrane protein in gram-negative bacteria: comparison with the ompA protein of *Escherichia coli*. *J Bacteriol* **143**, 906–13.

Beloin, C., Roux, A. & Ghigo, J. M. (2008). *Escherichia coli* biofilms. In *Bact Biofilms*, pp. 249–289. Edited by T. Romeo.

Berc, A., Petrinec, Z., Matasin, Z. & Kozarić, Z. (1999). *Yersinia ruckeri* septicaemia in experimentally infected carp (*Cyprinus carpio L.*) fingerlings. *Acta Vet Hung* **47**, 161–172.

Berven, F. S., Flikka, K., Jensen, H. B. & Eidhammer, I. (2004). BOMP: A program to predict integral B-barrel outer membrane proteins encoded within genomes of Gram-negative bacteria. *Nucleic Acids Res* **32**, 394–399.

Berven, F. S., Karlsen, O. A., Straume, A. H., Flikka, K., Murrell, J. C., Fjellbirkeland, A., Lillehaug, J. R., Eidhammer, I. & Jensen, H. B. (2006). Analysing the outer membrane subproteome of *Methylococcus capsulatus* (Bath) using proteomics and novel biocomputing tools. *Arch Microbiol* **184**, 362–377.

Blazer, V. S. (1991). Piscine macrophage function and nutritional influences: A review. *J Aquat Anim Health* **3**, 77–86.

Bogdanovich, T., Carniel, E., Fukushima, H. & Skurnik, M. (2003). Use of O-antigen gene cluster-specific PCRs for the identification and O-genotyping of *Yersinia pseudotuberculosis* and *Yersinia pestis*. *J Clin Microbiol* **41**, 5103–5112.

Bootland, L. M. & Leong, J. C. (1999). Infectious hematopoietic necrosis virus. In *Fish Dis Disord Vol 3; viral, Bact fungal Infect*, pp. 57–121. Edited by P. T. K. Woo & D. W. Bruno. CABI PUBLISHING.

Boyce, J. D., Cullen, P. A., Nguyen, V., Wilkie, I. & Adler, B. (2006). Analysis of the *Pasteurella multocida* outer membrane sub-proteome and its response to the in vivo environment of the natural host. *Proteomics* **6**, 870–80.

Boyd, J. M., Dacanay, A., Knickle, L. C., Touhami, A., Brown, L. L., Jericho, M. H., Johnson, S. C. & Reith, M. (2008). Contribution of type IV pili to the virulence of *Aeromonas salmonicida* subsp. *salmonicida* in Atlantic salmon (*Salmo salar* L.). *Infect Immun* **76**, 1445–1455.

Brambilla, L., Morán-Barrio, J. & Viale, A. M. (2014). Expression of the *Escherichia coli* ompW colicin S4 receptor gene is regulated by temperature and modulated by the H-NS

and StpA nucleoid-associated proteins. *FEMS Microbiol Lett* **352**, 238–244.

Braun, V., Ondraczek, R. & Hobbie, S. (1993). Activation and secretion of *Serratia* hemolysin. *Int J Med Microbiol* **278**, 306–15.

Bravo, S. & Kojagura, V. (2004). First isolation of *Yersinia ruckeri* from rainbow trout (*Oncorhynchus mykiss*) in Peru. *Bull Eur Assoc Fish Pathol* **24**, 104–108.

Bruno, D. W. & Munro, A. L. S. (1989). Immunity in Atlantic salmon, *Salmo salar L.*, fry following vaccination against *Yersinia ruckeri*, and the influence of body weight and infectious pancreatic necrosis virus (IPNV) on the detection of carriers. *Aquaculture* **81**, 205–211.

Buján, N., Hernández-Haro, C., Monteoliva, L., Gil, C. & Magariños, B. (2015). Comparative proteomic study of *Edwardsiella tarda* strains with different degrees of virulence. *J Proteomics*. Elsevier B.V.

Bullock, G. L., Stuckley, H. M. & Shotts, E. B. (1978). Enteric redmouth bacterium : comparison of isolates from different geographic areas. *J Fish Dis* **1**, 351–356.

Burguière, P., Auger, S., Hullo, M., Danchin, A. & Martin-verstraete, I. (2004). Three different systems participate in L-Cystine uptake in *Bacillus subtilis*. *J Bacteriol* **186**, 4875–4884.

Busch, R. A. & Lingg, A. J. (1975). Establishment of an asymptomatic carrier strate infection of Enteric redmouth disease (Hagerman Strain) in rainbow trout (*Salmo giardneri*). *Mar Fish Rev*.

Busch, R. (1978). Enteric redmouth disease (Hagerman strain). *Mar Fish Rev*.

Büttner, D. & Bonas, U. (2002). Port of entry - the type III secretion translocon. *Trends Microbiol* **10**, 186–92.

Cai, C.-S., Zhu, Y.-Z., Zhong, Y., Xin, X.-F., Jiang, X.-G., Lou, X.-L., He, P., Qin, J.-H., Zhao, G.-P. & other authors. **(2010).** Development of O-antigen gene cluster-specific PCRs for rapid typing six epidemic serogroups of *Leptospira* in China. *BMC Microbiol* **10**, 67.

Capkin, E. & Altinok, I. (2009). Effects of dietary probiotic supplementations on prevention／treatment of yersiniosis disease. *J Appl Microbiol* **106**, 1147–1153.

Carranza, P., Grunau, A., Schneider, T., Hartmann, I., Lehner, A., Stephan, R., Gehrig, P., Grossmann, J., Groebel, K. & other authors. **(2010).** A gel-free quantitative proteomics approach to investigate temperature adaptation of the food-borne pathogen *Cronobacter turicensis* 3032. *Proteomics* **10**, 3248–61.

Carson, J. & Wilson, T. (2009). Yersiniosis in fish. *Aust New Zeal Stand Diagnostic Proced* 1–19.

Cascales, E. & Christie, P. J. (2003). The versatile bacterial type IV secretion systems. *Nat Rev Microbiol* **1**, 137–49.

Chai, T. J. & Foulds, J. (1977). Purification of protein A, an outer membrane component missing in *Escherichia coli* K-12 ompA mutants. *Biochim Biophys Acta* **493**, 210–215.

Chandramouli, K. & Qian, P.-Y. (2009). Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. *Hum Genomics Proteomics* **2009**, 22.

Chen, P. E., Cook, C., Stewart, A. C., Nagarajan, N., Sommer, D. D., Pop, M., Thomason, B., Thomason, M. P. K., Lentz, S. & other authors. **(2010).** Genomic characterization of the *Yersinia* genus. *Genome Biol* **11**, R1.

Chen, Y., Stine, O. C., Badger, J. H., Gil, A. I., Nair, G. B., Nishibuchi, M. & Fouts, D. E. (2011). Comparative genomic analysis of *Vibrio parahaemolyticus*: serotype conversion

and virulence. *BMC Genomics* **12**, 294. BioMed Central Ltd.

Chiu, Y., Kuo, T.-Y., Lin, C.-C. & Chen, W.-J. (2011). Proteomic analysis reveals responsive proteins of *Vibrio parahaemolyticus* on exposure to cationic antimicrobial peptides. *J Appl Microbiol* **110**, 80–9.

Chng, S.-S., Ruiz, N., Chimalakonda, G., Silhavy, T. J. & Kahne, D. (2010). Characterization of the two-protein complex in *Escherichia coli* responsible for lipopolysaccharide assembly at the outer membrane. *Proc Natl Acad Sci U S A* **107**, 5363–5368.

Choi, D.-S., Kim, D.-K., Choi, S. J., Lee, J., Choi, J.-P., Rho, S., Park, S.-H., Kim, Y.-K., Hwang, D. & Gho, Y. S. (2011). Proteomic analysis of outer membrane vesicles derived from *Pseudomonas aeruginosa*. *Proteomics* **11**, 3424–9.

Chooneea, D., Karlsson, R., Encheva, V., Arnold, C., Appleton, H. & Shah, H. (2010). Elucidation of the outer membrane proteome of *Salmonella enterica* serovar *Typhimurium* utilising a lipid-based protein immobilization technique. *BMC Microbiol* **10**, 44.

Chowdhury, R., Sahu, G. K. & Das, J. (1996). Stress response in pathogenic bacteria. *J Biosci* **21**, 149–160.

Cipriano, R. C., Ruppenthal, T., Schill, W. B., Pyle, S. W. & Shotts, E. B. (1987). Susceptibility of brook trout (*Salvelinus fontinalis*) to experimental infections with different serotypes of *Yersinia ruckeri*. *Microbios Lett* **35**, 27–31.

Confer, A. W. & Ayalew, S. (2013). The OmpA family of proteins: Roles in bacterial pathogenesis and immunity. *Vet Microbiol* **163**, 207–222. Elsevier B.V.

Connolly, J. P., Comerci, D., Alefantis, T. G., Walz, A., Quan, M., Chafin, R., Grewal, P., Mujer, C. V, Ugalde, R. A. & Delvecchio, V. G. (2006). Proteomic analysis of *Brucella abortus* cell envelope and identification of immunogenic candidate proteins for vaccine development. *Proteomics* **6**, 3767–3780.

Coquet, L., Cosette, P., Quillet, L., Petit, F., Junter, G. A. & Jouenne, T. (2002). Occurrence and phenotypic characterization of *Yersinia ruckeri* strains with biofilm-forming capacity in a rainbow trout farm. *Appl Environ Microbiol* **68**, 470–475.

Coquet, L., Cosette, P., Dé, E., Galas, L., Vaudry, H., Rihouey, C., Lerouge, P., Junter, G. A. & Jouenne, T. (2005). Immobilization induces alterations in the outer membrane protein pattern of *Yersinia ruckeri*. *J Proteome Res* **4**, 1988–98.

Cordwell, S. J. (2006). Technologies for bacterial surface proteomics. *Curr Opin Microbiol* **9**, 320–9.

Cordwell, S. J., Len, A. C. L., Touma, R. G., Scott, N. E., Falconer, L., Jones, D., Connolly, A., Crossett, B. & Djordjevic, S. P. (2008). Identification of membrane-associated proteins from *Campylobacter jejuni* strains using complementary proteomics technologies. *Proteomics* **8**, 122–39.

Cossarini-Dunier, M. (1985). Indirect enzyme-linked immunosorbent assay (ELISA) to titrate rainbow trout serum antibodies against two pathogens: *Yersinia ruckeri* and Egtved virus. *Aquaculture*.

Costa, T. R. D., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M. & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol* **13**, 343–359. Nature Publishing Group.

Cotter, S. E., Surana, N. K. & St Geme, J. W. (2005). Trimeric autotransporters: a distinct subfamily of autotransporter proteins. *Trends Microbiol* **13**, 199–205.

Crosa, J. H., Mey, A. R. & Payne, S. M. (2004). Iron Transport in Bacteria. In *Iron Transp Bact*.

Cullender, T. C., Chassaing, B., Janzon, A., Kumar, K., Muller, C. E., Werner, J. J., Angenent, L. T., Bell, M. E., Hay, A. G. & other authors. **(2013).** Innate and adaptive immunity interact to quench microbiome flagellar motility in the gut. *Cell (host microbe)* **14**, 571–81. Elsevier Inc.

Al Dahouk, S., Jubier-Maurin, V., Scholz, H. C., Tomaso, H., Karges, W., Neubauer, H. & Köhler, S. (2008). Quantitative analysis of the intramacrophagic *Brucella suis* proteome reveals metabolic adaptation to late stage of cellular infection. *Proteomics* **8**, 3862–70.

Daligault, H. E., Davenport, K. W., Minogue, T. D., Broomall, S. M., Bruce, D. C., Chain, P. S., Coyne, S. R., Frey, K. G., Gibbons, H. S. & other authors. **(2014).** Whole-genome *Yersinia* sp. assemblies from 10 diverse strains. *Genome Announc* **2**, 2–3.

Dalmo, R. A., Ingebrigtsen, K. & Begwald, J. (1997). Non-specific defence mechanisms in fish, with particular reference to the reticuloendothelial system (RES). *J Fish Dis* **20**, 241–273.

Daly, J. G., Lindvik, B. & Stevenson, R. M. W. (1986). Serological heterogeneity of recent isolates of *Yersinia ruckeri* from Ontario and British Columbia. *Dis Aquat Organ* **1**, 151–153.

Davies, R. L. (1989). *Biochemical and cell-surface characteristics of Yersinia ruckeri in relation to the epizootology and pathogenesis of infections in fish*. University of Stirling, Scotland (PhD Thesis).

Davies, R. L. (1990). O-serotyping of *Yersinia ruckeri* with special emphasis on European isolates. *Vet Microbiol* **22**, 299–307.

Davies, R. L. (1991a). Virulence and serum-resistance in different clonal groups and serotypes of *Yersinia ruckeri*. *Vet Microbiol* **29**, 289–97.

Davies, R. L. (1991b). *Yersinia ruckeri* produces four iron-regulated outer membrane proteins but does not produce detectable siderophores. *J Fish Dis* **14**, 563–570.

Davies, R. L. (1991c). Outer membrane protein profiles of *Yersinia ruckeri*. *Vet Microbiol* **26**, 125–40.

Davies, R. L. (1991d). Clonal analysis of *Yersinia ruckeri* based on biotypes, serotypes and outer membrane protein types. *J Fish Dis* **14**, 221–228.

Davies, R. L. & Donachie, W. (1996). Intra-specific diversity and host specificity within *Pasteurella haemolytica* based on variation of capsular polysaccharide, lipopolysaccharide and outer-membrane proteins. *Microbiology* **142**, 1895–907.

Davies, R. L. & Frerichs, G. N. (1989). Morphological and biochemical differences among isolates of *Yersinia ruckeri* obtained from wide geographical areas. *J Fish Dis* **12**, 357–365.

Davies, R. L. & Lee, I. (2004). Sequence diversity and molecular evolution of the heat-modifiable outer membrane protein gene (ompA) of *Mannheimia (Pasteurella) haemolytica*, *Mannheimia glucosida*, and *Pasteurella trehalosi*. *J Bacteriol* **186**, 5741–5752.

Davies, R. L., Wall, R. A. & Borriello, S. P. (1990). Comparison of methods for the analysis of outer membrane antigens of *Neisseria meningitidis* by western blotting. *J Immunol Methods* **134**, 215–25.

Davies, R. L., Ali, Q., Parton, R., Coote, J. G., Gibbs, A. & Freer, J. H. (1991). Optimal conditions for the analysis of *Pasteurella haemolytica* lipopolysaccharide by sodium dodecyl sulphate-polyacrylamide gel electrophoresis. *FEMS Microbiol Lett* **69**, 23–8.

Davies, R. L., Parton, R., Coote, J. G., Gibbs, H. A. & Freer, J. H. (1992). Outer-membrane

protein and lipopolysaccharide variation in *Pasteurella haemolytica* serotype A1 under different growth conditions. *J Gen Microbiol* **138**, 909–22.

Davies, R. L., Parton, R., Coote, J. G., Gibbs, H. A. & Freer, J. H. (1994). Evaluation of different methods for the detection of outer membrane proteins and lipopolysaccharides of *Pasteurella haemolytica* by immunoblotting. *J Immunol Methods* **167**, 35–45.

Davies, R. L., MacCorquodale, R., Baillie, S. & Caffrey, B. (2003). Characterization and comparison of *Pasteurella multocida* strains associated with porcine pneumonia and atrophic rhinitis. *J Med Microbiol* **52**, 59–67.

Davies, R. L., MacCorquodale, R. & Reilly, S. (2004). Characterisation of bovine strains of *Pasteurella multocida* and comparison with isolates of avian, ovine and porcine origin. *Vet Microbiol* **99**, 145–58.

Debroy, C., Fratamico, P. M., Roberts, E., Davis, M. A. & Liu, Y. (2005). Development of PCR assays targeting genes in O-antigen gene clusters for detection and identification of *Escherichia coli* O45 and O55 serogroups. *Appl Environ Microbiol* **71**, 4919–4924.

Didierlaurent, A., Ferrero, I., Otten, L. A., Dubois, B., Reinhardt, M., Carlsen, H., Blomhoff, R., Akira, S., Kraehenbuhl, J.-P. & Sirard, J.-C. (2004). Flagellin promotes myeloid differentiation factor 88-dependent development of Th2-type response. *J Immunol* **172**, 6922–6930.

Diep, B. A., Phung, Q., Date, S., Arnott, D., Bakalarski, C., Xu, M., Nakamura, G., Swem, D. L., Alexander, M. K. & other authors. **(2014).** Identifying potential therapeutic targets of methicillin-resistant *Staphylococcus aureus* through in vivo proteomic analysis. *J Infect Dis* **209**, 1533–41.

DiNubile, M. J. (2012). Serotype replacement after pneumococcal vaccination. *Lancet* **379**, 1388; author reply 1388–9. Elsevier Ltd.

Dong, H., Xiang, Q., Gu, Y., Wang, Z., Paterson, N. G., Stansfeld, P. J., He, C., Zhang, Y., Wang, W. & Dong, C. (2014). Structural basis for outer membrane lipopolysaccharide insertion. *Nature* **511**, 52–6. Nature Publishing Group.

Durbin, M., McIntosh, D., Smith, P. D., Wardle, R. & Austin, B. (1999). Immunization against furunculosis in rainbow trout with iron-regulated outer membrane protein vaccines: relative efficacy of immersion, oral, and injection delivery. *J Aquat Anim Health* **11**, 68–75.

Eaves-Pyles, T., Murthy, K., Liaudet, L., Virag, L., Ross, G., Soriano, F. G., Szabo, C. & Salzman, A. L. (2001). Flagellin, a novel mediator of *Salmonella*-induced epithelial activation and systemic inflammation: I B degradation, induction of nitric oxide synthase, induction of proinflammatory mediators, and cardiovascular dysfunction. *J Immunol* **166**, 1248–1260.

Ebanks, R. O., Dacanay, A., Goguen, M., Pinto, D. M. & Ross, N. W. (2004). Differential proteomic analysis of *Aeromonas salmonicida* outer membrane proteins in response to low iron and <i>in vivo</u> growth conditions. *Proteomics* **4**, 1074–1085.

E-komon, T., Burchmore, R., Herzyk, P. & Davies, R. (2012). Predicting the outer membrane proteome of *Pasteurella multocida* based on consensus prediction enhanced by results integration and manual confirmation. *BMC Bioinformatics* **13**, 63.

Esteve, C., Alcaide, E., Canals, R., Merino, S., Blasco, D., Figueras, M. J. & Tomas, J. M. (2004). Pathogenic *Aeromonas hydrophila* serogroup O:14 and O:81 strains with an S-layer. *Appl Environ Microbiol* **70**, 5898–5904.

Esteve, C., Alcaide, E., Herraiz, S., Canals, R., Merino, S. & Tomás, J. M. (2007). First description of nonmotile *Vibrio vulnificus* strains virulent for eels. *FEMS Microbiol Lett* **266**, 90–97.

Evenhuis, J. P., Lapatra, S. E., Verner-Jeffreys, D. W., Dalsgaard, I. & Welch, T. J. (2009). Identification of flagellar motility genes in *Yersinia ruckeri* by transposon mutagenesis. *Appl Environ Microbiol* **75**, 6630–3.

Ewing, W. H., Ross, A. J., Brenner, D. J., Fanning, G. R. & Reed, W. (1978). *Yersinia ruckeri* sp. nov., the redmouth (RM) bacterium. *J Syst Bacteriol* **28**, 37–44.

FAO. **(2014).** *FAO Global Aquaculture Production Volume and Value Statistics Database Updated to 2012*.

Faraldo-Gómez, J. D. & Sansom, M. S. P. (2003). Acquisition of siderophores in gram-negative bacteria. *Nat Rev Mol Cell Biol* **4**, 105–16.

Farmer, J., Davis, B. R., Hickerman-Brenner, F. W., McWhorter, A., Huntley-Carter, G. ., Asbury, M. A., Riddle, C., Wathen-Grady, H. G., Elias, C. & other authors. **(1985).** Biochemical identification of new species and biogroups of Enterobacteriaceae isolated from clinical specimens. *J Clin Microbiol* **21**, 46–76.

Fernández, L., Lopez, J. R., Secades, P., Menendez, A. & Guijarro, J. A. (2003). *In vitro* and *in vivo* studies of the Yrp1 protease from *Yersinia ruckeri* and its role in protective immunity against Enteric redmouth disease of salmonids. *Appl Environ Microbiol* **69**, 7328–7335.

Fernández, L., Secades, P., Lopez, J. R., Márquez, I. & Guijarro, J. A. (2002). Isolation and analysis of a protease gene with an ABC transport system in the fish pathogen *Yersinia ruckeri*: insertional mutagenesis and involvement in virulence. *Microbiology* **148**, 2233–2243.

Fernández, L., Marquez, I., Guijarro, J. A. & Ma, I. (2004). Identification of specific *in vivo*-induced (*ivi*) genes in *Yersinia ruckeri* and analysis of Ruckerbactin, a catecholate siderophore iron acquisition system. *Appl Environ Microbiol* **70**, 5199–5207.

Fernández, L., Prieto, M. & Guijarro, J. A. (2007a). The iron- and temperature-regulated haemolysin YhlA is a virulence factor of *Yersinia ruckeri*. *Microbiology* **153**, 483–9.

Fernández, L., Méndez, J. & Guijarro, J. A. (2007b). Molecular virulence mechanisms of the fish pathogen *Yersinia ruckeri*. *Vet Microbiol* **125**, 1–10.

Filloux, A. (2010). A variety of bacterial pili involved in horizontal gene transfer. *J Bacteriol* **192**, 3243–3245.

Fitzgerald, C., Sherwood, R., Gheesling, L. L., Brenner, F. W. & Fields, P. I. (2003). Molecular analysis of the rfb O-antigen gene cluster of *Salmonella enterica* serogroup O:6,14 and development of a Serogroup-Specific PCR Assay. *Appl Environ Microbiol* **69**, 6099–6105.

Flores-Ramirez, G., Jankovicova, B., Bilkova, Z., Miernyk, J. A. & Skultety, L. (2014). Identification of *Coxiella burnetii* surface-exposed and cell envelope associated proteins using a combined bioinformatics plus proteomics strategy. *Proteomics* **14**, 1868–81.

Folsom, J. P., Parker, A. E. & Carlson, R. P. (2014). Physiological and proteomic analysis of *Escherichia coli* iron-limited chemostat growth. *J Bacteriol* **196**, 2748–61.

Forst, S., Delgado, J. & Inouye, M. (1989). Phosphorylation of OmpR by the osmosensor EnvZ modulates expression of the ompF and ompC genes in *Escherichia coli*. *Proc Natl Acad Sci U S A* **86**, 6052–6056.

Fouz, B., Zarza, C. & Amaro, C. (2006). First description of non-motile *Yersinia ruckeri* serovar I strains causing disease in rainbow trout, (*Oncorhynchus mykiss*, Walbaum), cultured in Spain. *J Fish Dis* **29**, 339–46.

Francis, S. M. (2011). Secretion systems and metabolism in the pathogenic *Yersiniae*. In

Stress response Pathog Bact, p. 185.

Fratamico, P., Bagi, L., Cray, W. & Narang, N. (2011). Detection by multiplex real-time polymerase chain reaction assays and isolation of Shiga toxin–producing *Escherichia coli* serogroups O26, O45, O103, O111, O121, and O145 in ground beef. *Foodborne Pathog Dis* **8**.

Fronzes, R., Christie, P. J. & Waksman, G. (2009). The structural biology of type IV secretion systems. *Nat Rev Microbiol* **7**, 703–714. Nature Publishing Group.

Fsihisdli, H., Kottwitzs, B. & Bremer, E. (1993). Single Amino Acid Substitutions Affecting the Substrate Specificity of the *Escherichia coli* K-12 Nucleoside-specific Tsx channel. *Mol Biol* **268**, 17495–17503.

Galdiero, S., Falanga, A., Cantisani, M., Tarallo, R., Della Pepa, M. E., D'Oriano, V. & Galdiero, M. (2012). Microbe-host interactions: structure and role of Gram-negative bacterial porins. *Curr Protein Pept Sci* **13**, 843–54.

Garcia, J. A., Dominguez, L., Larsen, J. L. & Pedersen, K. (1998). Ribotyping and plasmid profiling of *Yersinia ruckeri*. *J Appl Microbiol* **85**, 949–55.

Garrow, A. G., Agnew, A. & Westhead, D. R. (2005). TMB-Hunt: an amino acid composition based method to screen proteomes for beta-barrel transmembrane proteins. *BMC Bioinformatics* **6**, 56.

Gauthier, M. J., Munro, P. M. & Mohajer, S. (1987). Influence of salts and sodium chloride on the recovery of *Escherichia coli* from seawater. *Immunology* **15**, 5–10.

St. Geme, J. W. & Yeo, H. J. (2009). A prototype two-partner secretion pathway: the *Haemophilus influenzae* HMW1 and HMW2 adhesin systems. *Trends Microbiol* **17**, 355–360.

Gesslbauer, B., Poljak, A., Handwerker, C., Schüler, W., Schwendenwein, D., Weber, C., Lundberg, U., Meinke, A. & Kungl, A. J. (2012). Comparative membrane proteome analysis of three *Borrelia* species. *Proteomics* **12**, 845–58.

Gewirtz, A. T., Navas, T. A., Lyons, S., Godowski, P. J. & Madara, J. L. (2001). Cutting edge: bacterial flagellin activates basolaterally expressed TLR5 to induce epithelial proinflammatory gene expression. *J Immunol* **167**, 1882–1885.

Gibello, A., Blanco, M. M., Moreno, M. A., Cutuli, M. T. & Domínguez, L. (1999). Development of a PCR assay for detection of *Yersinia ruckeri* in tissues of inoculated and naturally infected trout. *Appl Environ Microbiol* **65**, 346–350.

Gil, F., Ipinza, F., Fuentes, J., Fumeron, R., Villarreal, J. M., Aspée, A., Mora, G. C., Vásquez, C. C. & Saavedra, C. (2007). The ompW (porin) gene mediates methyl viologen (paraquat) efflux in *Salmonella enterica* serovar Typhimurium. *Res Microbiol* **158**, 529–536.

Glenn, R. A., Taylor, P. W. & Hanson, K. C. (2011). The use of a real-time PCR primer/probe set to observe infectivity of *Yersinia ruckeri* in Chinook salmon, *Oncorhynchus tshawytscha* (Walbaum), and steelhead trout, *Oncorhynchus mykiss* (Walbaum). *J Fish Dis* **34**, 783–91.

Goel, A. K. & Jiang, S. C. (2010). Genetic determinants of virulence, antibiogram and altered biotype among the *Vibrio cholerae* O1 isolates from different cholera outbreaks in India. *Infect Genet Evol* **10**, 814–818. Elsevier B.V.

Görg, A., Weiss, W. & Dunn, M. J. (2004). Current two-dimensional electrophoresis technology for proteomics. *Proteomics* **4**, 3665–3685.

Gort, A. S. & Miller, V. L. (2000). Identification and characterization of *Yersinia enterocolitica*

genes induced during systemic infection. *Infect Immun* **68**, 6633–6642.

De Grandis, S. A. & Stevenson, R. M. W. (1982). Variations in plasmid profiles and growth characteristics of *Yersinia ruckeri* strains. *FEMS Microbiol Lett*.

De Grandis, S. A. & Stevenson, R. M. W. (1985). Antimicrobial susceptibility patterns and R plasmid-mediated resistance of the fish pathogen *Yersinia ruckeri*. *Antimicrob Agents Chemother* **27**.

De Grandis, S. A., Krell, P. J., Flett, D. E. & Stevenson, R. M. W. (1988). Deoxyribonucleic Acid Relatedness of Serovars of *Yersinia ruckeri*, the Enteric Redmouth Bacterium. *Int J Syst Bacteriol* **38**, 49–55.

Grant, J. R. & Stothard, P. (2008). The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Res* **36**, 181–184.

Green, M. & Austin, B. (1983). The identification of *Yersinia ruckeri* and its relationship to other representatives of the *Enterobacteriaceae*. *Aquaculture* **34**, 185–192.

Gu, W., Wang, X., Qiu, H., Luo, X., Xiao, D., Xiao, Y., Tang, L., Kan, B. & Jing, H. (2012). Comparative antigenic proteins and proteomics of pathogenic *Yersinia enterocolitica* bio-serotypes 1B/O: 8 and 2/O: 9 cultured at 25°C and 37°C. *Microbiol Immunol* **56**, 583–94.

Gu, Y., Stansfeld, P. J., Zeng, Y., Dong, H., Wang, W. & Dong, C. (2015). Lipopolysaccharide is inserted into the outer membrane through an intramembrane hole, a lumen gate, and the lateral opening of LptD. *Structure* **23**, 496–504.

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**, 1–11.

Guilvout, I., Quilici, M. L., Rabot, S., Lesel, R. & Mazigh, D. (1988). BamHI restriction endonuclease analysis of *Yersinia ruckeri* plasmids and their relatedness to the genus *Yersinia* 42- to 47-megadalton plasmid. *Appl Environ Microbiol* **54**, 2594–7.

Guina, T., Wu, M., Miller, S. I., Purvine, S. O., Yi, E. C., Eng, J., Goodlett, D. R., Aebersold, R., Ernst, R. K. & Lee, K. A. (2003). Proteomic analysis of *Pseudomonas aeruginosa* grown under magnesium limitation. *J Am Soc Mass Spectrom* **14**, 742–51.

Gunasena, D. K., Komrower, J. R. & Macintyre, S. (2003). The fish pathogen *Yersinia ruckeri* possesses a TTSS. In *The Genus Yersinia*, pp. 105–107. Edited by M. Skurnik. Kluwer Academic/Plenum Publishers.

Gunn, J. S. (2000). Mechanisms of bacterial resistance and response to bile. *Microbes Infect* **2**, 907–913.

Hagan, C. L., Silhavy, T. J. & Kahne, D. (2011). β-Barrel membrane protein assembly by the Bam Complex. *Annu Rev Biochem* **80**, 189–210.

Haig, S. J., Davies, R. L., Welch, T. J., Reese, R. A. & Verner-Jeffreys, D. W. (2011). Comparative susceptibility of Atlantic salmon and rainbow trout to *Yersinia ruckeri*: relationship to O antigen serotype and resistance to serum killing. *Vet Microbiol* **147**, 155–61. Elsevier B.V.

Hall, M., Chattaway, M. A., Reuter, S., Savin, C., Strauch, E., Carniel, E., Connor, T., Van Damme, I., Rajakaruna, L. & other authors. **(2015).** Use of whole-genus genome sequence data to develop a multilocus sequence typing tool that accurately identifies *Yersinia* isolates to the species and subspecies levels. *J Clin Microbiol* **53**, 35–42.

Hanage, W. P., Huang, S. S., Lipsitch, M., Bishop, C. J., Godoy, D., Pelton, S. I., Goldstein, R., Huot, H. & Finkelstein, J. A. (2007). Diversity and antibiotic resistance among nonvaccine serotypes of *Streptococcus pneumoniae* carriage isolates in the post-

heptavalent conjugate vaccine era. *J Infect Dis* **195**, 347–52.

Hantke, K. (2001). Iron and Metal Regulation in bacteria. *Curr Opin Microbiol* **4**, 172–177.

Håstein, T., Gudding, R. & Evensen, O. (2005). Bacterial vaccines for fish - an update of the current situation worldwide. *Dev Biol (Basel)*.

Hatfaludi, T., Al-Hasani, K., Boyce, J. D. & Adler, B. (2010). Outer membrane proteins of *Pasteurella multocida*. *Vet Microbiol* **144**, 1–17.

Hauberg, L., Schmidt, F., Scharf, C., Dörr, J., Völker, U. & Reinhold-Hurek, B. (2010). Proteomic characterization of a pilR regulatory mutant of *Azoarcus* sp. strain BH72 with the aid of gel-based and gel-free approaches. *Proteomics* **10**, 458–69.

Heffernan, E. J., Wu, L., Louie, J., Okamoto, S., Fierer, J. & Guiney, D. G. (1994). Specificity of the complement resistance and cell association phenotypes encoded by the outer membrane protein genes *rck* from *Salmonella typhimurium* and *ail* from *Yersinia enterocolitica*. *Infect Immun* **62**, 5183–5186.

Heinz, E., Pichler, P., Heinz, C., Op den Camp, H. J. M., Toenshoff, E. R., Ammerer, G., Mechtler, K., Wagner, M. & Horn, M. (2010). Proteomic analysis of the outer membrane of *Protochlamydia amoebophila* elementary bodies. *Proteomics* **10**, 4363–76.

Hellman, J., Loiselle, P. M., Tehan, M. M., Allaire, J. E., Boyle, L. a., Kurnick, J. T., Andrews, D. M., Kim, K. S. & Warren, H. S. (2000). Outer membrane protein A, peptidoglycan-associated lipoprotein, and murein lipoprotein are released by *Escherichia coli* bacteria into serum. *Infect Immun* **68**, 2566–2572.

Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C. & Ala'Aldeen, D. (2004). Type V Protein Secretion Pathway : the Autotransporter Story Type V Protein Secretion Pathway : the Autotransporter Story. *Microbiol Mol Biol Rev* **68**, 692–744.

Hitchcock, P. J. & Brown, T. M. (1983). Morphological heterogeneity among *Salmonella* lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. *J Bacteriol* **154**, 269–77.

Hitchcock, P. J., Leive, L., Mäkelä, P. H., Rietschel, E. T., Strittmatter, W. & Morrison, D. C. (1986). Lipopolysaccharide nomenclature - past, present, and future. *J Bacteriol* **166**, 699–705.

Hoiczyk, E., Roggenkamp, A., Reichenbecher, M., Lupas, A. & Heesemann, J. (2000). Structure and sequence analysis of *Yersinia* YadA and *Moraxella* UspAs reveal a novel class of adhesins. *EMBO J* **19**, 5989–5999.

Horne, M. & Barnes, A. (1998). Enteric redmouth disease. In *Fish Dis Disord Vol 3; viral, Bact fungal Infect*, 1st edn., pp. 455–479. Edited by P. T. K. Woo & D. W. Bruno. CABI PUBLISHING.

Horne, M. & Barnes, A. (1999). Enteric redmouth disease. In *Fish Dis Disord Vol 3; viral, Bact fungal Infect*, pp. 455 – 477.

Hounsome, J. D. A., Baillie, S., Noofeli, M., Riboldi-Tunnicliffe, A., Burchmore, R. J. S., Isaacs, N. W. & Davies, R. L. (2011). Outer membrane protein A of bovine and ovine isolates of *Mannheimia haemolytica* is surface exposed and contains host species-specific epitopes. *Infect Immun* **79**, 4332–41.

Hunter, V. A., Knittel, M. D. & Fryer, J. L. (1980). Stress-induced transmission of *Yersinia ruckeri* infection from carriers to recipient steelhead trout (*Salmo gairdneri*, Richardson). *J Fish Dis* **3**, 467–472.

Hurst, M. R. H., Becher, S. A., Young, S. D., Nelson, T. L. & Glare, T. R. (2011). *Yersinia entomophaga* sp. nov., isolated from the New Zealand grass grub (*Costelytra zealandica*).

Int J Syst Evol Microbiol **61**, 844–9.

Hynes, N. A., Furnes, C., Fredriksen, B. N., Winther, T., Bøgwald, J., Larsen, A. N. & Dalmo, R. A. (2011). Immune response of Atlantic salmon to recombinant flagellin. *Vaccine* **29**, 7678–87. Elsevier Ltd.

Imai, K., Asakawa, N., Tsuji, T., Akazawa, F., Ino, A., Sonoyama, M. & Mitaku, S. (2008). SOSUI-GramN: high performance prediction for sub-cellular localization of proteins in Gram-negative bacteria. *Bioinformation* **2**, 417–421.

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J. & Mann, M. (2005). Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* **4**, 1265–72.

Jagannadham, M. V, Abou-Eladab, E. F. & Kulkarni, H. M. (2011). Identification of outer membrane proteins from an Antarctic bacterium *Pseudomonas syringae* Lz4W. *Mol Cell Proteomics* **10**, M110.004549.

Jang, K.-S., Sweredoski, M. J., Graham, R. L. J., Hess, S. & Clemons, W. M. (2014). Comprehensive proteomic profiling of outer membrane vesicles from *Campylobacter jejuni*. *J Proteomics* **98**, 90–8. Elsevier B.V.

Jeannin, P., Magistrelli, G., Goetsch, L., Haeuw, J. F., Thieblemont, N., Bonnefoy, J. Y. & Delneste, Y. (2002). Outer membrane protein A (OmpA): A new pathogen-associated molecular pattern that interacts with antigen presenting cells - Impact on vaccine strategies. *Vaccine* **20**, 23–27.

Johansson, A., Claesson, R., Hänström, L. & Kalfas, S. (2003). Serum-mediated release of leukotoxin from the cell surface of the periodontal pathogen *Actinobacillus actinomycetemcomitans*. *Eur J Oral Sci* **111**, 209–215.

Johnson, S. L., Daligault, H. E., Davenport, K. W., Jaissle, J., Frey, K. G., Ladner, J. T., Broomall, S. M., Bishop-lilly, K. A., Bruce, D. C. & other authors. **(2015).** Thirty-two complete genome assemblies of nine *Yersinia* species, including *Y. pesits, Y. pseudotuberculosis*, and *Y. enterocolitica*. *Genome Announc* **3**, 3–4.

Johnson, T. J., Abrahante, J. E., Hunter, S. S., Hauglund, M., Tatum, F. M., Maheswaran, S. K. & Briggs, R. E. (2013). Comparative genome analysis of an avirulent and two virulent strains of avian *Pasteurella multocida* reveals candidate genes involved in fitness and pathogenicity. *BMC Microbiol* **13**, 106. BMC Microbiology.

Jovanovic, M., Lilic, M., Savic, D. J. & Javanovic, G. (2003). The LysR-type transcriptional regulator CysB controls the repression of hslJ transcription in *Escherichia coli*. *Microbiology* **149**, 3449–3459.

Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H. & Krogh, A. (2003). Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* **12**, 1652–1662.

Kanaujia, P. K., Bajaj, P., Kumar, S., Singhal, N. & Virdi, J. S. (2015). Proteomic analysis of *Yersinia enterocolitica* biovar 1A under iron-rich and iron-poor conditions indicate existence of efficiently regulated mechanisms of iron homeostasis. *J Proteomics* 1–11. Elsevier B.V.

Kao, D.-Y., Cheng, Y.-C., Kuo, T.-Y., Lin, S.-B., Lin, C.-C., Chow, L.-P. & Chen, W.-J. (2009). Salt-responsive outer membrane proteins of *Vibrio anguillarum* serotype O1 as revealed by comparative proteome analysis. *J Appl Microbiol* **106**, 2079–85.

Kawaoka, Y., Otsuki, K. & Tsubokura, M. (1983). Growth temperature-dependent variation in the bacteriophage-inactivating capacity and antigenicity of *Yersinia enterocolitica*

lipopolysaccharide. *J Gen Microbiol* **129**, 2739–2747.

Kawula, T. H., Lelivelt, M. J. & Orndorff, P. E. (1996). Using a new inbred fish model and cultured fish tissue cells to study *Aeromonas hydrophila* and *Yersinia ruckeri* pathogenesis. *Microb Pathog* **20**, 119–125.

Keeling, S. E., Johnston, C., Wallis, R., Brosnahan, C. L., Gudkovs, N. & McDonald, W. L. (2012). Development and validation of real-time PCR for the detection of *Yersinia ruckeri*. *J Fish Dis* **35**, 119–25.

Khushiramani, R. M., Maiti, B., Shekar, M., Girisha, S. K., Akash, N., Deepanjali, A., Karunasagar, I. & Karunasagar, I. (2012). Recombinant *Aeromonas hydrophila* outer membrane protein 48 (Omp48) induces a protective immune response against *Aeromonas hydrophila* and *Edwardsiella tarda*. *Res Microbiol* **163**, 286–291. Elsevier Masson SAS.

Kibenge, F. S. B., Munir, K., Kibenge, M. J. T. & Joseph, T. (2004). Infectious salmon anemia virus: causative agent , pathogenesis and immunity. *Anim Heal Res Rev* **5**, 65–78.

Kibenge, F. S. B., Godoy, M. G., Wang, Y., Kibenge, M. J. T., Gherardelli, V., Mansilla, S., Lisperger, A., Jarpa, M., Larroquete, G. & other authors. **(2009).** Infectious salmon anaemia virus (ISAV) isolated from the ISA disease outbreaks in Chile diverged from ISAV isolates from Norway around 1996 and was disseminated around 2005, based on surface glycoprotein gene sequences. *Virol J* **6**, 88.

Kim, K.-M., Lee, S.-G., Joo, J.-S., Kwon, Y.-C., Bea, D.-W., Song, J.-Y., Kang, H.-L., Lee, W.-K., Cho, M.-J. & other authors. **(2008).** Proteomic analysis of *Helicobacter pylori* J99 outer membrane protein by tandem mass spectrometry. *J Bacteriol Virol* **38**, 53.

Knowles, T. J., Scott-Tucker, A., Overduin, M. & Henderson, I. R. (2009). Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Microbiol* **7**, 206–214.

Kodama, H., Moustafa, M., Ishiguro, S., Mikami, T. & Izawa, H. (1984). Extracellular virulence factors of fish *Vibrio*: relationships between toxic material, hemolysin, and proteolytic enzyme. *Am J Vet Res* **45**, 2203–2207.

Koebnik, R., Locher, K. P. & Van Gelder, P. (2000). Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol* **37**, 239–53.

Koppe, B. & Harms, H. (1994). Antigenic determinants and specificity of antisera against acidophilic bacteria. *World J Microbiol Biotechnol* **10**, 154–158.

Kotetishvili, M., Kreger, A., Wauters, G., Jr, G. M., Sulakvelidze, A., Stine, O. C. & Morris, J. G. (2005). Multilocus sequence typing for studying genetic relationships among *Yersinia* species. *J Clin Microbiol* **43**, 2674–2684.

Kovacs-Simon, A., Titball, R. W. & Michell, S. L. (2011). Lipoproteins of bacterial pathogens. *Infect Immun* **79**, 548–561.

Krachler, A. M. & Orth, K. (2011). Functional characterization of the interaction between bacterial adhesin Multivalent Adhesion Molecule 7 (MAM7) protein and its host cell ligands. *J Biol Chem* **286**, 38939–38947.

Krachler, A. M., Ham, H. & Orth, K. (2011). Outer membrane adhesion factor multivalent adhesion molecule 7 initiates host cell binding during infection by Gram-negative pathogens. *Proc Natl Acad Sci U S A* **108**, 11614–11619.

Kruh, N. a, Troudt, J., Izzo, A., Prenni, J. & Dobos, K. M. (2010). Portrait of a pathogen: the *Mycobacterium tuberculosis* proteome in vivo. *PLoS One* **5**, e13938.

Kudva, I. T., Stanton, T. B. & Lippolis, J. D. (2014). The *Escherichia coli* O157:H7 bovine rumen fluid proteome reflects adaptive bacterial responses. *BMC Microbiol* **14**, 48.

Kulkarni, H. M., Swamy, C. V. B. & Jagannadham, M. V. **(2014).** Molecular characterization and functional analysis of outer membrane vesicles from the Antarctic bacterium *Pseudomonas syringae* suggest a possible response to environmental conditions. *J Proteome Res* **13**, 1345–58.

Kumar, G., Sharma, P., Rathore, G., Bisht, D. & Sengupta, U. (2010). Proteomic analysis of outer membrane proteins of *Edwardsiella tarda*. *J Appl Microbiol* **108**, 2214–21.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

LaFrentz, B. R., LaPatra, S. E., Call, D. R., Wiens, G. D. & Cain, K. D. (2009). Proteomic analysis of *Flavobacterium psychrophilum* cultured in vivo and in iron-limited media. *Dis Aquat Organ* **87**, 171–182.

Lai, E., Nair, U., Phadke, N. D. & Maddock, J. R. (2004). Proteomic screening and identification of differentially distributed membrane proteins in *Escherichia coli*. *Mol Microbiol* **52**, 1029–1044.

Lallier, R. & Daigneault, P. (1984). Antigenic differentiation of pili from non-virulent and fish-pathogenic strains of *Aeromonas hydrophila*. *J Fish Dis* **7**, 509–512.

Lan, C. & Igo, M. M. (1998). Differential Expression of the OmpF and OmpC Porin Proteins in *Escherichia coli* K-12 Depends upon the Level of Active OmpR. *J Bacteriol* **180**, 171–174.

Lawrenz, M. B., Lenz, J. D. & Miller, V. L. (2009). A novel autotransporter adhesin is required for efficient colonization during bubonic plague. *Infect Immun* **77**, 317–326.

Lawrenz, M. B., Pennington, J. & Miller, V. L. (2013). Acquisition of omptin reveals cryptic virulence function of autotransporter YapE in *Yersinia pestis*. *Mol Microbiol* **89**, 276–287.

Lee, E.-Y., Bang, J. Y., Park, G. W., Choi, D.-S., Kang, J. S., Kim, H.-J., Park, K.-S., Lee, J.-O., Kim, Y.-K. & other authors. **(2007).** Global proteomic profiling of native outer membrane vesicles derived from *Escherichia coli*. *Proteomics* **7**, 3143–53.

Lenz, J. D., Lawrenz, M. B., Cotter, D. G., Lane, M. C., Gonzalez, R. J., Palacios, M. & Miller, V. L. (2011). Expression during host infection and localization of *Yersinia pestis* autotransporter proteins. *J Bacteriol* **193**, 5936–5949.

Leo, J. C., Grin, I. & Linke, D. (2012). Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philos Trans R Soc B Biol Sci* **367**, 1088–1101.

Li, M., Shimada, T. & Morris, J. (2002). Evidence for the emergence of non-O1 and non-O139 *Vibrio cholerae* strains with pathogenic potential by exchange of O-antigen biosynthesis regions. *Infect Immun* **70**, 2441–2453.

Lim, J., Stones, D. H., Hawley, C. A., Watson, C. A. & Krachler, A. M. (2014). Multivalent adhesion molecule 7 clusters act as signaling platform for host cellular GTPase activation and facilitate epithelial barrier dysfunction. *PLoS Pathog* **10**, e1004421.

Lin, J., Huang, S. & Zhang, Q. (2002). Outer membrane proteins: key players for bacterial adaptation in host niches. *Microbes Infect* **4**, 325–31.

Liu, G. Y., Nie, P., Zhang, J. & Li, N. (2008). Proteomic analysis of the sarcosine-insoluble outer membrane fraction of *Flavobacterium columnare*. *J Fish Dis* **31**, 269–76.

Liu, X. & Ferenci, T. (2001). An analysis of multifactorial influences on the transcriptional control of ompF and ompC porin expression under nutrient limitation. *Microbiology* **147**, 2981–2989.

Loferer, H., Hammar, M. & Normark, S. (1997). Availability of the fibre subunit CsgA and the nucleator protein CsgB during assembly of fibronectin-binding curli is limited by the intracellular concentration of the novel lipoprotein CsgG. *Mol Microbiol* **26**, 11–23.

Lovely, J. E., Dannevig, B. H., Falk, K., Hutchin, L., MacKinnon, A. M., Melville, K. J., Rimstad, E. & Griffiths, S. G. (1999). First identification of infectious salmon anaemia virus in North America with haemorrhagic kidney syndrome. *Dis Aquat Organ* **35**, 145–148.

Luckey, M. (2014). Bundles and Barrels. In *Membr Struct Biol With Biochem Biophys Found*, 2nd edn., pp. 118–122. Cambridge University Press.

Lugtenberg, E. J. . & Peters, R. (1976). Distribution of lipids in cytoplasmic and outer membranes of *Escherichia coli* K12. *Biochim Biophys Acta* **441**, 38–47.

Ma, K., Khan, S. & Miyan, K. (2011). Aquaculture as a food production system : A review. *Biol Med* **3**, 291–302.

Maiden, M. C. J. (2006). Multilocus sequence typing of bacteria. *Annu Rev Microbiol* **60**, 561–588.

Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R. & Silhavy, T. J. (2006). YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol Microbiol* **61**, 151–164.

Marine Harvest Handbook. **(2012).** Salmon Farming Industry Handbook 2012.

Markwell, M., Haas, S., Bieber, L. & Tolbert, N. (1978). A modification of the Lowry procedure to simplify protein determination in membrane and lipoprotein samples. *Anal Biochem* **210**, 206–210.

Marshall, S. H., Conejeros, P., Zahr, M., Olivares, J., Gómez, F., Cataldo, P. & Henríquez, V. (2007). Immunological characterization of a bacterial protein isolated from salmonid fish naturally infected with *Piscirickettsia salmonis*. *Vaccine* **25**, 2095–102.

Martorana, A. M., Motta, S., Di Silvestre, D., Falchi, F., Dehò, G., Mauri, P., Sperandeo, P. & Polissi, A. (2014). Dissecting *Escherichia coli* outer membrane biogenesis using differential proteomics. *PLoS One* **9**, e100941.

Maskell, D., Frankel, G. & Dougan, G. (1993). Phase and antigenic variation - The impact on strategies for bacterial vaccine design. *Trends Biotechnol* **11**, 506–510.

Matsubara, M., Kitaoka, S. I., Takeda, S. I. & Mizuno, T. (2000). Tuning of the porin expression under anaerobic growth conditions by his-to-Asp cross-phosphorelay through both the EnvZ-osmosensor and ArcB-anaerosensor in *Escherichia coli*. *Genes Cells* **5**, 555–569.

Mattick, J. S. (2002). Type IV pili and twitching motility. *Annu Rev Microbiol* **56**, 289–314.

McCluskey, J., Gibbs, H. A. & Davies, R. L. (1994). Variation in outer-membrane protein and lipopolysaccharide profiles of *Pasteurella haemolytica* isolates of serotypes A1 and A2 obtained from pneumonic and healthy cattle. *Microbiology* **140 Pt 4**, 807–14.

Méndez, J., Fernández, L., Menéndez, A., Reimundo, P., Pérez-Pascual, D., Navais, R. & Guijarro, J. a. **(2009).** A chromosomally located traHIJKCLMN operon encoding a putative type IV secretion system is involved in the virulence of *Yersinia ruckeri*. *Appl Environ Microbiol* **75**, 937–45.

Merdanovic, M., Clausen, T., Kaiser, M., Huber, R. & Ehrmann, M. (2011). Protein quality control in the bacterial periplasm.

Michel, C., Faivre, B. & DeKinkelin, P. (1986). A clinical case of enteric redmouth in minnows

(*Pimephales promelas*) imported in Europe as bait-fish. *Bull Eur Assoc Fish Pathol* **6**, 7–9.

Miethke, M., Schmidt, S. & Marahiel, M. A. (2008). The major facilitator superfamily-type transporter YmfE and the multidrug-efflux activator Mta mediate bacillibactin secretion in *Bacillus subtilis*. *J Bacteriol* **190**, 5143–5152.

Miles, A. A., Misra, S. S. & Irwin, J. O. (1938). The estimation of the bactericidal power of the blood. *J Hyg (Lond)* **38**, 732–749.

Minnich, S. A. & Rohde, H. N. (2007). A rationale for repression and/or loss of motility by pathogenic *Yersinia* in the mammalian host. *Adv Exp Med Biol* **603**, 298–310.

Miyamoto, Y., Kato, T., Obara, Y., Akiyama, S., Takizawa, K. & Yamai, S. (1969). In vitro hemolytic characteristic of *Vibrio parahaemolyticus*: its close correlation with human pathogenicity. *J Bacteriol* **100**, 1147–1149.

Mjaaland, S., Rimstad, E. & Cunningham, C. O. (2002). Molecular diagnosis of infectious salmon anaemia. In *Mol diagnosis Salmon Dis*, pp. 1–22. Edited by C. O. Cunningham. Dordrecht: Kluwer Academic Publishers.

Molloy, M. P., Herbert, B. R., Slade, M. B., Rabilloud, T., Nouwens, A. S., Williams, K. L. & Gooley, a a. (2000). Proteomic analysis of the *Escherichia coli* outer membrane. *Eur J Biochem* **267**, 2871–81.

Morona, R., Kramer, C. & Henning, U. L. F. (1985). Bacteriophage receptor area of outer membrane protein OmpA of *Escherichia coli* K-12. *J Bacteriol* **164**, 539–543.

Moumène, A., Marcelino, I., Ventosa, M., Gros, O., Lefrançois, T., Vachiéry, N., Meyer, D. F. & Coelho, A. V. (2015). Proteomic profiling of the outer membrane fraction of the obligate intracellular bacterial pathogen *Ehrlichia ruminantium*. *PLoS One* **10**, e0116758.

Munro, P. M., Gauthier, M. J., Breittmayer, V. A. & Bongiovanni, J. (1989). Influence of osmoregulation processes on starvation survival of *Escherichia coli* in seawater. *Appl Environ Microbiol* **55**, 2017–2024.

Munro, P. M., Laumond, F. & Gauthier, M. J. (1987). A previous growth of enteric bacteria on a salted medium increases their survival in seawater. *Lett Appl Microbiol* **4**, 121–124.

Nabu, S., Lawung, R., Isarankura-Na-Ayudhya, P., Isarankura-Na-Ayudhya, C., Roytrakul, S. & Prachayasittikul, V. (2014). Reference map and comparative proteomic analysis of *Neisseria gonorrhoeae* displaying high resistance against spectinomycin. *J Med Microbiol* **63**, 371–85.

Nakatsuji, T., Shi, Y., Zhu, W., Huang, C.-P., Chen, Y.-R., Lee, D.-Y., Smith, J. W., Zouboulis, C. C., Gallo, R. L. & Huang, C.-M. (2008). Bioengineering a humanized acne microenvironment model: proteomics analysis of host responses to Propionibacterium acnes infection in vivo. *Proteomics* **8**, 3406–15.

Nandi, B., Nandy, R. K., Sarkar, A. & Ghose, A. C. (2005). Structural features, properties and regulation of the outer-membrane protein W (OmpW) of *Vibrio cholerae*. *Microbiology* **151**, 2975–2986.

Narita, S. I., Matsuyama, S. I. & Tokuda, H. (2004). Lipoprotein trafficking in *Escherichia coli*. *Arch Microbiol* **182**, 1–6.

Navas, E. & Bohle, H. (2014). Draft genome sequence of the fish pathogen *Yersinia ruckeri* strain 37551, serotype O1b, isolated from diseased, vaccinated Atlantic salmon (*Salmo salar*) in Chile. *Genome* **2**, 10–11.

Nelson, M. C., Lapatra, S. E. & Welch, T. J. (2015). Complete genome sequence of *Yersinia ruckeri* strain CSF007-82 , etiologic agent of red mouth disease in salmonid fish. *Genome Announc* **3**, 4–5.

Ni Bhriain, N., Dorman, C. J. & Higgins, C. F. (1989). An overlap between osmotic and anaerobic stress responses: a potential role for DNA supercoiling in the coordinate regulation of gene expression. *Mol Microbiol* **3**, 933–42.

Niemann, H. H., Schubert, W.-D. & Heinz, D. W. (2004). Adhesins and invasins of pathogenic bacteria: a structural view. *Microbes Infect* **6**, 101–112.

Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* **67**, 593–656.

Noga, E. J., Levine, J. F., Townsend, K., Bullis, R. A., Carlson, C. P. & Corbett, W. T. (1988). Kidney biopsy: a nonlethal method for diagnosing *Yersinia ruckeri* infection (Enteric redmouth disease) in rainbow trout (*Salmo gairdneri*). *Am J Vet Res* **49**, 363–365.

Noinaj, N., Kuszak, A. J., Gumbart, J. C., Lukacik, P., Chang, H., Easley, N. C., Lithgow, T. & Buchanan, S. K. (2013). Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* **501**, 385–90.

Norris, T. L. & Bäumler, A. J. (1999). Phase variation of the lpf operon is a mechanism to evade cross-immunity between *Salmonella* serotypes. *Proc Natl Acad Sci U S A* **96**, 13393–13398.

Nylund, A., Hovland, T., Watanabe, K. & Endresen, C. (1995). Presence of infectious salmon anaemia virus (ISAV) in tissues of Atlantic salmon, (*Salmo salar L.*), collected during three separate outbreaks of the disease. *J Fish Dis* **18**, 135–145.

O'Leary, P. J. (1977). *Enteric redmouth bacterium of salmonids: a biochemical and serological comparison of selected isolates*. Oregan State University (PhD Thesis).

O'Toole, G., Kaplan, H. B. & Kolter, R. (2000). Biofilm formation as microbial development. *Annu Rev Microbiol* **54**, 49–79.

O'Toole, G. A. & Kolter, R. (1998). Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Mol Microbiol* **30**, 295–304.

Ohtani, M., Villumsen, K. R., Strøm, H. K. & Raida, M. K. (2014). 3D Visualization of the initial *Yersinia ruckeri* infection toute in Rainbow trout (*Oncorhynchus mykiss*) by optical projection tomography. *PLoS One* **9**, e89672.

Okuda, S. & Tokuda, H. (2009). Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB. *Proc Natl Acad Sci U S A* **106**, 5877–5882.

Orme, R., Douglas, C. W. I., Rimmer, S. & Webb, M. (2006). Proteomic analysis of *Escherichia coli* biofilms reveals the overexpression of the outer membrane protein OmpA. *Proteomics* **6**, 4269–4277.

Ostland, V. E., Byrne, P. J., Hoover, G. & Ferguson, H. W. (2000). Necrotic myositis of rainbow trout, *Oncorhynchus mykiss* (Walbaum): Proteolytic characteristics of a crude extracellular preparation from *Flavobacterium psychrophilum*. *J Fish Dis* **23**, 329–336.

Ou, Y. Y., Gromiha, M. M., Chen, S. A. & Suwa, M. (2008). TMBETADISC-RBF: Discrimination of B-barrel membrane proteins using RBF networks and PSSM profiles. *Comput Biol Chem* **32**, 227–231.

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B. & other authors. **(2014).** The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* **42**, 206–214.

Owens, L. (2012). *Aquaculture: farming aquatic animals and plants*, Second. (J. S. Lucas & P.

C. Southgate, Eds.).

Pan, X., Yang, Y. & Zhang, J. (2014). Molecular basis of host specificity in human pathogenic bacteria. *Emerg Microbes Infect* **3**.

Papadioti, A., Markoutsa, S., Vranakis, I., Tselentis, Y., Karas, M., Psaroulaki, A. & Tsiotis, G. (2011). A proteomic approach to investigate the differential antigenic profile of two *Coxiella burnetii* strains. *J Proteomics* **74**, 1150–1159. Elsevier B.V.

Pasquali, C., Fialka, I. & Huber, L. (1997). Preparative two-dimensional gel electrophoresis of membrane proteins. *Electrophoresis* **18**, 2573–2581.

Pautsch, A. & Schulz, G. E. (2000). High-resolution structure of the OmpA membrane domain. *J Mol Biol* **298**, 273–282.

Peng, X., Xu, C., Ren, H. & Lin, X. (2005). Proteomic analysis of the sarcosine-insoluble outer membrane fraction of *Pseudomonas aeruginosa* responding to ampicilin, kanamycin, and tetracycline. *J Proteome Res* **4**, 2257–2265.

Perry, R. D. & Brubaker, R. R. (1979). Accumulation of iron by *Yersiniae*. *J Bacteriol* **137**, 1290–1298.

Pförtner, H., Wagner, J., Surmann, K., Hildebrandt, P., Ernst, S., Bernhardt, J., Schurmann, C., Gutjahr, M., Depke, M. & other authors. **(2013).** A proteomics workflow for quantitative and time-resolved analysis of adaptation reactions of internalized bacteria. *Methods* **61**, 244–50.

Pförtner, H., Burian, M. S., Michalik, S., Depke, M., Hildebrandt, P., Dhople, V. M., Pané-Farré, J., Hecker, M., Schmidt, F. & Völker, U. (2014). Activation of the alternative sigma factor SigB of *Staphylococcus aureus* following internalization by epithelial cells - an in vivo proteomics perspective. *Int J Med Microbiol* **304**, 177–87. Elsevier GmbH.

Phadke, N. D., Molloy, M. P., Steinhoff, S. A., Ulintz, P. J., Andrews, P. C. & Maddock, J. R. (2001). Analysis of the outer membrane proteome of *Caulobacter crescentus* by two-dimensional electrophoresis and mass spectrometry. *Proteomics* **1**, 705–720.

Pinsky, B. A. & Banaei, N. (2008). Multiplex real-time PCR assay for rapid identification of *Mycobacterium tuberculosis* complex members to the species level. *J Clin Microbiol* **46**, 2241–6.

Pizarro-Cerdá, J. & Cossart, P. (2006). Bacterial adhesion and entry into host cells. *Cell* **124**, 715–27.

Polissi, A. & Sperandeo, P. (2014). The lipopolysaccharide export pathway in *Escherichia coli*: Structure, organization and regulated assembly of the Lpt machinery. *Mar Drugs* **12**, 1023–1042.

Poobalane, S., Thompson, K. D., Diab, A., Ardó, L., Jeney, G. & Adams, A. (2008). Protein expression by *Aeromonas hydrophila* during growth *in vitro* and *in vivo*. *Microb Pathog* **45**, 60–69.

Pope, L. M., Reed, K. E., Payne, S. M. & Payne, S. M. (1995). Increased protein secretion and adherence to HeLa cells by *Shigella spp .* following growth in the presence of bile salts. *Microbiology* **63**, 3642–3648.

Post, G. (1987). Pathogenic *Yersinia ruckeri*, Enteric redmouth disease (Yersiniosis). In *Textb Fish Heal*, pp. 47–51. THF Publications.

Pratt, L. A. & Kolter, R. (1998). Genetic analysis of *Escherichia coli* biofilm formation: Roles of flagella, motility, chemotaxis and type I pili. *Mol Microbiol* **30**, 285–293.

Pratt, L. A., Hsing, W., Gibson, K. E. & Silhavy, T. J. (1996). From acids to osmZ: Multiple

factors influence synthesis of the OmpF and OmpC porins in *Escherichia coli*. *Mol Microbiol* **20**, 911–917.

Prouty, A. M. & Gunn, J. S. (2000). *Salmonella enterica* Serovar Typhimurium invasion Is repressed in the presence of bile. *Infect Immun* **68**, 6763–6769.

Pugsley, A. P. (1993). The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev* **57**, 50–108.

Pujol, C., Eugène, E., Marceau, M. & Nassif, X. (1999). The meningococcal PilT protein is required for induction of intimate attachment to epithelial cells following pilus-mediated adhesion. *Proc Natl Acad Sci U S A* **96**, 4017–4022.

Pujol, C. & Bliska, J. B. (2005). Turning *Yersinia pathogenesis* outside in: subversion of macrophage function by intracellular *Yersiniae*. *Clin Immunol* **114**, 216–26.

Pyle, S. W. & Schill, W. B. (1985). Rapid serological analysis of bacterial lipopolysaccharides by electrotransfer to nitrocellulose. *J Immunol Methods* **85**, 371–82.

Pyle, S. W., Ruppenthal, T., Cipriano, R. & Shotts, E. B. (1987). Further characterization of biochemical and serological characteristics of *Yersinia ruckeri* from different geographic areas. *Microbios Lett*.

Qi, W. (2002). Social and economic impacts of aquatic animal health problems in aquaculture in China. *FAO Fish Tech Pap* **2**, 55–61.

Raida, M. K., Larsen, J. L., Nielsen, M. E. & Buchmann, K. (2003). Short communication Enhanced resistance of rainbow trout , *Oncorhynchus mykiss* (Walbaum), against *Yersinia ruckeri* challenge following oral administration of *Bacillus subtilis* and *Bacillus licheniformis* (BioPlus2B). *J Fish Dis* **26**, 495–498.

Raida, M. K. & Buchmann, K. (2009). Innate immune response in rainbow trout (*Oncorhynchus mykiss*) against primary and secondary infections with *Yersinia ruckeri* O1. *Dev Comp Immunol* **33**, 35–45.

Rappuoli, R. (2000). Reverse vaccinology. *Curr Opin Microbiol* 445–450.

Rassam, P., Copeland, N. A., Birkholz, O., Tóth, C., Chavent, M., Duncan, A. L., Cross, S. J., Housden, N. G., Kaminska, R. & other authors. **(2015).** Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria. *Nature*.

Reading, N. C., Rasko, D. A., Torres, A. G. & Sperandio, V. (2009). The two-component system QseEF and the membrane protein QseG link adrenergic and stress sensing to bacterial pathogenesis. *Proc Natl Acad Sci U S A* **106**, 5889–5894.

Reddington, K., O'Grady, J., Dorai-Raj, S., Niemann, S., van Soolingen, D. & Barry, T. (2011). A novel multiplex real-time PCR for the identification of mycobacteria associated with zoonotic tuberculosis. *PLoS One* **6**, e23481.

Reyes, R. E., González, C. R., Jiménez, R. C., Herrera, M. O. & Andrade, A. A. (2012). Mechanisms of O-antigen structural variation of bacterial lipopolysaccharide (LPS). In *The Complex World of Polysaccharides*, pp. 71–98. Edited by D. N. Karunaratne. InTech.

Rhomberg, T. A., Karlberg, O., Mini, T., Zimny-Arndt, U., Wickenberg, U., Röttgen, M., Jungblut, P. R., Jenö, P., Andersson, S. G. E. & Dehio, C. (2004). Proteomic analysis of the sarcosine-insoluble outer membrane fraction of the bacterial pathogen *Bartonella henselae*. *Proteomics* **4**, 3021–33.

Ricci, D. P. & Silhavy, T. J. (2012). The Bam machine: A molecular cooper. *Biochim Biophys Acta - Biomembr* **1818**, 1067–1084. Elsevier B.V.

Ritter, A., Com, E., Bazire, A., Goncalves, M. D. S., Delage, L., Pennec, G. Le, Pineau, C.,

Dreanno, C., Compère, C. & Dufour, A. (2012). Proteomic studies highlight outer-membrane proteins related to biofilm development in the marine bacterium *Pseudoalteromonas* sp. D41. *Proteomics* **12**, 3180–3192.

Roberts, R. J. (2012). *Fish pathology*, 4th edn. (R. J. Roberts, Ed.). Wiley-Blackwell.

Rodgers, C. J. (1991). The usage of vaccination and antimicrobial agents for control of *Yersinia ruckeri*. *J Fish Dis* **14**, 291–301.

Rodgers, C. J. (1992). Development of a selective-differential medium for the isolation of *Yersinia ruckeri* and its application in epidemiological studies. *J Fish Dis* **15**, 243–254.

Romalde, J. L. & Toranzo, A. E. (1993). Pathological activities of *Yersinia ruckeri*, the enteric redmouth (ERM) bacterium. *FEMS Microbiol Lett* **112**, 291–9.

Romalde, J. L., Magarinos, B., Barja, J. L. & Toranzo, A. E. (1993). Antigenic and molecular characterization of *Yersinia ruckeri* proposal for a new intraspecies classification. *Syst Appl Microbiol* **16**, 411–419.

Ross, A. J. & Klontz, W. (1965). Oral immunization of rainbow trout (*Salmo gairdneri*) against an etiological agent of 'redmouth disease'. *J Fish Res Board Canada* **22**, 3–9.

Ross, A. J., Rucker, R. R. & Ewing, W. H. (1966). Description of a bacterium associated with redmouth disease of rainbow trout (*Salmo gairdneri*). *Can J Microbiol* **12**, 763–70.

Rozen, Y. & Belkin, S. (2001). Survival of enteric bacteria in seawater. *FEMS Microbiol Rev* **25**, 513–529.

Rucker, R. R. (1966). Redmouth disease of rainbow trout (*Salmo giardneri*). *Bull Off Int Epizoot* **65**, 825–830.

Ruiz, N., Kahne, D. & Silhavy, T. J. (2006). Advances in understanding bacterial outer-membrane biogenesis. *Nat Rev Microbiol* **4**, 57–66.

Ryckaert, J., Bossier, P., D'Herde, K., Diez-Fraile, A., Sorgeloos, P., Haesebrouck, F. & Pasmans, F. (2010). Persistence of *Yersinia ruckeri* in trout macrophages. *Fish Shellfish Immunol* **29**, 648–55. Elsevier Ltd.

Sakai, T., Kanai, K., Osatomi, K. & Yoshikoshi, K. (2003). Identification of a 19.3-kDa protein in MRHA-positive *Edwardsiella tarda*: Putative fimbrial major subunit. *FEMS Microbiol Lett* **226**, 127–133.

Saleh, M., Soliman, H. & El-Matbouli, M. (2008). Loop-mediated isothermal amplification as an emerging technology for detection of *Yersinia ruckeri* the causative agent of Enteric red mouth disease in fish. *BMC Vet Res* **4**, 31.

Samuel, G. & Reeves, P. (2003). Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydr Res* **338**, 2503–2519.

Sato, M., Arita, M., Honda, T. & Miwatani, T. (1989). Characterization of a pilus produced by *Aeromonas hydrophila*. *FEMS Microbiol Lett* **50**, 325–329.

Sauter, R. W., Meyer, E. A., Williams, C. & Celnik, B. (1985). *Etiology of early lifestage diseases: Final report*.

Schill, W. B., Phelps, S. R. & Pyle, S. W. (1984). Multilocus electrophoretic assessment of the genetic structure and diversity of *Yersinia ruckeri*. *Appl Environ Microbiol* **48**, 975–9.

Schirmer, T., Keller, T. A., Wang, Y. F. & Rosenbusch, J. P. (1995). Structural basis for sugar translocation through maltoporin channels at 3.1 A resolution. *Science* **267**, 512–514.

Schuhmacher, D. A. & Klose, K. E. (1999). Environmental signals modulate ToxT-dependent virulence factor expression in *Vibrio cholerae*. *J Bacteriol* **181**, 1508–1514.

Scott, C. J. W., Austin, B., Austin, D. A. & Morris, P. C. (2013). Non-adjuvanted flagellin elicits a non-specific protective immune response in rainbow trout (*Oncorhynchus mykiss*, Walbaum) towards bacterial infections. *Vaccine* **31**, 3262–7. Elsevier Ltd.

Secades, P. & Guijarro, J. A. (1999). Purification and characterization of an extracellular protease from the fish pathogen *Yersinia ruckeri* and effect of culture conditions on production. *Appl Environ Microbiol* **65**, 3969–3975.

Seib, K. L., Zhao, X. & Rappuoli, R. (2012). Developing vaccines in the era of genomics: a decade of reverse vaccinology. *Clin Microbiol Infect* **18 Suppl 5**, 109–16.

Sengupta, C., Ray, S. & Chowdhury, R. (2014). Fine tuning of virulence regulatory pathways in enteric bacteria in response to varying bile and oxygen concentrations in the gastrointestinal tract. *Gut Pathog* **6**, 38.

Sharma, R., Dill, B. D., Chourey, K., Shah, M., VerBerkmoes, N. C. & Hettich, R. L. (2012). Coupling a detergent lysis/cleanup methodology with intact protein fractionation for enhanced proteome characterization. *J Proteome Res* **11**, 6008–18.

Shehata, T. E. & Marr, A. G. (1971). Effect of Nutrient Concentration on the Growth of *Escherichia coli*. *J Bacteriol* **107**, 210–216.

Sherry, A. E., Inglis, N. F., Stevenson, A., Fraser-Pitt, D., Everest, P., Smith, D. G. E. & Roberts, M. (2011). Characterisation of proteins extracted from the surface of *Salmonella Typhimurium* grown under SPI-2-inducing conditions by LC-ESI/MS/MS sequencing. *Proteomics* **11**, 361–70.

Shevchuk, O., Jäger, J. & Steinert, M. (2011). Virulence properties of the *Legionella pneumophila* cell envelope. *Front Microbiol* **2**, 1–12.

Shi, L., Sohaskey, C. D., Kana, B. D., Dawes, S., North, R. J., Mizrahi, V. & Gennaro, M. L. (2005). Changes in energy metabolism of *Mycobacterium tuberculosis* in mouse lung and under in vitro conditions affecting aerobic respiration. *Proc Natl Acad Sci U S A* **102**, 15629–15634.

Singh, B., Fleury, C., Jalalvand, F. & Riesbeck, K. (2012). Human pathogens utilize host extracellular matrix proteins laminin and collagen for adhesion and invasion of the host. *FEMS Microbiol Rev* **36**, 1122–1180.

Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D. & Silhavy, T. J. (2007). Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in *Escherichia coli*. *Proc Natl Acad Sci U S A* **104**, 6400–6405.

Skurnik, M. & Bengoechea, J. A. (2009). Genetics and Regulation of Bacterial Lipopolysaccharide Synthesis. In *Bact Polysaccharides Curr Innov Futur Trends*, p. 27. Edited by M. Ullrich. Norfolk, UK: Caister Academic Press.

Smith, A. M., Goldring, O. L. & Dear, G. (1987). The production and methods of use of polyclonal antisera to the pathogenic organisms *Aeromonas salmonicida*, *Yersinia ruckeri* and *Renibacterium salmonicida*. *J Fish Biol* **31**, 225–226.

Snieszko, S. F. (1974). The effects of environmental stress on outbreaks of infectious diseases of fishes. *J Fish Biol* **6**, 197–208.

Sonntag, I., Schwarz, H. & Hirota, Y. (1978). Cell envelope and shape of *Escherichia coli* : multiple mutants missing the outer membrane lipoprotein and other major outer membrane proteins. *J Bacteriol* **136**, 280–285.

Steele-mortimer, O. (2009). The *Salmonella*-containing Vacuole – Moving with the Times.

Curr Opin Microbiol **11**, 38–45.

Stekhoven, D. J., Omasits, U., Quebatte, M., Dehio, C. & Ahrens, C. H. (2014). Proteome-wide identification of predominant subcellular protein localizations in a bacterial model organism. *J Proteomics* **99**, 123–37. Elsevier B.V.

Stevenson, P. (2007). *Closed Waters: the welfare of farmed atlantic salmon, rainbow trout, atlantic cod and atlantic halibut*.

Stevenson, R. M. W. (1984). Serological variation among *Yersinia ruckeri* strains. *J Fish Dis* **7**, 247–254.

Stevenson, R. M. W. & Airdrie, D. W. (1984). Isolation of *Yersinia ruckeri* Bacteriophages. *Appl Environ Microbiol* **47**, 1201–5.

Stevenson, R. M. W. & Daly, J. G. (1982). Biochemical and serological characteristics of Ontario isolates of *Yersinia ruckeri*. *Can J Fish Aquac Sci* 870–876.

Stevenson, R. M. W., Tremblay, O. & Kropinski, A. M. (2015). The genome of *Yersinia ruckeri* strain BC74.

Strepparava, N., Wahli, T., Segner, H. & Petrini, O. (2014). Detection and quantification of *Flavobacterium psychrophilum* in water and fish tissue samples by quantitative real time PCR. *BMC Microbiol* **14**, 105.

Sugawara, E. & Nikaido, H. (1992). Pore-forming activity of OmpA protein of *Escherichia coli*. *J Biol Chem* **267**, 2507–2511.

Sukumaran, S. K., Shimada, H., Nemani, V. & Prasadarao, N. V. **(2003).** Entry and intracellular replication of *Escherichia coli* K1 in macrophages require expression of outer membrane protein A **71**, 5951–5961.

Tallant, T., Deb, A., Kar, N., Lupica, J., de Veer, M. J. & DiDonato, J. A. (2004). Flagellin acting via TLR5 is the major activator of key signaling pathways leading to NF-kappa B and proinflammatory gene program activation in intestinal epithelial cells. *BMC Microbiol* **4**, 33.

Tamm, L. K., Hong, H. & Liang, B. (2004). Folding and assembly of B-barrel membrane proteins. *Biochim Biophys Acta - Biomembr* **1666**, 250–263.

Tanca, A., Biosa, G., Pagnozzi, D., Addis, M. F. & Uzzau, S. (2013). Comparison of detergent-based sample preparation workflows for LTQ-Orbitrap analysis of the *Escherichia coli* proteome. *Proteomics* **13**, 2597–607.

Tang, K. W. & Grossart, H.-P. (2007). Iron effects on colonization behavior, motility, and enzymatic activity of marine bacteria. *Can J Microbiol* **53**, 968–974.

Taniguchi, Y., Choi, P. J., Li, G., Chen, H., Babu, M., Hearn, J., Emili, A. & Xie, X. S. (2010). Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science (80-)* **329**, 533–539.

Tarr, P. I., Schoening, L. M., Yea, Y. L., Ward, T. R., Jelacic, S. & Whittam, T. S. (2000). Acquisition of the rfb-gnd cluster in evolution of *Escherichia coli* O55 and O157. *J Bacteriol* **182**, 6183–91.

Tatner, M. F. & Horne, M. (1985). The effects of vaccine dilution, length of immersion time, and booster vaccinations on the protection levels induced by direct immersion vaccination of brown trout, *Salmo trutta*, with *Yersinia ruckeri* (ERM) vaccine. *Aquaculture* **46**, 11–18.

Temprano, A., Yugueros, J., Hernanz, C., Sµnchez, M., Berzal, B. & Luengo, J. M. (2001). Rapid identification of *Yersinia ruckeri* by PCR amplification of yruI-yruR quorum sensing. *J Fish Dis* **24**, 253–261.

Temprano, A., Riaño, J., Yugueros, J., González, P., de Castro, L., Villena, A., Luengo, J. M. & Naharro, G. (2005). Potential use of a *Yersinia ruckeri* O1 auxotrophic aroA mutant as a live attenuated vaccine. *J Fish Dis* **28**, 419–27.

Thanassi, D. G. & Hultgren, S. J. (2000). Multiple pathways allow protein secretion across the bacterial outer membrane. *Curr Opin Cell Biol* **12**, 420–30.

Thorsen, B. K., Enger, O., Norland, S. & Hoff, K. A. (1992). Long-term starvation survival of *Yersinia ruckeri* at different salinities studied by microscopical and flow cytometric methods. *Appl Environ Microbiol* **58**, 1624–8.

Tinsley, J. W. (2010). *Studies on the pathogenicity of* Yersinia ruckeri *biotype 2 to rainbow trout (Oncorhynchus mykiss, Walbaum)*. Heriot-Watt University, Edinburgh, UK (PhD Thesis).

Tinsley, J. W., Austin, D. A., Lyndon, A. R. & Austin, B. (2011). Novel non-motile phenotypes of *Yersinia ruckeri* suggest expansion of the current clonal complex theory. *J Fish Dis* **34**, 311–7.

Tobback, E., Decostere, A., Hermans, K., Haesebrouck, F. & Chiers, K. (2007). Review of *Yersinia ruckeri* infections in salmonid fish. *J Fish Dis* **30**, 257–268.

Tobback, E. (2009). *Early pathogenesis of Yersinia ruckeri infections in rainbow trout (Oncorhynchus mykiss , Walbaum)*. Ghent University (PhD Thesis).

Tokuda, H. & Matsuyama, S. I. (2004). Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochim Biophys Acta - Mol Cell Res* **1693**, 5–13.

Topic Popovic, N., Coz-Rakovac, R. & Strunjak-Perovic, I. (2007). Commercial phenotypic tests (API 20E) in diagnosis of fish bacteria: A review. *Vet Med (Praha)* **52**, 49–53.

Toranzo, A. E. (2004). Report about fish bacterial diseases. *Options Méditerranéennes Série B Etudes Rech* **48**, 49–89.

Torres, A. G. & Kaper, J. B. (2003). Multiple elements controlling adherence of Enterohemorrhagic *Escherichia coli* O157 : H7 to HeLa cells. *Infect Immun* **71**, 4985–4995.

Tsai, C. M. & Frasch, C. E. (1982). A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. *Anal Biochem* **119**, 115–9.

Tsoi, S., Park, K. C., Kay, H. H., O'Brien, T. J., Podor, E., Sun, G., Douglas, S. E., Brown, L. L. & Johnson, S. C. (2006). Identification of a transcript encoding a soluble form of toll-like receptor 5 (TLR5) in Atlantic salmon during *Aeromonas salmonicida* infection. *Vet Immunol Immunopathol* **109**, 183–7.

Tsujita, T., Tsukada, H., Nakao, M., Oshiumi, H., Matsumoto, M. & Seya, T. (2004). Sensing bacterial flagellin by membrane and soluble orthologs of Toll-like receptor 5 in rainbow trout (*Onchorhynchus mykiss*). *J Biol Chem* **279**, 48588–97.

Uliczka, F., Kornprobst, T., Eitel, J., Schneider, D. & Dersch, P. (2009). Cell invasion of *Yersinia pseudotuberculosis* by invasin and YadA requires protein kinase C, phospholipase C-y1 and Akt kinase. *Cell Microbiol* **11**, 1782–1801.

Uliczka, F., Pisano, F., Schaake, J., Stolz, T., Rohde, M., Fruth, A., Strauch, E., Skurnik, M., Batzilla, J. & other authors. **(2011).** Unique cell adhesion and invasion properties of *Yersinia enterocolitica* O:3, the most frequent cause of human yersiniosis. *PLoS Pathog* **7**.

Vázquez-Juárez, R. C., Barrera-Saldaña, H. a., Hernández-Saavedra, N. Y., Gómez-Chiarri, M. & Ascencio, F. (2003). Molecular cloning, sequencing and characterization of omp48, the gene encoding for an antigenic outer membrane protein from *Aeromonas*

veronii. *J Appl Microbiol* **94**, 908–918.

Verner-Jeffreys, D. W., Haig, S. J., Welch, T. J., Pond, M. J., Davies, R. L. & Gardner, R. (2011). Characterisation of a serotype O1 *Yersinia ruckeri* isolate from the Isle of Man : further evidence that O antigen serotype is not a reliable indicator of virulence **31**, 86–91.

Verschuere, L., Rombaut, G., Sorgeloos, P. & Verstraete, W. (2000). Probiotic bacteria as biological control agents in aquaculture. *Microbiol Mol Biol Rev* **64**, 655–671.

Vogt, J. & Schulz, G. E. (1999). The structure of the outer membrane protein OmpX from *Escherichia coli* reveals possible mechanisms of virulence. *Structure* **7**, 1301–9.

Walker, P. J. & Winton, J. R. (2010). Emerging viral diseases of fish and shrimp. *Vet Res* **41**.

Wall, D. & Kaiser, D. (1999). Type IV pili and cell motility. *Mol Microbiol* **32**, 1–10.

Walters, M. S. & Mobley, H. L. T. (2009). Identification of uropathogenic *Escherichia coli* surface proteins by shotgun proteomics. *J Microbiol Methods* **78**, 131–135. Elsevier B.V.

Waltman, W. D. & Shotts, E. B. (1984). A medium for the isolation and differentiation of *Yersinia ruckeri*. *Can J Fish Aquac Sci* **41**.

Wang, C., Liu, Y., Li, H., Xu, W., Zhang, H. & Peng, X.-X. (2012). Identification of plasma-responsive outer membrane proteins and their vaccine potential in *Edwardsiella tarda* using proteomic approach. *J Proteomics* **75**, 1263–75. Elsevier B.V.

Wang, G., Ge, Z., Rasko, D. a. & Taylor, D. E. (2000). Lewis antigens in *Helicobacter pylori*: Biosynthesis and phase variation. *Mol Microbiol* **36**, 1187–1196.

Wang, K., Liu, T., Wang, J., Chen, D., Wu, X., Jiang, J. & Liu, J. (2015). Complete genome sequence of the fish pathogen *Yersinia ruckeri* strain SC09 , isolated from diseased *Ictalurus punctatus* in China. *Genome Announc* **3**, 2014–2015.

Watnick, P. I. & Kolter, R. (2012). Steps in the development of a *Vibrio cholerae* El Tor biofilm. *Mol Microbiol* **29**, 997–1003.

Watson, E., Alberdi, M. P., Inglis, N. F., Lainson, A., Porter, M. E., Manson, E., Imrie, L., Mclean, K. & Smith, D. G. E. (2014). Proteomic analysis of *Lawsonia intracellularis* reveals expression of outer membrane proteins during infection. *Vet Microbiol* **174**, 448–455. Elsevier B.V.

Webb, C. T., Heinz, E. & Lithgow, T. (2012). Evolution of the B-barrel assembly machinery. *Trends Microbiol* **20**, 612–620. Elsevier Ltd.

Webley, D. M. (1953). A simple method for producing microcultures in hanging drops with special reference to organisms utilizing oils. *J Gen Microbiol* **8**, 66–71.

Weinberg, E. D. (2004). Suppression of bacterial biofilm formation by iron limitation. *Med Hypotheses* **63**, 863–865.

Weinberger, D. M., Trzciński, K., Lu, Y.-J., Bogaert, D., Brandes, A., Galagan, J., Anderson, P. W., Malley, R. & Lipsitch, M. (2009). Pneumococcal capsular polysaccharide structure predicts serotype prevalence. *PLoS Pathog* **5**, e1000476.

Welch, T. J., Verner-Jeffreys, D. W., Dalsgaard, I., Wiklund, T., Evenhuis, J. P., Cabrera, J. A. G., Hinshaw, J. M., Drennan, J. D. & LaPatra, S. E. (2011). Independent emergence of *Yersinia ruckeri* biotype 2 in the United States and Europe. *Appl Environ Microbiol* **77**, 3493–9.

Wheeler, R. (2009). Outer membrane proteomics of *Pasteurella multocida* isolates to identify putative host-specificity determinants. *Biosci Horizons* **2**, 1–12.

Wheeler, R. W., Davies, R. L., Dalsgaard, I., Garcia, J., Welch, T. J., Wagley, S., Bateman,

K. S. & Verner-Jeffreys, D. W. (2009). *Yersinia ruckeri* biotype 2 isolates from mainland Europe and the UK likely represent different clonal groups. *Dis Aquat Organ* **84**, 25–33.

White, D. G., Wilson, R. A., San Gabriel, A., Saco, M. & Whittam, T. S. (1990). Genetic relationships among strains of avian *Escherichia coli* associated with swollen-head syndrome. *Infect Immun* **58**, 3613–3620.

Whitfield, C. (1995). Biosynthesis of lipopolysaccharide O antigens. *Trends Microbiol* **3**, 178–85.

Whittam, T. S. (1995). Genetic population structure and pathogenicity in enteric bacteria. In *Popul Genet Bact*, pp. 217–245. Edited by S. Baumberg, J. P. W. Young, E. M. H. Wellington & J. R. Saunders. Cambridge, United Kingdom: University Press.

Willems, R. J. L., Geuijen, C., Van Der Heide, H. G. J., Renauld, G., Bertin, P., Van Den Akker, W. M. R., Locht, C. & Mooi, F. R. (1994). Mutational analysis of the *Bordetella pertussis* fim/fha gene cluster: Identification of a gene with sequence similarities to haemolysin accessory genes involved in export of FHA. *Mol Microbiol* **11**, 337–347.

Williams, J. N., Skipp, P. J., Humphries, H. E., Christodoulides, M., O'Connor, C. D. & Heckels, J. E. (2007). Proteomic analysis of outer membranes and vesicles from wild-type serogroup B *Neisseria meningitidis* and a lipopolysaccharide-deficient mutant. *Infect Immun* **75**, 1364–72.

Willumsen, B. (1987). Birds and wild fish as potential vectors of *Yersinia ruckeri*. *J Fish Dis* **40**, 275–277.

Wisniewski, J. & Zougman, A. (2009). Universal sample preparation method for proteome analysis. *Nat Methods* **6**, 3–7.

Wooldridge, K. G. & Williams, P. H. (1993). Iron uptake mechanisms of pathogenic bacteria. *FEMS Microbiol Rev* **12**, 325–48.

Wu, C. C. & Yates, J. R. (2003). The application of mass spectrometry to membrane proteomics. *Nat Biotechnol* **21**, 262–7.

Wurpel, D. J., Moriel, D. G., Totsika, M., Easton, D. M. & Schembri, M. A. (2015). Comparative analysis of the uropathogenic *Escherichia coli* surface proteome by tandem mass-spectrometry of artificially induced outer membrane vesicles. *J Proteomics* **115**, 93–106. Elsevier B.V.

Xu, C., Wang, S., Ren, H., Lin, X., Wu, L. & Peng, X. (2005). Proteomic analysis on the expression of outer membrane proteins of *Vibrio alginolyticus* at different sodium concentrations. *Proteomics* **5**, 3142–52.

Xu, C., Lin, X., Ren, H., Zhang, Y., Wang, S. & Peng, X. (2006). Analysis of outer membrane proteome of *Escherichia coli* related to resistance to ampicillin and tetracycline. *Proteomics* **6**, 462–73.

Yadav, S. K., Sahoo, P. K. & Dixit, A. (2014). Characterization of immune response elicited by the recombinant outer membrane protein OmpF of *Aeromonas hydrophila*, a potential vaccine candidate in murine model. *Mol Biol Rep* **41**, 1837–1848.

Yamaguchi, K., Yu, F. & Inouye, M. (1988). A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. *Cell* **53**, 423–432.

Ying, T., Wang, H., Li, M., Wang, J., Wang, J., Shi, Z., Feng, E., Liu, X., Su, G. & other authors. **(2005).** Immunoproteomics of outer membrane proteins and extracellular proteins of *Shigella flexneri* 2a 2457T. *Proteomics* **5**, 4777–93.

Yoon, S. H., Han, M.-J., Lee, S. Y., Jeong, K. J. & Yoo, J.-S. (2003). Combined transcriptome and proteome analysis of *Escherichia coli* during high cell density culture. *Biotechnol*

Bioeng **81**, 753–767.

Yoshida, T., Qin, L., Egger, L. a. & Inouye, M. (2006). Transcription regulation of ompF and ompC by a single transcription factor, OmpR. *J Biol Chem* **281**, 17114–17123.

Yu, C.-S., Chen, Y.-C., Lu, C.-H. & Hwang, J.-K. (2006). Prediction of protein subcellular localization. *Proteins Struct Funct Bioinforma* **64**, 643–651.

Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., Dao, P., Cenk Sahinalp, S., Ester, M. & other authors. **(2010).** PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**, 1608–1615.

Yu, Y., Zeng, H., Lyons, S., Carlson, A., Merlin, D., Neish, A. S. & Gewirtz, A. T. (2003). TLR5-mediated activation of p38 MAPK regulates epithelial IL-8 expression via posttranscriptional mechanism. *Am J Physiol Gastrointest Liver Physiol* **285**, G282–90.

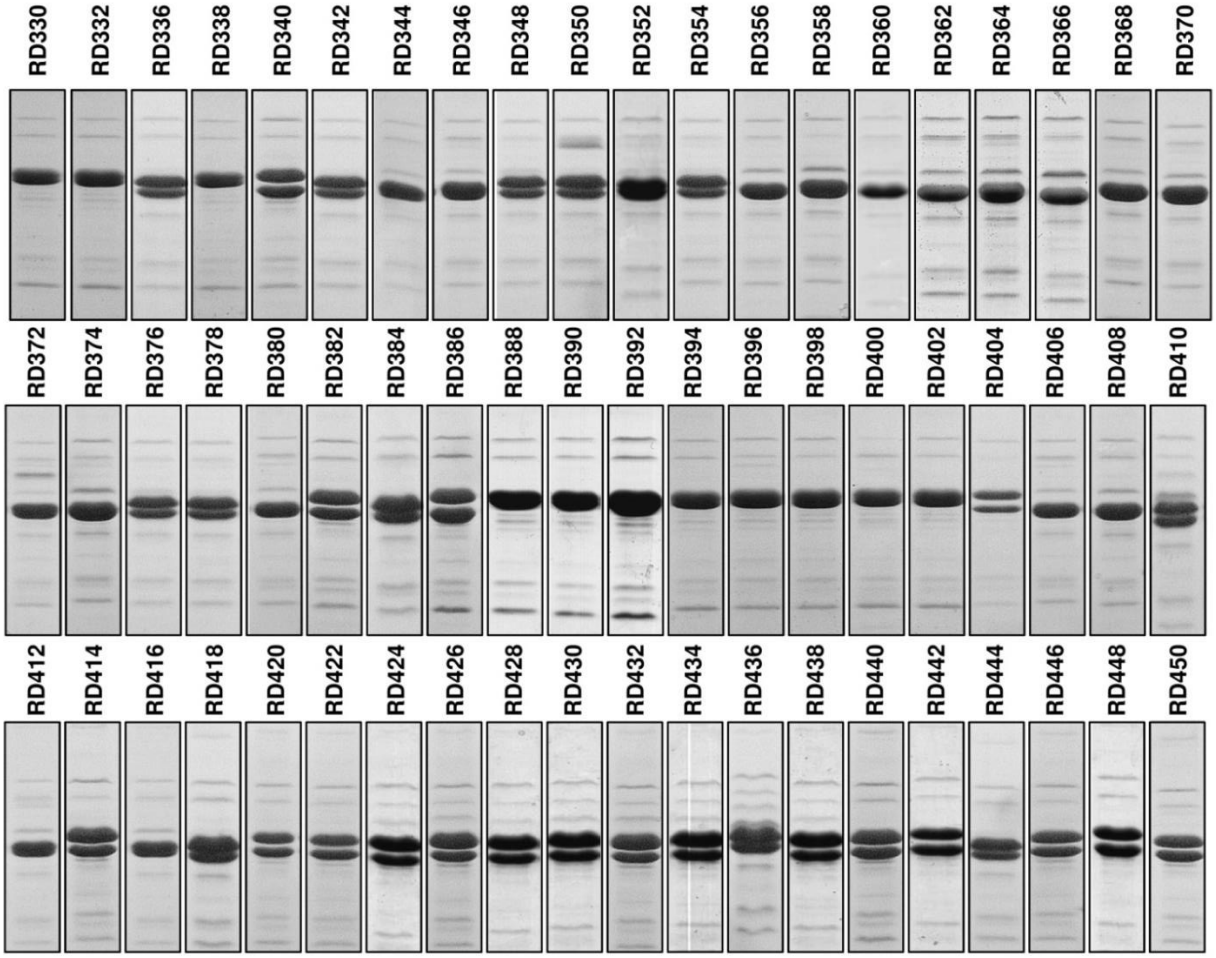
Zhang, L., Wen, Y., Li, Y., Wei, X., Yan, X., Wen, X., Wu, R., Huang, X., Huang, Y. & other authors. **(2014).** Comparative proteomic analysis of the membrane proteins of two *Haemophilus parasuis* strains to identify proteins that may help in habitat adaptation and pathogenesis. *Proteome Sci* **12**, 38.

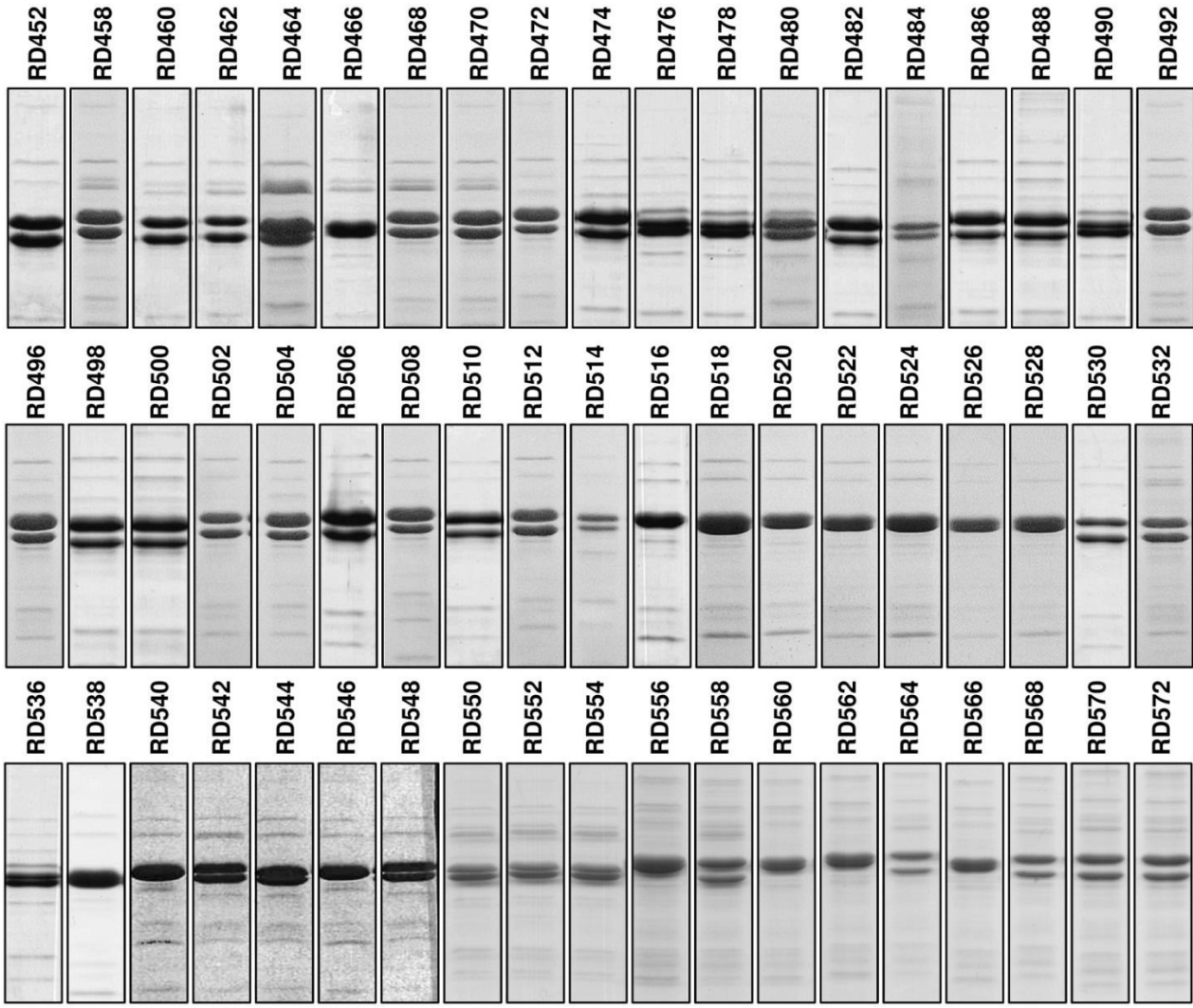
Zhang, W., Shao, J., Liu, G., Tang, F., Lu, Y., Zhai, Z., Wang, Y., Wu, Z., Yao, H. & Lu, C. (2011a). Immunoproteomic analysis of bacterial proteins of *Actinobacillus pleuropneumoniae* serotype 1. *Proteome Sci* **9**, 32. BioMed Central Ltd.

Zhang, Z., Niu, C., Storset, A., Bøgwald, J. & Dalmo, R. A. (2011b). Comparison of *Aeromonas salmonicida* resistant and susceptible salmon families: a high immune response is beneficial for the survival against *Aeromonas salmonicida* challenge. *Fish Shellfish Immunol* **31**, 1–9. Elsevier Ltd.

Zhou, L., Ying, W., Han, Y., Chen, M., Yan, Y., Li, L., Zhu, Z., Zheng, Z., Jia, W. & other authors. **(2012).** A proteome reference map and virulence factors analysis of *Yersinia pestis* 91001. *J Proteomics* **75**, 894–907. Elsevier B.V.

Chapter 8 Appendices




RD452
RD458
RD460
RD462
RD464
RD466
RD468
RD470
RD472
RD474
RD476
RD478
RD480
RD482
RD484
RD486
RD488
RD490
RD492
RD496
RD498
RD500
RD502
RD504
RD506
RD508
RD510
RD512
RD514
RD516
RD518
RD520
RD522
RD524
RD526
RD528
RD530
RD532
RD536
RD538
RD540
RD542
RD544
RD546
RD548
RD550
RD552
RD554
RD556
RD558
RD560
RD562
RD564
RD566
RD568
RD570
RD572

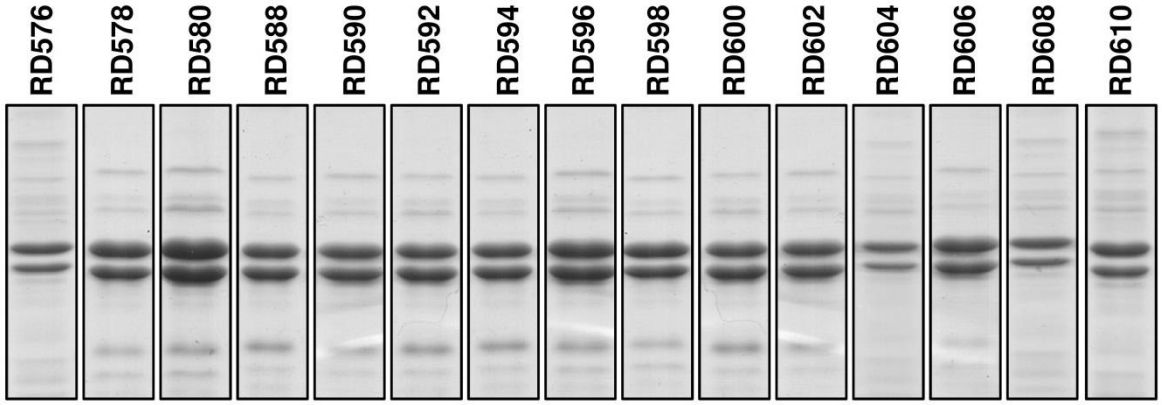


Fig. 8.1 OMP profiles of all isolates of *Y. ruckeri* included in this study.
The OM proteome of 135 isolates of *Y. ruckeri*, collected over a 14 year period (2001-2015) in the UK are presented in this figure. OMs were prepared by 0.5% *Sarkosyl* extraction and separated by SDS-PAGE. The region ranging from approximately 30 kDa to 60 kDa is shown, as this portion is used for OM-typing.

Table 8.1 Gel-free proteomics of isolates of *Y. ruckeri* under conditions that mimic *in vivo* and environmental growth.
Isolates RD6 **(A)**, RD28 **(B)**, RD64 **(C)** and RD124 **(D)** were recovered from rainbow trout, while isolates RD354 **(E)**, RD366 **(F)**, RD382 **(G)** and RD420 **(H)** were recovered from Atlantic salmon.

(A)

Accession	Protein	MW (Da)	RD6														
			Replicates					emPAI					Peptides				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
gi|238708193	Brauns (Lpp)	6740	3	3	3	3	3	33.6	33.7	24.0	12.8	6.3	37	59	54	37	33
gi|238707374	OmpA	37846	3	3	3	3	3	2.6	4.4	2.2	2.0	1.1	126	222	65	89	71
gi|238705740	Pal	18149	3	3	3	3	3	1.8	3.4	2.6	2.9	1.5	10	32	15	24	27
gi|238708170	SlyB	15382	3	3	3	3	3	1.0	2.1	2.0	3.6	0.2	5	21	8	18	1
gi|238706031	LptD	89769	3	3	2	3	1	0.1	0.2	0.1	0.3	0.0	9	11	8	13	10
gi|238706768	LptE	21133	3	3	3	3	-	0.2	0.6	0.3	0.4	-	2	8	2	6	-
gi|238706888	Slp Family	20689	3	3	3	3	-	0.4	0.5	0.7	0.4	-	4	7	6	4	-
gi|238706959	BamA/YaeT	88022	3	3	3	3	-	1.2	1.3	1.3	1.5	-	38	69	44	65	-
gi|238707860	BamB	42358	3	3	3	3	-	0.2	0.7	0.9	2.0	-	4	14	15	28	-
gi|238705699	BamC	38345	3	3	3	3	-	0.4	1.4	0.7	1.1	-	8	31	18	22	-
gi|238705321	BamD	27432	3	3	3	3	-	0.4	0.5	0.3	0.4	-	5	14	6	8	-
gi|238707927	BamE	9943	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708048	VacJ	28773	3	3	3	-	-	0.9	0.4	0.5	-	-	11	11	7	-	-
gi|238706986	MltA	42895	3	1	-	-	-	0.1	0.1	-	-	-	2	3	-	-	-
gi|238707253	MltC	39935	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706640	MltD	48282	-	-	-	1	-	-	-	-	0.1	-	-	-	-	3	-
gi|238707358	OmpF	40024	3	3	3	3	1	0.5	1.3	0.6	1.0	0.2	7	32	8	23	5
gi|238706321	TolC	53048	1	3	3	3	2	0.2	1.2	0.2	0.2	0.1	6	24	4	4	8
gi|238706508	OmpC.1	39947	3	3	2	3	-	0.2	0.3	0.2	0.4	-	3	9	3	10	-
gi|238707949	OmpC.2	40459	3	3	3	3	2	0.1	0.9	0.3	0.6	0.3	3	24	5	15	16
gi|238706822	OsmY	11124	3	3	2	2	-	0.8	3.1	0.7	1.0	-	6	20	2	6	-
gi|238706938	MetQ	29356	2	2	3	3	1	0.2	0.7	0.3	0.5	0.2	6	16	5	8	8
gi|238705360	TamA/YtfN	63394	2	3	2	3	-	0.1	0.1	0.2	0.2	-	8	5	7	10	-
gi|238706342	OM Siderophore receptor	73465	2	-	3	3	3	0.0	-	0.8	0.2	0.1	6	-	27	8	14
gi|238708209	HemR	72964	1	-	3	-	2	0.1	-	0.2	-	0.2	2	-	14	-	11
gi|238705851	FhuA	81724	1	-	3	1	1	0.1	-	1.2	0.0	0.4	6	-	56	6	29
gi|238706670	FiuA	77361	-	-	3	1	1	-	-	0.2	0.0	0.1	-	-	8	3	9
gi|238708284	Ton-B dependent receptor	76573	-	-	2	-	-	-	-	0.0	-	-	-	-	6	-	-
gi|238707022	HasR	89969	-	-	1	-	-	-	-	0.0	-	-	-	-	5	-	-
gi|238708022	OmpE	39925	3	3	-	1	-	0.1	0.1	-	0.1	-	3	3	-	5	-
gi|238707630	OprC	75132	2	-	2	-	-	0.0	-	0.0	-	-	3	-	4	-	-
gi|238706821	OmpW	23865	-	-	-	3	-	-	-	-	0.2	-	-	-		3	-
gi|238707009	BtuB	68957	1	-	-	1	-	0.1	-		0.2	-	2	-		3	-
gi|238708047	FadL	46035	-	1	-	1	-	-	0.1	-	0.3	-	-	3	-	8	-
gi|238705242	BcsC	126527	3	-	1	1	-	0.1	-	0.4	0.0	-	12	-	17	6	-
gi|238705556	Haemolysin activator	62594	2	2	1	1	-	0.1	0.1	0.0	0.1	-	4	3	-	2	-
gi|238705555	Hemolysin	48963	3	-	-	-	-	0.1	-	-	-	-	3	-	-	-	-
gi|238706248	Hemolysin activator	61801	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707066	Pep M37	50544	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705486	Phos. A1	32689	-	3	-	1	-	-	0.1	-	0.1	-	-	3	-	1	-
gi|238705976	Protease II	77905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|110007337	YhlB	62424	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705110	Possible Adhesin	155469	1	-	1	-	-	0.0	-	0.0	-	-	10	-	2	-	-
gi|238707370	PqiB/Mam7	59633	1	-	1	1	-	0.1	-	0.1	0.1	-	5	-	2	1	-
gi|238705561	AidA	44266	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705628	Fim. Usher	87773	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705013	Flagellin, partial	10509	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|498532474	Flagellin	43589	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706460	FlgE	44254	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706435	FliD	49054	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707768	YiaD	22348	3	3	3	3	2	0.3	0.8	0.2	0.7	0.4	7	8	3	16	11
gi|238707898	YfhG	35016	3	3	1	3	-	0.3	0.4	0.1	0.1	-	7	8	4	4	-
gi|238705190	YraP	20221	-	2	-	3	1	-	0.4	-	0.2	0.2	-	6	-	2	1
gi|238705961	CsgG	24712	2	1	3	1	-	0.2	0.1	0.1	0.3	-	10	4	3	4	-
gi|238707176	SecD	71766	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705788	TTSS protein	61602	1	-	-	-	-	0.1	-	-	-	-	2	-	-	-	-
gi|238705783	TTSS protein	45918	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705857	AmiB	62746	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705632	Blc	20801	-	-	1	-	-	-	-	0.2	-	-	-	-	2	-	-
gi|238705953	Mce-Like	91347	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706915	MipA	28023	1	1	2	-	-	0.1	0.3	0.1	-	-	2	5	3	-	-
gi|238707429	NagZ	99335	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707352	NlpD	34028	1	1	-	3	-	0.1	0.1	-	0.4	-	5	1	-	5	-
gi|238706941	NlpE/CutF	25245	-	3	1	1	-	-	0.1	0.7	0.1	-	-	2	5	2	-
gi|238708334	OmpX	17086	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|163883857	PtrA	107905	-	-	-	-	1	-	-	-	-	0.0	-	-	-	-	9
gi|238706761	RlpA	37591	-	-	-	-	1	-	-	-	-	0.1	-	-	-	-	6
gi|46369605	RupA	17881	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706180	Unknown	17544	-	-	-	1	-	-	-	-	0.2	-	-	-	-	2	-
gi|133904995	Unknown	200596	-	-	1	1	1	-	-	0.0	0.0	0.0	-	-	3	3	12
gi|238706101	Unknown	106480	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705250	Unknown	40804	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706025	Hypothetical	17964	3	-	2	1	-	0.2	-	0.2	0.2	-	4	-	3	8	-
gi|238705088	YeeJ	54634	3	-	2	-	-	0.1	-	0.1	-	-	3	-	1	-	-
gi|238706732	Hypothetical	31452	3	-	1	-	-	0.2	-	0.1	-	-	2	-	2	-	-
gi|238707408	Hypothetical	56778	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707442	Hypothetical	22794	-	-	-	2	-	-	-	-	0.2	-	-	-	-	1	-
gi|238707291	Hypothetical	54666	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

(B)

Accession	Protein	MW (Da)	RD28														
			Replicates					emPAI					Peptides				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
gi|238708193	Brauns (Lpp)	6740	3	3	3	3	3	28.5	30.3	53.3	21.8	18.6	42	46	72	43	53
gi|238707374	OmpA	37846	3	3	3	3	3	2.1	3.7	2.5	3.0	1.2	110	204	71	92	93
gi|238705740	Pal	18149	3	3	3	3	3	1.8	2.9	3.7	4.1	1.9	11	25	18	25	20
gi|238708170	SlyB	15382	3	3	3	3	1	1.2	1.4	2.4	2.9	2.3	6	11	8	14	10
gi|238706031	LptD	89769	2	2	3	3	1	0.0	0.1	0.1	0.2	0.1	8	7	5	7	9
gi|238706768	LptE	21133	2	3	2	3	-	0.2	0.4	0.2	0.4	-	2	5	2	4	-
gi|238706888	Slp Family	20689	3	1	3	3	1	0.3	0.2	0.8	0.3	0.6	3	1	6	3	3
gi|238706959	BamA/YaeT	88022	3	3	3	3	2	1.1	1.0	1.3	1.5	0.5	36	47	39	59	21
gi|238707860	BamB	42358	3	3	3	3	1	0.2	0.7	1.2	1.6	0.7	5	12	14	23	17
gi|238705699	BamC	38345	3	3	3	3	1	0.4	1.1	0.7	1.2	1.1	6	21	16	19	21
gi|238705321	BamD	27432	3	2	3	3	1	0.7	0.4	0.8	0.4	0.3	8	10	9	10	13
gi|238707927	BamE	9943	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708048	VacJ	28773	3	3	3	-	1	0.7	0.3	0.8	-	0.4	11	7	9	-	6
gi|238706986	MltA	42895	1	-	-	1	-	0.1	-	0.1	0.2	-	2	-	3	4	-
gi|238707253	MltC	39935	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706640	MltD	48282	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707358	OmpF	40024	2	3	3	3	1	0.2	0.5	0.4	0.8	0.4	3	13	6	10	8
gi|238706321	TolC	53048	2	3	3	3	2	0.1	0.6	0.2	0.3	0.4	2	15	6	7	11
gi|238706508	OmpC.1	39947	-	1	-	2	-	-	0.1	-	0.1	-	-	8	-	2	-
gi|238707949	OmpC.2	40459	-	2	1	3	3	-	0.5	0.4	0.4	0.4	-	9	6	7	13
gi|238706822	OsmY	11124	3	3	2	2	1	1.3	3.5	1.2	2.0	1.3	6	15	5	8	9
gi|238706938	MetQ	29356	3	3	3	3	3	0.2	0.5	0.4	0.6	0.4	4	11	3	10	10
gi|238705360	TamA/YtfN	63394	1	3	3	2	1	0.1	0.1	0.2	0.2	0.1	8	7	7	7	6
gi|238706342	OM Siderophore receptor	73465	3	-	3	3	3	0.0	-	1.0	0.1	0.1	4	-	31	3	17
gi|238708209	HemR	72964	-	-	3	-	2	-	-	0.4	-	0.1	-	-	17	-	6
gi|238705851	FhuA	81724	-	-	3	-	2	-	-	1.4	-	0.3	-	-	57	-	20
gi|238706670	FiuA	77361	-	-	3	-	1	-	-	0.1	-	0.1	-	-	7	-	9
gi|238708284	Ton-B dependent receptor	76573	-	-	2	-	-	-	-	0.1	-	-	-	-	4	-	-
gi|238707022	HasR	89969	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708022	OmpE	39925	-	1	1	1	-	-	0.2	0.1	0.1	-	-	6	3	4	-
gi|238707630	OprC	75132	1	-	2	-	1	0.0	-	0.2	-	0.0	4	-	6	-	4
gi|238706821	OmpW	23865	-	-	-	3	-	-	-	-	0.3	-	-	-	-	3	-
gi|238707009	BtuB	68957	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708047	FadL	46035	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705242	BcsC	126527	3	3	3	3	-	0.4	0.1	0.3	0.4	-	32	12	19	21	-
gi|238705556	Haemolysin activator	62594	-	-	-	1	-	-	-	-	0.1	-	-	-	-	5	-
gi|238705555	Hemolysin	48963	2	-	1	2	-	0.1	-	0.1	0.1	-	3	-	5	3	-
gi|238706248	Hemolysin activator	61801	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707066	Pep M37	50544	1	-	-	-	1	0.1	-	-	-	0.1	5	-	-	-	12
gi|238705486	Phos. A1	32689	-	2	-	1	-	-	0.1	-	0.1	-	-	4	-	4	-
gi|238705976	Protease II	77905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|110007337	YhlB	62424	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705110	Possible Adhesin	155469	1	-	-	-	1	0.0	-	-	-	0.0	7	-	-	-	32
gi|238707370	PqiB/Mam7	59633	1	1	1	-	-	0.1	0.1	0.1	-	-	5	5	3	-	-
gi|238705561	AidA	44266	-	1	-	-	-	-	0.1	-	-	-	-	3	-	-	-
gi|238705628	Fim. Usher	87773	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705013	Flagellin, partial	10509	3	3	3	1	-	3.0	2.4	0.7	0.8	-	10	8	2	3	-
gi|498532474	Flagellin	43589	3	3	3	1	-	1.3	2.2	0.5	0.8	-	25	37	8	13	-
gi|238706460	FlgE	44254	1	3	2	-	-	0.1	0.3	0.2	-	-	3	6	3	-	-
gi|238706435	FliD	49054	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707768	YiaD	22348	3	3	2	3	1	0.3	0.2	0.3	0.6	0.2	4	1	2	9	7
gi|238707898	YfhG	35016	3	3	2	3	1	0.1	0.2	0.1	0.3	0.2	4	6	2	5	3
gi|238705190	YraP	20221	1	2	2	1	1	0.2	0.2	0.2	0.4	0.6	1	1	1	2	6
gi|238705961	CsgG	24712	1	-	1	-	-	0.1	-	-	-	-	8	-	-	-	-
gi|238707176	SecD	71766	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705788	TTSS protein	61602	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705783	TTSS protein	45918	-	-	-	-	1	-	-	-	-	0.1	-	-	-	-	22
gi|238705857	AmiB	62746	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705632	Blc	20801	1	-	1	-	-	0.2	-	0.4	-	-	1	-	4	-	-
gi|238705953	Mce-Like	91347	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706915	MipA	28023	-	1	1	-	1	-	0.1	0.1	-	0.3	-	4	3	-	4
gi|238707429	NagZ	99335	-	-	1	-	-	-	-	0.0	-	-	-	-	9	-	-
gi|238707352	NlpD	34028	-	2	-	2	1	-	0.1	-	0.2	0.2	-	2	-	1	2
gi|238706941	NlpE/CutF	25245	-	2	2	1	1	-	0.1	0.4	0.5	0.1	-	2	3	7	2
gi|238708334	OmpX	17086	-	1	-	-	-	-	0.2	-	-	-	-	3	-	-	-
gi|163883857	PtrA	107905	-	-	-	-	2	-	-	-	-	0.0	-	-	-	-	16
gi|238706761	RlpA	37591	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|46369605	RupA	17881	-	-	-	-	-	-	-	0.8	-	-	-	-	9	-	-
gi|238706180	Unknown	17544	1	-	-	1	-	0.2	-	-	0.2	-	2	-	-	3	-
gi|133904995	Unknown	200596	-	-	1	-	-	-	-	-	-	0.0	-	-	-	-	0
gi|238706101	Unknown	106480	-	-	1	-	-	-	-	0.0	-	-	-	-	2	-	-
gi|238705250	Unknown	40804	-	1	-	-	-	-	0.1	0.1	-	-	-	2	2	-	-
gi|238706025	Hypothetical	17964	3	1	-	2	1	0.2	0.2	-	0.2	0.2	5	8	-	7	8
gi|238705088	YeeJ	54634	-	1	-	-	-	-	0.1	-	-	-	-	2	-	-	-
gi|238706732	Hypothetical	31452	3	-	1	-	-	0.1	-	0.1	-	-	2	-	3	-	-
gi|238707408	Hypothetical	56778	1	-	-	-	-	0.1	-	-	-	-	2	-	-	-	-
gi|238707442	Hypothetical	22794	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707291	Hypothetical	54666	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

(C)

Accession	Protein	MW (Da)	RD64														
			Replicates					emPAI					Peptides				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
gi|238708193	Brauns (Lpp)	6740	3	3	3	3	3	36.2	157.6	26.0	8.7	7.8	41	72	67	35	56
gi|238707374	OmpA	37846	3	3	3	3	3	2.2	6.5	2.5	4.2	3.4	96	238	69	129	145
gi|238705740	Pal	18149	3	3	3	3	3	2.3	3.5	3.1	3.2	2.3	8	27	21	24	46
gi|238708170	SlyB	15382	3	3	3	3	3	1.2	2.5	2.5	3.2	1.1	5	19	9	20	6
gi|238706031	LptD	89769	3	3	3	3	1	0.1	0.1	0.1	0.3	0.0	7	10	4	15	20
gi|238706768	LptE	21133	2	3	3	3	-	0.2	0.3	0.2	0.4	-	3	5	2	5	-
gi|238706888	Slp Family	20689	3	3	3	3	-	0.5	0.6	0.7	0.4	-	3	6	7	4	-
gi|238706959	BamA/YaeT	88022	3	3	3	3	1	1.1	1.3	1.1	1.4	0.1	34	64	44	58	19
gi|238707860	BamB	42358	3	3	3	3	-	0.2	1.0	0.8	1.3	-	6	20	15	20	-
gi|238705699	BamC	38345	3	3	3	3	1	0.5	1.7	0.7	0.9	0.1	8	34	13	19	4
gi|238705321	BamD	27432	3	3	3	3	1	0.8	0.6	0.5	0.8	0.1	8	17	9	11	2
gi|238707927	BamE	9943	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708048	VacJ	28773	3	3	3	-	1	0.7	0.8	0.7	-	0.1	10	16	9	-	1
gi|238706986	MltA	42895	3	-	2	1	-	0.2	-	0.1	0.3	-	4	-	3	4	-
gi|238707253	MltC	39935	3	-	-	1	-	0.1	-	-	0.2	-	1	-	-	3	-
gi|238706640	MltD	48282	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707358	OmpF	40024	1	3	3	3	3	0.2	0.4	0.3	0.7	0.2	5	8	6	15	9
gi|238706321	TolC	53048	3	3	3	3	3	0.1	0.8	0.1	0.4	0.2	2	22	5	12	13
gi|238706508	OmpC.1	39947	2	2	-	3	-	0.1	0.1	-	0.1	-	2	3	-	3	-
gi|238707949	OmpC.2	40459	-	3	2	3	3	-	0.8	0.3	0.7	0.7	-	23	5	17	26
gi|238706822	OsmY	11124	3	3	3	3	-	0.9	4.5	0.8	1.0	-	4	27	4	5	-
gi|238706938	MetQ	29356	1	2	3	2	2	0.4	0.1	0.2	0.4	0.3	5	6	3	4	6
gi|238705360	TamA/YtfN	63394	2	3	3	3	-	0.1	0.2	0.2	0.1	-	4	8	8	7	-
gi|238706342	OM Siderophore receptor	73465	3	-	3	3	2	0.0	-	0.9	0.1	0.0	4	-	31	2	8
gi|238708209	HemR	72964	-	-	3	-	3	-	-	0.5	-	0.2	-	-	25	-	15
gi|238705851	FhuA	81724	-	-	3	-	3	-	-	1.1	-	0.1	-	-	54	-	15
gi|238706670	FiuA	77361	-	-	1	-	1	-	-	0.1	-	0.1	-	-	11	-	11
gi|238708284	Ton-B dependent receptor	76573	-	-	2	-	1	-	-	0.1	-	0.0	-	-	4	-	4
gi|238707022	HasR	89969	-	-	-	-	-	-	-	0.0	-	-	-	-	5	-	-
gi|238708022	OmpE	39925	-	-	1	2	-	-	-	0.1	0.1	-	-	-	4	5	-
gi|238707630	OprC	75132	1	-	2	2	2	0.1	-	0.1	0.0	0.1	4	-	4	2	10
gi|238706821	OmpW	23865	1	1	-	2	-	0.1	0.1		0.3	-	2	4		4	-
gi|238707009	BtuB	68957	-	-	-	2	-	-	-	-	0.1	-	-	-	-	3	-
gi|238708047	FadL	46035	-	-	1	3	-	-	-	0.1	0.1	-	-	-	2	1	-
gi|238705242	BcsC	126527	3	3	3	3	-	0.3	0.2	0.3	0.2	-	24	25	19	19	-
gi|238705556	Haemolysin activator	62594	-	-	1	2	-	-	-	0.1	0.1	-	-	-	1	1	-
gi|238705555	Hemolysin	48963	2	-	-	2	-	0.1	-	-	0.1	-	4	-	-	3	-
gi|238706248	Hemolysin activator	61801	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707066	Pep M37	50544	3	1	-	-	-	0.1	0.1	-	-	-	7	7	-	-	-
gi|238705486	Phos. A1	32689	-	1	-	3	-	-	0.1	-	0.1	-	-	3	-	1	-
gi|238705976	Protease II	77905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|110007337	YhlB	62424	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705110	Possible Adhesin	155469	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707370	PqiB/Mam7	59633	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705561	AidA	44266	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705628	Fim. Usher	87773	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705013	Flagellin, partial	10509	3	3	1	1	-	16.9	5.0	-	1.3	-	32	13	-	7	-
gi|498532474	Flagellin	43589	3	3	1	1	-	7.5	5.6	0.1	1.2	-	77	65	2	21	-
gi|238706460	FlgE	44254	2	3	-	-	-	0.1	0.4	-	-	-	2	7	-	2	-
gi|238706435	FliD	49054	1	-	-	-	-	0.1	-	-	-	-	3	-	-	-	-
gi|238707768	YiaD	22348	3	3	3	3	-	0.2	0.3	0.2	0.5	-	3	4	2	12	-
gi|238707898	YfhG	35016	3	3	2	3	-	0.3	0.5	0.1	0.4	-	6	9	5	7	-
gi|238705190	YraP	20221	2	3	2	1	2	0.2	0.2	0.2	0.2	0.2	3	1	1	1	3
gi|238705961	CsgG	24712	2	-	1	-	1	0.1	-	0.1	-	0.1	9	-	4	-	19
gi|238707176	SecD	71766	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705788	TTSS protein	61602	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705783	TTSS protein	45918	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705857	AmiB	62746	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705632	Blc	20801	1	-	2	-	-	0.2	-	0.2	-	-	2	-	1	-	-
gi|238705953	Mce-Like	91347	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706915	MipA	28023	-	-	1	1	3	-	-	0.4	0.1	0.2	-	-	7	3	6
gi|238707429	NagZ	99335	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707352	NlpD	34028	-	2	1	3	1	-	0.1	0.3	0.3	0.1	-	3	4	4	3
gi|238706941	NlpE/CutF	25245	-	-	-	-	-	-	-	0.1	-	-	-	-	6	-	-
gi|238708334	OmpX	17086	-	1	-	1	2	-	0.2	-	0.4	0.3	-	1	-	2	3
gi|163883857	PtrA	107905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706761	RlpA	37591	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|46369605	RupA	17881	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706180	Unknown	17544	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|133904995	Unknown	200596	-	-	-	-	1	-	-	-	-	0.0	-	-	-	-	5
gi|238706101	Unknown	106480	1	-	-	-	-	0.0	-	-	-	-	3	-	-	-	-
gi|238705250	Unknown	40804	-	-	1	1	-	-	-	-	0.1	-	-	-	-	3	-
gi|238706025	Hypothetical	17964	2	1	1	3	2	0.2	0.2	0.2	0.2	0.2	7	8	5	9	11
gi|238705088	YeeJ	54634	-	-	-	-	-	-	-	0.1	-	-	-	-	4	-	-
gi|238706732	Hypothetical	31452	1	-	-	1	-	0.1	-	-	0.1	-	2	-	-	4	-
gi|238707408	Hypothetical	56778	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707442	Hypothetical	22794	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707291	Hypothetical	54666	1	-	-	-	-	0.1	-	-	-	-	5	-	-	-	-

(D)

Accession	Protein	MW (Da)	RD124														
			Replicates					emPAI					Peptides				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
gi|238708193	Brauns (Lpp)	6740	3	3	3	3	3	81.7	102.4	55.1	20.3	34.8	45	64	60	40	42
gi|238707374	OmpA	37846	3	3	3	3	3	1.8	4.3	2.7	3.6	2.3	97	207	64	103	156
gi|238705740	Pal	18149	3	3	3	3	3	1.5	3.1	2.6	3.5	2.8	10	30	16	23	23
gi|238708170	SlyB	15382	3	3	3	3	3	1.1	2.5	2.4	4.3	1.2	6	20	9	20	8
gi|238706031	LptD	89769	1	3	2	3	2	0.1	0.2	0.1	0.3	0.2	6	12	4	17	7
gi|238706768	LptE	21133	3	3	3	3	2	0.2	0.7	0.3	0.6	0.2	3	7	3	5	2
gi|238706888	Slp Family	20689	3	2	3	3	2	0.4	0.6	0.8	0.5	0.4	3	5	7	5	4
gi|238706959	BamA/YaeT	88022	3	3	3	3	3	0.9	1.3	0.9	2.0	0.7	31	64	38	64	43
gi|238707860	BamB	42358	3	3	3	3	3	0.3	0.6	0.4	1.2	0.6	7	15	10	21	17
gi|238705699	BamC	38345	3	3	3	3	3	0.7	1.1	0.7	1.1	0.7	10	26	15	20	19
gi|238705321	BamD	27432	3	3	3	3	3	0.8	0.4	0.8	0.4	0.2	8	10	8	7	9
gi|238707927	BamE	9943	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708048	VacJ	28773	3	3	3	-	3	0.5	0.3	0.4	-	0.3	8	8	6	-	5
gi|238706986	MltA	42895	2	-	-	1	2	0.1	-	-	0.3	0.1	3	-	-	7	1
gi|238707253	MltC	39935	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706640	MltD	48282	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707358	OmpF	40024	2	3	3	3	2	0.3	0.7	0.3	1.6	0.3	6	13	6	24	10
gi|238706321	TolC	53048	3	3	2	3	3	0.1	1.1	0.1	0.4	0.5	5	25	5	9	13
gi|238706508	OmpC.1	39947	3	2	3	3	-	0.2	0.3	0.1	0.6	-	5	8	3	14	-
gi|238707949	OmpC.2	40459	-	3	3	3	3	-	0.6	0.3	1.3	0.7	-	18	5	21	21
gi|238706822	OsmY	11124	1	3	2	2	3	0.7	2.6	1.1	0.8	1.7	6	11	5	3	9
gi|238706938	MetQ	29356	3	2	3	3	3	0.1	0.3	0.3	0.3	0.7	3	10	4	5	12
gi|238705360	TamA/YtfN	63394	2	3	3	3	3	0.1	0.2	0.2	0.2	0.2	7	10	10	7	8
gi|238706342	OM Siderophore receptor	73465	2	-	3	3	3	0.1	-	1.2	0.3	0.3	5	-	33	8	12
gi|238708209	HemR	72964	-	-	3	-	3	-	-	0.5	-	0.1	-	-	21	-	7
gi|238705851	FhuA	81724	1	-	3	1	3	0.0	-	1.1	0.2	0.3	5	-	55	16	12
gi|238706670	FiuA	77361	-	-	3	2	2	-	-	0.1	0.0	0.1	-	-	6	2	5
gi|238708284	Ton-B dependent receptor	76573	-	-	3	-	-	-	-	0.1	-	-	-	-	7	-	-
gi|238707022	HasR	89969	-	-	1	-	-	-	-	0.0	-	-	-	-	0	-	-
gi|238708022	OmpE	39925	-	2	2	2	1	-	0.1	0.1	0.2	0.1	-	2	3	5	3
gi|238707630	OprC	75132	-	-	2	-	3	-	-	0.0	-	0.1	-	-	4	-	7
gi|238706821	OmpW	23865	-	-	-	2	1	-	-	-	0.3	0.1	-	-	-	6	2
gi|238707009	BtuB	68957	1	1	-	2	-	0.1	0.1	-	0.1	-	1	1	-	4	-
gi|238708047	FadL	46035	-	-	-	2	-	-	-	-	0.3	-	-	-	-	7	-
gi|238705242	BcsC	126527	3	2	2	2	-	0.0	0.1	0.1	0.1	-	13	13	12	11	-
gi|238705556	Haemolysin activator	62594	1	-	1	1	-	0.1	-	0.1	0.1	-	4	-	2	4	-
gi|238705555	Hemolysin	48963	2	1	-	3	-	0.1	0.1	-	0.1	-	3	3	-	2	-
gi|238706248	Hemolysin activator	61801	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707066	Pep M37	50544	2	-	-	-	-	0.1	-	-	-	-	8	-	-	-	-
gi|238705486	Phos. A1	32689	-	2	-	2	-	-	0.1	-	0.1	-	-	4	-	4	-
gi|238705976	Protease II	77905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|110007337	YhlB	62424	-	-	2	-	-	-	-	0.1	-	-	-	-	3	-	-
gi|238705110	Possible Adhesin	155469	-	-	1	-	-	-	-	0.0	-	-	-	-	4	-	-
gi|238707370	PqiB/Mam7	59633	-	-	1	-	-	-	-	0.1	-	-	-	-	3	-	-
gi|238705561	AidA	44266	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705628	Fim. Usher	87773	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705013	Flagellin, partial	10509	3	2	-	-	-	25.6	0.5	-	-	-	27	2	-	-	-
gi|498532474	Flagellin	43589	3	1	-	-	-	4.5	0.2	-	-	-	56	3	-	-	-
gi|238706460	FlgE	44254	3	-	-	-	-	0.2	-	-	-	-	4	-	-	-	-
gi|238706435	FliD	49054	-	-	1	-	-	-	-	0.1	-	-	-	-	3	-	-
gi|238707768	YiaD	22348	3	3	3	3	3	0.3	0.8	0.3	0.7	0.3	5	8	3	14	6
gi|238707898	YfhG	35016	3	3	2	3	2	0.2	0.3	0.1	0.3	0.3	5	6	3	5	5
gi|238705190	YraP	20221	3	-	2	2	2	0.2	-	0.9	0.2	0.4	3	-	5	3	4
gi|238705961	CsgG	24712	1	-	-	-	-	0.1	-	-	-	-	11	-	-	-	-
gi|238707176	SecD	71766	1	-	-	-	-	0.1	-	-	-	-	8	-	-	-	-
gi|238705788	TTSS protein	61602	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705783	TTSS protein	45918	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705857	AmiB	62746	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705632	Blc	20801	-	-	-	-	3	-	-	-	-	0.2	-	-	-	-	1
gi|238705953	Mce-Like	91347	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706915	MipA	28023	-	-	1	2	2	-	-	0.3	0.3	0.1	-	-	3	5	2
gi|238707429	NagZ	99335	-	-	1	-	-	-	-	0.0	-	-	-	-	9	-	-
gi|238707352	NlpD	34028	1	1	2	3	3	0.1	0.1	0.1	0.2	0.3	6	2	4	2	4
gi|238706941	NlpE/CutF	25245	-	3	3	1	3	-	0.2	0.1	0.2	0.2	-	3	3	4	3
gi|238708334	OmpX	17086	-	1	-	-	-	-	0.2	-	-	-	-	1	-	-	-
gi|163883857	PtrA	107905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706761	RlpA	37591	-	-	1	-	-	-	-	0.1	-	-	-	-	2	-	-
gi|46369605	RupA	17881	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706180	Unknown	17544	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|133904995	Unknown	200596	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706101	Unknown	106480	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705250	Unknown	40804	-	-	-	2	-	-	-	-	0.1	-	-	-	-	2	-
gi|238706025	Hypothetical	17964	3	1	1	3	2	0.2	0.2	0.2	0.2	0.2	4	6	6	9	6
gi|238705088	YeeJ	54634	3	-	2	-	-	0.1	-	0.1	-	-	5	-	3	-	-
gi|238706732	Hypothetical	31452	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707408	Hypothetical	56778	1	-	-	-	-	0.1	-	-	-	-	2	-	-	-	-
gi|238707442	Hypothetical	22794	-	-	1	-	-	-	-	0.2	-	-	-	-	1	-	-
gi|238707291	Hypothetical	54666	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

(E)

Accession	Protein	MW (Da)	RD354														
			Replicates					emPAI					Peptides				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
gi|238708193	Brauns (Lpp)	6740	3	3	3	3	3	26.9	157.6	71.1	10.2	11.5	33	72	70	45	40
gi|238707374	OmpA	37846	3	3	3	3	3	2.5	6.0	3.2	7.2	2.3	135	232	71	141	164
gi|238705740	Pal	18149	3	3	3	3	3	2.5	5.5	2.6	4.2	2.3	11	28	20	26	24
gi|238708170	SlyB	15382	3	3	3	3	3	0.6	2.6	1.6	3.6	0.5	4	12	7	22	5
gi|238706031	LptD	89769	3	3	3	3	1	0.1	0.1	0.1	0.4	0.0	5	11	6	17	5
gi|238706768	LptE	21133	3	3	3	3	1	0.4	0.2	0.3	0.6	0.2	6	4	4	5	2
gi|238706888	Slp Family	20689	3	3	3	3	-	0.7	0.4	0.7	0.4	-	5	3	5	3	-
gi|238706959	BamA/YaeT	88022	3	3	3	3	3	1.0	0.8	1.2	1.4	0.4	41	47	45	54	31
gi|238707860	BamB	42358	3	3	3	3	2	0.4	0.5	0.7	1.0	0.5	9	14	15	21	14
gi|238705699	BamC	38345	3	3	3	3	2	0.8	1.3	0.7	1.4	0.5	11	26	15	24	18
gi|238705321	BamD	27432	3	3	3	3	2	1.3	0.5	0.9	1.0	0.1	12	14	9	12	9
gi|238707927	BamE	9943	1	-	-	-	-	0.4	-	-	-	-	1	-	-	-	-
gi|238708048	VacJ	28773	3	3	3	-	2	0.8	0.5	0.8	-	0.2	1-	12	9	-	4
gi|238706986	MltA	42895	3	1	2	1	-	0.1	0.1	0.1	0.4	-	6	2	4	5	-
gi|238707253	MltC	39935	2	-	-	-	-	0.1	-	-	-	-	2	-	-	-	-
gi|238706640	MltD	48282	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707358	OmpF	40024	2	3	3	3	1	0.2	0.5	0.4	1.2	0.1	5	14	8	24	4
gi|238706321	TolC	53048	2	3	3	3	3	0.1	1.1	0.1	0.4	0.2	4	28	4	10	8
gi|238706508	OmpC.1	39947	3	1	-	2	-	0.1	0.1	0.1	0.1	-	2	2	3	5	-
gi|238707949	OmpC.2	40459	-	3	2	3	1	-	0.3	0.3	0.9	0.4	-	9	3	19	20
gi|238706822	OsmY	11124	3	3	3	3	2	1.5	5.7	0.6	0.9	1.0	8	21	5	4	4
gi|238706938	MetQ	29356	1	1	2	1	2	0.1	0.1	0.2	0.1	0.7	4	4	4	3	18
gi|238705360	TamA/YtfN	63394	2	3	3	3	2	0.1	0.1	0.3	0.2	0.1	7	8	12	5	6
gi|238706342	OM Siderophore receptor	73465	3	-	3	1	3	0.1	-	1.0	0.0	0.1	6	-	36	1	11
gi|238708209	HemR	72964	-	-	3	-	3	-	-	0.4	-	0.1	-	-	21	-	6
gi|238705851	FhuA	81724	1	-	3	1	3	0.0	-	1.5	0.0	0.1	4	-	76	2	16
gi|238706670	FiuA	77361	-	-	1	-	2	-	-	0.2	-	0.0	-	-	8	-	7
gi|238708284	Ton-B dependent receptor	76573	-	-	3	-	-	-	-	0.0	-	-	-	-	2	-	-
gi|238707022	HasR	89969	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708022	OmpE	39925	-	1	2	1	-	-	0.1	0.1	0.1	0.0	-	6	3	4	0
gi|238707630	OprC	75132	3	-	2	-	2	0.0	-	0.1	-	0.0	4	-	7	-	5
gi|238706821	OmpW	23865	2	1	-	3	-	0.1	0.1	-	0.4	-	2	1	-	7	-
gi|238707009	BtuB	68957	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708047	FadL	46035	1	3	1	3	-	0.1	0.1	0.1	0.2	-	2	1	2	4	-
gi|238705242	BcsC	126527	3	3	3	3	-	0.2	0.1	0.2	0.2	-	27	16	18	15	-
gi|238705556	Haemolysin activator	62594	3	-	2	1	-	0.1	-	0.1	0.1	-	3	-	2	1	-
gi|238705555	Hemolysin	48963	2	3	-	3	-	0.1	0.1	-	0.1	-	4	3	-	2	-
gi|238706248	Hemolysin activator	61801	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707066	Pep M37	50544	2	-	-	-	-	0.1	-	-	-	-	6	-	-	-	-
gi|238705486	Phos. A1	32689	-	1	-	1	-	-	0.1	-	0.1	-	-	4	-	3	-
gi|238705976	Protease II	77905	1	-	1	-	-	0.0	-	0.0	-	-	4	-	2	-	-
gi|110007337	YhlB	62424	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705110	Possible Adhesin	155469	3	-	3	-	-	0.0	-	0.0	-	-	9	-	3	-	-
gi|238707370	PqiB/Mam7	59633	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705561	AidA	44266	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705628	Fim. Usher	87773	-	-	-	-	1	-	-	-	-	0.0	-	-	-	-	16
gi|238705013	Flagellin, partial	10509	3	3	-	-	-	10.8	0.8	-	-	-	18	3	-	-	-
gi|498532474	Flagellin	43589	3	3	-	-	-	3.0	1.1	-	-	-	44	21	-	-	-
gi|238706460	FlgE	44254	1	2	-	-	-	0.1	0.1	-	-	-	3	2	-	-	-
gi|238706435	FliD	49054	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707768	YiaD	22348	3	1	3	3	1	0.3	0.2	0.2	0.3	0.3	6	4	3	10	6
gi|238707898	YfhG	35016	3	3	2	3	-	0.3	0.4	0.1	0.4	0.0	6	6	2	7	0
gi|238705190	YraP	20221	1	2	2	2	2	0.2	0.2	0.2	0.2	0.2	3	2	1	2	2
gi|238705961	CsgG	24712	3	-	1	1	-	0.3	-	0.1	0.1	-	13	-	2	6	-
gi|238707176	SecD	71766	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705788	TTSS protein	61602	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705783	TTSS protein	45918	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705857	AmiB	62746	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705632	Blc	20801	1	-	2	-	1	0.2	-	0.4	-	0.4	1	-	3	-	2
gi|238705953	Mce-Like	91347	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706915	MipA	28023	-	1	1	2	2	-	0.1	0.3	0.1	0.1	-	1	3	2	1
gi|238707429	NagZ	99335	-	-	-	-	-	-	-	-	-	-	-	-	-	7	-
gi|238707352	NlpD	34028	1	3	-	3	1	0.1	0.1	-	0.2	0.1	4	2	-	4	4
gi|238706941	NlpE/CutF	25245	-	-	2	-	1	-	-	0.1	-	0.1	-	-	1	-	1
gi|238708334	OmpX	17086	-	1	-	3	-	-	0.2	-	0.3	-	-	1	-	3	-
gi|163883857	PtrA	107905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706761	RlpA	37591	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|46369605	RupA	17881	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706180	Unknown	17544	-	-	1	-	-	-	-	0.2	-	-	-	-	2	-	-
gi|133904995	Unknown	200596	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706101	Unknown	106480	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705250	Unknown	40804	-	-	-	1	-	-	-	-	0.1	-	-	-	-	4	-
gi|238706025	Hypothetical	17964	2	1	1	3	1	0.2	0.2	0.2	0.2	0.2	5	4	6	9	10
gi|238705088	YeeJ	54634	2	-	-	1	-	0.1	-	-	0.1	-	2	-	-	1	-
gi|238706732	Hypothetical	31452	3	1	1	-	-	0.1	0.2	0.1	-	-	2	6	3	-	-
gi|238707408	Hypothetical	56778	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707442	Hypothetical	22794	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707291	Hypothetical	54666	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

(F)

Accession	Protein	MW (Da)	RD366													
			Replicates					emPAI					Peptides				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
gi|238708193	Brauns (Lpp)	6740	3	3	3	3	3	53.2	135.1	36.3	23.8	23.6	37	64	58	41	30
gi|238707374	OmpA	37846	3	3	3	3	3	2.3	58.4	3.0	43.8	26.4	111	183	70	121	148
gi|238705740	Pal	18149	3	3	3	3	3	1.5	9.6	3.2	8.7	4.0	11	28	18	22	14
gi|238708170	SlyB	15382	3	3	3	3	3	0.6	2.6	2.5	8.0	1.6	5	8	9	19	6
gi|238706031	LptD	89769	1	1	2	3	1	0.0	0.0	0.1	5.8	0.1	6	9	3	14	6
gi|238706768	LptE	21133	1	2	2	3	1	0.2	0.2	0.3	2.2	0.2	4	2	2	4	2
gi|238706888	Slp Family	20689	2	2	3	3	1	0.3	1.1	0.7	1.5	0.6	4	1	7	3	3
gi|238706959	BamA/YaeT	88022	3	3	3	3	3	0.8	9.1	1.1	24.2	1.2	32	36	40	67	30
gi|238707860	BamB	42358	3	3	3	3	2	0.2	2.5	0.9	8.7	0.7	6	6	12	20	19
gi|238705699	BamC	38345	3	3	3	3	2	0.7	5.6	0.8	10.4	0.8	10	19	16	22	22
gi|238705321	BamD	27432	3	2	3	3	2	0.7	0.3	0.6	2.9	0.3	8	6	8	9	12
gi|238707927	BamE	9943	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-
gi|238708048	VacJ	28773	3	3	3	-	2	0.6	2.0	0.6	-	0.3	9	6	9	-	7
gi|238706986	MltA	42895	1	-	-	1	1	0.2	-	-	4.0	0.3	2	-	-	4	3
gi|238707253	MltC	39935	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706640	MltD	48282	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707358	OmpF	40024	2	2	2	3	1	0.3	0.4	0.2	6.8	4.0	3	7	7	18	4
gi|238706321	TolC	53048	2	3	3	3	3	0.1	3.2	0.1	3.5	1.1	3	18	3	11	5
gi|238706508	OmpC.1	39947	2	-	-	2	-	0.1	-	-	1.5	-	2	-	-	4	-
gi|238707949	OmpC.2	40459	-	3	3	3	3	-	2.2	0.2	4.2	3.9	-	6	3	21	9
gi|238706822	OsmY	11124	3	3	3	3	2	0.6	8.4	1.0	3.7	2.0	4	19	3	7	8
gi|238706938	MetQ	29356	2	2	3	3	3	0.5	5.3	0.4	3.7	0.7	7	11	6	11	8
gi|238705360	TamA/YtfN	63394	1	3	3	3	2	0.2	2.1	0.1	3.4	0.2	9	5	8	7	9
gi|238706342	OM Siderophore receptor	73465	2	-	3	2	2	0.0	-	0.8	0.6	0.3	6	-	25	2	10
gi|238708209	HemR	72964	-	-	3	-	2	-	-	0.4	-	0.2	-	-	17	-	12
gi|238705851	FhuA	81724	2	-	3	-	3	0.0	-	1.2	-	2.1	4	-	51	-	11
gi|238706670	FiuA	77361	-	-	3	-	2	-	-	0.2	-	0.0	-	-	8	-	3
gi|238708284	Ton-B dependent receptor	76573	-	-	2	-	-	-	-	0.1	-	-	-	-	3	-	-
gi|238707022	HasR	89969	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708022	OmpE	39925	-	-	-	3	-	-	-	-	1.4	0.0	-	-	-	5	0
gi|238707630	OprC	75132	-	-	3	1	2	-	-	0.1	0.0	0.0	-	-	4	3	3
gi|238706821	OmpW	23865	-	-	-	2	-	-	-	-	0.2	-	-	-	-	4	-
gi|238707009	BtuB	68957	-	-	-	-	-	-	-		-	-	-	-		-	-
gi|238708047	FadL	46035	-	1	-	2	-	-	0.1	-	2.5	-	-	1	-	3	-
gi|238705242	BcsC	126527	2	2	3	3	-	0.2	0.1	0.1	7.3	-	23	14	9	22	-
gi|238705556	Haemolysin activator	62594	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705555	Hemolysin	48963	-	-	-	1	-	-	-	-	2.0	-	-	-	-	2	-
gi|238706248	Hemolysin activator	61801	-	-	-	-	-	-	-		-	-	-	-		-	-
gi|238707066	Pep M37	50544	1	-	-	-	-	0.1	-	-	-	-	7	-	-	-	-
gi|238705486	Phos. A1	32689	-	3	-	1	-	-	0.7	-	0.1	-	-	2	-	1	-
gi|238705976	Protease II	77905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|110007337	YhlB	62424	-	-	3	-	-	-	-	0.1	-	-	-	-	2	-	-
gi|238705110	Possible Adhesin	155469	2	-	-	-	-	0.0	-	-	-	-	12	-	-	-	-
gi|238707370	PqiB/Mam7	59633	-	-	2	-	1	-	-	0.1	-	0.1	-	-	4	-	4
gi|238705561	AidA	44266	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705628	Fim. Usher	87773	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705013	Flagellin, partial	10509	3	3	-	-	-	10.8	6.4	-	-	-	21	9	-	-	-
gi|498532474	Flagellin	43589	3	3	-	-	-	2.5	18.3	-	-	-	35	34	-	-	-
gi|238706460	FlgE	44254	1	2	-	-	-	0.2	1.5	-	-	-	3	2	-	-	-
gi|238706435	FliD	49054	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707768	YiaD	22348	-	-	2	3	2	-	-	0.2	3.8	0.3	-	-	2	8	8
gi|238707898	YfhG	35016	2	-	1	3	1	0.1	-	0.1	2.1	0.1	4	-	1	5	2
gi|238705190	YraP	20221	2	1	1	2	2	0.2	2.0	0.2	1.1	0.8	1	2	1	2	6
gi|238705961	CsgG	24712	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707176	SecD	71766	1	-	-	-	-	0.1	-	-	-	-	7	-	-	-	-
gi|238705788	TTSS protein	61602	1	-	-	-	-	0.1	-	-	-	-	5	-	-	-	-
gi|238705783	TTSS protein	45918	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705857	AmiB	62746	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705632	Blc	20801	-	-	2	-	1	-	-	0.3	-	0.6	-	-	2	-	4
gi|238705953	Mce-Like	91347	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706915	MipA	28023	-	-	1	2	2	-	-	0.3	0.2	2.1	-	-	3	4	2
gi|238707429	NagZ	99335	-	1	-	-	-	-	6.0	-	-	-	-	6	-	-	-
gi|238707352	NlpD	34028	2	2	-	2	2	0.1	0.6	-	0.6	0.3	2	3	-	1	3
gi|238706941	NlpE/CutF	25245	-	2	3	3	2	-	0.6	0.4	1.1	0.2	-	2	4	2	4
gi|238708334	OmpX	17086	-	-	-	-	1	-	-	-	-	2.0	-	-	-	-	2
gi|163883857	PtrA	107905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706761	RlpA	37591	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|46369605	RupA	17881	1	-	-	-	-	0.2	-	-	-	-	4	-	-	-	-
gi|238706180	Unknown	17544	-	-	-	-	1	-	-	-	-	0.2	-	-	-	-	3
gi|133904995	Unknown	200596	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706101	Unknown	106480	-	-	-	-	-	-	-	-	-	-	-	-	-	4	-
gi|238705250	Unknown	40804	-	-	-	1	-	-	-	-	0.1	-	-	-	-	5	-
gi|238706025	Hypothetical	17964	1	1	1	2	1	0.2	0.2	0.2	4.6	0.2	6	7	2	8	1
gi|238705088	YeeJ	54634	1	-	1	-	1	0.1	-	0.1	-	0.1	4	-	2	-	2
gi|238706732	Hypothetical	31452	1	-	1	1	-	0.1	-	0.1	0.1	-	3	-	5	3	-
gi|238707408	Hypothetical	56778	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707442	Hypothetical	22794	-	-	-	1	-	-	-	-	0.2	-	-	-	-	2	-
gi|238707291	Hypothetical	54666	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

(G)

Accession	Protein	MW (Da)	RD382														
			Replicates					emPAI					Peptides				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
gi|238708193	Brauns (Lpp)	6740	3	3	3	3	3	17.1	100.9	30.0	7.3	12.4	30	73	57	30	44
gi|238707374	OmpA	37846	3	3	3	3	3	1.7	4.2	1.9	3.9	3.8	111	201	74	123	147
gi|238705740	Pal	18149	3	3	3	3	3	1.6	4.5	1.6	3.6	1.1	1-	32	12	24	23
gi|238708170	SlyB	15382	3	3	3	3	2	0.6	1.1	2.5	2.7	0.4	3	7	10	16	2
gi|238706031	LptD	89769	2	1	3	3	-	0.0	0.1	0.1	0.1	-	6	8	3	10	-
gi|238706768	LptE	21133	2	2	1	2	-	0.2	0.3	0.2	0.2	-	4	2	3	3	-
gi|238706888	Slp Family	20689	2	1	3	3	-	0.3	0.2	0.8	0.5	-	3	3	6	4	-
gi|238706959	BamA/YaeT	88022	3	3	3	3	-	0.7	0.8	1.0	1.6	-	36	49	46	64	-
gi|238707860	BamB	42358	3	3	3	3	-	0.1	0.6	0.7	1.6	-	6	11	16	24	-
gi|238705699	BamC	38345	3	3	3	3	-	0.5	0.8	0.7	1.4	-	13	22	13	20	-
gi|238705321	BamD	27432	2	2	3	3	-	0.6	0.4	0.6	0.4	-	9	8	8	8	-
gi|238707927	BamE	9943	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708048	VacJ	28773	3	3	3	-	-	0.5	0.5	0.8	-	-	1-	7	11	-	-
gi|238706986	MltA	42895	1	-	-	-	-	0.1	-	-	-	-	5	-	-	-	-
gi|238707253	MltC	39935	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706640	MltD	48282	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707358	OmpF	40024	2	3	2	3	3	0.1	0.3	0.1	0.7	0.6	2	6	4	12	21
gi|238706321	TolC	53048	1	3	3	3	3	0.1	0.8	0.1	0.6	0.1	2	21	3	11	6
gi|238706508	OmpC.1	39947	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707949	OmpC.2	40459	-	2	3	3	3	-	0.2	0.1	0.4	1.2	-	6	4	6	55
gi|238706822	OsmY	11124	3	3	3	3	-	0.4	3.0	0.5	0.8	-	2	15	2	3	-
gi|238706938	MetQ	29356	1	3	3	3	-	0.1	0.3	0.4	0.6	-	2	7	5	7	-
gi|238705360	TamA/YtfN	63394	-	2	3	3	-	-	0.1	0.2	0.2	-	-	6	7	6	-
gi|238706342	OM Siderophore receptor	73465	2	-	3	3	3	0.1	-	0.8	0.1	0.1	6	-	27	2	7
gi|238708209	HemR	72964	-	-	3	-	3	-	-	0.3	-	0.2	-	-	14	-	12
gi|238705851	FhuA	81724	2	-	3	1	3	0.0	-	1.0	0.1	0.2	4	-	46	6	16
gi|238706670	FiuA	77361	-	-	3	-	1	-	-	0.1	-	0.0	-	-	8	-	4
gi|238708284	Ton-B dependent receptor	76573	-	-	3	-	-	-	-	0.1	-	-	-	-	5	-	-
gi|238707022	HasR	89969	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708022	OmpE	39925	-	-	-	-	1	-	-	-	-	0.2	-	-	-	-	10
gi|238707630	OprC	75132	1	-	2	1	-	0.1	-	0.1	0.0	-	3	-	8	3	-
gi|238706821	OmpW	23865	-	-	-	3	-	-	-	-	0.2	-	-	-	-	3	-
gi|238707009	BtuB	68957	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708047	FadL	46035	-	-	-	1	-	-	-	-	0.1	-	-	-	-	1	-
gi|238705242	BcsC	126527	2	3	3	3	-	0.1	0.1	0.2	0.3	-	13	19	17	17	-
gi|238705556	Haemolysin activator	62594	1	-	1	-	-	0.1	-	0.1	-	-	3	-	3	-	-
gi|238705555	Hemolysin	48963	2	1	-	1	-	0.1	0.1	-	0.1	-	5	4	-	2	-
gi|238706248	Hemolysin activator	61801	1	-	-	-	-	0.1	-		-	-	3	-		-	-
gi|238707066	Pep M37	50544	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705486	Phos. A1	32689	-	1	-	-	-	-	0.1	-	-	-	-	1	-	-	-
gi|238705976	Protease II	77905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|110007337	YhlB	62424	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705110	Possible Adhesin	155469	2	-	1	-	-	0.0	-	0.0	-	-	13	-	3	-	-
gi|238707370	PqiB/Mam7	59633	-	-	1	1	-	-	-	0.2	0.1	-	-	-	6	1	-
gi|238705561	AidA	44266	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705628	Fim. Usher	87773	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705013	Flagellin, partial	10509	3	2	-	-	-	8.1	0.8	-	-	-	17	3	-	-	-
gi|498532474	Flagellin	43589	3	1	-	-	-	1.6	0.6	-	-	-	28	14	-	-	-
gi|238706460	FlgE	44254	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706435	FliD	49054	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707768	YiaD	22348	3	2	3	3	3	0.3	0.3	0.2	0.6	0.2	6	5	4	9	5
gi|238707898	YfhG	35016	2	1	2	3	-	0.1	0.2	0.1	0.3	-	4	5	3	4	-
gi|238705190	YraP	20221	2	-	2	-	-	0.2	-	0.3	-	-	1	-	2	-	-
gi|238705961	CsgG	24712	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707176	SecD	71766	1	-	-	1	-	0.1	-	-	0.1	-	3	-	-	2	-
gi|238705788	TTSS protein	61602	-	-	-	1	-	-	-	-	0.1	-	-	-	-	1	-
gi|238705783	TTSS protein	45918	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705857	AmiB	62746	-	-	1	-	-	-	-	0.1	-	-	-	-	4	-	-
gi|238705632	Blc	20801	-	-	3	-	-	-	-	0.2	-	-	-	-	2	-	-
gi|238705953	Mce-Like	91347	-	-	1	-	-	-	-	0.0	-	-	-	-	2	-	-
gi|238706915	MipA	28023	-	-	-	-	2	-	-	-	-	0.2	-	-	-	-	3
gi|238707429	NagZ	99335	-	-	1	-	-	-	-	0.0	-	-	-	-	8	-	-
gi|238707352	NlpD	34028	1	-	1	3	-	0.1	-	0.1	0.1	-	4	-	3	2	-
gi|238706941	NlpE/CutF	25245	-	1	3	2	-	-	0.2	0.4	0.2	-	-	2	3	3	-
gi|238708334	OmpX	17086	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|163883857	PtrA	107905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706761	RlpA	37591	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|46369605	RupA	17881	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706180	Unknown	17544	-	-	1	2	-	-	-	0.2	0.2	-	-	-	2	2	-
gi|133904995	Unknown	200596	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706101	Unknown	106480	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-
gi|238705250	Unknown	40804	-	2	-	1	-	-	0.1	-	0.1	-	-	3	-	5	-
gi|238706025	Hypothetical	17964	2	-	-	3	-	0.2	-	-	0.2	-	8	-	-	7	-
gi|238705088	YeeJ	54634	1	-	-	-	-	0.1	-	-	-	-	2	-	-	-	-
gi|238706732	Hypothetical	31452	-	-	1	-	-	-	-	0.1	-	-	-	-	5	-	-
gi|238707408	Hypothetical	56778	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707442	Hypothetical	22794	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707291	Hypothetical	54666	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

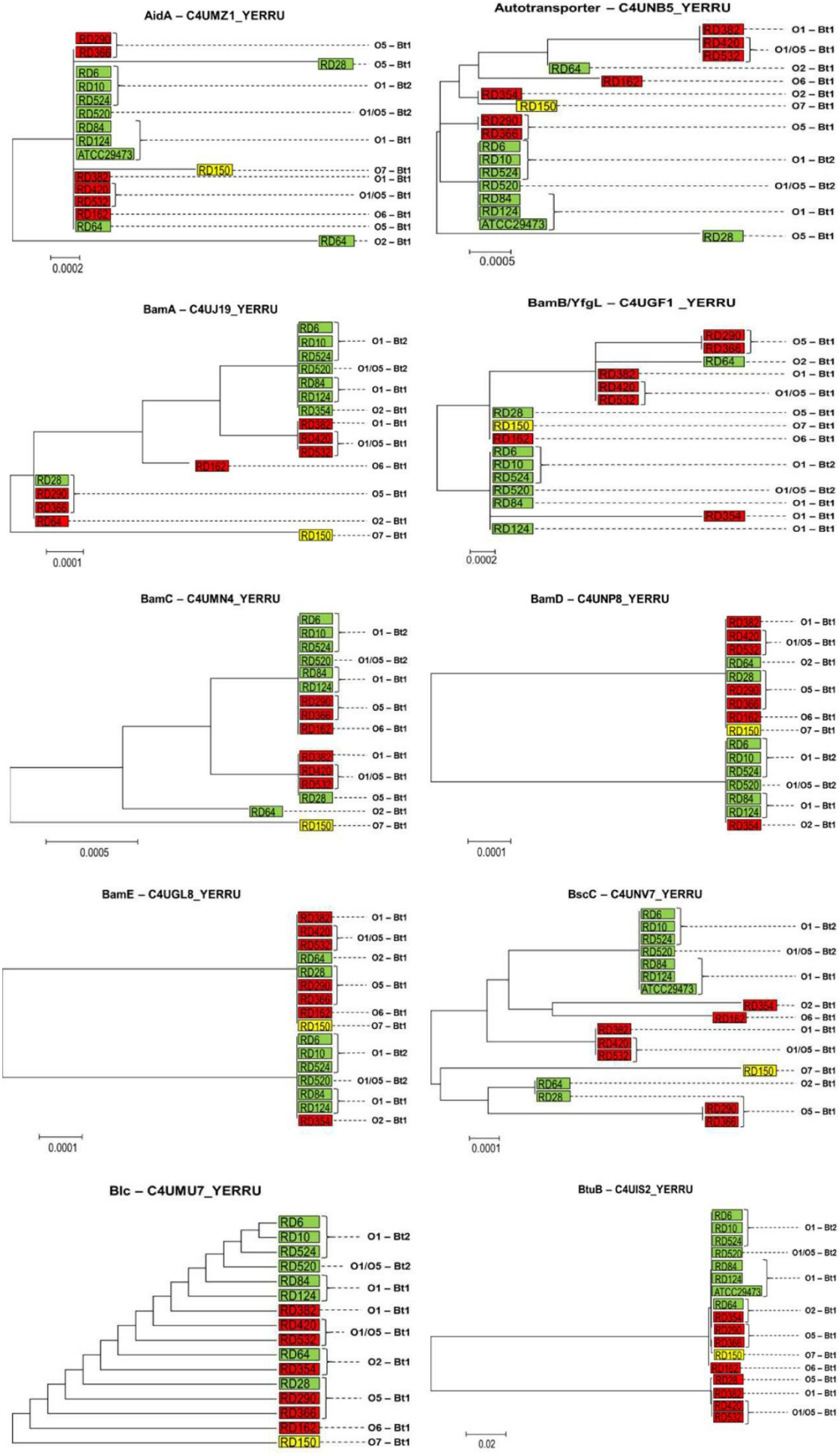
(H)

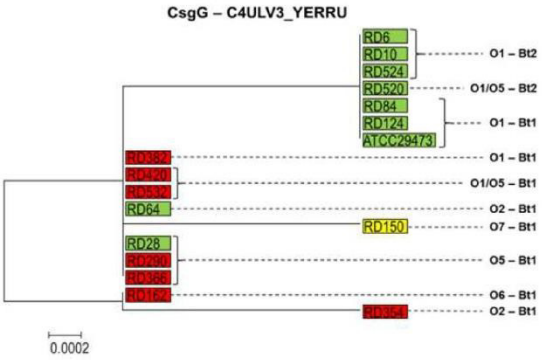
Accession	Protein	MW (Da)	RD420														
			Replicates					emPAI					Peptides				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
gi|238708193	Brauns (Lpp)	6740	3	3	3	3	3	30.3	212.9	39.1	19.0	8.7	40	67	57	37	33
gi|238707374	OmpA	37846	3	3	3	3	3	2.1	4.5	22.8	4.7	2.7	106	212	76	135	110
gi|238705740	Pal	18149	3	3	3	3	3	1.5	5.0	8.5	4.3	0.8	13	31	14	23	11
gi|238708170	SlyB	15382	3	3	3	3	-	1.5	1.5	4.5	3.6	-	7	10	9	16	-
gi|238706031	LptD	89769	1	2	2	3	-	0.1	0.1	2.5	0.1	-	6	7	4	7	-
gi|238706768	LptE	21133	1	2	2	3	-	0.2	0.5	2.6	0.2	-	3	4	1	4	-
gi|238706888	Slp Family	20689	2	2	3	3	-	0.4	0.4	1.7	0.4	-	3	3	6	3	-
gi|238706959	BamA/YaeT	88022	3	3	3	3	1	1.0	1.0	17.6	1.8	0.0	45	53	42	62	3
gi|238707860	BamB	42358	3	3	3	3	3	0.3	0.4	4.6	1.2	0.1	8	13	16	21	12
gi|238705699	BamC	38345	3	3	3	3	-	0.6	0.9	6.9	1.4	-	11	28	14	21	-
gi|238705321	BamD	27432	3	3	3	3	-	0.5	0.6	3.1	0.6	-	10	11	11	9	-
gi|238707927	BamE	9943	-	-	1	-	-	-	-	0.4	-	-	-	-	1	-	-
gi|238708048	VacJ	28773	3	3	3	-	-	0.5	0.5	3.1	-	-	10	10	8	-	-
gi|238706986	MltA	42895	-	-	-	1	-	-	-	-	0.2	-	-	-	-	3	-
gi|238707253	MltC	39935	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706640	MltD	48282	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707358	OmpF	40024	2	3	3	3	2	0.1	0.3	6.1	1.0	0.2	1	4	2	14	4
gi|238706321	TolC	53048	-	3	3	3	2	-	0.8	0.4	0.5	0.1	-	21	2	8	3
gi|238706508	OmpC.1	39947	-	1	-	1	-	-	0.1	-	0.1	-	-	3	-	3	-
gi|238707949	OmpC.2	40459	-	3	1	3	3	-	0.2	12.0	0.6	0.3	-	5	0	10	10
gi|238706822	OsmY	11124	2	3	3	2	3	0.3	4.1	1.8	2.0	0.3	4	20	4	7	1
gi|238706938	MetQ	29356	3	2	3	3	-	0.2	0.2	0.9	0.5	-	4	6	6	10	-
gi|238705360	TamA/YtfN	63394	3	3	3	3	-	0.1	0.2	2.1	0.3	-	10	10	7	10	-
gi|238706342	OM Siderophore receptor	73465	3	-	3	3	2	0.1	-	12.8	0.0	0.0	8	-	22	1	5
gi|238708209	HemR	72964	-	-	3	-	1	-	-	5.2	-	0.1	-	-	17	-	9
gi|238705851	FhuA	81724	2	-	3	-	1	0.0	-	21.0	-	0.0	9	-	49	-	2
gi|238706670	FiuA	77361	1	-	3	-	-	0.0	-	2.4	-	-	10	-	6	-	-
gi|238708284	Ton-B dependent receptor	76573	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707022	HasR	89969	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238708022	OmpE	39925	-	-	-	2	-	-	-	-	0.1	-	-	-	-	4	-
gi|238707630	OprC	75132	1	-	3	1	-	0.0	-	1.4	0.1	-	3	-	6	4	-
gi|238706821	OmpW	23865	-	-	-	2	-	-	-	-	0.2	-	-	-	-	2	-
gi|238707009	BtuB	68957	-	-	-	-	1	-	-	-	-	0.1	-	-	-	-	9
gi|238708047	FadL	46035	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705242	BcsC	126527	3	3	3	3	-	0.1	0.1	1.1	0.2	-	18	20	16	16	-
gi|238705556	Haemolysin activator	62594	1	-	-	1	-	0.1	-	-	0.1	-	4	-	-	2	-
gi|238705555	Hemolysin	48963	1	1	2	2	-	0.1	0.1	0.1	0.1	-	3	2	1	1	-
gi|238706248	Hemolysin activator	61801	-	-	-	-	-	-	-		-	-	-	-		-	-
gi|238707066	Pep M37	50544	1	-	-	-	-	0.1	-	-	-	-	6	-	-	-	-
gi|238705486	Phos. A1	32689	1	-	-	-	-	0.1	-	-	-	-	2	-	-	-	-
gi|238705976	Protease II	77905	-	-	1	-	-	-	-	0.1	-	-	-	-	2	-	-
gi|110007337	YhlB	62424	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705110	Possible Adhesin	155469	1	-	2	-	-	0.0	-	0.0	-	-	11	-	2	-	-
gi|238707370	PqiB/Mam7	59633	1	-	-	-	-	0.1	-	-	-	-	4	-	-	-	-
gi|238705561	AidA	44266	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705628	Fim. Usher	87773	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705013	Flagellin, partial	10509	3	1	-	-	-	6.0	0.3	-	-	-	15	2	-	-	-
gi|498532474	Flagellin	43589	3	1	-	-	-	1.7	0.7	-	-	-	30	12	-	-	-
gi|238706460	FlgE	44254	2	-	-	-	-	0.1	-	-	-	-	3	-	-	-	-
gi|238706435	FliD	49054	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707768	YiaD	22348	3	3	2	3	-	0.4	0.4	2.1	0.5	-	6	7	3	8	-
gi|238707898	YfhG	35016	2	3	1	3	-	0.1	0.3	4.0	0.2	-	4	4	0	4	-
gi|238705190	YraP	20221	3	1	2	1	2	0.2	0.2	0.6	0.2	0.2	2	1	2	1	2
gi|238705961	CsgG	24712	1	-	-	-	1	0.1	-	-	-	0.1	8	-	-	-	10
gi|238707176	SecD	71766	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705788	TTSS protein	61602	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705783	TTSS protein	45918	-	1	-	-	-	-	0.1	-	-	-	-	12	-	-	-
gi|238705857	AmiB	62746	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705632	Blc	20801	-	-	3	-	-	-	-	0.3	-	-	-	-	2	-	-
gi|238705953	Mce-Like	91347	-	-	1	-	-	-	-	0.0	-	-	-	-	6	-	-
gi|238706915	MipA	28023	-	-	-	1	-	-	-	-	0.1	-	-	-	-	5	-
gi|238707429	NagZ	99335	-	-	2	-	-	-	-	0.0	-	-	-	-	7	3	-
gi|238707352	NlpD	34028	1	-	1	2	-	0.1	-	1.6	0.2	-	3	-	2	3	-
gi|238706941	NlpE/CutF	25245	-	-	2	1	-	-	-	0.5	0.3	-	-	-	3	2	-
gi|238708334	OmpX	17086	-	-	-	1	-	-	-	-	0.2	-	-	-	-	1	-
gi|163883857	PtrA	107905	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706761	RlpA	37591	1	-	-	-	-	0.2	-	-	-	-	3	-	-	-	-
gi|46369605	RupA	17881	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706180	Unknown	17544	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|133904995	Unknown	200596	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238706101	Unknown	106480	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238705250	Unknown	40804	-	-	-	2	-	-	-	-	0.1	-	-	-	-	3	-
gi|238706025	Hypothetical	17964	2	1	-	2	-	0.2	0.2	-	0.2	-	4	6	-	7	-
gi|238705088	YeeJ	54634	1	1	-	1	-	0.1	0.1	-	0.1	-	7	3	-	2	-
gi|238706732	Hypothetical	31452	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707408	Hypothetical	56778	1	-	-	-	-	0.1	-	-	-	-	4	-	-	-	-
gi|238707442	Hypothetical	22794	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
gi|238707291	Hypothetical	54666	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

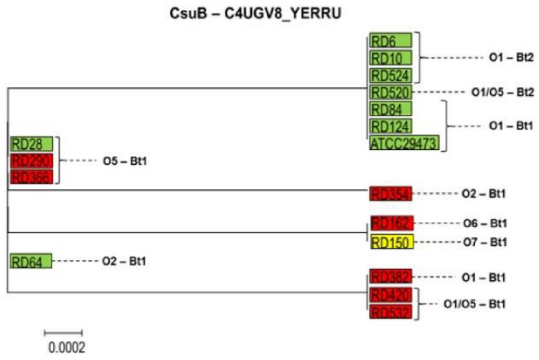
Table 8.2 Nucleotide and amino acid sequence analysis of predicted OMPs.
The NCBI and UniProt accession numbers relating to each protein are indicated. Amino acid and nucleotide variation is indicated as both individual amino acid/nucleotides and as a percentage of the protein/gene.

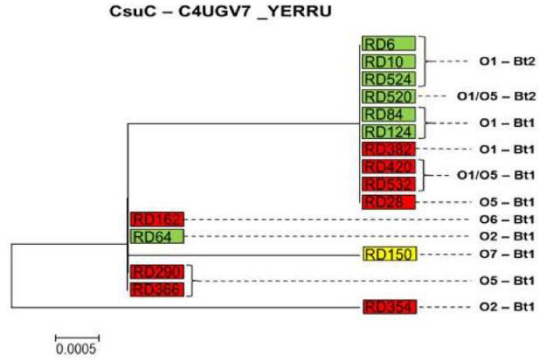
Accession		Protein		Amino Acid			Nucleotide			Genome presence															
Genome	UniProt	Name	MW (kDa)	Length (aa)	Variable sites	%	Length (nt)	Variable sites	%	RD6	RD10	RD28	RD64	RD84	RD124	RD520	RD524	RD162	RD290	RD354	RD366	RD382	RD420	RD532	RD150
gi|685058000	C4UMZ1_YERRU	AidA	44,293	397	2	0.50%	1191	6	0.50%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058889	C4UNB5_YERRU	Autotransporter	78,218	728	4	0.55%	2184	20	0.92%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059551	C4UJ19_YERRU	BamA	88,077	795	1	0.13%	2385	4	0.17%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060357	C4UGF1_YERRU	BamB/YfgL	42,328	393	2	0.51%	1179	6	0.51%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060479	C4UMN4_YERRU	BamC	38,311	349	1	0.29%	1047	1	0.10%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058394	C4UNP8_YERRU	BamD/YfiO	27,392	243	0	0.00%	729	1	0.14%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060222	C4UGL8_YERRU	BamE/SmpA	9,949	88	1	1.14%	264	1	0.38%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052900	C4UNV7_YERRU	BcsC	126,490	1,141	3	0.26%	3423	13	0.38%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058794	C4UMU7_YERRU	Blc	20,700	182	0	0.00%	546	0	0.00%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685055844	C4UIS2_YERRU	BtuB	68,999	622	137	22.03%	1887	437	23.16%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059979	C4ULV3_YERRU	CsgG	24,670	228	2	0.88%	684	4	0.58%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052782	C4UGV8_YERRU	CsuB	18,611	173	2	1.16%	519	4	0.77%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052971	C4UGV7_YERRU	CsuC	28,102	254	5	1.97%	762	8	1.05%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052972	C4UGV6_YERRU	CsuD	86,388	789	6	0.76%	2367	21	0.89%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731153986	C4UGV4_YERRU	CsuE	23,175	224	32	14.29%	672	76	11.31%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059974	C4UG91_YERRU	FadL	46,063	425	0	0.00%	1275	2	0.16%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058920	C4UM94_YERRU	FhuA	81,661	735	1	0.14%	2205	7	0.32%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058539	C4UGD2_YERRU	FimA	38,619	350	186	53.14%	1050	406	38.67%	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	+
gi|685052919	C4UGS5_YERRU	Fimbral family protein	34,279	308	2	0.65%	924	2	0.22%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058155	C4UGS4_YERRU	Fimbral Usher	88,931	808	1	0.12%	2409	8	0.33%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052781	C4UGD1_YERRU	FimD	92,257	845	9	1.07%	2535	14	0.55%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058140	C4UJX6_YERRU	FiuA	77,408	700	4	0.57%	2100	8	0.38%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060178	A0A094V4E9_YERRU	FlaA	44,042	369	125	33.88%	1107	332	29.99%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060359	C4UKI7_YERRU	FlgD	23,597	229	0	0.00%	687	3	0.44%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731156120	C4UKI6_YERRU	FlgE	44,281	427	4	0.94%	1281	6	0.47%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060177	C4UKI4_YERRU	FlgG	27,892	260	0	0.00%	780	3	0.38%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060129	C4UKI3_YERRU	FlgH	23,206	219	0	0.00%	657	5	0.76%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060451	C4UKI2_YERRU	FlgI	38,327	369	2	0.54%	1107	7	0.63%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060018	C4UKI1_YERRU	FlgK-Like	58,039	547	0	0.00%	1641	6	0.37%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059813	C4UKI0_YERRU	FlgL	34,501	322	0	0.00%	966	4	0.41%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731156090	C4UKG1_YERRU	FliD	49,084	469	0	0.00%	1407	3	0.21%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731156104	C4UKH4_YERRU	FliK	41,575	403	1	0.25%	1209	5	0.41%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|669788084	C4UNG7_YERRU	Glycoprotein receptor	16,528	159	1	0.63%	477	1	0.21%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058319	C4UIB8_YERRU	GspD	74,968	674	7	1.04%	2022	15	0.74%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685055851	C4UIT5_YERRU	HasR	90,025	803	3	0.37%	2409	15	0.62%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058718	C4UL77_YERRU	Hemolysin	157,845	1,464	60	4.10%	4392	90	2.05%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731154041	C4UNV8_YERRU	HmsP	75,524	666	6	0.90%	1998	10	0.50%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058205	C4UJP5_YERRU	Hypothetical 1	31,358	274	3	1.09%	822	12	1.46%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052857	C4UGU5_YERRU	Hypothetical 2	80,163	737	232	31.48%	2211	532	24.06%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052820	C4UGY2_YERRU	Hypothetical 3	74,053	671	4	0.60%	2013	9	0.45%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685056722	C4UMI8_YERRU	Hypothetical 4	24,405	221	156	70.59%	663	413	62.29%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731155684	A0A0B6FWT8_YERRU	Hypothetical 5	10,597	94	1	1.06%	282	3	1.06%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731156807	C4UIF9_YERRU	Hypothetical 6	26,431	241	0	0.00%	723	7	0.97%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|669790753	A0A085UA61_YERRU	Hypothetical 7	22,854	207	0	0.00%	621	1	0.16%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731156838	C4UIJ0_YERRU	Hypothetical 8	26,474	239	1	0.42%	717	5	0.70%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731155751	C4UFM9_YERRU	Hypothetical 9	12,247	113	0	0.00%	339	2	0.59%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060141	C4UFK3_YERRU	Hypothetical 10	12,150	112	0	0.00%	336	2	0.60%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|755306223	C4UJQ3_YERRU	Hypothetical 11	50,132	445	1	0.22%	1335	3	0.22%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|755306536	C4UIN2_YERRU	Hypothetical 12	10,806	97	1	1.03%	291	2	0.69%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058188	C4UNB4_YERRU	Hypothetical 13	35,297	309	0	0.00%	927	1	0.11%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058168	C4UM58_YERRU	Hypothetical 14	22,858	211	3	1.42%	633	6	0.95%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052862	C4UN28_YERRU	Invasin 2	93,186	842	0	0.00%	2526	1	0.04%	+	+	-	-	+	+	+	+	-	-	-	-	+	+	+	+
gi|685052574	C4UP93_YERRU	Invasin 3	54,668	501	1	0.20%	1503	2	0.13%	+	+	-	+	+	+	+	+	-	+	+	+	+	+	+	-
gi|685060527	C4UFT8_YERRU	Iron transport	72,952	674	4	0.59%	2022	12	0.59%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052613	C4UIQ8_YERRU	LamB	47,554	428	0	0.00%	1284	8	0.62%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|755306431	C4T4V0_YERIN	Lipoprotein 1	20,092	174	0	0.00%	522	1	0.19%	-	-	+	+	-	-	-	-	+	+	+	+	-	-	-	+
gi|685059823	C4ULV1_YERRU	Lipoprotein 2	22,244	209	1	0.48%	627	4	0.64%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058190	C4UL53_YERRU	LolB	24,168	207	1	0.48%	621	1	0.16%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052689	C4UP06_YERRU	LpoA	74,138	688	2	0.29%	2064	8	0.39%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058877	C4UHW9_YERRU	LpoB	20,654	195	1	0.51%	585	6	1.03%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060290	C4UFS2_YERRU	Lpp (Brauns)	6,687	62	0	0.00%	186	1	0.54%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059092	C4ULP1_YERRU	LptD	89,597	793	20	2.52%	2379	106	4.46%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685057941	C4UJT1_YERRU	LptE/RlpB	21,089	192	3	1.56%	576	9	1.56%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052647	C4UIR2_YERRU	MalE	43,177	398	0	0.00%	1194	1	0.08%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|669790527	C4UMT6_YERRU	Mem. protein 1	14,536	140	0	0.00%	420	1	0.24%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059512	C4UIZ8_YERRU	MetQ	29,316	271	0	0.00%	813	1	0.12%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058574	C4UJC3_YERRU	MipA	28,040	249	1	0.40%	747	5	0.67%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059422	C4UJ46_YERRU	MltA	42,866	389	5	1.29%	1167	8	0.69%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060287	C4UGA2_YERRU	MltB	39,236	354	1	0.28%	1062	8	0.75%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059461	C4UIJ5_YERRU	MltC	39,902	358	0	0.00%	1074	3	0.28%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731154743	C4UJU6_YERRU	MltD	48,312	435	1	0.23%	1305	4	0.31%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731155853	C4UKM4_YERRU	MltE	18,138	161	0	0.00%	483	0	0.00%	+	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
gi|685060195	C4UFR3_YERRU	NlpC	32,863	293	1	0.34%	879	5	0.57%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685057898	C4UK25_YERRU	NlpI	33,826	294	0	0.00%	882	2	0.23%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059281	C4UHY4_YERRU	OM Beta barrel	22,808	207	0	0.00%	621	1	0.16%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058433	C4UHR6_YERRU	OmpA	37,756	355	7	1.97%	1065	18	1.69%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058722	C4UFZ3_YERRU	OmpC.1	40,484	368	1	0.27%	1104	9	0.82%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059870	C4UKD7_YERRU	OmpC.2	39,971	365	0	0.00%	1095	3	0.27%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058036	C4UFF3_YERRU	OmpE	39,949	365	0	0.00%	900	2	0.22%	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+	-
gi|685058318	C4UHQ0_YERRU	OmpF	40,198	366	146	39.89%	1137	329	28.94%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058409	C4UNB2_YERRU	OmpL/KdgM	25,157	217	0	0.00%	651	4	0.61%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058624	C4UJH6_YERRU	OmpW	23,880	215	0	0.00%	645	0	0.00%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

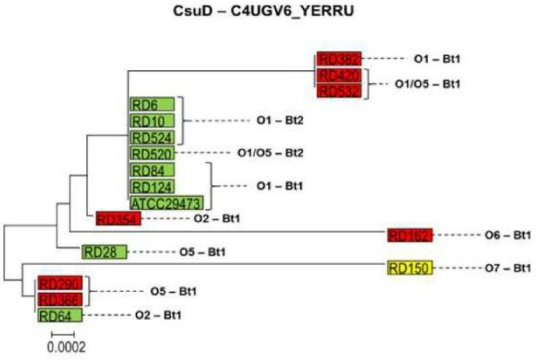
Accession		Protein		Amino Acid			Nucleotide			Genome presence																
Genome	UniProt	Name	MW (kDa)	Length (aa)	Variable sites	%	Length (nt)	Variable sites	%	RD6	RD10	RD28	RD64	RD84	RD124	RD520	RD524	RD162	RD290	RD354	RD366	RD382	RD420	RD532	RD150
gi|685057900	C4UEU8_YERRU	OmpX	17,097	161	1	0.62%	483	1	0.21%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052923	C4UH51_YERRU	OprC	75,008	686	1	0.15%	2058	8	0.39%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058306	C4UM23_YERRU	OprD/ChiP	49,223	439	0	0.00%	1317	10	0.76%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731155642	C4UJK4_YERRU	OsmB	6,799	71	0	0.00%	213	1	0.47%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731154694	C4UN57_YERRU	OutS	14,113	124	1	0.81%	372	4	1.08%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058131	C4UMF8_YERRU	Pal	18,103	167	0	0.00%	501	0	0.00%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685057903	C4UL41_YERRU	PefC	86,690	798	1	0.13%	2394	1	0.04%	+	+	-	-	+	+	+	+	+	-	+	-	+	+	+	-
gi|731157374	C4UIX9_YERRU	PepM37	50,517	457	2	0.44%	1371	6	0.44%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731155879	C4UFE1_YERRU	PhnE	34,286	306	2	0.65%	918	9	0.98%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685053812	C4UN64_YERRU	Phoslipase A1	32,709	284	1	0.35%	852	1	0.12%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060422	C4UGF6_YERRU	PilF	27,261	249	0	0.00%	747	2	0.27%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685053779	C4UKM0_YERRU	PilP	18,959	181	0	0.00%	543	0	0.00%	-	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
gi|685052780	C4UKM1_YERRU	PilQ	51,393	420	0	0.00%	1260	0	0.00%	+	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
gi|685053742	C4UKM6_YERRU	PilV	46,734	420	0	0.00%	1260	0	0.00%	+	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
gi|731155305	C4UHR2_YERRU	PqiB/Mam7	59,670	541	0	0.00%	1623	3	0.18%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058530	C4UKB8_YERRU	PrtF	50,752	448	1	0.22%	1344	6	0.45%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059441	A0A085U2W5_YERRU	RcsF	14,547	136	2	1.47%	408	1	0.25%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059034	C4UJS4_YERRU	RlpA	37,557	358	0	0.00%	1074	4	0.37%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685057886	C4UJG6_YERRU	SfmD	90,656	838	4	0.48%	2514	20	0.80%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058362	C4UPA8_YERRU	ShlA/HecA/FhaA family	32,521	303	1	0.33%	924	2	0.22%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685052665	C4UIR9_YERRU	ShlB	62,462	561	1	0.18%	1683	1	0.06%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059987	C4UKR5_YERRU	ShuA	73,397	660	3	0.45%	1980	10	0.51%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058700	C4UFB5_YERRU	Siderophore receptor	76,621	681	5	0.73%	2043	8	0.39%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059544	C4UJ18_YERRU	Skp	19,700	178	0	0.00%	534	0	0.00%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058965	C4UJ96_YERRU	Slp family	20,645	192	1	0.52%	576	3	0.52%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059839	C4UFP9_YERRU	SlyB	15,334	155	0	0.00%	465	4	0.86%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058941	C4UNT0_YERRU	Smp	25,754	232	0	0.00%	696	0	0.00%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059998	C4UKU0_YERRU	Spr	17,262	153	0	0.00%	459	0	0.00%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685057922	C4UHW8_YERRU	SteB	14,037	129	1	0.78%	387	1	0.26%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058085	C4UMU3_YERRU	SteJ	89,741	803	4	0.50%	2514	20	0.80%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731156910	C4UKV9_YERRU	T2/T3SS protein	47,342	430	3	0.70%	1290	5	0.39%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|755309125	C4UMD8_YERRU	T3SS protein	45,946	432	209	48.38%	1296	498	38.43%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|669788364	C4UKW6_YERRU	TadD	26,901	236	3	1.27%	708	6	0.85%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058056	C4UNK8_YERRU	TamA	63,376	569	3	0.53%	1707	8	0.47%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731154869	C4UHJ5_YERRU	TctC	35,930	326	0	0.00%	978	9	0.92%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059183	C4UKY9_YERRU	TolC	53,080	490	1	0.20%	1470	4	0.27%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685053786	C4UKN0_YERRU	TraL	9,748	94	0	0.00%	282	0	0.00%	+	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
gi|685053751	C4UKN3_YERRU	TraN-Like	27,705	255	0	0.00%	765	0	0.00%	+	+	-	-	+	+	+	+	-	-	-	-	-	-	-	-
gi|685058676	C4UHF4_YERRU	Tsx	29,140	259	0	0.00%	777	1	0.13%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059698	C4UG92_YERRU	VacJ	28,735	258	1	0.39%	774	3	0.39%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058261	C4UHF6_YERRU	YajG	20,929	192	1	0.52%	576	2	0.35%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059005	C4UM24_YERRU	YbfN	11,903	108	1	0.93%	324	4	1.23%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731155111	C4UF89_YERRU	YbhG/HlyD	35,741	328	0	0.00%	984	9	0.91%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|755306592	C4UL67_YERRU	YceB	21,178	191	0	0.00%	573	3	0.52%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731155365	CSF007_7410	YcfL	14,037	129	1	0.78%	387	1	0.26%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|755306197	C4UFJ9_YERRU	YdgD	31,330	285	1	0.35%	855	3	0.35%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731155707	C4UFI7_YERRU	YdgH	34,982	325	0	0.00%	975	1	0.10%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685058723	C4UNP0_YERRU	YdiY-like	27,933	255	0	0.00%	765	2	0.26%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|669790023	C4ULU5_YERRU	YebT	91,403	845	2	0.24%	2535	9	0.36%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060439	C4UG12_YERRU	YfaZ	17,593	169	0	0.00%	507	2	0.39%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685060131	C4UN85_YERRU	YfeY	20,753	194	0	0.00%	582	0	0.00%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731156502	C4UMN9_YERRU	YfgC	53,924	489	0	0.00%	1467	4	0.27%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|669788388	C4UKZ0_YERRU	YgiB	24,131	231	0	0.00%	693	3	0.43%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731157126	C4UIS0_YERRU	YhlA	169,632	1,630	11	0.67%	4890	21	0.43%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731153975	C4UGU4_YERRU	YiaD	22,248	219	2	0.91%	657	4	0.61%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|755306620	C4UGV3_YERRU	YiaF	26,618	239	0	0.00%	717	2	0.28%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|669788454	C4UIW5_YERRU	YiiQ	20,340	188	2	1.06%	564	5	0.89%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059872	C4UFK4_YERRU	YnbE	7,411	65	0	0.00%	195	2	1.03%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731157010	C4UP09_YERRU	YraP	20,176	192	0	0.00%	576	2	0.35%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|685059152	C4UME3_YERRU	YscC	61,640	558	2	0.36%	1674	6	0.36%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
gi|731154663	C4UME8_YERRU	YscJ	27,517	252	1	0.40%	756	6	0.79%	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

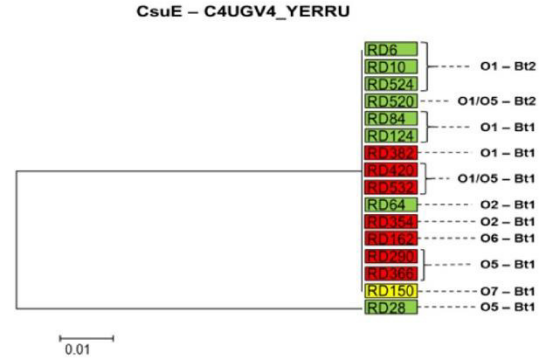

AidA – C4UMZ1_YERRU
RD290
RD366
RD28
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD150
RD382
RD420
RD532
RD162
RD64
O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O7 – Bt1
O6 – Bt1
O2 – Bt1
0.0002
Autotransporter – C4UNB5_YERRU
0.0005
BamA – C4UJ19_YERRU
0.0001
BamB/YfgL – C4UGF1 _YERRU
0.0002
BamC – C4UMN4_YERRU
0.0005
BamD – C4UNP8_YERRU
0.0001
BamE – C4UGL8_YERRU
0.0001
BscC – C4UNV7_YERRU
0.0001
Blc – C4UMU7_YERRU
BtuB – C4UIS2_YERRU
0.02

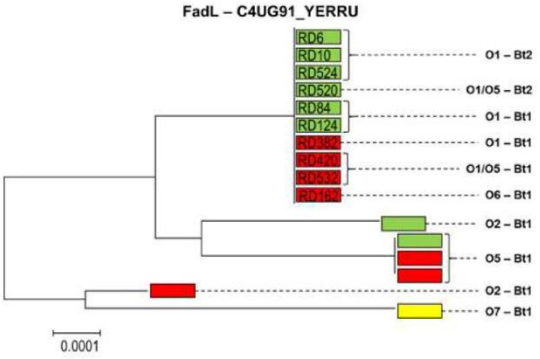

CsgG – C4ULV3_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD382
RD420
RD532
RD64
RD150
RD28
RD290
RD366
RD162
RD354
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O7 – Bt1
O5 – Bt1
O6 – Bt1
0.0002

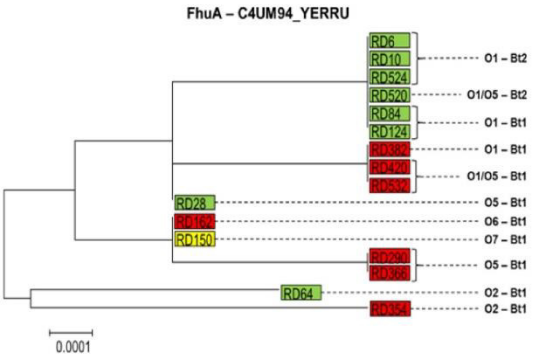

CsuB – C4UGV8_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD28
RD290
RD366
RD354
RD64
RD162
RD150
RD382
RD420
RD532
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O2 – Bt1
O6 – Bt1
O7 – Bt1
O1/O5 – Bt1
0.0002

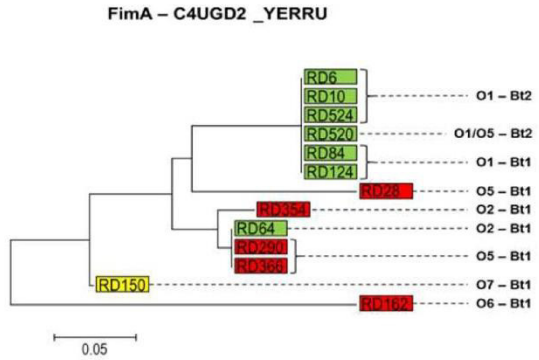

CsuC – C4UGV7 _YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD28
RD162
RD64
RD150
RD290
RD366
RD354
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O6 – Bt1
O2 – Bt1
O7 – Bt1
0.0005

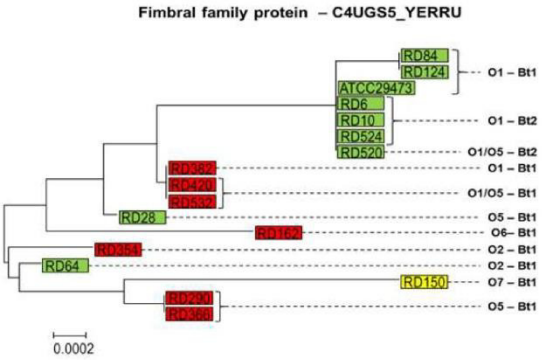

CsuD – C4UGV6_YERRU
RD382
RD420
RD532
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD354
RD162
RD28
RD150
RD290
RD366
RD64
O1 – Bt1
O1/O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O2 – Bt1
O6 – Bt1
O5 – Bt1
O7 – Bt1
0.0002

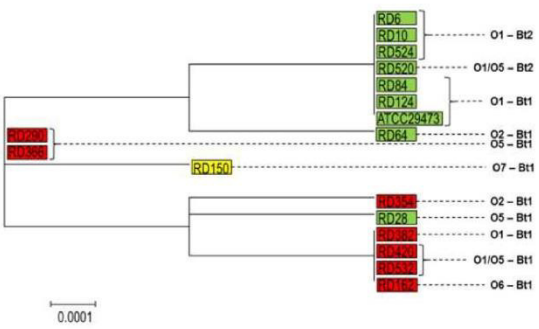

CsuE – C4UGV4_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD354
RD162
RD290
RD366
RD150
RD28
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O6 – Bt1
O5 – Bt1
O7 – Bt1
0.01

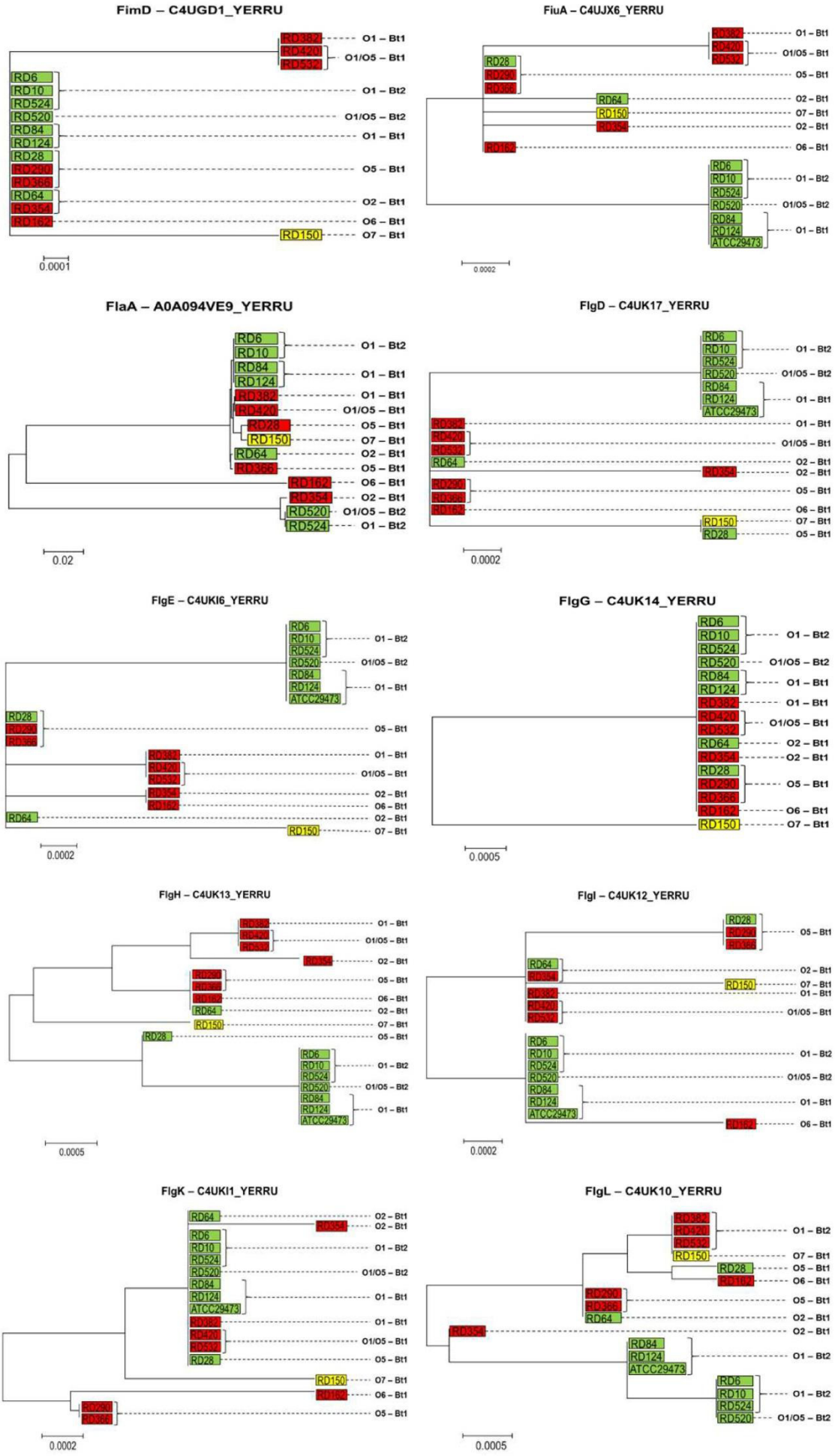

FadL – C4UG91_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD162
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1/O5 – Bt1
O6 – Bt1
O2 – Bt1
O5 – Bt1
O7 – Bt1
0.0001

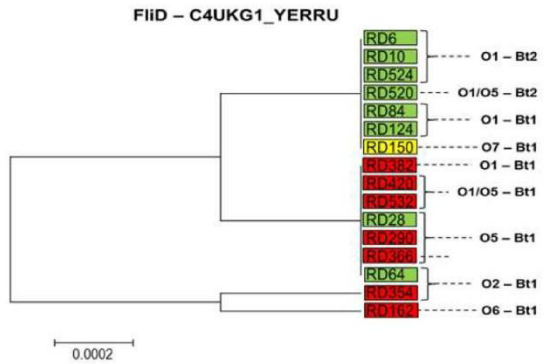

FhuA – C4UM94_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD28
RD162
RD150
RD290
RD366
RD64
RD354
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O6 – Bt1
O7 – Bt1
O2 – Bt1
0.0001

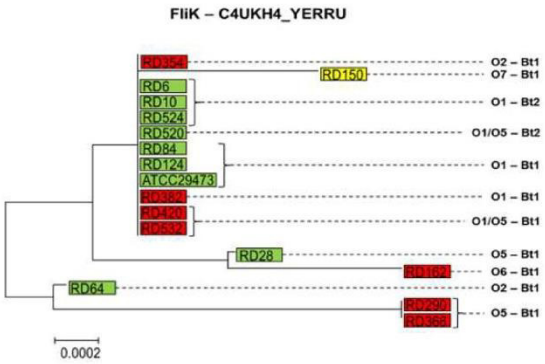

FimA – C4UGD2 _YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD28
RD354
RD64
RD290
RD366
RD150
RD162
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O2 – Bt1
O5 – Bt1
O7 – Bt1
O6 – Bt1
0.05

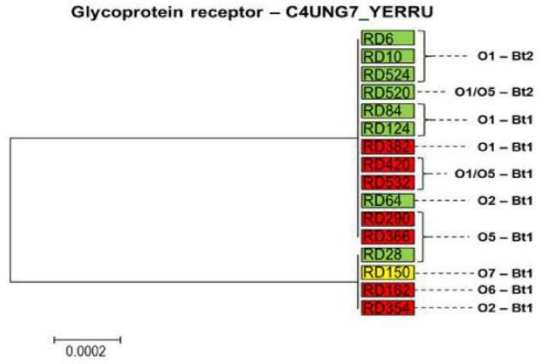

Fimbral family protein – C4UGS5_YERRU
RD84
RD124
ATCC29473
RD6
RD10
RD524
RD520
RD382
RD420
RD532
RD28
RD162
RD354
RD64
RD150
RD290
RD366
O1 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O6– Bt1
O2 – Bt1
O2 – Bt1
O7 – Bt1
O5 – Bt1
0.0002

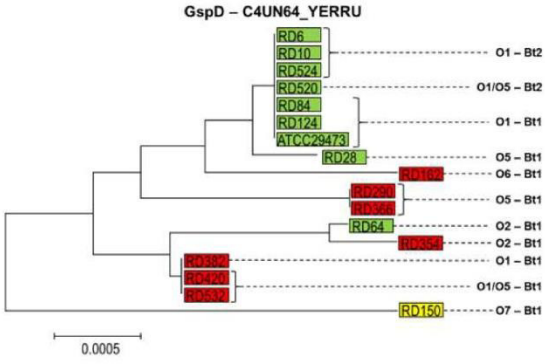

Fimbral Usher – C4UGS4_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD64
RD290
RD366
RD150
RD354
RD28
RD382
RD420
RD532
RD162
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O2 – Bt1
O5 – Bt1
O7 – Bt1
O2 – Bt1
O5 – Bt1
O1 – Bt1
O1/O5 – Bt1
O6 – Bt1
0.0001

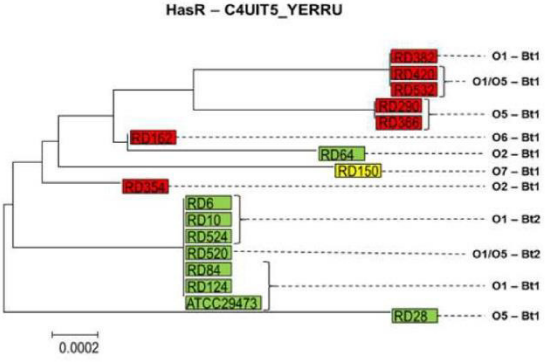

FimD – C4UGD1_YERRU
RD382 O1 – Bt1
RD420 O1/O5 – Bt1
RD532
RD6 O1 – Bt2
RD10
RD524
RD520 O1/O5 – Bt2
RD84 O1 – Bt1
RD124
RD28 O5 – Bt1
RD290
RD366
RD64 O2 – Bt1
RD354
RD162 O6 – Bt1
RD150 O7 – Bt1
0.0001
FiuA – C4UJX6_YERRU
RD382 O1 – Bt1
RD420 O1/O5 – Bt1
RD532
RD28 O5 – Bt1
RD290
RD366
RD64 O2 – Bt1
RD150 O7 – Bt1
RD354 O2 – Bt1
RD162 O6 – Bt1
RD6 O1 – Bt2
RD10
RD524
RD520 O1/O5 – Bt2
RD84 O1 – Bt1
RD124
ATCC29473
0.0002
FlaA – A0A094VE9_YERRU
RD6 O1 – Bt2
RD10
RD84 O1 – Bt1
RD124
RD382 O1 – Bt1
RD420 O1/O5 – Bt1
RD28 O5 – Bt1
RD150 O7 – Bt1
RD64 O2 – Bt1
RD366 O5 – Bt1
RD162 O6 – Bt1
RD354 O2 – Bt1
RD520 O1/O5 – Bt2
RD524 O1 – Bt2
0.02
FlgD – C4UK17_YERRU
RD6 O1 – Bt2
RD10
RD524
RD520 O1/O5 – Bt2
RD84 O1 – Bt1
RD124
ATCC29473
RD382 O1 – Bt1
RD420 O1/O5 – Bt1
RD532
RD64 O2 – Bt1
RD354 O2 – Bt1
RD290 O5 – Bt1
RD366
RD162 O6 – Bt1
RD150 O7 – Bt1
RD28 O5 – Bt1
0.0002
FlgE – C4UKI6_YERRU
RD6 O1 – Bt2
RD10
RD524
RD520 O1/O5 – Bt2
RD84 O1 – Bt1
RD124
ATCC29473
RD28 O5 – Bt1
RD290
RD366
RD382 O1 – Bt1
RD420 O1/O5 – Bt1
RD532
RD354 O2 – Bt1
RD162 O6 – Bt1
RD64 O2 – Bt1
RD150 O7 – Bt1
0.0002
FlgG – C4UK14_YERRU
RD6 O1 – Bt2
RD10
RD524
RD520 O1/O5 – Bt2
RD84 O1 – Bt1
RD124
RD382 O1 – Bt1
RD420 O1/O5 – Bt1
RD532
RD64 O2 – Bt1
RD354 O2 – Bt1
RD28 O5 – Bt1
RD290
RD366
RD162 O6 – Bt1
RD150 O7 – Bt1
0.0005
FlgH – C4UK13_YERRU
RD382 O1 – Bt1
RD420 O1/O5 – Bt1
RD532
RD354 O2 – Bt1
RD290 O5 – Bt1
RD366
RD162 O6 – Bt1
RD64 O2 – Bt1
RD150 O7 – Bt1
RD28 O5 – Bt1
RD6 O1 – Bt2
RD10
RD524
RD520 O1/O5 – Bt2
RD84 O1 – Bt1
RD124
ATCC29473
0.0005
FlgI – C4UK12_YERRU
RD28 O5 – Bt1
RD290
RD366
RD64 O2 – Bt1
RD354
RD150 O7 – Bt1
RD382 O1 – Bt1
RD420 O1/O5 – Bt1
RD532
RD6 O1 – Bt2
RD10
RD524
RD520 O1/O5 – Bt2
RD84 O1 – Bt1
RD124
ATCC29473
RD162 O6 – Bt1
0.0002
FlgK – C4UKI1_YERRU
RD64 O2 – Bt1
RD354 O2 – Bt1
RD6 O1 – Bt2
RD10
RD524
RD520 O1/O5 – Bt2
RD84 O1 – Bt1
RD124
ATCC29473
RD382 O1 – Bt1
RD420 O1/O5 – Bt1
RD532
RD28 O5 – Bt1
RD150 O7 – Bt1
RD162 O6 – Bt1
RD290 O5 – Bt1
RD366
0.0002
FlgL – C4UK10_YERRU
RD382 O1 – Bt2
RD420
RD532
RD150 O7 – Bt1
RD28 O5 – Bt1
RD162 O6 – Bt1
RD290 O5 – Bt1
RD366
RD64 O2 – Bt1
RD354 O2 – Bt1
RD84 O1 – Bt2
RD124
ATCC29473
RD6 O1 – Bt2
RD10
RD524
RD520 O1/O5 – Bt2
0.0005

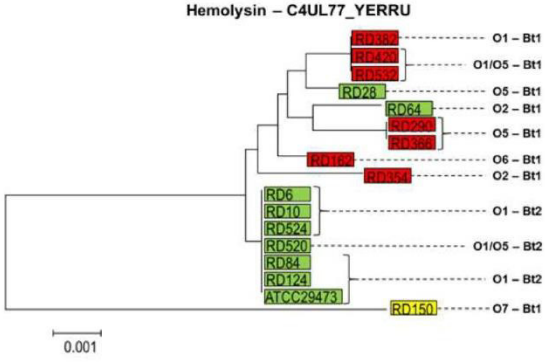

FliD – C4UKG1_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD150
RD382
RD420
RD532
RD28
RD290
RD366
RD64
RD354
RD162
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O7 – Bt1
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O2 – Bt1
O6 – Bt1
0.0002

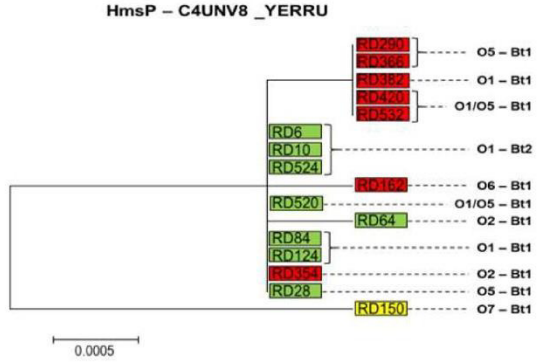

FliK – C4UKH4_YERRU
RD354
RD150
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD382
RD420
RD532
RD28
RD162
RD64
RD290
RD366
O2 – Bt1
O7 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O6 – Bt1
O2 – Bt1
O5 – Bt1
0.0002

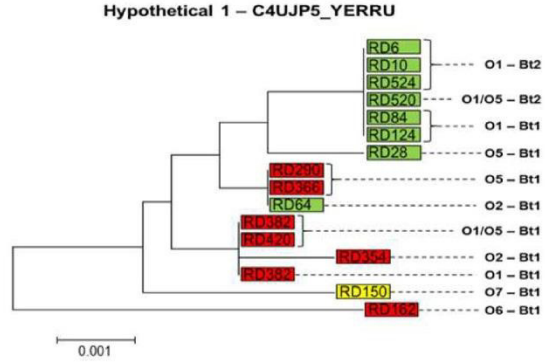

Glycoprotein receptor – C4UNG7_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD290
RD366
RD28
RD150
RD162
RD354
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O5 – Bt1
O7 – Bt1
O6 – Bt1
O2 – Bt1
0.0002

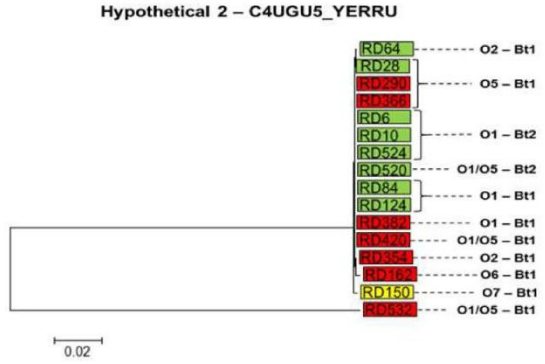

GspD – C4UN64_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD28
RD162
RD290
RD366
RD64
RD354
RD382
RD420
RD532
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O6 – Bt1
O5 – Bt1
O2 – Bt1
O2 – Bt1
O1 – Bt1
O1/O5 – Bt1
O7 – Bt1
0.0005

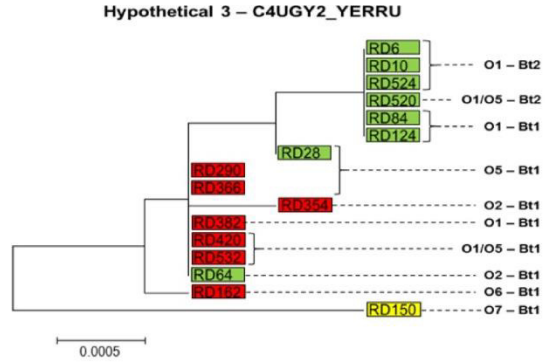

HasR – C4UIT5_YERRU
RD382
RD420
RD532
RD290
RD366
RD162
RD64
RD150
RD354
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD28
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O6 – Bt1
O2 – Bt1
O7 – Bt1
O2 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
0.0002

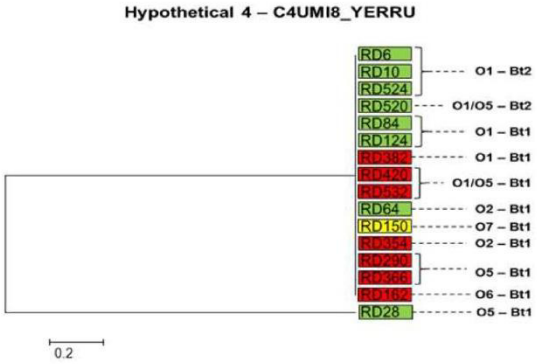

Hemolysin – C4UL77_YERRU
RD382
RD420
RD532
RD28
RD64
RD290
RD366
RD162
RD354
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD150
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O2 – Bt1
O5 – Bt1
O6 – Bt1
O2 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt2
O7 – Bt1
0.001

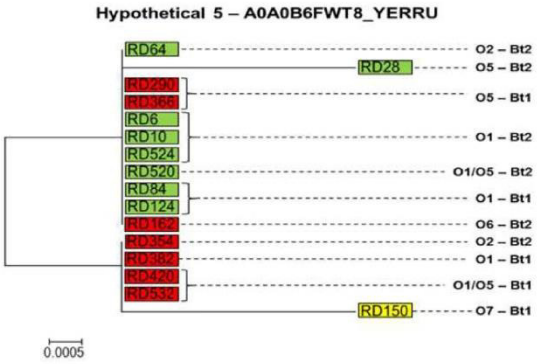

HmsP – C4UNV8 _YERRU
RD290
RD366
RD382
RD420
RD532
RD6
RD10
RD524
RD162
RD520
RD64
RD84
RD124
RD354
RD28
RD150
O5 – Bt1
O1 – Bt1
O1/O5 – Bt1
O1 – Bt2
O6 – Bt1
O1/O5 – Bt1
O2 – Bt1
O1 – Bt1
O2 – Bt1
O5 – Bt1
O7 – Bt1
0.0005

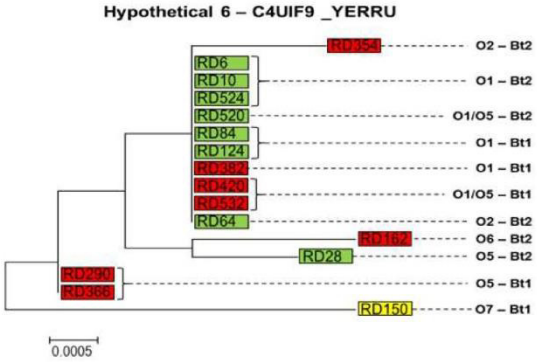

Hypothetical 1 – C4UJP5_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD28
RD290
RD366
RD64
RD382
RD420
RD354
RD382
RD150
RD162
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O5 – Bt1
O2 – Bt1
O1/O5 – Bt1
O2 – Bt1
O1 – Bt1
O7 – Bt1
O6 – Bt1
0.001

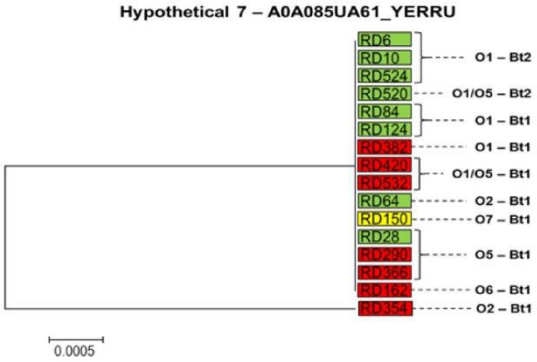

Hypothetical 2 – C4UGU5_YERRU
RD64
RD28
RD290
RD366
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD354
RD162
RD150
RD532
O2 – Bt1
O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O6 – Bt1
O7 – Bt1
O1/O5 – Bt1
0.02

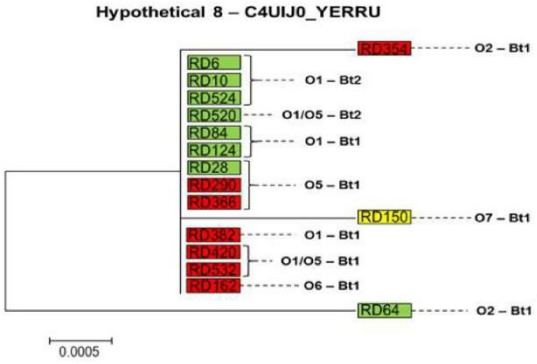

Hypothetical 3 – C4UGY2_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD28
RD290
RD366
RD354
RD382
RD420
RD532
RD64
RD162
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O2 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O6 – Bt1
O7 – Bt1
0.0005

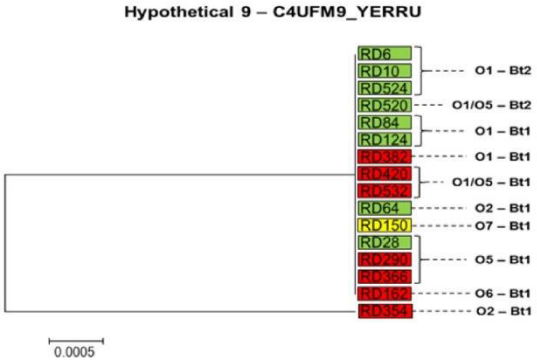

Hypothetical 4 – C4UMI8_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD150
RD354
RD290
RD366
RD162
RD28
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O7 – Bt1
O2 – Bt1
O5 – Bt1
O6 – Bt1
O5 – Bt1
0.2

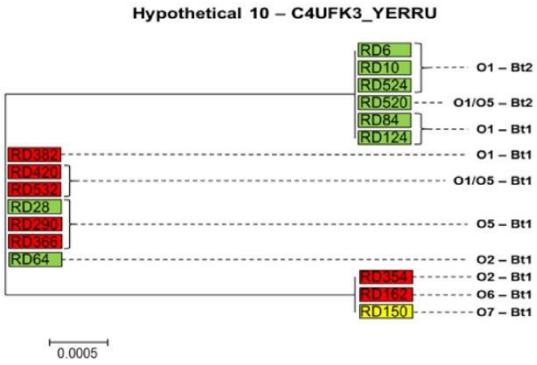

Hypothetical 5 – A0A0B6FWT8_YERRU
RD64
RD28
RD290
RD366
RD6
RD10
RD524
RD520
RD84
RD124
RD162
RD354
RD382
RD420
RD532
RD150
O2 – Bt2
O5 – Bt2
O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O6 – Bt2
O2 – Bt2
O1 – Bt1
O1/O5 – Bt1
O7 – Bt1
0.0005

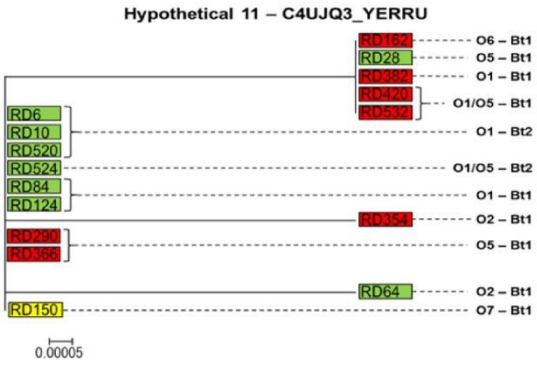

Hypothetical 6 – C4UIF9 _YERRU
RD354
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD162
RD28
RD290
RD366
RD150
O2 – Bt2
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt2
O6 – Bt2
O5 – Bt2
O5 – Bt1
O7 – Bt1
0.0005

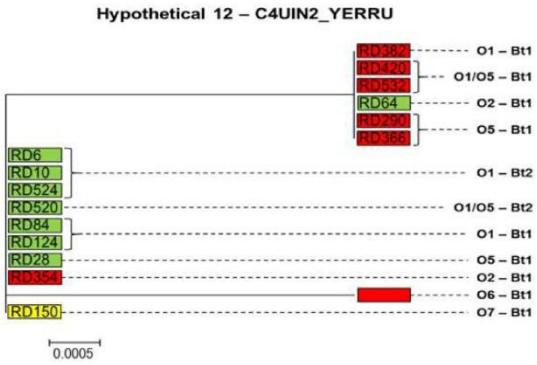

Hypothetical 7 – A0A085UA61_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD150
RD28
RD290
RD366
RD162
RD354
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O7 – Bt1
O5 – Bt1
O6 – Bt1
O2 – Bt1
0.0005

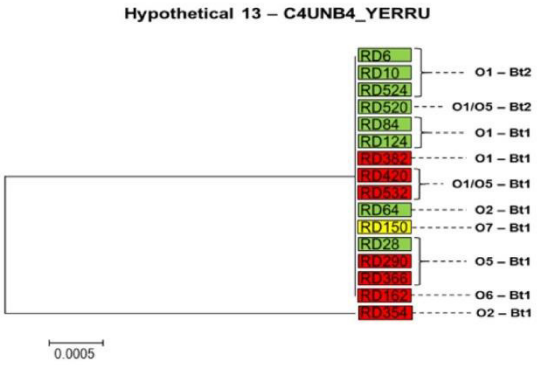

Hypothetical 8 – C4UIJ0_YERRU
RD354
RD6
RD10
RD524
RD520
RD84
RD124
RD28
RD290
RD366
RD150
RD382
RD420
RD532
RD162
RD64
O2 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O7 – Bt1
O1 – Bt1
O1/O5 – Bt1
O6 – Bt1
O2 – Bt1
0.0005

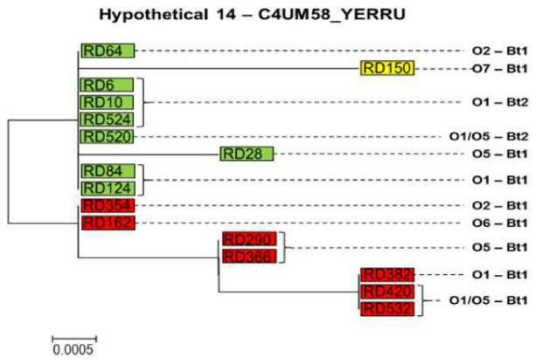

Hypothetical 9 – C4UFM9_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD150
RD28
RD290
RD366
RD162
RD354
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O7 – Bt1
O5 – Bt1
O6 – Bt1
O2 – Bt1
0.0005

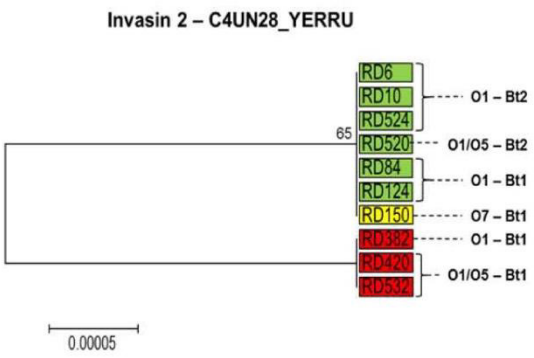

Hypothetical 10 – C4UFK3_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD28
RD290
RD366
RD64
RD354
RD162
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O2 – Bt1
O2 – Bt1
O6 – Bt1
O7 – Bt1
0.0005

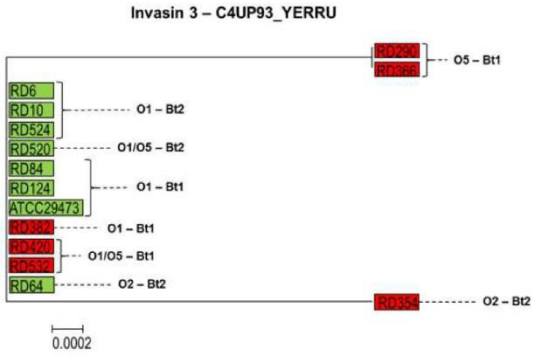

Hypothetical 11 – C4UJQ3_YERRU
RD162
RD28
RD382
RD420
RD532
RD6
RD10
RD520
RD524
RD84
RD124
RD354
RD290
RD366
RD64
RD150
O6 – Bt1
O5 – Bt1
O1 – Bt1
O1/O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O2 – Bt1
O5 – Bt1
O2 – Bt1
O7 – Bt1
0.00005

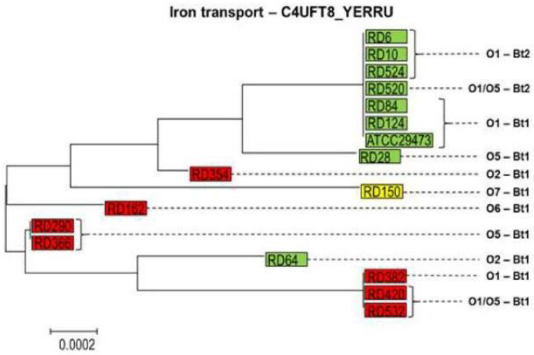

Hypothetical 12 – C4UIN2_YERRU
RD382
RD420
RD532
RD64
RD290
RD366
RD6
RD10
RD524
RD520
RD84
RD124
RD28
RD354
RD150
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O2 – Bt1
O6 – Bt1
O7 – Bt1
0.0005

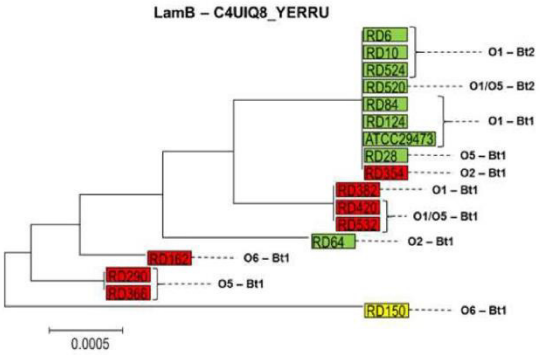

Hypothetical 13 – C4UNB4_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD150
RD28
RD290
RD366
RD162
RD354
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O7 – Bt1
O5 – Bt1
O6 – Bt1
O2 – Bt1
0.0005

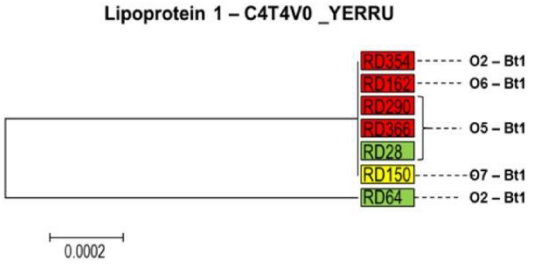

Hypothetical 14 – C4UM58_YERRU
RD64
RD150
RD6
RD10
RD524
RD520
RD28
RD84
RD124
RD354
RD162
RD290
RD366
RD382
RD420
RD532
O2 – Bt1
O7 – Bt1
O1 – Bt2
O1/O5 – Bt2
O5 – Bt1
O1 – Bt1
O6 – Bt1
0.0005

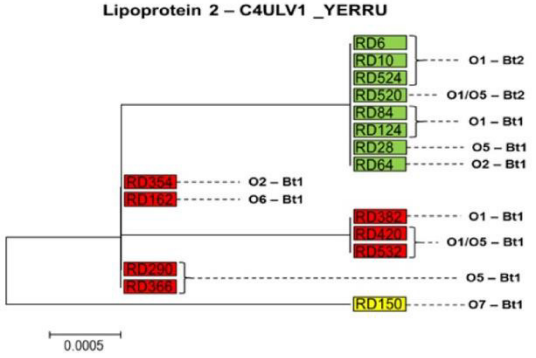

Invasin 2 – C4UN28_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD150
RD382
RD420
RD532
65
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O7 – Bt1
O1 – Bt1
O1/O5 – Bt1
0.00005

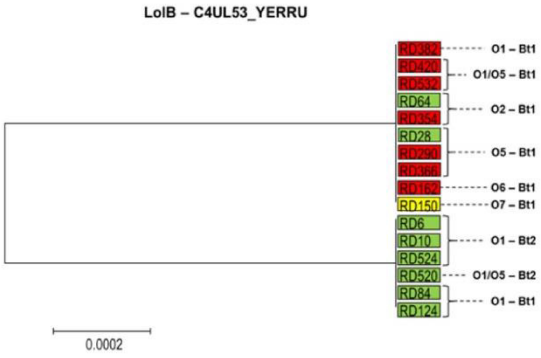

Invasin 3 – C4UP93_YERRU
RD290
RD366
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD382
RD420
RD532
RD64
RD354
O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt2
O2 – Bt2
0.0002

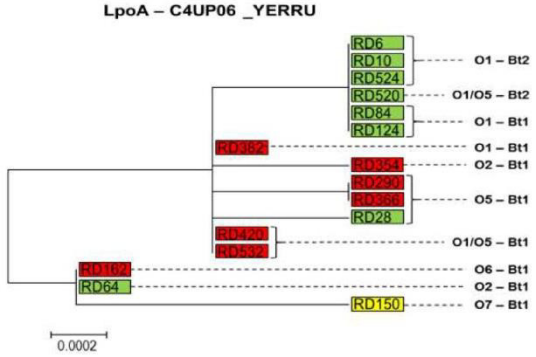

Iron transport – C4UFT8_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD28
RD354
RD150
RD162
RD290
RD366
RD64
RD382
RD420
RD532
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O2 – Bt1
O7 – Bt1
O6 – Bt1
O5 – Bt1
O2 – Bt1
O1 – Bt1
O1/O5 – Bt1
0.0002

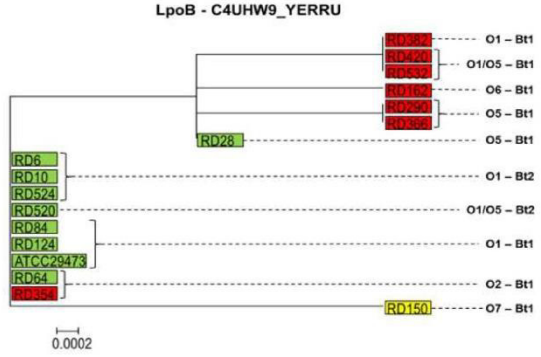

LamB – C4UIQ8_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD28
RD354
RD382
RD420
RD532
RD64
RD162
RD290
RD366
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O2 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O6 – Bt1
O5 – Bt1
O6 – Bt1
0.0005

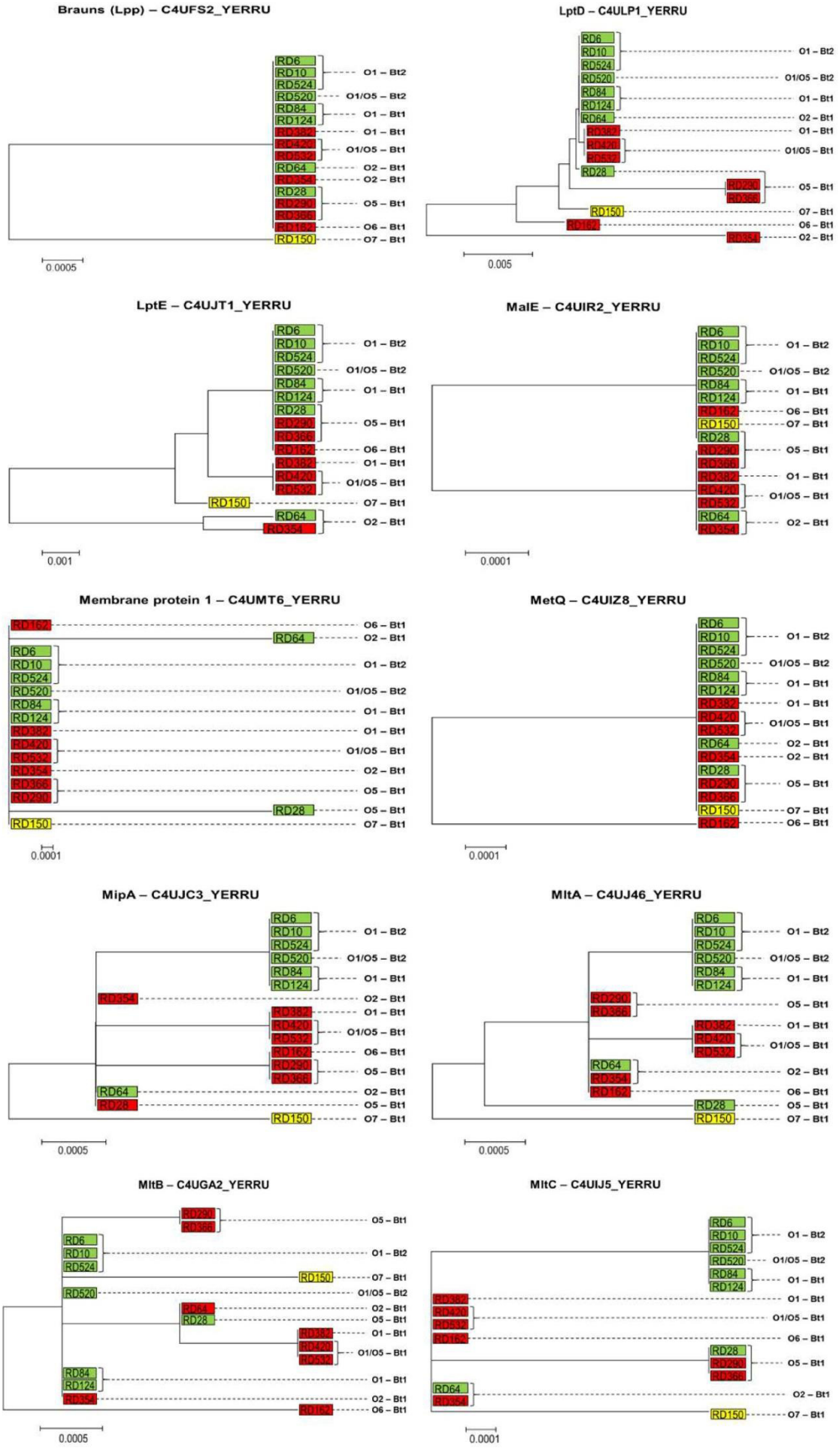

Lipoprotein 1 – C4T4V0 _YERRU
RD354
RD162
RD290
RD366
RD28
RD150
RD64
O2 – Bt1
O6 – Bt1
O5 – Bt1
O7 – Bt1
O2 – Bt1
0.0002

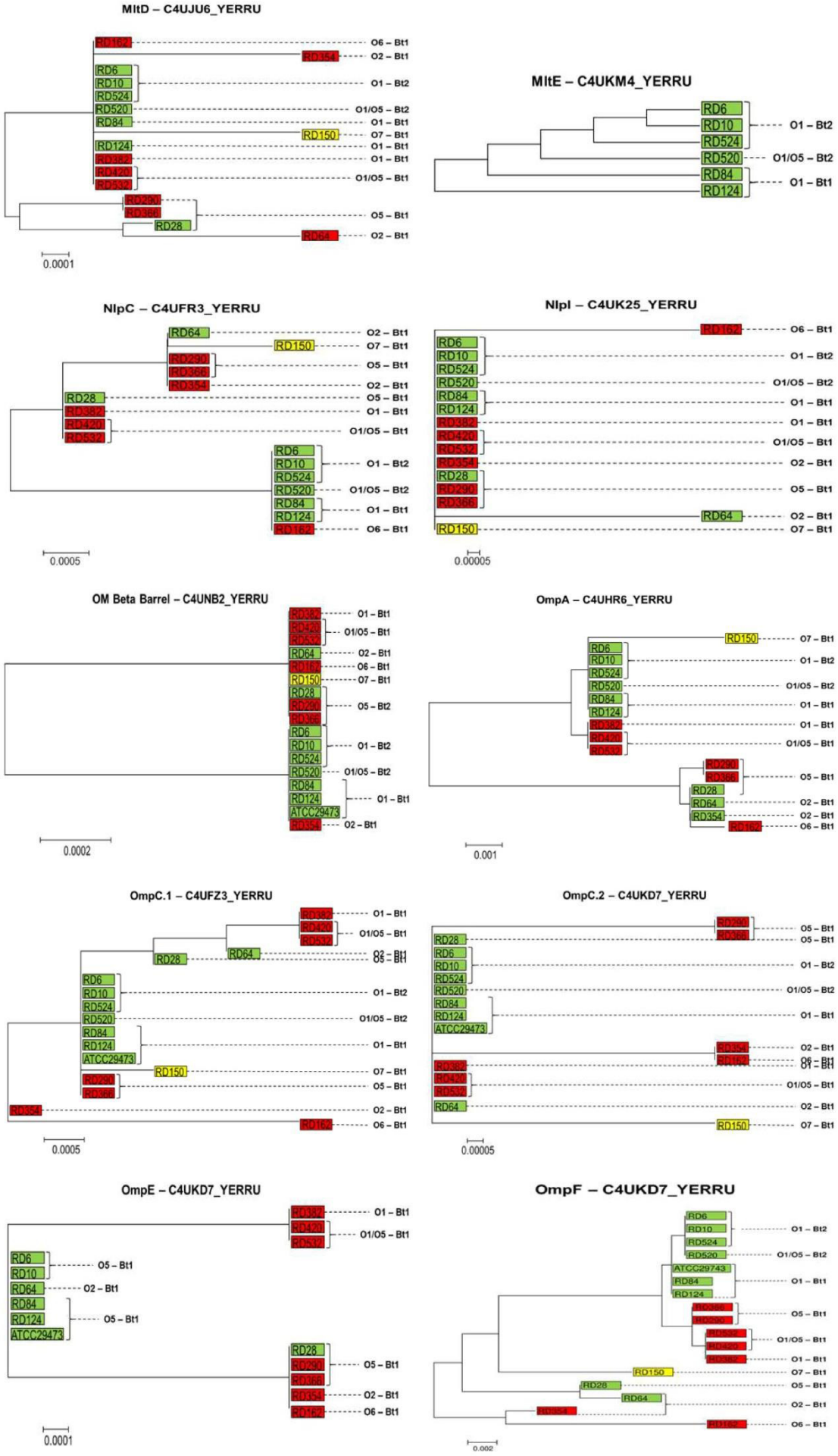

Lipoprotein 2 – C4ULV1 _YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD28
RD64
RD354
RD162
RD382
RD420
RD532
RD290
RD366
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O2 – Bt1
O2 – Bt1
O6 – Bt1
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O7 – Bt1
0.0005

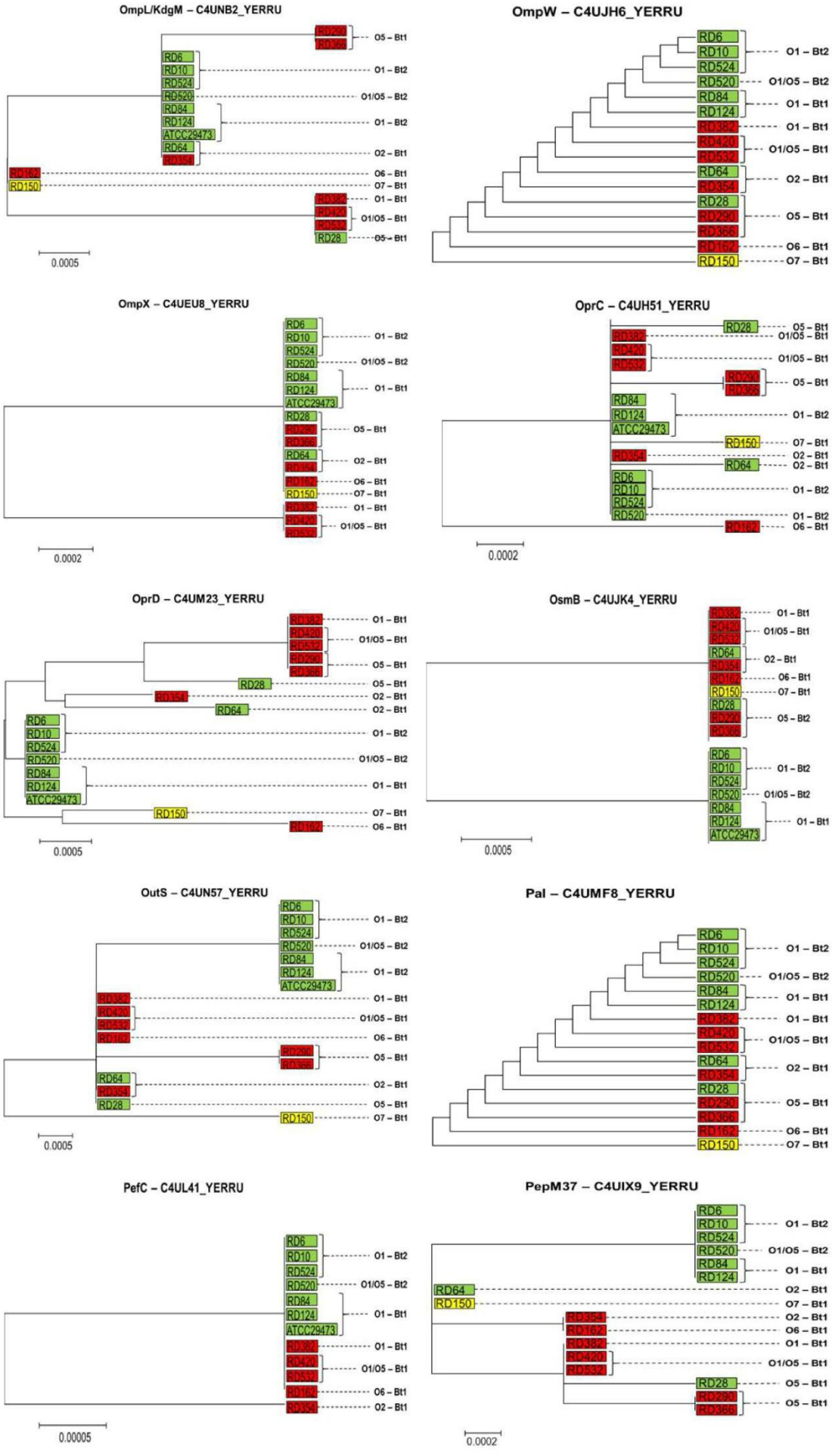

LolB – C4UL53_YERRU
RD382
RD420
RD532
RD64
RD354
RD28
RD290
RD366
RD162
RD150
RD6
RD10
RD524
RD520
RD84
RD124
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O5 – Bt1
O6 – Bt1
O7 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
0.0002


LpoA – C4UP06 _YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD354
RD290
RD366
RD28
RD420
RD532
RD162
RD64
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O2 – Bt1
O5 – Bt1
O1/O5 – Bt1
O6 – Bt1
O2 – Bt1
O7 – Bt1
0.0002

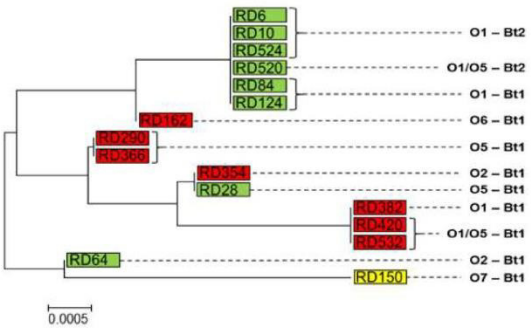

LpoB - C4UHW9_YERRU
RD382
RD420
RD532
RD162
RD290
RD366
RD28
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD64
RD354
RD150
O1 – Bt1
O1/O5 – Bt1
O6 – Bt1
O5 – Bt1
O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O2 – Bt1
O7 – Bt1
0.0002


Brauns (Lpp) – C4UFS2_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD354
RD28
RD290
RD366
RD162
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O2 – Bt1
O5 – Bt1
O6 – Bt1
O7 – Bt1
0.0005
LptD – C4ULP1_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD64
RD382
RD420
RD532
RD28
RD290
RD366
RD150
RD162
RD354
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O2 – Bt1
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O7 – Bt1
O6 – Bt1
O2 – Bt1
0.005
LptE – C4UJT1_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD28
RD290
RD366
RD162
RD382
RD420
RD532
RD150
RD64
RD354
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O6 – Bt1
O1 – Bt1
O1/O5 – Bt1
O7 – Bt1
O2 – Bt1
0.001
MalE – C4UIR2_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD162
RD150
RD28
RD290
RD366
RD382
RD420
RD532
RD64
RD354
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O6 – Bt1
O7 – Bt1
O5 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
0.0001
Membrane protein 1 – C4UMT6_YERRU
RD162
RD64
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD354
RD366
RD290
RD28
RD150
O6 – Bt1
O2 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O5 – Bt1
O5 – Bt1
O7 – Bt1
0.0001
MetQ – C4UIZ8_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD354
RD28
RD290
RD366
RD150
RD162
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O2 – Bt1
O5 – Bt1
O7 – Bt1
O6 – Bt1
0.0001
MipA – C4UJC3_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD354
RD382
RD420
RD532
RD162
RD290
RD366
RD64
RD28
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O2 – Bt1
O1 – Bt1
O1/O5 – Bt1
O6 – Bt1
O5 – Bt1
O2 – Bt1
O5 – Bt1
O7 – Bt1
0.0005
MltA – C4UJ46_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD290
RD366
RD382
RD420
RD532
RD64
RD354
RD162
RD28
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O6 – Bt1
O5 – Bt1
O7 – Bt1
0.0005
MltB – C4UGA2_YERRU
RD290
RD366
RD6
RD10
RD524
RD150
RD520
RD64
RD28
RD382
RD420
RD532
RD84
RD124
RD354
RD162
O5 – Bt1
O1 – Bt2
O7 – Bt1
O1/O5 – Bt2
O2 – Bt1
O5 – Bt1
O1 – Bt1
O1/O5 – Bt1
O1 – Bt1
O2 – Bt1
O6 – Bt1
0.0005
MltC – C4UIJ5_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD162
RD28
RD290
RD366
RD64
RD354
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O6 – Bt1
O5 – Bt1
O2 – Bt1
O7 – Bt1
0.0001

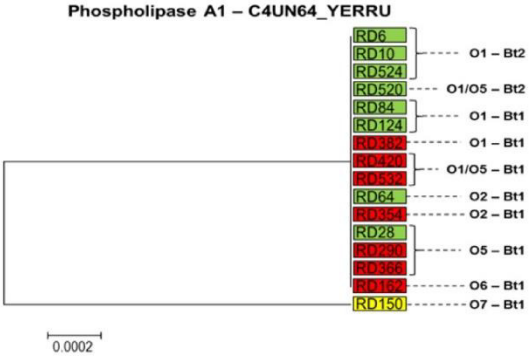

MltD – C4UJU6_YERRU
RD162
RD354
RD6
RD10
RD524
RD520
RD84
RD150
RD124
RD382
RD420
RD532
RD290
RD366
RD28
RD64
O6 – Bt1
O2 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O7 – Bt1
O1/O5 – Bt1
O5 – Bt1
0.0001
MltE – C4UKM4_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
NlpC – C4UFR3_YERRU
RD64
RD150
RD290
RD366
RD354
RD28
RD382
RD420
RD532
RD6
RD10
RD524
RD520
RD84
RD124
RD162
O2 – Bt1
O7 – Bt1
O5 – Bt1
O2 – Bt1
O5 – Bt1
O1 – Bt1
O1/O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O6 – Bt1
0.0005
NlpI – C4UK25_YERRU
RD162
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD354
RD28
RD290
RD366
RD64
RD150
O6 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O5 – Bt1
O2 – Bt1
O7 – Bt1
0.00005
OM Beta Barrel – C4UNB2_YERRU
RD382
RD420
RD532
RD64
RD162
RD150
RD28
RD290
RD366
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD354
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O6 – Bt1
O7 – Bt1
O5 – Bt2
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O2 – Bt1
0.0002
OmpA – C4UHR6_YERRU
RD150
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD290
RD366
RD28
RD64
RD354
RD162
O7 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O2 – Bt1
O2 – Bt1
O6 – Bt1
0.001
OmpC.1 – C4UFZ3_YERRU
RD382
RD420
RD532
RD64
RD28
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD150
RD290
RD366
RD354
RD162
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O7 – Bt1
O5 – Bt1
O2 – Bt1
O6 – Bt1
0.0005
OmpC.2 – C4UKD7_YERRU
RD290
RD366
RD28
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD354
RD162
RD382
RD420
RD532
RD64
RD150
O5 – Bt1
O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O2 – Bt1
O6 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O7 – Bt1
0.00005
OmpE – C4UKD7_YERRU
RD382
RD420
RD532
RD6
RD10
RD64
RD84
RD124
ATCC29473
RD28
RD290
RD366
RD354
RD162
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O2 – Bt1
O5 – Bt1
O5 – Bt1
O2 – Bt1
O6 – Bt1
0.0001
OmpF – C4UKD7_YERRU
RD6
RD10
RD524
RD520
ATCC29743
RD84
RD124
RD366
RD290
RD532
RD420
RD382
RD150
RD28
RD64
RD354
RD162
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O1/O5 – Bt1
O1 – Bt1
O7 – Bt1
O5 – Bt1
O2 – Bt1
O6 – Bt1
0.002


OmpL/KdgM – C4UNB2_YERRU
OmpW – C4UJH6_YERRU
OmpX – C4UEU8_YERRU
OprC – C4UH51_YERRU
OprD – C4UM23_YERRU
OsmB – C4UJK4_YERRU
OutS – C4UN57_YERRU
Pal – C4UMF8_YERRU
PefC – C4UL41_YERRU
PepM37 – C4UIX9_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD64
RD354
RD162
RD150
RD382
RD420
RD532
RD28
RD290
RD366
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O2 – Bt1
O6 – Bt1
O7 – Bt1
O5 – Bt1
O1/O5 – Bt1
0.0005
0.0002
0.00005

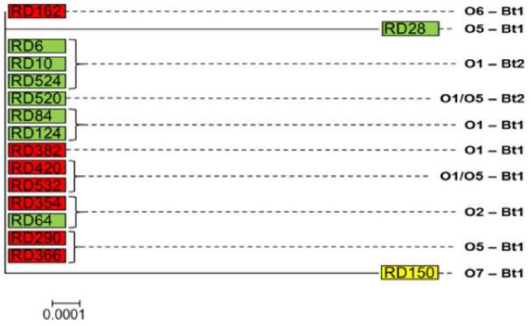
PhnE – C4UFE1_YERRU



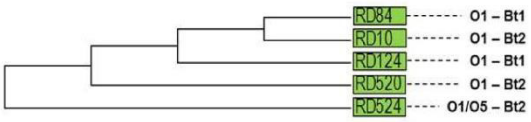
Phospholipase A1 – C4UN64_YERRU



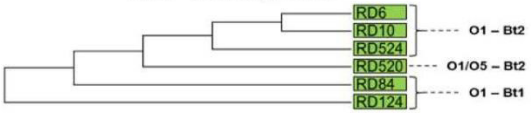
PilF – C4UGF6_YERRU



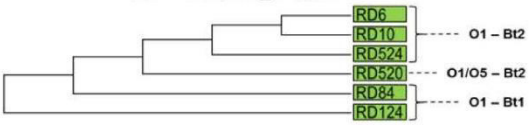
PilP – C4UKM0_YERRU



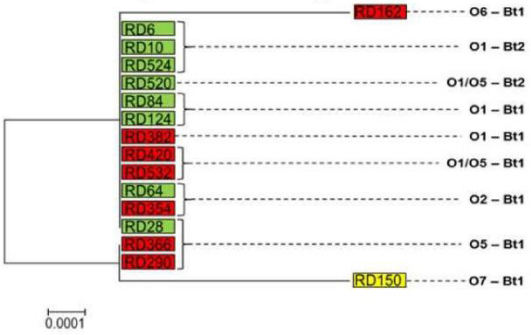
PilQ – C4UKM1_YERRU



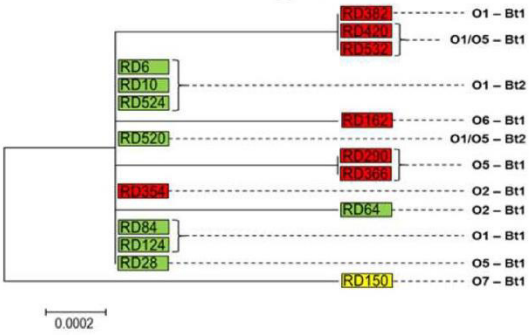
PilV – C4UKM6_YERRU



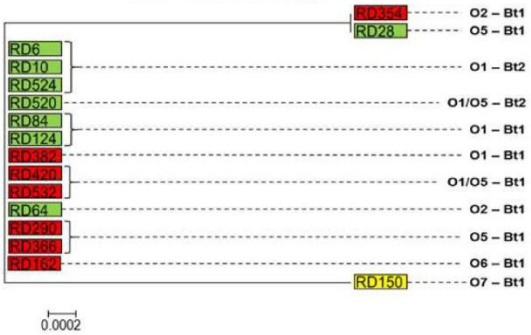
PqiB/Mam7 – C4UHR2_YERRU



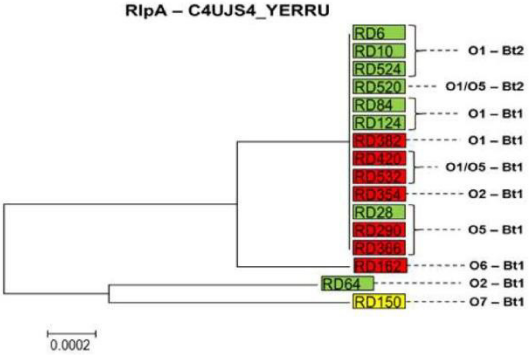
PrtF – C4UKB8_YERRU

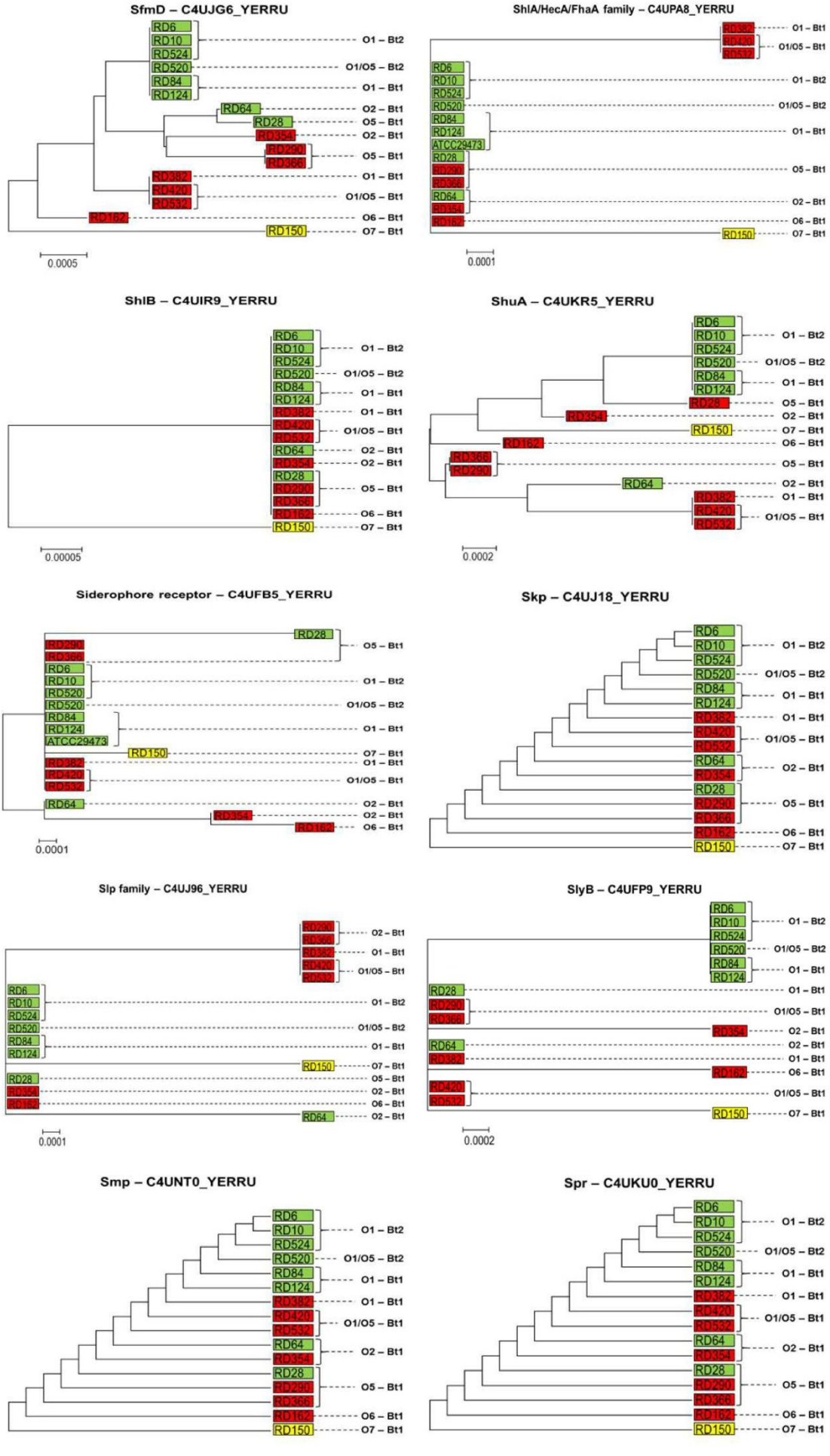


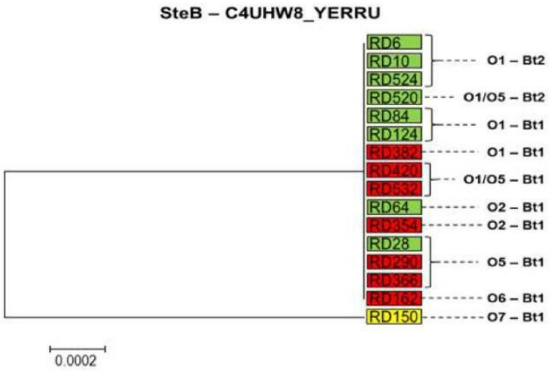
RcsF –A0A085U2W5_YERRU

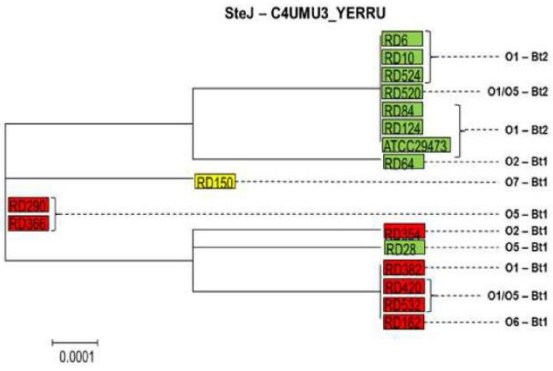


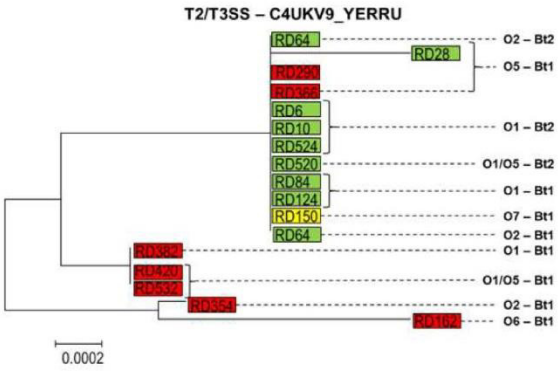
RlpA – C4UJS4_YERRU

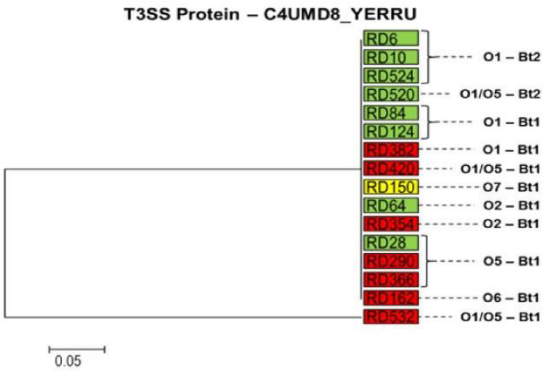


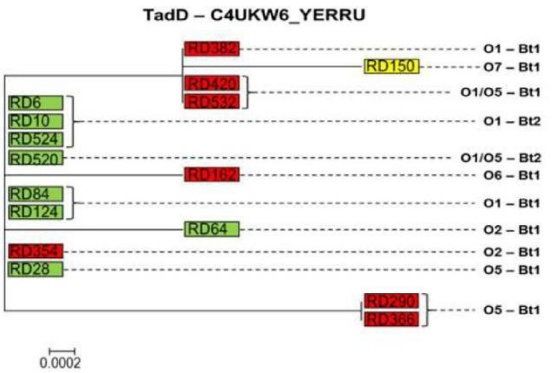

SfmD – C4UJG6_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD64
RD28
RD354
RD290
RD366
RD382
RD420
RD532
RD162
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O2 – Bt1
O5 – Bt1
O1/O5 – Bt1
O6 – Bt1
O7 – Bt1
0.0005
ShlA/HecA/FhaA family – C4UPA8_YERRU
ATCC29473
0.0001
ShlB – C4UIR9_YERRU
0.00005
ShuA – C4UKR5_YERRU
0.0002
Siderophore receptor – C4UFB5_YERRU
0.0001
Skp – C4UJ18_YERRU
Slp family – C4UJ96_YERRU
0.0001
SlyB – C4UFP9_YERRU
0.0002
Smp – C4UNT0_YERRU
Spr – C4UKU0_YERRU

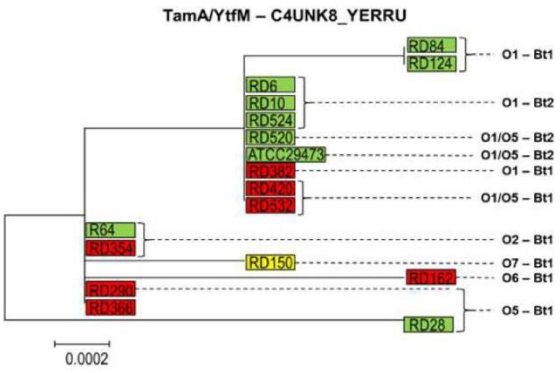

SteB – C4UHW8_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD354
RD28
RD290
RD366
RD162
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O2 – Bt1
O5 – Bt1
O6 – Bt1
O7 – Bt1
0.0002

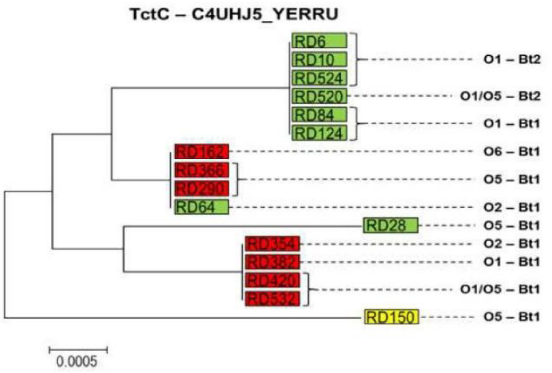

SteJ – C4UMU3_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD64
RD150
RD290
RD366
RD354
RD28
RD382
RD420
RD532
RD162
01 – Bt2
01/05 – Bt2
01 – Bt2
02 – Bt1
07 – Bt1
05 – Bt1
02 – Bt1
05 – Bt1
01 – Bt1
01/05 – Bt1
06 – Bt1
0.0001

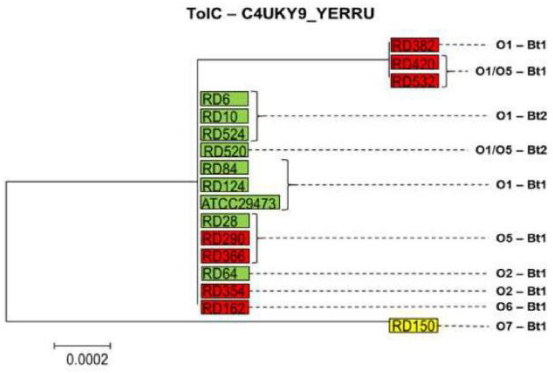

T2/T3SS – C4UKV9_YERRU
RD64
RD28
RD290
RD366
RD6
RD10
RD524
RD520
RD84
RD124
RD150
RD382
RD420
RD532
RD354
RD162
O2 – Bt2
O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O7 – Bt1
O2 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O6 – Bt1
0.0002

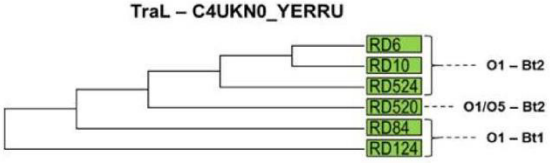

T3SS Protein – C4UMD8_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD150
RD64
RD354
RD28
RD290
RD366
RD162
RD532
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O7 – Bt1
O2 – Bt1
O2 – Bt1
O5 – Bt1
O6 – Bt1
O1/O5 – Bt1
0.05

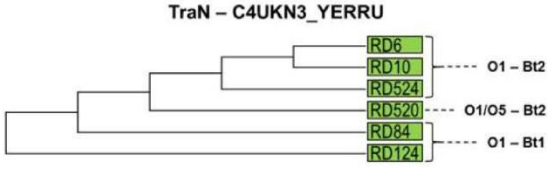

TadD – C4UKW6_YERRU
RD382
RD150
RD420
RD532
RD6
RD10
RD524
RD520
RD162
RD84
RD124
RD64
RD354
RD28
RD290
RD366
O1 – Bt1
O7 – Bt1
O1/O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O6 – Bt1
O1 – Bt1
O2 – Bt1
O2 – Bt1
O5 – Bt1
O5 – Bt1
0.0002

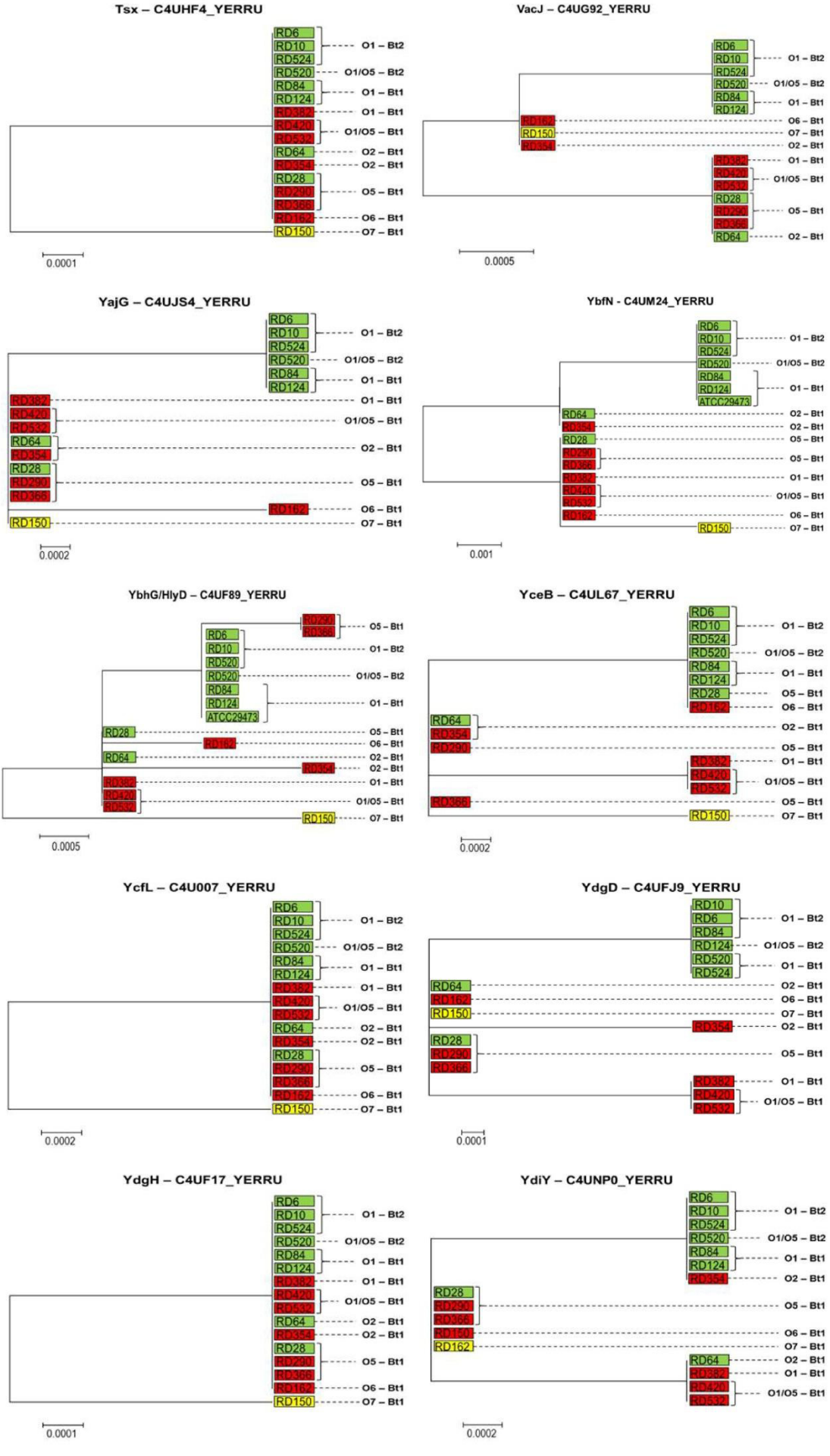

TamA/YtfM – C4UNK8_YERRU
RD84
RD124
RD6
RD10
RD524
RD520
ATCC29473
RD382
RD420
RD532
R64
RD354
RD150
RD162
RD290
RD366
RD28
O1 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O7 – Bt1
O6 – Bt1
O5 – Bt1
0.0002

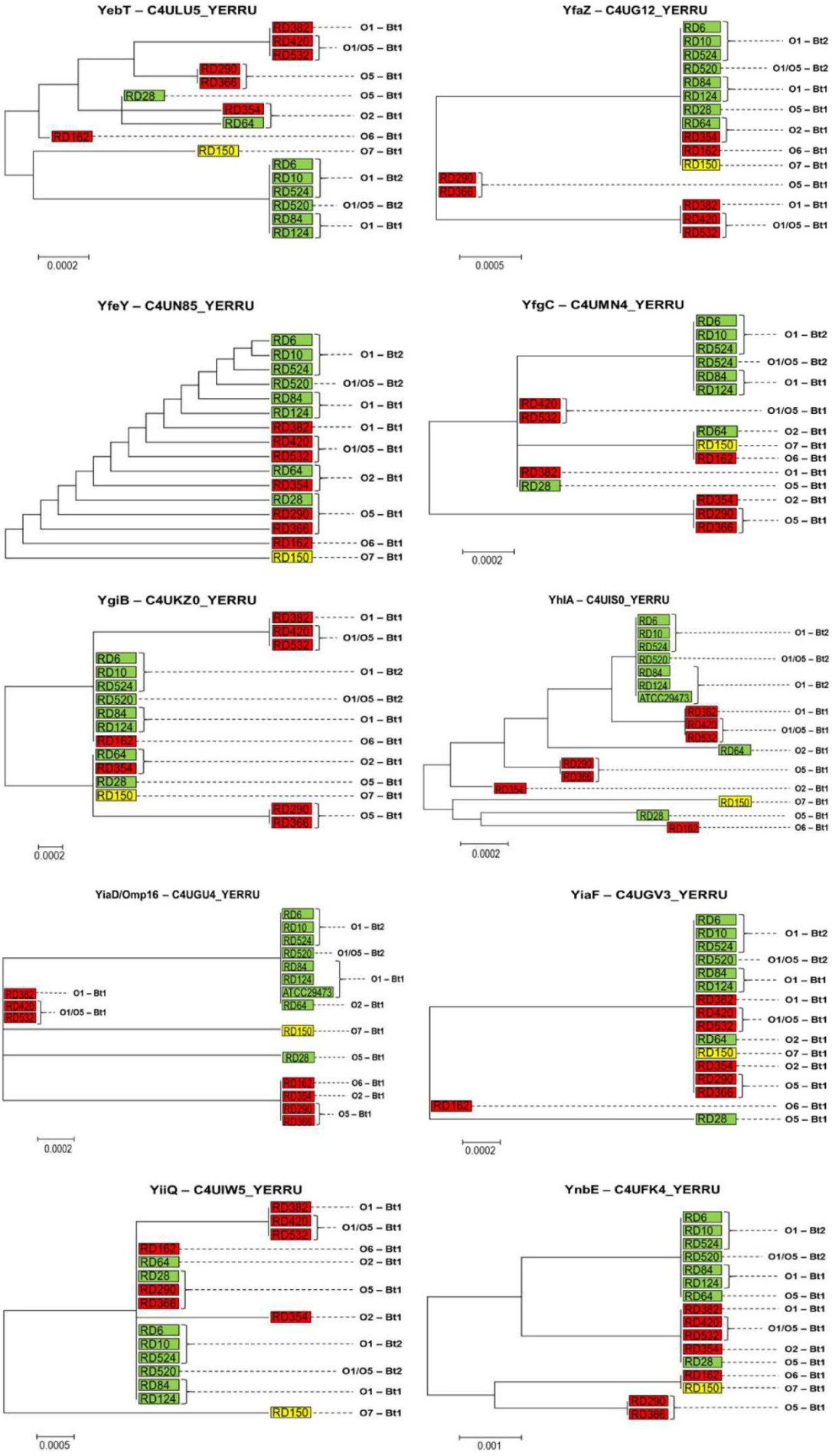

TctC – C4UHJ5_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD162
RD366
RD290
RD64
RD28
RD354
RD382
RD420
RD532
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O6 – Bt1
O5 – Bt1
O2 – Bt1
O5 – Bt1
O2 – Bt1
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
0.0005


TolC – C4UKY9_YERRU
RD382
RD420
RD532
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD28
RD290
RD366
RD64
RD354
RD162
RD150
O1 – Bt1
O1/O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O2 – Bt1
O2 – Bt1
O6 – Bt1
O7 – Bt1
0.0002


TraL – C4UKN0_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1


TraN – C4UKN3_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1


Tsx – C4UHF4_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD354
RD28
RD290
RD366
RD162
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O5 – Bt1
O6 – Bt1
O7 – Bt1
0.0001
VacJ – C4UG92_YERRU
0.0005
YajG – C4UJS4_YERRU
0.0002
YbfN - C4UM24_YERRU
ATCC29473
0.001
YbhG/HlyD – C4UF89_YERRU
0.0005
YceB – C4UL67_YERRU
0.0002
YcfL – C4U007_YERRU
0.0002
YdgD – C4UFJ9_YERRU
0.0001
YdgH – C4UF17_YERRU
0.0001
YdiY – C4UNP0_YERRU
0.0002


YebT – C4ULU5_YERRU
RD382
RD420
RD532
RD290
RD366
RD28
RD354
RD64
RD162
RD150
RD6
RD10
RD524
RD520
RD84
RD124
O1 – Bt1
O1/O5 – Bt1
O5 – Bt1
O2 – Bt1
O6 – Bt1
O7 – Bt1
O1 – Bt2
O1/O5 – Bt2
0.0002
YfaZ – C4UG12_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD28
RD64
RD354
RD162
RD150
RD290
RD366
RD382
RD420
RD532
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O2 – Bt1
O6 – Bt1
O7 – Bt1
0.0005
YfeY – C4UN85_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD354
RD28
RD290
RD366
RD162
RD150
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O2 – Bt1
O5 – Bt1
O6 – Bt1
O7 – Bt1
YfgC – C4UMN4_YERRU
RD6
RD10
RD524
RD524
RD84
RD124
RD420
RD532
RD64
RD150
RD162
RD382
RD28
RD354
RD290
RD366
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O7 – Bt1
O6 – Bt1
O5 – Bt1
0.0002
YgiB – C4UKZ0_YERRU
RD382
RD420
RD532
RD6
RD10
RD524
RD520
RD84
RD124
RD162
RD64
RD354
RD28
RD150
RD290
RD366
O1 – Bt1
O1/O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O6 – Bt1
O2 – Bt1
O7 – Bt1
0.0002
YhlA – C4UIS0_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD382
RD420
RD532
RD64
RD290
RD366
RD354
RD150
RD28
RD162
O1 – Bt2
O1/O5 – Bt2
O1 – Bt2
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O5 – Bt1
O2 – Bt1
O7 – Bt1
O5 – Bt1
O6 – Bt1
0.0002
YiaD/Omp16 – C4UGU4_YERRU
RD382
RD420
RD532
RD6
RD10
RD524
RD520
RD84
RD124
ATCC29473
RD64
RD150
RD28
RD162
RD354
RD290
RD366
O1 – Bt1
O1/O5 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O2 – Bt1
O7 – Bt1
O5 – Bt1
O6 – Bt1
O2 – Bt1
O5 – Bt1
0.0002
YiaF – C4UGV3_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD382
RD420
RD532
RD64
RD150
RD354
RD290
RD366
RD162
RD28
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O7 – Bt1
O2 – Bt1
O5 – Bt1
O6 – Bt1
O5 – Bt1
0.0002
YiiQ – C4UIW5_YERRU
RD382
RD420
RD532
RD162
RD64
RD28
RD290
RD366
RD354
RD6
RD10
RD524
RD520
RD84
RD124
RD150
O1 – Bt1
O1/O5 – Bt1
O6 – Bt1
O2 – Bt1
O5 – Bt1
O2 – Bt1
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O7 – Bt1
0.0005
YnbE – C4UFK4_YERRU
RD6
RD10
RD524
RD520
RD84
RD124
RD64
RD382
RD420
RD532
RD354
RD28
RD162
RD150
RD290
RD366
O1 – Bt2
O1/O5 – Bt2
O1 – Bt1
O5 – Bt1
O1 – Bt1
O1/O5 – Bt1
O2 – Bt1
O5 – Bt1
O6 – Bt1
O7 – Bt1
O5 – Bt1
0.001

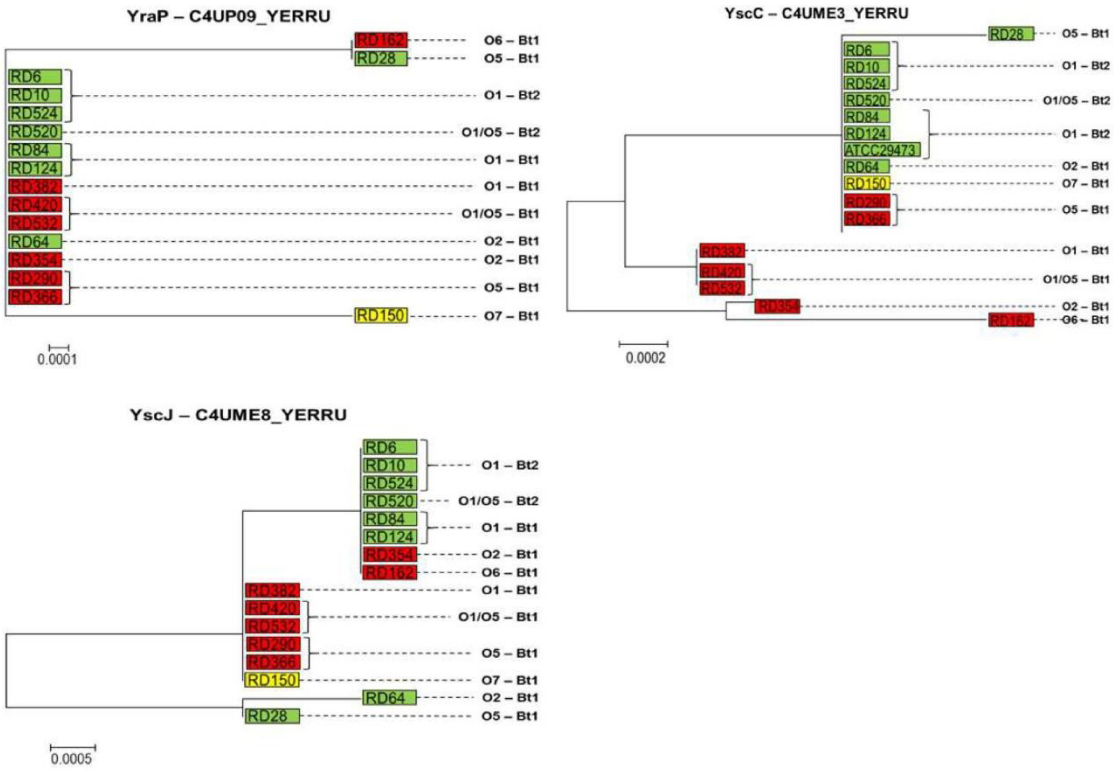


Fig. 8.2 Phylogenetic relatedness trees of predicted OMPs in *Y. ruckeri*.
The phylogenetic relatedness trees of 141 predicted OMPs are presented in this figure. Isolates coloured green were recovered from rainbow trout, isolates coloured red from Atlantic salmon and the isolate in yellow from European eel. The gene represented in each tree is shown above, with the UniProt accession number.